Currency Price Prediction using Machine Learning Algorithms

Kamini¹

Research Scholar Swami Vivekanand institute of engineering and Technology, Banur Punjab kaminidhiman140@gmail.com

Prince Sood²

Assistant professor Swami Vivekanand institute of engineering and Technology, Banur Punjab

Abstract

Bitcoin, being the most well-known blockchain technology, has recently received a lot of attention in the fields of economics and finance. The purpose of this dissertation is to determine if new machine-learning models can forecast better than traditional models. This research compares the accuracy of bitcoin price prediction using two different models in terms of forecasting errors: Long-Short Term Memory (LSTM) against Auto Regressive Integrated Moving Average (ARIMA), and Python routines were used. The Federal Reserve Economic Statistics was used to compile Bitcoin price data from 2017-06-18 to 2019-08-07. To compare the results of both models, the data was separated into two subgroups: training (83.5%) and testing (83.5%). (16.5 percent). In most cases, the literature shows that LSTM outperforms ARIMA. According to RMSE and MAE, In this dissertation, LSTM forecasts of bitcoin prices outperform average ARIMA predictions by 92 percent and 94 percent, respectively.

Keywords: ARIMA, LSTM, Bitcoin, Forecasting.

1. Introduction

Predicting economic and financial time series data is a difficult task due to unknown events and inadequate information in the economy in which we live. In chronological order, this condition leads to great volatility. To make more accurate forecasts, the authors have used increasingly sophisticated forecasting approaches over time.

The goal of this research is to see if machine learning models can make better predictions than traditional models in terms of prediction accuracy and low prediction error. Autoregressive integrated moving averages (ARIMA) and lengthy short-term memory will be investigated (LSTM), which is a type of recurrent neural network (RNN). In this paper, we run Python code1 on Spyder (Python 3.6) to estimate both models. What makes Python so special? Python is a dynamic freeware / open source language suitable for interactive development. It also supports libraries and is widely used in machine learning and data science. To run time-series programs in Python, NumPy for efficient array manipulation, Matplotlib for data plotting, Pandas for powerful data loading and processing, Statsmodels for statistical modeling, development and machine You will need to import some major libraries such as ScikitLearn for learning to practice in Python.

Bitcoin has established itself as the world's most important cryptocurrency in recent years and has attracted the attention of consumers, businesses and investors. Since Bitcoin networks are complex, we would like to look at some machine learning techniques for predicting Bitcoin.

The remaining chapters of this paper are divided into six sections. Chapter 2, which is divided into two parts, is dedicated to literature review. First, I will explain how Bitcoin and blockchain work, give an overview of what happened before this cryptocurrency was created, explain some of the strengths and weaknesses of the currency, and recently. Learn about the various other cryptocurrencies that have challenged Bitcoin. The second portion looks back at earlier Bitcoin research, including comparisons of traditional econometric models and innovative machine learning approaches. The predictive capabilities of ARIMA and LSTM models are discussed in Chapter 3. This section (RNN) also presents artificial neural networks (ANNs) and recurrent neural networks to help you better comprehend the LSTM technique (RNNs). The data utilised throughout the paper is reviewed in Chapter 4. In the fifth chapter, we will implement each model step by step to get Bitcoin Price Prediction according to the methodology. It also compares the predictions of the two models in terms of prediction error, particularly root mean square error (RMSE) and moving average error (MAE) (MAE). Chapter 6 summarises the findings, highlights the paper's main findings, and offers some areas for additional research.

2. Literature Review

According to Nathan Reiff's essay (Reiff, 2019), the Netherlands was one of the first countries to try to build a cryptocurrency in the late 1980s. Gas station theft was rampant, therefore groups of people tried to avoid spending cash by putting money into new smart cards.

Simultaneously, an American cryptographer named David Chaum devised digiCash, a new sort of e-commerce (Chaum, 1983). He has developed a "blind formula" for encrypting data transmitted between people. This gadget will allow you to securely transfer money between people simply by verifying the validity of your signature. This breakthrough has paved the way for future cryptocurrencies. In the 1990s, some companies used these fundamentals. PayPal was the company with the most lasting impact. This well-known company has revolutionized online payments between individuals. Individuals can send money quickly and securely via the Internet. The egold application was one of the most successful. Wei Dai created bMoney, an "anonymous decentralized electronic payment system" in 1998 (Dai, 1998). This concept was based on the use of digital pseudonyms for sending money over distributed networks. Instead of using a third party, the framework included a network contract. Despite the fact that it was not as popular as other cryptocurrencies, Nakamoto (2008) mentioned Bmoney in his bitcoin whitepaper.

Hashcash (Back, 1997), a prebitcoin digital currency developed in the mid1990s, was one of the most effective. It created and distributed new coins using proofofwork techniques.

Szabo (2008) suggested Bit Gold, which established a proof-of-work method that is employed in bitcoin's mining network in various aspects. In the following section, I'll go through how the most popular cryptocurrency, bitcoin, works.

3. Result analysis

The Autoregressive Integrated Moving Average (ARIMA) is a linear model that incorporates an AR, an MA, and an integrated component that differenciates the time series to convert it into a stationary process.

ARIMA is a modified version of the ARMA model that contains built-in components that are beneficial when the data is non-stationary5. ARIMA applies derivatives to the time series to remove transient values. ARIMA (,,) represents AR, MA, and derivatives in sequence. The Autocorrelation Function (ACF) examines how the series relates to itself at various lags. Helps determine the number of moving average (q) lags. The partial autocorrelation function (PACF) can be thought of as a time-series regression to the previous lag. Similarly, you can generate a virtual autoregressive clause sequence (p). You can also use the Akaike Information Criterion to select the optimal ARIMA model (AIC). After checking the residuals, you can proceed with the forecast calculation. This model is known for its accuracy in making predictions. However, because ARIMA is a linear model, it has some limitations when dealing with non-linear situations and may be suitable for shorter-term forecasts.

3.2.2 Non-linear machine learning models

3.2.2.1 ANN

ANNs are inspired by neural connections and aim to predict neural network behaviour. The more robust and self-adapting the ANNs are in forecasting non-linear difficulties, the better responses will be provided. Neurons, which are little processing units linked together by connections, make form a neural network. Weight refers to the strength of a connection between two neurons. A weight matrix or a weight vector can be used to express these weights. The input layer, hidden layer, and output layer are the three layers in which the neurons are organised. A multiple layer neural network is depicted in Figure 3.2.



<u>Figure 3.2</u> – Multi-layer perceptron⁶.

The propagation function accepts outputs from other nodes and converts them into neural input that can be processed by the activation function utilising linking weights, resulting in network input. It turns the network input as well as the previous state into a new activation state, with a threshold value being critical because it specifies the maximum gradient value of the activation function. As shown in figure 3.3, the most commonly used activation functions are the sigmoid (sig), hyperbolic tangent (tan), and a Rectifying Linear Unit (ReLu). The output function calculates the output value of a neuron based on its activation state.



*Figure 3.3 – Activation functions*⁷

It is possible to adjust the network to fit our needs depending on the learning strategy used. There are three distinct types of education: I Supervised learning: employing a collection of training datasets, learning from examples handed down by a supervisor.

ii) Unsupervised learning, which involves modelling data without assigning labels. Clustering and dimensional reduction can be used to create patterns with enough data; iii) reinforcement learning: applying the information acquired by observing how the environment reacts to action. This sort of machine learning is based on interactions with the environment in order to figure out which combination of actions produces the best results.

ANNs, on the other hand, are not ideal for sequential data because they lack the ability to represent dependencies because the network has no recall of past time steps.

3.2.2.2. RNN

Recurrent neural networks (RNNs) differ from normal neural networks in that they use a transition weight to transport information across time. According to this transition weight, the next state is reliant on the previous one. The model now has memory, as indicated by this. The data obtained earlier in the process is stored internally in the RNNs' hidden layers. The term "recurrent" refers to the fact that the model does the same thing for each element in the sequence, forecasting future values based on previous data. RNN is seen in Figure 3.4.



*<u>Figure 3.4.</u> Recurrent neural network with p time steps*⁸*.*

Long-range dependencies are difficult for RNNs to learn. Going too far back in time for time series forecasting is challenging because the information gathered in prior time steps either vanishes or becomes considerably magnified. These occurrences are referred to as vanishing or exploding gradients. These issues can be solved with Long Short-Term Memory (LSTM).

3.2.2.3. LSTM

The LSTM is a form of RNN with extra elements for memorising sequential input.

A critical component of LSTM is the cell state, which sends information across the sequence chain. It acts as the memory for the network. The cell state can really retain only the essential information in the sequence since information can be withdrawn or added via gates. The gates learn what knowledge is important to remember and what information they should forget during training. As a result, information from previous stages now influences later stages in the sequence.

4. Data

The bitcoin price data11 ('bprice') covers the time period "2017-06-18" to "2019-08-07."

Every day for a certain amount of time There are a total of 781 observations (n=781). It's worth noting that I just looked at data from the last two years to eliminate data from the early bitcoin years that was boring. Since 2017, the price of bitcoin has risen gradually.

5. Conclusion

Machine learning approaches have recently gotten a lot of press throughout the world. The primary goal of this study was to see if these new strategies were more effective than the old ones. To do so, I compared the accuracy of ARIMA and LSTM forecasts using a daily time series of Bitcoin prices.

To compare the two models, we divided the time series into training and test data at 83.5 percent and 16.5 percent, respectively. For the ARIMA model, ARIMA (1,1,1), which is based on the smallest AIC criterion, was optimal. This model correctly predicts the direction of travel. The anticipated values, on the other hand, are consistently lower than the actual data. The following model parameters were adjusted for LSTM: sequence length 265, 160 node LSTM layer, 0.5 dropout layer, 1,100 cycle high density layer, and 20% validation split. This model appears to be fairly realistic, with a smoother gradient than the actual data. In terms of error prediction, the results suggest that LSTM predictions are more accurate. On average, the RMSE and MAE are 92% and 94% better than the ARIMA model, respectively. These findings support the findings of the literature, as LSTM technology far exceeds ARIMA. Despite the fact that this paper achieves the desired results, there are some disadvantages to note. First of all, the smaller the dataset and the longer the time series of the LSTM algorithm, the better the performance. Second, Bitcoin prices are not seasonally adjusted. In some months, the Bitcoin price timeline shows seasonality, and even on weekends, users tend to make more transactions than on weekdays. After the seasonality is removed, the time series will be cleaner and predictive performance will be better. Third, with daily data, the series becomes more unstable than expected. As a result, weekly data can help enhance forecast accuracy.

The performance of simpler versions of neural networks, such as ANN, could be investigated in future extensions of the analysis presented in this paper. Using past data to forecast Bitcoin values for the next several days is also appealing. You can also make a model with moving windows. These models focus on creating a time series prediction structure that separates the data into multiple windows. Huisu et al. (2018) found that the rolling window LSTM outperforms the traditional LSTM in terms of validation errors. Not to mention, as previously indicated in Chapter 2.2, certain emotive components should be included in the machinelearning optimization. Despite the fact that the results were successfully handled, bitcoin prices are expected to be extremely difficult to forecast due to its high volatility and fundamentally speculative nature. Krugman mentions it (2018). "Bitcoin ... its price is almost completely speculative and very unpredictable ... who is responsible for the current price?" No one is aware of it. "Purely speculative Due to the inherent uncertainty surrounding variables, even the most sophisticated machine learning models always contain some error. There are cryptocurrencies in the next few years. As the image of Bitcoin deteriorates, new coins with better characteristics will emerge. Machine learning models, on the other hand, have a long way to go. There is still plenty of room to learn more about the pricing behavior of Bitcoin and other cryptocurrencies.

References

- Antonopoulos, A. M. 2014. Alternative Chains, Currencies and Applications, *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*, 219-233. O'Reilly Media, Inc.
- [2] Athey, S., Parashkevov, I., Sarukkai, V., & Xia, J. 2016. *Bitcoin Pricing, Adoption, and Usage: Theory and Evidence*. Working Paper No 17-033. Stanford Institute for Economic Policy Research, Stanford.
- [3] Back, A., 1997. A partial hash collision based postage scheme

[4] Barro, R. J. 1979. Money and the Price Level Under the Gold Standard. *The Economic Journal* .89 (353): 13–33.

- [5] Böhme, R., Christin, N., Edelman, B. & Moore, T. 2015. Bitcoin: Economics, Technology, and Governance. *Journal of Economic Perspectives*, *29*(2), 213–238.
- [6] Box, G. and Jenkins, G. 1970. Time Series Analysis: Forecasting and Control.

Holden-Day, San Francisco.

[7] Chaum, D., 1983. Blind signatures for untraceable payments. In: Chaum, D.,

Rivest, R.L., Sherman, A.T. (Eds.), Advances in Cryptology: Proceedings of

Crypto, vol. 82: 199-203. Springer.

[8] Ciaian, P., Rajcaniova, M., & Kancs, d'Artis. 2015. The economics of BitCoin price formation. *Applied Economics*, 48(19), 1799-1815.

[9] Crosby, M., Nachiappan, Pattanayak, P., Verma, S. & Kalyanaraman, V. 2016. Blockchain Technology: Beyond Bitcoin. *Applied Innovation Review*, (2), 5–20.

[10] Dai, W., 1998. *B-money*, s.l.: s.n. Retrieved from:

http://www.weidai.com/bmoney.txt (accessed 10.08.2019).

- [11] Dyhrberg, A.H. 2015. *Bitcoin, Gold and the Dollar a GARCH Volatility Analysis.* Working Paper Series No. 15/20, University College Dublin, UCD Centre for Economic Research, Dublin.
- [12] Elmasdotter, A., & Nyströmer, C. 2018. A comparative study between LSTM and ARIMA for sales forecasting in retail. Unpublished thesis. KTH Royal Institute of Technology, School of Electrical Engineering and Computer Science, Stockholm.
- [13] Hochreiter, S. & Schmidhuber, J. 1997. Long Short-Term Memory. Neural Computation. 9(8): 1735–1780.
- [14] Huisu, J., Ko, H., Lee, J. & Lee, W. 2018. Predicting Bitcoin Prices by Using Rolling Window LSTM model. Paper presented at Data Science Festival 2018, London, UK.
- [15] Hyndman,R.J & Athanasopoulos, G. 2018 Forecasting: Principles and Practice. Monash University, Australia: 2nd edn, Otexts.
- [16] Kingma, D & Ba, J. 2015. Adam: A Method for Stochastic Optimization.
- Published as a conference paper at ICLR, 2015, San Diego, USA.
- [17] Krugman. P. 2018. Bubble, Bubble, Fraud and Trouble. The New York Times.
- [18] Nian, L & Lee, D. 2015. Introduction to Bitcoin. In. D. Lee (Eds). Handbook of

Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big

Data. 5-29. Singapore: Elsevier.

- [19] Matta, M., Lunesu, I. & Marchesi, M. 2015. Bitcoin Spread Prediction Using Social and Web Search Media. Unpublished dissertation, Università degli Studi di Cagliari, Cagliari, Italy.
- [20] Mcnally, S. 2016. Predicting the price of Bitcoin using Machine Learning. Paper presented at 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP 2018), Cambridge, UK.

[21] Pal, D. A., & Prakash, D. P. 2017. Practical Time Series Analysis -Master Time

Series Data Processing, Visualization, and Modelling using Python. Birmingham: Packt Publishing.

[22] Nakamoto, S. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System.

[23] Schilling, L., & Uhlig, H. 2019. Some simple bitcoin economics. Journal of

Monetary Economics.

- [24] Schmidhuber, J., Gers, F.A. & Cummins, F. 1999. Learning to forget: Continual prediction with LSTM. Presented at International Conference on Artificial Neural Networks, Edinburgh, Scotland.
- [25] Shukla, N., & Fricklas, K. 2018. A peek into autoencoders. *Machine Learning with TensorFlow*. Shelter Island, New York: Manning Publications Co.
- [26] Stiglitz, J. 2019. '*We should shut down the cryptocurrencies.*' Retrieved from: <u>https://www.cnbc.com/2019/05/02/joseph-stiglitz-we-should-shutdown-the-</u> <u>cryptocurrencies.html (accessed 26.08.2019).</u>

- [27] Namin,S. S., & Namin, A. S. 2018. Forecasting Economics and Financial Time Series: ARIMA vs. LSTM. Unpublished dissertation, Texas Tech University, Texas, USA.
- [28] Nguyen, M. 2018. Ilustrated Guide to LSTM's and GRU's: A step by step explanation. Retrieved from: <u>https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21</u> (accessed 24.08.2019).

[29] Reiff, N. 2019. *Were there Cryptocurrencies before Bitcoin?* Retrieved from: <u>https://www.investopedia.com/tech/were-there-cryptocurrencies-bitcoin/</u> (accessed 09.08.2019).

[30] Stenqvist, E., & Lönnö, J. 2017. Predicting Bitcoin price fluctuation with Twitter sentiment analysis. Unpublished thesis. KTH Royal Institute of Technology, School of Electrical Engineering and Computer Science, Stockholm.

[31] Szabo, N., 2008. *Bit gold*, s.l.:s.n. Retrieved from:

http://unenumerated.blogspot.com/2005/12/bit-gold.html_(accessed 10.08.2019).

[32] Velde, F.R. (2013). Bitcoin: A primer. Chicago Fed Letter.

[33] Yermarck, D. 2015. Is Bitcoin a Real Currency? An Economic Appraisal. In: D.

Lee (Eds). Handbook of Digital Currency: Bitcoin, Innovation, Financial

- Instruments, and Big Data. 31-43. New York, USA: Elsevier.
- [34] Zhu, Y., Dickinson, D. & Li, J. 2017. Analysis on the influence factors of Bitcoin's price based on VEC model. *Financial Innovation*, 3:3
- [35] Zohar, A. 2015. Bitcoin: under the hood. The myths, the hype, and the true worth of bitcoins. *Communications of the Acm*, 58(9).