Emerging Trend of Transaction and Investment: Bitcoin Price Prediction using Machine Learning

Article *in* International Journal of Advanced Trends in Computer Science and Engineering · October 2020

DOI: 10.30534/ijatcse/2020/1591.42020

2 authors:

Eng Chuen Loh
Universiti Tun Hussein Onn Malaysia

Shuhaida Ismail
Universiti Teknologi Malaysia

ISSN 2278-3091
Volume 9, No.1.4, 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse1591.42020.pdf
https://doi.org/10.30534/ijatcse/2020/1591.42020

Emerging Trend of Transaction and Investment: Bitcoin Price Prediction using Machine Learning

Eng Chuen Loh[1], Shuhaida Ismail[2]
[1]Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia,
mclec95@gmail.com
[2]Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malayia, Johor, Malaysia,
shuhaida@uthm.edu.my

ABSTRACT

Bitcoin is the most popular cryptocurrency with the highest market value. It was said to have potential in changing the way of trading in future. However, Bitcoin price prediction is a hard task and difficult for investors to make a decision. This is caused by nonlinearity property of the Bitcoin price. Hence, a better forecasting method is essential to minimize the risk from inaccuracy decision. The aim of this paper is to first compare three different neural networks which are Feedforward Neural Network (FNN), Nonlinear Autoregressive with Exogenous Input (NARX) Neural Network and Nonlinear Autoregressive (NAR) Neural Network by obtaining the predicted result for each model. The best model is identified by evaluating the performance measurement of each model. After obtaining the best model, it is used to undergo 30 days ahead forecast. The result showed that the performance of NARX out-performed FNN and NAR. It is proven NARX is the suitable neural network to forecast Bitcoin price. The resulting model provides new insights into Bitcoin forecasting using NARX which directly benefits the investors and economists in lowering the risk of making the inaccurate decision when it comes to investing in Bitcoin.

Keywords: Bitcoin Price, Artificial Neural Network, Forecasting, Machine Learning

1. INTRODUCTION

Cryptocurrency has now been the hot topic worldwide. It is said to be the revolution of the currency transaction [1]. Bitcoin was first introduced by a pseudonym creator, Satoshi Nakamoto in late 2008 on the internet [2]. As Bitcoin price did not rely on any government bodies or agencies, therefore, Bitcoin has the feature of the decentralized system. This feature has influenced the demand for Bitcoin in the market. Moreover, this currency is built based on advanced mathematics and computer engineering, therefore, it is difficult to counterfeit the currency and the users are safe from cyber-attack which might cause losses or leakage of identity [3].

Bitcoin plays an important role in developing a brand-new decentralized currency transactions system while influencing the stock market direction. This is one of the factor that influence the users' decision on using this advanced transaction system [4]. However, Bitcoin prices are frequently fluctuating because of its virtual characteristics. Therefore, it is essential to develop a reliable model in order to enable investors and economists to forecast the Bitcoin price. Bitcoin price has complex and nonlinearity characteristics [5, 6]. Thus, classical forecasting methods such as Holt-Winters is not suitable to forecast Bitcoin price data because the methods require the assumption of linearity [7]. Therefore, a method which able to handle nonlinearity and complexity of the data are needed to forecast Bitcoin price.

This paper presents "Emerging Trend of Transaction and Investment: Bitcoin Price Prediction using Machine Learning". This study aims to forecast the future values of Bitcoin price using best model among three of the neural network which are Feedforward Neural Network (FNN), Nonlinear Autoregressive with Exogenous Input (NARX) Neural Network and Nonlinear Autoregressive (NAR) Neural Network. Indera *et al.* applied NARX in predicting Bitcoin price from their study whereas Almeida *et al.* suggested that FNN can be used in Bitcoin price prediction [1, 8]. However, there are no available literature on applying NAR in predicting Bitcoin price. Thus, from this study, these models can be compared and determine the best model. These models will then be evaluated using five different performance measurements which are Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), Mean Absolute Scaled Error (MASE) and Mean Forecast Error (MFE). The best model is chosen by comparing the performance measurement for each model for undergoing 30 days ahead forecast.

2. METHODOLOGY

There are three neural networks being used in this study which are FNN, NARX and NAR. The best model is determined by comparing the performance measurement of each neural network. Next, the best model is used to perform 30 days ahead forecast.

2.1 Data

Different websites offer different selling price for Bitcoin. For this research, the Bitcoin price data were collected from

Blockchain, which is the master ledger that records the original Bitcoin price. Blockchain is one of the big data that has to be retrieved cautiously. This because big data contain larger sizes compared to regular utilized programming [9].There were 1035 observations of daily Bitcoin price data starting from 1st July 2017 until 30thApril 2020 were used in this study. Aside from Bitcoin price data, there were several others daily data variables collected from Blockchain which includes hash rate, average block size, transaction cost, number of transactions, miner revenue and number of transactions per block. These are some of the influence factor mentioned by Kristoufek [10].

2.2Data Pre-processing

Different value range in the variables will directly influence the tendency and accuracy for the models especially for neural network [11]. Therefore, normalization method is applied in the analysis. The data used in the analysis are transformed using Min-Max normalization method which transforms the data into a defined range of 0 to 1 [12]. The equation of the Min-Max normalization method is shown in Equation (1).

$$x_i = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

where $x_i$is the normalized x value at time i; x is the actual x value; x_{min} is the minimum actual value; x_{max} is the maximum actual value.

2.3Feedforward Neural Network

Feedforward Neural Network is one of the basic forms of Artificial Neural Network which passes the information from the input layer directly to the output layer after undergoing activation function [13]. However, in order to handle nonlinear data, a hidden layer is needed to be inserted within the input and output layer [2].

The weight for each of the interconnection constantly changes based on the predetermined training algorithms. A basic three layers backpropagation FNN with α input nodes and β hidden nodes with one output node will produce the predicted output values based on the following equation:

$$\hat{y}_t = f\left[\sum_{j=1}^{\beta} w_j g\left(\sum_{i=1}^{\alpha} w_i x_i + w_i\theta\right) + w_j\theta\right] \tag{2}$$

$$w_{i+1} = w_i + \eta \hat{y}_i\left(y_i - \hat{y}_i\right) \tag{3}$$

where y is the target value, $\hat{y}_t$ is the predicted value, x is the input value, w_i is the connection weight between input and hidden layer nodes, w_j is the connection weight between hidden and output layer nodes, θ is the bias constant, η is the learning rate, $f(x)$ and $g(x)$ are the activation function.

In this research, Levenberg-Marquardt backpropagation algorithm is used as the training algorithm for all three of the neural network. Levenberg-Marquardt algorithm is modified from Gauss Newton method [14].

2.4 Nonlinear Autoregressive with Exogenous Input (NARX)

NARX is one of the recurrent neural network which requires information to loop back to be processed by the hidden layer. Unlike FNN, NARX considers that the target value as one of the input nodes. This model relates the current target value to either past target value or current and past target value of exogenous series. The equation of NARX model can be written as follows:

$$\hat{y}_t = f\left(y_{t-1}, y_{t-2}, ..., y_{t-n_y}, u_{t-1}, u_{t-2}, ..., u_{t-n_u}\right) \tag{4}$$

where $\hat{y}_t$ is the predicted value, y_{t-1} is the previous actual value and u_{t-1} is the previous external variable that influences y.

NARX network can be differentiated in parallel architecture and series-parallel architecture [15]. Parallel architecture loops back the predicted output as an input layer while series-parallel architecture do not loop back the predicted output which similar with FNN with target value as one of the input nodes.

2.5Nonlinear Autoregressive (NAR)

NAR neural network is another form of recurrent neural networks. The different between NAR and NARX is that NAR does not have exogenous input. NAR only loops back the information to the hidden layer. It is commonly used in forecasting nonlinear time series without taking account any influence variable [16]. The equation of NAR model can be written as follows:

$$\hat{y}_t = f\left(y_{t-1} + y_{t-2} + ... + y_{t-d}\right) \tag{5}$$

where $\hat{y}_t$ is the predicted value and d is the lagged value.

2.6 Training Algorithm Parameter

The parameter shown in Table 1 were used for FNN, NARX and NAR respectively.

Table 1: Training Parameters of each Neural Network

Parameter	FNN	NARX	NAR
Training algorithm	Levenberg Marquardt	Levenberg Marquardt	Levenberg Marquardt
Transfer function	log-sigmoid + linear	log-sigmoid + linear	log-sigmoid + linear
Maximum fail	500	500	500
Maximum epochs	10000	10000	10000
Learning rate, α	0.01	0.01	0.01
Performance goal	0	0	0
Minimum gradient	1.00×10^{-6}	1.00×10^{-6}	1.00×10^{-6}
μ	1.00×10^{-3}	1.00×10^{-3}	1.00×10^{-3}
Maximum μ	1.00×10^{10}	1.00×10^{10}	1.00×10^{10}

All neural networks were set with Levenberg Marquardt training algorithm and the transfer function of log-sigmoid function from input layer to hidden layer, linear function from hidden layer to output layer. Furthermore, the neural networks also been set with the maximum fail of 500 times when validating stage, maximum epochs of 10000 iterations, learning rate of 0.01 unit, performance goal of zero, minimum gradient of 1.00×10^{-6} unit, μ of 1.00×10^{-3} and the maximum μ of 1.00×10^{10}.

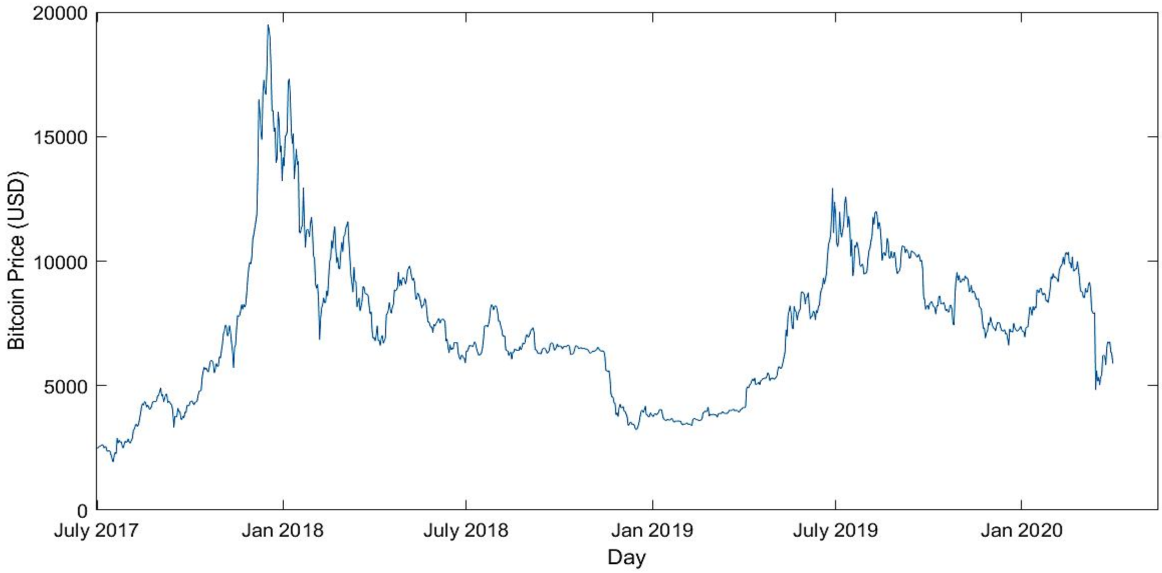

Figure 2: Time series plot for Bitcoin price

2.7 Forecast Accuracy

In this research, five types of forecast accuracy measurements are applied to evaluate the accuracy of the predicted output for the forecasting models. The measurements that used in this study are MAE, MFE, RMSE, MAPE and MASE, whereby, the best model is selected based on the smallest values for all measurements. The equations for each of the forecast accuracy are shown as follows:

$$MAE = n^{-1}\left[\sum_{i=1}^{n}\left|y_t - \hat{y}\right|\right] \quad (6)$$

$$MFE = n^{-1}\left[\sum_{i=1}^{n}\left(y_t - \hat{y}\right)\right] \quad (7)$$

$$RMSE = n^{-2}\left[\sum_{i=1}^{n}\left(y_t - \hat{y}\right)^2\right]^{-0.5} \quad (8)$$

$$MAPE = 100n^{-1}\left[\sum_{i=1}^{n}\left|\left(y_t - \hat{y}\right)\left(y_t\right)^{-1}\right|\right] \quad (9)$$

$$MASE = \sum_{i=1}^{n}\left|y_t - \hat{y}\right|\left[n\left(n-1\right)^{-1}\sum_{i=1}^{n}\left|y_t - \hat{y}\right|\right]^{-1} \quad (10)$$

where y_t is the actual values at time t; $y_{t-1}$is the actual values at time t-1; $\hat{y}$ is the predicted values at time t; $n$is the number of observations

3. RESULT AND DISCUSSIONS

Based on visual inspection of Figure 2, the series has shown to be nonlinear and non-stationary. However, proper statistical tests needed to carry out to prove the findings. Therefore, Augmented Dickey-Fuller (ADF) and Anderson Darling (AD)
statistical tests were performed to prove the stationarity and linearity properties in the data. After obtaining the result, it can be concluded that the data do not follow the normal distribution, hence nonlinearity does exist in the data.

Once the nonstationarity and nonlinear properties proved in the dataset, the dataset is ready to be used in prediction by using neural network method. In this study, three neural networks which are FNN, NARX and NAR will be used to obtain the predicted value. The predicted values were computed using forecast accuracy which are MAE, MAPE and RMSE, MASE and MFE. The forecast accuracy of the models is then tabulated in Table 2.

Table 2: Performance Measurement of each Neural Network

Parameter	FNN	NARX	NAR
MAE	143.586	23.415*	30.224
MAPE	53.421	14.433*	23.478
RMSE	341.974	141.941*	192.359
MASE	2.531	1.149*	1.524
MFE	1.211	-0.89*	-4.212

* Indicates the best result among all of the neural network.

From Table 2, the best performance measurements are from NARX. This model obtained MAE value of 23.415, MAPE value of 14.433, RMSE value of 141.941 and MASE value of 1.149. The lowest MAE and RMSE implies that NARX produced smaller error compared to FNN and NAR neural network. MAPE value of NARX falls in the category of good forecasting accuracy whereas MAPE value of FNN and shows that the models are over-forecasted. In overall, NARX neural network stands out among the three neural networks.

3.1 Multi-step Ahead Forecasting

multi-step ahead forecast. The forecasted Bitcoin Prices from 1st April 2020 to 30th April 2020 using NARX with direct strategy method are tabulated in Table 3.

Table 3:30 Days Ahead Forecast Values

Date	Price (USD)	Date	Price (USD)
1/4/2020	6038.76	17/4/2020	7439.71
2/4/2020	6011.60	18/4/2020	7468.65
3/4/2020	6177.53	19/4/2020	7560.85
4/4/2020	6198.23	20/4/2020	7446.12
5/4/2020	6203.66	21/4/2020	7344.79
6/4/2020	6336.83	22/4/2020	7392.94
7/4/2020	6384.44	23/4/2020	7521.42
8/4/2020	6437.48	24/4/2020	7562.99
9/4/2020	6522.37	25/4/2020	7521.40
10/4/2020	6624.66	26/4/2020	7624.66
11/4/2020	6678.17	27/4/2020	7513.78
12/4/2020	6720.42	28/4/2020	7623.60
13/4/2020	6719.25	29/4/2020	7793.89
14/4/2020	7046.84	30/4/2020	7849.42

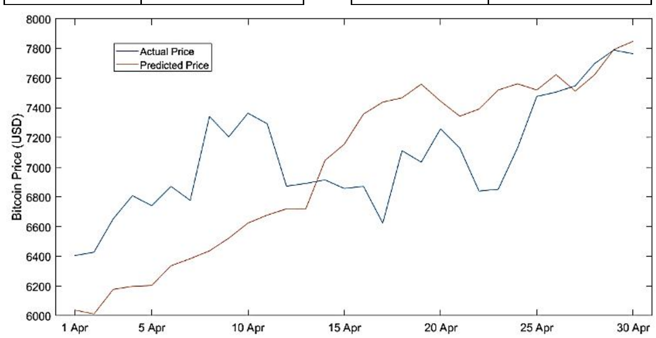

Figure 3: Validation of Forecasted Values

Based on observation, the neural network able to predict the Bitcoin price accurately from 1st April 2020 to 30th April 2020. The predicted price have the increment trend as same as the actual price. In the period of 27th to 30th April, the predicted prices are almost the same as the actual price. Based on this observation, it can be concluded that NARX model able to perform a 30 days ahead forecast for Bitcoin price.

4. CONCLUSION

This paper presents "Emerging Trend of Transaction and Investment: Bitcoin Price Prediction using Machine Learning".

NAR fall in the category inaccurate forecasting and reasonable forecasting respectively. MASE of NARX and NAR approached 1 which implies that the models slightly out-performed naïve model. MFE of FNN indicates that the model is slightly under-forecasted whereas MFE of NARX and NAR

30 days ahead forecast is carried out using direct strategy method after selecting the best model. The analysis of the comparison of performance measurement concludes that NARX has the best accuracy. Thus, it is selected to carry out

The ADF and AD tests indicate Bitcoin prices have the characteristics of nonlinear and nonstationary. Therefore, classical forecasting methods are not suitable to forecast Bitcoin price. NARX has the lowest error compared to FNN and NAR in MAE, MAPE and RMSE and MASE, with the values of 23.415, 14.433%, 141.941 and 1.149respectively. Thus, NARX is chosen as the best model to undergo 30 days ahead forecast. From the result of 30 days ahead forecast, there is an increasing trend from 6038.76 USD in 1stApril2020 to 7849.42 USD in 30thApril 2020.

By having the knowledge of how the Bitcoin price will be, investors and government authorities will have better decision making in investment on Bitcoin. Besides, understand the trend of Bitcoin price will enable the trading system to be more secure in the future. Thus, it would bring forwards the new era of transactions.

ACKNOWLEDGEMENT

This research was supported by the Universiti Tun Hussein Onn Malaysia and Ministry of Education Malaysia under the Fundamental Research Grant Scheme (FRGS) Vot K082.

REFERENCES

1. J. Almeida, S. Tata, A. Moser, and V. Smit.**Bitcoin prediction using ANN**,*Neural Network*, pp. 2-12, June 2015.
2. M. Negnevitsky.**Artificial Intelligence**, 2nd ed., England: Addison Wesley, 2003.
3. A. Gervais, G. O. Karame, V. Capkun, and S. Capkun.**Is Bitcoin a decentralized currency?**,*IEEE Security, and Privary*, Vol. 12, No. 3, pp. 54-60, June 2014. https://doi.org/10.1109/MSP.2014.49
4. B. Manoj, K. V. Sasikanth, M. V. Subbarao, and V. J. Prakash. **Analysis of data science with the use of big data.***International Journal of Advanced Trends in Computer Science and Engineering*, Vol. 7, No. 6, pp. 87-90, December 2018. https://doi.org/10.30534/ijatcse/2018/02762018
5. M. Polasik, A. Piotrowska, and T. P. Wisniewski.**Price fluctuations, and the use of Bitcoin: An empirical inquiry**,*International Journal of Electronic Commerce*, Vol. 20, No. 1, pp. 9-49, August 2015.

6. Q. Zhou, Q. Zhang, and Q. I. Zhang.**Agent-based simulation research on Bitcoin price fluctuation**,*2nd International Conference on Artificial Intelligence, and Engineering Application (AIEA 2017)*, Guilin, China,pp. 21-33, October 2017.
7. S. Mcnally.**Predicting the Price of Bitcoin using Machine Learning**,*2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP)*, Cambridge, United Kingdom, pp. 339-343,June 2018.
8. N. I. Indera, I. M. Yassin, A. Zabidi, and Z. I. Rizman.**Non-Linear Autoregressive with Exogeneous Input (NARX) Bitcoin price prediction model using PSO-optimized parameters, and moving average technical indicators**,*Journal of Fundamental, and Applied Science*, Vol. 9, No. 3s, pp. 791-808, January 2018.
https://doi.org/10.4314/jfas.v9i3s.61
9. Z. Rohimah, A. D. P. Elga, and G. Wang. **Factors that determine customer decisions in the use of E-payment in the culinary business**. *International Journal of Advanced Trends in Computer Science and Emerging*, Vol. 8, No. 6, pp. 3049-3052, December 2019.
https://doi.org/10.30534/ijatcse/2019/62862019
10. L. Kristoufek.**What are the main drivers of Bitcoin price? Evidence from wavelet coherence analysis**,*PloS ONE*, Vol. 10, No. 4, pp. 3-14, April 2015.
11. G. P. Zhang.**Neural Network for Time-Series Forecasting**, eds G. Rozenberg, T. Bäck, and J. N. Kok, *Handbook of Natural Computing*, Berlin, Heidelberg: Springer, pp. 461-477, 2012.
12. J. S. Chou, and J. P. P. Thedja.**Metaheuristic optimization within machine learning-based classification system for early warning related to geotechnical problems**, *Automation in Construction*, Vol. 68, pp. 65-80, March 2016.
13. I. Chibuisi, and O. Chukwuma.**Short term load forecasting using artificial neural network (ANN) for Umuahia, Nigeria**, *International Journal of Emerging Technologies in Engineering Research (IJETER)*, Vol. 3, No. 2, pp. 50-56, November 2015.
14. A. Reynaldi, S. Lukas, and H. Margaretha.**Backpropagation, and levenberg-marquardt algorithm for training finite element neural network**,*AMSS 6th European Modelling Symposium (EMS 2012)*, Malta: UKSim, pp. 89-94, January 2013.
15. H. Demuth, M. Beale, M. Hagan.**Neural Network Toolbox^{TM} User's Guide**, Mathwork, 2017.
16. Y. Kashyap, A. Bansal, and A. K. Sao.**Solar radiation forecasting with multiple parameters neural networks**,*Renewable, and Sustainable Energy Reviews*, Vol. 49, pp. 825-835, April 2015.
https://doi.org/10.1016/j.rser.2015.04.077