Name: Prajakta Ramesh chavan
Company name: Oasis infobyte
Date:24/02/2024

Car price prediction with machine learning:

Step 01: Download the dataset : kaggle.com

Step02: Importing the library as jupyter notebook

Overview:

To be able to predict used cars market value can help both buyers and sellers. There are lots of individuals who are interested in the used car market at some points in their life because they wanted to sell their car or buy a used car. In this process, it's a big corner to pay too much or sell less then it's market value.

In this Project, we are going to predict the Price of Used Cars using various features like Present_Price, Selling_Price, Kms_Driven, Fuel_Type, Year etc. The data used in this project was downloaded from Kaggle.

Data Cleaning:

We've seen that there are no missing values in the dataset. We've also seen that variables are in the correct format, except "symboling", which should rather be a categorical variable (so that dummy variable are created for the categories).We have also done data preprocessing on The variable "CarName" and created a new variable called as "car_company".

Data Preparation:

Let's now prepare the data for model building.

Split the data into X and y.

Conclusion:

I have selected the required used car prices dataset with needed features and parameters from Kaggle. Kaggle is an opensource Machine learning and data science platform which offers data and notebooks for data scientists and data analysts. The required data is cleaned and pre-processed used machine learning techniques before applying any algorithm for predicting the price. Then after pre-processing and cleaning the data first we need apply train test split to keep the data into two parts for training and validation using train and test data respectively. Then we must apply a simple linear regression model and predict the output and test its test and train accuracy with the help of roc auc score then again, we need to train and test it with multiple linear regression model and validate its accuracies. Then we need to use clustering methods and logistic regression methods and knn methods for predicting the output of car price. Also, we can use random forests and decision tree algorithms. At last, we need to compare all the accuracies of all the machine learning algorithms and choose the best algorithms for the prediction.