Time: 03 Hours **Marks: 80**

Note: 1. Question 1 is compulsory
2. Answer any three out of remaining five questions.
3. Assume any suitable data wherever required and justify the same.

Q1 a) Explain the Expectation Maximization Algorithm (EMA) [5]

b) Explain kernel functions and kernel trick [5]

c) What are the issues in decision tree learning? [5]

d) "Entropy is a thermodynamic function used to measure the disorder of a system in Chemistry." Clarify the concept of entropy in Machine Learning? [5]

Q2 a) Compare and contrast Linear and Logistic regressions with respect to their mechanisms of prediction. [10]

b) Consider 2-D dataset given in the table below. Construct a SVM classifier model. Given (2, 1), (2, -1) and (4, 0) as support vectors; estimate the parameters of the model and classify (4, 2). Why is SVM called as optimal binary hyper plane classifier? [10]

(X1, X2)	(1, -1)	(2, -1)	(5, -1)	(4, 0)	(6, 0)	(1, 1)	(2, 1)	(5, 1)
Class	C1	C1	C2	C2	C2	C1	C1	C2

Q3 a) You are given a data set on cancer detection. You have built a classification model and achieved an accuracy of 96%. Why shouldn't you be happy with your model performance? What can you do about it? [10]

b) What is a HMM? What are the issues in Hidden Markov Model (HMM)? [10]

Q4 a) You came to know that your model is suffering from low bias and high variance. Which algorithm should you use to tackle it? Why? [10]

b) Differentiate between simple linkage, average linkage and complete linkage algorithms. Use complete linkage algorithm to find the clusters from the following dataset. [10]

X	4	8	15	24	24
Y	4	4	8	4	12

Q5 a) Draw the block diagram of Error Back Propagation Algorithm and explain with flow chart the concept of Back Propagation. [10]

b) The following table consists of training data from an employee database. The data have been generalized. For example, "31 . . . 35" for age represents the age range of 31 to 35. For a given row entry, count represents the number of data tuples having the values for department, status, age, and salary given in that row. Let the status be the class-label attribute. [10]

(i) Design a multilayer feed-forward neural network for the given data. Label the nodes in the input and output layers.

(ii) Using the multilayer feed-forward neural network obtained in (i), show the weight values after one iteration of the back propagation algorithm, given the training instance "(sales, senior, 31 . . . 35, 46K . . . 50K)".

Assume initial weight values and biases. Assume learning rate to be 0.9. Use binary input and draw (*one input layer, one output layer and one hidden layer*) neural network. Solve the problem for one epoch.

department	*status*	*age*	*salary*	*count*
sales	senior	31 … 35	46K … 50K	30
sales	junior	26 … 30	26K … 30K	40
sales	junior	31 … 35	31K … 35K	40
systems	junior	21 … 25	46K … 50K	20
systems	senior	31 … 35	66K … 70K	5
systems	junior	26 … 30	46K … 50K	3
systems	senior	41 … 45	66K … 70K	3
marketing	senior	36 … 40	46K … 50K	10
marketing	junior	31 … 35	41K … 45K	4
secretary	senior	46 … 50	36K … 40K	4
secretary	junior	26 … 30	26K … 30K	6

Q6 Write short notes on any two of the following: [20]

a) Temporal Difference Learning in Reinforcement Learning

b) Over fitting' in Machine learning

c) Independent Component Analysis

***