[This question paper contains 6 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 9425 **HC**

Unique Paper Code : 32347607

Name of the Paper : Machine Learning

Name of the Course : **B.Sc. (H) Computer Science : DSE-3**

Semester : VI

Duration : 3 Hours Maximum Marks : 75

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Section **A** is compulsory. Attempt any **4** questions from Section **B.**
3. Use of scientific calculator is allowed.

Section – A (Compulsory)

1. (a) What is the difference between Supervised and Unsupervised machine learning techniques? (3)

   (b) Draw a diagram for a multi-layer perceptron? (3)

(c) What is the difference between Linear Regression and Logistic Regression? (3)

(d) Explain the Cost function for Linear regression. (3)

(e) Normalize the given data using mean normalization method.

12, 34, 45, 15, 40, 34 (3)

(f) How does the posterior probability of a class is computed by Naïve Byes classifier? (3)

(g) How can neural network be used to tackle three class problems? (3)

(h) Show that single layer perceptron can't solve XOR problem. (3)

(i) Give an expression of bipolar sigmoidal activation function. Also obtain first derivative of the function. (3)

(j) For a classification problem to classify 220 training instances into two classes **TRUE** and **FALSE**, the prediction pattern of a classifier is shown below :

1) 110 **TRUE** class instances classified as **TRUE**

2) 75 **TRUE** class instances classified as **FALSE**

3) 25 **FALSE** class instances classified as **TRUE**

4) 10 **FALSE** class instances classified as **FALSE**

Find the accuracy of this classifier. (4)

(k) List and explain applications of machine learning. (4)

Section – B

2. (a) Consider the following 10 training instances (4)

No.	Color	Type	Mileage	Tested
1	Blue	Sports	Average	Yes
2	Blue	SUV	Average	No
3	Blue	Sports	High	Yes
4	Pink	Sports	High	No
5	Pink	Sports	Average	Yes
6	Pink	SUV	Average	No
7	Blue	SUV	High	Yes
8	Red	SUV	High	No
9	Red	SUV	Average	Yes
10	Red	Sports	Average	Yes

Compute the following probabilities

1. P(Blue/Yes)

2. P(Sports/No)

3. P(Red/Yes)

4. P(Pink/Yes)

(b) Explain Gradient Descent algorithm for predicting parameters of multivariate Linear regression. (6)

3. (a) Calculate intercept and slope the following pairs of (x, y) training values : (2,6),(5,12),(8,15), (12, 23),(17,29). Predict the value of y for test data x=10. (5)

(b) Write Best Subset Selection algorithm. Comment on the complexity of this algorithm. (5)

4. (a) Why Linear regression cannot be applied for categorical data? Explain with a suitable example. (5)

(b) For the values of β_0 and β_1 as −2.16 and 0.425 for categorical predictor variable X and a categorical response variable Y respectively, apply logistic regression to find $\widehat{Pr}(Y = yes \mid x = yes)$ and $\widehat{Pr}(Y = yes \mid x = no)$. (5)

5. (a) What do you understand by over-fitting of a classifier? How regularization can be used to tackle the problem of over-fitting? (6)

(b) Write the expression of the cost function for logistic regression and explain it. (4)

6. (a) Find the linear regression coefficients using gradient descent method for the following dataset when learning rate = 0.2. Carry out the process for 2 iterations.

X	0	1	2	3	4
Y	3	4	5	4	6

(5)

(b) Explain polynomial regression using formal notation. (3)

(c) How does an Artificial Neural Network resemble functioning of brain? (2)

7. (a) Explain Backpropagation algorithm for multilayer perceptron. (6)

(b) Train a neural network for the following data with X1 and X2 as inputs and Y as output.

X1	X2	Y (output)
0	0	1
0	1	0
1	0	0
1	1	0

(4)