DOI: http://dx.doi.org/10.22441/fifo.2024.v16i1.010

Stock Prediction for Indonesia Stock Exchange with Long Short-Term Memory

Abdi Wahab*[1], Ali Herdian[2], Dian Wirawan[3], Yuwan Jumaryadi[4], Syamsir Alam[5], Andrew Fiade[6]

[1,2,3,4]Fakultas Ilmu Komputer, Universitas Mercu Buana, Jakarta, Indonesia
[5]Fakultas Teknik, Universitas Mercu Buana, Jakarta, Indonesia
[6]Fakultas Sains dan Teknologi, UIN Syarif Hidayatullah, Tangerang Selatan, Banten, Indonesia

*[1]abdi.wahab@mercubuana.ac.id, [2]ali.herdian@mercubuana.ac.id, [3]dian.wirawan@mercubuana.ac.id, [4]yuwan.jumaryadi@mercubuana.ac.id, [5]syamsir.alam1@mercubuana.ac.id, [6]andrew_fiade@uinjkt.ac.id

*) Corresponding author

(***received***: 05-06-24, ***revised***: 08-06-24, ***accepted***: 22-06-24)

Abstract

Predicting stock prices through different analyses and techniques is highly challenging. The task is complicated further by fluctuating market conditions and the impact of news, necessitating the consideration of numerous factors. The advancements in machine learning and deep learning have led many researchers to use algorithms like RNN with LSTM for predictions. In this study, we aim to predict stock prices on the Indonesia Stock Exchange using LSTM, focusing on optimizing the hidden layer and activation function. We focus on some stock data with good liquidation in the Indonesia Stock Exchange. The comparison performance between models proposed in this research will be the method in this research. The result showed that the LSTM model with hyperbolic tan activation method performed better than the LSTM model with sigmoid activation method. The future research based on this research, we can compare several other activation methods.

Keyword*: Stock Prediction, LSTM, Hidden Layer, Activation Function*

Abstrak

Memprediksi harga saham dengan berbagai macam analisa dan teknik sangatlah menantang. Pekerjaan menjadi lebih kompleks dengan kondisi pasar yang fluktuatif dan dampak dari berita-berita, sehingga memerlukan pertimbangan banyak faktor. Kemajuan dalam machine learning dan deep learning telah menyebabkan banyak peneliti menggunakan algoritma seperti RNN dengan LSTM untuk prediksi. Pada penelitian ini bertujuan untuk memprediksi harga saham di Bursa Efek Indonesia menggunakan LSTM dengan fokus pada optimasi lapisan tersembunyi dan fungsi aktivasi. Kami fokus pada beberapa data saham dengan likuidasi yang baik di Bursa Efek Indonesia. Perbandingan kinerja antar model yang diusulkan dalam penelitian ini akan menjadi metode dalam penelitian ini. Hasil penelitian menunjukkan bahwa model LSTM dengan metode aktivasi tan hiperbolik memiliki kinerja lebih baik dibandingkan model LSTM dengan metode aktivasi sigmoid. Penelitian selanjutnya berdasarkan penelitian ini, kita dapat membandingkan beberapa metode aktivasi lainnya.

Kata Kunci: Prediksi Saham, LSTM, Hidden Layer, Fungsi Aktifasi

I. Introduction

The capital market industry campaign that has been rife in the past few years has made many people aware of investing in the capital market. What's more, with the pandemic at the beginning of 2020, people started to like saving shares. According to Kustodian Sentral Efek Indonesia (KSEI), 2022, it is recorded that model market investors will reach 10 million by the end of 2022 [1]. It can be seen that many Indonesian people are starting to like investing in the capital market. One of the commodities that has also increased in the capital market is stocks, especially now that it is made easier by the existence of technology to make transactions in stock trading.

After making it easier to make transactions using information technology as well as the internet, the next important step is to determine what stocks to buy so that the investment made can develop the capital owned by investors. Various kinds of stock analysis techniques can be found on the internet and also from books related to stock analysis techniques. Some of these analytical techniques can be used to predict stock price movements, one of the techniques that can be used to predict stock price movements is using artificial intelligence. Several machine learning algorithms for making predictions are widely used as a tool to help predict stock movements.

In this study, we will try to predict stocks using the concept of deep learning, especially by using a Recurrent Neural Network (RNN) type neural network using Long Short-term Memory (LSTM). Several studies to predict stocks using LSTM were conducted by [2], where in their research they focused on improving the model by changing epochs. Then in [3], they made a prediction model with LSTM and the results were compared with other machine learning algorithms. Meanwhile, in [4], they created an LSTM model to predict stocks and improve model performance by increasing epochs. Paper from [5], they made stock price predictions using LSTM and Bi-Directional LSTM, and compared both models based on the epochs and the number of hidden layers. [6] proposed a prediction model with data from the Indian stock market, and it optimizes with a different number of hidden layers. The comparison of activation function in LSTM block for improving LSTM performance by [7], the comparison showed Elliots function had less error levels than another activation function. [8] focus on predicting ANTM.JK stock prices using LSTM during the COVID-19 pandemic. Another researcher in [9] developed a web based stock price prediction system using Flask Framework and LSTM algorithm. Comparing LSTM with another algorithm for predicting stock price done by [10], the result shows LSTM becomes the more accurate model. Another research for predicting stock price with LSTM is in [11], but the prediction is not only for close price, but also for open price, highest price and lowest price. The LSTM model is designed with multiple input and multiple associated output.

In the previous research described above, to improve the performance of the LSTM model that was created, an increase in epochs was carried out when training on the model occurred. An increase in the number of epochs will also affect the training time, the greater the epochs the longer the training time. In an LSTM model, apart from epochs which can be used to improve the performance of the predictive model, it is also possible to increase the hidden layer in the LSTM model being built. Apart from that, you can also use an activation function that is different from the activation function normally used in LSTM models with the hope that the change in the activation function can make the model work better.

Based on the description above, this research will try to see the correlation between the hidden layer and also the activation function in an LSTM model to improve predictions of stock movements. In this study, the LSTM model was built with different hidden layers and also different activation functions according to the scenario to be tested.

II. Research Methodology

Based on the explanation in section 1, in this study we will create an LSTM model with a hidden layer and also a different activation function. After making the models, we will try to compare the performance of these models. Does the change in the hidden layer and the activation function affect the performance of the model in predicting stock.

Long-short Term Memory (LSTM)

One of the weaknesses of the recurrent neural network model is its inability to maintain very long contexts. LSTM is deliberately designed to overcome the problem of long-term dependencies. Computation at the LSTM layer is more complex when compared to the RNN layer. This is because the LSTM tries to retain information received long before the current data is processed.

According to [12] one of the differences between LSTM and RNN is that the LSTM layer produces 2 outputs which we can think of as long-term memory and short-term memory. Therefore, the computational process in the LSTM layer also changes compared to the RNN layer.

Based on Figure 1, there are four gates in LSTM: input gate, forget gate, and output gate. And the activation method used in LSTM is hyperbola tan (tanh). The famous activation methods in LSTM are sigmoid function and hyperbola tan. We choose those activation function in this research to find between both of them which one is more effective for improving the performance LSTM block, especially for LSTM build for this case.

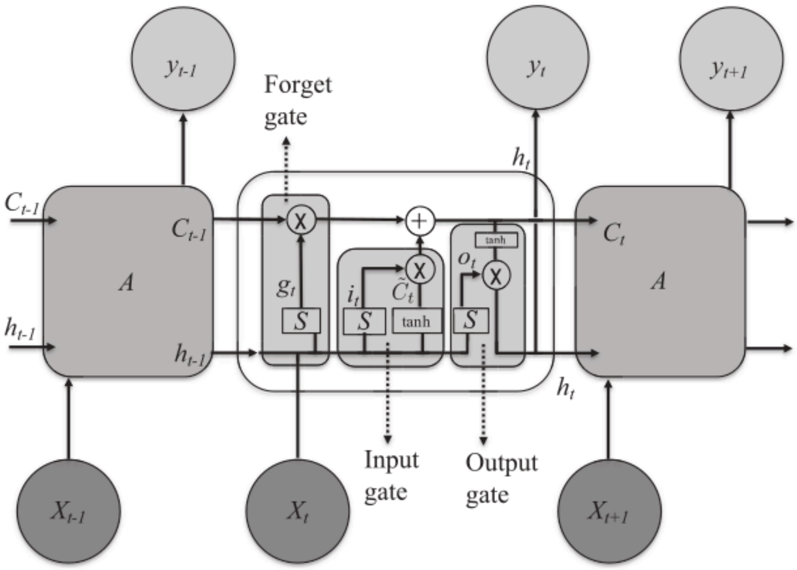

Figure 1. LSTM Architecture [13]

Indonesia Stock Market

Shares according to the Otoritas Jasa Keuangan (OJK) can be interpreted as a sign of the equity participation of a person or party (business entity) in a company or limited liability company. Shares can be purchased through the stock market through a broker registered on the Indonesia Stock Exchange (IDX), and the stock market is open from Monday to Friday from 09:00 to 15:00 GMT+7.

III. Proposed Model

In this research, we proposed some LSTM models for reaching our objectives. The models we proposed are shown in Figure 2.

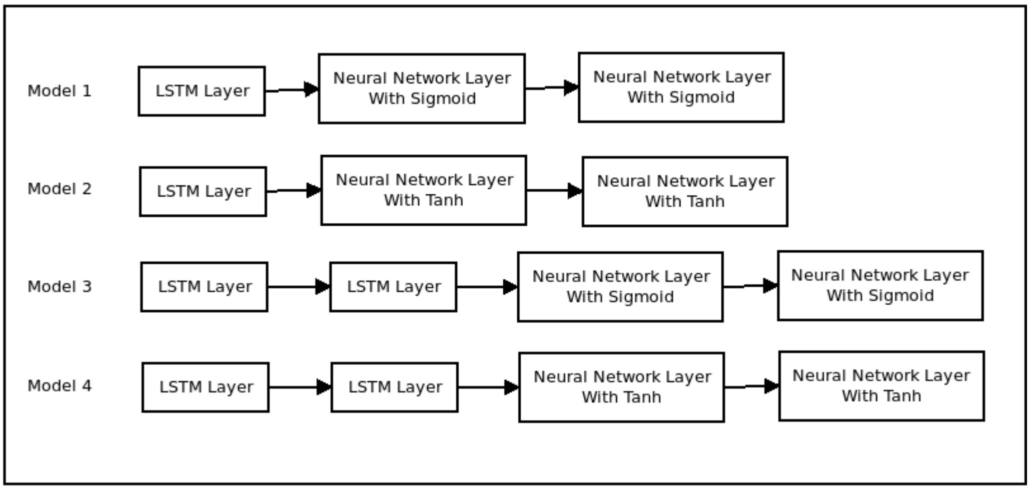

Figure 2. Proposed model

There are four LSTM models developed in this research. The first LSTM model has one layer of LSTM, and the activation function for the last layer is Sigmoid. The second model will be built with One Layer for LSTM and activation function will be used in this model are tanh. The third model will be built with two layers LSTM, and the activation function used in the last layer is sigmoid. The last model is in the same model with the third, but differences in activation function for the last model will be used, the tanh activation function will be used in this model. The models we developed in this study are an LSTM block with single input and single output. Table 1 and Table 2 show the summary of the LSTM models.

Table 1. Summary for LSTM model with one layer of LSTM

Layer (type)	Output Shape	Parameters
LSTM	(None, 60)	14.880
Dense (NN)	(None, 30)	1.830
Dense 1 (NN)	(None, 1)	31

Table 2. Summary for LSTM model with two layers of LSTM

Layer (type)	Output Shape	Parameters
LSTM	(None, 60, 60)	14.880
LSTM 1	(None, 30)	10.920
Dense (NN)	(None, 20)	620
Dense 1 (NN)	(None, 1)	21

Before the training process for the models, in neural networks, included in the LSTM model too, parameters for compiling must be set. The compile parameter for these models show in Table 3

Table 3. Parameters for compiling the models

Parameter	Value
Optimizer	Adam
Learning Rate	0.001
Dropout	0.0
Metric	Accuracy

In this research, based on Figure 2, we used two activation functions: the Sigmoid Function and Hyperbolic Tan. The Sigmoid function also known as Logistic function has an S-shaped curve that smoothly transitions from 0 to 1 as the input increases. Meanwhile the Hyperbolic Tan function also has an S-shaped curve, but it ranges between -1 and 1. The formula for the Sigmoid function and the Hyperbolic Tan function show in (1) and (2) respectively [12].

$$\sigma(z) = \frac{1}{(1 + exp(-z))} \quad (1)$$

$$tanh(z) = 2\sigma(2z) - 1 \quad (2)$$

The σ is the Sigmoid function, z in the input, and the tanh is the Hyperbolic Tan.

IV. Result and Discussion

Selection of Stock Companies

The selection of stocks in this study is focused on companies that are included in stocks that have good fluctuations. There are some stock indexes on the BEI (Bursa Efek Indonesia) website that categorize stock [14]. One category is LQ45, this index categorizes liquid stock for only 45 stocks. Some companies that have good fluctuations are companies in the banking sector and almost all the stocks are included in the LQ45 index. So in this study, we will try to take three examples of companies engaged in banking, namely Bank BRI, Bank BNI, and Bank Mandiri. The company codes for these three companies on the stock market are BBRI.JK, BBNI.JK and BMRI.JK.

Data Retrieval and Data Preprocessing

Retrieval of research data is assisted by a library in python, namely the yfinance library. This library helps in fetching daily stock data from Yahoo Finance. This library has also processed daily stock data in the form of a dataset, making it easier to retrieve stock data which will be used as input for the proposed models.

The stock data used in this study focuses on closing data from each day of stock data obtained. This is done, because closing data is data that describes the last condition of the company every day. As for the opening data, and the highest data, and the lowest data, it does not really describe the condition of the company. The range date for retrieving the stock data used in this research starts from 01 January 2016 until 31 December 2023.

DOI: http://dx.doi.org/10.22441/fifo.2024.v16i1.010

Before we trained the data, the preprocessing steps were done first, we checked for the empty values for the first step. The data downloaded with yfinance library is quite clean, so there is no activity for cleaning that must be done. The next step is scaling the data for getting optimum results when models train the dataset. We used a min max scaler for scaling the data, because it's suitable for data within a specific range [12]. For this research, we ignore the outliers in the dataset, because anomalies of data must be predicted by the proposed model.

Environment Setup

The research environment used in this study uses python 3, and several other libraries used in this study are as follows:

- Numpy
- Pandas
- Tensorflow
- Scikit-learn
- Matplotlib
- Jupyter notebook

We used a laptop without only one CUDA for training this models, because the data we used in this research was only text data, so we analyzed the training step with one CUDA engine.

Result of Model

The test results of the models tested in this study will be discussed in this subchapter. Prepared models will be run using data retrieved from Yahoo Finance as we describe in the data retrieval section. After the data set has been ready we start training and also calculating the performance of models. For training the models, we set the value epochs to 10, and the batch size to 1.

The performance testing for the models will use RMSE as the metrics. It is used because this research area is in regression [15] and for prediction of price so RMSE is suitable for evaluating the predictive models in this research. The formula for RMSE shown in Equation (3).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_o - y_i)^2} \quad (3)$$

The RMSE formula uses n as the number of observations, and y_o is the actual observed value, and y_i is the predicted value from the model. Based on the test results, the results of the test obtained in Table 4 for BBRI, Table 5 for BBNI, and Table 6 for BMRI.

Table 4. Test result for BBRI

Model	RMSE
Model 1	225.002
Model 2	152.401
Model 3	281.261
Model 4	190.552

Table 5. Test result for BBNI

Model	RMSE
Model 1	83.347
Model 2	63.616
Model 3	22.122
Model 4	89.172

Table 6. Test result for BMRI

Model	RMSE
Model 1	447.368
Model 2	292.39
Model 3	459.978
Model 4	245.644

The comparison between this result shown in Figure 3 as a chart line to show the performance between models tested in this research.

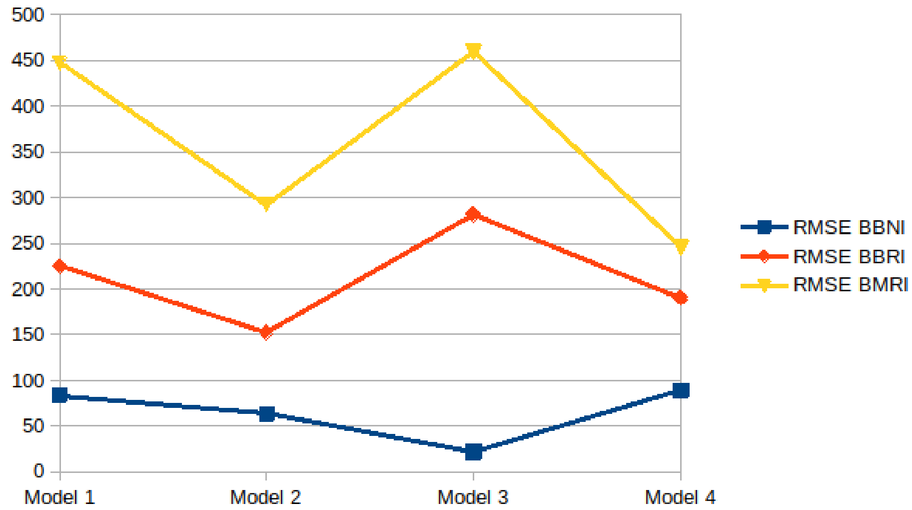

Figure 3. Comparison result of RMSE from models

From the test results obtained, the RMSE for BBRI and BMRI showed that models with 2 layers of LSTM (model 3 and model 4) got bigger RMSE than models with only 1 layer (model 1 and model 2). Different results showed in RMSE from BBNI, the models with 2 layers of LSTM got smaller RMSE values than models with 1 layer of LSTM. From this result, especially for this predicting case, the addition of LSTM layers in models did not impact performance of models. This happens because the input for the models is only 1 feature for this case, so 1 layer of LSTM is enough to handle this case.

Meanwhile the differences of activation function used in these models showed that tanh affecting the performance of models. The models with hyperbolic tan activation function perform better than models with sigmoid activation function. This happens because this case is a regression model, and hyperbolic tan performs better in the regression case.

The result test in [10] showed that LSTM became the best model for predicting stock data. This paper also uses the LSTM model but is a little bit different based on the testing method. In [10] testing the LSTM model with the various period time of stock data, meanwhile this research tested the LSTM model for comparing the activation function and number of hidden layers in LSTM.

The test results or predictions in graphical form can be seen in Figure 4 and Figure 5.

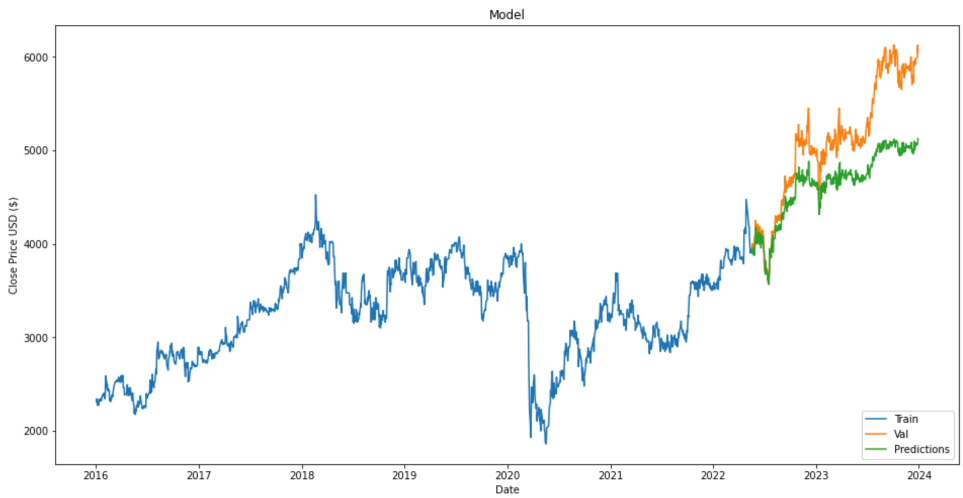

Figure 4. Prediction result from Model 1 with BMRI data

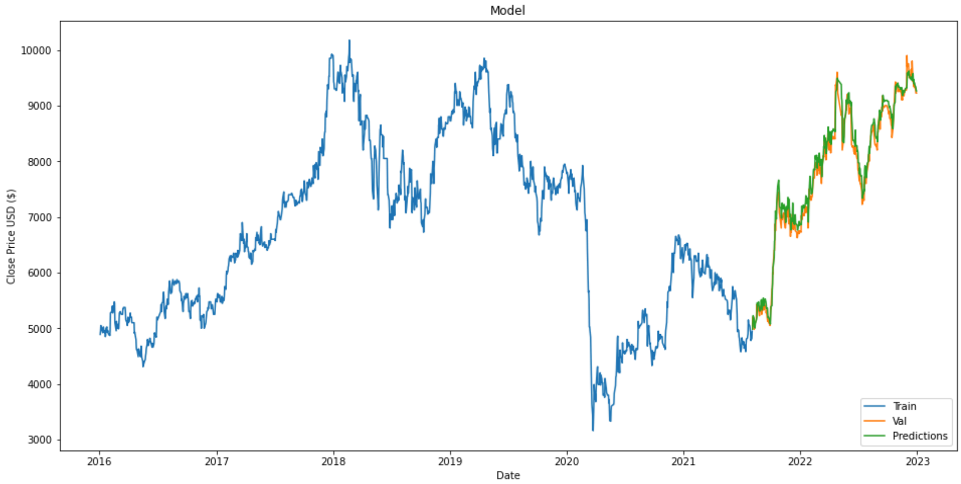

Figure 5. Prediction result from Model 1 with BBNI data

V. Conclusions

Based on the four models that have been built in this study, it shows that adding a new LSTM layer to an existing LSTM architecture does not always increase the value of model accuracy or the performance of the models. We must see the case first for developing the LSTM model before we decide how many layers of LSTM we must provide. Whereas the use of a different activation function in this research showed that hyperbolic tan performs better than sigmoid function, because the prediction is in a regression model . Suggestions in the future can use several more activation functions such as Elliot function, and this research can be developed for many features as input for the models.

References

[1] PT Kustodian Sentral Efek Indonesia, “Statistik Pasar Modal Indonesia Agustus 2023”. www.ksei.co.id. https://www.ksei.co.id/files/Statistik_Publik_Agustus_2023_v2.pdf (accessed 12 November, 2023).
[2] A. Moghar and M. Hamiche, “Stock Market Prediction Using LSTM Recurrent Neural Network,” Procedia Computer Science, vol. 170, pp. 1168–1173, 2020, doi: 10.1016/j.procs.2020.03.049.
[3] D.M.Q. Nelson, et al. "Stock Market’s Price Movement Prediction with LSTM Neural Networks" Conference Paper on International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 2017.
[4] M. Roondiwala, et al. "Predicting Stock Prices Using LSTM" Journal on International Journal of Science and Research (IJSR). Volume 6 Issue 4, April 2017.
[5] M.A.I. Sunny, et al. "Deep Learning-Based Stock Price Prediction Using LSTM and Bi-Directional LSTM Model" Conference Paper on Novel Intelligent and Leading Emerging Sciences Conference (NILES), 2020.
[6] A. Yadav, et al. "Optimizing LSTM for Time Series Prediction in Indian Stock Market" Conference Paper on International Conference on Computational Intellegence and Data Science (ICCIDS), 2019.
[7] A. Farzad, H. Mashayekhi, and H. Hassanpour, “A comparative performance analysis of different activation functions in LSTM networks for classification,” Neural Comput & Applic, vol. 31, no. 7, pp. 2507–2521, Jul. 2019, doi: 10.1007/s00521-017-3210-6.
[8] Khoirudin, P. T. Pungkasanti, and N. Wakhidah, “ANTM.JK Stock Price Prediction Using Long Short-Term Memory (LSTM) Method During COVID-19 Pandemic,” in 2023 International Conference on Technology, Engineering, and Computing Applications (ICTECA), Semarang, Indonesia: IEEE, Dec. 2023, pp. 1–5. doi: 10.1109/ICTECA60133.2023.10490699.
[9] K. Bagastio, R. S. Oetama, and A. Ramadhan, “Development of stock price prediction system using Flask framework and LSTM algorithm,” J. Infras. Policy. Dev., vol. 7, no. 3, Oct. 2023, doi: 10.24294/jipd.v7i3.2631.
[10] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana, and S. S., “Deep Learning for Stock Market Prediction,” Entropy, vol. 22, no. 8, p. 840, Jul. 2020, doi: 10.3390/e22080840.

[11] G. Ding and L. Qin, "Study on the prediction of stock price based on the associated network model of LSTM," Int. J. Mach. Learn. & Cyber., vol. 11, no. 6, pp. 1307–1317, Jun. 2020, doi: 10.1007/s13042-019-01041-1.
[12] A. Geron "Hands-On Machine Learning with Scikit Learn, Keras, and TensorFlow" O'Rielly, 2019.
[13] J. Rogel-Salazar "Advance Data Science and Analytics With Python" CRC Press, 2020.
[14] https://idx.co.id/id/data-pasar/data-saham/indeks-saham
[15] G. Athanasopoulos, and R.J. Hyndman, "Forecasting: Principles and Practice" Otext, 2018.