K. J. Somaiya Institute of Technology, Sion, Mumbai-22
(Autonomous College Affiliated to University of Mumbai)

Supplementary Exam Nov—Dec 2025 Jan 2026.

(B. Tech / M. Tech.) Program: Artificial Intelligence & Data Science Scheme :-III
Regular Examination: TY Semester: V
Course Code: AIC501 and Course Name: Machine Learning

Date of Exam: ~~24/11/2025~~ 28\01\2026   Duration: 02.5 Hours   Max. Marks: 60

Instructions:
(1) All questions are compulsory.
(2) Draw neat diagrams wherever applicable.
(3) Assume suitable data, if necessary.

Q. No.	Question	Max. Marks	CO	BT level
Q 1	**Solve any two questions out of three: (05 marks each)**	10		
a)	Write applications of machine learning in different domains. Elaborate with example, how machine learning is useful in solving the problem.		1	Un
b)	Write expression for hypothesis, cost function, and for parameters using gradient descent for multivariate linear regression. Explain each term in short.		2	Un
c)	How is the Bayesian Belief network different from the naïve Bayes classifier?		3	Un
Q 2	**Solve any two questions out of three: (05 marks each)**	10		

a)	Define
	   a) Accuracy
	   b) Precision
	   c) Recall
	   d) F1-Score
	   e) Specificity
			1	Un

b)	Explain the concept of "within-cluster sum of squares" (WCSS).		5	Un
c)	Differentiate between Supervised Learning and Unsupervised Learning		1	Un
Q.3	**Solve any two questions out of three. (10 marks each)**	20		
a)	A researcher records the **temperature (°C)** and the **average battery life (hours)** of a smartphone under different temperatures:		2	Ap

A researcher records the **temperature (°C)** and the **average battery life (hours)** of a smartphone under different temperatures:

Data Field	1	2	3	4	5	6
Temperature (oC)	10	15	20	25	30	35
Battery Life (hours)	14.5	13.8	13.2	12.5	11.8	11.2

1. Find the least squares regression line y=ax+b that predicts battery life from temperature.
2. Use this model to estimate the expected battery life when the temperature is **40°C**.

b)	Find the optimal hyperplane that separates the following two classes of data points: Class 1: {(1,2), (2,2), (1,1), (2,1)} Class 2: {(4,3), (5,2), (5,4), (6,3)}	4	Ap

| c) | Find the clusters using single link (nearest neighbour) technique. Use Euclidean distance and draw a dendrogram. | 5 | Ap |

Data Field	Q1	Q2	Q3	Q4	Q5	Q6
X	0.1	0.15	0.3	0.42	0.55	0.6
Y	0.25	0.45	0.35	0.2	0.4	0.25

Q.4   **Solve any two questions out of three. (10 marks each)**   20

| a) | Write expressions for hypothesis, cost function, and parameter using gradient descent for univariate linear regression. The dataset is a record of the square feet area and house price. Assume a learning rate of 0.01. Calculate values of θ_0 and θ_1 after 2 iterations using gradient descent optimization of linear regression. | 2 | Ap |

Data Field	1	2	3	4
Square Feet Area	1400	1600	1700	1875
House Price (Lakhs)	2	2.3	2.45	2.75

| b) | Construct a decision tree using the ID3 algorithm for the dataset given below:- | 3 | Ap |

Data Field	1	2	3	4	5	6
Classification	Pass	Fail	Pass	Fail	Pass	Fail
Study Hours	High	Low	Medium	Low	High	Medium
Attendance	Regular	Irregular	Regular	Regular	Irregular	Irregular

| c) | Use Principal Component Analysis (PCA) to arrive at the transformed matrix of the given data.

 $$B = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 5 & 4 & 2 & 1 \end{bmatrix}$$ | 6 | Ap |

****