On searching the best mode for forex forecasting: bidirectional long short-term memory default mode is not enough

Seng Hansun[1], Farica Perdana Putri[1], Abdul Q. M. Khaliq[2], Hugeng Hugeng[3]
[1]Informatics Department, Engineering and Informatics Faculty, Universitas Multimedia Nusantara, Tangerang, Indonesia
[2]Department of Mathematical Sciences and Center for Computational Science, Middle Tennessee State University, Murfreesboro, USA
[3]Electrical Engineering Department, Engineering Faculty, Universitas Tarumanagara, Jakarta, Indonesia

Article Info

Article history:

Received Jul 15, 2021
Revised Jun 22, 2022
Accepted Jul 21, 2022

Keywords:

Bidirectional long short-term memory
Currency pair
Deep learning
Forex forecasting
Long short-term memory

ABSTRACT

Presently, the Forex market has become the world's largest financial market with more than US$5 trillion daily volume. Therefore, it attracts many researchers to learn its traded currency pairs characteristics and predict their future values. Here, we propose simple three layers Bidirectional long short-term memory (Bi-LSTM) networks for Forex forecasting with four different merge modes. Moreover, the proposed model is also compared to the conventional long short-term memory (LSTM) networks with the same architecture. Five major Forex currency pairs, namely AUD/USD, EUR/USD, GBP/USD, USD/CHF, and USD/JPY, with more than ten years of historical records are considered in this study. It is revealed from the experimental results that among four available merge modes, the concatenation mode as the default merge mode in Bi-LSTM networks is actually the less preferred mode for Forex forecasting (Root mean square error 0.30685517, mean absolute error 0.27442235, mean absolute percentage error 0.827108%). Moreover, Bi-LSTM average mode gets the highest R^2 score that could achieve 89.579%. Therefore, the proposed three layers Bi-LSTM networks could provide a baseline result for developing a good trading strategy in Forex forecasting.

Corresponding Author:

Seng Hansun
Department of Informatics, Engineering and Informatics Faculty, Universitas Multimedia Nusantara
Jl. Scientia Boulevard Gading Serpong, Banten, Tangerang, Indonesia
Email: seng.hansun@lecturer.umn.ac.id

1. INTRODUCTION

One of the most popular fast-growing financial markets worldwide is the foreign exchange (Forex or FX) market. The main commodity to be traded in the Forex market is nonetheless the currency pairs. Although the Forex trading practice can be traced back to the time of Babylonians, it was in 1931 when the Forex market was born to replace the gold standard as a trading tool at that time. However, it was in the 1970s to be considered as the modern Forex trading era when the United States allowed its currency in the market [1].

Today, the Forex market has turned to be the world's largest financial market, with more than US$5 trillion total volume traded every day [2]. Therefore, it fascinates many researchers to learn the traded currency pairs' trends and predict their future values to gain more profit from the market. However, since the market is a decentralized market, which means that no single organization or institution controls the market, it is highly volatile and difficult to predict [3]. Various methods and approaches have been introduced to tackle this challenging problem that can be classified into two main groups, i.e., the fundamental analysis group and the technical analysis group. While the fundamental analysis relies on news related to the Forex market, such as the inflation rates, economic growth, et cetera, to predict the behaviour and trend of the market; the technical analysis relies heavily on the historical data of the market [3].

In the technical analysis domain, many techniques to predict the future values of Forex transaction data have been developed. Some researchers focused on the use of conventional and Statistical approaches, such as the moving averages method [4] and its variants [5], vector autoregression [6], and autoregressive integrated moving average (ARIMA) [7], [8]. Others applied more advanced techniques, including Machine Learning and Deep Learning methods.

In predicting the Forex rates of United States Dollar/Turkish Lira (USD/TRY), for example, Goncu [9] applied several regression-based machine learning methods, such as ridge regression, decision tree regression, support vector regression, and linear regression. He found that the ridge regression outperformed other studied methods. In another study, Ni *et al.* [10] proposed a convolutional recurrent neural networks (C-RNN) method, a combination of recurrent neural network and convolutional neural network, for Forex forecasting that could improve the prediction accuracy of considered Forex data. Dodevski *et al.* [11] also had tried to explore the capacity of the long short-term memory (LSTM) deep learning method to predict the exchange rates of Macedonian Denar against Euro (EUR/MKD). They found that the forecasted results were similar to the actual data and therefore, the LSTM method is potentially a useful method for forecasting the EUR/MKD exchange rate fluctuations. In a more recent study, Jung and Choi [12] had proposed a hybrid autoencoder-LSTM model for forecasting Forex volatility. Based on the empirical results, they concluded that the proposed method outperforms the traditional LSTM method.

In this study, we will try to predict the future values of Forex transaction data by using a more recent version of the long short-term memory network, known as the bidirectional LSTM (Bi-LSTM) method. In Bi-LSTM, the bidirectional structure of recurrent neural network (RNN) as proposed by Schuster and Paliwal [13] is employed. Moreover, we apply and compare several modes of the Bi-LSTM method to five major currency pairs in Forex data transactions, namely Australian Dollar/United States Dollar (AUD/USD), Euro/United States Dollar (EUR/USD), Great Britain Pound sterling/United States Dollar (GBP/USD), United States Dollar/Swiss Franc (USD/CHF), and United States Dollar/Japanese Yen (USD/JPY). In brief, some contributions that can be gained from this study are 1) a proposed simple three layers Bi-LSTM networks for Forex forecasting that could compete with other machine and deep learning approaches, 2) we compare and analyse the results of different merge modes available for Bi-LSTM networks, 3) we also compare and analyse the results between Bi-LSTM and conventional LSTM networks for Forex forecasting. Further description of Bi-LSTM as the main deep learning method employed here is given in the next section. The forecasting results of those five major currency pairs will be given and discussed next, followed by the Conclusion section at the end of this paper.

2. METHOD

In this section, first, we briefly describe the concept behind the conventional LSTM and Bi-LSTM methods. Next, three commonly used forecast error criteria and R^2 score is explained as the performance metrics used in this study. The three forecast error criteria are root means square error (RMSE), the mean absolute error (MAE), and the mean absolute percentage error (MAPE).

2.1. LSTM and Bi-LSTM

In the Deep Learning domain, one of the most popular methods applied in various fields, especially for solving regression tasks, is the LSTM networks. Yu *et al.* said LSTM is a special type of RNN that was proposed in 1997 by Hochreiter and Schmidhuber [14]. The main difference between RNN and feedforward neural networks (FNN) is the RNN’s ability to store past information by taking previous layers’ output as an input for the next layer in the networks [15]. However, because it only involves the output of the last layer in the networks, it cannot learn the long-term memory contained in the networks. Therefore, LSTM is introduced for solving the long-term dependency problem found in RNN.

LSTM networks consist of self-connected cells that are used to store the networks’ temporal state by using a three gates mechanism. These connected cells contain the cell state that could maintain the global information of the networks [15]. Figure 1 shows a cell in LSTM with its three gates, i.e., the input gate, the output gate, and the forget gate [16].

The forget gate can be found first. It controls how much information from the previous cell’s hidden state can be forgotten by using a sigmoid function as defined by (1). Next, the input gate can be found, which is used to control how much new information will be stored in the current cell state. The same sigmoid and a new tanh function as defined by (2) will be combined together for the input gate. Lastly, the output gate can be found, which is used to find the new value as the output of the current cell by utilizing the sigmoid function. Some formal Mathematical expressions for all equations in an LSTM cell can be represented as (3) to (8) [17], [18].

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (1)$$

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \quad (2)$$

$$f_t = \sigma\left(W_f h_{t-1} + U_f x_t + b_f\right) \quad (3)$$

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \quad (4)$$

$$\tilde{C}_t = \tanh(W_C h_{t-1} + U_C x_t + b_C) \quad (5)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (6)$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \quad (7)$$

$$h_t = o_t \odot \tanh\,(C_t) \quad (8)$$

f_t, i_t, o_t are the forget, input, and output gates values. C_t and $\tilde{C}_t$ are current and candidate cell state respectively. $W_f, W_i, W_C, W_o, U_f, U_i, U_C, U_o$ are the networks' weights, while b_f, b_i, b_C, b_o are bias values. h_t and h_{t-1} are current hidden and prior hidden state values. Lastly, x_t is a new input value for the current cell.

In this study, a more recent version of LSTM that employs the bidirectional structure of RNN is used. The method is popularly known as the Bi-LSTM networks. Rather than using one LSTM hidden layer, in Bi-LSTM, two hidden layers with similar output are used, but in inverse directions [19]. In the forward direction, Bi-LSTM networks will learn in increasing order of the input sequence, while in the backward direction, it will learn in decreasing order of the input sequence. When the networks have learned from both forward and backward LSTMs separately, their outputs then will be combined into one value by using any merge modes available for Bi-LSTM [20]. There are four different merge modes commonly used in the literature, namely addition, multiplication, concatenation, and average modes [21]. By using this approach, both the past and future information in the dataset can be preserved [22]. An illustration of LSTM versus Bi-LSTM architectures is shown in Figure 2.

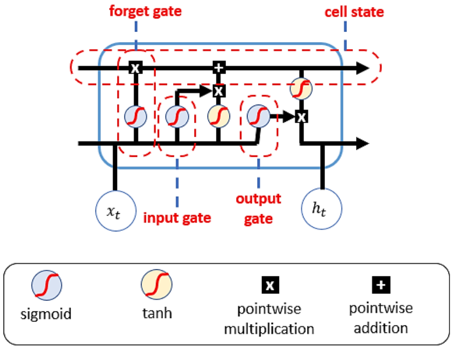

Figure 1. Illustration of an LSTM cell

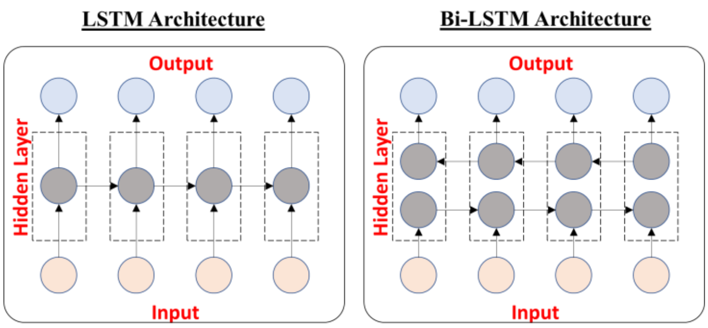

Figure 2. LSTM versus Bi-LSTM architectures [23]

2.2. Performance metrics

In the time series analysis domain, especially for predicting future time series data, there are several forecast error criteria popularly used in the literature. Here, we used three criteria, namely the RMSE, the MAE, and the MAPE. The degree of errors will be given in a unit value for RMSE and MAE, while for MAPE it will be given in a percentage value. All those three error criteria are represented in (9) to (11) where a smaller score implies better forecasting results for a particular forecasting method [24]–[26].

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t - F_t)^2} \quad (9)$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|Y_t - F_t| \tag{10}$$

$$MAPE = \left(\frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_t - F_t}{Y_t}\right|\right) \times 100\% \tag{11}$$

Here, the total number of data is denoted as n, the real observed value is Y_t, and the predicted value is F_t.

Moreover, we also calculated the coefficient of determination, which is usually denoted as R-squared (R^2) score. As an indication of goodness of fit, it shows how well new data samples can be predicted by the model compared to the explained variance. Commonly, R^2 has values that are ranging from 0% to 100%, where 0% means the response variable has no variability around its mean explained by the learned model and 100% means the response variable has all the variability around its mean [24]. However, it might be negative in practice since the learned model can be arbitrarily worse than a constant model. In (12) denotes the formal formulation for the R^2 score [27].

$$R^2 = 1 - \frac{\sum_{t=1}^{n}(Y_t - F_t)^2}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2} \tag{12}$$

where $\bar{Y} = \frac{1}{n}\sum_{t=1}^{n} Y_t$.

3. RESULTS AND DISCUSSION

In this section, the data source, pre-processing, and model development phase are briefly explained. Next, we show the prediction results of all considered Forex currency pairs in this study. Lastly, we will explain the performance results and analysis of this study.

3.1. Data source, pre-processing, and model development

In this study, our focus is to apply the Bi-LSTM Deep Learning method to five major currency pairs in Forex data transactions, namely the AUD/USD, EUR/USD, GBP/USD, USD/CHF, and USD/JPY. We collected the historical data of those five currency pairs from MetaTrader 5 desktop software. MetaTrader 5 is one of the largest trading platforms for Forex and Exchange markets. It is developed by MetaQuotes, a leading developer of software applications for financial markets [28]. We downloaded the last ten years of historical data starting from 3 January 2011 to 25 June 2021. Although there are several features available in the downloaded data, we only consider the ‘Close’ value of each dataset. Moreover, missing values in the dataset (if any) are handled by using simple imputation technique (replace the missing values with their last records). Except for USD/CHF with a total of 2,720 data points found, there are 2,721 data points for each other currency pair without any missing values found, which were further split into training (80%) and test (20%) sets. Therefore, there are 2,176 data points included in the training set (2,175 for USD/CHF) and 545 data points in the test set for each dataset. Figure 3 shows the overall process conducted in this study.

Although many argue that a more complex (deeper) network’s architecture allows for a more precise network that could discover complex patterns in a dataset [29], we argue that with proper parameters’ setting and simple network’s architecture, we can achieve similar results with those that applied more complex networks, especially in time series analysis domain. Therefore, in building the model, we proposed simple three layers Bi-LSTM networks that consist of a Bi-LSTM layer with 100 neurons, a dropout layer which will drop 10% of processed information to prevent overfitting, and a Dense layer with one neuron to represent the networks’ output. mean square error (MSE) and Adam optimizer were also incorporated to compile the networks. The model was trained for 20 epochs with 32 batch sizes each.

3.2. Forecasting results

After all those five Forex currency pairs were pre-processed, we used the proposed Bi-LSTM networks to train and build the best model for each currency pair. Moreover, there are four merge modes of Bi-LSTM networks that will be used in the training process, i.e., addition (sum), multiplication (mul), concatenation (concat), and average (ave). Another deep model development for LSTM networks was also trained in a similar fashion. The learned models then were tested on the test set of each currency pair. The forecasted results for each Forex currency pair are shown in Figure 4 to Figure 8 consecutively. The blue line depicts the actual closing prices of the Forex transaction data, while the red line depicts the forecasted closing prices.

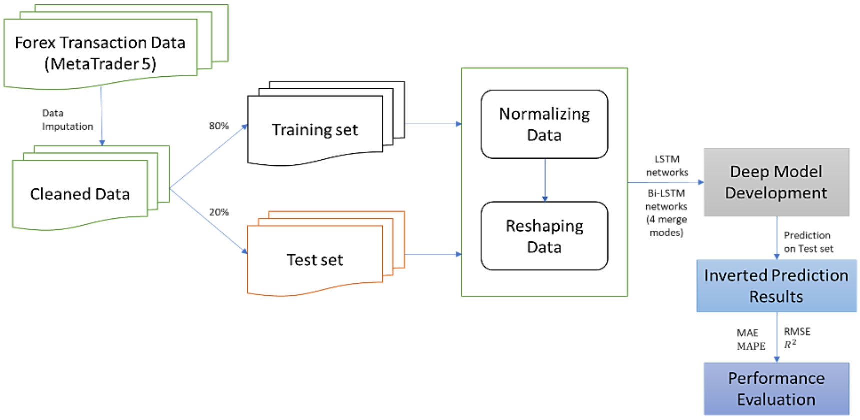

Figure 3. Overall research process

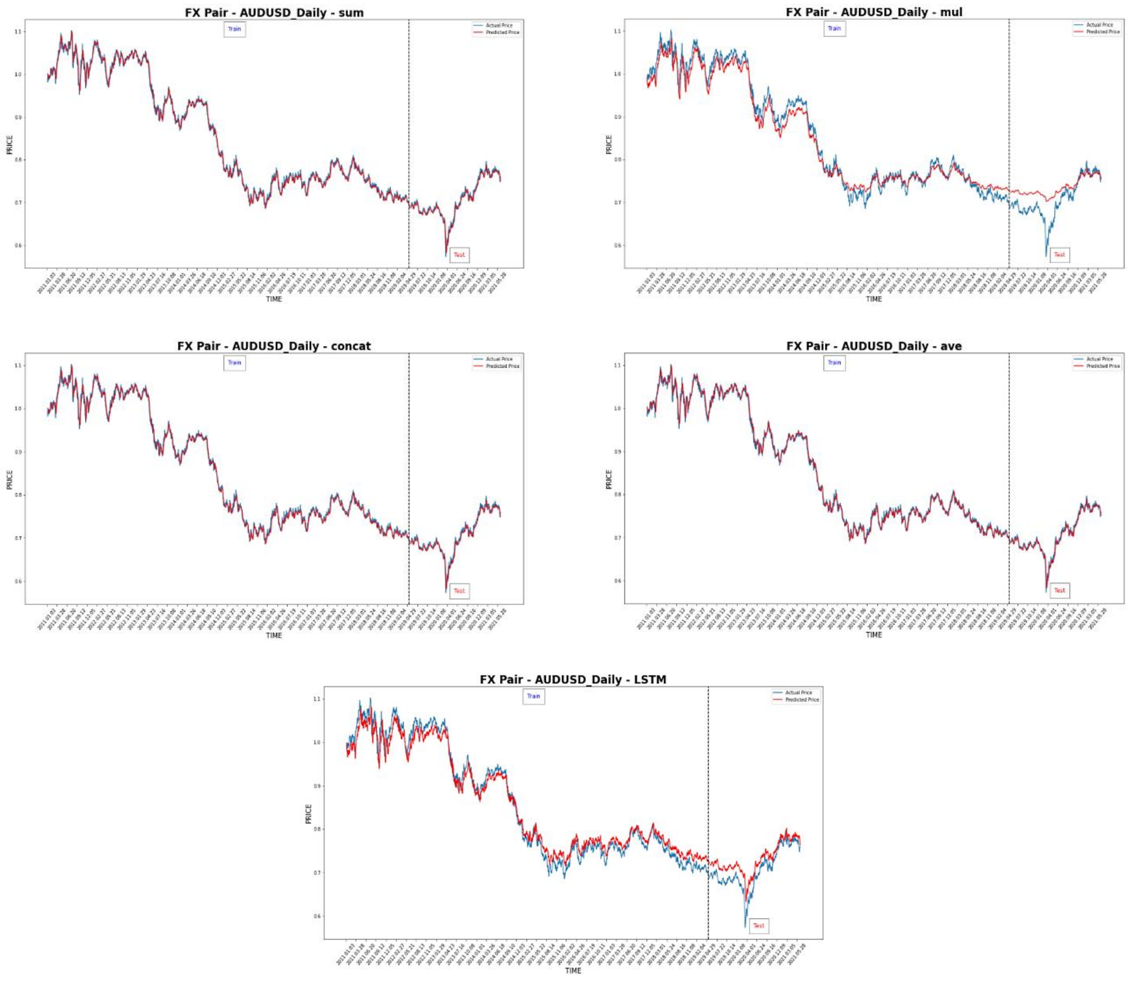

Figure 4. AUD/USD forecasting results for each model and its associating loss function plot

3.3. Performance results

In the performance analysis phase of the forecasting results for all Forex currency pairs and models learned, we used RMSE, MAE, and MAPE criteria. All of them are commonly used prediction error criteria where the model with a smaller value is better than the model with a larger value. The RMSE, MAE, and MAPE values for each Forex currency pair are displayed in Table 1.

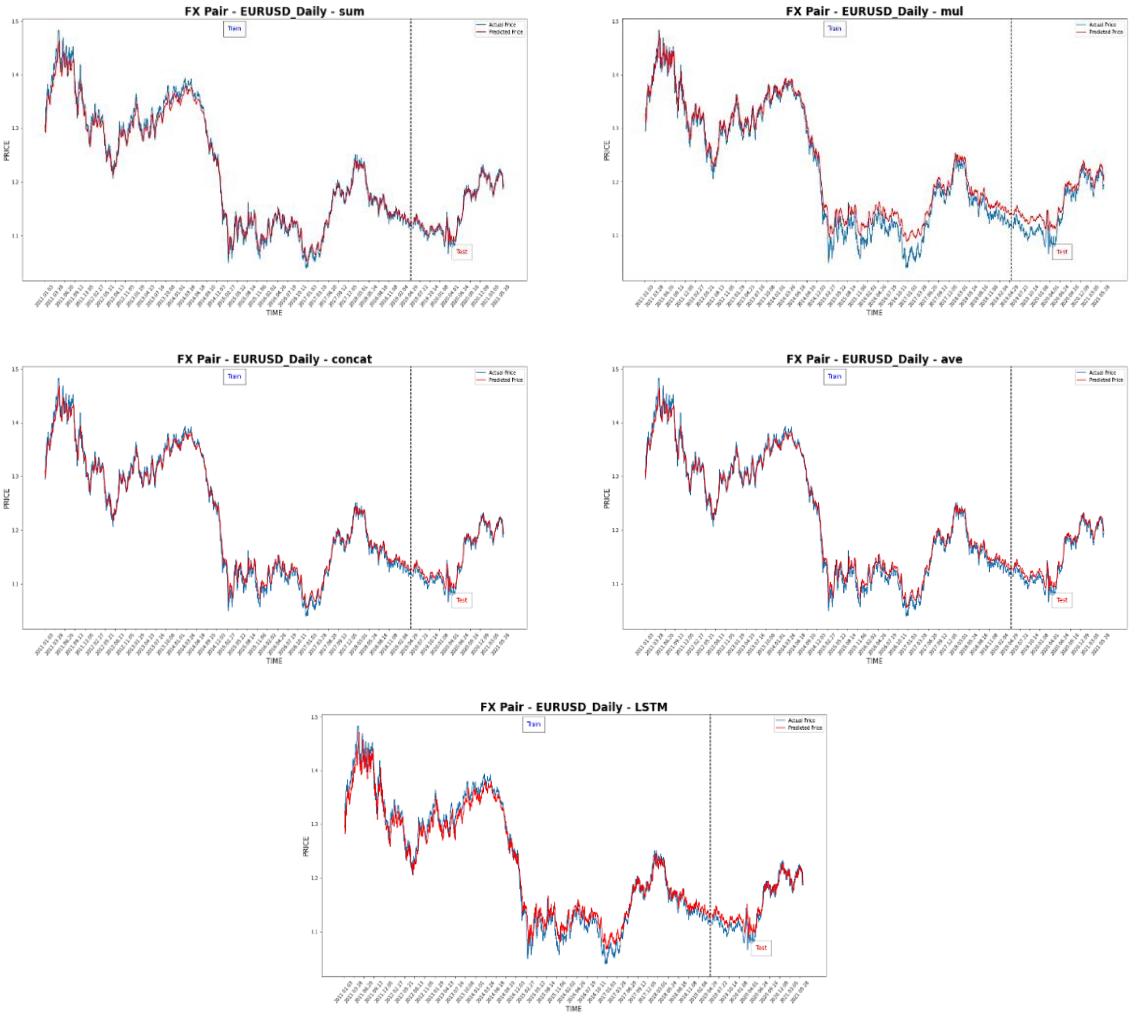

Figure 5. EUR/USD forecasting results for each model and its associating loss function plot

Table 1. Performance results of each currency pair with different merge modes in Bi-LSTM networks and LSTM networks

Currency pair	LSTM	Bi-LSTM			
		Addition	Multiplication	Concatenation	Average
		RMSE			
AUD/USD	0.02751067	0.00778919	0.03736315	0.00808630	0.00750842
EUR/USD	0.01425064	0.00805622	0.02037095	0.00938734	0.01058800
GBP/USD	0.03482888	0.01259609	0.01649531	0.01227448	0.01190310
USD/CHF	0.01599664	0.00796400	0.00750871	0.00809526	0.01143997
USD/JPY	1.11294160	1.27753880	0.88981270	1.49643250	1.08920100
Average	0.24110568	0.26278886	0.19431016	0.30685517	0.22612810
		MAE			
AUD/USD	0.02372061	0.00583180	0.02934104	0.00612274	0.00556377
EUR/USD	0.01142576	0.00629857	0.01802870	0.00742394	0.00863025
GBP/USD	0.02817570	0.00874773	0.01184306	0.00841569	0.00824152
USD/CHF	0.01299167	0.00632688	0.00591476	0.00645287	0.00991210
USD/JPY	0.80173770	1.11262920	0.63021020	1.34369650	0.90701800
Average	0.17561029	0.22796684	0.13906755	0.27442235	0.18787313
		MAPE			
AUD/USD	3.469731%	0.834453%	4.365431%	0.875546%	0.797978%
EUR/USD	1.012452%	0.553618%	1.598428%	0.656146%	0.762623%
GBP/USD	2.225188%	0.680627%	0.932230%	0.656030%	0.641959%
USD/CHF	1.405504%	0.677889%	0.630488%	0.690669%	1.061097%
USD/JPY	0.745092%	1.040946%	0.587187%	1.257148%	0.849830%
Average	1.771593%	0.757506%	1.622753%	0.827108%	0.822697%

As can be seen from Table 1, among four different merge modes tested for Bi-LSTM networks, the multiplication mode got the smallest RMSE and MAE values on average, while the addition mode got the smallest MAPE value on the average. In particular, based on the prediction error criteria, the addition mode is preferred for EUR/USD, the multiplication mode is preferred for USD/CHF and USD/JPY, and the average mode is preferred for AUD/USD and GBP/USD currency pairs. Meanwhile, the concatenation mode as the default merge mode used in the Bi-LSTM is actually the less preferred mode than other available merge modes for Bi-LSTM networks. It got the highest RMSE and MAE values on average, while for MAPE value the highest score is obtained by the conventional LSTM networks.

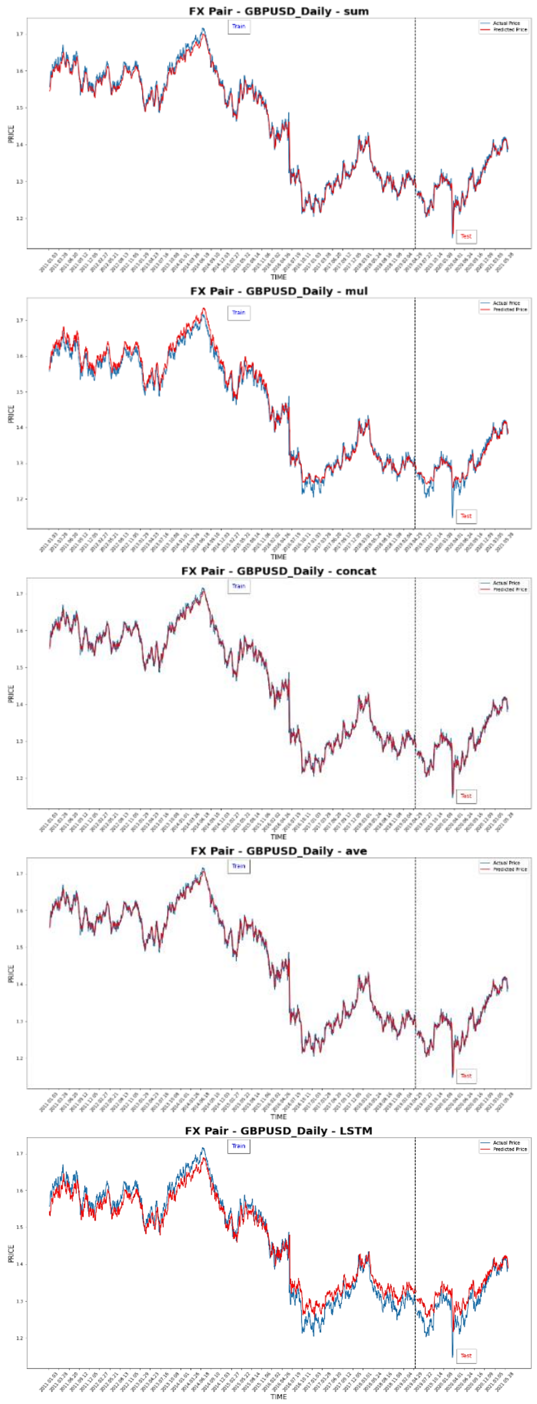

Figure 6. GBP/USD forecasting results for each model and its associating loss function plot

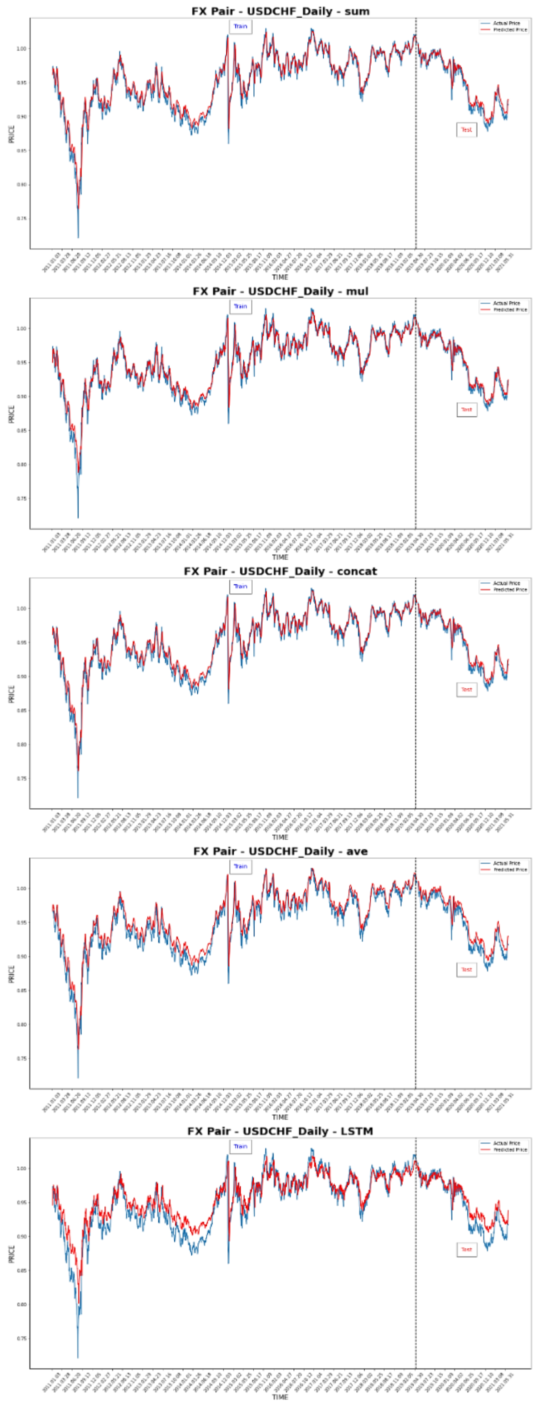

Figure 7. USD/CHF forecasting results for each model and its associating loss function plot

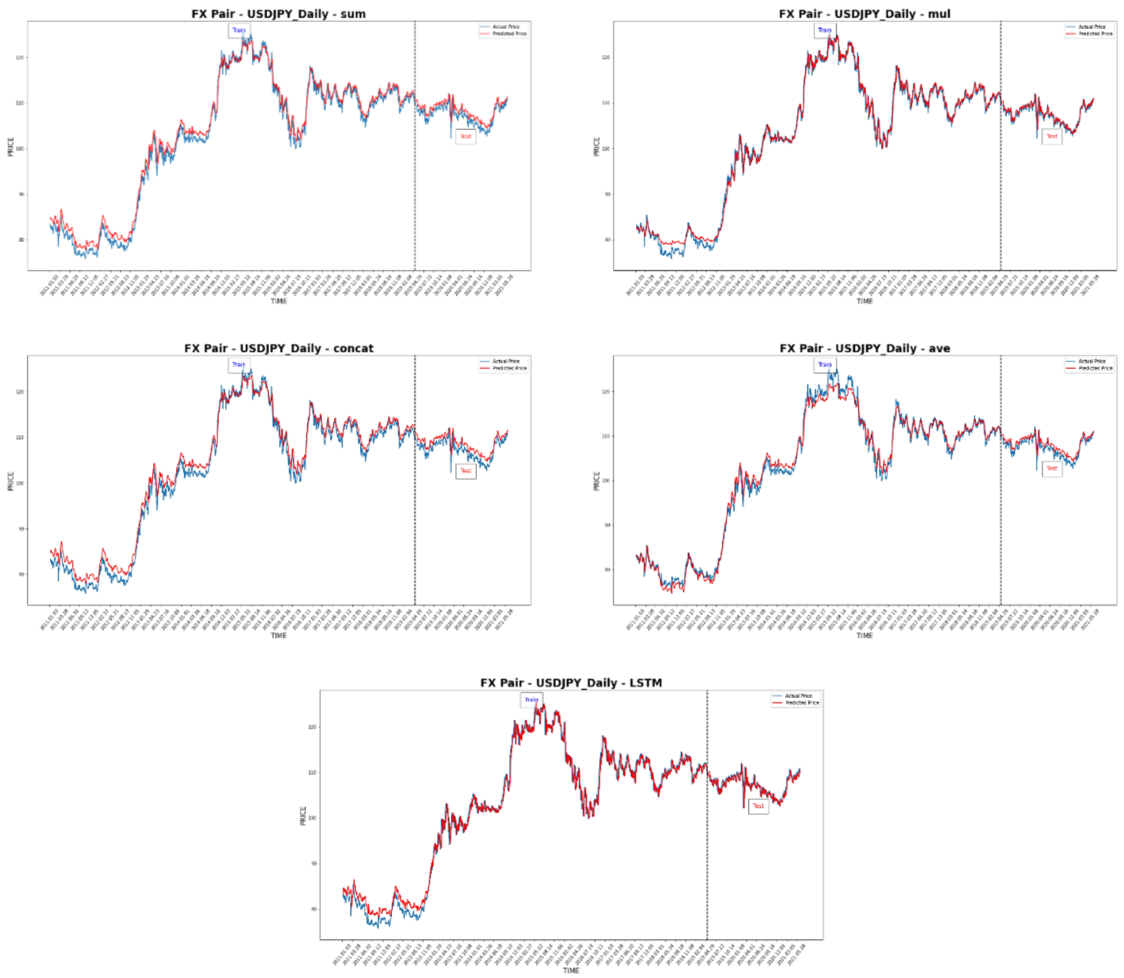

Figure 8. USD/JPY forecasting results for each model and its associating loss function plot

Furthermore, Table 2 shows the R-squared (R^2) score for each Forex currency pair with different merge modes in Bi-LSTM networks and LSTM networks. R^2 or Coefficient of Determination is widely used to estimate the prediction accuracy (performance) in regression tasks [30]. Similar to the previous results, based on the R^2 score, we found that the addition mode is preferred for EUR/USD, the multiplication mode is preferred for USD/CHF and USD/JPY, and the average mode is preferred for AUD/USD and GBP/USD currency pairs. However, since the average mode got the highest R^2 score of all considered Forex currency pairs on the average, it is the preferred mode among other Bi-LSTM merge modes to be used in Forex forecasting. The conventional LSTM itself got the lowest R^2 score of 73.885% among other approaches in this study, hence it is the less preferred method to be used.

Table 2. R^2 score of each currency pair on the test set

Currency pair	LSTM	Bi-LSTM			
		Addition	Multiplication	Concatenation	Average
R-squared					
AUD/USD	63.424770%	97.067967%	32.536056%	96.840022%	97.275528%
EUR/USD	90.639397%	97.008437%	80.872524%	95.938187%	94.832709%
GBP/USD	65.997989%	95.552693%	92.373122%	95.776894%	96.028576%
USD/CHF	81.493154%	95.412905%	95.922391%	95.260458%	90.534907%
USD/JPY	67.868403%	57.661473%	79.460741%	41.909943%	69.224609%
Average	73.884743%	88.540695%	76.232967%	85.145101%	89.579266%

Lastly, we also recorded the computation time for model development of each method during the training phase. Table 3 shows the recorded computation time. As can be seen from the results, the conventional LSTM method needs a shorter training time for model development than all other Bi-LSTM approaches. This finding can be understood because, unlike the LSTM method where the network is traversed once, Bi-LSTM propagates the network twice in both forward and backward directions.

Our study confirms the finding of another study by Siami-Namini *et al.* [31], where the Bi-LSTM excels the conventional LSTM in forecasting financial time series data. Although there is a trade-off between computation time and prediction performance of LSTM versus Bi-LSTM approaches, the differences in computation time are not significant. Moreover, the proposed three layers Bi-LSTM networks could achieve similar prediction performance results compared to other machine learning and deep learning methods that used deeper network's architecture, as reported in Aryal *et al.* [32] who used LSTM, convolutional neural networks (CNN), and temporal convolution networks (TCN); Qi *et al.* [33] who used RNN, LSTM, Bi-LSTM, and gated recurrent unit (GRU); and Dautel *et al.* [34] who employed FNN, RNN, LSTM, and GRU. Moreover, we could also try to compare the prediction results from this study with other Machine and Deep Learning methods commonly used in the literature, such as naïve Bayes [35], GRU [36], and Random Forest Regressor [37], a popular tree-based algorithm.

Table 3. Computation time during model development

Currency pair	LSTM	Bi-LSTM			
		Addition	Multiplication	Concatenation	Average
Training Time (seconds)					
AUD/USD	12.052388	19.268188	17.830289	17.756003	19.334207
EUR/USD	15.975887	19.946968	23.977128	20.143736	21.638367
GBP/USD	14.695210	28.014959	26.256105	24.679351	25.149869
USD/CHF	15.543050	25.277671	28.777432	25.630180	26.636279
USD/JPY	17.801251	21.298236	30.806167	29.044963	28.751902
Average	15.213557	22.761204	25.529424	23.450847	24.302125

4. CONCLUSION

In this study, we tried to predict the future values of five major Forex currency pairs, namely AUD/USD, EUR/USD, GBP/USD, USD/CHF, and USD/JPY, by using a well-known Deep Learning method, the Bi-LSTM method. Here, we proposed simple three layers of Bi-LSTM networks, which were further tested on four different merge modes available for the networks. Another test and comparison with the conventional LSTM network which has the same architecture were also conducted. Among available Bi-LSTM merge modes, we found that the default merge mode (concatenation) is actually the less preferred mode used for Forex forecasting. Specifically, based on all performance metrics used in this study (RMSE, MAE, MAPE, R^2), Bi-LSTM addition mode is most preferred for EUR/USD, Bi-LSTM multiplication mode is most preferred for USD/CHF and USD/JPY, and Bi-LSTM average mode is most preferred for AUD/USD and GBP/USD. It was also confirmed from the experimental results that the conventional LSTM could learn and build the model in a shorter time than the Bi-LSTM approaches. However, in general, its prediction performance is still inferior to Bi-LSTM. To get a better result, specific experiments on each currency pair should be conducted. Another study on the effect of different activation functions used in the model development could also be taken in the future.

ACKNOWLEDGEMENTS

Authors thank Universitas Multimedia Nusantara (Indonesia), Middle Tennessee State University (USA), and Universitas Tarumanagara (Indonesia) for the support provided in this study.

REFERENCES

[1] B. David, "The History of Forex," *Foundation Trading Knowledge*, pp. 1–4, 2018, [Online]. Available: https://www.dailyfx.com/education/beginner/history-of-forex.html.

[2] D. C. Yıldırım, I. H. Toroslu, and U. Fiore, "Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators," *Financial Innovation*, vol. 7, no. 1, 2021, doi: 10.1186/s40854-020-00220-2.

[3] S. Ahmed, S. U. Hassan, N. R. Aljohani, and R. Nawaz, "FLF-LSTM: A novel prediction system using Forex Loss Function," *Applied Soft Computing Journal*, vol. 97, 2020, doi: 10.1016/j.asoc.2020.106780.

[4] S. Hansun and M. B. Kristanda, "Performance analysis of conventional moving average methods in forex forecasting," *Proceeding of 2017 International Conference on Smart Cities, Automation and Intelligent Computing Systems, ICON-SONICS 2017*, vol. 2018-January, pp. 11–17, 2017, doi: 10.1109/ICON-SONICS.2017.8267814.

[5] S. Hansun and Subanar, "Brown's weighted exponential moving average implementation in forex forecasting," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 15, no. 3, pp. 1425–1432, 2017, doi: 10.12928/TELKOMNIKA.v15i3.5410.

[6] P. Taveeapiradeecharoen and N. Aunsri, "A Large Dimensional VAR Model with Time-Varying Parameters for Daily Forex Forecasting," *Wireless Personal Communications*, vol. 115, no. 4, pp. 3051–3072, 2020, doi: 10.1007/s11277-020-07531-8.
[7] C. U. Yıldıran and A. Fettahoğlu, "Forecasting USDTRY rate by ARIMA method," *Cogent Economics and Finance*, vol. 5, no. 1, 2017, doi: 10.1080/23322039.2017.1335968.
[8] R. J. Abreu, R. M. Souza, and J. G. Oliveira, "Applying singular spectrum analysis and arimagarch for forecasting eur/usd exchange rate," *Revista de Administracao Mackenzie*, vol. 20, no. 4, 2019, doi: 10.1590/1678-6971/eRAMF190146.
[9] A. Goncu, "Prediction of exchange rates with machine learning," *ACM International Conference Proceeding Series*, 2019, doi: 10.1145/3371425.3371448.
[10] L. Ni, Y. Li, X. Wang, J. Zhang, J. Yu, and C. Qi, "Forecasting of Forex Time Series Data Based on Deep Learning," *Procedia Computer Science*, vol. 147, pp. 647–652, 2019, doi: 10.1016/j.procs.2019.01.189.
[11] A. Dodevski, N. Koceska, and S. Koceski, "Forecasting Exchange Rate Between Macedonian Denar and Euro Using Deep Learning," *Journal of Applied Economics and Business*, vol. 6, no. 2, pp. 50–61, 2018, [Online]. Available: http://www.aebjournal.org/articles/0602/060204.pdf.
[12] G. Jung and S. Y. Choi, "Forecasting foreign exchange volatility using deep learning autoencoder-LSTM techniques," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/6647534.
[13] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997, doi: 10.1109/78.650093.
[14] L. Yu, S. Liang, R. Chen, and K. K. Lai, "Predicting monthly biofuel production using a hybrid ensemble forecasting methodology," *International Journal of Forecasting*, vol. 38, no. 1, pp. 3–20, 2022, doi: 10.1016/j.ijforecast.2019.08.014.
[15] J. Long, A. Q. M. Khaliq, and K. M. Furati, "Identification and prediction of time-varying parameters of COVID-19 model: a data-driven deep learning approach," *International Journal of Computer Mathematics*, vol. 98, no. 8, pp. 1617–1632, 2021, doi: 10.1080/00207160.2021.1929942.
[16] M. Phi, "Illustrated Guide to LSTM's and GRU's: A step by step explanation. Accessed May 2021.," *TowardsDataScience*, 2018, [Online]. Available: https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21.
[17] P. Wang, X. Zheng, G. Ai, D. Liu, and B. Zhu, "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 and Fractals*, vol. 140, 2020, doi: 10.1016/j.chaos.2020.110214.
[18] S. Wen *et al.*, "Memristive LSTM Network for Sentiment Analysis," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 3, pp. 1794–1804, 2021, doi: 10.1109/TSMC.2019.2906098.
[19] F. Shahid, A. Zameer, and M. Muneeb, "Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM," *Chaos, Solitons and Fractals*, vol. 140, 2020, doi: 10.1016/j.chaos.2020.110212.
[20] T. Le, M. T. Vo, B. Vo, E. Hwang, S. Rho, and S. W. Baik, "Improving electric energy consumption prediction using CNN and Bi-LSTM," *Applied Sciences (Switzerland)*, vol. 9, no. 20, 2019, doi: 10.3390/app9204237.
[21] J. A. Kumar and S. Abirami, "Ensemble application of bidirectional LSTM and GRU for aspect category detection with imbalanced data," *Neural Computing and Applications*, vol. 33, no. 21, pp. 14603–14621, 2021, doi: 10.1007/s00521-021-06100-9.
[22] A. Zeroual, F. Harrou, A. Dairi, and Y. Sun, "Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study," *Chaos, Solitons and Fractals*, vol. 140, 2020, doi: 10.1016/j.chaos.2020.110121.
[23] S. L. Shen, P. G. Atangana Njock, A. Zhou, and H. M. Lyu, "Dynamic prediction of jet grouted column diameter in soft soil using Bi-LSTM deep learning," *Acta Geotechnica*, vol. 16, no. 1, pp. 303–315, 2021, doi: 10.1007/s11440-020-01005-8.
[24] F. Rustam *et al.*, "COVID-19 Future Forecasting Using Supervised Machine Learning Models," *IEEE Access*, vol. 8, pp. 101489–101499, 2020, doi: 10.1109/ACCESS.2020.2997311.
[25] N. F. Omran, S. F. Abd-El Ghany, H. Saleh, A. A. Ali, A. Gumaei, and M. Al-Rakhami, "Applying Deep Learning Methods on Time-Series Data for Forecasting COVID-19 in Egypt, Kuwait, and Saudi Arabia," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/6686745.
[26] S. Subanar, C. R. Indrati, T. Gherman, S. Hansun, and V. Charles, "A Tuned Holt-Winters White-Box Model for COVID-19 Prediction," *International Journal of Management and Decision Making*, vol. 20, no. 1, p. 1, 2021, doi: 10.1504/ijmdm.2021.10034422.
[27] D. Zhang, "A Coefficient of Determination for Generalized Linear Models," *American Statistician*, vol. 71, no. 4, pp. 310–316, 2017, doi: 10.1080/00031305.2016.1256839.
[28] M. 5, "Company," 2021, [Online]. Available: https://www.metatrader5.com/en/company.
[29] A. Nazemi, H. Rezazadeh, F. J. Fabozzi, and M. Höchstötter, "Deep learning for modeling the collection rate for third-party buyers," *International Journal of Forecasting*, vol. 38, no. 1, pp. 240–252, 2022, doi: 10.1016/j.ijforecast.2021.03.013.
[30] T. A. Nguyen, H. B. Ly, and V. Q. Tran, "Investigation of ANN Architecture for Predicting Load-Carrying Capacity of Castellated Steel Beams," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/6697923.
[31] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The Performance of LSTM and BiLSTM in Forecasting Time Series," *Proceedings - 2019 IEEE International Conference on Big Data, Big Data 2019*, pp. 3285–3292, 2019, doi: 10.1109/BigData47090.2019.9005997.
[32] S. Aryal, D. Nadarajah, D. Kasthurirathna, L. Rupasinghe, and C. Jayawardena, "Comparative analysis of the application of Deep Learning techniques for Forex Rate prediction," *2019 International Conference on Advancements in Computing, ICAC 2019*, pp. 329–333, 2019, doi: 10.1109/ICAC49085.2019.9103428.
[33] L. Qi, M. Khushi, and J. Poon, "Event-Driven LSTM for Forex Price Prediction," *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering, CSDE 2020*, 2020, doi: 10.1109/CSDE50874.2020.9411540.
[34] A. J. Dautel, W. K. Härdle, S. Lessmann, and H.-V. Seow, "Forex exchange rate forecasting using deep recurrent neural networks," *Digital Finance*, vol. 2, no. 1–2, pp. 69–96, 2020, doi: 10.1007/s42521-020-00019-x.
[35] Y. Vernanda, S. Hansun, and M. B. Kristanda, "Indonesian language email spam detection using n-gram and naïve bayes algorithm," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 5, pp. 2012–2019, 2020, doi: 10.11591/eei.v9i5.2444.
[36] S. Krishnan, P. Magalingam, and R. Ibrahim, "Hybrid deep learning model using recurrent neural network and gated recurrent unit for heart disease prediction," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 6, pp. 5467–5476, 2021, doi: 10.11591/ijece.v11i6.pp5467-5476.
[37] M. A. Rasyidi, T. Bariyah, Y. I. Riskajaya, and A. D. Septyani, "Classification of handwritten javanese script using random forest algorithm," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 3, pp. 1308–1315, 2021, doi: 10.11591/eei.v10i3.3036.

BIOGRAPHIES OF AUTHORS

Seng Hansun received the Bc. degree in Mathematics (S.Si.) from Universitas Gadjah Mada, Yogyakarta, in 2008 and Master of Computer Science (M.Cs.) degree from the same university in 2011. Since then, he has been a Lecturer with the Computer Science Department, Universitas Multimedia Nusantara (UMN), Indonesia. He had been appointed as the Information and Communication Technology (ICT) Faculty Research Coordinator, Deputy Head of Computer Science Department, and Head of Informatics Department at UMN. He had published two books and more than 135 articles during his career as both academician and researcher at UMN. His research interests lately include computational science, soft computing methods, and internet and mobile technology in various fields, especially in the Medical Informatics area. He can be contacted at email: seng.hansun@lecturer.umn.ac.id.

Farica Perdana Putri her bachelor's degree in Computer Science from Universitas Multimedia Nusantara. She then continued her study in Computer Science at National Taipei University. She has been a lecturer and researcher at Universitas Multimedia Nusantara since 2017. She is also a member of APTIKOM and actively publishes many publications. Her research interests mainly focus on natural language processing, computer vision, and machine learning where she has successfully been granted some research grants from the government. Now, she is pursuing her PhD at Khalifa University, UAE, and doing research in deep learning and bioinformatics data search. She can be contacted at email: farica@umn.ac.id.

Abdul Q. M. Khaliq is a Professor in the Department of Mathematical Sciences at Middle Tennessee State University, USA. His area of expertise is Numerical Analysis and Scientific Computing, with specialization in the "Development of Stable and Efficient Numerical Methods for Large Scale Systems of Partial Differential Equations (PDEs)". His primary research interest lies at the interface of applied mathematics and computational science. His research includes the development of new and robust algorithms using partial differential equations with applications to complex financial and physical systems. He is the Editor-in-Chief of the International Journal of Computer Mathematics and actively contributes to several other journals as Associate Editor and Member of Editorial Board. He can be contacted at email: Abdul.Khaliq@mtsu.edu.

Hugeng Hugeng is an Associate Professor at Electrical Engineering Department, Universitas Tarumanagara, Jakarta, Indonesia. His current research interest includes audio signal processing, Internet of Things, image processing, speech recognition, deep learning, and machine learning. He was born in Lubuk Linggau, Indonesia. He worked as research assistant at RWTH Aachen University, Germany in 2001, and at TU Kaiserslautern, Germany in 2002-2005. He achieved his Ph.D. degree in Electrical Engineering from Universitas Indonesia in 2011. Dr. Hugeng has been actively involved in IEEE activities and since September 2016, he is appointed in a key position in IEEE SPS Indonesia Chapter as Technical Activities Coordinator. He serves currently as a Review Editor for IEEE Access. He is a senior member of IEEE and served as general chair of the 2017 IEEE ICON-SONICS. Since 2018, he is appointed as chairman of the Tarumanagara International Conference on the Applications of Technology and Engineering (TICATE) where its Proceedings is indexed by Scopus.com. He can be contacted at email: hugeng@ft.untar.ac.id.