Applied Mathematics and Computational Intelligence
Volume 12, No. 4, Nov 2023 [13-26]

CNN-LSTM Hybrid Model for Improving Bitcoin Price Prediction Results

Ferdiansyah[1], Raja Zahilah[2*], Sit Hajar[3], Deris Stiawan[4]

[1]Department of Informatics, Universitas Bina Darma, Palembang, Indonesia
[2,3]Department of Computer Science, Faculty of Computing, Universiti Teknologi Malaysia, Johor Bahru, Johor, Malaysia
[4]Faculty of Computer Knowledge, Universiti Sriwijaya, Bukit Besar, Palembang, Indonesia

*Corresponding Author: zahilah@utm.my

Received: 27 Oct 2022
Accepted: 8 Sept 2023

ABSTRACT

LSTM is a promising tool for predicting the stock exchange. Still, when the LSTM Model faces an anomaly problem with a dataset of Bitcoin that has hit more change in value by fluctuation, it can be a problem for producing good evaluation results such as RMSE. This research is an improvement over the discoveries of previous research. We tried another perspective besides using five years of historical data prices to predict a six-day value. We found that the results of RMSE were not very good but exhibited good results on MAPE as a comparison evaluation method. We are using the last six days to predict the next day. Logically, this dataset has good dataset stability, but the dataset has quite a significant minute-by-minute change in day-by-day value. Furthermore, CNN-LSTM was selected in this research to give another perspective and improve the results. The results were quite good and greatly improved previous research.

Keywords: *Cryptocurrency, Bitcoin prediction, Bitcoin Stock Market Prediction, CNN, LSTM.*

1 INTRODUCTION

The LSTM Models have been successfully applied in numerous cases and situations, especially in time series forecasting. Moreover, LSTM is a promising tool for predicting stock exchange, digital money, and cryptocurrency [1][2][3][4] as well as for short-term or long-term prediction. However, in various cases, there were also several problems, including that LSTM managed to predict well. Still, the evaluation results showed a significant enough error model, so improvements were needed from the model side to improve the evaluation results.

This research focuses on providing alternative solutions to Bitcoin prediction models using the previous LSTM models and improving the evaluation results. This study will refer to our previous research on Bitcoin prediction using LSTM with a case study of the yahoo-finance stock market in 2019, the data range will be different because it is no longer effective, and the value of Bitcoin has changed very significantly at the end of 2019. The Bitcoin value was recorded in the range of 7,200 USD in the middle of 2020 until now, and the fluctuation value of Bitcoin has risen significantly. In the beginning of June 2020, the value of Bitcoin is 9,656 USD. However, starting from October 2020, it was 11,296 USD. After that, the value of Bitcoin continued to rise to its peak on 15 April 2021, where the value of Bitcoin at that time was 63,314 USD for each Bitcoin. Then, it gradually decreased again by almost 10,000 USD; until finally on 6 June 2021 when we took the dataset for our research, the value of Bitcoin was 38,566 USD [5].

This study's endeavor makes some contributions in the domain of cryptocurrency prediction.

1) Initially, we proposed the Hybrid Algorithm CNN and LSTM model for daily cryptocurrency pricing.
2) To eliminate RMSE anomalies caused by changes due to variations from varied value fluctuations. For the dataset itself, the researcher changes the time span taken from the long-term model with five-year data. With Bitcoin fluctuations becoming increasingly erratic, to update the model to detect and take a view of a stable Bitcoin value with not too sharp fluctuations every month for a period of five years, so that historical values in the years behind it are difficult to predict the next Bitcoin value. With this initiative, the researchers changed the dataset model to weekly with a range of 9.585 lines with a record of the movement of the Bitcoin value per minute for 1 week, so that NaN values or data noise did not occur as much as in the five-year Bitcoin dataset.

The rest of the paper is organized as follows: Section 2 addresses relevant research, Section 3 discusses research technique, Section 4 discusses experimental findings and analysis, and Section 5 summarizes this study. Afterward, we must mention about many unknown factors, including political and economic turmoil at both the national and international scales, that affect stock prices. As a result, the LSTM price prediction of Bitcoin is not sufficient in making the decision to invest in Bitcoin, which requires another side for decision-making.

2 LITERATURE

According to the Papers using LSTM, Sean et al. [6] used Recurrent Neural Networks and Long Short-Term Memories to predict Bitcoin prices. Meanwhile, Ruchi Mittal et al. [7] used machine learning and past trends to suggest automating the process of predicting the price of cryptocurrency (daily trend). Recurrent Neural Network-Based Bitcoin Price Prediction Using Twitter Sentiment Analysis was proposed by Dibakar Raj Pant et al. The correlation between emotion and LSTM outcomes is pretty good. [8] Chih-Hung et al. [9] established a novel framework for predicting Bitcoin price using LSTM, proposing two different LSTM models (standard LSTM and LSTM with AR(2) model) with 208 records of dataset, comparing MSE, RMSE, MAE, and MAPE. Fei Qian et al.[3] developed a common stock market prediction model based on LSTM under different elements that affect the market. In this research, they picked three stocks with comparable tendencies. The LSTM prediction model worked perfectly.

Our previous model [10] is shown in Table 1. The algorithm has proven successful in predicting Bitcoin from Yahoo Finance's stock market. The time series model can produce results, and these results can be used to project future prices using the data split for training and testing described in the aforementioned article. Nonetheless, the RMSE is not acceptable, which is negative.

3 METHODOLOGY

Bitcoin prediction employs the same prediction technique as traditional stock exchange[11][12][13]. Furthermore, we can acquire different results by combining techniques, such as using time series data analysis, market technical analysis that uses historical data price, and combining them with some algorithm to enrich the results. Time series data analysis is on historical data which provides data by periods, namely by minutes, hourly, weekly, monthly, quarterly, and yearly, depending on the usage of researchers.

Table 1. LSTM results

No	Epoch	Model Dropout	RMSE Results
1	10	0	631.74963
2	100	0	455.98107
3	1,000	0	825.37505
4	200	0	360.64511
5	400	0	354.18368
6	500	0	288.59866
7	800	0	292.78967
8	2,000	0	477.91428
9	5,000	0	474930575

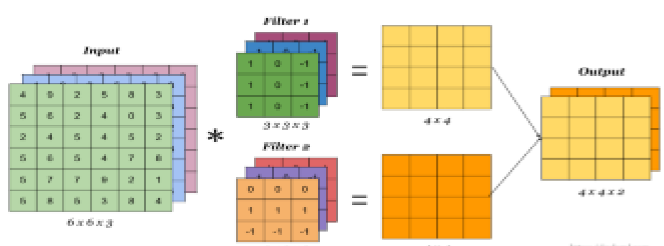


Figure 1. Convolutional Neural Network [14]

3.1 CNN

A Convolutional Neural Network (CNN) makes specific assumptions about the amount of the input data. As illustrated in Figure 1, CNN includes a layer of Rectified Linear Unit (ReLU) as an activation function, pooling, and is fully connected to the layers [15][16]. Pooling (POOL) aims to minimize the footprint of the network's representation by cutting down on the number of parameters and calculations needed. CNN, batch, normalization, ReLU, activation, and pooling were used sequentially to generate each pipeline. Each pipeline's output is aggregated into a single forecast [14][17]. There is a sub-sampling layer in the CNN layer that can be used to lower the background noise in the feature maps, and then there is a regression layer.

3.2 LSTM

LSTM (Long Short-Term Memory), which was developed by Hochreiter and Schmidhuber, is another module type for RNN (1997)[18]. It was subsequently refined and popularised by several researchers. The LSTM network (LSTM network) consists of modules with recurrent consistency, similar to RNN.

- i_i = The input gate signifies that the cell's information will be updated.
- f_t = The forget gate signifies that data should be removed from the cell.
- O_t = A measure of how much data is sent out of the system through the output gate.
- c_t = The potential value for the memory cell's states at time t.

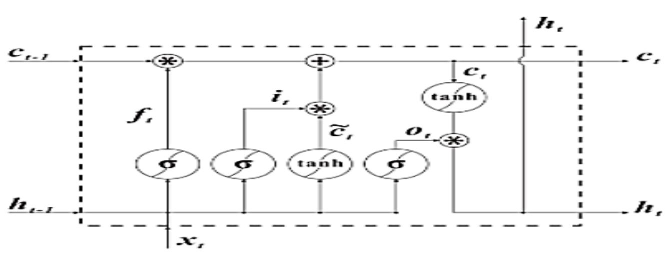


Figure 2. LSTM memory cell structure of the hidden layer [3]

- c_t = The state of the current memory cell at time t, which calculated by the combination of i_I and c_t f_t and c_{t-1} through element-wise multiapplication
- h_t = The average value of the output after going through the output gate.
- σ = is shorthand for a sigmoid function that takes a value and places it between -1 and 1, and its range is 0 to 1.

3.3 Propose Method

Experiments in Figure 3. were conducted with the same dataset used in this study, generated without using CNN, to determine the accuracy of the resulting model errors. The results obtained by MAPE 8,790 based on MAPE were deemed to be the most accurate. The obtained results were extremely precise, but the RMSE score did not reflect this. In this instance, a value close to 0 is optimal if the measurement is performed with RMSE. it is crucial to identify the most effective method for minimizing the consequence of RMSE.

Table 2. LSTM Results experiment using LSTM with minute by minute dataset.

No	Epoch	Training		Validation		Dev	
		MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
1	10	49.85	1902	50.67	1946	52.44	2082
2	100	10.09	3836	10.00	3844	10.19	4053
3	200	8.790	3346	8.863	3403	8.808	3501
4	500	9.948	3783	10.04	3857	10.09	4008
5	1000	10.43	3965	10.47	4020	10.46	4158
6	2000	10.25	3897	10.03	3966	10.33	4106
6	5000	10.18	3948	10.01	3965	10.68	4175

The framework explained in Figure 3 is the proposed model of the Hybrid CNN-LSTM algorithm, starting with the input dataset and calculations using windows function (windows size = 15 and batch size) based on further random tests. These inputs have better results and are optimal for computation, followed by a convolutional layer with two layers, namely the first layer using a kernel size of 5 and the second layer with a kernel size of 3. Then, the max pooling unit size is 2. For the use of LSTM, it is the same as using CNN before using 2 layers of LSTM with the same number of layers in each layer; the number of units per layer for the first layer is 128 units, and the second layer's number of units is 192, with mode sequences = true [15].

A flatten, which is used for input, separates the CNN-LSTM layers. For example, if it is applied to a layer that has an input form (batch_size, 2,2), the output form of that layer will be (batch_size, 4). Then, proceed with the Dense layer, which is a layer of deeply connected neural networks. This is the most common and frequently used coating. A solid layer that performs the operations below on the input and returns the output.

Dense in this model uses three solid layers with different units. In the first 3 layers, the number of units is 128, 31, and 1 with ReLU activated on 2 layers; the last layer is Dropout. Dropout is a technique used to prevent the model from overfitting. Dropout works by randomly setting the training exit edge of the hidden unit (neurons that make up the hidden layer) to 0 at each phase with input 0.2 and 0.1. The last step is the optimisation process with Adamax.

Furthermore, instead of using five years of historical data, we used the short-term seven-day data, for short-term prediction [19][20][21] because the fluctuation value of Bitcoin has risen and fallen spectacularly from the middle of 2020 until now.

We have been doing prediction trials for the long-term prediction, but based on the fluctuations of the existing circumstances. We can make predictions, however, since the purpose of this research is to improve on the findings of previous research, we tried to find another way by trying to make short-term predictions. Besides that, the error and percentage good accuracy cannot be expected from the previous long-term LSTM model. In the previous model, we can predict the time up to six days ahead, but we can only predict for the next day in this model.

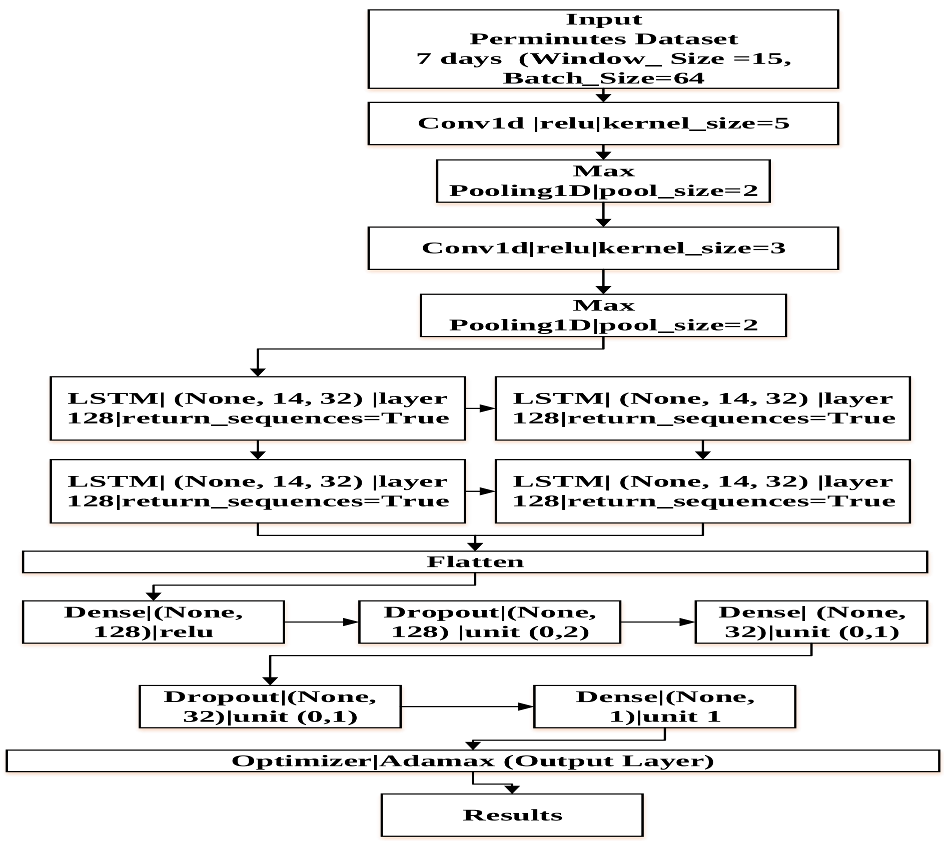


Figure 3. CNN-LSTM Model

3.4 Evaluation

a. RMSE

RMSE is Root Mean Square Error. RMSE will always be greater or equal to MAE. RMSE measures a model's continuous value prediction ability. The RMSE is in the same units as the dependent variable or target of your data, which helps determine if the error is substantial. RMSE decreases model performance. The higher the model's performance, the lower the RMSE [22][23][24].

$$\textbf{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2} \tag{1}$$

Where N is the total number of observations, x_i represents the actual value, whereas $\hat{x}_i$ represents the predicted value. The principal advantage of RMSE is that it penalises significant errors. Additionally, it scales the scores in the same units as the predicted values [25].

b. MAPE Mean Absolute Percentage Error (MAPE):

In addition to Mean Absolute Deviation (MAD) and Root Mean Squared Error (RMSE), MAPE measures relative error. MAPE is more instructive than MAD, since it expresses the proportion of error in estimating or forecasting actual results over a certain period, indicating whether the error is too high or too low. MAPE is the average absolute error multiplied by 100% [26][27].

$$\textbf{MAPE} = \sum_{t=1}^{n} \left| \frac{\gamma i - \widehat{yi}}{\widehat{yi}} \right| \times 100\% \tag{2}$$

Information:

MAPE = mean absolute percentage error

n = amount of data

y = actual yield value

ŷ = value of estimation result

Interpretation of MAPE Values:
Based on Lewis (1982) [28], MAPE values can be interpreted or interpreted into four categories:

< 10% = very accurate

10 – 20% = good

20 – 50% = reasonable

> 50% = inaccurate

The smaller the MAPE value, the smaller the error in the estimation results; on the contrary, the larger the MAPE value, the greater the error in the estimation results. The results of a prediction method have very good forecasting ability if the MAPE value is < 10% and have good predictive ability if the MAPE value is between 10% and 20%.

3.5 Dataset

Stock market data is gathered from Yahoo Finance using the CCC - CryptoCompare USD exchange rate. This study employed time-series data on this experimental dataset, which consisted of 9,585 rows and was stored in CSV format and represented prices in U.S. dollars over a seven-day period from 20 May 2021 to 26 May 2021. Dataset samples are displayed in Figure 4. Preprocessing dataset using windows function, windowing means to take a dataset and partition it into subsections (which increases the dimension shape of the dataset). In traditional machine learning, more input data tends to be better. However, in the time series, that might not be the case.

	Open	High	Low	Close	Adj Close	Volume	y
0	39169.38	39169.38	39169.38	39169.38	39169.38	0	39169.38
1	39196.16	39196.16	39196.16	39196.16	39196.16	0	39196.16
2	39153.21	39153.21	39153.21	39153.21	39153.21	0	39153.21
3	39027.18	39027.18	39027.18	39027.18	39027.18	102473728	39027.18
4	38966.34	38966.34	38966.34	38966.34	38966.34	0	38966.34

Figure 4. Sample dataset frame.

4 RESULTS AND DISCUSSIONS

Figures below show the pre-processing process to load datasets into machines and split data into training, validation, and development dataset. We divided the data into 80% for training, 15% for validation dataset, and 5% for dev test (test set), with all splits conducted in a single operation.

In this research, we used Anaconda and Python. For the simulation and visualization, we used Jupyter Notebook to display the results of our research. Table 2 shows some trial random epoch parameters to examine the parameter that produced the best result for this research,

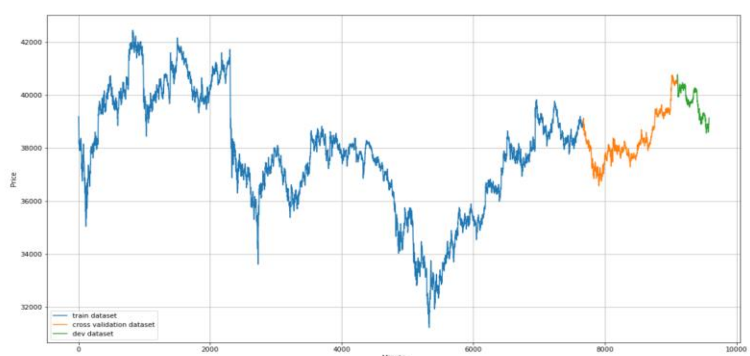


Figure 5. Split data

Table 3. Results

No	Epoch	Training		Validation		Dev	
		MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
1	10	3.421	1630	1.617	899	4.053	1680
2	100	0.729	365	0.444	223	0.381	196
3	200	0.550	280	0.450	207	0.304	155
4	500	0.699	318	0.717	300	0.609	266
5	1,000	0.361	174	0.346	151	0.323	144
6	2,000	0.355	172	0.215	172	0.393	172
6	5,000	0.481	212	0.517	216	0.681	278

Regarding the results above, epoch 1,000 is observed to have the best result out of other epochs that we have tried on experimental research. According to [28], MAPE score under < 10% is very accurate. The MAPE score for training is 0.361, validation is 0.346, dev is 0.323, and the RMSE result for training is 174, validation is 151, and dev is 144. For this, RMSE is the best and the smallest among other experiments. This is a significant improvement from previous research whereby the RMSE score is 288. The RMSE itself has the same units as the variable, which means there is no absolute good or bad threshold, for example, if the data range is 0 – 1,000 RMSE with a value of 0.7 is small, but 1 is not small anymore. One more example is if the prediction range is $500 – $1,000, then $15 is good, but overall, a value closer to 0 is better. The interpretation of multiple scores in a regression problem (i.e., RMSE, MAE, and MSE) all depends on your problem domain and what you think is acceptable. RMSE scores cannot be compared unless they come from the same regression problem. In this case, the researcher compares the MAPE score and sees the density of the prediction visualization results in Figure 6, 7, and 9 from all predictions, trainings, validations, and tests that have good results and are close to the actual Bitcoin value.

Based on Figure 8, the visualization describes about all results of prediction using training, validation, and development data. As the graphic or spectrum is very tight, it is hard to see the gap between the real value (blue line) and the prediction value (orange line). In the last row of dataset, the last close real value of Bitcoin on that day (May 26, 2021) was 39,391 USD. After that, we checked again on May 27, 2021, where the value of Bitcoin was 39,316 USD. The value was not too far, but instead slightly close with the value on the last close day. In Figure 9, the visualization displays 80% of the training result, and the result is close enough to prediction results.

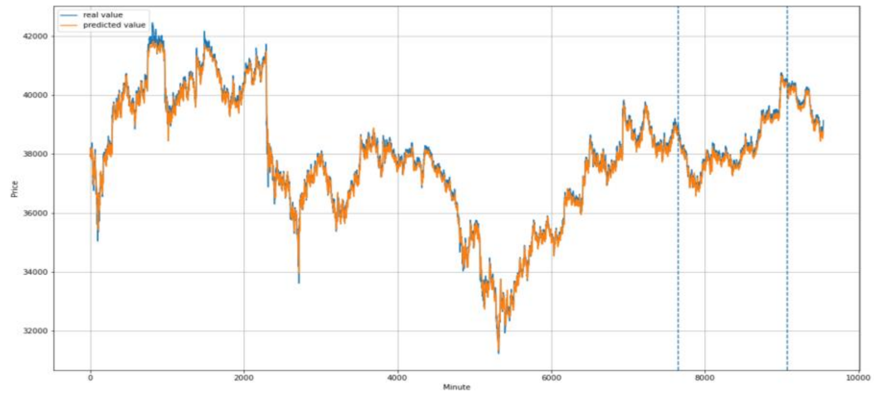


Figure 6. Prediction

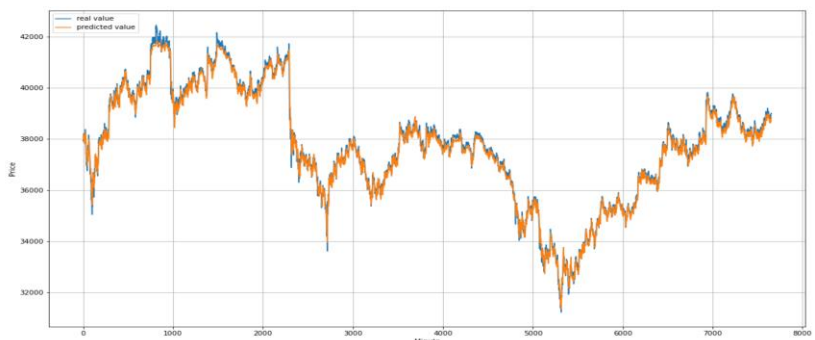
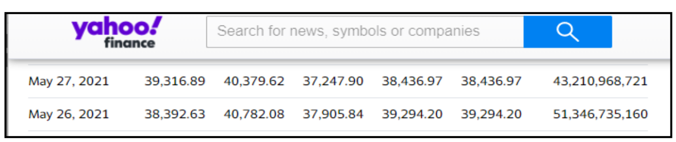


Figure 7. Training



May 27, 2021	39,316.89	40,379.62	37,247.90	38,436.97	38,436.97	43,210,968,721
May 26, 2021	38,392.63	40,782.08	37,905.84	39,294.20	39,294.20	51,346,735,160

Figure 8. Real value of Bitcoin on the Yahoo Finance Market [5]

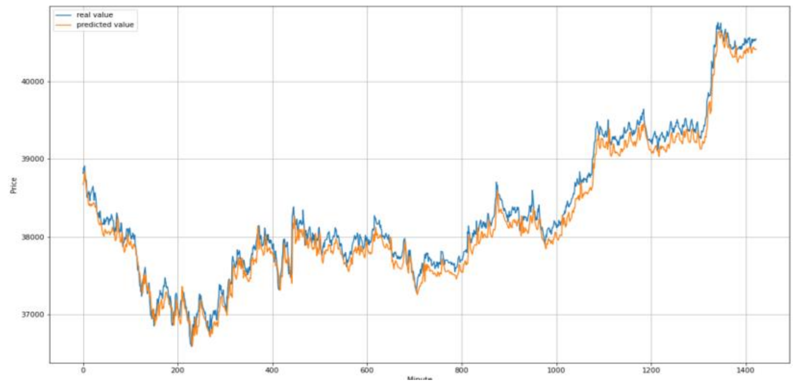


Figure 9. Validation

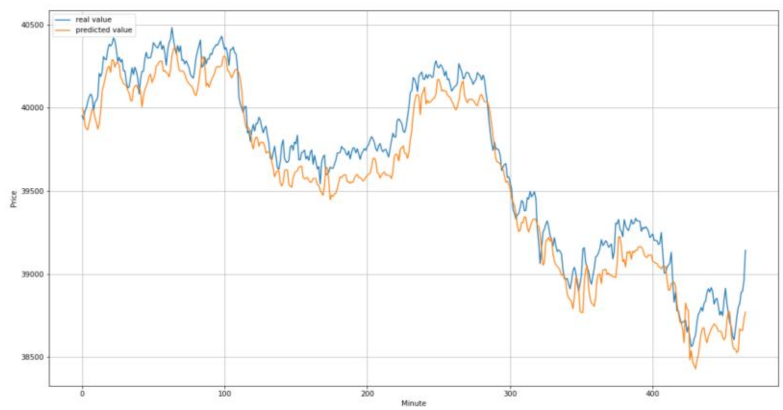


Figure 10. Dev Test

According to Figure 10, the visualization describes approximately 15% of the validation prediction result for the minutes 0 to 1,400. For the graphic or spectrum, we can see the gap between real value (blue line) and the prediction value (orange line) for the minutes 1,000 to 1,200. If calculated manually, minutes 0 to1,400 is on the first day of May 20, 2021, and the closing value on that day touched 41,539 USD. There is a slight gap, but the figure is still at 41,000 USD. Figure 9 explains the use of the remaining 5% dataset with the minute range of 0 – 400, meaning that the results are not too tight for the first eight hours of the same day due to the use of a small dataset. Then, the value of Bitcoin fluctuations that change every minute with a little data comparison causes the spectrum to be less good. The Bitcoin value at first minute to minute 400 is 39,678, while predictions based on the dev test itself are a little less good at around 38,700 USD.

5. CONCLUSION

In this paper, we have proposed a hybrid CNN-LSTM model for improvement and given another perspective on predicting Bitcoin. Previous research on Bitcoin prediction only used LSTM and five years format of historical data prices to predict for a few days, and the result of the model was good but the RMSE score still needed to be optimized. And in previous experiments using the same dataset as this study without using CNN, the results were the same, so a model that can optimize the measurement of the RMSE score is required. Several approaches have been exploited in this paper to predict Bitcoin price. The dataset was collected from the Yahoo Finance Market in a seven-day and per-minute format. In this paper, we have tried to optimize, maximize in some parts, and give comparative results using MAPE as compared with RMSE to see the results' accuracy. Furthermore, CNN-LSTM produced relatively better results on our dataset in this study. It is compared with previous LSTM research results. Moreover, the results may not be the best for RMSE, as we know the closer value to 0 is better, but we have significantly improved results. Our challenge ahead for future research is finding out how to solve anomaly problems with high fluctuation like Bitcoin, possibly by experimenting with optimization and fixing noise data that is not too much maximized in this model. The more often Bitcoin fluctuates, the harder it is to get a good final RMSE value.

ACKNOWLEDGEMENT

The authors would like to acknowledge the Ministry of Higher Education Malaysia for funding this research under the Fundamental Research Grant Scheme (FRGS) with the Registration Proposal No: FRGS/1/2022/ICT07/UTM/02/1 and the Universiti Teknologi Malaysia (UTM). We also would like to express our appreciation to the ISTEC-CoSQA 2023 for the opportunity to present our research findings.

REFERENCES

[1] A. Radityo, Q. Munajat, and I. Budi, “Prediction of Bitcoin exchange rate to American dollar using artificial neural network methods,” in *Advanced Computer Science and Information Systems (ICACSIS), 2017 International Conference on*, 433–438, 2017.

[2] D. M. Q. Nelson, A. C. M. Pereira, and R. A. de Oliveira, “Stock market's price movement prediction with LSTM neural networks,” in *2017 International joint conference on neural networks (IJCNN)*, 1419–1426, 2017.

[3] F. Qian and X. Chen, “Stock Prediction Based on LSTM under Different Stability,” in *2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, 483–486, 2019.

[4] J. Brownlee, “Time series prediction with lstm recurrent neural networks in python with keras,” *Available Mach. com*, 18, 2016.

[5] C.- CryptoCompare, “Yahoo Finance,” 2021. [Online]. Available: https://finance.yahoo.com/. [Accessed: 14-Feb-2022].

[6] S. McNally, J. Roche, and S. Caton, “Predicting the price of Bitcoin using Machine Learning,” in *Parallel, Distributed and Network-based Processing (PDP), 2018 26th Euromicro International*

Conference on, 339–343, 2018.

[7] R. Mittal, S. Arora, and M. P. S. Bhatia, “Automated Cryptocurrencies Prices Prediction Using Machine Learning,” 2018.

[8] D. R. Pant, P. Neupane, A. Poudel, A. K. Pokhrel, and B. K. Lama, “Recurrent Neural Network Based Bitcoin Price Prediction by Twitter Sentiment Analysis,” in *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, 128–132, 2018.

[9] C.-H. Wu, C.-C. Lu, Y.-F. Ma, and R.-S. Lu, “A New Forecasting Framework for Bitcoin Price with LSTM,” in *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, 168–175, 2018.

[10] F. Ferdiansyah, S. H. Othman, R. Z. R. M. Radzi, D. Stiawan, Y. Sazaki, and U. Ependi, “A LSTM-Method for Bitcoin Price Prediction: A Case Study Yahoo Finance Stock Market,” in *2019 International Conference on Electrical Engineering and Computer Science (ICECOS)*, 206–210, 2019.

[11] M. Jarrah and N. Salim, “A recurrent neural network and a discrete wavelet transform to predict the Saudi stock price trends,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 4, 155–162, 2019.

[12] Z. He, J. Zhou, H. N. Dai, and H. Wang, “Gold price forecast based on LSTM-CNN model,” *Proc. - IEEE 17th Int. Conf. Dependable, Auton. Secur. Comput. IEEE 17th Int. Conf. Pervasive Intell. Comput. IEEE 5th Int. Conf. Cloud Big Data Comput. 4th Cyber Sci.*, 1046–1053, 2019.

[13] L. Yu, S. Wang, and K. K. Lai, “A neural-network-based nonlinear metamodeling approach to financial time series forecasting,” *Appl. Soft Comput. J.*, vol. 9, no. 2, 563–574, 2009.

[14] S. A. Dwivedi, A. Attry, D. Parekh, and K. Singla, “Analysis and forecasting of Time-Series data using S-ARIMA, CNN and LSTM,” in *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 131–136, 2021.

[15] J. Eapen, D. Bein, and A. Verma, “Novel deep learning model with CNN and bi-directional LSTM for improved stock market index prediction,” in *2019 IEEE 9th annual computing and communication workshop and conference (CCWC)*, 264–270, 2019.

[16] W. Yu, W. Zhifei, W. Hongye, Z. Junfeng, and F. Ruilong, “Prediction of Passenger Flow Based on CNN-LSTM Hybrid Model,” *Proc. - 2019 12th Int. Symp. Comput. Intell. Des. Isc.*, 132–135, 2019.

[17] J. Rasheed, A. Jamil, A. Ali Hameed, M. Ilyas, A. Ozyavas, and N. Ajlouni, “Improving Stock Prediction Accuracy Using CNN and LSTM,” *2020 Int. Conf. Data Anal. Bus. Ind. W. Towar. a Sustain. Econ. ICDABI 2020*, 2020.

[18] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[19] P. Subagyo, “Forecasting: konsep dan aplikasi,” 2000.

[20] R. S. Russel, “dan Taylor, BW (2011). Operations Management: Along the Supply Chain.” NJ: Wiley.

[21] E. D. S. Mulyani, A. Bachtiar, D. Rifki, I. Yogaswara, N. S. Tyas, and others, “Prediksi Kurs Rupiah Terhadap Dollar Amerika Menggunakan Metode Trend Moment,” *INOVTEK Polbeng-Seri Inform.*, vol. 4, no. 2, 147–156, 2019.

[22] Squark, “ROOT MEAN SQUARE ERROR OR RMSE.” [Online]. Available: https://squarkai.com/root-mean-square-error-or-rmse/#.XWVCLegzZPY. [Accessed: 28-Aug-2019].

[23] N. Vandeput, “Forecast KPI: RMSE, MAE, MAPE & Bias,” 2019. [Online]. Available: https://medium.com/analytics-vidhya/forecast-kpi-rmse-mae-mape-bias-cdc5703d242d. [Accessed: 10-Apr-2020].

[24] B. Santoso, A. I. S. Azis, and others, *Machine Learning & Reasoning Fuzzy Logic Algoritma, Manual, Matlab, & Rapid Miner*. Deepublish, 2020.

[25] N. J. Salkind, *Encyclopedia of research design*, vol. 1. SAGE Publications, Inc., 2010.

[26] V. Gaspersz, “Production planning and inventory control,” *PT Gramedia Pustaka Umum, Jakarta*, 2004.

[27] U. Khair, H. Fahmi, S. Al Hakim, and R. Rahim, “Forecasting error calculation with mean absolute deviation and mean absolute percentage error,” in *Journal of Physics: Conference Series*, vol. 930, no. 1, 12002, 2017.

[28] C. D. Lewis, *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth-Heinemann, 1982.