Quantitative Economics and Management Studies (QEMS) *ISSN: 2722-6247 (online)*
Vol. 3 No. 6 (2022)
https://doi.org/10.35877/454RI.asci1118

BUMI Stock Price Prediction Using Long Short Term Memory (LSTM) With Three Hyperparameter Tuning Regression

Hartono

Widya Dharma University Pontianak, Indonesia

Abstract

Shares is one of investment with moderate to high level of risk profile and return rate, and is a legal ownership proof of a company (Ltd). To obtain an optimum return with possible low risk, price of shares can be predicted or analyzed via fundamental or technical analysis. One of the most popular technical analysis is by employing computer algorithms as a device to help making decision. Long short term memory (LSTM), as a deep improvement of recurrent neural networks (RNN), is a commonly used method to predict shares price. BUMI, as one of shares in the energy sector, is the most concerned shares due to the geopolitical conflict in Europe. This research aims to predict BUMI shares price using LSTM by considering three parameters, i.e. number of hidden layers, epoch, and learning rate. Linear regression is used to obtained optimum values of each parameter whichs further combined to yield the best performance of the algorithm. Based on the analysis, the combination of minimum and optimum parameters results in the best performance by RMSE value.

Keywords: LSTM; BUMI; shares; forecasting.

1. Introduction

Stock is an investment instrument with a medium to high risk profile and returns. According to Hermuningsih (Hermuningsih, 2012). Shares are securities that show the ownership of organizations, institutions, or individuals to a limited liability company (PT), both open and closed. In investing and trading stocks, institutions, organizations, and individuals can earn tens or even hundreds of times the value invested.

In the process of stock analysis, generally two types of approaches are used, namely fundamental and technical. Stock fundamental analysis is a method of analyzing stock potential by looking at the fundamental aspects of the company behind the shares being traded. This aspect can include the type of business, market capitalization, number of outstanding shares, level of liquidity, market valuation, and so on. In contrast, technical analysis is only based on data on stock price movements. Technical analysis can be done using stock price candle patterns, using indicators, or using programming language algorithms. The application of programming language algorithms in stock technical analysis is very beneficial because the automation process of historical checking can be carried out accurately (Purnama et al., 2017).

Programming algorithms (algo-trading) are growing rapidly in their use to predict stock price movements. One of the popular algorithms used is long short term memory (LSTM). This algorithm is included in the recurrent neural network (RNN) algorithm with multilevel learning layers to improve learning accuracy results. A comparative study of the regression method, LSTM and GRU has been carried out by (Sofi et al., 2021).

Various types of shares are traded on the Indonesia Stock Exchange (IDX), covering categories of raw, non-primary, primary, energy, finance, health, industry, infrastructure, property, technology, and transportation. The energy sector is a sector that is currently receiving attention due to global geopolitical conditions related to the conflict in Europe. One of the shares in the energy sector that is currently being considered is BUMI shares, which is proof of ownership of PT.

* Corresponding author
E-mail address: hartono@widyadharma.ac.id

BUMI Resources Tbk. One of the studies using LSTM to predict stock prices in the mining sector has also been carried out by (Julian & Pribadi, 2021). Agusta et al. (2021) using the LSTM algorithm to predict stock prices in the pharmaceutical sector. Arfan, & Lussiana, (2020) comparing support vector regression (SVR) and LSTM algorithms to predict stocks from the industrial, health, and non-primary sectors. Siahaan, (2017) has also combined the SVR and LSTM methods to assess the price of the composite stock price index (JCI).

The LSTM algorithm has been widely studied, but the process of determining the optimal parameters in the model is a very important part to obtain measurable model predictions. Determination of parameters by trial and error can be used, but will consume large computational resources and take a long time. In this study, a hyperparameter tuning method using three-parameter LSTM regression was studied. The three parameters include the number of hidden layers, epochs, and learning rate. Each parameter is optimized and then combined to get the best performance results.

2. Literature Review

2.1. Stocks

According to Hermuningsih, (2012), Stocks are securities that show the ownership of organizations, institutions, or individuals to a limited liability company (PT), both open and closed. Shares are an asset or investment instrument with a moderate to high risk profile with moderate to high returns, depending on the type of stock being considered. The returns that may be obtained from investing in stocks can range from tens to hundreds of times the value invested.

In Indonesia, stock trading activities are carried out on the Indonesia Stock Exchange (IDX), Jakarta. Shares can be classified into several general categories according to the line of business carried out by the related company, including raw, non-primary, primary, energy, finance, health, industry, infrastructure, property, technology, and transportation. One of the stocks in the energy category is BUMI. These shares are proof of ownership in PT. BUMI Resources Tbk.

PT. BUMI Resources Tbk. was established on June 26, 1973 and is engaged in the exploration and exploitation of coal and oil mine sites. The company's shares have been listed on the IDX since July 30, 1990, and currently has a market capitalization of around Rp. 14.4 trillion with the number of outstanding shares reaching 135 billion shares. At the time this research was conducted, the value of BUMI's shares was valued at Rp. 107 per share. In a period of five years, the highest price for this share was Rp. 338 per share.

In technical analysis of stock prices, especially using algorithms based on artificial neural networks, a sufficient amount of data is required. Limited data can cause the model built to have an overfit condition so it is not good for predicting data (Baek & Kim, 2018). This overfit condition can occur if the model has been trained in such a way that it follows the training data so that it can only predict based on the data that has been given (Mutasa et al., 2020).

2.2. Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) is one of the artificial neural network algorithms with a model architecture as shown in Figure 1. The input layer contains cells in charge of reading the data to be analyzed. Furthermore, each cell in the input layer will send information to the next layer, namely the hidden layer. Inside the hidden layer there are also several cells that are in charge of converting the input into something that can be used by the output layer. In this hidden layer the hidden features and patterns of the data are described. Furthermore, the processed cells in the hidden layer are forwarded to the output layer. However, before being used, this information is returned to the hidden layer to find a better solution. This scheme is a differentiating factor for the RNN method compared to other artificial neural network algorithms, where there is repetition/return of the process to get better results.

2.3. Long-Short Term Memory (LSTM)

In 1997, Schmidhuber Schmidhuber, (1997) proposed a development method of RNN called Long Short Term Memory (LSTM). The basic difference between LSTM and RNN is the number of hidden layers used. In the basic RNN, the number of hidden layers used is one layer. Meanwhile, the number of hidden layers in the LSTM is not limited according to the training needs to be carried out (Faurina et al., 2018).

Figure 2 shows the architecture of the hidden layer LSTM model. C_{t-1} and Ct are long-term memory in the (t-1) and t-th hidden layers. The variables xt are inputs, h_{t-1} and h_t can be short memory in the (t-1) and t-th hidden layers, and y_t is the output. In general, there are three gate stages in an LSTM, namely forget gate (f_t), input gate (i_t), and output gate (o_t). Each gate has its own task, where the forget gate functions to delete information from the cell, the input gate functions to determine and update the input value in memory, and the output gate functions to get the output of the input value stored in the cell.

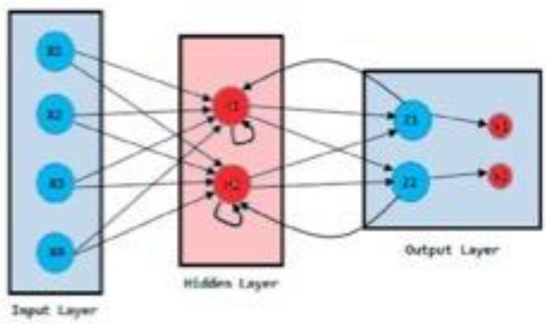


Figure 1. RNN model architecture

In Figure 2, the forget gate (f_t) will determine what information will be removed from the cell by assigning a value of 0 to not pass any information and a value of 1 to forward all received information. Forget gate is calculated based on Equation (1). The second process is the input gate (i_t) which functions to regulate the information to be stored in the cell. This process begins with selecting the information to be updated from the input gate, then adding new information from the ground layer. The input gate is calculated according to Equation (2), the calculation of the tanh layer is calculated based on Equation (3), and the addition of the tanh layer to the input gate which is then used to update the cell state is calculated based on Equation (4). The next step is to determine the output (o_t) obtained from the previous cell state which is processed through the soil layer and multiplied by the sigmoid factor (σ). The output obtained can be issued as output y_t or used as short-term memory (short term) in the next hidden layer. Likewise, C_t is used as long-term memory for the next hidden layer.

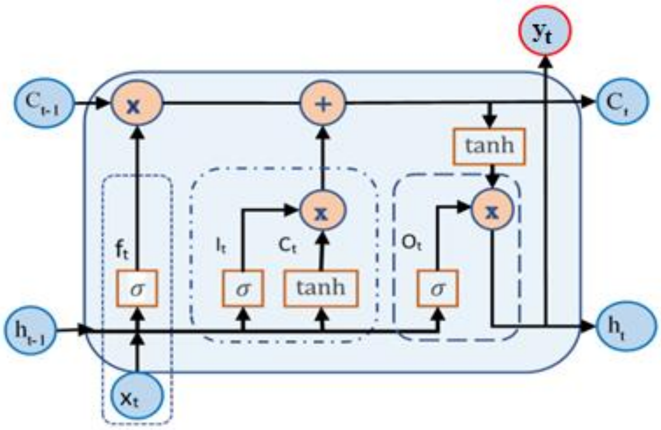


Figure 2. LSTM hidden layer model architecture

$$f_t = \sigma\left(W_{i_f}X_t + W_{hf}h_{(t-1)} + b_f\right) \quad (1)$$

$$i_t = \sigma\left(W_{ii}X_t + W_{hi}h_{(t-1)} + b_i\right) \quad (2)$$

$$g = tanh\left(W_{ig}X_t + W_{hg}h_{(t-1)} + b_g\right) \quad (3)$$

$$C_t = (f_t C_{t-1}) + (i_t g_t) \quad (4)$$

$$o_t = \sigma\left(W_{io}X_t + W_{ho}h_{(t-1)} + b_o\right) \quad (5)$$

$$h_t = o_t(tanh(\,c)_t) \quad (6)$$

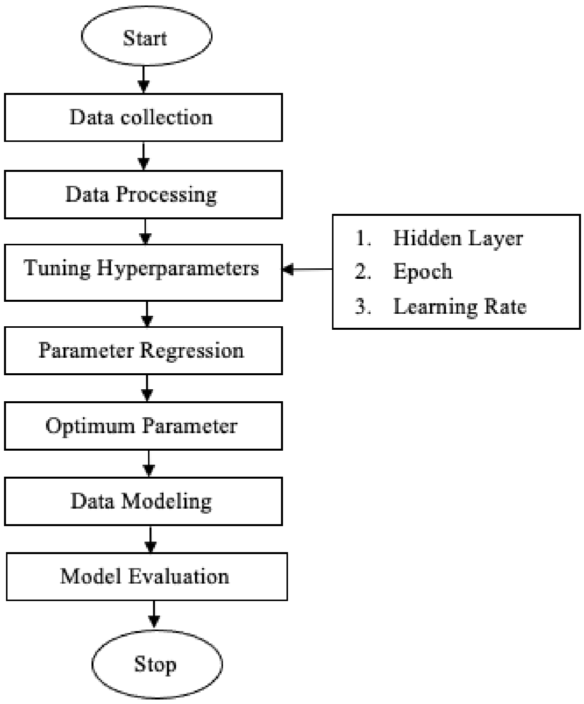
$$y_t = W_{yt}h_t + b_y \quad (7)$$

2.4. Root Mean Squared Error (RMSE)

Root mean squared error (RMSE) is one of the popular statistical instruments used to identify the extent of error or error in predicting the data against the original data. The RMSE value can be calculated using Equation (8). In addition to being presented in RMSE numerical values, prediction errors can also be described in the form of error graphs.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\overline{x_i} - x_i)^2}{N}} \qquad (8)$$

3. Research Method and Materials



The research flow used is shown in

Figure **3**. The study began with collecting data on BUMI's stock price movements. Prior to the analysis, the data obtained were cleaned first through a filtering process. After the data is cleaned, the hyperparameter tuning process is carried out to obtain the optimum parameter values. Regression model is used to obtain the optimum value of each parameter being reviewed. The optimum parameter candidates that have been obtained are then used in the data modeling process and finalized by model evaluation.

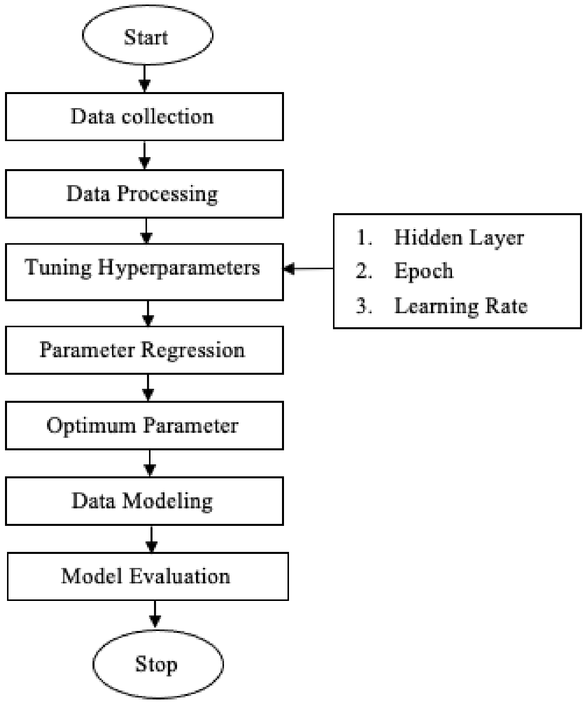


Figure 3. Research flow

3.1. BUMI Stock Dataset

BUMI's stock dataset was obtained from the Yahoo Finance website. The data used covers the movement of a five-year span, from July 1, 2017 to Jul 31, 2022, in a daily timeframe. The data obtained contains information on the date, opening price (open), highest price (high), lowest price (low), closing price (closed), adjusted closing price (adjusted close). In this study, the feature used for analysis is the closing price (close).

The data obtained is then filtered to remove invalid data, such as null data. After the filtering process, the length of the data is 1,274 (data on the day of recording). From this data, the proportion of 90% of training data and 10% of testing data is carried out. BUMI stock price movements within the selected range can be presented in

Figure **4**.

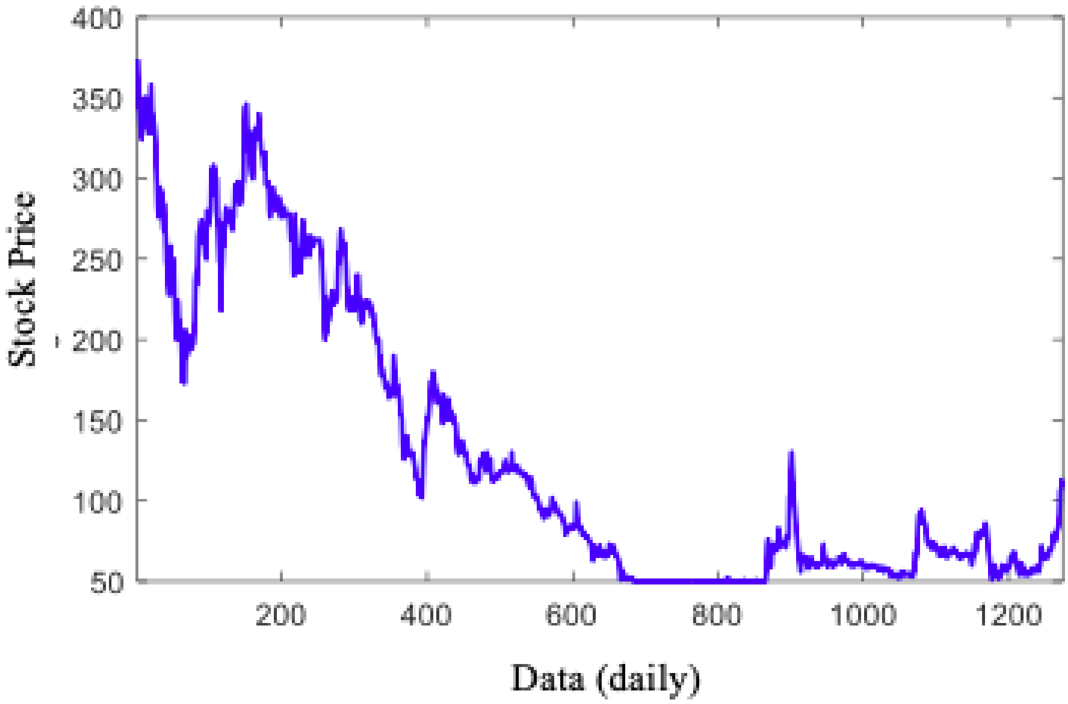


Figure 4. BUMI stock price movements in the selected five years

3.2. Tuning Hyperparameter Schematic

Hyperparameter is a parameter that will determine the optimization of the results of the analysis or prediction using the LSTM algorithm. In this study, the optimization scheme of each parameter is used, namely the number of hidden layers, the number of epochs, and the learning rate. The range used for each parameter is shown in Table 1. At each optimization stage of each parameter, the other parameters are locked with a fixed value of 20% of the maximum value. Each parameter is optimized using linear regression to get the optimum value of each parameter based on the RMSE value. In addition, the parameter with the smallest RMSE value is also taken into consideration in the final analysis.

Table 1. Parameter ranges in the tuning hyperparameter scheme

Parameter	Minimum value	Maximum value	Locked value
Hidden layer	1	25	5
Epoch	10	250	50
Learning rate	0,005	0,125	0,025

3.3. Final Analysis

Based on Table 1, for each parameter a dataset of parameters is obtained which will be considered in the final analysis. Thus obtained a total of eight dataset combinations. These eight combinations were re-analyzed to get the best performance based on the RMSE value. Comparison of the graph of BUMI's stock price movement and its predictions are then displayed based on the combination of these parameters.

4. Results and Discussion

Table 2 shows the performance of the LSTM algorithm with hidden layer parameters. The minimum RMSE value obtained is 4.1928 which is obtained from the combination of parameters for the number of hidden layers 22. The RMSE results shown in the table are then processed to obtain a regression function as shown in Figure 5. Based on the obtained regression function, the optimum value of RMSE in the hidden layer as much as 16.

Table 2. LSTM algorithm performance with hidden layer parameters

Hidden Layer	Epoch	Learning Rate	RMSE	Information
1	50	0,0250	76,2190	-
2	50	0,0250	18,3460	-
3	50	0,0250	18,4311	-
4	50	0,0250	5,2195	-
5	50	0,0250	4,2281	-
6	50	0,0250	4,6345	-
7	50	0,0250	5,6207	-
8	50	0,0250	6,6197	-
9	50	0,0250	5,0119	-
10	50	0,0250	5,7694	-
11	50	0,0250	6,2923	-
12	50	0,0250	4,9259	-
13	50	0,0250	5,2468	-
14	50	0,0250	5,8534	-
15	50	0,0250	4,8020	-
16	50	0,0250	4,7130	optimum
17	50	0,0250	5,6542	-
18	50	0,0250	5,9813	-
19	50	0,0250	4,6456	-
20	50	0,0250	4,5594	-
21	50	0,0250	5,2446	-
22	50	0,0250	4,1928	minimum
23	50	0,0250	4,8871	-
24	50	0,0250	5,0609	-
25	50	0,0250	4,4972	-

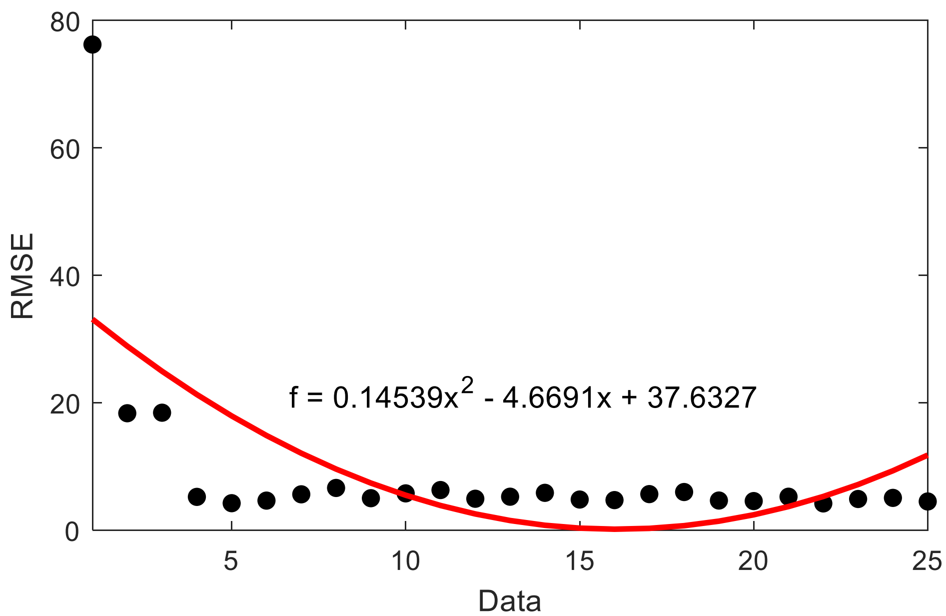


Figure 5. LSTM performance regression with hidden layer parameters

Table 3 shows the performance of the LSTM algorithm with the epoch parameter. The minimum RMSE value is obtained at the number of epochs of 240. Based on the regression equation shown in Figure 6, the optimum number of epochs is 160.

Table 3. Performance of LSTM algorithm with epoch parameter

Hidden Layer	Epoch	Learning Rate	RMSE	Information
5	10	0,0250	24,3469	-
5	20	0,0250	10,1422	-
5	30	0,0250	6,9108	-
5	40	0,0250	5,1065	-
5	50	0,0250	4,2281	-
5	60	0,0250	4,0576	-
5	70	0,0250	3,9111	-
5	80	0,0250	3,6243	-
5	90	0,0250	3,5636	-
5	100	0,0250	3,5189	-
5	110	0,0250	3,4250	-
5	120	0,0250	3,3943	-
5	130	0,0250	3,3698	-
5	140	0,0250	3,3404	-
5	150	0,0250	3,3449	-
5	160	0,0250	3,2754	Optimum
5	170	0,0250	3,3210	-
5	180	0,0250	3,3141	-
5	190	0,0250	3,3028	-
5	200	0,0250	3,3386	-
5	210	0,0250	3,2653	-
5	220	0,0250	3,3179	-
5	230	0,0250	3,2398	-
5	240	0,0250	3,2154	Minimum
5	250	0,0250	3,2211	-

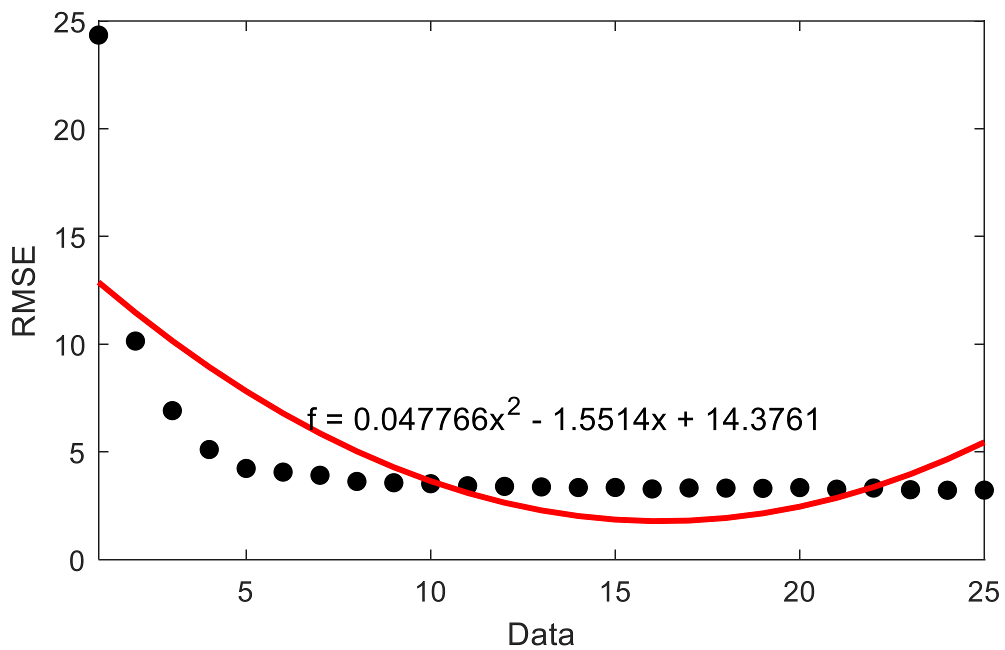


Figure 6. Regression of LSTM performance with parameter epoch

Table 4 shows the performance of the LSTM algorithm with the learning rate parameter. Based on the table, the minimum RMSE value is obtained at a learning rate of 0.1200. Figure 7 provides information that the optimum RMSE value is at a learning rate of 0.080.

Table 4. Performance of LSTM algorithm with learning rate parameter

Hidden Layer	Epoch	Learning Rate	RMSE	Information.
5	50	0,0050	24,6565	-
5	50	0,0100	9,0833	-
5	50	0,0150	6,1674	-
5	50	0,0200	4,9187	-
5	50	0,0250	4,2281	-
5	50	0,0300	4,0952	-
5	50	0,0350	4,1666	-
5	50	0,0400	3,8892	-
5	50	0,0450	4,0771	-
5	50	0,0500	3,8167	-
5	50	0,0550	3,8521	-
5	50	0,0600	3,9368	-
5	50	0,0650	4,2493	-
5	50	0,0700	3,7885	-
5	50	0,0750	3,7014	-
5	50	0,0800	4,2793	Optimum
5	50	0,0850	4,7555	-
5	50	0,0900	3,7162	-
5	50	0,0950	3,7858	-
5	50	0,1000	3,9891	-
5	50	0,1050	3,4985	-
5	50	0,1100	4,1314	-
5	50	0,1150	4,4049	-
5	50	0,1200	3,4379	Minimum
5	50	0,1250	3,6083	-

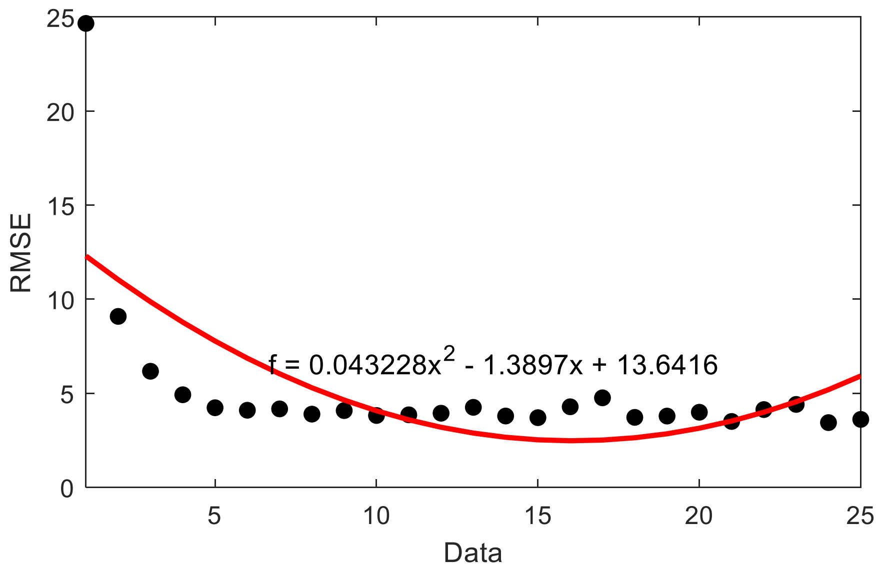


Figure 7. LSTM performance regression with learning rate parameter

Table 5 shows the performance of the LSTM algorithm with the optimum parameters that have been obtained. Therefore, for each parameter the value of the minimum RMSE and the optimum RMSE is considered, a number of eight optimum parameter combinations are obtained that can be used. Based on the table, the best value for hidden layer 2 is obtained, the number of epochs is 160, and the learning rate is 0.1200. These results indicate that not entirely the combination of the minimum or optimum RMSE values produces the best RMSE values. In this case the number of hidden layers and the best learning rate are values based on the minimum RMSE value, while the number of epochs is the contribution of the optimum RMSE parameter.

Table 5. LSTM algorithm performance with optimum parameters

Hidden Layer	Epoch	Learning Rate	RMSE
16	160	0,0800	3,4159
16	160	0,1200	3,6157
16	240	0,0800	3,1247
16	240	0,1200	3,3811
22	160	0,0800	3,1196
22	160	0,1200	3,0438
22	240	0,0800	3,2979
22	240	0,1200	3,0503

Figure 8 shows a comparison of observational data and predictions of BUMI's stock price using the optimum parameters from Table 5. It can be seen that the training model that has been developed successfully predicts the movement of BUMI's stock price on the specified testing data. Thus, this model can be used as a basis for making decisions to sell or buy BUMI shares. Figure 9 shows a graph of the model prediction error when compared to the testing data.

As an important note, the testing data in this study is 127 data, so the use of this model is valid enough to predict stock prices for the next three months, namely until November 2022 or towards the end of the year. After passing this timeframe, this model needs to be reconsidered according to price movements in the last five years.

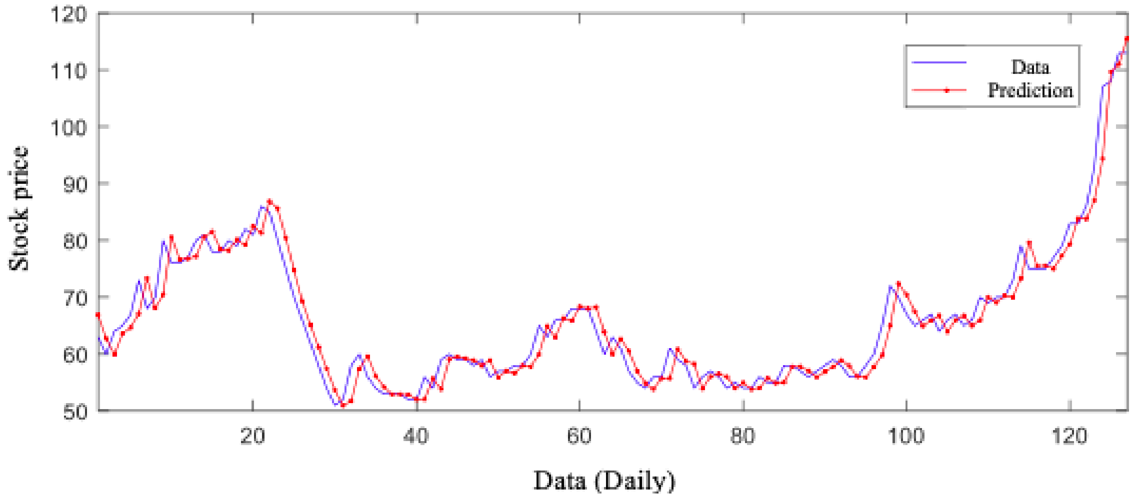


Figure 8. Comparison of observational data and predictions of BUMI's stock price

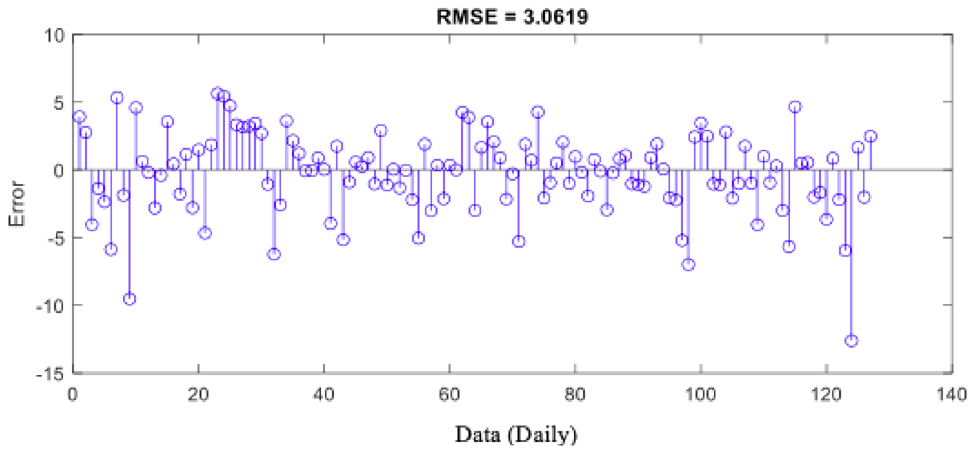


Figure 9. Comparison of observational data and predictions of BUMI's stock price

5. Conclusion

Based on the results of the analysis and discussion, the following conclusions are drawn; (1) The optimal value of each parameter in the hyperparameter tuning process can be done separately based on the minimum and optimum RMSE values; (2) The best performance of the LSTM algorithm to predict BUMI's stock price is obtained from a combination of parameters based on the minimum and optimum RMSE values, with a final RMSE of 3.0438; (3) The LSTM algorithm can be used to predict BUMI's stock price well.

References

Agusta, A., Ernawati, I., & Muliawati, A. (2021). Prediksi Pergerakan Harga Saham Pada Sektor Farmasi Menggunakan Algoritma Long Short-Term Memory. *Informatik: Jurnal Ilmu Komputer*, *17*(2), 164-173. doi: 10.52958/iftk.v17i2.3651.

Arfan, A., & Lussiana, E. T. P. (2020). Perbandingan Algoritma Long Short-Term Memory dengan SVR Pada Prediksi Harga Saham di Indonesia. *PETIR*, *13*(1), 33-43. doi: 10.33322/petir.v13i1.858.

Baek, Y., & Kim, H. Y. (2018). ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. *Expert Systems with Applications*, *113*, 457-480. https://doi.org/10.1016/j.eswa.2018.07.019

Examples: APA Style and we suggestion to using Mendeley.

Faurina, R., Winduratna, B., & Nugroho, P. (2018). Predicting stock movement using unidirectional LSTM and feature reduction: The case of an Indonesia stock. In *2018 International Conference on Electrical Engineering and Computer Science (ICEECS). Bali, Indonesia* (pp. 180-5).

Hermuningsih, S. (2012). Pengantar Pasar Modal Indonesia. *Yogyakarta: UPP STIM YKPN*.

Julian, R., & Pribadi, M. R. (2021). Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) Menggunakan Long Short Term Memory (LSTM). *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, *8*(3), 1570-1580. doi: 10.35957/jatisi.v8i3.1159.

Mutasa, S., Sun, S., & Ha, R. (2020). Understanding artificial intelligence based radiology studies: What is overfitting?. *Clinical Imaging*, *65*, 96-99. https://doi.org/10.1016/j.clinimag.2020.04.025

Purnama, R. B., Malyadi, M., & Novawati, N. R. (2017). Perancangan prediksi untuk menentukan indeks harga saham menggunakan jaringan syaraf tiruan. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 125-130. *Doi:Http://Dx.Doi.Org/10.22219/Kinetik.V2i2.207*

Schmidhuber, J. (1997). Discovering neural nets with low Kolmogorov complexity and high generalization capability. *Neural Networks*, *10*(5), 857-873. https://doi.org/10.1016/S0893-6080(96)00127-X

Siahaan, L. E. (2017). Prediksi Indeks Harga Saham dengan Metode Gabungan Support Vector Regression dan Jaringan Syaraf Tiruan. *Indonesia Journal on Computing (Indo-JC)*, *2*(1), 21-30 doi:10.21108/indojc.2017.21.45

Sofi, K., Sunge, A. S., Riady, S. R., & Kamalia, A. Z. (2021). Perbandingan Algoritma Linear Regression, LSTM, dan GRU dalam Memprediksi Harga Saham dengan Model Time Series. *PROSIDING SEMINASTIKA*, *3*(1), 39-46. doi: 10.47002/seminastika.v3i1.275.