

**BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI  
(END SEMESTER EXAMINATION)**

**CLASS: M.TECH  
BRANCH: CS/IT**

**SEMESTER : I  
SESSION : MO/19**

**SUBJECT: CS506 MACHINE LEARNING**

**TIME: 3 HOURS**

**FULL MARKS: 50**

**INSTRUCTIONS:**

1. The question paper contains 5 questions each of 10 marks and total 50 marks.
  2. Attempt all questions.
  3. The missing data, if any, may be assumed suitably.
  4. Before attempting the question paper, be sure that you have got the correct question paper.
  5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.
- 

- Q.1(a) (i) Describe the different types of machine learning. [3+2]  
(ii) Explain any one method to test the accuracy of a model.
- Q.1(b) Some randomly selected chocolate brands were tested for their fat contents (x) and calories (y). [2+1+2]  
The findings are documented in the table below:

Chocolate Brands	Fat contents	Calories (in 12 oz)
Diary Milk	4.70%	163
Snickers	6.70%	215
Munch	8.10%	222
Perk	4.15%	104
5 Star	5.10%	162
Gems	5.00%	158
Toblerone	5.00%	155
Amul	4.70%	158
Safari	6.20%	195

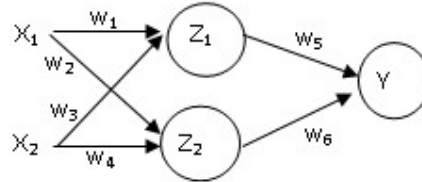
- (i) What linear regression equation best predicts the calories contained based on its fat contents?
- (ii) If a Chocolate brand has a fat content of 5.70% predict the calorie content?
- (iii) How well does the regression equation fit the data?
- Q.2(a) Illustrate how does overfitting occur in machine learning and how can it be avoided. [4]
- Q.2(b) Use the following dataset to construct a decision tree to predict whether a person is Happy (H) or Sad (S) based on the color of their shirt, whether they wear glasses and the number of SIM cards they have. [6]

Shirt Color	Wear Glasses	Number of SIM cards	Output
G	Y	2	S
G	N	2	S
G	N	2	S
B	N	2	S
B	N	2	H
R	N	2	H
R	N	2	H
R	N	2	H
R	Y	3	H

- Q.3(a) (i) Design a perceptron to implement AND function with bipolar inputs and targets. [2+2]  
(ii) Explain the working of a Recurrent Network.

**PTO**

- Q.3(b) Consider a two-layer feed-forward neural network that has the topology shown in the figure 1.  $X_1$  and  $X_2$  are the two inputs,  $Z_1$  and  $Z_2$  are the two hidden neurons,  $Y$  is the (single) output neuron,  $w_i, i=1..4$ , are the weights of the connections from the inputs to the hidden neurons and  $w_j, j=5..6$ , are the weights of the connections from the hidden neurons to the output neuron. Explain the first training iteration of the Backpropagation algorithm for the current network. [6]



- Q.4(a) Give an advantage of hierarchical clustering over K means clustering. [2]
- Q.4(b) (i) Given points  $A = (1,2)$ ,  $B = (2,2)$ ,  $C = (2, 1)$ ,  $D = (-1, 4)$ ,  $E = (-2, -1)$ ,  $F = (-1,-1)$ . Starting from initial clusters  $\text{Cluster1} = \{A\}$  which contains only the point A and  $\text{Cluster2} = \{D\}$  which contains only the point D, use the K-means clustering algorithm and report the final clusters. Use L1 distance as the distance between points which is given by  $d((x1, y1), (x2, y2)) = |x1 - x2| + |y1 - y2|$ . [4+4]
- (ii) Given 1-dimensional points  $A = 1$ ,  $B = 2$ ,  $C = 3$ ,  $D = 8$ ,  $E = 9$ ,  $F = 10$ . Compute single-link bottom-up hierarchical clustering using  $d(x, y) = |x - y|$  as the distance between points.
- Q.5(a) Illustrate how does an active learning model learn. [5]
- Q.5(b) Explain the AdaBoost Algorithm. [5]

::::04/12/2019::::E