COURSE CODE (CREDITS): 19B1WEC732 (3) MAX. MARKS: 25

COURSE NAME: Pattern Analysis in Machine Intelligence

COURSE INSTRUCTORS: Dr. Alok Kumar MAX. TIME: 1 Hour 30 Minutes

Note: (a)All questions are compulsory.

(b)Marks are indicated against each question in square brackets.

(c) The candidate is allowed to make Suitable numeric assumptions wherever required for

solving problems

Q. 1 How Bayes decision rule is helpful in pattern recognition. Derive the expression for Bayes decision for a vector X having n features. Show the probability of error occurs in the Bayes decision rule for feature vector Y, where feature vector Y belongs from binary class.

[CO3] **[4 Marks]**

Q.2 Consider the following datasets of the fruits which use features such as color, shape, and taste to predict the type of fruit. Apply the Naive Bayes algorithm to predict the type of fruit for a new observation with the following properties: Color = Orange, Shape = Round, Taste = Sweet.

[CO2, CO3] **[4 Marks]**

Color	Shape	Taste	Type
Red	Round	Sweet	Apple
Orange	Round	Sweet	Orange
Red	Oval	Sweet	Apple
Orange	Round	Tangy	Orange
Red	Round	Sweet	Apple
Orange	Oval	Tangy	Orange

Q.3 Consider the feature vectors X having two class (ω_1, and ω_2). Derive the expression for Risk Function. For the considered samples, how the decision comes in favor of either class ω_1 or class ω_2? [CO3, CO5] **[3 Marks]**

Q.4 What is discriminant function? What is the nature of discriminant function for different types of classifier? Derive the expression for discrimannat function for minimum error rate classifier. [CO5] **[3 Marks]**

Q.5 Explain the concept of feature importance in decision trees. How decision tree works in pattern recognition? Explicate the terms Pruning and Overfitting in the context of decision trees.

[CO2] **[3 Marks]**

Q.6 We have a dataset related to medical diagnosis to predict the diagnosis based on five features: Sore Threat, Fever, Swollen Glands, Congestion, Headache. Build a decision tree to predict which type of diagnosis will follow based on their features.   [CO1, CO2]   **[4 Marks]**

Sore Threat	Fever	Swollen Glands	Congestion	Headache	Diagnosis
Yes	Yes	Yes	Yes	Yes	Strep Throat
No	No	No	Yes	Yes	Allergy
Yes	Yes	No	Yes	No	Cold
Yes	No	Yes	No	No	Strep Throat
No	Yes	No	Yes	No	Cold
No	No	No	Yes	No	Allergy
No	No	Yes	No	No	Strep Throat
Yes	No	No	Yes	Yes	Allergy
No	Yes	No	Yes	Yes	Cold
yes	No	No	Yes	Yes	Cold

Q. 7. What do you mean by multivariate probability density function? Where it is used in pattern recognition? Assume a dataset of two features, X_1 and X_2, collected from a population and having the estimated mean (μ) and covariance matrix as follows (Σ):

$$\mu = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \qquad (\Sigma) = \begin{bmatrix} 2 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

For this given information answer the following questions.

a) Write down the expression for multivariable Gaussian probability density function (PDF)

b) Calculate the Mahalanobis distance for one of the generated samples

[CO3, CO5] **[4 Marks]**