Printed Pages – 6 Roll No. 20 0 0 32 0 2 10 45

C128571(022)

B. Tech. (Hon's) (Fifth Semester) Examination, Nov.-Dec. 2023

(AICTE Scheme)

(Data Science Engg. Branch)

PATTERN RECOGNITION and MACHINE LEARNING

Time Allowed : Three hours

Maximum Marks : 100

Minimum Pass Marks : 35

Note : *Each question contains four parts. Part (a) of each question is compulsory. Attempt any two parts from (b), (c) and (d) of each question. The figure in the right-hand margin indicates marks.*

Unit-I

1. (a) What is Learning? Design a learning system for checkers problem. 4

   (b) What do you mean by loss function machine learning? Explain loss function in linear regression. 8

(c) State Bayes theorem. Consider a scenario where your yearly checkup is done, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease, and that the test is 99% accurate (i.e., the probability of testing positive given that you have the disease is 0·99, as it is the probability of testing negative given that you have the disease). This is a rare disease, striking only one in 10,000 people. Why is it good news that the disease is rare? What are the chances that you actually have the disease? 8

(d) Define multi-dimensional space. Explain Normed vector and dot product space in brief. 8

Unit-II

2. (a) Explain Maximum entropy estimation in brief. 4

(b) Suppose we wish to include the continuous-valued attribute temperature in describing the training example days in the learning task of following table. Incorporate this continuous value to discrete value for constructing decision tree. 8

F	f1	f2	f3	f4	f5	f6
Temperature	40	48	60	72	80	90
Play Tannis	No	No	Yes	Yes	Yes	No

(c) Explain maximum likelihood estimation? Consider an unfair coin is flipped 100 times. 61 head are observed. The coin either has a probability 1/3, 1/2 and 2/3 of flipping a head each time. Find which of the three is MLE? 8

(d) What is Mutual information estimation? Explain Akaike and Bayesian information criterion. 8

Unit-III

3. (a) Briefly explain discriminant function. 4

(b) Suppose that the initial seeds are $A1$, $A4$ and $A7$. Run the k-means clustering for following 8 examples. Iterate this algorithm till convergence and mention centroid with cluster in each iteration. 8

$A1 = (2, 10)$, $A2 = (2, 5)$, $A3 = (8, 4)$, $A4 = (5, 8)$, $A5 = (7, 5)$, $A6 = (6, 4)$, $A7 = (1, 2)$, $A8 = (4, 9)$

(c) Explain Fishers linear discrimination brief. Discuss its merits over PCA. 8

(d) Explain Probabilistic Discriminative Models. Write down difference between Generative and Discriminative probabilistic model. 8

Unit-IV

4. (a) Briefly explain semi supervised algorithm with suitable example. 4

(b) Consider following data sample on effect of hours of mixing on temperature of wood pulp. Predict temperature of wood pulp for hour (x) = 13. 8

S. No.	1	2	3	4	5	6
Hours of mixing (X)	2	4	6	8	10	12
Temperature of wood pulp (Y)	21	27	29	64	86	92

(c) Consider following data sample with two cluster. Implement fuzzy c means clustering. Iterate this algorithm for at least two iteration. 8

	(1, 3)	(2, 5)	(4, 8)	(7, 9)
Cluster 1	0·8	0·7	0·2	0·1
Cluster 2	0·2	0·3	0·8	0·9

(d) What is rough k-means? Explain its algorithm. 8

Unit-V

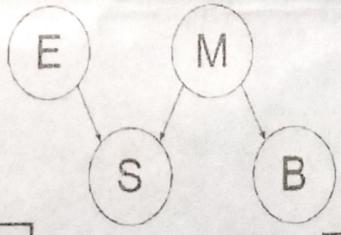
5. (a) What is two class problem? Explain different method to solve this problem in brief. 4

(b) What is Support Vector Machine? Derive an expression for maximum marginal hyperplane. 8

(c) What is Kernal function? Explain the linear and Fisher Kernal methods in machine learning. 8

(d) Explain conditional independence? Consider a scenario, where a smell of Sulphur (S) can be caused either by rotten eggs (E) or as a sign of the doom brought by the Mayan Apocalypse (M). The Mayan Apocalypse also causes the oceans to boil (B). The Bayesian network and corresponding conditional probability tables for this situation are shown below. 8

P (E)	
$+e$	0·4
$-e$	0·6



P (M)	
$+m$	0·1
$-m$	0·9

$P(S \mid E, M)$			
$+e$	$+m$	$+s$	1·0
$+e$	$+m$	$-s$	0·0
$+e$	$-m$	$+s$	0·8
$+e$	$-m$	$-s$	0·2
$-e$	$+m$	$+s$	0·3
$-e$	$+m$	$-s$	0·7
$-e$	$-m$	$+s$	0·1
$-e$	$-m$	$-s$	0·9

$P(B \mid M)$		
$+m$	$+b$	1·0
$+m$	$-b$	0·0
$-m$	$+b$	0·1
$-m$	$-b$	0·9

Then find :

(i) What is the probability that the oceans boil?

(ii) What is the probability that the Mayan Apocalypse is occurring, given that there is a smell of Sulphur, the oceans are boiling, and there are rotten eggs?

(iii) What is the probability that the Mayan Apocalypse is occurring, given that the oceans are boiling?