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

Question Paper Code : 11231

M.E./M.Tech. DEGREE EXAMINATION, NOVEMBER/DECEMBER 2014.

Elective

Computer Science Engineering

CS 9264/UCP 9164/CS 964/10244 CSE 51 — DATA WAREHOUSING AND DATA MINING

(Common to M.E. Software Engineering and M.Tech. Information Technology/ M.E. Networking and Internet Engineering/M.Tech. Multimedia Technologies/ M.Tech. Main Frames Technology)

(Regulation 2009)

Time : Three hours | Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. What is the need for data warehouse?

2. What is the use of back end process in data warehouse design?

3. What do you mean by 'noisy data'? Give an example.

4. What are support and confidence in association rule mining?

5. Differentiate prediction from classification.

6. Define Bayes theorem.

7. What is dissimilarity matrix?

8. Mention any four applications of cluster analysis.

9. What is a range query? Give an example.

10. Define web mining.

PART B — (5 × 16 = 80 marks)

11. (a) Explain Data Warehousing architecture with a neat sketch and examples. (16)

Or

(b) Explain OLAP operations and types of services in detail with necessary diagrams and examples. (16)

12. (a) Explain the various functionalities of data mining with respect to an example. (16)

Or

(b) Explain apriori algorithm for mining frequent item sets with an example. (16)

13. (a) (i) Describe the issues regarding preprocessing the data for classification and prediction. (8)

(ii) Explain the classification by decision tree induction. (8)

Or

(b) Compare the advantages and disadvantages of eager classification versus lazy classification. (16)

14. (a) (i) Discuss the different types of clustering methods. (8)

(ii) Describe the working of PAM (Partioning Around Medoids) algorithm. (8)

Or

(b) Describe the working of DBSCAN algorithm and explain the concept of clusters used in DBSCAN. (16)

15. (a) Explain the following:

(i) Spatial data mining.

(ii) Web mining. (8+8)

Or

(b) Explain the following:

(i) Multimedia data mining.

(ii) Text mining. (8+8)