TRUECAR

OCT 2021

Predicting Car Prices Using Linear Regression

Project Proposal

Proposed by:
Sultanah Aldossari
Rawan Alharbi

PROJECT PROPOSAL

PROPLEM STATEMENT

The purchase of a car can take a considerable amount of time. Deciding whether a car is worth the posted price when you see listings online can be difficult. It is possible to determine the actual worth of a car based on many factors, including the mileage, make, model, and year. The aim is to develop models based on our data to predict car prices using machine learning algorithms. We would like to know.

- Which variables are significant in predicting a car's price?
- How well those variables describe the price of a car

DATASET

A primary requirement for this project was to obtain data through web scraping. We used Selenium to obtain our data. Our data consist of 205 observations, 26 features, no null values. The column "**Price**" is the target variable and rest of the columns are independent variables. The independent variables are divided into Categorical and Numerical variables.

Numerical variables: wheelbase, car length, car width, car height, curb weight, engine size, bore ratio, stroke, compression ratio, horsepower, peak rpm, city mpg, highway mpg.

Categorical variables: symboling, fuel type, aspiration, door number, car body, drive wheel, engine location , engine type, cylinder number, fuel system, car name.

PROJECT PROPOSAL

TOOLS

To Achieve our goal. we will analyze and explore our data using **Python** and **Jupyter Notebook**.

LIBRARIES

- Sklearn
- NumPy
- Pandas
- Seaborn
- MatplotLib
- Selenium
- SciPy
- StatsModel.