GEORGIA INSTITUTE OF TECHNOLOGY
SCHOOL of ELECTRICAL and COMPUTER ENGINEERING

ECE 6254 Fall 2022
Project #4

Assigned: 20 Oct
Due Date: 27 Oct

Please contact the TAs for clarification on the instructions in the homework assignments.

PROBLEM 4.1:
In this problem we will explore nearest neighbor classification in Python.

The file `knn-example.py` provides a good start at this. You should be able to run this in the iPython environment simply by typing `run knn-example.py`. This uses the NumPy, Matplotlib, and scikit learn python packages. These should come included in the standard Anaconda distribution, but if you don't have them you will need to install them first.

The file begins by loading the appropriate packages and fixing the random seed so that your results will be repeatable. It then generates a simple 2-dimensional dataset with n datapoints from two possible classes. Next it builds a k-nearest neighbor classifier. Finally, it plots the results. Before going further, spend some time with this and try to understand what the code is doing.

In this problem I would like you to design a k-nearest neighbor classifier for several different values of n. In particular, I would like you to consider $n = 100, 500, 1000, 5000$. For each of these values of n, experiment with different choices of k and decide what the "best" choice of k is for each of these values of n (either based on the visual results, or using some quantitative method of your own devising). Provide a table showing your choices of k, and include a plot of the resulting classifier for each value of n.