Reg. No. : ☐☐☐☐☐☐☐☐☐☐☐☐☐

Question Paper Code : 31170

B.E./B.Tech. DEGREE EXAMINATION, MAY/JUNE 2013.

Seventh Semester

Computer Science and Engineering

080230056 — DATA WAREHOUSING AND DATA MINING

(Regulation 2008)

Time : Three hours   Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. What are dependent and independent data marts?

2. How is a data warehouse different from a database?

3. What is Smoothing?

4. How can task-relevant data be collected in a relational database?

5. Define anti-monotone property.

6. When we can say the association rules are interesting?

7. Define the concept of classification.

8. What are intervals scaled variables? Give examples.

9. What is meant by authoritative web pages?

10. What is Time Series Analysis?

PART B — (5 × 16 = 80 marks)

11. (a) Explain the design and construction of a data warehouse.   (16)

Or

    (b) (i) Discuss the various schematic representations in multidimensional model.   (10)

        (ii) Write the difference between ROLAP and MLOAP.   (6)

12. (a) Describe the methods for data integration and data transformation. Give examples. (16)

Or

(b) (i) Give an account on Data Mining Query Language (DMQL). (6)

(ii) Discuss why analytical characterization is needed and how it can be performed. (10)

13. (a) (i) Construct a FP Tree for the given database table 13a(i) and find the frequent patterns ($\sigma = 40\%$). (10)

Tid	Items
100	a c d f g o m p
200	a b c f l m o
300	g f h j o
400	b c k s p
500	a c e f l m n p

Table 13a(i)

(ii) Write the procedure for Generating Association rules from frequent item sets. Give an example. (6)

Or

(b) What are the things suffering the performance of apriori candidate generation Technique? Give few techniques to improve the efficiency of apriori algorithm. (16)

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

(ii) Write Bayes Theorem. (4)

Or

(b) Explain the following clustering methods in detail : (8)

(i) BIRCH

(ii) CURE. (8)

15. (a) Discuss the method of mining text databases in detail. (16)

Or

(b) (i) Describe the use of DBMiner. (8)

(ii) Explain how data mining is used for intrusion detection. (8)

———————

2   31170