T 3159

B.E./B.Tech. DEGREE EXAMINATION, APRIL/MAY 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. Write down the applications of data warehousing.

2. When is data mart appropriate?

3. What is Concept Hierarchy? Give an example.

4. What are the various forms of data preprocessing?

5. Write the two measures of Association Rule.

6. Define conditional pattern base.

7. List out the major strength of decision tree method.

8. Distinguish between classification and clustering.

9. Define a Spatial database.

10. List out any two various commercial data mining tools.

PART B — (5 × 16 = 80 marks)

11. (a) (i) With a neat sketch explain the architecture of a data warehouse. (10)

        (ii) Discuss the typical OLAP operations with an example. (6)

Or

   (b) (i) Discuss how computations can be performed efficiently on data cubes. (10)

        (ii) Write short notes on data warehouse meta data. (6)

12. (a) (i) Explain various methods of data cleaning in detail. (8)

        (ii) Give an account on Data Mining Query Language. (8)

Or

   (b) How is Attribute–Oriented Induction implemented? Explain in detail. (16)

13. (a) Write and explain the algorithm for mining frequent item sets without candidate generation. Give relevant example. (16)

Or

   (b) Discuss the approaches for mining multi level Association rules from the transactional databases. Give relevant example. (16)

14. (a) (i) Explain the algorithm for constructing a decision tree from training samples. (12)

        (ii) Explain Bayes Theorem. (4)

Or

   (b) Explain the following clustering methods in detail :

        (i) BIRCH (8)

        (ii) CURE. (8)

15. (a) What is a multimedia database? Explain the methods of mining multimedia database? (16)

Or

   (b) (i) Discuss the social impacts of data mining. (6)

        (ii) Discuss spatial data mining. (10)

T 3159