Jurnal Varian
Vol. 8, No. 3, Oktober 2025, pp 347–362
e-ISSN: 2581-2017

Application of ANN-GARCH on Volatility Analysis of Forecasting the Level of First Level Hospitalization Cost Claims (RITP) BPJS Health

Andi Daniah Pahrany[1], Melani Nur Wakhidah[1], Siti Mariam Binti Norrulashikin[2]
[1]Universitas Negeri Malang, Malang, Indonesia
[2]Universitas Teknologi Malaysia, Johor Bahru, Malaysia

Article Info

Article history:

Received : 07-31-2025
Revised : 10-14-2025
Accepted : 10-31-2025

Keywords:

Artificial Neural Network;
First-Level Inpatient Care;
Generalized Autoregressive Conditional Heteroscedasticity;
Volatility.

ABSTRACT

The volatility of First-Level Inpatient Care (RITP) claim costs poses a substantial challenge to BPJS Health's financial management, underscoring the need for accurate forecasting methods. This study employs Artificial Neural Network and Generalized Autoregressive Conditional Heteroscedasticity models to examine volatility dynamics and assess predictive performance. The results indicate that both models capture nonlinear patterns, heteroskedasticity, and temporal dependencies, with evidence that past fluctuations largely influence current volatility. Forecast accuracy is generally high, as reflected in the small discrepancies between predicted and actual values across most provinces. Nevertheless, the models exhibit limitations in capturing extreme peaks and troughs, where abrupt claim variations are not fully represented. These findings highlight the effectiveness of Artificial Neural Networks and Generalized Autoregressive Conditional Heteroscedasticity in modeling claim volatility, while emphasizing the need for model refinement, such as parameter optimization or integration with complementary approaches, to enhance forecasting reliability.

Accredited by Kemenristekdikti, Decree No: 200/M/KPT/2020
DOI: https://doi.org/10.30812/varian.v8i3.5422

Corresponding Author:

Andi Daniah Pahrany,
Department of Mathematics, Universitas Negeri Malang,
Email: andi.daniah,fmipa@um.ac.id

How to Cite:
Pahrany, A. D., Wakhidah, M. N., & Norrulashikin, S. M. B. (2025). Application of ANN-GARCH on Volatility Analysis of Forecasting the Level of First Level Hospitalization Cost Claims (RITP) BPJS Health. *Jurnal Varian*, *8*(3), 347–362.

A. INTRODUCTION

Health is a crucial aspect of human life that directly influences both quality of life and productivity. In Indonesia, more than 26% of the population reported health complaints in 2023, reflecting substantial demand for accessible, affordable healthcare services (Statistics Indonesia, 2023). To address this challenge, the government established the National Health Insurance (JKN) through BPJS Health under law No. 40 of 2004, which aims to ensure equitable health protection, including First-Level Inpatient Care (RITP) services (Yuwanita et al., 2022). RITP represents one of the service categories with relatively high claim rates, as it covers a wide range of essential treatments provided by primary healthcare facilities such as community health centers and clinics.

According to the Regulation of the Minister of Health of the Republic of Indonesia, first-level inpatient care refers to general individual healthcare services provided at treatment community health centers, involving at least one day of hospitalization for treatment, observation, or other medical procedures. In practice, RITP claims show significant fluctuations over time, indicating volatility in submissions by healthcare providers. Volatility itself is defined as a condition in which data experience considerable fluctuations, marked by sharp increases and decreases (Naik & Mohan, 2021). Such volatility complicates budgeting and financial planning within BPJS Health, highlighting the need for accurate forecasting methods to support sustainable management.

Previous studies provide valuable insights but leave important gaps. Research using GARCH Setiawan et al. (2021) concluded that volatility values were low and mainly influenced by prior volatility, without accounting for broader variance patterns. Kurniasari et al. (2023) demonstrated that ANN offers higher forecasting accuracy compared to GARCH. Meanwhile, Susila et al. (2023) showed that hybrid ARIMA–ANN models can improve prediction performance. Although these studies illustrate the usefulness of individual or hybrid approaches, they have not comprehensively compared ARIMA, GARCH, and ANN within the specific context of RITP claim forecasting. This creates a gap in understanding which method—or combination of methods—best captures both the linear and nonlinear characteristics, as well as the volatility inherent in RITP claim data.

To address this gap, the present study applies ARIMA, GARCH, and ANN models to analyze the volatility and forecasting accuracy of RITP claim costs in BPJS Health. ARIMA is applied to capture linear patterns and address non-stationarity in time series data (Kontopoulou et al., 2023). GARCH is applied to model volatility and heteroscedasticity, and ANN is utilized to capture nonlinear patterns beyond the scope of statistical models (Beck, 2018). By systematically evaluating these models, this research aims to identify the most effective approach for forecasting RITP claim costs, thereby contributing both theoretically to time-series forecasting literature and practically to improving financial planning within BPJS Health.

B. RESEARCH METHOD

This study employs time-series data on First-Level Inpatient Care (RITP) claim costs for the period 2016–2021, obtained from the JKN BPJS Health Statistics reports. The dataset was analyzed using ARIMA, ANN, and GARCH models, with data processing and computation conducted through Microsoft Excel, RStudio, and Minitab.

1. Autoregressive Integrated Moving Average (ARIMA)

The model was developed by George E.P. Box and Gwilym M. Jenkins in 1976. This statistical framework integrates Autoregressive (AR) and Moving Average (MA) components with differencing (I) to address non-stationary data. As noted by Perwitasari & Atikah (2020), a non-stationary time series can be modeled as an ARIMA process with parameters (p, d, q), representing the orders of the AR, differencing, and MA components, respectively. According to Kurniasari et al. (2023), the ARIMA (p,d,q) model is given in Equation (1):

$$y\prime_t = \emptyset_1 y\prime_{t-1} + \ldots + \emptyset_p y\prime_{t-p} + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q} + e_t \tag{1}$$

The $y\prime_t$ symbol is the value after differencing at time t, the $\emptyset_1, \ldots, \emptyset_p$ symbol is the AR coefficients, the $y\prime_{t-1}, \ldots, y\prime_{t-p}$ symbol is the lagged differenced data, the $\theta_1, \ldots, \theta_q$ symbol is the MA coefficients, the $e_{t-1}, \ldots, e_{t-q}$ symbol is the lagged errors, the e_t symbol is the error at time t, the p symbol is the order of the Autoregressive model, the q symbol is the order of the Moving Average model, the t symbol is the time or period in the time series.

The GARCH model will not perform well if the data are still non-stationary; therefore, the ARIMA model is used to address non-stationarity and linearity issues. The data processing method in the ARIMA model is as follows:

1. Stationarity Test

   The first step in determining the ARIMA model is to check whether the observed data is stationary. If the original data is not stationary, a differencing process is needed to achieve stationarity and handle trends (stationarity in the mean), or a transformation is applied to address non-constant variance (stationarity in the variance). With a significance level (α) of 5%, the criterion for the stationarity test is to reject H_0 if the p-value is less than α (0,05), indicating that the data is stationary. The hypotheses are:

   H_0 : $\theta = 0$, the data is non-stationary
   H_1 : $\theta \neq 0$, the data is stationary

2. Model Identification

   If the data's variance and mean are already stationary, the next step is to preliminarily identify the model by examining the values of p, d, and q. The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) are used to determine the values of p and q, while d indicates the number of differencing processes. Lags that exceed the significance bounds on the PACF plot indicate the order of the Autoregressive component. In contrast, lags that exceed the bounds on the ACF plot

DOI: https://doi.org/10.30812/varian.v8i3.5422

indicate the order of the Moving Average component. The Autocorrelation Function measures the linear relationship between X_t and X_{t+h} in a time series. According to Brockwell & Davis (2016), the ACF model is presented in Equation (2):

$$\rho\,(h) = \frac{Cov\,(\,X_t,\quad X_{t+h})}{Var(X_t)} \tag{2}$$

The $\rho\,(h)$ symbol is the autocorrelation coefficient at lag h, the X_t symbol is the value at time t, the X_{t+h} symbol is value at time $t+h$. The Partial Autocorrelation Function is used to assess the relationship between X_t and X_{t+h} after removing the effects of intermediate lags $X_{t+1},\ X_{t+2},\ \ldots,\ X_{t+h-1}$. According to Brockwell & Davis (2016), the PACF model is presented in Equation (3):

$$\emptyset_h = Corr\ \left(X_t,\ X_{t+h}\,|\,X_{t+1}\,,\ X_{t+2}\,,\ \ldots,\ X_{t+h-1}\right) \tag{3}$$

Where $\emptyset_h$ is the partial autocorrelation coefficient at lag h, while X_t is the value at time t, and X_{t+h} is the value at time $t+h$.

3. Parameter Estimation

This study uses Maximum Likelihood Estimation (MLE) to estimate the preliminary model because MLE produces parameter estimates that are close to the true values, and as the sample size increases, the estimates converge to the actual parameter values. The hypothesis for the significance test of the parameters is as follows:

H_0 : $\theta = 0$, the parameter is not significant
H_1 : $\theta \neq 0$, the parameter is significant

At a significance level $\alpha = 5\%$, the criterion for the parameter estimation test is to reject H_0 if p-value $< \alpha = 0.05$ indicating that the model is significant. MLE maximizes the likelihood function $L(\theta)$, where θ is the parameter to be estimated. The likelihood function $L(\theta)$ is the probability of observing the data $x_1,\ x_2,\ldots,\ x_n$ given the parameter θ expressed as according to Svalova et al. (2021):

$$L\,(\theta) = P\,(X = x_1,\ \ldots,\ X = x_n|\theta) = \sum_{i=1}^{n} f(x_i|\theta) \tag{4}$$

Where $f(x_i|\theta)$ is the probability function with parameter θ and n is the number of observations. The Probability Density Function (PDF) is differentiable with respect to θ, and the estimation is done by solving Equation (5):

$$\frac{\partial \ln L(\theta)}{\partial \theta} \tag{5}$$

$L(\theta)$ is the maximum likelihood function and θ is the parameter to be estimated. The estimate of θ is $\hat{\theta} = \hat{\theta}\ (X_1,\ X_2,\ \ldots,\ X_n)$, where $\hat{\theta}$ maximizes the likelihood function $L(\theta)$. $\hat{\theta}$ is called the Maximum Likelihood Estimator of θ.

4. Diagnostic Checking

Diagnostic checking involves residual analysis, including the white noise test and residual normality test, to ensure that the model is appropriate for use as a forecasting model. The selected model must satisfy the assumptions of white noise and residual normality. The hypotheses are:

H_0 : $\theta = 0$, does not satisfy the assumptions of white noise and residual normality
H_1 : $\theta \neq 0$, satisfies the assumptions of white noise and residual normality

At a significance level $\alpha = 5\%$, he criterion for the diagnostic checking test is to reject H_0 if p-value $>\ \alpha = 0,05$ indicating the model is significant. A commonly used method for the white noise test is the Ljung-Box Test, with the test statistic defined by to Kong & Lund (2023) the model is presented in Equation (6):

$$Q = n\,(n+2)\sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{n-k} \tag{6}$$

The Q symbol is the Ljung-Box test statistic, the n symbol is the number of observations, the $\hat{\rho}_k$ symbol is the residual autocorrelation at lag k, the m symbol is the number of lags tested. The residual normality test in this study is conducted using the Kolmogorov-Smirnov test, with the statistic defined by Monti et al. (2017) the model is given in Equation (7):

DOI: https://doi.org/10.30812/varian.v8i3.5422

$$D_{KS} = \sup_{x \in \mathbb{R}} |F_n(x) - f(x)| \tag{7}$$

$$\text{where,} \quad F_n(x) = \frac{1}{n} \sum_{i=1}^{n} IX_i \leq x \tag{8}$$

D_{KS} is the Kolmogorov-Smirnov test statistic, $F_n(x)$ is the Empirical Cumulative Distribution Function (ECDF) of the residual data, $f(x)$ is the Cumulative Distribution Function (CDF) of the normal distribution, and I is the indicator function.

5. Selection of the Best Model

The best model is selected by evaluating the Mean Squared Error (MSE). The model with the lowest MSE will be chosen as the optimal model. MSE is calculated by summing the squared differences between the original and forecast data across a set of processed data points. According to Hodson (2022), the MSE equation is presented in Equation (9):

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \widehat{Y}_i \right)^2 \tag{9}$$

Where Y_i is the original observed data, while $\widehat{Y}_i$ is the forecasted data, and n is the number of data points.

2. Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

The GARCH model was developed by Bollerslev in 1986 as a generalization of the Autoregressive Conditional Heteroscedasticity (ARCH) model. The ARCH model was introduced by Engle in 1982 as a time-series model that addresses heteroscedasticity. GARCH assumes that the variance of fluctuating data is influenced by p lagged squared residuals and q lagged variances from previous periods (Susanti et al., 2024). According to Petkov et al. (2024), the GARCH (p, q) model is presented in Equation (10):

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \ldots + \alpha_p \epsilon_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_q \sigma_{t-q}^2 \tag{10}$$

The σ_t^2 symbol is the volatility at time t, the ω symbol is the constant, the $\alpha_1, \ldots, \alpha_p$ symbol is the ARCH model parameters, the $\epsilon_{t-1}^2, \ldots, \epsilon_{t-p}^2$ symbol is the lagged squared residuals from time (t–1 to t–p), the $\beta_1, \ldots, \beta_q$ symbol is the GARCH model parameters, the $\sigma_{t-1}^2, \ldots, \sigma_{t-q}^2$ symbol is the lagged conditional variances from time (t–1 to t–q).

In the GARCH model, the PACF is used to assess the extent to which past volatility influences future volatility, assuming that the effects of intermediate lags are independent. PACF helps determine the order p of the GARCH model by identifying cutoff points at certain lags. Ali et al. (2022) highlight several advantages of the GARCH model:

a. It treats heteroscedasticity not as a problem but as a feature to build the model.
b. Besides forecasting based on variables, GARCH also forecasts based on variance.

Residual data are used for GARCH modeling because they still exhibit volatility or instability after stationarity has been achieved with ARIMA. GARCH captures the patterns of variance fluctuation in the residuals. The Lagrange Multiplier (LM) test is used to examine the homogeneity of residual variance, indicating the potential presence of ARCH/GARCH effects (Sari et al., 2024). Kanal et al. (2018) explain that the ARCH-LM test concept is based on the idea that residual variance depends not only on independent variables but also on the squared residuals from previous periods.

3. Artificial Neural Network (ANN)

Artificial Neural Network can be considered an artificial representation of the human brain that attempts to simulate the learning processes occurring in the human brain. ANN recognizes patterns in past data to learn from them, enabling it to make decisions on data that it has never encountered before. According to Sako et al. (2022), neural networks can be used for forecasting based on patterns in historical data. Kurniasari et al. (2023) argue that there are several types of ANN, but all share the same components. ANN consists of three layers: the input layer connected to the hidden layer, the hidden layer connected to the output layer, and neurons that connect each layer (Rahmawati & Lestari, 2019). Figure 1 present an illustration showing the similarity between the components of the human neural system and the ANN model.

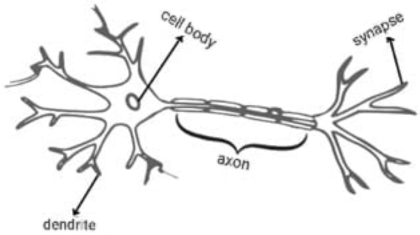

Figure 1. Human Neural Network and ANN

The components of an Artificial Neural Network, as shown in Figure 1, are as follows: The input is similar to dendrites in the human neural network; the input in an ANN function as the receiver of information from other neurons. An Artificial Neuron is the basic unit, called a neuron, that serves as the processing element within the network where all calculations take place. Weights in an ANN are a series of numbers that are crucial for optimizing the system and enabling it to process input data accurately to produce the desired output. Activation Function is a value that represents the result of mapping the function from the total sum received from all inputs to the neuron. Output is the solution to the problem, presented as numerical data.

ANN modeling uses actual RITP claim data. Suppose the modeling uses residual data from the ARIMA-GARCH model. In that case, the ANN model does not capture the original data pattern but only the error pattern from the ARIMA-GARCH model.

a. Backpropagation

The backpropagation algorithm, introduced in the 1970s and popularized by David Rumelhart, Geoffrey Hinton, and Ronald Williams in 1986 (Sekhar & Meghana, 2020) is a widely used method in ANN. ANN involves two main calculation processes: forward and backward. In forward calculation, the network computes the output and compares it to the desired target to determine the error. Backpropagation training then works backward, using the output error to adjust the network's weights and minimize it, aiming to reach the optimal minimum. This method typically involves one or more hidden layers and applies the backward propagation process to reduce errors continuously (Jauhari et al., 2016).

b. Data Scaling

Data scaling is the process of adjusting the range of numerical feature values in a dataset so that they share the same scale. There are several data scaling techniques; in this study, the author uses min-max normalization. Min-max normalization is a data scaling technique that transforms numerical data values to a specific range. In this study, the range is 0 to 1. The formula for min-max normalization is given in Equation (11):

$$X_{normalized} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{11}$$

Where X is the original value, X_{min} is the minimum value, and X_{max} is the maximum value.

c. Outlier

An outlier is a data point that has a value significantly different from other data points in a dataset. To handle outliers, this study uses the Interquartile Range (IQR) method, which replaces outlier values with the median.

d. Hidden Layer

The ability of a neural network to generalize is supported by its hidden layers. ANNs with one or two hidden layers are commonly used and have been shown to perform well. However, increasing the number of hidden layers can increase computation time and the risk of overfitting, which can reduce prediction accuracy (Mayatopani, 2021). Besides determining the number of hidden layers, it is also important to set the number of neurons in the model. If the number of neurons is too low, it can lead to underfitting. Conversely, if the number of neurons is too large, the model may experience overfitting.

An ANN model denoted as (2,1) has one input layer, two hidden layers, one neuron, and one output layer. The neural network model equation with one input layer and two hidden layers, according to Kurniasari et al. (2023) is as follows:

$$y_k = \sum_{l=1}^{2} w_{kl} \, . \, f\left[w_{jo} + \sum_{j=1}^{2} w_{kl} \, . \, f\left[v_{j0} + \sum_{i=1}^{2} x_i v_{ji}\right] + w_{k0}\right] \tag{12}$$

Where y_k is the output of the network for output neuron k, w_{kl} is the weight between neuron l in the hidden layer and neuron k in the output layer, f is the nonlinear activation function in the neural network, w_{jo} is the bias of the output layer neuron j, v_{j0} is the bias of neuron in hidden layer j, x_i is the input to the network, v_{ji} is the weight between input xi and neuron in hidden layer j, w_{k0} is the bias in the output layer neuron k

e. Forecasting

Forecasting is an effort to predict values or conditions that will occur in the future, based on the assumption that patterns or trends observed in the past will continue into the future (Makridakis et al., 2020). The goal of forecasting is to predict future values of a time series, x_{n+m}, $m = 1, 2, \ldots$, based on data collected up to the present, $x = x_n, x_{n-1}, \ldots, x_1$ (Shumway & Stoffer, 2017).

4. Methodology Flowchart

The methodology flowchart of this research is illustrated in Figure 2.

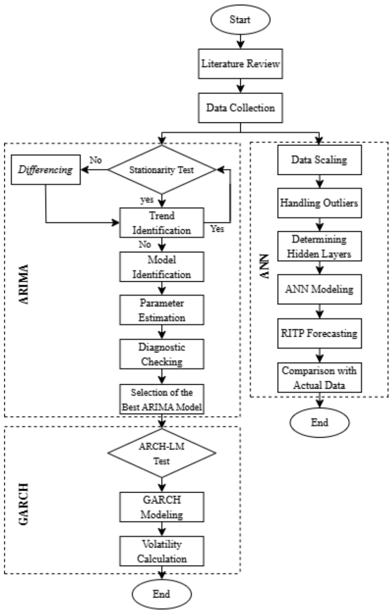

Figure 2. Research Methodology

C. RESULT AND DISCUSSION

1. Data Collection

The research data were obtained from the JKN BPJS Health Statistics book for the years 2016–2021. The data used consists of the claim rate for RITP costs in Indonesia from 2016 to 2021, and volatility analysis and short-term forecasting will be conducted. From Figure 3, the data are not stationary in terms of both the variance and the mean. The data also appears to have no seasonality and no trend. However, to gain a deeper understanding, the data will be further examined.

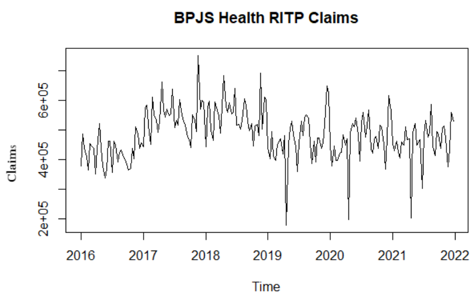

Figure 3. Plot of RITP Claim Data from BPJS Health for the Years 2016-2021

DOI: https://doi.org/10.30812/varian.v8i3.5422

2. Autoregressive Integrated Moving Average (ARIMA)

a. Data Stationarity Test

The stationarity test is used to determine whether the data contains seasonality and trend. Stationarity of the data can be verified by examining the data plot and by conducting the Augmented Dickey-Fuller (ADF) test. According to Maitra & Politis (2024), the ADF test is a statistical test used to assess the stationarity of the data's mean. Based on the ADF test, the p-value obtained was 0.1349, and Figure 3 indicates that the data is non-stationary in the mean. After applying first-order differencing, the p-value decreased to 0.01, and Figure 4 demonstrates that the data became stationary in the mean.

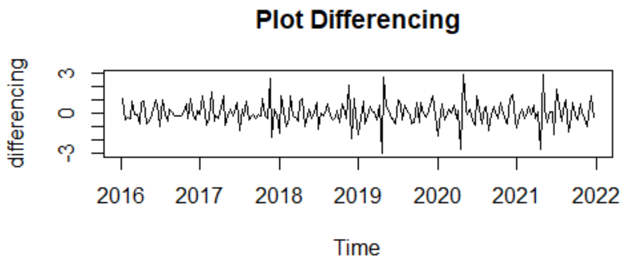

Figure 4. Plot of Differenced RITP Claim Data for 2016-2021

b. Trend Identifications

To prove the stationarity of the data with respect to the mean, refer to Figure 5. The data approach a mean of zero, and the plot shows no trend. That means that after one differencing process, the data are stationary in mean, as shown in Figure 5.

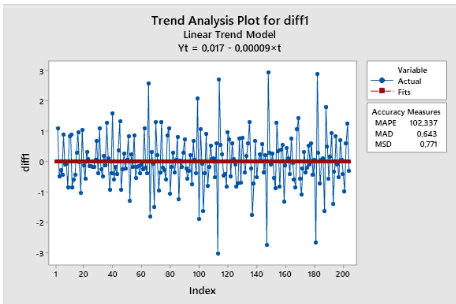

Figure 5. Trend Differencing Plot of RITP Claim Data for the Years 2026-2021

3. Trend Identifications

Non-stationary data can be modeled using ARIMA (p, d, q). To identify the ARIMA model, ACF and PACF are used. The p lag that exceeds the significance boundary in the PACF plot indicates the order of AR, while the q lag that crosses the boundary in the ACF plot shows the order of MA.

In Figure 6, there is a lag-2 that exceeds the boundary in the ACF plot, and in the PACF plot it can be said that the lags experience a "dies down" or a decline in correlation between lags to a value of zero or become insignificant. Thus, the AR order is 0 and the MA order is 2. The possible ARIMA (p, d, q) models that can be formed, based on Figure 6, include ARIMA (0,1,2), ARIMA (2,1,1), ARIMA (2,1,2), and ARIMA (2,1,0).

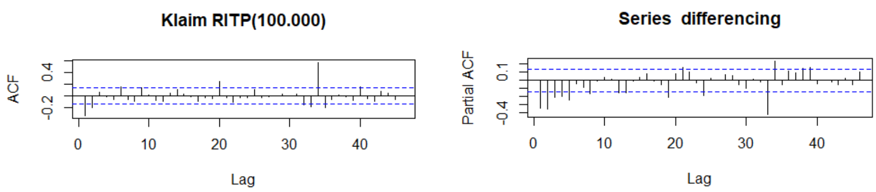

Figure 6. ACF and PACF Plots of RITP Claim Data

Vol. 8, No. 3, Oktober 2025, pp 347–362
DOI: https://doi.org/10.30812/varian.v8i3.5422

a. Parameter Estimation

The provisional model will be tested for parameter significance using the Maximum Likelihood method. The model is considered significant if the p-value is less than α. Based on Table 1, the significant models obtained are ARIMA (0,1,2) and ARIMA (2,1,0).

Table 1. Parameter Significance Test

Model	Parameter	Coefficient	p-value	Description
ARIMA (0,1,2)	MA(1)	-0.661849	2.2×10^{-16}	Significant
	MA(2)	-0.195358	0.008377	
ARIMA (2,1,2)	AR(1)	-0.039117	0.9405	Not Significant
	AR(2)	-0.099316	0.4716	
	MA(1)	-0.640929	0.2229	
	MA(2)	-0.182034	0.6985	
ARIMA (2,1,1)	AR(1)	0.162107	0.05062	Not Significant
	AR(2)	-0.138229	0.07886	
	MA(1)	-0.845083	2×10^{-16}	
ARIMA (2,1,0)	AR(1)	-0.464287	1.545×10^{-12}	Significant
	AR(2)	-0.357125	5.581×10^{-8}	

b. Diagnostic Checking

Diagnostic checking is conducted by assessing the model's adequacy with the White Noise test and the normality of the model residuals using the Kolmogorov-Smirnov test. The model is considered to pass the diagnostic checking if the p-value is greater than α. Based on Table 2, the model that passed the diagnostic checks is ARIMA(0,1,2).

Table 2. Parameter Significance Test

Model	Uji White Noise	Uji Kolmogorov-Smirnov	Result
ARIMA (0,1,2)	0.789	0.08926	Satisfied
ARIMA (2,1,0)	0.2871	0.006053	Not Satisfied

c. Selection of the Best ARIMA Model

After conducting several tests, the best model obtained is ARIMA (0,1,2). This can be seen in the following Table 3. ARIMA (0,1,2) is the best model because it has the highest log-likelihood, the smallest AIC, and the lowest MSE. Based on the data processing, the following is the equation for the ARIMA (0,1,2) model: $y_t = y_{t-1} - 0,6618e_{t-1} - 0,1954e_{t-2} + e_t$.

Table 3. Best ARIMA Model

Model	Log Likelihood	MSE	AIC
ARIMA (0,1,2)	-168.55	0.506	347.1
ARIMA (2,1,0)	-235.28	0.5934	476.55

4. Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

a. Test for Heteroscedasticity Effect

Heteroscedasticity can be tested using the ARCH-LM test on the residuals of the best-fitting ARIMA model. Based on Table 4, the test results reject H_0 and indicate heteroscedasticity in the data. To address heteroscedasticity, this study will incorporate a GARCH model.

Table 4. Heteroscedasticity Test

Lag	P-Value
4	0.00000
8	4.44×10^{-15}
12	2.30×10^{-7}
16	1.17×10^{-2}
20	3.64×10^{-2}
24	9.27×10^{-3}

b. GARCH Modeling

After determining the best ARIMA model and conducting the heteroscedasticity effect test (ARCH), GARCH modeling is performed to handle heteroscedasticity in the data. The models to be tested in this study are GARCH(1,0), GARCH(1,1), GARCH(1,2), and GARCH(2,1).

Table 5. GARCH Modeling

Model	Log Likelihood	AIC
GARCH (0,1)	-217.9896	2.1862
GARCH (1,1)	-217.5511	2.1917
GARCH (1,2)	-217.5441	2.2014
GARCH (2,1)	-217.5442	2.2014

In Table 5, the GARCH model parameters must be significant. In the GARCH(1,2) and GARCH(2,1) models, some parameters have undefined standard errors (NaN), indicating model instability. The GARCH(1,1) model is the best because it has the highest log-likelihood, a competitive AIC value, and significant parameters.

c. Volatility Calculation

The best model, GARCH(1,1), will be used to estimate RITP volatility. Based on the data processing, the following is the equation for the GARCH(1,1) model.

$$\sigma_t^2 = 0.00044443 + 0.00000001.\epsilon_{t-1}^2 + 0.99992691.\sigma_{t-1}^2$$

Based on Equation (10), the volatility value is $\sqrt{\sigma_t^2} = \sqrt{0.00088883} = 0.0298$

The GARCH(1,1) model has an MSE of 0.5502386. The parameter 0.99992691 is very large, indicating that future volatility is highly influenced by past volatility. If past volatility is high, future volatility will also be high, and vice versa.

5. Artificial Neural Network (ANN)

a. Data Scaling

After selecting the best GARCH model, the data will be reprocessed using an ANN with the backpropagation algorithm. This study applies the min-max normalization transformation method. In Figure 7, the min-max normalization results in values within the interval for both the lagged and original data. The mean value for the first lag is 0.5343751, and the mean value for the original data is 0.5356857. There is no significant difference between the minimum, first quartile, median, mean, third quartile, and maximum values between the lagged data and the original data.

```
   Lag 1...5          Klaim RITP...6
 Min.    :0.0000782   Min.    :0.0000782
 1st Qu.:0.4544567    1st Qu.:0.4562848
 Median :0.5298592    Median :0.5300348
 Mean   :0.5343751    Mean   :0.5356857
 3rd Qu.:0.6161709    3rd Qu.:0.6161709
 Max.   :0.9999166    Max.   :0.9999166
```

Figure 7. Normalization of RITP Claim Data

b. Handling Outliers

The outlier data in this study include Maluku in 2017 (7.52352) and in 2018 (6.92061). DKI Jakarta in 2018 with a value of 6.84733, in 2019 with a value of 1.77245, in 2020 with a value of 1.9513, and in 2021 with a value of 2.02249. Bali in 2021 had a value of 3.01654. These data were replaced with the median value of 4.84796.

c. Determining the Hidden Layer

Neural networks with one or two hidden layers are more commonly used and have been proven to perform well. Table 6 shows that using 2 hidden layers with 1 neuron is suitable for modeling RITP, as it yields the smallest error.

Table 6. Comparison of Hidden Layers

Hidden Layer	Neuron	Error
1	1	1.53693
1	2	1.633428
2	1	1.525658
2	2	1.622917

d. ANN Modeling

The ANN model will be used as the forecasting model for RITP. Figure 8 shows the ANN (2,1) model, which has one input layer, two hidden layers, one neuron, and one output layer. The MSE value of the ANN (2,1) model is 0.006663954. The data used to build the ANN model with one input layer and two hidden layers is the RITP claim data from Aceh 2016 to Aceh 2021. The input value for predicting RITP in North Sumatra 2021 is the RITP claim data for Aceh 2021. The RITP claim value for Aceh 2021 is 4.5278, and the normalized RITP claim value for Aceh 2021 is 0.479103.
The equation for the first hidden layer solution is as follows:

$$z_{net_j} = v_{j0} + \sum_{i=1}^{2} x_i v_{ji} = \begin{bmatrix} -4.23239 \\ -4.7763 \end{bmatrix} + \begin{bmatrix} 5.61428 \\ 7.09986 \end{bmatrix} [0.479103] = \begin{bmatrix} -4.23239 \\ -4.7763 \end{bmatrix} + \begin{bmatrix} 2.689818 \\ 0.340156 \end{bmatrix} = \begin{bmatrix} -1.542572 \\ -4.436144 \end{bmatrix}$$

Use the activation function

$$z_j = f(z_{net_j})$$

$$z_j = f(z_{net_j}) = \begin{Bmatrix} 0,01x \; ; x < 0 \\ x \; ; x \geq 0 \end{Bmatrix} = f\begin{pmatrix} -0.01542572 \\ -4.436144 \end{pmatrix} = \begin{bmatrix} -0.01542572 \\ -4.436144 \end{bmatrix}$$

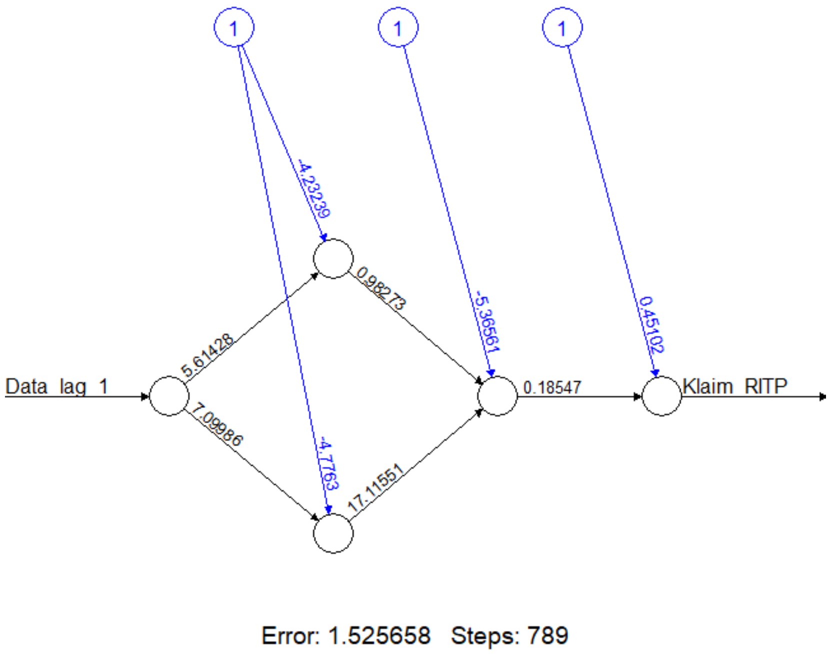

Figure 8. ANN (2,1) Model of RITP Claim Data 2016-2021

Use the output from the first hidden layer to calculate the second hidden layer:

$$q_{net_j} = w_{j0} + \textstyle\sum_{j=1}^{2} w_{kl} z_j = [-5.36561] + \left(\begin{bmatrix} 0.98273 & 17.11551 \end{bmatrix} \begin{bmatrix} -0.01542572 \\ -4.436144 \end{bmatrix} \right)$$

$$= [-5.36561] + [-75.94203]$$

$$q_j = f(q_{net_j}) = [-0.08130764]$$

DOI: https://doi.org/10.30812/varian.v8i3.5422

Use the output from the hidden layer and neuron to calculate the output layer:
$y_k = w_{ko} + \sum_{i=1}^{2} w_{kl} q_j = 0.45102 + ([0.18547]\,[-0.08130764]) = 0.43593987$
The obtained value will be normalized back to the original data scale using the equation:
$x = [y_k\,(x_{max} - x_{min})] + x_{min} = [0.43593987(7.52352 - 1.77245)] - 1.77245 = 4.279570$

The predicted RITP claim value for North Sumatra in 2021 using the backpropagation algorithm is 4.279570. This algorithm effectively minimizes error by optimally adjusting weights, resulting in highly accurate predictions, as evidenced by the forecasted value being very close to the actual RITP claim value of 4.29347.

e. Forecasting of First-Level Inpatient Claims

After obtaining the best ANN model, the forecasting results generated in RStudio are presented in Table 7. Based on Table 7, the forecasting results vary across regions, with Aceh in 2021 having a forecasted RITP claim value of IDR 448,923.3, and Lampung in 2021 having a forecasted RITP claim value of IDR 508,951.5.

Table 7. Forecast and Actual Data of RITP Claims for 2019-2021 (100,000)

Province	Forecast			Actual Data		
	2019	2020	2021	2019	2020	2021
Aceh	4.384017	4.474678	4.489233	4.40585	4.4917	4.5278
Sumatera Utara	4.014088	3.877205	4.226144	4.0167	3.77918	4.29347
Sumatera Barat	4.85697	4.349535	4.515743	4.93173	4.46841	4.60952
Riau	4.118263	3.979403	4.194207	4.14373	3.96038	4.25195
Jambi	3.975359	3.969601	4.038382	3.96361	3.95933	4.05848
Sumatera Selatan	4.386747	4.139211	4.471348	4.49794	4.21032	4.57869
Bengkulu	4.506938	4.179113	4.407674	4.5909	4.25048	4.48887
Lampung	4.639661	4.75926	5.089515	4.71051	4.83413	5.10441
Kep. Bangka Belitung	4.153892	4.405067	4.609522	4.1981	4.47474	4.66093
Kepulauan Riau	4.736017	4.621663	4.639821	4.81651	4.69964	4.70742
DKI Jakarta	4.77445	4.768785	4.769172	4.84796	4.84796	4.84796
Jawa Barat	4.25912	4.64041	4.670928	4.49765	4.90468	4.92624
Jawa Tengah	5.011712	5.228495	5.247391	5.03825	5.20824	5.22406
DI Yogyakarta	5.318462	5.132843	4.440585	5.28384	5.11795	4.49425
Jawa Timur	4.805615	5.441695	4.469414	4.83114	5.39352	4.55458
Banten	4.353031	4.820191	4.584262	4.42099	4.8407	4.6613
Bali	3.789796	3.982231	4.766856	3.58985	3.95296	4.84796
Nusa Tenggara Barat	4.431203	5.22611	4.65774	4.5652	5.24631	4.82451
Nusa Tenggara Timur	5.331319	5.638094	5.348507	5.30778	5.58038	5.31505
Kalimantan Barat	4.794727	4.719805	4.705131	4.82076	4.7454	4.74048
Kalimantan Tengah	5.467826	5.133685	4.87106	5.42175	5.1318	4.90786
Kalimantan Selatan	5.558669	5.732381	5.907674	5.50391	5.67479	5.86103
Kalimantan Timur	5.463683	4.316883	4.427873	5.41344	4.34804	4.45749
Kalimantan Utara	4.531478	4.156731	4.087832	4.57584	4.21653	4.12139
Sulawesi Utara	3.917624	4.579201	4.892112	3.85265	4.67358	4.9558
Sulawesi Tengah	4.488604	4.724302	4.835361	4.6072	4.78269	4.86849
Sulawesi Selatan	3.95965	4.311035	4.300656	3.92395	4.37816	4.36315
Sulawesi Tenggara	4.606939	5.137608	5.050567	4.71567	5.14996	5.07728
Gorontalo	4.660979	5.1063	5.141159	4.72414	5.09628	5.12822
Sulawesi Barat	4.309197	4.44993	4.681089	4.37871	4.50872	4.72542
Maluku	4.482642	3.820416	3.860891	4.57317	3.66042	3.74191
Maluku Utara	5.17747	4.611005	4.241518	5.17546	4.73557	4.34816
Papua Barat	6.344068	6.157084	5.654304	6.48087	6.16611	5.60632
Papua	6.110844	5.638341	5.344146	6.2496	5.60902	5.30234

f. Comparison with Actual Data

After obtaining the forecasting results from the ANN model, a comparison will be made to assess the accuracy of the forecasts relative to the actual data. The following shows the percentage differences between the forecasted RITP values and the actual RITP data:

DOI: https://doi.org/10.30812/varian.v8i3.5422

Table 8. Percentage Comparison of Actual Data for 2019-2021

Province	2019	2020	2021
Aceh	0.49%	0.38%	0.85%
Sumatera Utara	0.06%	2.53%	1.59%
Sumatera Barat	1.52%	2.73%	2.03%
Riau	0.62%	0.48%	1.38%
Jambi	0.29%	0.26%	0.49%
Sumatera Selatan	2.53%	1.72%	2.34%
Bengkulu	1.83%	1.71%	1.84%
Lampung	1.5%	1.55%	0.29%
Kep. Bangka Belitung	1.06%	1.58%	1.1%
Kepulauan Riau	1.67%	1.66%	1.44%
DKI Jakarta	1.54%	1.66%	1.65%
Jawa Barat	5.6%	5.39%	5.18%
Jawa Tengah	0.53%	0.38%	0.45%
DI Yogyakarta	0.65%	0.29%	1.21%
Jawa Timur	0.53%	0.89%	1.87%
Banten	1.56%	0.42%	1.65%
Bali	5.27%	0.74%	1.7%
Nusa Tenggara Barat	2.94%	0.38%	3.46%
Nusa Tenggara Timur	0.44%	1.03%	0.63%
Kalimantan Barat	0.54%	0.54%	0.75%
Kalimantan Tengah	0.85%	0.04%	0.75%
Kalimantan Selatan	0.99%	1.01%	0.79%
Kalimantan Timur	0.93%	0.72%	0.67%
Kalimantan Utara	0.97%	1.44%	0.82%
Sulawesi Utara	1.66%	2.02%	1.28%
Sulawesi Tengah	2.57%	1.22%	0.68%
Sulawesi Selatan	0.9%	1.56%	1.45%
Sulawesi Tenggara	2.31%	0.24%	0.53%
Gorontalo	1.34%	0.19%	0.25%
Sulawesi Barat	1.61%	1.32%	0.94%
Maluku	1.98%	4.18%	3.08%
Maluku Utara	0.04%	2.63%	2.51%
Papua Barat	2.11%	0.15%	0.86%
Papua	2.22%	0.52%	0.79%

Based on Table 8, the forecasting results do not differ significantly from the RITP claim data in most provinces. This indicates that the RITP's forecasting accuracy is quite good. Overall, the model used is capable of producing sufficiently accurate predictions for BPJS Kesehatan RITP claims. Figure 9 shows a plot comparing forecasted data with RITP claim data, where the blue line represents the actual RITP claim data and the red line represents the predictions.

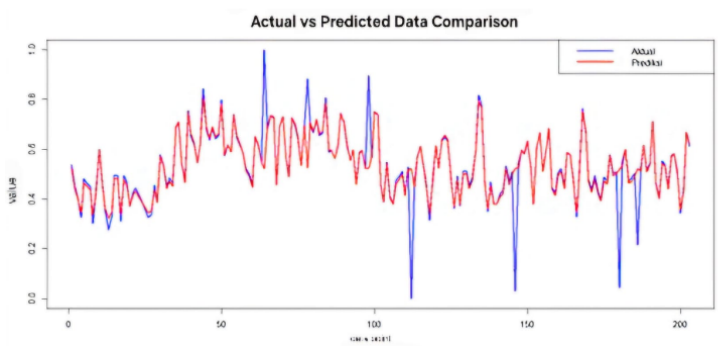

Figure 9. Forecasting of BPJS Kesehatan RITP Claims for 2016-2021

Based on Figure 9, the forecasted data and actual RITP claim data show that the ANN model captures the general pattern of RITP data, as indicated by the red prediction line closely following the blue actual data line. This study makes a significant contribution to the comparative analysis of ARIMA, GARCH, and ANN models in forecasting the volatility of RITP claim costs. The findings reveal that although these models can capture general patterns in the data, they exhibit limitations in representing extreme fluctuations. This outcome underscores the necessity for further model refinement, such as the inclusion of additional parameters or the development of hybrid modeling approaches, to enhance forecasting accuracy and reliability.

D. CONCLUSION AND SUGGESTION

Based on the volatility analysis of BPJS Kesehatan RITP claim data forecasts from 2016 to 2021, the best ARIMA model is ARIMA (0,1,2), which effectively captures the linear pattern in the data, with an MSE of 0.506. The GARCH model is used to address heteroscedasticity, and GARCH(1,1) is the best, with an MSE of 0.5502. This model captures volatility well and shows that claim volatility is strongly influenced by past volatility. The ANN model obtained is (2,1) with an MSE of 0.0066, indicating a very low prediction error and producing forecasts that are quite significant compared to the RITP claim data. However, there are some deviations at sharp peaks and valleys, where the prediction does not fully capture extreme fluctuations and the actual dynamics; thus, further evaluation of the forecasting model is needed.

For future research, it is recommended to use additional methods and variables, such as the LSTM method and BPJS policies, as well as economic conditions in each region. Based on the research results, BPJS Kesehatan should set aside appropriate reserve funds to anticipate an increase in claims in the future, allocate resources (medical personnel, healthcare facilities, and medicines) to regions experiencing an increase in RITP claims, and enhance disease prevention programs that can reduce the need for RITP, thereby lowering future RITP costs.

ACKNOWLEDGEMENT

The author would like to express sincere gratitude to the Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Malang (UM), for their continuous support throughout this research. Special thanks to all individuals and institutions who contributed valuable insights and resources during the study. This work would not have been possible without their cooperation and assistance.

DECLARATIONS

AUTHOR CONTIBUTION

First Author: Conceptualization, methodology, supervision, and funding acquisition. Second Author: Data analysis, software development, writing of the original draft, and validation. Third Author: Literature review, data curation, manuscript review, and refinement. All authors discussed the results and contributed to the final manuscript.

FUNDING STATEMENT

This research was self-funded by the author.

COMPETING INTEREST

The author declares no competing interests in this article.

REFERENCES

Ali, F., Suri, P., Kaur, T., & Bisht, D. (2022). Modelling time-varying volatility using GARCH models: Evidence from the Indian stock market. *F1000Research*, *11*, 1098. https://doi.org/10.12688/f1000research.124998.2

Beck, M. W. (2018). NeuralNetTools : Visualization and Analysis Tools for Neural Networks. *Journal of Statistical Software*, *85*(11). https://doi.org/10.18637/jss.v085.i11

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to Time Series and Forecasting*. Springer International Publishing. https://doi.org/10.1007/978-3-319-29854-2

Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development*, *15*(14), 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

Jauhari, D., Himawan, A., & Dewi, C. (2016). Prediksi Distribusi Air PDAM Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation Di PDAM Kota Malang. *Jurnal Teknologi Informasi dan Ilmu Komputer*, *3*(2), 83. https://doi.org/10.25126/jtiik.201632155

Kanal, F. A., Manurung, T., & Prang, J. D. (2018). Penerapan Model GARCH (Generalized Autoregressive Conditional Heteroscedasticity) dalam Menghitung Nilai Beta Saham Indeks PEFINDO25. *Jurnal Ilmiah Sains*, *18*(2), 67. https://doi.org/10.35799/jis.18.2.2018.19732

Kong, J., & Lund, R. (2023). Seasonal count time series. *Journal of Time Series Analysis*, *44*(1), 93–124. https://doi.org/10.1111/jtsa.12651

Kontopoulou, V. I., Panagopoulos, A. D., Kakkos, I., & Matsopoulos, G. K. (2023). A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks. *Future Internet*, *15*(8), 255. https://doi.org/10.3390/fi15080255

Kurniasari, D., Mukhlisin, Z., Wamiliana, W., & Warsono, W. (2023). Performance of the Accuracy of Forecasting the Consumer Price Index Using the GARCH and ANN Methods. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, *17*(2), 0931–0944. https://doi.org/10.30598/barekengvol17iss2pp0931-0944

Maitra, S., & Politis, D. N. (2024). Prepivoted Augmented Dickey-Fuller Test with Bootstrap-Assisted Lag Length Selection. *Stats*, *7*(4), 1226–1244. https://doi.org/10.3390/stats7040072

Makridakis, S., Hyndman, R. J., & Petropoulos, F. (2020). Forecasting in social settings: The state of the art. *International Journal of Forecasting*, *36*(1), 15–28. https://doi.org/10.1016/j.ijforecast.2019.05.011

Mayatopani, H. (2021). Implementation of ANN and GARCH for Stock Price Forecasting. *Journal of Applied Data Sciences*, *2*(4), 109–134. https://doi.org/10.47738/jads.v2i4.41

Monti, G. S., Mateu Figueras, G., Ortego Martínez, M. I., Pawlowsky Glahn, V., & Egozcue Rubí, J. J. (2017). Modified Kolmogorov-Smirnov test of goodness of fit, 152–158. https://hdl.handle.net/2117/105422

Naik, N., & Mohan, B. R. (2021). Stock Price Volatility Estimation Using Regime Switching Technique-Empirical Study on the Indian Stock Market. *Mathematics*, *9*(14), 1595. https://doi.org/10.3390/math9141595

Perwitasari, A., & Atikah, N. (2020). Model Fungsi Transfer Multi Input dalam Peramalan Penerimaan Pajak Hotel Kota Malang Tahun 2019. *Jurnal Kajian Matematika dan Aplikasinya (JKMA)*, *1*(1), 1–9. https://doi.org/10.17977/um055v1i12020p1-9

Petkov, P., Shopova, M., Varbanov, T., Ovchinnikov, E., & Lalev, A. (2024). Econometric Analysis of SOFIX Index with GARCH Models. *Journal of Risk and Financial Management*, *17*(8), 346. https://doi.org/10.3390/jrfm17080346

Rahmawati, N., & Lestari, T. E. (2019). Implementasi Model Fungsi Transfer dan Neural Network untuk Meramalkan Harga Penutupan Saham (Close Price). *Jurnal Matematika*, *9*(1), 11–25. https://doi.org/10.24843/JMAT.2019.v09.i01.p107

Sako, K., Mpinda, B. N., & Rodrigues, P. C. (2022). Neural Networks for Financial Time Series Forecasting. *Entropy*, *24*(5), 657. https://doi.org/10.3390/e24050657

Sari, H. R., Wahyuningsih, S., & Siringoringo, M. (2024). Indonesia Gold Price Forecasting Using ARIMA Model (0,1,1) - GARCH (1,0). *EKSPONENSIAL*, *15*(1), 1–10. https://doi.org/10.30872/eksponensial.v15i1.1265

Sekhar, C., & Meghana, P. S. (2020). A Study on Backpropagation in Artificial Neural Networks. *Asia-Pacific Journal of Neural Networks and Its Applications*, *4*(1), 21–28. https://doi.org/10.21742/AJNNIA.2020.4.1.03

Setiawan, B., Ben Abdallah, M., Fekete-Farkas, M., Nathan, R. J., & Zeman, Z. (2021). GARCH (1,1) Models and Analysis of Stock Market Turmoil during COVID-19 Outbreak in an Emerging and Developed Economy. *Journal of Risk and Financial Management*, *14*(12), 576. https://doi.org/10.3390/jrfm14120576

Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples*. Springer International Publishing. https://doi.org/10.1007/978-3-319-52452-8

Susanti, D., Maraya, N. S., Sukono, S., & Saputra, J. (2024). Prediction of Motor Vehicle Insurance Claims Using ARIMA-GARCH Models. *Operations Research: International Conference Series*, *5*(3), 86–92. https://doi.org/10.47194/orics.v5i3.331

Susila, M. R., Jamil, M., & Santoso, B. H. (2023). Akurasi Model Hybrid ARIMA-Artificial Neural Network dengan Model Non Hybrid pada Peramalan Peredaran Uang Elektronik di Indonesia. *Jambura Journal of Mathematics*, *5*(1), 46–58. https://doi.org/10.34312/jjom.v5i1.14889

Svalova, A., Walshaw, D., Lee, C., Demyanov, V., Parker, N. G., Povey, M. J., & Abbott, G. D. (2021). Estimating the asphaltene critical nanoaggregation concentration region using ultrasonic measurements and Bayesian inference. *Scientific Reports*, *11*(1), 6698. https://doi.org/10.1038/s41598-021-85926-8

Yuwanita, N., Adi, S., Mawarni, D., & Wardani, H. E. (2022). Analisis Faktor-Faktor Perilaku Pembayaran Iuran Program Jaminan Kesehatan Nasional Peserta Mandiri di Kecamatan Klojen Kota Malang. *Sport Science and Health*, *4*(12), 1059–1069. https://doi.org/10.17977/um062v4i122022p1059-1069

[This page intentionally left blank.]