

Nov – Dec 2025
(B. Tech.) Program: Computer Engineering Scheme : II/ IIB
Regular Examination: LY Semester: VII
Course Code: CE 701 and Course Name: Machine Learning
Date of Exam: 29/11/2025
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)	What is Bootstrap Aggregation? Explain its role in ensemble learning.		CO3	U													
b)	Alarm1 Alarm2 Burglary ii) Consider the given Bayesian network. Assume that: $P(\text{Alarm1}) = 0.1$ $P(\text{Alarm2}) = 0.2$ $P(\text{Burglary} \mid \text{Alarm1}, \text{Alarm2}) = 0.8$ $P(\text{Burglary} \mid \text{Alarm1}, \neg \text{Alarm2}) = 0.7$ $P(\text{Burglary} \mid \neg \text{Alarm1}, \text{Alarm2}) = 0.6$ $P(\text{Burglary} \mid \neg \text{Alarm1}, \neg \text{Alarm2}) = 0.5$ Calculate $P(\text{Alarm2} \mid \text{Burglary}, \text{Alarm1})$. Show all of your reasoning.		CO4	Ap													
c)	Explain Model based clustering with suitable examples and diagrams.		CO5	U													
Q 2	Solve any two questions out of three: (05 marks each)	10															
a)	Find the mean of squared error for the given predictions: <table><tr><th>Y</th><th>f(x)</th></tr><tr><td>1</td><td>2</td></tr><tr><td>2</td><td>3</td></tr><tr><td>4</td><td>5</td></tr><tr><td>8</td><td>9</td></tr><tr><td>16</td><td>15</td></tr><tr><td>32</td><td>31</td></tr></table>		Y	f(x)	1	2	2	3	4	5	8	9	16	15	32	31	CO5
Y	f(x)																
1	2																
2	3																
4	5																
8	9																
16	15																
32	31																

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

Nov – Dec 2025

(B. Tech.) Program: Computer Engineering Scheme : II/ IIB
Regular Examination: LY Semester: VII

Course Code: CE 701 and Course Name: Machine Learning

Max. Marks: 60

Date of Exam: 29/11/2025

Duration: 02.5 Hours

Date of Exam: _____

	Hint: Find the squared error for each prediction and take the mean of that.																																																																																													
b)	Explain kernels in SVM along with their applications like Linear Kernel, Polynomial Kernel and RBF Kernel.		CO4	U																																																																																										
c)	Compute following Metrics : Find Accuracy, Precision, Recall ,F1-Score , Specificity <table> <tr> <td></td> <td>Predicted Spam</td> <td>Predicted Not Spam</td> </tr> <tr> <td>Actual Spam</td> <td>90</td> <td>30</td> </tr> <tr> <td>Actual Not Spam</td> <td>20</td> <td>160</td> </tr> </table>		Predicted Spam	Predicted Not Spam	Actual Spam	90	30	Actual Not Spam	20	160		CO3	Ap																																																																																	
	Predicted Spam	Predicted Not Spam																																																																																												
Actual Spam	90	30																																																																																												
Actual Not Spam	20	160																																																																																												
Q.3	Solve any two questions out of three. (10 marks each)	20																																																																																												
a)	Apply CART decision tree (using Gini index) algorithm for the following table. Find the Gini index of age and income. Class-Labeled Training Tuples from the <i>AllElectronics</i> Customer Database <table> <tr> <th>RID</th> <th>age</th> <th>income</th> <th>student</th> <th>credit_rating</th> <th>Class: buys_computer</th> </tr> <tr><td>1</td><td>youth</td><td>high</td><td>no</td><td>fair</td><td>no</td></tr> <tr><td>2</td><td>youth</td><td>high</td><td>no</td><td>excellent</td><td>no</td></tr> <tr><td>3</td><td>middle_aged</td><td>high</td><td>no</td><td>fair</td><td>yes</td></tr> <tr><td>4</td><td>senior</td><td>medium</td><td>no</td><td>fair</td><td>yes</td></tr> <tr><td>5</td><td>senior</td><td>low</td><td>yes</td><td>fair</td><td>yes</td></tr> <tr><td>6</td><td>senior</td><td>low</td><td>yes</td><td>excellent</td><td>no</td></tr> <tr><td>7</td><td>middle_aged</td><td>low</td><td>yes</td><td>excellent</td><td>yes</td></tr> <tr><td>8</td><td>youth</td><td>medium</td><td>no</td><td>fair</td><td>no</td></tr> <tr><td>9</td><td>youth</td><td>low</td><td>yes</td><td>fair</td><td>yes</td></tr> <tr><td>10</td><td>senior</td><td>medium</td><td>yes</td><td>fair</td><td>yes</td></tr> <tr><td>11</td><td>youth</td><td>medium</td><td>yes</td><td>excellent</td><td>yes</td></tr> <tr><td>12</td><td>middle_aged</td><td>medium</td><td>no</td><td>excellent</td><td>yes</td></tr> <tr><td>13</td><td>middle_aged</td><td>high</td><td>yes</td><td>fair</td><td>yes</td></tr> <tr><td>14</td><td>senior</td><td>medium</td><td>no</td><td>excellent</td><td>no</td></tr> </table>	RID	age	income	student	credit_rating	Class: buys_computer	1	youth	high	no	fair	no	2	youth	high	no	excellent	no	3	middle_aged	high	no	fair	yes	4	senior	medium	no	fair	yes	5	senior	low	yes	fair	yes	6	senior	low	yes	excellent	no	7	middle_aged	low	yes	excellent	yes	8	youth	medium	no	fair	no	9	youth	low	yes	fair	yes	10	senior	medium	yes	fair	yes	11	youth	medium	yes	excellent	yes	12	middle_aged	medium	no	excellent	yes	13	middle_aged	high	yes	fair	yes	14	senior	medium	no	excellent	no		CO3	Ap
RID	age	income	student	credit_rating	Class: buys_computer																																																																																									
1	youth	high	no	fair	no																																																																																									
2	youth	high	no	excellent	no																																																																																									
3	middle_aged	high	no	fair	yes																																																																																									
4	senior	medium	no	fair	yes																																																																																									
5	senior	low	yes	fair	yes																																																																																									
6	senior	low	yes	excellent	no																																																																																									
7	middle_aged	low	yes	excellent	yes																																																																																									
8	youth	medium	no	fair	no																																																																																									
9	youth	low	yes	fair	yes																																																																																									
10	senior	medium	yes	fair	yes																																																																																									
11	youth	medium	yes	excellent	yes																																																																																									
12	middle_aged	medium	no	excellent	yes																																																																																									
13	middle_aged	high	yes	fair	yes																																																																																									
14	senior	medium	no	excellent	no																																																																																									
b)	Compare overfitting, underfitting & best fitting model w.r.t model complexity , Bias, Variance , Training error, Test Error (5 mk)		CO1	U																																																																																										

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

Nov – Dec 2025

(B. Tech.) Program: Computer Engineering Scheme : II/ IIB
Regular Examination: LY Semester: VII

Course Code: CE 701 and Course Name: Machine Learning

Duration: 02.5 Hours

Max. Marks: 60

Date of Exam: 29/11/2025

	For a document clustering case study, explain the step-by-step machine learning development process. (5 mk)			
c)	Describe graph-based clustering and explain the different ways of constructing similarity graphs (ϵ -neighborhood, k-NN graph, fully connected graph).		CO5	U
Q.4	Solve any two questions out of three. (10 marks each)	20		
a)	Differentiate between a hard margin SVM and soft margin SVM.(5mk) Describe Hinge loss used in SVM. (5mk)		CO4	U
b)	Explain Logistic Regression with suitable example (5mk) Explain ROC, AUC curve with respect to Logistic regression performance metrics with suitable examples. (5mk)		CO2	U
c)	Explain PCA, LDA, and SVD. Write a detailed comparison among the three with respect to: - Type of method - Objective - Mathematical method - Output components - Projection direction - Dimensionality reduction ability - Use case - Handle class information? - Applications		CO6	U
