Course Code : CST 404-2/CST 414-2 **CXDW/RW – 18 / 5106**

Seventh Semester B. E. (Computer Science and Engineering) Examination

Elective – II

MACHINE LEARNING

Time : 3 Hours] [Max. Marks : 60

Instructions to Candidates :—

(1) All questions carry marks as indicated against them.
(2) Assume suitable data wherever necessary.
(3) Illustrate your answers wherever necessary with the help of neat sketches.

1. Attempt any **Two** questions :

(a) Calculate the size of hypothesis space in the M-Disease Diagnosis learning task given in table 1.

ID	Headache	Fever	Vomiting	M-Disease
1.	F	T	F	T
2.	T	F	T	F
3.	T	F	T	F
4.	T	F	T	T

Table 1

How would the number of possible instances and possible hypotheses increase with the addition of attribute **Blood Pressure**, which can take on the values low, normal or high ? More generally, how does the number of possible instances and hypotheses grow with the addition of a new attribute A that takes k possible values. 5 (CO 1)

(b) Apply Candidate-Elimination Algorithm on training examples given in Table 1, and find out the class for ID 5 = (Headache = T, Fever = T, Vomiting = F) using majority voting. 5 (CO 1)

(c) Using ID3 algorithm, find out Decision tree for the data given in Table 1 and find out the class of ID 5 = (Headache = T, Fever = T, Vomiting = F). 5 (CO 1)

2. Attempt any **Two** questions :

(a) Describe Radial Basis Function (RBF) network. Solve X-NOR classification problem using RBF network. 5 (CO 2)

(b) Prove that conjunctions of Boolean Literals are PAC-Learnable. How many randomly drawn examples are sufficient for consistent learner to learn a target concept described by conjunction of up to 9 Boolean Literals, with a 90% probability that it will learn a hypothesis with error less than 0.05. 5 (CO 2)

(c) Give an example (in two dimensional space with coordinate values) in which k-Nearest Neighbors and Distance-Weighted k-Nearest Neighbors. Algorithms classify a query point in to two different classes. 5 (CO 2)

3. Attempt any **Two** questions :-

(a) Find derivatives of unipolar and bipolar sigmoid activation functions in terms of functions. 5 (CO 2)

(b) Derive Error Back Propagation rule for two layers (one hidden layer) feed forward neural network with unipolar activation function for both the layers. 5 (CO 2)

(c) Apply Perceptron learning algorithm for 2-input NOR classification up to four iterations, take initial weight vector $w = [0.1\ -0.1\ 0.2]^T$. 5 (CO 2)

4. Attempt any **Two** questions :

(a) Consider a medical diagnosis problem in which there are two alternative hypotheses : (1) that the patient; has a-particular form of cancer, and (2) that the patient does not. The available data is from a particular laboratory test with two possible outcomes: positive and negative. We have prior knowledge that over the entire population of people only .006 have this disease. Furthermore, the lab test is only an imperfect indicator of the disease. The test returns a correct positive result in only 99% of the cases in which the disease is actually present and a correct negative result in only 97% of the cases in which the disease is not present. In other cases, the test returns the opposite result. Suppose we now observe a new patient for whom the lab test returns a positive result. Should we diagnose the patient as having cancer or not ? Suppose the doctor decides to order a second test for the same patient from the same laboratory, and suppose the second test returns a positive result as well. What are the posterior

probabilities of cancer and not cancer following these two tests ? Assume that the two tests are independent. 5 (CO 3)

(b) An example of training data are as follows given in Table 2 :

X	Y	Z	Class
2	3	2	A
4	1	4	B
1	3	2	A
2	4	3	A
4	2	4	B
2	1	3	B
1	2	4	A
2	3	3	B

Table 2

Apply Naïve Bayes classifier to estimate the class for (x = 2, y = 3, z = 4). 5 (CO 3)

(c) Explain minimum length description principle. 5 (CO 2)

5. Attempt any **Two** questions :

(a) Describe Bayesian network using suitable example. Discuss the advantages of Bayesian networks. 5 (CO 3)

(b) Apply K-means clustering algorithm on following data and identify cluster for each individual. (assume k = 2)

Object	X Weight index	Y pH value
A	1	1
B	2	1
C	4	3
D	5	4

Table 3 5 (CO 3)

(c) Use following dataset and perform hierarchical clustering using complete-linkage algorithm. Show dendogram result.

	X	Y
1	4	4
2	8	4
3	15	8
4	24	4
5	24	12

Table 4 5 (CO 3)

6. Attempt any **Two** questions :

(a) Find α_+ and α_- of Linear SVM for following examples :

Positive Examples = {(–1, 2), (–3, 1), (–3, 3), (–4, 3)}

Negative Examples = {(2, –1), (3, –2), (4, –3), (3, –4)} 5 (CO 4)

(b) Explain with the help of example, bagging and boosting methods used in ensemble learning. 5 (CO 4)

(c) Explain Hidden Markov Models in machine learning; mention any one application of the same. 5 (CO 4)