ML Assignment-1

Nirish Singh (21053212)

Q1)

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
```

Q2)

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

Normalization: It helps in normalizing the cost function, making it easier tocompare across different datasets with varying numbers of samples. This normalization provides a clearer view of how well the model is performing.

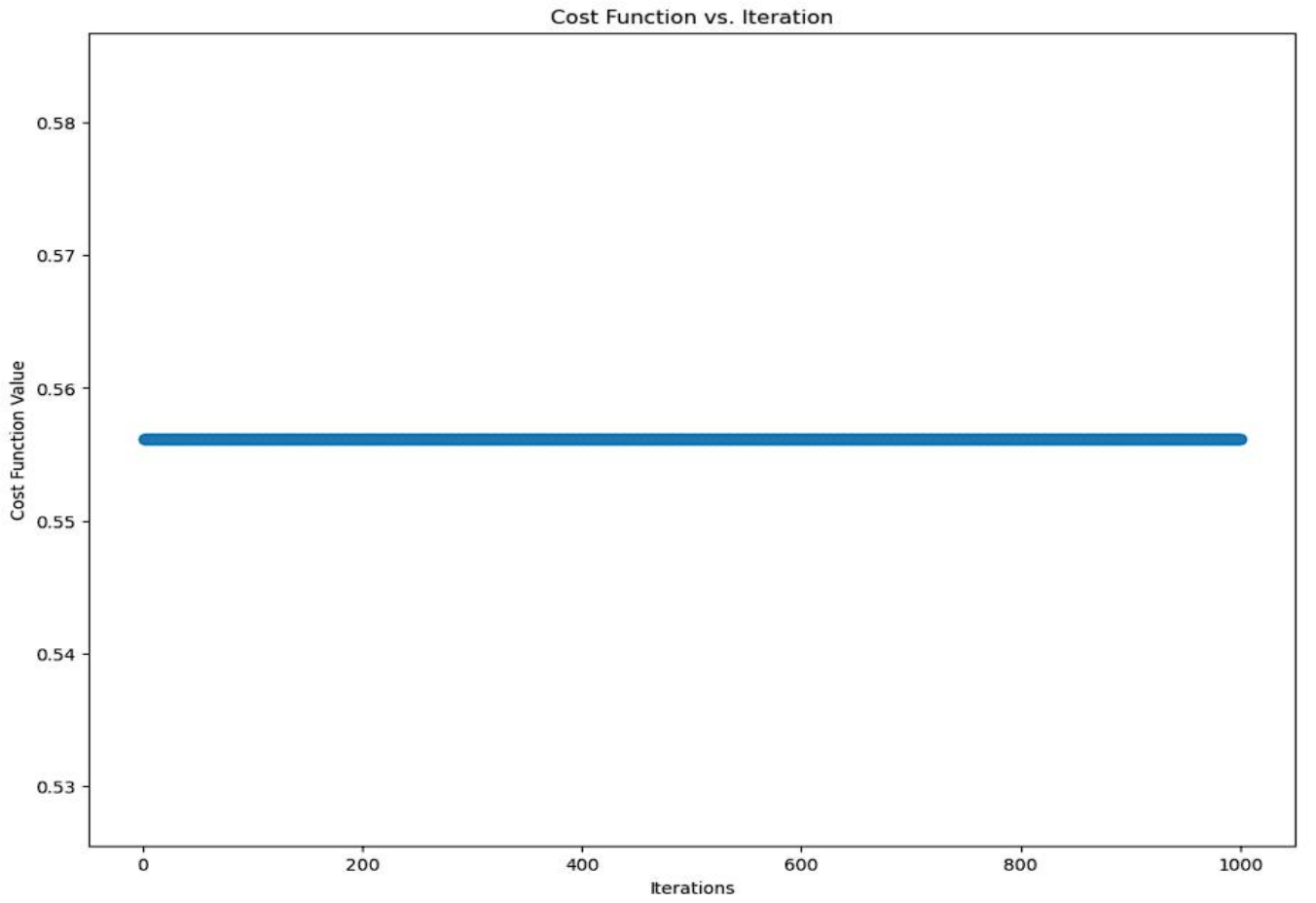
Stabilization: Averaging the cost aids in stabilizing and speeding up the gradientdescent optimization process. This is because smaller cost values (resulting
from averaging) can prevent the learning algorithm from oscillating excessively.

Mean Squared Error: The formula you provided is for calculating the mean squared error cost function, often used in regression problems. Averaging in this context is beneficial as it gives an idea of how off the predictions are, on average,across all data points.

ML Assignment-1

Nirish Singh (21053212)

Q3) & Q4)



ML Assignment-1

Nirish Singh (21053212)

Q5)

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)= -2.4064644777449575e-15

For l = 5

Cost function value: nan

Slope (m)= nan

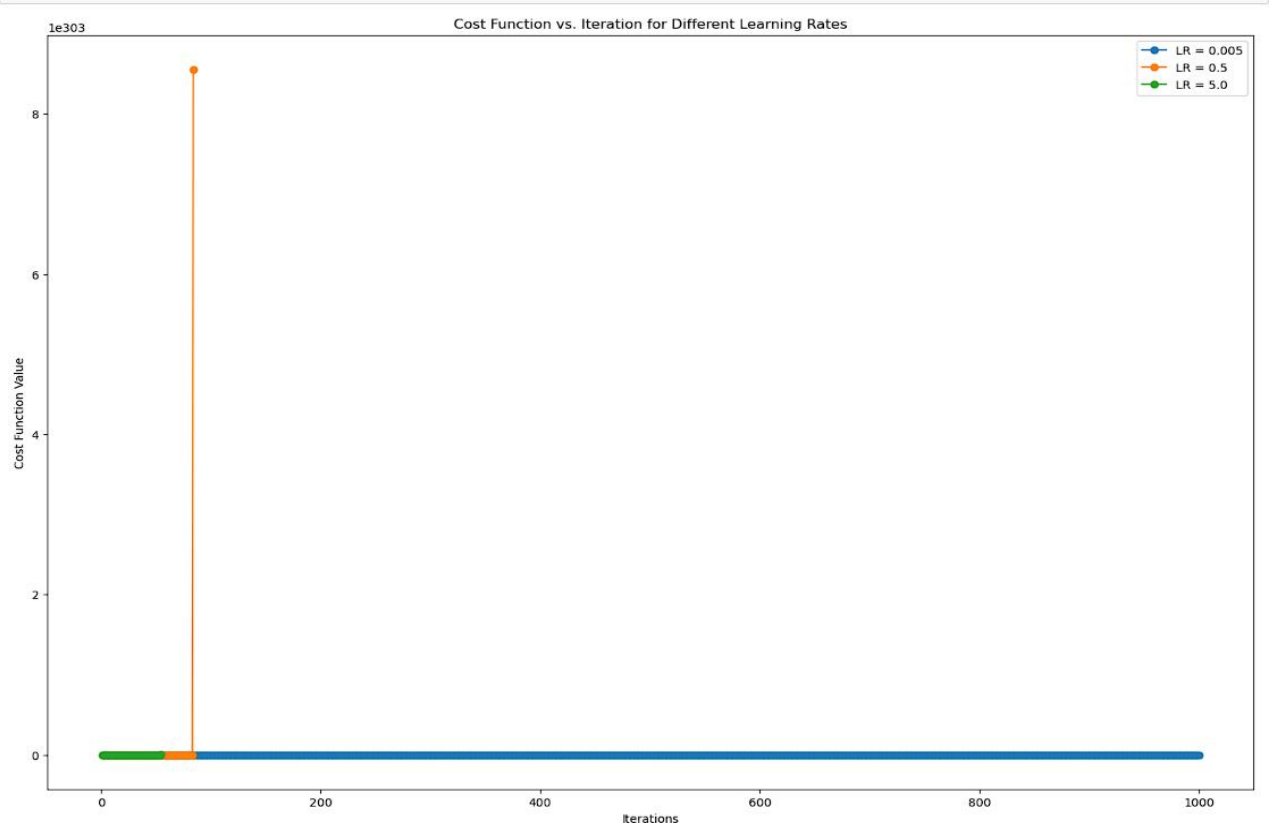
Intercept (c)= nan

A learning rate of 5 is relatively high and may lead to overshooting the minimum point ofthe 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.

ML Assignment-1

Nirish Singh (21053212)

Comparisons:

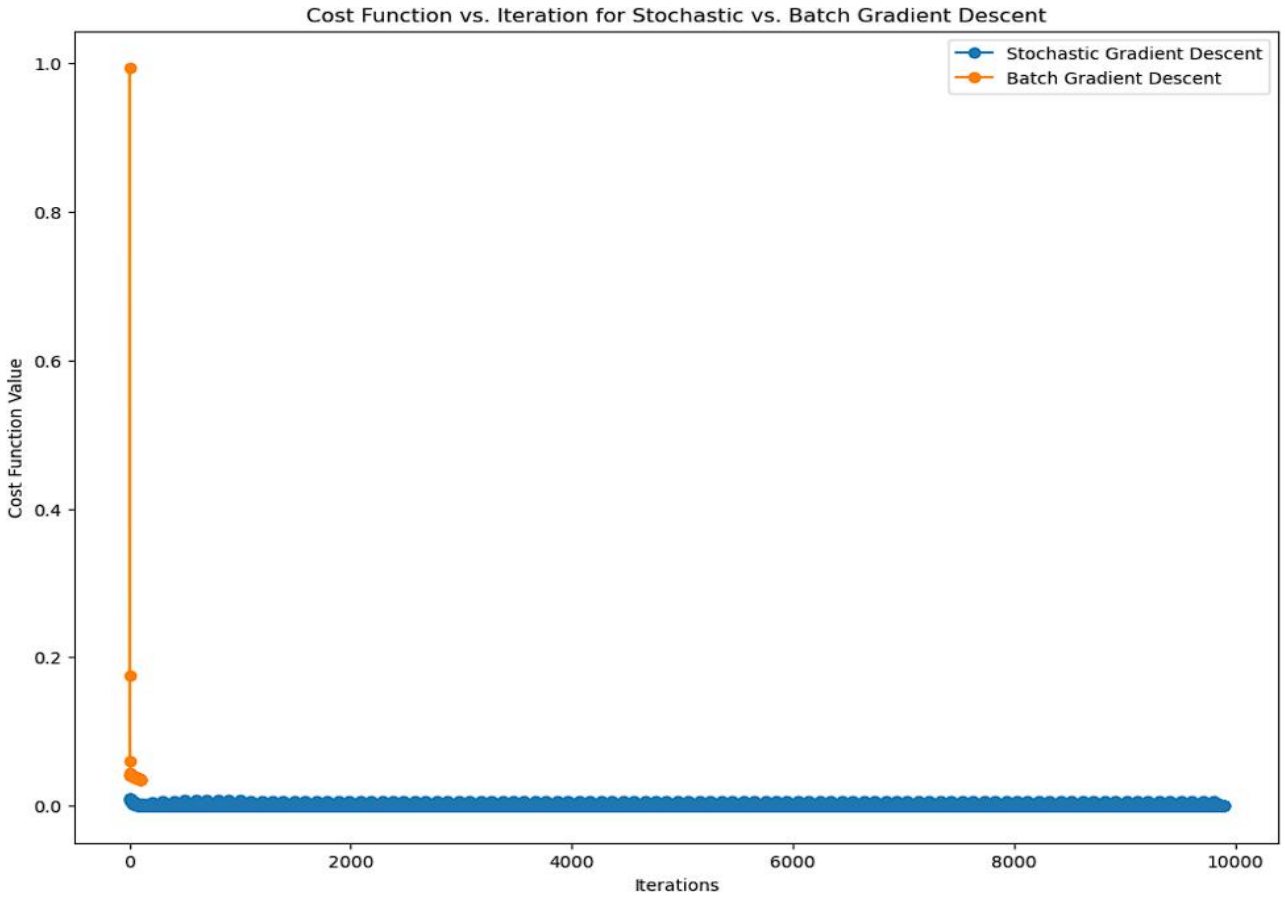


ML Assignment-1

Nirish Singh (21053212)

Q6)

Learning Rate = 0.01
Epochs = 100



The cost function is high for Batch Gradient Descent and Low and normal for StochasticGradient Descent.

GITHUB LINK:https://github.com/3212-NirishSingh/ML_Assignment