Name:Abdirahim Ahmed
Student ID: 801338570

Homework 2

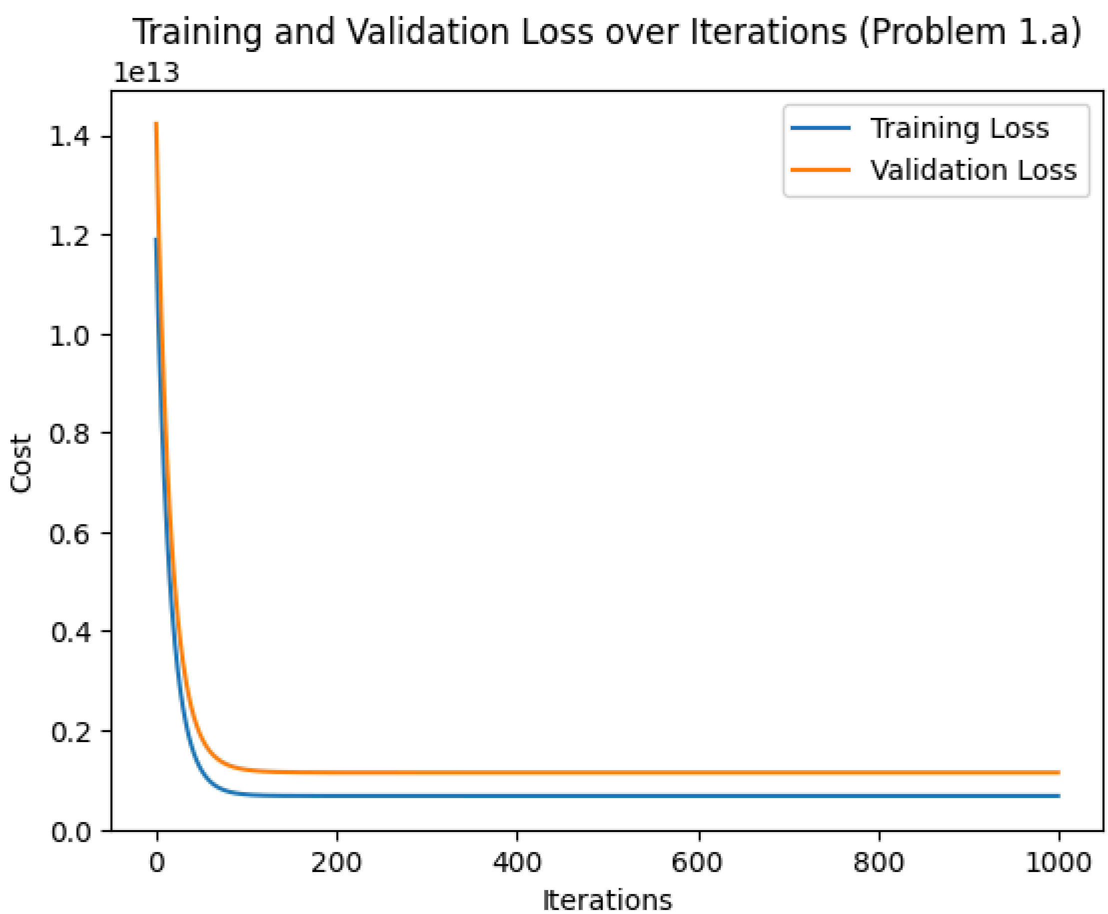
Github repository: https://github.com/Abdiirahim/ECGR-4105-Intro-to-ML/tree/main/HW2

Problem 1 (30 points)

1.a) Develop a gradient decent training and evaluation code, from scratch, that predicts housing price based on the following input variables:area, bedrooms, bathrooms, stories, parking

Identify the best parameters for your linear regression model, based on the above input variables.

Plot the training and validation losses (in a single graph, but two different lines). For the learning rate, explore different values between 0.1 and 0.01 (your choice). Initialize your parameters (thetas to zero). For the training iteration, choose what you believe fits the best.



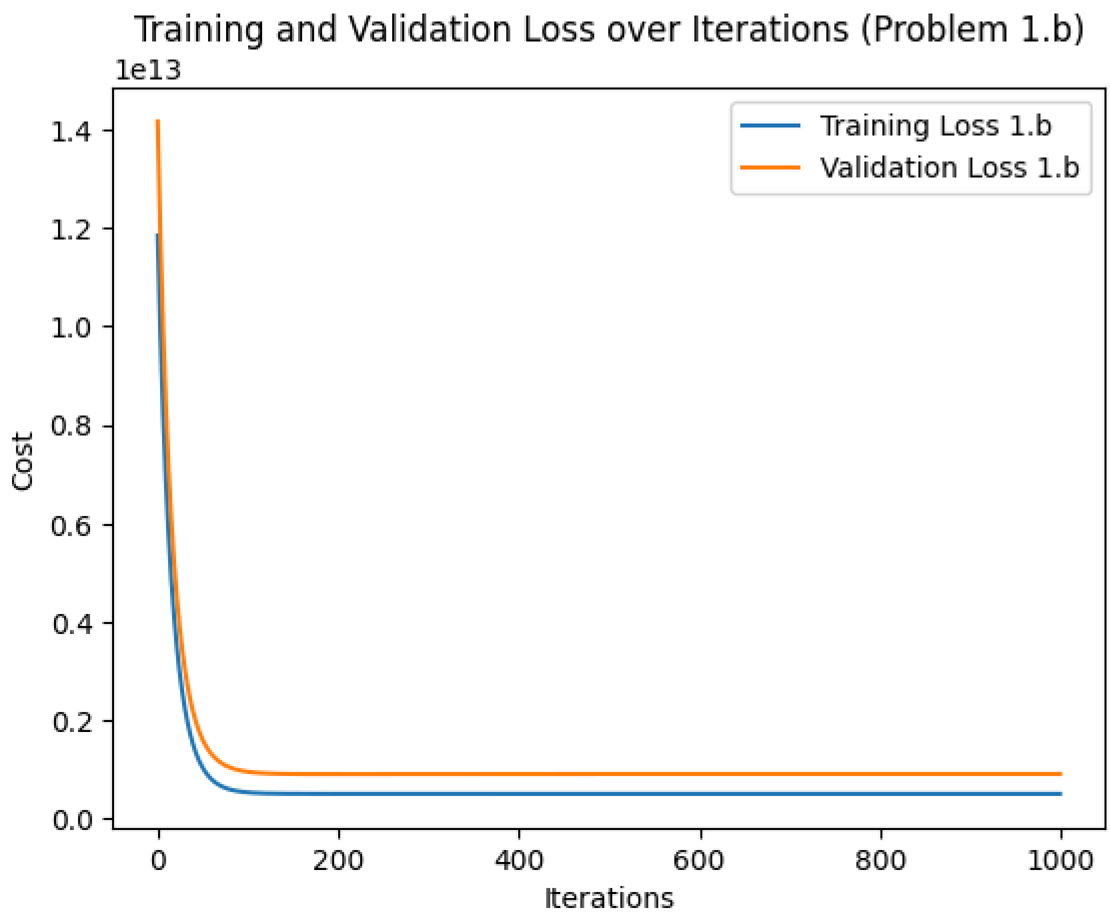
1.b) Develop a gradient decent training and evaluation code, from scratch, that predicts housing price based on the following input variables:

Name:Abdirahim Ahmed
Student ID: 801338570

Area, bedrooms, bathrooms, stories, mainroad, guestroom, basement, hotwaterheating, airconditioning, parking, prefarea

Identify the best parameters for your linear regression model, based on the above input variables.

Plot the training and validation losses (in a single graph, but two different lines) over your training iteration. Compare your linear regression model against problem 1 a. For the learning rate, explore different values between 0.1 and 0.01 (your choice). Initialize your parameters (thetas to zero). For the training iteration, choose what you believe fits the best.



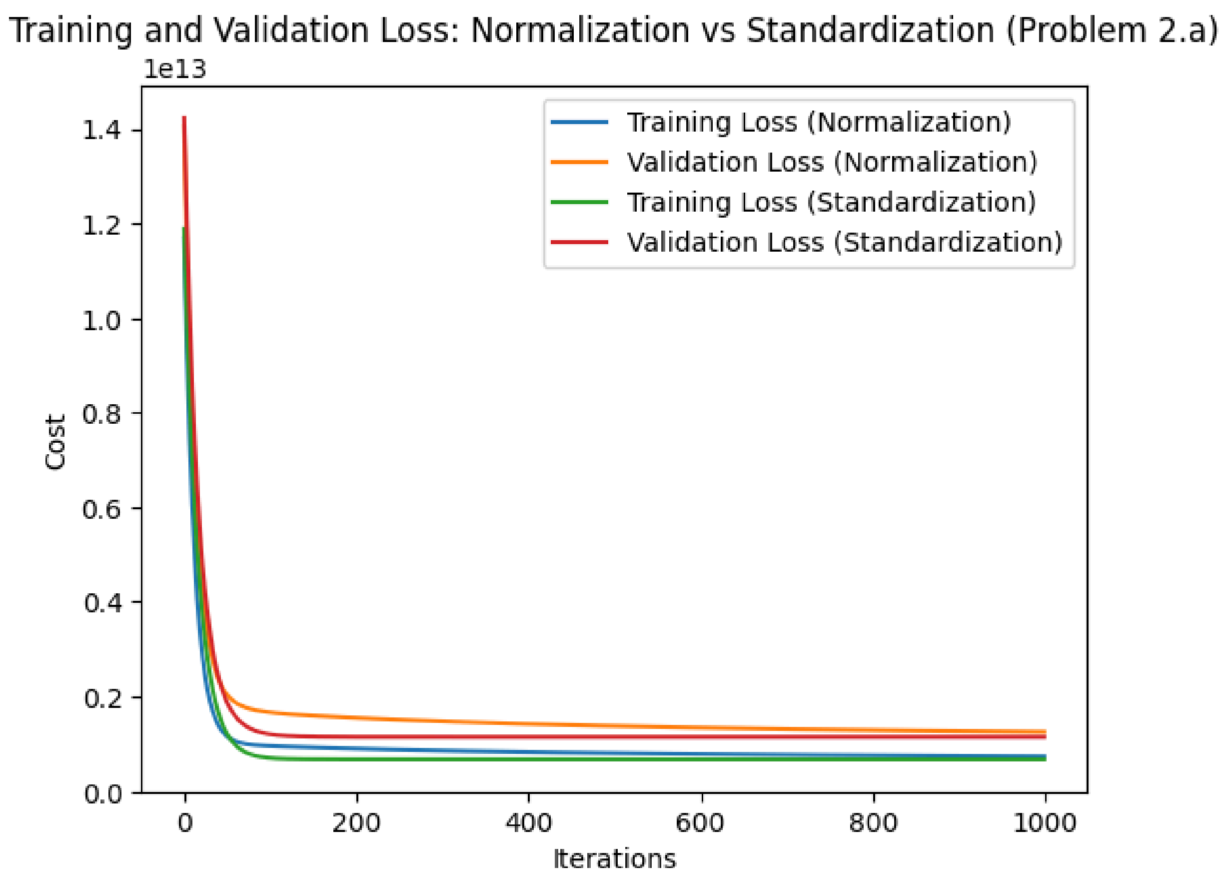
Problem 2 (30 points)

Name:Abdirahim Ahmed
Student ID: 801338570

2.a) Repeat problem 1 a, this time with input normalization and input standardization as part of your pre-processing logic. You need to perform two separate trainings for standardization and normalization. In both cases, you do not need to normalize the output!

Plot the training and validation losses for both training and validation set based on input standardization and input normalization. Compare your training accuracy between both scaling approaches as well as the baseline training in problem 1 a. Which input scaling achieves the best training? Explain your results. Standardization achieved the best training accuracy, with faster convergence and lower training and validation losses compared to both normalization and the baseline (no scaling).Normalization improved accuracy compared to the baseline but not as much as standardization.The baseline (no scaling) had the slowest convergence and the highest losses, showing that unscaled features negatively impact training.

Standardization worked best because it ensured that all input features had similar scales (mean 0 and standard deviation 1), allowing the gradient descent algorithm to perform more efficiently.

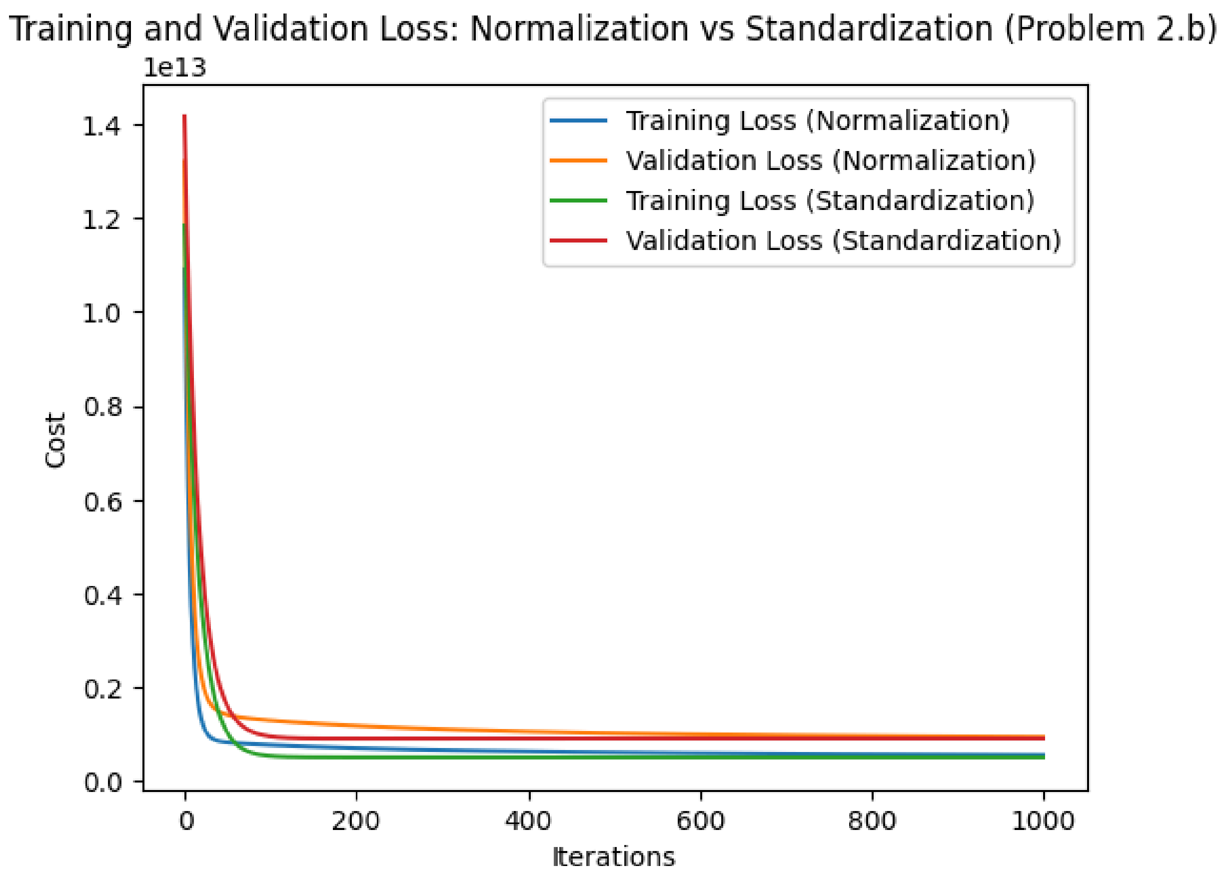


Name:Abdirahim Ahmed
Student ID: 801338570

2.b) Repeat problem 1 b, this time with input normalization and input standardization as part of your pre-processing logic. You need to perform two separate trainings for standardization and normalization. In both cases, you do not need to normalize the output!

Plot the training and validation losses for both training and validation sets based on input standardization and input normalization. Compare your training accuracy between both scaling approaches and the baseline training in problem 1 b. Which input scaling achieves the best training? Explain your results.

Standardization worked best because the feature set included variables with very different scales for area vs. features like air conditioning or parking. Standardization ensured that all features contributed more equally to the learning process, improving gradient descent efficiency and resulting in better model performance



Problem 3 (40 points)

Name:Abdirahim Ahmed
Student ID: 801338570

3.a) Repeat problem 2 a, this time by adding a parameters penalty to your loss function. Note that in this case, you need to modify the gradient descent logic for your training set, but you don't need to change your loss for the evaluation set.

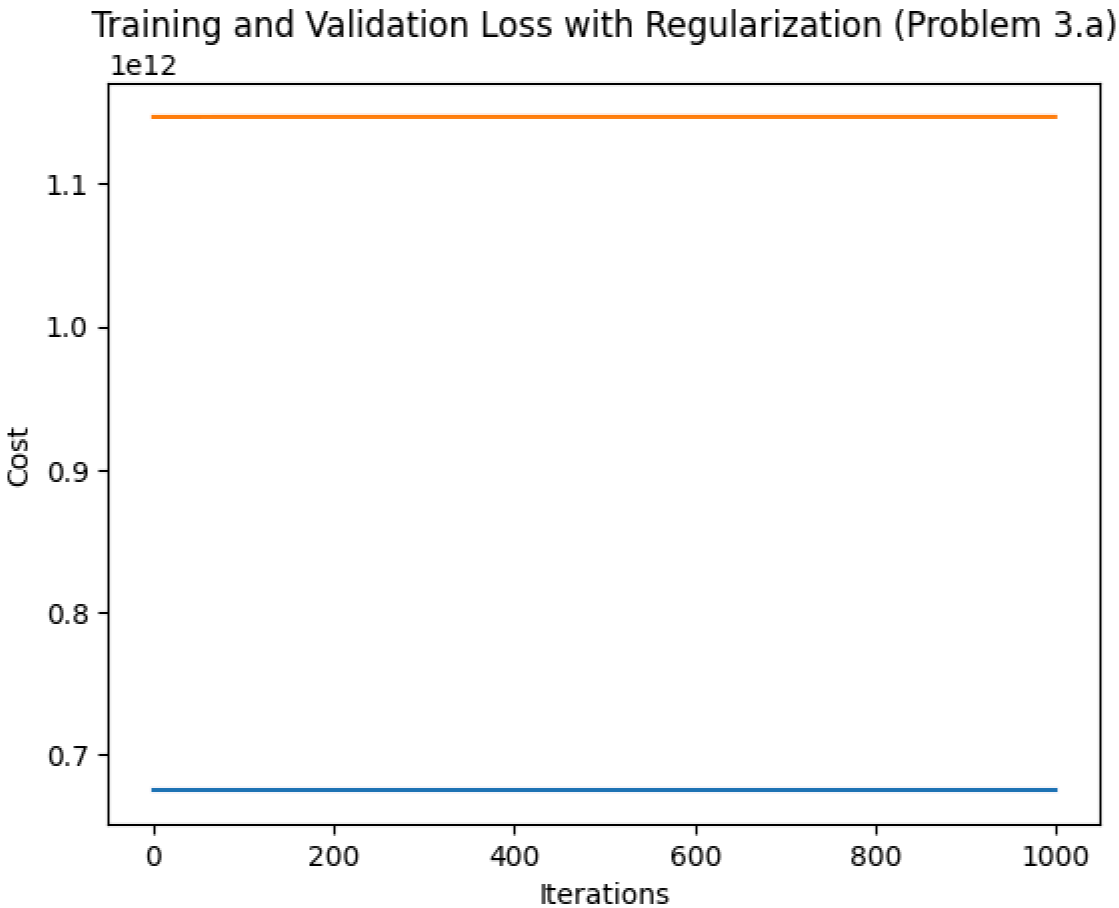
Plot your results (both training and evaluation losses) for the best input scaling approach (standardization or normalization). Explain your results and compare them against problem 2 a.

Standardization again achieved the best training accuracy, with the lowest training and validation losses compared to normalization and the baseline.
Normalization improved accuracy compared to the baseline but did not perform as well as standardization, though it still reduced both training and validation losses more effectively than the baseline.
The baseline (no scaling) showed slower convergence and higher losses, indicating less effective learning and poorer generalization.

Adding regularization reduced overfitting in both cases, as seen by the smaller gap between training and evaluation losses, especially in the standardized approach.

Explanation: The addition of a regularization penalty (which discourages large parameter values) improved the model's ability to generalize to new data by preventing overfitting. Standardization, combined with regularization, was the most effective because it ensured that all features contributed equally to learning, while regularization controlled model complexity, resulting in the best overall performance.

Name:Abdirahim Ahmed
Student ID: 801338570



b) Repeat problem 2 b, this time by adding a parameters penalty to your loss function. Note that in this case, you need to modify the gradient descent logic for your training set, but you don't need to change your loss for the evaluation set.
Plot your results (both training and evaluation losses) for the best input scaling approach (standardization or normalization). Explain your results and compare them against problem 2 b.

Standardization with regularization achieved the best results, showing lower training and evaluation losses compared to normalization and the baseline in Problem 2.b.
Normalization with regularization improved both training and evaluation performance compared to Problem 2.b, but it still did not perform as well as standardization with regularization.

Adding regularization significantly reduced overfitting, as seen by the reduced gap between training and evaluation losses, particularly in the standardized model.

Regularization helped the model generalize better by penalizing large parameter values, which reduced overfitting. Standardization with regularization performed the best because it ensured balanced feature scaling, while regularization controlled model complexity, leading to the most efficient and accurate learning. This combination resulted in the lowest overall losses.

Name:Abdirahim Ahmed
Student ID: 801338570