Machine Learning: Lab 4 – Linear Regression and Gradient Descent

Download the **"profits.txt"** file containing the profits of a restaurant for corresponding population in some cities.

Prerequisites: Python basics, numpy, pandas, matplotlib, etc.

Do NOT use sklearn library for this assignment.

Importing Data:

1. Import the profits data from the txt file. The first column contains the population of a city in 10,000s while second column contains the profit in $10,000s.

   A) show the dataframe details, rows, columns, dtypes and range of values.

   B) create a scatter plot showing the data points.

Linear Regression:

Create a class MyLinearRegression with the following functionality:

1. Create a constructor, that takes the input feature matrix X and the corresponding labels vector Y. These will be compulsory arguments. Check for valid dimensions of the dataframes and display appropriate error message.

2. Following Instance variables to be defined; learning rate, total iterations, model parameters, etc.

3. Method for computing cost function ($J(\Theta)$).

4. Method to implement "batch" gradient descent algorithm for specified number of iterations. Draw a plot showing the cost function values w.r.t. number of iterations.

5. A function that will show the 3D surface plot of parameters with respect to the cost function. Choose appropriate range of values of theta and compute corresponding values.

6. A function that will plot the a linear model for specified values of Θ on the scatter plot shown in question 1B above.

7. Find the gradients after the first iteration of gradient descent when all the rows are considered in the training set and starting value of $\Theta = (0, 0)$.

8. Draw another 3D surface plot of parameters with respect to the cost function after feature normalization.

(Optional)

Use Normal Equation to find the parameters Θ.