

Time: 3 hours

Max. Marks: 80

Note: 1. Question no.1 is compulsory.**2. Attempt any three out of remaining five.****3. Assumptions made should be clearly indicated.****4. Figures to the right indicates full marks.****5. Assume suitable data whenever necessary.****Q.1 Compulsory question. (20)**

A Differentiate between ER modelling vs Dimensional Modelling.
Design a Star schema for Library management system.

B Describe issues in data mining.

Q.2 (20)

A Explain need of strategic information. Also explain different features of Data Warehouse.

B An apparel company have sales department consider four dimensions namely Time, product, Store, promotion. The schema contains central act table with two measures dollar_cost and unit_sold.

Describe Slice, dice, Roll up, drill down for the given problem statement.

Q.3 (20)

A Explain different steps involved in data preprocessing.

B Describe k means algorithm. Apply k means algorithm {2,4,10,12,3,20,30,11,25} to form two clusters.

Q.4 (20)

A database has five transactions. Let minimum support be 60% and minimum confidence be 80. Find all frequent itemsets and strong association rules by Apriori algorithm.

T_ID	Items brought
T100	M,O,N,K,E,Y
T200	D,O,N,K,E,Y
T300	M,A,K,E
T400	M,U,C,K,Y
T500	C,O,K,E

B What is web mining? Explain types of web mining.

Q.5

(20)

Consider the following training data set.

Create classification model using decision tree.

A

Sr. no.	Income	Age	On house
1	V High	Young	Yes
2	High	Medium	Yes
3	Low	Young	Rented
4	High	Medium	Yes
5	V high	Medium	Yes
6	Medium	Young	Yes
7	High	Old	Yes
8	Medium	Medium	Rented
9	Low	Medium	Rented
10	Low	Old	Rented
11	High	Young	Yes
12	Medium	Old	Rented

B

Explain architecture of typical data mining system. Also explain data mining task primitives in brief.

Q.6

Write a short note on any four.

(20)

A

Differentiate between spatial vs classical data mining

B

Updates to the dimension tables

C

Market basket analysis with example

D

Applications of Data mining

E

MOLAP vs ROLAP
