

- (ii) Assume a total of 1,000 patients are tested for influenza; 900 are found to be healthy, while 100 are found to be sick. A test resulted in 60 being positive and 40 being negative for the sick persons. The same test was positive for 120 and negative for 780 in healthy adults. Construct a confusion matrix for the data and determine precision and recall. (6)
8. (i) What is regularization? What is the effect of the following on the model?
- (a) The regularization parameter (λ) is zero
- (b) The regularization parameter (λ) is very large (4)
- (ii) State the mathematical formulation of the SVM problem. Give an outline of the method used for solving the classification problem using SVM. (6)

[This question paper contains 6 printed pages.]

Your Roll No.....

Sr. No. of Question Paper : 1262

A

Unique Paper Code : 32347607

Name of the Paper : Machine Learning

Name of the Course : **B.Sc. (Hons.) Computer Science (LOCF)**

(Admission of 2019)

Semester : VI

Duration : 3 Hours

Maximum Marks : 75

Instructions for Candidates

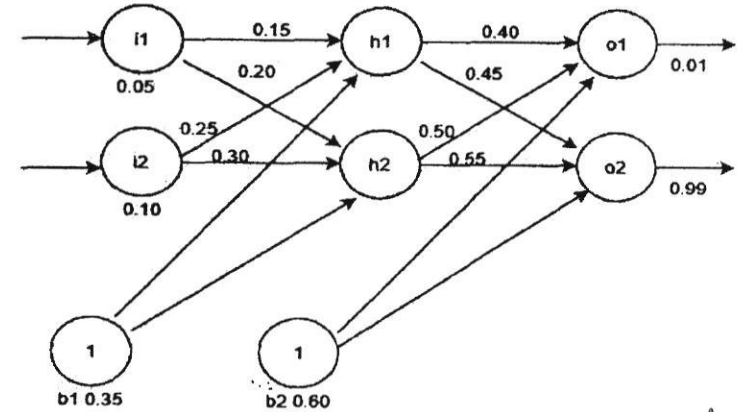
1. Write your Roll No. on the top immediately on receipt of this question paper.
2. **Section A** is compulsory.
3. Attempt any **4 (four)** questions from **Section B**.
4. Use of scientific calculator is allowed.

SECTION A

1. (i) Distinguish between supervised learning and unsupervised learning. Illustrate with an example. (5)
- (ii) Define Concept Learning. How the concept learning can be viewed as the task of searching? (5)

- (iii) Calculate the output y of a single layer neural network with three input neuron and one output neuron. The input feature vector is $(x_1, x_2, x_3) = (0.8, 0.6, 0.4)$ and weight values are $[w_1, w_2, w_3] = [0.2, 0.1, -0.3]$ and bias = 0.35. Use binary Sigmoid function as activation function. (5)
- (iv) Distinguish between overfitting and underfitting. How it can affect model generalization? (5)
- (v) Using an example discuss how new features can be constructed by forming Cartesian product of existing features. What are the implications of this approach? (5)
- (vi) Suppose that the probability of five events are $P(\text{first}) = 0.5$ and $P(\text{second}) = P(\text{third}) = P(\text{fourth}) = P(\text{Fifth}) = 0.125$. Calculate its entropy. (5)
- (vii) Use K-means clustering to cluster the following data into two groups :
- {2, 4, 10, 12, 3, 20, 30, 11, 25}
- Assume cluster centroid are $m_1=2$ and $m_2=4$. The distance function used is Euclidean distance. (5)

5. Consider the following neural network with initial weights, biases, and training input / outputs as mentioned.



Given the inputs $i_1 = 0.05$, $i_2 = 0.10$, determine the values of output nodes o_1 and o_2 . Also calculate the prediction error E_{total} if the actual output values o_1 and o_2 are 0.01 and 0.99 respectively. Use Sigmoid as the activation function for the hidden as well as the output layers. (10)

6. (i) What is the difference between K-means clustering and K-Nearest Neighbor classifier? (4)
- (ii) Discuss steps used by Principal Component Analysis to extract important features. (6)
7. (i) Given the set of values $X = (3, 9, 11, 5, 2)^T$ and $Y = (1, 8, 11, 4, 3)^T$. Evaluate the regression coefficients using ordinary least square method. (4)

SECTION B

2. (i) Find the least square regression line for the given dataset using the normal equation method. Show computation at each step.

x1	x2	y
1	9	14
2	1	7
3	2	12
4	3	16
5	4	20

(4)

- (ii) Consider the dataset given below having two input variables x_1 , x_2 and one output variable y . Update the coefficients θ_0 , θ_1 and θ_2 using gradient descent for the logistic regression model. Assume the learning rate = 0.3 and the initial values of coefficients as $\theta_0 = -0.5$, $\theta_1 = 1$, and $\theta_2 = -1$. Perform one iteration of gradient descent.

(6)

x1	1	3	8	6
x2	2	5	3	2
y	0	0	1	1

3. (i) State Bayes Theorem. (2)
- (ii) Consider the training data in the following table where Play is a class attribute.

Humidity	Sunny	Wind	Play
L	N	S	N
H	N	W	Y
H	N	W	N
H	Y	S	Y
H	N	W	Y
L	Y	S	N

Build a Naive Bayes Classifier using the above data. Estimate the class label for day (Humidity=L, Sunny=N, Wind=W) using the above classifier. (8)

4. (i) Differentiate between Standard and Stochastic Gradient Descent. (4)
- (ii) Identify the first splitting attribute for decision tree with the following dataset using ID3 algorithm :

Field	Experience	Hired
IT	Coding	No
IT	Coding	No
IT	Administration	Yes
IT	Administration	Yes
Business	Coding	Yes
Business	Coding	Yes
Business	Administration	No
Business	Administration	No

(6)