FACULTY OF INFORMATICS

B.E. IV/IV Year (IT) I Semester (Main) Examination, December 2010
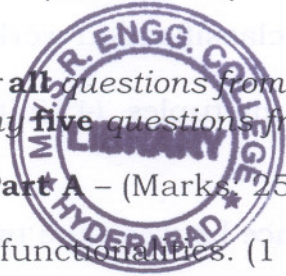
DATA WAREHOUSING & DATA MINING

(Elective – II)

Time : 3 Hours]   [Max. Marks : 75

Answer all questions from part A.
Answer any five questions from part B.

Part A – (Marks: 25)

1. Define any three Data mining functionalities. (1 Unit)   3

2. Why we preprocess the data? Give any 3 reasons. (1 Unit)   3

3. Why we kept Data warehouse separately from the operational databases?(2 Unit)2

4. Define the following terms. (2 Unit)   3

   (a) Frequent Itemsets   (b)   Frequent subsequences

   (c) Frequent Substructures.

5. Define classification and prediction. (3 Unit)   2

6. State the principle of Cluster Analysis. (4 Unit)   2

7. What is meant by single dimension Association Rule and Multi dimensional
   Association Rule? Give examples. (2 Unit)   3

8. What is meant by Supervised learning and unsupervised learning? (4 Unit) 2

9. Define Text mining and www mining. (5 Unit)   2

10. Define the following OLAP operations . (2 Unit)   3

    (a) Slice   (b) Dice   (c)   Pivot

Part B – (Marks: 50)

11. (a) Explain Data Mining functionalities. (1 Unit)   5

    (b) Explain any three Data cleaning techniques. (1 Unit)   5

12. (a) Explain three-tier data warehousing architecture. (2 Unit)   7

    (b) Explain data warehouse applications. (2 Unit)   3

13. (a) Explain Apriori Algorithm with example. (2 Unit)   6

(b) Explain different methods for improving the efficiency of Apriori Algorithm. (2 Unit)   4

14. (a) Describe the data classification process with a neat diagram. (3 Unit)   5

(b) How does the Naive Bayesian classification works? Explain. (3 Unit)   5

15. Given two objects represented by the tuples. (4 Unit)

(22,1,42,10) & (20,0,36,8)

(i) Compute the Euclidean distance between the two objects.   3

(ii) Compute the Manhanttan distance between the two objects.   3

(iii) Compute the Minkocoski distance between the two objects using p = 3.   4

16. (a) Explain Data Integration Techniques. (1 Unit)   5

(b) Explain different methods for data transformation. (1 Unit)   5

17. (a) Explain similarity search in Multimedia data. (5 Unit)   5

(b) Explain Mining Associations in multimedia Data. (5 Unit)   5