**Regression model – prediction of use car price in Saudi Arabia**

**Abstract**

Our project aimed to applied regression model to predict price of cars in Saudi country that change according to certain attributes and companies, this project may help people to have a clear vision about car prices and what the main factors that has highly influence. I used web scraping to obtain data from syarah web site, I applied multi leaner regression such as simple leaner regression ,Lasso and polynomial regression , R^2 in lasso and simple leaner regression are worst which represent under fitting prediction for them by building polynomial regression we get a good prediction model

**Design**

The main question of our work is how much the used car cost in Saudi Arabia After data scraped, I cleaned data from missing values and outlier, handling some feature, visualized distribution, and correlation of features and detect multicolleanity by VIP, we notice there are no leaner correlation between variables together and target as well. for that reason, we must build polynomial regression

**Data**

I obtain the dataset from sayarah web site which is reliable site in Saudi Arabia Columns: 6 features (Target value= price)

* + Data type: numerical and categorical attributes [year, engine size, gear type, option that represent if the car full or standard, fuel type , color ,name ,mileage ,Rows: more than 8000 , Data size 1000x8

**Molding**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Leaner reg** | **R^2** | **MAE** | **MSE** | **RMSR** |
| Training | 0.4425 | 0.1492 | 0.0393 | 0.1983 |
| Testing | -3.7696 | 96325 | 1148 | 10714 |
| **Lasso** | **R^2** | **MAE** | **MSE** | RMSR |
| Training | 0.0 | 0.2107 | 7518 | 8670 |
| Testing | -2.1235 | 71493 | 7518 | 8670 |
| **Polynomial reg (D=6)** | **R^2** | **MAE** | **MSE** | **RMSR** |
| Training | 0.7347 | 0.1492 | 0.0393 | 0.1983 |
| Testing | -3.7696 | 96325 | 1148 | 10714 |
| **Lasso** | R^2 | MAE | MSE | RMSR |

**Tools**

Main tool: python 3.9, seaborn, matplotlib, pandas, NumPy

Web scraping: beautiful soup, html5lib and requests