



Volume 11 Issue 1 Year 2026 | Page 320-330 ISSN: 2527-9866

Received: 26-01-2026 / Revised: 05-02-2026 / Accepted: 25-02-2026

A Comparative Analysis of Deep Learning Models for Stock Price Prediction

I Gusti Ayu Nandia Lestari¹, Deviana², Tubagus Mahendra Kusuma³

^{1,2,3}Institute of Technology and Business (ITB) STIKOM Bali, Denpasar, Indonesia, 15331

e-mail: nandia@stikom-bali.ac.id¹, deviana@stikom-bali.ac.id², mahendra_kusuma@stikom-bali.ac.id³

*Correspondence: nandia@stikom-bali.ac.id

Abstract: Indonesian equities exhibit high volatility and non-stationary dynamics, making consistent price forecasting difficult under realistic deployment settings. This study presents a comparative benchmark of Long Short-Term Memory (LSTM) and Bidirectional LSTM (Bi-LSTM) for one-step-ahead ($t+1$) stock price prediction using Walk-Forward Validation (WV) to preserve temporal causality and avoid optimistic single-split estimates. Historical data are retrieved from Yahoo Finance and modeled in a multivariate OHLCV setting (Open, High, Low, Close, Volume). After missing-value removal, feature standardization, and Min–Max scaling, the series is converted to supervised samples via a sliding window with $\text{lookback} = 30$ trading days; evaluation is focused on the Close variable. Model performance is assessed using MAE, RMSE, and R^2 , including inter-fold variability to quantify stability across market regimes. Across five Indonesian tickers (AGRO, ADES, ADMF, AALI, ADHI), LSTM consistently outperforms Bi-LSTM (5/5 tickers) in both MAE and RMSE, indicating that the added bidirectional complexity does not translate into improved out-of-sample forecasting under WV. The best error performance is achieved by LSTM on AGRO (MAE = 26.99, RMSE = 32.72), while the least-negative goodness-of-fit is observed on LSTM AALI ($R^2 = -0.63$), suggesting that both deep models may still underperform naïve baselines in several folds. Overall, the results support LSTM as a more stable and implementation-ready benchmark for Indonesian stock forecasting under time-aware evaluation, while highlighting the need for explicit baseline comparisons and stronger feature/target designs to improve out-of-sample generalization.

Keywords: Deep learning; LSTM; Bi-LSTM; Comparative analysis; Indonesian capital market.

1. Introduction

Indonesia's stock market plays a central role in capital mobilization and price discovery, while also providing an increasingly accessible investment channel for retail participants[1]. In line with capital market deepening, the retail investor base has expanded rapidly; market authorities report that the number of Single Investor Identifications (SID) surpassed 20 million by December 2025, with domestic investors dominating participation and a substantial proportion coming from younger age groups [2]. This development amplifies the need for reliable, data-driven decision support tools that can help investors and analysts navigate price dynamics that are often noisy, non-stationary, and sensitive to regime shifts.[3].

Deep learning models, particularly Recurrent Neural Network (RNN) variants such as Long Short-Term Memory (LSTM) and Bidirectional LSTM (Bi-LSTM), have been widely adopted for financial time series forecasting because they can capture nonlinear dependencies and long-range temporal structure. Nevertheless, model complexity alone does not guarantee real-world usefulness. Performance claims may be overstated when evaluation relies on a single train–test split or protocols that do not respect chronological ordering, potentially introducing look-ahead bias and optimistic error estimates. A time-aware evaluation framework is therefore crucial to ensure that reported performance reflects realistic deployment settings in which models are trained on past data and used to predict unseen future periods.

Prior studies have extensively examined LSTM/Bi-LSTM for stock price prediction, including within the Indonesian context. Some works explicitly evaluate Bi-LSTM on Indonesian stock prices, while others propose derivative architectures (e.g., stacked LSTM or CNN–BiLSTM) using historical data sources such as Yahoo Finance. Moreover, comparative studies in the academic literature indicate that relative model performance can vary across issuers—meaning that no single model consistently dominates across all stocks—highlighting the importance of rigorous and comparable evaluation designs [4].

This motivates the key research gap addressed in this study. First, time-aware benchmark studies that directly compare LSTM and Bi-LSTM under repeated chronological testing (e.g., walk-forward protocols) remain limited for the Indonesian market, despite its propensity for volatility and regime changes. Second, evaluations often emphasize point accuracy without assessing stability across periods, even though consistency across market conditions is critical for practical decision support. Third, many studies do not include simple yet essential baselines (e.g., naïve persistence) and therefore cannot clearly demonstrate whether deep learning models outperform straightforward forecasting strategies, especially in folds where negative R^2 may occur.

Accordingly, the novelty of this research lies in establishing a more realistic and reproducible benchmark for Indonesian stock price forecasting by: (i) applying Walk-Forward Validation (expanding/rolling windows) to compare LSTM and Bi-LSTM in a strictly time-aware setting; (ii) reporting both predictive accuracy (MAE, RMSE, R^2) and inter-fold stability to quantify robustness across market periods; and (iii) strengthening the comparative design through explicit baseline comparisons and broader ticker coverage to improve generalizability. As a result, the study aims to identify not only which model is more accurate, but also which is more reliable and implementation-ready for real-world forecasting in Indonesia’s capital market context.

2. Literature Review

The development of Artificial Intelligence (AI) has driven significant advances in financial analytics, particularly in stock price prediction. In early stages, stock forecasting largely relied on statistical methods such as ARIMA and linear regression. However, these approaches are constrained by linearity assumptions and their limited ability to capture market dynamics that are nonlinear, non-stationary, and noisy [5]. Consequently, deep learning has emerged as a widely adopted alternative, especially for complex time-series data. One of the most popular deep learning models for time-series forecasting is LSTM. LSTM is an extension of Recurrent Neural Networks (RNNs) designed to mitigate the vanishing gradient problem through gate mechanisms (forget, input, and output gates), enabling the model to retain long-term information [6]. In capital market applications, several studies report that LSTM can achieve higher predictive accuracy than conventional methods due to its ability to learn long-range and complex temporal patterns [7].

Nevertheless, LSTM is unidirectional, processing sequences only from past to future. To enrich temporal context, Bidirectional LSTM (Bi-LSTM) was developed to process sequences in both forward and backward directions, thereby producing a richer representation of temporal dependencies [8]. A number of studies indicate that Bi-LSTM can yield lower prediction errors than LSTM and ARIMA on certain financial datasets, particularly in global markets. Moreover, recent work has integrated attention mechanisms into Bi-LSTM to further improve performance while enhancing model interpretability [9].

3. Methods

A. System Overview

This workflow begins with a literature review and framework development to define the research scope and methodology can be seen in Figure 1. Next, historical stock data (OHLCV) are collected and then preprocessed through cleaning and normalization to ensure consistent input quality. An experimental design is established to compare LSTM and Bi-LSTM models, followed by model training using Walk-Forward Validation with an expanding-window scheme to simulate realistic time-series prediction.

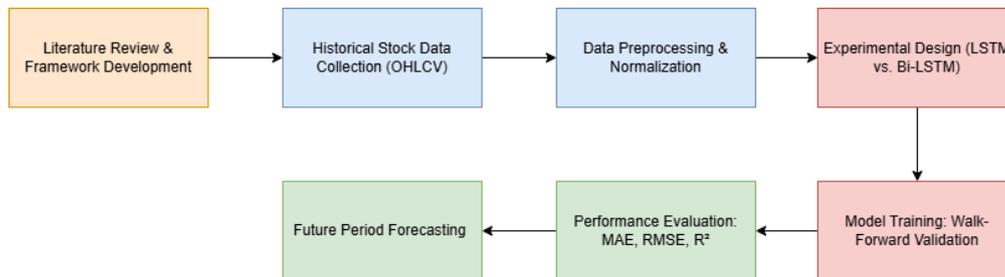


Figure 1. System Overview

B. Data Acquisition

This study uses historical Indonesian stock data retrieved from Yahoo Finance via the yfinance library can be seen in Table 1. To ensure reproducibility, the observation window is fixed from 2015-01-01 to 2026-01-28, and the data were downloaded on 2026-01-28. The dataset is collected at a daily frequency (trading days only) and OHLCV format, consisting of Open, High, Low, Close, and Volume.

Table 1. Data Description

| Variable | Description |
|----------|--|
| Open | The stock price at the start of the trading session (the first executed trade) for that day. It often reflects the market’s initial reaction to overnight news and early sentiment. |
| High | The highest price reached during the trading day. It indicates the maximum buying pressure and the upper bound of intraday price movement |
| Low | The lowest price reached during the trading day. It represents the strongest selling pressure and the lower bound of intraday price movement. |
| Close | The stock price at the end of the trading session (the last executed trade) for that day. It is commonly used as the primary forecasting target because it summarizes the day’s final market valuation. |
| Volume | The total number of shares traded during the day. Volume captures liquidity and market activity; spikes in volume often signal stronger interest, potential trend confirmation, or reactions to significant events [10]. |

To improve generalizability beyond a single issuer, the experiments are conducted on multiple Indonesian stocks rather than focusing on only one ticker. The ticker set is selected to represent varied issuer characteristics (e.g., sector diversity and trading activity), thereby providing a more robust comparative benchmark for LSTM vs. Bi-LSTM under Indonesian market dynamics. In this study, the evaluated tickers. can be seen in Table 2. Results are reported per-ticker and summarized across tickers to support more reliable conclusions about model behavior.

Table 2. Ticker Saham

| Symbol | Company Name |
|--------|-------------------------------------|
| ADES | PT Akasha Wira International Tbk |
| ADMF | PT Adira Dinamika Multi Finance Tbk |
| AALI | PT Astra Agro Lestari Tbk |
| ADHI | PT Adhi Karya (Persero) Tbk |
| AGRO | PT Bank Raya Indonesia Tbk |

The tickers were selected to represent diverse sectors and market characteristics in the Indonesian Stock Exchange, so that the comparison between LSTM and Bi-LSTM is not biased toward a single industry pattern. In addition, the chosen tickers provide consistent data availability over the study period and exhibit sufficient trading activity (liquidity), which reduces missing records and improves the reliability of OHLCV-based modeling and walk-forward evaluation. This selection strategy supports more generalizable conclusions by testing the models across stocks with different volatility and liquidity profiles.

C. Preprocessing

The preprocessing stage is conducted to ensure that the data are suitable for use as inputs to the time-series model :

1. OHLCV column selection and missing-value removal: The OHLCV features are selected, and rows containing missing values are removed using dropna() to prevent incomplete records from disrupting the sliding-window construction. can be seen in Figure 2.

```
Missing value per kolom
open      0
high      0
low       0
close     0
volume    0
dtype: int64
```

Figure 2. Checking Missing Values

2. Feature name standardization: The feature names are standardized to open, high, low, close, and volume to ensure consistent and unambiguous column referencing throughout the pipeline. can be seen in Table 3.

Table 3. Standardization of feature names

| Previous Feature Name | After Feature Name |
|-----------------------|--------------------|
| Open ANTM.JK | open |
| High ANTM.JK | high |
| Low ANTM.JK | low |
| Close ANTM.JK | close |
| Volume ANTM.JK | volume |

3. Normalization (scaling): The features are normalized using MinMaxScaler to map values into the [0,1] range can be seen in Table 4 and Table 5.

- During walk-forward validation, the scaler is fit only on the training data for each fold and then applied to the corresponding test segment, thereby preventing data leakage.
- During final training and forecasting, the scaler is fit on the full historical dataset before generating forecasts for future periods[11].

Table 4. Data before applying MinMaxScaler

| date | open | high | low | close | volume |
|------------|--------|--------|--------|--------|-----------|
| 2023-01-02 | 1990.0 | 2010.0 | 1975.0 | 1980.0 | 15045200 |
| 2023-01-03 | 1980.0 | 2020.0 | 1975.0 | 2000.0 | 45001200 |
| 2023-01-04 | 2020.0 | 2100.0 | 2010.0 | 2050.0 | 142169900 |
| 2023-01-05 | 2020.0 | 2050.0 | 1935.0 | 1965.0 | 130118600 |

Table 5. Data after applying MinMaxScaler

| date | open | high | low | close | volume |
|------------|----------|----------|----------|----------|----------|
| 2023-01-02 | 0.253205 | 0.247244 | 0.260163 | 0.250000 | 0.019691 |
| 2023-01-03 | 0.250000 | 0.250394 | 0.260163 | 0.256494 | 0.058897 |
| 2023-01-04 | 0.262821 | 0.275591 | 0.271545 | 0.272727 | 0.186069 |
| 2023-01-05 | 0.262821 | 0.259843 | 0.247154 | 0.245130 | 0.170297 |

D. Sliding Window / Lookback

To enable the LSTM/Bi-LSTM models to learn temporal patterns, the time-series data are transformed into a supervised learning format using a sliding-window approach can be seen in Table 6 and table 7.

Table 6. Example of Input Dataset

| open | high | low | close | volume |
|----------|----------|----------|----------|----------|
| 0.253205 | 0.247244 | 0.260163 | 0.250000 | 0.019691 |
| 0.250000 | 0.250394 | 0.260163 | 0.256494 | 0.058897 |
| 0.262821 | 0.275591 | 0.271545 | 0.272727 | 0.186069 |
| 0.262821 | 0.259843 | 0.247154 | 0.245130 | 0.170297 |
| 0.240385 | 0.239370 | 0.245528 | 0.251623 | 0.072864 |

Table 7. Target 1 step ahead

| open | high | low | close | volume |
|----------|----------|----------|----------|----------|
| 0.333333 | 0.322835 | 0.333333 | 0.324675 | 0.062123 |

E. Modeling

1. LSTM (Unidirectional)

LSTM is a variant of RNNs designed to model sequential (time-series) data more stably than standard RNNs. LSTM can retain important long-term information through a cell-state mechanism and three gates: the forget gate, input gate, and output gate. In the unidirectional setting, the sequence is processed in a single direction, from $t-k \rightarrow \dots \rightarrow t-1 \rightarrow t$ (past to future), which is well suited to stock price prediction tasks that rely on historical observations [12].

- 1) Forget gate

$$f_t = \sigma(w_f[h_t - 1, x_t] + b_f)$$

- 2) Input gate

$$i_t = \sigma(w_i[h_t - 1, x_t] + b_i)$$

- 3) Memory candidates

$$\tilde{c}_t = \tanh(w_c[h_t - 1, x_t] + b_c)$$

- 4) Update cell state

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

- 5) Output gate

$$o_t = \sigma(w_o[h_t - 1, x_t] + b_o)$$

2. Bi-LSTM (Bidirectional)

Bi-LSTM is an extension of LSTM that processes sequence data in two directions simultaneously, namely:

- Forward: from the start time to the end ($t = 1 \rightarrow T$). For the sequence x_1, x_2, \dots, x_T the forward mechanism follows the standard LSTM formula:

$$\vec{f}_t = \sigma(w_f[h_t - 1, x_t] + b_f)$$

- Backward: from end to beginning time($t = T \rightarrow 1$)

$$\overleftarrow{h}_t = LSTM_{backward}(x_t, h_{t+1}, c_{t+1})$$

- Two-way output merging

$$h_t = [\vec{h}_t; \overleftarrow{h}_t]$$

The outputs from these two directions are then combined typically via concatenation to produce a richer temporal representation. In practice, Bi-LSTM often improves the model’s ability to capture complex sequential patterns because it can “see” contextual information from both sides of the sequence. By processing data in both forward and backward directions, Bi-LSTM yields a more informative temporal representation [13]. However, this advantage comes at the cost of higher computational complexity.

3. Architecture and hyperparameters

The core architecture uses a single recurrent layer followed by dropout and a linear output layer. The model configuration is presented in Figure 3.

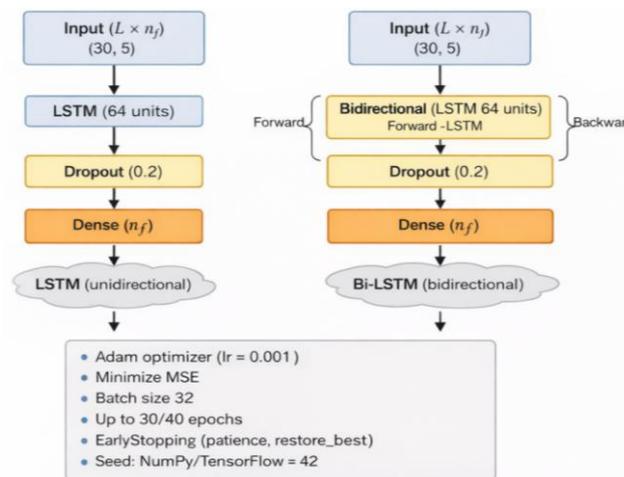


Figure 3. Architecture and hyperparameters

For LSTM, the network is: $Input(L \times n_f) \rightarrow LSTM(u = 64) \rightarrow Dropout(0.2) \rightarrow Dense(n_f)$. For Bi-LSTM, the recurrent layer is replaced with Bidirectional(LSTM(64)), keeping the remaining layers identical. Default LSTM nonlinearities are used (sigmoid gates and tanh state). Models are trained with Adam ($lr = 0.001$) to minimize MSE, using batch size 32. For WFV folds, training runs up to 30 epochs with early stopping (patience=5, restore_best_weights=True) and a 10% validation split; the final forecasting model is trained up to 40 epochs (patience=7) with a 30% validation split. Reproducibility is ensured by fixing NumPy and TensorFlow random seeds to 42[14].

F. Evaluation with Walk-Forward Validation

Walk-Forward Validation (WFV) is implemented using an expanding-window protocol. Let N be the number of observations, L the lookback length, and H the test horizon[15]. The initial training segment uses $L + H$ observations, and each fold evaluates the next H observations with a stride of H . Thus, fold k trains on $[1, \dots, tk]$ and tests on $[tk + 1, \dots, tk + H]$, where $t_1 = L + H$ and $t_{k+1} = tk + H$. The total number of folds is $K = \left(\frac{N-(L+H)}{H}\right)$. This time-aware

evaluation produces both average accuracy and inter-fold stability[16]. Evaluation is conducted using Walk-Forward Validation with an expanding-window mechanism:

- The model is trained on historical data up to a given time point, where the training set grows progressively across folds can be seen in Table 9.

Table 9. Metric Equations

| Metric | Equations |
|--------------------------------|---|
| MAE (Mean Absolute Error) | $MAE = \frac{1}{n} = \sum_{i=1}^n y_i - \hat{y}_i $ |
| RMSE (Root Mean Squared Error) | $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$ |
| R ² Score [17] | $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$ |

G. Future Period Forecast (Multi-step Forecast)

After the walk-forward benchmark is completed, a final training phase is performed using the entire historical dataset, followed by recursive multi-step forecasting for future periods. In this study, the forecasting horizon is set to 126 trading days (approximately six months of market sessions). The model first predicts $t + 1$, then the predicted value is fed back as part of the input to predict the next step, and this process is repeated iteratively until the full horizon is generated[18]. The output of this stage is a future prediction series (primarily for the Close variable) accompanied by visualizations that include: (1) the historical Actual Close, (2) the out-of-sample predictions obtained from walk-forward validation (aligned at the end of the historical period), and (3) the future forecast for the upcoming period. can be seen in Figures 4 and 5 predictions from several stock issuers.



Figure 4. visualization of ADES issuer stock predictions



Figure 5. visualization of AGRO issuer stock predictions

3. Results and Discussion

This section presents the results of the research implementation in accordance with the proposed methodology, covering the full pipeline from data acquisition and preprocessing to model evaluation. The study employs two deep learning approaches LSTM and Bi-LSTM and

evaluates their predictive performance using MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), and the R² score.

a) LSTM (Unidirectional)

Based on the Walk-Forward Validation (out-of-sample) evaluation, the LSTM model exhibits prediction errors that vary across issuers, as shown in Table 10.

- AGRO shows the lowest errors, with MAE = 26.99 and RMSE = 32.72. However, R² = -1.12 indicates that its performance still does not outperform a simple baseline in most folds.
- ADHI also falls into the lower–medium error group (MAE = 43.92, RMSE = 53.52), but R² = -2.12 suggests weak performance relative to the baseline.
- ADES and ADMF present relatively high errors (ADES: MAE = 119.26, RMSE = 153.74; ADMF: MAE = 116.77, RMSE = 141.25) and negative R² values (ADES: -2.78; ADMF: -1.21).
- AALI has the largest errors (MAE = 347.31, RMSE = 416.99), although its R² = -0.63 is closer to zero than several other tickers.

In terms of stability, the standard deviations for ADES, ADMF, and AALI are very large (e.g., ADES: std(MAE) = 170.74, std(RMSE) = 218.68; AALI: std(MAE) = 464.45), indicating that the model’s performance is not consistent across different testing periods.

Table 10. Prediction Results of the LSTM method

| | | mean | std |
|------|----------------|------------|------------|
| AGRO | MAE | 26.995517 | 39.104271 |
| | RMSE | 32.715390 | 45.084879 |
| | R ² | -1.117909 | 3.756582 |
| ADES | MAE | 119.259335 | 170.737177 |
| | RMSE | 153.743569 | 218.677417 |
| | R ² | -2.775203 | 18.738106 |
| ADMF | MAE | 116.774579 | 145.185283 |
| | RMSE | 141.245775 | 183.160405 |
| | R ² | -1.212801 | 5.330232 |
| AALI | MAE | 347.314173 | 464.453731 |
| | RMSE | 416.989248 | 513.615704 |
| | R ² | -0.631792 | 2.182792 |
| ADHI | MAE | 43.924427 | 33.625197 |
| | RMSE | 53.517842 | 39.278649 |
| | R ² | -2.116928 | 8.745164 |

b) Bi-LSTM (Bidirectional)

The Bi-LSTM model generally shows higher errors than the LSTM model for most tickers, as presented in Table 11.

- For AGRO, Bi-LSTM produces larger errors than LSTM (MAE 30.11 vs 26.99, RMSE 36.01 vs 32.72) and a more negative R² (-1.95).
- For ADES, Bi-LSTM is also worse (MAE 135.63, RMSE 169.50) compared to LSTM. Although the R² is slightly “better” numerically (-2.67 vs -2.78), it remains negative.
- For ADMF, Bi-LSTM yields higher errors (MAE 125.26, RMSE 151.11) and a more negative R² (-1.47).

- For AALI, Bi-LSTM again produces larger errors (MAE 376.13, RMSE 440.93) with a more negative R^2 (-1.00).
- For ADHI, Bi-LSTM also shows higher errors (MAE 53.41, RMSE 63.26), with R^2 still negative (-2.16).

Overall, Bi-LSTM does not demonstrate an advantage over LSTM under the out-of-sample evaluation setting for this set of issuers, either in terms of accuracy (MAE/RMSE) or stability (standard deviation).

Table 11. Bi-LSTM Prediction Results

| | | mean | std |
|------|-------|------------|------------|
| AGRO | MAE | 30.110744 | 40.966274 |
| | RMSE | 36.008367 | 47.423841 |
| | R^2 | -1.946079 | 5.021514 |
| ADES | MAE | 135.634083 | 202.080576 |
| | RMSE | 169.501736 | 242.175269 |
| | R^2 | -2.665584 | 10.189656 |
| ADMF | MAE | 125.264900 | 130.098650 |
| | RMSE | 151.113512 | 169.477961 |
| | R^2 | -1.471744 | 4.032255 |
| AALI | MAE | 376.134262 | 491.227164 |
| | RMSE | 440.926294 | 527.625358 |
| | R^2 | -0.999605 | 2.910991 |
| ADHI | MAE | 53.414824 | 46.941731 |
| | RMSE | 63.257310 | 52.660222 |
| | R^2 | -2.156311 | 4.997891 |

c) Discussion

In theory, Bi-LSTM can capture bidirectional context. However, in real-world time-series forecasting, future data are not available at prediction time. Under a Walk-Forward setting (training on the past → testing on the future), the advantage of “bidirectional context” often does not materialize, while model complexity increases. Bi-LSTM has more parameters, which makes it:

- more prone to overfitting when the dataset is relatively short and noisy,
- more sensitive to regime changes (trend shifts) and volatility spikes,
- and consequently, its out-of-sample errors can become larger than those of the simpler LSTM.

This is consistent with the results across multiple tickers: Bi-LSTM yields higher MAE/RMSE and more negative R^2 values. An R^2 value below 0 indicates that, on average, the model performs worse than a very simple baseline (e.g., predicting “tomorrow’s price = today’s price” / persistence, or using the mean). In highly volatile stock markets, naïve baselines can be difficult to outperform when:

- the market during the test periods is highly dynamic,
- an explicit baseline is required for a fair comparison, and
- the experimental design must be improved (e.g., features/targets/horizon/prediction strategy).
-

Therefore, a negative R^2 is not merely a sign that the model “fails,” but a strong signal that:

- the market during the test periods is highly dynamic,
- an explicit baseline is required for a fair comparison, and
- the experimental design must be improved (e.g., features/targets/horizon/prediction strategy).

Overall, across the five tested issuers (AGRO, ADES, ADMF, AALI, ADHI), the LSTM model consistently outperformed Bi-LSTM under Walk-Forward Validation (out-of-sample), as evidenced by lower MAE and RMSE for all tickers (5/5). This indicates that, under the current architecture configuration and data range, the added complexity of Bi-LSTM does not provide accuracy gains and tends to be less effective. Nevertheless, the consistently negative R^2 values for both models suggest that predictive performance still does not surpass simple baselines in many folds. Thus, the main takeaway at this stage is: LSTM is more suitable as a more stable deep-learning benchmark for the Indonesian stock market context, but stronger evaluation design (explicit baselines, more robust features/targets, and stability analysis across periods) is needed to improve out-of-sample performance and make the results more implementation-relevant.

4. Conclusions

Based on the Walk-Forward Validation (out-of-sample) evaluation on five tested issuers (AGRO, ADES, ADMF, AALI, and ADHI), the LSTM model consistently outperforms the Bi-LSTM model by producing lower MAE and RMSE across all tickers (5/5). The best performance (lowest error) is achieved by LSTM on the AGRO ticker, with MAE = 26.99 and RMSE = 32.72. Meanwhile, the best R^2 value (closest to zero) is obtained by LSTM on the AALI ticker, with $R^2 = -0.63$, although this still indicates that the model has not surpassed a simple baseline. Overall, these findings confirm that under the current research configuration, LSTM is more accurate and more suitable as a benchmark than Bi-LSTM. However, since R^2 remains negative for all issuers, further work is required by incorporating explicit baselines (e.g., naïve/persistence), expanding the feature set, and performing more systematic hyperparameter tuning and forecasting strategy refinement to achieve stronger and more realistic out-of-sample performance for deployment in the Indonesian stock market context.

References

- [1] R. Aprayuda and F. Misra, “Faktor Yang Mempengaruhi Keinginan Investasi Investor Muda di Pasar Modal Indonesia,” *E-Jurnal Akuntansi*, vol. 30, no. 5, p. 1084, May 2020, doi: 10.24843/EJA.2020.v30.i05.p02.
- [2] D. Selasi and R. Hernawati, “Perkembangan Pasar Modal Syariah di Indonesia: Potensi, Tantangan, dan Regulasi dalam Investasi Berbasis Syariah,” *Jurnal Ekonomi, Bisnis dan Manajemen*, vol. 3, no. 4, pp. 171–185, Dec. 2024, doi: 10.58192/ebismen.v3i4.2743.
- [3] 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, p. 47, Dec. 2021, doi: 10.1186/s40537-021-00430-0.
- [4] M. P. Halim, N. Yudistira, and C. Dewi, “Multicommodity Prices Prediction Using Multivariate Data-Driven Modeling: Indonesia Case,” *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 6, pp. 8106–8117, Dec. 2024, doi: 10.1109/TCSS.2022.3229772.
- [5] H. Mo, “Comparative Analysis of Linear Regression, Polynomial Regression, and ARIMA Model for Short-term Stock Price Forecasting,” *Advances in Economics, Management and Political Sciences*, vol. 49, no. 1, pp. 166–175, Dec. 2023, doi: 10.54254/2754-1169/49/20230509.

- [6] B. Lindemann, T. Müller, H. Vietz, N. Jazdi, and M. Weyrich, "A survey on long short-term memory networks for time series prediction," *Procedia CIRP*, vol. 99, pp. 650–655, 2021, doi: 10.1016/j.procir.2021.03.088.
- [7] M. M. Billah, A. Sultana, F. Bhuiyan, and M. G. Kaosar, "Stock price prediction: comparison of different moving average techniques using deep learning model," *Neural Comput. Appl.*, vol. 36, no. 11, pp. 5861–5871, Apr. 2024, doi: 10.1007/s00521-023-09369-0.
- [8] Y. Fan, Q. Tang, Y. Guo, and Y. Wei, "BiLSTM-MLAM: A Multi-Scale Time Series Prediction Model for Sensor Data Based on Bi-LSTM and Local Attention Mechanisms," *Sensors*, vol. 24, no. 12, p. 3962, Jun. 2024, doi: 10.3390/s24123962.
- [9] Q. Cheng, Y. Chen, Y. Xiao, H. Yin, and W. Liu, "A dual-stage attention-based Bi-LSTM network for multivariate time series prediction," *J. Supercomput.*, vol. 78, no. 14, pp. 16214–16235, Sep. 2022, doi: 10.1007/s11227-022-04506-3.
- [10] C. Fan and X. Zhang, "Stock price nowcasting and forecasting with deep learning," *J. Intell. Inf. Syst.*, vol. 63, no. 2, pp. 639–656, Apr. 2025, doi: 10.1007/s10844-024-00908-2.
- [11] P. P. G. D. Asanka, C. Rajapakshe, and M. Takahashi, "Harnessing Emerging Knowledge and Extensibility Techniques for Stationarity and Normalization in Multi-Agent Time Series Forecasting," *Engineering, Technology & Applied Science Research*, vol. 15, no. 5, pp. 26304–26309, Oct. 2025, doi: 10.48084/etasr.11329.
- [12] T. H. Putri *et al.*, "Fine-Tuning of Predictive Models CNN-LSTM and CONV-LSTM for Nowcasting PM_{2.5} Level," *IEEE Access*, vol. 12, pp. 28988–29003, 2024, doi: 10.1109/ACCESS.2024.3368034.
- [13] F. R. Alharbi and D. Csala, "Wind Speed and Solar Irradiance Prediction Using a Bidirectional Long Short-Term Memory Model Based on Neural Networks," *Energies (Basel)*, vol. 14, no. 20, p. 6501, Oct. 2021, doi: 10.3390/en14206501.
- [14] K. Astudillo, M. Flores, M. Soliz, G. Ferreira, and J. Varela-Aldás, "A Hybrid GAS-ATT-LSTM Architecture for Predicting Non-Stationary Financial Time Series," *Mathematics*, vol. 13, no. 14, p. 2300, Jul. 2025, doi: 10.3390/math13142300.
- [15] W. Sulandari, Y. Yudhanto, S. Subanti, E. Zukhronah, and M. Z. Subarkah, "Implementing Time Series Cross Validation to Evaluate the Forecasting Model Performance," *KnE Life Sciences*, Mar. 2024, doi: 10.18502/cls.v8i1.15584.
- [16] 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, Mar. 2025, doi: 10.1016/j.joitmc.2024.100458.
- [17] A. Vishwanath, M. Basheeruddin, and S. D, "Forecasting Sales Data Using Time Series Models and LSTM Model," in *2024 Second International Conference on Advances in Information Technology (ICAIT)*, IEEE, Jul. 2024, pp. 1–6. doi: 10.1109/ICAIT61638.2024.10690408.
- [18] Q. Yu *et al.*, "Dual-splitting conformal prediction for multi-step time series forecasting," *Appl. Soft Comput.*, vol. 184, p. 113825, Dec. 2025, doi: 10.1016/j.asoc.2025.113825.