



PRESIDENCY UNIVERSITY

BENGALURU

Roll No.																			
----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Mid - Term Examinations - March 2026

Date: 14 -03-2026

Time: 09.30am to 11.00am

School: SOCSE	Program: B.Tech		
Course Code: CAI2500	Course Name: Machine Learning		
Semester: IV/VI	Max Marks: 50	Weightage: 25%	

CO - Levels	C01	C02	C03	C04	C05
Marks	24	14	12	-	-

Instructions:

- (i) Read all questions carefully and answer accordingly.
- (ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks.

5Q x 2M=10M

1	Differentiate supervised and unsupervised methods of machine learning.	2 Marks	L2	C01
2	Find the RMSE of a trained simple linear regression model where $y=\{1.5, 2.0,4.5\}$ and $y' = \{1.0, 2.3, 4.0\}$	2 Marks	L3	C01
3	Differentiate Gaussian and Multinomial Naïve Bayes in terms of the feature type, assumptions of the input feature distribution and the key mechanism used.	2 Marks	L2	C02
4	What are soft margin SVM classifiers?	2 Marks	L1	C02
5	Define bias-variance tradeoff in machine learning.	2 Marks	L1	C03

Part B

Answer the Questions.

Total Marks 40M

6.	a.	Explain the step-by-step procedure of the kNN classification algorithm and its drawbacks.	10 Marks	L1	CO1
Or					
7.	a.	<p>Consider a multi-class problem where the set of class labels $Y = \{y_1, y_2, y_3\}$. A certain test instance is classified as (+, -, -) by all binary classifiers using the one-against-rest approach of multi-class classifiers. Answer the following:</p> <p style="margin-left: 40px;">i) How many binary classifiers are trained totally?</p> <p style="margin-left: 40px;">ii) Using a table clearly indicate the positive and negative examples of each binary classifier.</p> <p style="margin-left: 40px;">iii) Calculate the final predicted class of the test instance.</p>	10 Marks	L3	CO1

8.	a.	<p>A decision tree based classification model is to be trained using the following data set to predict the factors affecting sunburn. Using multi-way split on the attributes, gain and Gini Index as the measure of node impurity, find the root node of the tree. Clearly show the detailed working with the gain of each candidate splitting attribute. Draw the decision tree after the first iteration.</p> <table border="1" style="width: 100%; border-collapse: collapse; margin: 10px 0;"> <thead> <tr> <th style="width: 15%;">Hair</th> <th style="width: 15%;">Height</th> <th style="width: 15%;">Weight</th> <th style="width: 15%;">Lotion</th> <th style="width: 15%;">Sunburned</th> </tr> </thead> <tbody> <tr><td>Blonde</td><td>Average</td><td>Light</td><td>No</td><td>Yes</td></tr> <tr><td>Blonde</td><td>Tall</td><td>Average</td><td>Yes</td><td>No</td></tr> <tr><td>Brown</td><td>Short</td><td>Average</td><td>Yes</td><td>No</td></tr> <tr><td>Blonde</td><td>Short</td><td>Average</td><td>No</td><td>Yes</td></tr> <tr><td>Red</td><td>Average</td><td>Heavy</td><td>No</td><td>Yes</td></tr> <tr><td>Brown</td><td>Tall</td><td>Heavy</td><td>No</td><td>No</td></tr> <tr><td>Brown</td><td>Average</td><td>Heavy</td><td>No</td><td>No</td></tr> <tr><td>Blonde</td><td>Short</td><td>Light</td><td>Yes</td><td>No</td></tr> </tbody> </table>	Hair	Height	Weight	Lotion	Sunburned	Blonde	Average	Light	No	Yes	Blonde	Tall	Average	Yes	No	Brown	Short	Average	Yes	No	Blonde	Short	Average	No	Yes	Red	Average	Heavy	No	Yes	Brown	Tall	Heavy	No	No	Brown	Average	Heavy	No	No	Blonde	Short	Light	Yes	No	10 Marks	L3	CO1
Hair	Height	Weight	Lotion	Sunburned																																														
Blonde	Average	Light	No	Yes																																														
Blonde	Tall	Average	Yes	No																																														
Brown	Short	Average	Yes	No																																														
Blonde	Short	Average	No	Yes																																														
Red	Average	Heavy	No	Yes																																														
Brown	Tall	Heavy	No	No																																														
Brown	Average	Heavy	No	No																																														
Blonde	Short	Light	Yes	No																																														
Or																																																		

9.	a.	Using min-max normalization, transform the given values of salary = {1000, 2000, 3000, 9000} into the range 0 to 1. Write the transformed values of salary.	10 Marks	L3	C01
----	----	---	----------	----	-----

10.	a.	<p>For the training set given below, using Gaussian Naïve Bayes estimate the class of the given test data</p> <p>X = (Refund = Yes, Divorced, Taxable_Income = 120K).</p> <p>The sample mean and variance of Evade = 'Yes' for the given taxable income are {90,25} respectively and for Evade = 'No' the mean and variance are {110, 2975} respectively.</p> <table border="1"> <thead> <tr> <th>Tid</th> <th>Refund</th> <th>Marital Status</th> <th>Taxable Income</th> <th>Evade</th> </tr> </thead> <tbody> <tr><td>1</td><td>Yes</td><td>Single</td><td>125K</td><td>No</td></tr> <tr><td>2</td><td>No</td><td>Married</td><td>100K</td><td>No</td></tr> <tr><td>3</td><td>No</td><td>Single</td><td>70K</td><td>No</td></tr> <tr><td>4</td><td>Yes</td><td>Married</td><td>120K</td><td>No</td></tr> <tr><td>5</td><td>No</td><td>Divorced</td><td>95K</td><td>Yes</td></tr> <tr><td>6</td><td>No</td><td>Married</td><td>60K</td><td>No</td></tr> <tr><td>7</td><td>Yes</td><td>Divorced</td><td>220K</td><td>No</td></tr> <tr><td>8</td><td>No</td><td>Single</td><td>85K</td><td>Yes</td></tr> <tr><td>9</td><td>No</td><td>Married</td><td>75K</td><td>No</td></tr> <tr><td>10</td><td>No</td><td>Single</td><td>90K</td><td>Yes</td></tr> </tbody> </table>	Tid	Refund	Marital Status	Taxable Income	Evade	1	Yes	Single	125K	No	2	No	Married	100K	No	3	No	Single	70K	No	4	Yes	Married	120K	No	5	No	Divorced	95K	Yes	6	No	Married	60K	No	7	Yes	Divorced	220K	No	8	No	Single	85K	Yes	9	No	Married	75K	No	10	No	Single	90K	Yes	10 Marks	L3	C02
Tid	Refund	Marital Status	Taxable Income	Evade																																																								
1	Yes	Single	125K	No																																																								
2	No	Married	100K	No																																																								
3	No	Single	70K	No																																																								
4	Yes	Married	120K	No																																																								
5	No	Divorced	95K	Yes																																																								
6	No	Married	60K	No																																																								
7	Yes	Divorced	220K	No																																																								
8	No	Single	85K	Yes																																																								
9	No	Married	75K	No																																																								
10	No	Single	90K	Yes																																																								

Or

11.	a.	What is kernel trick in SVM classifiers? Explain the polynomial, RBF and Sigmoid kernels used in SVM with their mathematical representations.	10 Marks	L1	C02
-----	----	---	----------	----	-----

12.	a.	Explain how bagging and pasting approaches help to avoid overfitting with their step-by-step procedure.	10 Marks	L1	C03
-----	----	---	----------	----	-----

Or

13.	a.	Write the step-by-step procedure of the AdaBoost algorithm.	10 Marks	L1	C03
-----	----	---	----------	----	-----