

May 2018

Q. P. Code: 25961

1/2

B.E(Comp)-sem vii -CBSE -16/5/18

DWM

Time: 03 Hours

Marks: 80

- Note: 1. Question 1 is compulsory
2. Answer any three out of remaining questions.

Q1 A) i. Design star & snowflake schema for "Hotel Occupancy" considering [10] dimensions like Time, Hotel, Room, etc.

ii. Calculate the maximum number of base fact table records for the values given below:

Time period: 5 years

Hotels: 150

Rooms: 750 rooms in each Hotel (about 400 occupied in each hotel daily).

B) Explain Data mining as a step in KDD. Give the architecture of typical data mining [10] System.

Q2 A) The college wants to record the marks for the courses completed by students using [10] the dimensions: a) Course, b) Student, c) Time & a measure d) Aggregate marks.

Create a Cube and describe following OLAP operations:

i) Rollup ii) Drill down iii) Slice iv) Dice v) Pivot.

B) A simple example from the stock market involving only discrete ranges has profit [10] as categorical attribute, with values {up, down} and the training data is:

Age	Competition	Type	Profit
Old	Yes	Software	Down
Old	No	Software	Down
Old	No	Hardware	Down
Mid	Yes	Software	Down
Mid	Yes	Hardware	Down
Mid	No	Hardware	Up
Mid	No	Software	Up
New	Yes	Software	Up
New	No	Hardware	Up
New	No	Software	Up

Apply decision tree algorithm and show the generated rules.

Q3 A) Why naive Bayesian classification is called "naive"? Briefly outline the major ideas [10] of Naive Bayesian classification.

B) Discuss different steps involved in Data Pre-processing [10]

Q4 A) Explain ETL of data warehousing in detail. [10]

B) Find clusters using k -means clustering algorithm if we have several objects [10] (4 types of medicines) and each object have two attributes or features as shown in the table below. The goal is to group these objects into $k=2$ group of medicine

B. E (Comp) - sem VIII - CBSCS - 16/5/18

Q. P. Code: 25961

based on the two features (pH and weight index).

DWM

2f2

Object	Attribute 1(X) Weight Index	Attribute 2 (Y) pH
Medicine A	1	1
Medicine B	2	1
Medicine C	4	3
Medicine D	5	4

Q5 A) Explain Data Warehouse Architecture in detail. [10]

B) A database has five transactions. Let minimum support = 30% and minimum Confidence = 70%

- Find all frequent patterns using Apriori Algorithm.
- List strong association rules.

Transaction_Id	Items
A	1,3,4,6
B	2,3,5,7
C	1,2,3,5,8
D	2,5,9,10
E	1,4

Q6 Write short note on the following (Answer any FOUR) [20]

- Data warehouse design strategies
- Applications of Data Mining
- Role of metadata
- Multidimensional and multilevel association mining
- Hierarchical clustering

B.E (comp) sem vij - CBCS - 22/5/18 May 2018
Q.P. Code: 27190
Human Machine Interaction (Total Marks 80)

• Time: 3 Hours

Question no 1 is compulsory
solve any 3 from Q2 to Q6
Indicate your answers with neat sketch wherever necessary

Q1. ATTEMPT ANY FOUR

20

- a) What do you mean by response time? Explain.
- b) Explain three levels of processing in detail.
- c) State and explain different interview techniques.
- d) Explain steps in constructing persona.
- e) How one can create a dialogue with interface user? Explain with appropriate example.

Q2 a) What are various type of windows? Explain.

10

b) What are three levels of users? Explain how to accommodate them in user interface.

10

Q3 a) What are various factors to be considered for User Interface Design? Also give

10

suitable example for the same.

b) Differentiate between Qualitative and Quantitative Research

10

Q4 a) What are various methodologies adopted for Feedback and Guidance? Explain in detail.

10

b) Explain Various Menus in Human Machine Interface?

10

Q5 a) What do you mean by Keyboard Accelerators? Explain

10

b) Explain Gcal Directed Design Process in Detail.

10

Q6 Write Short Note on following

20

- a) Device Based Control
- b) Screen Based Control
- c) Statistical Graphics
- d) Graphics Icons and Images

3Hrs

PDS

Note: 1) Question no 1 is compulsory

2) Solve any three from remaining

- Q1 a) Explain Flynn's classification scheme (05)
b) Explain Data mapping and memory in array processor (05)
c) Explain desirable features of global scheduling algorithm (05)
d) Explain Berkeley physical clock algorithm (05)
- Q2 a) Explain different types of Hazards in Parallel System (10)
b) Explain Ricart-Agrawala algorithm for Mutual Exclusion (10)
- Q3 a) Give an example that can be solved effectively with SIMD architecture (10)
b) What are the different Architectural Model of Distributed System?
Explain with suitable diagram (10)
- Q4 a) Explain Hadoop Distributed File System (HDFS). (10)
b) Explain Software models supported by the distributed system (10)
- Q5 a) What is Remote Procedure Call? Explain the working of RPC in detail. (10)
b) What are different data centric consistency model? (10)
- Q6 a) Explain different load estimation policies and process transfer policies used by Load balancing algorithm (10)
b) Explain Bully Election Algorithm (10)

-x-

[Time: 3 Hours]

M L

[Marks:80]

Please check whether you have got the right question paper.

- N.B:
 - Question no 1 is compulsory
 - Attempt any three questions out of remaining five questions
 - Assume any suitable data wherever required but justify the same.

- Q.1 a. Define Machine Learning? Briefly explain the types of learning. 05
 b. What is independent component analysis? 05
 c. What are the issues in decision tree induction? 05
 d. What are the requirements of clustering algorithms? 05

- Q.2 a. The values of independent variable x and dependent value y are given below: 10

X	Y
0	2
1	3
2	5
3	4
4	6

- Find the least square regression line $y = ax + b$. Estimate the value of y when x is 10. 10
- b. What are the steps in designing a machine learning problem? Explain with the checkers problem. 10

- Q.3 a. For a SunBurn dataset given below, construct a decision tree 10

Name	Hair	Height	Weight	Location	Class
Sunita	Blonde	Average	Light	No	Yes
Anita	Blonde	Tall	Average	Yes	No
Kavita	Brown	Short	Average	Yes	No
Sushma	Blonde	Short	Average	No	Yes
Xavier	Red	Average	Heavy	No	Yes
Balaji	Brown	Tall	Heavy	No	No
Ramesh	Brown	Average	Heavy	No	No
Swetha	Blonde	Short	Light	Yes	No

- b. What is the goal of the Support Vector Machine (SVM)? How to compute the margin? 10

- Q.4* a. For the given set of points identify clusters using complete link and average link using agglomerative clustering. 10

	A	B
P1	1	1
P2	1.5	1.5
P3	5	5
P4	3	4
P5	4	4
P6	3	3.5

- b. What is the role of radial basis function in separating nonlinear patterns. 10

B.E (comp) sem-11 - CBCS - 08/05/18
Q.P. Code : 16172

(2)

- Q.5 a. Use Principal Component analysis (PCA) to arrive at the transformed matrix for the given matrix A. 10

$$A^T = \begin{matrix} 2 & 1 & 0 & -1 \\ 4 & 3 & 1 & 0.5 \end{matrix}$$

- b. What are the elements of reinforcement learning? 10

- Q.6* Write short notes on any two 20

- a. Logistic regression
b. Back propagation algorithm
c. Issues in Machine Learning