



SCHOOL OF COMPUTER SCIENCE ENGINEERING AND INFORMATION SYSTEMS
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025

SLOT: C1+TC1

Programme Name & Branch : B.Tech. Information Technology
Course Code & Course Name : BITE401L & Machine Learning
Faculty Name(s) : Dr. Selva Rani B, Dr. Prabhavathy P
Class Number(s) : VL2025260103200, VL2025260103192
Date of Examination : August 19, 2025
Exam Duration : 90 minutes

Maximum Marks: 50

General instructions:

- Answer All the Questions. You are expected to solve the **numerical problems step by step**.
- M-Max mark; CO-Course Outcome; BL-Blooms Taxonomy Level (1-Remember, 2-Understand, 3-Apply, 4-Analyse, 5-Evaluate, 6-Create)
- Course Outcomes:
CO1 – Understand the various machine learning approaches and data preprocessing techniques
CO2 – Develop the learning models for the problems using supervised and advanced learning-based approaches

Q. No	Question	M	CO	BL																									
1.	You have just joined the data science team at QuickBite, a fast-growing food delivery company. The management is eager to use machine learning to improve the customer experience, increase efficiency, and reduce fraud. They have outlined the following goals for you: a) Customer Rating Prediction QuickBite wants to be able to predict the rating (1–5 stars) a customer is likely to give before the order is delivered. Input data may include order details (items, quantity, price), past customer ratings, delivery time, weather conditions, and delivery partner's past performance. Goal: If the predicted rating is low, the system could take preventive actions—such as offering a coupon, prioritizing delivery speed, or changing the delivery partner b) Dish Recommendation QuickBite wants to recommend dishes to customers in real time based on their previous orders and the ordering habits of similar customers. Example: If a customer usually orders spicy noodles on weekends, the app might suggest a trending spicy ramen from a nearby restaurant they've never tried before. Goal: Increase order frequency and introduce customers to new restaurants. For each of the above goals, identify which type of machine learning would be most appropriate. Briefly explain why you chose that type, referring to the nature of the data and the desired outcome.	10	1	3																									
2.	Our institute wants to check if there is a relationship between students' degree and their preferred study environment. They survey 200 students and record the results as given in the table below. Compute the Chi-squared statistic and determine whether there is a significant relationship between Degree and Preferred Study Environment. <table border="1"> <thead> <tr> <th>Study Environment / Degree</th><th>BTech</th><th>MTech</th><th>MCA</th><th>Total</th></tr> </thead> <tbody> <tr> <td>Library</td><td>35</td><td>20</td><td>25</td><td>80</td></tr> <tr> <td>Food Court</td><td>25</td><td>30</td><td>20</td><td>75</td></tr> <tr> <td>Woodys</td><td>30</td><td>10</td><td>5</td><td>45</td></tr> <tr> <td>Total</td><td>90</td><td>60</td><td>50</td><td>200</td></tr> </tbody> </table>	Study Environment / Degree	BTech	MTech	MCA	Total	Library	35	20	25	80	Food Court	25	30	20	75	Woodys	30	10	5	45	Total	90	60	50	200	10	1	4
Study Environment / Degree	BTech	MTech	MCA	Total																									
Library	35	20	25	80																									
Food Court	25	30	20	75																									
Woodys	30	10	5	45																									
Total	90	60	50	200																									
3.	Does walking more steps per day relate to scoring a higher CGPA? A study was conducted with 7 VIT students tracking their average daily steps, average OTT watching hours and their CGPA for till now. Using the Pearson correlation coefficient, determine whether there is a relationship between physical activity (steps walked) and academic performance (CGPA).	10	1	4																									



SCHOOL OF COMPUTER SCIENCE ENGINEERING AND INFORMATION SYSTEMS
CONTINUOUS ASSESSMENT TEST - I
WINTER SEMESTER 2024-2025

SLOT: C1+TC1

Student	Steps Walked	OTT Hours	CGPA
S1	8000	2.0	8.4
S2	4500	5.0	7.1
S3	11000	1.5	8.9
S4	3000	6.0	6.8
S5	9500	2.5	8.6
S6	6000	4.5	7.4
S7	12500	1.0	9.0

4. The following table shows water quality data collected from 6 locations, with three measured parameters: pH, dissolved oxygen (DO in mg/L), and turbidity (NTU). Perform the first step involved in PCA, Standardization of the dataset using z-score method. Show all calculations clearly.

Location	pH	DO	Turbidity
Chennai	7.1	6.5	4.0
Bengaluru	6.8	6.8	4.5
Trivandrum	7.3	6.2	3.8
Hyderabad	7.0	6.9	4.2
Mumbai	6.9	6.6	4.1
Kolkata	7.2	6.4	3.9

10 1 3

[Hint: Attribute meanings: pH → acidity/basicity of water, Dissolved Oxygen (DO) → oxygen content (key for aquatic life), Turbidity → cloudiness of water (higher = more particles)]

5. You're a data analyst for a co-working café that wants to understand how hours spent working at the café influence a freelancer's weekly earnings. You collected data from 12 freelancers for one week.

- a) Fit a simple linear regression model predicting weekly earnings from hours at café.
b) Use the model to predict the expected weekly earnings for a freelancer who spends 8.5 hours at the café that week.

8
2

Freelancer	Hours at café per week (X)	Weekly earnings (€) (Y)
1	6	220
2	8	260
3	4	180
4	10	300
5	3	150
6	12	340
7	7	240
8	5	200
9	9	280
10	2	130
11	11	320
12	1	115

2 3