

**MACHINE LEARNING
(CBS3202)**

Time Allotted : 2½ hrs

Full Marks : 60

Figures out of the right margin indicate full marks.

*Candidates are required to answer Group A and
any 4 (four) from Group B to E, taking one from each group.*

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

Group – A

1. Answer any twelve:

12 × 1 = 12

Choose the correct alternative for the following

- (i) In a binary classification task, a model identifies 120 True Positives, 30 False Positives, and 50 False Negatives. What is the calculated F1-score for this model?
(a) 0.80 (b) 0.75 (c) 0.70 (d) 0.65
- (ii) Using k-Nearest Neighbors with k=3 and Euclidean distance, classify point P(2, 2). The neighbors are A(1, 1) [Class 0], B(2, 1) [Class 1], C(3, 3) [Class 1], and D(5, 5) [Class 0]. Which class is assigned?
(a) Class 0 (b) Class 1 (c) Class 2 (d) Undetermined
- (iii) Given the ReLU activation function $f(x) = \max(0, x)$, what is the derivative $f'(x)$ when $x = -5$?
(a) 0.5 (b) 0 (c) 1 (d) -5
- (iv) A dataset with 10 features is processed with PCA. The eigenvalues obtained are [5, 3, 1, 0.5, 0.2, 0.1, 0.1, 0.05, 0.03, 0.02]. How many components are needed to retain at least 80% of the variance?
(a) 1 (b) 2 (c) 3 (d) 4
- (v) In a 5-fold cross-validation on a dataset of 100 samples, how many samples are used for training in each iteration?
(a) 20 (b) 80 (c) 500 (d) 100
- (vi) A dataset has 4 positive examples and 4 negative examples. What is the initial Entropy of this dataset?
(a) 0.5 (b) 0.75 (c) 1.00 (d) 0.80
- (vii) A neural network has 3 input neurons, 2 hidden layers with 4 neurons each, and 2 output neurons. How many total connections (weights) are there, assuming fully connected layers?
(a) 26 (b) 36 (c) 34 (d) 32

- (viii) Which clustering algorithm is most likely to identify an 'outlier' as a point with a label of -1?
 (a) PCA (b) DBSCAN (c) Hierarchical (d) k-Means
- (ix) An SVM has a total of 100 data points. If 95 points have $\alpha_i = 0$, how many Support Vectors are there?
 (a) 0 (b) 100 (c) 5 (d) 95
- (x) Which of the following is a key parameter in gradient descent?
 (a) Cluster centroids (b) Decision tree depth
 (c) Learning rate (d) Number of output classes

Fill in the blanks with the correct word

- (xi) A commonly used technique to prevent overfitting by adding a penalty term to the loss function is called ____.
- (xii) Calculate the Manhattan distance between points (2, 3) and (5, 1) is ____.
- (xiii) ____ technique is considered 'Supervised' dimensionality reduction.
- (xiv) If your training error is 2% and your validation error is 15%, the estimated "Variance" of the model is ____.
- (xv) A convolutional layer has an input size of 32×32 , a filter size of 5×5 , and no padding with a stride of 1. The output size is ____.

Group - B

2. (a) Marks obtained by 10 students in the class test and semester examination in machine learning are provided in the following table. Estimate the marks a student may obtain in the semester examination when she got 20 in class test, using linear regression.

Sl. No.	Class Test Marks	Semester Marks	Sl. No.	Class Test Marks	Semester Marks
1	28	63	6	28	51
2	27	49	7	26	66
3	23	43	8	21	36
4	17	36	9	22	31
5	24	39	10	19	37

[[CO2](Apply/IOCQ)]

- (b) What is regularization? Explain L1 and L2-regularization in the formulation of objective functions in machine learning algorithms.

[[CO2](Understand/LOCQ)]

6 + (2 + 4) = 12

3. (a) "For unbalanced dataset, accuracy is not a good metric for classification" – justify with an example.
- (b) There are 550 samples of men or women in a given dataset. Out of 60% women, 75% are classified correctly, and the remaining women are classified as men by a

[[CO1](Analyse/HOCQ)]

machine learning algorithm. On the other hand, in the men category, 85% are classified as men, and the remaining are classified as women.

i. Derive the confusion matrix.

ii. Calculate accuracy, precision, recall and F1 score.

[[CO1](Apply/IOCQ)]

$$4 + [3 + (2 + 1 + 1 + 1)] = 12$$

Group - C

4. Consider a two-class two-dimensional dataset given below:

Class 1: (1, 2), (1, 1), (2, 2)

Class 2: (2, 0), (1, -2)

i. What is the decision boundary in case of SVM?

ii. What is the width of the margin?

iii. Find out the class of the unknown point (2, 3).

[[CO3](Apply/IOCQ)]

$$(6 + 4 + 2) = 12$$

5. The following table contains training examples for a classification problem. Build a decision tree that predicts the class label. Show each step of the computation and make the decision of the new data with attributes value (sunny, visit, poor and no).

# ex.	Weather	Parents	Cash	Exam	Decision
1	sunny	visit	rich	yes	cinema
2	sunny	no-visit	rich	no	tennis
3	windy	visit	rich	no	cinema
4	rainy	visit	poor	yes	cinema
5	rainy	no-visit	rich	no	stay-in
6	rainy	visit	poor	no	cinema
7	windy	no-visit	poor	yes	cinema
8	windy	no-visit	rich	yes	shopping
9	windy	visit	rich	no	cinema
10	sunny	no-visit	rich	no	tennis
11	sunny	no-visit	poor	yes	tennis

[[CO3](Apply/IOCQ)]

$$(6 + 4 + 2) = 12$$

Group - D

6. Apply hierarchical clustering using the complete linkage method to the following data points: A(2, 1), B(2, 2), C(1, 4), D(4, 1), E(3, 2), F(2, 3), G(1, 2).

(i) Compute the pair wise distance matrix.

(ii) Determine the three clusters formed at the appropriate level.

[[CO4](Apply/IOCQ)]

$$(2 + 10) = 12$$

7. Consider the following set of data points in a 2D space:

A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9).

Using the DBSCAN clustering algorithm with parameters $\epsilon = 2$ and MinPts = 2, answer the following:

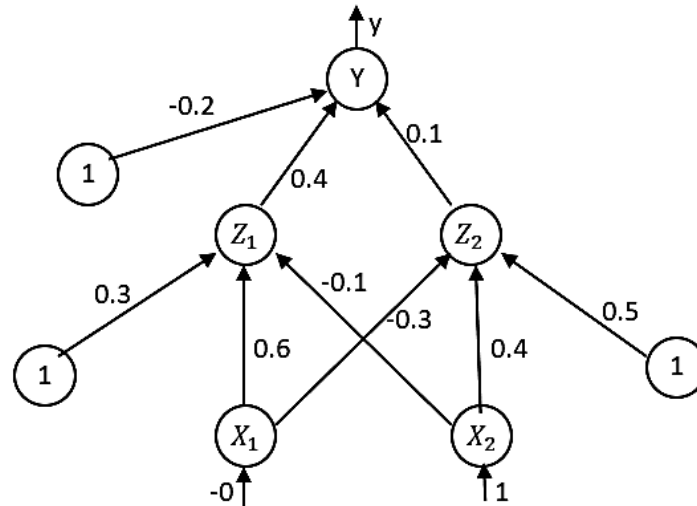
- (i) Identify all core points, border points, and noise points.
- (ii) Determine how many clusters are formed.

[[CO4](Apply/IOCQ)]

12

Group - E

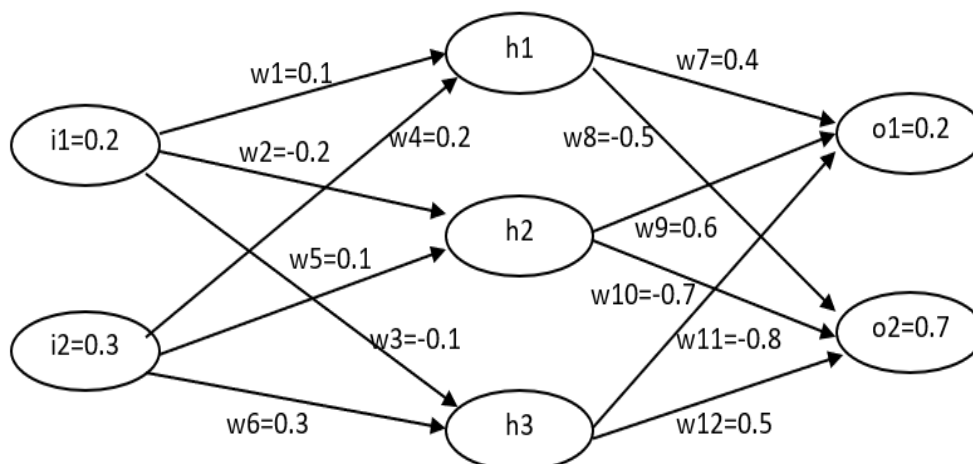
8. Using a back-propagation network, find the new weights shown in Figure below. It is presented with the input pattern $[0, 1]$ and the target output is 1. Use a learning rate = 0.25 and binary sigmoidal activation function.



[[CO5](Apply/IOCQ)]

12

9. Consider the following neural network with the input, output and weight parameters values shown in the diagram. The activation values in each neuron are calculated using the sigmoid activation function. Now, answer the following:
 - (i) For the given input i_1 and i_2 as shown in the diagram, compute the output of the hidden layer and output layer neurons.
 - (ii) Compute the error in the network with the initialized weight parameters shown in the diagram.
 - (iii) Update the weight parameters for w_7 to w_{12} using backpropagation algorithm in the first iteration. Consider learning rate as 0.01.



[[CO5](Apply/IOCQ)]

12

Cognition Level	LOCQ	IOCQ	HOCQ
Percentage distribution	6.25	89.58	4.17