CSE-O28 Project Report: Car Resale Value Prediction

Team Name: Fab 5

Team Members: J.Manisahithi – 18UK1A05F4

S.Pranathi – 18UK1A05H4

T.Manisha – 18UK1A05H6

J.Pausha – 18UK1A05D7

T.Harathi – 19UK5A0518

**INTRODUCTION**

Overview

The price of a new car in the industry is fixed by the manufacturer with some additional costs incurred by the Government in the form of taxes. So, customers buying a new car can be assured of the money they invest to be worthy. But, due to the increased prices of new cars and the financial incapability of the customers to buy them, Used Car sales are on a global increase. Considering the main factors which would affect the resale value of a vehicle a regression model is to be built that would give the nearest resale value of the vehicle. The focus of this project is developing machine learning models that can accurately predict the price of a used car based on its features. Here we are using the random forest regressor to predict the selling prices since this is a regression problem and it uses multiple decision trees.

Purpose

The purpose of this project is to predict the price of used cars by using a machine learning model. We are using random forest regressor because it uses multiple decision trees. We also predict the accuracy. Through this we can achieve best price for selling the used cars.

**LITERATURE SURVEY**

Existing problem

The approach to predict the price of used cars is Random forest regressor. Prediction based on the trees is more accurate because it takes into account many predictions. These algorithms are more stable because any changes in dataset can impact one tree but not the forest of trees.

Proposed solution

I propose random forest regression model because it is a powerful and versatile supervised machine learning algorithm that grows and combines multiple decision trees to create a forest. It can be used for both classification and regression problems.

**THEORITICAL ANALYSIS**

Hardware/software requirements

Hardware Requirements

* Operating System – Windows 7,8,10
* Processor – dual core 2.4 GHz (i7 series intel processor)
* RAM – 4GB

Software Requirements

* Google Colaboratory
* Anaconda
* Spyder
* Google
* MS Word

**EXPERIMENTAL INVESTIGATIONS**

Collected dataset from kaggle site. Did analysis on how to choose the best regression model, how to build it, how to evaluate it, how to test and train the data, how to deploy the model. After choosing the regression model, studied about it.

**FLOWCHART**

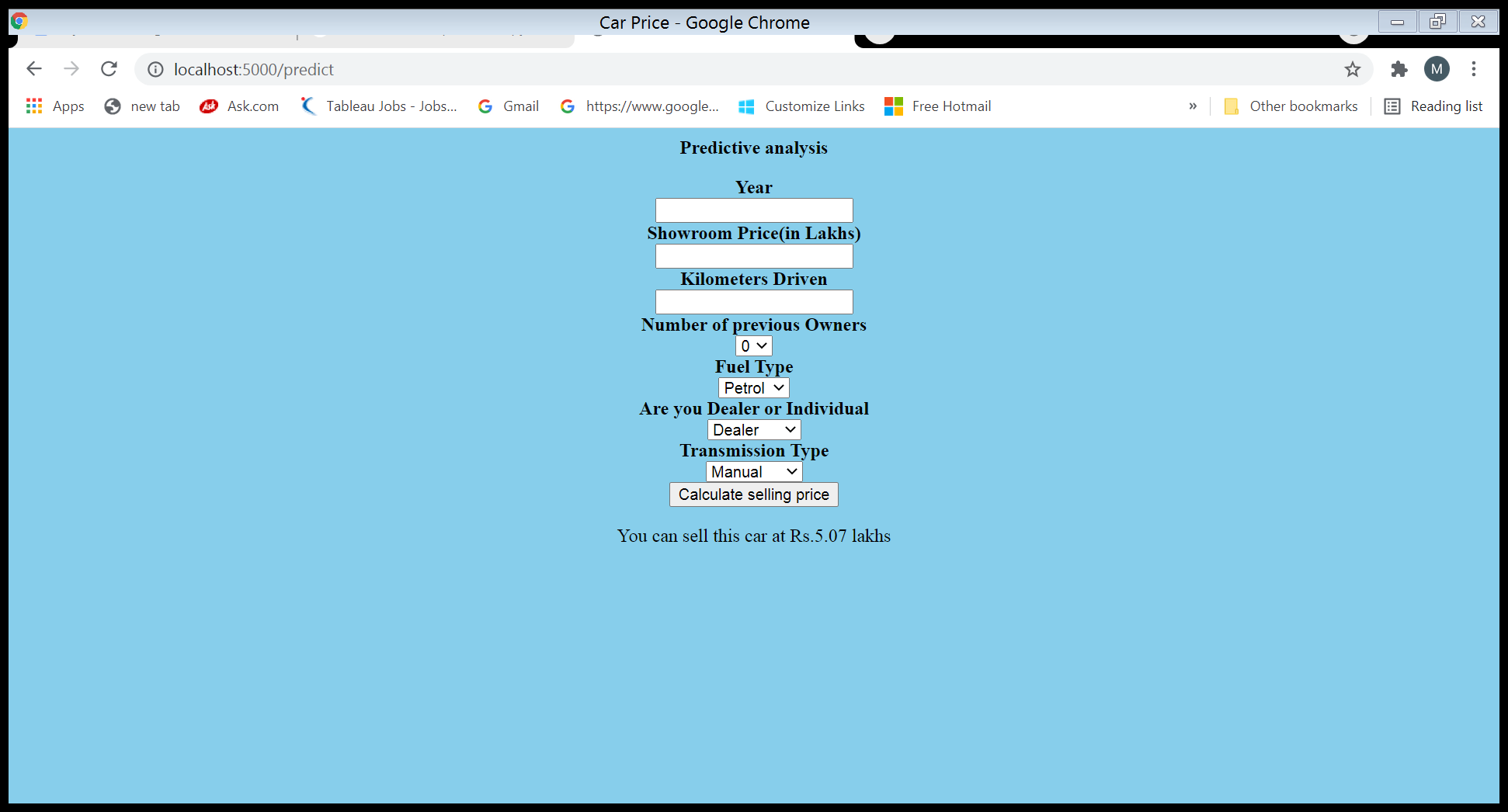
Collect the dataset

Data preprocesing

Choose the best model

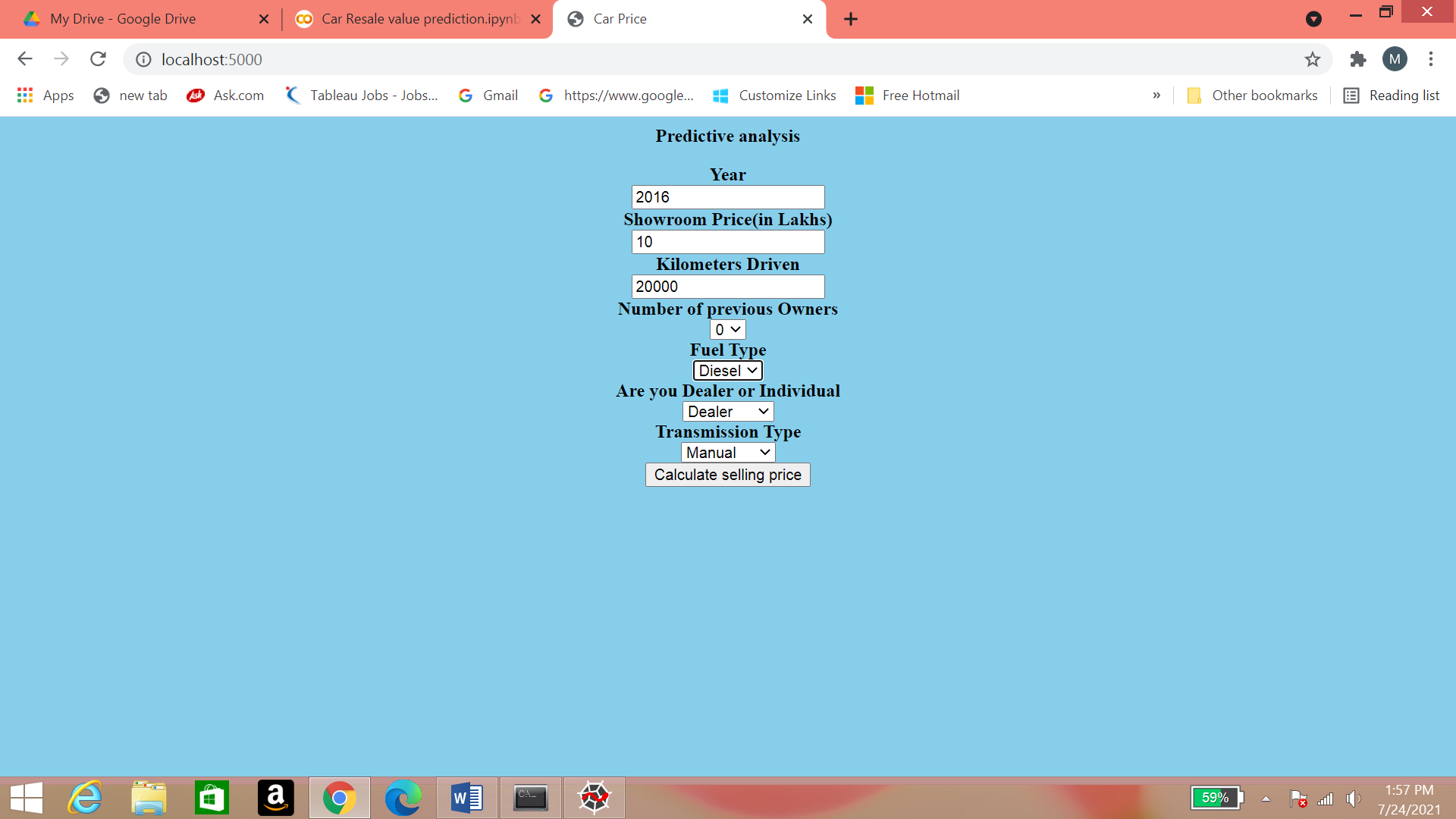
Evaluate the data

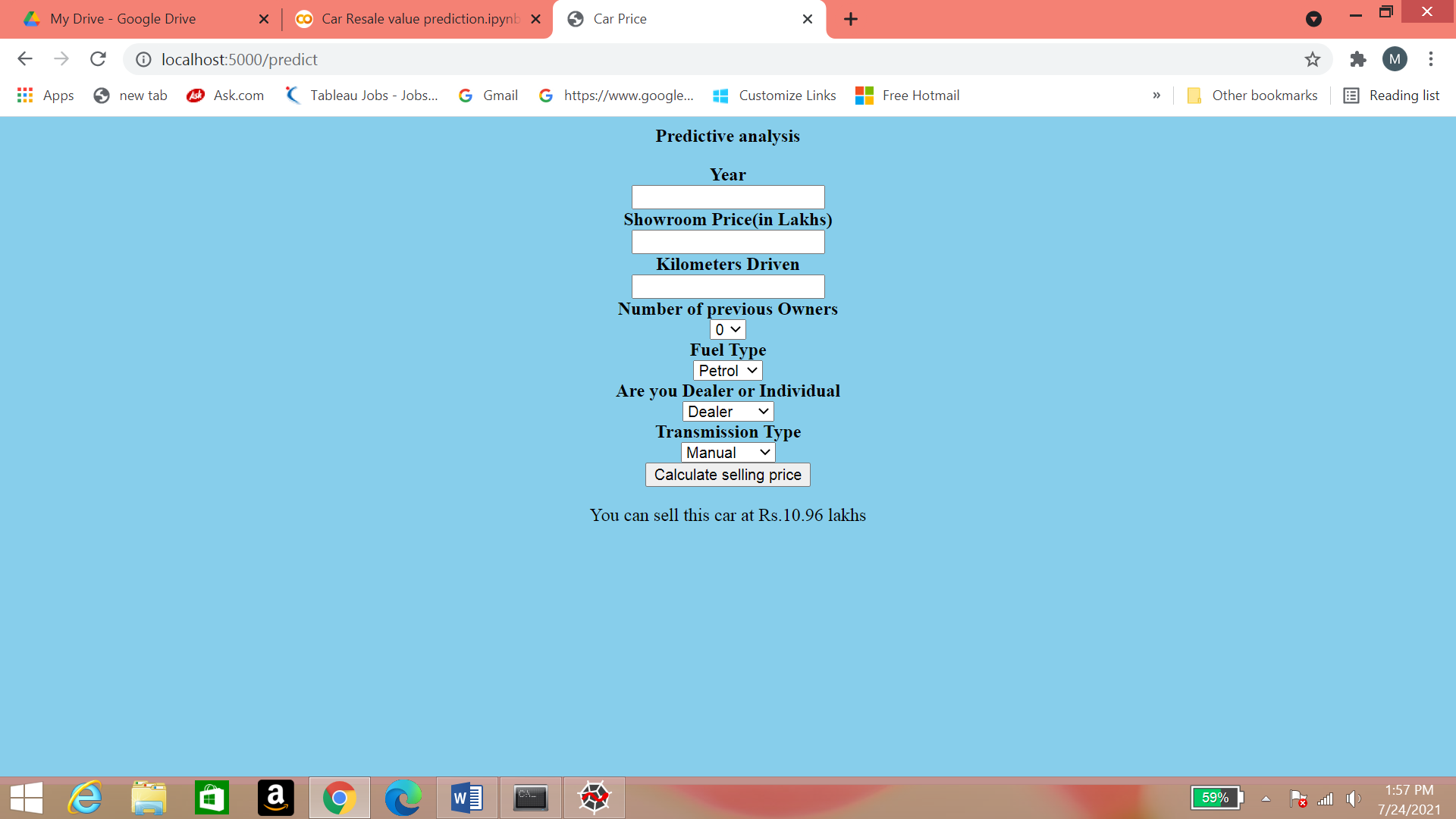
Save the model

Deploy the model****

**RESULTS**

After writing the code in flask, we need to create the html page also. Next we have run <http://localhost:5000> in chrome, so that we get the webpage we created. Insert the values. Calculate the selling price. We get the output.

****

****

**ADVANTAGES AND DISADVANTAGES OF PROPOSED SOLUTION**

ADVANTAGES

* Random Forest Algorithm is less prone to overfitting than decision tree and other algorithms
* Random Forest algorithm outputs the importance of features which is a very useful
* Random Forest can automatically handle missing values
* Random Forest can be used to solveboth classification as well as regression problems

DISADVANTAGES

* Random Forest algorithm may change considerably by a small change in the data
* Random Forest algorithm computations may go far more complex compared to other algorithms

**APPLICATIONS**

**Banking Industry**

* Credit card fraud detection
* Customer segmentation
* Predicting loan defaults

**Health care and Medicine**

* Cardiovascular disease prediction
* Diabetes prediction
* Breast cancer prediction

**STOCK MARKET**

* Stock market prediction
* Bitcoin price prediction

**CONCLUSION**

The increased prices of new cars and the financial incapability of the customers to buy them, Used Car sales are on a global increase. Therefore, there is an urgent need for a Used Car Price Prediction system which effectively determines the worthiness of the car using a variety of features. The proposed system will help to determine the accurate price of used car price prediction.

**FUTURE SCOPE**

In future this machine learning model may bind with various website which can provide real time data for price prediction. Also we may add large historical data of car price which can help to improve accuracy of the machine learning model. We can build an android app as user interface for interacting with user. For better performance, we plan to judiciously design deep learning network structures, use adaptive learning rates and train on clusters of data rather than the whole dataset.

**BIBILOGRAPHY**

1. <https://www.kaggle.com/nehalbirla/vehicle-dataset-from-cardekho?select=car+data.csv>
2. <https://www.analyticsvidhya.com/blog/2021/05/build-and-deploy-a-car-price-prediction-system>

**APPENDIX**

**SOURCE CODE**

GitHub link: <https://github.com/manisahithi1312/Car-Resale-Value-Prediction>