Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia

Vol. 16, 2023, page 25 – 35

Implementation of LSTM and MLP on Cryptocurrency Price Forecasting

Chan Mei Hua, Nur Arina Bazilah Kamisan*, Haliza Abd Rahman
Department of Mathematical Sciences, Faculty of Science, UTM, Skudai, Johor Bahru, Malaysia
*Corresponding author: nurarinabazilah@utm.my

Abstract
The purpose of this study is to investigate the application and efficiency of two machine learning models on cryptocurrency price forecasting. The Long-Short Term Memory model, a variant of recurrent neural networks known for its ability to capture long-term dependencies, is implemented and trained on the dataset. Similarly, the Multi-Layer Perceptron model, a traditional feedforward neural network, is trained using the same data. the dataset of one of the most well-known cryptocurrencies known as Bitcoin is used as the dataset to test on machine learning models. The objective of this study is to compare between the models using performance metrics with 120 days of future forecast. The LSTM outperforms the MLP model across different evaluation metrics. The results also suggest that LSTM model offer more accurate and reliable cryptocurrency price forecasts compared to MLP model, as evidenced by lower MAE and MAPE values. These insights can guide investors and traders in making informed decisions.

Keywords: Cryptocurrency; Bitcoin; machine learning; Long-Short Term Memory (LSTM); Multi-Layer Perceptron (MLP)

1. Introduction

Cryptocurrencies are digital or virtual currencies that use cryptography for security. They are decentralized and typically operate on a technology called blockchain, which is a distributed ledger that records all transactions across a network of computers (Amsyar et al., 2020). Cryptocurrencies are typically not controlled or regulated by any central authority unlike traditional currencies issued by governments. There is no organisation or agency contribute to the value of currency or issues more currency into the system (Tredinnick, 2019). Therefore, scarcity is maintained through the process of investing the cryptocurrency to valid transactions. This digital currency has many types such as Bitcoin, Ethereum, Litecoin, Ripple, Monero and many more (Amsyar et al., 2020). Cryptocurrency transaction is normally a deal from one individual to another individual online without involving third parties (Nasir et al., 2020). It provides secure and transparent transactions, offer potential anonymity for users, and enable cross-border transactions with low fees and fast settlement times. Additionally, cryptocurrencies have also attracted investors as speculative assets, as their values can be subject to significant volatility.

Machine learning is a branch of artificial intelligence (AI) and computer science that has the potential to tackle big data for which classical methods are not applicable (Bi et al., 2019). It also uses the way how humans learn and gradually improving its accuracy. This is where it merges information from neuroscience and biology, mathematics, statistics and physics to make computer learn. Additionally, machine learning is about making computers to adapt or modify their working whether or not these actions are making predictions or controlling a robot, and with that these actions are more accurate (Marsland, 2015). First and foremost, different feature selection techniques are used in Jiang (2021) to predict the forecast of cryptocurrency in Bitcoin. This paper presents a modified model of traditional K-Nearest Neighbour algorithm (KNN) and compare with logistic regression model and traditional KNN model. KNN algorithm is also one of the regression algorithms in machine learning. The improved model required combining the closing price of short-term trend into one sample. As a result, the improved KNN model has proved that it has more accurate price prediction with the lowest Root Mean Squared Error (RMSE) value than the other two models.

A comparative performance of cryptocurrency price forecasting is being studied further based on machine learning algorithms such as Support Vector Machine (SVM), Artificial Neural Networks (ANNs) and Deep Learning (DL) by Hitam & Ismail (2018). The authors concluded that SVM is the most reliable forecasting model for cryptocurrency with the least error. In addition, an optimised Support Vector Machine (SVM) based on Particle Swarm Optimisation (PSO) in short-form as SVM-PSO outperforms the single model from the previous study (N. A. Hitam et al., 2019).

Long-Short Term Memory (LSTM) are capable of learning long-term dependencies that is use to model time series data (Lazzeri, 2021). It helps to reduce the risk of losing important information where it maintains the error that can be backpropagated through time and layers. LSTM is also able to regulate the flow of information through the internal mechanisms called gates and cell state (Zhang et al., 2019). A related paper about LSTM model by Lahmiri & Bekiros (2019) investigated that digital currencies exhibit chaotic characteristics in which it enables deep learning LSTM neural networks have significantly higher efficient in forecasting compared to Generalised Regression Neural Networks (GRNN). Despite that LSTM took much longer time during convergence phase, whilst GRNN which is a one-pass algorithm utilised a Gaussian kernel deduced training less than a second but still LSTM outperforms. This is because deep learning neural systems are capable of extracting concealed information from the underlying signals by memorising short and long-term temporal data.

There is a demand of LSTM since it has greater accuracy of price prediction for cryptocurrency in most studies. This is especially notable that LSTM is unlike ARIMA and Facebook's Prophet (Fbprophet) does not require on certain information about the data such as date field or time series stationarity (Rathore et al., 2022). Therefore, this case study will explain the forecasting method of cryptocurrency on Bitcoin by utilising statistical technologies and LSTM model to obtain more accurate result.

The study by Albariqi & Winarko (2020) described the baseline models to predict both long-term and short-term Bitcoin price using two models; Multi-Layer Perceptron (MLP) and Recurrent Neural Networks (RNN). The dataset that was used for this study is 2-days period data from August 2010 until October 2017. This paper justified that both models are equally efficient when predicting long-term than short-term price change. However, the accuracy of MLP model outperforms RNN model where MLP is effective when predicting for the next 60-days price change while RNN has best accuracy for 56-days. Both MLP and RNN models have been formulated extensively and more advanced than previous research which makes their research studies to predict precise Bitcoin price. As a result, MLP has better performance than RNN model.

Other than the well-known and most used models such as Long-Short Term Memory (LSTM), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM) and Recurrent Neural Network (RNN). There are also improved and combined models between each other to represent the upgraded version model. For instance, Guo et al. (2021) proposed a new model WT-CATCN incorporated with Wavelet Transforms (WT) and Casual Multi-Head Attention (CA) Temporal Convolutional Network (TCN). This Bitcoin forecasting method was tested on a randomly chosen time period for around 15 months from the 15th of June 2017 to the 21st of September 2018. The performance of this model had been compared with other traditional models and deep learning models that are consist of the Autoregressive Integrated Moving Average

(ARIMA), Autoregressive Integrated Moving Average Exogenous (ARIMAX), Convolutional Neural Networks (CNN), Multi-Layer Perceptron (MLP), Long Short-Term Memory (LSTM), Sequence to Sequence (Seq2Seq), Bayesian Neural Networks (BNN) and State-Frequency Memory (SFM). The authors proved that WT-CATCN significantly performed over other models by 25% in terms of Root Mean Squared Error (RMSE).

The error metrics are referred to as performance indicator which is used to quantify the performance of a model. It also acts as an evaluation tool to quantitatively compare different models for forecasters to execute which models objectively produce the best results. Some of the error metrics used by forecasters to evaluate the prediction values are Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). Other deep learning models such as Long-Short Term Memory (LSTM), Bayesian Neural Network (BNN) and State-Frequency Memory (SFM) have better forecasting results in which SFM achieved the best performance with smaller RMSE value (Guo et al., 2021). However, SFM is less performed while comparing to the model for WT-CATCN proposed by Guo et al. (2021).

In addition, there were a study that relates to the factors affecting oil market and crypto-assets markets. The researchers have fitted copulas to pair before and after Covid-19 returns to analyse the changes in the dependence structure. They have utilised different types of copulas in the study and the findings on copulas is evaluated according to the Akaike Information Criterion (AIC) (Mzoughi et al., 2022). The finding supports that all markets show a significant long memory parameter and the observed changes on dependence structure provided valuable information on how Covid-19 pandemic affects inter-dependencies.

In this research, the machine learning models for Long-Short Term memory (LSTM) and Multi-Layer Perceptron (MLP) will be utilised to predict price change. The main contribution involves comparing the performance accuracy between Long-Short Term memory (LSTM) and Multi-Layer Perceptron (MLP) by using error metrics.

2.0 Materials and Methods

A. Long-Short Term Memory (LSTM)

LSTM, has been developed by Hochreiter and Schmidhuber in 1997 (Lindemann et al., 2021). LSTM is used to handle the vanishing gradient problem occurs with conventional Recurrent Neural Networks (RNN) in which it refers to the partial derivatives of loss function that vanishes after it approaches a value close to zero when there are more layers in the network.

The LSTM model has five required components to allow the model to process both long-term and short-term data. Each of the LSTM modules in the Figure 1 is a set of units that capture the data flow from one to another. These units connect from one module to another to collect the current data by transmitting past data on closing price.

The procedure of the LSTM walk through begins by deciding what information are supposed to omit from the cell state. The decision of passing through the data are made by the first sigmoid layer in the model that is forget gate, f_t. The gates are composed of sigmoid neural network layer and a pointwise operation. The purpose of sigmoid layer is to produce the numbers between zero and one where zero means let nothing through while one represents let everything through the gate. The forget gate, f_t layer looks at previous cell state, C_{t-1} and current input, X_t and produces a number between zero and one for each number in the previous cell state, C_{t-1}.

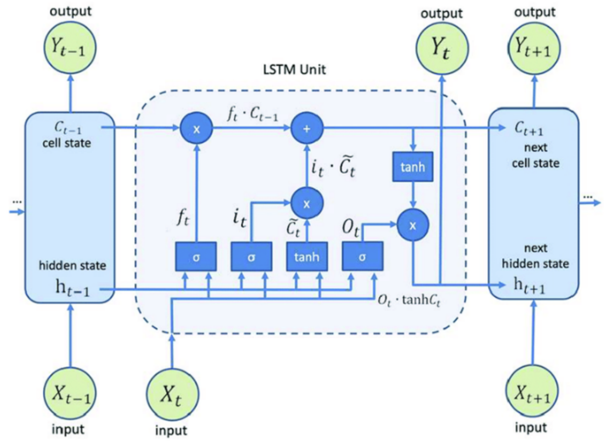

Figure 1 Structure of LSTM model.

Next, a sigmoid layer called input gate, i_t layer decides which values are to be updated along with the tanh layer creates a vector of new candidate values, $\widetilde{C}_t$. This step is to determine what information are going to store in the cell state.

It is needed to update the previous cell state, C_{t-1} to the current cell state, C_t. The previous steps in the LSTM unit have the tasks to complete. Therefore, the forget gate, f_t is multiplied with the previous cell state, C_{t-1}. This step is to forget the things that had been decided to be forgotten earlier. Then, it is added with the input gate, i_t multiplied by the new candidate values, $\widetilde{C}_t$. This is where to drop the old information and update the LSTM unit with new information.

Moving on, it is time to decide for the output values, Y_{t+1}. The last sigmoid layer called output gate layer, O_t is run to decide what part of the cell state to be processed. The cell state is then put through tanh layer to push the values between plus minus one and multiplied by the output of the sigmoid gate.

B. Multi-Layer Perceptron (MLP)

Multi-Layer Perceptron (MLP) is a continuation and variant of the original Perceptron model proposed by Rosenblatt in 1950 (Idrissi & Amine, 2016). It was found that the Perceptron was only capable of handling linearly arranged data, hence, MLP was introduced to overcome the problem. MLP is then capable of handling both linear and nonlinear data. It is widely used since the mapping is non-linear between inputs and output. The MLP model consist of three k ey components namely input layer, hidden layer and output layer. MLP model has one or more hidden layers between input and output layers. The connections between the input and output layers are always directed from lower layers to upper layers where the neurons are organised in layers (Idrissi & Amine, 2016).

Figure 2 is a simple neuron that has weighted input signals and produce an output signal using activation function. The input layer receives the input data and pass it to hidden layers. The feature in the input data in corresponded by nodes. The quantity of the nodes also equals to the number of features in the input data. The hidden layers act as transforming input data into output layer. The input layer connected by nodes in hidden layer and connected to other hidden layers through weights and biases. The single node works as multiply the input data to an assigned weight value and then add a bias before passing it to the next layer. The weight in artificial neural networks is the parameter that transforms input data within hidden layer network.

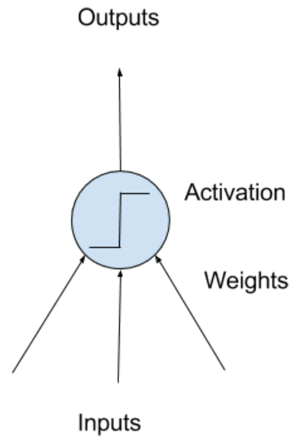

Figure 2 Single neuron model.

Biases, on the other hand, represents how far of the predictions are from the intended value. Some common activation function used in the hidden layers are tanh, sigmoid and rectified linear unit (ReLU). The last layer in MLP neural networks is output layer in which it receives the transformed representation of the input data from hidden layer to achieve the final output. The output layer also consists of nodes which is continuous and depends upon the tasks. It is important to note that the training process involves transforming the input data through the network, computing error, and then updating the weights and biases to reduce the error.

The difference between LSTM and MLP is that LSTM has great memory feature in the nodes whereas MLP has two algorithms namely feedforward and backpropagation algorithms where both of the algorithms work best together to minimise the error between its predictions and the actual outputs.

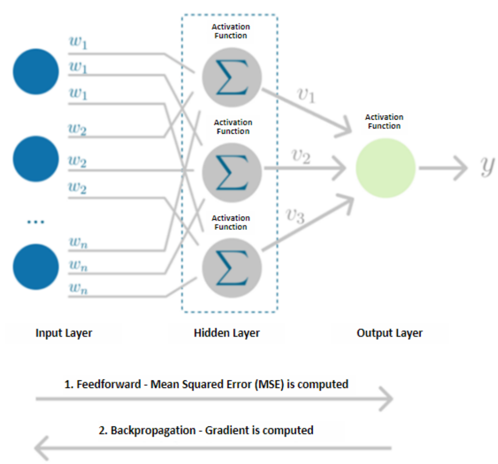

Figure 3 Structure of MLP model.

MLP is utilising the feedforward algorithms where the inputs are combined with the weights in a weighted sum and subjected to the activation function. It has the similar process in the perceptron of a single neuron model but MLP works a bit differently where the linear combination is propagated to the next layer. Each layer is feeding the next one and it will go all the way through the hidden layers to the output layer.

Besides that, MLP also use backpropagation algorithms in the model. It allows to iteratively adjust the weights in the network as to minimise the objective function. The function that combines inputs and weights in a neuron, for instance ReLU, must be differentiable. The method of backpropagation error makes it clear that the errors are sent backwards through the networks (Marsland, 2015). This is where Gradient Descent Method which is one of the optimisation functions is used in MLP to ensure the functions contain bounded derivatives. After the weighted sums are forwarded through all layers in each iteration, the gradient of the Mean Squared Error (MSE) is computed across all input and output pairs. Then, in the propagation stage, the weights of the first hidden layer are updated with the value of the gradient.

C. Error Metrics

The forecast of the model on price trends were evaluated using various error measurement formulas. In this study, three commonly used markers in time series have been utilized to compare the temporal similarities between actual and estimated price on cryptocurrency bitcoin. The performance of the model was measured by Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The evaluation of the models using these error metrics indicates that lower error which means the performance of the model is better than the other.

$$\mathrm{MAE} = \sum_{t=1}^{n} \frac{|y_t - \widehat{y_t}|}{n} \tag{1}$$

$$\mathrm{MAPE} = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - \widehat{y_t}}{y_t} \right| \tag{2}$$

where
$y_t =$ actual closing price
$\widehat{y_t} =$ predicted closing price
$n =$ number of samples for closing price

3.0 Result and Discussion

A. Data Collection

In this study, the dataset comes from the website Cryptocurrency Historical Prices (Rajkumar, 2021). The dataset consisting of 23 different types of cryptocurrencies. Each dataset consists of date of observation, opening, highest, lowest and closing price on the given day, volume of transactions on the given day as well as market capitalisation in USD. The most well-known cryptocurrency that is Bitcoin will be used. The data is also available from original website at Bitcoin.com where the extracted single.csv data was taken.

The price history is available on a daily basis contains record of 2991 days from the 29th of April 2013 to the 6th of July 2021. For this study, the training and testing dataset that will be used is from the 1st of January 2018 to the 31st of December 2020 which is three years and consists of 1096 days. The days onwards are used as actual value for forecasting while comparing with the predicted values by simulation.

Some of the important information on Bitcoin cryptocurrency are that “Date” represents the date of observation, “High” represents highest price on the given day, “Low” represents lowest price on the given day, “Open” represents opening price on the given day, “Close” represents closing price on the given day, “Volume” represents volume of transactions on the given day and lastly is “Marketcap” represents market capitalisation in USD. Moreover, “Date” is an ordinal type of string, “High”, “Low”, “Open” and “Close” are continuous numerical values as well as “Volume” and “Marketcap” are discrete numerical values.

B. Descriptive Analysis

The descriptive analysis of the Bitcoin prices can be seen as in Table 1 below. This analysis is based on the dataset from 2018 to 2020.

Table 1 Descriptive analysis of Bitcoin price variables in USD.

Variables	High	Low	Open	Close
Mean	8882.9598	8474.8128	8683.1758	8696.8507
Median	8522.8827	8143.9335	8322.4102	8321.3812
Standard Error	113.0383	106.2994	109.2809	110.7449
Standard Deviation	3742.2336	3519.1375	3617.8411	3666.3072
Kurtosis	5.3431	5.3307	5.0846	5.5014
Skewness	1.7121	1.6907	1.6589	1.7359
Range	25969.4989	25010.6884	25605.2990	25764.9582

The high price has the highest value among all the variables followed by close, open and low price. All of the variables have not much difference since the standard error which tells about the data that the sample mean would not vary much if were to repeat a study using new samples from within a single population. The standard deviation in the given dataset from 2018 to 2020 would explains that data would disperse about USD 3519 to USD 3742. Other measurement like kurtosis shows the highest in close price in which it tells that it tends to have heavy-tailed or outliers. On the other hand, dataset with low kurtosis tend to have light tails or lack of outliers. The same thing describes by skewness in the dataset is that close price has the highest skewness and it tends to distribute longer on the right side. This means that the outliers are further out towards the right and getting closer to the mean on the left. The descriptive analysis above also illustrated the range by subtraction between maximum price and minimum price, and this means high price is the highest. Consequently, this means that the prices have exhibit high volatility and variability.

Table 2 Descriptive analysis of Bitcoin closing price from 2018 to 2020.

Years	2018	2019	2020
Mean	7572.2989	7395.2463	11116.3781
Median	6906.9199	7824.2315	9713.4944
Standard Error	128.5244	138.1125	225.0708
Standard Deviation	2455.4555	2638.6351	4305.8588
Kurtosis	2.3075	-1.2250	3.4251
Skewness	1.1145	-0.0856	1.8089
Range	14290.2384	9616.7601	24030.9319

It can be observed from Table 2 that the values are slightly decreasing from 2018 to 2019 and then increase again to 2020 through the mean. Other than that, median, standard error and standard deviation are kept increasing from 2018 to 2020. However, kurtosis and skewness, respectively in 2019 shows negative value which means it has no obvious outliers in that year. The outliers tend to skew towards left and coming closer to the mean on the right side. Last but not least, the range value in above table shows that 2019 is the lowest as compared to 2018 and 2020. This can be explained that the difference between maximum and minimum value for 2019 is not high. This means that the dataset has exhibit the characteristics of nonlinearity.

C. Behaviour of Trend

By exploring the trends, it is allowable to get better feel about the data and find useful patterns in it. There are some specific uses of trends in Bitcoin. Investors can use trends in the price and volume of Bitcoin to make informed decisions about buying, selling or holding their Bitcoin. Traders can analyse price movements and identify patterns in the charts using trends, which can help them make more accurate predictions about the future price movements.

The close price is chosen over high, low and open price for the prediction and forecasting in this study. The closing price is often used in Bitcoin forecast models since it reflects the final traded price of the asset for the given period of time. While high and low prices do not represent the final price where Bitcoin is traded, they only provide meaningful information about the range of price movements during a time period. Similarly, the opening price does not gather full range of price movements occur during the time period, but it is essential to understand the initial price of the asset at the beginning stage.

Figure 4 Line chart of Bitcoin close price from 2018 to 2020.

The figure displays closing price movements with the x-axis representing the time period from 2018 to 2020 and the y-axis representing the closing price of Bitcoin in USD. It shows that the closing price fluctuated over the years, with spikes and dips in the Bitcoin closing price. However, despite the fluctuations, the overall trend of the closing price is upward. The line slopes upwards from left to right indicating that the Bitcoin price has been increasing over time. This upward trend suggests that the investors have been buying Bitcoin that causes the closing price driving up.

D. Performance of the Models

This study mainly focuses on two machine learning models to see the performance of the models on the chosen dataset. Clearly, the MAE and MAPE are among the error metrics used for this study to compare the efficiency between LSTM and MLP model.

Table 3 Error metrics values for LSTM and MLP.

Model	MAE	MAPE
LSTM	35616.8297	71.2826
MLP	58411122.7551	126186.3966

From Table 3, the MAE and MAPE values for LSTM and MLP models are 35616.8297 and 7128.26% as well as 5811122.7551 and 12618639.66%, respectively. Generally, a lower MAE and MAPE indicate a better fit of the model to the data, with smaller differences between predicted and actual values. As a result, it is certain that the LSTM model outperforms MLP model in this study with the least error have been calculated for the two-error metrics.

E. Forecasting Visualisation

The forecast for the price of Bitcoin refers to predictions or projections made by individuals or entities about the future value of Bitcoin. It is important to note that Bitcoin, like other cryptocurrencies, is highly volatile and subject to rapid price fluctuations. Therefore, forecast for Bitcoin price can vary significantly depending upon the source and the methodology used. It is worth mentioning that while forecasts can provide insights and guidance, but it is not guaranteed to be accurate.

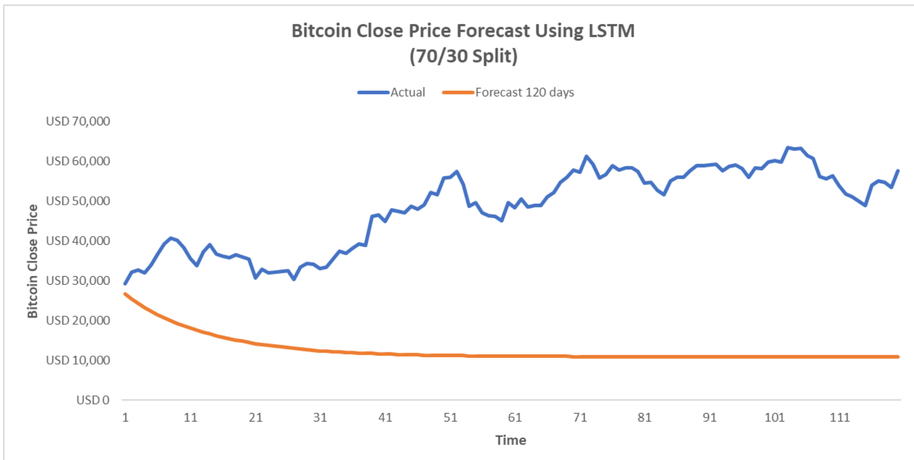

Figure 5 Bitcoin 120 days of future forecast for LSTM.

Figure 5 and 6 demonstrates Bitcoin forecast for LSTM and MLP with 120 days of future forecast. In each of the graphs, the x-axis represents the timeline from 1 to 120 days, while the y-axis represents the price of Bitcoin in USD. The graphs display historical data points from the past year since the 1st of January 2019 to 120 days onwards, showing the price fluctuations of Bitcoin. These data points provide a context for understanding the forecast. The graphs have an obvious trendline can be seen from the historical price data points. The line shows a gradual upward trend, suggesting that the price has been steadily increasing. By looking at the forecasted data, represented by orange colour in both graphs, they indicate the predicted future prices of Bitcoin. The forecast on LSTM model suggests that a continued downward trajectory, projecting lower prices in the coming months. However, MLP shows a continuous upward trajectory which is in contrast to the LSTM model.

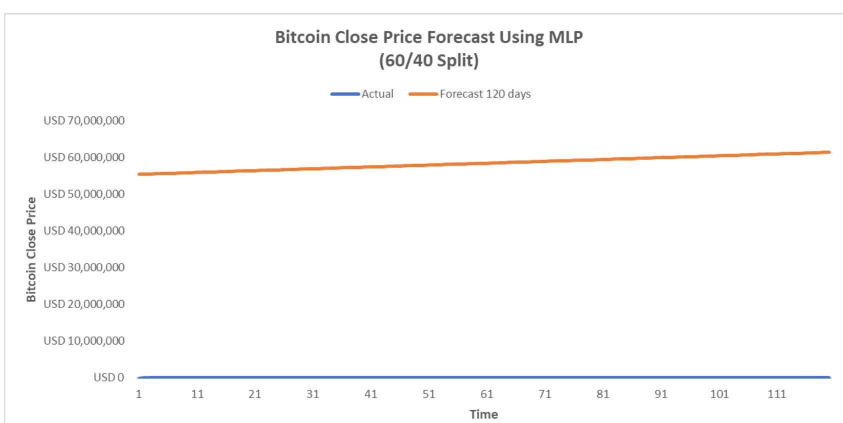

Figure 6 Bitcoin 120 days of future forecast for MLP.

Moreover, an unexpectable situation on the graphs that can be noticed is that the past historical data points for MLP model have un unclear pattern. This indicates that the historical data points are rather smaller than the forecasted data points. The forecast covers the next 120 days. It is important to note that cryptocurrency forecasts are not guaranteed predictions and should be viewed with caution. Various factors can impact the accuracy of the forecast, such as market volatility and unexpected events. All in all, it appears that price prediction of Bitcoin using LSTM model is more accurate although it has a downward trend from the beginning of the forecast. However, it is important to note that cryptocurrency market is highly volatile and influenced by various factors, so the forecast should be used as one of many tools to inform investment decision.

Conclusion

In this research, two mostly known machine learning models; Long-Short Term Memory (LSTM) and Multi-Layer Perceptron (MLP) were used to forecast day-to-day cryptocurrency Bitcoin close price. The day-to-day Bitcoin prices dataset taken from the website Cryptocurrency Historical Prices were analysed and the result were discussed.

Generally, this study is determining whether LSTM or MLP is better for forecasting Bitcoin close price. The error metrics such as Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) will be used in order to choose the most suitable model in forecasting. The smallest error among the models represent that it can give an accurate forecast and it is the best choice to the data for forecasting purposes in this study.

The error metrics results suggest that LSTM is better than MLP. As a result, the findings indicate that both LSTM and MLP models have the potential to effectively predict Bitcoin price. However, the LSTM model consistently outperforms the MLP model in terms of prediction accuracy and forecasting.

In future research, using the same datasets, it can consider to use other more advanced machine learning forecasting techniques for instance, Support Vector Machine (SVM) and Random Forest (RF). At the same time, it is best to compare the result of forecast together with the time series model such as Autoregressive Integrated Moving Average (ARIMA), Exponential Smoothing (ETS) and Vector Autoregression (VAR). By doing that, it enables the researchers to see whether machine learning or time series models that best suits for cryptocurrency Bitcoin price forecasting.

References

[1] Albariqi, R., & Winarko, E. (2020). Prediction of Bitcoin Price Change using Neural Networks. *Proceeding - ICoSTA 2020: 2020 International Conference on Smart Technology and Applications: Empowering Industrial IoT by Implementing Green Technology for Sustainable Development*. https://doi.org/10.1109/ICOSTA48221.2020.1570610936

[2] Amsyar, I., Christopher, E., Dithi, A., Khan, A. N., & Maulana, S. (2020). The Challenge of Cryptocurrency in the Era of the Digital Revolution: A Review of Systematic Literature. *Aptisi Transactions on Technopreneurship (ATT)*, *2*(2), 153–159. https://doi.org/10.34306/ATT.V2I2.96

[3] Bi, Q., Goodman, K. E., Kaminsky, J., & Lessler, J. (2019). Practice of Epidemiology What is Machine Learning? A Primer for the Epidemiologist. *American Journal of Epidemiology*, *188*(12), 2222–2239. https://doi.org/10.1093/aje/kwz189

[4] Guo, H., Zhang, D., Liu, S., Wang, L., & Ding, Y. (2021). Bitcoin price forecasting: A perspective of underlying blockchain transactions. *Decision Support Systems*, *151*, 113650. https://doi.org/10.1016/J.DSS.2021.113650

[5] Hitam, A., & Ismail, A. R. (2018). Comparative Performance of Machine Learning Algorithms for Cryptocurrency Forecasting. *Ind. J. Electr. Eng. Comput. Sci*, *11*(3), 1121–1128. http://doi.org/10.11591/ijeecs.v11.i3.pp1121-1128

[6] Hitam, N. A., Ritahani Ismail, A., & Saeed, F. (2019). An Optimized Support Vector Machine (SVM) based on Particle Swarm Optimization (PSO) for Cryptocurrency Forecasting. *Faisal Saeed c / Procedia Computer Science*, *163*, 0–000. https://doi.org/10.1016/j.procs.2019.12.125

[7] Idrissi, J., & Amine, M. (2016). Multilayer Perceptron: Architecture Optimization and Training. *International Journal of Interactive Multimedia and Artificial Intelligence*, *4*. https://doi.org/10.9781/ijimai.2016.415

[8] Jiang, H. (2021). Cryptocurrency price forecasting based on short-term trend KNN model. *Proceedings of 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology, ICCASIT 2021*, 1165–1169. https://doi.org/10.1109/ICCASIT53235.2021.9633738

[9] Lahmiri, S., & Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons and Fractals*, *118*, 35–40. https://doi.org/10.1016/j.chaos.2018.11.014

[10] Lazzeri, F. (2021). *Machine Learning for Time Series Forecasting with Python*. John Wiley & Sons.

[11] Lindemann, B., Maschler, B., Sahlab, N., & Weyrich, M. (2021). A Survey On Anomaly Detection For Technical Systems Using LSTM Networks. *Computers in Industry*, *131*(103498). https://doi.org/10.1016/J.COMPIND.2021.103498

[12] Marsland, S. (2015). Machine Learning: An Algorithmic Perspective. In T. Graepel & R. Herbrich (Eds.), *CRC Press* (Second Edition).

[13] Mzoughi, H., Ghabri, Y., & Guesmi, K. (2022). Crude oil prices and Crypto-assets: what are the effects of COVID-19 pandemic? *Article in International Journal of Energy Sector Management*. https://doi.org/10.1108/IJESM-10-2021-0016

[14] Nasir, A., Shaukat, K., Khan, K. I., Hameed, I. A., Alam, T. M., & Luo, S. (2020). What is Core and What Future Holds for Blockchain Technologies and Cryptocurrencies: A Bibliometric Analysis. *IEEE Access*, *9*, 989–1004. https://doi.org/10.1109/ACCESS.2020.3046931

[15] Rajkumar, S. (2021). *Cryptocurrency Historical Prices | Kaggle*. https://www.kaggle.com/datasets/sudalairajkumar/cryptocurrencypricehistory

[16] Rathore, R. K., Mishra, D., Mehra, P. S., Pal, O., HASHIM, A. S., Shapi'i, A., Ciano, T., & Shutaywi, M. (2022). Real-world model for bitcoin price prediction. *Information Processing & Management*, *59*(4), 102968. https://doi.org/10.1016/J.IPM.2022.102968

[17] Tredinnick, L. (2019). Cryptocurrencies and the Blockchain. *Business Information Review*, *36*(1), 39–44. https://doi.org/10.1177/0266382119836314

[18] Zhang, R., Chen, Z., Chen, S., Zheng, J., Büyüköztürk, O., & Sun, H. (2019). Deep long short-term memory networks for nonlinear structural seismic response prediction. *Computers & Structures*, *220*, 55–68. https://doi.org/10.1016/J.COMPSTRUC.2019.05.006