Reg. No. : | | | | | | | | | | | |

Y 1514

M.C.A. DEGREE EXAMINATION, AUGUST/SEPTEMBER 2008.

Elective

DMC 1628 — DATA WAREHOUSING AND DATA MINING

Time : Three hours Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. Define data mining.

2. Give the classification of Data Mining systems.

3. Why data preprocessing?

4. What do you mean by concept hierarchy?

5. Differentiate classification and clustering techniques.

6. Give the advantages of decision tree based classification systems.

7. Define OLAP.

8. Give the need for data integration.

9. Give atleast three applications of data mining.

10. Define web usage mining.

PART B — (5 × 16 = 80 marks)

11. (a) Briefly explain the statistical and database perspectives of data mining. Explain the functionalities of data mining and their relevant data mining techniques.

Or

(b) With the neat diagram, explain the KDD process. Explain each step of KDD process with their significance.

12. (a) Explain the data preprocessing step of data mining process in detail in its suitable examples.

Or

(b) Define association rule mining. Explain the Apriori-algorithm for mining association rules with the help of an example.

13. (a) Explain the Bayesian classification algorithms with a suitable example. Compare this approach with decision tree approach.

Or

(b) Explain any partitioning algorithm for cluster analysis with an example. Explain how this approach is different from hierarchical approach.

14. (a) In detail, explain the various components and their significance of data warehousing. Give the classification of various types of data warehouses.

Or

(b) Explain the implementation of data warehouses with a star schema model with suitable example. Compare this model with snow flake schema.

15. (a) Write short notes on DB-miner tool. Explain various business applications of Data Mining systems.

Or

(b) Explain the taxonomy of web mining. Compare web mining with information retrieval. Explain how spatial databases complicate the processes of data mining.