Stock Market Analysis using Supervised Machine Learning

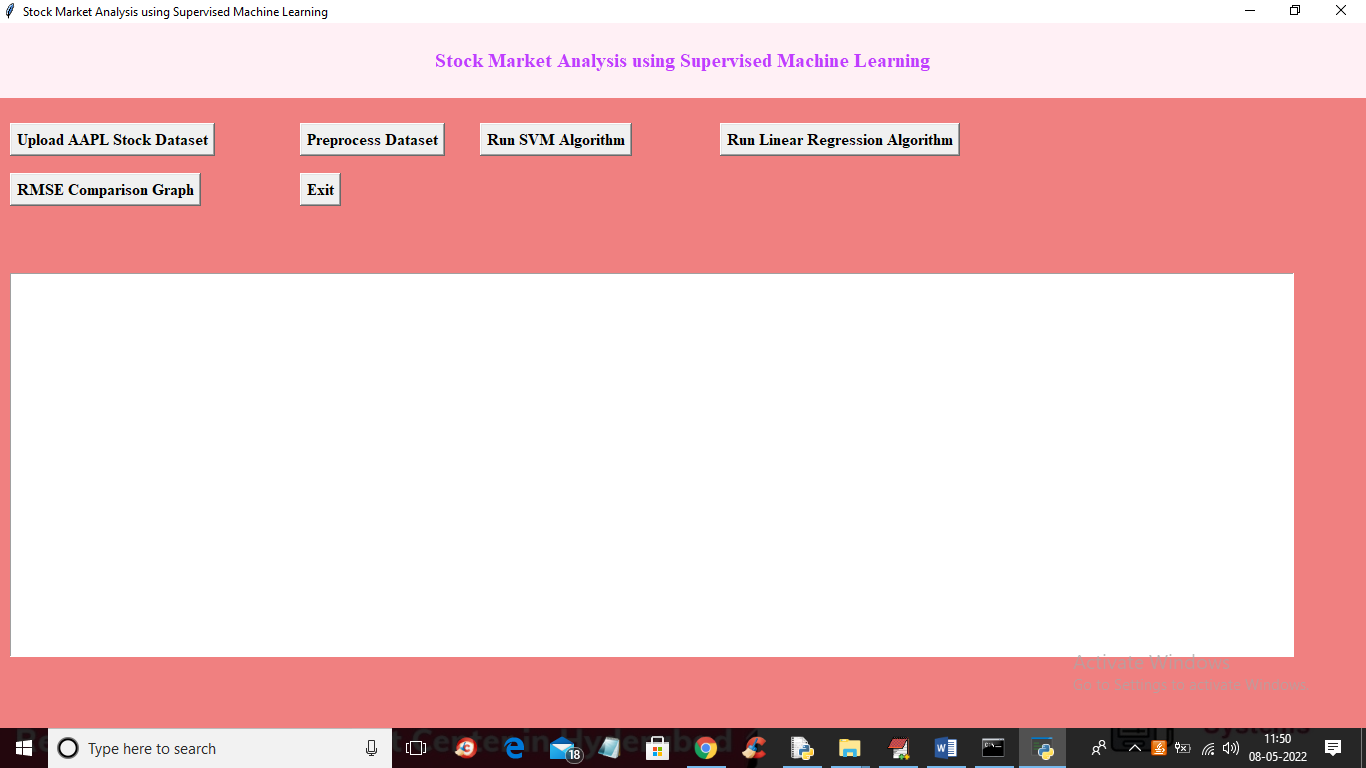
Now-a-days may peoples are investing their money in stock market and this market is one of the sophisticated market where many peoples can reach overnight and may loss also and to avoid such problem it’s become necessary to analyse market before investment but manual analysis often prove wrong so author of this paper is using Machine Learning algorithms to predict future prices as this algorithms prove their efficiency and accuracy of prediction in almost all fields such as health prediction, cyber-attack prediction, credit fraud prediction and many more. So author applying same algorithm on stock market using machine learning algorithms called Linear regression and SVM.

All machine learning algorithms get trained on past data and then generate a model and this model can be used to predict future data. So in this paper author training both algorithms using past AAPL stock data and then predicting future prices. To prove capability of this algorithm author is splitting dataset into train and test and then training algorithm with train data and then this trained model will be applied on test data to predict its future prices. The difference between original test data and predicted data will be consider as RMSE (root mean square error). The lower the RMSE the higher is the correct prediction and based on that accuracy will be calculated.

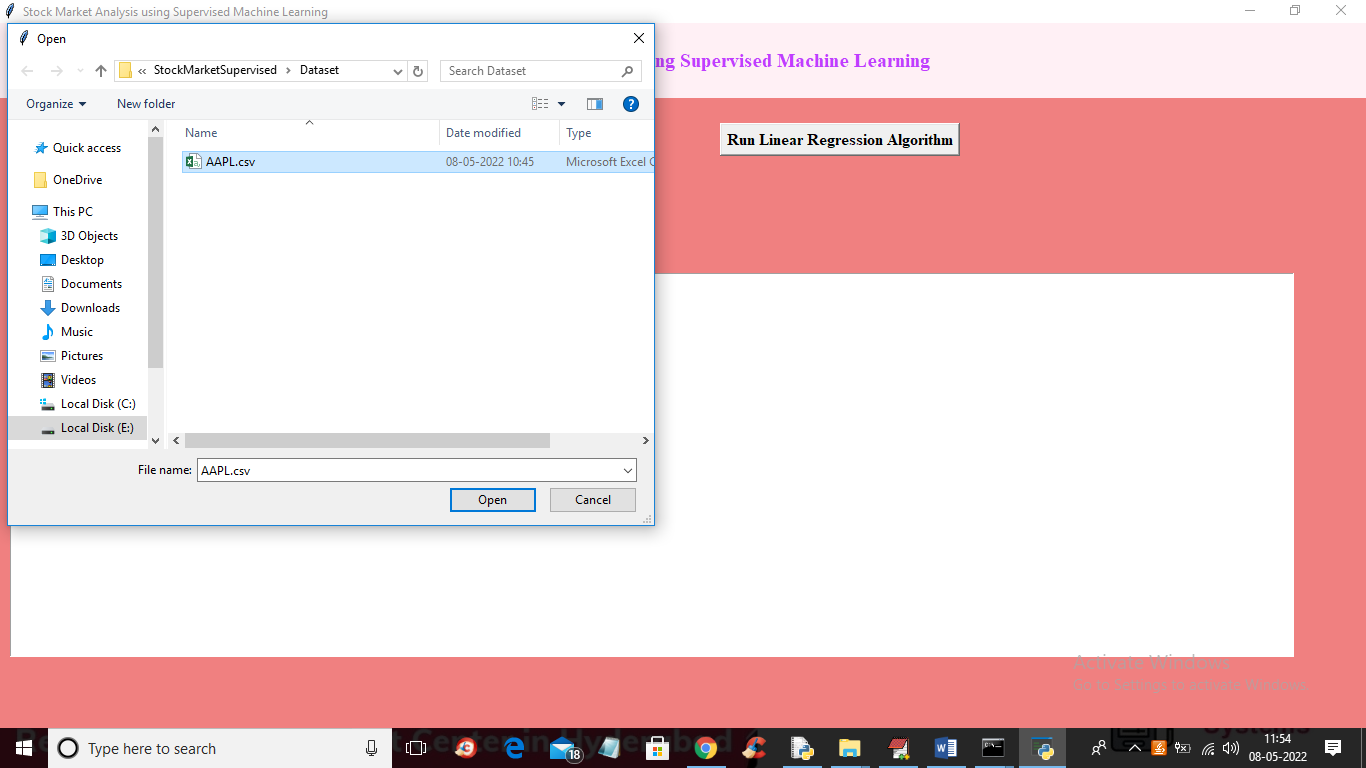
To implement this project we have used AAPL dataset.

SCREEN SHOTS

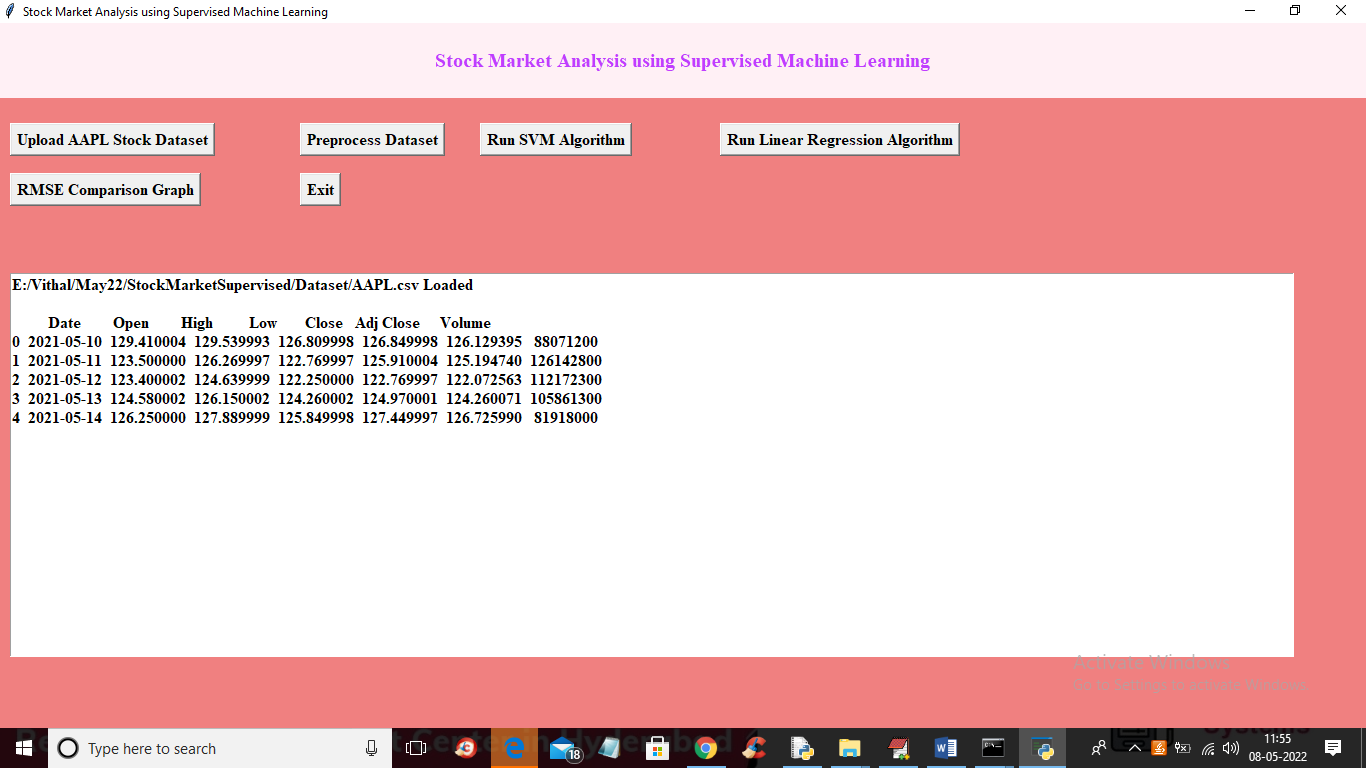
To run project double click on ‘run.bat’ file to get below screen



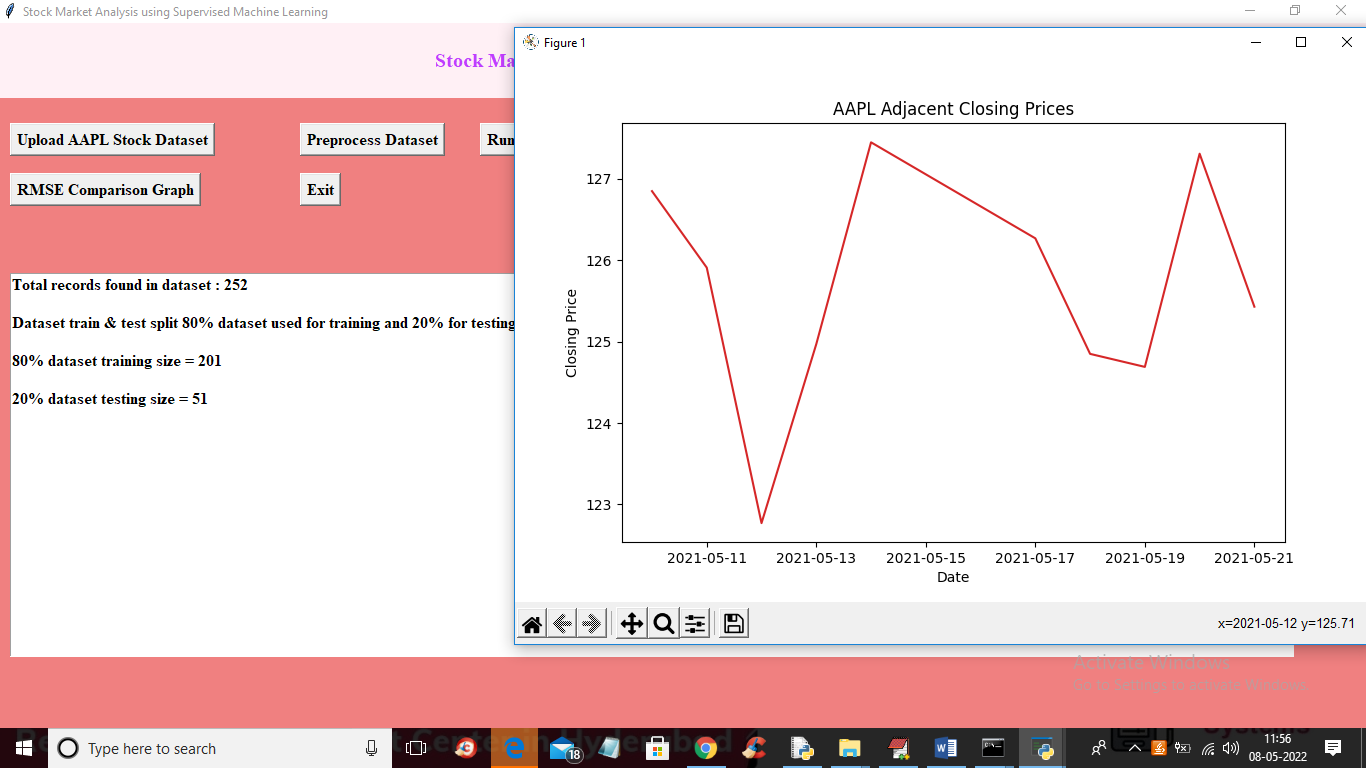
In above screen click on ‘Upload AAPL Stock Dataset’ button to upload dataset



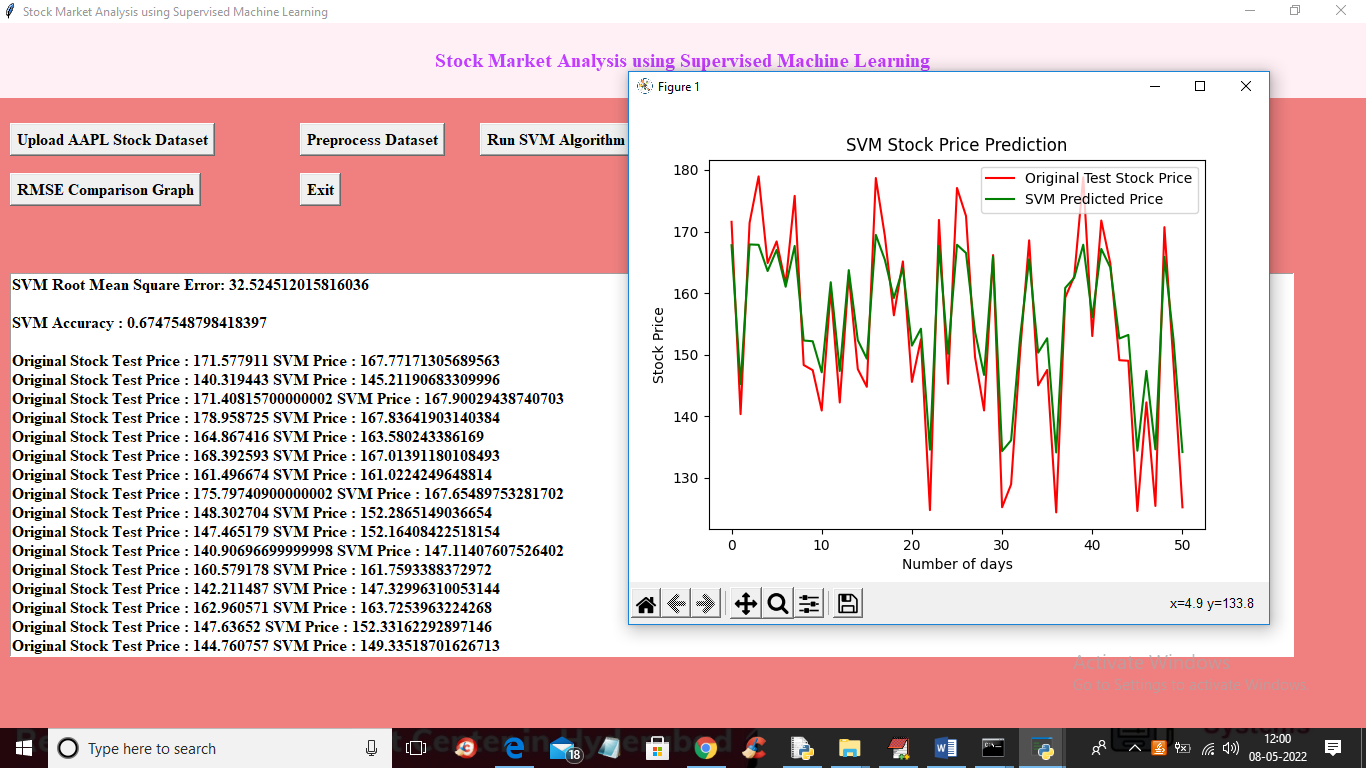
In above screen selecting and uploading ‘AAPL.csv’ dataset file and then click on ‘Open’ button to load dataset and to get below output



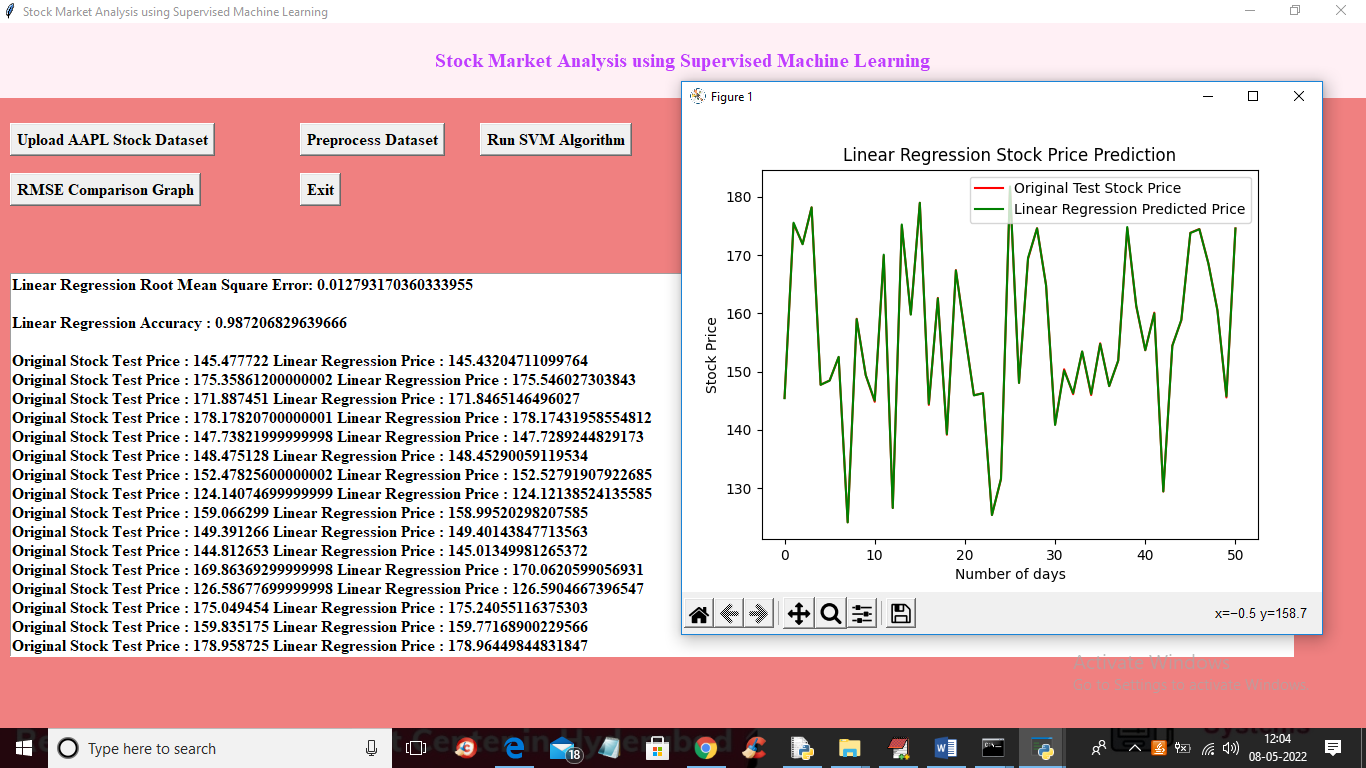
In above screen dataset loaded and now click on ‘Preprocess Dataset’ button to normalize dataset values and then split dataset into train and test part for training and testing machine learning algorithms



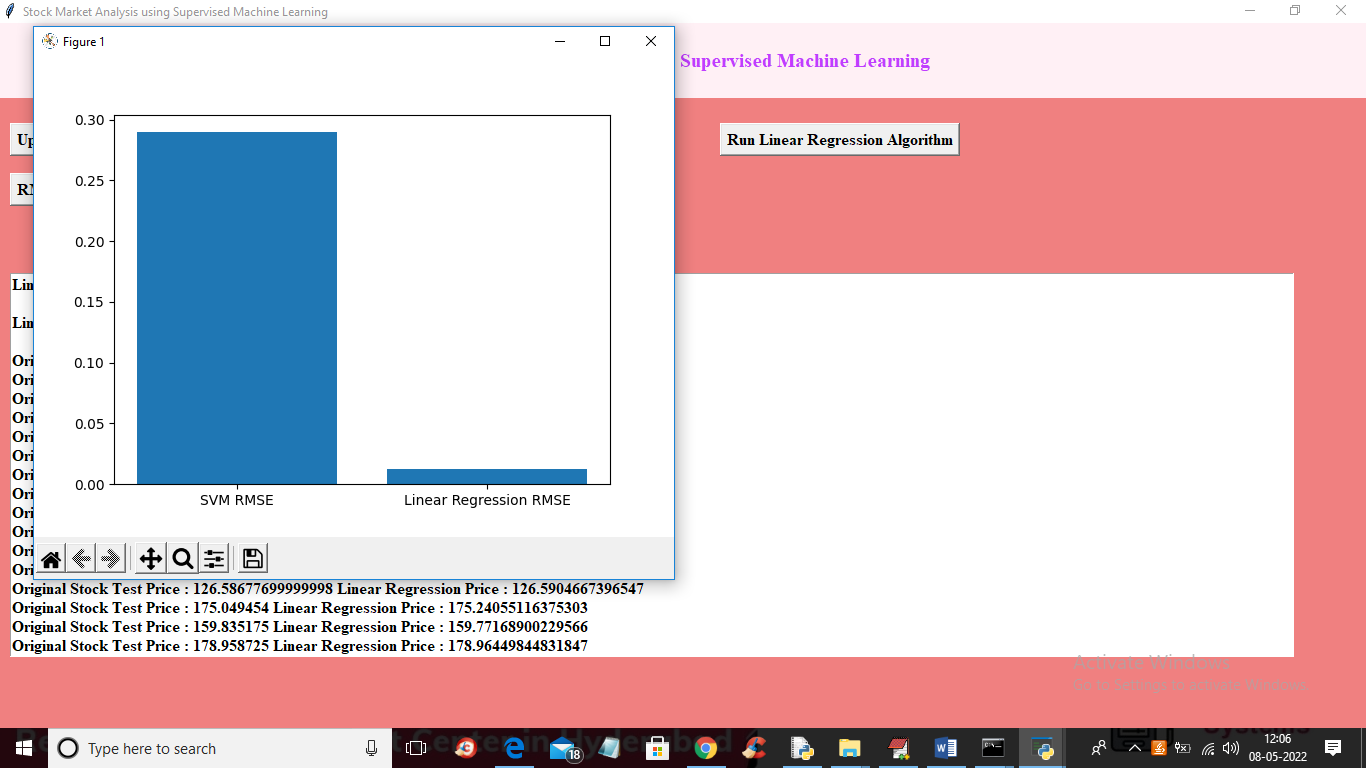
In above screen in text area we can see dataset contains 252 total records and using 201 records for training and 52 records for testing and in graph x-axis represents DATE and y-axis represents closing values on those dates and now close above graph and then click on ‘Run SVM Algorithm’ button to train SVM on training data and then test on test data and get below output



In above screen we got SVM RMSE as 32 and accuracy as 67 and then we can see original stock test PRICES and SVM predicted TEST prices and you can see difference between original and predicted prices and in graph x-axis represents number of days for future and y-axis represents stock prices and green line represents PREDICTED prices and red line is the original prices and SVM is not good in prediction so its accuracy is less and in graph we can see huge gap between predicted and original prices. Now close above graph and then click on ‘Run Linear Regression Algorithm’ button to get below output



In above screen with Linear regression we got RMSE as 0.012 and accuracy as 98% and in graph we can see both lines are fully overlapped so there is not much difference between original test and predicted prices so Linear regression is best in stock prices prediction and in text area you can see both original price and linear regression predicted price and you won’t find much difference between them. Now close above graph and then click on ‘RMSE Comparison Graph’ button to get below output



In above graph x-axis represents algorithm names and y-axis represents RMSE values and in both algorithms Linear regression got less RMSE error so linear regression is best