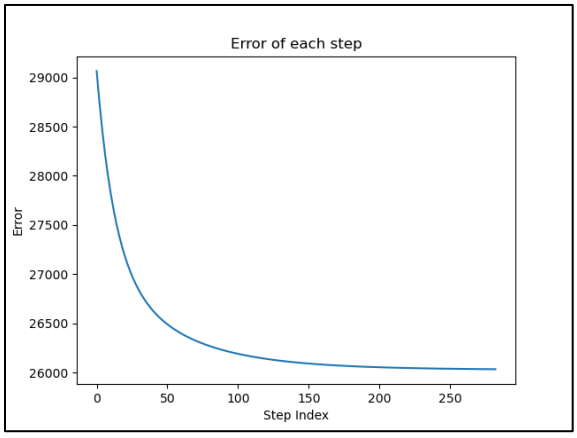
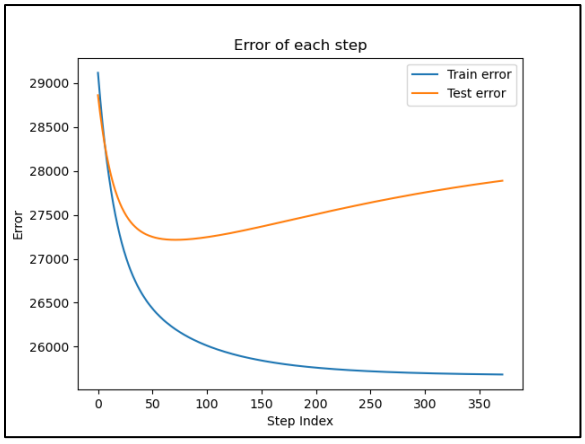
Applied Deep Learning – Homework 2

Task 1



Final error: 26033

Task 2



Final train error: 25682

Final test error: 27889

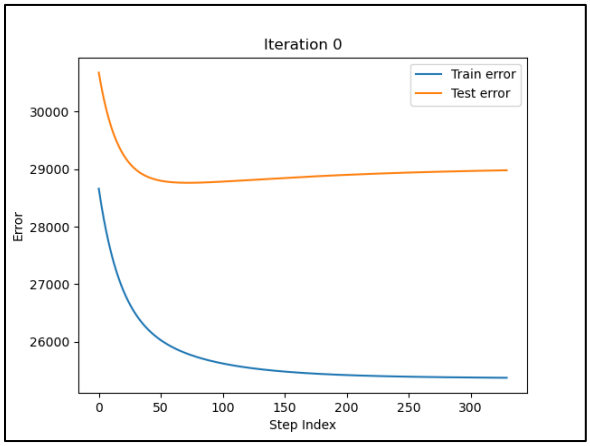
Based on the graph, our estimation is the algorithm **overfits** the data.

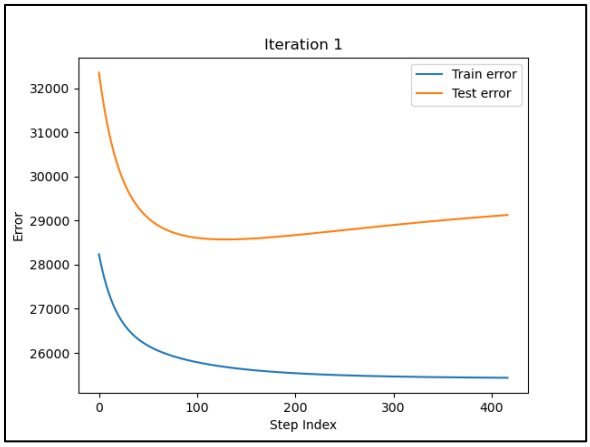
First, we can see the final train error is lower than the one we got when we used all the data (in Task 1). It implies the algorithm got result that is more tailored to the 80% training set, and less to the other 20% test set.

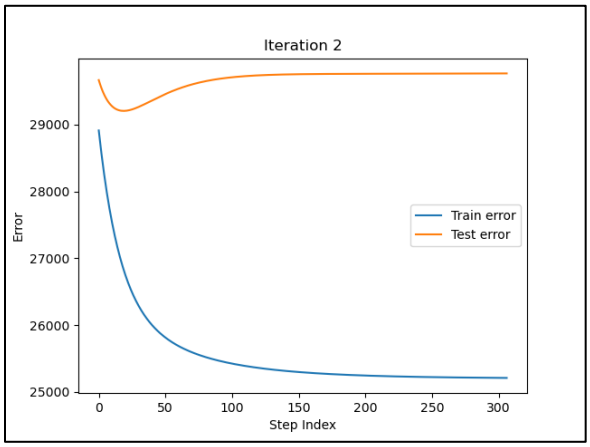
Second, the final test error is higher than the final train error - overfitting.

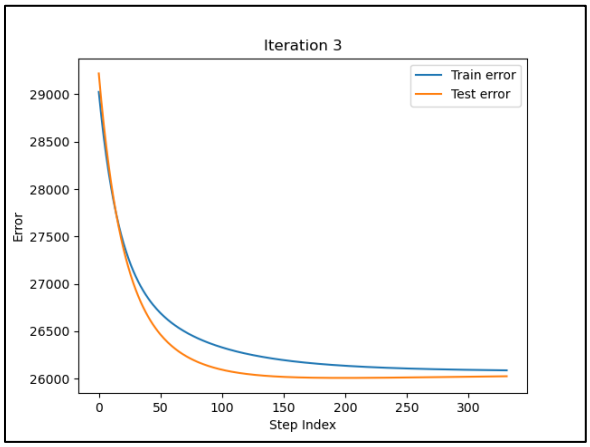
Additionally, we can see that the test error actually gets its lowest point fairly early, and raises from there. This can imply that the more we proceed towards the train data's minimum, the less generalized our model becomes and thus, the test error raises.

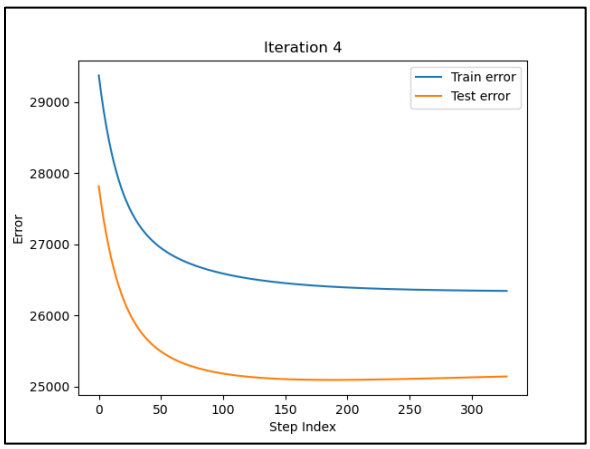
Task 3

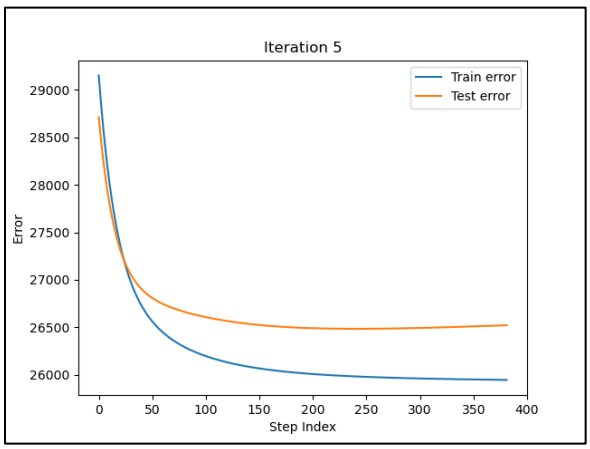

Iteration 0
Train error
Test error
Error
Step Index

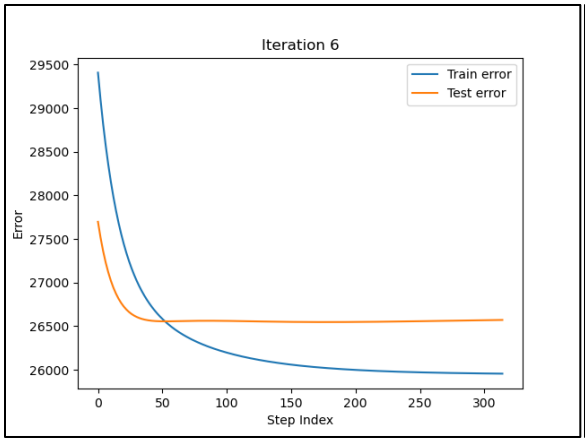

Iteration 1
Train error
Test error
Error
Step Index

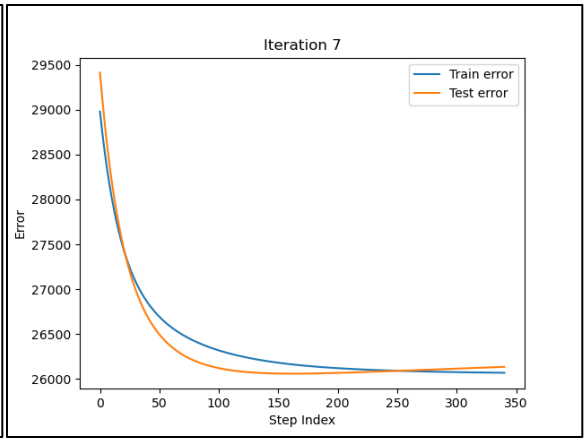

Iteration 2
Train error
Test error
Error
Step Index

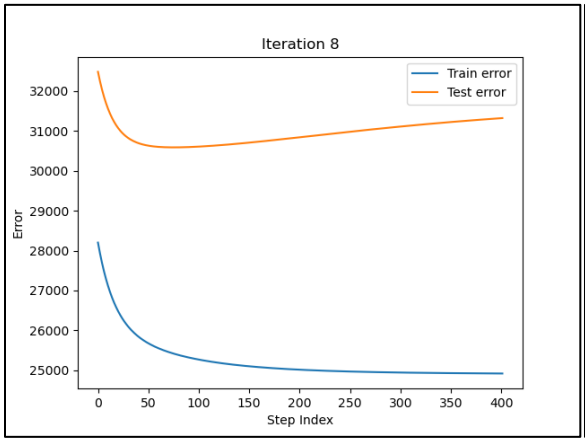

Iteration 3
Train error
Test error
Error
Step Index

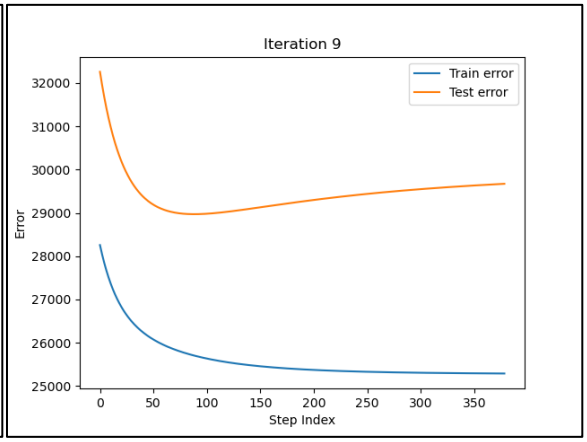

Iteration 4
Train error
Test error
Error
Step Index

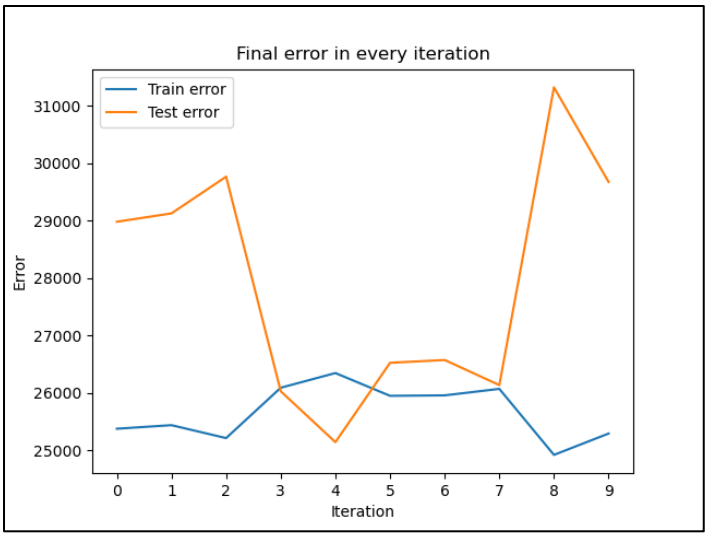

Iteration 5
Train error
Test error
Error
Step Index











Error	Train	Test
Min	24918 (iteration #8)	25140 (iteration #4)
Mean	25663	27926

From the results:

- In 1 iteration (iteration #4) the test error outperformed the train error.
  It can be partly explained by the relatively high error found on the train set in this iteration.
  Maybe the data contains several subpopulations that act slightly different,
  And in this specific train-test split points that acts “a bit weird” was over-represented on the train and under-represented on the test.
  In such a small dataset, changes of few samples between the train and test can have a great influence on the result.
- In 2 iterations (iterations #3, #7) the prediction was very similar in the train and the test – which is what we expect when there is no overfitting.
- In 7 iterations (iterations #1, #2, #5, #6, #8, #9, #10) the prediction worked better on train-set than the test-set – which means the algorithm **overfits** the data.

We can see that in most of the runs the algorithm tends to overfit the train data.

This is also reflected in the mean error – which is ~9% higher on the test data than the train data, as well as in the min error, that is lower in the train then in the test.