

August-December 2015 Semester

CS669: Pattern Recognition

Programming Assignment 2

Date: 22nd September, 2015

Datasets:

Dataset 1: 2-dimensional artificial data of 3 or 4 classes: Nonlinearly separable data set

Dataset 2: Real world data set:

- (a) Two dimensional speech dataset (used in Assignment 1)
- (b) Image dataset

Data of each class is given separately. For all datasets, 75% of data of a class is to be used as training data for that class, and the remaining data is to be used as test data for that class.

Note: Each batch of students must use the datasets identified for that batch

Classifiers to be built:

Bayes classifier using GMM on Dataset-1 and Dataset-2. GMM is built using the K-means clustering to initialize the parameters.

Perform the experiments on different number of mixtures of GMM

Report should include the results of studies presented in the following forms for each classifier and for each dataset:

1. Classification accuracy on test data
2. Confusion matrix based on the performance for test data
3. Decision region plot with the training data superposed (**only for Dataset-1 and Dataset 2(a)**)

Report should also include your observations about the performance. It should also include the observation on the nature of decision surface obtained for Dataset-1 and Dataset 2(a) in comparison with that of Assignment-1.

Deadline for submission: 04.00PM, Wednesday, 30th September 2015