

B.E. (Computer) Sem. VIII C8865

4/12/19

1/2

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. [10]
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?

(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 [10]
and achieved an accuracy of 96%. Why shouldn't you be happy with your model performance? What can you do about it?
- 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. [10]
Which algorithm should you use to tackle it? Why?
- b) Differentiate between simple linkage, average linkage and complete linkage [10]
algorithms. Use complete linkage algorithm to find the clusters from the following dataset.

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

B.E (Computer) Sem-VIII CBSEs

4/12/19

2/2

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.

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

- Temporal Difference Learning in Reinforcement Learning
- Over fitting in Machine learning
- Independent Component Analysis

Op lock 69460

B.E SEM-VIII CBSEGS Computer

(3 hours)

13/14/19

(1/1)

[Total Marks: 80]

NB :

- 1) Question No.1 is **compulsory**.
- 2) Attempt any **three** questions out of the remaining questions.
- 3) Make suitable assumptions wherever necessary.

- Q.1 Solve Any four (5 Marks X 4)
- a) Compare parallel and distributed system models by giving example of each. 5
 - b) State the goals of a distributed system. 5
 - c) Compare and contrast between message oriented and stream oriented communication. 5
 - d) Discuss Amdahl's law for measuring speed up performance of parallel system. 5
 - e) Enlist and discuss desirable features of global scheduling algorithm 5
- Q.2 a) Illustrate 4 stage pipeline architecture. 10
- b) What is Remote Procedure Call. Discuss the working of RPC in detail. 10
- Q.3 a) Discuss the role consistency in distributed system. What is the need of client centric consistency models. Explain any two data centric consistency models. 10
- b) Illustrate the implementation details of pipelined floating point adder. 10
- Q.4 a) Discuss the need of process migration. Explain the role of resource to process and process to resource binding in process migration. 10
- b) Explain Raymond's Tree based algorithm of token based distributed mutual exclusion. 10
- Q.5 a) Describe code migration issues in detail. 10
- b) Explain the load balancing approach. Explain static and dynamic load balancing algorithm. 10
- Q.6 Attempt any two (10X2) 20
- a) Pipeline hazards and techniques to eliminate those hazards
 - b) Lamport algorithm
 - c) Election Algorithm
 - d) Andrew File System(AFS)

Op Code 79170