



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 5, May 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Arima Model to Predict the Price of Bitcoin Using Machine Learning

A. Ramesh Reddy, K. Venkata Ramya

P.G. Student, Department of Master of Computer Applications, Annamacharya PG College of Computer Studies,
Rajampet, Andhra Pradesh, India

Assistant Professor, Department of Master of Computer Applications, Annamacharya PG College of Computer Studies,
Rajampet, Andhra Pradesh, India

ABSTRACT: The goal of this paper is to ascertain with what accuracy the direction of Bitcoin price in USD can be predicted. The price data is sourced from the Bitcoin Price Index. The task is achieved with varying degrees of success through the implementation of a Bayesian optimised recurrent neural network (RNN) and a Long Short-Term Memory (LSTM) network. The LSTM achieves the highest classification accuracy of 52% and a RMSE of 8%. The popular ARIMA model for time series forecasting is implemented as a comparison to the deep learning models. As expected, the non-linear deep learning methods outperform the ARIMA forecast which performs poorly. Finally, both deep learning models are benchmarked on both a GPU and a CPU with the training time on the GPU outperforming the CPU implementation by 67.7%. Deep learning models such as the RNN and LSTM are evidently effective for Bitcoin prediction with the LSTM more capable for recognising longer-term dependencies. However, a high variance task of this nature makes it difficult to transpire this into impressive validation results. As a result, it remains a difficult task. There is a fine line between overfitting a model and preventing it from learning sufficiently. Dropout is a valuable feature to assist in improving this. However, despite using Bayesian optimisation to optimize the selection of dropout it still couldn't guarantee good validation results. Despite the metrics of sensitivity, specificity and precision indicating good performance, the actual performance of the ARIMA forecast based on error was significantly worse than the neural network models. The LSTM outperformed the RNN marginally, but not significantly. However, the LSTM takes considerably longer to train.

I.INTRODUCTION

Bitcoin is the world's most valuable cryptocurrency and is traded on over 40 exchanges worldwide accepting over 30 different currencies. It has a current market capitalization of 9 billion USD according to <https://www.blockchain.info/> and sees over 250,000 transactions taking place per day. As a currency, Bitcoin offers a novel opportunity for price prediction due to its relatively young age and resulting volatility, which is far greater than that of fiat currencies. It is also unique in relation to traditional fiat currencies in terms of its open nature; no complete data exists regarding cash transactions or money in circulation for fiat currencies.

Prediction of mature financial markets such as the stock market has been researched at length. Bitcoin presents an interesting parallel to this as it is a time series prediction problem in a market still in its transient stage. Traditional time series prediction methods such as Holt-Winters exponential smoothing models rely on linear assumptions and require data that can be broken down into trend, seasonal and noise to be effective. This type of methodology is more suitable for a task such as forecasting sales where seasonal effects are present. Due to the lack of seasonality in the Bitcoin market and its high volatility, these methods are not very effective for this task. Given the complexity of the task, deep learning makes for an interesting technological solution based on its performance in similar areas. The recurrent neural network (RNN) and the long short term memory (LSTM) are favoured over the traditional multilayer perceptron (MLP) due to the temporal nature of Bitcoin data.

The aim of this paper is to investigate with what accuracy the price of Bitcoin can be predicted using machine learning and compare parallelisation methods executed on multi-core and GPU environments. This paper contributes in the following manner: of approximately 653 papers published on Bitcoin, only 7 (at the time of writing) are related to machine learning for prediction. To facilitate a comparison to more traditional approaches in financial forecasting, an ARIMA time series model is also developed for performance comparison purposes with the neural network models.

II. LITERATURE REVIEW

Greaves, A., & Au, B. (2015). Using the bitcoin transaction graph to predict the price of bitcoin. *No Data*.

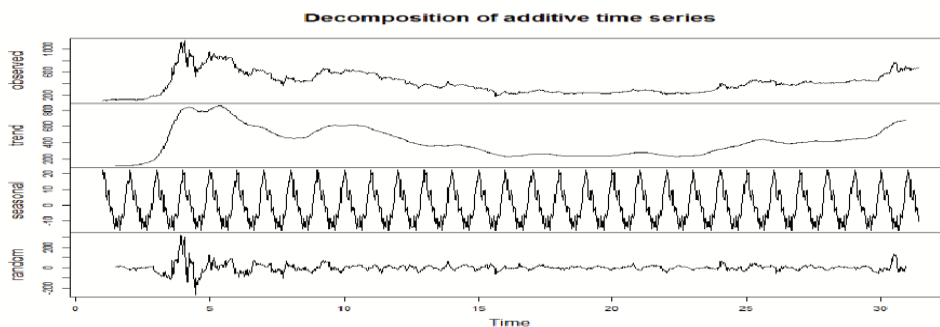
Bitcoin is the world’s leading cryptocurrency, allowing users to make transactions securely and anonymously over the Internet. In recent years, The Bitcoin ecosystem has gained the attention of consumers, businesses, investors and speculators alike. While there has been significant research done to analyze the network topology of the Bitcoin network, limited research has been performed to analyze the network’s influence on overall Bitcoin price. In this paper, we investigate the predictive power of blockchain network-based features on the future price of Bitcoin. As a result of blockchain-networkbased feature engineering and machine learning optimization, we obtain up-down Bitcoin price movement classification accuracy of roughly 55%.

Hayes, A. S. (2017). Cryptocurrency value formation: An empirical study leading to a cost of production model for valuing bitcoin. *Telematics and Informatics*, 34(7), 1308-1321.

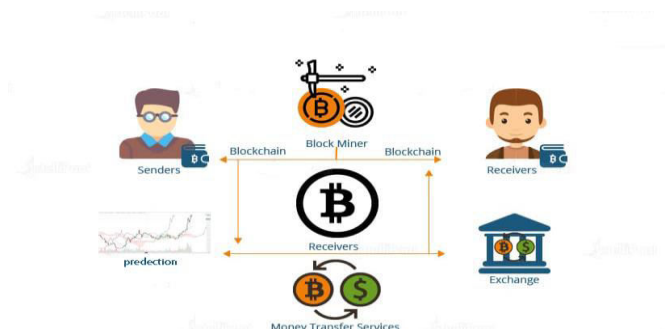
This paper aims to identify the likely determinants for cryptocurrency value formation, including for that of bitcoin. Due to Bitcoin’s growing popular appeal and merchant acceptance, it has become increasingly important to try to understand the factors that influence its value formation. Presently, the value of all bitcoins in existence represents approximately \$7 billion, and more than \$60 million of notional value changes hands each day. Having grown rapidly over the past few years, there is now a developing but vibrant marketplace for bitcoin, and a recognition of digital currencies as an emerging asset class. Not only is there a listed and over-the-counter market for bitcoin and other digital currencies, but also an emergent derivatives market. As such, the ability to value bitcoin and related cryptocurrencies is becoming critical to its establishment as a legitimate financial asset.

III. PROPOSED ALGORITHMS

This paper follows the CRISP data mining methodology. The motivation for CRISP-DM over the more traditional KDD [26] revolves around the business setting of the prediction task. The dataset Bitcoin dataset used, ranges from the 19th of August 2013 until the 19th of July 2016. A time series plot of this can be seen in Figure 1. Data from previous to August 2013 has been excluded as it no longer accurately represents the network. In addition to the Open, High, Low, Close (OHLC) data from CoinDesk, the difficulty and hash rate are taken from the Blockchain. The data was also standardized to give it a mean of 0 and standard deviation of 1. Standardization was chosen over normalization as it better suits the activation functions used by the deep learning models.



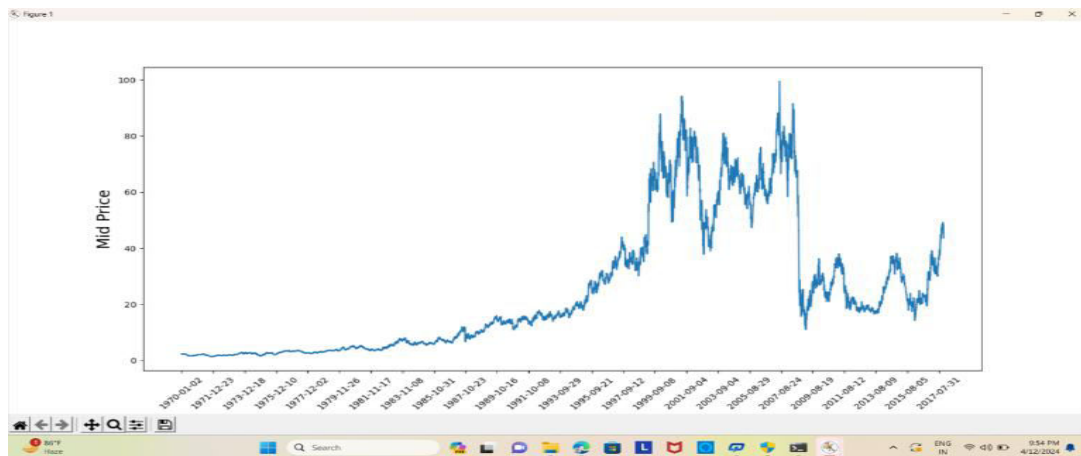
IV. RESEARCH METHODOLOGY



Feature engineering is the art of extracting useful patterns from data to make it easier for machine learning models to perform their predictions. It can be considered one of the most important parts of the data mining process in order to achieve good results in prediction tasks. Several papers in recent years have included indicators including the Simple Moving Average (SMA) for machine learning classification tasks. An example of an appropriate technical indicator is a SMA recording the average price over the previous x days, and is correspondingly included.

To evaluate which features to include, Boruta (a wrapper built around the random forest classification algorithm) was used. This is an ensemble method in which classification is performed by voting of multiple classifiers. The algorithm works on a similar principle as the random forest classifier. It adds randomness to the model and collects results from the ensemble of randomized samples to evaluate attributes and provides a clear view on which attributes are important. All features were deemed important to the model based on the random forest, with 5 day and 10 days (via SMA) the highest importance among the tested averages. The de-noised closing price was one of the most important variables also.

V. RESULT AND DISCUSSION



VI. CONCLUSION

Deep learning models such as the RNN and LSTM are evidently effective for Bitcoin prediction with the LSTM more capable for recognising longer-term dependencies. However, a high variance task of this nature makes it difficult to transpire this into impressive validation results. As a result it remains a difficult task. There is a fine line between overfitting a model and preventing it from learning sufficiently. Dropout is a valuable feature to assist in improving this. However, despite using Bayesian optimisation to optimize the selection of dropout it still couldn't guarantee good validation results. Despite the metrics of sensitivity, specificity and precision indicating good performance, the actual performance of the ARIMA forecast based on error was significantly worse than the neural network models. The LSTM outperformed the RNN marginally, but not significantly. However, the LSTM takes considerably longer to train.

REFERENCES

1. S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.
2. M. Briere, K. Oosterlinck, and A. Szafarz, "Virtual currency, tangible` return: Portfolio diversification with bitcoins," *Tangible Return: Portfolio Diversification with Bitcoins (September 12, 2013)*, 2013.
3. I. Kaastra and M. Boyd, "Designing a neural network for forecasting financial and economic time series," *Neurocomputing*, vol. 10, no. 3, pp. 215–236, 1996.
4. H. White, "Economic prediction using neural networks: The case of ibm daily stock returns," in *Neural Networks, 1988., IEEE International Conference on*. IEEE, 1988, pp. 451–458.
5. C. Chatfield and M. Yar, "Holt-winters forecasting: some practical issues," *The Statistician*, pp. 129–140, 1988.
6. B. Scott, "Bitcoin academic paper database," *suitpossum blog*, 2016.
7. M. D. Rechenhth, "Machine-learning classification techniques for the analysis and prediction of high-frequency stock direction," 2014.

8. D. Shah and K. Zhang, "Bayesian regression and bitcoin," in *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*. IEEE, 2014, pp. 409–414.
9. G. H. Chen, S. Nikolov, and D. Shah, "A latent source model for nonparametric time series classification," in *Advances in Neural Information Processing Systems*, 2013, pp. 1088–1096.
10. I. Georgoula, D. Pournarakis, C. Bilanakos, D. N. Sotiropoulos, and G. M. Giaglis, "Using time-series and sentiment analysis to detect the determinants of bitcoin prices," *Available at SSRN 2607167*, 2015.
11. M. Matta, I. Lunesu, and M. Marchesi, "Bitcoin spread prediction using social and web search media," *Proceedings of DeCAT*, 2015.
12. —, "The predictor impact of web search media on bitcoin trading volumes."
13. B. Gu, P. Konana, A. Liu, B. Rajagopalan, and J. Ghosh, "Identifying information in stock message boards and its implications for stock market efficiency," in *Workshop on Information Systems and Economics, Los Angeles, CA, 2006*.
14. A. Greaves and B. Au, "Using the bitcoin transaction graph to predict the price of bitcoin," 2015.
15. I. Madan, S. Saluja, and A. Zhao, "Automated bitcoin trading via machine learning algorithms," 2015.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details