A Research on Bitcoin Price Prediction Using Machine Learning Algorithms

Devansh Agarwal
11908637

Abstract

In this project, we proposed to predict the Bitcoin price accurately taking into consideration various parameters that affect the Bitcoin value. By gathering information from different reference papers and applying in real time, I found the advantages and disadvantages of bitcoin price prediction.

Each and every paper has its own set of methodologies of bitcoin price prediction. Many papers have accurate price but some other don't, but the time complexity is higher in those predictions, so to reduce the time complexity here in this paper we use a model linked to artificial intelligence named LSTM. The other papers used different models which do not have a great time management, but in LSTM finding of the results from a larger database is quick and fast. So, for this purpose we draw a comparison between other models and LSTM model, this survey paper helps the upcoming researchers to make an impact in their papers. The process happens in the paper is first moment of the research, we aim to understand and find daily trends in the Bitcoin market while gaining insight into optimal features surrounding Bitcoin price. Our data set consists of various features relating to the Bitcoin price and payment network over the course of every years, recorded daily. By preprocessing the dataset, we apply some data mining techniques to reduce the noise of data. Then the second moment of our research, using the available information, we will predict the sign of the daily price change with highest possible accuracy.

Index terms: Bitcoin, crypto currency, Decision Tree, K-Means Algorithm, Lasso Algorithm, nave Bayes algorithm, Prediction, Random Forest.

INTRODUCTION

The Main objective of this project is to predict the bitcoin using Machine Learning Algorithms. The model is based on long short-term memory (LSTM) recurrent neural networks. In all cases, we build investment portfolios based on the predictions and we compare their performance in terms of return on investment.

Machine learning and AI-assisted trading have attracted growing interest for the past few years. Here, we use this approach to test the hypothesis that the inefficiency of the cryptocurrency market can be exploited to generate abnormal profits. We analyze daily data for 1, 681 cryptocurrencies for the period between Sept 2014 and Mar 2022. We show that simple trading strategies assisted by state-of-the-art machine learning algorithms outperform standard benchmarks. Our results show that non-trivial, but ultimately simple, algorithmic mechanisms can help anticipate the short-term evolution of the cryptocurrency market.

The popularity of cryptocurrencies has skyrocketed in 2017 due to several consecutive months of super-exponential growth of their market capitalization. Today, there are more than 1, 500 actively traded cryptocurrencies capitalizing over $300 billion, with a peak of the market capitalization totaling more than $800 billion in Jan. 2018. Between 2.9 and 5.8 million privates as well as institutional investors are in the different transaction networks, according to a recent survey, and access to the market has become easier over time. Major cryptocurrencies can be bought using fiat currency in a number of online exchanges and then be used in their turn to buy less popular cryptocurrencies. The volume of daily exchanges is currently superior to $15 billion. Since 2017, over 170 hedge funds specialized in cryptocurrencies have emerged and bitcoin futures have been launched to address institutional demand for trading and hedging Bitcoin.

LITERATUREREVIEW

We have all considered where bitcoin costs will be one year, two years, five years or even 10 years from now. It's really difficult to anticipate however each and every one of us loves to do it. Tremendous measures of benefits can be made by purchasing and selling bitcoins, whenever done accurately. It has been proven to be a fortune for many people in the past and is still making them a lot of money today. But this doesn't come without its downside too. If not thought of and calculated properly, you can lose a lot of money too. You should have an incredible comprehension of how and precisely why bitcoin costs change (organic market, guidelines, news, and so forth), which implies you should realize how individuals make their bitcoin predictions. Considering these things (supply and demand, regulations, news, etc.), one must also think about the technology of bitcoin and its progress. This aside, we now have to deal with the technical parts using various algorithms and technologies which can predict precise bitcoin prices. Although we came across various models which are currently present like Biological neural networks. (BNN), Recurrent neural network (RNN), Long short-term memory (LSTM), Auto regressive integrated moving average (ARIMA), etc. with machine learning and deep neural network concepts. Normally a time series is a sequence of numbers along time. This is due to the fact that this being a time series data set, the overall data sets should be split into two parts: inputs and outputs. Moreover, LSTM is great in comparison with the classic statistics' linear models, since it can very easily handle multiple input forecasting problems.

In the approach which we are following, the LSTM will use the previous data to predict bitcoin prices 30 days ahead of its closing price. In the approach used by us, we implement Bayesian optimized Recurrent Neural Network (RNN) and a Long Short-Term Memory (LSTM) network. The highest classification accuracy is achieved by LSTM with the accuracy of 52% and a RMSE of 8%. Presently we execute the famous Auto backward incorporated moving normal (ARIMA) model for time arrangement gauging as a correlation with the profound learning models. The ARIMA forecast is out performed by the nonlinear deep learning methods which performed much better. Finally, both the profound learning models are benchmarked on both a GPU and CPU. The training time on the CPU is outflanked by the GPU execution by 67.7%. In the base papers selected by us, the author collected a data set of over 25 features relating to the bitcoin price and payment network over a period of twenty-four years, recorded on a daily basis were able to predict the sign of the daily bitcoin price change with an incredible accuracy of 98.7%.

Result and Discussion

In this section we show the results of our LSTM model. It was noted during training that the higher the batch size the worst the prediction on the test set. Of course, this is no wonder, since the more training, the more prone to overfitting the model becomes. While it is difficult to predict the price of Bitcoin, we see that features are critical to the algorithm, future work includes trying out the Gated Recurrent Unit version of RNN, as well as tuning, on existing hyper-parameters. Below we show the loss from

PREDICTION TECHNIQUES

LSTM IMPLEMENTATION

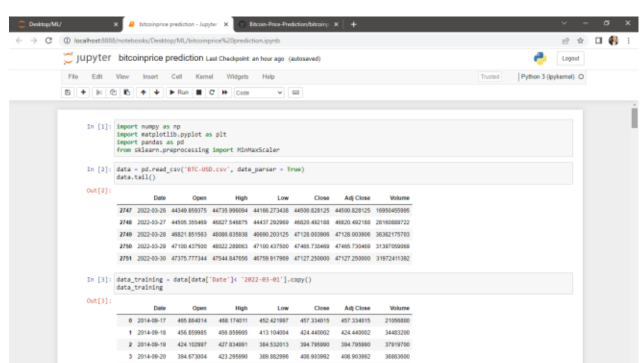
A key feature of feed networks is that they do not save memory. Therefore, each input is processed independently, without the saved state in the middle of the input. Given that we are dealing with a series of times where information from the previous Bitcoin price is required, we should keep track of future events. The building that provides this is the Recurrent neural network (RNN) associated with output has an automatic loop. So, the window we provide as input is processed sequentially rather than one step. However, when the measure of time (size of the window) The largest (most common) gradient is the smallest / largest, leading to a condition known as the disappearance/explosion of the gradient respectively. This problem occurs over time backpropagate optimizer and will activate the algorithm, while the tools are unlikely to change at all. RNN variation reduces the problem, i.e. LSTM and GRU. The LSTM layer adds other data-carrying cells to multiple timesteps. Cell status is a horizontal line from Ct-1 to Ct, and its value lies in capturing long-term or short-term memory. The LSTM effect is adjusted by the government to taste the cells. And this is important when it comes to predicting based on historical context, and not just lastly. LSTM networks can remember input using a loop. These logs are not in RNN. On the other hand, as time goes on, it is less likely that the next result will depend on installation being very old, so forgetting is necessary. LSTM achieves this by learning when to remember and when you forget, by their gates you forget. We will mention soon not to view LSTM as a black-box model.

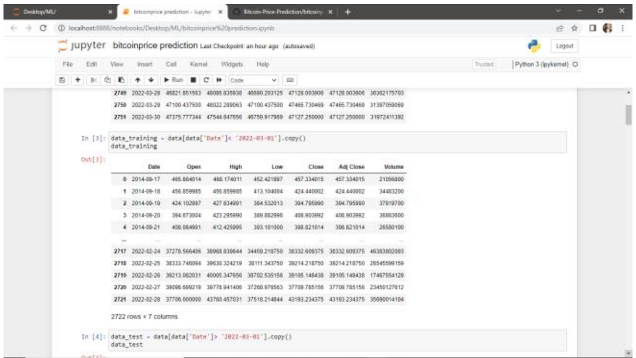
Forget gate: ft = σ(WfStt 1 + WfSt)

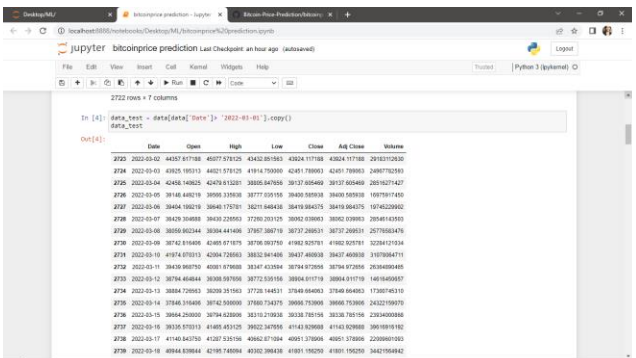
Input gate: it = σ(WiStt 1 + WiSt)

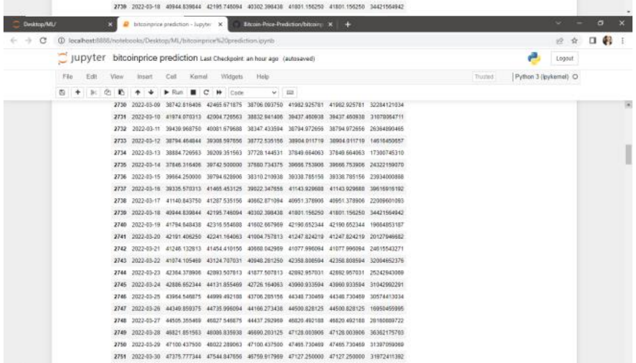
Output gate: ot = σ(WoStt 1 + WoSt).

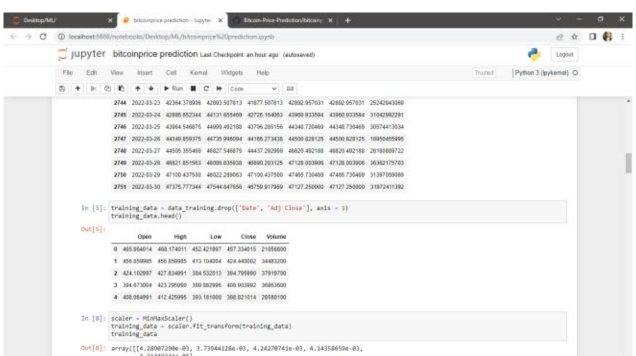
the Mean Absolute Error function, when using the model to predict the training and test data.

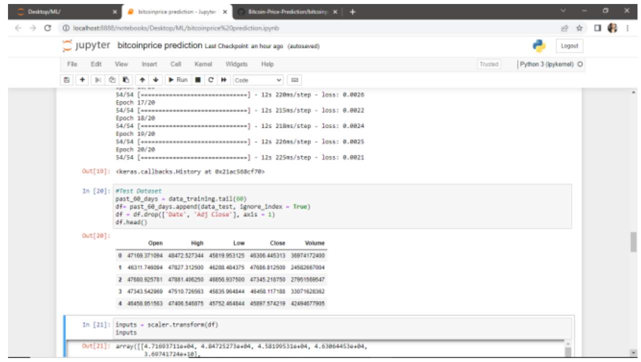


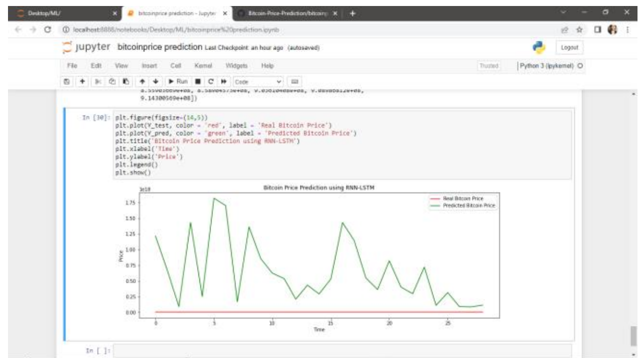












CONCLUSION

All in all, predicting a price-related variable is difficult given the multitude of forces impacting the market. Add to that, the fact that prices are by a large extent dependent on future prospects rather than historic data. However, using deep neural networks has provided us with a better understanding of Bitcoin, and LSTM architecture. The work in progress, includes implementing hyperparameter tuning, in order to get a more accurate network architecture. Also, other features can be considered (although from our experiments with Bitcoin, more features have not always led to better results). Microeconomic factors might be included in the model for a better predictive result. Anyway, maybe 6 Conclusions All in all, predicting a price-related variable is difficult given the multitude of forces impacting the market. Add to that, the fact that prices are to a large extent depended on future prospects rather than historic data. However, using deep neural networks has provided us with a better understanding of Bitcoin, and LSTM architecture. The work in progress, includes implementing hyperparameter tuning, in order to get a more accurate network architecture. Also, other features can be considered (although from our experiments with Bitcoin, more features have not always led to better results). Microeconomic factors might be included in the model for a better predictive result. Anyway, maybe the data we gathered for

Bitcoin, even though it has been collected through the years, might have become interesting, producing historic interpretations only in the last couple of years. Furthermore, a breakthrough evolution in peer-to-peer transactions is ongoing and transforming the landscape of payment services. While it seems, all doubts have not been settled, time might be perfect to act. We think it's difficult to give a mature thought on Bitcoin for the future

References

- Garcia, D. & Schweitzer, F. (2015). Social signals and algorithmic trading of Bitcoin. *Royal Society Open Science* 2(9), 150288. Retrieved from https://dx.doi.org/10.1098/rsos.150288 10.1098/rsos.150288

- Jang, H. & Lee, J. An Empirical Study on Modelling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Blockchain Information. *An Empirical Study on Modelling and Prediction of Bitcoin Prices with Bayesian Neural Networks based on Blockchain Information*.

- Kathyayini, R. S., Jyothi, D. G. & Crypto,". Currency Price Prediction using Machine Learning". *International Journal of Advanced Research in Computer and Communication Engineering*.

- Nakamoto, S. & Shah, D. (2017). Trading Bitcoin and Online Time Series Prediction. *NIPS 2016 Time Series Workshop*.

- Santhosh, &., S. G. &". A Relative Study on Bitcoin Mining. *Relative Study on Bitcoin Mining", Imperial Journal of Interdisciplinary Research*.

- Struga, K. & Qirici, O. Bitcoin Price Prediction with Neural Networks. *Bitcoin Price Prediction with Neural Networks*. Retrieved from http://ceur-ws.org/Vol-2280/paper-06.pdf