
Academia Open



By Universitas Muhammadiyah Sidoarjo

Table Of Contents

Journal Cover	1
Author[s] Statement	3
Editorial Team	4
Article information	5
Check this article update (crossmark)	5
Check this article impact	5
Cite this article.....	5
Title page	6
Article Title	6
Author information	6
Abstract	6
Article content	7

Originality Statement

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

Conflict of Interest Statement

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright Statement

Copyright © Author(s). This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Academia Open

Vol. 11 No. 1 (2026): June
DOI: 10.21070/acopen.11.2026.13521

EDITORIAL TEAM

Editor in Chief

Mochammad Tanzil Multazam, Universitas Muhammadiyah Sidoarjo, Indonesia

Managing Editor

Bobur Sobirov, Samarkand Institute of Economics and Service, Uzbekistan

Editors

Fika Megawati, Universitas Muhammadiyah Sidoarjo, Indonesia

Mahardika Darmawan Kusuma Wardana, Universitas Muhammadiyah Sidoarjo, Indonesia

Wiwit Wahyu Wijayanti, Universitas Muhammadiyah Sidoarjo, Indonesia

Farkhod Abdurakhmonov, Silk Road International Tourism University, Uzbekistan

Dr. Hindarto, Universitas Muhammadiyah Sidoarjo, Indonesia

Evi Rinata, Universitas Muhammadiyah Sidoarjo, Indonesia

M Faisal Amir, Universitas Muhammadiyah Sidoarjo, Indonesia

Dr. Hana Catur Wahyuni, Universitas Muhammadiyah Sidoarjo, Indonesia

Complete list of editorial team ([link](#))

Complete list of indexing services for this journal ([link](#))

How to submit to this journal ([link](#))

Academia Open

Vol. 11 No. 1 (2026): June
DOI: 10.21070/acopen.11.2026.13521

Article information

Check this article update (crossmark)



Check this article impact (*)



Save this article to Mendeley



(*) Time for indexing process is various, depends on indexing database platform

Hybrid Transformer-LSTM for Stock Price Prediction with Monte Carlo Testing of Loss Levels

David Andris Rizky Saputra, saputradavid573@gmail.com (*)
Informatics Engineering, Universitas PGRI Ronggolawe, Indonesia

Asfan Muqtadir, asfanme@gmail.com
Informatics Engineering, Universitas PGRI Ronggolawe, Indonesia

Andik Adi Suryanto, andikadisuryanto@gmail.com
Informatics Engineering, Universitas PGRI Ronggolawe, Indonesia

(*) Corresponding author

Abstract

General Background: Stock price prediction is a complex problem due to the non-linear, stochastic, and volatile characteristics of financial markets. **Specific Background:** Advanced deep learning approaches such as Long Short-Term Memory (LSTM) and Transformer architectures have been applied to capture sequential patterns and global dependencies in time-series financial data. **Knowledge Gap:** However, existing approaches often lack integration between accurate forecasting and quantitative risk measurement within a unified framework. **Aims:** This study proposes a Hybrid Transformer-LSTM model integrated with Monte Carlo simulation to provide both precise stock price prediction and risk evaluation. **Results:** Using historical daily stock price data of BMRI from March 2013 to March 2025 and incorporating technical indicators such as RSI and moving averages, the model achieved a Mean Absolute Percentage Error of 4.13% and a Mean Absolute Error of 246.35 Rupiah. Monte Carlo-based Value at Risk at a 99% confidence level estimated a potential maximum loss of 5.35%. **Novelty:** The study combines sequential learning, attention mechanisms, and probabilistic simulation in a single framework linking prediction accuracy with risk quantification. **Implications:** The proposed approach provides a comprehensive analytical basis for supporting investment decision-making through reliable forecasting and measurable downside risk estimation.

Highlights :

- Combined deep learning architecture produces low forecasting error on long-term historical data
- Probabilistic simulation quantifies maximum potential loss under high confidence level
- Integrated framework links predictive modeling with measurable investment risk

Keywords: Hybrid Transformer LSTM, Stock Price Prediction, Monte Carlo Value at Risk

Published date: 2026-04-22

Introduction

Investment in capital markets has become a crucial driver of modern economic growth, offering significant profit potential while simultaneously supporting capital formation and financial system stability. In recent years, there has been a notable increase in retail and novice investors, particularly among younger generations who increasingly recognize the importance of financial literacy for achieving long-term economic security. This growing participation reflects broader digital access to financial markets and heightened awareness of investment as a strategic tool for wealth accumulation. However, the expansion of investor participation also amplifies exposure to financial risk, especially in equity markets where potential returns are inherently proportional to potential losses.

Stock market investments are widely acknowledged as high-risk financial instruments due to the stochastic, non-linear, and highly volatile nature of stock price movements. Price dynamics are influenced by a complex interaction of internal factors, such as corporate fundamentals and managerial performance, as well as external factors including regulatory changes, macroeconomic conditions, geopolitical events, and investor sentiment. These intertwined influences result in chaotic and multifractal behaviors within financial time series, making stock price forecasting a computationally challenging yet essential task for informed investment decision-making [1], [2].

The inherent complexity and non-linearity of financial markets significantly limit the effectiveness of traditional forecasting approaches. Conventional techniques such as fundamental and technical analysis, along with classical statistical models, have long been employed to analyze market movements. Models like Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) rely on assumptions of stationarity and normality that are frequently violated in real-world financial data. Empirical evidence shows that stock price series often exhibit non-stationarity, volatility clustering, fat-tailed distributions, and asymmetric behaviors, causing traditional models to underperform, particularly during periods of market turbulence [3], [4], [5].

To mitigate these limitations, hybrid statistical models such as ARIMA-GARCH have been proposed to jointly model linear trends and conditional variance. While these hybrid approaches offer improved volatility estimation compared to standalone models, they remain constrained by sensitivity to parameter specification, limited adaptability to rapidly changing market conditions, and difficulties in capturing complex non-linear dependencies. Consequently, their predictive accuracy and robustness remain insufficient for highly dynamic and chaotic financial environments, highlighting the need for more flexible and expressive modeling paradigms [6], [7], [8].

In response to these challenges, deep learning has emerged as a dominant framework for financial time-series forecasting. Architectures such as Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Convolutional Neural Networks (CNNs) have demonstrated superior capability in modeling non-linear relationships and extracting hierarchical features from historical data. LSTM networks, in particular, are well-suited for sequential data due to their ability to capture long-term dependencies. Recent studies further demonstrate that hybrid and multi-architecture deep learning models outperform both traditional statistical methods and single-model deep learning approaches in highly volatile financial contexts [9], [10], [11].

Despite these advancements, accurate modeling of long-term dependencies and temporal dynamics remains a persistent challenge. Financial time series often exhibit long-range dependencies across multiple time horizons, requiring models that can simultaneously capture global contextual relationships and localized temporal patterns. Transformer architectures, with their self-attention mechanisms, have shown remarkable success in learning long-term dependencies by modeling relationships across entire sequences. However, Transformers alone may struggle with fine-grained temporal dynamics in noisy financial data, motivating the development of hybrid architectures that integrate attention-based models with recurrent networks such as LSTM to leverage complementary strengths [12], [13], [14].

Beyond predictive accuracy, a critical limitation of existing stock price prediction studies lies in their predominant focus on error minimization while neglecting explicit risk quantification. For investors, understanding potential downside risk is as important as forecasting expected returns. The absence of integrated risk analysis can result in overly optimistic or speculative investment decisions. Addressing this research gap, this study proposes a Hybrid Transformer-LSTM model integrated with Monte Carlo simulation to estimate Value at Risk (VaR) probabilistically. The model is evaluated using historical stock data of PT Bank Mandiri (Persero) Tbk (BMRI), a blue-chip stock characterized by high market capitalization and liquidity. This integrated framework aims to provide a more comprehensive investment decision-support system by jointly delivering accurate price predictions and measurable risk assessments.

Figure 1 shows the Hybrid Transformer-LSTM model research methodology for predicting stock Log Returns, which is then validated with VaR Monte Carlo risk analysis.

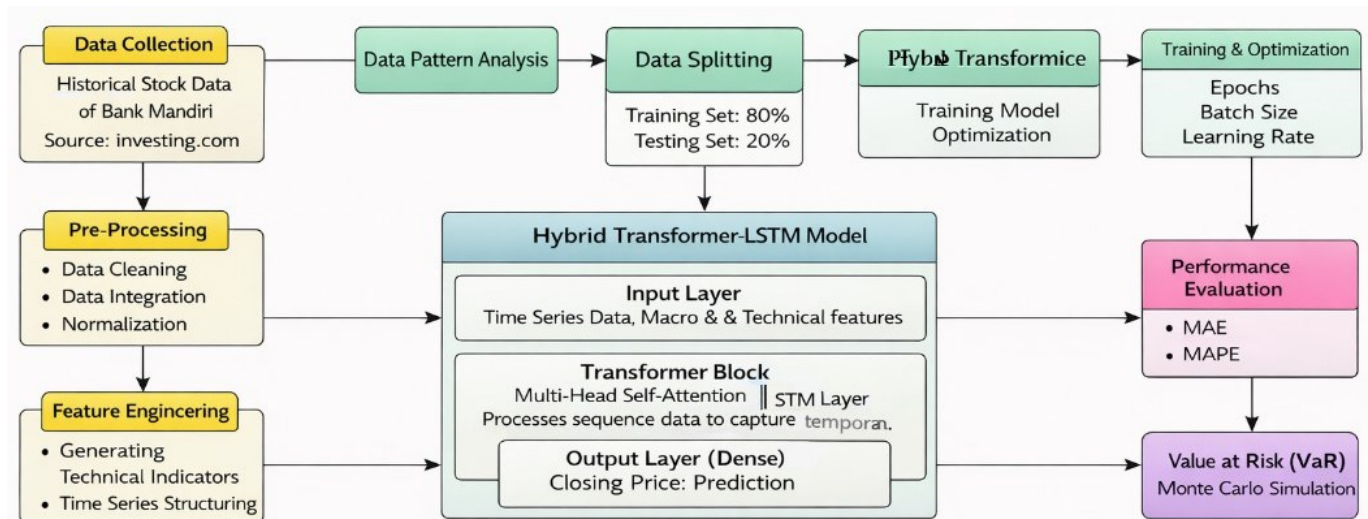


Figure 1. Research Method

In this approach, LSTM is used to extract temporal trends [15], which are then refined by the Multi-Head Attention mechanism in Transformer to capture data anomalies [16]. These features are combined through Global Average Pooling and Dense layers to produce an accurate Log Return prediction model. And the final stage is performance evaluation using the MAE and MAPE indicator metrics, which is followed by Value at Risk (VaR) risk estimation through Monte Carlo Simulation to map the maximum potential loss realistically [17].

1. Data Collection

The data used in this research is daily historical data on shares of PT Bank Mandiri Persero Tbk (BMRI) obtained from the web investing.com in the time period 01/03/2013 to 12/03/2025. The historical data consists of 3131 rows and consists of the variables "Date, Price, Open, High, Low, Vol."

Table 1. Bank Mandiri Historical Dataset

Date	Price	Open	High	Low	Vol.
01/03/2013	8,200	8,200	8,300	8,100	24.75M
01/04/2013	8,250	8,300	8,300	8,200	16.63M
01/07/2013	8,200	8,200	8,300	8,200	24.11M
01/08/2013	8,350	8,200	8,350	8,200	27.48M
01/09/2013	8,250	8,300	8,350	8,200	32.00M
...					
12/03/2025	4,850	4,850	4,870	4,830	30.13M
12/02/2025	4,850	4,920	4,940	4,820	104.01M
12/01/2025	4,860	4,850	4,910	4,820	58.60M
11/28/2025	4,830	4,800	4,880	4,800	128.78M
11/27/2025	4,860	5,025	5,025	4,850	170.21M

2. Data Preprocessing and Feature Engineering

Data Stage Data preprocessing is a fundamental step to transform raw data into a numerical format that can be studied optimally by Deep Learning architecture [16]. The quality of data representation at this stage plays a crucial role in accelerating model convergence and preventing bias in prediction results. The initial steps start with Feature Selection and Data Cleaning. This research utilizes the main attributes of daily trading activity, namely Open, High, Low, Close, Volume as multivariate input variables [15]. Data integration is ensured by checking for missing values to ensure the validity of the model input. Next, a normalization process was carried out using the Min-Max Scaler technique. This step is very important because the Gradient Descent based algorithm is very sensitive to data scales that have large variance, so equalization of the scale is needed so that the training process runs stably [18]. Normalization transforms all feature values into an interval range of [0,1] using the following equation:

$$X' = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (1)$$

Where x is the original value and x' is the normalized value.

After the data is normalized, the next stage is the formation of a sequential data structure using the Sliding Window method. This process changes the data dimensions from a 2-dimensional table format (rows, columns) to a 3-dimensional tensor format (Samples, Time Steps, Features) so that it can be processed by the LSTM layer [19]. This study sets the window size at 60 days. Finally, the dataset is divided into a train set and a test set with a ratio of 80:20 to ensure that model evaluation is carried out objectively on data that has never been seen before [20].

3. Data Pattern Analysis

Prior to the hybrid modeling process, historical stock data is analyzed to understand its volatility characteristics. This identification, based on technical analysis theory, serves as the basis for the LSTM layer to capture long-term trends and for the Transformer to assign attention weights to the most relevant technical signals to generate accurate predictions [21].

4. Model Architecture

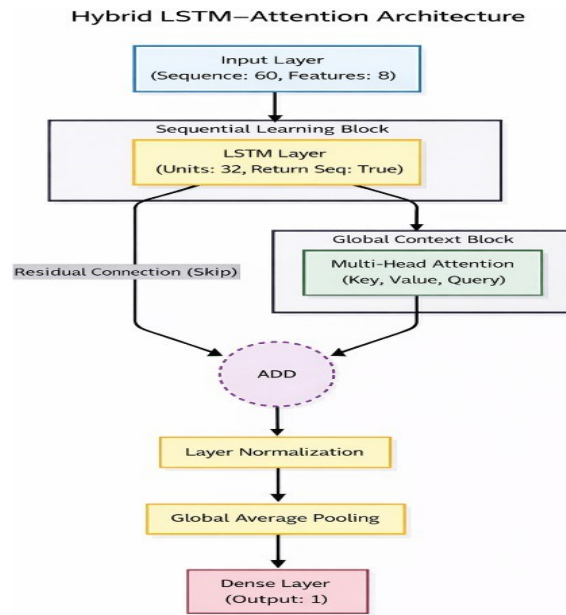


Figure 2. Architecture Hybrid LSTM-Transformer

The modeling stage is carried out by building a Hybrid Deep Learning architecture that integrates the Long Short-Term Memory (LSTM) mechanism sequentially with the Transformer block, as in Figure 2. This Hybrid approach is designed to acquire two characteristics of long-term temporal dependencies from time series data [15] while the transformer is tasked with identifying global relationships between features through an attention mechanism that is capable of capturing long-range correlations [16]. The model architecture is composed with dimensional tensor input (Samples, Time Steps, Features) the main computational process occurs in two core components as follows: information.

a. Long Short-Term Memory (LSTM)

This layer acts as an initial encoder that processes the data sequence using a gating mechanism to regulate the flow of relevant information [18]. Each LSTM cell has three main gates Input Gate (i_t), Forget Gate (f_t), and Output Gate (o_t). This mechanism allows the model to discard irrelevant information (such as market noise) and retain important trends to overcome the vanishing gradient problem. The update process of the hidden state (h_t) and cell state (c_t) at time step t is calculated based on the following equation:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (3)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (4)$$

$$h_t = o_t \odot \tanh(C_t) \quad (5)$$

Where σ is the sigmoid activation function, W is the weight matrix, b is the bias, and \odot denotes the element-wise product. The output of this layer is set with `return_sequences=True` to pass the temporal features to the transformer block [19].

b. Transformer (Multi-Head Attention)

The temporal features of the LSTM are then processed using the Multi-Head Self-Attention mechanism. This process allows the model to assign different weights (attention scores) to each time step within the input window, allowing the model to

focus on the most relevant market signals. The attention score is calculated using the Scaled Dot-Product Attention function, defined in the following equation:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (6)$$

Where Q (Query), K (Key), and V (Value) are the projection matrices of the input, and is a scaling factor to maintain gradient stability. The output of this block is then compacted through Global Average [22]. Pooling before entering the final training and evaluation stage

5. Performance Evaluation

After the training phase is completed, performance measurements are conducted to validate the performance of the Hybrid model in predicting new data. Evaluation is carried out by comparing the predicted value () with the actual value () in the testing set. To comprehensively measure the level of accuracy and error magnitude, this study uses two evaluation metrics, Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

a. Mean Absolute Error (MAE)

MAE is used to measure the average absolute difference between predicted and actual values. This metric describes how much the model's predictions differ. The smaller the MAE value, the more accurate the model's performance in stock prediction. MAE is defined by the equation

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

Where n is the number of data, the predicted result value with actual value [23].

b. Mean Absolute Percentage Error (MAPE)

In addition to nominal errors, evaluation is also carried out using Mean Absolute Percentage Error (MAPE) to measure relative accuracy in percentage form. MAPE is very effective in knowing how much the prediction deviates from the actual value, thus facilitating the interpretation of model performance regardless of the stock price scale [23]. The mathematical formulation for calculating MAPE is shown in the following equation:

(8)

The integration of these two metrics ensures the validity of the model both on a nominal and percentage scale.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (8)$$

c. Value at Risk (VaR) Monte Carlo Simulation- Based

The evaluation is then complemented by investment risk estimation using the Value at Risk (VaR) method based on Monte Carlo simulation. This simulation was run for 20,000 iterations to project the probability of future extreme losses based on the trend (drift) characteristics studied by the Hybrid model. The risk estimation formula is defined as follows:

$$VaR_a = P_0 - \text{Percentile}(S_t, 1 - a) \quad (9)$$

Where P_0 is the initial stock price, S_t is the distribution of simulated stock prices at the end of the period (T), a is the confidence level value [24]. This evaluation provides a statistically valid loss tolerance limit for investors.

Results and Discussion

A. Results

1. Data Analysis and Pre-processing

The initial discussion focused on analyzing the characteristics of the historical stock data used and the results of the feature engineering. The historical data used in this study is the stock price data of PT Bank Mandiri (Persero) Tbk (BMRI), obtained from investing.com. The data covers the daily trading period from March 1, 2013 to March 12, 2025, with a total of 3,131 rows of observations.

The main variables used include Date, Price, Open, High, Low, and Vol. Given the highly volatile nature of stock price movements and the influence of temporal trends, this study integrates additional features into the feature engineering process. Technical indicators such as the Relative Strength Index (RSI) and Moving Average are used to enrich the information learned by the model.

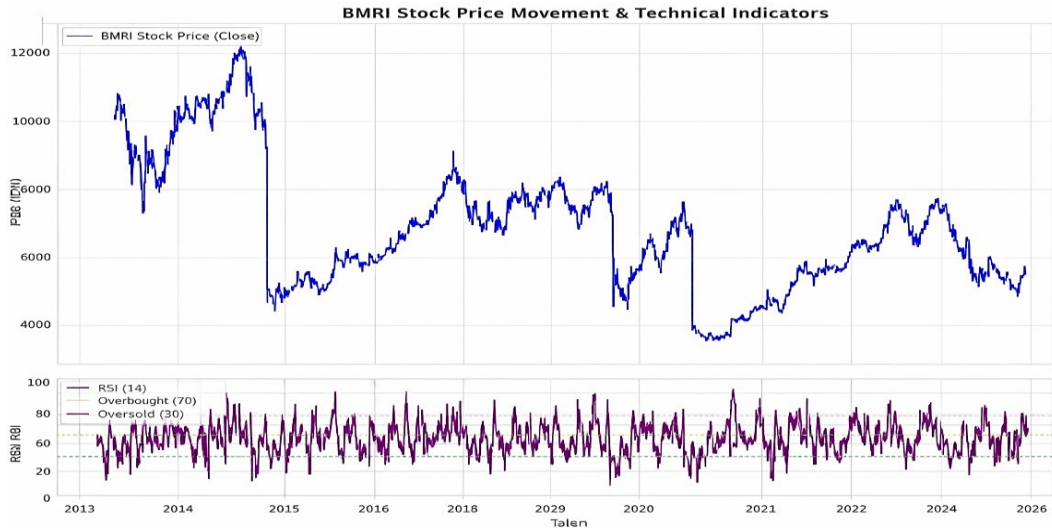


Figure 3. Visualization of BMRI stock price and RSI indicator

Based on Figure 3, it can be identified that the stock price data (top graph) is continuous and a time series with a dynamic trend. The RSI indicator (bottom graph) shows momentum alignment with price fluctuations, with areas above 70 indicating overbought conditions and areas below 30 indicating oversold conditions. This ensures that the technical features used are relevant to the prediction targets. To better understand the data structure, a seasonal decomposition was performed to separate the trend, seasonal, and residual (noise) components, as shown in Figure 4.

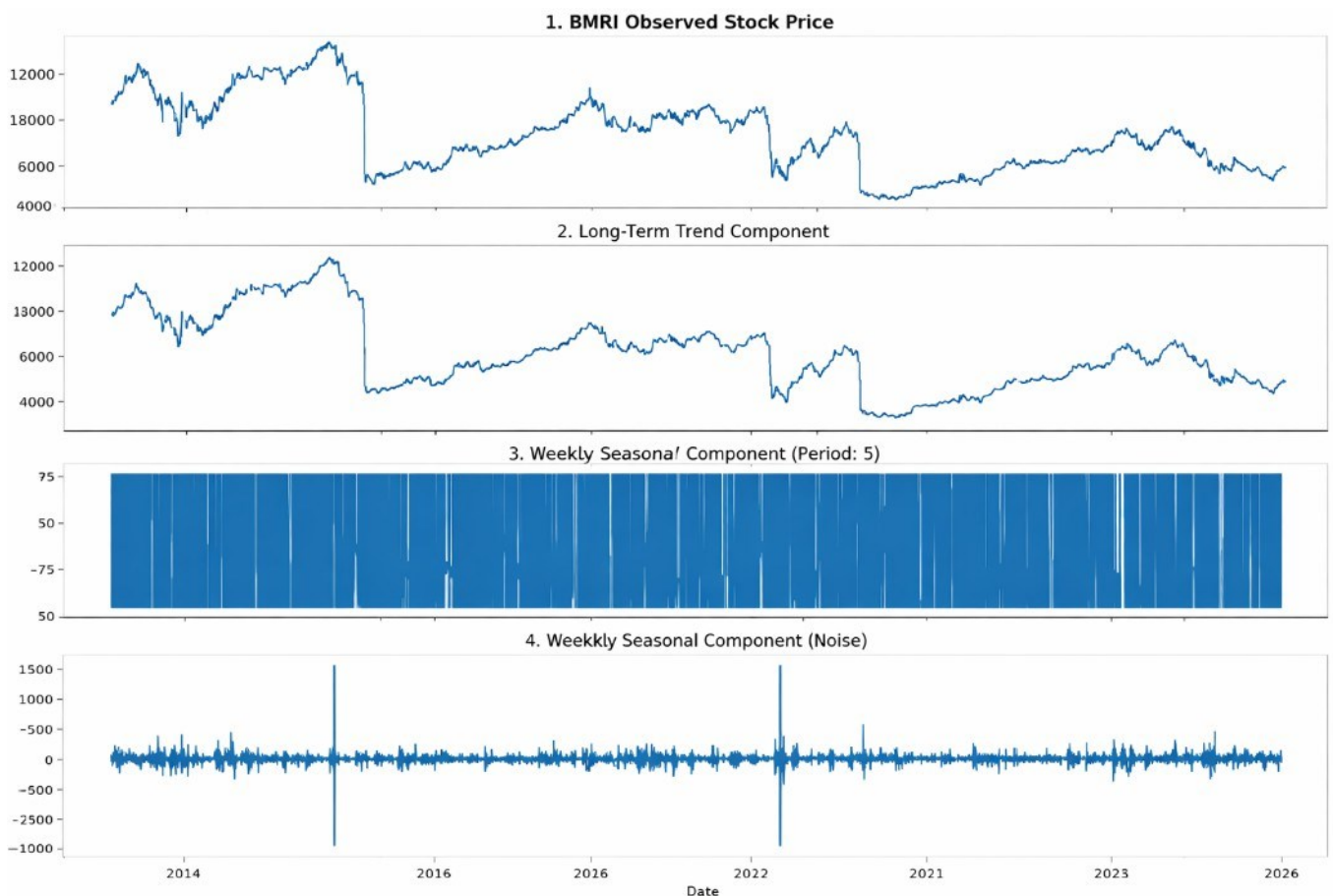


Figure 4. Decomposition of Trend, Seasonal, and Residual Components

Figure 4 shows that the BMRI stock price data contains a fluctuating long-term trend component and a periodically recurring seasonal component. The high fluctuation in the residual component indicates significant noise, making the use of a nonlinear model such as the Hybrid Transformer-LSTM more appropriate than a conventional linear model. Next, a correlation analysis was performed between variables to examine the relationship between input features and the prediction

targets, as shown in Figure 5.

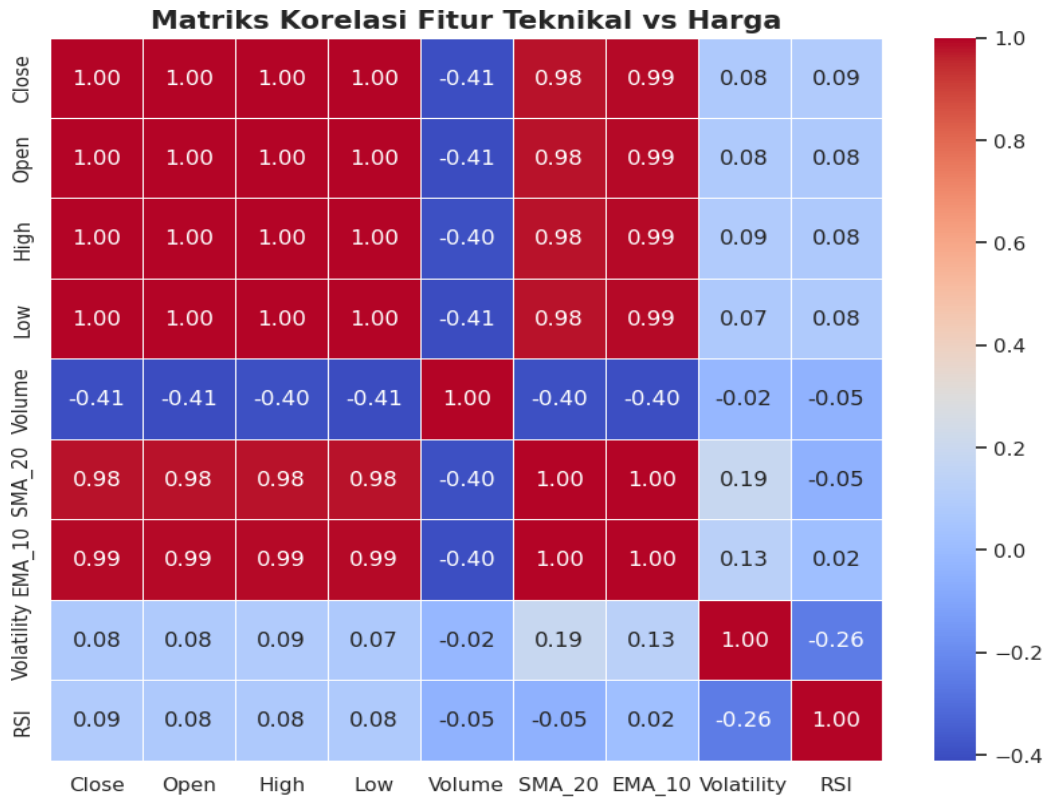


Figure 5. Correlation Matrix Between Features

Figure 5 shows that trend-based features such as the SMA and EMA have a very strong positive correlation (>0.98) with the closing price, indicating multicollinearity between price variables. Meanwhile, the RSI feature has a low correlation (0.09) with price. This value does not indicate the feature is useless, but rather indicates that the RSI carries independent information not represented by price data, thus enriching the model.

Next, the data is transformed using a scaling technique using a Min-Max Scaler to adjust the values of each feature to a range of 0 to 1. The goal is to reduce the impact of scale differences between features (for example, between a price in the thousands of rupiah and an RSI value in the tens of thousands). The final process at this stage is data splitting with an 80:20 ratio, with 80% as training data and 20% as test data.

Table 2. Data Splitting Output

X_train	(2427, 60, 8)
Y_train	2185
X_test	(562, 60, 8)
X_test	562

2. Modeling dan Evaluasi

The modeling phase is performed using the Hybrid Transformer-LSTM architecture.

Table 3. Hybrid Transformer-LSTM Model Architecture

Layer (type)	Output Shape	Param #	Connected to
input_layer (InputLayer)	(None, 60, 8)	0	-
lstm (LSTM)	(None, 60, 32)	5,248	input_layer[0][0]
multi_head_attenti... (MultiHeadAttentio...)	(None, 60, 32)	1,080	lstm[o][o], lstm[o][o]
add (Add)	(None, 60, 32)	0	lstm[o][o], multi_head_atten...

layer_normalization (LayerNormalizatio...	(None, 60, 32)	64	add[o][o]
global_average_poo... (GlobalAveragePool...	(None, 32)	0	layer_normalizat...
dense (Dense)	(None, 1)	33	global_average_p...

Based on Table 3, the model is designed with a complex sequential structure. The input data (shape: 60, 8) is first processed by an LSTM layer (32 units) to capture long-term dependencies in the time series data. The LSTM output is then passed to the Multi-Head Attention layer to focus weights on the most relevant features. The uniqueness of this architecture lies in the ADD layer, which functions as a residual connection (connecting the LSTM output directly with the Attention output). This mechanism aims to maintain the flow of information so that it doesn't get lost as the network deepens. Next, the data is normalized using Layer Normalization before finally entering the Dense layer to produce a single price prediction output. After the training phase is complete, model performance is measured using the MAE and MAPE evaluation metrics. The evaluation results on the test data are presented in:

Table 4. MAE and MAPE Evaluation Results

Mean Absolute Percentage Error	Mean Absolute Error
4.13%	246.35

Based on Table 4, the model produced a MAPE value of 4.13%. Referring to Lewis's (1982) criteria, a value below 10% is categorized as Highly Accurate Forecasting. This indicates the model's predictive ability is very accurate. Furthermore, the MAE value yielded 246.35. This value represents the average absolute deviation in rupiah, indicating that the model's predictions were only 246.35 Rupiah off the actual price. Considering that the average BMRI share price during the study period was around 5,500 Rupiah, this nominal error is considered very small and tolerable.

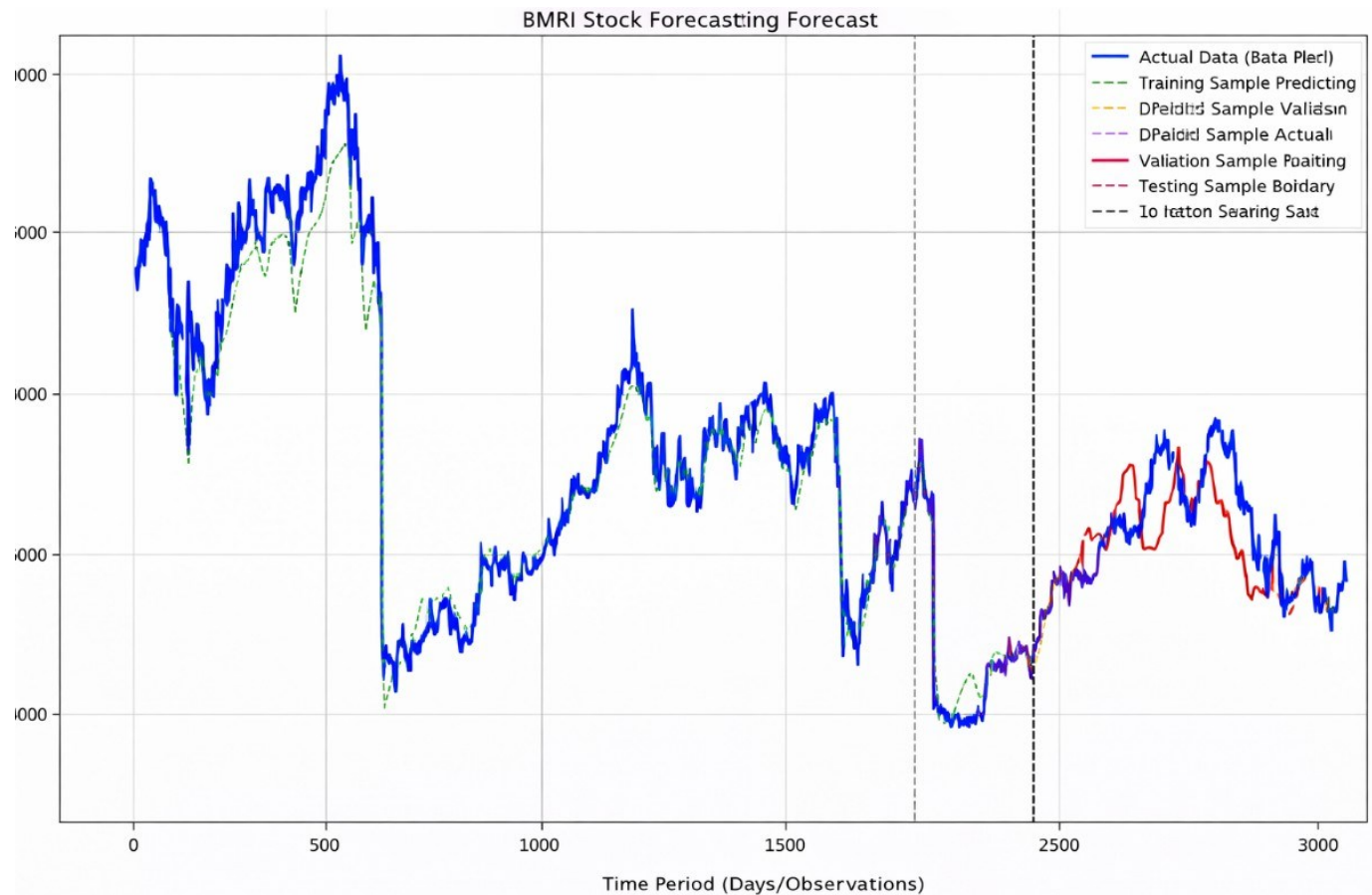


Figure 6. Performance Evaluation of Hybrid Transformer-LSTM

The validation results are confirmed through the visualization of the results presented in Figure 6. The graph shows that the prediction line is able to follow the actual price fluctuation pattern of BMRI shares very closely, which confirms that the model successfully captures the movement trend without experiencing significant overfitting.

3. Value at Risk (VaR) Monte Carlo

After the accuracy evaluation was completed, the research continued with an investment risk analysis using the Value at Risk (VaR) method based on Monte Carlo simulation. The simulation was run for 20,000 iterations to project the distribution of possible future prices.

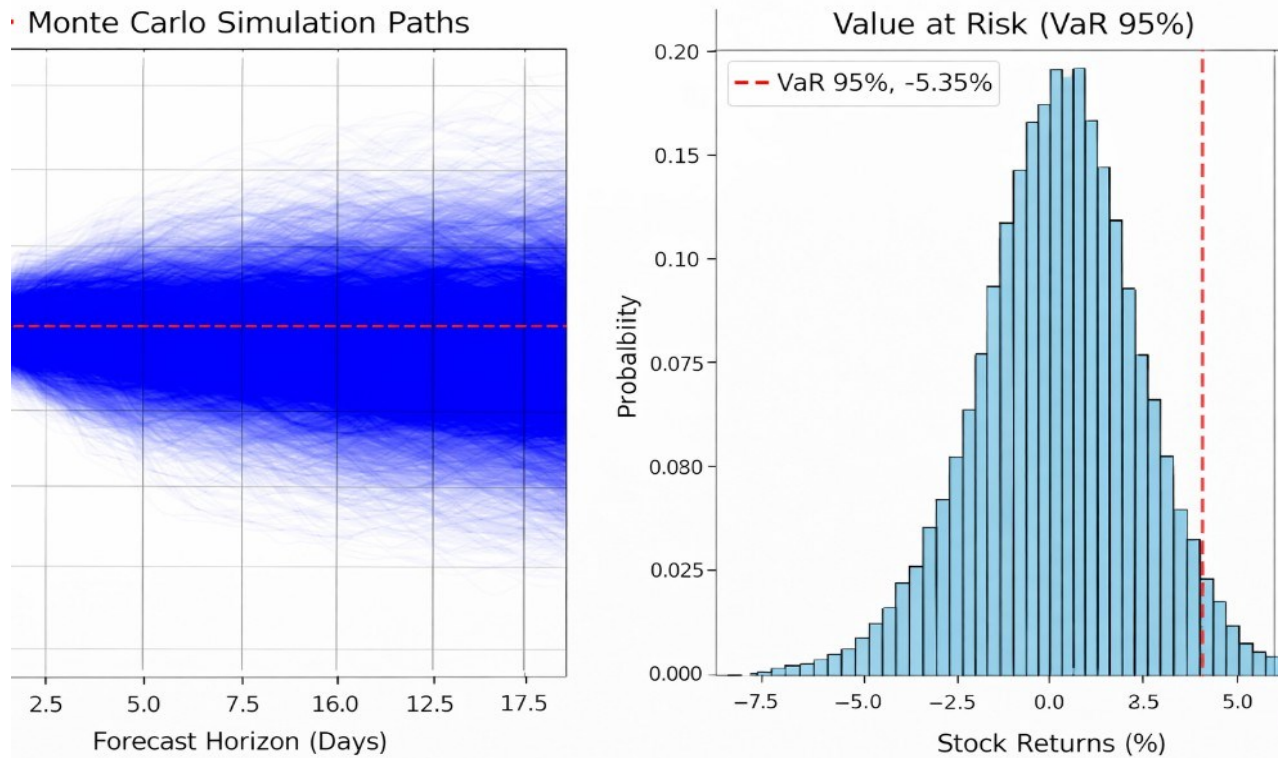


Figure 7. Value at Risk (VaR) Monte Carlo

The visualization results of the Monte Carlo simulation are presented in Figure 7. The left panel visualizes thousands of stock price movement scenario trajectories (transparent blue lines) for the next 20 days, while the right panel displays a histogram of the loss frequency distribution. To provide a more detailed picture of the resulting risk values, a quantitative summary of the simulation results is presented in table 5.

Table 5. Value at Risk (VaR)

Value at Risk (VaR) at 99% confidence level
5.35%

Referring to Table 5 and the distribution visualization above, the Value at Risk (VaR) calculation at a 99% confidence level yields a risk estimate of 5.35%, visually indicated by the red dotted line. This figure indicates that within the projection period, the probability of investors experiencing a loss greater than 5.35% is only 1%. This information provides valid quantitative risk limits for investors to use as a basis for loss mitigation in investment decision-making.

B. Discussion

The results of the data analysis and pre-processing stage confirm that BMRI stock prices exhibit strong non-linear, volatile, and time-dependent characteristics, which are typical of financial time series. The visualization of price movements and RSI indicators demonstrates that while price data capture long-term trends, momentum-based indicators such as RSI provide complementary information related to market sentiment and overbought or oversold conditions. This finding aligns with previous studies that emphasize the importance of incorporating technical indicators to enrich feature representation and improve predictive performance in stock forecasting models [10], [25]. Furthermore, the seasonal decomposition analysis reveals substantial noise in the residual component, reinforcing the argument that linear models are insufficient for capturing the complex dynamics of stock prices, as also noted by Zhang and Wen [2].

The correlation analysis further provides important insights into feature relationships. Trend-based indicators such as SMA and EMA show extremely high correlations with the closing price, indicating strong multicollinearity. While this confirms their relevance, it also suggests redundancy if used without careful modeling strategies. In contrast, RSI exhibits a weak linear correlation with price, yet this does not imply irrelevance. Instead, it indicates that RSI contributes orthogonal, non-linear information not captured by raw price features. This observation is consistent with findings by Chen et al, who argue that low-correlation features can still significantly enhance deep learning models by introducing diverse representations, [ISSN 2714-7444 \(online\)](https://doi.org/10.21070/acopen.11.2026.13521), <https://acopen.umsida.ac.id>, published by [Universitas Muhammadiyah Sidoarjo](https://www.muhammadiyah.ac.id)

particularly when attention-based mechanisms are employed [26].

The performance of the Hybrid Transformer–LSTM model demonstrates that integrating sequential learning and global attention mechanisms is highly effective for stock price prediction. The LSTM layer successfully captures temporal dependencies, while the Multi-Head Attention layer selectively emphasizes informative time steps and features. The inclusion of a residual connection (ADD layer) plays a crucial role in preserving information flow and mitigating gradient degradation, which is a known challenge in deep architectures. Similar architectural benefits have been reported in prior hybrid models combining recurrent networks and attention mechanisms, where residual learning improves both convergence stability and predictive accuracy [13], [27]. This confirms that the proposed architecture is well-suited to modeling complex financial time series.

Quantitatively, the model achieves a MAPE of 4.13%, which falls into the “highly accurate forecasting” category based on Lewis’s benchmark [28]. This level of accuracy is competitive with, and in some cases superior to, results reported in previous studies using standalone LSTM, CNN-LSTM, or ARIMA-based hybrid models, which typically report MAPE values ranging between 5% and 10% for stock price prediction [9], [29]. The low MAE value of 246.35 Rupiah further indicates that the model’s prediction errors are economically small relative to the average BMRI stock price. The close alignment between predicted and actual price trajectories observed in the validation visualization suggests that the model generalizes well and avoids overfitting, addressing a common limitation highlighted in earlier deep learning studies [30].

Beyond predictive accuracy, the integration of Monte Carlo–based Value at Risk (VaR) analysis represents a key contribution of this study. While many prior works focus solely on minimizing prediction error, they often neglect explicit risk quantification, limiting their practical relevance for investors. The VaR result of 5.35% at a 99% confidence level indicates a relatively low probability of extreme losses, providing a clear and interpretable risk boundary. This finding is consistent with studies that emphasize the effectiveness of Monte Carlo simulations for capturing uncertainty and tail risk in financial forecasting [31], [32]. By combining accurate price prediction with probabilistic risk estimation, this study advances existing literature and offers a more comprehensive decision-support framework for investment analysis.

Conclusion

This study demonstrates that the Hybrid Transformer–LSTM model is highly effective for predicting stock prices in a complex and volatile financial environment. By integrating sequential learning through LSTM with global dependency modeling via Multi-Head Attention and residual connections, the proposed architecture successfully captures both temporal patterns and contextual relationships in BMRI stock price data. The empirical results show strong predictive performance, as evidenced by a Mean Absolute Percentage Error (MAPE) of 4.13% and a low Mean Absolute Error (MAE) of 246.35 Rupiah, indicating that the model produces highly accurate and economically meaningful forecasts. These findings confirm that hybrid deep learning architectures outperform conventional and single-model approaches when applied to non-linear financial time series.

Beyond accuracy, this research highlights the importance of incorporating risk quantification into stock price prediction frameworks. The integration of Monte Carlo simulation for Value at Risk (VaR) estimation provides a probabilistic assessment of potential losses, offering practical insights for investment decision-making. The VaR result of 5.35% at the 99% confidence level indicates a relatively low probability of extreme downside risk within the forecast horizon. By combining precise price prediction with measurable risk evaluation, this study contributes to the development of more robust and investor-oriented forecasting systems, and provides a foundation for future research to extend the framework to multi-asset portfolios, longer forecasting horizons, or additional sources of market information.

References

1. C. Xu, J. Ke, Z. Peng, F. Wen, and Y. Duan, “Asymmetric Fractal Characteristics and Market Efficiency Analysis of Style Stock Indices,” *Entropy*, vol. 24, no. 7, p. 969, 2022, doi: 10.3390/e24070969.
2. S. Zhang and F. Wen, “Multifractal Behaviors of Stock Indices and Their Ability to Improve Forecasting in a Volatility Clustering Period,” *Entropy*, vol. 23, no. 8, p. 1018, 2021, doi: 10.3390/e23081018.
3. E. Y. Atanu, H. E. Ette, and E. Amos, “Comparative Performance of ARIMA and GARCH Model in Forecasting Crude Oil Price Data,” *Asian Journal of Probability and Statistics*, vol. 15, no. 4, pp. 251–275, 2021, doi: 10.9734/ajpas/2021/v15i430378.
4. Y. Xiang, “Using ARIMA-GARCH Model to Analyze Fluctuation Law of International Oil Price,” *Mathematical Problems in Engineering*, vol. 2022, pp. 1–7, 2022, doi: 10.1155/2022/3936414.
5. I. F. Amri, W. I. R. Sari, V. A. Widyasari, N. Nurohmah, and M. A. Haris, “The ARIMA-GARCH Method in Case Study Forecasting the Daily Stock Price Index of PT Jasa Marga (Persero),” *Eigen Mathematics Journal*, vol. 7, no. 1, pp. 25–33, 2024, doi: 10.29303/emj.v7i1.174.
6. F. T. A. Putri, E. Zukhronah, and H. Pratiwi, “Model ARIMA-GARCH Pada Peramalan Harga Saham PT Jasa Marga (Persero),” *Business Innovation and Entrepreneurship Journal*, vol. 3, no. 3, pp. 164–170, 2021, doi: 10.35899/biej.v3i3.308.
7. N. A. Aziz, S. N. M. Shafie, and M. N. A. Nafi, “Comparative Performance of ARIMA and GARCH Models in Modelling and Forecasting Volatility of Kuala Lumpur Composite Index,” *International Journal of Academic Research in Accounting, Finance and Management Sciences*, vol. 13, no. 1, 2023, doi: 10.6007/ijarafms/v13-i1/16213.
8. R. Oprasianti, D. Kusnandar, and W. Andani, “Stock Price Forecasting Using the Hybrid ARIMA-GARCH Model,” *Parameter Journal of Statistics*, vol. 4, no. 2, pp. 110–119, 2024, doi: 10.22487/27765660.2024.v4.i2.17162.
9. T. O. Kyaw, T. Shibayama, Y. Shibutani, and Y. Kotake, “Development of a Deep Learning Based Wave Forecasting Model Using LSTM Network,” *Coastal Engineering Proceedings*, no. 36, p. 31, 2020, doi:

10.9753/icce.v36v.waves.31.

10. S. Aryal, D. Nadarajah, L. Rupasinghe, C. Jayawardena, and D. Kasthurirathna, "Comparative Analysis of Deep Learning Models for Multi-Step Prediction of Financial Time Series," *Journal of Computer Science*, vol. 16, no. 10, pp. 1401–1416, 2020, doi: 10.3844/jcssp.2020.1401.1416.
11. C. Zhang, N. N. A. Sjarif, and R. Ibrahim, "Deep Learning Models for Price Forecasting of Financial Time Series: A Review of Recent Advancements 2020–2022," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 14, no. 1, 2023, doi: 10.1002/widm.1519.
12. Z. Mao and W. Chong, "Stock Price Index Prediction Based on SSA-BiGRU-GSCV Model From the Perspective of Long Memory," *Kybernetes*, vol. 53, no. 12, pp. 5905–5931, 2023, doi: 10.1108/K-02-2023-0286.
13. Y. Si, S. Nadarajah, Z. Zong-Xin, and C. Xu, "Modeling Opening Price Spread of Shanghai Composite Index Based on ARIMA-GRU/LSTM Hybrid Model," *PLOS ONE*, vol. 19, no. 3, p. e0299164, 2024, doi: 10.1371/journal.pone.0299164.
14. R. Zhao, Z. Lei, and Z. Zhao, "Application of Deep Learning Techniques in Stock Market Prediction and Investment Decision Making in Financial Management," *Frontiers in Energy Research*, vol. 12, 2024, doi: 10.3389/fenrg.2024.1376677.
15. M. Rizki et al., "Prediksi Harga Saham Bank BRI dan Bank BCA Dengan Menggunakan Model LSTM," *RIGGS Journal of Artificial Intelligence and Digital Business*, vol. 4, no. 2, pp. 4554–4560, 2025, doi: 10.31004/riggs.v4i2.1264.
16. S. Wang, "A Stock Price Prediction Method Based on BiLSTM and Improved Transformer," *IEEE Access*, vol. 11, pp. 104211–104223, 2023, doi: 10.1109/ACCESS.2023.3296308.
17. A. Humayrah and D. P. Sari, "Analisis Risiko Investasi Saham Tunggal Syariah Dengan Value at Risk Menggunakan Simulasi Monte Carlo," *Jurnal Matematika UNP*, vol. 8, no. 1, pp. 32–35, 2023.
18. A. Y. Febriyanti, D. A. Prasetya, and T. Trimono, "Stock Price Prediction and Risk Estimation Using Hybrid CNN- LSTM and VaR-ECF," *Jurnal Teknik Informatika*, vol. 6, no. 3, pp. 1539–1554, 2025, doi: 10.52436/1.jutif.2025.6.3.4648.
19. K. Cao, T. Zhang, and J. Huang, "Advanced Hybrid LSTM-Transformer Architecture for Real-Time Multi-Task Prediction in Engineering Systems," *Scientific Reports*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-55483-x.
20. D. I. Puteri, "Implementasi Long Short Term Memory (LSTM) dan Bidirectional LSTM Dalam Prediksi Harga Saham Syariah," *Euler Journal of Mathematics, Science and Technology*, vol. 11, no. 1, pp. 35–43, 2023, doi: 10.34312/euler.v11i1.19791.
21. M. Sangkala, "Analisis Teknikal Sebagai Dasar Pengambilan Keputusan Dalam Trading Saham Pada Bursa Efek Indonesia," *Gemilang Journal of Management and Accounting*, vol. 5, no. 2, pp. 652–660, 2025, doi: 10.56910/gemilang.v5i2.2530.
22. A. Vaswani et al., "Attention Is All You Need," *arXiv preprint arXiv:1706.03762*, 2017. [Online]. Available: <http://arxiv.org/abs/1706.03762>
23. N. Selayanti et al., "Prediksi Harga Penutupan Saham BBRI Dengan Model Hybrid LSTM-XGBoost," *Informatika: Journal of Informatics and Multimedia*, vol. 5, no. 1, pp. 52–64, 2025, doi: 10.51903/informatika.v5i1.1011.
24. B. N. Cecevic, L. Antic, and A. Jevtic, "Stock Price Prediction of the Largest Automotive Competitors Based on the Monte Carlo Method," *Economic Themes*, vol. 61, no. 3, pp. 419–441, 2023, doi: 10.2478/ethemes-2023-0022.
25. M. Li, Y. Zhu, Y. Shen, and M. Angelova, "Clustering-Enhanced Stock Price Prediction Using Deep Learning," *World Wide Web*, vol. 26, no. 1, pp. 207–232, 2022, doi: 10.1007/s11280-021-01003-0.
26. Z. Chen et al., "Multi-Scale Spatial Temporal Graph Convolutional Network for Skeleton-Based Action Recognition," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 2, pp. 1113–1122, 2021, doi: 10.1609/aaai.v35i2.16197.
27. Y. Li et al., "Empathy With Nature Promotes Pro-Environmental Attitudes in Preschool Children," *PsyCh Journal*, 2024, doi: 10.1002/pchj.735.
28. S. Lewis, "Providing a Platform for 'What Works': Platform-Based Governance and the Reshaping of Teacher Learning Through the OECD's PISA4U," *Comparative Education*, vol. 56, no. 4, pp. 484–502, 2020, doi: 10.1080/03050068.2020.1769926.
29. S. Cheng, "Heterogeneity in Stock Price Forecasting Based on the ARIMA-GARCH Model and PCA-LSTM Model," *Highlights in Science, Engineering and Technology*, vol. 88, pp. 39–46, 2024, doi: 10.54097/me8jvh37.
30. A. Casolaro et al., "Deep Learning for Time Series Forecasting: Advances and Open Problems," *Information*, vol. 14, no. 11, p. 598, 2023, doi: 10.3390/info14110598.
31. S.-H. Sung et al., "Cryptocurrency Log-Return Price Prediction Using Multivariate Time-Series Model," *Axioms*, vol. 11, no. 9, p. 448, 2022, doi: 10.3390/axioms11090448.
32. M. Faal and F. Almasganj, "ECG Signal Modeling Using Volatility Properties: Its Application in Sleep Apnea Syndrome," *Journal of Healthcare Engineering*, vol. 2021, pp. 1–12, 2021, doi: 10.1155/2021/4894501.