Ref.No :PG/CSE/PC/T/112A/2025

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

MACHINE LEARNING

Time: 3 hours Full Marks: 100

Answer Question No. **1 (Compulsory)** and any **Four** questions from the rest

1. a) How is a hypothesis space defined in the case of decision tree learning? If S is a collection of *m* training examples for a binary classification problem, what does it signify if entropy of S is 0.5?-Explain.
   b) Is KNN a non-linear classifier or a linear classifier? Give reasons.
   c)What are the drawbacks of the threshold classifier? Illustrate with an example.
   d)Compare a SVM classifier with an Artificial Neural Network with one hidden layer in terms of their merits and demerits.
   e)A learning algorithm learns a function that maps inputs to outputs. Is this statement true for unsupervised learning? Give reason in favour of your answer. 4 x 5 = **20 marks**

2. a) Consider the followingsentiment analysis data for training and testing Naive Bayes with add-one smoothing. This is a sentiment analysis domain with two classes positive (+) and negative (-). Find the prediction for the test example given in the following table using the trained Naïve Bayes Model ("Cat" means category).

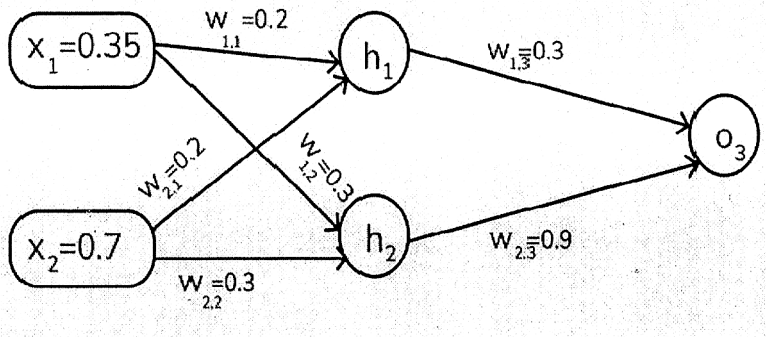
	Cat	Documents
Training	-	just plain boring
	-	entirely predictable and lacks energy
	-	no surprises and very few laughs
	+	very powerful
	+	the most fun film of the summer
Test	?	predictable with no fun

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

Match Number	Pitch	Format	Winner (Target)
1	S	T	Green
2	S	T	Blue
3	F	O	Blue
4	S	O	Blue
5	F	T	Green
6	F	O	Blue
7	S	O	Green
8	F	T	Blue
9	F	O	Blue
10	S	O	Green

10+10=20 marks

3. a) Explain why non-linearity should be introduced in an Artificial Neural Networks.
   b) The following is an Artificial Neural Networks, with sigmoid units in the hidden layer and the output layer.



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

X_1	X_2	Target output
0.35	0.7	**0.5**

Now show the forward pass to compute responses at h1, h2 and O3. Then use the backpropagation algorithm to compute $\nabla W_{2,3}$ and $\nabla W_{1,1}$ (Learning rate =1). 5 + 15 = **20 marks**

4. a) For a linear SVM, the equation of the separating hyperplane is: $W^T X + W_0 = 0$. Show that W can be written as the weighted sum of support vectors.
   b) With a suitable numerical example, show that classification performance of a KNN classifier depends on the choice of K, where K is the number of selected nearest neighbours.

   10 + 10 = **20 marks**

5. a) What is the difference between Bagging and Boosting? Why does boosting improve prediction accuracy?- Explain. Discuss the *Random Forest* Algorithm.
   b) Consider the following dataset and perform single linkage agglomerative clusteringalgorithm on data points. Represent the clustering results in the form of *Dendrogram*. Show each step of the algorithm and the necessary calculations.

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

(2+3+5) + 10 = **20 Marks**

6. a) Discuss the drawbacks of the K-means clustering algorithm.
   b) Differentiate (with examples) between supervised learning and reinforcement learning.
   c) Derive the Delta rule used to update weights of a linear artificial neuron.
   d) What is kernel trick? Explain why it is used with SVM.

5x 4 =**20 marks**

7. a) Calculate the total number of parameters for a Naïve Bayes model if n Boolean attributes are considered.
   b) Write short notes on: (i) Logistic regression model (ii) Reduced error pruning in decision tree learning
   c) Consider the values of the following metrics for the classification output of a binary classifier and calculate i) Macro F1-score (ii) Weighted F1-score

TP = 650, TN = 150, FP = 120, FN = 80.

5 + (7+3) + 5 = **20 Marks**