

22/12/2023

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

End Semester Exam

Nov-Dec 2023

(B.Tech/M.Tech.) Program: **M.Tech.**

Examination: FY Semester: I

Course Code: **PCEC102** and Course Name: **Machine Learning and Pattern Recognition**

Duration: 2.5 Hours

Max. Marks: 60

Instructions:

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

		Max. Marks	CO	BT level
Q 1	Solve any six questions out of eight:	12		
i)	What are the advantages of C4.5 over ID3 classification algorithm?	2	CO3	U
ii)	List different performance evaluation matrix for ML algorithms	2	CO4	U
iii)	Why KNN is a lazy learning and non-parametric algorithm.	2	CO1	U
iv)	Why Naïve Bayesian Classifier is called "naive" and "bayes" ?	2	CO4	U
v)	Explain Fisher Linear discriminant function.	2	CO2	U
vi)	Define association rule mining. List different algorithms.	2	CO6	U
vii)	Difference between Bagging and Boosting.	2	CO5	U
viii)	Explain R-squared in Linear regression with example.	2	CO3	U
Q.2	Solve any four questions out of six.	16		
i)	Differentiate between feature selection and feature extraction?	4	CO3	U
ii)	Calculate Precision, Recall and F1-score with example.	4	CO1	U
iii)	What is Single linkage, complete linkage and average linkage clustering?	4	CO4	U

iv)	Compare Supervised, unsupervised and reinforcement learning techniques.	4	CO1	U																																																							
v)	Explain Principal component analysis for dimensionality reduction.	4	CO2	U																																																							
vi)	Compare with example overfit, underfit & best fit models.	4	CO4	U																																																							
Q.3	Solve any two questions out of three.	16																																																									
i)	Perform Linear discriminant analysis for the given 2-D dataset <table border="1" style="margin-left: 20px;"> <tr> <td>C1 -> X1=(X1, X2)</td> <td>(4,2)</td> <td>(2,4)</td> <td>(2,3)</td> <td>(3,6)</td> <td>(4,4)</td> </tr> <tr> <td>C2 -> X2=(X1, X2)</td> <td>(9,10)</td> <td>(6,8)</td> <td>(9,5)</td> <td>(8,7)</td> <td>(10,8)</td> </tr> </table>	C1 -> X1=(X1, X2)	(4,2)	(2,4)	(2,3)	(3,6)	(4,4)	C2 -> X2=(X1, X2)	(9,10)	(6,8)	(9,5)	(8,7)	(10,8)	8	CO2	Ap																																											
C1 -> X1=(X1, X2)	(4,2)	(2,4)	(2,3)	(3,6)	(4,4)																																																						
C2 -> X2=(X1, X2)	(9,10)	(6,8)	(9,5)	(8,7)	(10,8)																																																						
ii)	What is Hyperplane and Support Vectors in the SVM algorithm? What is the use of SVM kernel?	8	CO4	U																																																							
iii)	Using the given training dataset classify the following tuple using Naïve Bayes Algorithm: (Homeowner = yes ; Status =Employed; Income=Average)	8	CO4	Ap																																																							
	<table border="1" style="margin-left: 20px;"> <thead> <tr> <th>Id</th> <th>Homeowner</th> <th>Status</th> <th>Income</th> <th>Defaulted</th> </tr> </thead> <tbody> <tr><td>1</td><td>Yes</td><td>Employed</td><td>High</td><td>No</td></tr> <tr><td>2</td><td>No</td><td>Business</td><td>Average</td><td>No</td></tr> <tr><td>3</td><td>No</td><td>Employed</td><td>Low</td><td>No</td></tr> <tr><td>4</td><td>Yes</td><td>Business</td><td>High</td><td>No</td></tr> <tr><td>5</td><td>No</td><td>Unemployed</td><td>Average</td><td>Yes</td></tr> <tr><td>6</td><td>No</td><td>Business</td><td>Low</td><td>No</td></tr> <tr><td>7</td><td>Yes</td><td>Unemployed</td><td>High</td><td>No</td></tr> <tr><td>8</td><td>No</td><td>Employed</td><td>Average</td><td>Yes</td></tr> <tr><td>9</td><td>No</td><td>Business</td><td>Low</td><td>No</td></tr> <tr><td>10</td><td>No</td><td>Employed</td><td>Average</td><td>Yes</td></tr> </tbody> </table>	Id	Homeowner	Status	Income	Defaulted	1	Yes	Employed	High	No	2	No	Business	Average	No	3	No	Employed	Low	No	4	Yes	Business	High	No	5	No	Unemployed	Average	Yes	6	No	Business	Low	No	7	Yes	Unemployed	High	No	8	No	Employed	Average	Yes	9	No	Business	Low	No	10	No	Employed	Average	Yes			
Id	Homeowner	Status	Income	Defaulted																																																							
1	Yes	Employed	High	No																																																							
2	No	Business	Average	No																																																							
3	No	Employed	Low	No																																																							
4	Yes	Business	High	No																																																							
5	No	Unemployed	Average	Yes																																																							
6	No	Business	Low	No																																																							
7	Yes	Unemployed	High	No																																																							
8	No	Employed	Average	Yes																																																							
9	No	Business	Low	No																																																							
10	No	Employed	Average	Yes																																																							
Q.4	Solve any two questions out of three.	16																																																									

i)	Describe the working of Random Forest machine learning algorithm with example.	8	CO5	U																																																																																																
ii)	<p>Apply CART(decision tree) algorithm for the following table. Solve upto second level of tree generation.</p> <p style="text-align: center;">Our Data: Loan Approval Prediction</p> <table border="1" data-bbox="240 437 1086 898"> <thead> <tr> <th>ID</th> <th>AGE</th> <th>JOB_STATUS</th> <th>OWNS_HOUSE</th> <th>CREDIT_RATING</th> <th>CLASS (Yes or No)</th> </tr> </thead> <tbody> <tr><td>1</td><td>Young</td><td>False</td><td>False</td><td>Fair</td><td>No</td></tr> <tr><td>2</td><td>Young</td><td>False</td><td>False</td><td>Good</td><td>No</td></tr> <tr><td>3</td><td>Young</td><td>True</td><td>False</td><td>Good</td><td>Yes</td></tr> <tr><td>4</td><td>Young</td><td>True</td><td>True</td><td>Fair</td><td>Yes</td></tr> <tr><td>5</td><td>Young</td><td>False</td><td>False</td><td>Fair</td><td>No</td></tr> <tr><td>6</td><td>Middle</td><td>False</td><td>False</td><td>Fair</td><td>No</td></tr> <tr><td>7</td><td>Middle</td><td>False</td><td>False</td><td>Good</td><td>No</td></tr> <tr><td>8</td><td>Middle</td><td>True</td><td>True</td><td>Good</td><td>Yes</td></tr> <tr><td>9</td><td>Middle</td><td>False</td><td>True</td><td>Excellent</td><td>Yes</td></tr> <tr><td>10</td><td>Middle</td><td>False</td><td>True</td><td>Excellent</td><td>Yes</td></tr> <tr><td>11</td><td>Old</td><td>False</td><td>True</td><td>Excellent</td><td>Yes</td></tr> <tr><td>12</td><td>Old</td><td>False</td><td>True</td><td>Good</td><td>Yes</td></tr> <tr><td>13</td><td>Old</td><td>True</td><td>False</td><td>Good</td><td>Yes</td></tr> <tr><td>14</td><td>Old</td><td>True</td><td>False</td><td>Excellent</td><td>Yes</td></tr> <tr><td>15</td><td>Old</td><td>False</td><td>False</td><td>Fair</td><td>No</td></tr> </tbody> </table>	ID	AGE	JOB_STATUS	OWNS_HOUSE	CREDIT_RATING	CLASS (Yes or No)	1	Young	False	False	Fair	No	2	Young	False	False	Good	No	3	Young	True	False	Good	Yes	4	Young	True	True	Fair	Yes	5	Young	False	False	Fair	No	6	Middle	False	False	Fair	No	7	Middle	False	False	Good	No	8	Middle	True	True	Good	Yes	9	Middle	False	True	Excellent	Yes	10	Middle	False	True	Excellent	Yes	11	Old	False	True	Excellent	Yes	12	Old	False	True	Good	Yes	13	Old	True	False	Good	Yes	14	Old	True	False	Excellent	Yes	15	Old	False	False	Fair	No	8	CO3	Ap
ID	AGE	JOB_STATUS	OWNS_HOUSE	CREDIT_RATING	CLASS (Yes or No)																																																																																															
1	Young	False	False	Fair	No																																																																																															
2	Young	False	False	Good	No																																																																																															
3	Young	True	False	Good	Yes																																																																																															
4	Young	True	True	Fair	Yes																																																																																															
5	Young	False	False	Fair	No																																																																																															
6	Middle	False	False	Fair	No																																																																																															
7	Middle	False	False	Good	No																																																																																															
8	Middle	True	True	Good	Yes																																																																																															
9	Middle	False	True	Excellent	Yes																																																																																															
10	Middle	False	True	Excellent	Yes																																																																																															
11	Old	False	True	Excellent	Yes																																																																																															
12	Old	False	True	Good	Yes																																																																																															
13	Old	True	False	Good	Yes																																																																																															
14	Old	True	False	Excellent	Yes																																																																																															
15	Old	False	False	Fair	No																																																																																															
iii)	What do you mean by Hypothesis space and inductive bias in machine learning. Also explain Model visualization?	8	CO1	U																																																																																																