Lab BE LIT/ Sem VII / Rev

ws Oct- 10 191

Con. 6110-10.

DWM & BI

(REVISED COURSE)

(3 Hours)

2-12-10

GT-8967

[Total Marks : 100

N.B.: (1) Question no. **1** is compulsory.

(2) Answer any **four** out of the remaining questions.

1.   One of India's large Retail Departmental Chains, with annual revenues touching $2.5 billion mark and having over 3600 employees working at diverse locations, was keenly interested in a business intelligence solution that can bring clear insights on operations and performance of departmental stores across the retail chain. The company needed to support a data warehouse that exceeds daily sales data from Point of Sales (POS) across all locations, with 80 million rows and 71 columns.

(a) List the dimensions and facts for above application.   05

(b) Design Star schema and snow flake schema for the above application   05

(C) Design a BI application which will provide Retail Chain Company with features
    and performance that meet their objectives using any data mining technique.   10

2. (a) What are the Major issues in Data mining.   06

    (b) Explain data Mining Task Primitives.   06

    (c) Explain Fact less fact table with suitable example.   08

3. (a) What is web mining and explain web content mining.   10

    (b) Explain any one method of hierarchical clustering with an example.   10

4. (a) A database has four transactions. Let minimum support and confidence is 50%.

Tid	Items Bought
1	A,B,D
2	A,D
3	A,C
4	B,D,E,F

Find out the frequent item sets and strong association rules for above example.   05

(a) Explain multi dimensional association rules with example.   05

(b) Explain multilevel association rules.   05

(c) Explain constraint based association rule mining.   05

5. (a) Describe Data Discretization, Summarization with an example   07

    (b) Explain Numerosity Reduction in data preprocessing   06

    (c) Explain BIRCH method of clustering with an example.   07

5. (a) Describe Data Discretization, Summarization with an example — 07
   (b) Explain Numerosity Reduction in data preprocessing — 07
   (c) Explain BIRCH method of clustering with an example. — 07

6. (a) Using given table . Create classification model using any algorithm and — 10
   hence classify following tuple.<income =medium, credit =good>

Transaction Id	Income	Credit	Decision
1	Very High	Excellent	Authorize
2	High	Good	Authorize
3	Medium	Excellent	Authorize
4	High	Good	Authorize
5	Very High	Good	Authorize
6	Medium	Excellent	Authorize
7	High	Bad	Request ID
8	Medium	Bad	Request ID
9	High	Bad	Reject
10	Low	Bad	Call Police

   (b) Explain K-means clustering and solve the following with K= 2 — 10
   {2,25,10,15,6,20 4,40}

7. Write short notes on (ANY TWO) :— — 20
   (a) Spatial OLAP
   (b) Association rule mining in Data stream
   (c) Approaches in Text Mining
   (d) Sequence Mining in transactional Data base