August-December 2018 Semester
CS669: Pattern Recognition
Programming Assignment 1

Date: 29th August, 2018

Datasets:
        Dataset 1: 2-dimensional artificial data of 3 or 4 classes:
                (a) Linearly separable data set
                (b) Nonlinearly separable data set
        Dataset 2: Real world data set

Data of each class is given separately. For all 2-dimensional artificial and real data in Dataset-1 and Dataset-2, 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

Assumption: Class-conditional densities are Gaussian

Classifiers to be built for each dataset:
        1. Bayes classifier
                (a) Covariance matrix for all the classes is the same and is $\sigma^2\mathbf{I}$
                    You can obtain the same Covariance matrix for all the classes by taking the average of Covariance matrices of all the classes. You can obtain the same variance by averaging all the variences.
                (b) Full Covariance matrix for all the classes is the same and is $\mathbf{\Sigma}$.
                    You can obtain the same Covariance matrix for all the classes by taking the average of Covariance matrices of all the classes.
                (c) Covariance matric is diagonal and is different for each class
                (d) Full Covariance matrix for each class is different

Report should include the results of studies presented in the following forms for each classifier and for each dataset:
For the Dataset-1 and Dataset-2:
        1. Classification accuracy, precision for every class, mean precision, recall for every class, mean recall, F-measure for every class and mean F-measure on test data
        2. Confusion matrix based on the performance for test data
        3. Constant density contour plot for all the classes together with the training data superposed
        4. Decision region plot for every pair of classes together with the training data superposed
        5. Decision region plot for all the classes together with the training data superposed

Report should also include your observations about the performance and the nature of decision surface for each classifier, and for each dataset.

Submit your code and report strictly in PDF form as one zip file in moodle. Name the zip file as Group<num>_Assignment1.zip. E.g. Group01_Assignment1.zip

Deadline for submission: 4.00PM, Sunday, 16th September 2018