Name : .....

Roll No. : .....

Invigilator's Signature : .....

CS/B.TECH (CSE)/SEM-8/CS-801F/2013

2013

PATTERN RECOGNITION

Time Allotted : 3 Hours Full Marks : 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

GROUP – A

(Multiple Choice Type Questions)

1. Choose the correct alternatives for the following : 10 × 1 = 10

i) Clustering algorithm usually employ

a) supervised learning

b) unsupervised learning

c) reinforcement learning

d) competitive learning.

ii) The likelihood of class w_1 and w_2 followed normal distribution $N(-0{\cdot}5, 2)$ and $N(0{\cdot}5, 2)$, respectively. For equal prior, a pattern $X = 1{\cdot}0$ belongs to

a) class w_1

b) class w_2

c) either class w_1 or class w_2

d) both the classes.

iii) If the covariance matrices for all of the classes are identical, then the discriminant functions will be

a) Linear b) Quadratic

c) Polynomial d) None of these.

iv) For uniform prior we can estimate the parameter of a density function by using

a) maximum likelihood (*ML*)

b) maximum a posteriority (*MAP*)

c) either *ML* or *MAP*

d) none of these.

v) *K*-Nearest Neighbor based classifier is

a) linear and optimal

b) linear and suboptimal

c) nonlinear and optimal

d) nonlinear and suboptimal.

vi) If P_{NN} is the classification error probability for the Nearest Neighbor rule and P_B is the Bayes error then

a) $P_B \leq P_{NN} \leq 2P_B$ b) $P_{NN} \geq P_{2B}$

c) $P_{NN} \leq P_{2B}$ d) $P_{NN} \leq P_B$..

vii) Gradient descent search is not applicable to find optima on a

a) rough surface

b) smooth surface

c) surface with single optima

d) surface with multiple optima.

viii) Perceptron is not able to implement

a) OR gate b) AND gate

c) XOR gate d) NOT gate.

ix) Given two fuzzy clusters A_1 and A_2. A data point X in two-class (fuzzy C-means clustering) then satisfies

a) $\mu_{A_1}(x) + \mu_{A_2}(x) = 1$

b) $\mu_{A_1}(x) + \mu_{A_2}(x) < 1$

c) $\mu_{A_1}(x) + \mu_{A_2}(x) > 1$

d) $\mu_{A_1}(x) + \mu_{A_2}(x) \leq 1$.

x) Principal component analysis is one important step in

a) Data dimension reduction

b) Data encryption

c) Noise filtering

d) Data communication.

GROUP – B

(Short Answer Type Questions)

Answer any *three* of the following. $3 \times 5 = 15$

2. Compare and contrast supervised and unsupervised learning.

3. Design a Bayes classifier in terms of a set of discriminant functions.

4. A sample from class-*A* is located at (*X*, *Y*, *Z*) = (1, 2, 3), a sample from class-*B* is at (7, 4, 5) and a sample from class-*C* is at (6, 2, 1). How would a sample at (3, 4, 5) be classified using the Nearest Neighbor technique and Euclidean distance ?

5. Write a short note on generalized linear discriminant function.

6. Consider the following proximity matrix :

$$P = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array} \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \begin{bmatrix} 0 & 6 & 8 & 2 & 7 \\ & 0 & 1 & 5 & 3 \\ & & 0 & 10 & 9 \\ & & & 0 & 4 \\ & & & & 0 \end{bmatrix} \end{array}$$

Draw the resulting dendrogram by applying single link clustering algorithm.

GROUP – C

(Long Answer Type Questions)

Answer any *three* of the following. 3 × 15 = 45

7. a) Describe the basic steps involved in the design of pattern recognition system.

b) What is maximum likelihood (ML) estimation ? Show that if the likelihood function is univariate Gaussian with unknowns the mean μ as well as variance σ^2, then ML estimate are given by

$$\mu = \frac{1}{N} \sum_{k=1}^{N} X_k \text{ , and } \sigma^2 = \frac{1}{N} \left(X_k - \mu \right)^2 ,$$

where X_k is the k^{th} pattern and N is the total number of training patterns.

c) Compare parametric and non-parametric technique.

6 + 5 + 4

8. a) What is Bayesian classifier ? Prove that it is an optimal classifier.

b) In a two class problem with single feature X the *pdf*'s are Gaussians with variance $\sigma^2 = \frac{1}{2}$ for both classes and mean value 0 and 1 respectively. If $P(w_1) = P(w_2) = \frac{1}{2}$, compute the threshold value X_0 for minimum error probability. 4 + 5 + 6

9. a) What is density estimation ? What are the necessary conditions for its convergence ?

b) Compare Parzen Windows and k-Nearest Neighbor density estimation technique.

c) What is perceptron ? Discuss briefly the perceptron based learning algorithm. 4 + 4 + 7

10. a) What is clustering ? Categorize the different clustering algorithms of the pattern recognition domain.

b) Explain Fuzzy-C-means clustering algorithm. Write a short note about its criterion function. 6 + 9

11. a) What is feature selection ? What is optimal and suboptimal feature subset selection ?

b) Explain one suboptimal feature subset selection technique.

c) What is feature generation ? Write a short note on principal component analysis. 4 + 5 + 6