*Predicting Stock Price using Machine Learning:*

*Tesla*

*Team 22:*

*Jayant Kumar*

*Varun Kandhari*

*Angad Singh Malhotra*

**Table of Contents:**

1. Introduction 2
2. Understanding the Model 4
3. Tesla Stock Price using ARIMA Model 6
4. Conclusion 13
5. References 13

*Introduction*

**What is the stock market?**

The stock market is a public market where one can buy and sell shares for publicly listed companies. The stocks, also known as equities, represent ownership in the company. The stock exchange acts as a mediator who allows the buying and selling of shares.

**Importance of Stock Market**

· Stock markets help companies to raise capital for growth purposes.

· It helps in generating personal wealth.

· It is a widely used source for people to invest money in companies with high growth potential.

· Stock markets also serve as an indicator of the state of the economy.

**Why is stock market prediction important?**

· Equity securities are one of the most traded securities as they have an attractive return and are a relatively liquid asset.

· Stock market prediction aims to determine the future movement of the stock value of a financial exchange. The accurate prediction of share price movement would eventually lead to more profit for investors.

· It contributes to the development of effective strategies for stock exchange transactions.

· There is an increasing number of people who are using deep learning and Neural networks to assist in predicting stock prices.

**Popularity of the Stock Market in the eyes of retail investors**

In recent times the stock market has seen an influx of new retail investors entering the stock market. Many factors have spurred the influx of new retail investors, including stimulus checks, the ability to save during pandemic lockdowns that curbed activities and the availability of free or low-cost trading. In addition, the performance of the so-called meme stocks driven in part by investors from r/WallStreetBets, a popular Reddit thread, and growth of cryptocurrencies such as dogecoin piqued interest for some.

The GameStop trading frenzy is the most explosive and apparent example of a shift in market power that has been building in plain sight for more than a year.

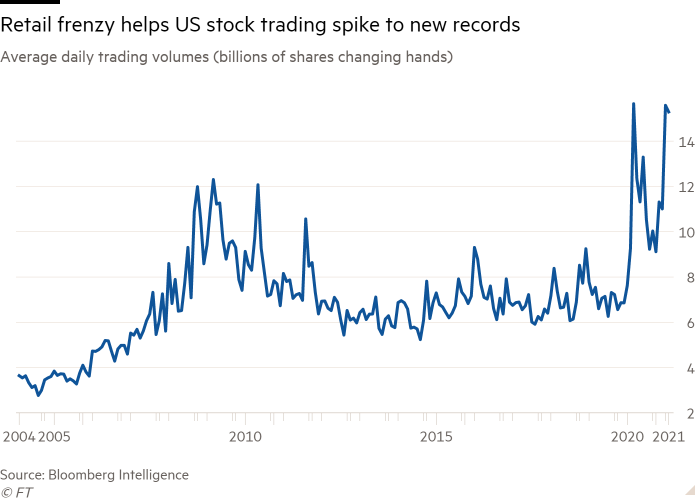
In the below chart we can see the increasing volume of trades by retail investors.

**Why Tesla Stock ?**

· Tesla has seen a high growth in its stock. In the last year alone, Tesla's share price has rocketed upward more than 700% and since it went public in 2010, the stock price has increase more than a whopping 20,000%

· Tesla has beaten the Wall Street delivery estimates multiple times in 2020, which has cultivated a lot of confidence in the shareholders.

· In recent times, there is a lot of interest in investors and fund managers in Tesla stock.





*Understanding the Model*

Predicting the Stock Market holds great significance as evidenced by the previous explanation in the study and to predict the future of the Stock Market or particular stock, in this case Tesla, Autoregressive Integrated Moving Average Model(ARIMA) is applied. As the name of the model suggests it’s the combination of Two models i.e. the Autoregressive Model and Moving Average Model, but another aspect to the model is the ‘Integrated’ which deals with non-stationarity issues in data. Before delving deeper into the technical aspects of the study, an understanding of components of the ARIMA model is necessary.

ARIMA Model = Autoregressive Model + Integrated + Moving Average Model

**Autoregressive Model:**

This model takes into account the impact of previous trends of the same variable on itself. As this study pertains to Tesla’s Stock Price, this model assumes that the previous trend of the stock price impacts the current and future movement in stock price of Tesla. Thus, it takes in lagged variables of Tesla Stock to a range that is permissible to predict the stock. The equation of the model:

is the constant and the to are the coefficients of the lagged variables, to are the lagged variables from time(t) -1 to time(t) - p. is the residuals/unobservables term which suggests the residual value at each predicted data point, meaning . Now the question comes how many lagged variables are to be selected for the model, this is answered by the Partial Correlation Function ( PCAF ) where the last number of lagged values is chosen which is above the threshold limit of PCAF.

**Moving Average Model:**

It’s not necessary that a variable wouldn’t have shock/seasonal periods or trends. Such shocks could either have a positive or a negative impact on the variable and sometimes they can have a large impact as well. So to account for them the Moving Average Model is implemented. Tesla’s Stock price previously had abnormal movements, one instance being the dip in stock price due to the split of the stock. It is imperative that such an anomaly is acknowledged if the stock price is to be predicted. Thus, the model sets a relation between such shocks with the dependent variable and produces coefficients as follows:

The Stock price Y is predicted at time t, where is the mean of the series and to are the shocks. to being the shock coefficients, that measure the relation of the shocks with the dependent variable.

**ARIMA Model:**

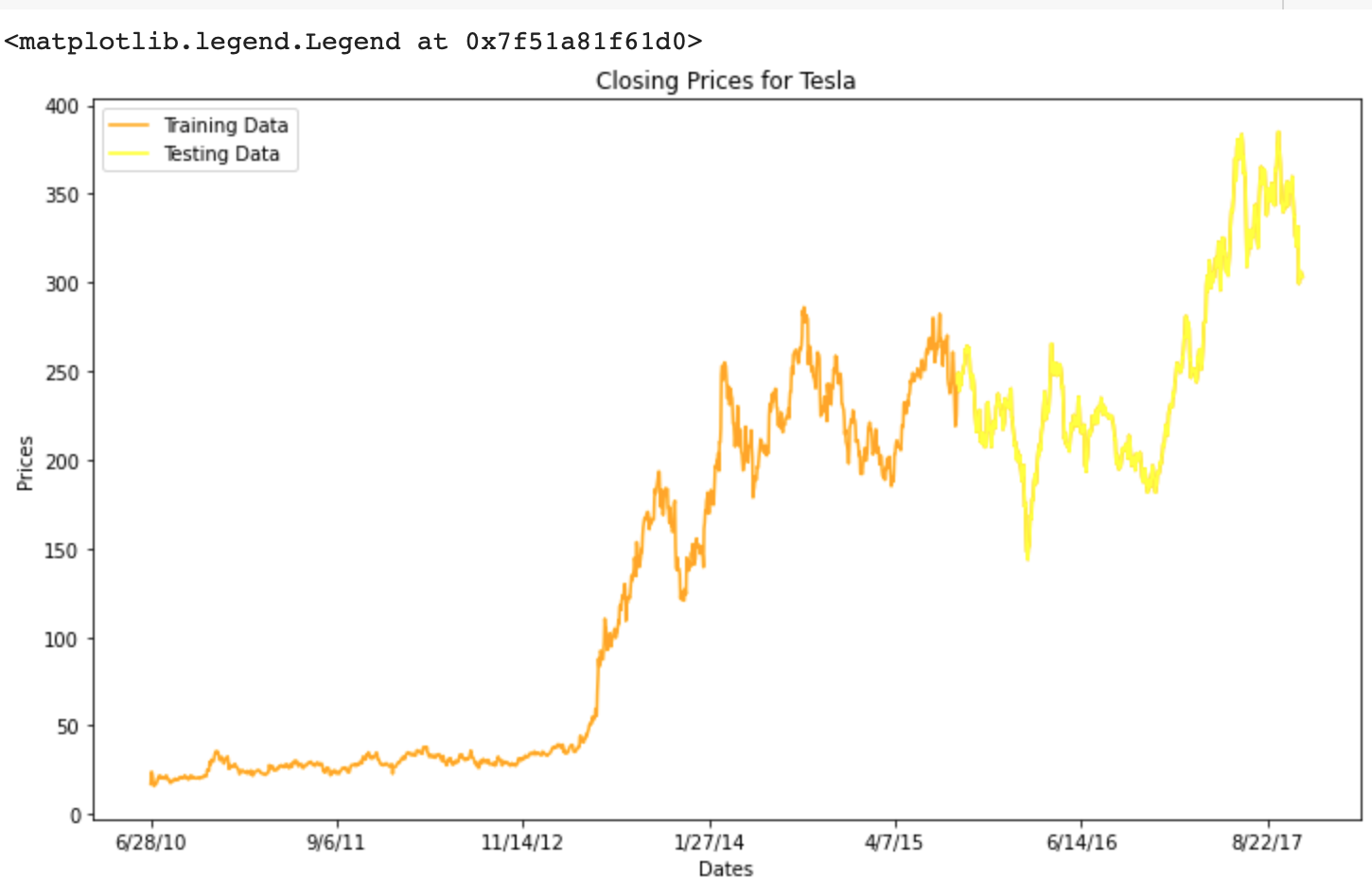
As previous pointed out the ARIMA model is a combination of the two models stated above and the equation after the amalgamation shall be as follows:

The value will be removed as the holds the information about the average value of the series.

But to use the ARIMA model some **assumptions/requirements** are to be satisfied:

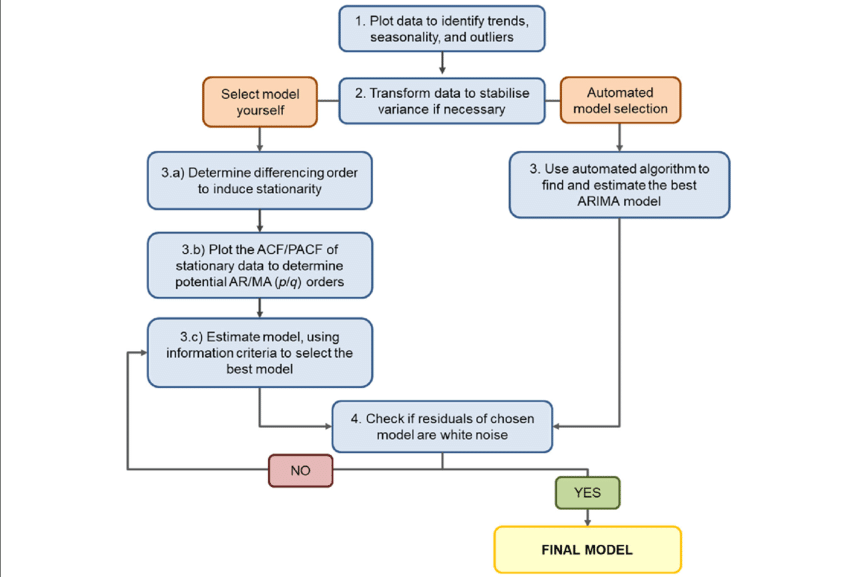
1. Univariate Variable - A single variable has to be studied by the ARIMA model, that is the model cannot have multiple Independent and Dependent variables. Just one Independent variable should be considered.
2. Stationarity in Data - The data should be stationary for the model to work. Stationarity means the mu and variance of across different periods in the series should be the same. For example, Tesla Stock price cannot have a different mu and variance for the year 2012 and 2013. If in case the mu and variance of aren’t the same, the data is made stationary by applying differencing. Differencing is done by deducting the current term with its previous term if Differencing is 1 and so on.
3. Linearity in Parameters- The parameters/coefficients that will predict the stock price at different values should be linear. For example, the nature of the parameters shouldn’t be quadratic.

To apply the model obviously these requirements will be taken care of. Firstly, the condition of the Univariate variable is satisfied as the variable that is considered for the study is just the Closing Stock Price of Tesla. The second condition of stationarity is something to be taken care of.



Evidenced in the plot of Closing Prices for Tesla, the stationarity assumption is obviously being broken because the level of price is drastically changing between year 2012 to 2014. To bring the mu and variance to the same level, differencing will be used here which will be dealt with later. Lastly, Linearity in parameters assumption will be automatically satisfied as the coefficients that will be generated after applying the model will be linear in nature.

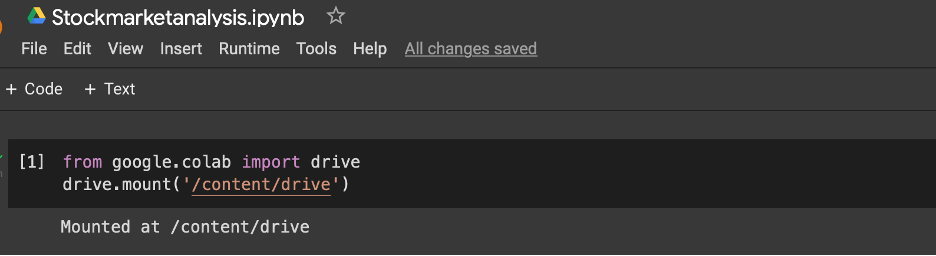
To summarize what was discussed before a flow chart is presented, which talks about the process of ARIMA Model and it’s requirements.



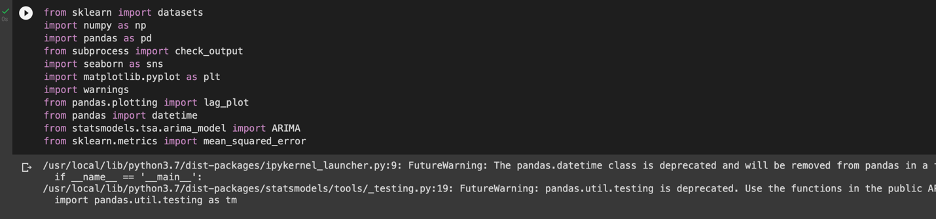
*Tesla Stock Price Prediction using ARIMA model*

We are using google collabo for this prediction. To run the program first we must mount the google drive which is shown below.

**Mounting the drive:**



**Importing the libraries:**

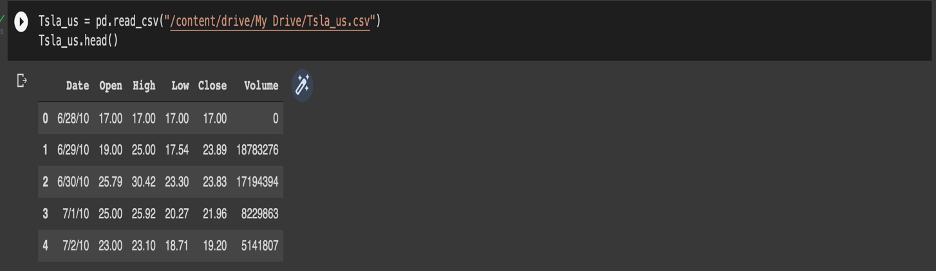


**Load the Training Dataset:**

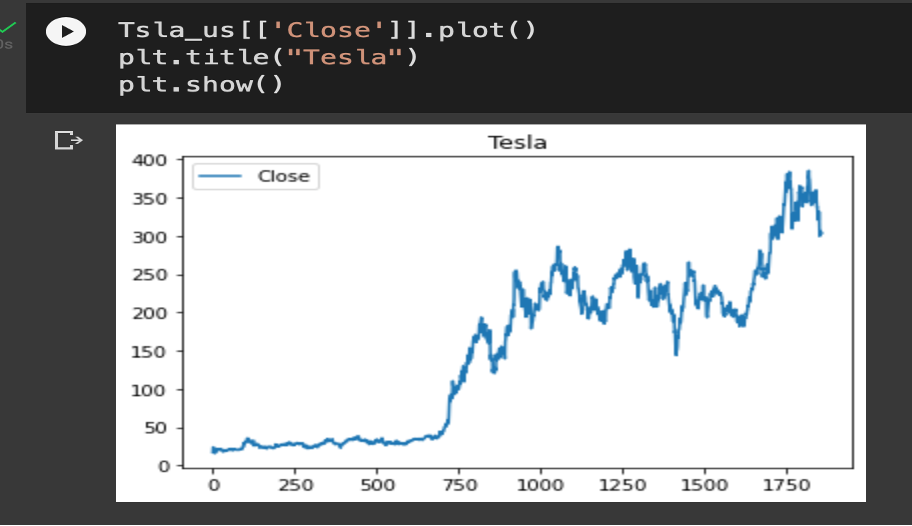
The tesla training data has been collected from below link:

● <https://finance.yahoo.com/quote/TSLA/>

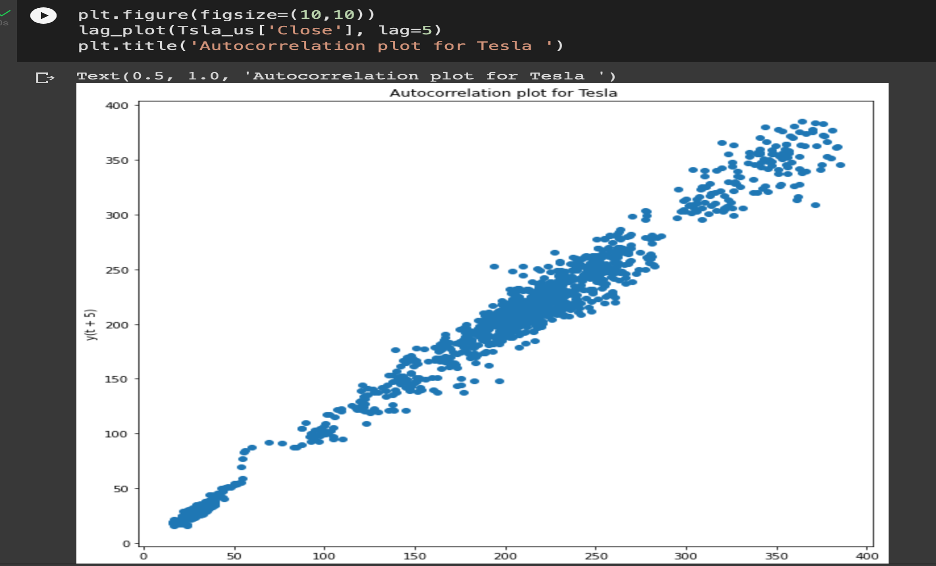
There are 5 columns. The Open column shows the price at which a stock opened to trade when the market opened on a specific date. The "Closing Price" column shows the price of an individual stock when the exchange closed the market for the day. The High column shows the highest price the stock traded during the period. The Low Column shows the lowest price for the period. Trading volume is the total volume of trading over a period.



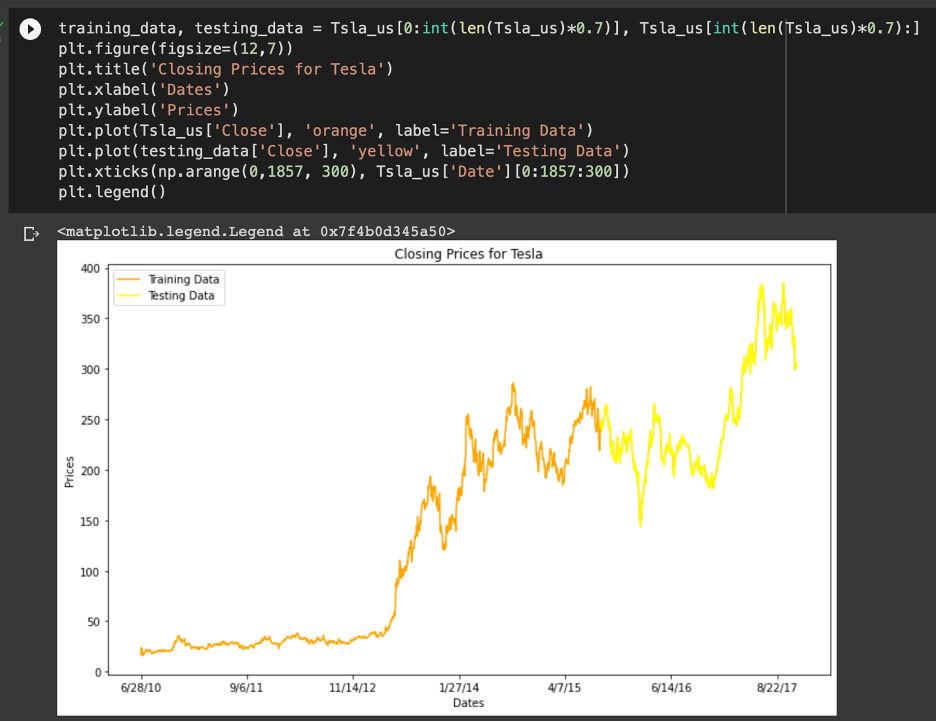
### **Use the Close Stock Price Column to Train Your Model:**



**Autocorrelation for close stock price column:**

****

### **Dividing the dataset in train and test:**

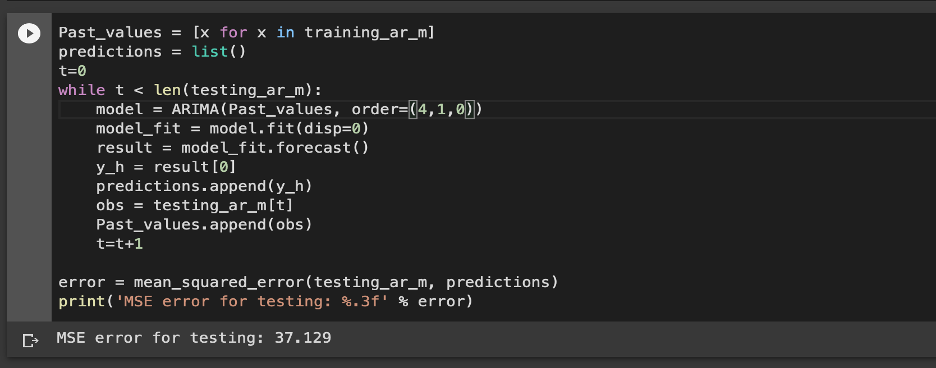
****

### **Building the Model by Importing the Crucial Libraries and Adding Different Layers to ARIMA:**

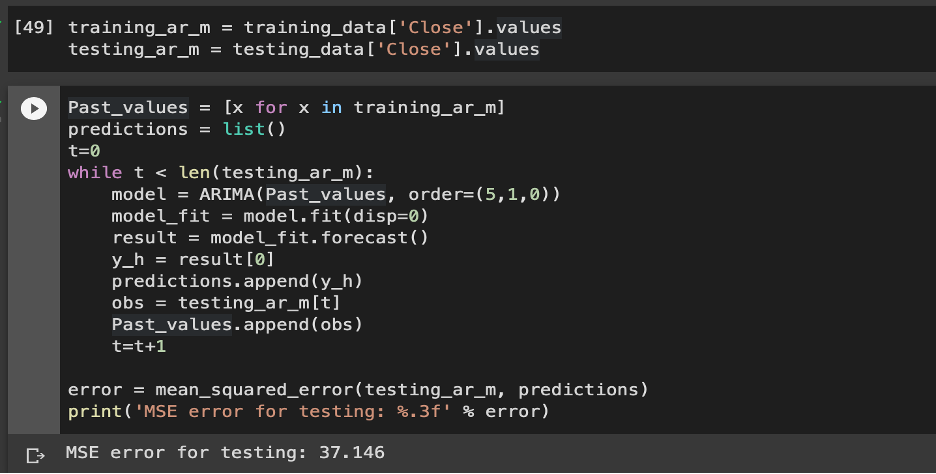
### In ARIMA model we have 3 parameters which are explain below:

* p: The number of lag observations included in the model, also called the lag order.
* d: The number of times that the raw observations are differenced, also called the degree of differencing.
* q: The size of the moving average window, also called the order of moving average.

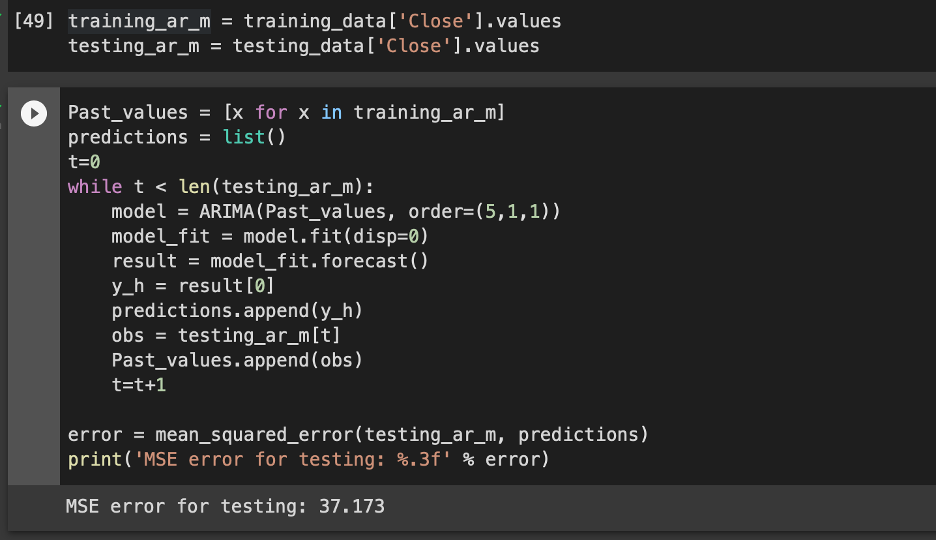
**Let’s take p=4, d=1, q=0:**

****

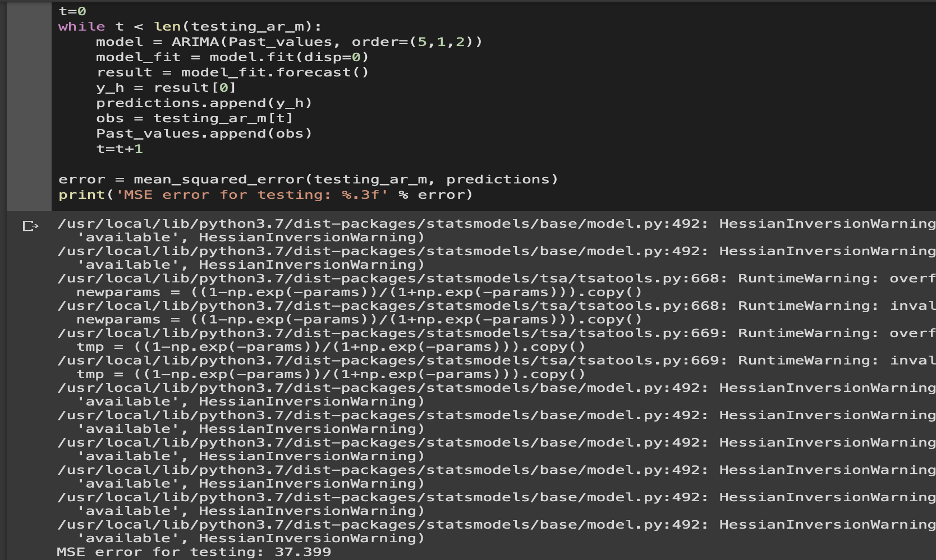
**p=5, d=1,q=0**

****

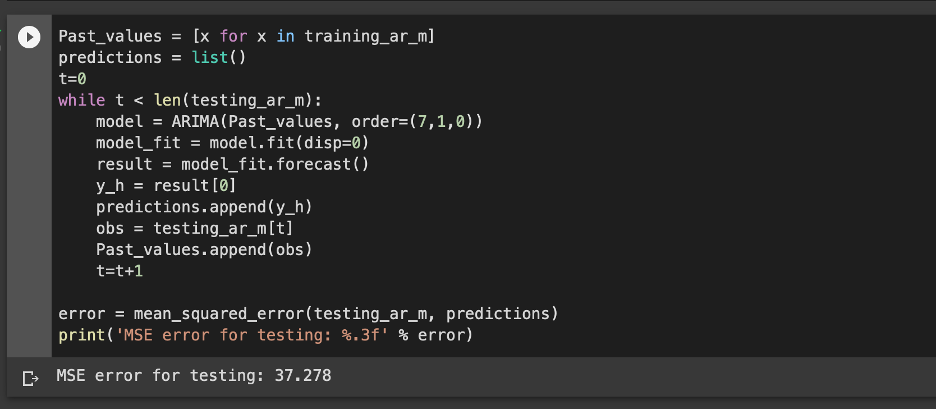
**p=5, d=1, q=1:**

****

**p=5, d=1, q=2:**

****

**p=7, d=1,q=0:**

****

From the above results we concluded, we have the best result with parameters **p=5, d=1, q=0** because we don't want our model to overfit as well as underfit.

### **Plotting the Actual and Predicted Prices for Tesla Stocks:**

****

### **Closing price prediction:**

****

*Conclusion*

The stock showcase plays a momentous part in our everyday lives. It may be a critical figure in a country's GDP development. In this instructional exercise, we learned the nuts and bolts of the stock showcase and how to perform stock price forecasts utilizing machine learning.

References

1. <https://docs.google.com/presentation/d/17rSgssG5q5pOeLtR-wabZ97Yr6pjF9zU/edit#slide=id.p5>
2. <https://www.ft.com/content/7a91e3ea-b9ec-4611-9a03-a8dd3b8bddb5>
3. <https://finance.yahoo.com/quote/TSLA/>
4. <https://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average>
5. <https://en.wikipedia.org/wiki/Moving-average_model>
6. <https://en.wikipedia.org/wiki/Autoregressive_model>