



SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
FALL SEMESTER 2025-2026

Programme Name & Branch : B.Tech(CSE)
Course Code and Course Name : BCSE209L & Machine Learning
Faculty Name(s) : SUDHA.S, SARAVANAKUMAR K, VISWANATHAN P, SUDHAKARAPANDIAN R
Class Number(s) : 2085, 2087, 2098, 2173
Date of Examination : 05/10/2025
Exam Duration : 90 minutes
Maximum Marks: 50

General instruction(s):

- Answer All Questions
- **Note:** Step by step answer required.
- Course Outcomes :
- CO2 - Apply appropriate algorithm to the data.
- CO3 - Analyze the results of algorithm and convert to appropriate information required for the real - time application.

Q. No	Question	Module	Marks	CO	BL
1.	Apply Agglomerative clustering using complete linkage method to create clusters from the data points (1,10), (9,7), (3,8), (3,4), (5,8), (4,3). Use Euclidean distance as a similarity measure to construct the initial distance matrix and draw a dendrogram.	4	10	CO2	BL3
2.	Given the data points A(3,7), B(4,6), C(5,5), D(6,4), E(7,3), F(6,2), G(7,2) and H(8,4). Cluster the data points using DBSCAN and find if there are any outliers. Take radius around each point, Eps=2.5 and minimum number of data points that should be around that points is MinPts=3.	4	10	CO2	BL3
3.	Construct a Self-Organizing Map for given five input vector of three samples: X1= [0.6 0.2 0.3 0.7] X2= [0.4 0.1 0.3 0.5] X3= [0.5 0.2 0.1 0.6] Number of clusters to be formed is 2 and the learning rate is 0.5. Initial weights are [1.0 0.4 0.7 0.3] [0.3 0.6 0.5 0.2]. Repeat the algorithm for one iteration.	4	10	CO3	BL3
4.	a) Consider a machine learning model with the following characteristics and explain how it relates to bias and variance. (i) High training error and High validation error, similar to the training error. (ii) Low training error and Low validation error, similar to the training error. (iii) High training error and Low validation error (iv) High training error and Validation error is higher than training error.	5	6	CO3	BL4
	b) Given Average intra-cluster distance (a) = 1.5 and Average inter-cluster distance (b) = 4.0 Calculate the silhouette score for a point and explain how it relates to measures the data points fits into its assigned cluster.		4		



SCHOOL OF COMPUTER SCIENCE AND ENGINEERING
CONTINUOUS ASSESSMENT TEST - II
FALL SEMESTER 2025-2026

5. Consider the following dataset.

Patient	Age	BMI	Glucose	Diabetes (+1/-1)
1	55	28	160	+1
2	40	22	110	-1
3	60	30	170	+1
4	35	20	100	-1
5	50	26	150	+1
6	45	24	130	-1
7	65	32	180	+1
8	42	23	120	-1

Apply the Adaboost Ensemble learning method to create a new dataset that will help to predict whether the patient has diabetes or not. Assume the random values are: 0.2, 0.38, 0.88, 0.40, 0.55, 0.63, 0.34 and 0.77. Consider a stump creation with the condition Age > 50 , BMI > 25 , Glucose > 140 then Predict is equal to +1 .

5

10

CO3

BL5
