DATA MINING (E-III)

Paper–CSE-416-N

Time Allowed : 3 Hours]   [Maximum Marks : 75

Note : Attempt **five** questions in all, selecting at least **one** question from each Unit. All questions carry equal marks.

UNIT–I

1. (a) Explain different techniques of data transformation through normalization.   5

   (b) Explain different data smoothing techniques.   5

   (c) Explain the steps to reduce dimentionality using Principal Components Analysis (PCA).   5

2. (a) Differentiate among enterprise warehouse, data mart and virtual warehouse models.   7

   (b) What is OLAP? Discuss typical OLAP operations.   8

UNIT–II

3. (a) Explain syntax of any four Data Mining Query Language (DMQL).   8

Learn Loner

(b) Explain data generalization technique using attribute induction. 7

4. Write notes on the following descriptive Statistical measures used for large databases :

(a) Boxplot analysis. 7

(b) Scatter plot. 8

UNIT–III

5. (a) How do you min closed and max patterns? Explain. 5

   (b) Write Apriori algorithm for discovering frequent itemsets for mining Boolean association rules. 10

6. (a) Explain constraint-based frequent pattern mining with an example. 7

   (b) Explain major steps of decision tree induction classifier. 8

UNIT–IV

7. (a) What do you mean by Cluster Analysis (CA)? Explain main requirements of CA. 8

   (b) Explain K-means partitioning method with an example. 7

Learn Loner

8. (a) What are symbolic sequences? Explain sequential pattern mining in symbolic sequences. 7

   (b) What are spatial and multimedia databases? Explain. 8

Learn Loner