Identifying explanatory factors of bitcoin price with artificial neural networks

Prosper Lamothe Fernández

Full Professor of Financial Economic and Accounting. University Autonoma of Madrid, Phone: +34 914974673, Email: prosper.lamothe@uam

Prosper Lamothe López

Rho Finanzas Partner, Phone: +34 917527676, Email: pll@rhofinanzas.com

Manuel A. Fernández Gámez (*)

Associate Professor of Financial Economic and Accounting. University of Málaga, Phone: +34 952131234, Email: mangel@uma.es

Sergio M. Fernández Miguélez

Ph. D. Mechanical Engineering and Energy Efficiency. University of Málaga, Phone: +34 951 952 350, Email: sergiofernandezmiguelez@uma.es

(*) Corresponding author

Identifying explanatory factors of bitcoin price with artificial neural networks

Abstract

This study aims to develop a new model that allows determining with high precision the factors that explain the price of bitcoin. To do this, an extensive database of variables related to bitcoin and artificial neural network techniques has been used. The results obtained have made it possible to identify that aspects related to the number of forum posts, the volume of transactions on the blockchain, and the hash rate provide an excellent strategy for predicting the price of bitcoin.

Keywords: Bitcoin, artificial neural networks, blockchain, cryptocurrency

1. Introduction

Bitcoin and the rest of the cryptocurrencies are setting an increasingly strong trend in all financial markets, and the financial literature is devoting more attention to their study (Nasir et al., 2019; Grinberg, 2012). Recent research has identified different factors that affect the price of cryptocurrencies, and especially bitcoin (Al-Khazali, Bouri & Roubaud, 2018; Civitarese, 2018). These factors are related to demand and supply, attractiveness on social networks, and macroeconomic and financial variables. However, the importance of these factors varies from study to study and the existing results demand measurement models with high levels of precision (Sin & Wang, 2017). To shed light on the importance of the factors that explain the price of bitcoin in the international market, this study has built an extensive database on bitcoin for the period 2011-2020 that includes all the variables that have been related to the currency in the previous literature. Subsequently, artificial neural network techniques have been applied to estimate the sensitivity of the variables regarding the price of bitcoin. Specifically, a multilayer perceptron model (MLP) has been developed, which has been one of the artificial intelligence techniques that have provided better prediction results in engineering and finance due to its ability to perform complex tasks, such as classification, pattern recognition, and predictions (Gupta & Wang, 2010). The results obtained with MLP have allowed estimating, with an accuracy of over 93%, an unpublished set of variables with a high impact on the price of bitcoin. In this way, the results provide important theoretical and practical contributions to the bitcoin literature, especially indicating new research perspectives that can be considered in cryptocurrency valuation models. Also offering an accurate guide to tailoring portfolios and risk management in bitcoin-related investment plans.

The rest of the study is organized as follows. After this introduction, an analysis of the previous literature on the variables related to the price of bitcoin is presented. The methodological aspects of the research are recorded below. Subsequently, the specifications of the variables and the data used are presented. The study ends with the presentation of the results obtained and the main conclusions.

2. Literature review

In the financial literature on cryptocurrencies, studies that try to determine the variables that affect the value of bitcoin stand out. This problem has been the subject of several lines of research that have addressed the analysis of different types of variables. First, the variable-based models on the dynamics between the demand and supply of bitcoin (Civitarese, 2018; Balcilar et al., 2017). Secondly, studies that relate the value of the currency with the attractiveness that it offers investors on social media (Nasir et al., 2019; Hayes, 2017). And thirdly, those that relate the value of bitcoin with macroeconomic and financial variables (Al-Khazali et al., 2018).

2.1. Bitcoin demand and supply

Since there is a definite limit on the amount of bitcoin offered in the market, previous research has attempted to determine the variables that directly impact the demand curve. On the one hand, variables related to market size stand out. The increasing completion of transactions tends to stimulate their adoption by other economic agents, driving the demand for bitcoins. The use of digital money and electronic commerce helps the development and adoption of virtual currencies by boosting their demand (Kristoufek, 2015; Polasik et al., 2015). In this sense, the number of available bitcoins is associated with a negative effect on the value of the currency (Ciaian, 2016a). Likewise, the number of addresses (virtual portfolios) is associated with a direct and positive effect on price (Civitarese, 2018; Ciaian, 2016a). And in the same way, the variables of the volume of daily transactions and transfers by network users have also been shown to have a direct impact on prices (Buchholz et al., 2012). Underlying these studies are the postulates of the quantitative theory of money (Lucas, 1980). This theory postulates that the value of the transactions carried out in an economy must be equal to the amount of money existing by the number of times that money changes hands, a variable that is called the speed of circulation of money. From its application to the field of cryptocurrencies, it follows that the price of the crypto asset is associated with volume measures such as the number of daily transactions and the number of addresses. On the other hand, it is also interesting to consider the influence of bitcoin's circulation speed, which has an inverse relationship with its value (Chen et al., 2020). In this sense, the role of the so-called holders is very important since their existence and financial capacity contribute to reducing the speed of circulation by “hoarding” bitcoins, reducing the possibilities of their transmission.

Other studies find results that qualify the previous conclusions. For example, Bouoiyour and Selmi (2015) indicate that the volume of the market affects the price of bitcoin only in the short term. And for their part, Balcilar et al. (2017) found that the effects of market size are not verified in periods of rising or falling prices that coincide with stressful situations. Finally, other technical and economic variables have also been related to the demand and supply of bitcoins. These include the so-called hash rate, the difficulty of mining a new block for the blockchain, the monetary speed of the circulation of bitcoin, and the commissions received by the miners (Matonis, 2012).

2.2. Attractive on social networks

An alternative model to explain the price of bitcoin is the so-called network economy, an emerging field in the new information society. Under this theory, the capitalization of the bitcoins issued would be the value of a network that basically depends on the number of users of it. Following this premise, previous literature has detected that variables related to attractiveness in social networks have a significant impact on the bitcoin value. For example, searches in electronic media to obtain information about bitcoin and its operation has been one of the most referenced attractive variables. Studies based on vector autoregression and vector error correction methodologies indicated that the number of searches on Google and Wikipedia has a strong temporal association with currency prices and that the public's interest in increasing knowledge about the operation of the asset showed a direct relationship with prices (Kristoufek 2015). Also Nasir et al. (2019), Davies (2014), and Polasik et al. (2015) found a direct and significant relationship between the search history of the term bitcoin on Google, Twitter, Wikipedia, and prices. However, the studies by Ciaian et al. (2016a) and Hayes (2017) found that the attractiveness factor, although significant, has lost relevance over time, due to the consolidation of the currency and the dissemination of knowledge about it. Furthermore, Bouoiyour and Selmi (2015) also found no evidence of the impact of Google searches on price in the long term.

2.3. Macroeconomic and financial variables

Different macroeconomic and financial variables such as the exchange rates of the US dollar, the price of gold and the stock indices have also been shown to have a significant impact on the bitcoin price over time. Zhu, Dickinson and Li (2017), Dyhrberg (2016), and Van Wijk (2013) noted that there was evidence of causality between the price of the dollar and the price of bitcoin, indicating that bitcoin could be used as a hedge for the risk of exposure to the dollar in the short term. For their part, Zhu et al. (2017) found a direct and significant relationship of the price of gold in the bitcoin value. Bouoiyour and Selmi (2015) and Van Wijk (2013) found that stock indices such as the Dow Jones and Shanghai Stock Exchange seemed to be positively correlated in the short and long term with the bitcoin price, and Kristoufek (2015) highlighted the general impact of macroeconomic variables in the prices.

The above conclusions about the impact of macroeconomic and financial variables on the bitcoin price are not out of the question either. Some authors have verified that macroeconomic and financial variables do not have a statistically significant influence on long-term bitcoin prices (Bouri et al., 2017; Chao et al., 2019; Ciaian et al., 2016a; Polasik et al., 2015). For example, Dyhrberg (2016) found that bitcoin has no correlation with the 100 largest companies listed on the London Stock Exchange. For their part, Ciaian et al. (2016b) pointed out that there was no significant statistical relevance regarding factors such as the Dow Jones index and oil prices. Polasik et al. (2015) concluded that the correlation between bitcoin returns and fluctuations in sovereign currencies was weak and statistically insignificant. Al-Khazali et al. (2018) argued that bitcoin is weakly related to macroeconomic variables due to bitcoin's volatility after surprise macro news. And according to Bouoiyour and Selmi (2015) and Kristoufek (2015), the gold price doesn't seem to be related to bitcoin's pricing either.

3. Empirical testing methods

Artificial neural networks have played a major role in empirical studies carried out in the last decade, boosted by the good results obtained and which, in many cases, improved those of the existing statistical models (Alaminos et al., 2018). The main advantages of artificial neural network models lie in the solution of problems regardless of their complexity, not requiring as if they need other statistical models, a linear relationship (Hetch-Nielsen, 1990). Furthermore, Nuñez de Castro and Von Zuben (1998) confirmed that learning in artificial neural networks constituted a special case of a functional approach where there is no assumption about the model underlying the data analyzed.

MLP is an artificial neural network model with one layer of input units, another layer of output units, and a number of intermediate layers called hidden layers, that have no connection to the outside. Each input unit is connected to the units of the hidden layer, and these in turn to those of the output layer. The network aims to establish a correspondence between a set of input data and a set of desired outputs, determining a function that correctly represents the learning patterns and allows a generalization process for data not analyzed during said learning (Heidari et al., 2019). Learning in MLP is a special case of a functional approach in which there is no need for previous hypotheses about the relationship between variables (Faris et al., 2016).

MLP uses expression [1] to provide a fit of W weights from a data set that minimizes learning error, $E(W)$.

$$\boldsymbol{min_W\, E(W) = min_W \sum_{i=1}^{p} \varepsilon(W, x_i}, y_i) \quad [1]$$

where $\{(x_1, y_1), (x_2, y_2) \ldots (x_p, y_p)\}$ represents the set of pairs of learning patterns, and $\varepsilon\,(W, X, Y)$ is the error function.

As a complement to the MLP model, the present study applies a sensitivity analysis of the independent variables. This sensitivity analysis aims to quantify the impact of the variables in explaining the problem under analysis. In this sense, one variable is considered more significant than another if the variance increases compared to the set of variables. For this, the Sobol method (1993) is used, which decomposes the variance of the total output $V(Y)$ according to the equations expressed in [2].

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>1} V_{ij} + \ldots + V_{12\ldots k}$$

$$V_i = V(E(Y|X_i) \quad [2]$$

$$V_{ij} = V(E(Y|X_i, X_j)) - V_i - V.$$

4. Variables and data

The present study uses a set of 22 variables selected from the previous bitcoin literature (Balcilar et al., 2017; Nasir et al., 2019; Zhu et al., 2017). The information corresponding to the variables refers to monthly average data for the period between January 2011 and April 2020. The independent variable is the bitcoin price index in US dollars. For their part, the independent variables refer to three aspects related to the bitcoin value. First, those of demand and supply (including variables of volume, transaction costs, and technology). Second, variables on attractiveness (online news and forums). And thirdly, a group of variables related to macroeconomic and financial factors (oil, gold, stock índices, and exchange rates). The description of the variables used in the present investigation appears in Table 1.

Table 1. Variables of possible influence on the bitcoin price

Variables	Description
a) Demand and supply	
Transaction value	Value of daily transactions
Number of bitcoins	Number of mined bitcoins currently circulating on the network
Bitcoin addresses	Number of unique bitcoin addresses used per day
Transaction volume	Number of transactions per day
Unspent transactions	Number of valid unspent transactions
Blockchain transactions	Number of transactions on blockchain
Blockchain addresses	Number of unique addresses used in blockchain
Block size	Average block size (in megabytes)
Miners reward	Block rewards paid to miners
Mining commissions	Average transaction fees (in USD)
Cost per transaction	Miners' income divided by the number of transactions
Difficulty	Difficulty mining a new blockchain block
Hash	Times a hash function can be calculated per second
Halving	Process of reducing the emission rate of new units
b) Attractive	
Forum posts	Number of new members in online Bitcoin forums
Forum members	New posts in online Bitcoin forums
c) Macroeconomic and financial	
Texas Oil	Oil Price (West Texas)
Brent Oil	Oil Price (Brent, London)
Dollar exchange rate	Exchange rate between the US dollar and the euro
Dow Jones	Dow Jones Index of the New York Stock Exchange
Gold	Gold price (in USD per troy ounce)

The data of the variables related to the demand and supply come from the information provided by the web www.blockchain.com and quandl.com. For its part, the information on attractiveness has been obtained from the web bitcointalk.org. And that of macroeconomic and financial variables from Eurostat, New York Stock Exchange, and web www.goldprice.org for exchange rates and oil and gold prices, respectively. Finally, 80% of the data has been used for training the MLP model and 20% of it has been reserved for testing the model.

5. Results

5.1. MLP model

Table 2 shows the characteristics of the MLP architecture developed in the present study. The best results have been obtained with a structure of 3 neurons in the hidden layer using the hyperbolic tangent activation function.

Table 2. MLP architecture

Parameters	Descripción
Number of neurons in the input layer	21
Number of neurons in the hidden layer	3
Number of neurons in the output layer	1
Scale change method of independent variables	Typified
Hidden layer activation function	Hyperbolic tangent
Error function	Sum of squares

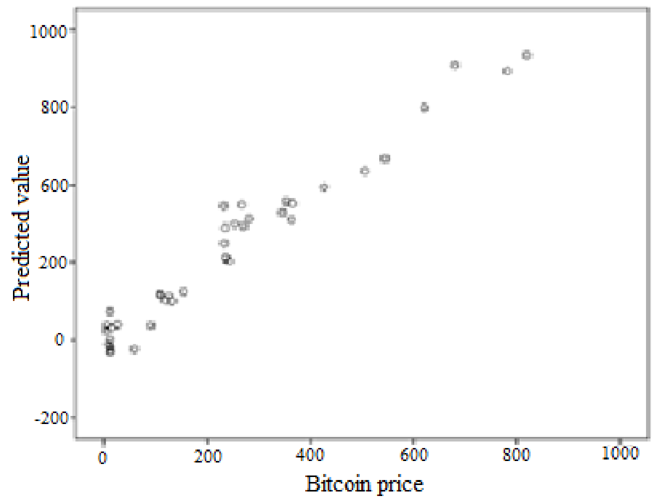
The prediction results obtained with the MLP model appear in Table 3 and Figure 1. These results highlight the high level of precision of the estimation with both the training data (98.58%) and the testing data (93.70%). The other indicators also suggest an acceptable fit of the model.

Table 3. Estimation results

Accuracy (%)		Relative error (%)		RMSE	
Training	Testing	Training	Testing	Training	Testing
98.589	93.7	1.411	6.3	0.037	0.053

RMSE: Square Root of the Mean Square Error

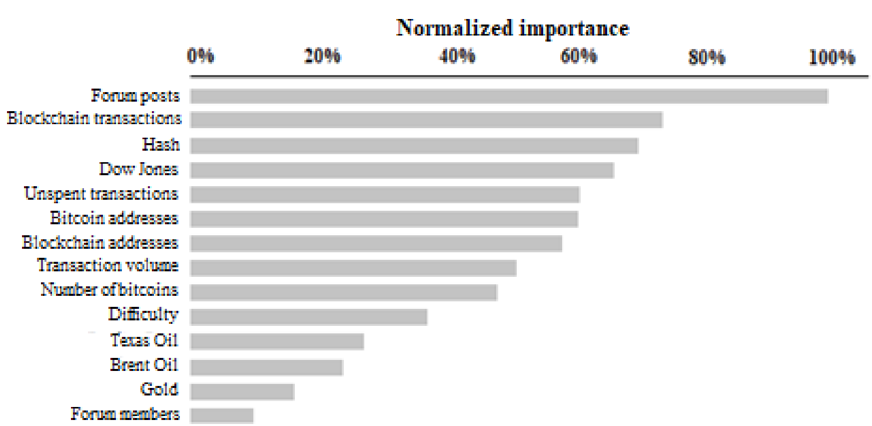
Figure 1. Predicted values



5.1. Sensitivity analysis

For its part, Figure 2 shows the sensitivity of the variables used in the study. The variables forum posts, blockchain transactions, and hash appear with the highest sensitivity values since their sensitivity coefficients exceed 60% in all cases. These results indicate that the incorporation of new participants in the forums related to bitcoin largely explains the levels of the price of the currency in the sample period. In addition, the volume of blockchain operations and the technical aspects related to the possibility that a hash function can be calculated per second are other important aspects to explain the price variations of bitcoin. On the other hand, there is a set of variables that also show significant sensitivity (normalized importance greater than 50%). These variables are related to the market capitalization indices (Dow Jones), with certain volumes of bitcoin operations (unspent transactions) and with the daily existence of a large number of unique addresses (bitcoin addresses and blockchain addresses).

Figure 2. Importance of independent variables



These results can be very interesting because they confirm the main theoretical approaches that try to explain the bitcoin value. Thus, the postulates of the network economy are fully justified as the posts in forums on bitcoin are the most normalized variable of the model. Also because the approach based on the quantitative theory of money is partially validated by the fact that the second variable by normalized importance is the volume of transactions in the blockchain. In addition, other indicators associated with the volume of transactions in the cryptocurrency appear as significant explanatory variables.

6. Discussion

In the present study, a MLP model has been developed that has allowed quantifying the importance of a set of variables to explain the bitcoin price. The results obtained indicate that the most sensitive aspect is related to the attractiveness of the currency on social networks. Thus, the number of posts in bitcoin forums has presented the highest sensitivity. This result is in line with those obtained in the studies by Kristoufek (2015) and Nasir et al. (2019), for which the public's interest in increasing knowledge about bitcoin shows a direct relationship with prices. However, our results differ from those obtained by Ciaian et al. (2016a) and Hayes (2017), for which the attractiveness factor seems to have lost relevance. Perhaps, the most recent sample period used in our study can gather the current significance of the attractive factor, confirming the importance of the network economy in justifying the bitcoin value.

On the other hand, certain variables related to the demand and supply of bitcoin have also shown great sensitivity in our results. Thus, the volume of transactions in the blockchain and the hash rate has reached sensitivities of over 65%. These results are in accordance with those obtained by Buchholz et al. (2012), referring to the fact that the volume of daily transactions by network users has also been shown to have a direct impact on prices, and those of Matonis (2012) regarding the importance of the hash rate. But our results are different from those of Bouoiyour and Selmi (2015), for whom the market volume affects only the bitcoin price in the short term, and those obtained by Balcilar et al. (2017) when they indicate that the size of the market does not have significant effects in periods of currency stress.

We have also been able to detect relative importance of macroeconomic and financial variables, especially in relation to the Dow Jones stock index. Other authors have also highlighted the significance of the stock indices (Bouoiyour & Selmi, 2015; Van Wijk, 2013).

Finally, another set of variables has presented little significance in our study. Such is the case of miners' rewards and commissions, block size, speed, and the euro/dollar exchange rate. These variables have been important in previous studies. For example, Matonis (2012) pointed out that the difficulty of mining a new block for blockchain, the speed of the circulation of the currency, and the commissions received by miners represented variables with an impact on the bitcoin price. On the contrary, our results have not captured this importance, perhaps because the wide set of explanatory variables used in the present study have provided an unprecedented combination of variables to explain the value of bitcoin in the market.

7. Conclusions and implications

The results of the present study have confirmed that the variables with the greatest impact on bitcoin prices are related to the number of forum posts, the volume of transactions on the blockchain, and the hash rate. In addition, that other sets of variables have also shown high sensitivity to currency prices. This is the case of the Dow Jones stock index, the number of daily unspent transactions, and the number of bitcoin and blockchain addresses. This set of variables make up a unique set of bitcoin price predictors that achieves an accuracy of over 93%.

Our study presents several contributions to the literature on bitcoin and cryptocurrencies. First, and from a theoretical perspective, it determines the variables with the greatest impact on bitcoin's price formation. Previous research has shown that variables related to demand and supply, attractiveness, and macroeconomic factors have also impacted the bitcoin price. However, this is the first study that incorporates a wide set of variables and manages to improve the precision results obtained in the previous literature. These conclusions open new research perspectives on evidence that can be considered in the valuation models of cryptocurrencies, especially those based on the quantitative theory of money and the network economy.

The present study also presents important practical implications for portfolio management. The set of variables with the greatest impact identified makes it possible to adapt investment plans based on the behavior of certain market factors. Also, reduce analysis time and risk of cryptocurrency related trades by providing an accurate guide to bitcoin valuation aspects. Bitcoin's specific sensitivity to certain factors is therefore a key variable for its use as an asset in diversified portfolios.

Last, the results of this study suggest future research on the cryptocurrency market. Given that the existing valuation models have only explained part of the behavior of prices, future studies could verify that other theoretical bodies would achieve greater predictive capacity by incorporating new variables not currently explored. Also, and given the excellent results achieved with MLP for the estimation of bitcoin prices, other studies could check if other artificial intelligence techniques would be adequate to predict the behavior of cryptocurrencies, among them ethereum, ripple or tether.

References

Alaminos, D., Becerra-Vicario, R., Fernández-Gámez, M. A., & Cisneros-Ruiz, A. J. (2019). Currency Crises Prediction using Deep Neural Decision Trees. *Applied Sciences*, 9, 5227.

Al-Khazali, O., Bouri, E., & Roubaud, D. (2018). The impact of positive and negative macroeconomic news surprises: gold versus Bitcoin. *Economics Bulleting*, 38(1), 373-382.

Balcilar, M., Bouri, E., Gupta, R., & Roubaud, D. (2017). Can volume predict bitcoin returns and volatility? A quantiles-based approach. *Economic Modelling*, 64, 74-81.

Bouoiyour, J., & Selmi, R. (2015). What does bitcoin look like? *Annals of Economics and Finance*, 16, 449-492.

Bouri, E., Jalkh, N., Molnár, P., & Roubaud, D. (2017). Bitcoin for energy commodities before and after the December 2013 crash: diversifier, hedge or safe haven? *Applied Economics*, 49(50), 5063-5073.

Buchholz, M., Delaney, J., Warren J., & Parker, J. (2012). *Bits and Bets, Information, Price Volatility, and Demand for BitCoin Economics*, 312. https://www.bitcointrading.com/pdf/bitsandbets.pdf.

Chao, X., Kou, G., Peng, Y., & Alsaadi, F. E. (2019). Behavior monitoring methods for trade-based money laundering integrating macro and micro prudential regulation: a case from China. *Technological and Economic Development of Economy*, 25(6), 1081-1096.

Chen, W., Zheng, Z., Ma, M., Wu, J., Zhou, Y., & Yao, J. (2020). Dependence structure between bitcoin price and its influence factors. *International Journal of Computational Science and Engineering*, 21(3). https://doi: 10.1504/IJCSE.2020.106058.

Ciaian, P., Miroslava, R., & Kancs, D. (2016a). Can bitcoin become a global currency? *IseB*, 14, 883-919.

Ciaian, P., Miroslava, R., & Kancs, D. (2016b) The economics of bitcoin price formation. *Applied Economics*, 48, 1799-1815.

Civitarese, J. (2018). *Does metcalfe's law explain bitcoin prices? A time series analysis*. https://ssrn.com/abstract=3107895.

Conley, J. P. (2017). *Blockchain and the Economics of Crypto-tokens and Initial Coin Offerings*. Vanderbilt University Department of Economics Working Papers 17-00008.

Davies, D. C. (2014). *The curious case of bitcoin: is bitcoin volatility driven by online search?* University of Victoria, Victoria.

Dyhrberg, A. H. (2016). Hedging capabilities of bitcoin. Is it the virtual gold? *Finance Research Letters*, 16, 139-144.

Faris, H., Aljarah, I., Al-Madi, N., & Mirjalili, S. (2016). Optimizing the learning process of feedforward neural networks using lightning search algorithm. *International Journal on Artificial Intelligence Tools*, 25, 1650033.

Grinberg, R. (2012). Bitcoin: An innovative alternative digital currency. *Hastings Science & Technology Law Journal*, 4(1), 159-208.

Gupta, S., & Wang, L. P. (2010). Stock Forecasting with Feedforward Neural Networks and Gradual Data Sub-Sampling. *Australian Journal of Intelligent Information Processing Systems*, 11, 14-17.

Hayes, A. (2017). *Cryptocurrency value formation: an empirical analysis leading to a cost of production model for valuing bitcoin*. Telematics Inform. https://ssrn.com/abstract=2648366.

Hecht-Nielsen, R. (1990). *Neurocomputing.* Addison Wesley: Nueva York.

Heidari, A. A., Faris, H., Aljarah, I., & Mirjalili, S. (2019). An efficient hybrid multilayer perceptron neural network with grasshopper optimization, *Soft Computing*, 23, 7941-7958.

Kristoufek, L. (2015) What Are the Main Drivers of the Bitcoin Price? Evidence from Wavelet Coherence Analysis. *PLoS ONE*, 10(4): e0123923.

Lucas, L. (1980). Two Ilustrations of the Quantity Theory of Money. *The American Economlc Review,* 70(5), 1005-1014.

Matonis, J. (2012). *Top 10 BitCoin Statistics*. Forbes. https://www.forbes.com/sites/jonmatonis/2012/07/31/top-10-bitcoin-statistics.

Nasir, M., Huynh, D, Nguyen, P., & Duong, D. (2019). Forecasting cryptocurrency returns and volume using search engines. *Financial Innovation*, 5, 2.

Nuñez de Castro, L., & Von Zuben, F. J. (1998). *Optimised Training Techniques for Feedforward Neural Networks*. Technical Report DCA-RT 03/98. Department of Computer Engineering and Industrial Automation. FEE/UNICAMP, Brasil.

Polasik, M., Piotrowska, A. I., Wisniewski, T. P., Kotkowski, R., & Lightfoot, G. (2015). Price fluctuations and the use of Bitcoin: an empirical inquiry. *International Journal of Electronic Commerce*, 20, 9-49.

Sin, E., & Wang, L. (2017). *Bitcoin Price Prediction Using Ensembles of Neural Networks*. International Conference on Natural Computation, Fuzzy Systems and Knowledge. Guilin, China. https://www.ntu.edu.sg/home/elpwang/PDF_web/08393351.pdf.

Sobol, I. M. (1993). Sensitivity estimates for non linear mathematical models. *Mathematical Modelling and Computational Experiment*, 1, 407-414.

Van Wijk, D. (2013). *What can be expected from the BitCoin?* Working Paper No. 345986. Rotterdam: Erasmus Rotterdam Universiteit.

Zhu, Y., Dickinson, D., & Li, J. (2017). Analysis on the influence factors of Bitcoin's price based on VEC model. *Financial Innovation*, 3, 3.