Bit Coin Price Prediction Using Time Series Analysis

B. Rama Devi, A. Sai Ganesh, K. Mukesh, M. Ramya, M. Sindhu

QIS College of Engineering and Technology (Autonomous) Vengamukkalapalem, Ongole, Andhra Pradesh

Abstract - 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 crypto currency market can be exploited to generate abnormal profits. We analyze daily data for 1, 681 crypto currencies for the period between Nov. 2015 and Apr. 2018. 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 crypto currency market. The popularity of crypto currencies 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 crypto currencies 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 of private 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 crypto currencies can be bought using fiat currency in a number of online exchanges and then be used in their turn to buy less popular crypto currencies. The volume of daily exchanges is currently superior to $15 billion. Since 2017, over 170 hedge funds specialized in crypto currencies have emerged and bitcoin futures have been launched to address institutional demand for trading and hedging Bitcoin.

I. INTRODUCTION AND EXISTING SYSTEM

In existing system, we analyzed stock markets prediction, suggests that these methods could be effective also in predicting crypto currencies prices. However, the application of machine learning algorithms to the crypto currency market has been limited so far to the analysis of Bitcoin prices, using random forests, Bayesian neural network, long short-term memory neural network and other algorithms. These studies were able to anticipate, to different degrees, the price fluctuations of Bitcoin, and revealed that best results were achieved by neural network-based algorithms. Deep reinforcement learning was showed to beat the uniform buy and hold strategy in predicting the prices of 12 crypto currencies over one year period.

Disadvantage

Other attempts to use machine learning to predict the prices of cryptocurrencies other than Bitcoin come from non-academic sources.

Most of these analyses focused on a limited number of currencies and did not provide benchmark comparisons for their results.

II. PROPOSED SYSTEM

Here, we test the performance of three models in predicting daily crypto currency price for 1,681 currencies.

Two of the models are based on Deep Neural Networks (DNN) and one is based on Linear Regression.

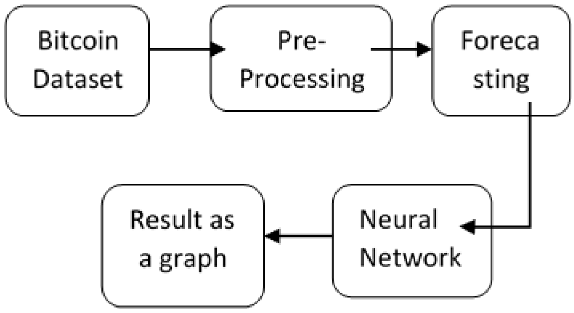
In all cases, we build investment portfolios based on the predictions and we compare their performance in terms of return on investment.

We find that all of the three models perform better than a baseline ‘simple moving average’ model where a currency’s price is predicted as the average price across the preceding days, and that the method based on Linear Regression and Deep neural networks systematically yields the best return on investment.

Advantage

1. We present and compare the results obtained with the three forecasting algorithms and the baseline method.
2. We predict the price of the currencies at day for all included between Jan, 1st 2016 and Apr 24th, 2018.
3. The analysis considers all currencies whose age is larger than 50 days since their first appearance and whose volume is larger than $100000.
4. To discount for the effect of the overall market movement (i.e., market growth, for most of the considered period), we consider crypto currencies prices expressed in Bitcoin.

System Architecture



System Requirements:

Software:

- Python
- Anaconda navigator

Hardware:

- Windows 7,8,10(64 bit)
- RAM 4GB

III. PURPOSE OF THE PROJECT

The purpose of this study is to find out with what accuracy the direction of the price of bitcoin can be predicted using machine learning methods. This is fundamentally time series prediction problem. While much research exists surrounding the use of different machine learning techniques for time series prediction, research in this area relating specifically to bitcoin is lacking. In addition, bitcoin as a currency is in a transient stage and as a result is considerably more volatile than other currencies such as the USD. Interestingly, it is the top performing currency out of the last five years. Thus, its prediction offers great potential and this provides motivation for research in this area. As evidenced by an analysis of the existing literature, running machine learning algorithms on a GPU as opposed to a CPU can offer significant performance improvements. This is explored by bench marking the training of the DNN and Linear Regression using both the GPU and CPU. This provides an answer to the sub research question. Finally in analyzin

g the chosen dependent variables, each variables importance is assessed using a random forest algorithm.

MODULES

1. **Dataset Collection**
2. **Dataset Preprocessing**
3. **Feature Extraction**
4. **Evaluation Model**

Dataset Collection

We first started with Bit coin market data that was publicly available on Kaggle2. The dataset consists of Bit coin historical data from December 1st, 2014 to January 8th, 2018 divided into one-minute increments. This time frame consists of 1,574,274 minutes. For each timestamp, the data included information on the opening value, the closing value, the highest value, the lowest value, the volume traded, and the weighted price. In an effort to iterate quickly and build an initial model, we opted to first analyze the polarity trends in the market. The dataset was labelled as true if the price went up at the end of the minute timestamp and false if it stayed the same or decreased.

Dataset Pre-processing

Scraping the data yields a 2D tensor of m samples by n features. We used a time-series transform to turn this into a set of windows data with window size w=50, yielding a 3D tensor of shape (m - w) samples by n features by w day window size. For instance, our first data point m=0 had a 2D tensor of m features for each of 0-49 days. Then, we normalized the data. Finally, we separated this into the input and output data by removing the last day and making it the output data. Refer to below for a visualization of this.

Feature Extraction

The obvious end-goal of creating a cryptocurrency based neural network is to predict price fluctuations in real time. With this goal in mind, we were eager to start with a highly temporally resolved dataset. If we could get information on a minute by minute or second by second timescale, we could do an even better job of predicting prices and staying ahead of the market. Furthermore, there would be millions of data points and that would be a dataset size that neural networks excel at. However, as we alluded to above, we realized that there also issues with highly resolved data. When looking at our minute dataset, we had an intuition that there would be no change on a minute timescale, or if there was change, that it was very small and noisy. The graph above shows that nearly all the 1.5 million minutes fall within the "third bin" which represented price changes below 0.003%. As a result, our model wouldn't be able to learn the price change since it would mostly be fitting to noisy data and any meaningful change would be drowned out. To note, at this early point in the project we had not actually binned our "y-values" yet but based on our intuition, we decided to convert the minute dataset into a daily dataset. The graph above is made after the fact to show the distribution.

Evaluation Model

3-layer bidirectional DNN to predict the closing price of Bit coin given a variety of data from previous days. Code for scraping crypto currency data is included, as well as the Linear Regression.

IV. CONCLUSION AND FUTURE WORK

Deep learning models such as the DNN and Linear Regression are evidently effectively learners on training data with the DNN more capable for recognizing longer term dependencies. However, a high variance task of this nature make it difficult to transpire this into impressive validation results. As a result it remains a difficult task. There is a fine line to balance between overfitting a model and preventing it from learning sufficiently. After establishing the learning framework and completing the normalization, we intend to use the two methods mentioned above and choose the best method to solve the Bitcoin prediction problem. In terms of the dataset, based on an analysis of the weights of the model the difficulty and hash rate variables could be considered for pruning. Deep learning models require a significant amount to data to learn effectively from. The dataset utilized contained 1066-time steps representing each day. If the granularity of data was changed to per minute this would provide 512,640 data points in a year. Data of this nature is not available for the past but is currently being gathered from CoinDesk on a daily basis for future use.

V. REFERENCES

[1] RakhiMahant, TrilokNathPandey, Alok Kumar Jagadev, and SatchidanandaDehuri ―Optimized Radial Basis Functional Neural Network for Stock Index Prediction,‖ International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) - 2016.

[2] Kai Chen, Yi Zhou and FangyanDai ―A LSTM-based method for stock returns prediction: A case study of China stock market,‖ IEEE International Conference on Big Data,2015.

[3] A.U.S.S Pradeep, SorenGoyal, J. A. Bloom, I. J. Cox, and M. Miller, ―Detection of statistical arbitrage using machine learning techniques in Indian Stock market,‖ IIT Kanpur, April 15, 2013.

[4] Prashant S. Chavan, Prof. Dr. Shrishail. T. Patil ―Parameters for Stock Market Prediction,‖ Prashant S Chavan et al, Int.J.Computer Technology & Applications, Vol 4 (2),337-340.

[5] NeelimaBudhani, Dr. C. K. Jha, Sandeep K. Budhani―Prediction of Stock Market Using Artificial Neural Network,‖ International Conference on Soft Computing Techniques for Engineering and Technology (ICSCTET)- 2014.

[6] SharvilKatariya, Saurabh Jain―Stock Price Trend Forecasting using Supervised Learning Methods.