STOCKX

Predicting Portfolio Performance

M V D Satya Swaroop [114388425]

Idea & Significance

- We intend to identify a optimal portfolio allocation for a given set of stocks using Efficient Frontier and Sharpe Ratio
- We used Linear Regression and Long and Short term neural networks to predict the portfolio's future returns given historical performance of stocks.
- We trained the models on past data of closing price of portfolio of optimal selection with minimal risk and high return. This would predict the stock performance in the future trading days.
- We can eventually use other variables such as news headlines, behavior patterns to train the model and predict further movements of portfolio in a event driven fashion.

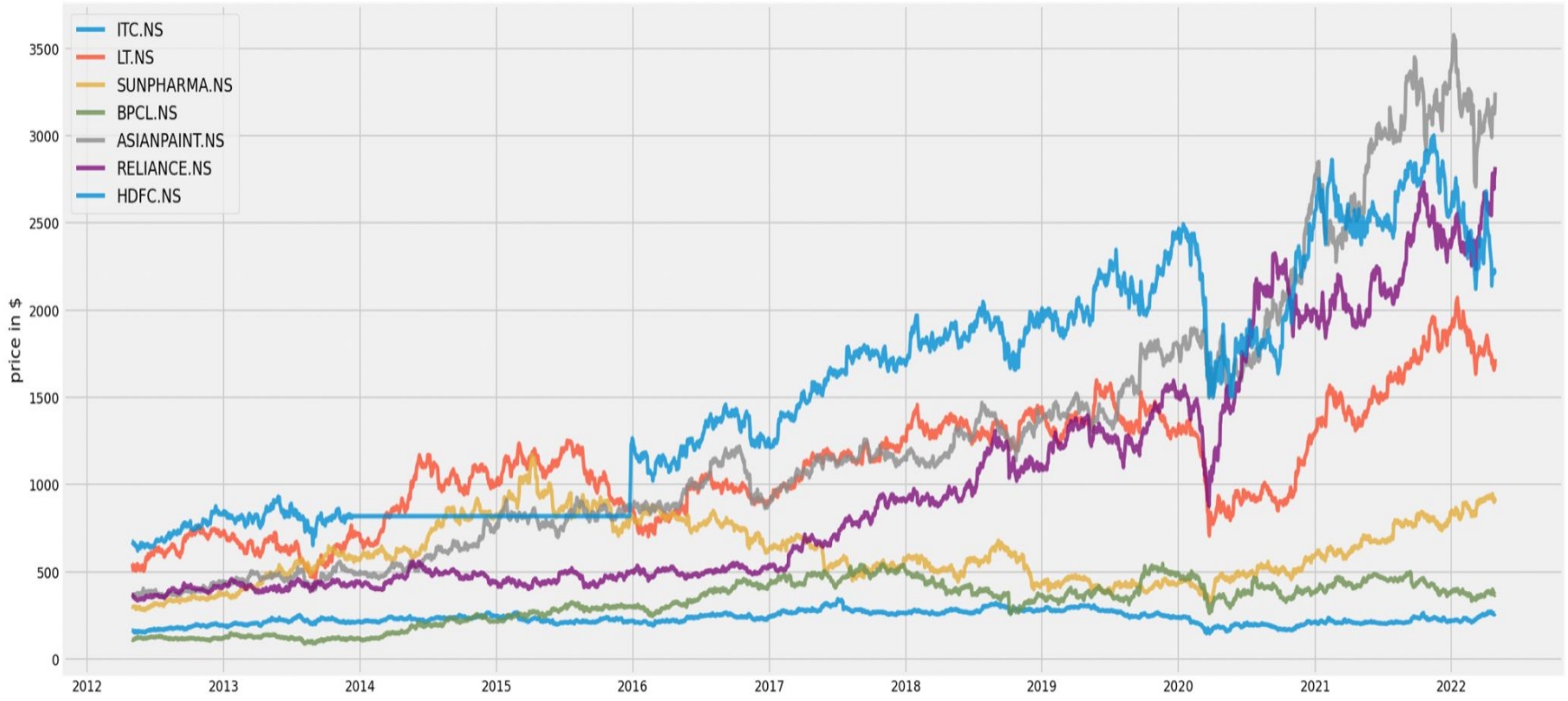
Team Expertise level and Learnings.

- I come from a computer science background good with developing web applications, using data structures and algorithms.
- In this project, I learned the theoretical concept behind Linear Regression and LSTM Neural Network. Also understood how to train the model and tune it in order to fit the model accurately on the training data. Also compute model metrics to measure the model performance on test data.
- Used Yahoo finance API to get stock data over the past decade, Also learned about risk adjusted return by Sharpe ratio and implementing efficient frontier to find optimal portfolio allocation.

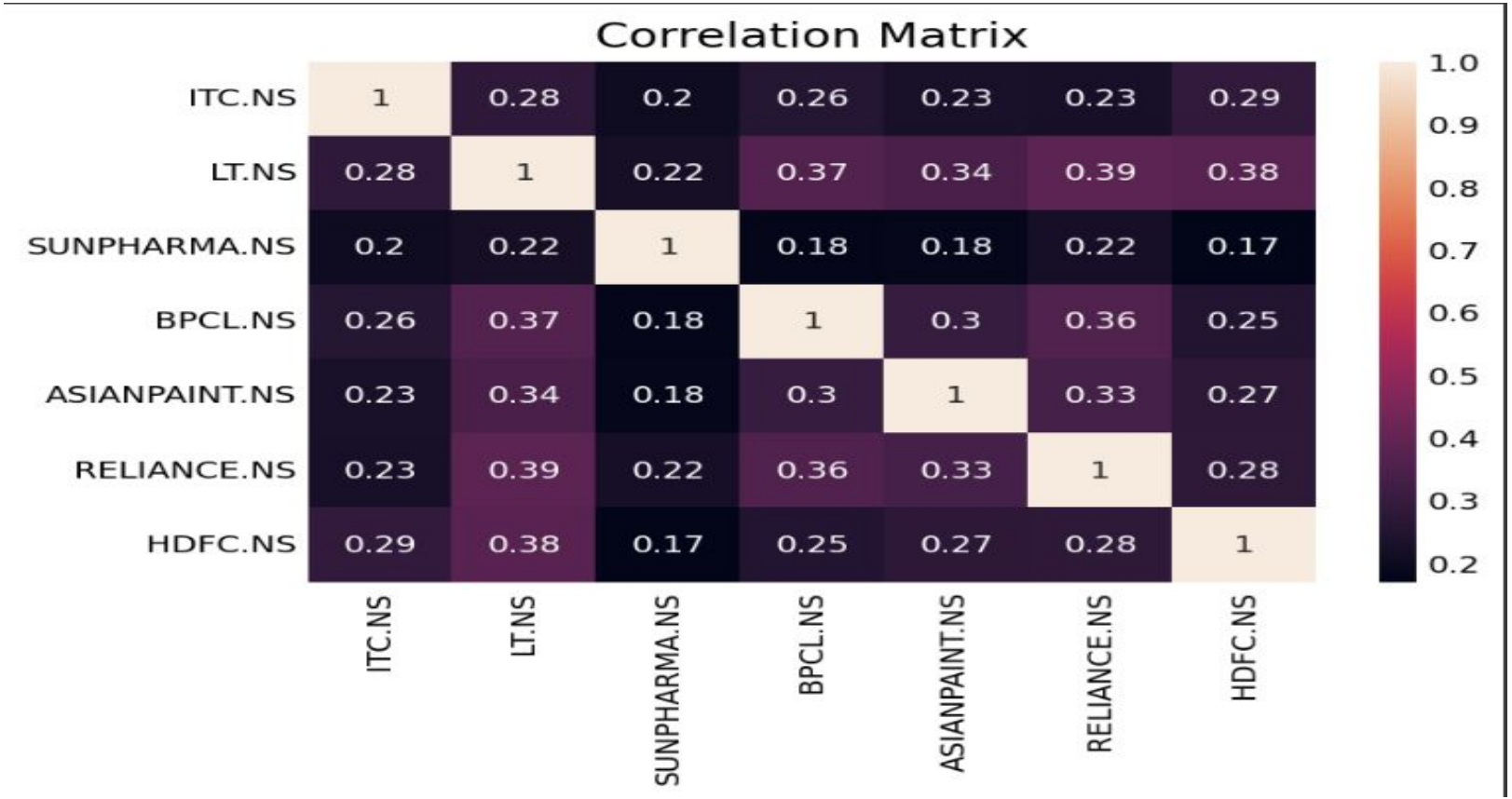
Techniques and Tools Used:

- Yahoo Finance Library
- Pandas and Numpy
- Matplotlib and Seaborn
- Sharpe Ratio & Efficient Frontier
- Linear Regression
- Recurrent Neural Networks: LSTM
- Sklearn
- Keras
- Machine Learning and Modularising Code

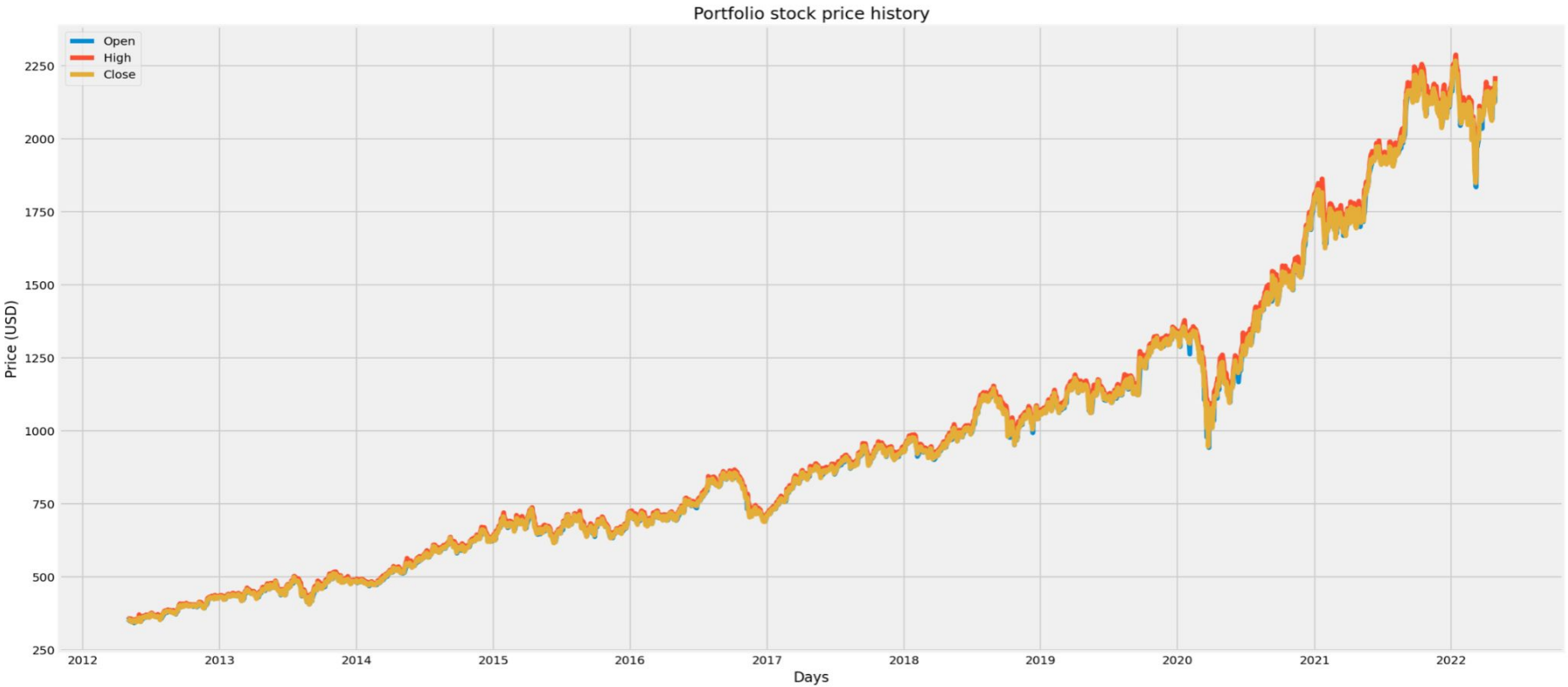
Stocks Performance Over the last Decade



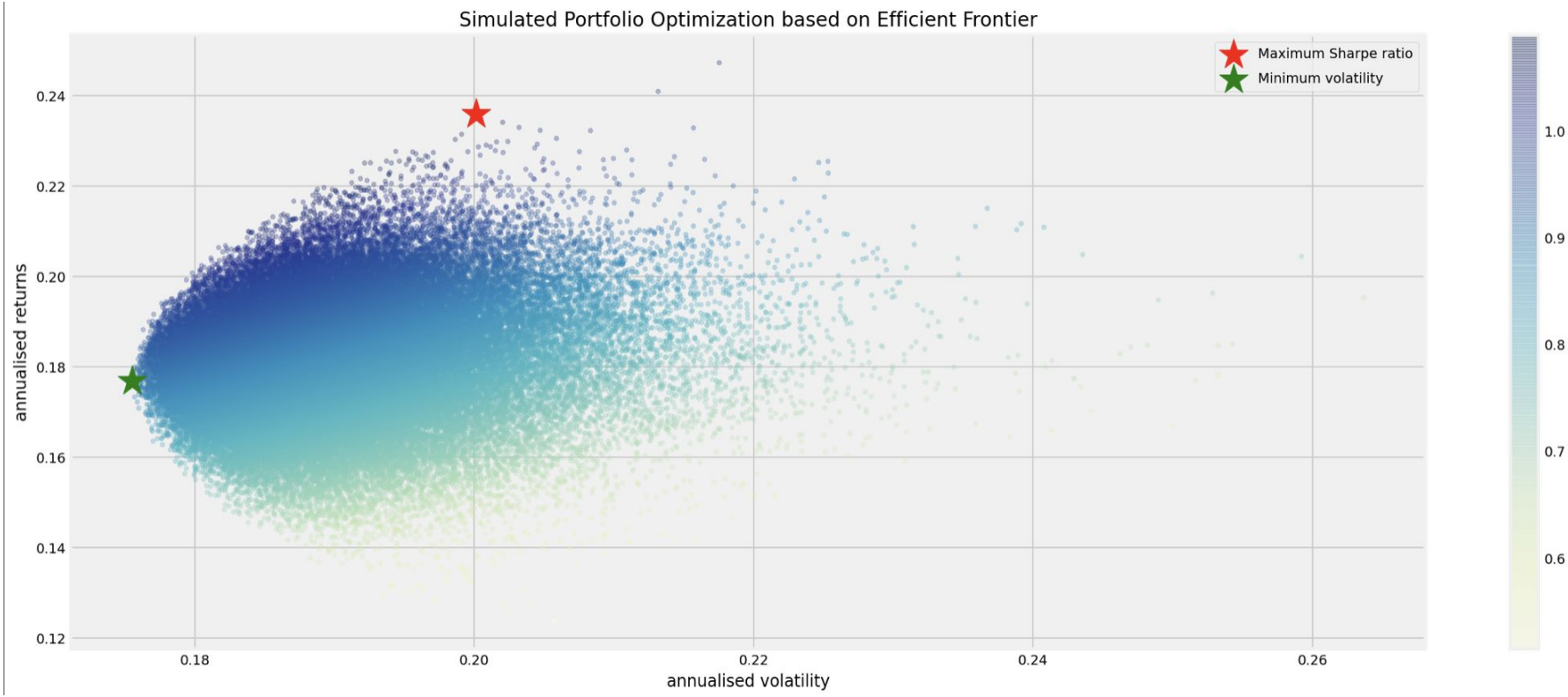
Correlation Matrix



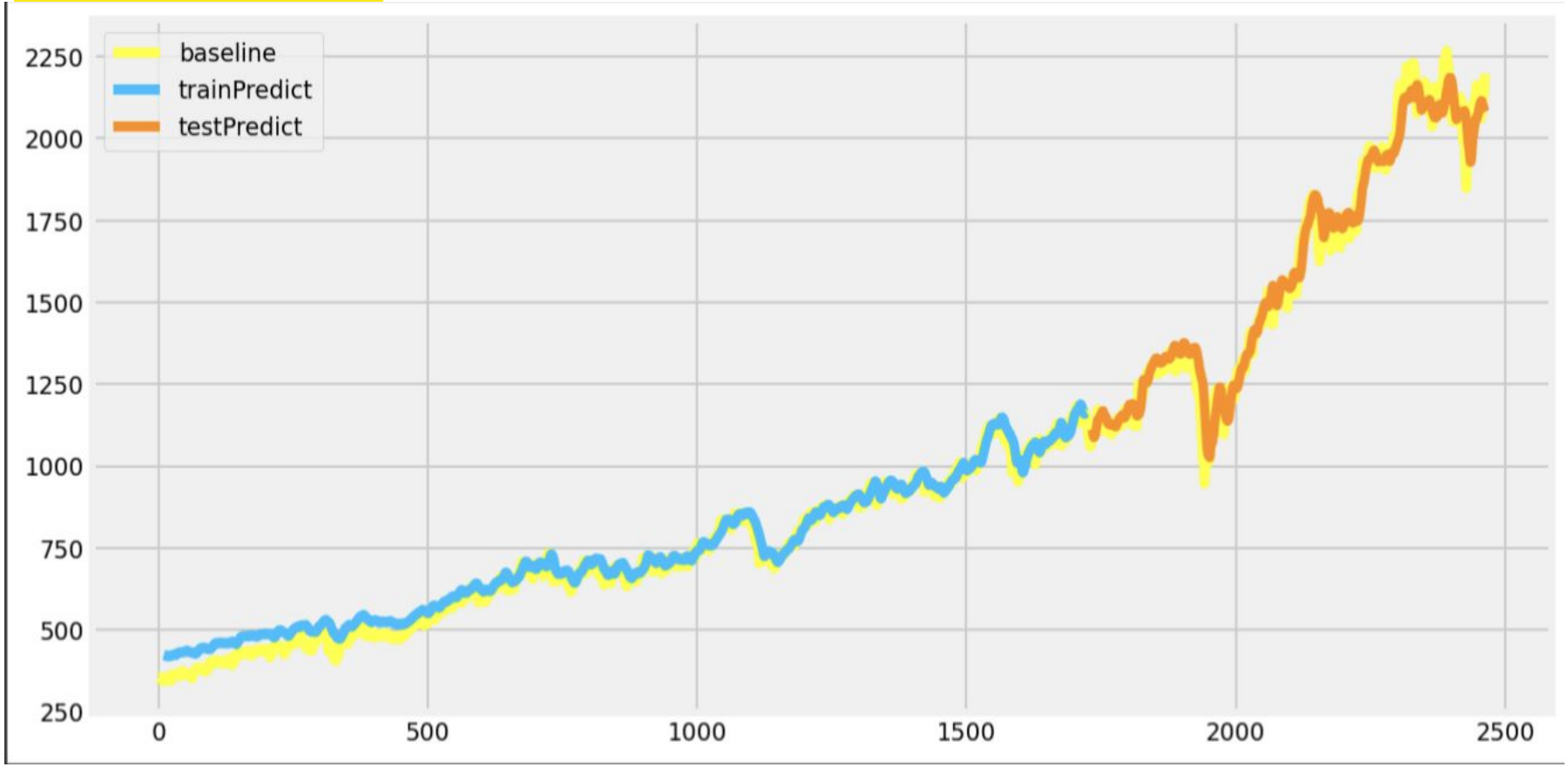
Portfolio Performance



Efficient Frontier



LSTM Prediction



Challenges/ Results and Outcomes.

While Linear Regression is performing on par with LSTM, LSTM is better at predicting both outliers and central values simultaneously. It's able to perform predictions accurately for longer extend to trading days. Hence this project accomplished to find optimal allocation of capital for stocks and tried to predict the portfolio's closing value in the future trading days.

LSTM Metrics: Prediction for 600 Trading days.

train_mse = 0.004419493016489919

test_mse = 0.012377929999732707

train_rmse = 0.06647926756884374

test_rmse = 0.11125614589645243

train_mae = 0.05310097557756166

test_mae = 0.08351598745942392

train_r2 = 0.9777689308163915

test_r2 = 0.978556546018192

Results of Linear Regression forecasting 600 trading days.

Prediction Using LR

Values Forecasted : 600

Model Name : LR

Train RMSE : 0.0032871086917017644

Train R2 Score : 0.9983526045346967

Test RMSE : 0.022054269727648443

Test R2 Score : 0.6521293413304413