ELL409
Assignment 1

1. **Gradient Descent**

   - You have been given a dataset of 10000 points.
   - You should remove any outliers, if required.
   - You have to then build a Linear Regression model to fit on the data where you have to minimize the Least Squares loss.
   - The Least Squares loss is defined as:

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

   - You should write the algorithm of Gradient Descent from scratch.
   - Gradient Descent: The goal is to minimize the cost function $J(\theta)$. The update rule for Gradient Descent is given by:

   $$\theta := \theta - \alpha \frac{\partial J(\theta)}{\partial \theta}$$

   where α is the learning rate.
   - Convergence Criteria: The algorithm converges when the change in the cost function $J(\theta)$ is less than a small value ϵ, i.e.,

   $$|J(\theta^{(t)}) - J(\theta^{(t-1)})| < \epsilon$$

   where $\theta^{(t)}$ and $\theta^{(t-1)}$ are the parameters at iterations t and $t-1$, respectively.

2. **Stochastic Gradient Descent (SGD)**

   - You should write the algorithm for Stochastic Gradient Descent from scratch.
   - Stochastic Gradient Descent (SGD): Unlike batch gradient descent, which uses the entire dataset, SGD updates the parameters for each training example $(x^{(i)}, y^{(i)})$:

   $$\theta := \theta - \alpha(h_\theta(x^{(i)}) - y^{(i)})x^{(i)}$$

   - Convergence Criteria for SGD: The algorithm converges when the moving average of the change in the cost function $J(\theta)$ is less than a small value ϵ, i.e.,

   $$\frac{1}{k} \sum_{i=1}^{k} |J(\theta^{(t-i)}) - J(\theta^{(t-i-1)})| < \epsilon$$

   where k is the number of past iterations considered in the moving average.

Submission Guidelines

- Your submission should include the Python script named `run.py`.
- The script should take the following arguments:
  - `data_path`: Path to the dataset.
  - `num_epochs`: Number of epochs to run the algorithm.
  - `batch_size`: Batch size for Stochastic Gradient Descent.
  - `learning_rate`: Learning rate for Gradient Descent.
- The command to run your script should be in the format:

```
python run.py --data_path <data_path> --num_epochs <
    num_epochs> --batch_size <batch_size> --
    learning_rate <learning_rate>
```

- The output should be the final parameters learned by the linear regression model.
- You should also include a report (not more than two pages) detailing:
  - Keep convergence criteria fixed at ϵ = 1e-5
  - The number of iterations it took to converge for varying learning rates. (lr=0.1, lr=0.01, lr=0.001)
  - The time for convergence for varying batch sizes.(1, 10, 100, 1000)
  - Training and validation loss plots.
- Submit the Python script, the report, and any other necessary files in a zipped folder named `assignment1_yourfirstname_entrynumber.zip`.