SOLUTION ARCHITECTURE

8. DATA SCALING

1. Scale the dataset using scaling metrices such that all the features have to values from 0 to 1

2. DATA PREPROCESSING & EXPLORATORY DATA ANALYSIS

1. Determine the total no.of.features & unique element in each features
2. Find the null values & missing data
3. Findout and eliminate non-representative samples and brought up the useful infomation about the dataset.

7. SPLIT THE PROCESSED DATA

1. In 80:20 ratio divide the data into train and test test

2. DATA ACQUISITION

1. Collect the data includes dofferent car features & prices
2. Data was collected from various website by web scrapping.

4. FURTHER DATA CLEANING & REMOVING OUTLIERS

1. Validate the consistencies in the data.
2. Remove outliers in the data

6. FEATURE SELECTION

1. Visualize correlations among different features & find linear correlation among every features

9. SELECT MODELS & TRAIN MODELS

1. Select regression (linear, lasso regression, decision tree, random forest & extreme gradient boosting are used for training.
2. Train the model by the selected regression on trainset.

1. PROBLEM FORMULATON

1. Predict the price of the second hand cars over the areas.
2. Predictions will be generated based on vehicles features (such as brand, model, year, mileage etc)

5. DATA ENCODING

1. Receive the features & scale the features by loading the scalar & convert them into a data frame.

10. MODEL EVALUATION & DEPLOY

1. Validate & save the best model.
2. Predict the price by loading the trained model.
3. Show the results on the webpage.