Journal of Economics, Finance and Management Studies

ISSN (print): 2644-0490, ISSN (online): 2644-0504

Volume 07 Issue 07 July 2024

Article DOI: 10.47191/jefms/v7-i7-94, Impact Factor: 8.044

Page No: 4777-4782

Sharia Bank of Indonesia Stock Price Prediction using Long Short-Term Memory

Fahmi Poernamawatie[1], I Nyoman Susipta[2], Dwi Winarno[3]

[1,2,3] Gajayana University, Malang

ABSTRACT: This study explored the applicability of Long Short-Term Memory (LSTM) networks for predicting the closing prices of Sharia Bank of Indonesia stock. Utilizing historical data and a rigorous hyperparameter tuning process, the LSTM model demonstrated exceptional accuracy in forecasting closing prices, achieving a Mean Absolute Percentage Error (MAPE) of 2.46%. This significantly surpasses the benchmark for "very good" forecasting performance. The findings underscore the potential of LSTM networks for accurate stock price prediction and provide a foundation for future research to incorporate additional financial indicators and explore more complex model architectures.

KEYWORDS: Stock Price Prediction, LSTM, Forecasting, Sharia Bank of Indonesia, Hyperparameter Tuning

I. INTRODUCTION

Equity instruments, commonly referred to as stocks, represent ownership units within a corporation. A company's share price can serve as a barometer of its overall performance. Sharia Bank of Indonesia, known as PT BSI Tbk, established in February 2021, is a prominent Indonesian Islamic bank resulting from the merger of three prior Islamic institutions: PT Bank Mandiri Syariah Tbk, PT Bank Rakyat Indonesia Syariah Tbk, and PT Bank Negara Indonesia Syariah Tbk (Hartini & Marhandrie, 2022). Since its inception, PT BSI Tbk's stock price has exhibited substantial volatility over the past two years. This phenomenon can potentially be ascribed to an initial overvaluation of its shares. Overvaluation is a situation where a stock's market price surpasses its intrinsic value, leading to inflated trading multiples (Nugraha & Sulasmiyati, 2017). A visual representation of PT BSI Tbk's stock price movements is depicted in Figure 1.

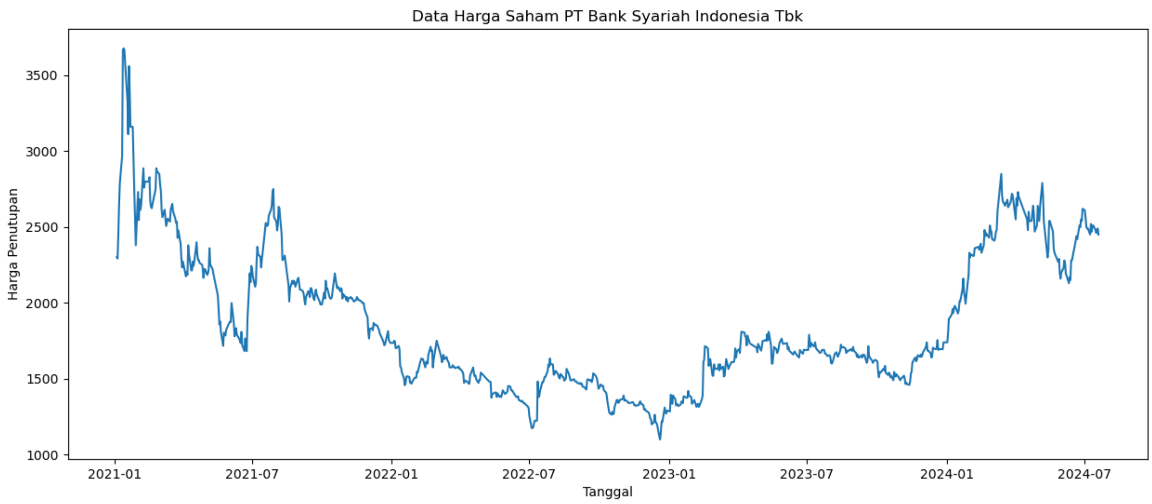

Figure 1. Closing prices of PT BSI Tbk.

The development of stock price models holds significant importance in the endeavor to forecast the future value of PT BSI Tbk shares. The generation of accurate predictions through such models can prove instrumental for both the company itself and its investors. From the company's perspective, these forecasts can serve as valuable insights guiding strategic decision-making processes aimed at mitigating stock price volatility. For investors, the models provide crucial input for informed investment decisions, including determinations regarding purchases, sales, or the retention of existing holdings within their portfolios.

Stock price data is inherently time-series in nature, making it amenable to forecasting techniques that utilize historical price data to predict future prices. This approach, supported by extensive research (Idrees et. al, 2019, Siami-Namini et. al, 2018, Gao et. al., 2018), involves analyzing the temporal patterns and trends within the data to establish predictive models. While ARIMA (Autoregressive Integrated Moving Average) has been employed for stock market prediction using time-series data (Idrees et al., 2019), it is not without its limitations such as its reliance on linear assumptions and its inability to capture complex market dynamics, especially during volatile periods or structural changes. Therefore, in this research, we employ Long Short-Term Memory (LSTM) networks, a type of recurrent neural network known for its ability to learn long-term dependencies and handle non-linear relationships in time series data.

The contribution of this research can be described as follows:

1. This research furthers the domain of stock price prediction by incorporating a meticulously curated dataset encompassing the most recent market intelligence. This facilitates the capture of the latest trends and market dynamics, demonstrably crucial for achieving superior forecasting accuracy.
2. This research delves into a comprehensive exploration of hyperparameter variations. This in-depth analysis aims to identify the model configuration that best aligns with the unique characteristics and complexities inherent to the most recent market data.

II. METHOD

DATASET COLLECTION

This research utilizes a dataset encompassing the historical closing prices of PT BSI Tbk curated from https://finance.yahoo.com/ since its initial public offering (IPO) on February 1st, 2021, extending through July 19th, 2024. The objective is to leverage this historical data to construct a predictive model for future closing prices. In essence, a single feature, representing the closing price, is employed for target variable prediction. Subsequently, the dataset undergoes a standard train-test split, allocating 80% of the data for model training and the remaining 20% for testing purposes.

DATA PREPROCESSING

Following data curation, meticulous preprocessing is essential to ensure the suitability of historical closing price data for subsequent analysis, particularly with Long Short-Term Memory (LSTM) networks. This preprocessing encompasses meticulous data cleaning, rigorous format standardization, and strategic normalization:

1. Data Cleaning: This initial stage involves the identification and rectification of potential missing values. A common approach, such as mean imputation, can be employed to address these missing data points.
2. Format Standardization: Here, data consistency is paramount. This may involve ensuring a uniform date format, verifying the data type of each column, and potentially addressing any inconsistencies in data representation.
3. Normalization: Given the inherent range of closing price data (e.g., 900-3700), normalization becomes crucial. LSTMs often exhibit improved training efficiency and convergence when the data is scaled to a specific range. This normalization ensures that all features contribute proportionally during the training process. Here, the data is normalized in (-1,1) range.

LSTM MODELLING

Initially proposed in (Hochreiter & Schmidhuber, 1997), LSTM are a type of Recurrent Neural Network (RNN) specifically designed to handle sequential data, like the stock prices. Their strength lies in capturing both long-term dependencies and short-term trends within the data, making them suitable for time series forecasting tasks. Figure 2 depicts the illustration of an LSTM module.

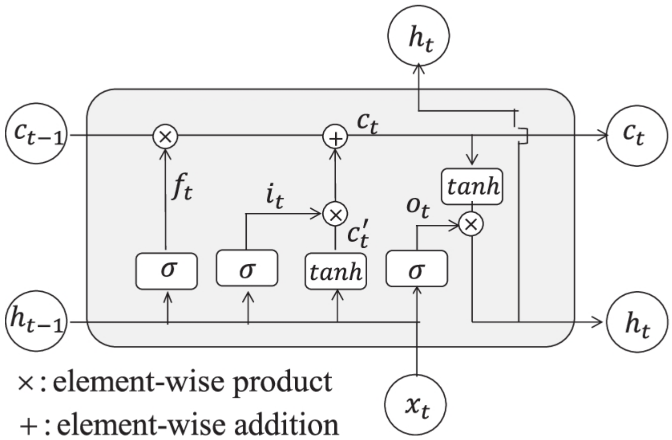

Figure 2. LSTM Module

LSTM are designed to mitigate the vanishing gradient problem. This problem hinders traditional RNNs in capturing long-term dependencies within sequential data. LSTMs address this by employing a gating mechanism that regulates information flow through the network. There are several core components of LSTM:

1. Cell State (C_t): The central element of an LSTM unit is the cell state. It acts as a memory lane, persisting information across sequential time steps (t). Unlike standard RNNs, LSTMs can selectively retain or discard information within the cell state.
2. Gates: These are specialized neural network layers that leverage sigmoid activation functions. Sigmoid functions map values between 0 and 1, effectively acting as gates that control information flow. LSTMs primarily utilize three distinct gates:
   a. Forget Gate (f_t): This gate determines the information to be forgotten from the preceding cell state (C_{t-1}). It takes the prior hidden state (h_{t-1}) and the current input (x_t) as inputs. The output is a vector of values between 0 and 1, with each element corresponding to a specific element in the cell state. Values closer to 1 signify "retain this information," while values closer to 0 signify "forget this information."
   b. Input Gate (i_t): This gate selects the pertinent new information from the current input (x_t) to be stored in the cell state. Similar to the forget gate, it takes ht-1 and xt as inputs and generates a vector of values between 0 and 1. Additionally, it creates a new "candidate state" ($C_t{}'$) that encompasses potential novel information.
   c. Output Gate (o_t): This gate determines what information from the current cell state (C_t) will be emitted as the hidden state (h_t) for the subsequent time step. It takes h_{t-1} and x_t as inputs and produces a vector of values between 0 and 1 for each element in C_t. This vector is then combined with a specific activation function (often tanh) to generate the final output gate.

Mathematically, the process inside of the LSTM model can be described as follow:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right)$$
$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$
$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$$
$$c_t = f_t \cdot c_{t-1} + i_t \cdot \sigma(W_c x_t + U_c h_{t-1} + b_c)$$
$$h_t = o_t \cdot \sigma(c_t)$$

Through these gates, LSTMs achieve controlled information flow: The forget gate acts as a filter, discarding irrelevant data from the cell state. The input gate selects valuable information from the current input and incorporates it into the cell state. The output gate determines the processed information from the cell state that is sent as output to the next time step. By selectively remembering and forgetting information, LSTMs can effectively learn long-term dependencies within sequential data sequences. This capability makes them powerful tools for various tasks, including machine translation, speech recognition, and time series forecasting.

MODEL TRAINING

In order to identify the model that achieves optimal generalization performance on the target data, a comprehensive hyperparameter search strategy was employed. This search strategy involved the systematic evaluation of a diverse range of hyperparameter configurations. The specific hyperparameters explored included epoch, sequence length, LSTM layers, hidden dimension, learning rate as detailed in table 1. In all of the experiments, Adam optimizer (Kingma & Ba, 2014) are used in the training process.

Table 1. Hyperparameters

Hyperparameters	Value Used in Experiment
Epoch	[20, 40, 60]
Sequence Length	[40, 60, 80]
LSTM Layers	[1, 2, 4]
Hidden Dimension	[16, 64, 128]
Learning Rate	[0.01, 0.001]

EVALUATION

Once trained using the 80% dataset, the model's performance will be evaluated on the 20% held-out test data. This assessment will involve metrics like Mean Absolute Percentage Error (MAPE) to gauge the model's effectiveness in predicting future closing prices. Mean Absolute Percentage Error (MAPE) is a widely used metric for evaluating the accuracy of forecasting models. The MAPE score can be calculated using the following equation where y_i is the actual data and $\widehat{y_i}$ is the prediction.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\widehat{y_i} - y_i}{y_i}\right|$$

It measures the average magnitude of the percentage errors between the predicted values and the actual values. A lower MAPE value indicates better predictive performance, as it suggests that the model's predictions are on average closer to the actual values.

Chang et al. (2007) proposed a classification scheme for interpreting MAPE values:

a. MAPE < 10%: Very good model performance
b. 10% ≤ MAPE ≤ 20%: Good model performance
c. MAPE > 20%: Poor model performance

III. RESULTS AND DISCUSSION

Following extensive training with a multitude of hyperparameter combinations meticulously outlined in Table 1, the LSTM model underwent a rigorous evaluation process. This evaluation leveraged a specifically curated 20% held-out test set, designed to assess the model's generalizability to unseen data. The test set encompassed real-world closing price data, spanning a substantial timeframe from June 11th, 2023, to July 19th, 2024. This time period incorporates a significant amount of market activity, allowing the model to be tested on diverse market conditions. To evaluate the model's accuracy, Mean Absolute Percentage Error (MAPE) was employed. MAPE measures the average discrepancy between the model's predicted closing prices and the actual closing prices, expressed as a percentage. A lower MAPE score signifies a higher degree of accuracy.

Table 2. MAPE Score for Several Hyperparameters Combinations

Sequence Length	Layer	Hidden Dimension	Epoch	Learning Rate	Test Set MAPE
40	1	64	60	0,01	0,024574121
40	1	16	20	0,001	0,18057415
40	2	64	20	0,001	0,2400006
60	1	128	40	0,001	0,17135365
60	2	128	20	0,01	0,12655123
60	4	128	60	0,001	0,1003337
80	1	16	20	0,001	0,16230309
80	4	64	20	0,01	0,23017143
80	4	128	60	0,001	0,10663771

Table 2 meticulously details the MAPE scores achieved by the various LSTM configurations tested during the hyperparameter tuning process. A meticulous analysis of Table 2 reveals that the optimal hyperparameter combination emerged after a comprehensive search. This winning configuration utilized a sequence length of 40. This implies that the model achieved the best performance by considering the past 40 closing prices when making predictions. This finding suggests that incorporating a historical window of this size provides the model with the most relevant information for accurate forecasting. The optimal configuration also employed a single LSTM layer with 64 hidden units. This architecture strikes a balance between model complexity and efficiency. Furthermore, a learning rate of 0.01 and 60 epochs of training were identified as the most suitable parameters for this specific task. The learning rate controls the magnitude of updates made to the model's weights during training, and 0.01 appears to be the optimal value for this scenario. Epochs signify the number of times the entire training dataset is passed through the model, and 60 epochs were found to be sufficient for the model to converge and achieve its best performance. This configuration achieved a remarkable MAPE score of 0.0246, which translates to a minuscule error of approximately 2.4%. The prediction result can be seen on figure 3.

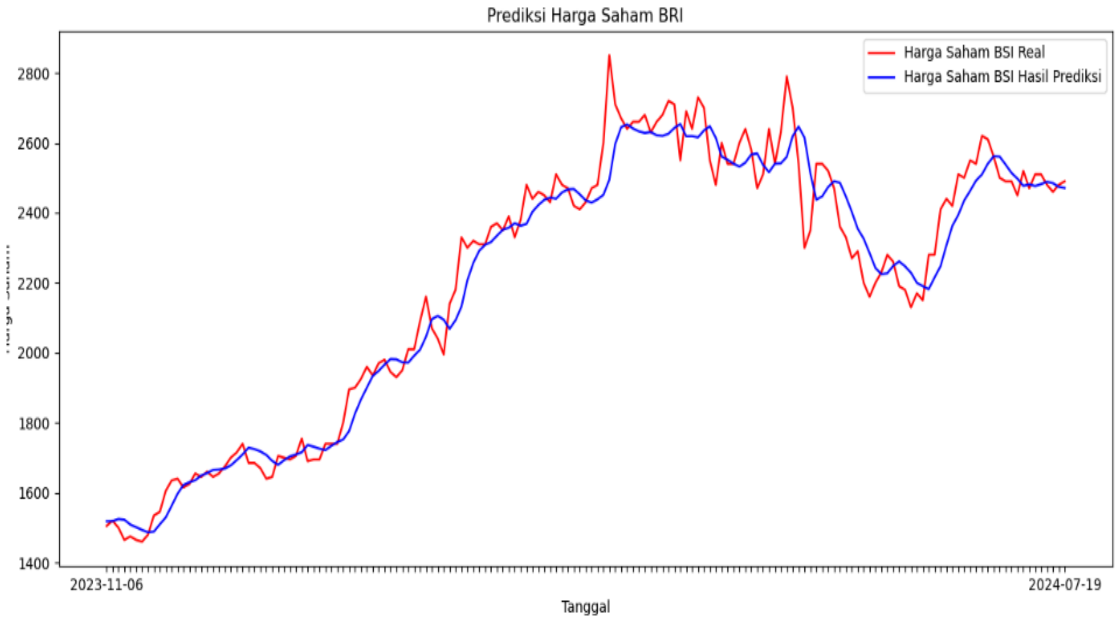

Figure 3. Prediction Result on The Test Set

This achievement is particularly noteworthy when considering the benchmark established by Chang et al. (2007). Their research suggests that a MAPE score below 10% signifies very good forecasting performance. In this context, the proposed LSTM model surpasses this benchmark by a significant margin. This outcome demonstrates a high degree of accuracy in predicting closing prices. With a MAPE score of only 2.46%, the model achieves a level of performance that can be confidently classified as "very good" according to established benchmarks. This success highlights the effectiveness of the hyperparameter tuning process and the LSTM architecture in creating a robust and accurate model for predicting closing prices.

IV. CONCLUSION

This study investigated the effectiveness of Long Short-Term Memory (LSTM) networks in predicting the closing prices of Sharia Bank of Indonesia stock. The LSTM model was trained on historical closing price data and evaluated on a held-out test set spanning a significant period (June 11th, 2023 - July 19th, 2024) to assess generalizability. Mean Absolute Percentage Error (MAPE) served as the evaluation metric, with lower scores indicating higher accuracy.

A rigorous hyperparameter tuning process identified an optimal configuration that achieved a MAPE score of 2.46%. This translates to a very high degree of accuracy, exceeding the benchmark of "very good" forecasting performance established by Chang et al. (2007) with a MAPE threshold of 10%. The success highlights the LSTM architecture's capability, along with the effectiveness of the hyperparameter tuning, in creating a robust and accurate model for predicting Sharia Bank of Indonesia's closing stock prices.

Future research directions could involve exploring the impact of incorporating additional financial indicators alongside historical closing prices. Additionally, investigating more complex LSTM architectures or ensemble methods with other forecasting techniques might be fruitful avenues for further improvement.

REFERENCES

1) Hartini, H., & Marhandrie, D. (2022). Pengaruh Profitabilitas, Risiko Finansial dan Harga Saham Terhadap Nilai Perusahaan Bank Syariah Indonesia (BSI) Di BEI Periode Tahun 2014-2021. Jurnal Ilmiah Ekotrans & Erudisi, Vol. 2, No. 1, Hal. 104-111

2) Nugraha, E. S., & Sulasmiyati, S. (2017). Analisis Nilai Intrinsik Saham Dengan Relative Valuation Techniques (Studi Pada Perusahaan Sub Sektor Rokok yang Terdaftar di Bursa Efek Indonesia Periode 2013–2016). Jurnal Administrasi Bisnis, Vol. 52, No. 1, Hal. 106-113.
3) S. M. Idrees, M. A. Alam, and P. Agarwal, "A Prediction Approach for Stock Market Volatility Based on Time Series Data," IEEE Access, vol. 7, pp. 17287–17298, 2019
4) S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A Comparison of ARIMA and LSTM in Forecasting Time Series," Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018, pp. 1394–1401, Jan. 2019
5) T. Gao, Y. Chai, and Y. Liu, "Applying long short term momory neural networks for predicting stock closing price," Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS, vol. 2017-November, pp. 575–578, Apr. 2018
6) Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
7) Chang, P. C., Wang, Y. W., and Liu, C. H. (2007). The development of a weighted evolving fuzzy neural network for PCB sales forecasting. Expert Systems with Applications, Vol. 32, No. 1, Hal. 86-96.