

# Predicting LQ45 Financial Sector Indices using RNN-LSTM

Seng Hansun (✉ [seng.hansun@lecturer.umn.ac.id](mailto:seng.hansun@lecturer.umn.ac.id))

Universitas Multimedia Nusantara <https://orcid.org/0000-0001-6619-9751>

Julio Christian Young

Universitas Multimedia Nusantara

---

## Research

**Keywords:** Deep Learning, LQ45 financial sector indices, LSTM networks, Prediction, Stock market

**Posted Date:** May 12th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-471941/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Journal of Big Data on July 30th, 2021. See the published version at <https://doi.org/10.1186/s40537-021-00495-x>.

## Predicting LQ45 Financial Sector Indices using RNN-LSTM

Seng Hansun, Julio Christian Young

Informatics Department, Universitas Multimedia Nusantara, Tangerang, Indonesia

Corresponding author: Seng Hansun, [seng.hansun@lecturer.umn.ac.id](mailto:seng.hansun@lecturer.umn.ac.id)

### Abstract

As one of the most popular financial market instruments, the stock has formed one of the most massive and complex financial markets in the world. It could handle millions of transactions within a short period and highly unpredictable. In this study, we aim to implement a famous Deep Learning method, namely the Long Short-Term Memory (LSTM) networks, for the stock price prediction. We limit the stocks to those that are included in the LQ45 financial sectors indices, i.e., BBKA, BBNI, BBRI, BBTN, BMRI, and BTPS. We propose to use a simple three layers LSTM network architecture in predicting the stocks' closing prices and found that the prediction results fall in the reasonable forecasting category. It is worthy to note that two of the considered stocks, i.e., BBKA and BMRI, have the lowest MAPE values at 19.1020 and 18.6135 which fall in the good forecasting results. Hence, the proposed LSTM model is recommended to be used on those two stocks.

### Keywords

Deep Learning, LQ45 financial sector indices, LSTM networks, Prediction, Stock market

### Introduction

Stock, as one form of securities, is a paper sheet as proof of someone's ownership in a company (1). Since it can be exchanged easily, it is popularly used as a kind of investment by many people nowadays. In fact, the stock has become one of the most popular financial market instruments (2) and has formed one of the most massive and complex financial markets in the world – the stock market. The market could handle millions of transactions in a short period and is an extremely dynamic environment (3). It also has nonlinear characteristics and could be highly unpredictable (4,5).

Many have tried to study and predict the stock's price movement which is considered as the crucial step in developing trading strategies (6). A good trading strategy could help traders in avoiding losses and gaining more profit. Different kinds of forecasting methods and approaches have been applied and developed in order to achieve that. Some have used conventional and technical analysis methods, such as Moving Average, ARIMA, and Stochastic Optimization method, but the accuracy results are not satisfactory enough (5). Hence, others move to a more emerging technic and approach, such as Machine Learning or Deep Learning methods, for the stock price prediction problem.

One of the most widely used Deep Learning methods, especially for time series analysis, is the Long Short-Term Memory (LSTM) networks. As a better version of Recurrent Neural Networks (RNN), LSTM has been used in many fields. Particularly in stock price prediction, LSTM had been proposed and applied by some notable researchers, such as Murtaza et al. (7), Nelson et al. (8), Faurina et al. (6), and Jin et al. (9). Murtaza et al. (7) had built a model and predict the stock returns of NIFTY 50 by using LSTM networks. Similarly, Nelson et al. (8) also had applied LSTM networks to predict the future trends

of stock prices based on its historical data. They applied the proposed LSTM networks on five different stocks from the Brazilian Stock Exchange and found that the outcomes are very promising. Faurina et al. (6) had tried a slightly different approach in predicting stock movement on the Indonesia market. They used three different dimensional reduction methods, namely Lasso, ElasticNet, and Principal Component Analysis (PCA), that were combined with the LSTM networks for the stock price prediction. They found that the PCA-LSTM combination outperforms the other two combinations. Lastly, Jin et al. (9) had recently published their findings in the stock closing price prediction. In their study, they incorporated the investors' emotional tendency with LSTM networks and found that the proposed model could improve the prediction results.

In this study, we aim to implement the famous Deep Learning methods, i.e., the LSTM networks for the stock price prediction, specifically for stocks that are included in the LQ45 indices. LQ45 consists of 45 companies' stocks that have high liquidity and huge market share; and therefore, considered to have a good financial status (10). The Indonesia Stock Exchange (IDX) has periodically assessed all companies' performance in the stock market and announced the report (LQ45) semi-annually (11). This report has been used by many stakeholders, including traders and investors, as one of their decision-making tools. Moreover, many studies also had used this report as their criteria in choosing the considered stocks for their evaluation, as can be found on Nurmalitasari et al. (12), Pantagama (11), and Pramanaswari and Yasa (13).

We will further discuss the research method used in this study, namely the LSTM networks in the following section. Some error criteria also will be described in the section. Moreover, the description of datasets being used in this study, the experimental results, and analysis will be discussed. Lastly, some concluding remarks and suggestions will be given in the last section of this paper.

## **Research Method**

A simple diagram, which shows the research methodology applied in this study is shown in Figure 1. We started the process by handling missing values in the stocks 'Closing' price with the data imputation method. Then, we divided the cleaned stocks' data into training and test set with an 80:20 ratio. On both sets, we conducted the feature scaling to get normalized training and test sets. Next, we reshaped both sets into 3D arrays that can be furthered process on the prepared LSTM networks. We built the Deep Learning model using three layers LSTM networks that consist of an LSTM layer, a Dropout layer to prevent overfitting, and a dense layer as the output layer of the networks. To get the prediction results, we used the built model on the reshaped test set and converted it back to the original scale. Lastly, we calculated the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) scores in the process of performance evaluation of the built model. A more detailed explanation of each step will be given in the following sections.

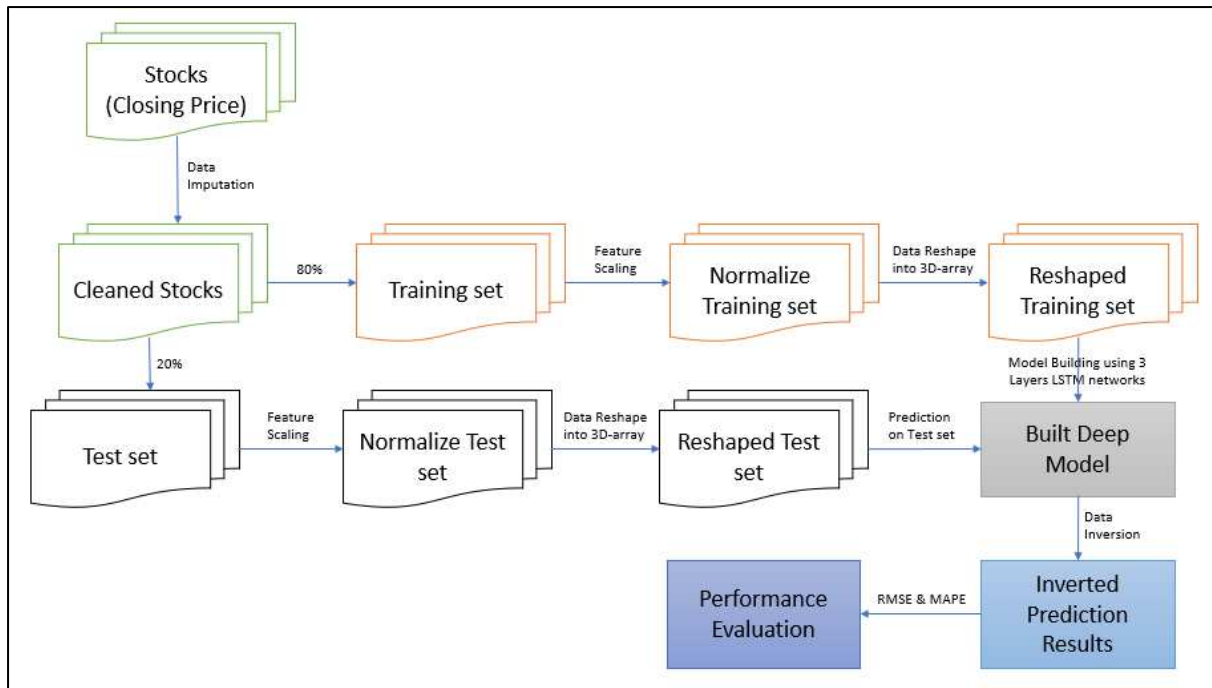


Figure 1. Research methodology

## RNN-LSTM

LSTM as an advanced soft computing method was first introduced by Hochreiter and Schmidhuber to tackle the limitation found in the conventional Recurrent Neural Networks (RNN) method (14). Typically, it consists of several LSTM cells that are self-connected and used to store the networks' temporal state by using three gates, namely the input, output, and forget gates (14). An illustration of an LSTM cell containing those three gates is depicted in Figure 2 (15).

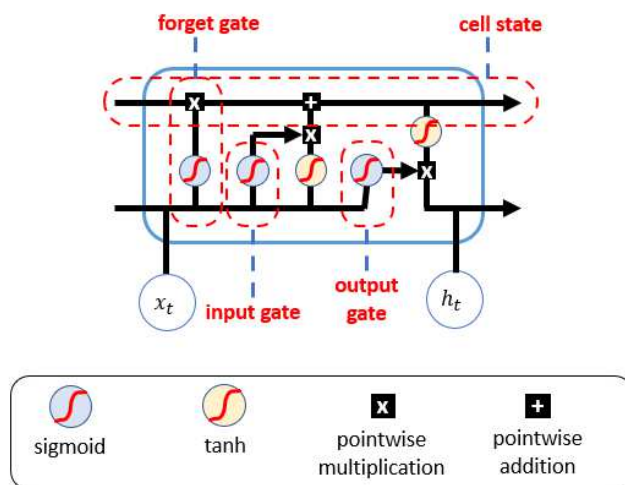


Figure 2. Three gates mechanism in an LSTM cell

The gate mechanism in an LSTM cell is used to control how much information can be passed throughout the networks. The forget gate can be found in the first part of the cell and is used to control how much the previous cell's hidden state could be forgotten. Next, the input gate is used to determine what new information will be stored in the current cell state, and finally, the output gate,

as its name inferred, is used to find the value we want to be the output of the current cell (16). Some equations related to this mechanism in an LSTM cell are given below.

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \quad (1)$$

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C h_{t-1} + U_C x_t + b_C) \quad (3)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \quad (5)$$

$$h_t = o_t \odot \tanh(C_t) \quad (6)$$

where  $f_t$  is the forget gate value at the current cell,  $i_t$  is the input gate value,  $C_t$  is the current cell state,  $\tilde{C}_t$  is the cell candidate value,  $o_t$  is the output gate value,  $W_f, W_i, W_C, W_o$  and  $U_f, U_i, U_C, U_o$  are the weights of the networks,  $b_f, b_i, b_C, b_o$  is the bias variable value,  $h_t$  is the current hidden state value,  $h_{t-1}$  is the prior hidden state value, and  $x_t$  is the new input value at the current cell.

### Error Criteria

There are two prediction error criteria being used as the performance evaluation metrics in this study, i.e., the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE) criteria. While RMSE will give the degree of error in a unit value, MAPE will give the degree of error in a percentage value. As explained by Shahid et al. (14) and Hansun et al. (17), both of them can be represented as

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2} \quad (7)$$

$$MAPE = \left( \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{Y_t} \right| \right) \cdot 100\% \quad (8)$$

where  $n$  is the total number of data,  $Y_t$  is the actual value, and  $F_t$  is the forecasted value.

### Result and Analysis

This section is divided into two sub-sections. Firstly, we describe the data source, pre-processing steps, and model development conducted in this study. Secondly, the experimental results for each considered stock in LQ45 indices will be given. Moreover, the evaluation results and analysis will be discussed following the experimental results.

#### Data Source, Pre-Processing, and Model Development

In this study, we used some stocks that are included in the LQ45 indices, as semi-annually reported by the Indonesia Stock Exchange (IDX). The last (Major Evaluation) Report is for the period February to July 2021 and can be downloaded from the IDX website (18). For simplicity, we will focus on some stocks on the list that are included in the financial sector, as shown in Table 1.

Table 1. LQ45 financial sector stocks

Code	Stock Name
BBCA	Bank Central Asia Tbk.
BBNI	Bank Negara Indonesia (Persero) Tbk.
BBRI	Bank Rakyat Indonesia (Persero) Tbk.

BBTN	Bank Tabungan Negara (Persero) Tbk.
BMRI	Bank Mandiri (Persero) Tbk.
BTPS	Bank BTPN Syariah Tbk.

Next, we collected the recorded daily stock prices for each considered stock from Yahoo! Finance (19). We chose to download the maximum data available in the data source, hence each stock could have a different number of data records. There are several features in the collected datasets, such as 'Date', 'Open', 'High', 'Low', 'Close', 'Adjusted Close', and 'Volume', but we will only consider the 'Close' value of each set. The downloaded data then were processed to handle any missing values using a simple data imputation method, by replacing the missing values with their last known records. Moreover, in the data splitting process, we used an 80:20 ratio for training and test sets of each considered stock. The resulted training and test sets are shown in Table 2.

Table 2. Training and test sets of LQ45 financial sector stocks

Code	All Data	Training Set	Test Set	# of Missing Values
BBCA	4,189	3,351	838	1 <sup>a</sup>
BBNI	4,309	3,447	862	4 <sup>b</sup>
BBRI	4,340	3,471	869	1 <sup>a</sup>
BBTN	2,811	2,248	563	1 <sup>a</sup>
BMRI	4,425	3,539	886	1 <sup>a</sup>
BTPS	741	592	149	2 <sup>c</sup>

<sup>a</sup> Missing value found on 2019-06-19

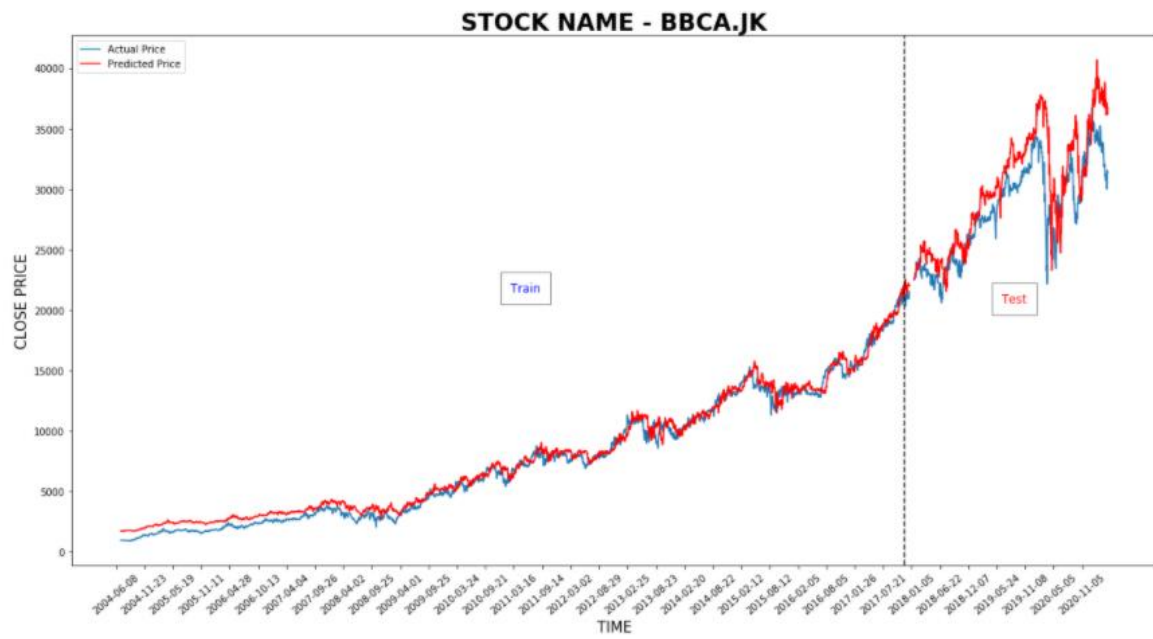
<sup>b</sup> Missing values found on 2006-09-29, 2006-10-02, 2006-10-03, and 2019-06-19

<sup>c</sup> Missing values found on 2018-05-10 and 2019-06-19

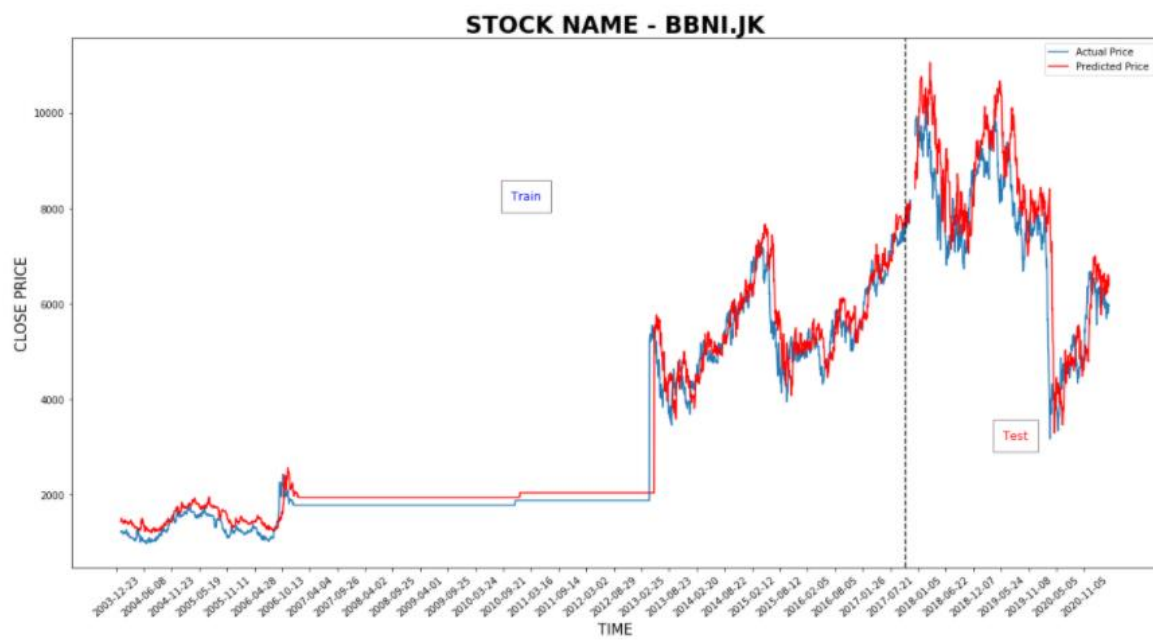
On both the training and test sets, we further processed them for data normalization using the feature scaling method. Then, we reshaped the data into a 3D-array shape that can be accepted by the LSTM model in Keras, a Deep Learning package for Python. Keras runs on top of the TensorFlow Machine Learning platform. In this study, we proposed a three layers LSTM network that contains an LSTM layer, a Dropout layer, and a Dense layer. There are 100 neurons being used in the LSTM layer, meanwhile, for the Dropout regularisation, we chose to drop 20% of the processed information in the networks to prevent overfitting. For the loss function in the networks, we used the simple Mean Square Error (MSE) with Adam optimizer. Moreover, we trained the model on the training set for 20 training epochs with a batch size of 32 each. The built model then will be used to predict the Closing price on the test set. However, we need to do the data inversion phase to convert the predicted results into the original scaling of the data. The prediction results based on the built model using the proposed LSTM networks are described in the following section.

## Experimental Results

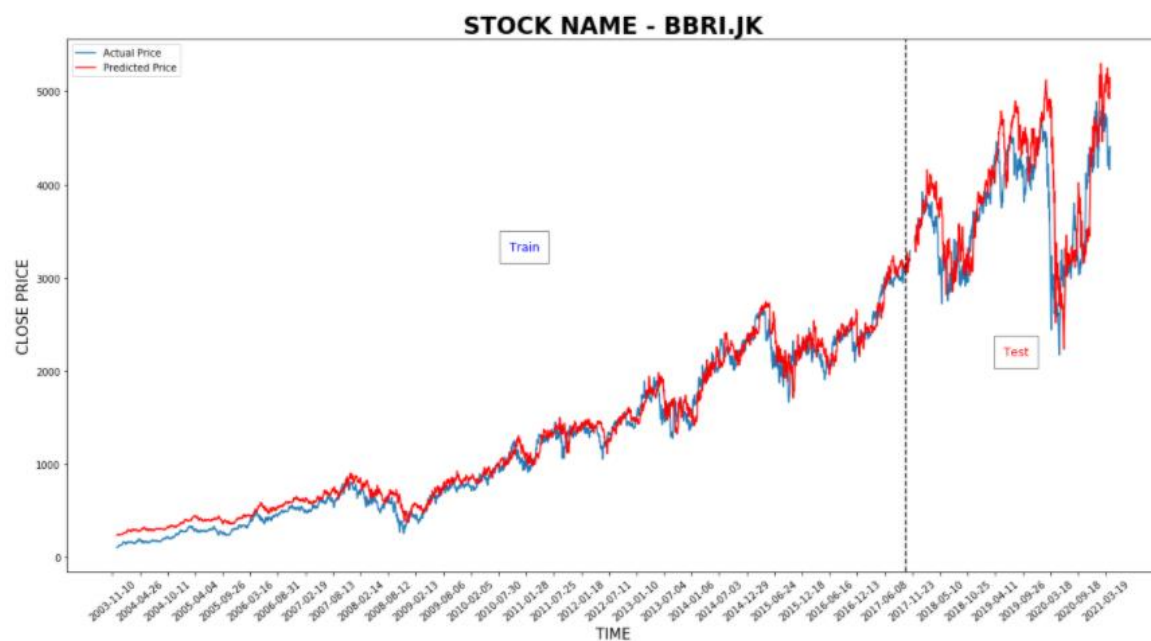
The prediction results on all considered stocks included in the list of LQ45 financial sector indices are shown in Figure 3. Furthermore, the loss function results obtained during the model training are shown in Figure 4. As can be inferred from Figure 3, the proposed LSTM model could predict all considered stocks data very well, and from Figure 4, we know that all built models have converged quite well and remained stable.



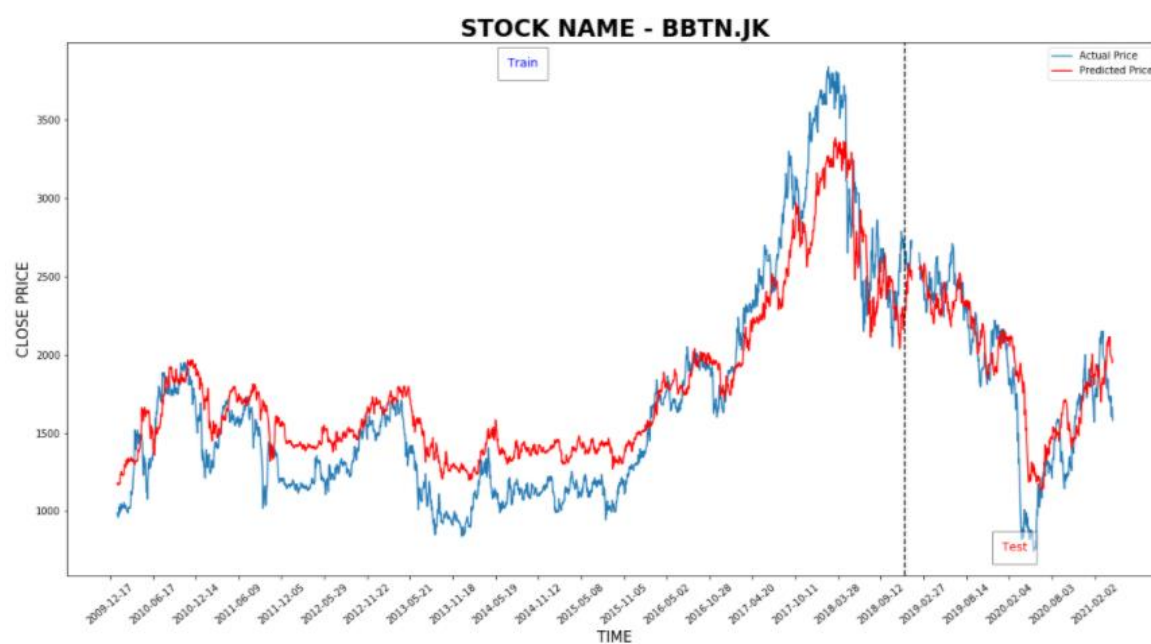
(a) BBKA



(b) BBNI

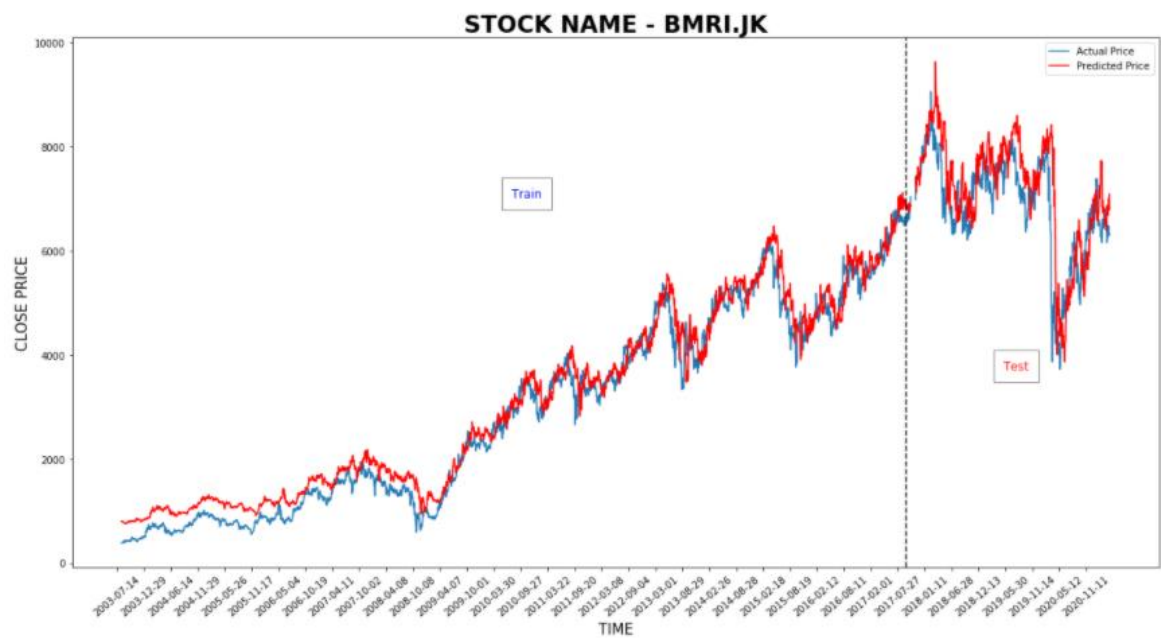


(c) BBRI

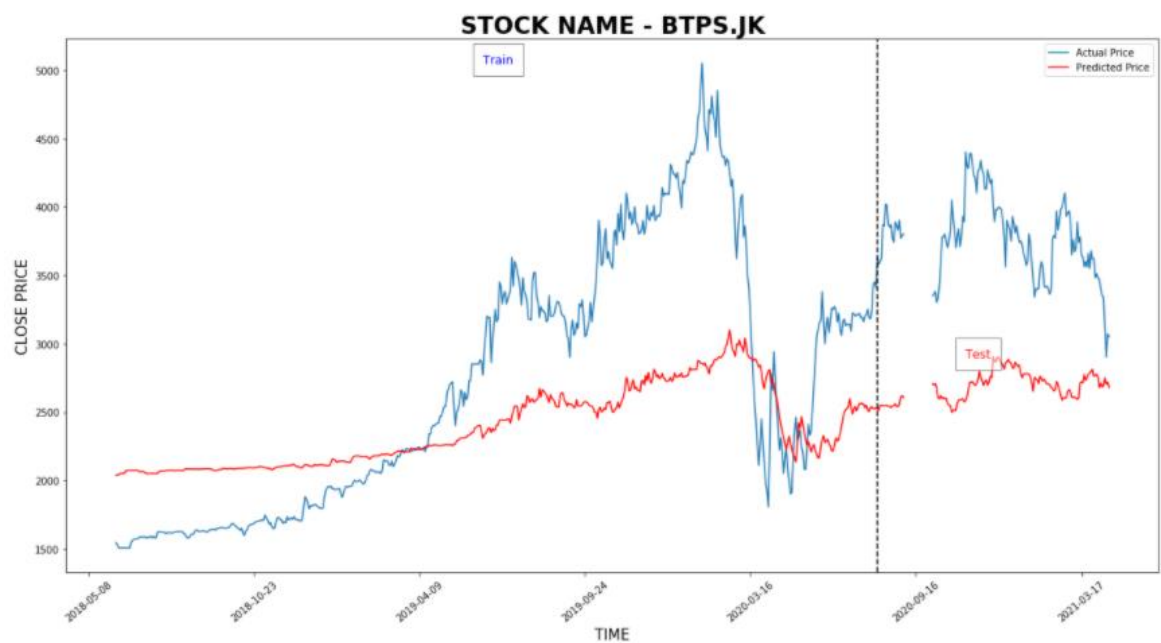


(d) BBTN



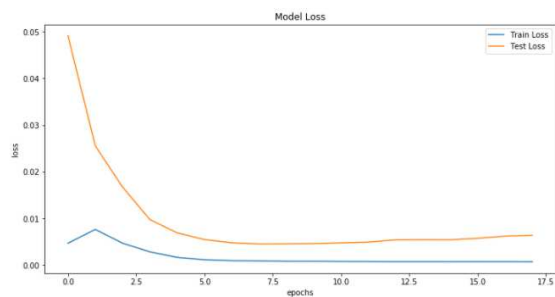


(e) BMRI

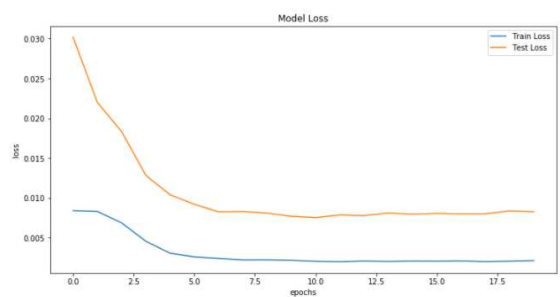


(f) BTPS

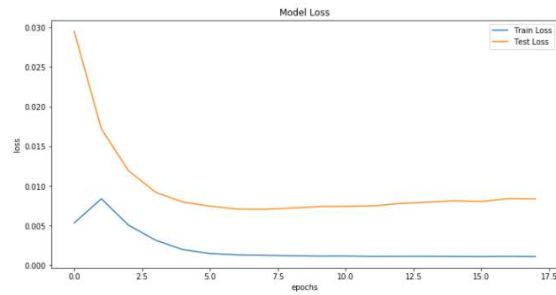
Figure 3. LQ45 financial sector indices' prediction results



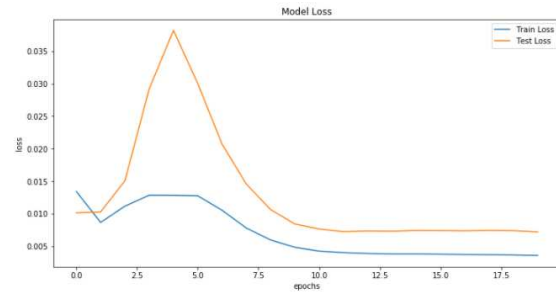
(a) BBCA



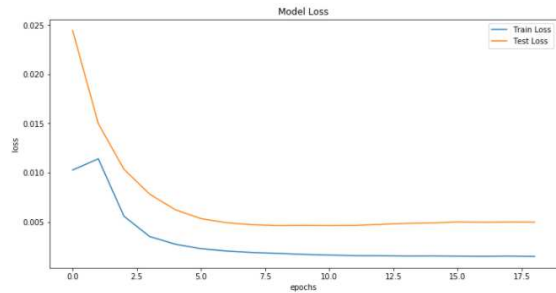
(b) BBNi



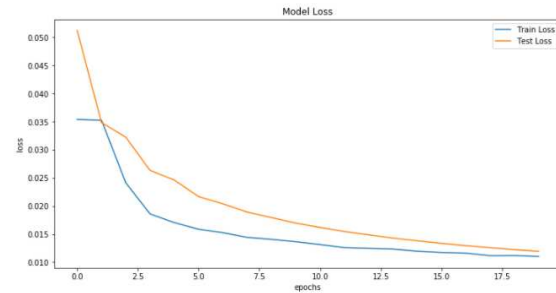
(c) BBRI



(d) BBTN



(e) BMRI



(f) BTPS

Figure 4. LQ45 financial sector indices' loss function plots

### Analysis

To get the performance results of the built model on the test set, we evaluated the prediction results by using the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE) criteria. Table 3 presents the RMSE and MAPE values for each considered stock in this study. As evinced from the results, the RMSE values are ranged from 266.1255 to 2878.4668, which depend not only on the prediction results but also on the magnitude of the stock's closing prices. Moreover, there are two stocks that have MAPE values under 20%, namely BMRI (18.6135) and BBKA (19.1020). The typical interpretation of MAPE values as stated by Moreno et al. (20) is highly accurate forecasting (<10), good forecasting (10-20), reasonable forecasting (20-50), and inaccurate forecasting (>50). Therefore, the prediction results on BMRI and BBKA stock prices fall in good forecasting, while all other prediction results have reasonable results.

Table 3. RMSE and MAPE values

Code	RMSE	MAPE
BBKA	2878.4668	19.1020
BBNI	994.6086	33.4879
BBRI	441.1157	20.6837
BBTN	266.1255	34.8298
BMRI	739.1828	18.6135
BTPS	1103.8376	27.5634

### Conclusion

In this study, we have successfully implemented the well-known Deep Learning methods, i.e., the LSTM networks for the stock price prediction, specifically for stocks that are included in the LQ45 financial sector indices. There are six stocks considered in this study, namely BBKA (Bank Central Asia Tbk.), BBNI (Bank Negara Indonesia Tbk.), BBRI (Bank Rakyat Indonesia Tbk.), BBTN (Bank Tabungan

Negara Tbk.), BMRI (Bank Mandiri Tbk.), and BTPS (Bank BTPN Syariah Tbk.). By using a simple three layers LSTM network architecture, we found that the proposed model gave reasonable prediction results for all considered stocks. For BBKA and BMRI, the model could get good prediction results, and therefore, it is more recommended to use the built model for those two stocks in the LQ45 financial sector indices.

There are also some limitations in this study that are worthy to be noted. As previously explained in the text, we only did simple data imputation technique to handle any missing values in the dataset. However, as can be seen in the prediction plot for BBNI in Figure 3, almost half of the stocks' closing prices in the training set are stagnant at the value of 1,800s. It seems that there are some periods when the stock's price data for BBNI in Yahoo! Finance are not recorded properly. This could affect the prediction results as shown by the high MAPE value, and therefore, some other pre-processing techniques need to be considered if the stock will be used in the future. Moreover, in this study, we only used two well-known forecast error criteria as the performance evaluation metrics, namely the RMSE and MAPE. Other criteria, including the R2 score or coefficient of determination, could also be applied to get better and comprehensive analysis results (21).

### **List of abbreviations**

BBKA: Bank Central Asia Tbk.

BBNI: Bank Negara Indonesia Tbk.

BBRI: Bank Rakyat Indonesia Tbk.

BBTN: Bank Tabungan Negara Tbk.

BMRI: Bank Mandiri Tbk.

BTPS: Bank BTPN Syariah Tbk.

IDX: Indonesia Stock Exchange

LSTM: Long Short-Term Memory

MAPE: Mean Absolute Percentage Error

MSE: Mean Square Error

RMSE: Root Mean Square Error

RNN: Recurrent Neural Networks

### **Declarations**

#### **Ethics approval and consent to participate**

Not applicable.

#### **Consent for publication**

Not applicable.

#### **Availability of data and materials**

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### **Competing interests**

The authors declare that they have no competing interests.

### **Funding**

No applicable.

### **Authors' contributions**

SH has initiated, conducted, and made the first draft of this paper. All authors have reviewed and contributed to the final version of the paper.

### **Acknowledgements**

The authors would like to thank the anonymous reviewers for their valuable comments and corrections. The authors would also like to extend our gratitude to Universitas Multimedia Nusantara for the support given during this study.

### **Authors' information**

Informatics Department, Universitas Multimedia Nusantara, Tangerang, Indonesia

Seng Hansun, Julio Christian Young

### **References**

1. Johan K, Young JC, Hansun S. LSTM-RNN automotive stock price prediction. *Int J Sci Technol Res.* 2019;8(9).
2. Meizir, Rikumahu B. Prediction of Agriculture and Mining Stock Value Listed in Kompas100 Index Using Artificial Neural Network Backpropagation. In: 2019 7th International Conference on Information and Communication Technology (ICoICT) [Internet]. Kuala Lumpur, Malaysia: IEEE; 2019. p. 1–5. Available from: <https://ieeexplore.ieee.org/document/8835284/>
3. Mootha S, Sridhar S, Seetharaman R, Chitrakala S. Stock Price Prediction using Bi-Directional LSTM based Sequence to Sequence Modeling and Multitask Learning. In: 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) [Internet]. New York, NY, USA: IEEE; 2020. p. 0078–86. Available from: <https://ieeexplore.ieee.org/document/9298066/>
4. Sridhar S, Mootha S, Subramanian S. Detection of Market Manipulation using Ensemble Neural Networks. In: 2020 International Conference on Intelligent Systems and Computer Vision (ISCV) [Internet]. Fez, Morocco: IEEE; 2020. p. 1–8. Available from: <https://ieeexplore.ieee.org/document/9204330/>
5. Liu S, Liao G, Ding Y. Stock Transaction Prediction Modeling and Analysis based on LSTM. In: 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA) [Internet]. Wuhan, China: IEEE; 2018. p. 2787–90. Available from: <https://ieeexplore.ieee.org/document/8398183/>
6. Faurina R, Winduratna B, Nugroho P. 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; 2018. p. 180–5.

7. Murtaza R, Patel H, Varma S. Predicting Stock Prices Using LSTM. *Int J Sci Res* [Internet]. 2017;6(4):1754–6. Available from: <https://www.ijsr.net/archive/v6i4/ART20172755.pdf>
8. Nelson DMQ, Pereira ACM, de Oliveira RA. Stock Market's Price Movement Prediction with LSTM Neural Networks. In: 2017 International Joint Conference on Neural Networks (IJCNN) [Internet]. Anchorage, AK, USA: IEEE; 2017. p. 1419–26. Available from: <http://ieeexplore.ieee.org/document/7966019/>
9. Jin Z, Yang Y, Liu Y. Stock Closing Price Prediction based on Sentiment Analysis and LSTM. *Neural Comput Appl* [Internet]. 2020 Jul 19;32(13):9713–29. Available from: <http://link.springer.com/10.1007/s00521-019-04504-2>
10. Tanuwijaya J, Hansun S. LQ45 Stock Index Prediction using k-Nearest Neighbors Regression. *Int J Recent Technol Eng* [Internet]. 2019 Sep 30;8(3):2388–91. Available from: <https://www.ijrte.org/wp-content/uploads/papers/v8i3/C4663098319.pdf>
11. Pantagama M, Rikumahu B. Indonesia Financial Sector Stock Prediction using Long Short-Term Memory Network Algorithm and Modeling (Study of Banking in August 2018 LQ45 Index). In: Anggadwita G, Martini E, editors. *Digital Economy for Customer Benefit and Business Fairness* [Internet]. Routledge; 2020. p. 159–64. Available from: <https://www.taylorfrancis.com/books/9781000070644>
12. Nurmalitasari, Sumarlinda S, Supriyanto N, Putri DK. LQ45 Stock Price Predictions Using The Deep Learning Method. *Int J Adv Res Publ* [Internet]. 2020;4(4):20–3. Available from: <http://www.ijarp.org/published-research-papers/apr2020/Lq45-Stock-Price-Predictions-Using-The-Deep-Learning-Method.pdf>
13. Pramanaswari ASI, Yasa GW. Graham & Dodd Theory in Stock Portfolio Performance in LQ 45 Index at Indonesia Stock Exchange. *Int Res J Manag IT Soc Sci* [Internet]. 2018;5(6):52–9. Available from: <http://sloap.org/journals/index.php/irjmis/article/view/338>
14. Shahid F, Zameer A, Muneeb M. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos, Solitons & Fractals* [Internet]. 2020 Nov;140:110212. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0960077920306081>
15. Phi M. Illustrated Guide to LSTM's and GRU's: A step by step explanation [Internet]. towards data science. 2018 [cited 2021 Apr 1]. Available from: <https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21>
16. Wang P, Zheng X, Ai G, Liu D, Zhu B. Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran. *Chaos, Solitons & Fractals* [Internet]. 2020 Nov;140:110214. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S096007792030610X>
17. Hansun S, Charles V, Gherman T, Subanar S, Indrati CR. A Tuned Holt-Winters White-Box Model for COVID-19 Prediction. *Int J Manag Decis Mak* [Internet]. 2021;20(1):1. Available from: <http://www.inderscience.com/link.php?id=10034422>
18. IDX. Indeks Saham [Internet]. Jakarta, Indonesia; 2021. Available from: <https://www.idx.co.id/data-pasar/data-saham/indeks-saham/>
19. Yahoo! Finance. Quotes [Internet]. 2021 [cited 2021 Apr 19]. Available from: <https://finance.yahoo.com/lookup>
20. Moreno JJM, Pol AP, Abad AS, Blasco BC. Using the R-MAPE Index as a Resistant Measure of Forecast Accuracy. *Psicothema* [Internet]. 2013;25(4):500–6. Available from: <http://www.psicothema.com/pdf/4144.pdf>

21. Budiharto W. Data Science Approach to Stock Prices Forecasting in Indonesia during Covid-19 using Long Short-Term Memory (LSTM). J Big Data [Internet]. 2021 Dec 11;8(1):47. Available from: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00430-0>

# Figures

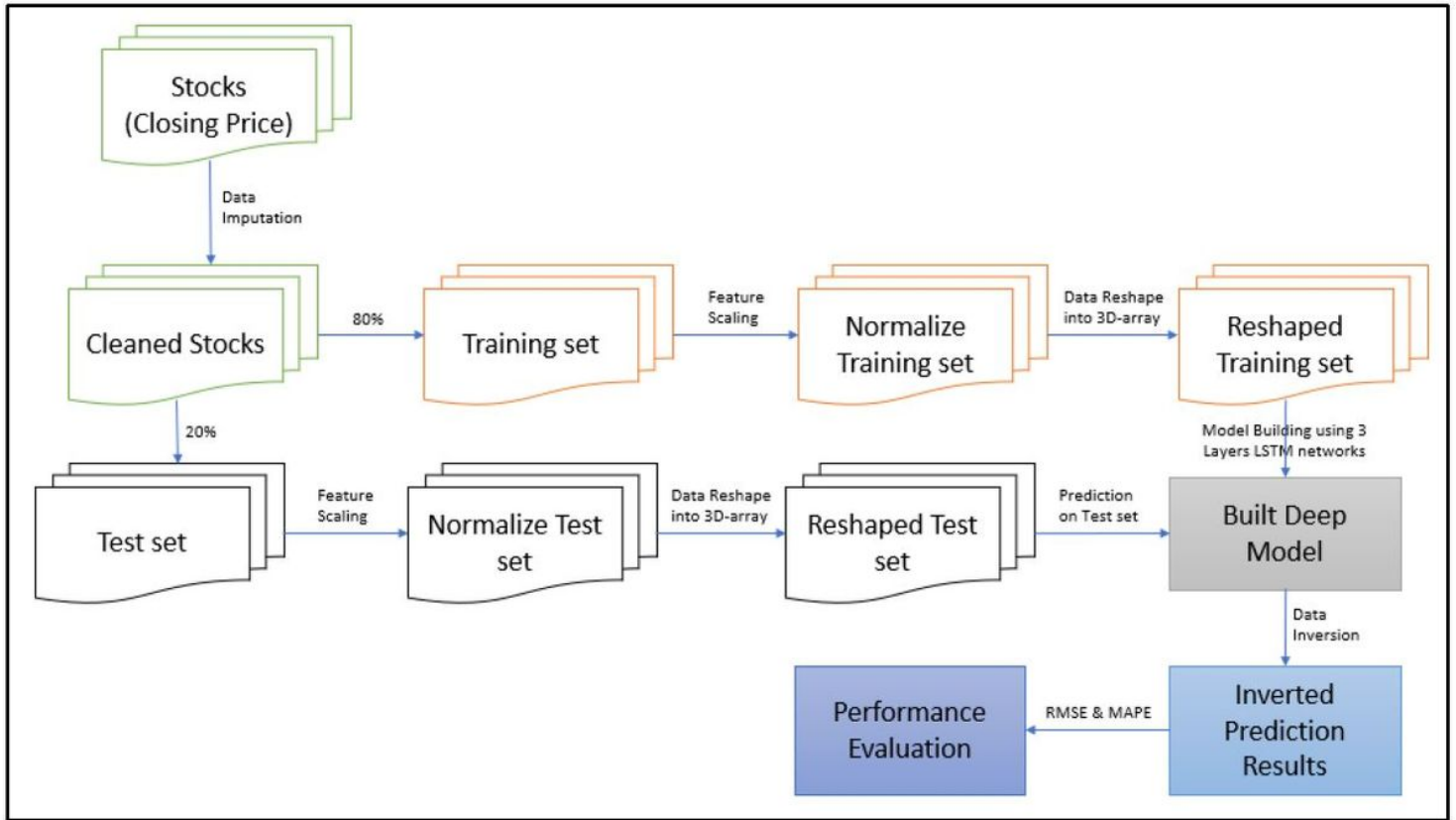


Figure 1

Research methodology

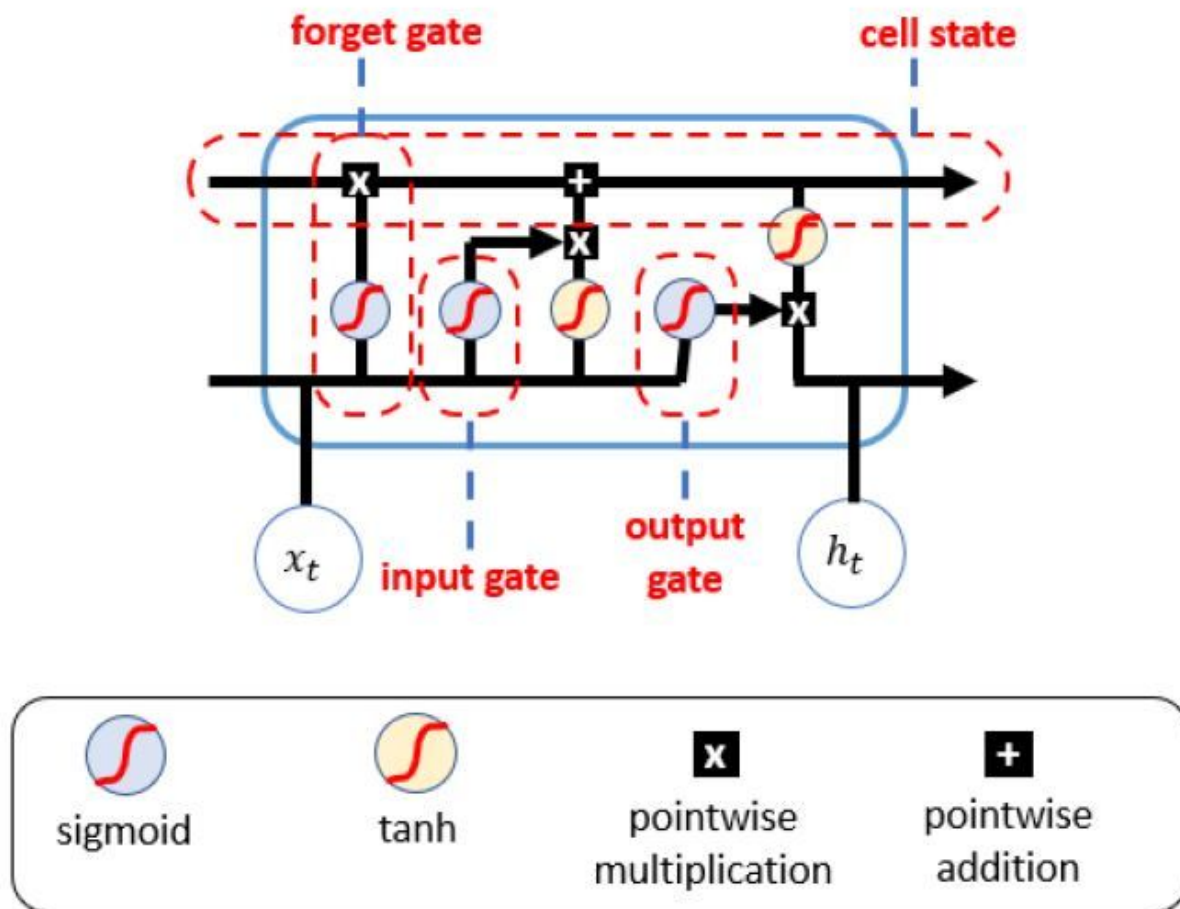


Figure 2

Three gates mechanism in an LSTM cell

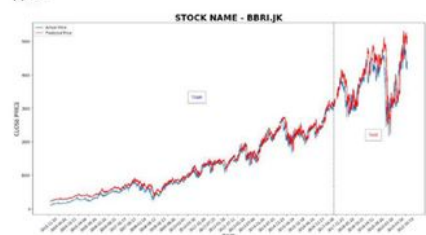




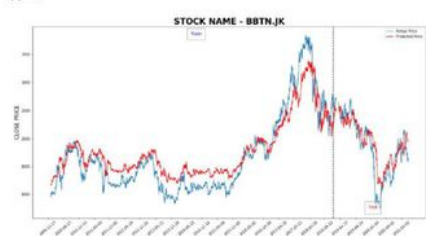
(a) BBKA



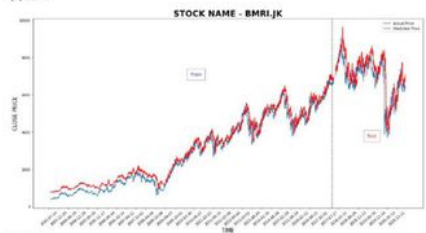
(b) BBNI



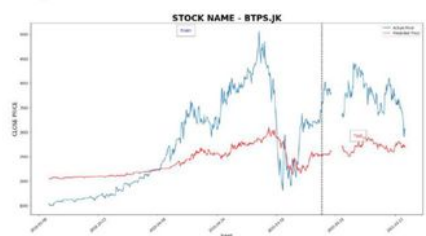
(c) BBRI



(d) BBTN



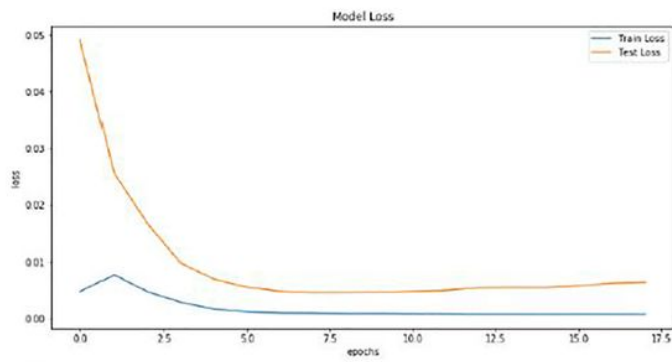
(e) BMRI



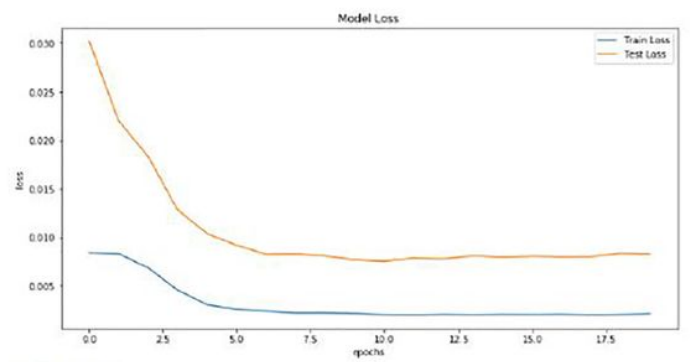
(f) BTPS

**Figure 3**

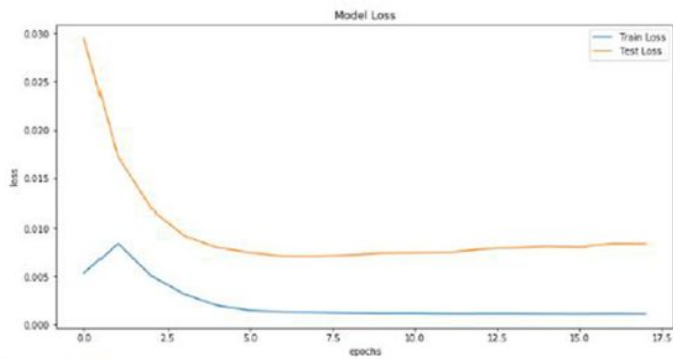
LQ45 financial sector indices' prediction results



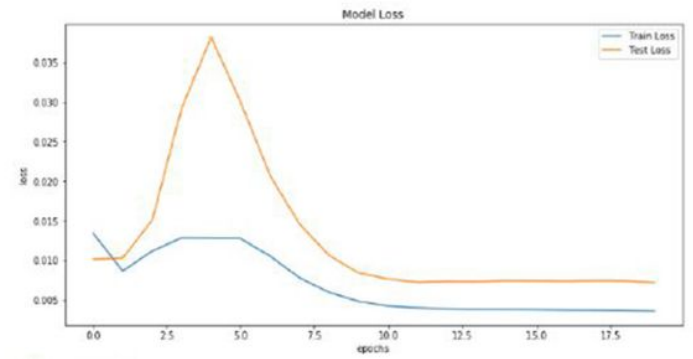
(a) BBCA



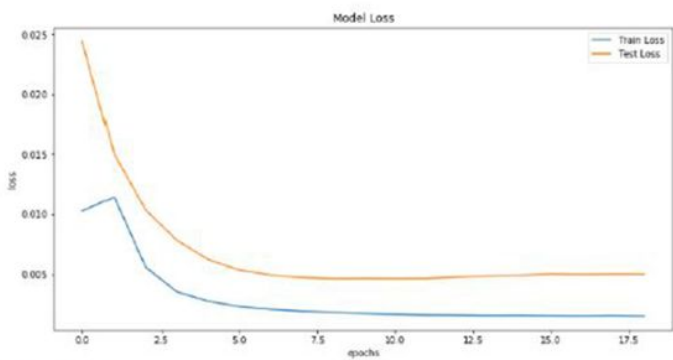
(b) BBNI



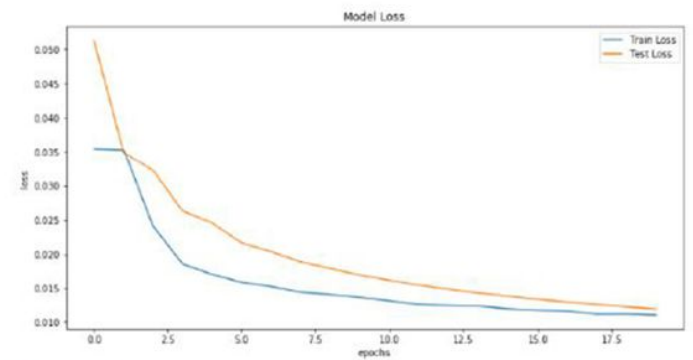
(c) BBRI



(d) BBTN



(e) BMRI



(f) BTPS

**Figure 4**

LQ45 financial sector indices' loss function plots