CS 4361/5361 Machine Learning
Fall 2020
Exercise 1 – Classifying the MNIST Dataset

In class, we discussed how we could design a digit classifier and evaluate its accuracy. In this exercise, we will implement two of the algorithms we considered.

First, we will split the data into a training set and a test set. We will use 90% of the data for training and 10% for testing.

Then we will implement the following 2 algorithms and evaluate their accuracies and running times:

Algorithm 1
1. Find the mean image in each of the classes m_0, m_1, …, m_9.
2. For every image x_i in the test set, find its Euclidean distance to each of the means, and assign it to the class of the closest mean.

Algorithm 2
1. For every image x_i in the test set, find its Euclidean distance to each image in the training set and assign it to the class of the closest image.