****

Shivaji University, Kolhapur

**A PROJECT REPORT ON**

“CRYPTO-CURRENCY PRICE PREDICTION”

**IN PARTIAL FULFILLMENT OF THE DEGREE**

Post-graduation diploma in data science

Department of Computer Science

**2021-2022**

**Under The Guidance of:**

**Dr. Urmila Pol**

**Project Member:**

**1) Gouri Mohite**

**2) Aishwarya Kolekar**

CERTIFICATE

This is to certify that Miss. Gouri H. Mohite & Miss. Aishwarya S. Kolekar, have a satisfactorily completed the project entitled as “Crypto-Currency Price Prediction” in partial fulfillment of Post-Graduation Diploma in Data Science during the academic year 2021-2022.

Place : Kolhapur

Date :

**H.O.D.**

Dr. R. K. Kamat

Dept. Of Computer Science,

Shivaji University, Kolhapur.

**Project Guide**

Dr. Urmila Pol

Dept. Of Computer Science,

Shivaji University, Kolhapur.

**Examiner**

DECLARATION

We hereby declare that, the project entitled ‘Crypto-Currency Price Prediction’ has not formed earlier the basis for the award of any degree of this or any other university or examination body.

Further we declared that we have not violated any of the provisions under copy write act.

Place : Kolhapur

Date :

Student Name Seat No. Signature

Miss. Gouri H. Mohite

Miss. Aishwarya S. Kolekar

INDEX

|  |  |  |
| --- | --- | --- |
| **Sr. No.** | **Content** | **Page No.** |
| 1. | Abstract | 05 |
| 2. | Introduction | 06 |
| 3. | Problem Statement | 07 |
| 4. | Dataset Description | 08 |
| 5. | Techniques | 09 |
| 6. | Summery Statistics | 12 |
| 7. | Sampling | 15 |
| 8. | Flowchart | 16 |
| 9. | Exploratory Data Analysis | 17 |
| 10. | Screenshots | 22 |
| 11. | System Requirements | 28 |
| 12. | Conclusion | 29 |
| 13. | References | 30 |

ABSTRACT

*Now-a-days more people wants to invest in different cryptocurrencies and make money out of it. About 20 million people now have invested their money into cryptocurrencies. but crypto have very ups and downs in its price daily and contains very high risk for investors cause it is unpredictable and no one can claims about its price and this is the main obstacle in predicting the future price of any cryptocurrency. So in this project we try to predict the price of cryptocurrency and we mainly focused on the top 5 currencies of today’s date namely Bitcoin, Etherium, Tether, USD-Coin and Binance Coin. For the forecasting purpose we have used different Statistical, Machine Learning and Deep Learning models such as Moving Average, ARIMA, Exponential Smoothing and LSTM.*

INTRODUCTION

*A cryptocurrency is a digital or virtual currency that is secured by cryptography, which makes it nearly impossible to counterfeit or double-spend. Cryptocurrency, sometimes called crypto-currency or crypto, is any form of currency that exists digitally or virtually and uses cryptography to secure transactions. Cryptocurrencies don't have a central issuing or regulating authority, instead using a decentralized system to record transactions and issue new units. Cryptocurrency is now trending method for making money. People do invest in crypto such as Bitcoin, Ethereum, Tether and so on. Today, there are over 18,000 cryptocurrencies in circulation. And among these cryptocurrencies Bitcoin is on high that is it have high cost now and it’s price is $29,683.72.*

*It is very difficult when we go for the forecasting the price of any crypto currency, because of unpredictable behaviour of it’s price. The prices of these currencies mainly depends on the daily news about crypto. The industries and authorities make various decision about various thinks it can be impact to the prices of crypto which leads to sudden increase or sudden decrease of its price.*

*So for the predictive purpose we have used various algorithms and few of them are able to predict the price of Cryptocurrency.*

PROBLEM STATEMENT

*Predicting the price of Cryptocurrency for the next one day or for next 10 days on the basis of previous Close prices with the help of Statistical tools, machine learning and deep learning algorithms. Which may help investors for the invest there money in crypto with minimum risk.*

DATASET DESCRIPTION

*We have collected the dataset from the Yahoofinance, yfinance is the python library which provides historical data of any cryptocurrency as well as stock.*

*Our dataset contains 7 columns namely Date, Open, High, Low, Close, Adj Close and Volume. The cryptocurrency data is the Time series data, we have collected the data of various currencies from yfinance from date of it come in existence to the today’s date. From yfinance library we are able to get the data of every day updated records. We have collected the data with time lapse of 1 day, the ‘Date’ column represents the dates in YYYY-MM-DD format. ‘Open’ column represents the daily opening price of the currency. ‘High’ column gives the highest value reached in that day by currency similarly ‘Low’ column gives the lowest value reached in that day by currency. As like column ‘Open’, column ‘Close’ gives the closing price of each day. The ‘Adj Close’ column price amends a stock's closing price to reflect that stock's value after accounting for any corporate actions. The column ‘volume’ gives the information about how currency traded over a price on that day.*

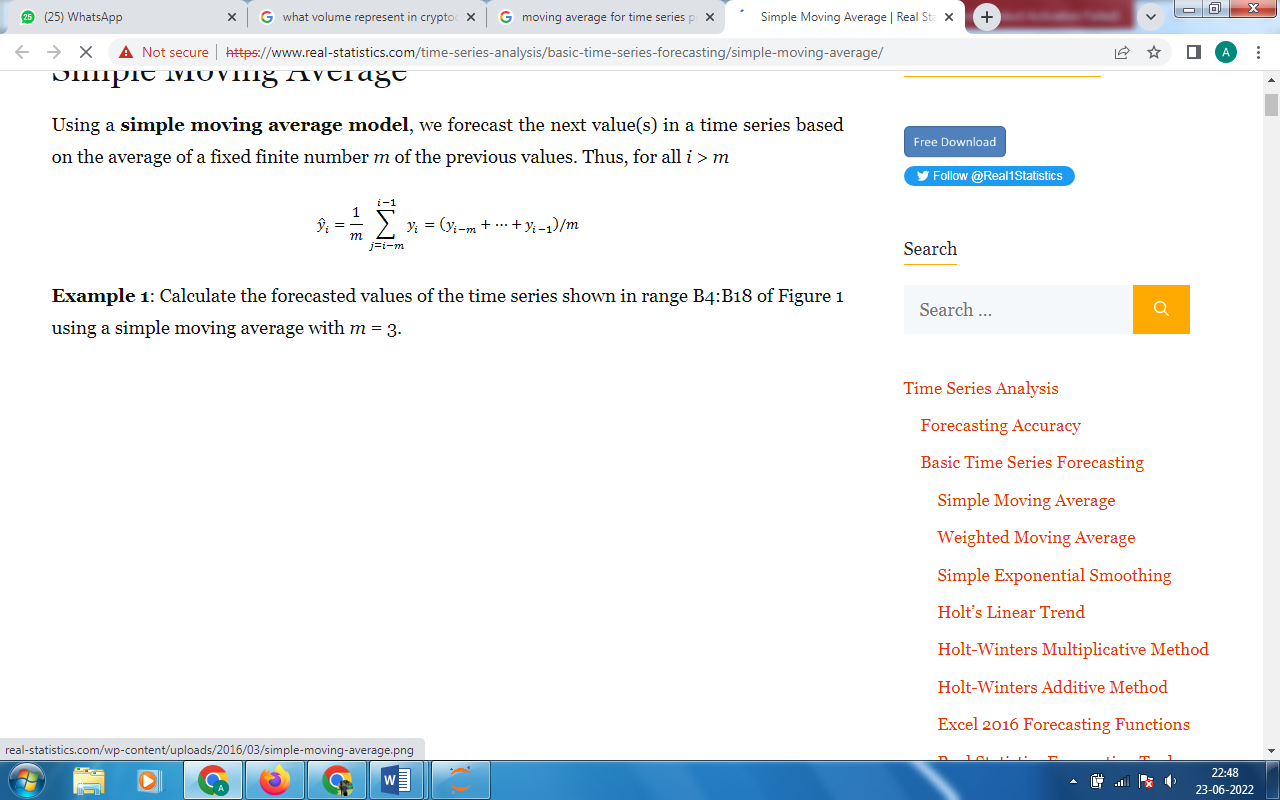
TECHNIQUES

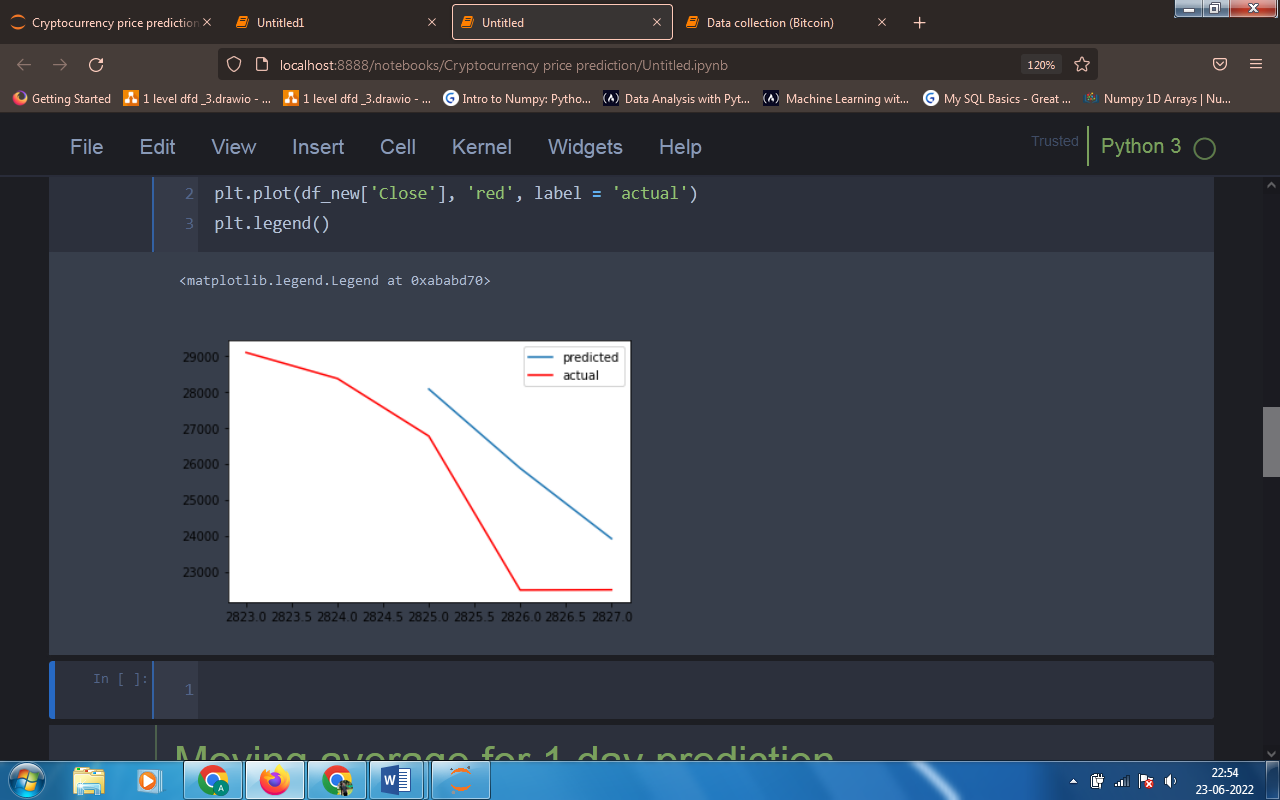
*In this project we used various Statistical, Machine Learning and Deep learning algorithms for prediction purpose.*

* *Moving Average* –

*Moving averages are a simple and common type of smoothing used in time series analysis and time series forecasting. Calculating a moving average involves creating a new series where the values are comprised of the average of raw observations in the original time series.*

*Moving averages have the property to reduce the amount of variation present in the data. In the case of time series, this property is used to eliminate fluctuations, and the process is called smoothing of time series. Formula for calculating moving average is given below.*





* *ARIMA* –

*An autoregressive integrated moving average, or ARIMA, is a statistical analysis model that uses time series data to either better understand the data set or to predict future trends. A statistical model is autoregressive if it predicts future values based on past values.*

*It's a model used in statistics and econometrics to measure events that happen over a period of time. The model is used to understand past data or predict future data in a series.*



*ARIMA model have 3 parameters p, d and q where,*

*p – order of auto-regression*

*d – order of integration*

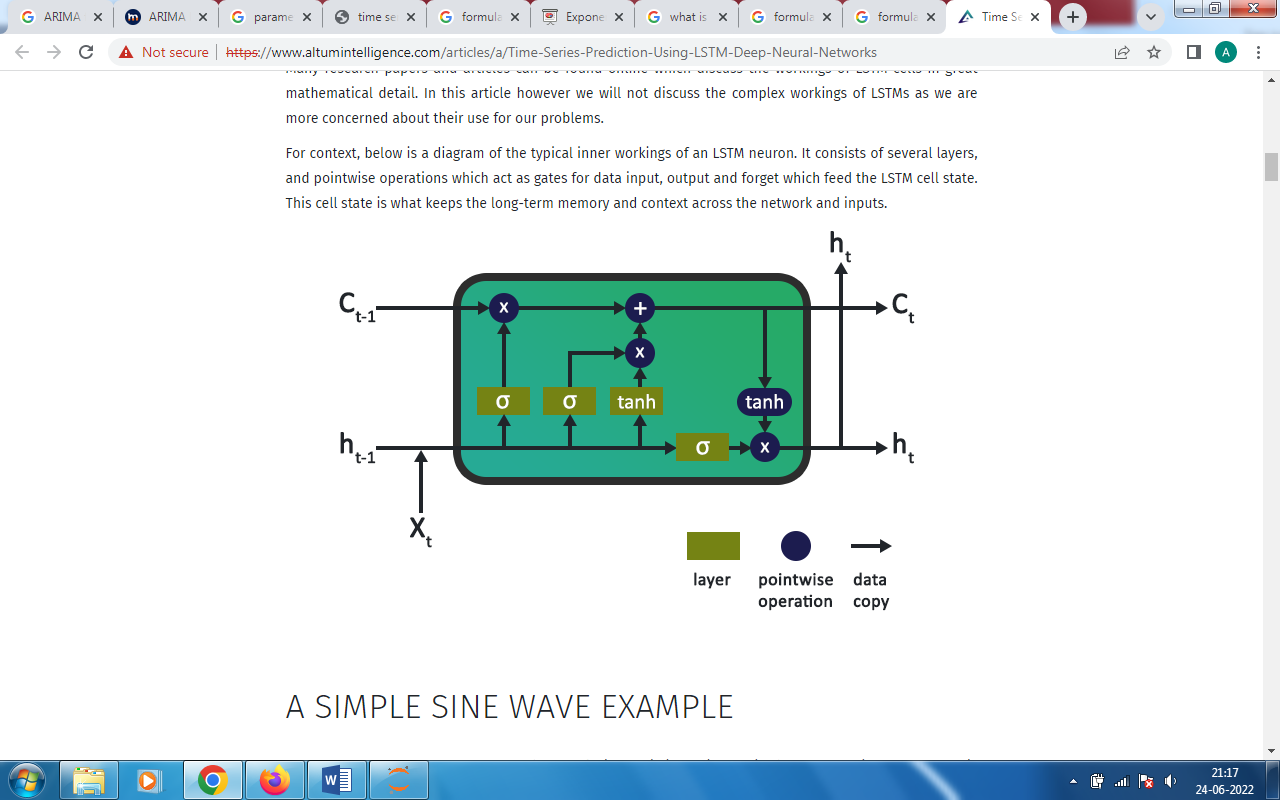
*q – order of moving average*

* *LSTM* –

*Long short-term memory (LSTM) is an artificial neural network used in the fields of artificial intelligence and deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections*

*LSTM cells are used in recurrent neural networks that learn to predict the future from sequences of variable lengths. Note that recurrent neural networks work with any kind of sequential data and, unlike ARIMA and Prophet, are not restricted to time series.*

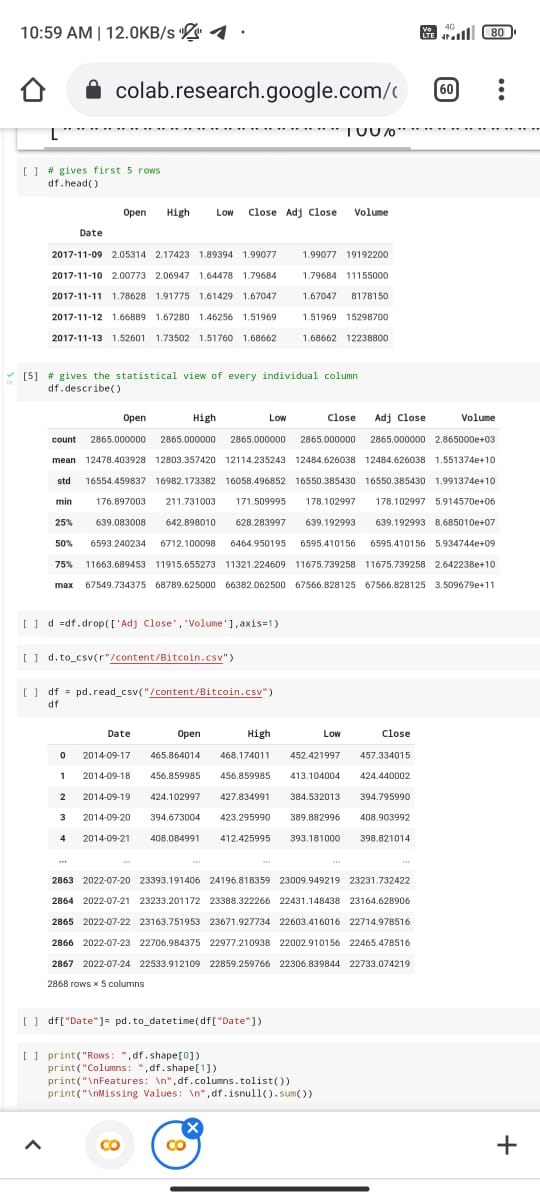
*One of the most advanced models out there to forecast time series is the Long Short-Term Memory (LSTM) Neural Network.*



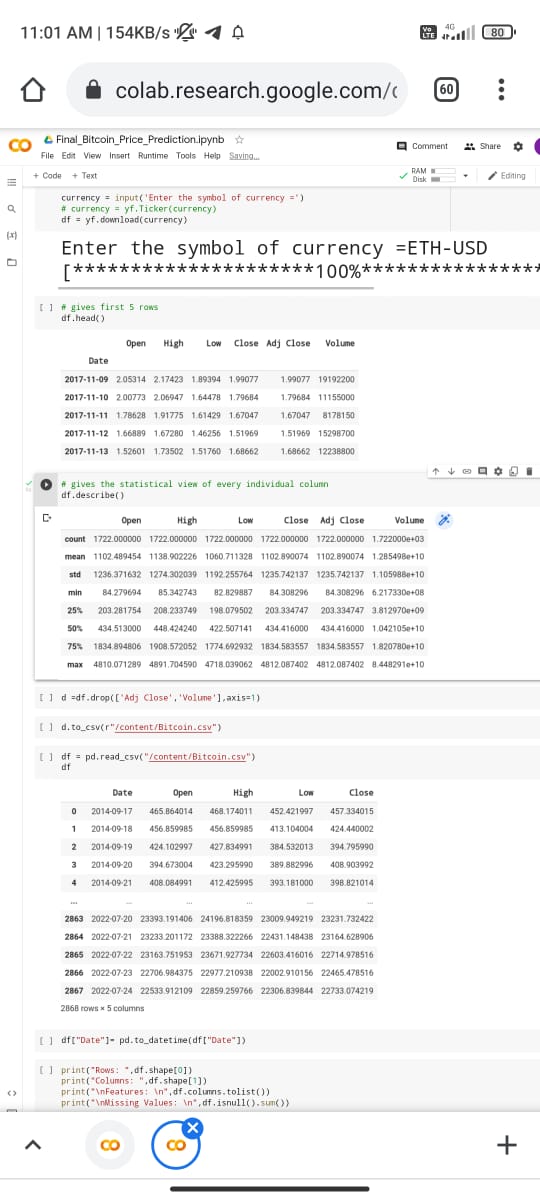
SUMMERY STATISTICS

*Summery statistics give the very clear idea about the data of numeric class. Summery statistics gives us the count, mean, standard deviation, minimum, first quartile, second quartile (median), third quartile and maximum value of each numeric data column.*

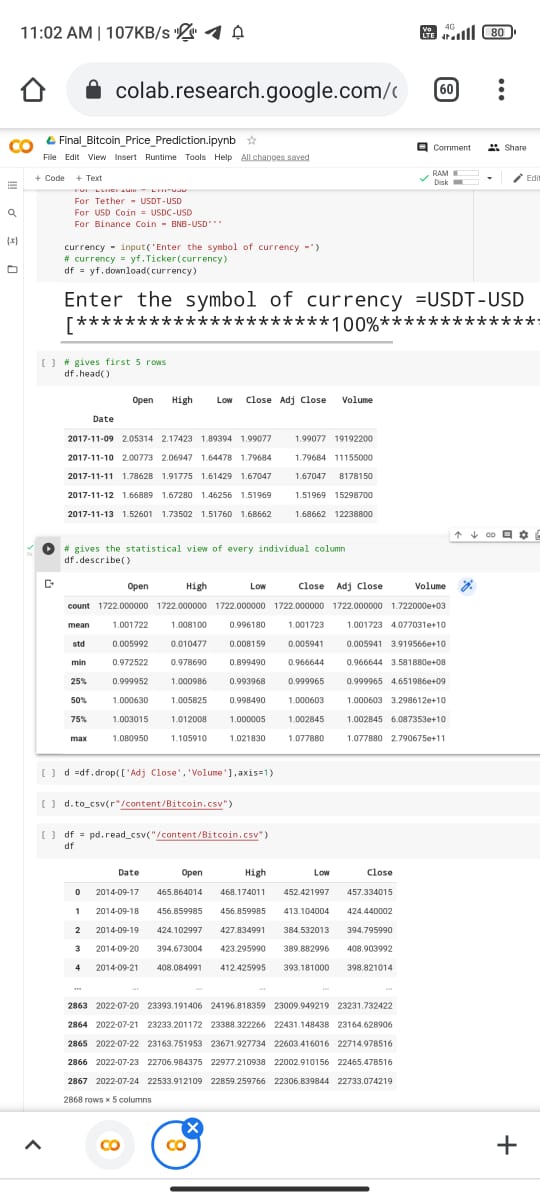
* *SUMMERY STATISTICS OF BITCOIN –*



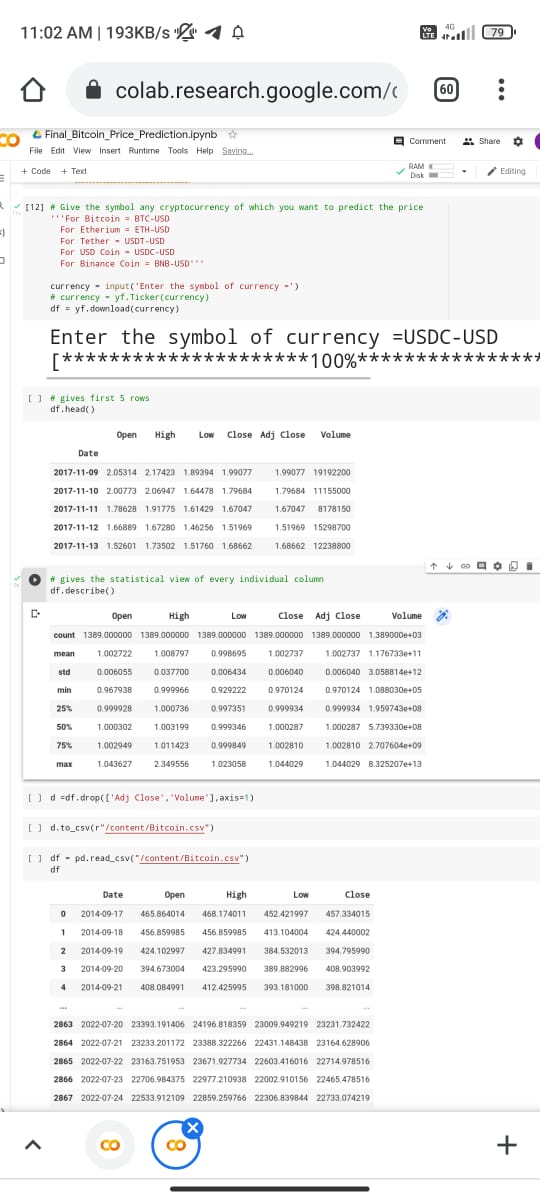
* *SUMMERY STATISTICS OF ETHERIUM –*



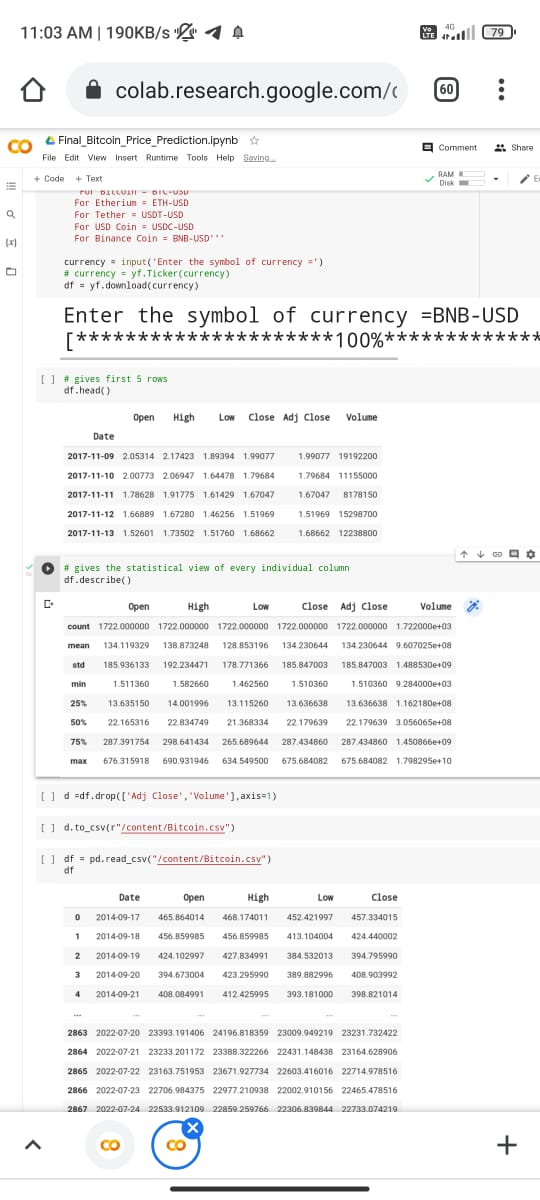
* *SUMMERY STATISTCS OF TETHER –*



* *SUMMERY STATISTICS OF USD COIN –*



* *SUMMERY STATISTICS OF BINANCE COIN –*



*To get this summery we have a ‘describe’ function in python.*

Sampling

*Our Dataset is varies from currency to currency and splitting of data set into training and testing is varies from algorithm to algorithm. In Moving Average technique we have used last 7 days records for training and not used testing data but rather we have directly predicted the next day price.*

*Then for ARIMA model we have used overall data excepting last 3 records used for training and tested on last that leftover 3 days records.*

*In LSTM method the whole data was provided to the LSTM model and predicted the next 7 days price of each column.*

FLOWCHART

**Data Collection:-**

**Data Visualization:-**

***Visualize the data to get more idea about data.***

**Train-Test Splitting:-**

***Used last 3 records for testing and other previous for training.***

**Moving Average:-**

***Used simple moving average for next one day prediction.***

**ARIMA:-**

***ARIMA model for prediction of next 3 days values.***

**LSTM:-**

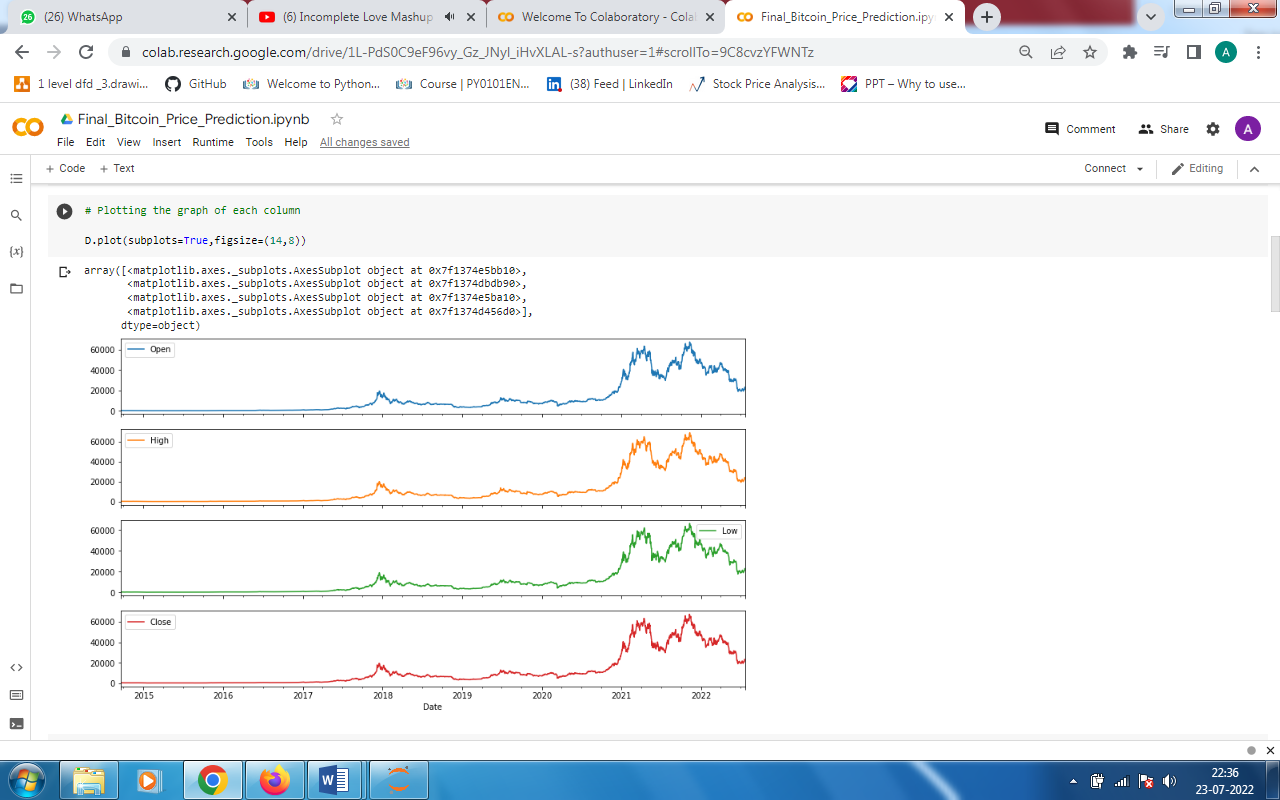
***Deep learning model for predicting next 7 days prices.***

***Data is collected from yfinance library.***

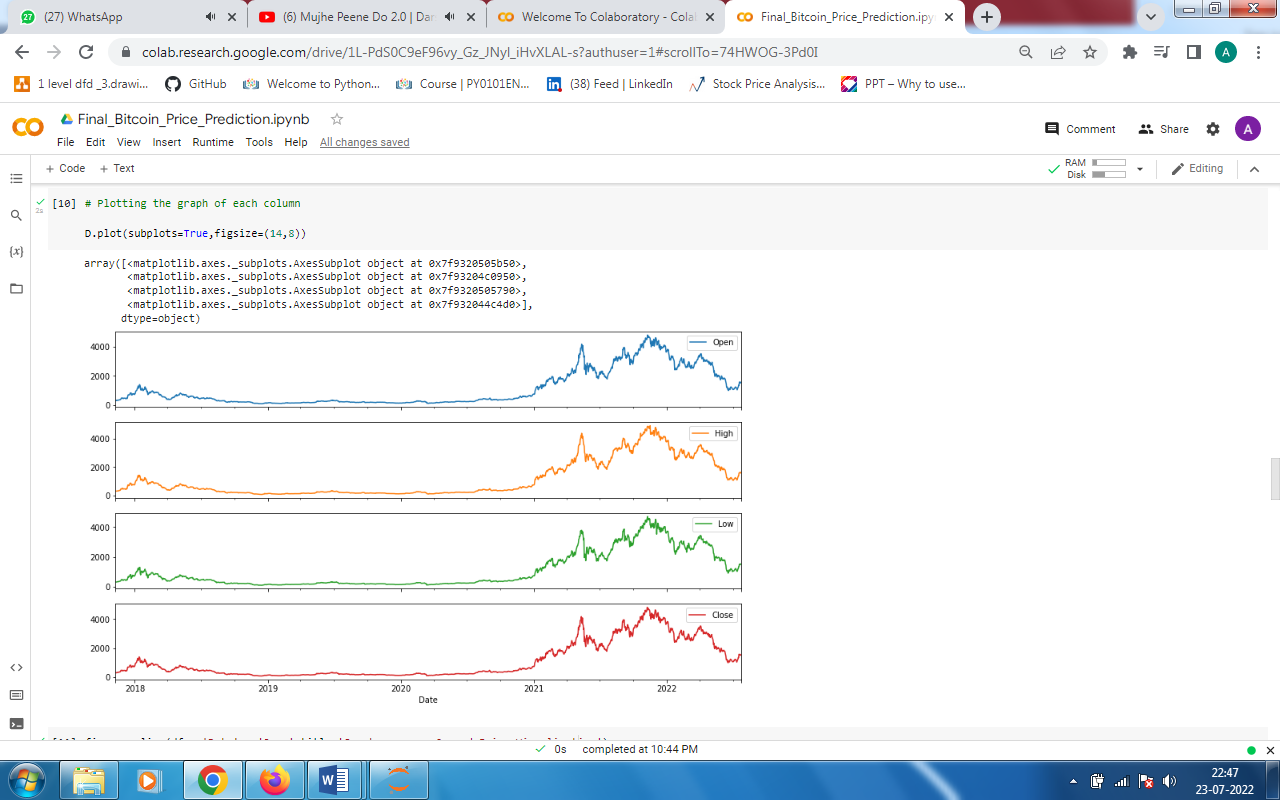
EXPLORATORY DATA ANALYSIS

*It always good for use visualize the data rather than using numbers for comparison. By visualizing the data we get more clear idea about the data. So data visualization we used Matplotlib,* *Ploty.express and Seaborn Libraries in Python.*

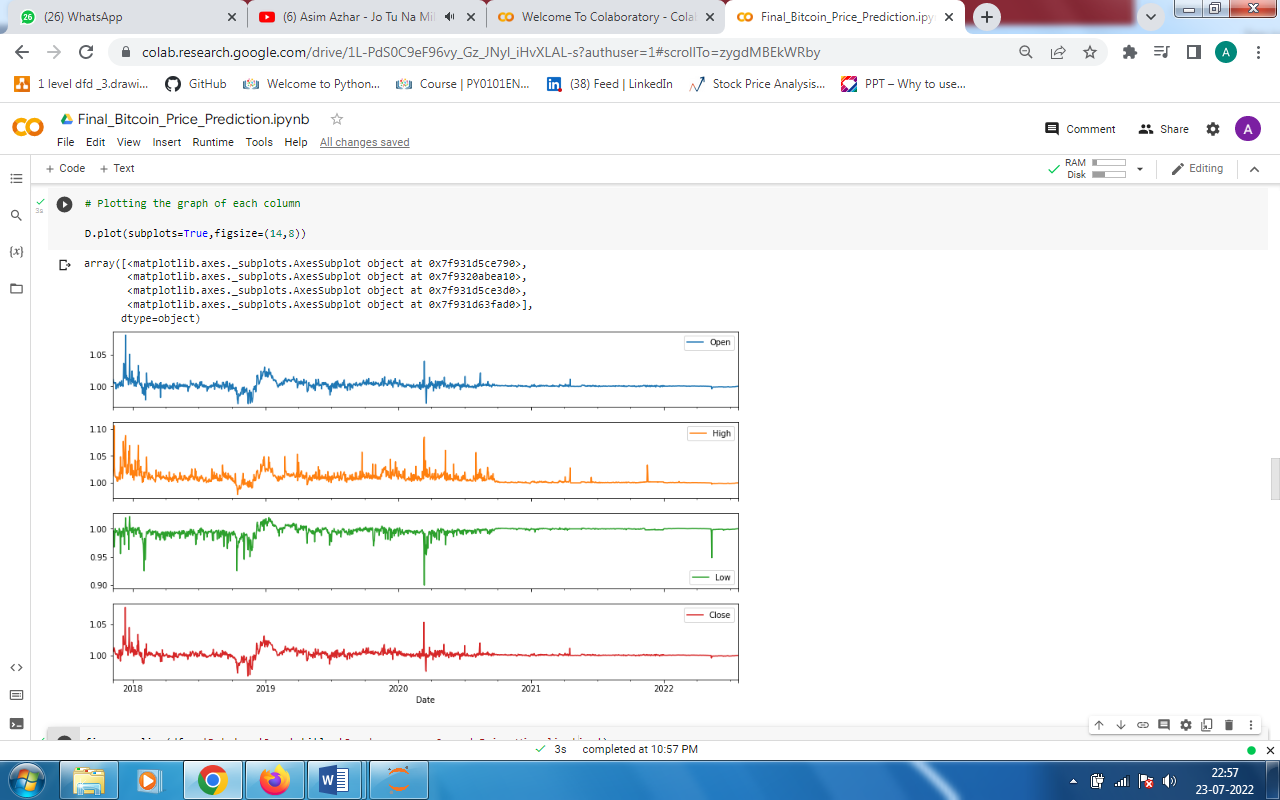
* *Line Plot of Bitcoin of previous records –*



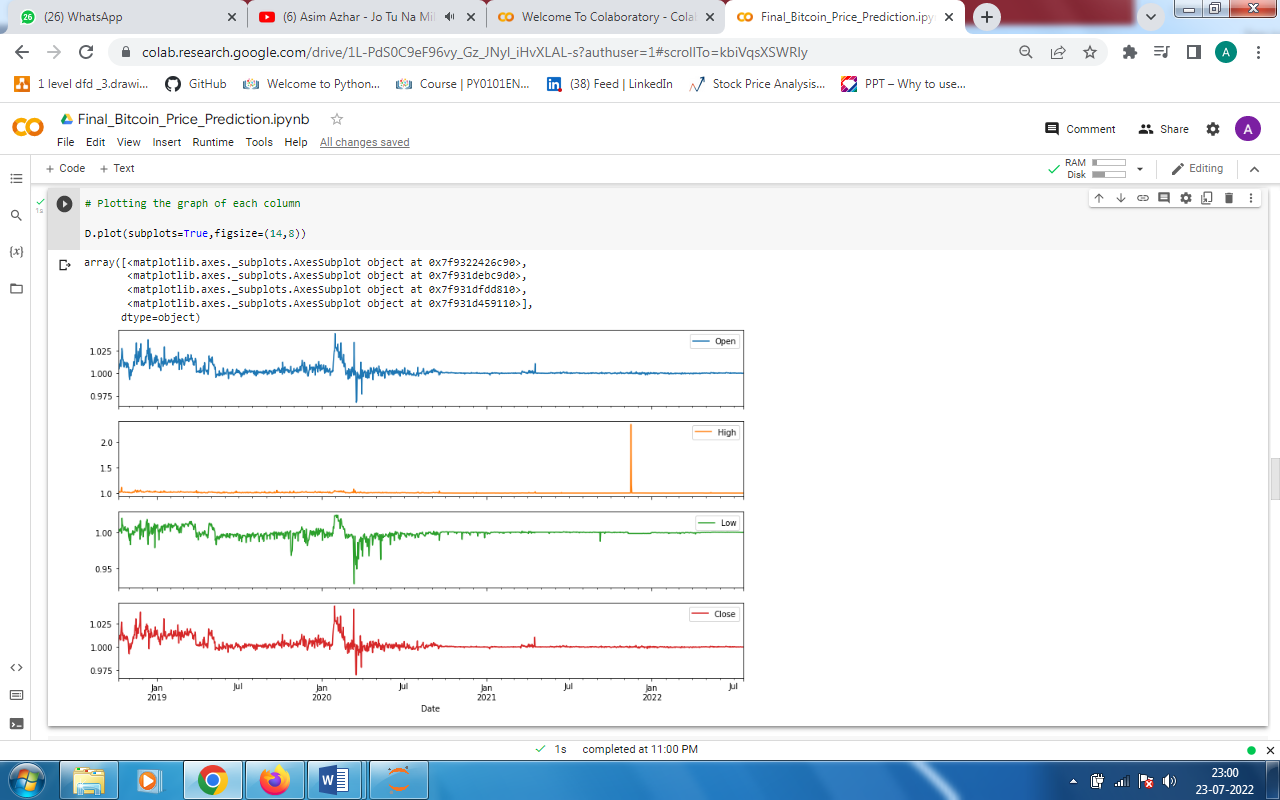
* *Line Plot of Etherium of previous records –*



* *Line Plot of Tether of previous records –*



* *Line Plot of USD Coin of previous records –*



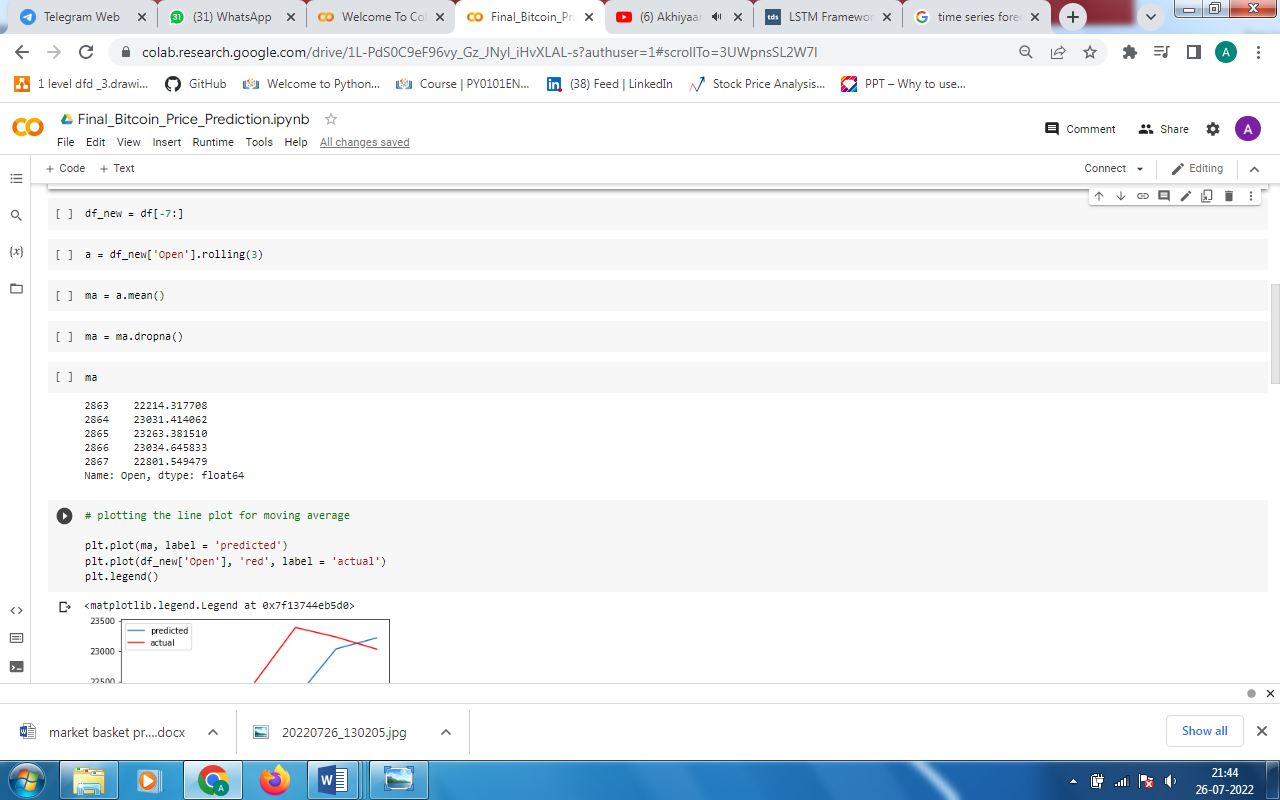
* *Line Plot of Binance Coin of previous records –*

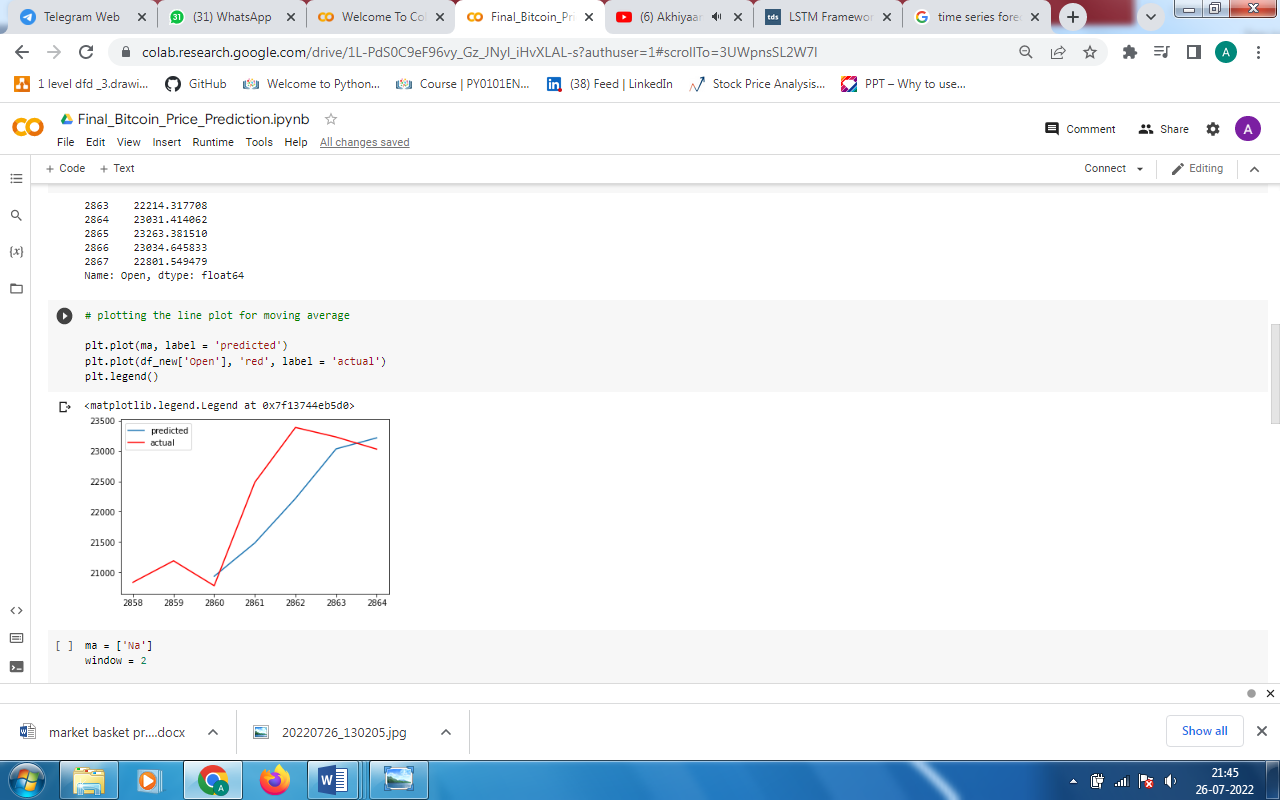


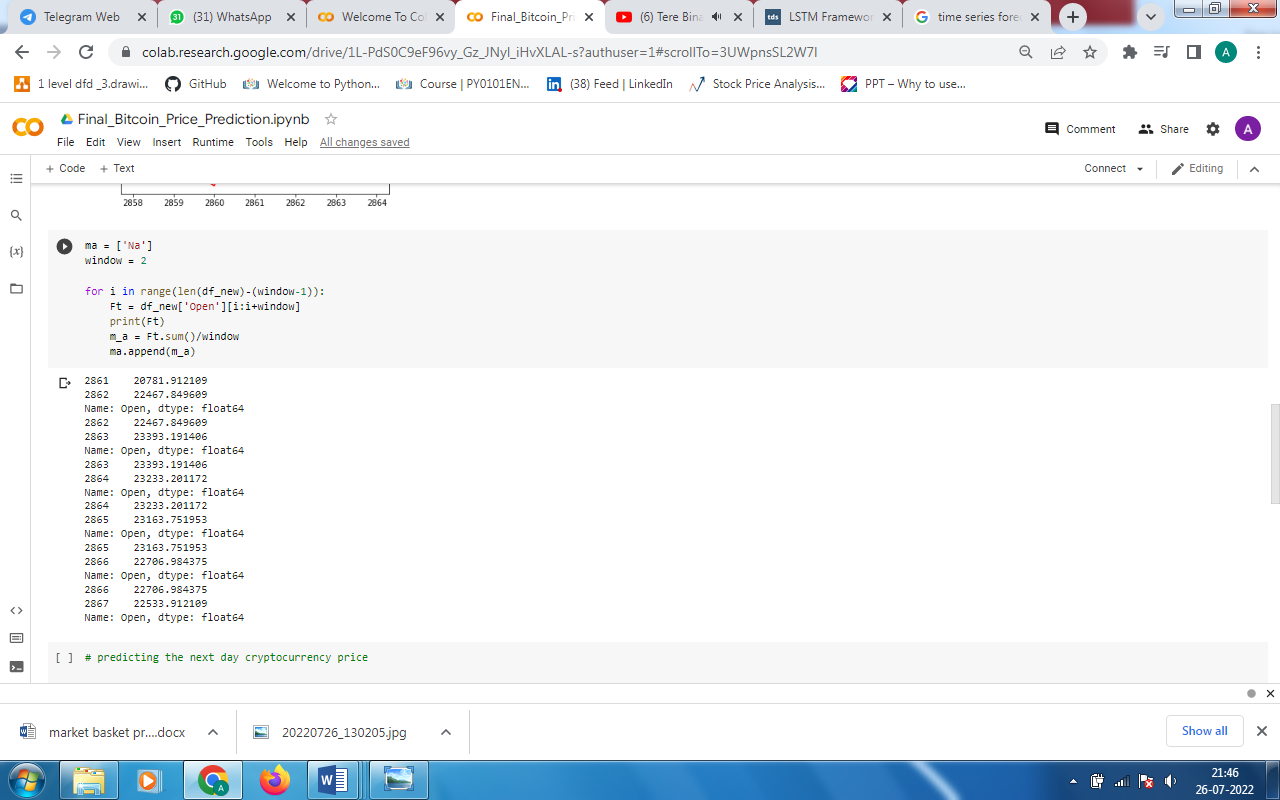
*Above graphs shows the 4 subplots of are ‘Open’, ‘High’, ‘Low’ and ‘Close’ respectively and of 5 different cryptocurrencies which have selected for the prediction purpose of prices for our project. So from the above graph we can get the idea of flow of history price variation sequence up to today’s date.*

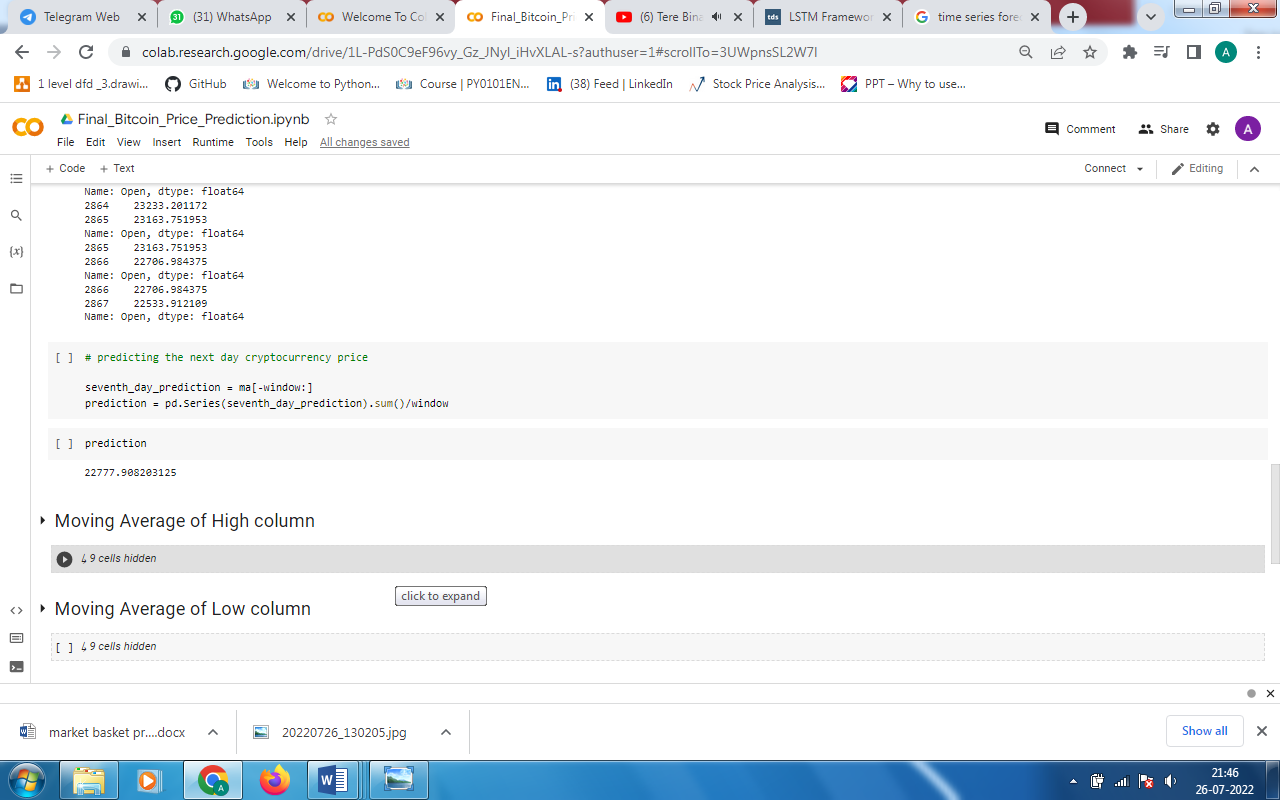
SCREENSHOTS

* *Code of Moving Average –*



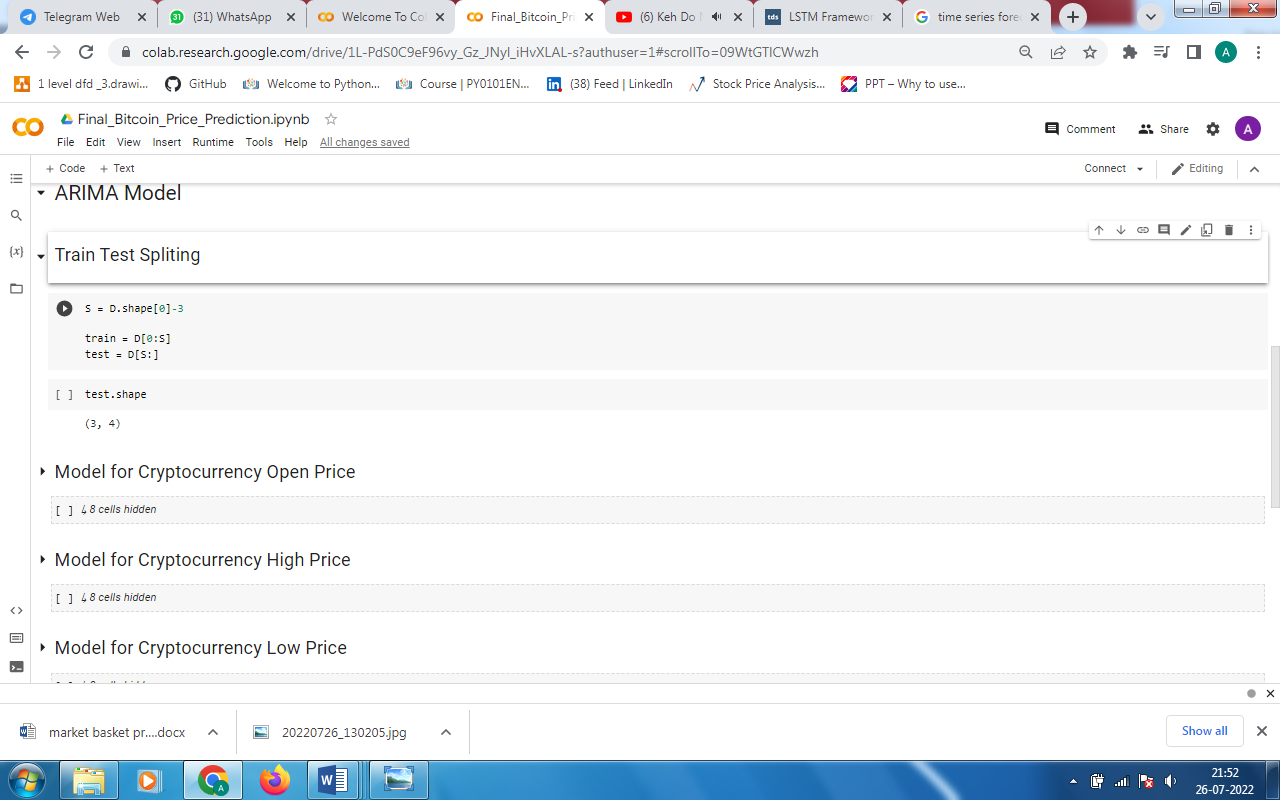


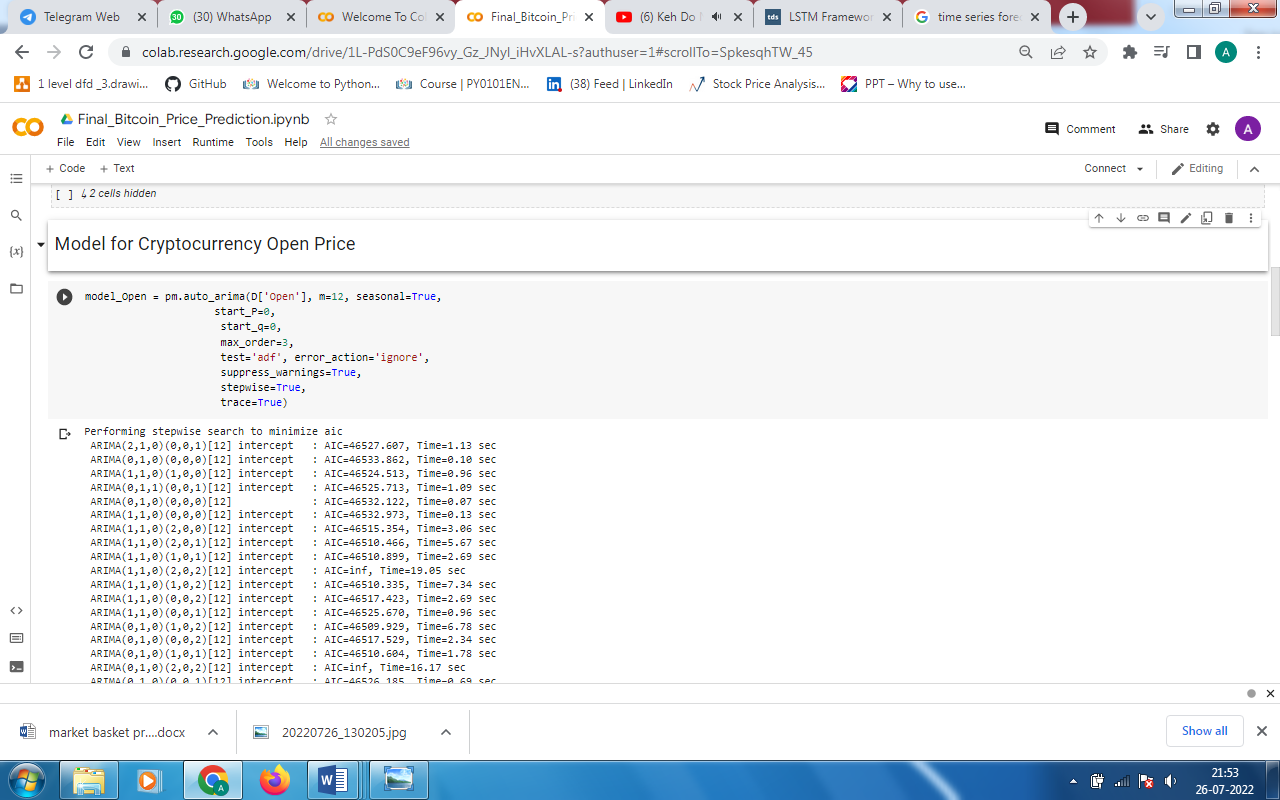


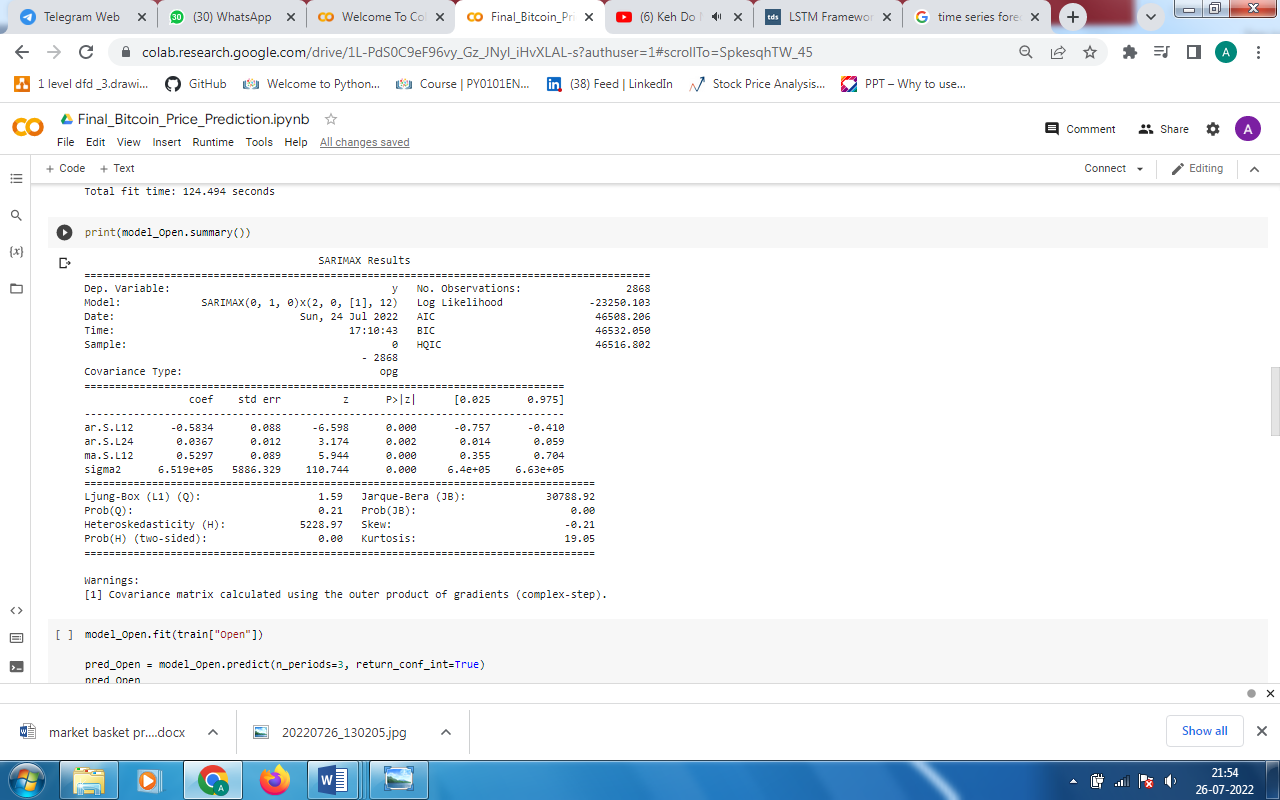


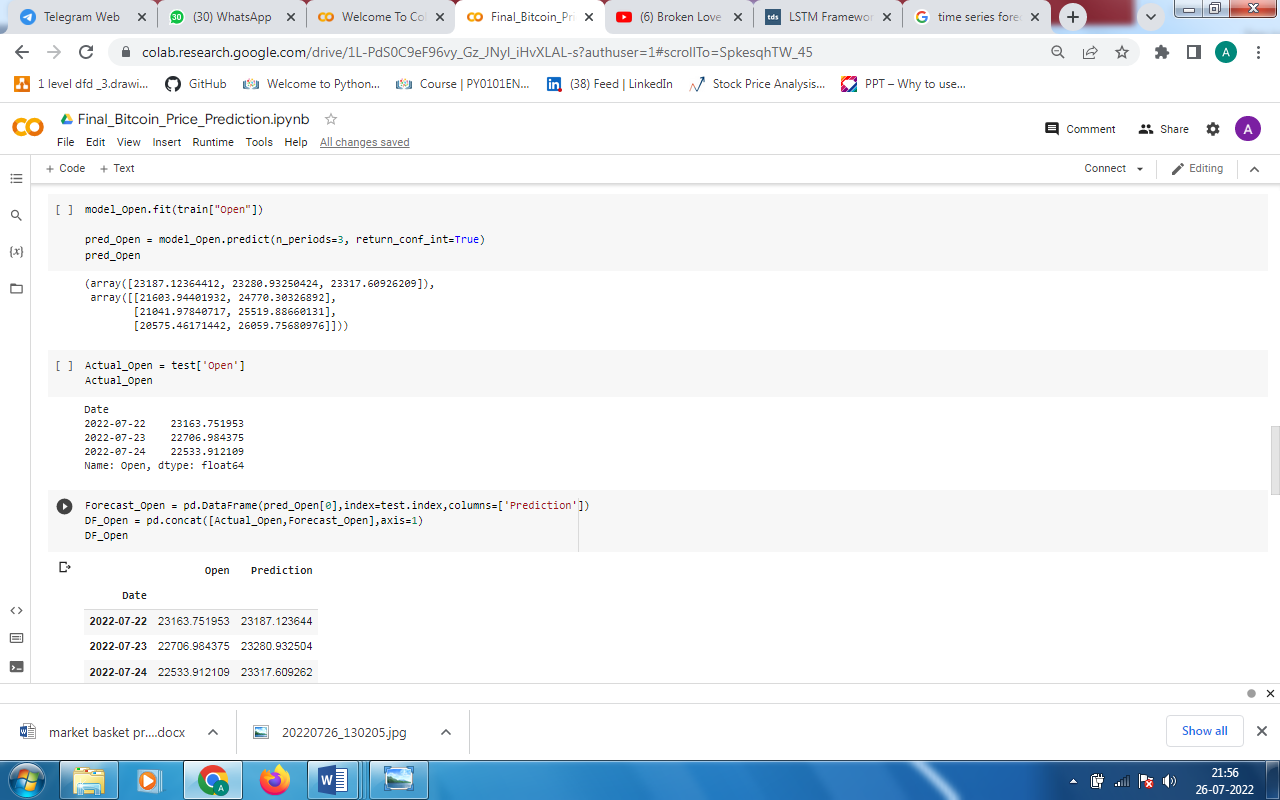
*The Moving Average technique is used for every four columns that is ‘Open’, ‘High’, ‘Low’ and ‘Close’ respectively. And given the next day price. In this technique we have used window is 2.*

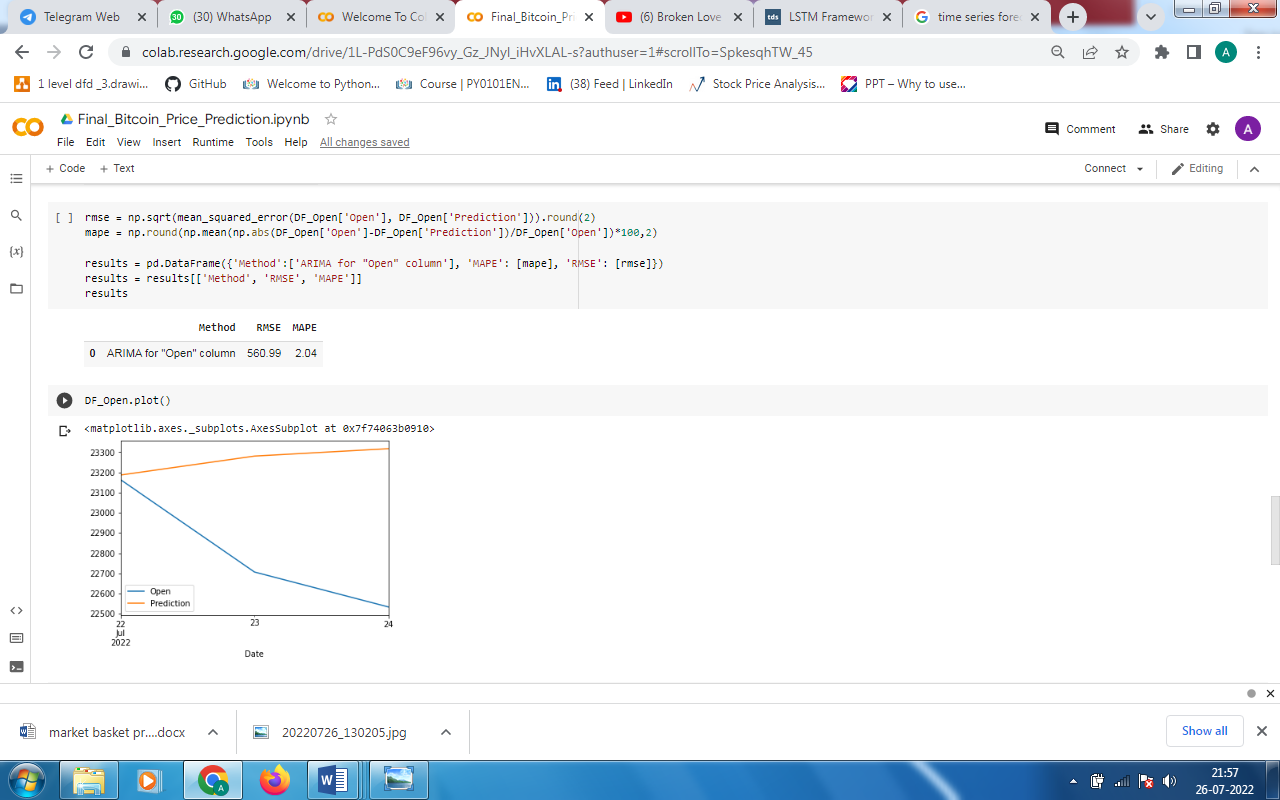
* *Code of ARIMA –*

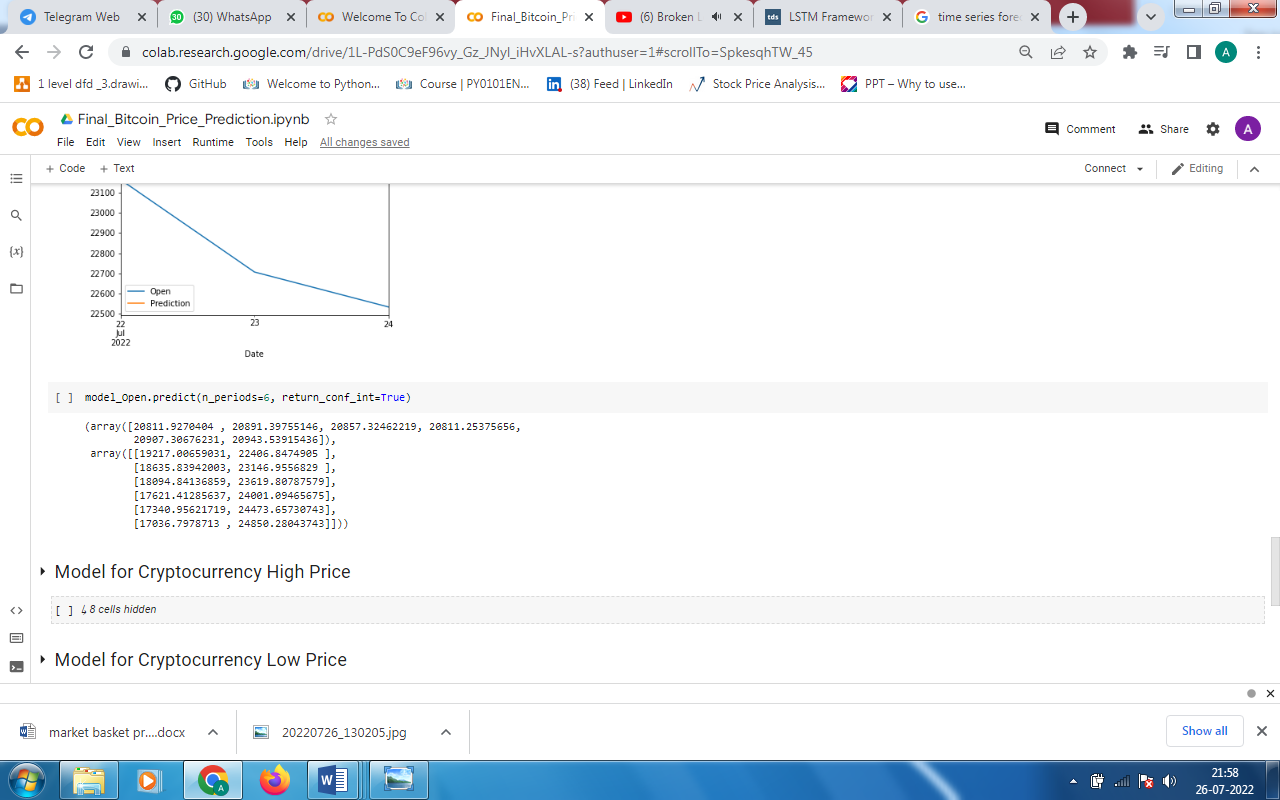






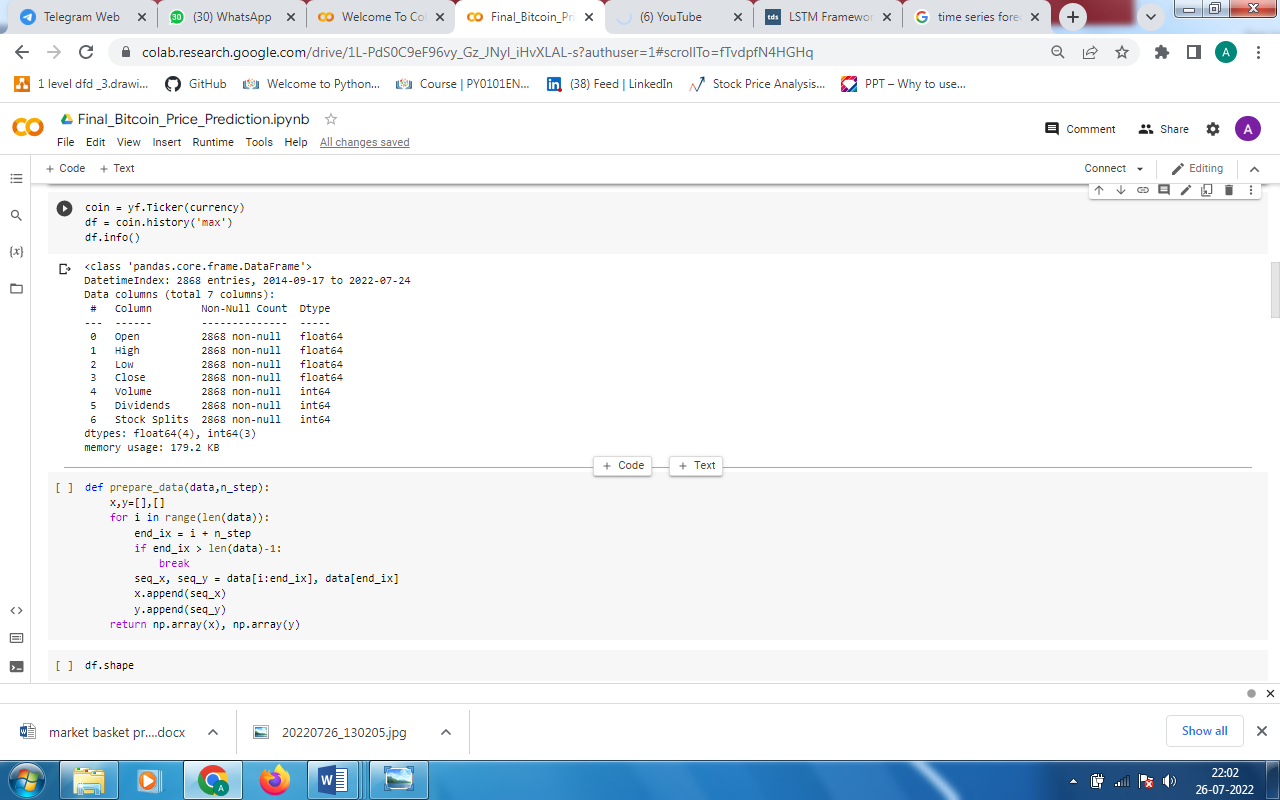


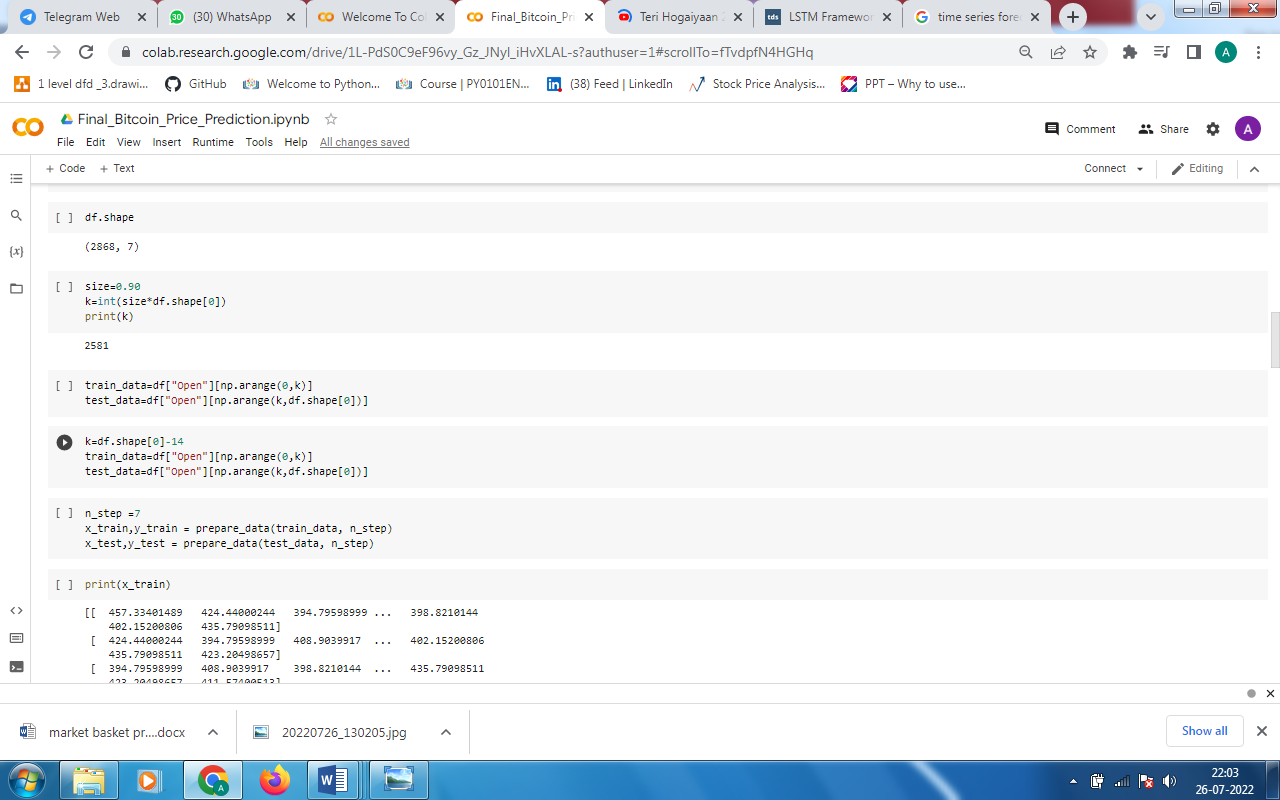


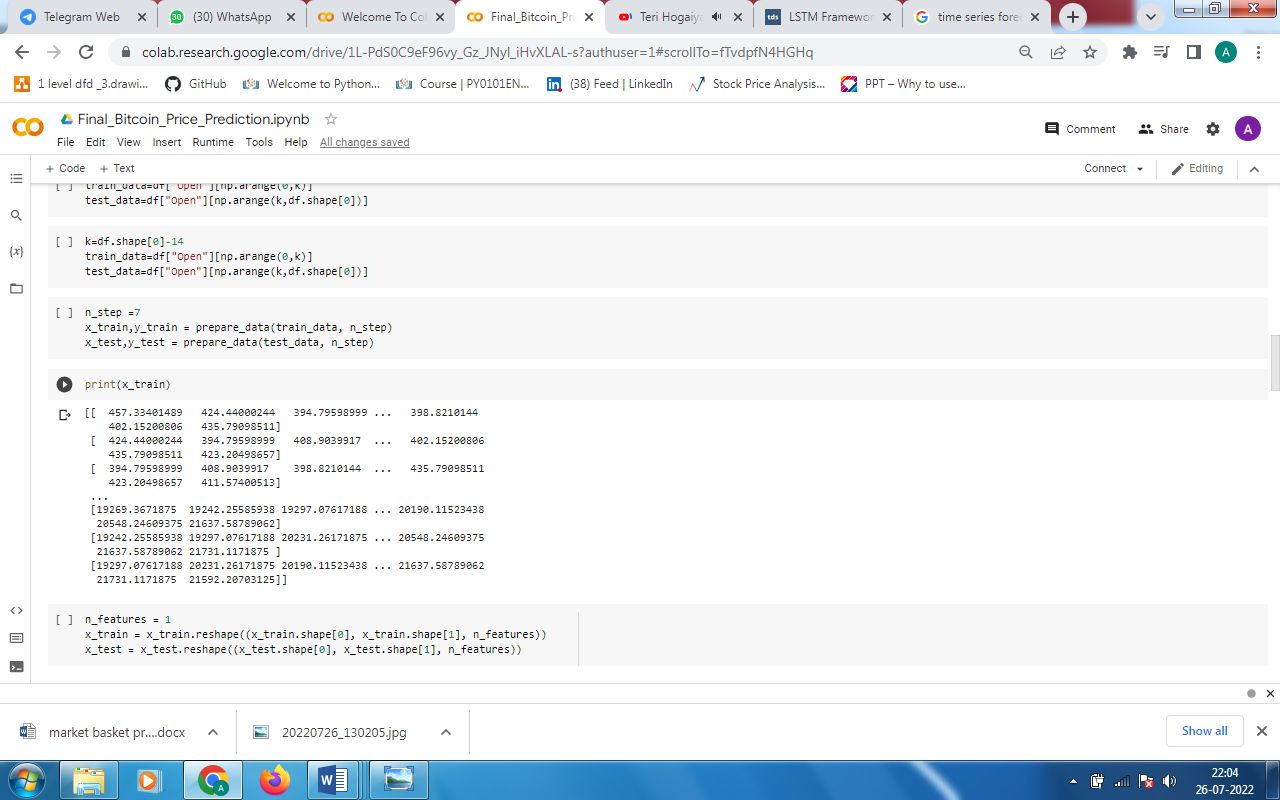


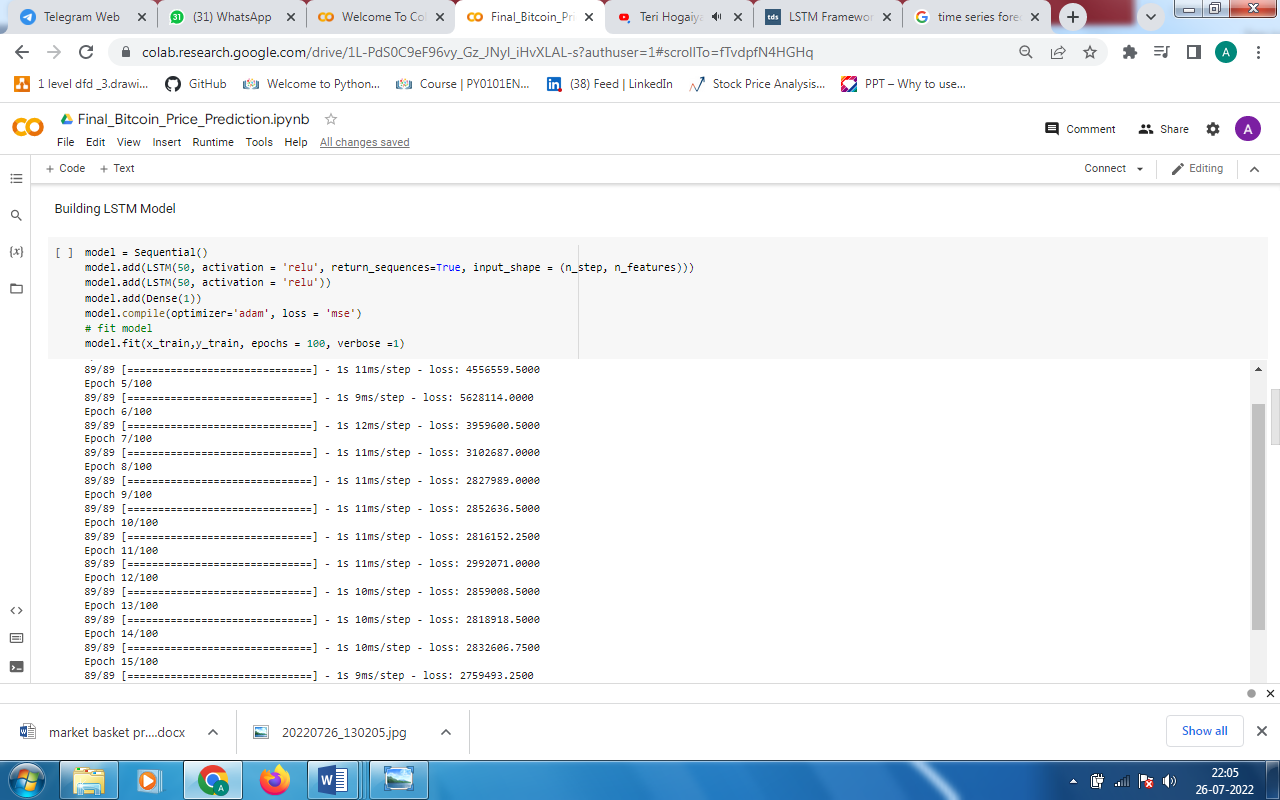
*The ARIMA technique is used for every four columns that is ‘Open’, ‘High’, ‘Low’ and ‘Close’ respectively. And given the next 3 days prices.*

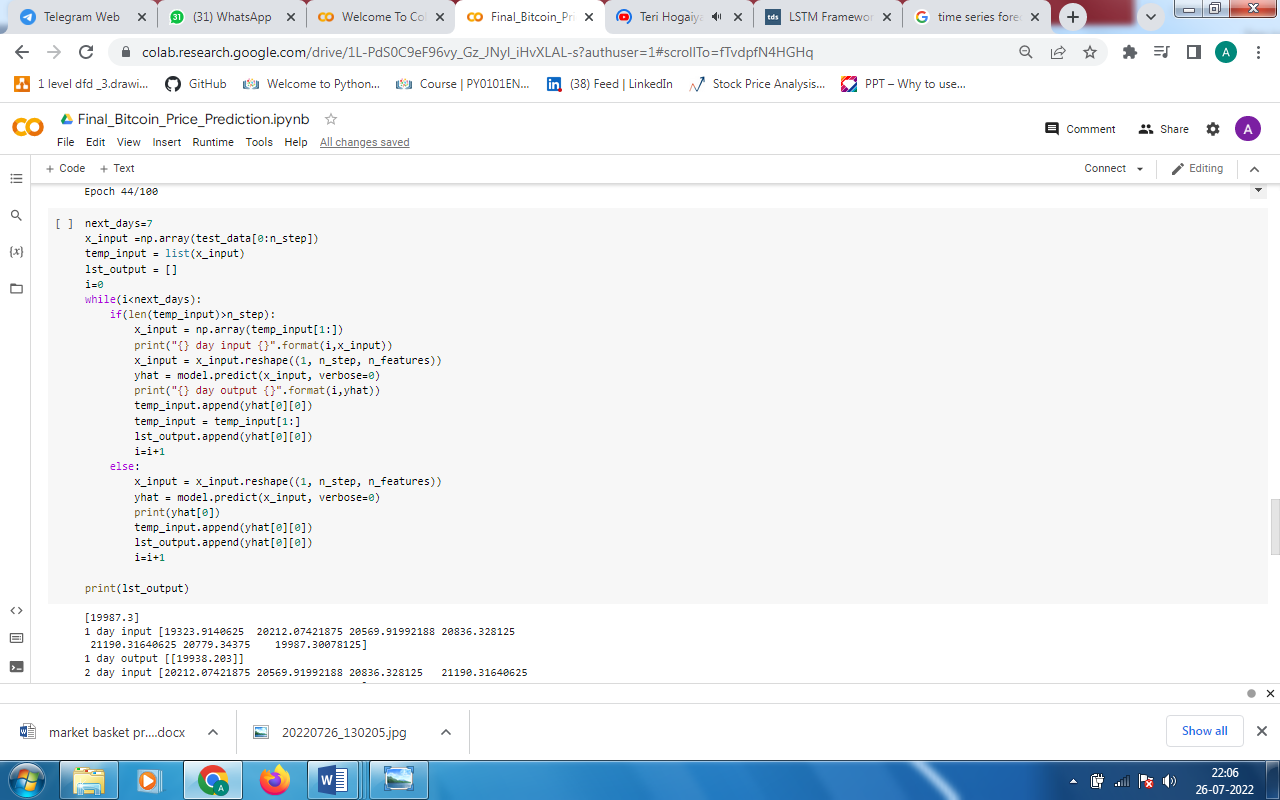
* *Code of LSTM –*

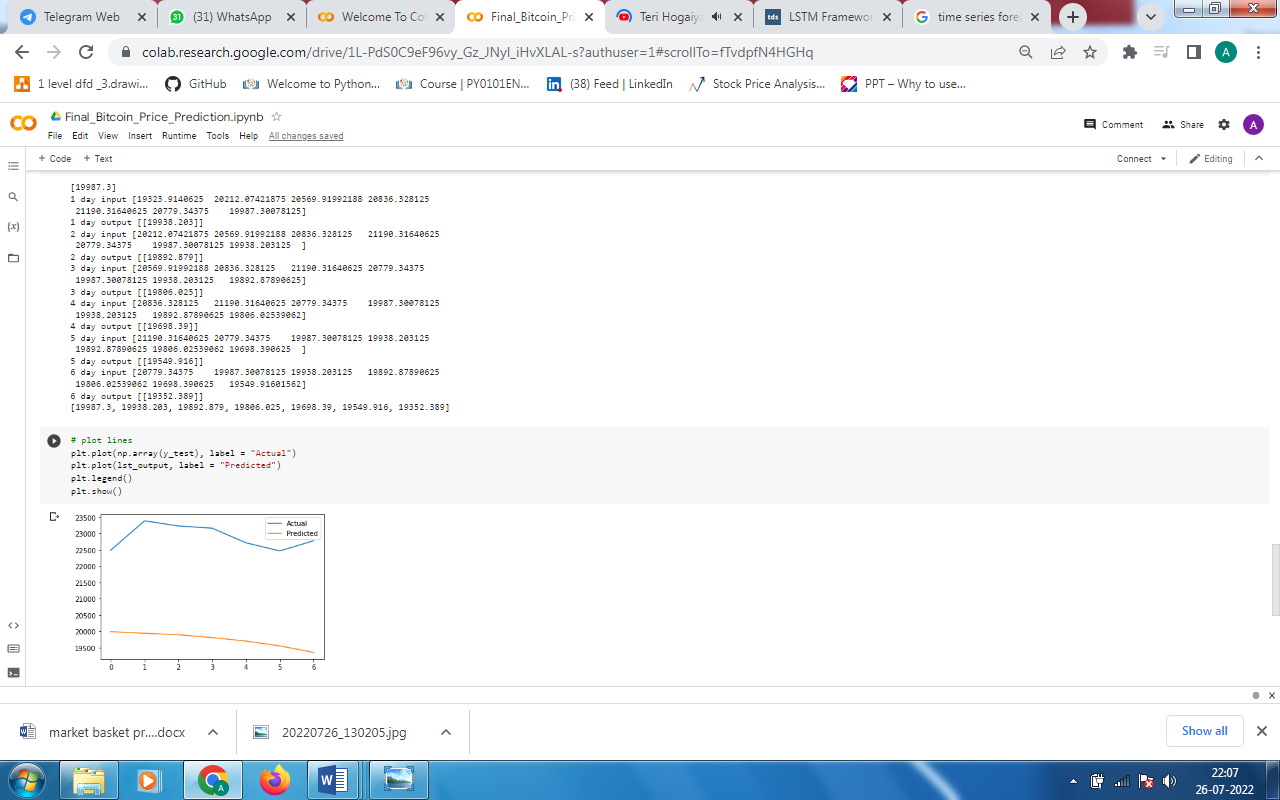












*The LSTM technique is used for every four columns that is ‘Open’, ‘High’, ‘Low’ and ‘Close’ respectively. And given the next 7 days prices.*

SYSTEM REQUIREMENTS

* *HARDWARE –*

1. *Graphic Card – NVIDIA : 2 GB| i5*
2. *RAM – 8GB*

* *SOFTWARE –*

1. *Windows10*
2. *Programming language- Python3*
3. *IDE – Jupyter*
4. *Google colab*

CONCLUSION

*From all over study of cryptocurrency price prediction methods we can conclude that the nature of crypto is very complex to understand and uncertain. In our project we have used Moving Average, ARIMA and LSTM models to prediction of prices. Among the three used models ARIMA gives the good accuracy and achieves the smallest mean squared error and mean absolute error on the test data, so ARIMA is used for more day price prediction as compared to LSTM and MA. But for just one day prediction MA also gives good result. And to get more correct prediction we provided the confidence interval for ARIMA model in which the price is mostly likely to lie. So we can prefer ARIMA model for forecasting purpose.*

REFERENCES

*1. Crypto Currency Prediction Model using ARIMA. Vidyulata, G., M., Mounika and N., Arpitha. s.l. : Turkish Journal of computer and mathematics education, 2020, Vol. 11.*

*2. Cryptocurrency price prediction using time series forecasting (ARIMA). Kynary, Sampat, Aanandhi, S. P. and Akhila, S. P. s.l. : International seminar on research of information technology and intelligent systems (ISRITI), 2021.*

*3. Time Series Forecasting Using LSTM Networks: A Symbolic Approach. Elsworth, Steven and Güttel, Stefan . s.l. : Cornell University, 2020.*

*4. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Zhang, Lei, Chimmula, Reddy and Kumar, Vinay . s.l. : ELSEVIER, 2020, Vol. 135.*