

M.E. COMPUTER SCIENCE AND ENGINEERING FIRST YEAR FIRST SEMESTER – 2025

MACHINE LEARNING

Time: 3 hours

Full Marks: 100

Answer any **FIVE** questions

1.

- a. What is information gain? Explain when it can be zero.
- b. What are the merits and demerits of the KNN algorithm? Is it a parametric or non-parametric learning algorithm?
- c. What are the differences between the Logistic Regression Model and the threshold classifier
- d. What are called support vectors? Why is kernel trick used with SVM?
- e. Explain drawbacks of the K-means clustering algorithm.

4 x 5 = 20 marks

2. a) Given the following short texts, each labelled with a genre, either **Class 1** or **Class 2**:

Text words	Class
AA, BB, CC, CC	Class 1
DD, EE, FF	Class 2
BB, GG, DD, AA, AA	Class 1
EE, FF, FF, AA	Class 2
GG, DD, FF, CC	Class 2

Using the above training dataset, compute the most likely class for a new text containing the words: (DD, BB, FF, GG). Assume a Naive Bayes classifier and use add-1 smoothing for the likelihood.

- b) What is called conditional independence? What is its significance in Bayesian Learning?

15 + 5 = 20 marks

[Turn over

3. a) Consider the following set of training examples. Use ID3 algorithm to learn a decision tree from the data (Show all necessary computations).

ID	TCVAD	GOUGA	EXCERPTING	INFECTED
1	NO	NO	NO	NO
2	YES	YES	YES	YES
3	YES	YES	NO	NO
4	YES	NO	YES	YES
5	YES	YES	YES	YES
6	NO	YES	NO	NO
7	YES	NO	YES	YES
8	YES	NO	YES	YES
9	NO	YES	YES	YES
10	YES	YES	NO	YES
11	NO	YES	NO	NO
12	NO	YES	YES	YES
13	NO	YES	YES	NO
14	YES	YES	NO	NO

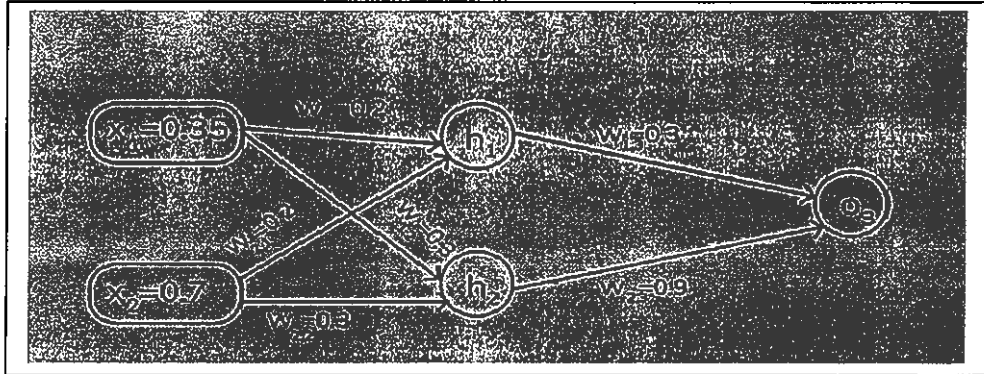
b) With a suitable example of classification output, show how the following evaluation measures are computed?

- (i) Precision (ii) Recall (iii) F-measure

15 + 5 = 20 marks

4. a) Why is deep learning taking off? Give at least three reasons.

b) The following is an Artificial Neural Networks, with sigmoid units in the hidden layer and the output layers.



Consider that the following training example is submitted to the net (as shown in the above figure).

X1	X2	Target output
0.35	0.7	0.7

Now show the forward pass to compute responses at h_1 , h_2 and O_3 . Then use the backpropagation algorithm to compute $\nabla W_{1,3}$ and $\nabla W_{2,1}$ (Learning rate = 1). 5 + 15 = 20 marks

5. a) For linear SVM, to find the maximum margin hyperplane ($w^T x + w_0 = 0$), we need to find w and w_0 by solving the a quadratic optimization problem. Show that total margin for a linear SVM is $\frac{2}{\|w\|}$.

b) Consider a dataset that contains two variables: height (cm) & weight (kg). Each point is classified as normal or underweight.

Weight(x2)	Height(y2)	Class
51	167	Underweight
62	182	Normal
69	175	Normal
64	173	Normal
65	170	Normal
56	174	Underweight
58	169	Normal
57	173	Normal
55	170	Normal

Using the above training data, classify the following new data point using the KNN algorithm when $K=1$ and $K=3$.

Weight(x2)	Height(y2)	Class
57	175	?

What is your observation on the results of KNN with two different values of K ?

5 + 15 = 20 marks

Signature of the question setter

Moderated by

6. a) Derive the weight update rule for the linear regression model. Consider MSE cost function.
b) What causes over-fitting in decision tree learning? Discuss one method for handling over-fitting in decision tree learning.
c) Differentiate (with examples) between supervised learning and reinforcement learning.

10 + 5 + 5 = 20 marks

7. a) Perform the K-means clustering algorithm on the following data to find 2 clusters. Show the clusters obtained after two iterations. Calculate MSE error after the end of second iteration.

Instances	X1	X2
1	2	4
2	2	1
3	8	4
4	6	6
5	8	6

- b) Discuss the single linkage hierarchical clustering with a simple example.

15+5 = 20 marks
