Novel Method For Car Price Prediction With Machine Learning Techniques.

Team Members:

- Yachamaneni Yashaswini
- Vishnuvardhan Rama

Texas A and M University – Corpus Christi .
 Department of computing science.
Instructor: Carlos Rubio-Medrano

Contents

INTRODUCTION

- Nowadays usage of own transport has increased as everyone is thinking of buying their own vehicle that may be a two-wheeler or four-wheeler instead of preferring public transport.

- Getting a used car is very easy as there are a lot of used car showrooms and some of the companies are selling them online.

- The goal is to develop a webpage to predict the price of the car using manual parameters and evaluating the accuracy of the proposed work.

PROPOSED WORK

- Here, Website is designed which helps to predict the car price using features such as year of manufactured, company, model, distance covered, size, type of fuel, mileage, horsepower and image.

- Our website provides the details of used-cars with a fixed price as-well-as description of each brand including its advantages.

- Machine learning algorithm and techniques were used to evaluate, how much accurate our proposed model is ?

FEATURES OF OUR MODEL
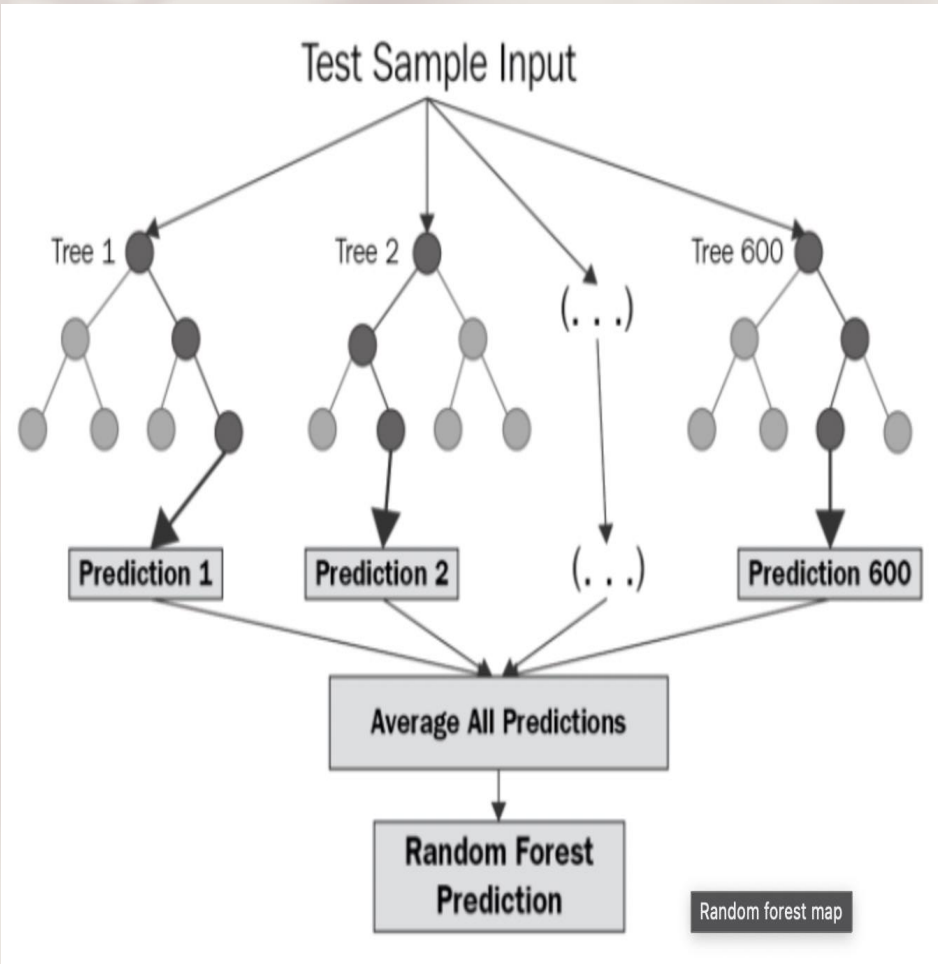
Price Prediction

Chat/Help section

Shop and Buy

Details of Brands

ALGORITHM

Random Forest Regressor:

❑ In automotive industry, machine learning is most often associated with product innovations. More than 78% of automotive companies invest in Machine Learning to regularly improve user experience.

❑ Linear Models perform poorly on data containing outliers. Linear models fail to perform on non-linear datasets. In such case, non-linear regression algorithms like random forest perform better in fitting non-linear data.
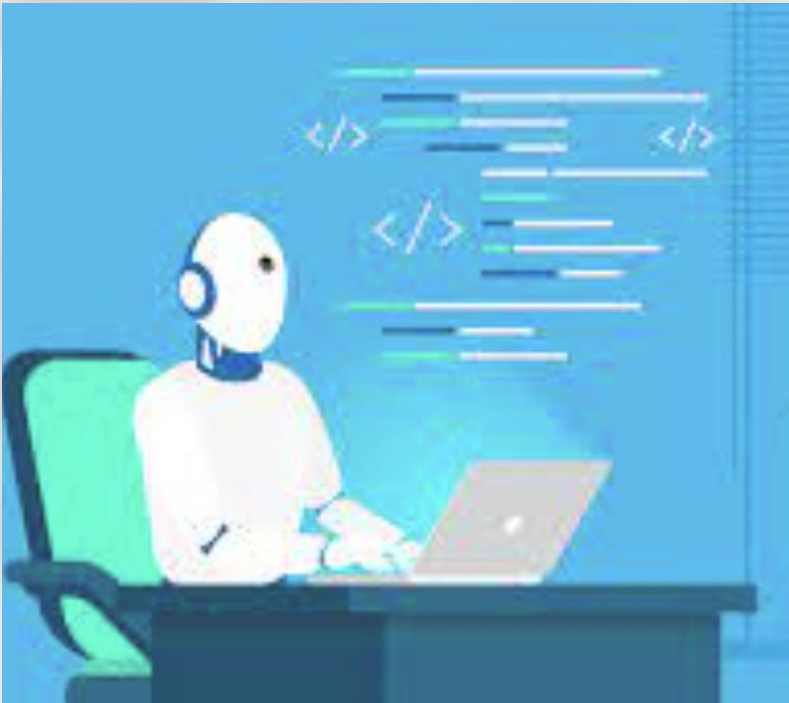
TECHNOLOGIES USED

- Python
- HTML
- CSS
- Java Script
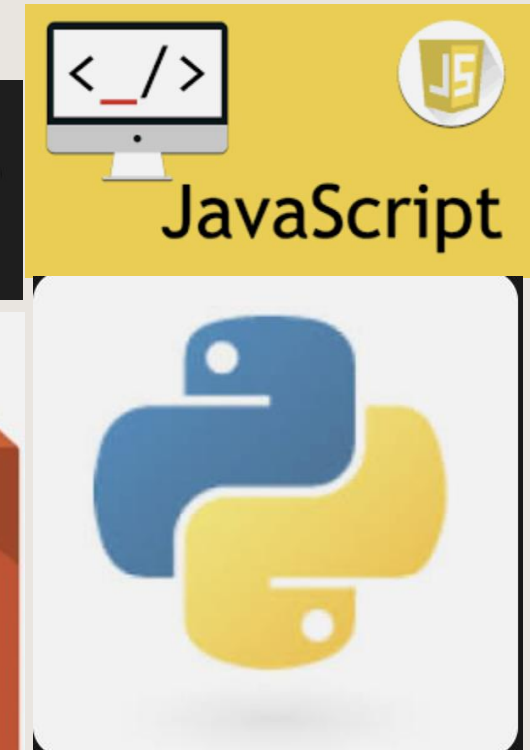
IMPLEMENTATION DETAILS

- Architecture is user friendly with input fields for the user to enter values or car properties.

- The approach was implemented by using sklearn, NumPy and pandas and trained our model based on car properties.

- By using sklearn libraries we trained our model and NumPy is used for working with arrays. It is also used for working with domain of linear algebra, Fourier transform and matrices.

- In extra-tree regressor algorithm, the class implements a meta-estimator that uses averaging to increase predictive accuracy and reduce overfitting.

EXPERIMENTS

➢ We first trained our model with dummy values and later trained the model with manual values given by the user.

➢ Initially data Analysis is performed to eliminate missing values and noisy data.

➢ Correlation Matrix: Visualizing the correlation is an effective way of determining dependencies.

➢ To identify distribution of single variable and relationship between two variables we performed 'Pair Plot'.

➢ Mean Square Error and Mean Absolute Error are used for performance evaluation.

➢ Also, plotted confusion matric with both dimensions (test_class and prediction_class).

RESULTS

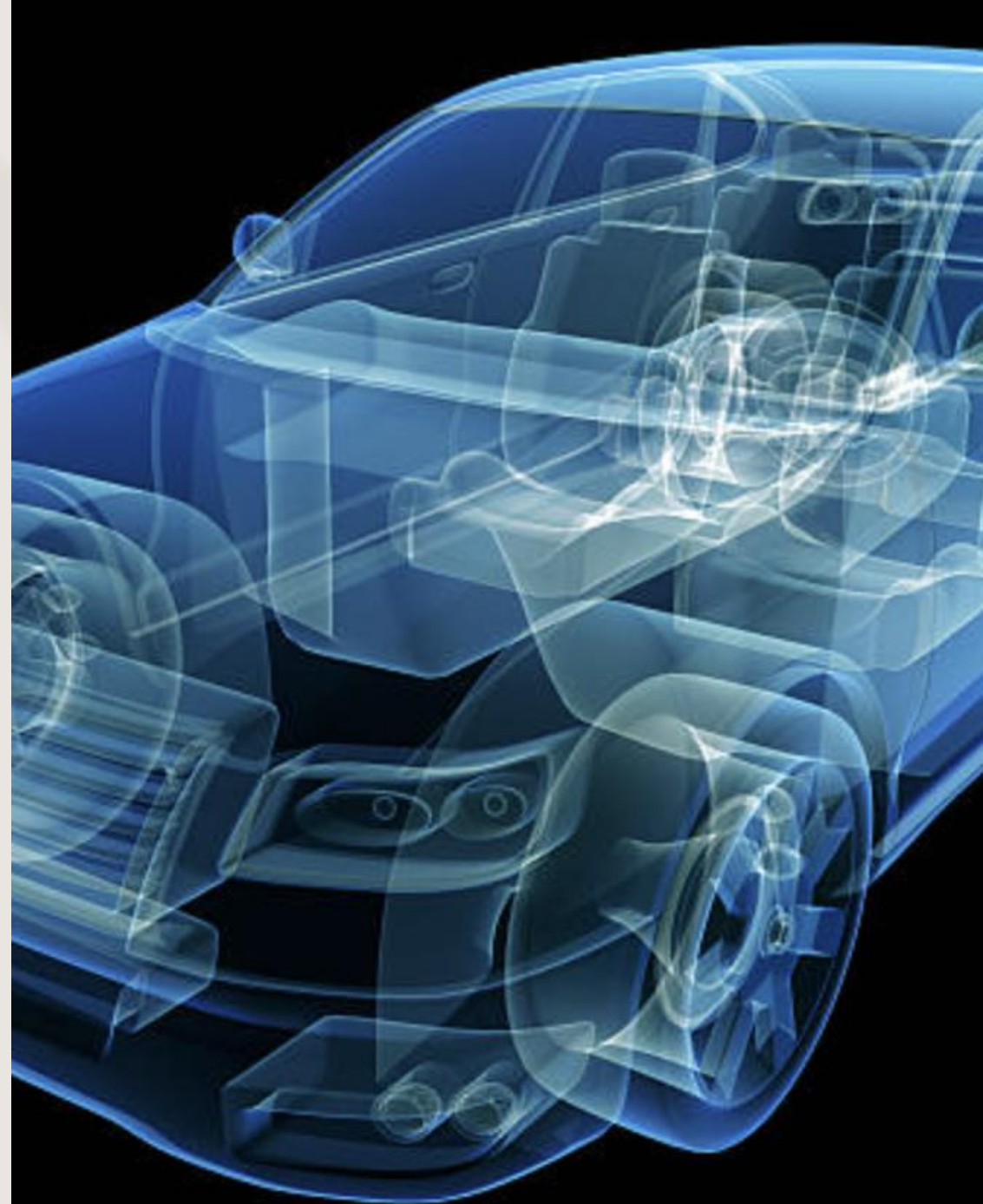Performance of Random Forest
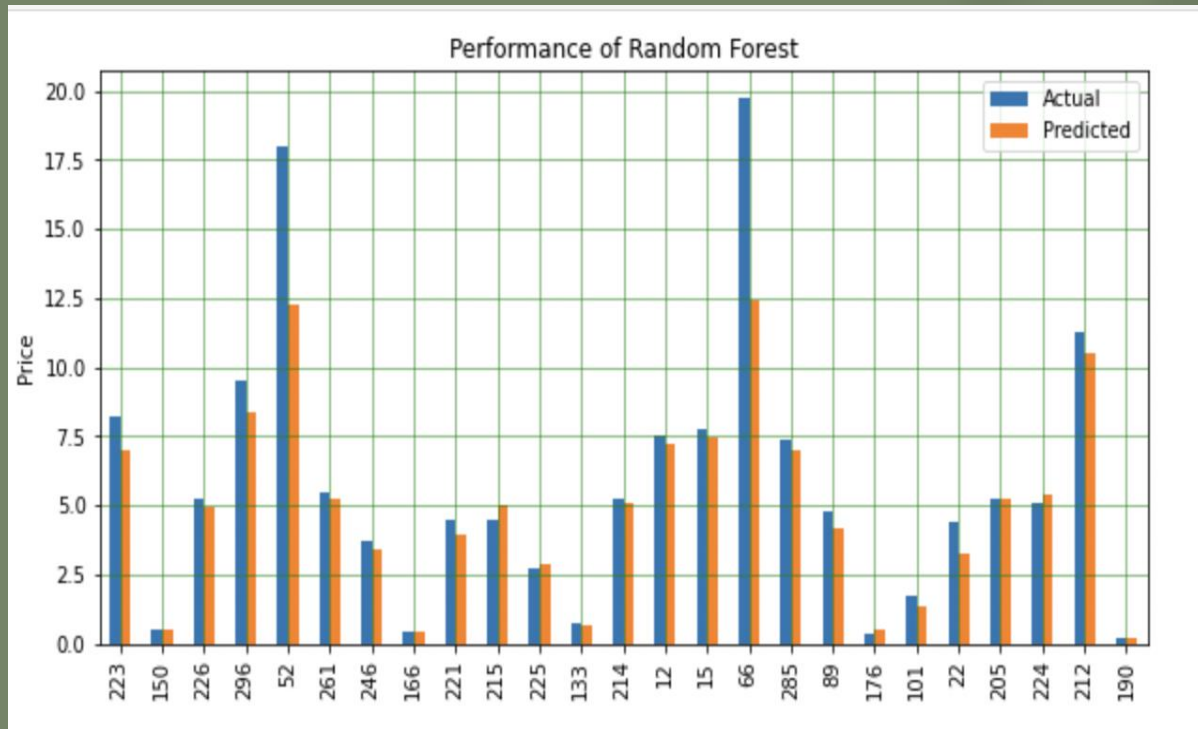
```
In [65]:   from sklearn.metrics import accuracy_score
           s = accuracy_score(y_test_classes,y_pred_classes)
           print("accuracy of the model : {}" . format(s))

           accuracy of the model : 0.9344262295081968
```

Final, accuracy of our proposed model was 0.93 or 93%.

FUTURE WORK

- The designers are trying to add the features like video and multiple images in their image in the long run for improving the quality of screening.

- The designed site is having some pros and cons. Pros like the user can go through different car models in a road and by input in the image they can track down the car price as well. There are some cons in detection of the image. These are the sections the designers should have to focus on the long run.

Demo

Thank you

ANY QUESTIONS ?