*Name* : ...............................................................

*Roll No.* : ...........................................................

*Invigilator's Signature :* ........................................

**CS/B.Tech/IT/New/SEM-6/IT-604C/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 any *ten* of the following :

$10 \times 1 = 10$

   i)   The most commonly used measure of similarity is the ................. .. . or its square.

   a)   Euclidean distance

   b)   City-b ock distance

   c)   Chebyshev's distance

   d)   Manhattan distance.

   ii)  ………………………… is frequently referred to as *k*-means clustering.

   a)   Non-hierarchical clustering

   b)   Optimizing partitioning

   c)   Divisive clustering

   d)   Agglomerative clustering.

CS/B.Tech/IT/New/SEM-6/IT-604C/2013

iii) Example of non-parametric density estimation methods is

a) Nearest Neighbours

b) Parzen Window

c) both (a) and (b)

d) none of these.

iv) The expression $i(N) = -\sum_{i=1}^{n} p(w_j) \log_2 p(w_j)$ is

a) Entropy impurity

b) Variance impurity

c) Information g in

d) None of these.

v) In case of a good clustering.

a) Internal ( within the cluster ) distance should be small.

b) External ( intra-cluster ) should be large.

c) Both (a) and (b)

d) None of these.

vi) Why are linearly separable problems of interest to neural network researchers ?

a) Because they are the only class of problems that a network can solve successfully.

b) Because they are the only class of problems a Perceptron can solve successfully.

c) Because they are the only mathem tical functions that are continuous.

d) Because they are th only mathematical functions you can draw.

vii) Linear Discriminant function is given by

a) $W^T X + w_0$

b) $W^T X - w_0$

c) $W^T X^T + w_0$

d) $W^T X^T - w_0$.

viii) Full form of iid is

a) individually and identically distributed

b) individually and independently distributed

c) independently and identically distributed

d) independent and identical distribution.

CS/B.Tech/IT/New/SEM-6/IT-604C/2013

ix) In Fisher's linear discriminant classifier, ………………
objective function ( $J(W)$ ) is maximized.

a) $\dfrac{(m_1 - m_0)^2}{s_0{}^2 - s_1{}^2}$

b) $\dfrac{(m_1 + m_0)^2}{s_0{}^2 - s_1{}^2}$

c) $\dfrac{(m_1 + m_0)^2}{s_0{}^2 + s_1{}^2}$

d) $\dfrac{(m_1 - m_0)^2}{s_0{}^2 + s_1{}^2}$ .

x) Which one of the following clasifiers is used in non-metric methods ?

a) K-nearest Ne ghbuor     b) Parzen Window

c) Decision Tree          d) None of these.

xi) The most natural representation of hierarchical c ustering is a corresponding tree, called

a) mendrogram

b) cendrogram

c) dendrogram

d) lendrogram.

xii) The conditional probability is given by

   a)    $P ( A / B ) = P ( A \cup B ) \, P ( B )$

   b)    $P ( A / B ) = P ( A \cap B ) \, P ( B )$

   c)    $P ( A / B ) = P ( A ) \; P ( B )$

   d)    None of these.

## GROUP – B
### ( Short Answer Type Questions )
Answer any *three* of the following.     $3 \times 5 = 15$

2. Using Bayesian approach, evolve expressions for mean and variances of samples dr wn from a univariate normal distribution.

3. What is the fundamental difference between maximum likelihood parameter estimation and Bayesian parameter estimation ?

4. Discuss supervised and unsupervised learning.

5. Describe the basic modules in designing a pattern recognition system.

6. Write a short note with diagrams on Decision trees which are non-linear, non-metric classifiers.

CS/B.Tech/IT/New/SEM-6/IT-604C/2013

**GROUP – C**
**( Long Answer Type Questions )**
Answer any *three* of the following.      3 × 15 = 45

7.   a)   Explain decision tree.

   b)   Suppose there is a student that decides whether or not to go into campus on any given day based on the weather, wake-up time and whether there is a seminar talk he is interested in attending. There are data collected from 13 days.

| Day | Wake-up | Have Talk | Weather | Go To School |
|-----|---------|-----------|---------|--------------|
| D1 | Normal | No | Sunny | No |
| D2 | Normal | No | Rain | No |
| D3 | Early | No | Sunny | Yes |
| D4 | Late | No | Sunny | Yes |
| D5 | Late | Yes | Sunny | Yes |
| D6 | Late | Y s | Rain | No |
| D7 | Early | Yes | Rain | Yes |
| D8 | Normal | No | Sunny | No |
| D9 | Norma | Yes | Sunny | Yes |
| D10 | L te | Yes | Sunny | Yes |
| D11 | Normal | Yes | Rain | Yes |
| D12 | Early | No | Rain | Yes |
| D13 | Early | Yes | Sunny | Yes |

   i)   Build a decision tree based on these observations, using entropy impurity.

   ii)   Show your work and the resulting tree.

   c)   Classify the following sample using above decision tree : Wake-up = Late, Have Talk = No and Weather = Rain.      2 + 10 + 3

CS/B.Tech/IT/New/SEM-6/IT-604C/2013

8. a) Indicating the conditions for convergence, describe the Parzen window method of density estimation.

b) To which category of clustering schemes does the $k$-means algorithm belong ? What is its major advantage ? Which are the factors that influence the computational duration of this algorithm ?

c) Write Expectation Minimization ( EM ) algorithm.

6 + 5 + 4

9. State the Bayes Rule and explain how it is applied to pattern classification problems. Show that in a multiclass classification task the Bayes decision rule minimizes the error probability.

10. According to records of WBUT, the probability that a student will engineer within four years is 0.8. The WBJEE scores of those who engineer within four years are normally distributed with mean 26 and standard deviation 2. The scores of those who do not engineer within four years are also normally distributed with mean 22 and s. d. 3. If a student with WBJEE score of 22 is admitted, then how most likely will assign the student to be an engineer within four years ?

CS/B.Tech/IT/New/SEM-6/IT-604C/2013

11. Write short notes on any *three* of the following :    3 × 5 = 15

    a)    Support Vector Machine

    b)    Hierarchical Clustering

    c)    Fisher discriminant analysis

    d)    Gaussian Mixture Model

    e)    K-nearest neighbour classifier.

———————