Forecasting Bitcoin Prices: ARIMA and Seasonal Decomposition Approach

Abinav Sudarshan R
Dr. Manu K S

ABSTRACT:
The research is done to forecast the Bitcoin prices, the study considered Bitcoin prices. The data has been collected in an hourly basis from 1st July 2017 to 1st July 2018. The data is of high frequency. The forecast has been made for a period of 12 months from 31st July 2018 to 31st July 2019. The forecasted data has two different periods, one to calculate the error and the other to have an idea of the future price changes. The study uses Autoregressive Integrated Moving Average(ARIMA) and Seasonal Decomposition method for forecasting the Bitcoin prices. This article also helps the investors to have an idea about the future prices of Bitcoin. The forecasted prices also act as a factor for investor decision making, based on the forecast further choices can be made based on the Bitcoin price changes and invest accordingly. The results of this paper have has sown that, between the two models used, ARIMA has resulted to be the best model compared to the seasonal decomposition model as, the Mean Absolute Error, Mean Absolute Percentage Error and Root Mean Square Error is lower in ARIMA forecasting model, which satisfies the objective of the paper and the best model is selected.

Key words: Bitcoin prices, ARIMA, Seasonal Decomposition, Mean Absolute Error, Mean Absolute Percentage Error, Root Mean Square Error.

I. INTRODUCTION:
A cryptocurrency is a intangible currency, which is used as a medium of exchange that uses the cryptography model to do financial transactions digitally. It is a decentralized control as opposed to any centralized works through distributed ledger technology, like a block chain, which is public financial transaction database. Decentralized cryptocurrency is produced by the system which develops the software at a rate which is defined when the system is created and publicly known to all. In a centralized banking system or economy they have lot of control over the financial tools used in the country, but in cases of cryptocurrencies it cannot be produced by governments or any boards, it is based on underlying technical system upon which decentralized cryptocurrencies are created. There are almost 1000’s of cryptocurrencies available in the world, and all of them are created in a similar manner. Out of all the cryptocurrencies, Bitcoin (₿) is the most famous and also has the highest market value compared to any cryptocurrency in the world. There are other cryptocurrencies in the world like Litecoin, Ethereum also known as Ether, Ripple, Dash, etc. these cryptocurrencies are a great alternative for a lot of investors to invest as they can use them to earn better returns than other investment avenues.

The price of Bitcoin is at 7,013.97 USD(INR= 2,44,961.36) and its market capitalisation is at 60.96 billion USD (INR=4.3 Lakh crore approx.), other currencies are very far behind by both pricing and market cap, Ether is currently traded at 226.81USD (INR= 16234.71) and dash at 169.52 USD (INR=12140.12), their respective market caps are valued at 22 billion USD and 1.56 billion USD. Few of the major countries which have allothyed cryptocurrency trading and are also regulated by their respective governments are China,
France, Germany, UK, Russia and USA. The current trends are the platforms which open for the investors to pull in their money to invest in the various cryptocurrency platforms and also ICO’s (Initial Coin Offering) are the trend changers in the field of cryptocurrency now. But the pre ICO’s perform better than ICO’s, also they face a regulatory pushback. The other challenges are Bad user experience who tend to cheat other investors, the availability of common regulatory bodies also the places which accept bitcoins are less, technical limitations also had few headaches to the currencies and finally it has a very high volatility. It is very important in few aspects as in the future it would be easier to use, it has low cost means of transaction, there are lots of safe trustable wallets, it helps in privacy and most importantly it has complete autonomy.

Cryptocurrencies are the latest trend in the world, but when it comes trading in it people hesitate a lot because of the risks involved in it and the other legal blocks which are exposed on it by different governments of different countries. The government in India for example has banned trading in cryptocurrency illegal. There are very huge risks as previously explained, it has high volatility, no regulatory body, its not globally accepted by all, there are still few technical inabilities which are faced by the traders or users. Cryptocurrencies as everyone say are the latest trend in the world, when it comes trading in it people hesitate a lot because of the risks involved in it and the other legal blocks which are exposed on it by different governments of different countries. The government in India for example has banned trading in cryptocurrency illegal. Before this ban the cryptocurrencies were being traded throughout the country for investment purposes and also for trading, few people used this currency to pay off business deals, but it was banned as the government could not regulate it and a lot of people who were interested in trading in the currencies were losing money because of the illegitimate transactions done by the regulators of the currency. So the government decided to take a step further and ban it, till the necessary steps can be taken to regulate it and use it as a tool for investment or exchange. After this there was a lot of illegal trading happening behind closed doors, because few traders have analysed the benefits involved in trading and also analyse the risks and returns involved in trading in cryptocurrencies. This research helps to analyse the risks and volatility of cryptocurrency and the benefits that can be gained due to trading in cryptocurrencies.

II. LITERATURE REVIEW:
(Isah, Kazeem, Raheem, Ibrahim D ,2018) This paper is to study the motivation of the investors interested in investing in cryptocurrencies. To validate this belief, the constructs a prescient model in which digital currencies are recognized as the indicators of US stock returns. Three outcomes radiated from their estimations. To start with, it has approved the significance of digital currencies in anticipating US stock costs; second, the cryptographic forms of money prescient model beats the regular time-arrangement models, third, their outcomes are powerful to various strategy for figure execution assessment measures and diverse sub-test periods. These outcomes have essential approach suggestions for the speculators and policymakers. (Catania, Leopoldo, Grassi, Stefano, Ravazzolo, Francesco ,2018) This paper contemplates the consistency of digital sorts of cash time arrangement. they give the impression of being at a couple of possibility univariate and variable models in purpose and thickness guaging of 4 of the foremost promoted arrangement: Bitcoin, Litecoin, Ripple and Ethereum. They apply tons of crypto indicators and rely upon Dynamic Model Averaging to consolidate a large arrangement of univariate Dynamic Linear Models and a couple of variable Vector Autoregressive models with numerous varieties of time
variety”. They notice factual large upgrades in purpose estimating once utilizing mixes of univariate models and in thickness determinative whereas counting on determination of variable models.

(Stavroyiannis Stavros, 2018) The purpose of this paper is to look at the value-at-risk and connected measures for the Bitcoin and to check the findings with customary and Poor’s SP500 Index, and therefore the gold damage statistic. A GJR- model has been enforced, within which the residuals follow the standardized Pearson type-IV distribution”. A large kind of value-at-risk measures and back testing criteria square measure enforced. (Tetsuya Takaishi, 2018) The study uses varied tools to spot applied math properties and multifractality of bitcoin at the influence of GBP-USD, at the time of Brexit. It uses varied applied math tools and conjointly , GJR, and R models. The results show whereas Brexit influenced the GBP-USD rate, Bitcoin was sturdy to Brexit. (Anshul Saxena, T.R. Sukumar, 2018) The study involves the statistic prediction of the bitcoin costs with improved potency mistreatment long STM techniques (LSTM) and compares its foregone conclusion with the standard methodology (ARIMA).

(Jonathan Rebane, Isak Karlsson, Stojan Denic, Panagiotis Papapetrou, 2018) This study compares the model performance of ARIMA to it of a seq2seq repeated deep multi-layer neural network (seq2seq) utilizing a varied choice of inputs sorts. The results demonstrate superior performance of seq2seq over ARIMA, for models generated throughout most of bitcoin value history”, with extra knowledge sources resulting in higher performance throughout less volatile value periods. (Leopoldo Catania, Stefano Grassi, Francesco Ravazzolo, 2018) This paper studies the sure thing of cryptocurrencies statistic. They compare many different univariate and variable models in purpose and density foretelling of 4 of the foremost capitalized series:” Bitcoin, Litecoin, Ripple and Ethereum. (Muhammad J Amjad, Devavrat Shah, 2018) They show that Bitcoin value knowledge exhibit fascinating properties like stationarity and compounding. Even so, some classical statistic prediction ways that exploit this behaviour, like ARIMA models”, produce poor predictions and additionally lack a probabilistic interpretation.

(Dian Utami Sutiksnso, Ansari Saleh Ahmar, Nuning Kurniasih, Eko Susanto, and Audrey Leiwakabessy, 2017) The article uses the ARIMA model to use the α-Sutte Indicator and ARIMA in foretelling knowledge. α-Sutte Indicator may be a new foretelling methodology that was developed in 2017 by Ansari Saleh Ahmar. to envision the accuracy of those ways, the foretelling results of the α-Sutte Indicator helped them to forecast the currency values”. (Jeffrey Chu, Stephen Chan, Saralees Nadarajah and Joerg Osterrieder, 2017) This article helps to calculate the volatility of cryptocurrencies through models,”it shows the assorted levels of volatility of various currencies. The results of the study show that cryptocurrencies like Bitcoin, Ethereum, Litecoin and plenty of others exhibit extreme volatility particularly once we scrutinize their inter-daily costs. (Nashirah Abu Bakar, Sofian Rosbi, 2017) This article calculates the volatility of Bitcoin through the statistical processes, it used various tools such as descriptive statistics in its research. Normality statistical checking is performed for data of exchange rate and rate of return for Bitcoin. The volatility dynamic behaviour of Bitcoin return is evaluated using box-whisker plot and statistical process control in this study.

III. OBJECTIVES OF THE STUDY:

1. To forecast Bitcoin prices using ARIMA and seasonal Decomposition models
2. To select the best forecasting model based on the forecasting performance in terms of lowest risk error.
Hypothesis:

H₀: The monthly Bitcoin prices are non-stationary.

H₁: The monthly Bitcoin prices are stationary.

IV. METHODOLOGY:

A. Data:

The data is for bitcoin prices and around 8,000 observations are collected. The hourly prices are considered in this study.

B. Sources of data:

Data are collected from Coindesk.com and Cryptodatadownload.com. These sources have the data range for the top cryptocurrencies being traded in the US/UK exchanges.

C. Period of study:

The period of study is for 1 year, as the data consists of hourly observations of Bitcoin in the US/UK exchange.

D. Analytical tools used:

Analance:

Analance Advanced Analytics, is the tool used, it coordinates with best-of-breed measurable investigation dialects like R and Python with a suite of pre-fabricated, zero-coding machine learning calculations to kick off examination. The procedure is totally consistent. Extra machine learning calculations can be added to target explicit business challenges.

 a. Augmented Dickey-Fuller (ADF):

The study used Augmented Dickey-Fuller (ADF) test to verify whether stock index prices are stationary or not.

 b. ARIMA model:

An autoregressive integrated moving average, or ARIMA, could be a applied mathematics analysis model that uses statistic information to either higher perceive the information set or to predict future trends. An autoregressive integrated moving average model could be a sort of multivariate analysis that gauges the strength of 1 variable quantity relative to alternative ever-changing variables. The model's goal is to predict future securities or monetary market moves by examining the variations between values within the series rather than through actual values.

An ARIMA model are often understood by outlining every of its elements as follows:

- Auto regression (AR) refers to a model that shows a ever-changing variable that regresses on its own lagged, or prior, values.
• Integrated (I) represents the differencing of raw observations to permit for the statistic to become stationary, i.e., information values are replaced by the distinction between the information values and also the previous values.

• Moving average (MA) incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations.

Calculation of ARIMA

\[
\left(1 - \sum_{i=1}^{q} \alpha_i L^i\right) X_t = \left(1 + \sum_{i=1}^{p} \theta_i L^i\right) \varepsilon_t
\]

Where; \( L \) = lag operator, \( \alpha_i \) = parameters of the autoregressive part of the model, \( \theta_i \) = parameters of the moving average part of the model, \( \varepsilon_t \) = Error term

c. Seasonal Decomposition Method:
Seasonal decomposition is fundamentally helpful for finding the statistic information, and finding historical changes over time, it can be employed in statements. To forecast a rotten statistic, we have a tendency to forecast the seasonal part, and also the seasonally adjusted part severally. It's typically assumed that the seasonal part is unchanging, or ever-changing extraordinarily slowly, therefore a seasonal naive methodology is employed for the seasonal part. The basic decomposition methodology consists of estimating the 5 elements of the model.

Calculation of Seasonal decomposition method

Additive Model

\[ y_t = T_t + C_t + S_t + I_t, \]

Multiplicative Model

\[ y_t = T_t \times C_t \times S_t \times I_t. \]

Where; \( T_t \) is the trend component at time \( t \), \( C_t \) is the cyclical component at time \( t \), \( S_t \) is the seasonal component at time \( t \), \( I_t \) is the irregular component at time \( t \).

d. Mean Absolute error is calculated using the following formula;

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - x_i|
\]

Where; \( y_i \) = actual value, \( x_i \) = forecasted value, \( n \) = number of observations

e. Mean Absolute Percentage Error is calculated using the following formula;

\[
M = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|
\]

Where; \( A_t \) = actual value, \( F_t \) = forecasted value, \( n \) = number of observations

f. Root Mean Square Error is calculated using the following formula;
Where; $P_i$ - actual value, $O_i$ - forecasted value, $n$ - number of observations

E. Basis of selection:

The basis of selection of cryptocurrency is popularity, value of the currency in the market and highly traded component in the market. The basis of USD, is that it is the most popularly used currency by the traders to trade the cryptocurrencies. The US/UK exchanges are the highly used exchanges compared to other exchanges. This currency is chosen on the basis of higher market cap and the availability of data for the chosen time frame.

V. DATA ANALYSIS AND INTERPRETATION:

Table (1): showing the descriptive statistics of the variable

<table>
<thead>
<tr>
<th>Mean</th>
<th>Median</th>
<th>Max</th>
<th>Min</th>
<th>SD</th>
<th>Sk</th>
<th>Kr</th>
<th>Jargue-Bera</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0001</td>
<td>0.0003</td>
<td>0.129</td>
<td>-0.096</td>
<td>0.0133</td>
<td>0.1349</td>
<td>10.86</td>
<td>21207</td>
<td>0</td>
</tr>
</tbody>
</table>

Table (1) shows the descriptive statistics. The basic features of the data could be explained through descriptive statistics. The samples that are chosen for the study are analysed for its measures and summaries through descriptive statistics. It forms the quantitative analysis of the data. The central tendency of the data is found. The standard deviation indicates the variations in Bitcoin prices over the period. Skewness indicates how the data differs or distorts from that of normal distribution.

Table (2): Showing stationarity results

<table>
<thead>
<tr>
<th>Level 1</th>
<th>t-statistics</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns</td>
<td>-70.785</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table (2), the data to be stationary the probability has to be less than 0.05%. The Bitcoin Price data has probability more than 0.05% at level. Therefore, we reject the null hypothesis and accept alternate hypothesis.

Table (3): showing the forecasted data at various frequency intervals using ARIMA method

<table>
<thead>
<tr>
<th>Date</th>
<th>Forecasting Values</th>
<th>LowerInterv al 80</th>
<th>HigherInterv al 80</th>
<th>LowerInterv al 95</th>
<th>HigherInterv al 95</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018-07-31</td>
<td>6417.60</td>
<td>5673.52</td>
<td>7161.68</td>
<td>5279.63</td>
<td>7555.57</td>
</tr>
<tr>
<td>2018-08-31</td>
<td>6418.75</td>
<td>5435.66</td>
<td>7401.84</td>
<td>4915.24</td>
<td>7922.26</td>
</tr>
<tr>
<td>2018-09-30</td>
<td>6418.76</td>
<td>5255.02</td>
<td>7582.50</td>
<td>4638.97</td>
<td>8198.55</td>
</tr>
<tr>
<td>2018-10-31</td>
<td>6418.76</td>
<td>5094.02</td>
<td>7743.50</td>
<td>4392.75</td>
<td>8444.77</td>
</tr>
<tr>
<td>2018-11-30</td>
<td>6418.76</td>
<td>4954.98</td>
<td>7882.54</td>
<td>4180.11</td>
<td>8657.41</td>
</tr>
<tr>
<td>2018-12-31</td>
<td>6418.76</td>
<td>4824.00</td>
<td>8013.53</td>
<td>3979.78</td>
<td>8857.74</td>
</tr>
<tr>
<td>2019-01-31</td>
<td>6418.76</td>
<td>4702.98</td>
<td>8134.54</td>
<td>3794.70</td>
<td>9042.82</td>
</tr>
<tr>
<td>2019-02-27</td>
<td>6418.76</td>
<td>4604.14</td>
<td>8233.38</td>
<td>3643.54</td>
<td>9193.98</td>
</tr>
<tr>
<td>2019-03-31</td>
<td>6418.76</td>
<td>4493.56</td>
<td>8343.96</td>
<td>3474.42</td>
<td>9363.10</td>
</tr>
<tr>
<td>2019-04-30</td>
<td>6418.76</td>
<td>4395.37</td>
<td>8442.15</td>
<td>3324.25</td>
<td>9513.27</td>
</tr>
</tbody>
</table>
The table (3) shows the forecasted value that is been generated by the tool, using ARIMA model, which has the lower interval and the higher interval with confidence level 80 and 95, time period being between July of 2018 to July of 2019.

**Figure (1):** showing forecasting chart of Bitcoin at various intervals under ARIMA model

![Forecasting Plot](image)

This figure (1) shows the plot of the forecasted value time period being between July of 2018 to July of 2019. Which includes the lower and higher interval with reference to confidence level 80 and 95.

**Table (4):** showing the forecasted data at various frequency intervals using Seasonal Decomposition method.

<table>
<thead>
<tr>
<th>Date</th>
<th>Forecasting Values</th>
<th>Lower Interval 80</th>
<th>Higher Interval 80</th>
<th>Lower Interval 95</th>
<th>Higher Interval 95</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018-07-31</td>
<td>6418.76</td>
<td>4298.68</td>
<td>8538.84</td>
<td>3176.38</td>
<td>9661.15</td>
</tr>
<tr>
<td>2019-06-30</td>
<td>6418.76</td>
<td>4209.13</td>
<td>8628.39</td>
<td>3039.43</td>
<td>9798.10</td>
</tr>
<tr>
<td>2019-07-31</td>
<td>6418.76</td>
<td>4177.20</td>
<td>8660.33</td>
<td>2990.58</td>
<td>9846.94</td>
</tr>
</tbody>
</table>

The table (4) is the forecasted value that is been generated by the tool, using Seasonal Decomposition, which has the lower interval and the higher interval with confidence level 80 and 95, time period being between July of 2018 to July of 2019.
Figure (2): showing forecasting chart of Bitcoin at various intervals under Seasonal Decomposition Method

The figure (2) shows the plot of the forecasted value time period being between July of 2018 to July of 2019. Which includes the lower and higher interval with reference to confidence level 80 and 95.

Table (5): showing the comparison of the forecasted value with the actual to find the error.

<table>
<thead>
<tr>
<th>Date</th>
<th>Forecasting Values(ARIMA)</th>
<th>Actual</th>
<th>Error</th>
<th>Forecasting Values(SDM)</th>
<th>Actual</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>31-07-2018</td>
<td>6417.60</td>
<td>8,184.70</td>
<td>-1,767.10</td>
<td>6392.24</td>
<td>8,184.70</td>
<td>-1,792.46</td>
</tr>
<tr>
<td>31-08-2018</td>
<td>6418.75</td>
<td>6938.91</td>
<td>-520.16</td>
<td>6399.48</td>
<td>6938.91</td>
<td>-539.43</td>
</tr>
<tr>
<td>30-09-2018</td>
<td>6418.76</td>
<td>6698.42</td>
<td>-279.66</td>
<td>6396.83</td>
<td>6698.42</td>
<td>-301.59</td>
</tr>
<tr>
<td>31-10-2018</td>
<td>6418.76</td>
<td>6276.46</td>
<td>142.30</td>
<td>6395.10</td>
<td>6276.46</td>
<td>118.64</td>
</tr>
<tr>
<td>30-11-2018</td>
<td>6418.76</td>
<td>4218.61</td>
<td>2,200.15</td>
<td>6399.48</td>
<td>4218.61</td>
<td>2,180.87</td>
</tr>
<tr>
<td>31-12-2018</td>
<td>6418.76</td>
<td>3689.56</td>
<td>2,729.20</td>
<td>6395.17</td>
<td>3689.56</td>
<td>2,705.61</td>
</tr>
<tr>
<td>31-01-2019</td>
<td>6418.76</td>
<td>3563.28</td>
<td>2,855.48</td>
<td>6397.00</td>
<td>3563.28</td>
<td>2,833.72</td>
</tr>
</tbody>
</table>

Table (5) compares the actual Bitcoin price with the forecasted price using two methods, ARIMA forecasting and the Seasonal decomposition forecasting. After the comparison the error value for both is found out. Where it was seen that the error in ARIMA method was lower than that of Seasonal Decomposition method, also it is found out that there is no constant increase or decrease in the prices.

Table (6) shows the various factors which show the tests validity

<table>
<thead>
<tr>
<th>Particulars</th>
<th>ARIMA</th>
<th>SEASONAL DECOMPOSITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Absolute Error</td>
<td>90.74%</td>
<td>91.61%</td>
</tr>
<tr>
<td>Mean Absolute Percentage Error</td>
<td>1.04%</td>
<td>1.05%</td>
</tr>
<tr>
<td>Root Mean Square Error</td>
<td>197.55%</td>
<td>199.31%</td>
</tr>
</tbody>
</table>

Table (6) show that the Mean Absolute Error is 90.74% in ARIMA, which is less that seasonal decomposition that has 91.61%, the Mean Absolute Percentage Error of seasonal decomposition 1.05% is slightly more than that of ARIMA 1.04%, whereas, the Root Mean Square Error is also less in ARIMA model. Which proves that when compared to the Seasonal Decomposition method ARIMA method has lesser error in the forecasting of the Bitcoin prices.
In this paper, the forecasted values are examined using two methods ARIMA and Seasonal Decomposition method. The interaction during the entire sample period (time period being between July of 2017 to July of 2018) is investigated. The trial period is from 31st July 2017 to 31st July 2018 is taken, and the forecast is made from 31st July 2018 to 31st July 2019. Where the actual data is compared with the forecasted values and the error is found out in both the models. The results of this paper have has sown that, between the two models used, ARIMA has resulted to be the best model compared to the seasonal decomposition model as, the Mean Absolute Error, Mean Absolute Percentage Error and Root Mean Square Error is lower in ARIMA forecasting model, which satisfies the objective of the paper and the best model is selected.

VI. CONCLUSION:

The research is done to forecast the Bitcoin prices, the study considered Bitcoin prices. The data has been collected in an hourly basis from 1st July 2017 to 1st July 2018. The data is of high frequency. The forecast has been made for a period of 12 months from 31st July 2018 to 31st July 2019. The forecasted data has two different periods, one to calculate the error and the other to have an idea of the future prices. The study used Autoregressive Integrated Moving Average(ARIMA) and Seasonal Decomposition method for forecasting the Bitcoin prices. The study found that ARIMA model has given better forecasting results compared to seasonal decomposition method in terms of forecasting error. The forecasted prices also act as a factor for investor decision making, based on the forecast further choices can be made based on the Bitcoin price changes and invest accordingly.

The study can be further carried on for various purposes, the scope for improvement always exists in research, it can be used for validating future prices also. i. The study can also been pursued to understand the volatility, fluctuations and vulnerability of the bitcoin prices. ii. The forecasted prices can be used for investment decisions and understanding the behaviour of the bitcoin prices in the future. iii. This research can be carried on bitcoin behaviour in the market and also the various price changes it experiences due to various factors influencing market value.

REFERENCES


