CRYPTOCURRENCY TRENDS AND PRICE PREDICTION USING PROPHET AND ARIMA TIME SERIES

PROJECT REPORT

Submitted by

GOWTHAM SAINI [RA1811031010070]

Under the Guidance of

Dr. M. Shobana

(Assistant Professor, Department of Networking and Communications)

in Partial Fulfillment of the requirements for the degree of

BACHELOR OF TECHNOLOGY
In
COMPUTER SCIENCE ENGINEERING
with specialization in INFORMATION TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

FACULTY OF ENGINEERING AND TECHNOLOGY

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY

KATTANKULATHUR- 603 203

APRIL 2022

Own Work Declaration

Department of Computer Science and Engineering

SRM Institute of Science & Technology

This sheet must be filled in (each box ticked to show that the condition has been met). It must be signed and dated along with your student registration number and included with all assignments you submit – work will not be marked unless this is done.

To be completed by the student for all assessments

Degree/ Course : **B. Tech/Computer Science and Engineering with specialization in Information Technology**

Student Name : **Gowtham Saini**

Registration Number : **RA1811031010070**

Title of Work : **Cryptocurrency trends and price prediction using Prophet and Arima time series**

I / We hereby certify that this assessment compiles with the University's Rules and Regulations relating to Academic misconduct and plagiarism**, as listed in the University Website, Regulations, and the Education Committee guidelines.

I / We confirm that all the work contained in this assessment is my / our own except where indicated, and that I / We have met the following conditions:

- Clearly references / listed all sources as appropriate
- Referenced and put in inverted commas all quoted text (from books, web, etc)
- Given the sources of all pictures, data etc. that are not my own
- Not made any use of the report(s) or essay(s) of any other student(s) either past or present
- Acknowledged in appropriate places any help that I have received from others (e.g. fellow students, technicians, statisticians, external sources)
- Compiled with any other plagiarism criteria specified in the Course handbook / University website

I understand that any false claim for this work will be penalised in accordance with the University policies and regulations.

DECLARATION:

I am aware of and understand the University's policy on Academic misconduct and plagiarism and I certify that this assessment is my / our own work, except where indicated by referring, and that I have followed the good academic practices noted above.

If you are working in a group, please write your registration numbers and sign with the date for every student in your group.

GOWTHAM SAINI
RA1811031010070
20.04.2022

SRM INSTITUTE OF SCIENCE AND TECHNOLOGY
KATTANKULATHUR – 603 203

BONAFIDE CERTIFICATE

Certified that this B.Tech project report titled "**Cryptocurrency trends and price prediction using Prophet and Arima time series"** is the bonafide work of **GOWTHAM SAINI (RA1811031010070)** who carried out the project work under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion for this or any other candidate.

Dr. M. Shobana

(Project Guide)
Assistant Professor
Department of
Networking and
Communications

Signature of the Internal Examiner

Signature of the External Examiner

ACKNOWLEDGEMENT

We express our humble gratitude to **C. Muthamizhchelvan**, Vice-Chancellor (I/C), SRM Institute of Science and Technology, for the facilities extended for the project work and his continued support.

We extend our sincere thanks to **Dr Revathi Venkataraman**, Professor & Chairperson, School of Computing, SRM Institute of Science and Technology, for his invaluable support.

We wish to thank **Dr B. Amutha**, Professor & Head, Department of Computer Science and Engineering, SRM Institute of Science and Technology, for her valuable suggestions and encouragement throughout the period of the project work.

We are extremely grateful to our Faculty Advisor **Dr.S. MURUGAANANDAM,** Assistant Professor, Department of Networking and Communication SRM Institute of Science and Technology, for their great support at all the stages of project work.

Our inexpressible respect and thanks to our guide, **Dr. M. Shobana,** Assistant Professor, Department of Networking and Communications, SRM Institute of Science and Technology, for providing me with an opportunity to pursue my project under his mentorship. He provided me with the freedom and support to explore the research topics of my interest. His passion for solving real problems and making a difference in the world has always been inspiring.

We sincerely thank the staff and students of the Computer Science and Engineering Department, SRM Institute of Science and Technology, for their help during my research. Finally, we would like to thank my parents, our family members and our friends for their unconditional love, constant support and encouragement.

GOWTHAM SAINI
(RA1811031010070)
20.04.2022

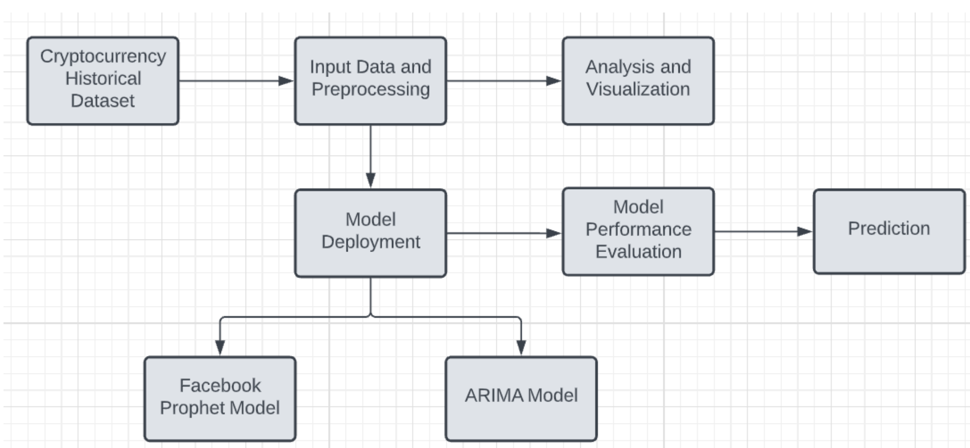
TABLE OF CONTENTS

ABSTRACT

Crypto currencies, also known as virtual currency, have received a lot of attention in recent years due to the success of the major crypto coin, Bitcoin. But nowadays, people are interested in more than just Bitcoin; there are a plethora of currencies available. There are around 2000 crypto currencies in all. People are frequently perplexed as to which coin to invest in for a lucrative future. This study takes the top 12 crypto currencies into account and forecasts the closing price of a crypto currency. Following investigation, the results revealed that Bitcoin, Tether, and Ethereum had the biggest volume of coin investment, with Bitcoin topping the closing value followed by Ethereum. The candlestick plot is used to visualise crypto currency trends. For the top three currencies stated above, the adjusted closing rate, which is the closing price of a bitcoin, is forecasted using two time series algorithms: ARIMA and Prophet Model. For both models, the R2 score has been calculated. Both models have nearly identical scores. The outcomes of this research may be used in the real world to choose the top coin to put a customer's money in. The code is modular in nature and can be simply altered to obtain the closing price for other currencies as well.

CHAPTER -1
INTRODUCTION

People always look for long term investments which would provide good profit for their money. According to a survey, about 10.2% of people worldwide own crypto currency which is a huge number. In India alone, 20 million have entered the crypto market. As exciting as it is to own virtual currency, it's also very risky to invest due to the volatility. This volatility startles many people from investing. It would be really helpful if they know the future trends of coins to better decide. There are multiple newsletters that suggest which crypto people need to invest in to make profit, however most of them are paid. Having the algorithm and workflow implemented would make it easier to analyze coins and make the prediction on adjusted closing price of coins with higher volume. This way people or investors can make better decisions about when and the amount to invest in a particular crypto coin. The paper presents machine learning models that can be used to estimate the closing price of a coin accurately and with easy to follow implementation. Initially, the data is collected on the top 12 crypto coins from an open source website [1]. This website contains data on all the crypto coins available till date. The reason for choosing this website as a data source is because Yahoo is considered to be a trusted data source for investment data, be it stock market or crypto market. Various techniques like area plot, candle plot, etc have been used to visualize the stock data. Analyzing the data helped decide what coins to do the prediction on and determine the best model with optimum performance. There are two models implemented in this paper: ARIMA and Prophet model. Data was preprocessed and prepared in accordance with both the models. A check was performed on the data for stationary properties which is a requirement for the ARIMA model. Overall, both the models have similar accuracy and give the same results. For the final model predictions, prophet model is chosen because the results are easy to interpret and prophet model can also easily detect the seasonal trends along with easy to understand parameters. This model property will come handy when the algorithm is scaled to other crypto coins which have uneven trends in volume and closing rate. This section briefly introduces the project and the next section contains the literature review where similar work by other researchers have been introduced. The coming sections succinctly explains the methodology and the experimental setup including the tests.



CHAPTER - 2
LITERATURE SURVEY

2.1 Yahoo Finance, portfolio management resources, international market data, social interaction and mortgage.

Blockchain technology, which is the underlying framework of cryptocurrencies, has gained a lot of attention and trust because it provides secure transactions and fast data transfer. It also provides authentication of a product and can act as a contract. Investing in crypto currencies has been a challenge for most people due to its volatility. Based on the study in [1], there are multiple factors that seem to have an effect on volatility and predictability of cryptos like trade volume, exchange rates, supply & demand, cost of transactions.

2.2 Mahir Iqbal, Muhammad Shuaib Iqbal, Fawwad Hassan Jaskani,*, Khurum Iqbal and Ali Hassan. Time-Series Prediction of Cryptocurrency Market using Machine Learning Techniques. Published on 07 July 2021 EAI Endorsed Transactions.

This paper suggests that crypto market related factors like trade volume and its uncertainty, volatility in coins are significant in determining coins like Bitcoin, Ethereum, Dash, Litecoin and Monero. Taking these factors into consideration, predicting the price of crypto currencies has been proposed by many research enthusiasts.

2.3 Xiaolei Sun, Mingxi Liu, Zeqian Sima. A Novel Cryptocurrency Price Trend Forecasting Model Based on LightGBM. December 2018. Research Gate.

This paper focusses on considering the right set of features that can be used for this prediction. This paper takes the deep learning approach to the problem by implementing algorithms like LSTM and RNN. Accuracy for single and multi-feature Machine Learning models is determined.

The plots for actual price versus predicted price in [3] show that the algorithms implemented in the paper do a good job in predicting prices and can be used to get prices for future values. This is an indication that now only mathematical analysis can determine the prices of crypto, Machine Learning and Deep Learning algorithms can also be used to predict prices with good accuracy. Due to the popularity of bitcoin, most of the research papers that were referred have predicted top coins like Bitcoin, Ethereum, etc using LSTM, RNN algorithm.

2.4 Bhanu PRAKASH Kolla, K L University. Predicting Crypto Currency Prices Using Machine Learning and Deep Learning Techniques. September 2020. Research Gate.

This paper proves that (Gated Recurrent Network) GRU is a better algorithm when compared to LSTM, bi-LSTM and RNN. The plots presented in this paper clearly depicts the closeness in predicted values to the true prices of crypto and lowest error rates with GRU model

2.5 Mohammad J. Hamayel and Amani Yousef Owda. A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms. MDPI

A similar approach is followed in [5] where recurrent network algorithms and an ensemble algorithm: gradient boosting classifier outperforms random classifiers. This paper uses technical features, sentiment features and blockchain based features. Results showed that technical features proved to be the most significant which resonates with the thoughts of financial analysts. Financial analysts also are relying on the technical features to determine if a stock or crypto is safe to buy or not. The paper also discusses whether Bitcoin can someday replace paper-based currency. The authors of the paper have mentioned that Bitcoin can possibly replace paper-based currency in developed countries but could prove to be a disaster in developing countries.

2.6 Mohammed khalid salman and Abdullahi Abdu Ibrahim. Price prediction of different cryptocurrencies using technical trade indicators and machine learning. IOP Conference Series.

Technical trade features have been used to develop an automated Machine Learning pipeline to predict Bitcoin’s prices with an accuracy of 94.89%. The paper also discusses whether Bitcoin can someday replace paper-based currency. The authors of the paper have mentioned that Bitcoin can possibly replace paper-based currency in developed countries but could prove to be a disaster in developing countries. Any new product must be accepted by both consumers and businesses in order to become a worldwide success. When we look at advanced machine learning methods for the transactions of paper-based currencies (such as US Dollar money) to cryptocurrencies for that matter, we see that it has been accepted and applied at the root level.

CHAPTER 3
METHODOLOGY

To meet the goals of this study, historical cryptocurrency prices were used to train two different models for three different types of cryptocurrencies named, Bitcoin, Ethereum and Tether. Firstly, we examined the top 12 cryptocurrencies based on their volume and the coin's closing price every 24 hours. We employed two separate models in this case: Facebook Prophet and the ARIMA model, which works well with time series data.

Facebook Prophet Model

The problem statement is implemented using Prophet, a Python module. The evaluation highlights characteristics as well as the seasonality of events. Pattern, seasonality, and holidays are three critical components in a time collection edition breakdown. The following formula can be used to combine these additives:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

The function g(t) represents non-periodic adjustments, the feature s(t) represents periodic adjustments, and the feature h(t) represents holiday outcomes. Any peculiar modifications that aren't handled by the version are represented by the error t.

The historical data of cryptocurrency comes with a lot of seasonality that has to be adjusted in order to create a better model as there are lot of effects that need to be handled before implementing the model. Prophet uses fourier series to give a flexible model for fitting and forecasting the impacts of seasonality. The following function approximates seasonal impacts s(t):

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right)\right)$$

With our time variable scaled in days, let P be the average duration of the time collection (e.g., P = 365.25 for yearly records or P = 7 for weekly recordings).

ARIMA Model

ARIMA – Autoregressive integrated moving average is a statistical model that works on a time series data set to either better comprehend the data or anticipate future trends. To successfully deploy the ARIMA model on a time series dataset, a few requirements must be satisfied.

- **Stationarity Check:**
  To get better results, it's advisable to examine whether the data we're dealing with is stationary or non-stationary. The data is said to be stationary when the properties of a time series do not rely on time and there is no trend in the data. The cryptocurrency dataset which was used gone through few tests to check its stationarity.

- **Augmented Dickey-Fuller test:**
  The Augmented Dickey Fuller test (ADF Test) is a typical statistical test used to determine whether or not a particular time series is stationary. When it comes to examining the stationary of a series, it is one of the most widely employed statistical tests.

```
x = Closing_price
result = adfuller(x)
print('ADF Statistic: %f' % result[0])
print('p-value: %f' % result[1])
print('Critical values:')
for key, value in result[4].items():
    print('\t%s: %.3f' % (key, value))

if result[0] < result[4]["5%"]:
    print ("Reject the Null Hypothesis - Time series is stationary")
else:
    print ("Failed to reject the Null Hypothesis - Time series is Non-Stationary")
```

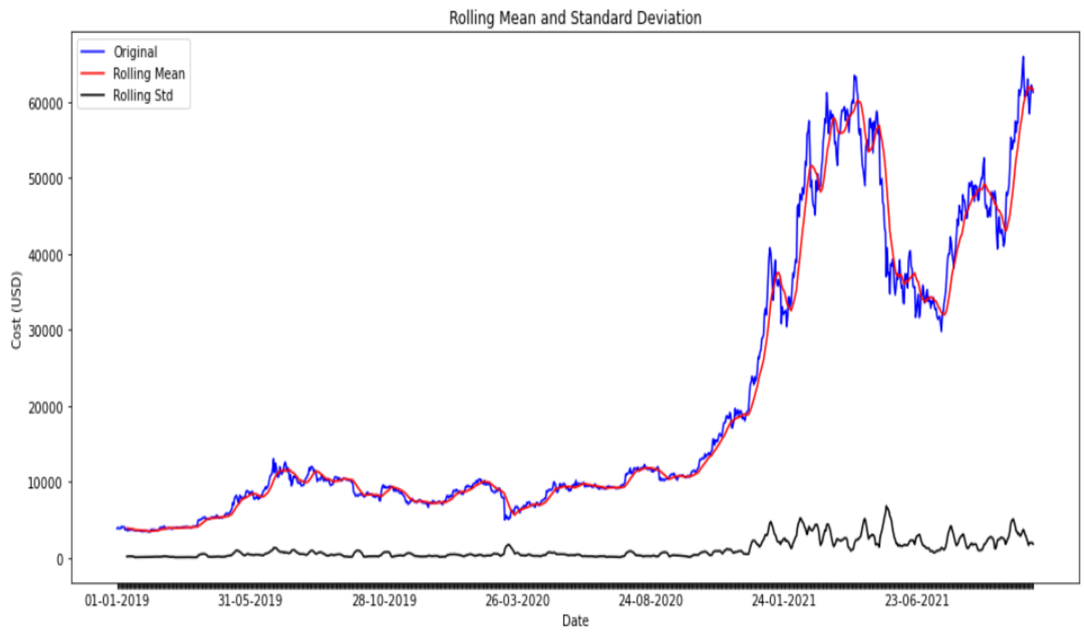
```
ADF Statistic: 0.107824
p-value: 0.966591
Critical values:
        1%: -3.437
        5%: -2.864
        10%: -2.568
Failed to reject the Null Hypothesis - Time series is Non-Stationary
```

If the test statistic value is less than the critical value then the null hypothesis (HO) is rejected, indicating that the time series lacks a unit root and is thus not stationary.

- **Rolling Statistics**

  The fundamental idea behind using the Rolling statistics test is that it visually depicts the rolling mean and rolling standard deviation to analyse the trend in the data by presenting both mean and standard deviation.

```
#Method 2: Rolling Statistics
rollmean = Closing_price.rolling(12).mean()

rollstd = Closing_price.rolling(12).std()
```



We could notice that the mean is constantly changing with the increase in time, and the standard deviation is likewise not constant, there it has been proved that the data is not stationary.

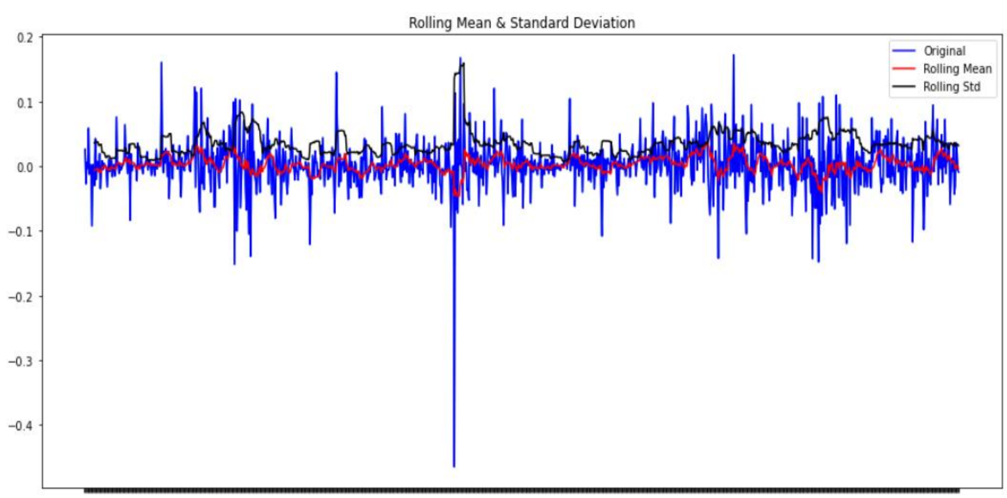
Converting the non-stationary data to stationary:

- **Differencing**

  The one efficient way to convert the data to stationary is by using differencing, it can eliminate the variations in the time series, alter the trends and also remove the seasonality by stabilising the mean. Stationarizing a time series by differencing (where necessary) is a key step in the ARIMA model fitting process.

  The focused value is the closing price of the cryptocurrency, if Xt signifies the values of the time series X at a period t, then the first difference of X calculated at period t is equal to Xt-Xt-1.

```
#Using Differencing
plt.figure(figsize=(16,7))
fig = plt.figure(1)
t_log_diff = t_log - t_log.shift(1)
plt.plot(t_log_diff)
```

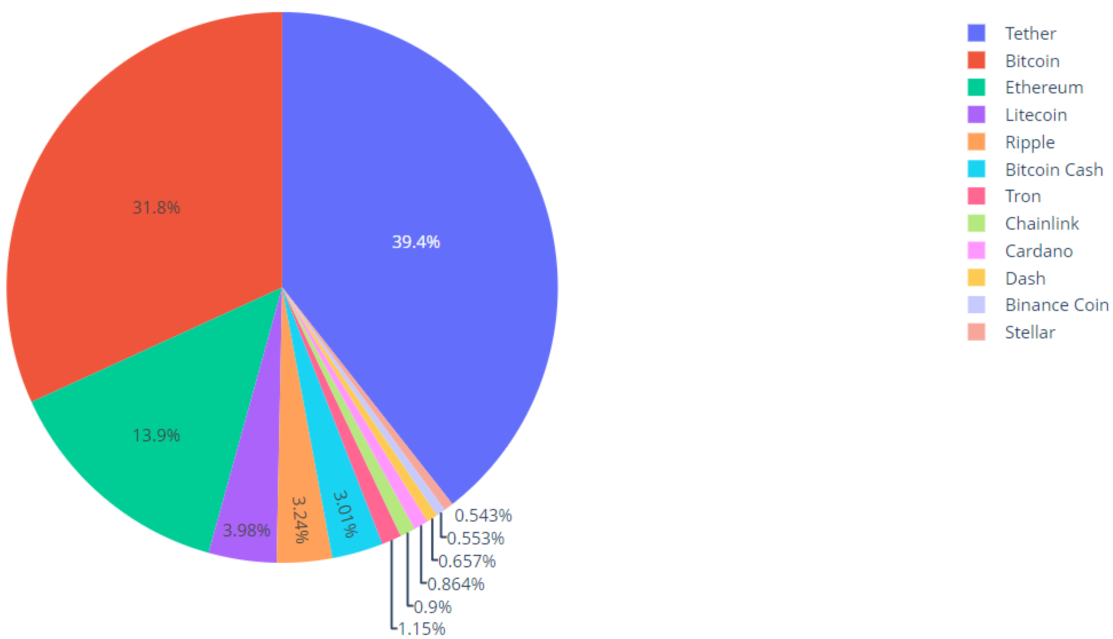


The above plot shows that there is no sudden trend in the data and the difference of the mean by considering two points is negligible, indicating that the time series is now stationary and that there is no seasonality in the time series, and we can now fit the data into the ARIMA model successfully.

CHAPTER - 4
RESULTS

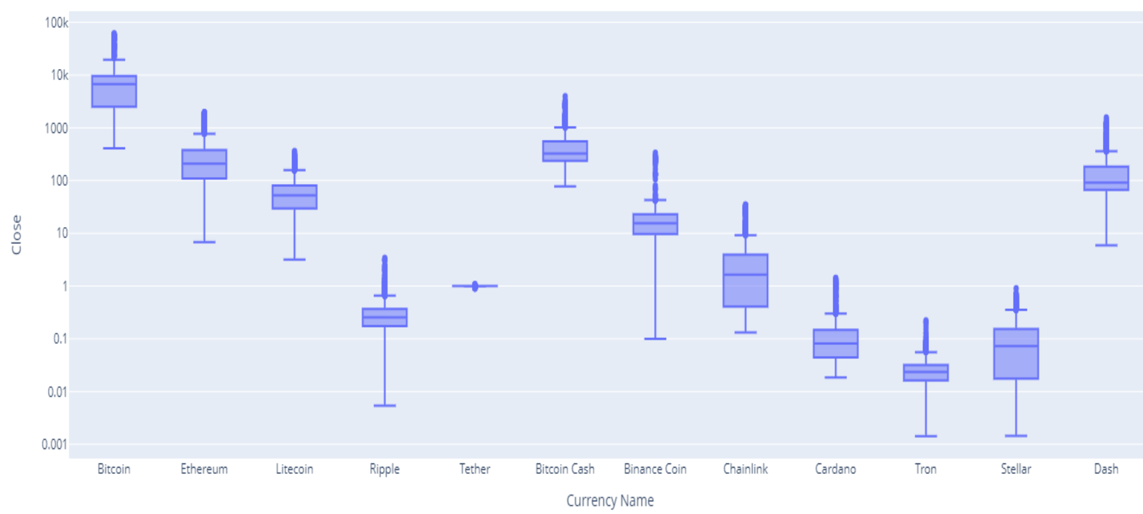
Data Analysis and Visualization:

The historical data for the top 12 cryptocurrencies was obtained from https://finance.yahoo.com/ which collected data from 2017 to 2021. The study was based on the crypto currency's closing price and volume, which refers to the number of coins traded on that given day

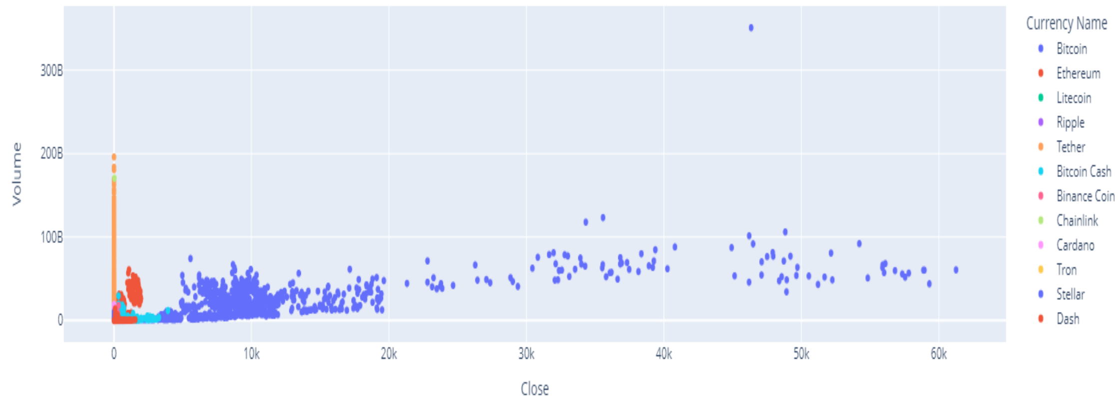


The above pie chart shows that the coins Bitcoin, Tether, and Ethereum have been highly traded in the market, and the majority of those investing in cryptocurrency are gravitating toward these three coins.

Two time series models were developed based on the aforementioned data to forecast the future prices of these coins, allowing consumers to make informed investment decisions and limit the chance of losing money. Aside from the pie chart, various additional plots were used to depict other qualities such as the opening and closing prices, such as the candle stick plot and box plot. A comparison of the top 12 cryptocurrencies was carried out using a scatter plot, which revealed that the closing price for bitcoin and Ethereum was the highest. The study also revealed that the majority of persons interested in cryptocurrencies invest in the coin Ethereum, despite the fact that the closing price of Ethereum has never surpassed $1.

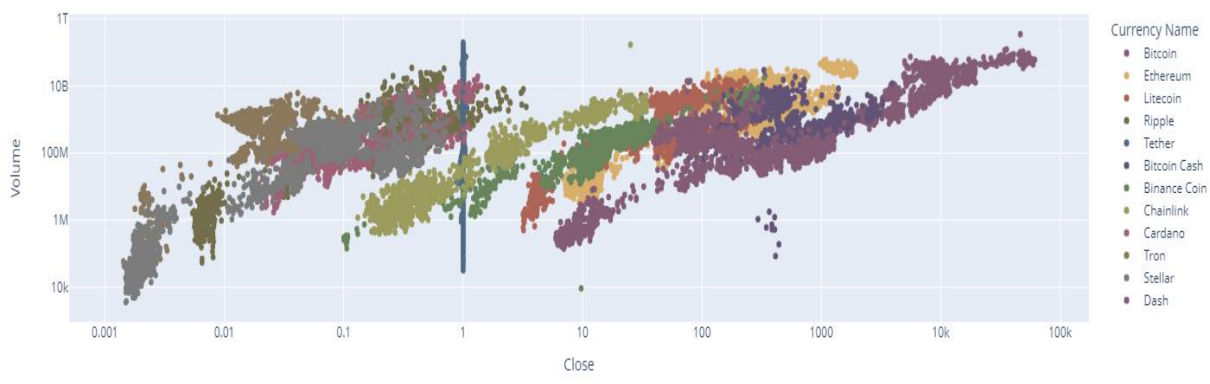


A boxplot is a standardised method of depicting data distribution based on a five-number summary ("minimum", first quartile (Q1), median, third quartile (Q3), and maximum"). It can provide information about your outliers and their values. The above box plot illustrates the first, median, and third quartiles of tether, which are nearly identical; it can be noted that there is no variation between the quartiles and there are few outliers. The Binance cryptocurrency has the highest number of outliers.



The following scatter plot compares currencies based on their closing price and volume to have a better understanding of the most invested coin among the top 12 cryptocurrencies and the coin with the greatest closing value in a given time period. In the above plot the volume of tether has remained the highest meaning that enough number of people have invested in Tether coin but it's also very clear that the closing price of Tether is almost constant. The highest and most varied closing price can be seen in the case of Bitcoin having the highest closing price reaching almost $60k USD.

According to the above analysis, the currency with the highest closing price is Bitcoin, while the coin with the highest volume is Tether.



The above plot depicts the highest, lowest, volume of the top 12 cryptocurrencies in order to get a better understanding of the coins.

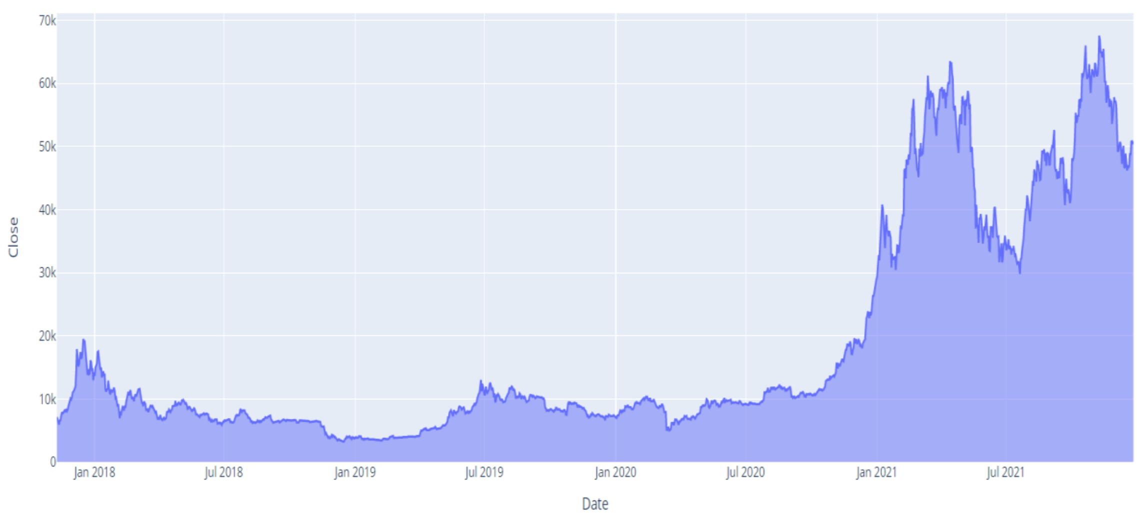
Implementing the Prophet model on the Bitcoin Dataset

Understanding the Dataset and it's features

	Date	Open	High	Low	Close	Adj Close	Volume
0	2017-11-09	7446.830078	7446.830078	7101.520020	7143.580078	7143.580078	3226249984
1	2017-11-10	7173.729980	7312.000000	6436.870117	6618.140137	6618.140137	5208249856
2	2017-11-11	6618.609863	6873.149902	6204.220215	6357.600098	6357.600098	4908680192
3	2017-11-12	6295.450195	6625.049805	5519.009766	5950.069824	5950.069824	8957349888
4	2017-11-13	5938.250000	6811.189941	5844.290039	6559.490234	6559.490234	6263249920
...	...	...	...	...	...	...	...
1504	2021-12-22	48937.097656	49544.796875	48450.941406	48628.511719	48628.511719	24447979559
1505	2021-12-23	48626.343750	51332.339844	48065.835938	50784.539063	50784.539063	28223878108
1506	2021-12-24	50806.050781	51814.027344	50514.496094	50822.195313	50822.195313	24367912228
1507	2021-12-25	50854.917969	51176.597656	50236.707031	50429.859375	50429.859375	19030650914
1508	2021-12-26	50428.691406	51196.378906	49623.105469	50809.515625	50809.515625	20964372926

1509 rows × 7 columns

The data shown above depicts the high, low, close, adjusted close, and volume of bitcoin during the last four years. The target variable in this case is the adjusted closing price of the day, on which the forecast will be made until 2023.



Candlestick plot demonstrate this sentiment by graphically showing the size of price movements with distinct colours. Candlesticks are used by traders to make trading decisions based on regularly recurring patterns that assist estimate the price's short-term direction.

The above candlestick plot illustrates the opening, closing, high, and low prices of the coins, with green plots indicating that the closing price was higher than the previous day and red plots indicating that the closing price was lower than the previous day.

Prophet is often used to handle with data that contains two columns: ds and y. For dates, the Pandas-friendly format for the ds (datestamp) column is YYYY-MM-DD and for timestamps, YYYY-MM-DD HH:MM:SS. The y column must be numeric and reflect the measurement we want to forecast; the methods require it to be converted to a datetime data type using the pandas function to datetime.

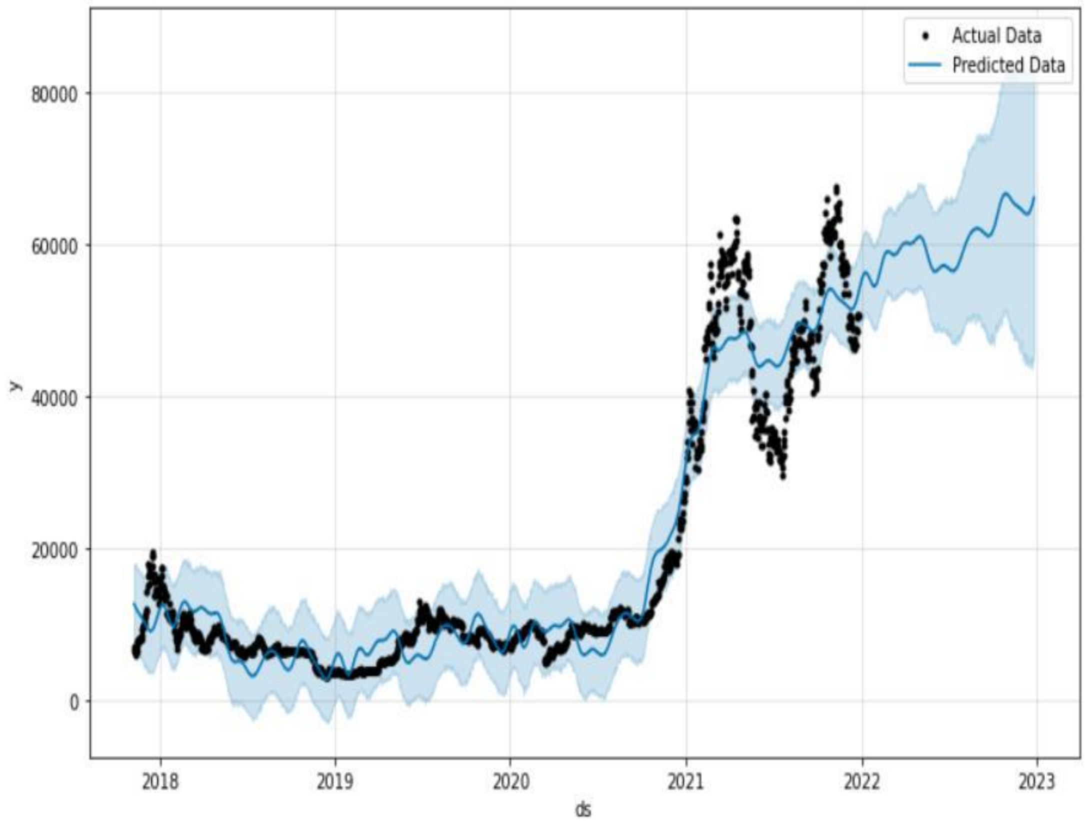
```
# Creating a new dataset that only contains two columns which is the date and the closing price of the bitcoin.
columns = ["Date","Close"]

df1= pd.DataFrame(df_bitcoin, columns = columns)
```

The input data for Bitcoin price prediction must be a data frame having two columns, "ds" and "y," where ds represent the date and y represents the closing price of the bitcoin.

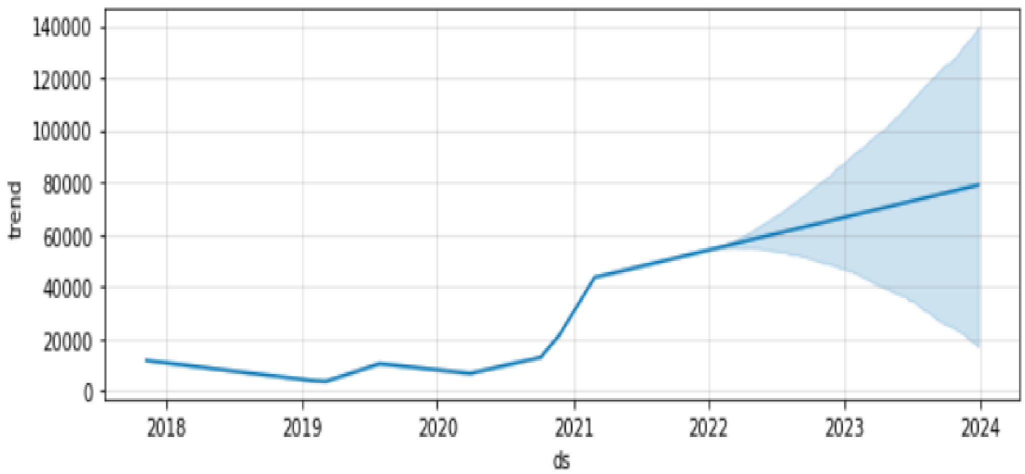
```
# The input must be a data frame with two columns 'ds' and 'y'# (ds is the date and y is the number of crimes). Let's adjust it.

prophet_df = df1.rename(columns = {'Date':'ds', 'Close':'y'})
```

	ds	y
0	2017-11-09	7143.5800
1	2017-11-10	6618.1401
2	2017-11-11	6357.6000
3	2017-11-12	5950.0698
4	2017-11-13	6559.4902

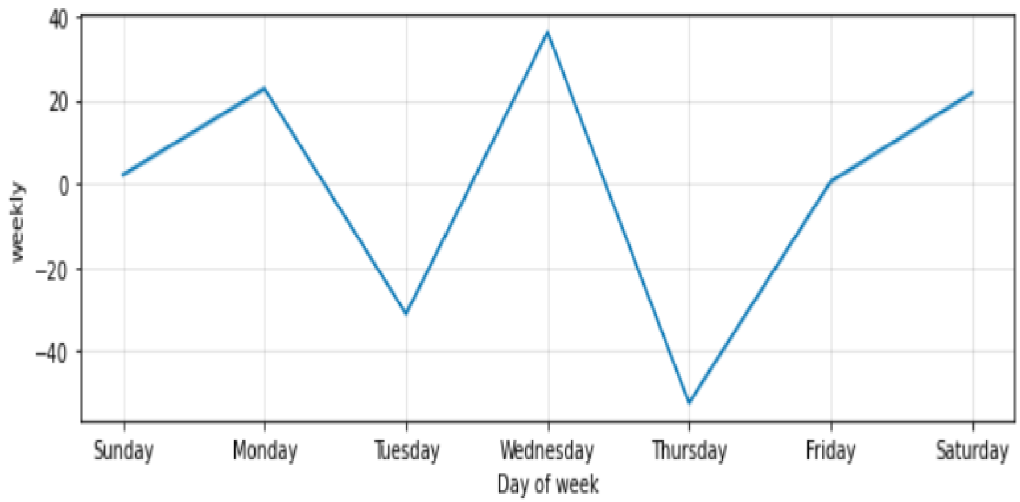
Prophet is a time series data forecasting model that can manage seasonality as well as holiday impacts on a monthly, weekly, and daily basis. When the data is time sensitive and has a lengthy history of seasonality, Prophet is the ideal model to use for predicting. According to Prophet's GitHub description, Facebook uses Prophet for a variety of reliable forecasts that are resistant to outliers and missing data. The model mentions projections for the future. Facebook Prophet also gives crime patterns on an annual, weekly, and monthly basis. These visual insights give an understanding of the data's underlying tendencies.



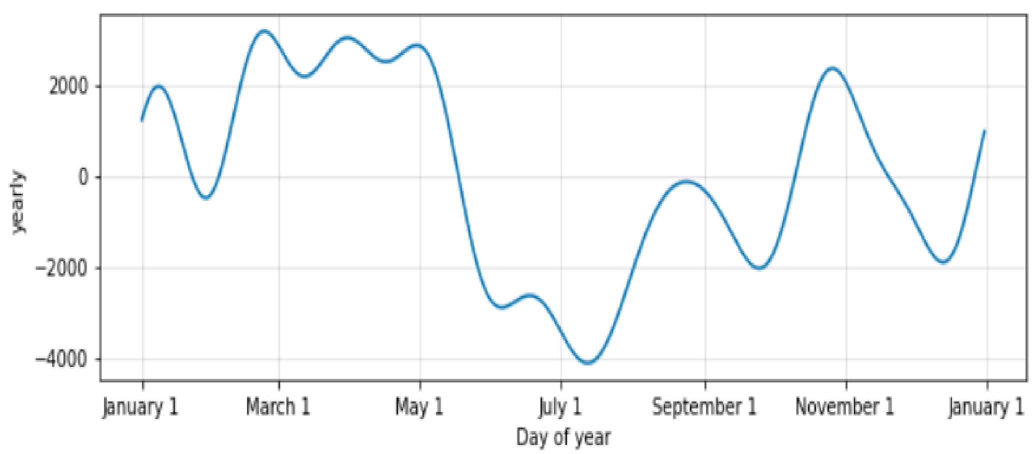
The graph above depicts the fundamental forecast, with light blue denoting the amount of uncertainty meaning the upper and the lower bound of the prices, dark blue representing the prediction, and black dots representing the actual data, ds represent the dates from 2018 to the year 2023, including future values, while y represents Bitcoin prices. The graph also depicts the top and lower bounds, which represent the maximum and least prices that bitcoin might reach.



140000
120000
100000
80000
60000
40000
20000
0
trend
2018
2019
2020
2021
2022
2023
2024
ds



40
20
0
−20
−40
weekly
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Day of week



2000
0
−2000
−4000
yearly
January 1
March 1
May 1
July 1
September 1
November 1
January 1
Day of year

	ds	yhat_lower	yhat_upper	yhat
0	2017-11-09	7395.627399	18237.104546	12748.756765
1	2017-11-10	7183.421457	18041.213646	12656.266457
2	2017-11-11	6826.171922	18027.212381	12535.912987
3	2017-11-12	7022.263958	18153.066358	12379.428551
4	2017-11-13	6800.484877	17525.308123	12268.191231
...	...	...	...	...
2234	2023-12-22	17684.093829	135862.497984	77616.669160
2235	2023-12-23	16756.280369	137568.147234	77861.022466
2236	2023-12-24	16950.266071	137471.564491	78080.206483
2237	2023-12-25	16305.770031	138962.507855	78352.313846
2238	2023-12-26	16887.851549	137289.638353	78560.034354

[2239 rows x 4 columns]

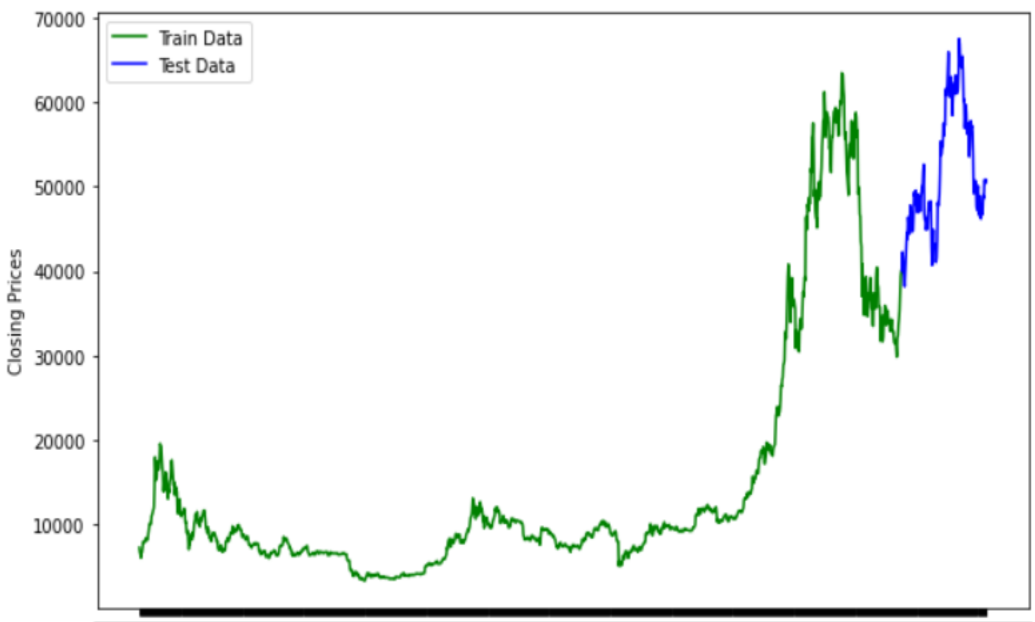
The table above displays the bitcoin's upper bound, lower bound, and predicted values. The expected value is depicted by yhat.

Implementation of ARIMA Model

Before fitting the data into the Arima model, various preconditions must be satisfied, such as the data must be stationary and there must be no seasonality in the data.

After meeting all of the requirements (detailed in the Methodology section), the data is divided into training and testing data. The train-test split is used to measure the performance of machine learning algorithms relevant to prediction-based Algorithms/Applications. This approach is a quick and simple procedure that allows us to compare our own machine learning model outcomes to machine results. Here 90% of the data is used to train, while 10% of the remaining data is utilised to test the findings.

```
#Plotting Training and Testing data
plt.figure(figsize=(10,6))
plt.xlabel('Dates')
plt.ylabel('Closing Prices')
plt.plot(alpha[0:to_row]['Adj Close'], 'green', label = 'Train Data')
plt.plot(alpha[to_row:]['Adj Close'], 'blue', label = 'Test Data')
plt.legend()
```

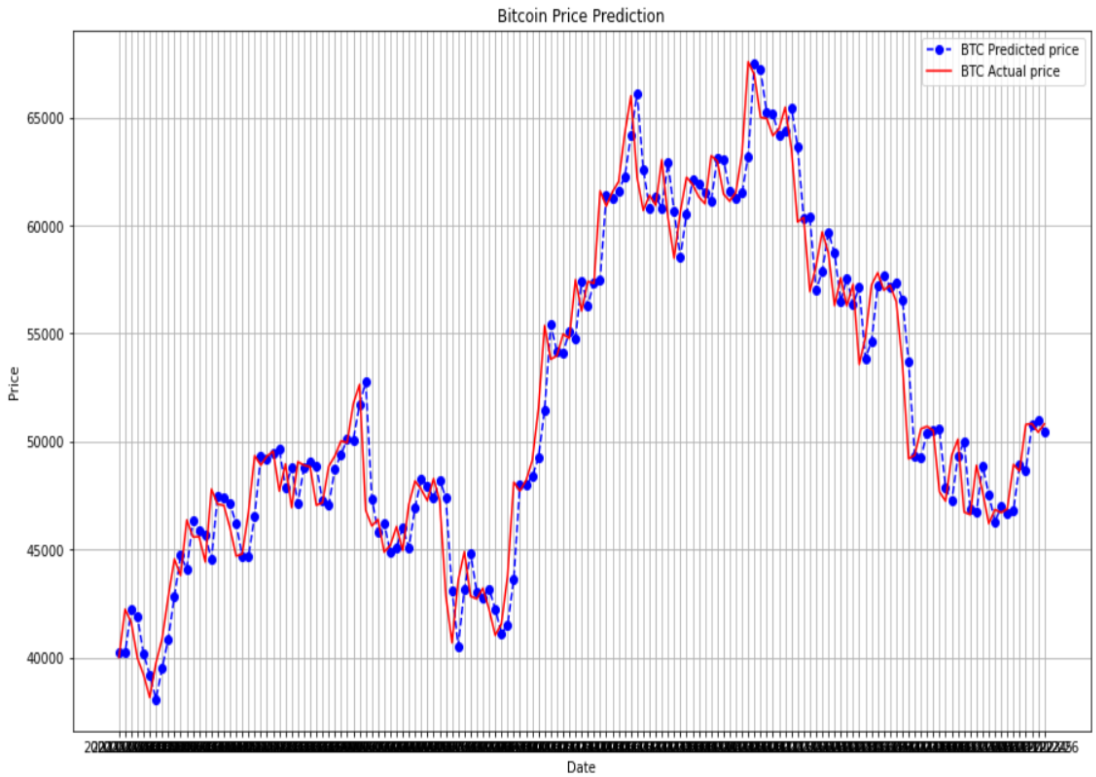


ARIMA models are commonly designated as ARIMA(p,d,q)(P,D,Q)m, where m refers to the number of seasons and the uppercase P,D,Q correspond to the autoregressive, differencing, and moving average components for the seasonal part of the ARIMA model. So the next stage in properly implementing the ARIMA Model is to determine the order that must be utilised and to go through the model summary in order to run the model flawlessly.

```
print(model_fit.summary())
```

```
                              ARIMA Model Results
==============================================================================
Dep. Variable:                    D.y   No. Observations:                 1507
Model:                 ARIMA(4, 1, 0)   Log Likelihood              -12550.955
Method:                       css-mle   S.D. of innovations           1001.741
Date:                Thu, 17 Mar 2022   AIC                          25113.910
Time:                        10:00:08   BIC                          25145.817
Sample:                             1   HQIC                         25125.794

==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         28.7235     26.453      1.086      0.278     -23.123      80.570
ar.L1.D.y     -0.0365      0.026     -1.420      0.156      -0.087       0.014
ar.L2.D.y      0.0016      0.026      0.063      0.950      -0.049       0.052
ar.L3.D.y      0.0063      0.026      0.245      0.806      -0.044       0.057
ar.L4.D.y      0.0532      0.026      2.068      0.039       0.003       0.104
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            2.0881           -0.0000j            2.0881           -0.0000
AR.2           -2.0680           -0.0000j            2.0680           -0.5000
AR.3           -0.0693           -2.0848j            2.0860           -0.2553
AR.4           -0.0693           +2.0848j            2.0860            0.2553
-----------------------------------------------------------------------------
```

```
# Plotting the predictions obtained from the model
plt.figure(figsize=(15,9))
plt.grid(True)
date_range = alpha.head(to_row)['Date']
date_range = alpha[to_row:].index
plt.plot(date_range, model_predictions, color = 'blue',
marker = 'o', linestyle = 'dashed', label = 'BTC Predicted
price')
plt.plot(date_range, testing_data, color = 'red', label =
'BTC Actual price')
plt.title('Bitcoin Price Prediction')
plt.xlabel('Date')
plt.ylabel('Price')
plt.legend()
plt.show()
```



The graph above represents the forecast, with the blue line representing the predicted price of bitcoin and the red line representing the actual price of bitcoin. It can be seen from the above graph that the predictions are very similar to the actual values of the coin and also follows the trend similar to the historical prices.

Accuracy measurement

To forecast the prices of cryptocurrencies, this programme uses techniques such as Facebook prophet and ARIMA model. The most important stage in this process is to measure the accuracy of the model to guarantee that the model is correct. Understanding the forecasting accuracy in comparison to real data is critical.

In forecasting, the expected value might be less or more than the actual number. The method used here to check the accuracy of the model is r2 score. The r2 score ranges from 0 to 100 percent. It has a tight relationship with the MSE. The fraction of the variation in the dependent variable that is predicted from the independent variable is denoted as r2 (s).

MODEL	**R2 Score**
Facebook Prophet	0.9398997074696636
ARIMA	0.9421699813469266

Both the models show similar results meaning both Prophet and ARIMA suits well on cryptocurrency price prediction.

CONCLUSION:

The report successfully proposes two time series Machine Learning models: ARIMA and Prophet model. Prior to deploying both models, data cleaning and data processing have been implemented to prepare data. The given data was converted to stationary for successful implementation of the ARIMA model. Analysis of top 12 crypto coins demonstrated that Tether has the highest trade volume but the closing price has been consistently $1 and Bitcoin had the second highest volume along with the maximum closing price, followed by Ethereum which aligns with what actually happened. Various visualizations throughout this study support the above statement. Both the ARIMA model and Prophet model have very similar performance with ARIMA model having slightly higher R-square score. R-square score for ARIMA is 94% and Prophet is 93%. However, the future predictions were chosen to be predicted using the Prophet model because of the model's simplicity and easy to understand methods and workflow. The code has been implemented in such a way that it can easily be reused to get the price predictions for other crypto coins as well.

REFERENCES:

[1] https://finance.yahoo.com/\

[2] Mahir Iqbal, Muhammad Shuaib Iqbal, Fawwad Hassan Jaskani,*, Khurum Iqbal and Ali Hassan. Time-Series Prediction of Cryptocurrency Market using Machine Learning Techniques. Published on 07 July 2021 EAI Endorsed Transactions.

[3] Xiaolei Sun, Mingxi Liu, Zeqian Sima. A Novel Cryptocurrency Price Trend Forecasting Model Based on LightGBM. December 2018. Research Gate.

[4] Bhanu PRAKASH Kolla, K L University. Predicting Crypto Currency Prices Using Machine Learning and Deep Learning Techniques. September 2020. Research Gate.

[5] Mohammad J. Hamayel and Amani Yousef Owda. A Novel Cryptocurrency Price Prediction Model Using GRU, LSTM and bi-LSTM Machine Learning Algorithms. MDPI

[6] PatrickJaquart, DavidDann, ChristofWeinhardt. Short-term bitcoin market prediction via machine learning. November 2021. The Journal of Finance and Data Science

[7] https://en.wikipedia.org/wiki/Technical_analysis

[8] Mohammed khalid salman and Abdullahi Abdu Ibrahim. PRICE PREDICTION OF DIFFERENT CRYPTOCURRENCIES USING TECHNICAL TRADE INDICATORS AND MACHINE LEARNING. IOP Conference Series