MACHINE LEARNING
(CSEN 5131)

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) Bayesian classifier is
- (a) a class of learning algorithm that tries to find an optimum classification of a set of examples using the probabilistic theory
- (b) any mechanism employed by a learning system to constrain the search space of a hypothesis
- (c) an approach to the design of learning algorithms that is inspired by the fact that when people encounter new situations, they often explain them by reference to familiar experiences, adapting the explanations to fit the new situation
- (d) none of these.

(ii) Pick the odd one out of the following options:
- (a) SVM
- (b) Multi layered Perceptron
- (c) DB-SCAN
- (d) Decision Tree.

(iii) In PCA transform the input data of certain dimension to a _____
- (a) data space of same dimension
- (b) data space of higher dimension
- (c) data space of smaller dimension
- (d) none of the above.

(iv) Which of the following is indicated by a hard margin in Support Vector Machine?
- (a) The SVM allows very low error in classification of training data
- (b) The SVM allows high amount of error in classification of training data
- (c) The SVM allows no error in classification of training data
- (d) None of the above.

(v) Which of the following is required by K-means clustering?
- (a) Defined distance metric
- (b) Number of clusters
- (c) Initial guess as to cluster centroids
- (d) All of the above.

(vi) Back propagation is a learning technique that adjusts weights in the neural network by propagating weight changes
- (a) forward from source to sink
- (b) backward from sink to source
- (c) forward from source to hidden nodes
- (d) backward from sink to hidden nodes.

(vii) Dimensionality reduction reduces the data set size by removing _____

(a) composite attributes (b) derived attributes
(c) relevant attributes (d) irrelevant attributes.

(viii) VC dimension of a hypothesis is n + 1, means,

(a) there are n + 1 point we cannot shattered
(b) there are n + 2 points we cannot shattered
(c) we cannot shatter any set of n + 1 points
(d) we cannot shatter any set of n + 2.

(ix) Consider the following data set, where x_1 and x_2 are input.

x_1	0	0	1	1
x_2	0	1	0	1
output	-1	+1	+1	-1

Minimum training error can be achieved by

(a) Single Layer Perceptron (b) Multi Layer Perceptron
(c) Both (a) and (b) (d) Neither (a) nor (b).

(x) In non-linear SVM, the kernel function $K(\boldsymbol{x_i}, \boldsymbol{x_j})$ has to satisfy the followings:

(a) $K(\boldsymbol{x_i}, \boldsymbol{x_j})$ is symmetric (b) $K(\boldsymbol{x_i}, \boldsymbol{x_j})$ is positive semi definite.
(c) Both (a) and (b) (d) None of the above.

Group – B

2. (a) Explain Naïve Bayes' Classification technique. [(CO1)(Remember/LOCQ)]

(b) Find the Class(X) using Naïve Bayes on the following Dataset, where, X = (Over 170CM; Eye=blue; Hair length=Long). Assume there are 2 classes as follows: Female and Male.

Sl No.	Over 170CM	Eye	Hair length	Gender
1	No	Blue	Short	Male
2	Yes	Brown	Long	Female
3	No	Blue	Long	Female
4	No	Blue	Long	Female
5	Yes	Brown	Short	Male
6	No	Blue	Long	Female
7	Yes	Brown	Short	Female
8	Yes	Blue	Long	Male

[(CO4),CO6)(Understand, Apply/IOCQ)]

4 + 8 = 12

3. (a) Describe the Perceptron Learning Algorithm (PLA) with appropriate example. [(CO1)(Remember/LOCQ)]

(b) Explain the working principle of PLA. [(CO3)(Explain/IOCQ)]

(c) Explain briefly the difference between Input space and feature space. Also explain how PLA can be used for nonlinear cases. [(CO2)(Apply/HOCQ)]

5 + 3 + 4 = 12

Group – C

4. (a) Explain the DBSCAN algorithm. [(CO1)(CO2)(CO3)(CO4)(Understand/IOCQ)]
   (b) Show all necessary steps to determine the clusters using DBSCAN algorithms for the following data points: A1 = (2,10), A2 = (2,5), A3 = (8,4), A4 = (5,8), A5 = (7,5), A6 = (6,4), A7 = (1,2), A8 = (4,9). Here, consider Epsilon as 2 and minpts as 2 for applying DBSCAN. [(CO2,CO4,CO5,CO6)(Apply/IOCQ)]

4 + 8 = 12

5. (a) Define minimum distance and maximum distances between two clusters. [(CO4)(Remember/LOCQ)]
   (b) Construct the dendrograms for the following proximity matrix using both minimum distance and maximum distances. Show all the steps. [(CO5)(Understand/LOCQ)]

	P1	P2	P3	P4	P5
P1	1.00	0.90	0.10	0.65	0.20
P2	0.90	1.00	0.70	0.60	0.50
P3	0.10	0.70	1.00	0.40	0.30
P4	0.65	0.60	0.40	1.00	0.80
P5	0.20	0.50	0.30	0.80	1.00

2 + 10 = 12

Group – D

6. (a) Explain the importance of VC dimension in machine learning. [(CO2,CO3,CO4)(Understand/LOCQ)]
   (b) Find the VC Dimension for the following hypotheses: Positive intervals
   f(x) = +1 for $P \leq x \leq Q$;
   otherwise f(x) = -1. [(CO2,CO3,CO4)(Analyze/IOCQ)]
   (c) Suppose 4 input data points X1, X2 ,X3, and X4 need to be classified into two classes +1 and -1. Determine a set or list of maximum possible dichotomies for these 4 data points, which can be obtained under the following three conditions:
   (i) break point is 2 (ii) break point is 3 (iii) break point is 4.
   [(CO2,CO3,CO4)(Analyze/IOCQ)]

3 + 3 + 6 = 12

7. (a) What is meant by hypothesis in the context of learning? Briefly discuss about Hoeffding's inequality. [(CO2,CO3,CO4)(Analyze/IOCQ)]
   (b) Explain the concept of dichotomy, growth function, break point in machine learning with the help of a suitable example. [(CO2,CO3,CO4)(Analyze/IOCQ)]

(2 + 4) + 6 = 12

Group – E

8. (a) Suppose a support vector machine for separating pluses from minuses finds a plus support vector at the point $x_1 = (1, 0)$, a minus support vector at $x_2 = (0, 1)$. You are to determine values for the classification vector w and the threshold value b. [(CO5)(Apply/IOCQ)]

(b) Construct the Lagrangian for the primal optimization problem in finding the support vectors for a two-class linearly separable classification problem.

[(CO4)(Understand/LOCQ)]

4 + 8 = 12

9. Write short notes on any 3 (three) of the followings: **(4 × 3) = 12**

(i) Data space and Feature space
(ii) Classification and Prediction
(iii) Under fitting and Over fitting
(iv) Precision and Recall. [(CO1)(Understand/HOCQ)]

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	35	48	17

Course Outcome (CO):

After the completion of the course students will be able to

1. Learn and understand various machine learning algorithms;
2. Understand complexity of Machine Learning algorithms and their limitations;
3. Compare and contrast various machine learning techniques and to get an insight of when to apply a particular machine learning approach;
4. Mathematically analyze various machine learning approaches and paradigms;
5. Apply common Machine Learning algorithms in practice and implementing their own;
6. Perform experiments in Machine Learning using real-world data

*LOCQ: Lower Order Cognitive Question; IOCQ: Intermediate Order Cognitive Question; HOCQ: Higher Order Cognitive Question.