RollNo. | | | | | | | | | | |

ANNA UNIVERSITY (UNIVERSITY DEPARTMENTS)
B.E. /B.Tech. / B. Arch. (Full Time) - END SEMESTER EXAMINATIONS, NOV. / DEC. 2023

INFORMATION TECHNOLOGY
Seventh Semester

IT5036 Machine Learning
(Regulation 2019)

Time:3hrs Max.Marks: 100

CO1	Choose and implement classification or regression algorithms for an application using an open source tool.
CO2	Implement probabilistic discriminative and generative algorithms for an application and analyze the results.
CO3	Use a tool to implement typical clustering algorithms for different types of applications.
CO4	Design and implement an HMM for a sequence model type of application.
CO5	Implement appropriate learning algorithms for any real time application using an open source tool.
CO6	Identify applications suitable for different types of machine learning with suitable justification.

BL – Bloom's Taxonomy Levels
(L1-Remembering, L2-Understanding, L3-Applying, L4-Analysing, L5-Evaluating, L6-Creating)

PART- A (10x2=20Marks)
(Answer all Questions)

Q. No.	Questions	Marks	CO	BL
1	What is the need for machine learning in real time applications?	**2**	1	2
2	Give the key differences between classification and regression.	**2**	1	2
3	Comment on bias-variance trade-off.	**2**	2	2
4	List the metrics used for validation of regression methods.	**2**	2	1
5	State the need for Expectation-Maximization algorithm.	**2**	3	2
6	Write briefly about LDA.	**2**	3	1
7	Differentiate Bayesian Network and Naive Bayes.	**2**	4	2
8	List out the characteristics of Markov Model.	**2**	4	1
9	How reinforcement learning is different from supervised and unsupervised learning?	**2**	5	2
10	Compare bootstrap and boosting.	**2**	5	2

PART- B (5x 13=65Marks)
(Restrict to a maximum of 2 subdivisions)

Q. No.	Questions	Marks	CO	BL
11 (a)	Consider the following dataset in which the week and number of working hours per week spent by a student in a library are tabulated. Apply linear regression technique to predict the no. of hours that will be spent by the student in the 7th and 9th week.	**13**	2	3

$X_{i\ (Week)}$	1	2	3	4	5
$Y_{i(Hours\ Spent)}$	12	18	22	28	35

OR

11 (b)	Determine the first splitting attribute for decision tree by using ID3 algorithm with the following training dataset.	**13**	2	3

S.No.	Assessment	Assignment	Project	Seminar	Result
1.	Good	Yes	Yes	Good	Pass

	2.	Average	Yes	No	Poor	Fail			
	3.	Good	No	Yes	Good	Pass			
	4.	Poor	No	No	Poor	Fail			
	5.	Good	Yes	Yes	Good	Pass			
	6.	Average	No	Yes	Good	Pass			
	7.	Good	No	No	Fair	Pass			
	8.	Poor	Yes	Yes	Good	Fail			
	9.	Average	No	No	Poor	Fail			
	10.	Good	Yes	Yes	Fair	Pass			

12 (a)	Consider the student performance training dataset given below: Classify the test data (7.5, 75, 7) using KNN algorithm (K = 3).	**13**	2	3

S.No.	CGPA	Assessment	Project	Result
1.	9.2	85	8	Pass
2.	8.0	80	7	Pass
3.	8.5	81	8	Pass
4.	6	45	5	Fail
5.	6.5	50	4	Fail
6.	8.2	72	7	Pass
7.	5.8	38	5	Fail
8.	8.9	91	9	Pass
9.	6.1	40	5	Fail
10.	7.0	82	8	Pass

OR

12 (b)	For the given points (3,1), (3,2) and (4,0) belong to class positive and points (1,0), (0,1) and (0,-1) belong to class negative, draw an optimal hyperplane to classify the points.	**13**	2	3
13 (a)	Use single and complete link agglomerative clustering to group the data described by the following distance matrix. Show the dendrograms.	**13**	3	3

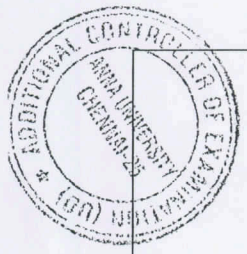
	A	B	C	D
A	0	1	4	5
B		0	2	6
C			0	3
D				0

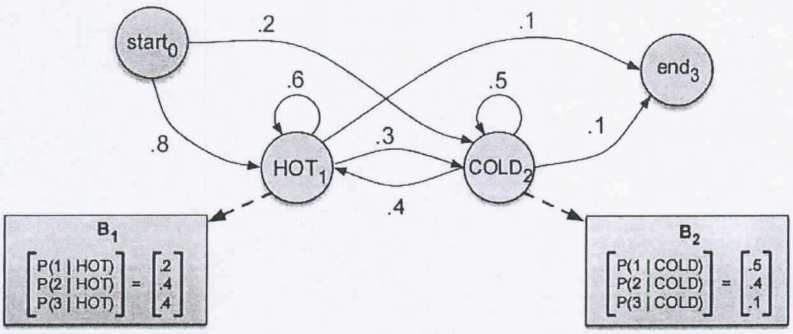
OR

13 (b)	Let the data points be (2,6) and (1,7). Apply PCA and find the transformed data.	**13**	3	3
14 (a)	Consider the training data set given in the Question 11(b). Predict the result of a student using Naive Bayes algorithm for the test data given below. X = (Assessment = Average, Assignment = Yes, Project = No and Seminar = Good)	**13**	4	3

OR

14 (b)	Consider the Hidden Markov Model given below with transition and emission probability.	**13**	4	3





Compute hidden state sequence for the given observation sequence '332'.

15 (a)	Design a Multi-layer feed-forward neural network with the backpropagation algorithm, whose class label is 1 and learning rate is 0.8. Use appropriate activation function. Draw the network neatly. Initial Input, weight and bias values:	**13**	5	3

X_1	X_2	X_3	W_{14}	W_{15}	W_{24}	W_{25}
1	1	0	0.2	-0.4	0.3	0.2

W_{34}	W_{35}	W_{46}	W_{56}	θ_4	θ_5	θ_6
-0.5	0.2	-0.2	-0.1	-0.5	0.3	0.1

OR

15 (b)	Apply convolution for the following image data using the mask and show the feature or activation map. Image Data Mask	**13**	5	3

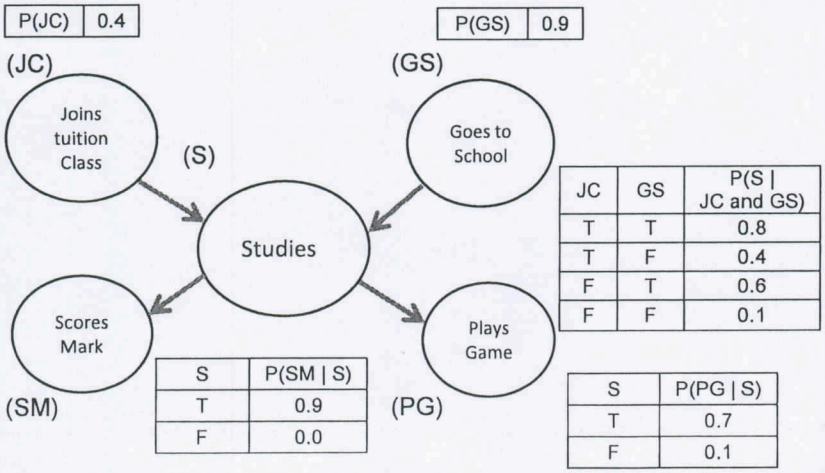
10	11	12	13
3	4	6	7
8	9	10	11
11	12	13	14

1	1
1	1

PART- C (1x 15=15Marks)

(Q.No.16 is compulsory)

Q. No.	Questions	Marks	CO	BL
16.	Consider the scenario shown in Figure below. Events that 'a student joins tuition classes' and goes to school daily' have a direct effect on how the 'student studies'. The event that the student has a direct effect on his/her scoring marks or playing games. What is the probability that he/she does not join tuition class, he goes to school daily, he studies, and he does not score marks?	**15**	4	3



JC	GS	P(S | JC and GS)
T	T	0.8
T	F	0.4
F	T	0.6
F	F	0.1

S	P(SM | S)
T	0.9
F	0.0

S	P(PG | S)
T	0.7
F	0.1