Project 1 - Two Category Classification Using Baysian Decision Rule (Due 09/11)

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.

- (5) Use maximum likelihood estimation to estimate the parameters of the Gaussian
- (30) Use MAP to derive your decision rules (try all three cases). Illustrate the three decision rules as well as the sample locations (use different symbols for different categories) on the same graph. Comment on the difference.
- (10) Try different prior probability distributions and evaluate the performance. Use classification accuracy as the performance metric.
- (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.
- (20) Use two-modal Gaussian to model the data set and compare the performance with that using the one-modal.

(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.