Evaluating Performance of LSTM and Transformer Models in Forecasting Bank Stocks in Indonesia

Panji Arisaputra1,a), Rezki Yunanda2,b), Agung Sudjatmoko3,c), and Adi Teguh Suprapto4,d)

1,2Software Engineering Program, Computer Science Department, Bina Nusantara University, Jakarta, Indonesia

3,4Business Management Program, Bina Nusantara University, Jakarta, Indonesia

a) Corresponding author: [panji.arisaputra@binus.ac.id](mailto:panji.arisaputra@binus.ac.id)

b) rezki.yunanda@binus.ac.id

c) agung.sudjatmoko@binus.edu

d) adi.suprapto@binus.edu

**Abstract.**  Stock price forecasting is critical for financial decision-making, particularly in volatile markets like Indonesia. This study evaluates the performance of Long Short-Term Memory (LSTM) and Transformer models in predicting stock prices of Indonesian banks. Both models are widely recognized for their effectiveness in handling sequential data, making them suitable for time series forecasting. The research objectives are to compare the models' accuracy and performance using historical stock price data. The study employs Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R-squared (R²), Mean Absolute Percentage Error (MAPE), and Median Absolute Error (MedAE) as evaluation metrics. The results indicate that the LSTM model generally outperforms the Transformer model, particularly in long-term stock price predictions, due to its architecture that better captures long-term dependencies in time series data. However, the Transformer model demonstrates its strength in capturing short-term patterns and dealing with stock price volatility, making it more effective for short-term forecasting. These findings highlight the importance of selecting the appropriate model based on the nature of the data and the forecasting horizon. In conclusion, while LSTM is better suited for long-term predictions, the Transformer model holds potential for short-term applications, particularly in environments with high market fluctuations.

**Keywords:** Stock Price Prediction, LSTM, Transformer Model, Machine Learning, Comparative Study

# INTRODUCTION

In today's financial landscape, accurate stock price predictions are essential for informed decision-making among investors and market analysts [1]. The advent of artificial intelligence, particularly machine learning techniques, has dramatically improved the sophistication and precision of financial market forecasting compared to traditional models [2, 3]. This study explores the application of two prominent machine learning models, Long Short-Term Memory (LSTM) and Transformer, to forecast the daily stock prices of 47 banks in Indonesia from July 11, 2022, to July 9, 2024.

LSTM and Transformer models are known for their ability to process time-series data, making them suitable for predicting stock prices, which are often influenced by temporal patterns. While both models have been applied in financial markets globally, research specifically targeting the Indonesian market remains limited [4]. This study addresses this gap by focusing on Indonesia's banking sector, which is marked by unique market dynamics and economic fluctuations.

Researchers have shown that Long Short-Term Memory (LSTM) and Transformer could greatly improve the accuracy of predicting financial time series [5-7]. This is attributed to its deep learning structure, which allows it to process input in a sequential manner and store information over long periods of time [8, 9]. However, research specifically addressing the Indonesian market and daily fluctuations within its banking sector remains sparse, indicating a significant gap that this study aims to fill.

The primary objective of this research is to compare the performance of LSTM and Transformer models in forecasting stock prices in Indonesia’s banking sector. The study employs rigorous evaluation metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R-squared (R²), Mean Absolute Percentage Error (MAPE), and Median Absolute Error (MedAE). Through these comparisons, we aim to identify which model is more effective in capturing the short-term and long-term trends in this volatile market. Specific questions include:

* Between LSTM and the Transformer model, which machine learning models are the most successful for this purpose?
* How do these models compare in accuracy and reliability?

Enhancing the accuracy of stock price predictions can significantly benefit investors and financial analysts by providing more reliable data for making investment decisions [1]. This research contributes to the field of financial forecasting by providing novel insights into the performance of advanced machine learning models within an emerging market. By specifically addressing the Indonesian banking sector, the study not only fills a gap in the existing literature but also offers practical implications for enhancing predictive accuracy in markets characterized by high volatility. Additionally, the findings can aid in the development of more robust and reliable machine learning models for time-series forecasting, contributing to the broader field of financial technology and data-driven decision-making.

## RELATED WORKS

Daniel Boyle and Jugal Kalita present a research paper titled "Spatiotemporal Transformer for Stock Movement Prediction," where they propose a hybrid model called Spatiotemporal Transformer-LSTM (STST) [10]. The purpose of this model is to improve the precision of predicting stock movements in U.S. stock exchanges. This study examines the integration of LSTM's temporal processing capabilities with Transformers' spatial relationship insights to enhance the comprehension of non-linear interactions in financial markets, hence offering a more sophisticated understanding of their functioning. The model has undergone testing using benchmark datasets such as ACL18 and KDD17. It has achieved accuracies of 63.707% and 56.879% respectively. Furthermore, it has proved its practical usefulness by outperforming the S&P 500 index in simulated trading scenarios, resulting in a remarkable annualized return of 31.24%. This research not only demonstrates the model's theoretical capabilities but also its potential for practical use, making it a valuable tool for investors seeking to utilize sophisticated analytics in stock market investments.

The paper named "Forecasting stock prices with long-short term memory neural network based on attention mechanism" is another research project [11]. Qiu, Wang, and Zhou investigate the utilization of LSTM neural networks augmented with an attention mechanism to boost the precision of stock price forecasts. The researchers utilize wavelet transform to preprocess the data, successfully removing noise from previous stock information to improve model accuracy. By conducting thorough trials on datasets such as S&P 500 and DJIA, the researchers show that their model, which combines LSTM with attention mechanisms and wavelet denoising, outperforms typical LSTM and GRU models in terms of prediction performance. More precisely, their model demonstrates a notably higher coefficient of determination (R² above 0.94) and a reduced mean square error (below 0.05). This indicates that the attention-based LSTM model is particularly proficient in handling the intricacies of financial time series data. This study presents strong evidence of the potential of sophisticated neural network designs in predicting financial outcomes. The findings give valuable insights that could enhance automated trading methods and risk management in highly unstable markets.

Wang and Yuan conduct a thorough examination to assess the efficacy of LSTM, Hidden Markov Model (HMM), and Transformer models in predicting stock prices within China's new energy vehicle industry. The models are thoroughly tested using data from the A-share market, with a specific focus on the new energy vehicles sector, to assess their forecast accuracy. Their research demonstrates that the Transformer model surpasses both LSTM and HMM in terms of Mean Absolute Percentage Error (MAPE) and Matthews Correlation Coefficient (MCC), indicating its superior ability to handle the intricacies of financial time series data. The study highlights the potential of utilizing advanced deep learning methods, specifically the Transformer model, to improve stock price prediction models. This offers useful information for investors dealing with the unpredictable new energy vehicle market [7].

Although machine learning models such as LSTM and Transformer have made significant progress in financial forecasting, the Indonesian banking system has distinct issues that have not been thoroughly investigated. Present study predominantly concentrates on wider market patterns and significant economies, resulting in a lack of localized investigations that consider the daily variations and distinctive attributes of Indonesian banks. The objective of this study is to fill these knowledge gaps by customizing models to gain a better understanding of and make more accurate predictions about the dynamics of the Indonesian market.

# RESEARCH METHOD

## DATA COLLECTION

The dataset for this study was meticulously compiled using comprehensive financial data sourced from Refinitiv, a renowned provider of financial market data and infrastructure. The data encompasses daily stock price movements of 47 Indonesian banks from July 11, 2022, to July 9, 2024. This timeframe was chosen to ensure a broad representation of market behaviors and economic cycles affecting the banking sector. Refinitiv was selected for its reliability and the depth of its financial databases, which are crucial for enhancing the accuracy of predictive models like LSTM and Transformer. This data will be used to compare the performance of these models in predicting stock price trends. The flowchart of research stage is shown in **FIGURE 1**.

A diagram of a model training

Description automatically generated

**FIGURE 1**. Research Stage

## DATA PREPROCESSING

The data preprocessing stage involves multiple crucial steps to ensure the quality and suitability of the data for model training. These steps include handling missing values, detecting and addressing outliers, normalizing the data, creating sequences for time series prediction, and splitting the data into training and testing sets [12].

The collected data was initially examined for any missing values or outliers. Missing values were handled using forward-filling or interpolation methods to maintain a continuous time series. Outliers were detected using statistical methods and addressed to prevent them from skewing the model's learning process.

Stock prices were normalized to a scale between 0 and 1 using the MinMaxScaler from the Scikit-Learn library. Normalization ensures that all input features are on the same scale, which is essential for the gradient-based optimization algorithms used in training deep learning models [13].

For time series prediction, sequences of historical prices were created as inputs to the models. Each sequence consisted of a fixed number of past daily closing prices, with a sequence length of 30 days. This means each input sequence to the model contains the closing prices for the past 30 days, and the corresponding target is the closing price of the next day.

The preprocessed data was divided into training and testing sets to ensure robust evaluation. 80% of the data was used for training, and the remaining 20% was reserved for testing. To leverage the diversity in stock price movements, data from multiple stocks were combined, creating comprehensive datasets. This approach ensures that models are exposed to a variety of price patterns during training, potentially enhancing their generalization capabilities [14, 15].

## TRAINING LSTM MODEL

The Long Short-Term Memory (LSTM) network, a type of recurrent neural network (RNN), is particularly effective for time series prediction tasks due to its ability to address the vanishing gradient problem [16]. The LSTM model architecture implemented in this study, using TensorFlow and Keras, consists of the following layers:

1. Input Layer

The input layer receives sequences of stock prices, each with a fixed length of 30 days, where each day is represented by a single feature (the closing price).

1. LSTM Layers

The core of the model consists of two stacked LSTM layers responsible for capturing the temporal dependencies and patterns in the input sequences.

1. First LSTM Layer: Contains 50 units. This layer is configured to return the full sequence of outputs for each input sequence, which is necessary for the second LSTM layer to process the full sequence of outputs from the first layer.
2. Second LSTM Layer: Also contains 50 units. This layer processes the sequence of outputs from the first LSTM layer and extracts higher-level temporal features. It is configured to return only the final output of the sequence, which is then passed to the Dense layer for the final prediction.
3. Dense Layer

The Dense layer is a fully connected layer that takes the final output from the second LSTM layer and produces a single output value, representing the predicted stock price for the next day. The Swish activation function is used in this layer for improved learning performance [17].

## TRAINING TRANSFORMER MODEL

Transformers have revolutionized sequence-to-sequence tasks, particularly in natural language processing, and their ability to handle long-range dependencies makes them an excellent candidate for time series forecasting [18]. The Transformer model used in this study is designed to capture complex temporal patterns in stock price data [19]. The architecture consists of several key components:

1. Input Layer

The input layer accepts sequences of stock prices, each with a length of 30 days, similar to the LSTM model. Each day is represented by a single feature (the closing price).

1. Transformer Encoder Layers

The core of the model consists of multiple stacked Transformer encoder layers. Each encoder layer includes multi-head self-attention mechanisms and feed-forward neural networks, which enable the model to focus on different parts of the input sequence and capture intricate temporal dependencies [20]. Each encoder layer includes the following key components:

1. Multi-Head Self-Attention Mechanism: This mechanism allows the model to focus on different parts of the input sequence simultaneously, capturing various temporal dependencies and relationships within the data. The self-attention mechanism computes a weighted sum of the input values, where the weights are determined by the similarity between the elements. The multi-head mechanism performs multiple attention functions, each focusing on different parts of the sequence, capturing various aspects of the data simultaneously.
2. Layer Normalization and Residual Connections: The Transformer encoder employs layer normalization and residual connections to improve training stability and performance. Layer normalization stabilizes and accelerates the training process, while residual connections help in training deeper models by adding the original input to the output of the layer, mitigating the vanishing gradient problem.
3. Feed-Forward Neural Networks: These layers process the outputs of the self-attention mechanism, enabling the model to learn complex patterns. After the attention mechanism, the output passes through a feed-forward neural network consisting of two dense layers with Swish activation, processing the attention outputs and enabling the model to learn complex patterns. Dropout layers are added to prevent overfitting by randomly setting a fraction of input units to zero during training.
4. Global Average Pooling Layer

After the Transformer encoder layers, a Global Average Pooling layer reduces the sequence of outputs to a fixed-size vector, providing a summary representation of the input sequence, regardless of the input sequence length.

1. Dense Layers

The pooled output is passed through several fully connected (Dense) layers, further processing the features extracted by the Transformer encoders and producing the final prediction. The Swish activation function is used for improved learning performance [17].

## MODEL COMPILATION AND TRAINING PROCESS

Both LSTM and Transformer models were compiled and trained using a standardized process to ensure consistent evaluation. The models utilized the Adam optimizer and Mean Squared Error (MSE) loss function. The Adam optimizer, known for its efficiency in training deep learning models, adapts the learning rate for each parameter. The MSE loss function measures the average squared difference between predicted and actual stock prices, providing a clear metric of prediction accuracy [21].

Training Steps:

* Training Data: The combined training dataset was used for model training, ensuring exposure to a variety of stock price patterns.
* Validation Split: A validation split of 20% was implemented to monitor the model's performance and prevent overfitting during training.
* Epochs and Batch Size: Both models were trained for 50 epochs with a batch size of 16, balancing computational efficiency and learning capability.
* Early Stopping: Early stopping was implemented to halt training when the validation loss stopped improving, thus preventing overfitting and reducing training time.

This training strategy, including the use of the Swish activation function, ensured that both LSTM and Transformer models effectively captured temporal dependencies in stock price data, facilitating accurate time series forecasting [8].

## MODEL EVALUATION

Both LSTM and Transformer models were evaluated on the combined and individual test dataset to assess their performance in predicting stock prices. The evaluation process involved generating predictions for the test set and calculating various performance metrics to provide a comprehensive comparison. The performance metrics used in this study include:

* Mean Squared Error (MSE): Measures the average of the squares of the errors between the predicted and actual values.
* Mean Absolute Error (MAE): Measures the average of the absolute differences between the predicted and actual values.
* Root Mean Squared Error (RMSE): The square root of the Mean Squared Error, providing an error metric in the same units as the original data.
* R-squared (R²): Indicates the proportion of the variance in the dependent variable that is predictable from the independent variables.
* Mean Absolute Percentage Error (MAPE): Measures the average of the absolute percentage differences between the predicted and actual values.
* Median Absolute Error (MedAE): Measures the median of the absolute differences between the predicted and actual values.

The evaluation process involved the following steps for both the LSTM and Transformer models:

1. Generating Predictions: Predictions were generated for the combined and individual test dataset using both models.
2. Inverse Transforming Predictions: The predicted values were inverse transformed to their original scale using the scalers fitted during the data preprocessing step.
3. Calculating Performance Metrics: The performance metrics were calculated by comparing the predicted values with the actual values from the test set.

## DATA PREDICTION AND VISUALIZATION

To assess the models' performance comprehensively, predicted stock prices were compared against actual stock prices for the test set. Visualization played a key role in understanding model behavior over time, including both training and test sets, as well as future predictions for each dataset.

1. Data Prediction:
2. Generating Predictions for the Training and Test Sets: Predictions were generated for both training and test datasets using the LSTM and Transformer models. These predictions were then inverse transformed to their original scale using the scalers fitted during data preprocessing, allowing for direct comparison with actual stock prices.
3. Generating Future Predictions: Predictions for the next 30 days were generated based on the last available data from the test set for each individual stock and the combined dataset. This demonstrates the models' capability to forecast future stock prices beyond the test period.
4. Visualization:

The visual comparison was provided to clearly illustrate the models' performance:

1. Training Set Predictions: Actual stock prices from the training set were plotted alongside predictions from both the LSTM and Transformer models for each individual stock. This comparison highlights how well each model learned from the training data.
2. Test Set Predictions: Actual stock prices from the test set were plotted alongside predictions from both models for each individual stock, showing the models' accuracy in predicting unseen data.
3. Future Predictions: Predicted stock prices for the next 30 days were plotted to showcase the forecasting capabilities of each model for individual stocks.

This comprehensive evaluation, including both statistical metrics and visualizations, provides detailed insights into the models' predictive performance and their potential for real-world application in stock price forecasting.

# RESULTS AND DISCUSSIONS

Both LSTM and Transformer models were evaluated on the combined dataset and 47 individual stock datasets using various performance metrics to assess accuracy and reliability.

1. Performance Metrics:

The following performance metrics were calculated for both the LSTM and Transformer models:

1. Mean Squared Error (MSE)

(1)

1. Mean Absolute Error (MAE)

(2)

1. Root Mean Squared Error (RMSE)

(3)

1. R-squared (R²)

(4)

1. Mean Absolute Percentage Error (MAPE)

(5)

1. Mean Absolute Percentage Error (MAPE)

(6)

1. Visual Comparison

The visualizations of the actual vs predicted stock prices for both models, including predictions on the training and test sets as well as future predictions, provided further insights into the models' performance.

1. Combined Dataset Performance

On **FIGURE 2**, the LSTM model demonstrated a closer fit to the actual stock prices on the combined test set, indicating strong learning. The Transformer model's predictions followed the general trends of the actual stock prices but showed higher deviations compared to the LSTM model.

1. Individual Stocks Performance

The individual stock visualizations highlighted the models' performance across different datasets on **FIGURE 3**. The LSTM model's future predictions for the next 30 days were consistent and closely aligned with the actual trends, suggesting better performance in capturing long-term dependencies. The Transformer model showed consistent trends but with larger deviations, suggesting potential issues in capturing fine-grained details.

LSTM

A screenshot of a graph

Description automatically generated

Transformer

A screen shot of a graph

Description automatically generated

**FIGURE 2**. Stock price prediction on combined test set



**FIGURE 3**. Examples of three (out of 47) individual data test set of stock price prediction

1. Performance Metrics Analysis

The calculated performance metrics provide insights into the accuracy and reliability of the models. Table I and Table II summarize the performance metrics for both the LSTM and Transformer models.

1. Prediction Accuracy:

The LSTM model achieved higher prediction accuracy on both the combined dataset and individual stocks, as evidenced by the lower MSE, MAE, and RMSE values. It demonstrated better performance in capturing fine-grained details and trends, likely due to its ability to effectively learn short-term dependencies.

The Transformer model, while capable of capturing overall trends, showed higher deviations, indicating challenges in capturing finer details [22].

1. Generalization Performance:

The R-squared values indicated strong generalization performance for both models, with the LSTM model significantly outperforming the Transformer model. This suggests that LSTMs may be better suited for capturing both short-term and long-term patterns in stock price data.

1. Forecasting Capability:

The future predictions for the next 30 days showed that the LSTM model could forecast stock prices with a higher degree of accuracy and consistency. The Transformer's future predictions were less stable, indicating potential limitations in capturing long-term trends.

1. Model Robustness:

The LSTM model demonstrated better robustness in handling different stock datasets, particularly those with higher volatility. This robustness makes the LSTM model a promising choice for stock price prediction in diverse market conditions.

**TABLE 1** and **2** provide a detailed comparison of the LSTM and Transformer models' performance across various metrics, highlighting the LSTM model's superior accuracy, generalization, and forecasting capabilities.

**TABLE 1**. Comparison of LSTM and Transformer Models Performance on Combined Test Set

|  |  |  |
| --- | --- | --- |
| **Metrics** | **LSTM Model Performance** | **Transformer Model Performance** |
| **MSE** | 755.92637 | 5706.88715 |
| **MAE** | 17.09573 | 57.06146 |
| **RMSE** | 27.49411 | 75.54394 |
| **R²** | 0.97756 | 0.81337 |
| **MAPE (%)** | 3.35473 | 11.73998 |
| **MedAE** | 10.43721 | 44.27827 |

The LSTM model consistently outperformed the Transformer model across all performance metrics, especially on the combined dataset. This can be attributed to LSTM's inherent ability to capture long-term dependencies in sequential data, which is crucial for accurate stock price prediction [23]. The Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R²) metrics all favor LSTM, particularly in long-term predictions, demonstrating its robustness in handling temporal dependencies. However, the Transformer model showed promise in short-term stock price forecasting due to its powerful attention mechanism, which allows it to capture intricate short-term patterns and volatility. This aligns with the findings of Wang and Yuan (2021), who reported that Transformers excel in capturing short-term fluctuations in complex datasets. Despite this, the Transformer model's overall performance lagged behind LSTM when evaluated on long-term data, indicating potential weaknesses in handling extended temporal dependencies in financial time series [24].

The visualizations provided on Figure. 2 and Figure. 3 further evidence of LSTM's superiority. On the combined dataset (Figure 2), the LSTM model closely tracked the actual stock prices, while the Transformer's predictions, though following general trends, exhibited larger deviations. This is consistent with studies such as Daniel Boyle and Jugal Kalita [10] who found that hybrid models like the Spatiotemporal Transformer-LSTM could mitigate some of these issues by combining the strengths of both models​.

Compared to similar research, this study's results align with prior findings that LSTM models perform better in long-term forecasting, particularly in volatile markets. For instance, in the work of Qiu, Wang, and Zhou [12], LSTM models enhanced with attention mechanisms achieved higher accuracy in long-term predictions. While the Transformer model's attention mechanism offers advantages in capturing short-term trends, it seems less capable of sustaining predictive accuracy over extended periods in comparison to LSTM [25].

Additionally, Qiu, Wang, and Zhou's research showed that Transformer models, when combined with wavelet denoising and attention mechanisms, can outperform basic LSTM models in handling complex time series. However, in this study, the LSTM's simplicity and focus on long-term memory made it better suited for the Indonesian banking sector's specific challenges, including its market volatility and regulatory complexities.**TABLE 2**. Comparison of LSTM (1) and Transformer (2) models Performance on Each Individual Data Test Set

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Data** | **MSE** | | **MAE** | | **RMSE** | | **R²** | | **MAPE (%)** | | **MedAE** | |
| ***1*** | ***2*** | ***1*** | ***2*** | ***1*** | ***2*** | ***1*** | ***2*** | ***1*** | ***2*** | ***1*** | ***2*** |
| **1** | **1317 AGRO** | 845.41 | 5592.38 | 19.44 | 65.01 | 29.07 | 74.78 | 0.92 | 0.48 | 3.05 | 9.43 | 13.72 | 66.27 |
| **2** | **1319 AGRS** | 7.27 | 50.65 | 2.02 | 5.30 | 2.69 | 7.11 | 0.92 | 0.49 | 1.92 | 4.96 | 1.48 | 3.91 |
| **3** | **1320 AMAR** | 54.80 | 739.35 | 5.66 | 23.42 | 7.40 | 27.19 | 0.91 | -0.16 | 2.13 | 9.00 | 4.43 | 19.45 |
| **4** | **1324 ARTO** | 101876.90 | 1406642.45 | 232.70 | 987.33 | 319.18 | 1186.01 | 0.97 | 0.64 | 3.25 | 12.01 | 184.80 | 1092.62 |
| **5** | **1325 BABP** | 15.14 | 78.96 | 2.28 | 7.72 | 3.89 | 8.88 | 0.85 | 0.26 | 1.95 | 6.51 | 1.33 | 7.08 |
| **6** | **1326 BACA** | 28.43 | 139.13 | 3.26 | 9.93 | 5.33 | 11.79 | 0.81 | 0.11 | 2.37 | 7.24 | 2.06 | 10.20 |
| **7** | **1327 BANK** | 2963.44 | 19641.58 | 38.38 | 125.19 | 54.43 | 140.14 | 0.93 | 0.55 | 2.42 | 7.65 | 23.44 | 115.42 |
| **8** | **1329 BBCA** | 12959.57 | 141173.27 | 92.14 | 306.10 | 113.84 | 375.73 | 0.95 | 0.47 | 1.13 | 3.92 | 79.76 | 293.06 |
| **9** | **1330 BBHI** | 11456.28 | 202255.05 | 75.92 | 350.33 | 107.03 | 449.72 | 0.97 | 0.53 | 3.01 | 12.13 | 65.66 | 251.17 |
| **10** | **1331 BBKP** | 12.27 | 214.95 | 2.45 | 12.94 | 3.50 | 14.66 | 0.97 | 0.58 | 1.48 | 7.62 | 1.74 | 11.17 |
| **11** | **1332 BBMD** | 1431.27 | 4554.31 | 25.79 | 34.97 | 37.83 | 67.48 | 0.72 | 0.13 | 1.27 | 1.66 | 18.44 | 21.35 |
| **12** | **1333 BBNI** | 4582.99 | 54754.93 | 50.87 | 192.25 | 67.69 | 233.99 | 0.94 | 0.35 | 1.20 | 4.69 | 37.45 | 147.05 |
| **13** | **1334 BBRI** | 3722.32 | 739.35 | 47.88 | 23.42 | 61.01 | 27.19 | 0.84 | -0.16 | 1.08 | 9.00 | 39.99 | 19.45 |
| **14** | **1337 BBSI** | 845.41 | 5592.38 | 19.44 | 65.01 | 29.07 | 74.78 | 0.92 | 0.48 | 3.05 | 9.43 | 13.72 | 66.27 |
| **15** | **1338 BBTN** | 485.50 | 2366.91 | 16.23 | 42.28 | 22.03 | 48.65 | 0.77 | -0.10 | 1.11 | 2.92 | 11.89 | 43.20 |
| **16** | **1340 BBYB** | 2119.12 | 17518.56 | 32.70 | 109.28 | 46.03 | 132.35 | 0.94 | 0.57 | 3.30 | 10.20 | 24.79 | 95.32 |
| **17** | **1341 BCIC** | 8.34 | 63.82 | 2.08 | 7.02 | 2.88 | 7.98 | 0.94 | 0.54 | 1.37 | 4.58 | 1.55 | 7.12 |
| **18** | **1342 BDMN** | 6860.58 | 73176.29 | 54.22 | 209.84 | 82.82 | 270.51 | 0.93 | 0.28 | 1.87 | 7.47 | 30.93 | 149.49 |
| **19** | **1344 BEKS** | 0.05 | 0.71 | 0.24 | 0.84 | 0.24 | 0.84 | 0.00 | 0.00 | 0.48 | 1.68 | 0.24 | 0.84 |
| **20** | **1345 BGTG** | 7.44 | 78.65 | 1.98 | 7.10 | 2.72 | 8.86 | 0.95 | 0.52 | 1.80 | 6.10 | 1.36 | 4.99 |
| **21** | **1346 BGTG** | 1641.80 | 5015.11 | 28.48 | 55.65 | 40.51 | 70.81 | 0.56 | -0.32 | 0.74 | 1.46 | 17.59 | 47.31 |
| **22** | **1346 BJBR** | 118.81 | 671.50 | 7.93 | 21.64 | 10.90 | 25.91 | 0.79 | -0.13 | 0.58 | 1.57 | 5.87 | 19.34 |
| **23** | **1348 BJTM** | 24.55 | 739.35 | 4.11 | 23.42 | 4.95 | 27.19 | 0.76 | -0.16 | 0.57 | 9.00 | 2.96 | 19.45 |
| **24** | **1348 BKSW** | 3.38 | 10.71 | 1.48 | 2.96 | 1.84 | 3.27 | 0.78 | 0.32 | 1.34 | 2.68 | 1.26 | 2.87 |
| **25** | **1349 BMAS** | 163.23 | 3660.19 | 10.06 | 42.21 | 12.77 | 60.49 | 0.28 | -14.92 | 1.99 | 8.42 | 8.02 | 23.57 |
| **26** | **1350 BMRI** | 6529.63 | 98493.12 | 60.61 | 282.37 | 80.80 | 313.83 | 0.96 | 0.45 | 1.36 | 6.57 | 53.03 | 303.10 |
| **27** | **1351 BNBA** | 3013.98 | 25088.51 | 37.01 | 135.09 | 54.89 | 158.39 | 0.91 | 0.27 | 2.21 | 7.98 | 27.63 | 122.26 |
| **28** | **1352 BNGA** | 291.63 | 5482.91 | 12.51 | 61.31 | 17.07 | 74.04 | 0.88 | -1.07 | 1.16 | 5.76 | 7.94 | 61.38 |
| **29** | **1353 BNII** | 7.23 | 47.69 | 2.02 | 5.98 | 2.68 | 6.90 | 0.91 | 0.42 | 0.78 | 2.29 | 1.20 | 5.65 |
| **30** | **1353 BNLI** | 164.89 | 1229.98 | 9.60 | 27.70 | 12.84 | 35.07 | 0.79 | -0.53 | 0.81 | 2.32 | 7.12 | 21.44 |
| **31** | **1355 BRIS** | 1442.31 | 4630.15 | 22.18 | 46.92 | 37.97 | 68.04 | 0.81 | 0.39 | 1.56 | 3.32 | 12.51 | 28.50 |
| **32** | **1356 BSIM** | 82.50 | 4843.97 | 7.20 | 49.55 | 9.08 | 69.59 | 0.85 | -7.72 | 1.18 | 7.96 | 5.46 | 26.43 |
| **33** | **1357 BSWD** | 109.19 | 739.35 | 10.44 | 23.42 | 10.44 | 27.19 | -5.28 | -0.16 | 0.57 | 9.00 | 10.44 | 19.45 |
| **34** | **1358 BTPN** | 1306.04 | 5806.77 | 19.06 | 53.85 | 36.13 | 76.20 | 0.79 | 0.06 | 0.75 | 2.14 | 11.21 | 35.63 |
| **35** | **1359 BTPS** | 6588.15 | 30998.56 | 59.41 | 156.20 | 81.16 | 176.06 | 0.75 | -0.17 | 2.09 | 5.43 | 44.28 | 138.57 |
| **36** | **1359 BVIC** | 11.05 | 40.67 | 2.39 | 5.52 | 3.32 | 6.37 | 0.63 | -0.34 | 1.79 | 4.09 | 1.70 | 5.27 |
| **37** | **1400 DNAR** | 20.52 | 69.11 | 3.03 | 6.71 | 4.53 | 8.31 | 0.72 | 0.08 | 1.71 | 3.74 | 1.96 | 6.16 |
| **38** | **1401 INPC** | 2.08 | 6.27 | 0.98 | 2.12 | 1.44 | 2.50 | 0.75 | 0.24 | 1.13 | 2.40 | 0.56 | 1.98 |
| **39** | **1402 MASB** | 4496.06 | 9249.83 | 35.54 | 63.15 | 67.05 | 96.17 | 0.62 | 0.22 | 1.01 | 1.83 | 12.43 | 34.91 |
| **40** | **1403 MAYA** | 55.27 | 474.08 | 5.53 | 15.39 | 7.43 | 21.77 | 0.85 | -0.25 | 1.65 | 4.53 | 3.67 | 11.67 |
| **41** | **1404 MCOR** | 1.32 | 10.11 | 0.95 | 2.43 | 1.14 | 3.18 | 0.85 | -0.14 | 1.06 | 2.64 | 0.61 | 1.90 |
| **42** | **1405 MEGA** | 13542.07 | 35440.07 | 78.42 | 152.28 | 116.37 | 188.25 | 0.70 | 0.21 | 1.47 | 2.89 | 46.08 | 122.95 |
| **43** | **1406 NISP** | 182.47 | 739.35 | 9.43 | 23.42 | 13.50 | 27.19 | 0.92 | -0.16 | 1.35 | 9.00 | 7.18 | 19.45 |
| **44** | **1406 NOBU** | 207.70 | 1679.43 | 9.76 | 34.04 | 14.41 | 40.98 | 0.85 | -0.18 | 1.77 | 6.37 | 7.30 | 24.81 |
| **45** | **1407 PNBN** | 7254.79 | 64233.04 | 61.27 | 197.18 | 85.17 | 253.44 | 0.90 | 0.15 | 2.95 | 9.44 | 41.85 | 169.25 |
| **46** | **1408 PNBS** | 7.31 | 36.55 | 1.98 | 5.11 | 2.70 | 6.04 | 0.84 | 0.22 | 2.50 | 6.47 | 1.37 | 4.73 |
| **47** | **1409 SDRA** | 36.23 | 73.24 | 4.55 | 7.19 | 6.01 | 8.55 | 0.56 | 0.12 | 0.81 | 1.28 | 3.25 | 7.55 |
| **Average** | | 4203.97 | 47550.71 | 26.20 | 86.98 | 36.93 | 106.90 | 0.68 | -0.35 | 1.62 | 5.72 | 19.22 | 78.96 |

# CONCLUSIONS

In conclusion, both LSTM and Transformer models proved to be effective for stock price prediction, with each model exhibiting unique strengths. The LSTM models are recommended for those interested in long-term investments, as they provide more reliable predictions over extended periods. In contrast, investors focused on short-term trades, particularly in highly volatile markets, may benefit from using Transformer models. These models' ability to capture immediate price fluctuations could inform high-frequency trading strategies.

To answer the question of which model is most successful for stock price prediction purposes, it can be concluded that the consistent performance of the LSTM model across the combined and individual datasets suggests that this model is more suitable for long-term stock price forecasting in the volatile Indonesian banking sector. However, the strength of the Transformer model in short-term forecasting highlights its potential in certain applications, such as high-frequency trading.

Future work could explore hybrid models that combine the strengths of both LSTM and Transformer architectures, potentially leading to even more accurate and robust stock price prediction models. Additionally, incorporating external factors such as market sentiment and macroeconomic indicators could further enhance the predictive power of these models.

# Acknowledgments

The Research and Technology Transfer Office (RTTO) of Bina Nusantara University is endorsing a paper titled "Evaluating Performance of LSTM and Transformer Models in Forecasting Bank Stocks in Indonesia" under the name of Penelitian Pemula Binus. The paper has been assigned the contract number 069A/VRRTT/III/2024 and was contracted on 18th March 2024.

# References

1. Zou, J., et al., *Stock market prediction via deep learning techniques: A survey.* arXiv preprint arXiv:2212.12717, 2022.

2. Mokhtari, S., K.K. Yen, and J. Liu, *Effectiveness of artificial intelligence in stock market prediction based on machine learning.* arXiv preprint arXiv:2107.01031, 2021.

3. Leung, C.K.-S., R.K. MacKinnon, and Y. Wang. *A machine learning approach for stock price prediction*.

4. Rouf, N., et al., *Stock market prediction using machine learning techniques: a decade survey on methodologies, recent developments, and future directions.* Electronics, 2021. **10**(21): p. 2717.

5. Mehtab, S., J. Sen, and A. Dutta. *Stock price prediction using machine learning and LSTM-based deep learning models*. Springer.

6. Lin, Z. *Comparative Study of LSTM and Transformer for A-Share Stock Price Prediction*. Atlantis Press.

7. Wang, Q. and Y. Yuan. *Stock Price Forecast: Comparison of LSTM, HMM, and Transformer*. Atlantis Press.

8. Tan, H.M., et al. *Comparing LSTM Models for Stock Market Prediction: A Case Study with Apple’s Historical Prices*. Springer.

9. Sonkavde, G., et al., *Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications.* International Journal of Financial Studies, 2023. **11**(3): p. 94.

10. Boyle, D. and J. Kalita, *Spatiotemporal Transformer for Stock Movement Prediction.* arXiv preprint arXiv:2305.03835, 2023.

11. Qiu, J., B. Wang, and C. Zhou, *Forecasting stock prices with long-short term memory neural network based on attention mechanism.* PloS one, 2020. **15**(1): p. e0227222.

12. Varoquaux, G. and O. Colliot, *Evaluating machine learning models and their diagnostic value.* Machine learning for brain disorders, 2023: p. 601-630.

13. Hani'ah, M., et al., *Google Trends and Technical Indicator based Machine Learning for Stock Market Prediction.* MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer, 2023. **22**(2): p. 271-284.

14. Puteri, D.I., *Implementasi Long Short Term Memory (LSTM) dan Bidirectional Long Short Term Memory (BiLSTM) Dalam Prediksi Harga Saham Syariah.* Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi, 2023. **11**(1): p. 35-43.

15. Peng, Z.-Y. and P.-C. Guo, *A data organization method for LSTM and transformer when predicting Chinese banking stock prices.* Discrete Dynamics in Nature and Society, 2022. **2022**(1): p. 7119678.

16. Hochreiter, S., *Long Short-term Memory.* Neural Computation MIT-Press, 1997.

17. Zeng, Z., et al., *Financial time series forecasting using cnn and transformer.* arXiv preprint arXiv:2304.04912, 2023.

18. Hadi, R.F., S. Sa'adah, and D. Adytia, *Forecasting of GPU Prices Using Transformer Method.* Jurnal Sisfokom (Sistem Informasi dan Komputer), 2023. **12**(1): p. 136-144.

19. Ramachandran, P., B. Zoph, and Q.V. Le, *Searching for activation functions.* arXiv preprint arXiv:1710.05941, 2017.

20. Vaswani, A., *Attention is all you need.* Advances in Neural Information Processing Systems, 2017.

21. Terven, J., et al., *Loss functions and metrics in deep learning. A review.* arXiv preprint arXiv:2307.02694, 2023.

22. Wang, C., et al., *Stock market index prediction using deep Transformer model.* Expert Systems with Applications, 2022. **208**: p. 118128.

23. Selvin, S., et al. *Stock price prediction using LSTM, RNN and CNN-sliding window model*. IEEE.

24. Nayak, P., et al. *Stock Price Prediction Using LSTM and SVM*. Springer.

25. Wang, S., *A stock price prediction method based on BiLSTM and improved transformer.* IEEE Access, 2023.