Name: Sai Vamsi Ogety
E-mail: Ogetyvamsi@gmail.com

A.S.U, Tempe, AZ,USA
Date: 03-08-2019

Edureka Techathon
Car Price Prediction Challenge

Intro: Buying a used car can be tough. The question of whether we are paying the right price? or are we paying more ? troubles lot of people in deciding to buy. Data science can tackle this issue by taking the leverage of the historical data.

Goal: To develop a model which will predict the car_price based on the input data.

Workflow: The Workflow i followed is pretty much how everyone do a data science project.

1. Understand the goal or objective first
2. Import the dataset
3. Data cleaning and Exploration
4. Baseline Model Building
5. Improving the Model or Build another model
6. Concluding the Results.

Libraries used :

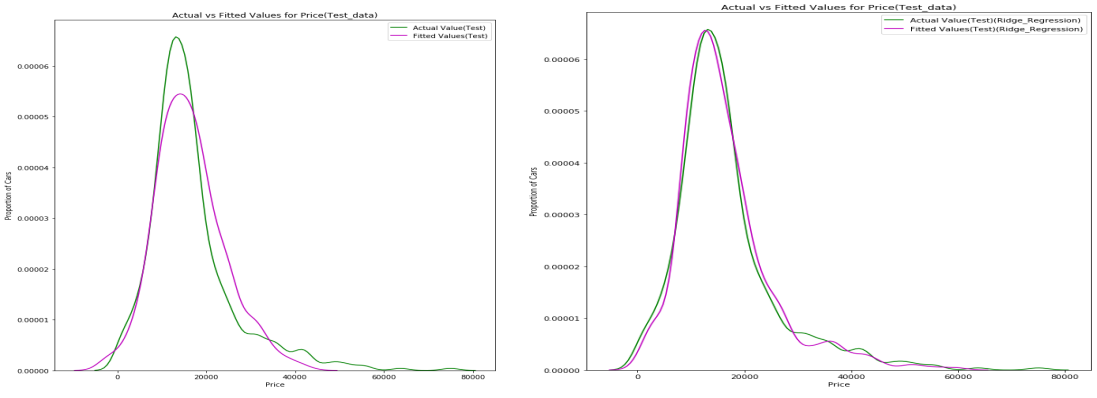
- Data_cleaning and Exploration :Numpy, Pandas
- Plotting: Seaborn,matplotlib.pyplot
- Machine learning : sklearn

Algorithms:
Linear Regression - Baseline Model
Polynomial Regression - Improved Baseline model
Ridge Regression - Final Model

I started with simple Linear regression as it is the most basic model and upon checking the distribution plots there is definitely some room for improvement.So I tried polynomial regression but the R-score is Negative which indicates the current model is actually worse than the previous. Finally went with Ridge regression since it just adds bias to the linear regression to reduce the effect of collinear features which makes the model difficult to learn the function.

Results:
The R-score of the Linear Regression : 0.7456099370960487
The R-score of the Polynomial Regression : -48.24927413031234
The R-score of the Linear Regression : 0.8517594846848306

Conclusion: Below figure is distribution of values predicted Vs Actual Values(Left-LinReg,Right-RidgeReg)



Ridge-Regression model pretty much predicts values which are closer to actual values.