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

July/Aug Nov—Dec 2024-25
(B. Tech / M. Tech) Program: B.Tech   Scheme I/II/IIB/III: IIB
Supplementary Regular Examination: TY Semester: VI
Course Code: AIC602 and Course Name: Machine Learning
Date of Exam: 22/05/2025 31/07/25 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)	Differentiate between the Supervised, Unsupervised and Reinforcement Learning with example.		CO1	Un
b)	Define Accuracy, Precision, Recall, F1-Score, and Specificity.		CO1	Un
c)	Explain logistic regression with an example.		CO2	Un
Q 2	**Solve any two questions out of three: (05 marks each)**	10		
a)	What is Principal Component Analysis (PCA), and what are its primary goals?		CO5	Un
b)	Discuss and articulate the role of learning rate (α) in training machine learning algorithms. How does the learning rate affect convergence? Support your answer with appropriate graphs.		CO4	Un
c)	Use **Naive Bayes Classifier** to classify the following tuple: X = (Outlook=Sunny, Temperature = Hot)		CO3	Ap
Q.3	**Solve any two questions out of three. (10 marks each)**	20		
a)	Explain in brief the working of SVM.		CO3	Un

Use **Naive Bayes Classifier** to classify the following tuple:
X = (Outlook=Sunny, Temperature = Hot)

Outlook	Temperature	Play tennis
Sunny	Hot	No
Sunny	Hot	No
Overcast	Hot	Yes
Rain	Mild	Yes
Rain	Cool	Yes
Overcast	Cool	Yes

July/Aug ~~Nov – Dec~~ 2024-25
(B. Tech / M. Tech.) Program: B.Tech   Scheme I/II/IIB/III: IIB
Supplementary ~~Regular~~ Examination: TY Semester: VI
Course Code: AIC602 and Course Name: Machine Learning
Date of Exam: ~~22/05/2025~~ 31/7/25 Duration: 02.5 Hours   Max. Marks: 60

											CO1	Un

b)	1. Explain ROC and AUC curve in ML with examples 2. A binary classification model is used to detect whether a patient has a rare genetic disorder (Positive cia ss) or not (Negative class). The model's performance on a test dataset is summarized as follows: True Negatives (TN): 95, True Positives (TP): 18, False Positives (FP): 4, False Negatives (FN): 13 Calculate the following performance metrics: Accuracy, Precision, Recall, F1-Score, Specificity.		

			CO3	Ap

c) What is the main difference between *k-Means* and *k-Nearest Neighbours*?

Age	20	32	18	29	47	45	46	48	45
Salary	86k	18k	82k	8k	25k	26k	28k	29k	22k
Bought	0	0	0	0	1	1	1	1	1

Using KNN, predict whether a person with age 30 and salary 67000 will buy a car or not. Consider k=3

Q.4	Solve any two questions out of three. (10 marks each)	20		
			CO4	Un
a)	Write a short note on derivative free optimization methods.			
			CO5	Ap

b) Use Principal Component Analysis (PCA) to arrive at the transformed matrix for the given data.

$$X = \begin{bmatrix} 2.5 & 2.4 \\ 0.5 & 0.7 \\ 2.2 & 2.9 \end{bmatrix}$$

How does PCA differ from other dimensionality reduction techniques like Linear Discriminant Analysis (LDA)?

		CO2	Un

c) What is pruning in Decision Trees? Build a decision tree for the given dataset using CART.

Instance	a1	a2	a3	Classification
1	True	Hot	High	No
2	True	Hot	High	No
3	False	Hot	High	Yes
4	False	Cool	Normal	Yes
5	False	Cool	Normal	Yes
6	True	Cool	High	No
7	True	Hot	High	No
8	True	Hot	Normal	Yes
9	False	Cool	Normal	Yes
10	False	Cool	High	Yes

****