**Multi Linear Regression**

In ML Algorithms we have :

1. Supervised Learning
2. Unsupervised Learning

In **Supervised Learning** we have :

1. Regression
2. Classification

Firstly, we discuss about different types of *Regressions*.

* 1. Linear Regression
     1. Simple Linear Regression
     2. Multi-linear Regression
  2. Logistic Regression
  3. Lasso Regression
  4. Ridge Regression

*In this repository we discuss about Multi Linear Regression*

**Multi-Linear Regression:**

It is applicable when relationship between input variable and output variable is linear, that is it should have positive or negative correlation between input and output variable. That can be known using 'Scatter plot'.When there are more then one independent variables and one dependent variable then that is Multi-Linear regression.

In case of this we use Pair plot to view relationship between variables.

Here X1,X2,X3,..... are 'independent variables' and Y is 'dependent variable', both X and Y are continuous.

After getting scatter plot we will get a best fit line using OLS (ordinary least squares) method

OLS method: we will find distance between actual and predicted value on the line and this is the error (e1), similarly for all data points e2,e3,.....

e1^2+e2^2+e3^2+...........+en^2=(error value)

which line will get this 'error value' as less that is best fit line.

Now using this best fit line we will build a model.

But there is a problem of 'Multi-collinearity' between input varaibles, due to which error of one variable effects other this can be known using R^2 as its value will be near to 0 => bad model.

So, now we do validation using 'VIF(Variance Inflation Factor)' ,by which we will know multi-collinear variables and can eliminate one of them to get better model. And also check for otliers if any using influence plot and delete them.

This process is done till we get:

1. R^2 - coefficient of determination

R^2 is nearly 1 => Good model

2. RMSE (Root Mean Square Error) should be less

**Data used:**

50\_startups: Using Multi-Linear model predicting profit depending upon R&D spend, Administration, Marketing Spend, state

Toyota Corolla: Predicting Price of Toyota depending on its Age, Manufacture date and so on

**Programming:**

Python

The Codes regarding this Multi-Linear Regression model with three different business problems *Cars MPG prediction* ,*Start-ups profit prediction*, *Toyota price prediction* with their datasets are present in this Repository in detail