ComS573 Machine Learning-Problem Set 1 Solution

Problem 1(a)

From the assumption:
$P(\omega_1) = \frac{1}{3}, P(\omega_2) = \frac{2}{3}$
$P(X, \omega_1) \sim N(1,1), P(X, \omega_2) \sim N(3,1)$

For minimum-error-rate classification, in order to decide ω_1,we need:
$P(\omega_1|x) > P(\omega_2|x)$

Based on Bayes rule, we have:
$P(X|\omega_1)P(\omega_1) > P(X|\omega_2)P(\omega_2)$

Based on assumption:
$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-1)^2} \times \frac{1}{3} > \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-3)^2} \times \frac{2}{3}$
$e^{-\frac{1}{2}(x-1)^2} > 2e^{-\frac{1}{2}(x-3)^2}$
$-\frac{1}{2}(x-1)^2 > ln(2) - \frac{1}{2}(x-3)^2$
$(x-3)^2 - (x-1)^2 > 2ln(2)$
$x < 2 - \frac{1}{2}ln(2)$
decide ω_1 if $x < 2 - \frac{1}{2}ln(2)$; otherwise decide ω_2.

Problem 1(b)
From the definition of λij be the loss incurred for deciding ω_i when the true category is ω_j. Such that λ11 = 0; λ22 = 0, which means there will be no loss if the decision match with the true category. So, we can define the action set is α1; α2, which αi means decide ω_i.
From the definition of conditional risk function:
$R(\alpha_i|x) = \sum_{j=1}^{c} \lambda(\alpha i \mid \omega j)\, P(\omega_1|x)$
to decide ω1, the risk function is:

$R(\alpha_1|x) < R(\alpha_2|x)$

which is equivalent rule, the formula is:

$\lambda_{11}P(\omega_1|x) + \lambda_{12}P(\omega_2|x) < \lambda_{21}P(\omega_1|x) + \lambda_{22}P(\omega_2|x)$
Based on Bayes rule, and multiply p(x) on both side:

$$\frac{P(x|\omega_1)}{P(x|\omega_2)} > \frac{\lambda_{12} - \lambda_{22}}{\lambda_{21} - \lambda_{11}} \times \frac{P(\omega_2)}{P(\omega_1)}$$

From the assumption:

$P(\omega_1) = \frac{1}{3}$, $P(\omega_2) = \frac{2}{3}$
$P(x, \omega_1) \sim N(1,1), P(x, \omega_2) \sim N(3,1)$
$\lambda_{11} = 0, \lambda_{12} = 2, \lambda_{21} = 1, \lambda_{22} = 0$
Such that, we can calculate this as:

$$\frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-1)^2}}{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-3)^2}} > \frac{2-0}{1-0} \times \frac{\frac{2}{3}}{\frac{1}{3}}$$

$$e^{-\frac{1}{2}(x-1)^2 + \frac{1}{2}(x-3)^2} > 4$$

$$x < 2 - \frac{ln4}{2}$$

Decide ω_1 if $x < 2 - \ ln2$; otherwise decide ω_2.

Problem 2

The minimum probability of error is achieved by the discriminant function, (from the lecture) we use the form: $g_j(x) = \ln P(x|\omega_j) + \ln P(\omega_j)$. In this case, as $\vec{x}$ is a vector of binary features, and each element of $\vec{x}$, xi is 1 if appears, and 0 if not. From the assmption, the estimate probability from training date, we have for p_{ij} :

$$P(x_i = 1|\omega_j)P(\omega_j) = p_{ij}$$
$$P(x_i = 0|\omega_j)P(\omega_j) = 1 - p_{ij}$$

such that $P(\vec{x}|\omega_j)$ can be simplify like this:

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

$= \sum_{i=1}^{d} x_i \ln P_{ij} + \sum_{i=1}^{d}(1 - x_i)\ ln\ (1 - P_{ij}) + \ln P(\omega_j)$.

$$= \sum_{i=1}^{d} \left(x_i \ln(P_{ij}) + (1 - x_i)\ln(1 - P_{ij})\right) + \ln P(\omega_j)$$

$$= \sum_{i=1}^{d} \left(x_i \ln \frac{P_{ij}}{1 - P_{ij}}\right) + \sum_{i=1}^{d} \left(\ln(1 - P_{ij})\right) + \ln P(\omega_j)$$

$= \sum_{i=1}^{d}(\text{ Cij } x_i + bj)$ where

$C_{ij} = \ln\frac{P_{ij}}{1-P_{ij}}$, $b_j = \sum_{i=1}^{d}\left(\ln(1-P_{ij})\right) + \ln P(\omega_j)$