CS 4361/5361 Machine Learning

Fall 2020

Final Exam Part 2

1. (50 points) Write a program to compare the accuracy of decision tree, k-nearest neighbor, and support vector classifier to classify the data contained in files *q1_X.npy* and *q1_y.npy*. You can use the *sklearn* implementations of the algorithms with default parameters. See the expected output for guidance.

2. (50 points) Write a program to compare the accuracy of multi-layer perceptron, k-nearest neighbor, and support vector classifier to classify the 28-by-28 images in the dataset in files *q2_X_train.npy*, *q2_X_test.npy*, *q2_y_train.npy*, and *q2_y_test.npy*. Perform the experiments listed below and look at the expected results for guidance.

   (a) Use only the original data to train.

   (b) Augment the training data by adding a mirror image of each image in the original training set.

   (c) Compress the data from the previous experiment using Principal Component Analysis, with 20 components.