

MANIPAL INSTITUTE OF TECHNOLOGY

(A constituent unit of MAHE, Manipal)

VI SEMESTER B.TECH. (MECHATRONICS ENGINEERING) END SEMESTER EXAMINATIONS, APRIL 2018

SUBJECT: MACHINE LEARNING [MTE 4025]

(20/04/2018)

Time: 3 Hours

MAX. MARKS: 50

Instructions to Candidates:

✤ Answer ALL questions.

Is not Employed

- Data not provided may be suitably assumed
- 1A. Bob needs to make a decision on whether to hire Martin as his secretary. He conducts an interview 6 for Martin and comes up with the following results:

The probability of the performance being good, fair or bad is shown below:

20

Class	Good			Fair		Bad		
P(Class)		0.2	0.4			0.4		
The class conditionals of Martin's employability are tabulated below:								
x Class	Good		Fair			Bad		
Employable	0.8		0.5			0.1		
Not Employable	0.2		0.5			0.9		
Bob does a risk analysis and comes up with the following data of whether or not to employ Martin:								
	Class	Good		Fair		Bad		
$\lambda(a_i Class)$								
Is Employed		0		5		10		

Based on this data can you help Bob decide whether he should hire Martin or not. Use Bayes' Decision theory.

5

- **1B.** Explain the motivation behind, and the usefulness of the technique of Principal Component Analysis. **4**
- 2A. List out the steps to fit the most optimal line to the following dataset using the Linear regression 7 model. Design the hypothesis, arrive at the cost function and apply the Gradient descent algorithm for one iteration to find the values of the parameters of your generated hypothesis.

	1 7		
Length of rice grain (cm)	Weight of rice grain (mg)	% of polishing	Price of 1 kg (Rs.)
1	10	100	70
0.5	5	20	40
0.75	6	33.7	20
0.33	2	50	50

2B. Given below is the list of items purchased from a Tesco outlet. Based upon the given data, perform **3** a Market Basket Analysis to answer the questions that follow:

Transaction 1	Window cleaner, Orange juice, Bananas
Transaction 2	Bananas, Apples, Soda
Transaction 3	Window cleaner, Orange juice, Apples, Soda, Detergent
Transaction 4	Orange juice, Bananas, Apples, Soda
Transaction 5	Window cleaner, Orange juice, Soda

a) In your opinion where should detergents be placed in order to maximize their sales?

b) Are window cleaning products purchased when detergents and orange juice are bought together?

- c) Is soda typically purchased with bananas?
- **3A.** Cite the differences between Probability and Likelihood when considering the process of calculation **2** of the Maximum Likelihood Estimate.
- **3B.** Describe the algorithm to find the Maximum Likelihood Estimate for a distribution.
- **3C.** Construct a neural network to predict the output of the following logic gate diagram.



- **4A.** Given the hyper plane defined by the line $y = x_1 2x_2$, where $y = W^T X = (1, -2)^T X$, are the following **4** points correctly predicted?
 - a) y = 1; x = (1, 0)
 - b) y = 1; x = (1,1)
- **4B.** Why is Naïve Bayes theorem, 'naïve'? Justify using an example.**4**
- **4C.** Differentiate between a Data matrix and Dissimilarity matrix using an example. **2**
- 5A. Elucidate the process of data mining using a flowchart.
- **5B.** Describe the process for finding the decision boundary in a Logistic Regression model for the **4** following hypotheses.

a) $h_{\theta}(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$

b) $h_{\theta}(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1^2 + \theta_4 x_2^2)$

5C. Justify how Ensemble learning helps in improving the accuracy of the predicted outputs. **3**

3

5

3