

LIB  
31/5/14 FN

Reg. No. :

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

**Question Paper Code : 61225**

M.E./M.Tech. DEGREE EXAMINATION, MAY/JUNE 2014.

Elective

Computer Science and Engineering (AUTC)

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

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

(Regulation 2009/2010)

Time : Three hours

Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. Is the Data warehouse a prerequisite for data mining? Does the data ware house help data mining, if so, in what ways?
2. Write down the six distinct steps that comprises the knowledge discovery process.
3. How is a data warehouse different from a database? How are they similar?
4. In what architectural component does OLAP fit in?
5. What is constraint based association mining?
6. Write about the two step process in association rule mining.
7. Why is tree pruning useful in decision tree induction?
8. Define data and task parallelism.
9. What do you mean by Harvest system?
10. What are spatial rules?

PART B — (5 × 16 = 80 marks)

11. (a) Define each of the following mining functionalities: Characterization, discrimination, association and correlation analysis, classification, regression, clustering and outlier analysis. Give examples of each data mining functionality using a real life database that you are familiar with. (16)

Or

- (b) A statistical perspective on data mining – Discuss in detail. (16)
12. (a) (i) What are the various data sources for the data warehouse? (8)
- (ii) What are the three major types of Meta data in a data warehouse? Briefly mention the purpose of each type. (8)

Or

- (b) Suppose that a data warehouse consists of the three dimensions time, doctor and charge, while charge is the fee that a doctor charges a patient for a visit
- (i) Enumerate three classes of schemas that are popularly used for modeling data warehouses. (8)
- (ii) Draw a schema diagram for the above data warehouse using one of the schema listed in. (8)
13. (a) Define classifications. Explain the different types of classification techniques with suitable example.

Or

- (b) What do you mean by data reduction? Explain Wavelet Transform, Principal component Analysis and parametric data reduction.
14. (a) Write the basic algorithm for inducing a decision tree training tuples. Also explain the three possibilities for partitioning tuples based on the splitting criterion.

Or

- (b) Mention all the advanced association rule technique. Briefly explain any three advanced association rule techniques.
15. (a) What do you mean by Web usage mining? Explain rule with examples.

Or

- (b) Briefly explain the categorization of major clustering methods.