JURNAL EMACS
[Engineering, MAthematics and Computer Science] Vol.8 No.1 January 2026: 79-85 e-ISSN: 2686-2573
DOI: 10.21512/emacsjournal.v8i1.15388

Effect of Price Volatility on LSTM Lookback Windows in Indonesian Banking Stocks

Joan Christina Bahagiono
Computer Science Department, School of Computer Science,
Bina Nusantara University,
Jakarta, Indonesia 11480
joan.bahagiono@binus.ac.id

Correspondence: joan.bahagiono@binus.ac.id

Abstract *– This study aims to explore how stock price volatility influences the sequence length or also known as lookback window hyperparameter of LSTM. This study uses a comparative approach to determine the relationship between stock prices volatility and the best lookback window to achieve the lowest error rate of an LSTM model in predicting stock prices. Nine selected stocks in the banking sector of Indonesia Stock Exchange were compared, ranging from relatively stable to volatile. The banking sector was used as it contains multiple stocks under the same sector that varies in price movement volatility. An aggregation was also conducted to produce grouped results. The results of this study highlighted the importance of hyperparameter tuning in LSTM especially in the lookback window hyperparameter. Shorter LSTM lookback window is well suited in low volatility stocks, with the lowest mean squared error rate of 0.030782 observed in this study at the 42 trading days lookback period. In contrast to that, highly volatile stocks exhibit a different pattern, where longer lookback period improves LSTM prediction performance, as demonstrated in this study through a 0.016001 mean squared error at the 252 trading days lookback period. The findings imply that high volatility stocks require longer temporal memory in the LSTM to capture complex and irregular price movements, whereas low volatility stocks are better modelled using shorter and more recent information.*

Keywords: *Long Short-Term Memory (LSTM); Stock Prices Prediction; Stock Prices Volatility; Indonesian Stock Exchange*

I. INTRODUCTION

Stocks listed on the Indonesia Stock Exchange (IDX) have varying levels of price movement, ranging from relatively stable trends to highly volatile fluctuations. Various machine learning techniques have been implemented in the past to predict stock prices (Soni et al., 2022), which support financial investment decision making processes (El Hajj & Hammoud, 2023) and even automating financial trading (Cohen, 2022). Models originating from the natural language processing field such as Long-Short Term Memory (LSTM) (Berrajaa, 2022) have also successfully been adapted to fit the domain of stock prices prediction (Deshpande, 2023). Due to its architecture, LSTM suits well in the domain of financial time series data prediction, including for other tasks similar to stock prices prediction such as gold prices prediction (Gong, 2024) and forex prices prediction (Ayitey Junior et al., 2022; Pahlevi et al., 2023). In the specific market of the Indonesian Stock Exchange, LSTM has also shown success in the task of stock prices prediction (Handharu Sworo & Hermawan, 2024; Joddy, 2025).

One important hyperparameter of LSTM is the sequence length, also known as lookback window, which determines the length of sequences used as training input to the model. Existing studies demonstrated the importance of lookback window tuning in different task-specific conditions to achieve high prediction accuracies (Xu, 2024). However, the main focus of these studies is how the hyperparameter-tuned prediction model affects the prediction accuracy, while the influence of underlying data characteristics such as data volatility has not yet been explored. A similar research gap is present in the domain of stock price prediction, where stock price movement volatility is acknowledged as one of the factors that may influence the lookback window size selection

(Mulyani & Ilham, 2025), but the direct relationship between them remains unexplored.

This study aims to bridge the research gap by examining how underlying data characteristics, specifically stock prices volatility, influences the lookback window hyperparameter determination to achieve low error rates of an LSTM model in predicting stock prices. This study brings novelty in taking a different approach of studying the underlying data characteristics in contrast to the common aproach of studying model architecture variations.

From a theoretical perspective, LSTM leverages a forget gate mechanism to selectively retain or discard contextual information (Öztürk, 2023), therefore highly volatile data can be a challenging aspect to the effectiveness of LSTM's architecture. In this study, a comparative analysis approach was taken between multiple banking sector stocks on IDX, ranging from relatively stable to fluctuative. The banking sector was selected as it contains multiple stocks under the same sector that varies in price movement volatility.

II. METHODS

This study carried out several methods to ensure stock prices volatility are well assessed and LSTM error rates results are compared across various banking sector stocks in the Indonesia Stock Exchange. In addition to that, several processes are also carried out to group the stocks to low volatility stock groups and high volatility stock groups for further comparison.

2.1. Dataset Gathering

Historical stock prices of nine banking sector stocks on the Indonesia Stock Exchange was retrieved from Yahoo Finance database programmatically through the yfinance Python library. The observation period spans for over five years, from January 1, 2021 to January 1, 2026. While the raw dataset includes open, high, low, close, and volume (OHLCV) data; the primary component analyzed in this study is the closing price. The dataset is then split into training and testing data.

2.2. Stock Volatility Calculation

To quantify the fluctuations of the selected stocks, a volatility measure is calculated for each stock by calculating the standard deviation of daily log returns of the respective stock:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad (1)$$

$$\sigma = std(r_t) \quad (2)$$

The daily stock volatility calculation result was also annualized based on a 252-day trading year:

$$\sigma_{annual} = \sigma \, . \sqrt{252} \quad (3)$$

Volatility is often calculated by the standard deviation of daily stock returns (Yan & Li, 2024). However, in this study, the standard deviation of daily logarithmic stock returns is used for calculation to mitigate skewness.

2.3. Stock Grouping based on Volatility

After stock volatility was calculated, stocks were grouped into two groups:

- Low Stock Price Volatility
- High Stock Price Volatility

The mean value of the annualized stock volatility measure is calculated and used as a threshold to split the two volatility groups. These stock groups are used to calculate aggregated results later in the study to supplement the individual stocks result in the study.

2.4. Long Short-Term Memory Model

The stock prices prediction utilizes a Long Short-Term Memory (LSTM) architecture, which is a type of recurrent neural network (RNN) (Wang et al., 2024) specifically designed to capture long-term dependencies and overcome the vanishing gradient problem of standard RNNs. For all experiments in this study, a two-layer LSTM model implemented using PyTorch is used:

Table 1. LSTM Architecture

Component	Specification
Input Dimension	1
Hidden Dimension	64
Number of LSTM Layers	2
Output Layer	1 (Fully Connected Layer)
Activation Functions	Tanh (state), Sigmoid (gates)
Loss Function	Mean Squared Error
Optimizer	Adam
Learning Rate	0.001
Number of Epochs	20

Prior to model training, a min-max scaler is used to normalize the data. After normalization, the same LSTM model architecture is used in all experiments to focus more on the hyperparameter tuning aspect and not on the model architecture. In addition, no regularization and batching is used in the LSTM architecture.

2.5. Lookback Window Selection

LSTM was tested on different sequence lengths representing lookback days and the mean squared error was plotted. Trading days were used as a measure instead of regular week lengths in order of ensuring domain suitability:

- 5 days: represents a trading week

- 10 days: represents two trading weeks (biweekly)
- 21 days: represents a trading month
- 42 days: represents two trading months
- 63 days: represents a trading quarter
- 126 days: represents half trading year
- 252 days: represents a trading year

2.6. Mean Squared Error (MSE)

To determine the accuracy of the LSTM models across different lookback windows, the Mean Square Error (MSE) is utilized as the primary loss metric:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{4}$$

2.7. Result Comparison

After separating training data and testing data, LSTM models were trained and tested with the various selected lookback windows to see which was the best lookback window for each stock. Initially, the experiment was conducted on each individual stock of the nine selected banking sector stock. The results were compared to see whether volatility affects the error rate on different lookback window periods of LSTM.

To further clarify the results, two stocks on the opposite end of the volatility difference spectrum were selected to carry out further comparative analysis. After that, RMSE was averaged for each stock groups. The aggregated RMSE results were compared to analyze whether the trends and results apply in an aggregated manner.

III. RESULTS AND DISCUSSION

The results were analyzed from multiple viewpoints, from a high level overview of how the lowest LSTM mean squared error vary among stocks to a comparison between the lowest and highest volatility stock. Besides individual stocks, two stock groups were also created to create an aggregated result to bring clarity to the results. Various charts were created to further explain the results including a trend analysis line chart to show underlying patterns of the volatiliy-based stock groups aggregated result.

Initially, volatility rates of each stock from the nine selected stocks from the banking sector of Indonesia Stock Exchange was calculated based on the calculation method defined in the methodology section. The banking sector stock volatility table shows the volatility rates of each stock sorted in increasing order.

Table 2. Banking Sector Stocks Volatility Calculation Results

No	IDX Stock Code	Stock Name	Volatility	Volatility Group
1.	BBCA	Bank Central Asia Tbk PT	0.2317	Low
2.	BNII	Bank Maybank Indonesia Tbk PT	0.2874	Low
3.	BBRI	Bank Rakyat Indonesia (Persero) Tbk PT	0.2886	Low
4.	BMRI	Bank Mandiri (Persero) Tbk PT	0.2956	Low
5.	BBNI	Bank Negara Indonesia (Persero) Tbk PT	0.3012	Low
6.	MEGA	Bank Mega Tbk PT	0.3451	Low
7.	BANK	Bank Aladin Syariah Tbk PT	0.5819	High
8.	ARTO	Bank Jago Tbk PT	0.6216	High
9.	BBYB	PT Bank Neo Commerce Tbk	0.7765	High

The mean volatility value across the selected stocks of 0.4144 was used as the threshold to split stocks into two volatility groups. The results shows that six stocks belong to the low volatility group while only three belong to the high volatility group. While the results of some major bank stocks such as BBCA, BBRI, BMRI, BBNI are as expected to be in the low volatility group, some bank stocks such as BNII and MEGA surprisingly also fall into this group. The three stocks that fall into the high volatility group were mainly categorized to belong in the digital banking sector of the Indonesia Stock Exchange. These three banks are expected to have high volatility since they are mostly recognized to still be in their youth phases and not as established as the major banks in Indonesia, although they are backed by major companies.

A volatility rate gap between MEGA (0.3451) and BANK (0.5819) was also encountered in the results, indicating a clear separation between low volatility and high volatility. Since the volatility rate was calculated using the annualized standard deviation of daily log returns, the difference between stock prices between banks were not an issue to be considered. Several different possible causes of the stock volatility rates gap include:

- Growth-based valuations: Unlike established banks that are valued on steady dividends and profits, newer banks are often valued on forward-looking metrics such as user growth and transaction volume. Since forward-looking metrics rely on predictions for future growth, any small changes in news reports such as a new partnerships and new investor fundings can cause major price adjustments.

- High sensitivity to interest rates: Often times, banks in the high volatility group rely highly on investor funding. When Indonesia's central bank (Bank Indonesia) adjusts the BI rate, it may cause direct impact on their funds and profit margins, leading to rapid reactions from the market causing high price volatility changes.
- Speculative retail trading: highly volatile stocks are favorites to short-term traders, causing even more volatile price changes and even unusual market activity alerts from the Indonesia Stock Exchange.

Table 3. LSTM Mean Squared Error Results on Different Lookback Days in Various Banking Sector Stocks

Stock Code	Lookback Days						
	5	10	21	42	63	126	252
BBCA	0.009541	0.011012	0.008585	0.01787	0.009978	0.010056	0.007784
BNII	0.094743	0.071134	0.063749	0.038459	0.033184	0.026862	0.013648
BBRI	0.021959	0.009174	0.007829	0.007929	0.004769	0.016674	0.038632
BMRI	0.007525	0.006334	0.012377	0.005323	0.004683	0.004204	0.011625
BBNI	0.00348	0.004736	0.003072	0.014866	0.007232	0.010639	0.006962
MEGA	0.20788	0.165452	0.149453	0.100245	0.139143	0.146725	0.22415
BANK	0.113246	0.074613	0.051781	0.033309	0.038354	0.041353	0.01484
ARTO	0.088232	0.09258	0.082356	0.074253	0.061199	0.05776	0.014977
BBYB	0.070504	0.024691	0.053616	0.048635	0.026307	0.031811	0.018185

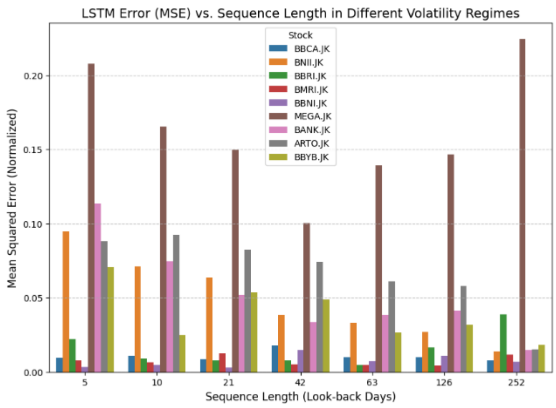

Figure 1. General Overview of LSTM Mean Squared Error on Different Lookback Days in Various Banking Sector Stock

The result displays a high level overview indicating which lookback days period has the lowest error rate across different banking sector stocks in the Indonesia Stock Exchange. On most stocks in the low volatility group, low to mid lookback days produced the lowest error rates of LSTM stock prices prediction. In contrast to that, on the high volatility group, higher lookback days meant better performance of the LSTM model. These results imply that:

- On low volatility stock groups, the LSTM model can learn well even on short time periods of data, as the stock prices trends are relatively stable to discern underlying patterns.
- On high volatility stock groups, a higher window of lookback days are beneficial to increasing accuracy, since underlying patterns are much harder to discern when stock prices trends are fluctuative.

However, the results contains an outlier, BNII. BNII is a member of the low volatility stock group but shows increasingly better results as lookback periods are increased. A possible reason for this outlier phenomenon could be that BNII is not as established as the other banks in the low volatility stock group, hence predicting stock prices on a higher lookback time period will result in better accuracy. Another interesting phenomenon shown in the results graph is that while MEGA bank stock has the lowest LSTM error on 42 days, which is considered a middle time period, the LSTM produces relatively higher error on MEGA compared to other banks. Similar to BNII, since MEGA is not as established as the other banks in the low volatility stock group, stock prices prediction may benefit from a higher lookback window. These outliers phenomenons helps disclose the true underlying nature of the banking sector stocks volatility rate alone cannot express.

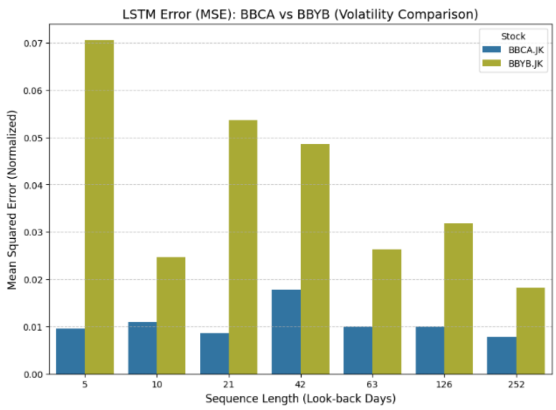

Figure 2. A Comparison of BBCA (Stable) and BBYB (Volatile) LSTM Mean Squared Error Rates across Different Lookback Days

To further show the contrast of low and high stock prices volatility impact to LSTM lookback window selection, two stocks from the end of each spectrum are selected and contrasted. BBCA represents the lowest volatility stock and BBYB represents the highest volatility stock. The LSTM shows the lowest error rate when 21 days were selected as lookback window for BBCA. Considering BBCA is a stable stock, 21 days were expected as a low to medium lookback window period for prices forecasting. This result further proves the implications of low volatility stocks have lower best lookback period for LSTM. BBYB shows good error rate on 10 lookback days, implying good short-term trading capabilities of LSTM. However, the best error rate lies on the longest period of 252 trading days, supporting the implications that high volatility stocks have higher best lookback period for LSTM. The decreasing trend of mean squred error rates as the lookback window increases also further supports the implications.

When two opposite ends of the spectrums are contrasted, clear patterns emerged. However, when all nine stocks are plotted in the same graph, trends are difficult to detect. To solve this issue, stocks are categorized into two main groups defined previously, with 0.5 as the midpoint. The average of the mean squared error from all stocks in each group is calculated to form an aggregated result.

Table 4. LSTM Mean Squared Error Rates on Different Lookback Days based on Stock Volatility Groups

Lookback Days	**MSE based on Volatility Groups**	
	Low	**High**
5	0.057521	0.090661
10	0.044640	0.063961
21	0.040844	0.062584
42	0.030782	0.052066
63	0.033165	0.041953
126	0.035860	0.043641
252	0.050467	0.016001

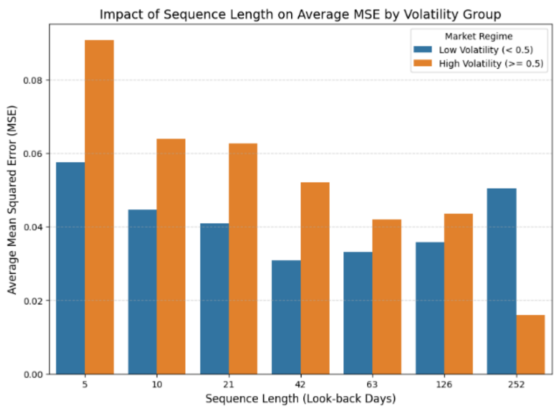

Figure 3. Average LSTM Mean Squared Error on Different Lookback Days in Low and High Volatility Banking Sector Stock Groups

The aggregated result imply that when grouped and averaged, the lowest LSTM mean squared error of the low volatility stock group happens on lower lookback days compared to the high volatility stock group. LSTM shows best performance in the highest lookback period of the high volatility stock group, further supporting this implication. One thing to note is that too little lookback days are also destructive to both the low volatility and high volatility stock groups, meaning hyperparameter fine tuning cannot be ommited in stock prices prediction tasks.

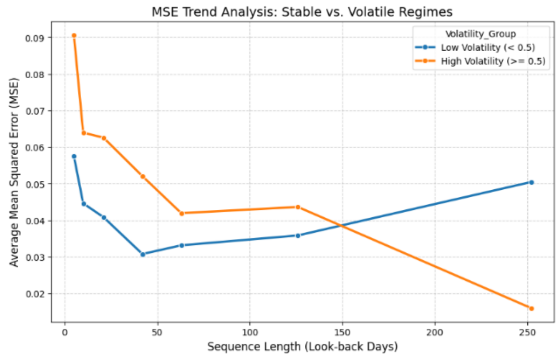

Figure 4. LSTM Mean Squared Error Trend Analysis in Low Volatility and High Volatility Banking Sector Stock Groups

A trend analysis was also conducted to analyze the mean-squared error trends further. The result shows that after slowly decreasing in LSTM error rates, the low volatility group experienced a dip when reaching the optimal point and then slowly rises as lookback period increases. This implies that stable stocks may have patterns that the LSTM finds hard to detect if learning period is too large. In contrast to that, the high volatility stock group has not yet reached its turning point but continues to decrease its error rates as lookback period increases. The high volatility stock group also starts with a higher error rate compared to the low volatility stock group, before managing to overtake and even go lower than the peak point of the low volatility stock group.

IV. CONCLUSION

The results of this study shows that short to medium (5–126 trading days) lookback period produces lower LSTM prediction errors for low volatility stocks. The best performance achieved in this study is observed at the 42 days lookback period resulting in 0.030782 mean squared error. In contrast to that, high volatility stocks exhibit a different pattern, where longer lookback period improves LSTM prediction performance, as demonstrated in this study through a 0.016001 mean squared error in a yearly time period (252 trading days). This contrast suggests that high volatility stocks require longer temporal memory to capture complex and irregular price movements, whereas low volatility stocks are better modelled using shorter and more recent information. However, there are some outlier conditions where this does not apply, supporting the argument of the importance of LSTM hyperparameter tuning for each task-specific conditions.

The key insight of this study is how data volatility affects the best sequence length hyperparameter of LSTM models. This study specificly analyzes the task of stock prices prediction of banking sector stocks in the Indonesia Stock Exchange. Data volatility in this study is defined as the stock prices volatility, while LSTM sequence length is defined as the lookback window period in trading days. Although this study is task specific, key takeaways can be applied to other tasks for future research through adjustments, as data volatility is observable in most tasks.

The limitations of this study is the stock selection is limited to banking sector stocks. Further research can be explored in other sectors of the market and other stock exchanges to achieve hollistic information. Furthermore, this study only used a simple LSTM architecture, which can be improved by replacing it with state of the art LSTM architectures optimized for stock prices prediction.

AUTHOR'S CONTRIBUTION

Conceptualization, Data curation, Formal analysis, Methodology, Validation, Visualization, Writing: Authors.

AVAILABILITY DATA AND MATERIALS

The dataset is provided from Yahoo Finance accessed through the yfinance module in Python.

REFERENCES

Ayitey Junior, M., Appiahene, P., & Appiah, O. (2022). Forex market forecasting with two-layer stacked Long Short-Term Memory neural network (LSTM) and correlation analysis. *Journal of Electrical Systems and Information Technology*, *9*(1). https://doi.org/10.1186/s43067-022-00054-1

Berrajaa, A. (2022). Natural Language Processing for the Analysis Sentiment using a LSTM Model. *(IJACSA) International Journal of Advanced Computer Science and Applications*, *13*(5). www.ijacsa.thesai.org

Cohen, G. (2022). Algorithmic Trading and Financial Forecasting Using Advanced Artificial Intelligence Methodologies. *Mathematics*, *10*(18). https://doi.org/10.3390/math10183302

Deshpande, V. (2023). Implementation of Long Short-Term Memory (LSTM) Networks for Stock Price Prediction. *Research Journal of Computer Systems and Engineering*, *4*(2), 60–72. https://doi.org/10.52710/rjcse.74

El Hajj, M., & Hammoud, J. (2023). Unveiling the Influence of Artificial Intelligence and Machine Learning on Financial Markets: A Comprehensive Analysis of AI Applications in Trading, Risk Management, and Financial Operations. *Journal of Risk and Financial Management*, *16*(10). https://doi.org/10.3390/jrfm16100434

Gong, W. (2024). Research on gold price forecasting based on lstm and linear regression. *SHS Web of Conferences*, *181*, 02005. https://doi.org/10.1051/shsconf/202418102005

Handharu Sworo, T., & Hermawan, A. (2024). Analysis and Prediction of Indonesia Stock Exchange (IDX) Stock Prices Using Long Short Term Memory (LSTM) Algorithm. *Journal of Computer Science and Technology Studies*. https://doi.org/10.32996/jcsts

Joddy, S. (2025). Comparative Analysis of CNN, LSTM, and CNN-LSTM for Indonesian Stock Prediction. *JURNAL EMACS (Engineering, MAthematics and Computer Science)*, *7*(3), 283–289. https://doi.org/10.21512/emacsjournal.v6

Mulyani, S., & Ilham, W. (2025). Application of LSTM-Based Deep Learning for Stock Return Prediction of DCII. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, *9*(4), 2336–2345. https://doi.org/10.33395/sinkron.v9i4.15353

Öztürk, M. M. (2023). Hyperparameter Optimization of a Parallelized LSTM for Time Series Prediction. *Vietnam Journal of Computer Science*, *10*(3), 303–328. https://doi.org/10.1142/S2196888823500033

Pahlevi, M. R., Kusrini, K., & Hidayat, T. (2023). Comparison of LSTM and GRU Models for Forex Prediction. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, *8*(4), 2254–2263. https://doi.org/10.33395/sinkron.v8i4.12709

Soni, P., Tewari, Y., & Krishnan, D. (2022). Machine Learning Approaches in Stock Price Prediction: A Systematic Review. *Journal of Physics: Conference Series*, *2161*(1). https://doi.org/10.1088/1742-6596/2161/1/012065

Wang, J., Hong, S., Dong, Y., Li, Z., & Hu, J. (2024). Predicting Stock Market Trends Using LSTM Networks: Overcoming RNN Limitations for Improved Financial Forecasting. *Journal of Computer Science and Software Applications*. https://www.mfacademia.org/

Xu, Z. (2024). Factors Influencing the Predictive Performance of the LSTM Model on Stock Prices and Its Application in Forecasting U.S. Technology Sector Stock Prices — Sample Time Span, Time Window Length, Feature Selection, and Prediction Days. *ACM International Conference Proceeding Series*, 370–375. https://doi.org/10.1145/3700906.3700964

Yan, K., & Li, Y. (2024). Machine learning-based analysis of volatility quantitative investment strategies for American financial stocks. *Quantitative Finance and Economics*, *8*(2), 364–386. https://doi.org/10.3934/QFE.2024014