CSCI 3302: Introduction to Robotics
Homework 5: Clustering and Classification

Due Date: Dec 7, 2019 @ 5:00pm

Your robot's LIDAR has returned a number of readings in the area ahead. In this assignment you will be implementing the machine learning tools necessary to 1) cluster unlabeled sensor observations to figure out which objects they might belong to (k-Means); and 2) to classify new readings according to existing "nearby" labeled data (K-Nearest Neighbor).

Using the provided base code, implement the functionality required for K-Means clustering (50%) and Weighted K-Nearest Neighbor classification (50%). You may use numpy or any math library you prefer, though this is not strictly necessary. Your algorithm does not need to use any sophisticated data structures (e.g., kd-trees) for nearest neighbor calculations: you are not being evaluated based on algorithm efficiency, though it shouldn't take more than a few minutes to run in the very worst case.

Caveat: You are not permitted to call k-Means or k-NN classifiers from other packages to implement your own (i.e., you may not just write a wrapper that calls sklearn's implementations).

The provided Python file will automatically output your k-Means cluster centers and assess your kNN classifier accuracy using Leave-One-Out-Cross Validation (https://en.wikipedia.org/wiki/Cross-validation_(statistics)#Leave-one-out_cross-validation).

Performance Characteristics:

Your kNN solution should get an accuracy above **60% for k=4**, while your visualization of k-Means should reasonably separate the four clusters into the original distributions that the points were sampled from. (But make sure that you don't hardcode for a particular value of **k** in your implementation!)

You are to complete this assignment on your own (without collaboration).

Anticipated time for completion: 120 minutes.
If you find yourself spending more than an hour for either part of this assignment, please reach out on Piazza and the course staff will be happy to help out with Python syntax/etc..

To Submit: Upload your fully implemented Homework5.py file, as well as the hw5_kmeans_*.pkl file containing your cluster centers to Canvas for full credit.