CS669: Pattern Recognition
Programming Assignment 2
Date: 26th August, 2013

Datasets:

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

75% of total 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 C
   (b) Covariance matrix for each class is different
2. Naive-Bayes classifier
   (a) Covariance matrix for all the classes is the same and is $\sigma^2 I$
   (b) Covariance matrix for all the classes is the same and is the C
   (c) 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:

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

Report should also include your observations about the performance and the nature of decision surface for each classifier, and for each dataset.
Deadline for submission of report: 4.00PM, Saturday, 14th September 2013