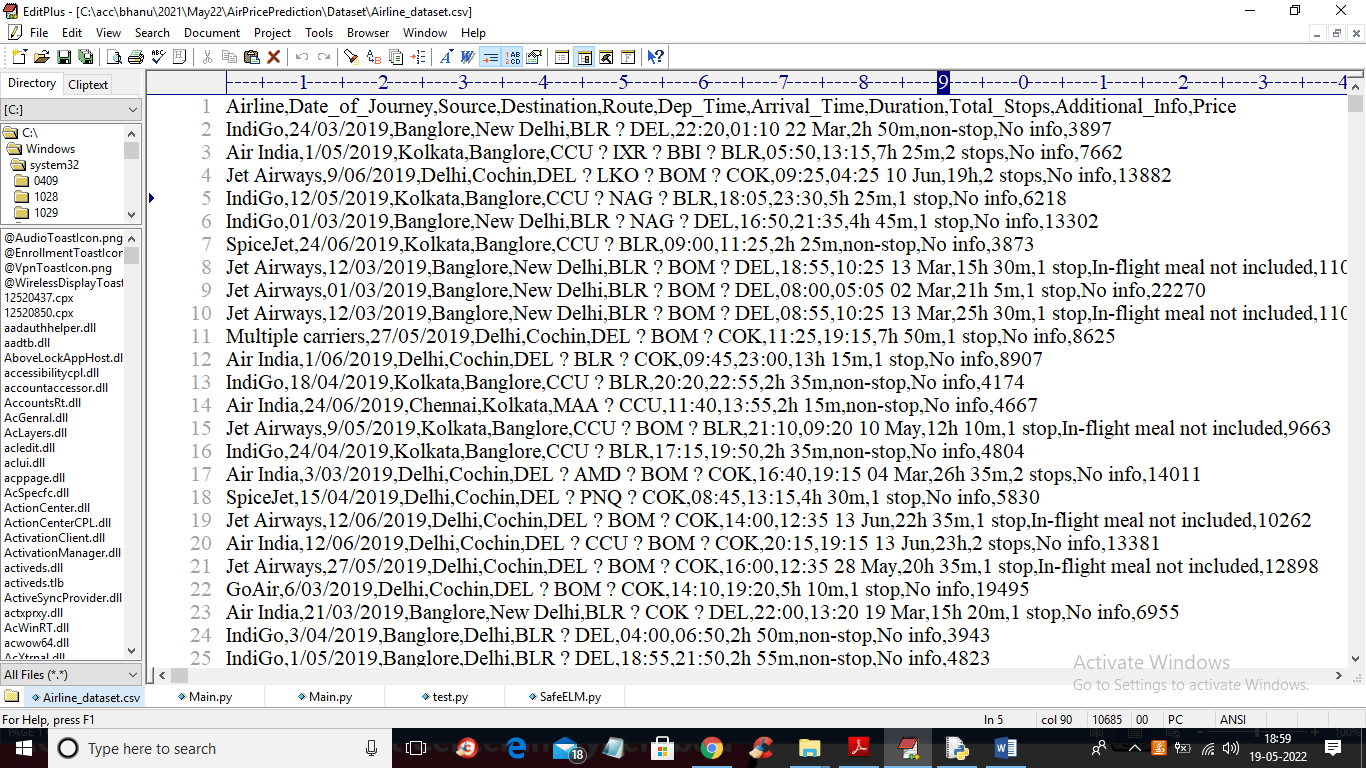
A Holistic Approach on Airfare Price Prediction Using Machine Learning Techniques

In this paper author is using various machine learning algorithms such as Polynomial and Linear SVM, Random Forest, Regression Tree, Bagging Regression, Linear Regression, MLP and ELM to predict air fare prices and evaluating each algorithm performance in terms of accuracy.

To train each algorithm author has used AIR FARE dataset from KAGGLE and this dataset can be downloaded from below link

<https://www.kaggle.com/code/vinayshaw/airfare-price-prediction/data>

Below screen showing dataset details



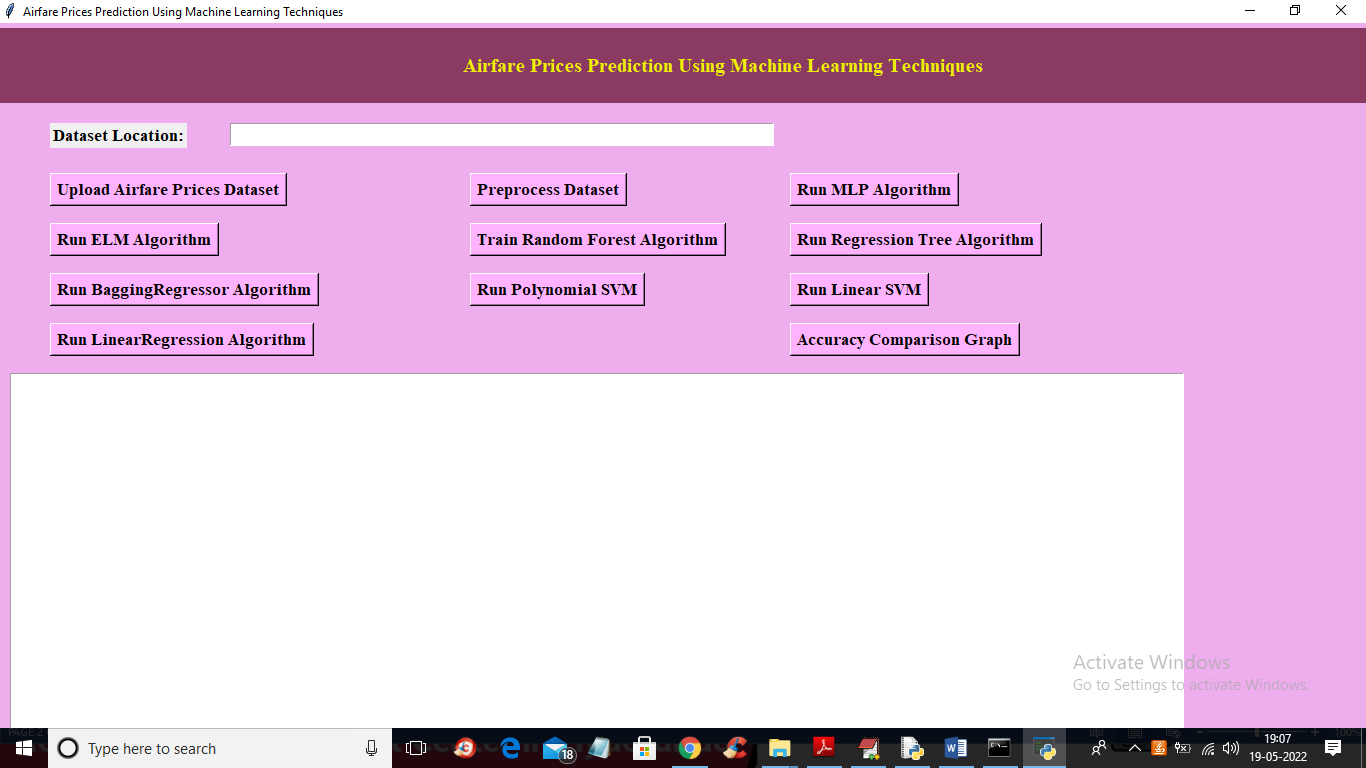
In above screen first row contains dataset column names and remaining are the dataset values and in last column we have PRICE which can be used as class label

To implement this project we have designed following modules

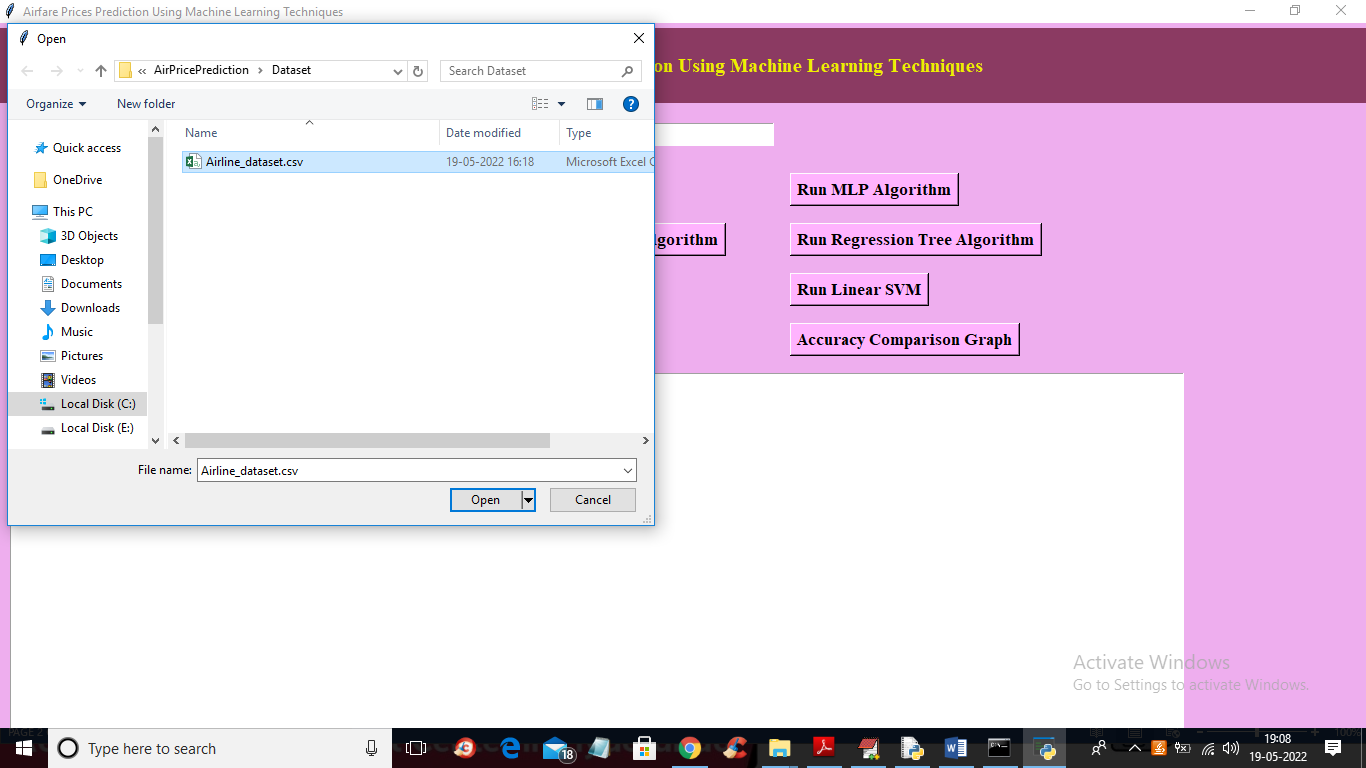
1. Upload Airfare Prices Dataset: using this module we will upload dataset to application
2. Preprocess Dataset: using this module we will removing missing values and then encode non-numeric string to numeric data by using Label Encoder class as machine learning accept only numeric data and above dataset contains some numeric and non-numeric data so by applying Preprocessing will convert that into numeric
3. Run MLP Algorithm: using this module we will split dataset into train and test and then trained algorithm by using TRAIN data and then apply test data to calculate its prediction accuracy. Similarly we will run all algorithms and compute their accuracy
4. Accuracy Comparison Graph: using this module we will plot accuracy comparison graph between all algorithms

SCREEN SHOTS

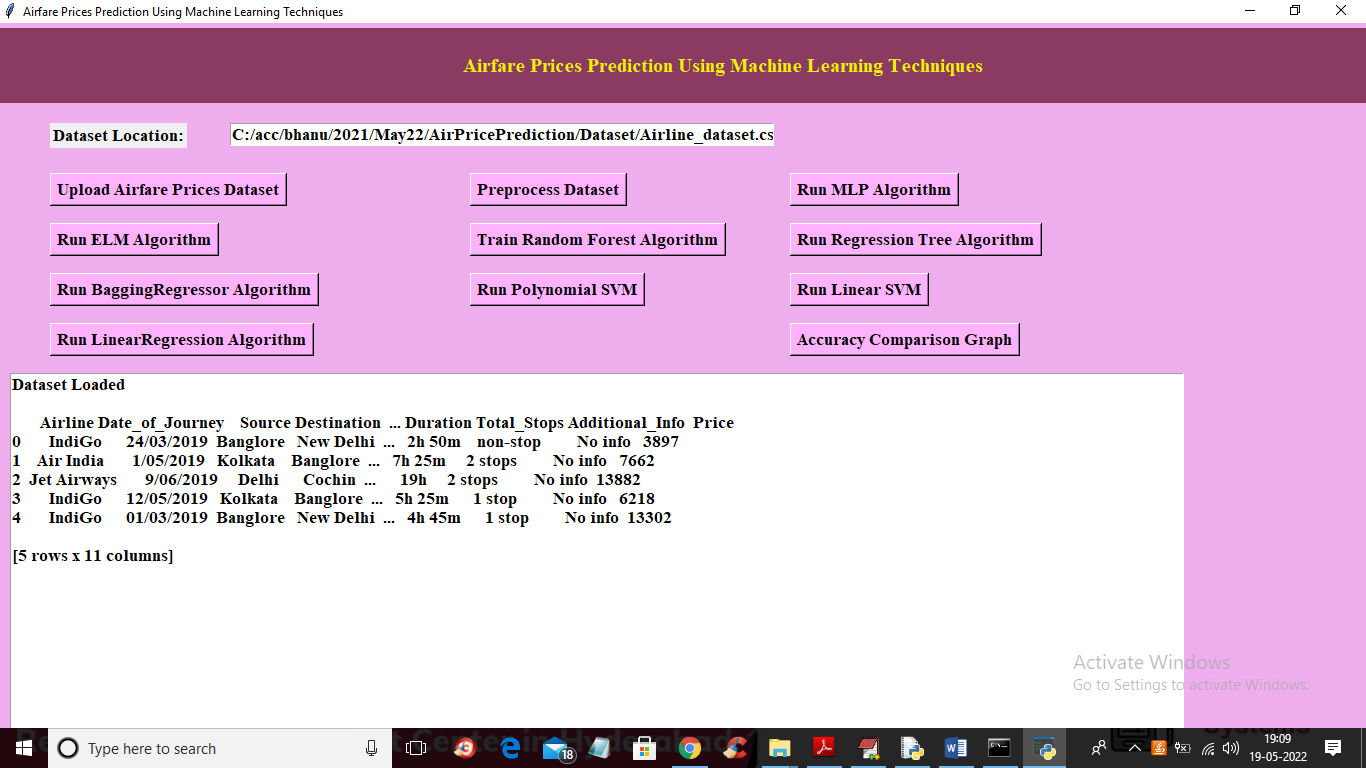
To run project double click on ‘run.bat’ file to get below output



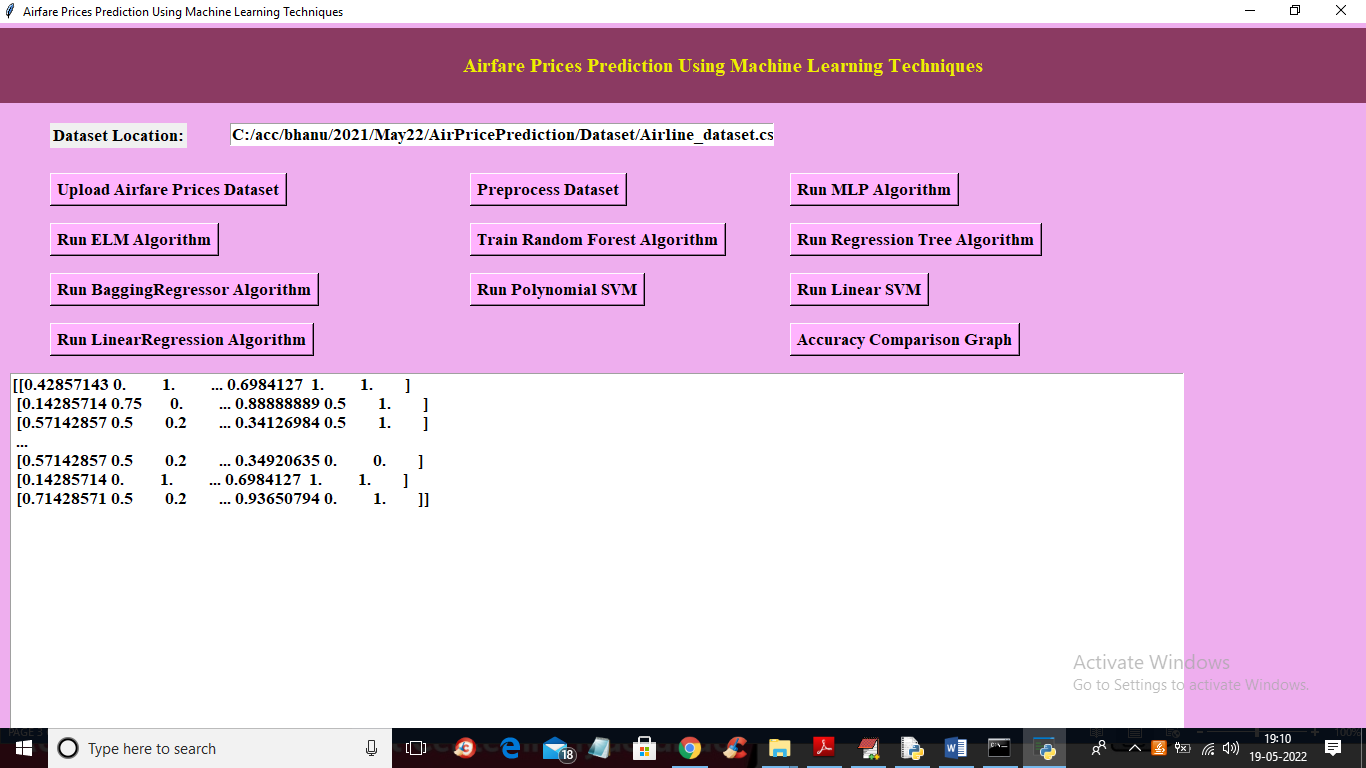
In above screen click on ‘Upload Airfare Prices Dataset’ button to upload dataset and get below output



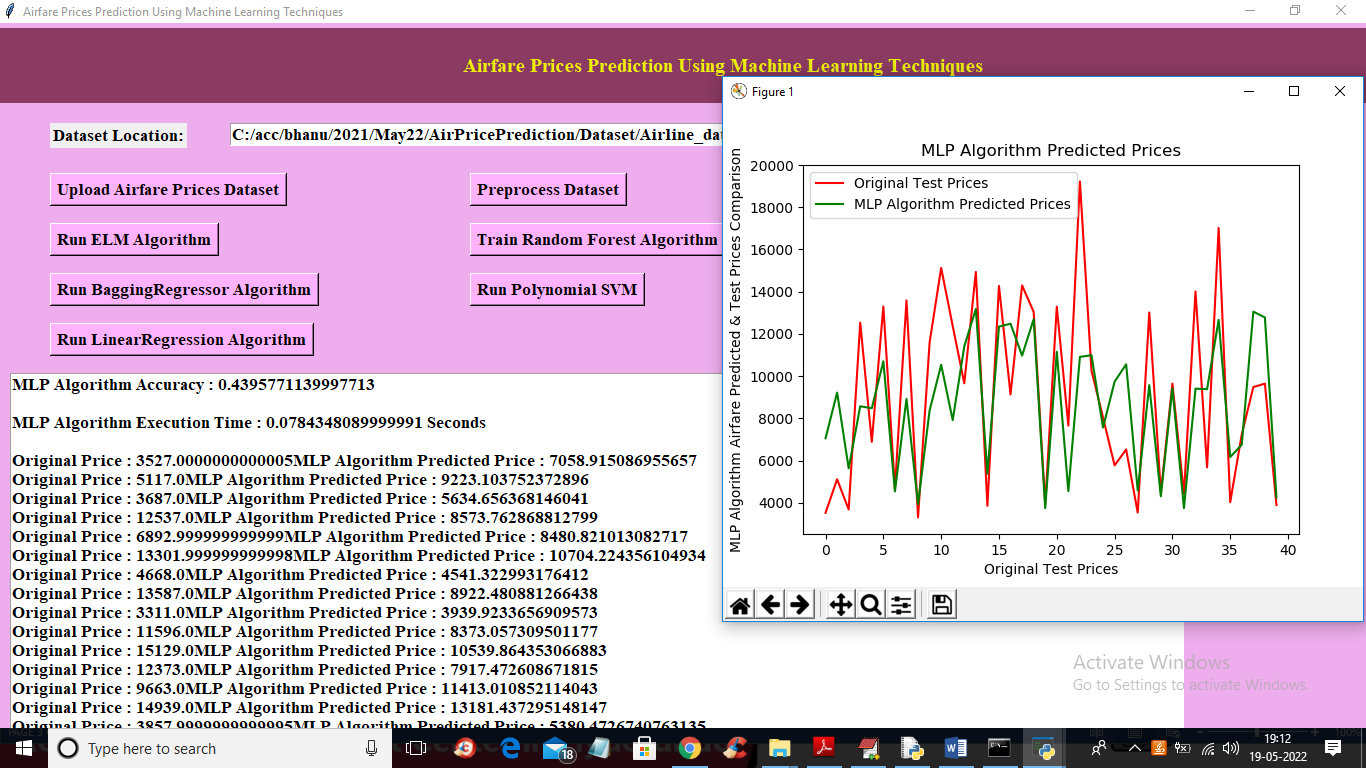
In above screen selecting and uploading dataset and then click on ‘Open’ button to load dataset and get below output



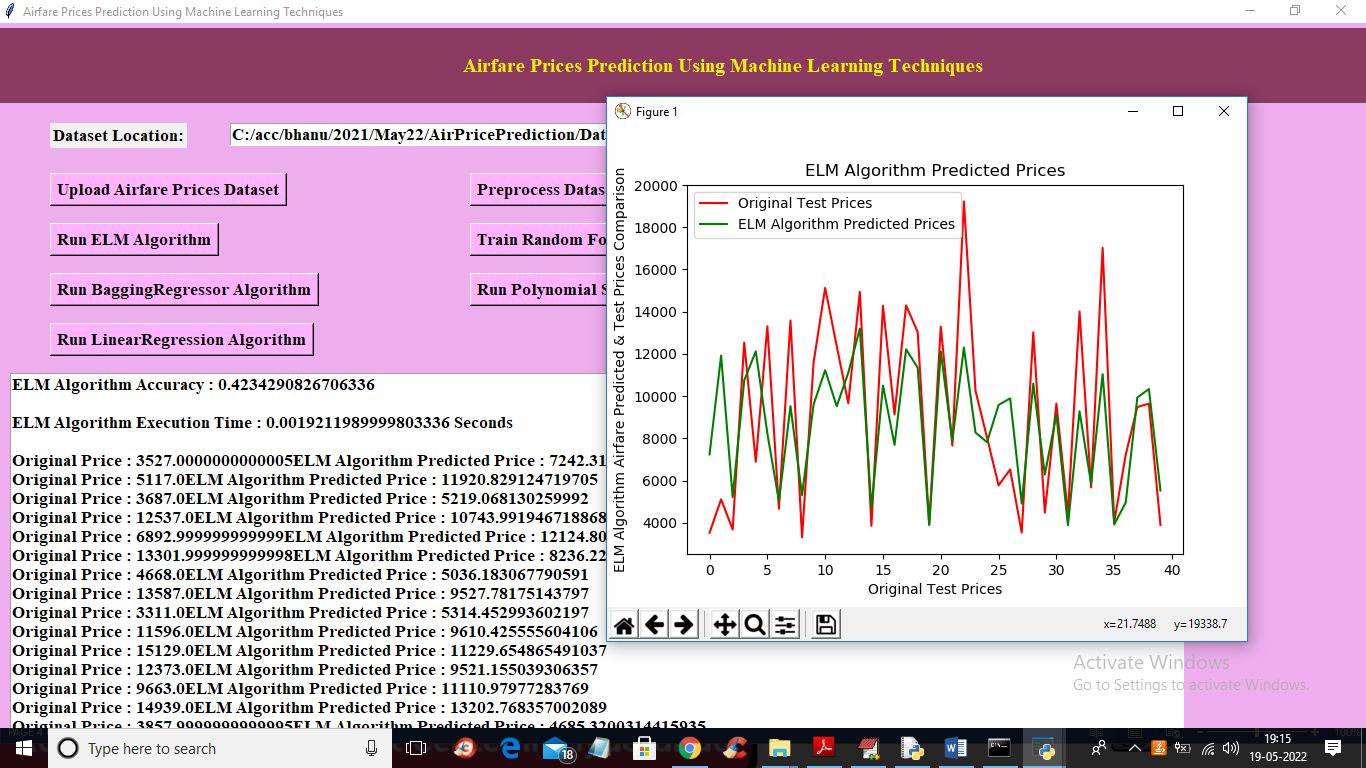
In above screen dataset loaded and its contains some non-numeric data so click on ‘Preprocess Dataset’ button to encode data to numeric and get below output



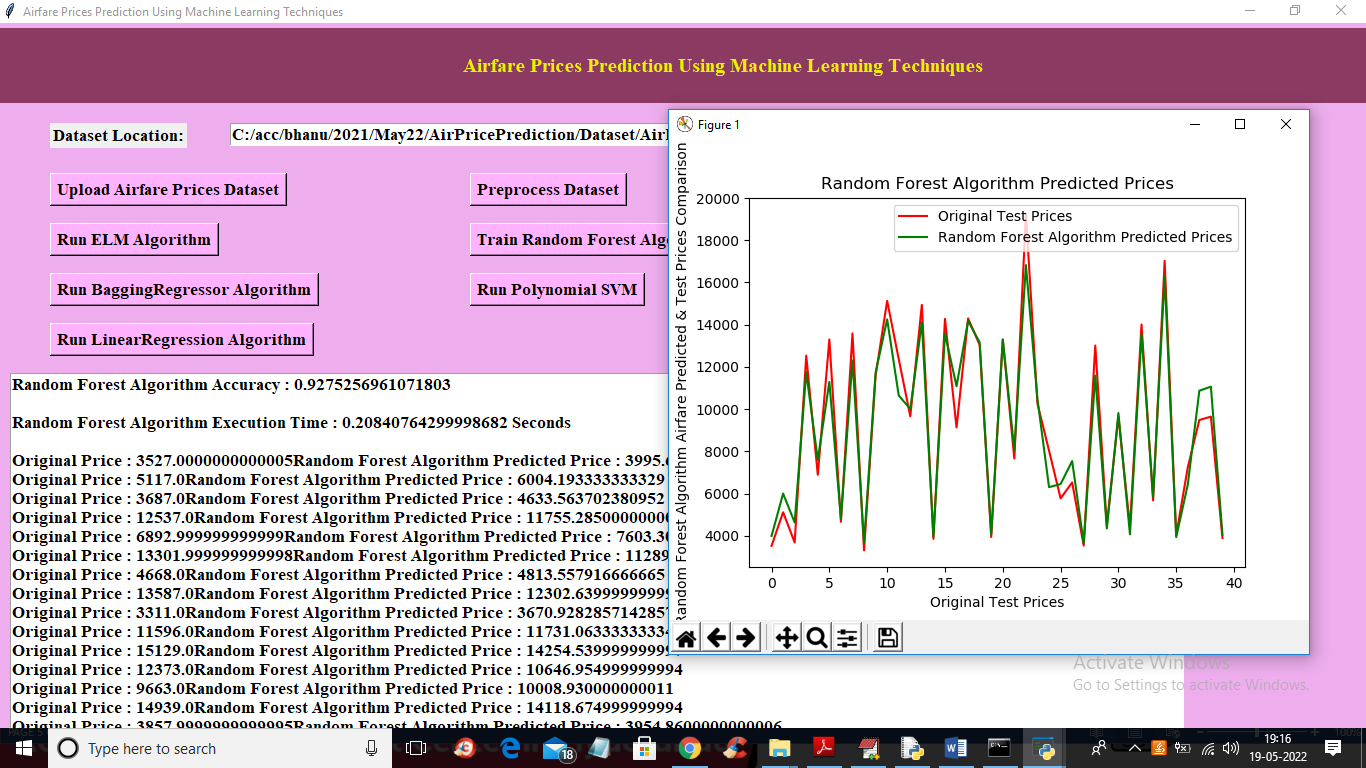
In above screen all dataset values converted to numeric and now click on ‘Run MLP Algorithm’ button to train MLP and apply trained model on test data to predict prices and calculate its prediction accuracy



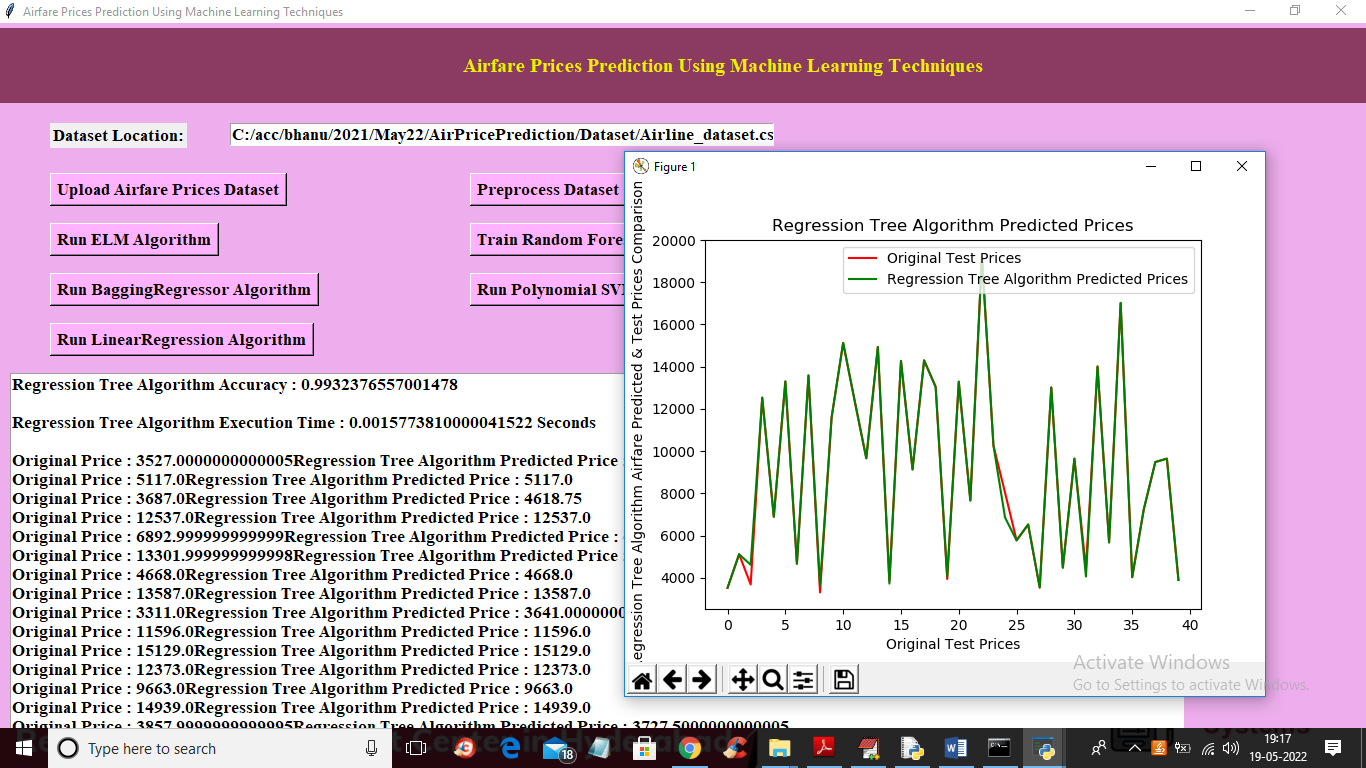
In above screen text area we can see MLP accuracy as 0.42% and its execution time is 0.07 seconds and then we can see original TEST prices and then prices predicted by MLP and same prices we are plotting in above graph. In graph X-axis represents DAYS and y-axis represents PRICES and red line represents original TEST prices and green line represents PREDICTED prices and we can see there is huge gap between both lines so its prediction is not accurate and due to that reason we got less accuracy and now close above graph and then click on ‘Run ELM Algorithm’ button to get below output



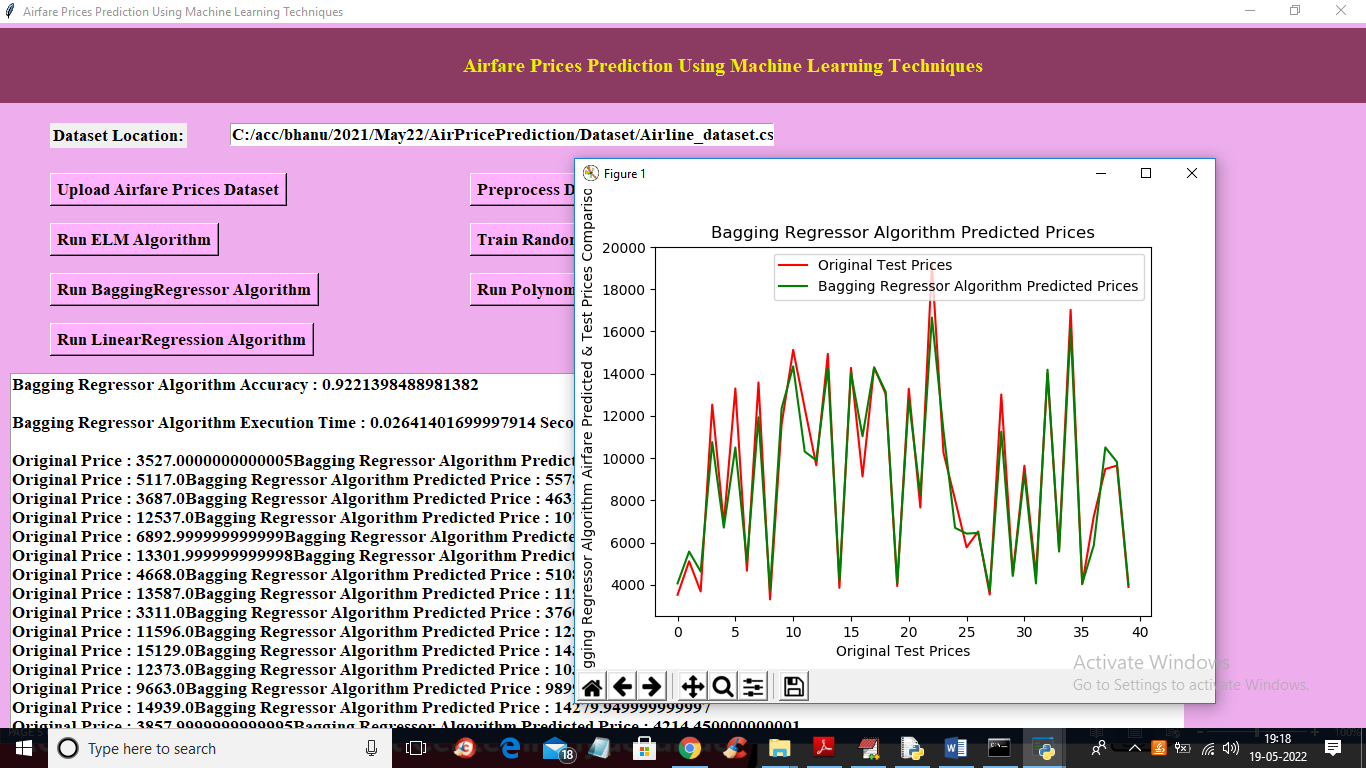
In above screen we can see ELM accuracy as 0.42 and its prediction is also not good and now click on Random Forest button to get below output



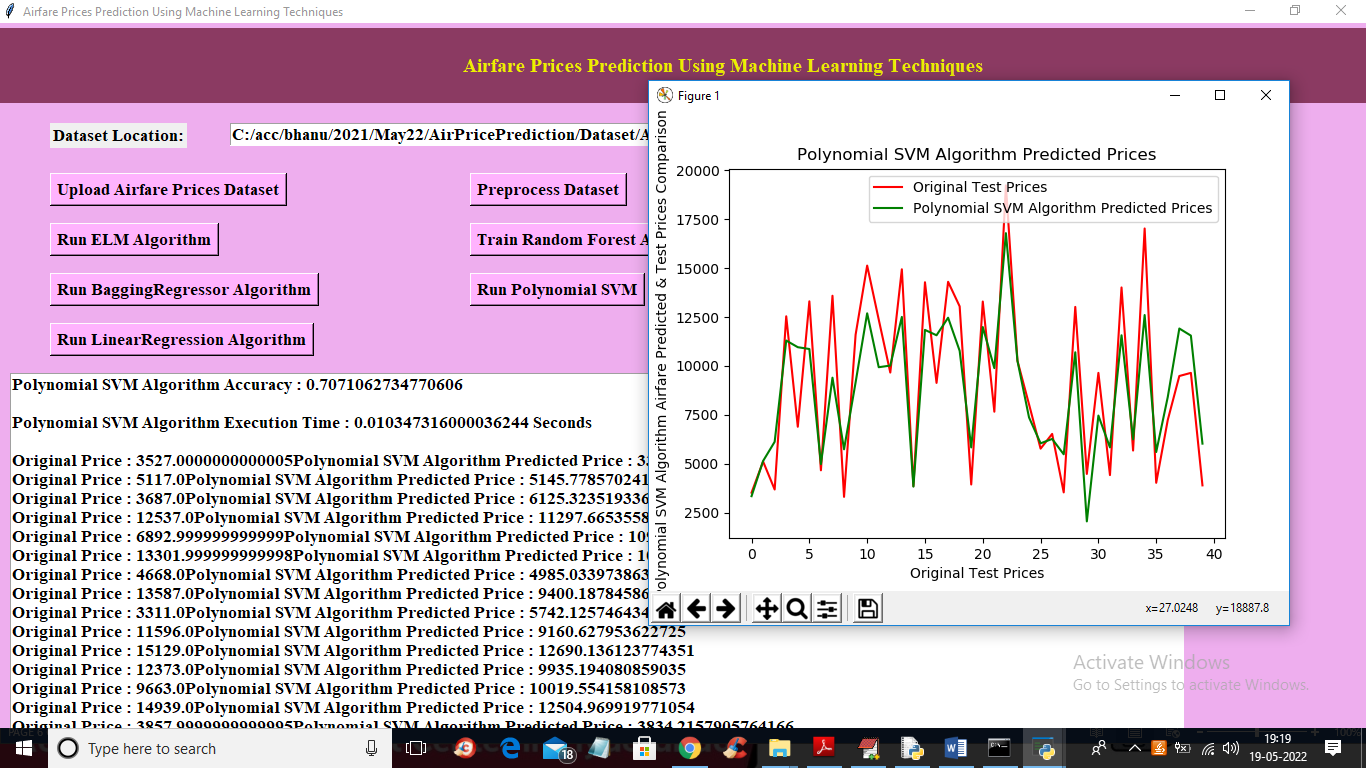
In above screen with Random Forest we got 0.92% accuracy and in graph we can see both lines are closely overlapping so between prediction price and test price there is not much difference so we can say random forest performance or its prediction is good and now close above graph and then click on ‘Regression Tree’ button to get below output



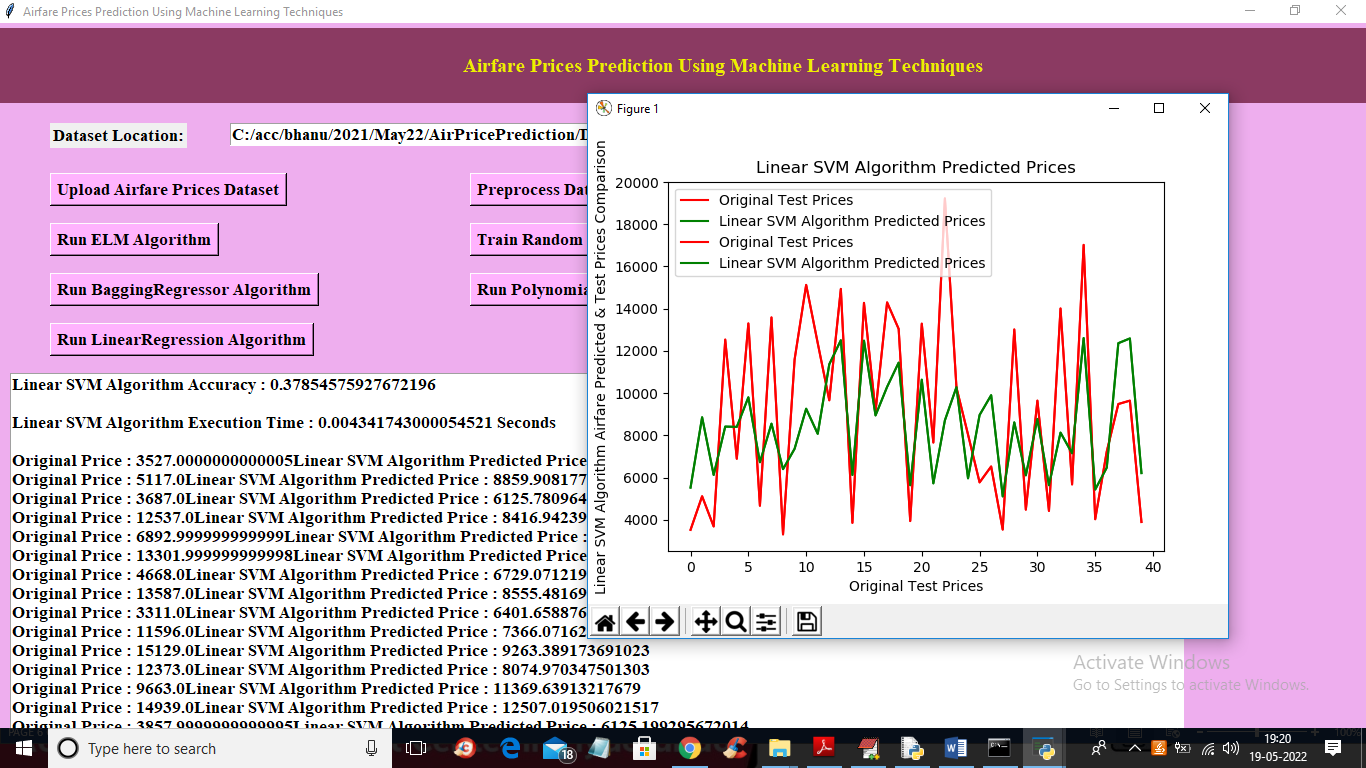
In above screen with Regression Tree we got 0.99% accuracy and in graph also we can see both line closely matching so its performance is good and now click on ‘Run Bagging Regression Algorithm’ button to get below output



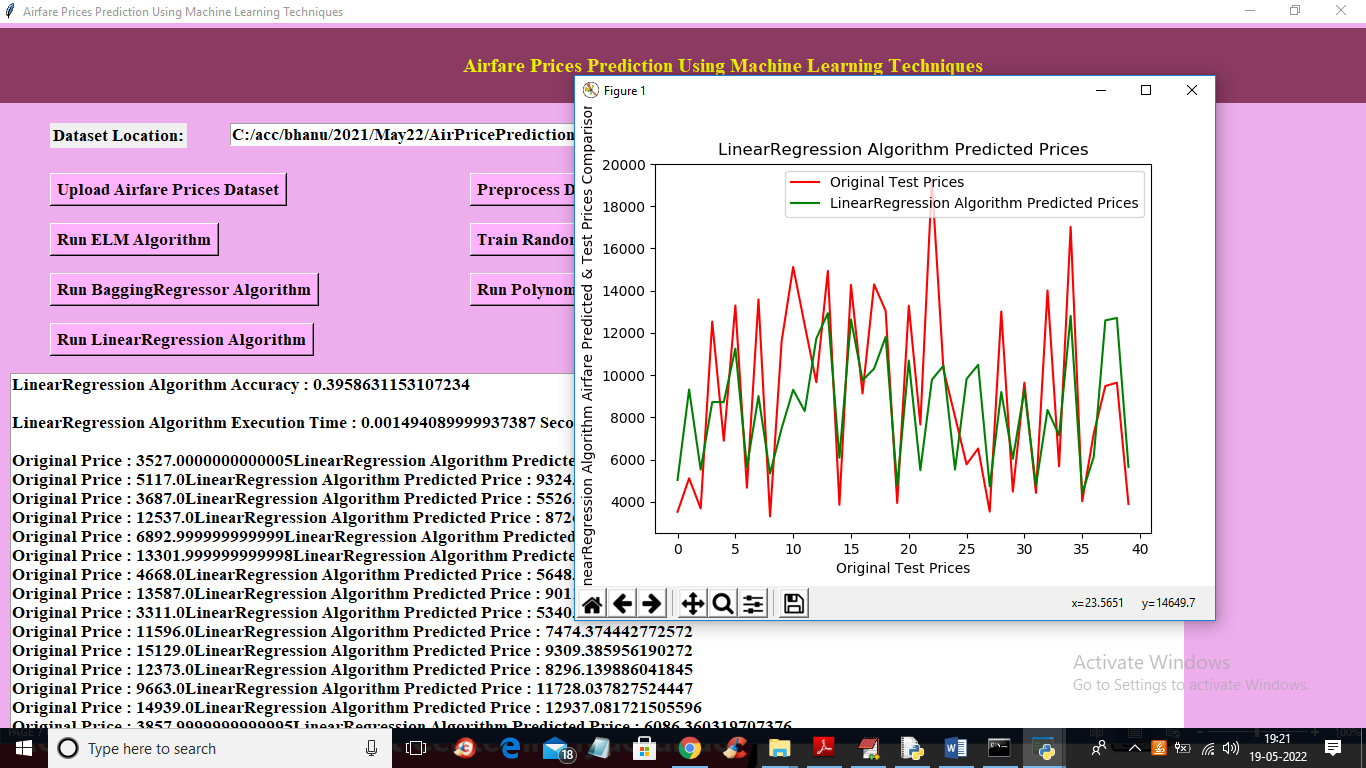
In above screen with Bagging Regression we got 0.92% accuracy and now click on ‘Run Polynomial SVM’ button to get below output



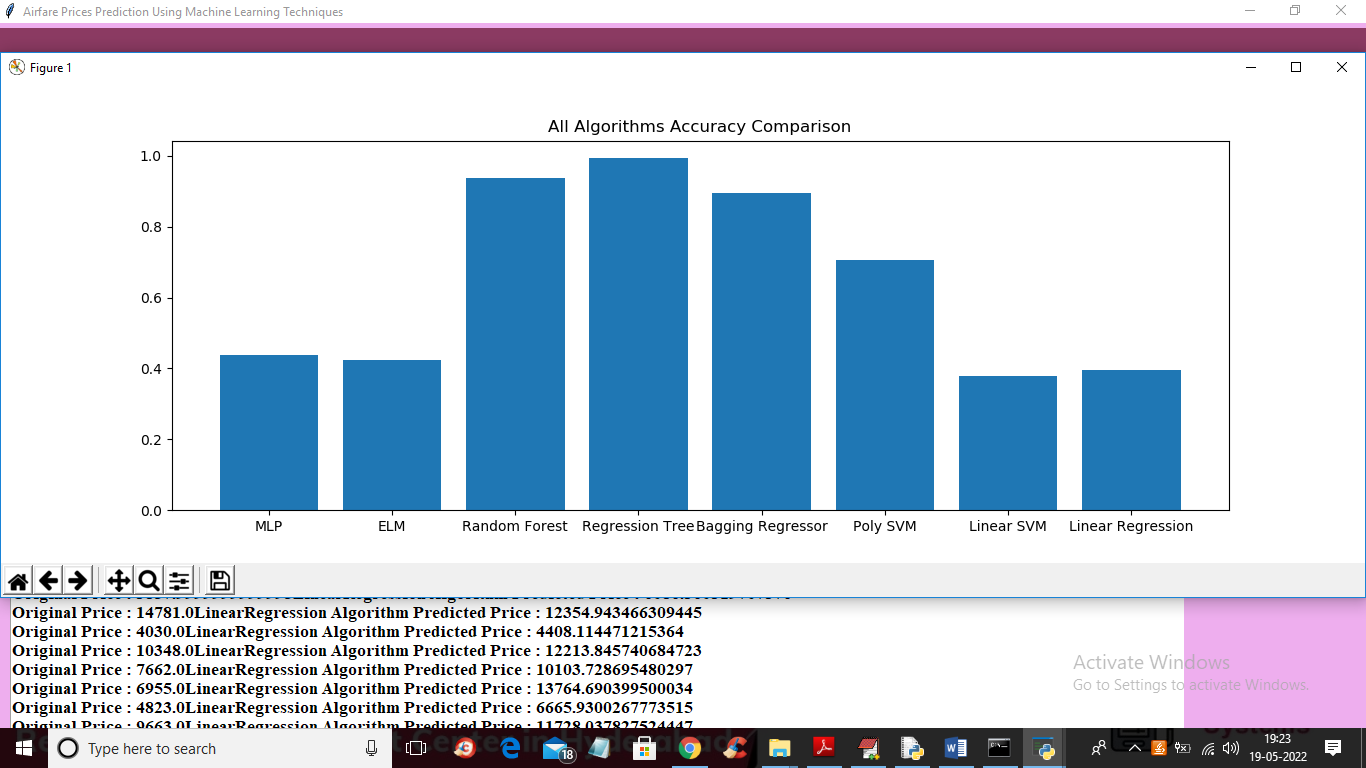
In above screen with Polynomial SVM’ we got 0.70% accuracy and now click on ‘Run Linear SVM Algorithm’ button to get below output



In above screen with Linear SVM we got 0.37% accuracy and now click on ‘Run Linear Regression Algorithm’ button to get below output



In above screen with Linear Regression we got 0.39% accuracy and now close above graph and then click on ‘Accuracy Comparison Graph’ button to get below graph



In above graph x-axis represents algorithm names and y-axis represents accuracy of those algorithms and in all algorithms Regression Tree got high accuracy