



VIT
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)
CHENNAI

Reg. Number: _____

Continuous Assessment Test (CAT) – I AUGUST 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
- Only non-programmable calculator without storage is permitted

Answer all questions

Q. No	Sub Sec.	Description	Marks	CO	BT Level
1	a)	An AgriTech company is collaborating with tomato farmers to maximize profitability by optimizing yield and ensuring high-quality produce for premium markets. They have collected the following data over multiple growing seasons: Case 1: Sensor data (Soil moisture, temperature, pH), Weather data (rainfall, humidity, sunlight hours), along with the corresponding yield (kg/hectare). Case 2: Drone-captured images that are annotated / labelled by the experts as (Very good, good, bad) based on the Tomato Quality. Your role is to design an appropriate ML model to help the AgriTech company. Discuss in detail about your recommended choice of ML algorithms for Case 1 and Case 2, separately. Also, justify your choice. [6 marks]	10	CO1	K3
	b)	Identify the suitable performance metrics for Case 1 and Case 2, also provide suitable justification for the choice of metrics with the necessary details pertaining to computing the metrics. [4 marks]			
2		While developing a medical diagnosis model, you have constructed two models, M1 and M2, by training on a balanced dataset. The model M1 appears to learn patient data too specifically, whereas the model M2 misses important disease patterns. Discuss the possible reasons for behaviour of individual models and propose ways to improve performance in both scenarios.	6	CO1	K5

Your task is to build a short text classifier using a probability based method, assuming that the words are independent.

Predict the category of the document D6, by illustrating the step-by-step procedure.

Doc ID	W1	W2	W3	Label
D1	University	lecture	exam	Education
D2	Movie	actor	concert	Entertainment
D3	Teacher	school	homework	Education
D4	Music	movie	concert	Entertainment
D5	Lecture	class	university	Education
D6	homework	lecture	university	?

10

CO4

K5

Using the data given below, train a single layered neural network. Assume the initial weights w_0 , w_1 , w_2 and the learning rate as follows: 0.3, -0.2, 0.1 and 0.1 with the step / threshold function as the activation function. Restrict to a maximum of 2 iterations.

With the above constructed model, predict whether the given species having the features $A_1=5.2$ and $A_2=1.6$, belong to Class 1 or Class 0.

Sample	A1	A2	Class Label
1	5.1	1.4	1
2	4.9	1.5	1
3	4.7	1.3	1
4	6.0	4.8	0
5	6.1	4.9	0
6	5.8	4.7	0

12

CO4

K3

Assume that you have built an image classifier that predicts one of the types of animals based on the maximum probability: Dog (0) / Cat (1) / Rabbit (2). For 10 samples, the true labels and predicted probabilities are given below.

12

CO2

K2

Sample No.	True Label	Predicted Probabilities (Dog / Cat / Rabbit)
1	0	(0.70, 0.20, 0.10)
2	1	(0.10, 0.80, 0.10)
3	2	(0.20, 0.30, 0.50)
4	0	(0.40, 0.50, 0.10)
5	1	(0.30, 0.60, 0.10)
6	2	(0.35, 0.50, 0.15)
7	0	(0.80, 0.15, 0.05)
8	1	(0.60, 0.30, 0.10)
9	2	(0.25, 0.55, 0.20)
10	0	(0.60, 0.10, 0.30)

Use the above information to perform the following.

- Construct the confusion matrix (rows - true labels and columns - predicted labels) [2 marks]
- Calculate the overall accuracy and class-wise accuracy [2 marks]
- For each class, determine the precision, recall and F1 [6 marks]
- Interpret the obtained results for each class and discuss why the accuracy metric alone is not sufficient to assess the model's performance. Do you think, it is a reliable model for identifying Rabbits ? Justify your answer. [2 marks]