Palestine Car price prediction

Student

Naser Shanti

Adham Soos

Supervisor

Dr. Adnan Salman

* Introduction: with a lot of thinking and analysis we decide to make a project that can predict the price of the car in Palestine , since its no similar program that do the Target , we choose a “شو ما بدك من فلسطين” as a source of our project , and we talk the features that provides as a features of our model , We obtained 200,000 car samples Some of the cars were used, and some were new.
* important Features: The most important features in determining the price of cars are found to be the engine power, the engine fuel type, and the distance that the car was made, which is sensical since they determine how well a car runs. Model types, the make, and the year come after and the color these have more to do with the brand and superficial appeal of the car, but are still deciding factors nonetheless.
* Team: Naser Shanti + Adham Soos
* plan to Complete the Task:

1. get the data from a Website, by Scrabing Technique used Selenium library that was provide by Python language.
2. Prepare and clean the missing data (noise data) ----> “hard part”
3. Split the hole data into 5 part , this steps is done to use the Online learning Technique since our Devices is not ready to deal with large number of this data
4. We use jupyter notebook as an environment of our work
5. Load the data by pandas library , Data shell be in (CSV) file
6. In this steps we want to split the data into Testing and training sets using the stratified Technique
7. Before applying regression models, lets look at the features and also relationship (Correlations) with each other by visually.
8. From the previous steps we can Experimenting with Attribute combinations
9. Check if there’s a Missing data in the Training sets and deal with it
10. We are going to change the categorical to numerical values to make it proper for regression models by One hot encoder and we will Prepare the data for a Machine learning Algorithm
11. Applying Regression Models , and Optimization Algorthim

* Output: The output for this Model is a Price for a Car ( prediction price)
* Techniques that used:

1. Selenium for Scrabing
2. Linear Regression
3. Grediant Descent as an Optimization Tecqinqe