

Forecasting bitcoin pricing with hybrid models: A review of the literature

Olvera-Juarez D.¹, Huerta-Manzanilla E.²

¹Facultad de Ingeniería, Universidad Autónoma de Querétaro, MX.
Email: denzelolvera@gmail.com

²Facultad de Ingeniería, Universidad Autónoma de Querétaro, MX.
Email:eric.huerta@uaq.mx

Abstract—The electronic transition has been gaining a large ground in recent decades due to the use of cryptocurrencies. One of the most popular is Bitcoin. It is open source, the transactions and the issuance of bitcoins occur collectively through the network. The analysis of the behavior of Bitcoin becomes a relevance to the prediction Price and achieve successful investments in it. This review is conducted for the analysis and comparison of the different prediction methods focused on the bitcoin price. An emphasis is placed on those who have a structure as the basis of the ARIMA model, then adding to the hybrid methods, which use neural networks to complete the method.

Keywords—Bitcoin, cryptocurrency, moving average, autoregressive, price prediction.

Objective. A review of the literature including ARIMA techniques applied in bitcoin forecasting is presented. A summary of metanalysis findings was prepared and a research agenda of potential further work is defined.

I. INTRODUCTION

In the last decades, the globalization and the technology brought great changes in several sectors, such as the economy and administration. One of those changes is electronic money, a new payment method. The cryptocurrency is an electronic currency, due it uses cryptographic tests to control the additional units and verify the transfer of assets. (Nakamoto, 2008).

The cryptocurrencies are a peer to peer version of commerce. The main advantage of these transactions is that payments can be sent from one user to another.

Due to the financial crisis of 2008, interest in cryptocurrencies returned. Cryptocurrencies may have the ability to face several problems relevant for fiat currency system, right at the beginning of the global financial crisis [1]. In fact, Bitcoin was born as a decentralized network and as a digital currency. Internet users split it by using a B to refer to the network.

Bitcoin technology uses cryptographic tests in its software to process transactions and verify the legitimacy of bitcoins and distributes the processing work among the network [2]. This was developed to avoid using trusted third parties, such as bank and cards.

At first Bitcoin operations, it was possible to make payments in the internet without restraint, and without the

costs of central authorities. This allows the behavior of bitcoin as an analogy of assets transference, retaining its value by itself. At the same time, the bitcoin achieves the economic definition of money: it is a mean of Exchange, unit of account and storage of value. [1].

II. PREDICTION TECHNIQUES

2.1 Autoregressive Integrated Moving Average (ARIMA).

The autoregressive integrated moving average (ARIMA) is the most common and widely used time series model. Due to its static properties this model is very important. [3].

This tool can develop several exponential smoothing models and could work in some types of time series, without losing the original characteristics or the time series.

The ARIMA model approach outperforms pure autoregressive series (AR), pure moving averages (MA) and combined AR and MA (ARMA) series models. An important lack of scope of these individual techniques is that they presuppose that the time series are linear (Zhang, 2003).

Using the linear model in the real world, complex processes cannot be represented and have successful results. The ARIMA model has an advantage, this model has individual components that describe trend, error and seasonality separately (p, d, q). That is why nonlinear models can be represented.

2.2 Recurrent Neural Network (RNN)

Neural networks predict the data of an observation along the spatial dimensions in which they occur. These can model the behavior of the observations due to the different learning they use on the existing data. An important way to deal with modeling complications with observations of erratic behavior is factoring. This translates the obstacle of modeling into a sequence problem. From the previously observed data, the network learns to predict the following data. An expressive sequence model is necessary to model non-linear correlations (Oord, Kalchbrenner, & Kavukcuoglu, 2016).

Recurrent Neural Networks (RNN) have a long history with a good performance in neural networks. Typically used in modeling sequential data such as voice recognition and handwriting. These are powerful tools that offer a compact and shared parameterization of conditional distributions series [4].

The prediction of time series data is considered a major problem in machine learning and artificial intelligence. The objective of statistical modeling of language is to predict the next word in the context of textual data; therefore, it deals with a problem of predicting sequential data when building language models [5]. The recurrent neuronal network (RNN) is a neuronal sequence model that obtains the last data in a specific process. This process includes processes such as language, voice recognition and machine translation [6].

Due to the learning ability of data observations, and non-parametric modeling, RNN becomes a very important complementary method to integrate with classical time series prediction methods such as ARIMA.

2.3 Learning machine (LM)

Machine learning models are specialized methods, developed from monolayer neural networks. One of the applications of learning machines is to analyze time series models. Within this field we can find models such as: Bayesian neural networks, multilayer perceptron, radial-based functions, generalized regression neural networks (also called kernel regression), CART regression trees, neighboring K-closest regression, Gauss processes and support vector regression [7].

III. SEARCHING METHODOLOGY

The development of the filtering methodology and the selection of the research keywords are shown below. The purpose is to give a general description of the focus of this review of the literature. The input method was based on

the one developed in the "GAMIFICATION IN HIGHER EDUCATION AND STEM: A SYSTEMATIC REVIEW OF LITERATURE" [8].

1. Objective of the review. Identify keywords and look for question options.
2. Filtering of relevant studies. The selection process began on March 3, 2016 and lasted approximately one month. The academic search service "Web of Science" and "Google Scholar" was used. This to ensure wide coverage, the written keywords were "ARIMA", "Bitcoin" and "Forecasting", looking for title and content fields. The deadline was advanced for 2015-2018. It was not necessary to include consideration of the language, since only documents written in English were found.
3. Selection of works. The filtering process is integrated into a data set of 171 articles. The first step of the selection came from the reading of abstracts, which made discard the first and the largest number of articles.
4. Data graphics. Several characteristics of the studies have been stratified, to obtain an overview for the reader and easily obtain a complete picture of the state of the art. A special approach is made on forecasting efficiency and how is the research behavior on the methods and combination used for this analysis.
5. Organize and report the results. The last stage presents the results of the meta-analysis, highlighting the benefits, limits and problems of each method and approach of the studies.

IV. METANALYSIS

As a result of the literature reviewed, the basic structure of the time series analysis method can be classified. A useful overview of the quantity and type of techniques used in bitcoin price forecasting is shown.

This analysis shows the accuracy of the price prediction methods. It is classified by the method plus another tool. The tool added to the ARIMA model modifies the behavior and skills of the original statistical method.

The literature reviewed shows that most authors begin work with conventional statistical methods for modeling the price of bitcoin. Such as AR, MA, or ARIMA. These tools are used as a starting point for add to neural networks or learning machines.

Table 1. Time series modeling methods used in reviewed literature from 2015 to 2018.

Author	ARIMA	AR	MA	LSTM	RNN	NNETAR	SUTTE	Empirical conditions	IRL
[9]	*								
[10]	*			*					
[11]	*	*	*						
[12]	*				*				
[13]	*					*	*		
[14]			*	*					
[15]	*				*				
[16]	*							*	
[11]	*	*	*						
[17]	*								*

Each method developed in a hybrid manner is not applied to the same study conditions. A crucial condition is the sampling interval. This affects LAG directly. Which allows to observe different behaviors in the closing price of bitcoin.

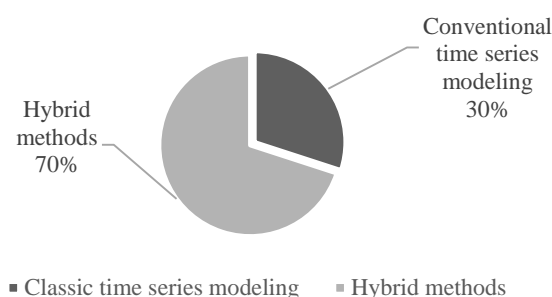


Fig. 1. Comparison between classic and pure statistical methods against hybrid methods which use Artificial Neural Networks.

At first sight, the difference between studies that use conventional statistical methods, 30% of the works being analyzed, against the other 70% are studies that use hybrid methods for the analysis of Bitcoin behavior.

METHODS FORECASTING ACCURACY

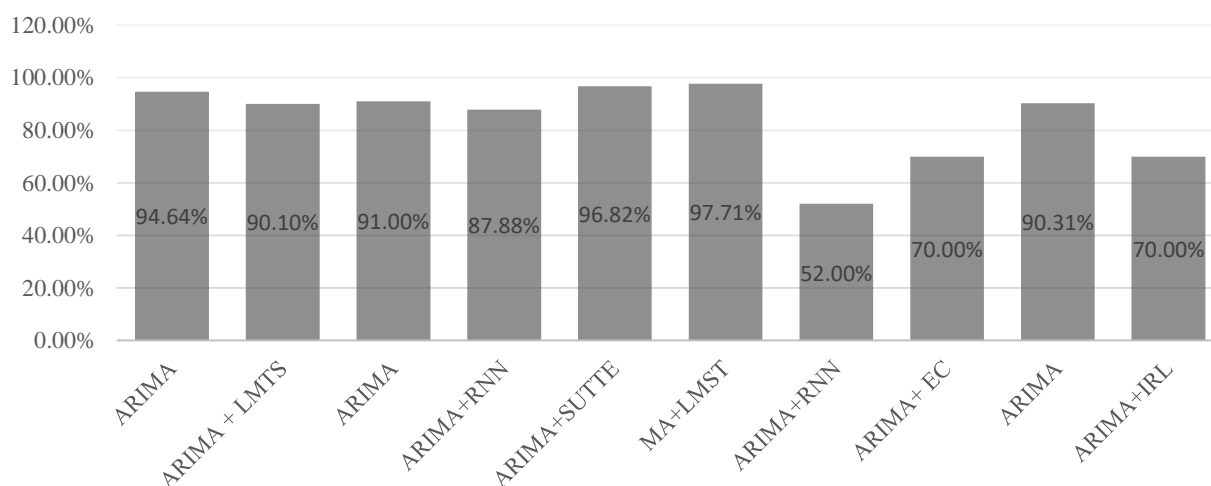


Fig. 2. Comparison of different methods and its prediction of bitcoin price accuracy.

V. DISCUSSION

ARIMA is not, by itself, the best way to model the behavior and prediction of bitcoin prices. The stationary

characteristics facilitate the ARIMA modeling process. Therefore, the data is pre-processed to make them

stationary, and then the values (p, d, q) are obtained to find the ARIMA model that minimizes the prediction MSE.

Different research suggests that predicting the price of bitcoin using its closing price history could result in large MSE values. This due to the vulnerability of bitcoin prices to large variations.

The high volatility environment creates a considerable error for the ARIMA model. Therefore, forecasting in a high volatility environment requires special consideration of error diagnoses. The forecasting approach used by the ARIMA method produces a reliable short-term model.

VI. CONCLUSION

The price of bitcoin during the sample period is a non-stationary time series, and the difference sequence cannot verify the specific type. Therefore, the appropriate ARIMA model cannot be found.

LSTM model make a strong framework with time series techniques. It can build an efficient time series prediction model without strict assumptions of data distribution.

LSTM and ANN provide a new forecast framework for bitcoin price prediction. Also became in tools for several behavior analysis. Industry instances such as medical data or financial time series data.

VII. RESEARCH AGENDA

Survey can continue to determine the factors that contribute to the volatility of the bitcoin exchange rate. In the same way the correlation of bitcoin with another currency can be considered as another area to analyze.

Research papers are important for predicting the bitcoin exchange rate in a high volatility environment. This information will help investors make predictions of the bitcoin exchange rate. For the same task, volatility should be monitored for trends and possible causes of this.

VIII. ACKNOWLEDGMENTS

The review would not have been these results without the stimulating contribution of CONACYT, for the financial support provided during the completion of this project with scholarship number 932840. Also want to thank to the Autonomous University of Querétaro for his very valuable support with the master's scholarship.

REFERENCES

- [1] L. P. Nian and D. L. K. Chuen, "Introduction to Bitcoin," in *Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data*, Elsevier Inc., 2015, pp. 5–30.
- [2] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," *J. Gen. Philos. Sci.*, vol. 39, no. 1, pp. 53–67, 2008.
- [3] P. G. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [4] M. Liang and X. Hu, "Recurrent convolutional neural network for object recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, no. Figure 1, pp. 3367–3375, 2015.
- [5] A. Ingolfsson and E. Sachs, "Stability and Sensitivity of an EWMA Controller," *J. Qual. Technol.*, vol. 25, no. 4, pp. 271–287, 1993.
- [6] W. Zaremba, I. Sutskever, and O. Vinyals, "Recurrent Neural Network Regularization," no. 2013, pp. 1–8, 2014.
- [7] N. K. Ahmed, A. F. Atiya, N. El Gayar, and H. El-Shishiny, "An empirical comparison of machine learning models for time series forecasting," *Econom. Rev.*, vol. 29, no. 5, pp. 594–621, 2010.
- [8] M. Ortiz, K. Chiluiza, and M. Valcke, "GAMIFICATION IN HIGHER EDUCATION AND STEM: A SYSTEMATIC REVIEW OF LITERATURE," in *EDULEARN16 Proceedings*, 2016, vol. 1, no. July, pp. 6548–6558.
- [9] N. A. Bakar and S. Rosbi, "Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Cryptocurrency Exchange Rate in High Volatility Environment: A New Insight of Bitcoin Transaction," *Int. J. Adv. Eng. Res. Sci.*, vol. 4, no. 11, pp. 130–137, 2017.
- [10] C. H. Wu, C. C. Lu, Y. F. Ma, and R. S. Lu, "A new forecasting framework for bitcoin price with LSTM," in *IEEE International Conference on Data Mining Workshops, ICDMW*, 2019, vol. 2018-Novem, pp. 168–175.
- [11] S. Roy, S. Nanjiba, and A. Chakrabarty, "Bitcoin Price Forecasting Using Time Series Analysis," in *2018 21st International Conference of Computer and Information Technology, ICCIT 2018*, 2019, pp. 1–5.
- [12] A. Azari, "Bitcoin Price Prediction: An ARIMA Approach," 2019.
- [13] D. U. Sutiksno, A. S. Ahmar, N. Kurniasih, E. Susanto, and A. Leiwakabessy, "Forecasting Historical Data of Bitcoin using ARIMA and α -Sutte Indicator," in *Journal of Physics: Conference Series*, 2018, vol. 1028, no. 1.
- [14] 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," *J. Fundam. Appl. Sci.*, vol. 9, no. 3S, p. 791, 2018.
- [15] J. Rebane and I. Karlsson, "Seq2Seq RNNs and ARIMA models for Cryptocurrency Prediction: A Comparative Study," no. August, pp. 2–6, 2018.
- [16] M. J. Amjad, "Trading Bitcoin and Online Time Series Prediction," *Time Ser. Work.*, pp. 1–15, 2016.
- [17] S. Roy and A. Chakrabarty, "Bitcoin Price Forecasting Based on Historical Data Submitted By : Supervisor :," no. 15101137, 2018.