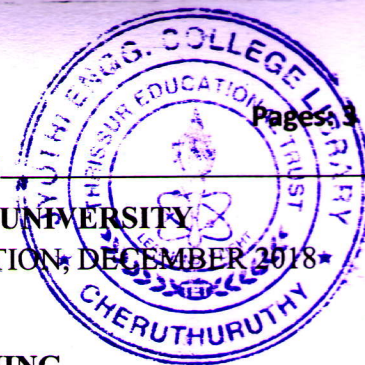


F

T7944

Reg No.: \_\_\_\_\_

Name: \_\_\_\_\_



SEVENTH SEMESTER B.TECH DEGREE EXAMINATION, DECEMBER 2018

Course Code: CS467

Course Name: MACHINE LEARNING

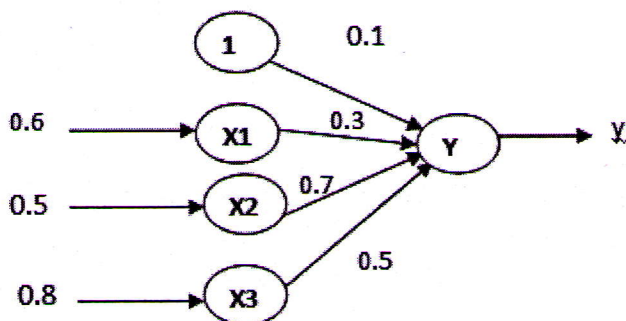
Max. Marks: 100

Duration: 3 Hours

**PART A***Answer all questions, each carries 4 marks.*

Marks

- 1 Distinguish between classification and regression with an example. (4)
  - 2 Define Hypothesis space and version space for a binary classification problem. (4)  
Determine the hypothesis space H and version space with respect to the following data D.
- |       |   |    |    |   |   |   |   |    |    |
|-------|---|----|----|---|---|---|---|----|----|
| x     | 2 | 11 | 17 | 0 | 1 | 5 | 7 | 13 | 20 |
| Class | 0 | 1  | 1  | 0 | 0 | 0 | 0 | 1  | 1  |
- 3 State Occam's razor principle. Illustrate its necessity in learning hypothesis. (4)
  - 4 Define the following terms (a) sensitivity (b) Specificity (c) Precision (d) Accuracy for a classification problem. (4)
  - 5 What is meant by k-fold cross validation. Given a data set with 1200 instances, how k-fold cross validation is done with k=1200. (4)
  - 6 Calculate the output of the following neuron Y if the activation function as (4)
    - (a) Binary sigmoid
    - (b) Bipolar sigmoid



- 7 Distinguish between bagging and boosting. (4)
- 8 Show that the function  $K(x, y) = (x \cdot y)^3$  is a kernel function where  $x \cdot y$  represents dot product of  $x$  and  $y$ . (4)
- 9 Briefly describe the concept of Expectation Maximization algorithm. (4)
- 10 Consider two data points in two dimensional A(5,8) and B(8,9). Calculate (4)
  - (a) City block distance (b) Chessboard distance between A & B.

**PART B***Answer any two full questions, each carries 9 marks.*

- 11 a) An open interval in  $\mathbb{R}$  is defined as  $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$ . It has two parameters  $a$  and  $b$ . Show that the set of all open intervals has a VC dimension of 2 (4)
- b) Compare Unsupervised Learning and Reinforcement Learning with examples. (5)

- 12 a) Discuss the necessity of dimensionality reduction in machine learning. (3)  
 b) Illustrate the idea of PCA for a two dimensional data using suitable diagrams. (6)
- 13 a) Let  $X = R^2$  and  $C$  be the set of all possible rectangles in two dimensional plane which are axis aligned (not rotated). Show that this concept class is PAC learnable. (6)  
 b) Describe the applications of machine learning in any three different domains. (3)

## PART C

*Answer any two full questions, each carries 9 marks.*

- 14 The following table consists of training data from an employee database. For a given row entry, count represents the number of data tuples having the values for department, status, age, and salary given in that row. Let status be the class label attribute. Given a data tuple having the values "systems", "31..35", and "46-50K" for the attributes department, age, and salary, respectively, what would a Naive Bayesian classification of the status for the tuple be? (9)

Department	Status	Age	Salary	Count
sales	senior	31..35	46K..50K	30
sales	junior	26..30	26K..30K	40
sales	junior	31..35	31K..35K	40
systems	junior	21..25	46K..50K	20
systems	senior	31..35	66K..70K	5
systems	junior	26..30	46K..50K	3
systems	senior	41..45	66K..70K	3
marketing	senior	36..40	46K..50K	10
marketing	junior	31..35	41K..45K	4
secretary	senior	46..50	36K..40K	4
marketing	junior	26..30	26K..30K	6

- 15 With the following data set, generate a decision tree and predict the class label for a data point with values <Female, 2, standard, high>. (9)

Gender	Car Ownership	Travel cost	Income level	Transport mode
Male	0	Cheap	Low	Bus
Male	1	Cheap	Medium	Bus
Female	0	Cheap	Low	Bus
Male	1	Cheap	Medium	Bus
Female	1	Expensive	High	Car
Male	2	Expensive	Medium	Car
Female	2	Expensive	High	Car
Female	1	Cheap	Medium	Train
Male	0	Standard	Medium	Train
Female	1	Standard	Medium	Train

- 16 a) Point out the benefits of pruning in decision tree induction. Explain different (5)

approaches to tree pruning.?

- b) Compute ML estimate for the parameter  $p$  in the binomial distribution whose probability function is (4)

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)} \quad x = 0, 1, 2, \dots, n.$$

### PART D

*Answer any two full questions, each carries 12 marks.*

- 17 a) Explain the basic problems associated with hidden markov model. (6)  
 b) Describe the significance of soft margin hyperplane and optimal separating hyperplane and explain how they are computed. (6)
- 18 a) Suppose that the datamining task is to cluster the following seven points (with  $(x,y)$  representing location) into two clusters  $A_1(1,1)$ ,  $A_2(1.5,2)$ ,  $A_3(3,4)$ ,  $A_4(5,7)$ ,  $A_5(3.5,5)$ ,  $A_6(4.5,5)$ ,  $A_7(3.5,4.5)$  The distance function is City block distance. Suppose initially we assign  $A_1, A_5$  as the centre for each cluster respectively. Using the K-means algorithm to find the three clusters and their centres after two round of execution. (6)  
 b) Give the significance of kernel trick in the context of support vector machine. Describe different types standard kernel functions. (6)
- 19 a) Describe any one technique for Density based clustering with necessary diagrams. (6)  
 b) Given the following distance matrix, construct the dendogram using single linkage, complete linkage and average linkage clustering algorithm. (6)

Item	A	B	C	D	E
A	0	2	3	3	4
B	2	0	3	5	4
C	3	3	0	2	6
D	3	5	2	0	4
E	4	4	6	4	0

\*\*\*\*