Code No. : **6477/N**

FACULTY OF ENGINEERING
B.E. 4/4 (CSE) (II Semester) (New) (Main) Examination, June 2010
DATA MINING

Time : 3 Hours]   [Max. Marks : 75

Note : Answer **all** questions from Part **A**, answer **any five** questions from Part **B**.

PART – A   (25 Marks)

1.	What are the steps in a KDD process ?	2
2.	Give various techniques for data transformation.	2
3.	What is information gain ?	2
4.	How do you perform class comparison ?	2
5.	Explain the criteria for classifying association rules.	3
6.	Write a short note on linear regression.	3
7.	Write a short note on estimation of classifier accuracy.	3
8.	How do you compare various classification methods ?	3
9.	Distinguish between data characterization and data discrimination.	2
10.	Write a short note on OLAP servers.	3

PART – B   (50 Marks)

11. Discuss the Apriori algorithm in detail.   **10**

12. a) Explain the architecture of a data mining system.   **5**
    b) Explain on what kind of data is data mining performed.   **5**

13. Explain the back propagation algorithm in detail.   **10**

14. Describe the CURE algorithm. What is its advantage ?   **10**

15. What are the various types of data in cluster analysis ? How do you compute their distances ?   **10**

16. a) Describe the process of a data warehouse design.   **5**
    b) What is a measure ? How are they computed ?   **5**

17. Describe in detail the various data reduction techniques.   **10**

(This paper contains 1 page)   1   **1,000**