# Machine Learning

## Assignment 1

**Submitted By:** Subhadeep Das
**Roll No.:** 2105931
**Section:** ML CSE - 3
**Branch:** Computer Science and Engineering
**Faculty:** Prof. Sohail Khan
**Semester :** 6th

**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.**

Ans. - Cost function value = 0.5561400558913063
Slope (m) = 0.6619554709942285
Intercept (c)= -2.4064644777449575e-15
the convergence criteria are based on the number of epochs (epochs).

```
Epoch 1000, Cost: 0.5561400558913063
Slope (m): 0.6619554709942285
Intercept (c): -2.4064644777449575e-15
```

**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?**

Ans.

Yes, averaging the cost in a cost function, as shown in the image, can have several advantages :

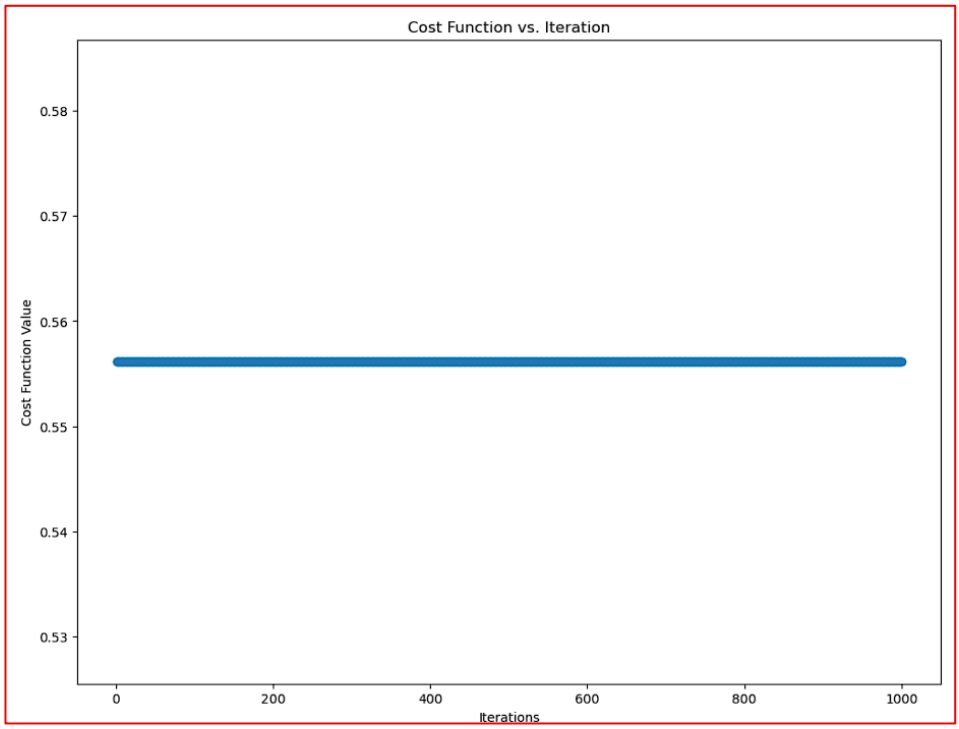
**Stabilization:** The process of averaging the cost contributes to the stabilization and acceleration of the gradient descent optimization. Smaller cost values resulting from averaging can prevent excessive oscillation in the learning algorithm.

**Normalization:** Facilitating the normalization of the cost function, averaging makes it simpler to compare across different datasets that have varying numbers of samples. This normalization provides a clearer perspective on the model's performance.

**Mean Squared Error:** The provided formula calculates the mean squared error cost function, commonly utilized in regression problems. Averaging in this context is advantageous as it provides an indication of the average prediction deviation across all data points.

**3 & 4. Plot cost function v/s iteration graph for the model in question 1. 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.**

## Ans.




**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 every iteration and write your observation.**

**Ans.** **For l = 0.005**

Cost function value : 0.5561400569026452

Slope (m): 0.6619238240082247

Intercept (c): -2.4148058692955056e-15

## For l = 0.5

Cost function value = 0.5561400558913063

Slope (m): 0.6619554709942285

Intercept (c): -2.4064644774449575e-15

## For l = 5

Cost fucntion value: nan

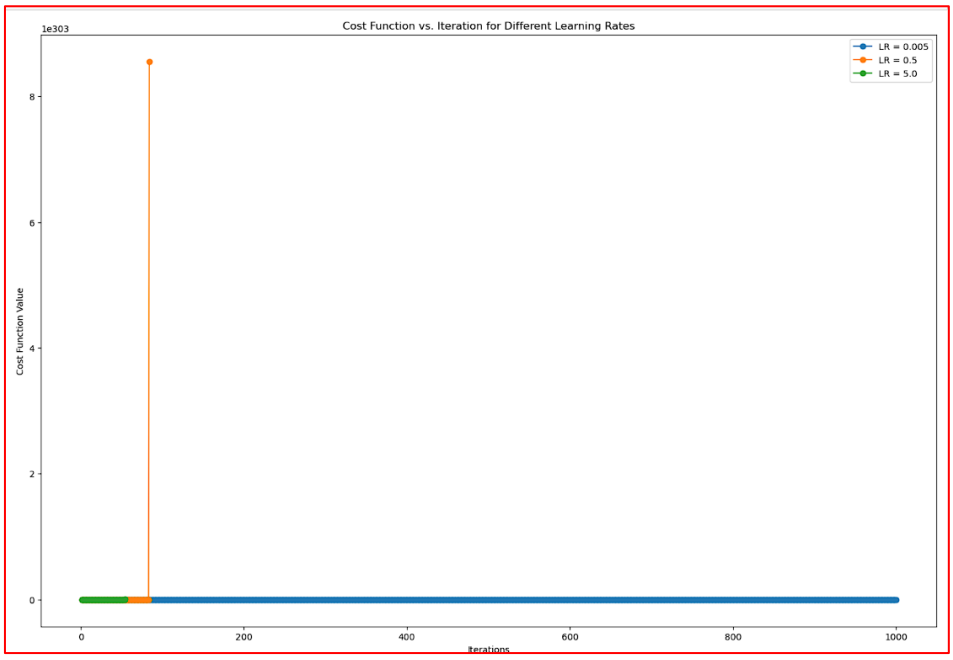
Slope (m): nan

Intercept (c): nan

A learning rate of 5 is relatively high and may lead to overshooting the minimum point of the cost function. In gradient descent, the learning rate determines the step size taken during each iteration to update the model parameters.

A very high learning rate can cause the algorithm to oscillate or diverge instead of converging to the minimum.

**Comparisons:**

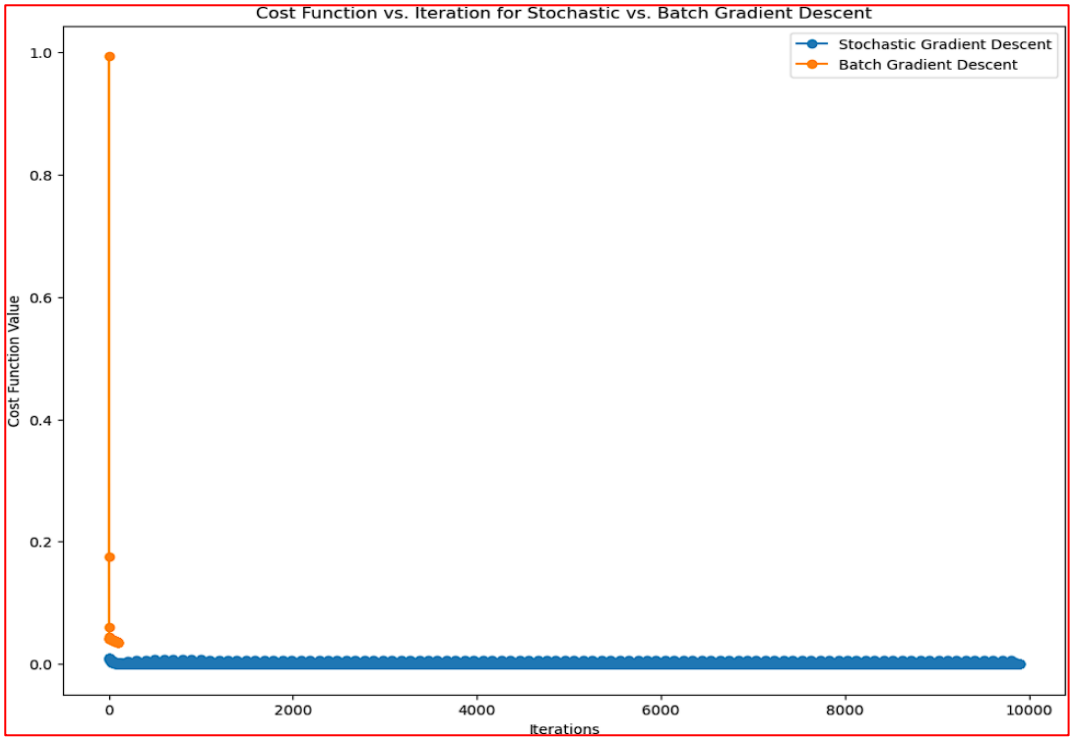


**6. Choose a suitable learning rate, then implement stochastic and min-batch gradient descent, plot the cost function against iteration, and observe how your cost function changes compared to batch gradient descent.**

**Ans.**

Learning Rate = 0.01

Epochs = 100




The cost function is low and normal for stochastic gradient descent and high for batch gradient descent .

Github Link – https://github.com/subhagittu/ML_Assignment_1

**Subhadeep Das (2105931)**