USED CARS PRICE PREDICTION

Certification Project

edureka!

edureka!

Domain – Price Prediction

Overview
Congratulations!!!! for making it so far. This is the Certification Project for Data with Python Certification Course here we will use all the concepts learned so far in this course.

Context
Riding the digital wave, India's used car market is set to grow at a compounded annual growth of 11% and is likely to touch sales of up to 8.3 million units by FY26 as more

people have been opting for pre-owned cars for personal mobility in the pandemic amid the ongoing supply shortages for manufacturing new cars.

The used car market in the country is expected to reach over 70 lakh vehicles by

2025-26, up from 38 lakh in 2020-21 as the Covid-19 pandemic, digitalization, changing demographics and aspirations, first-time buyers, and availability of financing options are acting as growth drivers, according to a report by OLX Autos and rating agency Crisil.

"MyCars" is a new-age startup laying foundations in the setting up a car resell domain and they are setting up a team of ML experts to make predictive models determine

the price of second-hand cars to optimize their revenue, you have joined as a new Data Scientist and your role is to create a model to determine the selling price of a used car.

Objective:
   ● Provide the best-performing model to determine the price of the used car.
   ● Providing the most important features which determine the price.

Data Description
The data provided consists of the following Data Dictionary

   ● Id: Unique ID assigned to a specific car.
   ● year: Manufacture year of the car.
   ● brand: Brand of the car.
   ● full_model_name: Model name includes other details such as engine capacity, transmission, etc., basically a detailed model name.
   ● model_name: Just the model name of the car.
   ● price: Sell price of the 2nd ownership car.
   ● distance_travelled(km): Distance traveled by car.
   ● fuel_type : Fuel engine type .
   ● city: City where the car is registered.
   ● car_age: Age of the car.

<u>Steps and Tasks:</u>

- Import libraries and load dataset
- Exploratory Data Analysis:
  - Univariate Analysis to understand the distribution of features.
  - Multivariate analysis to determine correlations.
- Layout regression experiment space.
- Publish the performance of all the models on the same hold-out/ test dataset.
- List out the most important features determining the price of used cars.