Synopsis of the B.Tech Minor Project Entitled

Exploring Computational Techniques to Predict Price of Bitcoin

Submitted by

Aditi Mittal(00901032015)

Vipasha Dhiman(04401032015)

Ashi Singh(04501032015)

Under the Supervision of

Dr. Chandra Prakash

Assistant Professor

Dept. of Information Technology

Department of Information Technology,

Indira Gandhi Delhi Technical University for Women, Delhi - 110006

September 2018

Chapter 1

INTRODUCTION

Research in prediction of price of stock market has been done in large numbers generally done using time series analysis. Another interesting addition to this kind of problem is of Bitcoin that was invented in 2009 by Satoshi Nakamoto. In contrast to traditional currencies that works with a central bank or single administrator, Bitcoin is completely decentralized. It can be sent from user to user on the peer-to-peer Bitcoin network. All Bitcoin transactions are posted in blocks to an open ledger known as the Blockchain to be verified by miners using cryptographic proof. Due to its open nature, usage of Bitcoin has increased sharply in last 8 years. When it was first released in 2009, its price was less than 1 cent. 8 years later in 2017, it reached an all-time high of around $20000! Figure 1.1 shows how largely the price of Bitcoin can vary in a single day. [3] Not only this, but the creation of Bitcoin has encouraged more than 1,600 different cryptocurrencies to enter the market to date. Thus, its prediction offers great potential and this provides motivation for research in the area.
The aim of this project is to explore machine learning techniques and with what accuracy can the price of Bitcoin be predicted. Techniques which are generally used for price prediction are Recurrent Neural Network(RNN), Long Short term Memory(LSTM) and Time Series Analysis [1].

Figure 1.1: Variation in price of Bitcoin in a day

Chapter 2

Literature Survey

Several attempts have been made previously to analyze the Bitcoin transaction network for the purpose of analysis.

2.1 Predicting the price of Bitcoin using Machine Learning [1]

The goal of this research was to ascertain with what accuracy can the price of Bitcoin be predicted using machine learning algorithms. The price data was sourced from the Bitcoin Price Index. They implemented Bayesian optimized recurrent neural network (RNN) and Long Short Term Memory (LSTM) network. The LSTM achieved the highest classification accuracy of 52 percent and a RMSE of 8 percent. The popular ARIMA model for time series forecasting was implemented as a comparison to the deep learning models which performed poorly. Wavelets were explored as part of the time series narrative but not implemented for prediction purposes. They also find out with what magnitude of performance improvement can be achieved from parallelization of algorithms on a GPU compared to a CPU. The GPU outperformed the CPU implementation by 67.7 percent.

2.2 An Empirical Study on Modeling and Prediction of Bitcoin Prices With Bayesian Neural Networks Based on Blockchain Information [7]

Bitcoin had attracted considerable attention in the fields of economics, cryptography,and computer science due to its inherent nature of combining encryption technology and monetary units. This paper revealed the effect of Bayesian neural networks(BNNs) by analyzing the time series of Bitcoin process. They conducted the study that compared the Bayesian neural network with other linear and non-linear benchmark models on modeling and predicting the Bitcoin process. The empirical studies showed that BNN performed well in predicting Bitcoin price time series and explained the high volatility of the recent Bitcoin price.

Given the data of the entire time range, experimental results show that the BNN model learned with the selected features effectively describes processes of Bitcoin log price and log volatility. Adoption of rollover framework experimentally demonstrates the predictive performance of BNN is better than other benchmark methods on log price and volatility processes of Bitcoin.

2.3 An experimental study of Bitcoin fluctuation using machine learning methods [8]

This paper focused on making the short-term prediction of the exchange price fluctuations (measured with volatility) towards the United States dollar for the Bitcoin market. The data of realized volatility was collected from one of the largest Bitcoin digital trading offices in 2016 and 2017 as well as order information. Experiments were performed to evaluate a variety of statistical and machine learning approaches.

2.4 Using Bitcoin Ledger Network Data to Predict the Price of Bitcoin [9]

The Bitcoin digital crypto-currency has evolved from basic financial technology to major currency. Every Bitcoin activity is tracked in a blockchain ledger which maintains a complete history of every Bitcoin transaction, offering a unique opportunity to analyze the evolving economy to its fullest. This work seeks to learn a model to predict the next-day trading price of Bitcoin by using data extracted from the transaction network and other economic indicators.

The results are reported by conducting a study on the included data using multiple common learning models . This study identified convolutional neural network to be an optimal model. This model is further refined in its training process, resulting in a final model which generalizes moderately well to testing conditions, though additional improvements are still required to be viable for trading.

Chapter 3

Objectives

Bitcoin prediction will help investors to do effective investments of their money. The prediction system will tell us at what time to withdraw the investment and at what time to make a new investment. It will tell us will Bitcoin crash or rise in near future.
When there is bad news published about Bitcoin, there are a lot more people selling Bitcoin than buying Bitcoin. These people sell their Bitcoin for lower prices than the current value so that they can sell it quickly. This causes the price to drop.
When there is good news about Bitcoin, there are more people buying Bitcoin than there are people selling it. These people buy Bitcoin for higher prices than the current values so that they can buy it quickly. This causes the price to rise.

When Bitcoin was created by Satoshi Nakamoto, he set a limit for how many Bitcoin can be made 21 million. This means that for as long as Bitcoin exists, there can only ever be 21 million no more. So, if the popularity of Bitcoin increases, so should the value.

Objectives of this work majorly includes:

1. To explore various machine learning techniques to predict Bitcoin price
2. Experimental validation of the predictive performance on near-future price of Bitcoin.

Chapter 4

Scope of the Work

For the prediction of bitcoin price, we have collected the data from bitcoin charts for time period February, 2010 to February, 2018. [5] Also we will be applying the computational techniques without considering the security aspects of the bitcoin. We won't be considering the sentiment of news published on daily basis. Figure 4.1 clearly shows that Google search for BTC-USD per day can alter the price of Bitcoin [2]. We won't be taking them into consideration.

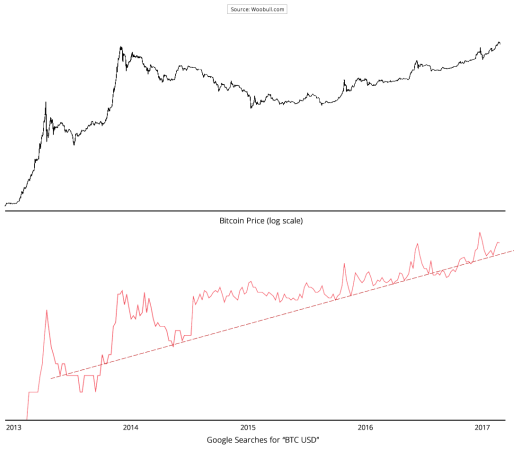

Figure 4.1: Variation in price of Bitcoin with the change in number of Google Search

The decision to limit the cryptocurrency to Bitcoin came down as Bitcoin being the most established cryptocurrency both in age and cryptocurrency market share, reflecting its acceptance in the publics eye. Although, the presented prediction model can be tweaked to any other cryptocurrency by providing the appropriate data. The accuracy estimations would have to be recomputed and would likely vary vastly.

Figure 4.2: Flow of the project

Chapter 5

Proposed methodology

1. Defining the research questions that need to be addressed.
2. Identifying the data source and analyzing the data- Data is collected from Kaggle, Home for Data Science
3. Pre-processing of data by applying techniques such as dimension reduction
4. Implementation of techniques from base paper:
   - Recurrent neural network (RNN)
   - Long short term memory (LSTM)
   - Time series analysis(ARIMA)
5. The performances of the modeling and forecasting are evaluated by root mean square error (RMSE) and mean absolute percentage error (MAPE). The formulas of two measurements are:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_{\mathrm{i}} - \dot{y}_{\mathrm{i}})^2}{N}}$$

$$MAPE = \frac{\sum_{i=1}^{N} \frac{y_{\mathrm{i}} - \dot{y}_{\mathrm{i}}}{y_{\mathrm{i}}}}{N}$$

   where N is the samples number, y_{i} is the i-th actual value, and $\dot{y}_{\mathrm{i}}$ is the i-th estimated value.
6. Exploring other techniques suitable for implementation (eg Singular spectrum analysis)

Recurrent Neural Networks (RNN) are the ones with an internal memory. Because of their internal memory, RNNs are able to remember important things about the input they received, which enables them to be very precise in predicting whats coming next. They produce predictive results in sequential data that other algorithms cant. In a RNN, the information cycles through a loop. When it makes a decision, it takes into consideration the current input and also what it has learned from the inputs it received previously.A usual RNN has a short-term memory.

Long Short-Term Memory (LSTM) networks are an extension for recurrent neural networks, which basically extends their memory. Therefore it is well suited to learn from important experiences that have very long time lags in between. The units of an LSTM are used as building units for the layers of a RNN, which is then often called an LSTM

network. LSTMs enable RNNs to remember their inputs over a long period of time. This is because LSTMs contain their information in a memory, that is much like the memory of a computer because the LSTM can read, write and delete information from its memory.In an LSTM you have three gates: input, forget and output gate. These gates determine whether or not to let new input in (input gate), delete the information because it isnt important (forget gate) or to let it impact the output at the current time step (output gate).The gates in a LSTM are analog, in the form of sigmoids, meaning that they range from 0 to 1. The fact that they are analog, enables them to do back propagation with it.

Time series analysis can be useful to see how a given asset, security or economic variable changes over time. It can also be used to examine how the changes associated with the chosen data point compare to shifts in other variables over the same time period. ARIMA used for stationary time series data. Time series, which contain trend and seasonal patterns, are also non-stationary in nature.In an autoregressive integrated moving average model, the data are differenced in order to make it stationary. A model that shows stationarity is one that shows there is constancy to the data over time. Most economic and market data show trends, so the purpose of differencing is to remove any trends or seasonal structures. When data show regular and predictable patterns that repeat over a calendar year, could negatively affect the regression model. If a trend appears and stationarity is not evident, many of the computations throughout the process cannot be made with great efficacy.

5.1 Gantt Chart along with deliverable

In this work, we will be focusing on analyzing and comparing the accuracy of three computational techniques namely RNN, LSTM and time series analysis. By the end of this semester, we will be able to implement the base paper and identify the extensions that can be applied in the further work.

S.No.	Tasks	Months		
		September	October	November
1	Literature review			
2	Defining the problem and Objective			
3	Data Collection			
4	Pre-processing			
5	Implementation			
6	Comparison of existing techniques			
7	Implementing other techniques			

Figure 5.1: Gantt chart for the tasks to be performed

Bibliography

[1] McNally, Sean & Roche, Jason & Caton, Simon. (2018). Predicting the Price of Bitcoin Using Machine Learning. 339-343. 10.1109/PDP2018.2018.00060.

[2] Willy Woo, Using Google Trends to Detect Bitcoin Price Bubbles "Coindesk, Mar 11, 2017. [Online]. Available: https://www.coindesk.com/using-google-trends-detect-bitcoin-price-bubbles. [Accessed: Sep. 17, 2018].

[3] "Bitcoin (USD) Price," coindesk. [Online]. Available: https://www.coindesk.com/price. [Accessed: Sept. 17, 2018].

[4] Montag Ali, This crypto-millionaire bought a Lamborghini for dollar 115 thanks to bitcoin CNBC, Feb 7, 2018. [Online]. Available: https://www.cnbc.com/2018/02/07/bitcoin-millionaires-are-buying-lamborghinis-with-cryptocurrency.htm. [Accessed: Sep. 18, 2018].

[5] SRK, Cryptocurrency Historical Prices Kaggle, Feb 2018. [Online]. Available:https://www.kaggle.com/sudalairajkumar/cryptocurrencypricehistor. [Accessed: Sep. 4, 2018]

[6] The Rickest Ricky, "Basic Time Series Analysis and Trading Strategy with Bitcoin Price Data", medium, Dec. 29,2017. [Online] Available:https://towardsdatascience.com/basic-time-series-analysis-and-trading-strategy-with-bitcoin-price-data-1a8f1a30f11

[7] Jang, Huisu, and Jaewook Lee. "An empirical study on modeling and prediction of bitcoin prices with bayesian neural networks based on blockchain information." IEEE Access 6 (2018): 5427-5437.

[8] Tian Guo and Nino Antulov-Fantulin. "An experimental study of Bitcoin fluctuation using machine learning methods."[Accessed: Sep. 18, 2018].

[9] Mern, John, Spenser Anderson, and John Poothokaran. "Using Bitcoin Ledger Network Data to Predict the Price of Bitcoin."[Accessed: Sep. 18, 2018].