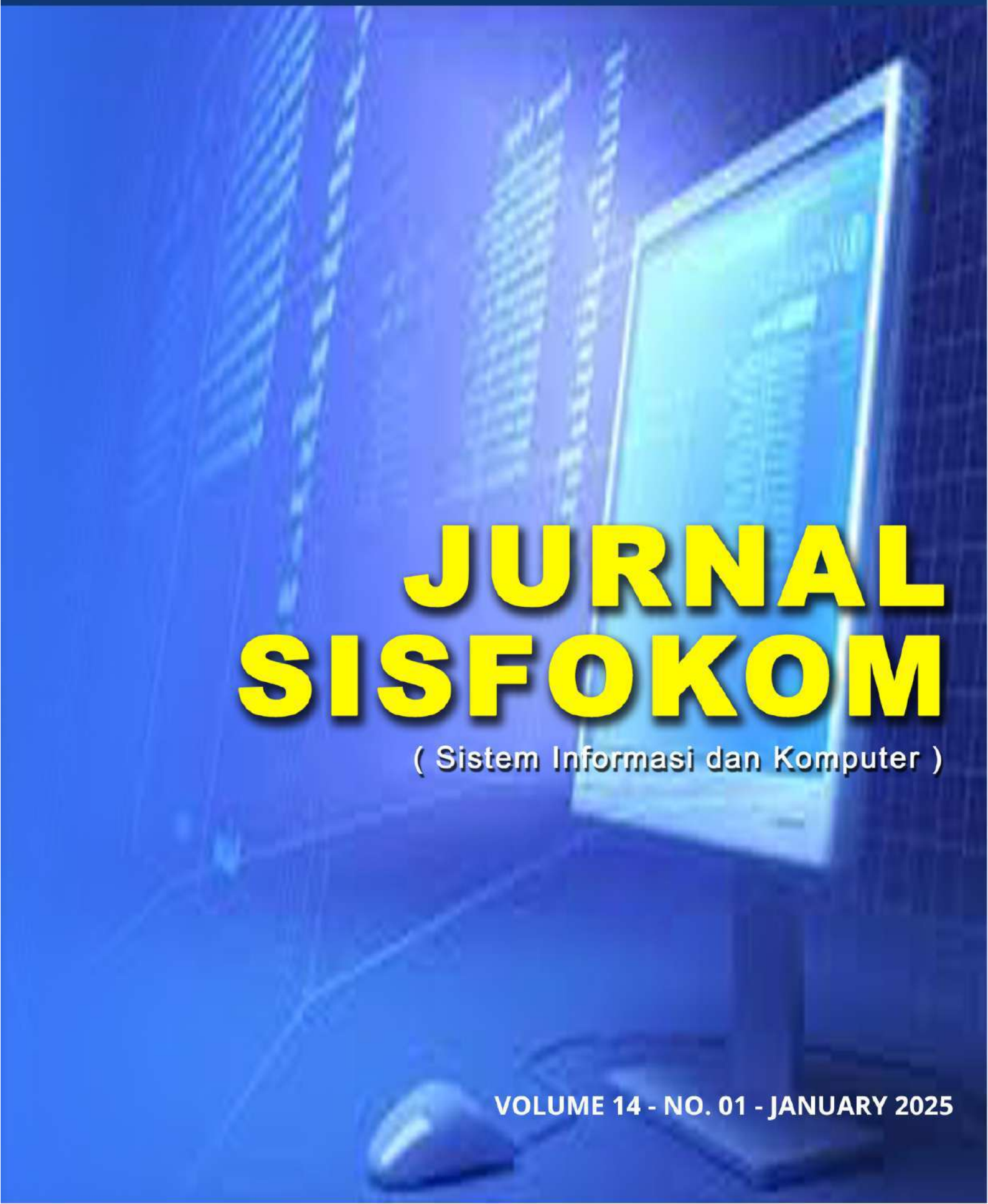
E-ISSN : 2581-0588
P-ISSN : 2301-7988

INSTITUT SAINS DAN BISNIS
ATMA LUHUR



JURNAL SISFOKOM

(Sistem Informasi dan Komputer)

VOLUME 14 - NO. 01 - JANUARY 2025

P-ISSN: 2301-7988
E-ISSN: 2581-0588

JURNAL SISFOKOM

(SISTEM INFORMASI DAN KOMPUTER)

Jurnal Sisfokom, an acronym for Journal of Information Systems and Computers, is a scholarly publication resulting from a collaborative effort between the academic community of ISB Atma Luhur and various higher education institutions across Indonesia. This journal serves as a vital platform for disseminating scientific articles from researchers, academics, and practitioners in the field of information technology. With a specific focus on information systems and computer science, Jurnal Sisfokom consistently presents high-quality papers, published four times a year in January, May, July, and November, ensuring a continuous and relevant flow of knowledge.

EDITORIAL BOARD

P-ISSN: 2301-7988
E-ISSN: 2581-0588

P-ISSN: 2301-7988
E-ISSN: 2581-0588

SECRETARIAT

LPPM Institut Sains dan Bisnis Atma Luhur
Jl. Jend. Sudirman, Selindung Baru, Pangkalpinang
Kepulauan Bangka Belitung - Indonesia
Telp. (0717) 433 506 Fax. (0717) 433 506
e-mail : sisfokom@atmaluhur.ac.id - lppm@atmaluhur.ac.id

P-ISSN: 2301-7988
E-ISSN: 2581-0588

EDITOR'S FOREWORD

Jurnal Sisfokom (Information Systems and Computers) stands as a distinguished scholarly publication, meticulously managed and proudly issued by the LPPM ISB Atma Luhur Pangkalpinang.

This particular edition, Volume 14 Number 01 – January 2025, exemplifies a robust collaborative endeavor, uniting the academic prowess of ISB Atma Luhur with a diverse array of esteemed universities throughout Indonesia.

The editorial board extends its profound gratitude for the invaluable participation and unwavering cooperation of our dedicated lecturers. Their significant contributions have been pivotal, enabling the timely and successful publication of Jurnal Sisfokom (Information Systems and Computers) Volume 14 Number 01 – January 2025, precisely in accordance with our meticulously laid plans.

Additionally, the editorial board extends its profound appreciation to the distinguished experts, both internal and external to ISB Atma Luhur, whose invaluable contributions in thoroughly assessing and meticulously refining the submitted manuscripts have been absolutely critical to the quality of this publication.

On this occasion, the editorial board cordially invites and extends the widest possible opportunity to all researchers, fellow lecturers, and discerning scholars/practitioners in both Information Systems and Informatics Engineering to contribute and publish their research findings through this esteemed journal.

Ultimately, the editorial board sincerely hopes that the scholarly articles published within this journal will yield substantial benefits for the entire academic community at ISB Atma Luhur, and profoundly contribute to the broader advancement of science and information technology.

Editorial,

P-ISSN: 2301-7988
E-ISSN: 2581-0588

TABLE OF CONTENTS

P-ISSN: 2301-7988
E-ISSN: 2581-0588

P-ISSN: 2301-7988
E-ISSN: 2581-0588

Prediction of Claim Fund Reserves in Insurance Companies Using the *ARIMA* Method

Goenawan Brotosaputro[1]*, Wiwin Windihastuty[2], Yohanes Setiawan Japriadi[3] , Rivai Ahsani[4]
Faculty of Information Technology [1], [3]
Atma Luhur Institute of Science and Business
Pangkalpinang, Indonesia

Faculty of Information Technology [2], [4]
Budi Luhur University
Jakarta, Indonesia
gbrotos@atmaluhur.ac.id[1], wiwin.windihastuty@budiluhur.ac.id[2], ysetiawanj@atmaluhur.ac.id[3],
2211600032@student.budiluhur.ac.id[4]

Abstract*— Insurance is a financial protection contract between a customer and an insurance company which is stated in the form of an insurance policy. Prediction of insurance claim reserve funds is necessary because the claim amount varies and the claim time can be the same. If at any time there is a claim that is so large that it exceeds the available claim reserve fund plus the claim occurs at the same time, it can cause the company to fail to pay the claim. This will certainly make the company's conduct decline, customer trust will be lost, and can cause the company to go bankrupt. The problem can be solved if the insurance company has sufficient claim fund reserves. Claim fund reserves are an important issue in insurance companies. This study aims to predict the claim fund reserves in insurance companies to anticipate varying claim amounts. Historical analysis of the value of claims with the ARIMA model approach is used to predict future claim values. We use claim value data that has been scaled in millions. 2020 to 2022 as training data and 2023 as test data. The Root Mean Square Error (RMSE) metric obtained is IDR 25,780.71; Mean Absolute Deviation (MAD) of IDR 14,421.89, and Mean Absolute Percentage Error (MAPE) of IDR 5,967.27; while the total actual claim value in 2023 is IDR 161,700.51 and the total predicted claim value is IDR 166,227.36; which means that an accuracy of 97% is obtained. The result of claim prediction value in one periodic year can give a favor to the management to make a decision, how much the claim funds should be prepared.

Keywords*— prediction; classification; claim; ARIMA; insurance

I. INTRODUCTION

In life, humans are not free from risks. Risk is any event that causes loss. Protection from various risks in everyday life has become a necessity. One institution that can provide this protection is an insurance company. Insurance is an agreement between the insurer and the insured to receive insurance premiums to replace or provide payment for losses experienced by the insured from an event.

Insurance promises protection to the insured against risks faced by individuals and organizations. Today, the insurance industry is experiencing very rapid growth. The insurance industry plays an increasingly important role in the Indonesian economy by providing benefits to both individuals and society. One of the characteristics of loss insurance is that the premium charged to policy holders depends on the number of claims and the size of the claims that have been submitted. The overall claim amount depends on the size of the claim and the frequency of claims in a particular period.

The number of claims submitted by policy holders represents the actual risk of the policy holder. Factors that cause the risk of damage or loss are disasters (perils) and dangers (hazards). This disaster factor can be influenced by the weather climate, which cannot be predicted or prevented. These factors will influence the number of claims.

Bankruptcy is the biggest risk that a company might experience. Technological developments, adaptability to the company's financial economic conditions are common factors causing bankruptcy. The company's financial condition has the largest percentage in influencing bankruptcy. To measure the level of risk and health of the company, an analysis of the company's financial statements needs to be carried out. From the analysis of the company's financial statements, it can be seen if the company is experiencing financial difficulties[1]. Bankruptcy is a significant event that can cause major losses to management, shareholders, employees, customers and the nation, so predicting company bankruptcy has become a hot topic for both industrial applications and academic research. In recent years, many studies have shown that machine learning techniques such as Artificial Neural Networks (ANN), Decision Tree (DT), Case Based Reasoning (CBR), Support Vector Machine (SVM) can be used as alternative methods for predicting company bankruptcy. Unlike statistical techniques, machine learning techniques do not assume a particular data distribution and automatically extract knowledge from training samples. Prediction performance depends on the details of the problem, the characteristics of the data structure used, the extent to which it is possible to separate classes using these characteristics, and the purpose of the classification[2]. Bankruptcy is a situation where a company experiences a shortage and insufficiency of funds to run or continue its

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

business, a more serious consequence of bankruptcy is in the form of business closure or liquidation[3]. Bankruptcy prediction is very important for all organizations and agencies because it has a major impact on the economy and price increases will cause many social problems[4]. Several insurance companies in Indonesia have defaulted, namely the failure of the obligation to pay claims to the insured. Many factors cause insurance companies to default, one of which is due to poor calculation of claim fund reserves[5].

This research will be predict claim fund reserves using the *ARIMA* method so that insurance companies avoid bankruptcy due to inability to pay when claims occur that exceed the available claim funds. The Arima-based method often used in predicts the stock price in a company like said in [6] and have the trend predictions better than other machine learning methods such as linear regression, random forest, decision tree and gradient boosting machine [7]. Measurements should be carried out on research results to determine whether the implementation of a method is appropriate, such as in [8] which produces a smaller MAPE using ARIMA compared to using Single Exponential Smoothing in forecasting Indonesian cocoa export. Metrics other than MAPE are Mean Square Error (MSE) and Mean Absolute Error which are also used in [9] to predict dengue fever cases using ARIMA methods and have a good result.

The *Autoregressive Integrated Moving average (ARIMA)* method or commonly called the Box-Jenkins method is a method that was intensively developed by George EP Box and Gwilym Jenkins in the 1970s. The Box-Jenkins model group included in this method includes: *Autoregressive (AR), Moving average (MA), autoregressive-moving average* (*ARMA*), and *Autoregressive Integrated Moving average (ARIMA)*. The single value of the time series cannot be predicted, but if an analysis is carried out on the entire value of the time series, it has a certain pattern[10].

The data to be processed is data that has a record of the time the claim occurred and the value of the claim (numerical). The data obtained is analyzed to determine the pattern of past data that has been collected using the *ARIMA* model which can predict short-term forecasting for non-stationary data. Time series models such as *ARIMA* attempt to predict future conditions by using historical data and predicting the future. *ARIMA* prediction results in several studies are suitable for predicting variables that are very sensitive to short-term changes.

The *ARIMA* model uses dependent variables and ignores independent variables to produce accurate short-term forecasting. This research calculates predicted claim values based on previous claim values recorded in time series form. This research can provide input to insurance companies as a reference in predicting the provision of claim funds so that they do not cause problems in the future for their customers.

II. METHODOLOGY

The methodology in this research uses a quantitative approach whose aim is to calculate various things in an effort to explain what is observed and is useful for generalizations, predictions, and causal explanations. Another characteristic is that data is obtained in a structured manner in numerical form and in large quantities[11]. Previous research shows that ARIMA can be used to predict many things and get good results, so in this research the ARIMA method will also be used to predict the reserve funds of claims at insurance companies.

A. Research Flow

In predicting claims fund reserves using the *ARIMA* model, there are several stages that need to be carried out. This stage includes data collection, data preparation, data transformation, data sharing, *ARIMA* modeling, evaluation and measurement of prediction results which can be seen in Fig. 1.

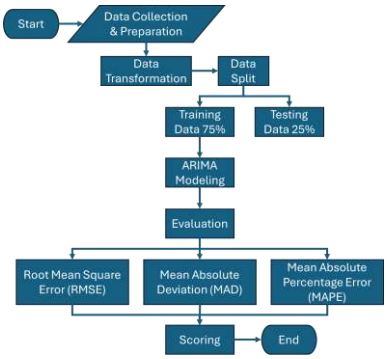


Fig. 1 Research Flow

B. Data Collection and Preparation

Claim value data for the analysis process was obtained from one of the leading insurance companies in Indonesia. The data obtained in Excel format is 8,301 rows with 25 columns which store claim data in a time series from 2019 to 2023. From the 25 existing fields, researchers only used the [Date of Loss, Currency and Gross Claim] fields with the aim of predicting claim value in the following times, so it is suspected that these 3 fields can be used to predict the value of the next claim which can ultimately predict the total claim fund reserves that must be provided by the insurance company. This raw data can be seen in Fig. 2. Data in 2019 that contained only 4 rows were eliminated because its not likely the others in 2020 until 2023.

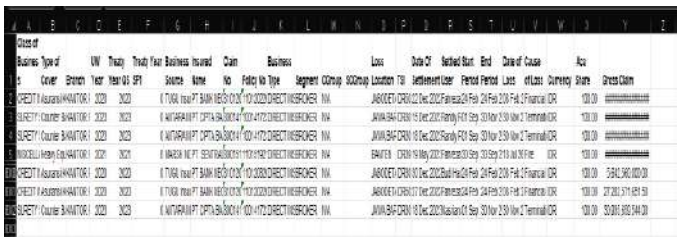


Fig. 2 Raw Data

C. Data Transformation

Based on observations and data analysis, we found that Currency consists of 5 different currencies. Standardization is needed for the differences in these currencies. Conversions were made from Australian Dollar, Euro, Singapore Dollar, and

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

US Dollar to Rupiah. We used historical currency data by utilizing the site https://www.xe.com/ and to simplify it, the exchange rate was done by searching for the average value history from the 1st of each month to the end date of it. The value obtained was used as the Gross Claim exchange rate in the same month and year and then the claim value in foreign currency was converted to Indonesian Rupiah. After the conversion was completed, data aggregation was made by calculating the claim value per month per year which was represented in the occurrences column and the total claim per month as shown in TABLE 1.

Some of total claims per month have very large value, so to makes the good visibility of data, these value will be scaled in millions.

ARIMA has ability to make predictions with a single historical data, so in this research we only use the Total Claims per month and month as a time series. Time series in the form of dates require complete dates, so that for each month the 1st is given, the data are ready to be applied to the *ARIMA* model is presented in *TABLE* 2.

D. Data Split

To validate whether the model can be applied, the data needs to be divided into training data and test data. The training data using three years of data claim value per month from January to December 2020 until 2022 (75%) and test data using one year claim value per month from January to December 2023 (25%).

TABLE 1. Number of incidents and total claims per month

	2020		2021	
Month	Occurences	Total Claim	Occurences	Total Claim
Jan	10	433,189,538.48	197	6,955,649,415.77
Feb	1	17,181,818.00	247	100,992,041,856.33
Mar	3	23,986,616.35	156	5,786,845,736.36
Apr	11	3,184,744,171.47	255	11,591,699,592.74
May	15	177,464,989.69	136	3,508,105,175.11
Jun	20	2,737,806,501.20	128	2,218,888,646.71
Jul	57	6,971,070,261.27	132	1,164,053,047.63
Aug	27	38,516,815,629.18	142	4,736,044,421.67
Sep	99	3,331,185,777.17	155	5,545,840,503.78
Oct	78	3,461,902,091.75	254	11,126,383,570.43
Nov	54	1,361,843,072.62	204	9,597,759,035.99
Dec	77	2,237,629,991.97	226	10,087,082,882.02
	2022		**2023**	
Month	Occurences	Total Claim	Occurences	Total Claim
Jan	222	8,267,771,542.23	173	4,430,835,869.42
Feb	165	2,982,610,888.24	196	5,029,537,017.84
Mar	222	9,291,233,284.94	177	6,899,123,050.30
Apr	177	2,497,450,290.58	167	5,631,003,547.19
May	240	52,063,370,024.67	166	5,284,908,740.93
Jun	246	104,268,400,239.60	154	5,932,366,203.28
Jul	213	36,285,032,455.71	415	12,875,165,010.95
Aug	172	3,296,534,460.00	506	12,543,437,955.69
Sep	156	22,788,867,600.88	516	3,753,778,425.75
Oct	176	3,402,032,912.49	365	1,140,997,927.52
Nov	429	6,824,707,679.67	121	98,159,305,889.88
Dec	235	12,356,519,724.81	4	20,047,980.45

TABLE 2. Claim by Month and Year (in Million)

Month	2020	2021	2022	2023
Jan	433.1895385	6955.649416	8267.771542	4430.835869
Feb	17.181818	100992.0419	2982.610888	5029.537018
Mar	23.98661635	5786.845736	9291.233285	6899.12305
Apr	3184.744171	11591.69959	2497.450291	5631.003547
May	177.4649897	3508.105175	52063.37003	5284.908741
Jun	2737.806501	2218.888647	104268.4002	5932.366203
Jul	6971.070261	1164.053048	36285.03246	12875.16501
Aug	38516.81563	4736.044422	3296.53446	12543.43796
Sep	3331.185777	5545.840504	22788.8676	3753.778426
Oct	3461.902092	11126.38357	3402.032912	1140.997928
Nov	1361.843073	9597.759036	6824.70768	98159.30589
Dec	2237.629992	10087.08288	12356.51973	20.04798045

E. ARIMA Model

The components of *ARIMA*, as the name suggests, are (1) *Autoregressive*, (2) *Integrated*, and (3) *Moving average*. The stages for doing *ARIMA* are *Autoregressive (AR)*, *Moving average (MA)*, *Autoregressive Moving average (ARMA)*, and finally *Autoregressive Integrated Moving average (ARIMA)*.

Autoregressive (AR)

The *autoregressive* model is a stationary model of time series data where the observation value at time *t* is influenced by the previous observation value. This model uses the *AR* order (*p*) or *ARIMA* model *(p,0,0)* expressed in equation *(1)*.

$$Yt = \beta 1\ Yt\text{-}1 + \beta 2\ Y\ t\text{-}2 + \ldots + \beta p\ Y\ t\text{-}p + et \qquad (1)$$

explanation:
Yt = The series values are stationary
βi = *i*-th *autoregressive* parameter
et = White Noise error value at time *t*

Independent variables are a series of values of similar variables in the last few *t* periods. While at is an error or residual unit that describes random disturbances that cannot be explained by the model. *Autoregressive* calculations can be done in the following process:

1) Determine the model that fits the time series.
2) Determine the value of the order p (determine the length of the equation formed)
3) Estimate the *autoregressive* coefficient values β1, β2, β3, βk

Moving average (MA)

The *Moving average (MA)* model shows observations at previously influenced times. The *moving average* is denoted in *MA (q)* or *ARIMA (0,0,q)* which is written in equation *(2).*

$$Yt = et - \beta 1\ et\text{-}1 - \beta 2\ et\text{-}2 - - \ldots - \beta q\ et\text{-}q \qquad (2)$$

explanation:
Yt = Stationary series values
βi = *Moving average* Parameter

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

et = White noise / error or residual unit

Autoregressive Moving average (ARMA)

The combination of the *Autoregressive (AR)* and *Moving average (MA)* models will form a new model, namely *ARMA (autoregressive moving average)* with the *ARMA* order *(p,q)*. The general form of the *ARMA* equation is a combination of the *AR* and *MA* equations which are denoted in equation *(3)*.

$$Y\,t = \beta 1\ Y\,t\text{-}1 + \beta 2\ Y\,t\text{-}2 + \ldots + \beta p\ Y\,t\text{-}p + et - \beta 1\ et\text{-}1 - \beta 2\ et\text{-}2 - \text{-} \ldots - \beta q\ et\text{-}q \quad (3)$$

explanation:
Yt = Stationary series values
βi = *Moving average* Parameter
et = White noise / error or residual unit

ARMA modeling has a basic theory of correlation and stationarity. This means that *ARMA* can be used when the time series has formed a stationary graph, or does not form an upward or downward trend. However, if the time series data is not stationary, then a differentiation process is needed to change the data to become stationary before it can be processed through *ARMA*.

Autoregressive Integrated Moving average (ARIMA)

The *AR*, *MA*, and *ARMA* models use the assumption that the time series data produced is already stationary. In reality, time series data is more often non-stationary. If the data is not stationary, the method used to make the data stationary is differencing for non-stationary data in the mean and the transformation process for non-stationary data in the variance. The general form of the *ARIMA* model can be expressed in equation *(4)*.

$$Zt = \emptyset 1 Zt{-}1 + \emptyset 2 Zt{-}2 + \cdots + \emptyset p Zt{-}p + at + \theta 1\, at{-}1 + \cdots + \theta p\, aq{-}1 \quad (4)$$

explanation:
Zt = data at time t, $t = 1,2,3, \ldots, n$
$Zt{-}i$ = data at time $t - i$, $i = 1,2,3, \ldots, p$
$Zt{-}i$ = error at period $t - i$, $i = 1,2,3, \ldots, q$
at = error at period t, $= 1,2,3, \ldots, n$
$\emptyset 0$ = *Autoregressive (AR)* model constants
$\emptyset i$ = coefficient of $Zt{-}i$ in the *Autoregressive (AR)* model
θi = coefficient of $at{-}i$ in the *Moving average (MA)* model

The general process for *ARIMA* modelling is: visualize the time series data, make the time series data stationary, plot the correlation and autocorrelation charts, construct the *ARIMA* model and use the model to make predictions. The establishment of the *ARIMA* model is that the time series are required to have stationarity. It can be judged intuitively by making the original sequence diagram, and the stationarity of the sequence can also be verified by the inspection method. If the sequence is not stable, one can use data transformation or difference to make the data stable [10].

1. Visualize the time series data

Time series data visualization is very important in *ARIMA* methods because it helps identify data patterns and trends, and ensures that the model built is accurate and relevant[12]. In addition, data visualization also helps identify the suitability of the *ARIMA* model to the existing data, thereby improving the quality of predictions.

From the historical data in *Fig. 3* it can be seen that in order to predict the claim value in the future, it is necessary to create a proper relationship between the claim value and time. The graphic image can be used to analyze data patterns and trends, whether stationary, non-stationary or seasonal and the hypothesis *H0* is taken: non-stationary data and *H1*: stationary data.

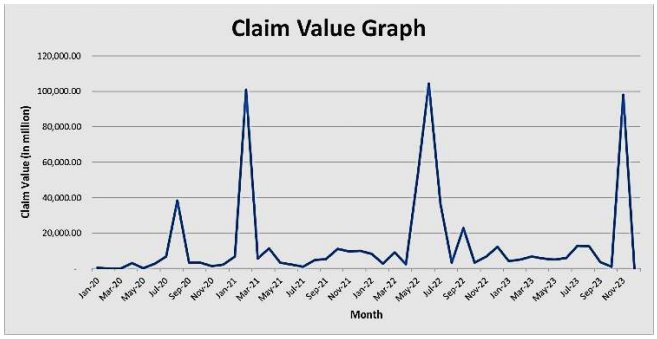
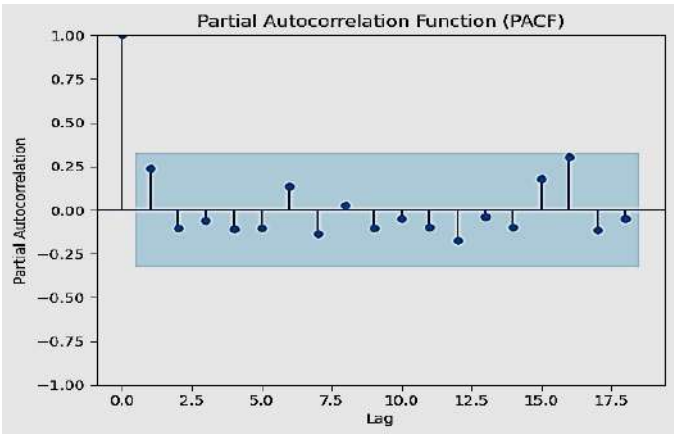


Fig. 3. Claim Value data visualization graph



2. Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) Plot Analysis

ACF and PACF are used to find the next parameters of *ARIMA*. Plot ACF (Autocorrelation Function) to determine the q value and PACF (Partial Autocorrelation Function) to determine the p value. In Fig. 5 it is known that the pattern decreases drastically after lag 1 and then the ACF value moves around zero. The appropriate model is MA(0) or the q value is 0 because it is not significant and there is no oscillation or cutting pattern at a certain lag. Because q is not significant, the value of q can also be 1. Meanwhile in Fig. 6 also shows a drastic decreasing pattern and then the PACF value moves around zero. The appropriate model is *AR*(1) or the p value is 1 which indicates a strong correlation between the current value and the previous value. The parameter p = 1, d = 0 and q

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

= 0 will be used as the first parameter to model ARIMA(1,0,0) and the result will be evaluated to find the optimum parameters.

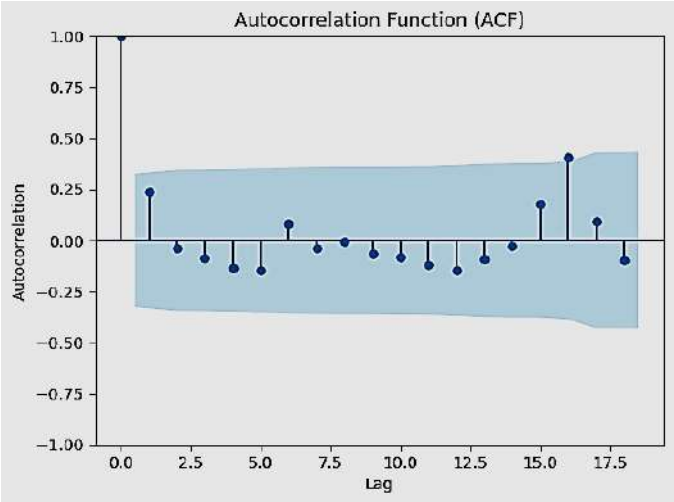


Fig. 5. Plot Autocorrelation (ACF)

F. Evaluation

After the values p, d, and q are obtained they will be used in the *ARIMA*(1,0,0) model. Based on training data and test data, training data from 2020 to 2023 was used and test data was data from 2023. The predicted claim value was obtained as in TABLE 3 and visualize in Fig. 4.

Based on the 2023 prediction results, several metrics were calculated, namely Root Mean Square Error (RMSE) of 25,780.71, Mean Absolute Deviation (MAD) of 14,421.89 and Mean Absolute Percentage Error (MAPE) of 5,967.27.

The ACF plot in *Fig. 5* shows that the residuals are randomly distributed, but there is an autocorrelation value that is outside the confidence interval limit (light blue area) at lag 16, which shows that the ARIMA(1,0,0) model is too simple, so it is necessary to modify the AR value or additional MA value to capture autocorrelation at lag 16.

Fig. 6. Plot Partial Autocorrelation (PACF)

TABLE 3. The 2023 Claim Fund Prediction with ARIMA(1,0,0)

Month	Claim	Forecast
Jan-23	4,430.84	13,532.02
Feb-23	5,029.54	13,807.24
Mar-23	6,899.12	13,871.68
Apr-23	5,631.00	13,886.77
May-23	5,284.91	13,890.30
Jun-23	5,932.37	13,891.13
Jul-23	12,875.17	13,891.32
Aug-23	12,543.44	13,891.37
Sep-23	3,753.78	13,891.38
Oct-23	1,141.00	13,891.38
Nov-23	98,159.31	13,891.38
Dec-23	20.05	13,891.38
Total	161,700.51	166,227.36

Experiments by changing the AR value or MA value several times obtained a fixed accuracy of 76% for predicting the value of claims accumulated over an annual period. However, if we look at the monthly period, at most 3 months the predicted value is close to the actual value as shown in *TABLE 4* and *Fig. 7* for January, February and August.

With the same metrics, the ARIMA(3,0,2) model obtains a Root Mean Square Error (RMSE) of 25,378.16, Mean Absolute Deviation (MAD) of 11,358.37 and Mean Absolute Percentage Error (MAPE) of 5,888.64.

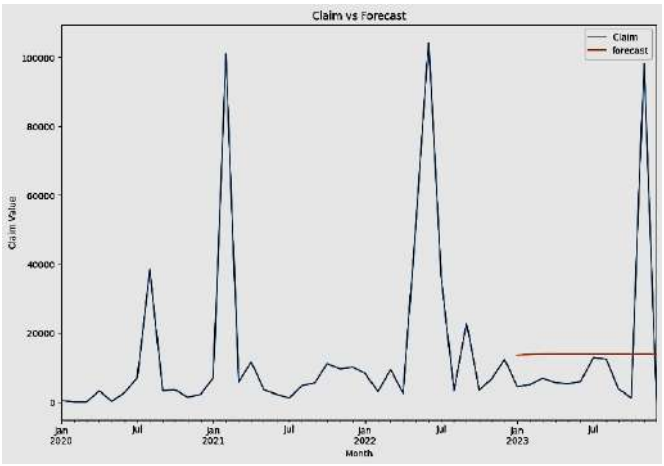


Fig. 4 Claim and Forecast comparison chart for 2023 with ARIMA(1,0,0)

TABLE 4. The 2023 Claim Fund Prediction with ARIMA(3,0,2)

Month	Claim	Forecast
Jan-23	4,430.84	5,275.16
Feb-23	5,029.54	5,796.02
Mar-23	6,899.12	5,898.88
Apr-23	5,631.00	8,541.00
May-23	5,284.91	9,022.58
Jun-23	5,932.37	11,161.42
Jul-23	12,875.17	11,169.29
Aug-23	12,543.44	12,694.75
Sep-23	3,753.78	12,358.17
Oct-23	1,141.00	13,477.99
Nov-23	98,159.31	12,976.87
Dec-23	20.05	13,851.67
Total	166,700.51	122,223.80

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

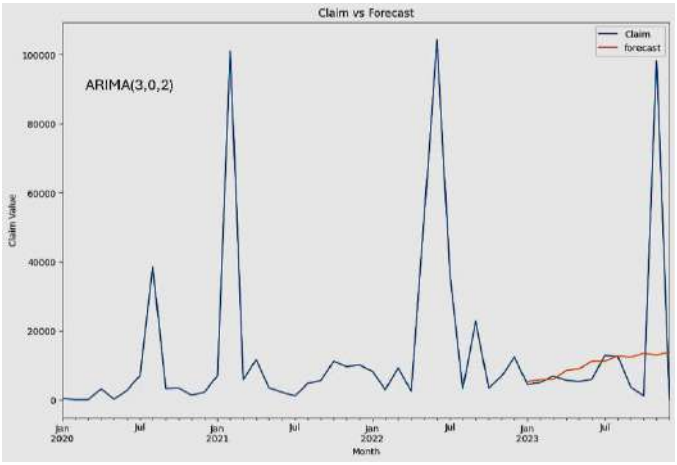


Fig. 7 Claim and Forecast comparison chart for 2023 with ARIMA(3,0,2)

Two experiments using ARIMA(1,0,0) and ARIMA(3,0,2) showed different prediction results for each month. The prediction value for each month in ARIMA(1,0,0) has almost the same value overall for 12 months, whereas in ARIMA(3,0,2) produces predictions with varying values in each month, especially in January, February and August whose value is close to the actual value.

III. Discussion And Results

Based on the Mean Absolute Percentage Error of ARIMA(1,0,0) model, a very large value is obtained, showing that the RMSE and MAPE values are greater than 50%, which means the model is inaccurate, this is due to the claim calculation which is calculated per month. However, if claims are calculated within 1 year, it can be calculated using TABLE 3 that the total actual claims in 2023 will be 161,700.51 and the total predicted claims in 2023 will be 166,227.36. This shows that if the insurance company will reserve funds for annual claims, this model can be used well, not for reserve monthly claim funds. The Mean Absolute Deviation (MAD) value of 14,421.89, compared with the actual average claim in 2023 which is 13,475.04 only has a difference of 946.85 or 7% also shows that this model suitable for predict the annual reserve claim funds.

Experiment by changing the parameters p, d and q becomes ARIMA(3,0,2) produce the metrics RMSE, MAD and MAPE becomes smaller than ARIMA(1,0,0), that are 25,378.16; 11,358.37 and 5,888.64 consecutively. This shows that the monthly prediction result better, but only for January, February and August. The total for one year shows the difference between actual and forecast claim is 39,476.71 or 24% deviated from the actual claim. Mean Absolute Deviation (MAD) has 11,358.37 which is deviated from the average actual claim 13,475.04 is 2,116.68 or 16%.

IV. Conclusion

Having sufficient claims reserves and predicting them accurately is an important step for insurance companies to manage their financial obligations, ensure business continuity, and fulfill obligations to policyholders. Without adequate claims reserves, insurance companies risk facing financial difficulties that can harm many parties.

The Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) metrics obtained from the *ARIMA*(1,0,0) model are greater than 50%, which means this model is not accurate, however, the Mean Absolute Deviation (MAD) metric The difference between the actual data was 7%, which means the model is very accurate. This contradiction is caused by the very fluctuating value of claims, especially in 2020 where the value was much smaller than the values in other years, this is thought to be due to the Covid-19 disaster at that time.

ARIMA(1,0,0) in the data studied shows better performance compared to ARIMA(3,0,2) which is shown by a difference of 3% in ARIMA(1,0,0) and 24% in ARIMA(3,2,0) for the total difference between predictions and actual values in one year, thus the parameters p = 1, d = 0 and q = 0 are suitable to be applied.

Comprehensively, based on the actual total claims in 2023 with the predicted total claims in 2023 using training data from 2020 to 2022, a difference of IDR 4,526.85 (in millions) is obtained or only 3% greater than the actual total claims or close to 97% accurate. This concludes that this model is suitable to be applied to predict the total value of claims annually, not monthly. The result of claim prediction value in one periodic year can give a favor to the management to make a decision, how much the claim funds should be prepared.

Acknowledgment

We express our deepest gratitude to the Institut Sains dan Bisnis Pangkalpinang and Universitas Budi Luhur Jakarta for providing the opportunity for their lecturers and students to implement *Tridarma Perguruan Tinggi* in the field of research. We also express special thanks to the Ministry of Education and Culture for the 2024 National Competitive Research Grant for Research on Affirmation Beginner Lecturers, who have funded this research until completion.

References

[1] A. W. Avrizal and E. Junianto, "Prediksi Kebangkrutan Perusahaan Menggunakan Algoritma Decision Tree Berbasis Teknik Resampling," *eProsiding Tek. Inform.*, vol. 4, no. 1, pp. 101–109, 2023, [Online]. Available: https://eprosiding.ars.ac.id/index.php/pti/article/view/919/403.

[2] H. Saleh, "Prediksi Kebangkrutan Perusahaan Menggunakan Algoritma C4.5 Berbasis Forward Selection," *Ilk. J. Ilm.*, vol. 9, no. 2, pp. 173–180, 2017, doi: 10.33096/ilkom.v9i2.97.173-180.

[3] S. Bahri and N. Widyawati, "Analisis Prediksi Kabangkrutan Pada Perusahaan Yang Di- Delisting Di BEI," *J. Ilmu dan Ris. Manaj.*, vol. 4, no. 8, pp. 3–22, 2015, [Online]. Available: http://journal.stieputrabangsa.ac.id/index.php/jimmba/index.

[4] R. A. Nurdini, Y. Priyadi, and . N., "Analisis Prediksi Kebangkrutan Perusahaan Menggunakan Artificial Neural Network Pada Sektor Pertambangan Batubara," *J. Sist. Inf. Bisnis*, vol. 8, no. 1, p. 107, 2018, doi: 10.21456/vol8iss1pp107-114.

[5] Y. Hikmah and I. R. Hikmah, "Perhitungan Cadangan Klaim Dengan Metode Chain Ladder Menggunakan Excel Dan Rstudio," *MAp (Mathematics Appl. J.*, vol. 4, no. 2, pp. 122–131, 2022, doi: 10.15548/map.v4i2.4837.

[6] D. M. Putri and Aghsilni, "Estimasi Model Terbaik Untuk Peramalan Harga Saham PT. Polychem Indonesia Tbk Dengan Arima," *MAp J.*, vol.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

1, no. 1, pp. 1–12, 2019.
[7] M. Jia, H. Xu, and S. Zhang, *Comparison Analysis of ARIMA and Machine Learning Methods for Predicting Trend of US Semiconductor Stocks*, vol. 1. Atlantis Press International BV, 2022.
[8] M. S. Fahmuddin and S. M. Sri Mustika, "Perbandingan Metode ARIMA dan Single Exponential Smoothing dalam Peramalan Nilai Ekspor Kakao Indonesia," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 5, no. 3, pp. 163–176, 2023, doi: 10.35580/variansiunm193.
[9] E. Munarsih and I. Saluza, "Comparison of exponential smoothing method and autoregressive integrated moving average (ARIMA) method in predicting dengue fever cases in the city of Palembang," *J. Phys. Conf. Ser.*, vol. 1521, no. 3, 2020, doi: 10.1088/1742-6596/1521/3/032100.
[10] X. Zhang, Q. Zhou, S. Weng, and H. Zhang, "ARIMA Model-Based Fire Rescue Prediction," *Sci. Program.*, vol. 2021, 2021, doi: 10.1155/2021/3212138.
[11] S. Macdonald and N. Headlam, *Research methods handbook : introductory guide to research methods for social research*. 2008.
[12] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley and Sons Inc., 2015.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

Feature Extraction using Histogram of Oriented Gradients and Moments with Random Forest Classification for Batik Pattern Detection

Wafiq Azizah[1][*], Soffiana Agustin[2]
Department of Informatics Engineering, Faculty of Engineering [1], [2]
Universitas Muhammadiyah Gresik
Gresik, Indonesia
wafiqbila@gmail.com[1], soffiana@umg.ac.id[2]

***Abstract*— The preservation of traditional batik patterns, often transmitted orally and through direct practice across generations, faces significant challenges in the modern era. Globalization introduces the risk of cultural homogenization, potentially diminishing the uniqueness and diversity of these patterns. Furthermore, the manual recognition of batik motifs is labor-intensive, time-consuming, and requires specialized expertise, rendering it unsuitable for large-scale preservation initiatives. Consequently, the development of technology-based solutions capable of documenting, analyzing, and recognizing batik patterns with efficiency and precision is imperative for safeguarding this cultural heritage. This study aims to address these challenges by developing an automated system for recognizing batik patterns, focusing on Javanese batik motifs—Kawung, Megamendung, and Parang—which serve as foundational designs for the evolution of batik in other regions. The proposed methodology integrates two feature extraction techniques, Histogram of Oriented Gradients (HOG) and Texture Moments, with the Random Forest machine learning algorithm. The research process encompasses four key stages: pre-processing, feature extraction, classification, and system evaluation, where the accuracy of individual and combined feature extraction methods is analyzed. Experimental results reveal that the HOG method achieves an accuracy of 78.99%, while the Texture Moments method yields 81.88%. Notably, the combination of these two methods enhances system performance, achieving the highest accuracy of 86.23%, representing a 4.65% improvement over the single methods. These findings underscore the efficacy of integrating HOG and Texture Moments with the Random Forest algorithm for automated batik pattern recognition.**

Keywords*— *Classification, Batik, HOG (Histogram of Oriented Gradients), Texture Moments, Random Forest

I. INTRODUCTION

Batik is one of Indonesia's cultural heritages, recognized globally as a symbol of diversity and the beauty of textile art. Each batik pattern carries profound philosophical values, reflecting local wisdom and traditions that have evolved over centuries [1]. With the advancement of technology, the use of computer-based methods has become increasingly important for automatically recognizing and classifying batik patterns. Image processing technology enables the efficient analysis of patterns and textures in batik designs [2, 3].

Despite its cultural significance, batik faces several challenges in preservation and innovation. One of the main challenges in batik image detection is the high complexity of patterns and the similarity between motifs, which often leads to identification errors. External factors such as variations in lighting, image rotation, and scale further complicate accurate recognition. According to previous studies [4], conventional techniques without specific texture feature extraction methods often fail to capture the intricate details of batik patterns. This limitation hinders the classification process, reducing the efficiency of production in the modern batik industry and the digital documentation of batik as cultural heritage.

To address these challenges, this study focuses on three widely recognized motifs: Kawung batik from Surakarta, Megamendung batik from Cirebon, and Parang batik from Yogyakarta. These motifs not only hold unique philosophical values but are also commonly used as foundational designs in modern batik creations. The accurate recognition of these motifs is critical to supporting their preservation and enabling their integration into contemporary batik production.

In the field of image processing, the Histogram of Oriented Gradients (HOG) method has been widely adopted for extracting directional and edge-based features from images [5]. HOG has demonstrated its effectiveness in detecting texture features and patterns in various applications, including batik motifs [6]. Additionally, Texture Moments provide complementary information by capturing the texture distribution and smoothness of patterns [7]. Combining these two feature extraction methods can enhance the ability to differentiate complex and similar batik patterns [8].

For the classification task, the Random Forest (RF) algorithm is particularly suitable due to its robustness in handling large datasets, managing noisy data, and producing high accuracy [9, 10]. Prior studies have shown the success of integrating HOG features with Random Forest classification in applications such as plant identification, facial recognition, and object detection [11].

This study aims to combine HOG and Texture Moments as feature extraction methods with the Random Forest algorithm to classify Indonesian batik patterns. By utilizing a processed batik image dataset, the research evaluates the effectiveness of

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2225, Copyright ©2025
Submitted : December 20, 2024, Revised : January 14, 2025, Accepted : January 21, 2025, Published : January 31, 2025

the proposed approach in recognizing batik motifs, particularly Kawung, Megamendung, and Parang. The outcomes of this study are expected to contribute to the preservation of Indonesia's cultural heritage by providing an efficient, accurate, and automated solution for batik pattern recognition.

II. RESEARCH METHODOLOGY

This research is designed to recognize batik types based on their motifs. The batik motifs studied are Kawung, Megamendung, and Parang. The research is conducted in several stages. The first stage is pre-processing, which simplifies the image to reduce computational costs. The second stage involves feature extraction, where two methods are applied: Histogram of Oriented Gradients (HOG) and texture feature extraction using moments. These extracted features are then classified using the Random Forest algorithm to identify the batik types. The final stage is system testing, where evaluation is performed using a confusion matrix. The research stages are illustrated in Figure 1.

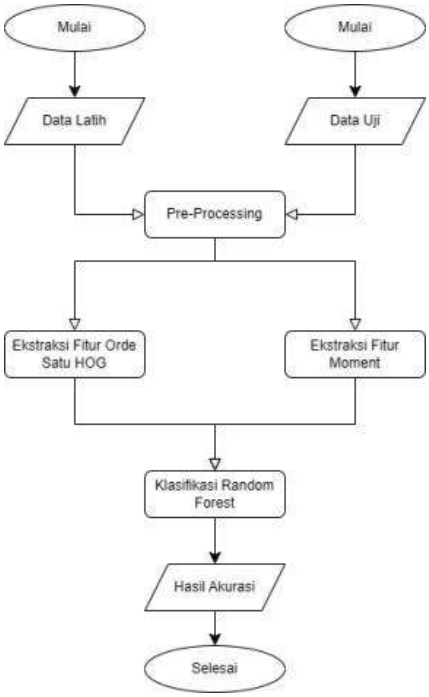


Fig. 1. Flowchart of the Batik Identification System Accuracy

A. *Pre-Processing*

This stage involves cropping and scaling to reduce the image size, which helps minimize computational load during processing. Once the images are resized uniformly, they are converted from RGB to grayscale.

B. *Feature Extraction*

The feature extraction process begins with preparing the batik image dataset, consisting of three main motifs: Parang, Kawung, and Megamendung. Each batik image is processed using two feature extraction methods: HOG and texture moments.

a. Texture Moments

The texture moments used in this study include statistical measures such as *mean, entropy, standard deviation, skewness, and kurtosis,* which can be mathematically defined as follows:

- *Mean*

  In Equation (1), the mean is obtained using xi to calculate the pixel intensity values, where N represents the total number of pixels in the image.

$$\mu = \frac{\sum x_i}{N} \quad (5)$$

- *Entopy*

  In Equation (2), p(xi) represents the probability of the pixel intensity xi.

$$H = -\sum p(x_i) log_2 p(x_i) \quad (6)$$

- *Standard Deviation*

  In Equation (3), the standard deviation is calculated based on the values of each data point (xi) and the Mean (μ).

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}} \quad (7)$$

- *Skewness dan Kurtosis*

  Skewness and kurtosis are calculated using standard statistical formulas, which measure the asymmetry of the distribution in Equation (4) and the sharpness of the distribution's peak in Equation (5).

$$Skewness = \left(\frac{x_i - \mu}{\sigma}\right)^3 \quad (8)$$

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2225, Copyright ©2025
Submitted : December 20, 2024, Revised : January 14, 2025, Accepted : January 21, 2025, Published : January 31, 2025

$$Kurtosis = \left(\frac{x_i - \mu}{\sigma}\right)^4 \quad (9)$$

b. Histogram of Oriented Gradients (HOG)

HOG is a highly effective feature extraction method for identifying patterns and textures in images. The HOG features capture the orientation and gradient magnitude information within an image, enabling the system to recognize the texture structure of batik images [12, 13, 14]. Mathematically, HOG can be explained through the following steps:

- ***Image Division****: The batik image is divided into several small grid cells (e.g., 8x8 pixels).*
- ***Gradient Calculation:*** *For each cell, the magnitude and orientation of the gradient are calculated using the Sobel operator, where Gx and Gy represent the image gradients in the horizontal and vertical directions, respectively.*

$$Gx = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (1)$$

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2225, Copyright ©2025
Submitted : December 20, 2024, Revised : January 14, 2025, Accepted : January 21, 2025, Published : January 31, 2025

$$Gy = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (2)$$

- ***Histogram Formation:*** *A gradient direction histogram is generated within each cell, indicating the distribution of orientation angles across the image. The magnitude and orientation are calculated using the following formulas:*

$$M = |Gx| + |Gy| \quad (3)$$

$$\theta = \arctan\left(\frac{Gy}{Gx}\right) \quad (4)$$

- ***HOG Feature Calculation:*** *After obtaining the HOG image, statistical values such as mean, standard deviation, skewness, and kurtosis are calculated to serve as the features derived from HOG.*

C. Random Forest Algorithm

After the feature extraction process, the next step is the classification of batik motifs using the Random Forest (RF) algorithm. Random Forest is an ensemble method that uses multiple decision trees to classify data [15, 16]. Each decision tree in RF is built using a random subset of the training data, and the final classification decision is made based on the majority decision from all the trees [17, 18].

Mathematically, Random Forest can be explained as follows:

1. ***Decision Tree Formation****: Each decision tree is created by selecting a random subset of the data and features.*
2. ***Classification Prediction****: For an input image with feature x, the following formula is used:*

$$\hat{y} = \frac{1}{n}\sum_{i=1}^{n} h_i(x) \quad (10)$$

This equation illustrates the basic principle of Random Forest in making predictions. By using multiple decision trees trained on random data subsets, Random Forest generates the final prediction by combining (averaging or majority voting) the results from all trees in the ensemble. This approach improves the accuracy and robustness of the model against overfitting, making it suitable for batik image data, which tends to have high variation.

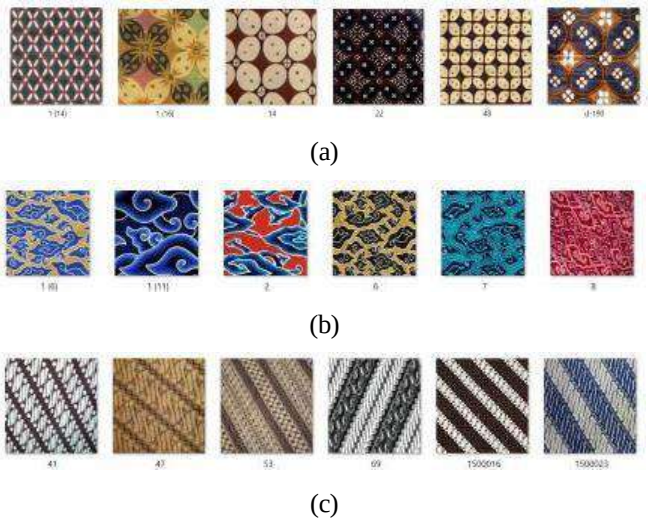
D. System Evaluation

The classification results from Random Forest will be compared with the true class labels to determine the accuracy of the results using the following formula:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Data}} \times 100\%$$

III. RESULTS AND DISCUSSION

The testing conducted in this study involves comparing the classification results from Random Forest using texture features, HOG, and the combination of both. The batik image dataset used in this research was sourced from the website https://www.kaggle.com/search?q=batik, with data collected from three main motifs: Parang, Kawung, and Megamendung. A total of 690 images were used in this study, divided into three classes: Parang (230 images), Kawung (230 images), and Megamendung (230 images), as shown in Figure 2 which devide into training data = 184 and testing data = 46.

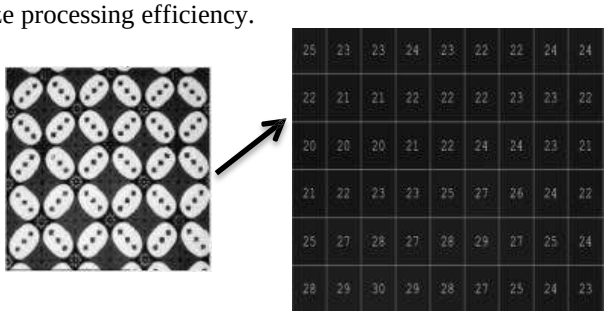


(a) (b) (c)

Fig. 2. Citra Data (a) Kawung (b) Megamendung (c) Parang

A. Pre-prosessing Stage

The dataset comprises images with varying dimensions, necessitating uniformity in size. This stage significantly reduces the image dimensions while standardizing them. Each image is resized to 244×244 pixels and subsequently converted to grayscale, as illustrated in Fig. 3. The grayscale conversion is performed to simplify computational complexity and optimize processing efficiency.



(a) (b)

Fig. 3. (a) Grayscale Citra Batik (b) Nilai Piksel Citra Batik

B. Feature Extraction Stage

a. Texture Feature Extraction

Texture feature extraction is calculated on the grayscale image. The texture values used are: Mean, Entropy, Standard Deviation, Skewness, and Kurtosis, calculated using equations (5 - 9). The extracted texture feature values are presented in Table 1.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2225, Copyright ©2025
Submitted : December 20, 2024, Revised : January 14, 2025, Accepted : January 21, 2025, Published : January 31, 2025

TABLE I. Texture Feature Extraction

FILE_NAME	Mean	Entropy	Std_Dev	Skewness	Kurtosis	Class
1 (14).png	112.62	6.78	60.17	1.17	-0.21	1
1 (16).png	125.62	7.23	45.22	-0.26	-1.01	1
1 (18).png	131.45	7.31	88.05	0.01	-1.78	1
1 (11).png	66.39	7.38	57.88	1.07	0.26	2
1 (6).png	128.04	6.97	53.25	0.26	-1.40	2
10.png	98.41	7.59	53.83	0.44	-0.81	2
1500016.png	103.47	7.05	93.24	0.65	-1.27	3
1500023.png	116.08	7.29	68.95	-0.03	116.08	3
1500028.png	125.87	7.33	45.17	0.21	125.87	3

The results of this analysis are derived from several batik images that have been processed through texture moment feature extraction. This extraction process provides a deeper insight into the characteristics of the patterns and textures of each batik motif, allowing for a better understanding of the complexity and variation present in the batik images. Images with a high mean brightness, high entropy, large standard deviation, as well as skewness and kurtosis indicating sharp variations in pixel intensity, generally represent more complex and diverse patterns

b. Histogram of Oriented Gradients (HOG) Feature Extraction

HOG feature extraction begins by calculating the intensity changes between adjacent pixels, resulting in two main components: magnitude and orientation. Magnitude measures the degree of brightness change between pixels, while orientation indicates the direction of that change. After calculating the magnitude and orientation, the gradient orientation values in degrees are obtained and used to compute HOG features for the image, such as Mean, Standard Deviation, Skewness, and Kurtosis. The method for calculating the batik image is demonstrated using an example of a grayscale image snippet from the Kawung batik, as shown in Fig. 4.

Fig. 4. Grayscale Batik Image

The area within the yellow box is a sub-matrix (kernel) of the main image. This sub-matrix is typically selected for local operation applications, such as edge detection, using equations (1) and (2). This can be seen in Fig. 5.

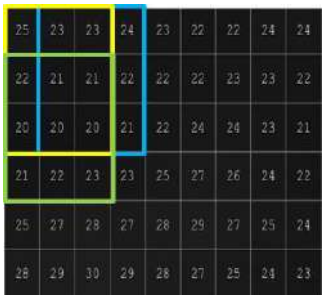


Fig. 5. Convolution Operation

The convolution operation is performed by sliding the Sobel kernel over the entire image, involving the multiplication of kernel elements with the corresponding sub-matrix pixel elements, and then summing the results. Below is the calculation for Gx (horizontal direction):

Matrix (1,1) = (−1×25) + (0×23) + (1×23) + (−2×22) + (0×21) + (2×21) + (−1×20) + (0×20) + (1×20) = -4

Matrix (1,2) = (−1×23) + (0×23) + (1×24) + (−2×21) + (0×21) + (2×22) + (−1×20) + (0×20) + (1×21) = 4

…..

Matrix (1,8) = (−1×22) + (0×21) + (1×21) + (−2×20) + (0×20) + (2×20) + (−1×21) + (0×22) + (1×23) = 1

….

After obtaining the values for Gx and Gy, the next step is to compute the magnitude, as seen in equation (3), the orientation in equation (4), and the calculation for the orientation angle $\theta'(x,y)$ for each pixel based on the gradient direction $\theta'(x, y)$derived from the orientation values. The degree of orientation is calculated using the following formula:

$$\theta'(x,y) = \begin{cases} \theta(x,y) + 180, & \text{jika } \theta(x,y) < 0 \\ \theta(x,y), & \text{jika } \theta(x,y) \geq 0 \end{cases}$$

This formula ensures that all gradient orientation values fall within the range of 0° to 180° (for unsigned gradients). If $\theta'(x, y)$ is negative, 180 is added to convert it to a positive value. If $\theta'(x,y))$ is already positive, no change is made to the orientation value, as presented in Table II.

TABLE II. Degree Orientation Results

99	108	112	108	84	59	45
72	56	45	63	69	108	156
75	84	77	61	81	130	144
70	88	0	63	153	8	0

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2225, Copyright ©2025
Submitted : December 20, 2024, Revised : January 14, 2025, Accepted : January 21, 2025, Published : January 31, 2025

99	108	112	108	84	59	45

This gradient orientation represents the direction of pixel intensity changes in an image, with higher values indicating sharp changes and lower values reflecting smooth transitions. The process of converting negative values into positive ones is done by adding 180°, ensuring all orientations fall within the 0° to 180° range.

Higher orientation values, such as 153° and 156°, indicate patterns with fine details or sharp transitions, while lower values, such as 45° and 56°, signify more regular textures. Gradient orientation plays a crucial role in detecting different batik patterns or textures, assisting in distinguishing between complex and simple motifs. Images with sharp gradients tend to have more intricate patterns, while low gradients indicate simpler patterns.

After obtaining the values for the Orientation Angle, the next step is to calculate the HOG features, which include Mean, Standard Deviation, Skewness, and Kurtosis, as outlined in equations (5), (7), (8), and (9).

TABLE III. HOG Feature Extraction

FILE_NAME	HOG Mean	HOG Std Dev	HOG Skewness	HOG Kurtosis	Class
1 (14).png	0.12	0.11	0.81	2.37	1
1 (16).png	0.14	0.08	0.34	1.97	1
1 (18).png	0.12	0.10	0.79	2.63	1
1 (11).png	0.11	0.12	0.88	2.40	2
1 (6).png	0.11	0.12	0.89	2.49	2
10.png	0.12	0.11	0.60	1.94	2
1500016.png	0.14	0.07	-0.07	2.15	3
1500023.png	0.12	0.11	0.38	1.59	3
1500028.png	0.14	0.09	0.18	1.78	3

Table III shows the results of HOG feature extraction, including Mean, Standard Deviation, Skewness, and Kurtosis of the batik images. The HOG Mean values range from 0.11 to 0.14, reflecting the average gradient intensity in the images, while the Standard Deviation indicates the variation in gradient distribution, ranging from 0.07 to 0.12. Skewness and Kurtosis reveal how the gradient distribution is skewed and sharp; most images have positive skewness, indicating a dominance of gradients in a specific direction, and high kurtosis, indicating a more concentrated gradient distribution.

The results indicate challenges in establishing direct thresholds between classes due to overlapping feature values across different classes, such as Mean, Standard Deviation, Skewness, and Kurtosis. To address this limitation, the Random Forest algorithm employs a decision tree approach that uses the average (mean) value of each class as the root of the decision tree.

This approach enables the algorithm to perform classification decisions in a more structured and balanced manner, effectively accommodating variations in feature values between classes. By using the mean value as the foundation for tree construction, Random Forest optimizes the classification process, even when direct thresholds between classes are difficult to determine.

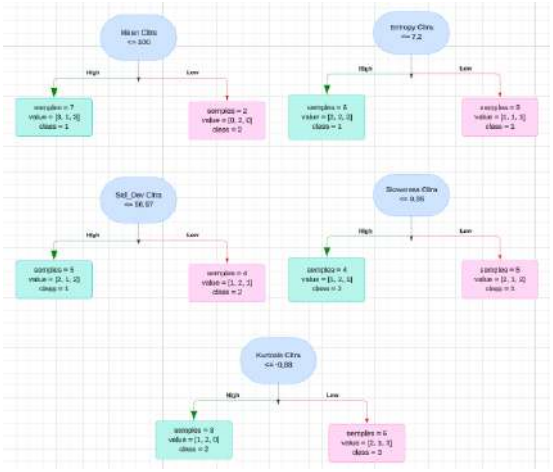
C. Random Forest Algorithm Classification Stage

Next, classification is performed using the Random Forest method to determine the correct and incorrect predictions as well as the precision [17]. In this stage, decision trees are created as follows:

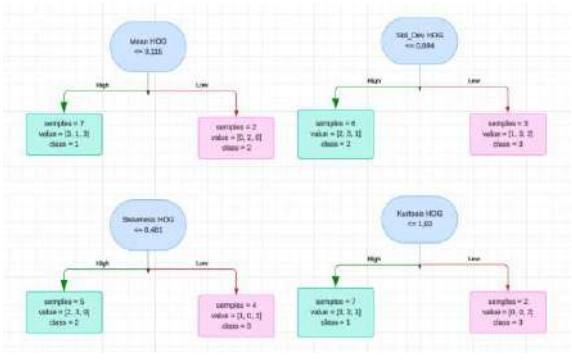
a. *Create thresholds based on the average of each feature value. If the value of a feature is lower than the average, it is categorized as "low," and if it is higher than the average, it is categorized as "high."*
b. *Performing the decision tree construction process to map classification results into more structured decisions. Each tree in the Random Forest performs classification and produces a class prediction.*

The predictions from all trees are then aggregated through a voting process, where the class with the highest number of predictions is selected as the final classification result [18]. This approach ensures that Random Forest can overcome the limitations of directly determining thresholds between classes while improving classification accuracy through the combined decisions of multiple trees.

Fig. 6 illustrate the results of the Random Forest construction.



(a)



(b)

Fig. 6. (a) Random Forest Results for Grayscale Image (b) Random Forest Results for HOG Image

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2225, Copyright ©2025
Submitted : December 20, 2024, Revised : January 14, 2025, Accepted : January 21, 2025, Published : January 31, 2025

Since each class consists of three images, accuracy is calculated based on the number of correct predictions (high samples) for each class. A prediction is considered correct if the number 3 appears in the image, representing the number of samples in each class. This strategy helps evaluate the accuracy of the Random Forest algorithm in classifying batik images in a more systematic manner as shown in Table.

TABLE IV. Random Forest (RF) Results

No.	NAMA_FILE	Class	RF Predict
1	1 (14).png	1	1
2	1 (16).png	1	1
3	1 (18).png	1	1
1	1 (11).png	2	2
2	1 (6).png	2	2
3	10.png	2	2
1	1500016.png	3	3
2	1500023.png	3	3
3	1500028.png	3	3

D. System Evaluation Stage

The system evaluation is performed by calculating the number of test data correctly classified. based on the feature extraction values from Texture, HOG, and their combination. Table V shows the accuracy results for each experiment.

TABLE V. System Accuracy Results

Feature Extraction	Correct Prediction	Incorrect Prediction	Accuracy
Texture (Moment)	113	25	81,88%
HOG	109	29	78,99%
Combined	119	19	86,23%

Based on the feature extraction methods using Texture (Moment), HOG, and their combination with Random Forest classification, the system's performance yields the following results:

The system using *Texture (Moment)* feature extraction achieved an accuracy of 81.88%. Although this method is capable of capturing texture information, its effectiveness is limited in identifying more complex batik patterns, which restricts its classification performance. The system with *HOG feature extraction* produced a higher accuracy of 78,99%. HOG excels at capturing gradients and orientations, which are crucial for recognizing patterns and contours in batik motifs. Additionally, HOG is more effective in handling sharper and clearer texture variations in batik images.

The *combination of Texture (Moment) and HOG* achieved the highest accuracy of 86.23%. This result demonstrates that combining the two methods significantly enhances classification accuracy. The combination leverages the strengths of both methods: Texture (Moment) captures fine pattern details, while HOG captures gradient orientations.

In addition to calculating the accuracy on the test data, this study also calculates the precision value of each class as presented in table VI. precision calculations are calculated as in the accuracy calculation but are carried out on each class of training data images.

TABLE VI. Precision Results

Class	Correct Prediction	Incorrect Prediction	Precision
1	41	5	89,13%
2	44	2	95,65%
3	33	13	71,74%

The table V presents the classification results for three batik image classes: class 1 (kawung batik image), class 2 (megamendung batik image), and class 3 (parang batik image). The "Class" column identifies the image categories, while the "Correct Predictions" and "Incorrect Predictions" columns record the number of correct and incorrect classifications. Precision is calculated by dividing the number of correct predictions by the total predictions for each class. Class 2 achieved the highest precision at 95.65%, indicating the most accurate classification. Class 1 recorded a precision of 89.13%, which is also relatively high. Class 3 had a lower precision of 71.74%, indicating a higher rate of misclassification. Overall, the system performed best on class 2 and worst on class 3.

And here are the results of the confusion matrix:

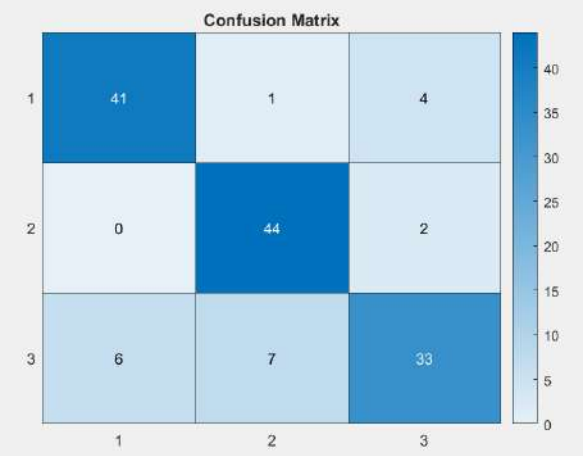


Fig. 7. (a) Confusion Matrix

The confusion matrix in the image above illustrates the model's performance in classifying three batik motifs: (1) Kawung, (2) Megamendung, and (3) Parang. The Kawung class has 41 data points correctly classified, but 1 data point was misclassified as Megamendung, and 4 data points were misclassified as Parang. The Megamendung class demonstrates the best performance with 44 data points correctly classified, while only 2 data points were misclassified as Parang, and none were misclassified as

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2225, Copyright ©2025
Submitted : December 20, 2024, Revised : January 14, 2025, Accepted : January 21, 2025, Published : January 31, 2025

Kawung. On the other hand, the Parang class has 33 data points correctly classified, but 6 data points were misclassified as Kawung, and 7 were misclassified as Megamendung. Overall, the model is more accurate in recognizing the Megamendung class compared to the others; however, there are still notable misclassifications, especially in the Parang class.

IV. CONCLUSION

The conclusions of this study indicate that the combination of HOG and texture moment feature extraction, followed by classification using the Random Forest algorithm, can deliver good results in recognizing Indonesian batik motifs. In the test results, the HOG method achieved an accuracy of 78.99%, while the Texture Moment method reached an accuracy of 81.88%. The combination of both feature extraction methods improved the system performance, achieving the highest accuracy of 86.23%, an increase of 4.65% compared to the texture method and an 8% improvement over the HOG feature extraction method. This demonstrates that both texture feature extraction and HOG can effectively recognize Batik motifs such as Parang, Kawung, and Megamendung.

V. SUGGESTIONS FOR DEVELOPMENT

The research results provide quite high accuracy with an accuracy of 86.23%. However, this accuracy can be improved by eliminating the resizing step which can result in loss of object details. In addition, other developments that can be done are by extracting features with other methods such as Local Binary Pattern (LBP), Gabor filter, Wavelet and so on which are texture-based. Research in Batik recognition is still wide open because of the many types of batik that exist both in Indonesia and abroad. This provides challenges for the future, especially in the field of image processing and computer vision.

REFERENCES

[1] S. Akbar, R. A. Fahreza, F. Ferdianto, D. F. Wulandari, and E. W. Astuti, "Klasifikasi Pes Planus Menggunakan Ekstraksi Fitur HOG dan BoF dengan Random Forest," *Journal Of Electronics And Instrumentation*, vol. 1, no. 1, pp. 1-10, 2024.

[2] S. D. Amanda, A. K. Ramadina, and S. P. Wulandari, "ANALISIS FAKTOR-FAKTOR KUALITAS PENDIDIKAN DI INDONESIA TAHUN 2023 MENGGUNAKAN ANALISIS FAKTOR," *SINDORO CENDEKIA PENDIDIKAN*, vol. 2, no. 1, pp. 50-60, 2023.

[3] T. H. Andika and N. S. Anisa, "Sistem Identifikasi Citra Daun Berbasis Segmentasi Dengan Menggunakan Metode K-Means Clustering," *Aisyah Journal of Informatics and Electrical Engineering*, vol. 5, no. 2, pp. 33-45, 2020.

[4] S. P. Bagu, N. A. Tanumihardja, and Michelle, "VISUALISASI BATIK PARANG YOGYAKARTA," *Jurnal Penelitian Multidisiplin*, vol. 9, no. 1, pp. 21-35, 2024.

[5] S. P. Bagu, N. A. Tanumihardja, and Michelle, "VISUALISASI BATIK PARANG YOGYAKARTA," *BEGIBUNG: Jurnal Penelitian Multidisiplin*, vol. 9, no. 1, pp. 21-35, 2024.

[6] K. R. Giri, W. Y. Putra, G. A. Utami, N. W. Utami, and D. G. Putra, "IDENTIFIKASI ARSITEKTUR KOLONIAL DALAM POLA TRANSFORMATIF REGIONALISME PADA RUMAH MAKAN DAPOER PEMUDA DI DENPASAR," *Prosiding Seminar Nasional Manajemen, Desain & Aplikasi Bisnis Teknologi (SENADA)*, vol. 2, no. 1, pp. 102-115, 2024.

[7] R. Handoko, D. Ramadhansyah, A. Asrofiq, A. Rahmaddeni, and Y. Yunefri, "ANALISIS SENTIMEN ULASAN PENUMPANG MASKAPAI PENERBANGAN DI INDONESIA DENGAN ALGORITMA RANDOM FOREST DAN KNN," *Jurnal Sistem Informasi*, vol. 10, no. 1, pp. 24-38, 2024.

[8] G. S. Hussein, S. Elseuofi, W. H. Dukhan, and A. H. Ali, "A Novel Method for Banknote Recognition Using a Combined Histogram of Oriented Gradients and Scale-Invariant Feature Transform," *Information Sciences Letters*, vol. 8, no. 4, pp. 200-215, 2023.

[9] I. Barkiah and A. Y. Ida, "Overcoming Overfitting Challenges with HOG Feature Extraction and XGBoost-Based Classification for Concrete Crack Monitoring," *INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS*, vol. 7, no. 3, pp. 110-125, 2023.

[10] A. Ikram, N. Mutmainna, J. Khasmawati, D. Wahyuli, I. W. Sudarsana, et al., "TRANFORMASI CALANTHE TRIPLICATA UNTUK BRANDING UNIK MOTIF BATIK SULAWESI TENGAH," *Jurnal Ilmiah Matematika dan Terapan*, vol. 6, no. 2, pp. 55-68, 2022.

[11] T. L. Kaunang, J. J. Tinangon, and V. Z. Tirayoh, "ANALISIS PENERAPAN SISTEM PENGENDALIAN MANAJEMEN UNTUK MENINGKATKAN KINERJA PERUSAHAAN PADA PT. PERUSAHAAN LISTRIK NEGARA (PERSERO) UNIT LAYANAN PELANGGAN MANADO SELATAN," *Jurnal EMBA*, vol. 8, no. 3, pp. 77-90, 2021.

[12] A. Khadafi and M. Iqbal, "Implementation of Random Forest for Motif Classification Based on Sift," *Jurnal Mantik*, vol. 5, no. 4, pp. 149-162, 2022.

[13] M. Muchtar and R. A. Muchtar, "Ekstraksi Fitur LBP Dan Moment Invariant Untuk Identifikasi Jenis Kayu Menggunakan SVM," *INTEGRASI FITUR WARNA, TEKSTUR DAN RENYI FRAKTAL UNTUK KLASIFIKASI PENYAKIT DAUN KENTANG MENGGUNAKAN LINEAR DISCRIMINANT ANALYSIS*, vol. 2, no. 2, pp. 140-150, 2024.

[14] N. Larasati, S. Devella, and M. E. Rivan, "Pengenalan ASL Menggunakan Metode Ekstraksi HOG Dan Klasifikasi Random Forest," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 7, no. 1, pp. 45-58, 2021.

[15] M. Meiriyamaa and Sudiadia, "Penerapan Algoritma Random Forest Untuk Klasifikasi Jenis Daun Herbal," *JTSI*, vol. 9, no. 2, pp. 22-35, 2022.

[16] N. M. Mitayani and D. A. Astuti, "PENGEMBANGAN BATIK KHAS KEBUDAYAAN INDONESIA DENGAN PENGARUH TEKNOLOGI," *e-journal unmas*, vol. 6, no. 3, pp. 110-125, 2024.

[17] A. Gumilang dan S. Agustin, "Deteksi Kepribadian Melalui Margin Pada Tulisan Tangan Menggunakan Random Forest," *Jurnal Inovtek Polbeng - Seri Informatika*, vol. 9, no. 1, pp. 311-325, 2024.

[18] S. Agustin, A. Sari, dan Ernawati, "Pengenalan Huruf Braille menggunakan Radially Average Power Spectrum dan Geometri," *Jurnal Inovtek Polbeng - Seri Informatika*, vol. 8, no. 1, pp. 25-35, 2023.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2225, Copyright ©2025
Submitted : December 20, 2024, Revised : January 14, 2025, Accepted : January 21, 2025, Published : January 31, 2025

Clustering OKU Timur Script Images using VGG Feature extraction and K-Means

Liu Toriko[1], Susan Dian Purnamasari[2]*, Yesi Novaria Kunang[3], Ilman Zuhri Yadi[4], Andri[5]
Intelligent Systems Research Group, Faculty of Science Technology [1], [2], [3], [4], [5]
Universitas Bina Darma
Palembang, Sumatera Selatan, Indonesia
liutoriko031102@gmail.com[1], susandian@binadarma.ac.id[2], yesinovariakunang@binadarma.ac.id[3],
ilmanzuhriyadi@binadarma.ac.id[4], andri@binadarma.ac.id[5]

***Abstract*— This study focuses on the utilization of clustering models to group manuscript images from the OKU Timur region based on specific characteristics. OKU Timur is rich in cultural heritage, including a unique writing system known as the OKU Timur script. The development of intelligent systems technology can be employed to recognize the OKU Timur script. For this purpose, a dataset of OKU Timur script is needed, which will later be used for classifying script images. One of the challenges in preparing the dataset is grouping a large number of script image samples according to the number of characters. A proposed solution in this research is to automatically group script images by applying the K-Means algorithm. The dataset comprises 2,280 images, representing 19 characters and 228 variations with different diacritics. Features are extracted using the VGG16 model, which are then clustered with the K-Means algorithm. Clustering performance is evaluated based on the percentage of correctly grouped characters. For 19 groups (character count), the model achieves an accuracy of 82.6%. For 228 groups (variations and diacritics), it correctly groups 48.16% of characters. Despite the challenges, the results demonstrate the model's potential for further refinement. This study's contribution lies in introducing an efficient clustering approach for cultural manuscripts, supporting digital preservation, and advancing automatic recognition of the OKU Timur script. These efforts aim to preserve the script for future generations.**

Keywords*— *OKU Timur Script, Clustering, K-Means, VGG16 Model, Manuscript Images

I. INTRODUCTION

OKU Timur Regency is located in South Sumatra Province and borders Ogan Ilir Regency, Ogan Komering Ulu Regency, South Ogan Komering Ulu Regency, and Lampung Province. This area is rich in historical sites, one of which is Aksara, which refers to letters or symbols that function as symbols of sound (phonemes) [1]. Aksara is also known as a "writing system." Over time, aksara has developed into a visual symbol system that appears on various media, such as paper, stone, trees, wood, or cloth, to convey elements of expression of a language [2]. Currently, the development of information technology is rapid and covers many aspects of life. This progress has led to the availability of extensive and diverse data, covering industries, the economy, science, technology, and various other fields [3]. The OKU Timur Script is an integral part of a unique cultural heritage, reflecting the identity and history of the people in the area.

Like many other ancient writing systems, the OKU Timur script holds significant cultural and historical value [4], [5]. The digital preservation of OKU Timur script is very important for several reasons. First, digitization allows wider access for future generations to learn and understand these characters, which may no longer be actively used in everyday communication [6], [7]. Second, by preserving the script in digital format, we can protect knowledge and culture that may be lost over time, especially as the number of speakers and active users of the script decreases [8].

In an effort to preserve the OKU Timur script, it is necessary to create intelligent applications that can automatically recognize the script, similar to those that researchers have previously developed for other types of scripts. [9], [10]. The initial effort to develop intelligent applications requires a dataset to train the model to recognize variants of the script. A specific challenge in preparing a large-scale dataset is the labor-intensive and error-prone process of manually sorting thousands of example images [11], [12]. To address this, the dataset preparation can be assisted by utilizing clustering algorithms that can automatically group character images into sets based on similarities [13]. In this context, clustering—through techniques such as the K-Means algorithm—plays an important role in the preservation process [14], [15]. Clustering provides a systematic way to organize and group images of OKU Timur script characters. By grouping characters based on visual similarities, we can identify and categorize characters more efficiently.

Clustering is the process of grouping data into various clusters where similar objects are placed in one cluster, while dissimilar objects are placed in different clusters [16]. Each cluster contains data that is as similar as possible to each other, and the degree of similarity is usually measured based on distance. Therefore, each object in one cluster must have similar characteristics, while objects in other clusters must have different characteristics. The goal of this clustering process is to organize the data into several groups, so that similar data is placed in one group while different data is placed in another group [17]. The basis of the clustering concept is to group a number of objects into clusters, with the aim of forming groups that have high similarity among the objects within them and

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2292, Copyright ©2025
Submitted : October 7, 2024, Revised : November 4, 2024, Accepted : November 25, 2024, Published : January 31, 2025

significant differences from the objects in other groups [18].

The effectiveness of the K-Means algorithm in clustering characters, as demonstrated in several previous studies, has inspired researchers to adopt this approach. For instance, the K-Means algorithm has been used to cluster Javanese script images [15], K-Means and spectral clustering for grouping Odia character images [19], clustering handwritten Indic scripts [14], and hierarchical K-Means for clustering Balinese script images [20]. Moreover, the K-Means algorithm is highly popular due to its ability to quickly cluster large amounts of data, including outliers [21], [22]. It also has relatively low computational complexity, making it efficient for application to large datasets [23], [24]. In this study, the researchers use the K-Means algorithm with a feature extraction process utilizing the VGG model.

The novelty of this research lies in the fact that there has been no prior study specifically focusing on the development of a clustering model for OKU Timur script images. The absence of prior research presents a unique challenge in finding suitable information for the development of this model. The development of this clustering model is necessary to make the system more efficient and accurate in grouping OKU Timur script images and to preserve the digital culture of using the OKU Timur script for future generations. Based on the existing issues, the researcher is conducting a study titled "Clustering OKU Timur script images using VGG feature extraction and K-Means"

II. RESEARCH METHODOLOGY

In the development of the clustering model for OKU Timur script images, this can be illustrated in the flowchart in Figure 1. The stages involve data collection, preprocessing, feature extraction, clustering model training, and evaluation of clustering results.

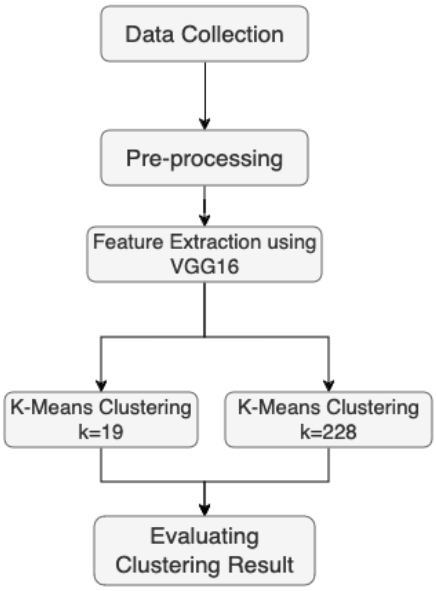


Figure 1. Flowchart

A. Data Collection

The data on the OKU Timur script was obtained from the Traditional Leader and Script Expert of OKU Timur. The data collection was previously conducted by another team from the ISRG (Intelligent System Research Group), and after the data was collected, there were 228 characters of the OKU Timur script. Before performing clustering, several processes must be carried out, such as preprocessing. At this stage, the data collection process for the OKU Timur script has involved local traditional leaders and script experts, covering information about the usage and characteristics of the script. Previously, the ISRG (Intelligent System Research Group) team had collected preliminary data regarding the OKU Timur script, including its use in cultural contexts and technical aspects.

Figure 2. Example filled Questionnaire for the OKU Timur Script

Additional data from the traditional leaders and script experts aims to deepen the existing understanding. The obtained data is then processed into a questionnaire to be distributed to respondents. The image below is an example of the OKU Timur script questionnaire that will be filled out by respondents, consisting of 12 punctuation marks of the OKU Timur script.

After being filled out by the respondents, the questionnaire aims to identify the uniqueness of each respondent's writing. This questionnaire contains 228 characters of the OKU Timur script and was completed by 102 respondents. Figure 2 is an example of the filled questionnaire from the respondent.

B. Preprocessing

In the preprocessing stage conducted on the raw data, the goal is to improve data quality so that clustering can be performed more effectively and yield good groups. The processes involved include downloading and extracting data, then saving, printing the path to the directory, and organizing, checking, and processing the list of image files to prepare them for the K-Means model training. Preprocessing is an important step in data clustering. In this study, 10 sample questionnaire forms were used, with a total of 228 character image types, resulting in a total of 2280 images collected. The number of images taken was then inputted as 2280 images. This process is conducted to prepare the data for use by the model. Several stages are performed during preprocessing, including: data

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2292, Copyright ©2025
Submitted : October 7, 2024, Revised : November 4, 2024, Accepted : November 25, 2024, Published : January 31, 2025

downloading and extraction zip data.

This research begins by downloading data files from an external URL using the 'wget' command. This command allows the system to automatically download files from the internet and save them in a specified location; in this case, the downloaded file is aksara_okut.zip

Next, the file extraction process is carried out using the zipfile module in Python. This module is used to read and extract content from ZIP files. The extraction process involves opening the downloaded file, after which all contents are extracted to a specified directory ('/content/'). After that, to ensure the extraction is successful, the program will display a list of files extracted from that folder.

C. Feature Extraction using VGG16 model

To transform raw image data into a more manageable and informative format, which facilitates effective clustering and analysis, feature extraction is performed. Feature extraction plays a critical role in the process of image clustering because it can reduce dimensions. Raw image data is typically high-dimensional and complex, making it difficult to process directly. Feature extraction reduces the dimensionality by transforming images into a lower-dimensional space while retaining important information. This simplifies the clustering process and can improve clustering performance. Effective feature extraction enhances the quality of clusters formed by algorithms like K-Means [25].

In feature extraction, a pre-trained *VGG16* model trained on the ImageNet dataset will be used. VGG is a convolutional neural network model for image recognition proposed by the Visual Geometry Group at the University of Oxford, where VGG16 refers to a VGG model with 16 weight layers [26]. On the VGG16 model but with the final layer (the classification layer) removed. Instead of utilizing the final fully connected layers for classification, features are extracted from one of the last convolutional layers (typically one of the layers before the global average pooling layer). The output from this layer is a high-level representation of the image features in a lower-dimensional space.

For more details on the feature extraction process, refer to Figure 3. The workflow begins with importing the necessary libraries, specifically Keras, particularly `keras.applications.vgg16`, along with any other required libraries. The next step is importing modules for using the VGG16 model.

```
model = VGG16(weights='imagenet', include_top=False)
```

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2292, Copyright ©2025
Submitted : October 7, 2024, Revised : November 4, 2024, Accepted : November 25, 2024, Published : January 31, 2025

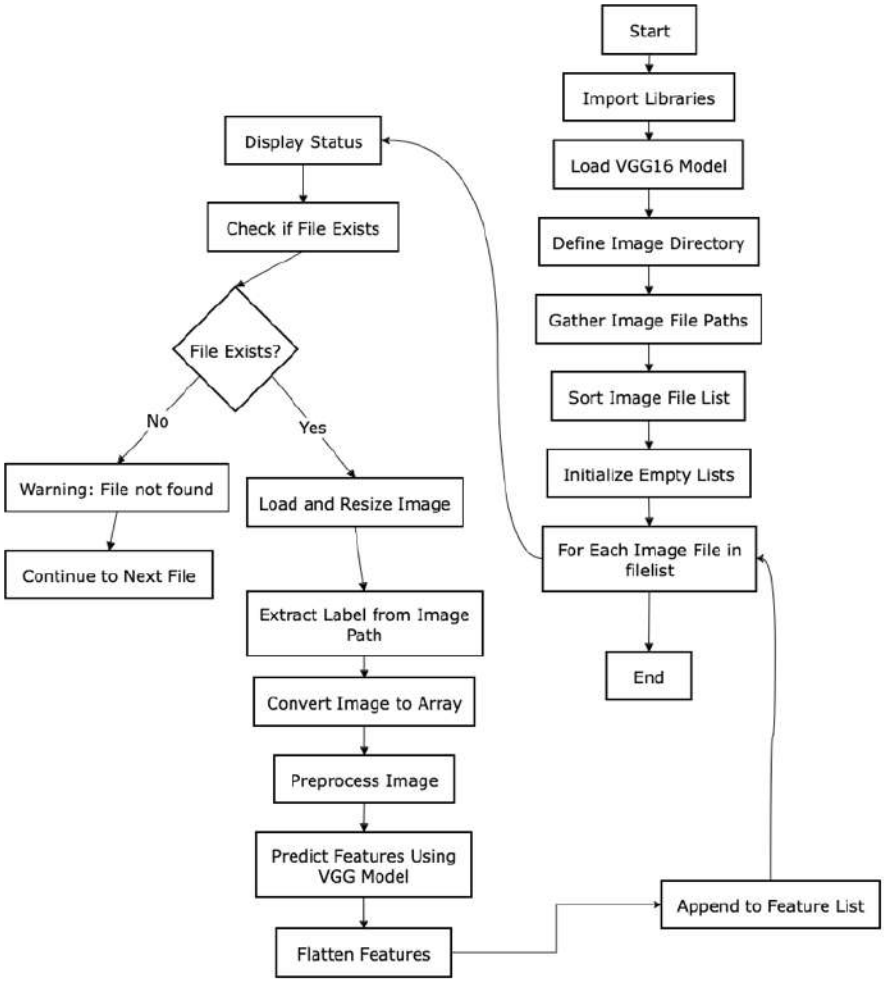


Figure 3. Workflow for feature extraction using the VGG16 model.

At this stage, the task is to create the VGG16 model without the final classification layer, so it can be used for feature extraction from images, such as clustering. In this process, there will be no output displayed because the VGG16 model here serves only as an additional feature to ensure the code runs smoothly without errors.

The next step involves preprocessing all files in the directory to change the file dimensions to match the input dimensions of the VGG16 model. Then, all images have their features extracted and stored in a flattened feature list until all images have been processed. The final result of this feature extraction process produces images with smaller sizes and dimensions corresponding to the feature extraction that was performed.

D. K-Means Clustering

Next, the stage of determining the number of clusters is conducted, as the researchers identified approximately 19 base characters of the OKU Timur script. Each character has 12 variants diacritics, resulting in a total of 228-character variants. Therefore, in the clustering, two types of cluster numbers are used: 19 and 228. These numbers are based on the number of base characters and the variant characters based on the variants.

After the data undergoes feature extraction using the VGG16 model, the clustering process is then performed using K-Means.

```
# Variables
number_clusters = 228

# Clustering
kmeans = KMeans(n_clusters=number_clusters, random_state=0).fit(np.array(featurelist))
```

The model suitable for clustering the OKU Timur script images is K-Means because it is simple, fast, efficient, and flexible in determining the number of clusters [21], [22]. In the Clustering Model Training, a K-Means model is created with the number of clusters specified by 'number_cluster' and 'random_state=0' for consistent results. The K-Means model is then trained on the image features stored in 'featurelist', which have been converted into a NumPy array. The purpose of this is to group images based on character similarity into the predetermined clusters.

E. Evaluating Clustering Result

To evaluate the clustering results, two approaches are used:

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2292, Copyright ©2025
Submitted : October 7, 2024, Revised : November 4, 2024, Accepted : November 25, 2024, Published : January 31, 2025

Check Cluster Consistency and Calinski-Harabasz Index.
Check Cluster Consistency: Review sample images from each cluster to ensure they are meaningful and consistent [27]. The main objective of this clustering process is to automatically group similar characters, so it is necessary to evaluate the homogeneity of the resulting clusters. To assess this, we calculate how many instances of the same character are correctly grouped together. An analysis is conducted to determine whether an ideal cluster should consist of a single character. If multiple characters are observed within a single cluster, it will be labeled as a cluster with the largest number of character variants or the major cluster. Other characters that fall into this cluster are considered misclassified characters. The final result will show the total number of character images that are accurately clustered compared to the total number of image data overall.

Calinski-Harabasz Index (CHI), also known as the Variance Ratio Criterion, is used to assess the quality of the cluster partitioning generated by the clustering algorithm [28]. This index is calculated by comparing the intra-cluster variance and the inter-cluster variance. The steps to calculate the Calinski-Harabasz Index are presented in equation (1).

$$\left[\text{CHI} = \frac{B_k/(k-1)}{W_k/(n-k)}\right] \quad (1)$$

where:

- (B_k) = *between-cluster variance*
- (W_k) = *within-cluster variance*
- (k) = *Number of clusters*
- (n) = *Total number of objects/data*

This stage involves analyzing the clustering results, observing how the data is grouped, and using the resulting dataset to train deep learning and transfer learning models for classifying the OKU Timur script images.

III. RESULT AND DISCUSSION

Figure 4 shows a visualization of how the image data consisting of 2,280 images undergoes feature extraction using the VGG16 model. The output from the feature extraction process comes from the last convolutional layer, resulting in an output tensor of size (7, 7, 512) for input images sized 224x224. The extracted feature tensors are then grouped using the K-Means method for the number of clusters = 19 and the number of clusters = 228.

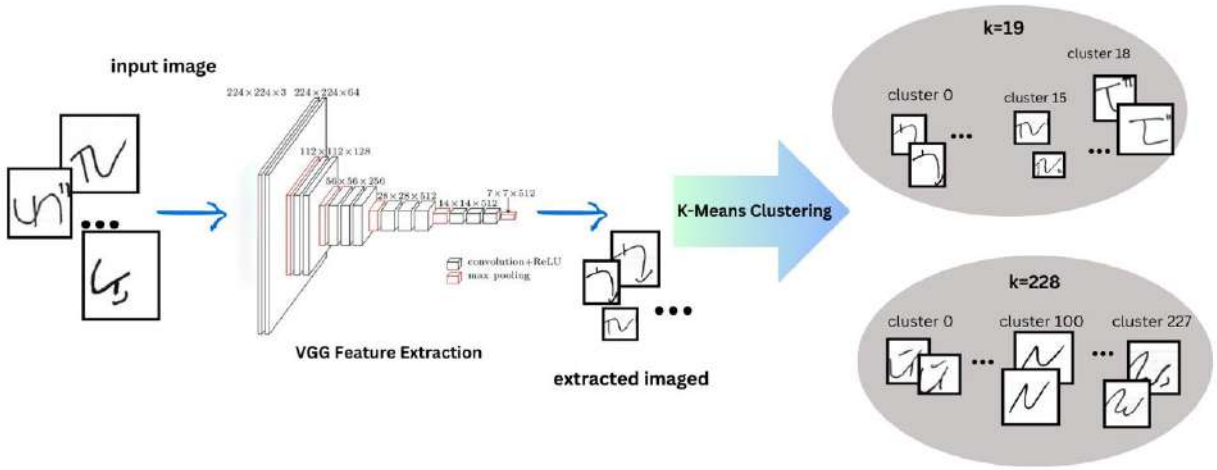


Figure 4. Image Feature extraction process and clustering image results with K-Means

A. *Clustering Result*

After performing clustering, the next step is to examine the results. The visual representation of images in each cluster provides strong visual confirmation of how these images are grouped by the model. This can help in understanding and validating whether the clustering has successfully separated the characters of the OKU Timur script into the predetermined groups. The clustering results for 19 clusters are presented in Table 1, while the clustering results for the number of classes = 228 are presented in Table 2.

TABLE I. CLUSTERING RESULTS UNTUK NUMBER CLUSTER= 9

Label cluster	Number of images	Mayor Images	Number of Major Images	Percentage of Major Class	Mis-Cluster
0	115	S	115	100%	-
1	182	R	130	71.4%	L, Ny, A
2	111	W	111	100%	-
3	211	J	116	55%	N, W, K
4	117	Ny	117	100%	-
5	109	Ng	109	100%	-
6	68	G	68	100%	-
7	112	L	112	100%	-
8	172	P	92	53.5%	L, R, Y, H
9	106	D	106	100%	-

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2292, Copyright ©2025
Submitted : October 7, 2024, Revised : November 4, 2024, Accepted : November 25, 2024, Published : January 31, 2025

10	121	Y	104	86%	B, P, M
11	121	M	109	90%	P, Ny
12	140	B	113	80.71%	H, L, N, P
13	51	C	51	100%	-
14	138	K	92	66.7%	N, L
15	100	A	92	92%	L, Y
16	70	C	70	100%	-
17	119	H	64	53.8%	N, G, Ng
18	117	T	112	95.7%	D, J, Ng
Total	**2280**		**1883**		

The description of the table 1, 2 are as follows :

- *Class Label is the name of the folder representing the class.*
- *Number of Images is the total number of sample images available in each class category for model testing.*
- *Major refers to the typeface and punctuation in each class.*
- *Number of Major Images is the highest number of images owned by a single class in the directory.*
- *Percentage of Major Class is the proportion of the number of images in the majority class to the total number of images across all classes, expressed as a percentage.*
- *Mis-cluster refers to the grouping error where an item is placed in the wrong group.*

In the process of clustering using the entire dataset, which consists of 2,280 images, no data partitioning is performed. This is because in clustering, the concept of separation between training, validation, and testing, as used in supervised learning models, is not implemented in the same way. Clustering is an unsupervised learning method, where no labels are used to provide context to the data. Therefore, traditional partitioning is not necessary [29].

For cluster consistency with the number of clusters = 19 (Table 1, table 3), it can be seen that the clustering model accurately grouped 1,883 images, which represents the number of major characters in one cluster. Major images represent those that are accurately clustered. Overall, the clustering model achieved an accuracy of 82.6% in grouping the images. Upon closer inspection, there are 3 fundamental OKU Timur characters with significant grouping errors of less than 60%, namely characters J, P, and H. On the other hand, the clustering results with the number of clusters = 228 show a consistency value of 48.16% (table 3) for grouping images of the same character into the same cluster.

TABLE II. CLUSTERING RESULTS UNTUK NUMBER CLUSTER= 228

Label cluster	Num. of images	Mayor Images	Num. of Major Images	%of Major Class	Mis-Cluster
0	26	Bang	16	61.5%	Rang, Ri
1	13	Hu	6	46.2%	Ha
2	5	Ngah	2	40%	Ngai, Ngau
3	6	Rai	2	33.3%	R, Ran
4	19	Wa	6	31.6%	Wi, War
5	21	Yu	10	47.6%	Ya, Yi, Yan
6	25	Ai	15	60%	Ao, An
7	9	Nyan	5	55.6%	Nyu, Nyar
8	5	Ng	3	60%	Ngan, Ngah
9	10	Jau	6	60%	Jan, Jar
10	8	Dan	3	37.5%	Do, Dau
11	9	Di	5	55.6%	Dah, Da
12	2	Car	1	50%	Co
13	11	P	8	72.7%	Pan
14	7	J	5	71.4%	Nyan
15	20	Lau	10	50%	Le, Lo
16	15	Ki	8	53.3%	Kar, Ka
17	34	Gi	9	26.5%	Gah, Gan
18	12	Ngar	7	58.3%	Ngang, Ngi
19	16	Pang	7	43.8%	Par, Pa, Po
20	8	T	4	50%	Ta, Tar, Tang
21	8	H	8	100%	-
22	6	Ro	4	66.7%	Re, Rar
23	13	Car	5	38.5%	Cang, C, Can
24	3	D	2	66.7	Dah
25	20	L	8	40%	La, Lar, Lan
26	14	A	7	50%	Ah, Ar
27	7	Sai	3	42.9%	Sah, Sau
28	3	Ngai	2	66.7%	Ngan
29	9	Ngang	4	44.4%	Ngan, Ngah, Nga
30	2	Ta	1	50%	Tan
31	16	Kai	6	37.5%	Kan, Ka, K, Ke
32	10	Rau	6	60%	Ro, Rah
33	17	Ng	7	41.2%	Ng, Ngau, Ngah
34	9	G	3	33.3%	Gah, Gang, Gu
35	23	Jai	7	30.4%	Jah, J, Jang
36	5	D	3	60%	Da, Dah
37	21	Y	7	33.3%	Ya, Yah, Yan
38	13	Mau	5	38.5%	Ma, Mah
39	4	Ca	2	50%	Cau, Cu
…	…	…	…	…	…
…	…	…	…	…	…
220	19	Wang	6	31.6%	Wa, Wah, War
221	1	An	1	100%	-
222	6	Ja	3	50%	Jan, Jang, Je
223	1	War	1	100%	-
224	10	Nyan	5	50%	Ny, Nyai, Nyang
225	11	Do	6	54.5%	Da, Dar, De
226	10	Hai	6	60%	Han, Hau, Ha
227	6	Nye	3	50%	Nya, Nyah, Nyai
Total	**2280**		**1098**		

B. Evaluation and Discussion

Table 3 shows a summary of the performance metrics from the image clustering process. For the consistency results, it can be observed that the number of character variations and the variation of characters in 228 clusters makes the clustering process challenging. Similarly, for the values of the Calinski-Harabasz Index (CHI), which indicate the quality of the cluster separation produced by the clustering model. For a CHI of 39.85 for 19 clusters, it indicates that with 19 clusters, the model produces a fairly good separation. A higher value suggests that the between-cluster variance is relatively large compared to the within-cluster variance, meaning the formed clusters are well-defined and separated. On the other hand, a CHI of 7.28 for 228 clusters is lower compared to the value for 19 clusters. This may indicate that with more clusters (228), the separation between

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2292, Copyright ©2025
Submitted : October 7, 2024, Revised : November 4, 2024, Accepted : November 25, 2024, Published : January 31, 2025

clusters becomes less clear, or that the clusters become too small, which can lead to redundancy or similarity between clusters. In general, a higher CHI value indicates better cluster quality. In this case, the model with 19 clusters appears to produce better separation compared to the model with 228 clusters

TABLE III. PERFORMANCE METRICS OF THE CLUSTERING RESULTS FOR OKU TIMUR SCRIPT

Number of cluster	Consistency	CHI
19	82.6%	39.85
228	48.16%	7.28

Several factors contribute to the significant errors in the clustering process in the context of clustering OKU Timur scripts, including the similarity in character shapes, particularly diacritics., which makes it difficult for the model to accurately cluster the glyph images (see example in Figure 5). Additionally, another factor is issues with the character's appearance that reduce the clarity of the letters in the writing. Furthermore, the quality of the images produced by the respondents is still lacking.

Figure 5. Examples of variant diacritic in OKU Timur script

Handling data quality is crucial for effective clustering and accurate results, particularly in tasks like clustering scripts or character recognition. To improve image data quality, several strategies can be employed for the development of future research. Data preprocessing techniques such as noise reduction using Gaussian or median filters can enhance clarity, while normalization of pixel values ensures uniform intensity. Resizing images to a fixed dimension is also important for consistency. Image enhancement methods, including contrast adjustment and binarization, can make characters more discernible, and sharpening techniques can enhance edge visibility. Additionally, data augmentation through transformations like rotation and scaling can create variations that help the model generalize better despite poor quality. Automated quality assessments and manual reviews can filter out low-quality images before clustering. Utilizing pre-trained models through transfer learning can adapt to varying image qualities, and ensuring high-quality image capture in future data collection is vital. Finally, outlier removal after initial clustering can mitigate the skewing effects of significantly different images. By implementing these strategies, the impact of poor image quality can be minimized, enhancing the effectiveness of clustering algorithms and leading to more accurate and meaningful results.

The main contribution of this research is the digital preservation of the OKU Timur script and the development of an automatic recognition system that can facilitate access and usage of the script in the future. This research paves the way for further studies in image processing and character recognition, which are crucial in preserving cultural heritage. As a recommendation, future research could explore alternative algorithms such as DBSCAN and Agglomerative Clustering, which may be more effective in handling irregular data distributions. Additionally, the use of optimization methods such as Grid Search or Random Search can help in determining better parameters for the clustering model.

To speed up processing time, the implementation of parallel processing or the use of GPUs can be a concrete solution. By leveraging this technology, it is expected that the time required to train the model and analyze the results can be minimized, allowing for the processing of larger and more complex datasets with greater efficiency. Overall, this research provides a strong foundation for further development in the automatic recognition of local cultural scripts and demonstrates the importance of technology in effectively preserving cultural heritage.

However, the challenges faced in clustering character images are quite complex, including variations in image quality, differences in writing styles, and the complexity of the character forms themselves. These uncertainties and variations can cause difficulties in the clustering process, where characters that appear similar may be expressed incorrectly, and vice versa. Therefore, an effective clustering development model not only helps in organizing data but also contributes to the digital preservation efforts of OKU Timur script. In this way, clustering serves not only as a technical tool but also as a link between technology and cultural preservation, ensuring that this heritage remains alive and accessible to future generations.

IV. CONCLUSION

This research successfully implemented a clustering model to group OKU Timur script images using feature extraction from the VGG16 model and the K-Means algorithm. The test results indicated that the model could group the majority of images with satisfactory accuracy, achieving 82.6% consistency in grouping major characters with 19 clusters. Evaluation metrics such as the Calinski-Harabasz Index provided insights into the quality of the cluster separation produced. However, challenges arose due to the similarity in character shapes and suboptimal image quality, leading to misclassification of some characters. Improvement strategies, such as applying image enhancement techniques and data augmentation, could be implemented to enhance future results. This research makes a significant contribution to the fields of image processing and character recognition, while also opening up opportunities for further studies to develop more effective and accurate clustering methods.

REFERENCES

[1] E. E. Panjaitan and N. Siregar, "THE IMPORTANCE OF LEARNING INDONESIAN LANGUAGE IN PRIMARY SCHOOL," *Ontol. J. PEMBELAJARAN DAN Ilm. Pendidik.*, vol. 2, no. 1, pp. 37–46, 2024.

[2] E. Roza, "Aksara Arab-Melayu di Nusantara dan Sumbangsihnya dalam Pengembangan Khazanah Intelektual," *Tsaqafah*, vol. 13, no. 1, pp. 177–204, 2017.

[3] G. Aceto, V. Persico, and A. Pescapé, "A survey on information and communication technologies for industry 4.0: State-of-the-art, taxonomies, perspectives, and challenges," *IEEE Commun. Surv. Tutor.*, vol. 21, no. 4, pp. 3467–3501, 2019.

[4] A. Drajat, E. W. Harahap, and others, "Rajah dan Spiritualitas Lokal dalam Hukum Islam; Studi Analisis Tafsir Hermeneutik," *Jurisprudensi J. Ilmu Syariah Perundang-Undangan Dan Ekon. Islam*, vol. 16, no. 1, pp. 225–240, 2024.

[5] L. Johanson, "The history of Turkic," in *The Turkic Languages*,

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2292, Copyright ©2025
Submitted : October 7, 2024, Revised : November 4, 2024, Accepted : November 25, 2024, Published : January 31, 2025

Routledge, 2021, pp. 83–120.
[6] C. Agus, S. R. Saktimulya, P. Dwiarso, B. Widodo, S. Rochmiyati, and M. Darmowiyono, "Revitalization of local traditional culture for sustainable development of national character building in Indonesia," *Innov. Tradit. Sustain. Dev.*, pp. 347–369, 2021.
[7] D. Iskandar, S. Hidayat, U. Jamaludin, and S. M. Leksono, "Javanese script digitalization and its utilization as learning media: an etnopedagogical approach," *Int. J. Math. Sci. Educ.*, vol. 1, no. 1, pp. 21–30, 2023.
[8] I. Siregar, "Papuan Tabla Language Preservation Strategy," *LingLit J. Sci. J. Linguist. Lit.*, vol. 3, no. 1, pp. 1–12, 2022.
[9] Y. N. Kunang, I. Z. Yadi, Mahmud, and M. Husin, "A New Deep Learning-Based Mobile Application for Komering Character Recognition," in *2022 5th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Yogyakarta, Indonesia: IEEE, Dec. 2022, pp. 294–299. doi: 10.1109/ISRITI56927.2022.10053072.
[10] T. P. Sari and Y. N. Kunang, "Pengembangan Aplikasi Transliterasi Teks Latin ke Aksara Ulu (Komering) Berbasis Web," *J. Process.*, vol. 18, no. 2, 2023.
[11] S. Huang, H. Wang, Y. Liu, X. Shi, and L. Jin, "OBC306: A large-scale oracle bone character recognition dataset," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, IEEE, 2019, pp. 681–688.
[12] R. Deng *et al.*, "Automatic Identification of Sea Rice Grains in Complex Field Environment Based on Deep Learning," *Agriculture*, vol. 14, no. 7, p. 1135, 2024.
[13] A. E. Ezugwu, A. K. Shukla, M. B. Agbaje, O. N. Oyelade, A. José-García, and J. O. Agushaka, "Automatic clustering algorithms: a systematic review and bibliometric analysis of relevant literature," *Neural Comput. Appl.*, vol. 33, pp. 6247–6306, 2021.
[14] I. Chatterjee, M. Ghosh, P. K. Singh, R. Sarkar, and M. Nasipuri, "A clustering-based feature selection framework for handwritten Indic script classification," *Expert Syst.*, vol. 36, no. 6, p. e12459, 2019.
[15] A. R. Widiarti, G. R. Prima, and C. K. Adi, "Preliminary research for provision of Javanese script image dataset from Javanese script printed book," in *AIP Conference Proceedings*, AIP Publishing, 2024.
[16] J. Oyelade *et al.*, "Data clustering: Algorithms and its applications," in *2019 19th international conference on computational science and its applications (ICCSA)*, IEEE, 2019, pp. 71–81.
[17] S. Setyaningtyas, B. I. Nugroho, and Z. Arif, "Tinjauan Pustaka Sistematis: Penerapan Data Mining Teknik Clustering Algoritma K-Means," *J. Teknoif Tek. Inform. Inst. Teknol. Padang*, vol. 10, no. 2, pp. 52–61, 2022.
[18] A. Ghosal, A. Nandy, A. K. Das, S. Goswami, and M. Panday, "A short review on different clustering techniques and their applications," *Emerg. Technol. Model. Graph. Proc. IEM Graph 2018*, pp. 69–83, 2020.
[19] S. Panda, M. Nayak, and A. K. Nayak, "Clustering of Odia character images using K-means algorithm and spectral clustering algorithm," in *ICICCT 2019–System Reliability, Quality Control, Safety, Maintenance and Management: Applications to Electrical, Electronics and Computer Science and Engineering*, Springer, 2020, pp. 55–64.
[20] A. R. Widiarti and C. K. Adi, "Clustering Balinese Script Image in Palm Leaf Using Hierarchical K-Means Algorithm," in *International Conference on Innovation in Science and Technology (ICIST 2020)*, Atlantis Press, 2021, pp. 38–42.
[21] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci.*, vol. 622, pp. 178–210, 2023.
[22] N. H. Shrifan, M. F. Akbar, and N. A. M. Isa, "An adaptive outlier removal aided k-means clustering algorithm," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 8, pp. 6365–6376, 2022.
[23] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data," *Inf. Sci.*, vol. 622, pp. 178–210, 2023.
[24] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," *Electronics*, vol. 9, no. 8, p. 1295, 2020.
[25] S. Sen, P. Chakraborty, S. Das, K. Pandey, and P. Narayana, "Investigation of Clustering Methods for SDSS Galaxy Images through Feature Extraction with VGG-16," in *2024 IEEE Space, Aerospace and Defence Conference (SPACE)*, IEEE, 2024, pp. 660–664.
[26] S. Tammina, "Transfer learning using vgg-16 with deep convolutional neural network for classifying images," *Int. J. Sci. Res. Publ. IJSRP*, vol. 9, no. 10, pp. 143–150, 2019.
[27] Y. Ren *et al.*, "Deep clustering: A comprehensive survey," *IEEE Trans. Neural Netw. Learn. Syst.*, 2024.
[28] I. Ioannou, C. Christophorou, P. Nagaradjane, and V. Vassiliou, "Performance Evaluation of Machine Learning Cluster Metrics for Mobile Network Augmentation," in *2024 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, IEEE, 2024, pp. 1–7.
[29] W. Bao, N. Lianju, and K. Yue, "Integration of unsupervised and supervised machine learning algorithms for credit risk assessment," *Expert Syst. Appl.*, vol. 128, pp. 301–315, 2019.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2292, Copyright ©2025
Submitted : October 7, 2024, Revised : November 4, 2024, Accepted : November 25, 2024, Published : January 31, 2025

Real-Time Vehicle Detection and Air Pollution Estimation Using YOLOv9

Hari Suparwito[1]*, Bernardus Hersa Galih Prakoso[2], Rosalia Arum Kumalasanti[3], Agnes Maria Polina[4]
Department of Informatics [1], [2], [3], [4]
Universitas Sanata Dharma
Sleman, Yogyakarta
shirsj@jesuits.net[1], Bernardus.hersa37@gmail.com[2], rosalia.santi@usd.ac.id[3], a.m.polina@usd.ac.id[4]

***Abstract*— Pollution of air, particularly in cities, is becoming an issue to be taken seriously owing to the health and environmental risks associated with it, and the major contributor to air pollution is car emissions. The objective of the study is to identify and classify vehicles such as motorbikes, cars, buses, trucks in order to monitor live traffic and potentially determine the extent to which the pollution level elevates, utilizing the YOLOv9 model. Traffic CCTV camera footage was gathered under a wide range of circumstances including different lighting and varying traffic intensity. Folders were particularly structured and images annotated, in the manner, which served the purpose of meeting the requirements of the YOLO structure. Once it was trained with a labeled dataset, the vehicle identification by YOLOv9 model was found to be quite satisfactory. Overall vehicle identification accuracy was calculated to be mAP50:95 of 0.826. In contrast, it had a harder time with smaller items like motorcycles, with a mAP50:95 of 0.682. Findings indicate that larger items were detected more than smaller items. Camera angles and the small size of the objects often make small objects appear to blend in to the background. This research indicates that AI can be of help when dealing with the urban structure. It offers a way of measuring traffic volume to predict the amount of CO emissions that can be avoided or controlled. The rest are keen in enhancing the effectiveness of recognizing small objects within the system and deploying it in multiple settings**.

Keywords*— *Air pollution, carbon monoxide, computer vision, object detection, yolov9

I. INTRODUCTION

Pollution is an issue of significant scientific concern due to its inherent negative health and environmental impacts [1]. In residential areas, pollution levels are closely linked to traffic patterns. Some pollutants are directly linked to vehicular traffic density [2]. More recent studies show an exponential increase in registered and traveled vehicles in residential areas [3]. In such a scenario of mixed development and environmental degradation, the need of the hour is to address these concerns using a positive approach. One such positive approach is to neutralize the traffic-related pollution to the best feasible.

Air pollution is regarded as one of the main environmental threats. The negative impacts of pollution are multifaceted, encompassing ecosystems, health, agricultural productivity, educational attainment, and worker productivity [4]. In many situations, transport significantly contributes to air pollution through vehicle emissions. Damage to human health is one of the major outcomes of vehicle pollution, as road traffic is deemed a principal source of fine particulate matter, which results from the incomplete combustion of fuels in vehicle engines [5]. The major contributions to air pollution levels are associated with exhaust damage and the generation of wear particulates in vehicle engines. Damage to the environment beyond vehicle emissions is assessed in related work and correlated with transportation infrastructure. In response, technology and the environment have been evaluated by some researchers in order to help minimize the amounts of pollution emitted into the atmosphere [6]. Various methodologies are used to monitor different pollutants and toxicants, as well as to extend the spectrum of monitoring purposes. However, existing literature has more deeply given attention to the technology of vehicle detection, paying less attention to the contribution of front vehicles to pollution build-up.

While there are various parameters taken into consideration under pollution, significant advances were made in monitoring vehicular emissions and accidents. Manufacturing has also been growing in different countries around the world, contributing to increasing air and noise pollution due to gaseous emissions from these industries, as well as increasing road traffic. Motorized road traffic has led to various problems such as fuel exhaust emissions leading to air pollution affecting both outer and in-cabin air quality, a noisy atmosphere with increased noise pollution, and traffic congestion leading to traffic accidents. Human activities, including the excessive use of motor vehicles, play a major role in environmental degradation and cause pollution. To control these factors, it is important to monitor the environment, and it is understood that the development of detection systems with an environment-friendly approach is an urgent need.

From the existing condition, there are possible factors contributing to the accumulation of air pollution, such as inadequate or unfound technology for detecting vehicles. Thus, uncovering where vehicles are crucially confined and exploring real-time locations give the opportunity for another study to be conducted. The rationale ultimately allows for the possibility to assess effective traffic control methodology. By reviewing the studies conducted, it was seen that no research has focused on developing an effective real-time approach to facilitate the process, as has been approached in this work. Therefore, the research question of monitoring using AI for vehicles is able to be tackled by designing a system that has the capability to monitor traffic.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2329, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Most of the primary purpose of air pollution research is to determine how much CO, CO_2, SO_2, and NO_2 can be directly measured in the open air. These pollutants degrade the quality of the air and are detrimental to human health [7, 8, 9, 10]. However, in this study, we used computer vision technology to count vehicles in four distinct categories: motorcycles, cars, buses, and trucks. We use the results of vehicle identification and enumeration to predict the presence of hazardous substances that contribute to air pollution. This instance focuses on a specific component of the atmosphere, carbon monoxide (CO).

Recently, deep learning has transformed various fields among the usage of AI, particularly in object detection [11]. A specific detection method performs detection instantly and accurately in one pass. This method can be used to detect objects. With its real-time detection rates, it was seen to perform better than two-stage detectors and single-shot detectors. The association of AI as a monitoring tool for environmental assessment should be able to augment existing plausible capabilities in the environmental domain; hence, this work aims to investigate the application of AI in the real-time detection of vehicles in a traffic lane using a specific version of the detection method. This investigation is also supported by a discussion about the contribution of pollution build-up.

The development of information technology and artificial intelligence has expanded since the 21st century as a result of extensive research and the increasing need for automation. Almost every industry today uses automation built with the help of computer vision and artificial intelligence. Computer vision has become one of the rapidly developing fields of knowledge and is reliably utilized in various industries. Computer vision has the potential to integrate human interaction with systems in a very modern way, ensuring that future technology will always be up-to-date. The advantage of computer vision lies in the extraction of information from images, videos, and other visual inputs, which can then be further processed according to needs [12]. Computer vision is a field of artificial intelligence that can recognize objects in its surroundings, and this will be utilized in the research.

Object detection is a subfield of computer vision that involves the identification and localization of an object in images or videos. Identifying objects requires domain-specific algorithms and the use of standard techniques that include algorithms like image segmentation, classification, or deep learning-based approaches, resulting in the emergence of a number of applications. Classification algorithms identify a class of an image that can be embedded with bounding boxes. Image segmentation enters smaller areas which, in a similar manner, classify images and then take the union of detected objects at different locations. At the same time, deep learning methods are focusing on classifying subparts of that image. On the contrary, with the advancement of technology, deep learning-based object detection algorithms have taken into consideration resources from a variety of fields that include computers, medical, structural, surveillance, home automation, and transportation applications. In addition, vehicle detection helps in traffic monitoring and management, automated parking systems, vehicle counting at tolls, and surveillance of attackers and terrorists, among others. In transportation, the detection framework must be accurate and timely as it impacts a vehicle's time, speed, and distance with respect to other vehicles.

The capabilities of computer vision are widely utilized in biometric identification, such as fingerprint detection, facial recognition, gestures, and other biometric objects. The increasing demand for technology has led to the involvement of computer vision in the development of the autonomous industry, namely the 4.0 industrial revolution. The 4.0 industrial revolution era emphasizes automation that collaborates with technology, making computer vision play a significant role in this era [13].

Real-time object detection is a tough and evolving area in computer vision and machine learning. In recent years, YOLO (You Only Look Once) has emerged as a leading model in this field. YOLO takes a different approach than traditional methods to train object detection models [14]. YOLO trains a single neural network to specify bounding boxes and class probabilities at the same time. Also, it doesn't rely on pre-defined object features like edge maps, HoG descriptors, or region proposals, which sets it apart from classic object detection or segmentation algorithms. The YOLO system can process images at about 45 frames per second while achieving accuracy that matches or beats most current state models for object detection [15]. One type of YOLO could look at 2 pictures every second and got a score of 57.9 mAP with a COCO model. Another type when changed a bit, could look at 61 pictures a second and did just as well as the ones before it. The newest type is much faster and more accurate. It can look at 65 pictures every second and is good at spotting things.

Because it's so good, YOLO is used a lot to watch traffic and in smart cities. But it might have trouble when there's a lot going on in the background or when there are small things that get in the way [16]. This capability allows us to control traffic flow and plan cities better. Smart video systems give steady info on traffic jams and car numbers. This data helps governments decide when to widen roads or fix them after breakdowns. People also use YOLO to watch traffic testing it to spot different kinds of vehicles [17]. The model can show how many cars there are how busy the roads are, what types of vehicles are around, and how fast they're going. This means YOLO could be used to manage air pollution by keeping an eye on various cars as they move. YOLO gives useful details by showing the user's surroundings to the program and working through the video feed. With this info, users can check traffic levels right away, guess car speeds, and follow air pollution amounts [18].

Computer vision has many uses besides traffic control and keeping self-driving cars safe, which are making big strides [19]. Cameras or CCTV watch traffic to see what's happening and catch rule-breakers. This watching also helps count how many cars pass by in a set time. Knowing the number of vehicles lets us figure out how much air pollution builds up at spots where cameras or CCTV are keeping an eye on things [20]. This study aims to create computer vision technology using YOLOv9 to spot air pollution by counting vehicles in a set time frame. We chose YOLOv9 because it was the stable version of the YOLO model at the time of this study. There is already YOLOv10, but we think that we will not use the latest

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2329, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

version, assuming that it may not be stable. The research splits motor vehicles into four groups: motorcycles, cars, buses, and trucks. The Transportation Department's CCTV captures traffic images or videos. The collected dataset then goes through processing with the chosen model. YOLOv9 handles the model testing in this research.

To detect and identify objects, the study looks at each model's confidence scores. The processed video will show confidence values for every type of motor vehicle. The types of motor vehicles that have been successfully recognized will also make it easy to count the number of vehicles passing through that point in a given time period. Based on the number of vehicles recognized by the system, the impact of air pollution at that point can also be calculated. This research uses an object detection approach to calculate exhaust gases or pollution. It is hoped that this approach can provide a solution for current traffic management.

This paper is an attempt to implement technology for vehicle detection and traffic analysis of a specific area. The primary objective of the implementation of the tool is to monitor the pollution levels in the study area and optimize the vehicular traffic sources of pollutants distributed in the respective area. Such a method would also generate data for traffic where exact vehicle counts can be detected and addressed, thus proposing a traffic management policy. With the development of the research findings and the consequent method, the environmental policy of the residential area could have been explored and made public in due course. Environmental studies and technology are the driving force for such innovation. The implementation of the tool thus developed in control systems can minimize the negative impact due to vehicular traffic

II. MATERIALS AND METHODS

A. Data Collection

In the related works, publicly available traffic camera and radar datasets or synthetic training data derived from real-world data are used for vehicle detection and traffic analysis. In this study, we adopted real-world vehicle and traffic data collection instead of using public data for the vehicle and traffic scenario in which we aim to analyze the impact of pollution on traffic. Real-world vehicle data are particularly essential for the testing and evaluation of our proposed model, as the testbed used in the data collection is representative of the research context of this paper. Thus, the vehicle data input of the carbon monoxide dataset are real-world data, instead of synthetic or public data that do not represent the underlying context of this research.

Figure 1. Daytime Image Example During Peak Hours

Figure 2. Daytime Image Example During The Off-Peak Hour Seasons

Moreover, collecting varied pollutant datasets captures the driving conditions under different situation and vehicle types. The video data is collected at one sample per minute when the traffic light is on the red and vehicles are stop. This study also presents traffic scenarios such as heavy and light traffic, night and day conditions, different speeds, and vehicle types, as above. in each video shot, done with a duration of about 10 minutes.

Figure 3. Night Time Image Example During Off-Peak Hours

The steps of the study following the diagram below,

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2329, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

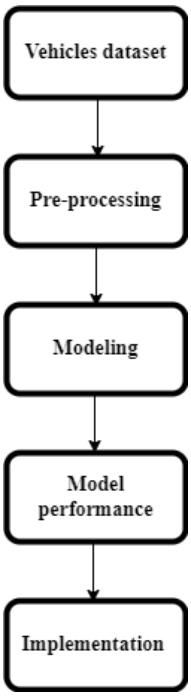


Figure 4. The Flow Diagram Process

in general, there are 5 stages starting from data collection and ending with real-world implementation.

Vehicles Dataset

This step involves collecting and preparing a dataset containing images or video frames of vehicles, such as cars, buses, motorcycles, and trucks. The dataset should include labelled data, where each vehicle is annotated with bounding boxes and corresponding class labels. High-quality and diverse data, encompassing different angles, lighting conditions, and vehicle types, is essential for training an effective model.

Pre-processing

Pre-processing ensures the data is prepared for training. This includes resizing images to a uniform dimension suitable for YOLO's input, namely 640x640. Annotation files are also converted into YOLO-compatible formats. These steps enhance the model's ability to generalize to unseen data. We did not conduct a noise and anomaly image pre-processing since we thought that YOLOv9 is good in handle this kind of images [21]. Similarly, we do not perform data augmentation because the data source is video images and runs continuously 24/7. Therefore, the amount of data obtained is more than enough, thus process augmentation is not necessary.

Modelling

YOLOv9, a state-of-the-art object detection algorithm, is used for vehicle detection and counting. In this project, the model is trained with the following parameters: Epochs, defines the number of complete passes through the dataset during training, is 50. Batch Size is equal to 32. Dropout, the parameter that prevents overfitting by randomly disabling neurons during training is 0.4. Other hyperparameters are kept at their default values. YOLOv9 processes the input images to detect and classify vehicles in real time with high accuracy.

Model Performance

After training, the model is evaluated on validation and test datasets using metrics such as precision, recall, F1-score, and mean Average Precision (mAP). The confusion matrix helps identify errors in classification. This step ensures the model performs well and identifies areas for improvement.

Implementation

The trained YOLOv9 model is deployed in a real-world scenario to count vehicles. This involves integrating the model with a live video feed or camera system, where it detects and counts vehicles frame by frame. Additional logic, such as tracking vehicle movements across frames, can be incorporated to avoid double counting and provide accurate results.

In the implementation step, we did objects counting. YOLOv9 offers a robust solution for real-time vehicle counting due to its high speed and accuracy. In this study, object counting is used to count the number of vehicles such as cars, buses, motorcycles, and trucks at a given time. From the results of the calculation of the number of vehicles, the amount of CO gas content in the air in the area will be predicted.

The most important thing in this study is to predict the amount of pollution, especially Carbon monoxide (CO) emissions, based on the number of vehicles at a certain time in a designated area. Each vehicle is measured and estimated its CO emission. We calculated the emission based on the following assumption, vehicles are in the idle machine condition for one minute, the machine capacities are 125 cc for motorcycles, 1500 cc for cars, 6000 cc for buses, and 11000 cc for trucks. According to Indonesian government, the CO emission threshold is 25 ppm or equal to 28640 ug/m^3. The following table shows the factor emission CO for each vehicle in Indonesia [22].

Table I. Vehicle CO Emission Factors

Vehicles	Average CO (gr/m)
Motorcycles	0.014
Cars	0.040
Buses	0.011
Trucks	0.008

Emission analysis was conducted to determine the number of emissions produced by motorized vehicles. Motor vehicle emissions on the road are caused by three factors namely the total volume of motorized vehicles, characteristics of motor vehicles and general traffic conditions [23]. The equation used in calculating the amount of motor vehicle emissions are [24]:

$$E = \sum_{i=1}^{n} L \; x \; Ni \; x \; Fi \quad (1)$$

where:

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2329, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

L = Length of road segment (km), N = Number of motorized vehicles of type i (vehicles/hour), F = CO emission factor of motorized vehicles of type i (gr/km), i = Type of motorized vehicle based on fuel type, and E = Total CO emissions per road segment (gr/hour).

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2329, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

III. RESULTS AND DISCUSSIONS

The performance of YOLOv9 for detecting vehicles in an urban area is summarized. We presented the results here.

Figure 5. Vehicles Detected During Day Time In Peak Hours

Table II. Number Of Vehicles In Figure 5

Class	Actual	Predicted
Motorcycles	45	39
Cars	39	35
Buses	1	2
Trucks	2	2

Figure 6. Vehicles Detected During Night Time In Peak Hours

Table III. Number Of Vehicles In Figure 6

Class	Actual	Predicted
Motorcycles	43	39
Cars	17	14
Buses	0	0
Trucks	0	0

Figure 7. Vehicles Detected During Night Time In Off Peak Hours

Table IV. Number Of Vehicles In Figure 7

Class	Actual	Predicted
Motorcycles	6	6
Cars	4	4
Buses	0	0
Trucks	0	0

Figures 5, 6, and 7 show the results of the YOLOv9 algorithm detection of several traffic conditions. While tables 2, 3, and 4 show the calculation of vehicles manually as well as from the calculation results of the YOLOv9 algorithm. from the figures and tables generated by the YOLOv9 algorithm, it shows that the YOLOv9 algorithm can be used and its accuracy is good enough to detect vehicles in all conditions. For example, in night conditions and crowded vehicles, YOLOv9 can predict the number of vehicles more than 80% correctly (table 3). Even in night situations with a small number of vehicles, YOLOv9 can guess 100% correctly (table4).

YOLOv9 models provided the calculation of the Confusion Matrix that is used as

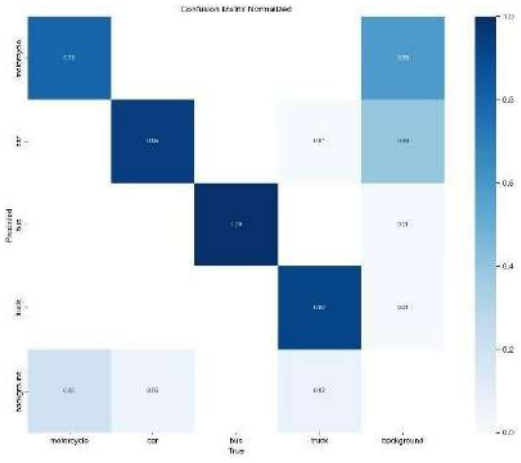


Figure 8. Confusion Matrix

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2329, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Also shown in figure 9 is the F1 confidence measurement as follows

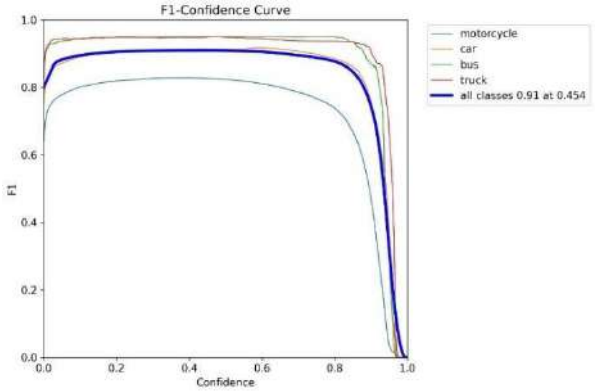


Figure 9. Confidence Curve

The confusion matrix shown in fig. 8 presents the normalized results of the classification model, where the model categorizes the image into one of the following classes: motorcycle, car, bus, truck, and background. The model accurately predicted motorcycles 94% of the time, with 6% being misclassified as cars. The model's performance for cars was quite good, with 98% accuracy, and only 2% misclassified as trucks.

Buses posed a bigger problem for the model achieving 88% accuracy. The model mistook 9% of buses for trucks and 3% for cars. Trucks reached 95% accuracy, but the model sometimes confused them with buses (3%) or cars (2%). The background class had the lowest accuracy at 80%, with 20% misclassified as cars. The diagonal lines running from top left to bottom right, show each class's accuracy pointing out correct predictions. On the other hand, cells outside the diagonal reveal the type and frequency of mistakes. The confusion matrix proves that the model excels at spotting cars, with high accuracy. It also does a good job identifying motorcycles and trucks. However, it tends to mix up buses and trucks, and it has trouble classifying the background class. This hints that the model might need more work to tell these classes apart better.

The F1-Confidence graphs (fig.9) depict the correlation between model's prediction confidence and the F1 scores averaged over each class as well as unique classes. The x-axis represents the confidence level of the model with values between 0 and 1. The y-axis considered the precision and recall characteristics of the model, and it is represented as F1 score. Each colored line corresponds to a particular class: a motorcycle, a car, a bus, and a truck. In the first, the thrilling blue line represents the combined F1 score of all classes. From the observations taken from the F1-Confidence Curve (fig. 9), it can be stipulated that for all the classes under review, the F1 score tends to increase with the increase in the confidence level. This means that as the confidence level increases, so does the efficiency of the model in making predictions. The model records high F1 scores for cars across all the confidences. This correlates with the observation made with the confusion matrix where car's accuracy is registered the highest. The F1 value obtained for motorcycles was satisfactory, though high throughout. Buses and trucks showed some fluctuations in F1 values, especially at lower confidence levels. This shows that the model may not be more effective when distinguishing these two classes when the confidence is less than high. The overall F1 score (blue line) tends to reach its maximum of about 0.91 when the confidence level is between 0.4 and 0.5. Thus, it can be stated confidently that the model is accurate as long as its confidence level exceeds 0.45.

Referring to all the image results, it can be concluded that YOLOv9 has limitations for small object detection. When compared between motorcycles and cars, it can be seen that motorcycle detection provides poor results compared to car detection. Motorcycle images that are too close and small objects become the same object as the background. in contrast to cars that are large enough, YOLOv9 can easily recognize and also the color of the car is more varied into clearly distinguishable objects. In order for YOLOv9 to be better at detecting small objects, there are several things that can be done including: increase image resolution, increase the number of grid cells and label small object accurately. In this study, we used 640x640 pixels of image resolution. We can use a higher resolution, such as 720x720 or even more, to improve YOLOv9's ability to detect small objects. As for grid cells, this study used 26x26 grid cells, but 32x32 is better. Another thing is to label the image more accurately by distinguishing each object, especially small and overlapping ones. Counting objects would be better if the objects were separate and not clustered in groups and stopped like those at a traffic light. Another thing that is also a supporting factor in object detection in this study is the CCTV layout which is quite high causing the angle of capture and object recognition to be a challenge. most small objects (motorbikes) will be recognized as background especially if the color of the motorcycle is similar to the background color.

Comparing the Convolution Matrix result, e.g., for a car the Convolution Matrix result is 0.86 with the comparison between the real count and the detection count by YOLOv9, 49:36 shows that the YOLOv9 model is close to correct and able to guess correctly. This study does not conclude that the area is polluted, but the results of this study can be used to predict the CO content in the air based on the count of vehicles detected by the YOLOv9 model. This approach has the potential to use object detection technology to regulate traffic and reduce potential air pollution in certain areas.

Next, we analyzed the image. The following table, table 6, shows the measurements of the mAP for the images recorded by CCTV.

Table V. Measurements of Map Based on Class

Class	Images	Instances	mAP50	mAP50:95
All	1473	28137	0.915	0.826
Motorcycles	1428	17012	0.836	0.682
Cars	1256	9937	0.919	0.845
Buses	232	400	0.944	0.867
Trucks	533	788	0.961	0.911

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2329, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Table 6 summarizes object detection performance across various classes using two key metrics: mAP50, which measures Mean Average Precision at a 50% Intersection over Union (IoU) threshold, and mAP50-95, which averages precision across IoU thresholds ranging from 50% to 95%. Focusing on mAP50-95, the overall detection performance across all classes is strong, with a score of 0.826. However, there are notable differences among the individual classes.

Motorcycles have the lowest mAP50-95 at 0.682, suggesting that detecting motorcycles is more challenging compared to other classes. This could be due to factors like their variability in appearance or smaller size. Cars perform significantly better, with an mAP50-95 of 0.845, indicating reliable detection accuracy for this class. Buses show slightly improved performance over cars, achieving an mAP50-95 of 0.867. Trucks, on the other hand, exhibit the highest detection accuracy with an impressive mAP50-95 of 0.911, making them the easiest class to detect accurately. In summary, the model performs well overall but shows a clear hierarchy in class performance. Trucks are detected with the highest accuracy, followed by buses and cars, while motorcycles present the greatest challenge, highlighting an area for potential improvement in detection algorithms.

Finally, we would calculate the CO emission content based on the number of motorized vehicles. Based on table 2, 3 and 4, we predicted the CO emission levels in the area on the specific time.

Table VI. total amount of CO emission level

Table	Values of CO (gr/m)
2	(39x0.014) + (35x0.04) + (2x0.011) + (2x0.008) =1.98
3	(39x0.014) + (14x0.04) =1.106
4	(6x0.014) + (4x0.04) = 0.24

Referring the assumption used in this study, we can conclude that during peak hours the CO emission content is 1.98 gr/m ≈ 198 gr/100m while at night or during non-peak hours the CO emission content decreases to around 1.06 gr/m ≈ 106 gr/100m (day time) and 0.24 gr/m ≈ 24 gr/100 m (night time).

Instead of focusing on chemicals in the air, this study advances traffic-related environmental studies by measuring CO emissions using computer vision technology. This work uses computer vision technology to propose a novel technique to estimate CO emissions based on vehicle recognition and counting. This technology enables real-time data collecting and analysis, resulting in a more dynamic and responsive understanding of transportation emissions. Unlike chemical content analysis, which provides a broad assessment of air quality, computer vision enables the accurate calculation of emissions based on the number and type of vehicles. This specificity may lead to more precise estimates of traffic-related emissions. Computer vision offers real-time data collecting, allowing for instant analysis and insights into traffic patterns and their environmental impact. This is very useful for conducting timely traffic control initiatives. We can potentially construct predictive models for transportation emissions by combining computer vision and emission computations, which can help with proactive policy creation. To summarize, this study improves the field by providing a more precise, rapid, and scalable method for measuring CO emissions from transportation. This provides us with vital information for controlling traffic and developing environmental policies.

The finding of the study also reveals that peak hours of the day have significantly higher CO emissions than non-peak hours. This could impact traffic management and environmental policies. Traffic management measures like adjusting traffic signal timings, adopting congestion pricing, and encouraging carpooling during peak hours can help reduce congestion and emissions. Encouraging public transit during peak hours can also reduce vehicle usage, resulting in lower CO emissions. Tougher emission restrictions for vehicles, particularly during peak hours, could be implemented. Real-time air quality monitoring systems can alert individuals to minimize outdoor activities or wear masks during high-emission periods. Long-term urban planning can create cities with less traffic congestion, such as mixed-use districts. Public education on peak hour emissions can lead to behavioral adjustments, such as avoiding unnecessary journeys or choosing cleaner transportation. These findings can guide targeted initiatives for better traffic management and reduced environmental impact, ultimately improving air quality and public health.

IV. CONCLUSION

The study explores the application of the YOLOv9 object detection model to monitor traffic and estimate air pollution in urban environments. Recognizing the significant environmental and health impacts of vehicular emissions, the research focuses on using computer vision technology to detect and classify vehicles in real-time, providing data on traffic density and pollution levels. The system was trained on a diverse dataset comprising various vehicle types under different traffic conditions and lighting scenarios. Results show the model effectively identifies and counts vehicles, with particularly high accuracy for larger vehicles like trucks and buses, 98% accuracy for car detection, 95% accuracy for truck detection and 88% accuracy for bus detection, but limited performance for smaller objects like motorcycles, which are often confused with the background as well as the camera distance and viewing angle of the CCTV camera.

The study on YOLOv9 for detecting vehicles and calculating pollution levels suggests future improvements. It suggests experimenting with advanced versions of YOLOv9, optimizing the model for unique locations or conditions, and integrating data from air quality sensors for improved emission estimates. Testing the system in various metropolitan environments and including additional pollutants like NOx, PM2.5, and PM10 can provide a more comprehensive picture of traffic-related environmental consequences. Collaborating with environmental scientists can also improve emission models and validate calculations using recent scientific findings. These recommendations aim to make vehicle detection and air pollution assessment technology more accurate and helpful. Specifically in this study, by estimating

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2329, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

carbon monoxide (CO) emissions based on vehicle counts, the study links object detection with air pollution management, providing a framework for real-time environmental monitoring. While the model achieves strong overall performance, challenges in detecting small objects and limitations due to camera angles highlight areas for improvement. This work underscores the potential of integrating AI with environmental policy to optimize traffic flow, reduce pollution, and enable more sustainable urban planning.

The findings pave the way for further research into improving detection accuracy and expanding applications to broader environmental monitoring scenarios.

ACKNOWLEDGMENT

We would like to thank the Intelligent System laboratory of the Informatics Department of Sanata Dharma University for allowing us to use the laboratory to do this research. All financial support for this research activity is funded by the Research and Service Institute (LPPM) of Sanata Dharma University.

REFERENCES

[1] D., Yadav, et al. "Environmental and health impacts of contaminants of emerging concerns: Recent treatment challenges and approaches." Chemosphere 272, 2021, p. 129492.

[2] J. Lu, B. Li, H. Li, A. Al-Barakani, "Expansion of city scale, traffic modes, traffic congestion, and air pollution." Cities 108, 2021, p. 102974.

[3] E. Barmpounakis, M. Montesinos-Ferrer, EJ. Gonzales, N. Geroliminis, "Empirical investigation of the emission-macroscopic fundamental diagram." Transportation Research Part D: Transport and Environment 101, 2021, p. 103090.

[4] W. Zhou, L. Mengmeng, V. Achal. "A comprehensive review on environmental and human health impacts of chemical pesticide usage." Emerging Contaminants, 2024, p. 100410.

[5] S. Aminzadegan, M. Shahriari, F. Mehranfar, B. Abramovic, "Factors affecting the emission of pollutants in different types of transportation: A literature review." Energy Reports 8, 2022, pp. 2508-2529.

[6] RM. Harrison, et al. "Non-exhaust vehicle emissions of particulate matter and VOC from road traffic: A review." Atmospheric Environment 262, 2021: 118592G.

[7] Singh, Dharmendra, Meenakshi Dahiya, Rahul Kumar, and Chintan Nanda. "Sensors and systems for air quality assessment monitoring and management: A review." *Journal of environmental management* 289, 2021: 112510.

[8] Jion, Most Mastura Munia Farjana, Jannatun Nahar Jannat, Md Yousuf Mia, Md Arfan Ali, Md Saiful Islam, Sobhy M. Ibrahim, Subodh Chandra Pal et al. "A critical review and prospect of NO2 and SO2 pollution over Asia: Hotspots, trends, and sources." *Science of The Total Environment* 876, 2023: 162851.

[9] Singh, Pratichi, and Deepak Yadav. "Link between air pollution and global climate change." In *Global climate change*, 2021, pp. 79-108.

[10] Hannun, Rafid M., and Ali H. Abdul Razzaq. "Air pollution resulted from coal, oil and gas firing in thermal power plants and treatment: a review." In *IOP Conference Series: Earth and Environmental Science*, vol. 1002, no. 1, 2022, p. 012008.

[11] AB. Amjoud, M. Amrouch. "Object detection using deep learning, CNNs and vision transformers: A review." *IEEE Access* 11, 2023, pp. 35479-35516.

[12] V. Wiley, T. Lucas, "Computer Vision and Image Processing: A Paper Review," *International Journal of Artificial Intelligence Research*, vol. 2, no. 1, 2018, pp. 28-36.

[13] Y. Fengwei, S. Gu. "Industry 4.0, a revolution that requires technology and national strategies." *Complex & Intelligent Systems* 7, 2021, pp. 1311-1325.

[14] S. Liang, et al., "Edge YOLO: Real-time intelligent object detection system based on edge-cloud cooperation in autonomous vehicles". IEEE Transactions on Intelligent Transportation Systems, 23(12), 2022, pp.25345-25360.

[15] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using YOLO: Challenges, architectural successors, datasets and applications," Multimedia Tools and Applications, vol. 82, no. 6, 2023, pp. 9243–9275.

[16] Y. Zhang et al., "Real-time vehicle detection based on improved YOLO v5," Sustainability, vol. 14, no. 19, 2022, p. 12274.

[17] C.-J. Lin, S.-Y. Jeng, and H.-W. Liao, "A real-time vehicle counting, speed estimation, and classification system based on virtual detection zone and YOLO," Mathematical Problems in Engineering, vol. 2021, no. 1, 2021, p. 1577614.

[18] MT. Mahmood, SR. Ahmed, MR. Ahmed. "Detection of vehicle with Infrared images in Road Traffic using YOLO computational mechanism." IOP Conference Series: Materials Science and Engineering. Vol. 928. No. 2. IOP Publishing, 2020.

[19] M. Maity, S. Banerjee, SS. Chaudhuri. "Faster r-cnn and yolo based vehicle detection: A survey." 2021 5th international conference on computing methodologies and communication (ICCMC). IEEE, 2021.

[20] U. Mittal, C. Priyanka, R. Tiwari. "EnsembleNet: A hybrid approach for vehicle detection and estimation of traffic density based on faster R-CNN and YOLO models." Neural Computing and Applications 35.6, 2023 pp. 4755-4774.

[21] Bakır, Halit, and Rezan Bakır. "Evaluating the robustness of yolo object detection algorithm in terms of detecting objects in noisy environment." *Journal of Scientific Reports-A* 054, 2023, pp. 1-25.

[22] R. P. Bachtera, H. S. Huboyo, and B. P. Samadikun, "Uji coba estimasi emisi kendaraan bermotor yang beroperasi di Kota Semarang berdasarkan umur dan jenis kendaraan dengan menggunakan perangkat lunak LEAP," Ph.D. dissertation, Diponegoro Univ., Semarang, Indonesia, 2017

[23] S. A. M. P. Ofrial, E. Ahyudanari, A. D. Syafei, "Estimation on the increasing value of CO based on the vehicle growth in Surabaya," Procedia - Social and Behavioral Sciences, vol. 227, 2016, pp. 410–416.

[24] Y. Sekaryadi and S. Wimpy, "Emisi kendaraan pada ruas jalan provinsi di Jawa Barat," Jurnal Himpunan Pengembangan Jalan Indonesia (HPJI), vol. 3, no. 1, 2017, pp. 29–36.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2329, Copyright ©2025
Submitted : December 14, 2024, Revised : January 15, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Enhancing Smart City Maturity Through Digital Transformation: A Success Factors Analysis

Widya Listianingsih[1], *Tony Dwi Susanto*[2]*
Department of Information System[1], [2]
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
6026221011@student.its.ac.id[1], tonydwisusanto@its.ac.id [2]

***Abstract*— Improving smart city maturity through digital transformation is becoming increasingly crucial in facing the challenges of rapid urbanization and the need for more efficient city governance. However, the lack of a unified understanding of key drivers and challenges in this domain has limited the effectiveness of existing strategies. This study aims to explore the role of digital transformation in improving Smart City maturity by identifying key success factors and best practices adopted by cities worldwide. This study used the systematic literature review (SLR) methodology based on the PRISMA framework, which included systematic steps in selecting, collecting, and analyzing relevant literature. The study results reveal six factors influencing Smart City maturity: Information and Communication Technology (ICT) infrastructure, data integration, government policies and strategic planning, stakeholder engagement, environmental sustainability, and innovation and human resource development. Unlike previous studies, this study synthesizes global best practices and success factors, offering actionable insights for policymakers and practitioners to design inclusive, sustainable, and forward-looking digital transformation strategies. Furthermore, the study underscores the need for context-specific research to optimize implementation and drive meaningful progress in diverse urban settings.**

Keywords—Digital Transformation, Smart City Maturity, Success Factor, Best Practice, PRISMA

I. INTRODUCTION

Digital transformation has become one of the key elements in addressing the challenges arising from rapid urbanization and the increasing complexity of urban governance in the modern era [1]. By leveraging digital technologies, especially Information and Communication Technology (ICT) infrastructure, cities worldwide can create innovative solutions to improve the efficiency of public services, environmental sustainability, and the community's quality of life [2]. Digital transformation supports daily city operations and paves the way for developing a mature, smart city. The smart city concept integrates advanced technologies, innovation management, and strategic policies to create an efficient, resilient, and sustainable urban system [3].

Rapid urbanization has put enormous pressure on city infrastructure, such as transportation, energy, and waste management, requiring a smarter and more integrated approach [4]. Smart cities are a promising solution that leverages digital technology and data to create cities more responsive to community needs. Digital transformation enables cities to manage resources more efficiently, reduce carbon emissions, increase community engagement, and optimize public services [3], [5]. However, the success of digital transformation depends not only on the adoption of technology but also on social and cultural readiness and collaboration between stakeholders such as the government, the private sector, and civil society.

While studies have explored individual aspects of digital transformation and smart cities, a more holistic approach is lacking to understand the relationship between the two concepts. Many studies focus on technological elements such as ICT infrastructure or data-driven applications but rarely link them to smart city maturity holistically. For instance, robust ICT infrastructure is a cornerstone for smart city development, enabling public data management, big data networks, IoT device deployment, AI applications, and cybersecurity [6], [7]. Complementing this, effective data integration is crucial for ensuring interoperability between public services, enhancing data-driven decision-making, and improving service quality [8], [9]. Other studies show that government policies and strategic planning further amplify these efforts, demonstrating the transformative impact of digitalization-focused initiatives in sectors like education and healthcare [10], [11]. In addition, stakeholder engagement is equally vital, as evidenced by initiatives such as online community consultations and public-private collaborations, which underline the importance of inclusivity and cross-sector cooperation [6], [8]. Moreover, environmental sustainability, supported by IoT and sensor technologies for resource management, highlights the need for long-term ecological solutions [7], [12].

Despite these advancements, there is an urgent need to identify broader success factors, including policy, governance, social innovation, and environmental sustainability. With a more integrated approach, the challenges hindering smart city development are hoped to be effectively addressed. Based on these considerations, this study seeks to answer the following research question: What are the key success factors of digital transformation in accelerating the maturity of smart cities?

This study aims to fill this gap by conducting a Systematic Literature Review (SLR) using the PRISMA framework [13].

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2326, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 21, 2025, Published : January 31, 2025

The PRISMA framework was selected for its rigor and transparency in synthesizing existing research. Its systematic approach ensures a comprehensive and unbiased selection, collection, and analysis of relevant literature, enabling the identification of patterns and gaps in the field. By leveraging this methodology, the study ensures the findings are grounded in a robust and reproducible evidence base. This study will identify various success factors and best practices that cities can adopt to accelerate their digital transformation process. By providing a more comprehensive understanding, the results of this study are expected to be the basis for developing more effective strategies for supporting innovative, inclusive, and sustainable smart cities.

II. RESEARCH METHOD

This study uses a Systematic Literature Review (SLR) approach to analyze relevant literature related to the contribution of digital transformation to smart city maturity. SLR was chosen because this method allows researchers to systematically identify, evaluate, and synthesize existing findings in the literature, thus providing a strong basis for drawing valid and relevant conclusions. The SLR process in this study follows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework [13], which provides structured guidance for each study stage, as seen in Fig 1.

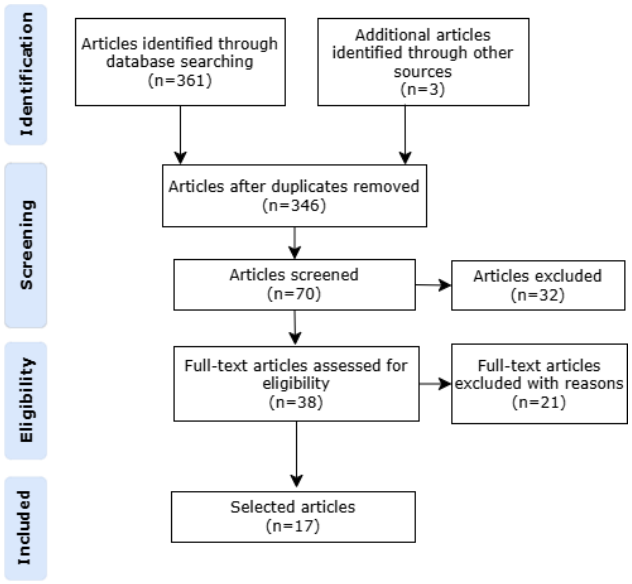


Figure 1. Prisma Work Flow

A. Literature Identification

The first stage of the review process involves identifying relevant literature on the contribution of digital transformation to smart city maturity. A systematic literature search was conducted across prominent academic databases, including Google Scholar, Scopus, IEEE Xplore, and ScienceDirect, to ensure comprehensive coverage of high-quality articles. The search was carried out using targeted keywords such as "digital transformation," "smart city," "maturity model," and "success factors," combined with Boolean operators (AND, OR) to refine the search. For example, one of the queries used was: ("Digital Transformation") AND ("Smart City") AND ("Maturity Model" OR "Success Factors"). This search was limited to articles published in the last 10 years (2014–2024) and written in English to ensure the relevance and accuracy of the findings. Only peer-reviewed articles were included to guarantee the credibility of the data.

B. Literature Screening

After identifying the literature, the next step is screening, ensuring that only relevant articles are considered for analysis. This process follows predefined inclusion and exclusion criteria to focus the review on quality studies. The literature screening includes two steps: first, reviewing titles, abstracts, and keywords to assess relevance; second, analyzing full-text articles to confirm their alignment with the study's focus. The inclusion criteria specify that the articles must be published in English, be peer-reviewed, and directly relate to digital transformation or smart city maturity models. Exclusion criteria include studies with a narrow geographic scope or articles that do not provide empirical data. A PRISMA flowchart was used to document this selection process, summarizing the number of articles screened, excluded, and ultimately included. Table 1 summarizes the inclusion and exclusion criteria used in literature screening:

TABLE I. CRITERIA SELECTION

Inclusion	Exclusion
Publication during the period of 2014 to 2024	Publications not written in English
Titles, abstracts, or keywords include search terms	Publications not subjected to the peer-review process
Publications are written in english	
The publication goes through the peer-review process.	

C. Eligibility Evaluation

Following the screening process, the next step is the eligibility evaluation, where the selected articles are reviewed in more detail to assess their relevance to the study's objectives. This evaluation aims to ensure that the articles provide valuable insights into the factors contributing to the success of digital transformation in the context of smart city maturity. Each article is evaluated based on the following criteria:

a) Key Success Factors: *The article should address specific factors that influence the successful implementation of digital transformation, such as ICT infrastructure, data integration, governance, innovation, stakeholder engagement, and sustainability.*

b) Contextual Relevance: The article should focus on the role of digital transformation in enhancing the maturity of smart cities rather than discussing only technological or generic city management aspects.

c) Empirical Focus: The article should provide empirical evidence, case studies, or data that illustrate the application of digital transformation success factors in real-world smart city contexts.

This evaluation ensures that only those articles that directly contribute to understanding the success factors for digital

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2326, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 21, 2025, Published : January 31, 2025

transformation in smart city maturity are included in the final analysis.

D. Data Extraction

Once the articles pass the eligibility evaluation, relevant data is extracted for further analysis. This phase systematically gathers information related to the identified success factors of digital transformation in smart cities. The data extraction follows a standardized template to maintain consistency and ensure that only the necessary and pertinent information is collected. The key data points extracted include:

a) Success Factors: Articles that identify or discuss key factors contributing to the success of digital transformation in smart cities are prioritized. These factors may include technology infrastructure, data-driven decision-making, stakeholder involvement, government policies, and sustainability practices.

b) Best Practices: Articles highlighting best practices in implementing digital transformation in smart cities are extracted to provide practical insights for other cities. These best practices may include successful governance models, stakeholder collaboration strategies, or technological implementations that have led to positive outcomes in specific cities or regions. Best practices offer valuable lessons for cities aiming to accelerate their journey toward smart city maturity.

c) Case Studies and Examples: Extracting case studies or practical examples where digital transformation success factors have been successfully implemented in cities. These case studies provide real-world evidence of how these factors are applied in the context of smart city development.

E. Data Analysis, Synthesis, and Validation

The selected literature was analyzed using a thematic approach to identify success factors in digital transformation that contribute to smart city maturity. This process involved thematic coding and identifying key recurring themes in the literature, such as the role of technology infrastructure, innovation governance, stakeholder collaboration, and strategy sustainability. These findings were then synthesized to provide more structured insights into the factors that support the development of mature, smart cities. After the analysis, the results are validated by comparing the main findings with relevant case studies or empirical reports. This step ensures the research findings are practically applicable and relevant to smart city development.

III. RESULT AND DISCUSSION

A. Result of Literature Review

Based on the analysis of 17 research papers, it was identified that the success of digital transformation in accelerating smart city maturity hinges on six key factors: ICT Infrastructure, Data Integration, Government Policies and Strategic Planning, Stakeholder Engagement, Environmental Sustainability, and Leadership and Human Resource Development. These factors have been demonstrated through best practices adopted by various cities globally, showcasing effective strategies to achieve higher maturity levels. While distinct in their focus, these factors are deeply interconnected, collectively shaping and driving the development of smart cities cohesively.

Figure 2 illustrates the six success factors essential for

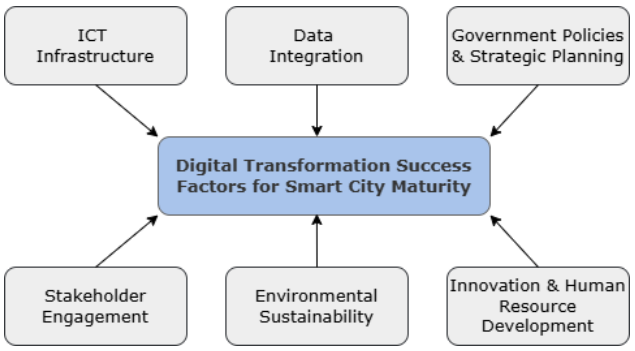


Figure 2. Digital Transformation Success Factors for Smart City Maturity

achieving Smart City maturity through digital transformation. These factors, ICT Infrastructure, Data Integration, Government Policies and Strategic Planning, Stakeholder Engagement, Environmental Sustainability, and Innovation and Human Resource Development, function as interconnected pillars that collectively drive the development of smart cities. Each factor contributes uniquely to the transformation process while supporting and reinforcing the others, ensuring a holistic and sustainable approach to urban growth.

ICT Infrastructure provides the foundational technology required to support innovations in public services, such as mobility, healthcare, and safety systems while enabling efficient communication and data-sharing processes [7]. Data Integration is a critical enabler, allowing disparate systems to interoperate and ensuring seamless coordination across sectors, thereby supporting efficient governance and sustainable resource management [9]. Government Policies & Strategic Planning are crucial in aligning technological investments with societal needs, ensuring that initiatives are feasible, forward-looking, and adaptable to future challenges [10]. Stakeholder Engagement is vital in fostering inclusivity and encouraging the participation of various groups, governments, private sectors, and communities in the development process. This ensures that smart city initiatives remain relevant and sustainable over time [8]. Environmental Sustainability emphasizes the importance of aligning digital transformation efforts with ecological considerations, leveraging technology to optimize resource management and reduce environmental impact [12]. Finally, Innovation & Human Resource Development ensures the capacity to manage, adapt, and sustain digital innovations by investing in skills, training, and visionary leadership capable of navigating technological and societal shifts [6], [14].

These factors do not operate in isolation but are highly interdependent. For example, robust ICT Infrastructure facilitates effective Data Integration, while supportive Government Policies drive Stakeholder Engagement and public-private partnerships. Similarly, a strong commitment to

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2326, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 21, 2025, Published : January 31, 2025

Environmental Sustainability requires capable leaders and skilled human resources to implement green technologies and practices. By integrating these six success factors, cities can address diverse challenges and achieve higher levels of smart city maturity, as depicted in Figure 2.

Table 2 categorizes and highlights the best implementation practices for the six success factors identified as essential for achieving Smart City maturity through digital transformation. These best practices are derived from cities worldwide that have successfully adopted and integrated innovative approaches to enhance their maturity levels.

TABLE II. DIGITAL TRANSFORMATION SUCCESS FACTORS FOR SMART CITY MATURITY

No.	Success Factor	Best Practice	City/ Country	Study
1.	ICT Infrastucture	Public data management, traffic management, and public service management	Bandung, Surabaya, Depok, Surakarta Indonesia Bad hersfeld, Germany & Brazil	[6], [8], [10], [12], [15]
		Application of IoT, big data, and cloud platforms to monitor vehicles, environment, and city conditions	Shenzhen & Hangzhou, China, Manado, Indonesia & Turkey	[7], [9], [16], [17]
		Development of internet networks, data centers, and cyber security to maintain system reliability and service integration.	Cerete, Columbia, Manado, Indonesia	[11], [18], [19]
		Use of AI and CPS technology to increase responsiveness to community needs	Surakarta, Indonesia	[6]
2.	Data Integration	Using middleware and open systems (open data) for interoperability between services	Depok, Indonesia	[8]
		Integrating data through big data analytics, cloud computing, and intelligent grid for operational efficiency	London, UK, Cerete, Columbia & China	[9], [14], [18]
		Provides a centralized data system and online portal that enables comprehensive integration of information	Brazil	[10], [20], [21]
3.	Government Policies & Strategic Planning	Develop strategic policies that support digitalization, including education, health, and security.	Shenzhen & Hangzhou, China, Manado, Indonesia & Brazil	[7], [10], [11]
		Using business cases as a guide in planning smart city initiatives	Depok, Indonesia	[8]
		Improving economic quality through the development of digital-based educational, tourism, and tourist attraction facilities	Surabaya & Bandung, Indonesia, Sao Paolo & Rio de Janeiro, Brazil	[12], [21], [22]
4	Stakeholder Engagement	Engaging citizens through online participation platforms such as Musrenbang and co-creation	Depok & Manado, Indonesia	[8], [11], [19], [20]
		Leveraging social media, hackathons, and feedback tools to design collaborative solutions	China, Bad hersfeld, Germany, Scheidam, Netherlands	[9], [15], [20], [23]
		Fostering collaboration between the public and private sectors, for example, through revitalizing digital markets	Surakarta, Indonesia	[6]
5.	Environmental Sustainability	Managing waste, air, and water pollution through environmentally friendly technologies	Bandung & Surabaya, Indonesia, Turkey & Shenzhen, China	[7], [12], [17]
		Monitoring air, water, and energy quality using IoT and environmental automation systems	Hangzhou, China	[7]
6.	Innovation & HR Development	Using innovative visualization and planning methods such as the Rich Picture method for smart city initiatives	Surakarta, Indonesia	[6]
		Improving human resource readiness with digital technology training and data-based leadership	Manado, Indonesia & London, UK	[14], [16]
		Managing infrastructure and public services adaptively with trained human resources	Manado, Indonesia	[11]

1) ICT Infrastructure

ICT infrastructure is the main foundation of the development of a smart city, and it supports the dimensions of smart mobility and smart living. Best practices in this case include public data management, big data network development, use of IoT devices, application of AI, and strengthening cybersecurity [6], [7], [9], [11], [19]. The significant impact of the application of this technology is seen in increasing the efficiency of public services, traffic management, and response to natural disasters [10], [12], [15], [16]. Examples of the best implementations can be found in Shenzhen, Hangzhou, and Surakarta, which have integrated digital systems to monitor city conditions in real-time using cloud platforms, environmental sensors, and IoT technology [6], [7]. This technology helps these cities manage traffic and provides valuable information for handling problems quickly and efficiently.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2326, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 21, 2025, Published : January 31, 2025

2) Data Integration

Data integration is essential to support interoperability between various public services, supporting the dimensions of Smart Governance and Smart Environment. Best practices in this regard include the use of middleware, big data, cloud computing, and smart grids [8], [9], [20]. Good data integration, such as that implemented through online portals, enables better coordination between sectors [21]. Cities such as Depok have successfully applied data integration to improve the quality of data-based decision-making [8]. In both cities, using middleware and big data analytics has increased data management efficiency and coordination between sectors, improving public service quality.

3) Government Policies & Strategic Planning

Government policies that support digitalization and strategic planning play a significant role in the success of smart cities, especially in the dimension of smart governance. Policies that support digital-based education, security, and health sectors are top priorities in various cities [10]. For example, in Brazil and Manado, policies focused on developing smart infrastructure and digitizing essential sectors such as education and health have created infrastructure planning that is more adaptive to technology [10], [11]. The impact of this policy is the creation of an ecosystem that is more responsive to the community's digital needs, as well as progress in building smarter and more efficient infrastructure.

4) Stakeholder Engagement

Stakeholder engagement is a key factor in ensuring the sustainability of Smart City initiatives. Best practices in this regard include online Musrenbang in Depok, co-creation through hackathons in Schiedam, and digital market revitalization in Surakarta through public-private collaboration [6], [8], [23]. These successes demonstrate the importance of community participation in decision-making and policy implementation. For example, in Depok, online Musrenbang allowed communities to participate directly in digital-based development planning [8]. In Schiedam, hackathons and co-creation became essential to engage communities in creating technology-based solutions [23]. Surakarta also successfully engaged the private sector in digital market revitalization, strengthening cross-sector collaboration and increasing community participation in smart city development [6].

5) Environmental Sustainability

Environmental sustainability is a significant focus in the Smart Environment dimension. Many cities have leveraged technology to monitor air, water, and energy quality more efficiently. Surabaya and Hangzhou, for example, have used sensor and IoT technology to monitor air pollution and water quality and optimize energy use [7], [12]. This technology contributes to better management of natural resources and helps reduce negative environmental impacts. Using this technology, both cities can more efficiently manage their environment, leading to long-term sustainability and a reduced ecological footprint.

6) Innovation and Human Resource Development

Visionary leadership and human resource capacity development are crucial in ensuring the success of Smart City projects, especially in the Smart People dimension. The cities of Surakarta demonstrate that human resource capacity development through digital training and innovative methods such as Rich Picture can enhance the ability of individuals and organizations to manage Smart City projects [6]. Strong leadership and investment in human resource training and development enable cities to be better prepared for the challenges of technological change and increase the effectiveness of implementing existing technology initiatives.

B. Comparison of Smart City Implementations: Local Context, Challenges, and Opportunities

A comparison of Smart City implementations in various cities shows that each region has a different approach tailored to its local characteristics and needs. Surabaya and Hangzhou, as examples of large cities, utilize high technology to address complex urban challenges, such as air pollution and traffic congestion. In both cities, the use of the Internet of Things (IoT), big data, and cloud platforms have enabled real-time traffic management and air quality monitoring [7], [12]. These technologies not only improve the operational efficiency of the city but also provide solutions to environmental and mobility problems that large cities often face. In contrast, smaller cities such as Schiedam and Depok rely on a more participatory approach in integrating data and involving the community in decision-making. In Schiedam, for example, co-creation through hackathons and citizen participation in online Musrenbang in Depok has contributed significantly to developing technology-based solutions tailored to the needs of the local community [8], [23].

Local context greatly influences how technology is implemented in each city, with large cities tending to rely on advanced technology solutions while smaller cities focus more on empowering citizens. Larger cities, with denser populations and more significant infrastructure challenges, tend to require more sophisticated systems, such as IoT and big data, to manage resources and solve big problems. Smaller cities, on the other hand, focus more on strengthening citizen participation and streamlining administrative processes through digital platforms. The main challenges facing cities transitioning to Smart Cities are budget constraints, lack of coordination between stakeholders, and resistance to technological change. However, there are also significant opportunities, primarily through public-private collaborations, that can open up access to new technologies and the potential for increased citizen engagement that is increasingly open through digital platforms.

IV. CONCLUSION

The findings of this study emphasize that the success of digital transformation in advancing smart city maturity is shaped by six critical factors: ICT infrastructure, data integration, government policies and strategic planning, stakeholder engagement, environmental sustainability, and innovation in human resource development. These factors are pivotal in enabling cities to enhance operational efficiency, sustainability, and responsiveness to urban challenges. Specifically, robust ICT infrastructure facilitates the

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2326, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 21, 2025, Published : January 31, 2025

deployment of smart technologies, while data integration enables real-time decision-making, essential for managing urban dynamics. Strong government policies and strategic planning create a supportive regulatory environment, guiding smart city initiatives. Engaging stakeholders, from citizens to the private sector, ensures that solutions are inclusive and well-adapted to the community's needs. Moreover, environmental sustainability addresses long-term urban viability, while innovation in human resource development ensures the availability of skilled personnel to manage and maintain smart city systems.

The best practices identified, such as using IoT and big data in large cities like Jakarta and Shenzhen and participatory initiatives like hackathons and online Musrenbang in cities like Depok and Schiedam, underscore the importance of context-specific approaches. Larger cities leverage advanced technological systems to tackle complex urban challenges, benefiting from their higher resource base. In contrast, smaller cities often prioritize community-driven, participatory methods, which foster inclusivity and adaptability in the face of limited resources. This contrast highlights the necessity of aligning smart city strategies with local needs, capacities, and socio-political contexts. While challenges such as budget constraints, stakeholder coordination, and resistance to change persist, opportunities for public-private partnerships and greater citizen engagement offer significant pathways for overcoming these barriers. By addressing these factors holistically and tailoring strategies to local conditions, cities can foster innovative, inclusive, and sustainable urban development.

REFERENCES

[1] B. Anthony Jnr, "Managing digital transformation of smart cities through enterprise architecture–a review and research agenda," *Enterp Inf Syst*, vol. 15, no. 3, pp. 299–331, 2021, doi: 10.1080/17517575.2020.1812006.

[2] B. A. Jnr, M. A. Majid, and A. Romli, "A Trivial Approach for Achieving Smart City: A Way Forward towards a Sustainable Society," *21st Saudi Computer Society National Computer Conference, NCC 2018*, pp. 1–6, 2018, doi: 10.1109/NCG.2018.8592999.

[3] S. Alawadhi *et al.*, "Building Understanding of Smart City Initiatives," in *LNCS*, 2012, pp. 40–53.

[4] P. C. Verhoef *et al.*, "Digital transformation: A multidisciplinary reflection and research agenda," *J Bus Res*, vol. 122, no. July 2018, pp. 889–901, 2021, doi: 10.1016/j.jbusres.2019.09.022.

[5] M. De Jong, S. Joss, D. Schraven, C. Zhan, and M. Weijnen, "Sustainable-smart-resilient-low carbon-eco-knowledge cities; Making sense of a multitude of concepts promoting sustainable urbanization," *J Clean Prod*, vol. 109, pp. 25–38, Dec. 2015, doi: 10.1016/j.jclepro.2015.02.004.

[6] R. A. Nugroho, S. G. Prakoso, K. N. Hidayati, and A. A. Rahmawati, "Smart Technology Maturity and Smart City Initiative: Is it inline? A case in Surakarta City," *APICS 2022 - 2022 1st International Conference on Smart Technology, Applied Informatics, and Engineering, Proceedings*, pp. 75–78, 2022, doi: 10.1109/APICS56469.2022.9918743.

[7] J. Liu *et al.*, "Towards sustainable smart cities: Maturity assessment and development pattern recognition in China," *J Clean Prod*, vol. 370, no. August, p. 133248, 2022, doi: 10.1016/j.jclepro.2022.133248.

[8] M. A. Juniawan, P. Shandhyaduhita, B. Purwandari, S. B. Yudhoatmojo, and M. A. A. Dewi, "Smart Government Assessment Using Scottish Smart City Maturity Model: A Case Study of Depok City Muhammad," in *ICACSIS*, IEEE, 2017, pp. 99–104.

[9] H. P. Lu, C. S. Chen, and H. Yu, "Technology roadmap for building a smart city: An exploring study on methodology," *Future Generation Computer Systems*, vol. 97, pp. 727–742, 2019, doi: 10.1016/j.future.2019.03.014.

[10] R. A. Afonso, K. Dos Santos Brito, C. H. Do Nascimento, V. C. Garcia, and A. Álvaro, "Brazilian smart cities: Using a maturity model to measure and compare inequality in cities," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, May 2015, pp. 230–238. doi: 10.1145/2757401.2757426.

[11] K. Pangauw, F. Usman, and A. Yudono, "Smart City Maturity Level and Quality of Life Index in Manado City," *International Journal of Science and Research (IJSR)*, vol. 10, no. 8, pp. 182–186, 2021, doi: 10.21275/SR21722000019.

[12] H. S. Firmansyah, H. Supangkat, A. A. Arman, and R. Adhitya, "Searching Smart City in Indonesia Through Maturity Model Analysis (Case Study in 10 Cities)," in *The International Conference on ICT for Smart Society (ICISS)*, Bandung, 2017.

[13] M. J. Page *et al.*, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *International Journal of Surgery*, vol. 88, Apr. 2021, doi: 10.1016/j.ijsu.2021.105906.

[14] F. V. Aragão *et al.*, "Smart Cities Maturity Model—A Multicriteria Approach," *Sustainability (Switzerland)*, vol. 15, no. 8, Apr. 2023, doi: 10.3390/su15086695.

[15] C. Anschütz, K. Ebner, and S. Smolnik, "Size does matter: A maturity model for the special needs of small and medium-sized smart cities," *Cities*, vol. 150, no. August 2023, p. 104998, 2024, doi: 10.1016/j.cities.2024.104998.

[16] M. Nur, J. R. Batmetan, and H. K. Manggopa, "Smart City Maturity Level Analysis Using ITIL Framework," vol. 299, no. Ictvet 2018, pp. 243–247, 2019, doi: 10.2991/ictvet-18.2019.55.

[17] H. Bayraktar, D. Y. Bayar, B. Kara, and G. Bilgin, "Leveraging maturity assessment to choose the right applications for smart cities: Turkey's approach," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 44, no. 4/W3, pp. 137–142, 2020, doi: 10.5194/isprs-archives-XLIV-4-W3-2020-137-2020.

[18] W. N. Bernal and K. L. G. Espitaleta, "Framework for developing an information technology maturity model for smart city services in emerging economies: (fsce2)," *Applied Sciences (Switzerland)*, vol. 11, no. 22, Nov. 2021, doi: 10.3390/app112210712.

[19] A. Damianou, A. Vayona, G. Demetriou, and V. Katos, "An actionable maturity planning model for smart, circular cities," *Cities*, vol. 140, no. September 2022, p. 104403, 2023, doi: 10.1016/j.cities.2023.104403.

[20] M. Lnenicka *et al.*, "Transparency of open data ecosystems in smart cities: Definition and assessment of the maturity of transparency in 22 smart cities," *Sustain Cities Soc*, vol. 82, no. February, p. 103906, 2022, doi: 10.1016/j.scs.2022.103906.

[21] S. De Santana, É. D. O. Nunes, S. De Santana, É. De Oliveira, N. Smart, and C. Evaluations, "Smart Cities Evaluations through SMM Framework - Sustainability Maturity Model To cite this version : HAL Id : hal-03133364 Smart Cities Evaluations through SMM Framework - Sustainability Maturity Model," 2021.

[22] E. da S. de Santana, É. de O. Nunes, D. C. Passos, and L. B. Santos, "SMM: A Maturity Model of Smart Cities Based on Sustainability Indicators of the ISO 37122," *International Journal of Advanced Engineering Research and Science*, vol. 6, no. 2, pp. 13–20, 2019, doi: 10.22161/ijaers.6.2.2.

[23] I. Bouzguenda, C. Alalouch, and N. Fava, "Examining digital participatory planning: Maturity assessment in a Small Dutch city," *J Clean Prod*, vol. 264, p. 121706, 2020, doi: 10.1016/j.jclepro.2020.121706.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2326, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 21, 2025, Published : January 31, 2025

KMS Protoype using Stohmaier Framework at Association of Indonesian Kindergarten Teachers in South Bangka Regency

Agustina Mardeka Raya[1], Rahmat Sulaiman[2]*
Information System, Faculty Of Information Technology [1]
Informatic Engineering, Faculty Of Information Technology [2]
ISB Atma Luhur
Pangkalpinang, Indonesia
agustinamardekaraya@atmaluhur.ac.id[1], rahmatsulaiman@atmaluhur.ac.id[1]

***Abstract*— IGTKI is a kindergarten teacher association organization that aims to realize the quality of early childhood education with a spirit of professional love and character to be able to face the era of globalization. However, this goal is not fully implemented because of the distance of each school is far, the results of the training are not socialized, the lack of knowledge to handle students with special needs so that the lack of knowledge sharing among PAUD teachers. To increase the knowledge sharing needed a Knowledge Management System application that can be a solution of knowledge sharing that is not hindered by the limitations of time and place. This study uses the Strohmaier Framework for modeling knowledge management systems. Techniques of Analysis and system design are carried out using the object oriented approach method Unifed Manipulation Language (UML). The suitability of the system to the business process was tested using the Forum Group Discussion (FGD) method which was validated using the Fit criteria from Strohmaier's theory which produced a value of 76%, this value shows that the Knowledge Infrastructure designed is in accordance with the business process. Meanwhile, user acceptance of the system was tested using User Acceptance Testing (UAT) in the form of a questionnaire which was calculated based on the Linkert scale, producing a value of 84%. From this value it can be concluded that the level of user acceptance at IGTKI is good for the KMS created.**
***Keywords*— *Knowledge Management System*, Framework Strohmaeir, *Knowledge Sharing*, IGTKI**

I. INTRODUCTION

Early childhood education is all efforts and actions taken by educators and parents in the process of caring for, nurturing and educating children by creating an aura and environment where children can explore experiences that provide them with the opportunity to know and understand learning experiences obtained from the environment. In relation to the term of educator in Early Childhood Education, based on Law Number 20 Article 40 Paragraph 2, it is stated that the obligations of educators are: (1) to create a meaningful, enjoyable, creative, dynamic and dialogical educational atmosphere; (2) to have a professional commitment to improving the quality of education; and (3) to provide examples and maintain the good name of the institution, profession and position in accordance with the trust given to them [1]. In order to carry out these obligations, educators must have a number of competencies. Competencies as learning agents in Early Childhood Education include: pedagogical competence, personality competence, professional competence, and social competence (Government Regulation no. 19 of 2005: National Education Standards Chapter VI).

IGTKI South Bangka Regency is an organization of Kindergarten teachers in South Bangka Regency. IGTKI is an independent organization that aims to improve the quality of Early Childhood Education assigned to PAUD teachers. To improve the competence of educators or PAUD teachers in South Bangka Regency, teachers are assigned to attend training conducted by the Education Office or other training institutions. IGTKI South Bangka also holds meetings every two months to discuss the implementation and problems in Early Childhood Education in the South Bangka Regency environment and to socialize the results of the training. Every teacher who participates in training is required to make a report on the results of the training provided by the training institution and share the knowledge gained with teachers who do not receive training. Only 10% of the 260 IGTKI South Bangka teachers often make reports on the results of the training.

The training report is only stored in the IGTKI South Bangka Regency secretariat. With the number of kindergartens as many as 53 schools in 8 sub-districts in South Bangka Regency, and not all teachers who participate in training share and socialize the knowledge gained from the training, making reports and knowledge gained from the training is not distributed as it should be. In addition, each teacher has their own way or method in stimulating, guiding, and providing learning activities to produce children's abilities and skills. Realizing this, based on (Permenpan, 2011). every government institution must be ready to utilize the wealth of knowledge, including learning from past experiences. In general, this is manifested in the form of rules and procedures in the organization. The obstacles often faced are that teachers' knowledge and experience are often not distributed, not well documented and may still be in the heads of each teacher.

Knowledge Management is an effort to improve an organization's ability to manage their intellectual assets, knowledge and experience. The aim is to utilize assets to achieve better organizational performance to accelerate the achievement of reform goals efficiently (Permenpan, 2011).

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2372, Copyright ©2025
Submitted : January 8, 2025, Revised : January 17, 2025, Accepted : January 26, 2025, Published : January 31, 2025

In research conducted by (Dewi Driyani and Dewi Mustari, 2015) regarding Web-based Learning Models for Kindergarten using the SECI method to produce a teaching and learning process system to help students get learning information and make it easier for teachers to share learning material information. The research entitled "learning system design using Knowledge Management System to improve the competency of early childhood Education Teachers" this research produces a knowledge management based learning system using the RAD (Rapid Application Development) model to analyze and design the system. Based on the description above, to solve teacher knowledge management problems, the author intends to create a "Knowledge Management System at IGTKI South Bangka Regency using the Strohmaier Framework".

II. METHODS

This study focuses on the formation of a knowledge architecture model based on Strohmaier's theory. This architecture model is formed based on the business processes that exist in the organization. Below is a picture of the stages of the research methodology.

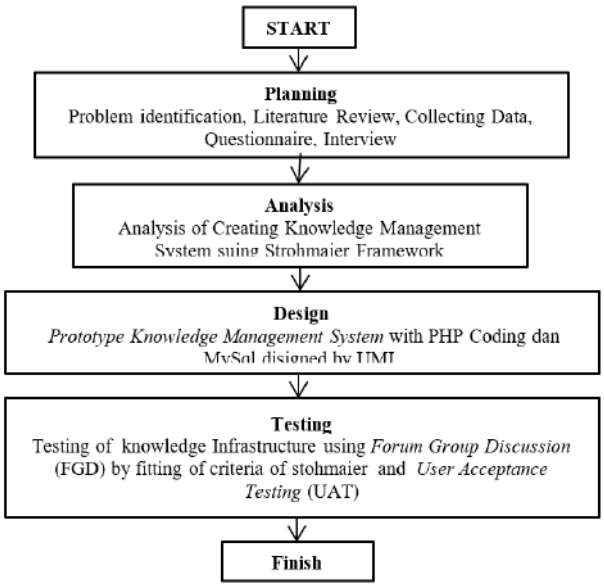


Figure 1. Methodology Research

1. Planning

The initial step taken at the planning stage was to identify problems by conducting interviews and questionnaires with PAUD teachers in South Bangka Regency as well as conducting literature studies in journals and books related to this research. The results obtained are an illustration of the problem formulation.

2. Analysis

At the analysis stage of this research using the Strohmaier Framework. The stages in this Framework are as follows:

a. Business process modeling. At this stage, business processes are used to create working models of organizational knowledge obtained from interviews. Structural modeling is illustrated using diagrams in UML.
b. Identify the relationship between business processes and knowledge. At this stage, several knowledge domains and organizational roles involved in the business process can be identified. In the knowledge identification process, the four processes used are knowledge creation, knowledge application, knowledge transfer and knowledge storage.
c. The development of infrastructure design knowledge is based on a knowledge process and consists of two elements.

1) The architectural knowledge infrastructure template is a system architecture that describes the basic layer of a knowledge management system technology solution

2) Design process to help design the knowledge infrastructure that supports the organization's knowledge business processes. In this design process, the knowledge infrastructure designer helps in integrating and identifying (knowledge) requirements into infrastructure design knowledge. This stage focuses on the process approach to design. This design process uses the knowledge infrastructure process design framework from Strohmaier.

3. Design

The system design technique used in the knowledge management system uses UML (Unified Manipulation Language) with PHP and MySql coding to create a Knowledge Management System Prototype.

4. Testing

The testing stage in this research was divided into two, namely:

a. User acceptance testing of the system uses the User Acceptance Test (UAT) method with Linkert scale calculations.
b. Testing the suitability of Knowledge Infrastructure Design with business processes is based on the Strohmaier Fit Criteria using the Forum Group Discussion (FGD) method. In this phase, it can be determined whether the design that has been created is suitable or not. The fit criteria were designed based on Strohmaier's KMS theory. The indicators used in this test are appropriate or not according to the Guttman scale calculation. The summary of fit criteria is as follows:

1. *Organizational roles must be able to access the knowledge they need.*
2. *The role of the organization must be able to provide / fill knowledge in the Knowledge Infrastructure.*
3. *Knowledge storage must be supported and managed by Knowledge Infrastructure.*
4. *Knowledge Infrastructure must handle knowledge transfer smoothly.*
5. *Knowledge Infrastructure must be able to facilitate knowledge collaboration between knowledge workers.*
6. *The collaboration model must also facilitate the need for collaboration anytime and anywhere (not limited to the workplace only).*

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2372, Copyright ©2025
Submitted : January 8, 2025, Revised : January 17, 2025, Accepted : January 26, 2025, Published : January 31, 2025

The FGD will begin with a focused discussion, the researcher will make a presentation and demo the KMS website that has been developed and explain each existing function based on the instruments that have been prepared. After paying attention and knowing how to operate this KMS website, the participants who attend will be given the opportunity to try this website directly. Next, the FGD participants provided information and responses via the form provided by the researcher before the respondents tried it on their respective PCs. Validation testing form with FGD is included in the attachment. In the FGD testing process, participants will be asked to provide responses in accordance with the hypothesis in this research.

Identification Of Business Processes and Knowledge

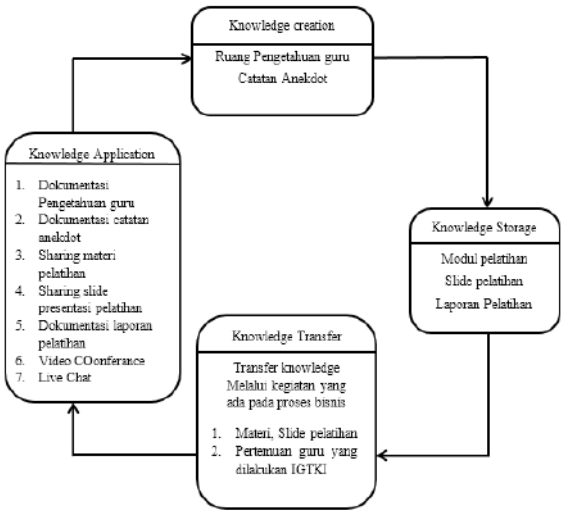


Figure 2. Design of Knowledge Process

This stage is the identification of knowledge

Domains and organizational roles involved in the business process. In the knowledge identification process, the four processes used are knowledge creation, knowledge application, knowledge transfer and knowledge storage. From this knowledge process, the fit criteria will be designed as a goal of knowledge infrastructure design. The fit criteria will be designed based on several KMS theories (Strohmaier, 2005) (Alvi, 2001). The summary of the fit criteria is as follows:

1. The role of the organization must be able to access the knowledge they need.

2. The role of the organization must be able to provide or fill the knowledge Knowledge infrastructure

3. Knowledge must be stored, supported and managed by the knowledge infrastructure.

4. Knowledge infrastructure must handle knowledge transfer. Knowledge infrastructure must be able to facilitate between knowledge and knowledge workers.

5. Collaboration must also facilitate the need for collaboration anytime and anywhere that is not limited to the office.

Design of Infrastructure Design based on fit criteria strohmaier

Table 1 Design of Infrastructure Design based on fit criteria strohmaier

Fit Criteria	Process Knowledge	IT Tools
The role of the organization must be able to provide or fill knowledge in Knowledge infrastructure	Knowledge Creation	Teacher's Knowledge Space, Anecdotal notes
The knowledge stored must be supported and managed by Knowledge infrastructure	Knowledge storage and Retrieval	Training materials, training slides, training reports
Knowledge must be transferable by Knowledge infrastructure	Knowledge transfer	Video Conference, Live Chat, Teacher's Knowledge Space

Knowledge Infrastructure Design

For the design of the system based on the flow of teacher knowledge business process design, there are two strategies related to how individuals (teachers) acquire and share the knowledge needed to perform their duties as teachers. This strategy will affect the design of the knowledge infrastructure. The strategies are as follows: Codification strategy: This strategy focuses on collecting, codifying and disseminating information (explicit). This strategy will be highly dependent on IT. One of the advantages is the reuse of existing knowledge.

The codification strategy in the IGTKI Kab. Bangka Selatan teacher business process can be seen in the table below. Personalization strategy: This strategy focuses on developing networks to connect teachers so that tacit knowledge can be shared. This strategy emphasizes dialogue between individuals. Personalization of teacher business processes can be seen in the table below.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2372, Copyright ©2025
Submitted : January 8, 2025, Revised : January 17, 2025, Accepted : January 26, 2025, Published : January 31, 2025

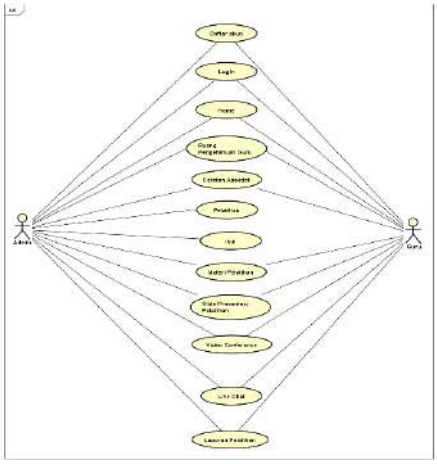


Figure 3. Usecase diagram of KMS IGTK

The image above is a description of what menus can be accessed by admins and teachers on the Knowledge Management System of IGTKI South Bangka Regency. Admin can use all menus according to their functions. Meanwhile, teachers cannot input training menus and participate.

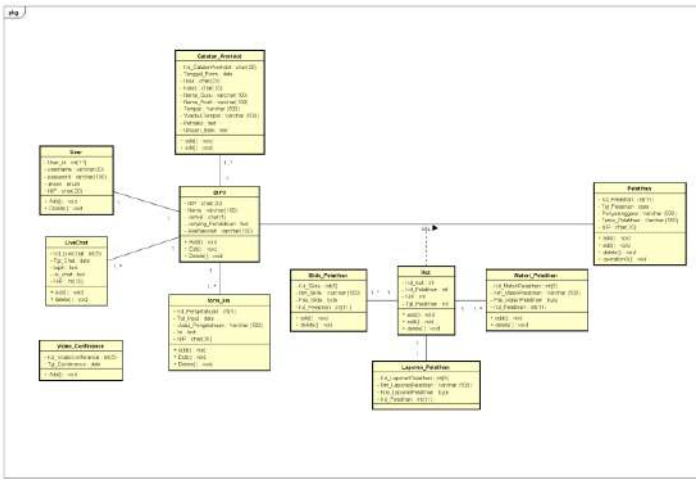


Figure 4. Class Diagram of KMS IGTKI

The Class Diagram above shows the classes of IGTKI teachers in South Bangka Regency and the system admin who is a user in this system who has relationships with several classes related to all knowledge stored in Knowledge Storage.

III. RESULT AND DISCUSSION

Based on the design of the Knowledge Infrastructure that has been made, testing is carried out. Testing uses Fit Criteria Validation to validate the design of the knowledge infrastructure that has been designed in the Forum Group Discussion (FGD). For user acceptance, the User Acceptance Test (UAT) questionnaire method is used for testing.

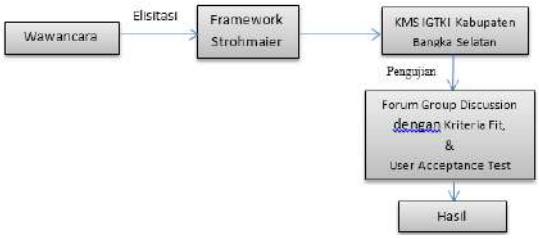


Figure 5. Testing

a. User Acceptance Test

User Acceptance Test is conducted using a questionnaire calculated based on the linkert scale. The weighting of the linkert scale is as follows. The respondents in this questionnaire were 7 respondents, which were then analyzed from each number to obtain an average scores

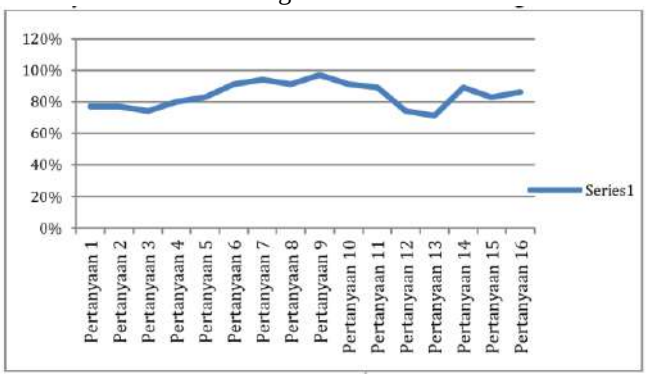


Figure 6. Questions Graph

Figure 5 is a graph of the percentage of each question that has been filled in by respondents in the User Acceptance Test questionnaire. The highest value is in question number 9. The results obtained from the calculations that have been carried out based on the linkert scale are 84%.

b. UAT Validity and Reliability Test

UAT questionnaire data was processed using SPSS to see the validity and reliability results. Table 4.12 provides information on the number of samples (n) of 7 respondents and produces a valid value of 100%.

Table I. Result of "Case Processing Summary" UAT

	N	%
Case Valid	7	100.0
Exclude d [a]	0	0
TOTAL	7	100.0

Listwise deletion based on all variables in the procedur

Table 4 describes the results of item-total statistics for UAT regarding the validity of questions in the UAT questionnaire with a total of 7 respondents, so the df value = 7-2 = 5.

Table II. Result of "Reliability Statistics"

Reliability Statistics	
Cronbach's Alpha	N of Items
.976	8

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2372, Copyright ©2025
Submitted : January 8, 2025, Revised : January 17, 2025, Accepted : January 26, 2025, Published : January 31, 2025

Table 2 is the result of the reliability test after conducting three validity tests. The final result of the Cronbach's Alpha value of this reliability test is 0.976. With a Cronbach's Alpha value of 0.976> 0.60, it can be concluded that the questionnaire items as a whole are reliable.

c. *Focus Group Discussion* (FGD)

Implementation of a Forum Group Discussion (FGD) aims to focus testing on functional requirements. Considering the long distance between sub-districts, 7 people were collected. The FGD was validated using the strohmaier fit criteria to carry out the suitability assessed based on the guttman scale calculation. The weighting of the guttman scale is as follows:

Table III. Guttman Scale

Answer	Score
Sesuai	1
Tidak Sesuai	0

Based on the results of the Guttman scale calculation for the assessment of the suitability of the system with the business process, the value obtained is in the good/adequate category based on the score interpretation

IV. CONCLUSION

The conclusion of the test results obtained in this study is based on the validation of the fit criteria, the knowledge infrastructure designed in accordance with the integrity of the knowledge process based on Strohmaier's theories calculated based on the Guttman scale with a result of 76%. The results of the User Acceptance Test show that the Knowledge Management System at the Kindergarten Teachers Association (IGTKI) of South Bangka Regency is functionally acceptable to users with a weight of 84%, this application can function to provide solutions for IGTKI Teachers of South Bangka Regency and can be accepted by teachers. The results of the two tests according to the interpretation of the score are in the good / feasible category. Based on the two tests, it proves that:

a. With the IGTKI Knowledge Management System of South Bangka Regency, teachers can share materials and results from training obtained from training institutions
b. With the Teacher Knowledge Room and video conferencing features, the IGTKI South Bangka Regency Knowledge Management System, teachers can share knowledge about the behavior of students with special needs, as well as knowledge gained from teaching experience.
c. With the upload file feature on the IGTKI Knowledge Management System of South Bangka Regency, it can improve and make it easier for teachers to document training reports

REFERENCES

[1] Y. N. Sujiono, "Konsep Dasar Pendidikan Anak Usia Dini," 1st ed., Jakarta: Indeks, 2009.

[2] A. Tiwana, *The Knowledge Management Toolkit: Orchestrating IT, Strategy, and Knowledge Platforms (2nd Edition)*. 2002.

[3] Sangkala, *Knowledge Management*. Jakarta: PT. RajaGrafindo Persada, 2007.

[4] M. Strohmaier and K. Tochtermann, "B-KIDE: A framework and a tool for business process-oriented knowledge infrastructure development," *Knowl. Process Manag.*, vol. 12, no. 3, pp. 171–189, 2005.

[5] P. L. Tobing, *Knowledge Management Society Indonesia*. Bandung, 2011.

[6] W. Febriantoro and K. Surendro, "Business Process Oriented Knowledge Management System Design (Case Study Development Center of Metrological Human Resources, Ministry of Trade Republic of Indonesia)," *Asian J. Technol. Manag.*, vol. 8, no. 2, pp. 83–95, 2015.

[7] M. Abdellatif and A. Kherbounche, "A Functional Approach Of Knowledge Management System Applied To Institutions Of Higher Education," *IEEE*, vol. 2, 2014.

[8] U. Maman, Y. Sugiarti, and S. Ratnawati, "Learning system design using knowledge management systems to improve the competency of early childhood education teachers," *Int. Conf. Cyber IT Serv. Manag.*, vol. 4, 2016.

[9] A. F. Blanco, M. L. S. Enhaluce, and F. J. G. Penalvo, "Knowledge Spirals in Higher Education Teaching Innovation," vol. 10, no. December, pp. 16–37, 2014.

[10] A.S, R. and Shalahuddin, M. (2015) 'Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek. Bandung : Informatika', *Jurnal Pilar Nusa Mandiri*.

[11] Agani , Samidi, H. (2014) 'Prototipe Knowledge Management System (KMS) Praktikum Perbankan Syariah Berbasis Web dan Android Dengan Kerangka Kerja "Tiwana" Pada Perguruan Tinggi: Strudi Kasus STEI SEBI', *Jurnal Telematika MKOM*, 6, p. 2.

[12] Budiaji, W. (2013) 'Skala Pengukuran dan Jumlah Respon Skala Likert (The Measurement Scale and The Number of Responses in Likert Scale)', *Ilmu Pertanian dan Perikanan*, 2(2), pp. 127–133.

[13] Driyani, Dewi dan Mustari, D. (2015) 'Model Pembelajaran Untuk Taman Kanak-kanank Berbasis Web', *ORBITH*, 11(3), pp. 161–166.

[14] G. Probst, Raub, Steffen, Romhardt, K. (2004) 'Managing Knowledge : Building For Success'.

[15] Malhotra, Y. (2004) 'Measuring Knowledge Assets of a Nation : Knowledge Systems for Development Measuring Knowledge Assets of a Nation ':, *Knowledge Creation Diffusion Utilization*.

[16] Paramita, A. and Kristiana, L. (2013) 'Teknik Focus Group Discussion dalam Penelitian Kualitatif', *Buletin Penelitian Sistem Kesehatan*, 16(2 Apr), pp. 117–127. doi: 10.22435/bpsk.v16i2.

[17] Probst, G., Raub, S. and Romhardt, K. (2001) 'Managing knowledge-Building blocks for success', *Organization Studies Berlin European Group for Organisational Studies*.

[18] Razmerita, Liana, Kathrin K., F. S. (2009) 'Personal knowledge management: The role of Web 2.0 tools for managing knowledge at individual and organisational levels', *Online information review*, 33(6), pp. 1021–1039.

[19] Saraswati, A., Hermadi, I. and Wijaya, S. H. (2019) 'Developmet Knowledge Management System Of Teacher's Competency', *Jurnal Pendidikan teknologi dan Kejuruan*, 25(1), pp. 124–131.

[20] Suharto, A. A. (2012) 'Pengaruh Kualitas Sumber Daya Manusia, KOmitmen Dan MOtivasi Terhadap Kinerja Pegawai Pada inspektorat kabupaten kediri', *Ilmu Manajemen*, 3, pp. 67–79.

[21] Technologies, I., Sein-echaluce, L. and Mathematics, A. (2014) 'Knowledge Spirals in Higher Education Teaching Innovation', 10(December), pp. 16–37. doi: 10.4018/ijkm.2014100102.

[22] Widayana (2005) *Knowledge Management : Meningkatkan Daya Saing Bisnis*. Jakarta: BayuMedia Publishing.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2372, Copyright ©2025
Submitted : January 8, 2025, Revised : January 17, 2025, Accepted : January 26, 2025, Published : January 31, 2025

VGG-16 Accuracy Optimization for Fingerprint Pattern Imager Classification

Agus Andreansyah[1], Julian Supardi[2] *
Department of Masters in Computer Science[1], [2]
Sriwijaya University
Palembang, Indonesia
agusandreansya@gmail.com[1] ,julian@unsri.ac.id[2]

***Abstract*— A fingerprint is a unique biometric identity commonly used as evidence in court. However, the quality of fingerprints can deteriorate due to external factors such as uneven surfaces, weather conditions, or distortion. This study uses the FVC2000 dataset and applies Convolutional Neural Networks (CNNs) to enhance and classify fingerprint images, focusing on patterns such as arches, loops, radial loops, ulnar loops, and twin loops. A novel aspect of this research is the optimization of the VGG-16 model by making specific adjustments to the hyperparameters, including setting the learning rate to 0.0001, using 50 epochs, and selecting a training-to-validation data split of 80%:10%. These adjustments were made to enhance the model's ability to classify complex and varied fingerprint patterns, which typically present challenges to standard CNN models. The results of the study show the highest accuracy of 100% on the test data with the optimized parameters.These findings demonstrate that the optimized VGG-16 model successfully classifies fingerprint images with optimal performance. The real-world implications of achieving 100% accuracy include an increase in the reliability of biometric identification systems, especially for forensic and security applications that require high accuracy to ensure accurate decisions. This study makes a significant contribution to the development of CNN-based fingerprint classification systems, offering a new approach that supports more reliable and precise biometric applications.**

Keywords*— *Fingerprint, Optimization, Classification, VGG-16. CNN

I. INTRODUCTION

Modernization in countries like Indonesia brings changes in customs, culture, and mindset. These changes can lead to both positive and negative effects, such as economic inequality, crime, and juvenile delinquency. Crime can be influenced by internal factors, like age, gender, education, and mentality, as well as external factors, such as time, place, and family conditions. One serious crime is murder, which causes both psychological and material harm. The police investigate to identify the perpetrators of such crimes. [1][2]

Fingerprint pattern identification is a very important aspect in the process of recognizing a person's identity. Fingerprints can also be valid physical evidence and be recognized in legal proceedings in court. However, fingerprint images found at crime scenes (crime scenes) often deteriorate in quality due to various factors. These factors include distortion due to being touched by another party, partially erased patterns, or skin conditions that have experienced wrinkles. [3][4]

To overcome this problem, efforts are needed to improve the quality of fingerprint images so that they can be processed and analyzed better. One effective method for improving image quality and classifying fingerprint patterns is the Convolutional Neural Network (CNN). This method is part of artificial intelligence (AI) that utilizes deep learning to analyze images with a high degree of accuracy. CNNs allow the analysis of fingerprint patterns even though the imagery is of poor quality due to conditions in the field. Thus, the application of CNNs can improve accuracy in the fingerprint identification process, which ultimately helps in the forensic investigation process.[5][6]

In the context of forensics, digital image processing plays a very important role. This processing not only increases the efficiency of law enforcement officials in obtaining more accurate evidence, but can also reduce the potential for errors in the judicial process. Through technologies such as CNNs, fingerprint image processing has become more sophisticated, resulting in more reliable and reliable results. [7][8][9][10]

CNN is a deep learning method inspired by how human brain cells function. It is effective for object classification, especially with large amounts of data. CNNs work by extracting and processing features from input images through multiple layers of neurons. This creates a more informative representation, making classification easier. [11][12][13]

In this study, the CNN method with the VGG-16 architecture will be used, an architecture that has been proven to be reliable in image analysis. To optimize its performance, some important parameters will be adjusted. These settings include:

1. The ratio of training data sharing and validation data to determine the optimal data composition.
2. The weight of the epoch value, which will be set to find the best number of epochs to improve the model's ability to generalize.
3. Learning rate, which is a parameter that controls how much

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2317, Copyright ©2025
Submitted : November 15, 2024, Revised : December 11, 2024, Received : December 12, 2024, Published : January 31, 2025

the model changes its weight during the training process.

In addition, this study will involve various experimental scenarios to test the influence of each parameter on the performance of the CNN model. The results of this experiment are expected to determine the most optimal configuration in improving the accuracy of fingerprint pattern classification. With this research, it is hoped that CNN-based deep learning technology can further contribute to supporting forensic efforts and improving justice in the legal process.

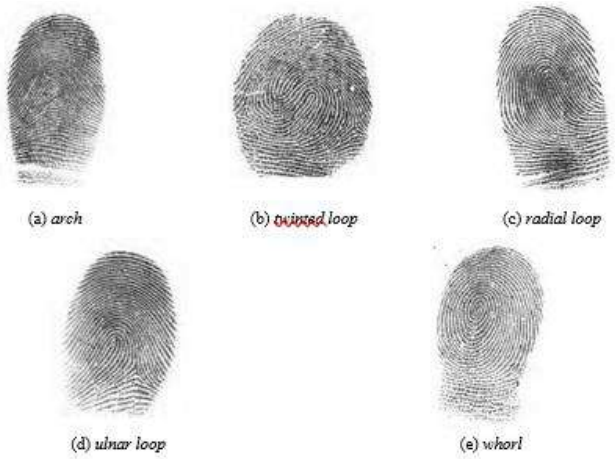


Fig. 1. Types of Fingerprint Image Patterns

Hadaris and Arisy Nabawai explained that dactyloscopy comes from two words in Greek, namely [6]*daktulos*, which means finger line, and *scopeoo*, which means to observe fingers. Therefore, it can be concluded that dactyloscopy is a science that develops the recognition of a person's identity through the observation of the unique pattern of fingerprint strokes found on the fingers and soles of the feet.

The fingerprint used in the identification process is a latent fingerprint obtained from the result of a reproduction that is deliberately taken or stamped with special ink left on an object that has been touched by the skin of the fingers or the soles of the feet. The selection of fingerprints is based on three axioms, namely:

1. Each pattern of a person's fingerprint has unique traits and different physical characteristics in each individual, so it will never be the same as the pattern belonging to another individual.

2. The fingerprint pattern begins to form from the time the fetus is approximately 120 (one hundred and twenty) days old in the womb and will not change until death.

3. Fingerprint strokes can be formulated based on several parameters, such as *core location*, *delta*, *ridge counting*, and *ridge tracing*, thus forming the formula for each fingerprint.

Every individual, whether human, animal, or other living beings, has a unique fingerprint. No two fingerprints are exactly the same, and they remain unchanged throughout life. Fingerprints begin to form around four months of age and continue to grow and change as we age. To capture latent fingerprint patterns, methods like applying oil, amino acid liquid, paint, blood, or nanomaterials are used on the surface, then transferred to paper or plastic.[14][15][16]

There are four main types of fingerprint patterns: ulnar and radial loops, whorls, arches, and tented arches. An arch has a raised center line that forms a slight outward curve. A whorl forms a circular pattern, while a loop creates a curved line that starts and ends on the same side. The point where lines meet in a loop is called the triradius.[2]

Based on the background discussed, this study will perform fingerprint pattern classification using the VGG-16 architecture method with five types of fingerprint patterns as output, namely arch, ulnar loop, radial loop, whorl, and twinned loop.

II. RESEARCH METHODS

This study proposes a fingerprint pattern image classification method using Convolutional Neural Network (CNN) with VGG-16 architecture. The fingerprint images in *.jpg format with dimensions of 256x256 pixels were obtained from the FVC 2000 Database (Fingerprint Verification and Competition). In the training process, fingerprint images will go through several stages, ranging from convolutional layers, max collection layers, dropouts, flattens, to solid layers, with the final output in the form of a classification of 5 types of fingerprint image patterns. Each stage in the CNN network serves to extract features from the image and improve the model's accuracy in recognizing fingerprint patterns. For more details, the flow diagram of this proposed method can be seen in Figures 2 and 3. [17][18]

Convolutional Neural Network (CNN) is a type of deep feed-forward artificial neural network architecture, by maintaining a hierarchical spatial structure to study the representation of internal features in processing grid data (such as images) and pattern recognition. CNNs are designed automatically from the input data through a convolution layer that serves to detect local features, such as edges, angles, or texture patterns in the image. CNNs also have a hidden layer that is only implicitly connected to a subset of neurons in the previous layer.[19]

[20] CNNs have three main types of layers, namely convolutional layers, pooling layers, and fully-connected layers. This architecture results in the extraction of hierarchical features or filters that have a specific purpose. Generally, the first layer is focused on identifying edges or color fluctuations, the second layer is used for shape identification, the next layer is used to study partial parts of the object, and the last layer is used for the identification of the object as a whole

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2317, Copyright ©2025
Submitted : November 15, 2024, Revised : December 11, 2024, Received : December 12, 2024, Published : January 31, 2025

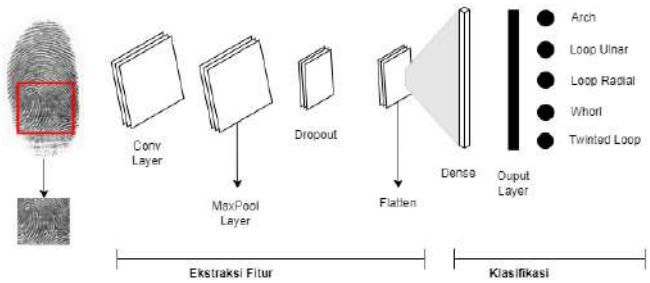


Fig. 2. VGG-16 Architecture Proposal

Based on the image 2, This method uses two stages of processing, namely feature extraction and classification. In the feature extraction process, this stage is responsible for retrieving important information from the input image, i.e. fingerprints. The fingerprint will pass through the first layer, the convolutional layer, which is responsible for picking up local features from the fingerprint image, such as a line or arch pattern. Furthermore, the max pooling layer is used to reduce the dimension of the feature, so that important information is preserved and the data size is reduced. Meanwhile, the dropout layer is used to prevent overfitting and convert the data to a one-dimensional size, which will be processed on a flattened or fully connected layer. The classification process is responsible for classifying the features that have been extracted into one of several fingerprint classes. The dense layer is the fully connected layer that connects all the neurons of the Flatten layer. The output layer is in charge of providing the final result of the classification. Each neuron in this layer represents a class of fingerprints. The fingerprint class consists of five types that are the target of classification, namely arch, ulnar loop, radial loop, whorl, and twinted loop. Of course, each class has a different pattern and is recognized based on the characteristics of each pattern. [21]

This architecture used in his research was explained to be able to achieve a high level of accuracy in the classification of image objects and tends to be very simple. There are several key features in this architecture:[18]

1. The layers in this architecture amount to 16 layers of convolution.
2. The processed kernel is 3x3 in size so there are not many parameters used.
3. In the feature map process, this architecture uses a 2x2 max pooling layer with 2 stride which of course facilitates the classification process and saves time.
4. Using ReLu on the input and output processes using softmax.

[22]Feature extraction is a technique to obtain features and datasets that are used for the storage of fingerprint pattern data later. Feature extraction on CNN consists of two layers, namely convolutional and sub-sampling which have hyperparameters, for hyperparameters applied to the CNN model of the VGG-16 architecture type can be seen in table 1 which is processed at the confiq stage to set all the parameters used in the built model.[23][24]

Table I. Architectural Parameters of VGG-16

Layer	Output Form	Param
VGG16	(7, 7, 512)	14.714.688
conv2d	(7, 7, 32)	147.488
max_pooling2d	(3, 3, 32)	0
Dropouts	(3, 3, 32)	0
flatten	(none, 288)	0
dense	(none, 5)	1.445

In the table 1, VGG 16 is a CNN architecture consisting of 16 layers that utilize transfer learning. The base layer of this VGG-16 model has dimensions of 7x7 with 512 channels. After extraction, it is followed by a convolution layer that has 32 filters for further processing. MaxPooling was used to reduce the spatial dimension from 7x7 to 3x3 while retaining 32 filters. Dropouts are used to prevent overfitting. In the Flatten layer, the data is converted from a 3x3x32 shape to a one-dimensional shape with 288 units. Dense is a fully connected layer that has 5 outputs in the form of a fingerprint image pattern class. [25]

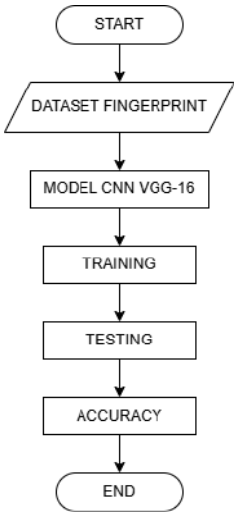


Fig. 3. Flowchart training & testing model

Based on image 3, The flowchart shown shows the steps in the fingerprint pattern recognition process using the Convolutional Neural Network (CNN) method with VGG-16 architecture. Here is an explanation of the steps on the method:

1. At the dataset stage, fingerprints are collected. This dataset can be a fingerprint image of various patterns (loops, circles, arches, or tent arches). The collected data is then prepared for use in the training and testing process of the model. This preparation typically involves several steps, such as pre-processing, image normalization, and data augmentation.
2. The dataset is ready to be fed into the VGG-16 model, which is one of the CNN architectures designed to handle image classification. VGG-16 uses multiple layers of convolution to extract features from fingerprint images, such as line patterns, curves, or circles. This model will process the input data to produce an optimal representation of the feature.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2317, Copyright ©2025
Submitted : November 15, 2024, Revised : December 11, 2024, Received : December 12, 2024, Published : January 31, 2025

3. At this stage, the VGG-16 model is trained using a dataset that has been divided into training data. During the training process, the model learns the patterns on the fingerprint by adjusting the weights and biases through an optimization algorithm. This process aims to minimize the prediction error measured through the loss function.
4. After training is complete, the model is tested using a test dataset (test data). This test dataset is typically not used during the training process, so the test results can represent the model's ability to classify fingerprint patterns on new data.
5. After testing, an evaluation stage is carried out to measure the performance of the model. One of the metrics used is accuracy, which is the ratio between the number of correct predictions and the total number of predictions. Accuracy is a key indicator of how well the model recognizes fingerprint patterns.
6. The process ends after the accuracy evaluation is complete. If the accuracy obtained is not optimal, the model can be improved by adjusting parameters, adding additional data, or changing the configuration at the training stage.

III. RESULTS AND DISCUSSION

The fingerprint pattern type image classification using CNN VGG-16 deep learning architecture obtained the following results:

A. Performance Metrics Results of the CNN Model

The performance metrics results of the CNN model use two types of data, namely secondary and primary data, to assess how well the CNN model performs in making predictions. Additionally, previous researchers also compared their results with those of others who applied their methods to the FVC database. Based on these benchmarks, this study achieved the best accuracy.

Table II. Benchmark Results on the FVC Dataset

Researcher's name	Accuracy (%)
Jeon, Wang-Su	97%
P. Nahar	90%
Ramesh Chandra Sahoo	97.5%
Reena Garg	91%
Yucel Cimtay*	98.5%
Diusulkan (VGG-16) lr 0.001	**97.5%**
Diusulkan (VGG-16) lr 0.0001	**100%**

The benchmark results on the FVC dataset show the accuracy levels of various CNN architectures, focusing on the performance of each architecture using the FVC data. The proposed study with the CNN architecture demonstrates competitive results compared to previous research. At a learning rate of 0.001, the VGG-16 architecture yields the most optimal result, comparable to the study by Ramesh Chandra Sahoo with an accuracy of 97.5%. Meanwhile, at a learning rate of 0.0001, the VGG-16 architecture achieved a perfect accuracy of 100%, while the other two architectures had the same accuracy of 97%.

B. Train and Validate Data Sharing Ratio Training Scenarios

Dividing data into data training (training) and data validation (validation) is an important step that greatly affects the accuracy level of the resulting model. Training data is used to train the model to recognize specific patterns in the dataset, while validation data is used to evaluate the model's performance during the training process, without affecting weight updates. In other words, validation data helps detect overfitting, which is a condition where the model is too adaptable to the training data so that its performance degrades on new data.

Table III. Share Data Ratio

Train Ratio : Validation	Accuracy (%)		
	Train	Validation	Testing
60% : 30%	88.25	92.78	93.28
70% : 20%	91.45	97.50	100.00
80%:10%	90.47	97.50	100.00

In the table 3 shows that the data division with a ratio of 70%:20% and 80%:10% results in the best accuracy on the VGG-16 model, with a final accuracy of 100%. The 60%:30% ratio results in lower accuracy than the other two ratios. The training process uses a learning level of 0.0001 with an adam optimizer and 50 epochs.

C. Training Scenarios with Different Epoch Counts

This study tries to apply various scenarios by using different numbers of epochs to train the model, while maintaining the use of Adam's optimizer and a learning rate value of 0.0001. Epoch refers to the number of complete cycles in which the entire training dataset is processed by the model during training. In this scenario, the number of epochs tested is 10, 30, and 50, with the aim of finding the best number of epochs that results in optimal weight in the fingerprint pattern image classification process.

The data sharing ratio used in this study is 80% for railway data, 10% for validation data, and 10% for test data. This ratio is chosen to ensure the model gets enough data for training, while providing sufficient validation data to evaluate the model's performance periodically during the training process. Data testing is used as a final test to assess the model's ability to recognize fingerprint patterns in new data that has never been seen before.

This experiment with different epochs aims to evaluate the influence of epochs on model performance. Too little age can lead to model underfitting, which is a condition in which the model fails to learn important patterns from the data. Conversely, too much age can lead to overfitting, which is a condition in which the model adapts too much to the training data so that its performance degrades on new data.

By using the Adam optimizer, the model is expected to achieve convergence faster and more efficiently. Adam is an optimization algorithm that is widely used in deep learning model training because of its ability to adaptively adjust the learning rate during the training process. The learning level

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2317, Copyright ©2025
Submitted : November 15, 2024, Revised : December 11, 2024, Received : December 12, 2024, Published : January 31, 2025

used in this study, which is 0.0001, was chosen because this value is often considered a stable conservative option to avoid weight changes that are too large or small in a single update step.

The results of this experiment will be analyzed to determine the best number of epochs that produce optimal accuracy in fingerprint pattern classification. By comparing the performance of the model in each epoch counting scenario, this study is expected to provide insight into the most effective parameter configuration for the VGG-16 model in the context of fingerprint classification. In addition, the study will also evaluate how the number of epochs affects training time, computational efficiency, and the model's ability to generalize new data.

Table IV. Comparison of the Number of Epoch

Epoch	Accuracy (%) 0.0001		
	Train	Validation	Testing
10	55.77	87.50	87.50
30	84.26	97.50	100.00
50	90.47	97.50	100.00

In the table 4, The best results in the table above were obtained using 50 epochs, with the accuracy of train data reaching 90.47%, validation data accuracy of 97.50%, and test data accuracy reaching 100%. This shows that the use of 50 epochs results in the most optimal model for fingerprint pattern image classification with very high accuracy at all stages of evaluation.

D. Training Scenarios with Different Learning Levels

The next experiment was carried out by training the CNN model using the VGG-16 architecture by applying different learning speed variations. The learning rate is an important parameter in the model training process, which determines how much the model's weight changes when making updates based on error gradients. A learning rate that is too small can cause the training process to be very slow, while a learning rate that is too large can cause the model to become unstable or even fail to achieve convergence.

In this scenario, the researcher used the Adam optimizer with a number of epochs of 50 and a data sharing ratio of 80% for the train data and 20% for the validation data. This ratio was chosen to provide a good balance between the amount of data used to train the model and the amount of data used to evaluate the model's performance periodically during training. The use of Adam's optimizer was chosen because of its ability to adjust the learning rate adaptively, so that it can help the model achieve optimal performance more efficiently. This study tested three different learning rate values, namely 0.0001, 0.001, and 0.01. Each value has a unique characteristic in influencing the update of the model's weights during the training process.

The main goal of this experiment is to determine the most optimal learning rate value, which is the value that provides the best balance between stability and convergence speed, and results in the highest accuracy in fingerprint pattern classification. The results of this scenario will be compared based on evaluation metrics such as accuracy, loss, and the model's ability to generalize validation data.

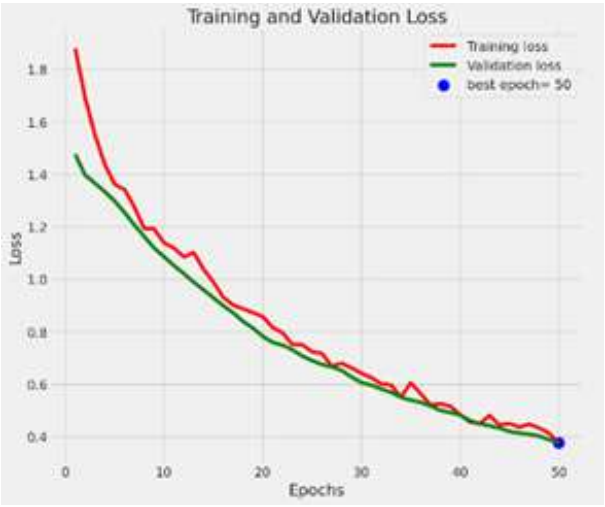
Table V. Comparison of Learning Levels

Learning Level	Accuracy (%)		
	Train	Validation	Testing
0.0001	90.47	97.50	100.00
0.01	87.49	87.50	97.50
0.001	99.69	95.00	97.50

Based on the table 5, experiments using learning rate variations show different accuracy results on the VGG-16 model with Adam optimizer and 50 epochs. At the learning level of 0.0001, the accuracy of the data set reached 90.47%, the validation accuracy was 97.50%, and the testing accuracy reached 100%, which is the best result in the test data. The learning level of 0.01 resulted in lower accuracy, namely 87.49% in train data, 87.50% in validation data, and 97.50% in test data. Meanwhile, the learning level of 0.001 provides the highest accuracy in the data series of 99.69%, with a validation accuracy of 95.00%, and a testing accuracy of 97.50%. From these results, it can be concluded that a learning rate of 0.0001 provides the best accuracy in the test data, while a learning rate of 0.001 indicates the best performance in the training data.

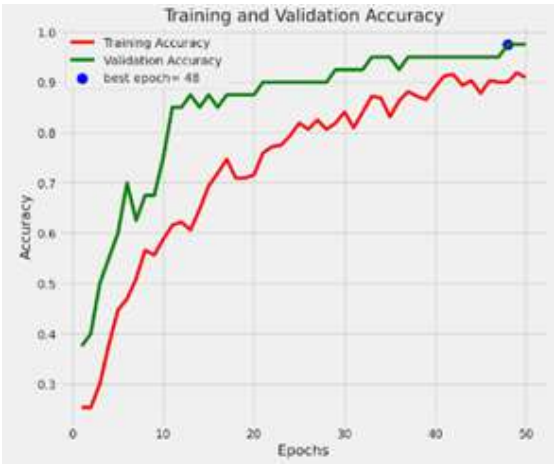
E. Best Results

Of the various experimental scenarios that have been carried out by changing various parameters, the best results are obtained on the CNN model with the VGG-16 architecture. The optimal configuration obtained is by using a data sharing ratio of 80%:10%, Adam optimizer, learning rate 0.0001, and training with 50 epochs. Based on this configuration, we can analyze the accuracy and loss graphs during the training process, as well as review the classification results using the confusion matrix to see the overall performance of the model in the fingerprint pattern image classification.



(a)

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2317, Copyright ©2025
Submitted : November 15, 2024, Revised : December 11, 2024, Received : December 12, 2024, Published : January 31, 2025



(b)

Fig. 4. A and B are Accuracy and loss charts

Overall, based on image 4 graph shows that the model managed to learn the data very well. This is demonstrated by a consistent decrease in loss value and a steady increase in accuracy during the training process. This pattern reflects that the model is progressively getting better at understanding the relationship between the input and output features of the target. In addition, there is no significant indication of overfitting, as the loss and accuracy lines for training and validation data remain close together during the training process. This consistency shows that the model not only performs well on the training data but is also able to maintain good generalization on the validation data.

The choice of using 50 epochs is based on several considerations. First, it helps prevent overfitting, as training the model for too many epochs could lead it to learn irrelevant details or noise from the data. Second, simpler models like VGG-16 often reach optimal performance with fewer epochs, so 50 epochs may be sufficient. Additionally, training deep learning models requires significant time and computational resources, and 50 epochs offer a balance between training time and achieving good results. Techniques like early stopping can also be used, which halts training once the model reaches its best performance, avoiding unnecessary training. Lastly, previous experiments may have shown that 50 epochs are effective for the given dataset, ensuring the model performs well without excessive training.

This solid result is further reinforced by evaluation using a confusion matrix, as can be seen in the image below. The confusion matrix provides a detailed picture of the model's ability to classify data into the correct classes, thus supporting a more in-depth analysis of the model's performance.

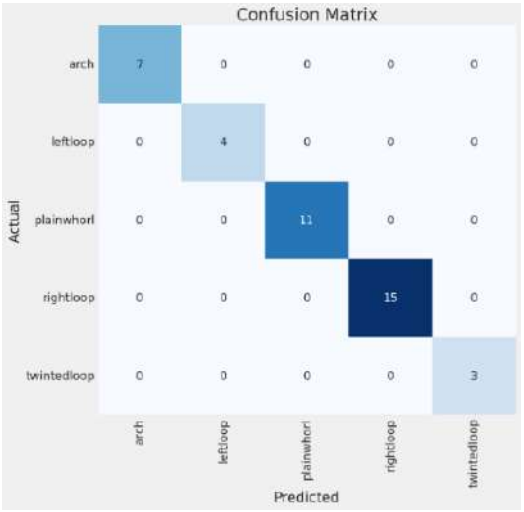


Fig. 5. Confusion matrix data testing

Based on image 5 this confusion matrix, the CNN model used with the VGG-16 architecture shows excellent performance and has a very high degree of accuracy in classifying fingerprint patterns. These results show that the model is able to recognize and group all fingerprint patterns precisely without classification errors, as indicated by the absence of error values in this matrix. With perfect accuracy, this model proves its ability to handle fingerprint data in this regard, making it very suitable for application in fingerprint classification systems. This success also demonstrates the effectiveness of the VGG-16 architecture in understanding the complex features of fingerprint patterns, thus providing confidence for the application of this model to biometric identification systems or similar applications in the future.

IV. CONCLUSION

Based on the results of several training scenarios using Convolutional Neural Network (CNN) with VGG-16 architecture, some interesting conclusions are obtained as follows:

1. The model trained using the VGG-16 architecture successfully classifies all test data with a total of 40 data. These results show that the model has excellent ability to recognize and classify fingerprint patterns into appropriate classes. This success gives an indication that the architecture used is able to capture complex patterns in fingerprint data, resulting in accurate predictions.

2. The best data sharing ratio obtained is 80% for training data, 10% for validation data, and the rest is used for testing. With this sharing, the model is able to learn effectively from the training data which includes most of the available data. The validation data of 10% provides an opportunity for the model to be evaluated periodically during the training process, so as to avoid overfitting. This proportion proved to be a balanced choice between the availability of sufficient training data and representative validation.

3. The highest accuracy was achieved when the model was

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2317, Copyright ©2025
Submitted : November 15, 2024, Revised : December 11, 2024, Received : December 12, 2024, Published : January 31, 2025

trained with a learning rate of 0.0001 and for 50 epochs. A learning rate of 0.0001 allows the model to update the weights gradually and carefully, thus avoiding the risk of overshooting. In addition, training for 50 epochs provides enough time for the model to learn patterns in the data without taking too long to avoid overfitting. This combination of parameters indicates an optimal balance between the speed of model convergence and generalization capabilities.

4. This fingerprint research can make a significant contribution to researchers or forensic applications. Additionally, the study suggests the use of additional architectures and more data to improve results and accuracy in fingerprint identification.

Overall, the results obtained from this experiment show that the VGG-16 architecture has excellent potential to be applied to fingerprint pattern classification tasks. With the appropriate parameter configuration and the right data sharing ratio, this model is able to show very satisfactory performance.

REFERENCE

[1] E. Iriyanto and H. Halif, "UNSUR RENCANA DALAM TINDAK PIDANA PEMBUNUHAN BERENCANA," *Jurnal Yudisial*, vol. 14, no. 1, p. 19, Apr. 2021, doi: 10.29123/jy.v14i1.402.

[2] Triwani, "Pemeriksaan Dermatoglifi sebagai Alat Identifikasi dan Diagnostik".

[3] Agus Andreansyah, Rika Favoria Gusa, and Muhammad Jumnahdi, "Pengenalan Pola Sidik Jari Menggunakan MultiClass Support Vector Machine," *ELKHA*, vol. 11, pp. 79–84, 2019.

[4] Didiek Wahju Indarta, "Fungsi Sidik Jari Dalam Proses Mengidentifikasi Korban dan Pelaku di Kepolisian Resort Bojonegoro," 2021.

[5] A. Z. Hibatullaha, M. F. Rahmanb, and A. P. Saric, "Pemanfaatan Metode Convolutional Neural Network (CNN) Dengan Arsitektur MobileNetV2 Untuk Penilaian Kelayakan Rumah," 2024. [Online]. Available: www.elektro.itn.ac.id

[6] M. Daktiloskopi, A. Tantangan, and D. Harapan, "BADAN PENGEMBANGAN SUMBER DAYA MANUSIA HUKUM DAN HAK ASASI MANUSIA KEMENTERIAN HUKUM DAN HAK ASASI MANUSIA REPUBLIK INDONESIA 2020 MODUL BEST PRACTICE," 2020.

[7] H. Rizkyawan, D. Juardi, J. Haerul Jaman, S. Karawang Jl HSRonggo Waluyo, K. Telukjambe Tim, and K. Karawang, "ANALISIS DIGITAL IMAGE FORENSIK TERHADAP KEASLIAN FILE IMAGE MELALUI WHATSAPP DENGAN METODE NIST," 2024.

[8] C. Militello, L. Rundo, S. Vitabile, and V. Conti, "Fingerprint classification based on deep learning approaches: Experimental findings and comparisons," *Symmetry (Basel)*, vol. 13, no. 5, May 2021, doi: 10.3390/sym13050750.

[9] Rohman Dijaya and Hamzah Setiawan, *Buku Ajar Pengolahan Citra Digital*. UMSIDA PRESS, 2023.

[10] M. Hirsi Mohamed, "Fingerprint Classification Using Deep Convolutional Neural Network," *Journal of Electrical and Electronic Engineering*, vol. 9, no. 5, p. 147, 2021, doi: 10.11648/j.jeee.20210905.11.

[11] F. Nurona Cahya *et al.*, "SISTEMASI: Jurnal Sistem Informasi Klasifikasi Penyakit Mata Menggunakan Convolutional Neural Network (CNN)," 2021. [Online]. Available: http://sistemasi.ftik.unisi.ac.id

[12] M. S. Liyananta, M. Shata, N. Latifah, F. Bimantoro, and T. Informatika, "Program Studi Teknik Informatika," 2024. [Online]. Available: https://www.kaggle.com/datasets/thomasdubail/brain-tumors-256x256

[13] J. Wira and G. Putra, "Pengenalan Konsep Pembelajaran Mesin dan Deep Learning Edisi 1.4 (17 Agustus 2020)," 2020.

[14] M. A. Satriawan and W. Widhiarso, "Klasifikasi Pengenalan Wajah Untuk Mengetahui Jenis Kelamin Menggunakan Metode Convolutional Neural Network," *Jurnal Algoritme*, vol. 4, no. 1, pp. 43–52, 2023, doi: 10.35957/algoritme.xxxx.

[15] U. Nur *et al.*, "BIO EDUCATIO (The Journal of Science and Biology Education) Identifikasi Populasi Pola Sidik Jari di Lingkungan II, Kelurahan Tembung, Kecamatan Tembung," vol. 8, no. 1, p. 69, 2023, doi: 10.31949/be.v6i2.3317.

[16] S. Syarif and M. Baharuddin, "PENERAPAN METODE CONVOLUTIONAL NEURAL NETWORK PADA FACE RECOGNITION UNTUK SMART LOKER," 2023.

[17] R. Garg, G. Singh, A. Singh, and M. P. Singh, "Fingerprint recognition using convolution neural network with inversion and augmented techniques," *Systems and Soft Computing*, vol. 6, Dec. 2024, doi: 10.1016/j.sasc.2024.200106.

[18] D. L. Andreea-Monica, S. Moldovanu, and L. Moraru, "A Fingerprint Matching Algorithm Using the Combination of Edge Features and Convolution Neural Networks," *Inventions*, vol. 7, no. 2, Jun. 2022, doi: 10.3390/inventions7020039.

[19] K. Azmi, S. Defit, and U. Putra Indonesia YPTK Padang Jl Raya Lubuk Begalung-Padang-Sumatera Barat, "Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat," vol. 16, no. 1, p. 2023, 2023.

[20] S. Trabelsi, D. Samai, F. Dornaika, A. Benlamoudi, K. Bensid, and A. Taleb-Ahmed, "Efficient palmprint biometric identification systems using deep learning and feature selection methods," *Neural Comput Appl*, vol. 34, no. 14, pp. 12119–12141, Jul. 2022, doi: 10.1007/s00521-022-07098-4.

[21] D. Sely Wita and D. Yanti Liliana, "Klasifikasi Identitas Dengan Citra Telapak Tangan Menggunakan Convolutional Neural Network (CNN)," *JURTI*, vol. 6, no. 1, 2022, [Online]. Available: https://www.kaggle.com/mahdieizadpanah/birjand-university-

[22] S. A. Suryaman, R. Magdalena, and S. Sa'idah, "Klasifikasi Cuaca Menggunakan Metode VGG-16, Principal Component Analysis Dan K-Nearest Neighbor," *Jurnal Ilmu Komputer dan Informatika*, vol. 1, no. 1, pp. 1–8, Aug. 2021, doi: 10.54082/jiki.1.

[23] A. Prayoga, Maimunah, P. Sukmasetya, Muhammad Resa Arif Yudianto, and Rofi Abul Hasani, "Arsitektur Convolutional Neural Network untuk Model Klasifikasi Citra Batik Yogyakarta," *Journal of Applied Computer Science and Technology*, vol. 4, no. 2, pp. 82–89, Nov. 2023, doi: 10.52158/jacost.v4i2.486.

[24] D. Irfan, R. Rosnelly, M. Wahyuni, J. T. Samudra, and A. Rangga, "PERBANDINGAN OPTIMASI SGD, ADADELTA, DAN ADAM DALAM KLASIFIKASI HYDRANGEA MENGGUNAKAN CNN," 2022. [Online]. Available: http://jurnal.goretanpena.com/index.php/JSSR

[25] R. A. Jasin and V. L. Santoso, "Pengembangan Arsitektur VGG16 dan DCNN7 pada Convolutional Neural Network dalam Melakukan Klasifikasi Pose Yoga," *Jurnal Sistem dan Teknologi Informasi (JustIN)*, vol. 11, no. 2, p. 314, Jul. 2023, doi: 10.26418/justin.v11i2.55533.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2317, Copyright ©2025
Submitted : November 15, 2024, Revised : December 11, 2024, Received : December 12, 2024, Published : January 31, 2025

Classification of User Expressions on Social Media Using LSTM and GRU Models

I Gede Putra Mas Yusadara[1]*, I Gusti Ayu Desi Saryanti[2]
Department of Information Systems, Faculty of Informatics and Computer[1], [2]
Institut Teknologi dan Bisnis STIKOM Bali
Bali, Indonesia
putramas@stikom-bali.ac.id[1], desi@stikom-bali.ac.id[2]

***Abstract*— Social media serves as a platform for sharing information. Through social media, users can interact with others and express their feelings and emotions. Therefore, emotion analysis plays a crucial role in understanding users' conditions regarding various issues and social events. This study aims to compare the performance of emotion classification models in analyzing and identifying users' emotions on social media. The research process includes data preprocessing, training, and model performance evaluation. The dataset used is derived from Twitter social media and is available on Kaggle. It consists of two main columns: text and label, with the latter categorized into six groups. The dataset undergoes several preprocessing techniques to ensure it is ready for model training. The model training process implements the architectures of LSTM and GRU to analyze the emotions contained within the text. The evaluation results show that the model achieves an accuracy of 93% for LSTM and 94% for GRU, indicating that the GRU model slightly outperforms the LSTM in classifying emotions in textual data. This research is expected to contribute to emotion analysis systems based on deep learning.**

***Keywords*— Emotion analysis; social media; sentiment classification; LSTM; GRU.**

I. INTRODUCTION

The advancement of technology today has significantly benefited various aspects of human life. One of the technologies currently evolving is social media, which facilitates communication and information sharing over the internet. Social media generates vast and diverse data daily from the interactions of millions of users on these platforms [1]. Through these interactions, users' emotional expressions can be observed. This makes emotion expression on social media a compelling aspect, as emotions can be identified through textual content. Therefore, this research develops a classification model to understand emotions in social media content, providing insights into users' conditions across various aspects of life, such as public opinion, brand sentiment, trend shifts, product marketing, and more [2].

One of the challenges in analyzing emotions on social media is the complexity of the data. Emotions are often not expressed explicitly and are influenced by various factors such as culture, communication styles, and others. This makes emotion classification a difficult task, requiring effective algorithms to recognize patterns in social media text. Previous studies have demonstrated that machine learning algorithms can be utilized to classify emotions in social media text [3][4][5]. One study conducted in 2019 focused on sentiment analysis of Twitter social media regarding fast-food restaurants. This study compared several machine learning methods for sentiment analysis, with the best results achieved using the bagging method [6]. Subsequently, in 2022, a study aimed to identify emotions in tweets on Twitter, categorizing them into six groups to analyze user behavior and public attitudes toward various global events. This research utilized the SVM method [7]. Finally, in 2024, a study developed a classification model to identify emotions expressed by Twitter users in their tweets, employing the BiLSTM method [8]. These existing studies serve as references for further research using two methods: Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), both of which fall under the category of artificial neural networks [9]. LSTM is used because of its ability to handle the vanishing gradient problem in long sequence data. Meanwhile, GRU is used due to its simpler and more efficient structure, with fewer parameters compared to LSTM. Both methods are capable of analyzing long-term dependencies in sequential data, making them suitable for emotion-related text analysis.

Therefore, this study provides a new contribution regarding the effectiveness of RNN models in emotion classification on social media. Scientifically, it offers insights into the emotion analysis process in the digital realm, capable of identifying patterns in social media interactions. The findings of this research can be useful for developing emotion analysis applications that can be implemented in fields such as marketing, social studies, and public policymaking. Understanding the emotional conditions of social media users can be utilized to devise data-driven strategies for decision-making.

This study aims to develop models using LSTM and GRU algorithms to classify emotions in social media text [10]. The data used to create these models was obtained from Kaggle, a publicly available dataset collected from the Twitter social media platform [11]. The dataset preprocessing, including various techniques to ensure all data contains essential information for building the models. Next, the models are trained using the architectures and parameters of both algorithms LSTM and GRU. The performance of the models is evaluated to assess their effectiveness. Evaluation techniques include metrics such as accuracy, precision, recall, and F1-score to determine the models' ability to classify emotions. The evaluation results of both models are then

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2318, Copyright ©2025
Submitted : January 6, 2025, Revised : January 14, 2025, Accepted : January 24, 2025, Published : January 31, 2025

compared to identify the most suitable algorithm for emotion classification on social media.

Overall, the results of this study are expected to contribute to understanding the emotional condition of users on social media. Additionally, by understanding the emotions expressed in social media interactions, this can be used as input for decision-making and improve the understanding of user communication dynamics. Furthermore, it can also offer potential for promoting products with more targeted marketing strategies through the application of emotion analysis and social studies conducted on social media.

II. RESEARCH METHODOLOGY

A. System Overview

Figure 1 illustrates the workflow followed in this study to develop an emotion classification model for social media users using the LSTM and GRU methods. The workflow begins with data collection from the Twitter social media platform, including text from users' tweets. After the data is collected, preprocessing is performed to prepare it for model development. The preprocessing steps include techniques such as tokenization (splitting the text into words), cleaning (removing punctuation, special characters, and irrelevant symbols), padding (standardizing text length), and lowercasing (converting all text to lowercase). These processes aim to ensure that the dataset is in the proper format and contains relevant information ready for model training.

Next, the preprocessed text data is converted into numerical vectors so that it can be understood by the model. After that, the dataset is split into two parts: the training dataset and the testing dataset. The training dataset is used to train the model using the LSTM and GRU methods [12][13]. Once the model is trained, evaluation is performed to measure how well the model predicts emotions. The evaluation process uses several metrics, including accuracy, precision, recall, and F1-score. Finally, the evaluation results of both models (LSTM and GRU) are compared to determine which model yields the best results in terms of accuracy and prediction performance.

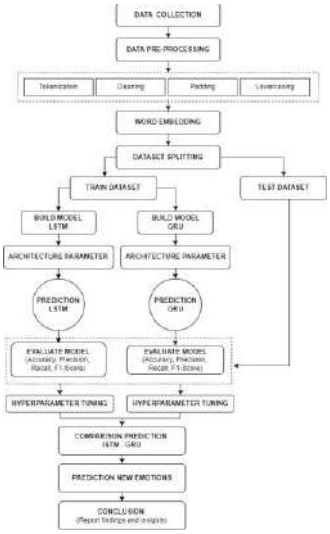


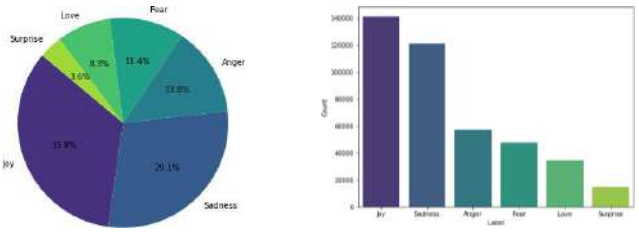
Fig. 1. System Overview

B. Dataset Description

The data used in this study was not directly measured by the researcher. It was obtained from the Kaggle platform, accessible to anyone via the website at www.kaggle.com. This dataset contains text from Twitter messages, labeled with six emotion categories: sadness, happiness, love, anger, fear, and surprise. The dataset consists of two columns: text and label. The text column contains the message data in string format, while the label column contains the emotion labels, which have been converted into numerical values ranging from 0 to 5, as shown in Table 1. Figure 2.a illustrates the distribution of data for each label, and the total number of data rows used is 416,809, with the number of data points for each label shown in Figure 2.b [14].

TABLE I. CATEGORY LABEL CODE

Label Code	Category
0	Sadness
1	Joy
2	Love
3	Anger
4	Fear
5	Surprise



(a) (b)

Fig. 2. Dataset visualization: (a) distribution of categories, (b) count of each category

C. Data Preprocessing

In this section, the data preprocessing steps performed in this study are explained. The initial process begins with splitting the dataset into two parts: 90% for training data and 10% for testing data, using the *train-test split* method. This division ensures that the majority of the data is used for training to effectively train the model. After that, the visualization of the most dominant words across all emotion categories is shown in Figure 3, which helps in understanding the data distribution and provides an initial insight into the emotional patterns emerging in each category [15].

Next, the tokenization process is performed to break the text into word units, making it easier to analyze words separately. This process also converts the text into numerical format by assigning an index to each word. Following that, padding is applied to the data so that all word sequences have uniform length, ensuring the compatibility of the data with the LSTM and GRU models used in this study [16][17].

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2318, Copyright ©2025
Submitted : January 6, 2025, Revised : January 14, 2025, Accepted : January 24, 2025, Published : January 31, 2025

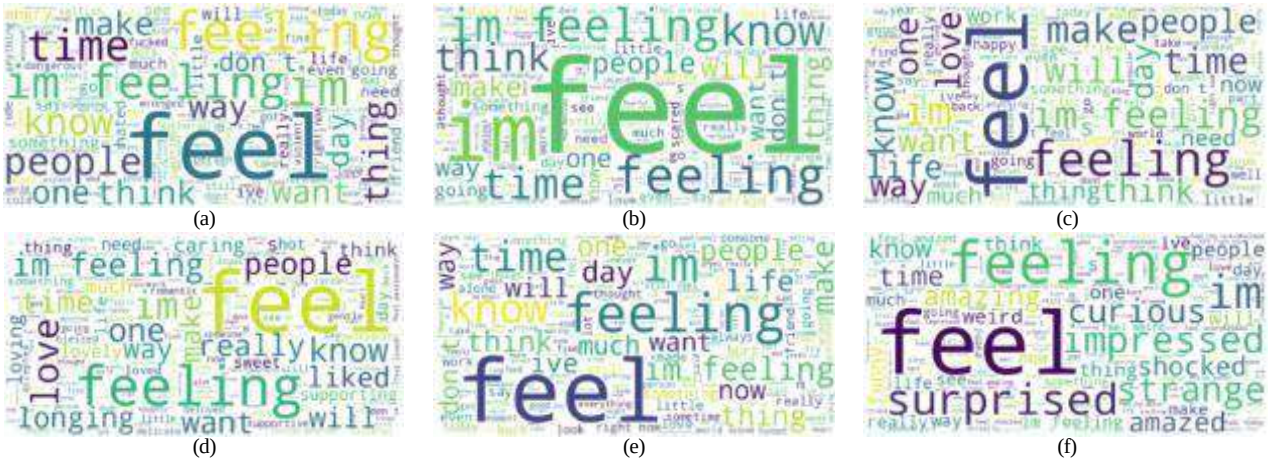


Fig. 3. Word cloud visualization from all categorie:(a) anger, (b) fear, (c) joy, (d) love, (e) sadness, (f) surprise

D. Long Short-Term Memory (LSTM)

The LSTM (Long Short-Term Memory) method is a type of Recurrent Neural Network (RNN) designed to address the vanishing gradient problem commonly encountered in traditional RNNs. In this study, LSTM is used to learn sequential data patterns, such as text. The LSTM process uses three main gates: the input gate, forget gate, and output gate. The input gate identifies important information from the words in the text sequence, which is stored in memory. The forget gate then discards irrelevant information that does not affect the classification process. Finally, the output gate integrates the information to establish relationships between words in the text sequence [18][19].

In its process, LSTM has two types of memory: cell state (long-term memory) and hidden state (short-term memory), which enable the network to learn long-term relationships between data in a time sequence [20]. The architecture of the LSTM model used in this study is explained in Table 2. The LSTM model starts with an Embedding Layer with a dimension of 128. It is followed by three LSTM layers: the first LSTM layer with 15 units, the second with 10 units, and the third with 5 units. After that, a Dropout layer with a ratio of 0.5 is used to reduce overfitting. Finally, the model ends with a Dense layer using 6 units and a Softmax activation function.

TABLE II. LSTM MODEL ARCHITECTURE

Layer	Properties
Embedding	input_dim=5,000, output_dim=128, input_length=max_length
1st LSTM	units=15, return_sequences=True
2nd LSTM	units=10, return_sequences=True
3rd LSTM	units=5, return_sequences=False
Dropout	rate=0.5
Dense	units=6, activation=Softmax

E. Gate Recurrent Unit (GRU)

The GRU method is a part or variation of LSTM that has a simpler architecture with only two types of gates: the reset gate and the update gate. The reset gate controls how much past information needs to be forgotten. Meanwhile, the update gate controls how much new information should be added and also determines how much old memory should be retained [21][22].

The GRU method is lighter compared to LSTM because GRU combines the cell state and hidden state into a single vector, making the computation process more efficient. Although the GRU architecture is simpler, it is still possible for GRU to provide better performance in emotion analysis and classification tasks. Table 3 explains the architecture used to build the GRU model with the same number of layers and units as the LSTM model. The number of units in each LSTM and GRU layer was selected based on experiments to achieve a balance between model complexity and performance. Gradual reduction of units improved generalization. A dropout rate of 0.5 was applied to prevent overfitting, while a Dense layer with 6 units and Softmax activation was used to classify data into 6 emotion categories.

TABLE III. GRU MODEL ARCHITECTURE

Layer	Properties
Embedding	input_dim=5,000, output_dim=128, input_length=max_length
1st GRU	units=15
2nd GRU	units=10
3rd GRU	units=5
Dropout	rate=0.5
Dense	units=6, activation=Softmax

F. Model Evaluation

The technique used for evaluation is the confusion matrix, which is a table containing four combinations of values between predicted and actual outcomes. These four combinations are True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Through these combinations, we can calculate accuracy, precision, recall, and F-measure. The calculations for these values can be explained as shown in Table 4 [23][24][25].

TABLE IV. [22]EVALUATION MATRIX

Matrix	Equation
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2318, Copyright ©2025
Submitted : January 6, 2025, Revised : January 14, 2025, Accepted : January 24, 2025, Published : January 31, 2025

F1-Score $$F_\beta = \frac{1}{\beta \times \frac{1}{precision} + (1-\beta) \times \frac{1}{recall}}$$

III. RESULT AND DISCUSSION

In this section, the results and testing of the model are explained based on the architecture designed using the LSTM and GRU methods. Tables 5 and 6 show the confusion matrix values for the LSTM and GRU models in classifying emotions. The categories Joy and Sadness have the highest accuracy values in both models, with correct predictions of 13,774 and 11,503 for LSTM and then 14,081 and 11,822 for GRU, respectively. Using these confusion matrix values, the model's performance in predicting each emotion category is calculated using accuracy, precision, recall, and F1-score.

Table 7 shows a comparison of the performance calculation results for accuracy between the LSTM and GRU models. The accuracy for the LSTM model is 93%, while the GRU model achieves 94%, which is slightly higher than that of LSTM. These results indicate that both models perform very well in emotion classification. Although the difference in results is minimal, GRU demonstrates a slight advantage in processing text data compared to LSTM.

Next, the performance results for precision, recall, and F1-score of the models are shown in Table 8, Table 9, and Table 10. It can be observed that the GRU model outperforms the LSTM model in several emotion categories. The precision value of GRU for the Love and Surprise categories reaches 100%, and the recall value for the Love category also reaches 100%. These results indicate that the GRU model has an advantage over LSTM in processing text data for emotion classification.

TABLE V. CONFUSION MATRIX – LSTM MODEL

	Anger	Fear	Joy	Love	Sadness	Surprise
Anger	**5,276**	206	26	0	156	0
Fear	138	**4,614**	15	0	39	43
Joy	10	15	**13,774**	280	25	38
Love	1	1	884	**2,609**	0	0
Sadness	263	208	22	3	**11,503**	13
Surprise	0	422	119	0	12	**963**

TABLE VI. CONFUSION MATRIX – GRU MODEL

	Anger	Fear	Joy	Love	Sadness	Surprise
Anger	**5,260**	139	25	0	243	0
Fear	172	**4,473**	7	0	197	0
Joy	24	12	**14,081**	5	16	4
Love	1	1	978	**2,515**	0	0
Sadness	112	64	13	1	**11,822**	0
Surprise	0	425	126	0	14	951

TABLE VII. MODEL ACCURACY PERFORMANCE

Method	Accuracy (%)
LSTM	93
GRU	94

TABLE VIII. MODEL PRECISION PERFORMANCE

Anger	93	94
Fear	84	87
Joy	93	92
Love	90	100
Sadness	98	96
Surprise	91	100

TABLE IX. MODEL RECALL PERFORMANCE

Label	LSTM (%) Recall	GRU (%) Recall
Anger	93	93
Fear	95	92
Joy	97	100
Love	75	72
Sadness	96	98
Surprise	64	63

TABLE X. MODEL F1-SCORE PERFORMANCE

Label	LSTM (%) F1-Score	GRU (%) F1-Score
Anger	93	94
Fear	89	90
Joy	95	96
Love	82	84
Sadness	97	97
Surprise	75	77

Figures 4 and 5 show the visualization of metrics for the LSTM and GRU models, including training accuracy and training loss. The training accuracy of the models (Figures 4.a and 5.a) shows an improvement as the number of epochs increases, indicating that the models are learning from the data used for training. Meanwhile, the training loss of the models (Figures 4.b and 5.b) shows a reduction in the model's error in performing the classification. Overall, these metrics demonstrate good performance of the LSTM and GRU models in emotion classification.

To obtain more comprehensive results, a comparison is made by testing several other machine learning methods. Table 11 presents a comparison of model performance based on evaluation metrics such as accuracy, recall, precision, and F1-score. Based on the results of the tests, it can be concluded that the GRU method achieved the highest accuracy compared to other methods in classifying emotions on social media.

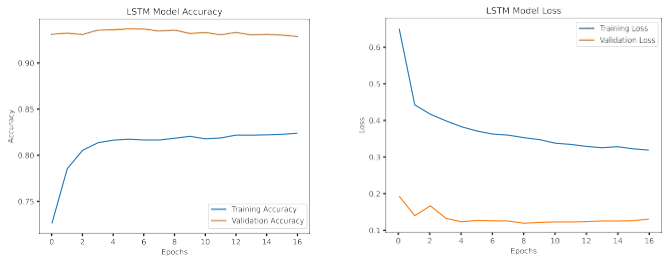
TABLE XI. COMPARISON OF MODEL PERFORMANCE

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
GRU	94	94	94	94
LSTM	93	93	93	93
BiLSTM	93	94	93	93
XGBoost	87	88	87	87
Random Forest	86	86	86	86
DT	84	84	84	84
Naive Bayes	83	85	83	82
AdaBoost	58	60	58	59
KNN	49	75	50	52

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2318, Copyright ©2025
Submitted : January 6, 2025, Revised : January 14, 2025, Accepted : January 24, 2025, Published : January 31, 2025

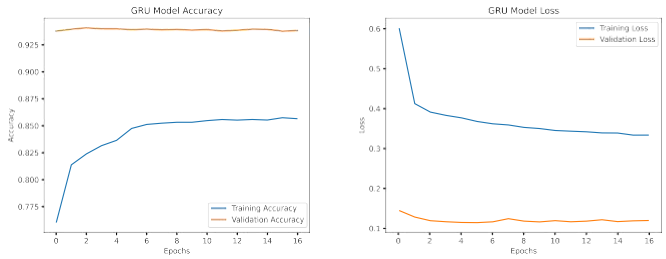
Label	LSTM (%) Precision	GRU (%) Precision

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2318, Copyright ©2025
Submitted : January 6, 2025, Revised : January 14, 2025, Accepted : January 24, 2025, Published : January 31, 2025



(a) (b)

Fig. 4. Visualization of metrics for LSTM: (a) training accuracy, (b) training loss



(a) (b)

Fig. 5. Visualization of metrics for GRU: (a) training accuracy, (b) training loss

ACKNOWLEDGMENT

I would like to express my sincere gratitude to ITB STIKOM Bali for their support throughout this research. I also wish to thank all of my colleagues for their continuous encouragement and assistance, which has been invaluable in the completion of this study.

REFERENCES

[1] D. Miller *et al.*, *How the World Changed Social Media*. 2016. doi: 10.14324/111.9781910634493.

[2] L. Minocha, P. Jain, A. Singh, and P. Pandey, "Social Media's Impact on Business and Society: A Study," in *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2022, vol. 1, pp. 2078–2081. doi: 10.1109/ICACCS54159.2022.9784959.

[3] C. Dayananda, H. P, S. B. Devamane, E. Aravind, I. D, and J. S, "Text Emotion Detection Using Machine Learning Algorithms," in *2023 International Conference on Computational Intelligence for Information, Security and Communication Applications (CIISCA)*, 2023, pp. 304–307. doi: 10.1109/CIISCA59740.2023.00065.

[4] S. M. Pardeshi, D. Chandra Jain, and R. B. Wagh, "A Qualified Study of Machine Learning Algorithms for Sentiment Analysis in Social Media," in *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, 2023, vol. 10, pp. 577–581. doi: 10.1109/UPCON59197.2023.10434795.

[5] M. K. Singh and S. Kumar, "Stress Detection During Social Interactions with Natural Language Processing and Machine Learning," in *2024 International Conference on Expert Clouds and Applications (ICOECA)*, 2024, pp. 297–301. doi: 10.1109/ICOECA62351.2024.00060.

[6] S. A. El Rahman, F. A. AlOtaibi, and W. A. AlShehri, "Sentiment Analysis of Twitter Data," in *2019 International Conference on Computer and Information Sciences (ICCIS)*, 2019, pp. 1–4. doi: 10.1109/ICCISci.2019.8716464.

[7] L. Sahu and B. Shah, "An Emotion based Sentiment Analysis on Twitter Dataset," in *2022 IEEE International Conference on Current Development in Engineering and Technology (CCET)*, 2022, pp. 1–4. doi: 10.1109/CCET56606.2022.10079995.

[8] K. Anbazhagan, S. Kurlekar, T. V Brindha, and D. Sudhish Reddy, "Twitter Based Emotion Recognition Using Bi-LSTM," in *2024 International Conference on Trends in Quantum Computing and Emerging Business Technologies*, 2024, pp. 1–5. doi: 10.1109/TQCEBT59414.2024.10545185.

[9] S. P. Ayunda, N. Qomariasih, R. B. Hadiprakoso, and H. Kabetta, "Comparative Analysis of Deep Learning Models for Web Defacement Detection Based on Textual Context," in *2023 IEEE International Conference on Cryptography, Informatics, and Cybersecurity (ICoCICs)*, 2023, pp. 287–291. doi: 10.1109/ICoCICs58778.2023.10276697.

[10] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao, "Deep Learning-Based Text Classification," *ACM Comput. Surv.*, vol. 54, no. 3, 2021, doi: 10.1145/3439726.

[11] S. Rajendar, S. Ranganathan, T. Lakshmanan, V. Nachimuthu, and S. Paramasivam, "Human Stress Detection on Twitter: A Hybrid Approach using LSTM and Natural Language Processing," in *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, 2023, pp. 714–719. doi: 10.1109/ICCES57224.2023.10192829.

[12] I. W. A. Suranata, "Pengembangan Model Prediksi Curah Hujan di Kota Denpasar Menggunakan Metode LSTM dan GRU," *J. Sist. dan Inform.*, vol. 18, no. 1, pp. 64–73, 2023, doi: 10.30864/jsi.v18i1.603.

[13] I. Wayan Aditya Suranata, I. Komang Agus Ady Aryanto, D. Maneetham, I. Nyoman Kusuma Wardana, and P. Nyoman Crisnapati, "Physical-Chemical Properties Prediction of Chao Phraya River using Deep Learning Methods," in *2024 10th International Conference on Smart Computing and Communication (ICSCC)*, 2024, pp. 601–607. doi: 10.1109/ICSCC62041.2024.10690816.

[14] N. Elgiriyewithana, "Emotions." Kaggle, 2024. doi: 10.34740/KAGGLE/DSV/7563141.

[15] V. Ibrahim, J. A. Bakar, N. H. Harun, and A. F. Abdulateef, "A word cloud model based on hate speech in an online social media environment," *Baghdad Sci. J.*, vol. 18, no. 2, pp. 937–946, 2021, doi: 10.21123/bsj.2021.18.2(Suppl.).0937.

[16] C. P. Chai, "Comparison of text preprocessing methods," *Nat. Lang. Eng.*, vol. 29, no. 3, pp. 509–553, 2023, doi: 10.1017/S1351324922000213.

[17] M. Sergii V. and N. Oleksandr V., "Data preprocessing and tokenization techniquesfortechnical Ukrainian texts," *Appl. Asp. Inf. Technol.*, vol. 6, no. 3, pp. 318–326, 2023, doi: 10.15276/aait.06.2023.22.

[18] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[19] Y. Luan and S. Lin, "Research on Text Classification Based on CNN and LSTM," in *2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 2019, pp. 352–355. doi: 10.1109/ICAICA.2019.8873454.

[20] D. M. Gunarto, S. Sa'adah, and D. Q. Utama, "Predicting Cryptocurrency Price Using RNN and LSTM Method," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 12, no. 1, pp. 1–8, 2023, doi: 10.32736/sisfokom.v12i1.1554.

[21] P. Eswaraiah and H. Syed, "A Hybrid Deep Learning GRU based Approach for Text Classification using Word Embedding," *EAI Endorsed Trans. Internet Things*, vol. 10, pp. 1–8, 2024, doi: 10.4108/eetiot.4590.

[22] S. Li and B. Gong, "Word embedding and text classification based on deep learning methods," *MATEC Web Conf.*, vol. 336, p. 06022, 2021, doi: 10.1051/matecconf/202133606022.

[23] I. G. Ayu Nandia Lestari, D. G. Hendra Divayana, and K. Y. Ernada Aryanto, "A Concentration Selection In Study Programs Using SMOTE Techniques With Ensemble Learning Algorithms," in *2023 5th International Conference on Cybernetics and Intelligent System (ICORIS)*, 2023, pp. 1–6. doi: 10.1109/ICORIS60118.2023.10352192.

[24] G. A. Pradipta, R. Wardoyo, A. Musdholifah, and I. N. H. Sanjaya, "Radius-SMOTE: A New Oversampling Technique of Minority Samples Based on Radius Distance for Learning from Imbalanced

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2318, Copyright ©2025
Submitted : January 6, 2025, Revised : January 14, 2025, Accepted : January 24, 2025, Published : January 31, 2025

Data," *IEEE Access*, vol. 9, pp. 74763–74777, 2021, doi: 10.1109/ACCESS.2021.3080316.

[25] P. D. R. Maulidiana, A. V. Vitianingsih, S. Kacung, A. L. Maukar, and D. Hermansyah, "Comparative Analysis of SVM and NB Algorithms in Evaluating Public Sentiment on Supreme Court Rulings," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 13, no. 2, pp. 230–236, 2024, doi: 10.32736/sisfokom.v13i2.2116.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2318, Copyright ©2025
Submitted : January 6, 2025, Revised : January 14, 2025, Accepted : January 24, 2025, Published : January 31, 2025

Post-Pandemic Usage of Collaboration Tools: Influencing Factors

Aulia Rido Salsabila[1], Nori Wilantika[2]*, Ibnu Santoso[3], Achmad Syahrul Choir[4]
Statistics Polytechnic College of Statistics[2]*,[3],[4]
Indonesian Central Statistics Agency[1]
Jakarta, Indonesia
aulia.rido@bps.go.id[1], wilantika@stis.ac.id[2]*, ibnu@stis.ac.id[3], madsyair@stis.ac.id[4]

***Abstract*—The distance existed due to the COVID-19 pandemic drove people to utilize collaboration tools to continue communication, coordination, and collaboration. The increasing use of collaboration tools during WFH impact on our way of working. The collaboration tools also offer efficiencies, allow workers to break silos, and increase the quality of communication at the company-wide level. As we recover from the pandemic, the government revoked the social distancing policy and it is assumed to influence the continued use of collaboration tools as people fully carry out their activities face-to-face again. This study aims to understand the continuance usage of collaboration tools after no more social distancing. This study also seeks to identify the factors influencing the ongoing use of collaboration tools by integrating the Technology Acceptance Model (TAM) and Expectation Confirmatory Model (ECM). The method of data analysis employed was the partial least squares structural equation model (PLS-SEM). The findings indicated that most of 437 respondents kept using collaboration tools after no more social distancing. However, there was a decrease in the frequency of use. Our study findings have also proved that Actual Continued Usage is influenced by Continuance Intention by 43%. Factor that influences continuance intention the most is the attitude toward using collaboration tools, which is influenced by users' perceived usefulness of the collaboration tools. The results of this study also support the integration of TAM and ECM to examine user intentions and behavior regarding the continuance use of a technology.**

Keywords—Social Distancing; Collaboration Tools; ECM; TAM; PLS-SEM

I. INTRODUCTION

On March 11, 2020, the World Health Organization (WHO) issued the Public Health Emergency of International Concern (PHEIC) status for COVID-19. The Indonesian government responded by enacting the Social Distancing policy on March 31, 2020, because the country had a comparatively high number of COVID-19 cases and deaths. This policy limited people's activities, including work and school activities. Throughout 2020 until the first half of 2022, teaching and learning activities and work activities in Indonesia were mainly carried out from home. Not only in Indonesia, since the COVID-19 pandemic started, several large companies in the United States, China, and Japan have also embraced Work From Home (WFH) [1].

The distance that existed during WFH due to the pandemic raises challenges to communication and coordination among workers. Therefore, people are looking for ways to utilize various software to continue communication, coordination, and collaboration while doing WFH. Software that enables two or more people to work together virtually on a project regardless of where they are physically located is called collaboration tools [2]. Collaboration tools are intended to make multi-person work easier. They enable people to work together on projects, share files, diagrams, photos, papers, and other materials, edit an object or file simultaneously, and view the most recent version of the editing or the same content [2]. Collaboration tools are divided into several types of categories, which are communication (e.g., Slack, Discord) [3], video conference (e.g., Zoom, Google Meet, Microsoft Teams) [3][4][5], cloud storage (e.g., Google Drive, Microsoft OneDrive, Dropbox) [4][6], document construction (e.g., Google Docs/Sheets/Slides) [5], project management (e.g. Jira, Trello, Asana) [2][3], online calendar (e.g. Google Calendar, Microsoft Outlook) [6][7], design (e.g. Figma, Canva, Sketch) [8][9], and software development (e.g. GitHub, GitLab, Bitbucket) [3].

Data shows that there has been an increase in the use of collaboration tools due to social distancing during the pandemic. A study [4] that surveyed German workers shows that 80.7% of respondents' digital tool usage increased during the home office or WFH due to the COVID-19 pandemic. The digital tools are video conference and project management tools [4]. Meanwhile, the software development team in the United States increasingly adopted Slack as a collaboration hub during remote work due to the COVID-19 pandemic [10]. In Indonesia, based on a survey conducted by [11], there has also been an increase in the duration of internet use during the pandemic to work or to do school from home, and the most used app for online meetings is Zoom.

The increasing use of collaboration tools triggered by the shift to WFH during the social distancing has had a remarkable impact on our way of working, especially in communicating and collaborating. Experiences of using collaboration tools make workers no longer rely on face-to-face interaction to discuss [10]. Collaboration tools allow workers to break silos [2] and increase the quality of communication at the company-wide level [6]. They also improve the habit of documenting, organizing, and sharing across the team [2]. Realizing the efficiencies offered by the remote-working concept using collaboration tools, many companies restructure their way of working and transform their working environment [4], as was done by the Indonesian Ministry of Finance, which adopted WFH during the pandemic to become permanent Flexible

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2295, Copyright ©2025
Submitted : October 10, 2024, Revised : November 10, 2024, Accepted : December 12, 2024, Published : January 31, 2025

Working Space and Hours as the new way of agile working [1]. In short, the sustainable use of collaboration tools brings many benefits to the company, mainly to stay competitive.

However, as the conditions of the COVID-19 pandemic got better, since the second half of 2022, the Indonesian government has allowed people to return to the office and school. In the end, on December 30, 2022, the Indonesian government revoked the Social Distancing policy, so people fully carry out their activities face-to-face again. Therefore, it comes to a question: will collaboration tools still be used after no more social distancing? Some researchers expect the increased use of technology during the pandemic will persist even after the pandemic has ended [4][10]. Nevertheless, several researchers argue that although people initially embrace technology, they may eventually stop using it [12]. Thus, this study investigates whether people continue to use collaboration tools even after no more social distancing due to the COVID-19 pandemic. Another purpose of this research is to determine what factors affect people's continued use of collaboration tools.

Previous research on collaboration tool continuance are limited by a particular category or software. For example, [13] and [14] restricted their study to the continuance use of video conferencing apps, [15] only focused on the continuance of cloud storage service, and [12][16] examined the continuance use of Google Docs only. Meanwhile, the use of collaboration tools can't be limited to specific categories because the use of collaboration tools usually combines several types, as mentioned in previous studies like [1],[4], or [10]. This research comprehensively examines the continuance of collaboration tools without being limited to specific categories or software. Additionally, this study looks at the ongoing use of collaboration tools to ascertain whether there have been any changes in their usage following the end of the COVID-19 pandemic. This research extends the literature regarding collaboration tools and the continuance usage of technology, especially thriving technologies during the COVID-19 pandemic.

Most studies on the continuance of collaboration tools use the Technology Acceptance Model (TAM) [12][17][15] and the Expectation Confirmatory Model (ECM) [13][16] as the research model. TAM is mainly used to determine factors contributing to users' initial acceptance and rejection of technology [18]. On the contrary, the ECM model focuses more on factors influencing user retention and loyalty to technology after the initial acceptance [19]. Although there are differences in perspective between TAM and ECM, both are designed to explain various aspects of user perception of system continuance. According to [19], TAM can also be used to examine sustainability intentions and behavior after the initial use of a technology. Moreover, there is an intersection between the ECM and TAM constructs, namely the Perceived Usefulness. Therefore, this research proposes the integration of ECM and TAM to examine the continued use of collaboration tools after social distancing. The integration is expected to provide additional information and increase understanding regarding the continuance use of a system [24].

II. METHODOLOGY

A. Research Model and Hypotheses

Several previous studies on the continuance of collaboration tools used the TAM as the research model, for example, research conducted by [12] to determine the factors that influence the continued use of Google Docs from the perspective of students in Taiwan. Using TAM as a basis, research [17] examines the influence of trust and risk variables on cloud storage usage, specifically Google Drive, in Indonesia. Then, a study by [15] discusses the intention to continue using cloud storage in Taiwan using the Task-Technology Fit (TTF) and TAM models and adds opinions of reference groups and privacy risk variables. Other research used the ECM as the research model. The study conducted by [13] used the ECM model with the addition of several other variables to determine the factors influencing the satisfaction and continuance of video conferencing in Indonesia. At the same time, research [16] used the ECM model by adding prior experience and IT skills variables to investigate factors influencing perceptions of acceptance of Google Docs with the case study in the United States.

However, several research combined TAM and ECM into one research model. Previous research by [20] used TAM and ECM integration to predict the intention of continuing web-based video conferencing for teaching in the post-COVID-19 period from the perspective of academic staff in the United Kingdom (UK). Other studies used ECM and TAM integration to understand the continuance of e-learning [21], ride-hailing apps [22], and cloud-based hospital information systems [23]. Although unrelated to collaboration tools, [21], [22], and [23] show the ability of TAM and ECM integration to explain various aspects of user perception regarding the continuance usage of a technology.

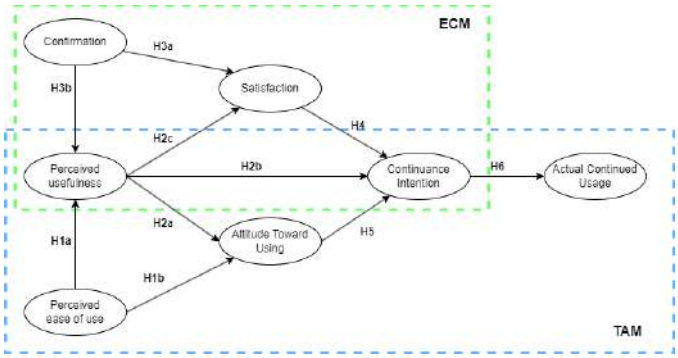


Fig. 1. Research Model.

The research model is based on the integration of the Technology Acceptance Model (TAM) and the Expectation Confirmation Model (ECM), as depicted in Figure 1. This model comprises key constructs such as Perceived Usefulness, Perceived Ease of Use, Attitude, and Continuance Intention, which are essential for understanding user behavior regarding collaboration tools. Based on the research model, we developed the following hypotheses.

H1a: Users' perceived ease of use positively affects their perceived usefulness of collaboration tools.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2295, Copyright ©2025
Submitted : October 10, 2024, Revised : November 10, 2024, Accepted : December 12, 2024, Published : January 31, 2025

According to [24][25][26], Perceived Usefulness is the measure of how users see the advantages of utilizing collaboration tools. At the same time, the term Perceived Ease of Use describes how users feel about how simple it is to use collaboration tools [24]. Perceived Ease of Use affects Perceived Usefulness in TAM. Theoretically, users consider technology to be beneficial when they find it convenient [24][25]. This hypothesis is also stated in [12][17][20].

H1b: Confirmation of the expected performance of collaboration tools positively affects users' perceived usefulness.

When a user compares the performance they expect from collaboration tools with the performance they actually experience, that user's suitability assessment is known as Confirmation [26]. According to ECM, Confirmation boosts the user's perception of its usefulness [26]. Previous research by [13] and [20] has also confirmed the impact of Confirmation on Perceived Usefulness.

H2a: Users' perceived ease of use positively affects their attitude toward using collaboration tools.

H2b: Users' perceived usefulness positively affects their attitude toward using collaboration tools.

Users' attitudes toward using collaboration tools are influenced by their salient beliefs about the outcomes of their actions [24]. The presumption is that users' perceived usefulness and ease of use positively affect their attitudes toward using the collaboration tools. If users' perceive the ease of use and advantages of using the collaboration tools, users' attitudes toward the tools will be more positive [12][17][24].

H3a: Perceived usefulness positively affects users' intention to use collaboration tools.

H3b: Attitude toward using collaboration tools positively affects users' intention to use them.

The user's intention to keep using collaboration tools is known as continuance intention [25][26]. Users' perceived usefulness and attitude toward using collaboration tools are believed to affect users' continuance intention directly. Users will intend to keep using collaboration tools if they perceive the advantages of the tools [13][15][20][27]. Similarly, the user's intention to continue will be influenced by their positive attitude toward using collaboration tools [12][17][25].

H3c: Users' satisfaction with collaboration tools positively affects their continuance intention to use the tools.

Satisfaction is a pleasurable or positive emotional state resulting from collaboration tools [26]. According to ECM, user satisfaction plays a significant role in determining the user's Continuance Intention [13][12][27].

H4a: Perceived usefulness positively affects users' satisfaction with collaboration tools.

H4b: Confirmation of the expected performance of collaboration tools positively affects users' satisfaction.

Perceived Usefulness is considered to affect Continuance Intention, and therefore, Perceived Usefulness is expected to influence Satisfaction as well [26]. In other words, if users feel that using collaboration tools is beneficial, they will be satisfied with them. This hypothesis also has been hypothesized by [13][20][27]. Besides that, the compliance between user expectations and user experiences when using collaboration tools is expected to increase user satisfaction [26]. The influence of Confirmation on Satisfaction has also been previously hypothesized by [13],[20], and [27].

H5: Users' continuance intention positively affects actual continued usage of collaboration tools.

The last hypothesis, Continuance Intention, significantly influences the Actual Continued Usage of collaboration tools [25]. If users intend to use a particular technology, they will use it [17][28].

B. Research Instrument

. Validated items from prior studies were utilized to define the measurement items for each construct in the research model. Each item was modified to align with the specific context of this study. All items were scored on a five-point Likert scale. To ensure clarity, a draft questionnaire was pre-tested with nine respondents similar to the target population. Detailed items are available in the Appendix.

C. Data Collection Method

This study targets Indonesians who have used collaborative tools since March 2020, with "After Social Distancing" defined as post-December 30, 2022, when the Indonesian government revoked the Social Distancing policy. Data were collected over 15 days from May 1 to May 15, 2023, using online questionnaires. A non-probability convenience sampling technique was employed due to the unknown population size, facilitating rapid participant recruitment [29].

For sample size determination, we follow the Ten Times rule for PLS-SEM [30][31], requiring a minimum of 30 respondents based on the maximum number of arrows to the latent variable. However, using Cochran's formula stated in formula (1), we calculated a minimum sample size of 385 to ensure statistical robustness.

$$n = \frac{(z_{\alpha^*/2})^2\, p\, q}{e^2} \quad (1)$$

where n is the minimum sample needed, z is the value in the Normal Distribution curve for deviation α^*, p is the probability of using collaboration tools, q is the probability of not using collaboration tools, and e is the margin of error [29]. The α^* deviation used in this study was 5%, so the z(0.025) was 1.96. The chance that someone uses and does not use collaboration tools is assumed to be 50%. The margin of error used is 5%.

D. Data Analysis Method

Data analysis was conducted using PLS-SEM, chosen for its effectiveness in evaluating complex models. The analysis involved two main stages: measurement model evaluation and structural model evaluation.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2295, Copyright ©2025
Submitted : October 10, 2024, Revised : November 10, 2024, Accepted : December 12, 2024, Published : January 31, 2025

Measurement Model Evaluation

This stage assesses the validity and reliability of the research model through four evaluations:

1. ***Indicator Reliability****: Assessed via Outer Loading, with acceptable values above 0.7 [31].*
2. ***Internal Consistency Reliability****: Evaluated using Cronbach's Alpha and Composite Reliability, both requiring a minimum threshold of 0.7 [30].*
3. ***Convergent Validity****: Measured by Average Variance Extracted (AVE), with an acceptance limit of 0.5 [30].*
4. ***Discriminant Validity****: Assessed through Cross Loading and Fornell-Larcker criteria [30].*

Structural Model Evaluation

The inner model evaluation explores construct relationships and includes:

1. ***Collinearity Assessment****: Evaluated by Variance Inflation Factor (VIF), with values below 5 indicating no multicollinearity.*
2. ***Path Coefficient Analysis****: Hypotheses are supported if t-statistics exceed critical values and p-values are below 0.05.*
3. ***Coefficient of Determination (R^2)****: Indicates prediction accuracy, with values of 0.75, 0.5, and 0.25 signifying strong, moderate, and weak correlations, respectively.*

III. Result and Discussion

A. Demographic Information

From the data collection stage, 437 responses were obtained, with 403 valid entries from 164 males and 237 females. Table 1 displays the demographic information of the participants. Most of the respondents, 75.4%, were between the ages of 18 and 25. Most respondents were also domiciled and worked/studied in Java (76.9%). From the occupation status, it can be seen from Table 1 that half of the respondents (59.6%) are workers who work as civil servants (excluding teachers/lecturers). From educational background, most respondents of this research have completed diploma 4 or a bachelor's degree.

TABLE I. Respondent Characteristics

Variable	Indicator	Frequency	Percentage (%)
Gender	Male	164	40.9
	Female	237	59.1
Age	<18 years old	6	1.5
	18-25 years old	304	75.4
	26-35 years old	48	11.9
	36-45 years old	20	5
	46-55 years old	19	4.7
	>55 years old	6	1.5
Domicile	Sumatera	40	9.9
	Java	310	76.9
	Bali & Nusa Tenggara	13	3.2
	Kalimantan	27	6.7
	Sulawesi	10	2.5
	Maluku	2	0.5
	Papua	1	0.2
Education	Elementary School	0	0
	Junior High School	0	0
	Senior High School	113	28
	Diploma 1/Diploma 3	39	9.7
	Diploma 4/Bachelor	219	54.3
	Magister	32	7.9
	Doctor/PhD	0	0
Occupation Status	Not Working	30	7.4
	Student	133	33
	Work	240	59.6
Occupation	Civil Servant (Exclude Teacher/Lecturer)	108	45
	Private Employees	54	22.5
	Entrepreneur	9	3.8
	Freelancer	21	8.8
	Teacher/Lecturer	29	12.1
	Other	19	7.9
Location of School/ College/ Work	Sumatera	35	9.4
	Java	287	76.9
	Bali & Nusa Tenggara	11	2.9
	Kalimantan	26	7
	Sulawesi	11	2.9
	Maluku	2	0.5
	Papua	1	0.3

B. Outer Model Evaluation

The purposes of evaluating the measurement or outer model is to assess the validity and reliability of the proposed research model. The evaluation consists of 4 stages: indicator reliability, internal consistency reliability, convergent validity, and discriminant validity. Table 2 shows the Outer Loading, Cronbach's Alpha, and Composite Reliability for the first, second, and third evaluation. Two indicators had outer loading values below 0.7, namely PEOU2 and CI2. Therefore, these two indicators should be considered for removal by analyzing the impact of removal on the Composite Reliability and AVE [30].

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2295, Copyright ©2025
Submitted : October 10, 2024, Revised : November 10, 2024, Accepted : December 12, 2024, Published : January 31, 2025

Because there was no significant change in the Composite Reliability and AVE values before and after removing the PEOU2 and CI2 indicators, shown in the last three columns in Table 2, these two indicators are still used in this study. Including PEOU2 and CI2 didn't impact the Composite Reliability and AVE values; for all constructs, Composite Reliability and AVE values have met the threshold. Moreover, several studies show that an indicator with an outer loading value below 0.7 is still acceptable as long as the value is above 0.5 [33][15][21].

TABLE II. OUTER LOADING, CR, CA, AND AVE VALUES

Variable	Indicator	Outer Loading	After PEOU2 and CI2 Removal		
			CR	CA	AVE
Perceived Ease of Use	PEOU1	0.833	0.835	0.884	0.604
	PEOU2	0.691			
	PEOU3	0.739			
	PEOU4	0.740			
	PEOU5	0.826			
	PEOU6	0.821			
Perceived Usefulness	PU1	0.764	0.894	0.919	0.656
	PU2	0.861			
	PU3	0.841			
	PU4	0.825			
	PU5	0.836			
	PU6	0.723			
Confirmation	CONF1	0.827	0.829	0.897	0.745
	CONF2	0.864			
	CONF3	0.897			
Satisfaction	SAT1	0.889	0.850	0.909	0.769
	SAT2	0.874			
	SAT3	0.869			
Attitude Toward Using	ATT1	0.872	0.804	0.884	0.718
	ATT2	0.864			
	ATT3	0.805			
Continuance Intention	CI1	0.802	0.800	0.883	0.715
	CI2	0.614			
	CI3	0.863			
	CI4	0.835			
Actual Continued Usage	ACU1	0.926	0.849	0.930	0.869
	ACU2	0.938			

Note: CR=Composite Reliability, CA=Cronbach's Alpha, the red color indicates the outer loading value below 0,7

TABLE III. FORNELL-LARCKER VALUE

	ACU	ATT	CI	CONF	PEOU	PU	SAT
ACU	0.932						
ATT	0.506	0.848					
CI	0.656	0.555	0.784				
CONF	0.327	0.576	0.398	0.863			
PEOU	0.396	0.550	0.335	0.539	0.777		
PU	0.437	0.600	0.476	0.585	0.613	0.810	
SAT	0.494	0.705	0.537	0.640	0.663	0.663	0.877

Last, the fourth evaluation was discriminant validity by looking at the Fornell-Larcker value. It is evident from Table 3 that every construct's Fornell-Larcker value in our research model satisfies the necessary criteria. In conclusion, from the four stages of measurement model evaluation, all measurement items of each construct in this research have good validity and reliability. The AVE value > 0.5 shows that all measurement items can explain more than 50% of the construct.

C. Inner Model Evaluation

The inner model describes how research constructs relate to one another [32]. The first step of inner model evaluation was checking collinearity by looking at the variance inflation factor (VIF) value. The VIF value must be below 5 to ensure no multicollinearity issues in the research model [32]. VIF values of all indicators in this study are below 5, meaning there is no multicollinearity problem in this research model.

The next step was assessing the path coefficient value to describe the hypothetical relationship between the research constructs [30]. Table 4 shows that the t-statistics of all variables are more than z(0.05)=1.64, and the the p-value of path coefficient values for all variables are less than 0,05, which means that the hypothetical relationship between the research constructs is statistically significant. Based on the evaluation of R2 involving 5 endogenous variables, it can be seen that the value of R2 is between 0.360 and 0.536 (Table 5). This value is considered a moderate to strong relationship [32]. The Q2 value in Table 5 also shows that the exogenous variables in this study have an excellent predictive relevance to the endogenous variables.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2295, Copyright ©2025
Submitted : October 10, 2024, Revised : November 10, 2024, Accepted : December 12, 2024, Published : January 31, 2025

TABLE IV. PATH COEFFICIENT VALUE

Hypothesis		Path Coefficient	t-statistics	p-value	Sig.
H1a	PEOU→PU	0.420	8.574	2.8422E-14	Yes
H1b	CONF→PU	0.358	6.879	9.0097E-12	Yes
H2a	PEOU→ATT	0.292	5.753	7.6349E-09	Yes
H2b	PU→ATT	0.421	8.096	2.8422E-14	Yes
H3a	PU→CI	0.140	2.313	0.01056	Yes
H3b	ATT→CI	0.314	4.875	7.3235E-07	Yes
H3c	SAT→CI	0.222	2.898	0.00196	Yes
H4a	PU→SAT	0.439	8.492	2.8422E-14	Yes
H4b	CONF→SAT	0.383	8.392	2.8422E-14	Yes
H5	CI→ACU	0.656	21.105	2.8422E-14	Yes

TABLE V. R^2 AND Q^2 VALUES

Variable	R^2	Q^2
Perceived Usefulness	0,467	0,300
Satisfaction	0,536	0,407
Attitude Toward Using	0,413	0,289
Continuance Intention	0,360	0,209
Actual Continued Usage	0,430	0,368

D. Use of Collaboration Tools

From 403 valid responses, 401 or 99.5% of respondents still use collaboration tools after revoking the social distancing policy. Only two respondents stopped using the collaboration tools because their jobs didn't need them. This data indicates that most people still sustainably use collaboration tools even after no longer engaging in social distancing. However, the collected data shows that the frequency of using collaboration tools decreases. After social distancing, most respondents use collaboration tools 2-3 times a week with a duration of less than 4 hours per use. This frequency has decreased significantly compared to the period when social distancing policy was still imposed, where most respondents usually used collaboration tools more than once a day for 4 to 9 hours. Although there is a decrease in the use of collaboration tools, in the future, collaboration tools may replace face-to-face as the primary media of collaboration or meetings. This presumption is indicated from the response to the CI2 measurement item: "I intend to continue using collaboration tools than any alternative (face-to-face)". Only 21.09% of all respondents answered Disagree and Very Disagree with that question item.

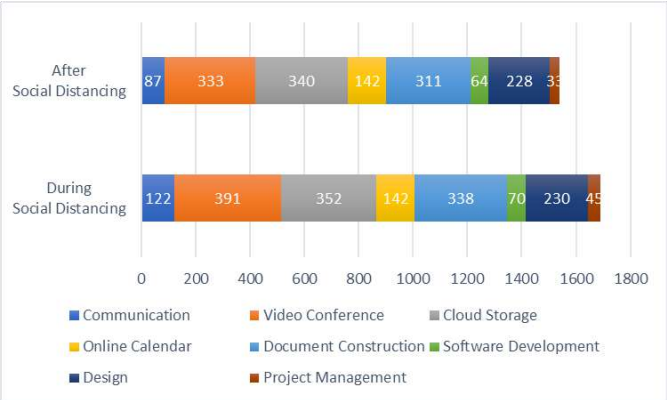


Fig. 2. Most Used Collaboration Tools Based on Category

Then, the categories of collaboration tools that respondents most frequently used were video conferencing, cloud storage, and document construction. As shown in Figure 2, there is no difference in the collaboration tools most frequently used during and after social distancing. More specifically, the most commonly used software in each video conferencing, cloud storage, and document construction category are Zoom, Google Drive, and Google Docs.

E. Factors Influencing Continuance Intention

The purpose of this study is also to find out the factors that influence the intention of the Indonesian people to continue using collaboration tools after social distancing due to the COVID-19 pandemic. This study proves that the use of collaboration tools or Actual Continued Usage (ACU) after social distancing was influenced by the user's Continuance Intention (CI) by 43% (can be seen from R2 in Table 5). Previously, in TAM, it was stated that people's intention affects their future acceptance of the systems [24]. Our research finding shows that intention also influences post-acceptance use as it affects the actual continued usage.

Meanwhile, the Continuance Intention itself is influenced by Perceived Usefulness (PU), Attitude Toward Using (ATT), and Satisfaction (SAT) towards the collaboration tools. In Table 4, among the path coefficient value of H3a, H3b, and H3c, the path coefficient value of H3b is the greatest, followed by H3c and H3a. It means that the Attitude Toward Using is the factor that most influences continuance intention to use collaboration tools, followed by Satisfaction and Perceived Usefulness. This finding is similar to previous research by [12] and [17], which explains that user attitudes affect the user's intention to continue using the system. The attitude toward using collaboration tools itself is influenced by Perceived Usefulness and Perceived Ease of Use (PEOU), which was also proven in research by [12] and [17]. Although both factors significantly and positively affect attitude toward using collaboration tools, users' perceived usefulness has a more significant role than users' perceived ease of use of using the collaboration tools.

Furthermore, Satisfaction is positively influenced by Perceived Usefulness and Confirmation (CONF). In this study, Perceived Usefulness has a more significant role in Satisfaction when compared to Confirmation, as can be seen from the path coefficient in Table 4. This result is following previous research

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2295, Copyright ©2025
Submitted : October 10, 2024, Revised : November 10, 2024, Accepted : December 12, 2024, Published : January 31, 2025

by [13],[20], and [27]. If the user feels that using collaboration tools is beneficial, then the user will be satisfied with the collaboration tools. Last, Perceived Usefulness itself is influenced by Perceived Ease of Use and Confirmation of 46.7%, as shown from R2 in Table 5. However, Perceived Ease of Use has a more significant influence than Confirmation, as shown in Table 4 by the greater path coefficient values. This means that if the user feels the easiness of using collaboration tools, then the user feels the benefits of using collaboration tools. Research using the TAM model that has been carried out by [12],[17],[20] also explains that perceived benefits are influenced by ease of use.

IV. CONCLUSION

This study aimed to examine the continuance of collaboration tool usage in Indonesia following the lifting of social distancing measures due to the COVID-19 pandemic. Our findings reveal that a significant majority of respondents (99.5%) continue to use these tools, albeit with a notable decrease in frequency and duration compared to the pandemic period. Specifically, respondents now use collaboration tools 2-3 times a week for less than 4 hours, whereas during the pandemic, most utilized them more than once a day for 4 to 9 hours.

The analysis indicates that users' continuance intention, which influences ongoing usage by 43%, is primarily driven by their attitudes towards these tools. Furthermore, while both perceived usefulness and perceived ease of use affect user attitudes, perceived usefulness plays a more critical role.

This research contributes to the integration of the Technology Acceptance Model (TAM) and the Expectation Confirmation Model (ECM) in understanding the sustainability of technology usage, particularly in the context of collaboration tools in Indonesia post-social distancing. However, it is important to note that the respondent distribution was concentrated in Java, suggesting a need for broader sampling in future research to enhance the generalizability of these findings. Future studies could explore different methodologies and geographic areas to provide a more comprehensive understanding of collaboration tool usage across Indonesia.

REFERENCES

[1] A. Rifai, M. S. Maarif, and A. Sukmawati, "Key to Succesful Implementation of Flexible Working Space as A New Normality in Public Organizations," *Bus. Rev. Case Stud.*, vol. 2, no. 1, pp. 23–35, Apr. 2021, doi: 10.17358/brcs.2.1.24.

[2] S. Q. Yang and L. Li, *Emerging Technologies for Librarians. A Practical Approach in Innovation.* Massachusetts: Elsevier Ltd., 2016. doi: 10.1016/b978-1-84334-788-0.00013-6.

[3] V. Jackson, A. Van der Hoek, R. Prikladnicki, and C. Ebert, "Collaboration Tools for Developers," *IEEE Softw.*, vol. 39, no. 2, pp. 7–15, Mar. 2022, doi: 10.1109/MS.2021.3132137.

[4] M. Schmidtner, C. Doering, and H. Timinger, "Agile Working During COVID-19 Pandemic," *IEEE Eng. Manag. Rev.*, vol. 49, no. 2, pp. 18–32, Jun. 2021, doi: 10.1109/EMR.2021.3069940.

[5] M. L. Rethlefsen, D. L. Rothman, and D. S. Mojon, "Collaboration Tools," in *Internet Cool Tools for Physicians*, no. August, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 119–128. doi: 10.1007/978-3-540-76382-6_17.

[6] M. Miljanic and N. Zaric, "Review of collaborative software applications and integration with standard collaboration tools," in *2020 24th International Conference on Information Technology (IT)*, Feb. 2020, no. November 2019, pp. 1–4. doi: 10.1109/IT48810.2020.9070420.

[7] A. Thayer, M. J. Bietz, K. Derthick, and C. P. Lee, "I love you, let's share calendars," in *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, Feb. 2012, pp. 749–758. doi: 10.1145/2145204.2145317.

[8] K. Tanner, "Visual Design Tools in Support of Novice Creativity," Standford University, 2019.

[9] P. Farmoudehyamcheh, "A Systems Approach to Graphic Design Practice Recommended Citation," Georgia Southern University, 2019. [Online]. Available: https://digitalcommons.georgiasouthern.edu/etd/2021

[10] V. Jackson, A. van der Hoek, and R. Prikladnicki, "Collaboration tool choices and use in remote software teams," in *Proceedings of the 15th International Conference on Cooperative and Human Aspects of Software Engineering*, May 2022, pp. 76–80. doi: 10.1145/3528579.3529171.

[11] Asosiasi Penyelenggara Jasa Internet Indonesia, "Indonesian Internet Profile 2022," Jakarta, 2022. [Online]. Available: apji.or.id

[12] Y.-M. Huang, "The factors that predispose students to continuously use cloud services: Social and technological perspectives," *Comput. Educ.*, vol. 97, pp. 86–96, Jun. 2016, doi: 10.1016/j.compedu.2016.02.016.

[13] A. I. Safira, P. W. Handayani, and A. A. Pinem, "The Meaning of User Satisfaction and Continuance Intentions with Video Conference Applications," in *2021 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS*, Oct. 2021, pp. 250–256. doi: 10.1109/ICIMCIS53775.2021.9699164.

[14] S. A. Nikou, "Web-based videoconferencing for teaching online: Continuance intention to use in the post-COVID-19 period," *Interact. Des. Archit.*, vol. 47, no. 47, pp. 123–143, Feb. 2021, doi: 10.55612/s-5002-047-006.

[15] H.-L. Yang and S.-L. Lin, "User continuance intention to use cloud storage service," *Comput. Human Behav.*, vol. 52, pp. 219–232, Nov. 2015, doi: 10.1016/j.chb.2015.05.057.

[16] X. Tan and Y. Kim, "User acceptance of SaaS-based collaboration tools: a case of Google Docs," *J. Enterp. Inf. Manag.*, vol. 28, no. 3, pp. 423–442, Apr. 2015, doi: 10.1108/JEIM-04-2014-0039.

[17] F. Gunadi, "Analisa Pengaruh Trust Dan Risk Berbasis Technology Acceptance Models (TAM) (Studi Kasus : Pengguna Google Drive)," *MULTINETICS*, vol. 6, no. 1, pp. 67–77, Aug. 2020, doi: 10.32722/multinetics.v6i1.2819.

[18] N. Marangunić and A. Granić, "Technology acceptance model: a literature review from 1986 to 2013," *Univers. Access Inf. Soc.*, vol. 14, no. 1, pp. 81–95, Mar. 2015, doi: 10.1007/s10209-014-0348-1.

[19] C. Liao, P. Palvia, and J.-L. Chen, "Information technology adoption behavior life cycle: Toward a Technology Continuance Theory (TCT)," *Int. J. Inf. Manage.*, vol. 29, no. 4, pp. 309–320, Aug. 2009, doi: 10.1016/j.ijinfomgt.2009.03.004.

[20] S. A. Nikou, "Web-based videoconferencing in online teaching during the COVID-19 pandemic: University students' perspectives," in *2021 International Conference on Advanced Learning Technologies (ICALT)*, Jul. 2021, pp. 431–435. doi: 10.1109/ICALT52272.2021.00137.

[21] M.-C. Lee, "Explaining and predicting users' continuance intention toward e-learning: An extension of the expectation–confirmation model," *Comput. Educ.*, vol. 54, no. 2, pp. 506–516, Feb. 2010, doi: 10.1016/j.compedu.2009.09.002.

[22] G. Malik and A. S. Rao, "Extended expectation-confirmation model to predict continued usage of ODR/ride hailing apps: role of perceived value and self-efficacy," *Inf. Technol. Tour.*, vol. 21, no. 4, pp. 461–482, Dec. 2019, doi: 10.1007/s40558-019-00152-3.

[23] Y.-M. Cheng, "Drivers of physicians' satisfaction and continuance intention toward the cloud-based hospital information system," *Kybernetes*, vol. 50, no. 2, pp. 413–442, Mar. 2021, doi: 10.1108/K-09-2019-0628.

[24] F. D. . Davis, R. P. . Bagozzi, and P. R. . Warshaw, "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models," *Manage. Sci.*, vol. 35, no. 8, pp. 982–1003, 1989.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2295, Copyright ©2025
Submitted : October 10, 2024, Revised : November 10, 2024, Accepted : December 12, 2024, Published : January 31, 2025

[25] F. D. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Q.*, vol. 13, no. 3, pp. 319–340, 1989, [Online]. Available: http://www.biodiversitylibrary.org/bibliography/33621

[26] A. Bhattacherjee, "Understanding Information Systems Continuance: An Expectation-Confirmation Model," *MIS Q.*, vol. 25, no. 3, pp. 351–370, 2001.

[27] A. P. Oghuma, C. F. Libaque-Saenz, S. F. Wong, and Y. Chang, "An expectation-confirmation model of continuance intention to use mobile instant messaging," *Telemat. Informatics*, vol. 33, no. 1, pp. 34–47, Feb. 2016, doi: 10.1016/j.tele.2015.05.006.

[28] N. Fathema, D. Shannon, and M. Ross, "Expanding The Technology Acceptance Model (TAM) to Examine Faculty Use of Learning Management Systems (LMSs) In Higher Education Institutions," *J. Online Learn. Teach.* , vol. 11, no. 2, pp. 210–233, 2015.

[29] M. Darwin *et al.*, *Metode Penelitian. Pendekatan Kuantitatif*. Bandung: CV Media Sains Indonesia, 2020.

[30] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) Third Edition*. Los Angeles: SAGE Publications, Inc., 2022.

[31] R. R. Marliana, "Partial Least Square-Structural Equation Modeling pada Hubungan Antara Tingkat Kepuasan Mahasiswa dan Kualitas Google Classroom Berdasarkan Metode Webqual 4.0," *J. Mat. Stat. dan Komputasi*, vol. 16, no. 2, p. 174, Dec. 2019, doi: 10.20956/jmsk.v16i2.7851.

[32] J. F. Hair, J. J. Risher, M. Sarstedt, and C. M. Ringle, "When to use and how to report the results of PLS-SEM," *Eur. Bus. Rev.*, vol. 31, no. 1, pp. 2–24, Jan. 2019, doi: 10.1108/EBR-11-2018-0203.

[33] C. Fornell and D. F. Larcker, "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *J. Mark. Res.*, vol. 18, no. 1, pp. 39–50, 1981.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2295, Copyright ©2025
Submitted : October 10, 2024, Revised : November 10, 2024, Accepted : December 12, 2024, Published : January 31, 2025

Temperature and Humidity Monitoring in Hydroponic Cultivation Based on Internet of Things: Dataset Development for Smart Agriculture

Simon Prananta Barus[1]*, Jeriko Ichtus Seo[2]
Informatics Study Program, Faculty of Science, Computers and Mathematics[1], [2]
Matana University
Serpong, Indonesia
simonpbarus@gmail.com[1], jerikoseo@gmail.com[2]

***Abstract*— This research is a continuation of the previous research, entitled "*Development of Hydroponic Application based on Web and Internet of Things for the Community to Monitor pH and Total Dissolved Solids.*" Not only pH and Total Dissolved Solids (TDS) need to be monitored, but also temperature and humidity. This research aims to produce a temperature and humidity monitoring application (in addition to pH and TDS which already exist) in hydroponic cultivation and complete the dataset that supports smart agriculture. The research method includes literature study, hardware development using NodeMCU ESP8266 microcontroller and DHT11 sensor, web-based software development with JavaScript on the Front-End side, PHP on the Back-End side, Apache as Web Server, and MySQL as database management system (DBMS), as well as the implementation stage, integration, system testing and report writing. The results of the research show that the developed system can monitor temperature and humidity in real-time with a good level of accuracy. Not only that, this system can produce a hydroponic dataset that includes temperature and humidity parameters, which can be used for data analysis and improvement of hydroponic management. Thus, this study successfully expanded the scope of the hydroponic monitoring system by adding temperature and humidity parameters. This study contributes to optimizing the hydroponic cultivation system and supporting the development of data-based smart agriculture. Further research will integrate more monitoring parameters, conduct direct hydroponic cultivation trials, and apply artificial intelligence such as machine learning and deep learning to improve efficiency and effectiveness in hydroponic cultivation.**

Keywords—Dataset, Humidity, Internet of Things, Monitoring, Temperature

I. INTRODUCTION

In the era of Industry 4.0, modern agriculture is undergoing significant transformation through the adoption of information technology to enhance efficiency and productivity. Among the most promising advancements shaping the future of farming is hydroponic cultivation [1]. This innovative method not only addresses the challenges posed by limited agricultural land, particularly in urban areas, but also provides a sustainable solution for crop production in environments where soil conditions are suboptimal for plant growth, such as vegetable farming [2].

Hydroponic cultivation plays a vital role in meeting the demand for vegetables at both local and national levels, contributing to the broader objective of strengthening food security [3]. Furthermore, this agricultural approach aligns with Sustainable Development Goal (SDG) 2 — to eradicate hunger, achieve food security, improve nutrition, and promote sustainable agricultural practices [4].

Ensuring the success of hydroponic cultivation relies heavily on the ability to continuously monitor and maintain optimal environmental parameters. Extensive researches have identified critical ranges for key variables such as pH, electrical conductivity (EC), temperature (both nutrient solution and ambient), aeration or dissolved oxygen (DO), growing media or substrates, lighting and artificial illumination, relative humidity, CO2 concentration, and plant-specific requirements [5].

This research builds upon prior research entitled "*Development of Hydroponic Application Based on Web and Internet of Things for the Community to Monitor pH and Total Dissolved Solids*" [6]. The previous research successfully developed an Internet of Things (IoT)-based system capable of monitoring pH and TDS levels. However, temperature and humidity monitoring are not included, leaving a significant gap in ensuring comprehensive environmental control. Temperature and humidity significantly influence plant growth and productivity [7]. Rapid undetected changes in temperature or humidity can lead to reduced quality of plant growth, resulting in reduced quantity and quality of crop yields.

To mitigate such risks, there is a pressing need for an accurate and real-time environmental monitoring system capable of safeguarding optimal growth conditions while generating historical datasets. These datasets not only facilitate comprehensive analysis but also serve as valuable inputs for artificial intelligence-driven agricultural innovations. In addition, the results of the review of articles in journals accredited by the Ministry of Research and Technology stated that there were no articles that specifically developed smart agriculture. [8] Thus, an accurate and real-time environmental monitoring system is needed for hydroponic cultivation to ensure that plant growth conditions remain optimal and that their history is recorded to support the realization of smart agriculture based on data.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2311, Copyright ©2025
Submitted : December 24, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

To address these limitations, this follow-up research seeks to develop an IoT-based monitoring system capable of tracking temperature and humidity, thereby complementing the previous study. This system is anticipated to deliver real-time data, facilitating the optimization of hydroponic management while establishing an integrated dataset to support artificial intelligence-driven analysis—an essential component as smart farming continues to evolve [9].

The Internet of Things (IoT) presents significant potential to enhance the efficiency and accuracy of monitoring and data collection processes. In this study, temperature and humidity sensors will be integrated with a microcontroller to capture and transmit data in real-time to a web application. The collected data will be systematically stored in a structured database, forming the foundation for a dataset to be further analyzed.

The development of this system not only addresses the demand for more comprehensive environmental monitoring but also represents a critical step toward building a data-driven smart hydroponic framework. This framework will possess the adaptability to respond to environmental changes, contributing to the advancement of intelligent and sustainable agricultural practices. Thus, this research brings new innovations in IoT-based hydroponic management, by expanding the scope of monitoring, developing a dataset-based system, and paving the way for the application of artificial intelligence in hydroponics. This research not only improves the efficiency and effectiveness of hydroponic farming, but also becomes a strategic step towards a more sustainable and technology-based smart agriculture.

II. RESEARCH METHOD

This research method consists of six phases: literature study, assembly of hardware, application development, implementation and integration, testing, and finally report writing, as shown in Figure 1.

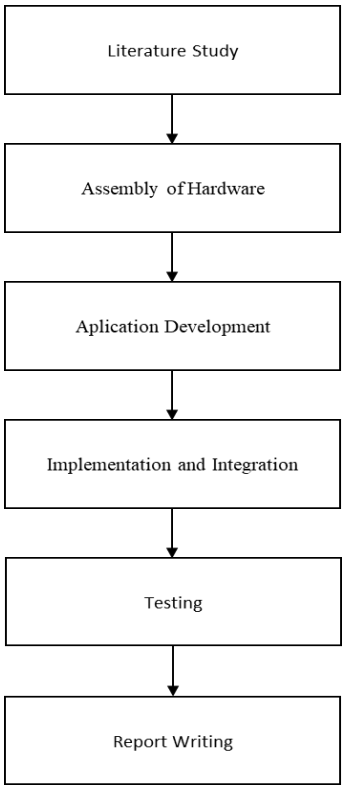


Fig. 1. Research method phases

A. *Literature Study*

Conducting a literature study to understand the basic principles of hydroponic farming, temperature and humidity measurement, Internet of Things (IoT), previous related studies, and application development.

B. *Assembly of Hardware*

The hardware used includes the DHT11 sensor (Figure 2) and the NodeMCU ESP8266 microcontroller (Figure 3). The DHT11 sensor is useful for measuring air temperature and humidity. DHT11 is chosen because it is inexpensive and sufficiently accurate, in addition to being readily available and widely used.

Fig. 2. DHT11 sensor

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2311, Copyright ©2025
Submitted : December 24, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Fig. 3. NodeMCU version 1 ESP8266

The NodeMCU ESP8266 microcontroller is used to connect to Wireless Fidelity (Wi-Fi), enabling internet connectivity as part of the Internet of Things (IoT). It communicates with the sensor and the web server. The NodeMCU ESP8266 is chosen because it is affordable, easily obtainable, and meets the needs of this research.

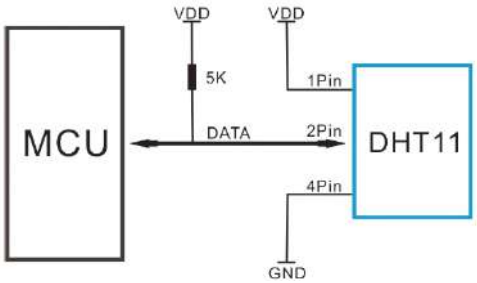


Fig. 4. Typical Application [10]

In Figure 4, the method of connecting the DHT11 sensor to the Micro-computer Unit (MCU) or single-chip computer is shown. The DHT11 sensor is connected to the NodeMCU, and a program is created to facilitate communication between the DHT11 sensor and the web server. The microcontroller collects data from the DHT11 sensor and sends it to the web server where the web application resides.

The program is developed using the Arduino IDE and written in the C++ programming language. It is subsequently embedded into the ESP8266 microcontroller, which then executes the instructions as specified within the program.

C. Aplication Development

The application development adheres to the Patas model, as illustrated in Figure 5. This model comprises six phases: user requirements, selection of devices, modification, evaluation, implementation, and maintenance. The Patas model is selected due to its suitability for this research, which involves a small team, demands rapid completion, and is conducted on a limited scale.

The developed application is a web-based system for monitoring temperature and humidity, serving as an extension of the previous research application that monitored pH and TDS levels. This application displays sensor data and humidity measurements collected from the DHT11 sensor, which are stored in a database. The data is presented in tabular form, while summarized information, including average values, is visualized on a dashboard.

The application is composed of two sides: the Front-End and Back-End. The Front-End is developed using HTML, CSS, and JavaScript, ensuring an interactive and user-friendly interface. The Back-End is built with PHP, managing the logic and data processing. MySQL is used as the database management system (DBMS), with Apache serving as the web server. The development process encompasses database design, user interface creation, coding, and the implementation of features to monitor and record temperature and humidity data.

The application is designed to be fully responsive, allowing seamless access through web browsers on various devices, including personal computers, laptops, smartphones, and tablets, thereby ensuring flexibility and ease of use across different platforms.

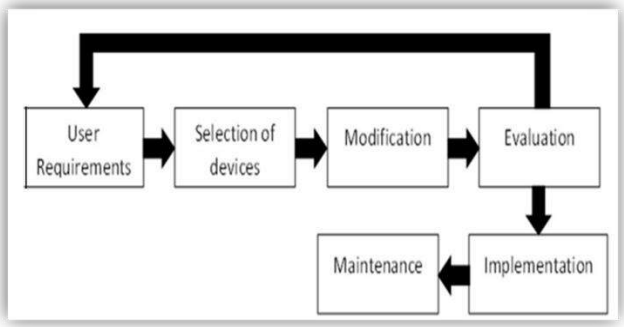


Fig. 5. Patas model [11]

The database design has been enhanced by incorporating temperature and humidity attributes into the existing schema from previous research. This expansion enables the recording of four essential parameters: pH, TDS, temperature, and humidity. In parallel, the user interface has been refined to accommodate the newly added temperature and humidity metrics, complemented by corresponding dashboard displays to ensure comprehensive data visualization.

D. Implementation and Integration

This phase involves the implementation and seamless integration of both hardware—comprising the DHT11 sensor and the NodeMCU ESP8266 microcontroller—and software, specifically the web-based monitoring application. The process adheres to the architectural framework depicted in Figure 6. The monitoring and data collection application based on web is deployed on an Apache web server and connected to a MySQL database. The NodeMCU ESP8266 is configured and establishes communication with the Apache web server over the Internet, ensuring continuous data transmission and synchronization between the hardware and software components.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2311, Copyright ©2025
Submitted : December 24, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Fig. 6. Architecture for data collection and monitoring of hydroponic cultivation

E. Testing

The monitoring application, integrated with sensors and the microcontroller, undergoes rigorous testing to ensure seamless hardware-software connectivity. This testing process verifies that measurements are consistently recorded in the database at regular intervals and that temperature and humidity data are accurately presented in both tabular and dashboard formats. If a problem occurs, resolve it immediately.

Figure 7 illustrates the testing environment, where the DHT11 sensor is positioned within a hydroponic container utilizing a wick system, while the NodeMCU ESP8266 is securely enclosed. The DHT11 sensor continuously captures ambient temperature and humidity levels at the hydroponic cultivation site. The NodeMCU ESP8266 transmits the collected data via Wi-Fi to an Apache web server, where it is systematically stored in a MySQL database. Users can monitor this data in real-time through a web application accessible on various devices, including personal computers, laptops, smartphones, and tablets.

To ensure the longevity and reliability of the system, the NodeMCU ESP8266 is enclosed in a protective casing, shielding it from water, dust, heat, and other environmental factors that could compromise its functionality and disrupt the temperature and humidity monitoring process.

Fig. 7. DHT11 sensor and ESP8266 Node MCU wrapped

F. Report Writing

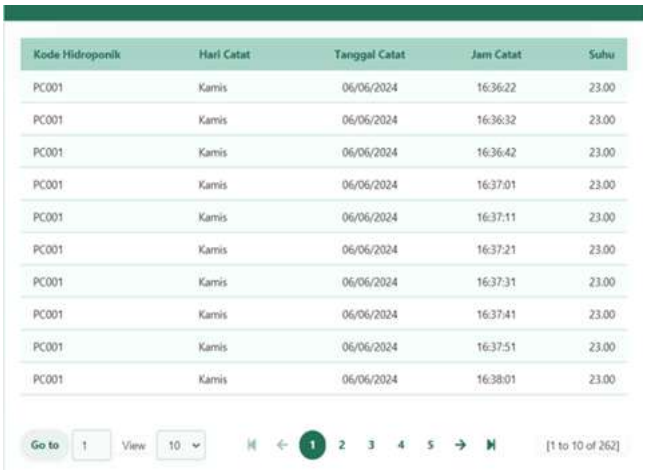
The research report is compiled to comprehensively document the steps undertaken, the findings obtained, and the recommendations for future development. This documentation serves as a valuable resource for understanding and replicating the research. The entire process and its outcomes are presented in the form of a scientific report and an article, which includes an introduction, research methodology, results and discussion, and conclusion.

III. RESULT AND DISCUSSION

The research results demonstrate that the IoT-based monitoring system developed can effectively monitor temperature and humidity parameters in real-time. The system utilizes a DHT11 sensor connected to the NodeMCU ESP8266, which continuously transmits data via Wi-Fi to a web server hosting the monitoring application. The collected data offers a comprehensive overview of the environmental conditions (temperature and humidity) surrounding the hydroponic cultivation.

The monitoring was conducted on June 6, 2024, from 16:36:22 to 17:55:40. During this study, temperature and humidity data were recorded every 10 seconds. The temperature recordings are depicted in Figure 8, while the humidity readings are shown in Figure 9. The average temperature and humidity during this period were 23.16°C and 66.67%, respectively. These average values are also displayed on the dashboard (Figure 10). This monitoring can be performed anytime and from any location, if the device used is connected to the Internet. Various devices, such as personal computers, laptops, smartphones, and tablets, can access the monitoring application via web browsers like Google Chrome and Microsoft Edge.

By utilizing the DHT11 sensor connected to the NodeMCU ESP8266, the IoT system can generate extensive data. For instance, if the vegetable harvest period spans 25 days and data is recorded every second, it would result in 2,160,000 data points for temperature and humidity. This system not only generates large volumes of data but also facilitates remote monitoring due to its Internet connectivity.

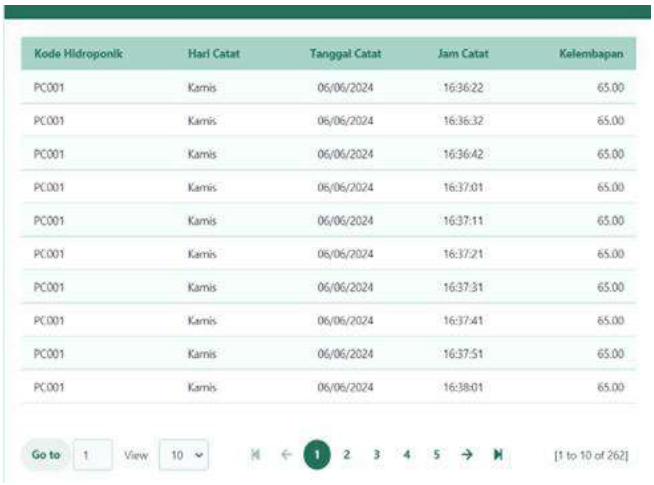


Kode Hidroponik	Hari Catat	Tanggal Catat	Jam Catat	Suhu
PC001	Kamis	06/06/2024	16:36:22	23.00
PC001	Kamis	06/06/2024	16:36:32	23.00
PC001	Kamis	06/06/2024	16:36:42	23.00
PC001	Kamis	06/06/2024	16:37:01	23.00
PC001	Kamis	06/06/2024	16:37:11	23.00
PC001	Kamis	06/06/2024	16:37:21	23.00
PC001	Kamis	06/06/2024	16:37:31	23.00
PC001	Kamis	06/06/2024	16:37:41	23.00
PC001	Kamis	06/06/2024	16:37:51	23.00
PC001	Kamis	06/06/2024	16:38:01	23.00

Go to 1 View 10 1 2 3 4 5 [1 to 10 of 262]

Fig. 8. Temperature Monitoring Results

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2311, Copyright ©2025
Submitted : December 24, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025



Kode Hidroponik	Hari Catat	Tanggal Catat	Jam Catat	Kelembapan
PC001	Kamis	06/06/2024	16:36:22	65.00
PC001	Kamis	06/06/2024	16:36:32	65.00
PC001	Kamis	06/06/2024	16:36:42	65.00
PC001	Kamis	06/06/2024	16:37:01	65.00
PC001	Kamis	06/06/2024	16:37:11	65.00
PC001	Kamis	06/06/2024	16:37:21	65.00
PC001	Kamis	06/06/2024	16:37:31	65.00
PC001	Kamis	06/06/2024	16:37:41	65.00
PC001	Kamis	06/06/2024	16:37:51	65.00
PC001	Kamis	06/06/2024	16:38:01	65.00

Go to 1 View 10 1 2 3 4 5 [1 to 10 of 262]

Fig. 9. Humidity Monitoring Results

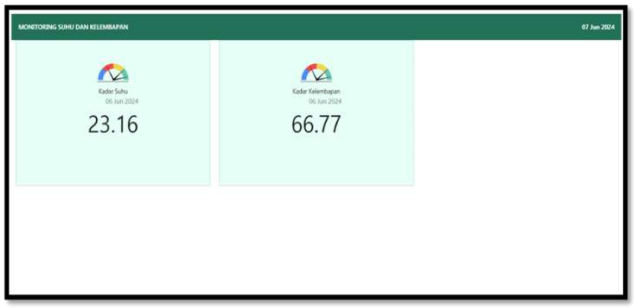


Fig. 10. Dashboard for Average Temperature and Humidity

The primary implications of this research include improved efficiency in hydroponic management through data-driven decision-making. The generated dataset not only aids in the optimization of the cultivation environment but also opens opportunities for the development of data-driven artificial intelligence models, such as neural networks, machine learning, and deep learning, which can enhance prediction, classification, and control mechanisms. In addition, the developed system has direct practical benefits for hydroponic farmers, such as real-time automatic monitoring (reducing the risk of environmental fluctuations that can affect plant growth), data-driven decision-making (allowing farmers to optimize the use of resources such as water and nutrients) and increased cultivation efficiency (ensuring consistent yields and high-quality products through precise environmental control).

Numerous previous researches related to the monitoring of temperature and humidity in various contexts have predominantly focused on non-artificial intelligence (AI) approaches [12], [13], [14]. However, some have begun to incorporate AI techniques, such as Fuzzy Logic [15], [16]. In general, prior research has remained partial and not data-oriented, meaning it has neither generated nor utilized datasets to facilitate effective and efficient smart hydroponic farming. Therefore, it is essential that, in addition to monitoring plant conditions, the system also records plant growth, creating datasets that can be employed for analysis, AI applications, or serve as a reference model for future cultivation aimed at achieving optimal harvest quality.

The research findings reveal that, within the study area, the temperature and humidity levels were approximately 23°C and 66%, respectively, in a hydroponic space of less than 1 m². For larger cultivation areas, additional DHT11 sensors would be required to ensure accurate readings across various sections. For more precise temperature and humidity measurements, the DHT11 sensor could be replaced with more advanced options such as the DHT22, SHT3x Series, or BME280 sensors. Furthermore, to accommodate broader coverage, microcontrollers like Arduino or mini-computers such as Raspberry Pi could be utilized. Additionally, careful attention must be given to power supply requirements to maintain a stable electricity flow and ensure a reliable, continuous Wi-Fi connection.

IV. CONCLUSION

This research has successfully developed a web-based and Internet of Things (IoT) real-time temperature and humidity monitoring and recording application in hydroponic cultivation. The integration of the DHT11 sensor and the NodeMCU ESP8266 microcontroller allows continuous data collection, resulting in a structured data set that supports monitoring the cultivation environment. System testing shows the efficiency and effectiveness of real-time monitoring, ensuring stable environmental conditions for hydroponic cultivation, and improving the accuracy of decision making in hydroponic management.

This research makes a significant contribution to data-driven hydroponic management and smart farming by addressing the existing research gap. Unlike previous researches that primarily focused on pH and Total Dissolved Solids (TDS) monitoring, this research extends the monitoring system by incorporating temperature and humidity tracking. This comprehensive dataset serves as a foundation for advanced Artificial Intelligence applications, enabling predictive modeling, automated environmental control, and optimized cultivation strategies.

Further more, it is essential to compare the performance of other sensors, integrate findings with previous research, and incorporate additional sensors, such as light and CO2 sensors, to broaden the scope of environmental data. Monitoring should also include visual data on plant growth to further enrich the dataset. Moreover, the development of artificial intelligence algorithms capable of providing recommendations, predictions, and classifications will become a strategic innovation, enhancing the competitiveness of IoT-based hydroponic technology in the future. Further hydroponic cultivation of plants, such as kale, celery, lettuce, and others, should be conducted to observe actual outcomes and refine dataset management practices.

This research lays the foundation for the future of smart hydroponic systems. IoT-based hydroponics will play a vital role in improving sustainability (food security), efficiency, and productivity in modern agriculture.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2311, Copyright ©2025
Submitted : December 24, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

ACKNOWLEDGMENT

We gratefully acknowledge the research funding and support provided by the University of Matana.

REFERENCES

[1] G. Rajaseger et al., "Hydroponics: current trends in sustainable crop production," Bioinformation, vol. 19, no. 9, pp. 925–938, Sep. 2023, doi: 10.6026/97320630019925.

[2] S. Barus, M. Tangdililing, P. A. Widjaya, S. Sujanto, and N. Jong, "PENERAPAN METODE HIDROPONIK PADA DESA BENTENG MAMULLU TANAH TORAJA UNTUK MENGATASI KEBUTUHAN SAYUR MAYUR", Universal Raharja Community (URNITY Journal), vol. 3, no. 1, pp. 31-36, Feb. 2023.

[3] P. Digital, "Kementan Dukung Ketahanan Pangan melalui Hidroponik," Tempo, Jun. 15, 2022. Available: https://bisnis.tempo.co/read/1602065/kementan-dukung-ketahanan-pangan-melalui-hidroponik.

[4] "Goal 2 | Department of Economic and Social Affairs." Available: https://sdgs.un.org/goals/goal2.

[5] U. Shareef, A. U. Rehman, and R. Ahmad, "A Systematic Literature Review on Parameters Optimization for Smart Hydroponic Systems," AI, vol. 5, no. 3, pp. 1517–1533, Aug. 2024, doi: 10.3390/ai5030073

[6] R. Raimarda, "Pengaruh Suhu dan Kelembapan pada Tumbuhan Halaman all - Kompas.com," KOMPAS.com, Oct. 09, 2020. Available: https://www.kompas.com/skola/read/2020/10/09/223258169/pengaruh-suhu-dan-kelembapan-pada-tumbuhan?page=all

[7] S. P. Barus and Y. A. Dhamma, "Development of Hydroponic Application based on Web and Internet of Things for The Community to Monitor pH and Total Dissolved Solids," International Journal of Research in Community Service, vol. 5, no. 3, pp. 130–137, Jul. 2024, doi: 10.46336/ijrcs.v5i3.708.

[8] D. Y. Setyawan, W. Warsito, R. Marjunus, and S. Sumaryo, "Automasi dan Internet of Things (IoT) pada Pertanian Cerdas: review artikel pada Jurnal Terakreditasi Kemenristek," Setyawan | Prosiding Semnastek, Jun. 26, 2024. https://jurnal.umj.ac.id/index.php/semnastek/article/view/22702/10492 .

[9] G. S. Patel, A. Rai, N. N. Das, and R. P. Singh, *Smart Agriculture: Emerging Pedagogies of Deep Learning, Machine Learning and Internet of Things*. CRC Press, 2021.

[10] alldatasheet.com, "DHT11 Download." Available: https://www.alldatasheet.com/datasheet-pdf/download/1440068/ETC/DHT11.html .

[11] S. P. Barus, "Implementation of the PATAS model in the development of the Matana University Graduation Information System," JISA(Jurnal Informatika Dan Sains), vol. 5, no. 2, pp. 116–119, Dec. 2022, doi: 10.31326/jisa.v5i2.1327.

[12] B. Haryanto, N. Ismail, and E. J. Pristianto, "Sistem Monitoring Suhu dan Kelembapan Secara Nirkabel pada Budidaya Tanaman Hidroponik," JTERA (Jurnal Teknologi Rekayasa), vol. 3, no. 1, p. 47, Jun. 2018, doi: 10.31544/jtera.v3.i1.2018.47-54.

[13] R. M. Abdurrohman, "Prototipe monitoring suhu dan kelembapan secara realtime," Journal ICTEE, vol. 4, no. 2, p. 29, Aug. 2023, doi: 10.33365/jictee.v4i2.3158.

[14] D. C. M. Wijaya, H. Khariono, M. R. Abrori, R. A. Fernanda, and H. A. Kusuma, "Sistem Pemantauan Suhu dan Kelembapan Udara Pada Tanaman Hias Janda Bolong Terintegrasi," Informatik Jurnal Ilmu Komputer, vol. 17, no. 3, p. 174, Jan. 2022, doi: 10.52958/iftk.v17i3.3436.

[15] S. K. Risandriya, "Pemantauan dan Pengendalian Kelembapan, Suhu, dan Intensitas Cahaya Tanaman Tomat dengan Logika Fuzzy Berbasis IoT," Journal of Applied Electrical Engineering, vol. 3, no. 1, pp. 9–14, Jun. 2019, doi: 10.30871/jaee.v3i1.1394.

[16] S. D. Putra, H. Heriansyah, E. F. Cahyadi, K. Anggriani, and M. H. I. S. Jaya, "Development of smart hydroponics system using AI-based sensing," JURNAL INFOTEL, vol. 16, no. 3, Aug. 2024, doi: 10.20895/infotel.v16i3.1190.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2311, Copyright ©2025
Submitted : December 24, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Implementation of Doodle Jump Game Based on Accelerometer Sensor and Kalman Filter

Edghar Danuwijaya[1], Yohanes Armenian Putra[2], Hilal Hudan Nuha[3]*, Ikke Dian Oktaviani[4]
School of Computing[1], [2], [3], [4]
Telkom University
Bandung, Indonesia
edghardanu@student.telkomuniversity.ac.id[1], yohanesar@student.telkomuniversity.ac.id[2],
hilal@sojc.telkomuniversity.ac.id[3],oktavianiid@telkomuniversity.ac.id[4]

***Abstract*— The doodle jump game is a video game with a jumping game model assisted by accelerometer sensor technology. Placing the accelerometer sensor in the doodle jump game is a very appropriate solution to determine the accuracy of the values on the sensor. The accelerometer sensor can be measured in real time, however applying a small force to the sensor can result in interference with measurement accuracy. Therefore, creating the measurement results you need using filters can help reduce noise. The method used to use this filter is the Kalman Filter algorithm. The use of the Kalman Filter method can provide a stable level of accuracy in the movements of the main characters in the game and the accelerometer sensor so that it can become a precise algorithm. Apart from that, the use of the Kalman Filter as a tool or method for measuring numbers to provide a solution to improve the design of the previous developer.**

Keywords*— *Doodle Jump, Accelerometer, Noise, Kalman Filter, Accuration.

I. INTRODUCTION

As technology begins to improve, so does the need for distance estimation; the development of this technology can optimize the performance of applications or systems. However, sensor readings can affect the accuracy of measurement results [1, 2, 3]. For example, accelerometer readings can reduce accuracy and are only useful in certain applications. It is therefore important to select the correct optimization technique to provide an accurate and useful value. The use of filters in the sensor will reduce the occurrence and increase the accuracy of the data [4, 5, 6].

Over the years there have been several technological developments including force plates, motion capture systems and accelerometers [7, 8]. These technologies have been used to assess vertical jump performance characteristics such as jump height, peak concentric force and impulse[9, 10]. However, they are not suitable for field evaluation due to their low power, high cost and the need for specialized computer equipment for data collection and analysis [11, 12].

Doodle Jump was created by Igor and Marko Pusenjak and published by the American studio Lima Sky. At the time of its launch, the Doodle Jump game was experiencing a surge, with more than 25,000 copies of the game being sold every day for a period of 4 months, and the game was released on various platforms [13, 14, 15]. Doodle Jump is a video game with a jumping game model using accelerometer technology. The placement of accelerometer sensors in the game is a very appropriate solution to know the accuracy of the value on the sensor [7, 8]. The sensor can be measured in real time with a small force applied, causing interference in the measurement. To obtain reliable measurement results, it is therefore necessary to use filters to reduce the noise [16, 17].

The Kalman filter can be used as a parameter estimation system with the aim of minimizing the error and noise in the estimation and, under certain conditions, minimizing the covariance error. Kalman filters have many applications in science and technology, such as tracking and monitoring vehicles, especially aircraft and spacecraft. It has evolved from optimal vehicle estimation to automation, positioning, target tracking, communications and signal processing, digital image processing, speech signal processing, earthquake prediction and many other fields. Therefore, the filter can be used as one of the tools that have general information, control and process automation [18].

From the above statement may require some effective, flexible and concrete solutions in its implementation, with this development of the doodle jump game can have a positive impact on these problems, especially the Kalman filter method that can minimize the game error by providing accurate values to correct the game error [19, 20]. The accelerometer sensor technology and the Kalman filter method in the Doodle Jump game can help to minimize errors and stabilize the character at the start of the game, making it a material for evaluating and analyzing problems that will be able to provide solutions to users[21, 22].

Fajar Irvansyah, Setiawansyah and Muhaqiqin, in his research entitled 'ANDROID-BASED HAIR SURGE SERVICE BOOKING APPLICATION' the development of the Android system has a very significant technological development so that it cannot be separated from the role of AOSP (Android Open-Source Project) which has the authority to develop the Android operating system and is directly managed by Google [23].

Android developer, in his research entitled 'Platform Architecture' Doodle Jump has a game with a type 2 game of jumping or moving places with the aim of getting as many points as possible. In the Doodle Jump game, the character itself

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2259, Copyright ©2025
Submitted : December 21, 2024, Revised : January 5, 2025 , Accepted : January 15, 2025, Published : January 31, 2025

has a role or character that almost resembles a cartoon alien and is given small clouds on the game screen that act as a background or wallpaper in the game [24].

R. I. Alfian, A. Ma'Arif and S. Sunardi, in a study entitled "Noise reduction in the accelerometer and gyroscope sensor with the Kalman filter algorithm" The Kalman filter process consists of two stages, namely the "predict" stage, which uses estimates from the previous time to predict current events, and the "update" stage, which is used to correct the first stage prediction and provide a more accurate time estimate. This algorithm aims to control systems that are susceptible to noise disturbances with the aim of reducing errors and monitoring time under predetermined conditions [25].

Based on the background, the problem formulation in this final project can be summarized as follows: how to implement an accelerometer sensor in the Doodle Jump game, how to stabilize the movement of the main character in the game, and what is the accuracy of the Kalman filter in enhancing the gameplay experience of Doodle Jump.

The objectives of this study aim to address the formulated problems as follows: to analyze the motion stability of the main character in the Doodle Jump game, to stabilize the character's movement within the game, and to determine the accuracy level of the Kalman filter in improving the gameplay experience.

The topics and limitations of this research are as follows: this study will focus on analyzing the motion stability of the main character in the Doodle Jump game, with the analysis being conducted by the author. Additionally, the research will concentrate on the analysis process using the Kalman filter method.

II. RESEARCH METHOD

A. Flowchart System

The system that has been built uses a flowchart to assist researchers in their research, as described in Fig 1.

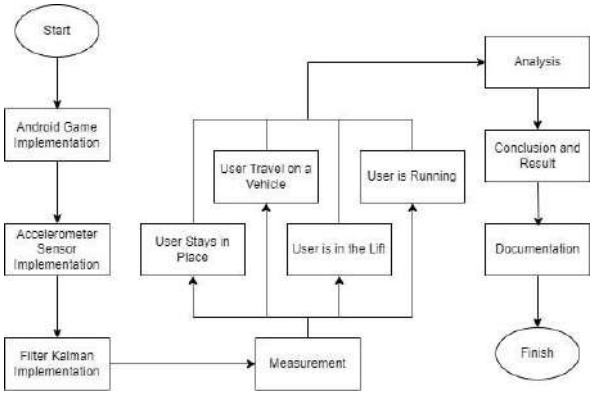


Fig. 1. Flowchart Diagram

Based on the Fig 1, it can be concluded that the flow of game development by researchers starts from game implementation using Android Java language, accelerometer sensor implementation, Kalman filter implementation, measurement, to case studies conducted daily by researchers on the application of doodle jump games. The implementation carried out by the researchers can produce accurate data as a comparison of accuracy with previous algorithms, therefore accuracy comparison is the final part of the research.

1) Implementation of the game on Android

The game observed in the research is a doodle jump game based on the Java programming language and supported by the Android Studio IDE software.

2) Accelerometer Sensor Implementation

The sensor is inbuilt by Android in the previous system found in the Doodle Jump game that has been made. The accelerometer sensor functions as providing movement to the game character. In addition, the movement of the game character will produce an output value that can be known and analysed by researchers who focus on the level of accuracy of the game character.

3) Filter Kalman Implementation

The mathematical calculation of the Kalman filter will be implemented in the Doodle Jump game by integrating it using the Java programming language. The output generated by the accelerometer sensor is used as input for the calculation of the Kalman filter method.

4) Measurement

Consists of five parameters that are used as measurements in the Doodle Jump game. In addition, there is an accelerometer sensor to provide movement to the appropriate character as directed by the user so that it will produce a user input value.

5) Case Study

The main case study as a variable control in the measurement of parameters consisting of four, namely:

- Neutral/no interference, the user will stay in place without moving while operating the game.
- Interference from the acceleration of the environment around the user, the user will ride in a vehicle as a passenger while operating the game.
- Gravity interference, the user will be in a lift while using the game.
- User acceleration interference, the user will be walking while using the game.

6) Analysis

This analysis presents the test results and the comparison with and without the Kalman filter implementation in the game.

7) Conclusion and Result

Conclusions can be drawn from the results of the analyses carried out.

8) Documentation

Report the results of the research carried out by the researcher to analyze the results of the entire case study, so that it can be used as a reference material to solve the solution to the game previously made and know the implementation.

B. Research Diagram Illustration

System research on Doodle Jump games, which can be

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2259, Copyright ©2025
Submitted : December 21, 2024, Revised : January 5, 2025 , Accepted : January 15, 2025, Published : January 31, 2025

interpreted in the following block diagram.

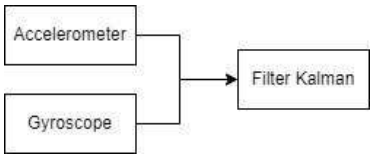


Fig. 2. Doodle Jump Diagram

Fig 2 shows a smartphone with an accelerometer sensor to control the character in the game Doodle Jump. In addition, the sensor can be played by tilting it to the right and left or called a gyroscope. In addition, the use of Kalman filter because of comparison numbers from the accelerometer sensor so that it can be analyzed in the analysis based on the test results.

C. Use case Diagram

The diagram below provides an overview of the program flow as interpreted by the researchers.

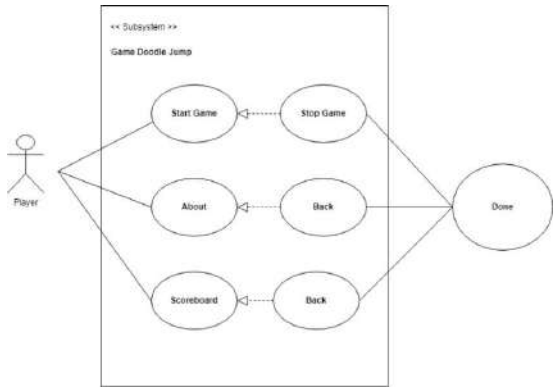


Fig. 3. Use case Diagram

Fig 3 illustrates the interaction between players (actors) and various features in the game. A detailed explanation of the figure follows.

- Player: The player is the main actor interacting with the system. The player can choose different options available in the game.
- Start game: The player can start the game by selecting the 'Start Game' option.
- Stop Game: Once the game has started, the player can stop the game by selecting 'Stop Game'.
- About: The player can select the 'About' option to get more information about the game.
- Scoreboard: Players can view the highest score by selecting the 'Scoreboard' option.
- Back: From the 'About' and 'Scoreboard' menus, players can return to the main menu by selecting the 'Back' option.
- Done: All player activity, whether starting the game, viewing About the Game or viewing the Scoreboard, ends at a point called 'Done'.

III. RESULT AND DISCUSSION

A. Research Diagram Illustration

1) Accelerometer Sensor Result

a) User Stays in Place

The user stays in place while operating the game and requires 10 trials to generate a number as a reference for comparing accuracy without using the Kalman filter. The detection results in the table below are the tilt results obtained from the user remaining in place under seated conditions. As shown in table 1.

TABLE I. TESTING WHEN THE USER IS STATIONARY USING THE ACCELEROMETER SENSOR

Test Scenario	Tilt	Detection Result	Tilt	Detection Result
Test 1	Right	39,89	Left	-31,09
Test 2	Right	45,42	Left	-42,95
Test 3	Right	42,51	Left	-39,22
Test 4	Right	30,19	Left	-43,05
Test 5	Right	30,35	Left	-44,15
Test 6	Right	39,50	Left	-38,52
Test 7	Right	30,97	Left	-36,06
Test 8	Right	41,39	Left	-41,74
Test 9	Right	32,08	Left	-43,36
Test 10	Right	31,44	Left	-39,22

b) User Travel on Vehicle

The user rides on a motorized vehicle as a passenger during operation and requires up to 10 tests to produce figures as a reference for comparing accuracy without using the Kalman filter. The detection results in this table below are the slope results obtained from the user riding on a motorized vehicle at a speed of 40km/h. As shown in the following table. As shown in the following table 2.

TABLE II. TESTING WHEN THE USER IS RIDING IN A VEHICLE USING THE ACCELEROMETER SENSOR

Test Scenario	Tilt	Detection Result	Tilt	Detection Result
Test 1	Right	58,10	Left	-56,65
Test 2	Right	62,42	Left	-53,81
Test 3	Right	57,55	Left	-52,96
Test 4	Right	57,37	Left	-58,51
Test 5	Right	64,16	Left	-50,10
Test 6	Right	60,40	Left	-60,50
Test 7	Right	52,94	Left	-61,29
Test 8	Right	51,95	Left	-50,54
Test 9	Right	49,96	Left	-61,58
Test 10	Right	56,38	Left	-56,66

c) User is in the lift

The user is in the lift while playing the game and needs 10

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2259, Copyright ©2025
Submitted : December 21, 2024, Revised : January 5, 2025 , Accepted : January 15, 2025, Published : January 31, 2025

tests to produce numbers as a reference for comparing accuracy without using the Kalman filter. The detection results in this table below are the tilt results obtained from the user in the lift. As shown in the following table 3.

TABLE III. TESTING WHEN THE USER IS IN THE LIFT USING THE ACCELEROMETER SENSOR

Test Scenario	Tilt	Detection Result	Tilt	Detection Result
Test 1	Right	30,97	Left	-36,22
Test 2	Right	41,95	Left	-34,95
Test 3	Right	42,30	Left	-40,28
Test 4	Right	45,67	Left	-50,10
Test 5	Right	33,22	Left	-43,86
Test 6	Right	39,27	Left	-30,21
Test 7	Right	45,20	Left	-44,23
Test 8	Right	30,03	Left	-38,34
Test 9	Right	30,23	Left	-41,01
Test 10	Right	30,24	Left	-43,25

d) User is Running

The user walks while the game is running and needs 10 trials to produce numbers as a reference for comparing accuracy without using the Kalman filter. The detection results in this table below are the tilt results obtained from the running user. As shown in the following table 4.

TABLE IV. TESTING WHEN THE USER IS RUNNING USING THE ACCELEROMETER SENSOR

Test Scenario	Tilt	Detection Result	Tilt	Detection Result
Test 1	Right	87,31	Left	-89,49
Test 2	Right	70,01	Left	-70,56
Test 3	Right	74,30	Left	-76,50
Test 4	Right	73,56	Left	-89,88
Test 5	Right	71,70	Left	-86,94
Test 6	Right	89,10	Left	-78,90
Test 7	Right	73,56	Left	-80,56
Test 8	Right	80,20	Left	-89,85
Test 9	Right	83,95	Left	-74,67
Test 10	Right	89,77	Left	-75,26

2) Filter Kalman Testing

a) User Stays in Place

The user remains in place while operating the game and requires 10 trials to produce numbers as a reference for comparing accuracy using the Kalman filter. The detection results in the table below are the tilt results obtained from the user remaining in place under seated conditions. As shown in table 5.

TABLE V. TESTING WHEN THE USER IS STATIONARY USING A KALMAN FILTER

Test Scenario	Tilt	Detection Result	Tilt	Detection Result
Test 1	Right	30,84	Left	-26,81
Test 2	Right	27,50	Left	-24,13
Test 3	Right	24,81	Left	-28,85
Test 4	Right	30,01	Left	-22,39
Test 5	Right	29,21	Left	-27,43
Test 6	Right	30,40	Left	-27,75
Test 7	Right	31,02	Left	-27,25
Test 8	Right	31,09	Left	-25,50
Test 9	Right	22,75	Left	-26,31
Test 10	Right	24,61	Left	-26,77

b) User Travel on Vehicle

The user rides on a motorized vehicle as a passenger during operation and requires 10 tests to produce figures as a reference for comparing accuracy using the Kalman filter. The detection results in this table below are the tilt results obtained from the user riding on a motorized vehicle at a speed of 40km/h. As shown in the following table. As shown in table 6.

TABLE VI. TESTING WHEN THE USER IS RIDING IN A VEHICLE USING A KALMAN FILTER

Test Scenario	Tilt	Detection Result	Tilt	Detection Result
Test 1	Right	27,40	Left	-34,83
Test 2	Right	28,52	Left	-29,85
Test 3	Right	30,45	Left	-31,30
Test 4	Right	33,18	Left	-39,45
Test 5	Right	28,44	Left	-36,98
Test 6	Right	32,92	Left	-20,42
Test 7	Right	28,31	Left	-28,05
Test 8	Right	33,45	Left	-34,78
Test 9	Right	37,40	Left	-32,93
Test 10	Right	35,76	Left	-33,62

c) User is in the lift

The user is in the lift while playing the game and needs 10 tests to produce numbers as a reference for the accuracy comparison using the Kalman filter. The detection results in this table below are the tilt results obtained from the user in the lift. As shown in table 7.

TABLE VII. TESTING WHEN THE USER IS IN THE LIFT USING A KALMAN FILTER

Test Scenario	Tilt	Detection Result	Tilt	Detection Result
Test 1	Right	35,96	Left	-30,40
Test 2	Right	38,94	Left	-45,58
Test 3	Right	37,38	Left	-47,69
Test 4	Right	35,86	Left	-48,57
Test 5	Right	36,84	Left	-44,75
Test 6	Right	38,15	Left	-45,81
Test 7	Right	35,60	Left	-47,30

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2259, Copyright ©2025
Submitted : December 21, 2024, Revised : January 5, 2025 , Accepted : January 15, 2025, Published : January 31, 2025

Test 8	Right	37,55	Left	-43,13
Test 9	Right	34,25	Left	-48,30
Test 10	Right	34,39	Left	-44,61

d) User is Running

The user runs while playing the game and needs 10 tests to produce numbers as a reference for the accuracy comparison using the Kalman filter. The detection results in this table below are the slope results obtained from the running user. As shown in table 8.

TABLE VIII. TESTING WHEN THE USER IS RUNNING USING A KALMAN FILTER

Test Scenario	Tilt	Detection Result	Tilt	Detection Result
Test 1	Right	34,15	Left	-39,53
Test 2	Right	40,72	Left	-38,52
Test 3	Right	37,00	Left	-19,25
Test 4	Right	32,53	Left	-42,30
Test 5	Right	30,83	Left	-39,68
Test 6	Right	29,81	Left	-38,49
Test 7	Right	32,79	Left	-22,94
Test 8	Right	24,83	Left	-38,03
Test 9	Right	34,43	Left	-39,76
Test 10	Right	34,06	Left	-43,67

B. Discussion

In the accelerometer sensor and Kalman filter method, it can be seen the difference in the value of the accuracy results obtained by researchers, as described as follows.

- The use of the performance of the accelerometer sensor and the application of Kalman filter in the condition of the user staying in place has an accuracy difference of 20.70% for the right tilt detection results and 33.25% for the left tilt detection results.
- Using the performance of the accelerometer sensor and applying the Kalman filter in the condition of the user riding on a motorized vehicle has an accuracy difference of 14.29% for the right tilt detection results and 24.47% for the left tilt detection results.
- Using the performance of the accelerometer sensor and applying the Kalman filter in the state of the user in the lift has an accuracy difference of 5.27% for the right tilt detection results and 5.83% for the left tilt detection results.
- The use of accelerometer sensor performance and the application of Kalman filter in the walking user condition has an accuracy difference of 46.23% for right tilt detection results and 48.89% for left tilt detection results.

C. Doodle Game Simulation

Then, for the Doodle Game simulation built in this study, it can be seen in the following fig 4

Fig. 4. Doodle Game Simulation

IV. CONCLUSION

This research has solved the problem of the accuracy of the accelerometer sensor data by implementing it into the Kalman filter method. This research has obtained information on accuracy problems based on initial testing of the accelerometer sensor. In making the table can obtain information on accuracy data that has been compared by researchers, the information compared is accelerometer sensor data and Kalman filter. This research has successfully influenced the accuracy of the accelerometer sensor, as an example of using the performance of the accelerometer sensor and the application of Kalman filter has an accuracy difference of 20.70% for the right tilt detection results and 33.25% for the left tilt detection results. This indicates that the application of the Kalman filter to the Doodle Jump game can have a significant effect.

ACKNOWLEDGMENT

The financial support of the Indonesia's DRTPM, DITJEN DIKTIRISTEK, KEMDIKBUDRISTEK through grant 106/E5/PG.02.00.PL/2024, 043/SP2H /RT-MONO/LL4/2024, and 077/LIT07/PPM-LIT/2024 is hereby acknowledged and appreciated.

REFERENCES

[1] Ben, Lonnqvist., Zhengqing, Wu., Michael, H., Herzog. (2023). 1. Latent Noise Segmentation: How Neural Noise Leads to the Emergence of Segmentation and Grouping. arXiv.org, doi: 10.48550/arxiv.2309.16515
[2] Ahmet, Kağan, Kaya. (2023). 2. Noise-Robust Diffusion Based Semantic Segmentation. doi: 10.1109/siu59756.2023.10223870
[3] Noisy image segmentation model combining PDE and improved variational level set method. Academic Journal of Mathematical Sciences, doi: 10.25236/ajms.2022.030106
[4] Yan, Fang., Feng, Zhu., Bowen, Cheng., Luoqi, Liu., Yao, Zhao., Yunchao, Wei. (2023). 4. Locating Noise is Halfway Denoising for Semi-Supervised Segmentation. doi: 10.1109/iccv51070.2023.01523
[5] Eden, Grad., Moshe, Kimhi., Lion, Halika., Chaim, Baskin. (2024). 5. Benchmarking Label Noise in Instance Segmentation: Spatial Noise Matters. doi: 10.48550/arxiv.2406.10891
[6] A. Abadleh, B. M. Al-Mahadeen, R. M. AlNaimat, and O. Lasassmeh, "Noise segmentation for step detection and distance estimation using smartphone sensor data," Wireless Networks, vol. 27, no. 4, pp. 2337–2346, May 2021, doi: 10.1007/s11276-021-02588-0.
[7] Damjana, Cabarkapa., Dimitrije, Cabarkapa., Nicolas, M., Philipp., Andrew, C., Fry. (2023). 1. Impact of the Anatomical Accelerometer Placement on Vertical Jump Performance Characteristics. Sports, doi: 10.3390/sports11040092
[8] Dimitrije, Cabarkapa., Andrew, C., Fry., Matthew, J., Hermes. (2021). 2. Accuracy of an Experimental Accelerometer for Assessing Countermovement Vertical Jump Height. doi: 10.18060/24831

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2259, Copyright ©2025
Submitted : December 21, 2024, Revised : January 5, 2025 , Accepted : January 15, 2025, Published : January 31, 2025

[9] Dennis, Moschina., Tobias, Röddiger., Michael, Beigl. (2023). 3. Vertical Jump Test Using an Earable Accelerometer. doi: 10.1145/3594739.3610669
[10] M.A., Choukou., Guillaume, Laffaye., Redha, Taiar. (2014). 4. Reliability and validity of an accelerometric system for assessing vertical jumping performance. Biology of Sport, doi: 10.5604/20831862.1086733
[11] Giacomo, Villa., Serena, Cerfoglio., M., Galli., Veronica, Cimolin. (2024). 5. Impact of IMU Placement on Vertical Jump Height Estimation: A Comparative Analysis in Countermovement Jumps. The Star, doi: 10.1109/star62027.2024.10635965
[12] D. V. Cabarkapa, D. Cabarkapa, N. M. Philipp, and A. C. Fry, "Impact of the Anatomical Accelerometer Placement on Vertical Jump Performance Characteristics," Sports, vol. 11, no. 4, Apr. 2023, doi: 10.3390/sports11040092.
[13] Anita, Teli., Sheetal, Harkuni., Deepti, M., Kadeangadi., Madhumati, J, Patil. (2023). 2. An assessment of the effectiveness and perception of the doodle – Video reinforced teaching method. BLDE university journal of health sciences, doi: 10.4103/bjhs.bjhs_207_22
[14] Jaimie, Carlson., Jason, Friedman., Christopher, Kim., Cynthia, Sung. (2020). 3. REBOund: Untethered Origami Jumping Robot with Controllable Jump Height. doi: 10.1109/ICRA40945.2020.9196534
[15] Dyaneshwar, Bavakar., Ramesh, Shahabade., Vishal, Vilas, Shinde., Bhargav, Vivek, Modak., Manas, Narendra, Telavane., Shantanu, Parameswara. (2024). 5. A Proposed Approach for AI Doodle Generation with a Hybrid Intelligent Agent. International Journal For Multidisciplinary Research, doi: 10.36948/ijfmr.2024.v06i02.17871
[16] Amulya Raveendra Katti; Anamay Sarkar; Pranav Mallikarjuna Swamy; Riya Kothari; Shenoy Pratik Gurudat, "Reinforcement Learning and Adversarial Attacks on Player Model with Doodle Jump."
[17] S. M. Irsyad, A. Basuki, and B. S. B. Dewantara, "Rancang Bangun AirMouse Menggunakan Sarung Tangan Bersensor Berbasis ESP32," Jurnal Rekayasa Elektrika, vol. 18, no. 3, Sep. 2022, doi: 10.17529/jre.v18i3.25816.
[18] M. Khodarahmi and V. Maihami, "A Review on Kalman Filter Models," Archives of Computational Methods in Engineering, vol. 30, no. 1. Springer Science and Business Media B.V., pp. 727–747, Jan. 01, 2023. doi: 10.1007/s11831-022-09815-7.
[19] Manuel, Baltieri., Takuya, Isomura. (2021). 3. Kalman filters as the steady-state solution of gradient descent on variational free energy. arXiv: Neurons and Cognition,
[20] Alejandro, Juárez-Lora., Luis, M., García-Sebastián., V., Ponce., E., Rubio-Espino., Herón, Molina-Lozano., Juan, Humberto, Sossa, Azuela. (2022). 2. Implementation of Kalman Filtering with Spiking Neural Networks. Sensors, doi: 10.3390/s22228845
[21] Thomas, James. (2019). 5. The Kalman Filter. doi: 10.1007/978-3-030-31934-2_5
[22] Lyakhov, P., Kalita, D., Bergerman, M. (2022). Hardware Implementation of the Kalman Filter for Video Signal Processing. In: Tchernykh, A., Alikhanov, A., Babenko, M., Samoylenko, I. (eds) Mathematics and its Applications in New Computer Systems. MANCS 2021. Lecture Notes in Networks and Systems, vol 424. Springer, Cham. https://doi.org/10.1007/978-3-030-97020-8_12
[23] Fajar Irvansyah; Setiawansyah; Muhaqiqin, "APLIKASI PEMESANAN JASA CUKUR RAMBUT BERBASIS ANDROID," Bandarlampung, 1, 2020.
[24] Developer Android, "Arsitektur Platform," developer.android.com.
[25] R. I. Alfian, A. Ma'Arif, and S. Sunardi, "Noise reduction in the accelerometer and gyroscope sensor with the Kalman filter algorithm," Journal of Robotics and Control (JRC), vol. 2, no. 3, pp. 180–189, May 2021, doi: 10.18196/jrc.2375

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2259, Copyright ©2025
Submitted : December 21, 2024, Revised : January 5, 2025 , Accepted : January 15, 2025, Published : January 31, 2025

Weather Monitoring and Classification Tools Using Fuzzy Logic Method Based on Internet of Things for Agriculture

Muhammad Alfazri Avindra [1], Ahmad Taqwa [2], Suroso [3]
Department of Electrical Engineering, Study Program of Telecommuncation Engineering[1], [2], [3]
Sriwijaya State Polytechnic
Palembang, Indonesia
alfazriimuhammadd@gmail.com [1], taqwa@polsri.ac.id [2], osorus11@gmail.com [3]

Agriculture is a crucial sector that faces significant challenges due to climate change, such as altered rainfall patterns, increased temperatures, and extreme weather, which threaten productivity. This research aims to design and develop a weather monitoring and classification tool utilizing the Fuzzy Logic method based on the Internet of Things (IoT). The system integrates sensors for temperature, humidity, air pressure, wind speed, and rainfall, processing the data through Fuzzy Logic to classify weather into sunny, cloudy, light rain, moderate rain, and heavy rain. The results are accessible in real-time via an LCD screen and the Blynk application. Achieving 100% accuracy in monitored conditions, the system operates continuously for 24 hours, powered by solar energy. Unlike traditional systems, this tool provides innovative solutions to weather-related agricultural challenges, enabling farmers to make informed decisions and mitigate losses due to extreme weather. This research contributes to advancing IoT applications in precision agriculture, providing a reliable, sustainable, and scalable solution to weather monitoring.

Keywords—Fuzzy Logic, Internet Of Things, Precision Agriculture, Solar Energy, Weather Monitoring

I. INTRODUCTION

The agricultural sector faces growing challenges due to climate change, as evidenced by shifting rainfall patterns, rising temperatures, and extreme weather events. These changes significantly threaten agricultural productivity, requiring innovative approaches to enhance resilience and sustainability in farming. Integrating the Internet of Things (IoT) into agriculture has proven to be a crucial strategy in tackling these issues. IoT applications enable real-time data collection and analysis, equipping farmers with the tools to make informed decisions and mitigate the impacts of climate variability on crop production[1], [2], [3].

Recent research emphasizes the transformative role of IoT in agricultural practices, especially in weather monitoring and management. For example, smart weather stations equipped with sensors allow continuous tracking of environmental factors like temperature, humidity, and rainfall [4], [5]. These systems not only deliver timely data but also enhance the precision of weather forecasts, which is critical for effective farm management. Additionally, the use of fuzzy logic to process this data allows for categorizing weather conditions, thereby enabling farmers to make proactive and well-informed decisions [6].

Despite advancements in IoT-based weather monitoring, there is a gap in integrating fuzzy logic methods to enhance the interpretation of collected data. Conventional systems often rely on binary classifications, which may fail to capture the complexity of weather patterns. This research aims to address this limitation by developing a fuzzy logic-based classification system that divides weather into five categories: sunny, cloudy, light rain, moderate rain, and heavy rain. This approach not only improves weather prediction accuracy but also equips farmers with actionable insights tailored to specific weather conditions[7].

This study's uniqueness lies in its focus on both technological and practical aspects of agricultural weather monitoring. By using solar energy to power the system, the proposed tool ensures sustainable energy use and continuous operation. Real-time weather data is made accessible via an LCD screen and the Blynk application [8]. The integration of renewable energy with IoT technology marks a significant advancement in precision agriculture, offering a scalable solution adaptable to diverse farming environments [9].

Furthermore, the study contributes to IoT applications in agriculture by demonstrating the effectiveness of combining fuzzy logic with real-time monitoring. The findings aim to provide farmers with valuable insights to reduce losses caused by extreme weather events. By addressing the need for more advanced data interpretation methods, this research seeks to strengthen agricultural resilience in the context of climate change [10].

In summary, the proposed weather monitoring and classification tool represents a major advancement in applying IoT technologies to agriculture. By leveraging fuzzy logic for data classification and incorporating renewable energy, this research addresses immediate agricultural challenges while promoting sustainable sector development. The subsequent sections of this paper will detail the methodology, results, and implications of these findings for the future of agricultural practices.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2330, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

II. METHODOLOGY

This research consists of several stages which will produce a system that can work properly. The overall research framework process is :

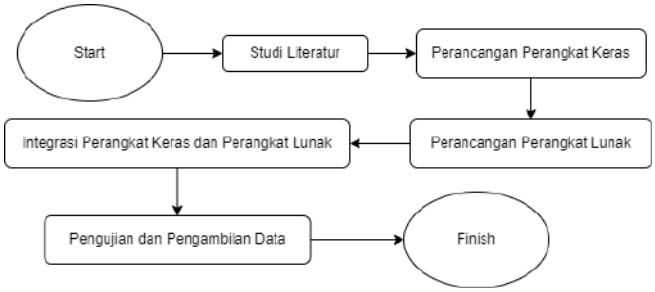


Fig. 1. Research Farmework

A. Literature Study

This study draws from diverse literature to establish the relevance of IoT and fuzzy logic in modern agriculture. Previous works have demonstrated the potential of IoT in real-time environmental monitoring, while fuzzy logic has proven effective in managing uncertainty in agricultural systems. However, few studies have explored the combined application of these technologies in weather classification for agricultural purposes.

B. Hardware Design

Hardware design is the design of the device to be made, the components used must be considered to avoid damage during system testing. These components should have characteristics that are in accordance with the needs of making tools. The hardware design begins with the block diagram in Figure

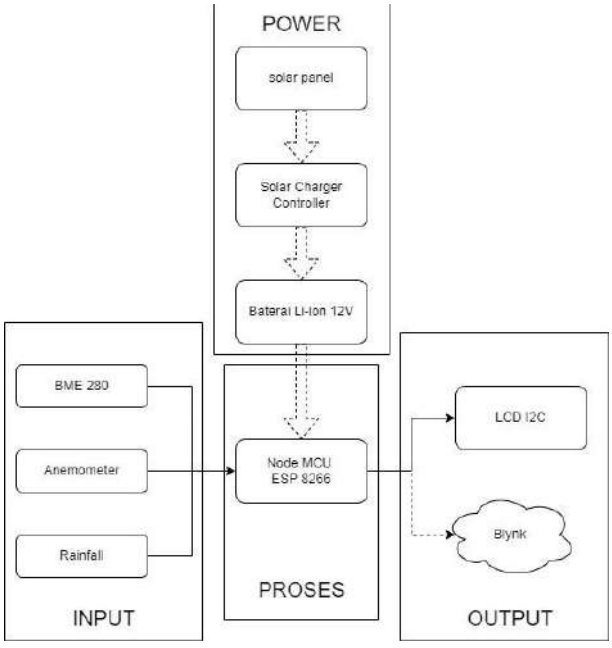


Fig. 2. Hardware Block Diagram

The following is an explanation of the block diagram above:

Solar panels are utilized to harness solar energy and convert it into electricity. The energy generated is regulated and stored in a 12V lithium-ion battery via a solar charge controller, ensuring safe and efficient charging. This battery serves as a stable power source for the system during periods without sunlight, such as nighttime or cloudy conditions. The ESP8266 module, a microcontroller with Wi-Fi capabilities, processes data from various sensors and sends it to the Blynk application for remote monitoring. Additionally, the ESP8266 powers an I2C LCD screen to display weather data locally. Sensors integrated into the system include the BME280 for measuring temperature, humidity, and air pressure, an anemometer for wind speed, and a rainfall sensor for precipitation levels.

C. Software Design

The software for this system is developed using the Arduino IDE to enable monitoring of weather parameters and facilitate weather classification. The Blynk application plays a crucial role by providing a user-friendly interface to display the results of real-time weather monitoring and classification. This combination of hardware and software ensures accurate data collection, processing, and convenient access to weather information both locally and remotely.

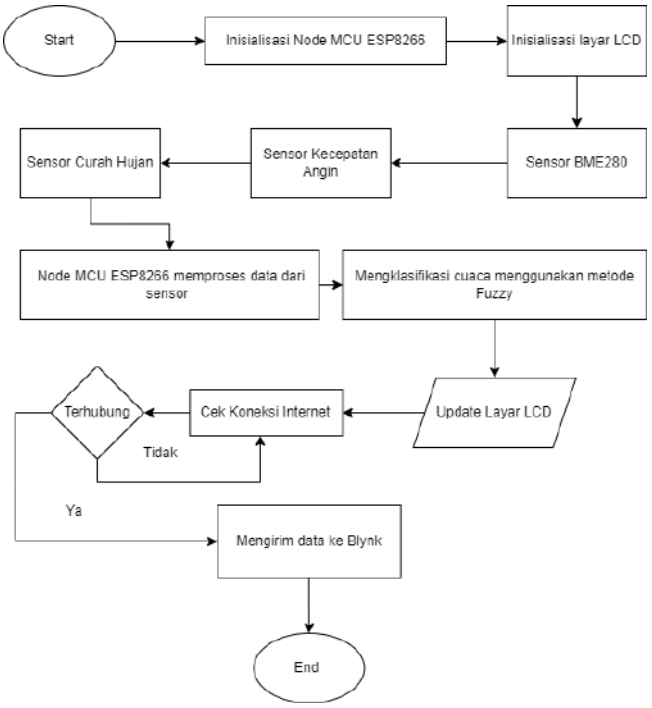


Fig. 3. Software Flowchart

The process begins with system initialization. The Node MCU ESP8266, serving as the microcontroller for data processing and communication between sensor devices, is initialized. Subsequently, the system initializes the LCD screen, which displays weather data. If the LCD screen is not connected, the initialization process will repeat.

Next, the BME280 sensor is activated to measure environmental parameters such as air pressure, temperature, and humidity, which are essential for weather classification.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2330, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Additionally, a wind speed sensor is used to measure wind velocity, providing critical data for more accurate weather classification, particularly for analyzing wind conditions. The rainfall sensor is also checked for its connection, as it measures rainfall intensity. If the sensor is not connected, the system will repeatedly check the connection.

Once data from the BME280 and wind speed sensors are obtained, the Node MCU ESP8266 processes the data. The processed data is then analyzed using the Fuzzy Logic method to classify weather conditions. This method is well-suited for situations involving uncertainty, such as weather classification. The classification results are displayed on the LCD screen, allowing users to view the weather conditions directly.

The system then checks the internet connection of the Node MCU ESP8266. If there is no connection, the process of checking the internet connection will repeat. If connected, the system sends the weather classification data to the Blynk platform. Blynk enables users to monitor weather conditions remotely via a smartphone application or web interface. Finally, the process concludes.

D. Method Development

Fuzzy logic is one of the artificial intelligence methods that can handle uncertain or ambiguous data. In weather measurement systems, fuzzy logic can be used to process sensor data that has high variability such as temperature, humidity, air pressure, wind speed, and rainfall, so as to produce information that is easier to understand such as weather classification which is divided into 5 namely light rain, medium rain, heavy rain, cloudy, and sunny.

1) Design System Fuzzy Logic

TABLE i INPUT AND OUTPUT CLASSIFICATION

Input	**Variabel**	**Output**
Temperature (Suhu)	Dingin, Normal, Panas	Hujan Ringan, Hujan Sedang, Hujan Deras, Mendung, dan Cerah
Humidity (Kelembapan)	Rendah, Sedang, Tinggi	
Pressure (Tekanan)	Rendah, Sedang, Tinggi	
WindSpeed (kecepatan angin)	Pelan, Sedang, Cepat	
Rainfall (Curah hujan)	Tidak Ada, Sedikit, Sedang, Banyak	

2) Sensor variable membership division

TABLE ii TEMPERATURE MEMBERSHIP FUNCTION

Variabel Suhu	**Membership Function**
Dingin	0 – 20° C
Normal	20 – 30° C
Panas	30 – 40° C

TABLE iii HUMIDITY MEMBERSHIP FUNCTION

Variabel Kelembapan	**Membership Function**
Rendah	0 – 40 %
Sedang	40 – 70 %
Tinggi	70 – 100 %

TABLE iv PRESSURE MEMBERSHIP FUNCTION

Variabel Tekanan Udara	**Membership Function**
Rendah	900 – 1010 hPa
Sedang	1010 – 1030 hPa
Tinggi	1030 – 1100 hPa

TABLE v WINSPEED MEMBERSHIP FUNCTION

Variabel Kecepatan Angin	**Membership Function**
Pelan	0 – 5 m/s
Sedang	5 – 8 m/s
Kencang	8 – 30 m/s

TABLE vi RAINFALL MEMBERSHIP FUNCTION

Variabel Curah Hujan	**Memberhsip Function**
Tidak ada	0 mm
Sedikit	0 – 20 mm
Sedang	20 – 50 mm
Banyak	50 – 100 mm

3) Weather classification using fuzzy rules

Fuzzy rules specify how input variables are related to output variables. These rules are expressed in the form “If ... then ...”. Here are some of the rules used:

a) Sunny weather output is determined if Temperature is normal and hot, Humidity is low and moderate, Air Pressure is High, Wind Speed is Slow, and Precipitation is Absent.

b) Cloudy weather output is determined if Temperature is cold, Humidity is high, Air Pressure is medium and low, Wind Speed is medium and Strong, and Precipitation is Absent.

c) Light Rain weather output is determined if Rainfall is slight.

d) Moderate Rain weather output is determined if Rainfall is moderate.

e) Heavy Rain weather output is specified if there is a lot of Rainfall.

Validation was conducted using historical weather data, and the results of the fuzzy classification were compared with official data from national weather agencies to ensure accuracy.

Each sensor was calibrated prior to deployment. The BME280 sensor’s accuracy for temperature, humidity, and air pressure measurements was validated using a laboratory-grade reference device. The anemometer’s wind speed readings were tested against a professional-grade wind tunnel, and the rainfall sensor was validated by simulating different precipitation levels in a controlled environment. The overall system accuracy was found to be within acceptable ranges, achieving a classification

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2330, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

success rate of 100% during field tests. Field testing was conducted over seven days in an open area. The tool's performance was evaluated under various weather conditions, and data was logged for analysis. The results showed high reliability and consistency, with real-time monitoring and classification outputs aligning with actual weather conditions.

III. RESULT AND DISCUSSIONS

Weather monitoring and classification system using fuzzy logic method based on IOT is a system that serves to know the weather conditions in realtime that can be seen on the I2C LCD screen and Blynk application. The results of the design are divided into two parts, namely the design of the hardware circuit design, then the results of the software design.

A. Hardware Design Result

The hardware design for this fuzzy logic-based weather measurement tool consists of several main components that work together to collect and process weather data. This system uses solar panels as the main energy source that converts solar energy into electrical energy with a specification of 20 WP. The electrical energy generated is then regulated by the Solar Charger Controller. This module serves to charge the 18650 battery efficiently, using a linear charging method with a maximum current of 30 A. The energy stored in the 18650 12V battery (with a capacity of 12800mAh) is used to provide power for the entire system. The ESP8266 microcontroller acts as the center of data processing and communication in this system. The ESP8266 is in charge of collecting data from various sensors and sending the data to the Blynk platform via WiFi connection. The sensors used include a BME280 sensor that measures temperature, humidity, and air pressure, an anemometer to measure wind speed, and a rainfall sensor that measures rainfall intensity. The data collected by the ESP8266 is processed using the fuzzy logic method to produce an output that represents the weather comfort level. The processed data is then displayed on an I2C LCD, making it easy to monitor weather conditions locally.

Fig. 4. Hardware Design Result

In Figure 4, it can be seen that the monitoring device circuit uses a 170 cm high pole to support each sensor - sensor, solar panel, and controller box. The selection of poles with a height of 170 cm is because the anemometer sensor and solar panel must be in a high place so as not to be obstructed by objects and shadows that can cause the monitoring tool to not work optimally.

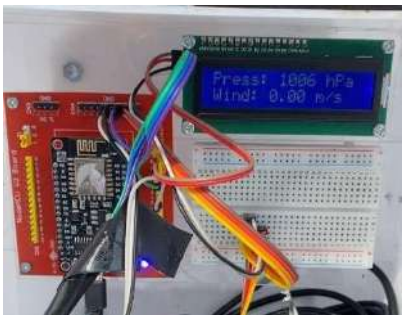


Fig. 5. Device circuit inside the box

In Figure 5, the controller cover box with size P(20cm) x L(20cm) x T(15cm) is used to cover and protect the nodeMCU module, I2C LCD, and protoboard from rain.

B. Software Suite Results

The results of the software suite are displayed on the Blynk application. The following table shows the detection results of several sensors and weather classification.

Fig. 6. View on the Blynk app

From Figure 6, it can be seen that the Weather Station device displays the results of sensor readings - sensors sent data by

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2330, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

NodeMCU in the form of gauges and graphs. Monitoring using the Blynk application can function properly

C. Tool Testing Results

The testing process of the Internet Of Things (IOT) based weather monitoring and classification tool was carried out at Jalan Kelapa Gading Blok II-A in an open area. Testing was carried out for 7 consecutive days from June 10 - June 16, 2024. In testing this tool there are 3 sensors, namely the BME280 sensor (temperature, humidity, and air pressure), Anemometer sensor (Wind Speed), and Rainfall sensor (rainfall) which are used fuzzy methods to classify weather. The results of the monitoring tool testing carried out can be seen in the following table.

TABLE vii TOOL TESTING RESULTS

Date	*Temperature* (°C)	*Humidity* (%)	*Pressure* (hPa)	*Windspeed* (m/s)	*Rainfall* (mm/mnt)	Cuaca	Cuaca Sebenarnya
10/06/2024	29,04	72,6	1009,76	0,24	0	Cerah	Cerah
	Normal	Tinggi	Rendah	Pelan	Tidak ada		
11/06/2024	33,67	76,28	1010,15	0,17	0	Cerah	Cerah
	Panas	Tinggi	Sedang	Pelan	Tidak ada		
12/06/2024	33,4	75,24	1010,15	0,16	0	Cerah	Cerah
	Panas	Tinggi	Sedang	Pelan	Tidak ada		
13/06/2024	33,67	75,82	1010,15	0	0	Cerah	Cerah
	Panas	Tinggi	Sedang	Pelan	Tidak ada		
14/06/2024	32,34	75,59	1010,38	0,15	0	Cerah	Cerah
	Panas	Tinggi	Sedang	Pelan	Tidak ada		
15/06/2024	32,13	80	1010,15	1,53	0	Cerah	Cerah
	Panas	Tinggi	Sedang	Pelan	Tidak ada		
16/06/2024	31,82	76,28	1010,15	0,52	0	Cerah	Cerah
	Panas	Tinggi	Sedang	Pelan	Tidak ada		

From table , it can be seen that the monitoring tool testing was carried out from Sunday to Saturday. During the 7 consecutive days of testing, the reading results of the weather monitoring tool on Sunday received an accuracy of 100% because the weather and sensor readings from Monday - Sunday were in accordance with data from BMKG. So it can be stated that the accuracy of the monitoring tool reaches 100%, which means that the monitoring tool can function properly.

D. Analysis of Monitoring tool Test Results

The test results of the Internet of Things (IoT)-based weather monitoring tool using the Fuzzy Logic method have shown satisfactory performance during the 7-day test period from June 10 to June 16, 2024. Here are some analysis points from the test results:

Sensor Performance and Reading Accuracy: This tool uses three main sensors: BME280 for temperature, humidity, and air pressure, Anemometer for wind speed, and Rainfall sensor for rainfall. These sensors were able to provide consistent and accurate weather data throughout the test period. From the comparison results between the tool data and the actual weather recorded, the accuracy of the monitoring tool reached 100%, which shows that this tool can function properly and accurately.

Use of the Fuzzy Logic Method: The application of the Fuzzy Logic method in weather classification has proven to be effective. Fuzzy Logic is able to handle the variability of sensor data and produce weather classifications that are close to actual conditions. The test results show that this tool can classify the weather into sunny, cloudy, and rainy well.

Energy Efficiency: The monitoring device uses solar panels as the main power source and batteries as energy storage. Tests were conducted for 7 days non-stop in an open area, and the device continued to operate well even when the solar panel did not get the maximum energy from the sun, such as at night or in cloudy weather. This shows that the power management system works efficiently, ensuring continuity of operation.

This tool requires a connection to the internet so that the sensor reading data can be sent to the blynk server, but when taking data in the middle of the farm still requires an internet connection via Hotspot tethering from a cellphone so that it cannot be left behind because the distance from the farm to the residence is far while the range of cellphone hotspots is usually around 10 meters. Sometimes also the signal in the middle of the farm is very bad so it takes longer for the tool to send the readings to the Blynk server.

Use of Blynk App and LCD: Weather data can be monitored in real-time through the Blynk app and I2C LCD display. The display on the Blynk app shows sensor readings in gauge and graph form, which makes it easy for users to monitor weather conditions. The LCD display also clearly shows temperature, humidity, air pressure, wind speed, and rainfall data.

Testing and Environmental Conditions: Tests were conducted in an open area, ensuring that the tool can operate in varied environmental conditions. The test conditions were continuous for 7 days from June 10 to June 16, proving the durability and reliability of the device in various weather conditions.

IV. CONCLUSION

Based on the research and testing that has been carried out on weather monitoring and classification tools using the Fuzzy Logic method based on Internet Of Things (IOT), it can be concluded that:

The design of weather monitoring and classification tool using fuzzy logic method based on Internet of Things uses several sensors, namely BME 280 sensor that reads temperature, humidity, and air pressure, wind speed sensor, and rainfall sensor. The sensor reading data is processed by the MCU Node and then displayed on the LCD screen and Blynk application.

The application of the fuzzy logic method to the output of each sensor can be done to classify the weather so that it can help agriculture in determining the rice planting schedule in order to increase production.

Integration between monitoring results on the tool and weather classification using fuzzy logic method is done in Arduino IDE software using C++ programming language so that the accuracy reaches 100%. This tool can also be monitored remotely using the Blynk application.

Because this tool requires a connection to the internet so that

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2330, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

the sensor reading data can be sent to the blynk server, it is necessary to use Long Range (LoRa) communication that can send data remotely without using the internet. It can also use the internet network that is currently booming, namely starlink, because starlink uses satellite internet so that the internet can run smoothly if the tool is placed in the middle of a large farm.

REFERENCES

[1] M. N. O. Sadiku, T. J. Ashaolu, A. Ajayi-Majebi, and S. M. Musa, "Internet of Things in Agriculture: A Primer," *International Journal Of Scientific Advances*, vol. 2, no. 2, pp. 215–220, 2021, doi: 10.51542/ijscia.v2i2.24.

[2] G. Vitali, M. Francia, M. Golfarelli, and M. Canavari, "Crop Management with the IoT: An Interdisciplinary Survey," *Agronomy*, vol. 11, no. 1, pp. 1–18, 2021, doi: 10.3390/agronomy11010181.

[3] D. M. Quan and H. V. Trần Thị, "Potential and Trends of IoT Application in Agriculture in Vietnam," *Journal of Social Science*, vol. 4, no. 3, pp. 666–676, 2023, doi: 10.46799/jss.v4i3.256.

[4] N. V. G. Deekshithulu, J. Mali, V. V. Krishna, and D. Surekha, "Design, Development and Implementation of Real Time Canal and Weather Monitoring Devices," *Journal of Experimental Agriculture International*, vol. 43, no. 11, pp. 1–13, 2021, doi: 10.9734/jeai/2021/v43i1130752.

[5] F. N. Ibraheem, S. N. Abdulrazzaq, I. Fathi, and Q. Ali, "High-Resolution and Secure IoT-Based Weather Station Design," *International Journal of Safety and Security Engineering*, vol. 14, no. 1, pp. 249–258, 2024, doi: 10.18280/ijsse.140125.

[6] M. Pawar, V. More, P. Jalke, and Dr. A. R. Nichal, "IoT Based Smart Agriculture Monitoring, Weather Control and Irrigation System," *International Journal of Research Publication and Reviews*, vol. 03, no. 12, pp. 1688–1694, 2022, doi: 10.55248/gengpi.2022.31251.

[7] N. Setyanugraha, S. Al Aziz, I. W. Harmoko, and F. Fianti, "Study of a Weather Prediction System Based on Fuzzy Logic Using Mamdani and Sugeno Methods," *Physics Communication*, vol. 6, no. 2, pp. 61–70, 2022, doi: 10.15294/physcomm.v6i2.39703.

[8] Y. Wang and Y. Yang, "A Novel Secure and Energy-efficient Routing Method for the Agricultural Internet of Things Using Whale Optimization Algorithm," *Journal of Cyber Security and Mobility*, vol. 13, no. 4, pp. 725–750, 2024, doi: 10.13052/jcsm2245-1439.1347.

[9] P. Vaishnavi *et al.*, "An Efficient Agricultural Data Analysis Based On IOT," *MATEC Web of Conferences*, vol. 392, p. 01098, 2024, doi: 10.1051/matecconf/202439201098.

[10] P. Zhang, X. Chen, S. Li, C. Zhang, and Y. Hu, "Development of the Internet of Smart Orchard Things Based on Multi-Sensors and LoRa Technology," *Intelligent and Converged Networks*, vol. 4, no. 4, pp. 342–354, 2023, doi: 10.23919/ICN.2023.0028.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2330, Copyright ©2025
Submitted : December 13, 2024, Revised : January 17, 2025, Accepted : January 23, 2025, Published : January 31, 2025

Comparative Analysis of RESTful, GraphQL, and gRPC APIs: Perfomance Insight from Load and Stress Testing

Steven Chandra[1]*, Ahmad Farisi[2]
Faculty of Computer Science and Engineering [1], [2]
Universitas Multi Data Palembang
Palembang, Indonesia
stevenchandrafei@mhs.mdp.ac.id[1], ahmadfarisi@mdp.ac.id [2]

Backend constitutes a critical component of digital infrastructure, responsible for processing business logic, managing data, and facilitating communication between software systems. APIs serve as the interface that enables software interaction and plays a pivotal role in backend operations. This study investigates the performance of three API architectures: RESTful, GraphQL, and gRPC. The experimental approach involves the implementation of Load Testing and Stress Testing to assess the performance of these architectures. The experiment utilizes a dedicated server and client hardware to simulate real-world conditions, with parameters such as CPU usage, memory usage, response time, load time, latency, success rate, and failure rate evaluated using a dataset comprising 1,000 rows of student-related records. Result show that RESTful achieves the highest total request but exhibit greater resource consumption and a higher failure rate. GraphQL demonstrated better CPU and memory efficiency with strong stability, though it has higher latency and slower response times. gRPC strikes a balance with a moderate latency and resource usage, albeit with slightly higher memory consumption under stress. By presenting a comprehensive analysis of each API architecture, this study contributes a comprehensive performance analysis under practical testing scenarios giving developers and system architect with data-driven guidance for selecting API architecture to their application needs. RESTful is well suited for high-throughput scenarios with less critical operations, GraphQL excels in resource efficiency and stability, and gRPC offers balanced performance across diverse workloads.

Keywords— API Architecture, gRPC, GraphQL, Restful, Load Testing, Stress Testing

I. INTRODUCTION

One of the most critical elements in digital infrastructure is the backend. The backend is a system component that works behind the scenes, focusing on business logic processing, data management, and is responsible for managing servers and databases [1], [2]. An efficient and reliable backend is essential to support complex operations, such as data processing and real-time updates, especially for applications that handle large volumes of information or need to respond directly to user interactions [3]. In its operations, the backend is supported by APIs that function as interfaces enabling two software components to communicate with each other.

Application Programming Interface (API) is an interface comprising a set of instructions organized in a library [4]. According to [5], an API is code that connects one application to another, providing all the necessary permissions for two software programs to communicate. APIs allow various systems, whether desktop or web, to interact and exchange data with servers or databases without requiring an additional backend, thus simplifying application integration and development. Additionally, API architecture is a critical aspect in determining how an API is organized and implemented.

In recent years, several API architectures have been developed to meet the needs of applications and developers in various scenarios, such as REST, GraphQL, and gRPC. REST (Representational State Transfer) is an architectural style for distributed systems that separates the interface on the client side and business logic on the server side to achieve anarchic scalability in line with internet growth [6]. GraphQL, developed by Facebook in 2012 and publicly released in 2015, is a dynamic single-endpoint query language for interacting with APIs [7]. gRPC is an open-source Remote Procedural Call (RPC) framework developed by Google [8].

With the growing demand for applications to handle user requests and large data volumes quickly and responsively, it is crucial for developers to choose the right API architecture according to application requirements. Each API architecture has its functions and use cases that influence application performance. Several previous studies have compared the performance and effectiveness of API architectures in various scenarios, such as the study Evaluation of Microservices Communication while Decomposing Monoliths [9], which focused on evaluating microservices technologies like HTTP REST, RabbitMQ, Kafka, gRPC, and GraphQL. This study used criteria such as latency, throughput, message size, and memory consumption to test the performance of these technologies. The results showed that each technology has its strengths and weaknesses, such as HTTP REST being simpler and more efficient for direct communication needs, while RabbitMQ and Kafka, which use message brokers for asynchronous communication, are better suited for architectures requiring high availability and loose coupling between services. However, this study had limitations in terms of parameter scope, device types, testing tools, programming

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2315, Copyright ©2025
Submitted : January 11, 2025, Revised : January 23, 2025, Accepted : January 26, 2025, Published : January 31, 2025

languages, and the complexity of the data used.

Another relevant study, Implementation Comparison of GraphQL and REST API Methods on Node.js Technology by [10], compared two commonly used API architectures in application development: REST API and GraphQL on Node.js. GraphQL was found to be more flexible and efficient than REST API as it allows clients to customize the data they need to display. This study compared the performance of the two API architectures using parameters such as response time and scalability. However, the testing in this study had some shortcomings, including limited testing parameters, the use of simulated data that did not represent real-world complexities, and a lack of repeated testing to ensure consistency of results.

Another study, Analysis of the Effectiveness Comparison Between RESTful and gRPC Architectures in Web Service Implementation [6], used parameters such as response time, response size, CPU usage, throughput, and load time. The study indicated that gRPC has more stable and faster response times as data volume increases compared to RESTful. It also stated that gRPC has lower CPU usage and response sizes compared to RESTful.

Previous studies often focus on limited research parameters, use of simulated data does not represent real-world complexity, and unequal comparison such as [6], [9], [10] [11], and [12] making it difficult to draw definite conclusions for real-world applications. To addresses these gaps, this study aims to comprehensively comparing RESTful, GraphQL, and gRPC in Load Testing and Stress Testing scenario by using seven parameters such as CPU usage, response time, latency, memory usage, loading time, success rate, and failure rate. These seven parameters overlap with those from previous studies which typically examines only two or three parameters, thus enabling a more thorough and realistic evaluation of API architectures.

II. RESEARCH METHODS

This study uses quantitative True-Experimental to compare the performance between three API architectures, namely RESTful, GraphQL, and gRPC. This method was chosen because it allows full control with research variables and ensures high internal validity in performance testing. This study aims to analyze the performance of RESTful, GraphQL, and gRPC API architectures.

A. Research Flow

The flow of the research experiment can be seen in Figure 1 below.

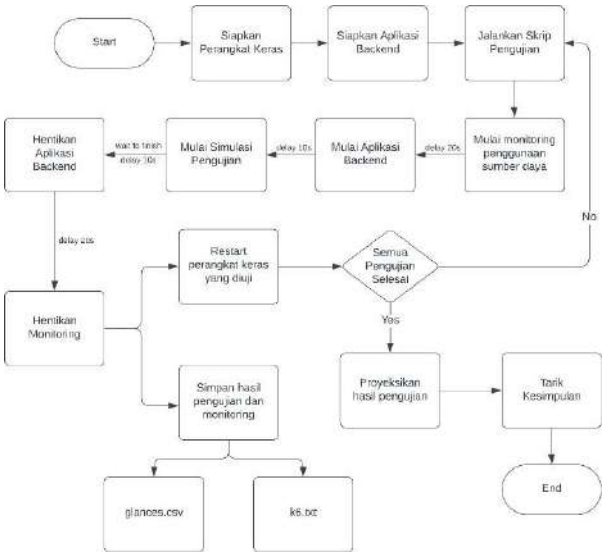


Fig. 1. Research Flow

The experiment begins with preparing the necessary hardware. Afterward, the backend application is configured with required settings. The test script is then executed to start the simulation and data collection process. First, Glances which is be used for monitoring resource usage is started. After a 20-second dela, the backend application is launched, followed by another 10-second delay before starting the K6 simulation script. The simulations are carried out until completion, with a duration of 6 minutes for Load Testing and 23 minutes for Stress Testing. Once the simulation is finished there is a 10-second delay before shutting down the backend application, followed by a 20-second delay to stop the resource monitoring. The results of the test and monitoring are stored in files glances.csv for resource monitoring logs and k6.txt for load test result logs generated by K6. This data is then analyzed to project test results and draw conclusions regarding system performance. The experimentation process is considered complete once all tests are finished and conclusions are drawn based on the collected data.

B. Research Scope

This research uses Load Testing and Stress Testing to compare the performance between three API architectures namely RESTful, GraphQL, and gRPC. This method was chosen because it allows full control with the research variables and ensures high internal validity in performance testing such as CPU usage, response time, latency, memory usage, load time, success rate, and failure rate. data used for testing consists of student records joined with tables for students, accounts, institutions, and study programs, totaling 1,000 rows of data. This choice of 1,000 rows of data balances complexity and resource constraints, providing a manageable sample that reveals key performance behaviors without overwhelming the test environment. It also provided a solid baseline for assessing how each API architecture handle typical loads before scaling up to a larger dataset if needed. By maintaining a controlled dataset size this research can focus on the architectural differences on performance, ensuring that trends remain

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2315, Copyright ©2025
Submitted : January 11, 2025, Revised : January 23, 2025, Accepted : January 26, 2025, Published : January 31, 2025

attributable to the API design rather than external factors introduced by excessive data volumes. The evaluation are conducted by using two dedicated hardware to act as server and client for testing to simulate a real world scenario, the hardware specification can be seen in Table I.

TABLE I. SPESIFICATION TABLE

Spesification	Hardware	
	Server	Client
CPU	Intel Core i7-7700HQ	AMD Ryzen 5 7535HS
RAM	20 GB DDR4	16 GB DDR5
Operating System	Ubuntu Server 24.04.1 LTS	Windows 11 23H2

III. RESULT AND DISCUSSION

The following experimental results that have been carried out can be seen in the following 3.1 and 3.2.

3. 1. Load Test

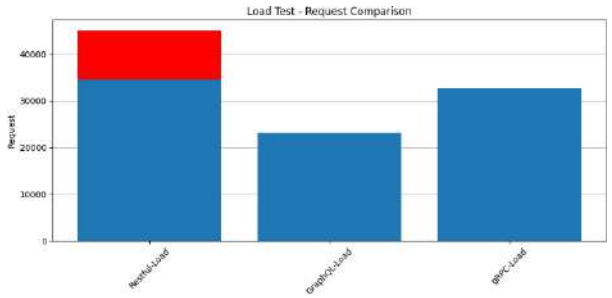


Fig. 2. Load Testing Results – Total Comparison

In Figure 2, the results of Load Testing total requests show that Restful architecture has the largest total requests compared to GraphQL and gRPC during the same testing period. However, Restful has a higher request failure rate than the other two architectures. GraphQL processes fewer requests than Restful, while gRPC performance is close to Restful with no failed requests like Restful.

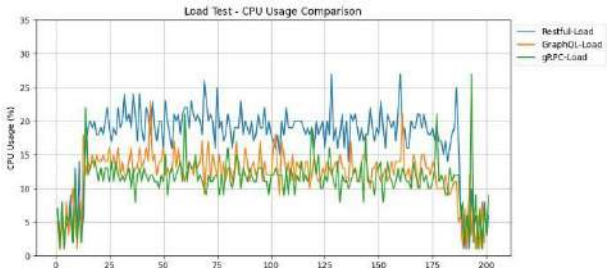


Fig. 3. Load Testing Results – CPU Usage

In Figure 3, the results of Load Testing CPU usage show that Restful uses higher CPU resources than other architectures. GraphQL shows a lower and stable CPU usage with a range of 15%. While gRPC has lower CPU usage than GraphQL and Restful it has some spikes in CPU usage especially in the few seconds before the test ends, this may possibly due to protocol buffer encoding/decoding and connection handling toward the end of the test.

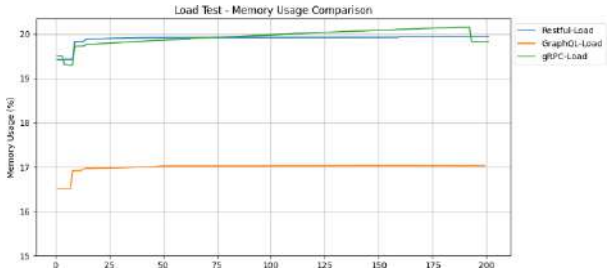


Fig. 4. Load Testing Results – Memory Usage

In Figure 4, the simulated Load Testing results of memory usage show GraphQL has lower memory usage at 17%. While Restful and gRPC have identical memory usage at around 20%. This shows GraphQL is more efficient in memory usage than Restful and gRPC.

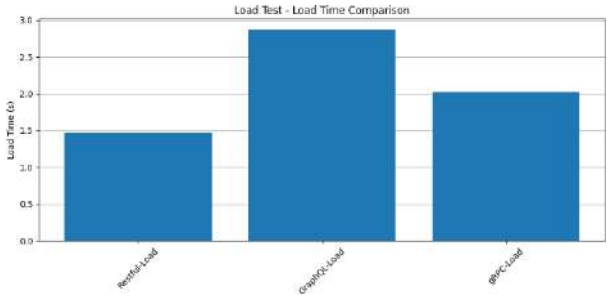


Fig. 5. Load Testing Results – Load Time Comparison

In Figure 5, shows a comparison of the average load time of the API architectures during Load Testing. GraphQL has the highest average load time at around 2.8 seconds, followed by gRPC with an average of around 2 seconds. While Restful has the lowest average load time, which is 1.4 seconds. These results show that Restful tends to have a faster load time response than GraphQL and gRPC.

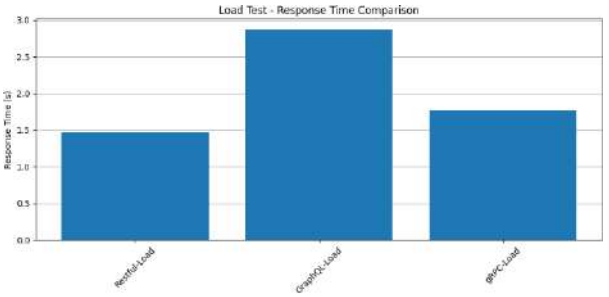


Fig. 6. Load Testing Results – Response Time Comparison

In Figure 6, the results of the average response time in Load Testing show a similar pattern to the average load time. GraphQL stands out with an average response time of 2.8 seconds, while gRPC has an average response time of 1.7 seconds. Restful shows the lowest response time of under 1.6 seconds compared to the other two architectures.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2315, Copyright ©2025
Submitted : January 11, 2025, Revised : January 23, 2025, Accepted : January 26, 2025, Published : January 31, 2025

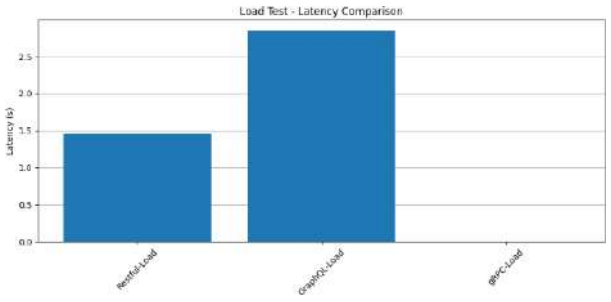


Fig. 7. Load Testing Results – Latency Comparison

In Figure 7, shows the average latency results on Load Testing. The graph shows that GraphQL shows the highest latency by reaching 2.85 seconds followed by Restful with a latency of about 1.5 seconds. Meanwhile, gRPC latency is written as 0 because it cannot be measured using the k6 testing tool.

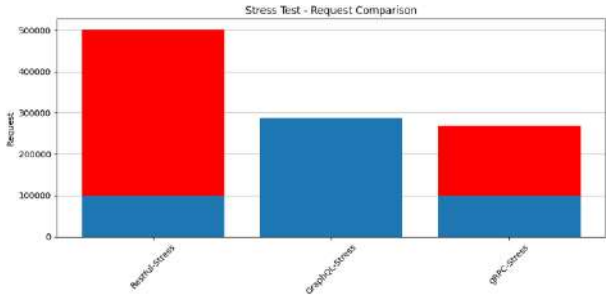


Fig. 8. Load Testing Results – Total Comparison

In Figure 8, the total request Stress Testing results show Restful architecture has higher performance with 500,000 total requests, but has a low success rate at 20%. gRPC also follows a similar pattern with 267,000 total requests with 37% success rate. In contrast, GraphQL showed higher total requests than gRPC, having a 100% success rate out of 287,000 requests.

To facilitate easier comparison, Table II summarize key perfomance metrics for RESTful, GraphQL, and gRPC under Load Testing.

TABLE II. LOAD TEST PERFOMANCE SUMMARY

Metric	RESTful	GraphQL	gRPC
Request Processed	44,976	23,055	32,668
Succes Rate	76.64%	100%	100%
Failure Rate	23.36%	0%	0%
Avg CPU Usage	Highest	Moderate	Lowest Average, occasional spike
Memory Usage	~20%	17%	~20%
Load Time	1,47ss	2,87s	2,02s
Response Time	1,47s	2,87s	1,77s
Latency	1,46s	2,85s	Not measured by K6

3. 2. Stress Test

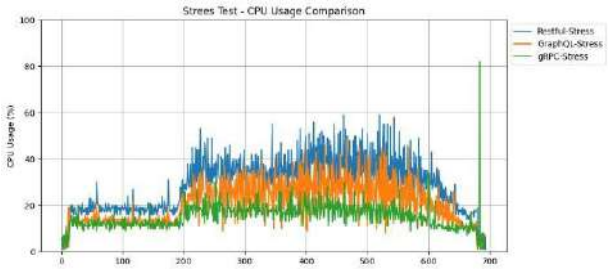


Fig. 9. Stress Testing Results – CPU Usage

In Figure 9, Stress Testing CPU usage also shows a similar trend to Load Testing. Restful tends to use higher CPU than GraphQL and gRPC. While GraphQL has relatively lower CPU usage than Restful. gRPC shows lower CPU usage than the other two architectures, with some spikes in CPU usage at some periods this is especially noticeable in the moments before the test is completed.

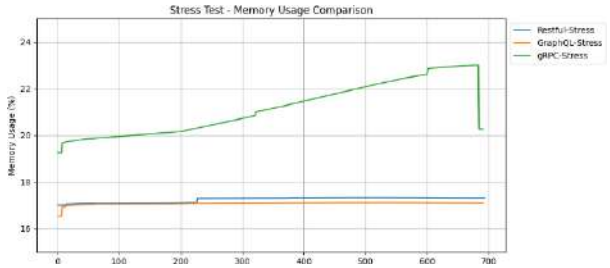


Fig. 10. Stress Testing Results – Memory Usage

In Figure 10, shows the results of Stress Testing memory usage shows a different pattern compared to the previous Load Testing. gRPC shows higher memory usage than Restful and GraphQL memory usage which tends to be stable and almost the same. This indicates that gRPC utilizes more memory than Restful and GraphQL in handling stress testing conditions.

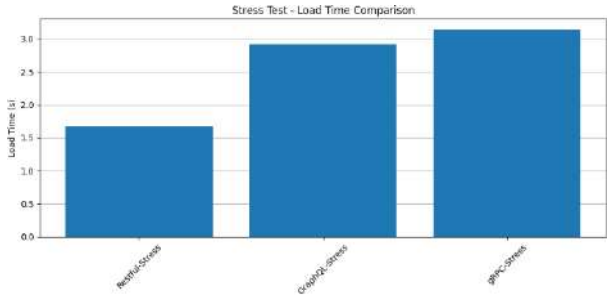


Fig. 11. Stress Testing Results – Load Time Comparison

In Figure 11, shows the average load time results for Restful, GraphQL, and gRPC. Restful has the lowest average load time at around 1.5 seconds. While GraphQL and gRPC show high average load times around 3 seconds, with gRPC having the highest load time of 3.14 seconds.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2315, Copyright ©2025
Submitted : January 11, 2025, Revised : January 23, 2025, Accepted : January 26, 2025, Published : January 31, 2025

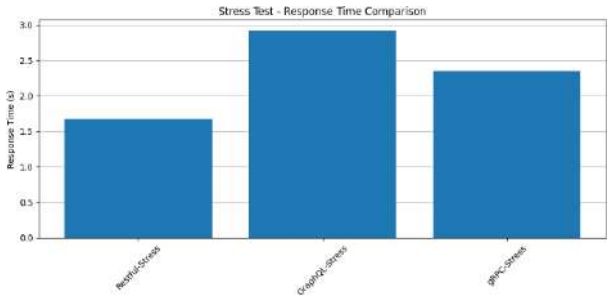


Fig. 12. Stress Testing Results – Response Time Comparison

In Figure 12, shows the results of the average response time during stress testing. Restful has the lowest response time at 1.6 seconds, while GraphQL has the highest response time at 2.9 seconds. gRPC has an intermediate response time of 2.3 seconds.

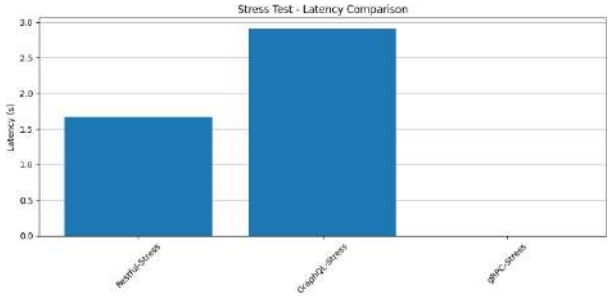


Fig. 13. Stress Testing Results – Latency Comparison

In Figure 13, shows the average latency results of Restful, GraphQL, and gRPC during Stress Testing. The test results are consistent with the Load Testing test, where GraphQL shows the highest average latency at 2.9 seconds while Restful shows an average latency of 1.6 seconds. Meanwhile, gRPC latency could not be measured by the k6 testing tool.

To facilitate easier comparison, Table III summarize key perfomance metrics for RESTful, GraphQL, and gRPC under Stress Testing.

TABLE III. STRESS TEST PERFOMANCE SUMMARY

Metric	RESTful	GraphQL	gRPC
Request Processed	501,460	287,603	267,594
Succes Rate	80.21%	100%	37.27%
Failure Rate	19.79%	0%	62.73%
Avg CPU Usage	Highest	Moderate	Lowest Average, occasional spike
Memory Usage	~20%	17%	~20%
Load Time	1,67ss	2,92s	3,14s
Response Time	1,67s	2,92s	2,35s
Latency	1,67s	2,91s	Not measured by K6

IV. CONCLUSION

The research result of both types of tests shows consistent results, Restful architecture shows the ability to execute more requests in the same period of time, but the success rate is lower than GraphQL and gRPC. GraphQL shows better CPU and memory utilization efficiency and is stable in both tests, while gRPC offers a balance between performance and resources, although under sustained stress conditions gRPC memory utilization is higher than other architectures. It is important to note however that these experiments arise from limited a relatively limited dataset (1,000 rows) and a specific hardware configuration, which could affect generalizability to real-world applications with larger databases or different infrastructure. Future research and experiment therefore could investigate this API architecture with more substantial datasets, potentially in tens of thousands of records, and explore variations in programming language and API architectures. Additionally future research could integrate more comprehensive gRPC monitoring tools, detailed time-series monitoring, and granular profiling to pinpoint the exact cause of gRPC resource spike.

REFERENCES

[1] D. A. Hutomo Putra, E. Darwiyanto, and R. Nurtantyana, "Development of Backend Admin Dashboard for Business Project Monitoring using Scrum Method," *Indonesia Journal On Computing (Indo-JC)*, vol. 9, no. 2, pp. 118–133, 2024, doi: 10.34818/indojc.2024.9.2.969.

[2] Humdiana and Julieca, "Implementation of Full Stack Web Development for Data Admin on Social Media Buzzbuddies," *International Journal of Social Science (IJSS)*, vol. 2, no. 2, pp. 1535–1544, Aug. 2022, doi: 10.53625/ijss.v2i2.3093.

[3] M. J. Jamshed o'g'li, "The Significance of Backend Development in Modern Web Applications," *International Journal of Scientific Researchers*, vol. 8, no. 1, pp. 737–741, 2024.

[4] N. K. Dwi Sabrina, D. Pramana, and T. M. Kusuma, "Implementation of Golang and ReactJS in the COVID-19 Vaccination Reservation System," *ADI Journal on Recent Innovation (AJRI)*, vol. 5, no. 1, pp. 1–12, Feb. 2023, doi: 10.34306/ajri.v5i1.877.

[5] A. F. Rochim, T. N. Wijaya, and D. Eridani, "A Citation Data Collector Tool of Author's Profiles in Scopus Based on Web and Application Programming Interface (API)," *IOP Conf Ser Mater Sci Eng*, vol. 1077, no. 1, p. 012017, Feb. 2021, doi: 10.1088/1757-899x/1077/1/012017.

[6] M. I. Yanuardi, Aminudin, and M. Faiqurahman, "Analisis Perbandingan Efektivitas Arsitektur RESTful dan Arsitektur gRPC pada Implementasi Web Service," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 10, no. 2, pp. 333–341, Aug. 2024.

[7] S. Guha and S. Majumder, "A Comparative Study between Graph-QL & Restful Services in API Management of Stateless Architectures," *International Journal on Web Service Computing*, vol. 11, no. 2, pp. 1–16, Jun. 2020, doi: 10.5121/ijwsc.2020.11201.

[8] N. Jagnik, "Highly Performant Python Services using gRPC and AsyncIO," *Journal of Scientific and Engineering Research*, vol. 8, no. 1, pp. 225–229, 2021, [Online]. Available: www.jsaer.com

[9] J. Kazanavičius and D. Mažeika, "Evaluation Of Microservice Communication While Decomposing Monoliths," *Computing and Informatics*, vol. 42, pp. 1–36, 2023, doi: 10.31577/cai.

[10] Muntahanah, Y. Darmi, and K. Pinandita, "Implementasi Perbandingan Metode GraphQL dan REST API pada Teknologi Nodejs," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 7, no. 1, pp. 25–34, 2024.

[11] M. N. Hedelin, "Benchmarking and Performance Analysis of Communication Protocols: A Comparative Case Study of gRPC, REST, and SOAP," KTH Royal Institute of Technology, 2024. Accessed: Nov. 03, 2024. [Online]. Available: https://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A1887929&dswid=-8588

[12] Goriparthi, "Streamlining API Development: A Comparative Analysis of GraphQL and RESTful Web Services," *International Journal of Data Analytics Research and Development*, vol. 2, no. 1, pp. 59–71, 2024, [Online]. Available: https://iaeme.com/Home/issue/IJDARD?Volume=2&Issue=1

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2315, Copyright ©2025
Submitted : January 11, 2025, Revised : January 23, 2025, Accepted : January 26, 2025, Published : January 31, 2025

Business Intelligence Model of Regional Hospitals using HGOD Discovery

Hengki [1]*, Rahmat Gernowo [2], Oky Dwi Nurhayati [3]
Doctoral Program of Information System, School of postgraduate studies [1], [2], [3]
Diponegoro University
Semarang, Indonesia
hengki@atmaluhur.ac.id [1], rahmatgernowo@lecturer.undip.ac.id [2], Okydwin@gmail.com [3]

***Abstract*— Based on data from the Regional General Hospital in the Bangka Belitung Islands province, the Gross Death Rate (GDR) is the general death rate for every 1000 patients discharged of 108,430 compared to the health department standard of <45. The Net Death Rate (NDR) is the death rate 48 hours after being treated for every 1000 patients discharged of 67,388 compared to the health department standard of <25. TOI (Turn Over Interval) is the average turnover period of days where a bed is unoccupied from being filled to the next time it is filled of 19,832 days compared to the health department standard of 1 to 3 days. The solution offered by the researcher develops Business Intelligence (BI) optimization with a new model called the HGO (Hierarchy, Governance, Outlook) Discovery approach as a framework model for developing business intelligence for regional general hospitals in Indonesia. This model is expected to be able to solve or reduce the dimensional problems that exist in hospitals, namely the main patient management, HR Key Resources, and the quality of inpatient health services. The HGO Discovery approach is able to find patterns in a series of events called sequences by sorting the work patterns that exist in the hospital so that the business process of regional general hospitals is faster and more interactive in decision making. The Business Intelligence approach carried out by regional hospitals with HGOD is expected to make patient health services more integrated through the hierarchy of patient services, governance and outlook in decision making.**

Keywords*— *Business Intelligence, HGOD Discovery, Hospital, Optimatization, Governance.

I. INTRODUCTION

Business intelligence covers almost all information system ecosystems today, but not all sectors utilize Business Intelligence, such as the Regional General Hospital Information System (SIRSUD). Business Intelligence (BI) uses techniques for consolidating data, storing, analyzing data, designing data logic in decision making in multidimensional data using data mining. Currently, regional general hospitals are the main and mainstay facilities used by people in Indonesia. The business intelligence function organizes a decision support system where this system converts and extracts data contained in the hospital (implicit data and explicit data) into knowledge management or knowledge. Business intelligence technology has the potential to develop and explain complex problems in business processes in organizations [1]. Hospitals have an obligation to be fast and agile in making decisions regarding patients and processing data in an important role in the process of taking patient action using technology so that the accumulation of patient data is quickly obtained to extract timely, credible and fully researched information regarding patient data at hospital [2]. Business intelligence is not a product tool but a new approach proposed to make difficult decisions in business as quickly as possible [3]. Business Intelligence (BI) will become an important investment and have an important role in today's hospital management, technology implementation This is very challenging and requires resources, strategy, management and skills [4].

Regional General Hospital patients who generally use BPJS health generally have problems with slow registration and treatment processes at the hospital, researchers have even seen how many cases resulted from delays in treatment causing patients to die. After the outbreak of Covid-19, hospitals had to adopt technology and systems to address the slowness of hospital services that had occurred so far [5]. The BPJS feature used by Indonesian society still has various problems that cause patients to queue for hours at regional public hospitals. Apart from that, in general there is no Business Intelligence technology, especially a decision-making feature that contains patient history where the user of the history system is the patient. Most hospitals, especially the emergency department, have inefficiencies in human resource management and data processing, resulting in slow follow-up by the hospital front line to assist their work [6]. Usually regional hospital system users currently try various support patterns to view patient history which results in the process of searching or mining patient data being long and repetitive and the current approach to the data display process is not interactive.

Here are some of the original things in this research as follows : Developing a Business Intelligence model with the HGOD (Hierarchy, Governance, Outlook) Discovery approach, which previously hospitals used more of the decision-making Simulation approach or model which had the weakness of a gap between medical resources and patient needs. Developing a model that is integrated into a framework with 3 approaches, namely HGO (Hierarchy, Governance, Outlook)

There is a gap between medical resources and patient needs, managers need to obtain productivity information to optimize

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2320, Copyright ©2025
Submitted : November 22, 2024, Revised : December 13, 2024, Accepted : December 18, 2024, Published : January 31,2025

resource allocation. The value of this study is that the information provided by the dashboard allows hospital managers to respond quickly. We recommend that the study can integrate more data (such as temperature data, national death population data records), so that it is covered for hospital operating cost control and to estimate patient needs.

II. LITERATURE REVIEW

Smart Business Intelligence is a set of models and technologies designed to efficiently extract and process useful information from data. its application to integrate and operate with various institutional or institutional data sources. Apart from that, BI (Business Intelligence) is a step in the decision-making process that is supported by integration and analysis of organizational data resources. BI (Business Intelligence) plays an increasingly important role in several types of institutions and institutions because various kinds of information have been identified as the most valuable asset in an institution and a fundamental resource.

Business Intelligence has an important objective, namely to provide services with quality organizational intelligence and the purpose of this research is to assess organizational intelligence in hospitals which is carried out based on Karl Albrecht's organizational intelligence components. Research tries to provide service quality status based on these components [7].

Healthcare systems face enormous challenges, essentially due to the amount of data generated every day in the hospital environment, which forces entities to think about how to organize and use the same data.

The business ecosystem becomes complex in industrial development. Therefore, to provide a fast response in this dynamic era, companies need innovation and advanced technology. In this context, Smart Business Intelligence (SBI) is needed to process information and make the right decisions at the level of all institutional lines or institutions. This technology is applied in an organization, it can provide several benefits such as architecture, efficient information, and customer data management. With this approach, agencies can get a clearer picture of how important BI is in all agency sectors Currently, the number of studies at this level is growing, with a focus on innovations to be implemented, so that the same sector can adopt new methodologies, architectures and technologies that allow more efficient support of existing hospital processes, as well as their results to be provided for all professionals involved in this field. In this research, an Adaptive Business Intelligence architecture is proposed, whose contribution is supported by the realization of an adequate conceptual and technological framework that describes its development at various levels. Thus, the possibility of modernization of several work approaches begins, with the introduction of an architecture capable of contributing to several factors, both at the clinical and administrative level, meeting the needs of the hospital system, regarding design, development, implementation and demonstration of the results [8].

Business Intelligence can provide effective and useful insights for investors and business owners to utilize more appropriate BI tools and functions to achieve more ideal organizational excellence. It also allows managers to better understand the application of BI functions in the process of achieving specified managerial support benefits [9].

The current implementation of Business Intelligence has an impact and influences the current economy, intangible assets have gained considerable appreciation for the continuity of organizations and institutions. Therefore, important parts of business value, such as intellectual property strategy, play a vital role in defining, creating and maintaining a superior business strategy. This allows the creation of value and strengthening various strategic aspects for business continuity. The implementation of a combination of BI (Business Intelligence) and AI (Atificial Intelligence) has emerged as a new approach to assessing and creating real-time technologies such as weather technology and capital markets throughout the world. Then there is also Big Data technology which is able to create value and show the potential for transformation in organizations and for improving business processes.

Each recognized hospital patient management unit (PMU) has focused its efforts on improving clinical patient care, with a process approach, analyzing everything from adult emergency crowding to length of stay in clinical services. This results in many patients waiting for beds in emergency services. PMUs do not have a business intelligence (BI) platform that provides real-time information, creating a blind search problem. The aim is to demonstrate the need for a BI platform using Artificial Intelligence (AI) to analyze in real-time relevant information for decision making.

Hospital Management Information System (SIM-HM) application, which can manage daily operational data, but the information produced is still very limited, so the use of technology and information is not optimal. The role of IT can be increased by implementing a system that can extract and change information business from existing operational data, so that later it can provide support for strategic business decisions in hospitals. The Business Intelligence Road Map approach is an approach to describe the steps and procedures for developing and implementing Business Intelligence [10].

There is a gap between medical resources and patient needs, Hospital Managers need to obtain productivity information to optimize resource allocation [11]. The Hospital Business Intelligence produced lacks analysis, especially for conducting analysis and reporting [12]. The findings show that all three factors of organization, process, and technology equally influence the implementation of business intelligence. However, it is not explained what the results of the information system/application follow-up from the research findings area[13].

Business Intelligence developed and used in hospitals, and user surveys showed positive results. In addition, top management support and involvement in the development of HBIS were found to be critical success factors, and the implementation of the system enabled the hospital to significantly improve the performance of managerial indicators. However, it does not contain key human resource resources managed to assist BI technology to improve the Hospital's

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2320, Copyright ©2025
Submitted : November 22, 2024, Revised : December 13, 2024, Accepted : December 18, 2024, Published : January 31,2025

services [14].

The design science research methodology was developed and used in hospitals, and user surveys showed positive results. Additionally, top management support and involvement in the development of HBIS was found to be a critical success factor, and implementation of the system enabled the hospital to significantly improve performance on managerial indicators. However, it does not contain key HR resources that are managed to help BI technology improve hospital services [15].

The healthcare system faces enormous challenges, essentially due to the amount of data generated every day in the hospital environment, which forces concentration on how to organize and use the same data. Currently, the number of studies at this level is growing, with a focus on the innovations to be implemented, so that the same sector can adopt new methodologies, architectures and technologies that allow more efficient support of existing hospital processes, as well as the results provided for all professionals involved in this field. In this research, an Adaptive Business Intelligence architecture is proposed, whose contribution is supported by the realization of an adequate conceptual and technological framework that describes its development at various levels. Thus, the possibility of modernization of some work approaches begins, with the introduction of an architecture capable of contributing to several factors, both at the clinical and administrative level, meeting the needs of the hospital system, regarding design, development, implementation and meaning of the results.

III. RESEARCH METHODOLOGY

HGOD Discovery is a approach in hospitals referring to the balancing and monitoring mechanisms that form the process of making important decisions in serving Public Health. Hospital management is responsible for ensuring the best performance for patients. Therefore, it is important to understand the characteristics of organizational and provide services to contribute positively to the outlook and performance of the hospital.

Hierarchy (H) is a sequence of processes based on highest to lowest priority. Usually, an organizational hierarchy consists of various levels, starting from the top to the bottom. Hierarchy in this case is a focus on activity processes that are focused on urgent and non-urgent matters. A hierarchy can connect parts either directly or indirectly or either vertically or horizontally. The existing levels will be arranged based on predetermined business processes, resulting in an arrangement of processes that are considered priorities. Hierarchies help in organization, decision making, and resource allocation, and provide a clear structure for understanding relationships between entities.

The hospital governance (G) framework is becoming increasingly important considering the rapid development of health services in general and hospital services in particular. Governance in hospitals refers to the balancing and monitoring mechanisms that form important decision-making processes in serving public health. Hospital management is responsible for ensuring the best performance for patients. Therefore, it is important to understand the characteristics of organizational governance, and provide services to make a positive contribution to the hospital's outlook and performance.

- *Implementing the organization and coordination of the formulation of technical policies for the RSUD Strategic Plan (Renstra) in line with the strategic plan of the Provincial Government which is used as a reference guideline in carrying out tasks.*
- *Implementing the organization and coordination of the formulation of policies for the vision, mission and objectives of the RSUD based on references from the Government's vision and mission*
- *Implementing and coordinating work to the hierarchical structure in accordance with the duties and responsibilities of each for the quality of service at the RSUD UPTD.*

Outlook (O) is an attitude or view towards a situation. In this research, you can have an optimistic or pessimistic outlook about the future. Decision outlook refers to the view, prospect, or expectation of the outcome or consequences of a decision taken. This can include how the decision is expected to affect the future in a particular context.

- *Impact Projection: Estimating the short-term and long-term impacts of the decision on the organization, individual, or situation.*
- *Risks and Opportunities: Identifying potential risks and opportunities that may arise as a result of the decision.*
- *Strategy and Action Plan: Developing strategies to address challenges and take advantage of opportunities based on projections from the decision.*
- *Outcome Evaluation: Measuring and evaluating actual results against initial projections to assess the success of the decision.*
- *External Context: Considering external factors that influence the Decision.*

Conducting Business Intelligence updates and building the existing Regional General Hospital Information System with a Dimensional approach, namely Hierarchy, Governance and Outlook. The steps for completing HGO Discovery are as follows:

$$HGOd_{ij} = \begin{bmatrix} \frac{(Xhg_{ij})}{(Max\ X_{ij})} \\ \\ \frac{(Min\ X_{ij})}{(Xhg_{ij})} \end{bmatrix}$$

The following are existing policies in regional general hospitals, namely: Implementing the implementation and coordination of the preparation of technical policy formulations for the Regional General Hospital Strategic Plan (Renstra) that are in line with the strategic plan that serves as a reference guideline in carrying out tasks. Implementing the implementation and coordination of the preparation of policy formulations for the vision, mission and objectives of the Regional General Hospital based on references from the Government's vision and mission. Implementing and coordinating work to the hierarchical structure in accordance

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2320, Copyright ©2025
Submitted : November 22, 2024, Revised : December 13, 2024, Accepted : December 18, 2024, Published : January 31,2025

with the duties and responsibilities of each for the quality of service at the Regional General Hospital.

This model performs simulations with 2 aspects, namely positive (+) and negative (-) with the result that the value 0.000 is the highest value and the value 1 is the lowest value.

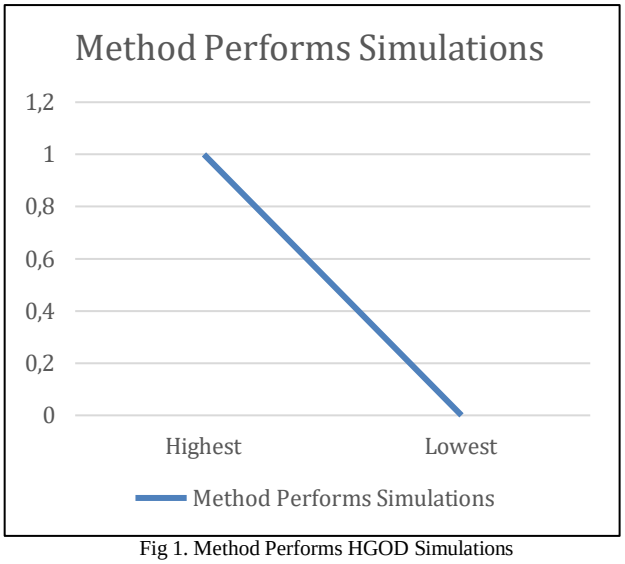


Fig 1. Method Performs HGOD Simulations

Above the image of the results of 0.000 is the highest value or low risk and 1 is the lowest or high risk value. The results of the equation above provide a new picture or perspective on hospital, how an intelligent business is able to provide priorities for what will be done.

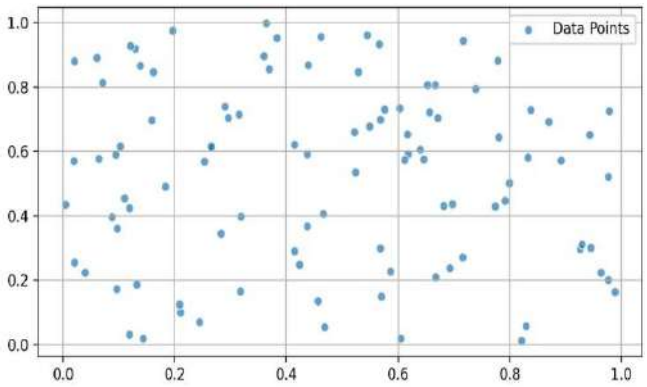


Fig 2. Values of HGOD Model

The following describes how the performance of this model from start to finish produces a dashboard Business Intelligence.

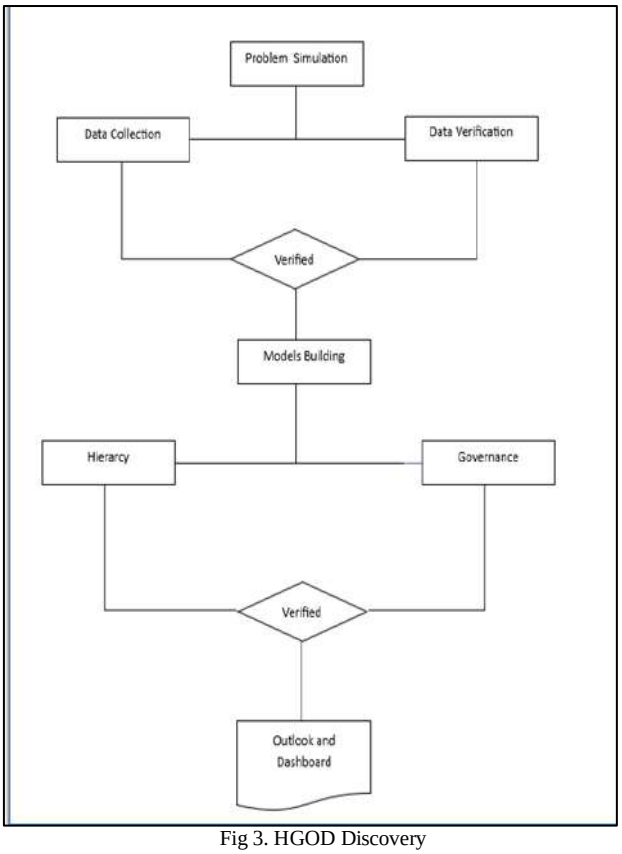


Fig 3. HGOD Discovery

HGOD discovery works starts from problem sumulation where the regional hospital performs data collection and data verification. after that, the system will build building models associated with priority hierarchy and HGOD so as to produce information in the form of outlook and dashboard. The dashboard produced by the data and information processing process which will become a decision-making technique that will be implemented by the general hospital.

Table 1. Comparison of Business Intelligence Model

No	Reference	WT	LS	CO	DD
1	DES (Discrete-Event Simulation)	X	X		X
2	DSRM (The design science research methodology)	X	X		
3	HBIS (hospital-based business intelligence system)	X		X	
4	MCS (Cross-Sectional)	X			
5	GRD (Group Related Diagnostics)	X	X		X
6	ABS (Agent-Based Simulation)	X	X		
7	HGOD Discovery	X	X	X	X

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2320, Copyright ©2025
Submitted : November 22, 2024, Revised : December 13, 2024, Accepted : December 18, 2024, Published : January 31,2025

The identified gaps show that there are 4 focal points that are used as a comparison with other methods, namely WT (Wait Time), LS (Length of Stay), CO (Cost), and DD (Door to Doctor). In these 4 points, each method has its own characteristics that you can see in the table above for comparison.

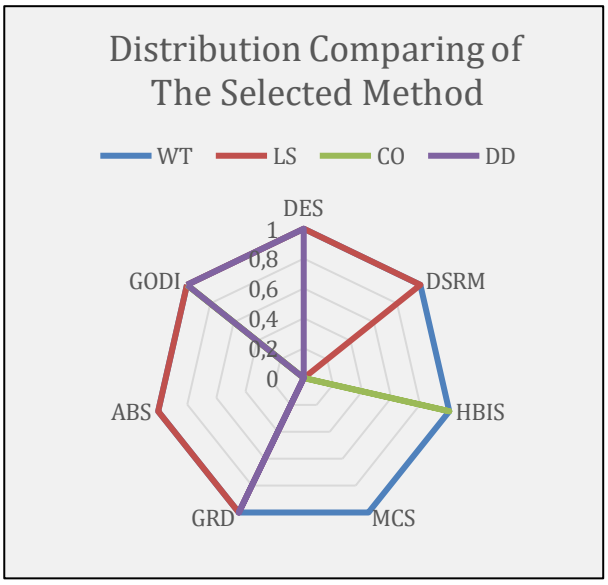


Fig 4. Distribtion Comparing Method

From the comparison results, DES (Discrete-Event Simulation) has 3 points, DSRM (The design science research methodology) has 2 points, HBIS (hospital-based business intelligence system) has 2 points, MCS (Cross-Sectional) has 1 point, GRD (Group Related Diagnostics) has 3 points, ABS (Agent-Based Simulation) has 2 points and HGOD Discovery has 4 points.

In this finding, the focus becomes clear that the increasing importance of simulation analysis in improving hospital services with business intelligence. The finding is the importance of a method that considers and emphasizes what happens in a hospital system so that patient management is maximized and the involvement of health care professionals in validation and verification simulations. In addition, this study proposes a HGOD Discovery framework that offers valuable insights into management, technology, human resources, and policies.

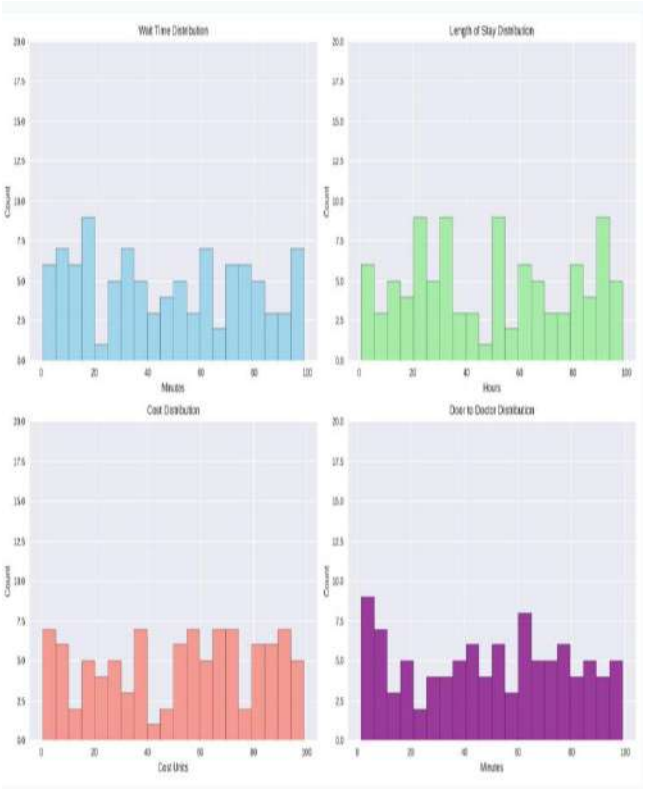


Fig 5. Simulation dashboard criteria HGOD

The following is a simulation of the dashboard graph for WT (Wait Time), IS (Length of Stay), CO (Cost), and DD (Door to Doctor) with a value range of 0 to 100.

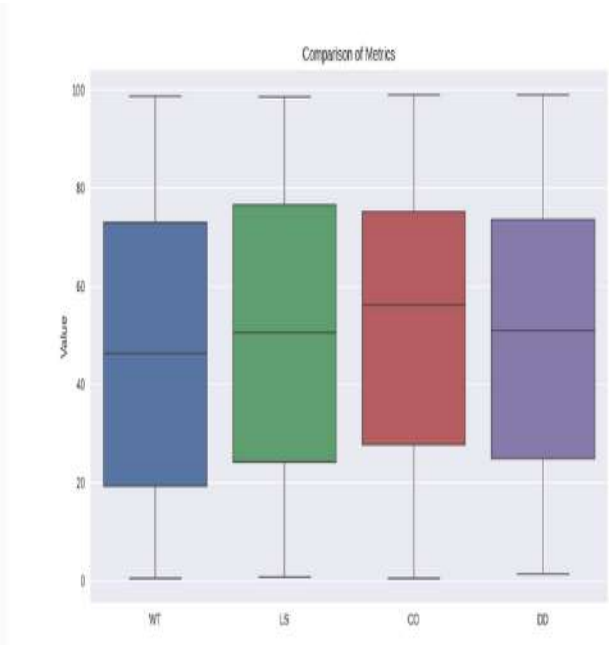


Fig 6. Simulation Comparing of Metirc

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2320, Copyright ©2025
Submitted : November 22, 2024, Revised : December 13, 2024, Accepted : December 18, 2024, Published : January 31,2025

IV. CONCLUSION

From the research conducted, several important conclusions were obtained, including:

1. The hospital HGOD Discovery Model is becoming increasingly important given the rapid development of health services in general and hospital services in particular. HGOD in hospitals refers to the balancing and oversight mechanisms that shape the process of making important decisions in serving Public Health.
2. This method performs simulations with 2 aspects, namely positive (+) and negative (-) with the result that the value 0.000 is the highest value and the value 1 is the lowest value. Above the image of the results of 0.000 is the highest value or low risk and 1 is the lowest or high risk value. The results of the equation above provide a new picture or perspective on hospital, how an intelligent business is able to provide priorities for what will be done.
3. The identified gaps show that there are 4 focal points that are used as a comparison with other methods, namely WT (Wait Time), LS (Length of Stay), CO (Cost), and DD (Door to Doctor). In these 4 points, each method has its own characteristics that you can see in the table above for comparison.

ACKNOWLEDGMENT

We extend our appreciation to the management and staff of regional hospital who provided valuable insights and data throughout the research. Their commitment to improving healthcare services through the use of modern business intelligence techniques was instrumental in shaping the outcomes of this research

REFERENCES

[1] Abdulsattar A.H, (2022), Learning Methods of Business Intelligence and Group Related Diagnostics on Patient Management by Using Artificial Dynamic System, Journal of Nanomaterials, Volume 2022

[2] Riallen L, Andrew C, (2016), Novel function discovery through sequence and structural data mining, Current Opinion in Structure Biology, Volume 38, Juni, Page 53-61

[3] Celia M, Q Ramos, (2022), Business Intelligence Approach and Sentiment Analysis as a Management Strategy Applied to Study Customer Satisfaction in the Hospitality Sector, Advances in Tourism, Technology and System, page 537-547

[4] Cheng Che S, Chang Rey A, Hsu Ching J, Chang I C, (2017), How business intelligence maturity enabling hospital agility, Telematics and Informatics, Volume 34, Issue 1, page 450-456

[5] Escher A, Hainch N, Boll D, (2016), Business Intelligence in Hospital Management, Radiology Managament, europepmc, page 47

[6] Hao Yun kao et al, (2016), Design and evaluation of hospital-based business intelligence system (HBIS): A foundation for design science research methodology, Computer in Human Behavior, Volume 62, page 495-505

[7] Tofighi S dkk, (2014), The Survey of Business Intelligence in the Baqiyatallah Hospital: Our Experience and Literature Review, Intenational Journal of Medicine Reviews, Tehran, 2014

[8] Jaou L, Braga J, Santos F.M, (2021), Adaptive Business Intelligence platform and its contribution as a support in the evolution of Hospital 4.0, Procedia Computer Science, Volume 184, page 905-910

[9] Rouhani, S., Ashrafi, A., Zare Ravasan, A. and Afshari, S. (2016), The impact model of business intelligence on decision support and organizational benefits, Journal of Enterprise Information Management, Vol. 29 No. 1, pp. 19-50

[10] Meza S, Ricky A, Derisma, (2017), Pengembangan Model Business Intelligence Manajemen Rumah Sakit untuk Peningkatan Mutu Pelayanan (Studi Kasus : Semen Padang Hospital), Jurnal Edukasi dan Penelitian Informatika (JEPIN), Volume 3, No 2,

[11] Zheng W, Jim Wu YC, Chen L, (2018), Business intelligence for patient-centeredness: A systematic review, Journal Telematics and Informatics, Volume 35, Issue 4, Page 665-676

[12] Hsu PC, Huang WN, Kuo KM, Yeh YT, (2022), Application of Business Intelligence in Decision Support to Hospital Management: An Example of Outpatient Clinic Schedule Arrangement, Europepmc, Page 1050-1051

[13] L. Hong, M. Luo, R. Wang, P. Lu, W. Lu, and L. Lu, (2019), Big data in health care: applications and challenges, Data and Information Management, vol. 2, no. 3, page. 175–197

[14] Naderinejad M, Jafar Tarokh M, Poorebrahimi A (2014), Recognition and Ranking Critical Success Factors of Business Intelligence in Hospitals Case Study: Hasheminejad Hospital, arvix computer science, 1405.4597.

[15] Kao Hao Y, Yu Min C, Masud M, Wu Wen H, Chen Li J, Wu Yen C, (2016), Design and evaluation of hospital-based business intelligence system (HBIS): A foundation for design science researchmethodology, Computers in Human Behavior, Volume 62, page 495-405

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2320, Copyright ©2025
Submitted : November 22, 2024, Revised : December 13, 2024, Accepted : December 18, 2024, Published : January 31,2025

ISO Technology Analysis with Extended TAM : A Case Study in PT Ebako Nusantara

Nindhia Hutagaol[1]*, Faizal Asrul Pasaribu[2], Jumadi Simangunsong[3]
Department of Informatic Engineering [1], [3]
Department of Information System [2]
Universitas Asa Indonesia
Jakarta, Indonesia
nindhia.hutagaol@asaindo.ac.id[1],faizal@asaindo.ac.id [2],jumadi@asaindo.ac.id [3]

***Abstract*— Implementing information systems in manufacturing aims to improve operational efficiency and service quality. This study evaluates the acceptance of Internal Service Order (ISO) applications with the Extended Technology Acceptance Model (ETAM), that adds variables such as information quality, system quality and user habits. PLS-SEM analysis of 31 respondents found that user habits significantly influenced perceived ease of use (77.8%), which in turn influenced perceived usefulness (86.4%) and user attitude (62%). However, information quality did not significantly influence user habits, suggesting a need to improve information detail. These findings will help develop a TAM model for Indonesian furniture companies. The study recommends the improvement of information accuracy, the development of real-time notification features, and user training to increase adoption and operational efficiency This study provides guidance for organizations to optimize the application of technology in manufacturing operations.**

Keywords*— *ETAM, Internal Service Order, PLS-SEM, Manufacturing, Technology Acceptance.

I. INTRODUCTION

The increasing growth of technology in today's competitive industrial market encourages institutions to adopt innovative applications in various sectors, including the manufacturing industry. In recent years, many studies have explored adopting and accepting information technology applications in different business contexts using the Technology Acceptance Model (TAM). Chotijah & Retrialisca studied user readiness in adopting technology in the furniture sector in the Java area and found that the perception of inconvenience and unsafe factors from a user is an important issue faced in adopting technology [1]. This approach was also utilized in the furniture industry in Jepara, Indonesia to explore how users accept technology to improve operational efficiency and local competitiveness. By using TAM dimensions, this study shows how new technology can be accepted by users if it is able to deliver tangible benefits in terms of improved operational efficiency and convenience, such as real-time tracking and an intuitive user interface[2]. This research strengthens the validity of TAM in the acceptance of Distributed Ledger Technology in manufacturing which identifies that ease of use of technology (PEOU) significantly influences usability and attitude towards technology[3]. Similarly, this study found that technology usefulness is the most influential factor affecting the intention to use systems such as an Enterprise Resource Planning (ERP) system in apparel manufacturing in Java, reinforcing the importance of technology benefits to users[4]. In the same context, Arif evaluated the manufacturing module of an ERP system in an aluminium company and identified usability and ease of use as the main drivers that influence users' intention to adopt new technology. Although user-perceived risk was not significant to usage intention, this research demonstrates that managing operational risk is a key driver of digital transformation success.[5].

Ultimately, the furniture industry believes that technology can help streamline internal processes, such as operational service management, which can generate significant profits, reduce costs, and improve efficiency and decision-making. One company that has used technology in its business processes is PT Ebako Nusantara, which is involved in the cross-country export business of wooden furniture manufacturing. The technology used is the Internal Service Order (ISO) application. One of the technological innovations that optimize operations and improve services to be efficient is the Internal Service Order (ISO) application[6]. This application became a necessary tool in a furniture manufacturing company to facilitate the management of requests in internal units and improve operational workflow. The manufacturing industry's complex production processes and diverse operational needs create a complex environment to understand and adapt to the use of Internal Service Order (ISO) applications. Streamlining internal processes such as operational service management in the furniture manufacturing industry can bring significant benefits, reducing costs and improving efficiency. More importantly, it can improve decision-making and provide a more confident outlook for the future. In this case, the important role of organizational culture, employee training and education as well as management support, becomes a collaborative process to support the acceptance of this technology in the manufacturing business environment.

This study aimed to analyze the acceptance of the Internal Service Order (ISO) application in a wooden furniture

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2231, Copyright ©2025
Submitted : January 6, 2025, Revised : January 16, 2025, Accepted : January 24, 2025, Published : January 31, 2025

manufacturing company, PT Ebako Nusantara, using an extended TAM framework. In this case, the researcher presents insights into effective technology adoption strategies in the manufacturing sector. Hopefully, the findings can contribute to the development of the Technology Acceptance Model and suggest ways to improve technology adoption in manufacturing.

II. LITERATURE REVIEW

In the industrial world, the pressures of globalization, business competition, and the need for high efficiency are driving various industries to innovate and improve operational performance. The furniture industry, a dynamic sector, is a prime example of this adaptability. It constantly evolves, quickly adapting to consumer preferences, technological innovations, and market challenges. As furniture trends follow lifestyle and interior design, manufacturers demonstrate their resilience by adapting to remain competitive. For example, innovations in materials and production technologies, such as automated technologies and 3D printing, enable greater efficiency and flexibility in responding to changing market demands [7]. The development of textile production and tracking systems that automate processes and improve data management [8] is another example of this adaptability. In addition, the implementation of technology such as the Internal Service Order (ISO) application makes it easier for various departments to perform operational services with mutual integration [6] , including one of them in manufacturing. This Internal Service Order is one of the important components in the operational framework used to manage and track internal service requests within an organization such as maintenance, logistics, and quality control. The process in this type of system is designed to ensure that every service required during production runs according to schedule, thus meeting production targets on time and efficiently [9]. Effective task scheduling in ISO is key to maintaining smooth operations and minimizing delays affecting overall productivity.

In its utilization, the Internal Service Order (ISO) application applied to manufacturing companies requires evaluation to understand the extent to which employees accept and implement this technology. The evaluation of application usage is not just a formality, but a crucial step to optimize operational efficiency, improve product quality, and increase customer satisfaction. The evaluation process is conducted in various ways from design to production management and customer relations. The importance of systematic evaluation cannot be overstated, as it helps companies identify strengths, weaknesses, and areas for improvement, thereby encouraging renewal and adaptation to evolving market dynamics [10].

One of the relevant methods for assessing the acceptance of this technology is the traditional TAM. It is commonly known as a framework used to identify the factors influencing how individuals intend to use and behave when using new technology. Furthermore, this model is even used in developing innovative products in the manufacturing industry, where perceived ease of use is highly influential on usability[11]. This relationship confirms the important role of user interface and training in increasing technology acceptance. In her research, Christine extended this method to include external factors relevant to the business context [12]. In furniture manufacturing, the study showed that factors such as supplier support, management efficiency, and facility conditions strongly influence the acceptance and use of ISO applications. Meanwhile, from a managerial perspective, Jacky [13] extended this model. He found that the management perspective has not been explored in the manufacturing industries, although this role is essential in applying technology in all sectors. Another study [5] modified this model to assess digital transformation in aluminum companies by adding the risk perception of technology adoption. Other factors [14], such as system quality, also need to be integrated into the TAM approach to understand organizations' sustainability intentions when using technology services. This researcher found it to play a role in improving users' perceptions of technology services.

A. Model Technology Acceptance (TAM)

This model is one of the widely used theoretical frameworks to measure how new technology is accepted with two main drivers, i.e., the usefulness and ease of use of technology (Fig. 1). Perceived usefulness (PU) refers to a high level of user confidence that technology can improve performance or productivity, and ease of use (PEU) to determine user perceptions of the ease in using technology. Both factors are believed to influence user behavior in using technology, thus creating interest in accepting and using it.

Technology Acceptance Modelling (TAM) is not just a theoretical construct, but a practical tool that can be used to predict the future application of technology. It has been successfully applied in various fields, making it a valuable asset for understanding and predicting technology adoption [15].

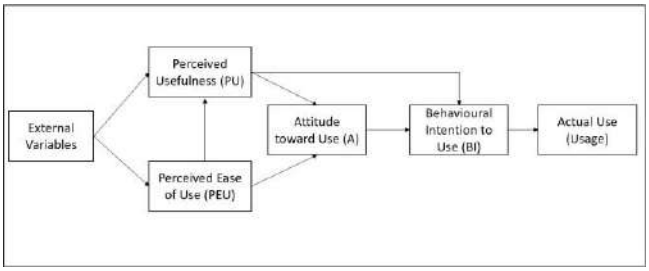


Fig. 1. Framework TAM [15]

Several researchers have used TAM to study technology acceptance in various fields. For example, Hu used it to study the development of Digital Trade technology among Chinese manufacturers, while Cao applied it to understand the acceptance of automated manufacturing technology in China [16] [17].

A study by Laumer et al.[18], found that information quality plays an important role in determining system user satisfaction. Information quality, which includes completeness, relevance, timeliness, and usefulness, has a greater impact. In addition, system quality (complexity, flexibility, navigation, and reliability) has the greatest impact on user satisfaction. This research confirms that information quality, particularly task

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2231, Copyright ©2025
Submitted : January 6, 2025, Revised : January 16, 2025, Accepted : January 24, 2025, Published : January 31, 2025

context, should be an organizational priority to improve user satisfaction, prevent workarounds and maximize the benefits of information systems. The study analyzed acceptance using a modified TAM with cultural dimensions and found that the cultural dimension 'uncertainty avoidance' has a significantly influence on perceived usefulness.[19]. High levels of uncertainty avoidance tend to favour structure, regulation and routine to reduce uncertainty in the work environment. high levels of uncertainty avoidance see systems as useful tools if they believe they can provide predictability and stability in day-to-day operations. Another study also used this concept to comprehend the enablers of acceptance of one telemedicine service from 275 participants and found that one of the factors that increased acceptance of this technology was the level of trust in the technology service. [20]. In the educational environment, Lukman applied this method to evaluate the academic system of a private institution. He found that attitudes towards system use significantly impacted technology acceptance [21]. Lenni found in her research that as long as digital technology is easy to use, it will tend to be adopted, especially in tourism. The study also highlights the importance of innovation in driving technology adoption, and shows that easy usage directly increases people's intention to use the technology [22]. This study extends the traditional TAM by including external variables such as perceived mobility, personal habits, and perceived experience with financial services technology. They found that personal habits have a stronger direct influence on technology adoption intentions than perceived mobility[23].

B. Information System Success Model (ISSM)

The ISSM model, a practical and influential tool in evaluating the success of information systems or technology, has been in use since its inception in 1992 [24]. Initially, the model comprised six key dimensions as shown in Fig.2 including information quality, system quality, system usage, user satisfaction, individual impact, and organizational impact. Over the following decade, further research [25], led to the addition of two new dimensions by DeLone: system service quality (Service Quality) and user interest in applying the system or technology (Intention to Use). These additions combined individual and organizational influences into a tangible benefit dimension (Net Benefit).

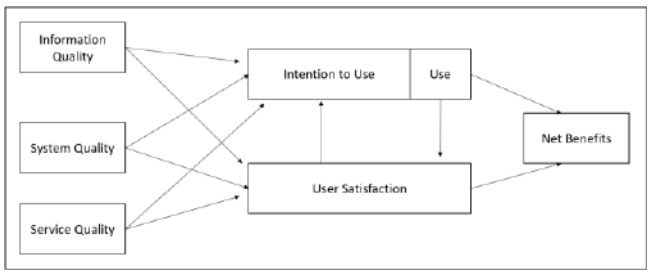


Fig. 2. D&M Information Success Model [25]

Previous researchers have used the ISSM model framework extensively to measure and identify factors affecting technology application acceptance. For example, the model was used to examine a mobile-based academic information system usage in education. According to this study, information quality, system quality, and perceived usefulness influenced users to continue using the technology. Meanwhile, service quality affects perceived usefulness but otherwise has no effect on the level of ease of use[26]. Similarly, in Ashfaq's research, this model was used and found that quality system information and services are very important to improve user experience and encourage the adoption of AI-based chatbot e-service technology[27]. This practical application underscores the relevance of this model in real-world scenarios.

C. Partial Least Squares Methods (PLS)

The PLS method is a commonly used analysis tool due to its ability to provide reliable evaluation results and work efficiently with small data sample sizes. Its unique feature is the ease with which it handles reflective and formative measurement models, making it applicable in a variety of research situations [28]. The PLS method works with two sets of components: a predictor variable component and a response variable component. Each predictor variable predicts the response variable construct by analyzing the data pattern, which indicates the most relevant relationship between the independent and dependent variables.

In addition to maximizing the correlation between predictor and response variables, PLS reduces the dimensionality of the data. This reduction results in a more straightforward and understandable model, providing ease and clarity in the analysis process. [29].

From the literature review conducted, this research will combine the TAM and ISSM approaches as an Extended Technology Acceptance Model (ETAM) by integrating external variables such as information quality, system quality, and user habits into the implementation of ISO applications. This additional dimension is to analyze the factors that influence technology acceptance, so that users' understanding is more comprehensive regarding the relationship between user perceptions and actual technology use. This approach not only considers the influence of technology convenience and usability but also information and quality systems and user habits, which significantly impact technology adoption. Using Partial Least Squares (PLS) analysis, this study will show the impact of these three factors and provide insights to identify areas for improvement and optimize the success of technology adoption.

III. RESEARCH METHOD

This study analyses and identifies the factors that shape and influence user interest in using the Internal Service Order (ISO) application. Based on the review of literature studies that have been conducted, three external variables will be added to the TAM model framework. These variables are adapted from the ISSM model: information quality, system quality, and user culture (habit) variables are added.

As shown in Figure 3, these three external variables along with other variables, namely the usefulness of technology, ease of use of technology, user attitudes towards technology, user intention to apply technology affect the dependent variable, namely the actual usage variable.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2231, Copyright ©2025
Submitted : January 6, 2025, Revised : January 16, 2025, Accepted : January 24, 2025, Published : January 31, 2025

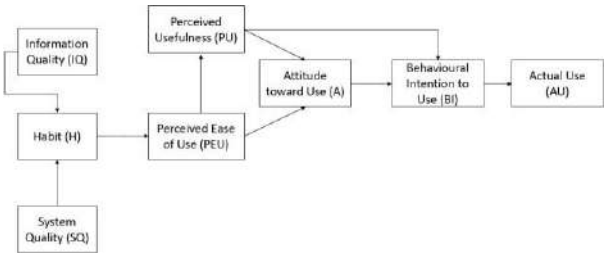


Fig. 3. Extended Technology Acceptance Model

a) Information Quality (IQ) is a crucial factor that determines how information is received and perceived by users when using a system or technology. In this model, it is not just assumed, but it is proven that information quality significantly impacts the system quality (SQ) of the Internal Service Order (ISO) application service.
***H1**: Information quality (IQ) positively and significantly impacts user habits (H) of the Internal Service Order application.*

b) System Quality (SQ), which represents the ability of the system or technology to perform interactions or operations, plays a critical role in the performance of information systems. This is an important measure for information system performance.
***H2**: System quality (SQ) is not just a factor, but a potential game-changer that impacts user habits (H) in the application of the Internal Service Order application.*

c) User habit (H) has a cognitive effect on behavioral interest and are proven that they significantly impact the ease of use factor[30]. This is a vital aspect for user behavior in the Internal Service Order (ISO) application.
***H3**: User habits (H) positively significantly impact the perceived ease of using the Internal Service Order application. (PEU)*

d) Perceived Ease of Use (PEU) determines the level of user confidence in the ease of use of technology and this factor contributes to improving user performance.
***H4**: PEU significantly affects the perceived usefulness of the Internal Service Order application (PU).*
***H5**: PEU significantly positively affects user attitudes while adapting the Internal Service Order application (A).*

e) Perceived Usefulness describes user confidence where the Internal Service Order application provides benefits during use.
***H6**: Perceived Usefulness (PU) positively and significantly affects user interest (IU) in the Internal Service Order application.*
***H7**: Perceived Usefulness (PU) significantly impacts user attitudes (A) to continue using the Internal Service Order application.*

f) Individual attitude towards using the system / Attitude towards Use (A) is essential in predicting users' behavioral interest in the system or technology services.
***H8**: A significantly affects user behavioral interest in using Internal Service Order (IU) services.*

g) The behavioral interest factor using the system / Behavioral Intention to Use (IU) is used as an indicator to measure the user's desire to take a specific action.
***H9** : User behavioral interest (IU) strongly influences the use of Internal Service Order (AU) services.*

h) Actual Use (AU) of Internal Service Order services in this case is affected by ease of use (PEU), usefulness (PU), individual attitudes (A) and behavioral interest (BI) of users towards the use of Internal Service Order services.

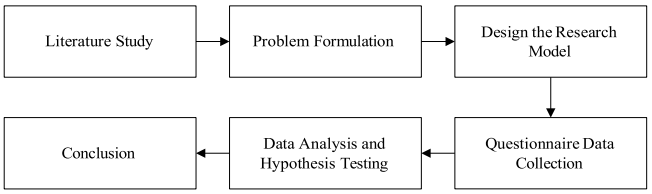


Fig. 4. Research Method

The systematic stages of this research were carried out in a structured manner (Figure 4), including:

1) Literature Study
   Researchers meticulously traced literature or previous research related to the research topic. This comprehensive stage provides a deep understanding of the context of the problem and the methods used in the study, ensuring the audience is well-informed.
2) Problem Formulation
   The researcher plays a pivotal role in guiding the research by formulating a problem that will find its solution at the end of the research. This stage sets the focus and direction for the study.
3) Design the Research Model
   Researchers design research models that describe the relation among variables or factors that have impact on the acceptance of Internal Service Order application technology. This stage helps the reader understand how the research is structured.
4) Questionnaire Data Collection
   At this stage, researchers compile questionnaires and document the results of participants' answers to be processed and analyzed at the next stage.
5) Data Analysis and Hypothesis Testing
   This process involves analysis of participant data and statistical test analysis using quantitative methods.
6) Conclusion
   Researchers explain the findings and relate the findings to what researchers have found before.

This study was conducted at PT Ebako Nusantara, which utilized the Internal Service Order (ISO) application by conducting a survey using a questionnaire instrument. Data collection began in September - October 2024, involving 31 employees from all units who actively applied for production machine repair services as research respondents. The questionnaire was designed using the Google Form feature and distributed online to respondents. the proposed model is examined by analyzing the entire data using the PLS method

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2231, Copyright ©2025
Submitted : January 6, 2025, Revised : January 16, 2025, Accepted : January 24, 2025, Published : January 31, 2025

and validate the statistically determined factors to test the research hypothesis. The data analysis process was conducted using SmartPLS software by extracting respondents' answers based on the modified instrument developed in previous research [10].

The questionnaire chart includes two parts of the instrument: the first demographic characteristics containing respondent information such as gender, age, and work unit. The next part consists of question items, where respondents will specify their level of experience according to a 5-point Likert scale with following questionnaires (Table 1) items below:

TABLE I. CONSTRUCT ITEMS

Variables	Statements
Information Quality (IQ)	1. The information available on this application helps me complete my work more efficiently
	2. The information available is always accurate
	3. The information available is complete and contains everything I need
	4. Information is constantly updated on time
Habit (H)	1. I use this application regularly to complete tasks
	2. I am comfortable and used to using this application.
	3. I believe this application can be used to improve collaboration between teams
	4. This application will be helpful for future sustainability
System Quality (SQ)	1. This application is easy to access and use at any time.
	2. This application rarely experiences interruptions or system problems.
	3. This application works quickly and efficiently when it's needed.
Perceived usefulness (PU)	1. This application can increase company productivity.
	2. This application can improve the quality of operational services.
	3. This application is useful for maintaining work performance faster.
	4. This application can reduce operational costs.
Perceived ease of use (PEU)	1. I can easily interact with the ISO application
	2. I believe this application is easy to learn and use
Attitude toward Use (A)	1. I feel this application is handy for me
	2. It is enjoyable for me to use this application
Intention to use (IU)	1. I am interested in using this application often in the future
	2. I like using this application because it fits the prevailing culture in my work environment
Actual use (AU)	1. My unit has actively used this application.
	2. My unit has started using this application for operational services.
	3. The company has substantially used this application in manufacturing.

IV. RESULT AND ANALYSIS

The research data are obtained from a survey by distributing questionnaires to employees of the Maintenance Division and Production Division at the Ebako company, a total of 31 people who directly used the Internal Service Order (ISO) application. This unique perspective provides a detailed understanding of the application's usage. Table 2 shows the distribution of demographic data and characteristics of respondents. The fact that 68% of the respondents were male is significant, as it defines that male workers mostly use ISO applications in the manufacturing work environment. This is possible because this research case study is specifically a unit involving technical work where the study found a very high level of male participation in this domain [31]. This finding could have implications for gender diversity in technical roles within the manufacturing industry.

The ISO technology adaptation process can be categorized as still having potential because most of the respondents are of productive age (31-50 years old) although the involvement of young age groups is minimal. In addition, the limited academic background also affects the respondents' ability to understand and use the application optimally. From the respondents' last level of education, 61% of them mostly have a high school education, which reflects that the proportion of respondents with higher education is also very minimal. This indicates the need for a competency improvement strategy, especially technological literacy, to support better and optimal adaptation. This strategy could involve targeted training programs or educational initiatives to enhance the technical skills of the workforce.

The data in Table 2 shows that the Internal Service Order (ISO) application has been functioning and running at the operational level as seen by 90% of employees at the technician level who have actively used this technology. However, managerial-level participation can also be an issue that hinders strategic supervision if their role is minimal towards the application. This means that if managers are not actively involved in the use and oversight of the ISO application, it could lead to a lack of strategic direction and control over its implementation and use.

TABLE II. USER'S DEMOGRAPHICS

Characteristics	Maintenance Unit	Production Unit	Total	Percentage
Gender				
Male	16	5	21	68%
Femal		10	10	32%
Age				
17 – 30	3	5	8	26%
31 – 40	6	5	11	35%
41 – 50	6	5	11	35%
51 – 60	1		1	3%
Last Education				
Senior High School equivalent	9	10	19	61%
Diploma	1	4	5	16%
Bachelor	4	1	5	16%
Master	2		2	6%
Job Level				
Manager	2	1	3	10%
Officer / Technician	14	14	28	90%

In the analysis process, all respondent data is processed using SmartPLS to test data validity and data reliability. From the data in Table 3, all variable indicators were found with a Cronbach's value > 0.60 and a correlation between statement items ≥ 0.7. From these findings, each questionnaire statement item in this study was concluded to be reliable for use as a research instrument. The composite reliability test value (CR) also proves and strengthens the test results that all variables in the model meet the validity criteria seen from the correlation

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2231, Copyright ©2025
Submitted : January 6, 2025, Revised : January 16, 2025, Accepted : January 24, 2025, Published : January 31, 2025

value in Table 4.

TABLE III. VALIDITY AND RELIABILITY RESULTS

Indicator	Average Score	Item Correlation	Result	Cronbach Alpha	Result
IQ1	4.516	0.785	valid	0.889	Reliable
IQ2	4.129	0.940	valid		
IQ3	4.000	0.886	valid		
IQ4	3.935	0.852	valid		
H1	4.335	0.857	valid	0.838	Reliable
H2	4.290	0.846	valid		
H3	4.387	0.875	valid		
H4	4.290	0.701	valid		
SQ1	4.129	0.805	valid	0.781	Reliable
SQ2	3.935	0.795	valid		
SQ3	4.258	0.891	valid		
PU1	4.452	0.872	valid	0.912	Reliable
PU2	4.387	0.938	valid		
PU3	4.452	0.926	valid		
PU4	4.258	0.820	valid		
PEU1	4.484	0.918	valid	0.827	Reliable
PEU2	4.419	0.929	valid		
A1	4.419	0.939	valid	0.870	Reliable
A2	4.226	0.943	valid		
IU1	4.355	0.970	valid	0.935	Reliable
IU2	4.129	0.967	valid		
AU1	4.226	0.917	valid	0.836	Reliable
AU2	4.194	0.929	Valid		
AU3	4.129	0.753	Valid		

In Table 4, respondents' feedback on the adaptation of the use of the Internal Service Order (ISO) application from each variable, namely attitude towards system use (A), actual use (AU), Habit (H), information quality (IQ), user intention (IU), perceived ease of system use (PEU), perceived usefulness (PU), and system quality (SQ) shows that the composite reliability estimate results obtained from the bootstrapping process are stable and consistent with each variable value <0.05.

TABLE IV. COMPOSITE RELIABILITY (CR)

Variable	Composite Reliability Value
Actual Use (AU)	0.903
Attitude towards Use (A)	0.939
Habit (H)	0.893
Information Quality (IQ)	0.924
User Intention (IU)	0.968
Perceived Ease of Use (PEU)	0.920
Perceived Usefulness (PU)	0.938
System Quality (SQ)	0.870

TABLE V. DISCRIMINANT VALIDITY (DV)

Var	Discriminant Validity							
	AU	A	H	IQ	IU	PEU	PU	SQ
AU	0.870							
A	0.724	0.941						
H	0.695	0.728	0.823					
IQ	0.582	0.798	0.781	0.868				
IU	0.732	0.933	0.764	0.842	0.969			
PEU	0.773	0.808	0.778	0.721	0.811	0.923		
PU	0.741	0.752	0.769	0.714	0.791	0.864	0.890	
SQ	0.571	0.744	0.799	0.813	0.764	0.736	0.714	0.832

The empirical research model analysis results by applying the Partial Least Square (PLS) approach are shown in Figure 5 and Table 5. Figure 5 shows the correlation between variables, with each indicator representing latent variables. In the model used, three external factors will be analyzed and tested for their influence on the acceptance of ISO application technology at PT Ebako Nusantara: information quality factors on the system, system quality factors, and user habits. Each factor is positively correlated with indicators of perceived ease of use of the system. Several hypotheses in the research model (Figure 3) show that eight of the nine hypotheses tested have a significant impact with a p-value <0.05.

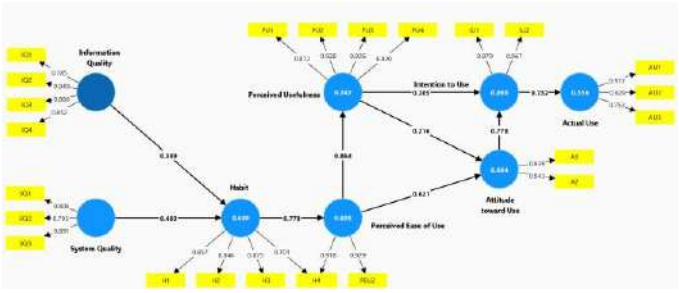


Fig. 5. Correlation Variables in extended TAM

First, hypothesis H1 (IQ → H) is rejected as it shows information quality does not impact on user habits (p-value 0.062) even though information quality (IQ) is positively correlated to application user habits (0.389). According to a survey conducted on all ISO application user employees, 19% of employees still feel inaccurate information, such as unclear details and images of product damage, notification problems, and lack of information on the progress of ongoing repairs. This allows the influence of the information quality factor to be relatively low for users as a reason to get used to using the ISO application. On the other hand, the p-value of hypothesis H2 (SQ → H) is 0.020, which indicates there is a significant effect of system quality on user habit. This finding strengthens the hypothesis that ISO application quality factors directly contribute to improving user habits. The p-value of H2 is also in line with the positive correlation between the system quality variable and the user habit factor (0.483), meaning that the Internal Service Order (ISO) application is included in the system quality factor, including reliability both from the stability, speed, and security of the system to encourage each employee to use the system more often and habitually.

The data testing results also prove the hypothesis H3 (H→PEU) has a significant effect (coefficient value 0.778, p-value 0.000), so it is concluded that user habits (H) have a dramatic effect on user experience when adapting to new technology. In this case, employees who are accustomed to using ISO applications will find it easier to use the application when used on an ongoing basis. This user habit will ultimately affect the ease of use and automatically also affect the perceived usefulness of the application (coefficient value 0.864 with t-statistic 17.474 and p-value 0.000) and even user attitudes (coefficient value 0.621, t-statistic = 2.481, p-value = 0.013). This user habit may also result from encouragement from the work environment, including people who are considered important in the organization, to adopt the technology[32]. This means that in addition to the application being perceived as easy

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2231, Copyright ©2025
Submitted : January 6, 2025, Revised : January 16, 2025, Accepted : January 24, 2025, Published : January 31, 2025

to use, the role of superiors or colleagues who initiate the use of this application can also affect the perceived ease of use and usefulness of technology. These results support that hypotheses H4 (PEU → PU) and H5 (PEU→A) are also valid where ease of use positively affects application usability and user attitude.

However, the results of testing the perceived usefulness factor in hypotheses H6 (PU→IU) and H7 (PU→A) reveal a different effect. The perceived usefulness factor has a significant but relatively low impact, with a coefficient value of 0.205 and a p-value of 0.014. This suggests that perceived usefulness has a smaller influence on user intentions and does not even show significance on user attitudes (coefficient 0.216, t-statistic = 0.947, p-value = 0.344), leading to the rejection of H7.

TABLE VI. DIRECT EFFECT ANALYSIS RESULTS

Hypothesis	*Coefficient*	t-statistic	p-value	Results
H1 : IQ→H	0.389	1.873	0.062	rejected
H2 : SQ→H	0.483	2.337	0.020	accepted
H3 : H→ PEU	0.778	10.131	0.000	accepted
H4 : PEU→PU	0.864	17.474	0.000	accepted
H5 : PEU→A	0.621	2.481	0.013	accepted
H6 : PU→IU	0.205	2.453	0.014	accepted
H7 : PU→A	0.216	0.947	0.344	rejected
H8 : A→ IU	0.778	9.291	0.000	accepted
H9 : IU→AU	0.732	7.854	0.000	accepted

This implies that users' attitudes towards technology are influenced not only by perceived usefulness but also by other variables such as trust in the system, user satisfaction, and previous interaction experience. In a manufacturing environment, the benefits of technology (such as increased productivity) may not be perceived directly by individuals, but rather at an organizational level. This may mean that perceived benefits are not directly related to individual attitudes towards the technology.

Meanwhile, the results of hypotheses H8 (A→IU) and H9 (IU→AU) are both significant, with coefficients of 0.778 and 0.732 respectively (p-value 0.000). These results indicate that the user attitude factor has a strong influence on usage intention, and user intention directly affects the actual usage of the application. This analysis aligns with the Technology Acceptance Model modeling concept, which defines positive user attitudes toward technology as driving the technology adoption process.

In conclusion, the three external variables added to the traditional TAM significantly influence Internal Service Order (ISO) application usage. The application is proven to be used on an ongoing basis, especially in the Production Division and Maintenance Division at PT Ebako Nusantara, which significantly improves the performance of each employee involved. The success of these factors affects the perceived ease of adopting the ISO application, thereby triggering the user's intention to continue using the application. Notably, ISO application users at PT Ebako Nusantara are predominantly men aged 31-50, with an average high school education. Given this demographic, it is crucial for the company to design an intuitive and user-friendly interface for users over 40 and to simplify the features to accommodate their potentially limited technical understanding. This will support more optimal adaptation and enhance the application's usability.

The high and low quality of information in ISO applications is measured by relevance, accuracy, and availability. This application has good information quality but does not affect user habits. This is due to differences in user education levels that affect user understanding of the information in the application, as well as the availability of some information that has not been detailed to answer user needs so that it tends to affect the confidence of some users in the usefulness of this ISO application. This possibility can be seen from the IQ4 statement in Table 3 (mean = 3.935), which received a relatively low response, indicating that users feel the information available is not correct. This fact was also found in previous research [33], which explains that quality information on a system becomes a factor that significantly affects users with high experience and sufficient educational competence. If the quality of information (in this case, such as repair work time, complete detailed information about machine or device damage in production operations, and other important information) is accurate and relevant, it can increase the perception that technology helps maximize user performance.

On the other hand, the highest statement response (mean = 4.484) is that it is easy for users to interact with the ISO application, indicating that the design of the application design is adequate so that users can easily interact with this application. However, it is necessary to consider evaluating the technical infrastructure to speed up the application response time. In addition, based on the survey 25% of employees suggested the addition of several features such as the provision of notifications and the progress feature of the ongoing repair process where these two features help provide information for employees to achieve the completion of work targets. Overall, this study's results answer hypotheses in the proposed ETAM model, a model that extends the traditional TAM by adding three additional variables that are positively correlated and influential to the TAM model. However, the insignificant hypothesis H7 indicates the relationship between perceived usefulness and user attitude requires further analysis in the future by considering additional variables such as trust. Perceived usefulness and satisfaction need to be considered from the user's point of view to find out the benefits that are directly felt by the users one of which is affected by the quality of information. This will increase the user's trust that the system provides significant value, thus increasing the user's use of the system.

The factors analyzed include information quality, system quality and user habits, but do not consider other factors such as confidence in the technology, management support or individual motivation. As a result, the model does not fully capture the complex dynamics of technology adoption in a manufacturing environment. This study was also limited to one manufacturing company in Indonesia, so the results are highly contextual and may not apply to other companies with different organizational structures, work cultures or technologies in other regions. In addition, the small sample size of 31 respondents also reduces the ability to generalize the results to a wider population.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2231, Copyright ©2025
Submitted : January 6, 2025, Revised : January 16, 2025, Accepted : January 24, 2025, Published : January 31, 2025

The adoption ISO applications in manufacturing can be a solution that improves operational performance. This is achieved by ensuring the quality of the information and systems developed and, importantly, by encouraging employees to use the system through regular and continuous training. This training helps build user habits when applying the technology, improving the overall user experience and system adoption.

V. CONCLUSION

This research goal is to extend and evaluate a robust model that empirically elucidates how information quality factors in a system, system quality factors, and user habit factors influence user perspectives on system use. This model, we believe, has the potential to significantly influence user adoption of a system in a sustainable manner, thereby making a substantial contribution to the field of information systems and technology adoption.

System quality positively influences user habits in using technology to determine its usefulness. On the other hand, the results show that the quality of information does not significantly affect users' habits of using the application when the quality of the system is simple and makes it easy for users with a wide age range to get used to new technology. In this case, the external factor of user habits is a good predictor in measuring the ease of use manufacturing systems. This factor becomes very important if the quality of the qualified system supports it. The Internal Service Order (ISO) application is still a new system in PT Ebako Nusantara, so several functions and features still require users to perform new tasks, such as monitoring the repair of production machines.

Some specific suggestions that are applied to improve the quality of information in ISO applications include improving the accuracy of information by adding detailed descriptions to machine damage information. Details include not only images, but also full technical explanations so that users can properly understand the condition of the damage. In addition, repair status updates should be provided in real time to increase transparency in the repair process. This can be supported by implementing automatic notification features such as repair deadlines and the status of completed or ongoing repairs. User training also plays a key role in improving application usage. This should be carried out to improve technological literacy, especially for employees with lower levels of education. On the other hand, to ensure that the application remains relevant to user needs, regular evaluation by the management team through surveys or periodic discussions is also necessary. This allows organisations to identify missing or missing features in the app, which can then be incorporated into the development of new features, ensuring that the app continues to evolve as needed. In addition, an analytics dashboard that summarises key data such as average repair time and service request status can help users monitor the performance of the app. In the future, the app can also integrate AI technology to predict potential breakdowns based on usage and repair history, so that preventative action can be taken earlier.

Moreover, it's crucial that application developers review and enhance the system to ensure that ISO applications can effectively meet the needs of their users. This is particularly important in terms of maximizing the system's flexibility, accuracy, and security. Additionally, our proposed model suggests the inclusion of trust factors as new predictors to test system usage and user satisfaction, further underlining the need for continuous system improvements.

REFERENCES

[1] U. Chotijah and F. Retrialisca, "Analysis of Information Technology Readiness in Furniture Business in Indonesia," *Indones. J. Inf. Syst.*, vol. 3, no. 1, pp. 14–22, 2020, doi: 10.24002/ijis.v3i1.3470.

[2] T. Suryawijaya and M. Safitri, "Angkut.in: The Future of Trucking by Digital Product Design Journey," *Ajie*, vol. 08, no. 01, pp. 30–39, 2024, doi: 10.20885/ajie.vol8.iss1.art5.

[3] N. Ullah, W. Mugahed Al-rahmi, and A. Alkhalifah, "Predictors for distributed ledger technology adoption: integrating three traditional adoption theories for manufacturing and service operations," *Prod. Manuf. Res.*, vol. 9, no. 1, pp. 178–205, 2021, doi: 10.1080/21693277.2021.1976963.

[4] A. Syafira, W. Puspitasari, and W. Witjaksono, "Analysis and Evaluation of User Acceptance on the Use of Enterprise Resource Planning (ERP) Systems at PT Trisco Tailored Apparel Manufacturing," *IJAIT (International J. Appl. Inf. Technol.*, vol. 04, no. 01, p. 17, 2021, doi: 10.25124/ijait.v4i01.2852.

[5] A. Saifunnasrullah and K. Budiman, "Digital Transformation Analysis in the Manufacturing Module in Aluminium Companies using the TAM Method," *J. Adv. Inf. Syst. Technol.*, vol. 5, no. 1, pp. 49–63, May 2023, doi: 10.15294/jaist.v5i1.66567.

[6] J. Simangunsong, N. Hutagaol, and F. A. Pasaribu, "Implementasi Aplikasi Internal Service Order (ISO) Berbasis Web pada Perusahaan Manufaktur Furniture," *J. Inform. J. Pengemb. IT*, vol. 9, no. 2, pp. 151–163, 2024, doi: 10.30591/jpit.v9i2.6813.

[7] E. Skorupińska, M. Hitka, and M. Sydor, "Surveying Quality Management Methodologies in Wooden Furniture Production," *Systems*, vol. 12, no. 2, pp. 1–17, 2024, doi: 10.3390/systems12020051.

[8] N. Okyay, S. Kes, F. Işık, and A. T. Uzun, "Development of Smart Software Program for Production and Order Tracking in the Textile Industry," *Eur. J. Res. Dev.*, vol. 4, no. 2, pp. 168–173, 2024, doi: 10.56038/ejrnd.v4i2.449.

[9] G. Ahn and S. Hur, "Multiobjective Real-Time Scheduling of Tasks in Cloud Manufacturing with Genetic Algorithm," *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/1305849.

[10] T. P. Mabotja and Winiswa Mavutha, "Effect of Lean Six Sigma on order fulfilment process: evidence from manufacturing companies in Gauteng, South Africa," *Int. J. Res. Bus. Soc. Sci. (2147- 4478)*, vol. 13, no. 3, pp. 54–65, 2024, doi: 10.20525/ijrbs.v13i3.2952.

[11] M. Hsu, "An Analysis of Intention to Use in Innovative Product Development Model through TAM Model," *EURASIA J. Math. Sci. Technol. Educ.*, vol. 12, no. 3, pp. 487–501, Jun. 2016, doi: 10.12973/eurasia.2016.1229a.

[12] Christine and N. Legowo, "Evaluation of Policy Processing System Using Extended UTAUT Method at General Insurance Company," *Proc. 2018 Int. Conf. Inf. Manag. Technol. ICIMTech 2018*, no. September, pp. 172–177, 2018, doi: 10.1109/ICIMTech.2018.8528113.

[13] J. Chin and S. C. Lin, "A behavioral model of managerial perspectives regarding technology acceptance in building energy management systems," *Sustain.*, vol. 8, no. 7, 2016, doi: 10.3390/su8070641.

[14] M. Ramkumar, T. Schoenherr, S. M. Wagner, and M. Jenamani, "International Journal of Production Economics Q-TAM : A quality technology acceptance model for predicting organizational buyers ' continuance intentions for e-procurement services," *Intern. J. Prod. Econ.*, vol. 216, no. March 2018, pp. 333–348, 2019, doi: 10.1016/j.ijpe.2019.06.003.

[15] M. Turner, B. Kitchenham, P. Brereton, S. Charters, and D. Budgen, "Does the technology acceptance model predict actual use? A systematic literature review," *Inf. Softw. Technol.*, vol. 52, no. 5, pp. 463–479, 2010, doi: 10.1016/j.infsof.2009.11.005.

[16] S. Hu, "A Conceptual Development of Digital Trade Adoption among

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2231, Copyright ©2025
Submitted : January 6, 2025, Revised : January 16, 2025, Accepted : January 24, 2025, Published : January 31, 2025

Manufacturing Companies in China: A Technology Acceptance Model Approach," *J. Digit. Realis. Mastery*, vol. 3, no. 02, pp. 76–80, 2024, doi: 10.56982/dream.v3i02.213.

[17] D. Cao, H. Tao, Y. Wang, A. Tarhini, and S. Xia, "Acceptance of automation manufacturing technology in China: an examination of perceived norm and organizational efficacy," *Prod. Plan. Control*, vol. 31, no. 8, pp. 660–672, 2020, doi: 10.1080/09537287.2019.1669091.

[18] S. Laumer, C. Maier, and T. Weitzel, "Information quality, user satisfaction, and the manifestation of workarounds: A qualitative and quantitative study of enterprise content management system users," *Eur. J. Inf. Syst.*, vol. 26, no. 4, pp. 333–360, 2017, doi: 10.1057/s41303-016-0029-7.

[19] W. Kurniawan and D. N. Utama, "Culture-Based TAM for Analyzing the Implementation of Building Management System," *J. Syst. Manag. Sci.*, vol. 13, no. 6, pp. 368–383, 2023, doi: 10.33168/JSMS.2023.0622.

[20] S. A. Kamal, M. Shafiq, and P. Kakria, "Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM)," *Technol. Soc.*, vol. 60, no. March 2019, 2020, doi: 10.1016/j.techsoc.2019.101212.

[21] L. Reza, S. Sunardi, and H. Herman, "Penilaian Sistem Informasi Akademik Dengan Metode Technology Acceptance Model," *Fountain Informatics J.*, vol. 7, no. 1, p. 31, 2021, doi: 10.21111/fij.v7i1.6393.

[22] L. Yusrini, N. Eviana, Maeenuddin, R. Dorbala, and R. Darmadi, "Understanding the Acceptance of Digital Marketing Among Tour Operators: An Empirical Study," *African J. Hosp. Tour. Leis.*, vol. 13, no. 1, pp. 183–191, 2024, doi: 10.46222/ajhtl.19770720.496.

[23] Y. S. Yen and F. S. Wu, "Predicting the adoption of mobile financial services: The impacts of perceived mobility and personal habit," *Comput. Human Behav.*, vol. 65, pp. 31–42, 2016, doi: 10.1016/j.chb.2016.08.017.

[24] S. Petter, W. Delone, and E. R. McLean, "Information systems success: The quest for the independent variables," *J. Manag. Inf. Syst.*, vol. 29, no. 4, pp. 7–62, 2013, doi: 10.2753/MIS0742-1222290401.

[25] W. H. DeLone and E. R. McLean, "The DeLone and McLean model of information systems success: A ten-year update," *J. Manag. Inf. Syst.*, vol. 19, no. 4, pp. 9–30, 2003, doi: 10.1080/07421222.2003.11045748.

[26] N. Aeni Hidayah, N. Hasanati, R. Novela Putri, K. Fiqry Musa, Z. Nihayah, and A. Muin, "Analysis Using the Technology Acceptance Model (TAM) and DeLone McLean Information System (DM IS) Success Model of AIS Mobile User Acceptance," *2020 8th Int. Conf. Cyber IT Serv. Manag. CITSM 2020*, 2020, doi: 10.1109/CITSM50537.2020.9268859.

[27] M. Ashfaq, J. Yun, S. Yu, and S. M. C. Loureiro, "I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents," *Telemat. Informatics*, vol. 54, no. April, p. 101473, 2020, doi: 10.1016/j.tele.2020.101473.

[28] E. Edeh, W.-J. Lo, and J. Khojasteh, *Review of Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*, vol. 30, no. 1. 2023. doi: 10.1080/10705511.2022.2108813.

[29] K. Kumar, "Partial Least Square (PLS) Analysis," *Resonance*, vol. 26, no. 3, pp. 429–442, 2021, doi: 10.1007/s12045-021-1140-1.

[30] F. Rafique, Hamaad., Shamim, Azra., Anwar, "Investigating Acceprance of Mobile Library Application with Extended Rechnology Acceptance Model (TAM)," *Elzeiver Ltd*, pp. 1–21, 2019.

[31] J. E. Yavorsky and J. Dill, "Unemployment and men's entrance into female-dominated jobs," *Soc. Sci. Res.*, vol. 85, p. 102373, 2020, doi: 10.1016/j.ssresearch.2019.102373.

[32] S. P. Mohanty, A Bhijit;Mishra, "An Investigation on the Application of the Technology Acceptance Model (TAM) to Evaluate the Adoption of ERP Solutions in MSMEs of Odisha," *Int. J. Sci. Res.*, vol. 7, no. 11, pp. 94–101, 2018, doi: 10.21275/ART20192478.

[33] S. Chatterjee, N. P. Rana, Y. K. Dwivedi, and A. M. Baabdullah, "Understanding AI adoption in manufacturing and production firms using an integrated TAM-TOE model," *Technol. Forecast. Soc. Change*, vol. 170, no. November 2020, p. 120880, 2021, doi: 10.1016/j.techfore.2021.120880.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2231, Copyright ©2025
Submitted : January 6, 2025, Revised : January 16, 2025, Accepted : January 24, 2025, Published : January 31, 2025

Comparison of the Performance of the C.45 Algorithm with Naive Bayes in Analyzing Book Borrowing at the Library Pringsewu Muhammadiyah University

Dani Wilian[1]*, Sriyanto[2]
Departmen of Computer Science[1], [2]
Institut Informatika dan Bisnis Darmajaya
Lampung, Indonesia
wilian.2321210009p@mail.darmajaya.ac.id[1], sriyanto@darmajaya.ac.id [2]

***Abstract* — This study examines the effectiveness of the Naïve Bayes and C4.5 algorithms in analyzing book borrowing patterns at the Pringsewu Muhammadiyah University Library. As libraries increasingly serve as vital educational hubs, understanding user borrowing behavior is essential for effective collection management and service enhancement. The research follows the Cross-Industry Standard Process for Data Mining (CRISP-DM), which includes stages of business understanding, data understanding, preparation, modeling, evaluation, and implementation. A dataset consisting of 5,586 records and ten attributes related to book lending was utilized, with comprehensive data cleaning and preprocessing conducted. The performance of both algorithms was assessed using K-fold cross-validation, yielding an accuracy of 96.26% for C4.5, compared to 91.44% for Naïve Bayes. These results demonstrate that C4.5 is more adept at capturing complex relationships within the data, providing deeper insights into user preferences and enhancing library services. This research underscores the potential of data mining techniques to optimize library management and proposes avenues for future investigation, such as exploring advanced machine learning algorithms and expanding datasets for use in broader library contexts.**

Keywords*— *Book Borrowing Patterns, C4.5, Naive Bayes, Datamining

I. INTRODUCTION

In today's digital era, the library does not function as a place to store my books, but also as an important source of information and education for students, lecturers and researchers. With the increasing number of books and other materials available, and the diversity of users' information needs, managing library collections effectively has become a significant challenge[1]. One important aspect in library management is understanding book borrowing patterns by users.

Book borrowing patterns can provide valuable insight into user preferences, collection usage trends, and information needs that may be unmet[2]. However, with large data volumes and complexity, manual analysis becomes less efficient and prone to errors[3]. Muhammadiyah Pringsewu University as a higher education institution also faces challenges in managing its library so that it can meet the information needs of the entire academic community. This is where data mining plays an important role. By utilizing data mining to analyze book borrowing patterns, libraries can be more effective in developing collection management strategies and improving service quality. In recent years, data mining has become a powerful tool for enhancing library management by providing detailed analyses of user behavior and collection utilization. Several studies have examined various data mining techniques, such as association rule mining, clustering, and classification algorithms, to identify hidden patterns in library usage data. Researchers have found that algorithms like Naïve Bayes, Decision and K-means clustering can offer valuable insights into user preferences and improve services by enabling targeted recommendations and collection development. Despite these advancements, gaps remain in exploring the performance differences among these algorithms specifically in university libraries, as well as in developing strategies for real-time, automated recommendations. This research aims to bridge these gaps by focusing on the comparative effectiveness of Naïve Bayes and C4.5 in the context of a university library.

Data mining is a technique that can be used to extract valuable information from large and complex data[4][5]. In a library context, data mining can help identify hidden patterns in book lending data, which can then be used to improve library services, such as collection management, personalization of book recommendations, and budget planning for procuring new books[6].

To analyze book borrowing patterns, various data mining algorithms can be used, including the Naïve Bayes and C4.5 algorithms[7]. These two algorithms have different approaches to processing data and making predictions, so it is important to compare to determine which algorithm is more effective in a particular context. Naïve Bayes is an algorithm based on Bayes' theorem, which uses probability to make predictions[8]. This algorithm is simple, fast, and frequently used in text classification and pattern analysis. C4.5 is a decision tree-based machine learning algorithm that generates decision trees from training data[9]. These algorithms are known for their ability to handle varying data and provide easy-to-understand interpretations in the form of decision trees.

The application of data mining in library systems can

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2300, Copyright ©2025
Submitted : October 14, 2024, Revised : November 12, 2024, Accepted : November 14, 2024, Published : January 31, 2025

help identify hidden book borrowing patterns and provide more personalized recommendations to library users[10]. Analyzing user behavior through data mining techniques can improve library collection management by understanding user preferences based on lending data[11]. The Naïve Bayes algorithm has advantages in text classification, especially because of its simplicity and speed in processing data, which makes it suitable for applications with large data volumes[12]. Naïve Bayes is effective in library book recommendation systems because this algorithm can accurately predict book categories of interest based on borrowing history[13]. The C4.5 algorithm is very effective in producing decision trees that are easy to interpret, especially in complex data analysis such as book borrowing patterns in libraries[14]. The C4.5 algorithm is able to handle varied data well and provides more interpretive results than other algorithms in analyzing library user behavior.

Therefore, this research aims to compare the performance of the Naïve Bayes and C4.5 algorithms in analyzing book borrowing patterns at the Pringsewu Muhammadiyah University Library. It is hoped that the results of this research can provide recommendations regarding the most appropriate methods to be implemented in library information systems in the future.

II. METHODOLOGY

A. Research Stage

The method used in this research follows the stages of the Cross-Industry Standard Process for Data Mining (CRISP-DM) model[15]. The research stages can be seen in Figure 1 below:

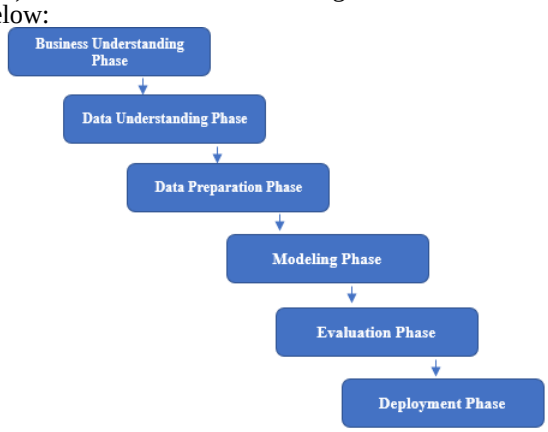


Figure 1 Research Flow

- *Business Understanding Phase*

At this stage the focus is on the research objective, namely to find out the best algorithm for analyzing book borrowing patterns in libraries by looking at what books are the most popular among borrowers, which students from study programs borrow books most often, then what are the book borrowing patterns over a certain period of time , so that the best model is obtained to fulfill the research objectives.

- *Data Understanding Phase*

The data that will be used in the research is Pringsewu Muhammadiyah University library information system data. There are 10 attributes that will be used in this research

Namely Borrower ID, Gender, Membership Type, Copy Code, Book Category, Length of Membership, Department/Faculty, Borrow Date, Return Date, Status.

- *Data Preparation*

Data preparation is one of the important stages in the CRISP-DM process which ensures that the data to be used for analysis is in optimal condition. Prepare data for analysis. This includes data cleaning, feature selection, and data transformation. After the data preparation stage is complete, the data will be ready to be used in the modeling stage. A careful data preparation process is essential to ensure that the analysis and models built are accurate and relevant.

- *Modeling (Modeling Phase)*

The algorithms used in this research are the C4.5 and Naive Bayes algorithms to classify book borrowing patterns at Pringsewu Muhammadiyah University and to obtain a model or function to describe graduation predictions by comparing the C4.5 and Naive Bayes algorithms.

- *Evaluation Phase (Evaluation Phase)*

At this stage, the performance evaluation of the two algorithms, namely the C4.5 and Naive Bayes Algorithms, is carried out by comparing the results of the average values of accuracy, recall and error rate contained in the confusion matrix table.

- *Deployment Phase (Deployment Phase)*

After the evaluation stage where the results of a model are assessed in detail, the model performance is monitored periodically and adjusted if necessary. Apart from that, adjustments were also made to the model so that it could produce results that were in line with the initial target of this CRISP-DM stage.

B. Data Collection Stage

Data collection methods are an important thing in research and are strategies or methods used by researchers to collect the data needed in their research. The data collection methods used in this research are:

- *Literature Review (Research Library)*
  The literature review is carried out by reading, quoting and making notes sourced from library materials that support and are related to research, in this case regarding C4.5 and Naive Bayes Algorithm data mining.
- *Field Studies (Field Research)*
  In this research, data collection through documents was carried out by studying the facts or data in the Pringsewu Muhammadiyah University Library information system.

C. Experiment Stages

This research will be carried out by applying two methods, namely the C4.5 Algorithm and Naive Bayes to classify book borrowing patterns.

- Decision tree (C4.5)

C4.5 is a collection of algorithms for classification techniques in machine learning and data mining. The goal is supervised learning, where each tuple in the data set can be described by a set of attribute values, and each tuple belongs to one of many different and incompatible classes[16]. The goal of C4.5 is to learn mappings from attribute values to categories that can be used to categorize unknown items into new categories. J. Rossi Quinlan suggests C4.5 based on ID3. A decision tree is constructed using the ID3 algorithm. A decision tree is a tree structure that is like a flowchart, with each internal node (nonleaf node) representing a test on an attribute, each branch representing a test result, and each leaf node holding a class label. After building a decision tree

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2300, Copyright ©2025
Submitted : October 14, 2024, Revised : November 12, 2024, Accepted : November 14, 2024, Published : January 31, 2025

for tuples that do not provide classification labels, we select a path from the root node to the leaf node, and the path stores the prediction information of the tuple. Decision trees have the advantage of not requiring domain information or parameter configuration, making them ideal for exploratory information mining[17].

The C4.5 algorithm is based on ID3 added to continuous attributes, attribute values, and information processing, by generating a tree to build a pruning decision tree in two stages. For each attribute, with the C4.5 algorithm information calculation, we can find out the Gain Ratio, the rate of information acquisition. Finally, it is selected with the highest level of information gain from the given test set attributes to organize the branch. According to the test attribute values using a recursive algorithm, obtain an initial decision tree. The computational formula related to the C4.5 algorithm is as follows[18]. First, the expectation value required for sample classification is given as follows: Determine the root of the tree by calculating the highest gain value of each attribute or the lowest entropy index value. Previously, the entropy index value was calculated using the formula:

$$Entropy\ (i) = \sum_{j=1}^{m} \quad f(i,j).\ 2f[(i,j)] \tag{1}$$

Gain value using the formula:

$$gain = -\sum_{i=1}^{p} \quad .\ IE(i) \tag{2}$$

To calculate the gain ratio, you need to know a new term called Split Information with the formula:

$$SplitInformation = -\sum_{t=1}^{c} \frac{S1}{S} log2\ \frac{S1}{S} \tag{3}$$

Next, calculate the gain ratio

$$Gainratio(S,A) = \frac{\text{Gain(S.A)}}{\text{SplitInformation (S,A)}} \tag{4}$$

Repeat step 2 until all records have been split. The decision tree splitting process ends when:

1. *All tuples in node record m are of the same class.*
2. *The attributes in the dataset are not further divided.*
3. *An empty branch has no records*

- **Naïve Bayes**

Naive Bayes is a probabilistic classification algorithm that utilizes Bayes' theorem to classify data. This algorithm is called "naive" because it assumes that all features in the dataset are independent of each other when assigned a particular class. Although this assumption is often unrealistic in practice, Naive Bayes remains popular and effective for many classification tasks[19].

Bayes' theorem is the basic principle of this algorithm. This theorem can be used to update the probability of a hypothesis based on new evidence. In the context of classification, the hypothesis is the class we are trying to predict, and new evidence is the observed features of the data.

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \tag{5}$$

Where:

- *P(C/X): Posterior probability of the class C is given a feature X*
- *P(X/C): Feature probability X is given a class C (likelihood).*
- *P(C): Class prior probability C*
- *P(X): Marginal probability of a feature X*

D. Performance Evaluation

- K-fold Cross Validation

k-fold cross-validation is a technique for validating the accuracy of a model built on a certain data set, which divides the data set into two parts, namely training data and testing data. For prediction problems, the model is usually given a dataset of known data to train on (training dataset) and unknown data (or first-time data) to test the model (called validation). or test data)[20]. The goal of cross-validation is to test a model's ability to predict new data that was not used in its evaluation, to flag problems such as overfitting or selection bias, and to provide insight into how the model generalizes to independent data. set (i.e. unknown dataset, e.g. problem).

- Confusion matrices

Confusion matrix is a very popular measure used when solving classification problems. It can be applied to binary classification as well as to multiclass classification problems. This matrix is used to evaluate the performance of the method used after classification. The confusion matrix represents TP values that are correctly classified, FP values in the relevant class when they should be in other classes, and FN values in other classes when they should be in the relevant class and TN values that are correctly classified in other classes[21]. The most frequently used performance metrics for classification according to these values are accuracy (ACC), precision (P), sensitivity (Sn), specificity (Sp), and F-score values. The calculation of these performance metrics according to the values in the confusion matrix is made according to Eq.

$$ACC = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{6}$$

$$P = \frac{\text{TP}}{\text{TP+FP}} \tag{7}$$

$$Sn = \frac{\text{TP}}{\text{TP+FN}} \tag{8}$$

$$Sp = \frac{\text{TN}}{\text{TN+FP}} \tag{9}$$

$$F - score = 2\text{x}\frac{\text{P x Sn}}{\text{P+Sn}} \tag{10}$$

III. RESULTS AND DISCUSSION

A. Analysis and Preprocessing Data

Before applying the algorithm, data collected from the Pringsewu Muhammadiyah University Library was analyzed and processed first. The dataset consists of 5618 data with ten attributes including Borrower ID, Gender, Membership Type, Copy Code, Book Category,

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2300, Copyright ©2025
Submitted : October 14, 2024, Revised : November 12, 2024, Accepted : November 14, 2024, Published : January 31, 2025

Membership Length, Department/Faculty, Borrowing Date, Return Date, and Status. A total of 170 records were removed due to missing values and duplicates identified. For example, all records with an incomplete Borrower ID or missing Return Date will be excluded from the data set. This data can be seen in Figure 2 below.

Figure 2 Dataset

B. Modeling and Evaluation

Both algorithms (C4.5 and Naïve Bayes) were implemented using the preprocessed dataset. The performance of each algorithm was evaluated using K-fold cross-validation (with k=10) to ensure reliable accuracy metrics.

Algorithms C4.5

The application of data in Rapidminer for analyzing book borrowing using the C4.5 algorithm is shown in Figure 3 below:

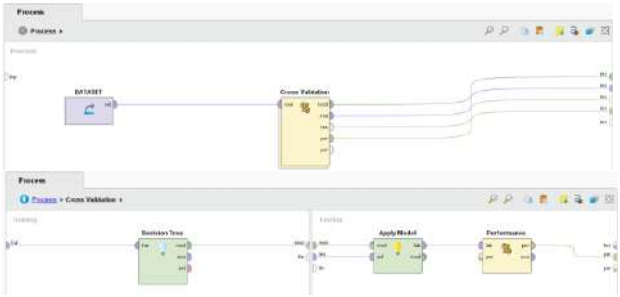


Figure 3 Testing scheme with C4.5

In Figure 3, the prepared dataset is applied to the Rapidminer application by conducting experiments using cross validation which can directly divide the data into training data and testing data because the data used is supervised and the algorithm used is C4.5. We can see the experimental results in Figure 4 below.

Figure 4 Accuracy Results with the C4.5 Algorithm

In figure 4 The model achieved an overall accuracy of 96.26%. This high accuracy indicates that the model correctly classifies approximately 96 out of every 100 instances, suggesting effective performance in distinguishing between the classes. Decision tree generated by the C4.5 algorithm for classifying book borrowing patterns we can see the experimental results in Figure 5 below.

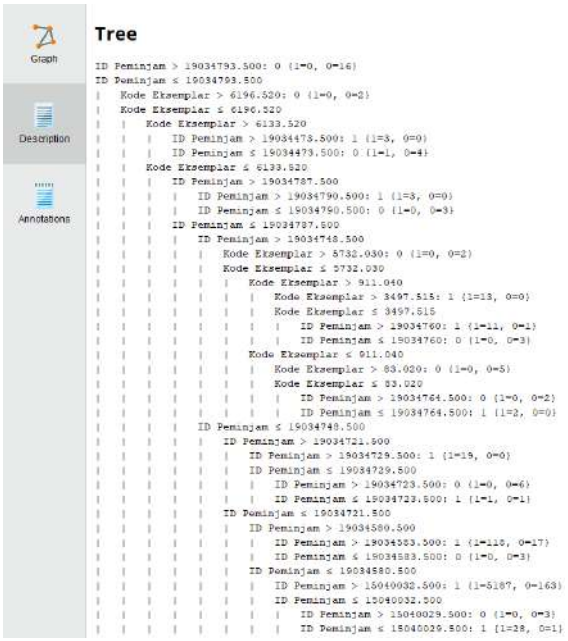


Figure 5 Decision tree

Figure 5 above depicts the decision tree structure used to classify Borrower IDs and Exemplar Codes based on various conditions. Each branch of the tree represents a decision or comparison made against the Borrower ID and Exemplar Code. For example, first a separation is carried out based on Borrower ID > 19034793.500, which then leads to further grouping based on Exemplar Code and Borrower ID values. Each branch contains information regarding the number of borrowers who meet or do not meet the given conditions, with the numbers in brackets indicating the number of borrowers in each category.

This decision tree structure helps in understanding various factors such as "Exemplar Code", "Borrower ID", etc. that influence borrowing patterns and helps librarians in making data-based decisions regarding book management. This can guide collection development, help identify popular books, and understand borrowing behavior at the

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2300, Copyright ©2025
Submitted : October 14, 2024, Revised : November 12, 2024, Accepted : November 14, 2024, Published : January 31, 2025

Pringsewu Muhammadiyah University Library.

Algorithms Naïve Bayes

The application of data in Rapidminer for analyzing book borrowing using the Naïve Bayes algorithm is shown in Figure 6 below.

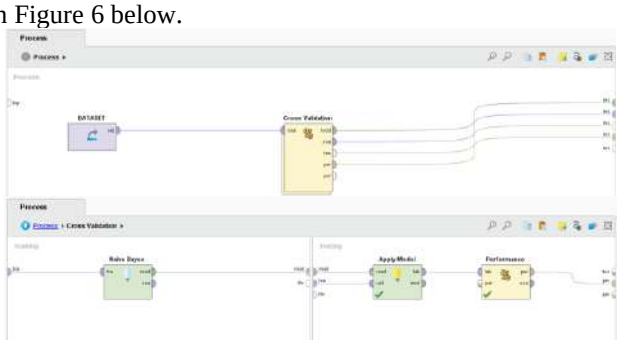


Figure 6 Testing scheme with Naïve Bayes

The experiment in Figure 3.5 is a validation technique for dividing training data and testing data using cross validation techniques. From this experiment, we got the results that we can see in Figure 7 below

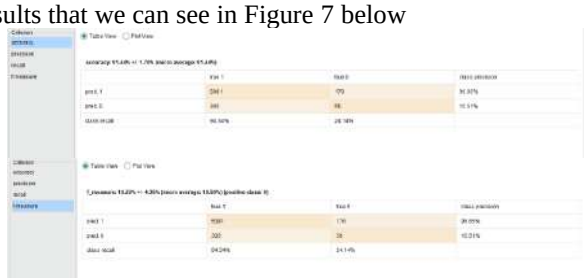


Figure 7 Accuracy Results with the Naïve Bayes Algorithm

In figure 7 The model achieved an overall accuracy of 91.44%. Apart from accuracy in the Naive Bayes algorithm, there is SimpleDistribution where The focus of this model is on the label attribute Status Pengembalian. This attribute likely classifies whether a book has been returned on time or not, reflected by binary classes (1 for returned on time, 0 for not returned on time). we got the results that we can see in Figure 8 below.

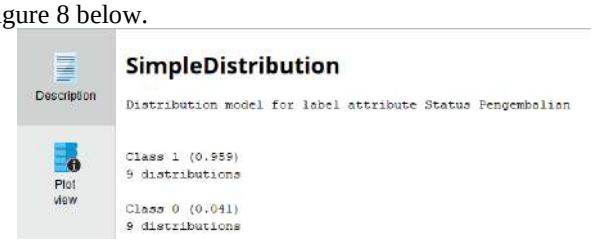


Figure 8 Sample Distribution

In figure 8 The model indicates that Class 1 has a probability of 0.959, suggesting that 95.9% of the instances in the dataset are classified as this class. This is a strong indication that most books are returned on time. In contrast, Class 0 has a probability of 0.041, indicating that only 4.1% of the instances are classified as not returned on time. This suggests that overdue returns are significantly less frequent compared to timely returns. Both classes have 9 distributions listed, which means the model used nine different feature distributions (attributes) to help classify the return status of the books.

C. Results

The results of the model performance are summarized in the following

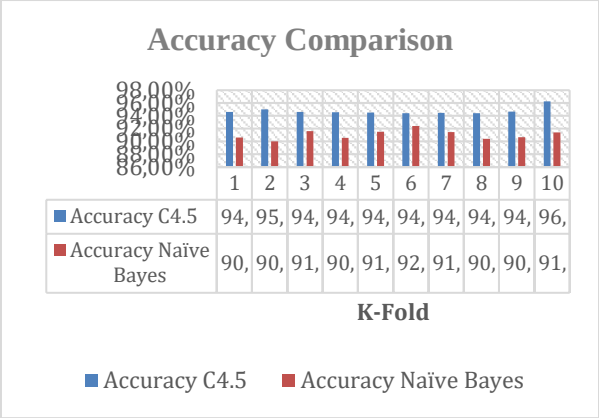


Figure 10 algorithm performance comparison results

Based on the table and diagram image above, it shows that changing the number of K-Folds in Cross Validation when carrying out classification will produce different accuracy so that we can produce the best accuracy. We can see that with K-Fold 10, the accuracy and precision results are the highest, in contrast to the recall results which are not higher than other uses of K-Fold. C4.5 achieved an accuracy of 96.26%, indicating it correctly classified a significant majority of the borrowing patterns. In contrast, Naïve Bayes, with an accuracy of 91.44%, showed a slightly lower capability in predicting user behavior.

D. Discussion

The C4.5 algorithm outperformed the Naïve Bayes algorithm across all evaluated metrics. This can be attributed to C4.5's ability to model complex relationships in the data through its decision tree structure. The tree structure allows for capturing interactions between features, which is particularly useful when analyzing book borrowing patterns that may be influenced by multiple factors (e.g., book categories, user demographics). C4.5 achieved an accuracy of 96.26%, indicating it correctly classified a significant majority of the borrowing patterns. In contrast, Naïve Bayes, with an accuracy of 91.44%, showed a slightly lower capability in predicting user behavior.

The findings emphasize the importance of utilizing data mining techniques for effective library management. By implementing the C4.5 algorithm, the Pringsewu Muhammadiyah University Library can enhance its collection management strategies, personalize recommendations for users, and allocate resources more efficiently based on borrowing trends.

IV. CONCLUSION

The study successfully demonstrated the application of the C4.5 and Naïve Bayes algorithms in analyzing book borrowing patterns at the Pringsewu Muhammadiyah University Library. The results showed that the C4.5 algorithm outperformed Naïve Bayes across all evaluation

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2300, Copyright ©2025
Submitted : October 14, 2024, Revised : November 12, 2024, Accepted : November 14, 2024, Published : January 31, 2025

metrics, including accuracy, precision, recall, and F-score. Specifically, C4.5 achieved an accuracy of 96.26%, while Naïve Bayes reached 91.44%.

This indicates that C4.5 is more effective in capturing complex relationships in the data and predicting user borrowing behavior. The findings of this research underscore the importance of employing data mining methodologies in library management. By implementing the C4.5 algorithm, the Pringsewu Muhammadiyah University Library can improve its services, such as personalized book recommendations, targeted marketing efforts, and more efficient budgeting for new acquisitions. Understanding borrowing patterns also aids in developing strategies to meet the evolving information needs of the academic community.

While this study provides valuable insights, it also highlights several areas for future research. Subsequent studies could explore the application of other advanced machine learning algorithms, such as Random Forest or Support Vector Machines, to compare their performance against C4.5 and Naïve Bayes. Additionally, expanding the dataset to include a more extended time frame and more diverse attributes could enhance the robustness of the analysis and its applicability across different library settings.

ACKNOWLEDGMENT

The author would like to thank all parties who have provided support and assistance in preparing this journal, especially the supervisors and colleagues who have provided valuable input as well as their beloved family who always provide support.

REFERENCES

[1] L. Yulita, A. S. Sunge, and N. Nurhidayanti, "Optimasi Algoritma Genetika Dalam Memprediksi Minat Baca Siswa Pada Perpustakaan SMK Negeri 1 Gantar Dengan Metode Decision Tree."

[2] E. Irfiani, Y. Kusnadi, S. Sunarti, and F. Handayanna, "Implementasi Data Mining dalam Mengklasifikasi Minat Baca Pada Perpustakaan Daerah Menggunakan Algoritma C4.5," *JOINS (Journal of Information System)*, vol. 8, no. 2, pp. 106–114, Nov. 2023, doi: 10.33633/joins.v8i2.8004.

[3] T. T. Sang Nguyen, "Model-based book recommender systems using Naïve Bayes enhanced with optimal feature selection," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, 2019, pp. 217–222. doi: 10.1145/3316615.3316727.

[4] L. Feng, "Research on Higher Education Evaluation and Decision-Making Based on Data Mining," *Sci Program*, vol. 2021, 2021, doi: 10.1155/2021/6195067.

[5] D. Berrar, "Bayes' Theorem and Naive Bayes Classifier," in *Encyclopedia of Bioinformatics and Computational Biology*, S. Ranganathan, M. Gribskov, K. Nakai, and C. Schönbach, Eds., Oxford: Academic Press, 2019, pp. 403–412. doi: https://doi.org/10.1016/B978-0-12-809633-8.20473-1.

[6] J. Moolayil, *Learn Keras for Deep Neural Networks*. Apress, 2019. doi: 10.1007/978-1-4842-4240-7.

[7] S. Xuanyuan, S. Xuanyuan, and Y. Yue, "Application of C4.5 Algorithm in Insurance and Financial Services Using Data Mining Methods," *Mobile Information Systems*, vol. 2022, 2022, doi: 10.1155/2022/5670784.

[8] X. Xiongjun and D. Lv, "The Evaluation of Music Teaching in Colleges and Universities Based on Machine Learning," *Journal of Mathematics*, vol. 2022, 2022, doi: 10.1155/2022/2678303.

[9] S. and Communication Networks, "Retracted: Analysis and Application of Data Mining Technology for College English Education Integration," *Security and Communication Networks*, vol. 2024, pp. 1–1, Jan. 2024, doi: 10.1155/2024/9836129.

[10] W. Liu *et al.*, "Using a classification model for determining the value of liver radiological reports of patients with colorectal cancer," *Front Oncol*, vol. 12, Nov. 2022, doi: 10.3389/fonc.2022.913806.

[11] P. V. Ngoc, C. V. T. Ngoc, T. V. T. Ngoc, and D. N. Duy, "A C4.5 algorithm for english emotional classification," *Evolving Systems*, vol. 10, no. 3, pp. 425–451, Sep. 2019, doi: 10.1007/s12530-017-9180-1.

[12] J. Wang, "Application of C4.5 Decision Tree Algorithm for Evaluating the College Music Education," *Mobile Information Systems*, vol. 2022, 2022, doi: 10.1155/2022/7442352.

[13] Y. A. Alsariera, Y. Baashar, G. Alkawsi, A. Mustafa, A. A. Alkahtani, and N. Ali, "Assessment and Evaluation of Different Machine Learning Algorithms for Predicting Student Performance," 2022, *Hindawi Limited*. doi: 10.1155/2022/4151487.

[14] X. Yang and J. Ge, "Predicting Student Learning Effectiveness in Higher Education Based on Big Data Analysis," *Mobile Information Systems*, vol. 2022, 2022, doi: 10.1155/2022/8409780.

[15] Parteek Bhatia, "Data Mining and Data Warehousing," 2019.

[16] B. Mahesh, "Machine Learning Algorithms - A Review," *International Journal of Science and Research (IJSR)*, vol. 9, no. 1, pp. 381–386, Jan. 2020, doi: 10.21275/art20203995.

[17] T. Sinta Peringkat *et al.*, "Komparasi Algoritma Decision Tree, Naive Bayes Dan K-Nearest Neighbor Untuk Memprediksi Mahasiswa Lulus Tepat Waktu," 2020, [Online]. Available: www.bri-institute.ac.id

[18] S. Alim, "Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model K-Nearest Neighbor, Decision Tree Serta Naive Bayes Orange Data Mining Implementation For Student Graduation Classification Using K-Nearest Neighbor, Decision Tree And Naive Bayes Models," 2021.

[19] D. Lianda and N. Surya Atmaja, "Prediksi Data Buku Favorit Menggunakan Metode Naïve Bayes (Studi Kasus: Universitas Dehasen Bengkulu)," 2021. [Online]. Available: www.ejournal.unib.ac.id/index.php/pseudocode

[20] Nurmalitasari, Z. Awang Long, and M. Faizuddin Mohd Noor, "Factors Influencing Dropout Students in Higher Education," *Educ Res Int*, vol. 2023, 2023, doi: 10.1155/2023/7704142.

[21] S. A. Alwarthan, N. Aslam, and I. U. Khan, "Predicting Student Academic Performance at Higher Education Using Data Mining: A Systematic Review," 2022, *Hindawi Limited*. doi: 10.1155/2022/8924028.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2300, Copyright ©2025
Submitted : October 14, 2024, Revised : November 12, 2024, Accepted : November 14, 2024, Published : January 31, 2025

Evaluating GRU Algorithm and Double Moving Average for Predicting USDT Prices: A Case Study 2017-2024

Rahmat Rizky[1]*, Munirul Ula[2], Zara Yunizar[3]
Dept. of Informatics Engineering, Universitas Malikussaleh [1], [2], [3]
Lhokseumawe, Indonesia
rahmat.200170114@mhs.unimal.ac.id [1], munirulula@unimal.ac.id [2], zarayunizar@unimal.ac.id [3]

Abstract*— The cryptocurrency market is highly volatile, requiring advanced analytical methods for accurate price forecasting. This study evaluates the effectiveness of Gated Recurrent Units (GRU) and Double Moving Average (DMA) in predicting USDT (Tether Coin) prices using historical data from 2017 to 2024, sourced from Investing.com. Implemented in Jupyter Notebook, the research explores the strengths of each method in analyzing market fluctuations and price trends. GRU, a deep learning-based recurrent neural network, processes sequential data using a gating mechanism, making it effective for capturing short-term price dynamics. DMA, in contrast, is a statistical method that filters market noise to identify long-term trends, making it more reliable for stable market conditions. Performance evaluation shows DMA achieving lower errors (MAE: 5.494, MAPE: 0.0339%) than GRU (MAE: 5.984, MAPE: 0.0369%), suggesting higher accuracy for trend-based predictions. However, GRU's lower RMSE (8.531 vs. 8.715 for DMA) indicates better adaptability to sudden price fluctuations, making it more responsive to volatile markets. A hybrid approach combining GRU and DMA reveals their complementary strengths—DMA's minimal bias (-0.0013% MPE) supports stable trend analysis, while GRU's slight positive bias (0.0286% MPE) captures short-term fluctuations. Additionally, a comparison with Long Short-Term Memory (LSTM) demonstrates its superior predictive accuracy, outperforming both GRU (MAE: 5.98, RMSE: 8.53) and DMA (MAE: 5.49, RMSE: 8.72) with the lowest MAE (4.31), MAPE (0.027%), and RMSE (5.64), alongside minimal bias (MPE: 0.007%). This study highlights the need for integrating multiple forecasting techniques in cryptocurrency price prediction. While DMA is well-suited for stable trends and GRU excels in volatile conditions, LSTM outperforms both, reinforcing the effectiveness of deep learning for financial time-series forecasting.

Keywords: Cryptocurrency, Forecasting, Historical Data, Jupyter Notebook

I. INTRODUCTION

The cryptocurrency market is characterized by significant volatility and unpredictability, presenting both challenges and opportunities for market participants. These rapid price fluctuations necessitate sophisticated analytical tools and predictive models to support informed decision-making in trading and investment activities [1][2][3] .

The selection of Gated Recurrent Units (GRU) and Double Moving Average (DMA) methods for USDT price prediction is fundamentally driven by their complementary strengths in addressing the complexities of cryptocurrency markets. GRU demonstrates exceptional capability in processing sequential data and identifying intricate patterns within cryptocurrency price movements. Its sophisticated architecture enables the retention of critical long-term historical data while maintaining sensitivity to emerging market trends [4][5]. The method's superior performance in managing the vanishing gradient problem, coupled with its computational efficiency, makes it particularly well-suited for analyzing the dynamic nature of cryptocurrency markets [6].

The implementation of DMA methodology complements the GRU approach through its effective filtering of market noise and precise identification of price trends via dual-timeframe analysis. This method has proven especially valuable in generating accurate trading signals for USDT, which exhibits subtle price variations despite its stablecoin designation [7][8]. Research by Sari has demonstrated DMA's reliability in projection accuracy based on historical data analysis, validating its applicability to USDT price movement analysis [9].

The strategic integration of GRU and DMA methodologies creates a comprehensive analytical framework that leverages their respective strengths. While GRU excels in deep pattern recognition and sequential data processing, DMA provides structured trend analysis, forming a synergistic approach to USDT price prediction. This combined methodology aims to deliver enhanced predictive accuracy and deeper market insights, ultimately providing traders and investors with a more robust foundation for strategic decision-making.

USDT (Tether), as a prominent stablecoin, maintains a value pegged to the US dollar through reserve asset backing equivalent to its circulating supply. Its widespread adoption in cross-border transactions, cryptocurrency exchange trading, and decentralized finance (DeFi) ecosystems underscores the importance of accurate price prediction models [10][11]. This

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2328, Copyright ©2025
Submitted : December 2, 2024, Revised : January 8, 2025, Accepted : January 28, 2025, Published : January 31, 2025

research utilizes historical data from 2017 to 2024 to develop a precise predictive model, aiming to enhance the quality of investment decisions in the USDT market through data-driven analysis.

II. LITERATURE REVIEW

A. Cryptocurrency

Cryptocurrency is a digital or virtual form of currency that employs cryptography to secure transactions and control the creation of new units. Built on decentralized technology, commonly referred to as blockchain, it allows for the verification of transactions and maintenance of an ownership ledger without relying on a central authority. Since the advent of Bitcoin in 2008, cryptocurrencies have become integral to global financial markets, reaching a market capitalization exceeding $600 billion. However, debates persist regarding the role and value of cryptocurrencies, such as Bitcoin, and their correlation with global economic indices [12][13].

B. Investment in Cryptocurrency

Cryptocurrency investments are gaining traction due to speculative opportunities and their potential as a store of value. However, these markets are highly volatile, with prices heavily influenced by market demand, regulatory announcements, and technological innovations [14][15].

Formulas Used

Volatility (σ):

$$\sigma - \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (r_i - \bar{r})^2} \quad (1)$$

Explanation :
σ : Standard deviation (volatility)
N : Number of data points
r_i : Return at time i
$\bar{r}$: Mean return

Expected Return ($E[R]$) :

$$E[R] - \sum_{i=1}^{N} P_i \cdot r_i \quad (2)$$

Explanation :
$E[R]$: Expected return
P_i : Probability of outcome i
r_i : Return at outcome i

C. USDT(Teather coin)

USDT, a stablecoin, maintains its value by pegging it 1:1 with the US dollar. This stability has made it a preferred cryptocurrency for transactions in a volatile market [16]. Tether's value stability can be modeled as:

$$V_{\text{USDT}} - \frac{\text{Reserves}}{\text{Circulating Supply}} \quad (3)$$

Explanation :
$V_{\text{USDT:}}$ Value of USDT
Reserves: Assets backing USDT
Circulating Supply: Total USDT units in circulation

D. Gated Recurrent Units (GRU)

GRU is a simplified Recurrent Neural Network (RNN) model effective for sequential data predictions, such as time series forecasting. Unlike LSTM, GRU uses two gates—the update gate and the reset gate—to control the flow of information between time steps [17][18].

Update Gate (z_t) :

$$z_t - \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (4)$$

Explanation :
z_t : Update gate value at time t
σ : Sigmoid activation function
W_z : Weight matrix for the update gate
h_{t-1} : Previous hidden state
x_t : Input at time t
b_z : Bias term

Reset Gate (r_t) :

$$r_t - \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (5)$$

Explanation :
r_t : Reset gate value at time t
W_r : Weight matrix for the reset gate
b_r : Bias term

New Memory Content $(\bar{h}_t)$:

$$\bar{h}_t - \tanh(W \cdot [r_t * h_{t-1}, x_t] + b) \quad (6)$$

Explanation :
$\bar{h}_t$: New memory content
tanh: Hyperbolic tangent activation function
W: Weight matrix for new memory content
r_t : Reset gate value
*: Element-wise multiplication
b: Bias term

Final Hidden State (h_t) :

$$h_t - (1 - z_t) * h_{t-1} + z_t * \bar{h}_t \quad (7)$$

Explanation of Symbols:
h_t : Final hidden state at time t

E. Double Moving Average (DMA)

The Double Moving Average (DMA) method is used for smoothing time series data and identifying price trends [19]. It calculates two moving averages with different periods:

Short-Term Moving Average (SMA):

$$\text{SMA}_{\text{short}} - \frac{\sum_{i-1}^{n} P_i}{n} \quad (8)$$

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2328, Copyright ©2025
Submitted : December 2, 2024, Revised : January 8, 2025, Accepted : January 28, 2025, Published : January 31, 2025

Explanation of Symbols:
SMA $_{short:}$ Short-term moving average
P_i: Price at time i
n : Short-term period

Long-Term Moving Average (SMA):

$$\text{SMA}_{\text{long}} - \frac{\sum_{i-1}^{m} P_i}{m} \quad (9)$$

Explanation of Symbols:
SMA $_{long:}$: Long-term moving average
m : Long-term period

Signal:
A buy signal occurs when SMA $_{short} >$ SMA $_{long,}$ and a sell signal occurs when SMA $_{short} <$ SMA$_{long}$.

F. Technical Analysis

Technical analysis uses mathematical models to predict trends and guide investment decisions. A common tool is the Moving Average Convergence Divergence (MACD) [20]:
Formulas Used

MACD:

$$\text{MACD} - \text{EMA}_{\text{short}} - \text{EMA}_{\text{long}} \quad (10)$$

Explanation of Symbols:
MACD: Moving Average Convergence Divergence
EMA$_{short}$: Short-term exponential moving average
EMA$_{long:}$ Long-term exponential moving average

G. In Newer Pesearch

The previous study in "*A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms*" by Mohammad J. Hamayel and Amani Yousef Owda, explores the implementation of three machine learning algorithms—GRU, LSTM, and Bi-LSTM—for predicting cryptocurrency prices, including Bitcoin (BTC), Ethereum (ETH), and Litecoin (LTC), using historical data. The results indicate that GRU outperforms LSTM and Bi-LSTM in terms of prediction accuracy, achieving lower Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) across all tested cryptocurrencies. The dataset spans daily data from January 2018 to June 2021, with performance evaluated using RMSE and MAPE metrics. GRU is proven to be the most reliable and efficient model for price prediction in highly volatile markets. The study also recommends future exploration of additional factors, such as social media sentiment analysis and trading volume, to further enhance prediction accuracy.

In similar new research, The study, titled *"Methods of Forecasting the Prices of Cryptocurrency on the Financial Markets"* by Yurii Pronchakov and Oleg Bugaienko, analyzes various moving average techniques for cryptocurrency price forecasting, including Simple Moving Average (SMA), Exponential Moving Average (EMA), and Weighted Moving Average (WMA). The research shows that SMA has the lowest mean-square deviation (0.46), making it the most stable method, while EMA, with a deviation of 0.50, is more responsive to short-term price changes, which is advantageous for volatile markets. WMA has a higher deviation of 0.69. The authors recommend using EMA for short-term analysis and SMA for long-term trends. The study concludes that integrating these techniques into real-time financial tools, such as mobile applications, could enhance decision-making for cryptocurrency traders.

III. RESEARCH AND METHOD

This study was conducted in Lhokseumawe City from December 2023 until its completion, as part of the graduation requirements for the Informatics Engineering Program at Universitas Malikussaleh. It focused on developing predictive models for USDT (Tether Coin) prices using historical data from 2017 to 2024, sourced from Investing.com and supplemented by the Stock Market API (api.marketstack.com). Data from January 2024 was reserved for model testing to assess the predictive accuracy of the proposed methods. Additionally, secondary data from academic journals and books was utilized to provide theoretical context and insights into cryptocurrency market behavior and price dynamics.

The research methodology included a literature review and data analysis. The literature review examined prior studies, academic journals, and relevant publications to establish a theoretical framework emphasizing the unique characteristics of USDT and its role in the cryptocurrency market. Data analysis involved preprocessing the historical price dataset through cleaning, outlier detection, and normalization using Min-Max Scaling to ensure the data was suitable for machine learning models. Two predictive methods were applied: Double Moving Average (DMA), which calculated short-term (10-day) and long-term (50-day) moving averages to identify trends and market signals, and Gated Recurrent Units (GRU), a sequential modeling approach leveraging TensorFlow and Keras for time-series prediction.

Performance evaluation of the predictive models employed metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to quantify accuracy. DMA demonstrated effectiveness in identifying long-term trends, while GRU excelled at capturing short-term market fluctuations. The results, visualized through comparative charts and tables, revealed distinct strengths for each method, offering a comprehensive understanding of USDT price trends.

This study focuses solely on evaluating two predictive methods, namely Double Moving Average (DMA) and Gated Recurrent Units (GRU), for modeling USDT prices. It does not compare other methods that could potentially be applied to cryptocurrency price analysis, nor does it aim to develop new techniques that might outperform the two methods under investigation. The primary objective of this research is to assess the effectiveness of these existing methods in predicting USDT

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2328, Copyright ©2025
Submitted : December 2, 2024, Revised : January 8, 2025, Accepted : January 28, 2025, Published : January 31, 2025

prices using historical data from 2017 to 2024. Therefore, the findings are not intended to compare a range of approaches or to introduce innovative new methods, but rather to evaluate how well DMA and GRU perform in analyzing time-series data for cryptocurrency prices, specifically USDT.

Despite the promising results obtained from the GRU and DMA methods, there are several limitations that must be considered. First, both methods rely on historical data patterns to make predictions, which can be problematic in the context of the cryptocurrency market, known for its high volatility and dynamic nature. Since cryptocurrency prices often experience sudden shifts due to various factors, these methods may struggle to capture emerging trends or adapt to sudden market changes. Additionally, the data used in this study is based on daily closing prices, which presents another limitation. While historical data spanning multiple years provides valuable insights, it only captures a snapshot of the market at a single point in time each day. Cryptocurrency prices, however, can fluctuate dramatically within a single day, and the daily resolution of data may not fully account for these intra-day changes, limiting the accuracy of predictions. As such, while GRU and DMA are effective within the constraints of their data and modeling techniques, they cannot be regarded as absolute references for predicting cryptocurrency prices due to the constantly evolving nature of the market and the limitations in data granularity.

IV. RESEARCH RESULT

The results and discussion of the research conducted on predicting the cryptocurrency USDT (Tether) prices using the Gated Recurrent Units (GRU) and Double Moving Average (DMA) algorithms are presented in this chapter. It will provide a detailed explanation of the data description used, the calculation process, the results of the algorithm implementation, as well as a comprehensive analysis of the obtained outcomes. This chapter also includes tables and charts to support the explanation and help readers understand the prediction patterns generated from the implemented models.
The process begins with the collection of daily USDT prices in IDR, followed by preprocessing to ensure the data is clean and consistent. The first step involves calculating a three-day moving average to identify short-term trends. Next, a six-day moving average is computed using the results of the first moving average to identify long-term trends. The level and trend components are then determined based on the differences between these two moving averages. Finally, price predictions for the next five days are generated by adding the level component to the trend component, multiplied by the number of days to be forecasted. Data collection is a critical initial step in implementing the Double Moving Average (DMA) method for predicting cryptocurrency prices, specifically USDT in IDR. Daily price data for USDT is sourced from reliable providers offering real-time or historical information. Once the data is gathered, preprocessing ensures that the dataset is clean and consistent.

Moving Average (MA) calculations are widely employed in financial market analysis to identify price trends, including daily price assessments. An MA is computed as the average of prices over a specific period, offering a smoother representation of price movements and filtering out temporary fluctuations. And The Implementation Flow is :

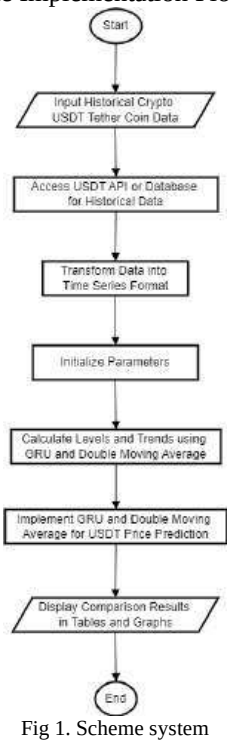


Fig 1. Scheme system

Details :

- *Start*
- *Input historical data for Crypto USDT (Tether Coin).*
- *Access the USDT API or Database for historical data on USDT.*
- *Transform the data into a time series format.*
- *Initialize parameters.*
- *Calculate levels and trends using GRU and Double Moving Average.*
- *Implement Gated Recurrent Units (GRU) and Double Moving Average to predict USDT prices.*
- *Display the comparison results in the form of tables and graphs.*
- *End process*

Based on the diagram below, the implementation process of the Gated Recurrent Units (GRU) and Double Moving Average (DMA) methods for cryptocurrency price prediction begins with loading historical USDT data from 2017 to 2024. The data is cleaned, normalized, and split into training and testing sets. For the DMA method, short-term and long-term moving averages are calculated to identify level and trend components, which are used to generate price predictions. Meanwhile, the data is transformed into a time-series format to train the GRU model. After predictions are generated from both methods, their performance is evaluated using metrics such as MAE, RMSE, and MAPE. Based on the comparison, the more reliable method is selected: DMA is recommended for stable trends, while GRU is suited for volatile market conditions. The final results are saved and visualized to support investment decisions.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2328, Copyright ©2025
Submitted : December 2, 2024, Revised : January 8, 2025, Accepted : January 28, 2025, Published : January 31, 2025

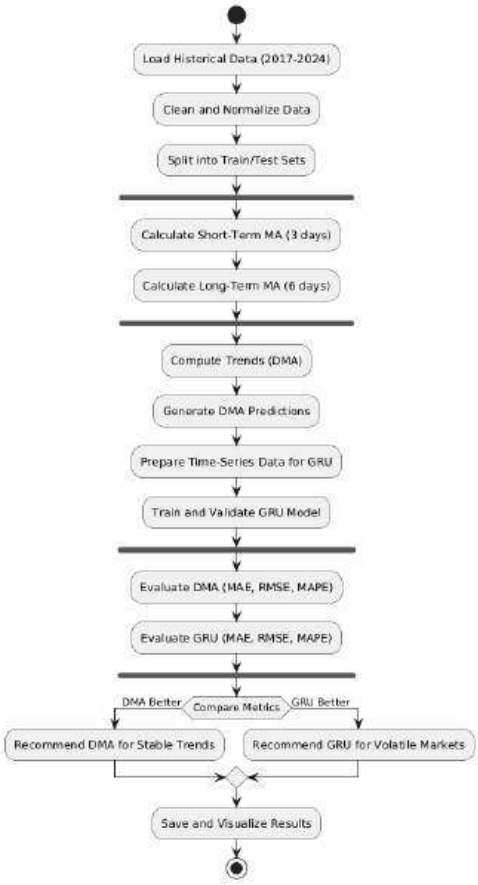


Fig 2 Data Flow

A. Training Data

The results of the research conducted, including data analysis and the application of prediction algorithms, will be discussed in this chapter. The results from the use of the Gated Recurrent Units (GRU) and Double Moving Average (DMA) algorithms will be presented to evaluate the effectiveness of the models in predicting the price of the cryptocurrency USDT (Tether). This analysis will also include a comparison between the predicted results and the actual data to assess the accuracy of the developed models.

Table I. Tether Coin USDT Historical Data (2017-2024)

No	Date	close	Open	High	low
1	14/04/2017	16.028,36	16.093,13	16.127,13	15.688,33
2	15/04/2017	16.023,50	16.028,36	16.028,36	15.667,28
3	16/04/2017	15.406,59	16.023,50	16.023,50	15.406,59
4	17/04/2017	15.207,43	15.406,59	15.702,90	14.914,36
5	18/04/2017	15.301,35	15.207,43	15.589,56	14.907,88
6	19/04/2017	15.317,54	15.301,35	15.461,65	15.175,05
7	20/04/2017	15.366,11	15.317,54	15.461,65	14.896,55
8	21/04/2017	15.160,48	15.366,11	15.382,31	14.572,71
9	22/04/2017	14.815,59	15.160,48	15.382,31	14.815,59
….	….	….	….	….	….
….	….	….	….	….	….
10	12/07/2024	16.198,38	16.196,76	16.201,62	16.196,76
11	13/07/2024	16.200,00	16.198,38	16.203,23	16.195,14
12	14/07/2024	16.201,62	16.200,00	16.204,85	16.198,38
13	15/07/2024	16.200,00	16.201,62	16.208,09	16.196,76
14	16/07/2024	16.200,00	16.200,00	16.203,23	16.195,14
15	17/07/2024	16.191,90	16.200,00	16.200,00	16.191,90
16	18/07/2024	16.188,66	16.191,90	16.195,14	16.187,04
17	19/07/2024	16.188,66	16.188,66	16.191,90	16.187,04

The table above presents the historical price data of USDT (Tether) used as the training data in this study. The data includes the closing price, opening price, highest price, and lowest price over the period from April 14, 2017, to July 19, 2024. This data is used to train the prediction model, where price movement patterns are analyzed using the Gated Recurrent Units (GRU) and Double Moving Average (DMA) algorithms. Each row represents daily data collected sequentially to capture trends and fluctuations over the given period. This information is crucial in helping the model learn and generate more accurate predictions for the future.*Identify the Headings.*

B. Testing Data

Table II. Tether Coin USDT test Historical Data (2017-2024)

No	Date	Close	Open	High	Low
1	01/01/2020	16.387,82	16.345,72	16.397,54	16.318,20
2	02/01/2020	16.277,72	16.455,83	16.462,30	16.274,48
3	03/01/2020	16.255,05	16.277,72	16.358,68	16.248,57
4	04/01/2020	16.272,86	16.295,53	16.342,48	16.271,24
5	05/01/2020	16.243,71	16.272,86	16.326,29	16.242,09
6	06/01/2020	16.258,29	16.258,29	16.279,34	16.240,48
7	07/01/2020	16.217,81	16.279,34	16.284,19	16.212,95
8	08/01/2020	16.212,95	16.217,81	16.238,86	16.209,71
9	09/01/2020	16.230,76	16.212,95	16.245,33	16.212,95
….	….	….	….	….	….
….	….	….	….	….	….
10	12/07/2024	16.198,38	16.196,76	16.201,62	16.196,76
11	13/07/2024	16.200,00	16.198,38	16.203,23	16.195,14
12	14/07/2024	16.201,62	16.200,00	16.204,85	16.198,38
13	15/07/2024	16.200,00	16.201,62	16.208,09	16.196,76
14	16/07/2024	16.200,00	16.200,00	16.203,23	16.195,14
15	17/07/2024	16.191,90	16.200,00	16.200,00	16.191,90
16	18/07/2024	16.188,66	16.191,90	16.195,14	16.187,04
17	19/07/2024	16.188,66	16.188,66	16.191,90	16.187,04

The table above presents the test data used in this research, containing historical USDT (Tether) prices from the period of January 1, 2020, to July 19, 2024. The data includes the closing price, opening price, highest price, and lowest price for each day. This test data is used to evaluate and validate the accuracy of the prediction model that has been trained on prior data. The prediction models, both using the Gated Recurrent Units (GRU) and Double Moving Average (DMA) algorithms, will be tested against this data to assess how well the models can predict the actual prices based on the historical patterns already analyzed. The predicted results from the models will then be compared with the actual values from the test data to evaluate the effectiveness and performance of the developed models.

Table III. Comparison Results of GRU and DMA Metrics

Metric	DMA	GRU
MAE	5.494285	5.983858
RMSE	8.715295	8.530839
MAPE	0.033918	0.036935
MPE	-0.001262	0.028594

The comparison between Double Moving Average (DMA) and Gated Recurrent Unit (GRU) in predicting closing prices

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2328, Copyright ©2025
Submitted : December 2, 2024, Revised : January 8, 2025, Accepted : January 28, 2025, Published : January 31, 2025

reveals that both methods have their own strengths. Overall, DMA performs better in terms of absolute error metrics, with MAE at 5.494 and MAPE at 0.0339%, compared to GRU .

(MAE at 5.983 and MAPE at 0.0369%). This indicates that DMA provides more stable predictions with lower average errors.

However, GRU excels in handling large errors, as reflected in its lower RMSE (8.5308) compared to DMA (8.7153). Additionally, GRU is more responsive to rapid price fluctuations, making it more suitable for use in dynamic and volatile market conditions.

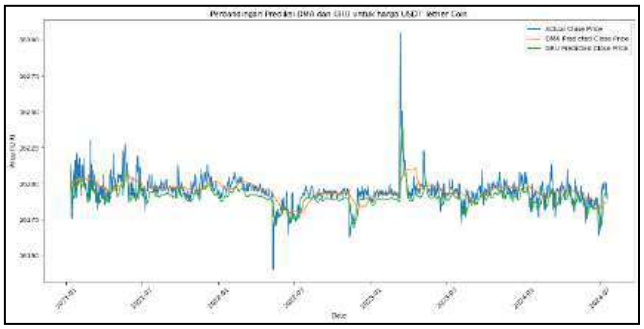


Fig 3. Comparison graph between original prices and DMA prediction results

The comparison between GRU and DMA reveals distinct strengths for each method. Overall, DMA demonstrates lower average errors (MAE: 5.494285, MAPE: 0.033918%) compared to GRU (MAE: 5.983858, MAPE: 0.036935%), making it more accurate on average. However, GRU is better at handling large errors, as indicated by its lower RMSE (8.530839 vs. DMA's 8.715295). In terms of bias, DMA has minimal systematic bias (MPE: -0.001262%), while GRU exhibits a slight positive bias (MPE: 0.028594%). DMA is more stable and suited for long-term trend analysis, but less responsive to rapid price changes. In contrast, GRU excels in capturing daily fluctuations, making it ideal for volatile market conditions. Both methods have their merits: DMA for stable, long-term predictions, and GRU for dynamic, high-volatility environments, with the choice depending on specific analytical needs.

In conclusion, DMA is better suited for long-term trend analysis with greater stability, while GRU is more effective at addressing significant and rapid price changes. The choice of method depends on the specific context of the analysis and the need for accurate price predictions. In difference with other research, *"Forecasting the Prices of Cryptocurrency on the Financial Markets"* by Yurii Pronchakov and Oleg Bugaienko and journal *"Parent Coin-Based Cryptocurrency Price Prediction Using Regression Techniques"* by Jg Aravindan and Ram Sankara.The first referenced study focuses on applying various moving average techniques, including SMA, EMA, and WMA, to forecast cryptocurrency prices.

It emphasizes that EMA performs best in volatile conditions due to its responsiveness to short-term changes, while SMA excels in long-term trend stability. However, the study does not explore the combination of moving averages with advanced machine learning models such as GRU, limiting its ability to handle dynamic and highly volatile market conditions comprehensively. In contrast, your research integrates DMA, a specific variation of moving averages, with GRU, offering a hybrid approach to balance trend stability with responsiveness to rapid fluctuations.

The second referenced study introduces regression-based methods for price prediction using parent coin relationships and machine learning models like Decision Trees and Ridge Regressors. While these methods are effective for capturing interdependencies among cryptocurrencies, they lack the temporal dynamics needed for precise predictions in real-time or short-term fluctuations. Compared to your study, which combines DMA for trend identification with GRU's strength in modeling sequential data, your approach provides a more holistic view by addressing both trend stability and temporal volatility, ensuring better adaptability in real-world scenarios.

This research builds upon these methodologies by not only validating the effectiveness of DMA in capturing trends but also integrating it with GRU to enhance predictions in volatile market conditions. The results from your study show that DMA has lower average errors (MAE: 5.494285, MAPE: 0.033918%) and stability in long-term analysis compared to GRU, which excels in handling large errors and rapid fluctuations. This hybrid model addresses the gaps in both referenced studies, as it successfully combines the strengths of traditional moving average techniques and advanced machine learning algorithms. This innovation positions your research as a significant step forward in cryptocurrency price forecasting, offering a balanced and adaptable solution for both stable and volatile market environments.

Comparison With LSTM Algorithm

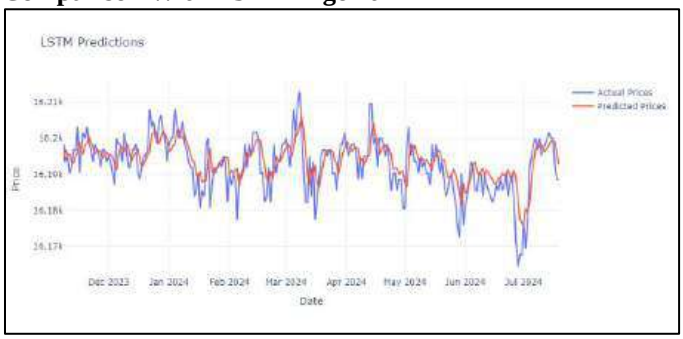


Fig 4. Comparison With LSTM

The comparison between the manual LSTM model, GRU, and DMA highlights the distinct strengths of each method. The manual LSTM model demonstrates superior accuracy, as evidenced by its lower MAE (4.3143) and MAPE (0.0266%) compared to GRU (MAE: 5.9839, MAPE: 0.0369%) and DMA (MAE: 5.4943, MAPE: 0.0339%).

This indicates that the LSTM model performs better in predicting average values with minimal deviations. Additionally, the RMSE of the manual LSTM model (5.6415) is significantly lower than both GRU (8.5308) and DMA (8.7153), showcasing its ability to handle large prediction errors more effectively. In terms of bias, the LSTM model's MPE

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2328, Copyright ©2025
Submitted : December 2, 2024, Revised : January 8, 2025, Accepted : January 28, 2025, Published : January 31, 2025

(0.0072%) is close to DMA (-0.0013%) and notably better than GRU (0.0286%), highlighting its minimal systematic bias.

While GRU and DMA excel in specific contexts, the manual LSTM model appears to combine their strengths. DMA is often preferred for its stability and suitability for long-term trend analysis, but it struggles with rapid market changes. GRU, on the other hand, excels in dynamic, high-volatility environments but introduces slightly higher average errors and bias. The manual LSTM model strikes a balance by providing both accurate average predictions and effective error handling, making it a robust choice for both stable and volatile market conditions. This versatility positions the LSTM model as a strong contender in predictive analytics across diverse scenarios.

V. CONCLUSION

The comparison between the Gated Recurrent Unit (GRU) and Double Moving Average (DMA) methods for predicting USDT (Tether Coin) prices highlights the unique strengths of each approach. GRU demonstrates superior performance in capturing price dynamics and trends in volatile markets, evidenced by its lower Root Mean Squared Error (RMSE) of 8.530839, indicating better handling of significant deviations. Conversely, DMA offers more stable predictions with a lower Mean Absolute Error (MAE) of 5.494285 and a Mean Absolute Percentage Error (MAPE) of 0.033918%, showcasing its ability to minimize average errors and filter out market noise effectively. When it comes to practical applications, GRU proves to be highly effective in responding to real-time market volatility, accurately forecasting rapid price changes.

In contrast, DMA is better suited for identifying and analyzing long-term market trends, making it particularly beneficial for traders seeking stability and consistent insights. Therefore, the choice between GRU and DMA depends on the specific objective, whether it is to adapt to short-term price fluctuations or to gain a comprehensive understanding of broader market movements.Future research should address the

limitations of this study by incorporating more granular data, such as intra-day price fluctuations, to better capture rapid market changes. Additionally, integrating external factors like social media sentiment analysis, macroeconomic indicators, and trading volume could enhance the predictive capability of GRU and DMA. Exploring hybrid models that combine these techniques with other machine learning approaches, such as transformers or ensemble models, may further improve accuracy in both stable and volatile market conditions. Finally, real-time testing on live markets with dynamic updates can validate the practicality and scalability of this approach for broader adoption in cryptocurrency trading platforms.

REFERENCES

[1] V. S. Bhilawadikar and E. Garg, "Investment attitude of millennials towards cryptocurrencies," *Soc. Bus.*, vol. 10, no. 4, pp. 383–410, 2020, doi: 10.1362/204440820x15929907056652.

[2] A. Kiky and D. Saputera, "Investasi Cryptocurrencies. Bertahan dengan Kerugian Cryptocurrencies?," *J. Perspekt.*, vol. 21, no. 2, pp. 120–128, 2023, doi: 10.31294/jp.v21i2.15657.

[3] M. Widyastuti and Y. B. Hermanto, "Cryptocurrency Analysis of Indonesian Market Education Facilities," *Int. J. Econ. Bussiness Account. Res.*, vol. 5, no. 2, pp. 534–546, 2021, [Online]. Available: http://jurnal.stie-aas.ac.id/index.php/IJEBAR/article/view/2321

[4] D. N. L. Tobing, "Indihome Product Sales Forecasting with the Double Moving Average and Double Exponential Smoothing Methods on PT. Telkom Witel Sumut Pematang Siantar," *Formosa J. Sci. Technol.*, vol. 1, no. 8, pp. 1201–1222, 2022, doi: 10.55927/fjst.v1i8.2281.

[5] M. Qamal, S. Hasan, and T. Wulandari, "Prediksi Jumlah Mahasiswa Baru Universitas Malikussaleh Menggunakan Metode Dma (Double Moving Average)," *TECHSI - J. Tek. Inform.*, vol. 11, no. 3, p. 353, 2019, doi: 10.29103/techsi.v11i3.1491.

[6] R. Dey and F. M. Salem, "Gate-Variants of Gated Recurrent Unit (GRU)," *Midwest Symp. Circuits Syst. Inst. Electr. Electron. Eng. Inc.*, vol. 784, no. 2017, pp. 1597–1600, 2017.

[7] H. Jdi and N. Falih, "Comparison of time series temperature prediction with auto-regressive integrated moving average and recurrent neural network," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 2, pp. 1770–1778, 2024, doi: 10.11591/ijece.v14i2.pp1770-1778.

[8] M. I. A. PRAMUDYA, "Penerapan Market Basket Analysis Menggunakan Proses Knowledge Discovery in Database (Kdd) Sebagai Strategi Penjualan …," *Repository.Radenintan.Ac.Id*, 2022.

[9] F. Sari, R. Fauziah, and H. D. ElIyany, "Prediction of Goods Stock System in Santi Electronic Shop With Double Moving Average Method," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 8, no. 3, pp. 375–382, 2022, doi: 10.33330/jurteksi.v8i3.1765.

[10] C. A. Anisiuba *et al.*, "Analysis of Cryptocurrency Dynamics in the Emerging Market Economies: Does Reinforcement or Substitution Effect Prevail?," *SAGE Open*, vol. 11, no. 1, 2021, doi: 10.1177/21582440211002516.

[11] A. Ferreira, "The Curious Case of Stablecoins-Balancing Risks and Rewards?," *J. Int. Econ. Law*, vol. 24, no. 4, pp. 755–778, 2021, doi: 10.1093/jiel/jgab036.

[12] M. A. Fauzi, N. Paiman, and Z. Othman, "Bitcoin and cryptocurrency: Challenges, opportunities and future works," *J. Asian Financ. Econ. Bus.*, vol. 7, no. 8, pp. 695–704, 2020, doi: 10.13106/JAFEB.2020.VOL7.NO8.695.

[13] D. Molitor, W. Raghupathi, V. Raghupathi, and A. Saharia, "Understanding Cryptocurrency," *Int. J. Blockchain Appl. Secur. Comput.*, vol. 1, no. 1, pp. 1–25, 2023, doi: 10.4018/ijbasc.331079.

[14] J. Almeida and T. C. Gonçalves, "A systematic literature review of investor behavior in the cryptocurrency markets," *J. Behav. Exp. Financ.*, vol. 37, p. 100785, 2023, doi: 10.1016/j.jbef.2022.100785.

[15] I. G. A. D. Perayunda and L. P. Mahyuni, "Faktor-Faktor Yang Mempengaruhi Keputusan Investasi Cryptocurrency Pada Kaum Milenial," *EKUITAS (Jurnal Ekon. dan Keuangan)*, vol. 6, no. 3, pp. 351–372, 2022, doi: 10.24034/j25485024.y2022.v6.i3.5224.

[16] J. Barth, P. Gardin, and B. Nguyen, "Stablecoins and short-term funding markets," 2023.

[17] J. A. Ripto, "Penerapan Gated Recurrent Unit untuk Prediksi Pergerakan Harga Saham pada Bursa Efek Indonesia," 2022, [Online]. Available: http://repository.ithb.ac.id/id/eprint/38/%0Ahttp://repository.ithb.ac.id/id/eprint/38/8/1118032_Paper-TA.pdf

[18] D. A. Aquinaldo, "Perbandingan Metode Multivariatif GRU dan VAR Berdasarkan Sentimen Investor dan Nilai Kurs Dollar Untuk Prediksi Harga Saham," vol. 3, no. April 2005, pp. 10238–10257, 2023.

[19] R. Rachman, "Penerapan Metode Moving Average Dan Exponential Smoothing Pada Peramalan Produksi Industri Garment," *J. Inform.*, vol. 5, no. 2, pp. 211–220, 2018, doi: 10.31311/ji.v5i2.3309.

[20] K. Mahendra, N. Satyahadewi, and H. Perdana, "Analisis Teknikal Saham Menggunakan Indikator Moving Average Convergence Divergence (Macd)," *Bimaster Bul. Ilm. Mat. Stat. dan Ter.*, vol. 11, no. 1, pp. 51–58, 2022, [Online]. Available: https://jurnal.untan.ac.id/index.php/jbmstr/article/view/51602

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2328, Copyright ©2025
Submitted : December 2, 2024, Revised : January 8, 2025, Accepted : January 28, 2025, Published : January 31, 2025

The Effect of Chatbot Usage on Customer Satisfaction: A Quantitative Study of Shopee, Tokopedia, and Lazada Using SmartPLS

Mira Afrina[1]*, Naretha Kawadha Pasemah Gumay[2], Ardina Ariani[3] Mukhlis Febriady[4]
Department of Information Systems, Faculty of Computer Science, Sriwijaya University [1],[2]
Department of Informatics Engineering, Faculty of Computer Science, Sriwijaya University [3]
Institute of Technology and Science Nahdlatul Ulama Sriwijaya South Sumatera [4]
miraafrina81@gmail.com[1], narethakawadha@unsri.ac.id[2], ardinaariani@unsri.ac.id[3], mfrebriady5@gmail.com[4]

***Abstract*— With the increasing growth of e-commerce, it is important to identify the features available in e-commerce applications that can provide customer satisfaction. One of the features in e-commerce is the chatbot. Chatbots in e-commerce can provide various services to users, such as assistance in product search, ordering, product information, payment processing, customer support, and more. This research aims to analyze and understand how the response quality of each chatbot in e-commerce platforms such as Shopee, Tokopedia, and Lazada affects e-commerce user satisfaction. This study employs a quantitative methodology, integrating data analysis conducted through the SmartPLS 4.1 software. The research results show that the chatbot in Shopee platform has a impact on customer satisfaction. The same goes for chatbot in Tokopedia platform, but there are two variables that do not have a direct impact, there are information quality and waiting time. Meanwhile, chatbot in Lazada platform does not affect customer satisfaction. The findings of this research should reveal new strategies for leveraging chatbot technology to better satisfy customers in e-commerce environments, as well as lay the groundwork for further research on how artificial intelligence can shape customer experiences in the future.**

Keywords*— Chatbot, E-Commerce, User Satisfaction, Customer Service, SmartPLS

I. INTRODUCTION

The increasing utilization of communication technology and online platforms in economic activities has become an indicator of the development of the digital economy in Indonesia at present. The movement of the digitalization era began with the Industrial Revolution 3.0, where the role of humans was replaced by the advent of computers. And now the world has just entered the Industrial Revolution 4.0. This era is marked by the ability to retrieve or exchange data in real time with an internet connection [1]. The digital economy is capable of transforming the economic activities of society and businesses from initially manual processes to fully automated ones.

The growth of e-commerce is rapidly advancing, marked by the growth of online transaction volumes, market expansion into various product types, enhanced user experiences through personalization and technology, increased popularity of shopping via mobile devices, data analytics, emphasis on data security and privacy, and a focus on sustainable business practices [2]. Based on data compiled by Databoks, the Shopee and Tokopedia platforms still lead the e-commerce market in Indonesia, followed by Lazada in third place. Based on the average number of visitors per month, Shopee reaches 195.8 million visitors each month, followed by Tokopedia with 104.2 million visitors, and Lazada with a total of 63.5 million visitors per month.

Chatbots have ability to handle multiple user issues simultaneously and save on customer service costs [3]. Chatbots are no longer just considered assistants, their interaction approach has brought them closer to users as friendly companions [4][5]. Chatbots can be a digital era solution to address the issue of improving customer service [7].

The research conducted focuses on analyzing how the use of chatbots from each e-commerce platform, such as Shopee, Tokopedia, and Lazada affects e-commerce user satisfaction.

Although there have been several previous studies discussing the use of chatbots, research comparing their impact across major e-commerce platforms in Indonesia remains limited. For instance, studies by Dey *et al.* (2020) and Wibowo *et al.* (2020) focus solely on the design and implementation of chatbots, particularly regarding the quality of information provided to customers [8][10].

II. THEORETICAL FOUNDATION

A. Customer Service

In order to have loyal customers, the service industry, especially e-commerce, must meet the service quality expected by customers, so that customers can feel satisfied and have trust in the services provided [23]. With the large number of customers shopping on e-commerce platforms, it is necessary to support it with good service quality so that customers feel satisfied after making online transactions through the e-commerce platform [34]. In the development of the business world, service quality is one way to achieve excellence by consistently applying the right service quality and product quality, which can later influence customer satisfaction and create customer loyalty [10].

The research by Rita *et al.* (2019), when consumers are

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2312, Copyright ©2025
Submitted : January 16, 2025, Revised : January 21, 2025, Accepted : January 28, 2025, Published : January 31, 2025

satisfied with the products or services provided by an online store, they are likely to revisit the store and purchase products from the same store again. This study also proves that consumers will spread information about their satisfaction through word of mouth to relatives or close friends. [12].

B. E-Commerce

The use of the internet is not only limited to the exchange of information but has also expanded into economic activities such as buying and selling. Currently, the buying and selling process can be conducted through electronic devices such as smartphones, utilizing the electronic network known as E-Commerce [13]. Currently, e-commerce has entered and developed significantly in Indonesia and is used by the public, such as Shopee, Tokopedia, and Lazada.

1. Shopee

Shopee is one of the e-commerce platforms that has successfully established dominance in the Indonesian market. Shopee is a marketplace founded by Forrest Li, which was first launched to the public in 2015. Shopee could only used and accessed in Indonesia since December 2015 [14]. Shopee is a mobile marketplace application with a Consumer-to-Consumer (C2C) model, which ensures a secure and convenient way to buy and sell. It features a wide range of products that cater to everyday needs, including clothing, electronics, beauty products, home appliances, and shopping vouchers. Furthermore, Shopee frequently offers substantial promotions to entice customers, with one of the most favored being free shipping [12].

2. Tokopedia

Tokopedia is an online marketplace that was officially introduced to the public in 2009, under the management of PT Tokopedia. To attract more consumers, Tokopedia employs diverse marketing strategies, one of which is collaborating with several public figures as their brand ambassadors. Among the various online marketplaces available, Tokopedia is an Indonesian-made marketplace that is highly favored by the Indonesian public [15].

3. Lazada

Lazada Group was founded in 2012 by Rocket Internet, a global investment and technology startup development company. Lazada is a platform that acts as an intermediary between sellers and buyers using an internet connection. This provides an opportunity for sellers to transition from traditional sales to digital sales. Meanwhile, it also offers an online shopping experience for buyers instead of traditional shopping [16].

C. Information Quality

High-quality information is more valuable to consumers, while information is considered low-quality if it does not provide benefits to consumers [26]. The quality of information displayed on e-commerce platforms helps predict product quality and availability, provides accurate information, clear and detailed transactions, as well as reliable information that meets user expectation [27].

D. User Experience

User experience emerges as a promising element in the analysis of consumption related to technology, especially in connection with the development of digital products [28]. User experience is also an important part of buying and selling transactions in e-commerce, which will influence consumer purchasing interest [29]. User experience is defined as an individual's perception and response from using a particular product, system, or service [30]. Purchasing behavior will be heavily influenced by the user experience in the future. As consumers evaluate online shopping experiences, they will take into account their perception of product information, payment methods, delivery, services available, features, risks involved, and privacy and security concerns [32]. This emphasizes that satisfied users are likely to become loyal customers. In other words, improving comfort and user experience on e-commerce platforms is a strategic tool to extend business relationships with customers and encourage more frequent purchases [31].

E. Waiting Time

Waiting time encompasses the total period customers must wait from the point of ordering until they obtain their desired product or service [32]. The implementation of chatbots can usually resolve customer issues without wasting the customer's time. One of the benefits of using chatbots is that they can provide instant and automatic responses to user inquiries. With this capability, they can help reduce user wait times and expedite problem resolution, making them a highly useful feature in customer service on an online platform such as e-commerce [33]. Research by Ramadhani *et al.* (2024) indicates that the waiting time of chatbots can affect customer satisfaction. This can be caused by the time gap between when consumers request information and when they receive the desired information, which is a crucial period in the considering stage [26].

F. User Satisfaction

Consumer interest in using e-commerce can be influenced by satisfaction in conducting online transactions, which is a primary indicator for consumers to prefer one marketplace over another [17]. Setyaningsih explains that satisfaction is the feeling experienced when the outcome of evaluating what has been received aligns with expectations, encompassing both the purchase decision and the needs and desires associated with the purchase [17].

The level of satisfaction can vary between one customer and another. Customer expectations can be influenced by previous purchase experiences, opinions from close acquaintances, and the information provided [38]. As written by marketing experts Kotler and Keller in their book, successful sellers are those who carefully reinforce customer satisfaction [18].

G. Chatbot

Chatbots are built using AI and Machine Learning technologies, enabling them to behave like humans. With the availability of data possessed by the chatbot, it is capable of providing the best recommendations to customers and can answer follow-up questions that encourage further explanations [19][20]. According to Følstad *et al.* (2021), customers are able to interact with chatbots to access information and services utilizing everyday language, which has been applied across various sectors, including customer service, e-commerce,

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2312, Copyright ©2025
Submitted : January 16, 2025, Revised : January 21, 2025, Accepted : January 28, 2025, Published : January 31, 2025

healthcare, and education [21].

Chatbots, often referred to as conversational interfaces, offer an innovative approach to interacting with computer systems. In the past, finding answers to questions from software typically required the use of search engines. However, it is possible for users to communicate with chatbots in the same manner as they would with a human. In addition to voice chatbots, chatbots are now also being used with computer chat platforms at an increasing rate [22].

H. SmartPLS

One of the latest and easiest applications for processing statistical data is SmartPLS [32]. SmartPLS, or Smart Partial Least Square, is one of the tools for processing statistical data. One of the uses of SmartPLS is to test the relationships between variables, namely between latent variables, latent variables with indicators, or latent variables with manifests. This software is recommended for a limited sample size [23].

III. METHODOLOGY

1. Research Type

This research uses a quantitative approach that will specifically contribute by analyzing the impact of chatbot usage in e-commerce (Shopee, Tokopedia, and Lazada) on customer satisfaction using the parametric procedure of PLS-MGA. PLS-MGA is considered more appropriate for exploring and comparing differences between two or more variables [27]. The proposed research framework follows these steps:

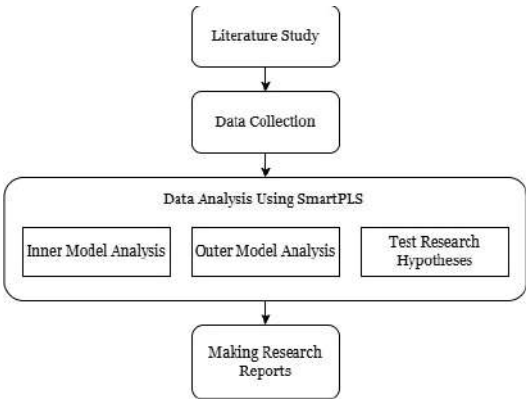


Fig. 1. *Research Framework.*

The first step is to gather information relevant to the research topic from various sources, such as books, articles, journals, and previous studies. The next step involves collecting data by distributing questionnaires to the research respondents. Once the data distribution results are gathered, the next step is to analyze the data using the SmartPLS 4.0.9.5 application. The final step is to create a research report by compiling all the data and information obtained throughout the research process.

2. Population and Sample

A population generally refers to a group of objects or subjects possessing specific qualities and characteristics identified by the researcher for analysis and conclusion drawing [36]. In this study, the research population included all students from the Faculty of Computer Science at Sriwijaya University.

Whereas a sample is a part of a population [36]. In the framework of this research, the sample includes users of the e-commerce platforms Shopee, Tokopedia, and Lazada who meet the necessary criteria for the study such as having an e-commerce account (Shopee, Tokopedia, Lazada), actively using e-commerce for at least the last 1 month, actively interacting with chatbots. The total sample required is 100 respondents, determined using the Slovin's formula as follows:

$$n = \frac{N}{1+Ne^2} \quad (1)$$

3. Research Instruments

The research instruments used in this study are information quality, user experience, wait time, customer service, and customer satisfaction.

TABLE I. RESEARCH VARIABLES

Variable	Code	Questions
Informasion Quality	KI.1	How important is the accuracy of the information and answers provided by chatbot for your satisfaction?
	KI.2	How easy is it for you to find the information you need through chatbot?
	KI.3	How relevant is the information provided by chatbot to your needs?
	KI.4	Is the information provided by chatbot easy to understand?
	KI.5	How do you assess the consistency of the answers provided by chatbot in various conversations or questions?
User Experience	PP.1	How often do you use chatbots when shopping online?
	PP.2	How was your experience using chatbot?
	PP.3	Did chatbot help you resolve the problems or questions you raised?
	PP.4	Does chatbot provide good service during the interaction?
	PP.5	How do you assess the ease of use of chatbot?
Waiting Time	WT.1	How important is the response speed of chatbot for your satisfaction in online shopping?
	WT.2	How quickly does chatbot respond to your questions?
	WT.3	Is chatbot available at all times when you need it?
	WT.4	Do you feel that the response time of chatbot is adequate?
	WT.5	How do you rate the response speed of chatbot compared to human-based customer service in responding to your questions?
Customer Service	LP.1	How well does chatbot understand your questions or requests?
	LP.2	How easy is it for you to find chatbot feature on the e-commerce platform?
	LP.3	Is chatbot interface easy for you to understand?
	LP.4	Does your interaction with chatbot make you feel satisfied with the customer service of this platform?
	LP.5	Does chatbot use clear and easy-to-understand language?
Customer Satisfaction	KP.1	How satisfied are you with chatbot's ability to provide information and answer your questions?
	KP.2	To what extent do you feel chatbot influences your purchasing decisions?
	KP.3	Can chatbot enhance your shopping experience?
	KP.4	Does chatbot help you resolve the issues you encounter while shopping?

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2312, Copyright ©2025
Submitted : January 16, 2025, Revised : January 21, 2025, Accepted : January 28, 2025, Published : January 31, 2025

	KP.5	How would you assess your overall satisfaction with the service offered by the chatbot?

IV. RESULT

A. Outer Model

1) Outer Loadings and Convergend Validity (AVE)

An outer loading score exceeding 0.70 (>0.70) indicates that the validity indicator is acceptable. To meet the minimum criteria, AVE must be greater than 0.50 (>0.50). If the outer loading value is below 0.70 but remains above the minimum threshold of 0.40 (>0.40), and the AVE value exceeds 0.50 (>0.50), the indicator can still be considered valid. If less than 0.40, it must be removed [25].

TABLE III. OUTER LOADINGS AND AVE

Laten Variable	Construck Variable	Loading (>0.70)	AVE (>0.5)
Information Quality	KI.1	0.742	0.567
	KI.2	0.823	
	KI.3	0.720	
	KI.4	0.702	
	KI.5	0.772	
User Experience	PP.1	0.794	0.584
	PP.2	0.778	
	PP.3	0.746	
	PP.4	0.757	
	PP.5	0.742	
Waiting Time	WT.1	0.783	0.609
	WT.2	0.803	
	WT.3	0.817	
	WT.4	0.735	
	WT.5	0.762	
Customer Service	LP.1	0.790	0.586
	LP.2	0.761	
	LP.3	0.729	
	LP.4	0.791	
	LP.5	0.755	
Customer Satisfaction	KP.1	0.752	0.577
	KP.2	0.746	
	KP.3	0.756	
	KP.4	0.791	
	KP.5	0.753	

Construct loading values are higher than 0.70, and AVE values all exceed 0.50. Therefore, the factor loading values and AVE of all variables and indicators meets the validity criteria.

2) Construct Reliability

To ensure the reliability of constructs, both composite reliability and Cronbach's alpha should exceed 0.70. A construct variable is considered reliable if both above this threshold [25].

TABLE IIIII. CONTRUCT RELIABILITY

Laten Variable	Cronbach's Alpha	Composite Reliability (rho_c)	Information
Information Quality	0.810	0.867	Reliable
User Experience	0.823	0.875	Reliable
Waiting Time	0.840	0.886	Reliable
Customer Service	0.825	0.876	Reliable
Customer Satisfaction	0.817	0.872	Reliable

According to the test results, leads to these conclusions that both values for all variables exceed 0.70. Based on this assessment, the research model is suitable for further analysis.

3) Discriminant Validity Heterotrait Monotrait (HTMT)

HTMT is the ratio of the inter-construct correlation to the intra-construct correlation. According to the HTMT method, the disattenuated correlation between two constructs can be calculated if both are perfectly measured. A construct variable is considered a latent variable if its value is less than 0.90, and each construct variable can form a latent variable if its value is less than 0.90 [25].

TABLE IVV. HETEROTRAIT MONOTRAIT

Path Coefficient	HTMT Ratio
Information Quality <-> Customer Satisfaction	0.242
Customer Service <-> Customer Satisfaction	0.412
Customer Service <-> Information Quality	0.066
User Experience <-> Customer Satisfaction	0.387
User Experience <-> Information Quality	0.133
User Experience <-> Customer Service	0.101
Waiting Time <-> Customer Satisfaction	0.343
Waiting Time <-> Information Quality	0.139
Waiting Time <-> Customer Service	0.163
Waiting Time <-> User Experience	0.144

HTMT calculation value is less than 0.90. This value indicates that each construct variable is capable of forming its own latent variable and satisfies the HTMT criteria.

B. Inner Model

1) Collinearity Assessment

In order to test a model further, the VIF value should be below 5.0, which indicates that the model is free from multicollinearity issues among all predictors and responses [25].

TABLE V. COLLINEARITY ASSESSMENT VIF

Patch Coefficient	VIF	Multikoleniaritas
Information Quality -> Customer Satisfacttion	1.019	No
Customer Service -> Customer Satisaction	1.022	No
User Experience -> Customer Satisfaction	1.024	No
Waiting Time -> Customer Satisfaction	1.041	No

From the table above, it is evident that the VIF values for each construct variable are below 5.0 (<5.0). Therefore, all variables do not show any indications of multicollinearity and are suitable for use in further analysis.

2) Coefficient of Determination (R^2)

The R^2 is used to evaluate the precision of predictions. Generally, an R^2 value of 0.75 signifies high predictive accuracy, 0.50 indicates moderate, and 0.25 implies low [25].

TABLE VI. COEFFICIENT OF DETERMINATION

Latent Variable	R^2	R^2 Adjusted	Information
Customer Satisfaction	0.277	0.268	Moderate

According to the test results, leads to these conclusions that

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2312, Copyright ©2025
Submitted : January 16, 2025, Revised : January 21, 2025, Accepted : January 28, 2025, Published : January 31, 2025

the accuracy of the User Satisfaction model estimation is 0.277. Based on this value, it has a Moderate estimation accuracy. Predictive Relevance (Q^2).

The blindfolding procedure is applied. Q^2 value of 0.02 (small), 0.15 (moderate), and 0.35 (high) predictive relevance [25].

TABLE VII. PREDICTIVE RELEVANCE

Latent Variable	SSO	SSE	Q^2 (=1-SSE/SSO)
Customer Satisfaction	1500	1272.721	0.152
Information Quality	1500	1500	0
Customer Service	1500	1500	0
User Experience	1500	1500	0
Waiting Time	1500	1500	0

According to the test results, leads to these conclusions that the predictive relevance Q^2 value for the User Satisfaction construct model is influenced by Information Quality, User Experience, Wait Time, and Customer Service at a rate of 0.152, which is classified as having moderate predictive relevance.

3) Effect Size (f^2)

In contrast to R^2, f^2 is specifically concerned with each exogenous variable. Generally, a value of 0.02 (small), 0.15 (medium), 0.35 (large) effect size. Here is f^2 value table [25].

TABLE VIII. EFFECT SIZE

Path Coefficient	f-square	Information
Information Quality -> Customer Satisfaction	0.031	Small
Customer Service -> Customer Satisfaction	0.119	Small
User Experience -> Customer Satisfaction	0.094	Small
Waiting Time -> Customer Satisfaction	0.054	Small

According to the test results, leads to these conclusions:

- *Information Quality -> Customer Satisfaction has a constructive model value with an affect size f^2 of 0.031, which is classified as having a small estimated value.*
- *Customer Service -> Customer Satisfaction has a constructive model affect size f^2 value of 0.119 and is categorized as having a small estimated value.*
- *User Experience -> Customer Satisfaction has a constructive model affect size f^2 value of 0.094 and is classified as having a small estimated value.*
- *Waiting Time -> Customer Satisfaction has a constructive model affect size f^2 value of 0.054 and is classified as having a small estimated value.*

C. Hypothesis Testing of the Research Model

The analysis of structural model coefficients is conducted to test hypotheses by identifying which relationships have a significant impact. If p-value is less than the significance level (0.05), the relationship is considered significant. Conversely, a relationship is considered not significant when the p-value exceeds 0.05 [25].

TABLE IX. HYPOTHESIS TESTING

Hypothesis	Path Coefficient	Original Sample (O)	T Stat	P Values	Information
H1	Information Quality -> Customer Satisfaction	0.150	2.518	0.012	Accepted
H2	User Experience -> Customer Satisfaction	0.264	4.435	0.000	Accepted
H3	Waiting Time -> Customer Satisfaction	0.201	3.240	0.001	Accepted
H4	Customer Service -> Customer Satisfaction	0.296	5.480	0.000	Accepted

From the test results in the table above, leads to these conclusions:

- *Information Quality significant positive impact on Customer Satisfaction. Therefore, **H1 is accepted** and H0 is rejected.*
- *Experience has a significant impact on Customer Satisfaction. Therefore, **H2 is accepted** and H0 is rejected.*
- *Waiting Time has a significant positive effect on Customer Satisfaction. Therefore, **H3 is accepted** and H0 is rejected.*
- *Customer Service has a significant positive impact on Customer Satisfaction. Therefore, **H4 is accepted** and H0 is rejected.*

D. Multigroup Analysis (PLS-MGA)

Multigroup Analysis (MGA) in the context of PLS-SEM is a technique used to evaluate differences in structural relationships between sample groups. The goal is to determine whether there are significant differences in the influence between latent variables across different groups, such as based on age, gender, geographic region, or other characteristics. PLS-MGA allows researchers to compare PLS-SEM models across different groups and test parameter differences, such as path coefficient values, to determine whether these differences are statistically significant. The testing criterion is that there is a significant difference if the p-value < a (0.05) [25]. The results of MGA-PLS can be seen as follows:

TABLE X. MULTIGROUP ANALYSIS

Path Coefficient	2-tailed (Lazada v Shopee) p value	2-tailed (Lazada v Tokopedia) p value	2-tailed (Shopee v Tokopedia) p value
Information Quality -> Customer Satisfaction	0.318	0.829	0.135
User Experience -> Customer Satisfaction	0.787	0.251	0.251
Waiting Time -> Customer Satisfaction	0.092	0.345	0.001
Customer Service -> Cusomer Satisfaction	0.634	0.013	0.057

According to the test results, leads to these conclusions:

Lazada v Shopee

- *There is no significant differences in the Direct Influence Path Coefficient between Information Quality and Customer Satisfaction, with p-value 0.318, greater than 0.05.*

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2312, Copyright ©2025
Submitted : January 16, 2025, Revised : January 21, 2025, Accepted : January 28, 2025, Published : January 31, 2025

- *There is no significant differences in the Direct Influence Path Coefficient between User Experience and Customer Satisfaction, with p-value 0.787, greater than 0.05.*
- *There is no significant differences in the Direct Influence Path Coefficient between Waiting Time and Custome Satisfaction, with p-value 0.787, greater than 0.05.*
- *There is no significant differences in the Direct Influence Path Coefficient between Customer Service and Customer Satisfaction, with p-value 0.787, greater than 0.05.*

Lazada v Tokopedia

- *There is no significant differences in the Direct Influence Path Coefficient between Information Quality and Customer Satisfaction, with p-value 0.829, greater than 0.05.*
- *There is no significant differences in the Direct Influence Path Coefficient between User Experience and Customer Satisfaction, with p-value 0.251, greater than 0.05.*
- *There is no significant differences in the Direct Influence Path Coefficient between Waiting Time and Customer Satisfaction, with p-value 0.345, greater than 0.05.*
- *There is a significant differences in the Direct Influence Path Coefficient between Customer Service and Customer Satisfaction, with p-value 0.013, less than 0.05.*

Shopee v Tokopedia

- *There is no significant differences in the Direct Influence Path Coefficient between Information Quality and Customer Satisfaction, with p-value 0.135, greater than 0.05.*
- *There is no significant differences in the Direct Influence Path Coefficient between User Experience and Customer Satisfaction, with p-value 0.251, greater than 0.05.*
- *There is a significant differences in the Direct Influence Path Coefficient between Waiting Time and Customer Satisfaction, with p-value 0.001, less than 0.05.*
- *There is no significant differences in the Direct Influence Path Coefficient between Customer Service and Customer Satisfaction, with by p-value 0.057, greater than 0.05.*

E. Multigroup Analysis Hypothesis Testing

1) Shopee

TABLE XI. MULTIGROUP SHOPEE

Hypothesis	Path Coefficient	Original Sample (O)	T Stat	P Values	Information
H1	Information Quality -> Customer Satisfaction	0.295	2.850	0.004	Accepted
H2	User Experience -> Customer Satisfaction	0.296	4.223	0.000	Accepted
H3	Waiting Time -> Customer Satisfaction	0.400	4.344	0.000	Accepted
H4	Customer Service -> Customer Satisfaction	0.265	3.034	0.002	Accepted

From the test results in the table above, leads to these conclusions:

- *There is a significant positive influence of Information Quality on Customer Satisfaction.* ***H1 is accepted*** *and H0 is rejected.*
- *There is a significant positive influence of User Experience on Customer Satisfaction. Therefore,* ***H2 is accepted*** *and H0 is rejected.*
- *There is a significant positive influence of Waiting Time on Customer Satisfaction. Therefore,* ***H3 is accepted*** *and H0 is rejected.*
- *Customer Service has a significant positive impact on Customer Satisfaction. Consequently,* ***H4 is accepted*** *while H0 is rejected.*

2) Tokopedia

TABLE XII. MULTIGROUP ANALYSIS TOKOPEDIA

Hypothesis	Path Coefficient	Original Sample (O)	T Stat	P Values	Information
H1	Information Quality -> Customer Satisfaction	0.083	0.761	0.447	Rejected
H2	User Experience -> Customer Satisfaction	0.439	4.354	0.000	Accepted
H3	Waiting Time -> Customer Satisfaction	0.009	0.099	0.921	Rejected
H4	Customer Service -> Customer Satisfaction	0.502	5.818	0.000	Accepted

According to the test results, leads to these conclusions:

- *There is a positive but insignificant of Information Quality on Customer Satisfaction. Therefore,* ***H1 is rejected*** *and H0 is accepted.*
- *There is a significant positive influence of User Experience on Customer Satisfaction. Therefore,* ***H2 is accepted*** *and H0 is rejected.*
- *There is a positive but insignificant effect of Waiting Time on Customer Satisfaction. Therefore,* ***H3 is rejected*** *and H0 is accepted.*
- *There is a significant positive influence of Customer Service on Customer Satisfaction. Therefore, H4 is accepted and H0 is rejected.*

3) Lazada

TABLE XIII. MULTIGROUP ANALYSIS LAZADA

Hypothesis	Path Coefficient	Original Sample (O)	T Stat	P Values	Information
H1	Information Quality -> Customer Satisfaction	0.096	0.557	0.578	Rejected
H2	User Experience -> Customer Satisfaction	0.167	0.753	0.451	Rejected
H3	Waiting Time -> Customer Satisfaction	0.150	0.989	0.323	Rejected
H4	Customer Service -> Customer	0.197	1.728	0.084	Rejected

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2312, Copyright ©2025
Submitted : January 16, 2025, Revised : January 21, 2025, Accepted : January 28, 2025, Published : January 31, 2025

	Satisfaction				

According to the test results, leads to these conclusions:

- *There is a positive but insignificant effect of Information Quality on Customer Satisfaction. Therefore,* ***H1 is rejected*** *and H0 is accepted.*
- *There is a positive but insignificant effect of User Experience on Customer Satisfaction. Therefore,* ***H2 is rejected*** *and H0 is accepted.*
- *There is a positive but insignificant effect of Waiting Time on Customer Satisfaction. Therefore,* ***H3 is rejected*** *and H0 is accepted.*
- *There is a positive but insignificant effect of Customer Service on Customer Satisfaction. Therefore,* ***H4 is rejected*** *and H0 is accepted.*

According to the test results, leads to these conclusions that Shopee chatbot service has an impact on user satisfaction. This indicates that the service provided by the Shopee chatbot is already running effectively, starting from the quality of information provided, user experience, relatively faster response time, and the features offered are quite complex and easy to use.

On the other hand, Tokopedia chatbot shows the same thing. However, there are still shortcomings in the quality of the information provided, as well as a relatively long wait time for responses.

Meanwhile, chatbot on Lazada does not affect user satisfaction, which could be due to the quality of the information provided not matching user complaints, the user experience not being very good while using the chatbot service, the relatively longer wait time for responses, and the features provided still having shortcomings. Additionally, this could also be due to the lack of user interest in using the Lazada platform, especially the chatbot service.

V. DISCUSSIONS

This study is expected to make a significant contribution to the fields of digital marketing and customer service, particularly in the context of chatbot usage on e-commerce platforms. While many previous studies have examined chatbot usage in general, this research delves deeper into understanding how the effectiveness of chatbots can vary across leading e-commerce platforms in Indonesia, such as Shopee, Tokopedia, and Lazada.

This study highlights that chatbots can play a significant role in enhancing customer satisfaction on e-commerce platforms, especially on Shopee and Tokopedia. The findings reveal that Shopee's chatbot consistently has a positive impact on customer satisfaction.

Meanwhile, Tokopedia demonstrates a more limited positive impact, with two variables that are waiting time and information quality that showing no significant positive influence. In the context of e-commerce, waiting time is often associated with customer service response speed. With respect to waiting time, e-commerce platforms like Tokopedia have largely automated their customer service processes using chatbot features. According to Song et al. (2019), an artificial intelligence assistant (chatbot) primarily functions to automatically respond to customer inquiries [35]. Based on this, customers may have become accustomed to minimal waiting times in e-commerce transactions, making this factor less influential in shaping their satisfaction.

Additionally, information quality in e-commerce encompasses the accuracy, completeness, and relevance of the product information provided. The insignificance of information quality in Tokopedia's chatbot could be attributed to customers prioritizing other features. Customers may rely more on reviews and ratings from other buyers as their primary source of information rather than the product descriptions provided by sellers. As research by Sudirjo et al. (2023), consumers often seek information in various ways, including through previous customer reviews of purchased products [36]. This reliance on social proof may diminish the importance of information quality as a determinant of satisfaction. Therefore, Tokopedia should consider optimizing its chatbot, both in terms of its ability to answer inquiries and its response speed, to maximize customer satisfaction.

For Lazada, this study found no significant impact of chatbot usage on customer satisfaction. User expectations play a critical role in shaping perceptions of chatbot effectiveness. Modern e-commerce users often expect chatbots to efficiently resolve complex issues, yet many chatbots are limited to providing predefined responses or handling basic queries. As Reinkemeier and Gnewuch (2022) emphasize, users value chatbots more when they exhibit a high degree of interactivity and adaptability [37], attributes that may not be consistently present in Lazada's chatbot.

Therefore, Lazada should evaluate and potentially revise its strategy for utilizing chatbots. One possible approach is to enhance the quality of chatbot interactions, improve the quality of information provided, and ensure faster and more accurate responses.

Thus, it is expected that e-commerce companies can utilize these findings to design more effective customer service strategies, focusing on enhancing digital interactions to boost customer loyalty and satisfaction.

VI. CONCLUSION

After conducting tests, analysis, and discussions to determine the impact of chatbot usage on e-commerce users (Shopee, Tokopedia, and Lazada), it can be concluded that Chatbot service on Shopee platform for all its variables significantly shows a positive impact on user satisfaction. This is consistent with the test results showing P Values for the information quality variable at 0.004, user experience at 0.000, wait time at 0.000, and customer service at 0.002. The values of each of these variables are less than 0.05. Besides that, chatbot service on Tokopedia platform also significantly shows a positive influence on user satisfaction for the user experience and customer service variables. This is evidenced by the P values of the user experience variable being 0.000 and the customer service variable being 0.000. The values of both variables are less than 0.05. Meanwhile, the information quality and waiting time variables show no significant influence on user satisfaction. However, chatbot service on Lazada platform

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2312, Copyright ©2025
Submitted : January 16, 2025, Revised : January 21, 2025, Accepted : January 28, 2025, Published : January 31, 2025

for all its variables shows an not significant effect on user satisfaction. This is indicated by the P values of each variable being greater than 0.05.

This study focuses only on three major e-commerce platforms in Indonesia (Tokopedia, Shopee, and Lazada), which means the results may not be generalizable to other e-commerce platforms with different characteristics. Additionally, the variables examined include information quality, user experience, waiting time, and customer service in relation to customer satisfaction. However, other factors such as price, promotions, user reviews, and additional elements that may also influence customer satisfaction have not been explored in this study. This study should reveal new strategies by connecting the variables of waiting time, user experience, information quality, and customer service with customer satisfaction in the context of e-commerce chatbots. Future research could further examine the implementation of chatbots as a critical element in customer experience management. Furthermore, future research are expected to evaluate additional variables that were not included in this research.

REFERENCES

[1] Handayani, N. L. P. (2022). E-Commerce Sebagai Penunjang Ekonomi Digital di Jawa Tengah. Jurnal Ilmiah Manajemen, Bisnis Dan Kewirausahaan, 2(1), 9–14. https://doi.org/10.55606/jurimbik.v2i1.103.

[2] Bawack, R. E., Wamba, S. F., Carillo, K. D. A., & Akter, S. (2022). Artificial intelligence in E-Commerce: a bibliometric study and literature review. In Electronic Markets (Vol. 32, Issue 1). Springer Berlin Heidelberg. https://doi.org/10.1007/s12525-022-00537-z.

[3] Madasamy, S. and Aquilanz, L. L. C. (2023) 'The Evolution of Chatbots: Cloud and Ai Synergy in Banking Customer Interactions', Journal of Emerging Technologies and Innovative Research, 10(October 2023), pp. 127–137.

[4] Adamopoulou, E., & Moussiades, L. (2020). An Overview of Chatbot Technology. In IFIP Advances in Information and Communication Technology: Vol. 584 IFIP. Springer International Publishing. https://doi.org/10.1007/978-3-030-49186-4_31.

[5] Zumstein, D. and Hundertmark, S. (2020) 'Chatbots – An Interactive Technology for Personalized Communication , Chatbots – An Interactive Technology For Personalized Communication ', IADIS International Journal, 15(February 2018), pp. 96–109.

[6] Trivedi, A., Gor, V., & Thakkar, Z. (2019). Chatbot generation and integration: A review. International Journal of Advance Research, 5(2), 1308–1311. www.IJARIIT.com

[7] Sabila, A. R. and Kusumaningrum, L. (2020). Analisis Kualitas Layanan E-Commerce Shopee Dalam Meningkatkan Kepercayaan Dan Kepuasan Pelanggan (Studi Kasus: Pelanggan Shopee Di Kota Tangerang 2020). Jurnal Ilmiah Sekolah Tinggi Teknologi Informasi NIIT, 16(2), pp. 72–80.

[8] Dey, N., Mahelle, P., Mohd Shafi, P., Kimabahune, V. and Hassanien, A. E. (2020) Internet of Things, Smart Computing and Technology: A Roadmap Ahead, Studies in Systems, Decision and Control.

[9] Indah Yani, M. and Sugiyanto, S. (2022). Pengaruh Kualitas Layanan, Kepercayaan, Citra Merek Dan Kepuasan Pelanggan Terhadap Loyalitas Pelanggan Edwin Jeans Indonesia Pada E-Commerce Shopee. SINOMIKA Journal: Publikasi Ilmiah Bidang Ekonomi dan Akuntansi, 1(3), pp. 443–472.

[10] Wibowo, B., Clarissa, H. and Suhartono, D. (2020) 'The Application of Chatbot for Customer Service in E-Commerce', Engineering, MAthematics and Computer Science (EMACS) Journal, 2(3), pp. 91–95.

[11] Rita, P., Oliveira, T. and Farisa, A. (2019). The impact of e-service quality and customer satisfaction on customer behavior in online shopping. Heliyon. Elsevier Ltd, 5(10), p. e02690.

[12] Sim, L. S., Christopher, M., Nafitra, H. D., Panjaitan, J. S. V., & Rakhmawati, N. A. (2023). Pemilihan Platform E-Commerce Pada Kalangan Mahasiswa Its Dengan Menggunakan Metode Weighted Product. Jurnal Teknoinfo, 17(1), 36. https://doi.org/10.33365/jti.v17i1.2252.

[13] Chong, D., & Ali, H. (2022). Competitive Strategy, Competitive Advantages, dan Marketing Performance pada E-Commerce Shopee Indonesia. Jurnal Kewirausahaan Dan Multi Talenta, 1(1), 1–13. https://doi.org/10.38035/jkmt.v1i1.7.

[14] Sasmita, A., Ambarita, Y. M., & Putri, A. M. (2021). Strategi Pemasaran Tokopedia dalam Persaingan Antar E-Commerce dengan Analisis SWOT. Jptam, 5(2), 3397–3404. https://jptam.org/index.php/jptam/article/view/1403

[15] Edeh, F. O., Quttainah, M. A., Syed Badrulzaman, S. N., Kee, D. M. H., Mohd Zamri, S., Tariq Mahmud, T., Izni Shafiee, S. S., & Awoke, V. C. (2021). Factors Contributing to Online Purchase Intention amongst Malaysian Consumers: A Case of Lazada. International Journal of Accounting & Finance in Asia Pasific, 4(2), 33–45. https://doi.org/10.32535/ijafap.v4i2.1113.

[16] Nurdiansyah, A., Pratiwi, A., & Kaunaini, B. (2022). Literature Review Pengaruh Kepercayaan, Kemudahan dan Kepuasan. Jurnal Ilmu Multidisiplin, 1(1), 297–303.

[17] Masyhuri, M. (2022). Key Drivers of Customer Satisfaction on the E-Commerce Business. East Asian Journal of Multidisciplinary Research, 1(4), 657–670. https://doi.org/10.55927/eajmr.v1i4.405.

[18] Pallathadka, H., Ramirez-Asis, E. H., Loli-Poma, T. P., Kaliyaperumal, K., Ventayen, R. J. M., & Naved, M. (2023). Applications of artificial intelligence in business management, e-commerce and finance. Materials Today: Proceedings, 80(xxxx), 2610–2613. https://doi.org/10.1016/j.matpr.2021.06.419.

[19] Talanquer, V. (2023) 'Interview with the Chatbot: How Does It Reason?', Journal of Chemical Education, 100(8), pp. 2821–2824.

[20] Følstad, A., Araujo, T., Law, E. L. C., Brandtzaeg, P. B., Papadopoulos, S., Reis, L., Baez, M., Laban, G., McAllister, P., Ischen, C., Wald, R., Catania, F., Meyer von Wolff, R., Hobert, S., & Luger, E. (2021). Future directions for chatbot research: an interdisciplinary research agenda. Computing, 103(12), 2915–2942. https://doi.org/10.1007/s00607-021-01016-7.

[21] Ansari, M., & Parbulkar, M. S. (2021). Intelligent Chatbot. 9(4), 79–82.

[22] Permata, R. A., Syaidatussalihah and Abdurahim (2023). Analisis Data Penelitian Kesehatan: Perbandingan Hasil antara SmartPLS, R dan IBM SPSS untuk Jumlah Sampel Kecil. JSN : Jurnal Sains Natural, 1(1), pp. 17–22.

[23] Gajewska, T., Zimon, D., Kaczor, G. and Madzík, P. (2020) 'The impact of the level of customer satisfaction on the quality of e-commerce services', International Journal of Productivity and Performance Management, 69(4), pp. 666–684.

[24] Hair, J.F., Hult, G.T.M., Ringle, C.M. & Sarstedt, M. 2022. Partial Least Squares Structural Equation Modeling (Pls-Sem) Third Edition. Third Edit ed. Angewandte Chemie International Edition, 6(11), 951–952., Los Angeles: SAGE Publications, Inc.

[25] Ismawati, A. (2020). Pengaruh Kualitas Informasi Terhadap Keputusan Pembelian Produk Online Pada Shopee. Prisma (Platform Riset Mahasiswa Akuntansi), 01(06), 51–62. https://ojs.stiesa.ac.id/index.php/prisma/article/view/692.

[26] Yolandita, C., Nursanta, E., Widiyarsih, W., & Masitoh, S. (2022). Analisis Pengaruh Daya Tarik Iklan, Kualitas Informasi, Kemudahaan dan Kepercayaan Terhadap Keputusan Pembelian Marketplace Shopee. Jurnal Sosial Teknologi, 2(6), 479–486. https://doi.org/10.59188/jurnalsostech.v2i6.343.

[27] Maulida, N., & Jaya, U. A. (2024). Pengaruh Personalisasi dan Kecerdasan Buatan (AI) Terhadap Loyalitas Pelanggan Dalam Industri 6.0 Pada Platform E-Commerce di Sukabumi. Jurnal Ekonomi, Manajemen Dan Akuntansi, 1192, 304–317.

[28] Zuhdi, D. A., & Yasya, W. (2023). Pengaruh Kepercayaan dan Pengalaman Pembelian Online terhadap Minat Beli di E-Commerce Shopee. Jurnal Interaksi: Jurnal Ilmu Komunikasi, 7(2), 261–277. https://doi.org/10.30596/ji.v7i2.14761.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2312, Copyright ©2025
Submitted : January 16, 2025, Revised : January 21, 2025, Accepted : January 28, 2025, Published : January 31, 2025

[29] Maulidiyah, N. F. A., Singasatia, D., & Sunandar, M. A. (2022). Analisis Pengaruh User Experience Terhadap Kepuasan Pengguna Mobile Application VLive Menggunakan Model SCSI. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 2(2), 28–34. https://doi.org/10.57152/malcom.v2i2.398.

[30] Susiyanto, D., & Naningsih. (2024). Pengaruh Artificial IntelligenceTerhadap E Business Dan E Commerce. Jurnal Ekonomi, Manajemen Dan Akuntasi, 2(1), 482–486. http://jurnal.kolibi.org/index.php/neraca.

[31] Kapojos, A. L. N., Saerang, D. P. E., & Gunawan, E. M. (2023). the Effect of E-Service Quality and Waiting Time Towards Customer Satisfaction At Pizza Hut Restaurant in Manado Town Square. Jurnal EMBA : Jurnal Riset Ekonomi, Manajemen, Bisnis Dan Akuntansi, 11(4), 439–449. https://doi.org/10.35794/emba.v11i4.51230.

[32] Fauzan, M., Wicaksana, M. P., & Rahardandi, P. G. (2024). Literatur Review Penggunaan Chatbot Untuk Layanan Informasi. Journal Of Social Science Research, 4(4), 8316–8323.

[33] Ramadhani, D. R., Birawa, M. S., Sholikah, D. S., & Prabandanu, R. M. A. A. H. (2024). Analisis Pengaruh Kualitas Pelayanan Berbasis Chatbot Terhadap Kepuasan Pelanggan Dalam Transaksi Online Dengan Objek Aplikasi Shopee. Journal of Exploratory Dynamic Problems, 1, 104–115.

[34] Pasaribu, F., Sari, W. P., Ni Bulan, T. R. and Astuty, W. (2022) 'The effect of e-commerce service quality on customer satisfaction, trust and loyalty', International Journal of Data and Network Science, 6(4), pp. 1077–1084.

[35] Song, X., Yang, S., Huang, Z. and Huang, T. (2019) 'The Application of Artificial Intelligence in Electronic Commerce', Journal of Physics: Conference Series, 1302(3).

[36] Sudirjo, F., Ratnawati, Hadiyati, R., Sutaguna, I. N. T. and Yusuf, M. (2023) 'The Influence Of Online Customer Reviews And E-Service Quality On Buying Decisions In Electronic Commerce Ratnawati 2 Politeknik Pertanian Negeri Pangkep 2', Journal of Management and Creative Business (JMCBUS), 1(2), pp. 156–181.

[37] Reinkemeier, F. and Gnewuch, U. (2022) 'Designing effective conversational repair strategies for chatbots', 30th European Conference on Information Systems (ECIS), (June), pp. 1–17.

[38] Ingaldi, M. and Ulewicz, R. (2019) 'How to make e-commerce more successful by use of Kano's model to assess customer satisfaction in terms of sustainable development', Sustainability (Switzerland), 11(18), pp. 1–22.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2312, Copyright ©2025
Submitted : January 16, 2025, Revised : January 21, 2025, Accepted : January 28, 2025, Published : January 31, 2025

Analysis to Predict the Number of New Students At UNU Pasuruan using Arima Method

Fachri Ayudi Fitrony[1]*, Laksmita Dewi Supraba [2], Tessa Rantung [3], I Made Artha Agastya [4], Kusrini [5]
Faculty of Computer Science
University of Amikom, Yogyakarta, Indonesia
fahri24@students.amikom.ac.id[1], laksmitasupraba@students.amikom.ac.id[2], tessarantung@students.amikom.ac.id[3], artha.agastya@amikom.ac.id [5], kusrini@amikom.ac.id[4]

Abstract*—New student admission is an important aspect in higher education management, including Nahdlatul Ulama University (UNU) Pasuruan. Relevant prediction of total new students is needed to support resource planning such as teaching staff, facilities, and budget. This study aims to evaluate the historical pattern of new student admissions at UNU Pasuruan and predict the number of new students in the coming years using the ARIMA (Auto Regressive Integrated Moving Average) method. The data used is historical data on new student admissions in the last five years, which is analyzed to identify trends, seasonality, and fluctuation patterns. The analysis is performed using statistical software such as Python to improve the accuracy and efficiency of the process. This study approach includes several main steps, namely collecting historical data on the number of new students, testing stationarity using the Augmented Dickey-Fuller (ADF) test, identifying model parameters through ACF and PACF graphs, and estimating ARIMA model parameters. The resulting model is evaluated using prediction error metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). The study findings describe that the ARIMA model (6,0,1) produces an RMSE value of 21.88 and a MAPE of 0.2%. In addition to having the smallest error score, the ARIMA model (6,0,1) also has the smallest AIC score of the various models that can be used for predictions, which is 447.44 and the largest log likelihood value, which is -214.72. The largest prediction of the number of new students is in July, which is 92.72 and the smallest in February, which is 24.43. This prediction is expected to help university management in optimizing resource planning, increasing management efficiency, and anticipating fluctuations in the number of new students in the future. This study offers new findings in the form of the use of predictive models based on historical data to support strategic decision-making, such as resource allocation and promotion planning. With these results, universities can anticipate changes in the number of enrollments more effectively, which were previously only done based on subjective estimates. The model built can also be applied to similar datasets in the future with appropriate parameter adjustments.

Keywords— New Student Prediction, ARIMA, Stationarity

I. INTRODUCTION

Higher education is an educational institution that aims to educate students to become productive members of society by acquiring knowledge and skills in various fields, including the world of work [1] In the context of higher education, students play a very important role in achieving educational goals [2]. In order to complete their education quickly and obtain a suitable career, prospective students tend to choose the best colleges. When choosing a college, they consider various factors such as the facilities offered, the academic performance of the college and the reputation of the institution [3]. To maximise academic services to students and even prospective students, the university needs to prepare various things both in terms of physical and non-physical, such as facilities in each study program, accreditation, teaching staff and other things.

In order to become one of the best universities that is the choice and destination of prospective students to come to study, the University of Nahdlatul Ulama (UNU) Pasuruan needs a strategy in order to prepare various things needed by the university to develop even better. It is also based on its goal to produce young people who excel in academics and have a strong moral and spiritual character. This university offers various study programs and continues to improve the quality of education and its competitiveness.

One of the strategies taken by UNU Pasuruan is predicting the number of new students each year, which is affected by external aspects including economic conditions and education policies, as well as internal factors such as campus reputation [4]. Forecasting the number of new students is essential to ensure that resource needs can be anticipated [5].

To take into account the forecasting of the number of new students we apply one of the data analysis technologies, namely the ARIMA method, which predicts trends based on historical data [2]. The use of ARIMA for prediction of the number of new students at UNU Pasuruan can help understand fluctuations in enrolment which is important for managerial decision making. In addition to researching and predicting the number of incoming students at UNU Pasuruan, by using ARIMA we also provide recommendations for human resource planning and management.

For the application of this arima method, commonly used time series analysis models are the Auto Regression (AR) model, Moving Average (MA) model, Autoregression Moving Average (ARMA) model, Auto Regression Integration Moving Average (ARIMA) model and SARIMA model [6]. Among these methods, ARIMA is an effective approach to analyzing time series data because of its ability to handle both stationary and non-stationary data [7]. ARIMA models can identify trends, seasonality, and fluctuations in data, and provide a solid

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2251, Copyright ©2025
Submitted : January 9, 2025, Revised : January 20, 2025, Accepted : January 28, 2025, Published : January 31, 2025

foundation for accurate forecasting and decision making. ARIMA also has high flexibility in handling time series data with various characteristics [8]. ARIMA can be more reliable in time and resource-limited interval forecasting when compared to other forecasting methods [9].

The reason we chose the ARIMA method for prediction analysis of the number of new students at UNU Pasuruan is because this method is very suitable for analysing time series data. ARIMA has the ability to identify patterns in historical data, including trends, seasonality, or random fluctuations, which often occur in new student enrolment data. In addition, ARIMA is flexible because its components (AR, I, MA) can be adjusted to the characteristics of the data, including nonstationary data [5] [8]. This ability allows the model to provide more accurate prediction results, which can help understand new student enrolment patterns in more depth [5].

In addition to flexibility and accuracy, the ARIMA method is also easy to implement with the help of statistical software such as Python that we used in this research, making it a practical choice for data-driven research [3]. ARIMA has been widely used in educational contexts, including student enrolment prediction, making it relevant for use in this study. With more precise prediction results, this research is expected to help UNU Pasuruan in resource planning, promotion strategies, and capacity management more effectively.

II. RESEARCH METHODS

The research method used uses the steps in the ARIMA method, with a flow as in Figure 1.

1) Time series data collection in the form of a dataset of the number of new students over 5 years.
2) Stationary inspection using Augmented Dickey-Fuller (ADF) test.
3) If the data is not stationary, differentiation must be carried out so that the data becomes stationary.
4) If the information is stationary, it will proceed to the next stage.
5) Using the pmdarima library in Python to identify the model to be analyzed.
6) The best model will be obtained based on the lowest AIC and BIC values.
7) Validation is performed to check the accuracy of the selected mathematical model.
8) If validation fails to meet the most appropriate model parameters, a repeat will be performed to select a different model using the appropriate lag.
9) If the validation is correct, proceed to the next stage.
10) Forecasting is done to evaluate the total number of incoming students in the coming period.

Stationary data is time series data that has statistical properties that are constant over time [5]. This means that the mean, variance and covariance of the data do not change significantly over time. In time series analysis, stationarity is an important prerequisite because many predictive methods, including ARIMA, assume that the data has statistically stable properties. Stationary data usually show no upward or downward trend patterns, and data fluctuations appear to be consistent around the average [10] [11].

Indicators of successful stationary data can be seen in several ways. Visually, a graph of stationary data will show a flat pattern with no clear trend, with uniform fluctuations [11]. Statistically, tests such as ADF can be used to confirm stationarity, where a p-value <0.05 indicates that the data is stationary. In addition, the Autocorrelation Function (ACF) graph for stationary data will show rapidly decreasing autocorrelation after a few lags. If the data is not stationary, differencing techniques are often applied to achieve stationarity, which is characterised by a more stable data pattern after the process. Stationarity is important to ensure the results of analyses and predictions are more accurate and reliable [12].

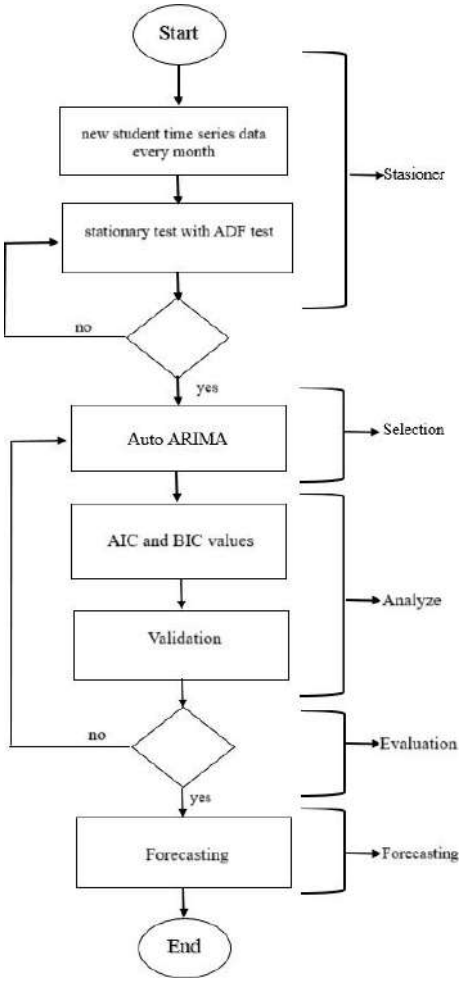


Fig. 1. Research Methods

A. Data and Variables

This research was conducted in the period from November 2024 to January 2025. The source of research data came from the New Student Admissions Bureau (PMB) located on Jl. Warungdowo, Pasuruan Regency, and connected to Nahdlatul Ulama Pasuruan University. The data collection technique in this study is a series of steps to collect relevant information to support the research objectives [11]. The methods used include document studies, surveys, interviews, literature reviews, and observations. Specifically, this study relies on document study methodology to collect findings from various media sources,

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2251, Copyright ©2025
Submitted : January 9, 2025, Revised : January 20, 2025, Accepted : January 28, 2025, Published : January 31, 2025

such as books, images, videos, and databases. Data obtained from the PMB Bureau became the main basis in this study for further analysis.

Quantitative data and qualitative data are the two main categories of information collection used in research [10]. The main difference between the two lies in the form and method of processing: quantitative data is presented in numerical form, which allows for more precise and measurable processing and analysis, while qualitative data is descriptive and cannot be measured with numbers [13]. For example, in this study, the materials used are quantitative data, and can be seen in Table 1, which contains information in the form of numbers and statistics that will be analyzed in more detail.

TABLE I. STUDENT DATA PER MONTH FOR 5 YEARS

No	Number of New Students
1	8
2	24
3	34
4	55
5	72
6	75
7	79
8	82
9	88
10	90
……..	……
59	89
60	118

B. ARIMA Model

ARIMA is a technique used to predict future events by utilizing past data to produce short-term forecasts [2]. ARIMA has the advantage of being able to predict with an effective level of accuracy for use in short-term forecasting. In addition, ARIMA is also flexible and fast [14]. Arima combines AR and MA models to be able to make predictions on time series data. AR explains the model regarding the dependent variable that is affected by the variable in the previous time. The AR model is described with AR(p) or ARIMA (p,0,0) through the stage of forming equation (1)

$$Y_t = \emptyset 1.Y_{t-1} + \emptyset 2.Y_{t-2} + \cdots + \emptyset p.Y_{t-p} + wt \quad (1)$$

Where:
Yt: value of variable X at time t
$\emptyset$: p-th Autoregressive Parameter
wt *: error value at time t*

Moving Average (MA) is a stage that describes in an explicit context the correlation of dependence on error scores in sequence. The MA model is described with MA(q) or Arima (0,0,q) by forming equation (2).

$$Y_t = wt + \theta 1.wt-1 + \theta 2.wt-2 + \cdots + \theta q.wt-p \quad (2)$$

Where :
Yt*: value of variable X at time t*
θq: qth moving average parameter
wt*: error value at time t*

AR and MA can be combined into ARMA, namely (p,q) or Arima (p,0,q) by forming equation (3).

$$Y_t = \emptyset 1.Y_{t-1} + \emptyset 2.Y_{t-2} + \cdots + \emptyset p.Y_{t-p} + wt + \theta 1.wt-1 + \theta 2.wt-2 + \cdots + \theta q.wt-p \quad (3)$$

Where :
Yt*: Value of variable Z at time t*
$\emptyset p$: p-th AR parameter
θq: qth MA parameter
wt*: error value at time t*

Meanwhile, the ARIMA model is a technique that integrates the AR model and the MA model with the form ARIMA (p, d, q) where the order p represents the AR operator, the order d represents differencing, and the order q represents the MA operator. The ARIMA (p, d, q) model forms equation (4).

$$Y_t = (1+\emptyset 1).t-1 + (\emptyset 1-\emptyset 2).t-2 + \cdots + (\emptyset p-\emptyset p-1).t-p + wt - \theta q.wt-1 - \cdots - \theta q.wt-p \quad (4)$$

Where :
Yt*: Value of variable Z at time t*
$\emptyset p$: p-th AR parameter
θq: qth MA parameter
wt*: error value at time t*

The forecasting carried out in this study uses the ARIMA method using Python, with several libraries in it to assist in this study, namely the Matplotlib, PMdarima, SicklitLearn, Pandas, Numpy and other libraries. [12]. The steps of applying the ARIMA method in this study are described in Figure 1.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2251, Copyright ©2025
Submitted : January 9, 2025, Revised : January 20, 2025, Accepted : January 28, 2025, Published : January 31, 2025

III. RESULTS AND DISCUSSION

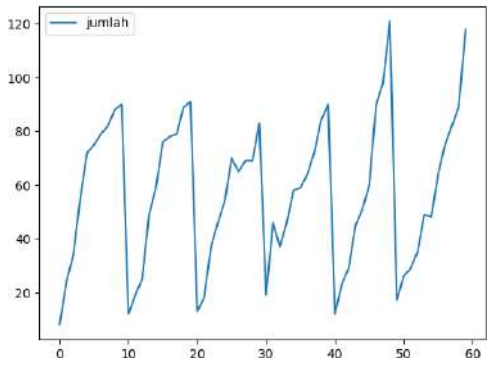


Fig. 2. New Student Time Series Graph

In Figure 2, the time series graph shows the number of students per month throughout the period 2019 - 2023 with a total of 60 datasets on the (X) axis and 120 students on the (Y) axis. Figure 2 shows that the data needs to be ensured to be stationary in mean and variance, this is done to see whether the ARIMA method has been fulfilled to be carried out in this study. This test is useful for finding out whether a time series has a unit root, which indicates whether or not there is a trend or pattern in the data. The ADF test carried out with Drift and Trend produces a low p-value score (<= 0.1) with a value of 0.044, these findings illustrate stationary data. Because the p-value score is very low (<= 0.01) for testing, it can be concluded.

The ACF or Autocorrelation Function test in Figure 4 is used to calculate and display the autocorrelation function of the time series stored from the decade rainfall data. 17.5 iterations were performed looking at the ACF value. The ACF value helps in identifying patterns or correlation structures in the time series, which can provide insight into the nature of the data and help in modeling and prediction. From the results of the ACF value, 2 lags are seen that are out of the standard significance of the values 1.00 and 0.50 which means they contain MA (1) and MA (2).

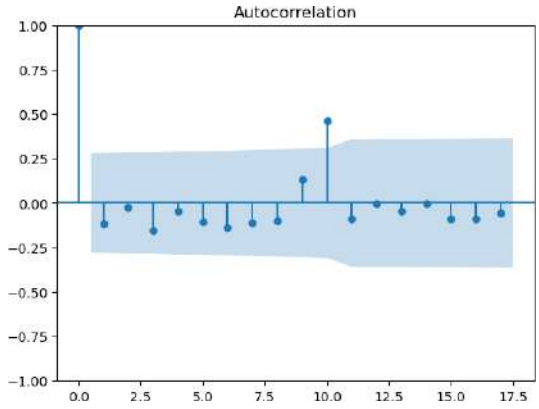


Fig. 3. ACF Plot of New Student Data

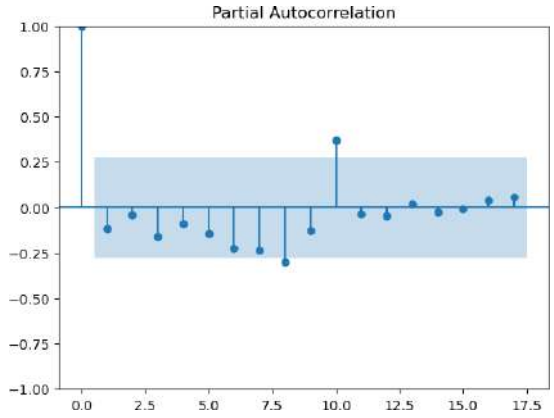


Fig. 4. PACF Plot of New Student Data

The PACF or Partial Autocorrelation Function test in Table 1 is used to calculate and display the partial correlation function of the time series stored in the Data object of the number of new students. The results of this test will provide a partial correlation value for each lag. PACF is a method used to identify a direct relationship between one time point and a specific time point in a time series, after eliminating the influence of correlation on previous lags. PACF analysis is useful in determining the optimal order of the ARIMA model, which is often applied to time series modeling. From the results of this PACF test, it can be seen that there are 3 lags that are outside the significance limits of the values 1.00, -0.25 and 0.35 which have the meaning of autoregressive or AR (1) AR (2) AR (3).

TABLE II. AUTO ARIMA RESULTS

ARIMA (0,0,0) (0,0,0) [0]	: AIC=670.794, Time=0.01 sec
ARIMA (0,0,1)(0,0,0)[0]	: AIC=626.368, Time=0.03 sec
ARIMA (0,0,2)(0,0,0)[0]	: AIC=601.122, Time=0.05 sec
ARIMA (0,0,3)(0,0,0)[0]	: AIC=592.650, Time=0.13 sec
ARIMA (0,0,4)(0,0,0)[0]	: AIC=588.802, Time=0.18 sec
ARIMA (0,0,5)(0,0,0)[0]	: AIC=infc, Time=0.21 sec
ARIMA (1,0,0)(0,0,0)[0]	: AIC=567.388, Time=0.02 sec
ARIMA (1,0,1)(0,0,0)[0]	: AIC=568.932, Time=0.04 sec
ARIMA (1,0,2)(0,0,0)[0]	: AIC=570.915, Time=0.05 sec
ARIMA (1,0,3)(0,0,0)[0]	: AIC=infc, Time=0.13 sec
ARIMA (1,0,4)(0,0,0)[0]	: AIC=infc, Time=0.18 sec
ARIMA (2,0,0)(0,0,0)[0]	: AIC=568.936, Time=0.03 sec
ARIMA (2,0,1)(0,0,0)[0]	: AIC= infc, Time=0.13 sec
ARIMA (2,0,2)(0,0,0)[0]	: AIC= infc, Time=0.15 sec
ARIMA (2,0,3)(0,0,0)[0]	: AIC= infc, Time=0.17 sec
ARIMA (3,0,0)(0,0,0)[0]	: AIC=570.935, Time=0.04 sec
ARIMA (3,0,1)(0,0,0)[0]	: AIC=572.655, Time=0.16 sec
ARIMA (3,0,2)(0,0,0)[0]	: AIC=infc, Time=0.17 sec
ARIMA (4,0,0)(0,0,0)[0]	: AIC=572.061, Time=0.04 sec

SARIMAX Result

Dep. Vaiable:	jumlah	No. Observations:	50
Model:	ARIMA (6,0,1)	Log Likehood	-214.732
Date:	Sat, 04 Jan 2025	AIC	447.446
Time:	07:39:50	BIC	464.654
Sample:	0	HQIC	453.999

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2251, Copyright ©2025
Submitted : January 9, 2025, Revised : January 20, 2025, Accepted : January 28, 2025, Published : January 31, 2025

In Figure 6, Auto Arima is run to create an Arima model to predict the number of new students. From this command, several iterations are carried out which are used to find the most optimal model based on the AIC score. Several models that were tried include ARIMA (2,0,2), ARIMA (0,0,0), ARIMA (1,0,0), ARIMA (0,0,1), ARIMA (0,0,0) (without mean), ARIMA (2,0,0), ARIMA (3,0,0), ARIMA (6,0,1), and so on. The model that has a smaller AIC score is declared better. The best model is obtained with the smallest AIC value among the other orders, namely ARIMA (6,0,1) with parameters ARIMA (6,0,1) with non-zero mean, Model coefficient ar1 = 0.6015, ma2 = -0.6801, std err = 0.833, mean = 163.524 with model performance Residual variance (sigma ^ 2) = 286.1147, Log likelihood = -214.72, AIC = 447.44 and BIC = 464.65.

SARIMAX Results

Dep. Variable:	jumlah	No. Observations:	50
Model:	ARIMA(6, 0, 1)	Log Likelihood	-214.723
Date:	Sat, 04 Jan 2025	AIC	447.446
Time:	07:39:50	BIC	464.654
Sample:	0	HQIC	453.999
	- 50		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
const	56.5149	0.833	67.879	0.000	54.883	58.147
ar.L1	0.6015	0.265	2.273	0.023	0.083	1.120
ar.L2	-0.2384	0.359	-0.664	0.507	-0.942	0.465
ar.L3	-0.2913	0.311	-0.938	0.348	-0.900	0.318
ar.L4	-0.0112	0.372	-0.030	0.976	-0.740	0.718
ar.L5	-0.1609	0.362	-0.444	0.657	-0.870	0.549
ar.L6	-0.3418	0.286	-1.194	0.232	-0.903	0.219
ma.L1	-0.6801	0.251	-2.710	0.007	-1.172	-0.188
sigma2	286.1147	62.547	4.574	0.000	163.524	408.705

Ljung-Box (L1) (Q):	0.20	Jarque-Bera (JB):	1.20
Prob(Q):	0.66	Prob(JB):	0.55
Heteroskedasticity (H):	1.49	Skew:	0.31
Prob(H) (two-sided):	0.42	Kurtosis:	3.44

Fig. 5. ARIMA (6,0,1) Results

This model also has a small AIC and BIC value so that the better the model is in explaining the data, then the high Log likelihood value indicates that the model can explain the data well. So overall, this model provides a good representation of the data on the number of new students, and the coefficients obtained can be used to make predictions or for further analysis. The last step in time series analysis to predict the number of new students is to forecast. Forecasting is done as much as 10 Prediction data or about 10 months into the future. This is done to find out the predicted number of new students in the following year.

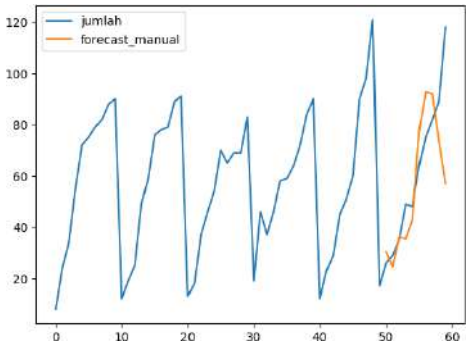


Fig. 6. Results of Forecasting the Number of New Students

Previous studies have shown that the ARIMA model can be effectively used to predict the number of new students with various parameters, such as ARIMA (2,1,1) and ARIMA (11,0,12), which were selected based on the lowest accuracy and error rates [15][16]. However, recent studies using the ARIMA (6,0,1) model show higher AIC and BIC values, indicating that parameter selection can still be optimized. In addition, the low statistical significance of some lags in this model can affect the accuracy of the prediction results. This difference emphasizes the need for a more optimal development model to improve the accuracy of estimating the number of new students. Meanwhile, the results of forecasting the number of new students show that the estimation of the new student registration period can follow the pattern of the training data used. In addition, the ten test data shown in Table 2 show an increasing trend that is in accordance with the movement pattern of the graph, so that the validity of the model in describing historical trends can be confirmed.

50	30.465696
51	24.437012
52	36.102347
53	35.364126
54	42.616382
55	77.199688
56	92.724535
57	91.896221
58	73.671824
59	57.082380

Fig. 7. 10 data on forecasting the number of new students

The final step in forecasting is to evaluate the model used by measuring MAE, MAPE and RMSE. In Figure 10, it is known that the ARIMA (6,0,1) model produces MAE = 14,621, MAPE = 0.2 and RMSE = 21,884.

TABLE III. MAE, MAPE AND RMSE EVALUATION RESULTS

MAE – manual	14.62167618080127
MAPE – manual	0.20027750011448164
RMSE – manual	21.8840908325154497

IV. CONCLUSION

This study proves that the ARIMA method is effective in predicting the number of new students based on historical new student admission data. This prediction provides a strong foundation for UNU Pasuruan in developing new student admission strategies and better resource management. The resulting model is Model (6,0,1) based on auto ARIMA which is known to have the smallest AIC score of various models that can be applied in making predictions, namely 447.44 and the highest log likelihood score, namely -214.723. So the ARIMA model (6,0,1) is considered to be the most optimal step in carrying out the estimation in the context of the analysis carried out with the prediction results for the next period from the data obtained by the researcher, namely: January (30.46), February (24.43), March (36.10), April (35.36), May (42.61), June (77.19), July (92.72), August (91.89), September (73.67), October (57.08).

ARIMA models have advantages in analysing and predicting time series data, but their application to different

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2251, Copyright ©2025
Submitted : January 9, 2025, Revised : January 20, 2025, Accepted : January 28, 2025, Published : January 31, 2025

datasets or longer periods has limitations that need to be considered [17][18]. One of the main challenges is the **sensitivity of ARIMA to non-stationary data**[19]. ARIMA requires the data to be stationary before the model is built, so if the new dataset has complex patterns of trends, seasonality, or fluctuations that cannot be normalised through differencing, the model may not provide optimal results. This can be a problem when the data used has a highly dynamic pattern or is influenced by significant external factors [20].

In addition, ARIMA is less able to capture non-linear patterns that may exist in time series data. The model is based on linear relationships between historical data, so complex patterns of relationships, such as variable interactions or sudden changes, are often poorly represented. In the context of longer periods, ARIMA also tends to become less accurate as the assumption that past patterns will continue may not hold in the future. As a result, these models need to be evaluated and updated regularly to ensure their performance remains relevant to the latest data. For very different datasets or longer periods, combining ARIMA with other models, such as machine learning-based models, can help capture more complex patterns.

For further research, it is recommended that researchers consider external factors that may affect the results, such as changes in education policy and economic conditions. Both of these factors have the potential to have a significant impact on the variables studied, so by including these factors, the accuracy of predictions in the study can be improved. Taking these external factors into account will provide a more comprehensive and realistic picture, and help in producing more relevant and applicable findings in a broader context.

REFERENCES

[1] A. P. T. N. &. L. D. H. Yordan, "Peramalan Penerimaan Mahasiswa Baru Universitas Samudra Menggunakan Metode Regresi Linear Sederhana.," Jurnal Teknik Informatika (Jtifa), P. 21–27 , 2019.

[2] M. &. Y. R. Rianto, "Analisis Runtun Waktu Untuk Memprediksi Jumlah Mahasiswa Baru Dengan Model Random Forest.," Paradigma – Jurnal Komputer Dan Informatika, P. 23(1), 2021.

[3] T. A. S. D. S. M. Dimri, "Time Series Analysis Of Climate Variables Using Seasonal Arima Approach," Journal Of Earth System Science, , 2020.

[4] S. Aminah, "Modeling And Forecasting University Admissions Trends Using Arima," International Journal Of Advanced Trends In Computer Science And Engineering, 2022.

[5] S. Kornelio, "Forecasting Students' Enrollment In Tanzania Government Primary Schools From 2021 To 2035 Using Arima Model.," International Journal Of Curriculum And Instruction, P. 162–174, 2024.

[6] Adeline Purisima Dela Cruz, "Higher Education Institution (Hei) Enrollment Forecasting Using Data Mining Technique.," International Journal Of Advanced Trends In Computer Science And Engineering, Pp. 2060-2064., 2020.

[7] D. Mirna Nachouki, "Student Course Grade Prediction Using The Random Forest Algorithm: Analysis Of Predictors' Importance. Trends In Neuroscience And Education.," Journal Of Cloud Computing, 2023.

[8] T. Hailemeskel Abebe, " Time Series Analysis Of Monthly Average Temperature And Rainfall Using Seasonal Arima Model (In Case Of Ambo Area, Ethiopia). International Journal Of Theoretical And Applied Mathematics,," Sistem Pendukung Keputusan Untuk Umkm. [Ebook]. Ubx Press., Pp. 6(5), Hal.76., 2020.

[9] Y. D. D. D. Lai, " Use Of The Autoregressive Integrated Moving Average (Arima) Model To Forecast Nearterm Regional Temperature And Precipitation. Weather And Forecasting,," P. 959–976. , 2020.

[10] S. Aminah, "Modeling And Forecasting University Admissions Trends Using Arima," International Journal Of Theoretical And Applied Mathematics, 2022.

[11] D. Jiang Li, "Wavelet Transforms Based Arima-Xgboost Hybrid Method For Layer Actions Response Time Prediction Of Cloud Gis Services," Journal Of Cloud Computing, 2023.

[12] D. Yan Pan, "A Study On The Prediction Of Book Borrowing Based On Arima-Svr Model.," Procedia Computer Science, Pp. 93-102 , 2021.

[13] R. Kumar, "Predicting Academic Performance By Data Mining Methods," Educational Economics, Pp. 405-419, 2021.

[14] D. Mirna Nachouki, "Student Course Grade Prediction Using The Random Forest Algorithm: Analysis Of Predictors' Importance.," Trends In Neuroscience And Education. , 2023.

[15] A. U. Jamila, B. M. Siregar, And R. Yunis, “Analisis Runtun Waktu Untuk Memprediksi Jumlah Mahasiswa Baru Dengan Model Arima,” Paradigma - Jurnal Komputer Dan Informatika, Vol. 23, No. 1, Mar. 2021, Doi: 10.31294/P.V23i1.9758.

[16] B. H. Mawaridi And M. Faisal, “Prediksi Mahasiswa Baru Universitas Papua Menggunakan Autoregressive Integrated Moving Average,” Jurnal Informatika, Vol. 10, No. 2, Pp. 157–162, Oct. 2023, Doi: 10.31294/Inf.V10i2.16000.

[17] S. A. Sinaga, "Implementasi Metode Arima (Autoregressive Moving Average) Untuk Prediksi Penjualan Mobil," Seminar Nasional Teknologi Dan Informatika (Snti), Universitas Budi Darma, Medan, Indonesia, 2023.

[18] D. A. Fajari Dan M. F., "Aplikasi Metode Arima Dalam Meramalkan Rata-Rata Harga Beras Di Tingkat Perdagangan Besar (Grosir) Indonesia," Jurnal Agroteknologi, Vol. 6, No. 1, Hal. 1-10, 2022, Doi: 10.31316/Agroteknologi.V6i1.1766.

[19] Farahdinna, F. (2022). Forecasting System Prototype Amount Of Printing Raw Material Inventory Needs Using The Autoregressive Integrated Moving Average (Arima) Method. Jurnal Mantik, 6(2), 1964-1970

[20] R N Putri1, M Usman, Warsono, Widiarti, E Virginia. . (2021). Modeling Autoregressive Integrated Moving Average (Arima) And Forecasting Of Pt Unilever Indonesia Tbk Share Prices During The Covid-19 Pandemic Period. Journal Of Physics: Conference Series, 1751(1), 012027. Https://Doi.Org/10.1088/1742-6596/1751/1/012027

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2251, Copyright ©2025
Submitted : January 9, 2025, Revised : January 20, 2025, Accepted : January 28, 2025, Published : January 31, 2025