

CBCS SCHEME

USN

--	--	--	--	--	--	--	--	--	--

BAI602

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2026 Machine Learning - I

Time: 3 hrs.

Max. Marks: 100

*Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M : Marks, L: Bloom's level, C: Course outcomes.*

Module – 1				M	L	C
Q.1	a.	Draw the chart showing the types of machine learning and state each, in brief.		10	L1	CO1
	b.	i) Apply min-max procedure to normalize the set $V = \{42, 35, 28, 16\}$ to the new range 0 to 1. ii) Apply Z-score normalization to the marks list : $M = \{40, 30, 20\}$		10	L3	CO2
OR						
Q.2	a.	Describe any four characteristics of Big Data and the four types of Data Analytics.		10	L1	CO1
	b.	i) Suppose a linear regression algorithm generates a model : $y = 0.54 + 0.66x$ for data pertaining to weekly sales data of a product. Here x is the week and y represents the sales data. Predict the sales for the 5 th and 8 th weeks. ii) Consider the following set : $S = \{11, 13, 18, 21, 23, 25, 27, 30, 33\}$ Apply various binning techniques using three bins and show the result.		10	L3	CO2
Module – 2						
Q.3	a.	Discuss the following distributions along with their parameters : i) Binomial Distribution ii) Exponential Distribution iii) Poisson Distribution		10	L2	CO2

	b.	<table><tr><th colspan="3">PREDICTED LABEL</th></tr><tr><th rowspan="2">TRUE LABEL</th><th></th><th>NEGATIVE</th><th>POSITIVE</th></tr><tr><th>NEGATIVE</th><td>55 (TN)</td><td>5(FP)</td></tr><tr><th></th><th>POSITIVE</th><td>10 (FN)</td><td>30 (TP)</td></tr></table> Given the above confusion matrix compute: accuracy, sensitivity, precision, recall and FI-score.	PREDICTED LABEL			TRUE LABEL		NEGATIVE	POSITIVE	NEGATIVE	55 (TN)	5(FP)		POSITIVE	10 (FN)	30 (TP)	10	L3	CO2							
PREDICTED LABEL																										
TRUE LABEL		NEGATIVE	POSITIVE																							
	NEGATIVE	55 (TN)	5(FP)																							
	POSITIVE	10 (FN)	30 (TP)																							
OR																										
Q.4	a.	Explain the following : i) Wrapper based methods for feature engineering ii) Singular Value Decomposition (SVD)	10	L2	CO2																					
	b.	i) Solve the following set of equations using Gaussian Elimination Method: $3x_1 + 6x_2 = 9$ $4x_1 + 10x_2 = 14$ ii) Find the covariance between the two features for the following bivariate data : <table><tr><td>Age</td><td>10</td><td>20</td><td>30</td><td>40</td><td>50</td></tr><tr><td>Weight</td><td>25</td><td>40</td><td>60</td><td>80</td><td>70</td></tr></table>	Age	10	20	30	40	50	Weight	25	40	60	80	70	10	L3	CO2									
Age	10	20	30	40	50																					
Weight	25	40	60	80	70																					
Module – 3																										
Q.5	a.	Draw the chart of types of Regression Methods and explain each, in brief. Also discuss the limitations of Regression Method.	10	L2	CO3																					
	b.	Apply the K-Nearest Neighbours (KNN) algoirhtm with K = 3 for the sample data shown in the below Table Q5(b). Predict the class for the test instance (4, 4) Table Q5(b) Sample Data <table><tr><td>x</td><td>y</td><td>Class</td></tr><tr><td>3</td><td>1</td><td>A</td></tr><tr><td>5</td><td>2</td><td>A</td></tr><tr><td>4</td><td>3</td><td>A</td></tr><tr><td>7</td><td>6</td><td>B</td></tr><tr><td>6</td><td>7</td><td>B</td></tr><tr><td>3</td><td>3</td><td>B</td></tr></table>	x	y	Class	3	1	A	5	2	A	4	3	A	7	6	B	6	7	B	3	3	B	10	L3	CO3
x	y	Class																								
3	1	A																								
5	2	A																								
4	3	A																								
7	6	B																								
6	7	B																								
3	3	B																								

OR

OR																							
Q.6	a.	Explain Instance based learning and compare it with model based learning.	10	L2	CO3																		
	b.	Suppose the five week's sales data (in thousands) is as given below : <table><tr><td>Week</td><td>:</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr><tr><td>Sales</td><td>:</td><td>1</td><td>4</td><td>9</td><td>16</td><td>25</td></tr></table> Develop a linear regression model to fit the above data. Predict the sales for the 8 th and 11 th weeks.	Week	:	1	2	3	4	5	Sales	:	1	4	9	16	25	10	L3	CO3				
Week	:	1	2	3	4	5																	
Sales	:	1	4	9	16	25																	
Module – 4																							
Q.7	a.	Discuss the terms : Prior probability, Likelihood probability and Posterior probability. Explain the Naïve Bayes algorithm for classification.	10	L2	CO4																		
	b.	Construct a decision tree using ID3 algorithm for the HR dataset in Table Q7(b). Predict whether a person who has completed High School and is performing medium would be promoted or not. <table><tr><th colspan="3">Table Q7(b) HR Dataset</th></tr><tr><th>Education Level</th><th>Performance Review</th><th>Target (Promotion)</th></tr><tr><td>High School</td><td>Poor</td><td>No</td></tr><tr><td>College</td><td>Medium</td><td>Yes</td></tr><tr><td>High School</td><td>Good</td><td>Yes</td></tr><tr><td>College</td><td>Poor</td><td>No</td></tr></table>	Table Q7(b) HR Dataset			Education Level	Performance Review	Target (Promotion)	High School	Poor	No	College	Medium	Yes	High School	Good	Yes	College	Poor	No	10	L3	CO3
Table Q7(b) HR Dataset																							
Education Level	Performance Review	Target (Promotion)																					
High School	Poor	No																					
College	Medium	Yes																					
High School	Good	Yes																					
College	Poor	No																					
OR																							
Q.8	a.	Discuss the drawback of ID3 algorithm. Explain how it is addressed in C4.5 algorithm. Also write the C4.5 algorithm.	10	L2	CO3																		
	b.	Assess a Person's performance using the Naïve Bayes classifier to the HR data in Table Q7(b). Predict Whether a person who has completed High School and is performing medium would be promoted or not.	10	L3	CO4																		

Module – 5					
Q.9	a.	Compare clustering and classification. Write the K-means clustering algorithm.	10	L2	CO5
	b.	Design a simple perceptron to perform the Boolean AND function. Use the following initial parameters : Initial weights : $w_1 = 0.4, w_2 = 0.3$, Bias : $b = -0.4$, Learning rate $\eta = 0.1$, Activation function : step function $y = \begin{cases} 1 & z \geq 0 \\ 0 & z < 0 \end{cases}$	10	L3	CO4
OR					
Q.10	a.	What is an Artificial Neural Network (ANN)? Discuss any three applications, advantages and limitations of ANN.	10	L2	CO4
	b.	i) Suppose the coordinates of the objects are (0, 3) and (5, 8). Compute the Euclidean, Manhattan and Chebyshev distances between the objects ii) If the given vectors are $x = (1, 0, 0)$ and $y = (1, 1, 1)$ then find the SMC and Jaccard coefficient. SMC Simple Matching Coefficient.	10	L3	CO5
