EE E6887 Statistical Pattern Recognition

Homework #4

Due Date: Oct. 19th 2005 Wed. 1pm

Please complete both problems.

P.1 (Nearest Neighbor)

In this problem, we would like to get familiar with the procedure of computing the error probability of 1-nearest neighbor. Consider data samples from the following two distributions. Assume the two classes have equal priors, i.e., $P(\omega_1) = P(\omega_2) = 0.5$

$$p(x \mid \omega_1) = \begin{cases} 2x & \text{for } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad p(x \mid \omega_2) = \begin{cases} 2(1-x) & \text{for } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) derive the Bayesian Decision Rule and its probability of classification.

(b) suppose we have one single training sample from class ω_1 and one single training sample from class ω_2. Now given a randomly selected test sample, we would like to use 1-nearest neighbor classifier to classify the test data. What's the probability of classification error of such 1-NN classifier?

P.2 (Distance Metrics)

Computing distances in a high-dimensional feature space sometimes could be cost prohibitive. One popular trick is to compute a certain distance in a lower dimension space as a pre-filtering step.

Assume $\vec{x} = \{x_1, x_2, \ldots, x_d\}$ and $\vec{y} = \{y_1, y_2, \ldots, y_d\}$ are two feature vectors in a d-dimensional space. Prove that

$$\left(\frac{1}{\sqrt{d}} \sum_{i=1}^{d} x_i - \frac{1}{\sqrt{d}} \sum_{i=1}^{d} y_i \right)^2 \le \sum_{i=1}^{d} (x_i - y_i)^2 .$$

Namely, the distance between the scaled means of two vectors is less than their L_2 distance. Discuss how we may use this property to reduce the computational complexity of the process of finding the nearest neighbor point.