Vol. 7, No. 1, 2025, pp. 1-12.

Application of the ARIMA-GARCH Model for Forecasting Indonesia's Monthly Inflation Rate

Anisa¹, Ira Yudistira², Tony Yulianto³

^{1,2,3}Universitas Islam Madura

¹Corresponding author: <u>annisasarimin29@gmail.com</u>

Abstract. Inflation is one of the important aspects that is used as a benchmark to see economic growth and economic conditions in each country. Inflation has resulted in increasing public expenditure in meeting basic needs. Inflation must be controlled to maintain the economic stability of a country, including Indonesia. Therefore, there is a need for a model that can forecast the inflation rate in Indonesia. The aim of this research is to create a model that can predict future inflation levels so that it can help the government in determining policies related to controlling inflation in Indonesia. The data used is monthly inflation data in Indonesia for 19 years from March 2007-October 2023 in percentage form. The forecasting model used in this study is the ARIMA-GARCH model. The ARIMA model is a time series model used to forecast future data based on past data. While GARCH is a time series model used to overcome heteroscedasticity in the ARIMA model. Inflation data will be modeled using the ARIMA model and then continued by modeling the residuals using the GARCH model if heteroscedasticity occurs in the ARIMA model residuals. Based on data analysis that has been done, the best model for inflation forecasting cases in Indonesia is the ARIMA (2,0,2) - GARCH (0,1) model with a MAPE value of 17.78%.

Keywords: Inflation, Forecasting, ARIMA, GARCH

1 Introduction

Inflation is an indicator to see the economic stability of a region. Inflation shows the general development of prices of goods and services which is calculated from the consumer price index. Inflation rates greatly influence people's purchasing power and the amount of goods produced. Inflation is one of the obstacles to the process of economic development in order to improve people's standard of living as measured by the high and low income of the population each year or per capita income [1]. The continuous increase in inflation results in several economic problems such as an unstable economy, slow economic growth, a decline in the value of the currency which indirectly affects global trade activities [2].

Inflation also affects the number of unemployed. If inflation rises, the number of unemployed will increase and on the other hand, if economic growth rises, the number of unemployed decreases [3]. The inflation rate in Indonesia is measured from the percentage change in the Consumer Price Index (CPI) and at the beginning of the month announced to the public by the Central Statistics Agency (BPS) [1]. BPS announced Indonesia's inflation or CPI of 2.28% in September 2023 on an annual basis. BPS

Received: July 21, 2024 Accepted: January 14, 2025 recorded monthly inflation of 0.19%. The inflation rate has increased compared to the inflation position in the previous month [4].

Inflation causes people's spending to meet basic needs to increase. The problem of inflation must be controlled to maintain the economic stability of a country, including Indonesia. Therefore, there is a need for a model that can predict the level of Indonesian inflation in the future so that the government can immediately take policies to overcome its effects. Time Series Analysis is a quantitative forecasting analysis method that considers time. In this analysis, data is collected periodically based on time sequence to determine past data patterns. Time series forecasting techniques are divided into two parts. First, the forecasting model is based on statistical mathematical models such as moving average, exponential smoothing, regression and ARIMA (Box Jenkins). Second, forecasting models based on artificial intelligence such as neural networks, genetic algorithms, classification and hybrids [5]. ARIMA is a combination of two models, namely the Autoregression (AR) and Moving Average (MA) models [6]. The ARIMA model is most often used in forecasting financial data, because in some cases this model produces better estimates [5]. A good ARIMA model requires heteroscedasticity-free residuals. The GARCH model is a forecasting model used to improve ARIMA models whose residuals are indicated to contain heteroscedasticity. GARCH is used to model residuals from ARIMA models that contain heteroscedasticity. The complete model of the ARIMA-GARCH model is a combination of the ARIMA model obtained and the GARCH model from the residuals of the ARIMA model.

Research on forecasting using ARIMA-GRACH has been carried out by Yolanda et al (2017) regarding the Application of the ARIMA-GARCH Model to Predict BRI Bank Share Prices, where the results of their research show that the best model obtained for BRI bank share prices is ARIMA(2,1, 1)-GARCH(2,2). This model has a coefficient of determination or (R-squared) value of 0.99916 or 99.91%. Gam et al (2022) have also researchedVolatility Analysis and Inflation Forecasting in North Maluku Using the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) Model which produces the results of ARCH/GARCH analysis of monthly inflation data in North Maluku producing the best model for estimating inflation volatility, namely the GARCH (1.0) model. This model provides information about the level of movement of the inflation rate in the period January 2013 to August 2022, where the results show that the volatility value of the inflation rate moves with unstable fluctuations as well as in forecasting for the next 4 months. Other research has also been carried out by Sirius et al (2023) regarding Implementation of the ARIMA-GARCH Method for Forecasting Yen to Rupiah Currency Conversion. The results of this research show that the most appropriate model for analyzing the data in this research is the ARIMA(3,1,3)-GARCH(1,1) model. This is because the model can overcome homogeneity in the existing data. The conversion of Yen to Rupiah from August 2022 to July 2023 can be predicted to have fluctuations, and will result in a MAPE value of 19.29%. Based on this background, researchers will forecast inflation in Indonesia using the ARIMA-GARCH method.

2 Methods

2.1 Research data

The data used in this research is secondary data obtained from the Indonesian Data Bank website. The data used is monthly inflation data in Indonesia for 19 years from March 2007- October 2023 in percentage form.

2.2 Research procedure

The stages of data analysis carried out in this research are as follows:

2.2.1 Data Stationarity test

Time series data can be said to be stationary if there is no sharp increase or decrease in the data [7]. Stationarity can be divided into two, there are stationarity in the mean and stationarity in variance. The data stationarity test can be carried out using the ADF (Augmented Dicky Fuller) test. The equation used in the test is defined below.

$$\Delta Z_t = (\rho - 1)Z_{t-1} + \mu_t$$

where $\delta = (\rho - 1)$ and $\Delta Z_t = Z_t - Z_{t-1}$ with Z_t is data in the periode. The hypothesis tested is

 H_0 : $\delta = 0$ (data contains unit roots/data is not stationary)

$$H_1: \delta < 0$$
 (data does not contain unit roots/stationary data)

The test statistic used to test the hypothesis is $ADF = \frac{\hat{\delta}}{se(\hat{\delta})}$ where $\hat{\delta}$ is the estimated value of δ and $se(\hat{\delta})$ is the standard error of $\hat{\delta}$. Reject H_0 if the ADF value is smaller than the ADF critical value and it can be concluded that the data is stationary. The ADF critical value is found using the *t* distribution table [8].

2.2.2 Identify ARIMA Models

Model identification is used as a reference in selecting the best model. In this research, the best model selection criterion used is Akaike's Information Criterion (AIC), which defined in Equation 1. The AIC procedure is used to evaluate how well a temporary model compares with the actual model by looking at the difference between the expected values of the actual model and the interim model [9].

$$AIC = n\ln\left|\widehat{\sum_{p}}\right| + 2pr^{2} \tag{1}$$

Where *n* is the number of observations, *r* is the dimension of the process vector x_t , *p* is the lag used and $|\widehat{\Sigma_p}|$ is the determinant of the covariance. The AIC value is used to select a fit autoregressive model, so that the model that has the smallest AIC value is the best model for predicting values simultaneously.

The ARIMA model is a univariate time series model and is a combination of autoregressive (AR) and moving average (MA). The form of the ARIMA model is defined in Equation 2.

$$Z_{t} = (1 - \phi_{1})Z_{t-1} + (\phi_{1} - \phi_{2})Z_{t-2} + \dots + (\phi_{p} - \phi_{p-1})Z_{t-p} + \theta_{0} - \theta_{1}\gamma_{t-1}$$

$$-\theta_{2}\gamma_{t-2} - \dots - \theta_{q}\gamma_{t-q}$$
(2)

where:

 $\begin{array}{ll} Z_t &= \text{Data in } t \\ \gamma_{t-1} &= \text{Error in period period } t \\ \phi_1, \phi_2, \cdots, \phi_{p-1} &= \text{AR model parameters} \\ \theta_0, \theta_1, \cdots, \theta_q &= \text{MA model parameters} \end{array}$

2.2.3 Estimation of ARIMA Model Parameters

Estimates of parameters in sample data can be used to determine inferences about β and σ^2 Where β is the average of a data sample and σ^2 is the variance of the data. These inferences may take specific points (point estimates) or determine a range of parameter values (interval estimates) [10].

2.2.4 Parameter Significance Test

According to Wei in 2006, The parameter significance test is carried out after obtaining the estimated value of the parameter. Parameter significance tests are carried out using a significance level of 5% [9]. The hypothesis in this test is:

 H_0 : parameter is not significant H_1 : Significant Parameters

The test statistics used are

$$t = \frac{\alpha}{SE(\alpha)} \tag{3}$$

Where α in Equation 3 is the estimated parameter value and SE is the Standard Error. The decision making criterion is to reject H_0 if $|t| > t_{\frac{\alpha}{2},df}$ or p - value < 5%.

2.2.5 Model Verification

Test the white noise assumption, namely that there is no autocorrelation with the previous data residuals and follows a normal distribution. The test statistic used in this test is the Ljung-Box test statistic, the hypothesis formulation used is:

 H_0 : residuals are not autocorrelated H_1 : residuals are autocorrelated

The level of significance used is 5%. The Ljung-Box formula is expressed as follows:

$$Q = n(n+2)\sum_{k=1}^{K} (n-k)^{-1} \hat{\rho}_k^2$$
(4)

Where *n* and K in Equation (4) is the number of observations in the time series and the number of lags tested respectively, *k* is the lag difference and $\hat{\rho}_k$ is the value of the autocorrelation coefficient at lag-*k*. The decision criteria used are reject H_0 if $H_0 > x_{\alpha,db}^2$. Do not reject H_0 if $H_0 > x_{\alpha,db}^2$ with degree of freedom table (db) = K - p or p - value < 5% [9].

2.2.6 GARCH Modeling

According to Lutkepohl in 2004, this model is a development of the ARCH model, which was developed by Bollerslev and Taylor in 1986. This model was created to prevent too high orders based on the principle of parsimony selection where the simpler a statistical model is, including the number of dependent (influenced) variables which contains a lot of information that can explain the model, makes the statistical model even better. This principle also allows the GARCH method to make better forecasts with fewer variable values and avoid overfitting [11]. This method can be formulated as:

$$\sigma_{t}^{2} = \alpha_{0} + \alpha_{1}\varepsilon_{t-1}^{2} + \dots + \alpha_{p}\varepsilon_{t-p}^{2} + \beta_{1}\sigma_{t-1}^{2} + \dots + \beta_{q}\sigma_{t-q}^{2}$$
$$= \alpha_{0} + \sum_{i=1}^{p} \alpha_{i}\varepsilon_{t-i}^{2} + \sum_{j=1}^{q} \beta_{j}\sigma_{t-j}^{2}$$
(5)

where σ_t^2 in Equation (5) is the variance in period t, α_0 is a constant, $\alpha_1, \alpha_2, ..., \alpha_p$ is an indicator of ARCH, $\sigma_1, \sigma_2, ..., \sigma_q$ is an indicator of GARCH, ε_{t-p}^2 as the conditional variance value in period t - q where t moves from 1 to q.

2.2.7 ARIMA-GARCH Modelling

The ARIMA model is a forecasting model used to predict future data based on past data. The Arima model is said to be good when the resulting residuals meet the white noise assumption, that is, the residuals do not contain heteroscedasticity. When the ARIMA model residuals contain heteroscedasticity, they can be overcome with GARCH modeling. GARCH modeling is carried out by modeling the residuals of the resulting ARIMA model [11]. ARIMA-GARCH is a combination of ARIMA in Equation (2) and GARCH models in Equation (5) therefore the model form is:

$$Z_{t} = ((1 - \phi_{1})Z_{t-1} + (\phi_{1} - \phi_{2})Z_{t-2} + \dots + (\phi_{p} - \phi_{p-1})Z_{t-p} + \theta_{0} - \theta_{1}\gamma_{t-1} - \theta_{2}\gamma_{t-2} - \dots - \theta_{q}\gamma_{t-q}) - \alpha_{0} - \sum_{i=1}^{p} \alpha_{i}\varepsilon_{t-i}^{2} - \sum_{j=1}^{q} \beta_{j}\sigma_{t-j}^{2}$$
(6)

2.2.8 Mean Absolute Percentage Error (MAPE).

The MAPE criterion has advantages over the Mean Absolute Deviation (MAD) because MAPE states the percentage of error of projected results against actual observations over a certain period that will provide information that the percentage of error is too high or too low. Systematically, MAPE is expressed as follows [12].

$$MAPE = \frac{1}{2} \sum_{t=1}^{n} \left| \frac{Z_t - \widehat{Z_t}}{Z_t} \right| \times 100\%$$
(7)

Where *n* is number of data, Z_t is actual data in the tth period, \hat{Z}_t is forecasting data in the tth period. Furthermore, the MAPE value in Equation 7 can be interpreted according to the range of values that can be used as measurement material regarding the capabilities of the resulting model. The interpretation of MAPE values can be seen in Table 1.

MAPE	Interpretation	
MAPE < 10%	The prediction results are classified as very accurate	
$10 \le MAPE \le 20\%$	The prediction results are accurate	
$20 < MAPE \le 50\%$	The predicted results are feasible	
MAPE > 50%	The predicted results are poor	

Table 1 Prediction Results Criteria

3 Result And Discussion

Data analysis was carried out using R software. Based on the monthly inflation data plot in Indonesia analyzed from March 2007 to October 2023, it can be seen that inflation fluctuates every month, the lowest value reaching 1.32% in August 2020 and the highest 12.14% in September 2008. A fluctuating and erratic rate of inflation will result in relative price changes at the general price level. This is very dangerous because in a market economic system, the price level is a signal for households and for businesses about the balance of economic resource allocation in an economy [13]. Inflation developments for the period from March 2007 to October 2023 can be seen in Figure 1.

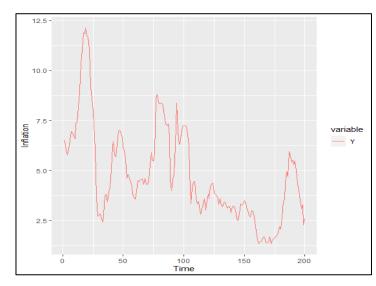


Figure 1. Inflation for the period March 2007 to October 2023

3.1 Data Stationary Test

The first stage in time series modeling is the data stationarity test. The stationary test was performed using the Augmented Dicky-Fuller (ADF) root test unit test. Data is said to be stationary if if the ADF value< the ADF Critical Value value (at a real level of 1%, 5%, 10%) and it is said that the data is not stationary if the ADF value> the ADF Critical value (at the real level of 1%, 5%, 10%). The ADF test results can be seen in Table 2 below. Based on Table 2, it can be seen that the ADF value < ADF Critical Value so it can be concluded that the data is stationary.

Table 2. The Results of ADF Test			
ADF value	ADF Critical Value	p −value	
-4.564	-3.463	0.01	

Another way to determine data stationary is to look at the p-value generated in the ADF test. Based on the results of the stationary test, inflation data can get a p-value of 0.01, meaning that the data is stationary at real level $\alpha = 1\%$ so that inflation data does not require differentensing.

3.2 Identify ARIMA Models (*p*, *d*, *q*)

ARIMA model identification is performed to determine the optimal order of the model. Inflation data does not require differencing so that order d = 0. While the order p and q are determined by looking at the smallest AIC value produced by the model. The model that has the smallest AIC value is a better model so that the order in the model is the optimal model order. The following AIC values of several models can be seen in Table 3. Based on the AIC values in Table 3, the best model is the ARIMA (2,0,2). This model has the smallest AIC value of 326.988.

lodel Value
AIC
328.100
329.199
330.551
332.193
328.738
326.988
328.513
330.390

T-LL 2 ATC M. 1.1 V.L.

ARIMA (2, 0, 2) means order p = 2, d = 0, and q = 2. The order p = 2 indicates that inflation data for periods t - 1 and t - 2 have an influence on forecasting inflation data for period t. The order d = 0 indicates that there is no data differentiation because the data is stationary. The order q = 2 indicates that the residuals for periods t - 1 and t - 12 have an influence on forecasting inflation data for period t.

3.3 **Parameter Estimation and Significance Test**

After obtaining the best model, the next stage is to estimate the parameters of the model and test the significance of these parameters to see whether the data has a significant effect or not. The estimated parameter results can be seen in Table 4.

Parameters ϕ_1 and ϕ_2 are Autoregressive parameters while parameters θ_1 and θ_2 are Moving Average parameters. Based on Table 4, it can be concluded that all parameters are significant because the p-value is less than 0.05. This indicates that the data in periods t-1 and t-2 have a significant effect on the results of subsequent data forecasting. Apart from that, the residuals in the t-1 and t-2 periods also have a significant effect.

Model	Parameter	Estimate	p – value
	Øo	4.766	9.873×10^{-13}
	Ø1	1.776	$< 2.2 \times 10^{-16}$
ARIMA (2,0,2)	Ø ₂	-0.796	$< 2.2 \times 10^{-16}$
	θ_1	-0.420	0.0005318
	$ heta_2$	-0.219	0.0078376

Table 4. Results of Estimation and Significance Test of Model Parameters

Based on the parameter estimation results, an ARIMA(2,0,2) forecasting model can be obtained below.

$$Z_t = 4.766 + (1 - 1.776)Z_{t-1} + (1.776 + 0.796)Z_{t-2} + 0.420 a_{t-1} + 0.219a_{t-2}$$

= 4.766 - 0.776Z_{t-1} + 2.572Z_{t-2} + 0.420 a_{t-1} + 0.219a_{t-2}

3.4 Model Verification

The final step in ARIMA modeling is model verification. Model verification is carried out to see if the model has a random residual (white noise) so that the model is a model that is able to explain the data well. The results of the model verification using the Ljung-Box test can be seen in Table 5.

I able 3. White Noise Test Results			
Lags	Statistics	Df	p – value
4	8.184	0	0.000000
5	9.566	1	0.001982
6	10.425	2	0.005447
7	11.445	3	0.009546
8	11.891	4	0.018177
9	12.745	5	0.025885
10	12.869	6	0.045157

Table 5. White Noise Test Results

The results of the white noise test on all tested lags resulted in a p-value of less than 0.05. This shows that the residual is not random and is indicated to contain heteroscedasticity, causing the model to be unable to explain the data properly. This can be overcome by modeling residuals using the GARCH model.

3.5 GARCH Modeling (p,q)

GARCH modeling is performed on residual data. The first step is to determine the optimal order of the model by looking at the smallest AIC value from several simulated models. The AIC values of several models can be seen in Table 6.

Table 6. GARCH Model AIC Value		
Model	AIC	
GARCH (0,1)	286.334	
GARCH (0,2)	286.815	
GARCH (0,3)	298.557	
GARCH (0,4)	297.828	
GARCH (1,1)	287.310	
GARCH (1,2)	290.091	
GARCH (1,3)	300.139	
GARCH (1,4)	299.743	

Based on Table 6 it can be seen that the best model is GARCH (0.1) with an AIC value of 286.334.

The results of parameter estimation and significance test from the GARCH(0.1) model can be seen in Table 7.

Table 7. Results of GARCH (0,1) Parameter Estimation and Significance Test

Model	Parameter	Estimate	p – value
GARCH(0,1)	a_0	0.139	2.83×10^{-11}
	a_1	0.769	8.14×10^{-07}

Based on Table 7 above, it can be seen that all parameters have a significant effect because the resulting p-value is less than 0.05. The GARCH(0,1) model formed is $\sigma_t^2 = 0.139 + 0.769\sigma_{t-1}^2$. The model diagnostic test produced a p-value > 0.05, namely 0.596. This shows that the residual model meets the white noise assumption. The final model obtained is ARIMA(2,0,2) - GARCH(0,1).

$$\begin{split} Z_t &= 4.766 - 0.776Z_{t-1} + 2.572Z_{t-2} + 0.420 \ a_{t-1} + 0.219a_{t-2} \\ &- (0.139 + 0.769\sigma_{t-1}^2) \\ &= 4.627 - 0.776Z_{t-1} + 2.572Z_{t-2} + 0.420 \ a_{t-1} + 0.219a_{t-2} - 0.769\sigma_{t-1}^2 \end{split}$$

This model has a MAPE of 17.78%. Based on the mape value criteria, it can be said that this model is classified as accurate [14]. The goodness of the model can also be seen from the forecasting results, where the forecasting results are not much different from the actual data. A plot of actual values and forecasting results can be seen in Figure 2.

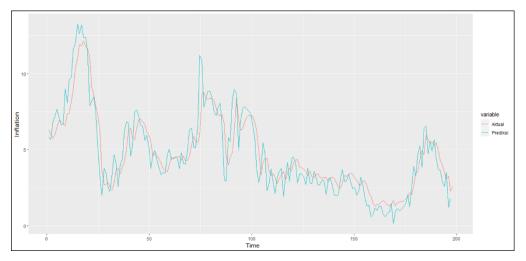


Figure 2. Plot of Actual Data and Forecasting Results

4 Conclusion

Inflation data in Indonesia from March 2007 to October 2023 experienced significant fluctuations with the lowest value reaching in August 2020 (32%) and the highest in September 2008 (12.14%). Based on the data analysis carried out, the inflation forecasting model in Indonesia is the ARIMA GARCH model, namely ARIMA(2,0,2) - GARCH(0,1)

$$Z_t = 4.276 - 0.776Z_{t-1} + 2.572Z_{t-2} + 0.420 a_{t-1} + 0.219a_{t-2} - 0.769\sigma_{t-1}^2$$

This model has a MAPE value of 17.78%.

5 References

[1] Armi A. E., Kridalaksana A. H. & Arifin Z., 2019, *Peramalan Angka Inflasi Kota Samarinda Menggunakan Metode Double Exponential Smoothing, Studi Kasus: Badan Pusat Statistik Kota Samarinda*, Jurnal Ilmiah Ilmu Komputer, **14**(1), 21-26.

- [2] Alfiyatin A. N., Mahmudy W.F., Ananda C.I.F. & Anggodo Y.P., 2019, Penerapan Extreme Learning Machine (Elm) Untuk Peramalan, Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), 6(2), 179-186.
- [3] Purba W., Nainggolan P. & Panjaitan P.D., 2022, Analisis Pengaruh Inflasi dan Pertumbuhan Ekonomi Terhadap Pengangguran di Provinsi Sumatera Utara, Jurnal Ekonomi Pembangunan, **4**(1), 62-74.
- [4] Katerina P., 2018, Chapter 10 Theory and Theoretical Frameworks, *How to Write About Economics and Public Policy*, Academic Press, pp. 181-206, ISBN 9780128130100.
- [5] Faustina R. S., Agoestanto A. & Hendikawati P., 2017, *Model Hybrid Arima-Garch Untuk Estimasi Volatilitas Harga Emas Menggunakan Software R*, UNNES Journal of Mathematics, **6**(1), 12-24.
- [6] Asmarani T. E., 2023, *Peramalan Inflasi dengan Menggunakan Metode Arima: Studi di Indonesia*, Journal on Education, **5**(2), 4684-4692.
- [7] Sanchia N. G., 2020, Implementasi Model ARIMA-GRACH Menggunakan Metode Maximum Likelihood, Undergraduate thesis, Mathematics, Universitas Islam Negeri Maulana Malik Ibrahim, Malang.
- [8] Aktivani S., 2020, *Uji Stasioneritas Data Inflasi Kota Padang Periode 2014-2019*, STATISTIKA Forum Teori dan Aplikasi Statistika, **20**(2), 83-90.
- [9] Qobilyati N., 2022, *Penerapan ARIMA Ensembel Dalam Peramalan Laju Inflasi Di Indonesia*, Undergraduate thesis, Mathematics, Universitas Lampung, Lampung.
- [10] Purwinda I., 2018, Estimasi Parameter Model Vector Autoregressive Menggunakan Metode Ordinary Least Square, Undergraduate thesis, Mathematics, Universitas Islam Negeri Maulana Malik Ibrahim, Malang.
- [11] Sirius B., Rahayu W. & Mahatma Y., 2023, Implementasi Metode ARIMA-GARCH Terhadap Peramalan Konversi Mata Uang Yen ke Rupiah, JMT (Jurnal Matematika dan Terapan), 5(2), 86 – 96.
- [12] Nabillah I. & Ranggadara I., 2020, *Mean Absolute Percentage Error untuk Evaluasi Hasil*, Journal of Information System, **5**(2), 250-255.
- [13] Aldila R. F., 2016, Dampak Inflasi Terhadap Anggaran Pendapatan Dan Belanja Daerah (Apbd) Pada Pemerintah Kota Tasikmalaya, Sekolah Tinggi Ilmu Ekonomi (STIE) Ekuitas, Bandung.
- [14] Rosa D. U., Alan M. S., Nurhidayah, Hardianti W., Rosana & Ramadhan S., 2019, Metode Exponential Smoothing dalam memproyeksikan jumlah penduduk miskin di Nusa Tenggara Barat, Jurnal Pemikiran dan Penelitian Pendidikan Matematika, 2(1), 42-53.