Roll No: | | | | | | | | | | | |

PRESIDENCY UNIVERSITY

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

Presidency University Act, 2013 of the Karnataka Act No. 41 of 2013 | Established under Section 2(f) of UGC Act, 1956
Approved by AICTE, New Delhi | Approved By BCI
Bengaluru

Mid Term Examinations - October 2026

Date: 07/10/2026 **Time:** 09:30 AM - 11:00 AM

Course Code: CAI2500 **Course Name:** Machine Learning

Semester: Five Semester **Max. Marks:** 50 **Weightage:** 50%

CO - Levels	CO1	CO2
Marks	46	44

PART A: Answer Following Questions. 10 M

Q.N.	Questions	M	CO	BT
1	Define Machine Learning and state the difference between supervised and unsupervised learning	2	CO1	BT2
2	A model predicts values [5, 8] for actual values [4, 10]. Calculate the MSE and MAE.	2	CO1	BT3
3	A logistic regression model gives z = 1.5 for a sample. Calculate the predicted probability and the class using a threshold of 0.5	2	CO1	BT3
4	A word occurs 4 times in documents belonging to Class A. The total number of words in Class A is 20, and the vocabulary contains 5 words. Using Laplace smoothing, calculate the probability of the word given Class A.	2	CO2	BT3
5	Given two vectors X= (1,2), Z= (3,4). Calculate their dot product, which is required for evaluating a polynomial kernel.	2	CO2	BT3

PART B : Answer Any 1 Following Questions. 10 M

Q.N.	Questions	M	CO	BT
6	The following dataset contains missing values: **Age Income Experience** **25 30,000 2** **30 40,000 5** **28 Missing 4** **35 50,000 Missing** **32 45,000 8** Using mean imputation, calculate the missing values for Income and Experience. Also show the intermediate calculations and the resulting completed dataset.	10	CO1	BT3

7	A dataset contains the following categorical features: Customer Gender Payment Mode City C1 Male Cash Bangalore C2 Female UPI Chennai C3 Male Card Bangalore C4 Female UPI Delhi Apply appropriate encoding techniques to convert the categorical features into numerical form: 1. Perform Label Encoding for Gender. 2. Perform One-Hot Encoding for Payment Mode. 3. Perform One-Hot Encoding for City. 4. Show the final numerical representation of all four records.	10	CO1	BT3

PART C: Answer Any 1 Following Questions. 10 M

Q.N.	Questions	M	CO	BT
8	A bank uses a logistic regression model to predict whether a customer will take up a loan offer (Y = 1) or not (Y = 0). The model uses two inputs: the customer's age (in years) and annual income (in thousands). The intercept is −4, the coefficient for age is 0.08, and the coefficient for annual income is 0.05. A customer is aged 40 and has an annual income of 60 thousand. 1. Calculate the value of the linear score z for this customer. 2. Calculate the predicted probability that the customer takes up the loan offer. Using a classification threshold of 0.5, determine the predicted class.	10	CO1	BT3
9	A three-class classification model (Class A, Class B, Class C) produces the following raw scores for an input sample: • Class A: 2 • Class B: 1 • Class C: 0 1. Calculate the Softmax probability for each class. 2. Determine the predicted class. 3. If the actual class is Class A, calculate the Cross-Entropy Loss (use the natural logarithm).	10	CO1	BT3

PART D: Answer Any 1 Following Questions. 10 M

Q.N.	Questions	M	CO	BT
10	A university uses an automated system to identify students who may be at risk of failing a course. Based on historical data, the following probabilities are known: • Probability that a randomly selected student fails the course, P(F)=0.20 • Probability that the system generates an Alert given that the student fails, P(A|F) =0.80 • Probability that the system generates an Alert given that the student passes, P(A|P) =0.10 If a student receives an Alert, apply Bayes' theorem to determine the probability that the student will actually fail the course. Show all necessary steps, including the calculation of the total probability of receiving an Alert.	10	CO2	BT3

11	A text classifier distinguishes between Sports and Politics documents. The vocabulary consists of: {goal, team, vote, election}. The word counts are: Word Sports Politics goal 4 1 team 3 1 vote 1 4 election 0 5 Total word counts are Sports = 8, Politics = 11. Assume: P(Sports)=P(Politics)=0.5P(Sports)=P(Politics)=0.5. Apply Multinomial Naïve Bayes with Laplace smoothing to classify the document: "goal team vote". Show all intermediate probability calculations.	10	CO2	BT3

PART E: Answer Any 1 Following Questions. 10 M				
Q.N.	**Questions**	**M**	**CO**	**BT**
12	An SVM classifier produces the following results on 100 test samples: • True Positive = 40 • True Negative = 35 • False Positive = 10 • False Negative = 15 Apply the appropriate evaluation measures to calculate: 1. Accuracy 2. Precision 3. Recall 4. F1-score 5. Specificity Based on the calculated values, identify which metric would be useful when the primary concern is detecting as many actual positive cases as possible.	10	CO2	BT3
13	A weather prediction system uses Gaussian Naïve Bayes to classify a day as either Sunny or Rainy based on temperature and humidity. The following parameters are given: Class Prior Temperature (μ,σ) Humidity (μ,σ) Sunny 0.6 (30, 3) (55, 8) Rainy 0.4 (24, 4) (80, 10) For a day with temperature = 27°C and humidity = 70%, use the Gaussian probability density function to calculate the probability for each class and determine whether the day is Sunny or Rainy.	10	CO2	BT3