Assignment 6

Problem 1

In many classification problems, one has the option either of assigning x to class j or, if you are too uncertain, of choosing a "reject" option. If the cost for rejects is less than the cost of falsely classifying the object, it may be the optimal action. Let α_i mean you choose action i, for $i = 1 : C+1$, where C is the number of classes and $C+1$ is the reject action. Let $Y = j$ be the true (but unknown) state of nature. Define the loss function as follows

$$\lambda(\alpha_i|Y=j) = \begin{cases} 0 & \text{if } i = j \text{ and } i, j \in \{1, \ldots, C\} \\ \lambda_r & \text{if } i = C+1 \\ \lambda_s, & \text{otherwise} \end{cases}$$

In other words, you incur 0 loss if you correctly classify, you incur λ_r loss (cost) if you choose the reject option, and you incur λ_s loss (cost) if you make a substitution error (misclassification). Both λ_r and λ_s are positive numbers.

1. Show that when we decide to choose a class (and not reject), we always pick the most probable one.
2. Show that the minimum risk is obtained if we decide to pick the most probable class $j_{max} = \arg\max_j p(Y = j|x)$ and if $p(Y = j_{max}|x) \geq 1 - \frac{\lambda_r}{\lambda_s}$; otherwise we decide to reject.
3. Describe qualitatively what happens as λ_r/λ_s is increased from 0 to 1 (i.e., the relative cost of rejection increases).

Problem 2: Mixture of multivariate Bernoullis EM

Consider a unsupervised mixture of multivariate Bernoullis model. For each data example n, we have D binary values x_{nj}, we assume there is an unobserved cluster indicator $y_n \in \{1 \ldots K\}$, and that $P(x_{nj} \sim \text{Ber}(\mu_{y_n,j}))$. In this problem, you will show how to fit this model using EM.

1. Using the definition of arbitrary distributions $q_n(y_n)$, write the evidence lower bound $Q(\theta, \{q_n\})$. Simplify as much as you can.
2. Using Bayes rule, write an expression that can be used to calculate the cluster responsibility $r_{nk} = P(y_n = k|\theta)$
3. Show that the M step for ML estimation of a mixture of multivariate Bernoullis is given by

$$\mu_{kj} = \frac{\sum_n r_{nk} x_{nj}}{\sum_n r_{nk}}$$

Assignment 6

4. Show that the M step for MAP estimation of a mixture of multivariate Bernoullis with a $Beta(\alpha, \beta)$ prior is given by

$$\mu_{kj} = \frac{(\sum_n r_{nk} x_{nj}) + \alpha - 1}{(\sum_n r_{nk}) + \alpha + \beta - 2}$$