ComS 573: Machine Learning
Spring 2012

Homework 2
Due Friday, January 27, 2012 in class

Note: Please do not hesitate to contact the instructor or TA if you have difficulty understanding or getting started with solving any of the problems.

1. (20 pts.) Consider a two-category classification problem with one-dimensional feature x. Suppose that the class-conditional probability densities are uniform distribution

$$p(x|\omega_1) \sim U(0,1) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

$$p(x|\omega_2) \sim U(0,2) = \begin{cases} 1/2 & 0 \le x \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

Assume that we observe a feature value $0 \le x \le 2$. Give the Bayes decision rule for the following two cases

(a) $P(\omega_1) = P(\omega_2) = 1/2$

(b) $P(\omega_1) = 1/4$, $P(\omega_2) = 3/4$

2. (20 pts.) In many pattern classification problems one has the option either to assign the pattern to one of c classes, or to *reject* it as being unrecognizable. Let α_i, for $i = 1, \dots, c$, be the action "decide ω_i" and α_{c+1} be the action "reject". If the cost for rejects is not too high, rejection may be a desirable action. Let the loss function be

$$\lambda(\alpha_i|\omega_j) = \begin{cases} 0 & i = j \quad i, j = 1, \dots, c \\ \lambda_r & i = c+1 \\ \lambda_s & \text{otherwise} \end{cases}$$

where λ_r is the loss incurred for choosing the $(c+1)$th action, rejection, and λ_s is the loss incurred for making any substitution error. Recall that the conditional risk is given by

$$R(\alpha_i|\vec{x}) = \sum_{j=1}^{c} \lambda(\alpha_i|\omega_j) P(\omega_j|\vec{x}).$$

Show that the minimum risk is obtained if we decide ω_i if $P(\omega_i|\vec{x}) \ge P(\omega_j|\vec{x})$ for all j and if $P(\omega_i|\vec{x}) \ge 1 - \lambda_r/\lambda_s$, and reject otherwise. What happens if $\lambda_r = 0$? What happens if $\lambda_r > \lambda_s$?

3. (20 pts.) Let the components of the feature vector $\vec{x} = (x_1, \ldots, x_d)^t$ be binary-valued (0 or 1), and let $P(\omega_j)$ be the prior probability for the class ω_j and $j = 1, \ldots, c$. Now define

$$p_{ij} = P(x_i = 1|\omega_j) \quad i = 1, \ldots, d, \quad j = 1, \ldots, c,$$

with the components of $\vec{x}$ being statistically independent given the category (i.e., $P(\vec{x}|\omega_j) = \prod_{i=1}^{d} P(x_i|\omega_j)$). Show that the minimum probability of error is achieved by the following decision rule: Decide ω_k if $g_k(\vec{x}) \geq g_j(\vec{x})$ for all j and k, where the discriminant function is given by

$$g_j(\vec{x}) = \sum_{i=1}^{d} x_i \ln \frac{p_{ij}}{1 - p_{ij}} + \sum_{i=1}^{d} \ln(1 - p_{ij}) + \ln P(\omega_j).$$

(The discriminant functions for Binary Independence Models.)