Roll No. ……
Printed Pages : 3

10212-R

MCA/M- 13
DATA WAREHOUSING & DATA MINING
Paper- MCA- 402

Time allowed : 3 hours] [*Maximum marks : 80*

Note : Attempt five questions in all, selecting at least one question from each unit.
. Question No. 1 is compulsory.

Compulsory question

1. Write short notes on the following:

(a) Illustrate Drill down and Roll-up operation on data cube.
(b) What do you mean by Data mart? Discuss.
(c) Write short note on Loose Coupling and Tight Coupling.
(d) Differentiate summarization from discrimination.
(e) Write a note on Market Basket Analysis.
(f) Discuss join and Prune for association rule mining.
(g) Explain Nominal & Binary Variables.
(h) Briefly explain DBSCAN algorithm.

UNIT-I

2. (a) What do you mean by data warehouse schema? Draw and explain the various type of schemas of data warehouse.
(b) How partial materialization is important for data warehouse implementation?

3. (a) Define data warehouse. What are the characteristics of a good warehouse design ? Discuss the growth structure of data warehouse.
(b) Differentiate between training data and test data. How fact and dimension tables are identified and generated?

UNIT-II

4. (a) Why we perform data pre-processing? Discuss the various methods associated with data integration and transformation.
(b) List and describe the five primitives for specifying a data mining task.
5. (a) Identify, which part of data mining life cycle is most important for an analyst To identify trend/pattern.
(b) Why attribute relevance analysis is performed? Discuss the methodology for applying attribute relevance analysis attribute relevance analysis with suitable example.

UNIT-III

6. (a) Discuss the issues regarding classification and prediction.
(b) Explain classification by Back propagation.
7. For the input database given below generate association rules with minimum

Support =2 using Apriori and FP tree algorithms:

Tid	Items
1	Pen, Pencil, Cutter
2	Pencil, Sharpener, Cutter
3	Pen, Pencil, Eraser, Bread
4	Pen, Pencil, Sharpener
5	Pen, Eraser, Cutter
6	Pencil, Eraser, Bread
7	Pen, Pencil, Eraser
8	Pen, Pencil, Eraser, Cutter
9	Pen, Pencil, Eraser, Bread

UNIT-IV

8. (a) What data mining tool does? Draw a comparative chart of data mining tool And discuss them in term of accuracy, efficiency and complexity.

(b) How outlier mining differs from clustering?

9. What do you mean by clustering? Discuss the important issues to be addressed By a data clustering system. Explain similarity and distance measures for clustering algorithms.