Course Code : CST 414-2 **KOLP/RW – 19 / 9158**

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

Elective – II

MACHINE LEARNING

Time : 3 Hours] [Max. Marks : 60

Instructions to Candidates :—

(1) Assume suitable data wherever necessary.

(2) Each question carries marks as indicated against them.

1. (a) Apply FIND – S algorithm on following dataset to find maximally specific hypothesis.

hair	body	likesSimon	pose	smile	smart	Target Concept **C(d)**
blond	thin	Yes	Arrogant	Toothy	No	1
brown	thin	No	Natural	Pleasant	Yes	0
blond	plump	Yes	Goofy	Pleasant	No	1
black	thin	No	Arrogant	None	No	0
blond	plump	No	Natural	Toothy	Yes	0

3 (CO 1)

(b) Consider the following dataset of houses represented by five training examples. The target attribute is "Acceptable", this can have values Yes or No, and this to be predicted based on the features of the house :

House	Furniture	Nos. Rooms	New Kitchen	Acceptable
1	No	3	Yes	Yes
2	Yes	3	No	No
3	No	4	No	Yes
4	No	3	No	No
5	Yes	4	No	Yes

Construct the decision tree from the given examples that would be learned by the ID3 algorithm. 7 (CO 1)

2. (a) Consider following dataset and use K – Nearest Neighbor algorithm (k = 3) to predict the weight for test instance, <height = 5·5, age = 38> :

ID#	Height	Age	Weight
1	5	45	77
2	5·11	26	47
3	5·6	30	55
4	5·9	34	59
5	4·8	40	72
6	5·8	36	60
7	5·3	19	40

5 (CO 2)

(b) "Conjunction of Boolean literals is PAC learnable". Justify your answer with a suitable example. 5 (CO 2)

OR

(c) What do you mean by VC dimension ? What does it represent ? Give the VC dimension for following cases with justification :

(i) Specify VC – dimension of intervals in R. The target function is specified by an interval, and labels any example positive iff it lies inside that interval.

(ii) Conjunction of n Boolean literals.

(iii) Linear Support Vector Machines in d – dimensional space. 5 (CO 2)

3. (a) Describe the Back propagation algorithm used to train Neural Networks. 5 (CO 2)

(b) Design the two layer network of perceptrons that implement A XOR B. 5 (CO 2)

4. (a) Let Boolean random variables B stand for "has breast cancer" and M stand for "mammography test in positive". A research study has produced the following three observations :

- The prior probability of having breast cancer is 0·01.
- The probability of testing positive when you have breast cancer is 90%.
- The probability of testing negative when you do not have breast cancer is 89·9%.

(i) Calculate the prior probability of having a positive mammography test P(M).

(ii) If a patient has a positive mammography test, what is the probability that she has breast cancer ? Solve P(B|M) [using Bayes formula]. 5 (CO 3)

(b) What is MDL ? Derive an expression for minimum description length principle. 5 (CO 2)

5. (a) Apply k – means algorithm on given data for k = 2, Use C_1 (2, 4) and C_2 (6, 3) as initial cluster centers.

Data : a (2, 4), b (3, 3), c (5, 5), d (6, 3), e (4, 3), f (6, 6). 5 (CO 3)

(b) Explain Expectation Maximization algorithm with its advantages. 5 (CO 3)

6. (a) Find $\alpha+$ and $\alpha-$ of linear support vector machine for following examples :

Positive examples : {(–1, 2), (–3, 1), (–3, 3) (–4, 3), (–1, –1)}

Negative examples : {(1, 1), (3, –2), (4, –3) (3, –4)} 5 (CO 4)

(b) Describe Hidden Markov model with suitable example. 5 (CO 4)

OR

(c) Explain Ensemble learning with suitable example. 5 (CO 4)