E-ISSN : 2798 - 4664

INTERNATIONAL
Applied Business and Engineering
Conference
2021

Pekanbaru
August, 25th 2021

PROCEEDING

Organized by

Politeknik Caltex Riau

In collaboration with:

POLMANBABEL · polibatam · POLITEKNIK NEGERI PADANG · POLITEKNIK MALAYSIA Mersing · POLITEKNIK MALAYSIA Mukah

PEKANBARU, 2021

PROCEEDING

INTERNATIONAL ABEC 2021
International Applied Business And Engineering Conference 2021

Topic:

Technology Acceleration and Digital Readiness Towards New Era of Creative Economy and MSME (Micro, Small and Medium Enterprises)

Steering Commite:

Dr. Mohammad Yanuar Hariyawan, S.T., M.T.
Retno Tri Wahyuni, S.T., M.T.

Conference Chair:

Anggy Trisnadoli, S.ST., M.T.

Politeknik Caltex Riau, Pekanbaru-Riau
25 August 2021

E-ISSN : 2798 - 4664

PROCEEDING

INTERNATIONAL ABEC 2021
International Applied Business And Engineering Conference 2021

Topic:
Technology Acceleration and Digital Readiness Towards New Era of Creative Economy and MSME (Micro, Small and Medium Enterprises)

Editor in chief:
Yuliska, S.T., M.Eng

Editors:
Daniel Sutopo Pamungkas, Ph.D., IPM.
Ibnu Surya, S.T, M.T.
Meliza Putriyanti Zifi, S.E., M.Acc.
Muhammad Diono, S.S.T., M.T.
Dr. Ir. Yuhefizar, S.Kom., M.Kom., IPM

Politeknik Caltex Riau, Pekanbaru-Riau
25 August 2021

Preface: Proceedings of International Applied Business and Engineering Conference 2021 (IABEC 2021)

This International Applied Business and Engineering Conference 2021 (IABEC 2021) held in Politeknik Caltex Riau (PCR), Pekanbaru, Indonesia. International Applied Business and Engineering Conference 2021 with the topic "Technology Acceleration and Digital Readiness Towards New Era of Creative Economy and MSME (Micro, Small and Medium Enterprises)" intended to confer, share and innovate on recent developments and trends in applied business engineering for the research community. Experts from different parts of the globe dealing in Applied Business & Engineering attended IABEC 2021. This Conference provided an international forum to present, discuss and exchange innovative ideas and recent developments in the field of Applied Business and Engineering.

On this occasion, 3 distinguished keynotes speakers had delivered their outstanding research works in various fields of Applied Business and Engineering. There were 40 oral presentations by participants from 5 countries: Indonesia, Malaysia, Japan, Taiwan, and Italy. The oral presentations brought great opportunity to share participant's recent research works knowledge with each other graciously.

Efforts taken by reviewers contributed to improve the quality of papers provided constructive critical comments, improvements, and corrections to the authors are gratefully appreciated. We are very grateful to keynote speakers, session chairs, reviewers, participants, committee, and student volunteers who selflessly contributed to the success of this Conference. Last but not least, we are thankful to all the authors who submitted papers, because of which the Conference become a story of success. It was the quality of their presentations and their passion to communicate with the other participants that really made this conference series a grant success.

Anggy Trisnadoli, S.ST., M.T.
Conference Chair-IABEC 2021

TABLE OF CONTENTS

THE COMMITTEE OF IABEC 2021

Dr. Mohammad Yanuar Hariyawan, S.T., M.T.	Steering Commite
Heri Subagiyo, S.T., M.T.	Steering Commite
Wahyuni Khabzli, S.T., M.T.	Steering Commite
Retno Tri Wahyuni, S.T., M.T.	Steering Commite
Anggy Trisnadoli, S.S.T., M.T	Conference Chair
Yuliska, S.T., M.Eng	Technical Conference Program (International)
M. Suhairi, S.Pd., M.T	Technical Conference Program (Domestic)
Mona Elviyenti, S.Si., M.Si.	Secretariat
Puja Hanifah, S.S.T., M.MSI.	Secretariat
Leyon Priyan Gusdi, A.Md.	Member
Fifitri Ali, S.S.T.,M.Sc.	Member
Mardhiah Fadli, S.T., M.T.	Member
Khairul Umam Syaliman, S.T., M.Kom.	Member
Hamid Azwar, S.T., M.T.	Member
Arif Gunawan, S.T., M.T.	Member
Roni Novison, S.T., M.T.	Member
Sandy Setyo Nugroho, A.Md	Member
Muhammad Saputra, S.ST	Member

THE REVIEWER TEAM

Anggy Trisnadoli, S.ST., M.T	Politeknik Caltex Riau, Indonesia
Ts. Dr Bong Siaw Wee	Politeknik Mukah Sarawak, Malaysia
Ts. Dr Chen Wong Keong	Politeknik Mukah Sarawak, Malaysia
Dr. Dalrino, ST., MT	Politeknik Negeri Padang, Indonesia
Daniel Sutopo Pamungkas, Ph.D., IPM.	Politeknik Negeri Batam, Indonesia
Dr. Erlin, M.Kom	Institut Bisnis dan Teknologi Pelita Indonesia, Indonesia
Dr. Evizal, ST, M.Eng	Universitas Islam Riau, Indonesia
Dr. Habsah Binti Mohamad Sabli	Politeknik Mukah Sarawak, Malaysia
Dr. Hendriko, M.Eng.	Politeknik Caltex Riau, Indonesia
Humaira	Politeknik Negeri Padang, Indonesia
Dr. Mas Nordiana Rusli	Politeknik Mukah Sarawak, Malaysia
Dr. Mohammad Yanuar Hariyawan, S.T., M.T.	Politeknik Caltex Riau, Indonesia
Muhammad Sufyan Safwan bin Mohamad Basir	Politeknik Mukah Sarawak, Malaysia
Dr. Okfalisa, ST, M.Sc	Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
Qoriatul Fitriyah, S.T., M.Sc.	Politeknik Negeri Batam, Indonesia
Satria Perdana Arifin, S.T., M.T.I	Politeknik Caltex Riau, Indonesia
Ts. Shanusi Bin Ahmad	Kolej Vokasional Sultan Ahmad Shah, Malaysia
Sri Rizki Putri Primandari, M.T., Ph.D	Universitas Negeri Padang, Indonesia
Warnia Nengsih, S.Kom., M.Kom	Politeknik Caltex Riau, Indonesia
Yoanda Alim Syahbana, S.T., M.Sc.	Gifu University, Japan
Dr. Yohana Dewi Lulu Widyasari, S.Si., M.T.	Politeknik Caltex Riau, Indonesia

KEYNOTE SPEAKERS

Dr. H. Sandiaga Salahuddin Uno
MINISTER OF TOURISM AND
CREATIVE ECONOMY, INDONESIA

Dr. Ramesh Zaidi Rozan
ASSOCIATE PROFESSOR
INFORMATION SYSTEMS AZMAN HASHIM
INTERNTIONAL BUSINESS SCHOOL
UNIVERSITI TEKNOLOGI MALAYSIA

Dr. Dadang Syarif Sihabudin Sahid
ASSOCIATE PROFESSOR
ON INFORMATION TECHNOLOGY DEPARTMENT,
POLITEKNIK CALTEX RIAU, INDONESIA

COMPUTER SECTION

Invited Speaker :
Dr. Roslina, M.I.T

"Characteristics of Cash Waqf Perpetrator in the Digitalization Era"

POLITEKNIK NEGERI MEDAN

Price Difference Embedded Multivariate Long Sort-Term Memory for Stock Movement Prediction

1st Wawan Yunanto
Department of Computer Science and Information Engineering
National Taiwan University of Science and Technology
Taipei, Taiwan
d10615811@mail.ntust.edu.tw

***Abstract*—The course of the capital market is complicated, unpredictable, and volatile for investors to formulate. Fundamental and technical analysis are the main common approaches to predict stock prices used by the economic experts nowadays. The fundamental aspect is determined by the internal factor of the companies, but the technical one is clearly represented on a daily basis in the stock market. Stock price is not the only important item for investors to make investment policy. The fluctuation in the trading floor has become the most important issue to be considered. In this research, we propose a prediction framework, namely Price Difference Embedded Multivariate Long Sort-Term Memory (PDEM-LSTM), to combine stock price and movement prediction into a single pipeline. We employ recurrent deep learning modeling technics into stock market forecasting since there are sequential properties in the technical components. Our work solely based on these sequences or timeseries features to simplify the experimental setting and more focus on the improvement compared to most previous studies. Our benchmark compares results from univariate scheme on the same sequences with 3 difference features which are current day and next day price along with the price difference between those days. We use 5 stock issuers from 5 different stock indices and the market data taken from January 1, 2000, to December 31, 2020. The results showed that price difference feature embedded into LSTM in multivariate setting greatly improve stock movement prediction without degrading stock price forecasting too much. It is simple and robust; it can be attached on most stock prediction techniques in the feature engineering phase.**

Keywords—stock price, stock movement, LSTM, univariate, multivariate

I. INTRODUCTION

Capital market is a high-risk investment due to its broad factors that can affect the profit taking strategy. There are two major aspects noticed by investors before they make trading decisions. Company performance become the first aspect to make a good justification about how secure the investment will be [1]. These fundamental statistics tend to attract long term investors to spend their capital and participate in every decision making made by the company. On the other hand, there are sort-term investors who pay attention more on the second aspect, namely technical aspect, in trading. Technical aspect mostly driven by external factors, such as: price trend, investor sentiment, political situation, or event global pandemic just like we are all experiencing today. Hence, investors and traders are prone to severe monetary loss if their do not maintain the financial risk management carefully. Conversely, careful data-driven investment decisions can maximize profits and minimize loss greatly [2].

The vast majority research in financial studies rely on the sequential data that widely available for public. In stock market there are several time-series features to choose for building robust prediction method for traders. The most common features are stock prices, e.g., opening and closing price. Prediction the next day price seems to be the most important thing for investors because in most people's view money is closely related to price. The problem is that even the lowest error price prediction for the next day does not give us any clue whether we should buy or sell. In a statistical perspective, low error means a good prediction but what the investors really need is the direction of the next day price, is it better or worse than today's price. If next day price predicted to be lower than today, regardless the error, traders will likely sell their share and vice versa, as simple as that. Therefore, prediction of stock movement direction considered to be significant in financial studies and plays a key role in determining to buy and sell stocks [3].

Unlike stock price prediction, which focus on predicting the next day price by minimizing the mean squared error and usually resulting a really low errors, stock movement prediction is much harder to cope so that it can be implement in trading strategy. We cannot just use the price differences between real and predicted prices to justify the stock movement direction because oftentimes they are pointing to the opposite direction. Some researchers engineered new features [4], utilized technical analysis [2], or even formulated a new loss function (in deep learning) to get better stock movement prediction [5]. In this study, we focus on how to build a simple framework to get significant improvement, only using single time-series feature in the stock market data. Our goal is building a method that can be generalized to be use in any stock movement prediction to improve the prediction result. We are not designing a production level end to end method but more on the enhancement of any deep learning-based predictor using our strategy.

In brief, the main contributions of this work can be summarized as follows:

1. We extend the use of price differences and embed them in LSTM-based predictor in multivariate manner.
2. We improve stock movement prediction result while keeping stock price prediction errors at the acceptable level using single pipeline network.
3. We provide extensive experiments in univariate manner as well to benchmark our proposed framework against common prediction scenarios.

The rest of paper is organized as follows. Previous works in this area are introduce in Section 2. In Section 3 we explain about the methodology of this research. Our results and its analysis are provided in Section 4. Finally, the conclusion is drawn in Section 5.

II. Related Work

Stock movement prediction is widely considered challenging due to the high stochasticity of the stock signal and sometimes chaotic. Basically, it is a time-series problem where the temporal dependencies are crucial [6]. Research in stock market field, most of previous works are usually a comprehensive research to get better result for traders. They use all features available, fundamental and technical analysis, or even external data such as social media.

Technical indicators become the most popular features to use by researcher to forecast next day stock price and movement. These features used by predictors, mostly machine learning models, to extract more features in order to improve prediction ability [7]. Shallow model machine learning methods such as k-Nearest Neighbor (kNN), Support Vector Machine (SVM), and Random Forest are often utilized to capture the effect of using the features from the technical analysis and price movement. These models use the deterministic trends of stock price index movement to learn the correlation between current and previous stock data, but still get undesirable outcome [4].

A deep learning approach has been used in several research in this area. The usage of Convolutional Neural Network combined with LSTM has been proposed in [3]. Although this method performs well in detecting stock price but struggle in forecasting stock movement. The best performance is only 0.6144 in average on 5 market indices using their proposed network with CNN3D-DR+LSTM. An extensive deep learning model based on Factorization Machine and Attention Mechanism was utilized to predict stock price movement as presented in [5]. The idea of this paper is to construct a convolutional neural network graph based on a deep factorization machine and attention mechanism (FA-CNN) to improve the prediction accuracy of stock price movement via enhanced feature learning. The average accuracy obtained is below 60%, approximately equal to others. Other researchers try to involve external aspects related to investment field, namely social media comments and company correlations, in order to overcome the non-stationary nature of stock market [8]. They best accuracy is 0.608 on MAN-SF with Fundamental Analysis models. It is slightly better than the others but still not far from flip of a coin.

One of the advantages in using deep learning method is that we can design our input vector to be suitable to our need. Combination of time-series features and engineered features can capitalize the prediction power of the model in multivariate scheme. Our work is building a feature engineering method to benefit from this advantage.

III. Methodology

A. Dataset

In this research, we utilize 5 stock issuers namely Abbott Laboratories (ABT), Walmart Inc. (WMT), Intel Corporation (INTC), Boeing Company (BA) and HSBC HOLDINGS (0005.HK), from these international stock market indices respectively: S&P 500, DJIA, NASDAQ, NYSE and HKEX. We collect the data between January 1, 2000, to December 31, 2020, from Yahoo Finance (https://finance.yahoo.com). As mentioned earlier, our work focusses on finding performance improvement by employing time-series feature contained in stock market data, that is closing stock price.

Our proposed experiment scheme is a multivariate LSTM that takes multi-dimensional vectors as input. Before creating the input sequences, we preprocess the original dataset to generate next day closing price signal and extract the price movement by capturing the difference between today and next day prices. We keep the price difference values as signed real number and not convert them to boolean because we want our model to not only learning the movement but also the magnitude of the movement. When the next day price is going up, we want to know how much it is, likewise if next price is going down. As a result, our engineered dataset is 3-dimensional with each one of the features hold a sequential property.

Let i be the trading day index, P be today closing price, Q be next day closing price and D be the difference between next day and today's price. We engineer the new features as follows:

$$P = \{p_1, p_2, p_3, \dots, p_{i-1}\} \quad (1)$$

$$q_i = p_{i+1}, q \in Q \quad (2)$$

$$d_i = q_i - p_i = p_{i+1} - p_i, d \in D \quad (3)$$

We utilize holdout method to partition the dataset into 3 parts, which are train, validation and test set, with each portion is 80%, 10% and 10% respectively. We build and train the predictor using train set and observe the result for convergence using validation set. Finally, we test our method and analysis the result using test set, to represent data that have not seen before by the predictor.

B. Experimental Setup

Recently, deep learning gains better performance in various tasks compared to more conventional methods due to improved computational power, breakthrough abilities to learn non-linear

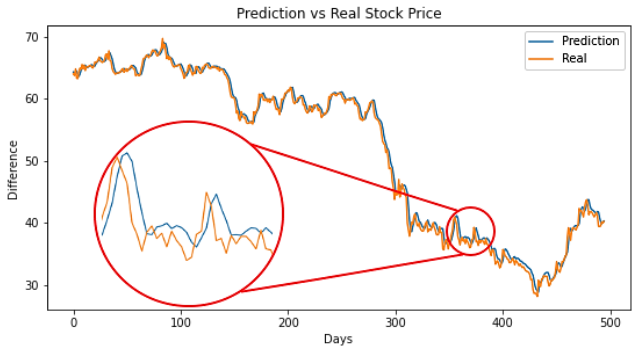


Fig. 1. Univariate LSTM prediction of the HSBC HOLDINGS closing day price.

connection and feature extraction in high-dimensional space. LSTM in particular, designed to learn the sequential observations from previous time steps and capture future trends. Its feedback connection not only can process a single data point, but also entire sequences of data. LSTM introduces the memory mechanism to enables long sequence dependency between time steps [9]. Sequence or time-series forecasting like stock price and stock movement prediction is greatly benefit from the use of LSTM.

In this study, 2 experimental setups are tested and compared, namely univariate and multivariate manner. We use simple univariate LSTM using 3 sequential data points that we have in our modified dataset as the baseline. It represents any method that use the same setting in their experiment. Although we utilize deep learning who has the capability to take multidimensional input, univariate scenario is a good way to represent shallow learning methods with limitation in their input vector shape. In this scenario, we examine a learning method that solely based on the raw sequential property, i.e., today and next day stock price.

In stock price prediction, a simple univariate LSTM without any complicated adjustment like regularization, with only today stock price alone as its dataset, already achieve a good fit, with very low error. In Fig. 1, stock price prediction for HSBC HOLDINGS in Hong Kong's stock index, from 495 days prior to the end of year 2020, shows a very promising result, with 1.002840 mean squared error. However, as we can see in a zoomed region of the graph, the prediction approximates the real price with delayed pattern, means that although the graph is very close, it does not match for the same day trading. Moreover, there are known evidence that for any given trading day we will likely have opposite direction of price movement. This is reinforced by the movement prediction based on this graph, it is only 0.56, not very different than random guessing. It will not good enough to earn lots of profit in real trading.

In addition, we also use the generated price difference data in this univariate scenario to give a broader view about this baseline measurement. The result is almost the same, it only gives random guessing range prediction. We presume that learning solely based on the individual sequential property only capture the temporal feature, connection between current and previous values and exclude the spatial feature, the other property that might present in the surroundings. The model only cares about the statistical values of the input sequence when learning how to minimize error. This particular goal is well achieved, shown by the low error value, but the other goal, detecting price movement, is fall miserably due to lack of movement sequences to be learn simultaneously.

Deep Learning known for its ability to take input in almost any shape. This advantage can be used to create a mechanism how a sequence learning method, let say LSTM, learns not only from the temporal property but also from the spatial property as well. Normally, to predict next value in a long sequence, we need to consider the temporal property given by the historical factor in the sequential feature. The model will find the sequential pattern given by the training set as its knowledge and then use it to predict the next value. Since our augmented data is multidimensional as in (1), (2) and (3), we believe that the spatial properties are beneficial to build a robust stock movement prediction model.

Our proposed method is different from others in term of managing multi-features in building model. Mostly, all previous works mentioned previously use these features in a conventional supervised manner [4][5][8]. They accommodate all features and pass them to the training pipeline without considering the relationship of these features to each other. In machine learning, more features mean more possibility to get better or worse, not to mention the impact caused by the computational cost. Surely, they have curate and evaluate the pros and cons in features selection in their studies to make the most out of spatial properties among features.

We take the same direction as the previous researchers did but with different feature selection, and then make prediction of stock price and stock movement in one single pipeline. We emphasize the power of temporal property in the sequence P and Q with our generated spatial property, i.e., price difference D. Our proposed framework is trying to overcome the delayed pattern problem depicted in Fig. 1. In multivariate, we try to find the prior knowledge in the previous sequence, termed temporal features, and at the same time, expand the search area to the surrounding environment, let say the next feature found, and use them to predict the next value.

We do not regard prediction of stock movement as a supervised binary classification task since it will only predict the trends, it might be similar to the real movement but not align at any given day. Besides, we still have to build another model to predict the stock price. The goal of our proposed framework is to make predicted stock prices as close as possible to the real prices without any delayed movement pattern. We calculate the ups and downs movement based on the predicted prices, and not by building a new model, all task in a single pipeline. We formulate our predicted value $\hat{Y}$ at day i as follows:

$$(\hat{y}_i \approx y_i) \wedge (\hat{y}_i^{movement} = y_i^{movement}) \quad (4)$$

There are two ways to determine $Y^{movement}$, i.e., look-behind and look-ahead price difference. At any given day i, we want to determine the price movement, if we compare today's price against previous day's price, we term it as look-behind price difference. This is not very intuitive since what we are looking for is next day price, but for benchmarking purpose we will accommodate this scenario. Look-ahead price difference seems more suitable for our problem since we consider next day price to determine the price difference. Every data point in the sequence will have a look-behind price difference feature except the first, and vice versa, the last data point will have no look-ahead price difference feature. Whichever scenario that we use in our model, our input sequences will be 1 less than overall sequences from the dataset.

Although we already have equation for the price difference D, in practice we need to extent the formulation in (3) to get $Y^{movement}$. We use binary to indicate movement, 0 for down and 1 for up. In look-behind scenario, we determine stock price movement as follows:

$$y_i{}^{movement} = \begin{cases} 0, & p_i < p_{i-1} \\ 1, & p_i \geq p_{i-1} \end{cases} \quad (5)$$

In line with (5), here is the stock price movement in look-ahead scenario:

$$y_i{}^{movement} = \begin{cases} 0, & p_i > p_{i+1} \\ 1, & p_i \leq p_{i+1} \end{cases} \quad (6)$$

We put it all together in a multivariate LSTM graph. The spatial feature, which is price difference, is embedded to the temporal feature, which is stock price, to capitalize the predictive power of the resulting model. Our input is a 2D tensor with time steps 60 which consist of 2 sequences. We use simple LSTM layer with 60 nodes and Dropout 0.2 to avoid overfitting [10]. Since this is basically a regression task, we utilize MSE (Mean Squared Error) as our loss function and Adam optimizer with learning rate 0.001 to control our model convergence [11]. All the comparative models were implemented in slightly the same setting, except the Look-ahead scenario.

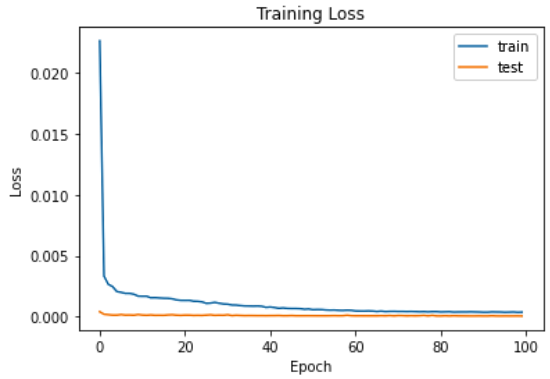


Fig. 2. Look-ahead PDEM-LSTM Training Loss

Price difference in Look-ahead scenario obtained by comparing today's price with next day's price. We can easily get the next day value from our dataset, but this is not applicable in the real world. Let say today is Tuesday and we need to predict Wednesday's price. We create a sequence consist of price historical data and its Look-ahead price difference. The problem is that we cannot calculate price difference for today because we have not known Wednesday's price yet. Just like we mentioned before, the last data point in Look-ahead scenario do not have price difference. In fact, Wednesday's price is the goal that we trying to predict. So, we need to make adjustment in the implementation. We shift the input sequence to exclude the last data point, which is Tuesday's price. Monday's price will be the last data point in our input sequence to predict Wednesday's price, there are 2 days gap. However, we need to maintain 1 day gap in our prediction to align with our input sequence. We employ sequence to sequence prediction, 60-time steps input sequence will be used to predict 2-time steps result. We adjust the final layer of our model with 2 nodes fully connected layer. The first node will contain Tuesday's predicted price that we will ignore since we already have Tuesday's real price. And the second will contain Wednesday's price, the one that we record to calculate the stock movement accuracy.

IV. Result and Analysis

A. Metric

Our proposed framework is actually a regression model, so the suitable performance metric is MSE. However, the price predicted is only an intermediate goal to be further used for determining its movement each day, as in (6). The MSE is used to ensure that our price prediction is acceptable, not too far away from the actual stock price. Afterwards, Movement Accuracy (MA) is calculated by comparing $y_i{}^{movement}$ and $\hat{y}_i{}^{movement}$ as follows:

$$MA = \frac{\sum_{i=1}^{n} [\hat{y}_i^{movement} = y_i^{movement}]}{n} \times 100\% \quad (7)$$

We utilize Iverson bracket notation to count how many predicted movements match the real stock movements, all the way from day 1 to day n.

B. Model Convergence

We observe the learning curve of our model using loss graph at training phase. The train and validation/test loss should converge after a certain number of epochs. We need to ensure that our model has a decreasing loss and constant slope by the end of the training process. Fig. 2 shows our train loss decrease as intended and converge along with the test/validation loss at the end of the training phase. Ideally, the validation loss will be higher than the train loss, unlike our loss graph, but our model is considered acceptable since we do not tune the hyperparameter excessively. After all, both losses are converged eventually. The gap between these losses describes the generalization ability of the model. There is no sign of overfitting, our model is ready to predict data that has never seen before.

Fig. 3 depicts Look-ahead PDEM-LSTM for 0005.HK stock price. Our proposed model manages to successfully predict the stock prices with very low error. Although the error considered lower than the other models, but the value difference is extremely small, so it does not upgrade the stock price prediction significantly. The prediction graph matches the real price perfectly, no sign of excessive delayed movements. The numbers shown in Table I confirm that while maintaining price prediction as precise as possible, Look-ahead PDEM-LSTM surprisingly achieve a very good accuracy on stock movement prediction as well. All of this is obtained with single integrated pipeline, no extra computational cost. Furthermore, this great performance also applies to other datasets that we explore in this study.

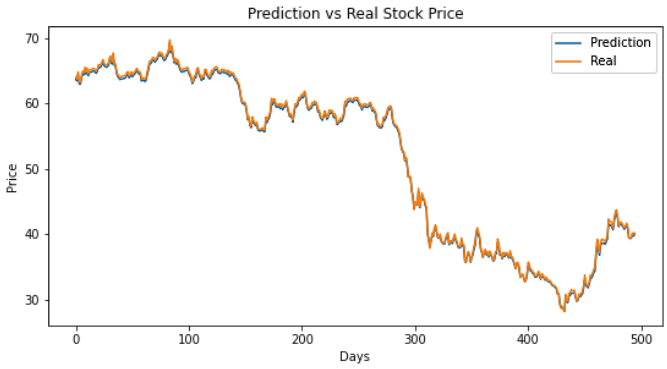


Fig. 3. Look-ahead PDEM-LSTM prediction of the HSBC HOLDINGS closing day price.

C. Comparative Study

Stock movement prediction is considered a very challenging task. A small portion of improvement can lead to a significant potential profit. In this field, the accuracy of 56% is generally reported as a satisfying result [12]. Most of the previous studies accommodate as many factors as possible in their models, namely fundamental analysis, technical analysis, political situation, social media comments and more. This complex information truly influences the model performance. This is the best practice to build an end-to-end production level application. Meanwhile, this research does not go to that direction. We focus on maximizing the use of temporal and spatial features resided in the stock market datasets. We may not have an end-to-end production level application, but we still give a noteworthy contribution in this research area.

Our main focus is the improvement that our proposed framework can achieve with only using simple model. Table I exhibits the overall performance conducted in our experiments using 0005.HK stock price. Sequential properties in price data alone does not perform well enough to predict stock movement although the prediction error considered very low as shown in the first and second row. The movement accuracy score is in par with most of the previous studies. The third row tells the performance of our engineered feature in univariate manner to do the same task. Although we make an exception, the MSE noted is not related to stock price but price difference instead. Anyway, the movement accuracy is slightly better than other univariate models.

The combination of temporal and spatial features in multivariate manner makes a notable improvement. It is a surprise that with more data involved, the results are very different. Look-behind scenario struggle to perform better even get the worst performance. The idea to engineer price difference using previous day's price does not affect movement accuracy at all. The accuracy of 55% is considered no improvement whatsoever, it is still on the same level as other previous studies. This is allegedly because in Look-behind scenario, the price difference does not describe the future values that our model should train on. We determine the price difference by looking at the past, the previous day's price. Therefore, the model learning direction points to the opposite direction, generate more delayed movements despite the low price MSE. Our model should learn the future, it must be taught using future directed features, price difference should consider the next day and not the previous day. Look-ahead scenario is appropriate for this task since it is created with future direction in mind.

TABLE I. MOVEMENT ACCURACY ON VARIOUS DATASET USING SVM

Dataset	Movement Accuracy	Price MSE
S&P 500: ABT	46.31%	7.241039
DJIA: WMT	46.51%	11.67176
NASDAQ: INTC	48.70%	2.013702
NYSE: BA	44.91%	80.23296
HKEX: 0005.HK	55.42%	0.904179

LSTM is known for its ability to deliver previous knowledge from the past sequence and use it to predict current value based on current input sequence. As a baseline, we build another predictor with slightly the same approach to benchmark our proposed method. We utilize SVM with RBF (Radial Basis Function) kernel. This kernel accepts two parameters, the first is gamma which defines how far the influence of a single example reaches, means that this parameter act just like the short-term memory in LSTM in carrying previous knowledge to predict the current situation. The second parameter determines the margin being accepted as a regularization function in SVM.

Table I shows the overall results on various datasets that we examine in this research. We can draw several conclusions based on these numbers. In term of MSE, SVM did a good job in predicting the stock price although the Price MSEs are no match against more advance method like LSTM. In term of movement accuracy, SVM experienced delayed pattern, so the result is slightly equal to random guess, somewhat the same as experienced by Univariate LSTM.

TABLE II. 0005.HK MOVEMENT ACCURACY AND PRICE MSE

Model	Movement Accuracy	Price MSE
Univariate – Today's Price	56.48%	1.002840
Univariate – Next Day's Price	56.48%	1.424277
Univariate – Price Difference	64.98%	0.544506*
Look-behind PDEM-LSTM	55.87%	0.070927
Look-ahead PDEM-LSTM	**97.77%**	**0.070927**

[a.] Third row represents Price Difference MSE, not Price MSE*

TABLE III. MOVEMENT ACCURACY ON VARIOUS DATASET

Dataset	Movement Accuracy	Price MSE
S&P 500: ABT	97.18%	0.278237
DJIA: WMT	97.99%	0.265340
NASDAQ: INTC	98.99%	0.017432
NYSE: BA	96.38%	4.184101
HKEX: 0005.HK	97.77%	0.070927

Last row of Table II indicates the best performance of our proposed model without sacrificing price MSE too much. It is harder to implement than the others, but the result is very promising. The accuracy of 97% is a huge improvement and everyone should consider this approach to predict stock movement. This simple framework can be attached easily to whatever model that one already has, and it is ready to be developed further to become production level application.

Generalization ability is very important since no one knows what kind of data will be processed by our model. A good model must generalize better even with a broad kind of data throw at it. We have to ensure that our model behave as expected, not only good in predicting test set but also good in predicting arbitrary input. Table III shows the performance of our model against data from 5 stock issuers in 5 stock indices. The performance remains excellent in all datasets that we use, this indicates that our model has a sufficient generalization ability. It is simple, robust, yet easy to implement.

V. CONCLUSION

In this paper, we proposed Price Embedded Multivariate Long Short-Term Memory to predict stock movement using only feature that has sequential property in it. The main contributions of this work can be summarized as follows:

1. Our engineered price difference feature makes a huge impact on the movement accuracy using Look-ahead scenario and embed them in Multivariate LSTM predictor.
2. Look-ahead PDEM-LSTM performs surprisingly good at predicting stock movement while maintain predicted price error at an acceptable level.
3. Our model not only perform great but also generalize better against all datasets that we use in this research.

One of the reasons for this good performance is that our datasets contain sufficient amount of data, from 2000 to 2020. We have not test against stock data from new companies which definitely much less. In the future, we will examine our model using short sequences so it will be useful to be implemented using relatively newly listed companies which only has a very little stock data.

ACKNOWLEDGMENT

The author would like to thank Google LLC for providing a helpful IDE. Most of the work in this research is conducted on Google Colaboratory.

REFERENCES

[1] K. Adam, A. Marcet, and J. P. Nicoli, “Stock Market Volatility and Learning”, The Journal of Finance, Vol. 71, pp. 33-82, 2016

[2] M. J. S. de Souza, D. G. F. Ramos, M. G. Pena, V. A. Sobreiro, and H. Kimura, “Examination of the profitability of technical analysis based on moving average strategies in brics”, . Examination of the profitability of technical, Vol. 4, 2018

[3] C. Yang, J. Zhai, and G. Tao, “Deep Learning for Price Movement Prediction Using Convolutional Neural Network and Long Short-Term Memory”, Mathematical Problems in Engineering, 2020

[4] J. Patel, S. Shah, P. Thakkar, and K. Kotecha, “Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques,” Expert Systems with Applications, vol. 42, no. 1, pp. 259–268, 2015

[5] X. Zhang, S. Liu, and X. Zheng, “Stock Price Movement Prediction Based on a Deep Factorization Machine and the Attention Mechanism”, Mathematics, 2021

[6] Y. Xu, and S. B. Cohen, “Stock Movement Prediction from Tweets and Historical Prices”, Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Long Papers), pp. 1970-1979, 2018

[7] Y. Shynkevich, T. M. McGinnity, S. A. Coleman, A. Belatreche, and Y. Li, “Forecasting price movements using technical indicators: investigating the impact of varying input window length,” Neurocomputing, vol. 264, pp. 71–88, 2017

[8] R. Sawhney, S. Agarwal, A. Wadhwa, and R. R. Shah, “Deep Attentive Learning for Stock Movement Prediction From Social Media Text and Company Correlations”, Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, pp. 8415–8426, 2020

[9] S. Hochreiter, J. Schmidhuber, ”Long Short-Term Memory,” Journal of Neural Computation, Vol. 9, 1997

[10] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting” J. Mach. Learn. Res. 15, pp. 1929–1958, 2014

[11] D. P. Kingma, J. Ba, “Adam: A Method for Stochastic Optimization”, International Conference for Learning Representations, 2015.

[12] T. H. Nguyen and K. Shirai, “Topic modeling based sentiment analysis on social media for stock market prediction”, In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Vol. 1, pp. 1354–1364.

Adaptation of Riau Province Folklore to a 3D Animation with Unwrapping Materials and Folley Effect Method

1st Anggy Trisnadoli
Information Technology Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
anggy@pcr.ac.id

2nd Fitri Andriani
Information Technology Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
andriani@alumni.pcr.ac.id

***Abstract*—Folklore is a story that originates from the community and develop in the community in the past which is characteristic of each nation. In the Rokan Hilir area there is temple that is said to be a legacy of the Green Princess and can be used as historical evidence that this story really exists. The story of Putri Hijau is a story that develop in Pekaitan, Rokan Hilir, Riau. But now, to find out how the story of the Green Princess is very difficult because there is no visual media, even written literature is also difficult to find, and it is feared that this story will be forgotten over time. Even today not all people know this story. Therefore, visual media in the form of 3D animation was made so that people can find out about this story and have visual media that can be used to preserve the folklore in their area. This 3D animation was created using the method of unwrapping materials and foley effects. The results obtained after making this 3D animated film are using the unwrapping method resulting in a more detail, neat texture, so it fits into an object that is the focus of the camera. And by using the Foley effect method, a more detailed and natural sound is produced. In addition, this fil is also made using a blender application and follows the stages in making animation.**

Keywords—Folklore, 3D Animation, Unwrapping Materials, Foley Effect

I. INTRODUCTION

Animation is a process for moving an object. In this case, the animation creation process can be done using various technologies, including CGI (Computer generated Image), Motion picture, 3D animation editor software, special effects, and 3D models [1]. In addition in making animated films, of course, storytelling and foley effects are needed. Storytelling is a way to convey a story to an audience, either in the form of words, images, photos, or sounds. Meanwhile, foley effect is a process of making audio to make the sounds in a film more detailed. To be able to produce animated films with a clear appearance and detailed sound, the unwrapping materials and foley effect methods are applied. Where the materials unwrapping technique that functions to regulate the coordinate mapping of an object so that it can place or map 3D textures precisely according to the design that has been designed so that 3D object will have a texture that is almost like original when compared to texturing without applying coordinate mapping. The textured results might look sloppy and the foley effect is an audio creation to make the sounds more detailed.

The story of Green Princess is a folk tale that originated from Tanah Putih, Rokan Hilir and has begun to be forgotten for now and one way to introduce it is to make a visual form so that it can be seen by people from various circles. Therefore, the authors propose to make a 3D animated film that tells the story of Green Princess which can display objects with a clear appearance and a more detailed sound. This is so that the audience can enjoy good quality films. In addition, this animated film was created using the Blender application. Based on the proposed proposal, it is hope that the surrounding community can find out about the story of Green Princess in the Tanah Putih area, Rokan Hilir, which beginning to be forgotten and there is no difference in the delivery of one person another.

II. 3D ANIMASTION DEVELOPMENT TECHNIQUES

A. Unwrappping Materials

Unwrapping materials is a modifier that functions to manage coordinate mapping an object [2]. Coordinate mapping is required for object that have mapped material (Textured material). The unwrapping modifier allows controlling how the map is applied to the selected sub objects [3]. When a model is created using polygons for the 3D model, UV coordinate can be generated for each point in the mesh. One way of 3D modeler is to define a triangular mesh at the seams, this automatically placing the triangles on a flat page. If the mesh is a UV sphere, for example, the modeler might turn it into a rectangular projection. Once he model is opened, the draftsman can paint triangles or textured squares on each individual face, using the open mesh as a template. When the images are created, each triangle will map to the exact texture of the unwrapping pattern choosing the use of unwrapping

materials can help achieve a more realistic three-dimensional model result target [4].

B. Foley Effect

Foley effect is a process of making audio to make the sounds in a film more detailed, foley effect which is one of the sound elements in the film has a fairly vital role as a medium for delivering non-verbal information in an auditive manner. This information will then form the audiences auditive perception which has an impact on the audience on the meaning of a scene or film as a whole [4]. Examples of the sound of feet, doors, cloth scraping, or hitting. The first person to apply this method in a film is Jack Donovan Foley (1891 – 1967). Foley sound is usually recorded in a studio called the foley stage. A foley artist watches the film to synchronize while recording the required sound. For example in making the sound of footsteps. Foley sound effect is the most responsible and realistic sound effect for applying pressure in a film.

Foley is the art of manipulating sound in everyday life into a film, video, game so that the film or video looks more natural. Foley itself is a very important element in the postproduction process in terms of audio in a film, game and video. Without foley, a film will feel empty and empty. The artist who did the dialogue in the film seemed to be doing dialogue in a vacuum without foley. With the existence of foley, the film becomes more alive because consciously or not the audience wants detailed sound from the rubbing of clothes, furniture, and so on. In addition, foley can also add a more impression to a reaction, for example in an action film where there is a hit scene or in a comedy film that has sounds that appear to be jokes. Foley is usually also used to replace unwanted noises in a film, for example the sound of a motorized vehicle that enters during shooting. The parts of work in foley are:

- Foley artist : The person who works as a foley artist has a duty to mimic the behaviour of the player from the film that you want in foley,
- Foley Operators : to record foley performed by foley artist, this foley crew is also in charge of mixing sounds that have been obtained during foley.
- Foley Stages : the room where foley artists perform foley.
- Foley Operator Room : a place for foley operators to record foley activities performed by foley artists.

III. METHODOLOGY

A. Modelling

The character and property modelling in this study comes from the book by Sudarno [5]. Here are some character models created using texture unwrapping materials.

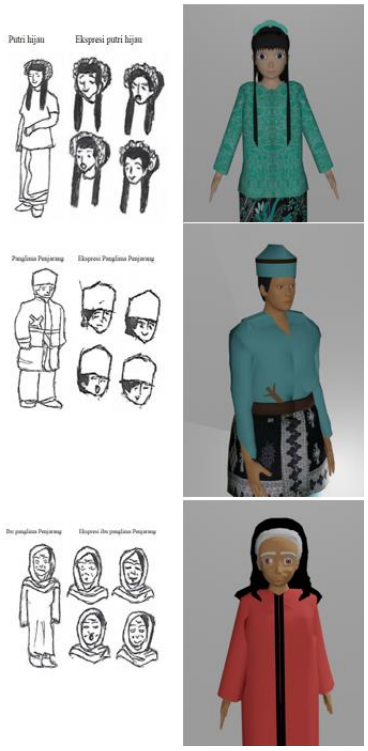


Fig. 1. The Sketch and Concept Art

B. Texturing

After presenting the previous 3D results, we will discuss the texturing process by unwrapping materials. The texturing will be done on the cylinder object because it is a difficult object to attach a texture mapping UV. If texturing is not done by technique whatever it will be obtained the following texturing.

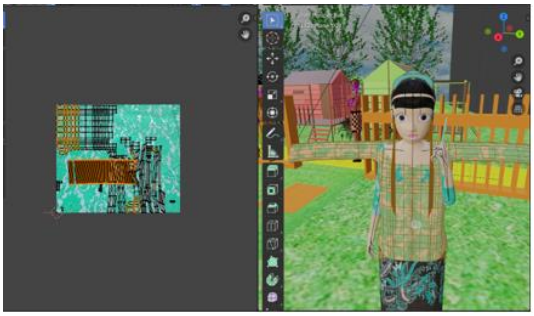


Fig. 2. The results of texturing without using technique

If you pay close attention, texturing without using any technique will cause the texturing to be messy and look untidy. The following will discuss the texturing process using unwrapping materials.

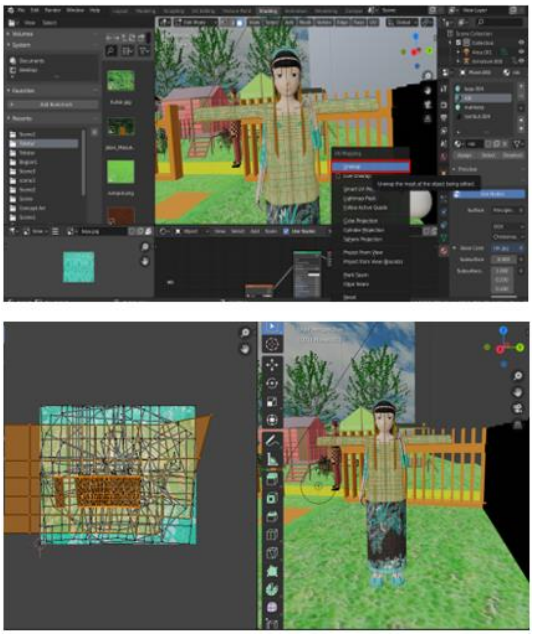


Fig. 3. Mapping process with unwrapping materials

Texturing done without using any technique will look messy and irregular in texture. While the texturing done by using unwrapping materials looks tidier and more organized, the object of the image that you want to text is also more detailed and clearer.

Fig. 4. The results are textured using unwrapping materials

IV. Analysis Of Evaluation Results

After the film product has been made, the film will be validated. And the validation process was carried out by the Riau Malay Customary Institution. The validation process is carried out providing a checklist document containing statement relating to the content and validity of the story which will be answered with yes or no answers. Following are the results of the checklist document related to film validation.

From this test, the results obtained were 100% yes which means this film already has the truth or validity of the story, content that is worth seeing, and a plot that is easy to understand. Furthermore, to test the unwrapping materials method, it was done by comparing 2 objects made with different methods, namely the unwrapping materials and UV Mapping methods. After the two objects were made, several people were asked to make comparisons of the results of these two methods so that the writer would draw conclusion.

And for the foley effect the test will be carried out by displaying videos made without foley effects and animated films that have applied the foley effect technique later. Respondents for conducting this test were student majoring in information technology, especially the informatics engineering study program who understands about making animate film with techniques. This is done so that the process of comparing the 2 methods is carried out by people who know how the process of comparing the textures of the UV Mapping and Unwrapping materials methods so that respondents have a reason why they choose this method so that later they can give their opinion about these two methods. The following is a chart related to the unwrapping materials test.

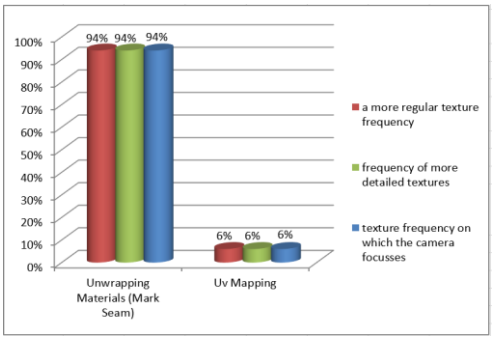


Fig. 5. Respondents for a more regular textures

Based on fig 4, the value is 94% where the unwrapping materials method produces a more regular, detailed texture, and is suitable for the texturing that is the focus of the camera. Meanwhile, UV mapping produces a value of 6%. From the tests that have been described, it is concluded that the unwrapping materials method is suitable for texturing with more regular, detailed variables and becomes the focus of the camera. In addition to the above explanation, next is a graph of the results of respondents' opinions on the unwrapping materials and UV mapping method.

Statement :

P1 : Unwrapping materials technique is better than UV Mapping technique

P2 : UV mapping is generally good for objects that are not the focus of the camera, but if the object is the focus of the camera, you should use Unwrapping materials because of the detail of the textures.

P3 : UV Mapping technique is better than unwrapping materials

P4 : UV Mapping shows texture images to be more detailed than unwrapping materials

P5 : Unwrapping materials look tidier than UV Mapping

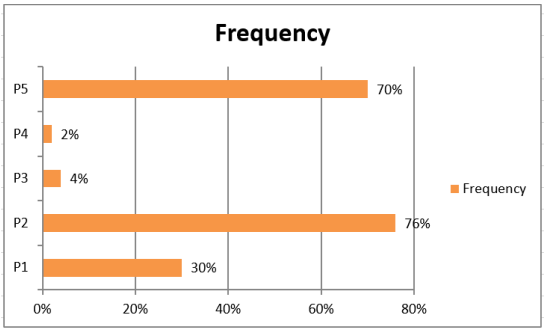


Fig. 6. Validator on UV Mapping and Unwrapping Materials Method

Based on the test chart above regarding the respondents opinion about the UV mapping and unwrapping materials method, it can be seen that there are 70% of respondents who choose that "Unwrapping materials look neat when compared to UV Mapping" this statement is called P5, then 2% of respondents chose that Texturing with UV mapping shows texture images to be more detailed than Unwrapping materials "This statement is called P4, there are 4% of respondents who think that" UV Mapping technique is better than Unwrapping materials "This statement is called P3, 76% of respondents stated that" Texturing with UV mapping in general, it is good to use for objects that are not the focus of the camera, but if the object is the focus of the camera, you should use Unwrapping materials because of the detail of the textures. better than UV Mapping "this statement is called P1. From the five statements above, the highest 3 percentages were obtained, namely 76% for P2, 70% for P5, and 30% for P1. From the results obtained, it can be concluded that the UV mapping method is suitable for objects that are not the focus of the camera or just ordinary texturing, while unwrapping materials are suitable for the texturing of objects that are the focus of the camera because they are tidier and better.

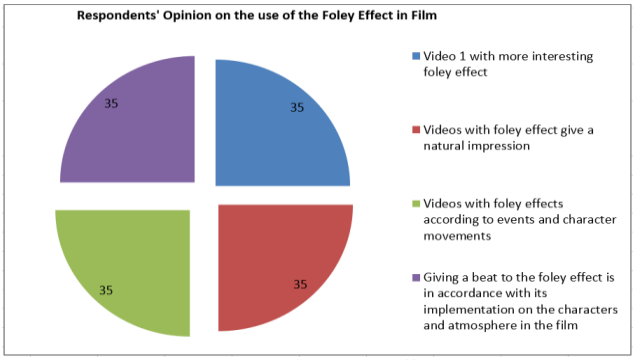


Fig. 7. Responden Opinion to use of the Foley Effect in Film

The graphic above is the answer from 35 respondents, where all 35 respondents chose that the video with the foley effect is more attractive, can give a natural impression, according to events, events, and character movements, and the beats are in accordance with the implementation of the characters and atmosphere in the film. Because of the 35 respondents all of them support the use of the foley effect method, it can be concluded that 100% of the use of the method with the foley effect is better than without the foley effect.

V. CONCLUSSION AND FUTURE WORKS

From this research, we can conclude that the texturing with unwrapping materials got 94% results related to texture, detail, and objects that fit into the focus of the camera compared to the uv mapping value of 6%. In addition, unwrapping materials gets 70% results where the texture of the object is neater than UV mapping, and for the details of the UV mapping method it gets 2% results, which means that this value is lower when compared to the details in the unwrapping materials method, besides that there is 76% result that the object by using the texture method unwrapping materials suitable for objects that are the focus of the camera. And there is a value of 30% that texturing with UV mapping is better than unwrapping. So it can be seen that the unwrapping materials method is more detailed, neat, and this object is suitable to be the focus of the camera. Meanwhile, texturing with UV mapping is suitable for objects that do not pay attention to detail or are not the focus of the camera.

The use of the foley effect method gets 100% results which can be proven by the answers of respondents from 35 people all of whom supports the use of the foley effect method which makes the film more attractive because it can give a natural impression according to events, events and character movements, and the beats are in accordance with the implementation of the characters and atmosphere in the film.

REFERENCES

[1] S. B. Zaharuddin G. Djalle Heni Hendraman, "The Making of 3D Animation Movie Using 3DStudioMax.," *Bandung: Bandung Informatics.*, 2006.

[2] H. A. Soma, Dasar-dasar Modeling dan Animasi dengan 3ds max, Jakarta: PenerbitPT Elex Media Komputindo, 2007.

[3] B. S. Nugraha, "Three-Dimensional Model Texturing Using Seamless Unwrapping Material Method," STMIK Amikom, Yogyakarta, 2015.

[4] P. M. Nugraha, "Penerapan Sound Effect Dengan Gaya Hyperreality Dalam Film Fiksi Malam Minggu Kliwon," *Artikel Jurnal*, . 1, 2019.

[5] E. Y. Sudarno, Putri Hijau di Pekaitan, Yogyakarta: Adi Cita, 2006.

[6] M. Andrianto Tjourdry, "Utilizaion of UVW Mapping Techniques For 3D Car View Design," *Informatics Engineering*, 2014.

[7] N. Putu Eka Suputra, "Design of 3D Animated Film The Story of King 1 Rajapala," *SPEKCTRUM*, 2015.

[8] H. Hendratman, "The Magic Of Blender 3D Modelling," *Bandung: Bandung Informatics.*, 2015.

[9] R. &. B. C. J. Thompson, Grammar of the Shot, 2009.

Implementation of Web Service Host to Host Payment for Pajak Bumi dan Bangunan PBB Pekanbaru City

1st Wawan Kurniawan
Departement of Computer
Politeknik Caltex Riau
Pekanbaru, Indonesia
wedy20s2tk@mahasiswa.pcr.ac.id

2nd Dadang Syarif Sihabudin Sahid
Departement of Computer
Politeknik Caltex Riau
Pekanbaru, Indonesia
dadang@pcr.ac.id

***Abstract*—Web service is an application of a collection of data (database), software (software) or part of software that can be accessed remotely by various devices with a certain intermediary. Web service can also be interpreted as a method of exchanging data, regardless of where a database is embedded, in what language an application that consumes data is made, and on what platform a data is consumed. Web services are able to support interoperability. So that the web service is able to become a bridge between the various existing systems.**

Keywords— Web Service, Host to Host, Json, Rest API, HTTP.

I. INTRODUCTION

Information systems and computer technology are developing very rapidly in line with the magnitude of the need for information. Changes and dynamics of society that are getting faster along with the times and technology demand quality information that is fast, precise and accurate. Information systems are one example of a rapidly growing technology product that can assist humans in processing data and presenting quality information so that currently almost all fields of work use information systems.

Every company agency and government certainly needs an information system in carrying out activities its work so that it is more organized and directed with a more efficient time. Likewise with the city of Pekanbaru, into order realize the vision of a civilized Smart City. Where one of the indicators is Smart Governance (transparent, informative and responsive government) information systems are a necessity for every Regional Apparatus Organization.

Pekanbaru City Regional Revenue Agency is one of the Regional Apparatus Organizations which one of its duties is to collect Land and Building Tax in the Pekanbaru City area. The PBB payment process by involving banks has made it easier to process financial data for the Government. But because the bank payment data directly access the database without using the method web service. This is a problem because the bank should not have direct access to the database because the database is data that is very risky to be accessed by the public.

Several studies to integrate information systems have been carried out, one of which uses web service-based technology [1] ,[2], [3], [4], [5], [6]. The architecture of the REST method is the configuration with the best latency value to be implemented in the data integration process [7]. Apart from that, REST is also better than SOAP in response time and response data size [8].

Based on this background, in this study, technology was REST-based web service applied to integrate the PBB payment information system managed by banking with the PBB information system managed by the Pekanbaru City Regional Revenue Agency. With this technology, data communication is more secure because the mechanism is not a combination between systems but the provision of specific data access services (Data on Land and Building Tax Payments).

II. RELATED RESEARCH

A. Web Service

Studies related to web services have been carried out. Rohmat Gunawan, Alam Rahmatulloh in a paper using a REST-based web service method, from the results of his research that results in an analysis of the test results is that the Web service exchanges data using XML format over a network that utilizes the standard internet or intranet protocol, namely HTTP [9].

Other research related to payments using the web service that Randi Rizala, Nature Rahmatulloh journal conducts research using methods Rest API, while the conclusions obtained from this study is the format of data exchange between systems using JSON format stateless (stateless), making it easier to be accessed by any programming language, architecture or different operating system [10].

B. Payment Host to Host

A study related to host to host payments, namely Iis Pradesan, this study aims to design and implement a Web Service using the H2H method from banking, while the research method used is descriptive with the RUP information system development methodology [11].

III. THEORETICAL BACKGROUND

A. Web Service

W3C defines a web service as a software system designed to support communication and interaction between machine to machine (Machine to Machine) through a network (network). Web Services also include Web APIs that can be accessed over a network such as the internet and executed via a remote system according to the requested service. The definition of Web Service according to the W3C also includes many different systems, but in general it is more concerned with the client and server communicating using XML that meets the SOAP (Simple Object Access Protocol) standard. The general assumption is that in terminology there is a description of the machine whose services are provided by the server, or the same as the concept of WSDL. WSDL is not a standard of SOAP but is an absolute requirement for automatic client-side in Java and .NET SOAP frameworks. Some industry organizations such as WS-I claim both SOAP and WSDL as the definition of Web Service essence.

B. API

API is a software interface that consists of a collection of instructions stored in the form of a library and describes how a piece of software can interact with other software. This explanation can be exemplified by analogy if a house is to be built. By hiring a contractor who can handle different parts, the homeowner can give the contractor the tasks the contractor needs to do without having to know how the contractor gets the job done. From this analogy, the house is the software that will be made, and the contractor is the API that works on certain parts of the software without having to know the procedure for doing the work [15].

C. Rest API

REST (Representational State Transfer) is an architectural method of communication that is often applied in the development of web-based services. REST architecture, which is generally run via HTTP (Hypertext Transfer Protocol), involves reading a specific web page that contains an XML or JSON file. This file describes and contains the content to be presented. After going through a certain definition process, consumers will be able to access the intended application interface. The specialty of REST lies in the interaction between the client and server which is facilitated by a number of operational types (verbs) and Universal Resource Identifiers (URIs) that are unique to each resource. Each of the GET, POST, PUT and DELETE verbs has a special operational meaning to avoid ambiguity. REST is often used in mobile applications, social networking websites, mashup tools, and automated business processes. The decoupled REST architecture and the light communication load between producers and consumers make it popular in the world of cloud-based APIs, such as those provided by Amazon, Microsoft, and Google. Web-based services that use such a REST architecture are called RESTful APIs (Applications). Programming Interfaces) or REST APIs [12][13].

D. Json

JSON (JavaScript Object Notation) is a lightweight data interchange format, readable and writable by humans, as well as easily translated and made *(generated)* by the computer. This format is based on part of the JavaScript Programming Language, Standard ECMA-262 3rd Edition - December 1999. JSON is a text format that does not depend on language programming and because it uses a language style commonly used by C family programmers including C, C++, C#, Java, JavaScript, Perl, Python etc. Because of these properties, it makes JSON ideal as a data-exchange language.

JSON is made up of two structures:

- A collection of name/value pairs. In some languages, this is represented as an object (*object*), record (*record*), structure (*struct*), dictionary (*dictionary*), hash table (*hash table*), keyed list (*keyed list*), or *associative array*.
- An(ordered list of values an ordered list of values). In most languages, this is expressed as an array (array), vector (Vector), lists (list), or sequences (sequences).[14]

IV. THE PROPOSED MODEL

There are four stages carried out in this study, namely: data collection, system analysis and software requirements, design, implementation and testing as shown in Figure 1

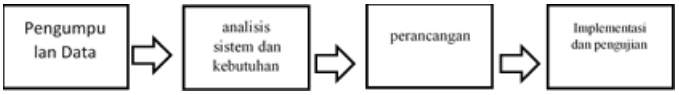


Fig. 1. Stages Carried Out in Study

A. Data Collection

At this stage data collection related to the system to be integrated. An overview of data related to the academic fee payment system is shown in Table I.

TABLE I. DATA AVAILABLE IN THE PBB SYSTEM

No	Name Data	Description
1	SPPT	Determination of Taxpayers
2	Payment of SPPT	Tax Payment List
3	Proof of Payment	List of Payment Proofs

B. Analysis of System and Equipment Needs Software

Based on the analysis of the academic fee payment process carried out by taxpayers, there are 3 entities involved, namely: taxpayers, finance and banking departments. The business process for paying PBB payments is shown in Figure 2.

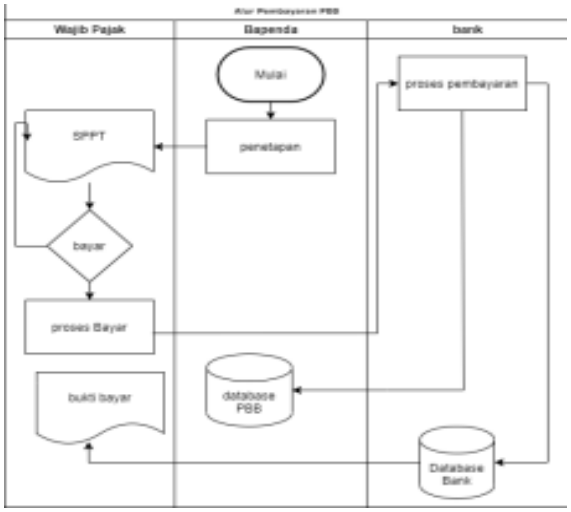


Fig. 2. Flowchart of the ongoing Land and Building Tax Payment Process

In Figure 2, the stages of the current PBB payment process are shown. The explanation of these stages is as follows:

1. The Regional Revenue Agency issues SPPT for taxpayers to make payments immediately.
2. Payment slips or SPPT received by taxpayers can make payments through banks.
3. Banks receive PBB payment fees and provide proof of payment to taxpayers.
4. The bank flags payments directly to the PBB database.

From Figure 2, it can be seen that there are deficiencies in the current system, namely the bank should not be allowed to access the database directly.

C. *Design*

a) System Architecture Design

There are several main components involved in the system designed, including: user, server, database, service and connection. This system is designed to be able to integrate PBB payment systems stored in banks and integrated host to host. The system architecture designed to support system integration is shown in Figure 3.

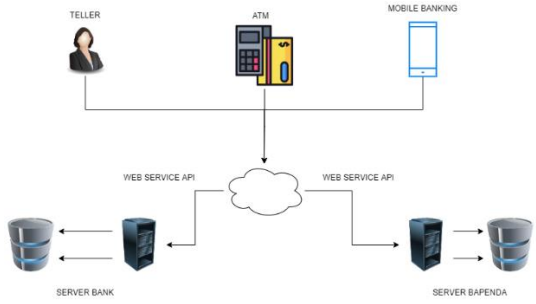


Fig. 3. Web Service API Architecture

In Figure 3 it can be seen, the service used for system integration is placed between the *Server* (BANK) and the *Server* (Bapenda) which is connected by connection. The next stage is the design *service* that will be implemented.

b) Service Design

There are two *services* main that will be built to perform system integration. *The service* is designed to be able to *query data*, *Payments*.

1) Query Data

TABLE II. QUERY DATA

URI Format	:	http://{web}/API-PBB/inquiry_Tagihan
Request Method	:	(GET)

parameter	:	{
		"nop":"14711000040011xxxx",
		"year":"2020"
		}

Description

a. Web: Is the web address where the API stored
b. Body: Parameter requested

2) Payment

TABLE III. PAYMENT

Format URI	:	http://{web}/API-PBB/inquiry_Tagihan
Request Method	:	(GET)

parameter	:	{
		"nop":"14711000040011xxxx",
		"year":"2020",
		"noref":"xxxx112204",
		"merchant":"xxxx",
		"receive_date":"20201203",
		"ship_date":"20201203"
		}

Description

a. Web: Is the web address where the API is stored
b. Body: Parameters requested, from the bank

V. IMPLEMENTATION AND TESTING

At this stage the implementation of the design *service* that has been done previously is carried out. The Design is *service* implemented into programming using the programming PHP language.

A. *Implementation*

Implementation of design that has been done in the previous stage implementation *(coding)* into the PHP programming language and data exchange using format *JavaScript Object Notation* (JSON). Pieces *source code of web service module* API is shown in Figure 4.

Fig. 4. Chunks of *Source Code* the Bill Inquiry Web *Service API*

```
<?php
ini_set('max_execution_time', 600); //300 seconds = 5 minutes
include_once"kon.php";
include_once"afwa.php";

 function curl($url, $data){
    $ch = curl_init();
    curl_setopt($ch, CURLOPT_URL, $url);
    curl_setopt($ch, CURLOPT_CUSTOMREQUEST, "POST");
    curl_setopt($ch, CURLOPT_POSTFIELDS, $data);
    curl_setopt($ch, CURLOPT_RETURNTRANSFER, 1);
    $output = curl_exec($ch);
    curl_close($ch);
    return $output;
}
$alamatapinya    = "https://monitoring-pbb.pekanbaru.go.id/ws/postbayar";

date_default_timezone_set("Asia/Jakarta");
$konfig     = new konfigurasi();
$tanggal    = date("Y-m-d");
$jam        = date("H:i:s");
$upat       = date("Y-m-d H:i:s");
$getip      = $_SERVER['REMOTE_ADDR'];
$tanggalbayar_log   = date("Y-m-d H:i:s");
$getip              = $_SERVER['REMOTE_ADDR'];

$arr        = json_decode(file_get_contents("php://input"));
if (empty($arr)){ // chek array request
    $M_SPPT['status']   = '12';
    $M_SPPT['message']  = "Array tidak dimasukkan";
    $response['Result'] = $M_SPPT;
}else{
    $headers = apache_request_headers();
    if(!isset( $headers['Authorization'])){ //chek autorization header jika tidak ada
        $M_SPPT['status']   = '220402';
        $M_SPPT['message']  = "Authorization is null";
        $response['Result'] = $M_SPPT;
    }else{ //chek autorization header jika ada
        $pecah = explode(".", $headers['Authorization']);
        $aatanggal      = $pecah[1];
        $base64UrlPayload   = str_replace(['+', '/', '='], ['-', '_', ''], base64_decode($aatanggal));
        $tanggal_jwt = substr($base64UrlPayload,-12,-2);
        if($tanggal_jwt==$tanggal){ // chek tanggal header jika sama

                $nop        = $arr->nop;
                $tahun      = $arr->tahun;
                $noref          = $arr->noref;
                $merchant       = $arr->merchant;
                $tanggal_terima = $arr->tanggal_terima;
                $tanggal_kirim  = $arr->tanggal_kirim;
                $tNop           = $nop;
                $kasirbayar     = "BRK000000000000000";
                $KD_PROPINSI    = substr($tNop,0,2);
```

Fig. 5. Chunks Of *Source Code Modules API Web Service* Payments

B. *Test*

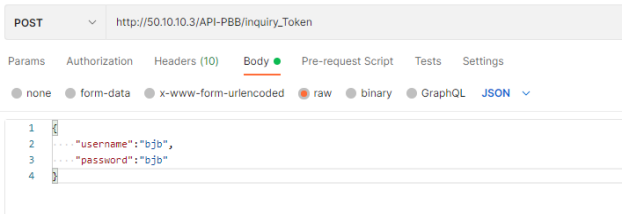


Fig. 6. Generate Token Inquiry UN data

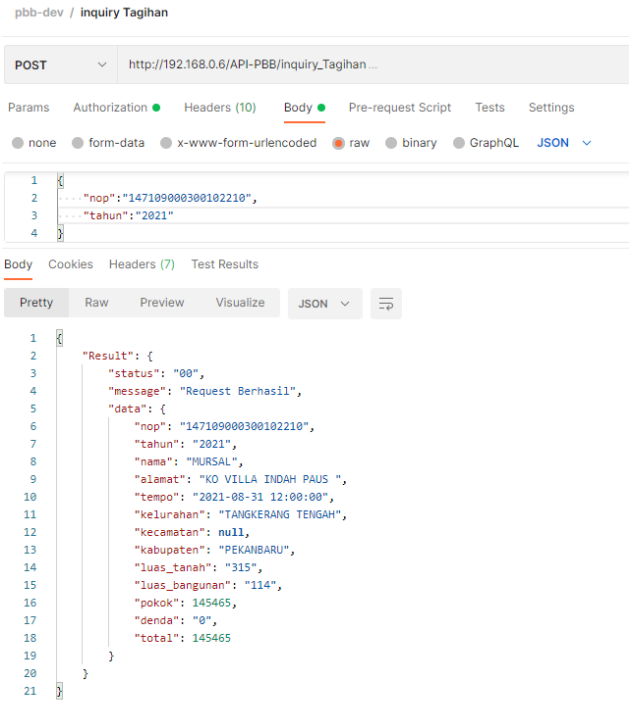


Fig. 7. Generate Token Inquiry UN data

C. *Analyst Testing Results*

After testing, The following is an analysis of the results of the implementation and tests that have been carried out.

1. The PBB payment application on abasis *Host to Host* by implementing technology *web service* has been running according to system requirements.
2. The test system is done in 2 stages: The first stage of performed *testing* on a *web service* that has been built. The second stage is *testing the* application that accesses the *service*, through *interface* the application *Postman*.
3. *Web services* exchange data using JSON format over a network that utilizes the standard internet or intranet protocol, namely HTTP.

VI. Conclusion

A system has been designed that can integrate the PBB payment information system managed by banking with the Land and Building Tax information system managed by the Pekanbaru City Revenue Agency by applying web service technology. Data exchange between systems using the format *JavaScript Object Notation* (JSON). The security of data exchange on the host-to-host system is better than the previous system, because the bank no longer has direct access to the Land and Building Tax database of the Regional Revenue Agency. Meanwhile, to improve the security aspect when exchanging data, an authentication process can be added by applying data encryption and integrating json data with the ISO8583 method.

References

[1] Ratih Titi Komala Sari "Applied PBB Electronic Payment With Host To Host Iso 8583 In Local Government", 2017

[2] IP & BS Officer, “Web Service Implementation In Academic Data Integration With Dikti Database Replication,” *Web Service Implementation. On Integr. Contract Data. With Dikti Database Replica*, vol. 40, no. 1, p. 11, 2017.

[3] R. Hidayat and A. Ashari, “Application of Web Service Technology for Integration of Health Center and Hospital Services,” *Berk. MIPA, (23)I*, pp. 64–77, 2013.

[4] D. Irawati, A. Ashari, and K. Mustofa, “Utilization of Web Services in Data Integration Prototypes of Vocational High School Students Recipients of Student Special Assistance Funds,” in *the 2013 National Informatics Seminar (semnasIF 2013)*, 2013, pp. 12–19.

[5] A. Adi, "Utilization of Web Service as Pharmacy Data Integration at RSU Banyumas (Web Service Useness as a Pharmacy Data Integration in RSU Banyumas)," vol. II, pp. 231–238,2013.

[6] B. Santosa, DB Prasetyo, and Y. Pungki, “Integration of online stores using technology web service,” *Telematics*, vol. Vol. 8 No., pp. 33–42, 2011.

[7] Yogiswara, Wiyono, and HS Dahlan, “Web Service Performance in Data Integration Processes,” *Eeccis*, vol. 1, no. 1, pp. 73–78, 2014.

[8] RK Sungkur and S. Daiboo, “SOREST, A Novel Framework Combining SOAP and REST for Implementing Web Services,” *SOREST, A Nov. Frames. Comb. SOAP REST Implements. WebServer.*, pp. 22–34, 2015.

[9] Rohmat Gunawan, Alam Rahmatulloh "Implementation of Web Services in the Host-To-Host System for Payment of Academic Fees", 2018 p-ISSN: 2301-4652 / e-ISSN: 2503-068X

[10] Randi Rizala, Alam Rahmatulloh "RESTful Web Service for Integration of Academic Systems and Libraries of the University of Struggle' 2019 Scientific Journal of Informatics

[11] Iis Pradesan, "Design and Implementation of Payment Information Systems in Higher Education Using Web Services Method H2H from Banking" 2017 Jatisi, Vol. 4 No. 1 September 2017

[12] https://ekajogja.com/

[13] Ali Firdaus 1, Slamet Widodo 2, Adi Sutrisman3, Sutan Gading Fadillah Nasution 4, Rina Mardiana 5 "DESIGN AND BUILD A LIBRARY INFORMATION SYSTEM USING WEB SERVICE IN POLSRI COMPUTER ENGINEERING DEPARTMENT", 2019 Journal of Informatics, Volume 5 No.2, July-December 2019 ISSN:2407-1730

[14] https://www. json.org/json-id.html

[15] J. Webber, S. Parastatidis, and I. Robinson. 2010. REST in Practice. UnitedStates: O'Reilly Media, Inc.

Implementation And Performance Analysis Development Security Operations (DevSecOps) using Static Analysis and Security Testing (SAST)

1st Wedy Freddy Santoso
Departement of Computer
Politeknik Caltex Riau
Pekanbaru, Indonesia
wedy20s2tk@mahasiswa.pcr.ac.id

2nd Dadang Syarif Sihabuan Sahid
Departement of Computer
Politeknik Caltex Riau
Pekanbaru, Indonesia
dangsyarif@gmail.com

***Abstract*—DevSecOps solves the problem by integrating the security of development operations through various development life cycles. benefits, implementation and challenges during the process. in addition to many documented web hacks. For the scope of work reported that the focus is on two widely used digital library systems: DSpace and Greenstone, in performing Static Application Security Testing (SAST) in addition to more traditional port scanning. Weaknesses were found and details how to make improvements to both systems to make them more secure. can ensure by considering more broadly on the forms of security problems found, to assist the development of software architecture in the future.**

Keywords—Development, operations Security, Architecture, Application.

I. INTRODUCTION

Every individual and organization to understand the importance of cyber security. In 2019, for example, it is estimated that globally a business becomes a victim of ransomware attacks every 14 seconds and is expected to increase to every 11 seconds by 2021 [1]. Regarding personal data breaches, in the same year, cyber security firm UpGuard reported on the discovery of a staggering 550 million Facebook records, totaling 146 GB in size, which had been left arbitrarily exposed on Amazon Cloud Services by Cultura Colectiva, a third-party vendor to the media giant. social [2].

Cybersecurity has become increasingly popular in recent years. Computers and other forms of electronic devices have undoubtedly made many daily tasks easier to complete, but the prevalence of digital platforms that make this possible also increases the risk, with hackers constantly trying to exploit flaws in a system and tamper with personal information, extorting money, subsequently engage in other malicious acts [3].

Cyber security is a field of ICT that is responsible for the protection of information assets, through Protection against threats that harm information, stored and transported by interconnected information systems [3], If cyber security is carried out after development is complete, the system is built in an insecure manner by bugs that are hard to fix. However, when security teams share knowledge and provide tools for team development and operations, the latter can modify systems and applications accordingly [4].

DevSecOps is about breaking security, further passing knowledge to different teams, and ensuring that security is implemented at the right level and at the right time [2]. [5]. DevSecOps can be defined as an approach to improve and accelerate the delivery of business value by making dev and ops team collaboration effective.

II. RELATED RESEARCH

A. *Development Security Operations (DevSeeOps)*

Previous research on DevSecOps, revealed that culture, automation, measurement and sharing (CAMS) are important factors to consider, in a fashion similar to DevOps. So, organizations cannot simply buy or lease their services to DevOps, and the same goes for DevSecOps. In fact, culture has been recognized as an important part of both, but DevSecOps emphasizes the importance of creating security [8].

B. *Static Analysis and Security Testing (SAST)*

Several studies have also shown that better ratios for identifying right and wrong can be obtained by combining different types of methods to take advantage of different synergies. This shows how a combination of methods can reduce the identification of positive (true vulnerabilities detected) and negative (true vulnerabilities not found). The analyzed work concluded that any security vulnerabilities included in the AST tool report, including manual reviews, should be verified. positive identification is actually harmless and can be corrected in security analysis. However, negative identification is more difficult to find if the previous method does not have the ability to detect it, causing a real danger. these methods include the use of static white box security analysis (SAST), dynamic black box security analysis (DAST), or ar interactive white box security analysis (IAST). Manual analysis requires highly specialized staff and time. To perform a web application security analysis, using any method, it is necessary to cover the entire attack that accesses

parts and layers of the application and use methods to automate security analysis as much as possible [9].

TABLE I. SAST VULNERABILITIES CATEGORIES

CWE	Detections
CWE-79 XSS persistent	44
CWE-79 XSS reflected	244
CWE-94 dangerous file inclusion	20
CWE-676 dangerous function	2
CWE-95 code injection	5
CWE-91 JSON injection	2
CWE-321 hardcoded encryption key	1
CWE-601 open redirect	16
CWE-259 hardcoded password	2
CWE-22 path manipulation	48
CWE-359 privacy violation	17
CWE-89 SQL injection	1
CWE-497 system information leak	1
Total	**403**

III. THEORETICAL BACKGROUND

A. *Development Security Operations (DevSecOps)*

DevSecOps is relatively new to the field of information security. The fundamental idea aligns with the concept of having security as an integral part of software development principles, processes and methodologies. The DevOps model is rapidly being adopted by the technology industry to support the need to develop and release core business systems and applications to customers in a much faster and reliable manner than the software development life cycle (SDLC) model traditionally followed. The security industry has adapted to the DevOps demand by introducing relevant processes in the form of DevSecOps principles and methods without affecting the original intent of DevOps. The author will review how security processes can be effectively embedded in a DevOps model to improve the success of IT projects within an organization. However, this article is not meant to review how to adopt DevOps to its advantage over traditional approaches[9].

B. *Analysis Security Testing*

There are various types of testing techniques that auditors or security analysts can choose from performing web application security analysis, static white box security analysis (SAST), dynamic black box security analysis (DAST) or interactive white box security analysis (IAST) techniques [10]. The OWASP Security Testing Guide Methodology v4.1 [11] suggests that to perform complex web application security analysis it is necessary to be automated as best as possible using static, dynamic and interactive analytical testing tools, including manual checks to find more actual attacks.

C. *Static Analysis Security Testing*

SAST tools will perform a security analysis of the source code of an application program, starting with determining whether an application program is ready for use or not [12]. Apart from these problems, the SAST tool is considered the most important security activity in SSDLC [13].

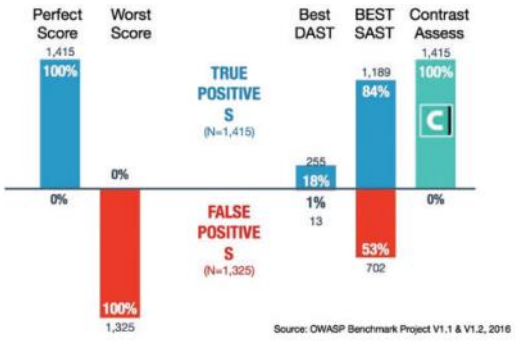


Fig. 1. SSDLC security techniques [15]

The SAST tool analyzes the entire application covering all attacks. Static analysis requires completing a manual audit of the results to discard positive identifications and find negative identifications (much more). However, some assert that different SAST tools have different algorithms by design as Interpretation. Therefore, combining SAST tools can find different types of vulnerabilities and therefore get better combination results [16].

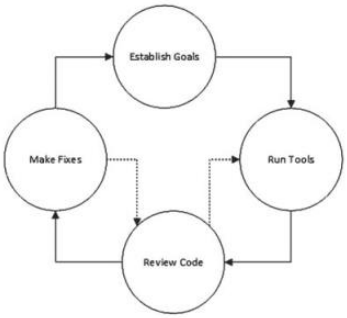


Fig. 2. Static analysis process [17]

IV. THE PROPOSED MODEL

This research is based on increasing security in the development of Information Technology, with the right way to take corrective steps in using various types of tools to detect security vulnerabilities through the following three steps: [3]

a. Find areas where security fails to reach in building reliable monitoring and warning systems.
b. Take measurable action to reduce or eliminate the negative impact of security issues detected on both application platforms.
c. Evaluate the results of each corrective action by comparing the situation before and after each phase.

A. *Monitoring and Warning*

Good monitoring to monitor various types of attacks against all application programs used, in order to have a reliable validation by recording all states that exist in each phase, but also plays an important role in providing clear transparency in the functionality of all security measures implemented. In this phase, we will install a metric scraper on each asset depending on the functional type and security metrics we are looking for that specific asset. Accurately running a system of alerts and rules based on metrics that the monitoring system collects from different assets is also another important thing that we should use to create a

successful DevSecOps file in the exchange of corporate information. [18]

B. Application Protection

Software composition analysis can be applied holistically to ensure that each open source dependency has a compatible license and is free of vulnerabilities. A behavioral by-product of this is that developers feel a sense of ownership over the security of their applications, getting direct feedback on the relative security of the code they write.

Once the code is checked and generated, you can start using security integration testing. Running code in an isolated container sandbox enables automated testing of things like network calls, input validation, and authorization. This test results in rapid feedback, enabling rapid iteration and triage for any identified issues, causing minimal disruption to the overall flow. If things like unexplained network calls or unclean input occur, the test fails, and the pipeline generates actionable feedback in the form of reporting and notification to the relevant team.

This can happen if the AST tool has a large capacity image or has to store data in a container. But the fact is that AST tools do take some time to run, and they can slow down the overall CI/CD pipeline. Interestingly, the CI/CD pipeline was never conceptualized with security in mind, but rather speed and convenience[19]. Here are good steps:

a. Hosting code written in secure and reliable source code, such as GitHub and Git, which can control the version files of the code sent to the associated repository and allow to quickly roll back to a previous version of the code if the code doesn't want to be pushed to the repository.

b. Run a static code analyzer against all newly written code and third-party libraries they use; This scan should also run automatically on each application deployment. As soon as vulnerabilities are detected during application deployment, the deployment work should end immediately and not go to production. The result of this type of failed implementation should be announced to developers and product managers, immediate feedback.

c. For individual developers about failed implementations can be a great source of truth to understand why application implementations fail. Loosely integrating with Jenkins in the job deployment process can fulfill this goal. Having detailed information about the results of running the static code analyzer with new application code in each new deployment will increase the visibility in the background of the deployment pipeline to a decent level

V. CONCLUSION

Using a variety of tools to detect different security vulnerabilities helps developers and organizations to safely release applications, reducing the time and resources that must later be devoted to fixing errors. During the secure software development cycle of an application where vulnerability detection tools help integrate security. Results Correlation between tools with different types is still an aspect that is not too broad. For this reason, it is necessary to develop a methodology or software that allows custom made automatically or semi-automatically for the evaluation and correlation of the results obtained with several different types of tools. It is very important to develop representative tests to perform a series of vulnerability tests included in OWASP and then combine them with SAST tools.

REFERENCES

[1] Issie Lapowsky. 2019. In Latest Facebook Data Exposure, History Repeats Itself. Wired (March 2019).

[2] Issie Lapowsky. 2019. In Latest Facebook Data Exposure, History Repeats Itself. Wired (March 2019).

[3] Matt Powell. 2019. 11 Eye Opening Cyber-Security Statistics for 2019. CPO Magazine (June 2019).

[4] K. Carter, ‘‘Francois Raynaud on DevSecOps,’’ IEEE Softw., vol. 34, no. 5, pp. 93–96, Sep. 2017. doi: 10.1109/MS.2017.3571578.

[5] Pulasthi Perera, Roshali Silva, and Indika Perera. 2017. Improve software quality through practicing DevOps. In 2017 Seventeenth International Conference on Advances in ICT for Emerging Regions (ICTer), 1-6.DOI:https://doi.org/10.1109/ICTER.2017.8257807

[6] Information Systems Audit and Control Association (ISACA). Accessed: May 20, 2017. [Online]. Available: http://www.isaca.org/KnowledgeCenter/Documents/Glossary/Cybersecurity_Fundamentals_glossary.pdf

[7] V. Mohan and L. B. Othmane, ‘‘SecDevOps: Is it a marketing buzzword?- Mapping research on security in DevOps,’’ in Proc. 11th Int. Conf. Availability, Rel. Secur. (ARES), Salzburg, Austria, Aug./Sep. 2016, pp. 542–547. doi: 10.1109/ARES.2016.92.

[8] Pulasthi Perera, Roshali Silva, and Indika Perera. 2017. Improve software quality through practicing DevOps. In 2017 Seventeenth International Conference on Advances in ICT for Emerging Regions (ICTer), 1–6. DOI:https://doi.org/10.1109/ICTER.2017.8257807

[9] Higuera, J.B.; Aramburu, C.A.; Higuera, J.-R.B.; Sicilia, M.-A.; Montalvo, J.A.S. Systematic Approach to Malware Analysis (SAMA). Appl. Sci. 2020, 10, 1360.

[10] Jeganathan S, 2019 DevSecOps: A Systemic Approach for Secure Software Development, ISSA Journal

[11] Felderer, M.; Büchler, M.; Johns, M.; Brucker, A.D.; Breu, R.; Pretschner, A. Security Testing: A Survey. In Advances in Computers; Elsevier: Cambridge, MA, USA, 2016.

[12] OWASP Foundation. OWASP Testing Guide, 2020. Available online: https://owasp.org/www-project-websecurity-testing-guide/ (accessed on 9 mei 2021).

[13] Sipser, M. Introduction to the Theory of Computation, 2nd ed.; Thomson Course Technology: Boston, MA, USA, 2006.

[14] Mohino, J.D.V.; Higuera, J.B.; Higuera, J.-R.B.; Montalvo, J.A.S.; Higuera, B.; Mohino, D.V.; Montalvo, J.A.S. The Application of a New Secure Software Development Life Cycle (S-SDLC) with Agile Methodologies. Electronics 2019, 8, 1218

[15] CMS Critic Awards (2020). CMS critic awards, https://www.cmscritic.com/awards/#bestopen-source-cms.

[16] Nunes, P.; Medeiros, I.; Fonseca, J.C.; Neves, N.; Correia, M.; Vieira, M. An empirical study on combining diverse static analysis tools for web security vulnerabilities based on development scenarios. Computing 2018, 101, 161–185.

[17] Chess, B., West, J. (2013). Secure programming with static analysis, USA: Addison-Wesley.

[18] Srinivasu N, On the viability of adaptive paris metro pricing in agent based model for federated clouds, International Journal of Recent Technology and Engineering (2018)

[19] Jyoti B. Kulkarni, Dr. Manna Sheela Rani Chetty, “Depth Map Generation from Stereoscopic Images Using Stereo Matching on GPGPU”, Journal of Advanced Research in dynamical and Control Systems, Volume 9, ISSN 1943- 023X, Special Issue – 02 / 2017.

Investigating the Effect of Climate on National Rice Production using Machine Learning

1st Rizkyana
Departement of Computer
Politeknik Caltex Riau
Pekanbaru, Indonesia
rizkyana20s2tk@mahasiswa.pcr.ac.id

2nd Yohana Dewi Lulu Widyasari
Departement of Computer
Politeknik Caltex Riau
Pekanbaru, Indonesia
yohana@pcr.ac.id

***Abstract*—This paper examines the influence of climate on national rice productivity using K-Mean Clustering through the Fact Constellation scheme. The research was conducted to see the effect of climate on national rice productivity, especially on the island of Sumatra. For the clustering, the popular machine-learning tool K-Nime was used whose visualization feature was principally useful to determine the patterns, dependencies, and relationships of rice yield on different climate and soil factors of rice production. Research shows that several climate components such as temperature, humidity and solar radiation have a very strong influence on national rice productivity.**

Keywords—K-Means Clustering, Fact Constelation, K-nime, Visualization, Climate, National Rice Productivity

I. INTRODUCTION

Indonesia is an agrarian country with a very wide agricultural land reaching 36,817,086 ha. Of the total area of agricultural land, the land used for rice fields in 2018 was 11.377.934 ha, then in 2019 it was recorded that the area had decreased to 10,677 887 ha. [1,2]

National rice production in 2018 was recorded at 59.200.534 tons of dry milled rice which is equivalent to 33.942.865 tons of rice. In 2019, rice production decreased to 54.604.033 tons, equivalent to 31.313.034 tons of rice. Statistical data from BPS Indonesia shows that there are ups and downs in rice production from year to year [1,2].

Rice is the main food of the Indonesian people. With a population of 270.2 million people, the need for rice is also very large and is increasing from year to year in line with the increase in population every time.

Given the importance of rice for national food security, efforts to increase rice production must be one of the focuses of development. This research is very important to investigate which factors have an influence on production which are responsible for increasing rice productivity in certain area.

It is evident that determining the most significant factors of rice production will be greatly beneficial not only to the consumers of the crops but indirectly, also to the producers as well. The farmers often do not get the proper price for their hard work in the harvests because of the "middle-men" who act as the connecting string between the root level producers of the crops and the potential buyers of bulk quantities of those crops. If such entrepreneurs could know which locations of rice productions have a greater impact on their annual rice yield based on controllable factors like use of fertilizers and other non-climatic factors, then they might want to invest in those areas. This research tries to fill-in those gaps between producers and consumers so that both the parties can be benefitted. This paper conveys the findings in five sections. The first section, Introduction, works as the introduction to the research"s domain and establishes its significance. The second section, Related Works, explores the domain of rice production in Sumatera and reconnoiters the different approaches commonly taken by other researchers to attain results similar to or within the domain of this research. This section also contains brief discussion on drawbacks of those approaches and what alternatives to them could be used. The third section, Proposed Methodology, portrays the proposed solution to the problem established in the first section, along with an elaborate discussion on how the data were pre-processed and then how they were analyzed. The fourth section, Results and Discussion, describes the findings from the third section and discusses them in light of the Related Works for each variety of rice. Finally, the fifth section, Conclusion, and Future Works, wraps up the research and critically discusses the findings along with providing other scopes of future works where the findings from this research may be applied directly to or implied in an ancillary form.

II. DATA, MACHINE LEARNING AND AGRICULTURE

A. Utilization of Big Data Analytics and Machine Learning for Agricultural Data Processing

Agricultural developments where large amounts of data processing are needed with many variables are the impetus for utilizing big data analytics and machine learning to solve them. Dynamically changing agricultural conditions such as climate, soil conditions and agricultural production should be understood to gain a better understanding. This dynamically changing data certainly produces data in large quantities and variations. Big Data and Machine Learning as high technology in the field of data processing has the advantage of describing, measuring and understanding agricultural processes using data. However, there are several challenges to integrating this technology, so it is necessary to make some machine learning adaptations to use it with Big Data. [3]. The effective and efficient use of Big Data, Machine Learning and Internet of Things computerized technology in rice smart farming systems (Rice Smart Farming) in every production and post-production activity of rice is a very crucial step that will transform traditional rice cultivation practices into a new

perspective. new careful and smart rice cultivation [4]. Big data and machine learning are used for agriculture in the following ways [5]:

1. Weather and Climate Change
2. Agricultural Land
3. Animal Research
4. Crops
5. Soil
6. Weeds
7. Food avibility anda Food Security
8. Biodiversity
9. Farmer decision making
10. Farmer's insurance dan finance
11. Remote sensing

B. Fact Constelation

A data warehouse conceptual model or dimensional model is a logical design that represents data in a standardized form, and supports quick access to data. There are several terms related to this model, namely fact table, dimension table, and hierarchy. Fact table is the main table that contains a collection of the primary keys of other tables. Every fact table in the conceptual data warehouse model has a composite key, and vice versa, every table that has a composite key is a fact table. In other words, any table that shows a many-to-many relationship must be a fact table. A dimension table is a simpler table in which there is a primary key that corresponds to one of the composite keys in the fact table. The hierarchy defines the order of mapping from lower-level concepts to higher-level, more general concepts.[6]. There are several models that are often used, namely the star scheme, the snowflake scheme and the most popularly used is the fact costelation scheme. The fact constellation schema is very flexible but can sometimes be difficult to setup and support. The fact constellation schema allows a dimension table to be related to many fact tables so that complex systems are the most frequently use.

C. K-Mean and Clustering

The k-means algorithm is a widely used method that starts with the initial partition of the data and then converges repeatedly to a local solution by reducing the number of squared errors (SSE). Clustering is a data mining technique used to analyze data that has variations. and the number of lots. In the Clustering process, data is grouped into a cluster, so that it contains data that is as similar as possible and different from other cluster objects. Researchers will use a combination of the K-Means method with elbows to increase the efficiency and performance of effective k-means in processing data in large numbers. K-Means Clustering is sensitive to the selection of the initial position of the midpoint of the cluster. So choosing the initial position of the midpoint of the bad cluster will result in the K-Means Clustering algorithm which produces high errors and poor cluster results. means has a problem determining the number of clu the best ster so that it is often combined with the Elbow method to find the best number of clusters in the K-means method. Based on the results obtained from the process of determining the best number of clusters with the elbow method, it can produce the same number of K clusters in different amounts of data. The result of determining the best number of clusters with the elbow method will be the default characteristic for the process in the case study [15].

D. Related Work

An end-to-end model based on deep learning fusion is able to accurately predict rice yields for 81 districts in the Guangxi Zhuang Autonomous Region, China, using a combination of time series meteorological data and area data. Model testing shows that the model is accurate for predicting rice yields in summer and winter [8]. Machine Learning is also used to predict the application yield level using the Naif Bayes method and the K-Nearest method on certain plants, certain places and certain times [9]. Process-based and data-based models can be used to provide early warning to anticipate rice disease outbreaks and detect their presence, thus supporting fungicide applications [10]. These data-driven models are derived from machine learning methods, are viable alternatives to process-based approaches and – in cases where training data sets are available – offer potentially greater adaptability to new contexts. The application of Machine Learning to predict rice yields was carried out in 27 districts in the state of Maharashtra, India using the WEKA tool for SMO Classification [11]. Parameters considered for the study were rainfall, minimum temperature, mean temperature, maximum temperature and evapotranspiration of reference plants, area, production and yield for the Kharif season (June to November) for 1998 to 2002. The experimental results show that the technical performance others on the same dataset are much better than SMO. Machine Learning system is proposed for two crops (maize and soybean) and two geographies (Brazil and US) which is able to provide estimates of preseason yields, which are formed by a neural network in which the inputs are treated separately [12]. This particular architecture is trained with historical data for several soil properties, rainfall, minimum and maximum temperatures against historical yield labels in the two countries. System testing shows results that are comparable to existing methods.

III. METHODOLOGY

The data in this study were obtained from various sources. The condition of the data, the preprocessing of the data and how this data is processed are explained as follows.

A. Structure of Collected Data

Rice data are taken from the Central Statistics Agency (BPS). The data taken are data on the area of rice fields in each province in Indonesia, data on total rice production, and rice productivity per hectare. The range of data taken starts from 2019 and 2020 [1,2].

Climatic data taken in the form of data on rainfall, temperature, humidity and duration of sunlight. This data is recorded at every weather observer station (BMG station) as many as 84 stations spread throughout Indonesia from 2019 to 2020. The data used in this study are data contained in the publication of Climate Statistics, Plant Pest Organisms and Climate Change Impacts in 2017-2020 of the Ministry of Agriculture in 2020 [13]

B. Data Cleaning and Preprocesing

Data on rice production, paddy field area and productivity per hectare starts from 2019 to 2020. The climate data is also taken in the same time span. To synchronize the data, data

sorting and initial data processing are carried out [14] such as rice production statistics in 2019 and 2020 which are processed from initial production data every month and sometimes the accumulation of several months is processed and added up to become one year's rice production. Especially in 2020, the data recorded is only from January to April so it needs to be processed or deleted.

The same thing is also done on climate data where climate records are taken from various weather observation stations spread across all provinces in Indonesia. Climate data per province is taken as the average of climate measurement data by several weather observation stations in that province. The monthly climate data recorded are converted into annual average data so that they are in sync with rice production data. The steps are as follows:

1. Data preprocessing is a process or step taken to make data raw data into quality data. First step, check first the data using the K-nime application:

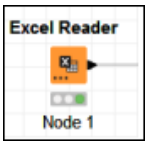


Fig. 1. Data Preprocessing

2. Data preprocessing is an early data mining technique for converting raw data or commonly known as raw data collected from various sources information that is cleaner and can be used for further processing. The goal is to transform data into a format that is easier to process and effective for user needs. The steps taken are as follows :

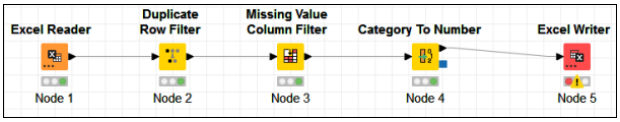


Fig. 2. The Steps Data Processing

3. So that the results look good, then I determine the K range first. Based on the elbow plot, the results obtained are very clear elbow bends as follows :

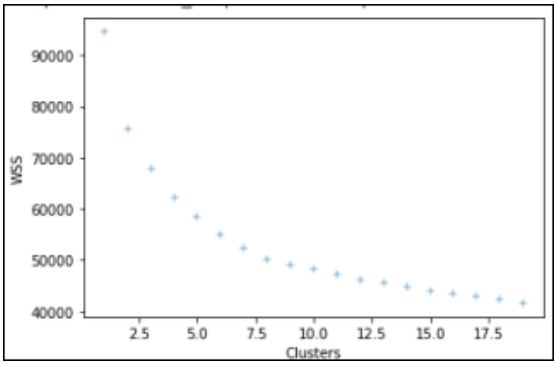


Fig. 3. The K Range First

4. Before the classification process is carried out using the KNN algorithm, then the preprocessing process is carried out, namely changing the cluster value from the cluster category which is still of type string is modified to be numeric. In this experiment, in the K-nime application, change the Category To Number as follows:

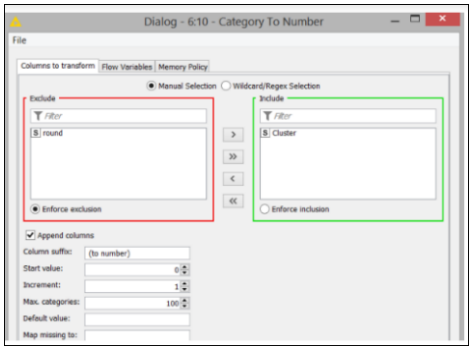


Fig. 4. Category To Number

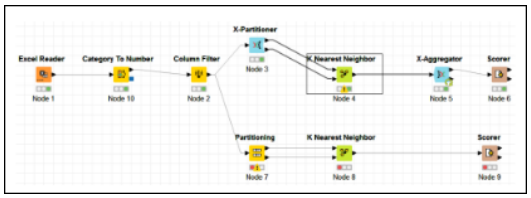


Fig. 5. The Category Of Number

5. The next process, is the validation process:
The validation process uses Cross Validation on the KNN algorithm, where the excel file used is excel which has been processed by k-means from K=2 to K=10. Then do the preprocessing process using node partitioning with 80% division for training data and 20% for data testing. With neighbor value k = 2

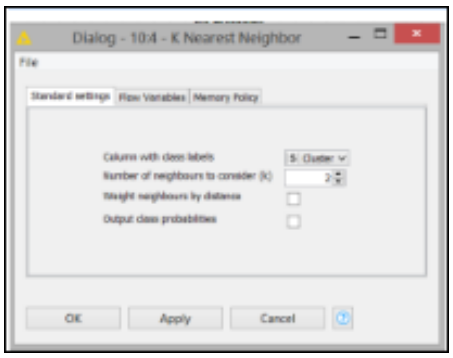


Fig. 6. Validation Process

6. The validation process uses K-fold validation on KNN, where the excel file used is excel which has been processed by k-means with K=2 up to K=10. In the K-fold validation process, K-fold = 10 is carried out and random sampling.

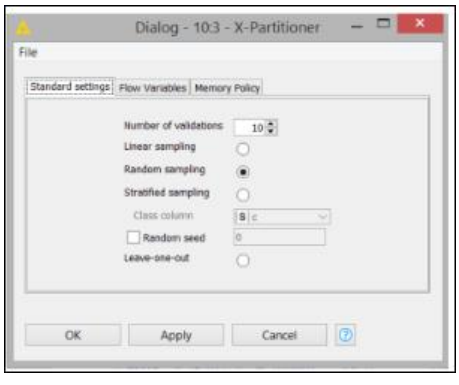


Fig. 7. Validation Process Uses K-Fold

With Neighbor K=2

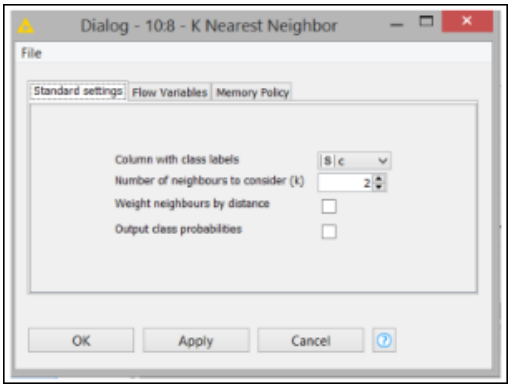


Fig. 7. Validation Process Uses K-Fold With Neighbor K=2

7. On data analysis to look for hidden patterns or groupings in the data. To create a cluster using the K-means algorithm.

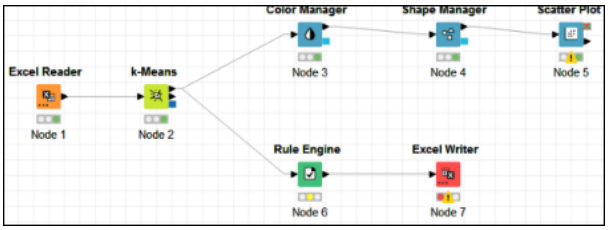


Fig. 8. Cluster

C. Fact Constelation Scheme and Data warehousing

Data related to rice production that is synchronized so that it is homogeneous is then compiled into a data warehouse using the Fact Constellation scheme. Figure 1 shows a part of the schematic design of this research Fact Constellation. There are two dimensions in this scheme, namely the location dimension and the time dimension and 10 fact tables that are adjusted to the number of provincial locations analyzed. Included in the fact table are rainfall, temperature, humidity and duration of sunlight because it is tempo-spatial which is always dynamic and changes all the time [8].

IV. Results and Discussion

Data related to rice production that is synchronized so that it is homogeneous is then compiled into a data warehouse using the Fact Constellation scheme. Figure 1 shows a part of the Fact Constellation schema that has been designed.

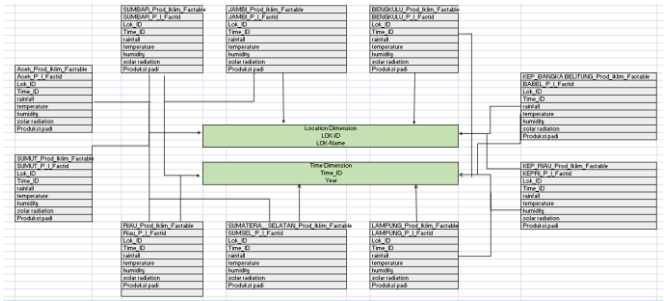


Fig. 9. Fact Constelation Scheme

The results of the visualizations were as follows.

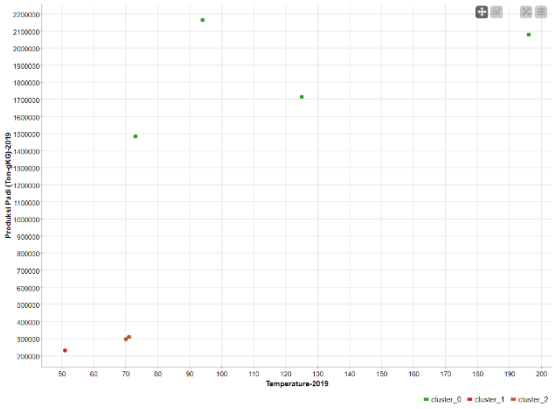


Fig. 10. Effect of Temperature on Rice Production in 2019

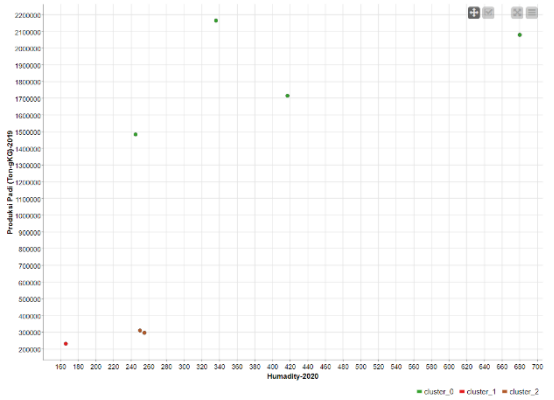


Fig. 11. Effect of Humidity on Rice Production in 2019

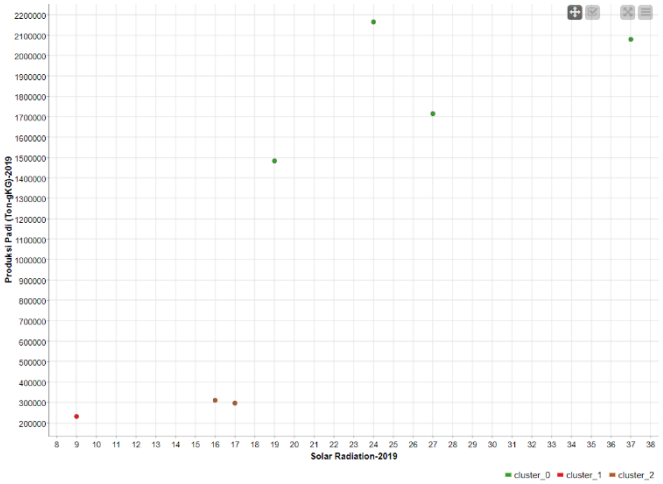


Fig. 12. Effect of Solar Radiation on Production Rice in 2019

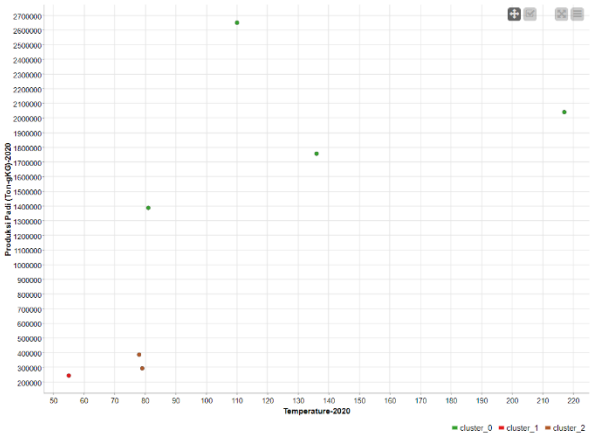


Fig. 13. Effect of Temperature on Rice Production in 2020

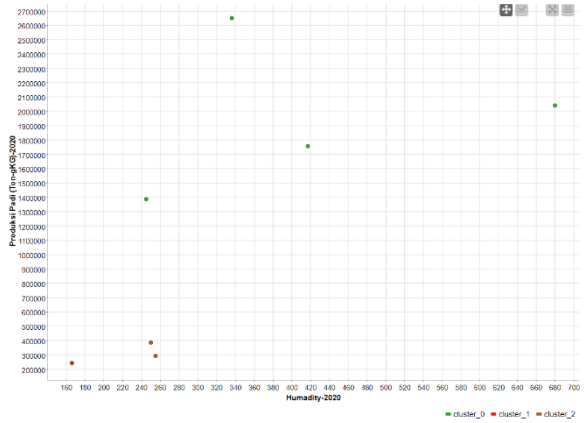


Fig. 14. Effect of Humidity on Rice Production in 2020

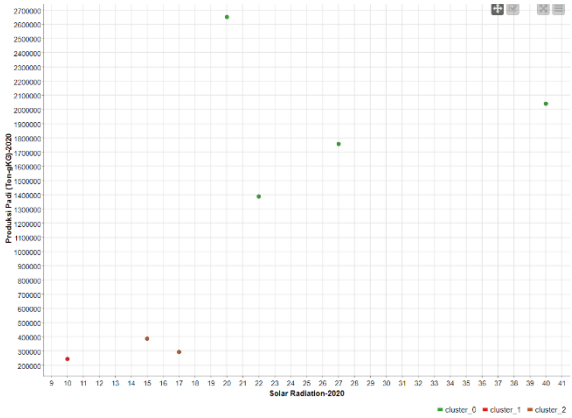


Fig. 15. Effect of Solar Radiation on Rice Production in 2020

As it can be seen from the fig. 2 to 7, even though the clustering algorithm simply created K clusters of data points based on their relative characteristics, using the visualization tool in K-Nime, we could individually correlate the variation of each factor, whether it be climate or soil factor, against rice yield values and directly deduce their relationship with each other from the visualizations themselves. In each of the visualizations, the x-axis represented a factor that would influence the rice production and on the y-axis are the yield values. The clustered data points have been classified based on their yield values and thus have been colored according to the yield value they represent for easier understanding of relationship between the factor under observation and the yield, such that, for data points lower yield values the color of the centroid will be more towards „blue‟ while for data points that correspond to a higher yield value will have a color more towards the yellow end of the spectrum. The color spectrum has been shown for each visualization under “Class Colour‟.

From the cluster visualization of the influence of climate on rice production, it can be analyzed that the climate factor (weather); which is represented by temperature, humidity and solar radiation have a proportional impact throughout 2019 and 2020. During 2019 and 2020, temperature did not experience a significant change in pattern beyond the ability of rice plants to grow so that it had a negative effect on rice productivity. also not visible. In this temperature cluster, it is observed that the temperature in these years has a linear impact on rice productivity. There was a slight change in the length of irradiation and humidity in 2019 and 2020, both of which had an observable impact on rice productivity. The duration of solar radiation has decreased which has a negative effect on rice productivity. The effect of solar radiation is also proportional to humidity so that the impact caused by changes in humidity is also proportional to changes in solar radiation.

V. Conclusion

From the various descriptions above, it can be concluded that the use of machine learning for rice farming has been successfully carried out. The study above shows that the climate in this case temperature, humidity and duration of irradiation has a proportional influence on rice productivity on the island of Sumatra. This can be seen from the cluster visualization of the influence of climate on rice production, climatic factors (weather) can be analyzed; represented by temperature, humidity and solar radiation. Research on the factors that affect rice productivity is still very wide open, for example research on the effect of rainfall, soil acidity, nutrient levels in the soil, dry season, and others.

References

[1] BPS dan BPPT Indonesia, “Ringkasan Eksekutif Luas Panen Dan Produksi Padi Di Indonesia 2020.”

[2] Badan Pusat Statistik Indonesia, “Indikator Pertanian 2020,”

[3] J. Clerk Cravero, Ania; Sepúlveda, Samuel. 2021. "Use and Adaptations of Machine Learning in Big Data—Applications in Real Cases in Agriculture" Electronics 10, no. 5: 552, https://doi.org/10.3390/electronics10050552Maxwell,

[4] R. Alfred, J. H. Obit, C. P. -Y. Chin, H. Haviluddin and Y. Lim, "Towards Paddy Rice Smart Farming: A Review on Big Data, Machine Learning, and Rice Production Tasks," in IEEE Access, vol. 9, pp. 50358-50380, 2021, doi: 10.1109/ACCESS.2021.3069449

[5] Andreas Kamilaris, Andreas Kartakoullis, Francesc X. Prenafeta-Boldú, “A review on the practice of big data analysis in agriculture,” Computers and Electronics in Agriculture, Volume 143, 2017, Pages 23-37, https://doi.org/10.1016/j.compag.2017.09.037.

[6] Doro Edi , Stevalin Betshani, “Analisis Data dengan Menggunakan ERD dan Model Konseptual Data Warehouse ,” Jurnal Informatika, Vol.5, No. 1, Juni 2009: 71 – 85.

[7] M A Syakur, B K Khotimah, E M S Rochman1 and B D Satoto, “ Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster”, IOP Conference Series: Materials Science and Engineering, Volume 336, The 2nd International Conference on Vocational Education and Electrical Engineering (ICVEE) 9 November 2017, Surabaya, Indonesia

[8] Zheng Chu, Jiong Yu, An end-to-end model for rice yield prediction using deep learning fusion, Computers and Electronics in Agriculture, Volume 174, 2020

[9] R. Medar, V. S. Rajpurohit and S. Shweta, "Crop Yield Prediction using Machine Learning Techniques," *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, 2019, pp. 1-5, doi: 10.1109/I2CT45611.2019.9033611.

[10] Nettleton, D.F., Katsantonis, D., Kalaitzidis, A. *et al*. Predicting rice blast disease: machine learning versus process-based models. *BMC Bioinformatics* **20,** 514 (2019). https://doi.org/10.1186/s12859-019-3065-1

[11] N. Gandhi, L. J. Armstrong, O. Petkar and A. K. Tripathy, "Rice crop yield prediction in India using support vector machines," 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2016, pp. 1-5, doi: 10.1109/JCSSE.2016.7748856.

[12] Igor Oliveira, Renato L. F. Cunha, Bruno Silva, Marco A. S. Netto IBM Research , "A Scalable Machine Learning System for Pre-Season Agriculture Yield Forecast", 2018 IEEE 14th International Conference on e-Science, pp:423-430

[13] Kementrian Pertanian, "Statistik Iklim,Organisme Pengganggu Tanaman dan Dampak Perubahan Iklim tahun 2020,"

[14] Fahad Ahmed, Dip Nandi, Mashiour Rahman, and Khandaker Tabin Hasan , "Investigating Factors that Influence Rice Yields of Bangladesh using Data," I.J. Modern Education and Computer Science, 2017, 3, 36-47, DOI: 10.5815/ijmecs.2017.03.05

[15] Hussain, S. F., & Haris, M. (2018). A k-means based Co-clustering (kCC) Algorithm for Sparse, High Dimensional Data. ," Iternational Journal Expert Systems with Applications, pp 22-26. doi:10.1016/j.eswa.2018.09.006

Credit Scoring Models and Applications Based on Personality Predictions Using Twitter Data and Debtor Big Data at PT. Bank Riau Kepri

1st Sugianto
Magister Terapan Teknik Komputer
Politeknik Caltex Riau
Pekanbaru, Indonesia
Sugianto20s2tk@mahasiswa.pcr.ac.id

2rd Dadang Syarif Sihabudin Sahid
Magister Terapan Teknik Komputer
Politeknik Caltex Riau
Pekanbaru, Indonesia
dadang@pcr.ac.id

3rd Juni Nurma Sari
Magister Terapan Teknik Komputer
Politeknik Caltex Riau
Pekanbaru, Indonesia
juni@pcr.ac.id

4rd Yohana Dewi Lulu Widyasari *Magister Terapan Teknik Komputer*
Politeknik Caltex Riau
Pekanbaru, Indonesia
yohana@pcr.ac.id

***Abstract*—The development of social media in Indonesia is very fast, even the latest data shows that social media users in Indonesia have increased from year to year. The total active social media is 160 million or 59% of the total Indonesian population aged 16 to 64 years. 99% of social media users via mobile. The most widely used social media is the Twitter platform. Rapid development, many business lines are starting to use social media analysis to see the personality of users. This phenomenon is called personality analysis by utilizing Big Data. In the internal of Bank Riau Kepri itself, there are no tools that can be used to analyze a person, including the analysis of prospective debtors.**
Therefore, debtor data in the Core Banking System at Bank Riau Kepri internal and tweet data on the twitter platform will be analyzed using Big Data using a machine learning model with the application of the Decision Tree and Random Forest algorithms. This analysis aims to see the personality of prospective debtors by utilizing the Twitter platform social data media combined with big data from Bank Riau Kepri debtors to see the character, capacity, and capital. After the analysis is done, testing is done on the model built by performing Split validation and cross validation to determine the level of model accuracy.
The end result will help to see the credit analysis of the prospective debtor, which will be visualized in the application in the form of credit scoring. Credit scoring using algorithms combined with Big Data shows a very good level of accuracy, as evidenced by several previous studies.

Keywords—credit scoring, Decision Tree and Random Forest, big data, tools

I. INTRODUCTION

In Indonesia, the development of social media is very fast, even the latest data shows that social media users in Indonesia have increased from year to year. Social media users in Indonesia are 16 to 64 years old. For information, the average time Indonesian people spend accessing social media is 3 hours 26 minutes. The total number of active social media users is 160 million or 59% of the total population of Indonesia. 99% of social media users surf via mobile [1].

With the rapid development of social media, millions and even billions of data are created every day, this phenomenon is called Big Data. The big data phenomenon is now being used and developed by various large companies in the world, one of which is in the banking sector in supporting business aspects of the company, such as: customer relationships, customer segmentation, credit loan assessment, and marketing strategy [2].

Changes in the credit process were followed by changes to the credit system policy (credit policy), as well as changes to the organization and human resources both at the Head Office and at branches. This change then becomes the background of this research, namely whether the implementation of Credit Scoring can improve credit quality? Increasing the effectiveness of the credit process, especially in terms of controlling the emergence of non-performing loans at PT. Bank Riau Kepri. What is the credit scoring model and application like at PT. Bank Riau Kepri by utilizing social media data, especially Twitter to predict personality combined with big data of debtors at PT. Bank Riau Kepri.

II. THEORY

A. Credit

According to Thomas. S, et al. The term credit comes from the Greek "Credere" which means trust, therefore the basis of credit is trust. A person or all agencies that provide credit (creditors) believe that the recipient of credit (debtor) in the future is able to fulfill everything that has been promised, in the form of goods, money or services [3]

B. Principles of 5C Analysis on Credit

Before a credit facility is given to a prospective credit recipient, the bank must feel sure that the credit given will actually return. This is because the bank wants to minimize the risks that arise. This confidence is obtained from the results of the credit assessment before the credit is disbursed. In providing credit, banks must also consider several things related to the good faith (willingness to pay) and ability to pay (ability to pay) of the customer to repay the loan along with the interest. Banks in disbursing credit adhere to the principle of prudence. Before the credit is approved for issuance, the credit officer (credit analyst) at the bank must perform a credit analysis first. The purpose of this credit analysis is to convince the bank that the prospective debtor is really

trustworthy. One way of credit analysis is to apply the 5C analysis principle which includes Character, Capacity, Condition, and Capital Collateral [12].

C. Principles of 5P Analysis on Credit

The application of the 5P analysis principle in credit analysis is as follows:

1) Personality

The bank looks for data on the personality of the prospective debtor such as his life history, hobbies, family circumstances, social standing, and other things that are closely related to the personality of the borrower.

2) Purpose

The bank looks for data about the purpose or need for using credit

3) Prospect

The bank looks for data about the future expectations of the borrower's line of business or business activities

4) Payment

Banks are looking for data on how the estimated loan repayments will be provided

5) Party

Party (class) of prospective bank borrowers need to classify prospective debtors into several groups according to character, capacity and capital. This classification will give direction to the bank's analysis of how it should behave [14] [15] [16].

D. Principles of 3R Analysis on Credit

The application of the principle of 3R analysis in credit analysis is as follows:

1) Returns

Returns means an assessment of the results to be achieved by the debtor's business after being assisted by loans from creditors. Whether the results achieved can cover for loan repayments and the business run by the debtor continues to grow or not?

2) Repayment

The creditor must assess how long the debtor's business can repay the loan in accordance with the repayment capacity and whether the credit must be paid in installments or paid off all at once at the end of the period.

3) Risk bearing ability

What is meant is that the creditor must know and assess the extent to which the debtor's business is able to bear the risk of failure if something unexpected happens. By having strong capital, the debtor's business will usually be stronger in facing or winning the competition with other parties. In addition, the ability to bear risks is also on the creditor's side, namely by asking for guarantees or collateral from the debtor [17] [18]

E. Credit Scoring

Credit scoring is one of the assessment methods used by financial institutions such as banks to help make decisions in approving credit applications from borrowers by looking at the risks involved, such as whether the borrower will repay the loan according to the agreement made or not [4].

F. FICO Score

The FICO Score is a three-digit number based on the information in the debtor's credit report. This helps the lender determine how likely the debtor is to repay the loan [13]. When a debtor applies for credit, the lender needs a quick and consistent way to decide whether or not to lend money to the debtor. A FICO score not only helps lenders make smarter and faster decisions about who they lend money to, it also helps people like debtors get access to fair and fast credit when debtors need it.

G. Social Media

Social media is an online platform that is used to build social networks or social relationships between users through various things they have in common, such as: interest in something, various kinds of activities, backgrounds, or similar life connections [5]. Meanwhile, according to El Badawy & Hashem, social media is defined as a platform that provides an opportunity for an individual to interact, communicate in two directions, and share opinions with other social media users [6].

H. Big Data

According to O'Reilly, big data is a collection of data that exceeds processing capacity and its structure is incompatible with conventional database architectures because the data is too large, moves too fast, and has to choose an alternative way to process it to get the value from the data [7]. Big data is described by Zhan & Fang as any data source that has at least three unique characteristics, or what is commonly referred to as 3 V's, namely: volume, velocity, variety [8].

I. Decision Tree

The explanation of the decision tree according to Swamy & Hanumanthappa, is an algorithm model that is represented as a two-way binary poshon which can show the value of the target variable (output) that can be predicted with a set of predictor variables (input) [9].

J. Bank

Banking is a mediation institution whose main task is to collect public funds and distribute in the form of credit [10].

III. RESEARCH METHODS

The framework used in this study was adapted from the previous research model from Guo et al [11].

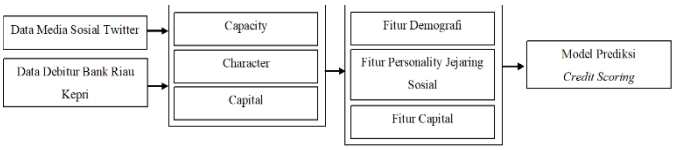


Fig. 1. Thought Framework, Source: Guo et al [11]

From the picture above the pipeline of this research framework consists of four steps:

A. Preprocessing Data

For data pre-processing, this study removed duplicate tweet records, tokenized text data, and filtered users with too few tweets. To be more specific, set a minimum limit of up to 21 tweets per user in experimental evaluation. The user dataset whose tweet count is lower than 21 can be leveraged to train Level 1 classifiers, which further results in high-level features

for users with more than 21 tweets. After preprocessing the data, the user demographics, tweet data, and network data for each user were obtained from the social data.

B. Low-level Feature Extraction

In this step, extract features from the three data aspect features available in our data set: demographic features, tweet features, and network features as well as ngram features and topics for assessing personality. As illustrated in the Figure, the low level feature is used as a classifier input for stacking as well as as an input. How to extract these features and detailed analysis of these features are described.

C. High-level Feature Generation

For some of the features for Credit Scoring features the stacking technique, which trains the classifier with a sample data set. Classifier predictions learned on other training data, such as a predicted label or a probability estimate, are used as a high-level feature. Apart from the formally defined ngram features, we also include a topic feature as well as a low-level feature from Step 2 for stacking to some extent for final prediction.

D. Ensemble Learning-based Prediction

To integrate various types of features into a unified credit scoring model, choose the machine learning decision tree method as the final prediction classifier for its outstanding speed, stability, and accuracy in performance. As mentioned in Step 3, use stacking to deal with the diversity and heterogeneity of social features, which is also a type of ensemble learning method. As illustrated in the section below, the classifier's inputs are of low and high level features.

TABLE I. TESTING METHOD

Tested Object	Technique	Purpose	Instrument
Model validation	Split validation (Training Error & Test Error)	Performs simple validation by randomly dividing the dataset into two separate data — training data & test data	Training error is obtained by calculating the misclassification of the model on the same data as the model being trained
	Cross validation	Performs iterative validation where the dataset is divided into many subsets (sets) of training & validation data	• k-fold cross validation • View the accuracy value of each data being tested
Web application system	Black Box Testing	In this test, the system will be tested to check whether the system can function properly or not	• Document (Software Requirements Specification) • Credit Scoring System
System Functionality Questionnaire	User Acceptance Test	In this test, it will be carried out in certain areas such as branch offices to check whether this system can be used properly	• User Acceptance Test Questionnaire • User / System User

IV. RESULTS AND DISCUSSION

A. Establishing a Credit Scoring Model

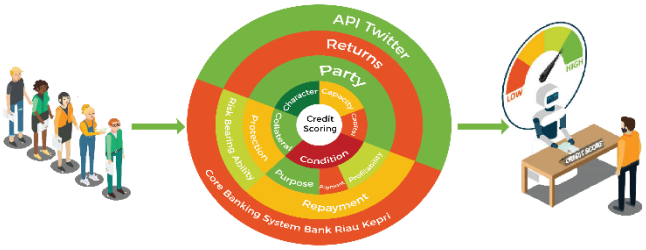


Fig. 2. Credit Scoring Model Design

From the picture above, the stages of forming the Credit Scoring Model are as follows:

1) Input Prospective Debtor Data into the System

The first step is to input debtor data into the Credit Scoring System. The data to be input is the identity data of the prospective debtor, income data, prospective collateral, and Twitter social media accounts.

2) Process Data Level 1

At this level, the data of prospective debtors that are input will be processed in 2 stages.

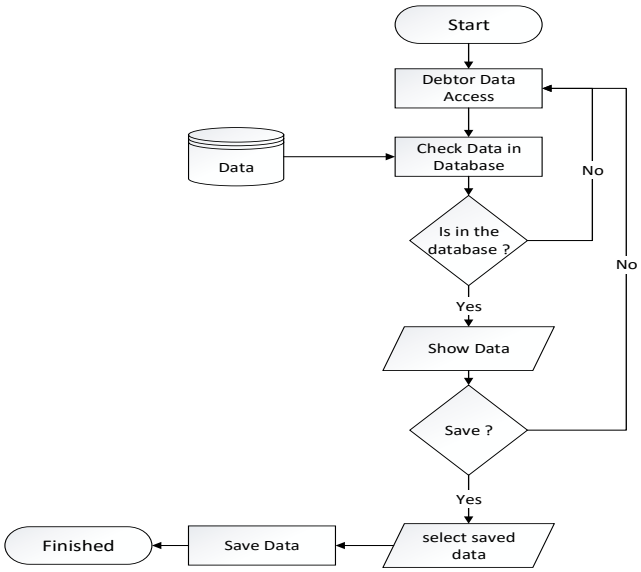


Fig. 3. Process flow of data retrieval to core banking system

First, personal identity data, income data, and prospective collateral will be checked using a machine learning model to

the big data Core Banking System of Bank Riau Kepri. The data that will be used at this stage is debtor data taken from the credit debtor database at PT. Bank Riau Kepri with a period of 2010 to 2021. The machine learning model will provide output whether the prospective debtor is recommended or not.

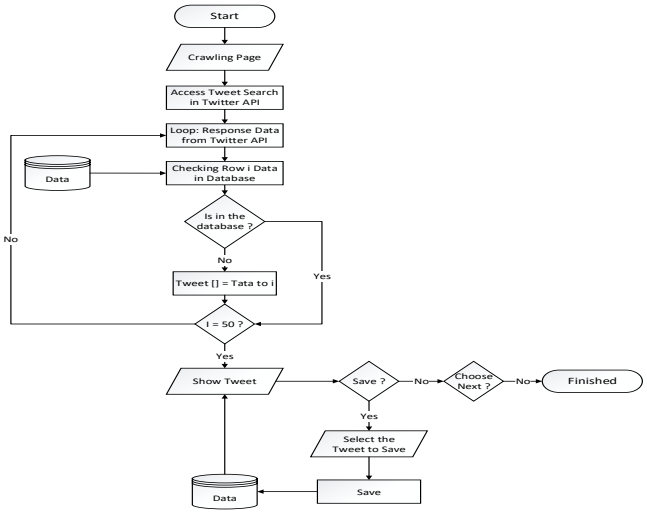


Fig. 4. The flow of the data retrieval process to the Twitter API

Furthermore, the Twitter social media account data will be checked into the Twitter database using the Twitter API (application programming interface). The output of this stage is in the form of recommendations from machine learning models that will provide recommendations or not.

3) Process Data Level 2

At level 2, the process at level one will check the database using a machine learning model that uses indicators of Risk Banking Ability and Repayment. For twitter social media accounts a machine learning model that uses the Returns indicator.

4) Process Data Level 3

At this stage, the stage 3 process will use deeper indicators to derive machine learning models. Data at level 2 Risk Banking Ability using Protection and Repayment indicators will use 3 indicators, namely: propose, payment, and profitability. While Returns uses the party indicator.

5) Process Data Level 4

At this stage, the data process at this stage will use a deeper approach to obtain machine learning models. Data at level 3 Protection will use the collateral indicator to get the model. Meanwhile, propose, payment, and profitability will use condition indicators to get the best model. Meanwhile, the party in terms of tweet data will use the Character, Capacity, and Capital indicators to recommend its machine learning model.

B. Establishment of a Credit Rating Model

TABLE II. ESTABLISHMENT OF A CREDIT RATING MODEL

Model	Fitur Input
Model 1	Demographic Features
Model 2	Social Network Personality Features
Model 3	Capital features
Model 4	Demographic Features + Social Network Personality + Capital

This study classifies the data features into five types of models as shown in the description. After the data has gone through the preprocessing process, the next step is to determine the creditworthiness threshold. An account can be said to be creditworthy if it has an aggregation value of attribute value weights in the credit scoring score that exceeds the specified threshold. Then, the data is processed using decision tree and random forest algorithms to evaluate the performance of its performance metrics.

C. Prediction Model

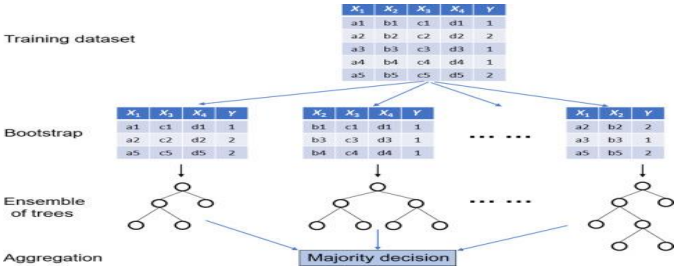


Fig. 5. Prediction Model

From the picture above, it is hoped that the results can compare between the two algorithms tested and the four models being tested. So it is expected to get the highest level of accuracy. Based on the results of the comparison of the four types of models that have been formed, model 4 which is processed using the random forest and decision tree algorithms has the highest performance evaluation value of performance metrics when compared to other models. From these results, it can be concluded as a whole that model 4, which consists of input data in the form of a combination of demographic features, social network personality, and capital which is processed using random forest and decision tree algorithms is the best model for credit scoring analysis based on social media personality data.

D. Architecture Technology Big Data

This study divides three important parts of the Big Data Architecture Technology, first, namely the data source which consists of the Core Banking System of Bank Riau Kepri and the API from Twitter data. The second part is at the analytical data stage on the big data system, namely ingestion and ETL, BI Analytics, and Advance Analytics. The last part is the presentation layer. This layer is where the analytical process takes place and determines the final value of the credit score.

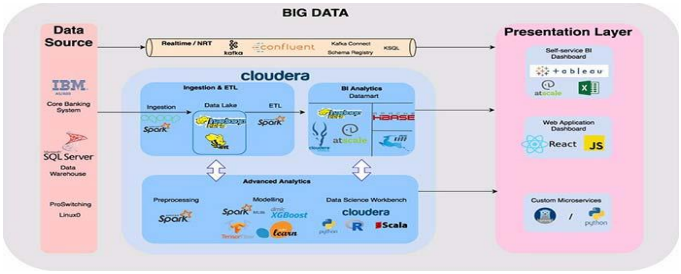


Fig. 6. Architecture Technology Big Data

The results of this study indicate that by utilizing digital data traces and content generated by each user of Twitter social media data by processing account profile data and tweets which are converted into social media personality data and

credit debtor data at PT. Bank Riau Kepri which is converted into demographic and capital data can provide useful insights and benefits for business activities, especially financial institutions such as banks in conducting credit scoring analysis in determining a person's creditworthiness.

V. CONCLUSION

From the research that has been done, the following conclusions can be drawn:

1. The process of establishing a credit scoring model based on social media data can be started by collecting and processing twitter account profile data into social media personality and credit debtor data at PT. Bank Riau Kepri which is converted into demographic and capital data, each of which has an insight that can represent the three credit principles proposed by the researchers, namely: character, capacity, and capital.
2. Based on the results of previous studies, it can be concluded that the appropriate predictive model to be used in conducting credit scoring based on personality data from social media and credit debtor data is to use the random forest algorithm and the decision tree shows the highest performance evaluation value of performance metrics.
3. Based on the results of previous studies, to test the level of accuracy of the model that was built, Split validation (Training Error & Test Error) and cross validation were carried out to determine the level of model accuracy..

REFERENCES

[1] D. H. Jayani, "10 Media Sosial yang Paling Sering Digunakan di Indonesia," *Databoks; Katadata.co.id*, pp. 1–1, 2020, [Online]. Available: https://databoks.katadata.co.id/datapublish/2020/02/26/10-media-sosial-yang-paling-sering-digunakan-di-indonesia.

[2] A. Ramaquita, A. Alamsyah, S. Prodi, M. Bisnis, F. E. Bisnis, and U. Telkom, "Model Credit Scoring Berdasarkan Data Demografi Dan Jejaring Sosial Di Media Sosial (Studi Kasus : Linkedin) (Credit Scoring Model Based On Demographic And Social Network Data In Social Media (Case Study : Linkedin))," vol. 7, no. 2, pp. 2214–2219, 2020.

[3] Grewal, D., & Levy, M. (2012). Marketing. New York: Mc Graw Hill

[4] Hand, D. J., & Henley, W. E. (1997). Statistical Classification Methods in Consumer Credit Scoring: a Review. 1.

[5] Akram, W., & Kumar, R. (2017). A Study on Positive and Negative Effects of Social Media on Society . International Journal of Computer Sciences and Engineering, 1.

[6] El-Badawy, T. A., & Hashem, Y. (2015). The Impact of Social Media on the Academic Development of School Students. International Journal of Business Administration, 1-2

[7] O'Reilly. (2012). Big Data Now: 2012 Edition. Sebastopol: O'Reilly Media Inc.

[8] Zhan, J., & Fang, X. (2011). Social Computing: The State of The Art. International Journal Social Computing and Cyber-Physical Systems, 1-12.

[9] Swamy, M. N., & Hanumanthappa, M. (2012). Predicting Academic Success from Student Enrolment Data Using Decision Tree Technique. International Journal of Applied Information Systems (IJAIS), 2.

[10] U.-U. R. I. N. 10 T. 1998 TENTANG and P. A. U.-U. N. 7 T. 1992 T. PERBANKAN, "Undang-Undang Republik Indonesia No. 10 Tahun 1998 Tentang Perbankan," *Lembaran Negara Republik Indonesia*, p. 182, 1998, [Online]. Available: http://www.bphn.go.id/data/documents/98uu010.pdf.

[11] Guo, G., Zhu, F., Chen, E., Liu, Q., & Wu, L. (2016). From Footprint to Evidence: An Exploratory Study of Mining Social Data for Credit Scoring. Research Collection School of Information Systems Singapore Management University, 1-40

[12] A. Lailiyah, "Urgensi Analisa 5C Pada Pemberian Kredit Perbankan Untuk Meminimalisir Resiko," *Yuridika*, vol. 29, no. 2, pp. 217–232, 2014, doi: 10.20473/ydk.v29i2.368

[13] FICO® Score, "What is a FICO® Score?," *FICO Scores by Percent of Scorable Population*, pp. 1–7, 2021, [Online]. Available: https://www.cadbury.co.uk/products

[14] K. Farisi and M. Saifi, "DALAM RANGKA MENDUKUNG PENGENDALIAN KREDIT BANK Pada PT . KBPR Pancadana Batu," *Jurnal Administrasi Bisnis*, vol. 30, no. 1, pp. 117–124, 2017.

[15] A. Puspitaningtyas, "Analisis Prinsip 5C Dan 7P Pada Penyaluran Kredit Di Pt. Bpr Antar Rumeksa Arta Karanganyar," *Fakultas Ekonomi Universitas Sebelas Maret Surakarta*, pp. 1–73, 2012.

[16] S. Suyitno and budi agus riswandi, "Peranan Klausul Standar Baku dalam Perjanjian Kredit Bank," *Jurnal Hukum IUS QUIA IUSTUM*, vol. 7, no. 15, pp. 175–182, 2000, doi: 10.20885/iustum.vol7.iss15.art13.

[17] S. Suyitno and budi agus riswandi, "Peranan Klausul Standar Baku dalam Perjanjian Kredit Bank," *Jurnal Hukum IUS QUIA IUSTUM*, vol. 7, no. 15, pp. 175–182, 2000, doi: 10.20885/iustum.vol7.iss15.art13.

[18] I. Universal and P. Release, "Analisis Kredit , Pintu Masuk Pencairan Pinjaman Analisis Kredit : Cermat dan Teliti," pp. 1–10, 2021.

AdaBoost integration with Genetic Algorithm for Psychological Aptitude Result Interpretation Model

1st Bayu Hanif Pratama
Department of Information Technology
Politeknik Caltex Riau
Pekanbaru, Indonesia
bayu20s2tk@mahasiswa.pcr.ac.id

2nd Dadang Syarif Sihabudin Sahid
Department of Information Technology
Politeknik Caltex Riau
Pekanbaru, Indonesia
dadang@pcr.ac.id

***Abstract*—SMA Darma Yudha Pekanbaru has a special program to facilitate students in knowing their interests and talents. The process is carried out by conducting special tests that can only be carried out by a certified psychologist. The psychologist is an educational worker under the division of guidance and counseling. The results of the test will be interpreted by psychologists with an output in the form of interests that can be used to determine the selection of majors in further study in college. Moreover, the results of the interpretation also contain recommendations of interest and talent for a career after the learner graduated from college. However, the recommendations of interest and talent are the result of unilateral analysis by psychologists without a supporting device to confirm the truth and accuracy. In this case the author will conduct research in the form of analysis of the results of interpretation issued by the psychologist of course with a variety of assessment instruments and not only that, to further ensure the results of interpretation, the author conducts analysis using 2 (two) machine learning methods and will then be done in order to get the best results. In this study, the authors used machine learning by comparing the results of analysis from 2 (two) methods namely Naïve Bayes and Decision Trees Classifier which then the classification results will be improved with AdaBoost.**

Keywords—Machine Learning, Supervised Learning, Classification, AdaBoost, Genetic Algorithm.

I. INTRODUCTION

Aptitude assessment and intervention plays sustainable and distinguished roles for every type and phase of evaluation for kids and teenagers concerning to their learning and behaviour issues. Aptitude, intelligence, and achievement as psychological constructs or types of tests are not easily distinguished It is simply distinguished that an achievement test describes people's present status, and aptitude test predict their future behaviour and ability tests asses their innate potential. Both aptitude and intelligence are enduring traits of individual, not easily modified by experiences and special training. In some cases both aptitude and intelligence-tests results were regarded as indications of innate capacity [1].

A psychological testing is a series of for an individual's different abilities, such as their aptitude in a particular field, cognitive functions like memory and spatial recognition, or even traits. These tests are based on scientifically tested psychological theories. The test is administered by the school and the result is interpreted by a licensed and nationally certified school psychologist, with a specialization in multicultural school psychology.

This research aims to develop a model that has been built on previous research namely Ada Boost – Genetic Algorithm [2]. Development focuses on optimizing the process of split datasets into data trains and test data. Optimization uses two algorithms that have the best performance in the classification, namely naïve bayes and decision trees. The accuracy of the predictions of the two algorithms will be compared to determine the quality of the model being worked on.

II. RELATED WORK

Ahmed Sharaf Elden et al [2], implement and test The Ada/ GA algorithm on ASSISTments dataset. The results showed that this algorithm has improved the detection accuracy as well as it reduces the complexity of computation. The accuracy value of the proposed algorithm prediction is 82.07% which is not much different from the application of a single AdaBoost algorithm, it is 81.85%.

Achmad bisri dan Rinna Rachmatika [3], use a sampling technique, SMOTE (synthetic Minority Over-Sampling Technique) and bagging technique as an ensemble in the Gradient Boosted Trees (GBT) classification method for handling the class imbalance problem. The proposed method is able to provide significant results with an accuracy of 80.57% and an AUC of 0.858, in the category of good classification.

Saifudin [4], A research on the selection of prospective new students by proposing several data mining methods and the one with the best value is the Support Vector Machine (SVM) method with an accuracy of 65%. Model testing with a 10-fold cross validation technique is implemented in this work. With this validation technique, the split process between training data and testing data should be a much better way. However, the accuracy of the predictions of each algorithm is not more than 65%.

Al-Radaideh et al [5], proposed a decision tree model with three different classification methods (ID3, C4.5, and nave Bayes) which allows students to predict the final grade in a course under study. From the test results they found that the decision tree model is the best of several models ever.

Dekker et al [6], presented a case study to predict students' drop out after the first semester of their studies or even before they enter the study program as well as identifying success-factors specific to the EE program by demonstrating the effectiveness of several classification techniques and cost-sensitive datasets. The experimental results found that using simple classification (J48, CART) gave satisfactory results compared to other algorithms such as Bayes Net or JRip.

Kalles and Perrakeas [7], studied the performance of different machine learning techniques (decision tree, neural network, nave Bayes, instance-based learning, logistic regression and support vector machine). By comparing those techniques with genetic algorithm based on decision tree induction, they can analyze the students' academic performance with students' homework as the paramete. They also got short rules to explain and predict success/failure on students' exams.

Based on the literacy review conducted, we found 2 (two) algorithms that will be developed with previous research, namely ada-ga model by focusing on optimizing the training process and testing datasets. These algorithms are naïve bayes and decision trees classifiers.

III. RESEARCH METHODOLOGY

This research methodology systematically defined the experiment models with research stages from data collection method, data processing, proposed models, model experiment, and evaluation and result validation.

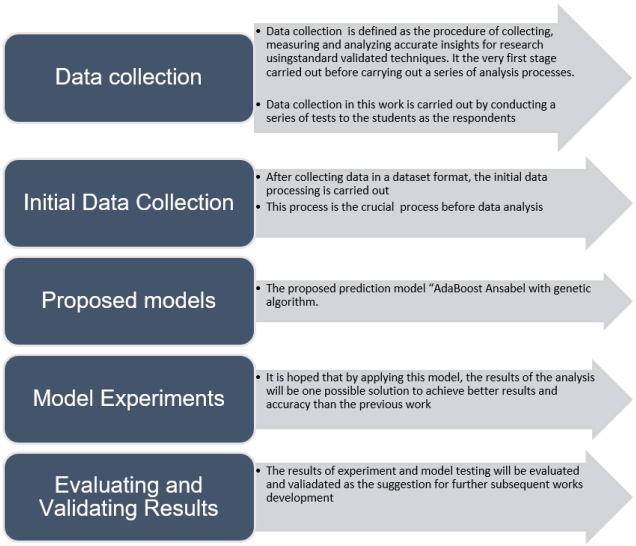


Fig. 1. Flowchart of Methodology Research

Figure 1 describes the stages of work. Starting from the stage of data collection, initial data processing, applying the proposed model, testing and finally is the process of evaluation and validation.

IV. THE PROPOSED MODEL

The objectives of this paper are framed to analyse and predict of the psychological test results by constructing a predictive model using AdaBoost algorithm and genetic algorithm, then validating developed model with original datasets and reliable sources. The next subsections will describe the AdaBoost algorithm, the genetic algorithm and a brief description of the proposed algorithm.

The researcher in this field decided to use the classification techniques, decision tree (DT) dan naïve bayes (NB) due to some advantages they may have over traditional statistical models. Mostly, DT has advantages over traditional statistics on two issues: Primarily, they can handle a large number of predictor variables, far more than the traditional statistics. Moreover, the DM techniques are non-parametric and capture nonlinear relationships and complex interactions between predictors and dependent variable [8].

NB is a classification method with a simple probabilistic-based prediction technique which refers to Bayes' theory by using strong assumptions (naives) and based on probability functions for each instance in mapping attribute classification on a stable efficiency and low complexity system [9][10]

A. *Boosting and AdaBoost*

Boosting is a common machine learning algorithm that increases accuracy of Learning algorithm. It is a widely used and powerful prediction technique due to sequentially builds an ensemble of weak classifiers. In boosting, a weak classifier is a model for binary classification that performs slightly better than random guessing. Formally, a weak classifier achieves slightly better than 50 percent accuracy on the training data. Weak classifier sets are built repeatedly from training data more than thousands of iterations. At each iteration, the training data are re-weighed on how good they are classified (greater weight is given to the classification error sample). Weights are calculated for weak classifiers based on their classification accuracy. The weighted predictions of the weak classifiers are combined by voting to calculate the final outcome prediction [11].

AdaBoost is the most common optimization algorithm for binary classification proposed by Freund and Schapire [12]. It takes as input a training set "S" of "m" sample ($S = \{(x_1,y_1),\ldots,(x_m,y_m)\}$), where each instance of x_i is a vector of attribute values that belongs to a domain or instance space X, and each label y_i is the class label associated with xi that belongs to a finite label space $Y=\{-1, +1\}$for binary classification problems.

The following figure is a general illustration of the AdaBoost algorithm for binary classification problem.

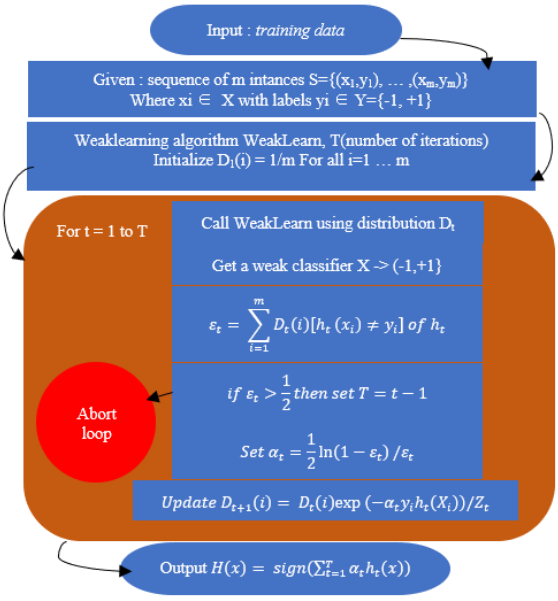


Fig. 2. A generalized version of the AdaBoost Algorithm

Ada Boost weights the training sample with a probability distribution of $D_t(x)$ in each iteration. WeakLearn algorithm is applied to generate ht classifier with error rate ϵt on training sample. The effect of changing the weights is placing more error classifiers in the final stage. This process continues during the T round until the of the final classifier H, is constructed by weighting the weak classifier $h_1, h_2, \ldots, h_T$. Each classifier is weighted according to its accuracy of the distribution Dt that it was trained on [12]. The weak classifier is the core of an AdaBoost algorithm. In this work, classification and regression tree (CART) algorithm, proposed by Breiman et al. [13], was used as WeakLearn to AdaBoost algorithm.

B. *Genetic Algorithm*

Genetic algorithm (GA)is an evolutionary based stochastic optimization algorithm with a global search potential proposed by Holland (1973) [14]. GA is among the most successful class of algorithms under EAs(Evolutionary Algorithms) which are inspired by the evolutionary ideas of natural selection. Because of its outstanding performance with optimization, GA has been regarded as a function optimizer.

The algorithm starts by initializing the solution population (chromosomes) and comprises representation of the problem usually in the form of a bit vector. Chromosomes evolve through successive iterations called generations. During each generation, the chromosomes are evaluated, using some measures of fitness (using an appropriate fitness function suitable for the problem). To create the next generation, new chromosomes, called offspring, are formed by combining two chromosomes from the current generation using the crossover operator or modifying the chromosomes using the mutation operator. A new generation is formed by voting; fitter chromosomes have a higher probability of being selected. After several generations, the algorithm encounters the best chromosome, which hopefully represents an optimal or suboptimal solution to the problem. The three main genetic operators in GA involve selection, crossover, and mutation[15].

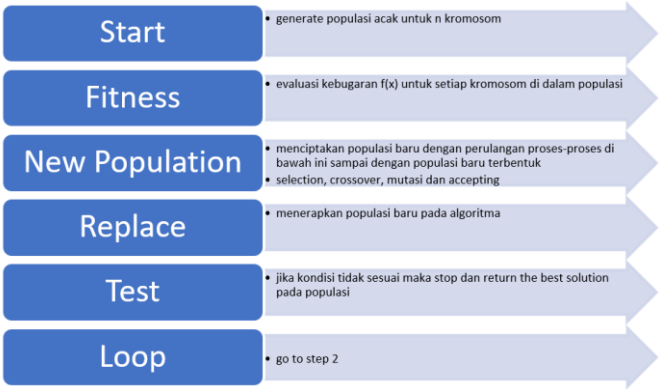


Fig. 3. The outline of the GA algorithm

C. *Overview of the Proposed Model: AdaBoost-GA*

Freund dan Schapire [12] concluded that the AdaBoost algorithm is less prone to overfitting problems compared to most learning algorithms, because it increases sensitivity to data and noise outliers. Thus, mislabelled cases or outliers may cause the overfitting problems, for the new classifier to focus more on those observations that have been misclassified, resulting in a large number of weak classifiers to achieve better performance [11].

In this study, a new boosting algorithm called "Ada-GA" which strengthen the classification of the decision tree and nave Bayes algorithms. In the previous work, applied another classification algorithm after the output was corroborated by AdaBoost – GA [2].

Input:a set S of m instances: $S=\{(x_1,y_1), \ldots, (x_m,y_m)\}$ where $x_i \in X$ with labels $y_i \in Y=\{-1, +1\}$, P (population size), G (maximum number of generations), T(initial number of weak learners).

Initialize: a randomly generated population of m solutions (consists of a set of Tweak learners withtheir weightsproduced by AdaBoost)

Evolve: fork = 1, 2, …, G.
1-Generate a population of bit strings b
2- Evaluate the fitness of the solutions:$f(b)=w_1 * (1-L/T)+w_2 * (1/E_b^s)$
where b isthe evolved best individual, w_1, w_2 are fitness weights ,$E_b^s$is the validation error and $L=\sum_{i=1}^{T} b_i$
3-Use vector b to update the weak classifiers (T)and their weights.
4-Produce new generation of solutions using the genetic operations (selection, mutation and crossover).
5- If end condition, stop, and return the best solution; else loop

Output:final hypothesis (with optimized classifiers and their weights).

Fig. 4. The Proposed Procedure of Ada-GA

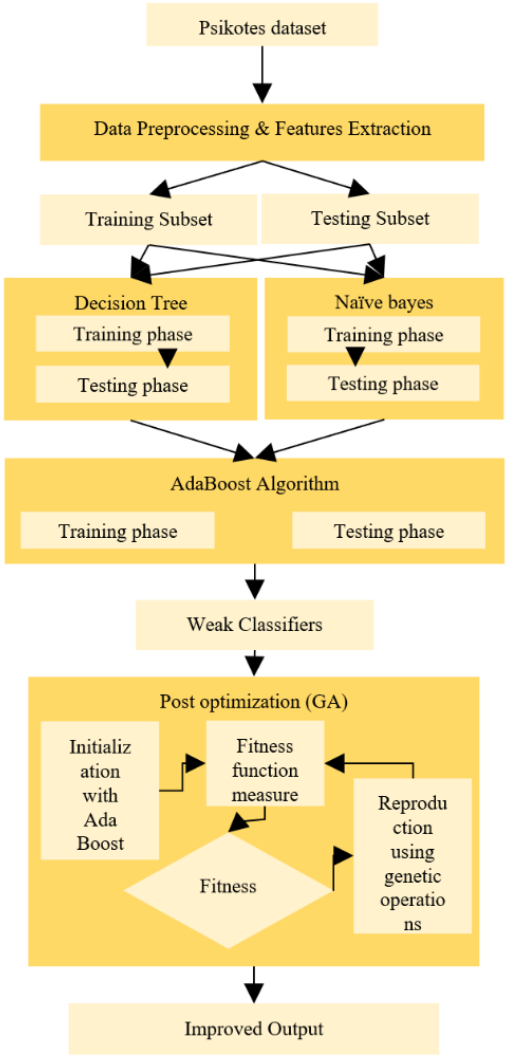


Fig. 5. The Structure of the Proposed Model Ada-GA

The structure of "Ada-GA" is detailed in Figure 4 which consists of three following phases:

1. Pre-processing and feature extraction phase: This phase is specifically to split the dataset into training and testing randomly.
2. Pre-processing and feature extraction phase: This phase is specifically to split the dataset into training and testing randomly.
3. The difference between this work and the previous work is that there is a combination of higher values two algorithms, Decision Tree and Naïve Bayes, and a comparative analysis of both models.
4. Post optimization procedure phase: this phase is composed of three parts:
   a. Initialization with AdaBoost
   b. Fitness function, and
   c. GA

V. CONCLUSIONS

Modeling for optimization in the training and testing process carried out in this study can produce a higher score than in the previous work of 82.07%. We are very optimistic because based on the reference classification algorithm that has always been at the top in terms of performance are decision trees and naïve bayes.

REFERENCES

[1] G. Goldstein and M. Hersen, *Handbook of Psychological Assessment 3rd Edition*, 3rd ed. Pergamon, 2000.

[2] A. ElDen, M. A. Moustafa, H. M. Harb, and A. H.Emara, "Adaboost Ensemble with Simple Genetic Algorithm for Student Prediction Model," *Int. J. Comput. Sci. Inf. Technol.*, vol. 5, no. 2, pp. 73–85, 2013, doi: 10.5121/ijcsit.2013.5207.

[3] A. Bisri and R. Rachmatika, "PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN TEKNIK MACHINE LEARNING PADA LEVEL DATA UNTUK MENANGANI KETIDAKSEIMBANGAN KELAS," *Direktorat Ris. dan Pengabdi. Masy. Direktorat Jenderal Penguatan Ris. dan Pengemb. Kementrian Riset, Teknol. dan Pendidik. Tinggi*, 2019.

[4] A. Saifudin, "Metode Data Mining Untuk Seleksi Calon Mahasiswa," vol. 10, no. 1, pp. 25–36, 2018.

[5] Q. A. Al-Radaideh and E. Al-Shawakfa, "Mining Student Data Using Decision Trees," *Expert Syst. Appl.*, vol. 40, no. 2, pp. 1–18, 2015, [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2008.02.021%0Ahttp://mechanicaldesign.asmedigitalcollection.asme.org/article.aspx?doi=10.1115/1.4026094%0Awww.wiley.com%0Ahttp://www.sciencedirect.com/science/article/pii/S22125671150.

[6] G. W. Dekker, M. Pechenizkiy, and J. M. Vleeshouwers, "Predicting students drop out: A case study," *EDM'09 - Educ. Data Min. 2009 2nd Int. Conf. Educ. Data Min.*, no. January, pp. 41–50, 2009.

[7] D. Kalles and C. Pierrakeas, *Analyzing student performance in distance learning with genetic algorithms and decision trees*, vol. 20, no. 8. 2006.

[8] Z. J. Kovacic, "Early Prediction of Student Success: Mining Students Enrolment Data," *Proc. 2010 InSITE Conf.*, pp. 647–665, 2010, doi: 10.28945/1281.

[9] T. D. Salma and Y. S. Nugroho, "Sistem Rekomendasi Pemilihan Sekolah Menengah Tingkat Atas Menggunakan Metode Naive Bayes," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 2, no. 2, p. 85, 2016, doi: 10.23917/khif.v2i2.2306.

[10] N. Ailmi, Z. Saharuna, E. Tungadi, and A. M. Learning, "Metode Klasifikasi Pada Aplikasi Pendukung Keputusan Untuk Pemilihan Unit Kegiatan Mahasiswa," *Pros. Semin. Nas. Tek. Elektro dan Inform. 2020*, pp. 142–147, 2020.

[11] DONG-YOP OH, "GA-BOOST: A GENETIC ALGORITHM FOR ROBUST BOOSTING," 2012.

[12] Y. Freund, R. E. Schapire, and M. Hill, "Experiments with a New Boosting Algorithm Rooms f 2B-428 , 2A-424 g," 1996.

[13] L. Breiman, J. H. Friedman, R. A. Olshen, Charles, and J. Stone., *Classification and Regression Trees*, vol. 10, no. 4. Brooks/Cole Publishing, 1984.

[14] J. H. Holland, "The Optimal Allocation of Trials," *Adapt. Nat. Artif. Syst.*, vol. 2, no. 2, pp. 88–105, 2019, doi: 10.7551/mitpress/1090.003.0008.

[15] B. S and S. S. Sathy, "A Survey of Bio inspired Optimization Algorithms," *Int. J. Soft Comput. Eng.*, vol. 2, no. 7, pp. 2323–2339, 2012, doi: 10.1007/s11269-015-0943-9.

Rice Quality Detection Based on Digital Image Using Classification Method

1st Sellya Meizenty
Master of Applied Computer Engineering
Politeknik Caltex Riau
Pekanbaru, Riau, Indonesia
Sellya20s2tk@mahasiswa.pcr.ac.id

2nd Dadang Syarif Sihabudin Sahid
Master of Applied Computer Engineering
Politeknik Caltex Riau
Pekanbaru, Riau, Indonesia
dadang@pcr.ac.id

3rd Juni Nurma Sari
Master of Applied Computer Engineering
Politeknik Caltex Riau
Pekanbaru, Riau, Indonesia
juni@pcr.ac.id

***Abstract*—Rice is one of the staples that is included in the consistent list of staple food commodities (Bapok), currently some irresponsible people make the rice more durable, fragrant and whiter. Many assume that the rice is clean, odorless, and has a high price is rice with good quality and vice versa. From the existing problems the author wants to help the community to better determine good quality rice and good for consumption.**

This research will create a system that can recognize the type of rice based on the image of the rice. Rice data that has been collected will be sampled and trained using the K-Nearest Neighbors (k-NN) method where this method is used for the classification of the shortest distance calculation which will produce a class in the form of rice data classes, while to obtain parameter values from the rice image using the extraction feature. RGB color average (Red, Green, and Blue) and to get results with a good level of accuracy will use K-Fold Validation.

Keywords—Rice, RGB, K-Nearest Neighbors (k-NN), K-Fold Validation.

I. INTRODUCTION

Rice is one of the staple ingredients included in the consistent list of staple food commodities (Bapok). Staples play an important role in economic, social, and even political aspects. [15] Rice contains a lot of carbohydrates in the human body. Rice is a food that comes from the rice plant that has been separated from the skin. According to [11] 75% of the daily caloric intake of people in Asian countries comes from rice. More than 59% of the world's population depends on rice as the main source of calories

The current problem is that some irresponsible people make the rice more durable, fragrant and also whiter. So the researchers want to help the community to be more proactive about the problems that occur, namely finding the fact that a number of traders and rice entrepreneurs have found the fact that there is a bleaching agent attached to rice. so that it is known that there are fraudulent practices of a number of traders in an effort to increase prices. [16]

Many assume that rice that is clean, odorless, and has a high price is rice of good quality and vice versa, in previous studies it was assumed that rice that looks very clean, expensive and does not smell is rice of good quality.[9]

From the existing problems, the author wants to help the community to better determine good quality and good rice for consumption. This study will create a system that can recognize the type of rice based on the image of the rice, then the data that has been collected will be sampled and trained using the K-Nearest Neighbors (k-NN) method and to get the results a good level of accuracy will use k-fold validation. ice is one of the staple ingredients included in the consistent list of staple food commodities (Bapok). Staples play an important role in economic, social, and even political aspects. [15]

Rice contains a lot of carbohydrates in the human body. Rice is a food that comes from the rice plant that has been separated from the skin. According to [11] 75% of the daily caloric intake of people in Asian countries comes from rice. More than 59% of the world's population depends on rice as the main source of calories

The current problem is that some irresponsible people make the rice more durable, fragrant and also whiter. So the researchers want to help the community to be more proactive about the problems that occur, namely finding the fact that a number of traders and rice entrepreneurs have found the fact that there is a bleaching agent attached to rice. so that it is known that there are fraudulent practices of a number of traders in an effort to increase prices. [16] Many assume that rice that is clean, odorless, and has a high price is rice of good quality and vice versa, in previous studies it was assumed that rice that looks very clean, expensive and does not smell is rice of good quality.[9]

From the existing problems, the author wants to help the community to better determine good quality and good rice for consumption. This study will create a system that can recognize the type of rice based on the image of the rice, then the data that has been collected will be sampled and trained using the K-Nearest Neighbors (k-NN) method and to get the results a good level of accuracy will use k-fold validation.

II. LITERATURE REVIEW

Rice contains a lot of carbohydrates in the human body. Rice is a food that comes from the rice plant that has been separated from the skin. According to Marjuki, 75% of the daily caloric intake of people in these Asian countries comes from rice. More than 59% of the world's population depends on rice as the main source of calories [11].

The current problem is that some irresponsible people make the rice more durable, fragrant and also whiter. So the researchers want to help the community to be more proactive about the problems that occur, namely finding the fact that a number of traders and rice entrepreneurs have found the fact that there is a bleaching agent attached to rice. so that it is known that there are fraudulent practices of a number of traders in an effort to increase prices. [16].

Some researchers want to test it visually to find out the quality of rice is good and with good quality standards as well. In another study explained that currently testing is done manually so it is prone to errors. There is research on identifying rice quality with digital images [4]

This testing process is seen from the white, clean and intact value of a rice, by analyzing the value of HSV (Hue, Saturation, Value) to analyze the value of white and clean, while to analyze the value of whole rice, the object area is used. Then they took a sample of 30 data to form a decision tree classification method with the ID3 (Interative Dichotomiser Tree) model, there are 3 classification classes, namely good, poor, bad. The results of the test using the k-fold cross validation method with k = 5 obtained an accuracy of 96.67.

Other researchers studied the characteristics of rice based on analysis of image processing and artificial neural networks [1], the quality process will be seen from head rice, broken rice, groat rice and unhusked rice. researchers use digital image processing to minimize the use of inputs and parameters using index B, roundness, area, length and saturation,

To estimate brown rice, yellow rice, green rice and other foreign rice, the index parameters R, G, B, Roudness and area can be used. These values will be trained and grouped based on the type of each rice so that researchers find differences in shape, size and color so as to make different accuracy values. The results of the training between head rice, broken rice, groat rice and unhulled rice were tested with 5 input parameters the results showed a good value of 97.14% and the validation was 96.74% then for the training results from brown rice, yellow rice and green rice were 98.55 % and the validation is 90.48%.

Subsequent research discusses the same thing, namely detecting rice quality by using image segmentation based on rice grain fractions and color distribution, according to [9]. His research has factors to determine the quality of rice such as grain, non-uniform color, odor and others. This study uses the outsu feature to determine the number of broken rice grains and color distribution to determine color uniformity, this researcher uses the k-nearst neighbor classification method which has a value of 99.87 and the test results will use k-fold validation with k = 10.

The next research is on digital image processing using the SUSAN detection and Neurofuzzy methods. According to [2] the quality component of rice in his research uses a digital image processing recognition approach where the results will be able to recognize the quality components of rice which include broken grains, groats, sosoh degrees, and yellow grains. and get a score for success. The test performs blob detection on rice images with a distance of 18 cm from the camera. The sample will be tested using two methods, namely grain length using the SUSAN method, while for texture using the Neurofuzzy method.

Susan's method will include grayscale, limiting thresholding, while the Neurofuzzy method includes image extraction with GLCM to divide sata samples, training data and testing data, the resulting outputs are (good premium, bad premium, good medium, bad medium, and economical). The results obtained from this study are that the training data got good premium results, which resulted in sensitivity 51.962 %, specificity 40.151%, and accuracy 45.345%, while for testing the results were sensitivity 48.387%, specificity 38.376%, and accuracy 42.640%.

III. RESEARCH METHODS

A. *Data Collection (Data Collection)*

The data used is from Bulog in Pekanbaru, Riau. Several samples of all types of rice from good quality to poor quality will be taken.

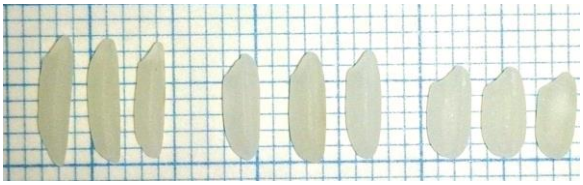


Fig. 1. Several Sample of Rice

Rice as the main food in the form of whole grains, so that shape and appearance are the first characteristics that consumers choose to choose and buy rice. The shape of the rice is a character caused by the operation of the milling process which is a combination of the type and capability of the machine, the operator's competence and the quality of the milled grain. Aspects of the shape and appearance of rice consist of length (size) and grain shape. In general, the length measurement consists of the complete length and shape based on the standard rice plant breeding process in general [7].

B. *Training Data*

K-Nearest Neighbor (KNN) belongs to the instance-based learning group. This algorithm is also a lazy learning technique. KNN is done by looking for groups of k objects in the training data that are closest (similar) to the objects in the new data or testing data.

There are many ways to measure the proximity between new data and old data (training data), including the Euclidean distance and the Manhattan distance (city block distance), the most commonly used is the Euclidean distance.

$$D(A,B)\sqrt{\sum_{k=1}^{d}(Ak-Bk)^2} \quad (1)$$

Information :

- D : proximity
- A: training data
- A: data testing
- d : the number of individual attributes between 1 to d
- f : attribute similitary function k between cases
- A and case B
- k = individual attributes between 1 to d

In finding the original nearest neighbor of the query point, K-Nearest Neighbors (k-NN) designs the nearest neighbor with the sum of the weights in each class to the query point. The winning class is determined from the smallest number of weights. The steps to calculate the K Nearest Neighbor method include:

1. Specifies the parameter (number of closest neighbors).
2. Calculate the square of the Euclidean distance (query instance) of each object to the given sample data using equation 1.
3. Then sort the objects into groups that have the smallest Euclidean distance.
4. Collect category (Nearest Neighbor Classification)
5. By using the category of Nearest Neighbor which is the majority, it can predict the value of the calculated query instance.

The image is a continuous function of the light intensity in the two-dimensional plane. The light source illuminates the object and reflects back part of the light beam [13] The image as the output of a data recording system can be optical, analog, or digital. Images that can be directly stored on a storage medium are called digital images. Digital images can be processed by a computer [10]

A digital image forms a two-dimensional field for image processing to be carried out. coordinates (x,y) represent the position of the coordinates in the Cartesian system where the horizontal axis is expressed as the x-axis, and the vertical axis is expressed as the y-axis. The light intensity function at the coordinates $(x,)$ is symbolized by $f(x,y)$ [7] Because light is a form of energy, the light intensity is between 0 and infinity, as shown in equation 4.

$$0 < f(x,y) < \infty \quad (2)$$

The intensity function (x,y) is the product of the amount of light (x,y) or illumination from the light source that illuminates the object or illumination, with the degree of the object's ability (x,y) to reflect light by the object or reflection, according to equation 5. The range of values $(x,)$ is [0,], while the range of values (x,y) is [0,1].

$$f(x,y) = i(x,y)\ .\ r(x,y) \quad (3)$$

C. Digital Image Data Acquisition

The digital image data of rice that has been collected for the testing phase is processed into textual data. The data acquired are the amount of rice, white value, net value, and rice delivery value.

1) Binary Image Segmentation

The digital image data of rice that has been collected for the testing phase is processed into textual data. The data acquired were the amount of rice, white value, net value, and whole rice value. [4].

2) Acquisition of Rice Grains

First, binary image segmentation is performed. The purpose of segmentation is to get a simple representation of an image so that it is easier to process. Segmentation is done by converting the RGB (Red, Green, Blue) rice image into a grayscale image first. Converting an RGB image to grayscale is done with the following equation: [4]

$$Grayscale = \frac{R+G+B}{3} \quad (4)$$

After being converted into grayscale image data, then it is converted to a binary image with the middle value threshold of the gray value in the image. After obtaining the binary image, segmentation is carried out by separating black pixels as background and white as objects. To facilitate object analysis for the next stage, data is collected on the location of the coordinates of each segmented object region. [4]

3) White Rice Value Acquisition

In the process of acquiring the value of white rice, analysis of the value of Hue, Saturation, and Value is carried out. The initial RGB image is converted into RGB form for the analysis process. Each value of Hue, Saturation, and Value is taken and analyzed according to the specified threshold according to the standard from the rice warehouse. The analysis process is carried out on each grain of rice according to the coordinates of the segmentation object. After all objects have been analyzed, the value is labeled as white or not white. Then calculated the percentage of objects that are white. If the number of white rice grains is less than 75% of the entire rice picture, then the rice is categorized as not white [4].

4) Acquisition of Rice Net Value

In the process of acquiring the net value of rice, the process carried out to obtain the net value is to analyze the value of Hue with a predetermined threshold according to the standards of the rice warehouse. The initial RGB image is converted into HSV format for the analysis process. The analysis process is carried out on each grain of rice according to the coordinates of the segmentation object. After all the objects have been analyzed, they are labeled as clean or unclean. Then the percentage of objects that are net worth is calculated. If the number of clean rice grains is less than 75% of the entire rice picture, then the rice is categorized as unclean. [4]

5) Rice Whole Value Acquisition

In the process of acquiring the value of the integrity of rice, it is necessary to check the area of each object first. The area calculation is done by counting the number of Mpixel in each object region. Then each object is labeled intact or incomplete value. Determination of the whole value of rice is done by comparing the object area with the standard rice area determined according to the standard from the rice warehouse. If the number of whole grains of rice is less than 75% of the whole picture of rice, then the rice is categorized as incomplete. [4]

IV. The Proposed Model

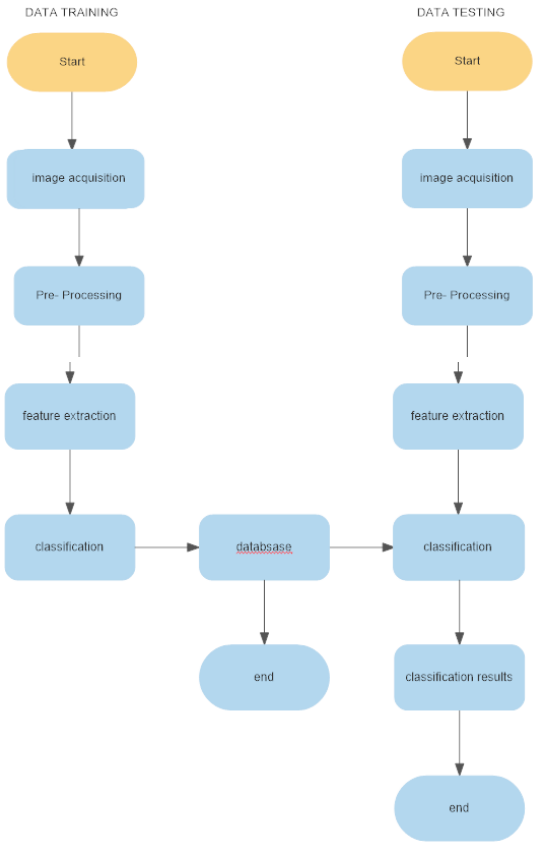


Fig. 2. Proposed Model

A. Starting from image acquisition

Then preparing the data to be tested which will then be processed. Then the parameters are taken to distinguish between other objects. Then it is entered into the system according to the existing group or characteristics. At this stage the database should be classified again and at the final stage the results of the classification will be tested with other classifications so as to get the result which is the best method..

Acknowledgment

Finally, I would like to thank all those who played a role in supporting the success of this research. This paper is far from perfect, but it is hoped that it will be useful for readers, for that the author needs criticism and suggestions from readers so that this paper is better in the future.

References

[1] Agus Supriatna Soemantri, ED (2017). Identification of Physical Quality of Rice Using Image Processing Technology and Artificial Neural Networks. Journal of Agricultural Postharvest Research (Agricultural Research and Development Agency).

[2] Amal, I. (2017). Digital Image Processing Using Susan Method. Journal of Mathematics, No.6 Vol 4.

[3] Anggraini, AP (2020, 05 18). Get to know the various types of rice and their nutritional content. Retrieved From Kompas.Com: Https://Health.Kompas.Com/Read/2020/05/18/030000368/Mengenal-Berbagai-Jenis-Beras-Dan-Kandungan-Gizinya?Page=All

[4] Arissa Aprilia Nurcahyani, RS (2015). Identification of Rice Quality With Digital Image. Scientific Journal of Informatics.

[5] Aruan, RW (2020). Comparison of the Pseudo K-Nearest Neighbor Method and the K-Nearest Neighbor Method on the Classification of Beef and Pork. Uty Journal.

[6] BV, VD (2017, 06 28). Rice. Retrieved From Indonesia Investments:Https://Www.IndonesiaInvestments.Com/Id/Bisnis/Komoditas/Beras/Item183

[7] Bbpadi, BB (2016, 01 11). Rice Quality Aspects. Retrieved From Rice Plant Science and Technology Source: Https://Bbpadi.Litbang.Pertanian.Go.Id/Index.Php/Info-Berita/Info-Teknologi/Aspek-Mutu-Beras

[8] Blp, BL (2013, 07 17-23). Harvest And Post-Harvest Rice. Agroinnovation (Sinar Tani), Pp. 6-16.

[9] Eko Supriyadi, AB (2020). Rice Quality Detection Using Image Segmentation Based On Grain Fraction And Color Distribution. Jire (Journal of Informatics & Electronic Engineering), 20-29.

[10] Coverage6. (2007, 128). Retrieved From Liputan 6: Https://Www.Liputan6.Com/News/Read/136577/Beras-Kita-Berchlorin

[11] Marjuki, FA (2008). Analysis of Factors Affecting Rice Prices in Indonesia 1981-2006.

[12] Modern, KT (2019, 11 4). Harvest and Post-Harvest Handling of Rice. Retrieved From Http://Cybex.Pertanian.Go.Id/:Http://Cybex.Pertanian.Go.Id/Mobile/Artikel/79243/Penanganan-Panen-Dan-Pasca-Panen-Padi/

[13] Nurcahyani, AA (2015). Identification of Rice Quality With Digital Image. Scientific Journal of Informatics.

[14] Pamungkas, A. (2017, August 26). Matlab Programming. Retrieved From Pemrogramanmatlab.Com/2017/07/26/Pengolahan-Citra-Digital/

[15] Prabowo, DW (2014). Clustering of Staple Food Commodities Using Analytical Hierarchy Process Method. Jurnal.Kemendag.Go.Id/.

[16] Sinabang, M. (2018, May). Analysis of Chlorine (Cl) Content in Rice and Knowledge Level of Sellers at Afternoon Market, Padang Bulan Village, Medan City in 2018.RetrievedFromHttp://Repositori.Usu.Ac.Id/Bitstream/Handle/123456789/11446/141000423.Pdf?Sequence =1&Isallowed=Y.

[17] Sofia Saida, D. (2019). Rice Quality Identification Using K-Nearest Neighbor And Support Vector Machine Method. Journal of Telecommunications, Electronics, Computing and Control.

[18] Ugm, T. (2017, 08 11). Paca Harvest Techniques. Retrieved From Post-Harvest Engineering Agricultural and Biosystem EngineeringUniversitasGadjahMada:Https://Pascapanen.Tp.Ugm.Ac.Id/2017/08/11/Teknik-Pascapanen-Padi./

The FMEA Traditional Modifications (FMEA Improvement) in IT Risk Assessment

1st Nina Fadilah Najwa
Information System Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
nina@pcr.ac.id

2nd Apol Pribadi Subriadi
Information System Department
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
apolpribadi@gmail.com

3rd Okfalisa
Informatics Department
UIN Sultan Syarif Kasim
Pekanbaru, Indonesia
okfalisa@gmail.com

4th Eki Saputra
Information System Department
UIN Sultan Syarif Kasim
Pekanbaru, Indonesia
eki@uin-suska.ac.id

5th Muhammad Ariful Furqon
Informatics Department
Universitas Jember
Jember, Indonesia
ariful.furqon@unej.ac.id

***Abstract*—Failure mode effect and analysis (FMEA) was an essential technique in risk management. Various studies have revealed the shortcomings of FMEA, one of them is about subjectivity and the consistency of the results issues. This research contributed to the FMEA improvement as the improvement Model to minimize traditional FMEA consistency issues. Based on previous research, the present study will discuss in more detail the FMEA improvement method. The research questions formulated for this research are: (1) How was the use of the FMEA improvement model in the case study; and (2) What were the implications of the FMEA improvement model. The FMEA Improvement Model consisted of four main stages, namely: (1) the determination of risk assessment requirements; (2) risk identification; (3) risk analysis; and (4) evaluation and the recommendation control documentation. From the results of risk level categorization, one risk was at the highest level. This risk was classified as severe because it often happens almost every day. Two risks were at the medium to the high level, three risks were at the medium level, 13 risks were at the low to medium level, and 18 risks were low. Therefore, this paper's contribution was to guide researchers or IT experts to apply the FMEA improvement.**

Keywords—FMEA Improvement, IT Risk, Risk Assessment

I. INTRODUCTION

Information technology is part of the information system [1]. In information systems and technology, risk management is carried out for information security. An essential aspect of information security was the aspect of the CIA triad (*confidentiality*, *integrity*, and *availability*) [2]. Failure mode effect and analysis (FMEA) was an essential technique in risk management that is often used in risk assessment [3]. Using FMEA, the failure modes' criticality was assessed to select only the critical system-level failure that might be happened [4]. FMEA provided a standard structure and language used in various organizations and different risks, including IT risk. The FMEA used in the IT field was to determine the information security risks of sensitive data [5]. Besides, FMEA can analyze the causes and effects of longer lead times in software development [6].

Using the FMEA has proven to reduce the cycle of warranty costs and the minimal costs to prevent compared to fixing problems that have already occurred [7]. FMEA was a commonly used method and could be well documented [8]. Previous researches criticized the limitation or weaknesses of FMEA traditional [9]. The FMEA weaknesses occurred mainly when risk assessment. There were the subjectivity and inconsistency element; the RPN's potential value is unsustainable; there are duplicate RPN values; the RPN was not recommended for use in risk assessment [7]. Conventional FMEA also had considerable shortages due to various reasons for which uncertainty and ambiguity are significant [10]. Traditionally, FMEA only considered the impact of a failure of a system. The strategies needed to be examined in defining risks and in calculating risks [11]. The results of risk assessment using FMEA contained consistency and subjectivity issues [3],[12], [13], [14]. Differences in risk ratings could occur in the prevention or focus on handling [7].

The causes of those limitations were identified and analyzed, then provided recommendations solutions. FMEA Improvement was an improved model of traditional FMEA weaknesses that had been defined by [15]. The research examined the consistency results using Traditional FMEA and FMEA Improvement. The study results proved that the results of risk measurement with improvement FMEA were more consistent than FMEA traditional. Consistency testing was carried out on two different teams and the same case study. However, the research did not discuss how to use FMEA Improvement in detail step and its implications. Thus, this paper explained the detailed stages of using FMEA Improvement in a case study.

Based on previous research, the present study discussed in more detail the FMEA Improvement method. The research questions formulated for this research are: (1) How was the use

of the FMEA Improvement model in the case study; and (2) What were the implications of the FMEA Improvement model. Therefore, this research paper's contribution was to guide researchers or IT experts to apply the FMEA improvement based on the previous work [15]. The theoretical contribution did an empirical study of the FMEA Improvement and analyzed the implications of FMEA improvement.

II. THEORETICAL BACKGROUND

A. *Information Technology/Information Security Risk*

The risk was predicted from possible threats from specific potential vulnerabilities and impacts on the organization [16]. IT risks were related to threats and dangers due to intensive IT usage, which might cause undesirable or unexpected damage [17]. The source of threats originated from one or more potential attackers with a set of one or more system conditions that provided attacker motivation [18]. Sources of threats consisted of internal threats and external threats. Internal threats were a potential possibility of attacks from internal organizations. External threats were a likely possibility of attacks from external organizations [19].

The organization must determine the IT security capabilities that IT systems must possess in the face of real-world threats. Most organizations had a strict budget for IT security. Expenditures for IT security must be thoroughly reviewed, such as other management decisions. The risk management methodology was well structured and could help management identify appropriate controls to provide essential security capabilities for the organization's mission [16]. IT risk management was meant to protect IT assets such as data, hardware, software, personal, and facilities from all threats [20]. Risk management could minimize the costs incurred if it turned out that the risk event occurred.

B. *Failure Mode Effect and Analysis (FMEA)*

FMEA as a tool in risk management. FMEA is used to identify the potential failure in the process, product, or service. The most method type in risk management was qualitative and descriptive. FMEA is included in the semi-quantitative method [21]. FMEA implementation used the risk rank technique called *Risk Priority Number* (RPN). RPN resulted from the risk assessment with three parameters (severity, occurrence, and detection) [22]. The RPN assessment was based on experts' experience and cognitive skills at gathering data to make risk assessments [23].

FMEA was popular in the risk management method because it could be used in various types of level management and organizations. Technician and non-technician could use the FMEA easily because the FMEA method used understandable language [5]. The FMEA team analyzed failure modes and identified the severity impacts, levels of occurrence, and detection of potential failures. Based on the most critical and possible failure impacts, rank the of each failure [24]. The FMEA result helped the managers and technicians identify failure modes and identify the causes and mitigation of the risk [25].

FMEA was a bottom-up method that is less structured and requires more expert knowledge[4]. The general methodology of FMEA were [5] (1) Review the process or product, (2) Brainstorm potential failure modes, (3) List potential effects of each failure mode (4) Assign a severity ranking for each effect, (5) Assign an occurrence ranking for each failure mode, (6) Assign a detection ranking for each failure mode and effect, (7) Calculate the risk priority number for each effect, (8) Prioritize the failure modes for action, (9) Take action to eliminate or reduce the high-risk failure modes, (10) Calculate the resulting RPN as the failure modes reduced or eliminated.

On the contrary, FMEA had limitations or weaknesses that previous studies have criticized because of several shortcomings. The RPN score does not consider the relative importance of severity, occurrence, and detected [26],[27],[28]. Furthermore, FMEA had uncertainty in the complex system to model the domain experts' subjective opinion[29]. Conventional FMEA also has considerable shortages due to various reasons which uncertainty and ambiguity are significant [10].

III. METHODOLOGY

The research methodology is divided into four stages. The stages were: (1) the formulation of research questions; (2) research objects; (3) FMEA improvement; (4) and implementation of FMEA improvement in the case, then explanation of the results and discussion. The block diagram in figure 1 shows the methodology.

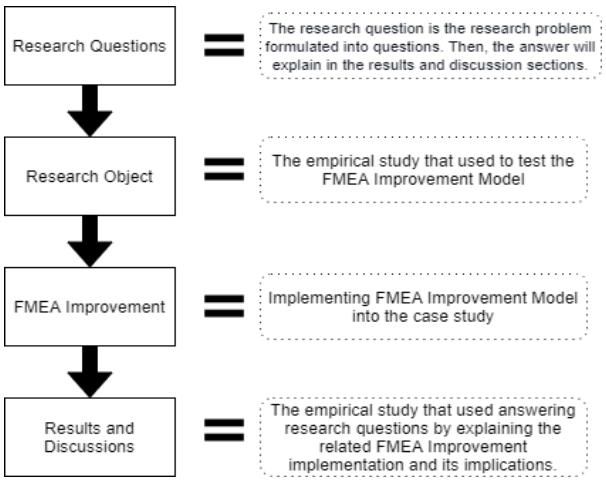


Fig. 1. Methodology diagram

A. *Research Questions*

Based on previous research, this research aims to explain how to implement the FMEA Improvement model and describe the implications of FMEA Improvement. The implication of FMEA Improvement included the explanation of the weaknesses of FMEA traditional. The research questions this research were: (1) How was the use of the FMEA Improvement model in a case study; and (2) What were the implications of the FMEA Improvement model. The research questions were used to describe the procedure to implementations of FMEA Improvement in the study case. Those findings and explain the implication of FMEA

improvement will be discussed further in the discussion section.

B. Research Object

The research object was conducted in the Regional Office of the Ministry of Religious Affairs in Riau province. The empirical study of the FMEA Improvement Model will follow the methodology described in this study.

C. FMEA Modification (FMEA Improvement)

FMEA Improvement Model (Figure 2) was the result of the synthesis of FMEA traditional. The previous study results proved that the results of risk measurement with improvement FMEA were more consistent than FMEA traditional[15]. The consistency of FMEA improvement and FMEA traditional examined by action research with two cycles. The first cycle of action research examined the FMEA traditional consistency using two teams in the same case study. The second action research examined the FMEA Improvement using two teams in the same case study. Based on the gaps analysis, inconsistency will reduce when using FMEA improvement.

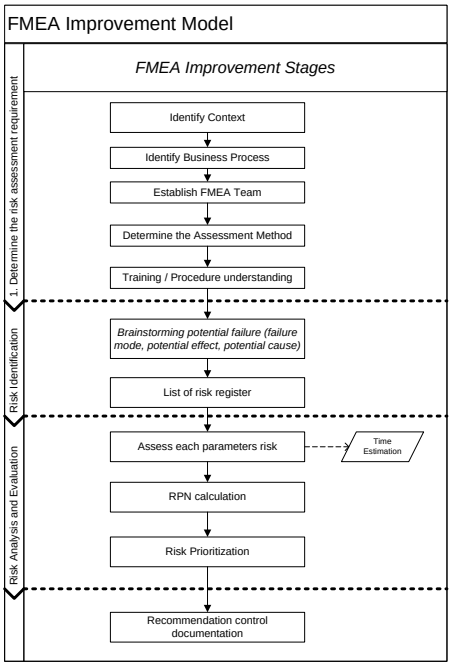


Fig. 2. FMEA Improvement Model

1) Identify context

Context identification includes the declared purpose, scope, and performance indicators. The *critical asset in the organization must be identified and related to the information security aspect. Then, the performance indicator was the organization's commitment to assessing the risk according to the procedure of the FMEA Improvement model.*

2) Identify Business Process

The business process identification helped the team FMEA identify the vulnerability and source of threat in each organization's business process.

3) Establish FMEA Team

The procedure for establishing the FMEA team is based on team size, team member expertise, and team coordinator [15]. In this case study, we created the team size was consist of 3 people. The team member's knowledge was two expert technicians in the FMEA or IT risk assessment and an internal team member who knows the organization well.

4) Determine assessment model

At this stage, using the modified FMEA improvement document [15]. The criteria scale for each parameter (severity and occurrence) was a 1-5 range scale. Also, the severity scale is divided into three impacts/severity (operational, media attention, regulations). The Occurrence criteria scale is the instrument for assessing the frequency of potential failure. The determination of the risk level for the action plan is based on the inherent risk level. Each risk had an RPN score by the multiple of the parameter value (severity x occurrence). Then, risk level criteria consisted of 5 risk level categorizations. The organization's risk level criteria could be customized depending on the organization's risk appetite and risk tolerance [15].

5) Training/procedure understanding

In the training phase, the lecturer driver the team member to understand the FMEA improvement process.

6) Brainstorming

Brainstorming potential failures had previously been identified as critical assets. Identifying critical assets was carried out to determine the information technology assets the case study owned were to examine. The flow of the brainstorming process is shown in fig. 3.

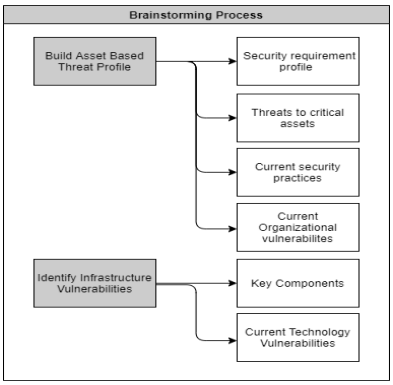


Fig. 3. Brainstorming Process

The first step was to build an asset-based threat profile. The activity in building asset-based threat profiles was to identify the security requirement profile, threats, current security practice, and ongoing organizational vulnerabilities. Identifying infrastructure vulnerabilities was to know the critical component used to support the system and identify present technology vulnerabilities. The result of the information obtained from the brainstorming stage was a risk register.

7) List of the risk register

The activity in this step was to fill out the document FMEA improvement.

8) Assess each parameters risk

The assignment of risk parameters did base on the criteria scale of FMEA improvement. The time estimation in risk assessment was less than 90 minutes.

9) RPN calculation

RPN calculation was the result of multiply the severity and occurrence value.

10) Risk prioritization

This risk level determined the risk that could be ignored or accepted, eliminated the source of the threat, mitigated risk, or monitoring the source of the threat.

11) Recommendation control documentation

This step was documentation for risk evaluation on the sustainability of risk assessments that had carried out control recommendations.

IV. Result and Discussion

A. RQ1. FMEA Improvement in a case study

The implementation of the FMEA improvement model in the IT risk assessment case study consists of ten stages.

1) Identify Context

The purpose of IT risk measurement was to implement FMEA Improvement in the case study of the Regional Office of the Ministry of Religion in Riau Province. The scope of the determination of the risk team involved the parties involved in risk management activities who knew about the agency's critical assets and understood the use of FMEA. The performance indicators are based on the suitability of the procedure implementation time with the procedure maker (agency)—the coordinator based on the researcher's control. The object of risk measurement focused on the Integrated Hajj Computerized System in the Field of Hajj and Umrah Administration, Regional Office of the Ministry of Religion of Riau Province.

The result of the business process analysis was a list of critical assets that need risk measurement. The list of critical assets were categorized by hardware, software, people, data, and network [30]. Descriptions of critical assets are obtained from the results of interviews that have been carried out regarding critical IT assets in the Field of Hajj and Umrah Administration.

The Critical Asset Components identified are described in Table I.

TABLE I. Critical Asset List

Asset	Asset Detail
Hardware	Server, PC, Intranet and Internet Networks device, Printer/Scanner.
Software	Antivirus, PC / Server Operation System, JRE, Microsoft Office
People	Head of Division, Head of Section, Staff, and Operators
Data	Hajj pilgrimage data, departure schedules, portion numbers, cancellation of pilgrims, financial and audit, travel, hajj officers, KBIH.
Network	Internet, intranet

2) Identify Business Process

At the provincial level, the business processes were a valuable system for monitoring pilgrims' registration and departure to pilgrims' return. The business process included the registration process of Hajj Plus, validation or examination of Hajj documents, cancellation of Hajj, monitoring (the number of pilgrims, the number of the registration of prospective pilgrims per day).

3) Establish FMEA Team

The team consisted of a Head of System Section, an IT practitioner, and one person as the team coordinator.

4) Determine assessment model

Used of FMEA Improvement documents proposed by [15].

5) Training/ procedure understanding

The training mechanism was explained in table 2, as follow:

TABLE II. Training Procedure

Pre-Training	Training
1. Preparation of training modules/materials 2. Presenters 3. Participants 4. Training schedule 5. Supporting equipment	Submission of material and demo filling the FMEA score.

6) Brainstorming Potential Failure (failure mode, potential cause, potential effect)

The brainstorming process is carried out to obtain a list of risks, which later will make it easier for the team to compile a risk register based on the FMEA improvement document. The first activity was identifying the security requirements profile. Each critical asset list (Table I) identified its security needs based on three security principles: confidentiality, integrity, and availability. The second activity was identifying the threat (internal and external).

The results were a list of threats in critical asset aspects. The third activity was identifying the current security practice. Case studies need to implement several strategies to prevent risk. The strategies were: (1) used the CCTV and room partitions; (2) reminders to turn off IT devices; (3) use one-way entrances; (4) use of VPN networks; (5) strategic environmental conditions; (6) restrictions on the system access; and (7) change passwords regularly at the central office. The fourth activity was identifying the current organizational vulnerabilities. The results were: (1) centralized server; (2) The server location is on the ground floor; (3) lack of IT risks awareness by agencies; (4) lack of maintenance and monitoring; (5) Lack of UPS and generator sets, (6) there is no user guidance system, (7) lack of employee training.

The main component explained the use of infrastructure and information that were available in the case studies. The main components were the server, VPN network, Personal Computer, antivirus, JRE Software, Operating System, and People. The current technology vulnerabilities were threats that existed in each of the main components.

7) List of the risk register

At this stage, it filled the risks identified in the FMEA Improvement document [15].

8) Assess each parameters risk

Teams followed the direction of the FMEA team coordinator. According to the research scenarios designed, the assessment duration was limited to less than 90 minutes.

9) RPN calculation

Based on the FMEA model modification, the critical parameters in risk analysis were severity and occurrence. Following the determination valuation method, the calculation of the value of the RPN by multiplying the value of severity (Sev) and occurrence (Occ).

10) Risk prioritization

The results shown in Table III, one risk was at the highest level(H) with the risk code 'NT03'. The risk of network connectivity decreases due to network failure. This risk was classified as severe because it often happens almost every day, and employees feel uncomfortable with the situation. Then, two risks were at the medium to the high level(MH), three risks were at the medium level(M), 13 risks were at the low to medium level(LM)l, and 18 risks were at the low level(L).

TABLE III. RISK PRIORITIZATION

Code	**(Risk)** ***Potential Failure Mode***	**(Threat) Potential causes/mechanisms**	**Level**
NT03	Network connectivity decreases	Network failure	H
DA07	Corrupt Data	Less optimal internet network	MH
NT01	Network connection lost	Network failure	MH
HW04	*Server* down	server access at the same time / DDOS attack	M
SW02	Virus attack	Antivirus is not able to detect and prevent incoming viruses	M
DA03	Data/information breach	Dissemination of confidential information by employees (share passwords)	M
PP03	Falsification or abuse of access rights	Collaboration with outside parties to falsify signatures recorded on the system	LM
DA02	The spread of confidential information	Misuse of access rights	LM
PP01	Human Failure	Error input data and using system devices	LM
PP02	Human Failure	Human Resources are less competent	LM
HW01	Server fire	*The server is overheating*	LM
HW12	Loss of PC components	Theft	LM
HW17	Loss of network device components	Theft	LM
HW08	The computer cannot be used	Error in computer configuration	LM
HW09	The computer cannot be used	The software license used has exceeded the time limit	LM
HW13	Access PC information illegally	Guard is weak, and or the computer is not given a password.	LM
HW14	Network failure	Damage to network infrastructure	LM
HW15	Network failure	Manipulation of network configurations.	LM
SW01	Software failure	The software license used has exceeded the time limit	LM
HW02	*Server fire*	Short-circuiting (*power failure*)	L
HW03	*Server overheat*	Air Conditioner does not function in the server room	L
HW05	Server failure	Lack of routine controlling and maintenance processes	L
DA04	Data does not match the system with physical data	Data input error	L
HW07	Computer Damage	Virus attack	L
HW10	The computer cannot be used	Natural disasters (fire, flood, lightning)	L
HW20	Loss of printer/scanner	Theft	L
DA01	Full capacity	Lack of memory control	L
NT02	Network connection lost	Network device damage/power failure	L
HW06	Server failure	Natural disasters such as being hit by building debris (server located on the lower floor)	L
HW11	PC out of date	Outdated technology used	L
HW16	Network Device Damage	*the force of nature and or animals*	L
HW18	Printer/scanner damage	*Maintenance and control that is not routine.*	L
HW19	Printer/scanner damage	*Force of nature*	L

SW03	System failure	The system still has security holes	L
DA05	Lost Data	Software and network failure	L
DA06	*Cybercrime* (hacker attack)	Lack of security on the system (firewall)	L
NT04	IP address error	*Human error*	L

11) Recommendation control documentation

Regarding the level of risk, the risks are classified as low to medium, and no action is taken. The medium to high and high action plan must exist. Action plans depended on the ability of the organization to deal with risk issues at that level. Risk elimination is carried out to eliminate the risk threat directly. Mitigation is carried out to manage risk by providing control over risks (treatment). While risk transfer was the transfer of risk to a third party, the organization could use outsourcing services to handle the risk.

B. RQ2. The implications of the FMEA Improvement model

The FMEA is used widely and commonly carried out in risk management. FMEA used FMEA documents (worksheets) to analyze, consistency, and documentation structure [31]. FMEA Weakness identification was carried out based on FMEA documents, which analyzed FMEA weakness points in each risk assessment process [12]. The following was the identification of FMEA weaknesses based on previous research, namely:

1) Risk register (the potential failure, causes, and control)

Failure mode was an observation of the impact of a failure or risk. This failure would affect the organization. Bias focuses on low RPN values, but the risk of the consequences of considerable severity must be prevented. If the risk of a slight severity is calculated with a high frequency of occurrence, it would be a significant risk [32].

The risk register's weakness was that the three risk factors are difficult to evaluate correctly [33]. Many risks with different risk scenarios also influenced difficulties in assessing risk appropriately. The many variations of risk scenarios produced identical RPN values. FMEA did not allow the assessor to distinguish between different risk implications [34].

2) Assessment and criteria scale (severity, occurrence, and detection)

A risk assessment by FMEA used risk factors or parameters (*severity, occurrence, and detection).* Team FMEA assessed each risk register by scale criteria of parameters. Scale criteria had a range of 5, 7, or 10, similar to Likert's [33]. In determining this scale, there is no specific procedure. Risk assessment was according to the customization of the risk gauge. However, the most commonly used is a scale of 1 to 10. Other criteria scales used a 1-3 or 1-5 level. The difference criteria scales can be used for the three parameters [8].

This scale criterion becomes a problem if unclear definitions and vague boundaries. The research conducted by [35] did the modification or improvement methodology in FMEA scale criteria. The modification of scale criteria could minimize the limitation of FMEA. Generally, team FMEA used a 1-10 scale in risk assessment. The disadvantage of the range scale 1-10 was time-consuming in risk assessment. The team thought longer in determining the right scale because of the many considerations of the correct numbers. The research defined the scale (1,3,9) in giving the value of each parameter with a level (high, medium, low). The limitation of the range scale could make the FMEA a faster method and be more effective and high-reliability. It was not easy for the expert to provide the correct input values for risk factors [33]. These criteria scale significantly affected the results obtained on the RPN risk register. The value of the criteria scale was practical to identify the most severe risks for corrective action. The three parameters were not the same weight as each other in terms of risk. This distortion is exacerbated by the non-linear nature of individual rating scales (1-10) [36].

3) FMEA Team

There was a subjectivity issue because of human error and bias in risk assessment. The subjectivity was the limitation of FMEA because it was carried out based on human emotions and thoughts. There were probably doubts about the accuracy of the result. The FMEA team would feel it difficult to determine the different opinions that occurred in the risk assessment. There were variables needed to calculate risk figures that were inappropriate and doubtful [37].

Another weakness was the differences of opinion in determining the value of each FMEA parameter. The same RPN value might happen because there was no ability of the FMEA team to articulate risk implications [34]. FMEA team consisted of a multidisciplinary team, thus it needs time to understand the process analyzed adequately [38]. There were no clear guidelines or procedures regarding the team members. FMEA was a more explicit method than other methods [31].

4) RPN (Risk Priority Number)

Risks with the highest RPN score were assumed to be essential risks and got the most top priority. [36]. Traditionally, the decision to improve an operational process is based on the prioritization of the RPN. This method was compelling and helpful in conducting risk assessments. RPN scores also had disadvantages, namely [39]:

- RPN scores are not considered the relative importance of each parameter.
- The identical RPN scores came from the combination parameters that had different risk implications.
- It was not easy to assess the parameters.
- The formulation of RPN still needed more deeply discussion because it was not keen to evaluate the critical factors.

RPN problems due to the same value parameters were significant to evaluate. Based on the literature review conducted by [33], the weakness of FMEA are that the relative importance of parameters was not considered. The conventional RPN did not weigh because the parameters (*severity*, *occurrence*, dan *detection*) had the same importance level. A large number of combination variations parameters scale could cause the duplicate RPN value. Thus, the equal RPN value for low and high implications is risk-targeted for mitigation [40].

The combination parameters might produce the duplicate RPN but had essence values that might be different and hidden [33]. The severity and the occurrence were the primary keys compared to detection parameters [11]. For example, there was a risk of a damaged system. There were two components of RPN, which had the same value, namely RPN with a value of 100 (RPN1 = 10(S) x 5(O) x 2(D), RPN2= 10 x 2 x 5). From both RPNs, the priority mitigation was the same. However, priority mitigation was for the top risk level. The first risk had the occurrence value was the primary key in the example of the problem. The detection factor was not crucial in the measurement. Whereas, in that case, the severity value was equal to 10.

According to [36], the weakness of RPN was a simple method or formulation. The subjectivity issue in risk assessment could affect the inconsistency result of the RPN. Scenario results from RPN could benefit from combining the parameters with a low RPN value but had a dangerous potential impact. For example, in the calculation of RPN, the failure mode (risk), which had a very high severity, a low rate of occurrence, and has a very high detection value (in scale: Severity 9, occurrence three and detection 2) with RPN score was 54. In comparison, other risks had an RPN value of 120 with severity 4, occurrence 5, and detection 6. The risk of having an RPN 54 should have a high priority than the RPN, whose value is 120.

On the other hand, RPN evaluation and prioritization are widely criticized by researchers. There was a condition in which a team had a different view in providing an assessment. So, it recommended taking the average value without considering the three values of the parameter. The FMEA team gave top priority to failure modes that had a higher severity. In practical situations, all parameters were equally important, and it must evaluate the same weight in determining RPN priorities [41].

The research examined the consistency results using Traditional FMEA and FMEA Improvement[15]. The study results prove that the results of risk measurement with improvement FMEA are more consistent than FMEA traditional. Two different teams and the same case study were carried out on Consistency testing.

V. CONTRIBUTION IMPLICATION

FMEA Improvement results from synthesizing the FMEA traditional method that gave the strategies in process FMEA[15]. FMEA improvement had more details and procedures than FMEA traditional. The time estimation for team FMEA assigned the severity and occurrence parameters was no more than 90 minutes. The categorization risk effect divides into service/operational risk, media attention risk, and regulatory risk, aligned with the severity criteria scale. Thus, the risk assessed was suitable with the description of the severity parameter.

FMEA improvement has added a source of threat column in document FMEA. There were three types of source threats (people, process, and technology)in the source of the threat column. Then, the criteria scale used in FMEA Improvement was five range scales. The primary critical parameters in FMEA improvement used severity and occurrence parameters. Besides, the detection variable described in current compensating controls (Compensate Vulnerability) consists of preventive control and detective control. This paper presented the detailed steps of FMEA improvement and explained the weaknesses of Traditional FMEA. The use of FMEA Improvement in the case studies examined in this study results in more systematic and easily applied measurements.

As a practical contribution, FMEA improvement was a guidance for the FMEA team in IT risk assessment. The modification of FMEA (FMEA improvement) did not remove the characteristic of FMEA (ease of use). FMEA improvement gave the strategies and specific steps in IT risk assessment. As a future researcher, FMEA improvement needs more exploration and criticism from many points of view.

VI. CONCLUSIONS AND FUTURE WORKS

FMEA improvement consists of four main stages: determining risk assessment requirements, risk identification, risk analysis and evaluation, and recommendation control. Based on the results of risk level categorization, one risk is at the highest level. This risk was classified as severe because it often happens almost every day. Then, two risks were at the medium to the high level, three risks were at the medium level, 13 risks were at the low to medium level, and 18 risks were at the low level. In subsequent studies, it can test the FMEA Improvement model at other agencies or companies. The critical analysis of FMEA Improvement is also an exciting topic for future works. FMEA improvement also can be combined with other risk management methods to obtain a more effective IT risk assessment.

REFERENCES

[1] Sidorova, Evangelopoulos, Valacich, and Ramakrishnan, "Uncovering the Intellectual Core of the Information Systems Discipline," *MIS Q.*, vol. 32, no. 3, p. 467, 2008.

[2] M. E. Whitman and H. J. Mattord, "Principles of information security," *Course Technol.*, pp. 1–617, 2012.

[3] S. Gary Teng, S. M. Ho, D. Shumar, and P. C. Liu, "Implementing FMEA in a collaborative supply chain environment," *Int. J. Qual. Reliab. Manag.*, vol. 23, no. 2, pp. 179–196, 2006.

[4] J. F. W. Peeters, R. J. I. Basten, and T. Tinga, "Improving failure analysis efficiency by combining FTA and FMEA in a recursive manner," *Reliab. Eng. Syst. Saf.*, vol. 172, no. April, pp. 36–44, 2018.

[5] R. E. McDermott, R. J. Mikulak, and M. R. Beauregard, *The Basic of FMEA*, 2nd ed. New York: Taylor & Francis Group, 2009.

[6] A. Pai H, S. H S, S. Soman, and D. P. K. Pareek, "Analysis of Causes and Effects of Longer Lead Time in Software Process Using FMEA," *SSRN Electron. J.*, no. May, 2019.

[7] C. S. Carlson, "Understanding and Applying the Fundamentals of FMEAs," *2014 Annu. Reliab. Maintainab. Symp.*, p. 12, 2014.

[8] J. F. van Leeuwen *et al.*, “Risk analysis by FMEA as an element of analytical validation,” *J. Pharm. Biomed. Anal.*, vol. 50, no. 5, pp. 1085–1087, 2009.
[9] A. P. Subriadi, N. F. Najwa, B. D. Cahyabuana, and V. Lukitosari, "The consistency of using failure mode effect analysis (FMEA) on risk assessment of information technology," in *2018 International Seminar on Research of Information Technology and Intelligent Systems, ISRITI 2018*, 2018, pp. 61–66.
[10] M. Yazdi, "Improving failure mode and effect analysis (FMEA) with consideration of uncertainty handling as an interactive approach," *Int. J. Interact. Des. Manuf.*, vol. 13, no. 2, pp. 441–458, 2019.
[11] N. Xiao, H. Z. Huang, Y. Li, L. He, and T. Jin, "Multiple failure modes analysis and weighted risk priority number evaluation in FMEA," *Eng. Fail. Anal.*, vol. 18, no. 4, pp. 1162–1170, 2011.
[12] C. Estorilio and R. K. Posso, "The reduction of irregularities in the use of 'process FMEA,'" *Int. J. Qual. Reliab. Manag.*, vol. 27, no. 6, pp. 721–733, 2010.
[13] D. M. Barends, M. T. Oldenhof, M. J. Vredenbregt, and M. J. Nauta, "Risk analysis of analytical validations by probabilistic modification of FMEA," *J. Pharm. Biomed. Anal.*, vol. 64–65, pp. 82–86, 2012.
[14] M. T. Oldenhof *et al.*, "Consistency of FMEA used in the validation of analytical procedures," *J. Pharm. Biomed. Anal.*, vol. 54, no. 3, pp. 592–595, 2011.
[15] A. P. Subriadi and N. F. Najwa, "The consistency analysis of failure mode and effect analysis (FMEA) in information technology risk assessment," *Heliyon*, vol. 6, no. 1, p. e03161, 2020.
[16] G. Stoneburner, A. Goguen, and A. Feringa, "Risk Management Guide for Information Technology Systems Recommendations of the National Institute of Standards and Technology," *Nist Spec. Publ. 800*, vol. 30, no. July, 2002.
[17] M. Spremic and M. Popovic, "Emerging issues in IT Governance : implementing the corporate IT risks management model," *Wseas Trans. Syst.*, vol. 7, no. 3, pp. 219–228, 2008.
[18] D. G. Firesmith, "Common Concepts Underlying Safety, Security, and Survivability Engineering," no. December, 2003.
[19] K. D. Loch, H. H. Carr, and M. E. Warkentin, "Threats to Information Systems: Today's Reality, Yesterday's Understanding," *MIS Q.*, vol. 16, no. 2, p. 173, Jun. 1992.
[20] K. Bandyopadhyay, P. P. Mykytyn, and K. Mykytyn, "A framework for integrated risk management in information technology," *Manag. Decis.*, vol. 37, no. 5, pp. 437–445, 2011.
[21] L. K. H. Lai and K. S. Chin, "Development of a Failure Mode and Effects Analysis Based Risk Assessment Tool for Information Security," *Ind. Eng. Manag. Syst.*, vol. 13, no. 1, pp. 87–100, 2014.
[22] K. Claxton and N. M. Campbell-Allen, "Failure modes effects analysis (FMEA) for review of a diagnostic genetic laboratory process.," *Int. J. Qual. Reliab. Manag.*, vol. 34, no. 2, pp. 265–277, 2017.
[23] C. Özkan, "Fuzzy FMEA Aplication Combined with Fuzzy Cognitive Maps to Manage the Risks of a Software Project," *Eur. J. Eng. Form. Sci.*, vol. 2, no. 2, 2018.
[24] L. S. Lipol and J. Haq, "Risk analysis method : FMEA / FMECA in the organizations.," *Int. J. Basic Appl. Sci.*, vol. 11, no. 5, pp. 5–74, 2011.
[25] R. K. Sharma and P. Sharma, "System failure behavior and maintenance decision making using, RCA, FMEA and FM," *J. Qual. Maint. Eng.*, vol. 16, no. 1, pp. 64–88, 2010.
[26] H. W. Lo and J. J. H. Liou, "A novel multiple-criteria decision-making-based FMEA model for risk assessment," *Appl. Soft Comput. J.*, vol. 73, pp. 684–696, 2018.
[27] W. Wang, X. Liu, Y. Qin, and Y. Fu, "A risk evaluation and prioritization method for FMEA with prospect theory and Choquet integral," *Saf. Sci.*, vol. 110, no. December 2017, pp. 152–163, 2018.
[28] H. C. Liu, L. E. Wang, Z. Li, and Y. P. Hu, "Improving risk evaluation in FMEA with cloud model and hierarchical TOPSIS method," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 1, pp. 84–95, 2019.
[29] Z. Li and L. Chen, "A novel evidential FMEA method by integrating fuzzy belief structure and grey relational projection method," *Eng. Appl. Artif. Intell.*, vol. 77, no. October 2018, pp. 136–147, 2019.
[30] I. Desy, B. Cahyo, H. Hanim, and M. Astuti, “PENILAIAN RISIKO KEAMANAN INFORMASI MENGGUNAKAN METODE FAILURE MODE AND EFFECTS ANALYSIS DI DIVISI TI PT . BANK XYZ SURABAYA,” *Semin. Nas. Sist. Inf. Indones.*, no. September, 2014.
[31] C. Raspotnig and A. Opdahl, "Comparing risk identification techniques for safety and security requirements," *J. Syst. Softw.*, vol. 86, no. 4, pp. 1124–1151, 2013.
[32] I. Cameron *et al.*, "Process hazard analysis, hazard identification and scenario definition: Are the conventional tools sufficient, or should and can we do much better?," *Process Saf. Environ. Prot.*, vol. 110, pp. 53–70, 2017.
[33] H. Liu, L. Liu, and N. Liu, "Expert Systems with Applications Risk evaluation approaches in failure mode and effects analysis : A literature review," *Expert Syst. Appl.*, vol. 40, no. 2, pp. 828–838, 2013.
[34] R. Sawhney, K. Subburaman, C. Sonntag, P. Rao Venkateswara Rao, and C. Capizzi, "A modified FMEA approach to enhance reliability of lean systems," *Int. J. Qual. Reliab. Manag.*, vol. 27, no. 7, pp. 832–855, 2010.
[35] C. Paciarotti, G. Mazzuto, and D. D’Ettorre, “A revised FMEA application to the quality control management,” *Int. J. Qual. Reliab. Manag.*, vol. 31, no. 7, pp. 788–810, 2014.
[36] R. N. Sankar and B. S. Prabhu, "Modified approach for prioritization of failures in a system failure mode and effects analysis," *Int. J. Qual. Reliab. Manag.*, vol. 18, no. 3, pp. 324–336, 2001.
[37] M. Kakvan, M. A. Mohyeddin, and H. Gharaee, "Risk evaluation of IT service providers using FMEA model in combination with Multi-Criteria Decision-Making Models and ITIL framework," *2014 7th Int. Symp. Telecommun. IST 2014*, pp. 873–878, 2014.
[38] K. Jain, "use of Failure Mode Effect Analysis (FMEA) to Improve Medication Management Process," *Int. J. Health Care Qual. Assur.*, vol. 30, no. 2, 2017.
[39] K. C. Chai, C. H. Jong, K. M. Tay, and C. P. Lim, "A Perceptual Computing-based Method to Prioritize Failure Modes in Failure Mode and Effect Analysis and Its Application to Edible Bird Nest Farming," *Appl. Soft Comput. J.*, 2016.
[40] M. Banghart, "Utilizing Confidence Bounds in Failure Mode Effects Analysis (FMEA) Hazard Risk Assessment," 2014.
[41] N. Sellappan, D. Nagarajan, and K. Palanikumar, "Evaluation of risk priority number (RPN) in design failure modes and effects analysis (DFMEA) using factor analysis," *Int. J. Appl. Eng. Res.*, vol. 10, no. 14, pp. 34194–34198, 2015.

Implementation of Hill Climbing Algorithm on Tourist Attraction Android Based Application

1th Ramadhani, Shumaya Resty
Information of Technology
Politeknik Caltex Riau
Pekanbaru, Indonesia
shumaya@pcr.ac.id

2th Kreshna, Jan Alif
Information of Technology
Politeknik Caltex Riau
Pekanbaru, Indonesia
janalif@pcr.ac.id

3rd Rianda, Septian
Information of Techology
Politeknik Caltex Riau
Pekanbaru, Indonesia
septianrianda@alumni.pcr.ac.id

***Abstract*—West Sumatra is one province that has been a popular destination among tourists. West Sumatra is a strategic area that holds natural beauty that is no less great, and this is evenly distributed throughout the region. Data shows that the number of tourist visits continues to increase every year before the pandemic hits Indonesia. Travelers often use google maps to find routes to destinations. However, with little knowledge about the area, people only visit familiar tourist locations. Meanwhile, West Sumatra offers plenty of beautiful places. With those problems in mind, a mobile application is created to help those visitors searching the nearest tourist spots around them in real-time. Not only that, but the system also gives information about the list of places worth visiting in West Sumatra. The system implemented a hill-climbing algorithm to find the shortest path. Based on the evaluation, the application is well accepted for people who want to visit West Sumatra. According to survey conducted on this research, overall 94.6% of users agree that this app is useful. The app is handy and fast enough to display the desired result to the users.**

Keywords—hill climbing, android, LBS, West Sumatra, cyclomatic complexity

I. INTRODUCTION

Travelling is one of the activities that most people would love to do. Indonesia is a country that offers many wonderful places and views for domestic and international tourists to explore. In 2019 itself, there was around 1.3777.067 foreign tourist that visits Indonesia through international airport [1]. Although the number experienced a slight decline compared to 2018, which is around 2.03%, it is still appreciated. West Sumatra is one of the tourist destinations that are popular in Sumatra.

According to the Ministry of Tourism, the average number of international tourists visiting West Sumatra is around five thousand people each month [2]. Most of them may spend one to three days minimum in this province to enjoy the scenery [3]. This number shows that West Sumatra is attractive as it brings beautiful scenery, delicious food, and a friendly community for people worldwide.

Lack of tourism information is one of the reasons why the visitor number in West Sumatra remains low. The hidden gem spots are unreachable and remain unknown, while that could have been a potential attraction to many tourists. Aside from that, many travelers wasted their time finding the route to their destination and ended up missing the spots on the way. They visit only the most popular places without looking at other locations, especially in West Sumatra. People use Google Maps to find locations, yet it displays only specific information and route. In contrast, the information of the spots nearby is needed as well to travelers.

Certainly, by this fact, the need for a system that describes tourist spots is necessary. The system's goal is to find the nearest tourist attractions from the user position as well as the route to the location. Additionally, the system could provide a piece of clear information about the tourist objects. Thus, users can decide which place they would like to visit. Furthermore, the system could be handy and simple; thus, users feel comfortable using it while traveling. By having this system, travelers could maximize their trip to West Sumatra.

This application is designed using the Simple Hill Climbing method and using the Global Positioning System (GPS) function on the smartphone to detect user location. The Simple Hill Climbing method is relatively simple compared to other search optimization methods. However, the execution speed of this method could be a point to be considered, since the algorithm is implemented in mobile phones.

II. LITERATURE REVIEWS

A. Previous Research Studies

The Hill climbing search algorithm is commonly used to find the shortest path to the destination. There have been numerous studies related to this algorithm. One of the studies uses this method to calculate the distance between attractions. Not only that, but researchers also used the method to optimize the route to the 5 specific popular locations in every province of Indonesia. The study shows that the accuracy of determination from the research is 93.3% [4].

Another implementation that uses this method is to determine the shortest path for kerosene distribution. The system is an android based application for the agents and customers in Ambon city. Customers can see where the

nearest agents are around them. Meanwhile, the agents get the optimal route to distribute the products around the city [5].

Interestingly, some studies also show that hill-climbing could be implemented in Games. The research used simple hill-climbing in crossword puzzles to get a wording pattern in empty boxes. Hence, the more wording combination created to fill the boxes, the solution could be optimal. This algorithm is implemented automatically to create computer games [6].

B. Hill Climbing Search Algorithm

One of mathematical optimization method is called hill climbing for local search [7]. Hill climbing is one of the heuristics techniques to find rapid solutions compared to classic methods. Using the heuristic technique, hill climbing could reduce the amount of time needed on iteration processes to solve problems. However, it may not produce the best solution.

In heuristic, hill climbing consists of two types, simple hill-climbing, and steepest-ascent hill climbing. The algorithm that is used in this study is the simple hill climbing and implemented to find the nearest tourist location in West Sumatra. This method searches for better solutions in its surrounding by evaluating the current state. When the goal state finds that the heuristic value is less than the initial, the iteration process stops. The process may keep going when there is no better result, and there are operators left. Here is the pseudocode of this algorithm.

```
INITIALIZE initial state, current state, goal state
IF initial state != goal state
      WHILE neighbors != 0
        current state = neighbors
          IF current state < initial state
            THEN initial state = current state
          END
      END WHILE
END

IF initial state = goal state
RETURN "Success"
END
```

Hill climbing algorithm is straightforward to understand. Moreover, the execution process is fast and easily adapt to the case study [8]. Therefore, the consideration of implementing this algorithm in a mobile application is acceptable to get a fast solution.

C. Android

Android is an operating system that generally runs on mobile devices. Android was originally developed by a company called Android Inc. which Google later purchased in 2005. Devices that embed the Android operating system were first launched in 2007, around 13 years ago [9]. This Android-based operating system is currently prevalent and has become a popular choice for users of smartphones, tablet PCs, and Smart TVs. The development of Android opens up huge market opportunities for mobile app developers to compete in developing various types of mobile applications that run on this operating system.

D. Location Based Services

Location-Based Services (LBS) is the concept of an application that uses a combination of geographical location and services [10]. This service could be accessed in almost all mobile devices through network in order to get spatial coordinates.

Even though this service is popular and provides benefits to applications, collecting user locations as information is quite sensitive. Systems should not store the private data arbitrarily or without a permit. Hence the usage of these services should be restricted or need proper security improvement [11].

This implementation uses LBS from mobile devices to get the user's location by taking advantage of GPS. Application may ask user permission to enable GPS. When declined, then users cannot use the searching feature. Nevertheless, this application will not save user location.

III. RESEARCH METHODS

A. Implementation of Hill Climbing

This application compares the distance obtained from the user's location to the nearest tourist location from the user. This application will ask for tourist data that has been inputted into firebase. After the data is obtained, the data is stored into an Array List which later will be used as a comparison of the distance between tours hence the process will be faster rather than requesting from database.

After the data has been taken, the next stage is finding the nearest tourist attraction using the Simple Hill Climbing algorithm. At this stage, the user's location is the initial node, and the data is collected in the destination node. Then a comparison of the initial node with the existing tourist data nodes is carried out. After getting the results using the Simple Hill-Climbing algorithm, the route will be created using Google Direction to find the shortest route. The Simple Hill-Climbing algorithm in this application is written using the Java programming language. By this way, the algorithm could work efficiently before displaying the result as markers on screen.

B. Data Collection Method

The tourism spot data is obtained from the official web of West Sumatra is https://sumbar. travel/ and https://dispar.sumbarprov.go.id/. Hence the data could be more valid as it is registered and publicly known by people. There are 37 tourist attractions in West Sumatra spread from one district to another. Some of the most popular objects are jam gadang, lembah anai, lembah arau, air manis beach and more. The data obtained will be used to reference the list of tourist locations in the android application.

IV. IMPLEMENTATION

This chapter explains how the system is implemented. It contains the system architecture, flowchart, requirements, and the result of a mobile-based application.

A. Architecture System

Fig. 1 is the architecture of the mobile application. The app gets the user's location using GPS and saves it into Firebase. The next step is implementing the simple hill-climbing

algorithm to calculate and find the shortest path between objects. Lastly, the output is mapped using google maps on android.

Fig. 1. System Architecture

This app needs an internet connection to access objects' locations through google API and Firebase. There is no local storage used in this mobile app. The application used google API to retrieve additional information about the objects.

B. *Flowchart System*

Fig. 2 below shows the flow of the android based mobile application. Once the app is on, it asks the user to activate the GPS. This process is needed to get the current user position. Application retrieves all attractions in West Sumatra and information and display it on the map on the application.

The system will implement the algorithm after all location data is obtained from Firebase and API. The algorithm calculates the distance between one point to another and finds the shortest path. The location range is set to five kilometers to prevent such a far location from arising on the app. When the result shows no data found, it displays a notification and asks the user to set another starting position.

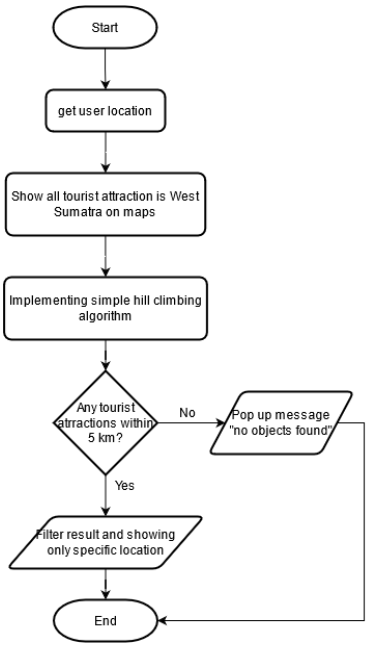


Fig. 2. Flowchart of the system

C. *Usecase*

Usecase diagram as it shows in Fig. 3 is made according to the user's requirements. The stakeholder of this mobile app is the tourists or visitors who want to visit West Sumatra.

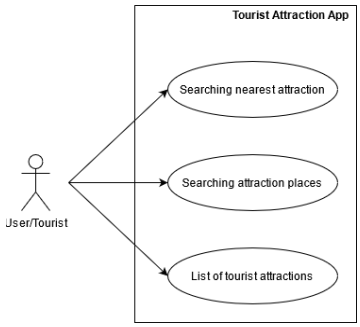


Fig. 3. Usecase application

The focus of the mobile app is to automatically find the nearest place within the designated range to the user position. Besides that, users can also find specific places by inputting the location into the system. As an additional feature for new visitors, the app also provides a list of tourist attractions in West Sumatra.

D. *Interface*

Pictures below display the interface of android based application. Fig. 4 is the page to find the tourism spot nearby.

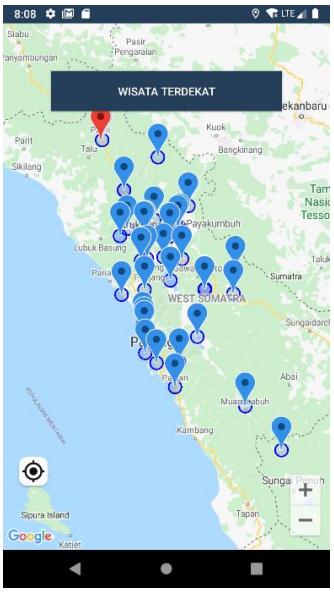


Fig. 4. Finding nearest location page

GPS detects user position automatically once it is activated. When the user presses the button, the app displays the result within the area, as shown below. The red marker is the user location, and the blue is the spot (Fig. 5).

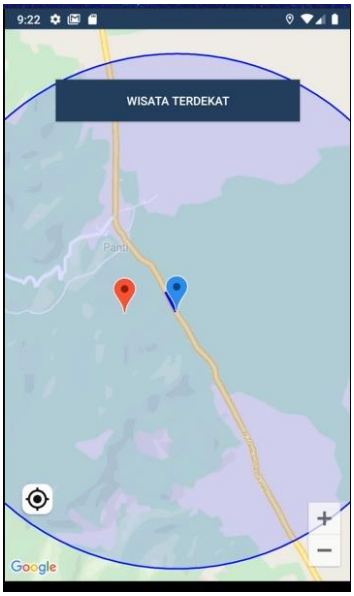


Fig. 5. The nearest objects within range

This page is used to search tourist locations that have been stored in Firebase as a dropdown list. User selects location and map shows the way to specific input by giving route. Application displays the coordinates as it shown in Fig. 6.

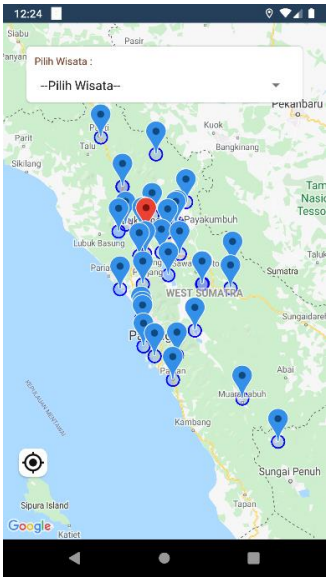


Fig. 6. Searching location page

This page in Fig. 7 displays all tourism spots as a list. Users can click on an item and get all information and description that the user might need.

Fig. 7. List of tourist attractions

If the user clicks on each item, the app displays details and information about the objects (Fig. 8). The user could read the story behind the tourist spot and take a look at the gallery it provides.

Fig. 8. Details page

V. EVALUATION

The researcher conducted several evaluations for this implementation, such as counting time execution, checking the algorithm's complexity, and giving questionnaires to several users.

A. Computational Time

Testing on-time execution is mandatory since the application could load numerous object locations in a short amount of time. The evaluation of this testing is to ensure that a simple hill-climbing algorithm works in fair time to the phone user. This testing used all data locations stored in Firebase. However, the tourism spot in West Sumatra that the writer found is limited, which is less than 30 locations.

Therefore, to test loading time for the app, dummy data is added into storage from a thousand up to a million points. Here are the details of the average loading time:

TABLE I. TABLE 1. LOADING TIME

No. of Iteration	*Total Data Coordinates*	*Avg. Loading Time (s)*
1.	1.000	0.003
2.	10.000	0.004
3.	100.000	0.015
4.	1.000.000	0.111

The data above shows that it takes a longer time to execute expandable data from storage. It needs time to retrieve all data from the cloud and then process it using a hill-climbing algorithm before inserting the result as a marker to the map. Nevertheless, the result of this evaluation is still tolerable, especially for mobile apps. The average time remains under 0.5 second for around a million-location data, hence could be considered hence it has less impact on user experience.

B. Cyclomatic Complex (CC)

This Cyclomatic Complexity test aims to assess whether the code of the Simple Hill-Climbing algorithm has been written well. The metrics could help identify the part of the code that can potentially lead to error and possibly lead to ineffectiveness of algorithm implementation [12].

The test using embedded Plugins from Android Studio, namely Metrics Reloaded [13]. This plugin calculates the value of Cyclomatic Complexity for all projects in Android. The v(G) is the cyclometric complexity, ev(G) is the essential cyclometric complexity, and iv(G) is the Module Design Complexity Metric.

TABLE II. TABLE 2. RESULT OF METRICS RELOADED ON SIMPLE HILL CLIMBING

method	*ev(G)*	*iv(G)*	*v(G)*
LokasiTerdekat.simpleHillClimbing()	1	4	4

Based on the results above, v(G) = 4 is obtained for the Simple Hill Climbing algorithm. The result shows that the code implementation has low complexity [14]. However, this metrics (CC) has disadvantages that cannot measure the satisfaction of code understandability [15].

C. Questionnaire Evaluation

The researcher surveyed this mobile app to gather knowledge about user response. This survey was conducted on 30 respondents by using online media, Google Forms. The age of respondents is above 17 years old and loves to travel, especially to West Sumatra. More than half of respondents (73.3%) agree that this mobile application is helpful, specifically for tourists who never visit West Sumatra before.

Application features are also expected to meet user requirements. The survey shows that 50% of respondents agree that the features provided were sufficient, and the rest expect more additional focus besides tourism spots. For example, adding culinary and shopping locations.

92.6% of respondents recognize that the app is easy to understand. They were able to operate and navigate through the app without any problems. Moreover, the information provided in the mobile app is clear and straightforward. Therefore, the users do not need assistance when using the application.

VI. CONCLUSION AND FUTURE WORK

According to the implementation and evaluation above, a simple hill algorithm is a one of the proper optimization algorithms for finding the nearest tourist locations. The system is a mobile-based application; hence requires the enhancement of the searching speed before displaying the result to the user. According to the evaluation on computation time, the algorithm only needs less than 0.2 seconds to get the result from a huge amount of data and put the marker on the screen. Therefore, the user may get the information needed immediately.

The algorithm is also implemented with a good coding style showed by evaluating the outcome of the complexity metrics. However, there is still uncertainty on the code understandability since the evaluation did not measure the satisfaction of the code readability.

This algorithm could produce better results for future reference to find and recommend locations when combined with artificial intelligence methods. Therefore, the travelers could have recommendations of which place to visit. Additionally, the application could also cover the tourist locations and accommodation places, culinary, and added real-time traffic.

REFERENCES

[1] K. P. dan E. Kreatif, "Data Kunjungan Wisatawan Mancanegara Bulanan Tahun 2019," 2019. https://kemenparekraf.go.id/post/data-kunjungan-wisatawan-mancanegara-bulanan-tahun-2019 (accessed Jul. 12, 2021).

[2] S. Provinsi, "Perkembangan Pariwisata dan Transportasi Sumatera Barat Agustus 2019," 2019. https://sumbarprov.go.id/home/news/17484-perkembangan-pariwisata-dan-transportasi-sumatera-barat-agustus-2019 (accessed Jul. 12, 2021).

[3] P. Ekonomi and S. Barat, "Potensi Daya Tarik Obyek Pariwisata Dalam Pembangunan Ekonomi Sumatera Barat," *Economica*, vol. 1, no. 1, pp. 1–15, 2012, doi: 10.22202/economica.2012.v1.i1.101.

[4] D. Fitriati and N. M. Nessrayasa, "Implementasi Algoritma Hill Climbing Pada Penentuan Jarak Terpendek Kota Wisata Di Indonesia," *J. Ris. Inform.*, vol. 1, no. 3, pp. 127–132, Jul. 2019, doi: 10.34288/JRI.V1I3.35.

[5] R. G. Elly, "Penerapan Metode Simple Hill Climbing Untuk Penentuan Jalur Terpendek Penyaluran Minyak Tanah Pada Agen Berbasis Android - Akakom Repository," *Digital Library STMIKAKAKOM*, 2017. https://eprints.akakom.ac.id/3741/ (accessed Jul. 12, 2021).

[6] P. M. Hasugian, "Implementasi Algoritma Simple Hill Climbing Pada Aplikasi Game Tts (Teka – Teki Silang)," *J. Mantik Penusa*, vol. 19, no. 1, Jun. 2016, Accessed: Jul. 12, 2021. [Online]. Available: http://www.e-jurnal.pelitanusantara.ac.id/index.php/mantik/article/view/232.

[7] A. Schaerf and A. Meisels, "Solving Employee Timetabling Problems by Generalized Local Search," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1792 LNAI, pp. 380–389, 1999, doi: 10.1007/3-540-46238-4_33.

[8] M. Shehab, A. T. Khader, and M. Laouchedi, "A Hybrid Method Based on Cuckoo Search Algorithm for Global Optimization Problems," *J. Inf. Commun. Technol.*, vol. 17, no. 3, Jul. 2018, doi: 10.32890/JICT2018.17.3.4.

[9] "What is Android | Android." https://www.android.com/what-is-android/ (accessed Jun. 10, 2020).

[10] A. V. Jochen Schiller, "Location-Based Services - Google Buku," in *Location-Based Services*, San Francisco: Morgan Kaufmann

Publishers, 2004.

[11] N. Fei, Y. Zhuang, J. Gu, J. Cao, and L. Yang, "Privacy-preserving relative location based services for mobile users," *China Commun.*, vol. 12, no. 5, pp. 152–161, May 2015, doi: 10.1109/CC.2015.7112037.

[12] L. Dr. Rosenberg, T. Hammer, and J. Shaw, "Software metrics and reliability," *Proc. 9th Int. Symp. Softw. Reliab. Eng.*, pp. 1–8, 1998.

[13] B. Wang, "An Android Studio Plugin for Calculating and Measuring Code Complexity Metrics in Android Applications," Towson University, Maryland, 2016.

[14] "PMD - Rule Set: Code Size Rules." https://pmd.sourceforge.io/pmd-4.3.0/rules/codesize.html (accessed Jul. 14, 2021).

[15] G. A. Campbell, "Cognitive Complexity," Switzerland, 2016. doi: 10.21075/kacsn.2009.11.3.237.

Design and Implementation: JavaFX Face Detection with Scene Builder and NetBeans IDE

1st Syefrida Yulina
Information System Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
syefrida@pcr.ac.id

***Abstract*—JavaFX face detection is an application widely used to detect and recognize faces in digital images. One of the challenging problems in the image processing is how to develop and design an automatic face recognition application using JavaFX technology. JavaFX is a library of Java that is used to build rich internet applications (RIA) which can run across several platforms such as Desktops, Mobile Devices, TVs, Tablets, etc. Design and implementation of this application applies the concept of model-view-controller using framework Scene Builder and Netbeans IDE. Scene Builder is used as a tool to add components of GUI in the view that can produce document called FXML. Netbeans IDE is an integrated development environment for FXML document editing and maintaining the connection between view and controller. In this research stages of creating JavaFX Face Detection application are started with requirements identification, followed by design of UI components in Scene Graph, integration of the scene builder panels, and then making controller. An application is implemented using the stages of JavaFX.**

Keywords—Face Detection, FXML, Netbeans IDE, Scene Builder, Scene Graph

I. INTRODUCTION

Face detection applications are very necessary in the process of face recognizing. It can gain informations from image or video by using computer algorithms. Many previous studies have developed face detection application using certain programming language but lack of them using Java as a programming language to develop and implement an application based on graphical user interface.

JavaFX is next generation client application platform build on Java. It is a set of graphics and media packages that enables developers to design, create, test, debug, and deploy rich client applications that operate consistently across diverse platforms[1][2]. The interface and implementation of JavaFX application are defined separately from its behavior. By using *model-view-controller* technique, the controllers handle interactions of interface, while the views contain visual attribute/graphical components of interface[3][4]. JavaFX controller are written in Java programming language, and views in the framework are declared in FXML documents[5], where it is written in custom markup language based on the Extensible Markup Language(XML). A controller in JavaFX is used to make GUI interaction in the view class. The views are represented in a tree structure called a scene graph. Scene graph maintains the nodes (graphical components) in a branch node or the leaf node. The first node in tree is called the root node.

Three attributes to make interaction between nodes: 1) *fx:controller* attribute; 2) *fx:id* attribute; and 3) various *event handler* attributes. The *fx:controller* attribute is used to associate the controller with the view by setting the root node's value of the attribute to the name of the controller class. The *fx:id* attributes link field declaration in the controller with their corresponding component instances in the view, to enable programmatic manipulation. *Event handler* attributes assign method declaration in the controller as the recipients of control flow when events are fired by view components. Event handler are required to take a single argument of a type extending javafx.event.Event.

FXML is a textual data format, can be edited in text editor called Scene Builder. JavaFX scene builder enables to quickly design JavaFX application by dragging a UI component from a library of UI components and dropping it into a content view area[6][7][8]. The FXML code for UI layout is automatically generated. Scene Builder can be used as a standalone design tool and can be used in conjuction with Java IDEs. IDE can write, build, and run the controlller source code. JavaFX Scene Builder includes key features such as: a drag and drop WYSIWYG interface, tight interaction with the Netbeans IDE, automatic FXML code generation, live editing anf preview features, access to the complete JavaFX GUI controls library, ability to add custom GUI components to the library, 3D support, support for Rich Text, JavaFX Scene Builder kit, CSS support and cross-platform support.

In this paper, we design an application for face detection with Scene Builder, which integrated with Netbeans IDE. The integration enables Scene Builder by opening FXML document, run the application, and generate controller source file.

II. RESEARCH METHOD

A. Requirements Identification

This research use Netbeans IDE which facilitate JavaFX application development. It builds the connection between views and controller, and the mechanism for source code analysis and manipulation. The link between view and its controller is divided into three parts (Figure 1)

- The *fx:controller* attribute is used to associate the view and controller. It is set as the scene graph's root node. The scene graph for face detection application is shown in figure 2.
- The *fx:id* attributes are declared as component instance in the view from field declaration in the controller. Three *fx:id* are used: btnDetectFace, preProcessImage, originalImage
- *Event handler* attributes is assigned from various component events in the controller. onAction( ) handle the process from source image to detect the face/faces.

Atributes in FXML represent a property of each class instance which it has static property and event handler.

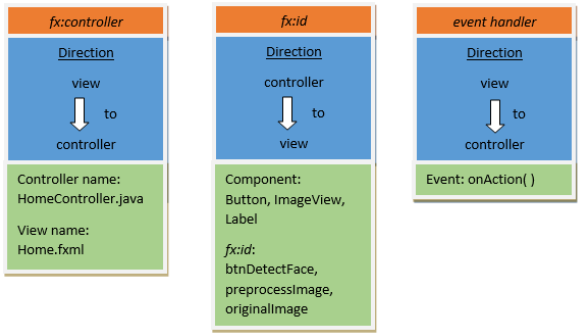


Fig. 1. The Requirements Identification for Face Detection Application

B. Design UI Component in Scene Graph

To build this application, the hierarchical scene graph (Figure 2) should first show the BorderPane layout as a root node, and then it is followed by HBox layout and VBox layout in the next level (child node). Each child node contains UI components such as: ImageView, Label, and Button (leaf node). The implementation of this scene graph can be seen in Figure 5.

BorderPane lays out the children in top, left, right, bottom and center position. The center node in BorderPane is filled the HBox layout which position all its child nodes in horizontal row. VBox is the child nodes of HBox. Each of VBox contains two child nodes: ImageView and Label in first of VBox, two Buttons in the second VBox, and ImageView and Label for the last VBox.

The ImageView component is a node used for displaying original image and result image after face detected. The Button component is a node which can respond to mouse event by implementing an event handler to process the mouse event. The view for this JavaFX face detection is shown in figure 3.

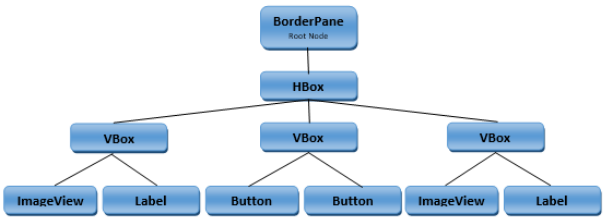


Fig. 2. Scene Graph for UI Components

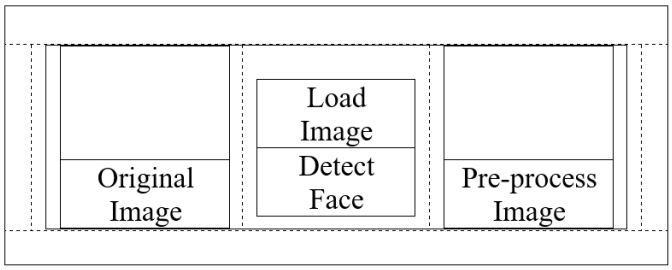


Fig. 3. The View of Application in Scene Builder

C. Integration of the scene builder panels (FXML & View)

Scene builder's library, document and content panels are extracted to Netbeans view in FXML document. Library panel display a collection of the GUI components that are used to build the view. Document panel is comprised of two sub-panels: the Hierarchy and the Controller panel. The Hierarchy panel display the view components in a tree structure. The Controller panel manage the view to its controller, which has an input field for the controller's name, as well as an overview of all the *fx:id* declaration in the view. The content panel display a static preview of view's content will look like when rendered in a running application. The FXML document loads component instances declared in the view into the controller's field declaration.

D. The Controller

Scene builder component implements the connection with controller. The controller input field and event handler (in *fx:id*'s) are declared with the notation @FXML in Java class. This notation integrates scene builder's component to be recognized by its controller. The controller enables field and event handler access through FXML loader by importing javafx.fxml.FXML. The FXML loader is responsible for loading the FXML source file and returning the value of each component in the scene graph.

The input fields such as: origrinalImage and preprocessImage attibutes are declared by importing package javafx.scene.image.ImageView, while btnDetectImage and btnPreprocess attributes import package javafx.scene.control.Button. Event handler utilized to process images to be able to detect faces is onAction( ) to call method loadImage( ) dan detectAndDisplay( ). The method of loadImage( ) requests user to select image file to be processed, and then the method of preProcess( ) will process the images that have been input to be grayscale image by using the library open source version 3.4 from OpenCV. The method of detectAndDisplay( ) will process images that have been input for faces to be able to be detected by using the feature of Haar Cascade Classifier. This feature is obtained from the open source library of OpenCV by importing the package org.opencv.objdetect.CascadeClassifier.

III. RESULT AND DISCUSSION

In this research, JavaFX application to detect faces in images has been designed and implemented by using Scene Builder and Netbeans IDE (shown in Figure 4). The preliminary stage of designing this application is by determining its requirements such as *fx:controller*, *fx:id*, and *event handler*. What comes next is that every single GUI component needed is designed into a scene graph to view the component structure used in the display. The integration of the display in scene builder uses file FXML which is recognized by controller so that every component can be declared and processed in the programming syntax. In the controller, various libraries required have been imported from JavaFX and library OpenCV.

Fig. 4. JavaFX Face Detection Application

The controller contains input field and event handler using notation @FXML, which integrate the component to be recognized. It shown in Figure 5.

```
public class HomeController {

    @FXML
    private ImageView originalImage, preprocessImage;

    @FXML
    private Button btnDetectImage, btnPreprocess;

    protected void init( ) {…}

    @FXML
    protected void loadImage( ) throws IOException {…}

    @FXML
    protected void detectAndDisplay( ) {
        String xmlFile = "D:/haarcascade_frontalface_alt2.xml";
        CascadeClassifier classifier = new CascadeClassifier(xmlFile);

        MatOfRect faces = new MatOfRect( );
        classifier.detectMultiScale(imageGray, faces);
        JOptionPane.showMessageDialog(null, String.format("Detected %s faces", faces.toArray( ).length));

        for (Rect : faces.toArray( )) {
            Imgproc.rectangle(imageGray, new Point(rect.x, rect.y), new Point(rect.x + rect.width, rect.y +
            rect.height), new Scalar(0, 0, 255), 3);}
            preProcess( );
            }

    @FXML
    protected void preProcess( ) {…}

    public void setStage(Stage stage) {…}

    private void updateImageView(ImageView view, Image image) {…}
}
```

Fig. 5. HomeController.java

While controllers can be easy to write event handlers in script, it is preferable to define complex application logic in a compiled using Java language. The *fx:controller* attribute allows a caller to link a "controller" class with an FXML document. It is a compiled class that implements the "code behind".

The view from scene builder is generated to Netbeans in FXML code (Figure 6 and Figure 7). In FXML, all classes are imported such as java.lang package to following processing import the VBox, ImageView and Button classes. The root node assigns the *fx:controller* to connect the view and the controller. The HBox layout as the child node on the level 1 is used to arrange the other series of nodes in a single row, and the VBox layout as the child node on the level 2 is used to arrange the other nodes in single column. The VBox layout contains UI components with their *fx:id*.

Each property for an object in FXML is being set. Two ImageView components have *fx:id* originalImage and fx:id preprocessImage. The *fx:id* original image is then assigned to contain the original image from the file chooser in which *fx:id* preprocessImage is assigned to show the preprocessed image. One button component has *fx:id* btnDetectFace which is declared to process the image and detect the face(s). In document panel, the controller is setting up the value of controller class: HomeController.java.

```
<?xml version="1.0" encoding="UTF-8"?>

<?import javafx.geometry.*?>
<?import javafx.scene.control.*?>
<?import javafx.scene.image.*?>
<?import java.lang.*?>
<?import javafx.scene.layout.*?>

<BorderPane maxHeight="-Infinity"
            maxWidth="-Infinity"
            minHeight="-Infinity"
            minWidth="-Infinity"
            xmlns="http://javafx.com/javafx/8" xmlns:fx="http://javafx.com/fxml/1"
            fx:controller="projectfriver.HomeController">
  <center>
    <HBox alignment="CENTER" spacing="20.0" BorderPane.alignment="CENTER">
      <children>
        <VBox alignment="CENTER" spacing="20.0">
          <children>
            <ImageView fx:id="originalImage" fitHeight="150.0" fitWidth="200.0" pickOnBounds="true"
preserveRatio="true">
              <VBox.margin>
                <Insets left="30.0" right="30.0" />
              </VBox.margin></ImageView>
            <Label text="Original Image" />
          </children>
          <HBox.margin>
            <Insets />
          </HBox.margin>
        </VBox>
        <VBox alignment="CENTER" spacing="20.0">
          <children>
            <Button mnemonicParsing="false" onAction="#loadImage" text="Load Image" />
            <Button fx:id="btnDetectFace" mnemonicParsing="false" onAction="#detectAndDisplay"
text="Detect Face" />
          </children>
          <HBox.margin>
            <Insets />
          </HBox.margin>
        </VBox>
        <VBox alignment="CENTER" spacing="20.0">
          <children>
            <ImageView fx:id="preprocessImage" fitHeight="150.0" fitWidth="200.0" pickOnBounds="true"
preserveRatio="true" />
            <Label text="Pre-processed Image" />
          </children>
          <HBox.margin>
            <Insets left="30.0" right="30.0" />
          </HBox.margin>
        </VBox>
      </children>
      <BorderPane.margin>
        <Insets bottom="70.0" left="30.0" right="30.0" />
      </BorderPane.margin></HBox>
  </center>
</BorderPane>
```

Fig. 6. Home.FXML

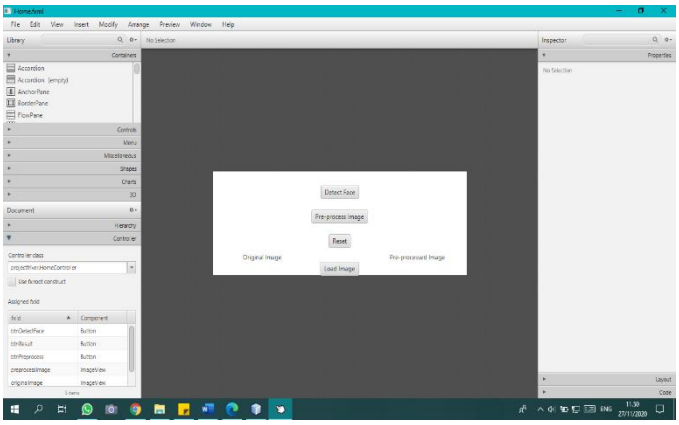


Fig. 7. Home.FXML in Scene Builder

IV. CONCLUSION

In the stage of UI component integration in the display on the controller, there are several things that have to be fulfilled,

such as: assigning fx:controller, fx:id and event handler in UI component must be performed in document panel in Scene Builder so that the display can be recognized by the controller; the use of @FXML notation must be added to the attribute declaration and method to process the input; the FXML document loader should be assigned with the controller's name. JavaFX face detection is an application to detect face using image as an input. It is develop and implement based on the stages: requirements identification, design of UI components in Scene Graph which simplify developer to design based on hierarchy of components, integration of the scene builder panels, and controller.

REFERENCES

[1] O. Corporation, "Introduction to FXML," 2012. [Online]. Available: https://docs.oracle.com/javafx/2/api/javafx/fxml/doc-files/introduction_to_fxml.html [Accessed: 20-Mar-2020].

[2] O. Corporation, "JavaFX: Getting Started with JavaFX," 2014. [Online]. Available: https://docs.oracle.com/javase/8/javafx/get-started-tutorial/index.html [Accessed: 20-Mar-2020].

[3] G. Kruk, O. Alves, L. Molinari, and E. Roux, "Best practices for efficient development of JavaFX applications," JACoWPublishing, pp. 1078–1083, 2018, doi: 10.18429/JACoW-ICALEPCS2017-THAPL02.

[4] I. Science, "Comparison and implementation of graph visualization algorithms using JavaFX," 2016.

[5] A. Pomaroli, JavaFX Programming CookBook. Hot Recipes for JavaFX Development. Exelixis Media, 2016.

[6] H. Qluon, "Scene Builder Documentation," 2016. [Online]. Available: https://docs.gluonhq.com/scenebuilder/ [Accessed: 20-Mar-2020].

[7] T. Ask, "Tobias Ask Master ' s thesis sb4e : an open source integration of the Scene Builder GUI editor into the Eclipse IDE Facilitating JavaFX application development with an extension to the IDE," no. June, 2019.

[8] A. Alkhars and W. Mahmoud, "Cross-Platform Desktop Development (JavaFX vs. Electron)," 2017.

Face Recognition for an Attendance System

1st Daniel Sutopo Pamungkas
Mechatronics Departement
Politeknik Negeri Batam
Indonesia
daniel@polibatam.ac.id

2nd Frederik Mahardika
Mechatronics Departement
Politeknik Negeri Batam
Indonesia
frederickmahardika@gmail.com

***Abstract*—In the pandemic era, the attendance system needs to be contactless. Moreover, this system needs to be more automatic compare to the existing systems. This paper introduces the visual system using face recognition. The Haar cascade method and Local Binary pattern histogram algorithm is used to recognize the user's identity. To capture the face, a webcam is used. This system enables detection and identifies the users. It also stores the time that check-in and checks out of the users automatically. The proposed system adequate to detect face up to 55 cm in the low lighting condition. Furthermore, this system enables to detection of multiple users in one frame.**

Keywords—attendance system, Haar-like feature, face recognition, LBP

I. INTRODUCTION

To record the attendance using the manual method is not a simple activity also time consumed. To overcome these problems, an automatic system is needed. Moreover, Covid-19 change many habits. One of them is to touch something because this virus could pass to the human through touching. In the previous attendances' methods, there are still need contact to the device, for example, using fingerprint [1], or RFID card [2]

Face recognition has the potential to be used for automatic and contactless attendance systems. There are many methods to detect face and recognize the face.

In [3], the authors propose a face recognition-based attendance system. They utilize the Viola-Jones feature and Histogram of Oriented Gradients (HOG) algorithms, combined with the classifier. The Support Vector Machine (SVM) is used for the classification for their system. Their experiments varying several scenarios for example change the distances between subject with the system, illumination of the environment, are there any obstacles also several poses. These conditions are considered. They used Signal to Noise Ratio (SNR) to analyze the results using MATLAB. Another research in [4], the authors introduce biometric recognition algorithm which are the Eigenface and Fisher face method. They applied their algorithm in Open CV 2.4.8. They compare the Receiver Operating Characteristics (ROC) curve. Moreover, they were implementing these algorithms on the attendance system. They shared that the ROC curve proves that the Eigenface achieves better results than the Fisher's face. the tactic implemented using the Eigenface algorithm achieves with an rate of accuracy between 70% to 90%

Other researchers used a technique which can be apply in an attendance system. This system is applying in a class. They utilized the face recognition technique. They combine the Discrete Wavelet Transforms (DWT) method with a Discrete Cosine Transform (DCT) algorithm [5]. These combined method were accustomed to extract the features of the subject's face. Afterward, the authors classify the subjects applying the Radial Basis Function (RBF). This authors reports that these method has an accuracy rate around 82%. In [6], the authors employ the Haar Like to detect the face with Local Binary Pattern Histogram (LBPH) methods for identify the subjects. This method enables to detection and recognizes the user well [6]. For this study, for the attendee system the Haar and LBPH methods is used.

To present a comprehensive information, this manuscript is arranged as follows: the objectives of the next part is to produce the proposed system to acknowledge the user also therefore the database system. Then continued with the third section. This part, which provides the tests on the proposed algorithm by describing the system's potentials in enabling to spot the face and identification the subject. In the final section presents the concluding remarks also the future work of this study.

II. METHOD

Figure 1 shows the configuration of the hardware of the system. A High Definition Logitech Webcam with type C310 is used. This device is located on the wall. The camera Function is to capture the condition on the front of the wall. The signal from the webcam is proceed to the PC. This information is processed by the computer. When there is/are humans, then the computer will locate the face(s). If the face(s) has already been in the database, then the system will recognize and put the time stamp in the database. Those processes are in the computer utilizing the Haar Cascade algorithm to detect the face then the Local Binary Pattern Histogram method to recognize the user.

Figure 2 shows the flowchart of the system. This system needs a database which is containing the data of the users. Faces of the users are put in the database. If the user is detected first time that day, the system will write to the database that the user is signed in. If the user has been seen yet, the system will report to the database that the user is sign out.

Webcam has transferred the information to the PC. Then the data is converted to a grayscale image. The Haar Cascade methods are used to detect the face of the user. If the face of the subject is located, then this proposed algorithm will tag the face of the subject. To focus the subject face from other images than the face, a green box is used in this system. The math function of the Haar method is the same as the Fourier function. [7-8].

The next step is a cascade classifier process. This process will obtain more accurate results. This process is to calculate the Haar feature in the repetitive process. If the result did not meet the standard, then rejected. While if the results meet the criteria, then proceed to the next step. This process is repeated three times [9-10].

The next step after the face is obtained to recognize the face of the user. The LBPH algorithm is used to do this task. The trained data will be matched with the result of the previous process. LBPH will extract the value of the image histogram. The method of this algorithm is to confront the worth of the binary pixel in central image with values of eight binary pixels around the central image. When the value of the result is more than or equal to null, then the value is assigned to one; other than that, the value is null. Afterward, the values in binary form are set clockwise or vice versa. After that the binary values form is changed into form of decimal. This process is to substitute the value of the pixel of the image center [3].

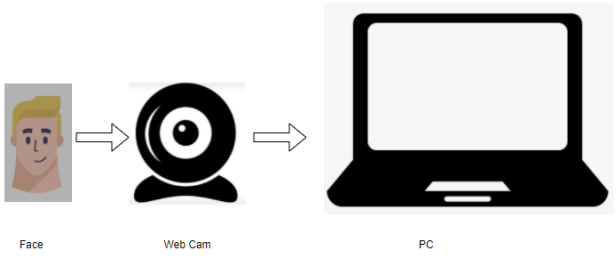


Fig. 1. The system's block diagram

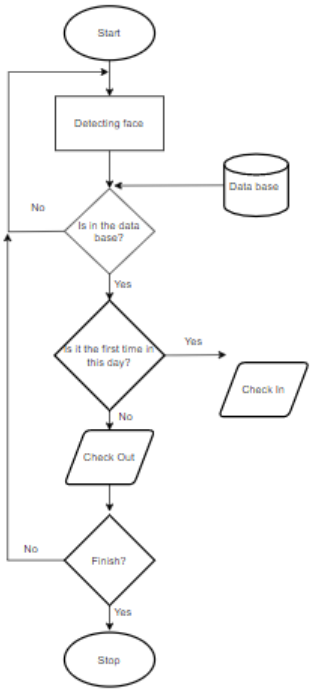


Fig. 2. Software system flowchart

Thereafter arranging binarization clockwise process, then proceed to the next action. When a binary threshold box is worth one enters the binary value in keeping with the rank. In the other hand, when the worth is null, then the output is also null. Finally, the LBP values is added. To comparing the owner's face an equation is used to urge the approach histogram value. This value is used as a prediction value to spot the owner of the face.

This value is employed as a comparison between data of the face contained within the database with the data which detected by the webcam. The database of the faces needs to be stored first. This database will enable for system to identify the owner of the faces. These faces will be converted to histogram values or LBPH. To be able to recognize the face well, the system needs a minimum of 20 input images from the webcam. These images must be trained first. During the training process, the images on the database will be extracted. These images will change to histogram values. Each image will have their own histogram value These values will be stored in form of array data. These values will be together with the positive identification of every subject. On the face detection process, the webcam will obtain the face image input detected is going to be known histogram value. Therefore, the method enable to compare the input image histogram value which come from the webcam with the database of the histogram faces values. This process will enable to identify of the face of the owner by the histogram value which closest to the histogram value within the database.

To the popularity face process in real-time, the LBPH algorithm is applied. During this step, the worth of histogram value within the database face is going to be compared with the number of histogram values of the image detected directly by the camera (real-time). To induce an identical image with the worth the database has stored, it's necessary to match two histograms between the detected image and, therefore, the image within the database and find the space of the closest histogram value. Thus, the output of the algorithm is that the positive identification of every image that's changed with the name of the face owner.

III. Results and Discussion

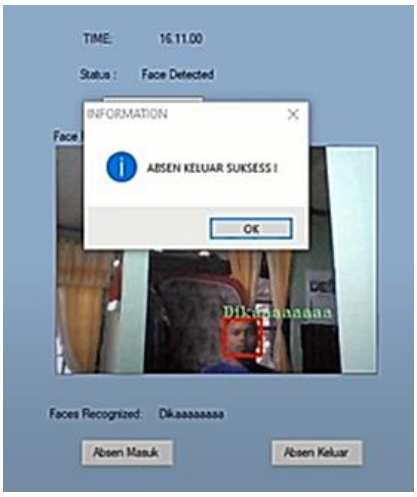
To test the ability of the system, five subjects are used as a sample of the users. These are consisting of three men and two women. The age is between 22-25 years old. The system is used in the different distances and ambient brightness. The results of these experiments are presented in Table 1. In the bright environment, the system enables to recognize of the face, and the user is also successful in signing in or sign out in the 55 cm to 180 cm range. While in the dim situation, the system is also capable of doing its task with a narrow distance. However, the distance is only around 50 cm from the camera. If the system is able to detect the face, then the system will be capable of doing the absent to the subject.

Figure 3a shows bright lighting while the subject is standing around 180 cm from the camera. In this figure, the system enables the detection of the face of the subject. Moreover, this system is capable to sign out the subject where the subject has sign in before. Figure 3 b gives an example capturing the subject in the dim situation at a close distance. A Subject is at around 50 cm around the camera. This system is adequate to sign in the subject to the system.

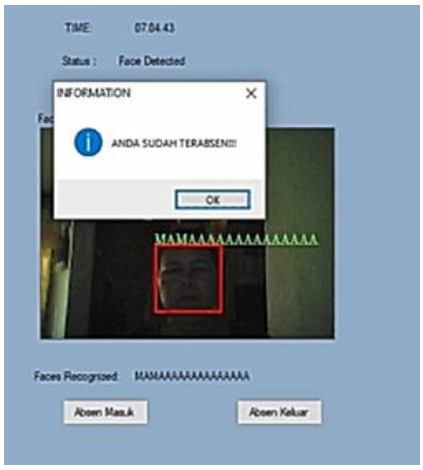
TABLE I. Results

Distance	Brightness		Sign In/Out
	Light	*Dim*	
30 cm	Not detected	Not detected	not successful
55 cm	detected	detected	successful
85 cm	detected	Not detected	not successful

180 cm	detected	Not detected	not successful
200 cm	Not detected	Not detected	not successful



(a)



(b)

Fig. 3. The results of the system (a) in bright 180 cm distance and (b) in dim 55 cm distance

IV. CONCLUSION

This study is applying the Haar cascade and LBPB methods in the absence system. Based on the results of the research data above, it can be concluded that the lighting factor significantly affects the accuracy of face recognition and detection; the more significant the brightness level of a place, the easier the system will be to recognize the object to be identified. The indicator of the success of this final project is to be able to make a face attendance application without touching it with a retrieval time of not less than 3 seconds, as well as with various lighting and the success rate of taking face attendance is with a maximum distance of 180 cm and a minimum distance of 55 cm. This system is very well used in recent times. In the current corona pandemic, this system will minimize the touch of a hand on a device.

ACKNOWLEDGMENT

This research was fully funded by Politeknik Negeri Batam.

REFERENCES

[1] S. Hapani, et al. "Automated Attendance System Using Image Processing." 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA). IEEE, 2018.

[2] S. Akbar, et al. "Face Recognition and RFID Verified Attendance System." 2018 International Conference on Computing,Electronics & Communications Engineering (iCCECE). IEEE, 2018.

[3] H. Rathod, et al. "Automated attendance system using machine learning approach." 2017 International Conference on Nascent Technologies in Engineering (ICNTE). IEEE, 2017

[4] A.R. Siswanto, A.S. Septian, and G. Maulahikmah "Implementation of face recognition algorithm for biometrics based time attendance system." 2014 International Conference on ICT For Smart Society (ICISS). 2014.

[5] K. Okokpujie, et al. "Design and implementation of a student attendance system using iris biometric recognition." 2017 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, 2017

[6] D. S. Pamungkas and S. Al-Aidid, "Detection, Recognition, and Tracking Face Using 2 DoF Robot with Haar LBP Histogram," *2018 International Conference on Applied Engineering (ICAE)*, 2018, pp. 1-5, doi: 10.1109/INCAE.2018.8579409.

[7] Hengliang Tang, Yanfeng Sun, Baocai Yin and Yun Ge, "Face recognition based on Haar LBP histogram," 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), Chengdu, 2010, pp. V6-235-V6-238.

[8] Rahim, Abdur "Face Recognition using Local Binary Patterns (LBP)." Global Journal of Computer Science and Technology Graphics & Vision, Pabna University of Science and Technology Bangladesh, 2013.

[9] M. Da'san, A. Alqudah and O. Debeir, "Face detection using Viola and Jones method and neural networks," 2015 International Conference on Information and Communication Technology Research (ICTRC), Abu Dhabi, 2015, pp. 40-43.

[10] Q. Li et al., "Multi-view face detector using a single cascade classifier," 2016 10th International Conference on Software, Knowledge, Information Management & Applications (SKIMA), Chengdu, 2016, pp.464-468

Mining Student's Reviews to Obtain Their Perception toward College Department Performance

1st Yuliska
Informatics Engineering
Politeknik Caltex Riau
Pekanbaru, Indonesia
yuliska@pcr.ac.id

2nd Dini Hidayatul Qudsi
Information System
Politeknik Caltex Riau
Pekanbaru, Indonesia
dinihq@pcr.ac.id

3rd Lia Anggraini
Electrical Engineering and Informatics
Bandung Institute of Technology
Bandung, Indonesia
23520037@std.stei.itb.ac.id

4th Juanda Hakim Lubis
Informatics Engineering
Universitas Medan Area
Medan, Indonesia
juandahakim@gmail.com

Abstract*—Student's perception toward department performance can be crucial, therefore, it can be used to evaluate the department outcome and take an immediate action to improve its management. This study applies sentiment analysis and topic modeling to the student's reviews of college department at Politeknik Caltex Riau in order to mine student's perception for seven college departments performance. Sentiment analysis with Support Vector Machine (SVM) is employed to obtain student's sentiment. There are 3 types of sentiments to be analyzed; positive, negative and neutral. Topic modeling with Latent Dirichlet Allocation (LDA) is also carried out to get some important keywords in the student's reviews. Our experiments show that Positive is the most prominent sentiment in the student's reviews while LDA reveals some important topics toward preferences.

Keywords—Student's Review, Student's Perception, Department Performance, Sentiment Analysis, Topic Modeling

I. INTRODUCTION

As technology develops, more and more users use the internet as a source of information. The role of the internet as a source of information makes users use the internet to share their personal opinions in the form of reviews. The review can be in the form of a user's personal opinion about the quality of a movie, products, hotel services and services of an institution. Many websites provide a questionnaire feature so that users can express their opinions. The results of the review will then be used for better decision making [1].

In this study, we utilize student's reviews on the performance of seven college departments at Politeknik Caltex Riau (PCR) to mine student's perception toward those departments. The student's reviews on the performance of each department can reach 1500 reviews; with the number of words can be up to 150 of each review. Therefore, mining these many and long reviews becomes essential because it allows the department to improve their performance based on the mining result.

The mining process is performed using sentiment analysis and topic modeling. We carry out Sentiment analysis with Support Vector Machine (SVM) [2] to get 3 types of student's sentiments in the review, namely, positive, negative and neutral. The Latent Dirichlet Allocation (LDA) [3] is then used to reveals some important keywords hidden on the student's review.

II. LITERATURE REVIEW

A. Sentiment Analysis

Sentiment analysis is a process that analyzes and detects the sentiment of a text input that has positive, negative or neutral sentiments [4]. However, until now, the sentiments that can be detected have become more diverse and detailed and are not limited to only positive and negative, but also happiness, sadness, anger, fear, disgust and surprised [5].

Sentiment analysis can be used to monitor the performance of a product or institution's services. By applying sentiment analysis, product developers and service owners can easily find out whether a product or service is received positively by customers or vice versa.

B. Support Vector Machine (SVM)

Support Vector Machine is one of the powerful machine learning techniques for data and text classification. It is also a supervised method, so we need labeled dataset to train a text classifier which based on Support Vector Machine.

Suppose we have some data points which each belongs to one of 2 classes, and the goal is to predict which class those points are belong. Support Vector Machine works by building some hyperplanes between those points that might classify the data points. An optimal hyperplane is the one with the largest margin or separation between the two classes.

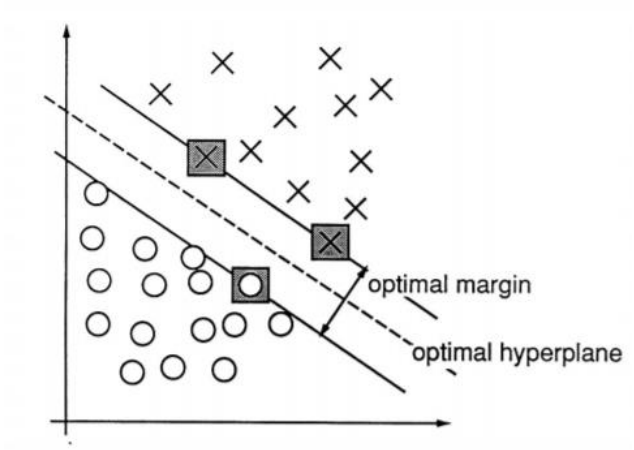


Fig. 1. An Optimal Hypeplane [2]

C. Latent Dirichlet Allocation (LDA)

Latent Dirichlet Allocation is a method which is largely used for topic modeling. It is a technique used to find some topics or hidden keywords on a collection of text document. Suppose we have following documents which contains some words:

Document1: $word_1$, $word_2$, $word_3$, …, $word_n$
Document2: $word_1$, $word_2$, $word_3$, …, $word_n$
Document3: $word_1$, $word_2$, $word_3$, …, $word_n$

The words on each document are ordered by the frequency of its occurrence. With Latent Dirichlet Allocation, we can figure out the words which belong to different topics, as can be seen in Table I below:

TABLE I. WORDS WITH PROBABILITY SCORE ON DIFFERENT TOPICS

Topic	$Word_1$	$Word_2$	$Word_3$	$Word_n$
Topic1	0.23	0.001	0.019	0.007
Topic2	0.015	0.49	0.02	0.011
Topic3	0.010	0.40	0.021	0.001

III. EXPERIMENT

In this section, we describe some stages we conducted during the experiment, the datasets, and our experimental settings.

A. Dataset

We used Quality Assurance (QA) dataset of the department performance at Politeknik Caltex Riau (PCR) in the second semester of the Academic Year of 2018/2019. It is generated from the PCR Planning, Development and Quality Assurance Department, called BP3M. Elaborately, the data contained of reviews from students on the performance of seven college departments at PCR. The seven departments consist of Academic and Student Administration (BAAK), Information System and Technology (BSTI), Student department, Cooperation, Finance, Library, and Infrastructure department. The number of reviews in each department is from 1470 to 1500, but only 1200 reviews will be used, while the remaining 300 reviews are eliminated based on the number of words contained in it. The number of words in the reviews ranges from 2 words to 150 words. Table II shows an example of reviews in one of the departments.

TABLE II. AN EXAMPLE OF REVIEWS OF DEPARTMENTS

Number	Reviews
1	Pelayanan Sangat baik dan diberikan arahan dengan baik dan ramah ☺☺
2	Terapkan pembayaran ansuran agar mahasiswa yang tidak mampu dapat kemudahan
3	Jangan sampai antrian panjang menanti
4	Pelayanan yang d berikan sudah baik, penjaga nya juga ramah, dalam menindaklanjuti pembludakan pembayaran serentak yang dilakukan mahasiswa harus d perbaiki.
5	Saran dari saya supaya tempat pembayaran SPP di tambah karna menurut saya kalau 1 bank masi kurang cukup, sebab admin bank akan kerepotan mengurus mahasiswa yang mengurus SPP tersebut
6	Bagian keuangan sudah memberikan pelayanan dan tanggapan yang memuaskan

B. Automatic Labelling

To be able to use the data explained in sub section A to train SVM model, we employed an automatic labeling with a tool called *SentiStrength* [6]. *SentiStrength* gives label to the review based on the sentiment score. For instance, if the total scores of words which have positive sentiment greater than the score of words with negative and neutral sentiment, then the review is labeled as positive review. This scenario is also applied to negative and neutral label.

In this study, we take 200 reviews from each department for training purpose. Therefore, there are 1400 reviews used for training. Subsequently, we perform automatic labeling to give sentiment label on each review.

C. Text Preprocessing

This study uses *NLTK (Natural Language ToolKit)* as a library for word processing. The preprocessing stages applied in this study are as follows:

1) Case Folding: The process of converting the uppercase to lowercase.

TABLE III. CASE FOLDING

Review	*Case Folding*
Pelayanan Sangat baik dan diberikan arahan dengan baik dan ramah ☺☺	pelayanan sangat baik dan diberikan arahan dengan baik dan ramah ☺☺

2) Stop word Removal: The process of eliminating the unnecessary words.

TABLE IV. STOPWORD REMOVAL

Review	*Stopword Removal*
pelayanan sangat baik dan diberikan arahan dengan baik dan ramah ☺☺	pelayanan sangat baik diberikan arahan baik ramah ☺☺

3) Emoticon Removal: The process of removing the emoticons in the text.

TABLE VI. EMOTICON REMOVAL

Review	*Emoticon Removal*
pelayanan sangat baik dan diberikan arahan dengan baik dan ramah ☺☺	pelayanan sangat baik diberikan arahan baik ramah

4) Tokenization: The process of tokenizing sentence into a collection of words.

TABLE VII. TOKENIZATION

Review	*Tokenization*
pelayanan sangat baik diberikan arahan baik ramah	[“pelayanan”, “sangat”, “baik”, “diberikan”, “arahan”, “baik”, “ramah”]

D. *Experimental Setting*

We use a library called *Scikit-Learn* to train SVM model, with the kernel is set to ‘linier’, regularization parameter is set to 1.0, and the remaining parameters are set to default. SVM model is trained using the data explained in sub section B, and the model will be employed to predict the sentiment of 1000 reviews on each department. In the next stage, to perform Latent Dirichlet Allocation, we still utilize *Scikit-Learn* Library with all parameters is set to default. For word representation, we use TF-IDF representation. In addition, because data training is imbalance, we apply oversampling to balance it.

IV. RESULT & DISCUSSION

This section reports on our experimental results and discusses the reason behind those results.

A. *Sentiment Analysis Result*

As we explained in section III, we used 1400 labeled review for training purpose. During training, the SVM model obtained the accuracy of 96%. However, we have to mention that although the accuracy of the model is good, we found that the model do not perform effectively in predicting negative sentiment. We then use this model to predict the 1000 reviews on each department. The results are shown below:

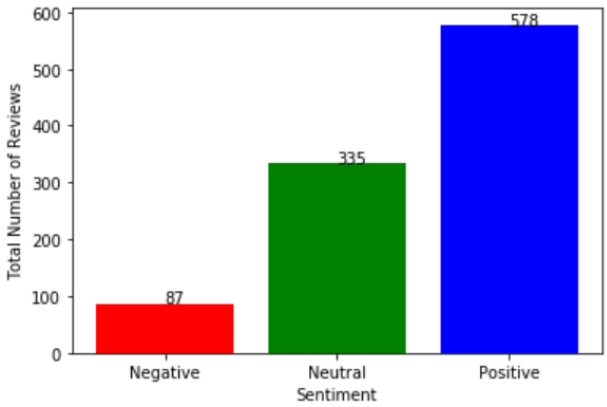


Fig. 2. Sentiment Analysis of BAAK

Figure 1 shows the review sentiments for Academic and Student Administration (BAAK) department. It shows that reviews with positive sentiment are the most prominent. There are 578 positive reviews, 335 neutral reviews, and 87 negative reviews.

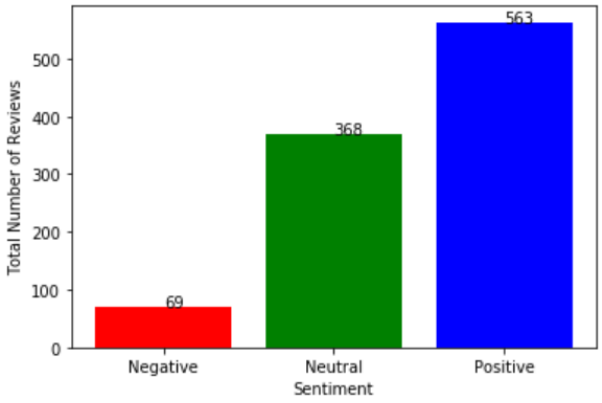


Fig. 3. Sentiment Analysis of BSTI

The review sentiments for Information System and Technology (BSTI) department can be seen in Figure 2. It shows that positive sentiments are also the most notable review. In this department, there are 563 positive reviews, 368 neutral reviews, and 69 negative reviews. It can be seen that Information System and Technology (BSTI) department has less positive and negative reviews than Academic and Student Administration (BAAK) department.

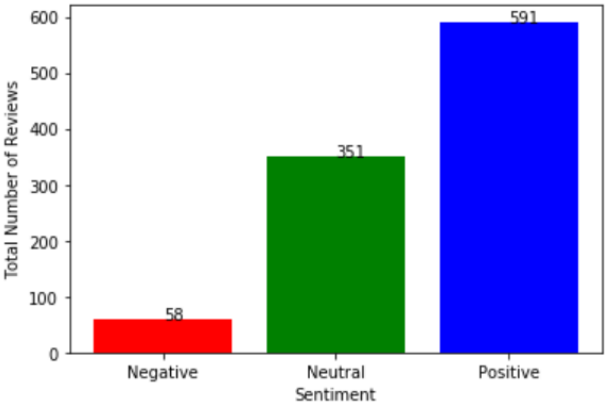


Fig. 4. Sentiment Analysis of Student Department

Moreover, the review sentiments of Student Department have the same trend with Academic and Student Administration (BAAK) and Information System and Technology (BSTI) departments which the positive sentiment is the most significant, and the negative reviews have the least total number. This department has 591 positive reviews, 351 neutral reviews, and 58 negative reviews. It can be observed that this department has more positive reviews and less negative reviews than the previous two departments.

After that, we analyze the sentiment analysis result of Finance department. As can be seen in Figure 4, the Finance Department also has the same trend with previous departments. In this department, reviews with positive sentiment have the highest number and reviews with negative sentiment come with the least number. There 601 positive reviews, 347 neutral reviews and 52 negative reviews.

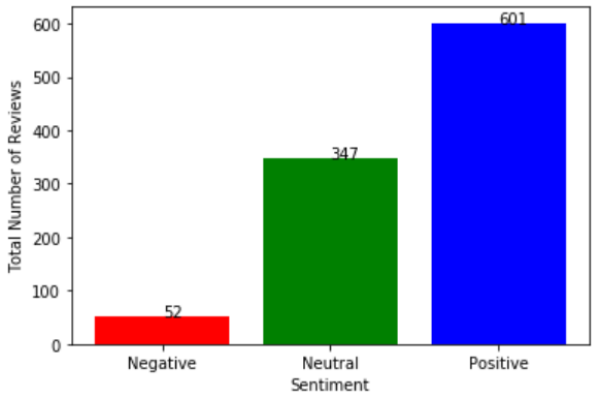


Fig. 5. Sentiment Analysis of Finance Department

Figure 5 shows the review sentiments for Cooperation Department. Expectedly, this department has the same trend with the sentiment analysis of previous departments. In this department, the positive sentiments also become the most striking review with the total number of 600. Additionally, with 41 negative reviews, this department has the least number of negative sentiments than the previous departments. The neutral sentiment comes with the number of 359 reviews.

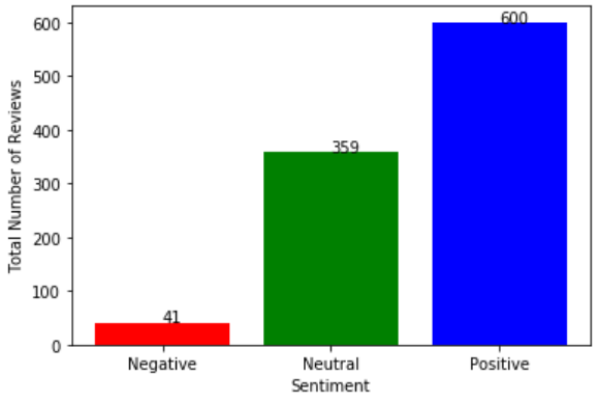


Fig. 6. Sentiment Analysis of Cooperation Department

Figure 6 and 7 display the review sentiments for Infrastructure Department and Library, respectively. These departments have the same trend that the positive reviews have the highest number. It is followed by neutral reviews in the second high number and negative reviews with the least number. As can be seen in Figure 6, there are 90 negative reviews in Infrastructure department, which is the highest among 7 departments.

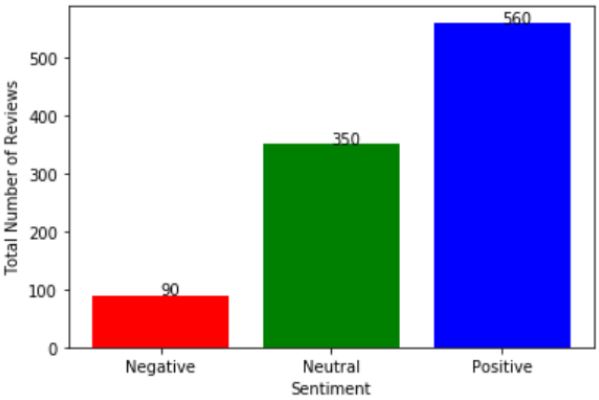


Fig. 7. Sentiment Analysis of Infrastructure Department

Furthermore, we analyze the reviews which have positive sentiments in 7 departments, and we found that all departments have similar reviews for positive sentiments. For positive sentiment, “good service” is the common topic.

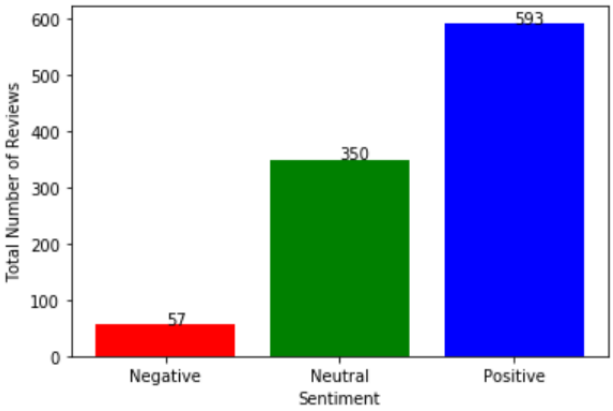


Fig. 8. Sentiment Analysis of Library

B. *Topic Modelling Result*

We perform topic modeling on the positive reviews with Latent Dirichlet Allocation (LDA) with the number of topics is set to 10 topics. Figure 8 shows the information about topic word weight or words with probability score in 10 different topics.

	akademik	alhamdulillah	baak	bagus	belajar	berani
Topic0	0.100007	21.400820	0.100022	0.100086	0.100017	0.100009
Topic1	0.100019	0.100011	0.100020	0.100024	0.100016	0.100010
Topic2	0.110206	0.100021	0.147426	506.581892	0.100012	0.100010
Topic3	0.100009	0.100009	0.100018	0.100020	0.100012	0.100010
Topic4	0.100026	0.100008	62.656136	0.100035	0.100013	0.100008
Topic5	0.100013	0.100008	0.101251	0.100024	0.100015	0.100008
Topic6	0.100008	0.100009	0.100083	0.100025	0.100019	0.100008
Topic7	25.028356	0.100010	35.942779	0.100014	0.100033	0.100008
Topic8	0.100012	0.100009	0.100016	0.100010	0.100017	0.100008
Topic9	0.100015	0.100009	0.100013	0.100142	14.953149	11.360504

Fig. 9. Topic Word Weight

The top 10 topic keywords in 10 different topics can be observed in Table VII.

TABLE VIII. TOP 10 TOPIC KEYWORDS

Topic	***Keywords***
Topic 1	Mantap, semangat, ruangan, alhamdulillah, bersih, lumayan, oke, konsisten, layanannya, puas
Topic 2	Terimakasih, sarana, sukses, prasana, ditingkatkan, membantu, cepat, diharapkan, maju, keramahan
Topic 3	Kedepannya, bagus, semoga, layanan, memuaskan, puas, kemahasiswaan, pelayanan, baak, sistem
Topic 4	Terimakasih, pelayanannya, maju, kinerja, semoga, mohon, kedepannya, keuangan, sistem, tolong
Topic 5	Pelayanan, baak, sistem, tingkatkan, informasi, industri, senyum, mudah, dipertahankan, teknologi
Topic 6	Keuangan, saran, terbaik, tolong, fasilitas, layanan, suka, pelayanan, mahasiswa, semoga
Topic 7	Perpustakaan, kerjasama, rapi, kedepannya, buku, keren, sopan, melayani, industri, terimakasih

Topic 8	Mahasiswa, ramah, baak, melayani, proses, akademik, pengelola, mohon, pembelajaran, sesuai
Topic 9	Semoga, nyaman, kampus, berkembang, depannya, senang, penilaian, pcr, layanan, perpusatakaan
Topic 10	Pcr, kedepan, pertahankan, lanjutkan, kerja, kembangkan, belajar, ruang, berani, meningkatkan

The common keywords on the 10 topics are "puas", "bagus", "ramah", and "terimakasih". The first, the second and the last topics are mainly related to the service of infrastructure department. The third is notably focused on the service of Student Department. Furthermore, the fourth and the sixth topic are mainly related to the Service of Finance Department. The fifth and the eight topics consist of the keywords related to the service of Academic and Student Administration (BAAK) department. Lastly, the seventh and the ninth topics are discussed about the service of Library. As can be seen in Table VII, all keywords on each department show that the students were satisfied with the service of all departments. From the keywords, in addition, we can also conclude that the students hoped for the department to improve their performance and service in the future.

V. CONCLUSIONS & FUTURE WORK

Our experiment implemented sentiment analysis with Support Vector Machine and Topic Modeling with Latent Dirichlet Allocation to mine student's perception toward performance of college departments. Our findings are summarized as follow:

- Sentiment analysis can be utilized to discover the sentiment expressed by students through their reviews on the performance of college departments. Based on the sentiment analysis results, it can be concluded that the Finance and Cooperation departments have the highest number positive reviews and has the least number of negative reviews, respectively.
- Topic modeling can be used to identify some hidden keywords associated to some topics. With Topic modeling, we can find the characteristic of each department through the keywords displayed on each topic.

We perform sentiment analysis with Support Vector Machine (SVM). The model has great accuracy, but has poor performance in predicting negative sentiments. Thus, utilizing another machine or deep learning model might result in better and more accurate performance.

REFERENCES

[1] Raut, V. B., & Londhe, D. D. (2014). Opinion mining and summarization of hotel reviews. *Proceedings - 2014 6th International Conference on Computational Intelligence and Communication Networks, CICN 2014*, 556–559. https://doi.org/10.1109/CICN.2014.126

[2] Cortes, C., Vapnik, V. Support-vector networks. Mach Learn 20, 273–297 (1995). https://doi.org/10.1007/BF00994018

[3] Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3, 993--1022. doi: http://dx.doi.org/10.1162/jmlr.2003.3.4-5.993

[4] Gupta, P., Tiwari, R., & Robert, N. (2016). Sentiment analysis and text summarization of online reviews: A survey. *International Conference on Communication and Signal Processing, ICCSP 2016*, 241–245. https://doi.org/10.1109/ICCSP.2016.7754131

[5] Seyeditabari, A., Tabari, N., & Zadrozny, W. (2018). *Emotion Detection in Text: a Review*. http://arxiv.org/abs/1806.00674

[6] Wahid, D. H., & SN, A. (2016). Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, *10*(2), 207. https://doi.org/10.22146/ijccs.16625

Design and Development of Information System for Administration and Monitoring of Authentic Deed Services for Notary and Conveyancing Using the Prototyping Method

1st Rekyahim
Information Technology Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
rekyahim17si@mahasiswa.pcr.ac.id

2nd Indah Lestari
Information Technology Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
Indah@pcr.ac.id

***Abstract*—Notary Asep Sudrajat, SH., M.Kn. is a company that provides legal services in making authentic deeds and other legally regulated authorities. Currently, Notary Asep Sudrajat, SH., M.Kn has many requests for authentic deed processing, which is currently the filing process and also for submitting information, clients must come to the office or call the notary to find out the stages of transactions that the client has submitted at this time. The notary must look for files to provide information on the progress of deed processing to clients. Sometimes the delivery of information still occurs problems because the deed completion process takes about 1 to 3 months, this can make it difficult for the notary to find the file. From these problems, an Authentic Deed Service Administration and Monitoring Information System will be built. This information system will be developed based on the website in the notary section with the PHP and MYSQL programming language as the database and the android application will be used by the client. As well as the development method using the prototyping method with 2 iterations. Based on the test results, the system built was in accordance with the requests and needs of the Notary Asep Sudrajat, SH., M.kn. The application also runs well with usability testing with the SUS method resulting in a total value of 84, these results indicate that the application has been accepted by the notary client.**

Keywords—Notary Asep Sudrajat, SH., M.Kn, Authentic Deed, PHP, MYSQL, Android, Prototyping.

I. INTRODUCTION

Notary Asep Sudrajat, SH., M.Kn. is a company that provides legal services in making authentic deeds and other authorities that have been legally regulated.

The obstacle during the company's existence was during the authentic deed service process, in which the submission of requirements and knowing the status of the deed were still manual and not yet systematic. The submission of the requirements is carried out by coming directly to the office, of course there are still many clients who do not know the requirements for managing an authentic deed. To find out the status of the deed, the client must call a notary to confirm where the deed is. There are still problems in submitting information because the processing time for the completion of the deed takes about 1 to 3 months. This can make it difficult for the notary to find the file.

From these problems, an Administration Information System and Monitoring Services for Notary was built. This system is designed to make it easier for the notary to view client file submissions, manage file layout, and also provide information on the status of the deed submitted by the client. In addition, this system can make it easier for clients to see the status of the management of the proposed deed. This system is built on a website platform on the part of a notary using the Codeigniter framework and MYSQL as the database. The android application will be used by the client to view the deed transactions developed with the Java programming language, the Android Studio IDE. The method used in this system is to use the prototyping method because it is very well used to resolve misunderstandings between users and analysts due to users not being able to clearly define their needs. The android application will be used by the client to view deed transactions developed with java programming, the Android Studio IDE [3].

II. LITERATUR REVIEW

A. Notary and Conveyancer

According to Article 15 of the Law on Notary Positions number 30 of 2004, what is meant by a notary is a public official who is only authorized to make an authentic deed regarding all acts, agreements, and stipulations required by a general regulation or by interested parties who are required to be stated in a authentic deed, guaranteeing the certainty of the date, keeping the deed and providing Grosse, copies and

tations, all of which are made by a general regulation and are not assigned or excluded to officials or other people.

The definition of Conveyancer is based on the sound of article 1 paragraph 1 of Government Regulation No. 37 of 1998 concerning the Position Regulation of Land Deed Making Officials that what is meant by Conveyancer or Land Deed Making Officials are public officials who are given the authority to make authentic deeds regarding certain legal actions regarding rights. on land or Ownership of Flat Units. Certain legal actions as evidenced by the deed of the Land Deed Official, namely buying and selling, grants, and granting mortgage rights [9].

The management of the name transfer of land certificates consists of a deed of sale and purchase, a deed of grant and a deed of granting mortgage rights. The management of the Notary Deed consists of a power of attorney to sell, a fiduciary guarantee deed and a Limited Liability Company / Limited Liability Company deed [6].

B. Prototyping

Prototyping is a software development method that is suitable for use in software. With this prototyping method, developers and customers can interact with each other during the system creation process [8].

The prototyping stages are as follows:

- Identifying Users: At this stage of gathering needs, the author interviewed the parties concerned to find out about the problems and existing business processes, then got an idea to create the system needed.
- Developing Prototype: Build a prototype by making a temporary design that focuses on presenting input and output.
- Determining whether the Prototype is Acceptable: Evaluate the system prototype that has been made, whether it is in accordance with what is desired. If yes, then the next step is coding the system. If not, a revision will be made to the system prototype that has been made

C. Usability Testing

Usability testing is carried out if the user interface of the application is important and must be specific to a particular type of user. Usability testing is the process of working with end-users directly or indirectly to assess how users perceive a software package and how they interact with it. This process will dismantle the user's trouble areas. The main purpose of usability testing is to improve the usability of a product. For each test, it must have a specific and stated purpose when planning to test [4].

Usability testing is done by using a questionnaire technique for clients who try android applications. In testing this system, the method used is the system usability scale (SUS). SUS can be used to perform independent technology testing on hardware, software, websites, and even mobile devices [10].

In conducting the test, SUS has 10 instruments as shown in Table I The calculation of the application test results is carried out by following the following rules:

- For each statement with an odd number, the respondent's answer scale is reduced by 1.
- Each statement with an even number is 5 minus the respondent's answer scale.
- Summing up the scale of respondents' answers and multiplying by 2.5.
- Then to find the average SUS score, it can be done by dividing the number of SUS scores by the number of respondents.

TABLE I. SUS TESTING INSTRUMENTS

No	Statement	Scale
1	I think that I would like to use this system frequently	1 to 5
2	I found the system unnecessarily complex	1 to 5
3	I thought the system was easy to use	1 to 5
4	I think that I would need the support of a technical person to be able to use this system.	1 to 5
5	I found the various functions in this system were well integrated	1 to 5
6	I thought there was too much inconsistency in this system	1 to 5
7	I would imagine that most people would learn to use this system very quickly	1 to 5
8	I found the system very cumbersome to use	1 to 5
9	I felt very confident using the system	1 to 5
10	I needed to learn a lot of things before I could get going with this system	1 to 5

Determination of the grade on the assessment results seen from Acceptability, grade scale, an adjective rating, which consists of the level of user acceptance there are three categories, namely not acceptable, marginal and acceptable. The grade level has six scales, namely A, B, C, D, E and F. And from the adjective rating it consists of worst imaginable, poor, ok, good, excellent and best imaginable as shown in Figure 1.

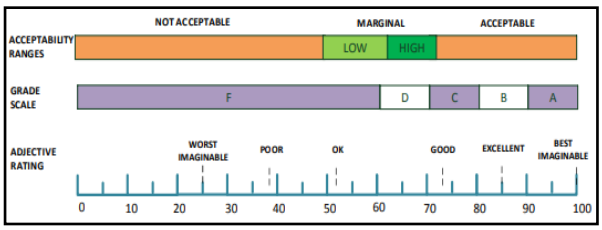


Fig. 1. SUS Rating Scale

III. RESULT AND DISCUSSION

A. Business process of Notary Asep Sudrajat S.H., M.kn

A business process is a set of activities that are interconnected with each other to produce an output that can support the goals of an organization. A business process is a sequence of activities that occur from start to finish and provide the required results for the user.

The land title deed business process starts from the client submits an application for a land deed. After that the notary will check the land certificate, if the certificate is approved the notary will make a land deed and notify the client to come to the notary to sign the deed approval. After that, the notary will provide the signed land deed document along with the client's

personal data to the National Land Agency for the processing of the land title transfer documents which has an estimated time wait about 1 month or 2 months. After land title transfer documents finished the notary returned to contact the client to take a certificate of land it as shown in Figure 2.

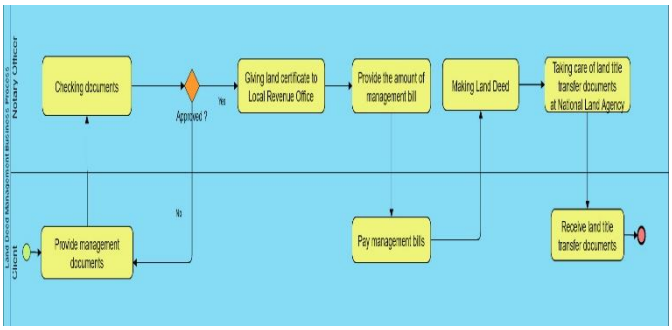


Fig. 2. Business Process Diagram for Land Deed Management

The notarial deed business process starts when the client submits the documents for the deed man agement requirements to the notary. If the requirements submitted are complete, then the notary will give a fee for handling the deed. Then the notary will make a notarial deed according to the type of deed and the client comes to sign the deed. After the signature process is complete, the notary will give the deed to the client as shown in Figure 3.

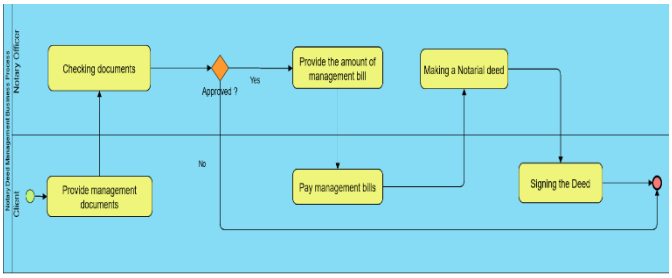


Fig. 3. Business Process Diagram of Notarial deed Management

B. Usecase Diagram

Figure 4 it shows the use case diagrams on the website have 3 actors on the, namely Owner, Employee, and user. Employees have access to confirm files, manage management history, manage work schedules, manage deed storage, add invoices and verify payments. The owner has access to add employees, view management history, and view deed storage. Users have access to view management requirements and download applications. In Figure 5 Use case diagrams on android have actors, namely clients who have access rights to add management requests, view management history, and upload proof of payment.

Fig. 4. Use case Diagrams on Website

Fig. 5. Use case Diagrams on Implementation

This app enabled the client to can add requests, view the progress of the deed history, and pay bills for management. Figure 6 shows the interface for an authentic deed application by uploading a picture of the required documents. Authentic Deed History Page is a page to display the progress of the submitted deed as shown in Figure 7. Figure 8 shows a list of bills added by a notary, clients can also pay by uploading proof of payment.

Fig. 6. Authentic Deed Request Page

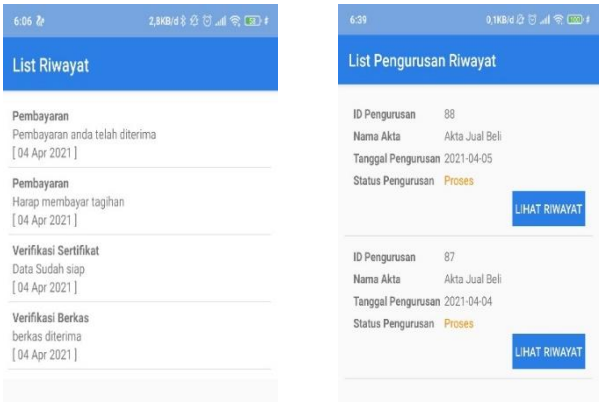


Fig. 7. Authentic Deed History Page

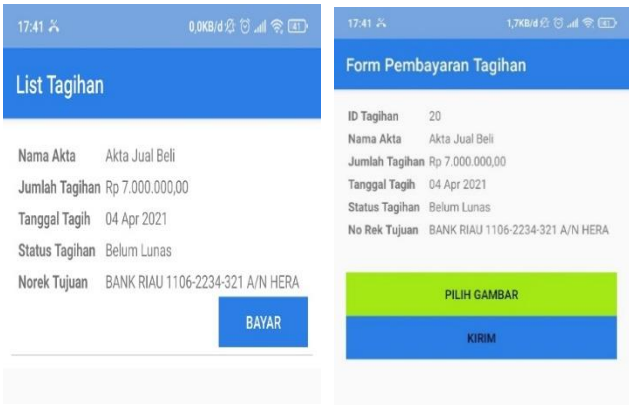


Fig. 8. Billing and Payment Page

This website allows employees to verify deed submission documents, add deed progress and payment verification. Figure 9 shows the deed history list page, on this page employees can add a deed history as shown in Figure 10. Employees can add authentic deed management invoices and client payment verification by clients which can be seen in Figure 11 – 12. In Figure 13 contains the schedule of client management. On the schedule page, notary employees can view and confirm schedules and search features. Notary employees can add data and edit work schedules. On the deed storage page, notary employees can view and delete deed storage and the search feature. Notary employees can add to the deed storage. Notary employees can change the deed storage data as shown Figure 14. Users can download the application and see the requirements for submitting a deed on the home page as shown in Figure 15.

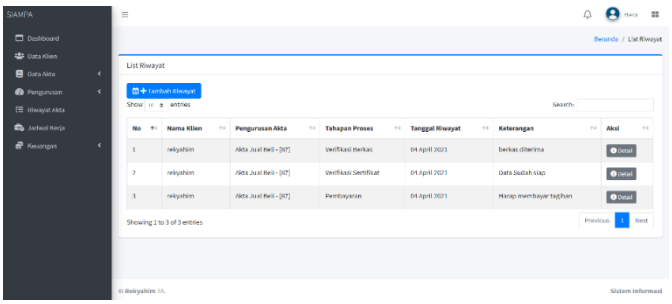


Fig. 9. Billing and Payment Page

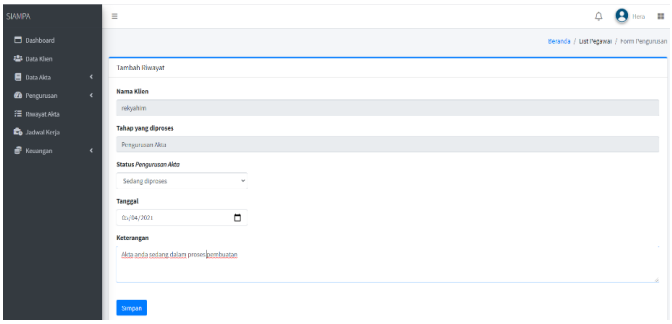


Fig. 10. Billing and Payment Page

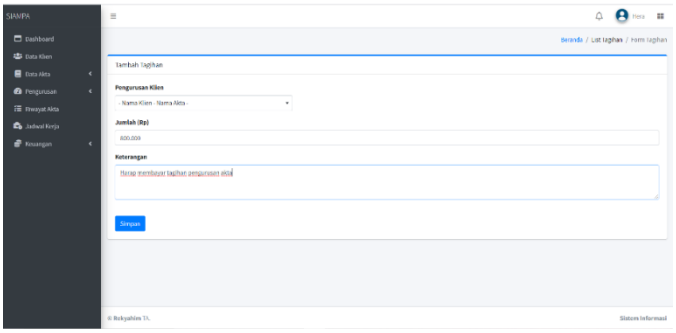


Fig. 11. Billing and Payment Page

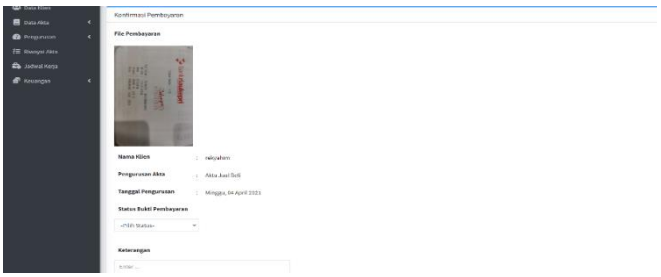


Fig. 12. Billing and Payment Page

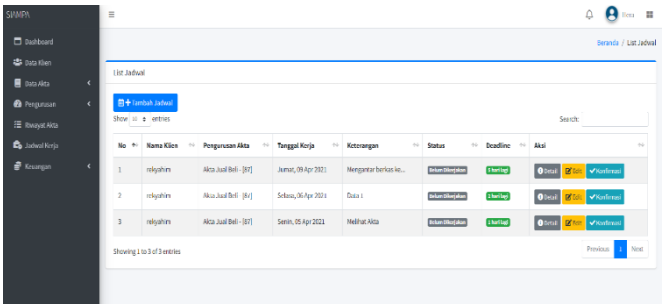


Fig. 13. Work Schedule Page

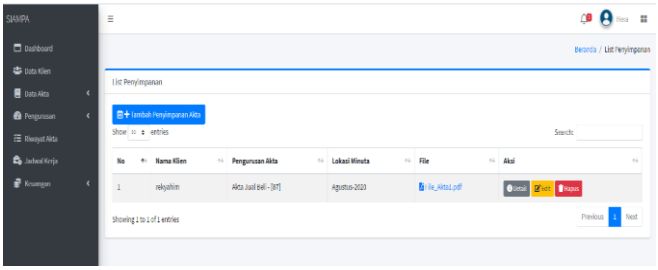


Fig. 14. Deed Storage Page

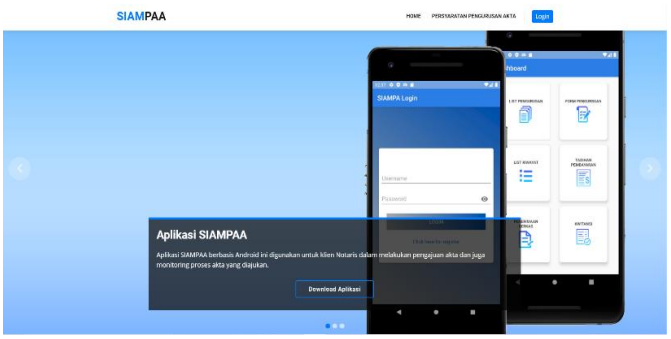


Fig. 15. Homepage

IV. Analysis of Evaluation Results

A. Prototype testing

Based on the testing of prototyping is carried to the notary on 27 July 2020 and 7 September 2020. The prototyping test of the web and android application parts was completed in 2 iterations and the prototype was tested according to the needs of the notary Asep Sudrajat S.H., M.Kn. Prototype build test results in the 1st iteration stage. There is an addition to the Android application, namely the addition of features so that clients can register and submit directly through the Android application and manage work schedules by employees via the web. The test results of the 2nd iteration prototype build show that the application and web have been accepted by the user.

Based on the results of the testing , the system can provide efficiencies to clients when doing filing and monitoring the maintenance of the deed of authentic that previous clients have to come to the notary converted into submission through application android. This web system also provides efficiency in the performance of notary employees in managing the client's deed which includes every progress of the deed management. The previous progress submission was manually converted into a website so that there would be no inconsistency when sending the progress of the deed management to the client.

B. User Acceptence Testing

User Acceptance Testing is only done once after the system is built because each iteration of the prototyping stages of the customer test drives has been carried out and has been approved.

C. Usability Testing

Testing usability testing on the application is done by giving the task and try the application to the 10 respondents. In Table I it can be seen that the Average overall score is 84. The results of the Application Assessment for Adjective Rating get good results. Grade Scale gets Grade B results, and Acceptability Range gets Acceptable results.

The SUS score can indicate the level of user acceptance SUS scores should be worth more than 70 (Brook, 2013) to be included into the category Acceptable. Based on the results of calculations SUS show that the application entered into in the inappropriate or can be received by the respondent who has been testing the application.

TABLE II. SUS Value Calculation on Android

Name	Statement										Score
	1	2	3	4	5	6	7	8	9	10	
Resp't A	4	3	4	3	2	4	4	3	3	2	80
Resp't B	3	4	4	2	4	4	3	4	4	4	90
Resp't C	4	4	3	2	4	3	4	3	3	3	83
Resp't D	4	3	3	4	3	3	3	4	3	3	83
Resp't E	4	3	4	3	4	3	2	4	4	3	85
Resp't F	3	3	3	4	3	4	4	3	4	3	85
Resp't G	4	3	4	3	2	4	4	3	3	2	83
Resp't H	3	4	4	2	4	4	3	4	4	4	88
Resp't I	4	4	3	2	4	3	4	3	3	3	80
Resp't J	3	3	4	3	3	3	3	4	4	3	83
Average overall score of SUS											84

Testing usability testing on a website is done by giving the task and try the application to the three respondents who are party notary.

Based on Table II, the results of testing the usability of the system website which carried on the notary obtained result average overall score is 99 which indicates that the assessment website to Adjective Rating got results Excellent, Grade Scale got the results of Grade A, and acceptability Range gets results Acceptable.

TABLE III. SUS Value Calculation on Website

Name	Statement										Score
	1	2	3	4	5	6	7	8	9	10	
Resp't A	4	4	4	4	4	4	4	4	4	4	100
Resp't B	4	4	4	4	4	4	4	4	4	4	100
Resp't C	4	4	4	4	4	4	4	4	4	3	98
Average overall score of SUS											99

V. Conclusion

The conclusions obtained from the analysis of the system built are as follows:

- Prototype testing on the web and android applications is completed in 2 iterations and prototypes are tested according to user requirements. So that the system can provide efficiency to clients when submitting and monitoring the management of authentic deeds through the android application. The web system also provides efficiency in the performance of notary employees in processing the client's deed management into the website which will avoid inconsistencies when submitting the progress of the deed management to the client.
- Based on User Acceptance Testing (UAT), the system that was built was in accordance with the needs of the notary Asep Sudrajat S.H., M.Kn. and the system is running well.

- Based on the results of usability testing with SUS calculations, the system can be operated properly by the notary on the website and the client on the application.

REFERENCES

[1] Anhar, PHP & MySql Secara Otodidak, Jakarta : MediaKita, 2010.

[2] Bangor, A., Kortum, P., and Miller, J., Determining what individual SUS scores mean Adding an adjective rating scale, 2009

[3] Crystian, C., Sistem Informasi Manajemen Bridal Pada Clara Sky Power Dengan Metode Prototyping, Tugas Akhir, Jurusan Teknologi Industri , Sistem Informasi: Politeknik Caltex Riau, 2018.

[4] Dumas, J. S., & Redish, J., A Practical Guide to Usability Testing,"Maryland: Intellect Books, 1999.

[5] Herlinah, & Musliadi, Pemrograman Aplikasi Android dengan Android Studio, Photoshop, dan Audition, Jakarta Pusat : Elex Media Komputindo, 2019

[6] Moechthar, O, Dasar-dasar Teknik Pembuatan Akta, Surabaya: Airlangga University Press, 2017

[7] Perry, W. E, Effective Methods for Software Testing, Third Edition. indianapolis: Wiley, 2006

[8] Pressman, S. Roger. 2005. E-book Software Engineering: a Practitioner's Approach. Seventh Edition.

[9] Santoso, U. 2017. Pejabat Pembuat Akta Tanah: Perspektif Regulasi, Wewenang, dan Sifat Akta. Jakarta Timur: Prenada Media. 2017

[10] Sauro, J. 2011. A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. North Charleston SC: CreateSpace Independent Publishing Platform.

Detection of Congested Traffic Flow during Road Construction using Improved Background Subtraction with Two Levels RoI Definition

1st Yoanda Alim Syahbana
Graduate School of Engineering
Gifu University
Gifu, Japan
yoanda@pcr.ac.id

2nd Yokota Yasunari
Department of Electrical, Electronics and Computer Engineering,
Faculty of Engineering
Gifu University
Gifu, Japan
ykt@gifu-u.ac.jp

***Abstract*—This study is aimed to detect traffic congestion that may occur during roadblocks of road construction. We improved the background subtraction method by considering Region of Interest (RoI) in the video frame to detect the congestion. The proposed method has experimented with video test material that shows traffic condition in the road construction site. The performance of the proposed method is evaluated using Confusion Matrix by comparing the result of the experiment with ground truth obtained visually. As a benchmarking process, the performance is also compared with the conventional background subtraction method. The result shows that the proposed method can achieve an accuracy of 83.2% for video from the first camera and 82.3% for video from the second camera. In comparison, the conventional background subtraction method only achieves 49.8% for video from the first camera and 0% for video from the second camera. Based on this evaluation, the proposed method can support implementation of efficient traffic control using adaptive traffic light that is equipped with camera.**

Keywords—traffic surveillance, traffic congestion detection, background subtraction, stationary foreground object, Region of Interest (RoI)

I. INTRODUCTION

Road traffic management plays an essential role in the smart city concept. The management collects traffic information and analyzes driving trends in the area. Based on the information, the management controls the signal of the traffic light to adapt to the fluctuation of the traffic situation. In addition, the management also provides information for the driver so that the driving time will be shortened and traffic congestion can be avoided. In terms of traffic safety, the management also ensures traffic accident monitoring and contributes to better comfortability for the road traffic environment.

The video tracking system is the most popular approach to understand traffic conditions in road traffic management [1]. It uses a camera for surveillance purposes. Video frame from the camera is processed to identify and track a moving object. The moving object is further classified as a vehicle, and movement of the vehicle is tracked frame by frame. Instead of using electronic sensors embedded in the roadside, a camera provides a non-invasive approach and is easy to install [2], [3].

The background of this study is to enable efficient traffic control using adaptive traffic light equipped with a camera as illustrated in Fig. 1. Mainly, this study is focused on one-sided alternating traffic sections such as during roadblocks. Camera records traffic activity and sends the captured scene to CPU. CPU performs decision-making process to control the traffic light based on traffic information. In order to implement the system, some objectives are important to be achieved.

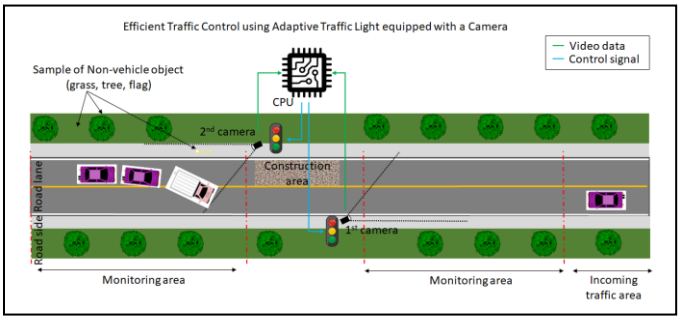


Fig. 1. Illustration of the road construction site.

First, the system needs to detect incoming traffic as early as possible in order sends control signal to the traffic light. Late detection of incoming traffic leads to late signaling of traffic light and this condition is dangerous for high-speed traffic. We had proposed method to achieve this objective in [4]. The result shows that the proposed method based on vanishing point concept has earlier detection of incoming traffic than detection using R-CNN method.

Second, the decision-making process is vital to control timing of traffic light signal based on traffic condition in monitoring area. Ideally, efficient traffic control is achieved when the traffic light signal can make minimum waiting time for both side of traffics. The waiting time is determined by length queuing traffic and waiting duration that can be defined as congestion level. However, since the monitoring area is an unpredictable outdoor environment, monitoring the traffic condition need to consider existence of non-vehicle object such as grass, tree, and flag. For example, existence of flag

may masked the traffic position that cause error in traffic tracking. In addition, movement of non-vehicle objects can bias traffic detection and its condition (stopped or moving).

Therefore, this paper is aimed to report our research in supporting to achieve the second objective. We proposed a method to detect congestion conditions in the traffic. As test material, we use the same video used in [4] that shows the traffic controller managing the traffic flow near the road construction site. In this study case, only one traffic should be allowed at a time. The traffic stopped when the traffic controller raises the red flag. Then the traffic will continue when the traffic controller raises the white flag. To evaluate the proposed method's performance, we compare the timing of detected congestion from the proposed method with the timing of stopped traffic in the video. Timing of stopped traffic is obtained by watching every second of the video test material. Fig. 2 shows sample of captured scenes of the video test material.

Fig. 2. Captured scene in the monitoring area.

The paper is organized into four sections, including this introduction section. Section 2 presents the design of the proposed method. It includes six essential steps of the proposed method. Section 3 explains the design of the experiment and the result of the method evaluation. Finally, Section 4 concludes the paper with a research conclusion and suggestions for future work.

II. PROPOSED METHOD

A. *Design of the Proposed Method*

Application for traffic detection in this study case requires a method to detect the traffic as Foreground Object (FO) in the scene. Most of the widely used methods are called background subtraction [5]. Background subtraction method can be further classified based on mathematical concept, machine learning concept, signal processing model, and classification model [6]. Basic techniques such as temporal median, temporal histogram, and filter are most employed for primary application. It is because the basic technique has low computation and minimum memory requirement.

This study case focused on a particular type of FO namely, Stationary Foreground Object (SFO). SFO is defined as an object that stops and remains static for several consecutive frames [7]. In this study case, SFO represents congested traffic that stops because the traffic controller raised a red flag as a stop signal.

Processing video frame to detect the congestion deals with some challenges. In this study case, the outdoor environment with a variation of sunlight exposure and wind that shakes shrub, grasses, trees, and flags influence the detection of traffic that passes the roadway. In addition, the passing traffic also moves with a variation of speed and direction. Using conventional background subtraction method [5], [8] leads to false detection of congestion.

Therefore, we improved this conventional method considering Region of Interest (RoI) existence in the video frame. The RoI improves the detection quality by focusing the detection in the roadway area only. This study proposed two levels RoI definition to focus the detection in roadway activity only. By this proposed RoI definition, unnecessary FO candidates such as cloud, shrub, grass, tree, and flag can occur. The first level of RoI is aimed to define the main region where traffic movement exists. This first level of RoI is defined in the video before the traffic control is applied. Then, the second level of RoI refines the first level of RoI to focus on the region where the traffic is stopped during the traffic control. This second level of RoI is defined in video when traffic control has been conducted.

Fig. 3 presents the design of the proposed method. First, the background subtraction process is performed for a sequence of video frames. This step detects FO in the video frame. Second, the first level RoI is defined based on the difference of FO from frame to frame. The difference represents the movement of the FO. Then, all of the FO movement is aggregated to obtain the region where the main movement exists in the captured scene. Third, background subtraction is performed once again, especially on RoI defined from the first step. Fourth, the FO detected from this background subtraction is further processed to obtain the second level RoI. In this step, FO that remains static for five sequential frames is defined as SFO. Fifth, the congestion metric is calculated to obtain the congestion timing of the SFO. Finally, resuming time is calculated by calculating the SFO difference from the sequential frame. Resume time is decided based on the maximum difference that represents the movement of the congested car.

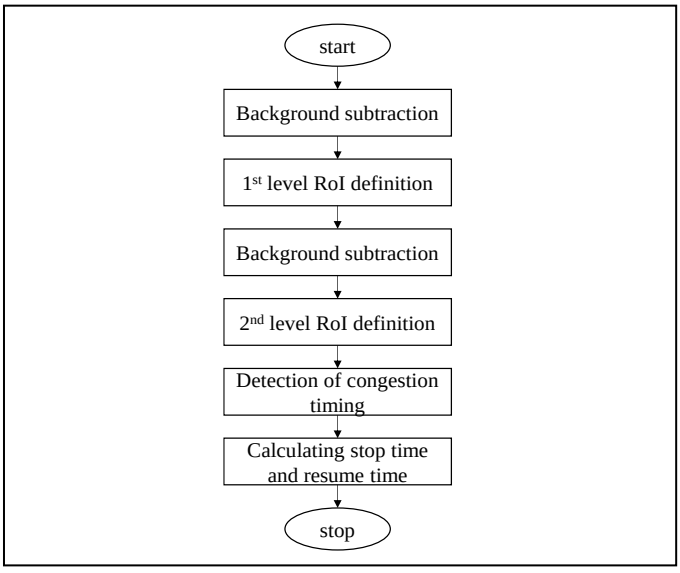


Fig. 3. Design of the proposed method.

The proposed method is designed based on the study case shown in Fig. 1. The main element of the method is using two levels of RoI to enclose the region where the congestion possibly happens. For different study cases with different conditions of the captured scene, the designed method can be adjusted. For example, other methods of FO detection other than background subtraction, such as Gaussian Mixture Model (GMM) or Neural Network, can be implemented with trade-off between accuracy and computation time. In addition, the thresholding process that is predefined in this method can be adjusted based on the condition of the captured scene. The design of the proposed method is also limited to processing traffic video recorded in daylight and fine weather.

B. Definition of Video Frame Variable

Each video frame is represented as $\boldsymbol{I}(x,y,t) \in \{0,1,\ldots,255\}$ where $x \in \{1,2,\ldots,N_x\}$, $y \in \{1,2,\ldots,N_y\}$, and $t \in \{1,2,\ldots,T\}$ with N_x and N_y are width and height of video frame and T is total of the video frame. This study downsamples video frames based on video frame rate to obtain frames for every second. Therefore, $t \in \{20/V_{fps}, 40/V_{fps}, \ldots, T/V_{fps}\}$ where V_{fps} is video frame rate.

C. Background Subtraction

The Background subtraction method is used to detect FO in the evaluated video frame. First, the background frame is generated from the median value of the sequential frame before and after the evaluated video frame, $\boldsymbol{I}(x,y,t)$. Background frame is calculated using Eq. (1)

$$\boldsymbol{I_{BG}}(x,y,t) = \text{median}\big(\boldsymbol{I}(x,y,t-r), \ldots \boldsymbol{I}(x,y,t), \ldots, \boldsymbol{I}(x,y,t+r)\big) \quad (1)$$

where $\boldsymbol{I_{BG}}(x,y,t)$ and r are generated background and range of sequential frame calculated using median, respectively. In this study, r is predefined using Eq. (2)

$$r = \frac{T}{2} - 1 \quad (2)$$

where r is defined based on observation from FO detection quality.

Background subtraction is calculated using Eq. (3)

$$\boldsymbol{I_{FO}}(x,y,t) = |\boldsymbol{I}(x,y,t) - \boldsymbol{I_{BG}}(x,y,t)| \quad (3)$$

where $\boldsymbol{I_{FO}}(x,y,t)$ is detected FO candidate from background subtraction calculation. The FO result includes noise from a variety of road textures and shadows under the car. In addition, the noise also can be a small insignificant object that is far away from the camera. The FO candidate result is filtered from noise using thresholding, blurring, and morphology process to remove the noise.

First, $\boldsymbol{I_{FO}}(x,y,t)$ that consist of $I_R(x,y,t)$, $I_G(x,y,t)$, and $I_B(x,y,t)$ is converted to grayscale image. Grayscale conversion is calculated using Eq. (4)

$$I_{FOgray}(x,y,t) = \frac{I_R(x,y,t) + I_G(x,y,t) + I_B(x,y,t)}{3} \quad (4)$$

where $I_{FOgray}(x,y,t)$ is grayscale conversion result from $\boldsymbol{I_{FO}}(x,y,t)$. To remove noise caused by road texture and shadow variation, $I_{FOgray}(x,y,t)$ with low intensity value lower than 30 is removed from $I_{FOgray}(x,y,t)$. The selection of 30 as a threshold value is predefined in this study. Removal of this type of noise is calculated using Eq. (5)

$$I_{FO}(x,y,t) = \begin{cases} 1, I_{FOgray}(x,y,t) > 30 \\ 0, otherwise \end{cases} \quad (5)$$

where $I_{FO}(x,y,t)$ is detected FO after $I_{FOgray}(x,y,t)$ thresholding. In addition, the morphology process, including blurring and dilation, is also applied to remove the noise.

D. First Level RoI Definition

Since RoI's purpose is to localize the main region where the traffic movement exists, this study uses the frame difference method to obtain moving FO. This step has also been used in [4] for the different research objective as previously mentioned in Section 1. Frame difference is calculated using Eq. (6)

$$I_{FD}(x,y,t) = |I_{FO}(x,y,t) - I_{FO}(x,y,t+1)| \quad (6)$$

where $I_{FD}(x,y,t)$ is detected moving FO. Then, $I_{FD}(x,y,t)$ from all t is summed up to obtain a single frame that shows the number of the region where moving FO existed.

In addition, thresholding is also applied to filter regions with little movement. The selection of 20 as a threshold value is predefined in this study. The process is calculated using Eq. (7)

$$ROI_{first}(x,y) = \begin{cases} 1, \sum_{t=1}^{T} I_{FD}(x,y,t) > 20 \\ 0, otherwise \end{cases} \quad (7)$$

where $ROI_{first}(x,y)$ represents the first level RoI. Among the existing region in $ROI_{first}(x,y)$, the largest region is defined as the first level of RoI because it covers most roadway areas. Fig. 4 shows $ROI_{first}(x,y)$ results from the first camera and second camera.

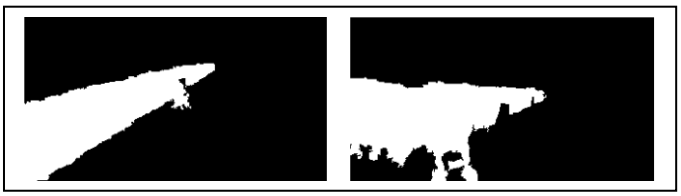


Fig. 4. First level of RoI from the first camera (left figure) and the second camera (right figure).

E. Second Level RoI Definition

The second level of RoI refines the first level of RoI to focus on the region where the traffic is stopped during the traffic control. Background subtraction processes are repeated, and $I_{FO}(x,y,t)$ result is masked using $ROI_{first}(x,y)$. The masking process is calculated using Eq. (8)

$$I_{FO}(x,y,t) = I_{FO}(x,y,t) \cap ROI_{first}(x,y) \quad (8)$$

where $I_{FO}(x,y,t)$ is redefined as a masked result of the process to avoid notation complexity. This study defines SFO as $I_{FO}(x,y,t)$ that remains unchanged for five frames. The SFO is calculated using Eq. (9)

$$I_{SFO}(x,y,t) = I_{FO}(x,y,t-2) \cap I_{FO}(x,y,t-1) \cap I_{FO}(x,y,t) \cap I_{FO}(x,y,t+1) \cap I_{FO}(x,y,t+2), \quad (9)$$

where $I_{SFO}(x,y,t)$ is detected FO that remain static for five sequences of the frame.

Similarly, (7) is modified and used to obtain the second level of RoI. The process is calculated using Eq. (10)

$$ROI_{second}(x,y) = \begin{cases} 1, \sum_{t=1}^{T} I_{SFO}(x,y,t) > 10 \\ 0, otherwise \end{cases} \quad (10)$$

where $ROI_{second}(x,y)$ represent the second level RoI. $ROI_{second}(x,y)$ shows the region where SFO mainly exists, including stopped vehicles and standing traffic controllers. Refining this result, the selection of RoI in this step is based on the width of the RoI candidate. As observed in this study, the width of the stopped vehicle is wider than the width of the standing traffic controller. Fig. 5 shows $ROI_{second}(x,y)$ results from the first camera and second camera. The white region in the figure represents the second level of RoI.

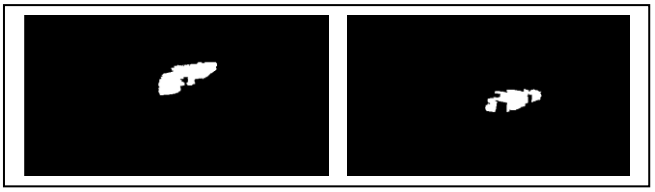


Fig. 5. Second level of RoI from the first camera (left figure) and the second camera (right figure).

The first level of RoI is defined using the video before the traffic control is applied. Therefore, movement activity is well captured. On the contrary, the second level of RoI is defined using the video after the traffic control is applied. Following these two processes, the rest of the video is processed using the background subtraction method. In addition, $I_{FO}(x,y,t)$ result is masked using $ROI_{second}(x,y)$. The masking process is calculated using Eq. (11)

$$I_{FO}(x,y,t) = I_{FO}(x,y,t) \cap ROI_{second}(x,y) \quad (11)$$

where $I_{FO}(x,y,t)$ is redefined as a masked result of the process to avoid notation complexity. Then, (9) is used to calculate masked $I_{SFO}(x,y,t)$ as a static FO that stopped for minimum of five frames.

F. Detection of Congestion Timing

In order to obtain timing information of significant SFO, $I_{SFO}(x,y,t)$ is summed up for all of its y. The aim is to obtain the maximum horizontal width that is defined as the traffic congestion candidate. This approach is selected because most of the traffic movement is mainly progressed in the horizontal width of the video frame. This process is calculated using Eq. (12)

$$length_{SFO}(x,t) = \sum_{y=1}^{N_y} I_{SFO}(x,y,t) \quad (12)$$

where $length_{SFO}(x,t)$ is the length of SFO. Then, timing of the congestion candidate is calculated using Eq. (13)

$$congestion(t) = \begin{cases} \sum_{x=1}^{N_x} length_{SFO}(x,t), length_{SFO}(x,t) > 20 \\ 0, otherwise \end{cases} \quad (13)$$

where $congestion(t)$ is a metric to measure the congestion candidate.

G. Calculating Stop Time and Resume Time

The timing represents the stop time of the front-most vehicle. However, it does not precisely provide information on the resume time. Therefore, frame difference is applied again for all $I_{SFO}(x,y,t)$ inside the timing. It also aimed to check whether the candidate is valid congestion or only slowed traffic. The process is calculated using Eq. (14)

$$I_{FD-SFO}(x,y,t) = |I_{SFO}(x,y,t) - I_{SFO}(x,y,t+1)| \quad (14)$$

where $I_{FD-SFO}(x,y,t)$ is absolute difference of $I_{SFO}(x,y,t)$. Then, $I_{FD-SFO}(x,y,t)$ is further processed to determine the congestion state of the traffic. The congestion state is calculated using Eq. (15)

$$congestion_{sum}(t) = \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} I_{FD-SFO}(x,y,t) \quad (15)$$

where $congestion_{sum}(t)$ is the congestion state that represents how the $I_{SFO}(x,y,t)$ changed inside the timing.

III. Experiment and Evaluation

A. Video Test Material

Video test material is obtained from In-Luck Company. This company provides services of security, including traffic control in road construction. The video test material shows two captured scenes from both sides of the roadway. It shows traffic flow that will pass the construction site controlled by the traffic controller. Captured scene from the first and the second camera shows the roadway and the traffic controller. The scene also shows the environment surrounding the road, such as the parking area and rice field. In addition, the captured scene also shows a waving flag on the roadside. These non-vehicle objects such as shrubs, grasses, trees, flags, and their movements influence traffic detection.

Sample of video test material is selected from the first and the second camera. The video has been synchronized to align the recorded duration between the cameras. Four sample videos recorded from 09:10 a.m. to 09:30 a.m from each captured scene from the first and the second camera are used for the experiment and evaluation. Each of the videos has a total of T=300 frames. The videos have N_x=640 and N_y=360 with V_{fps}=20.

B. Experiment Setup

The proposed method is implemented using Matlab R2020b (64 bit) with the operating system Microsoft Windows 10 Pro. The method has experimented on Intel core i9 processor with 32 GB memory. This study also utilizes parallel processing to optimize the processing time of the video frame.

C. Experiment Result

Fig. 6 shows $congestion(t)$ as a result of (13) for the video test material that has been processed by the proposed method. The vertical axis of Fig. 6 represents the congestion metric that shows the length of congested traffic. This metric along the horizontal axis represents the congestion, especially for the initial stop time of the congested traffic. As highlighted by the red ellipse, congested traffic is detected from the start of the video until frame t=25 in the first camera. After the congested traffic is detected in the first camera, the proposed method also detects congested traffic in the second camera from t=14 to t=50. Therefore, this timing pattern shows that the roadway is only accessed one at a time.

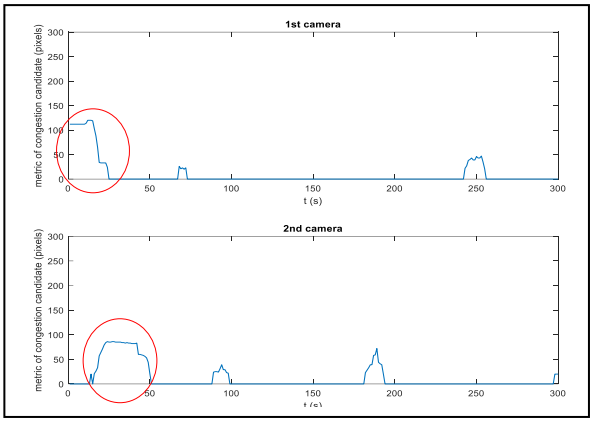


Fig. 6. Candidate of congestion timing.

Following the result discussed in Fig. 6, the proposed method processes the result to detect the resume time of the congested traffic. Eq. (14) and Eq. (15) are used to determine the resume time. For example, Fig. 7 shows the result of $congestion_{sum}(t)$ for t=1 to t=25 from the first camera. The peak of $congestion_{sum}(t)$ highlighted by the red circle represents the resume time of the congested traffic. In this example, t=17 is the resume time.

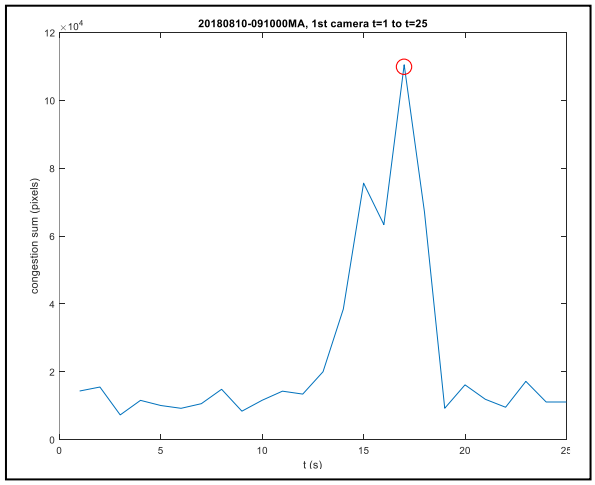


Fig. 7. Result of resume time calculation from the proposed method.

D. Evaluation

This study uses Confusion Matrix to evaluate the performance of the proposed method. Confusion Matrix compares congestion timing from the proposed method with congestion timing from the ground truth. The ground truth is obtained by watching the video visually and record the timing manually. Table I shows the rule of Confusion Matrix that is used to evaluate the performance of the proposed method. Confusion Matrix compares congested and uncongested timing between the proposed method and the ground truth.

TABLE I. RULE OF CONFUSION MATRIX

		Proposed Method	
		Congested	**Uncongested**
Ground Truth	**Congested**	True Positive (TP)	False Negative (FN)
	Uncongested	False Positive (FP)	True Negative (TN)

Table II summarizes the result of the Confusion Matrix calculation from each sample of video test material that has been averaged.

TABLE II. RESULT OF CONFUSION MATRIX ELEMENT (%)

		TP	FP	FN	TN
Proposed method	First camera	74	0.15	16.65	9.12
	Second Camera	82.25	17.75	0	0
Conventional background subtraction method	First camera	31	43.15	7.675	18.12
	Second Camera	0	100	0	0

Based on the rule, the accuracy of the proposed method is calculated using

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (16)$$

where *Accuracy* is metric to evaluate performance of the proposed method.

Based on Confusion Matrix calculation, the proposed method can achieve an accuracy of 83.2% for video from the first camera and 82.3% for video from the second camera. In comparison, processing video using the conventional background subtraction method only achieves 49.2% for video from the first camera and 0% for video from the second camera. The existence of a waving flag causes a 0% result for video from the second camera detected as SFO.

The performance of the proposed method has been experimented on the captured scene, as shown in Fig. 1. Therefore, the design of the proposed method is fine-tuned with the study case and has some limitations. The proposed method's design has not considered a variation of captured scene conditions influenced by camera position, angle, and lens. For example, a high position capturing camera provides a different angle to capture traffic congestion. In this condition, conversion of world space to camera space is required. In addition, the captured scene is also influenced by the perspective projection that also needs to be considered. The study case also shows a straight road lane without an intersection. The existence of more than one road lane, intersection and curved road will influence estimated RoI and calculation of Eq. (15).

IV. CONCLUSION AND FUTURE WORK

A. Conclusion

This study has proposed a method to detect congested traffic that is happened during road construction. The proposed method improved the existing background subtraction method with two levels of RoI definition. The improvement aims to minimize the influence of non-vehicle objects in the captured scene and focus the detection on roadway areas only. The experiment of the proposed method has been conducted using video test material obtained from the In-Luck company. Performance evaluation using Confusion Matrix also has been performed. The result shows that the proposed method achieves better accuracy in detecting congestion than the conventional background subtraction method. Based on the benchmarking result, non-vehicle object movement such as waving flags influences congestion

detection. In conclusion, the proposed method can support road traffic management in detecting traffic congestion, especially in temporary surveillance such as in road construction.

B. Future Work

It is interesting to evaluate the further performance of the proposed method in a variety of environmental conditions, such as roads with more than one road lane, road intersection, and different weather conditions. It is also intriguing to implement other FO methods with a combination of two levels of RoI definition. In addition, it is also possible to continue to other objective of the research as mentioned in Section 1 that is designing a decision making process to achieve efficient traffic control using adaptive traffic light.

REFERENCES

[1] Y. B. B. Pethe, "An implementation of Moving Object Detection , Tracking and Counting Objects for Traffic Surveillance System," 2011, doi: 10.1109/CICN.2011.28.

[2] M. Fathy and M. Y. Siyal, "A window-based image processing technique for quantitative and qualitative analysis of road traffic parameters," *IEEE Trans. Veh. Technol.*, vol. 47, no. 4, pp. 1342–1349, 1998, doi: 10.1109/25.728525.

[3] D. M. Ha, J. M. Lee, and Y. D. Kim, "Neural-edge-based vehicle detection and traffic parameter extraction," *Image Vis. Comput.*, vol. 22, no. 11, pp. 899–907, 2004, doi: 10.1016/j.imavis.2004.05.006.

[4] Y. A. Syahbana and Y. Yokota, "Early Detection of Incoming Traffic for Automatic Traffic Light Signaling during Roadblock using Vanishing Point-Guided Object Detection and Tracking," in The SICE Annual Conference 2021, Tokyo, Japan, 2021.

[5] A. Sobral and A. Vacavant, "A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos," *Comput. Vis. Image Underst.*, vol. 122, pp. 4–21, 2014, doi: 10.1016/j.cviu.2013.12.005.

[6] B. Garcia-Garcia, T. Bouwmans, and A. J. R. Silva, "Background subtraction in real applications: Challenges, current models and future directions," *Comput. Sci. Rev.*, vol. 35, p. 100204, 2020, doi: 10.1016/j.cosrev.2019.100204.

[7] C. Cuevas, R. Martínez, and N. García, "Detection of stationary foreground objects: A survey," *Comput. Vis. Image Underst.*, vol. 152, pp. 41–57, 2016, doi: 10.1016/j.cviu.2016.07.001.

[8] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: An overview," *Comput. Sci. Rev.*, vol. 11–12, no. October, pp. 31–66, 2014, doi: 10.1016/j.cosrev.2014.04.001.

Performance Improvement and User Satisfaction Analysis on Android-Based Educational Mobile Game

1st Anggy Trisnadoli
Information Technology Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
anggy@pcr.ac.id

2nd Kelvin
Information Technology Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
kelvin17ti@mahasiswa.pcr.ac.id

***Abstract*—Tourism is a sector currently in the spotlight by many parties, including the government. This is because tourism attracts tourists. This tourism can drive the economy and empower the community. However, the lack of promotion from the government and the community makes tourism in Riau Province less well known. Therefore, in 2016 the educational game "Ayo Wisata ke Riau" was developed, aiming to provide education about tourism in Riau Province. However, the developed game has several problems, such as content that is not up to date and inappropriate, the size of the game is too large, and the display is not attractive. Therefore, a re-engineering process with a User Experience (UX) approach is needed to overcome these problems. After the re-engineering process, Usability testing and User Experience Questionnaires were carried out to test user satisfaction and experience using the application. Testing is done by doing pre-test and post-test. The pre-test is a test by playing games that have been developed in 2016, and a post-test is a test by playing games after the re-engineering process is carried out. Based on the results of the tests carried out, the level of user satisfaction increased by 49.7%, and the results of the evaluation of UX measurements showed positive evaluation results, which meant that users were satisfied in using the application.**

Keywords—Riau Province Tourism, Educational Mobile Game, User Experience

I. INTRODUCTION

Nowadays, technology was developed very rapidly and cannot be avoided. Technology development plays an essential role in human life because technology makes it easier for humans to carry out daily activities. One of the positive impacts of the rapid development of technology is that it is easier for people to get the information they need to learn many new things.

The use of technology such as games as a medium of education about tourism in Riau Province is still rare. Games can be a medium to provide learning to players through games that are easy to understand. Educational themed games can provide learning to players, so players can learn new things from the games they play.

Tourism is a sector currently in the spotlight by many parties, one of which is the government. This is because tourism attracts tourists from within and outside the Riau Province. Thus tourism can drive the economy and empower the community. However, the lack of promotion from the government and the community makes tourism in Riau Province less well known.

Febriano et al [1] have developed a mobile educational game, "Ayo Wisata ke Riau". This game aims to provide education about tourism in Riau Province to the public. The game runs on android-based mobile devices and has a 2D display. However, in the game, several problems cause user satisfaction in using the application to be only 82.49%.

Problems that arise starting from the size of the game that is too large, the display is less attractive, and the content is not up to date. Therefore it is necessary to re-fix the game. The problems that arise in the game will be discussed in the Focus Group Discussion (FGD). The results of the FGD are used as a reference for the redevelopment of tourism educational mobile games in Riau Province (Kreshna et al., 2020), resulting in an optimization of application improvements based on the user's point of view (Trisnadoli et al., 2021). However, in its implementation, UX improvement and user satisfaction measurement are still not very visible.

Based on the above problems, an analysis of the performance improvement of the mobile educational game "Ayo Wisata ke Riau" was carried out. Performance improvements are carried out using an approach to the User Experience of players when using the application. It is hoped that with the performance improvements made, the level of user satisfaction can increase.

II. RESEARCH METHODOLOGY

The stages of the development of the 'Ayo Wisata ke Riau' mobile game are:

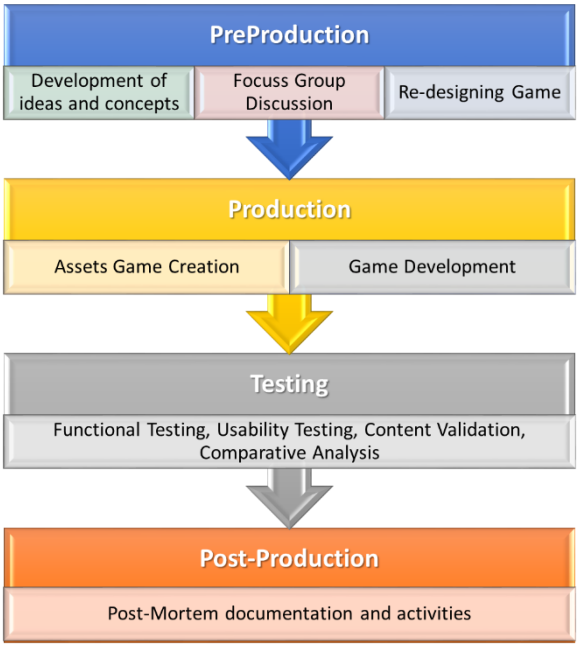


Fig. 1. Game Development Stages

There are several stages in game development, from game ideas to playable games, namely Pre-production, Production, Testing, and Post-production.

1) Pre-production

Pre-production is the stage to determine the game to be developed and project planning. This stage focuses on finding ideas, game concepts, and game development designs.

2) Production

Production is the main stage in game development. This stage implements the ideas, concepts, and game designs that have been made in the pre-production stage.

3) Testing

After the production stage is complete, the game that has been developed will be tested. Tests are carried out to ensure the game can run well, look for potential bugs, and fix bugs.

4) Post-production

Post-production is the stage for making documentation and post-mortem activities.

The testing methods used in this study are:

1) Blackbox Testing

Blackbox testing is a type of software testing on system functionality. The purpose of this test is to test each function of the game according to its function or not.

2) Content Validation

Content validation is required so that the content presented in the game corresponds to the actual situation. In this test, tourism experts are needed so that the content presented is valid. The expert who will validate the content presented in the game is the head of tourism resource development. If inappropriate content is found, a review will be carried out on the content presented. If appropriate, the test is declared valid.

3) Usability Test

Usability Test is a type of software testing conducted to measure the level of user satisfaction. Players will provide feedback by filling out the questionnaire given after playing the game that has been developed. The questionnaire uses a Likert scale and has several criteria that players will answer.

4) User Experience Questionnaire

The User Experience Questionnaire is a test that assesses user experience in using the application. Players will provide feedback by filling in the UEQ given after playing the game that has been developed. Six factors will be measured in the UEQ, namely Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty.

The research method at least describes the approach used in the study, the population, and the research sample explains the operational definition of variables along with data measurement tools or methods of collecting data and data analysis methods.

If the data measurement tool uses a questionnaire, it is necessary to include the results of the validity and reliability test of the research instrument.

III. GAME PRODUCT DEVELOPMENT RESULTS

The following image is the initial view when we run the game. There is a Play button to process the following display and a settings button for settings in the game. Players must press the play button to move to the Riau Province map page.

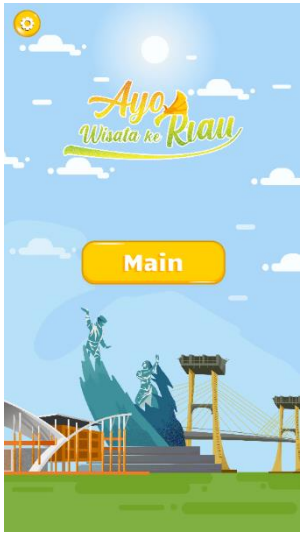


Fig. 2. Main Menu of Game Product

Fig. 2 displays the Riau Province map page that appears when the Main button is pressed. In this scene, the player can choose the area to start the game. Each selected area has a different type of game, including quiz games, guessing pictures, and puzzles. Below is a selection of games for each region in Riau Province. Regional selection is based on the district/city because the area is the main area of the Riau province

TABLE I. GAME OPTIONS OF GAMEPLAY

No.	Game Options	City / Regency
1	Questions	Pekanbaru City
		Kampar Regency
		Pelalawan Regency
		Dubai City
2	Guess the picture	Rokan Hulu Regency
		Rokan Hilir Regency
		Kuantan Singingi Regency
		Siak Regency
3	Puzzle	Bengkalis Regency
		Kepulauan Meranti Regency
		Indragiri Hulu Regency
		Indragiri Hilir Regency

Fig. 3. Main Map of Riau Province as in Gameplay

Fig. 4 shows that there are three game pages as quizzes, guess pictures, and puzzles. Each player will be given ten different challenges and a playing time limit. Each completed challenge will get a score of 10 points. In each game, the player will be given the correct answer if the playing time has run out or the answer given is wrong.

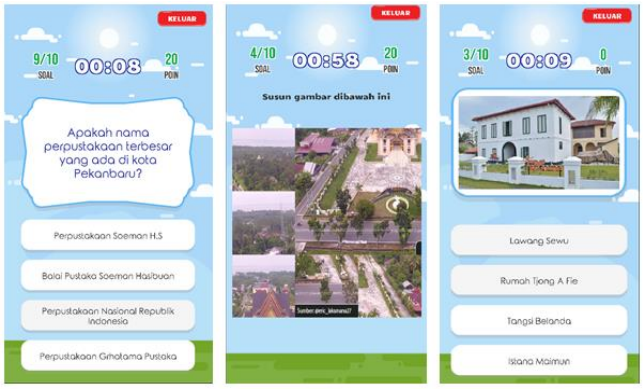


Fig. 4. Gameplay in Ayo Wisata ke Riau Mobile Game

After the game is over, the final score will appear. Players will be asked to enter a name so that the score can be saved. The saved score will be displayed on the scoreboard.

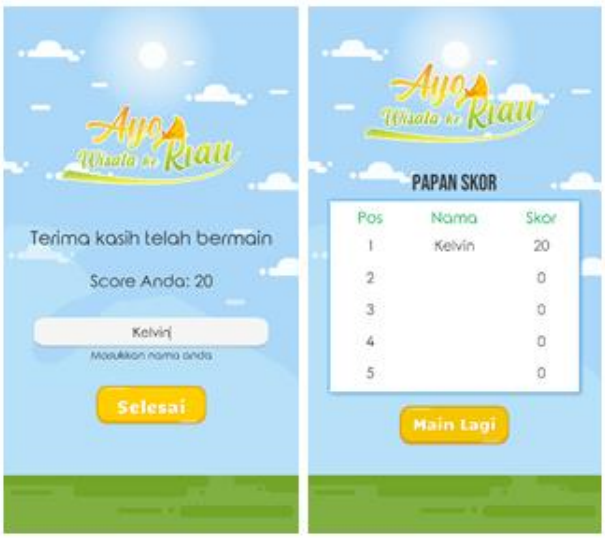


Fig. 5. Scoreboard of Game

IV. ANALYSIS OF TESTING RESULTS

A. *Analysis of the results of the Black Box Test*

Blackbox testing was carried out to test the functional game 'Ayo Wisata ke Riau' that had been built. Tests are carried out to check whether all game functionalities can run as expected or not. There are several test cases carried out with different scenarios to test all game functionality. The test results show that all game functionality runs according to the expected results based on the tests carried out.

TABLE II. BLACK BOX TEST RESULTS

No	*Test Case*	*Scenario*	*Expected Result*	*Actual Result*	Status
1	Fungsi menu main	Pemain membuka permainan lalu menekan tombol "Main"	Menuju ke *scene* peta Provinsi Riau	Menuju ke *scene* peta Provinsi Riau	Valid
2	Fungsi keluar permainan	Pemain menekan tombol Keluar lalu memilih "Ya"	Pemain keluar dari game	Pemain keluar dari game	Valid
		Pemain menekan tombol Keluar lalu memilih "Tidak"	Pemain tetap berada pada game	Pemain tetap berada pada game	Valid
3	Fungsi menu bantuan	Pemain menekan tombol *Setting* lalu memilih menu "Bantuan"	*Pop-up* bantuan muncul	*Pop-up* bantuan muncul	Valid
4	Mulai permainan	Pemain memilih Kabupaten/Kota pada peta Provinsi Riau lalu menekan "Lanjut"	Pindah ke *scene* permainan	Pemain masuk ke *scene* permainan	Valid
5	*Gameplay* permainan kuis	Pemain membuka halaman permainan kuis lalu menjawab soal dengan Benar	*Game* akan memberitahu bahwa jawaban benar dan skor bertambah	Jawaban benar dan skor bertambah	Valid
		Pemain membuka halaman permainan kuis lalu menjawab soal dengan Salah	*Game* akan memberitahu jawaban salah dan benar	*Game* memberitahu jawaban salah dan benar	Valid
		Pemain membuka halaman permainan kuis lalu tidak menjawab soal hingga waktu habis	*Game* akan memuat soal selanjutnya	*Game* memuat soal selanjutnya	Valid
6	*Gameplay* permainan tebak gambar	Pemain membuka halaman permainan tebak gambar lalu menebak gambar dengan Benar	*Game* akan memberitahu bahwa jawaban benar dan skor bertambah	Jawaban benar dan skor bertambah	Valid
		Pemain membuka halaman permainan tebak gambar lalu menebak gambar dengan Salah	*Game* akan memberitahu jawaban salah dan benar	*Game* memberitahu jawaban salah dan benar	Valid
		Pemain membuka halaman permainan tebak gambar lalu tidak menebak gambar hingga waktu habis	*Game* akan memuat soal selanjutnya	*Game* memuat soal selanjutnya	Valid
7	*Gameplay* permainan *puzzle*	Pemain membuka halaman permainan *puzzle* lalu menyusun puzzle hingga selesai sebelum waktu habis	Skor pemain akan bertambah	Skor bertambah	Valid
		Pemain membuka halaman permainan *puzzle* lalu tidak menyusun *puzzle* hingga selesai sampai waktu habis	Pemain akan diberi gambar utuh dan memuat *puzzle* selanjutnya	Pemain diberi gambar utuh dan memuat *puzzle* selanjutnya	Valid

B. *Analysis of Usability Testing results*

Based on the data recapitulation of usability test results, it addressed 62 children aged 12-19 years. There are 15 assessment criteria, each of which is given five answer options, namely "Strongly Disagree", "Disagree", "Hesitating", "Agree", and "Strongly Agree". Each answer has different points ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

The collected test data will be processed using the Likert scale rule. After the data is processed, the results are obtained as in Table II. The table below shows the percentage is in the range of 80%-100%, which means "Very Good".

TABLE III. USABILITY TEST RESULTS

No.	Kriteria	Pilihan Jawaban					Score	Percentage (%)
		SS	S	RG	TS	STS		
1	Pemain mudah untuk memahami permainan	47	15	0	0	0	295	95.16%
2	Pemain mendapatkan pengetahuan dari permainan	46	16	0	0	0	294	94.84%
3	Pemain merasa terhibur dengan permainan	40	20	2	0	0	286	92.26%
4	Pemain akan merekomendasikan permainan ini kepada teman dan kerabat	24	33	5	0	0	267	86.13%
5	Tampilan permainan nyaman untuk dilihat	37	24	1	0	0	284	91.61%
6	Audio pada permainan sudah nyaman didengar	36	23	3	0	0	281	90.65%
7	Pengontrolan permainan sudah sesuai dengan standar	36	26	0	0	0	284	91.61%
8	Bahasa yang digunakan mudah dipahami	43	18	1	0	0	290	93.55%
9	Fitur pada permainan sudah lengkap	37	19	5	1	0	278	89.68%
10	Pengalaman bermain menyenangkan	31	29	2	0	0	277	89.35%
11	Konten pariwisata pada game sudah terbaru	34	25	3	0	0	279	90.00%
12	Tampilan permainan sesuai dengan layar perangkat *mobile*	46	16	0	0	0	294	94.84%
13	Dengan adanya skor dan waktu menjawab, pemain merasa tertantang untuk bermain	36	25	1	0	0	283	91.29%
14	Dengannya adanya penjelasan mengenai daerah yang dipilih, pemain dapat lebih mengenal daerah tersebut	39	21	2	0	0	285	91.94%
15	Dengan adanya soal yang bervariasi pemain tidak merasa bosan saat bermain	34	26	2	0	0	280	90.32%

C. *Analysis of the results of the User Experience Questionnaire Test*

In UEQ, there are six factors to be measured: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty. The results obtained from these criteria will be processed to obtain the average value as in table III. In the UEQ, it is stated that the average value of each factor is in the range of -3 to +3. An average value of -3 indicates a negative evaluation result, while an average value of +3 indicates a positive evaluation result. The average value between -0.8 to +0.8 indicates a neutral evaluation result.

TABLE IV. USER EXPERIENCE QUESTIONNAIRE RESULTS

Faktor	Pre-Test	Post-Test
Attractiveness	-1.683	2.311
Perspicuity	-0.725	2.142
Efficiency	-1.692	2.317
Dependability	-1.558	2.083
Stimulation	-1.750	2.117
Novelty	-1.533	1.800

It can be seen in Table II the comparison of the results of the pre-test and post-test conducted to measure user experience. In the pre-test, the attractiveness, efficiency, dependability, stimulation, and novelty factors were negative, and perspicuity factors were neutral. In the post-test, all the measured factors showed positive evaluation results. If the results of the two tests are compared, it can be seen that there is an increase in the results of the evaluation of UX measurements after the re-engineering process is carried out.

When visualized in graphical form, the following is a description of the results of the tests carried out.

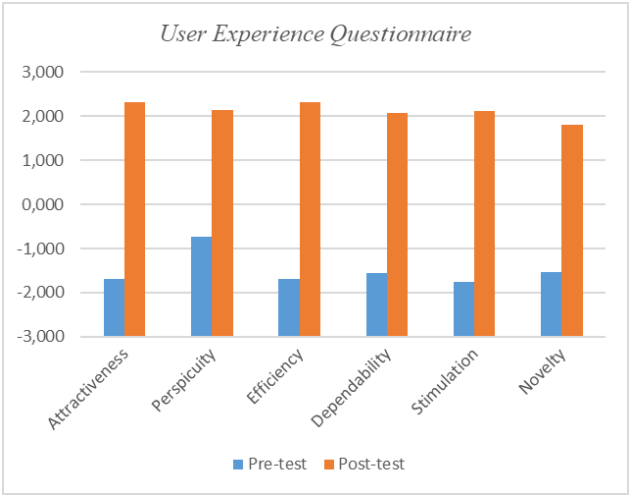


Fig. 6. UEQ Hasil Results Comparison

V. CONCLUSIONS

Based on all phases of the activities that have been carried out, in this study can be concluded that from the results of usability testing, the level of user satisfaction is in the range of 80%-100%, which means Very Good. Moreover, based on the User Experience Questionnaire results, which was conducted to measure user experience when playing the game 'Ayo Wisata ke Riau' the scale is 0.8 up to 3, which means the user experience while playing is very good. So it can be concluded that the 'Ayo Wisata ke Riau' game meets the quality needs of tourism educational mobile games in Riau Province.

So this research proves that the process of re-engineering the 'Ayo Wisata ke Riau' game increases user satisfaction by 49.7% and improves user experience in using the application.

REFERENCES

[1] Cahyadi, R. Pengertian Edukasi. Retrieved from Medium: https://medium.com/@tibalagiblog/pengertian-edukasi-329d0a154d2, 2019

[2] Candra Agustina, T. W. Aplikasi Game Pendidikan Berbasis Android Untuk Memperkenalkan Pakaian Adat Indonesia. Indonesian Journal on Software Engineering, 2015

[3] Febriano, E., Trisnadoli, A., Putra, Y. E. Pembangunan Mobile Game Edukasi Kebudayaan dan Pariwisata Provinsi Riau. Jurnal Aksara Komputer Terapan, vol. 6, no 2, 2017 (*references*)

[4] Himatyar, M. Analisis Pengembangan Game Edukasi 'Indonesiaku'. Yogyakarta: Universitas Negeri Yogyakarta, 2015

[5] Husniah, L. Gamification and GDLC (Game Development Life Cycle) Application for Designing the Sumbawa Folklore Game "The Legend of Tanjung Menangis (Crying Cape)". Malang: Universitas Muhammadiyah Malang, 2018

[6] Intan Trivena Maria Daeng, N. M.. Penggunaan Smartphone Dalam Menunjang Aktivitas. Acta Diurna, 2017

[7] Kreshna, J. A., Trisnadoli, A. Pengumpulan Kebutuhan Kualitas Dari Pengguna Dalam Rangka Reengineering Aplikasi Permainan Edukasi 'Ayo Wisata ke Riau'. Seminar Nasional Teknologi Informasi & Ilmu Komputer (SEMASTER) 2020, Pekanbaru, 2020

[8] Rido Ramadan, Y. W. Game Development Life Cycle Guidelines. ICACSIS, 95-97, 2013

[9] TechTerms. Android. Retrieved from TechTerms: https://techterms.com/definition/android, 2016

[10] Trisnadoli, A., & Kreshna, J. A. Optimization of Educational Mobile Game Design 'Ayo Wisata ke Riau' based on User's Perspective. IT Journal Research and Development, 6(1), 52 - 59. https://doi.org/10.25299/itjrd.2021.vol6(1).5766, 2021

[11] Zamojc, I. Introduction to Unity3D. Retrieved from Envatotuts+: https://code.tutsplus.com/tutorials/introduction-to-unity3d--mobile-10752, 2012

UKT (Single Tuition) Classification Prediction uses MKNN (K-Nearest Neighbor Modification) Algorithm

1st Dziki Adli
Department of Information Technology
Politeknik Caltex Riau
Pekanbaru, Indonesia
dziki20s2tk@mahasiswa.pcr.ac.id

2nd Dadang Syarif Sihabudin Sahid
Department of Information Technology
Politeknik Caltex Riau
Pekanbaru, Indonesia
dadang@pcr.ac.id

***Abstract*—Islamic University of Sultan Syarif Kasim (*UIN SUSKA*) Riau Province has used an information system, namely the *Sistem Registrasi* (SIREG) to facilitate the registration of prospective students and also SIREG also provides a decision on determining the *UKT* of students who have been declared graduated. But there has never been an evaluation of SIREG's accuracy in determining the *UKT*. From these problems, a model is needed to be implemented to facilitate *SIREG* officers in conducting classifications to establish *UKT* new students. Using the MKNN method and supported by the K-Fold Cross Validation validation method, the classification results get an accuracy value of 71%.**

Keywords—Classification, KNN, K-Fold Cross Validation

I. INTRODUCTION

One of the public universities in Indonesia and has used the UKT system is UIN Suska Riau, the UKT payment system has been implemented in UIN Suska Riau since the 2014/2015 school year which has been implemented in the SIREG application. The process for determining the UKT group requires rigor and time, as student data will be compared to UKT criteria one by one. The decision-making system in determining existing UKT groups uses 5 criteria in assessing the ability of parents of students.

MKNN is an algorithm developed from the KNN algorithm, the MKNN algorithm adds a new process to perform classification i.e. calculation of validity values to consider validity between training data and weighted voting calculations to calculate the weight of each nearby. The addition of 2 new processes in the MKNN is expected to correct any errors in the KNN process. Based on the research will be carried out algorithm classification prediction using K-NN Modification on UKT UIN Suska data because it has been done in previous research for UKT predictions at Riau University using KNN. It is expected that by using the MKNN algorithm, better accuracy results are obtained.

II. RELATED WORK

Sukamto, 2020 [1], the research conducted resulted in an accuracy value from classification using KNN algorithm obtained by 84.21% to predict the UKT that will be paid by prospective students, especially the S1 study program of FMIPA Information System, Riau University. The criteria used are gross income, tuition insurers, the number of dependents listed in the electricity card, the status of residence, the state of the walls of the residence, the state of the roof of the residence, the total area of land ownership and the cost of electricity usage a month. The UKT groups are UKT1, UKT2, UKT3, UKT4, UKT5 and UKT6. The data used is students of S1 FMIPA Universitas Riau in the class of 2016, 2017, and 2018. The ratio for training data and test data is 90%: 10%.

Okfalisa, 2018 [2], research was conducted using the Dataset of the Hope Family Program Implementation Unit using 7,395 records of data has been compared KNN with MKNN using 3 types of K-Fold Cross Validation method and k = 10. Confusion Matrix calculations in Cross 2 get the highest average accuracy value yield of 93.945%. While in the accuracy comparison between KNN and MKKN it was found that KNN has the highest accuracy value of 94.95% and the average accuracy is 93.94%. Using MKKN obtained the highest accuracy of 99.51% and an average accuracy of 99.20%. So it can be said that using MKNN is better 5-7% than KNN.

Lestari, 2017 [3], this research was conducted for the classification of Acceptance of Toyota Astra scholarships by applying knn and naïve bayes algorithms. Then the values of the two algorithms will then be compared. KNN has a higher accuracy value than the accuracy value of the naïve bayes algorithm.

Parvin, 2008 [4], this study compared two algorithms that are one clump namely KNN and MKNN. The development that occurs is performance improvement that is a kind of preprocessing on data training. Add a new value called "validity" to train the sample. Validity takes into account the

stability and reliability value of each data train against its neighbors' data.

Pisarenko, 2021 [5], the proposed method mknn is applied to analyze seismic intensity in two seisgomenic regions. The graph of increased seismic activity can be identified by the MKNN and some of the quantitative statistical characteristics of this graph will be determined and discussed.

Masoodi, 2018 [6], mknn algorithm applied to tracking systems. Research was also conducted on determining the value of K. Simulation of the application of algorithms in case studies has errors in Euclideans below 1 meter for an area of 100 square meters or about 1%. With the same problem, other models are also applied but have a complicated system. Another advantage of MKNN is simplicity.

Of the overall related work that has been described in detail, the author again made sure to apply MKNN to the UKT classification case study. Given the validity of the new value enhanced in MKNN [4] the study applied validation with K-Fold Cross Validation.

III. THEORETICAL FOUNDATION

A. UKT

The Ministry of Education and Culture of the Republic of Indonesia stipulates the Regulation of the Minister on Single Tuition (UKT) which began to be implemented in the academic year 2013/2014. Single tuition is a single tuition fee that is borne by each Prospective New Student based on his or her economic ability. Each State University has a different UKT rate, this is influenced by the regional tentacle and its study program (Permendikbud No.55, 2013).

B. Data Normalization

Data normalization is required in data preprocessing tools used in data mining systems and calculations to narrow the range of data training and distribution of data evenly, in normalization there are several normalization techniques such as min-max normalization, z-score normalization, decimal scalling and sigmoidal normalization. While in this study the normalization of data used is min-max normalization. Min-max normalization is a transformation of the value of an existing data with the smallest range value (min) of 0 and the largest value (max) of 1, in the equation of Min-Max Normalization shown in Equation (1). [7]

$$V' = \frac{V - minA}{maxA - minA} . (new_{maxA} - new_{minA}) + new_{minA} \quad (1)$$

V' = The value of the new data results from normalization
V = Values of data before normalizing newmax
A = Newmin's latest maximum value limit
A = Latest minimum value limit
Max A = maximum value in the column
Min A = minimum value in the column

C. K-Nearest Neighbor Algorithm

The K-NN method is one of the simplest and most intuitive algorithmic techniques in the field of statistical discrimination. The process of calculating euclidean distance in this algorithm is to first create a dataset from training data that the class has known first, the next step is that the dataset of the test data will be classified based on the closest distance of each training data and depending on the k value used. Euclidean equation to perform calculations of the distance between the test data point and the training data point where i is a representation of the attribute value and n dimensions of the attribute shown in the equation (2) [8].

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (2)$$

d(x,y) = Euclidean distance between training data point x and test data y
x_i = Sample training data
y_i = Test data
n = Attribute dimensions

D. Modified K-Nearest Neighbor Algorithm

This MKNN method is the result of a modification with the goal of optimizing the KNN method, in the MKNN method is divided into two processes, namely, first is the validation of training data and then second is the KNN weighting process or weight voting. Classification is performed on test data based on the highest-grade weight in class K training data that has been validated at the closest distance. Unlike KNN, the KNN method does not go through the process of validation of training data. This method of training data validation can maximize training data with high validity and has a close proximity to test data.[4].

E. Validity of data

The process of data validity must be passed in the MKNN process, the validity of each point is calculated according to its nearest neighbor based on k. the validity of this training data is to function to know the number of points with the same table for all the data in the training data. In this process all function values in s depend on the nearest neighbor is equal value or not on training data. 12 Equation of data validity shown in equation (3)[4].

$$validitas(x) = \frac{1}{k}\sum_{i=1}^{k} S(label(x))(Ni(x))) \quad (3)$$

K = number of closest points
Label(x) = class label x
Ni(x) = nearest point class label x
S = 1 when class is the same or worth 0 when the class is not the same

F. Weight voting

In MKNN each data is calculated in weight. Weight voting is useful on training data with high validity and the closest distance to the test data. The first work of weight voting is to calculate the weight of each neighbor, both the validity of the sample of training data multiplied by the Euclidean weight. The multiplication process can reduce weaknesses in each data that has distance problems with weight in the outlier. Weight voting equation shown in equation (4)[4]

$$W(i) = validitas(i) * \frac{1}{d_e + a} \quad (4)$$

W(i) = calculation of weight voting to i
Validitas(i) = validity value to i
d_e = Euclidean distance training data and test data
α = alpha value

G. K-Fold Cross Validation

Cross-validation or rotation estimation is a model validation technique to assess how statistical analysis results will generalize independent data sets. This technique is primarily used to predict models and estimate how accurate a predictive model is when executed in practice. One technique of cross validation is k-fold cross validation, which breaks down data into k parts of the same size. The use of k-fold cross validation to eliminate biases in data. Training and testing are done k times. In the first experiment, the S1 subset was treated as testing data and the other subset was treated as training data, in the second trial the S1 subset, S3,... Sk becomes training data and S2 becomes testing data, and as accurate as it is[9].

H. Confusion Matrix

Confusion matrix is a method that is usually used to perform accuracy calculations on the concept of data mining. Measurement of the performance of a classification system is important. The performance of a classification system describes how well the system classifies data.

Basically, the confusion matrix contains information that compares the results of classifications carried out by the system with the results of classifications that should be.

IV. RESEARCH METHODS

The stages for completing the research are shown in Figure 1.

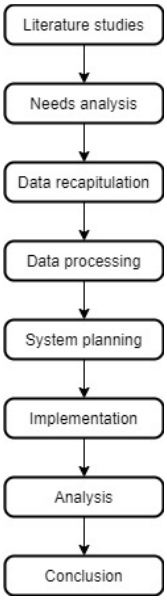


Fig. 1. Methodology flow diagram

In this research begins by collecting then studying the study of the literature used, the next step is to analyze the needs needed, collect then process the data used, design the system to be created, implement the design that has been made, test the system that has been made, analyze the results obtained based on the results of the test, the last step is to make conclusions from the research that has been done.

A. Data sources and research variable

Data is sourced from the database system SIREG UIN SUSKA Riau in 2020 with the amount of data 6000 student data. Variables used in existing weighting are:

- Amount of parents' gross income
- Number of unmarried children
- Number of vehicles
- Number of houses
- Amount of land

While additional variables that will be predicted using MKNN plus 3 variables are:

- Monthly expenses
- Electricity costs per month
- Size of electricity meter

B. Process Flow Diagram

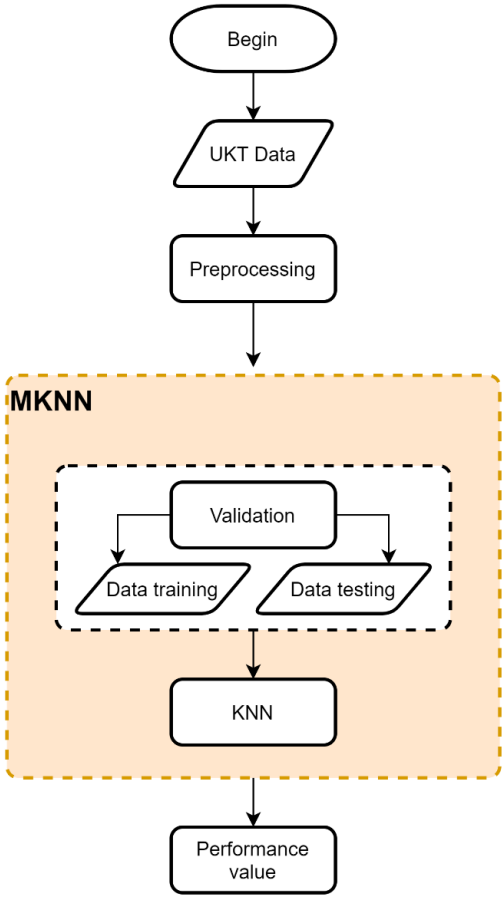


Fig. 2. System process flow diagram

Figure 2. Explain about the stages of the system that is done, ranging from preprocessing, validation, process with the main algorithm, namely MKNN and accuracy assessment.

V. RESULTS AND DISCUSSIONS

1) Preprocessing process

At the preprocessing stage, data transformation is carried out, namely:

- values in parent earnings column.
- value in the monthly electricity cost column.
- values in the field of land
- value on wattage column of electricity meter

2) Validation with Cross Validation, K=4 and with dataset sharing 80:20

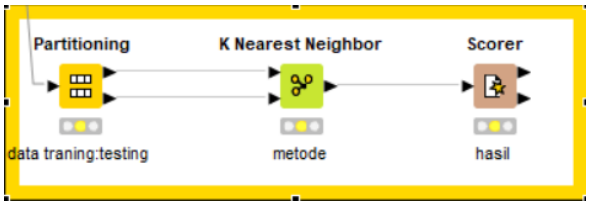


Fig. 3. Validation process using cross validation

3) Validation with K-Fold Cross Validation, K=4 and with dataset sharing 80:20

Fig. 4. Validation process using K-Fold cross validation

Figures 3 and 4 are stages of MKNN because they have been modified with the addition of validation methods in the process of separating data into training and testing.

4) Confusion matrix results

a) With Cross Validation

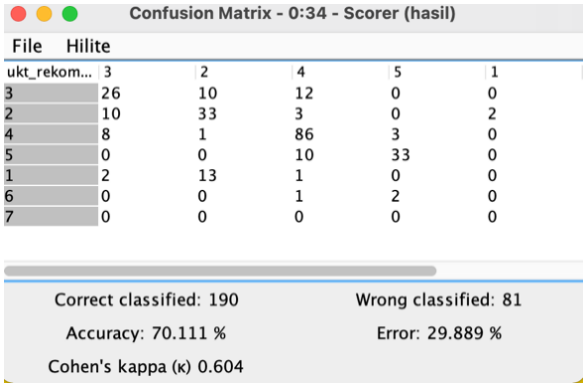


Fig. 5. Accuracy uses confusion matrix for cross validation

b) With K-Fold Cross Validation

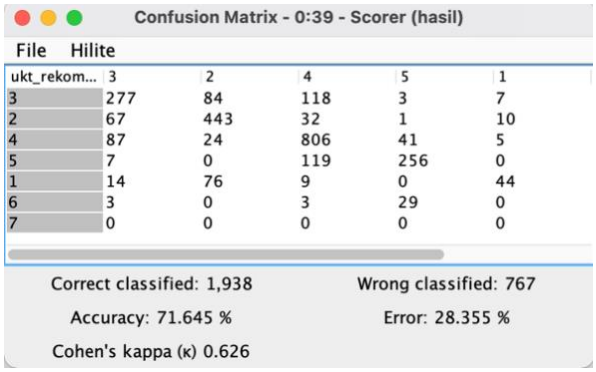


Fig. 6. Accuracy uses confusion matrix for cross validation

Figures 5 and 6, describe the results of the performance of systems designed using different validation processes. K-Fold and Cross Validation have the advantages of each performance, but are varied from K more K-Fold Cross Validation.[10]. MKNN is an optimization algorithm from conventional KNN algorithms by adding validation methods successfully get maximum performance values. MKNN with cross validation managed to get a value of 70.11% and MKNN with K-Fold Cross Validation managed to increase the value to 71.64%..

VI. CONCLUSION

The system that has been built to determine the cost of each student's UKT, SIREG, still needs to be done quality analysis gradually and in-depth. Because based on the analysis using the 2020 tuition fee dataset, the analysis model applied can measure quality based on the classification generated by the SIREG application. By applying the MKNN algorithm, the accuracy value of the decisions generated by SIREG earns 71.64% points. This value is relatively good considering that the dataset used only the dataset in the last 1 year.

REFERENCES

[1] S. Sukamto, Y. Adriyani, and R. Aulia, "Prediksi Kelompok UKT Mahasiswa Menggunakan Algoritma K-Nearest Neighbor," *JUITA J. Inform.*, vol. 8, no. 1, p. 121, 2020, doi: 10.30595/juita.v8i1.6267.

[2] Okfalisa, I. Gazalba, Mustakim, and N. G. I. Reza, "Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification," *Proc. - 2017 2nd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2017*, vol. 2018-Janua, pp. 294–298, 2018, doi: 10.1109/ICITISEE.2017.8285514.

[3] D. R. D. Lestari, "Perbandingan Klasifikasi Beasiswa Toyota Astra Menggunakan K-Nearest Neighbor Classifier dan Nive Bayes Sebagai Penentu Metode Klasifikasi Pada Sistem Pendukung Keputusan Penerimaan Beasiswa Toyota Astra (Studi : Institut Teknologi Sepuluh Nopember)," *Inst. Teknol. Sepuluh Nop.*, pp. 9–10, 2017, [Online]. Available: http://repository.its.ac.id/id/eprint/42490.

[4] H. Parvin, H. Alizadeh, and B. Minaei-bidgoli, "MKNN : Modified K-Nearest Neighbor," *Proc. World Congr. Eng. Comput. Sci. WCECS*, pp. 22–25, 2008.

[5] V. F. Pisarenko and D. V. Pisarenko, "A Modified k-Nearest-Neighbors Method and Its Application to Estimation of Seismic Intensity," *Pure Appl. Geophys.*, Apr. 2021, doi: 10.1007/s00024-021-02717-y.

[6] M. Masoodi, E. A. Sekehravani, and M. Maesoumi, "Rssi-Based Modified K-Nearest Neighbors Algorithm for Indoor Target Tracking," *Far East J. Electron. Commun.*, vol. 18, no. 2, pp. 345–356, 2018, doi: 10.17654/ec018020345.

[7] S. Jain and K. Asawa, "Modeling of emotion elicitation conditions for a cognitive-emotive architecture," *Cogn. Syst. Res.*, vol. 55, no. June, pp. 60–76, 2019, doi: 10.1016/j.cogsys.2018.12.012.

[8] K. Schliep, K. Hechenbichler, and A. Lizee, "Weighted k-Nearest Neighbors," *2016*, vol. 399, p. 15, 2016, [Online]. Available: https://cran.r-project.org/web/packages/kknn/kknn.pdf.

[9] M. Bramer, *Principles of Data Mining*, no. January 2007. 2007.

[10] U. B. Hatta, "JURNAL IPTEKS TERAPAN Research of Applied Science and Education V15.i1 (34-47)," vol. 1, no. March, pp. 34–47, 2021.

BUSINESS SECTION

Invited Speaker :
Arniati, SE., MSi., PhD. AK,. CA,. CPA.

"Annual Personal Income Tax Electronic Reporting for Learning"

POLITEKNIK NEGERI BATAM

Job Satisfaction among Academic Staff in TVET Institutions

1st Mohammad Fardillah Wahi
Mathematics, Science and Computer Department
Politeknik Mukah
Mukah, Sarawak, Malaysia
fardillah@pmu.edu.my

2nd Habsah Mohamad Sabli
Commerce Department
Politeknik Mukah
Mukah, Sarawak, Malaysia
habsah@pmu.edu.my

3rd Sari Lestari Zainal Ridho
Politeknik Negeri Sriwijaya
Palembang, Indonesia
sarilestari@polsri.ac.id

***Abstract*—The scenario of the Covid 19 pandemic and movement control order (MCO) force closing all the higher institutions in Malaysia. Due to this situation, online learning implements by educators at home plus household matter demotivated and lead to a decrease in job satisfaction. Therefore, the purpose of this study to identify the relationships concerning recognition, promotion and also to identify if there exists a differences job satisfaction of male and female among academic staff in Politeknik Sarawak which include Politeknik Kuching Sarawak (PKS), Politeknik Mukah Sarawak (PMU) and lastly Politeknik Metro Betong (PMB). Data was collected from 324 academic staff and analyzed using correlation and Mann-Whitney U test analysis. There is a significant relationship between promotion with job satisfaction compared to recognition. The female consists of (48.1%), and male (51.9%). Only the recognitions have significantly corelated with job satisfaction compare to promotions. Based on the Mann-Whiney U test there is no different of job satisfactions between genders. This outcome of this research can benefit the decision-makers such as governments, institutions, and policymakers to give suitable recognition and promotion to educators to enhance their job satisfaction.**

Keywords—Job Satisfaction, educators, recognition, promotion, Covid 19 pandemic

I. INTRODUCTION

The Covid 19 pandemic, needs urgently changing in educational systems environments which educators and students need to adopt a technological environment for online learning and teaching [1]. A prior finding in [2] studies stated that a low mastery of ICT skills among teachers has caused them incapable to apply ICT in their teaching and learning activities. This can reduce the motivation and job satisfaction among educators which did not have any knowledge and skills toward online learning. Teaching online classes during Covid 19 pandemic has a prominent influence on job satisfaction among the 252 educators in secondary school Dharmapuri District, Tamil Nadu [3]. [4] The factors include environment, remuneration, promotion chances, working conditions, job security, and coworkers were identified to be the six criteria that can affect the staff academic satisfaction level in a previous study of job satisfaction in Bahawalpur Colleges, Pakistan. According to the study's findings, professors reported higher levels of job satisfaction than research assistants. Job satisfaction among educators also influences by demographic factors include age, gender, and length of services [5-7]. Academicians' job satisfaction levels were in moderate level according to a study [8] on the relationship between job satisfactions. Recognition given to individuals or groups of employees can also encourage them to finish their jobs on time or even before the dateline.

Past research found that employers can significantly impact job satisfaction when their people do well in terms of relationship, and quality and quantity of projects or programs. [9] individuals or groups are entitled to receive recognition after executing innovative performances or managed to simplify work processes within an organization. Furthermore, one of the most significant aspects of job satisfaction is promotion. According to a study conducted by [10], academics are more motivated and devoted to performing their jobs and being satisfied with their work if they are offered the option for advancement based on great performance. The importance of job satisfaction among educators leads to this study. Inconsistent findings and lack of research on this matter attract authors to do this research. Hence, this paper is an attempt to identify the relationships concerning recognition, promotion and also to identify if there exists a differences job satisfaction between genders among academic staff in Politeknik Sarawak. The findings can help the institutions, government to implement policies to help educators facing the situations.

II. LITERATURE REVIEW

Workers who are satisfied with their jobs are more motivated, put in more effort, and are more likely to perform well than those who are not [11-12]. Satisfied workers are more productive and will remain longer in the company compare to dissatisfied workers who will be less productive and more inclined to quit [13]. At least three outcomes (retention, attrition, and absence) and at least three major influences (demographic variables, job role-related characteristics, and work experiences) have been investigated [14]. Age, education level, gender, marital status, and tenure are all factors that influence job satisfaction [15-17]. Result survey from the academicians of Khyber Pakhtunkhwa, a

province of Pakistan shows females are more satisfied, age group more than 40 years old with the length of services more than 20 years score shows the highest mean score of job satisfaction. This support by [18], which the age, commitment, and length of services have positive relationships with job satisfaction. Work from home (WFH) during MCO, forces educators to fully adopt information technology without any help or facilities provided [19], emphasis on top management to provided facilitating conditions and system quality which are the factors that motivate academics to use online learning including the provision of computer support and the new IT innovation adoption by the organization. The institutions of higher education in Portugal state that recognition is an important element to determine the levels of satisfaction among lecturers during their time serving in higher education institutions [20]. Recognition of creative ideas may have implications on the organization in the future [21]. [9] claimed that recognition can lead to problem-solving in a creative organization. Recognition is one of the best ways to increase educators' quality of teaching [22]. In terms of education, educators should prepare well beforehand for class sessions by preparing notes, screen projectors, internet, and e-learning whenever applicable. Therefore, recognition can be given to employees in the form of praise or awards, either achievement certificates or rewards, bonuses, and increases in salary as showing that positive values have a good impact on the organization.

Furthermore, opportunities for advancement may arise based on seniority, ability, or both. Seniority is determined by the length of time spent in the organization, but it can also be determined by the date of service, the departments or units served, and the specific types of work performed [23]. Several methods, such as structured career path, occupational restructuring, research development, job postings, and career advancement, can be used to promote worker advancement [24]. Managers, on the other hand, can decide on promotions based on recommendations from the employee's supervisor, records appraisal, experience, company objectives, and educational background [25]. There is a positive relationship between promotion and job satisfaction among Pakistani university teachers [26]. [27] discovered a relationship between employee job satisfaction and promotion. Once recruited as an educator, young teachers and researchers may benefit from the change in career path. Promotion is a key issue that people consider before entering any profession. In the situation where avenues exist for individuals to progress in a job by following the academic ladder or ranks, they will be ever willing to enter and stay.

III. Research Design and Methodology

To collect data for this study, a survey was used. The questionnaire instrument was divided into two sections. The first section focused on gathering demographic information from respondents, includes age, gender, marital status, salary, education level, and length of service. The second section includes 21 items related to recognition, promotions, and job satisfaction. For these 21 statements, this study used a Likert scale of 1 for "strongly agree" to 5 for "strongly disagree". The questionnaire was developed in English adopt from job satisfaction [28, 29] recognition, [22] and promotion from [10]. The questionnaires distributed to 500 lecturers at Politeknik Sarawak (Politeknik Kuching Sarawak (PKS), Politeknik Mukah Sarawak (PMU), and Politeknik Metro Betong Sarawak (PMB) in August 2020 via email and WhatsApp's between colleagues. This study followed Krajie and Morgan's [30] proposed sampling size to obtain the target respondents for the survey. SPSS Version 26 was used for this purpose. The reliability tests on the identified variables for this study were found to be acceptable and reliable as the Cronbach's alpha values for all the variables are 0.614 above 0.60 [31].

IV. Results and Discussion

TABLE I. Respondents' Demographic Profiles

Characteristics		N	(%)
Sex	Male	168	57.9
	Female	156	48.1
Status	Married	164	50.6
	Single	160	49.4
Age	26 until 35 years	57	17.6
	36 until 45 years	177	54.6
	Above 46 years	71	21.9
Salary	Below Rm2500.00	26	5.9
	RM2501 until RM3500.00	82	8.0
	RM3501 until RM4500.00	106	25.3
	above RM4501.00	110	32.7
Education Level	Diploma	16	4.9
	Degree	188	58.0
	Master	108	33.3
	PHD	12	3.70
Length of Svc	Less than 5 years	43	13.3
	6 to 10 years	180	55.6
	11 until 15 years	30	9.3
	Above 16 years	71	21.9

Table 1 shows the frequency and percentage distribution of respondents by gender. A total of 168 respondents (51.9%) are male while 156 (48.1%) are female. Table I shows the frequency and percentage distribution of marital status. A total of 164 people (50.6%) respondents is married while 160 respondents (49.9%) are still single. Overall, the respondents are aged between 26 until 35 years old, with a total of 177 respondents (54.6%). This is followed by 57 respondents (17.6%) aged less than 25 years. The number of respondents under 36 until 45 years is 71 people (21.9%) and the lowest percentage is for respondents aged above 46 years (6%). The respondents are mostly freshly graduating when they started working in the Sarawak Politeknik since they opened around the year 1990 to 2005. About 26 people (8.0%) has salaries below RM2500.00 and 82 people (25.3%) are paid between

RM2501 until RM3500.00. Additional percentages showed a total of 106 people (32.7%) have salaries between RM3501 until RM4500.00 and 26 people (8.0%) earn salaries above RM4501.00. Salary increments are based on grade and experience. The results from the study showed that 43 respondents (13.3%) have a length of services of less than 5 years, followed by 6 to 10 years tenure with 180 respondents (55.6%). Furthermore, 30 respondents (9.3%) have 11 to 15 years of length of service and 70 people (21.6%) have worked more than 16 years. This is followed by 9 people (8.9%) that have 9 to 12 years of work experience and 3 people (3%) that have taught for more than 12 years. lastly, a total of 188 respondents (58.0%) has a degree, as this is the minimum requirement for teaching posts. This is followed by 108 respondents (33.3%), who have master's degrees. 16 respondents (4.9%) are diploma holders, while12 respondent (3.70%) has a Ph.D.

TABLE II. PEARSON CORRELATIONS FOR IDENTIFIED FACTORS AND JOB SATISFACTION

Variables	Significant value (p)	Pearson Correlation (rs)
Recognition	0.000	0.250
Promotion	0.093	0.093

[a.] Significant value with $p \leq 0.05$

The analysis from Pearson Correlation coefficient test, this study found that only recognition has significant positive relationships with job satisfaction. with a correlation coefficient of 0.050 and its p-value is smaller than 0.05; and promotions have not significantly correlate of 0.093 and its p-value is higher than 0.05.

TABLE III. JOB SATISFACTION BETWEEN GENDERS (MANN- WHITNEY U TEST)

	Job satisfaction	
Mann- Whitney U		12.962.000
Wilcoxon W		27.158.000
Z		-1.69
Asymp. Sig. (2-tailed)		.866
Sum Of rank	Male	27158.00
	Female	25492.00
Mean Rank	Male	161.65
	Female	163.41

Referring to Table III, because the two gender groups were not normally distributed, the Mann Whitney U test was used to determine whether there is a difference in job satisfaction between male and female academic staff. The result indicates assymp. sig (2 tailed) 0.866 (unsignificant there no difference of job satisfaction between genders.

V. CONCLUSION

The primary objective is to identify the relationships between recognition, promotion with job satisfaction among academic staff in Politeknik Sarawak and to identify if there exists a difference job satisfaction of male and female among academic staff in Politeknik Sarawak. Firstly, the analysis shows that there is a significant relationship between the levels of job satisfaction with recognition among academic staff in Politeknik as the correlation value is semi-weak and positive. This parallel with finding from [22], recognition can be given to employees in the form of praise or awards, either achievement certificates or rewards, bonuses, and increases in salary as showing that positive values have a good impact on the organization. Secondly, the test was conducted to examine the Mann-Whitney U test result showed that. The result indicates there no difference of job satisfaction between genders. This finding support by [32] whereby female staff is more satisfied than their opposite. Further research should be conducted due to the limitations in this research, for instance due to movement control order (MCO) due to Covid 19 pandemic. Hence these limitations can be elaborate in carried out the next research agenda, as the effort to enhance the contributions to the knowledge and policy implementations.

ACKNOWLEDGMENT

We acknowledge the Politeknik Mukah, Sarawak, Malaysia and Politeknik Negeri Sriwijaya, Indonesia for motivation support and provision of enabling working environment. Also appreciate the conference organizer (ABEC 2021) Politeknik Caltex Riau, Indonesia given opportunity for us to join the conference.

REFERENCES

[1] Tsai, C.-C., Chai, C. S., Wong, B. K. S., Hong, H.-Y., & Tan, S. C. (2013). Positioning design epistemology and its applications in education technology. Educational Technology & Society, 16, 81–90.

[2] Awaatif, A. & Norizan, E. (2011). Effect of Courseware (courseware) With Concept Map on Student Achievement. Asia Pacific Journal of Educators and Education, 26 (1), 51-70.

[3] Gopinath, R. (2020 b), "The Influence of Demographic Factors on the Job Involvement, Organizational Commitment and Job Satisfaction of Academic Leaders in the Tamil Nadu Universities", European Journal of Molecular & Clinical Medicine, Vol. 7, No. 3, pp. 5056-5067.

[4] Saba, I. (2011). Measuring the Job Satisfaction Level of the Academic Staff in Bahawalpur Colleges, 1(1).

[5] Hagedorn, L. S. (2000). What contributes to job satisfaction among faculty and staff: New directions for institutional research. San Francisco, CA: Jossey-Bass.

[6] Santhapparaj, A.S. and Alam, S.S. (2005), "Job satisfaction among academic staff in private universities in Malaysia", Journal of Social Sciences, Vol. 1 No. 2, pp. 72-6.

[7] Bilimoria, D., Perry, S.R., Liang, X., Stoller, E.P., Higgins, P. and Taylor, C. (2006), "How do female and male faculty members construct job satisfaction? The roles of perceived institutional leadership and mentoring and their mediating processes", Journal of Technology Transfer, Vol. 31, pp. 355-65

[8] Toker, B. (2011). Job satisfaction of academic staff: an empirical study on Turkey. Quality Assurance in Education, 19(2), 156–169. https://doi.org/10.1108/09684881111125050

[9] S. Chatterjee, S. Chakraborty, S. Sarker, S. Sarker, F.Y. Lau, Examining the success factors for mobile work in healthcare: a deductive study, Decision Support Systems 46 (3), 2009, pp. 620–633.

[10] Baloch, Q.B., (2009) "Effects of Job Satisfaction on Employees Motivation & Turn over Intentions". Journal of Managerial Sciences Volume II, Number I.

[11] Heryanto, H. (2019). The effect of work motivation and work environment on performance with satisfaction as intervening variables education personnel rectorate Andalas University. Archives of Business Research, 7(2), 103–120

[12] Rubel, M. R. B., Kee, D. M. H., &Rimi, N. N. (2020). High-performance work practices and medical professionals' work

outcomes: the mediating effect of perceived organizational support. Journal of Advances in Management Research, 0972–7981.

[13] Macerich R, Ives C, McGreal R. Measuring use and creation of open educational resources in higher education. Int Rev Res Open Distance Learn. 2013;14(4):90-103. doi:10.1002/

[14] Perrachione, B. A., Rosser, V. J., & Petersen, G. J. (2008). Why do they stay? Elementary teachers' perceptions of job satisfaction and retention. The Professional Educator, 32(2), 1.

[15] Ronen and Pines (2008). Gender differences in engineers' burnout Type: Equal Opportunities International, 2 7: (8).

[16] Kabungaidze, T. Mahlatshana, N (2013). The Impact of Job Satisfaction and Some Demographic Variables on Employee Turnover Intentions. International Journal of Business Administration. 4(1).

[17] Hayat, Y., Mughal, Abdul, D., & Busari, H. (2015). Personal characteristics, job satisfaction and turnover intention of lecturers of advance learning institutions of Khyber Pakhtunkhwa province Pakistan. International Researchers, 4(4).

[18] Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988). "SERVQUAL: a multi-item scale for measuring consumer perceptions of the service quality", Journal of Retailing, Vol. 64, No. 1, pp. 12- 40

[19] Sormunen, K. (2020). From inclusive practices to personal strategies. Nordic Studies in Science Education, 16(2), 234–234.

[20] Machado-Taylor and J. B. Ferreira, "What factors of satisfaction and motivation are affecting the development of the academic career in Portuguese higher education institutions?" vol. 45, no. 1, pp. 33–44, 2011.

[21] Kassim, E. S., Zamzuri, N. H., Shahrom, M., & Nasir, H. M. (2010). E-learning use among academics: Motivations and cognitive style. Proceeding of the 3rd International Conference on Information and Communication Technology for the Moslem World: ICT Connecting Cultures, ICT4M 2010, July 2019.

[22] Nijhof, A., Krabbendam, K. and Looise, J.K. (2002), "Innovation through exemptions: building upon the existing creativity of employees", Tec novation, Vol. 22 No. 11, pp. 675-83

[23] Jackson C. The impact of Ip on TQM and staff motivation in general practice. Int J Heal Care Qual Assur 13/2. 2000;13(2):53-54.

[24] Mouawad M, Kleiner B (1996). New developments in customer service training. Managing Service Quality, 6(2): 49-56

[25] Gruenfeld LW, Mac Eachron AE. Relationship between age, socioeconomic status, and field independence. Percept Mot Skills. 1975 Oct;41(2):449-50. doi: 10.2466/pms.1975.41.2.449. PMID: 1187300.

[26] Rubel, M. R. B., Kee, D. M. H., &Rimi, N. N. (2020). High-performance work practices and medical professionals' work outcomes: the mediating effect of perceived organizational support. Journal of Advances in Management Research, 0972–7981.

[27] Tessema MT, Soeters JL (2006). Challenges and prospects of HRM in developing countries: Testing the HRM-performance link in Eritrean civil service. Int. J. Hum. Res. Manage., 17(1): 86-105.

[28] Dahler, S. A., & Luyten, H. (2006). Faculty of Behavioural Sciences Towards Academic Staff Satisfaction of Working Condition of Master' s Thesis, (August).

[29] Mustapha, N., & Yu Ghee, W. (2013). Examining Faculty Workload as a Single Antecedent of Job Satisfaction among Higher Public Education Staff in Kelantan, Malaysia. Business and Management Horizons, 1(1), 10–16.

[30] Krejcie, R. V., & Morgan, D. W. (1970). Determining Sample Size for Research Activities. Educational and Psychological Measurement, 30(3), 607–610.

[31] I. Helga, "Chapter 3 - Demonstration Chapter 3 – Demonstration," no. 2003, 2002, pp. 37–57.

[32] G. O. Oriarewo, K. C. Agbim, and Z. B. Owutuamor, "Job rewards as correlates of job satisfaction : Empirical evidence from the Nigerian banking sector," Int. J. Eng. Sci., vol. 2, no. 8, pp. 62–68, 2013.

Development of To-Do List and Monetary Management System

1st Diana anak Ringgau
Department of Electrical Engineering
Politeknik Mukah
Mukah, Sarawak, Malaysia
dianaringgau@gmail.com

2nd Chen Wong Keong
Department of Mechanical Engineering
Politeknik Mukah
Mukah, Sarawak, Malaysia
chenwk2012@yahoo.com

3rd Bong Siaw Wee
Department of Electrical Engineering
Politeknik Mukah
Mukah, Sarawak, Malaysia
shaweibong2016@gmail.com

***Abstract*—Currently, most of the available "To-Do List" Applications on Android or iOS are just consisted of a "To-Do list" and integrated with an electronic calendar. These applications are just suitable for individual use for keeping alert daily "To-Do List" and its functionality is limited to update the status of each task. The other limitation of these applications is the "To-Do List" could not be easily shared and updated if involved other group members. Therefore, a new "Group To-Do List" plus Monetary Management System application was proposed. The proposed application can be shared and updated by group members of interest in communicating and monitoring "To Do List' and fund management. This application was developed based on Rapid Application Development Method (RAD) by understanding the natural practices of task management and quantities of tasks taken on by team members. Therefore, this app can be expanded and customized to integrate other attractive functions based on user requests. This new application is suitable for family members, project group members, or even individuals from all walks of life who are interested in managing a total integrated "To-Do List" plus finance management system. Some of the important features of this application are its ability to identify tasks that need to be completed on daily basis with indications of which tasks need financial support, and which tasks could generate income. An interactive table and bar charts have summarized all of these tasks with details of expenditure/income and are integrated into an electronic calendar. The status of these tasks could be updated by group members with ease on individual tablets/handphones/computers/laptops. This application can be expanded/customized to integrate other attractive functions based on user requests.**

Keywords—group, to-do list, finance management system

I. INTRODUCTION

In the modern era, technology devices have become mobile portable, and networked to the point that they have become pervasive in everyday life [1]. The knowledge expansion system is quickly changing due to the rapid development of mobile devices based on next-generation information technology (IT) convergence technology, and information acquiring is progressing actively and swiftly as well [2-4]. From notepads to digital to-do list apps, the world has changed drastically [5]. Fully-enabled, powerful to-do list apps not only offer quality assistance for task management but improve your productivity too. There are lots of to-do list apps available. Some of these apps are free, some are fee-based, but the primary objective of all of them is to offer classic task scheduling assistance to help users to manage and finish the tasks with ease and on time. Statistics suggest that there are 17 million to-do list software applications and websites available but not all of them will be useful for users. There are tools available on how to organize one's time and prioritize work, however, many people worry about whether they are prioritizing and meeting their many tasks effectively. Thus, To-do lists are more than simply a list that gives an impression of the amount of work to be done; they are also a resource for consultation, linking to project objects or task objects, and indicating the status of tasks. Moreover, since the beginning of the year 2020, the world tackles the highly infectious respiratory disease called Covid-19. Many industries had declared work-from-home to aid with the containment of the coronavirus and to prevent its spread from overburdening healthcare systems. The invisibility of home productivity and its infrastructure is supported by traditional productivity measurement methods, thus an app is needed to manage and monitor non-face-to-face or work-from-home team members.

Currently, most of the available to-do list apps on Android or iOS are just consisted of to-do list apps and integrated with electronic calendars [6, 7]. These applications are just suitable for individual use for keeping alert daily to-do list apps and their functionality is limited to update the status of each task. The other limitation of these applications is the to-do list apps could not be easily shared and updated if involved other group members. The technology acceptance model is used to predict an individual's intention to use the to-do list application. We believe that developing an application with user engagement would better fit user requirements and capabilities. Users can choose the best to-do list software to suit their needs by answering questions shown in Fig. 1. It may be easy to find an answer to a specific question, a solution to the overarching problem may be more elusive. Based on Fig. 1, mobile apps have an important role because they allow the user to assign tasks personally or by a group [8]. User and group members are also able to use e-mail, electronic instance messages, and database management systems for task completion's notification method. Measuring Perceived usefulness (U) will involve respondent's agreement that using the application with such features should make them easier to manage their mini-project. Moreover, the respondent believed that data-driven apps can bring many opportunities to support the need of users by allowing others to add additional details for better decision-making. The data or task can be recorded via a cloud database system [9] making it easier for the user to access the

app. Thus, using the application should enhance their effectiveness on the tasks. Usefulness (U) is the user's subjective probability that utilizing the app would improve his or her tasks performance within an organizational environment.

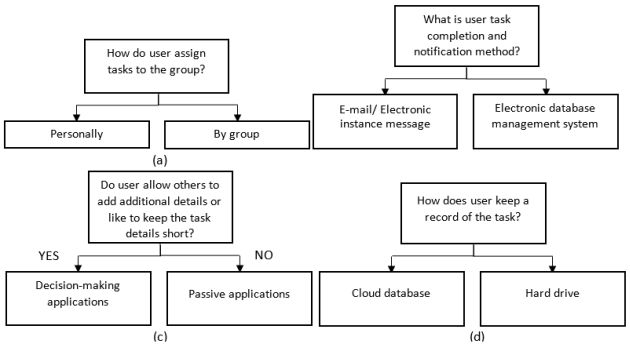


Fig. 1. Identifying user's most basic task management

A brief online training session is compulsory for the proposed app's user. The training sessions should demonstrate how to use the application. This is to give the user some opportunity to interact with the application and to facilitate understanding of the app. Measuring one of the Perceived ease of use (EOU)'s elements will involve the respondent's agreement that the training sessions could lead a user to operate the application easily. Ease of use (EOU) is the degree to which the user expects the proposed app to be effort-free. However, (U) is affected by (EOU). Both EOU and U can predict user's attitude (A), which is defined as the user's evaluation of the desirability of utilizing the app. (A) and (U) have an impact on a user's intention to utilize the proposed to-do list application.

Therefore, a new "Group To-Do List" plus Monetary Management System or iDo$$ application was proposed. A systematic approach to address these issues and present a solution – A cloud-based database system, with a custom made graphical user interface (GUI), that any group member can access from their mobile phone at any time, powered by AppSheet (AppSheet Seattle, Washington, USA), Google Forms® and Google Sheets® (Google, Mountain View, California, USA) that gets instantly updated upon data entry and store data on Google Drive® in a retrievable format for data analysis [10]. The proposed application can be shared and updated by group members of interest in communicating and monitoring "To Do List' and fund management. This application is developed with user-friendly application development software and the data is stored in free google drive. This new application is suitable for family members, project group members, or even individuals from all walks of life who are interested in managing a total integrated "To-Do List" plus finance management system.

Some of the important features of this application are its ability to identify tasks that need to be completed on daily basis with indications of which tasks need financial support, and which tasks could generate income. An interactive table and bar charts have summarized all of these tasks with details of expenditure/income and are integrated into an electronic calendar. The status of these tasks could be updated by group members with ease on an individual mobile smartphone. A group member can also seek consultation from their group leader or manager based on the indications of their task's status. Besides that, the apps also allowing other group members to add additional details for better decision-making.

By understanding the natural practices of task management and the quantities of tasks taken on by team members, the proposed application should be expanded and customized to integrate other attractive functions based on user requests. This proposal gives an easy understanding of how to create a database management system with AppSheet and Google Sheets® in a step-wise manner. An app called iDO$$ was fabricated to manage and monitor non-face-to-face or work from home team members. The app was implemented based on Fig. 1 and the prediction of a user's intention to use the proposed app. A survey was conducted among students at Mukah Polytechnic to identify the effectiveness of the app.

II. METHOD

In application development, basically can be divided into three approaches, there are waterfall, rapid application development (RAD), and agile approach. In this application, rapid application development (RAD) is selected as the main methodology because of its flexibility to produce a working version of the application as quickly as possible, and then to continuously finetune the application. RAD is willing to change or upgrading the application that is suited to the needs of the user. Following is the detailed explanation of the rapid application development approach provided.

A new Google account was created for this purpose. Google spreadsheets® were used to design the data tables and were linked to AppSheet® software to generate the graphical user interface of the mobile app of the database. Appearance and features of the App were designed through options provided by the AppSheet® without writing codes using computer language shown in Fig. 2. Once the app was ready, the email address and password were shared among the authorized team members and the mobile app was installed into their phones.

Fig. 2. Three important steps for To-Do List and Monetary Management system App development

This section will contain a step-by-step guide to create a custom-made mobile app to enter personal/ group data on a mobile phone using Google services and the AppSheet platform.

Step 1: Establish a Google Account:

It is advisable to have a separate Google account for the management of the database management system so that the password can be shared among the members of your team who are certified to handle data. A Google account could be created by visiting http://account.google .com/.

Step 2: Creating a Google spreadsheet based on the necessary data fields:

Once the Google account is created, log on to it and access the Google Apps grid on the top right-hand corner of the screen and select Google drive from it shown in Fig. 3. Once

the Google Drive page opens, click on the +New button on the top left-hand corner shown in Fig. 4 and select Google Sheets and then, blank spreadsheet. Rename the Google spreadsheet [11, 12]. Then, Design the data tables which are needed to be included in the database within the Google spreadsheet. Use the "+" button on the bottom left-hand corner in the Google spreadsheet window to create several data sheets (data tables) within the user's spreadsheet. Designing these tables is the key step to create the database as well as the Mobile App of the database. The data sheets/tables you will be designing this way will be the data fields of the App you will be creating later

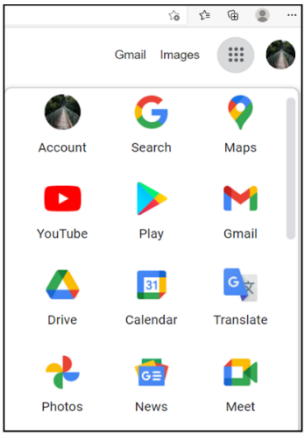


Fig. 3. Opening a Google apps

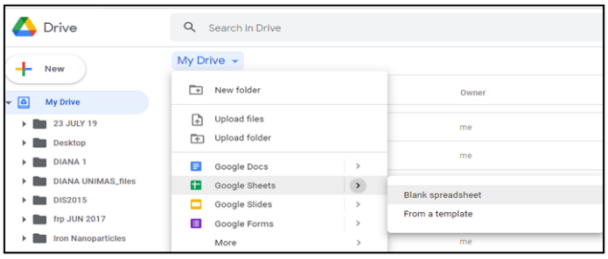
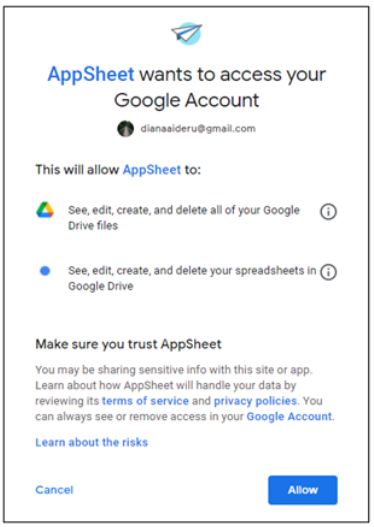


Fig. 4. Creating a Google spreadsheet

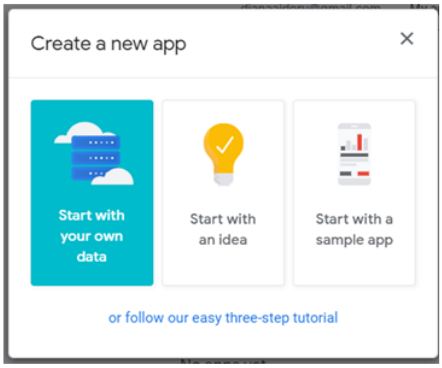
Step 3: Link the Google Spreadsheet to AppSheet to create the Graphical User Interface of the mobile app.

Upon finalizing the Google spreadsheet, it should then be linked to AppSheet software to create the Graphical user interface of the mobile app. Go to https://www.appsheet.com/ to get the free version of the AppSheet software shown in Fig. 5. Then, click the button "Start for free".

Then, Select Google sheets and forms. Then Sign in to the newly created Google account (As mentioned in step 1) with the e-mail address and password. Then, Fig. 5(a) shows giving access to AppSheet to use the Google account by clicking the "Allow" button. Then, AppSheet Window will appear and select "Create a New App". Once you click that, three options will be provided, and select "With Your Own data" shown in Fig. 5(b)



(a)



(b)

Fig. 5. Link the Google Spreadsheet to AppSheet to create the Graphical User Interface of the mobile app.

Step 4: Modify the GUI of the Mobile App using AppSheet.

Then, Make use of the various options available such as "Data", "UX" to build up the GUI of the newly created app. Users can use the "Data" option, to add new tables to the database.

- Select "Data"
- Click "+ Add New Table"
- Select the Google spreadsheet from the drop-down menu in the "Source ID "field.
- Select the Table/Sheet from the "Worksheet name/Qualifier" drop-down menu.
- Click "Add this table"

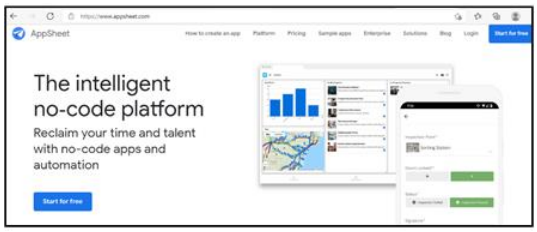


Fig. 6. Homepage AppSheet Software

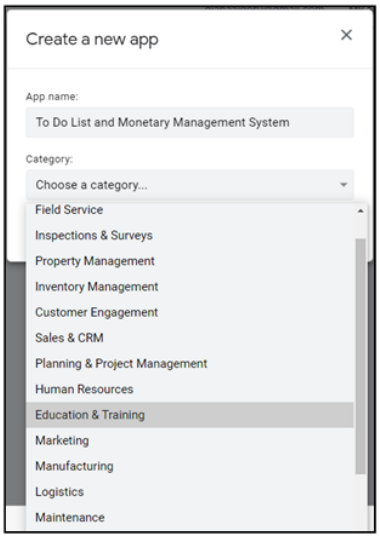


Fig. 7. Naming app and select related category.

Once the user adds a new table, select the table and click "Updates, Adds, Deletes" from the "Are updates allowed" menu as in Fig. 9. Always click "Save" after each change which user has done in the AppSheet.

The AppSheet web-based platform allows the users to change the external appearance of the app by adding color themes, custom app logos, launch images, and background images [13]. User can explore the "UX", "Data" and "Behavior" options on AppSheet to further customize the app to user own taste. App formulas, Slices, References between Tables, and incorporating them into the app would further enhance the capabilities of the app, but it is beyond the scope of this proposal. The AppSheet web-based platform will open up in a new tab on the user web browser with a basic mobile app interface displayed on the right-hand side shown in Fig. 8.

The present app interface in Fig. 8 shows the "Form View" of the app. It is based on the Google Spreadsheet that the user-created earlier. Users can click the "+" golden button to add new data. The added data records of individual patients will appear as in Figure 8. Users can see the data by clicking the data records name in the form view of the app.

This form view shows "Name" – as a short text field, "Amount of income" as a numeric field, "Expenses" –as a multiple choice. Once user-entered data, click "Save". When the user clicks the "Plan" tab, he can enter data into the table. Upon populating the spreadsheet, which is saved in Google Drive, the user can access it anytime for data analysis. The Google spreadsheet can be connected to SPSS through Microsoft Excel.

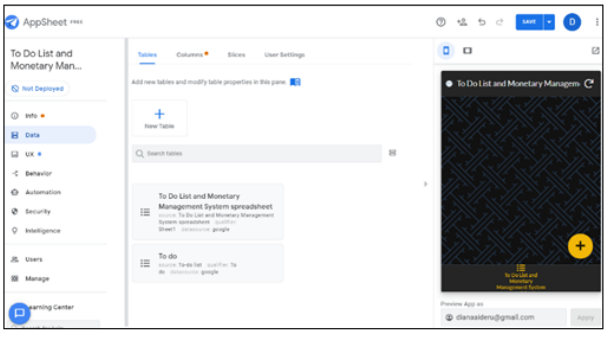


Fig. 8. GUI of the new app

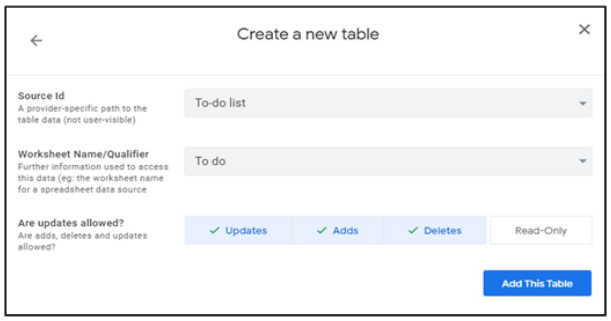


Fig. 9. Create a new table

Step 5: Finalizing and launching the mobile app among authorized team members.

Once the app is ready to be shared with the user team, head over to the "Users" section of the AppSheet web platform and enter the email addresses of the team members and send invites. When the team members receive the invites to their corresponding email accounts, they should click the link on the invite. This will prompt them to install the AppSheet App on their phones from the App Store (iOS) or Play Store (Android) on their devices. Once it is installed, open the AppSheet app on their phones and log in using their own Google email address or the Google email address that was created earlier to create the database. The team members will now be prompted to install your database apps shortcut to the home screen of their mobile device. Maintaining the database in the cloud can be considered a limitation as it requires an internet connection always, although it is what makes this approach more secure and available across multiple devices at the same time.

The mobile app was launched and introduced to students of Mukah Polytechnic who took Innovation and Invention Club (MPU24021). In this course, team members were expected to demonstrate communication skills, leadership, teamwork, problem-solving skills as well as organize activities such as planning by writing papers, appointing committees, conducting meetings, implementing activities, writing activities, and financial reports. However, due to the pandemic, team members had to experience with non–face-to-face coordination for project or task management. We have provided online training sessions for iDO$$'s users. All users had attended the training sessions demonstrating the use of the application. Survey questions on the effectiveness of the app were constructed and distributed to all users at the end of the academic June 2020 session. The survey was distributed among iDO$$' users. The respondents consisted of seven (7) mini-project groups. Each group consisted of a team leader and four (4) group members.

III. Results And Analysis

This app enabled the team members to add, edit and view information. Fig. 10 shows the appearance of the front page on the smartphone.

Fig. 10. Short description and explanation of the developed App at the About Function

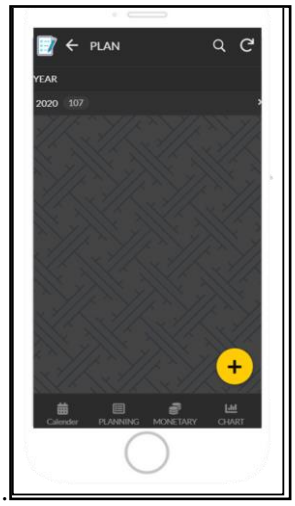


Fig. 11. Main user interface with fout main buttons at the bottom of the app, there are calender, planning, monetary, and chart

Fig 11 shows the user interface of this application. At the bottom of the application, there are four main buttons. The first button is the calendar button which integrates all the events/ tasks into one integrated and synchronized calendar as shown in Fig. 12. The second button is the planning button which was used to input all the important tasks into this cloud database by more than one user in different users' handphones/ personal computers as shown in Fig. 13-16. The third button is the monetary button used to summarize all the income and expenses in an interactive interface as shown in Fig. 17. The fourth button summarized all the tasks and related to their income/ expenses in an interactive graph as shown in Figure 18.

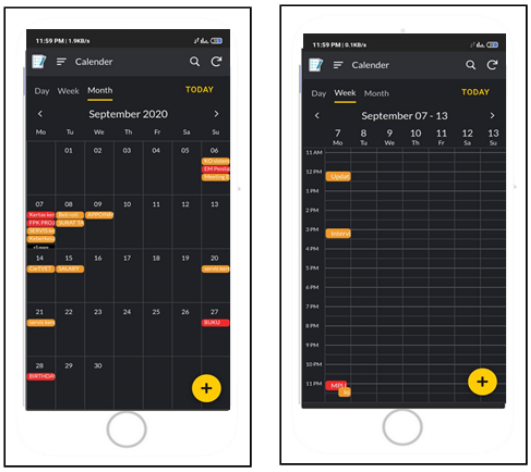


Fig. 12. An integrated data shown in monthly and weekly calender (Calender button)

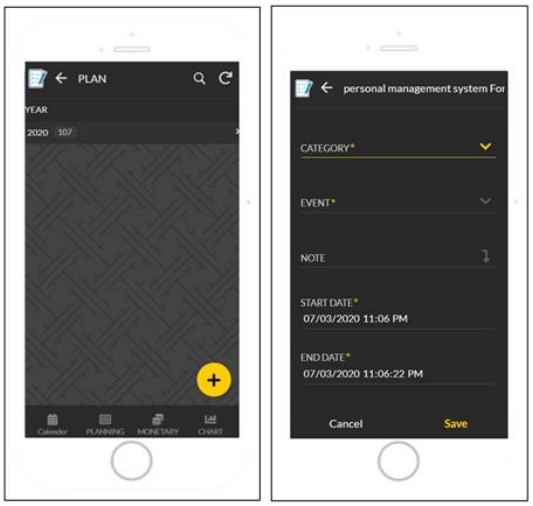


Fig. 13. Press on the "+" button to add a new task to the "To-Do List" (Planning button)

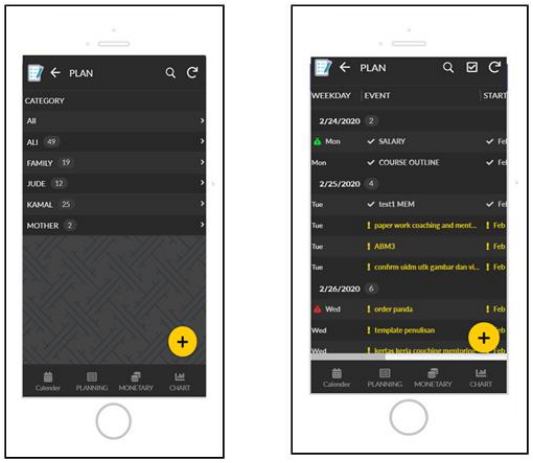


Fig. 14. Users could view several lists in each individual and its detail (Planning button)

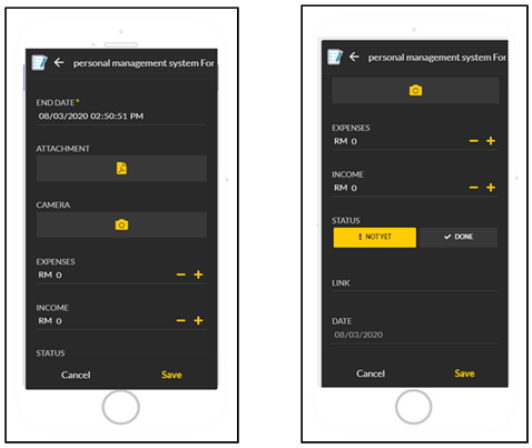


Fig. 15. Users could insert files, pictures, expenses, income, and links to the list (Planning button)

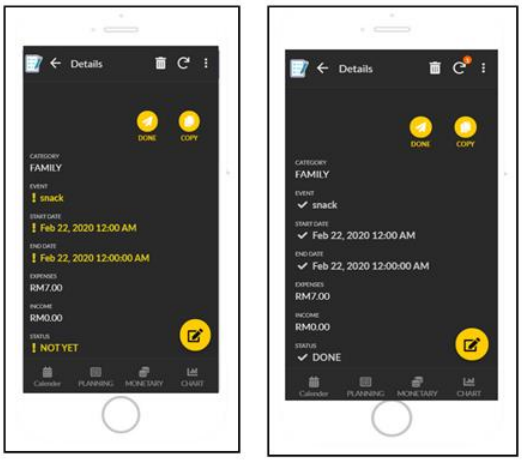


Fig. 16. Users could update the status of the list by simply press the "DONE" button (form ! To √) (Planning button)

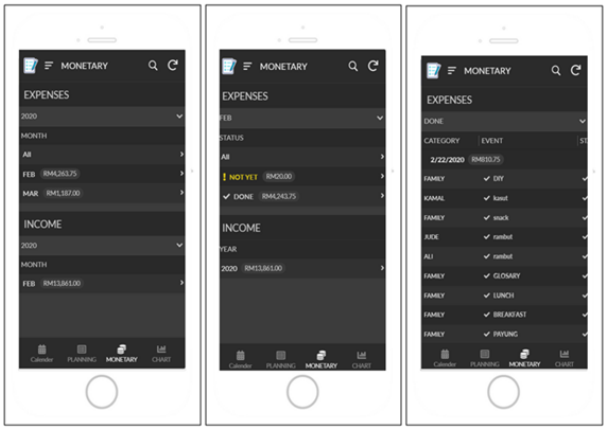


Fig. 17. Monetary button shows the expenses and incomes in detail (Monetary button)

Fig. 18. An interactive chart to display the details of expenses and incomes (Chart button)

The survey was distributed among iDO$$' users at Mukah Polytechnic. The respondents consisted of seven (7) managers or group leader that managed seven (7) mini project groups. Each group consisted of five (5) members. The survey questions are divided into two (2) parts, namely part I: To-do List application (iDO$$)'s Function and part II: To-do List application (iDO$$)'s Features. All sections included 21 Likert -scale questions, in which the alternative responses were (1) to strongly disagree, (2) to disagree, (3) to be neutral, (4) to agree, and (5) to strongly agree. Descriptive statistics are used to measure the mean and to summarize the data collection. The rating scale questionnaire was designed to facilitate the data evaluation in the analysis stages. Thus, the score helps in identifying the level of functionality and the quality of features of the application. Furthermore, total average score mean values and ranking were analyzed using Landell's scale (1997). Landell's classification scale, described low as around 1.0 – 2.33, moderate around 2.34 – 3.68, whereas high is classified around 3.69 – 5.00. Therefore, Landell's classification scale was used to determine the scale of all the variables from the rankings such as the functionality and the quality of features of the application.

TABLE I. DATA ANALYSIS AND FINDINGS FOR THE FUNCTIONALITY OF THE APPLICATION

No	items	Mean	Interpretation
Part I: To-do List (iDO$$)'s Function			
1	The app is capable to record tasks that need to be carried out easily by users as well as group managers.	4.55	Good/Agree
2	Built-in electronic calendar assist me to identify the appropriate date and period to complete the task	4.36	Good/Agree
3	The application displays information on assignments given by the project manager to group members.	4.45	Good/Agree
4	The application notifies me of assigned tasks given by the manager to me and my members.	4.45	Good/Agree
5	This app helps me to assess priority tasks.	4.45	Good/Agree
6	The app is capable to monitor and display the status of tasks that have	4.36	Good/Agree

	not been completed by my members even without having to meet face-to-face.		
7	This app is capable to evaluate the division of tasks is done fairly and transparently among members.	4.45	Good/Agree
8	The scope of tasks on this application can be added according to my group's requirements/ objectives	4.36	Good/Agree
9	This app helps me to update the status of my assignment in real-time to the club advisor/ chairman.	4.55	Good/Agree
10	The app shares information on names, emails, and mobile phone numbers that can be contacted by each group member.	4.45	Good/Agree
11	This application helps me to ascertain the problems that will be faced if a task that has priority is not solved first.	4.18	Good/Agree
12	This application is capable to calculate total allocation received by the highest committees as well as budget allocations in each group	4.55	Good/Agree
13	This app helps me to manage and record the financial status of my group quickly.	4.36	Good/Agree
14	This app will notify me of financial problems that will be faced if one task is overspending.	4.36	Good/Agree
15	This app helps me to organize a program.	4.36	Good/Agree
16	This application can record tasks that need to be carried out by users as well as group managers	4.45	Good/Agree
Overall average Mean Part I		**4.42**	**Good/Agree**

The highest mean score was 4.55 given by respondents for questions number 1, 9, and 12. The respondents are very aware that the To-do List app may display the completion status directly to the project manager. The status can be shared among other group members so that the next task or link can be performed quickly. The app was also capable to calculate total allocation received by the highest committees as well as budget allocations in each group. However, question number 11 gain the lowest mean score, 4.18. Certain respondents were not aware that the app can be able to help them in identifying an encountered problem if a priority task is not completed first. Questions number 2, 8, and 13 got the same mean score of 4.36, where these three items indicate that it is likely that the respondents do not yet have practical experiences with the application or are not well-informed about the app thoroughly and in-depth. This situation may have occurred due to the absence of a user manual, although all respondents had attended brief training sessions demonstrating the use of the application.

TABLE II. DATA ANALYSIS AND FINDINGS FOR THE FEATURES OF THE APPLICATION

No	items	Mean	Interpretation
Part II: To-do List app (iDO$$)' Features			
1	I am not charged a fee for using this app.	4.64	Good/Agree
2	This app can operate using android and ios systems.	4.55	Good/Agree
3	This application can be used on any device such as smartphones and PCs, laptops, and tablets	4.45	Good/Agree
4	This application can be used easily (user-friendly) without the need for any expert help.	4.45	Good/Agree
5	This application is beneficial to manage a project	4.55	Good/Agree
	Overall average Mean Part II	**4.53**	**Good/Agree**

Based on Table 2, the highest mean score was 4.64 given by respondents for questions number 1. Marketability of iDO$$ App to clients, this new application is suitable to family members, project group members, student study groups, or even individuals from all walks of life who are interested in managing a total integrated "To-Do List" plus finance management system in a group with minimal or no cost at all. Questions number 3 and 4 got the same mean score of 4.45. This innovation project is a new platform to gather all the related users to contribute "To-Do List" and "Monetary Management". Users can easily upload and share the latest task with other users or group members. Users could access this app easily through a variety of devices, such as smartphones, tablets, personal digital assistants (PDAs), i-pads, mobile phones, and others. This app can be achieved anywhere and at any time without limit. By using this app can reduce the use of raw materials, cost savings, and be environmentally friendly. The app developers could receive feedback from time to time from users for further improvement of this app. However, some of the respondents may not be well-informed about the app. All respondents were between 18 to 21 years old and not familiar with nor used any To-do list application to manage mini-project efficiently.

The To-do list app was helpful in industries with work-from-home policies to contain the spread of COVID-19. The app created home productivity visible to team leaders and other group members. The management and non-face-to-face monitoring became more transparent. For example, the app provides real-time notification of the task and budgets that have been used and recorded for future audits. Applications have also been developed to allow a group member to stay in touch with colleagues. However, users involved in sharing data across the app should be able to handle group members' data with care.

IV. CONCLUSION

Based on Table I and Table II, the overall average mean for the app's functionality was 4.42 and the quality of features was 4.53 respectively, which shows that the mobile app has been tested as an effective solution for non-face-to-face projects or task management apps among group members. The app was useful to support the needs of users by allowing others to add additional details for better decision-making or a resource for consultation. The application was used for real-time online monitoring of activities planning such as appointing committees, task delegation, conducting meetings, budget/expenses updates and tasks progress. This app-enabled the team members to add, edit and view data from these tables. Telephone calls could be taken and group members' locations could be accessed through the app. Since the database is accessible through phones, group members' past tasks could be reviewed and new records can be entered. These spreadsheets could be linked to SPSS software to analyze data for research purposes. Even though this provides adequate confidentiality and security to data, all the authorized team

members who share the app have unrestricted access to add, edit or remove data.

REFERENCES

[1] S. Caballé, F. Xhafa, and L. Barolli, "Using mobile devices to support online collaborative learning," Mobile information systems, vol. 6, no. 1, pp. 27-47, 2010.

[2] Y. Chen and D.Z. Wang, "Knowledge expansion over probabilistic knowledge bases," in Proceedings of the 2014 ACM SIGMOD international conference on Management of data. 2014.

[3] M. Craglia, M. F. Goodchild, A. Annoni, G. Camara, M. F. Gould, W. Kuhn, D. Mark, L. Masser, D. Maguire and S. Liang, "Next-generation digital earth," 2008.

[4] M. S. Hossain, C. Xu, Y. Li, A. K. Pathan, J. Bilbao, W. Zeng, and A. El-Saddik, "Impact of Next-Generation Mobile Technologies on IoT-Cloud Convergence," IEEE Communications Magazine, vol. 55, no. 1, pp. 18-19, 2017.

[5] Miller, M., Cloud computing: Web-based applications that change the way you work and collaborate online2008: Que publishing.

[6] E. Ramadani, A. Memeti, F. Imeri, N. Sejfuli-Ramadani and F. Idrizi, "Social media-based app organizing daily events," in 2017 6th Mediterranean Conference on Embedded Computing (MECO). 2017. IEEE.

[7] M. S. Ahmad, N. E. Musa, R. Nadarajah, R. Hassan and N. E. Othman, "Comparison between android and iOS Operating System in terms of security," in 2013 8th International Conference on Information Technology in Asia (CITA). 2013. IEEE.

[8] N. Yorke-Smith, S. Saadati, K. L. Myers and D. N. Morley, "The design of a proactive personal agent for task management," International Journal on Artificial Intelligence Tools, vol. 21, no. 01, pp. 1250004, 2012.

[9] X. Cheng, L. Fang, L. Yang and S. Cui, "Mobile big data: The fuel for data-driven wireless," IEEE Internet of things Journal, vol. 4, no. 5, pp. 1489-1516, 2017.

[10] C. Ilacqua, C., H. Cronstrom, and J. Richardson, Learning Qlik Sense®: The Official Guide2015: Packt Publishing Ltd.

[11] M. Miller, Using Google Spreadsheets (Digital Short Cut)2002: Pearson Education.

[12] I. Lamont, Google Drive and Docs in 30 Minutes: The unofficial guide to the new Google Drive, Docs, Sheets & Slides2015: i30 Media Corporation.

[13] A. J. Quinn and B.B. Bederson, Appsheet: Efficient use of web workers to support decision making.

Hospital Resource Management Interoperability for Pandemic Management: Research Development

1st Riky Perdana
Economic and Business Faculty
Padjadjaran University
Indonesia
rikyperdana@umri.ac.id

2nd Dwi Kartini
Economic and Business Faculty
Padjadjaran University
Indonesia
d.kartini@unpad.ac.id

3rd Yudi Azis
Economic and Business Faculty
Padjadjaran University
Indonesia
yudi.azis@unpad.ac.id

4th Ummi Kaltum
Economic and Business Faculty
Padjadjaran University
Indonesia
umi.kaltum@unpad.ac.id

***Abstract*—Pandemic is a medical disaster that affects almost all countries in the world. The front line for handling this pandemic is none other than the hospital as a service provider and the government as the coordinator. Despite the maximum efforts of each party, gaps are still found between the capacity of medical facilities and the large number of requests for patients. One of the shortcomings observed by the researchers was that there was no information exchange protocol that allowed the government and all participating hospitals to exchange information on the actual availability of resources, which allowed all patients to be directed quickly to health facilities that still had service capacity. This study attempts to design a hospital resource information exchange protocol that can be used to handle the distribution of pandemic patients more quickly and accurately. The research method used is the Expert System Development Life Cycle. The testing conducted with White Box Testing Method which done by the developer. The results of this study are in the form of 3 things, including: 1. Cross-hospital interoperability network topology that can be mediated by the government or certain agents, 2. Standard data structures that can be transmitted via a web service, 3. Standard Operating Procedures that can guide all participants to perform exchange the following resource data with patient distribution.**

Keyword—interoperability, healthcare management, pandemic

I. INTRODUCTION

A pandemic is a situation where an infectious disease spreads on a large scale and is no longer able to be contained effectively. An epidemic condition whose spread can no longer be controlled at the regional level will be declared a pandemic by the international agency WHO (World Health Organization). Handling a pandemic situation involving many countries certainly requires handling efforts that involve cooperation between countries as well. For every country affected by the pandemic, the respective governments must coordinate closely with regard to the count of victims, their status, their location, available resources, the number of patients, patients currently being hospitalized, patients who have been cured, patients who have died, and future predictions of the situation. [14]

Apart from the role of the government and health service providers, mitigating the pandemic situation also requires community participation to minimize the impact of spread through social distancing. According to the latest reports on COVID-19 at the time of writing, almost every country in the world has been affected by the disease. Even though it originated in the city of Wuhan in China, the disease managed to escape from observation due to the lack of preparation and vigilance of the government and the community. Other countries that have failed to prevent the entry of the disease are also responsible for the outbreak of the pandemic within their respective countries, which can be seen from reports that the epicenter of the pandemic can move from one country to another even though they are not geographically adjacent.

In managing this pandemic situation, the hospital can rely on the role of the management information system to be able to identify and manage patient information, whether infected with a pandemic disease or not and organize special resources for handling the pandemic to ensure that all patients received can still be served with the capacity there is. The majority of the information systems available in the field can vary according to the needs and capabilities of the hospital in developing or acquiring them from the party when they are [15].

The information managed by the hospital management information system is very diverse, starting from patient data, employee data, doctors and nurses, data on supporting resources such as laboratory and radiology, drug data and medical devices. The diversity of these resources makes the design of the information system that is implemented also varies between hospitals. This minimizes the chance for some hospitals to be able to establish a coherent suite of data to deal with a pandemic situation. According to the latest report at the time of writing, many hospitals were found that were no longer able to handle the number of requests for pandemic patient care, while several other hospitals were still unemployed pandemic management resources.

This indicates that it is important to have continuous cooperation between hospitals that can be mediated by cross-hospital information system technology that allows them to exchange information about the availability of resources. In the case of Indonesia, efforts to coordinate the handling of pandemic patients have been attempted through the pandemic management task force program by collecting resource data from each participating hospital and manually directing the patient to a referral hospital. Although this step can be considered good, it is not optimal for a pandemic situation where the number of victims and their distribution has

exceeded the management capacity of the appointed task force.

So it is important for researchers to find a method with integrated technology media to facilitate each participating hospital to be able to independently and collectively manage information on the availability of medical resources for handling pandemic situations. This study aims to produce an interoperability design for health service resource information that can be applied generically to general health services and specifically for handling pandemic situations.

II. Literature Review

Several previous studies have tried to involve the integration of information on hospital resources, especially for cases of handling pandemic situations in South Africa. Research sponsored by Dramowski (2020) shows that there are three types of interventions that can be carried out by the government and hospitals, such as: Administrative Intervention, Engineering Intervention, and Protective Equipment Intervention. These findings can help governments and hospitals to be able to implement emergency regulations for handling pandemic situations that can minimize the risk of bad treatment and seek adequate resources by sharing information across hospitals. What is considered still lacking from this study is the absence of concrete guidelines on how best to integrate information across hospitals, which allows them to exchange information about the availability of resources and readiness for handling future pandemic patients.

An example of coordinating the availability of hospital resources for handling a pandemic in Indonesia is carried out by the COVID-19 Handling Task Force which actively collects information from all hospitals designated as referral for pandemic patient services and coordinates patient absorption rates for each of these hospitals, and transferring patient referrals to hospitals that still have service capacity. This effort is also assisted by BPJS Kesehatan (Badan Pelayanan Jaminan Kesehatan) the Social Security Administering Body which is a state-owned company by opening up information on the availability of inpatient beds for patients who need inpatient services. As far as this writing is written, these two instruments are still the mainstay of the government and Indonesian society for handling pandemic victims. Although there have been continued efforts of the government to develop the coordination of hospital information systems, but still ad-hoc or situational which among participating hospitals have not fully implement the integration of information resources independently and autonomous (can not exchange information directly).

III. Research Method

In order to develop a system which may be used for cross-coordinating resources between bodies of organization, particularly in healthcare service, researcher should look for a method which satisfy following characteristics:

1. Allow the development of a system that accommodates cross-organizational coordination
2. Help developer reduce the risk of mismanagement or misconduct by human resources
3. Ease the phases from planning, prototyping, testing and up to release.

According to prior references and consideration, the researcher decided to observe the Expert System Development Life Cycle Method proposed by Agarwal (1990) with following detailed stages:

1. Assessment
   a. Defining the problem
   b. Defining the general purpose and scope of the system
   c. Verifying suitability Expert systems withproblems
2. Knowledge Acquisition
   a. Determine sources of knowledge
   b. Get knowledge related to issues to be discussed
   c. Conduct interviews withexperts
3. Design
   a. Build design concepts
   b. Determine development strategies
   c. Choose development platforms
4. Testing
   a. Perform system simulations
   b. Conduct conformity testing
5. Documentation
   a. Make information structure
   b. guides Make workflow guides
6. Maintenance
   a. Makesystem maintenance manual

A. Assesment

1) Defining the Problem

Setiaji (2014) research findings identified several problems and challenges in implementing healthcare information system in Indonesia which are considerably similar to the observed problems by the researcher on the field, like the following:

- There is a service capacity gap between hospitals designated to serve pandemic patients
- There are hospitals whose patient care requests exceed available capacity
- There are hospitals that has a capacity that exceeds the level of patient service requests
- Patients who fail to find a hospital that can serve do not know how to find an alternative hospital
- Each hospital management cannot know in real time the service capacity of other hospitals
- The task force is formed by the government is overwhelmed in coordinating supply and demand between hospitals and patients

2) Defining research objectives and scope

- Generating a workflow design for data resource integration across hospitals
- Generating an information technology infrastructure design to support integration
- Generating a standard data format for the exchange of hospital resource information

3) Verifying the suitability of the expert system with the problem

According to the Ministry of Health of the Republic of Indonesia (2019), the availability of a database of positive victim information, the number of deaths, patients who are being treated, patients who are under monitoring and quarantine, along with the hospital and all its resources are the means that It is important to support the government in responding to a pandemic and reducing the number of victims. This research is not to ensure the adequacy of information on health facilities for the response to a pandemic, but rather to seek open access to information resources across hospitals to support a better distribution of patient handling burdens.

B. Knowledge Acquisition

1) Determining sources of knowledge

To obtain appropriate study material for this research, it is important for researchers to observe the parties involved in handling a pandemic situation, from strategic to technical levels, such as:

- Previous research on the integration of resource information across hospitals, especially regarding pandemic handling
- The Ministry of Health and its task force and policies.
- Several referral hospitals that are directly involved in the technical handling of pandemic patients.
- Some open source software designs specifically for data integration across organizations, especially in health services

2) Gaining knowledge related to the discussion

- Conducting a literature review to study previous findings about interoperability across hospitals
- Conducting questions and answers with the COVID-19 handling task force and reviewing their SOP (Standard Operational Procedure) documents
- Conducting questions and answers with referral hospital officers directly handle patient administration
- Conduct in-depth review of open source software that provides data integration

3) Conducting interviews with experts

Researchers have tried to communicate with several colleagues of the Ministry of Health and from the results of the conversation it was found that every day at certain hours the officers collect and collect data on the number of victims per day which they can get from hospitals, community health centers, and BPJS Health (Social Security Administering Body). This information is collected to be able to calculate the cumulative number of victims such as how many victims today, how many positive patients today, how many deaths so far, how many patients died per day, how many patients are being hospitalized / quarantined, how many the number of patients who do self-isolation, how many and how many people are being monitored, the identities of all patients who are known to be positive for infection, and family data of patients who may be infected.

Provided with all the raw data, the government can process the available quantitative data to produce information such as the growth trend in the number of positive victims for future predictions, data on the geographic distribution of positive COVID-19 patients, tracking the potential distribution both within the neighborhood and between regions, knowing the ratio of differences between the number of patients recovered and who died, and many other forms of conclusions that can be revealed through data and statistics.

Provided with descriptive and analytical information, the government can be more confident in making decisions such as which regions and community groups should receive special attention, which hospitals should concentrate on aid funds, medical personnel, and all supporting facilities, and other forms of policies that allow reduction. the trend of increasing casualties and reducing the likelihood of positive patients dying. As far as this research has been written, the task force established by the government must manually collect data along with its updates. Current information cannot be obtained in real time.

Researchers have tried to communicate with administrative officers who are responsible for organizing hospital resources, both medical personnel such as doctors and nurses, as well as supporting facilities such as rooms, medical equipment, equipment, medicines, and personal protective equipment. In coordination with the government, the referral hospital must always record and make daily reports about patients, doctors and nurses on duty, resources used, estimated costs that must be billed to patients, insurance, or the government, and all other information that can be used by government and outsiders to channel resource assistance.

For hospitals that have reached their maximum capacity, the government will be notified that the hospital can no longer accommodate additional referral patients and it is hoped that the task force can help these new patients to be directed to other hospitals that still have service capacity, according to the referral hospital list data. that the government owns.

IV. Research Findings

A. Design

1) Building design concepts

To be able to make a good system design, it is important for researchers to identify who are the parties who participate and interact with the system, including:

- Hospitals. In this integrated system design, the hospital will always update the data related to the list of resources they have, the patients that are being treated, and the number and types of resources that are being used for patient care. In addition, the status of the resources that are being used, are in need of repair, are over / under capacity, and others.
- Government. The task force for handling COVID-19 which was formed by the Ministry of Health was given the role and responsibility to provide a

media for collecting data that could be used to coordinate the distribution of patient handling burdens.

Regarding infrastructure design, it is important for all participants and government agencies to agree on what integration platform will be used together in the long term, such as process flows, integration data formats and connection channels that can be used for data transmission. For example, the data format can use JSON / XML, data transmission can use REST / SOAP. Because there are differences in infrastructure design between hospitals, each participant should be free to manage their own information technology infrastructure. The topology of this integration design is:

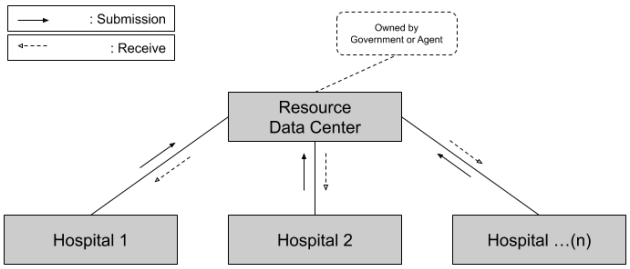


Fig. 1. System Integration Topology

In the chart above we can see an information communication topology design involving all parties who will participate in the integration of cross-hospital data resources. The first party is the government represented by the Ministry of Health or a specific agency that can be assigned responsibility for managing the hospital resource data center. In this case the government must provide a complete server facility with a database where all resource information will be stored. The government can determine its own technology specifications to be used while the choice of technology can support data integration objectives, such as servers, operating systems, databases, programming materials, software, and network providers.

At the bottom of the chart is the interaction scheme of each participating hospital. Each hospital must provide its own technology infrastructure whose specifications can be determined independently according to the hospital's capacity, while being able to support collective data integration. The line connecting each hospital with the data center indicates that each participating hospital must send data updates to the data center to then be managed and informed to all other participating hospitals. With this topology, each participant can know the actual situation of the resources at other hospitals.

Interoperability of information across organizations requires standardization of data formats and structures to be exchanged [10]. This study does not aim to dictate to the government, hospitals, or other parties but rather recommends interoperability designs produced through scientific methods. To be able to produce an ideal information structure, it is important for researchers to identify all elements of information that must be involved regarding resources for handling pandemic patients.

The general hospital resource information is described in the following table:

TABLE I. HOSPITAL RESOURCE INFORMATION ELEMENT INFORMATION

Element	Description
Origin	Is information that indicates the source of resources such as the name of the hospital, hospital code, hospital address, and other additional information
Type	Is information that indicates the general type / category of resources such as rooms, medical devices (ventilator, oxygen cylinder), etc.
Name	Is information that shows the unique identity of these resources such as the name of the room, the name of the bed / ward, the name of the medical device in question, including specifications and brands.
Amount	Is information that shows the quantity of resources in units the smallest number such as the number of beds / wards available and the number of medical devices in the name of the resource.
Status	Is information that shows the factual and actual status of these resources in the hospital such as' used ',' available ',' under repair ',' damaged ', or other status

Provided with knowledge of elements For this information the researcher can compile a standard data structure for exchange in JSON format as follows:

```
{
  "origin": 160014, # Unique code of participating hospitals
  "type": 2, # Reference: 1. Bed / ward, 2. Medical Devices, 3. Other
  "name": "Ventilator", # Can add information on specifications and brands
  "amount": 14, # The smallest number of units of the resource referred to
  "status": 2 # Ref: 1. Available, 2. Used, 3 Repair, 4. Broken, 5. Others
}
```

The design of the above data structure requires consensus between the government as the provider of the data center and all hospital participants, which means that the data design can be changed according to agreement. After the data structure is mutually agreed upon and becomes the standard for exchange, all participants must always follow the standard to maintain consistency in the data that will be exchanged [13].

After designing the infrastructure topology and data structure standards for exchange, the next researcher's task is to design a process flow that can guide all parties involved to carry out interoperability of information resources across hospitals [2]. As previously mentioned, the two groups of parties involved in this interoperability effort are the government as the liaison and temporary coordinator of the hospital as the owner and manager of their respective resources. In this study, researchers will use a Business

Process Management framework to help researchers formulate and design a good process flow.

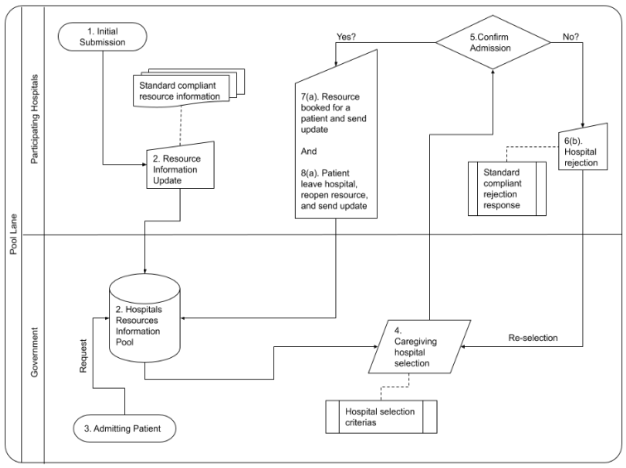


Fig. 2. Interoperability Process Flow of Hospital Information Resources (Riky, 2020)

In the chart above is shown a 'pool' consisting of two 'lanes' represented by the government and the hospital. The details of each step in the process include:

- Initial Submission

Is the initial process in which each participating hospital sends information on the list of resources they have to a data center whose formatting is adjusted to the standards set by consensus beforehand.

- Resource Information Update

It is a process whereby each participating hospital updates their list of resource information to a data center whose format is in accordance with established standards. Partially, the hospital can fill in this information manually, but it is better if it is automated by each information system to ensure that the actuality of data can be maintained while avoiding / reducing the potential for human error.

- Hospital Resource Information Pooling

Is the stage where a central server provided by the government passively collects all initiative submissions or updates sent by each participating hospital. At this stage the data center can send data updating information to all other participating hospitals both passively and actively through mutually agreed data distribution methods, so that the hospital can also know the actual situation of the availability of other hospitals. The government can design its own strategy and visualization software for all the collected data, as well as for each participating hospital.

- Request for Patient Admission

Is the stage where there are newly infected positive patients who need assistance from the government to which hospital these patients should be taken. This patient admission can come from the patient initiative, the health facility, the local level, or from a mass clinical trial conducted by the government on a community group.

- Selection of referral hospital

It is the stage where representatives of the government make the selection of the hospital to which the patient will be directed. The task force established by the government can determine its own screening criteria and priority scale for the best ranking. For example, the selection of a referral hospital for patients who are geographically closest to the patient to be directed. Or choose a hospital based on the level of patient emergency, where more critical patients can be directed to a referral hospital with more complete means of care. The results of the screening and sorting can be used as a basis for consideration by the task force to then send notification of requests for new patient care to several alternative hospitals that the patient will go to.

- Response to service requests

Through a mutually agreed data communication method between the data center system and the hospital, participating hospitals will passively receive notification requests for patient services from the data center and are asked to respond to the patient's admission approval. The organization can independently decide whether to accept these patients or not according to certain criteria that they can determine independently. The branch of the decision is to accept or reject.

- Rejection by the hospital

The hospital can make response decisions with rejection of patients on the basis of certain considerations by sending back request rejection information which can be attached with details such as the date, person in charge, and reasons for rejection. The quick response from the task force coordinator will allow the task force to immediately select an alternative hospital that can handle it.

- Hospital acceptance

Participating hospitals decide to accept requests for new patient care. With this the hospital is ready to receive these patients and updates the list of availability information to be sent to the data center, where the cycle of information dissemination from the data center to all referral hospitals will be repeated and allows each referral hospital to know the changes in the situation of the latest collective resource availability.

Similar to the design of information technology infrastructure and standard data format structures, the design of this process flow is open and can be adjusted according to mutual agreement between the government as the coordinator and each hospital as a participant.

2) Determining a Platform Development Strategy

In order to realize the previously mentioned design, it is important for researchers to determine the information interoperability system development strategy. In this case the researcher tries to make an example of a server-side application that can be used to test the concepts that have been designed. The programming language used by researchers is Javascript, the engine used is NodeJS 18.4.1, the database used is MongoDB, and the interface used to access the application is REST [11]. This server-side application is deliberately designed without a graphical display because it only acts as a web service. The system logic is made simple where there are several lines of code with the CRUD (Create, Read, Update, Delete) function to perform database operations.

B. Simulation and Testing

1) Perform system simulation

Using a simple prototype that the researcher developed independently, the researcher succeeded in exchanging data between the client device and the server via the REST interface. In this case, both government-owned servers as data centers and each hospital-owned server are web services that are connected to each other and free each participant to choose their own platform while the API (Application Programming Interface) used remains REST [12]. Researchers have successfully tested data transmission from the hospital server to the data center and the data center sends data updating information to all other hospital servers that are connected.

2) Conducting a suitability test

The researcher brought the server application prototype to a private hospital in Riau-Indonesia Province to be tested. provided with inpatient room data, medical devices, and other facilities at their disposal, the researchers documented it in JSON (Javascript Object Notation) format and tried to exchange data through the web service of each server. As a result, resource data can be collected on a data center server and distributed on several sample servers on other computers. With this, researchers believe that the design of the hospital resource information interoperability system for handling pandemic situations has been successfully simulated and is feasible to implement.

C. Documentation

1) Making information structure guidelines

Based on the design points described earlier, the researcher has drafted a textual guide on information infrastructure design, information structure standards for interoperability, information exchange process flow, and the design of a web service application example to simulate the data exchange. The total number of manual pages is 12 pages, accompanied by descriptions and illustrated narration.

D. Maintenance

Maintenance of the hospital's resource information interoperability system for handling pandemic situations is a shared responsibility to be designed, developed, managed, and maintained. Carrying out their respective roles in accordance with the agreed rules will ensure that the interoperability system is viable and reliable in the long term and will play a major role in a critical period of a pandemic. Meanwhile, the design, development, management and maintenance of each hospital's web service is the responsibility of each participating hospital, and this study does not attempt to dictate how each hospital's best practices are to do so.

V. Conclusion

Starting from the identification of the gap in the availability of resources in several hospitals observed in Indonesia, the researcher concludes that the main problem lies in the lack of opportunities for hospitals to exchange information on the availability of resources in real-time.

There have been several previous studies where the researchers put forward arguments related to the problem of coordinating the availability of hospital resources for handling the pandemic along with managerial solutions. While the researchers here identify that the problem lies in the coordination facilities between hospitals, especially those based on information technology.

Based on the researcher's study of various development research methods, the Expert System Development Life Cycle is the most capable method to help researchers answer these problems. The result of this study is a concept of a hospital resource information exchange protocol that can cover the gap between the availability of service resources and the demand for services at the hospitals involved in the exchange protocol.

The results of this study have been tested for feasibility through the White Box Testing method where the test subjects are software and application developers where the concept can be run. The test results show that the concept can be implemented if each participant in the system agrees on what data structure will be exchanged.

The shortcoming of this research is the lack of Black Box Testing which involves participants from the hospital or the government who can act as a hub in this data exchange concept.

References

[1] Agarwal, Ritu, and Mohan Tanniru. Systems Development Life-cycle For Expert Systems. Journal of Knowledge Based System, 1990.

[2] Begoyan, A. 2007. An overview of interoperability standards for electronic health records. USA: society for design and process science.

[3] Conger, Joan. Collaborative Electronic Resource Management: From Acquisitions to Assessment. Sterling Publisher, 2004.

[4] Dramowski, A., and M. Zunza. South African healthcare workers and COVID-19: A shared responsibility to protect a precious and limited resource. South African Medical Journal, 2020.

[5] Kumar, Amit. Big media healthcare data processing in cloud: a collaborative resource management perspective. Journal of Cluster Computing, 2017.

[6] Manring, Nancy J. Collaborative Resource Management: Organizational Benefits and Individual Costs. Journal of Administration and Society, 1998.

[7] Schruijer, Sandra. Schruijer, Sandra. 2020. The Dynamics of Interorganizational Collaborative Relationships: Introduction. Journal of Administrative Science, 2020.

[8] Hari Setiaji, Fathul Wahid. 2014. Masalah Implementasi Sistem Informasi Rumah Sakit: Pelajaran dari Beberapa Proyek. Seminar Nasional Informatika Medis (SNIMed). 2015.

[9] Tsinakis, M., and D. G. Katehakis. Information Infrastructure For An Integrated Healthcare Services Network Proceedings. 2000. IEEE EMBS International Conference on Information Technology Applications in Biomedicine, 2000.

[10] Oemig, F., & Snelick, R. 2016. Healthcare interoperability standards compliance handbook. Cham: Springer international publishing AG.

[11] Kasthurirathne, S. N., Mamlin, B., Kumara, H., Grieve, G., & Biondich, P. 2015. Enabling better interoperability for healthcare: lessons in developing a standards based application programing interface for electronic medical record systems. Journal of medical systems, 39(11), 1-8.

[12] Braunstein, M. L. 2018. Healthcare in the age of interoperability: the promise of fast healthcare interoperability resources. IEEE pulse, 9(6), 24-27.

[13] Kaur, K., & Rani, R. 2015. Managing data in healthcare information systems: many models, one solution. Computer, 48(3), 52-59.

[14] Li, J., & Ray, P. 2010. Applications of E-Health for pandemic management. In The 12th IEEE International Conference on e-Health Networking, Applications and Services (pp. 391-398). IEEE.

[15] Baveja, A., Kapoor, A., & Melamed, B. 2020. Stopping Covid-19: A pandemic-management service value chain approach. Annals of Operations Research, 1.

The Relationship of Intellectual Capital Disclosure on Financial Statements of Plantation Companies on Tax Paying Compliance

1st Raya Puspita Sari Hasibuan
Accounting Department
Politeknik Negeri Medan
Medan, Indonesia
rayahasibuan@polmed.ac.id

2nd Anita Putri
Accounting Department
Politeknik Negeri Medan
Medan, Indonesia
anitaputri48@polmed.ac.id

3rd William Caesar
Accounting Department
Politeknik Negeri Medan
Medan, Indonesia
williamhrp@students.polmed.ac.id

***Abstract*—The long-term goal of implementing this research is to determine the relationship between the independent variable and the dependent variable and the existence of moderating variables that can strengthen or weaken the relationship between these variables. Where the independent variables in this study are the application of accounting conservatism, firm size, firm age, concentration of ownership and independent commissioners, while the dependent variable is tax compliance and the moderating variable is intellectual capital disclosure. The object of this research is all plantation companies whose financial statements have been audited from 2016-2020 and are listed on the Indonesia Stock Exchange. The analytical tool used in this research is SmartPLS. This research was conducted by downloading the financial statements of the plantation companies on the IDX website, namely www.idx.co.id. The number of companies that were sampled in this study were 75 plantation companies according to predetermined criteria. The result of this study is that all the independent variables contained in this study have no effect on the dependent variable, meaning that there is a discrepancy between the theory previously described and the results of this study. While the moderating variable in this study was able to strengthen the relationship between the dependent variable and the independent variable. Thus the results of this study need to be evaluated better with respect to the use of appropriate variables, suitability of the use of analytical tools, as well as the object under study.**

***Keywords*—Intellectual Capital, Tax Compliance, Plantation Companies.**

I. INTRODUCTION

The rise of globalization in the world of accounting and a business-based economy is a process of transformation, capitalization and transfer from a science to something material. For example, humans, existing assets are human bodies, while intellectual capital is the mindset of humans, so the human obligation is to pay taxes in order to contribute to the welfare of society. So that if the human body consists of a body but does not have the ability to think brilliantly, it will certainly weaken the human and of course will not be able to operate anything properly. What this means is that every human being is a resource who is given the ability to have intellectual capital which is considered an asset that should increase awareness in the process of paying taxes.

One of the problems that must be reported in the financial statements to increase the usability of the company is the disclosure of intellectual capital (Chrisdianto, 2009). Intellectual capital is used for all non-tangible or non-physical assets and resources of an organization, which includes processes, innovation capacity, patterns, and invisible knowledge of its members and collaboration networks and organizational relationships (Cut Zurnali, 2003). 2008).

Basically, the emergence of the phenomenon of intellectual capital today is caused by the rapid development of accounting science in relation to intangible assets. The mindset of accountants is also constantly evolving. Where it can be understood that an intangible asset is an asset that is not visible but its usefulness can be felt. Tangible assets have a very special nature, namely non-monetary, cannot be seen but can be felt well, even though they do not have a physical form but can be used to produce something of value, profits, increase the company's capital, and of course can be used by various parties to each other's interests.

Furthermore, there are several things that are regulated in detail in PSAK 19, including expenses that cannot be capitalized, the acquisition price of these intangible assets is determined using historical costs, besides that research and development costs are always needed which can be understood under intellectual capital This is certainly greatly influenced by the times, especially the development of accounting science. A company should always pay attention to various provisions that have been set by accounting standards when its assets are included in intangible assets. Where it can be understood that the intangible assets owned should indeed have a clear disclosure of the value owned by these assets.

Research (Ulum, Ghozali and Chariri, 2016), states that company performance has a positive effect on intellectual capital. Medium (Kuryanto, 2007), states that disclosure of intellectual capital has a negative effect on company

performance. Research (Nugroho, 2012) which examines the effect of firm size, firm age, independent commissioners, leverage, and concentration of ownership on intellectual capital disclosure. The sample of this study uses 2010 manufacturing companies. The results show that there is no effect of firm size, firm age, independent commissioners, leverage, and concentration of ownership on intellectual capital disclosure either simultaneously or partially.The reason that supports this research in relation to intellectual capital is always carried out is as a form of researcher participation in dealing with a problem that is essentially considered quite crucial, while efforts to improve taxpayer compliance are always encouraged so it is very important to carry out this research again Basically the disclosure of intellectual capital will certainly affect the value or amount of taxes that must be paid to the government. So the title of the research that will be raised by the researcher in connection with this topic is "The Relationship of Disclosure of Intellectual Capital in the Financial Statements of Plantation Companies to Compliance with Paying Taxes".

II. Formulation of The Problem

The following are some of the main problems that must be analyzed in this research, including:

- Does the application of accounting conservatism, company size, company age, ownership concentration, and independent commissioners partially affect tax compliance?
- Is the disclosure of intellectual capital able to moderate the relationship between the application of accounting conservatism, company size, company age, concentration of ownership, and independent commissioners with tax compliance?

III. Agency Theory

Agency theory provides the view that there is a contractual relationship between the agency and the principal. In this agency theory there is a view that there are certain interests owned by the principal where everything is not only measured by material, but also considers many things such as working environment conditions, comfort with coworkers, interaction with superiors and so on. The agency has full authority over the company's operational access so that it is free to act. While the principal focuses on what they will earn or get from the investment process in the company, the principal is not actively involved in the company's operations. The consequence is that the agent is positioned and required to be able to follow the will/desire of the principal, thus the occurrence of a conflict of interest will be at risk or also called the agency problem.

IV. Stakeholder Theory

Stakeholder Theory explains that the company is also responsible to the company and is also responsible to the shareholders. Stakeholders are groups or individuals who have a strategic role to achieve the company's goals or targets, in other words, these stakeholders have a major influence in the company's activities, the process to achieve the company's goals or targets is on their shoulders, therefore their presence is very important in the company and is taken into consideration if you want to disclose information in stakeholder financial statements. The importance of the existence of stakeholders will also increase the defense of the existence of the company.

V. Compliance Paying Taxes

Compliance with paying taxes is the ability of a taxpayer to show his good faith and maintain his ethics as an obedient Indonesian citizen in paying taxes. Thus, compliance with paying taxes is an attitude that deserves appreciation in order to increase the income owned by the State, where the benefits will also return to the community. There are two types of taxpayer compliance, namely formal tax compliance, this type of tax is more directing taxpayers towards formal matters, such as compliance with laws and regulations related to taxes, having a TIN, being able to report SPT on time and so on. While material tax compliance is more towards personal honesty and awareness to carry out the SPT filling process correctly and in accordance with the existing facts.

VI. Intellectual Capital

The association in the field of accounting also certainly takes over in connection with the issue of intangible assets which explains that intellectual capital is part of intangible assets. Basically intellectual capital is considered as a capital owned by a human being who will participate in the progress of a company. Said or disclosed as an intangible asset but has a fairly special value. Thus, the simple analogy is that intellectual capital will make a fairly good contribution to a company if the existence of intellectual capital can be identified and appreciated well by the company concerned.

VII. Application of Accounting Conservatism

Accounting conservatism in a company is basically contained in different levels. Several things that affect the percentage of conservatism in financial statements include the commitment of management and internal companies to disclose financial information that is transparent, measurable, accurate and not manipulated. Conservatism can be measured on an accrual basis, where when it has a negative value, the profits owned by the company can be classified as conservative, which is because the company's profits are lower when compared to cash inflows.

VIII. Company Size

Firm size or better known as company size gives meaning in relation to the size of the company's assets, both tangible assets and intangible assets. For example, a banking company that is classified in the category of giant/big company that has a large number of assets, which can be understood as a large company. Where, the company has a certain effort to make the process of finding, obtaining, developing, utilizing, maintaining resources, and revealing strategic resources will be maximized. Basically, the amount owned by firm sixze can be known through the total assets, total sales and the company's ability to market capitalization. So that the size of the company is dependent on the ownership of all its assets which can be measured by its ability to expand and maintain the company.

IX. Company Age

In essence, the age of the company is the age owned by the company, starting from the company's establishment until the research activity is carried out. Where, the age of the company can show the extent or how long the company has the ability to survive in the face of market share. The longer the company has been in existence, of course, becomes a benchmark that the company is able to stand up to survive in the long term, the more it has a solid defense and of course has the ability to develop very well. Thus, this becomes a very good parameter for an investor to invest. The identification is that a company that has been around for a longer time will certainly be more transparent in disclosing its intellectual capital even in detail because it will certainly participate in providing useful value added.

X. Ownership Concentration

Ownership is meant that the ownership of shares is controlled by a certain part or group of parties which causes the ownership to be sufficient to contribute to a company. The simple analogy is that the concentration of ownership is basically owned by interested parties, where the shares they have are properly allocated in a company, even though these shares cannot be specifically protected by a State. Because these shares are privately owned, all risks must be borne individually, including when the company in question experiences many problems.

XI. Independent Commissioner

Independent commissioners are expected to be able to act independently within a company without having to have a sense of siding with the company's management. This is because with the existence of an independent commissioner, the company can be better monitored, monitored and developed with the capabilities possessed or the authority possessed by the independent commissioner. This means that there is no partisanship that should not be done specifically to the company's internal parties which will later cause an imbalance between company interests and personal interests. Thus, independent commissioners are expected to provide perfect support for the progress of the company. The existence of an independent commissioner in a company is really needed.

XII. Hypothesis Development

From several explanations related to the previous research, the following are some research hypotheses that can be derived:

1. The application of accounting conservatism, firm size, firm age, concentration of ownership, and independent commissioners partially affect tax compliance.
2. Disclosure of intellectual capital is able to moderate the relationship between the application of accounting conservatism, firm size, firm age, concentration of ownership, and independent commissioners with tax compliance.

XIII. Research Model

The following is the research model that is described, namely through the conceptual framework of the research described using SmartPLS.

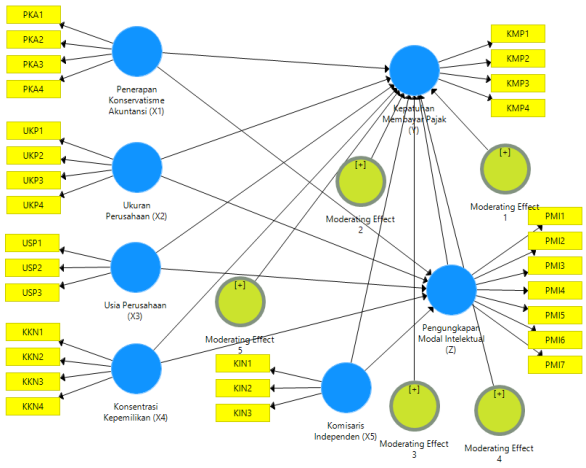


Fig. 1. Research model

XIV. Data Analysis Results

The results of data analysis from this study are as follows, where calculations and tests have been carried out in accordance with the predetermined data analysis method, including using the outer model test, inner model and also testing the moderating variable. The following presents the results of data processing that has been carried out where the conceptual framework that has been processed is as follows. Where from the conceptual framework we can see the results of the variables that do have a direct or indirect influence, and whether these variables are indeed able to moderate or strengthen the relationship between one variable and another and the following will be explained in more detail.

XV. Evaluation of Measurement (Outer) Model

A. *Validity test*

An indicator is declared valid if it has a loading factor above 0.5 for the intended construct. The SmartPLS output for the loading factor gives the following results, including:

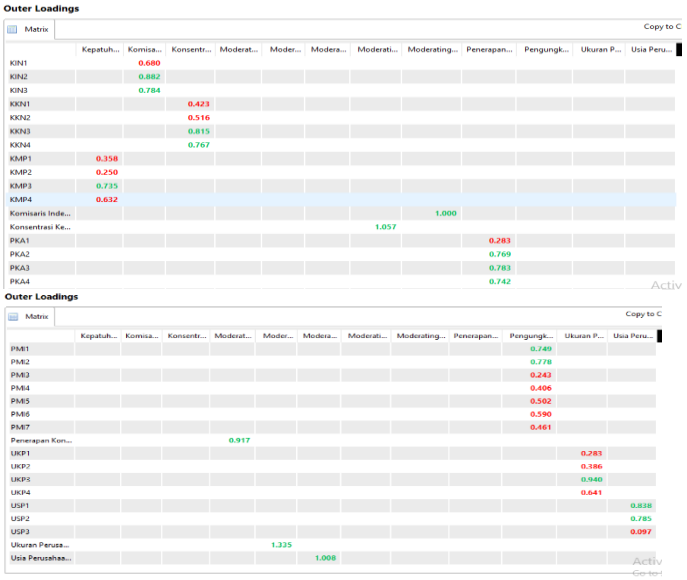


Fig. 2. The loading factor

Based on the loading factor table, it can be seen that the validity test for reflective indicators uses a correlation

between item scores and construct scores. Measurements with reflective indicators show a change in an indicator in a construct if other indicators in the same construct change or are removed from the model. Reflective indicators are suitable for measuring perception so that this study uses reflective. The table above shows that the green indicators have a loading factor value that is more than the recommended one or in other words, has a value above 0.5. However, there are also indicators that already have a loading factor value above 0.5, which is 0.6 but still feels inadequate but is assumed not to be too small. This means that in this study there are still indicators that are not feasible to use or will be re-evaluated by researchers. The highest loading factor value is found in the UKP3 company size indicator, which is 0.940. This means that this indicator has met the level of convergent validity. While some other indicators that show red are still not valid. However, the value that is owned is not too far away. Where it can be seen that the lowest loading factor value is 0.097 which is contained in the USP3 indicator.

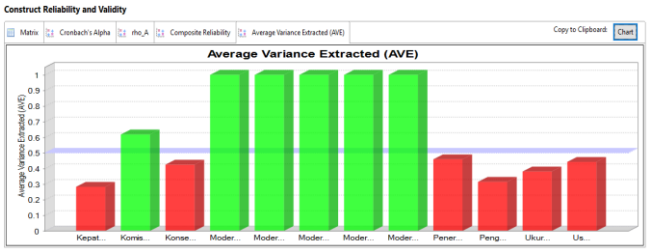


Fig. 3. The AVE value

The figure above shows the AVE values which are very diverse, starting from the AVE values that are above 0.5 to the AVE values that are below 0.5. Where it can be seen that the highest AVE value is found in the moderating variable, which is 1,000, while the lowest AVE value, which is below 0.5, which has a red color, according to the table above, has a value of 0.283, which is for the compliance construct of paying taxes. Where it means that it is necessary to re-evaluate the various indicators used for the construct.

B. *Reliability Test*

The reliability test is carried out by looking at the composite reliability value of the indicator block that measures the construct. The results of composite reliability will show a satisfactory value if the value is above 0.7. The following shows the composite reliability value using a graph in accordance with the results of this study.

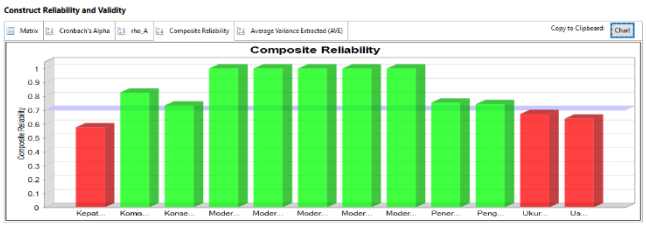


Fig. 4. The composite reliability value

The figure above shows that the composite reliability value for the majority of the constructs is above the value of 0.7 although not all of them. In some of the constructs above, there are also some whose values are below 0.7 including tax compliance, company size and company age. Thus, based on the graph that has been presented, all green constructs are constructs that do meet the criteria value, which is above 0.7. The highest construct value is 1,000 and the lowest value is below 0.7, which is approximately 0.5. The highest value is reinforced by moderating values, while the lowest value is the value of paying tax compliance constructs.

C. *Structural Model Testing (Inner Model)*

After the estimated model meets the Outer Model criteria, the next step is to test the structural model (Inner model). Here are the R-Square values in the construct:

R Square

Matrix | R Square | R Square Adjusted

	R Square	R Square Adjus...
Kepatuhan Me...	0.281	0.155
Pengungkapan...	0.271	0.218

Fig. 5. R-Square values

The figure above gives a value of 0.281 for the KMP construct, namely tax compliance, which means that all constructs in this study are able to represent that value, which is 28.1% in relation to its effect on the ability to pay taxes. Furthermore, it can be seen that the R Square Adjusted value is 0.155 or 15.5%. In addition, because this study uses a moderating construct, it can be seen that the R Square value is 0.271 or 27.1% which is the disclosure of intellectual capital (PMI) and the Adjusted R Square is 0.218 or 21.8%. In addition, there are also graphs presented in this study in relation to the R Square and R Square Adjusted values. Hypothesis testing is as follows:

	Original Sa...	Sample Me...	Standard Deviat...	T Statistics (O...	P Values
Komisaris Independen (X5) -> Kepatuhan Membayar Pajak (Y)	0.071	0.078	0.192	0.368	0.713
Komisaris Independen (X5) -> Pengungkapan Modal Intelektual (Z)	0.326	0.314	0.171	1.905	0.057
Konsentrasi Kepemilikan (X4) -> Kepatuhan Membayar Pajak (Y)	0.117	-0.010	0.260	0.451	0.652
Konsentrasi Kepemilikan (X4) -> Pengungkapan Modal Intelektual (Z)	0.159	0.158	0.217	0.732	0.465
Moderating Effect 1 -> Kepatuhan Membayar Pajak (Y)	-0.101	-0.037	0.190	0.535	0.593
Moderating Effect 2 -> Kepatuhan Membayar Pajak (Y)	0.038	0.044	0.191	0.200	0.842
Moderating Effect 3 -> Kepatuhan Membayar Pajak (Y)	0.011	-0.013	0.205	0.053	0.958
Moderating Effect 4 -> Kepatuhan Membayar Pajak (Y)	-0.074	-0.014	0.177	0.415	0.678
Moderating Effect 5 -> Kepatuhan Membayar Pajak (Y)	-0.106	-0.042	0.196	0.542	0.588
Penerapan Konservatisme Akuntansi (X1) -> Kepatuhan Membayar Pajak (Y)	0.056	-0.004	0.224	0.250	0.803
Penerapan Konservatisme Akuntansi (X1) -> Pengungkapan Modal Intelektual (Z)	0.158	0.120	0.145	1.090	0.276
Pengungkapan Modal Intelektual (Z) -> Kepatuhan Membayar Pajak (Y)	0.261	0.197	0.226	1.157	0.248
Ukuran Perusahaan (X2) -> Kepatuhan Membayar Pajak (Y)	-0.239	-0.206	0.301	0.795	0.427
Ukuran Perusahaan (X2) -> Pengungkapan Modal Intelektual (Z)	0.038	-0.085	0.243	0.156	0.876
Usia Perusahaan (X3) -> Kepatuhan Membayar Pajak (Y)	-0.099	0.094	0.336	0.295	0.768
Usia Perusahaan (X3) -> Pengungkapan Modal Intelektual (Z)	-0.258	-0.131	0.198	1.303	0.193

Fig. 6. Hypothesis testing

The table above shows that the relationship between PKA (X1) and KMP (Y) is significant with a T-statistic of 0.250 (< 1.665). The original sample estimate value is positive, which is 0.056 which indicates that the direction of the relationship between PKA and KMP is positive. Thus the hypothesis in this study which states that the application of accounting conservatism affects tax compliance cannot be accepted.

The table above shows that the relationship between UKP (X2) and KMP (Y) is significant with a T-statistic of 0.795 (< 1.665). The original sample estimate value is negative, which is -0.239 which indicates that the direction of the relationship between UKP and KMP is negative. Thus the hypothesis in this study which states that firm size affects tax compliance cannot be accepted.

The table above shows that the relationship between USP (X3) and KMP (Y) is significant with a T-statistic of 0.295 (< 1.665). The original sample estimate value is negative, which is -0.099 which indicates that the direction of the relationship between USP and KMP is negative. Thus the hypothesis in this study which states that company age affects tax compliance cannot be accepted.

The table above shows that the relationship between KKN (X4) and KMP (Y) is significant with a T-statistic of 0.451 (< 1.665). The original sample estimate value is positive, namely 0.117 which indicates that the direction of the relationship between KKN and KMP is positive. Thus the hypothesis in this study which states that ownership concentration affects tax compliance cannot be accepted.

The table above shows that the relationship between KIN (X5) and KMP (Y) is significant with a T-statistic of 0.368 (< 1.665). The original sample estimate value is positive, which is 0.071 which indicates that the direction of the relationship between KIN and KMP is positive. Thus the hypothesis in this study which states that independent commissioners affect tax compliance cannot be accepted.

The table above shows that the relationship between PKA (X1) and PMI (Z) is significant with a T-statistic of 1.090 (< 1.665). The original sample estimate value is positive, namely 0.158 which indicates that the direction of the relationship between PKA and PMI is positive. Thus the hypothesis in this study which states that the application of accounting conservatism affects the disclosure of intellectual capital cannot be accepted.

The table above shows that the relationship between UKP (X2) and PMI (Z) is significant with a T-statistic of 0.156 (< 1.665). The original sample estimate value is positive, which is 0.038 which indicates that the direction of the relationship between UKP and PMI is positive. Thus the hypothesis in this study which states that firm size affects the disclosure of intellectual capital cannot be accepted.

The table above shows that the relationship between USP (X3) and PMI (Z) is significant with a T-statistic of 1.303 (< 1.665). The original sample estimate value is negative, which is -0.258 which indicates that the direction of the relationship between USP and PMI is negative. Thus the hypothesis in this study which states that the age of the company affects the disclosure of intellectual capital cannot be accepted.

The table above shows that the relationship between KKN (X4) and PMI (Z) is significant with a T-statistic of 0.732 (< 1.665). The original sample estimate value is positive, namely 0.159 which indicates that the direction of the relationship between KKN and PMI is positive. Thus the hypothesis in this study which states that the concentration of ownership affects the disclosure of intellectual capital cannot be accepted.

The table above shows that the relationship between KIN (X5) and PMI (Z) is significant with a T-statistic of 1.905 (< 1.665). The original sample estimate value is positive, namely 0.326 which indicates that the direction of the relationship between KIN and PMI is positive. Thus, the hypothesis in this study which states that independent commissioners affect the disclosure of intellectual capital can be accepted.

Then the table above shows that the relationship between moderating effect 1 to moderating effect 5 with KMP (Y) is significant with T-statistics of 0.535, 0.200, 0.053, 0.415, and 0.542 (< 1.665). The original sample estimate values are negative and positive, each of which is -0.101, 0.038, 0.011, -0.074 and -0.106 which indicates that the direction of the relationship between the moderating effects 1 to 5 and the KMP is negative and positive. Thus the hypothesis in this study which states that the disclosure of intellectual capital in this study is able to strengthen the relationship between all independent variables and the dependent variable can be accepted.

Based on the original sample estimate value, it is obtained that the highest value affecting tax compliance (KMP) is the disclosure of intellectual capital (PMI) which is 0.261. This shows that the disclosure of intellectual capital has an influence on tax compliance. Furthermore, from all the independent variables and moderating variables, it turns out that the moderating variable does strengthen the relationship between several independent variables and the dependent variable. Thus the disclosure of intellectual capital is the most dominant variable in influencing tax compliance. While the least dominant variable is the age of the company with the smallest original sample estimate, which is -0.099.

XVI. CONCLUSION

Based on the research that has been done, several conclusions can be obtained from this research activity including:

- All of the hypotheses in this study are basically unacceptable, this is because in this study the T-statistical value produced is overall below the t-table value, so all hypotheses are rejected However, basically this is also related to the data that is processed on it and also the number of samples used.
- The moderating variable in this study is basically able to strengthen the relationship between all independent variables and the dependent variable.
- The values of R Square and R Square Adjusted are also not very large, but the meaning is that these values are sufficient to represent the independent variables even though they are only around 20% and above.

XVII. SUGGESTION

- It is necessary to carry out further evaluation in connection with the selection of independent variables, dependent variables and so on in connection with further research and adjust to the theory used in the study.

- A deeper understanding is needed regarding the selection of indicators in the variables and constructs that will be determined in a study.
- The next research is expected to be able to give a better contribution in scientific publications in an effort to improve the quality of publications for lecturers and also improve the institution's achievements in terms of scientific publications.

REFERENCES

[1] Ulum, I. (2016). Intellectual capital performance of the banking sector in Indonesia. Journal of Accounting and Finance, 10(2), 77-84.

[2] Lestari, N., & Sapitri, R. C. (2016). The influence of intellectual capital on firm value. Journal of Accounting, Economics and Business Management, 4(1), 28-33.

[3] Nugroho, A. (2012). Factors Affecting Intellectual Capital Disclosure (ICD). Accounting Analysis Journal, 1(2).

[4] Ulum, I. (2013). Intellectual capital performance measurement model with iB-VAIC in Islamic banking. INFERENCES: Journal of Religious Social Research, 7(1), 185-206.

[5] Istanti, Sri Layla Wahyu. 2009. Factors Affecting Disclosure.

[6] Voluntary Intellectual Capital. Thesis. Semarang. Diponegoro University. Khlifi, F. And A. Bouri. 2010. Corporate Disclosure and Firm Characteristics: A Puzzling Relationship. Journal of Accounting – Business & Management 17 (1): 62-89.

[7] Jacub, J. O. S. (2012). The Influence of Intellectual Capital and Disclosure on Firm Value (Empirical Study on Pharmaceutical Companies on the IDX). Scientific Journal of Accounting Students, 1(4), 96-100.

[8] Kartika, M. (2013). The Influence of Intellectual Capital on the Profitability of Banking Companies Listed on the Indonesia Stock Exchange in 2007-2011. Business Accounting Review, 1(2), 14-25.

[9] Suhardjanto, D. Dan M. Wardhani. 2010. "Intellectual Capital Disclosure Practices of Companies Listed on the Indonesian Stock Exchange", Indonesian Journal of Accounting and Auditing 14 No.1 : 71-85.

[10] Wijaya, N. (2012). The influence of Intellectual Capital on financial performance and market value of banking companies using the Value Added Intellectual Coefficient method. Journal of Business and Accounting, 14(3), 157-180.

[11] Artinah, B., & Muslih, A. (2016). The effect of intellectual capital on capital gains (empirical study of banking companies listed on the Indonesian stock exchange). Scientific journal of business and finance, 1(1).

Factors Influencing Purchase Intention Towards Eco-Friendly Products

1st Julianti Samsudin
Jabatan Perdagangan
Politeknik Sultan Salahuddin Abdul Aziz Shah, Shah Alam
Selangor, Malaysia
julianti.samsudin@gmail.com

2rd Lilis Seri Yana Sirun
Jabatan Perdagangan
Politeknik Sultan Salahuddin Abdul Aziz Shah, Shah Alam
Selangor, Malaysia
lisyana98@gmail.com

3rd Nor Vera Zuhairah
Jabatan Perdagangan
Politeknik Sultan Salahuddin Abdul Aziz Shah, Shah Alam
Selangor, Malaysia
nverazuhairah01@gmail.com

4th Yuvaneswary A/P Maniamaran.
Jabatan Perdagangan
Politeknik Sultan Salahuddin Abdul Aziz Shah, Shah Alam
Selangor, Malaysia
yuvaneswary3325@gmail.com

***Abstract*—The greatest global challenges during this modern era is to integrate environmental sustainability as healthy environments will make the people around the world healthy. It is not only for the people but also to the animals and living things. The twenty-first century threats to our environment puts wildlife populations at risk. The world is facing a lot of environmental issues today and one of the main causes is the use of non-eco-friendly packaging materials for various consumer goods is one of the problems that the world is facing today and one of the main reason is due to the widespread use of non-eco-friendly packaging of food and other goods. Since the non-eco-friendly products has been used a lot, this study aiming to identify factors that influencing purchase intention towards eco-friendly products especially in Shah Alam area. Quantitative research was adopted for this research as the research strategy. Questionnaire-survey has been used in this research with 382 sets of questionnaires distributed to the respondents to get the data required. The results show a positive influence with consumer awareness being the most influencing factors followed by environmental knowledge and product attributes. Lastly, this research will be able to help both consumers and firm to identify what's best for the environment to create a better world with the use of more eco-friendly products.**

Keywords—purchase intention, eco-friendly, product, consumer awareness, product attribute

I. INTRODUCTION

The environment is our basic life support system and is composed of living beings, physical surroundings, and climatic conditions. Healthy environments also support healthy wildlife. Twenty-first century threats to our environment including invasive species, diseases, pollution, and a warming climate are putting wildlife populations at risk. The world is facing a host of environmental issues today and one of the main causes is the use of non-eco-friendly packaging materials for various consumer goods is one of the problems that the world is facing today and one of the main reason is due to the widespread use of non-eco-friendly packaging of food and other goods. These packaging materials emit toxic pollutants or poisonous gases that affect human health and the environment. Eco-friendly products are products that are not harmful to the environment. These are products made from organic and all-natural ingredients. They also come in recyclable compostable, or biodegradable packaging.

In Malaysia, a recent study has confirmed that 38,000 tonnes of environmental waste have been generated every day without considering the labour cost, social cost and environmental cost Greenpeace Malaysia (2019). Malaysia imported ½ a million metric tons of environmental waste from the United State between January to July 2018 Greenpeace Malaysia (2019) because of it is a source of revenue for the government according to Greenpeace East Asia Ananth Lakshmi, Greenpeace Malaysia. (2019). Only 9 per cent from environmental waste have been recycled, 12 per cent incinerated and 79 per cent end up in landfills and natural environment according to Leoi, S. L. (2019). However, the government is coming up with implementing policies coupled with a greater awareness campaign to stimulate sustainable consumption among its nationals and consumer have begun to notice their purchasing behaviour is having a direct impact on environmental challenges according to Yahya, W.K., (2019). A shift in the attitude level and awareness level has to exist first in order to implement a swift in green behaviour.

This study was conducted to identify factors that influencing purchase intention towards eco-friendly products. It also aims to determine the dominant factors of purchase intention towards eco-friendly products and to examine the relationship between the factors (Consumer Awareness, Environmental Knowledge and Product Attributes) and purchase intention.

II. LITERATURE REVIEW

A. Dependent and independent variable

Purchase intention is the preference of consumer to buy the product or service (Keller, 2001). To describe it in another words, purchase intention has another aspect that the consumer will purchase a product after evaluation. There are a lot of factors that influence the consumer's intention when selecting the product and services they wanted to buy and they need. The decision to buy the product depends on the consumer's purchase intention. There are many external factors that cause them to intent purchase intention such as consumer awareness, environmental knowledge and product attributes.

The increased awareness about the environment friendly products made a large number of consumers to boycott such manufacturer and companies who were not implementing green marketing practices. Many firms had adopted this green

marketing concept to target such consumers who were environmentally conscious and were willing to pay more for the green products. Thus, green marketing had become a source sustainable of competitive advantage for many firms (Chen and Chai, 2010).

Mostafa, M.M, (2007) showed that environmental knowledge is positively linked with attitude towards green products which further influences their purchase intention. Consumers with a high level of environmental knowledge have a much better pro-environmental attitude and have a stronger intent to purchase green products for consumption Huang et al., (2014).

Some studies have stated that some of the product offerings such as value and quality, price, features of product are the main indicators to determine the customers understanding of green product concepts (Lee et al., 2010). The credibility of the green products is one of the elements that boost the purchasing powers of the customers because the performance and quality must reach their expectation (Ng et al, 2013).

B. *Hypotheses and Framework*

- H1: There is relationship between consumer awareness of eco-friendly products and purchase intention
- H2: There is relationship between environmental knowledge of eco-friendly products and purchase intention
- H3: There is relationship between product attributes and purchase intention.

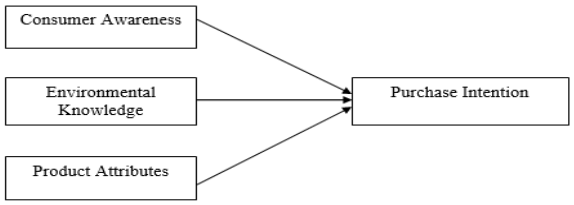


Fig. 1. Research framework

III. RESEARCH METHODOLOGY

A. *Sampling and Measurement Instruments*

The survey was conducted through an online survey by using google form link and a total 382 survey questionnaires are collected within Shah Alam residents in Selangor, Malaysia. The questionnaire design consists two part: Section A demographic factors of respondents such as gender, age, monthly income and educational information. Section B are questions regarding purchase intention towards eco-friendly products.

B. *Data Analysis*

The data was analyzed using Statistical Package of the Social Science (SPSS), employed both the descriptive and inferential analysis techniques (correlation analysis, multiple regression analysis and ANOVA). The statistical analysis was used is SPSS package 22.0.

IV. RESULTS AND DISCUSSION

A. *Respondents' Demographic Analysis*

There were 382 respondents of the questionnaire that all indicated their demographics. The demographic profile such as their gender, age, education level and monthly income has been analyzed and illustrated as shown in the Table I below.

TABLE I. RESPONDENT DEMOGRAPHIC PROFILE

Respondent's Demographic	Percentage (%)
Gender	
Male	47.1%
Female	52.9%
Age	
18-25 years old	59.9%
26-35 years old	9.7%
36-45 years old	9.7%
46-55 years old	12.3%
56 years and above	8.4%
Education	
SPM	30.4%
Diploma	41.6%
Degree	17.3%
Masters	6.5%
PhD	4.2%
Monthly Income	
Less than RM1500	56.3%
RM1600-RM3000	18.3%
RM3100-RM5000	13.1%
RM5100-RM7000	5.2%
RM7100 and above	7.1%

B. *Descriptive Analysis*

Table II indicates total average mean for independent variables. The highest mean is product attributes with average mean value at 4.17 followed by consumer awareness with mean value 4.16 and environmental knowledge with mean value 4.15. It shown that product attributes are the most dominant factors towards purchase intention of eco-friendly products.

TABLE II. DESCRIPTIVE ANALYSIS

	Number of items	Total Mean	Standard deviation
Consumer Awareness	5	4.16	0.8392
Environmental Knowledge	5	4.15	0.8165
Product Attributes	5	4.17	1.2102

C. *Correlation Test*

Pearson correlation test were used to examine the relationship between factors (consumer awareness, environmental knowledge, product attributes) and purchase intention towards eco-friendly products. According to the analysis, all the factors are significantly giving impact on consumer purchase intention towards eco-friendly products.

TABLE III. PEARSON'S CORRELATION COEFFICIENTS OF THE STUDY VARIABLES

	PA	CA	EK	PI
Product Attributes	1.00			
Consumer Awareness	0.715**	1.00		
Environmental Knowledge	0.738**	0.889**	1.00	
Purchase Intention	0.688**	0.908**	0.815**	1.00

**. Correlation is significant at the *p<.05, **p<.01 level.
PI = Purchase Intention
CA = Consumer Awareness
EK = Environmental Knowledge
PA = Product Attributes

D. Regression Test

In this study, multiple regression analysis is used to provide pattern or relationship between the factors and outcome factor (purchase intention). Table IV, indicates a summary of the results. R Square value is 0.827 and in can be interpreted that 82.7% of the variance in purchase intention of consumers is significantly explained by the independent variables of consumer awareness, environmental knowledge, and product attributes.

TABLE IV. MODEL SUMMARY

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.909[a]	0.827	0.826	1.51122

a. Predictors: (Constant), Consumer Awareness, Environmental Knowledge, Product Attributes
b. Dependent Variable: Purchase Intention towards Eco-friendly Products

The ANOVA table below is indicating the significant value is less than α alpha 0.05 level with value of 0.000. Therefore, hypothesis is accepted.

TABLE V. ANOVA[b]

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	4131.340	3	1377.113	602.998	.000[a]
Residual	863.267	378	2.284		
Total	4994.607	381			

a. Predictors: (Constant), Consumer Awareness, Environmental Knowledge, Product Attributes
b. Dependent Variable: Purchase Intention towards Eco-friendly Products

Generally, consumer awareness is the most influence independent variable among three variables, the reason is that it has the highest standardized beta coefficients with the value of 0.908. Based on the result of multiple regression analysis in Table VI below, consumer awareness, environmental knowledge, product attributes interpret that there are the significant positive relationship with purchase intention towards eco-friendly products. The reason is that all of their P- value regard with 0.000 which is less than significant value of 0.05.

TABLE VI. RESULT OF MULTIPLE REGRESSION ANALYSIS

Variable	Standardized Coefficient Beta (β)	Sig.
Consumer Awareness	0.908	0.000
Environmental Knowledge	0.815	0.000
Product Attributes	0.688	0.000

Note: *p< .05, **p< .01

The study has shown that consumer's purchase intention towards eco-friendly products has a positive relationship with consumer awareness. The increased consumer awareness about the eco-friendly products made a large number of consumers to boycott non eco-friendly products and the firms after knowing the effects. This is in the same lane as research that are conducted by (Chen and Chai, 2010), (Maniatis, 2015; Tseng & Hung, 2013) and (Rex & Baumann, 2007, p. 567). Thus, many firms are working hard to adopt this eco-friendly concept into their products to target such consumers who were environmentally conscious and were willing to pay more for more eco-friendly products.

Next, there is also a positive relationship between consumer's purchase intention towards eco-friendly products and environmental knowledge. This can happen to both the consumer and the producer of the product itself it's because if consumer are looking for the best product and find out about the benefit it gives to the environment, they might choose the product. Same goes to the firm, they will promote more about environment after knowing that it gives benefits to not just to the environment but to gain profits. This is in sync with past literature from Wang, P., Liu, Q., Qi, Y., (2014) and Huang et al., 2014; Rokicka, (2002).

There is a positive relationship between consumer's purchase intention towards eco-friendly products and product attributes. It is because nowadays attributes such as the features of the product whether they are energy or water efficient, whether they use healthy and nontoxic materials, whether they are made from recycled or renewable sources are also important in sync with how important the product offerings such as value, quality and price. Past research that can relate to this are (Lee et al., 2010), (Napolitano, 2010) and (Bui, 2005).

TABLE VII. TEST OF SIGNIFICANT

Hypothesis	Result	Supported /Rejected
H1: There is a relationship between consumer awareness and purchase intention.	Beta: 0.908	SUPPORTED
H2: There is a relationship between environmental knowledge and purchase intention.	Beta: 0.815	SUPPORTED
H3: There is a relationship between product attributes and purchase intention.	Beta: 0.688	SUPPORTED

V. CONCLUSION

The findings of the research conclude that consumer awareness, environmental knowledge and product attribute can affect the purchase intention of the customers strongly. This study provides insights that these factors could influence the performance of business on eco-friendly products. In conclusion, this study would be able to provide a reference or guideline for industries related to eco-friendly products in strengthening their sustainability in the industry.

REFERENCES

[1] Greenpeace Malaysia. (2019). The Recycling Myth - Malaysia and the broken global recycling system. https://www.greenpeace.org/international/press-release/19566/recycling-from-developed-world-dumped-in- Malaysia-and-left-to-rot/

[2] Leoi, S. L. (2019). Malaysia generates 38,000 tonnes of waste every day. https://www.star2.com/living/2019/05/21/waste-environmentissue/#It9EOPmu6rQU1PIq.99

[3] Yahya, W.K., (2019). Making sustainable consumption work. https://www.nst.com.my/opinion/columnists/2019/02/461780/making-sustainable-consumption-work

[4] Keller, K. L. (2001). Building customer-based brand equity: A blueprint for creating strong brands (pp. 3-27). Cambridge, MA: Marketing Science Institute.

[5] Chen, T. B., & Chai, L. T. (2010). Attitude towards the environment and green products: Consumers' perspective. Management Science and Engineering, 4(2), 27–39.

[6] Mostafa, M. M. (2007). A hierarchical analysis of the green consciousness of the Egyptian consumer. Psychology & Marketing, 24(5): 445–473.

[7] Huang, Y. C., Yang, M., & Wang, Y. C. (2014). Effects of green brand on green purchase intention. Marketing Intelligence & Planning, 32(3), 250–268.

[8] Lee, J.S., Hsu, L.T., Han, H. and Kim, Y. (2010) 'Understanding how consumers view green hotels: how a hotel's green image can influence behavior intention', Journal of Sustainable Tourism, September, Vol. 18, No. 7, pp.901–914

[9] Ng, P.F., Butt, M.M., Khong, K.W. and Ong, F.S. (2013) 'Antecedents of green brand equity: an integrated approach', Journal of Business Ethic, pp.1–13.

[10] P. Maniatis (2015). Investigating factors influencing consumer decision-making while choosing green products. Journal of Cleaner Production, 1 (2015) (2015), pp. 1-14

[11] Tseng, S.-C., Hung, S.-W. (2013). A framework identifying the gaps between customers' expectations and their perceptions in green products. Journal of Cleaner Production 59, 174–184

[12] Rex, E., & Baumann, H. (2007). Beyond ecolabels: what green marketing can learn from conventional marketing. Journal of Clean Production, 15, 567-576.

[13] Wang P., Liu Q., Qia Y. (2014). Factors influencing sustainable consumption behaviors: a survey of the rural residents in China. J. Clean. Prod. 63 152–165. 10.1016/j.jclepro.2013.05.007

[14] Rokicka, E. (2002). Attitudes towards natural environment, International Journal of Sociology, 32(2): 78–90.

[15] Napolitano, F. (2010) 'Cheese liking and consumer willingness to pay as affected by information about organic production', Journal of Dairy Research, Vol. 77, No. 3, p.280.

[16] Bui, M.H. (2005) Environmental Marketing: A Model of Consumer Behaviour, Loyola University New Orleans, USA.

Financial Incentives and Students Achievement: A Qualitative Study on Students' Success in Engineering Program

1st Nur Hidayah Mohamed Yunus
Mathematics, Science and Computer Department,
Politeknik Mukah Sarawak
Malaysia
hidayah.yunus@pmu.edu.my

2nd Aini Nazura Paimin
Faculty of Technical and Vocational Education,
Universiti Tun Hussein Onn
Malaysia
aininazura@uthm.edu.my

***Abstract*— Issues on retention in higher education have been discussed from variety of perspectives and is complex to be understood. There is an urge need to understand how some engineering students could survive until final year while the others not. Financial incentive is suggested as an important external motivational factor that could influence student's achievement in the program. These beliefs have led to the use of financial incentives (along with the intrinsic incentives already presumed to be present) to motivate students. This study explores the role of financial incentives in motivating positive emotions of students in collectivist context, Malaysia. The main aim of this study is to qualitatively explore how financial incentive contributes to students' academic achievement. 24 final year engineering undergraduates were purposely selected from four Malaysian Technical Universities Network (MTUN) who obtaining a CGPA of at least 3.60 and above were selected as research informants. Data were collected using semi-structured interviews, integrated with two creative methods namely graph sketching and photo-card activity. Data were analysed using a thematic coding approach and the coding process was assisted using Atlas.Ti software. Findings of the study discussed how education loan and scholarship play role in.**

Keywords—Financial Incentive, External Motivation, Student Success, Higher Education

I. INTRODUCTION

Education is synonymous with rapid country development all around the world. [1] states that the engineering sector is one of the important areas that is strongly related to a country's development. The increasing numbers of student enrolments in engineering program from both public and private institutions indicate increasing trend by years [2]. However, a high rate of enrolment does not guarantee a high return in terms of number of successful engineering graduates.

Statistics shows that the number of students enrolling in Bachelor's engineering degree programs is higher than the number of students who graduated in the consequent years [3]. This rate can be supported by a statistical report by National Education Statistics from Ministry of Higher Education [4] which showed that the number of engineering students who dropped out before the end of the study period was fluctuated between 15 and 30 percent for each of the respective years.

Some of the previous research also reported that a huge number of engineering students dropout from engineering programs after first year of study [5-7]. Based on the dropout issue, there has been increasing interest from universities counterpart in understanding the behavior of the successful students. The purpose of this study is to explore about students' experiences related to their academic achievement and factors that influence them to survive from the first year to the final year specifically in the financial incentives.

II. LITERATURE REVIEW

Previous research have discussed several factors that influence student academic achievement in engineering programs. These factors will be discussed further in this section.

A. Educational Financing for Higher Education in Malaysia

Financial assistance has become essential for students who wish to pursue their studies at higher education level [8]. Regularly, Malaysia's tertiary education funding comes from various sources such as personal or self-financing, government or private funding. Due to financial crisis, it is becoming hard to obtain a scholarship and there are usually two main sources of financial supports, namely education loans and scholarships.

National Higher Education Fund Corporation (NHEFC) (also known as Perbadanan Tabung Pendidikan Tinggi National, PTPTN) is an agency responsible to provide and administering education loans to students who pursuing their studies at public and private institutions in Malaysia. Apart from the PTPTN loan, students are also rely on the second funding options, scholarships which comes from either government institutions or private agencies such as Maybank, MARA, CIMB Bank, Bank Rakyat, SHELL, Khazanah, Petronas, and many more to complete their studies in higher education institutions [9-10] .

Settlement of PTPTN educational loan is compulsory for those who funded by PTPTN funding according to the terms of the arrangement [11]. However, there is also an incentive provided by PTPTN where students could apply for an exemption if they manage to obtain a Bachelor's Degree with First Class Honours results (CGPA for first class are varies by

institution) and graduate within the scheduled time (graduate on time). Meanwhile, there is also a situation where the funding is paid according to yearly basis where each students will be assessed every semester and students must meet the minimum CGPA requirements that have been set (for example, GCPA 3.0), otherwise the scholarship funding will be ceased the next semester if they did not fulfil the minimum requirement.

B. Motivation

Motivation is the driving factor of the will and desire to succeed or achieve something [12]. Motivation can also be considered as a condition where a person plans towards achieving success to avoid failure [13-15]. [16] stated that students' success in higher education can be driven by motivational factors, namely the intention from the heart that drives a person to make an action that includes all kinds of urges, stimuli, needs and desires to achieve a goal. In the context of this study, extrinsic factors such as financial incentives refer to external supports that is believe to influence engineering students' inner motivations to persist and bouncing back from any adverse situation (resilient) throughout their study period.

C. Theory of Reasoned Action (TRA)

The theory of reasoned action model (TRA) connects individual mental states namely belief, attitude, intention to form a specific behaviour [17-19]. There are several elements of intention namely desire, resourcefulness, initiative and persistence [20]. [21] Intention is believed as the main determinant of behaviours where an individual's actions can directly influenced by the formation of intention within oneself. This view is in line with [22] study, who provide evidence to the direct relationship between resourcefulness and academic achievement of engineering students at an Australian university. According to the original theory, a person's positive or negative belief about performing a certain behaviour (which is known as behavioural belief) can influence emotional states which consequently leads to intentional behaviours. An individual will intend to perform a particular behaviour when he or she positively evaluates that the intended behaviour will have a positive impact on him or herself. In the theory of planned behaviour model (TPB), normative beliefs and subjective norm are proposed as the contributing factors of intention prior to form an action [23]. In this situation, an individual's beliefs to act is also shaped by their perception on the external motivation factors such as community, culture, and environment.

D. Incentives

An incentive is a material or rewards given or offered as an encouragement to a person for the purpose of motivation or encouragement or inspiration that indirectly leads to individual change in behaviours as an effort to obtain the incentive [24]. [25] study stated that a person who controls his or her motivation to obtain an incentive known as controlled extrinsic motivation which refers to the means of a person to control his or her behaviours to get the desired incentive.

III. METHODOLOGY

The qualitative case studies approach was chosen as the study design considering the nature of the study that are more compatible with informants' experiences. The elaboration of findings is also clearer, concrete and easy to understand. According to [26-28], case study inquiries are contextual specific focusing on the researcher's answers for each questions to the informants. Qualitative researchers themselves play role as research instruments who responsible for the entire process of data collection and analysis [29-31].

In this research, semi-structured interviews were conducted with 24 undergraduate engineering students at four technical universities in Malaysia namely Universiti Tun Hussein Onn Malaysia (UTHM), Universiti Technical Malaysia Melaka (UTeM), Universiti Malaysia Pahang (UMP) and Universiti Malaysia Perlis (UniMAP). All the informants were final year engineering students specialised in one of three main fields of engineering; Civil, Electrical and Mechanical, and must obtaining a CGPA of at least 3.60 and above.

The purposive sampling selection in this study was made by using referral and chain referral techniques. According to [32-33], data from informants chosen via purposive sampling ensures richness and completeness of data for the phenomena under study. Each interviews took place individually between 00:52:04 to 01:48:10 hours and was recorded with informants' consent. Other than using semi-structured interviews, graft sketching and photo card activity were an alternative approach of collecting data in this study. Ethical considerations were also applied and a full consent was obtained from the Universities, the faculties and the informants themselves before data were collected. After the interview data were transcribed and transform into a transcript, thematic analysis was undertaken to identify common themes using the ATLAS.ti software.

IV. FINDINGS

Two themes were proposed under financial incentives that are education loans and scholarships. Financial incentives can comes from any sources of fundings that encourage students to strive and thrive to obtain excellent results with various form of intentions that are directed towards securing or preserving the incentives.

A. Education Loan

From the results obtained, all the informants who had educational loans stated that they strived their best to obtain good results for the purpose of getting the loan exemption. Informants also strived to maintain their first-class GPA to waive their loan commitment. Here are some statements from the informants regarding their intentions to obtain the financial incentives.

> *I received an education loan from PTPTN, I strived to maintain high GPA starting from the first semester to obtain exemption of the loan payment (Informant 7).*

> *I remain perform my best until now (final year) because I want to maintain my exam results and most importantly, I want to obtain exemption from PTPTN (Informant 17).*

> *Although my motivation level is decreasing every semester, I still remain until now (final year) and strive to get a first-class degree to obtain education loan payment exemption from PTPTN (Informant 10).*

> *I work hard to foster interest in this field so that I can score good results in the exams so that my studies are funded by PTPTN (Informant 20).*

I always work hard in all my subjects so that I can obtain the best exam results and qualify for loan payment exemption from PTPTN (Informant 3).

My main goal throughout my studies is to obtain a loan payment exemption from PTPTN. Therefore, I have to work hard to be able to get first-class results as required (Informant 18).

All these statements have proved that financial commitment can be one of motivational factors that influence students' efforts to maintain and achieve good results with a main goal to obtain an exemption for the loan commitment. The result of this study is in line with [34-36], who also provide very similar outcomes in their study. Across the three decades, financial factor play important role as a driven factors that could influence students' intention to achive and maintain better performance in engineering program. Furthermore, [37- 39], claimed that when the rate of payment for free education fees increases, the rate of student attendance also increases. This reflects that students can be more motivated to participate in the learning process and strive to meet the academic requirements in order to obtain the assured incentives.

B. Scholarship

The results showed that only six informants who did not receive any loans from PTPTN instead, they received scholarships from either the state governments or private companies. The following excerpts are some of the informant's statements regarding their scholarship;

I was funded by a scholarship from the Sabah state government. I force myself to remain until final year and strive my best to get a good GPA to make sure that I can continually get the funding (Informant 9).

Because I being bonded by a scholarship agreement, I had to work hard each semester to earn a first-class GPA (Informant 14).

I am so determine to obtain the Dean's lists for each semester to ensure the continuity of my scholarship funding. Therefore, I always forced myself to foster deep interest in any of the engineering topics eventhough this field was not my first choice (Informant 2).

All informants who get the scholarships stated that they need to strive hard to maintain a high GPA and to ensure the continuity of the scholarship throughout the years of study otherwise, their sponsorship will be terminated. This results in line with the Theory of Reasoned Action (TRA) where individuals will set an intentions in the first place and striving persistently towards achieving the desired goal [19]. This indicates that the insistence from their sponsors, eventough seems to be stressful, can be a good motivation for students to persistently put high commitment in ther study. Such an external force, can be a promising approach to ensure that students always strive their best to maintain good performance until graduation. According to [40], efforts made to move an individual toward one or more specific goals and to ensure the attainment of incentives or rewards are a form of extrinsic motivations.

The findings of this study are in line with [41-42], that the provision of incentives in the form of finance has a positive effect on student achievement at the tertiary level. In addition, the findings also clearly showed that each informant has set the beliefs that if they could succeed in getting excellent results, they are also able to obtain financial incentives or loan exemption [43]. This findings provide support to the notion that a person's behavior and actions are depending on establishment of intentions and beliefs [17&43].

V. CONCLUSION

In overall, findings of the current study demonstrates the role of financal invcentives as an extrinsic driven factors of success behaviours where informants demonstrated specific behaviours towards goal attainment, such as to secure scholarship funding or to get exemption for PTPTN loan. This study highlighted that the two main factors as important supportive factors to ensure good academic performance. Students strive to achieve high results with the purpose to obtain educational scholarship or loan, getting loan exemption and to ensure continuity of their scholarship funding. Therefore, students need to clearly establish learning intentions and goals so that every effort they made could potentially leads them towards the desired achievement. Based on these findings, these researchers propose a potential relationships between financial incentives, desire, persistence and student achievement for a further investigation.

ACKNOWLEDGMENT

The researchers thank the Ministry of Higher Education for this research funding (FRGS 1605), MTUN universities, the engineering faculties and students for their cooperation in this research.

REFERENCES

[1] M. Sas, K. Ponnet, G. Reniers, and W. Hardyns, "The Role of Education in the Prevention of Radicalization and Violent Extremism in Developing Countries," *Sustain. Educ. Approaches*, 2020.

[2] C. Faye and D. Sharpe, "Academic Motivation in University : The Role of Basic Psychological Needs and Academic Motivation in University : The Role of Basic Psychological Needs and Identity Formation," no. October 2008, 2014.

[3] Kementerian Pengajian Tinggi Malaysia., "Statistik Pendidikan Tinggi 2016. Bab 1 : Makro Institusi Pendidikan Tinggi," 2016. [Online]. Available: http://mohe.gov.my/muat-turun/awam/statistik/2016-statistik/401-bab-1-makro-institusi-pendidikan-tinggi-

[4] Kementerian Pengajian Tinggi Malaysia., "Statistik Pendidikan Tinggi 2019. Bab 1: Makro Institusi Pendidikan Tinggi," 2019. [Online]. Available: http://mohe.gov.my/muat-turun/awam/statistik/2019-1/612-bab-1-makro-institusi-pendidikan-tinggi-3

[5] M. Van Den Bogaard, "Explaining student success in engineering education in Delft University of Technology : A synthesis of literature," *Eur. J. Eng. Educ.*, vol. 37, no. 1, 2012.

[6] M. Pinxten, T. D. Laet, S. V. Soom, and G. Langle, "Fighting increasing dropout rates in the STEM field: The European readySTEMgo Project," *43rd Annu. SEFI Conf.*, 2015.

[7] L. J. Ortiz, V. A. Rua, C. P. Bilbao, and F. M. Casadesus, "University student retention: Best time and data to identify undergraduate students at risk of dropout," *Innov. Educ. Teach. Int.*, vol. 57, no. 1, 2020.

[8] A. H. A. Hanafi, W. R. M. A. Shahimi, M. A. Anuar, K. Y. Chin, and S. S. Hwei, "Determinants of borrower's intention to repay the educational loan : PLS-SEM method," *Int. J. Accounting, Financ. Bus.*, vol. 3, no. 7, pp. 27–38, 2020.

[9] N. R. Zainal and N. Ismail, "Debt Composition of University Graduates and their Attitude towards Education Loan," *J. ASIAN Behav. Stud.*, vol. 2, no. 4, pp. 41–47, 2017.

[10] N. B. Zakaria, N. Mohamed, D. Daud, and A. M. Ismail, "Study Loan Defaults Among Tertiary Graduates," *Int. J. Financ. Res.*, vol. 11, no. Special issue, 2020.

[11] N. D. Zamro, "Antecedents of the educational loan repayment among the POLIMAS students," *J. Appl. Soc. Psychol.*, 2020.

[12] K. S. Atman, “Goal Accomplishment Style and Psychological Type Cultural Variations,” *Psychol. Type Cult. West Multicult. Res. Symp.*, no. 1, pp. 207–220, 1993.

[13] C. P. Cerasoli, J. M. Nicklin, and M. T. Ford, “Intrinsic Motivation and Extrinsic Incentives Jointly Predict Performance : A 40-Year Meta-Analysis,” in *Psycological Bulletin*, Advance online publication, 2014.

[14] W. C. C. Pinder, *Work motivation in organizational behavior (2nd ed.)*. New York, NY: Psychology Press, 2011.

[15] R. M. Ryan and E. L. Deci, “Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions,” *Contemp. Educ. Psychol.*, vol. 25, pp. 54–67, 2000.

[16] M. K. Baharuddin and Z. M. Ashari, “Relationship between Shame Emotions toward Attachment Styles and Achievement Motivation among Students : Meta Analysis Study,” *Sains Humanika UTM*, vol. 2, pp. 15–18, 2018.

[17] C. Nisson and A. Earl, “The Theories of Reasoned Action and Planned Behavior,” in *The Wiley Encyclopedia Psycology*, S. K, M. L. Robbins, and L. M. Cohen, Eds. 2020.

[18] I. Azjen, *Attitude, persoality and behaviour*. New York: Open University Press., 2005.

[19] M. Fishbein and I. Ajzen, *Beliefs, attitude, intention and behaviour: an introduction to theory and research*. Canada: Addison-Wesley Publishing Company, 1975.

[20] S. Foen Ng, G. J. Confessore, and M. Abdullah, “Learner autonomy coaching: enhancing learning and academic success,” *Int. J. Mentor. Coach. Educ.*, vol. 1, no. 3, pp. 191–204, 2012.

[21] J. G. Confessore and E. Park, “Factor Validation of the Learner Autonomy Profile, version 3.0 and Extraction of The Short Form,” *Int. J. Self-Directed Learn.*, vol. 1, no. 1, pp. 39–58, 2004.

[22] A. N. Paimin, “The success factors of engineering students: Learning strategies, interests and intentions,” 2014.

[23] I. Ajzen, “The theory of planned behaviour: Reactions and reflections,” *Psychol. Heal.*, vol. 26, no. 9, pp. 1113–1127, 2011.

[24] D. Goldhaber and J. Cowan, “Do bonuses affect teacher staffing and student achievement in high poverty schools? Evidence from an incentive for national board certified teachers in Washington State,” *J. Econ. Psychol.*, 2018.

[25] J. Reeve and S. H. Cheon, “Teachers become more autonomy supportive after they believe it is easy to do,” *Psychol. Sport Exerc.*, 2016.

[26] N. K. Denzin and Y. S. Lincoln, *Introduction: The discipline and practice of qualitative research*, 4th ed. Thousand Oaks, CA: SAGE Publication, Inc, 2011.

[27] J. W. Creswell and C. N. Poth, *Qualitative inquiry and research design: Choosing among five approaches*, 4th ed. United kingdom, London: SAGE Publication, Inc, 2018.

[28] J. Farquhar, N. Michels, and J. Robson, “Triangulation in industrial qualitative case study research: Widening the scope.,” *Ind. Mark. Manag.*, 2020.

[29] D. Silverman, *Doing qualitative research q practical handbook (Fourth Edition)*. SAGE Publication, Inc., 2013.

[30] I. Seidman, *Interviewing as Qualitative Research: A Guide for Researchers in Education and the Social Sciences*. New York, NY: Teachers College Press, 2013.

[31] S. B. Merriam and E. J. Tisdell, *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass, 2016.

[32] A. Bagheri and M. Saadati, “Exploring the Effectiveness of Chain Referral Methods in Sampling Hidden Populations,” *Indian J. Sci. Technol.*, 2015.

[33] E. B. Andrea, “Sampling Method,” *J. Hum. Lact.*, 2020.

[34] D. Z. Norqulova, “The importance of incentives in the education system,” *Int. Soc. Humanit. Res.*, 2021.

[35] R. E. Slavin, “Can financial incentives enhance educational outcomes? Evidence from international experiments,” *Educ. Res. Rev.*, 2010.

[36] M. J. Johnson and S. D. Sheppard, “Students entering and exiting the engineering pipeline-identifying key decision points and trends,” *32nd Annu. Front. Educ.*, vol. 3, p. S3C–13–S3C–19, 2002.

[37] J. Meece and C. Agger, “Oxford Research Encyclopedia of Education Achievement Motivation in Education Gender-Related Trends in Academic Achievement and Educational Attainment in the United States and Abroad,” no. March 2019, pp. 1–21, 2018.

[38] T. Gunnes, L. J. Kirbeboen, and M. Ronning, “Financial incentives and study duration in higher education,” *Labour Econ.*, 2013.

[39] J. Bishop, “Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn - Externalities, Information Problems and Peer Pressure,” *Handb. Econ. Educ.*, vol. 2, pp. 909–944, 2006.

[40] E. E. M. Meens, *Motivation Individual differences in students’ educational choice and study success*. 2018.

[41] W. Dixon, “Predicting student retention using scholarship and grant aid,” PhD : Liberty University, Lynchburg, 2018.

[42] C. P. Veenstra, E. L. Dey, and G. D. Herrin, “A model for freshman engineering retention,” *Adv. Eng. Educ.*, vol. 1, no. 3, pp. 1–23, 2009.

[43] K. S. N. H. Mohd Razali, “Kajian terhadap strategi pembelajaran, emosi, konatif dan sasaran pencapaian pelajar kejuruteraan di Universiti Teknikal Malaysia,” Universiti Tun Hussein Onn Malaysia, 2017.

[44] S. C. Lin and C. Chen, *Application of theory of planned behavior on the study of workplace dishonesty*. Manila, Philippines.: IAC S IT Press, 2011.

Development of DPB6043 E-Notes Business Project Mobile Applications

1st Habsah Mohamad Sabli[1]
Commerce Department
Politeknik Mukah
Sarawak, Malaysia
habsah@pmu.edu.my

2nd Siti Khadijah Sebli Joney
Commerce Department
Politeknik Mukah
Sarawak, Malaysia
khadijah.seblijoney@pmu.edu.my

3rd Mohamad Shafiq Mohamed Nasaruddin
Commerce Department
Politeknik Mukah
Sarawak, Malaysia
mshafiq@pmu.edu.my

***Abstract*—The Covid-19 pandemic and Movement Control Order (MCO) resulted in the closure of all higher education institutions throughout the country to prevent the spread of this contagious outbreak. The teaching and learning process becomes fully online learning due to this situation. Students living in rural areas face the problem of no internet access. Hence, the purpose of the innovation DPB6043 E-Notes in the form of mobile application to assist students who faced problem with online learning during Covid-19 pandemic. This mobile application was developed using Thunkable and Google Sites based on TAM theory, CBAAM, and ADDIE model. Data were obtained using the survey research method. A total of 33 students from the final semester of Diploma in Business Studies (DPM) program from the Commerce Department, Politeknik Mukah Sarawak participated in the survey. The survey findings further revealed that DPB6043 E-Notes improves students' understanding and is very important for online learning with the highest mean = 4.484. The outcome of this research can benefit the decision-makers of higher institutions in Mukah Sarawak regard to the way to enhance m-learning and promote effective teaching and learning activities as well as strengthening the quality of learning delivery.**

Keywords—mobile application, business project, student, Covid-19, DPB6043

I. INTRODUCTION

DPB6043 Business Project course is a compulsory course that must be taken by students of Diploma in Business Studies (DPM) at Politeknik Mukah and Polytechnic Malaysia. This course covers knowledge in the basic areas of theoretical and practical research that emphasizes the concept of teamwork in the production of project output via face-to-face learning. This course is divided into two choices i.e., innovation on product or services or conducting case studies. Covid-19 pandemic and MCO resulted in the closure of all higher education institutions throughout the country to prevent the spread of this pandemic. According to UNESCO (2020), approximately 320 million learners in India are infected, with approximately 34 million enrolled in tertiary education [1]. In terms of education, how institutions and stakeholders adapted and faced challenges to the new scenario created by the Covid-19 pandemic [2-4] as well as training strategies and innovation experiences [5] have been published. Internet connections problems also can affect online learning. Students living in rural areas face the problem of no internet network and need to go to the city to get an internet line that requires a high cost. Hence, the use of information technology plays an important role in the teaching and learning process facing current situations [6]. There is a lack of agreement on the critical challenges and methods that shape the successful use of e-learning systems during the Covid-19 pandemic; thus, a clear gap in knowledge on the critical challenges and methods of e-learning usage during this pandemic has been identified. In line with the current presence of mobile devices, the field of education can ease the constraints that existed before and be able to equip students with interest in the DPB6043 Business Project course. Hence, the purpose of the innovation DPB6043-Notes in the form of mobile application is to assist students who faced problem with online learning. take DPB6043 Business Project course during Covid-19 pandemics. The use of mobile applications is a new learning method created to match the development of a borderless information world. It was also created to reduce the burden of students who have to carry the burden of bringing thick and heavy reference books to lectures. There are several rationales for the development of mobile apps innovation materials in the DPB6043 Business Project module, among which are that students can easily read and be carried anywhere as the notes are in their mobile applications only. In addition, this innovation is to save money where students do not need to print reference materials, but they can use them anytime and anywhere from their mobile devices. The mobile app has many advantages among which it makes it easier for students to read notes without printed books.

II. LITERATURE REVIEW

The developments of mobile applications DPB6043 E-Notes based on TAM theory. The Technology Acceptance Model (TAM) has been regarded as one of the most fundamental and influential theories in predicting m-learning adoption. The technology acceptance model identifies the relationships between external variables, perceived usefulness, perceived ease of use, and behavioral attitude [7].

The degree to which an individual believes that using a particular system will be free of physical and mental effort and will allow them to complete tasks easily was defined as perceived ease of use [8]. The ease of use associated with the themed-learning system is regarded as an important factor in encouraging users to adopt mobile learning. Furthermore, the contents of DPB6043 E-Notes are based on the Computer-Based Assessment (CBAAM) model [9]. The Computer-Based Assessment Acceptance Model (CBAAM) is a model that explains learners' intentions to use a computer-based assessment satisfactorily. The CBAAM employs nine major variables, including perceived usefulness, perceived ease of use, perceived playfulness, perceived importance, social influence, facilitating conditions, perceived content, goal expectancy, and computer self-efficacy, to determine the intention to used [10]. The Perceived Content (PC) of a CBAAM is defined by two factors: 1) the students' perceptions of the course content and 2) the students' perceptions of the questions asked during the CBAAM. As a result, learners first assess the content based on their prior knowledge, experience gained during the course, such as if it is difficult, interesting, and useful, and secondly, during the CBAAM if the questions are clear and understandable [11].The ADDIE model was chosen because of its systematic generic approach to instructional design, which simplifies the instructional framework for designers or researchers, ensuring the effectiveness of instructional products with creative processes [12]. The ADDIE instructional model [13] describes the analysis, design, development, implementation, and evaluation of instruction. The subject of learning is established through analysis, the method of learning is established through design, the instructional materials are chosen and created through development, the materials and activities are implemented through implementation, and the impact of instruction is established through evaluation. Each phase contained in the ADDIE model has specific steps and procedures.

M-Learning, or mobile learning, is frequently associated with the use of mobile technology, particularly the mobile phone [14, 15]. According to [15], the term "mobile" indicates to it is both portable and personal. [16] describes m-learning as "wireless learning", a subset of e-learning that emphasizes the use of personal communication devices. M-learning, according to [17], solves the problems of insufficient internet access, frequent power outages, and limited PC support and availability. It has a particularly strong impact in isolated and rural places and is bolstered by the mobile phone networks' potential and skill. M-learning is a type of teaching and learning that takes place on mobile devices such as phones, PDAs, and other similar devices, allowing learners to access material and conduct learning at any time and from any location [18].

III. RESEARCH DESIGN AND METHODOLOGY

A. *Innovation Implementation Methods*

Mobile applications DPB6043 E-Notes development developed through Thunkable software. Then, after the e-content has been developed, the software will produce a single-user version application (standalone) in the form of APK format. This APK format can be installed on any smartphone such as iPhone, Samsung, Oppo, and others. In addition, the development of e-Content involves minimum storage and low data usage allow users to freely start learning as quickly as possible. In addition to the application of support activities, information in e-Content is also integrated with user understanding without altering the original meaning of the text. Users are also free to go to any part without the following order and without finishing reading on one display in advance. Mobile application software uses Thunkable and Google Sites to upload the notes DPB6043 Business Project done by lecturers. After that, lecturers will publish on the Google Play store to be downloaded by students. After completing the process, students were taught to download DPB6043 E-Notes mobile application software on their smartphones for free. This innovative material is very easy to use once it is uploaded into student-owned mobile devices. Students can also take it anywhere and it does not require an internet line. It is also used during the teaching and learning process in the classroom.

The application development process of DPB6043 E-Notes is based on the Perceived Ease of Use used in the Theory Acceptance Model (TAM). DPB6043 E-Notes content development process is content based on a computer-based acceptance model (CBAAM). Each phase contained in the ADDIE model has specific steps and procedures. According to [19] analysis at the first stage involves needs analysis, student and lecturer analysis, context analysis, and content analysis. Then, the output result for the analysis phase acts as input to the design phase. In addition to the design model, builders also combine multimedia elements from YouTube to attract students to be more focus and interested with the online notes. Multimedia contains text, audio, video, and motion. Text-based instructional contents are being phased out in favor of content that is auditory, visible, and has a high level of interaction with the user as the internet infrastructure evolves in the twenty-first century[20].

B. *Innovation Development Design*

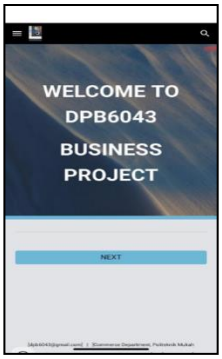


Fig. 1. Front screen

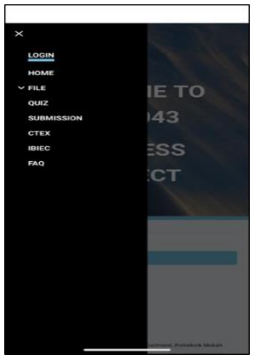


Fig. 2. Side menu

Fig. 3. Home screen

Fig. 4. Control screen

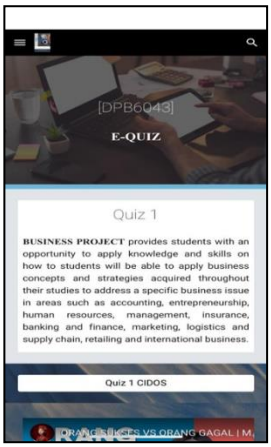


Fig. 5. Quiz screen

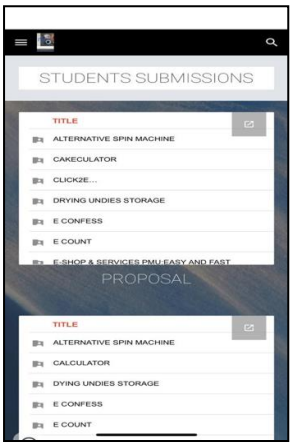


Fig. 6. Submission screen

IV. FINDING OF STUDY AND DISCUSSION

The scope is focused on DPM5 class students who took the subject of DPB6043 Business Project Session December 2020, which is a total of 33 students. Further, under time and budget constraints, the chosen sample is justified. Due to time and cost constraints, convenience sampling is used in this research. This study deployed a survey method to obtain data. Refer to tables below; the questionnaire instrument contained three sections. The first section focused on obtaining respondents' demographic profiles such as age, gender, and area. The second section in the questionnaire is the DPB6043 E-Notes Business Project application. There were 15 questions in this section to identify the perception of students. The questionnaire adapts from [21]. Section C involved questionnaire is DPB6043 E-Notes Business Project contents which have seven questions. Questions were slightly modified to suit the context of mobile technology and application for college students adapts from [10]. The relative importance of each construct was assessed by a 5-point Likert scale from one being "strongly disagree" to five being "strongly agreed".

TABLE I. RELIABILITY TEST ITEM QUESTIONNAIRE EFFECTIVENESS OF MOBILE APPS DPB6043 BUSINESS PROJECT FOR DPM5 STUDENTS

Reliability Statistics	
Cronbach's Alpha	***No of Items***
.840	22

The result of the Cronbach Alpha reliability test was 0.840 at a good level exceeding 0.60. This shows respondents understand the question of the distributed survey.

TABLE II. THE DEMOGRAPHIC BACKGROUND OF THE RESPONDENTS

Respondent Profile		**n**	**(%)**
Age	21 to 22	16	59
	23 to 24	10	37
	25 and above	1	3.0
Gender	Men	5	15.2
	Women	28	84.8
Area	Kuching	1	3.0
	Sibu	5	15.2
	Sri Aman	1	3.0
	Miri	3	9.1
	Bintulu	5	15.2
	Kapit	7	21.2
	Semarahan	1	3.0
	Tatau	1	3.0
	Debak	1	3.0
	Tawau	1	3.0

Refer to Table II, 33 DPM5 students took the DPB6043 course as respondents. The majority of respondents aged between 21 to 22 years of age (59 %), followed by 23 to 24 years (37%) and over 25 years (3%) up there is only one. The respondents were 28 women (84.8%), while men were only 5 (15.2%) only. Most of the respondents lived in Kapit District (21.2%) and other districts such as Kuching, Sri Aman, Kota Semarahan, Tatau, Debak, and Tawau (3%).

TABLE III. MEAN AND STANDARD DEVIATION FOR DPB6043 BUSINESS PROJECT INNOVATION APPLICATION FOR DPM5

Descriptive Statistics				
No.	***Item***	***n***	***Mean***	***Std. Deviation***
1.	Using DPB6043 E-Notes is better than the traditional method.	22	4.000	.5590
2.	DPB6043 E-Note design is good.	22	4.030	.6839
3.	DPB6043 E-Note system is user friendly.	22	4.000	.5000
4.	DPB6043 E-Note interface is very attractive.	22	4.000	.7071
5.	DPB6043 E-Notes important for online learning.	22	4.484	.6671
6.	DPB6043 E-Notes are easy to install.	22	4.242	.7512
7.	Using DPB6043 E-Notes easy to upload and download documents.	22	4.394	.7044
8.	DPB6043 E-Notes can be used offline.	22	4.212	.7809
9.	I can use DPB6043 E-Notes anytime.	22	3.818	.9170
10.	DPB6043 E-Notes can be used by android and iOS systems.	22	4.000	.7071
11.	DPB6043 E-Notes can be installed using a smartphone, tablet, and laptop.	22	4.152	.7124
12.	ICT skills and knowledge are important in DPB6043 E-Notes.	22	4.394	.7044
13.	DPB6043 E-Notes useful during Covid-19 pandemic.	22	4.333	.4787
14.	DPB6043 E-Notes are free and have no fees.	22	4.424	.6628
15.	Overall DPB6043 E-Notes easy to use.	22	4.273	.5740

As shown in Table III, the level of mobile application usage among DPM5 students in PMU is as high as the mean is between 3.50 and 4.49. For this category, the highest mean value of 4.484, E-Notes is very important for online learning and teaching during the Covid-19 pandemic. Overall, the majority of respondents agreed that the use of mobile applications in facilitating teaching and learning activities at PMU was enjoyable (min = 4.273). However, the lowest mean value in this category is related to the use of mobile application E-Notes DPB6043 anytime by phone (min = 3.818).

TABLE IV. MEAN AND STANDARD DEVIATION FOR DPB6043 BUSINESS PROJECT INNOVATION APPLICATION FOR DPM5 STUDENTS

Descriptive Statistics				
No.	***Item***	***n***	***Mean***	***Std. Deviation***
1.	DPB6043 E-Notes improve students understanding of the subject.	22	4.484	.5075
2.	The contents DPB6043 E-Notes follows the syllabus.	22	4.181	.6825
3.	Using DPB6043 E-Notes is more interesting to study.	22	4.151	.7124
4.	I'm more focused on the study, using DPB6043 E-Notes.	22	3.909	.7230
5.	I'm more confident answering the quizzes after using DPB6043 E-Notes.	22	3.939	.6585
6.	I appreciate all notes and videos in DPB6043 E-Notes.	22	4.090	.5789
7.	Overall, the content DPB6043 E-Notes is very informative.	22	4.484	.5075

As shown in Table IV, the level of mobile application usage among DPM5 students at PMU is moderately high with the mean values being between 3.50 and 4.49. For this category, the highest mean value (min=4.48) i.e. can increase students' knowledge and strengthen existing knowledge in students. Overall, the majority of respondents agreed on the contents of the mobile application DPB6043 E-Notes in facilitating teaching and learning activities at PMU (min= 4.484). The contents of this mobile application follow the syllabus content i.e. (min= 4.181). However, the lowest mean value in this category is that students are more focused on using DPB6043 E-Notes via easy-to-use smartphones (min = 4.909).

V. DISCUSSION AND SUGGESTION FOR IMPROVEMENT

Results found that the students were very satisfied with the application developed. This shows that there is a very positive potential among students to use mobile applications in DPB6043 Business Project online learning. The results support the finding from [22] explained that the aspect of user acceptance depends on how easily users feel and can be followed by users when using any type of application. This high user-friendly aspect is also in line with the [11] study, which also depends on how long the user uses the app and adapts to the application. Overall, the content DPB6043 E-Notes is very informative and can improve students understanding of the subject. This supports the findings from [23]the content has a positive impact on students adopting mobile learning. The contents must be informative, interesting, and follow the syllabus. Before the use of the DPB6043 Business Project mobile application, students were only provided with notes in the form of PowerPoint slides or hardcopy notes for reference during the teaching and learning process. The problem faced by the students is, they are easily bored with the existing form of notes and difficult to carry where they go. With the development of innovative materials the DPB6043 E-Notes mobile application can add and foster student interest in DPB6043 Business Project courses and so on can make students more proactive in the classroom.

REFERENCES

[1] UNESCO. (2020). COVID-19 Educational Disruption and Response. from https://en.unesco.org/covid19/educationresponse.

[2] Bao, W. 2020. "COVID-19 and Online Teaching in Higher Education: A Case Study of Peking University

[3] Flores, M. A., and M. Gago. 2020. "Teacher Education in Times of COVID-19 Pandemic in Portugal: National, Institutional and Pedagogical Responses." Journal of Education for Teaching, Advance online publication. doi:10.1080/02607476.2020.1799709

[4] Quezada, R. L., C. Talbot, and K. B. Quezada-Parker. 2020. "From Bricks and Mortar to Remote Teaching: A Teacher Education Programme's Response to COVID-19." Journal of Education for Teaching, Advance online publication. doi:10.1080/02607476.2020.1801330.

[5] Ferdig, R. E., E. Baumgartner, R. Hartshorne, R. Kaplan-Rakowski, and C. Mouza, Eds. 2020. Teaching, Technology, and Teacher Education during the COVID-19 Pandemic: Stories from the Field. Association for the Advancement of Computing in Education (AACE).

[6] N. Kapasia et al., "Impact of lockdown on learning status of undergraduate and postgraduate students during COVID-19 pandemic in West Bengal, India," Child. Youth Serv. Rev., vol. 116, no. June, p. 105194, 2020.

[7] Davis. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

[8] A. Harchay, A. Berguiga, L. Cheniti-Belcadhi, and R. Braham, "Student perception of mobile self-assessment: An evaluation of the technology acceptance model.," Interact. Des. Archit., no. 41, pp. 109–124, 2019.

[9] C. N. Moridis, V. Terzis, and A. A. Economides, "The effect of instant emotions on behavioral intention to use a computer-based assessment system," IEEE Glob. Eng. Educ. Conf. EDUCON, no. April, pp. 1457–1462, 2017.

[10] Terzis, V., & Economides, A. A. (2011). The acceptance and use of computer-based assessment. Computers & Education, 56(4), 1032–1044.

[11] Seels, B.B.; Richey, R.C. Instructional Technology: The Definition and Domains of the Field, 1994 ed.; Information Age Publishing: Charlotte, NC, USA, 2012.

[12] College Station (2001). ADDIE Instructional Design Model. Texas Copyright 2001 © LOT All rights reserved. L:\htms\training\handouts\pf_files\addie.doc.

[13] Elias, T. (2011), Universal Instructional Design Principles for Mobile Learning International Review of Research in Open and Distance Learning, Vol. 12(2). Retrieved June 19, 2012, from http://www.irrodl.org/index.php/irrodl/article/view/9 65/1675

[14] Cavus, N, Bicen, H., & Akcil, U. (2008). The Opinions of Information Technology Students on Using Mobile Learning. Online Submission, (2003), 23–25.

[15] Naismith, L., Lonsdale, P., Vavoula, G., Sharples, M., (2005). Literature Review in Mobile Technologies and Learning. NESTA Futurelab Series.

[16] Traxler, J. (2007). Defining, Discussing, and Evaluating Mobile Learning: The moving finger writes 8(2).

[17] Yamaguchi, T. (2005). Vocabulary learning with a mobile phone. Program of the 10th Anniversary Conference of Pan-Pacific Association of Applied Linguistics, Edinburgh, UK.

[18] Lehner, F.; Nosekabel, H. The Role of Mobile Devices in E-Learning First Experiences with a Wireless E-Learning Environment. In Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education, Vaxjo, Sweden, 30 August 2002.

[19] Pollara, P. (1999). Mobile Learning in Higher Education: A Glimpse and A Comparison of Student and Faculty Readiness, Attitudes and Perceptions. Neurosurgery, 45(3), 975–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17690829

[20] S. Gümüş and M. R. Okur, "Using multimedia objects in online learning environment," *Procedia - Soc. Behav. Sci.*, vol. 2, no. 2, pp. 5157–5161, 2010, doi: 10.1016/j.sbspro.2010.03.838.

[21] Fetaji, Majlinda, Suzana Loskovska, Bekim Fetaji, and Mirlinda Ebibi. "Combining virtual learning environment and integrated development environment to enhance e-learning." In Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on, pp. 319-324. IEEE, 2007.

[22] N. S. Yahaya and S. N. A. Salam, "Mobile Learning Application for Children: Belajar Bersama Dino," Procedia - Soc. Behav. Sci., vol. 155, no. October, pp. 398–404, 2014, doi: 10.1016/j.sbspro.2014.10.312.

[23] Wang, Y.-S., Wu, M.-C., & Wang, H.-Y. (2009). Investigating the determinants and age and gender differences in the acceptance of mobile learning. British Journal of Educational Technology, 40(1), 92–118.

The Influence of Financial Ratio on Changes in Profit in Mining Companies Listed on The Indonesia Stock Exchange (Bei)

1st Sastra Karo-Karo
Accounting Department
Politeknik Negeri Medan
Medan, Indonesia

2nd Hartono Ginting
Accounting Department
Politeknik Negeri Medan
Medan, Indonesia

***Abstract*— The purpose of this study was to determine the effect of financial ratios consisting of quick ratio, return on equity, and debt to assets ratio on profit changes in mining companies listed on the Indonesia Stock Exchange. The population in this study were all companies engaged in the coal mining sector which were listed on the Indonesia Stock Exchange until 2019, while the sampling in this study use purposive sampling method with the criteria of coal mining companies that published their financial reports in 2017 to 2019 and have made profits in the 2017 – 2019 financial year period as many as 22 companies. The data testing technique uses multiple linear regression and partial hypothesis testing with a significance level of alpha = 5%. The results of the research show that the financial ratios consisting of the quick ratio, return on equity, and debt to assets ratio has no effect on changes in profits in coal mining companies. In conclusion, this research supports the research conducted by Justina (2015) and Lestari (2020) because Justina and Lestari's research results showed that financial ratios have no effect on changes in earnings and are contradictory to research done by Prameswari (2018). Prameswari's research results conclude that the current ratio, debt to equity ratios affects profit growth.**

Keywords— financial ratios and profit changes.

I. INTRODUCTION

Competition in the business world and industry in all sectors, at this time and especially during this pandemic is very tight, for that management and all employees who work in the company are required to improve their performance so that the companies they manage can survive. The company is said to survive if the company can survive in uncertain economic conditions. This can be seen from its ability to fulfill financial obligations and carry out its operations properly and can maintain the continuity of its business development from time to time. In general, the public measures the success of a company based on the company's ability as seen from its performance The company's performance can be assessed through financial statements that are presented every period. To realize this performance, it can be done by analyzing financial statements using financial ratios. Financial ratios are a comparison of numbers from the estimates contained in the balance sheet and income statement. With the financial ratios, it can be seen whether the company's financial condition is safe or not and how the company's profit growth in the future. This financial ratio analysis can also be used by management to predict the health of the company operations and company performance, as well as potential investor, need to conduct an analysis to determine which company shares will be selected for their fund to be invested. The amount of profit earned by a company can be seen from the company's financial statements in the Profit/Loss Statement. Every company expects an increase in profit from the previous period and the company's profit from one period to the next period usually changes. Profit is an increase in economic benefits during an accounting period in the form of an increase in assets or a decrease in liabilities which results in an increase in equity that is not derived from investment contributions. The presentation of profit information is a very important part of the company's operations because the level of profit can describe the company's overall achievements. The greater the level of profit of a company, the better the productivity of the company, so that it will increase stakeholder confidence in the company. Justina (2015) conducted research on textile and garment companies found that financial ratios as measured by current ratio and net profit equity and total assets turnover had no effect on profit growth (Prameswari,2018) concluded that the current ratio, debt to equity has an effect on profit growth while total asset turnover has no effect on profit growth in the banking sector. Lestri (2020) conducted research on advertising, printing, and non-building construction companies and conclude that financial ratios consisting of the current ratio, total assets turnover, debt to equity, and return on assets did not simultaneously affect profit changes. The phenomenon of the inconsistency of research results that have been carried out as described above is the attraction of researchers to replicate research by conducting the same research on mining companies. This study aims to conduct further testing of empirical findings regarding financial ratios, especially regarding their influence in predicting company profits in the future. The selection of changes in corporate profits as a predicted phenomenon in this study is based on the reason that similar studies still do not provide consistent results. If the influence of financial ratios can be used as a prediction of future earnings changes, this research is certainly quite useful knowledge for users of financial

statements who have real or potential interests in a company, and otherwise, the influence of financial ratios is not significant enough in predicting earning changes. In the future, the results of this study will strengthen evidence about the inconsistency of previous empirical findings.

This research is a replication of Lestari's research (2020), the difference between the research that the author did with this research is that the data for the research year studied are 2017, 2018, and 2019 and the variables used are quick ratio (QR), return on equity (ROE) and debt to assets ratio (DAR) and the companies studied are companies engaged in mining, while Lestari uses research data from 2014 – 2018 and the ratios used are current ratio (CR), total assets turn over (TOTA), debt to equity ratio (DER) and return on assets (ROA) and the companies studied so far are companies in the advertising, printing, media subsector and non-building construction subsector companies.

Based on the description described above, the formulation of the problem in this study is whether the financial ratios consisting of the quick ratio (QR), return of equity (ROE), and debt to asset ratio (DAR) affect changes in profits in mining companies on the Indonesia Stock Exchange?. While the purpose of the study is to determine the financial ratios consisting of quick ratio (QR), return on equity ratio (ROE), and debt to asset ratio (DAR) that affects the changes in earnings of mining companies in the Indonesian Stock Exchange.

A. Understanding Financial Ratios and Financial Ratio Analysis

Financial ratios are a ratio calculation using financial statements as a measuring tool in assessing the financial condition and performance of the company [1] .

According to [11]), financial ratios are indexes that have a relevant and significant relationship between two numbers in financial statement items by comparing these numbers in one period or several periods in order to help evaluate a financial report

While financial ratio analysis is the process of observing indexes related to accounting in financial statements such as balance sheets, income statements, and cash flow statements with the aim of assessing the financial performance of a company. This analysis is used to provide an overview of information about the company's financial position and performance that can be used as a guide in making business decisions [13].

Analyzing financial statements means making comparisons to the numbers in the financial statements to obtain a good and precise understanding and understanding of the financial statements. Financial statement analysis aims to determine whether the financial condition as the result of the company's operations and financial progress is satisfactory or unsatisfactory. The financial analysis is carried out by measuring the relationship between the elements of the financial statements and how these elements change over time from year to year and also to find out the direction of the company's development progress [1].

From the understanding described above, the conclusion is that financial ratio analysis can be useful to evaluate the financial condition, company performance, and the health condition of the company concerned.

B. Relationship of Financial Ratios with Changes in Profit

A change in profit is an increase or decrease in the profit earned by the company compared to the previous year. Accounting profit is the difference between the measurement of income and costs, where the main focus of financial statements is profit. The definition of operating profit is the difference between the realized income arising from transactions during a period and the costs associated with that income [11].

All companies have the main goal in their business to generate maximum profit. Profit is very important because it is for the survival of the company going forward. "High *profits* are a sign that consumers want more industrial output and high profits provide incentives for companies to increase output and more companies will enter the industry in the long run.

At any given period the company's profit will show its activities. By knowing the changes in profit each period, it can be used as a benchmark to determine the company's development [9]. Financial ratios obtained from financial statements can be used for investors in estimating the calculation of cash receipts from distributed profits, interest in the future. Dividends received for investors will depend on the amount of profit in the future so that predictions of company profits by analyzing financial ratios are very necessary to do.

II. RESEARCH METHODS

The population in this study were coal mining companies listed on the Indonesia Stock Exchange until 2019 amounted to 25 companies, while the sample in this study was conducted using a purposive sampling method with the criteria of the companies used being coal mining companies that published financial statements in the 2017 – 2019 period and the company made a profit in the period amounted to 22 companies. The type of data in this study is a panel or pooled data consisting of a times series, namely the annual period 2017 – 2019 so the number of samples is 66 with an error rate (alpha) = 5%.

The variable used in this study is the quick ratio which is used to measure current assets minus inventory and prepayments compared to short-term liabilities. The next variable is the return on equity ratio, which is the company's ability to generate net operating profit to total equity. The debt to assets ratio is the ratio of long-term liabilities compared to company assets. The change in profit is the year's profit minus last year's profit compared to last year's profit.

The research model is in figure 1.

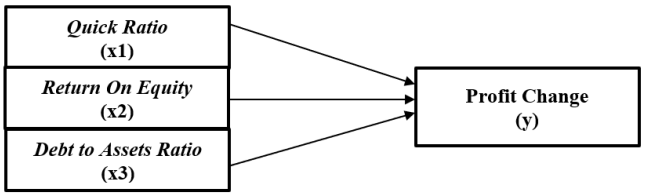


Fig. 1. Research Model

Data collection techniques in this study are library research and documentation. The data in this study were obtained from the Indonesia Stock Exchange through the website www.idx.co.id. The data analysis method used in this study was multiple linear regression using SPSS software.

III. Results and Discussion

A. Research Result

The purpose of descriptive statistics is to describe data from research results. The results of descriptive statistical analysis from this study are as follows:

TABLE I. Descriptive Statistics

	Mean	Std. Deviation	N
Profit Change	1,2283	8,84060	66
QR	802,0809	5033,94465	66
ROE	198,4936	1106,46771	66
DAR	17,3510	21,18128	66

Based on Table I mention above, it is known that the mean or average of profit changes is 1.2283 with a standard deviation of 8.84060. Thus, the mean value is smaller than the standard deviation value, and this indicates that the earnings change variable data used in this study is inefficient.

B. Hypothesis Test Results

Hypothesis testing in this study uses the t-test, which aims to see the effect of the independent variables, namely QR, ROE, and DAR on the dependent variable of earnings changes.

TABLE II. Multiple Regression Test Results

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.
		B	*Std. Error*	*Beta*		
1	(Constant)	2,336	1,457		1,603	,114
	QR	-6,71E.5	,000	-,038	-303	,763
	ROE	,000	,001	,033	,251	,802
	DAR	-,064	,055	-,153	-1,156	,252

From the multiple regression test shown on Table-2 above, the following multiple regression equation is obtained: Y=2,336-6,71E.5 QR + 0,000 ROE – 0,064 DAR.

IV. Discussion

From the multiple regression equation above, it can be seen that the constant of 2,336 states that if QR=0, ROE=0 DAR=0 then the value of the change in profit is 2,336.

The Quick ratio has a regression coefficient of -6,71E.5 stating that every 1% addition of the Quick ratio with the assumption that the value of the coefficients of other variables remains constant, will decrease the change in profit by 6,71E.5 but on the contrary if the Quick ratio decreases by 1% assuming the value the coefficients of other variables are fixed, the change in profit are predicted to increase by 6,71E.5. ROE has a regression coefficient of 0,000 stating that every 1% addition of ROE with the assumption that the value of coefficient of other variables is fixed it will reduce the change in profit by 0,000 but on the contrary if the Quick ratio decreases by 1% assuming the coefficient value of other variables remain, the change in profit is predicted to increase by 0,000. DAR has a regression coefficient of -0,064 which state that every 1% addition of ROE with the assumption that the coefficients of other variables remain constant, will reduce profit changes by 0,064, but on the other hand, variables remain constant, it will reduce profit changes by 0.064, but on the other hand, if the Quick ratio decreases by 1% assuming the coefficients of other variables remain, the profit changes are predicted to experience an increase of 0,064.

Looking at the results of the t-test shown on Table II, it can be explained that there is no independent variable that has an influence on changes in earnings because the significance value of QR = 0,763; ROE = 0,802, and DAR = 0,252 is greater than the alpha value of 5% or 0.05.

V. Conclusion

From the discussion above, it can be concluded that the financial ratios consisting of Quick Ratio, Return on Equity, and Debt to Assets Ratio have no effect on changes in profits in coal mining companies. This research supports research done by Justina (2015), and research done by Lestari (2020) and contradicts the research done by Prameswari (2018).

VI. Suggestion

Based on the conclusions that have been stated above, the researcher suggests for the next researcher to replace the independent variables to be studied, such as CAR, BOPO, or other variables and the company being studied is not a mining company.

References

[1] M. Andayani, M, and L. Ardini, L," Analisis Rasio Likuiditas dan Rasio Profitabilitas Terhadap Perubahan Laba". Jurnal Ilmu dan Riset Akuntansi, Volume 5, No. 7, Juli 2016.

[2] I. Ghozali, I, " Aplikasi Analisis Multivariate Dengan Program SPSS". Semarang, Badan Penerbit Universitas Diponegoro, 2016

[3] S.S. Harahap, "*Analisis Kritis atas Laporan Keuangan*", Jakarta : Rajawali Pers, 2015

[4] Hery, "*Analisis Laporan Keuangan Integrated AND Comprehensive*" Edition: Jakarta, PT Grafindo Persada, 2017

[5] T.L .Justina, , "*Pengaruh Rasio Likuiditas, Solvabilitas, Aktivitas, Profitabilitas dan Produk Domestik Bruto Terhadap Pertumbuhan Laba Pada Perusahaan Tekstil dan Garmen Yang Terdaftar di BEI*", JRKA, Volume 1, Isue 2, Agustus 2015, (60-70)

[6] Kasmir, " *Analisis Laporan Keuangan*", Jakarta, Rajawali Pers, 2015

[7] D. E. Kieso., J.J. Weygand., and T.D. Warfield, "*Akuntansi Keuangan Menengah*",Jakarta, Salemba Empat , 2017

[8] M. P. Lestari, "*Pengaruh Rasio Keuangan Terhadap Perubahan LabaPadaPerusahaan Sub Sektor Advertising, Printing, Media dan Pada Perusahaan SubSektor Konstruksi Non Bangunan Yang Terdaftar di BEI Periode 2014-2018*", Skripsi, Universitas Andalas, Padang, 2020

[9] A.A. Permata and S.R. Fuadati, *" Pengaruh Likuiditas dan Profitabilitas Terhadap Laba pada Perusahaan Retail Trade*". Jurnal Ilmu dan Riset Manajemen, Volume 5, Nomor 4, 2016

[10] I.P. Prameswari, "*Pengaruh Rasio Keuangan Terhadap Pertumbuhan Laba Pada Sektor Perbankan yang terdaftar di BEI*", Skripsi, UNSRI, Palembang, 2018

[11] I. L. K. Safitri, Sektor Industri Konsumsi yang Terdaftar di Bursa Efek Indonesia (Studi Kasus pada Perusahaan Kalbe Farma TBK Periode 2007-2014), "*Jurnal Akuntansi dan Bisnis*", Volume 2, Nomor 2 Nopember. 2016

[12] Sugiyono, "*Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif dan R&D, Bandung,* Alfabeta. 2017

[13] W. N. Utami, Pengertian, Fungsi dan Jenis Analisis Rasio Keuangan, "*Jurnal Enterpreneur*", Januari 5, 2020

Analysis of Customer Loyalty and The Influence of Product Attributes, Promotions and Religious Commitments on The Decision to Choose Saving Products in Sharia Bank Sumut

1st Muhammad Zuardi
Department of Accounting
Politeknik Negeri Medan
Medan, Indonesia
muhammadzuardi@gmail.com

2nd Muslim Marpaung
Department of Accounting
Politeknik Negeri Medan
Medan, Indonesia

3rd Irma Suryani Lubis
Department of Accounting
Politeknik Negeri Medan
Medan, Indonesia

***Abstract*—Savings are deposits that can only be withdrawn according to certain agreed conditions, but cannot be withdrawn by cheque, bilyet, demand deposit, and or other equivalent instruments. If the customer wants to take his deposit, he can come directly to the bank with a passbook, withdrawal slip, or through an ATM facility. Deposit customers have more consideration factors than financing customers, who tend to only focus on the amount of financing that must be paid. Financing customers are selected strictly by customers, in contrast to TPF customers who tend to act as selectors of the bank to be selected. This has an impact on bank policies to pay attention to many aspects in increasing the number of customers. Islamic banks as banks that display religious symbols have a target market for Muslim community groups in the process of collecting and distributing funds. Religion is one factor that is highly considered by consumers to decide to become a customer of a sharia bank, there are still many reasons and other factors that can influence consumer considerations to want to become a customer of a sharia bank. One of them comes from the marketing strategy carried out by Islamic banks through the marketing mix strategy.**

Keywords—Customer Loyalty, Product Attributes, Promotion, Religious Commitment, Decision to Choose a Savings Product.

I. PRELIMINARY

Customers who are disloyal to the bank will have a negative impacts, ranging from the bank's increasingly bad image to the possibility of a rush or massive withdrawal of funds in one period of time. Rush will cause a decrease in bank liquidity until finally there can be liquidation.

Having loyal customers is the ultimate goal of all companies. Eventhough the bank's services already considered quality, but customers are not necessarily loyal to a bank. The quality of bank services is sometimes not in line with the level of customer loyalty of a bank. Many companies do not know that consumer loyalty is formed through several stages, starting from finding potential customers to the establishment of customer advocates that will bring benefits to the company. (Wattimury, 2012).

Customer loyalty will be seen from the acquisition of third party funds. Funds that come from the general public or often called third party funds in the form of deposits, which traditionally usually consist of current deposits (demand deposits), time deposits and savings. Funds derived financial institutions, both in the form of banks and non-banks, are obtained by banks as short-term and long-term loans in accordance with the needs of the loan bank (Suyatno, 2004).

In Table I. the following is a list of Islamic banks that have the highest level of loyalty in 2019.

TABLE I. ISLAMIC BANK LOYALTY RATING IN 2019

No	Bank Name
1	Bank Muamalat
2	BNI Syariah
3	BRI Syariah
4	Bank Mandiri Syariah

Real customer loyalty is created if the customer becomes an advocate for the company without incentives though. Customer loyalty shows a desire to keep funds contimuously. Of course this is different if the customer is not loyal and makes a withdrawal of funds or closes the account. If done in large quantities and by many customers will cause a rush. Before the rush there will appear symptoms that indicate a less competitive bank. In 2020 there is rush money or chaotic bank PT Bank Bukopin Tbk, as a result many clients to withdraw funds on a large scale, whereas the Financial Services Authority has never released performance is unsatisfactory from the bank, but the issue of the rush to lowom).

Increasing customer loyalty will increase the acquisition of third party funds (Sari and Prijanto, 2016). North Sumatra is one of the provinces that has experienced an increase in third party funds from Sharia banks operating in this province. The expansion of the sharia banking business in North Sumatra (North Sumatra) in 2015 grew positively, which was marked by an increase in assets, third party funds and bank

financing. Sharia banking assets in the third quarter increased 3.68 percent compared to the second quarter or to Rp9.58 trillion. The increase in Sharia banking assets occurred amid the development of conventional banking and the presence of many non-bank financial institutions. Third party funds (DPK) and the financing of Sharia banking to grow positively. Sharia banking deposits amounted to Rp 5.69 trillion, up 3.83 percent from the second quarter, while financing also increased 3.07 percent or Rp. 7.38 trillion. The increase in assets, deposits and Sharia banking financing in North Sumatra this year shows a higher level of public trust (Regional Economic Study, 2015).

North Sumatra is the province with the largest population outside Java, this shows the market potential for banking products, especially Sharia banking. Seeing the increasing potential of North Sumatra, Sharia banks and sharia business units have added branch offices and sub-branches at the district/city and sub-district levels in North Sumatra. Savings and financing products are aggressively offered to prospective customers so that the bank's performance will continue to grow. Many factors affect the performance of Islamic banking, apart from being influenced by the internal factors of the bank itself, the performance of Sharia banking is also influenced by other monetary and financial indicators. For the implementation of the intermediation function itself, Sharia banks are still good with a high financing to deposit ratio (FDR) position. Thus, this also affects the collection of third party funds (Hasibuan, 2006).

The source of third party funds is arguably the core of the operational activities of the bank. Without this flow of funds that is smooth and growing, the continuity of the bank's business could be threatened with failure or liquidation.

TABLE II. MARKET SHARE OF BANK SUMUT SYARIAH THIRD PARTY FUNDS

Year	Third Party Funds (Thousands of Rp)		TPF Market Share (%)
	North Sumatra	*Islamic Bank of North Sumatra*	Islamic Bank of North Sumatra
2015	8 .371.000.000	682.726.997	8,16
2016	10.260.000.000	694.323.880	6,77
2017	11.837.000.000	936.124.749	7,91
2018	12.773.000.000	910.362.757	7,13
2019	14.806.000.000	1.061.002.050	7,17

Table II. shows an increase in DPK of Sharia banks in North Sumatra every year, in 2019, the number of TPF reached Rp. 14,806 Trillion. Bank Sumut Syariah also experienced an increase in TPF, except in 2018. In terms of market share, Bank Sumut Syariah only controlled 7.17% of DPK for Sharia banks in North Sumatra, the highest market share of Bank Sumut Syariah was 8.16% in 2015. There are 8 (eight) Sharia banks operating in North Sumatra, with a market share of 7 to 8%, show that the competitiveness of Bank Sumut Syariah is still not optimal.

The establishment of Bank Sumut Syariah is based on the religious culture of the people of North Sumatra, especially Muslims who are increasingly aware of the importance of carrying out their teachings in all aspects of life, including in the economic field. Since the issuance of Law No.10 T ear 1998 provides an opportunity for conventional banks to establish Sharia. Sharia Business Unit is a work unit at the head office of a conventional commercial bank that functions as the main office of a sharia branch office and or sharia unit, or a work unit at a conventional foreign bank branch office that functions as the main office of a sharia sub-branch office and or sharia unit (Soemitra, 2010). PT. Bank Sumut established a Sharia Business Unit under the name Bank Sumut Syariah. In line with the passage of time, until 2019 Bank Sumut Syariah had 22 operational offices consisting of 5 Branch offices and 17 Sub-Branch offices spread across Medan and other big cities in North Sumatra . This is certainly not in line with expectations. The people of North Sumatra, which is inhabited by 80% of the population are Muslims, and Bank Sumut Syariah as a regional bank were unable to increase market share as expected.

The potential for fund raising products from savings products is still very potential. Challenge faced by the Bank of North Sumatra Sharia in the year 2020 increasingly complex , because it is not only trying to attract funds from the people of North Sumatra are Muslims but also compete with eight (8) Islamic banks continue to expand by setting up branch offices and branch offices Devel tu in districts/cities in North Sumatra.

The acquisition of third party funds from the savings products of Bank Sumut Syariah is still lower than Time Deposits and Current Accounts, this can be seen in Table III.

TABLE III. COMPOSITION OF THIRD PARTY FUNDS FOR BANK SUMUT SYARIAH YEAR 2015 – 2019

Description	Thousands of Rupiah				
	Dec-15	*Dec-16*	*Dec-17*	*Dec-18*	*Dec-19*
Third-party funds	682,726,997	694,323,880	936,124,749	910,362,757	1.061.002.050
Current Account	12,643,185	12,777,347	11,934,703	16,608,105	148,724,407
Wadiah Current Account	11,906,028	12,032,368	11,257,084	15,665,143	15,028,966
Mudharabah Current Account	737.157	744,979	677,619	942,962	133,695.441
Savings	80,557,562	85,764,573	87,493,164	89,769,924	110,812,498
Wadiah Savings	9,121,859	9,218,655	8,654,372	8,879,577	10.156.114
Mudharabah Savings	71,396,076	72,153,690	69,638,823	71,450,975	91,217,012
Mudharabah Plus Savings	24,885	25,149	51	52	52
Mudharabah Savings iB Plan	14,742	14,899	23,530	24,143	24,143
Mudharabah Savings Simple IB	-	32.082	44,932	46,101	46,101
BSS Ceria Mudharabah Savings	-	4,320,098	2,782,153	2,854,550	2,854,550

Priority iB Savings			6,3 49,303	6,5 14,525	6, 514,52 5
Deposit	589,52 6,249	595,78 1,960	836,69 6,882	803.98 4728	801. 465.14 6
1 Month Mudharabah Deposit	412,53 5.868	416,91 3,460	719,06 9.582	690,95 6,277	686, 706.14 6
Mudharabah Deposit 3 Months	135,34 5,294	136,78 1,500	63,5 69,500	61,0 84,137	49, 631,00 0
Mudharabah Deposit 6 Months	5,9 71,633	6,0 35,000	6,6 59,000	6,3 98,655	6, 804,00 0
Mudharabah Deposit 12 Months	35,6 73,454	36,0 52,000	47,3 98,800	45,5 45,659	58, 324,00 0

In terms of fund raising, Table III. shows the total funds collected by Bank Sumut Syariah in 2015 amounted to Rp. 682.72 billion, then in 2019 it exceeded Rp. 1 Trillion. Of the three types of savings products, Savings products showed the lowest growth compared to Current Accounts and Time Deposits.

In general, Bank Sumut Syariah deposits have increased, but the increase in deposits has been dominated by deposits since 2015, while savings from 2015 to 2018 were higher than current accounts, in 2019 they were even lower than current accounts. Third party funds obtained from time deposits are funds that provide higher fees than savings and current accounts, so that banks certainly expect the increase in deposits from savings to be more dominant than deposits. Savings are deposits which can only be withdrawn according to certain agreed conditions, but cannot be withdrawn by check, bilyet, demand deposit, and or other equivalent instruments. If the customer wants to take his deposit, he can come directly to the bank with a passbook, withdrawal slip, or through an ATM facility. Deposit customers have more consideration factors than financing customers, who tend to only focus on the amount of the financing burden that must be paid. Financing customers are selected strictly by customers, in contrast to TPF customers who tend to act as selectors of the bank to be selected. This has an impact on bank policies to pay attention to many aspects in increasing the number of customers.

Sharia banks as banks that display religious symbols have a target market for Muslim community groups in the process of collecting and distributing funds. Religion is one factor that is highly considered by consumers to decide to become a customer of a sharia bank, there are still many reasons and other factors that can influence consumer considerations to want to become a customer of a sharia bank. One of them comes from the marketing strategy carried out by Islamic banks through the marketing mix strategy. Which includes product policies, prices, promotions, places and distribution channels, employee services, service processes, and the physical form of the Sharia bank office itself. So that from the marketing mix, customers can be influenced to want to relate to Islamic banks by becoming customers of Sharia banks. The marketing mix is considered to be able to influence consumer considerations to want to become customers of Islamic banks (Yulianto et al., 2010).

Research results Sartika et al. (2011) stated that there is a relationship between religious commitment and the decision to use the services of Islamic Bank products, the results of this study are reinforced by Metawa and Almossawi who state that the basic consideration of a customer choosing Sharia Bank products is religious commitment. In addition to religious commitment, the marketing mix is a factor that influences customers in choosing Islamic banks.

Payne (2007) states that the service marketing mix will affect consumer decisions in buying a product. Kustiningsih (2014) explains that the marketing mix which consists of products, prices, promotions, places, people, processes, and customer service, is one of the factors that causes someone to become a customer of a sharia bank. The marketing mix policies carried out by each Islamic bank operating in North Sumatra are very diverse and adaptive to the behavior of the community.

Decisions to choose savings products are actions that are directly involved in obtaining, determining products and services, including the process of making savings decisions and following these actions. Products and promotions as part of the marketing mix show a positive and significant effect between products and promotions partially and jointly on saving decisions (Nugroho et al., 2014).

The marketing mix as a form of implementing marketing strategies will also affect customer loyalty. The results showed that simultaneously and every element of the mix had a significant effect on consumer loyalty. The right marketing mix policy will have a positive effect on increasing consumer loyalty. The results of Tambunan and Nasution's research also show that it is simultaneously known that product, service, promotion, location, and credibility variables have a significant effect on customer decisions to save at the bank. Improving product and service quality, promotion, bank location and bank credibility will increase the number of savings customers (Tambunan and Nasution, 2013).

In the product marketing mix elements, there are product attribute components, which are elements that become the development or differentiator of a product, thus providing added value, benefits and being considered in making purchasing decisions.

Product attributes have a major influence on the buyer's perception of the product. In addition to distinguishing a product from other products, product attributes must also be able to be an attraction for consumers. This is because the physical attributes of the product bring various kinds of benefits that buyers need and want. The ability of management to position the product through the product attributes that are owned appropriately in the market is one of the determining factors for the success of a product in the market. If a product has product attributes or properties that are in accordance with what is expected by consumers, then the product is considered suitable by consumers. Such a product will be a successful product. In a study it was stated that the attributes of savings products influenced the decision to choose savings (Karmila, 2017).

From the description on the background, the development of Islamic banking in North Sumatra which continues to grow, but has not been optimally utilized by the savings products of the Sharia Bank of North Sumatra. The decision to choose Bank Sumut Syariah savings products is allegedly influenced by the marketing mix and religious commitment. The

elements of the marketing mix consist of product, price, distribution, promotion, participants, physical evidence, process and publication (Melewar and Saunders, 2000). If the decision to choose savings products increases, this indicates an increase in customer loyalty.

II. RESEARCH METHOD

Research is a systematic investigation process aimed at providing information to solve problems. The type of research used is correlational research. This research was conducted with the aim of detecting the extent to which variations in a factor are related or correlated with one or more other factors based on the correlation coefficient (Sinulingga, 2011).

A. Research Location

This research was conducted in all Branches of Bank of North Sumatra Sharia s Mangasi North Sumatra Province. The research period is 3 (three) months, starting from June to August 2021.

B. Parameter Measurement and Observation

1) Population and Sample

The population in this study are customers who have used the savings products of Bank Sumut Syariah. The sampling technique used is purposive sampling . Sampling by determining the criteria made by the researcher (Sugiyono, 2016) . In this study, the criteria set were customers who were over 18 years old and Muslim.

In this study, the sample size is based on the opinion of Hair (2008) , which states that the ratio between the number of subjects and the number of independent variables in multivariate analysis is recommended around 15 to 20 subjects. Based on the framework of thought there are as many as 5 (five) variables, so the number of subjects is at least 7 5 to 10 0 respondents. In this study, it is planned to distribute as many as 2 20 questionnaires, so that this number can meet the requirements that have been set.

2) Questionnaire Measurement Scale

The questionnaire was prepared using a Likert scale measurement scale. Questionnaires were distributed using Google Forms. This scale is a psychometric scale used in questionnaires and is one of the techniques that can be used in evaluating programs, planning policies and community attitudes. With a Likert scale, the variables to be measured are translated into indicator variables, the Likert scale is included in the Ordinal scale group. Then the indicator is used as a starting point for compiling instrument items which can be in the form of questions or statements. The answer to each instrument item using a Likert Scale has a gradation from very positive to very negative, which can be in the form of words including: Strongly Agree (SS), Agree (S), Neutral (N), Disagree (TS) and Strongly Disagree agree (STS). Likert scale is used to measure attitudes, opinions, and perceptions of a person or group of people about social phenomena.

C. Research Model

The Research Model is a systematic framework for thinking. The frame of mind this will be very helpful in preparing the mindset that leads to the conclusion , as shown in Fig. 1.

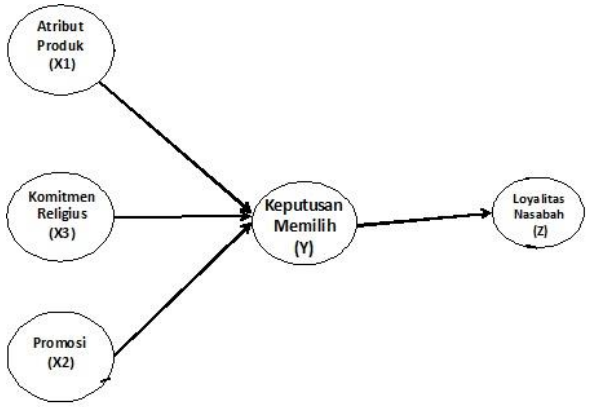


Fig. 1. Research Model

The concept of thought that is used as the basis for carrying out this research is by considering the existing conditions and the limitations of the problems discussed and processed with an approach of several mutually supportive disciplines so that a conclusion can be drawn that can be accounted for. This study consists of 3 (three) exogenous variables, namely Product Attributes (X 1), Promotion (X 2) and Religious Commitment (X 3) and the intervening variable in this study is the Decision to Choose Savings Product Services at Bank Sumut Syariah (Y) , while Customer Loyalty is an endogenous variable (Z) . Each independent variable in this study will determine how much influence it has on the decision to choose a savings product at Bank Sumut Syariah through correlation testing and the Structural Equation Model.

D. Variable Operational Definition

Operationalization of variables is an action in making boundaries that will be used in the analysis. In Table IV. the definitions and indicators of the research variables are described.

TABLE IV. VARIABLE OPERATIONAL DEFINITION

Variable	Definition	Dimension	Indicator
Attribute Product (X_1)	Any service or attribute that provides benefits to Sharia bank customers , both internally and externally	Quality product	According to sharia principles
			Easy account opening procedure
		Feature product	There is no element of usury
			Equipped with various facilities
		Style and Product Design	Atm card display is attractive
			Neat printouts or savings account prints
Promotion (X_2)	The form of communication used by banks to inform, persuade or remind people about Sharia bank service products	Advertising	Brochure Usage
			Use of Electronic Media
		Promotion Sale	Product socialization
			Sending product information to customers
		Connection	CSR Program

		Public	Community empowerment program
		Sale Personal	*Teller* service
			Service *customer service*
Commitment Religious (X_3)	There is an internalization of religious teachings within the individual. Internalization of religious teachings will foster confidence and obedience to the religion adopted.	Trust	Religion as the basis for making decisions
			Considerations of scholars or community leaders
		Practical	The form of obedience to religion
			Easy process to open an account
		Experience	Ask those who have used
			Ask the bank before deciding
		Knowledge	Mem read about usury
			Inquire about the product to the bank
		Ethical	Ethics of choosing a product
			Ethics in everyday life

TABLE V. OPERATIONAL DEFINITIONS OF VARIABLES (CONTINUED)

Variable	Definition	Dimension	Indicator
Decision Choose (Y)	The process of formulating various alternative actions to make a choice on one particular alternative to determine the choice of Sharia bank savings services	Needs introduction	Knowing various Sharia bank products
			Knowing the difference between conventional bank products and Sharia banks
		Information search	View *the* bank's *website* or social media
			Asking users of Sharia bank products
		Alternative evaluation	Comparing conventional bank products with Sharia banks
			Comparing products between Sharia banks
		Decision to choose	Choose after considering
			Choose after getting support from the closest people
		Post-Selecting behavior	Keep using
			Recommendations to those closest to you
Loyalty Customer (Z)	Customer loyalty after experiencing service is expressed in behavior to use the bank's services and reflects the existence of a long-term bond between the bank and the customer.	Keep saving	Carry out savings activities at least once a month
			Use an atm card at least once a month
		Using additional facilities	Using the main product support facilities
			Intend to use other Islamic bank products
		Do not use other Sharia bank products	Only use one sharia savings product
			Do not use
		It's not easy to be attracted to similar products from other Sharia banks	Frequently introduce products
			Be proud of the products used

E. Data Collection and Analysis Techniques

1) Data collection technique

- List of Questions (Questionnaire) given to the savings customers of Bank Sumut Syariah which were selected as samples using a Likert scale which is an Ordinal Scale group . Open questions were also given to customers of Sharia banks who were not customers of Bank Sumut Syariah, questions were asked about the reasons why customers chose the Sharia bank.
- Documentation Study, by collecting and studying data and information from news sources, Bank Sumut Annual Report .

2) Data Analysis

This research uses a quantitative approach in data analysis. Quantitative data analysis in this study consisted of descriptive statistical analysis and statistical inferential analysis, using Structural Equation Modeling (SEM) Prior to the discussion, an evaluation of the SEM assumptions was made, namely normality, outliers, and linearity tests. SEM is a structural equation model variable that is used to examine the relationship between complexes, both recursive and non-recursive to obtain a comprehensive picture of the entire model. (Trianto, 2016).

SEM is a statistical technique that has the ability to analyze the pattern of relationships between latent constructs and their indicators, latent constructs with each other, as well as direct measurement errors (Yamin dan Kurniawan, 2009).

There are two important reasons underlying the use of SEM. First, SEM has the ability to estimate the relationship between variables that are multiple relationships. This relationship is formed in a structural model (the relationship between the dependent and independent latent constructs). Second, SEM has the ability to describe the pattern of the relationship between the latent construct (unobserved) and the manifest variable (indicator variable) (Yamin dan Kurniawan, 2009).

III. RESULTS AND DISCUSSION

A. Data analysis method

In this study, the data analysis method used was structural equation modeling- partial least squares (SEM-PLS) using SmartPLS software . Mahmud and Ratmono (2013:6) stated that in its development, SEM was divided into two types, namely covariance-based SEM (CB-SEM) and variance-based SEM or partial least squares (SEM-PLS). CB-SEM developed in the 1970s pioneered by Karl Joreskog as a Lisrel software developer. Meanwhile, SEM-PLS developed after CB-SEM and was pioneered by Herman Wold (academic supervisor of Karl Joreskog). The following are some examples of software from CB-SEM and SEM-PLS (Mahmud dan Ratmono, 2013:6-7).

TABLE VI. SOME EXAMPLES OF SOFTWARE FROM CB-SEM AND SEM-PLS

Software CB-SEM	Software SEM-PLS
LISREL	SmartPLS
Amos	WarpPLS
EQS	PLS-Graph
Mplus	Visual-PLS
STATCAL	STATCAL

Mahmud and Ratmono (2013:7) stated that SEM-PLS can work efficiently with small sample sizes and complex models. In addition, the assumption of data distribution in SEM-PLS is relatively looser than that of CB-SEM. Estimation with CB-SEM requires a series of assumptions that must be met such as multivariate data normality, minimum sample size, homoscedasticity, and so on. Mahfud and Ratmono (2013:8) state that the estimation results of the two are not much different so that SEM-PLS can be a good proxy for CB-SEM. SEM-PLS can still produce estimates even for small sample sizes and deviations from the assumption of multivariate normality.

SEM-PLS can therefore be viewed as a nonparametric approach to CB-SEM. In addition, when the assumptions of CB-SEM are not met, then SEM-PLS can be the right method for theory testing.

Mahfud and Ratmono (2013:9-13) state that if the data meets the CB-SEM assumptions correctly, such as the minimum sample size and normal distribution, then choose CB-SEM. If not, select SEM-PLS. SEM-PLS is a nonparametric approach; can work well even for extreme abnormal data.

B. Outer Model Evaluation (Measurement Model) : Validity and Reliability Testing

Convergent validity is part of the measurement model which in SEM-PLS is usually referred to as the outer model while in covariance-based SEM it is called confirmatory factor analysis (CFA) (Mahfud and Ratmono, 2013:64). There are two criteria to assess whether the outer model (measurement model) meets the requirements of convergent validity for reflective constructs, namely (1) loading must be above 0.7 and (2) significant p- value (<0.05) (Hair et al. in Mahfud and Ratmono, 2013:65). However, in some cases, loading requirements above 0.7 are often not met, especially for newly developed questionnaires. Therefore, loading between 0.40-0.70 must be considered to be maintained (Mahfud and Ratmono, 2013:66). Indicators with loadings below 0.40 should be removed from the model. However, for indicators with loadings between 0.40 and 0.70, we should analyze the impact of the decision to delete these indicators on average variance extracted (AVE) and composite reliability . We can remove the indicator with a loading between 0.40 and 0.70 if the indicator can increase the average variance extracted (AVE) and composite reliability above its limit (threshold) (Mahfud and Ratmono, 2013:67). The limit value of AVE is 0.50 and composite reliability is 0.7. Another consideration in removing indicators is their impact on the content validity of the construct. Indicators with small loadings are sometimes maintained because they contribute to the validity of the construct content (Mahfud and Ratmono, 2013:67). Table VII and Fig. 2 present the loading values for each indicator.

TABLE VII. VALIDITY TEST BASED ON LOADING FACTOR

	Product Attributes (X1)	Decision to Choose Sharia Bank Products (Y)	Religious Commitment (X3)	Customer Loyalty (Z)	Promotion (X2)
X11	0.974				
X12	0.966				
X13	0.964				
X14	0.975				
X15	0.957				
X17	0.985				
X21					0.944
X22					0.950
X23					0.920
X24					0.963
X25					0.957
X26					0.924
X27					0.975
X28					0.939
X31			0.929		
X310			0.977		
X32			0.946		
X33			0.950		
X34			0.969		
X35			0.939		
X36			0.963		
X37			0.946		
X38			0.956		
X39			0.967		
Y1		0.962			
Y10		0.957			
Y2		0.934			
Y3		0.940			
Y4		0.897			
Y5		0.948			
Y6		0.963			
Y7		0.955			
Y8		0.942			
Y9		0.970			
Z1				0.957	
Z2				0.953	
Z3				0.896	
Z4				0.946	
Z5				0.948	
Z6				0.934	
Z7				0.919	

Z8				0.965	

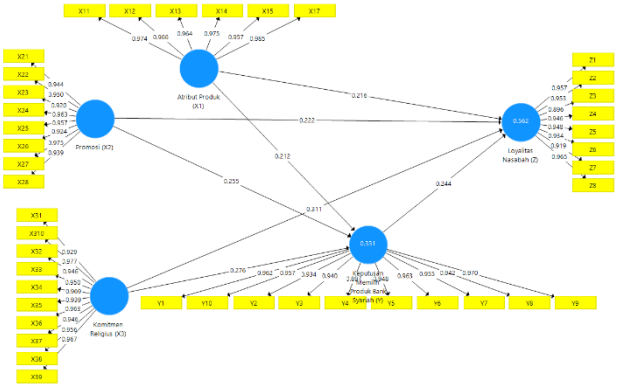


Fig. 2. Path Diagram based on Value of Loading Factor

Based on the validity test as loading factor in Table VII , it is known that all loading values are > 0, 7 , which means that they have met the validity requirements based on the loading value. Furthermore, validity testing is carried out based on the average variance extracted (AVE) value.

TABLE VIII. VALIDITY TESTING BASED ON AVERAGE VARIANCE EXTRACTED (AVE)

	Average Variance Extracted (AVE)
Product Attributes (X1)	0.941
Decision to Choose Islamic Bank Products (Y)	0.897
Religious Commitment (X3)	0.911
Customer Loyalty (Z)	0.884
Promotion (X2)	0.896

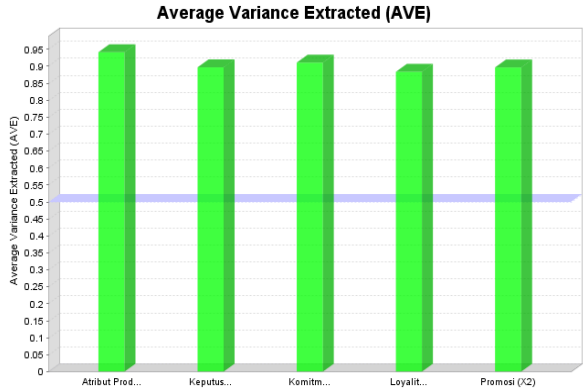


Fig. 3. Validity Testing based on Average Variance Extracted (AVE)

The recommended AVE value is above 0.5 (Mahfud and Ratmono, 2013:67). It is known that all AVE values are > 0.5, which means that they have met the validity requirements based on the AVE . Furthermore , the reliability test was carried out based on the composite reliability (CR) value.

TABLE IX. RELIABILITY TESTING BASED ON COMPOSITE RELIABILITY (CR)

	Composite Reliability
Product Attributes (X1)	0.990
Decision to Choose Islamic Bank Products (Y)	0.989
Religious Commitment (X3)	0.990
Customer Loyalty (Z)	0.984
Promotion (X2)	0.986

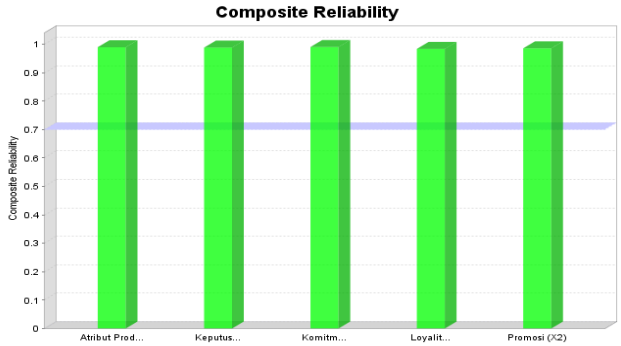


Fig. 4. Reliability Testing based on Composite Reliability (CR)

Value CR suggested is above 0, 7 (Mahfud and Ratmono, 2013: 67). It is known that all CR values are > 0, 7 , which means that they have fulfilled the reliability requirements based on CR. Furthermore, reliability testing was carried out based on the value of Cronbach's alpha (CA).

TABLE X. RELIABILITY TESTING BASED ON CRONBACH'S ALPHA (CA)

	Cronbach's Alpha
Product Attributes (X1)	0.988
Decision to Choose Islamic Bank Products (Y)	0.987
Religious Commitment (X3)	0.989
Customer Loyalty (Z)	0.981
Promotion (X2)	0.983

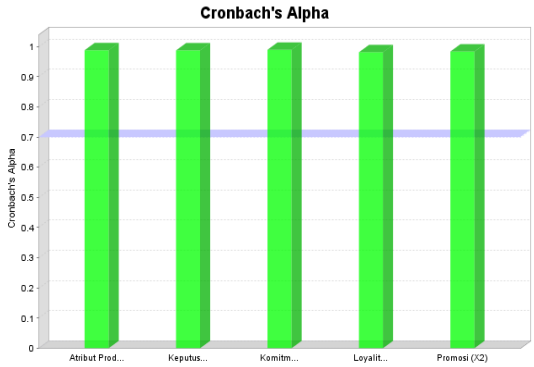


Fig. 5. Reliability Testing based on Cronbach's Alpha (CA)

Value CA suggested is above 0, 7 (Mahfud and Ratmono, 2013: 67). Known throughout the value CA > 0, 7 , which means it meets the requirements of reliability by Cronbach's alpha.

C. Effect Significance Test (Hypothesis Testing)

Table XI. presented the results for hypothesis testing.

TABLE XI. HYPOTHESIS TESTING

	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)	P Values

Product Attributes (X1) -> Decision to Choose Sharia Bank Products (Y)	0.212	0.208	0.067	3.145	0.002
Product Attributes (X1) -> Customer Loyalty (Z)	0.216	0.213	0.055	3.940	0.000
Decision to ChooseSharia Bank Products (Y) -> Customer Loyalty (Z)	0.244	0.239	0.062	3.938	0.000
Religious Commitment (X3) -> Decision to Choose Sharia Bank Products (Y)	0.276	0.272	0.071	3.868	0.000
Religious Commitment (X3) -> Customer Loyalty (Z)	0.311	0.315	0.068	4.585	0.000
Promotion (X2) -> Decision to Choose Sharia Bank Products (Y)	0.255	0.256	0.064	4004	0.000
Promotion (X2) -> Customer Loyalty (Z)	0.222	0.221	0.055	4.026	0.000

$$Y = 0{,}212X1 + 0{,}255X2 + 0{,}276X3 \quad (1)$$

$$Z = 0{,}216X1 + 0{,}222X2 + 0{,}311X3 + 0{,}244Y \quad (2)$$

Based on the results in Table XI, the results obtained are:

- The product attribute (X1) has a positive effect on the decision to choose Sharia bank products (Y) with a path coefficient value of 0.212 and is significant with a P-Values = 0.002 <0.05.
- Product attributes (X1) have a positive effect on customer loyalty (Z), with a path coefficient value of 0.216 and significant with a P-Values = 0.000 <0.05.
- Promotion (X2) has a positive effect on the decision to choose Sharia bank products (Y), with a path coefficient value of 0.255 and significant with P-Values = 0.000 <0.05.
- Promotion (X2) has a positive effect on customer loyalty (Z), with a path coefficient value of 0.222 and significant with a P-Values = 0.000 <0.05.
- Religious commitment (X3) has a positive effect on the decision to choose Sharia bank products (Y) with a path coefficient value of 0.276 and is significant with a P-Values = 0.000 <0.05.
- Religious commitment (X3) is positive on customer loyalty (Z), with a path coefficient value of 0.311 and significant with a P-Values = 0.000 <0.05.
- The decision to choose Sharia bank products (Y) has a positive effect on customer loyalty (Z), with a path coefficient value of 0.244 and significant with a P-Values = 0.000 <0.05.

Table XII presents the results of the coefficient of determination (r- square).

TABLE XII. VALUE OF COEFFICIENT OF DETERMINATION

	R Square
Decision to Choose Sharia Bank Products (Y)	0.331
Customer Loyalty (Z)	0.562

Based on Table XII:

- It is known that the coefficient of determination for the decision variable to choose Sharia bank products (Y) is 0.331, which means product attributes (X1), promotion (X2), religious commitment (X3) are able to influence the decision to choose Sharia bank products (Y) of 33.1 %.
- It is known that the coefficient of determination for the customer loyalty variable (Z) is 0.562, which means product attributes (X1), promotion (X2), religious commitment (X3), the decision to choose Sharia bank products (Y) can affect customer loyalty (Z) by 56.2%.

IV. CONCLUSION

From the results of the discussion, it is concluded:

- The product attribute (X1) has a positive effect on the decision to choose Sharia bank products (Y) with a path coefficient value of 0.212 and is significant with a P-Values = 0.002 <0.05.
- Product attributes (X1) have a positive effect on customer loyalty (Z), with a path coefficient value of 0.216 and significant with a P-Values = 0.000 <0.05.
- Promotion (X2) has a positive effect on the decision to choose Sharia bank products (Y), with a path coefficient value of 0.255 and significant with P-Values = 0.000 <0.05.
- Promotion (X2) has a positive effect on customer loyalty (Z), with a path coefficient value of 0.222 and significant with a P-Values = 0.000 <0.05.
- Religious commitment (X3) has a positive effect on the decision to choose Sharia bank products (Y) with a path coefficient value of 0.276 and is significant with a P-Values = 0.000 <0.05.
- Religious commitment (X3) is positive on customer loyalty (Z), with a path coefficient value of 0.311 and significant with a P-Values = 0.000 <0.05.
- The decision to choose Sharia bank products (Y) has a positive effect on customer loyalty (Z), with a path coefficient value of 0.244 and significant with a P-Values = 0.000 <0.05.

REFERENCES

[1] Al Fauzan, Syaikh Shalih bin Fauzan , 2018, *Loyalitas dalam Islam*, Akafa Press, Jakarta

[2] Alma,Buchari, 2007, *Manajemen Pemasaran dan Pemasaran Jasa*, Edisi Revisi,. CV. Alfabeta, Bandung

[3] Antonio, M. Syafi'i, 2001, Bank Syariah Dari Teori Kepraktik, 2001, Gema Insani. Press, Jakarta

[4] Arikunto, Suharsimi, 2006, *Prosedur Penelitian Suatu Pendekatan Praktik*, Rineka Cipta, Jakarta

[5] Asnawi, Nur dan Muhammad Asnan Fanani, 2017, *Pemasaran Syariah ; Teori, Filosofi dan Isu-Isu Kontemporer*, Rajawali Pers, Jakarta

[6] Frida, Sri Waqoyah Nining, 2004, *Mengklasifikasikan Konsumen Dengan Analisis Cluster Dalam Menentukan Market Segnentation (Studi Kasus pada PT. Alfa Retailindo Surakarta)*, Tesis, Universitas Muhammadiyah Surakarta, Solo

[7] Ferdinand, Augusty, 2006, *Structural Equation Modelling dalam Penelitian Manajemen*, Badan Penerbit Universitas Diponegoro, Semarang

[8] Ghozali, Imam, 2012, *Aplikasi Analisis Multivariate dengan Program SPSS*, Badan Penerbit Universitas Diponegoro, Semarang

[9] Ghozali, Imam. 2013. *Aplikasi Analisis Multivariate dengan Program IBM SPSS19*, Penerbit Universitas Diponegoro, Semarang

[10] Gio, P.U., Elly Rosmaini, 2015, *Belajar Olah Data dengan SPSS, MINITAB, R, MICROSOFT EXCEL, EVIEWS*, LISREL, AMOS, dan SMARTPLS, USU Press

[11] Gio, P. U., & Caraka, R. E., 2018, *PEDOMAN DASAR MENGOLAH DATA DENGAN PROGRAM APLIKASI STATISTIKA STATCAL*

[12] Gio, P. U., & Caraka, R. E., 2021, *Regresi logistik biner dengan STATCAL Disertai Perbandingan Hasil dengan SPSS*, Stata, EViews & JASP, USUpress

[13] Griffin, Jill, 2013, *Customer Loyalty*, Erlangga, Jakarta

[14] Hair dkk., 2013, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, Sage

[15] Hair, J.F. JR., Anderson, R.E, Tatham, R.L. & Black, W.C., 1998, *Multivariate Data Analysis*, Fifth Edition, Prentice Hall, New Jersey

[16] Hasibuan, Melayu SP., 2006, *Dasar-Dasar Perbankan*, Bumi Aksara, Jakarta

[17] Karim, Adiwarman, 2004, *Bank Islam: Fiqih dan Keuangan*, Gema Insani Press, Jakarta

[18] Karim, Adiwarman, 2008, *Bank Islam*, Kencana, Jakarta

[19] Kertajaya, Hermawan dan Muhammad Syakir Sula, 2008, *Syari'ah Marketing*,Cetakan ke-4, Mizan, Jakarta

[20] Kotler, Philip. 2003, *Manajemen Pemasaran*, Terjemahan: Hendra Teguh, Ronny A. Rusli dan Benjamin Molan, Edisi Milenium, Penerbit PT. Indeks, Jakarta

[21] Kotler, Philip *dan Gary* Armstrong, 2008, *Prinsip – Prinsip Pemasaran*, (Alih bahasa Bob. Sabran), Jilid 1. Edisi keduabelas, Erlangga, Jakarta

[22] Lovelock, Christopher, Jochen Wirtz, Jacky Mussry, 2010, *Pemasaran Jasa ; Manusia, Teknologi, Strategi, Perspektif Indonesia*, Alih Bahasa oleh Dian Wulandari dan Devri Barnadi Putera, Erlangga, Jakarta

[23] Lupiyoadi, Rambat, 2013, *Manajemen Pemasaran Jasa: Teori dan Praktek*, Edisi Pertama, Penerbit Salemba Empat, Jakarta

[24] Mahfud Sholihin & Dwi Ratmono, 2013, *Analisis SEM-PLS dengan WarpPLS 3.0*, Penerbit Andi

[25] Payne, Adrian, 2007, *The Essence of Service Marketing*, Diterjemahkan oleh Fandy Tjiptono, Edisi II, Penerbit Andi, Yogyakarta

[26] Poerwadarminta, W.J.S, 2007,*Kamus Umum Bahasa Indonesia*. Cetakan ke-10. Balai Pustaka, Jakarta.

[27] Priyatno, Dwi, 2008, *Mandiri Belajar SPSS Untuk Analisis Data & Uji Statistik*, MediaKom, Jakarta

[28] Santoso, Muhammad, 2015, *Loyalitas Nasabah PD Bank Perkreditan Rakyat Wilayah Cirebon*, Deepublish, Yogyakarta

[29] Sekaran, Uma, 2006, *Metodologi Penelitian Bisnis*, Edisi 4, Buku 1, Penerjemah : Kwan Men Yon, Salemba Empat, Jakarta

[30] Sinulingga, Sukaria, 2011, *Metode Penelitian*, USU Press, Medan

[31] Singarimbun, Masri dan Sofian Effendi. 1995. *Metode Penelitian Survey*. LP3ES, Jakarta

[32] Situmorang, Syafrizal Helmi dan Muslich Lufti, 2008, *Analisis Data Penelitian (Menggunakan Program SPSS)*, USU Press, Medan

[33] *Soemitra, Andri*, 2010, *Bank &Lembaga Keuangan Syariah*, Kencana. Prenada Media Group, Jakarta

[34] Stanton, William J., 2003, *Pemasaran*, Alih Bahasa oleh Sadu Sundaru. Jilid Satu. Edisi Kesepuluh. Jakarta : Erlangga, Jakarta

[35] Sumarwan, Ujang, 2012, *Perilaku Konsumen*, Ghalia, Bandung

[36] Sugiyono, 2006, *Metode Penelitian Bisnis*, Alfabeta, Bandung

[37] Suyatno,Thomas, 2004, Kelembagaan Perbankan, 2004, Edisi Kedua, Gramedia Pustaka Utama, Jakarta

[38] Swastha, Basu,2006, *Azas-Azas Marketing*, Edisi Ketiga, Penerbit BPFE, Yogyakarta

[39] Tjiptono, Fandy, 2002, *Strategi Pemasaran*, Penerbit Andi, Yogyakarta

[40] Trianto, Budi, 2016, *Riset Modeling*, Edisi Kedua, Adh Dhua Insstitute, Pekanbaru

[41] Umar, Husein, 2005. *Metode Riset Bisnis*, CetakanPertama, Gramedia Pustaka Utama, Jakarta

[42] *Yazid*, 2003, *Pemasaran* Jasa, Edisi Kedua, Ekonisia Fakultas Ekonomi UII,. Yogyakarta

[43] Yamin, Sofyan dan Heri Kurniawan,2009,*Structural Equation Modeling dengan Lisrel – PLS*, Salemba Empat, Jakarta

[44] Alserhan, Baker Ahmad, 2011, *The Principles of Islamic Marketing*, Gower PublishingLimited, England

[45] Caraka, Rezzy Eko, Youngjo Lee, Rung Ching Chen, Toni Toharudin, Prana Ugiana Gio, Robert Kurniawan, and Bens Pardamean. "*Cluster Around Latent Variable for Vulnerability Towards Natural Hazards*, Non-Natural Hazards, Social Hazards in West Papua." IEEE Access 9 (2020): 1972-1986

[46] Cronin Jr., J. Joseph and Steven A. Taylor, *Measuring Service Quality: A Reexamination and Extension*, Vol. 56, No. 3,Journal of Marketing

[47] Dharmayanti, Diah, 2016, *Analisis Dampak Service Performance dan Kepuasan Sebagai Moderating Variable Terhadap Loyalitas Nasabah*, Jurnal Manajemen Pemasaran,1006, Vol 1. No. 1, April

[48] Dian Utami Sutiksno et al 2018 J. Phys.: Conf. Ser. 1028 012244

[49] Firmansyah, 2008, *Faktor-Faktor yang Berhubungan Dengan Loyalitas Nasabah Bank Syariah*, Fakultas Syariah dan Hukum, Universitas Islam Negeri Syarif Hidayatullah, Jakarta

[50] Foster, Bob, 2010, *Pengaruh Kinerja Bauran Pemasaran Jasa Terhadap Loyalitas Nasabah Penabung*, Strategic, Volume 9, Nomor 17, Februari, Jakarta

[51] Khotimah, Heniffa Nurul, 2013, Analisis Pemilihan Bank Syariah Dengan Pendekatan *Analythical Hierarchy Process***,** Jurnal Akuntansi Unesa | Vol 1, No 2**,** Semarang

[52] Kustiningsih, Eka Wundi Lilis, 2014, *Variabel-Variabel Yang Mempengaruhi Keputusan Menjadi Nasabah Tabungan Faedah Bank BRI Syariah Cabang Samarinda*, 2014, eJournal Ilmu Administrasi Bisnis, Samarinda

[53] Mahyuni dan M. Riza Firdaus, 2014, *Analisa Faktor-Faktor yang Mempengaruhi Loyalitas Nasabah Pada Bank Muamalat Indonesia Cabang Banjarmasin*, 2014, Jurnal Wawasan Manajemen, Vol. 2, Nomor 2. Bulan Juni, Banjarmasin

[54] Maski, Ghozali, 2010, *Analisis Keputusan Nasabah Menabung: Pendekatan Komponen dan Model Logistik Studi Pada Bank Syariah di Malang*, Journal of Indonesian Applied Economics, Vol. 4 No. 1 Mei, Malang

[55] Melewar, T.C. and John Saunders, 2000,*Global Corporate Visual Identity Systems : Using an Extended Marketing Mix*,European Journal of Marketing, Vol. 34 No. 5/6, MCB University Press

[56] Metawa, Saad A. and Mohammaed Almossawi, 1998, *Banking behavior of Islamic bank customers: perspectives and implications*, International Journal of Bank Marketing, MCB University Press, Hlm. 2

[57] Muhmin, Abdul, 2002, *Effects of Supplier's Marketing Program Variables on Indistrial Buyer's Relationship Satisfaction on Commitment*, Journal of Bussines & Indistrial Marketing, Vol. 7, p 637-651

[58] Nugroho, Robertus Andy, Nawazirul Lubis, Apriatni EP. 2014. Pengaruh Produk dan Promosi Terhadap Keputusan Menabung Tabungan Simpedes di PT. Bank Rakyat Indonesia (Persero) Tbk. Cabang Semarang Pattimura. JIAB, Vol. 3 Nomor 4. Semarang

[59] Nurhidayah, 2013, *Analisis Kualitas Jasa Terhadap Kepuasan dan Pengaruhnya Terhadap Loyalitas Nasabah Debitur Kupedes Pada Bank BRI Unit Banjar Selatan Cabang Banjarmasin*, Jurnal Spread, Volume 3 Nomor 2 Bulan Oktober, Banjarmasin

[60] Rahmatullah, 2009, Pengaruh Komunikasi Pemasaran Terhadap Loyalitas Nasabah (Studi Kasus Bank XYZ Syariah Cabang Bogor), Islamic Finance & Business Review Vol. 4 No. 2 Agustus-Desember, Bogor

[61] Rejeki, Denik Sri, Achmad Fauzi dan Edy Yulianto, 2015, Pengaruh Green Marketing Pada Keputusan Pembelian dan Implikasinya Terhadap Loyalitas Pelanggan, Jurnal Administrasi Bisnis JA, Vol. 26 No. 1 September, Malang

[62] Roziq, Ahmad dan Rinanda Fitri Diptyanit, 2013, Variabel Penentu dalam Keputusan Memilih Tabungan *Mudharabah* Pada Bank Syariah Mandiri Cabang Jember, JEAM Vol XII No. 1/2013, Jember

[63] Salim, PeterdanYennySalim, 2002,*KamusBahasaIndonesiaKontemporer*,Modern EnglishPress,Jakarta

[64] Sartika, Dewi, Ali Mubarak, dan Indari Larasati, 2011, *Hubungan Antara "Religious Commitment" dengan Keputusan Mengunakan Jasa Bank Syariah pada Dosen Unisba*, Prosiding Seminar Nasional Penelitian dan PKM: Sosial, Ekonomi, dan Humaniora

[65] Saulatya, Syarifah, 2010, *Pengaruh Atribut Produk Islam Tabungan Muamalat Umroh Terhadap Loyalitas Nasabah Bank Muamalat Cabang Bandung*, repository.unisba.ac.id, Bandung

[66] Inayah, Nurul dan Sri Sudiarti,2017,*Analisis keputusan nasabah menabung di Bank Syariah*,At-Tawassuth, Vol. 2, No. 1, 2017: 191-214, Medan

[67] Djunaedi, 2016, *Pengaruh Corporate Sosial Responsibility (CSR), dan Kualitas produk terhadap citra bank keputusan menabung di BNI Syariah Kota Kediri*,JMM17 Jurnal Ilmu Ekonomi & Manajemen, September 2016, Vol. 3 No.2. hal. 104 – 119

[68] Shofwa, Yoiz, 2016, *Pengaruh kualitas produk dan religiusitas terhadap keputusan nasabah produk simpanan pada BSM cabang Purwokerto*,el-JIZYA, Jurnal Ekonomi Islam (Islamic Economic Journal), Vol. 4, No. 1, Januari-Juni 2016

[69] Ali, Karnila. 2011. Pengaruh Promosi Dan Atribut Produk Terhadap Keputusan Nasabah Dalam Memilih Tabungan Ekasave Pada Bank Eka Kantor Pusat Metro.

[70] Ika, Dewi Nur, 2013, Pengaruh Kualitas Layanan, Nilai Nasabah, Dan Atribut Produk Islam Terhadap Kepuasan Nasabah.

[71] Limbong, Novi Septiani dan Ani Solihat. 2013. Pengaruh atribut produk tabungan Terhadap Loyalitas Nasabah,Jurnal Ecodemica, Vol 1. No.2., Jakarta.

[72] Rahmatullah, 2009, Pengaruh komunikasi pemasaran terhadap loyalitas nasabah, Tazkia, Vol 4, No. 2, Agustus-Desember 2009.

[73] Hidayati, Amalia Nuril, 2012. Pengaruh faktor Marketing Mix dan faktor syariah terhadap Keputusan menjadi Nasabah Lembaga Keuangan Mikro Syariah

[74] Anabas, Safitri, Pengaruh produk, harga, promosi, tempat, proses, orang, dan bukti fisik terhadap loyalitas nasabah Tabungan Simpedes pada PT.Bank Rakyat Indonesia (Persero) Tbk. Cabang Sintang, Jurnal Manajemen Motivasi 673-679

[75] Akbar, Dinul Arfian, 2016, *Bauran pemasaran jasa terhadap loyalitas nasabah*, I-Economic, Vol.2, No.1, Juli 2016

[76] Irmansyah, Noor, Emy Rahmawati dan Maryono, 2015, Pengaruh Bauran Promosi Terhadap Keputusan Nasabah Menabung Simpeda pada Bank Kalsel Unit Gambut, Jurnal Bisnis dan Pembangunan, Vol 4, No.1, Juli-Desember 2015, Banjarmasin

[77] Koyong, Melka Dodai, Sontje Somayku dan Jhony Kalangi, 2016, *Pengaruh promosi terhadap loyalitas nasabah di PT. Bank Mandiri Kantor Cabang Dotulolong Lasut Manado*, Jurnal Administasi Bisnis, Manado

[78] Nugroho, Robertus Andy, Nawazirul Lubis dan Apriatni EP. 2016. *Pengaruh produk dan promosi terhadap keputusan menabung Tabungan Simpedes di PT Bank Rakyat Indonesia (Persero) Tbk. Cabang Semarang Pattimura*, Ambon

[79] Krismanto, Adi, 2009, *Analisis Faktor-Faktor yang Mempengaruhi Loyalitas Nasabah (Studi Kasus Pada PT. Bank Rakyat Indonesia Cabang Semarang Pattimura)*, http://eprints.undip.ac.id/16154/1/Adi_Krismanto.pdf, Semarang

[80] Astuti, Ambar Kusuma dan Agustini Dyah Respati, 2015, *Pengaruh Promosi Dan Kinerja Pelayanan Terhadap Loyalitas Nasabah Dengan Kepuasan Nasabah Sebagai Pemoderasi Pada Bank BUMN di DIY*, Fakultas Bisnis, Universitas Kristen Duta Kencana, Yogyakarta

[81] Selang, Christian A.D., 2013, *Bauran Pemasaran Pengaruhnya Terhadap Loyalitas Konsumen Pada Fresh Mart Bahu Mall Manado* Jurnal EMBA, Vol.1No.3Juni, Manado

[82] Selnes, Fred, 2003, *An Examination of The effect of Product performance on Brand Reputation, Satisfaction and Loyalty*, European Journal of Marketing Vol 27, No. 9, p 19-35

[83] Tambunan, Monang Ranto dan Inggrita Gusti Sari Nasution, 2013, Analisis Faktor-Faktor Yang Mempengaruhi Keputusan Nasabah Menabung di Bank BCA Kota Medan, Jurnal Ekonomi dan Keuangan, Vol.1 No. 3. Medan

[84] Wattimury, Ferry, 2012, Analisis Kualitas Pelayanan Terhadap Loyalitas Nasabah Pada Bank Mandiri Cabang Pattimura Ambon, Jurnal Benchmark, Volume 1 No. 1 November, Ambon

[85] Yulianto, Firman dan Agung Yuniarinto dan Surachman, 2010, Analisis Pengaruh Faktor Bauran Pemasaran Terhadap Pertimbangan Nas abah dalam Memilih Bank Syariah di Kota Medan. Jurnal Wacana, Vol. 13 No.4, Medan

[86] Erfandi, 2009, *Pengetahuan dan Faktor-Faktor yang Mempengaruhi*, http://forbetterhealth.wordpress.com/2009/04/19/pengetahuan-dan-faktor-faktor-yang-mempengaruhi/, diakses pada tanggal 10 Desember 2016

[87] Fahrudin, Ari, 2014, *Pengaruh Kualitas Produk, Harga Dan Keputusan Pembelian Terhadap Loyalitas Konsumen Handphone Blackberry*, https://eprints.unisbank.ac.id/2311/, diakses pada tanggal 17 Desember 2016

[88] Kartika Sari dan Budi Prijanto, *Pengaruh Tingkat Loyalitas Nasabah Pada Dana Pihak Ketiga Industri Perbankan di Indonesia*, https://eprints.unisbank.ac.id/191/1/artikel-27.pdf, pada tanggal 21 Desember 2016, jam 18.31 Wib

[89] http://m.tempo.co/read/news/2012/10/04/087433709/Pertumbuhan-Lambat-Bank-Syariah-Perlu-Dukungan, 2012, Pertumbuhan Lambat, Bank Syariah Perlu Dukungan

[90] http://bisnis.news.viva.co.id/news/read/59103-bank-syariah-perlu-ekspansi-ke-infrastruktur, 2009, Ekonomi Syariah Perlu Ekspansi ke Infrastruktur

[91] https://www.mandirisyariah.co.id/news-update/edukasi-syariah/kenapa-harus-beralih-ke-bank-syariah-ini-alasanya

[92] https://www.cnbcindonesia.com/market/20200630194002-17-169235/kejar-provokator-rush-money-bukopin-ojk-libatkan-kepolisian

[93] https://www.simulasikredit.com/apa-itu-rush-money-dan-mengapa-berbahaya-bagi-dunia-perbankan/

[94] http://news.unair.ac.id/2020/01/24/kualitas-layanan-dan-loyalitas-pelanggan-pada-perbankan-syariah-di-indonesia/

Analysis of The Influence of Leadership Style on Job Achievement at Madani Shariah Medan Hotel with Motivation as Intervening Variable

1st Hubbul Wathan
Accounting Department
Politeknik Negeri Medan
Medan, Indonesia
wathanhubbul@polmed.ac.id

2nd Muhammad Zuardi
Accounting Department
Politeknik Negeri Medan
Medan, Indonesia

3rd Rehulina Bangun
Accounting Department
Politeknik Negeri Medan
Medan, Indonesia

***Abstract*—Hotel Madani Syariah is one of the hotels that uses management principles based on sharia. In the management aspect, the organization has the principle of mutual benefit between the leader and the led. This study aims to determine the performance of employees at the Madani Syariah Hotel Medan. Leaders have the right to be obeyed as long as they do not violate the Shari'a, as well as employees have the right to be prospered with compensation according to their field of work. Employees as important assets in the success of the company must be maintained and maintained like company assets. Human resources are the main factor that is very important for an organization. Human resources cannot be separated from the process of how to achieve organizational goals and human resources can support the achievement of these goals cannot be separated from how the leadership style is applied to create employee motivation and produce good work performance. The conclusion in this research are the leadership style applied by the manager of Hotel Madani Syariah Medan tends to be a leadership style called the inviting leader as well, with high motivation and achievement of employees. As for the advice, the direct supervisor of Hotel Madani Syariah Medan should pay more attention to the feelings of subordinates in actions and actions.**

Keywords—Leadership Style, Work Motivation, Work Achievement, Performance

I. INTRODUCTION

To achieve a company's goals, the alignment of human resources and objectives must be considered. How the management of human resources supports the achievement of the company's targets cannot be separated from the influence of the leaders in the company. Leaders must be able to give good influence to human resources both for individuals and for task completion. A leader cannot be separated from the applied leadership style. The leadership style essentially aims to encourage employee work motivation, employee achievement, and high employee work productivity, in order to achieve maximum organizational goals (Rivai, 2014). Good influence in the application of leadership style will have an impact on employee motivation in completing tasks and in achieving company goals.

Motivation can be said as a driving force from within and within the subject to carry out certain activities in order to achieve a goal. Even motivation can be interpreted as an internal condition (preparedness). Starting from the word "motive", then motivation can be interpreted as a driving force that has become active. The motives become active at certain times, especially when the need to achieve goals is felt/urgent. In this case the better the leadership style in a company with the support of good motivation, the employee's work achievement also increases so as to achieve the goals desired by the company (Robinns, 2006:238). Work achievement is defined as a person's success in carrying out a job and is the result to be achieved by someone according to the size applicable to the job in question. In general, work performance is influenced by two things, namely individual factors and situational factors (Daft, 2006: 371).

Work achievement is an evaluation of the work carried out through direct superiors, co-workers, oneself or direct subordinates. To get work achievement in a company, motivation and support from the company's leadership style in the workplace must be needed. To get the job achievement is not easy employee must show the best ability to corporate leaders, then with the style of good leadership in leading companies with good support to employees result is a good achievement and in accordance with the objectives desired by these companies (Simamora, 2014 :339).

Hotel Madani Syariah is one of the hotels that uses management principles based on sharia. In the management aspect, the organization has the principle of mutual benefit between the leader and the led. Leaders have the right to be obeyed as long as they do not violate the Shari'a, as well as employees have the right to be prospered with compensation according to their field of work. Employees as important assets in the success of the company must be maintained and maintained like company assets. Human resources are the main factor that is very important for an organization. Human resources cannot be separated from the process of how to achieve organizational goals and human resources can support the achievement of these goals cannot be separated from how the leadership style is applied to create employee motivation and produce good work achievement (Mathis and Jackson, 2006: 3).

In this chapter the author will analyze the research data with the aim of knowing how the relationship between leadership style, motivation and employee achievement is. The research data were obtained by distributing questionnaires to non-managerial permanent employees at

Madani Syariah Hotel Medan. An initial survey on employee Syariah Hotel Madani Medan in January 2021, the result of data with the number of employees as much as 5 2 men and women who work at the hotel. The results of interviews with employees as many as 10 people consisting of 7 men and 3 women about leadership style, motivation and employee achievement. Employees stated that they were not satisfied with the leadership style carried out by the hotel because the leader rarely came to visit the hotel and chat with the hotel employees, but employees were motivated to work at the hotel because these employees had to earn a living for their families and employees also had to demonstrate their work achievement so that the employee can be promoted.

Based on the background of the problem employee performance is not optimal, a research will be conducted on the influence of leadership style on employee achievement at Madani Syariah Hotel Medan with motivation as an intervening variable.

The framework of thought is a conceptual model of how the theoretical basis that has been described relates logically to various factors identified as important problems (Sekaran, 2006). A good model can explain the relationship between research variables, namely independent variables and dependent variables (Ferdinand, 2006). The following is the research design (framework) of this study.

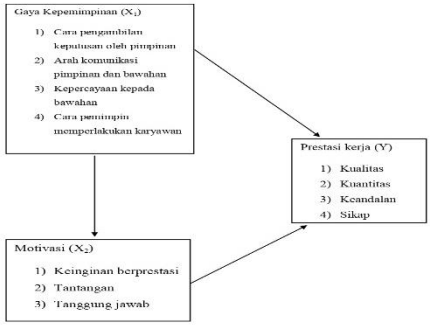


Fig. 1. The framework

II. RESEARCH METHODS

A. Research Type

Type of research in this study is research quantitative descriptive, the study aimed to describe the state of an independent variable (Sugiyono, 2007) . Approach to research through surveys, research survey is an investigation conducted research to get the facts of the existing symptoms and seek factual information to get to the truth (Sinuli n gga, 2011).

B. Population

The population in this study were employees working at the Madani Syariah Hotel during the research period, the number of employees working until April 30, 2021 was 110 employees. The type of sample used is a saturated sample, or the entire population in this study. Based on expert opinion as proposed by Gay in Hasan (2014) "The minimum sample size acceptable can be seen based on the design or the research methods used. If the research design is descriptive, then the minimum sample is 30".

C. Flowchart Research

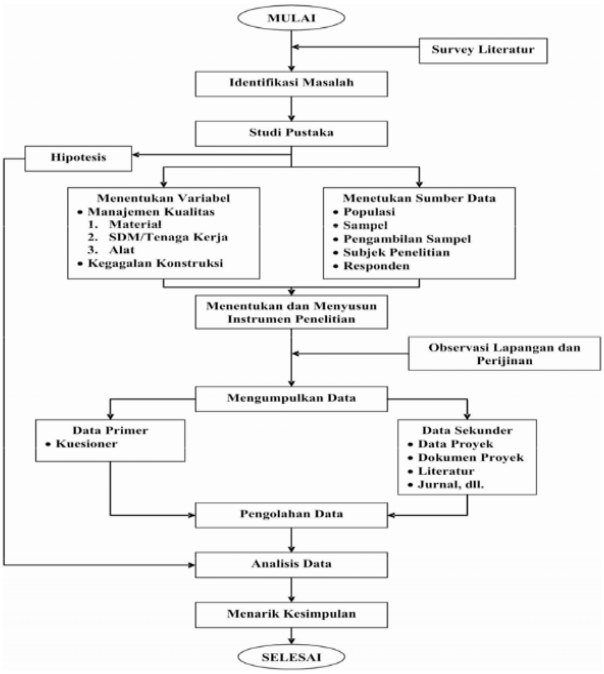


Fig. 2. Flowchart

The research stages are described in general as follows:

- Literature Survey

This stage is collecting literature and information related to the research title.

- Identification of problems

Identify what problems will be discussed related to quality management and construction failure based on the literature and information that has been obtained.

- Literature review

Studying the literature that will be used as a theoretical study in this research.

- Hypothesis

The initial question raised is whether there is a relationship between quality management and construction failure and how big the relationship is.

- Determining Variables and Data Sources

Determine the variables of quality management and construction failure with management aspects constrained, namely human resources, materials and equipment. Then determine what kind of data is needed based on the population, sample and sampling method. Then determine the research subjects and respondents.

- Determine and Develop Research Instruments (Questionnaire)

This stage is the determination of the research instrument by using a questionnaire. The preparation of this questionnaire is divided into 4 parts, namely the identity of the data source, qualitative, quantitative and essays. Then compiled in a google form to be distributed online.

- Field Observation and Licensing

Field observations were carried out with resource persons, namely employees of Hotel Madani Syariah Medan.

- Collecting data

Distribute questionnaires to respondents. This is done in conjunction with observation and licensing to save time, cost and effort.

- Data processing

Data processing consisted of giving variable codes, tabulations, calculations using the SPSS version 24 program and then a second tabulation were carried out.

- Data analysis

Analyze the results of data processing based on the results of research and existing theories.

- Draw a conclusion

Conclusions are drawn based on data analysis and checked whether they are in accordance with the aims and objectives of the study.

D. Data Analysis

The data analysis carried out in this study is path analysis, the intervening variable which is an intermediate variable that functions to connect the independent variable with the dependent variable. The research path diagram is described in Figure 3.

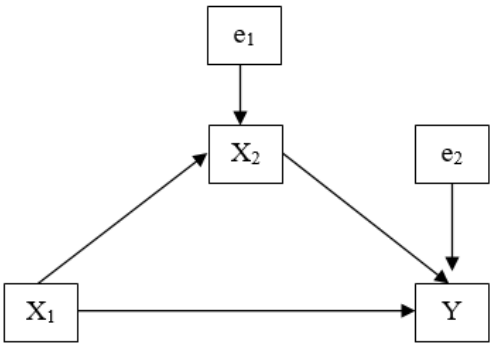


Fig. 3. Research Path Diagram

Description :

X 1 = Leadership Style

X 2 = Motivation

Y = Work Performance

Based on Fig. 3 the path analysis model can be explained, that leadership style can have a direct effect on work performance, but can also have an indirect effect, namely through the motivation variable first. To test the intervening variables used path analysis.

III. RESULTS AND DISCUSSION

A. Research Result

This research was conducted during July 2021. In this study, the data used were 32 employees at Madani Syariah Hotel Medan. The questionnaire consists of 3 groups of statements, namely group I to obtain data on leadership style, group II to obtain data on motivation and group III to obtain data on employee performance which is filled out by the direct supervisor.

1) Characteristics of Respondents

a) Gender

The table below shows the gender composition of Hotel Madani Syariah Medan employees.

TABLE I. GENDER FREQUENCY DISTRIBUTION OF EMPLOYEES WORKING AT HOTEL MADANI SYARIAH MEDAN

No	Gender	Frequency	Percentage
1	Men	15	46,9
2	Women	17	53,1
	Total	32	100

From the table above, it can be seen that of the 32 respondents, most of the respondents were 17 women (53.1%), while 15 men (46.9%).

b) Age

The table below shows the age composition of Hotel Madani Syariah Medan employees.

TABLE II. AGE FREQUENCY DISTRIBUTION OF EMPLOYEES WORKING AT HOTEL MADANI SYARIAH MEDAN

No	Age	Frequency	Percentage
1	20-29 YEARS	17	53.1
2	30-39 YEARS	10	31.2
3	>40 YEARS	5	15.7
	Total	32	100

c) Marital Status

The table below shows the composition of employee status at Madani Syariah Hotel Medan.

TABLE III. FREQUENCY DISTRIBUTION OF MARITAL STATUS OF EMPLOYEES WHO WORK AT HOTEL MADANI SYARIAH MEDAN

No	Status	Frequency	Percentage
1	Marry	12	37.5
2	Not Married	20	62.5
	Total	32	100

From the table above, it can be seen that of the 32 respondents, most of the respondents were unmarried as many as 20 people (62.5%), while as many as 12 (37.5%) were married.

d) Last Education

The table below shows the composition of the last education of Hotel Madani Syariah Medan employees.

TABLE IV. DISTRIBUTION OF THE FREQUENCY OF THE LAST EDUCATION OF EMPLOYEES WORKING AT MADANI SYARIAH HOTEL MEDAN

No	Last Education	Frequency	Percentage
1	Junior High School	5	15.7

2	Senior High School	10	31.2
3	College	27	53.1
	Total	32	100

From the table above, it can be seen that of the 32 respondents, most of the respondents had tertiary education as many as 17 people (53.1%), while a small portion had junior high school education as many as 5 (15.7%).

2) Data Description

a) Leadership style

The author calculates the total value of the group I questionnaire to answer the first problem, namely what is the leadership style of Hotel Madani Syariah Medan. After knowing the total value of the group I questionnaire, the author entered the total score of the respondents into the table of types of leadership styles studied. The value of the questionnaire is known by dividing it into 4 class scales of leadership style on a data scale of 1-4.

Where 1 shows that the leadership style used is absolute mastery which in absolute control the leader gives full guidance and strict supervision of employees with the assumption that the best way to motivate employees is to give fear, threats, and punishment. While the grade 4 scale is an inviting leadership style and which this leadership style aims to make the organization run well through the real participation of employees.

The leadership style data scale is as follows:

TABLE V. DISTRIBUTION OF THE FREQUENCY OF THE LAST EDUCATION OF EMPLOYEES WORKING AT MADANI SYARIAH HOTEL MEDAN

Data Scale	Class	Category
1	1-1.99	Absolute Master
2	2-2.99	Semi-Absolute Mastery
3	3-3.99	Leadership Advisor
4	04-May	Participants

(Source: Results of Primary Data Processing, 2021)

b) Employee motivation

The author calculates the total value of the group II questionnaire to answer the second problem, namely how high the motivation of the employees of Hotel Madani Syariah Medan is. The author includes the total score of respondents in the employee motivation table at Madani Syariah Hotel Medan under study. The value of the questionnaire is known by observing the Likert scale as follows:

TABLE VI. MOTIVATION SCALE TABLE

Data Scale	Class	Category
1	1-2.32	Low Motivation
2	2.23-3.65	Medium Motivation
3	3.66-5	High Motivation

(Source: Results of Primary Data Processing, 2021)

c) To determine the level of employee achievement, the group III questionnaire was used. The first step is to calculate each item with a Likert scale, then calculate the total value of the questionnaire. The author includes the total score of respondents in the table of employee achievement at the Madani Syariah Hotel Medan under study. The value of the questionnaire is known by paying attention to the Likert scale as follows:

TABLE VII. WORK PERFORMANCE SCALE TABLE

Data Scale	Class	Category
1	1-1, 7	Very Low
2	1.8-2.5	Low
3	2.6-3.3	Currently
4	3.4-4.1	Tall
5	4.2-5	Very High

(Source: Results of Primary Data Processing, 2021)

3) Data Analysis

a) Validity test results

Validity testing was conducted to determine whether the questionnaire used was valid or not to be used as a research questionnaire. Validity testing is carried out using Product Moment (corrected item-total correlation) in the SPSS program. The questionnaire is said to be valid if the corrected item-total correlation > r table. Testing the group I questionnaire, consisting of 4 indicators and 12 statement items, the indicators are the way of making decisions of the leader which consists of 3 statements, The direction of communication between leaders and subordinates consists of 3 statements, Trust in subordinates consists of 3 statements, The way the leader treats employees consists of 3 statements. By using the product moment formula, at a significant level (a) of 5% and degrees of freedom (df) N-2 = 30. The results of the validity test can be seen in the table:

TABLE VIII. THE RESULTS OF THE LEADERSHIP STYLE VALIDITY TEST AT THE MADANI SYARIAH HOTEL MEDAN

Item No	Variable	rcount	rtabel 5% Df= N-2	Status
1	How to make decisions by leaders	0.316	0.3	valid
2	How to make decisions by leaders	0.606	0.3	valid
3	How to make decisions by leaders	0.391	0.3	valid
4	Direction of leader and subordinate communication	0.394	0.3	valid
5	Direction of leader and subordinate communication	0.52	0.3	valid
6	Direction of leader and subordinate communication	0.43	0.3	valid
7	Trust in subordinates	0.333	0.3	valid
8	Trust in subordinates	0.38	0.3	valid
9	Trust in subordinates	0.444	0.3	valid
10	How leaders treat employees	0.422	0.3	valid

11	How leaders treat employees	0.48	0.3	valid
12	How leaders treat employees	0.338	0.3	valid

The product moment correlation for items 1-12 is greater than r table 5% = 0.30, then the questionnaire items 1-12 are declared valid for use in the study.

B. Discussion

Discussion of the results of data analysis has been processed.

1) Medani Syariah Madani Hotel Leadership Style

From the results of the study, it can be seen that the average leadership style score is 4.06 with an interval class of 4-5 is the inviting leadership style as well. In making decisions, the head of Hotel Madani Syariah Medan provides opportunities for employees to express their opinions. Opinions that have been expressed by employees will be input for decision making by the head of Hotel Jentra Dagen. Good communication between superiors and subordinates applied by Hotel Madani Syariah Medan is that the leader when meeting with subordinates always greets, as well as superiors to subordinates who have good relationships so that in dealing with problems they can be resolved properly. Trust from superiors to subordinates is implemented by Hotel Madani Syariah Medan so that employees can complete their tasks in their own way and believe in the results they are doing. Leaders and superiors of Hotel Madani Syariah Medan always regard employees as co-workers, not as machines that are only ordered to. And always pay attention to the feelings of employees in acting, and provide tolerance for small problems that arise generated by employees.

2) Motivation of Medan Islamic Madani Hotel Employees

Based on the results of the research conducted, it can be seen that the average score of the questionnaire is 4.26, the average is in the high motivation group with a high score of 3.66-5. From the results of the above calculations, it is concluded that the motivation of the employees of Madani Syariah Hotel Medan is high. Employees who have the high work motivation tend to have a high desire for achievement. Employees with a high desire for achievement tend to want to have a better performance than the previous one.

Employees with high motivation have a desire to always learn new things and like challenges, because with challenges, employees will feel their abilities are honed. Employees with high motivation always carry out their assigned responsibilities well. The greater the responsibility given, the employees will feel challenged to carry out these responsibilities.

3) Employee achievement at Madani Syariah Hotel Medan

From the results of the research conducted, it can be seen that the average score of the questionnaire is 4.26, the average is in the group score of 4.2-5 is very high work performance.

4) Motivation does not mediate the influence of leadership style on employee performance.

The results of this study indicate that motivation does not mediate the influence of leadership style on work achievement, which means that the direct influence of leadership style on work achievement is greater than the indirect effect mediated by motivation, this is evidenced by the direct path coefficient value greater than the indirect path coefficient. directly with motivation as an intervening variable on the influence of leadership style on work achievement with a coefficient of 0.289 obtained from the leadership style variable X motivation, while the coefficient of direct influence of leadership style on work achievement is 0.635.

The thing that causes motivation does not mediate the influence of leadership style on work achievement is natural because employees are required to work in accordance with the SOP (standard operating system) in the company. High or low personal work motivation owned by employees, with SOP employees really have to carry out their duties with applicable SOPs and cannot carry out tasks outside of existing SOPs, which means that employees cannot work outside of existing SOPs and personal motivation is not so applies with the SOP.

C. Achieved Outcomes

1) Research Output

- The output of this study is in accordance with the results of previous research by Edu Dermantio In with the title The Relationship Between Motivation and Work Performance, Case Study on Employees of the Bogor Journal Newspaper 2005. The results showed that the motivation of Bogor Journal employees was high, it was seen from the average score. average for each section in the Bogor Journal. Based on the results of the correlation test, the general analysis of indicators is closely related to employee work motivation so that these indicators reflect the influence in the formation of Bogor Journal employee work motivation. The equation of this research is both about employee motivation and work achievement.
- While the difference with the research above, the author examines the relationship of leadership style with employee motivation and work achievement as well as with different research places and times. In addition, the results of previous research by Ricky Irawan entitled the relationship between leadership style, work motivation, and employee achievement with a case study on employees of Natasha Skin Care Medan 2008. This study was conducted to determine the relationship between leadership style, work motivation, and employee achievement. The first analysis shows that the leadership style of the head of the Natasha Skin Care Medan branch is democratic. The second analysis shows that Natasha Skin Care Medan employees have high work motivation. The third analysis shows that there is a positive relationship between leadership style and work motivation. The fourth analysis

shows that there is a positive relationship between work motivation and employee achievement.

- And previous research by Mariana with the title of analyzing the relationship between a manager's leadership style and enthusiasm, motivation, and work achievement of employees in a case study at PT. Sun Putra Pratama Corporation Palembang 2003. The results of the first analysis show that the leadership style applied is a supportive leadership style. The results of the second analysis show that there is a relationship between a manager's leadership style and employee morale, there is a relationship between a manager's leadership style and employee motivation, and there is a relationship between a manager's leadership style and employee achievement.

2) Achieved Outcomes

- The direct supervisor of Hotel Madani Syariah Medan should be closer to employees/subordinates to create closer relationships as colleagues
- The direct supervisor of Hotel Madani Syariah Medan emphasizes good cooperation with subordinates in problem solving.

The direct supervisor of Hotel Madani Syariah Medan should pay more attention to the feelings of subordinates in actions and actions.

IV. CONCLUSION

- The leadership style applied by the manager of Hotel Madani Syariah Medan tends to be a leadership style called the inviting leader as well.
- Employee motivation at Madani Syariah Hotel Medan is high.
- Employee achievement is very high.
- Motivation does not mediate the influence of leadership style on work achievement, which means that the direct influence of leadership style on work achievement is greater than the indirect effect mediated by motivation, this is evidenced by the value of the direct path coefficient being greater than the indirect path coefficient with motivation. as an intervening variable on the influence of leadership style on work achievement with a coefficient of 0.289.

REFERENCES

[1] Augusty, Ferdinand. 2006. Metodologi Penelitian Manajemen: Pedoman Penelitian untuk Skripsi, Tesis dan Disertai Ilmu Manajemen.Semarang: Universitas Diponegoro.

[2] Daft, Richard L.2006. Management. Jakarta: Erlangga.

[3] Mathis, Robert L dan Jhon H. Jackson.2006. Human Resource Management. Jakarta:Salemba.

[4] Rivai, Veithzal.2014. Manajemen Sumber Daya Manusia untuk Perusahaan,Edisi ke 6, PT. Raja Grafido Persada,Depok 16956.

[5] Robinns, Stephen P.2006. *Perilaku Organisasi*. Jakarta: PT. Indeks Gramedia.

[6] Simamora, Henry. 2014. Manajemen Sumber Daya Manusia. Medan: STIEYKPN.

[7] Sinulingga, Sukaria.(2011). *Metode Penelitian*.Medan: USU Press.

[8] Sekaran, Uma.(2006), Metodologi Penelitian Untuk Bisnis,Jakarta : Salemba Empat.

[9] Sugiyono, 2007. Metode Penelitian Kuantitatif Kualitatif dan R&d. Bandung: Alfabeta.

Market Function through Small Family Farms-Middleman Interaction, and Its Relation to Food Security: Evidence from Indonesia

1st Endar Purnawan
Agriculture Food and Environment Department
University of Pisa
Pisa, Italy
e.purnawan@studenti.unipi.it

2nd Natalia Gabriela Molina Navas
Agriculture Food and Environment Department
University of Pisa
Pisa, Italy
n.molinanavas@studenti.unipi.it

3rd Gianluca Brunori
Agriculture Food and Environment Department
University of Pisa
Pisa, Italy
gianluca.brunori@unipi.it

4th Yuliska
Informatics Engineering Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
yuliska@pcr.ac.id

***Abstract*—This research addresses the interaction between small family farms and middleman and its relation to food security. Based on a qualitative method analysis derives from farmer and intermediaries interviews, this research sheds light on four types of middleman in this area: conventional middleman, rice middleman, banana middleman, and rubber middleman. It reveals factors motivating farmers to market crops through a middleman such as best crops selling price, the quantity of the crops, a new market for subsistence crops, and tied relationship. The result indicates that each middleman's role differs depending on the interaction, where most of them are simply a way to connecting farmers to the final market to earn cash quickly. Taken together, the research findings highlight the role of each middleman, including "Bertauke" interaction, which is uniquely proven to contribute to food availability and food accessibility of the small farm households from one harvest-time to another, before recommending farmers market in order to benefit farmers and the community for better local food security.**

Keywords—Small Family Farms, Middleman, Market Access, Farmers Market, Food Security

I. INTRODUCTION

Lack of formal connections between small farm holders and market actors tends to limit the access to information that farmers require to sell their produce [3]. Markelova et al. asserted that if small producers act collectively, they may actually reduce transaction costs, get updated market information, and enhance their bargaining power with buyers [6]. Nevertheless, collective actions from smallholders are a challenge due to their poor organization capacity and low level of trust amongst them [4].

Middleman has an integral role in linking smallholder farmers to traders and final markets [1]. Rubinstein and Wolinsky [14] asserted that middlemen are a time-saving institution since they shorten the negotiation time between sellers and buyers in a transaction. In Indonesia, most farmers are small-scale farmers, wherein middleman has a positive role because most farmers in Indonesia are fragmented [8]. If they distribute their agricultural products directly to consumers, it will cause variation in prices and costs distribution will become more expensive due to irregular quantities [7]. Middleman may, however, reduce the profitability of farmers in the long-run [7]. Middleman often compromises the efficiency of distributing agricultural products by decreasing prices at the level of farmers [13] [19].

Broader than economic aspect, middleman may also function as a social network structure, as defined by group-membership [10] and institutionalization of group relations [12]. An economic transaction that involves a middleman signifies a socially-tied relationship, where close knit communities, such as those based on kinship and geographical proximity, are characterized by strong social ties [10].

This paper probes into the following: 1) the characteristics of farmers and the type of middleman, 2) factors that influence farmers' decision of marketing crops through middleman, and lastly, 3) the role of middleman on food security among family farmers in the context of Indonesia.

II. RESEARCH METHODOLOGY

This study was conducted in Sekayam sub-district located at Sanggau district, West Kalimantan province in Indonesia. The qualitative data were analyzed using inductive and descriptive analyses to obtain in-depth and accurate outcomes [11]. In total, 21 informants were employed for this study, comprising of 6 middlemen and 15 farmers. In-depth interviews and observations were the data collection techniques used in this study. The data were gathered between April and August 2019. Table I below shows indicator used in the interviews. This study was conducted in Sekayam sub-district located at Sanggau district, West Kalimantan province in Indonesia. The qualitative data were analyzed using inductive and descriptive analyses to obtain in-depth and accurate outcomes [11]. In total, 21 informants were employed for this study, comprising of 6 middlemen and 15 farmers. In-depth interviews and observations were the data collection techniques used in this study. The data were

gathered between April and August 2019. Table I below shows indicator used in the interviews.

TABLE I. INDICATORS USED DURING INTERVIEWS

Indicators			
No	**Farmer**	**No**	**Middleman**
1.	Information of the respondent and the household	1.	Information of the respondent
2.	Land for farming, types of plants, and amounts of crops harvested each year	2.	Crops/ food items to accommodate/ sell? Where to get?
3.	Crops consumed, sold, and used for agricultural inputs?	3.	How much bought and sold? Purchase and selling price?
4.	Where the crops were sold?	4.	Where to sell?
5.	Determining factors in choosing food to consume and special moment for food?	5.	Number of intermediaries or traders in the area?
6.	Experienced difficulty accessing food (e.g., no money to buy food) and strategies to deal with it?	6.	The quantity of crops sold in this hamlet/village in one month?
		7.	The obstacles faced in becoming a middleman
		8.	The supporting factors in the business

Source: Proceed by authors

The stages of data analysis in this study adhered to that prescribed by Neuman, as follows [11]: Data 1 referred to raw data and the researcher's experience based on observation and interview in the field. Data 2 denoted data recording and physical document derived from observation in the form of visual recording and from the interviews in the form of sound recording. Data 3 derived from data selection or the final report from data 2 and other processed resources (agency documents and literature). Figure 1 illustrates the simple steps embedded for the process of data analysis.

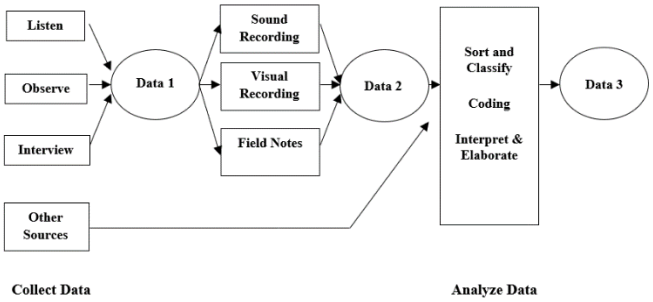


Fig. 1. Data Analysis Process (Modified from [11])

a. Characteristics of Farmers and Type of Middleman

1) Characteristics of Farmers

This study was conducted in Sekayam Sub-district located at Sanggau District, West Kalimantan Province, Indonesia. The farmers in this area spent 32.48% only of their revenue for food from their income. Nevertheless, they spent nothing (or in small amount) for rice as staple food and vegetables. Rice, rubber, oil palm, pepper, banana, vegetables, and tropical fruits were the types of crops planted by the farmers in the study area. In average, they managed 2.68 ha of farm field per household, and they planted any commodity with good selling price to sustain their income. Farmers in this area planted a range of crops because planting only one crop reduced the variety of food for consumption, while increasing the risk of crop failure, as Von Braun and Kennedy asserted that farmers select crop mix to gain food security and maximum return on scarce resources [20].

2) Different Types of Middleman in the Research Area

The 4 types of middleman identified in this study area are as follows:

a) Conventional Middleman

This type of middleman is a normal middleman, whereby farmers can sell their crops to them, in order to gain cash. Most of them accepted pepper, palm oil fruits, and rubber from farmers, who then brought the crops to a wholesale (for pepper) or to a factory (for oil palm fruits and rubber). These middleman gained profit from the margin selling price between farmer and wholesaler/factories.

b) Rice Middleman

There is a rice middleman in Berungkat village as a result of a program namely *Pengembangan Usaha Pangan Masyarakat* (PUPM) that supported farmer shop (TTI), which was managed by the chief of joint group of farmers in Berungkat village. They accepted grain from farmers, process it, and then sold it to community. Prior to PUPM program, farmers keep their grain for their own consumption and sell it directly to consumers in small amount. This middleman motivated the farmers to produce more grain, as he could absorb their crops in large quantity.

c) Banana Middleman in Ruis Hamlet

This middleman initially had a part-time job harvesting bananas for several years. When the farmers need to harvest their crops, they would pay a part-time worker to help them harvest bananas in their farm. When the worker had sufficient money to buy bananas from the farmers, including harvesting cost, he served as a middleman. The interaction between the farmers and the worker had motivated him to become a banana middleman.

d) Rubber Middleman in Ruis hamlet

As a conventional middleman, a rubber middleman in Ruis hamlet (part of Bungkang village) had unique tie with the farmers. Local people called it "bertauke", which reflects their interaction; where farmers can borrow food items from the middleman and repay after selling their rubber. The middleman had a small grocery that sold food items to the farmers.

III. RESULTS AND DISCUSSION

A. Farmers' Decision of Marketing Crops through Middleman

Based on field findings, several reasons that drove the farmers' decisions on marketing their crops were noted, such as selling price, crops quantity, new market, and tied relationship.

1) Best Selling Price of Crops

A farmer (farmer 2) said that if they had a chance to reach the border line between countries to get better selling price, they would do it. The farmer said that the border between Indonesia and Malaysia used to be free for transaction, so they could freely and directly sell pepper and cocoa at the border gate. At that time, they used motorcycle to bring crops to the border gate, where a middleman from

Malaysia would be waiting for them. This shows that farmers try to get the best selling price of their crops, even though they need extra effort to obtain it.

2) The Quantity of Crops

The farmers sold small amounts of crops to small middleman in their hamlet. A farmer (farmer 12) said that when they had a small amount of pepper crop, they would sell it directly to a middleman in their hamlet. However, when they had more pepper crop, they sometimes went to a wholesaler in Nekan. Transportation and time costs were weighed in if they had decided to go to wholesaler, in order to gain better selling price that also could cover the transaction cost they spent. As for the oil palm and rubber farmers, low quantity of crops was related to lack of access to market. These commodities were sold to factory; more crops were sold directly to the factory for better selling price (oil palm or rubber factory). However, small family farms sold their little amount of crops to middleman in their community, who served as a link for farmer and trader [14]. The farmers need not to keep their crops to meet certain amount, thus being able to gain cash money more quickly.

3) New Market for Subsistence Crops

A rice middleman started buying rice grain from the farmers. Typically, rice was only for household consumption or sold in small amount directly to consumers. However, the rice middleman in Rintau hamlet (part of Berungkat village) bought rice grain from farmers under the PUPM program in the area. The farmer (farmer 2) said that the existence of those middlemen gave a great help to rice farmers in Rintau hamlet, because now they can sell their grain in large amount to gain money. This obviously motivated the rice farmers to increase their productivity. As Sadler explained, market availability and access motivate farmers to not only farming for subsistence purpose but also gain income from the farm activity [15].

4) Tied Relationship

The three examples for this farmer-middleman relationship (see [9] [11]) particular point identified in the research area are as follows: A farmer in Berungkat hamlet claimed that he sold oil palm to a middleman in Lubuk Sabuk village, who are his cousin named Pak Abang Roni. The farmer also said that two middlemen hold and buy palm fruits from farmers in the area. The relationship between banana farmer and banana middleman in Rusi hamlet is portrayed as follows, "In the beginning, I worked as a banana harvester in someone else's banana plantation, because I was the one who harvested most of the banana yields here. Finally, I have been known and trusted by farmers who own banana plantations. Once, a large banana wholesaler at Balai Karangan asked me if I was willing to supply bananas for him. We negotiated the price of bananas. Upon agreeing with the price, I approached the banana farmers and told them that I wanted to buy the bananas that I had harvested for them immediately. So, I cropped directly for them, and then I bought the bananas directly for 1,500 rupiah per kg. The farmers received the harvest directly without the need to think about the costs of harvesting and transportation. The farmers agreed, and finally, I turned into a banana middleman in this hamlet. I harvest and buy bananas directly from the farmers, and then, I sell them to large intermediaries at Balai Karangan. Once, a farmer came to my house to offered me the job of harvesting bananas in his plantation for 300 rupiah per kg of bananas, including the costs of harvesting and transportation" (Middleman 5).

The last form of relationship is "Bertauke", between rubber farmers and a middleman in Ruis hamlet. In local language, "Bertauke" means an interaction, where farmers can borrow food items from the middleman and only pay for the borrowed items when they have rubber to sell. Normally, the middleman should have a small grocery that sells food items to the community. A farmer in Ruis Hamlet (farmer 11) said that they sold rubber products to Mr Indra, who is the only rubber middleman in Ruis hamlet. The farmer added that they could sell their rubber to any middleman outside the hamlet for better price, but they would not. Mr. Indra is his relative and they always owe goods from his small grocery.

This interaction was further explained by the middleman, "In this hamlet, it seems that I am the only one who accommodates rubber from farmers. I also have a small grocery store, where farmers sell rubber and buy things from me. They are accustomed to exchanging rubber with goods; bartering. So, they first take goods from me, and later sell their rubber to me, or pay off their debts with rubber. For other intermediaries around here, they take the sap from farmers in other villages. So, the latex that enters my place is approximately the amount of latex produced by the sap farmers in Ruis, which is about 3 tons a month. In this hamlet, 51 farmers routinely take goods and sell rubber to me. I sell food items commonly bartered by farmers, such as rice, edible oil, sugar, chicken, and fish" (middleman 2).

Figure 2 shows the interaction of farmers with different types of middleman, along with the reasons of farmers' decision marketing crops through middleman discussed above.

B. The Role of Middlemen on Food Security of Family Farmers

The term 'intermediary' is used commonly in business sector, including agriculture, with both positive and negative connotations [9]. Intermediaries only increase personal profits without increasing the added value of the products by using limitations of market information of farmers as producers [16]. Middlemen make profit by reducing the prices of farmers as low as possible [13] [16] [18]. Farmers, at times, have no other choice because agricultural products rot easily and are bulky, thus it is better for farmers to sell them to middlemen [8]. Most Indonesian farmers (93%) were small family farmers [2]. This signifies the positive role of intermediaries as most farmers were fragmented. However, this positive role on food security might not be equal for every middleman. Hence, the following explains the contribution of different middlemen to food security of small family farms:

1) Conventional Middleman

In this study, all the farmers sold their cash crops to middlemen as explained by a farmer (farmer 4). The farmer stated that the yields of oil palm plantations in there were all sold to a middleman in Berungkat Hamlet named Mr. Ja. He accommodated from farmers and then sold palm fruits to the factory in Kembayan. Mr Ja also held rubber. Another farmer

claimed, “We sell pepper to intermediaries in Balai Karangan, while we sell rubber to a local intermediary in this village.” (farmer 11). Farmers gained cash quickly when they sold crops to middleman, in which they need not pay additional distribution cost but only pay a small amount due to the lower price offered by the middleman.

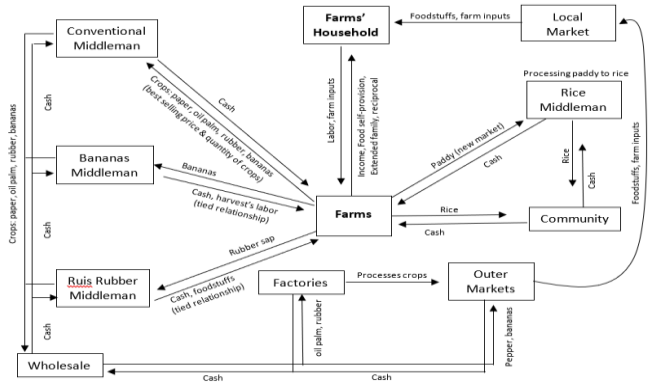


Fig. 2. Model of interaction (farmers and middlemen)

A middleman (Middleman 2) said that he used to sell pepper and cocoa to a wholesaler in Engkahan village named Mr. Mus. He even accepted the cocoa price of 5,000 rupiah because cocoa fruits do not last long. He stated that if the fruit shrinks, they can sell only 40% to the wholesaler. Linking gaining cash quickly with food security at household level is another point. However, earning extra income need not necessarily affect food consumption, as farmers need to meet many daily needs and save their money for future needs.

2) Rice Middleman (TTI) in Bungkang Village Supported by PUPM Program

The rice middleman (TTI) in the area was supported by the government via PUPM program to ascertain food availability and food access for the community. The middleman (Middleman 1) claimed that they were assisted by the government through PUPM funds for rice availability. The goal was to make rice available to the community in affordable and accessible price. So, they were given capital to buy grain from farmers at their price and sell rice to people at an amount determined by the government. As farmers in this area plant rice only for self-provision purpose, this program as a new market motivated farmers to produce rice more than what they need [13]. A farmer (farmer 2) mentioned that for them who are from Berungkat hamlet, all farmers sold grain to a middleman in Rintau. This is good because the farmers could sell 5 to 6 sacks of crops to restore their operational costs of paddy planting. The farmer added that the middleman helped the rice farmers a lot. Likewise, the chief of TTI (Middleman 1) explained that this condition stimulated the productivity of paddy farming in this region because they tried planting twice or thrice a year to gain more productivity. This program benefits the community, as they can buy rice at a low price (9,000 rupiah per kg), where previously they bought rice for 13,000 rupiah per kg. The TTI production was equal to 11.87% and 17.46% of rice needed in Bungkang village for 2018 and 2019, respectively.

3) Banana middleman in Ruis hamlet

Banana middleman in this hamlet built long-term relationship with the farmers. The middleman said, “I have been working for bananas for eight years, since 2011. In the beginning, I worked as a banana harvester in someone else's banana plantation. Once a farmer came to my house to offer me the job of harvesting bananas at his plantation with a wage of 300 rupiah per kg of bananas, including harvesting and transportation costs. I began gaining trust from farmers who owned banana plantations. One day, a large banana wholesaler at Balai Karangan asked me if I was willing to supply bananas for him. After negotiating and agreeing with the price of bananas, I began approaching the banana farmers to buy their bananas that I had harvested for them immediately. So, I bought the bananas directly for 1,500 rupiah per kg. The farmers received the harvest directly without the need to think about harvesting and transportation costs. The farmers agreed, and finally, I have turned into a banana wholesaler in Ruis Hamlet. I directly harvest and buy bananas from the farmers, and then, sell them to large intermediaries at Balai Karangan. The profit that I took is only 500 rupiah per kg banana. So, if the selling price from farmers is 1,500 rupiah per kg of bananas, then I would sell it for 2,000 rupiah per kg” (Middleman 5).

Clearly, the middleman simplified the banana market process for farmers from the long market chain. Nevertheless, in order to determine the effect of this interaction on food security at household level, more effort is required. This is because, the community in Ruis preferred saving their money than adding variety to their consumption.

4) Rubber middleman in Ruis hamlet

The rubber middleman in Ruis hamlet built a unique relationship with the local farmers. First, the middleman had a small shop that sold many food items, where farmers could borrow the food items, and later pay with their crops. A farmer (farmer 12) mentioned that when they did not have money but they needed something, they could go to the middleman and borrowed items from his shop. The farmer lent the middleman goods because they always repaid by selling their crops to the middleman to pay off the debt. As this happened rather frequently, the middleman contributed to provide food items for the farmers in this hamlet. The farmer further explained that every month, they borrowed goods from the store. After a week or two, they would sell their produce to pay off the debt. They usually borrowed food items twice a month. The local people called this relationship “Bertauke”, as the farmers sold their crops and borrowed food items from the shop owner/middleman of their crops. If the farmers had no rice at home or no side dish to consume, they would go to the middleman and borrow food items, which are paid later. “There has never been a concern about lack of food in our house, thank God” (farmer 12).

The middleman (middleman 2) in this hamlet explained the situation very well; he stated that in this hamlet, he was the only one who accommodates rubber latex. He had a small grocery store, where farmers sold rubber and buy things from him. They were accustomed to exchanging rubber with goods; bartering. So, they first took products, owing him goods, then sold latex to him or paid off their debt with latex. For other intermediaries who were around the area, they took the latex from farmers in other villages. So, the latex that entered his

place was approximately the amount of latex produced by the sap farmers in Ruis, which was about 3 tons per month. He added that at present, he takes latex from farmers at 7,000 rupiah, while the selling price at the warehouse is 7,800 rupiah, so the profit gained is only 800 rupiah per kg of latex. He sold rubber to latex warehouses in two places, Buyung's place in Balai Karangan and Herman's in Kenaman hamlet. They are big middlemen who buy latex from small middlemen like him. When he lacked of financial capital, he borrowed money from the rubber sap container where he sells the sap. Normally, loans to rubber sap warehouse were given only half of their sales. Monthly income from collecting the sap is around 1.2 million rupiah, but sometimes less than that due to the unusual amount of sap depreciation. Not all latex that comes from farmers was clean. Sometimes the farmers mix something into the latex that would be dried so that it could add to the weight of the latex that he bought. Depreciation when selling to warehouse from his experience was 5-15% per ton latex. So, from the purchase of 1-ton of sap, he only got sales of 850-950 kg. He marketed this latex from the farmers to the warehouse or large container, which was then sent to a sap factory in Pontianak. He sold all kinds of food items commonly bartered by farmers, such as rice, chicken, fish, eggs, cooking oil, and sugar. He bought them at Balai Karangan. He estimated this trading profit at around 2 million rupiah per month.

The second reason of this unique tied relationship is the relationship itself, stemming from the extended family relationship shared between the farmers and the middleman. Despite the potential to gain better rubber selling price with other middlemen, the farmers chose to keep this "bertauke" tied with the middleman due to family ties. This established the position of the middleman among the local farmers in this hamlet, "My supporting factor is the support from farmers who always supply latex to me due to the barter system we use between food and agricultural products. Another factor is my wife; she has many farming families who continue selling their agrarian produce to me. Actually, 51 farmers in this village routinely take goods and sell latex to me" (Middleman 2).

Clearly, the tied relationship between farmers and middleman had uniquely contributed to food availability, food accessibility, and food utilization of the small farm households from one harvest to another. Essentially, the interaction contributes to food stability during the cycling of food scarcity for the farmers. Table II lists the reasons on why the farmers sold crops to middlemen and the contribution of these middlemen to farmers.

TABLE II. HOW MIDDLEMEN BENEFIT FARMERS

No	Type of middleman	Reason to sell	Benefit
1.	Conventional middleman	Best selling price Quantity of crops	Gain cash quickly, and do not need to pay additional distribution cost, or only need to pay in small amount.
2.	Rice middleman	New market	Before, farmers in this area planted rice only for self-provision. This middleman helped the farmers to earn extra income from paddy farming, and motivated the farmers to produce rice more than they needed.
3.	Banana middleman	Tied relationship (middleman and harvest labor)	The farmers received cash directly without the need to think about harvesting and transportation costs. Simplified the banana market process for farmers.
4.	Rubber middleman in Ruis hamlet	Tied relationship (extended family and "bertauke")	Uniquely contributes to food availability, food accessibility, and food utilization of the small farm households from one harvest time to another. The interaction contributes to food stability during the cycling of food coping strategy.

Source: Proceed by authors

b. The Need for a Farmers Market Initiation

In former parts, it is clearly seen that the small farmers are very dependent to middlemen. Although the farmers want to get the best price in selling their crops, they often end up selling it to middlemen in their area with cheaper price because of their limited access to the market and the small quantity of their crops. Tied relationship and "bertauke" culture are also additional factor for farmers to sell their crops to middlemen. In addition, the lack of information about the value of the product and market also play a significant role to this kind of interaction. Hence, a farmers market needs to be initiated.

Farmers market is still an unfamiliar concept for farmers in Sekayam Sub-district and even for most small farmers in Indonesia. But, the initiation of farmers market can be a great option for small farmers to market their crops with the best price. As Singleton *et al*. explained that to provide more benefits to smallholders, there must be innovation in terms of marketing arrangements, where farmer or producer organizations can play a better role in this new arrangement. For example, farmers' markets can be a solution to break the long food chain; at the same time, farmers' markets can also be a place for nutrition intervention [17]. Furthermore, there are several reasons why farmers' markets are ideal places for nutrition and food security programs, such as good interactions between buyers and sellers, which open up opportunities for buyers to find out more information about the food to be purchased, food and drink which are sold fresh and healthy and involve food processing skills by involving food ingredients and other healthy food ingredients [15]. Even farmers' markets are said to contribute positively to alleviating food desert challenges and can impact meeting social justice goals [5]. Those explained that farmers market provides numerous advantages for small farmers from minimizing their interaction to middlemen, asses to the best price, to gain more information about the value of their products.

IV. CONCLUSION

In light of small family farms within this study context, the intermediaries had a role in connecting farmers to the final market, where farmers could earn some quick cash. However, the farmers received lower prices because the intermediaries

took advantage of the difference in price obtained from farmer price and selling price to large intermediaries or factories.

The farmer-middlemen interaction known as "Bertauke" proved the contribution of food security for farmers. Another intermediary provided staple food (rice) to the community at a lower price. Other intermediaries were essentially a way for farmers to gain cash. However, the link between receiving cash and family food security demands more investigation. The availability of money might not directly guarantee increment in family consumption, as other needs of farmer families must be met. The farmers preferred saving money to prepare for difficult times than significantly changing their diet.

This paper strongly recommends farmers market initiation to benefit farmers, for better local food security either for the small family farms and the community at the same time.

ACKNOWLEDGMENT

This work was supported and funded by Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, Republic of Indonesia.

REFERENCES

[1] G. K. Abebe, J. Bijman, A. Royer, "Are middlemen facilitators or barriers to improve smallholders' welfare in rural economies? Empirical evidence from Ethiopia", Journal of Rural Studies, Vol. 43, 2016, pp. 203-213. https://doi.org/10.1016/j.jrurstud.2015.12.004

[2] FAO, "Smallholders data portrait" (available at www.fao.org/family-farming/data-sources/dataportrait/farm-size/en), 2018.

[3] A. Hall, "Challenges to strengthening agricultural innovation systems: where do we go from here?", In: Scoones, I & Thompson, J (Eds.) Farmer First Revisited: innovation for agricultural research and development, Practical Action Publishing, Rugby, 2009, pp.30-38.

[4] J. Howells, "Intermediation and the role of intermediaries in innovation", Research Policy, Vol. 35, 2006, pp. 715-728. https://doi.org/10.1016/j.respol.2006.03.005.

[5] K. Larsen & J. Gilliland, "A farmers' market in a food desert: Evaluating impacts on the price and availability of healthy food", Health & Place, Vol. 15(4), 2009, pp. 1158–1162. https://doi.org/10.1016/j.healthplace.2009.06.007

[6] H. Markelova, R. Meinzen-Dick, J. Hellin, S. Dohrn, "Collective action for smallholder market access", Food Policy, Vol. 34 (2009) 1–7, 2009, pp. 1-7. https://doi.org/10.1016/j.foodpol.2008.10.001

[7] G. Mejía & C. García-Díaz, "Market-level effects of firm-level adaptation and intermediation in networked markets of fresh foods: A case study in Colombia", Agricultural Systems, Vol. 160 (October 2015), 2018, pp. 132–142. https://doi.org/10.1016/j.agsy.2017.06.003

[8] J. S. Meliala, M. Hubeis, S. Jahroh, and A. Maulana, "Aggregator Business as an Intermediary in Agriculture: A Literature Review", In Proceedings of the 4th Sriwijaya Economics, Accounting, and Business Conference (SEABC), 2018, pages 56-64. ISBN: 978-989-758-387-2

[9] D. D. Monieson, "Explorations & insights a historical survey concerning marketing middlemen as producers of value", Journal of Historical Research in Marketing, Vol. 2(2), 2010, pp. 218–226. doi:10.1108/17557501011042560.

[10] K. Munshi, "Community networks and the process of development", J. Econ. Perspect., Vol. 28, 2014, pp. 49-76.

[11] W. Lawrence Neuman, "Social Research Methods: Qualitative and Quantitative Approaches", Sixth Edition, United States of America: Pearson Education, Inc, 2006.

[12] A. Portes, "Social capital: its origins and applications in modern sociology", Annu. Rev. Sociol, Vol. 24, 1998, pp. 1-24.

[13] R. Ranjan, "Challenges to farm produce marketing: A model of bargaining between farmers and middlemen under risk", Journal of Agricultural and Resource Economics, Vol. 42(3), 2017, pp. 386–405.

[14] A. Rubinstein & A. Wolinsky, "Middlemen", Q. J. Econ, Vol. 102, 1987, pp. 581-593.

[15] R. C. Sadler, "Strengthening the core, improving access: Bringing healthy food downtown via a farmers' market move", Applied Geography, Vol. 67, 2016, pp. 119–128. https://doi.org/10.1016/j.apgeog.2015.12.010

[16] T. Shankar, K. M. Singh, and S. Dwiwedi, "An Analysis on Problems of Vegetables Marketing in Farmers' Market of Jharkhand: A Case Study in Ranchi District", Vol. 62(1), 2017, pp. 175–183. doi:10.5958/2230-7311.2017.00054.X.

[17] C. R. Singleton, B. Sen, and O. Affuso, "Disparities in the Availability of Farmers Markets in the United States", Environmental Justice, Vol. 8(4), 2015, pp. 135–143. https://doi.org/10.1089/env.2015.0011

[18] J. P. Simon, P. J. Benghozi, and E. Salvador, "The new middlemen of the digital age: The case of cinema", Info, Vol. 17(6), 2015, pp. 97–115. doi: 10.1108/info-04-2015- 0023.

[19] S.K. Tapsavi, "Uttarakhand state cooperative federation: can it help the horticulture farmers?: Case Analysis 1", The Journal of Business Perspective, Vol. 13(2l), 2009, pp. 63-64

[20] J. Von Braun & E. Kennedy, "Commercialization of subsistence agriculture: income and nutritional effects in developing countries", Vol 1, Washington DC: International Food Policy Research Institute, 1986.

The Effects of Destination Attributes and Benefits on Decision Making of MICE Stakeholders in Medan City

1st Nursiah Fitri, S.E., M.Si.
Business Management Study Programme
Politeknik Negeri Medan
Medan, Indonesia
nursiahfitri@polmed.ac.id

2nd Jumjuma, S.E., M.Si
Business Adminstration Study Program
Politeknik Negeri Medan
Medan, Indonesia
jumjuma@polmed.ac.id

3rd Faulina, S.E., M.Si.
MICE Study Programme
Politeknik Negeri Medan
Medan, Indonesia
Faulina.19640414@polmed.ac.id

***Abstract*—Medan City is one of the government-appointed MICE-destination. Harsh and strict competition amongst tourist destinations demands support, participation, and investment from stakeholders. This study is focused more on the perceptions of stakeholders towards the destination attributes of Medan City and the benefit of MICE activities on decision making. The population used in this study is MICE stakeholders in Medan City, consisting of planners who have experience as EOs (Event Organizers), travel agencies, hotels, venues, and the government tourism office. Using the Lameshow formula, a sample of 96 people was determined. The sampling technique used in this study was the purposive sampling technique. PLS (Partial Least Square) structural equation modelling was used for the analysis. The results showed that destination attributes had a positive but not significant effect on decision making. Destination attributes have a positive and significant effect on benefits. Benefits have a positive and significant effect on decision-making. Destination attributes through benefits have a positive and significant effect on decision-making.**

Keywords—Benefits, Decision Making, Destination Attributes, MICE Stakeholders

I. INTRODUCTION

The MICE (Meetings, Incentives, Conventions, and Exhibitions) sector is one of the fastest-growing tourism market segments and was greatly affected during the Covid-19 pandemic. The potential loss during 2020 is estimated to be around Rp. 2.69 trillion - Rp. 6.94 trillion. This is because around 96.43 percent of events in 17 provinces in Indonesia had to be postponed during the pandemic. Data from ICCA (International Congress and Convention Association) as of February 27, 2020, recorded that 12 events were postponed, and two were canceled [1]. In the New Normal era, this sector must be able to rise by aligning strict health protocols in organizing events and using information technology in their activities.

Medan City is one of the government-appointed MICE-destination. Medan City is strategically located on the Strait of Malacca, which is the entrance to western Indonesia. This has permitted Medan City to be the center for tourism and trading on a domestic and international scale. The city of Medan is close to Kuala Lumpur and Singapore, which are the world's MICE cities. The government's plan to increase national and international MICE activities is faced with firm competition between destinations. For this reason, it is necessary to make decisions (decision making) by stakeholders in the form of support, participation, and investment in the MICE sector.

A lot of researches on stakeholder behavior have been carried out including the behavior of MICE stakeholders, consisting of hospitality workers and tour officers [2]. Zhiyi & Hanyu examined the impact of the transformation of the destination governance structure on stakeholder perceptions [3]. The needs and expectations of MICE stakeholders for tourism development in the Barents region were carried out by Konovalenko [4]. The decision-making to organize a MICE event is a difficult process for event organizers in a MICE destination. Destination attributes play an important role in this decision-making process. This variable is considered an attracting factor that can inspire delegates and event planners to a destination [5], [6].

MICE activities are recognized to have benefits in various aspects which are referred to as the multiplier effect. Various studies showed that the benefits resulting from holding MICE activities are the basis for investing in this business [7]–[12].

Therefore, this research is more focused on stakeholder perceptions of destination attributes and the benefits of MICE activities on decision-making in supporting and investing in the MICE sector. This is important considering the intense competition between destinations to attract participants to carry out MICE activities. So that in the post-Covid-19 pandemic, Medan City can be transformed as an icon of destinations in Indonesia.

II. MICE STAKEHOLDERS

Biset in Azheri stated that Stakeholders are those who have an interest and are motivated by profit expectations [13]. The company is not an entity that only operates for its own interests but must provide benefits to its stakeholders [14].

Stakeholder groups are categorized into three groups, which are primary, secondary, and other groups. The primary

group consists of city officials, marketing organizations, competitors, tourism attraction operations, service companies, tourists, restaurants, and hotels. Meanwhile, the secondary group consists of chambers of commerce, incentive planners, and community groups. Other groups are made of property owners, local businesses, coastal area managers, and employees [15]. In the context of this study, the selected stakeholder category is business owners consisting of hotels, venues, EOs, and travel tours.

III. Destination Attributes

Destination attributes are deemed critical by event planners, associations, attendees, and destination hosts. Competition between destinations in helding MICE events has increased, which led to the need for identifying key criteria to satisfy clients and meet their needs and expectations. Push factors are wants, needs, and perceptions that affect the clients, while attracting factors are destination attributes [16].

Destination attributes are factors that attract tourists to a destination [17]. The dimensions of the destination attributes are Amenity, Accessibility, Accountability, Affordability, Attractions, and Activity. Amenity is described as adequate facilities for conventions and exhibitions, and meeting room facilities and their ability to have certain basic services such as fire, police, water, etc. Accessibility refers to the ease with which participants can travel to and from the event location by considering time and effort. Accountability refers to the ability of the host destination to provide overall quality to MICE tourism participants in terms of customs, telecommunications, health care, and qualified employees. Affordability refers to the overall price or cost of participating in an event. Attractions refer to the ability of the host goal to provide attractions and points of interest to meeting participants. Activity refers to recreational activities before, during, and after meeting events [5].

IV. MICE Benefits

MICE benefits can be described as acknowledgment of the benefits and contribution of MICE activities in the form of investment incentives that encourage investment. In a behavioral study, attitudes toward recognizing the socioeconomic impact of MICE business and the contribution of MICE were found to significantly influence the behavior of respondents [7].

MICE benefits to various aspects will foster curiosity about how potential investors make decisions to invest in this sector. The literature shows that the benefits resulting from holding MICE are valuable. This is the basis for doing business in this sector.

Dimensions of these benefits include facilitating access to new technologies, attracting high-spending visitors, providing high yields and returns per capita, improving international economic relations, creating more economic multiplier effects and competitiveness, and can occur outside of peak season [7].

MICE delegates stay longer with large expenses [18]. Furthermore, MICE contributes to community development, urban renewal, and the growth of national identity [19]. Kay emphasizes the profit criteria and motives behind the establishment of convention and exhibition centers; These include airport repairs, subway systems, highways, host city renovations, parks, and other various urban renewal schemes [20].

V. Decision Making

Investment in decision-making is considered a critical factor for any business entity that affects its prosperity, competitiveness, and long-term performance [21]. Theoretically, the investment decision process goes through four main things, namely the initiation phase and initial thinking, investigation, evaluation, and final decision[22]. For Sykianakis & Bellas, there are five stages: introduction, diagnosis, screening, development and design, and negotiation stages [23]. More comprehensively, Niskanen & Niskanen describe seven phases of the investment process: recognition, search, information search, selection, funding, and implementation and monitoring of investment projects [24].

The stages of the decision-making process will differ sequentially according to the nature of the project. For tourism, including MICE, recognition, search, information retrieval, diagnosis, investigation, evaluation, and screening will be the necessary stages before the final decision. These phases can be combined into fewer ones. However, there has been no research to examine the effect of these phases on investment decision-making in tourism, especially MICE.

For this reason, the phase that precedes decision-making is represented by stakeholder perceptions of destination attributes and recognition of the benefits of MICE. The availability of information about destination attributes allows investors to choose alternatives to maximize profits. Meanwhile, the lack of information affects the benefits and increases the risk that can lead to bankruptcy. Therefore, decision-making ultimately refers to the dimensions of action taken after analyzing the destination attributes and benefits of MICE in the form of decisions to support, participate and invest in MICE activities [15].

VI. Relationship between Variables

A. The Effect of Destination Attributes on MICE Benefits.

Getz & Page stated that MICE tourism, an event-based business, requires facilities that are part of MICE destination attributes [25]. The completeness of these facilities will shape the image of the destination and will further provide benefits for tourist visits to carry out MICE activities at the destination. Thus destination attributes play a crucial role in winning the competition to seize MICE activities and provide multiplier effect benefits to various sectors in a MICE-destination.

H1: Destination attributes affect the MICE benefits.

B. The Effect of MICE Benefits on Decision Making

Leong defines MICE as a type of tourism in which participants are gathered to achieve certain goals. It is known as the event business in Australia, the MICE industry in Europe, and MICE tourism in Asia and North America [8]. The benefits derived from organizing MICE activities are recognized as affecting decision-making. On the other hand, benefits are the initial stage of the investment decision process which is described as a driving factor because it includes some form of incentives. Decision-making depends on investment return rate, risks, and expected costs [26].

H2: MICE Benefits has a positive and significant effect on Decision-Making.

C. The Effect of Destination Attributes on Stakeholders' Decision Making

Before making a decision, stakeholders need to conduct an analysis and feasibility study of the project to be implemented. This is done so that the success of the implemented project can be predicted. For this reason, stakeholders of MICE activities can conduct analysis and feasibility studies through information about destinations such as the destination attribute.

Decision-making in investment in the MICE sector should be made after analyzing a detailed feasibility study of the project in question. The availability of information about destinations along with destination attributes allows investors to choose alternatives to maximize profits. Investors need to base their investment decisions on complete and accurate data, while a lack of information will affect the investment benefits and increase the risk of loss. These risks are related to uncertainty and insufficient information that can even waste company resources and lead to bankruptcy [27].

H3: Destination attributes have a positive and significant impact on investment decision-making in the MICE sector.

D. The Effect of Destination Attributes on Decision Making through the MICE Benefits.

Good destination attributes will bring benefits to the MICE implementation so that it will affect stakeholders in making decisions to invest in this sector. The benefits of MICE activities are providing access to new technologies, attracting high-spending visitors, providing high yields and returns per capita, increasing international economic contacts. Moreover, the benefits of MICE activities also are creating more economic multiplier effects and competitiveness, bringing more visits outside of peak seasons, causing visitors to stay longer and to have higher expenses than ordinary tourists (Huo, 2014). Furthermore, MICE contributes to community building, urban renewal, and the growth of national identity. Through MICE contribution benefits, it will affect decision-making [19], [20].

H4: Destination attributes have a positive and significant impact on stakeholder decision-making through the MICE benefits.

VII. Research Methods

A survey with a quantitative approach was conducted in this study. The independent variable is the destination attributes and the moderating variable is benefits. Moreover, The dependent variable is decision-making. The research population used in this study were stakeholders represented by both government and privately owned travel agencies, tour operators, investors, restaurant owners, event organizers, hoteliers, and other parties who have held MICE activities in Medan City. The number of samples, determined using the Lemeshow formula, was 96 people. Furthermore, the data were analyzed using the PLS path model.

VIII. Data Analysis and Results

A. Descriptive Analysis

Descriptive analysis was conducted to determine respondents' perceptions of each variable. The average respondent's response to the destination attributes is in the category of 3.78 or is at a good level or has met 75.6% of the specified criteria. Respondents' perceptions of the benefits of organizing MICE in Medan City are in the category of 3.90 or are at the good category or reaching 78% of the specified criteria. Respondents' perceptions of Decision Making on the implementation of MICE in the city of Medan are in the category of 4.20 or are at the good category or have reached 84% of the specified criteria.

B. Analysis of the Measurement Model.

1) Validity and Reliability Test

Analysis of the measurement model is carried out by conducting validity and reliability tests (outer model), consisting of Convergent validity, average variance extracted (AVE), to see composite reliability, discriminant validity. Convergent validity can be seen in the outer loading. The results of the outer loadings for measuring the convergent validity of the measurement model (instruments) show a score above 0.50-0.70, except for indicators Y1.1, Y1.8, and Y1.9. Therefore, the three indicators were removed and the measurement of the model was carried out again. Composite reliability is above 0.6 with an alpha above 0.5. Thus, the measurement model is valid and reliable. Convergent validity and AVE can be seen in Table II.

Discriminant validity can be assessed from the correlation value of the variable with itself and with other variables. If the correlation value against the variable itself is higher than the correlation value with other variables, then the discriminant variable validity is satisfied.

TABLE I. Correlation Value

	Destination Attributes	**Benefits**	**Decision Making**
Destination Attributes	0.726		
Benefits	0.604	0.784	
Decision Making	0.469	0.619	0.857

From the Table I, it can be seen that all variables have the highest correlation on themselves compared to the correlations on other variables. Thus, the discriminant validity requirements in this study are met.

TABLE II. Convergent Validity and AVE

No.	**Variable**	**Indicator**	**Loading Factor**	**Alpha**	**Results**	**Composite Reliability**	**Alpha**	**Results**
1	Destination Attributes (X)	$X_{1.1}$	0,775	0,000	Valid			
		$X_{1.2}$	0,610	0,000	Valid			
		$X_{1.3}$	0,772	0,000	Valid			
		$X_{1.4}$	0,718	0,000	Valid			
		$X_{1.5}$	0,741	0,000	Valid	0,847	0,527	Reliable
2	Benefits (Y)	$Y_{1.2}$	0,773	0,000	Valid			
		$Y_{1.3}$	0,740	0,000	Valid			
		$Y_{1.4}$	0,795	0,000	Valid			

		$Y_{1.5}$	0,698	0,000	Valid			
		$Y_{1.6}$	0,834	0,000	Valid			
		$Y_{1.7}$	0,891	0,000	Valid			
		$Y_{1.10}$	0,741	0,000	Valid	0,917	0,615	Reliable
3	Decision Making (Z)	$Z_{1.1}$	0,846	0,000	Valid			
		$Z_{1.2}$	0,927	0,000	Valid			
		$Z_{1.3}$	0,792	0,000	Valid	0,892	9,734	Reliable

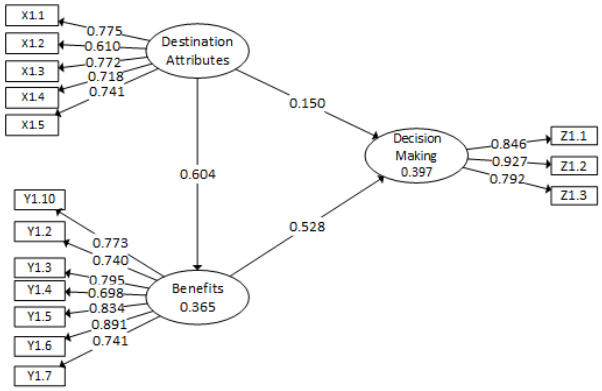


Fig. 1. PLS Research Model

2) Multicollinearity Test

Multicollinearity occurs when the VIF (Variance Influence Factor) value is > 10. It can be seen from Table III that there is no multicollinearity, because all VIF values are < 10.

TABLE III. VIF VALUE

Indicator	VIF	Indicator	VIF	Indicator	VIF
$X_{1.1}$	1,682	$Y_{1.2}$	1,934	$Y_{1.7}$	2,196
$X_{1.2}$	1,237	$Y_{1.3}$	3,034	$Y_{1.10}$	2,653
$X_{1.3}$	3,034	$Y_{1.4}$	2,053	$Z_{1.1}$	1,856
$X_{1.4}$	2,053	$Y_{1.5}$	3,360	$Z_{1.2}$	2.458
$X_{1.5}$	1,643	$Y_{1.6}$	3,655	$Z_{1.3}$	1,733

IX. STRUCTURAL MODEL

Measurement of the structural model or the measurement of the inner model was carried out to see the relationship between the construct, significance value and R-square of the research model. The results of the research model can be seen as shown in Fig.1.

The destination attributes have a positive and significant direct effect on benefits (0.604; P-value <0.05). The hypothesis is accepted. The destination attribute has a positive but not significant direct effect on decision-making (0.150; P-value>0.05). The hypothesis is rejected. Benefit has a positive and significant direct effect on decision-making (0.528; p-value <0.05). The hypothesis is accepted. Destination attributes have a positive and significant influence through benefits on decision making (0.319; p-value <0.05). The hypothesis is accepted. Furthermore, in the figure above, it can also be seen that the coefficient of determination (R^2) of the endogenous benefit and decision-making variables is 0.365 and 0.397. This indicates that the model is categorized as moderate in explaining the variation of the proposed model.

X. DISCUSSION AND CONCLUSIONS

The results of the study reveal that stakeholder decision-making to support, participate and invest in Medan City as one of the tourist destinations in Indonesia is influenced by destination attributes thorough knowledge of the benefits of MICE activities. Therefore, to increase the desire to invest in the MICE sector, it is necessary to disseminate information to stakeholders regarding the benefits of MICE events to create a multiplier effect for the regional economy. In addition, the government also needs to complete and improve the attributes of destinations in Medan City in an effort to increase the desire of stakeholders to invest in the MICE sector. The positive attitude of respondents in making decisions to support, participate and invest in the MICE sector gives hope that businesses in the MICE sector will continue to receive support during and after the COVID-19 pandemic.

The insignificant relationship between destination attributes and decision-making could be due to the Covid-19 problem that has an influence on decision-making in the MICE business. Therefore, strict rules are needed for the CHSE (Clean, Heathy, Safety, Environment sustainability) program in the new normal period.

Stakeholders' positive perception of the destination attributes of Medan City is the growth of stakeholder confidence in the attributes of the Medan City destination so that the government needs to strengthen destination attributes such as facilities for venues, quality of human resources, competitive prices, ease of transportation, and the ability to provide quality service for MICE participants in customs, telecommunications, health care, and eligible employees. Both the government and private sectors need to encourage the growth of professionals in the MICE field with the availability of PCO (professional Conference Organizer), PEO (Professional Exhibition Organizer), CVB (Conference Visitor Berau) so that they are able to promote destinations, build databases and information about MICE services and organizations in Indonesia. Medan city.

ACKNOWLEDGEMENTS

The authors wish to thank State Polytechnique of Medan who have funded this study. The authors would also like to express special gratitude towards the Director and the Unit of Reseach Development and Outreach Program (P3M) of State Polytechnique of Medan.

REFERENCES

[1] K. Safitri, "Industri MICE Diprediksi Rugi hingga Rp 7 Triliun Akibat Covid-19," 2020. https://money.kompas.com/read/2020/06/03/095847926/industri-mice-diprediksi-rugi-hingga-rp-7-triliun-akibat-covid-19. (accessed Mar. 22, 2021).

[2] R. Chavan and S. S. Bhola, "Gap Analysis of stakeholders' perception in tourism industry," MERC Glob. Int. J. Manag., vol. 1, no. 1, 2013.

[3] W. Zhiyi and Z. Hanyu, "The impacts of transformation of the destination governance structure on the perceptions of stakeholders-the case study in Simian Mountain," 2011.

[4] E. Konovalenko, "Assessing stakeholder needs and expectations for tourism development in the Barents Region: An analysis," Lapl. Inst. Tour. Res. Educ., 2012.

[5] L. K. Chiu and O. A. Ananzeh, "The role of MICE destination attributes on forming Jordan touristic image," Acad. Res. Int., vol. 3, no. 1, p. 267, 2012.

[6] L. K. Chiu and O. A. Ananzeh, "Evaluating the Relationship between the Role of Promotional Tools in MICE Tourism and the Formation of the Touristic Image of Jordan," Acad. Tur. Innov. J., vol. 5, no. 1, pp. 59–73, 2012.Y. Huo, "Meeting planners' perception on convention

destination attributes: Empirical evidence from six major Asian convention cities," J. Bus. Inq., vol. 13, no. 2, pp. 74–84, 2014.

[7] A. M. W. Leong, "Macau MICE tourism with a closer review of the Las Vegas and Atlantic City experiences," J. Macau Univ. Sci. Technol., vol. 1, no. 2, pp. 55–70, 2007.

[8] C. K. H. Lau, S. Milne, and C. S. Johnston, "MICE, ICT and local economic development: The case of Te Kahurangi, New Zealand," in Journal of Convention & Event Tourism, 2005, vol. 7, no. 1, pp. 61–75.

[9] C. M. Rogerson, "Unpacking tourism SMMEs in South Africa: structure, support needs and policy response," Dev. South. Afr., vol. 22, no. 5, pp. 623–642, 2005.

[10] Y. Yoon, D. Gursoy, and J. S. Chen, "Validating a tourism development theory with structural equation modeling," Tour. Manag., vol. 22, no. 4, pp. 363–372, 2001.

[11] M. Lawrence and V. McCabe, "Managing conferences in regional areas: a practical evaluation in conference management," Int. J. Contemp. Hosp. Manag., 2001.

[12] B. Azheri, Corporate social responsibility: Dari voluntary menjadi mandatory. RajaGrafindo Persada, 2011.

[13] A. Nicolaides, "Tourism Stakeholder Theory in practice: instrumental business grounds, fundamental normative demands or a descriptive application?," 2015.

[14] O. Alananzeh, M. Al-Badarneh, A. Al-Mkhadmeh, and O. Jawabreh, "Factors influencing MICE tourism stakeholders' decision making: The case of Aqaba in Jordan," J. Conv. Event Tour., vol. 20, no. 1, Jan. 2019, doi: 10.1080/15470148.2018.1526152.

[15] C. Jönsson and D. Devonish, "Does nationality, gender, and age affect travel motivation? A case of visitors to the Caribbean island of Barbados," J. Travel Tour. Mark., vol. 25, no. 3–4, pp. 398–408, 2008.

[16] N. Fitri, "Penerapan Konsep Quadruple Helix untuk mengembangkan Kreatifitas dan Inovasi Industri MICE," 2020.

[17] M. J. Lee and K.-J. Back, "Effects of destination image on meeting participation intentions: Empirical findings from a professional association and its annual convention," Serv. Ind. J., vol. 27, no. 1, pp. 59–73, 2007.

[18] D. Getz, "Event tourism: Definition, evolution, and research," Tour. Manag., vol. 29, no. 3, pp. 403–428, 2008.

[19] A. L. K. Kay, "China's convention and exhibition center boom," in Journal of Convention & Event Tourism, 2005, vol. 7, no. 1, pp. 5–22.

[20] M. Merková, J. Drábek, and D. Jelačić, "Application of Risk Analysis in Business Investment Decision-Making," Drv. Ind., vol. 64, no. 4, 2013, doi: 10.5552/drind.2013.1317.

[21] H. Wei and C. Christodoulou, "An examination of strategic foreign direct investment decision processes: the case of Taiwanese manufacturing SMEs," Manag. Decis., 1997.

[22] N. Sykianakis and A. Bellas, "The foreign direct investment decision-making process," Manag. Audit. J., 2005.

[23] M. Niskanen and J. Niskanen, "The determinants of firm growth in small and micro firms-Evidence on relationship lending effects," Available SSRN 874927, 2007.

[24] D. Getz and S. J. Page, "Progress and prospects for event tourism research," Tour. Manag., vol. 52, pp. 593–631, 2016.

[25] A. Virlics, "Investment decision making and risk," Procedia Econ. Financ., vol. 6, pp. 169–177, 2013.

[26] I. KOLENKA and I. HAJDÚCHOVÁ, Financovanie podniku II., 3rd ed. Zvolen: Technická univerzita vo Zvolene, 2008.

ENGINEERING SECTION

Invited Speaker :
Dr. Eng. Noerdin Basir

"Implementation of Automatic identification system (ais) in supporting the marine infrastructure development policies in the Malacca Straits"

POLITEKNIK NEGERI BENGKALIS

Design and Implementation of an Arduino Based Smart Fingerprint Authentication System for Key Security Locker

1st Bong Siaw Wee
Department of Electrical Engineering
Politeknik Mukah
Mukah, Sarawak, Malaysia
shaweibong2016@gmail.com

Abstract*—Human beings have always prioritized security. Simple mechanical locks with the key as the authentication factor were among the earliest forms of security. A key is a tiny, shaped piece of metal with incisions made to match the wards of a certain lock that is placed into the lock and twisted to open or close it. People tend to misplace the keys and it is troublesome for other users to use the keys because it is hard to find the right key. Moreover, the keys are without any protection. Anyone besides the user can easily access the keys. Therefore, an Arduino Based Smart Fingerprint Authentication System for Key Security Locker was proposed and developed in this research. The Key Security Locker System is designed to make sure that the keys are stored in a more organized and effective location. Biometrics is the science of determining an individual's identification based on physical, chemical, or behavioral characteristics. The importance of biometrics in modern society has been strengthened by the necessity for large-scale identity management systems, the operation of which is dependent on the precise assessment of an individual's identity in the context of many applications. This technology recognizes authorized personal's unique fingerprints and allows them access. To utilize the fingerprint scanner, the user must place their finger on it. The event will capture new human minutiae through fingerprint scanning. These new minutiae will be compared to those in the database to determine if the individual is authorized or non-authorized. The Liquid-Crystal Display (LCD) will show "Fingerprint Match" if the fingerprint matches the fingerprint in the database. Then, the microcontroller will instruct relay and solenoid lock to unlock the locker door. If the fingerprint is not matched, the LCD will appear "Not Matching". Then, the locker door will remain locked. Hence, the proposed system provides better security, higher efficiency, and in many instances, increased user convenience due to it being built based on a biometric system.

Keywords—Arduino UNO, Security, Key Box Locker, Biometrics, Fingerprint Authentication System

I. INTRODUCTION

Security has always been a concern of paramount importance to human beings [1]. When it comes to security, it is one of the most pressing problems in today's hectic, competitive environment, where a human being is unable to offer protection for his possessions daily. Keeping our staff safe has always been at the top of our priority list. Today's security includes a wide range of software and hardware, such as web-based security services, biometrics, and personal gadgets with built-in security levels [2].

A traditional lock with a key was the favored locking mechanism. This is an old-fashioned locking mechanism with a few flaws, such as the key being easily copied and the lock is easily opened by an unauthorized person [3]. Several approaches have been reported in the literature for security solutions such as a radio-frequency identification (RFID) card, keypad, pin, password, or Internet of Things (IoT) by unlocking the device with a mobile phone [4-8]. These systems have the same advantages and disadvantages, and this form of security lock may open the system from any security level. To unlock the system, the user either enters a Personal Identification Number (PIN) or swipes an RFID card [9]. This system lacks a security level chain, which would improve security. An RFID card, which stands for radio frequency identification, can be used to unlock the system. The scanner scans the radio frequency to determine if the identity is allowed or unauthorized. However, the primary drawback of this approach is that the passwords could be hacked and a card may be stolen or lost. As a result, the user must handle it with caution [10]. Besides, this system also has no alert system in case of a break-in, or an unauthorized person tries to unlock the door.

IoT is an abbreviation for the internet of things, which is used indoor locks through a wireless connection. The user can utilize IoT-enabled applications on his smartphone to unlock the door lock. With a single touch, the user may simply open or lock the system However, IoT is supported only by an internet connection [11, 12]. The most striking problem for those models operating with smartphones is the risk of being unable to open the door if the smartphone's battery runs out. Unfortunately, the autonomy of these gadgets is limited, and whenever the user is unable to charge their smartphone, so the user will face the risk of being unable to open the door.

Biometrics provides a natural and dependable approach to some elements of identity management by employing fully automated or semi-automated systems to recognize persons based on biological features [13, 14]. It is possible to create an identity based on who you are rather than what you own, such as an identification (ID) card, or what you know, such as passwords, by utilizing biometrics. Biometrics may be used to augment ID cards and passwords in some applications, adding

an extra layer of protection. This type of setup is sometimes referred to as a dual-factor authentication method.

Therefore, an Arduino Based Smart Fingerprint Authentication System for Key Security Locker was proposed and developed in this research. The proposed method is more efficient due to biometric identifiers cannot be readily misplaced, faked, or shared, they are regarded more trustworthy for person recognition than the traditional token (e.g. keys or ID cards) or knowledge (e.g. password or PIN) based techniques. Fingerprint technology is the most generally recognized and established biometric approach. It is the most user-friendly and puts a better level of protection at your fingertips. Fingerprint recognition improves security, increases efficiency, and increases consumer convenience. Furthermore, fingerprint recognition is the most commonly used biometric feature. It is often assumed that each finger has a distinct pattern. Given that there are around 6.5 billion live individuals in the world, and assuming that each person has ten fingers, there are 65 billion distinct fingers! Fingerprints were originally used as a form of identification over a century ago.

Fingerprint recognition is one of the most secure systems due to everyone has various forms of fingerprints, fingerprint identification. Therefore, unauthorized access can be restricted by creating a lock that saves the fingerprints of one or more authorized users and opens the system when a match is detected. The Arduino serves as a data storage device for users. the skin on our hands and soles has a flow-like pattern of ridges on each fingertip that is unique and unchangeable, biometric authorization shows to be one of the greatest features. As a result, fingerprints are a one-of-a-kind form of identification for everyone. The usage of fingerprint scanners in contemporary hand-held devices such as mobile phones and computers demonstrates their popularity and dependability.

II. METHOD

A. *Block Diagram*

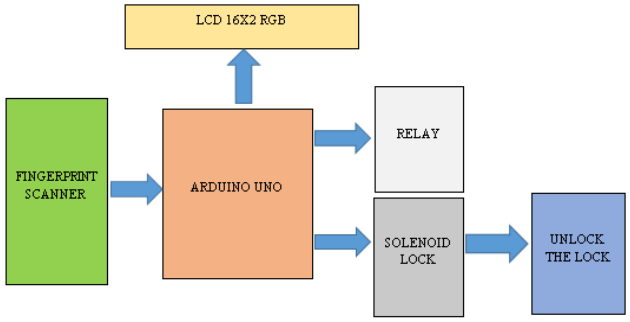


Fig. 1. Block diagram

The block diagram of an Arduino-based Smart Fingerprint Authentication System for Key Security Locker is shown in Fig. 1. To utilize the fingerprint scanner, the user must place their finger on it. The fingerprint scanner subsequently captures the fingerprint picture and sends it to the Arduino Uno. The Arduino Uno will search for the original fingerprint data that is already on the board. If the fingerprint matches one previously saved in Arduino, the Liquid-Crystal Display (LCD) will display "Fingerprint Match." The microcontroller will then instruct the relay and solenoid lock to open the locker door. If the fingerprint is not matched, the LCD will appear “Not Matching”. Then, the locker door will remain locked.

The Arduino is an open-source physical computing platform based on a single microcontroller board. The Arduino is used when there are interactions between inputs and outputs. It has 14 digital Input/Output (I/O) pins, six of which can be used as pulse width modulation (PWM) outputs, six analog inputs, a reset button, a power connector, a USB connection, and other features. It includes everything needed to support the microcontroller; simply connect it to a personal computer (PC) through a USB cable and power it up using an Alternating Current (AC)-to-Direct Current (DC) adapter or battery. It is used to control the output based on the input directives, for as utilizing a switch to control a light or motor.

The optical finger reader sensor, commonly known as a fingerprint reader. It is a piece of technological equipment that captures a digital image of the fingerprint pattern. The captured image is known as a live scan, which is then digitally processed. The distinctive characteristics of the fingerprint are extracted, and a fingerprint biometric template is generated. This biometric template has been saved and will be used for matching in the future [15].

The relay is a switch that operates electrically and consists of two main parts, namely electromagnetic (coil) and mechanical (set of switches). The electromagnet requires a low voltage to activate, which will provide via the Arduino. When once triggered, it will pull the contact to complete the high voltage circuit.

Solenoids are electromagnets that are made out of a huge copper wire coil with an armature in the middle. The slug is drawn into the center of the coil as it is charged. This allows the solenoid to pull from only one end.

An LCD is an electrical display module that employs liquid crystals to generate a visible image. The primary advantages of utilizing this module are its low cost, ease of programming, animations, and the fact that there are no restrictions on displaying unique characters, special animations, and so on.

B. *Fingerprint Authentication System*

One of the most fundamental biometric traits is fingerprinting. Dactyloscopy is the field of study that deals with fingerprints. The papillary lines on the inside of human fingers are the subject of this study. Every person's papillary lines, as well as their form, course, and orientation, are unique. It is feasible to establish numerous basic patterns that assist to categorise all of the forms based on the shapes created by the papillary lines. Four patterns are used as a criterion for classifying each fingerprint. As seen in Fig. 2, Arch, loop, and whorl are three primary fingerprint pattern ridges. The following is an explanation for the three primary fingerprint pattern ridges:-

- Arch: The ridges enter from one side of the finger, ascend in the center to create an arc, and then leave from the other side.
- Loop: The ridges enter from one side of the finger, curve, and leave on the same side.
- Whorl: Ridges on the finger develop in a circular pattern around a central point. Ridges develop in a circular manner around a finger in the whorl pattern.

Fig. 2. Types of the fingerprint pattern [16]

C. Experimental Setup

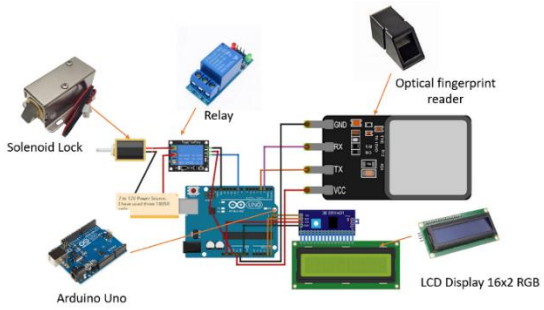


Fig. 3. Experimental setup for the proposed system

Fig. 3 presents the experimental setup for an Arduino Based Smart Fingerprint Authentication System for Key Security Locker. Based on Fig. 4, the optical fingerprint is as an input and is connected to the Arduino board. Pin 2 is assigned to the transmitter (Tx), whereas Pin 3 is assigned to the receiver (Rx). The Arduino Uno supplies 5v to power the fingerprint reader. Fig. 5 shows the connecting of Vcc to 3.3v, the LCD 16x2 is linked. Then, connect the Serial Clock (SCL) to pin A5 and the Serial Data (SDA) to pin A4.

ARDUINO Pin	Finger Sensor
2	Green (Tx)
3	White (Rx)
5V	Red (VCC)
GND	Black (GND)

Fig. 4. The connection between an Arduino board and an optical fingerprint reader

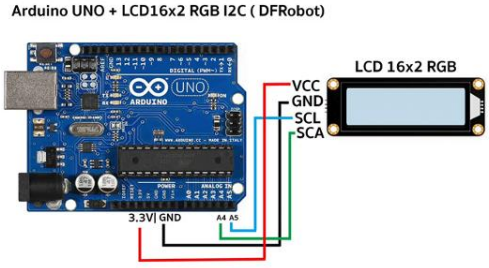


Fig. 5. Connection LCD 16x2 and Arduino board

Fig. 6 shows the connection between the Arduino board with the relay module and the solenoid lock. The relay module acts as an on/off switch. IN1 is connected to pin 9 and 5v is supplied to the relay module from the Arduino board. Besides, NO is connected to the solenoid lock-in anode polarity. The negative polarity of the solenoid lock and pin COM are both connected to the power source 12V.

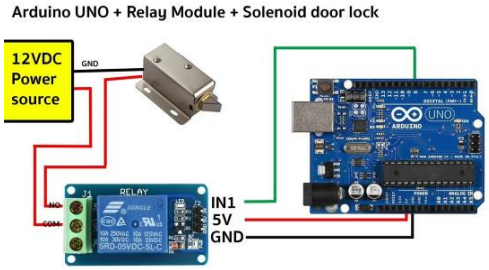


Fig. 6. The connection between an Arduino with a relay module, and a solenoid lock

Fig. 7 depicts the flow chart for this proposed system. When the user places a finger on the scanner, the fingerprint sensor illuminates the surface of the finger and records the minutiae using a charge-coupled device (CCD) with an 8 bit per dot resolution of 500 dpi (dots per inch) in grey-scale. The collected data is saved in a local database once it has been transformed into a digital signal. Every time a fingerprint is scanned, a new human minutia is captured. These new details will be compared to those in the database to determine if the individual is permitted or not.

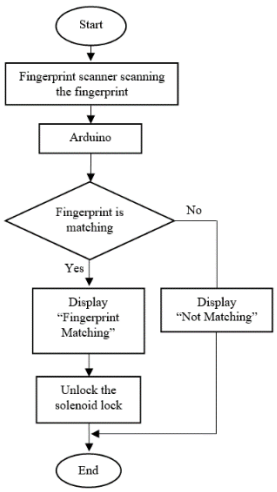


Fig. 7. Flow chart of the proposed system

D. Software Arduino (IDE)

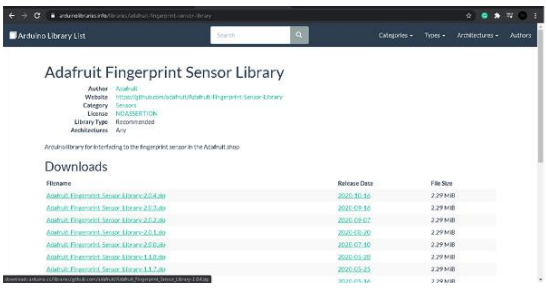


Fig. 8. Adafruit fingerprint library in Google

The Arduino Adafruit fingerprint library can be downloaded from the Google website as shown in Fig. 8. After downloading the library, copy and paste the Adafruit fingerprint library into the Arduino library. Fig. 9 shows the coding for Adafruit_Fingerprint.h while Fig. 10 shows the coding for the proposed system in the Arduino software.

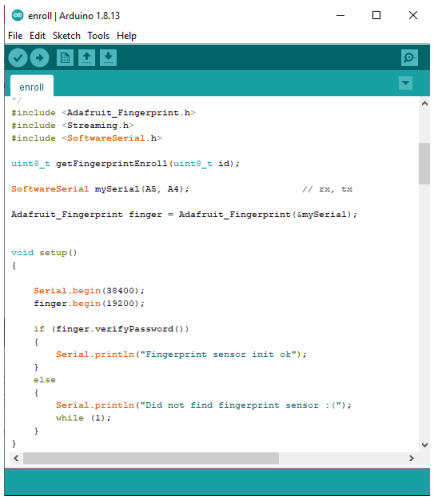


Fig. 9. Coding for Adafruit_Fingerprint.h

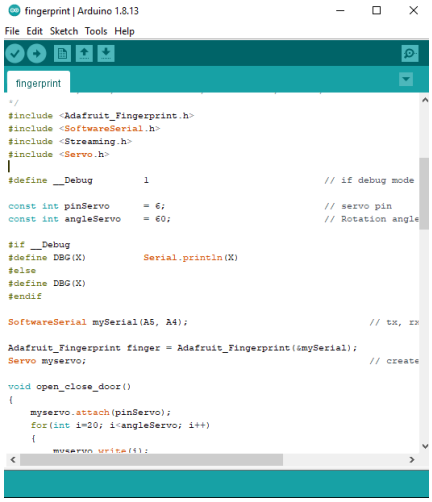


Fig. 10. Coding for Fingerprint

III. RESULTS AND DISCUSSION

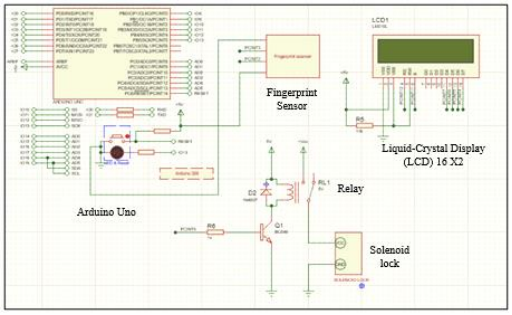


Fig. 11. Schematic design

The schematic diagram for an Arduino-based Smart Fingerprint Authentication System for Key Security Locker can be found in Fig. 11. The input voltage for Arduino Uno is 5V while the solenoid electric lock requires 12V. The Arduino Uno MCU board serves as the circuit's brain. It contains 14 digital I/O pins, six analog inputs, 32k flash memory, a 16MHz crystal oscillator, a USB connection, a power connector, an In-Circuit Serial Programming (ICSP) header, and a reset button, among other features. It can be programmed using Arduino Integrated Development Environment (IDE) software. The fingerprint sensor is an optical fingerprint scanner that also serves as an input. The user can store fingerprint data in the module and set it for identification in 1:1 or 1:N mode. The sensor's Tx and Rx pins are linked to Arduino digital pins 2 and 3 for serial communication. The LCD is used to display the output messages. As a result, a 5V Relay (RL1) is needed to operate the lock. The Normally Open (N/O) contacts of RL1 are linked to Ground through CON3 (GND). Connector CON3 is used to connect an electronic door-lock solenoid. It's an electromagnet with a large coil of copper wire in the center and an armature in the center. The slug is drawn into the center of the coil when it is electrified. The solenoid can now be moved to one end.

(a) Front view (b) Side view

Fig. 12. Front view and side view of the proposed system

Fig. 12 shows the front view and side view of an Arduino Based Smart Fingerprint Authentication System for Key Security Locker. The first step when accessing this Key Security Locker System, the user needs to place their fingerprint on the optical fingerprint scanner. The optical fingerprint scanner will be scan the fingerprint.

(a) Fingerprint scanner (b) Fingerprint matching

Fig. 13. Fingerprint matching

Based on Fig. 13, the fingerprint scanning event will capture a new humans minutiae. These new minutiae will be compared to those in the database to determine whether the person is authorized or non-authorized. If the fingerprint matches one in the database, the Liquid Crystal Display (LCD) will display "Fingerprint Match.". Then, the microcontroller will instruct relay and solenoid lock to unlock the locker door as shown in Fig. 14.

(a)

(b)

Fig. 14. When the fingerprint is matching, a solenoid lock will unlock the locker

If the fingerprint is not matched, the LCD will appear "Not Matching". Then, the locker door will remain locked as shown in Fig. 15.

Fig. 15. Not matching

Table I shows that an Arduino Based Smart Fingerprint Authentication System for Key Security Locker is a locking system that uses a fingerprint sensor module to authenticate the user's fingerprint. Arduino is used to power the fingerprint sensor module. The Fingerprint module determines whether or not a particular fingerprint is allowed. The locking system uses the user's fingerprint to open the system.

TABLE I. RESULTS FOR AN ARDUINO BASED SMART FINGERPRINT AUTHENTICATION SYSTEM FOR KEY SECURITY LOCKER

Finger-print	Type of Fingerprints	Matching/ Not Matching	Database	Key Security Locker Door
1	Tented Arch	Not Matching	No	Close
2	Central Pocket Loop	Matching	Yes	Open
3	Plain Left Loop	Matching	Yes	Open
4	Plain Loop	Not Matching	No	Close

Based on Table II, the common comparison is the Key Security Locker System using a fingerprint lock while product A still uses a traditional lock that uses a key. Besides, the Key Security Locker has a high level of security due to it using a fingerprint system that allows a valid fingerprint to access it. The Key Security Locker is more expensive than product A since it uses electronic components and coding to construct it. Furthermore, the Key Security Locker uses a power supply while a traditional lock does not. This locker has been developed with a more modern and stylish design.

TABLE II. COMPARISON OF PROPOSED SYSTEM WITH TRADITIONAL LOCK

Product	Product A	Proposed System
Type of lock	Traditional lock (using key)	Fingerprint lock
Security	Low-level security/easy to access	High-level security/difficult to access
Power Supply	Do not need supply	Need power supply
Design	Traditional look	Modern and stylish

An Arduino Based Smart Fingerprint Authentication System for Key Security Locker is designed to ensure the key storage is in a more organized and effective location. Fingerprints are unique to each person and cannot be lost or stolen, making them extremely accurate and dependable. This suggested device can safely store a large number of keys in one location. This technology recognizes authorized personal's unique fingerprints and allows them access.

IV. CONCLUSION

An Arduino-based Smart Fingerprint Authentication System for Key Security Lockers has been created, which offers improved security, efficiency, and user convenience in many cases. By using this innovative product, the keys can be stored in a more organized and secure location. Only authorized users can open the door of this key security locker to access the key insides of this locker. Thus, with biometric technologies like fingerprint scanning, authentication can be made more secure and convenient. Biometric technology's use will continue to expand in the future, and it will be employed in even more sectors that affect our daily lives.

REFERENCES

[1] H. Hassan, R. A. Bakar, and A. T. F. Mokhtar, "Face recognition based on auto-switching magnetic door lock system using microcontroller," in 2012 International Conference on System Engineering and Technology (ICSET), IEEE, 2012.

[2] Y. T. Park, P. Sthapit, and J.-Y. Pyun, "Smart digital door lock for the home automation," in TENCON 2009-2009 IEEE Region 10 Conference, IEEE, 2009.

[3] Keriya, F.R.B., "Development of IoT–based locking system," B.S. dissestation, Dept. Faculty of Electrical and Electronic Engineering (UTHM), Johor, Malaysia 2019.

[4] Z. Mumtaz, Z. Ilyas, A. Sohaib, S. Ullah, and H. A. Madni, "Design and implementation of user-friendly and low-cost multiple-application system for smart city using microcontrollers," arXiv preprint arXiv:2010.07016, 2020.

[5] Y. Motwani, S. Seth, D. Dixit, A. Bagubali, and R. Rajesh, "Multifactor door locking systems: A review," Materials Today: Proceedings, 2021.

[6] S. Goyal, P. Desai, and V. Swaminathan, "Multi-level security embedded with surveillance system," IEEE Sensors Journal, vol. 17, no. 22, pp. 7497-7501, 2017.

[7] S. Hossain, M. I. Ahmed, and M. N. Mostakim, "A prototype of automated vault locker solution for industrial application," in 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), IEEE, pp. 1-7, 2019.

[8] A. M. A. Bakry, and D. R. Rajaa, “Smart phone-Arduino based of smart door lock/unlock using RC4 stream cipher implemented in smart home,” International Journal of Advanced Computer Technology, vol. 5, no. 5, pp. 14-18, 2016.

[9] G. K. Verma, and P. Tripathi, “A digital security system with door lock system using RFID technology,” International Journal of Computer Applications, vol. 5, no. 11, pp. 6-8, 2010.

[10] N. Meenakshi, M. Monish, KJ. Dikshit, and S. Bharath, “Arduino based smart fingerprint authentication system,” in 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), IEEE, pp. 1-7, 2019.

[11] T. Adiono, S. Fuada, S. F. Anindya, I. G. Purwanda, and M. Y. Fathany, “IoT-enabled door lock system,” International Journal of Advanced Computer Science and Applications, vol. 10, no. 5, pp. 445-449, 2019.

[12] O. Doh, and I. Ha, “A digital door lock system for the internet of things with improved security and usability,” Advanced Science and Technology Letters, vol. 109(Security, Reliability and Safety 2015), pp. 33-38, 2015.

[13] T. Tanjin, and S. Akter, “Design and implementation of Arduino based home security system,” International Research Journal of Modernization in Engineering Technology and Science, vol. 3, pp. 1335-1338, 2018.

[14] H. S. Eldo, “Fingerprint based security system for ATM,” International Research Journal of Engineering and Technology, vol. 6, pp. 850 – 854, 2019.

[15] F. Hidayanti, F. Rahmah, and A. Wiryawan, “Design of motorcycle security system with fingerprint sensor using arduino uno microcontroller,” International Journal of Advanced Science and Technology, vol. 29, no. 5, pp. 4374-4391, 2020.

[16] S. B. Ahmed, and M. I. Razzak, “The minutiae based latent fingerprint recognition system: A comprehensive survey,” Proceedings of the International Conference on Internet of things and Cloud Computing, 2016.

Design and Implementation of an Arduino-Based Body Temperature and Pulse Rate Monitoring System

1st Bong Siaw Wee
Department of Electrical Engineering
Politeknik Mukah
Mukah, Sarawak, Malaysia
shaweibong2016@gmail.com

2nd Kong Shuet Fang
Department of Electrical Engineering
Politeknik Mukah
Mukah, Sarawak, Malaysia
shuetfang0829@gmail.com

Abstract*—Health is an important aspect of human existence, and vital indicators such as blood pressure, body temperature, respiration rate, and pulse rate/heart rate are widely known in the medical field. According to the World Health Organization (WHO), 2,200 people die in the United States (U.S.) each day as a result of heart disease. Each year, heart disease and stroke cost about 312.6 billion in health-care expenses. Especially during COVID 19 Pandemic outbreak, rural citizens need a more effective and affordable way for personal healthcare monitoring system before they decide to travel to the nearby medical center for medical assistance. Therefore, the Arduino-Based Body Temperature and Pulse Rate Monitoring System was proposed and developed in this project. The Arduino Software was used to write the programming for this project, which interfaced with the Arduino UNO board. The temperature sensor detects the temperature while the heartbeat sensor counts the pulse rate for a certain interval of time and calculates beats per minute. Both data are delivered to the Arduino UNO for transmission to reception. The results of the measurements will be shown on the Liquid-Crystal Display (LCD) screen. The measured data will be entered into the body temperature and pulse rate app. This App is a platform that presents a graph of the data so that the user can quickly see if the measurements have changed significantly. This app was created utilizing the AppSheet platform's user-friendly features and can be used on a number of platforms, including phones, tablets, and web browsers. The user can retrieve their measurement data anytime and anywhere so that early detection of the heart disease or symptoms of COVID-19. The typical normal temperature of the human body is between 36°C and 37 °C. Low fever is defined as 37.3°C-38 °C, whereas severe fever is defined as 38.1-41 °C. Another function for this innovation project can be used to measure the pulse rate. The normal pulse rate is 60 to 100 beats per minute. People who exceed 100 beats per minute are very dangerous, so they need to refer to the nearest clinic for further medical consultation.

Keywords—Pulse Rate, Heart disease, Body Temperature, Arduino UNO, COVID 19

I. INTRODUCTION

The fundamental criteria of normal levels of the body's vital signs, such as heart rate and body temperature, are used to assess physical health. The heart is the most essential organ in the human body since it is the primary organ responsible for blood circulation throughout the body [1]. Heart disease is the primary cause of mortality worldwide, according to the World Health Organization (WHO) [2]. Heart disease and stroke cost for approximately 312.6 billion dollars in health-care costs per year. Chest pain or discomfort, feeling weak, light-headed, or faint, and shortness of breath are all common symptoms of a heart attack. Pressure, squeezing, fullness, or discomfort may be felt. The advent of testing enables patients to better monitor their health. It is appropriate for individuals suffering from heart disease, diabetes, high blood pressure, and other conditions. Medical practitioners utilize heart rate to monitor patients' physical health, such as athletes who wish to monitor their heart rate to get the most out of their training. Despite huge differences in temperature outside the body, the human body's tendency is to regulate the temperature within a small, safe range. Cardiovascular disease refers to a group of illnesses that affect the heart or blood arteries. Coronary artery disease (CAD), such as angina and myocardial infarction, is a kind of CVD (commonly known as a heart attack). The underlying processes differ according to the illness. High blood pressure, smoking, diabetes, a lack of exercise, obesity, high blood cholesterol, a poor diet, and excessive alcohol consumption are all risk factors for heart disease.

People spend a lot of money to be healthy. Unfortunately, when things are non-invertible, individuals always discover that it is too late to obtain severe medical care [3]. Many people can be treated if early action are are taken in a timely way. However, most medical products on the market today have these major drawbacks, as well as limits in flexibility and portability. Furthermore, certain technologies [4-6] are available on the market that can give raw medical data computations for patients and clinicians, but the patients unable to converts med

ical measurement into the meaningful diagnoses due to lack of medical background and human literacy [7].

The need for a home-based self-health monitoring system has increased significantly in recent years. This may be achieved by employing a low-power, cost-effective, accurate, and easy-to-use gadget capable of measuring, displaying, and alerting the patient to vital data [8-10]. Therefore, an Arduino-Based Body Temperature and Pulse Rate Monitoring System was proposed and developed in this research. The benefits of this initiative are that treatment may be given to patients in priority to the ailment they have as compared to other patients,

and they can be hospitalized in emergency situations. Body temperature, heart rate, and fall detection are all vital indicators. The mass-weighted average temperature of body tissues and skin temperature are measured. Especially for rural citizens, they need a more effective and affordable apparatus for self-monitoring heartbeat and temperature because rural areas are hardly approached by medical personnel. Rural citizens need a more effective and affordable way for personal healthcare monitoring system before they decide to travel to the nearby medical center for medical assistance, especially during COVID 19 pandemic outbreak.

The paper is organized as follows: Sect. II describes the technique used to create this invention, which includes the prototype model research approach, system architecture and design, Arduino software, experimental setup, and design of the Body Temperature and Pulse Rate App. The findings of the Arduino-Based Body Temperature and Pulse Rate Monitoring System are evaluated in Section III. Finally, we conclude the paper in Sect. IV.

II. METHOD

A. *Prototype Model Research Method*

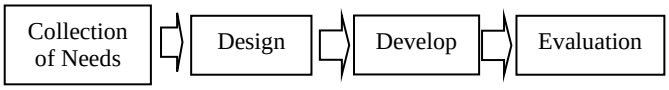


Fig.1. Prototype Model Research Method

Fig. 1 shows the prototype model research method for this paper. The demands of consumers for body temperature and pulse rate measurement devices are surveyed in this phase through conversations with sources and direct observation. Determine what instruments are necessary to address the problem after recognizing it. The gathering of needs is also carried out in the design of tools, which comprise any needed components.

This is the stage at which a tool is created to help with the problem of developing an Arduino-Based Body Temperature and Pulse Rate Monitoring System. The design addresses how the tools function, tool workflows, and code design, as well as the development of a series of tools that will lead to the production of tools in the form of prototyping [11].

Once the process of building a tool in the form of a prototype has been finished, the proposed system is presented to the user for an initial evaluation. It assists in evaluating the strengths and weaknesses of the working model, which includes numerous instruments that have performed well or poorly and may be a solution or not in order to create or repair this health monitoring system. The prototype should then be refined based on the user's comments and ideas. Based on the final prototype, the final monitoring system is built, extensively tested, and deployed to production.

B. *System Architecture and Design*

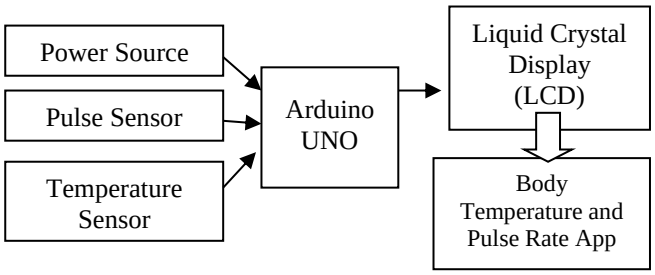


Fig 2. Health Monitoring System Block Diagram

The proposed health monitoring system may use the patient's fingertip to measure the patient's heart rate and surface body temperature and display the findings on the LCD screen. This block diagram consists of three parts, which are the transmission part, receiving part, and central processing unit as shown in Fig. 2. The diagram of the transmitter is composed of the controller. This controller is equipped with a capacity and the temperature sensor is connected. The controller is used to control the body temperature and heart rate sensors. The LCD is also linked to the body temperature, pulse rate, and Radio Frequency (RF) transmitter, which sends data to the controller's receiving circuit. The schematic design of the receiver is made up of a microcontroller that is powered by a power source and connected to the controller. A rack circuit to receive the signal transmitted by the transmitter and processed by the controller and then display on the LCD screen.

C. *Software Arduino (IDE)*

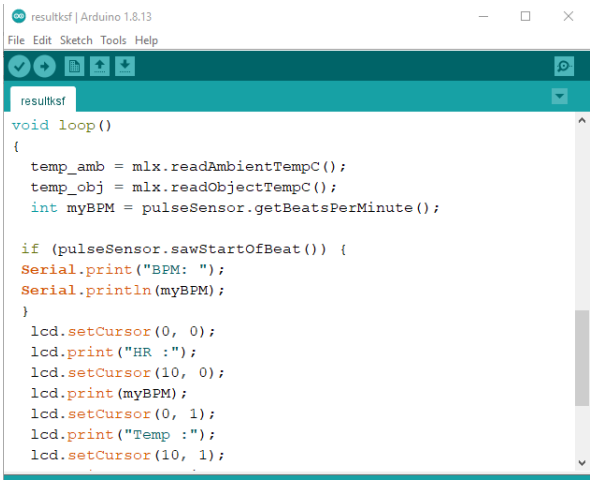


Fig. 3. Software development

The Arduino Integrated Development Environment (IDE) is a cross-platform application (for Windows, macOS, and Linux) with built-in C and C++ functions, as seen in Fig. 3. With the help of third-party cores, it is used to write and upload programmes to Arduino compatible boards as well as other vendor development boards [12]. The IDE's source code is available under the GNU General Public License, version 2 [13].

D. Experimental Setup

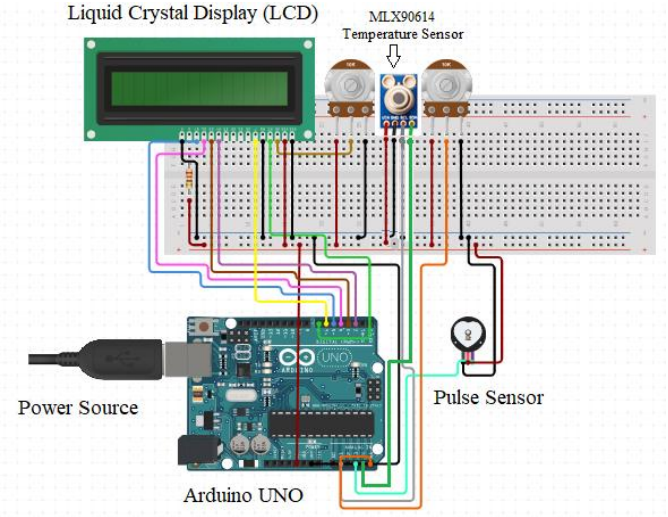


Fig. 4. Experimental setup for the proposed system

Fig. 4 presents the experimental setup for an Arduino-Based Body Temperature and Pulse Rate Monitoring System. The Arduino Uno is a microcontroller board that is open-source and operates on the Microchip ATmega328 microprocessor. This microcontroller serves as a CPU for measuring the user's body temperature and pulse rate. The power supply is utilized to power the entire circuit.

The MLX90614 is an infrared thermometer that monitors temperature without the use of any contact. The TO-39 container contains both the IR-sensitive thermopile detector chip and the signal conditioning ASIC. The MLX90614 includes a low noise amplifier, a 17-bit ADC, and a powerful DSP unit, all of which contribute to outstanding thermometer accuracy and resolution. The 10-bit PWM is developed to continuously transmit the observed temperature in the range of -20 to 120°C, with a standard output resolution of 0.14°C. Pulse Sensor is a well-designed, Arduino-compatible heart-rate sensor. The sensor is attached to a fingertip or earlobe and communicates with the Arduino via jumper wires. It also includes free open-source monitoring software for graphing your pulse in real time. The Pulse sensor has three pins; connect the pulse sensor's 5V and ground pins to the Arduino's 5V and ground, and the signal pin to the Arduino's A0. A Liquid-Crystal Display (LCD) is a flat-panel display or other electronically modulated optical device that uses liquid crystals' light-modulating capabilities in combination with polarizers. It is composed of two different states of matter: solid and liquid. On an LCD, a liquid crystal is utilized to generate a visible picture.

The circuit diagram of this system includes the transmitter (TX) section and receiver (RX) section. The temperature and heartbeat of the user's body, as well as the data sensed by the sensor, are transmitted to the ATmega328 in the TX portion. The transmitted data can be encoded into serial data over the air using a Radio Frequency (RF) module. The patient's body temperature is shown on the LCD through an antenna mounted at the end of a transmitter, and data from the transmitter is transmitted to the receiving end. The Arduino then processes the body temperature and pulse rate data and sends it to the LCD. The measured data will be entered into the body temperature and pulse rate app by the user. This App is a platform that presents a graph of the data so that the user can quickly see if the measurements have changed significantly. This App will be discussed in detail in sub-section E. The flow chart for this proposed system is depicted in full in Fig. 5.

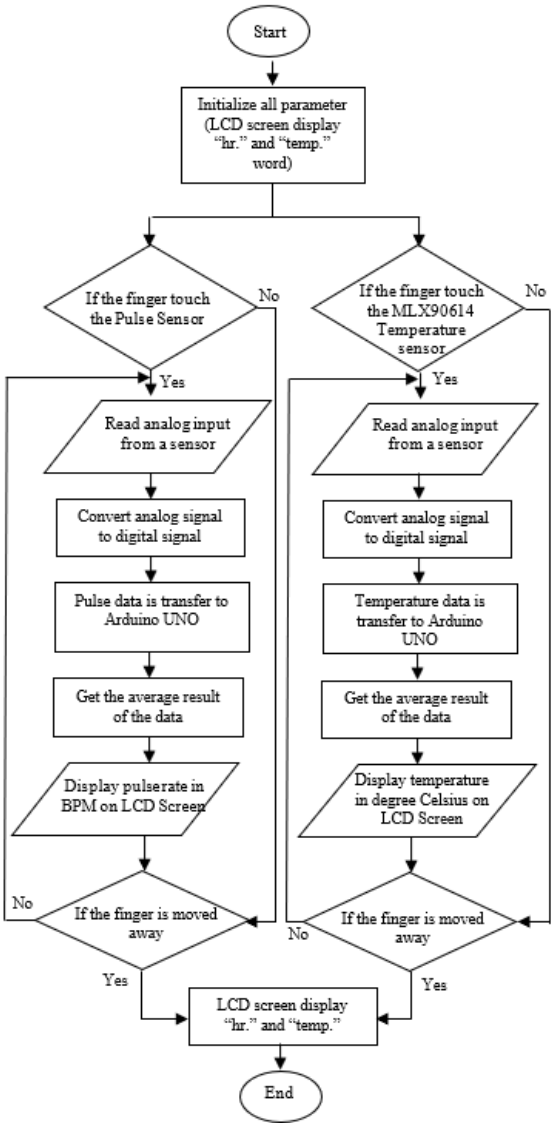


Fig. 5. Flow chart of the proposed system

E. Design of Body Temperature and Pulse Rate App

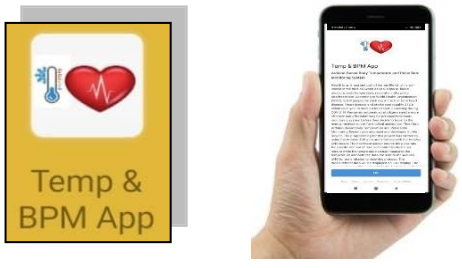


Fig. 6. Body Temperature and Pulse Rate App

Fig. 6 shows the design Body Temperature and Pulse Rate App. This App is created with the AppSheet Platform, as illustrated in Fig. 7, and consists of three (3) major steps:

- Step 1: First, all of this app's data is saved on the cloud, and Google Drive (specifically, spreadsheets) was chosen as the database storage design.
- Step 2: Second, the AppSheet Platform has been used to connect the data and develop the app with popular features in this platform.
- Step 3: Publish the app instantly and share the app with customers. The implication of this app could support a variety of devices: phones, tablets, and web browsers.

Fig. 7. Three important steps for Smart Educational App development

III. RESULTS AND DISCUSSION

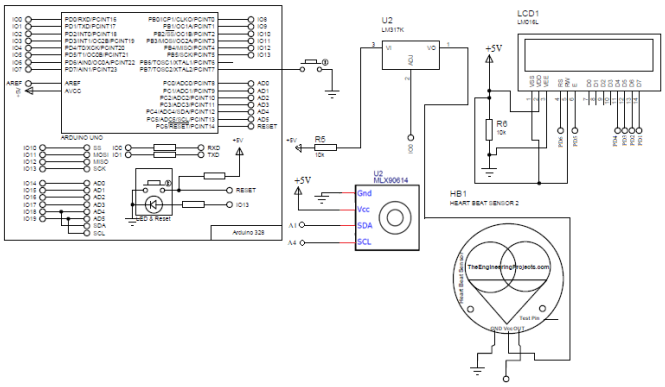
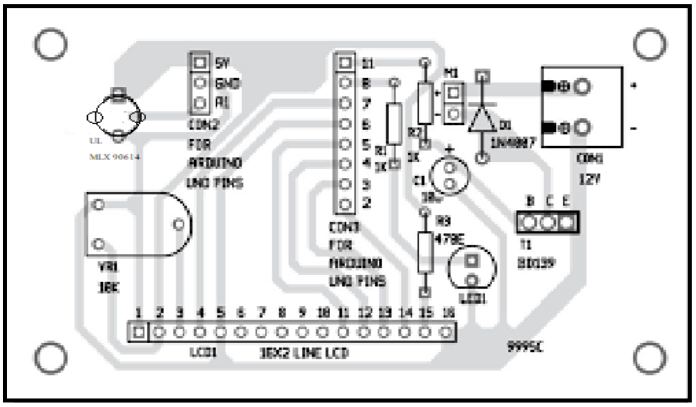
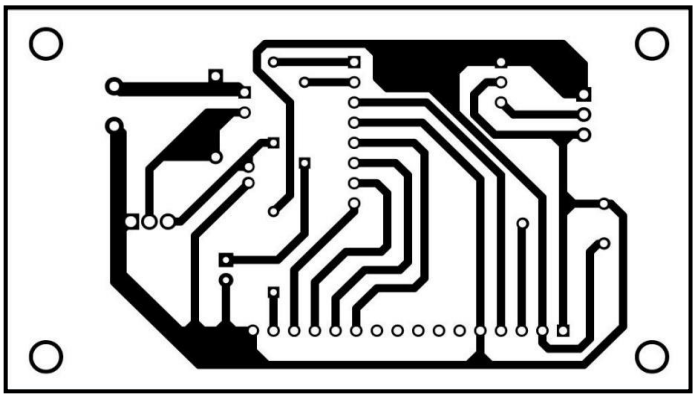


Fig. 8. Schematic Design

Fig. 8 shows the schematic design for an Arduino-Based Body Temperature and Pulse Rate Monitoring System. When the battery is inserted into the circuit, the Arduino begins reading the pulse rate from the pulse sensor and the ambient temperature from the MLX 90614 temperature sensor. An infrared LED and a phototransistor in the pulse sensor aid in detecting the pulse at the tip of the finger or earlobe. Its IR LED glows whenever it senses a pulse. The phototransistor detects the flash of the IR LED, and its resistance varies when the pulse is altered. The average adult's heart rate ranges between 60 and 100 beats per minute. To detect beats per minute (BPM), an interrupt is first configured to activate every 2 milliseconds. As a result, the Arduino's sampling rate for detecting pulses is 500 Hz. This sample rate is enough for detecting any pulse rate. Fig. 9 shows the PCB design for this proposed project.



(a) Top view



(b) Botton view

Fig. 9. Top and botton view of PCB Design

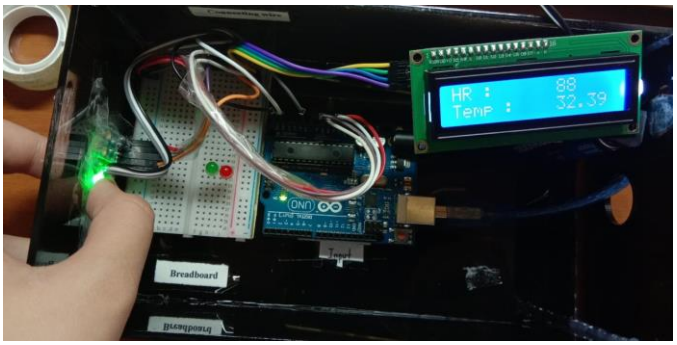
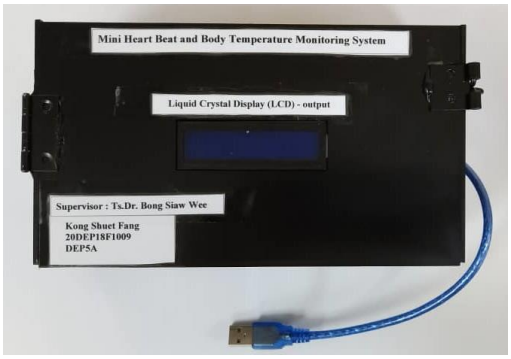
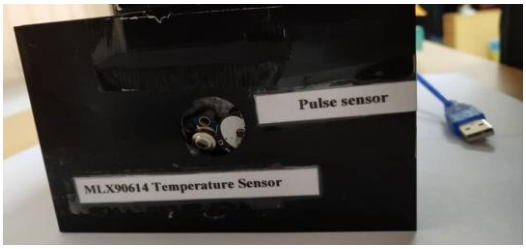


Fig. 10. The functionality of this proposed system

The functioning of this suggested system is depicted in Fig. 10. This device properly measures heart rate and body temperature. The results of the measurements are shown on the LCD panel. Fig. 11 depicts the top view, side view, and front view of this system.



(a) Top view



(b) Side view

(c) Front view

Fig. 11. Top view, side view and front view

The instruction user handbook for this proposed health monitoring system is shown in Fig. 12. There are four steps to use this system as shows bellows:-

Step 1: Connect the USB to the power bank first.
Step 2: Turn on the computer by pressing the "ON" button.
Step 3: Touch the pulse and temperature sensors with your finger.
Step 4: Read data from the LCD Display panel about your pulse and body temperature.

Fig. 12. Instruction user manual

According to the findings of pulse sensor testing in the figures above, each heart in the human body does not experience the same stress because various things affect it, both normal and unhealthy body circumstances, so the graph experiment and bpm results will change continuously. Tables 1 and 2 show the results of pulse sensor and body temperature testing using the LCD user interface.

TABLE I. PULSE SENSOR TEST RESULTS

Testing	BPM calculation results	Result
1	101	Normal
2	98	Normal
3	100	Normal
4	99	Normal

TABLE II. BODY TEMPERATURE TEST RESULTS

Testing	Body Temperature (℃)	Result
1	36.7	Normal
2	36.2	Normal
3	36.6	Normal
4	36.4	Normal

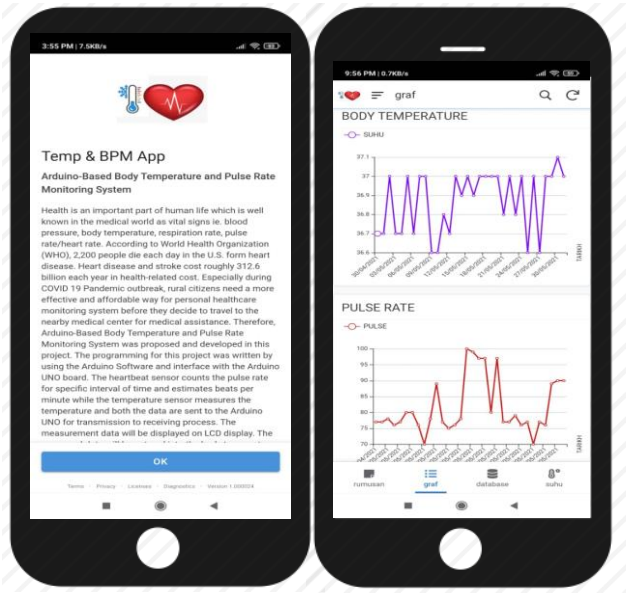


Fig. 13. Measurement graph from Body Temperature and Pulse Rate App

As illustrated in Fig. 13, the Body Temperature and Pulse Rate App is a platform that presents a data graph. As a result, it is simple for the user to identify any major changes in the measurements. This app was created with the AppSheet platform's user-friendly features and can be used on a number of platforms, including phones, tablets, and web browsers.

The user can retrieve their measurement data anytime and anywhere so that early detection of the heart disease or symptoms of COVID-19. Normally, the average normal temperature of the human body is between 36 °C to 37.2 °C. However, 37.3 - 38 °C is low fever, and 38.1- 41 ° C is high fever. Another function for this innovation project is can be used to measure the pulse rate. The normal pulse rate is 60 to 100 beats per minute. People who exceed 100 beats per minute are very dangerous, so they need to refer to the nearest clinic for further medical consultation.

IV. CONCLUSION

An Arduino-Based Body Temperature and Pulse Rate Monitoring System has been developed as a user-friendly product that able to measure the heartbeat and body temperature accurately. This project is capable of monitoring

patients' heart rates and body temperatures at any time and from any location. The system calculates the pulse rate, beats per minute, and body temperature, and displays the physiological data on the LCD. The product price is less than other medical products. Thus, this product is affordable for the low-income group and suitable for residents living in rural areas. The user can detect heart disease or covid-19 earlier, then can make treatment earlier.

REFERENCES

[1] N. H. Wijaya, F. A. Fauzi, E. T. Helmy, P. T. Nguyen, and R. A. Atmoko, "The Design of Heart Rate Detector and Body Temperature Measurement Device Using ATMega16," Journal of Robotics and Control (JRC), vol. 1, no. 2, pp. 40-43, 2020.

[2] W. H. Organization, Cardiovascular diseases, 2021 [Accessed 29 June 2021], Available from: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(DVDs).

[3] D. Vasava and S. Deshpande, "Portable device to monitor vital health parameters," IRJET, 2016.

[4] F. Landreani, A. Martín-Yebra, C. Casellato, C. Frigo, E. Pavan, P. F. Migeotte, and E. G. Caiani, "Beat-to-beat heart rate detection by smartphone's accelerometers: validation with ECG," in 2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC), IEEE, 2016.

[5] U. Gogate, M. Marathe, J. Mourya, and N. Mohan, "Android-based health monitoring system for cardiac patients. International Research Journal of Engineering and Technology," vol. 4, no. 04, pp. 1628-1634, 2017.

[6] S. Chatterjee, P. G. Jose, P. Basak, A. Athar, B. Aravind, R. S. Beed, and R. Biswas, "Microcontroller based automated life savior-Medisûr. in Computational Science and Engineering," Proceedings of the International Conference on Computational Science and Engineering (Beliaghata, Kolkata, India, 4-6 October 2016). 2016. CRC Press.

[7] S. Gupta, A. Talwariya, and P. Singh, "Development of Arduino-Based Compact Heart Pulse and Body Temperature Monitoring Embedded System for Better Performance," in Performance Management of Integrated Systems and its Applications in Software Engineering2020, Springer. pp. 189-197.

[8] B. Karthik, L. D. Parameswari, R. Harshini, and A. Akshaya, "Survey on IoT & Arduino based patient health monitoring system," International Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 3, no. 1, pp. 1414-1417, 2018.

[9] A. Lahamage, S. Dabake, D. Kharat, A. Gharat, and N. Kulkarni, "Remote Healthcare Monitoring System," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 7, no. 4, 2018.

[10] M. S. Mardianto, A. I. Saputra, C. Sukma, and A. Nasrulloh, "Infant Incubator Temperature Controlled and Infant Body Temperature Monitor using Arduino Mega2560 and ADS1232," International Journal of Computer Techniques, vol. 6, no. 6, 2019.

[11] Y. Devis, Y. Irawan, F. Zoromi, and M. R. Amartha, "Monitoring System of Heart Rate, Temperature and Infusion in Patients Based on Microcontroller (Arduino Uno)," in Journal of Physics: Conference Series, IOP Publishing, 2021.

[12] A. A. Galadima, "Arduino as a learning tool," in 2014 11th International Conference on Electronics, Computer and Computation (ICECCO), IEEE, 2014.

[13] S. Monk, "Programming Arduino: getting started with sketches 2016: McGraw-Hill Education.

Thermal Annealing Effect of Strontium Stannate Thin Film Grown by an RF Magnetron Sputtering

1st Yusmar Palapa Wijaya
Faculty of Electrical and Engineering
Univeristy Tun Hussein Onn
Johor, Malaysia
yusmar.wijaya@gmail.com

2nd Fatin Nor Ahmad
Faculty of Electrical and Engineering
Univeristy Tun Hussein Onn
Johor, Malaysia
fatinnorahmad@gmail.com

3rd Khairul Anuar Mohamad
Faculty of Electrical and Engineering
Univeristy Tun Hussein Onn
Johor, Malaysia
khairulam@uthm.edu.my

***Abstract*—Strontium Stannate ($SrSnO_3$ or SSO) thin films were deposited on the ITO substrate using an RF magnetron sputtering method. Thermal annealing has been treated to the samples by 400ºC, 600ºC and 800ºC using a furnace. Structure evolution has been observed by X-Ray diffraction (XRD), atomic force microscopy (AFM) and ultraviolet visible (UV-VIS). A decrease full width at half maximum (FWHM) at (002) peak observed. The amplitude's reduction from 0.472 to 0.354 for as deposited and as annealed sample, respectively, indicating an improvement of crystallite size by 30%. Furthermore, an absorbed wavelength by transmittance exhibited at 380 nm and 600 nm. The energy bandgap of the annealed thin film at 400ºC, 600ºC and 800ºC were 3.33 eV, 3.28 eV and 3.20 eV, respectively.**

Keywords—$SrSnO_3$, Thermal annealing, RF magnetron sputtering, energy bandgap.

I. INTRODUCTION

Strontium Stannate (SrSnO3) has been emerged as a semiconductor material that can potentially be used in wide range of optoelectronics application. The $SrSnO_3$ is also known as SSO perovskite, can be used in photovoltaic cells, light emitting diode (LED), near infra-red (NIR) emission, photo catalysis, photo sensor, power electronics and double layer transistor [1]. As perovskite is an alkaline-earth material, where an SSO crystal structure has the ABX_3 formula in SnO_3 or stannate base beside barium stannate ($BaSnO_3$) and calsium stannate ($CaSnO_3$). Along with attractive new perovskite transparent conductive oxide (TCO), SSO typically has wide bandgap over 3 eV, low resistance as low as 10^{-5} Ω.cm^{-1} and transmittance around 90% [2].

Various synthesis techniques of SSO perovskite in previous research has been reported such as thermal evaporation, mechanical grinding, co-precipitation, modified Pechini method, hydrothermal, solid state reaction and sol-gel process [3]–[5]. Muthukutty et. al. have concluded that co-precipitation was a simple, fast and handy technique to prepare the uniform structure in the range of nanometer to micron[1]. Despite using synthesis mentioned above, dry-processed RF magnetron sputtering has been considered as an effective and relatively simple technique for preparing SSO film. The sputtering equipment has been a high-priced tool due to system's complexity in conducting plasma circumstance for ion plating to substrate. Meanwhile, to improve the morphology of semiconductor material thin film, thermal treatment or annealing widely applied as typical method [6]. Gul et. al. have demonstrated annealing treatment of sol-gel processed SSO thin film[7].

In this paper, SSO thin film were deposited on an ITO glass substrate using an RF magnetron sputtering from an $SrSnO_3$-compound target under argon and oxygen atmosphere with a 1-hour deposition time. Thermal annealing was treated to the samples at different temperatures of 400ºC, 600ºC and 800ºC by a furnace. The purpose of this study is to investigate the structure morphology, electrical and optical behaviour based on different thermal treatment temperature films..

II. METHODOLOGY

A. Thin Film fabrication

In preparation of the sample, indium thin oxide (In_2O_3:Sn) or ITO coated glass in 2 × 2 cm2 was used as a substrate. The sequential cleaning of substrate was done using acetone, methanol and de-ionized water to ITO/glass in an ultrasonic bath for 15 minutes each. The moisture of the substrate were dehumidified using a furnace in 70ºC ambience. In deposition process, the RF magnetron sputtering was set up with 100W of RF power sources, frequency of 1365 Hz, chamber base pressure of 7.7 × 10^{-6} Torr and working pressure 5 mTorr. Pure argon and oxygen reactive gas are composed of 75% and 25%, respectively. The distance between target and substrate was set to 120 mm non perpendicular height. The target was SSO depleted compound (99.9% purity, 3 inch in diameter and 0.125 inch in thickness) by Plasmaterials. One cathode connected to RF power source was assigned and its angle of target to the base is 45º degrees. Programmable furnace Carbolite ELF-1100ºC 14 L was applied for temperature treatment to the thin films. Fig. 1 depicts the process of samples annealing after RF magnetron sputtering.

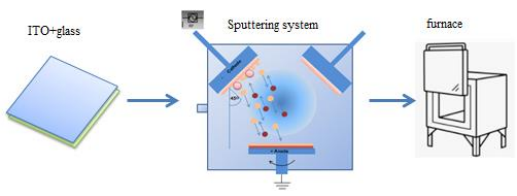


Fig. 1 ITO+glass sputtered samples annealing process

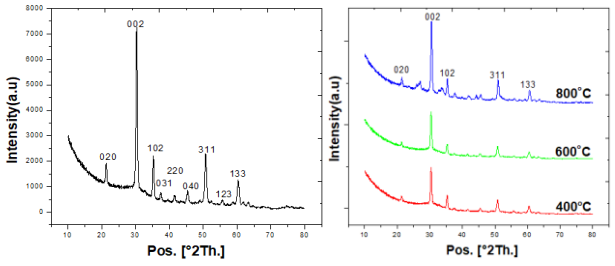
B. *Measurement and Characterization*

There are four samples of SSO thin films, which indicated sample 1 as-deposited, annealed sample 2 as-annealed at 400ºC, sample 3 as-annealed at 600ºC and sample 4 as-annealed at 800ºC. Each annealed samples were placed in a furnace for 1 hour. The structure of as-deposited and annealed SSO films were characterized X-Ray diffraction (XRD) employing Cu-Kα (PANalatycal X'Pert Pro XRD System, α = 0.152 nm) radiation and data analysis using X'Pert HighScore. Hitachi 5100N atomic force microscope (AFM) was used for surface morphology observation. UV-VIS Shimadzu was employed to measure transmittance at wavelength range 200-1000 nm using ITO glass as a base and four point probe Pro 4 was used for electrical properties.

III. RESULT AND DISCUSSION

A. *Morphology*

The XRD spectra of the prepared films are shown in Fig. 2(a) and 1(b). The polycrystalline $SrSnO_3$ films can be observed and the peaks namely (020), (002), (102), (031), (220), (311), (123) and (133) using the ICSD database. After the annealing process, as shown in Fig. 2**Error! Reference source not found.** (b) at 400ºC affects the decrease of peak (031), (220), (040) and (123). After 600ºC annealing, the enhancement of crystallinity begin to fashion, the intensity of amplitude peak (020), (102), (311) and (133) follow to decrease which means peak (002) at 2θ (=30.47) has been dominant. From an atomic perspective, the arrangement of atoms $SrSnO_3$ leads to high crystalline film. However, upper temperature at 800ºC has attracted other peaks to enhance, such as (020), (102) and (311) as compared to thermal study analysis by Mary C.F Alves et. al. which synthesized by calcination of powder. The sputtering technique shows preferable consistency of peaks appearance since as deposited. It has been reported that annealing at 700ºC, the peak (002) starts formed after amorphous phase at lower degrees by calcination technique [8]. Difficulties in forming high intensity of peaks have been reported by annealed sol-gel thin film [7].



(a) (b)

Fig. 2 (a) XRD pattern of as-deposited RF magnetron sputtered SrSnO3 on ITO substrate (b) After annealing at different temperatures

The lattice behaviour of an atom could be determined by crystallinity value parameter which can be affected by intrinsic or extrinsic factors. The crystal size usually obtained by full width at half maximum (FWHM) and Scherrer's equation [9]. Table I**Error! Reference source not found.** summarizes position 2θ, FWHM, d-spacing or interplanar spacing and crystallite size of the peak (002) for every samples only assumed without microstrain. It shows a decrase in FWHM value from as-deposited 0.472 to post-annealing 0.354. Crystallite size increases as-deposited 198.25 Å to annealed ±266.30 Å which indicated improvement of crystallinity has undergone. Improvement of crystallinity could be determined by decreasing FWHM value and increasing grain size [10].

TABLE I. CRYSTAL COMPARISON AMONG DIFFERENT ANNEALING TEMPERATURE AT (002) ORIENTATION

Sample	2θ ±0.001	FWHM ±0.001	d-spacing (Å)	Crystallite Size(Å)
As deposited	30.35	0.472	2.9444	18.21
400 ºC	30.51	0.354	2.9298	24.29
600 ºC	30.47	0.354	2.9332	24.29
800 ºC	30.46	0.354	2.9344	24.29

The AFM examined the surface morphology of the films by observation area of 1 × 1 μm^2. Fig. 3(a) depicts a 3D surface on as-deposited $SrSnO_3$ on an ITO substrate. As deposited sample has roughness (R_a) of 1.714 nm with a grain size 110 nm, and the post-annealed samples are 0.9174 nm, 1.761 nm and 4.833 nm. After annealed at 400ºC, 600ºC and 800ºC, it shows grain size 56.89 nm, 80.13 nm and 30.14 nm, respectively. The higher temperature on annealing process usually affects the increase of grain size, there was exception at 800ºC annealed sample which decreased grain size since over-limit temperature could be semi melting on the material.

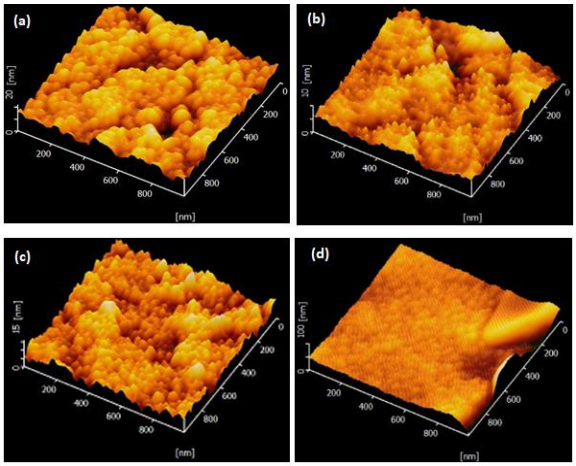


Fig. 3 3D surface topographies of SrSnO3 on different annealed film (a) as-deposited (b) 400oC. (c) 600oC (d) 800oC

Fig. 3 (b)-(d) shows post-annealed samples at different temperatures in 1 hour each using a Carbolite furnace. As depicted in Fig. , the surface of samples has undergone microstructure change. It has probably been high forced microstructure arrangement by too high temperature. Remarkably, at 800ºC, the film became smooth surface slightly different the surface change from as deposited, at 400ºC and 600ºC.

B. *Electrical and Optical Properties*

Fig. depicts annealing effect to resistivity of three samples with different temperatures. Generally, the resistivity of the as-deposited film decreases with annealing temperature. Thus, electrical conductivity is robustly affected by tunneling the charge carriers through the barrier of grain boundary and re-crystallization during annealing, as reported by Ahmed et. al. on (ITO) post-annealing resistivity analysis [11].

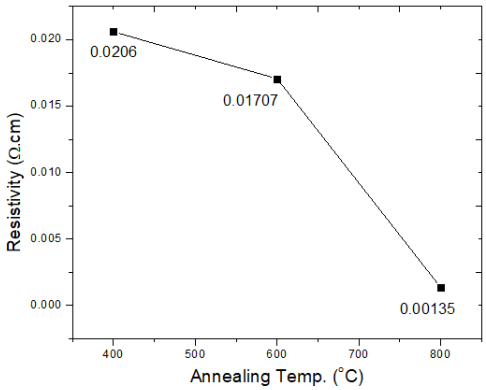


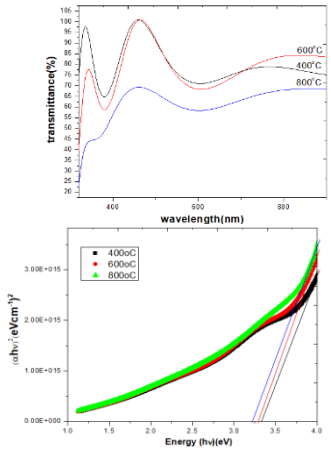
Fig. 4 Sheet resistivity of different annealing temperature

Tabel 2 summarizes the electrical and optical parameters of post-annealed samples. From Fig. 4 and Table II, the value of sheet resistivity decreases due to increasing annealing temperature.

TABLE II. SHEET RESISTANCE AND ABSORBED WAVELENGTH BASED ON TRANSMITTANCE

Post-annealed Sample	Sheet resistvity (Ω.cm)	1st Absorbed Wavelength (nm)	2nd Absorbed Wavelength (nm)
400ºC	0.0206	378	601
600ºC	0.01707	380	606
800ºC	0.00135	365	598

The intensity of light transmitted through heat-treated film was measured using UV-vis. As annealing effect, transmittance percentage of higher temperature applied was decreased and a absorbed the first wavelength shifting from 378 nm to 365 nm and second wavelength shifting from 601 nm to 598 nm as shown in Fig. 5 and Table II.



(a) (b)

Fig. 5 (a)Transmittance of SrSnO3 films by temperature treatment and (b)Energy band gap

Fig. 5(b) depicts plots of $(\alpha hv)^2$ vs. hv for SSO thin film thickness 203 nm after being heated at different temperatures for 1 hour each. The plot $(\alpha hv)^2 = f(hv)$ in the observed area was linear function. By Tauc plot, the energy gap point of film annealing at 400ºC, 600ºC, 800ºC obtain 3.33 eV, 3.28 eV and 3.20 eV, respectively. Thermal treatment has shifted the width of the material energy bandgap since increasing temperature accelerate oxygen to evaporate during the process, more realigning and robust interaction between film and substrate [11], [12]. This wide bandgap (>3 eV) exhibits optical transparency with electrical conductivity [7].

IV. CONCLUSION

This experiment has successfully demonstrated thermal annealing treatment by three different annealed temperatures of 400ºC, 600ºC and 800ºC. By analyzing microstructure evolution using XRD data, FWHM decreased value of (002) peak has been shown by annealing process from 0.472 to 0.354 and its crystallite size has improved by 30% without microstrain assumption. The trend of their sheet resistivity showed decreasing value due to increasing temperature. The wavelength absorbed by post-annealed sample exhibits at 380 nm and 600 nm. The increasing temperature treated on the sample has implied to oxygen evaporation that affected the energy gap between conduction band and valence band approaching. The energy bandgap of annealed thin film of SrSnO3 at 400ºC, 600ºC and 800ºC were 3.33 eV, 3.28 eV and 3.20 eV, respectively. Based on their morphology and optical properties, the optimum preference was 600ºC annealed sample.

ACKNOWLEDGMENT

This work is supported by University Tun Hussein Onn Malaysia (UTHM) under Fundamental Research Grant Scheme (FRGS-K086) fund. The author would like to thank Microelectronic and Nanotechnology-Shamsuddin Research Centre, University Tun Hussein Onn Malaysia for providing a great laboratory and facilities.

REFERENCES

[1] B. Muthukutty, R. Karthik, and S. Chen, "stannate (SrSnO 3) and its potential as an electrode material for the enhanced sensing of anti-inflammatory drug mesalamine in biological samples," 2019.

[2] A. B. A. Rahman, M. S. Sarjadi, A. Alias, and M. A. Ibrahim3, "Fabrication of Stannate Perovskite Structure as Optoelectronics Material : An Overview Fabrication of Stannate Perovskite Optoelectronics Material : An Overview Structure as," *J. Phys. Conf. Ser.*, vol. 1358, 2019.

[3] M. A. Riza, S. Sepeai, N. A. Ludin, M. A. M. Teridi, and M. A. Ibrahim, "Synthesis and Characterizzation Of SrSnO3 Using Different Synthesis Methods," *Malaysian J. Anal. Sci.*, vol. 23, no. 1, pp. 100–108, 2019.

[4] A. Rita, F. Alves, A. De Meireles, and E. Longo, "Chemistry SrSnO 3 perovskite obtained by the modi fi ed Pechini method — Insights about its photocatalytic activity," *J. Photochem. Photobiol. A*, vol. 369, no. August 2018, pp. 181–188, 2019.

[5] S. Moshtaghi, S. Gholamrezaei, and M. Salavati, "New controllable procedure for preparation of SrSnO 3 nanostructures : photo-degradation of azo dyes and photovoltaic measurement," *J. Mater. Sci. Mater. Electron.*, 2015.

[6] H. Wang, C. Liao, Y. Chueh, C. Lai, P. Chou, and S. Ting, "Crystallinity improvement of ZnO thin film by hierarchical thermal annealing," vol. 3, no. 2, pp. 1381–1385, 2013.

[7] E. Gul *et al.*, "The influence of thermal processing on microstructure of sol–gel-derived SrSnO3 thin films," *J. Mater. Sci.*, vol. 52, no. 21, pp. 12624–12634, 2017.

[8] M. C. F. Alves, S. C. Souza, E. C. Paris, S. J. G. Gomes, and L. R. M, "Thermal analysis applied in the crystallization study of SrSnO3," *Therm. Anal Calorim*, vol. 97, pp. 179–183, 2009.

[9] S. M. Londoño-restrepo, R. Jeronimo-cruz, B. M. Millán-malo, E. M. Rivera-muñoz, and M. E. Rodriguez-garcía, "Effect of the Nano Crystal Size on the X-ray Diffraction Patterns of Biogenic Hydroxyapatite from Human , Bovine , and Porcine Bones," no. October 2018, pp. 1–12, 2019.

[10] A. H. O. Alkhayatt, I. A. D. Al-hussainy, and O. A. C. Al-rikaby, "Annealing Effect on the Structural and Optical Properties of Sol- Gel Deposited Nanocrystalline CdO Thin Films," vol. 34, pp. 1–8, 2014.

[11] N. M. Ahmed, F. A. Sabah, H. I. Abdulgafour, A. Alsadig, A.

Sulieman, and M. Alkhoaryef, “The e ff ect of post annealing temperature on grain size of indium-tin-oxide for optical and electrical properties improvement,” *Results Phys.*, vol. 13, no. December 2018, p. 102159, 2019.

[12] M. Abdulhur *et al.*, “Effect of annealing temperature on optical properties of TiO 2 18 NR-T type thin film Effect of annealing temperature on optical properties of TiO 2 18 NR-T type thin film.”

Development of Real-Time Dissolved Oxygen Optimizer for Caged Fish Farm

1st Mohd Sani Said
Director Office
Politeknik Mukah
Sarawak, Malaysia
mdsanisaid@gmail.com

2nd Muhammad Sufyan Safwan Mohamad Basir
Department of Electrical Engineering
Politeknik Mukah
Sarawak, Malaysia
sufyan@pmu.edu.my

3rd Azhar Abdullah
Department of Electrical Engineering
Politeknik Kota Bharu
Kelantan, Malaysia
azharabdullah@pkb.edu.my

***Abstract*—This paper presents the development of real-time dissolved oxygen (DO) optimiser for caged fish farm. The prototype is for maintaining the DO level especially for sea basses, red snappers and groupers. An EZO-DO sensor senses the DO level, which then is sent to an Arduino microcontroller to process the data gathered. The prototype works by triggering the DO water pump if the DO level is less than 3 mg/L, and if the DO level is less than 4.5 mg/L, the paddle wheels will spin. The proposed DO optimiser is powered by a solar panel, hence it is suitable to be implemented in remote areas that are far from lagoon. The proposed DO optimiser is expected to maintain the DO level within 3 – 7 mg/L. Future work will focus on the system performance.**

Keywords—dissolved oxygen optimiser, caged fish farm, Arduino

I. INTRODUCTION

The continuous improvement of water quality variables for dissolved oxygen (DO) is significant to ensure the sustenance of commercial aquaculture industry. Maintaining the optimum water quality for fish growth can reduce the number of fish dead, hence increase the profit for firm farming [1]. Poor management of fish farming may result in the spread of viral diseases in fish and significant aquaculture losses [2]. Reduction in fish supply due to low DO, leading to fish kill, caused estimated losses of about RM 0.1 million in 2003, RM 2 million in 2007, RM 0.5 million in 2009, RM 0.02 million in 2012, and RM 0.02 million in 2014, as reported by Department of Fisheries (DOF) in Kelantan.

Previous studies have shown that increasing the DO contributes to more fish production. There are many techniques to increase the DO while reducing the water temperature. Authors in [3] implemented a microscopic bubble generating system to increase the DO level; for this case in Kusuura Bay, Japan. Three micro-bubble generators had been placed at a water depth between 3, 7, and 14 meters. The reliability of the system was determined based on the maximum concentrations, which were 6.51 mg/L at 3 meters depth, and 6.76 mg/L at 7 meters and 14 meters depth, from the horizontal distribution of DO. When the concept was changed to vertical distribution, the DO level decreased about 1 mg/L by each depth. Authors in [4] implemented an embedded microcontroller (MSP430 series) to monitor the DO, temperature, and pH level of the fish farm aquaculture environment. The novelty of the system was indicated by a wireless sensor using ZigBee. Besides, the reading measured could be monitored using mobile phone, developed using Android SDK software. However, the authors did not implement the proposed system in real application, therefore the reliability of the system cannot be obtained.

In studying the growth production of fish in Mymensingh, Bangladesh, authors in [5] conducted an experiment at two treated tilapia ponds. One pond was treated by aeration through using a blower, while the other pond was non-aerated. Aeration was done three times a day for 3 hours, and the measurements were taken from May to September 2016. The outcome revealed that there was slightly higher DO, with an average of 5 mg/L in the aerated pond, compared to the non-aerated pond. In addition, the temperature in the aerated pond was slightly lower than the non-aerated pond with a difference of 0.24^0C. In Ibadan, Nigeria, the authors in [6] measured DO at a catfish farm during dry and wet seasons. The results indicated 8.01 mg/L of DO with temperature of 24^0C during dry season and DO value of 8.33 mg/L with temperature of 22^0C during the wet season. In relation to Malaysian season, heavy rain may contribute to increase of DO level, but for cases like flood in certain states (Kuala Lumpur, Kedah, Penang and Perlis) that normally happens from December to January [7], the contaminated water may reduce the DO.

Sung et al [8] made improvement on wireless monitoring system as reported in [4] by enhancing the solar system. Improvement in graphical user interface (GUI) was made, by presenting measurement in graph to ease users in understanding the trend, especially DO and temperature. The data measured was stored in Cloud storage, and any user is able to retrieve the data in real-time. The benefit of using solar panel system is that the system can be applied in remote areas, which is beneficial for fish farms that are usually located in rural area. For an early prototype, authors in [9] developed a Wi-Fi wireless transfer automation fish farming using embedded system. Sensors were used to monitor pH, temperature, and quality of water. The authors found that when the pH value was less than 4.5, the fish died. To prevent this, the proposed system automatically triggered the DC motor and aerator to ensure DO kept optimum. However, the reliability of the system was not tested since it was still a prototype.

This study is focused on caged fish farm at an estuary in Tumpat, north-east of Peninsula Malaysia. The farm supplies fish like sea basses, red snappers, and groupers for local and imports. The proposed system is supplied with 24V solar panel, and the energy is stored inside a battery. An EZO-DO sensor is used to measure the DO level of fish farm. The pump that stores the DO water as well as paddles is triggered when the sensor detects DO level less than 3 mg/L. The use of paddles is to increase and keep the DO level to be within 3 to 7 mg/L since it is recommended for caged fish production in Malaysia [10]. The effectiveness of this method has been verified, as in the monitoring results disclosed.

II. METHOD TO INCREASE DISSOLVED OXYGEN

Dissolved oxygen is important for optimum water quality for caged fish [11]. The level of dissolved oxygen in the water can be increased by several ways, either by using aerator, mini bubble injection, or chemical compound [12]. The most basic method is by installing a mechanical water splash to regulate the water frequently and periodically, controlled by a simple timer. Mini bubbles which are rich in oxygen can also be injected into water in a periodically manner.

The aforementioned methods are the most basic ways to increase the dissolved oxygen level in the pond/fish cage. To increase the efficiency, usually there are more than one mechanical water splash installed in the fish cage. However, there are some of the drawbacks with these mechanisms; one of which is the lack of monitoring system used to measure the level of dissolved oxygen over time. This kind of monitoring system is important especially when there is a sudden depletion of dissolved oxygen of pollution, algae bloom, or other reasons [13]. Besides that, motor used to turn the water splash is energy consuming [14]. Therefore, a new renewable energy technology can be integrated with this basic method. Fig. 1 shows the flow diagram of proposed design system to control the DO level in caged fish farm industry.

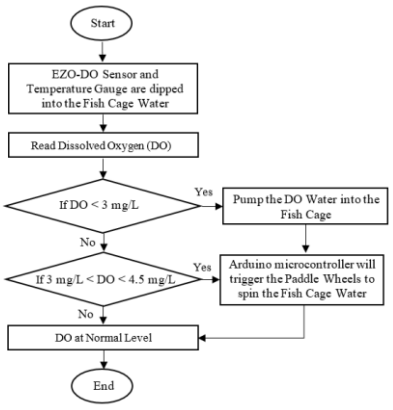


Fig. 1. Flowchart of the proposed design system

A. *Design of Solar Powered Automatic Dissolved Oxygen Optimiser for Caged Fish Farm Industry*

Fig. 2 shows the design of solar powered automatic dissolved oxygen optimiser for caged fish farm industry. The EZO-DO sensor and temperature gauge are dipped into the fish cage water. Supplying DO from aerated water flowing down from two aeration tanks, the DO optimiser can be stationed far into the lagoon since it is solar powered. If the sensor indicates that DO concentration in the cage water is below 3 mg, an extra DO water will be pumped/supplied. Subsequent low DO concentration after the DO water infiltration will trigger the paddle wheels to spin until the DO reaches normal level.

An optimiser has been designed to adapt to the present condition at the estuary as well as to optimise the DO, within level of 3 to 7 mg/L to avoid hypoxic or oversaturation. The present estuary condition has issues such as fluctuating tides and flood which make it difficult to control the DO level, no device is available to continuously monitor and control the DO in caged fish farm, mud flooding the cages due to changing tides and also covering fish gill which suffocate them from breathing, as well as clogging of the barnacles during flood.

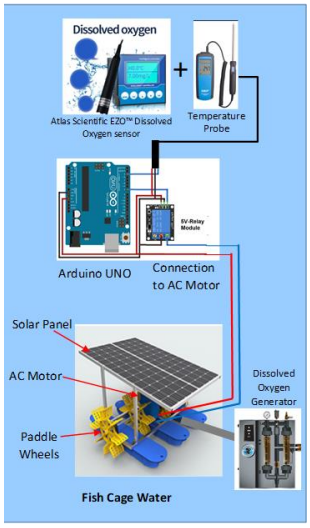


Fig. 2. Design of solar powered automatic dissolved oxygen optimiser for caged fish farm industry

As shown in Fig. 3, the proposed solar powered automatic fish DO optimiser has components like solar panel, AC motor, Arduino UNO, relay, battery, and inverter. The working procedure starts with solar panel cable connected to the 2 terminals of solar panel, then the main switch is turned on, followed by switching on of the voltmeter and inverter switches. Two 12 V batteries are used to store the energy to ensure the continuity of supply. This system is connected to 240 VAC output terminals. Then, the mode switch is turned on to measure both solar panel and battery voltages.

The Atlas scientific EZO-DO sensor, which is a galvanic dissolved oxygen probe, consists of a polytetrafluoroethylene membrane, an anode bathed in an electrolyte, and a cathode [15]. Oxygen molecules defuse through the probes membrane at a constant rate. Once the oxygen molecules have crossed the membrane, they are reduced at the cathode and a small voltage is produced. If no oxygen molecule is present, the probe will determine output as 0 mV. As the oxygen increases, so does the mV output from the probe. Each probe will output different voltage in the presence of oxygen. The only thing that is constant is 0mV = 0 oxygen.

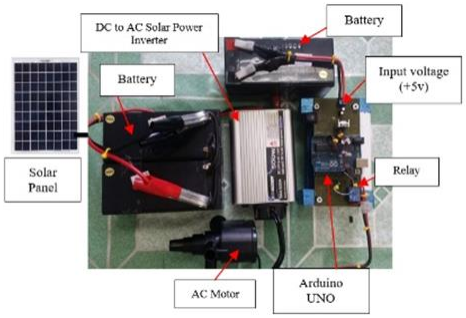


Fig. 3. Electronic apparatus for the proposed system

The sensor is dipped into the water where there are visible dead fish. If the system detects DO level with reading below 3 mg/L, it will trigger the upward water pump to pump water up into the upper aeration tank. To ensure the reliability of measurements, the DO sensor was calibrated by using a 0 mg/L DO solution. The value measured by a sensor should indicate almost similar value. The water in the upper tank will flow down into a lower aeration tank through series of physical obstacles to create rapid-like effects to supply oxygen into the water, which flows into the lower aeration tank. Then, dissolved oxygenated water is pumped into a network of water pipes to distribute the DO water evenly throughout the whole cage complex. If DO level remains below 3 mg/L after 20 minutes of DO water pumping into the cage water, the sensor system will trigger the paddle wheels to spin for better aeration for the fish. The fish cage can be moved from one location to another to get a better position for DO supply in case of need or in case of flood in the area, since it is solar powered that can generate its own power anywhere.

III. Proposed Do Optimiser

Based on literature, the DO indication differs by country, thus different value. As shown in Table I, the optimum DO in Malaysia is between 3 to 7 mg/L. Meanwhile, countries like Philippines and Australia require DO more than 5 mg/L for the fish to survive. If the value of DO is too low, the fish and other aquatic organisms cannot survive. Temperature also plays a vital role in maintaining the DO, by means the lower the temperature, the higher oxygen content it can hold [16]. In accordance, the temperature inside caged fish farm in Malaysia should be kept within 20 to 30 0C.

TABLE I. Optimised Parameters for Cages Fish Production in Different Countries

Water Quality Parameters	Temperature (^{0}C)	Dissolved Oxygen (mg/L)
India [17]	21 - 33	4.0 – 10.0
Australia [18]	-	> 5.0
The Philippines [18]	-	5.0
Malaysia [10]	Normal + 2	3.0 – 7.0

The proposed DO optimiser as shown in Fig. 4 is significant for the fishing industry since it can help to regulate the oxygen quality in estuary area especially at caged fish farm. The proposed system can be stationed far into the lagoon since it is fully solar powered. It is not only capable of solving the low DO problem in the farm fish cage, but also improve water clarity. Consequently, the use of DO optimiser may multiply the income of caged fish farmers. In addition to intelligent and efficient DO control, the proposed system can reduce the mud from covering fish's gill. Besides its ability to reduce barnacles from clogging the cages, lowering production and maintenance cost, the DO optimiser may give a significant impact on the social and economic development of the fishing industry, in addition to ensuring environmental sustainability in the estuary areas.

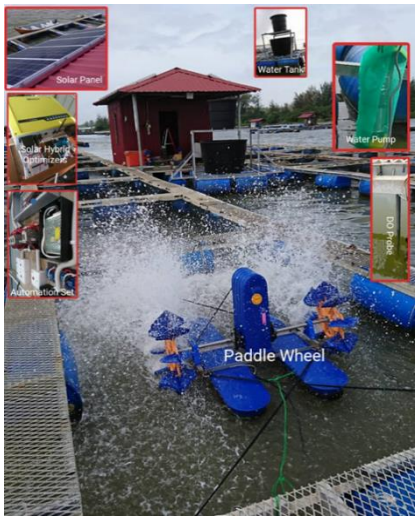


Fig. 4. Dissolved oxygen optimiser prototype in full operation due to low DO concentration

Technically, with the proposed DO optimiser, the DO level could be maintained more than 3 mg/L. The proposed system operates automatically based on threshold value (3 mg/L), by means the paddle wheels will start spin and the DO water is pumped to the farm fish cage if the DO level is less than that. Last but not least, there are some improvements that can be added to the prototype. The author discovered that the aeration tanks were not up to the original design to produce water filled with high concentration of dissolved oxygen due to its high production cost. Therefore, the upper tank can be raised to optimum to maximize dissolved oxygen level into the poured water. Also, without a proper shield, this might lead to a lot of DO water wastage. Secondly, the lower tank had to be filled with needle-like pokers pointing upwards to break the poured water to create rapid-like effect in the effort to maximize oxygen getting into the water.

IV. Conclusion

The proposed DO optimiser functions to stabilize the dissolved oxygen (DO) level in the water of caged fish farm. An EZO-DO sensor is placed deep inside the water and the value is recorded for 24 hours daily. The proposed DO optimiser developed will automatically trigger the pump from the stored DO water as well as rotate the paddles if the DO value drops to less than 3 mg/L. From the outcomes obtained, the proposed prototype is indeed capable to maintain the DO value to more than 3 mg/L. Accordingly, the DO will rise to the optimised value (within 3 to 7 mg/L), hence the percentage of dead fish can be reduced.

ACKNOWLEDGMENT

The authors acknowledge the support from Politeknik Mukah Sarawak, Politeknik Kota Bharu, and Ministry of Higher Education Malaysia for providing the financial support under TVET Applied Research Grant Scheme/T-ARGS (No. 2008/19).

REFERENCES

[1] Bergheim, A., Gausen, M., Næss, A., Hølland, P. M., Krogedal, P., & Crampton, V. (2006). A newly developed oxygen injection system for cage farms. *Aquacultural engineering*, *34*(1), 40-46.

[2] Abdel-Tawwab, M., Monier, M. N., Hoseinifar, S. H., & Faggio, C. (2019). Fish response to hypoxia stress: growth, physiological, and immunological biomarkers. *Fish physiology and biochemistry*, *45*(3), 997-1013.

[3] Endo, A., Srithongouthai, S., Nashiki, H., Teshiba, I., Iwasaki, T., Hama, D., & Tsutsumi, H. (2008). DO-increasing effects of a microscopic bubble generating system in a fish farm. *Marine pollution bulletin*, *57*(1-5), 78-85.

[4] Chen, J. H., Sung, W. T., & Lin, G. Y. (2015, October). Automated monitoring system for the fish farm aquaculture environment. In *2015 IEEE International Conference on Systems, Man, and Cybernetics* (pp. 1161-1166). IEEE.

[5] Sultana, T., Haque, M. M., Salam, M. A., & Alam, M. M. (2017). Effect of aeration on growth and production of fish in intensive aquaculture system in earthen ponds. *Journal of the Bangladesh Agricultural University*, *15*(1), 113-122.

[6] Olaifa, F. E., Olaifa, A. K., Adelaja, A. A., & Owolabi, A. G. (2004). Heavy metal contamination of Clarias gariepinus from a lake and fish farm in Ibadan, Nigeria. *African Journal of Biomedical Research*, *7*(3).

[7] Chan, N. W. (1997). Increasing flood risk in Malaysia: causes and solutions. *Disaster Prevention and Management: An International Journal*.

[8] Sung, W. T., Chen, J. H., & Wang, H. C. (2014, April). Remote fish aquaculture monitoring system based on wireless transmission technology. In *2014 International Conference on Information Science, Electronics and Electrical Engineering* (Vol. 1, pp. 540-544). IEEE.

[9] Kiruthika, S. U., Kanaga, S. R., & Jaichandran, R. (2017). IoT based automation of fish farming. *Journal of Adv Research in Dynamical Control Systems*, *9*(1).

[10] Harun, Z., Reda, E., & Hashim, H. (2018, March). Real time fish pond monitoring and automation using Arduino. In *IOP Conference Series: Materials Science and Engineering* (Vol. 340, No. 1, p. 012014). IOP Publishing.

[11] Degefu, F., Mengistu, S., & Schagerl, M. (2011). Influence of fish cage farming on water quality and plankton in fish ponds: A case study in the Rift Valley and North Shoa reservoirs, Ethiopia. *Aquaculture*, *316*(1-4), 129-135.

[12] Roberts, R. J. (2012). *Fish pathology*. John Wiley & Sons.

[13] Cox, B. A. (2003). A review of dissolved oxygen modelling techniques for lowland rivers. *Science of the Total Environment*, *314*, 303-334.

[14] Ghosh, M., Ghosh, A., & Roy, A. (2020). Renewable and sustainable materials in automotive industry.

[15] EZO-DO™. 2019. Embedded Dissolved Oxygen Circuit. *Atlas Scientific Environmental Robotics* v5.3 Revised 8/13/19: Jordan Press.

[16] Ismail, R., Shafinah, K., & Latif, K. (2020, May). A Proposed Model of Fishpond Water Quality Measurement and Monitoring System based on Internet of Things (IoT). In *IOP Conference Series: Earth and Environmental Science* (Vol. 494, No. 1, p. 012016). IOP Publishing.

[17] Raju, K. R. S. R., & Varma, G. H. K. (2017, January). Knowledge based real time monitoring system for aquaculture using IoT. In *2017 IEEE 7th international advance computing conference (IACC)* (pp. 318-321). IEEE.

[18] PHILMINAQ, "Water Quality Criteria and Standards for Freshwater and Marine Aquaculture Abbreviations and Acronyms," *Mitigating impact from Aquac. Philipp.*, p. 34, 2014.

Synthesis of Materials Al(OH)^3 Based on Sand in Batam Island as a Geopolymer for Brick Materials

1st Ulfa Nabilla
Mechanical Engineering
Politeknik Caltex Riau
Pekanbaru, Indonesia
Ulfa@alumni.pcr.ac.id

2nd Dianita Wardani
Piping Engineering
Politeknik Perkapalan
Negeri Surabaya
Surabaya, Indonesia
Dianitawardani@ppns.ac.id

3rd Amnur Akhyan
Mechanical Engineering
Politeknik Caltex Riau
Pekanbaru, Indonesia
Akhyan@pcr.ac.id

4th Roni Novison
Mechanical Engineering
Politeknik Caltex Riau
Pekanbaru, Indonesia
Roni@pcr.ac.id

Abstract*—Bauxite sand is beach sand on the island of Batam which contains a lot of kaolin, this is the result of the FTIR test that has been tested. Kaolin has a good composition of alumina (Al_2O_3) and silica (SiO_2) compounds to produce geopolymer materials for quality building materials. The geopolymer material synthesis process uses materials, namely Batam island bauxite sand, fly ash from burning sugarcane waste, and redmud and then mixed with wet materials such as NaOH and Sodium silica which are varied as much as 7% and 5%. After the geopolymer material, XRD testing was carried out to determine the angle 2 on the geopolymer and continued with the process of making bricks with a mixture of geopolymer and clay materials and then molding them with brick molds manually. The bricks were given curing or heating treatment for 24 hours with temperature variations of 80°C and 90°C then aging or left in the open for 14 days, then the bricks will be tested by testing hardness, compressive strength and impact testing at the Mechanical Engineering workshop Caltex Riau Polytechnic. The results of the tests that have been carried out are found that the compressive strength of the geopolymer brick composition 1 at 90°C is the brick that has the highest compressive strength value of 12.84 kg/cm². Impact testing is to see the toughness value of a material increases the impact value, the object is getting more, in this test geopolymer brick composition 1 + 90°C and geopolymer brick Composition 2 + 80°C are materials that are not or easily brittle with an impact price of 0.0063 J/mm² and 0.0071 J/mm². In the Brinell hardness test, the highest value was found by geopolymer bricks with a composition of 1 + 90°C, after three experiments the average HB value was 83.21. is the highest HB value of other composition geopolymer bricks and ordinary bricks.

Keywords—Geopolymer, Bauxite, Sodium Silicate, Redmud, Curing, Aging

I. INTRODUCTION

Sand is a material used as a building material to glue cement, besides that sand is also the main ingredient for making bricks. Batam City, which is part of the Riau Archipelago Province, has a land area of 715 km, while the total area reaches 1,575 km. The material produced on the island of Batam contains bauxite (kaolin) which is good for building materials.

After conducting the FTIR test at one of the universities in Java, the compositions contained in the Batam Island sand are Polyster with kaolin filler, CH_3 COHCICOCH_3 (3-chloro-2,4-pentanedione), C_5 H_8 (pentane), C_2 H_4 O_3 (2 - Hydroxyethanoic acis), C_6 H_14 (hexane), C_3 H_60 (propane), C_6 H_12 (hexene), C_7 H_12 (heptuna), C_4 H_10 O_2 (1,2-butabediol). As according to previous research, Kaolin is found in the form of sediments that are still mixed with quartz sand. Quartz sand is found along the coast. This sand contains 60% alumina clay in the form of bauxite deposits resulting from weathering of granite. The nature of bauxite can not absorb water and can not expand when in contact with water. Bauxite has a residue called bauxite residue which is very useful for the manufacture of quality building materials because it has a high compressive strength. (Muchtar Aziz and Azhari.2014)

The manufacture of bauxite-based geopolymer materials in this activity is intended for building materials, namely bricks, because according to previous research the kaolin content which has alumina and alkali is suitable for making geopolymer building materials. Testing the characteristics of the geopolymer specimens produced includes FTIR, XRD, compressive strength, tensile, impact and hardness.

II. RESEARCH METHODS

A. Sample Test

Bauxite sand from the island of Batam was tested to determine the composition of any compounds contained in the sand itself. Sand testing process, namely the first stage of sample preparation (bauxite sand) then an FTIR test is carried out to determine the composition of the bauxite sand, then complete.

B. Geopolymer material manufacture

The manufacture of geopolymer material is carried out by starting with washing of bauxite sand, then dry mixing, then wet mixing, then XRD test is carried out to get 2 theta angle of the geopolymer itself.

C. Brick Making

The manufacture of bricks begins with molding the bricks, namely geopolymer material mixed with clay then curing at a temperature of 80°C and 90°C and aging for 14 days, then testing is carried out namely hardness, compressive, tensile, and impact tests.

D. *Geopolymer Design*

The process of making geopolymer materials with several compositions, namely bauxite, bauxite residue, fly ash, NaOH, and sodium silicate. The amount of sodium silicate was varied by 5% and 7% to see the difference in compressive strength that would occur in the bricks. The weight of the geopolymer material itself is 700 grams.

TABLE I. SHOWS THE COMPOSITION OF GEOPOLYMER MATERIALS.

Percobaan	Bauksit	Residu bauksit	Abu terbang	NaOH	Sodium silikat
Komposisi 1	35%	15%	15%	28%	7%
Komposisi 2	35%	15%	17%	28%	5%

E. *NaOH*

Sodium hydroxide functions to react with the aluminum (Al) and silicate (Si) elements contained in fly ash, so as to produce strong polymer bonds. The NaOH content in the geopolymer is 28% of the 700 grams of the geopolymer. So, 200 grams of NaOH will be used in the manufacture of geopolymer materials.

F. *Brick Composition Design*

The process of making bricks uses a composition of 60% geopolymer and 40% clay. Temperature variations that will be used in the curing process are 80°C and 90°C to see the compressive strength that will occur in the bricks.

TABLE II. SHOWS 4 SAMPLES THAT WILL BE MADE INTO BRICKS WITH VARIATIONS IN TEMPERATURE AND SODIUM SILICATE OF 7% AND 5%.

Perco baan	Bahan Geopolimer		Tanah liat	Tempera tur
	Komposisi 1	*Komposisi 2*		
1	60 %	-	40 %	90°C
2	-	60 %	40 %	90°C
3	60 %	-	40 %	80°C
4	-	60 %	40 %	80°C

G. *Brick Test Design*

The bricks that have been molded are then subjected to a curing process for 24 hours and aging for 14 days, then the bricks are tested for compression, tensile, impact and hardness to find out what the resistance value of the bricks is.

III. TESTING AND ANALYSIS

A. *FTIR Testing on Bauxite*

Batam island bauxite is the main material in the manufacture of geopolymer materials. Before the geopolymer synthesis process is carried out, we must know the composition of what compounds are present in the bauxite sand of Batam Island. To find out the composition of the compounds contained in the bauxite sand on the island of Batam, the FTIR test was carried out. This FTIR test was carried out at the material testing laboratory at ITS. FTIR (Fourier Transform Infra-Red) characterization was carried out to determine the type of functional group bonding in a compound. FTIR is a type of spectroscopy based on the vibration of a molecule. The working principle of FTIR uses the Fourier transform method to measure the absorption of the infrared spectrum emitted from the source to the test material at various wave numbers.The results of the FTIR test on batam island bauxite are shown in Figure 4.1 below.

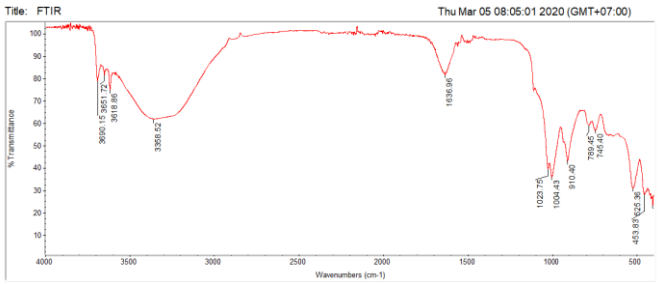


Fig. 1. FTIR Graphic

Figure 1 is an FTIR graphic image of the Batam Island bauxite sand test. In this FTIR test, it is carried out to determine the composition of the kaolin content needed in the manufacture of geopolymers. This graph has two axes, the x-axis and the y-axis where the x-axis is the wave number and the y-axis is the intensity angle. Based on the graph above, the graph has several peaks. In this case, the geopolymer requires kaolin compounds which contain alumina and silica elements, both of which are the basic elements of the manufacture of geopolymer materials. The kaolin is formed at a wave number of 1023.75 cm-1 with a peak at an intensity angle of 37.39°. Thus, Batam island's bauxite sand is declared good for use in the manufacture of geopolymers.

B. *Manufacturing of Geopolymer Materials*

1) *Geopolymer Synthesis*

The manufacture of geopolymer material is carried out by starting with the washing of bauxite sand, then dry mixing is carried out, this dry mixing consists of several materials, namely bauxite, the bauxite used for this synthesis process is bauxite originating from the island of Batam. Bauxite contains alumina and silica which are good for the manufacture of building materials such as bricks. The next ingredient is fly ash which comes from the combustion of sugarcane waste, then another material is used, namely bauxite residue (redmud), then wet mixing, namely mixing with a wet material consisting of sodium silica which is varied by 7% and 5%, then there is NaOH which functions to react the silica and alumina elements in fly ash, after doing dry and wet mixing, mix the ingredients manually until they are well mixed, after the geopolymer material is done, do an XRD test to see the peaks of the 2 theta angle of the geopolymer material.

The geopolymer synthesis process uses several materials, namely bauxite, the bauxite used for this synthesis process is bauxite originating from the island of Batam. Bauxite contains alumina and silica which are good for the manufacture of building materials such as bricks. The next ingredient is fly ash which comes from the burning of sugarcane waste, then another material is used, namely bauxite residue (redmud), then there is sodium silica which is varied by 7% and 5%, then there is NaOH which functions to react silica and alumina elements. on fly ash.

The process of making geopolymer material is done by manual stirring. In the manufacture of this geopolymer there are two compositions with a weight of 700 grams of the geopolymer material. The first composition consists of 35%

bauxite, 15% bauxite residue, 15% fly ash, 7% sodium silica and 28% NaOH. The difference between the first and second compositions is 17% fly ash and 5% sodium silica. In the process of making geopolymer materials, water is added so that the geopolymer has a soft texture and is easily mixed with clay.

Fig. 2. Clay

C. *XRD Test Result*

After making the geopolymer material, we carry out XRD testing which serves to see if the geopolymer material that has been made has been formed at an angle of 2θ in the geopolymer range. According to the journal geopolymer is usually formed at an angle of 11°-40° 2θ.

XRD testing is carried out to determine changes and elemental content that occurs due to the treatment given to the sample, with the information that the y-axis is the intensity produced and the x-axis is the 2θ angle formed when X-Ray rays are reflected on the test sample. XRD testing is carried out in the materials laboratory at ITS. Figure 4.9 is a graph of XRD geopolymer, by heating the geopolymer material at 90°C for 24 hours, the XRD geopolymer sample will show the 2 sudut angle it will produce. Seen in Figure 4.9 on the XRD graph, in the graph there are several peaks produced, the highest peak is located at an angle of 2θ with a value of 26.6 with an intensity of 100%. According to the journal geopolymer is usually formed at an angle of 11°-40° 2θ. In this test, the peak is formed at an angle of 26.6° 2θ, which means that this angle is included in the Geopolymer range. high crystallinity and yields a suitable fraction angle at a good geopolymer angle.

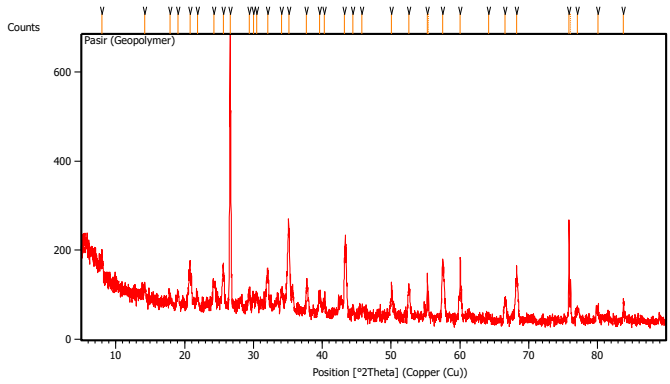


Fig. 3. XRD Geopolymer Test Graph

With the value of the angle formed already included in the geopolymer range, the geopolymer material that has been tested by XRD, which will be used for the manufacture of brick base materials already meets the standards of geopolymer material.

D. *Brick Making Process*

Making bricks is carried out in several stages, the first stage is mixing 60% geopolymer material and 40% clay with a weight of 1,200 grams. The second stage is the geopolymer and clay materials are stirred until well mixed and then inserted into the brick mold

E. *Geopolymer Bricks Result*

Here is the result of making geopolymer bricks. Figure 4.6 shows bricks that have undergone treatment. Geopolymer bricks are different from bricks usually in terms of color and are also different because the composition of these bricks does not only contain clay like ordinary bricks that we often encounter.

Fig. 4. Geopolymer Brick

After the geopolymer brick has been aged for 14 days, mark each brick with a different composition and temperature for Brinell hardness testing, impact testing, and compression testing. Figures 4.7, 4.8, 4.9, are examples of brick samples that were tested:

Fig.5. Brinell . test sample

Fig. 6. Sampel uji impak

Fig. 7. sampel uji kuat tekan

F. Brinell Hardness Test Result

The Brinell test is carried out to determine the hardness value of an object in the form of resistance to a steel ball (indenter) that presses on the surface of the test object. In the Brinell test for ordinary bricks and geopolymer for comparison of hardness in bricks, it is proven by the HB value for each of these bricks. The HB value is the hardness value in the Brinell test. In this test, a 5 mm indenter is used, with a load of 62.5 kg, and during loading, the load is held for 30 seconds.

From the results of the Brinell hardness test, it can be analyzed that each point of data collection on the surface of the sample experiences a difference in the HBW value. This difference in value may be due to uneven mixing of geopolymer and clay materials, but the difference in values is not too significant. It can be analyzed on the data that the smaller the diameter produced in the Brinell test, the greater the HBW value, and vice versa if the diameter produced in the Brinell test is large, the HBW value will be smaller. The HBW value is the hardness value in the Brinell test. If the resulting HBW value is large, the harder the material will be. From the table above, it can be seen that the hardest geopolymer brick is a brick made of composition 1 with a composition of sodium silica 7% with a curing temperature of 90°C with the highest HBW value of 84.60 because it contains sodium silica with a higher presence of 7% added with the curing temperature is 90°C, while the material that has a soft surface is geopolymer brick which is made of composition 2 containing 5% sodium silica and curing temperature 80°C temperature is due to sodium silica with a low percentage of 5% and a lower curing temperature of 80° C.

For comparisons between ordinary bricks or bricks made of clay with geopolymer bricks, geopolymer bricks with a composition of 7% sodium silica content and a curing temperature of 90°C have a higher HBW value than ordinary bricks, by looking at the data in the table above it can be concluded that geopolymer brick with composition 1 with a sodium silica content of 7% and a temperature of 90°C is twice as hard as ordinary brick because it contains sodium silica, NaOH and fly ash which makes the geopolymer material harder.

G. Analysis and Impact Testing

Impact testing is a material or object resistance test against shock loads. In this impact test, the amount of energy absorbed by the material for fracture is a measure of the impact resistance or toughness of the material. A material is said to be tough if it has the ability to absorb large shock loads without cracking or deforming easily.

From the data generated by the impact test, it can be analyzed that the material does not have a significant difference, it can be seen from the data generated on each test object that is not too far away. The geopolymer brick that has the greatest resistance to shock loads is the brick containing composition 2 with 5% silica content and a curing temperature of 80°C, where the energy value obtained is 115 Joules with an impact price of 0.0071 J/mm², if we compare it with Ordinary bricks have a clay composition of 83 Joules of energy with an impact price of 0.005 J/mm², then geopolymer bricks are more resilient than ordinary bricks, because they contain sodium silica, NaOH and fly ash which help strengthen the material properties of the bricks.

From the results of the calculation of the impact price, we can see that the higher the impact price, the tougher the material is, therefore we can see that the geopolymer brick composition 1 has a sodium silica content of 7% and a curing temperature of 90°C and the geopolymer brick composition 2 contains sodium. silica 5% and curing temperature 80°C is a tough material or not easily brittle with impact prices of 0.0063 J/mm² and 0.0071 J/mm², while geopolymer bricks Composition 1 contains 7% sodium silica and curing temperature 80°C, geopolymer bricks Composition 2 contains 5% sodium silica and a curing temperature of 90°C, and ordinary bricks with clay composition have properties that are more brittle with impact values of 0.0054 J/mm², 0.0057 J/mm² and 0.005 J/mm².

TABLE III. ANALYSIS AND IMPACT TESTING

Number	Sampel	Effort (joule)	Picture
1	Composition 1 Temperature 90°C	104	
2	Composition 1 Temperature 80°C	88	

3	Composition 2 Temperature 90°C	93	
4	Composition 2 Temperature 80°C	115	
5	ordinary brick	82	

H. Analysis and Compressive Test

The compressive strength test is a test to determine the amount of compressive load on an object, from knowing the compressive load generated on the test object, we can also know the value of the compressive strength. The value of the compressive strength of the brick is needed to determine the maximum strength of an object to withstand pressure or an object until it cracks and breaks. The quality of bricks is usually indicated by the size of the compressive strength

TABLE IV. THE QUALITY OF BRICKS

Number	Item	Age	Weight (kg)	cross section (mm²)	compressive load (kN)	conversion (kg)	compressive strength (kg/cm²)
1	K190	14	1.1	162	20.6	2080.6	12.84
2	K180	14	1.1	162	19.8	1999.8	12.34
3	K290	14	1.1	162	15.4	1555.4	9.6
4	K280	14	1.1	162	16.0	1616	9.97
5	ordinary brick	14	1.1	162	14.8	1494.8	9.22

The following is a graph that is formed on the compressive strength test of geopolymer brick and ordinary brick:

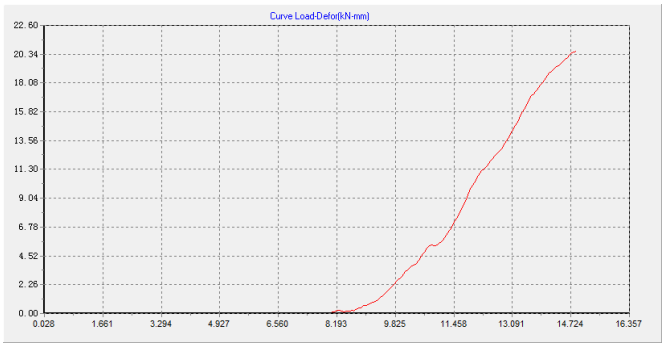


Fig. 8. Graphical image of geopolymer brick composition 1 with a curing temperature of 90°C

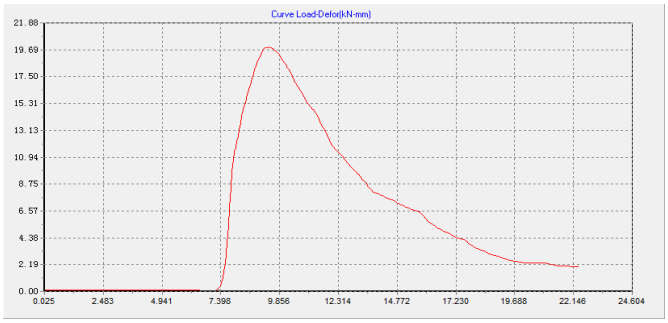


Fig. 9. Graphical image of geopolymer brick composition 1 with a curing temperature of 80°C

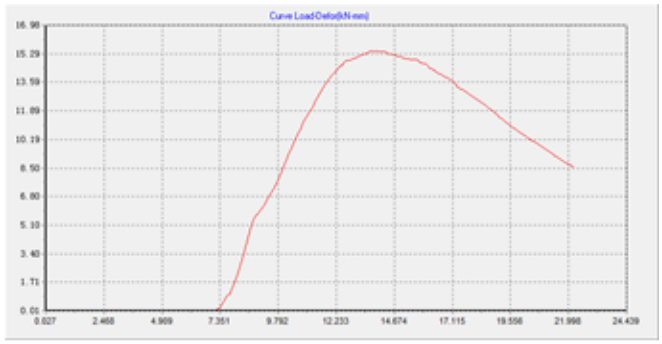


Fig. 10. Graphical image of geopolymer brick composition 2 with a curing temperature of 90°C

Fig. 11. Graphical image of geopolymer brick composition 2 with a curing temperature of 80°C

Fig. 12. Drawings of ordinary brick graphics

From the data generated by the compressive strength test, it can be analyzed that the temperature curing material greatly affects the value of the compressive load on the test object as seen from the data generated on each test object and the graph formed on each test object. The geopolymer brick that has the highest compressive strength value is composition 1 brick with a composition of sodium silica 7% with a curing temperature of 90°C, the value is 12.84 kg/cm² which means that it exceeds the compressive strength value of ordinary brick which is 9.22 kg/cm², while we know that the temperature The curing of geopolymer brick is only 90°C in just 24 hours of heating, while the ordinary brick is burned in a furnace with a heating of more than 1000°C for several days, so from the test data the compressive strength of geopolymer brick is stronger than ordinary brick.

From the calculation results of the compressive strength value, we can know that the greater the compressive strength value, the stronger the material is, therefore we can know that geopolymer brick composition 1 contains 7% sodium silica and a curing temperature of 90°C with a compressive strength value of 12.84 kg. /cm² and geopolymer brick Composition 1 contains 7% sodium silica and a curing temperature of 80°C with a compressive strength value of 12.34 kg/cm², is a material that has a compressive strength value greater than that of geopolymer brick composition 2 containing 5% sodium silica and a curing temperature of 90 °C with a compressive strength value of 9.6 kg/cm², composition 2 contains 5% sodium silica and a curing temperature of 80°C with a compressive strength value of 9.97kg/cm², as well as ordinary brick with a compressive strength value of 9.22 kg/cm². Thus the composition of sodium silica, NaOH, fly ash and curing temperature greatly affect the compressive strength of a brick.

IV. CONCLUSION

From the results of the activities carried out, starting from the preparation of tools and materials, making geopolymer materials, mixing geopolymer materials and clay, brick molding, curing treatment for 24 hours and aging for 14 days and carrying out tests, until the test data is generated, conclusions can be drawn as follows. following:

- Bauxite sand on the island of Batam has good potential in the manufacture of geopolymer materials as proven in the FTIR test that contains kaolin which is an element of the manufacture of geopolymers.
- Geopolymer made from bauxite sand on the island of Batam in the XRD test got the highest peak at an angle of 2θ with a value of 26.6 with an intensity of 100%. According to the journal, geopolymers are usually formed at an angle of 11°-40° 2θ. In this test, the peak is formed at an angle of 26.6° 2θ, which means that this angle is included in the Geopolymer range.
- In the Brinell test, geopolymer bricks with composition 1 contain 7% sodium silica and a curing temperature of 90°C proved to have the highest Brinell hardness value (HB) compared to other bricks with an average HB value of 83.21. The HB value is the hardness value in the test. Brinell, which means the higher the HB value in the Brinell test on a test object, the harder the test object.
- Impact testing is carried out to see the toughness properties of the material, on geopolymer bricks that have the greatest value of resistance to shock loads are bricks containing a composition of 2, 5% sodium silica content and a curing temperature of 80°C, where the energy value obtained is of 115 joules with an impact price of 000.71 J/mm², if we compare it with ordinary bricks which have an energy clay composition of 83 joules with an impact price of 0.0005 J/mm², the geopolymer brick is tougher than ordinary brick.
- Based on the compressive strength test, the brick that has the highest compressive strength value is geopolymer brick with composition 1 containing 7% sodium silica and a curing temperature of 90°C with a compressive strength value of 12.84 kg/cm², here it is proven that the higher the curing temperature and the addition of sodium silica, affect the value of the compressive strength of the test object.
- From the tests that have been carried out, the Brinell hardness test, impact test, and compressive strength test, geopolymer bricks have greater hardness, toughness and strength values than ordinary bricks because they contain bauxite, sodium silica, NaOH and fly ash which make geopolymer material has a higher hardness value.

REFERENCES

[1] Aziz Muchtar and Azhari. Manufacture of Bauxite Residue Based Geopolymer Materials For Building Materials.Journal of Mineral and Coal Technology. 2014. 10(1)

[2] Aziz Muchtar. Characterization of Dregs Minerals for Making Geopolymer Building Materials. Journal of Mineral and Coal Technology. 2012.8(2)

[3] Aziz Muchtar. Concept of Utilization and Processing of Mineral Dregs, Case Study of West Kalimantan Bauxite Dregs Processing Plan. Journal of Mineral and Coal Technology. 2012. 8(1)

[4] Dieta Wulansari, et al. Characterization of Mineral Content in Bauxite Semi-Quantitative XRD Method in the Tayan Mine Area, West Kalimantan. 2016

[5] Devi Silfia Istiqomah, et al. 2016. Synthesis of Al_2O_3 Nanoparticles From Bauxite Ores For Applications In Radiator Models.3(2)

[6] Buddha Maryanti, et al. Effect of Alkalization of Coconut Fiber-Polyester Composite on Tensile Strength. 2011. 2(2)

[7] Yopi Handoyo. Design of Impact Test Equipment Charpy Method Capacity of 100 Joules. 2013. 1(2)

[8] Gusti Rusydi Furqon S, et al. Analysis of Hardness Tests on ST 60 Steel Shafts With Different Cooling Media. 2016. 1(2)

[9] Nurul Latifah Khoiriyah and Putri Maisytoh. Characteristics of Geopolymer Mortar With Oven Treatment with Various Curing Time Variations. 2016. 15(1)

[10] Safe and Amun Amri. Making geopolymer paving blocks from coal fly ash waste (Fly Ash) in Pematang Kapau village, Tenayan Raya Pekanbaru. 2019.

Prediction of the Second Transition Point of Tool Wear Phase Using Vibratory Signal Analysis (Z-ROT)

1st Nur Adilla Kasim
Department of Mechanical Engineering
Politeknik Mukah
Sarawak, Malaysia
nur.adilla@pmu.edu.my

2nd Mohd Ghafran Mohamed
Department of Mechanical Engineering
Politeknik Mukah
Sarawak, Malaysia
mohd.ghafran@pmu.edu.my

3rd Mohd Zaki Bin Nuawi
Department of Mechanical and Manufacturing Engineering
Universiti Kebangsaan Malaysia
Selangor, Malaysia
mzm@ukm.edu.my

***Abstract*—Early intervention to change worn cutting tool before its failure could avoid unexpected machine downtime. A mathematical based predictive model is employed to estimate early tool failure using vibratory signal. The statistical-based signal analysis technique as wear tracking analysis is applied in the predictive model to outline the data pattern concerning wear and number of cutting. The signal analysis based on the changes in the vibration signatures that captured from accelerometer during the milling operation throughout the tool life. A significant correlation between the tool flank wear and the statistical index has achieved. The tool life as a function of the acceleration amplitude of assimilated vibrations. Selected curve fitting equations are considered to decide the transition point between the steady state and failure region. The result shows a significant expectation of determining the second transition point with estimate value of 0.235mm below the rapid wear (<0.25mm). The accuracy, reliability and robustness of the predicted transition point were then parallel against another sensing elements where it predicts almost the same transition point. The de-termination of the second transition point will assist the preparation to anticipate the tool to be bro-ken. The results reflected that the model gives reasonable estimation of tool life and the transition points at which changes of the region transpire.**

Keywords—Cutting Tool Wear, Signal Analysis, Z-rot, Piezofilm-Based Sensor, Vibration

I. INTRODUCTION

The variation in cutting condition influence and alters the momentary and unpredictable cutting dynamic process and thus affects the cutting tool condition and process stability. As vibration be-come the primary effect, cutting tool condition and machine tools ultimately pay the price with worn tool eventually fails and broken machines. A vibratory system in machining operations consist the machine tool, cutting tool, workpiece and cutting conditions create a complex dynamic behaviour [1]. Vibration signals are one of the most widely analysed because they provide a thorough in-sight into the metal cutting process and considerably less complex in nature, more inclusive, and convenient [2]. The metal cutting process naturally produces vibration signal apart from mechanical vibration harvest from the tool wear and the cutting conditions such as breakage [3]. Many researchers present vibration analysis for tool condition monitoring. Dimla and Lister [4] found that vibration signal is most tool wear sensitive to measure vibration signals in time and frequency domain analysis to predict tool wear. Chen et al. [5], Wang et al. [6], Elangoyan et al. [7], Rao et al. [8] and Rajesh and Namboothi [9] used vibration signals to measure reliability and wear correlation. Therefore make a prediction and develop tool condition monitoring successfully comply with selected method respectively. Recently, Aghdam et al. [10] captured wear sensitive features and derived from autoregressive moving average (ARMA) model of the recorded signals. The outputs of ARMA metric can also be used to provide reliable predictions for the tool.

The determination of transition point between mild and severe wear was the starting point. Mild wear is considered as acceptable wear state whereas the transition to severe conditions often represents a change to commercially unacceptable situations [11]. Therefore, the wear map was developed [12]–[15]. The well-known tool wear progression in machining describe the flank wear versus cutting time. The wear initially increases rapidly and later on gradually reduces to a constant rate until tool failure is reached. The wear rate was almost constant where there was no obvious transition from the steady-state in the failure region [16]. The point between the steady-state and failure region is known as the second transition point, STP.

Significant changes in the tool wear rate indicate the position of the transition time between the states. The values of the maximum wear corresponding to the first and second transition times are considered as the wear state criteria. The first and second criteria were found to be in the range of 0.05 – 0.1mm and 0.15 – 1.00mm, respectively, depending on the type of the operation [17]. Previous researchers have classified flank wear into three conditions in their tool wear monitoring studies using different analyses such as using wavelet analysis [18], fuzzy logic [19] and neural networks [20]. It consists of Phase 1 flank wear (VB = 0 - 0.15mm) were classified as a normal phase. Phase 2 flank wear (VB = 015 - 0.25mm), classified as medium abnormal phase. Lastly, phase 3 flank wear (VB = 0.25 - 0.30mm), classified as critical abnormal

phase. The three-period curve of flank face wear value increases from the initial wear (normal phase), and the wear rate is kept at a high level. It is followed by the steady normal wear (medium abnormal phase), where the wear rate is decreased. When the rapid wear takes place, the wear rate is the highest compared with the initial and normal periods, and therefore, the cutting stability is low. Based on the phenomenon, the ideal tool life should be before the rapid wear [21] where the wear value maximum of 0.250mm where STP is located. These works have been carried out to develop a more favourable tool wear tracking using vibration signals and eventually to predict and estimate the second transition point from the steady-state region into the failure region in tool wear monitoring. Until this paper is written, there is no specific and well-established method to determine the STP. This paper is offering an alternative way by using the previous developed statistical analysis to track the severity of wear index. The index then become input to the sequence of two lower order of polynomials in search for the second intersection point that resemble the STP in tool wear phase.

II. Materials and Method

An innovative, integrated rotating dynamometer was designed and constructed by Rizal to measure the cutting force in a wireless environment system [22]. This dynamometer utilised strain gauge that is mounted on legged cross beam transducer to measure three components of cutting force based on a rotating cutting force system as in Figure 1. Namely, main cutting force, F_c, thrust force, F_t and perpendicular cutting force, F_{cn}. Meanwhile an accelerometer, a thermocouple and strain gauge to measure torque are also built to the rotating dynamometer to build a multi sensor system.

This sensor system is then used for tool wear monitoring in milling P20+Ni tool steel using end milling cutting tool insert was tungsten carbide with multi-layer PVD TiAlN/AlCrN grade ACP200 (Code: AXMT170504PEER-G). A milling process experimental is prepared with various 2^3 full factorial combinations of cutting speed (200 and 373 m/min), feed rate (0.10 and 0.20 mm/tooth), radial depth of cut (0.4 and 0.6 mm) and axial depth of cut is kept constant at 1mm.

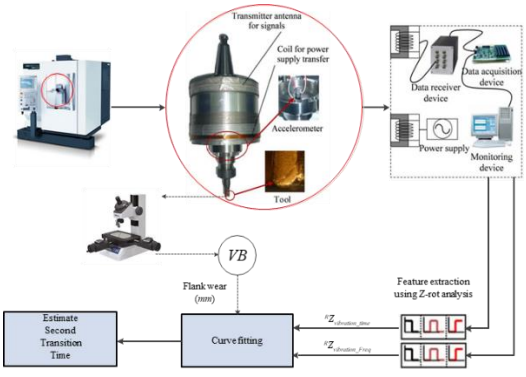


Fig. 1. Experimantal setup

III. Statistical Analysis Method

The analysis produces Z-rot index as an input parameter that employs mathematical and statistical features (mean, standard deviation and Kurtosis) from signal data (such as force signal and vibration) [23]. The development of Z-rot predictive model (ZrPM) starts with the Z-rot tracking analysis method. Afterwards, ZrPM is determined to resolve the unclear transition point from the second region to the third region. The transition point will be assessed to several-underlined curve fit-ting accordingly. The algorithm is summarized as presented in Figure 1.

A. Z-rotation Method

The selected features are suggested to among the best features to study wear correlation with signal amplitude [24]. It is based on a signal element variance scattering around its mean centroid. The method exhibits data pattern in defining the randomness of data features over the whole lifetime to diagnose inferences and expected to have more sensitivity toward amplitude and anomalies changes in a signal. These interpretations are beneficial for prediction and decision making such as in machine learning adaptation. It is also expected to be able to improve the wear progression curve which was unable to exhibit the three typical wear region for the cutting speed more than 120mm/min [25]. Compute the distance, r, for each data variable by subtracting the mean data, $\overline{sig}_y$ from the data variable, sig_y to generate a zero-mean distribution [23].

$$r = (sig_y - \overline{sig}_y) \quad (1)$$

Based on the standard deviation and kurtosis value obtained, will gives the index, Z-rot, that indicates the current condition and records the wear evolution of cutting tool wear during the cutting process [23].

$$^{R}Z = \frac{1}{N}\sqrt{\sigma_r{}^4 K_r} \quad (2)$$

Z-rot is a tracking analysis kurtosis-based use to track the severity of wear. It is expected to show a healthy relationship over the wear evolution.

B. Z-rot predictive model (ZrPM)

Several numbers of complex damage phenomena are happening within these stages (steady state and failure region) and require better understanding. The events are the likelihood to produce different wear progression scenarios. The wear progression scenario is significantly essential to predict the length (time) and propagation rate of each stage to predict the second transition point of a remaining lifetime as depicted in Figure 2.

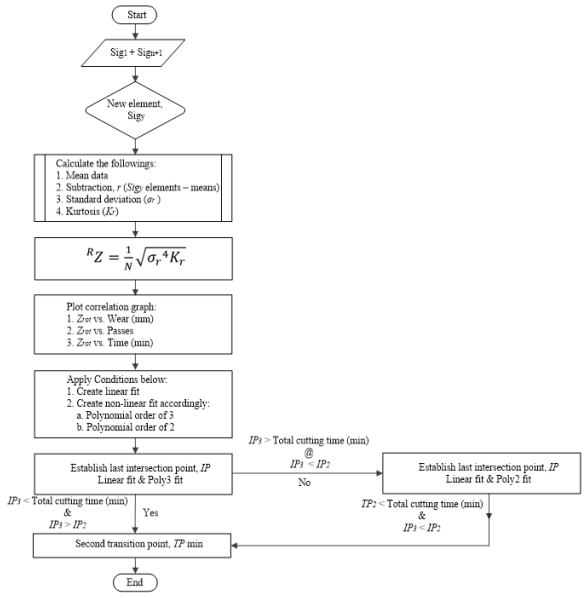


Fig. 2. Flowchart of the Z-rot predictive model (ZrPM)

The first attempt in spotting the time of the second transition point (TP_{STP}), plot a linear fit (Eq. 3) on the scattered data points followed by 2nd (Eq. 4) and 3rd (Eq. 5) order of polynomial regressions to fit the trend line:

$$f(x) = a_1 x + a_0 \quad (3)$$

$$g(x) = a_2 x^2 + a_1 x + a_0 \quad (4)$$

$$h(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0 \quad (5)$$

The decision to decide over the TP_{STP} is determining by the last intersection point between the linear and one of both polynomial regressions. The intersection point should not be extended too far outside the data. The choice is according to conditions:

- At first, consider the last intersection point (TP_3) between linear fit and the 3rd order of the polynomial. The intersection point must be within or under the total cutting time ($T_{Total\ cutting\ time}$) and has a bigger value from the other resulting last intersection point (TP_2) between the linear and 2nd order of the polynomial. The expression as stated:

$$TP_3 < \max(T_{Total\ cutting}) \;\&\&\; TP_3 > TP_2 \quad (6)$$

- If the intersection point does not comply either one or both conditions above, the system will return to consider the TP_2 resulted from crossing point between the linear and 2nd order of the polynomial.

$$TP2 < \max(T_{Total\ cutting}) \quad (7)$$

IV. Results and Dicussion

A. Vibration Signal

Figure 3 and Figure 4 shows the time domain and frequency domain of vibration accelerations of the cutting process detected during the advancement of tool wear. It can be seen that dominant frequency activities occur at relatively low and middle-frequency regions. The high-frequency activities occur at low-frequency regions covering up to 500Hz and contain the most condition indicating information about the cutting process.

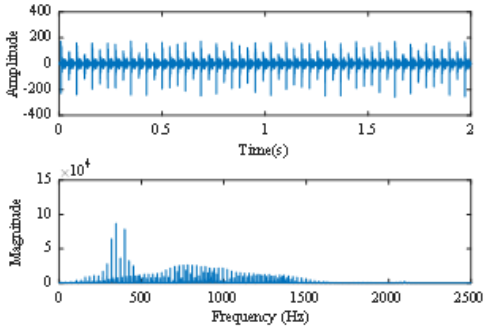


Fig. 3. Representative examples of Time domain and FFT analysis for the cutting vibrations (v = 200m/min, f = 0.1mm and d = 0.6mm) for experiment 2 run 1

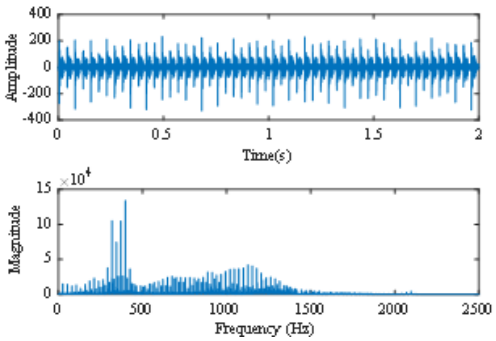


Fig. 4. Representative examples of Time domain and FFT analysis for the cutting vibrations (v = 200m/min, f = 0.1mm and d = 0.6mm) for experiment 2 run 23

The magnitudes are also gradually increased with the advancement of tool wear and notably increased when the flank wear as in Fig. 3 has significantly developed from 0.001mm (run 1) to 0.131mm (run 23). The other frequency activities take place around after 500Hz up to 1500Hz which is the reflection of the damped natural frequency of the tool-workpiece system [26]. It can be seen in all experimental sets that characteristics of the frequency components located at the high-frequency region change with the advancement of wear. They occupy a larger frequency span around range between 500Hz and 1500Hz and their amplitudes rise when the severity of wear is increased.

B. Flank Wear Response

Figure 5 illustrates the statistic (Z-rot index) of the detected vibration acceleration signals at the mentioned cutting condition concerning the wear. During the very early phase of wear development, the amplitude of the vibration acceleration is slightly increased which is correspondingly reflected by the Z-rot. However, it is sometimes reduced when the wear starts developing on the tool's cutting edges which is also indicated as reductions in the Z-rot index as well as in signal amplitude. Nevertheless, the amplitudes of the ensuing Z-rot are then gradually increased with the advancement of wear where amplitude variations are also observed in the frequency domain. When the wear is fully developed over the flank surfaces at the end of wear test, the amplitude of the vibration signal magnitude represented by Z-rot index is notably increased, and the symptoms of tool wear are favorably revealed in the vibration signal plotting alongside with other sensors related. Estimation the Second Transition Point (STP)

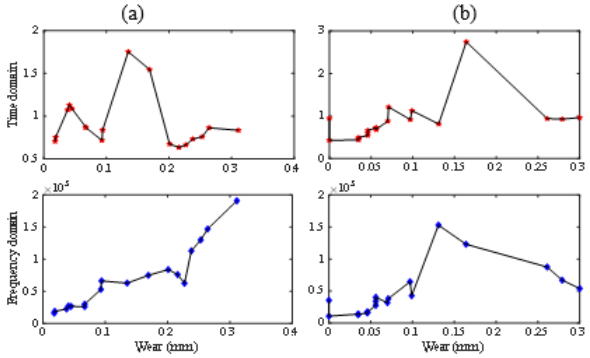


Fig. 5. Z-rot index of the vibration accelerations magnitude of cutting process during the advancement of wear: a) Experimental set 5 – v = 375m/min, f = 0.1mm, d = 0.4mm. b) Experimental set 6 – v = 375m/min, f = 0.1mm, d = 0.6mm.

C. *Estimation of the Second Transition Point (STP)*

The index was plotted using MATLAB software to see the variation within a certain time. The results of the machining tests and the analyses were portrayed together on the same graph for comparison and better understanding. The tool wear mechanisms trending is presented in figure 6.

Flank wear is the major failure pattern regardless the cutting speed. However, the tool wear is more critical at 375m/min compared with that at 200m/min. It is difficult to express the failure numerically, and therefore, the data are hard to collect within the limited experiments, resulting in that the rake face wear is out of the consideration of tool life evaluation in this case. The example of every last intersection point between linear and the polynomial regression fitting respectively of Z-rot vs cutting time are shown in Figure 4. In the time domain and frequency domain extracted of an experiment set 1, the second transition point is achieved at about 58.53s and 57.65s (approximate wear VB = 0.250mm) separately before the tool begins to total failure. While, with different speed level, the second transition point is considered around 15.82s and 14.4s (approximate wear VB = 0.250mm).

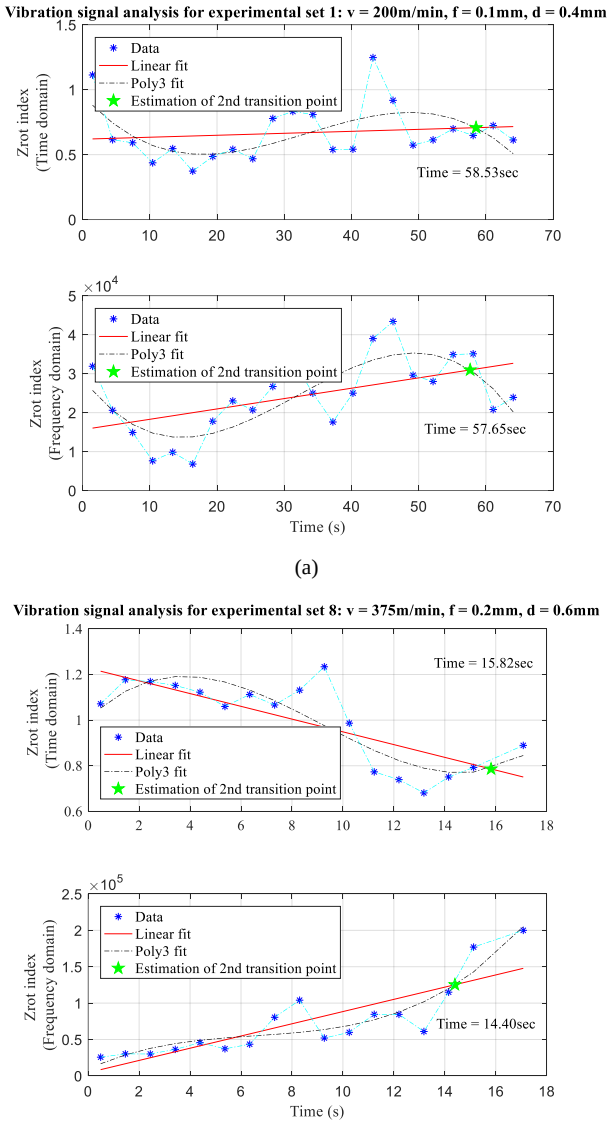


Fig. 6. Z-rot index progression for two experiment conditions and the fittings of polynomial regression of 2nd order and 3rd order

Both experiment condition was assigned to the maximum fit of 3rd order of the polynomial. As the 2nd order of polynomial would be detecting the STP earlier than what 3rd order of polynomial has. Comparing with the cutting speed 200m/min, the time to reach second transition point wear rate is faster at 375m/min approximately after 15s, and the wear value reaches the wear criterion rapidly after 17second. The overall result arrangement is in Table I.

D. *The Relationship Between Vibration and Cutting Force*

The signal data is analyzed in time domain using Z-rot analysis and the illustration as in Figure 7.

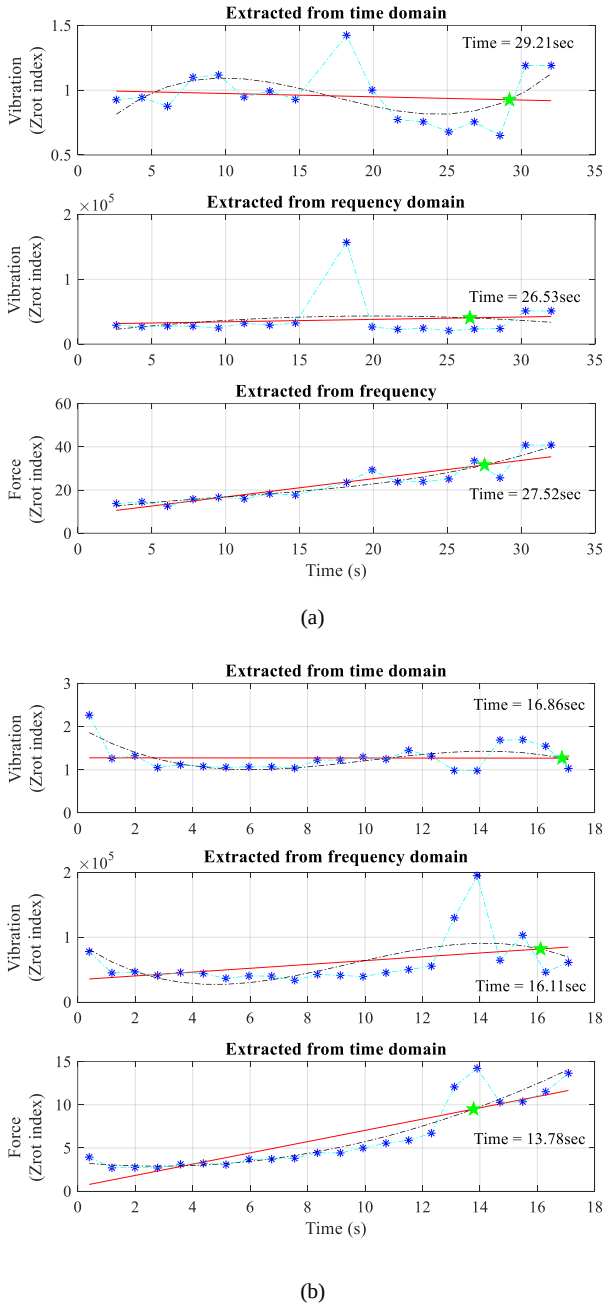


Fig. 7. a) Experimental set 4 - v = 200m/min, f = 0.2mm, d = 0.6mm. b) Experimental set 7 - v = 375m/min, f = 0.2mm, d = 0.4mm

Result validation of the model for estimating the transition times between the second and third states using vibration signal data, Z-rot prediction method also done on force signal data extracted simultaneously from the same experimental set.

The force signal data contain three components of cutting force based on a rotating cutting force system.

From the example observations and analyses made in Figure 7, it is clear that the cutting vibrations do not necessarily have the same varying pattern as that of the cutting forces in machining with either a sharp tool or a worn tool. In machining with a sharp tool and with a worn tool, the cutting forces can be very close to each of experimental sets, but the vibration magnitude can be very different. In addition, larger cutting forces do not necessarily lead to larger vibration amplitudes. For example, Figure 5b shows the time domain resultant force in Z-rot index in machining with a sharp tool is higher than that in time domain Z-rot index for vibration acceleration. In summary, the cutting forces are determined by material property, tool geometry, the cutting conditions, and so on. The cutting forces not only determine the cutting vibration, but also by the structural rigidity (such as damping and stiffness) of the tool–work–machine system [27].

E. Overall Estimation of the Second Transition Point

The outline of the overall result to estimate STP for the various condition, different sensors signal data (vibration, force components and torque) where vibration both in the time domain and in frequency domain while force and torque were analysed in time domain respectively as in Table I.

TABLE I. OVERAL TESTIMATION OF THE SECOND TRANSLASITION POINTS FOR VARIOUS MACHINING CONDITION

Set	Total cutting time (s)	Predicted STP of Average	Set
1	64.06	56.71	~ 0.210
2	64.06	52.02	~ 0.230
3	32.03	27.08	~ 0.210
4	32.03	28.05	~ 0.240
5	34.16	30.66	~ 0.250
6	34.16	28.79	~ 0.240
7	17.08	15.57	~ 0.250
8	17.08	15.53	~ 0.250

Based on the observation made above, the predicted STP values are mostly similar at certain machining conditions. The predicted second transition point has an average of predicted wear 0.235mm as the ideal tool life should be before the rapid wear [22] as the maximum wear land width (VB) at steady state region is 0.250mm. The new statistical feature, Z-rot index have been carried out to as characteristic feature for tool wear tracking using vibration signals and eventually, ZrPM was able to predict and estimate the second transition point from the steady-state region into the failure region in tool wear monitoring.

V. CONCLUSION

The resulting trend of the Z-rot index analysis is useful in determining the second transition point in wear phase. Using vibration signal, ZrPM identified the specific safe cutting time in every test sample. The method is proposed as an alternative autonomous technique to monitor cutting tool wear progression during the machining process. ZrPM is time savvy and successful for observing the tool wear phase and avoiding traditional direct tool wear observation.

REFERENCES

[1] G. H. Lim, “Tool-wear monitoring in machine turning,” J. Mater. Process. Technol., vol. 51, no. 1–4, pp. 25–36.

[2] M. S. H. Bhuiyan, I. A. Choudhury, and Y. Nukman, “Tool Condition Monitoring using Acoustic Emission and Vibration Signature in Turning,” in Proceedings of the World Congress on Engineering 2012 Vol III WCE 2012, July 4 - 6, 2012, London, U.K., 2012.

[3] D. E. Dimla, “Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods,” Int. J. Mach. Tools Manuf., vol. 40, no. 8, pp. 1073–1098, Jun. 2000, doi: 10.1016/S0890-6955(99)00122-4.

[4] D. E. Dimla and P. M. Lister, “On-line metal cutting tool condition monitoring.,” Int. J. Mach. Tools Manuf., vol. 40, no. 5, pp. 739–768, Apr. 2000, doi: 10.1016/S0890-6955(99)00084-X.

[5] B. Chen, X. Chen, B. Li, Z. He, H. Cao, and G. Cai, “Reliability estimation for cutting tools based on logistic regression model using vibration signals,” Mech. Syst. Signal Process., vol. 255, pp. 2526–2537, 2011.

[6] G. F. Wang, Y. W. Yang, Y. C. Zhang, and Q. L. Xie, “Vibration sensor based tool condition monitoring using support vector machine and locality preserving projection.,” Sensors and Actuators, vol. 209, pp. 24–32, 2015.

[7] M. Elangovan, V. Sugumaran, and S. Ramachandran, K. i. Ravikumar, “Effect of SVM kernel functions on classification of vibration signals of a single point cutting tool.,” Expert Syst. Appl., vol. 38, pp. 15202–15207, 2011.

[8] K. V. Rao, B. S. N. Murthy, and N. M. Rao, “Cutting tool condition monitoring by analyzing surface roughness, work piece vibration and volume of metal removed for AISI 1040 steel in boring,” Measurement, vol. 46, pp. 4075–4084, 2013.

[9] V. G. Rajesh and V. N. N. Namboothiri, “Flank wear detection of cutting tool inserts in turning operation: application of nonlinear time series analysis,” Soft Comput., vol. 14, pp. 913–919, 2010.

[10] B. H. Aghdam, M. Vahdati, and H. M. Sadeghi, “Vibration-based estimation of tool major flank wear in a turning process using ARMA models,” Int J Adv Manuf Technol, vol. 76, pp. 1631–1642, 2015.

[11] J. A. Williams, “Wear modelling: analytical, computational and mapping: a continuum mechanics approach,” Wear, vol. 225, pp. 1–17, 1999, doi: 10.1016/S0043-1648(99)00060-5.

[12] S. . Lim, “Recent developments in wear-mechanism maps,” Tribol. Int., vol. 31, no. 1, pp. 87–97, 1998, doi: 10.1016/S0301-679X(98)00011-5.

[13] M. F. Ashby and S. C. Lim, “Wear-mechanism maps,” Scr. Metall. Mater., vol. 24, no. 5, pp. 805–810, May 1990, doi: 10.1016/0956-716X(90)90116-X.

[14] A. Cantizano, A. Carnicero, and G. Zavarise, “Numerical simulation of wear-mechanism maps,” Comput. Mater. Sci., vol. 25, no. 1, pp. 54–60, 2002, doi: 10.1016/S0927-0256(02)00249-5.

[15] R. Bosman and D. J. Schipper, “Mild Wear Prediction of Boundary-Lubricated Contacts,” Tribol. Lett., vol. 42, no. 2, pp. 169–178, 2011, doi: 10.1007/s11249-011-9760-3.

[16] H. Z. Li, H. Zeng, and X. Q. Chen, “An experimental study of tool wear and cutting force variation in the end milling of Inconel 718 with coated carbide inserts,” J. Mater. Process. Technol., vol. 180, no. 1, pp. 296–304, 2006, doi: 10.1016/j.jmatprotec.2006.07.009.

[17] V. P. Astakhov, “The Assessment of Cutting Tool Wear,” Int. J. Mach. Tools Manuf, vol. 44, no. 6, pp. 637–647, 2004.

[18] K. Zhu, Y. S. Wong, and G. S. Hong, “Multi-category micro-milling tool wear monitoring with continuous hidden Markov models,” Mech. Syst. Signal Process., vol. 23, no. 2, pp. 547–560, Feb. 2009, doi: 10.1016/j.ymssp.2008.04.010.

[19] G. Wang, L. Qian, and Z. Guo, “Continuous tool wear prediction based on Gaussian mixture regression model,” Int. J. Adv. Manuf. Technol., vol. 66, no. 9, pp. 1921–1929, 2013, doi: 10.1007/s00170-012-4470-z.

[20] S. Cho, S. Binsaeid, and S. Asfour, “Design of multisensor fusion-based tool condition monitoring system in end milling,” Int. J. Adv. Manuf. Technol., vol. 46, no. 5, pp. 681–694, 2010, doi: 10.1007/s00170-009-2110-z.

[21] W. Ji, J. Shi, X. Liu, L. Wang, and Steven Y. Liang, “A Novel Approach of Tool Wear Evaluation,” J. Manuf. Sci. Eng., vol. 139, 2017.

[22] M. Rizal, J. A. Ghani, M. Z. Nuawi, and C. H. C. Haron, "A Wireless System and Embedded Sensors on Spindle Rotating Tool for Condition Monitoring," Adv. Sci. Lett., vol. 20, no. 10–12, pp. 1829–1832, 2014.

[23] N. A. Kasim, M. Nuawi, J. Ghani, M. Rizal, M. A. F. Ahmad, and C. Haron, "Cutting tool wear progression index via signal element variance," J. Mech. Eng. Sci., vol. 13, pp. 4596–4612, 2019, doi: 10.15282/jmes.13.1.2019.17.0387_rfseq1.

[24] P. Krishnakumar, K. Rameshkumar, and K. I. Ramachandran, "Tool Wear Condition Prediction Using Vibration Signals in High Speed Machining (HSM) of Titanium (Ti-6Al-4V) Alloy," Procedia Comput. Sci., vol. 50, pp. 270–275, 2015, doi: 10.1016/j.procs.2015.04.049.

[25] Z. M. Nuawi, F. Lamin, M. N. M. J., N. Jamaluddin, S. Abdullah, and E. N. C. K., "Integration of I-kaz Coefficient and Taylor Tool Life Curve for Tool Wear Progression Monitoring in Machining Process," Int. J. Mech., vol. 1, no. 4, pp. 44–50, 2007.

[26] I. Yesilyurt and H. Ozturk, "Tool condition monitoring in milling using vibration analysis," Int. J. Prod. Res., vol. 45, no. 4, pp. 1013–1028, 2007.

[27] N. Fang, P. S. Pai, and S. Mosquea, "A comparative study of sharp and round-edge tools in machining with built-up edge formation: cutting forces, cutting vibrations, and neural network modeling," Int J Adv Manuf Technol, vol. 53, pp. 899–910, 2011.

Electronics Diagnose for Jaundice Disease Integrated with Informative SMS Notification: Diagnose Jaundice Expert System (DIJAD)

1st Fazeha Nasya Abd Malik
Mathematics, Science and Computer Department,
Politeknik Mukah Sarawak
Malaysia
fazehanasya@pmu.edu.my[1]

2nd Muhammad Thariq A Razak
Information Technology and Telecomunication
Politeknik Mukah Sarawak,
Malaysia
m.thariq@pmu.edu.my[2]

***Abstract*—Medical expert system is very important nowadays to detect any disease as early as we can. Electronics Diagnose For Jaundice Disease Integrated With Informative SMS Notification: Diagnose Jaundice Expert System (DIJAD) has been designed and developed to help early detection of jaundice by identifying the symptoms statistically after diagnosing the patient, based on the information available in its database. The system is conducted to detect the type of the jaundice faster based on the initial sign or symptom of the disease. It will provide the result of diagnosing jaundice and the nearest clinic via Short Message Service (SMS) to the patient. The interview were conducted to validate the functionality and effectiveness of the DIJAD, two method is used which is user test and system function test. The result indicated that 5 random user 2 doctors agree that the system is user-friendly and easy to use. Other than that, the result from system function test. It can be concluded that the system is fully functioning, and no error occurred. This system uses Forward Chaining method. Forward Chaining is method that is driven by the data where the tracking stars from the observation of the input information and the try to describe the conclusions. DIJAD developed using Hypertext Preprocessor (PHP) for the programming language in conjunction with MySQL as a database system. The outcome of DIJAD able to facilitate community in diagnose early jaundice disease that can save time and give positive impact in our effort to prevent jaundice to become worst.**

Keywords—Expert System, jaundice, symptoms, disease, treatment, Short Message Service

I. INTRODUCTION

Jaundice is occurred when the skin, nails, tongue and the white eyes turns into yellow colour. This disease appeared when there is excessive amount of bilirubin in blood of human being of all ages. Jaundice indicates the difficulty with the liver. It also known as one of sign liver problem. The causes of jaundice are classified into three types pre-hepatic, intra-hepatic and post-hepatic.

This disease not only experience by the adults but also endure by new born babies. According to a research, 60% normal new born babies and 80% premature babies suffered from jaundice [1]. Neonatal jaundice is very common, and it typically appears in first week of the new born. Neonatal hyperbilirubinemia is classified as unconjugated or conjugated. Jaundice goes away when baby's liver develops and they begins to feed, which helps bilirubin pass through the body.

DIJAD has the information of jaundice and early treatment suggestions for the patients. At the moment, the awareness about jaundice is considered low in our society. DIJAD is the online system platform where its users can find all the information about jaundice as a one stop center. Many adults would have not been aware if they have jaundice symptoms. Early detection could help the patients to seek appropriate treatment and eventually would reduce the implication of this disease. Early precaution for the patient via Electronics Diagnose for Jaundice Disease Integrated with Informative SMS Notification called DIJAD is the best way for early treatment before it become worst.

The role and importance of healthcare systems to improve quality of life and social welfare in a society have been well recognized. Attention should be given to raise awareness and implementing appropriate measures to improve health care. Therefore, a computer based system is developed to serve as an alternative for people to self-diagnose their health status based on given symptoms. This strategy should be emphasized so that people can utilize the information correctly as a reference to enjoy healthier life. Hence, DIJAD is developed based on expert system technique. Inference technique is employed in the system to enable information about treatment of the diseases based on given symptoms. Web based technology is used as a platform to disseminate the information to users in order for them to optimize the information appropriately. This system will benefit people who wish to increase health awareness and seek expert knowledge on the diseases by performing self-diagnosis for early disease detection.

Thus, the objective of the system are to design and develop web based Diagnose Jaundice Expert System (DIJAD) that can diagnosing the type of jaundice. Other than that, to

validate the functionality of DIAD by implementing System Function Test and User Testing. Besides that, scope of the system user able to access the DIJAD at any place and any time, making diagnosing easier and reducing the time for medical expert in consultation. Medical experts can gain access to the DIJAD and update useful possible information related to system on their electronic devices.

Remaining of this paper is organized as follows: Section 2 describes study background for literature review that includes artificial intelligence, expert system, forward chaining method and comparison between existing system; Section 3 discusses the methodology; Section 4 presents finding and analysis; Section 5 concludes this study.

II. LITERATURE REVIEW

In this part will discuss about artificial intelligence, expert system, forward chaining and comparison with existing system as per below details.

A. Artificial Intelligence

In artificial intelligence, an expert system is computer system that emulates the decision-making ability of human expert. Electronic Diagnose for Jaundice Disease Integrated with Informative SMS Notification: Diagnose Jaundice Expert System (DIJAD) are designed based on medical expert knowledge, represented mainly as if-then rules rather than through conventional procedure code. Many other AI applications were employed in various healthcare sectors, like Radiology, Screening and Disease Diagnosis. Several hospitals including Mayo Clinic, USA and the National Health Service, UK have developed their own Intelligent system [2,3], as well as Google [4] and IBM's [5] contributions to healthcare technology advancements.

B. Expert System

An expert system can be modified based on current changes. There are three main component in this system; knowledge base, inference engine and user interface [6]. Much like other artificial intelligence systems, expert system's knowledge may be enhanced with add-ons to the knowledge base, or additions to the rules.

Knowledge Base in expert system consist the data and information about custom domain. This knowledge base is important in modern intelligent system and application in information strategy, planning, design, scheduling, error monitoring, diagnosing and so on [7]. Collected data and knowledge will be transformed knowledge representation. There are the fundamental types of knowledge representation technique; production rules, semantic network, frame representation, logical representation and hybrid representation.

Inference engine applies to run operation and processing the knowledge. It will determined when the fact or rule will be applied. User Interface allowed the user and expert system exchange information in two way communication. It is function to obtain the graphical information from user and translate the input into the machine language that understand by the system. As a result, user will receive the advice or suggestion

C. Forward Chaining

Forward chaining system is the system that works from known fact to the hypothesis or goal. This system in fact part IF towards the goal THEN as shown in Figure 1.

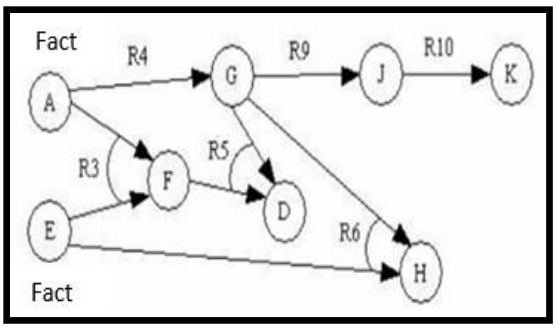


Fig. 1. Forward chaining technique

Electronics Diagnose for Jaundice Disease Integrated with Informative SMS Notification are used the forward chaining technique. This technique selected instead of backward chaining technique. In view of the forward chaining is the system where it works from the recognized fact to the hypothesis or goal.

D. Comparison Between Existing System

The first expert system called MYCIN was introduced in the early 1970s after almost 6 years of development at Stanford University, USA [8]. A few expert system had been listed to compare about the rule, programming language, module, technique used and others specifications. It has shown in Table I.

1) Diagnose Lungs Disease towards Children Expert System.

This expert system is limited to lungs disease towards children [9]. The development of this system is based on the web. Two programming language were used in this system. The result that can be obtained from this expert system are the type of the disease, the explanation of the disease and the treatment suggest to the patient. Logical rule IF-THEN is used for the diagnose result.

2) Diagnose Diabetes Expert System.

This expert system using the Clips programming language [10]. It is develop by providing the diagnose modul that consist some questions based on the diabetes symptom. The questionnaire included in this system based on yes or no answer. In addition, the evaluation based on the laboratory test is possible to be included in this system.

3) Diagnose Migraine Expert System.

The development of this expert system is the prototype that assist doctor to diagnose migraine [11]. This system applied by using knowledge input process to identify the main source of the application where the doctor will be impersonate as the domain expert.

4) Suggestion for Expert System. Diagnose Jaundice Expert System (DIJAD).

This system developed to diagnose the jaundice in Malaysia with additional Informative SMS Notification to

the patients. It will carried out the knowledge regarding jaundice towards new born babies and adults. This system is not only giving the information about jaundice and the type of the jaundice, but also the information of the nearest hospital or clinic will be provided via Short Message Service (SMS). Comparison between all the system as per shown below.

TABLE I. COMPARISON OF EXPER SYSTEM

Specification	Name of Expert System			
	Diagnose Lungs Disease towards Children Expert System	Diagnose Diabetes Expert System	Diagnose Migraine Expert System	Diagnose Jaundice Expert System
Rule	IF-THEN Logic	IF-THEN Logic	IF-THEN Logic	IF-THEN Logic
Programming Language	HTML and PHP	Clips	Clips	PHP
Login Module	Developer and User Login	Develop er and User Login	Developer and User Login	Developer and User Login
Technique	Forward chaining and backward chaining	Forward chaining	CLIPS	Forward chaining
Supported Answer	Yes/No	Yes/No	Yes/No	Yes/No
Inference Engine	Knowledge Base	Clip	Knowledge Base	Knowledge Base
Notification	None	None	None	SMS

According to the Table I, all system are using the expert system technique. The software applied to develop this system are *Adobe Dreamweaver*, *Visual Basic 6.0* and *Java Netbeans*. For DIJAD, *Adobe Dreamweaver* software and *MySQL* database will be used. Login Module designed towards the system administrator and user.

This system also using the forward chaining technique to find the result based on the symptoms given. The different between the existing expert system and the suggested expert system is, the Electronics Diagnose for Jaundice Disease Integrated with Informative SMS Notification consist the SMS Notification. This includes the information about jaundice and nearest hospital or clinic for further treatment.

III. METHODOLOGY

In this section, the methodology for completing the expert system development will be discuss. This system applied the qualitative approach to analyse the quality of the object [12].

The effective way to understand the function and the rule of certain matter is from interviewing an expert in specific field [13]. Thus, the interview between the doctor who diagnose infectious disease, consultation internal medicine and the doctor specialist of children, consultation paediatrician, had been held during proposal stage. All issues or problem related to jaundice and the treatment needed had been ask to the both Specialist Doctor from Putra Specialist Hospital, Batu Pahat, Johor.

Technical issues concern the knowledge had been implemented to describe and design the knowledge system based [14]. Final phase specifically the maintenance phase will be dismissed as it is only practice by the certain organisation only [15]. Table II shows the methodology process in Knowledge Base.

TABLE II. METHODOLOGY PROCESS OF KNOWLEDGE BASE

Phase	Activities	Outcomes
Problem assessment system	Determine the problem. Identify the main engagement for the project. State the objective of the project.	Objective and scope of the project identified. Hardware and software had been determined.
	Write the proposal	Proposal had been verified.
Knowledge Acquisition	Collect and analyse the data and knowledge. Prepare the main concept for the system design.	The symptoms of the disease gathered for analysing. Studied the equivalent system function which is diagnose diabetes Expert System
	Analyse the information and system requirement and user desire.	Select the suitable technique that is forward chaining technique
Prototype system development	Select the requirement to develop the expert system. Modify the data and knowledge representation. Design and implement the prototype system. Examine the prototype with the certain case.	Design the Data Flow Diagram (DFD) and Entity Relationship Diagram (ERD) to explain the function. Design the user interface, database and the accomplishment of the physical architecture system design
Completion of system development	Collect the additional data and knowledge. Develop the user interface. Complete the entire system.	System prototype. The reliable information shown and delivered to user. The interface is user-friendly and easy to be applied.
	Module coding and testing	
Assessment and review	Evaluate the system based on the criteria and performance.	Electronics Diagnose for Jaundice Disease Integrated with Informative SMS Notification had been reviewed to confirm the objective are

		fulfilled

IV. FINDING AND ANALYSIS

Analysis and system design are the details study to determine the problem and the objective of the system. The accumulate data and information are beneficial for the system development. Designing stages divided into 2 type; interface design and database design.

A. *Interface design*

The interface design applied to make sure interaction of the user to control and display the system. This interface must be practical to ensure it is suitable to the user requirement.

Fig. 2. Interface of DIJAD homepage

B. *Database design*

Database structure were built based on the Entity Relationship Diagram (ERD). This system using *MySQL* database software. This design is the logical data structure where all the data will be stored by the administrator.

- *Entity Relationship Diagram (ERD).* ERD portrayed the relationship between the entities in database.
- *Relationship Scheme.* It is the written description respected to the data in the database. The entities as shown below;

1. Symptom
   (id_simptom(pk),id_penyakit,id_pentadbir,simptom,peratus)
2. Disease
   (id_penyakit(pk),id_pentadbir,id_kategori,peratus,keterangan,pencegahan,rawatan,penyakit)
3. Administrator
   (id_pentadbir(pk),id_group,email,nama,jawatan,katalaluan,namapengguna)
4. Category
   (id_kategori(pk),kategori)
5. Result
   (id_keputusan(pk),nama,notelefon,jantina,kategori,poskod,id_penyakit,tarikh_create)

C. *Electronics Diagnose for Jaundice Disease Integrated with Informative SMS Notification: DIJAD. IF THEN structure.*

The interactive system design created to ensure it will the beneficial to user. Criteria decision method is important to generate the appropriate criteria and factor which is relevant to the problemsolving decision.

1) *Inference Network*

This technique is the list of jaundice symptoms diagram which identify the type of jaundice.

2) *Rule Set*

Rule set for Electronics Diagnose for Jaundice Disease Integrated with Informative SMS Notification is using IF THEN rule. This rule related to the symptoms of all types of jaundice suffered by newborn and adult. Knowledge representation schemes is a common rule. The expert system had been commercialized based on rule-base systems.

```
IF Warna urin kuning, coklat
OR Badan letih dan lemah
OR Mendapat lelah atau asma
OR Kulit semakin pucat
ATAU Kulit gatal
ATAU Mata dan kulit berwana kuning
MAKA Pre-hepatic
```

Fig. 3. IF THEN Structure

3) *Database Relationships*

This is the important element to make sure the information management restored in database. The information are possible to be added, withdraw and updated. Figure 4.0 shows the coding regarding the database;

```
<?php
$db_host = 'localhost';
$db_name = 'sistempakardijad';
$db_user = 'root';
$db_pass = '';

$db_link = mysql_connect($db_host, $db_user, $db_pass);
mysql_select_db($db_name); ?>
```

Fig. 4. The relationship of database and Interface in programming

D. *Electronics Diagnose for Jaundice Disease Integrated with Informative SMS Notification: DIJAD Implementation.*

In this phase, the database programming allow the user manage the information in database. This is the important element to make sure the information management restored in database. The information are possible to be added, withdraw and updated.

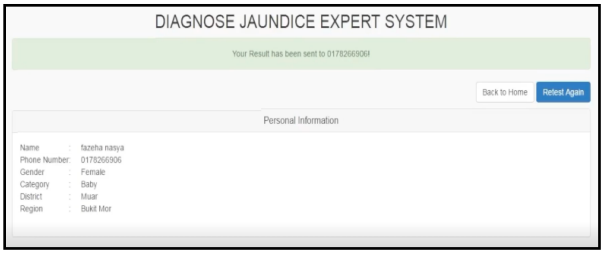


Fig. 5. Result was sent via SMS

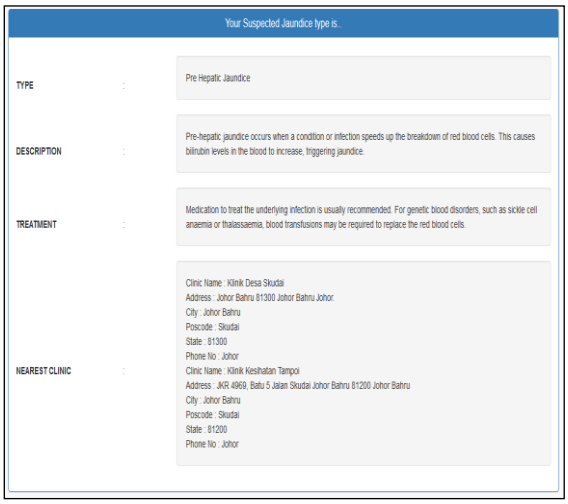


Fig. 6. Diagnose jaundice result details

The result consist the type jaundice disease of new born babies or adult been diagnosed, details about type of jaundice, suggested treatment and the nearest clinic they can visit. This information will be send to patient via SMS notification simultaneously.

The analysis of the system are divided into two part; System function test and user test. This analysis is important to ensure the system provide the feedback and reliable to the user.

E. System function test

For this section, the characteristics and the function will be analysed to verify the information is exact as the data listed before the system developed. The summary of the system function test shown in Table III.

TABLE III. SYSTEM FUNCTION TEST

No.	Experiment Plan	Expected Output	Actual Output
1	User will click the 'Next' button once register	First category question and answer displayed	First category question and answer displayed
2	User will click 'Next' button	Next category question displayed	Next category question displayed
3	User will click the 'Back to Home' button	Diagnose process ended	Diagnose process ended
4	User will click the 'Diagnose Again'	First category question and answer displayed	First category question and answer displayed
5	Diagnose Result	Type of jaundice, explanation, treatment and nearest hospital or clinic displayed. On the same time the SMS Notification will be send to the user	Type of jaundice, explanation, treatment and nearest hospital or clinic displayed. On the same time the SMS Notification will be send to the user

F. User Test

This test is to make sure the objectives of the system fulfilled. The test had been run by five users and a doctor. The feedback analysis shown in the Figure 7 based on the questionnaire towards the user. It shows that 40% of user mostly approved this user-friendly system while another 60% approved this user friendly system.

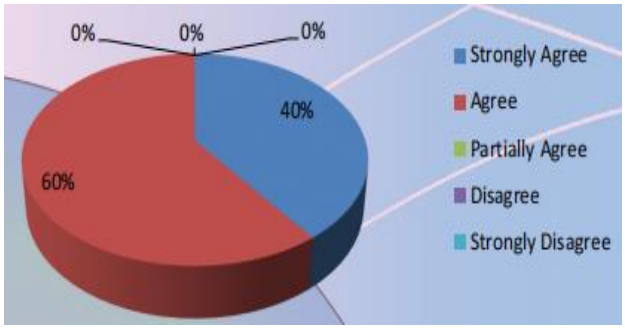


Fig. 7. The feedback of questionnaires

In general, the achievement of the project and suggestion to enhance the productivity of the system;

- Help users to determine whether they suffered from jaundice, suggest the appropriate treatment and provide the information of the nearest clinic.
- Deliver the information related the type of jaundice. Result of diagnose jaundice will receive via SMS to user.
- The system design is easy to manage and user-friendly. User easily to get start diagnose in the system.
- The developer of this system qualified to update the system occasionally.
- To provide the graphical report that shown the statistic of the jaundice that infected in community for beneficial research.

V. CONCLUSION

This system point out that Electronics Diagnose for Jaundice Disease Integrated with Informative SMS Notification: Diagnose Jaundice Expert System (DIJAD) help the patient to diagnose the type of the jaundice, the suitable treatment and recommend the nearest clinic for further consultation and proper treatment. In consequence, this system will help the patient to aware and in the same time able to decrease the risk of the jaundice. The result from the system function test and user test indicates that the system are fully functioning, no logic error occurs and result from questionnaire proved that mostly user are agree that the system are user friendly and able to help user to identify jaundice disease by using IF-ELSE statements.

ACKNOWLEDGMENT

We acknowledge the Politeknik Mukah and Universiti Tun Hussein Onn Malaysia for provision and support for this project. This paper would not have been possible

without the exceptional support of the supervisor, family and friends. The encouragement are much appreciated and strength.

REFERENCES

[1] Asmunni Yahya, 2010. *Masalah Kuning Bayi.* [Accessed 28 June 2020], Available from: http://www.dbmhospital.com.my/home/morearticle.asp?articleid=20.

[2] C.Ross and E. Brodwin. "At Mayo Clinic, AI engineers face an 'acid test' – STAT," Dec. 18, 2019. [Accessed 2 Feb 2021], Available from: https://www.statnews.com/2019/12/18/mayo-clinic-artificial-intelligence-acid-test/

[3] "Artificial Intelligence: How to get it right". [Accessed 2 Feb 2021], Available from: https://www.nhsx.nhs.uk/ai-lab/explore-all-resources/understand-ai/artificial-intelligence-how-get-it-right/

[4] CB Insights, "How Google Plans to Use AI To Reinvent The $3 Trillion US Healthcare Industry." Apr. 19,2018. [Accessed 2 Feb 2021], Available from: https://www.cbinsights.com/research/report/google-strategy-healthcare/

[5] Deepak Kumar, "Four ways in which Watson is transforming the healthcare sector". [Accessed 2 Feb 2021], Available from: https://www.healthcareglobal.com/technology-and-ai-3/four-ways-which-watson-transforming-healthcare-sector

[6] Darlington K, 2000. *The Essence of Expert Systems.* Prentice Hall.

[7] Abraham. A, 2005. *Rule-based Expert Systems.* Oklahoma State University, Stillwater, OK, USA.

[8] Stanford University. Computer Science Dept. Heuristic Programming Project and E.H Shortlife, *MYCIN: a Knowledge-based Computer Program Applied to Infectious Diseases.* 1977.

[9] Septiana Irwanti, 2009. *Sistem Pakar Diagnosis Penyakit Paru Pada Anak Berbasis Web.* Universitas Sebelas Maret : Diploma Tesis

[10] Zeki, Malakooti, Ataeipoor and Tabibi, 2012. An Expert System for Diabetes Diagnosis. *American Academic & Scholarly Research Journal, 4(5),* 1-13.

[11] Kopec, Shagas, Selman, Reinharth dan Tamang, 2004. Development of an Expert System for Aiding Migraine Diagnosis. *The Journal on Information Technology in Healthcare, 2(5),* 355-364.

[12] Reaves, C.C. 1992. *Quantitiative research for the behavioural sciences.* New York, NY: John Wiley & Sons.

[13] Hakim, Catherine, 1987. *Research design: strategies and choices in the design of social research.* Contemporary Social Research Series (13). London, UK : Allen and Unwin.

[14] Feigenbaum, McCorduck. 1983. *'Computers in Your Future'.* Dicapai pada September 1, 2015, dari http://www.nybooks.com/articles/archives/1983/dec/08/computers-in-your-future/

[15] Negnevitsky, 2005. *Artificial Intelligence: A Guide to Intelligent System.* Second Edition. USA: Addison Wesley.

Advantages of Hyperspectral over RGB image on Land Cover Classification

1st Cheng Hao Ko
Graduate Institute of Automation and Control
National Taiwan University of Science and Technology
Taipei, Taiwan
kevin.ko@mail.ntust.edu.tw

2nd Jajang Jaenudin*
Graduate Institute of Automation and Control
National Taiwan University of Science and Technology
Taipei, Taiwan
Mechatronics Department
Politeknik Caltex Riau
Pekanbaru, Indonesia
jajang@pcr.ac.id

3rd Vincentius Christian Bintang
Graduate Institute of Automation and Control
National Taiwan University of Science and Technology
Taipei, Taiwan
bintangvincent@gmail.com

4th Jih-Run Tsai
National Space Organization (NSPO)
Taiwan
Hsinchu, Taiwan
jrtsai@nspo.narl.org.tw

5th Shin Fan Lin
National Space Organization (NSPO)
Taiwan
Hsinchu, Taiwan
sflin@nspo.narl.org.tw

6th Jiun-Kai Tseng
Graduate Institute of Automation and Control
National Taiwan University of Science and Technology
Taipei, Taiwan
tseng.kai@gmail.com

***Abstract*—Crop identification and land cover estimation are essential for farming and land management practices in the precision agriculture field. Conventional measurements are expensive and time-consuming and thus cannot be treated as appropriate for large areas. An automatic crop or land classification should be applied to overcome these problems. Therefore, high-quality data availability is required to feed the classification tools. To fulfill the needs, we have used an airborne system for collecting in the Taiwan agriculture area. A VNIR hyperspectral image has been proven to significantly increasing accuracy compared to an RGB image. With simple discriminant algorithm LD and QD, the classification accuracy of VNIR images reaches 88.14 % and 92.02%, respectively. Meanwhile, RGB images attain 52.73% and 52.27%.**

Keywords—RGB Image, Hyperspectral Image, Separability, Classification, Accuracy.

I. INTRODUCTION

Agriculture is an important socio-economic element that needs to develop to sustain national food security continually. From a business, perspective is essential to know to forecast the supply and demand of the particular farm commodity, plan landscape management, and regulate the market price [1].

Remote sensing is one of the technical methods in precision agriculture mapping to forecast the farm product in industry 4.0 emergence [2]. This technique can tackle ineffective and inefficient traditional ways because the direct field investigation by measuring the farm landscape manually. one way to utilize technological method development to optimize the cost-benefit process in production forecasting. Hence the intelligent agriculture innovation can contribute.

One standard method for automatic crop yield estimation uses airborne remote sensing and classification methods to identify the crop species. The emerging of remote sensing is supported by the technology development of existing imaging sensors. The hyperspectral imaging (HSI) sensor can offer broader information compare to Multispectral (MS) and RGB sensors [3][4]. As the state-of-the-art data cube with a broader band, HSI can provide richer features that enhance the classification performance [5].

Remote sensing using UAV or satellite [6] as an imaging device platform combined with machine learning methods is standard in object detection and mapping [7]. The technique of taking pictures using an airborne system allows hyperspectral cameras to be transported, which is not yet possible with UAVs. In addition to having a vast spectrum from VNIR to SWIR with a very high resolution (3-10 nm) per band, the resulting image also has a competitive spatial resolution compared to UAV systems. Increasing the number of features will increase class separability, wherein images with limited features between one class and another class with similar characteristics cannot be separated/classified. However, with the addition of feature data, the differences in characteristics between classes become more significant. With the bursting of features in hyperspectral data, of course, it will not be very easy for computation if using complex classification algorithms [8]. Therefore, selecting a simple algorithm such as the discriminant classifier is intended to make the training and testing process more efficient [9]. The

research aims to provide direction on the importance of hyperspectral data in mapping a specific area. Even with a simple algorithm method, the classification performance can achieve a reasonably high performance compared to using only RGB images.

The research aims to provide direction on the importance of hyperspectral data in mapping a particular area. Even with a simple algorithm method, the classification performance can achieve a reasonably high performance compared to using only RGB images. Therefore, in this works, it will be shown that replacing RGB data with hyperspectral data can improve the performance of a structurally simple classifier.

The problem of mapping agricultural land is that it is ineffective and inefficient if the investigation is carried out manually, directly in the field. It takes a long time and a very high cost to map a large area. The use of imaging with aerial photos is beneficial for the land cover classification process. However, imaging consisting only of the RGB spectrum is not reliable enough when classified using a simple classification method. Reliable imaging with more spatial resolution and spectrum is required. The classification accuracy of RGB images may be improved by choosing a more complex algorithm such as SVM [10] or even by deep learning [11]. However, this method will not be effective because the spectral features are minimal. In this work, many methods for getting Non-RGB images with various spatial and spectral resolutions, such as satellite, airborne, or unmanned aerial vehicles (UAV). The platform is used a manned plane that carries the hyperspectral imaging system to do so farmland landscape scanning mission.

The drawback of using hyperspectral data compared to RGB is the redundancy of feature information contained within it [12]. Thus, it is necessary to unmix or reduce the dimensional spectral features to gain better separability. Principal component analysis (PCA) is the standard method to reduce the spectral into its most significant feature is principal component analysis (PCA). This method significantly improves spectral classification by finding the most principle band to differentiate between similar classes. These steps are only crucial for sophisticated methods. This works omit this procedure and include all bands because of the simplicity of discriminant analysis.

Therefore, in this presentation, it will be shown that replacing RGB data with hyperspectral data can improve the performance of a classifier that is structurally very simple.

II. Study Area and Data Set

A. Study Area

The study area is part of Yunlin County (southwest Taiwan), located between 120°20'58"E–120°21'6"E and 23°44'22" N –23°44'29" N (around 59.624 m^2). RGB and VNIR hyperspectral image on study area shown in Fig. 1. Some different crops and soil cover the area.

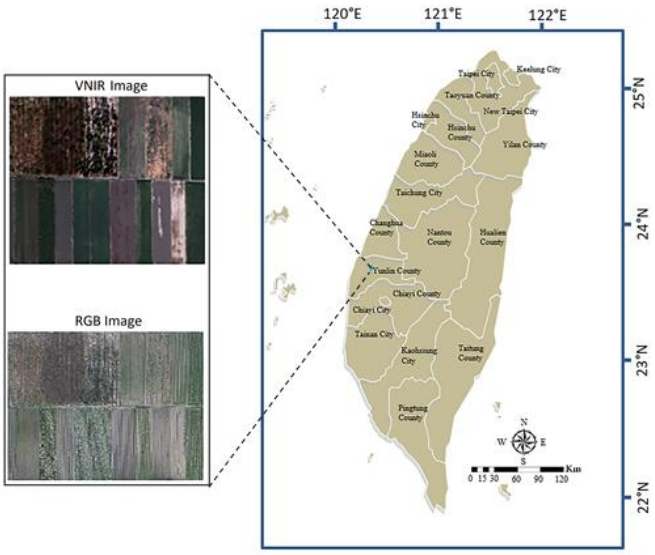


Fig. 1. VNIR and RGB image on the study area

B. Data Set

The UAV system has taken the RGB image. Meanwhile, the hyperspectral is collected by the airborne system. The Taiwan Agriculture Research Institute (TARI) served the ground truth or label by direct field investigation. The label is shown in Fig. 2.

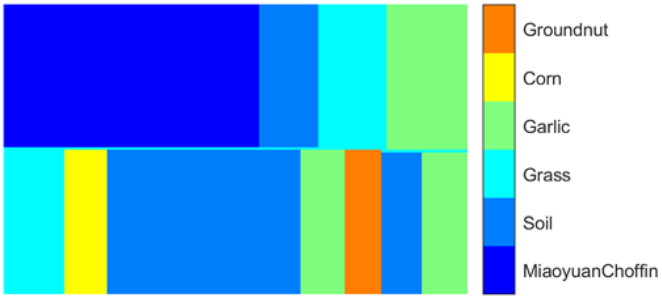


Fig. 2. Label for training and testing data set

The data set spatial size is 331-pixel x 367 pixels, a total of 121.477 pixels consisting of 6 different imbalanced classes. The data has been split 30% for training and 70% for testing. Class pixel distribution is shown in Table I.

TABLE I. Class Pixel Distribution

No.	Class Name	Number of Pixels	Pixel for Training	Pixel for Testing
1	Miaoyuan	33.128	10.100	23.028
2	Soil	38.169	11.600	26.569
3	Grass	17.518	5.100	12.418
4	Garlic	22.267	6.750	15.517
5	Corn	5.610	1.700	3.910
6	Groundnut	4.785	1.450	3.335
Total pixels		121.477	36.700	84.777

III. Data Characteristic

An RGB (red, green, blue) image is a three-dimensional byte array that explicitly stores a color value for each pixel. RGB image arrays are three channels of color information.

Meanwhile, a hyperspectral image consists of hundreds or thousands of narrower bands (10-20 nm). Fig. 3 shows the spectrum of RGB and hyperspectral images. Jeffries-Matisuta value is used to check the separability between classes of both RGB and hyperspectral images.

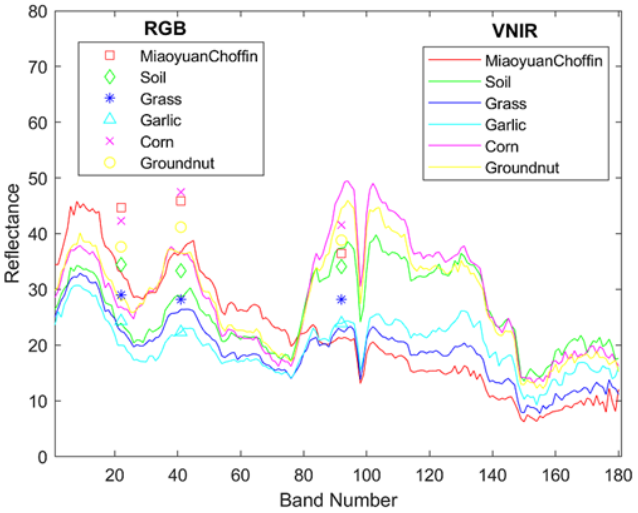


Fig. 3. RGB and VNIR spectral signature

The Jeffries–Matusita (J–M) distance method [13][14] is suitable to describe the class separability of multispectral or hyperspectral images. This method is based on Bhattacharyya distance, one of the most common ways to predetermine the statistical separability of two classes of materials. This distance measure in terms of only the first two moments of these two density functions is

$$B=\frac{1}{8}\left[\mu_x-\mu_y\right]^T\left[\frac{\Sigma_x+\Sigma_y}{2}\right]^{-1}\left[\mu_x-\mu_y\right]+\frac{1}{2}\ln\frac{\left|\frac{1}{2}\left[\Sigma_x+\Sigma_y\right]\right|}{\sqrt{\left|\Sigma_x\right|\left|\Sigma_y\right|}} \quad (1)$$

where $\sum$ is the covariance matrix, $\Sigma=\sigma^2$

Large values of B imply small upper limits on the Bayes error and hence good separability. As a measure of separability, the Bhattacharyya distance has the disadvantage of growing even after the classes have become so well separated that any classification procedure could distinguish them. The Jeffries–Matusita (J-M) distance measures the separability of two classes on a more convenient scale $[0-2]$ in terms of B:

$$J=2\left(1-e^{-B}\right) \quad (2)$$

as B continues to grow, the measure saturates at the value 2. Calculating by estimating the class means and covariance matrices. Table II shows that the RGB image's J-M value is relatively lower than two, which means the RGB separability is sub-optimal. On the other hand, the J-M value of the VNIR image is closest to 2, mean very high separability. These data characteristics will affect the classification accuracy.

TABLE II. J-M DISTANCE OF RGB/VNIR IMAGES

Classes	Miaoyuan Chooffin	Soil	Grass	Garlic	Corn	Ground nut
Miaoyuan Chooffin	0 / 0	0,65 / 2	0,48 / 2	1,05 / 2	0,49 / 2	1,2 / 2
Soil	0,65 / 2	0 / 0	0,24 / 1,99	0,51 / 1,98	0,33 / 2	0,8 / 2
Grass	0,48 / 2	0,24 /	0 / 0	0,52 / 2	0,25 / 2	0,82 / 2
Garlic	1,05 / 2	0,51 /	0,52 / 2	0 / 0	0,34 / 2	1,19 / 2
Corn	0,49 / 2	0,33 / 2	0,25 / 2	0,34 / 2	0 / 0	1,12 / 2
Groundnut	1,2 / 2	0,8 / 2	0,82 / 2	1,19 / 2	1,12 / 2	0 / 0

IV. DISCRIMINANT CLASSIFIER

In this experiment, we use a discriminant classifier because of the simplicity of its algorithm structure—there are two kinds of kernel: linear and quadratic.

A. *Linear Discriminant*

LD is a fast and straightforward statistical method to separate two or multiple classes by fitting the estimation model into the data based on its gaussian distribution. The decision boundaries are placed between the gaussian distribution of different categories as a line[15]. The algorithm also uses the dimensionality reduction method by maximizing the increase in class separability. The sparse feature can make the classification easier due to it being linearly separable[16]. The distribution characteristics can be measure using the covariance and means of all data.

LD can be used as a supervised classification[15], [17] tool for *k* classes. The key steps assume a rule to divide the scattering data into *k* number of areas that belong to each observed data. The Maximum likelihood to allocate the data is used:

$$j=\arg\max\ f_i(x) \quad (3)$$

or Bayesian rules

$$j=\arg\max\ \pi_i f_i(x) \quad (4)$$

From the datasets, the discriminant function for all classes for each observation can be written as:

$$\delta_i(x)=\log f_i(x)+\log\pi_i \quad (5)$$

with the assumption that all classes are distributed with the same covariance matrix, the discriminant for the decision boundary can be written as a linear function:

$$\delta_k(x)=x^T\Sigma^{-1}\mu_k-\frac{1}{2}\mu_k^T\Sigma^{-1}\mu_k+\log\pi \quad (6)$$

B. *Quadratic Discriminat*

For the quadratic analysis (QDA), the decision boundary is nonlinear because the quadratic terms remain, so the discriminant function can be rewritten as follow:

$$\delta_k(x) = -\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x-\mu_k)^T \Sigma_K^{-1}(x-\mu_k) + \log \pi_k$$

(7)

the quadratic type tends more flexibly than a linear one.

V. RESULT AND DISCUSSION

In this paper, the performance classification indicated by three kinds of accuracy used: Overall Accuracy (OA), Average Accuracy (AA), and Cohen's kappa coefficient (κ), consecutively defined by

$$OA = \frac{1}{N}\sum_{i=1}^{r} n_i$$

(8)

$$AA = \frac{1}{k}\sum_{c=1}^{k} A_c$$

(9)

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

(10)

where,

$$P_o = OA \text{ and } P_e = \frac{1}{N^2}\sum_k n_{k1} n_{k2}$$

N: Total sample
n_i: Correct prediction of pixel i
r: Total correct prediction
k: Number of classes
A_c: Accuracy of Class c
n_{ki}: Number of items classified as k by rater i

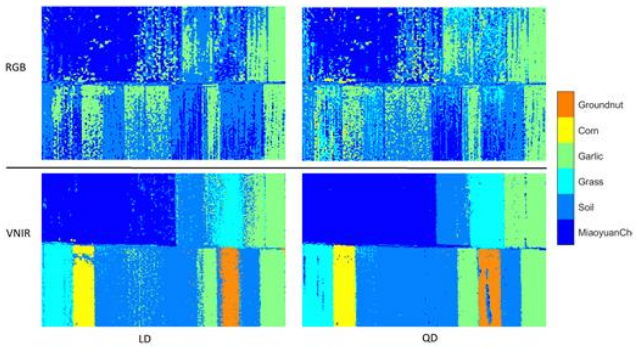


Fig. 4. Map classification results

Classification Mapping is shown in Fig.4. The experimental results show a significantly increased accuracy for both AA, OA, and Kappa on using a 180 band VNIR image compared to a three-band RGB image. As seen in the RGB image, the groundnut and corn classes cannot be detected by the LD at all. Using QD, groundnut is still not detected, and corn has started to be detected by a few pixels. Misclassification occurs in several areas. OA accuracy rate is still shallow at around 50%. The detail of each class's accuracy for each image type with respective methods is shown in Table III.

TABLE III. PER CLASS ACCURACY

Image Type	Classifier	Classes					
		Miaoyuan Choffin	Soil	Grass	Garlic	Corn	Ground nut
RGB	LD	75.88	64.1	0,49	65,3	0	0
	QD	69,16	54,6	27,0	66,3	5,17	0
VNIR	LD	90,32	93,8	79,8	82,7	85,1	87,20
	QD	95,36	97,0	89,4	92,3	94,0	85,58

However, using VNIR image data, simple LD and QD algorithms have detected all existing classes with reasonable accuracy, as shown in Fig.5. With the addition of so many bands, the discrimination function between classes becomes more apparent, increasing the separability, making it easier for the classifier to separate. The level of accuracy is relatively high in the range of numbers, more than 90%.

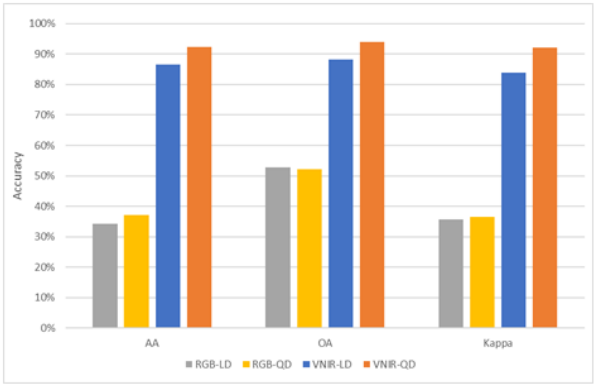


Fig. 5. Performance of RGB vs. VNIR image classification by LD and QD

TABLE IV. METHODS TIME COSTS

Image Type	Classifier	Training Time	Testing Time
RGB	LD	1.64	0.14
	QD	0.77	0.06
VNIR	LD	8.62	2.17
	QD	8.62	2.29

The combination between data type and the methods applied shows in Table IV that discriminant analysis has rapid execution with an average of 8 seconds for training time with a sample size of 36,700 and 2 seconds for testing with a sampling size of 84,777. Even with RGB data is significantly lower than VNIR however, the hyperspectral data time processing is quite acceptable under 10 seconds. Thus it indicates the computation efficiency and rapid performance for classification using discriminant analysis.

VI. CONCLUSION

Hyperspectral VNIR data was very significant in increasing the accuracy of land cover classification compared to RGB images. Abundant features can increase separability between classes. However, high separability does not guarantee high accuracy for classes with a limited number of training samples. The type of kernel/filter selected is also very influential on the classification results. With RGB image, the Quadratic kernel has not increased the accuracy value

compared to the Linear kernel. However, with VNIR data, there was a very significant increase. Meanwhile, training and testing time is still fast. Kernel development in discriminant analysis can further improve classification accuracy while maintaining short training and testing times.

ACKNOWLEDGMENT

The authors would like to acknowledge NSPO Taiwan for their support and funding of this research and TARI for providing ground truth data of our study area.

REFERENCES

[1] S. J. Scherr, S. Shames, and R. Friedman, "From climate-smart agriculture to climate-smart landscapes," pp. 1–15, 2012.

[2] J. Ruan *et al.*, "Agriculture IoT: Emerging Trends, Cooperation Networks, and Outlook," *IEEE Wirel. Commun.*, vol. 26, no. 6, pp. 56–63, 2019, DOI: 10.1109/MWC.001.1900096.

[3] E. Honkavaara *et al.*, "Remote Sensing of 3-D Geometry and Surface Moisture of a Peat Production Area Using Hyperspectral Frame Cameras in Visible to Short-Wave Infrared Spectral Ranges Onboard a Small Unmanned Airborne Vehicle (UAV)," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5440–5454, 2016, DOI: 10.1109/TGRS.2016.2565471.

[4] P. Singh *et al.*, *Hyperspectral remote sensing in precision agriculture: present status, challenges, and future trends*. LTD, 2020.

[5] P. Ghamisi *et al.*, "Advances in Hyperspectral Image and Signal Processing: A Comprehensive Overview of the State of the Art," *IEEE Geosci. Remote Sens. Mag.*, vol. 5, no. 4, pp. 37–78, 2017, DOI: 10.1109/MGRS.2017.2762087.

[6] Z. Sun, D. Wang, and G. Zhong, "A review of crop classification using satellite-based polarimetric SAR imagery," *2018 7th Int. Conf. Agro-Geoinformatics, Agro-Geoinformatics 2018*, pp. 3–7, 2018, DOI: 10.1109/Agro-Geoinformatics.2018.8476020.

[7] P. Aposporis, "Object Detection Methods for Improving UAV Autonomy and Remote Sensing Applications," *Proc. 2020 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2020*, pp. 845–853, 2020, DOI: 10.1109/ASONAM49781.2020.9381377.

[8] G. Taskin, H. Kaya, and L. Bruzzone, "Feature selection based on high dimensional model representation for hyperspectral images," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2918–2928, 2017, DOI: 10.1109/TIP.2017.2687128.

[9] H. Yuan, Y. Y. Tang, Y. Lu, L. Yang, and H. Luo, "Spectral-spatial classification of hyperspectral image based on discriminant analysis," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2035–2043, 2014, DOI: 10.1109/JSTARS.2013.2290316.

[10] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3. pp. 247–259, May 2011, DOI: 10.1016/j.isprsjprs.2010.11.001.

[11] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, no. March, pp. 166–177, 2019, DOI: 10.1016/j.isprsjprs.2019.04.015.

[12] S. Sattar, H. Ahmad Khan, and K. Khurshid, "Optimized class-separability in hyperspectral images," *Int. Geosci. Remote Sens. Symp.*, vol. 2016-November, pp. 2711–2714, 2016, DOI: 10.1109/IGARSS.2016.7729700.

[13] M. Dabboor, S. Howell, M. Shokr, and J. Yackel, "The Jeffries–Matusita distance for the case of complex Wishart distribution as a separability criterion for fully polarimetric SAR data," *Int. J. Remote Sens.*, vol. 35, no. 19, pp. 6859–6873, 2014, DOI: 10.1080/01431161.2014.960614.

[14] P. Kumar, R. Prasad, A. Choudhary, V. N. Mishra, D. K. Gupta, and P. K. Srivastava, "A statistical significance of differences in classification accuracy of crop types using different classification algorithms," *Geocarto Int.*, 2017, DOI: 10.1080/10106049.2015.1132483.

[15] B. Ghojogh and M. Crowley, "Linear and Quadratic Discriminant Analysis: Tutorial," *arXiv*, no. 4. pp. 1–16, 2019, [Online]. Available: https://arxiv.org/pdf/1906.02590.pdf.

[16] C. Xia, S. Yang, M. Huang, Q. Zhu, Y. Guo, and J. Qin, "Maize seed classification using hyperspectral image coupled with multi-linear discriminant analysis," *Infrared Phys. Technol.*, vol. 103, no. July, p. 103077, 2019, DOI: 10.1016/j.infrared.2019.103077.

[17] Q. Du and R. Nekovei, "Implementation of real-time constrained linear discriminant analysis to remote sensing image classification," *Pattern Recognit.*, vol. 38, no. 4, pp. 459–471, 2005, DOI: 10.1016/j.patcog.2004.09.008.

Review on Protection Against Transients and Surges in Power System

1st Norliza
Jabatan Kejuruteraan Elektrik
Politeknik Mukah
Sarawak, Malaysia
norliza.razak@pmu.edu.my

2nd Hamadi
Jabatan Kejuruteraan Elektrik
Politeknik Mukah
Sarawak, Malaysia
hamadi@pmu.edu.my

***Abstract*—The most essential issue for electricity providers to address is power distribution breakdowns. A surge is a voltage or current wave that occurs suddenly. It happened within the power system in a very brief period of time, usually a few milliseconds. Furthermore, a surge in the electrical system can be caused by lightning, utility switching, including capacitor switching, equipment switching, and switching inductive loads within a facility. Surge protection systems are widely used around the world. The term "surge protection device" (SPD) refers to devices that protect electronics from surges and transients. We use the term SPD to refer to devices that safeguard systems and equipment from electrical transients in this study. The method for selecting SPD to protect a power system must adhere to the IEC 62305 guideline. The purpose of this article, from the standpoint of the power system, is to review the protection against transient and surge using appropriate methods and adhering to all standards in order to satisfy the requirements.**

Keywords—Lightning Protection, Surge Protective Device(SPD), Standard.

I. INTRODUCTION

Customers are directly affected by power supply disruptions since power system distribution equipment is positioned on the terminal side of power networks. The most essential issue for electricity providers to address is power distribution breakdowns. Lightning is a common cause of 6.6-kV high-voltage overhead distribution line failures [1].

A surge is a voltage or current wave that occurs suddenly. It happened within the power system in a very brief period of time, usually a few milliseconds. Furthermore, a surge in the power system is caused by lightning, utility switching, including capacitor switching, equipment switching, and switching inductive loads within a facility [2].

Surge protection systems are widely utilized around the world. Furthermore, each country's protective system has a different name. Surge Protective Device (SPD) is the general name for devices that protect equipment from surges and transients. It is mostly used in Europe. Meanwhile, it is known as Transient Voltage Suppressors in the United States and several other nations (TVS). TVS is a word used in some cultures to refer to a transient protector constructed of a specific component (Zen er Diode). The devices that protect systems and equipment from electrical transients are referred to as SPDs in this paper [3].

The method for selecting SPD to protect a power system must adhere to the IEC 62305 guideline. In December 2010, the IEC 62305 Ed 2 Lightning Protection series of publications was issued as an International Standard. A new instructive Annex D is added, which contains information on factors to consider while choosing SPD [4-5].

The purpose of this article, from the standpoint of the power system, is to review the protection against transient and surge using appropriate methods and adhering to all standards in order to satisfy the requirements.

II. SOURCES OF SURGES

A surge is a voltage or current wave that occurs suddenly. It happened within the power system in a very brief period of time, usually a few milliseconds. Furthermore, there is a surge in the power system as a result of: (a) Lightning strike (b) Utility switching, such as capacitor switching, equipment switching, and switching inductive loads in the facility are all examples of utility switching. J. In his paper, A. Milke discusses the effects of these various surge sources, which is presented in Table I [10].

TABLE I. SOURCES OF SURGES [10]

Source of surge	Peak voltage magnitude	Frequency of occurrence	Comments
Lightning	ranging from 1,000 to >40,000 V, with a mean of around 20,000 V	Depending on the location, once a week to once in a lifetime	The size of the stroke is determined by its closeness to the facility and its electrical system coupling. Flashover is unlikely to cause a facility's voltages to exceed 6,000 volts

System switching and the utility capacitor	Up to 1,300 V on a 480 V system	Never more than once a day, depending on utility.	Capacitors may or may not be installed in the area.
Facility equipment switching	A 480 V system can deliver up to 2,000 V.	Several times a day	When compared to lightning-induced transients, the magnitude is minor, yet switching might happen often.

A. Lightning

The most prevalent origin of external surge, which generates a voltage surge in transmission lines, is lightning from the atmosphere. Because of dispersed line conductance and stray capacitance, this lightning surge travels to both the source and load sides of the transmission line. The surge is travelling at light speed. The surge-impedance changes at the transmission line's ends, and the wave is reflected back. The surge wave continues to go back and forth until the energy of the surge is absorbed by line resistance. As a result, the line's voltage rises to several times its rated voltage [6]. The majority of non-programmed electric system disruptions are caused by lightning strikes. According to international statistics, lightning strikes that strike overhead power lines cause roughly 65 percent of line disruptions. Because the amount of power carried by transmission lines is increasing, electric systems are being pushed to their limitations, and unplanned line disruptions are increasing the danger of instability [7].

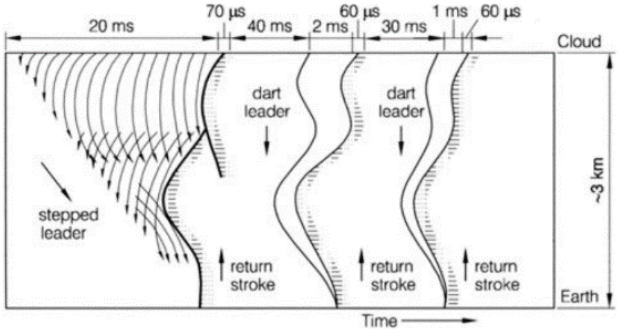


Fig. 1. Mechanism of a lightning discharge to earth [7].

The lightning discharge is caused by a streamer from the cloud that flows in a sequence of steps toward the earth (see fig. 1). It's referred to as a stepped leader. The steps are separated by a time interval of 40 to 100 j.LS between them. This leader divides into branches as it proceeds, increasing the luminance of the ionised way [7].

In general, the behaviour of lightning is currently fairly predictable, although precise knowledge of specific instances is not. There are two types of lightning protection available: (a) direct impacts related to the lightning current's energy, heating, flash, and ignition, and (b) indirect effects related to produce over voltages in surrounding electrical and electronic systems [8].

When compared to the distance stroke, the induce voltage caused by lightning can have a significant impact on electrical components. Figure 2 demonstrates that these induced voltages can be rather significant, which explains why equipment can be destroyed by lightning flashes that are relatively far away [9].

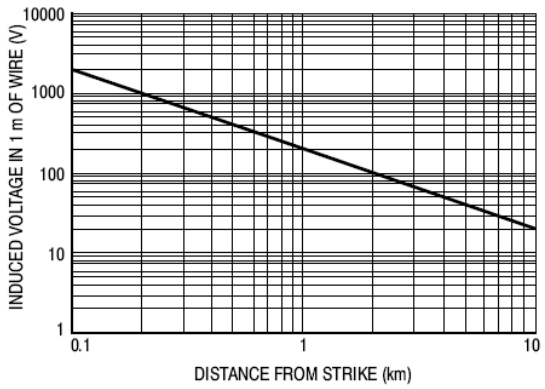


Fig. 2. Voltage Induced by Nearby Lightning Strike [9].

Furthermore, a lightning-induced surge can attack a power system in a relatively short period of time, but the damage is significant due to the impulsive transient's high amplitude. Microseconds or milliseconds are widely used to measure the duration of a lightning strike. However, serious system damage might occur during this brief window. A typical lightning surge current is depicted in Fig. 3. In the relationship between time and current for lightning, we discovered that when time is short, the quantity of current increases dramatically, and when time is longer, current decrees [10].

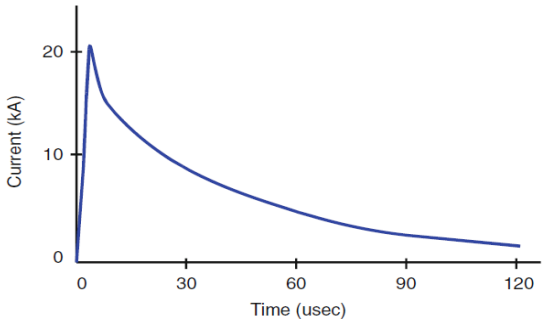


Fig. 3. Typical lightning surge current [10].

B. Utility Switching

Utility switching refers to the use of a capacitor as a switching device in general. It also changes the utility configurations. For a small amount of time, each switching operation can cause a transient that exceeds the equipment voltage specifications. Utility switching produces a momentary transient with the same amplitude as a lightning strike, but it can cause long-term damage to electrical equipment.

When the voltage across a capacitor is 0 before it is switched into the circuit, a process surge voltage occurs. When a capacitor is switched, the system voltage is applied to the capacitor's zero voltage, causing a brief short circuit across the capacitor. The bus voltage drops to 0 volts for a brief duration at the capacitor site. After the initial step change, the voltage rebounds and then overshoots as the system returns to its steady state value. The system then oscillates until the voltage is dampened to its steady-state value. The peak occurs during the initial oscillation period [10].

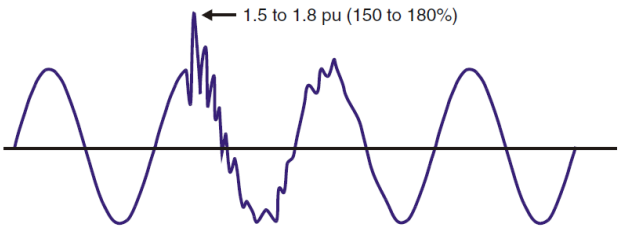


Fig. 4. Voltage waveform for capacitor switching transient [10].

A capacitor switching transient waveform is seen in Fig. 4. Switching-induced oscillations can cause capacitors within a facility to resonate, magnifying the peak voltage and lengthening the time it takes for the voltage to recover to normal [10].

C. Facility Internal Switching

An inductive discharge of energy in switched equipment created the voltage surge. Even little system changes, such as lighting loads, might result in a massive inductive surge. This type of switching is responsible for the vast majority of switching transients. In contrast to lightning-induced surges, the size of this sort of surge is far smaller [10].

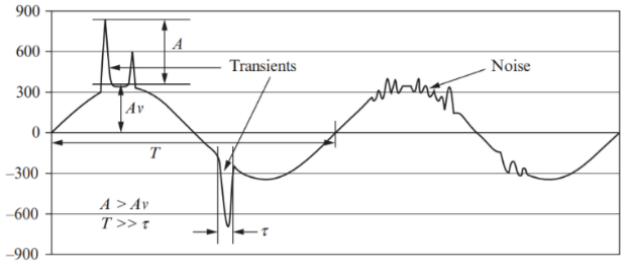


Fig. 5. Transients and noise superimposed on the waveform [10].

As shown in Fig. 5, voltage transients spikes in the supply waveform due to high inductive loads such as air conditioners, electric power tools, machinery, and elevators are caused by on and off switching. In general, the more energy transferred into the circuit by the transient surge, the greater the risk of damage [16].

III. DAMAGES DUE TO SURGES

Electrical surges that occur in electrical systems can be dangerous. Either to a system or a person. The most serious injuries caused by lightning surges conveyed by incoming lines are injuries caused by stepping or touching voltages, harmful sparks causing fire, and damage to the structure's interior electrical and electronic systems [13].

Surges created by flashes on the incoming line are the most common cause of damage, whereas surges induced on the line are the most common cause of internal system failures. Surge risk R should be compared to a tolerable value to identify situations where preventative measures should be implemented. When RT and protective measures are required, they should be given.

$$R = F \times L > R_T \quad (1)$$

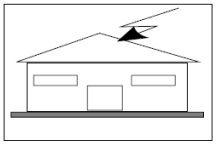
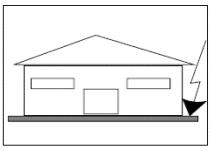
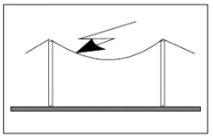
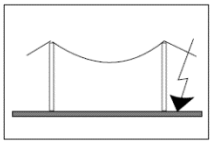
Where:

- $F = N \times P$ the number of times a year that a damage occurs;
- N the annual number of harmful events (i.e., the number of surges of a certain amplitude);
- P the likelihood that a harmful event may cause damage to the structure being served;
- L related to the worth (people and products) of the structure to be protected, the amount of loss caused by a certain form of harm (injury to human life or failure of electrical and electronic equipment) caused by a dangerous occurrence.

Protection against lightning (IEC 62305-2) Depending on the features, a lightning flash could cause structural damage, according to the report. Type of structure, contents and application, type of service, and protection measures supplied are some of the most important characteristics [14].

Table II shows the sources of harm, types of damage, and types of loss based on the point of impact.

TABLE II. SOURCES OF DAMAGE, TYPES OF DAMAGE AND TYPES OF LOSS ACCORDING TO THE POINT OF STRIKE [14]

Lightning flash	Source of damage	Type of damage	Type of loss
	S1	D1 D2 D3	L1, L4 L1, L2, L3, L4 L1, L2, L4
	S2	D3	L1, L2, L4
	S3	D1 D2 D3	L1, L4 L1, L2, L3, L4 L1, L2, L4
	S4	D3	L1, L2, L4

Where:

S1: A structure appears in a flash.
S2: Near a structure, a flash occurs.
S3: flashes to a single line
S4: flashes in the vicinity of a line
D1: electric shock causes harm to biological creatures,
D2: physical harm,
D3: electrical and electronic system failure
L1: a human life has been lost (including lasting harm);
L2: a reduction in public service;
L3: cultural heritage is being lost;
L4: economic value loss

Table II show that, damage cause by lightning strike is different according to area hit. Direct connection to appliance or building cause more damage than nearby hit.

IV. METHOD OF PROTECTION

Many devices have been developed to deal with electrical surges. Its design is determined by its intended use. Surge Protective Device (SPD) is the popular name for a device that protects against electrical surges [12].

A. *Protection Circuit Model of SPD*

Figure 6 depicts a typical SPD protection circuit for a low-voltage distribution system, where Ig represents incoming surge current and ZL represents load impedance [15].

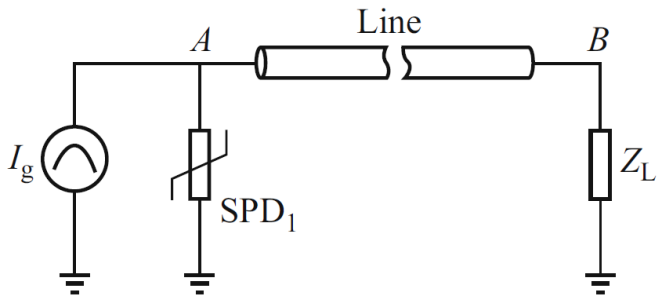


Fig. 6. Protection circuit of SPD [15].

The residual voltage of SPD1 produces a flat headed voltage wave when it is turned on. Furthermore, because SPD1's turned-on resistance Rs is significantly lower than the connecting line's wave impedance and the load impedance, the residual voltage of SPD1 can be employed in the protection circuit in series with the turned-on resistance Rs [15].

B. *Types Of SPD Systems*

There are two types of SPD systems for lightning protection technology: MOV and TSG.

MOV is a voltage-dependent non-linear resistor with metal oxide varistor plates placed between two metal plates that function as electrodes as a lightning dissipation element. MOVs feature a high nonlinear coefficient, low leakage current, excellent lightning dissipation, and a fast response time. As a result, MOV at the structure's entrance point is adequate for structures in metropolitan regions with an induced voltage tolerance of 8/20µs created by indirect lightning strikes.

The TSG device is made by Gap Technology. A spark is created between the spaces in the spark gap, which is how this device works. TSG offers benefits such as high lightning current dissipation, overvoltage impulse impact, and the flexibility to function in a variety of electrical systems. When TSG technology is employed for overvoltage protection, the surge energy dissipation performance improves. As a result, TSG is suitable for protecting structures in suburban regions that are susceptible to direct lightning strikes, lightning currents of 10/350µs, and huge lightning current amplitudes [11].

MOV is more effective than TSG at protecting against overvoltage, according to a study on residual voltage across the load, as illustrated in Fig. 7 below [11].

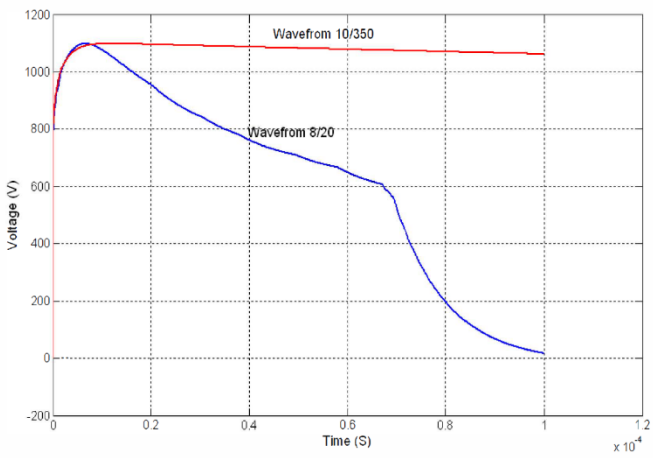


Fig. 7. Residual voltage across the load, SPD type MOV and TSG respectively [11].

C. *Levels of Surge Protection*

SPD protection is employed differently depending on what it is protecting. The level varies depending on the application, but in general, protection levels are categorised into three categories: main protection, secondary protection, and board-level protection [18].

Power and data lines exposed to the outdoors, service entry, and AC distribution panels are all subject to primary protection. Equipment inputs, such as electricity from long branch circuits, internal data lines, PBX, wall sockets, and lines with primary protection at a substantial distance from the equipment, require secondary protection. Internal to the equipment, board-level protection is used to defend against residual transients from earlier stages of protection, system induced transients, and Electro Static Discharge (ESD).

V. STANDARD

The performance and safety tests of the SPD employed are part of the standard for surge prevention. To overcome surge energy or excessive overvoltage, the protection mechanism employed must be trustworthy [19].

Electrical equipment used to be certified by nationally recognised testing facilities in several countries. The majority of these national standards are based on the Worldwide Electrotechnical Commission's (IEC) international standards (IEC). Meanwhile, standards have been produced by the American National Standards Institute (ANSI) and the Institution of Electrical and Electronic Engineers (IEEE) in North America [17].

TABLE III. STANDARD TEST AGENCIES AND THEIR RECOMMENDED STANDARDS FOR SPDS [17]

Standard Agency	SPD - Related Standards	Remark
IEC[a]	IEC 61000-4-5	Create a model to replicate surges so that you can see if your equipment can withstand them.
	IEC 60364 IEC 61643-11 IEC 61643-12n	General guidelines for product selection and use

UL	UL 1449 UL 1449 3rd Ed	The most important consideration is safety.
ANSI/IEEE[b]	IEEE C62.41.1	Low-voltage networks are at risk of transient over voltages.
	IEEE C62.41.2 IEEE C62.45 IEEE C62.62 IEEE C62.72	Types of transients and the environment during a surge Equipment linked to the low voltage network is tested against transients using this method. SPD performance is compared using tests and ratings for SPDs installed on the load side of service equipment. SPDs are installed in the electrical power distribution equipment.
NEC	Article 280 Article 285	Surge arrestors are used on wiring systems that are over 1000 volts. SPDs are chosen and installed under certain parameters.
NEMA[c]	LS1 Low-voltage surge protective devices	Specific SPD performance is provided. This information can be used to compare actual SPD test results.

Table III shows a number of standards for standard test agencies conducting a variety of SPD tests. Depending on how SPDs are used, some bodies advocated different regulatory standards.

National and international standards, according to the Phonex Contact 2016 guidebook, are responsible for providing a guide to constructing a lightning and surge protection concept as well as the design of specific protective devices [21].

IEC 62305 [1] [2] [3] [4]: Protection against Lightning is the most important standard in surge protection. According to Brandon, the European TC 81 Technical committee of the IEC established the IEC 62305 lightning protection standard, which is made up of four parts [20]:

- IEC 62305-1 : Protection Against Lightning Part 1: General Principles
- IEC 62305-2 : Protection Against Lightning Part 2: Risk Management
- IEC 62305-3 : Protection Against Lightning Part 3: Physical Damage to Structures and Life Hazard
- IEC 62305-4 : Protection Against Lightning Part 4: Electrical and Electronic Systems within Structures

IEC 60364-4-44: Protection for safety – Protection against voltage disturbances and electromagnetic disturbances in switching operations is another common protection standard [22].

VI. CONCLUSION

The power system must be protected against transients and surges in order to protect electrical appliances. Electrical surges were caused by lightning, utility switching, such as capacitor switching, and equipment switching inductive loads.

Surge damage is extremely dangerous to both people and the environment. The damage caused by lightning surge varies depending on where the strike occurred.

Surges Protective Devices must be used to prevent electric surges from causing damage to electrical appliances (SPD). IEC 62305 specifies how SPD should be used in electrical appliances. Lightning protection and IEC 60364-4-44: Protection against voltage and electromagnetic disturbances in switching operations.

REFERENCES

[1] S. MATSUURAJ *et al*., "Lightning Surge Characteristics of an Actual Distribution Line and Validation of a Distribution Line Model for Lightning Overvoltage Studies" *Electrical Engineering in Japan*, Vol. 173, No. 1, 2010.

[2] I. Daut *et al., "*The Effects of Harmonic Components on Transformer Losses of Sinusoidal Source Supplying Non-Linear Loads" *Am. J. Applied Sci.*, 3 (12): 2131-2133, 2006

[3] Suresh Mikkili and Anup Kumar Panda, "Power Quality Issues and Solutions – Review" *Int. J. Emerg. Electr. Power Syst*. 16(4): 357–384J, 2015.

[4] A. Kalaira *et al., "*Review of harmonic analysis, modeling and mitigation techniques*" Renewable and Sustainable Energy Reviews 78: 1152–1187, 2017*

[5] A. Kannan1 *et al*., "A Review of Power Quality Standards, Electrical software tools, Issues and Solutions" *ICRESE'13,* 2013.

[6] *Heavy Electricals Limited: Handbook of Switchgears. SURGE PROTECTION,* McGraw-Hill Global Education Holdings , New York,2007, Chapter 9.

[7] *Overhead Power Lines, Planning, Design, Construction*, Springer, New York,2003, pp. 99-113.

[8] Francois D. Martzloff, "*TRANSIENT OVERVOLTAGE PROTECTION SEMINAR"* NBS Forum, 2014.

[9] *Transient Overvoltage Protection*, Semiconductor Components Industries, LLC, 2008.

[10] *Eddie Davis, Nick Kooiman and Kylash Viswanathan, "Data Assessment for Electrical Surge Protective Devices", New York, USA,* SpringerBriefs in Fire, 2015, pp 1-36.

[11] Le Quang Trung, Quyen Huy Anh and Vu Phan Tu, "Surge Protection on the Power Lines, Considering the Influenced Factors*", IEEE, ICSSE 2017.*

[12] Chandima Gomes and Vernon Cooray, "SURGE PROTECTION OF LOW VOLTAGE SYSTEMS AND RELEVANT STANDARDS", *ICLP September 13th -17th, 2010.*

[13] F. Fiamingo, C. Mazzetti and G.B. Lo Piparo, "QUALITY OF THE SUPPLIED ELECTRIC SERVICE: A TOOL TO EVALUATE THE NEED OF PROTECTION AGAINST LIGHTNING SURGES", *ICLP September 13th -17th, 2010.*

[14] IEC INTERNATIONAL STANDARD for Protection against lightning –Part 2: Risk management, IEC 62305-2, 2010.

[15] D. ZHANG *et al.* "Protection Distance of Surge Protective Devices in Low Voltage Distribution Systems", *J. Shanghai Jiaotong Univ.* (Sci.), 13(4): 477–481, 2008.

[16] N. Kularatna *et al.* "Background to Surge Protection", *Design Of Transient Protection Systems,* Elsevier Science, 2019, ch 3, pp 1-15.

[17] N. Kularatna *et al.* "Background to Surge Protection", *Design Of Transient Protection Systems,* Elsevier Science, 2019, ch 2, pp 17-27.

[18] N. Kularatna *et al.* "Background to Surge Protection", *Design Of Transient Protection Systems,* Elsevier Science, 2019, ch 10, pp 175-198.

[19] A. ROUSSEAU *et al. "*SURGE PROTECTION"J. Webster (ed.), John Wiley & Sons, Inc. 1999, pp 153-165.

[20] T. Brandon, “Adoption of IEC 62305 as the Basis for One Major U.S. Electric Utility’s Lightning Protection Standard”, *25th International Lightning Detection Conference,* pp 1-6, March 2018.

[21] *The basics of surge protection*, PHOENIX CONTACT GmbH & Co. KG, Germany, 2016, pp 1-60.

[22] IEC INTERNATIONAL STANDARD for Protection for safety – Protection against voltage disturbances and electromagnetic disturbances, IEC 60364-4-44, 2003.

Comparison of Global and Regional Tidal Models at Sekupang Tidal Station

1st Siti Noor Chayati
Geomatics Engineering, Dept of Informatic Engineering
Politeknik Negeri Batam
Batam, Indonesia
sitinoorchayati@polibatam.ac.id

2nd Sefi Nuraeni
Geomatics Engineering, Dept of Informatic Engineering
Politeknik Negeri Batam
Batam, Indonesia
sefinur.ani@gmail.com

***Abstract*—Indonesia is an archipelagic country with a total marine area of 5.9 million km², consisting of 3.2 million km² of territorial waters and 2.7 km² of Exclusive Economic Zone waters, not including the continental shelf. With the vast waters in Indonesia, sufficient information about the tides is needed. However, in reality the number of tidal stations targeted by the Geospatial Information Agency (BIG), which is as many as 400 stations, has not been fulfilled. The data used in this study include tidal data at Sekupang station, global model data TPXO9 and regional model data from BIG. The components studied are the main components M2, K2, O1, and S1 using the Least Squares Adjustment method. The minimum discrepancy value is shown by the TPXO9 model, which means that the TPXO9 global model is a more suitable model to be used as a tidal prediction model in the waters of Batam Island.**

Keywords—tides, TPXO9, regional model, harmonic constant

I. INTRODUCTION

Indonesia is the largest archipelagic country in the world. The total area of Indonesia's sea area is 5.9 million km², consisting of 3.2 million km² of territorial waters and 2.7 km² of waters of the Exclusive Economic Zone, not including the continental shelf. Judging from the total area of the area, most of the development will be carried out in coastal areas [1]. Tidal data contains information regarding the highest and lowest seawater estimates that can be used as a reference for the development of coastal areas and the construction of piers and ports. Therefore, sufficient tide stations are needed to provide tidal information throughout Indonesian waters. Indonesia until 2018 only had 139 tidal stations from a target of 400 stations under the coordination of the Geospatial Information Agency [2]. Due to the inadequacy of tidal data from tidal stations, it is necessary to use another alternative, namely by using a tidal prediction model to meet the need for tidal information. Some of the tide prediction models that are often used are the global TPXO prediction model and the regional model by BIG.

Tides are rhythmic fluctuations (movements up and down) of sea level due to the attraction of objects in the sky, especially the moon and sun, to the mass of sea water on earth [3]. In the earth-moon system, the tidal generating forces are caused by the resultant forces that cause the tides, namely, the centrifugal force of the earth-moon system (Fs) and the moon's gravitational force (Fb) [4] . The characteristics of the tides in some areas may vary. This is not only influenced by the attraction of the moon, but also by the morphology of the seabed, the shape of the coastline, and the characteristics of the waters themselves. In an area of water there can be two or one tides. [5] The types of tides in a water area can be categorized into four types, including:

- Semi-diurnal tide
- Diurnal tide
- Mixed prevailing diurnal tide
- Mixed tide prevailing semi diurnal

Variations in sea level at a particular location are expressed as a result of the superposition of various waves of tidal harmonic constants. Determination of the value of the change in amplitude and phase delay due to the attractive force of celestial bodies against the earth's equilibrium condition will later be expressed in a constant. These constants are referred to as harmonic components. [6] The dominant tidal constituents are the semidiurnal constituents M2, S2, N2, and S2, with periods of 12.42, 12.00, 12.66, and 11.97 h, respectively, and the diurnal constituents, K1, O1, P1, Q1, and S1, with periods of 23.93, 25.82, 24.07, 26.87, and 24.00 h, respectively.

Least Square Method is a formula used to get the best approximation value with minimal error. The goal is to use the least squares method so that the approximate value obtained is as close as possible to the actual situation in the field. In the Least Square method, suitability with field data is defined by the situation, where the square integral of the value of the difference in water level elevation results from calculations and measurements and minimal measurements (Least of square of error) [7]. The principle of this method is to minimize the difference between the composite signal and the size signal. [8] If *ht* is the water level observation data at time *t* and *h(t)* is the predicted water level, in the Least Squares method, the square of the difference between the observations and the model must be minimal. Therefore,

$$\sum_{k=1}^{n}[h_t - h(t)]^2 = minimal \quad (1)$$

The global tide model provides tidal modeling in all seas on he earth's surface, both in wide oceans and in coastal oceans which is built using altimetry satellite data for a certain period and several tide stations in the field. The TPXO-Atlas incorporates all local models except the Mediterranean Sea plus the Baltic Sea model. The Atlas model will fit coastal tidal stations significantly than the base model however still worse than the local model due to the smaller resolution [9].

In addition to the global tidal model, there is a regional tidal model in the territory of Indonesia, namely the tidal model made by BIG. The harmonic constants provided include K1, K2, M2, N2, O1, P1, Q1, and S2. The regional tidal model was created for a particular water area. The accuracy is increased in these water areas when compared to the global tidal model. The regional tidal model was obtained from several altimetry data and other datasets such as assimilated observational data at tidal stations.

II. METHODOLOGY

A. Tidal Data

In this study, 3 tidal data were used, namely tidal data for the Sekupang tidal station, regional model data by BIG, and TPXO9.1 global model data. Tidal data at the Sekupang station used is hourly observation data. The BIG regional model data is downloaded from the http://tides.big.go.id/pasut/konstanta/ page. The harmonic constants provided include the constants K1, K2, M2, N2, O1, P1, and S1. This data file has *.nc format (NetCDF.File) which contains Amplitude and Phase data. The TPXO9 model data is downloaded from the official website of OSU (Oregan State University) Tidal Data Inversion at http://volkov.oce.orst.edu/tides/. There are 4 data obtained in the extracted file, namely grid_tpxo9, h_tpxo9, u_tpxo9, and data model_tpxo9.

B. Research Method

The design activities are presented in Fig 1. In general, the research stages include tidal data collection, tidal data processing by handling blank data, data quality control, tidal harmonic analysis, and calculating RMS, RSS, RSSIQ, and D values.

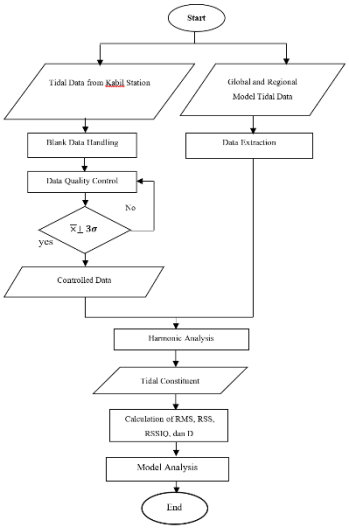


Fig. 1. Research design

C. Data Quality Control

After checking and handling tidal data, then the data will be quality controlled to see the quality of tidal data, by knowing the outlier or spike data, i.e. tidal data that is strange or out of most data, and offset data, i.e. tidal data that has a different height reference in the observation range. the same one. The tidal data quality control can be done graphically and numerically.

Numerical quality control is carried out by calculating the standard deviation/standard deviation to determine the value of the tidal error. The global test using a confidence level of 99.7% or 3σ is used to check tidal data.

Meanwhile, graphical quality control is also carried out to check the tidal data whether there are spikes or outliers. Fig 2 shows the results of data quality control graphically.

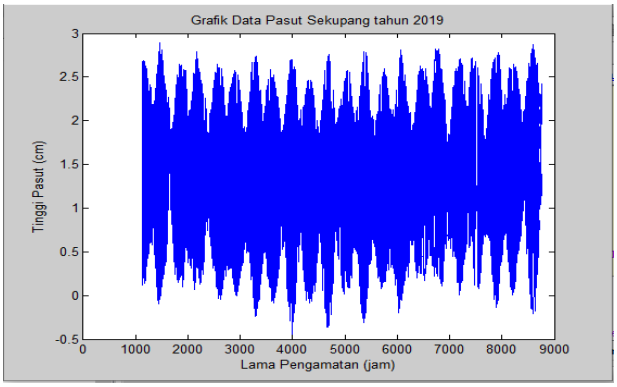


Fig. 2. Tidal data after quality control

D. Harmonic Analysis

Harmonic analysis was carried out to obtain the harmonic components present in the water level signal, in the form of amplitude and phase. In this research, we calculated the constituent of M2, S2, K1, and O1.

The analytical equation with the least squares count can be seen in equation 2[8].

$$h(t) + v(t_n) = hm + \sum_{i=1}^{k} A_i \cos(\omega_i t - g_i) \quad (2)$$

From equation 2, where:

$$A_i \cos g_i = A_r \; dan \; A_i \sin g_i = B_r \quad (3)$$

Therefore,

$$h(t) + v(t_n) = hm + \sum_{i=1}^{k} A_r \cos \omega_i t + \sum_{i=1}^{k} B_r \sin \omega_i t \quad (4)$$

where and is the i-th harmonic constant, k is the tidal component and is the hourly observation time.

The magnitude of the mean water level calculated by equation 1 is close to the observed tidal elevation as a function of time if it meets the requirements of the law of least squares, namely the sum of the squares of the minimum residue. This condition is then derived against A_r and B_r. Based on the least squares method, the completion of the harmonic analysis of the least squares method can be described as follows [8]:

- Equation of sea level observation L=AX

- The correction equation V=AX-L, shown in equation 5.

$$v(t_n) = hm + \sum_{i=1}^{k} A_r \cos \omega_i t + \sum_{i=1}^{k} B_r \sin \omega_i t \quad (5)$$

The amplitude and phase of the tidal components are determined by equations 6 and 7.

$$A_i = \sqrt{Ar_i} + Br_i \quad (6)$$

$$g_i = \frac{Ar_i}{Br_i} \quad (7)$$

The design of the tidal observation matrix is determined in equations 8, 9, 10 and 11.

$$_nA_k = \begin{bmatrix} 1 \cos \omega_1 t_1 \sin \omega_2 t_1 \dots \dots & \cos \omega_k t_1 \sin \omega_1 t_1 \dots . \sin \omega_k t_1 \\ 1 \cos \omega_1 t_1 \sin \omega_2 t_1 \dots \dots & \cos \omega_k t_1 \sin \omega_1 t_1 \dots . \sin \omega_k t_1 \\ \vdots & \vdots \\ 1 \cos \omega_1 t_n \sin \omega_2 t_n \dots \dots & \cos \omega_k t_n \sin \omega_1 t_n \dots . \sin \omega_k t \end{bmatrix} \quad (8)$$

$$\mathrm{L} = \begin{matrix} h_1 \\ \vdots \\ h_n \end{matrix} \quad (9)$$

$$\mathrm{X} = (A^T PA)^{-1}(A^T PL) \quad (10)$$

$$_kX_i = \begin{bmatrix} h_0 \\ A_i \\ \vdots \\ A_k \\ B_1 \\ \vdots \\ B_k \end{bmatrix} \quad (11)$$

Where,

L : sea level data

A : coefficient matrix

V : correction value

A_r : parameter A of the tidal-forming component

Br : parameter B of the tidal-forming component

Ω : angular velocity of harmonic wave

t : observation time

E. Tidal Model Evaluation

The evaluation of the tidal model was carried out to determine the value of D in each tidal model. The smallest D value is the best value, which means that the model is the most suitable for use in these waters. To calculate the RMS (Root Mean Squares) value between each tidal harmonic constant from the station data and the model data, equation (12)[10]:

$$\sqrt{\frac{1}{2N}} \sum_{i=1}^{N} \left\{ \left[h_1^{sol}(i,j) - h_1^{ref}(i,j) \right]^2 + \left[h_2^{sol}(i,j) - h_2^{ref}(i,j) \right]^2 \right\} \quad (12)$$

The RSS (Root Sum of Squares) value is the value of all effects of the tidal harmonic constant for the model with tidal stations. Using equation (13)[10].

$$\mathrm{RSS} = \sqrt{\sum_{j=1}^{n} RMS \frac{2}{j}} \quad (13)$$

The RSSIQ value is an estimate of all errors (errors) between the model and the tidal station obtained previously from the RSS value. Using equation (14) and equation (15) for D value [10]:

$$\mathrm{RSSIQ} = \sqrt{\frac{i}{2N}} \sum_{j=1}^{N} \sum_{i=1}^{N} \left\{ \left(h_1^{ref}(i,j)^2 + h_2^{ref}(i,j)^2 \right) \right\} \quad (14)$$

$$\mathrm{D} = \frac{RSS}{RSSIQ} \times 100\% \quad (15)$$

III. RESULT AND DISCUSSION

A. Data Quality Control

Based on the results of checking the tidal station data in Table I, it is known that the amount of data rejected by the Sekupang tidal station is 14.46% while the percentage of data accepted is 85.56%.

TABLE I. TIDAL DATA AT SEKUPANG STATION

STDEV	0,11291923
Total Data	8761
Accepted Data	7495 (85,55%)
Rejected Data	1266 (14,46%)

B. Tidal Constituent

From the data in Table II, it is known that the amplitude value for each tidal component has a significant status. The highest amplitude value is found in the tidal constant M2, it indicates that the area is dominated by the main semidiurnal tide which is more dominant than the main diurnal harmonic constant value.

TABLE II. MAIN TIDAL CONSTITUENT FROM SEKUPANG TIDAL STATION

Constituent	Amplitude (cm)	Status
O1	0,2269	Significant
K1	0,2775	Significant
M2	0,7929	Significant
S2	0,3261	Significant

In Table III, each model provides a different value for the harmonic constant even at the same point location. This difference is because the data used in the formation of the model is different. As in the BIG model, this model is made from the TPXO model which is assimilated and validated using tidal observation data. The difference in constant values between models is only in the millimetre to centimetre fraction so that it does not provide a large difference in the range of amplitude values or when used for correction.

TABLE III. MAIN TIDAL CONSTUTUENT FROM TIDAL MODEL DATA

Constituent	Amplitude of Tidal Model Data (cm)	
	Global Model TPXO9	***Regional Model BIG***
O1	0,1852	0,2578
K1	0,1643	0,2649

Constituent	Amplitude of Tidal Model Data (cm)	
	Global Model TPXO9	*Regional Model BIG*
M2	0,4942	0,5131
S2	0,2158	0,1855

C. Model Evaluation

TABLE IV. TIDAL DATA AT SEKUPANG STATION

Regional Model				
Constituent	***RMS***	***RSS***	***RSSIQ***	***D***
M2	0,378	0,420	0,665	63,137
S2	0,043			
K1	0,167			
O1	0,059			
Global Model TPX09				
M2	0,299	0,348	0,665	52,364
S2	0,110			
K1	0,113			
O1	0,084			

In Table IV and Fig 3, it is known that the highest constant value in the two tidal models, namely the M2 constant in the regional model is at a value of 0,378 while for the global model TPXO9 is 0,299. The same thing happened to KI value, with 0,167 and 0,113 for regional model dan global model respectively. On the other hand, RMS value for S2 and O1 of regional model has lower value than the global model TPX09.

The TPXO9 global prediction model has a resolution accuracy of 1/30°×1/6° which is assimilated data. This also affects the resulting elevation results. The minimum error value shows the tidal prediction model has a small error rate and a high level of accuracy.

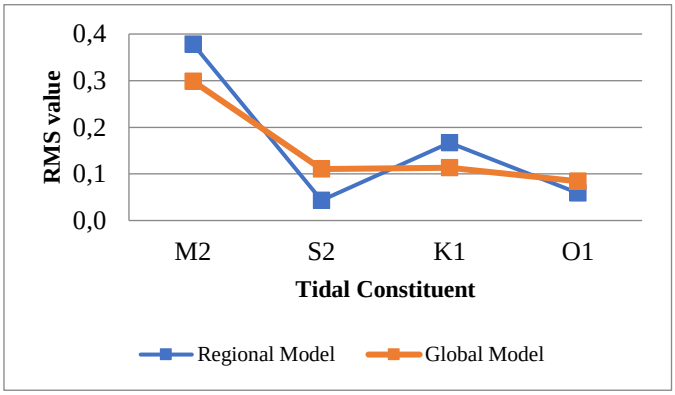


Fig. 3. Comparison of RMS value for both data models

Table IV shows that the RSS value for regional model is bigger than the value of global model TPX09. The RSSIQ value in both data models has a value of 0.665 which indicates that the estimates of all errors between the tidal model and the Sekupang tidal station obtained previously from the RSS value are not significantly different.

The discrepancy value obtained shows different values between models. The best model is the model that has the smaller D value. This means that the difference between the model's harmonic constant and the tidal observation data owned by the model is smaller. From the calculation results, it is known that the D values of the regional model and global models are 63,137 and 52,364, respectively, which means that the smaller D value is the TPXO9 model. This is possible because the TPXO9 model has a denser grid so that it can perform better data interpolation.

IV. CONCLUSION

From the research result, it is known that the dominant tidal component is the M2 tidal constituent, this indicates that the area is dominated by the semidiurnal tide which is more dominant than the value of the diurnal harmonic constituent. The minimum discrepancy value is shown by the TPXO9 model, which means that the TPXO9 global model is a more suitable model to be used as a tidal prediction model in the waters of Batam Island.

ACKNOWLEDGMENT

Acknowledgments are addressed to the Geospatial Information Agency that has been willing to assist in providing data for the completion of this research.

REFERENCES

[1] Lasabuda, R. (2013). Laboratorium Pengelolaan Wilayah Pesisir Terpadu, FPIK UNSRAT.Regional Development in Coastal and Ocean in Archipelago Perspective , 1-2.

[2] Pusat Jaring Kontrol Geodesi dan Geodinamika, BAKOSURTANAL, (2008), Prediksi Pasang Surut 2008, Laporan

[3] Triatmodjo, B. (1996). Pelabuhan. Yogyakarta: Beta Offset.

[4] Poerbondono, & Djunarsjah, E. (2005). Survei Hidrografi. Bandung: Pt. Refika Aditama

[5] Defant. (1958). Ebb and flow; the tides of earth, air, and water. Michigan: Ann Arbor,University of Michigan Press.

[6] G. Shanmugam. (2021). Mass Transport, Gravity Flows, and Bottom Currents - Downslope and Alongslope Processes and Deposits. Department of Earth and Environmental Sciences, The University of Texas at Arlington, Arlington, Texas, United States.

[7] Poerbandono dan Djunarsjah, E., 2005, Survei Hidrografi, Bandung: PT Refika Aditama.

[8] Supriyono, Pranowo, W. S., Rawi, S., Herunadi, B. 2015. Analisa dan Perhitungan Prediksi Pasang Surut Menggunakan Metode Admiralty dan Least Square (Studi Kasus Perairan Tarakan dan Balikpapan). Jurnal Program Studi S1-Hidrografi. STTAL.

[9] Rufaida, N. H. (2008). Perbandingan Metode Least Square (Program World Tides dan Program Tifa) dengan Metode Admiralty Dalam Analisis Pasut. Bandung: Istitut Teknologi Bandung

[10] Fok, H. S, dkk, (2010), Evaluation of ocean tide models used for Jason-2 altimetry corrections, Marine Geodesy 33.S1: 285-303

9th Applied Business and Engineering Conference 2021

Pekanbaru - August, 25th 2021

Organized by

Politeknik Caltex Riau