EE 626 : Pattern Recognition and Machine Learning

Mid-Semester Examination

Duration: 2 hours **Date: Feb 29' 2024** **Marks: 35**

Please outline all the steps systematically to get full marks.

1. Perform a DTW match between the temporal sequence of vectors

$$\mathbf{X} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \text{ and}$$

$$\mathbf{Y} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

   (a) Fill up the elements of the cost matrix of size 6×4 by using the Euclidean distance.

   (b) Fill up the elements of the accumulator matrix of size 6×4 by using the cost matrix in part (a).

   (c) Determine the DTW cost between **X** and **Y**.

   (d) Write down the indices of the warping path obtained.

[2 + 3 + 1 + 2 = 8 marks]

2. (a) Suppose that we are given a sequence of training samples $\{x_1, x_2.....x_N\}$, that are generated from a univariate normal distribution $N(\mu, 1)$. It is known that our prior knowledge on the mean μ follows a normal distribution $N(3, 2)$. Compute the MAP estimate of the mean μ. The normal distribution $N(\mu, \sigma^2)$ is given by

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

   (b) Consider 2 matrices $\mathbf{A} = \mathbf{X}\mathbf{X}^\mathbf{T}$ and $\mathbf{B} = \mathbf{X}^\mathbf{T}\mathbf{X}$, where **X** is of size $m \times n$ ($m \geq n$). Derive a relationship between the eigenvalues and eigenvectors of **A** and **B**.

[4 + 2.5 =6.5 marks]

3. (a) Consider a training set with 100 positive and 150 negative samples. Based on the decision based on values of an attribute, we split it to 3 nodes with details as:

   - Node 1 contains 10 positive and 30 negative samples
   - Node 2 contains 10 positive and 100 negative samples

- Node 3 contains 80 positive and 20 negative samples

Compute the information gain resulting from the split

(b) Given the 2-dimensional data for two classes ω_1 and ω_2 :

$$\omega_1 : \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \end{bmatrix} \right\} \qquad \omega_2 : \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Compute the optimal weight vector **w** (projection line) in a reduced dimension using the Fisher criterion.

[4 + 4 = 8 marks]

4. (a) In a 3 class, 2 dimensional problem, the feature vectors for class ω_1 , ω_2 and ω_3 are normally distributed with covariance matrices : $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ respectively. The respective mean vectors μ_1 , μ_2 and μ_3 (for class ω_1 , ω_2 and ω_3) are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$. Assuming that the three classes are equi probable classify the feature vector $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$, according to the Bayes minimum probability classifier. Specify the maximum posterior probability value. The normal distribution $N(\mu, \Sigma)$ in d-dimensions is given by :

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{0.5}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^T\Sigma^{-1}(\mathbf{x}-\mu)}$$

[4 marks]

(b) Consider a one-dimensional two-category classification problem with prior $P(\omega_1) =$ 0.6 and continuous class conditional densities of the form

$$p(x|\omega_i) = \frac{x}{\theta_i^2} e^{-\frac{x^2}{2\theta_i^2}} \qquad x \geq 0$$

The parameters θ_i for $i = 1, 2$ are positive but unknown. In addition, the misclassification of samples of ω_1 to ω_2 ($P(\omega_2|\omega_1)$) incurs a loss twice to that of ω_2 to ω_1 ($P(\omega_1|\omega_2)$). Correct classifications are assigned zero loss.

(i) The following i.i.d training observations were collected: $\mathbb{D}_1 =\{1, 2, 5\}$ and $\mathbb{D}_2 =$ $\{3, 6, 7\}$ for ω_1 and ω_2, respectively. Using this information, find the maximum-likelihood values for θ_1 and θ_2.

(ii) Based on your answer to part (i), determine the Bayes threshold (decision boundary) x^*.

[4 + 4.5 = 8.5 marks]