Assignment 3

Simple Linear Regression

In this assignment, you need to implement a simple linear regression model using gradient descent algorithm on the following database.

Database
Independent/Predictor Variable -

xxxxxxxz

Dependent/Response Variable -

zzzzzzzz

You are required to use the following Cost function

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta\left(x^{(i)}\right) - y^{(i)} \right)^2$$

Answer the following questions based on your observations
   1. Use linear regression to fit a straight line to the given database. Set your learning rate to 0.5. What are the cost function value and learning parameters values after convergence? Also, mention the convergence criteria you used.

   2. The cost function that we are using in this assignment is different than the one we used in class. Can you think of the advantage of averaging the cost?

3. Plot cost function v/s iteration graph for the model in question 1 for first 50 iterations.

4. Plot the given dataset on a graph and also print the straight line you obtained in question 1 to show how it fits the data.

5. Test your regression model with the learning rates
   lr = 0.005, lr = 0.5, lr = 5
   For each learning rate, plot a graph showing how the cost function changes for the first 50 iterations and write your observation.

6. Choose a suitable learning rate, then implement stochastic and min-batch gradient descent, plot the cost function against iteration, and observe how your cost function changes compared to batch gradient descent.

Note

1. Use batch gradient descent unless mentioned in the question.
2. Your code should be in Python.
3. Normalize the data before training.
4. You are required to implement the linear regression model yourself. You are not allowed to use libraries like sci-kit learn, Tensorflow, Pytorch, etc, and use their built-in linear regression models.
5. You may use libraries like numpy, pandas, matplotlib, etc to read and manipulate the dataset and plot the graphs.
6. Your plots should be labeled properly.
7. Normalize your predictor before starting training the model.

How to Submit?
1. You need a submit a single PDF report file and the name of that file should be your roll numbers.
2. Your report should have answers to all the questions with the plots. Upload your code on github and mention the link in the report. Do not copy paste code in the report.