Machine Learning & Machine Learning (extended)

(Practice Exercise)

This is ungraded exercise which you're strongly encouraged to attempt and solve, which will help your understanding of relationship (dependence) between attributes, their effect on modelling the decision boundary, and Bayesian classification. These exercises will be solved during next Friday's lecture (i.e. 11th Nov 2016).

Bayesian Classification and the Role of Covariance

Question: What is the effect of covariance between attributes and the resultant shape modelled for the point cloud (i.e. the training samples within a class)? You can study the effect under various scenarios: (i) increase in covariance, (ii) decrease in covariance, (iii) positive covariance, (iv) negative covariance, (v) zero covariance, (vi) identical covariance for all classes, (vii) arbitrary (i.e. non-identical) covariance for different classes, etc. You may find the following MATLAB's functions useful to determine the covariance between attributes and studying the effect of covariance on the shape of point cloud: cov, mvnrnd, mvnpdf.

Question: Consider a 2 class Bayesian classifier which is trained from the past examples to predict the target labels by computing the posterior estimates. However, this classifier can be considered as a model that draws a decision boundary (linear or non-linear) between the classes such that this boundary separates the training samples. Answer the following questions about the Gaussian classifier:

(i) Is it possible for a Gaussian classifier to implement a non-linear decision boundary? If so, draw an example and suggest the shape of this non-linear decision boundary. If not, explain why not.

(ii) How about a Gaussian Naive Bayes classifier? Justify your answer.

Hints:

With multivariate Gaussian, the following cases can be considered:

a) statistically independent attributes, identically distributed Gaussian for each class (i.e. same variance for each class, and 0 entries at the non-diagonal location in the covariance matrix)

b) identical covariance for each class (i.e. $\Sigma = \Sigma_1 = \Sigma_2$)

c) arbitrary (non-identical) covariance for each class (i.e. $\Sigma_1 \neq \Sigma_2$)

There are two ways to attempt this problem.

In an informal way, you can attempt this problem by considering the shape of each class (for example by drawing the density contours around training samples) under the influence of covariance matrix.

In a more formal (and mathematical way), you can consider the maximum a posteriori estimate and determining the shape of the decision boundary in each of the above cases.