

Group 04

Clement Justus S

Mohan Kumar R

Mohammed Saqlain

Prajwal Borse

## **Car Price Prediction**

### **Background of the project:**

The main concept behind this car price prediction, in which each car's rate is sold based on 'Symboling', 'Car Name', 'fuel type', 'aspiration', 'door number', 'car body', 'drive wheel', 'engine location', 'wheelbase', 'car length', 'car width', 'car height', 'curb weight', 'engine type', 'cylinder number', 'engine size', 'fuel system', 'bore ratio', 'stroke', 'compression ratio', 'horse power', 'peak rpm', 'city mpg', 'highway mpg', 'price' based on all these categories the cars are sold to customers.

We are required to model the price of cars with the available independent variables. It will be used by the management to understand how exactly the prices vary with the independent variables. They can accordingly manipulate the design of the cars, the business strategy etc. to meet certain price levels. Further, the model will be a good way for management to understand the pricing dynamics of a new market.

Using so many other technologies in project we can predict the particular car can be sold for this particular price.

So, to be clear, this deployed application will provide you with the approximate selling price for your car based on the fuel type, years of service, showroom price, the number of previous owners, kilometres driven, if dealer/individual, and finally if the transmission type is manual/automatic. Deciding whether a used car is worth the posted price when you see listings online can be difficult. Several factors, including mileage, make, model, year, etc. can influence the actual worth of a car. From the perspective of a seller, it is also a dilemma to price a car. Based on existing data,

### **Goal of the Project:**

The aim is to use machine learning algorithms to develop models for predicting car prices

The focus of this project is developing machine learning models that can accurately predict the price of a car based on its features, in order to make informed purchases. We implement and evaluate various learning methods on a dataset consisting of the sale prices of different makes and models across cities.

You as a Data scientist are required to apply some data science techniques for the price of cars with the available independent variables. That should help the management to understand how exactly the prices vary with the independent variables. They can accordingly manipulate the design of the cars, the business strategy etc. to meet certain price levels.

## **Objectives & Motivation**

- To develop a efficient and effective model which predicts the price of a car according to user's inputs.
- To achieve good accuracy about the used car details, it's features and pricing.
- To develop a User Interface( UI ) which is user-friendly and takes input from the user and predicts the price.

Deciding whether a car is worth the posted price when you see listings online can be difficult. Several factors, including mileage, make, model, year, etc. can influence the actual worth of a car. From the perspective of a seller, it is also a dilemma to price of a car Based on existing data, the aim is to use machine learning algorithms to develop models for predicting car prices.

## **Problem Statement:**

A Chinese automobile company Geely Auto aspires to enter the US market by setting up their manufacturing unit there and producing cars locally to give competition to their US and European counterparts.

They have contracted an automobile consulting company to understand the factors on which the pricing of cars depends. Specifically, they want to understand the factors affecting the pricing of cars in the American market, since those may be very different from the Chinese market. The company wants to know:

Which variables are significant in predicting the price of a car

How well those variables describe the price of a car

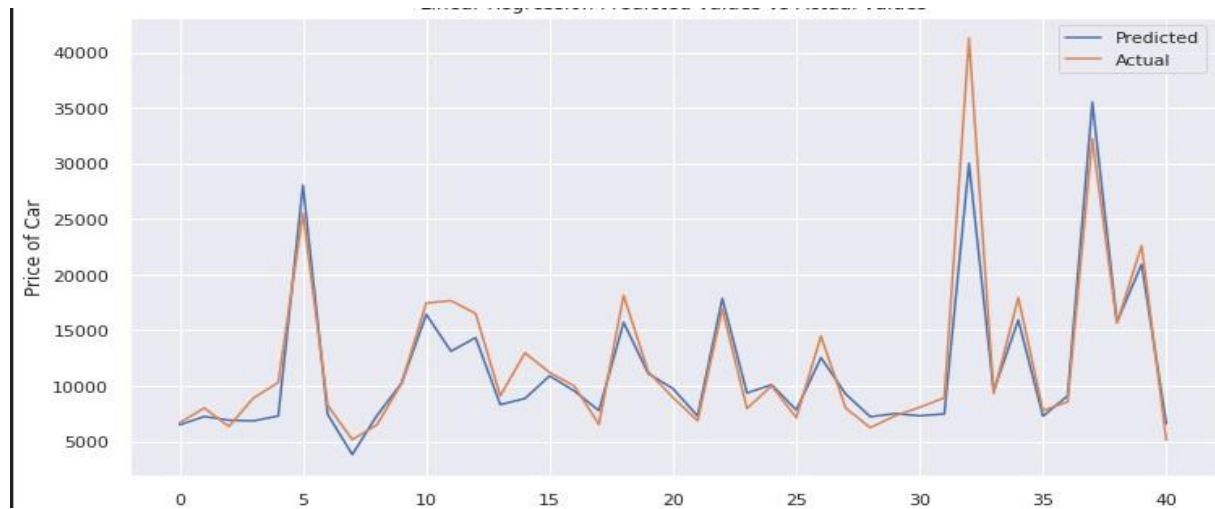
Based on various market surveys, the consulting firm has gathered a large data set of different types of cars across the America market.

## **Solution/Proposed system:**

The process starts by collecting the dataset. The next step is to do Data Preprocessing which includes Data cleaning, Data reduction, Data Transformation. Then, using various machine learning algorithms we will predict the price. The algorithms involve Linear Regression, Ridge Regression and Lasso Regression. The best model which predicts the most accurate price is selected. After selection of the best model the predicted price is displayed to the user according to user's inputs. User can give input through website to for used car price prediction to machine learning model.

## Linear Regression

Linear Regression attempt to model the relationship between two variables by fitting a line equation to observed data. The other is considered to be dependent variable.

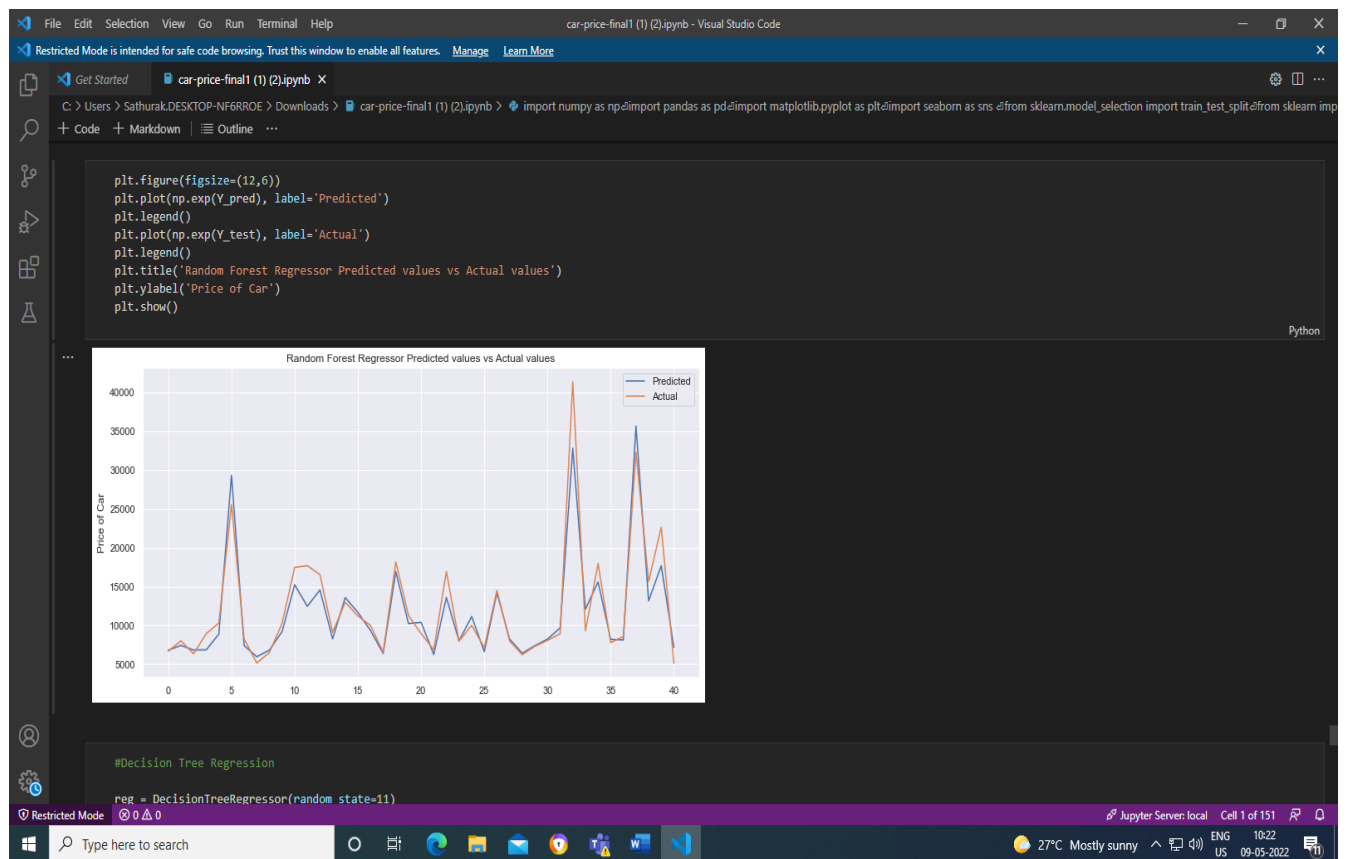


Linear regression is useful for finding relationship between multiple continuous variables

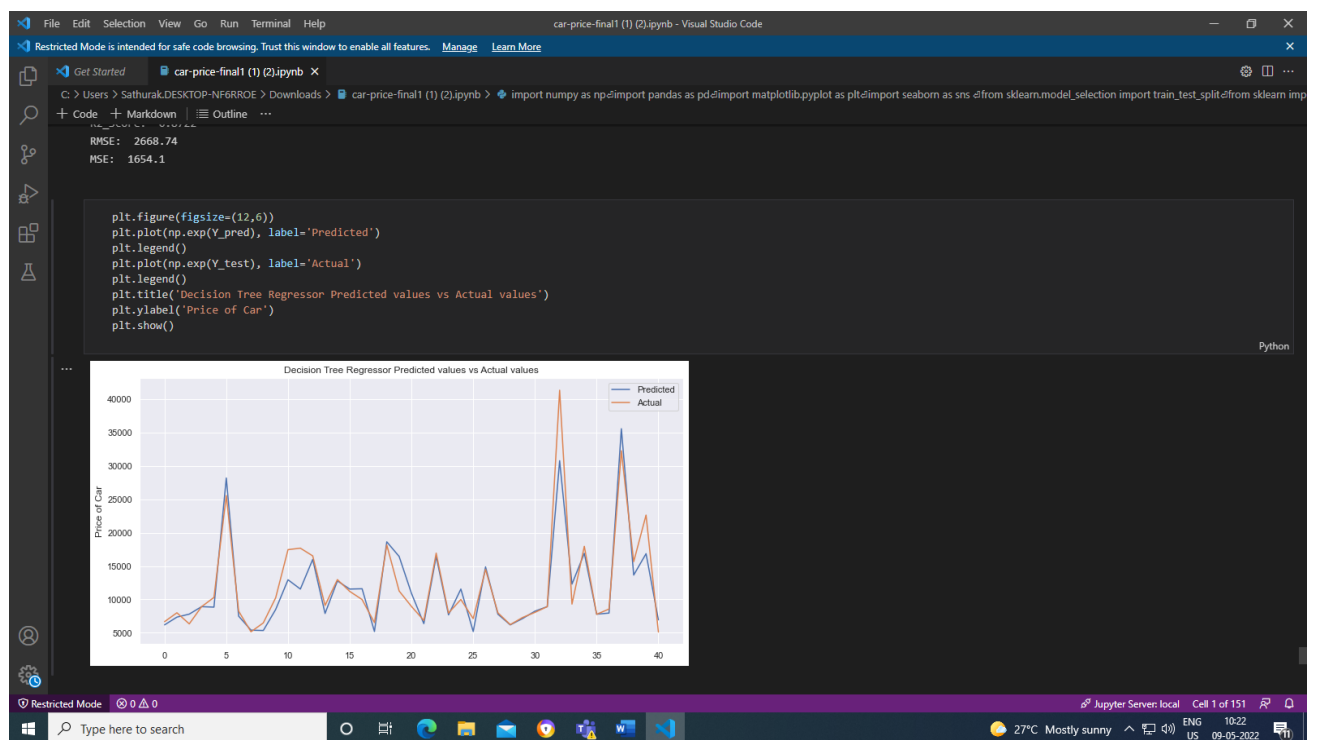
There are multiple independent variables and single dependent variable =  $m_1X_1 + m_2X_2 + \dots + m_nX_n + b$ ,  $m_1, m_2, m_3, \dots$  -> slope  $b$  -> y intercept

$X_1, X_2, X_3, \dots$  -> independent variables  $y$  -> dependent variables.

## Random forest Regression



## Decision Tree Regression



## **FUTURE SCOPE**

In future this machine learning model may bind with various website which can provide real time data for priceprediction. 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.

## **Requirements**

Hardware requirements:

- Operating system- Windows 7,8,10
- Processor- dual core 2.4 GHz (i5 or i7 series Intel processor or equivalent AMD)
- RAM-4GB

Software Requirements:

- Python Pycharm PIP 2.7
- Jupyter Notebook Chrome



