BITCOIN PRICE FORECASTING BASED ON HISTORICAL DATA

BRAC
UNIVERSITY

Inspiring Excellence

SUBMISSION DATE: 22.07.18

SUBMITTED BY:

Shaily Roy (15101137)

Samiha Nanjiba (15101134)

Department of Computer Science and Engineering

SUPERVISOR:

Amitabha Chakrabarty, Ph.D

Assistant Professor

Department of Computer Science and Engineering

DECLARATION

We, hereby declare that this thesis is based on results we have found ourselves. The materials of work found by other researchers and sources are properly acknowledged and mentioned by reference.

Signature of Supervisor | **Signature of Authors**

Amitabha Chakrabarty, Ph.D
Assistant Professor
Department of Computer Science and Engineering
BRAC University

Shaily Roy (15101137)

Samiha Nanjiba (15101134)

ABSTRACT

Over the past few years, Bitcoin has been a topic of interest of many, from academic researchers to trade investors. Bitcoin is the first as well as the most popular cryptocurrency till date. Since its launch in 2009, it has become widely popular amongst various kinds of people for its trading system without the need of a third party and also due to high volatility of Bitcoin price. In this thesis, our aim is to be able to propose a suitable model that can predict the market price of Bitcoin best by applying a few statistical analysis. We have used Time series method specially Autoregressive Integrated Moving Average (ARIMA) model which can absolutely be called "learning algorithms" and be considered as a part of machine learning (ML) similarly with respect to regression. The work, at last could acquire the accuracy for deciding volatility in weighted costs, with an exactness of 91%.

Keywords: Bitcoin, Time series Analysis, Regression, Machine Learning, ARIMA, Historical data.

ACKNOWLEDGEMENT

We are grateful to our supervisor, Dr. Amitabha Chakrabarty, for his massive help and thoughts all through the work. He made us to escape our usual ranges of familiarity and to push the limit. Without his guidance, it would have been impossible to get introduced to the amazing research prospect of Time Series Analysis.

Additionally, we want to thank Dr. Mahbub Alam Majumdar for his continuous support and guidance when we needed that the most. His classes as well as his support for his all students is a wellspring of motivation for us. Also, we are thankful to our Seniors Pranjal Chakraborty, Tasnia Ashrafi Heya and Dhrubo Hasan who helped us with their experience with Machine Learning and report writing which helped us to make right choices.

Nevertheless, we would like to express our gratitude to Department of Computer Science and Engineering, BRAC University and our educators for assisting us with all the fundamental help.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
Introduction

Bitcoin is a cryptocurrency which has developed as a decentralized, peer-to-peer payment system to encourage safer online transactions without the need of a third party [1]. The transactions, issuing and storing is independent of any central bank or server as it is based on a public distributive network known as blockchain. It utilizes cryptographic techniques and depends on an open-source algorithm which checks decentralized exchanges and controls the formation of new Bitcoin. Regardless of the substantial vacillations of Bitcoin prices (particularly during 2013 and late 2017) and the massive growth in the capitalization of the related market, the condemnations about illicit uses and social concerns, it has still managed to draw the attention of many investors, such as China who is buying Bitcoin, seeing this as an opportunity of investments [2], as well as researchers in the scientific community to study and understand the market in order to predict the worth of Bitcoin.

Recently, Bitcoin has been the topic of many discussion, consequently earning it the acceptance that it has now. As a result, the previous year has been the time when it has seen the most noticeable increase in price, by 1600 US dollar [3], by the end of the year 2017. Hence, this volatility in the cryptocurrency market and the escalation in the popularity of various cryptocurrencies lead us to further explore this area, particularly Bitcoin, since it is the first digital currency and the leading one up to this date. In this paper, we study various models and apply them in order to forecast the Bitcoin prices by using time-series analysis techniques, as well as incorporating machine learning methods to further analyze the performance of our models.

1.1 Motivation

Bitcoin and blockchain are fairly innovative and evolving concepts of this time. Based on this, many other cryptocurrencies have emerged over the years. Moreover, this new technology has opened up new horizons and possibilities to many financiers and Bitcoin miners, who are now speculating on the prospect of whether this can be a new form of asset and financial securities because mining, trading and transactions are secured and reliable. Having said that, since the Bitcoin market is erratic, it is challenging for many to keep track of the market movements.

Therefore, it enthused us work on this project aiming to predict the Bitcoin price with respect to US Dollar with the purpose of aiding these financial enthusiasts by providing them with reliable information which will help them make trading decisions.

1.2 Objective

The year 2017 has been the year of Bitcoin boom. It is reported that the price in USD has almost doubled to about 950 USD at the beginning of the year from that of 2016 and later while towards the end, during the mid of December on 2017, the price eventually doubled again [4], reaching its peak to about 19,000 USD [5]. However, starting this year, the price of Bitcoin gradually started to fall. Since the market is so volatile, yet these cryptocurrencies represent a different type of asset class and are being accepted as the medium of legal trade by many. Hence, the aim of this thesis is to study different models, use them to predict the Bitcoin price in USD and analyze their performance on a given set of data in order to assist its traders and investors to evaluate the risk on financing on Bitcoin.

1.3 Methodology

To implement this thesis proposal, at first we collected the historical data of Bitcoin. Afterwards, we applied some of the time-series analysis techniques and machine learning models to acquire our desired results. We have split the dataset into two parts, the training set that we used to train our models and the test set that we used to predict the closing price of Bitcoin. We have made prediction for the next 10 days and also predicted randomly.

1.4 Thesis Outline

This report is organized as follows:

- **Chapter 1** is an introduction of our overall thesis work, including a short, general idea of the main concept, the motivation, objective and he methodology used in this paper.
- **Chapter 2** summarizes the background of theories and methods that we have used in this thesis.
- **Chapter 3** provides the literature review of the works previously done in this topic.
- **Chapter 4** contains the forecasting methodology and our proposed model.

- **Chapter 5** discusses the results obtained through our work.
- **Chapter 6** concludes the report, mentioning about the future plan of this research and the references used throughout the whole paper.

CHAPTER 2

Background Analysis

2.1 Bitcoin

Among all the available cryptocurrencies, Bitcoin is the most widely established digital currency which was first created by Satoshi Nakamoto in 2009. Initially, it was termed as a decentralized, independent electronic cash-system [6]. Bitcoin became very popular within a few years for its efficiency and for being a cheaper way to transfer money across borders. However, there is a proverb that "all the coins have two sides". Therefore, despite its usefulness, Bitcoin is also criticized for its facility to provide money-laundering, black-market transactions such as "The Silk Road", and terrorist financing. Several researchers think that as an ever increasing number of individuals are using Bitcoin exchange after some time, the unpredictability is diminishing gradually and the time is close when it will develop, consequently making the need to additionally contemplate [7].

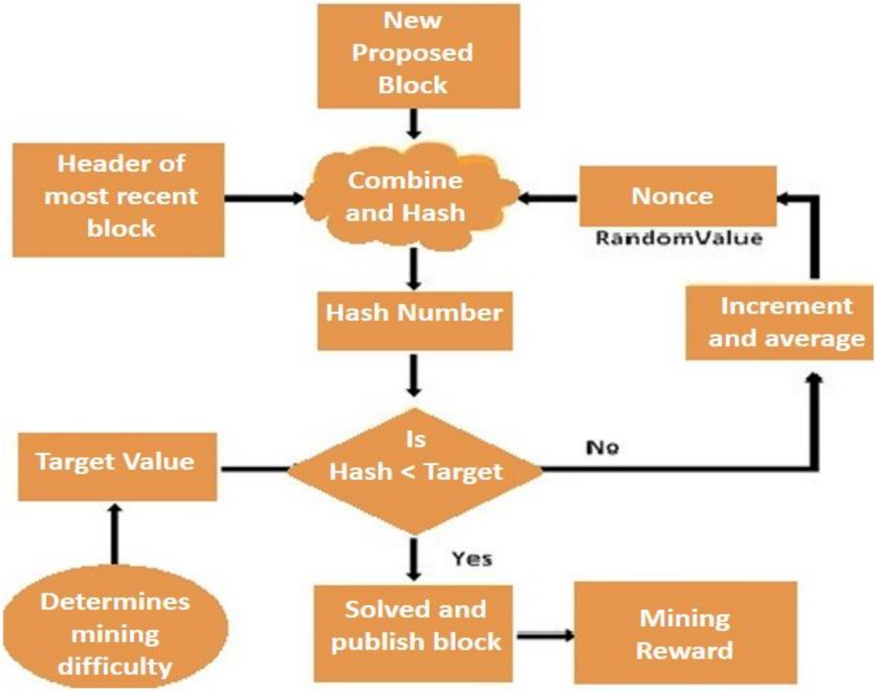


Figure 2.1 Bitcoin mining using Block chain [8]

Bitcoin uses the concept of blockchain in order to mining. Bitcoin transactions are secured by blockchain technology, creating blocks in an open ledger with transaction information encrypted and verified by miners who lend out their computing power in return of a small reward in Bitcoin. The flowchart of Figure 2.1 shows the complete process of Bitcoin mining [8]. The blockchain is a log of all transactions that were ever verified on the Bitcoin network. The transactions are collected in blocks, which are found approximately every ten minutes in a random process called mining. As transactions transfer ownership of Bitcoin balances, each of these blocks represents an update of the user's balances on the network.

2.2 Typical Bitcoin Prediction Method

Predicting is a typical statistical function in business, where it illuminates choices about the booking of creation, transportation and faculty, and gives a manual for long haul key arranging. In this kernel we are foreseeing the future estimation of Bitcoins. The forecasting strategies depend to a great extent on what type of datasets are accessible.

If there are no information accessible, or if the information accessible are not pertinent to the conjectures, at that point qualitative techniques must be utilized. Otherwise, we use quantitative technique. Quantitative technique can be connected when two conditions are fulfilled:

- Numerical data about the past is accessible.
- It is sensible to expect that a few parts of the past examples will proceed into what is to come.

2.3 Time Series Analysis

In the Time Series approach of Bitcoin price analysis and forecasting, it is considered that the Bitcoin price incorporates all important and available information of the financial market. By trying out the underlying structure and function that produced the past observations of the price, time series analysis of Bitcoin prices aims to forecast the future prices or trends. Time Series Analysis may be considered as a mathematical version of technical analysis which uses statistical tools to extract meaningful information from a given historical price series and makes predictions on the basis of those information. We chose Time series analysis because of two technical characteristics of any financial data:

- **Price Moves in Trends**

  The price movements follow a trend and when such trend has been established, the likelihood of the future Bitcoin prices to be in that same direction increases.

- **History Tends to Repeat Itself**

  The patterns of value movement in the previous tends to rehash itself in the present as participants tend to respond in a steady manner to similar stimuli constantly. Time series analysis is not quite the same as other machine learning models in light of the fact that:

  - It is time subordinate. Thus, the fundamental presumption of a straight relapse demonstrate that the perceptions are independent does not hold for this situation.
  - Alongside an expanding or diminishing pattern, most time series have some type of regularity patterns, i.e. varieties particular to a specific time period.

Time series analysis is a measurable strategy that arrangements with time series information, or pattern examination [9]. A time series is a series of data points indexed (or listed or graphed) in time order. Time series data should be treated differently to other types of data. One of the statistical data assumptions is its independence. Independence means the value of one observation does not influence or affect the value of other observations. But in time series data, each data point is close together in time, and they are not completely independent of their adjacent values. So we need a different approach to model time series data.

2.3.1 Terms and Ideas

- **Stationarity:** Shows the mean estimation of the arrangement that remaining parts consistent over an era; if past impacts gather and the qualities increment toward boundlessness, at that point stationarity is not met. The principal suspicion is that the series are stationary. Basically, this implies the arrangement are ordinarily circulated and the mean and change are steady finished quite a while period.
- **Reliance:** Dependence alludes to the relationship of two perceptions with a similar variable, at earlier time focuses.
- **Differencing:** Used to make the arrangement stationary, to De-slant, and to control the auto-relationships; be that as time series analysis do not require differencing and over-differenced arrangement can deliver incorrect assessments.

- **Determination:** May include the testing of the direct or non-straight connections of ward factors by utilizing models, for example, ARIMA, ARCH, GARCH, VAR, Co-incorporation, and so on. It refers to isolating a period arrangement into drift, occasional impacts, and remaining variability assumptions
- **Exponential smoothing in time series analysis:** This technique predicts the one next period esteem in view of the past and current esteem [10]. It includes averaging of information with the end goal that the nonsystematic parts of every individual case or perception counteract each other. The exponential smoothing technique is utilized to foresee the fleeting predication. Alpha, Gamma, Phi, and Delta are the parameters that gauge the impact of the time series information. Alpha is utilized when regularity is absent in information. Gamma is utilized when an arrangement has a pattern in information. Delta is utilized when regularity cycles are available in information. A model is connected by the example of the information. Bend fitting in time series examination: Curve fitting relapse is utilized when information is in a non-straight relationship. The accompanying condition demonstrates the non-straight conduct:

  Subordinate variable, where case is the consecutive case number. Bend fitting can be performed by choosing "relapse" from the examination menu and after that choosing "bend estimation" from the relapse choice. At that point select "needed bend straight," "power," "quadratic," "cubic," "opposite," "calculated," "exponential," or "other." We accept that there is no anomaly in the arrangement. Anomalies may influence conclusions emphatically and can be deluding.

2.3.2 ARIMA

ARIMA remains for auto-regressive integrated moving average. This technique is otherwise called the Box-Jenkins strategy [11].

- **ID of ARIMA parameters**
  - **Autoregressive part:** AR remains for autoregressive. Autoregressive parameter is meant by p. At the point when p =0, it implies that there is no auto-relationship in the arrangement. Whenever p=1, it implies that the arrangement auto-connection is till one slack.

- **Integrated:** In ARIMA time series analysis, coordinated is indicated by d. Mix is the reverse of differencing. Whenever d=0, it implies the arrangement is stationary and we do not have to take its distinction. Whenever d=1, it implies that the arrangement is not stationary and to make it stationary, we have to take the principal contrast. Whenever d=2, it implies that the arrangement has been differenced twice. Generally, in excess of two time contrast is not dependable.
- **Moving average:** MA remains for moving the normal, which is meant by q. In ARIMA, moving normal q=1 implies that it is a mistake term and there is auto-connection with one slack. So as to test regardless of whether the arrangement and their blunder term is auto related, we as a rule utilize W-D test, ACF, and PACF.
- **Uncorrelated arbitrary blunder:** We accept that the mistake term is arbitrarily dispersed and the1 mean and change are consistent over an era. The Durbin-Watson test is the standard test for connected mistakes.
- **Arbitrary stuns (an irregular mistake segment):** If stuns are available, they are thought to be arbitrarily disseminated with a mean of 0 and a steady change.

A stochastic process is a sequence of random variables, represented as follows:

$\{Y^t$: t = 0, $\pm 1, \ \pm 2\}$…………………………………….………...…………[i]

Stochastic processes are used to model observed time series

To summarize, a time series can either be continuous or discrete. It is said to be continuous if it consists of observations taken continuously in time. If the observations are taken in specific equal time intervals, then it is said to be discrete. In this paper, we will be working with discrete time series. In the time series, the distance between any two consecutive time points must be the same and each time point must have at most one observation.

That is, if the series is an observation of monthly data, then it must have the observations of every month; and for each month, only one observation should be taken. Figure 2.2 shows the historical weighted price of Bitcoin from July, 2013 to July, 2017. As we will consider market capital as our main feature, this feature is also shown in the y-axis of figure 2.2[12].

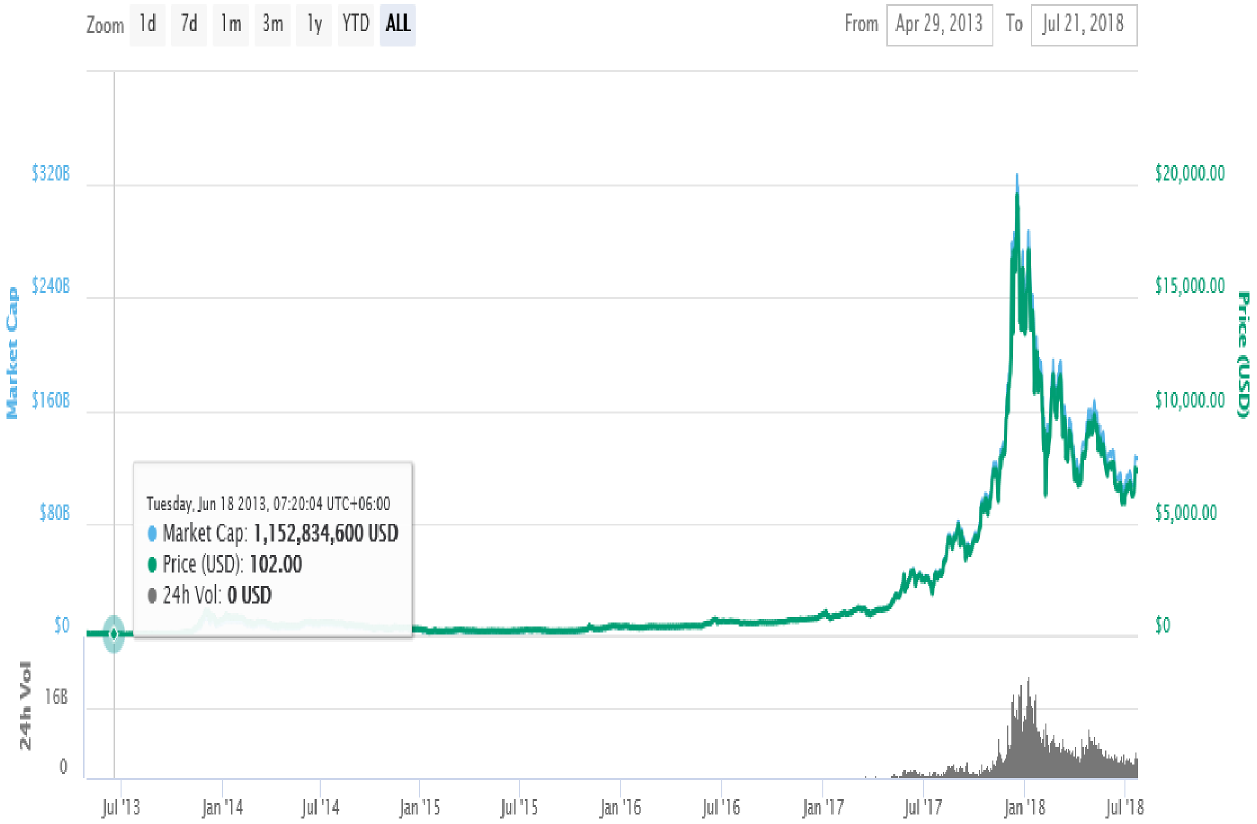


Figure 2.2 Bitcoin Historical Data [12]

Time Series Analysis consists of various techniques of analyzing the time series data with the aim of extracting significant statistics and other important features of data, usually in order to make forecasts of future values based on the past observations. Therefore, during time series analysis, the order of the observations must be maintained. Otherwise, the very meaning of the data would change [13].

2.3.3 Trend Estimation

In stationary time series, we assume that the mean function is constant. However, in practicality, that is never the case and so we often need to consider the mean functions to be simple functions of time or trends. These trends can either be stochastic or deterministic. Stochastic trends are impossible to model because they tend to show completely different characteristics with every simulation, e.g. the random walk model.

On the other hand, a deterministic trend can be modeled using deterministic functions to represent them. For example, a possible model of a time series with deterministic trend could be,

$Y_t = \mu_t + X_t$……………………………...…………………………………………..[ii]

Here,

μ_t = a deterministic function

x_t = unobserved deviations from μ_t, having zero mean

We might consider μ_t to be periodic. We could also assume it to be a linear or a quadratic function of time. However, it must be kept in mind that whenever we are stating that E (x_t) = 0, we are assuming that the trend μ_t will last forever.

2.4 Machine Learning

Different machine learning calculations are being applied for Bitcoin volatility prediction. Among them Sentiment Analysis, Support Vector Machines, Linear relapse, Online Learning, master Weighting and Prediction utilizing Decision Stumps are noteworthy. These machine learning algorithms are used based on same assumption as that of technical analysis. It is expected that [14] relevant information embedded in them. Utilizing machine learning methods alone, it will not be possible to get a good outcome. Therefore, a hybrid of several algorithms or a hybrid of algorithms and some other analysis technique can be used to get a better result. However, same hybrid algorithm will not work similarly for different cryptocurrency's price prediction.

2.4.1 Logistic Regression

Logistic regression [15] is a statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible out-comes). In logistic regression, the dependent variable is binary or dichotomous, i.e. it only contains data coded as 1 and 0. 1 is for true and success, 0 is for false and failure.

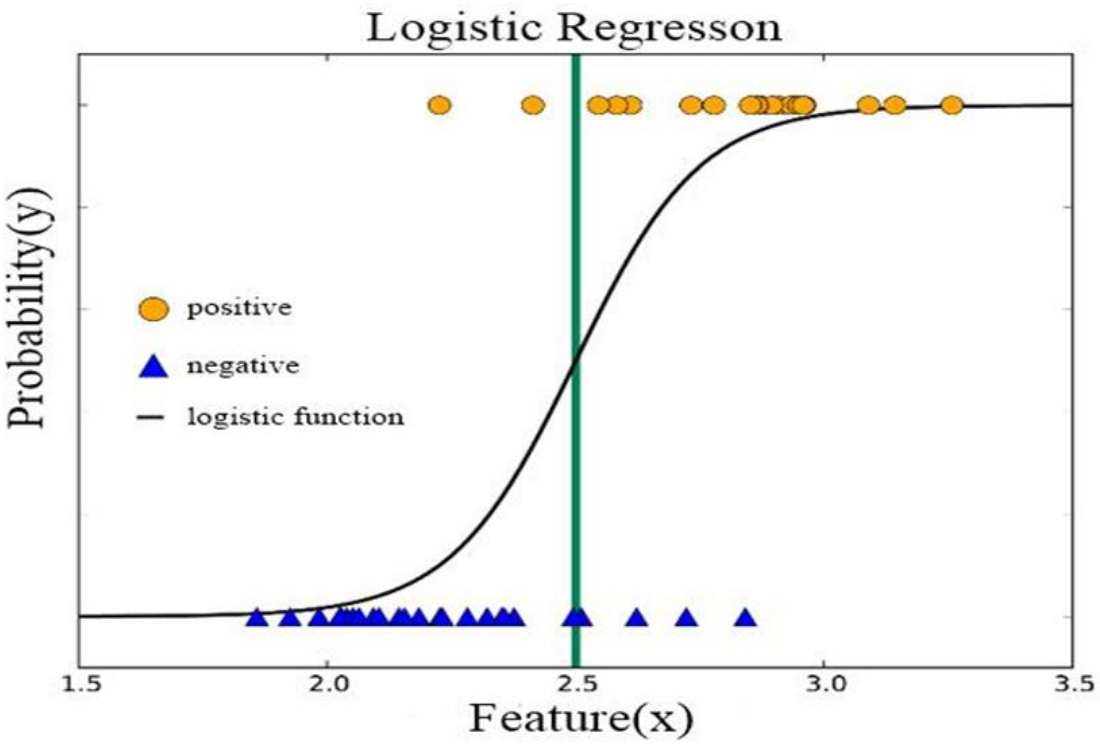


Figure 2.3: Logistic Regression [15]

The main goal of logistic regression is to find the best fitted model to explain the relationship between the dependent variable and a set of independent variable. This regression model is also known as binary logistic model as the outcome is 1 or 0. It allows one to say that the presence of a risk factor increases the odds of a given outcome by a specific factor. To illustrate logistic regression let us look at Figure 2.3. In the given diagram, it is illustrated how to train a 1-dimensional classifier. Here, the training data is divided into two segments, positive (in orange circles) and negative (in blue triangle), where a positive score means a probability of 1 and negative means probability of 0. The Black line represents the decision boundary of the logistic regression which separates the data into two classes. The logistic regression can be understood simply as finding the β parameters that best fit:

$$y = \begin{cases} 1, & if\ \beta_0 + \beta_1 x + \in \\ 0, & otherwise \end{cases} \cdots\cdots\cdots\cdots\cdots\cdots[iii]$$

Where, ∈ is an error distributed by the standard logistic regression. Logistic regression is named at the core of the method logistic function. Logistic function is S-shaped curve that can take any real value number and then, map it to a value that is between 0 and 1. But never exactly the limit values.

2.4.2 Linear Regression

A linear approach that would model the relationship between a scalar dependent variable and one or more independent variables [16]. The independent variables are also known as, explanatory variables. If the number of independent variable is one, then it is known as simple linear regression whereas for multiple number of independent or explanatory variables, the model is called multiple linear regression. It is the most basic form of regression.

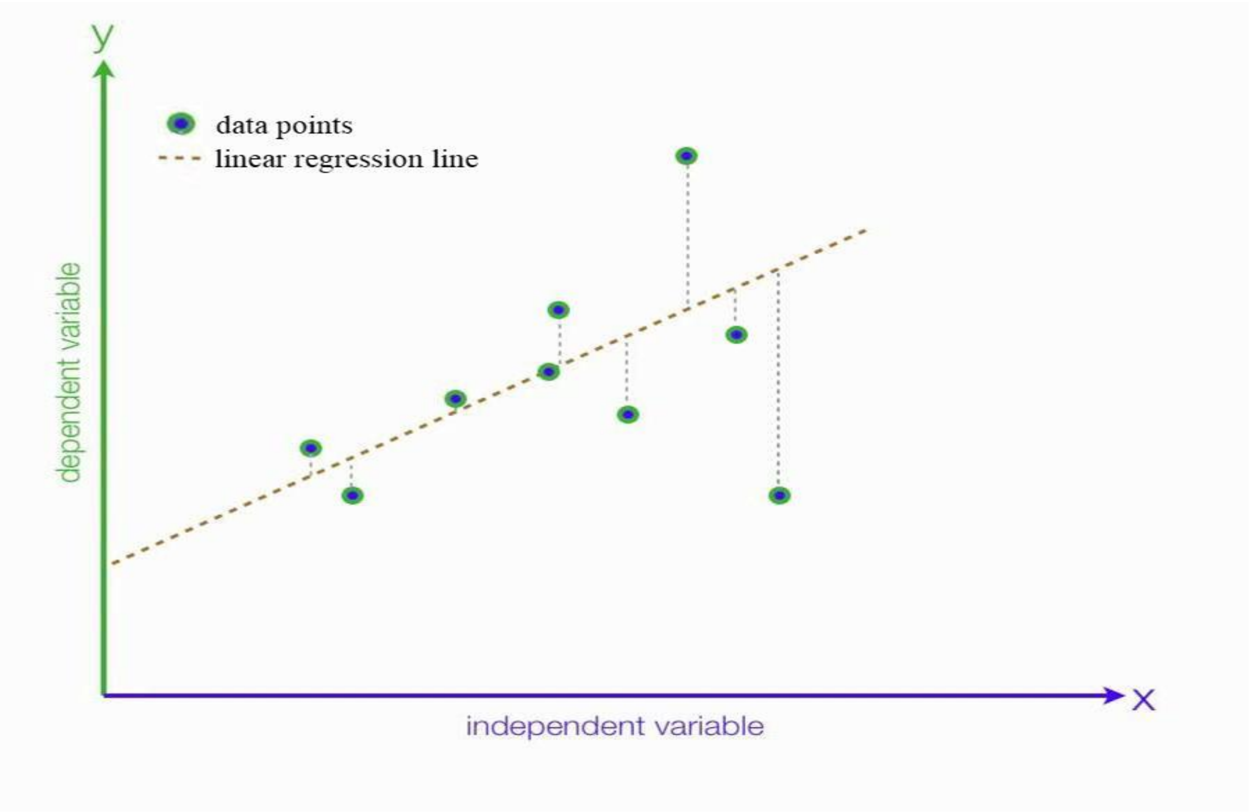


Figure 2.4 Linear Regression [16]

To illustrate simple regression model, let us take a look at the given 2.4 figure, where x denotes the independent variable and y denotes the dependent variable. The green dots on the graph are the scattered plots of the data points for which there would occur changes to the value of dependent variable y based on value of x. The dotted line represents the regression line also known as best fitted line [17]. In simple linear regression the data set would be modelled as:

Y= $\beta_0 + \beta_1 x$……………………………………….…………………...[iv]

Where y is the predicted output and x is the independent variable or explanatory variable. β_0 and β_1 are the coefficient that would move the regression line to be best fitted. As β_0 determines where the line intercepts the y-axis in the graph, it is call the intercept. β_1 Is called the slope as it determines the slope of the regression line.

2.5 Root Mean Squared Error

The root-mean-square deviation (RMSD) or root-mean-square error (RMSE) (or sometimes root-mean-squared error) is a frequently used measure of the differences between values (sample or population values) predicted by a model or an estimator and the values observed. The RMSD represents the square root of the second sample moment of the differences between predicted values and observed values. These individual differences are called residuals when the calculations are performed over the data sample that was used for estimation and are called prediction errors when computed out-of-sample. The RMSD serves to aggregate the magnitudes of the errors in predictions for various times into a single measure of predictive power. RMSD is a measure of accuracy, to compare forecasting errors of different models for a particular dataset and not between datasets, as it is scale-dependent [18].

RMSE = $\sqrt{\frac{\sum_{t=1}^{T}(y_0 - y_t)^2}{T}}$ ……………………………………………….………[v]

The RMSD of predicted values y_0 for times *t* of a regression's dependent variable y_t with variables observed over *T* times, is computed for *T* different predictions as the square root of the mean of the squares of the deviations.

CHAPTER 3

Literature Review

In order to complete the thesis, we had to have a clear idea about Bitcoin and time series analysis. To understand the economic and technological perception of Bitcoin, [19] is really a great journal. They states that Bitcoin depends on two basic technological innovations from cryptography: public-private key cryptography to store and burn through cash and the other one is cryptographic validation of transaction. To know more about Time Series Analysis, [20] and [21] are great resources. In these books, the distinctive themes of time series analysis were clarified in incredible detail with straightforward examples. We learned python programming from online course of Coursera [22].

However, a large number of research has been conducted on the fluctuation of Bitcoin price using Time Series Analysis and machine learning. C. Chatfield and M. Yar predicted Bitcoin price using twitter sentiment data. They worked on Valence Aware Dictionary and sentiment reasoner (VADER) which is a combined lexicon based approach to measure the individual tweet sentiment intensity and then grouped the scores into time-series. They have proposed a prediction model and evaluated aggregating tweet sentiments over a 30 minute period along with 4 shifts forward with a constant change in sentiment resulted a 79% of accuracy [23].

Additionally, Shai Ben-David, in his understanding machine learning book [24], talked about realizability assumption, in which true error for dataset in a particular algorithm should be zero. In order to explain that he stated about Empirical risk minimization (ERM) hypothesis class which means having samples with minimum error assuring minimization of the sample error with unknown distribution and unknown function. But for some cases, ERM does not represent the real world as the hypothesis class is created by choosing training and testing data from the same particular area which leads to data overfitting and poor performance of the algorithm in global applications. In that case, they mentioned PAC (Probability of Approximately correct) learnability and Agnostic PAC learnability in their book. PAC learnability supports the concept of realizability assumption and Agnostic PAC learnability states that true error for a hypothesis class, H is less than the summation of sample error and accuracy parameter. Basically accuracy parameter is a parameter to forgive a learning algorithm for having minor mistakes. They also introduced linear multiclass predictions, then they have explained the use of family of linear predictors even if the number of classes is extremely large with an adequate structure on the problem.

Nevertheless, A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau analyzed sentiment feeds on Twitter Data and introduced POS-specific prior polarity features and explored the use of a tree kernel to remove the need for tedious feature engineering. Their introduced features and tree kernel performed approximately at the same level, both outperforming the state-of-art baseline. [25]. Y. Kim in his article "Convolutional neural networks for sentence classification", proposed a simple one-layer CNN that achieved state-of-the-art results across several data sets. The very strong results achieved with this comparatively simple CNN architecture suggest that it may serve as a drop-in replacement for well-established baseline models, such as SVM [26] or logistic regression. Kim defined a one-layer CNN architecture that uses pre-trained word vectors as inputs, which may be treated as task-specific or static vectors. Efthymios Kouloumpis, Theresa Wilson, Johanna Moore [27] have used Hashtag and Emoticon in tweets to predict sentiment using n-gram, lexicon, parts-of-speech, micro-blogging as features. T. Joachims worked on an unsupervised classification algorithm called PMI_IR (Turney 2001), using the association of adjectives in reviews [28].

Our thesis is mainly inspired by [29]'s work. They detected the determinants of Bitcoin price fluctuation using time series analysis. The created relation between fundamental economic and technical factors with Bitcoin price volatility. They explained that Bitcoin is kind of exchange medium which should be valued based on its supply and demand curve interaction. According to Fisher's equation, MV=PT, M is value of money, V is velocity of money, P is price level and T is the size of underlying economy. The ostensible supply of Bitcoins is given by M=P^BB (where P^B is the cost of Bitcoins and B is the supply of Bitcoins available for use), in this way inferring P^B=PT/VB. In this manner, the balance cost of Bitcoins (i.e. the cost evening out request and supply) ought to be emphatically identified with the general value level (P) and the measure of the Bitcoin economy (T) however contrarily identified with the aggregate load of Bitcoins available for use (B).

CHAPTER 4

Forecasting methodology

In this section, we will talk about the time Series analysis approach that we have utilized as a part of our venture. We attempt to outline what we have gained from the books, address notes and other online materials specified in literature survey section of this paper. With a view to clarifying the themes, we have utilized the dataset on the yearly breadth of BITCOIN from 2013 to 2017 collected from Kaggle [33]. We utilized the python programming as our apparatus to play out the examination on the informational collection. We will discuss about Python in detail later on.

We are trying to predict the Closing price of Bitcoin, and so we will create a new object eliminating other columns as follows:

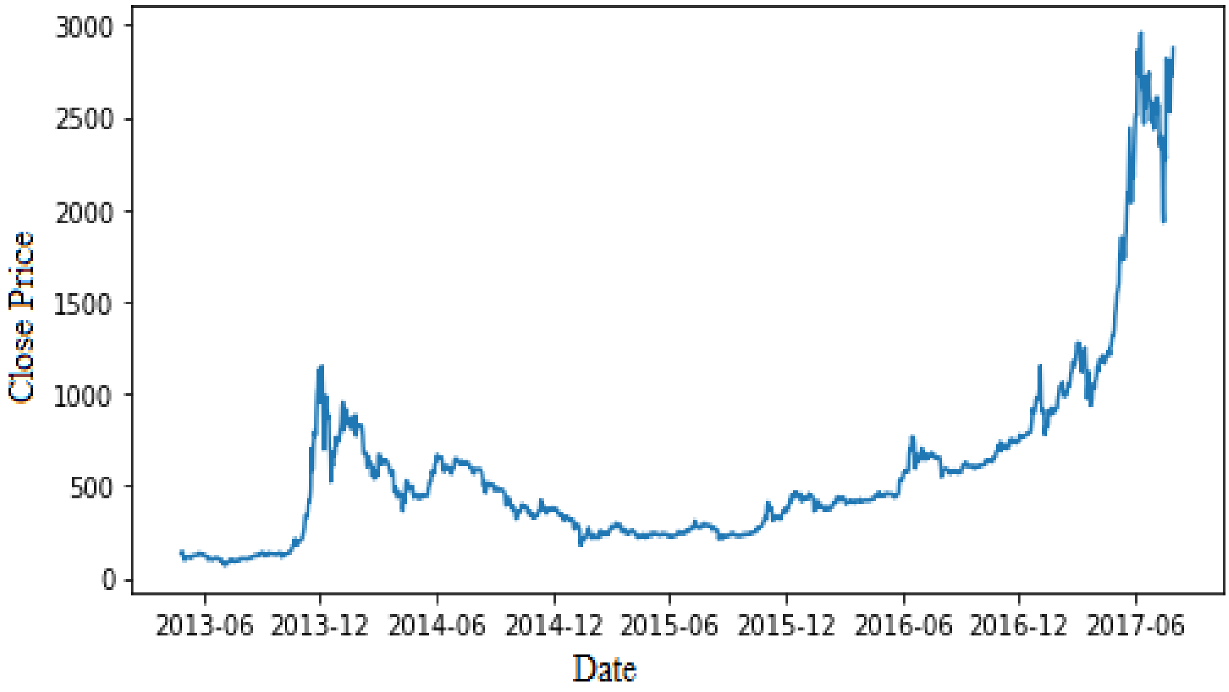


Figure 4.1: Bitcoin close price

In Figure 4.1, using date as index the series is plotted with Date on x axis and closing price on y axis. Here we can see that from June, 2013 to December, 2013 value Bitcoin price was increasing, on the other hand, from 2014 to 2015 it was decreasing gradually. In 2017 the demand of Bitcoin was very high, therefore the graph went maximum in July, 2017. This is the proof of the volatility of the close price of Bitcoin which can be visualized by the scattered graph.

4.1 Time Series Analysis

For time series analysis, firstly we imported all the python library such as numPy, pandas, matplotlib, prophet etc that we need to process the data and evaluated the result. After reading the input data from CSV file, we started with Rolling mean and Standard deviation.

- **Rolling mean and Standard deviation**

By plotting rolling statistics, we could plot the moving average or moving variance and see if it varied with time.

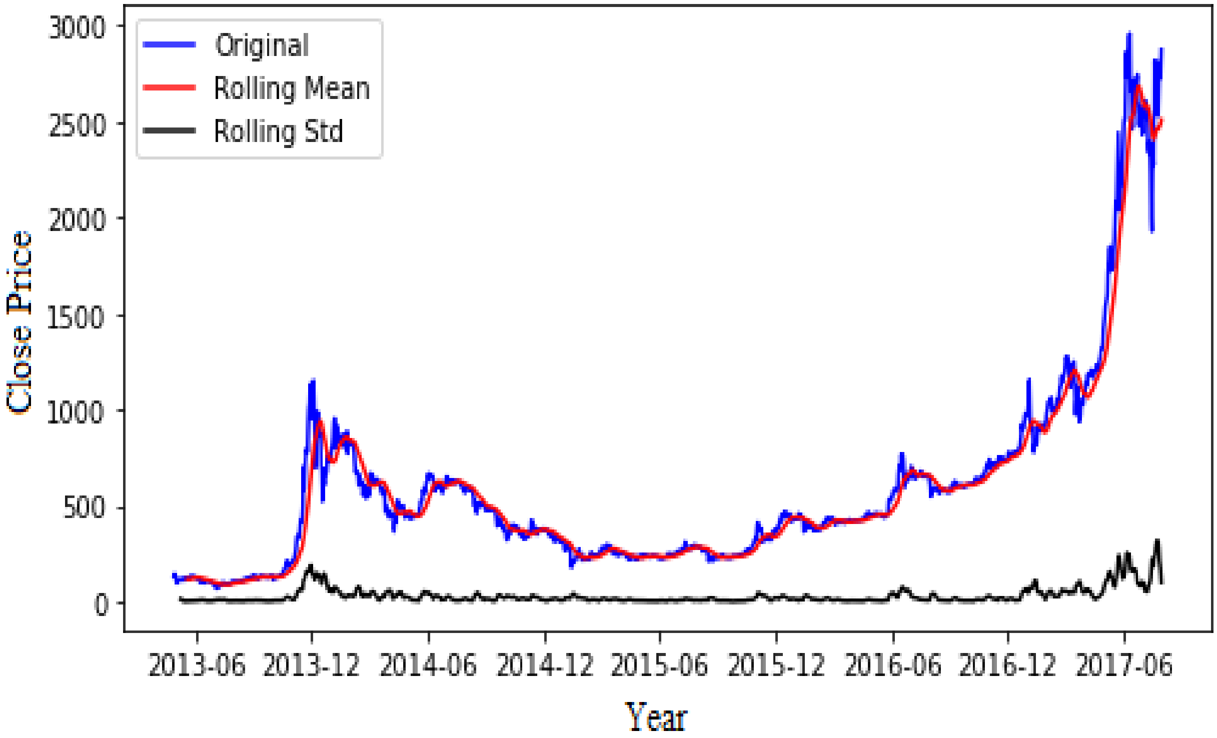


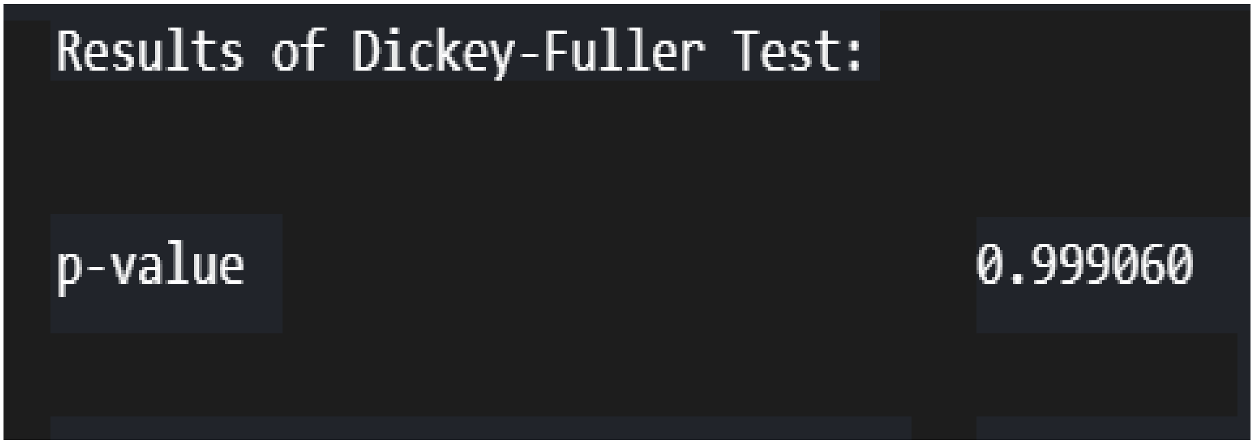
Figure 4.2: Rolling mean and Standard deviation

By moving average/variance we can say that at any instant time ‘t’, we will take the average or variance of the last year, i.e. last 12 months. But again this is more of a visual technique. The

Figure 4.2 shows the relation between rolling mean and rolling standard deviation where x axis and y axis denote date and weighted price of Bitcoin respectively. Blue line shows the original value, red line is for rolling mean and black line indicates rolling standard. However, the test statistic is very large compared to the critical value. The above series shows that rolling mean value varies with time which means data is not stationary. Therefore, from the figure we can assume that initially we have a large difference between rolling mean and rolling standard which should be lessened in order to increase stationarity. Because one the most important theories of time series analysis are dataset have to be stationary.

Dickey-Fuller Test

The Augmented Dickey Fuller test is a sort of factual test called a unit root test. The instinct behind a unit root test is that it decides how unequivocally a time series is characterized by a pattern. There are no. of unit root tests and ADF (Augmented Dickey-Fuller Test) is a standout amongst the most generally utilized.



```
Results of Dickey-Fuller Test:

p-value                          0.999060
```

Figure 4.3: Dickey-Fuller Test Result

From the figure 4.3, interpretation of p value is [30]

- **p value** > **0.05:** Accepts the Null Hypothesis (H0), the data has a unit root and is non-stationary.
- **p value** < = **0.05:** Rejects the Null Hypothesis (H0), the data is stationary.

Therefore, we can assume that the data we have used is not stationary as value of p is greater than 0.05. But to apply time series analysis we need to make it stationary. In order to do that, we need to eliminate the trend and seasonality from the series.

- **Eliminating Trend**

One of the primary traps to decrease trend can be Transformation. For instance, for this situation we can obviously observe that there is a critical positive trend. So we can apply transformation which punish higher qualities more than smaller qualities. These can be taking a log, square root, solid shape root, and so forth. We used Log transformation and differencing methods in order to deleting the trend for simplicity.

- **Log Transformation**

We transformed the data to remove the nonlinearity. We used both log transformation and differencing for getting better result. Log-transformed data follows a normal or near normal distribution. In this case, Logarithmic (LOG) transformation is used to unskew highly skewed data.

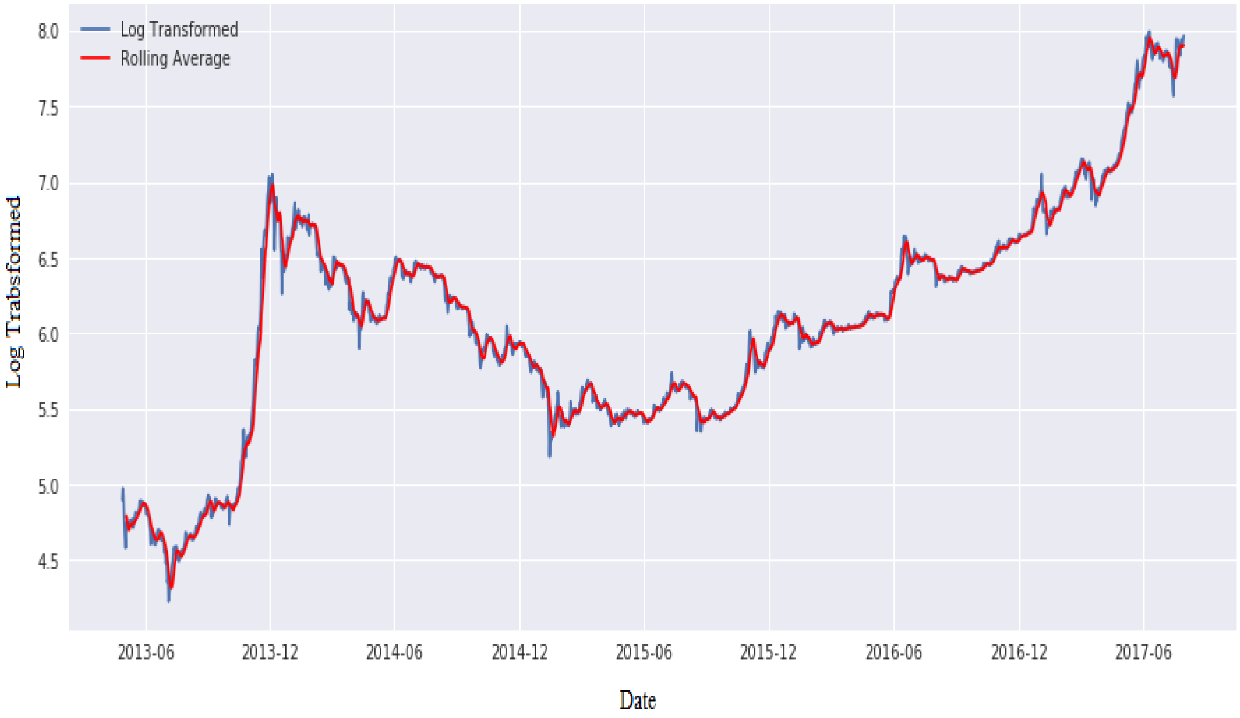
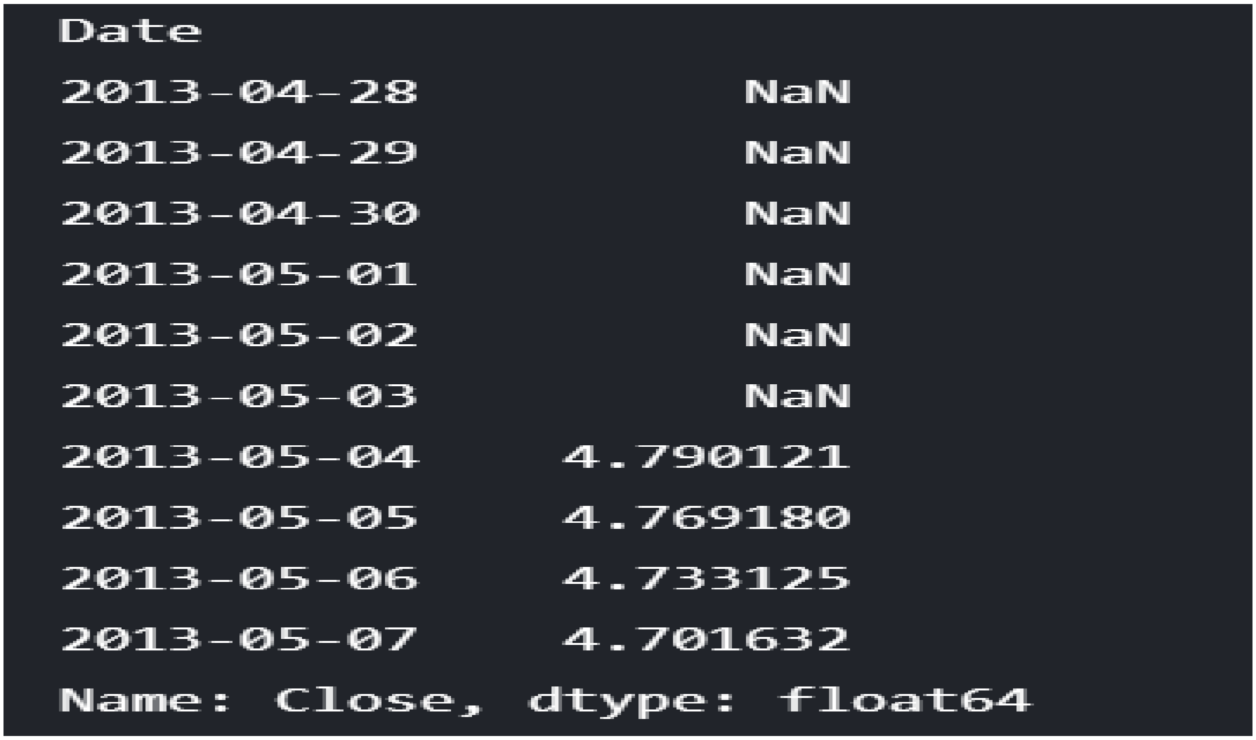


Figure 4.4: Transforming data for trend deletion

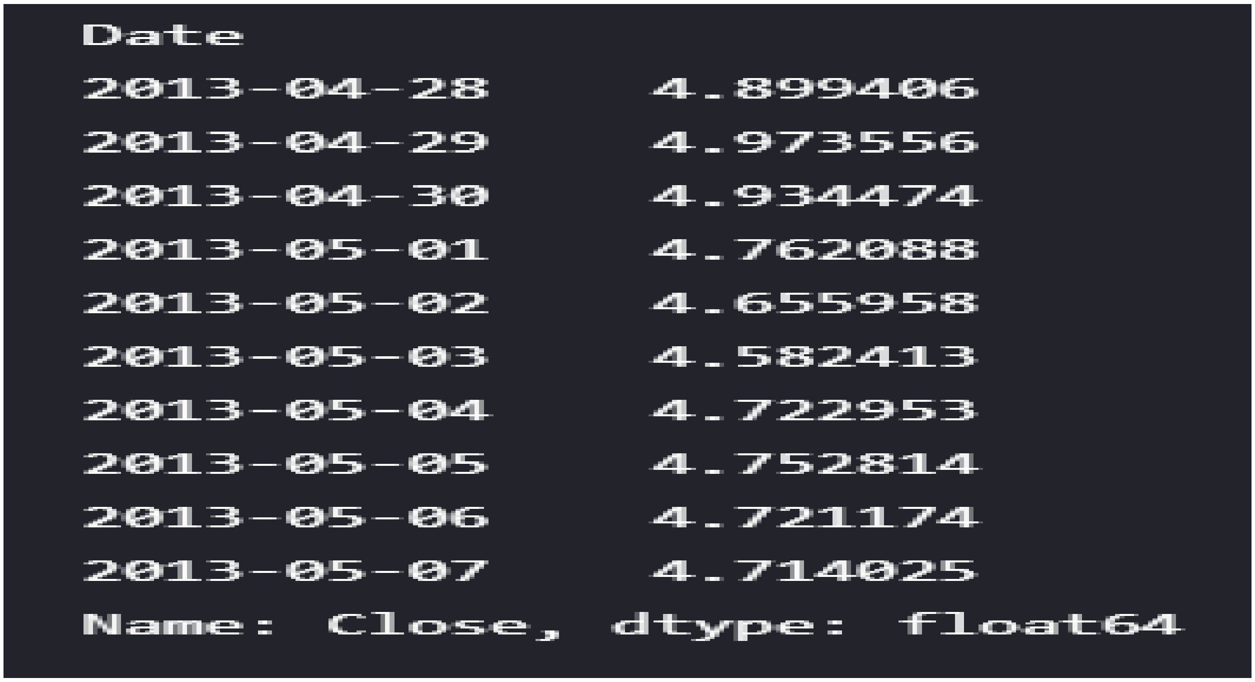
Figure 4.4 shows the graph of Log transformed price of Bitcoin which deletes the trend from dataset. Here, graph of Bitcoin price has been upper bounded and the prices are converted to lower values. By decreasing the output values, we can decrease the difference which increases stationarity. Here, red line and blue line represents rolling average and log transformation respectively. Since we are taking average of last 7 values, rolling mean is not defined for first 7 values. This can be observed as Figure 4.5.



```
Date
2013-04-28          NaN
2013-04-29          NaN
2013-04-30          NaN
2013-05-01          NaN
2013-05-02          NaN
2013-05-03          NaN
2013-05-04     4.790121
2013-05-05     4.769180
2013-05-06     4.733125
2013-05-07     4.701632
Name: Close, dtype: float64
```

Figure 4.5: With Missing Data

To solve this problem we subtract this Rolling average from the original log transformed series which is shown in Figure 4.6. This is mainly used for replacing NAN by average values in the data frame which is initially generated for missing data (shown in figure 4.5) to avoid any errors in the future. Then we get a log transformed graph like Figure 4.7. The figure shows Log transformed price on Y axis and Date in x axis which is taken from Figure 4.6. As we have eliminated the missing data problem, fluctuation between actual price and log transformed price has been decreased. Thus helping in forecasting process but after dickey fuller test the value of p was still greater than 0.05. Therefore we decided to use differencing for adjusting the seasonality and improving stationarity



```
Date
2013-04-28      4.899406
2013-04-29      4.973556
2013-04-30      4.934474
2013-05-01      4.762088
2013-05-02      4.655958
2013-05-03      4.582413
2013-05-04      4.722953
2013-05-05      4.752814
2013-05-06      4.721174
2013-05-07      4.714025
Name: Close, dtype: float64
```

Figure 4.6: Data after processing (resolving missing data)

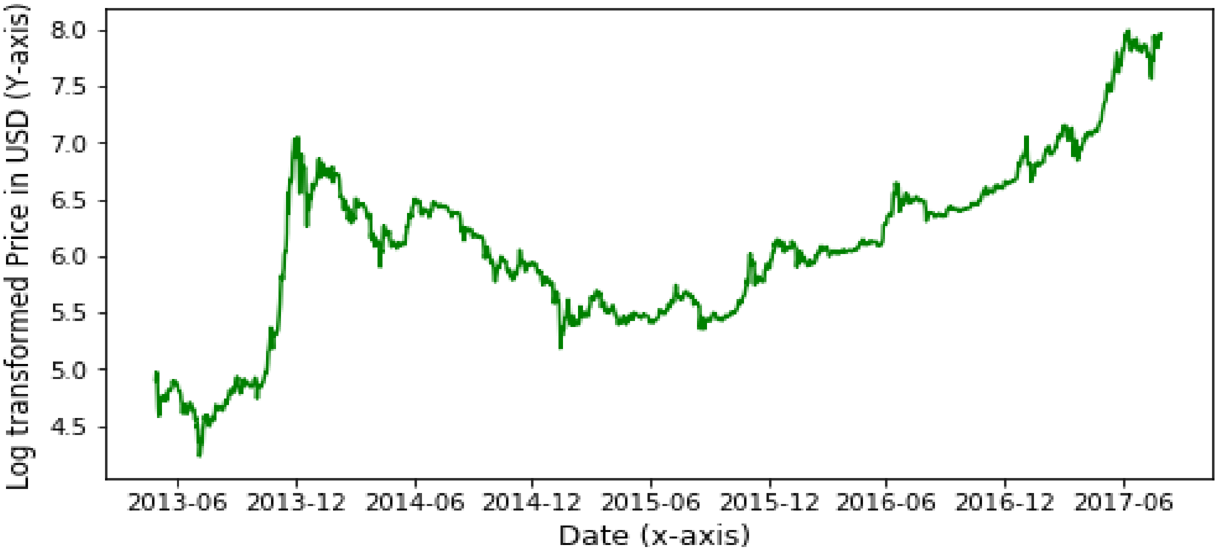


Figure 4.7: Log transformed graph After Processing

- **Differencing**

Unfortunately, data arising from many studies do not approximate the log-normal distribution so applying this transformation does not reduce the skewness of the distribution, so we used differencing. Differencing is performed by subtracting the previous observation from the current observation. Seasonal segments in a Time series can be eliminated by utilizing differencing. In that case, if there is a regular part at level of multi month, at that point it can be expelled on a perception today by subtracting the incentive from a month ago.

- **Cycle Observation**

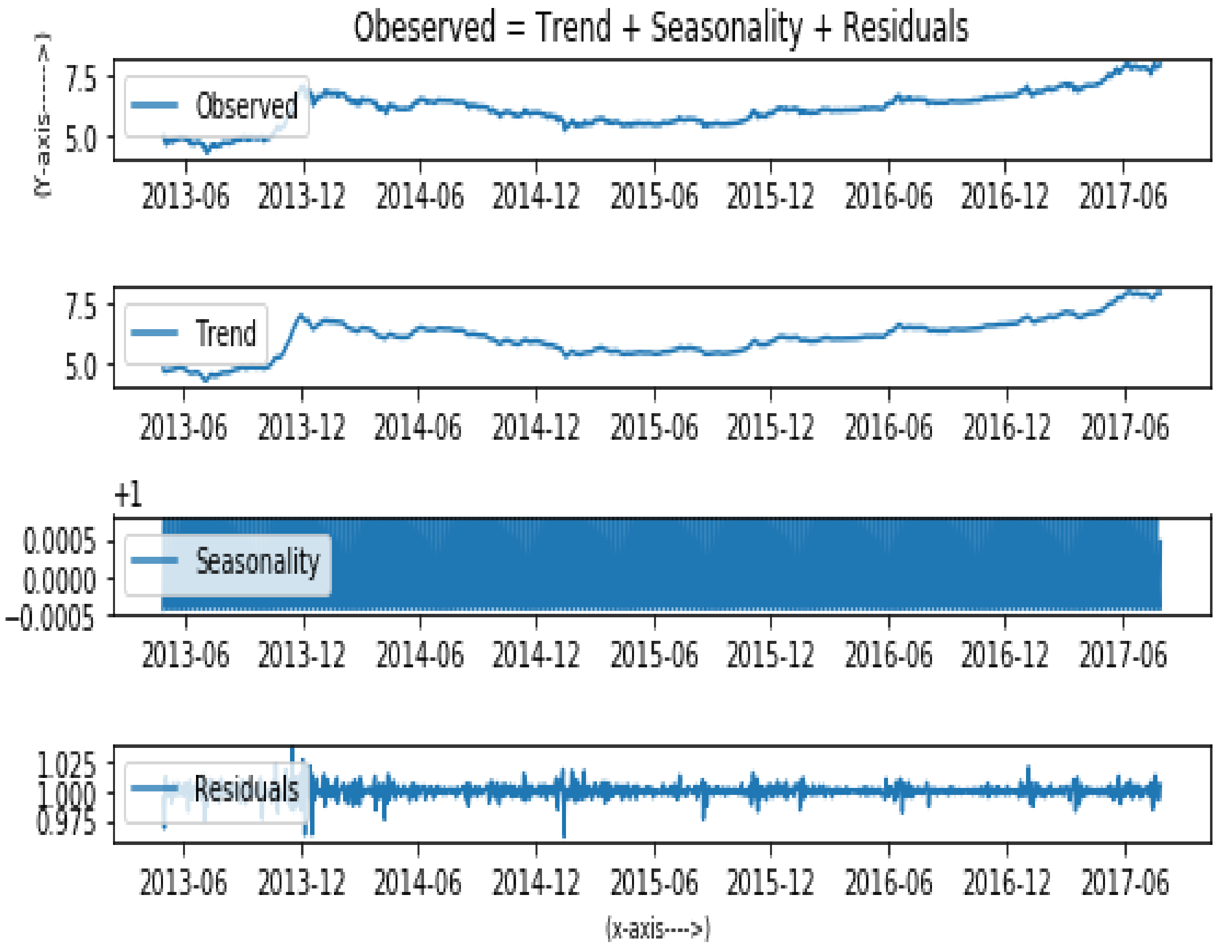


Figure 4.8: Cycle Observation

After applying log transformation and differencing, we observed the data based on its seasonality, trend and residuals which is shown in Figure 4.8. The observed graph is basically the combined result of trend, seasonality and residuals. Trend, seasonality and residuals were estimated for every 6 month data from 2013 to 2017. These methods basically detect the pattern which repeats in a certain period of time. Trend was estimated in a range of 5 to 7.5, seasonality's range was -.0005 to +.0005. Similarly, residuals was measured with a scale of .975 to 1.025. Finishing these steps, we were assured to use this data for Time series model.

- **Autocorrelation**

Autocorrelation is a measure of the inside connection inside a time series. It is a method for estimating and clarifying interior relationship between perceptions in a period series [31].

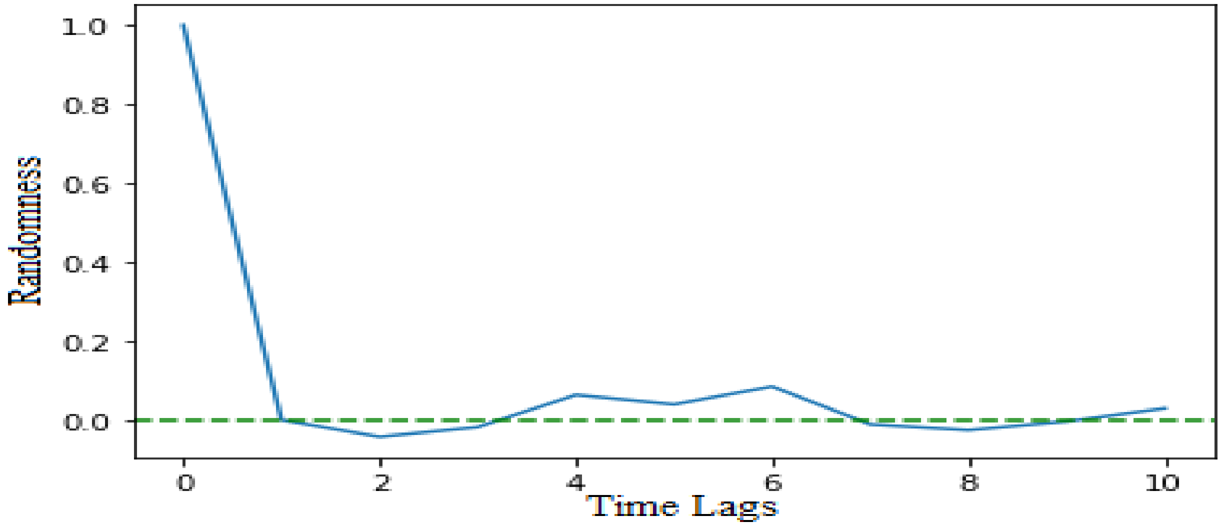


Figure 4.9: Autocorrelation

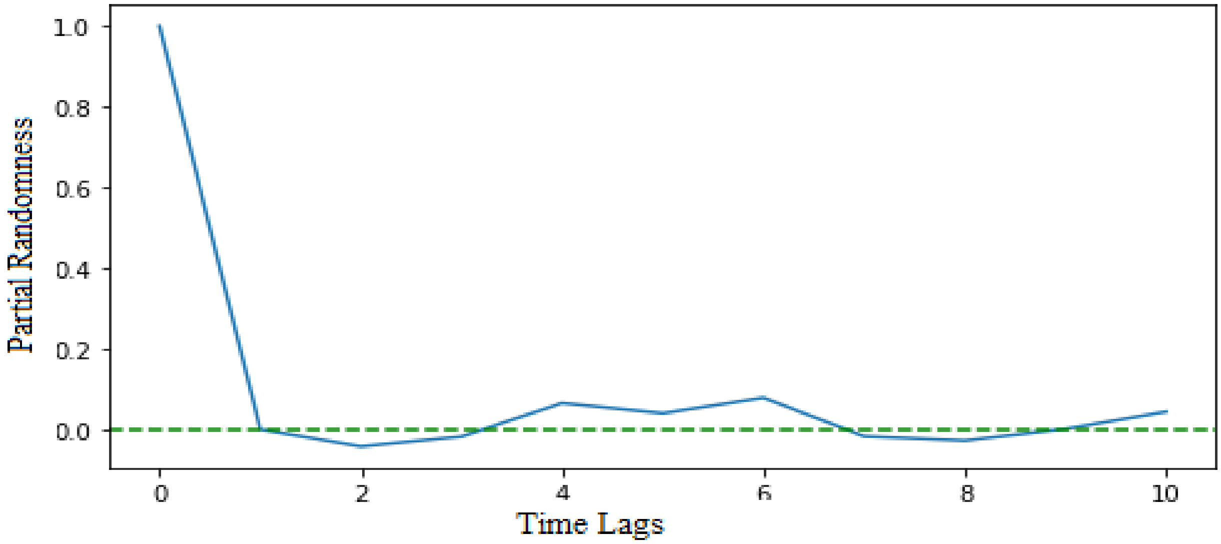


Figure 4.10: Partial Autocorrelation

Autocorrelation is formed by autocorrelation coefficient [31]

$$R_h = C_h / C_0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots [vi]$$

Where C_h is the autocovariance function, C_0 is the variance function and R_h is between minus 1 to plus 1.

It is used for checking randomness in a data set. This randomness is ascertained by computing autocorrelations for data values at varying time lags. Basically it finds out the relation between features and the output. Here, our Bitcoin features were market capital, volume in BTC and volume in USD price. Figure 4.9 shows the correlation of features and time lags of our data set. Y axis is the measurement of the randomness and x axis represents the time lags. Additionaly, Figure 4.10 represents the partial randomness between features and output. Partial randomness is needed to know in order to selecting a model.

- **Model Selection**

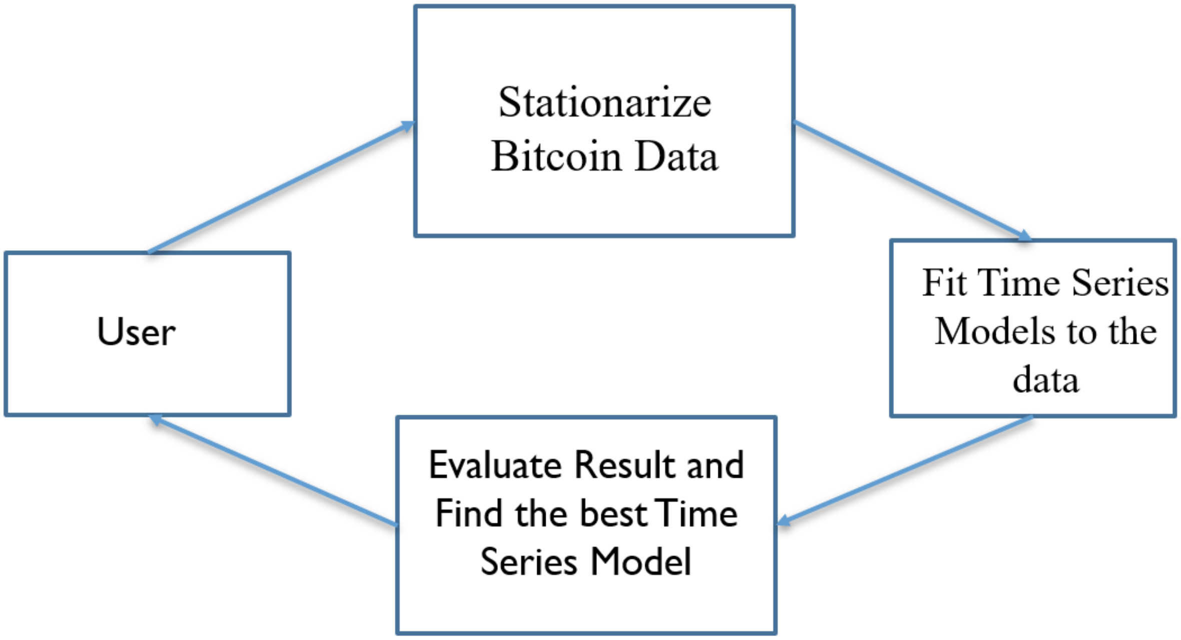

Stationarize Bitcoin Data
Fit Time Series Models to the data
Evaluate Result and Find the best Time Series Model
User

Figure 4.11: Data Flow of Time Series Model

After processing the data, we applied time series models Autoregressive integrated moving average model (ARIMA), Autoregressive model (AR) and Moving average model (MA) in our dataset and plotted the resultant graph which will be discussed in the next Chapter 5. Based on the accuracy of the models, we chose ARIMA to predict Bitcoin price as our data fitted well in it. Figure 4.11 indicates the data flow of a time series model. Historical data will be processed in test and train set. Test set will be stored in time series actual data and train set will be fit in ARIMA model with all available features in the dataset and then our model will send the result of test data to user. User will compare the trained result with the actual test data that was previously stored in database. Our course of work will be to find the best model for price prediction of Bitcoin in the first place. We will collect data from reliable source and then process the data according to our model with the best performance. After that, we will input the data to our model and analyze the historical price of Bitcoin. We will then try to correlate the day wise closing price with other features that were given in the dataset and find out the pattern of forecasting method. After forecasting the price of next days we will output the value through python ide to user.

4.2 Machine Learning

Though sometimes time series analysis is said to be a part of machine learning, we considered it as a different section. We trained our data and evaluated the test data with two machine learning algorithms.

- **Linear Regression**

While applying linear regression model in our data set, the challenge was to find out the correlation between price and the features of our dataset as it was hard to find linear relation with the price volatility and its features. We considered date, market capital and volume as our main features for this model. Then we used ML Studio [32] in order to correlate and predicting the price of Bitcoin.

- **Logistic Regression**

In logistic regression, instead of modeling the probability as a linear function, we model it using the logistic function. Therefore, it gives better accuracy than a linear regression when the prediction is not based on any binary decision. Basically, Figure 4.12 summarizes the whole process of applying machine learning in Bitcoin price prediction. Firstly, we imported CSV file from computer and then converted it to dataset. After that, we selected the features that we want to use to predict the weighted price and then split the data into test and train model. We trained the dataset with Machine learning algorithms and then evaluated the result with the test model and found out the accuracy.

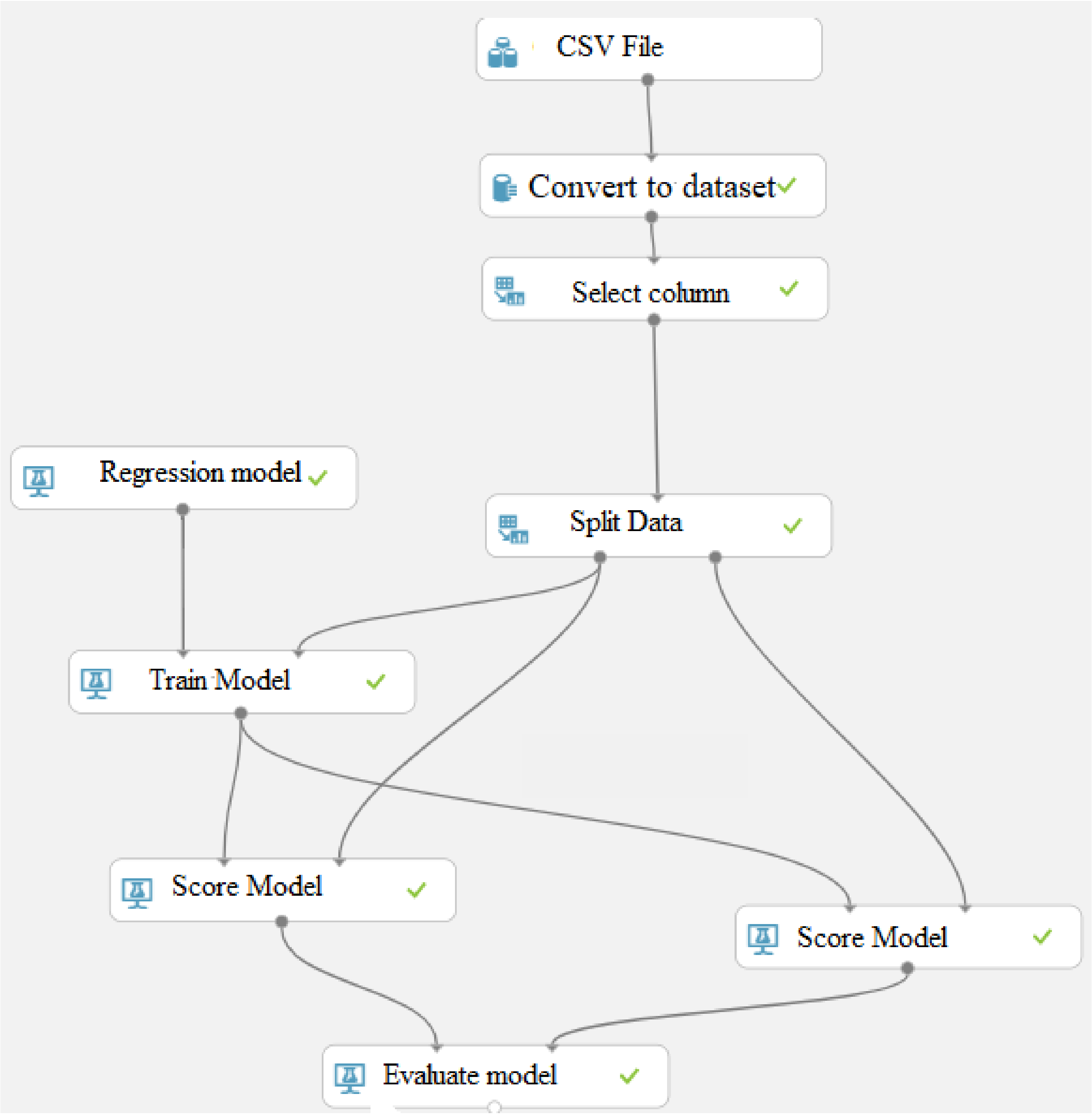


Figure 4.12: Work Flow of Machine Learning

4.3 Proposed Model

Figure 4.13 shows how our model will be implemented. User gives input, then our Python code in anaconda extracts the features from the dataset and executes several functions on them to generate the outputs. The outputs are then saved in a directory of the Computer by Python, which are then fetched and shown to the users on the kernel.

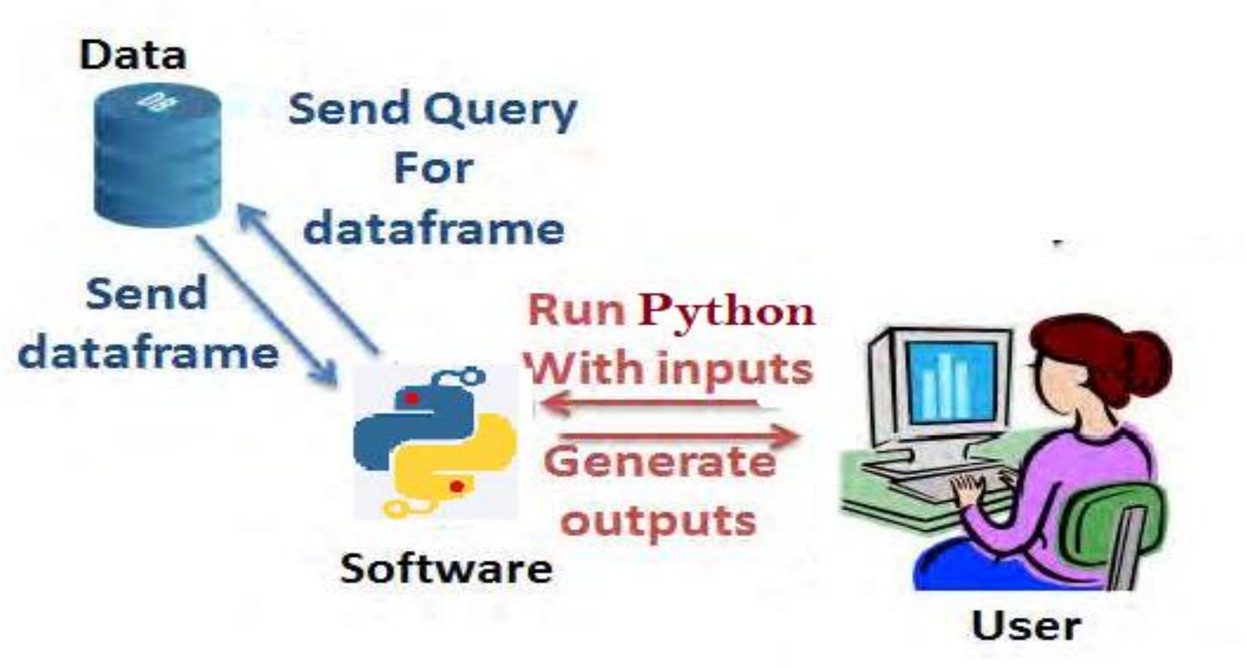


Figure 4.13: System Diagram

4.4 Dataset

In our dataset of Figure 4.14, we currently have the historical daily price from 28th April, 2013 to 31st July, 2017. This dataset has the following features:

- **Date :** date of observation
- **Open :** Opening price on the given day
- **High :** Highest price on the given day
- **Low :** Lowest price on the given day
- **Close :** Closing price on the given day
- **Volume :** Volume of transactions on the given day
- **Market Cap:** Market capitalization in USD

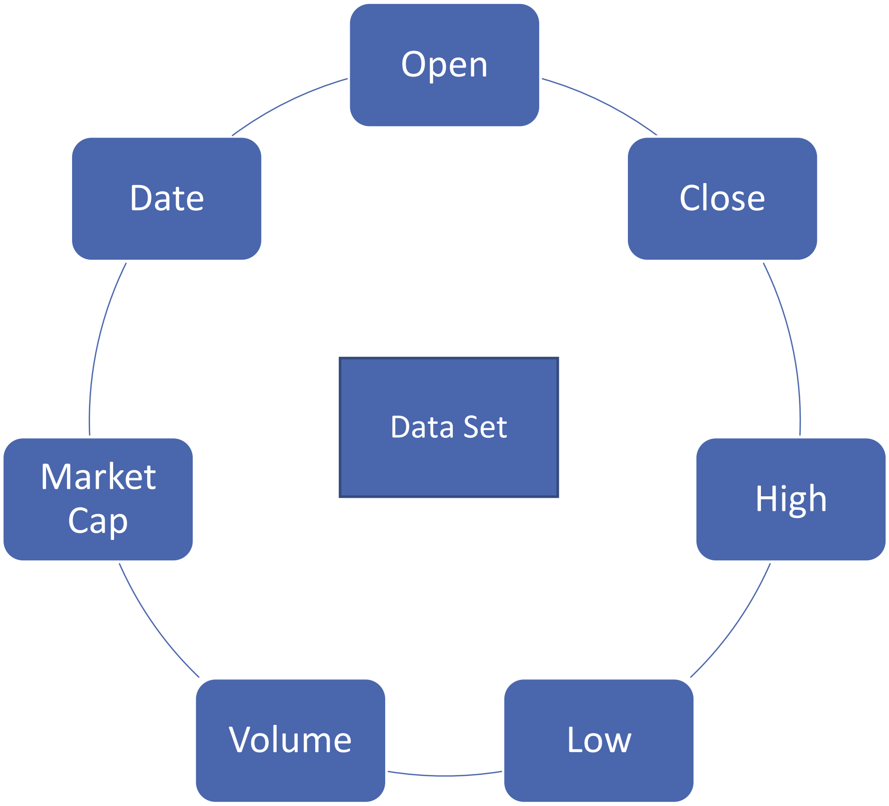


Figure 4.14: Features of Dataset

Dataset was collected from Kaggle and coindesk. [33].

4.5 Python

Python is a programming language and software environment which is widely used for statistical computations. It is not only used by statisticians to perform data analysis, but it also helps in developing statistical software. To learn the Python language for our project, we used the Spyder software which is an IDE for using Python.

- **Data Frames**

Data frames in Python are used to store data tables. A data frame is actually a list of vectors of equal length. In our project, we used Python to collect data from user directory, to store the data in the database and to extract required data from database. All these were done by the use of data frames. The data from user were collected as data csv file and then converted to data frame. We then appended the data from the data frame to the existing table in kernel using Python.

- **Pandas**

Pandas is a Python module that contains high-level data structures and tools designed for fast and easy data analysis operations. Pandas is built on NumPy and makes it easy to use in NumPy-centric applications, such as data structures with labelled axes. Explicit data alignment prevents common errors that result from misaligned data coming in from different sources [34]. We used Pandas for reading datasets and converting it to data frame [35].

- **NumPy**

NumPy is an open source extension module for Python. It provides fast precompiled functions for numerical routines. It is very easy to work with large multidimensional arrays and matrices using NumPy [36]. We used Numpy in our code for mathematical calculations such as sqrt, exp and so on.

- **MatPlotlib**

Matplotlib is a low-level library for making two-dimensional outlines and charts [36]. With its assistance, you can assemble assorted outlines, from histograms and scatter plots to non-Cartesian directions diagrams. Additionally, numerous mainstream plotting libraries are intended to work in conjunction with matplotlib. We used it for visualization purpose.

CHAPTER 5

Experimental Result

Evaluation enables us to test the model against the information that has never been utilized for the training. This metric enables us to perceive how the model perform against information that has not yet observed. This is intended to speak to how the model could perform in this present reality. When we are finished with Evaluation it is conceivable that we might want to check whether we can additionally endorse our preparation in any capacity. We can do this by tuning a few parameters. One case of parameter we can tune is how frequently we can go through the preparation set amid preparing demonstrating the information different circumstances. By doing that we will possibly prompt higher exactness. Another parameter characterizes how far we move the line amid each progression in light of the data from the past training step. These qualities assume the part of how exact and idealize the model can move toward becoming and to what extent the preparation takes. For more troublesome models starting conditions can assume an urgent part also in deciding the aftereffect of the training session. We can likewise observe the distinctions relying upon whether a model begins off preparing with beginning qualities 0's versus some dissemination of those qualities. As we can see there are many consideration in the phase of training and we will determine what makes a model good for us.

We have tried to train several different models and compare their results. We used root mean squared method for calculating the accuracy. We did not mention any other accuracy calculating method because we got less error in RMSE than others. These results were obtained using the following hardware: 4-core CPU, 16 GB RAM and by training each model ten times with different random states. We have tried Autoregressive integrated moving average model (ARIMA), Autoregressive model (AR), Moving Average model (MA), Linear regression and Logistic regression. As we can see, the best result is obtained by using the Autoregressive integrated moving average model (ARIMA). Nevertheless, linear regression shows the worst results both in performance and training time.

Firstly, we divided the data into two parts. 80% for training and 20% for testing randomly. We trained our model using these 80% random data and then our models predicted with maximum error of 50%.

A bar chart for comparison is as follows:

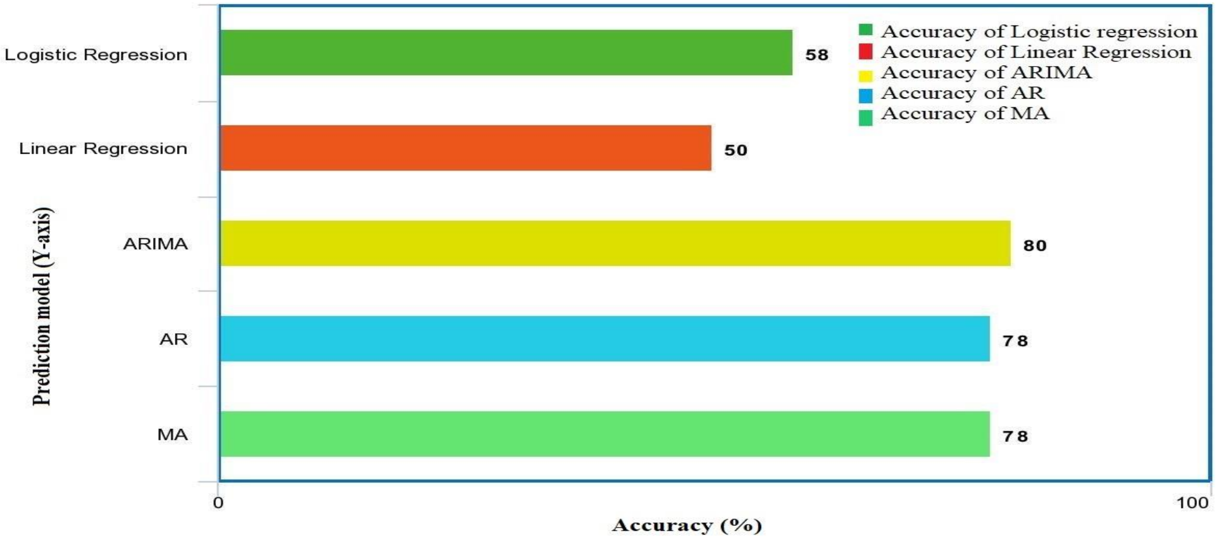


Figure 5.1: Bar charts representing differences among different models

To improve our accuracy we tried to predict the next 10 days value rather than finding the random value from the trained dataset and we find a slight boost in performance.

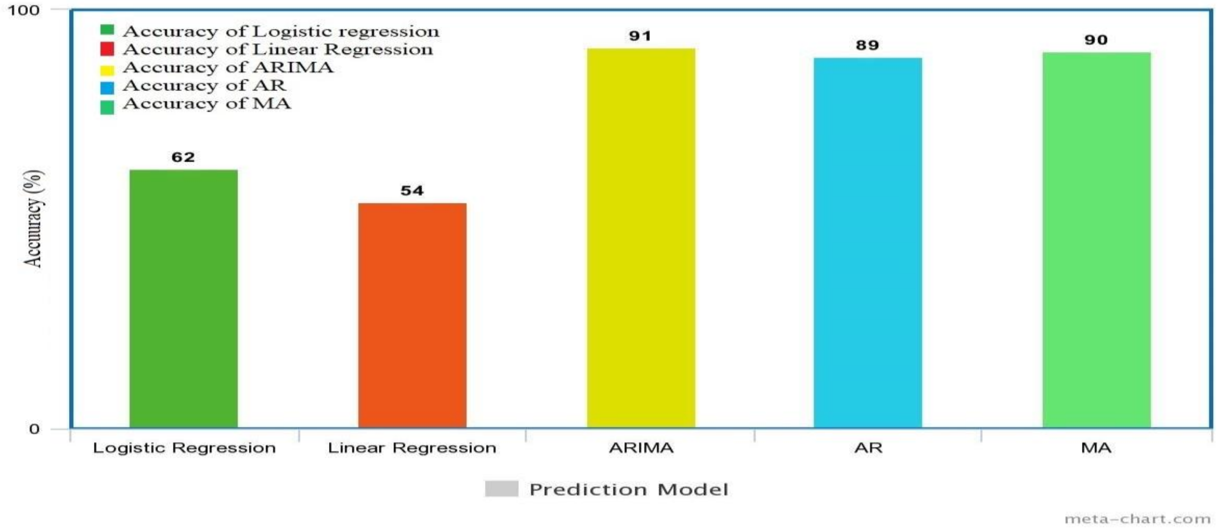


Figure 5.2: Bar charts representing differences among different prediction models after predicting next 10 days data

Here, green represents logistic regression, red is for linear regression, yellow indicates ARIMA model and lastly blue and light green represent AR and MA model respectively for both Figure of 5.1 and 5.2. Prediction model and Accuracy level are represented by x and y axis. From these figures, we can notice that AR and MA model gives almost same accuracy as ARIMA though these models do not have the integrated portion. ARIMA gives the best performance because our dataset fitted well in this model. Linear regression performed very poor because price and date do not have a linear correlation. The fluctuation is scattered. Therefore, scaling the data to linear reduced its accuracy.

Table 5.1 shows a summary of our work that we have done in our thesis project.

Table 5.1: Comparison of different algorithms

Prediction Model	Accuracy
Logistic Regression	62%
Linear Regression	54%
ARIMA	91%
AR	89%
MA	90%

Figure 5.3 predicts Bitcoin price by using ARIMA model which gave the best result among all the algorithms we applied. Here, x –axis represents the year and y-axis represents the weighted price of Bitcoin. Blue line is for actual weighted price and red indicates the predicted Bitcoin price using ARIMA model. Basically we got this output by splitting the data set randomly into train and test part. We have calculated the difference between actual value and predicted value by converting the each predicted values from pandas to dataframe and then calling the root mean squared function, we got the accuracy of this model which was around 80%.

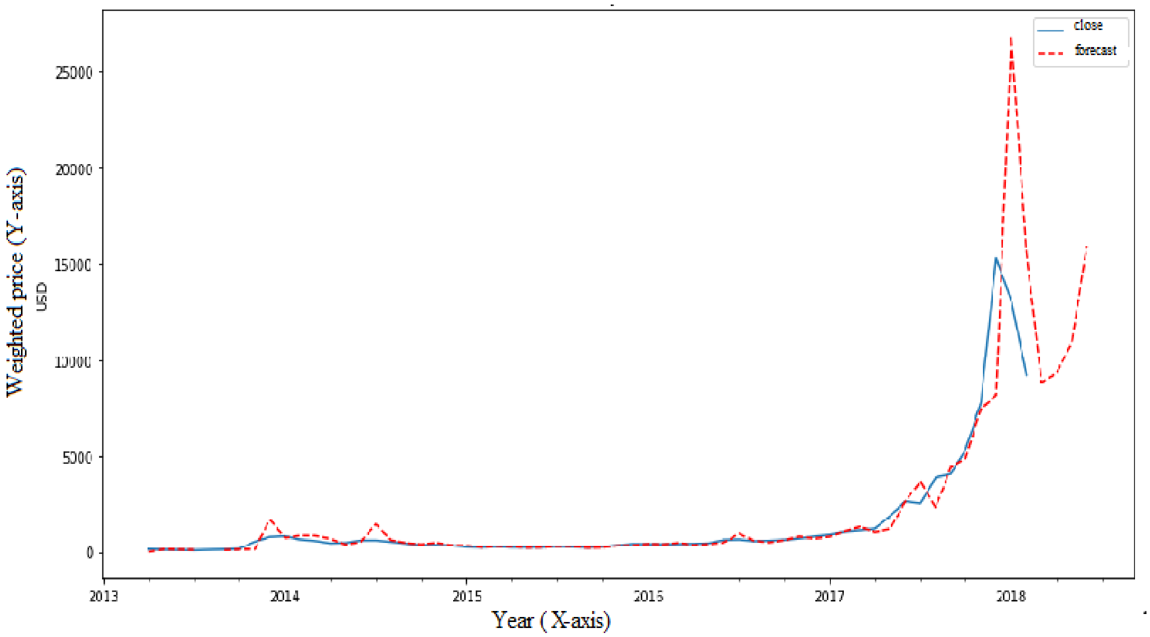


Figure 5.3. Prediction using ARIMA

Figure 5.4 shows the graph of MA model which is calculated by month. We had applied AR model as well, which had given us very similar return to MA model with good risk assessment. We had the output for both graphs and till 2017 both performed very similarly.

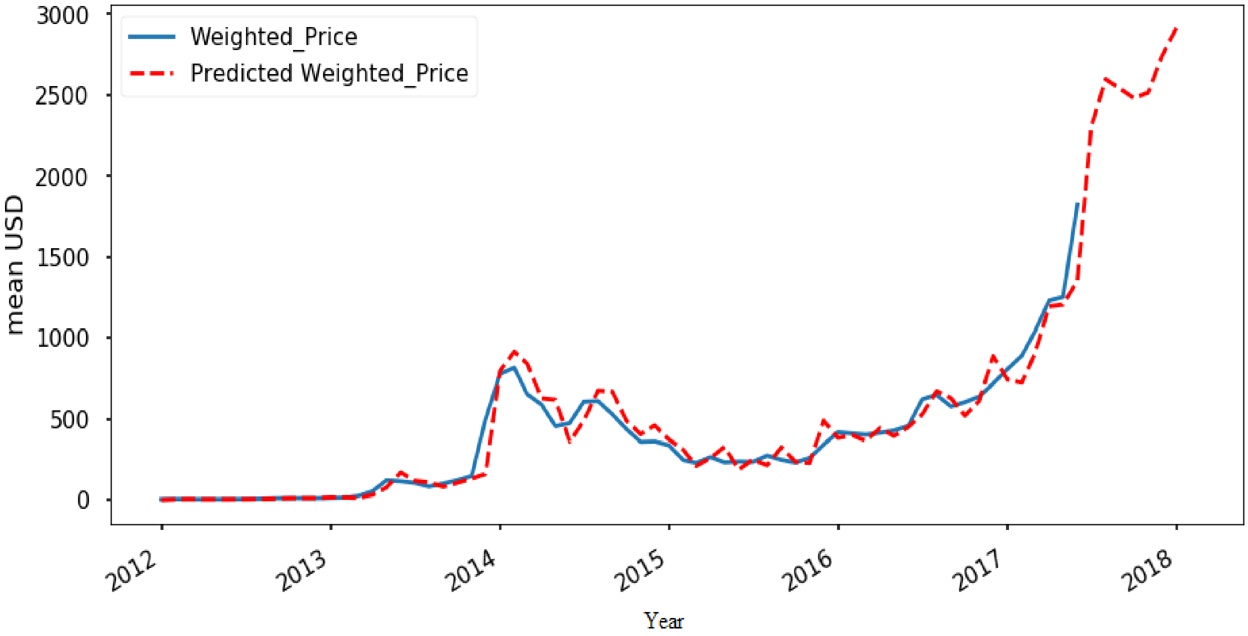


Figure 5.4: Prediction using MA

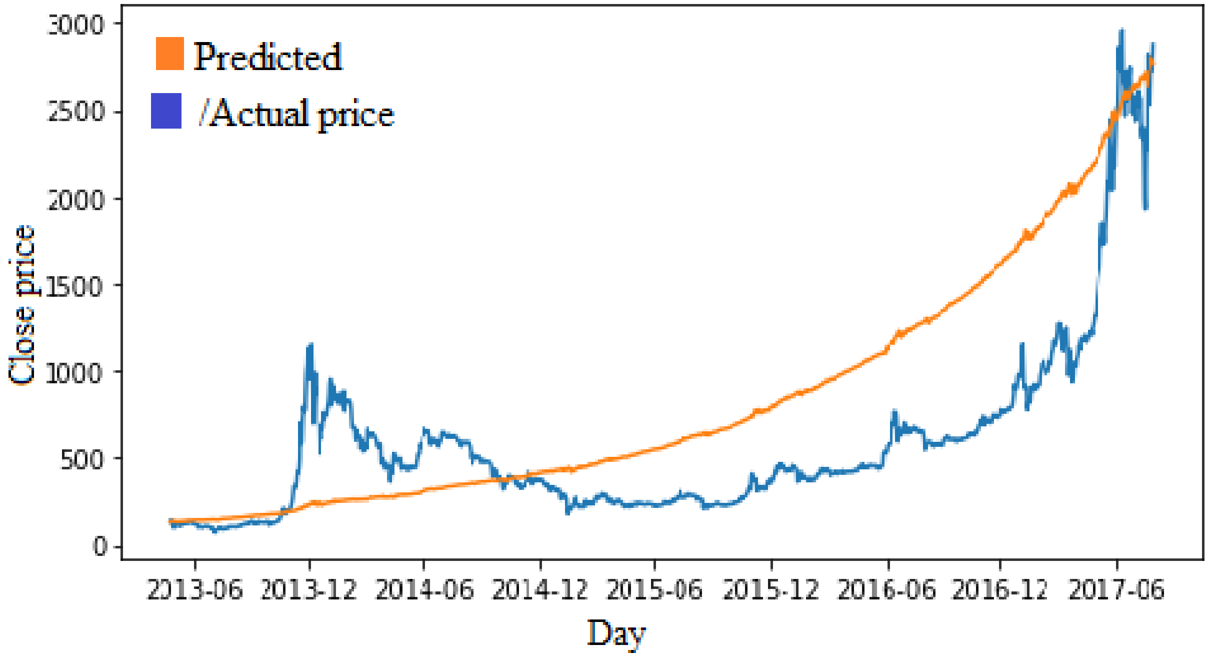


Figure 5.5: Prediction using Logistic Regression

Figure 5.5 illustrates the prediction graph of Logistic regression from July, 2013 to July, 2017. Here, red line indicates the predicted value and, blue line represents actual price of Bitcoin. As we could not have a good correlation with the price and the other features, the result was not up to the mark.

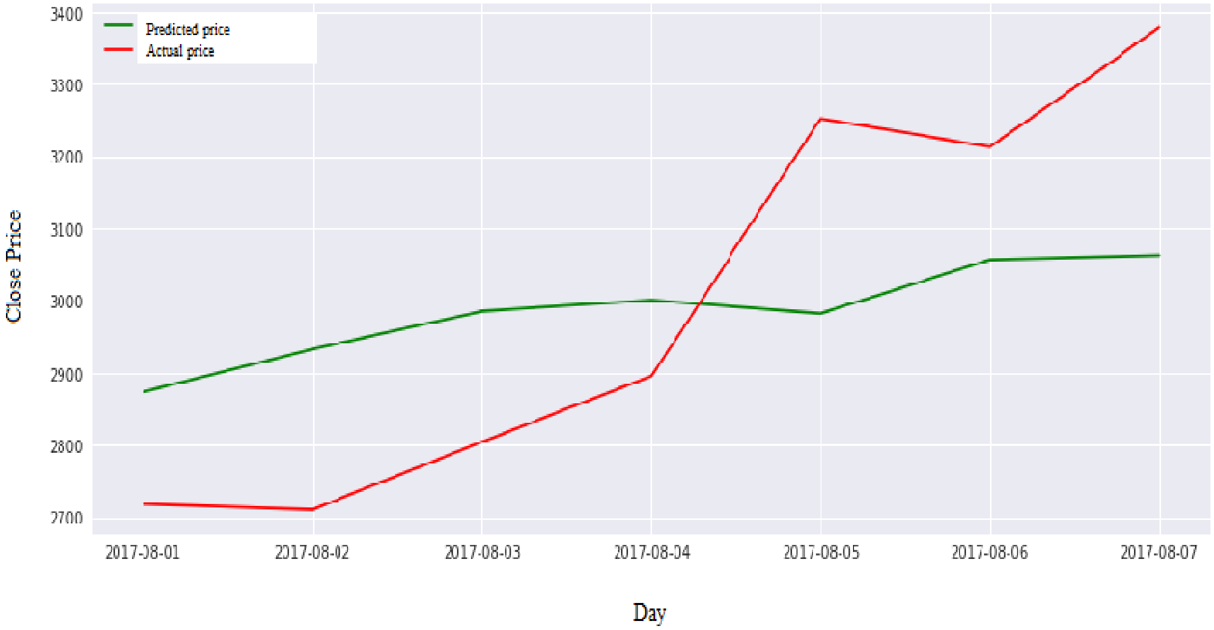


Figure 5.6: Prediction using Linear Regression

We found out that linear regression model works better for Bitcoin day to day prediction. Figure 5.6 shows the prediction graph of linear regression where green line is the actual value and red line is predicted. We can see that the difference between predicted value and actual value is not so high. The reason behind this accuracy is in this time period, Bitcoin price did not fluctuated too much. So every day's price had almost linear relation with its previous day's price. Therefore, linear regression worked better in this case.

When the price of bitcoin was stable just like 2013 to 2016 in the above graphs of 5.3, 5.4, 5.5, 5.6, our time series models could determine the closing price quite precisely. Moreover, from the starting of 2017, there was a huge instability in bitcoin price, the price was increasing very quickly, still our models could recognize it and predict the prices following the trend. But in case of machine learning models logistic and linear, it did not perform very well compared to the time series models. Therefore, we didn't get good accuracy from these machine learning models.

Lastly, after all the result the conclusion is that with the regression algorithm, it cannot be predicted how much the price of Bitcoin changes. But with the promising result of ARIMA, AR and MA model, it can be predicted whether the Bitcoin price will be increased and decreased which will be a great decision making component for the investors.

CHAPTER 6

Conclusion and Future Work

6.1 Conclusion

This task was attempted with a view to adding to the financial innovation area by enabling financial specialists or any invested individual to figure out how to dissect Bitcoin information and furthermore to utilize that learning to construct their own models. With a specific end goal to convey this undertaking to realization, we began by concentrate the different techniques for securities exchange expectation. We at that point worked with the time series analysis of investigation and forecasting of financial data. We examined the diverse ARIMA models, basing on them. While working with ARIMA models, we understood that utilizing just ARIMA models cannot prompt a precise forecast. More factors should be considered. In this thesis, we have analyzed the price of Bitcoin with respect to the US Dollar using some of the popular time-series models and incorporating machine learning to test and train our dataset. We have collected the historical data of Bitcoin over a time interval, applied different statistical techniques to get the predicted values of Bitcoin in USD on random days on our test data points and also over the next 10 days.

After analyzing the performance of these algorithms used for this thesis task, we were able to achieve a better accuracy on our dataset with ARIMA with accuracy of 91% compared to the other models that we have used.

6.2 Future Work

We worked with just the day by day shutting costs of Bitcoin information, though Bitcoin gets refreshed each moment. The end costs get refreshed by the day's end when the market is shut. Inside this time allotment, a great deal can happen, none of which gets incorporated into our investigation. On the off chance that we had minutely information in our grasp, we could have possessed the capacity to t models all the more proficiently to the information and in this way we could have become better forecasts.

Later on, we plan to gather live information and process them continuously to offer yields to the clients. If we had minutely data in our hands, we could have been able to fit models more efficiently to the data and thus we could have gotten better predictions. In the future, we intend to collect live data and process them in real time to give outputs to the users. Additionally, this research subject merely addresses just a few possible methods to obtain the predicted result of Bitcoin price. Nonetheless, there is a lot of potential to further expand this research and possibly improve the findings by applying other techniques, such as measuring the volatility of the Bitcoin market using GARCH-MIDAS analysis as done in [37]. Furthermore, user data on several social media or other online platforms can also be used to find correlation between the user sentiment and the fluctuation in the price of Bitcoin.

References

[1] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” 2008.

[2] S. Chan, J. Chu, S. Nadarajah, and J. Osterrieder, “A Statistical Analysis of Cryptocurrencies,” *Journal of Risk and Financial Management*, vol. 10, no. 2, pp. 12, May 2017 [Online]. Available at: http://dx.doi.org/10.3390/jrfm10020012

[3] Y. B. Kim, J. G. Kim, W. Kim, J. H. Im, T. H. Kim, S. J. Kang, C. H. Kim, “Predicting Fluctuations in Cryptocurrency Transactions Based on User Comments and Replies,” PLoS ONE, vol. 11, no. 2, August 2016. [Online] Available at: https://doi.org/10.1371/journal.pone.0161197 [Accessed: July 19, 2005].

[4] “Bitcoin (BTC) price, charts, market cap, and other metrics | CoinMarketCap” *Coinmarketcap.com*.[Online]. Available at: https://coinmarketcap.com/currencies/Bitcoin/ [Accessed: Jul. 19, 2018]

[5] O. Poyser, “Exploring the determinants of Bitcoin’s price: an application of Bayesian Structural Time Series,” Thesis, Autonomous University of Barcelona, Barcelona, Spain, 2017 [Online]. Available at: https://arxiv.org/ftp/arxiv/papers/1706/1706.01437.pdf

[6] M. Shen, W. Zhao and N. De, “CoinDesk - Leader in blockchain news,” *CoinDesk.com*, 2018 [Online]. Available at: https://www.coindesk.com/ [Accessed: 13 Jul. 2018].

[7] R. Böhme, N. Christin, B. Edelman, and T. Moore, "Bitcoin: Economics, Technology, and Governance," *Journal of Economic Perspectives*, vol. 29, no. 2, pp. 213-38, 2015 [Online]. Available at: http://dx.doi.org/10.1257/jep.29.2.213

[8] “Bitcoin Mining,” *Steemit.com* [Online]. Available at: https://steemit.com/Bitcoin/@tturpin/Bitcoin-mining [Accessed: Jul. 19, 2018]

[9] A. Asthana, “Bitcoin Price Prediction Using Time Series Forecasting,” *Towards Data Science*, June 27, 2018 [Online] Available at: https://towardsdatascience.com/Bitcoin-price-prediction-using-time-series-forecasting-9f468f7174d3 [Accessed: July 14, 2018].

[10] Bollerslev, T., Kretschmer, U., Pigorsch, C., & Tauchen, G. (2009). “A discrete-time model for daily S & P500 returns and realized variations: Jumps and leverage effects,” *Journal of Econometrics, 150*(2), 151-166.

[11] A. A. Adebiyi, A. O. Adewumi and C. K. Ayo. “Stock Price Prediction Using the ARIMA

Model," *in Proceedings of the 2014 UKSIM-AMSS 16th International Conference on Computer Modeling and Simulation, 26-28 March 2014, Cambridge, UK* [Online]. Available: https://ieeexplore.ieee.org [Accessed: 19 Jul. 2018]

[12] J. D. Cryer and K.-S. Chan. *Time Series Analysis with Applications in R*, 2nd ed. New York: Springer Science+Business Media, 2008.

[13] A. Kravets, "Cryptocurrency Markets Are Broken," *Cointelegraph.com*, para. 2, Dec 24, 2017. [Online]. Available at: https://cointelegraph.com/news/crypto-markets-are-broken. [Accessed: Jul. 14, 2018].

[14] E. L. Hamaker, C. V. Dolan, and P. C. M. Molenaar, (2005). "Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis," *Multivariate Behavioral Research*, vol. 40, no. 2, pp. 207-233, April 2005 [Online]. Available at: https://doi.org/10.1207/s15327906mbr4002_3

[15] D. W. J. Hosmer, S. Lemeshow, R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. New Jersey: John Wiley & Sons, 2013.

[16] P. D. Turney, "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews," *in Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), 7-12 July 2002, Philadelphia, Pennsylvania*. [Online]. Available at: https://dl.acm.org [Accessed: 19 Jul. 2018]

[17] C. Bishop, "Pattern Recognition and Machine Learning," Springer, 2006.

[18] C. Chatfield and M. Yar, "Holt-Winters Forecasting: Some Practical Issues," *Journal of the Royal Statistical Society. Series D (The Statistician)*, vol. 37, no. 2, pp. 129-40, 1988 [Online]. Available at: https://www.jstor.org/stable/2348687

[19] A. Gray, "This is what's driving the Bitcoin boom," *World Economic Forum*, para. 1, May 17, 2017. [Online] Available at: https://www.weforum.org/agenda/2017/05/what-s-driving-the-Bitcoin-boom-here-s-what-you-need-to-know/ [Accessed: July 19, 2018]

[20] "Time Series Introduction," *People.maths.bris.ac.uk*, 2018 [Online] Available at: https://people.maths.bris.ac.uk/~magpn/Research/LSTS/STSIntro.html [Accessed: July 14, 2018].

[21] P. J. Brockwell, and R. A. Davis, *Time Series: Theory and Methods*, 2nd ed. New York: Springer-Verlag, 1987. [Online] Available at: Springer e-book.

[22] C. Guestrin, and E. Fox, “Machine Learning Foundations: A Case Study Approach” *Coursera.org*. Available at: https://www.coursera.org/learn/ml-foundations/

[23] C. Chatfield, *The Analysis of Time Series: Theory and Practice*. New York: Chapman and Hall, 1975.

[24] S.-S. Shai, and B.-D. Shai, *Understanding Machine Learning*. New York: Cambridge University Press, 2014.

[25] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment Analysis of Twitter Data," *in Proceedings of the LSM ’11 Proceedings of the Workshop on Languages in Social Media, 23 June 2011, Portland, Oregon* [Online] Available: https://dl.acm.org [Accessed: 19 Jul. 2018]

[26] Y. Kim, "Convolutional neural networks for sentence classification," *in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 25-29 October 2014, Doha, Qatar*. [Online]. Available: https://arxiv.org/abs/1408.5882 [Accessed: 19 Jul. 2018]

[27] E. Kouloumpis, T. Wilson, J. D. Moore, “Twitter Sentiment Analysis: The Good the Bad and the OMG!,” *in the Proceedings of the 5th International AAAI Conference on Weblogs and Social Media, 17-21 July 2011, Barcelona, Spain*. [Online]. Available at: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2857/3251 [Accessed: 19 Jul. 2018]

[28] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features," *in European Conference on Machine Learning, ECML 1998*, C. Nédellec, C. Rouveirol, Eds. Berling: Springer, pp. 137-142.

[29] I. Georgoula, D. Pournarakis, C. Bilanakos, N. D. Sotiropoulos, and G. M. Giaglis, “Using Time-Series and Sentiment Analysis to Detect the Determinants of Bitcoin Prices,” 2015 [Online] Available at: http://dx.doi.org/10.2139/ssrn.2607167 [Accessed: July 19, 2018].

[30] “Time Series Analysis,” *statisticssolutions.com*. [Online]. Available at: http://www.statisticssolutions.com/time-series-analysis/ [Accessed: July 19, 2018].

[31] “Autocorrelation Plot,” *Itl.nist.gov*, vol. 1.3, no. 3.1, 2018 [Online] Available at: https://www.itl.nist.gov/div898/handbook/eda/section3/autocopl.htm [Accessed 21 Jul. 2018].

[32] "Microsoft Azure Machine Learning Studio," *Studio.azureml.net*, 2018 [Online]. Available at: https://studio.azureml.net/ [Accessed: July 14, 2018].

[33] Praneeth, "Bitcoin price prediction with ARIMA," *Kaggle.com*, Dec 2017 [Online] Available at: https://www.kaggle.com/praneethji/Bitcoin-price-prediction-with-arima/data [Accessed: July 13, 2018].

[34] S. Achanta, "Big Data Made Simple - One source. Many perspectives. (2018). Ten handy python libraries for (aspiring) data scientists," *bigdata-madesimple.com*, Oct 28, 2017 [Online]. Available at: http://bigdata-madesimple.com/ten-handy-python-libraries-for-aspiring-data-scientists/ [Accessed: July 13, 2018].

[35] J, Crosbie, and C. F. Sharpley, "DMITSA: A simplified interrupted time-series analysis program," *Behavior Research Methods, Instruments & Computers*, vol. 21, no. 6, pp. 639-642, June 1989.

[36] I. Bobriakov, "Top 20 Python libraries for data science in 2018," *medium.com*, June 11, 2018 [Online] Available at: https://medium.com/activewizards-machine-learning-company/top-20-python-libraries-for-data-science-in-2018-2ae7d1db8049 [Accessed: July 13, 2018].

[37] C. Conrad, A. Custovic, and E. Ghysels, "Long- and Short-Term Cryptocurrency Volatility Components: A GARCH-MIDAS Analysis," *Journal of Risk and Financial Management*, vol. 11, no. 2, pp. 23, May 2018. [Online] Available at: http://dx.doi.org/10.3390/jrfm11020023 [Accessed: July 19, 2005].