Assignment 4

Calvin Tran

All code for this assignment can be found in the file named Assignment\_4.ipynb found with this assignment.

# Exercise 1

The implementation of the logistic regression algorithm can be found in Assignment\_4.ipynb. A step size of 0.1 was used and a tolerance of 10e-7 was used for convergence criteria. With these values, the GD loop converged after 93 epochs.

The plot for test error and training error vs the number of epochs can be found below:

A screen shot of a graph

Description automatically generated

The plot for test loss and training loss vs the number of epochs can be found below:

A graph on a white surface

Description automatically generated

The final test error using logistic regression is 46.34%. Using the sklearn implementations of the other classifiers, we can quickly determine the test errors of the other classifiers:

* MED: 46.34%
* GED: 46.34%
* kNN: 52.20%

As seen above, the test errors for each classifier is almost exactly the same, with the exception of the kNN classifier, performing worse than the other two. The poor performance of all of these classifiers with respect to the test error can be explained by a couple of reasons. For the MED and GED classifiers, they are most effective when the classes are well-separated. In this case, there may be a lot of input points that seem similar, especially after transforming the 784x1 vectors to 2x1 vectors. The same applies to the kNN and logistic regression classifier, where many of the inputs may seem to overlap after transforming the data.