Project Description:-

The following project aims to predict mileage per gallon (mps) using various technical specifications (features) as input to the regression algorithms.

Database Description:-

The data is technical spec of cars. This dataset is a slightly modified version of the dataset provided in the StatLib library. In line with the use by Ross Quinlan (1993) in predicting the attribute "mpg", 8 of the original instances were removed because they had unknown values for the "mpg" attribute. The original dataset is available in the file "auto-mpg.data-original".

"The data concerns city-cycle fuel consumption in miles per gallon, to be predicted in terms of 3 multivalued discrete and 5 continuous attributes." (Quinlan, 1993)

Number of Instances: 398

Number of Attributes: 9 including the class attribute

Attribute Information:-

Mpg: continuous cylinders: multi-valued discrete displacement: continuous horsepower: continuous weight: continuous acceleration: continuous model year: multi-valued discrete origin: multi-valued discrete car name: string (unique for each instance) Missing Attribute Values: horsepower has 6 missing values.

Libraries Involved:-

1. pandas
2. Numpy
3. Seaborn
4. Matplotlib

Steps Involved:-

1. Importing the libraries
2. Loading the dataset
3. Data Preprocessing
4. Study Correlation
5. Univariate Analysis
6. Bivariate Analysis
7. train and test data split
8. Building the model

Machine Learning Steps Involved:-

1. Linear Regression
2. Polynomial Regression