ISSN 2278-3091
Volume 9, No.1.1, 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse6291.12020.pdf
https://doi.org/10.30534/ijatcse/2020/6291.12020

Bitcoin Price Forecasting using Different Artificial Neural Network and Training Algorithm

Loh Eng Chuen[1], Shuhaida Ismail[2], Aida Mustapha[3]
[1, 2] Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Malaysia
shuhaida@uthm.edu.my
[3] Faculty of Computer Science and Information technology, Universiti Tun Hussein Onn Malaysia, Malaysia.

ABSTRACT

Bitcoin gain popularity day by day. Economists anticipate that Bitcoin might one day replace current transaction method. However, Bitcoin price is hard and difficult for investors to predict and make decision when investing. One of the reason is that Bitcoin price has the nonlinearity property as the price of Bitcoin fluctuated a lot. Thus, a better forecasting method is needed to minimize the risk from inaccuracy decision. The aim of this paper is to find the best model to predict Bitcoin price using two different Neural Network which are Feedforward Neural Network (FNN) and Nonlinear Autoregressive (NAR) Neural Network. The NN models are tested with two different training algorithm which are Levenberg-Marquardt (LM) and Scaled Conjugate Gradient (SCG) backpropagation training algorithm. The best model is identified by evaluating the performance measurements of each model. The result showed that the performance of NAR with LM training algorithm out-performed other models. It is proven NAR with LM training algorithm is the suitable neural network to predict Bitcoin price. The resulting model provides new insights into Bitcoin forecasting using NAR which directly benefits the investors and economists in lowering the risk of making wrong decision when it comes to invest in Bitcoin.

Key words : Bitcoin Price; Artificial Neural Network; Forecasting

1. INTRODUCTION

Bitcoin gain popularity day by day. Economists anticipate that Bitcoin might one day replace current transaction method. However, Bitcoin price is hard and difficult for investors to predict and make decision when investing. One of the reason is that Bitcoin price has the nonlinearity property as the price of Bitcoin fluctuated a lot. Thus, a better forecasting method is needed to minimize the risk from inaccuracy decision. The aim of this paper is to find the best model to predict Bitcoin price using two different Neural Network which are Feedforward Neural Network (FNN) and Nonlinear Autoregressive (NAR) Neural Network. The NN models are tested with two different training algorithm which are Levenberg-Marquardt (LM) and Scaled Conjugate Gradient (SCG) backpropagation training algorithm. The best model is identified by evaluating the performance measurements of each model. The result showed that the performance of NAR with LM training algorithm out-performed other models. It is proven NAR with LM training algorithm is the suitable neural network to predict Bitcoin price. The resulting model provides new insights into Bitcoin forecasting using NAR which directly benefits the investors and economists in lowering the risk of making wrong decision when it comes to invest in Bitcoin**.**

2. METHODOLOGY

There are two NN models used in this study known as Feedforward Neural Network (FNN) and Nonlinear Autoregressive (NAR) Neural Network. The NN models are trained with two different training algorithms which are Levenberg-Marquardt (LM) and Scaled Conjugate Gradient (SCG) training algorithm. The best model is determined by comparing the performance measurement of each model.

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. There are 2435 observations of daily Bitcoin price data starting from 1st January 2012 until 31st August 2018 used in this study. Aside from Bitcoin price data, there are several others daily data variables collected from Blockchain which includes hash rate, average block size, transaction cost, numbers of transactions, miner revenue and number of transaction per block. These are some of the influence factor mentioned by Kristoufek [6].

2.2. Feedforward 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 [7]. However, in order to handle nonlinear data, a hidden layer is needed to be inserted within the input and output layer [8]. Fig. 1 illustrates the model of the FNN.

The weight for each of the interconnection constantly changes based on the predetermined training algorithms.
In this study, a basic three layers backpropagation feedforward neural network with α input nodes, β hidden nodes and one output node are used. The predicted output values are obtained from the Equation (1).

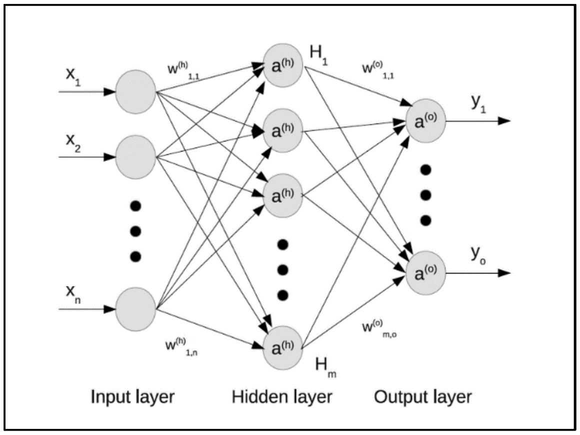

Figure 1: Architecture of FNN [9]

$$y_t = f\left[\sum_{j=1}^{\beta} w_j g\left(\sum_{i=1}^{\alpha} w_i x_i + w_i\theta\right) + w_j\theta\right] \qquad (1)$$

where y_t is the output value at actual time t; x_t is the input value at actual time t; $w_i$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; $f(x)$ and $g(x)$ are the activation functions; $i = j = 1, 2, 3, \dots, n$.

2.3. Nonlinear Autoregressive (NAR)

NAR neural network is another form of recurrent neural network. 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 [10]. The equation of the NAR model is as follows:

$$\hat{y}(t) = f\left[y(t-1) + y(t-2) + \dots + y(t-d)\right] \qquad (2)$$

where $\hat{y}(t)$ is the next value of predicted output value.

The architecture of the NAR neural network is shown in Fig. 2.

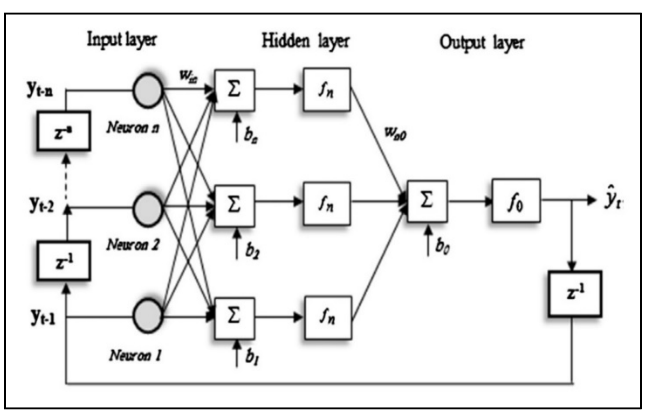

Figure 2: Architecture of NAR Model [11]

2.4. Parameters Setting for NN Models

The parameter shown in Table 1 were used for FNN and NAR. All model are set with the same parameter in order to obtain a fair result.

All neural networks are 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 have also been set with maximum fail of 500 times when validating stage, maximum epochs of 10000 iteration, learning rate of 0.01 unit, performance goal of 0, minimum gradient of 1.00 x 10^{-6} unit, μ of 1.00 x 10^{-3} and the maximum μ of 1.00 x 10^{10}.

Table 1: Training Parameters of each model

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

2.5. Data Pre-processing

Different value range in the variables will directly influence the tendency and accuracy for the models especially for NN [12]. Therefore, normalization method is applied in the analysis. The data used in the analysis are normalized using Min-Max normalization method which transforms the data into a defined range of 0 to 1 [13]. The equation of the Min-Max normalization method is shown in Equation (3).

$$x_i = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (3)$$

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

2.6. Forecast Accuracy

In this research, five forecasting accuracy measurements are applied to evaluate the accuracy and performance of the predicted output for all models. The measurements that used in this study are Mean Absolute Error (MAE), Mean Forecast Error (MFE), Root Mean Square Error (RMSE), Mean Absolute Error (MAPE) and Mean Absolute Scaled Error (MASE). The criterion of the best model is based on the smallest obtained values for all measurements. The equation for each of the forecast accuracy are shown as follows:

$$MAE = \frac{\sum_{i=1}^{n} \left| y_t - \hat{y}_t \right|}{n} \quad (4)$$

$$MFE = \frac{\sum_{i=1}^{n} \left(y_t - \hat{y}_t \right)}{n} \quad (5)$$

$$RMSE = \frac{1}{n^2} \sqrt{\sum_{i=1}^{n} (y_t - \hat{y}_t)^2} \quad (6)$$

$$MAPE = \frac{100}{n} \sum_{t=1}^{n} \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (7)$$

$$MASE = \frac{\sum_{i=1}^{n} \left| y_t - \hat{y}_t \right|}{\frac{n}{n-1} \sum_{i=2}^{n} \left| y_t - y_{t-1} \right|} \quad (8)$$

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

3. RESULT AND DISCUSSIONS

Fig. 3.1 shows the time series plot Bitcoin price. Based on visual inspection 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. ADF Test showed a logical result of value 0 with the *p*-value of 0.8401 ($p > 0.05$). It indicates that this test fails to reject the null hypothesis of a unit root is equal to 0, suggesting that the data is not stationary. Meanwhile, the obtained output for AD test shows a logical result of value 1 with *p*-value approximate to 0.0005 which indicates that the null hypothesis is rejected. Thus, it can be concluded that the data does not follow the normal distribution, hence nonlinearity does exist in the data. Pieces of evidence indicate that Bitcoin time series data is proven to be non-stationary and nonlinear, hence, fulfill the assumption of neural network.

Once the non-stationarity and nonlinear properties proved in the dataset, the dataset is ready to be used in prediction by using neural network method.

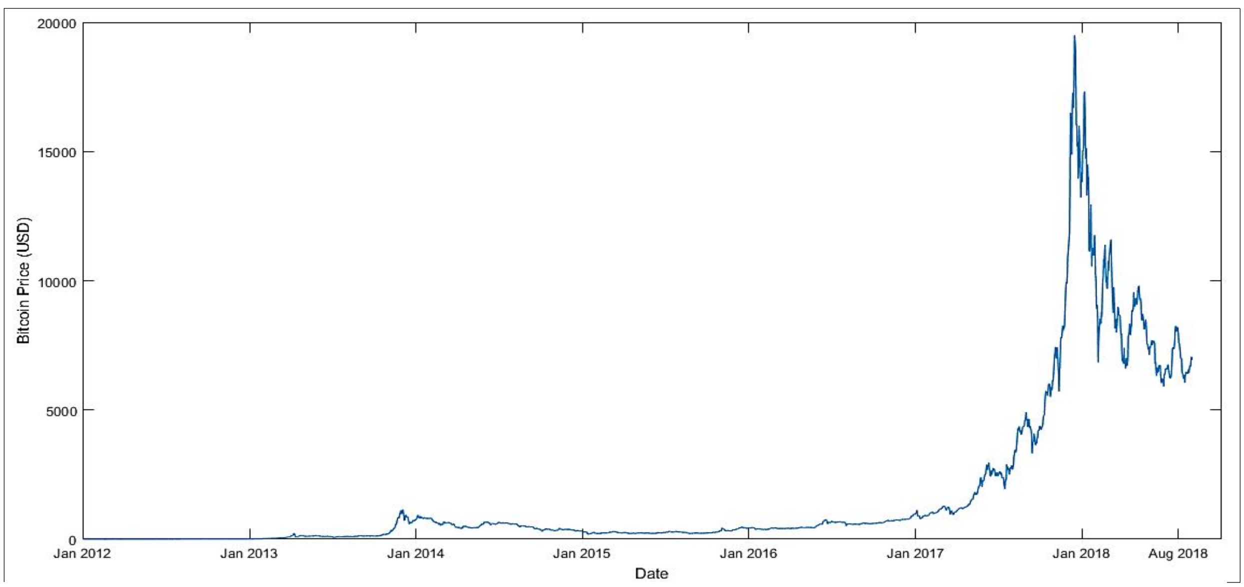

Figure 3: Time Series Plot of Bitcoin Price Daily Data

The predicted value versus actual value of each model is illustrated in Fig. 4 to Fig. 7. Visual inspection indicates that the fluctuation of the predicted values for both FNN models are huge and the fluctuation for FNN with SCG training algorithm is the worst. However, the predicted values for both NAR model are very close to the actual value.

The predicted values were computed using forecast accuracy which are MAE, MAPE and RMSE, MASE and MFE. The forecast accuracies of the models are then tabulated in Table 2.

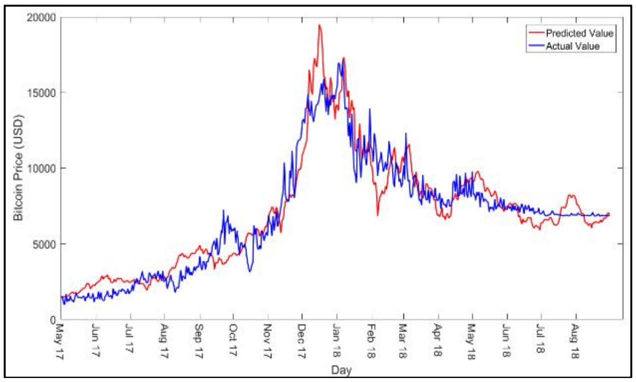

Figure 4: Predicted vs. Actual value (FNN-LM)

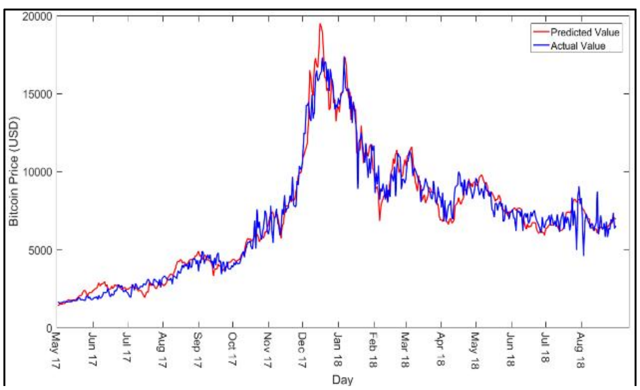

Figure 5: Predicted vs. Actual value (FNN-SCG)

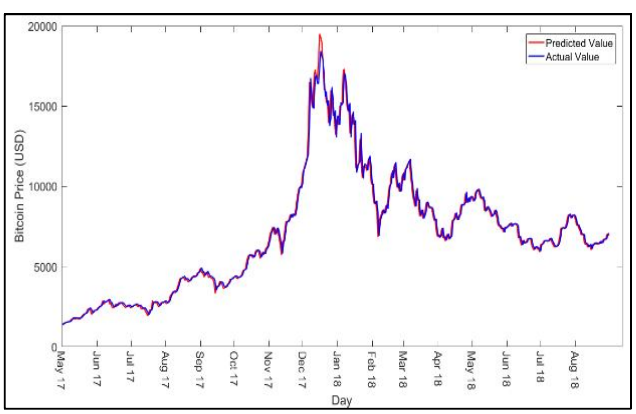

Figure 6: Predicted vs. Actual value (NAR-LM)

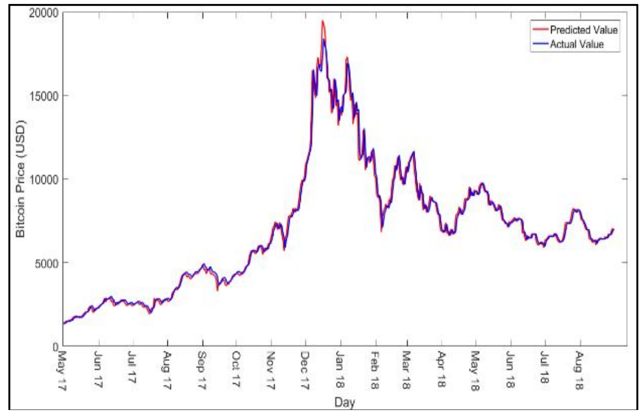

Figure 7: Predicted vs. Actual value (NAR-SCG)

Table 2: Performance Measurement of each Neural Network

Parameter	FNN		NAR	
	LM	**SCG**	**LM**	**SCG**
MAE	173.11	267.55	***65.87**	69.15
MAPE	73.10	78.60	***39.71**	46.81
RMSE	372.30	573.23	***202.34**	205.65
MASE	2.87	4.43	***1.09**	1.14
MFE	***0.35**	8.06	-6.53	0.63

Asterisk sign '*' represent the best result among all other models.

From Table 2, the best performance measurements are obtained from NAR model with the values of 65.87, 39.71, 202.34 and 1.09 for MAE, RMSE, MAPE and MASE, respectively. The lowest MAE and RMSE implies that NAR produced smaller error compared to FNN. MAPE value of NAR with LM and SCG training algorithms falls in the category of reasonable forecasting accuracy whereas MAPE value of FNN fall in the category inaccurate forecasting accuracy. Meanwhile, MASE of NAR for both training algorithms approached 1 which implies that the models slightly out-performed naïve model. The analysis showed MFE for FNN with LM training algorithm and NAR with SCG training algorithm indicate that the models are slightly under-forecasted whereas MFE for FNN with SCG training algorithm and NAR with LM training algorithm shows that the models are over-forecasted.

4. CONCLUSION

The ADF and AD tests indicate Bitcoin price has the characteristics of nonlinear and non-stationary. Therefore, classical forecasting methods are not suitable to forecast Bitcoin price as the classical forecasting methods require to fulfill the linearity and stationary assumption. NAR has the lowest error compared to FNN in term of MAE, MAPE and RMSE and MASE, with the values of 65.878, 39.708%, 202.337 and 1.090 respectively. Thus NAR is the best model when dealing with Bitcoin price prediction.

However, there are some limitations in forecasting the Bitcoin price data using ANN. This is because that it is hard to explain how ANN produces a solution. Besides, the network structure of ANN contributes significant effect on the result. Thus, determining the suitable network structure is essential through many times of trial and error, which in result of consuming large amount of time. Furthermore, ANN is limited to numerical based information, thus it cannot process the information such as news of Bitcoin, global comment and other non-numerical information.

Besides, aside from internal factors of the Bitcoin system, external factors such as global trend, Bitcoin news, latest events and more are also might influence the price of the Bitcoin price [6]. Moreover, another limitation is that the Bitcoin price data has to be up to date in order to achieve better accuracy in prediction its price.
It is recommended that further research should be taken into account of the optimal network structure in order to achieve better accuracy.

Besides, a hybrid model of quantitative forecasting approaches with qualitative forecasting approaches is recommended so that all factors can be included in the forecasting.

ACKNOWLEDGEMENT

This research is funded by Universiti Tun Hussein Onn Malaysia grant scheme under vote number H073.

REFERENCES

1. Gervais A, Karame GO, Capkun V & Capkun S (2014), Is Bitcoin a Decentralized Currency? IEEE Security and Privacy, Vol. 12, No. 3, p. 54 – 60. https://doi.org/10.1109/MSP.2014.49
2. Polasik M, Piotrowska A & Wisniewski TP (2015), ***Price Fluctuations and the Use of Bitcoin*** : ***An Empirical Inquiry.*** International Journal of Electronic Commerce, Vol. 20, No. 1, pp. 9 – 49.
3. Nakamoto S. (2008). Bitcoin: ***A Peer-to-Peer Electronic Cash System, ava***ilable online: https://bitcoin.org, last visit: 16.9.2018
4. Zyskind G, Nathan O & Pentland AS (2015), Decentralizing Privacy: ***Using Blockchain to Protect Personal Data***. Proceedings **-** 2015 IEEE Security and Privacy Workshops, SPW 2015, pp. 180 – 184.
5. Almeida J, Tata S, Moser A & Smit V (2015***), Bitcoin Prediction using*** ANN, Neural Network.
6. Kristoufek L (2015), ***What are the Main Drivers of Bitcoin Price?*** Evidence from Wavelet Coherence Analysis, PloS ONE, Vol. 10, No. 4, p. 1 – 15.
7. Rosenblatt F (1961), ***Principle of Neurodynamics. Perceptrons and the Theory of Brain Mechanism, Buffalo,*** New York: Cornell Aeroneurical Lab Inc. https://doi.org/10.21236/AD0256582
8. Negnevitsky M (2003), ***Artificial Intelligence***, 2nd ed., England: Addison Wesley.
9. Parvizi, M, Paegert M & Stassun KG (2014), ***The EB Factory Project. II. Validation with the Kepler Field in Preparation for K2 and TESS***. Astronomical Journal, Vol. 148, No. 6, pp. 125.
10. Kashyap Y, Bansal A & Sao AK (2015***), Solar Radiation Forecasting with Multiple Parameters Neural Networks, Renewable and Sustainable Energy Reviews***, Vol. 9, pp. 825 – 835.
11. Benmouiza K & Cheknane A (2016), ***Small-scale Solar Radiation Forecasting Using ARMA and Nonlinear Autoregressive Neural Network Models***. Theoretical and Applied Climatology, Vol. 124, No. 3–4, pp. 945 – 958.
12. Zhang GP (2012), ***Neural Network for Time-Series Forecasting, eds Rozenberg G, Bäck T & Kok JN, in Handbook of Natural Computing***, Berlin, Heidelberg: Springer, pp. 461 – 477
13. Chou JS & Thedja JPP (2016), ***Metaheuristic optimization within machine learning- based classification system for early warning related to geotechnical problems,*** Automation in Construction, Vol. 68, pp. 65 – 80. https://doi.org/10.1016/j.autcon.2016.03.015