Project 1 - Two Category Classification Using Baysian Decision Rule (Due 02/06)

Matlab Code for Plotting:

- plotsynth.m
- twomodal.m

Objectives:

Design a decision rule on a synthetic data set with two categories. Assume the probability density is Gaussian.

Data set used:

Download *synth.tr* (the training set) and *synth.te* (the test set) from Ripley's Pattern Recognition and Neural Networks (link provided on the course website).

(80) Basic requirements:

Use synth.tr to train your decision rule, and use synth.te to test the decision rule.

- (10/5) Use maximum likelihood estimation to estimate the parameters of Gaussian
- (45/30) Use discriminant function (try all three cases; note that Case III is actually the MAP method) to derive your decision rule. Illustrate the three decision rules (i.e., decision boundaries) as well as the sample locations in the same graph and comment on the difference.
- (10/10) Try different prior probability distributions and evaluate the performance.
- (15/15) Evaluate the performance of your decision rule extensively. Some methods include calculation and comparison of the classification accuracy of applying different decision rules on the testing set.

- (+15/20) Use two-modal Gaussian to model the data set and compare the performance with that using the one-modal.
- (+15/+15) Use Baysian learning to estimate the parameters of Gaussian and compare performance.

(20) Report

Each project requires a formal and comprehensive report. Reporting is especially important to graduate students. Here's a suggested outline for your reference.