August-December 2015 Semester
CS669: Pattern Recognition
Programming Assignment 4

Date: 02nd November, 2015

Datasets:

Dataset 1: 2-dimensional artificial data of 3 or 4 classes:
- (a) Linearly separable dataset
- (b) Nonlinearly separable data set **(Only for SVM)**

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 data in Dataset, 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: (1) Each batch of students must use the datasets identified for that batch
(**2) Concatenate the local feature vectors in each image to form a supervector. Now each image will become a vector with 36*23-dimension.**

Classifiers to be built:

1. Bayes classifier using unimodal Gaussian distribution on Dataset-1(a) and Dataset-2.
2. Bayes classifier using GMM on Dataset-2. GMM is built using the K-means clustering to initialize the parameters.
3. Build Bayes classifier using unimodal Gaussian distribution on the 1-dimensional representation of Dataset-1(a) and Dataset-2(a) obtaind using PCA.
4. Build Bayes classifier using unimodal Gaussian distribution and GMM on the reduced dimensional representations of Dataset-2 obtaind using PCA.
5. Fisher linear discriminant analysis (FDA) based classifier on Dataset-1 and Dataset-2. Use both Bayes classifier using unomodal Gaussian and GMM.
6. Perceptron-based classifier on Dataset-1.
7. SVM-based classifier using (a) linear kernel, (b) polynomial kernel and (c) Gaussin/RBF kernel on Dataset-1 and Dataset-2
8. SVM-based classifier using (a) linear kernel, (b) polynomial kernel and (c) Gaussin/RBF kernel on the reduced dimensional representations of Dataset-2 obtaind using PCA.

Perform the experiments on different values of *K* in GMM for different reduced dimensions. Also perform experiments on different values of SVM and kernel parameters for different reduced dimensions.

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. Observation on the nature of decision surface obtained for Dataset-1 and Dataset 2(a) **for SVM** in comparision with that of Assignment-1
4. Plot of 1-dimensional and 2-dimensional reduced dimensional representations.
5. Comparision with all the classifiers

Report should also include your observations about the performance.

Deadline for submission: 04.00PM, Sunday, 15th November 2015