

**PATTERN RECOGNITION
(CSEN 4244)**

Time Allotted : 3 hrs

Full Marks : 70

Figures out of the right margin indicate full marks.

Candidates are required to answer Group A and any 5 (five) from Group B to E, taking at least one from each group.

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

**Group – A
(Multiple Choice Type Questions)**

1. Choose the correct alternative for the following: **10 × 1 = 10**
- (i) Clustering algorithm usually employ
(a) supervised learning (b) unsupervised learning
(c) reinforcement learning (d) None of the above.
- (ii) Single Perceptron is not able to implement
(a) OR gate (b) AND gate (c) XOR gate (d) NOT gate.
- (iii) Principal component analysis is one important step in
(a) Data dimension reduction (b) Data encryption
(c) Noise filtering (d) Data communication.
- (iv) Fuzzy c-means algorithm can be categorised as _____ technique
(a) non-hierarchical clustering (b) linear regression
(c) agglomerative clustering (d) logistic regression.
- (v) In a k-class data set the expression : $-\sum_{j=0}^k [p(w_j) \log_2 p(w_j)]$ denotes
(a) Entropy impurity (b) Information gain
(c) GINI impurity index (d) None of these.
- (vi) The objective of a good clustering technique
(a) Inter-cluster distance should be minimized
(b) Intra-cluster distance should be minimized
(c) Neither option (a) nor (b)
(d) Both option (a) and (b).
- (vii) Which of the following option is true about k-NN algorithm?
(a) It can be used for classification
(b) It can be used for regression
(c) It can be used in both classification and regression
(d) None of these.

- (viii) Which of the following can act as the best possible termination condition in fuzzy c-Means clustering algorithm?
- (a) For a fixed number of iterations
 - (b) Until the cluster centroids do not change
 - (c) Centroids of each cluster changes alternatively between successive iterations
 - (d) Variance of each cluster changes frequently between successive iterations.
- (ix) Consider the following two statements:
Statement 1: Principal Component Analysis is an unsupervised method
Statement 2: It searches for the directions that data have the largest variance.
- (a) Only Statement 1 is correct
 - (b) Only Statement 2 is correct
 - (c) Both statements are incorrect
 - (d) Both statements are correct
- (x) Gradient descent algorithm is generally used
- (a) in reducing loss function, as used in data classification or regression
 - (b) when data points are hierarchically clustered
 - (c) when data points are non-hierarchically clustered
 - (d) None of these.

Group – B

2. (a) Describe the basic steps involved in the design of a pattern recognition system.
- (b) 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 Neighbour technique and Euclidean distance?

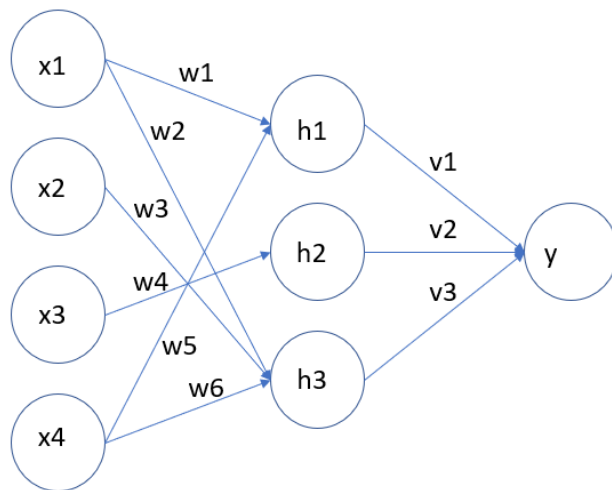
6 + 6 = 12

3. (a) What is feature selection?
- (b) Compare parametric and non parametric learning approach, in terms of their advantages and drawbacks.
- (c) What are the differences between metric and non-metric measures? Give one example of a measure which qualifies as a metric.

2 + 4 + (3 + 3) = 12

Group – C

4. (a) Why Multi Layered Perceptron model is called a feed forward network model? What is the importance of bias input connection in a single perceptron model?
- (b) Briefly explain how gradient descent algorithm can reduce the error in misclassification of data with the help of an example.
- (c) A scientist has designed 2-Layer Neural Network with restricted connections from one layer to another. The network has been shown in the below diagram, where $x_1 - x_4$ are the input nodes, h_1 are the hidden nodes and y is the output node. The weights between each layer are assigned randomly.



The weights assignment are as follows,

$w_1 = 0.3$, $w_2 = 0.5$, $w_3 = -0.41$, $w_4 = 0.23$, $w_5 = 0.81$, $w_6 = -0.92$

$v_1 = 0.11$, $v_2 = 0.28$, $v_3 = -0.35$

Using the concept of feed forward network calculate the output node y with input values as 1,0,2,1.

Also specify the final output if the network contains the following threshold,

$$\text{Output} = \begin{cases} 1 & \text{if } \sum w_i x_i > t \\ 0 & \text{otherwise} \end{cases}$$

$$(1 + 1) + 4 + 6 = 12$$

5. (a) What is information gain?

Discuss in detail the steps of creating the **root node** of a decision tree using the following credit card fraudulence data, where information about fraud and honest class of credit card users are provided:

User #	Country	Avg. delay of bill payment	User class
U1	Pakistan	>12 months	Fraud
U2	India	<6 months	Honest
U3	Bhutan	Between 6 to 12 months	Honest
U4	USA	>12 months	Fraud
U5	USA	Between 6 to 12 months	Honest
U6	USA	>12 months	Fraud
U7	Pakistan	>12 months	Fraud
U8	Pakistan	>12 months	Fraud
U9	Pakistan	Between 6 to 12 months	Honest
U10	Bhutan	<6 months	Honest
U11	Bhutan	<6 months	Honest
U12	Bhutan	<6 months	Honest
U13	India	>12 months	Fraud
U14	India	Between 6 to 12 months	Honest
U15	India	<6 months	Honest
U16	USA	>12 months	Fraud
U17	USA	<6 months	Honest
U18	India	>12 months	Fraud

- (b) Briefly explain Bayesian classification technique?
- (c) Discuss about decision boundaries in context of classification?

$$(1 + 5) + 4 + 2 = 12$$

Group – D

6. (a) What is the necessity of clustering technique? Write down the k-means algorithm? Mention any two challenging tasks while implementing k-means algorithm? What happens if you decide to have n number of cluster (having n data points). What is the preferred number of clusters used in general with respect to n data points?
- (b) You are provided with n = 6 objects and number of cluster = 2 along with the following data,

X	Y	C1	C2
1	6	0.8	0.2
2	5	0.9	0.1
3	8	0.7	0.3
4	4	0.3	0.7
5	7	0.5	0.5
6	9	0.2	0.8

What will be the updated cluster membership value after 1st iteration.

$$(1 + 2 + 2 + 1 + 1) + 5 = 12$$

7. (a) What is the different classification of data points in the db-scan technique? Explain briefly with an example. What is the advantage of db-scan over other classification techniques?
- (b) What is hierarchical clustering? Explain agglomerative clustering with the help of an example.

$$(3 + 2 + 2) + 5 = 12$$

Group – E

8. (a) What is dimension reduction of data set? Briefly discuss about the various approaches of dimension reduction.
- (b) From the given square matrix,

504	360	180
360	360	0
180	0	720

Compute its Eigen value and EigenVector.

$$(1 + 5) + (3 + 3) = 12$$

9. (a) Briefly explain how Principal Component Analysis can be used for data reduction.
- (b) Write notes on **any two** of the following topics:
- (i) Training and Testing Data set

B.TECH/CSE/8TH SEM/CSEN 4244/2021

- (ii) Data Clustering
- (iii) Linear and Non-linear class of data.

$$6 + (2 \times 3) = 12$$

Department & Section	Submission Link
CSE A + B + C	https://classroom.google.com/c/Mjk3MjQxODE3OTEy/a/MzYwNjczNzQ3NDU3/details