CSE 566 Midterm 1

Olcay Taner YILDIZ

I. Question

In K-fold cross-validation, the dataset X is divided randomly into K equal sized parts, $X_i, i = 1, 2, \ldots, K$. To generate each train and test set pair, we keep one of the K parts out as the test set, and combine the remaining $K - 1$ parts to form the training set. On the other hand, classes must be represented in the right proportions not to disturb the class prior probabilities (**stratification**). Propose an algorithm to do K-fold cross-validation with stratification.

II. Question

Let x have an exponential density, for all x≥0

$$p(x|\theta) = \theta e^{-\theta x}, \tag{1}$$

Suppose that n samples $x_1, x_2, \ldots, x_n$ are drawn independently according to $p(x|\theta)$. Show that the maximum likelihood estimate for θ is given by

$$\hat{\theta} = \frac{1}{\frac{1}{n}\sum_{k=1}^{n} x_k} \tag{2}$$

III. Question

Suppose that the data mining task is to cluster the following eight points (with (x, y) representing location) into three clusters. A_1(2, 10), A_2(2, 5), A_3(8,4), A_4(5, 8), A_5(7,5), A_6(6,4), A_7(1,2), A_8(4,9). The distance function is Euclidian distance. Suppose initially we assign A_1, A_4, and A_7 as the center of each cluster, respectively. Run the k-means algorithm one epoch on this data.

IV. Question

Let say we have two variables x_1 and x_2 and we want to make a quadratic fit using them, namely

$$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1 x_2 + w_4 (x_1)^2 + w_5 (x_2)^2$$

Given sample data X, how can we find optimum w_i's?

V. Question

Give an example feature selection run on 5 feature dataset, where the best feature subset found by Forward Selection and Backward Elimination is $\{F_2, F_3, F_5\}$. On the other hand, arrange the error rates of the Feature subsets so that the global optimum is not $\{F_2, F_3, F_5\}$.

VI. Question

In many pattern classification problems one has the option either to assign the pattern to one of K classes, or to reject it as being unrecognizable. If the cost for rejects is not too high, rejection may be a desirable action. Let

$$\begin{aligned} \lambda(\alpha_i|C_j) &= 0 \quad i = j \\ \lambda(\alpha_i|C_j) &= \lambda_r \quad i = K + 1 \\ \lambda(\alpha_i|C_j) &= \lambda_s \quad otherwise \end{aligned}$$

where λ_r is the loss incurred for choosing the (K + 1) action, rejection, and λ_s is the loss incurred for making any error. Show that the minimum risk is obtained if we decide C_i if $P(C_i|x) \leq P(C_j|x)$ for all j and if $P(C_i|x) \leq 1 - \frac{\lambda_r}{\lambda_s}$, and reject otherwise.