Name: Faisal Mirza   faisal.a.mirza11@gmail.com   GitHub: faisalmirza11

Project Proposal Template

Background:

Carsforsale is a website that sells all kinds of new and used cars, it also allows owners of cars to post and sell their cars through the website, and due to the increased sales and demand of SUV's and crossovers over sedans and other types of vehicles in the recent years, Carsforsale wants to know if a used SUV posted for sale on their website has a fair price or not. So by using the current used SUV's prices on the website (assuming all have a fair price) we will predict if a new post has a fair price or not, which will allow Carsforsale to clean their search results for users of the website.

Question/need:

Framing question: Does a new post of a used SUV for sale have a fair price or not.?

Who benefits: Carsforsale, or a website for selling used cars, and also people who are in the market to buy used cars.

Data Description:

The dataset will be collected/scraped from Carsforsale.com, and it will contain all kinds of used SUVs posted on the website for sale. And some of this data will include:

- The car's brand name, model, year of production.
- The mileage of the car.
- The price of the car.
- The car's engine, drivetrain.

Scope:

- We will attempt to collect data of roughly 1000 cars (rows)
- Each row will have about 18 columns

Tools:

Technologies: Python, Jupiter notebook

Libraries: BeautifulSoup, Selenium, Sklearn, statsmodel, Numpy, Pandas, Matplotlib, Seaborn.