

# CS4361/5361 Machine Learning

Fall 2020

## Exercise PCA and k-NN

1. Experiment with PCA preprocessing to determine the minimum number of principal components required to classify the MNIST dataset using the *sklearn* implementation of k-nearest neighbors with:
  - a. 90% accuracy
  - b. 95% accuracy
  - c. 97% accuracy

Warnings/hints:

- Do not evaluate accuracy on training set!
- The default parameters for training are not necessarily the best
- Remove columns with zero variance prior to computing PCA