

--	--	--	--	--	--	--	--

# B.Tech. Degree V Semester Regular Examination November 2021

## CS 19-202-0507 MACHINE LEARNING (2019 Scheme)

Time: 3 Hours

Maximum Marks: 60

### PART A

(Answer ALL questions)

(8 × 3 = 24)

- I. (a) Identify procedures for learning multiple classes.
- (b) Discuss underfitting and overfitting with an example.
- (c) Describe EM algorithm for Gaussian mixtures.
- (d) Compare Feature Extraction and Feature Selection techniques.
- (e) Why is kernel density estimation used?
- (f) What is the significance of soft margin hyperplane?
- (g) Differentiate bagging and boosting.
- (h) Discuss Markov process with an example.

### PART B

(4 × 12 = 48)

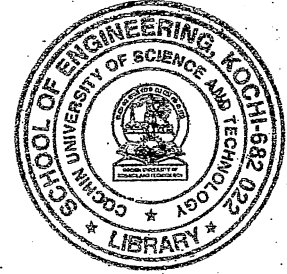
- II. (a) Identify the types of learning in machine learning with suitable example. (9)
- (b) Differentiate regression and interpolation. (3)

OR

- III. Consider two random variables X and Y with the joint PMF given below. (12)

Table: Joint PMF of X and Y

	Y = 0	Y = 1
X = 0	$\frac{1}{5}$	$\frac{2}{5}$
X = 1	$\frac{2}{5}$	0



- (i) Find the linear MMSE estimator of X given Y, ( $\hat{X}_L$ ).
- (ii) Find the MMSE estimator of X given Y, ( $\hat{X}_M$ ).
- (iii) Find the MSE of  $\hat{X}_M$ .

- IV. (a) Cluster the following eight points (with (x, y) representing locations) into three clusters: (6)

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

Initial cluster centers are: A1(2, 10), A4(5, 8) and A7(1, 2).

The distance function between two points a = (x1, y1) and b = (x2, y2) is defined as-

$$P(a, b) = |x_2 - x_1| + |y_2 - y_1|$$

Use K-Means Algorithm to find the three cluster centers after the second iteration.

- (b) Explain the procedure for the computation of the principal components of the data. (6)

OR

(P.T.O.)

BTS-V(R)-11.21-0970

- V. Given the dataset  $\{a, b, c, d, e\}$  and the following distance matrix, construct a dendrogram by complete linkage hierarchical clustering using the agglomerative (12)

	a	b	c	d	e
a	0	1	2	2	3
b	1	0	2	4	3
c	2	2	0	1	5
d	2	4	1	0	3
e	3	3	5	3	0

- VI. State the mathematical formulation of the SVM problem. Give an outline of the method for solving the problem. (12)

OR

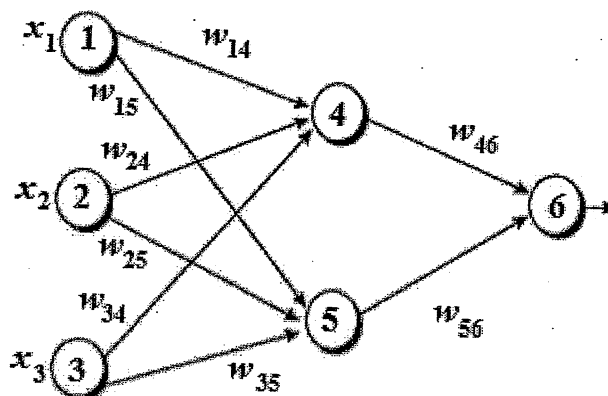
- VII. (a) Identify the first splitting attribute for decision tree by using ID3 algorithm with the following dataset. (8)

Major	Experience	Tie	Hired?
CS	Programming	pretty	NO
CS	Programming	Pretty	NO
CS	Management	Pretty	YES
CS	Management	Ugly	YES
Business	Programming	pretty	YES
Business	Programming	ugly	YES
Business	Management	pretty	NO
Business	Management	pretty	NO

- (b) What are the benefits of pruning in decision tree induction? Explain different approaches to tree pruning. (4)
- VIII. (a) Differentiate model based and temporal based learning. (8)
- (b) Discuss the importance of Radial Basis Functions in learning. (4)

OR

- IX. Implement back propagation algorithm in the below network and update the weights. (12)



$$x_1 = x_3 = 1, x_2 = 0, w_{14} = 0.2, w_{15} = -0.3, w_{25} = 0.1, w_{24} = 0.4, \\ w_{34} = -0.5, w_{35} = 0.2, w_{46} = -0.3, w_{56} = -0.2, h_4 = -0.4, h_5 = 0.2, \\ o_6 = 0.1$$