Name: Abdirahim Ahmed
Student ID: 801338570

Link to Github repository:https://github.com/Abdiirahim/ECGR-4105-Intro-to-ML/tree/main/Homework1

Homework 1

In this homework, we will use a dataset provided: “D3.csv” Let the first three columns of the data set be separate explanatory variables x1, x2, x3. Let the fourth column be the dependent variable Y. (Note: You cannot use the built-in function from ML libraries for gradient descent, you have to implement it yourself.)

Problem 1 (40 points)

Develop a code that runs linear regression with gradient descent algorithm for each of the explanatory variables in isolation. In this case, you assume that in each iteration, only one explanatory variable (either X1, or X2, or X3) is explaining the output. Basically, you need to do three different training, one per each explanatory variable. For the learning rate, explore different values between 0.1 and 0.01 (your choice). Initialize your parameters to zero (theta to zero).

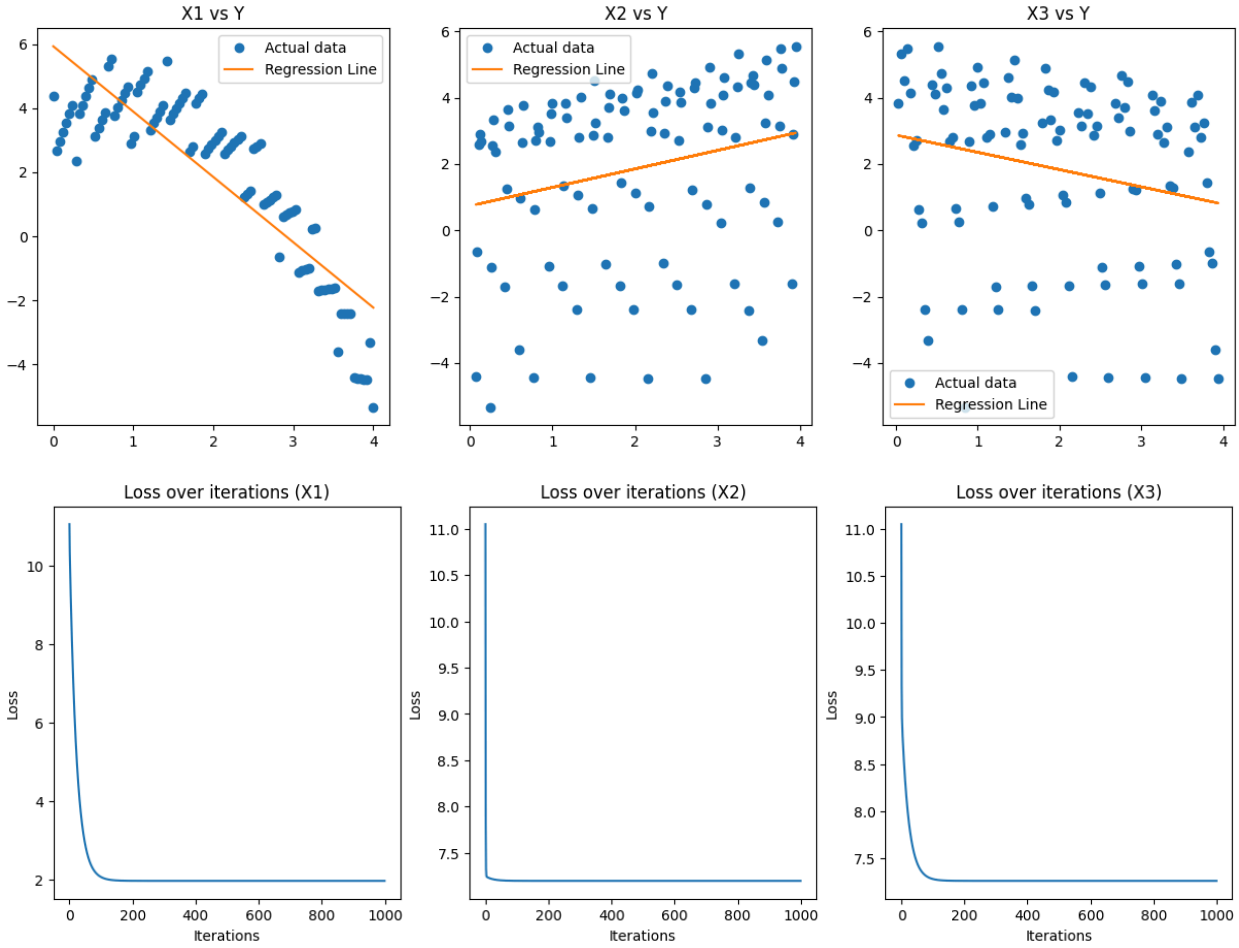
Report the linear model you found for each explanatory variable.

{'X1': (-2.038336633229477, 5.9279489169790756),
 'X2': (0.5576076103651677, 0.7360604300111252),
 'X3': (-0.5204828841600003, 2.8714221036339524)}

X1: Y=5.928-2.038X1
X2: Y=0.736+0.558X2
X3: Y=2.871-0.520X3

Plot the final regression model and loss over the iteration per each explanatory variable.

Name: Abdirahim Ahmed
Student ID: 801338570



Which explanatory variable has the lower loss (cost) for explaining the output (Y)?

X1 has the lowest final loss (cost), with a value of 1.970

Based on your training observations, describe the impact of the different learning rates on the final loss and number of training iteration.

Learning rate = 0.01
A learning rate of 0.05 was found to be quite effective, leading to a smooth and stable reduction in the loss as the model converged. On the other hand, a higher learning rate, such as around 0.1, often caused the loss to fluctuate or diverge, preventing the model from converging properly. Conversely, a lower learning rate, like 0.01, while stable, required significantly more iterations to reach convergence, making the training process slower.

Problem 2 (60 points)

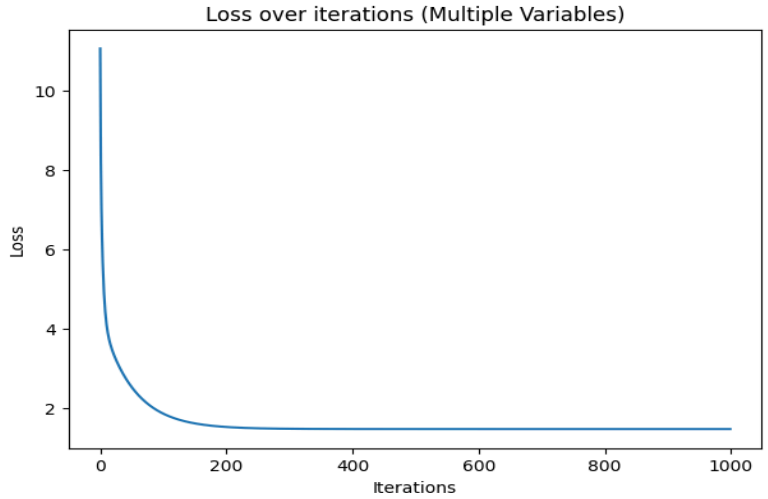
Name: Abdirahim Ahmed
Student ID: 801338570

This time, run linear regression with gradient descent algorithm using all three explanatory variables. For the learning rate, explore different values between 0.1 and 0.01 (your choice). Initialize your parameters (theta to zero).

Report the final linear model you found the best.

Given this array([5.31393577, -2.00368658, 0.53260157, -0.26556795])
Based on the final lear model using all three explanatory variables
Y=5.314-2.004X1+0.533X2-0.266X3 is the best

Plot loss over the iteration.
array([5.31393577, -2.00368658, 0.53260157, -0.26556795])



Based on your training observations, describe the impact of the different learning rates on the final loss and number of training iteration.

With a learning rate of 0.05, the loss decreased smoothly over 1000 iterations. A higher rate, like 0.1, caused errors and instability. A lower rate would need more iterations to achieve similar results.

Predict the value of y for new (X1, X2, X3) values (1, 1, 1), for (2, 0, 4), and for (3, 2, 1)

array([3.57728282, 0.24429082, 0.10251123])

Y values based on the new x-values

X1	X2	X3	Y
1	1	1	3.57728282

Name: Abdirahim Ahmed
Student ID: 801338570

2	0	4	0.24429082
3	2	1	0.10251123