

Application of the Hybrid ARIMA-ANN Method to Predict the Stock Price of PT Telekomunikasi Indonesia (TLKM)

Inaya Zalsabilah*, Ansari Saleh Ahmar, & Isma Muthahharah

Department of Statistics, Faculty of Mathematics and Natural Sciences, Makassar State University, Indonesia

Keywords: Prediction, ARIMA, ANN, Hybrid, Stock Price.

Abstract:

Stock prices fluctuate over time, making them difficult to predict accurately. One commonly used method is the Autoregressive Integrated Moving Average (ARIMA), but it has limitations in capturing nonlinear patterns. To overcome this, an Artificial Neural Network (ANN) was used to create a hybrid ARIMA-ANN model. This study aims to determine the best model and compare the accuracy between the ARIMA model and the hybrid model in predicting the stock price of PT Telekomunikasi Indonesia (Persero) Tbk (TLKM). The data used are daily stock closing price data for the period June 2020 to June 2025. The best ARIMA model obtained is ARIMA (2,1,3), then the residuals are modeled using ANN with the (1-3-1) architecture. The results show that the ARIMA model produces a MAPE of 1.3179% and an RMSE of 60.52, while the hybrid ARIMA-ANN model produces a MAPE of 1.2941% and an RMSE of 59.01. Thus, the hybrid model is more effective in improving stock price prediction performance.

1. Introduction

The capital market plays a crucial role in the economy as a means of investment and fundraising for companies. One of the most popular instruments is stocks, which offer high potential returns but also carry significant risks. Stock price movements are dynamic and influenced by various factors such as economic conditions, company performance, and market sentiment, making them difficult to accurately predict (Zhou et al., 2018). Therefore, the ability to analyze and predict stock prices is crucial for investors in decision-making.

In time series analysis, the Autoregressive Integrated Moving Average (ARIMA) model is used due to its ability to model linear patterns. However, this model has limitations in capturing nonlinear patterns that often appear in stock price data. On the other hand, Artificial Neural Networks (ANN) are capable of modeling nonlinear relationships effectively, but are less than optimal in capturing linear patterns (Alsuwaylimi, 2023). These limitations encourage the use of a hybrid approach that combines the advantages of both methods.

The hybrid ARIMA-ANN model combines ARIMA's ability to model linear components with ANN's ability to capture nonlinear patterns in residuals. Several studies have shown that hybrid models can improve accuracy compared to single models (Amrullah, 2023; Alrweili & Fawzy, 2022). Therefore, this approach is relevant for application to stock data with complex characteristics.

PT Telekomunikasi Indonesia (Persero) Tbk (TLKM) is a blue-chip company with high liquidity in the Indonesian capital market, making its stock price movements a significant concern for investors. Based on this, this study aims to apply the hybrid ARIMA-ANN method to predict TLKM's stock price and compare the accuracy of the hybrid model

* Corresponding author.

E-mail address: inayazalsabilah22@gmail.com



with that of a single ARIMA model. The results are expected to contribute to the development of more accurate prediction methods for financial time series data.

2. Literature Review

2.1. Forecasting

Forecasting is a technique for estimating a value in the future by taking into account past and current data (Aswi & Sukarna, 2006).

2.2. ARIMA Method

The ARIMA model is a time series model that is not stationary with respect to the mean and requires a differencing process of d times to become stationary. In general, the ARIMA (p, d, q) model is formulated as follows: (Aswi & Sukarna, 2006):

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t \quad (2.1)$$

with:

p, d, q = autoregressive (AR), differencing order, non-seasonal moving average (MA)
 B = Backshift operator
 $\phi_p(B)$ = non-seasonal AR operator = $(1 - \phi_1 B - \phi_2 B - \dots - \phi_p B^p)$
 $(1 - B)^d$ = non-seasonal differencing order
 $\theta_q(B)$ = non-seasonal MA operator = $(1 - \theta_1 B - \theta_2 B - \dots - \theta_q B^q)$

2.2.1. Model Identification

Since the ARIMA model can only be applied to stationary time series, an essential first step is to test the data for stationarity. If the data is identified as non-stationary, it is necessary to first perform a stationarity process for both the mean and variance. Next, identify the ARIMA model by examining the ACF and PACF plots of the stationary data to determine the initial model (AR, MA, ARMA) using Table 2.1 below (Aswi & Sukarna, 2006).

Table 2. 1 ACF and PACF forms of stationary ARIMA models

Model	ACF	PACF
AR (p)	<i>dies down</i>	<i>cut off after lag p</i>
MA (q)	<i>cut off after lag q</i>	<i>dies down</i>
ARMA (p, q)	<i>dies down</i>	<i>dies down</i>

2.2.2. Parameter Estimation

Once the ARIMA model is identified, the next step is parameter estimation to obtain the coefficient values for each model component. Various methods can be used to estimate these parameters, including the method of moments, the least squares method, and the maximum likelihood method (Aswi & Sukarna, 2006).

2.2.3. Diagnostic Examination

Tests are conducted to verify the model's adequacy. Diagnostic tests can be divided into two parts: parameter significance tests and model fit tests (including white noise and normally distributed residuals) (Aswi & Sukarna, 2006).

2.2.4. Parameter Significance Test

In general, suppose ϕ is a parameter in the ARIMA Box-Jenkins model and $\hat{\phi}$ is the estimated value of the parameter, and $SE(\hat{\phi})$ is the standard error of the estimated value of $\hat{\phi}$ then testing of the model parameters can be done as follows (Aswi & Sukarna, 2006).

1. Hypothesis

$H_0 : \hat{\phi} = 0$ (model parameters are not significant)

$H_1 : \hat{\phi} \neq 0$ (significant model parameters)

2. Statistical Test

$$t = \frac{\hat{\phi}}{SE(\hat{\phi})} \quad (2.2)$$

$$SE(\hat{\phi}) = \sqrt{\frac{1}{n}} \quad (2.3)$$

3. Rejection Area

H_0 rejected if $|t| > t_{\frac{\alpha}{2}, df=n-n_p}$; n_p = number of parameters or by using values *sig (p-value)*, namely reject H_0 if value *(p-value)* $< \alpha(0,05)$.

2.2.5. Model Fit Test

Model suitability tests include white noise assumption tests and normally distributed residual tests. (Aswi & Sukarna, 2006).

2.2.5.1. Residual White Noise Test

A model is white noise, meaning that the residual values of the model meet the assumptions of homogeneous residual variation and are independent (no correlation between residuals) (Nofinda & Wahyuningsih, 2012). The test usually used is the Ljung-Box test. The steps for conducting the Ljung-Box test are as follows.

1. Hypothesis

H_0 : there is no correlation of residuals (white noise)

H_1 : there is a correlation of residuals (not white noise)

2. Statistical Test

$$Q = n(n+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{n-k}, K < n \quad (2.4)$$

3. Rejection Area

Reject H_0 jika $|Q| \geq \chi_{\alpha; (K-p)}^2$ or *p-value* $\geq \alpha$ This means there is a correlation between the residuals (not white noise), if the value *(p-value)* $< \alpha(0,05)$.

2.2.5.2. Test for Normally Distributed Residuals

This assumption test aims to determine whether the data meets the assumption of normality. One method for conducting this test is the Shapiro-Wilk test. The steps for the Shapiro-Wilk test are as follows (Ahadi & Zain, 2023).

1. Hypothesis

H_0 : residuals are normally distributed

H_1 : residuals are not normally distributed

2. Statistical Test

$$W = \frac{(\sum_1^n a_i x_i)^2}{\sum_1^n (x_i - \bar{x})^2} \quad (2.5)$$

3. Rejection Area

The W test statistic yields a reject decision H_0 jika $W_{hitung} > W_{\alpha,n}$ or by using *sig (p-value)*, namely reject H_0 if $p\text{-value} < \alpha$ (0,05).

2.2.6. BDS Test (Brock-Dechert-Scheinkman)

The BDS test is a non-parametric statistical test used to test the hypothesis that a time series is *independent and identically distributed* (IID).

1. Hypothesis

H_0 : The data is Independent and Identically Distributed (IID) or random

H_1 : The data is not IID (there is a non-linear pattern/structure)

2. Statistical Test

$$W_m(\epsilon) = \sqrt{n} \frac{C_m(\epsilon) - C_1(\epsilon)^m}{\sigma_m(\epsilon)} \quad (2.6)$$

3. Rejection Area

If $p\text{-value} < \alpha(0,05)$, then H_0 is rejected. This shows that the data (or model residuals) still have a significant non-linear structure.

2.2.7. Selection of the Best ARIMA Model

Akaike's Information Criterion (AIC) is a criterion for selecting the best model introduced by Akaike in 1973 by considering the number of parameters in the model. The AIC criterion can be formulated as follows (Aswi & Sukarna, 2006).

$$AIC = n \times \ln (SSE / n) + 2f + n \times \ln (2\pi) \quad (2.7)$$

2.2.8. ARIMA Forecasting

The forecasting stage can be carried out if all model parameters are significant and all assumptions are met.

2.3. Artificial Neural Network (ANN)

Artificial Neural Network (ANN), originally developed by Warren McCulloch and Walter Pitts in 1943, ANN is an information processing paradigm inspired by the biological nervous system, specifically the workings of the human brain. The fundamental concept of ANN lies in its information processing structure, which consists of many interconnected processing elements (often called neurons). These elements operate in parallel and simultaneously to solve specific problems (Sutojo et al., 2011). Broadly speaking, the Artificial Neural Network (ANN) architecture consists of three main layers: the input layer, the hidden layer, and the output layer (Alsuwaylimi, 2023).

2.3.1. Activation Function

The activation function is a key component of each neuron in an artificial neural network, determining whether or not it will be activated. A commonly used activation function is the Tan Hyperbolic Function (TanH). The Tan Hyperbolic (TanH) activation function maps the output of each neuron into a range of values between -1 and 1 (Aprilia, 2024).

2.3.2. Backpropagation Algorithm

The backpropagation algorithm is designed to identify optimal functions, whether maximal or minimal, through an iterative process. This algorithm is widely applied in various control applications because its training procedure is based on a simple interconnection principle: if the network output is inaccurate, the connection weights are corrected. The ability of backpropagation to correct weights not only in the output layer but also in the hidden layers makes it an effective tool (Abishek et al., 2012). Training with the backpropagation algorithm generally involves three main stages: feedforward, backpropagation, and weight and bias modification (Meidyta, 2021).

2.4. Hybrid ARIMA-ANN

The ARIMA and ANN models are hybrid models used to address linear and nonlinear problems (Zhang, 2001). Therefore, the hybrid model forecast can be written as follows.

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \quad (2.8)$$

Where $\hat{L}_t$ is a linear component modeled with ARIMA, while $\hat{N}_t$ is a nonlinear component modeled with ANN.

2.5. Model Evaluation

Measuring the accuracy of a forecast period can be done by finding the difference (forecast error) between the forecast data and the actual data. This can be seen in the decreasing error (Vivianti et al., 2020). Measuring the error rate can be tested using the Mean Absolute Percentage Error (MAPE) and Mean Squared Error (MSE). The MAPE value can be calculated using the following formula (Wulandari et al., 2021).

$$MAPE = \left(\frac{1}{n} \sum_{t=1}^n \left| \frac{x_t - \hat{x}_t}{x_t} \right| \right) \times 100\% \quad (2.9)$$

RMSE calculated based on the sum of the differences between forecast data and actual data.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (A_i - F_i)^2}{n}} \quad (2.10)$$

2.6. Shares of PT Telekomunikasi Indonesia (Persero) Tbk

Stocks are an investment instrument that reflects ownership of a company and provides the right to dividends and decision-making through the General Meeting of Shareholders (GMS). In an investment context, stocks have return and

risk characteristics that must be optimally managed (Puspitasari & Pramesti, 2018). PT Telkom Indonesia (Persero) Tbk is one of the largest telecommunications companies in Indonesia with a broad service reach and a significant customer base. The company first went public through an Initial Public Offering (IPO) on November 14, 1995, and since then, its shares have been actively traded on the Indonesia Stock Exchange and several international exchanges. Currently, PT Telkom's share ownership structure is dominated by the Government of the Republic of Indonesia at 52.09%, while the remainder is owned by the public (Oktarina, 2023).

3. Research Methods

The research used is quantitative. The data used in this study is secondary data in the form of daily stock closing prices. The object of the study is the shares of PT. Telekomunikasi Indonesia (Persero) Tbk (TLKM). TLKM stock closing price data was collected daily for the period from June 1 2020, to June 30 2025, with a total of 1,220 observations adjusted to reflect stock exchange trading days. The data was obtained from the financial website Yahoo Finance via the page www.finance.yahoo.com.

The stages carried out in analyzing data in this study are as follows.

1. Descriptive analysis and data visualization through descriptive statistical tables and time series plots.
2. ARIMA Method
 - a. Stationarity testing is performed by performing stationarity tests for the mean (Augmented Dickey-Fuller) and variance (Box-Cox). If the data is not stationary, transformation and differencing are performed.
 - b. Identify nonlinear components in the data (BDS Test I).
 - c. Identify ARIMA models by analyzing ACF and PACF plots of stationary data.
 - d. Estimate ARIMA parameters by constructing and estimating the parameters of the best ARIMA model.
 - e. The estimated ARIMA model is then tested for its feasibility through residual diagnostic tests, including ARIMA model diagnostic tests using white noise and residual normality tests.
 - f. Select the best model.
 - g. Identify ARIMA Residual Nonlinear Patterns (BDS Test II).
3. ANN Method
 - a. Extract residuals by calculating the residual value (e_t) as the difference between the actual value and the fitting (in-sample prediction) ARIMA results as input data for the ANN.
 - b. Normalize residuals using the Min-Max transformation to the range [0, 1] to match the activation function in the ANN.
 - c. Supervised ANN data generation.
 - d. Development of an ANN (MLP) architecture by constructing, training, and optimizing a Multi-Layer Perceptron (MLP) architecture with a hidden layer configuration and tanh activation function.
 - e. Prediction and Denormalization of ANN Output.
4. Formation of a *Hybrid* ARIMA–ANN model.
5. Evaluation of model performance.
6. Interpret the results by drawing conclusions based on the results of the accuracy evaluation and visualization of the comparison between actual data and predicted data.

4. Results and Discussion

4.1. Descriptive Statistical Analysis

Descriptive analysis is used to determine the general description of stock data from the company PT. Telekomunikasi Indonesia (Persero) Tbk (TLKM) from June 1 2020 to June 30 2025. The results of the descriptive analysis of PT Telekomunikasi Indonesia's stock closing price data are presented in Table 4.1.

Table 4. 1 Descriptive Statistics of PT Telkom Indonesia's Closing Stock Price

Characteristics	Value (Rupiah)
Minimum	2290,00
Maximum	4770,00
Average	3532,00
Standard Deviation	596,52

Table 4.1 shows that the lowest closing price of PT Telkom Indonesia shares was 2290.00 and the highest was 4770.00, with an average of 3532.00 and a standard deviation of 596.52. The following is a time series plot of PT Telkom Indonesia's closing price data, shown in Figure 4.1.

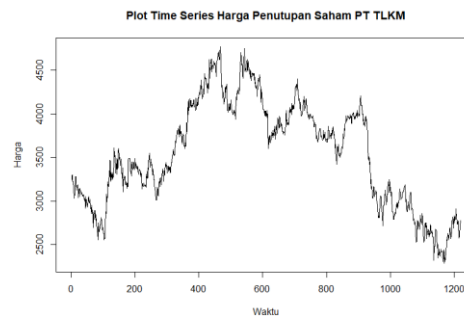


Figure 4. 1 Time Series Plot of Closing Price Data of PT Telkom Indonesia Shares

Based on Figure 4.1, the closing price movement of PT Telekomunikasi Indonesia (Persero) Tbk shares exhibits a fluctuating pattern with a changing trend. At the beginning of the period, the stock price experienced a decline, followed by a significant upward trend until it reached a peak. Afterward, the price fluctuated again with a downward trend before finally moving relatively stable with a slight increase towards the end of the observation period. This pattern indicates that the data is dynamic and unstable.

4.2. Stages of PT Telkom Indonesia Stock Price Prediction Using Autoregressive Integrated Moving Average (ARIMA)

4.2.1. Stationerity

Based on Figure 4.1, the Augmented Dickey-Fuller (ADF) test shows that the data is not stationary in the mean at the initial conditions ($p\text{-value} > 0,05$). After differencing once, the data becomes stationary ($p\text{-value} < 0,05$). Furthermore based on the Box-Cox transformation, the value is obtained $\lambda = 0,8575577$ approaching 1, so the data can be considered stationary in variance. The plot of the differencing data is shown in Figure 4.2.

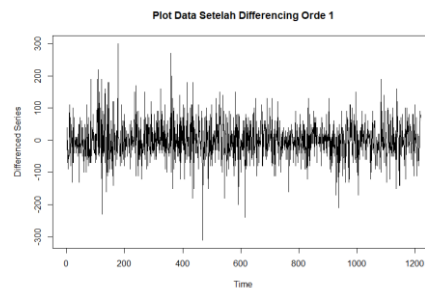


Figure 4. 2 Time Series Plot of Differentiating Results

4.2.2. Identification of Non-Linear Structures in Data (BDS Test I)

The initial BDS test was conducted on data that had undergone one differencing process ($d = 1$).. This test aimed to identify the presence of nonlinear patterns in PT Telkom Indonesia's stock price data before linear modeling was performed. The results of the BDS test on PT Telkom Indonesia's differencing stock price data are presented in Table 4.2.

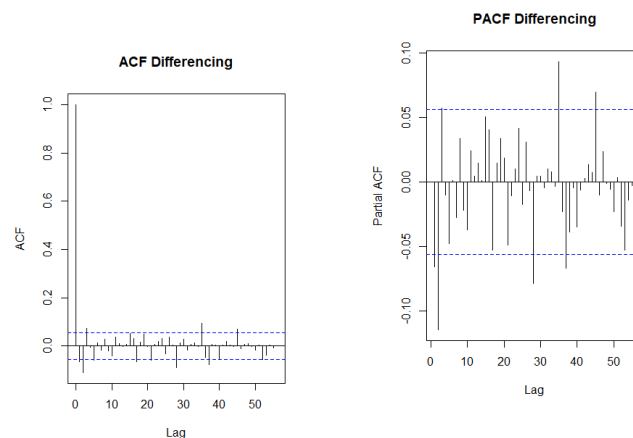
Table 4. 2 BDS Test Results on PT Telkom Indonesia Stock Price Differentiation Data

<i>Embedding Dimension (m)</i>	<i>BDS Statistic (Z)</i>	<i>P-value</i>	<i>Information</i>
2	2,6703	0,0076	Significant
3	3,6071	0,0003	Significant

Based on Table 4.2, the test results for dimensions 2 and 3 show that the statistical values are significantly above the critical value of the standard normal distribution ($Z > 1,96$).

4.2.3. Model Identification

In identifying the ARIMA model, start by identifying the Autocorrelation Function (ACF) plot and the Partial



Autocorrelation Function (PACF) plot presented in Figure 4.3.

Figure 4. 3 ACF and PACF plots

Figure 4.3 displays the ACF and PACF plots of the stationary data used to identify the ARIMA model. The ACF plot shows significant autocorrelation up to the 3rd lag, indicating an MA component up to the 3rd order, while the PACF plot is significant up to the 2nd lag, indicating an AR component up to the 2nd order. Based on these results, several ARIMA models were tested, including ARIMA (1,1,1), (2,1,1), (1,1,2), (2,1,2), (1,1,3), and (2,1,3).

4.2.4. Parameter Estimation

Parameter significance testing is carried out based on the p-value with a significance level of $\alpha=(0.05)$. The results of testing and estimating the parameters of the ARIMA (1,1,1), ARIMA (2,1,1), and ARIMA (2,1,3) models are presented in Table 4.3.

Table 4. 3 ARIMA Model Parameter Estimation

Model	Parameter	Estimation	P-Value	Parameter Significance
ARIMA (1,1,1)	ϕ_1	0,47246	0,019735 *	Significant
	θ_1	-0,55968	0,003344 **	Significant
ARIMA (2,1,1)	ϕ_1	-0,384640	0,01191 *	Significant
	ϕ_2	-0,139281	0,00000115***	Significant
	θ_1	0,315879	0,03904 *	Significant
ARIMA (2,1,3)	ϕ_1	-1,253932	0,0001 ***	Significant
	ϕ_2	-0,654306	0,0001 ***	Significant
	θ_1	1,190735	0,0001 ***	Significant
	θ_2	0,466719	0.0004048 ***	Significant
	θ_3	-0,119717	0.0007105 ***	Significant

Based on Table 4.3, the three ARIMA models tested, namely ARIMA (1,1,1), ARIMA (2,1,1), and ARIMA (2,1,3), meet the parameter significance criteria.

4.2.5. Model Diagnostics

The model suitability test includes a white noise residual test and a normal distribution assumption test. The results of the white noise residual test using the Ljung-Box test for the ARIMA (1,1,1), ARIMA (2,1,1), and ARIMA (2,1,3) models can be seen in Table 4.4.

Table 4. 4 White Noise Assumption Test Results

Model	Ljung-Box	P-Value
ARIMA (1,1,1)	0,22458	0,6356
ARIMA (2,1,1)	0,0041451	0,9487
ARIMA (2,1,3)	0,00025876	0,9872

Based on the results of the Ljung-Box test obtained, it can be seen in Table 4.4 that the p-value of all the models above is greater α (0,05) This means that the residuals from all models meet the white noise assumption. The next step is to test the normal distribution assumption. This test is performed using the Shapiro-Wilk test. The results of the Shapiro-Wilk test for the ARIMA (1,1,1), ARIMA (2,1,1), and ARIMA (2,1,3) models can be seen in Table 4.5.

Table 4. 5 Shapiro Wilk Test Results

Model	Shapiro Wilk	P-Value
ARIMA (1,1,1)	0,98012	0,000000000006644
ARIMA (2,1,1)	0,98017	0,000000000006975
ARIMA (2,1,3)	0,98083	0,00000000001246

Based on the Shapiro-Wilk test in Table 4.5, all models yielded p-values < 0.05 , indicating that the residuals were not normally distributed. This indicates a nonlinear pattern that the ARIMA model could not capture. Therefore, an Artificial Neural Network (ANN) method was used in the hybrid model to model the residuals and improve prediction accuracy.

4.2.6. Selecting the Best Model

The selection of the best model based on the smallest AIC value of all models can be seen in Table 4.6.

Table 4. 6 AIC Value of ARIMA Model

Model	AIC Value
ARIMA (1,1,1)	13486,39
ARIMA (2,1,1)	13475,74
ARIMA (2,1,3)	13474,52

From Table 4.6 it can be seen that the ARIMA (2,1,3) model has the smallest AIC value, namely: 13474.52, so the ARIMA model is used for modeling with the following equation.

$$Y_t = Y_{t-1} - 1,253932(Y_{t-1} - Y_{t-2}) - 0,654306(Y_{t-2} - Y_{t-3}) + \varepsilon_t + 1,190735\varepsilon_{t-1} + 0,466719\varepsilon_{t-2} + 0,119717\varepsilon_{t-3}$$

4.2.7. BDS Test Results on ARIMA Residuals (2,1,3)

The results of the BDS test on the residuals of the best linear model can be seen in Table 4.7.

Table 4. 7 BDS Test Results on ARIMA Residuals (2,1,3)

Embedding Dimension (m)	BDS Statistic (Z)	P-value	Information
2	2,0806	0,0375	Significant
3	2,9594	0,0031	Significant

Based on the Brock–Dechert–Scheinkman (BDS) test in Table 4.7, the p-value is < 0.05 for several embedding dimensions. This indicates that the residuals of the ARIMA (2,1,3) model still contain nonlinear patterns. Therefore, a further approach using Artificial Neural Networks (ANN) in a hybrid model is needed to improve modeling capabilities.

4.3. Artificial Neural Network (ANN) Modeling on ARIMA Residuals

4.3.1. Normalization of PT Telkom Indonesia's Residual Shares

The residuals from the ARIMA model were normalized using the Z-Score method to achieve a mean close to zero and a standard deviation of one. The residual values before and after normalization using the Z-Score method are presented in Table 4.8.

Table 4. 8 Example of ARIMA Residuals Before and After Normalization

Period (t)	ARIMA Residuals (e_t)	Normalized Residuals (e_t^*)
1	3,249998	0,0612
2	39,517555	0,6605
3	12,483474	0,2138
4	-64,477687	-1,0579
5	-5,524731	-0,0838
6	-97,618632	-1,6055
$\vdots$	$\vdots$	$\vdots$
1220	78,0326356315973	1,2969

4.3.2. Supervised Data Generation Using Lag Input

The ANN model requires data in the form of supervised learning, consisting of input and target variables. In this study, one input variable is used, the residual from the previous period, $e(t-1)$, while the residual from period t is used as the network's output target. An example of lag variable formation can be seen in Table 4.9.

Table 4. 9 Example of Normalized ARIMA Residual Data with Lag

Observation to-	$e(t)$	$e(t-1)$
1	0,6605	0,0612
2	0,2138	0,6605
3	-1,0579	0,2138
4	-0,0838	-1,0579
5	-1,6055	-0,0838
$\vdots$	$\vdots$	$\vdots$
1219	1,2969	1,5924

4.3.3. ANN Training and Testing

The Artificial Neural Network (ANN) model used is a feedforward neural network with a backpropagation algorithm. Modeling is carried out using a trial and error approach to determine the best architecture based on the number of neurons in the hidden layer. In this study, several network configurations were tested, namely the (1–2–1) and (1–3–1) architectures. The activation function used in the hidden layer is hyperbolic tangent (tanh), while the output layer uses a linear function. The stepmax parameter is set at 10^7 to ensure the training process reaches convergence. Evaluation of the performance of each ANN architecture is carried out using the MAPE and RMSE error measures. The test results of several ANN architecture configurations can be seen in Table 4.10.

Table 4. 10 Experimental Results of ANN Architecture on Hybrid ARIMA–ANN Model

Number of Inputs	Hidden Layer	Activation Function	Stepmax	MAPE (%)	RMSE
1	2	tanh	1×10^7	1,296561	59,22
1	3	tanh	1×10^7	1,294106	59,01

Based on the model selection criteria using minimum MAPE and RMSE values, the best ANN architecture in this study is the (1–3–1) model, which is then used as a nonlinear component in the formation of a hybrid ARIMA–ANN model. The residual prediction results obtained from the ANN network are first returned to their original scale through a denormalization process.

4.4. Formation of the ARIMA-ANN Hybrid Model

The hybrid model is built by combining the forecasting results of the linear components from the ARIMA model and the forecasting results of the nonlinear components from the ANN, according to the equation:

$$\hat{Y}_t = \widehat{ARIMA}_t + \widehat{ANN}_t$$

So the final model can be written as:

$$\hat{Y}_t = Y_{t-1} - 1,253932\Delta Y_{t-1} - 0,654306\Delta Y_{t-2} + \varepsilon_t + 1,190735\varepsilon_{t-1} + 0,466719\varepsilon_{t-2} + 0,119717\varepsilon_{t-3} + f(e_{t-1})$$

4.5. Comparison of ARIMA and Hybrid ARIMA-ANN Model Performance

To determine the effectiveness of using the hybrid model, a performance comparison was carried out between the ARIMA model and the ARIMA-ANN hybrid model using the MAPE and RMSE measures, which can be seen in Table 4.11.

Table 4. 11 Comparison of ARIMA and Hybrid ARIMA-ANN Model Performance

Model	MAPE (%)	RMSE
ARIMA (2,1,3)	1,318771	60,52
Hybrid ARIMA-ANN	1,294106	59,01

Based on Table 4.11, it can be seen that the hybrid ARIMA-ANN model produces smaller MAPE and RMSE values than the ARIMA model. This decrease in error values indicates that adding the ANN component to the ARIMA residuals can improve forecasting accuracy. To support the quantitative analysis, a visual comparison is performed between the actual data, ARIMA forecasting results, and the best hybrid ARIMA-ANN model forecasting results, which can be seen in Figure 4.6.

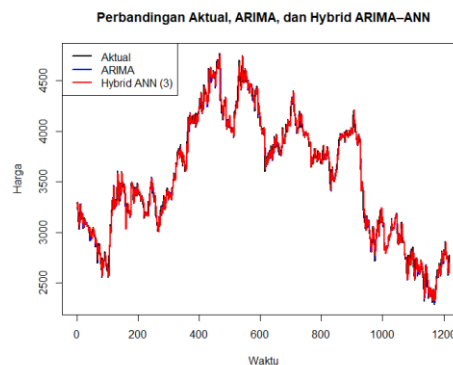


Figure 4. 1 Forecast Plot of Hybrid ARIMA-ANN Model and ARIMA Model

4.6. ARIMA and Hybrid ARIMA-ANN Forecasting

Forecasting using the ARIMA (2,1,3) model was performed for the next five periods and produced relatively stable predicted values with a slight downward trend. In the hybrid ARIMA-ANN model, forecasting is performed in two stages: modeling the linear component using ARIMA and the nonlinear component using ANN based on the ARIMA residuals. The final results are obtained by combining the two components and are presented in Table 4.12.

Table 4. 12 Five-Day Stock Price Forecast Results

Period	ARIMA (2,1,3) Forecasting	Hybrid ARIMA-ANN Forecasting
$t + 1$	2768,25	2766,45
$t + 2$	2762,12	2760,87
$t + 3$	2768,15	2766,89
$t + 4$	2764,60	2763,34
$t + 5$	2765,11	2763,85

Based on Table 4.12, the forecasting results show that the Hybrid ARIMA-ANN model produces slightly different prediction values than the ARIMA (2,1,3) model. This relatively small difference indicates that ANN acts as a correction to the ARIMA results. Furthermore, based on the MAPE and RMSE values, the hybrid model provides better accuracy, resulting in more optimal prediction results.

5. Conclusion

Based on the research results, the Hybrid ARIMA-ANN model is proven to have better performance than the single ARIMA model in predicting the closing price of PT Telekomunikasi Indonesia (Persero) Tbk shares. The best ARIMA model obtained is ARIMA (2,1,3) with a MAPE value of 1.317903% and an RMSE of 60.52. After being combined with ANN, the hybrid model with the (1-3-1) architecture produces an increase in accuracy with a MAPE value of 1.294106% and an RMSE of 59.01. These results indicate that the addition of ANN is able to capture nonlinear patterns that are not accommodated by ARIMA. In addition, the results of the prediction of the next five periods show that the hybrid model produces more stable values than the single ARIMA. Thus, the Hybrid ARIMA-ANN method is effective for increasing the accuracy of stock price predictions.

References

- Abishek, K., Khairwa, A., Pratap, T., & Prakash, S. (2012). A stock market prediction model using Artificial Neural Network. *Proceedings of 3rd International Conference on Computing, Communication and Networking Technologies (ICCCNT 2012)*, 1-5.
- Ahadi, G. D., & Zain, N. N. L. E. (2023). Pemeriksaan uji kenormalan dengan Kolmogorov-Smirnov, Anderson-Darling dan Shapiro-Wilk. *Eigen Mathematics Journal*, 6(1), 11-19. <https://doi.org/10.29303/emj.v6i1.131>
- Alrweili, H., & Fawzy, H. (2022). Forecasting crude oil prices using an ARIMA-ANN hybrid model. *Journal of Statistics Applications and Probability*, 11(3), 845-855. <https://doi.org/10.18576/jsap/110308>
- Alsawaylimi, A. A. (2023). Comparison of ARIMA, ANN and hybrid ARIMA-ANN models for time series forecasting. *Information Sciences Letters*, 12(2), 1003-1016. <https://doi.org/10.18576/isl/120238>
- Amrullah, A., Soesanto, O., & Maisarah, M. (2022). Penerapan metode hybrid ARIMA-ANN untuk memprediksi harga saham PT. BNI (Persero) Tbk. *RAGAM: Journal of Statistics & Its Application*, 1(1), 52. <https://doi.org/10.20527/ragam.v1i1.7328>
- Aprilia, A. (2024). *Penerapan metode hybrid Autoregressive Integrated Moving Average (ARIMA)-Artificial Neural Network (ANN) dalam meramalkan tingkat inflasi di Indonesia* [Skripsi, Universitas Negeri Makassar].
- Aswi, & Sukarna. (2006). *Analisis deret waktu teori & aplikasi*. Andira Publisher.
- Meidyta, R. (2021). *Prakiraan harga saham menggunakan Artificial Neural Network dengan algoritma backpropagation (Studi kasus: Harga penutupan saham bulanan PT. Indofood)* [Skripsi, UIN Syarif Hidayatullah]. Repository UIN Jakarta. <https://repository.uinjkt.ac.id/dspace/handle/123456789/56308>
- Nofinda, L., & Wahyuningsih, N. (2012). Peramalan kunjungan wisata dengan pendekatan model SARIMA. *Jurnal*

Sains Dan Seni ITS, 1(1), 29–33.

- Oktarina, A. T. (2023). *Pengaruh kinerja keuangan terhadap harga saham pada perusahaan telekomunikasi yang terdaftar di Bursa Efek Indonesia* [Skripsi, Universitas Negeri Makassar].
- Puspitasari, R., & Pramesti, D. (2018). Analisis resiko dan tingkat pengembalian saham terhadap portofolio optimal saham (Studi kasus pada 8 saham dari LQ-45). *Journal Ilmiah Rangga Gading*, 11(1), 17–21.
- Sutojo, T., Mulyanto, E., & Suhartono, V. (2011). *Kecerdasan Buatan*. Penerbit Andi.
- Vivianti, Aidid, M. K., & Nusrang, M. (2020). Implementasi metode fuzzy time series untuk peramalan jumlah pengunjung di Benteng Fort Rotterdam. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 2(1), 12. <https://doi.org/10.35580/variantsium12895>
- Wulandari, S. S., Sufri, & Yurinanda, S. (2021). Penerapan metode ARIMA dalam memprediksi fluktuasi harga saham PT Bank Central Asia Tbk. *BUANA Matematika: Jurnal Ilmiah Matematika Dan Pendidikan Matematika*, 11(1), 53–68.
- Xu, Y., Ke, Z., Xie, C., & Zhou, W. (2018). Dynamic Evolution Analysis of Stock Price Fluctuation and Its Control. *Complexity*, 2018. <https://doi.org/10.1155/2018/5728090>
- Zhang, P. G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. [https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)