C 3173

B.E./B.Tech. DEGREE EXAMINATION, NOVEMBER/DECEMBER 2008.

Sixth Semester

Computer Science and Engineering

CS 1004 — DATA WAREHOUSING AND MINING

(Regulation 2004)

Time : Three hours   Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1.   Define meta data used in Data ware house environment.

2.   Define Outlier Mining.

3.   What do you mean by discretization?

4.   What are the Data Mining primitives?

5.   What is pattern?

6.   What is OLAP? Explain.

7.   What are the methods used to solve the classification problem?

8.   What is the use of 'decision tree'?

9.   Define Join indexes.

10. What is the need for spatial databases?

11. (a) Explain the characteristics of data stored in the data ware house. Describe the architecture of data warehouse.

Or

(b) Explain about VLDB and how indexing improves the performance.

12. (a) (i) Explain the steps of Datamining process with neat diagram. (8)

(ii) Differentiate between predictive and descriptive data mining. (8)

Or

(b) Explain the need for data processing and different forms of data processing giving an example for each.

13. (a) What is association rule mining? How are the multilevel association rules framed from transaction databases?

Or

(b) (i) Explain about apriori algorithm.

(ii) Explain how Back Propagation algorithm can be used in data mining.

14. (a) Explain the classification method by Decision tree induction approach and explain an algorithm for inducing a decision tree from training sample.

Or

(b) Discuss about the keyword based association analysis and document classification with an example.

15. (a) Explain the descriptive mining of complex data objects.

Or

(b) Write short notes on : (8 + 8)

(i) Multimedia databases.

(ii) Spatial Databases.

———————