DMPA LAB PROJECT REPORT

Portfolio Optimization with Validation Using Time Series Forecasting and Deep Learning

TEAM MEMBERS:

60) AKSHAT TULSANI   - 170953250
69) AASHAL KAMDAR   - 170953290
71) AKASH HEGDE   - 170953300

Overview -

Portfolio Optimization is the process of selecting the best asset distribution out of all the portfolios being considered. The objective is to achieve a trade-off between maximizing expected returns and minimizing the risk and volatility. Stock and financial markets tend to be unpredictable and due to these characteristics, the data is turbulent in nature which makes it hard to model. Machine Learning, Deep Learning and Time Series Forecasting methods are required to model such data. These methods have the potential to ease the whole process of analyzing data, spotting patterns and generating outputs which can help a trader or broker to make a particular decision or validate it.

The model which has the least Root Mean Square Error for all the stocks, is used to predict on future prices so that we can validate that the prediction made by the Markowitz Portfolio Selection Model is appropriate.

Collection of data -

Dataset was taken from Yahoo Finance. The dataset contains the following attributes:-
A] Date
B] Open (the price at which the stock opened that day)
C] High (the highest price which the stock reached that day)
D] Low (the lowest price which the stock reached that day)
E] Close (the price at which the stock closed that day)
F] Adjusted Close (amends the Close price to reflect any corporate actions that might have been taken after the stock market closes for the day)
G] Volume (the amount of stocks traded that day)

We choose Adjusted Close for all the analysis done because it is the best reflection of all the actions taken after the market closes.

Here we have considered 3 stocks for analysis -

1] ICICI Bank (ICICIBANK.BO)

2] Hindustan Unilever Ltd (HINDUNILVR.BO)

3] Larsen & Toubro (LT.BO)

The bracket contains the ticker symbol (how the stock is represented on the stock exchange, in this case, the Bombay Stock Exchange as indicated by BO).
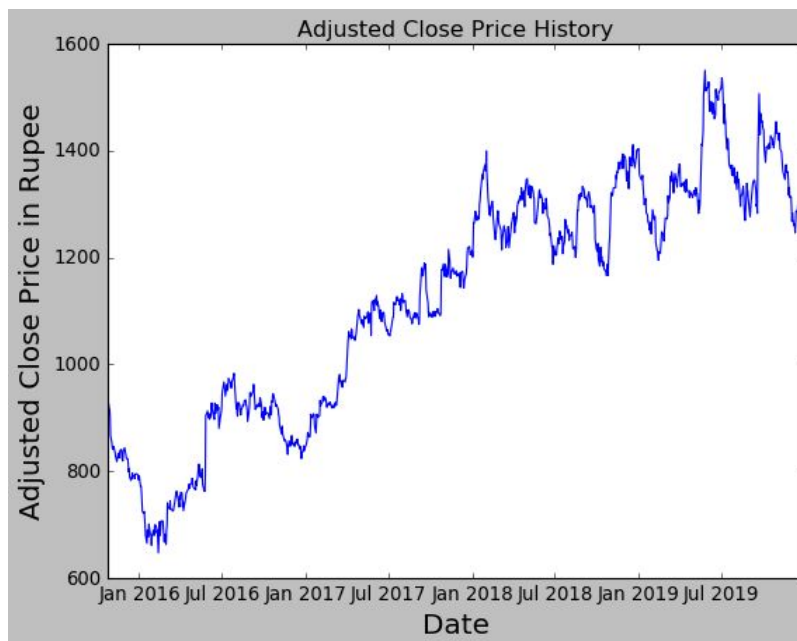
First, we do the Portfolio Optimization using the Markowitz Portfolio Selection Model, then we validate the allocation using Time Series and Deep Learning Models.

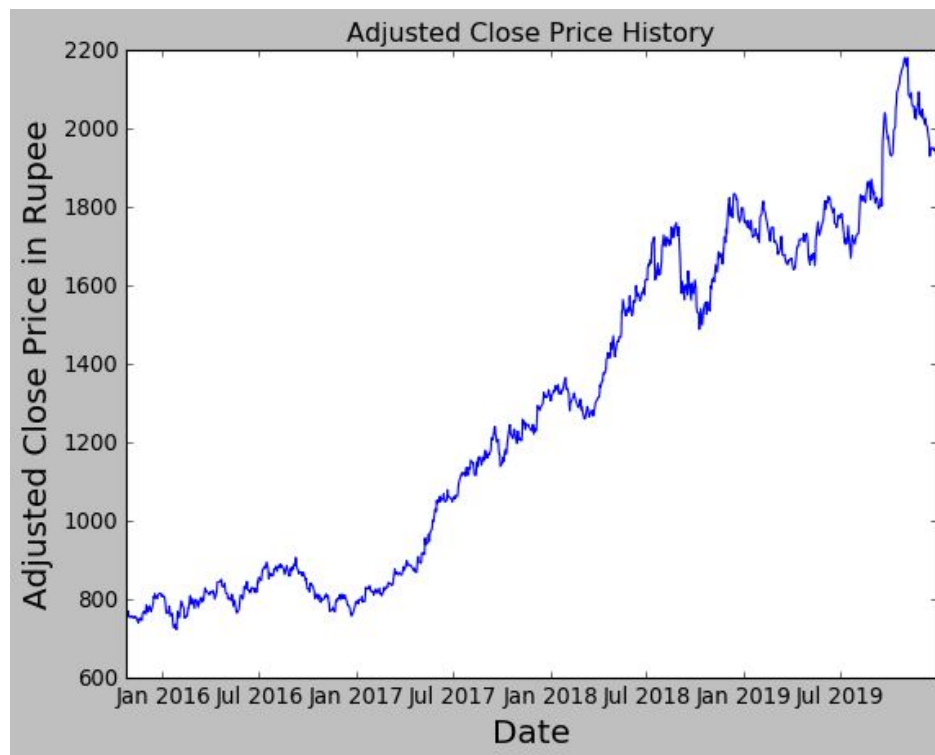The graph for the Adjusted close of each stock is as follows -

1] ICICI Bank


2] L & T

3] Hindustan Unilever Ltd

Portfolio Optimization -

We find the logarithmic return of each stock (for the adjusted close price), and then concat it all into one dataframe.

Date	LT	ICICI	HU
2015-10-23	NaN	NaN	NaN
2015-10-26	-0.001431	-0.003323	0.008711
2015-10-27	-0.009222	-0.005275	0.005376
2015-10-28	-0.001942	-0.043956	0.007367
2015-10-29	-0.013365	-0.000182	-0.019581

The below picture depicts the histogram for the logarithmic return for each stock. We can see that each stock almost follows a normal distribution hence it is preferable to use logarithmic returns as compared to arithmetic returns.
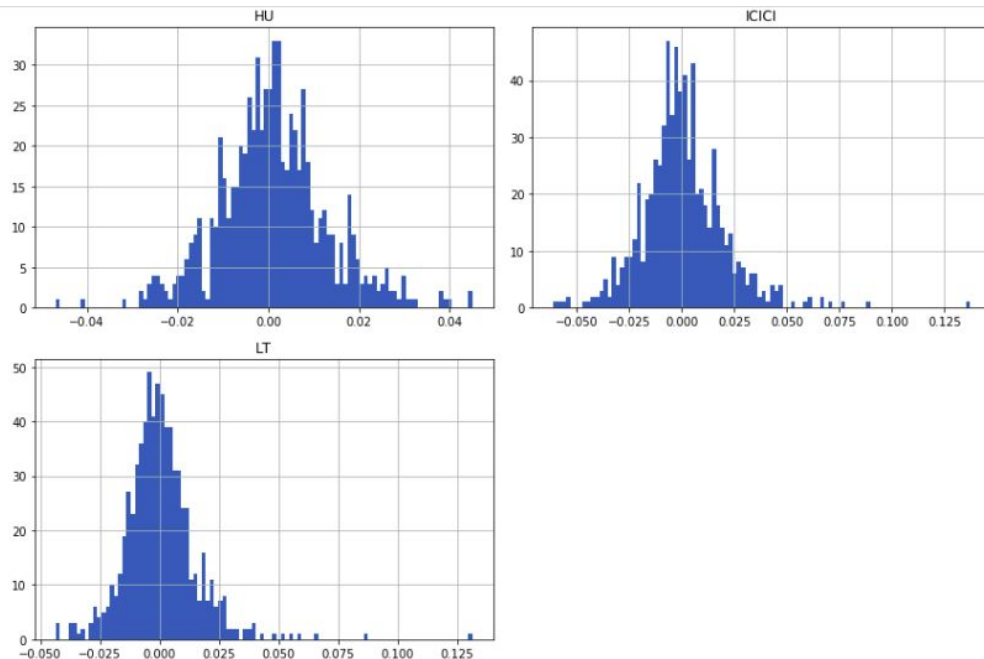
The results are as follows -

```
In [27]: sharpe_arr.max()
Out[27]: 1.384789610792806

In [28]: sharpe_arr.argmax()
Out[28]: 15956

In [29]: all_weights[15956,:]
Out[29]: array([0.0881561, 0.0010194, 0.9108245])

In [31]: max_sr_ret = ret_arr[15956]
         max_sr_vol = vol_arr[15956]

         print(max_sr_ret)
         print(max_sr_vol)

         0.24948628894543434
         0.18016187224469496
```

In cell [27] we find the maximum Sharpe ratio. Sharpe ratio is used to help investors understand the return of an investment compared to its risk. The more the Sharpe ratio, the more profitable an investment is. Here the Sharpe ratio we get is 1.38.

Cell [29] tells us the allocation of the stocks in our portfolio-
L & T - 8.8%
ICICI Bank - 0.12 %
Hindustan Unilever Ltd - 91.08%

The maximum returns we can get is 24.9% with a risk(volatility) of 18%.
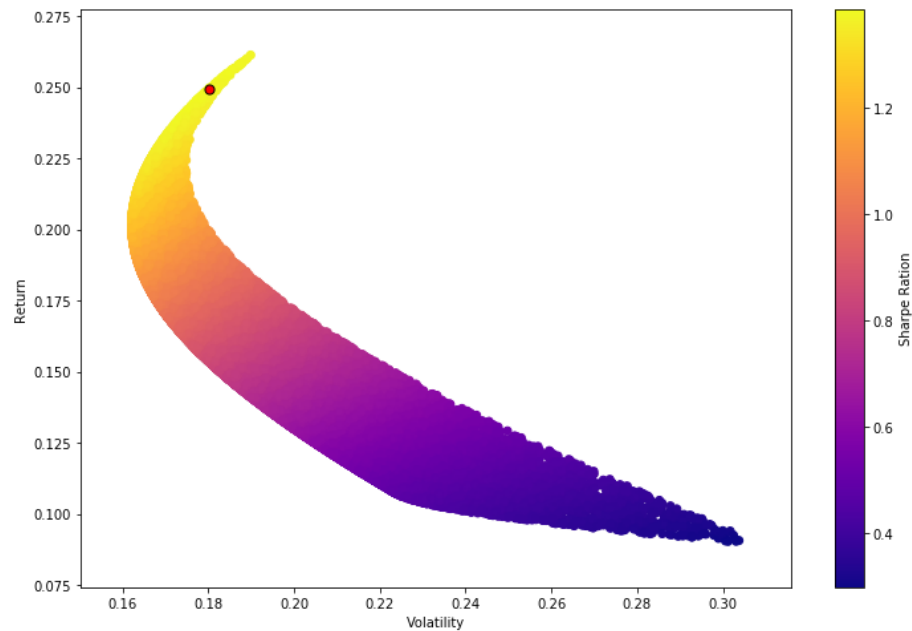
The red dot on the above graph is the depiction of best return vs volatility i.e., Sharpe ratio.

Efficient Frontier Curve - The efficient frontier is the set of optimal portfolios that offer the highest expected return for a defined level of risk or the lowest risk for a given level of expected return. Portfolios that lie below the efficient frontier are sub-optimal because they do not provide enough return for the level of risk
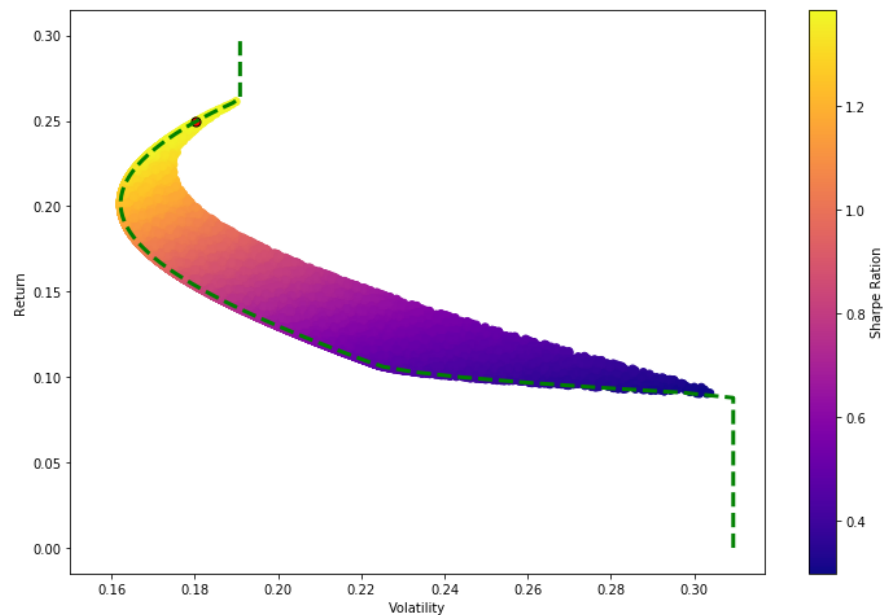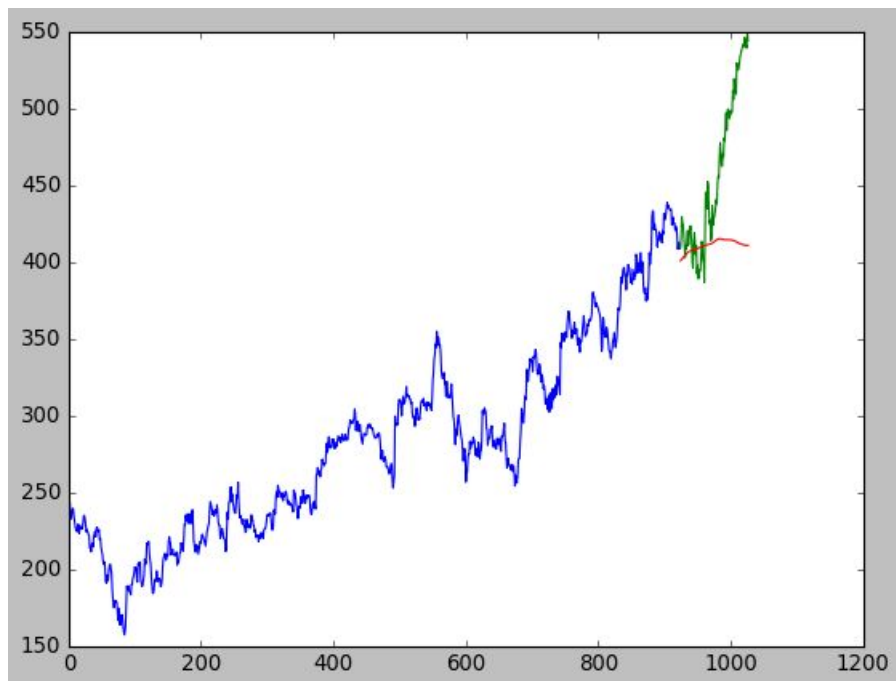
I. Moving Average :

Comparison of actual stock value to that of predicted using moving average for L&T:
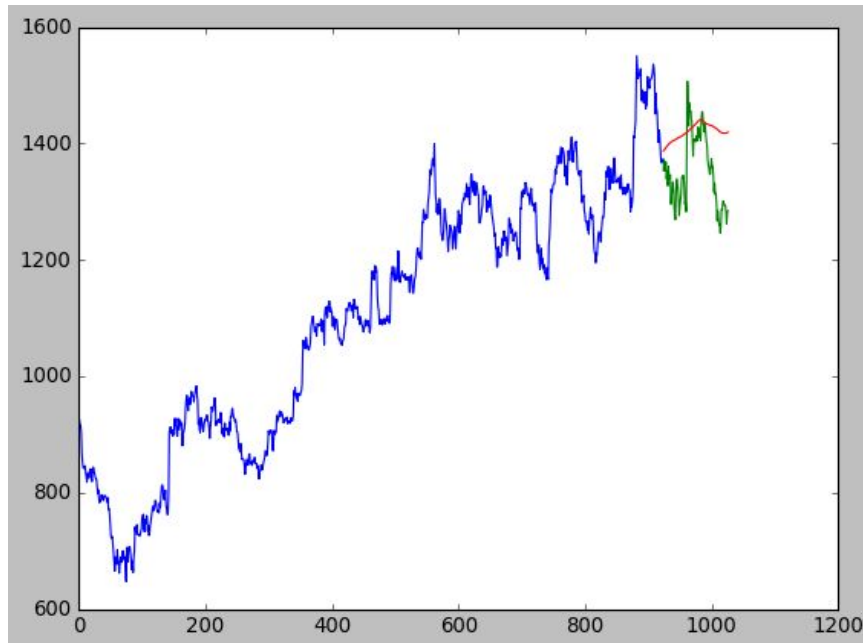
Comparison of actual stock value to that of predicted using moving average for ICICI:

Comparison of actual stock value to that of predicted using moving average for Hindustan Unilever:
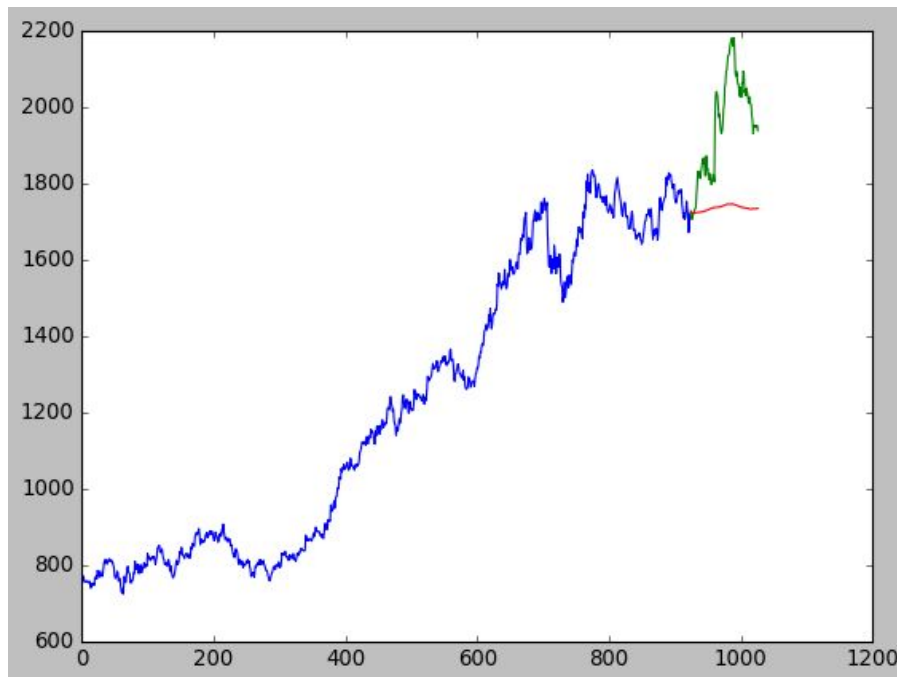
II. LSTM (Long short-term memory):

Model summary-

Model: "sequential_2"

Layer (type)	Output Shape	Param #
lstm_4 (LSTM)	(None, 60, 70)	20160
dropout_3 (Dropout)	(None, 60, 70)	0
lstm_5 (LSTM)	(None, 60, 70)	39480
dropout_4 (Dropout)	(None, 60, 70)	0
lstm_6 (LSTM)	(None, 70)	39480
dense_2 (Dense)	(None, 1)	71

Total params: 99,191
Trainable params: 99,191
Non-trainable params: 0

1. L & T assessment Using LSTM -

Mean Absolute Error Graph based on the model trained using LSTM for L&T stocks
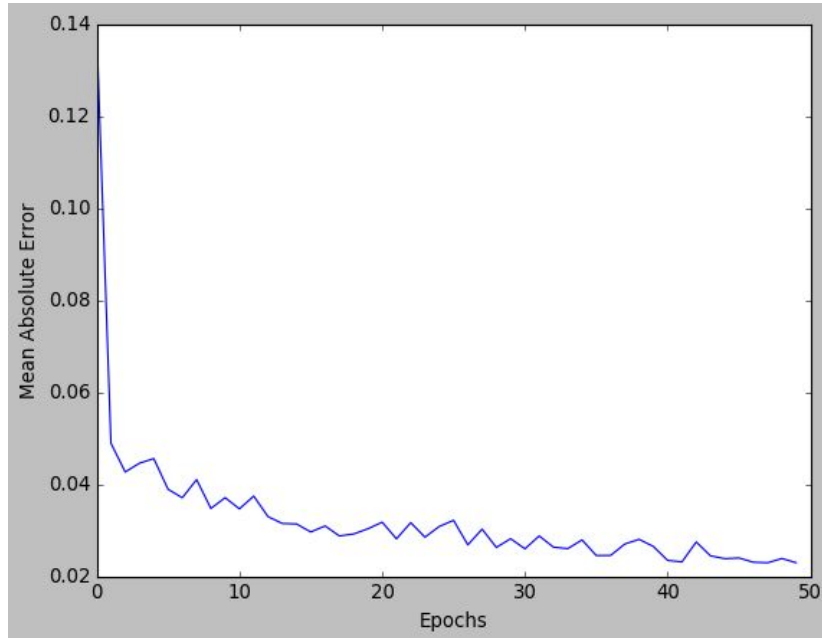
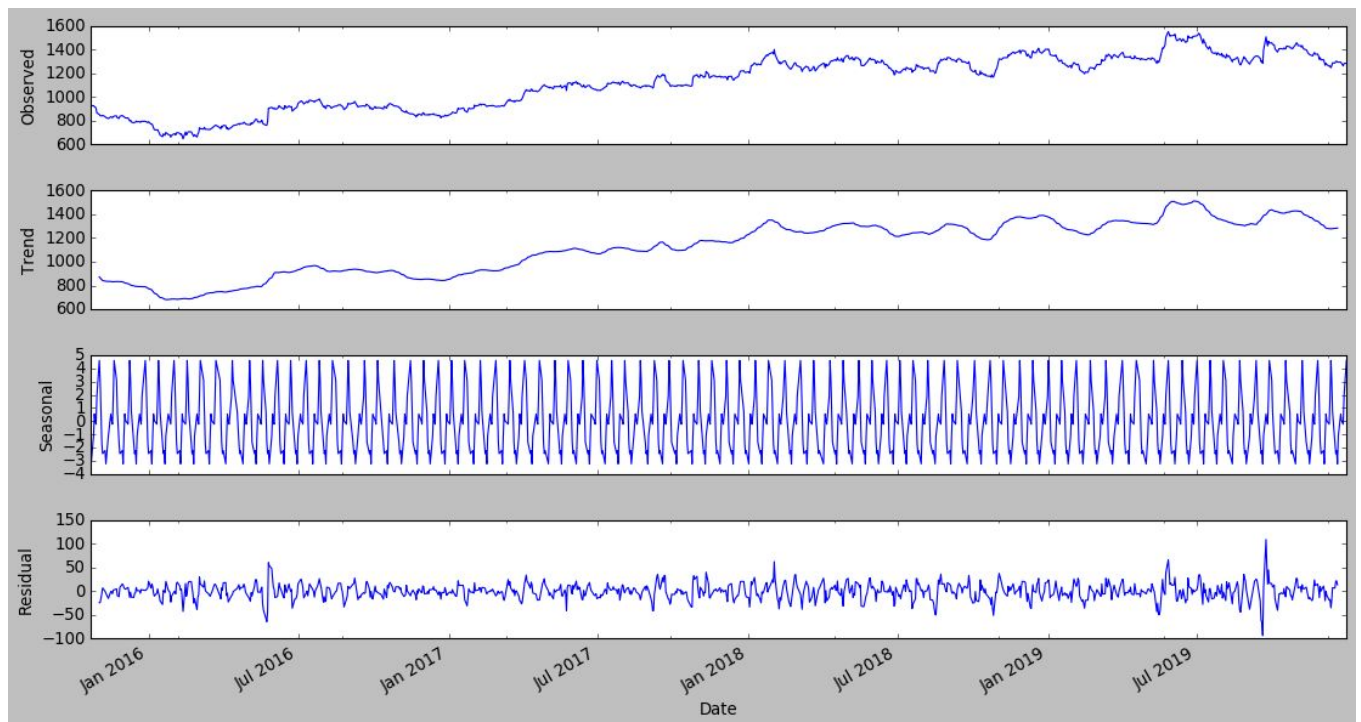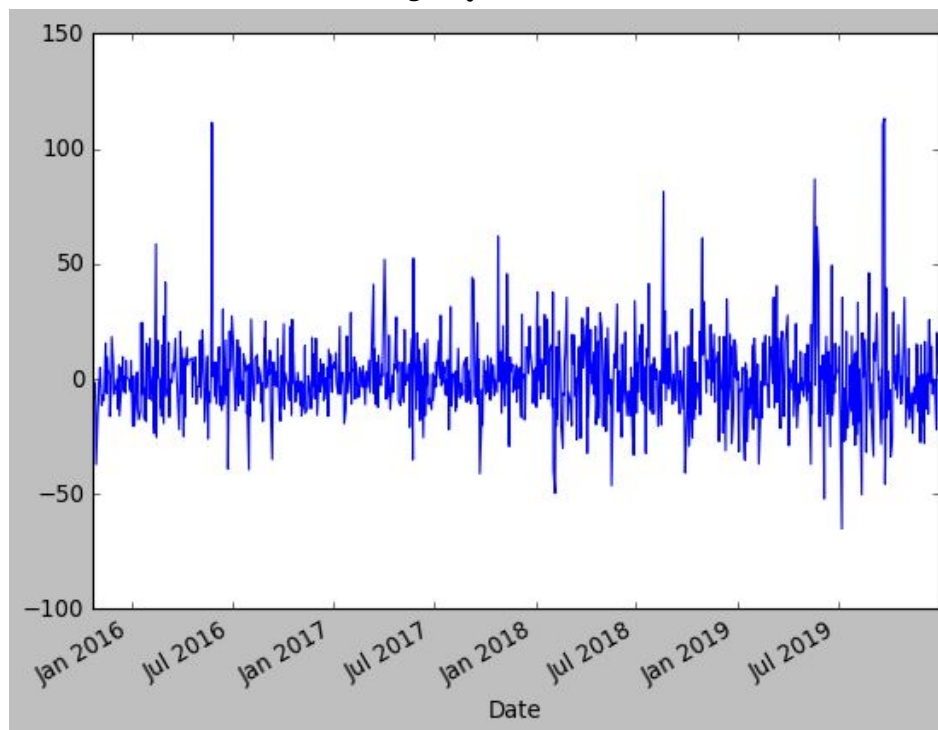Prediction and actual change in stock values for L&T stocks:
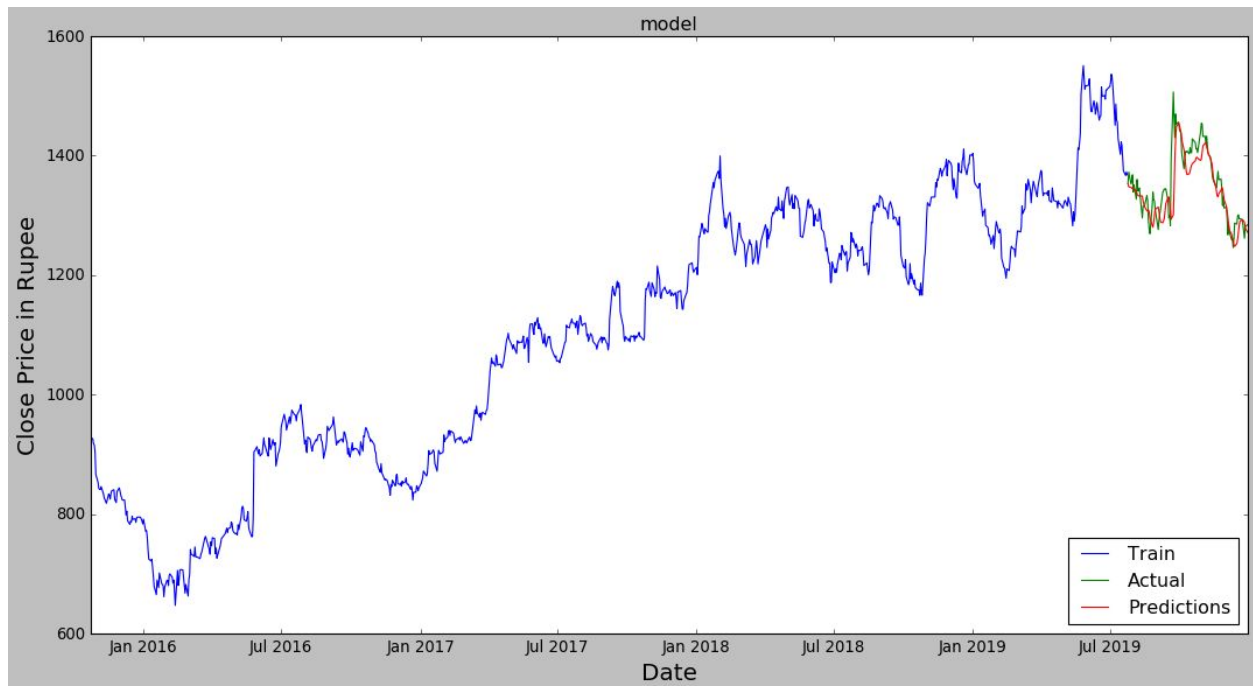
2. ICICI assessment using LSTM -

Mean Absolute Error Graph based on the model trained using LSTM for ICICI stocks:
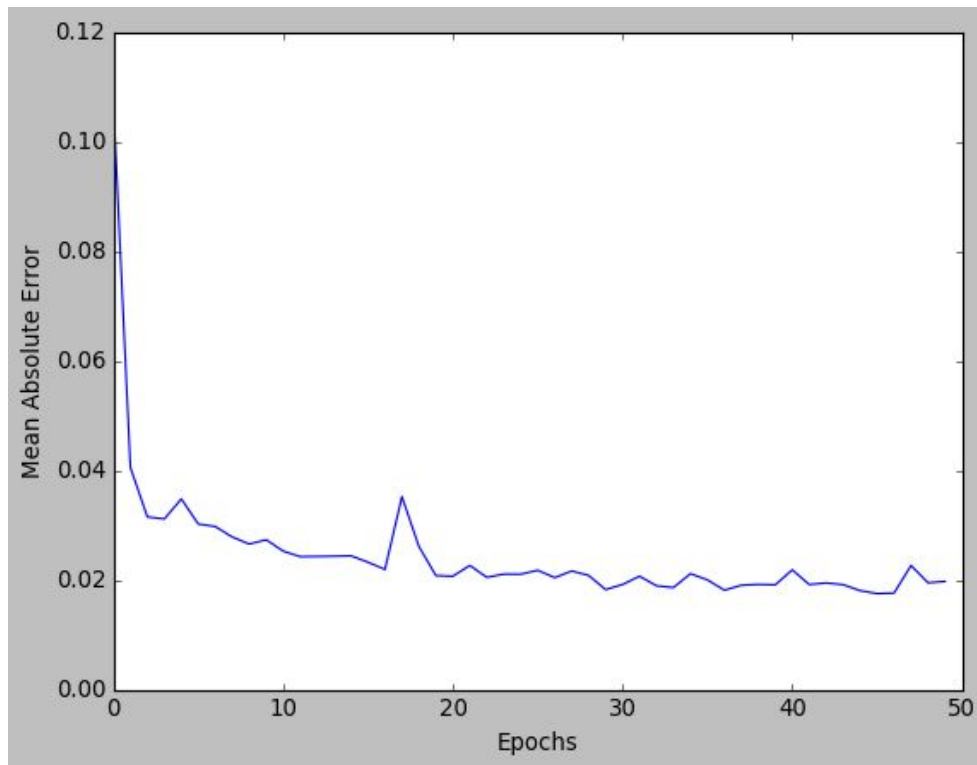
Prediction and actual change in stock values for ICICI stocks:
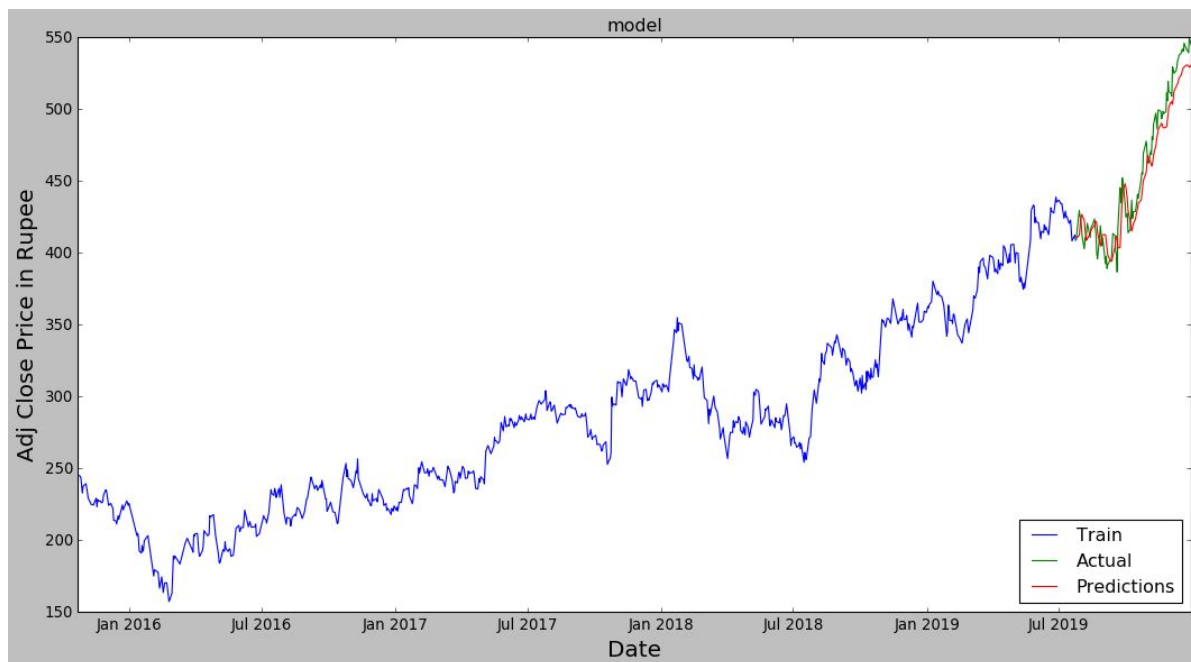
3. Hindustan Unilever Ltd assessment using LSTM -

Mean Absolute Error Graph based on the model trained using LSTM for HUL stocks
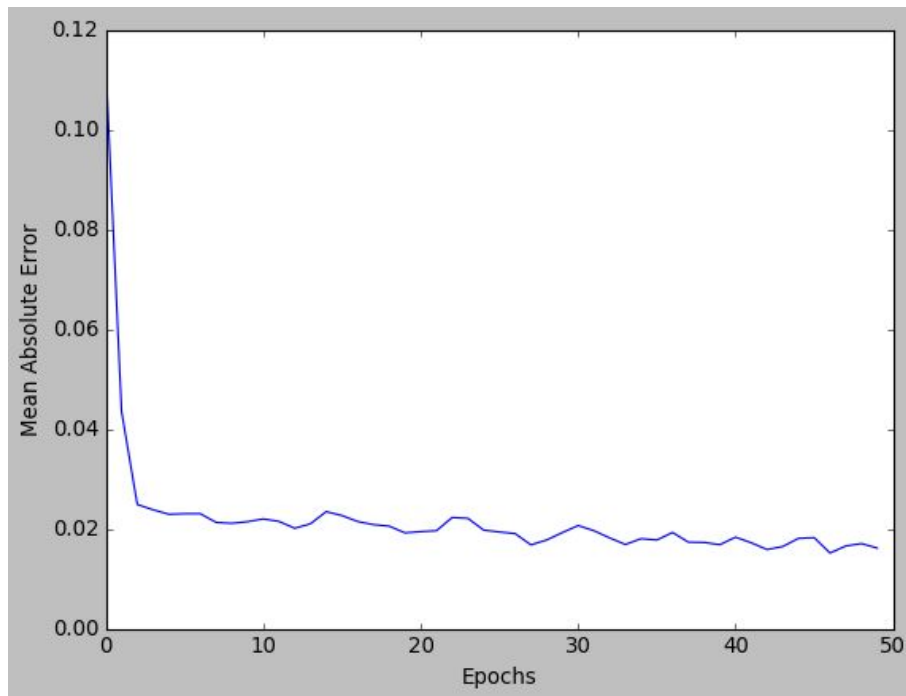
Prediction and actual change in stock values for HUL stocks:
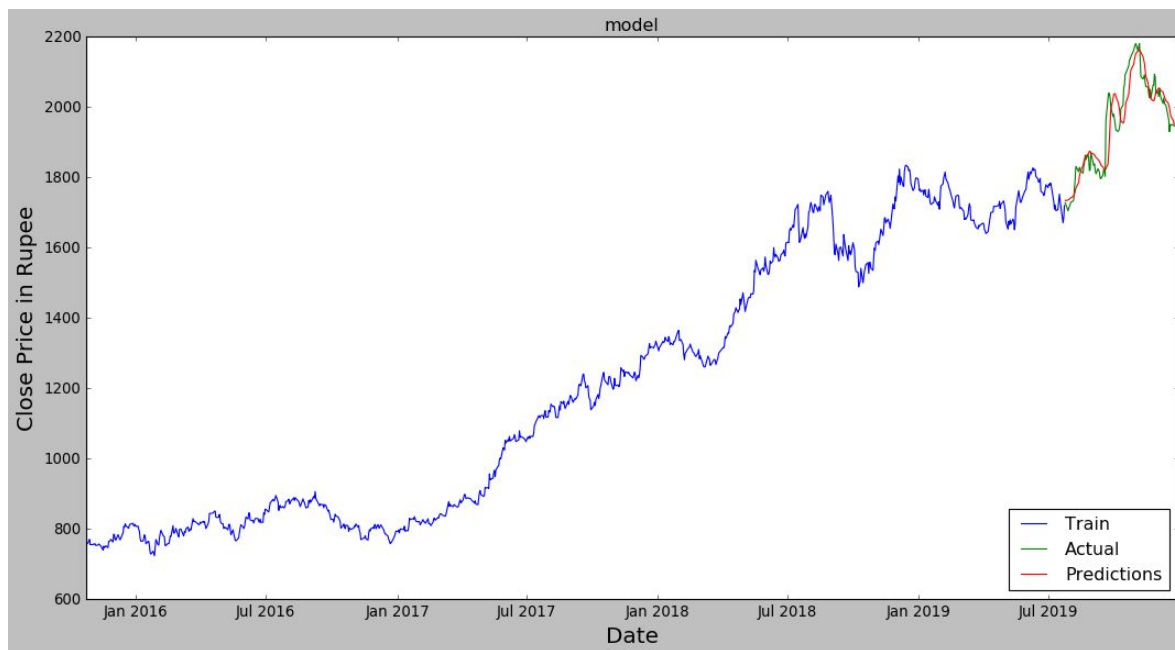
III. ARIMA

(Auto-Regressive Integrated Moving Average)

1. **Larsen & Toubro assessment using ARIMA**

Error, Trend and Seasonality (ETS)-

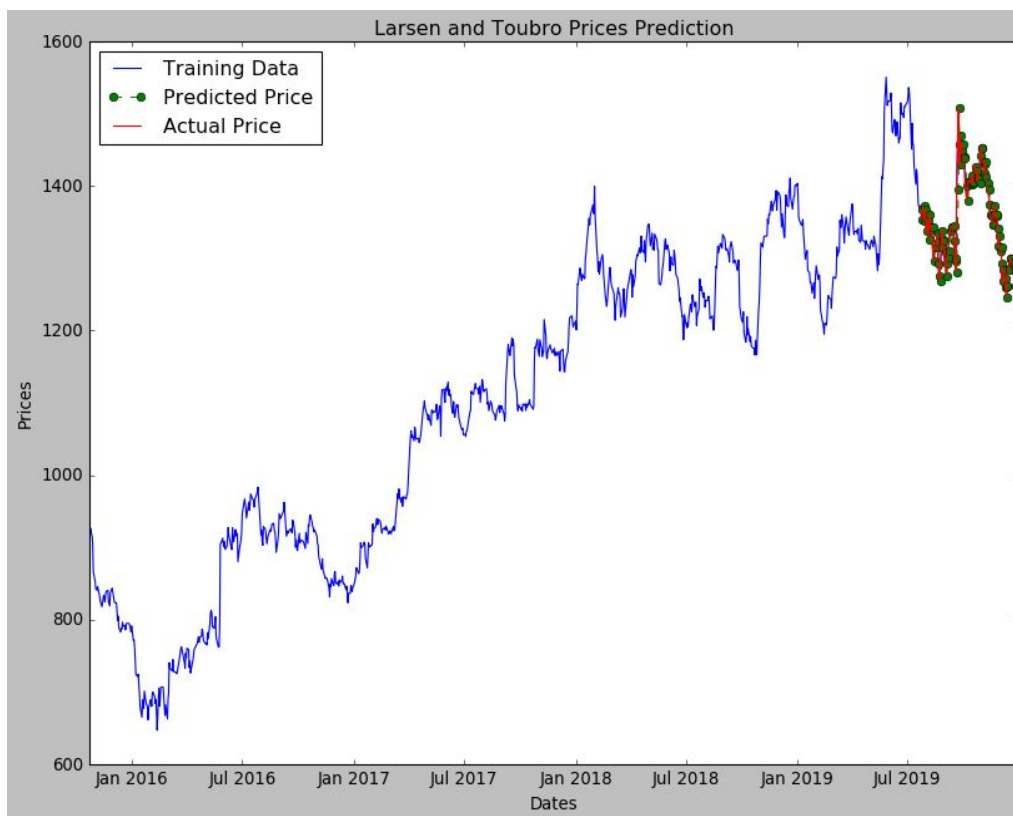Graph of first difference (After shifting adjusted close once) -
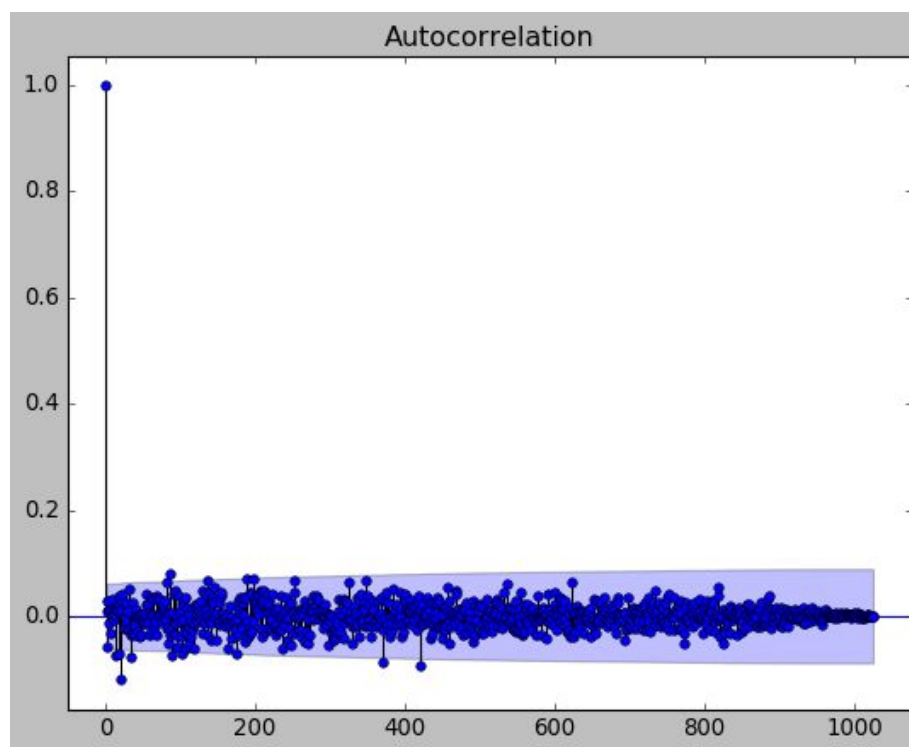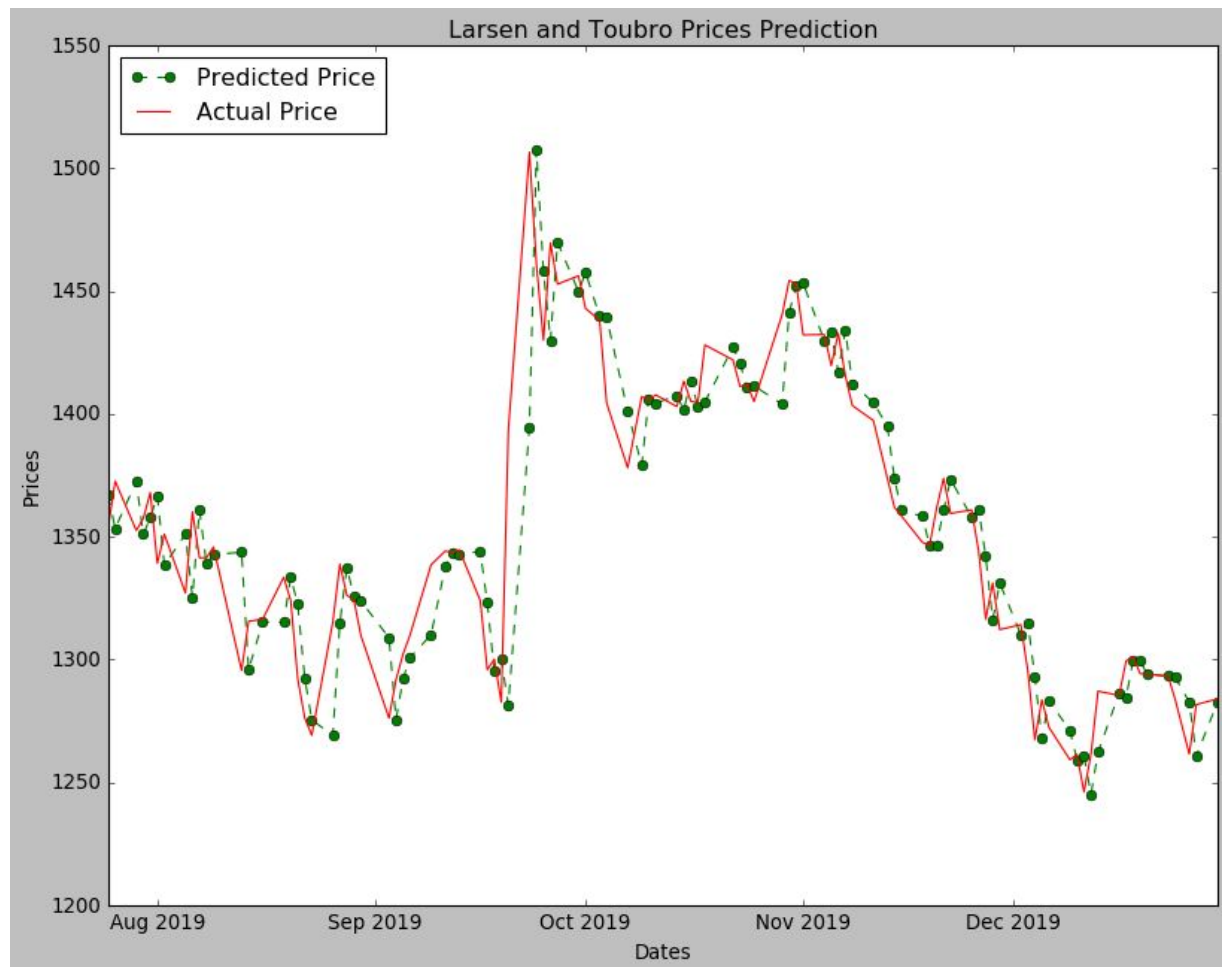

Autocorrelation Plot -

Partial Autocorrelation Plot -

Graph of the training and testing data-

L & T Prices

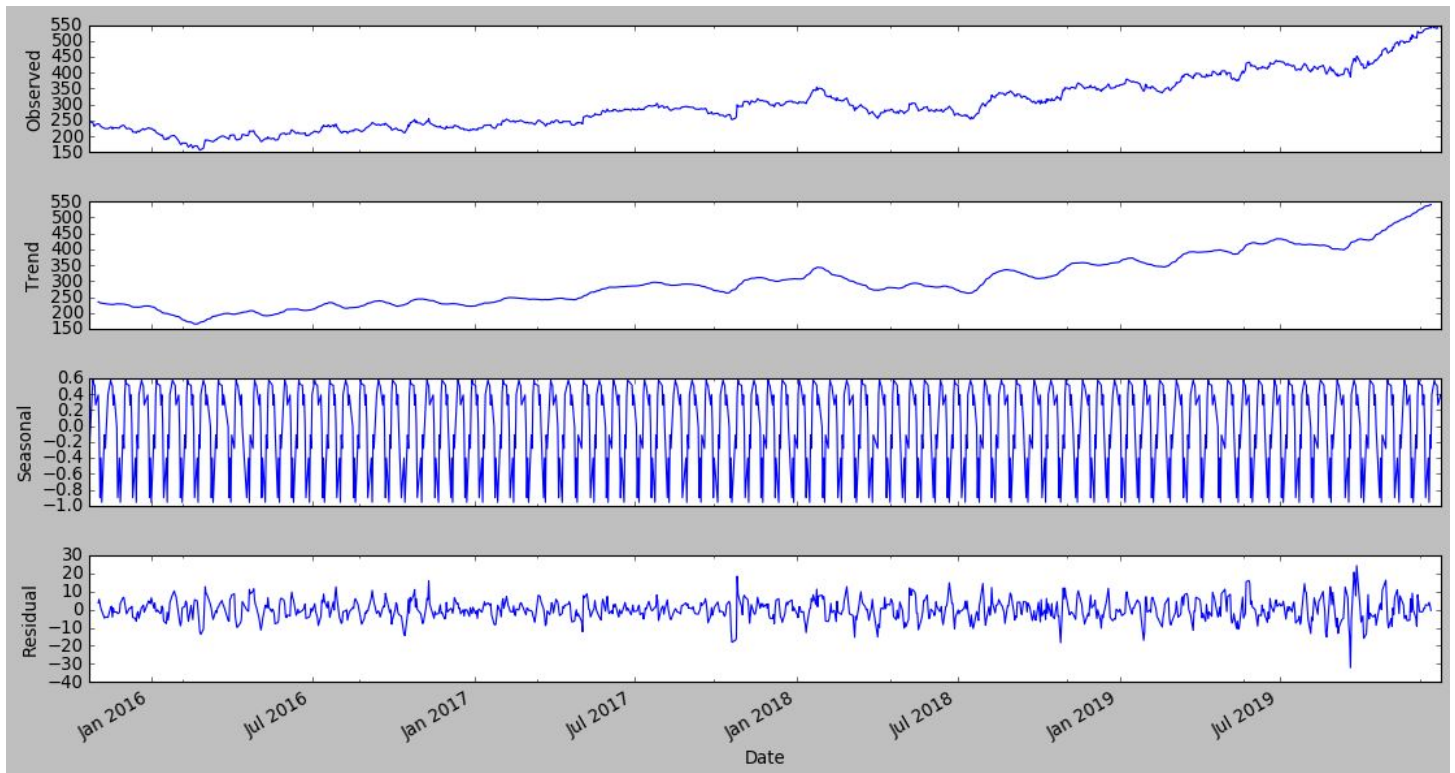Prediction and actual change in stock values for L & T stocks -

Larsen and Toubro Prices Prediction
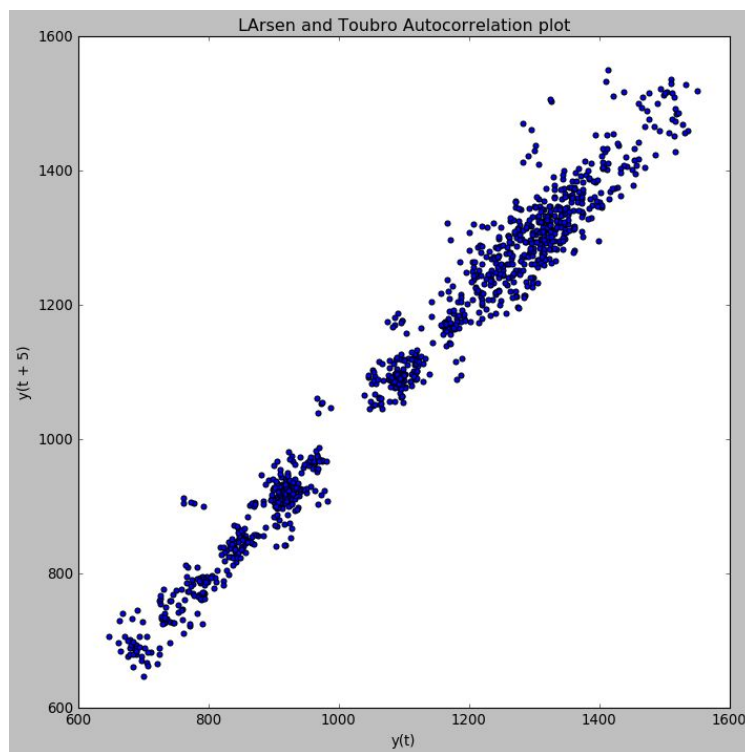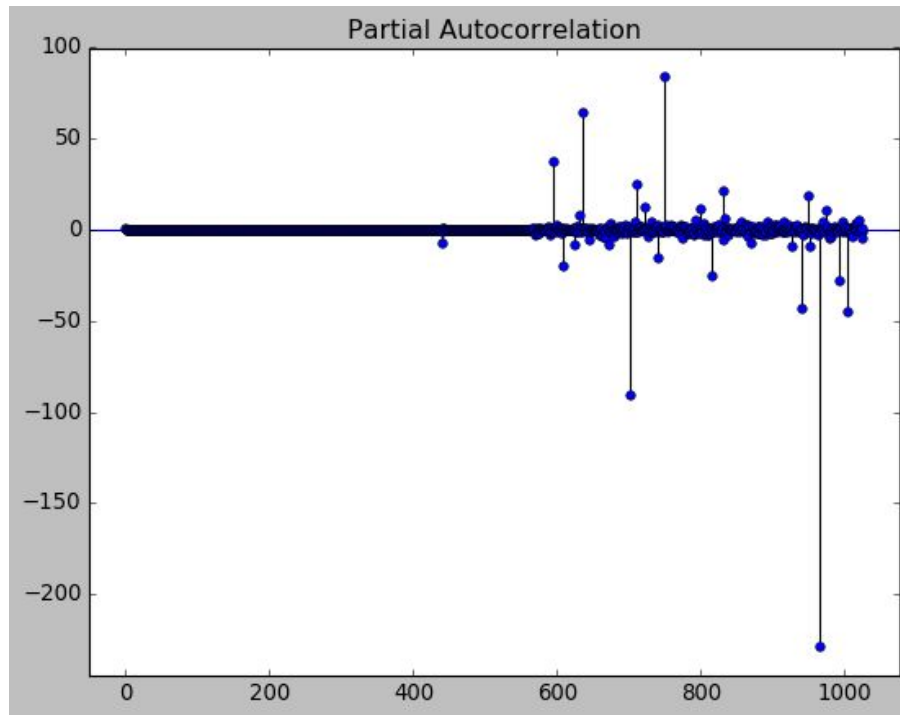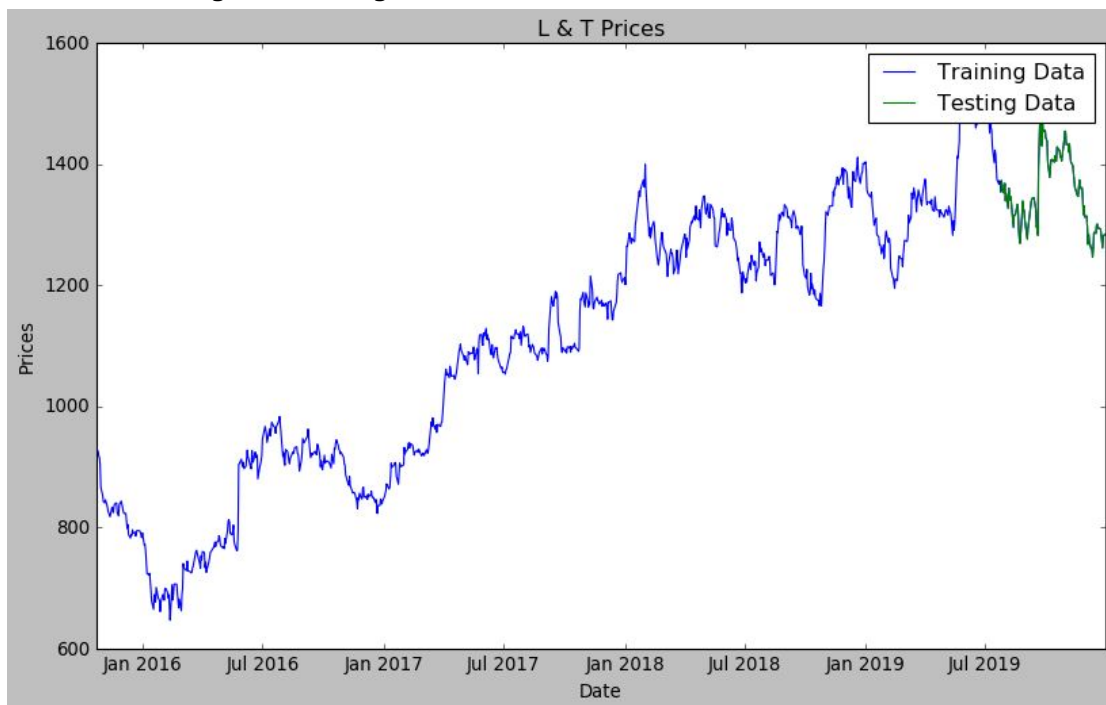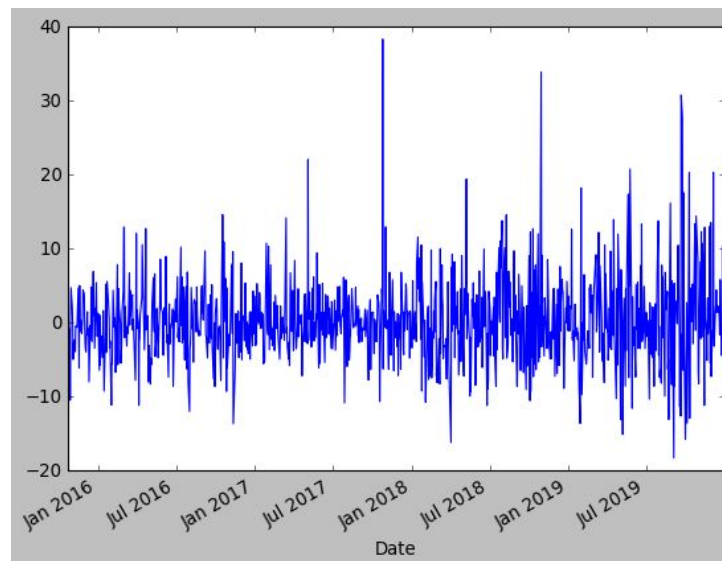
Close up of Predicted and Actual Price -
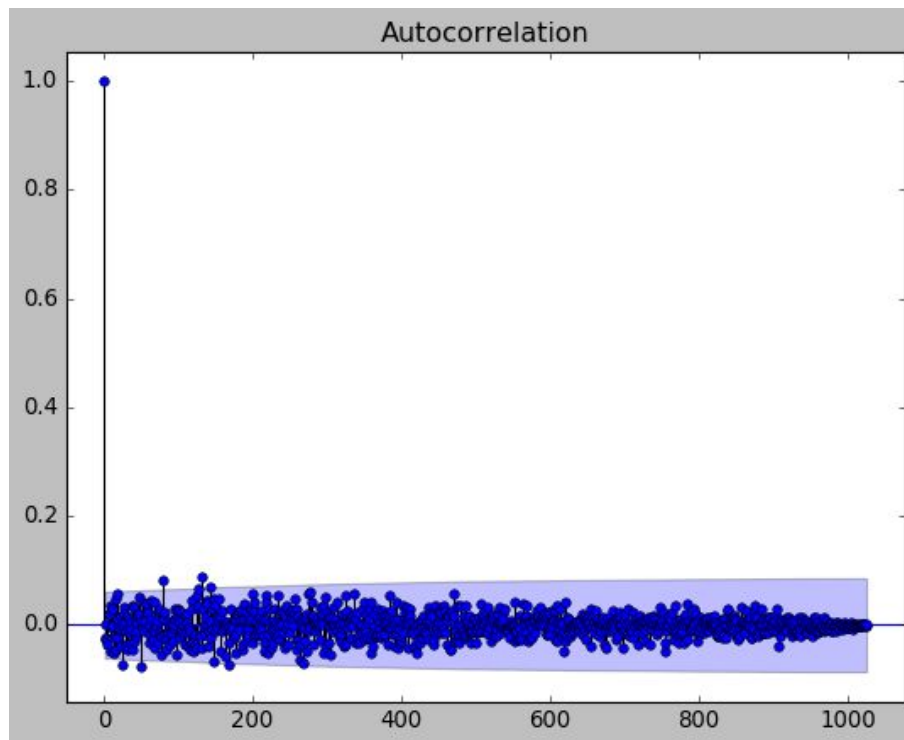
2. ICICI assessment using ARIMA:
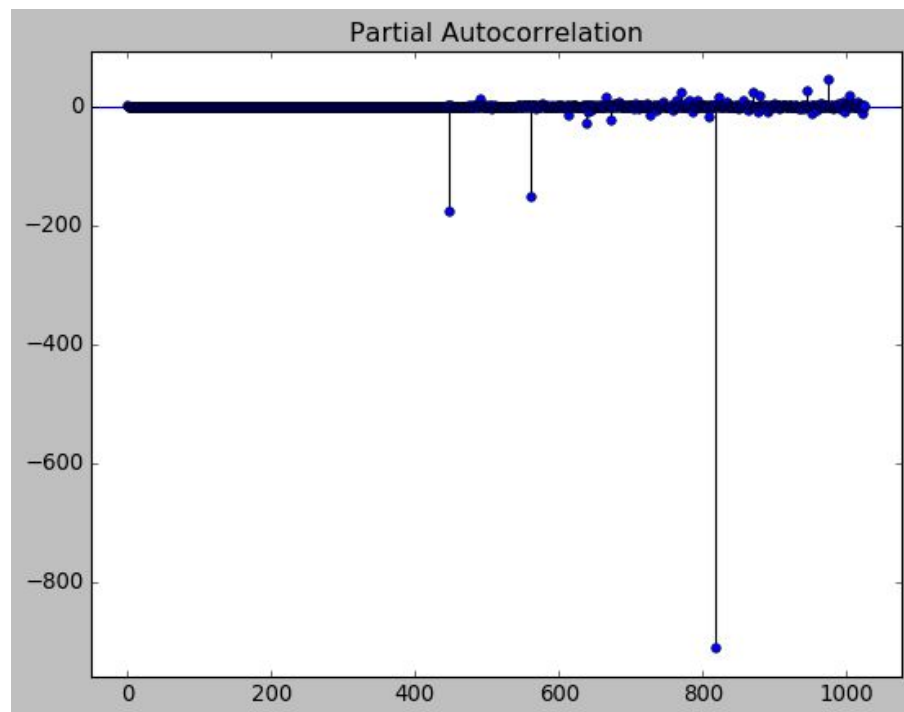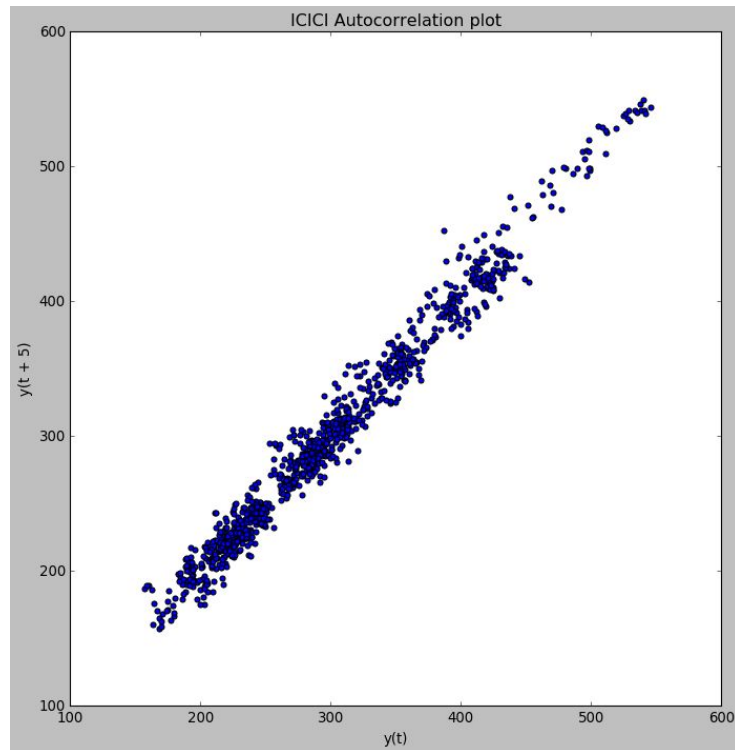
Error, Trend and Seasonality Graph-

Graph of first difference (After shifting adjusted close once) -

Autocorrelation Plot-
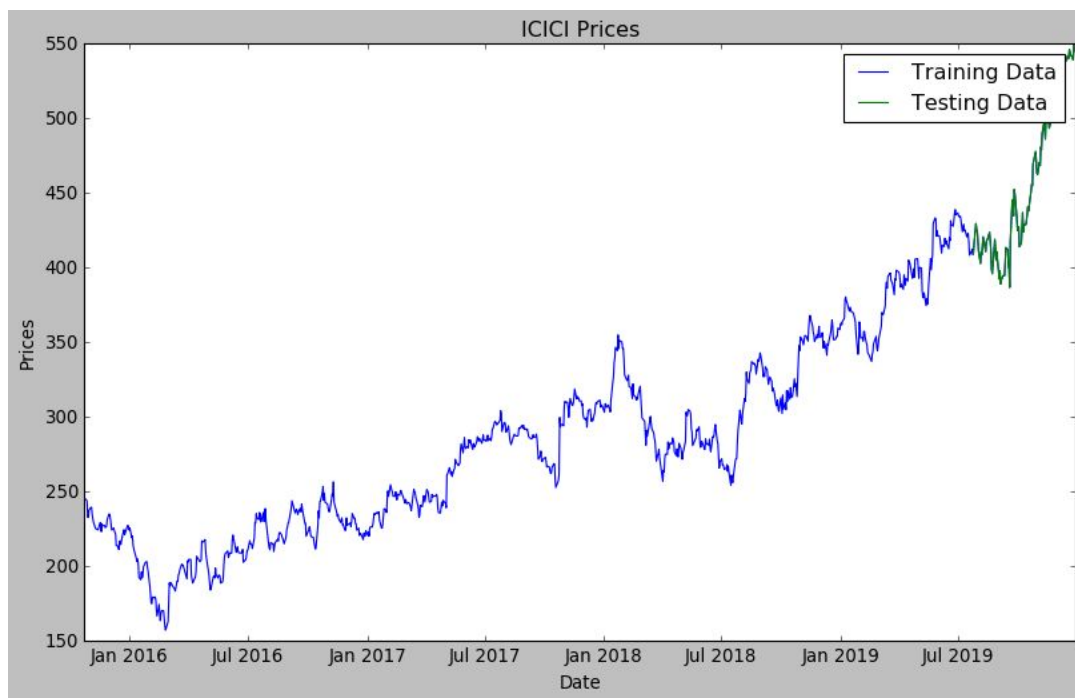

Partial Autocorrelation Plot -
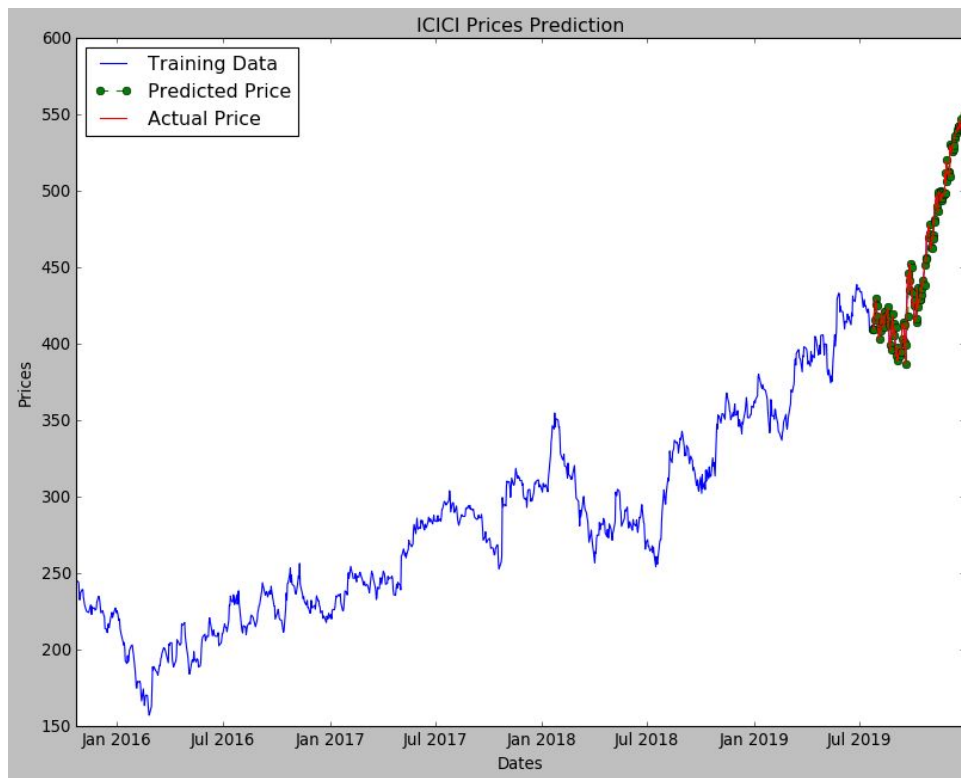
ICICI Autocorrelation plot

Graph of the training and testing data-

ICICI Prices

Prediction and actual change in stock values for ICICI stocks -

Close up of Predicted and Actual Price-

3.Hindustan Lever assessment using ARIMA:

Error, Trend and Seasonality Graphs-

Graph of first difference (After shifting adjusted close once) -

AutoCorrelation Plot -
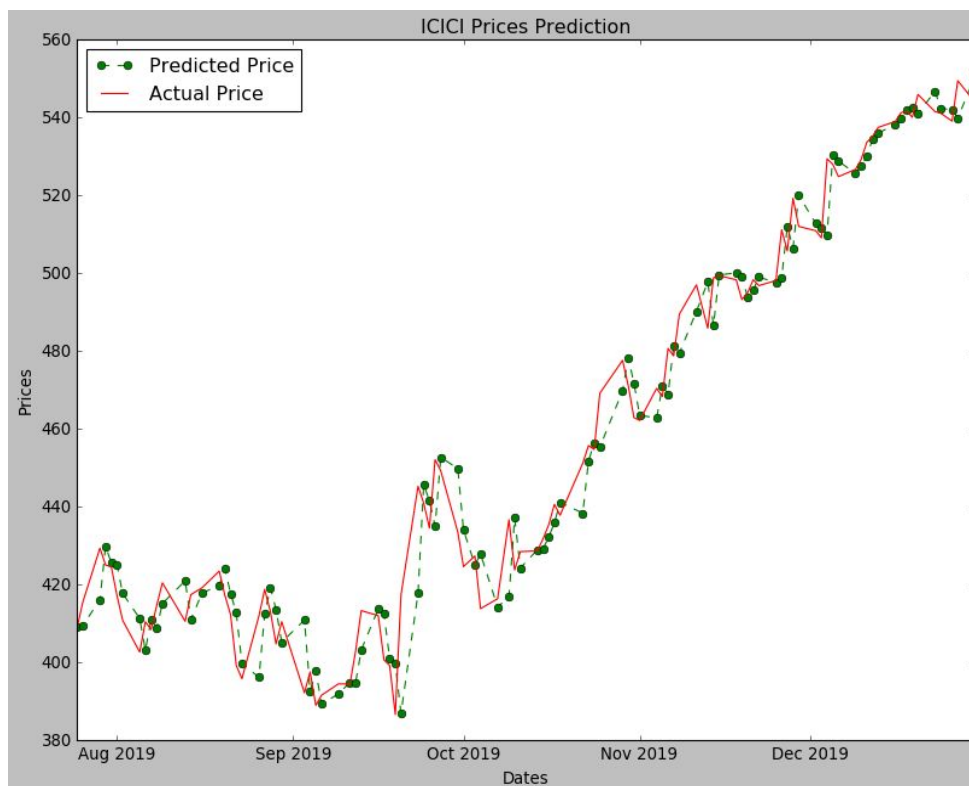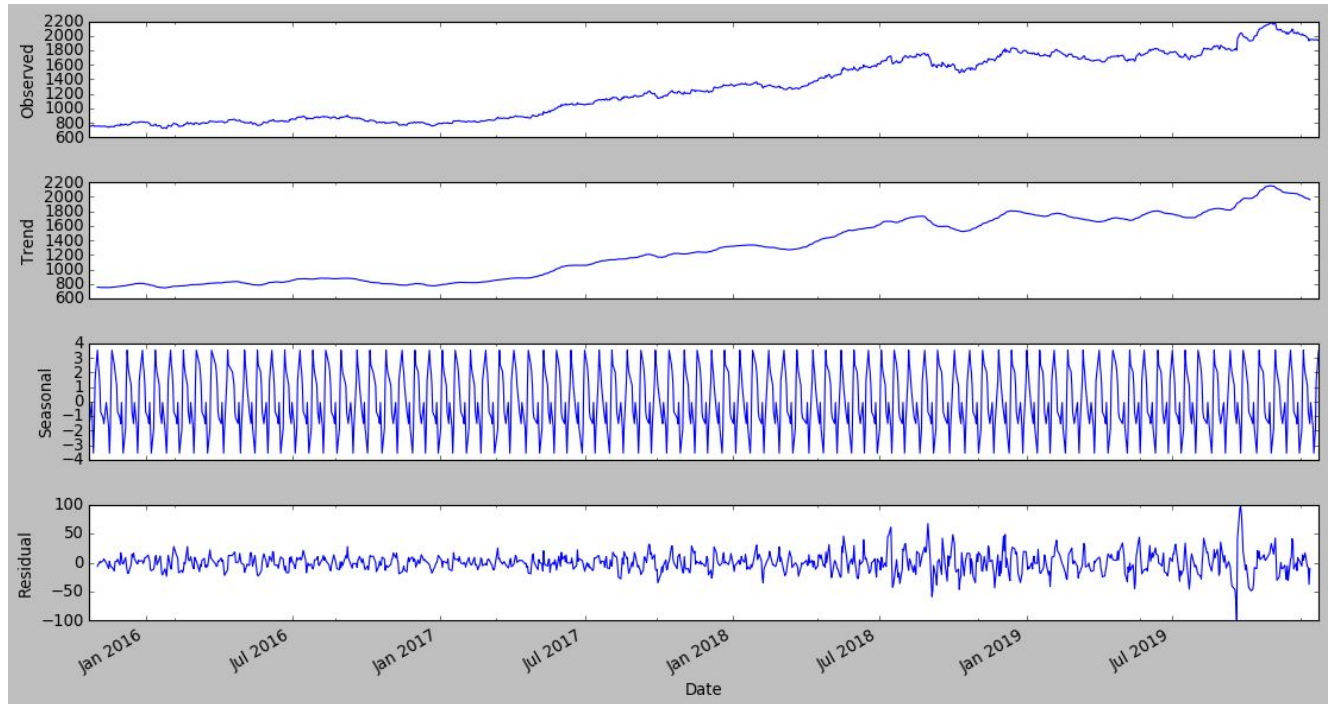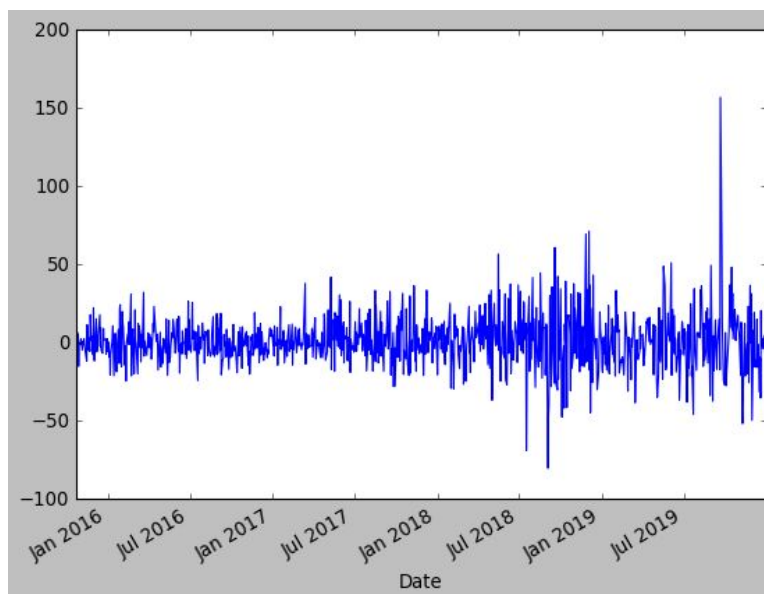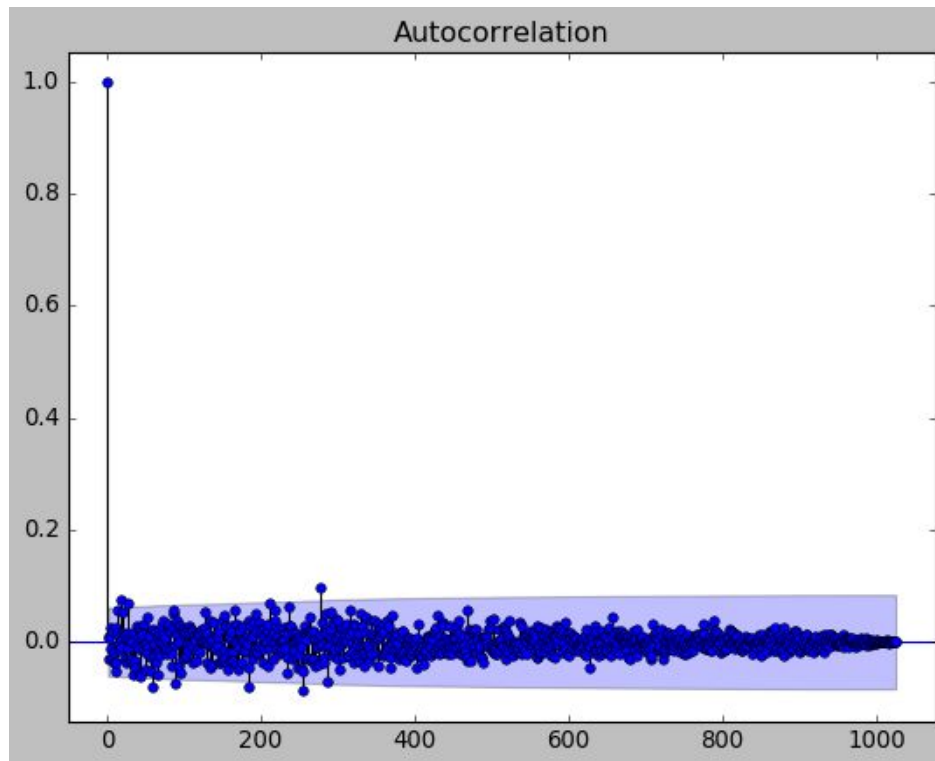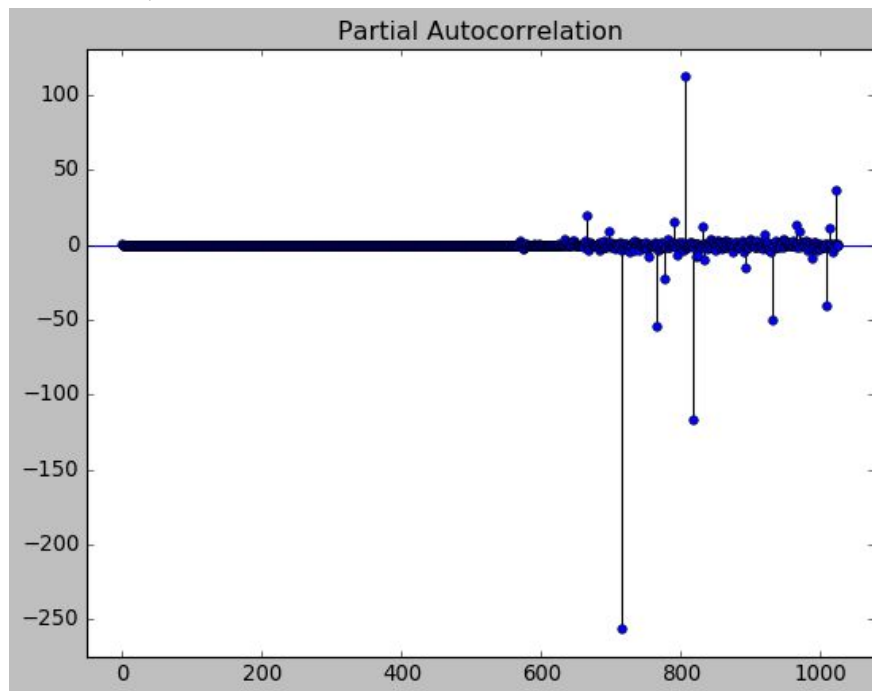

Partial Autocorrelation plot-
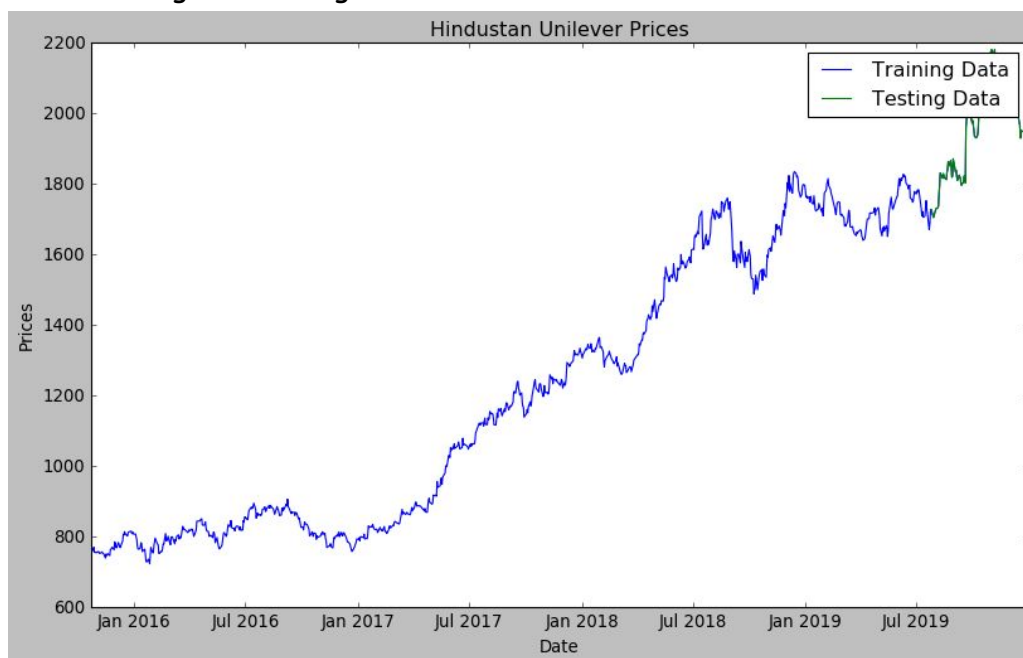
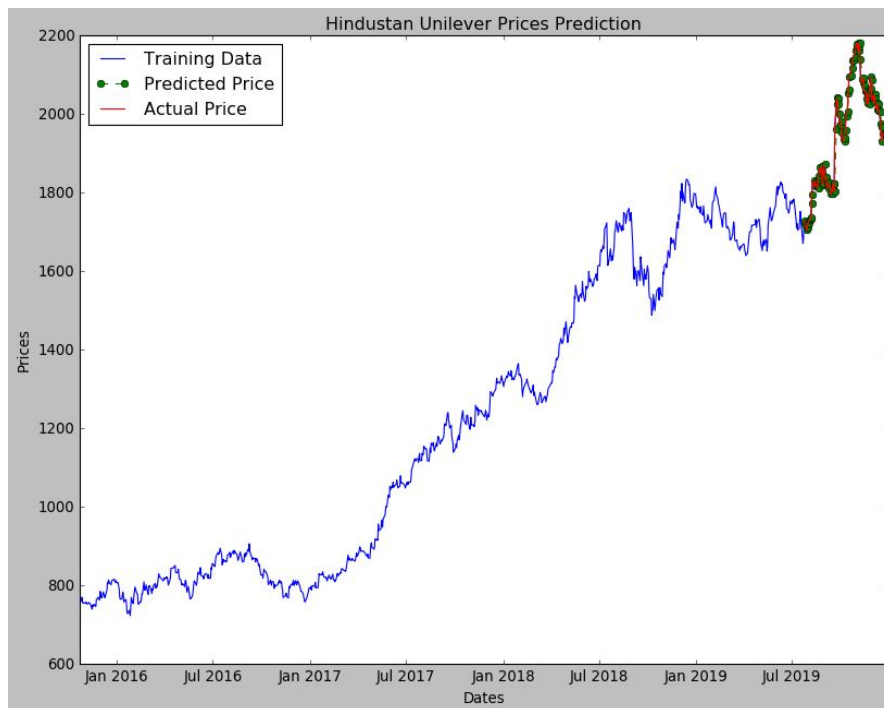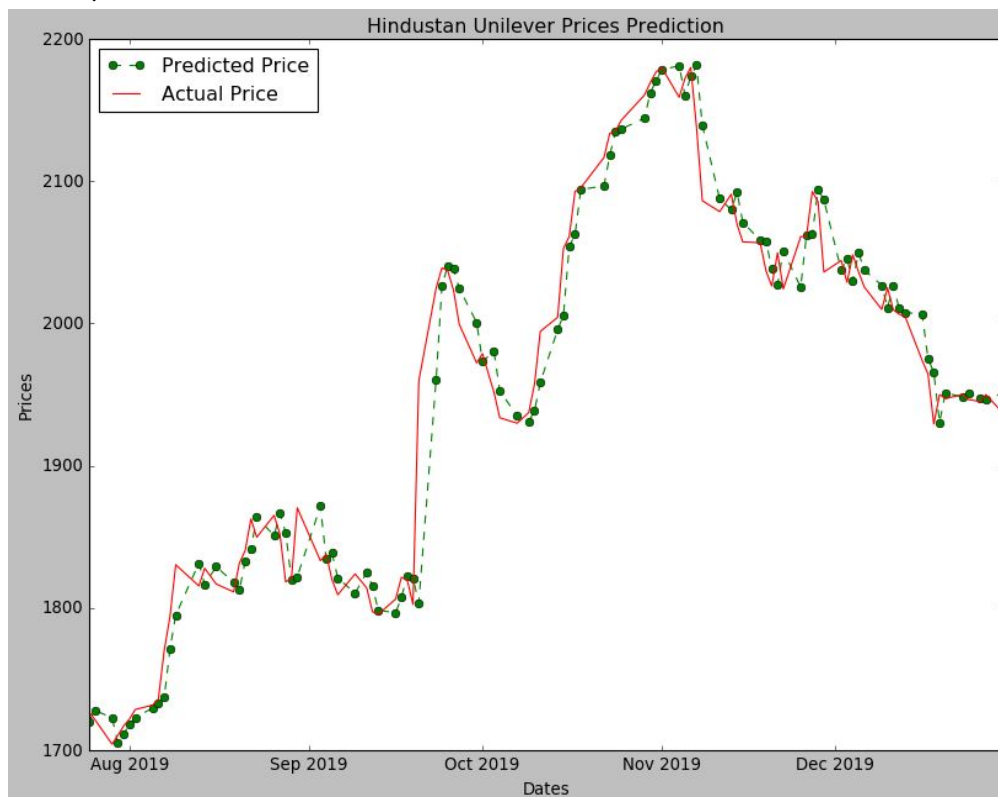Hindustan Unilever Autocorrelation plot

Graph of the training and testing data-
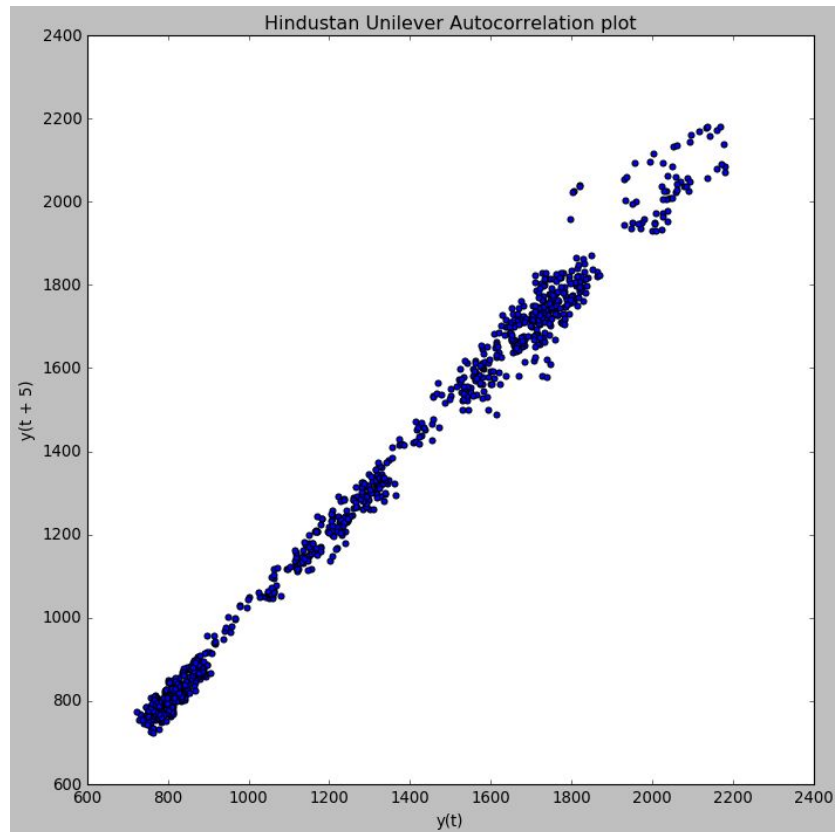
Hindustan Unilever Prices

Prediction and actual change in stock values for Hindustan Unilever Ltd stocks -

Closeup of Predicted and Actual Price -

Results:-

Root Mean Square Error (RMSE) is the parameter which we have chosen to judge the performance of the model. It is a measure of how spread out the residuals are. In other words, it tells us how concentrated the data is around the line of best fit. Therefore, smaller the RMSE, better is the performance of the model.

	Moving Average	LSTM	ARIMA
ICICI	67.20	13.57	8.78
Larsen & Toubro	87.45	34.98	23.86
Hindustan Unilever	252.56	50.73	25.60

Below are the results of performing the Markowitz Portfolio Selection Model-
L & T - 8.8%
ICICI Bank - 0.12 %
Hindustan Unilever Ltd - 91.08%

Conclusion -

Thus, we can see that the ARIMA model performs the best. And after observing the predicted values and plotting them against the actual values, we can confirm the Portfolio Allocation is optimal because we can see that Hindustan Unilever Ltd stocks are increasing in price in the future and hence it can be allotted most of the portfolio percentage as it will be significant in giving us great returns. The model does not allot 0% to ICICI (the stock which has been allotted the least amount of portfolio because its value decreases) because the whole point of portfolio allocation and optimization is that all the stocks or securities are allotted some percentage of the portfolio so as to promote diversification so that if one industry is affected by some bad returns, the other stocks of other industry can offset those losses for profit.

Github Link for the code:

https://github.com/aashalkamdar/DMPA-Lab-Project
(P.S. If code doesn't load in the first try, please try reloading as a jupyter notebook takes some attempts to load in Github)