Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.6, December 2025, pp. 3575~3587
e-ISSN: 2548-6861

Optimized LSTM with TSCV for Forecasting Indonesian Bank Stocks

Rizka Mars Salsabila [1]*, Amiq Fahmi [2]*, Farrikh Al Zami[3]*
* Faculty of Computer Science, Universitas Dian Nuswantoro
112202206946@mhs.dinus.ac.id [1], amiq.fahmi@dsn.dinus.ac.id [2], alzami@dsn.dinus.ac.id [3]

Article Info

Article history:

Received 2025-09-23
Revised 2025-11-12
Accepted 2025-11-14

Keyword:

Hyperparameter Optimization, OHLCV, Long Short-Term Memory, Time Series Cross-Validation, Volatility.

ABSTRACT

Volatility in financial markets presents complex forecasting challenges for investors, particularly within emerging economies such as Indonesia. This study proposes an optimized Long Short-Term Memory (LSTM) model for forecasting the stock prices of five significant Indonesian banks: BBCA, BBRI, BMRI, BBNI, and BBTN, utilizing daily OHLCV data (Open, High, Low, Close, Volume) and technical indicators from 2020 to 2025. The dataset comprises over 6,000 daily records, segmented using a sliding window approach to preserve temporal structure and enhance learning efficiency. Concurrently, the model architecture comprising dual LSTM layers with dropout regularization was refined through systematic hyperparameter tuning to enhance predictive performance. Model evaluation employed 5-fold Time Series Cross-Validation (TSCV), a sequential validation technique that mitigates data leakage and explicitly overcomes the limitations of conventional k-fold methods by preserving chronological integrity. Performance metrics included MSE, RMSE, MAE, R^2, and MAPE. The experiment results demonstrate the model's robustness in capturing long-term dependencies within financial time series. BBCA and BMRI achieved superior accuracy ($R^2 > 0.95$), with BBCA recording the lowest MAPE of 2.34%. Despite market fluctuations, the model maintained consistent reliability across all test folds. This study overcomes a methodological limitation by integrating LSTM with TSCV in expanding markets, offering actionable insights for investors, analysts, and policymakers, and serving as a reference for adaptive AI-based, more informed forecasting tools. Moreover, the proposed framework holds promise for broader application across other financial sectors and regional markets with similar volatility characteristics.

I. Introduction

The capital market holds a vital role in a nation's financial system by mobilizing public funds and allocating them to entities needing capital [1]. In Indonesia, stock market activity has significantly expanded in recent years, marked by a sharp increase in retail investor participation and consistently high daily transaction volumes at the Indonesia Stock Exchange (IDX) [2]. However, stock price movements remain highly volatile, influenced by internal corporate performance, macroeconomic policies, government interventions, and global developments [3]. This volatility challenges investors, who rely on accurate analytical insights and reliable forecasts to make well-informed decisions [4].

Although conventional forecasting techniques are widely used and effective for many applications, they often fail to capture nonlinear relationships and long-term dependencies among explanatory variables. These limitations reduce predictive accuracy and increase exposure to investment risk [5]. Recent advances in artificial intelligence (AI), especially in machine learning (ML) and deep learning (DL), present promising alternatives. In particular, the Long Short-Term Memory (LSTM) network, an enhanced form of Recurrent Neural Networks (RNNs), has demonstrated strong capabilities in modeling sequential data and preserving temporal dependencies [6]. These characteristics make LSTM particularly suitable for capturing the dynamic behavior of stock prices [7].

Empirical research has consistently confirmed the effectiveness of LSTM-based approaches in forecasting Indonesian stock prices, particularly due to their strength in learning sequential patterns and handling non-linear market behavior. For instance, one study integrated anomaly detection with LSTM for leading banking stocks (BBCA, BBRI, BMRI) and achieved an exceptionally low Mean Squared Error (MSE) of 0.003 for BBCA, indicating strong predictive accuracy under volatile conditions [8]. Another investigation using an enhanced LSTM model reported promising results, with BBCA recording a Mean Absolute Percentage Error (MAPE) of 0.0099 and a Root Mean Squared Error (RMSE) of 128.02 [9]. In contrast, research on the mining sector employing ETSFormer with Time Series Cross-Validation (TSCV) revealed difficulties in handling short-term price fluctuations [10]. Similarly, studies conducted during the COVID-19 pandemic showed inconsistent RMSE values for stocks such as BBCA and BBRI [11]. Previous studies have shown that the ARIMA model applied to the LQ45 index produced an average MAPE of 10.09%, with considerable variation in performance across different stocks, where BBCA recorded the lowest error of 2.18% [12]. Meanwhile, the study conducted by Beno Jange (2022) using the XGBoost algorithm to predict BBCA stock prices demonstrated better performance, achieving a MAPE of 4.01% [13].

Despite promising outcomes reported in previous studies, important methodological limitations and unexplored areas remain open and unaddressed [14]. As far as the literature reveals, the existing literature has yet to examine the optimization of LSTM hyperparameter combinations for forecasting Indonesian banking stocks using TSCV [15]. This validation technique is particularly well-suited for sequential data, as it mitigates overfitting (data leakage) and enables a more reliable assessment of model performance [16].

The present study addresses this methodological shortcoming by constructing and validating an optimized LSTM-based framework to refine the accuracy of stock price prediction for five major Indonesian banks: BBCA, BBRI, BMRI, BBNI, and BBTN. In addition to hyperparameter tuning to identify the best model configuration, this research compares predictive performance across the selected banks. The study contributes to the growing literature on deep learning-based financial forecasting and offers practical guidance for investors, market analysts, and policymakers. It also serves as a reference for technology developers seeking to build AI-based forecasting tools that are adaptive to dynamic market conditions and relevant to emerging financial markets.

II. METHODS

This section presents a systematic overview of the research stages, as illustrated in Figure 1, highlighting the methodological approach followed throughout the study.

This study adopts a quantitative research design, utilizing advanced deep learning techniques, specifically the Long Short-Term Memory (LSTM) architecture, to forecast stock price movements within Indonesia's banking sector. The methodological workflow is systematically illustrated through a process diagram, comprising key stages: data collection, cleansing and preprocessing, feature engineering, model development via hyperparameter optimization, and performance evaluation using a time series split cross-validation strategy.

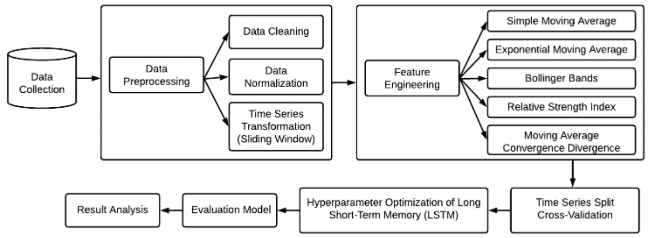

Figure 1. Research Flow Diagram

A. *Data Collection*

This study utilizes daily stock price time series obtained from prominent publicly traded companies banking stocks listed on the Indonesia Stock Exchange (IDX) [17]. The analysis focuses on five major state-owned banks: PT Bank Central Asia Tbk (BBCA), PT Bank Rakyat Indonesia Tbk (BBRI), PT Bank Mandiri Tbk (BMRI), PT Bank Negara Indonesia Tbk (BBNI), and PT Bank Tabungan Negara Tbk (BBTN). Historical data were retrieved from Yahoo Finance using the yfinance Python library, covering the period from January 1, 2020, to September 1, 2025. The dataset includes daily OHLCV (Open, High, Low, Close, and Volume) values essential for capturing market dynamics. During the observation period, 6,670 equities were actively traded on the IDX, with the selected banking stocks serving as the primary objects of analysis.

B. *Data Preprocessing*

The data preprocessing procedure in this study comprises three primary stages:

1) *Data Cleansing*: The dataset is carefully examined to detect and remove missing values and duplicate entries [18]. This step is critical to prevent distortions and biases that may compromise the accuracy and reliability of model training.

2) *Data Normalization:* All numerical features are preprocessed through Min-Max normalization, effectively transforming the raw values into a normalized range between 0 and 1, thereby preserving relative relationships while preventing the dominance of variables with larger magnitudes. This transformation facilitates faster convergence during model training and enhances

computational efficiency [19]. Normalization is performed using Eq. (1).

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

Where:
X'= Normalized value
X= Original value before normalization
Xmin= Minimum value variable X
Xmax= Maximum value variable X

3) *Time Series Transformation (Sliding Window):* A sliding window approach is employed to model temporal dependencies within the historical dataset. This technique divides the time series into overlapping sequences of a predetermined length, enabling the model to extract insights from recent trends and forecast subsequent stock prices [7]. By organizing the input data in this manner, the LSTM model is able to effectively capture sequential patterns and short-term market fluctuations [18].

C. Feature Engineering

To achieve higher predictive fidelity and improve the generalization capacity of the model, a set of widely adopted technical indicators is integrated into the dataset during the feature engineering phase. These indicators include the Simple Moving Average (SMA), Exponential Moving Average (EMA), Relative Strength Index (RSI), Bollinger Bands (BB), and Moving Average Convergence Divergence (MACD) [20]. Each indicator reflects a specific dimension of market behavior, including trend direction, momentum, and volatility, thereby supplying the model with enhanced input features that facilitate learning of complex patterns in stock price dynamics.

1) *Simple Moving Average (SMA):* In technical analysis, the SMA is a fundamental tool used to smooth short-term volatility and reveal long-term price trends. It is computed by averaging a specified number of past closing prices over a defined time window, as shown in Equation (2).

$$SMA_t = \frac{1}{N} \sum_{i=t-N+1}^{t} Close_i \quad (2)$$

Where:
SMA_t= Moving average value at time -t
N= Number of observation periods
$Close_i$= Closing price at time -i

2) *Exponential Moving Average (EMA):* Building on the SMA, the EMA incorporates an exponential weighting mechanism that emphasizes recent observations. This approach enhances its sensitivity to rapid price movements, rendering it a useful tool for analyzing markets with significant short-term volatility. The calculation of the EMA is presented in Equation (3).

$$EMA_t = \alpha \cdot Close_t + (1-\alpha) \cdot EMA_{t-1} \quad (3)$$

Where:
α= Smoothing or weighting factor
$Close_t$= Closing price at period -t
EMA_{t-1}= EMA in the previous period

3) *Relative Strength Index (RSI):* The Relative Strength Index (RSI) operates as a momentum oscillator designed to measure both the speed and amplitude of price fluctuations. By identifying periods when market conditions approach overbought or oversold thresholds, the RSI provides insights into potential price reversals and allows for assessment of the durability of prevailing market trends. The RSI is calculated using the formula presented in Equation (4).

$$RSI = 100 - \frac{100}{1+RS}, \quad RS = \frac{\text{Average Gain}}{\text{Average Loss}} \quad (4)$$

Where:
Average Gain= Average price increase over
Average Loss= Average price decrease over

4) *Bollinger Bands (BB):* In technical analysis, Bollinger Bands (BB) are employed as volatility indicators, consisting of two bands plotted at a fixed distance from a central moving average. The upper band reflects the moving average plus a specified multiple of the standard deviation, whereas the lower band represents the equivalent deviation below. Such a configuration enables analysts to discern shifts in market volatility and to anticipate possible reversals or breakout events. The mathematical formulation of Bollinger Bands is presented in Eq. (5).

$$UpperBand_t = SMA_t + k \cdot \sigma_t, \quad (5)$$
$$LowerBand_t = SMA_t - k \cdot \sigma_t$$

Where:
SMA_t= Simple moving average at time t
σ_t= Standard deviation of price at time t
k= Factor applied to the standard deviation

5) *Moving Average Convergence Divergence (MACD):* The MACD is a trend-following momentum indicator derived from the difference between two Exponential Moving Averages (EMAs) calculated over distinct timeframes, typically a short-term and a long-term period. To enhance its interpretability, a Signal Line, defined as the exponential moving average (EMA) of the MACD values, is employed to detect potential trend reversals and to validate shifts in market momentum. The mathematical formulation of MACD and its Signal Line is presented in Equation (6).

$$MACD_t = EMA_{short}(t) - EMA_{long}(t) \quad (6)$$

Where:
EMA_{short}= Exponential moving average short period
EMA_{long}= Exponential moving average long period

D. Time Series Split Cross-Validation

In recognition of the sequential dependencies that characterize financial market data, this study employs Time Series Split Cross-Validation (TSCV). Unlike conventional k-fold cross-validation, TSCV preserves the temporal order of observations, making it particularly suitable for time-dependent datasets. In this approach, the data is partitioned chronologically, This approach ensures that training is conducted on earlier segments of the series and validation on later portions, thereby strengthening the model's capacity to learn temporal dependencies and yielding a more rigorous and realistic assessment of its predictive performance [21].

E. Long Short-Term Memory Modeling

The predictive framework uses a sequential deep learning architecture comprising two stacked LSTM layersThe architecture specifies the number of hidden units per LSTM layer to optimize the extraction of temporal relationships within financial time series. Furthermore, dropout regularization is incorporated between layers to minimize overfitting tendencies and enhance the model's generalization performance. The final output is generated through a fully connected dense layer, transforming the learned temporal representations into the target prediction.

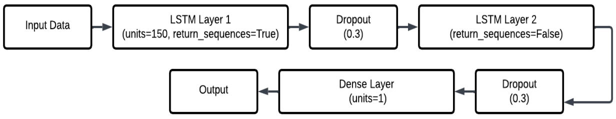

Figure 2. LSTM Model Architecture

The model processes normalized historical stock price data structured within a predefined lookback window. Input sequences are initially passed through an LSTM layer, followed by dropout regularization to address overfitting. It is succeeded by a second LSTM layer and an additional dropout layer, enhancing the model's generalization capability. The final output is produced through a fully connected dense layer with a single neuron responsible for generating the predicted stock price. Model training uses the Adam optimization algorithm, with MSE as the loss function to evaluate regression performance [6].

F. Hyperparameter Optimization

The process of selecting and fine-tuning hyperparameters is fundamental to enhancing the performance of an LSTM model, as these parameters govern both the learning behavior and the predictive capability of the network. In this study, particular attention is directed toward optimizing critical hyperparameters including the number of LSTM units, the lookback window length, the dropout rate, the batch size, and the total number of training epochs. Systematic experimentation and validation are conducted to identify the optimal configuration that yields the most robust and generalizable predictive outcomes [22].

1) Number of LSTM Units: The number of LSTM units, defined as the hidden neurons within each layer, governs the model's capacity to capture temporal dynamics. Although increasing the units allows the network to learn more complex structures, it may also induce overfitting. To balance accuracy with generalization, this study employs a data-driven approach to hyperparameter tuning instead of predetermined settings.

2) Lookback Window Length: The lookback window length defines the span of historical data the model utilizes to generate future price predictions. It is critical to select an appropriate window size, as it influences the model's ability to capture relevant temporal patterns. A longer window may incorporate broader market trends, but risks diluting short-term signals and increasing computational overhead due to redundant information. Conversely, a shorter window may enhance responsiveness to recent fluctuations but potentially overlook meaningful long-term dependencies. Therefore, this study emphasizes a balanced configuration to ensure predictive precision and computational efficiency.

3) Dropout Rate: Dropout is a regularization method employed to mitigate overfitting in neural network training. During each iteration, a specified fraction of neurons is randomly deactivated according to a predetermined probability, which serves to reduce co-adaptation among network units and promote more robust feature learning. The adopted mechanism encourages the model to learn varied and stable feature encodings rather than depending on a limited set of pathways. This reduction in pathway-specific reliance mitigates overfitting and leads to demonstrably better generalization when evaluated on out-of-sample data. The dropout rate must be carefully tuned to balance model complexity and predictive stability.

4) Batch Size: Refers to the quantity of training instances employed to fit the model propagated through the network during each feedforward and backpropagation cycle. Smaller batch sizes introduce higher gradient variance, enabling the model to adapt more rapidly to fluctuations in the data, which may be beneficial in volatile environments. However, larger batch sizes tend to produce more stable and consistent gradient updates, improving convergence reliability at the cost of increased computational demand. This study explores a range of batch sizes to identify an optimal trade-off between training stability and responsiveness.

5) Number of Epochs: In machine learning contexts, an epoch denotes a full traversal of the training dataset, typically segmented into batches to support iterative parameter updates. An inadequate epoch configuration can result in underfitting, indicating insufficient learning of data regularities, whereas an excessive number may lead to overfitting, whereby the model achieves high training accuracy but performs inadequately on unseen data.

Consequently, the choice of epoch number should be optimized with reference to convergence dynamics, commonly observed through validation loss trajectories across training cycles.

G. Model Performance Evaluation

The proposed model's performance is assessed using a comprehensive set of error metrics commonly employed in time-series forecasting. These include Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), and Mean Absolute Percentage Error (MAPE). Each metric provides a distinct perspective on the model's predictive accuracy, robustness, and generalization capability, thereby enabling a multidimensional evaluation of its effectiveness in forecasting stock price movements [9].

1) *Mean Squared Error (MSE):* The Mean Squared Error (MSE) serves as a fundamental criterion in regression modeling, calculated as the average of squared deviations between predicted outputs and observed data. Its formulation disproportionately penalizes larger errors, thus ensuring that significant deviations are emphasized in the evaluation process. Models with lower MSE values are generally regarded as demonstrating stronger accuracy and generalization ability [23], as shown in Eq. (7).

$$\text{MSE}=\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2 \qquad (7)$$

Where:
γ_i = Observed value for the i-th observation
$\hat{\gamma}_i$ = Predicted value for the i-th observation
n= Total number of observations

2) *Root Mean Squared Error (RMSE):* RMSE measures the square root of the average squared deviations between predicted and observed values, providing an indicator of prediction accuracy in the data's original scale. Its primary advantage lies in its unit consistency with the target variable, facilitating intuitive interpretation, particularly in financial forecasting contexts such as stock price prediction. A lower RMSE value reflects higher predictive accuracy and indicates that the model's outputs closely approximate the observed data [23], as shown in Eq. (8).

$$\text{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}\left(y_i-\hat{y}_i\right)^2} \qquad (8)$$

Where:
γ_i = Observed value for the i-th observation
$\hat{\gamma}_i$ = Predicted value for the i-th observation
n= Total number of observations

3) *Mean Absolute Error (MAE):* MAE represents the mean of the absolute differences between predicted and actual values, providing a direction-independent assessment of prediction error. Unlike metrics that square deviations, MAE is less sensitive to extreme outliers, providing a robust assessment of overall predictive accuracy. Lower MAE values reflect improved model precision and greater consistency in forecasting performance [23], as formulated in Eq. (9).

$$\text{MAE}=\frac{1}{n}\sum_{i=1}^{n}\left|y_i-\hat{y}_i\right| \qquad (9)$$

Where:
γ_i = Observed value for the i-th observation
$\hat{\gamma}_i$ = Predicted value for the i-th observation
n= Total number of observations

4) *Coefficient of Determination (R^2):* The R^2 score quantifies the extent to which the variability in the observed data can be accounted for by the model's predictions. It indicates the model's explanatory power, with values ranging from 0 to 1. An R^2 coefficient approaching unity denotes a pronounced alignment between model predictions and empirical observations, thereby evidencing the model's efficacy in capturing and explaining the intrinsic variability and structure of the dataset. This metric is beneficial for assessing goodness-of-fit in regression tasks [23], as formulated in Eq. (10).

$$.R^2=1-\frac{\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2} \qquad (10)$$

Where:
γ_i = Observed value for the i-th observation
$\hat{\gamma}_i$ = Predicted value for the i-th observation
n= Total number of observations

5) *Mean Absolute Percentage Error (MAPE):* Expressed in percentage form, this metric provides a scale-free evaluation of forecasting error, making it highly interpretable across diverse contexts. It is obtained by computing the average of the absolute differences between predicted and observed values, subsequently normalized by the corresponding actual observations. As such, it facilitates robust comparisons of predictive accuracy across datasets that differ in units or magnitude. A lower MAPE value reflects superior predictive performance, while higher values indicate greater divergence between model outputs and observed data [23], as formulated in Eq. (11).

$$\text{MAPE}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\gamma_i-\hat{\gamma}_i}{\gamma_i}\right|\times 100\% \qquad (11)$$

Where:
γ_i = Observed value for the i-th observation
$\hat{\gamma}_i$ = Predicted value for the i-th observation
n= Total number of observations

III. Result and Discussion

This section presents the empirical results derived from the processed dataset, which form the foundation for subsequent analysis and interpretation. The model's predictive performance is assessed through previously defined evaluation metrics, allowing for a thorough and systematic appraisal of its forecasting accuracy. Comparative insights are drawn to highlight the effectiveness of the proposed architecture, and the implications of hyperparameter configurations are discussed in relation to forecasting accuracy, generalization, and computational efficiency.

To ensure the robustness of the proposed predictive framework, the methodological design was deliberately structured to capture variations observed during both stable and volatile market environments. The dataset utilized spans several consecutive years of trading observations, notably including the COVID-19 period (2020), which represented a phase of substantial market turbulence and irregular behavior. Although the research did not implement a separate, predefined stress-testing mechanism, the inclusion of data from such unstable intervals effectively allowed the model to learn and generalize from diverse market dynamics. In addition, the application of Time-Series Cross-Validation (TSCV) provided a systematic approach for assessing model performance across different temporal partitions, thereby reinforcing the model's robustness against abrupt structural and behavioral shifts within financial markets. Looking ahead, future studies may refine this approach by incorporating dedicated stress-testing analyses under specific macroeconomic or sectoral disruptions, such as monetary tightening episodes or sudden liquidity shocks, to further evaluate the model's resilience and adaptive capacity.

A. Dataset Description

This study utilizes historical stock price data from five major Indonesian banking institutions, BBCA, BBRI, BMRI, BBNI and BBTN from January 1, 2020, to September, 2025.

The selection of BBCA, BBRI, BMRI, BBNI, and BBTN was based on their economic significance and representativeness within Indonesia's banking sector. These institutions are among the largest and most liquid constituents of the IDX Financials Index, accounting for a substantial share of market capitalization and trading activity. Their inclusion captures both systemic stability and market diversity, as they differ in terms of ownership structure, asset scale, and risk exposure, making them suitable benchmarks for comparative forecasting performance within the national financial market.

B. Data Sanitization and Feature Development

Before analysis, the raw dataset exhibited missing values and duplicate records, necessitating a rigorous data cleansing process. Subsequently, Feature engineering techniques were employed to improve the dataset's ability to support accurate predictions. The input variables comprised standard OHLCV (Open, High, Low, Close, Volume) data augmented by a set of technical indicators, including the Simple Moving Average (SMA), Exponential Moving Average (EMA), Bollinger Bands, Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and the MACD Signal Line.

A sliding window technique was employed to capture temporal dependencies with a retrospective duration of 90 days. This transformation yielded a final dataset of 1,244 samples, each representing a 90-day sequence and encompassing 13 input features. Summary tables detailing the raw data characteristics for each bank are presented in the subsequent section to provide contextual grounding for model development and evaluation.

As shown in Table I – Table V, the observation period for all five banking stocks was standardized to span from January 2020 to September 2025, ensuring uniformity in the number of trading days across samples. Although minor discrepancies exist due to temporary market suspensions and variations in liquidity, these differences were statistically insignificant and did not affect the training consistency of the model. Consequently, each banks dataset contains approximately the same volume of daily records, allowing for a fair comparison of forecasting accuracy and model robustness across the selected institutions.

TABLE I
BBCA Stock

Date	Open	High	Low	Close	Volume
2020-01-02 00:00	5825.27	5899.24	5812.23	5820.93	49,445,000
2020-01-03 00:00	5873.13	5916.64	5851.38	5916.64	47,755,500
2020-01-06 00:00	5847.02	5873.13	5820.93	5860.08	27,300,000

TABLE II
BBRI Stock

Date	Open	High	Low	Close	Volume
2020-01-02 00:00	2954.13	2960.84	2927.27	2960.84	45,886,302
2020-01-03 00:00	2967.56	2980.99	2947.42	2967.56	91,189,705
2020-01-06 00:00	2927.27	2947.42	2900.42	2933.99	48,648,450

The model's performance was evaluated using a 5-fold TSCV strategy. This validation technique preserves the temporal structure of the dataset by ensuring that training is consistently performed on earlier observations, while testing is conducted on subsequent data points. Such chronological integrity is essential in time-series forecasting, as it prevents data leakage and ensures that future information does not inadvertently influence the training process. Compared to conventional k-fold cross-validation, TSCV offers a more robust and realistic assessment of predictive performance in sequential data environments.

TABLE III
BMRI STOCK

Date	Open	High	Low	Close	Volume
2020-01-02 00:00	2772.93	2809.06	2763.90	2800.03	37,379,800
2020-01-03 00:00	2800.03	2827.12	2754.86	2790.99	70,294,600
2020-01-06 00:00	2763.90	2772.93	2718.73	2745.83	61,892,000

TABLE IV
BBNI STOCK

Date	Open	High	Low	Close	Volume
2020-01-02 00:00	3127.14	3127.14	3087.43	3087.43	18,602,600
2020-01-03 00:00	3097.36	3127.14	3037.80	3097.36	32,251,400
2020-01-06 00:00	3077.50	3077.50	3017.94	3027.87	26,249,200

TABLE V
BBTN STOCK

Date	Open	High	Low	Close	Volume
2020-01-02 00:00	3087.43	3127.14	3087.43	1631.68	6116029
2020-01-03 00:00	3097.36	3127.14	3037.80	1647.00	32,251,400
2020-01-06 00:00	3027.87	3077.50	3017.94	1631.68	26,249,200

C. Hyperparameter Optimization

Prior to finalizing the model configuration, a Random Search approach was employed to identify optimal hyperparameter settings. This method enables efficient exploration of the hyperparameter space by randomly sampling combinations, thereby reducing computational burden compared to exhaustive grid search. The results indicated that LSTM unit sizes of [50, 100, 150] consistently yielded superior performance across validation folds, outperforming alternative configurations. Accordingly, this setup was adopted as the core architecture for subsequent experiments. The model design was carefully structured to establish a robust foundation for comparative analysis and performance benchmarking.

D. Comparative Analysis

This study employs an LSTM architecture combined with TSCV to evaluate the predictive performance across five major banking stocks listed on the Indonesia Stock Exchange (IDX): BBCA, BBRI, BMRI, BBNI, and BBTN. The TSCV framework ensures temporal integrity by validating the model across sequential data splits, thereby enhancing the reliability of performance assessment. The data splitting procedure in this study followed a fully sequential approach. Model training and validation were performed using Time Series Cross-Validation (TSCV) with expanding folds, ensuring that each fold respected the temporal ordering of the data. No random partitioning was used at any stage. After completing the TSCV process, a separate chronological holdout set was used solely for the final out-of-sample test to evaluate the model's generalization performance.

1) *BBCA Stock*

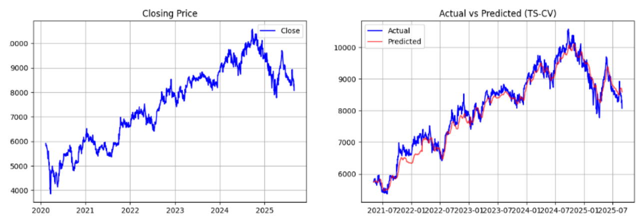

Figure 3. BBCA Stock Chart

Figure 3 illustrates the BBCA plot exhibits a clear upward trajectory across the 2020–2025 period, with comparatively mild intra-period fluctuations. The prediction line closely tracks the observed price series for most of the timeline, indicating that the model captures the stock's dominant trend and short-term momentum with a high degree of visual fidelity. Only at a few sharp inflection points does the forecast diverge noticeably from the actual series, suggesting a limited lag in reacting to abrupt market moves but strong alignment for medium- to long-term behavior.

The evaluation results presented in Table VI indicate that the BBCA prediction model exhibits a high level of consistency and reliability, as reflected by an average R^2 value of 0.9631. Although Fold 4 and Fold 5 show slightly lower R^2 scores compared to the other folds, the model overall maintains a commendable degree of accuracy, with a mean MAPE of 2.34%. Such a low error margin demonstrates the model's robustness in capturing the underlying patterns of stock price movements. These findings suggest that the model performs effectively across varying data partitions, maintaining stability even under fluctuating market conditions.

TABLE VI
BBCA FOLD EVALUATION

Fold	MSE	RMSE	MAE	R^2	MAPE
1	58,685.47	242.25	193.40	0.8312	0.02949
2	48,941.63	221.23	182.42	0.8082	0.02428
3	15,570.89	124.78	97.01	0.8466	0.01178
4	68,098.28	260.96	220.21	0.7430	0.02304
5	93,566.94	305.89	251.17	0.7624	0.02817
Mean	**56,972.64**	**238.69**	**188.84**	**0.9631**	**0.0234**

TABLE VII
COMPARASION OF ACTUAL AND PREDICTED BBCA

Date	Actual	Predicted	Price Difference
2021-05-07	5,761.67	5,757.83	3.85
2021-05-10	5,779.68	5,767.46	12.22
2021-05-11	5,833.70	5,776.96	56.73
2021-05-17	5,851.70	5,791.03	60.68
2021-05-18	5,752.67	5,806.30	-53.62

As shown in Table VII, the predicted BBCA stock prices closely align with the actual values, indicating strong model performance. Most deviations remain relatively small, suggesting consistent accuracy in stable market conditions. Minor fluctuations, such as on May 17 and May 18, reflect the models sensitivity to short-term market movements. Overall, the model effectively tracks price trends with limited prediction error.

TABLE VIII
HOLDOUT (CHRONOLOGICAL) EVALUATION FOR BBCA

MSE	RMSE	MAE	R^2	MAPE
87617.84	296.00	253.33	0.692	0.0288

As shown in Table VIII, the holdout evaluation provides an assessment of the model's performance on out-of-sample data, which was not used during training. The results indicate that the model successfully captures long-term price trends of BBCA, though prediction errors increase in periods of high volatility. An R^2 value of 0.692 suggests that approximately 69.2% of the price variance is explained by the model. Moreover, a MAPE of 2.88% confirms a reasonably accurate predictive performance.

In conclusion, the combination of TS-CV and holdout evaluations highlights the robustness of the LSTM model in forecasting BBCA stock prices, showing optimal performance for medium-term trends while demonstrating limited responsiveness to abrupt price spikes.

Figure 4 illustrates the SHAP analysis reveals that the Signal, BB_upper, and Low features exert the strongest influence on the BBCA stock price prediction model.

As depicted in Figure 5, the historical closing price of BBRI has exhibited a steady upward trajectory accompanied by pronounced volatility in recent years. The right panel of the figure compares actual price movements and those predicted by the LSTM model, revealing that the model effectively captures the overarching trend. While discrepancies are observed during abrupt price surges or declines, the predicted curve generally aligns with the actual data. These findings suggest that the model demonstrates strong capability in modeling long-term temporal patterns, albeit with limited responsiveness to short-term fluctuations in stock prices.

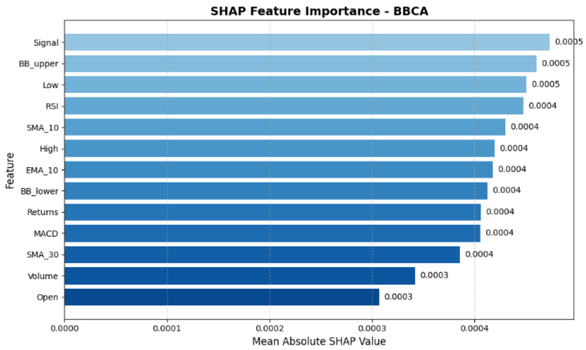

Figure 4. SHAP BBCA

2) *BBRI Stock*

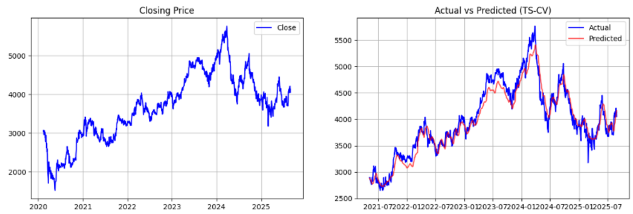

Figure 5. BBRI Stock Chart

TABLE IX
BBRI FOLD EVALUATION

Fold	MSE	RMSE	MAE	R^2	MAPE
1	18,082.13	134.47	109.77	0.7553	0.0344
2	12,761.97	112.97	87.98	0.6161	0.0234
3	8,935.91	94.53	72.92	0.9315	0.0163
4	55,817.10	236.26	201.25	0.7360	0.0420
5	18,081.02	134.47	109.52	0.7382	0.0282
Mean	**22,735.62**	**150.78**	**116.29**	**0.9472**	**0.0289**

TABLE X
COMPARASION OF ACTUAL AND PREDICTED BBRI

Date	Actual	Predicted	Price Difference
2021-05-07	2,893.35	2,905.66	-12.31
2021-05-10	2,893.35	2,889.17	4.18
2021-05-11	2,850.49	2,872.97	-22.48
2021-05-17	2,786.19	2,859.57	-73.39
2021-05-18	2,786.19	2,848.27	-62.08

As presented in Table IX, the BBRI model demonstrates fluctuating predictive performance across folds. The lowest errors are recorded in Fold 3, indicating strong model accuracy, while Fold 4 shows the highest error values, likely due to increased market volatility. Despite these variations, the model maintains a low mean MAPE of 2.89%, signifying good overall predictive stability. Hence, the model is effective in capturing the general trend of BBRI's stock prices across different time segments.

BBRI stock yielded reliable forecast outcomes, as evidenced by the relatively low mean price difference reported in Table X. The model demonstrated strong capability in capturing short-term market movements, effectively aligning with transient fluctuations in the stock's trajectory. However, compared to BBCA, the predictive outputs for BBRI exhibited greater volatility, indicating a higher degree of sensitivity to abrupt market changes. It suggests that while the model performs well in identifying immediate trends, its stability may vary depending on the underlying asset's volatility profile.

TABLE XI
HOLDOUT (CHRONOLOGICAL) EVALUATION FOR BBRI

MSE	RMSE	MAE	R^2	MAPE
14,826.57	121.76	96.80	0.6747	0.0253

As shown in Table IX, the holdout evaluation for the BBRI model demonstrates satisfactory predictive accuracy. The model achieves an R^2 of 0.6747, indicating that it explains a substantial portion of the variance in actual stock prices. With an RMSE of 121.76 and a MAPE of 2.53%, the prediction errors remain relatively low. Overall, the model performs reliably in forecasting unseen (out-of-sample) data within acceptable error margins.

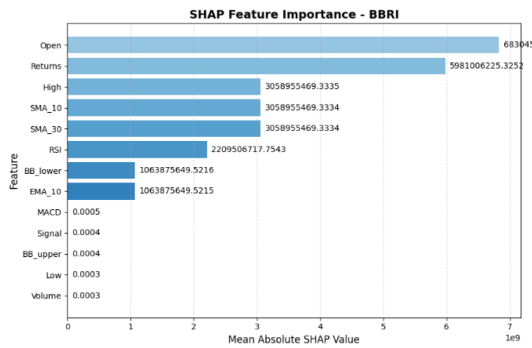

Figure 6. SHAP BBRI

Figure 6 illustrates the SHAP analysis for the BBRI stock prediction model indicates that the Open and Returns features have the highest impact on model performance. These variables significantly influence the prediction results, suggesting that the model heavily relies on recent price movements and daily opening prices.

3) *BMRI Stock*

As illustrated in Figure 7, the model's predictions for BMRI exhibit a consistently stable trend, particularly during consolidation phases where price movements are gradual and range-bound. The LSTM architecture accurately captures these slow dynamics, aligning closely with actual price behavior. However, its responsiveness diminishes during abrupt upward shifts, indicating a lag in adapting to rapid market changes. It suggests that the model is better suited for forecasting steady price trajectories rather than reacting to short-term volatility.

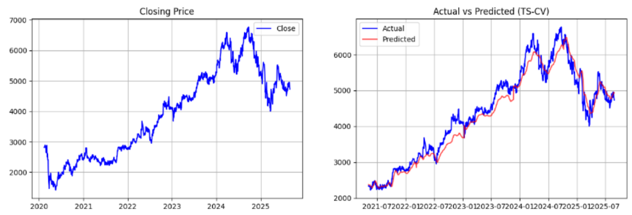

Figure 7. BMRI Stock Chart

TABLE XII
BMRI FOLD EVALUATION

Fold	MSE	RMSE	MAE	R^2	MAPE
1	11,390.45	106.73	87.34	0.8433	0.0325
2	46,496.48	215.63	179.24	0.7184	0.0471
3	26,117.50	161.61	133.36	0.8489	0.0280
4	79,057.48	281.17	239.91	0.6578	0.0398
5	78,551.25	280.27	230.28	0.7648	0.0468
Mean	**48,322.63**	**219.82**	**174.02**	**0.9681**	**0.0388**

Table XII indicates that BMRI demonstrates noticeable variation in predictive performance across the five folds. Fold 1 shows solid accuracy with a MAPE of 3.25%, and Fold 3 records the lowest error at 2.80%. In contrast, Fold 2 and Fold 5 display higher error levels, with MAPE values of 4.71% and 4.68%, respectively. This variation suggests that although the model is capable of capturing BMRI's overall price movement patterns, its sensitivity to differing market conditions remains evident. While the model performs well in tracking long-term trends, its ability to handle short-term fluctuations or periods of increased volatility appears less robust.

TABLE XIII
COMPARASION OF ACTUAL AND PREDICTED BMRI

Date	Actual	Predicted	Price Difference
2021-05-07	2,335.32	2,376.20	-40.87
2021-05-10	2,345.14	2,370.88	-25.74
2021-05-11	2,315.70	2,366.94	-51.24
2021-05-17	2,325.51	2,362.74	-37.23
2021-05-18	2,315.70	2,359.69	-43.99

Table XIII shows the prediction results for BMRI, highlighting the model's relatively consistent performance during periods of price consolidation. Across the selected dates, the predicted values remain close to the actual prices, with only moderate deviations observed. This indicates that the model is capable of capturing gradual, stable movements in the market. However, when prices shift more abruptly, the gap between predicted and actual values tends to widen, reflecting the model's reduced responsiveness to sudden market changes. While the LSTM architecture effectively learns long-term temporal patterns, its ability to adapt to short-term volatility remains limited.

TABLE XIV
HOLDOUT (CHRONOLOGICAL) EVALUATION FOR BMRI

MSE	RMSE	MAE	R^2	MAPE
99,896.04	316.06	265.50	0.4925	0.0545

As shown in Table XIV, the BMRI model exhibits moderate predictive performance, with an R^2 value of 0.4925, indicating that nearly half of the price variance is explained by the model. The relatively high RMSE and MAPE (5.45%) suggest that prediction errors increase under volatile market conditions. Nonetheless, the model captures the general stock price direction adequately.

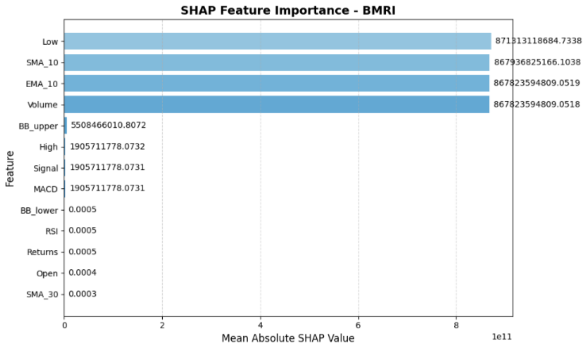

Figure 8. SHAP BMRI

As illustrated in Figure 8 the SHAP analysis for the BMRI stock prediction model shows that the Low, SMA_10, EMA_10, and Volume features have the highest influence on model performance.

4) *BBNI Stock*

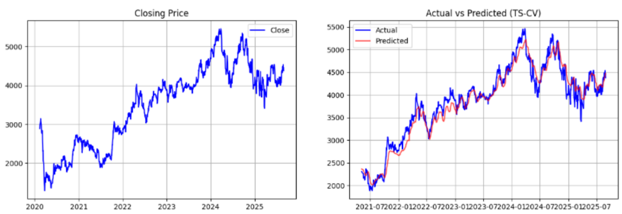

Figure 9. BBNI Stock Chart

Figure 9 overlays actual BBNI prices with model predictions, revealing generally strong correspondence between the two series. Minor deviations occur during periods of elevated volatility, yet predicted values consistently reflect the stocks directional movement. This observation, combined with the fold-wise evaluation metrics, suggests that the model can produce reliable forecasts under typical market conditions, albeit with reduced precision during extreme fluctuations.

TABLE XV
BBNI FOLD EVALUATION

Fold	MSE	RMSE	MAE	R^2	MAPE
1	11,202.48	105.84	81.30	0.9383	0.0316
2	28,535.08	168.92	144.15	0.5499	0.0389
3	6,152.37	78.44	59.97	0.8762	0.0148
4	44,739.10	211.52	173.42	0.6478	0.0362
5	40,080.94	200.20	156.07	0.6429	0.0378
Mean	**26,141.99**	**161.68**	**122.98**	**0.9610**	**0.0319**

Table XV indicates that BBNI's predictive performance varies across folds, with error metrics showing a moderately wide range. Fold 3 demonstrates the highest accuracy, reflected by the lowest MAPE of 1.48% and the smallest RMSE of 78.44. In contrast, Folds 4 and 5 record higher error levels, with RMSE values exceeding 200 and MAPE values around 3.6–3.8%. Despite this variation, the model maintains generally stable performance, with mean MAPE remaining relatively low at 3.19%. These results suggest that while the model is capable of generating reliable forecasts for BBNI, its accuracy may still be influenced by fluctuations in market conditions, especially during periods of elevated volatility.

TABLE XVI
COMPARASION OF ACTUAL AND PREDICTED BBNI

Date	Actual	Predicted	Price Difference
2021-05-07	2,308.60	2,333.32	-24.71
2021-05-10	2,298.30	2,324.04	-25.74
2021-05-11	2,277.68	2,313.08	-35.39
2021-05-17	2,267.38	2,303.59	-36.21
2021-05-18	2,246.77	2,300.44	-53.68

Table XVI shows a strong alignment between the actual and predicted BBNI prices, with only moderate deviations observed across the selected dates. The predicted values consistently remain slightly above the actual prices, resulting in relatively small price differences that indicate stable model performance. While discrepancies may become more pronounced during periods of heightened market volatility, the LSTM model still succeeds in capturing the overall direction and underlying trend of the stock's movement. The consistently low error metrics further support the model's reliability in generating accurate forecasts under normal market conditions.

TABLE XVII
HOLDOUT (CHRONOLOGICAL) EVALUATION BBNI

MSE	RMSE	MAE	R^2	MAPE
20,946.89	144.73	114.01	0.6438	0.0278

As presented in Table XVII, the BBNI model achieves solid predictive accuracy with an R^2 of 0.6438, indicating strong explanatory power. The low MAPE (2.78%) reflects that the predicted prices closely align with actual values. This demonstrates the model's reliability in capturing stock price movements with reasonable precision.

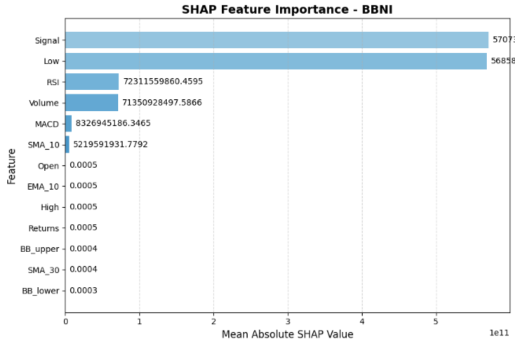

Figure 8. SHAP BBNI

Figure 8 shows the SHAP analysis for the BBTN stock prediction model indicates that the Signal and Low features have the most significant impact on the model's output.

5) *BBTN Stock*

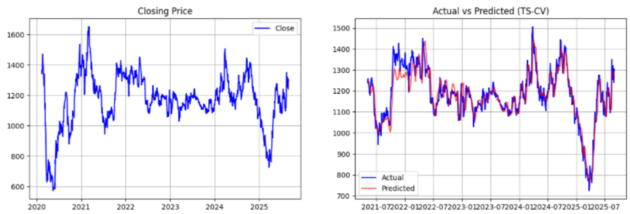

Figure 9. BBTN Stock Chart

Figure 9 presents the prediction outcomes for BBTN, which exhibit slightly lower accuracy than the other banking stocks analyzed in this study. This performance discrepancy is likely attributable to BBTN's elevated volatility, which poses greater challenges for temporal modeling. Despite these limitations, the LSTM model successfully captures the overarching trend, indicating its capacity to generalize long-term price movements even under unstable market conditions.

Table XVIII shows that BBTN's predictive performance varies across folds, with RMSE ranging from 31.10 to 55.84. Despite this variation, all MAPE values remain below 3.9%, indicating that the model maintains a satisfactory level of accuracy. Overall, the results suggest that the model reliably captures medium-term price movements, even amid shifts in market conditions.

Table XIX shows that although noticeable gaps exist between the actual and predicted BBTN prices, the model still captures the overall trend. Larger deviations appear during more volatile periods, indicating reduced responsiveness to sudden price changes, yet the general trajectory remains accurately reflected.

TABLE XVIII
BBTN FOLD EVALUATION

Fold	MSE	RMSE	MAE	R^2	MAPE
1	3,118.34	55.84	45.78	0.8225	0.0386
2	2,299.47	47.95	36.98	0.6620	0.0305
3	966.96	31.10	23.73	0.4366	0.0209
4	1,529.48	39.11	27.95	0.8081	0.0220
5	1,455.71	38.15	30.09	0.9501	0.0279
Mean	**1,873.99**	**43.29**	**32.91**	**0.8773**	**0.0280**

TABLE XIX
COMPARASION OF ACTUAL AND PREDICTED BBTN

Date	Actual	Predicted	Price Difference
2021-05-07	1,239.76	1,272.75	-32.99
2021-05-10	1,258.95	1,272.50	-13.55
2021-05-11	1,251.28	1,273.00	-21.72
2021-05-17	1,216.73	1,273.44	-56.71
2021-05-18	1,209.06	1,269.83	-60.78

TABLE XX
HOLDOUT (CHRONOLOGICAL) EVALUATION BBTN

MSE	RMSE	MAE	R^2	MAPE
1,397.41	37.38	28.88	0.9419	0.0274

As indicated in Table XX, the BBTN model performs exceptionally well with an R^2 of 0.9419, meaning it explains over 94% of the variance in actual stock prices. The small error values (RMSE, MAE, and MAPE) confirm the model's high precision and stability. Overall, this model provides the most accurate and consistent predictions among all evaluated datasets.

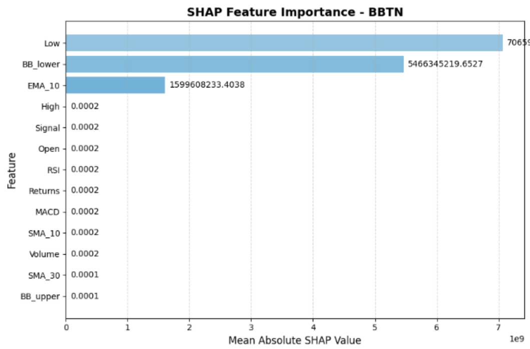

Figure 10. SHAP BBTN

Figure 10 the SHAP analysis for the BBTN stock prediction model shows that the Low, BB_lower, and EMA_10 features have the highest influence on the model's predictions.

The performance metrics in Table XXI confirm the strong forecasting capability of the LSTM model across all five banking stocks. With MAPE values consistently below 5%, the model demonstrates high predictive accuracy. BBCA and BMRI stand out with the highest R^2 scores above 0.96, reflecting the model's strong ability to capture long-term price movements. Overall, the results indicate that the LSTM framework performs reliably in modeling stock price dynamics within the Indonesian banking sector.

TABLE XXI
COMPARATIVE PERFORMANCE OF LSTM PREDICTIONS

Stock	MAE	MAPE	RMSE	R^2
BBCA	188.84	2.34%	238.69	0.9631
BBRI	116.29	2.89%	150.78	0.9472
BMRI	174.02	3.88%	219.82	0.9681
BBNI	122.98	3.19%	161.68	0.9610
BBTN	32.91	2.80%	43.29	0.8773

Table XXI shows positive results are also seen for BBRI and BBNI, though both exhibit slightly higher sensitivity to market fluctuations, indicating moderate exposure to volatility-related forecasting risk. In comparison, BBTN poses a greater modeling challenge due to its more pronounced price variability, reflected in the lower R^2 value of 0.8773.

The variation in predictive accuracy observed among the five banking stocks is closely linked to differences in their fundamental economic and market characteristics. Banks with larger capitalization, such as BBCA and BBRI, typically exhibit higher liquidity levels and more consistent trading activity, resulting in smoother price movements that the model can learn with greater precision. In contrast, institutions like BBTN, which operate with relatively lower liquidity and higher volatility, present more irregular price dynamics that challenge the model's forecasting stability. This pattern aligns with established evidence in financial econometrics, where asset liquidity, volatility, and capitalization are recognized as key determinants of price predictability. Accordingly, the observed disparities in accuracy should not be viewed as methodological shortcomings but rather as reflections of the inherent heterogeneity and structural diversity within Indonesia's banking market.

These findings underscore the critical role of price stability in enhancing model reliability. The LSTM architecture demonstrates superior capability in learning long-term sequential patterns, while its responsiveness to short-term stochastic changes remains limited. Nonetheless, the model's ability to preserve temporal dependencies offers practical value for investors and financial analysts, providing a robust framework for informed and strategically conservative investment decision-making.

In addition to delivering strong predictive accuracy, the LSTM–TSCV model also offers concise financial insights. The SHAP analyses indicate that price-based features such as Low, trading volume, and trend indicators including SMA, EMA, and Bollinger Bands play the most influential roles in shaping forecast outcomes, particularly for BBCA and BBRI. The operational results show that BUY signals appear when the predicted price exceeds the current market value, while SELL signals occur when the model anticipates limited or negative movement, as seen for BBTN. These patterns are consistent with core principles of technical and behavioral finance, underscoring the importance of liquidity, trend strength, and investor trading behavior in short-term price dynamics.

TABLE XXII
OPERATIONAL IMPLEMENTATION RESULTS BASED ON LSTM FORECASTS

Bank	Current Price	Predicted Price	Expected Change (%)	Result
BBCA	8075.00	8650.90	7.13	BUY
BBRI	4050.00	4182.65	3.28	BUY
BMRI	4730.00	4950.75	4.67	BUY
BBNI	4380.00	4515.98	3.10	BUY
BBTN	1300.00	1299.34	-0.05	SELL

The operational application of the LSTM–TSCV model is shown in Table XXII. The table provides an overview of how the model's prediction results can be used for stock trading recommendations by comparing the latest actual prices with the predicted values. The Expected Change (%) indicates the percentage difference between the prices and serves as the basis for determining trading actions. A positive percentage indicates a potential price increase, corresponding to a BUY signal, while a negative percentage suggests a possible decline, resulting in a SELL recommendation.

In this analysis, the model indicates potential price growth for BBCA, BBRI, BMRI, and BBNI, resulting in BUY signals. In contrast, BBTN shows a minor decrease and is categorized as a SELL. These findings demonstrate that the developed model can be operationalized within a stock recommendation or decision-support dashboard, providing investors and analysts with quantitative insights for informed short-term trading decisions.

TABLE XXIII
BASELINE COMPARASION XGBOOST

Bank	MAE	RMSE	MAPE	R^2
BBCA	586.0	791.2	6.12%	-0.7022
BBRI	126.4	202.7	2.69%	0.8474
BMRI	925.6	1112.0	16.08%	-2.0281
BBNI	366.9	507.0	7.64%	-0.4354
BBTN	17.4	23.9	1.51%	0.9698

To establish a non-sequential baseline, XGBoost was implemented using the same feature set and sliding-window configuration as the proposed LSTM–TSCV model. As shown in Table XXIII, XGBoost delivered competitive performance for BBRI and BBTN, achieving high R^2 values of 0.8474 and 0.9698, respectively. However, its performance deteriorated sharply for BBCA, BMRI, and BBNI, where the R^2 values were negative, indicating that the model performed worse than a simple mean predictor. These results highlight the limitations of tree-based models in capturing temporal dependencies and complex nonlinear patterns in stock price movements. Consequently, the findings reinforce the suitability of the LSTM–TSCV approach, which provides more stable and accurate predictions across all five banking stocks.

IV. Conclusion

This study investigated the forecasting of stock price movements in the highly volatile Indonesian banking sector using LSTM networks, enhanced through hyperparameter optimization and TSCV. The empirical findings confirm that LSTM models effectively capture long-term dependencies within sequential financial data, yielding reliable and interpretable predictions.

BBCA and BMRI exhibited the most stable performance among the evaluated stocks, with R^2 values exceeding 0.95 and consistently low error rates. BBCA, in particular, achieved the lowest MAPE of 2.34%, highlighting the model's precision in tracking long-term price trajectories. Although the model encountered challenges during periods of extreme volatility, it maintained robust accuracy across most conditions, reinforcing its potential as a dependable forecasting framework.

Overall, the study successfully validated the applicability of LSTM for predicting Indonesian banking stock prices, emphasizing the importance of hyperparameter tuning and TSCV in ensuring methodological rigor. The findings theoretically and practically support investors, market analysts, and policymakers in making data-driven decisions while enriching the knowledge in AI-driven financial forecasting. Furthermore, the study establishes a foundation for future improvements in predictive modeling, particularly in enhancing responsiveness to short-term market fluctuations and volatility.

References

[1] B. Y. Dwiandiyanta, R. Hartanto, and R. Ferdiana, "Optimization of Stock Predictions on Indonesia Stock Exchange: A New Hybrid Deep Learning Method," *Engineering, Technology and Applied Science Research*, vol. 15, no. 1, pp. 19370–19379, Feb. 2025, doi: 10.48084/etasr.9363.

[2] M. Azka Atqiya, L. Septem Riza, and A. Anisyah, "Stock Price Prediction Using the ETSFormer Model Case Study: PTBA," vol. 5, no. 2, 2025, doi: 10.47709/brilliance.v5i2.6729.

[3] 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 (Jutif)*, vol. 6, no. 3, pp. 1539–1554, 2025, doi: 10.52436/1.jutif.2025.6.3.4648.

[4] "Tren Investasi Saham di 2025: Apa yang Harus Dipertimbangkan Sebelum Berinvestasi ?" Accessed: Sep. 22, 2025. [Online]. Available: https://online.binus.ac.id/accounting/2025/04/17/tren-investasi-saham-di-2025-apa-yang-harus-dipertimbangkan-sebelum-berinvestasi/

[5] D. Satria, "Predicting Banking Stock Prices Using Rnn, Lstm, And GRU Approach," *Applied Computer Science*, vol. 19, no. 1, pp. 82–94, 2023, doi: 10.35784/acs-2023-06.

[6] "Everything about LSTM and its working." Accessed: Sep. 22, 2025. [Online]. Available: https://medium.com/@abhishekjainindore24/everything-about-lstm-and-its-working-f25f1cc30482

[7] W. Budiharto, "Data science approach to stock prices forecasting in Indonesia during Covid-19 using Long Short-Term Memory (LSTM)," *J Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-021-00430-0.

[8] S. A. S. Purwandhani, A. A. N. Sajiatmoko, and C. S. K. Aditya, "Optimizing Indonesian Banking Stock Predictions with DBSCAN and LSTM," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 3, pp. 1173–1188, 2025, doi: 10.52436/1.jutif.2025.6.3.4439.

[9] F. Santosa, E. Oktafanda, H. Setiawan, and A. Latif, "Advanced Long Short-Term Memory (LSTM) Models for Forecasting Indonesian Stock Prices," *Jurnal Galaksi*, vol. 1, no. 3, pp. 198–208, 2024, doi: 10.70103/galaksi.v1i3.42.

[10] W. Li and K. L. E. Law, "Deep Learning Models for Time Series Forecasting: A Review," 2024, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/ACCESS.2024.3422528.

[11] W. Hastomo, A. Satyo, B. Karno, N. Kalbuana, E. Nisfiani -4, and L. Etp -, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Optimasi Deep Learning untuk Prediksi Saham di Masa Pandemi Covid-19," vol. 7, Aug. 2021, doi: 10.26418/jp.v7i2.47411.

[12] R. Gunawan Santosa, A. R. Chrismanto, and Y. Lukito, "Stocks Forecasting Exploration on Lq45 Index Using Arima (P,D,Q) Model," *J Theor Appl Inf Technol*, vol. 15, no. 13, 2022, [Online]. Available: www.jatit.org

[13] B. Jange, "Prediksi Harga Saham Bank BCA Menggunakan XGBoost," *ARBITRASE: Journal of Economics and Accounting*, vol. 3, no. 2, pp. 231–237, Nov. 2022, doi: 10.47065/arbitrase.v3i2.495.

[14] T. Purwa, U. Nafngiyana, and S. Suhartono, "Comparison of Arima, Transfer Function And Var Models For Forecasting Cpi, Stock Prices, And Indonesian Exchange Rate: Accuracy Vs. Explainability," *MEDIA STATISTIKA*, vol. 13, no. 1, pp. 1–12, Jun. 2020, doi: 10.14710/medstat.13.1.1-12.

[15] E. P. Wahyuddin, R. E. Caraka, R. Kurniawan, W. Caesarendra, P. U. Gio, and B. Pardamean, "Improved LSTM hyperparameters alongside sentiment walk-forward validation for time series prediction," *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 11, no. 1, p. 100458, 2025, doi: 10.1016/j.joitmc.2024.100458.

[16] "How-To: Cross Validation with Time Series Data." Accessed: Sep. 22, 2025. [Online]. Available: https://towardsdatascience.com/how-to-cross-validation-with-time-series-data-9802a06272c6/

[17] "Bursa Efek Jakarta & Indonesia | Pasar Keuangan Indonesia." Accessed: Sep. 22, 2025. [Online]. Available: https://www.idnfinancials.com/id/

[18] A. Aldoseri, K. N. Al-Khalifa, and A. M. Hamouda, "Re-Thinking Data Strategy and Integration for Artificial Intelligence: Concepts, Opportunities, and Challenges," Jun. 01, 2023, *MDPI*. doi: 10.3390/app13127082.

[19] "Data Pre-Processing with Sklearn using Standard and Minmax scaler." Accessed: Sep. 22, 2025. [Online]. Available: https://www.geeksforgeeks.org/machine-learning/data-pre-processing-wit-sklearn-using-standard-and-minmax-scaler/

[20] W. Li, L. Zhang, and X. Zhang, "Predicting Multi-Indicator Stock Time Series using Convolutional Neural Networks based on Feature Engineering," *International Journal of Computers Communications & Control*, vol. 20, no. 5, 2025, doi: 10.15837/ijccc.2025.5.6774.

[21] "Time Series Cross-Validation (tsCV)." Accessed: Sep. 22, 2025. [Online]. Available: https://schneppat.com/time-series-cross-validation_tscv.html

[22] V. Deswal, "Hyperparameter Tuning of the LSTM model for Stock Price Prediction," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4, pp. 705–712, 2024.

[23] B. Tanuwijaya *et al.*, "A Novel Single Valued Neutrosophic Hesitant Fuzzy Time Series Model: Applications in Indonesian and Argentinian Stock Index Forecasting," *IEEE Access*, vol. 8, pp. 60126–60141, 2020, doi: 10.1109/ACCESS.2020.2982825.