



VIT

Vellore Institute of Technology
Approved to be a University under Section 3 of UGC Act, 1956
CHENNAI

Reg. Number: _____

Continuous Assessment Test (CAT) – II OCTOBER 2025

Programme	: B.Tech CSE (AI & ML Specilization)	Semester	: FS 2025-26
Course Code & Course Title	: BCSE 209L	Class Number	: CH2025260102135, CH2025260102137, CH2025260102133, CH2025260102129, CH2025260102131
Faculty	: Dr. Bhargavi R, Dr. Prakash P, Dr. Rajalakshmi R, Dr. Sajidha S A, Dr. Syed Ibrahim S P	Slot	: A2 + TA2
Duration	: 90 minutes	Max. Mark	: 50

General Instructions:

- Write only your registration number on the question paper in the box provided and do not write other information

Answer all questions

Q. No	Sub Sec.	Description	Marks	CO CO4	BT Level K4
1		a) A retail company wants to group its customers based on the following categorical features: • Region (North, South, East, West) • Preferred Payment Method (Cash, Card, UPI, Wallet) • Shopping Frequency (Low, Medium, High) • Loyalty Membership (Yes, No). Apply an appropriate algorithm and illustrate the steps in detail to group the customers [5 Marks] b) Design a simple multilayer perceptron (MLP) to predict student academic performance based on two normalized input features: study hours ($x_1 = 0.6$) and previous exam score ($x_2 = 0.4$). The network should have one hidden layer with a single neuron and a single output neuron, using sigmoid activation functions throughout. Given initial weights $w_1 = 0.3$ (input₁ to hidden), $w_2 = 0.2$ (input₂ to hidden), and $v = 0.4$ (hidden to output), along with biases $b_h = 0.1$ (hidden layer) and $b_o = 0.05$ (output layer), with a learning rate of $\eta = 0.5$ and target output $y = 1$ (indicating pass). i) Perform one forward pass and compute the predicted output $\hat{y}$ [3 marks] ii) Calculate the mean squared error loss between the predicted and target values. [2 marks] iii) Apply backpropagation and update all parameters using gradient descent by illustrating the steps involved [5 marks]	15		

Consider a two-class classification problem with a dataset that is perfectly separable by a linear boundary. You are given a choice between a Hard-Margin SVM and a Soft-Margin SVM with a very high value of the penalty parameter C .

- How would you compare the decision boundaries and margins of these two models? [4 Marks]
- If you add a single, highly influential outlier within the opposite class's data, how would each model's decision boundary likely to change? Justify your answer. [4 Marks]
- Based on your analysis, elaborate which SVM (soft margin / hard margin) will be more robust and practical choice for real-world engineering problems. [3 Marks]
- Discuss the impact on the margin in the following two cases
 - Removing the support vectors from the data [2 Marks]
 - Removing the non-support vectors from the data [2 Marks]

15

CO1

K3

2

A wildlife researcher is tracking animal sightings in a forest. They have recorded the GPS coordinates (latitude and longitude) of 8 sightings of a rare species. The researcher wants to group sightings that are close together in space, where a group should have at least 2 points within a distance of 2 units from each other. Points that don't belong to any group should be treated as isolated.

Sighting	X	Y
1	1	2
2	2	2
3	2	3
4	8	7
5	8	8
6	25	80
7	24	79
8	25	81

- Using an appropriate algorithm, identify the clusters formed and noise points, if any. [7 Marks]
- Discuss why some sightings are grouped together and why others are classified as noise. [3 Marks]

10

CO2

K4

3

Consider a Kohonen Self-organizing network that can organize/group the data into five clusters. Consider a two dimensional input vector (x_1, x_2) . Let the weights from first and second input nodes to all the five output nodes be $[0.3, 0.2, 0.1, 0.8, 0.4]$ and $[0.5, 0.6, 0.7, 0.9, 0.2]$, respectively.

10

4

a) Find the cluster unit close (BMU) to the data point (0.2, 0.4) using a learning rate of 0.2. Determine the updated weights. [5 marks]

b) Now consider another input data (0.6, 0.6). With a learning rate of 0.1, find the winning cluster for this input. Determine the updated weights. [5 marks]

Note: Use the square of the Euclidean distance to find the cluster distances.

CO3

K3

*****All the best *****

Ad. Bannu