Roll No. ..... Total Pages : 3

OMCA/M-20 10650

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. All questions carry equal marks.

Compulsory Question

1. Answer the following questions in brief : 2×8=16

(a) Distinguish between roll-up and drill down OLAP operations.

(b) Define metadata and data mart.

(c) "Whether all patterns are interesting", Comment on the statement.

(d) Write a note on DMQL.

(e) What is support and confidence? Describe.

(f) What do you understand by data mining prediction techniques?

(g) Write a brief note on density-based clustering methods.

(h) What is meant by outliers? Write its different types.

UNIT–I

2. (a) What is Data warehouse? Discuss its important characteristics. Compare among OLAP and OLTP. 8

(b) Why Data warehouse schema is important to understand? Outline snow-flake schema with example. 8

3. (a) Outline and explain each component of three-tier architecture of Data warehouse. 8

(b) Explore the methods for implementation of Data warehouse systems. 8

UNIT–II

4. (a) Under which criteria a Data mining system may be classified? Explain the Life cycle of Knowledge discovery process. 8

(b) Define Data mining? What are data mining primitives? Discuss the system architectures for data mining. 8

5. (a) Discuss the factors which influence the quality of Data. How data can be cleaned and reduced? 8

(b) Differentiate between Descriptive and Predictive data mining models. 8

UNIT–III

6. What do you mean by Association rule mining? Discuss the algorithm of Apriori for mining association rules in transactional database with candidate key and without candidate key generation. 16

7. What is meant by Classification? Which data mining methods are comes under supervised learning? How the classifier accuracy determined? Also explain it various types. 16

UNIT–IV

8. Define Clustering. Discuss the desired requirements of Cluster analysis. Explain hierarchical and partitioning technique of the cluster analysis with appropriate distance function. 16

9. (a) Draw a comparative analysis on any seven Data mining tools. 8

(b) Describe the use of Data mining in Biology, city planning, sports and financial sectors. 8