Roll No……. **10212**

MCA/M-14

DATA WAREHOUSING AND 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. All questions carry equal marks.

(Compulsory Question)

1. (a) Differentiate between OLAP and OLTP.

(b) Write a note on snowflake schema.

(c) Differentiate between No Coupling and Tight Coupling.

(d) What do you mean by interesting pattern?

(e) Discuss Market basket analysis.

(f) What do you mean by information gain?

(g) Briefly explain data and dissimilarity matrix.

(h) What do you mean by outliers? 8x3=24

UNIT -I

2. (a) Differentiate between Training data and Test data. How fact and dimension tables are identified and generated.

(b) How you will differentiate between large dataset and high dimensional Dataset? What are the various OLAP operations available for visualization of data cubes? 7+7=14

3. (a) Discuss in brief the steps for construction and implementation of Data warehouse.

(b) What do you mean by data warehouse? Discuss the various component of

data warehouse architecture. How data warehouse and data mining integration is achieved? 7+7=14

UNIT—II

4. (a) What are the different kinds of data? Discuss Data mining systems on the basis of various classifications.

   (b) What kind of problems we generally face in data? For which problems data cleaning is appropriate? Discuss the various methods for data cleaning. 7+7=14

5. What do you mean by Concept Description? Why we perform attribute relevance analysis? Explain its methods. 14

UNIT-III

6. What do you mean by association rule mining? Discuss the algorithm of Apriori for mining association rules in transactional databases. 14

7. (a) Discuss the issues regarding classification and prediction.

   (b) Explain classification by Back propagation with suitable example. 7+7=14

UNIT-IV

8. (a) Discuss the application where data mining may be applied? What kind of answer we can and we cannot expect from a data mining algorithm application?

   (b) What is the role of Data Mining tools in Data mining applications? List out the names of various DM tools. 7+7=14

9. Distinguish between clustering and classification tasks. What distance measure is appropriate for categorical data for the clustering task? 14