

FACULTY OF ENGINEERING**B.E. IV/IV Year (CSE) II Semester (Main) Examination, May/June, 2011****DATA MINING**

Time : 3 Hours]

[Max. Marks : 75

*Answer all questions from Part A.
Answer any five questions from Part B.*

Part A – (Marks : 25)

- | | |
|--|---|
| 1. Explain the need for preprocessing the data. | 2 |
| 2. Differentiate between ROLAP and OLAP. | 3 |
| 3. What is meant by DMQL? | 2 |
| 4. Define attribute generalization. | 3 |
| 5. What is meant by correlation analysis? | 2 |
| 6. Define (a) Itemset, (b) Frequent Itemset (c) Candidate Set. | 3 |
| 7. Differentiate between classification and prediction. | 2 |
| 8. How is classifier accuracy measured ? | 3 |
| 9. Define clustering with an example. | 3 |
| 10. What is meant by Multimedia mining? | 2 |

**Part B – (Marks : 50)**

- | | |
|---|----|
| 11. (a) Discuss about 3-tier architecture of data warehouse. | 5 |
| (b) Explain about Data Mining Functionalities. | 5 |
| 12. Explain the need to perform Attribute Relevance Analysis. Discuss briefly the various methods of Attribute Relevance Analysis. | 10 |
| 13. Explain Apriori algorithm with a suitable example. | 10 |
| 14. Explain the classification process by Decision Tree with an example. | 10 |
| 15. (a) Given two objects (22,1,42,10) and (20,0,36,8). Compute Euclidean distance and Minkowski distance (with p=3) between 2 objects. | 10 |
| (b) Explain the K – Means clustering algorithm with an example. | 10 |

16. Write short notes on :

10

- (a) Data Reduction technique.
- (b) Multilevel association Rules Mining.

17. Write short notes on any **two** :

10

- (a) Mining text databases
- (b) Data Cleaning
- (c) DMQL Syntax for DM primitive "Kind of Knowledge to be mined".