44 (6) DMAW 6·1 (N/O)

2020

DATA MINING AND WAREHOUSING

Paper : 6·1·2

(New & Old Course)

Full Marks : 40

Time : Two hours

**The figures in the margin indicate
full marks for the questions.**

Answer **any four** questions.

1. (a) What is data warehouse ? Why data warehouse is subject-oriented, time-varying and integrated ? 5

   (b) What is OLAP operation ? Write the difference between R-OLAP and M-OLAP. 2+3=5

2. What are the different stages of KDD ? How is data mining different from KDD ? Give a brief account of data mining techniques. 3+3+4=10

3. Discuss the concepts of frequent itemsets, support and confidence. Define border-set. 10

4. Explain Apriori Algorithm with suitable example. 10

5. Discuss the working of PAM algorithm. Compare its performance with CLARA and CLARANS. 6+4=10

6. Explain the following concepts in the context of DBSCAN (any five) 2×5=10

   (i) E-neighbourhood of an object

   (ii) Core-object

   (iii) Directly Density-Reachable

   (iv) Density-Reachable

   (v) Density-Connected

   (vi) Noise

7. What is a decision tree ? Write down the advantages and shortcomings of decision tree classification. Explain splitting attribute and splitting criterion. 2+4+4=10

8. Write short notes on : (any two) 5×2=10

   (i) Text mining

   (ii) Bitmap indexing

   (iii) Supervised and unsupervised learning

   (iv) CART

   (v) Web mining

   (vi) Spatial mining

   (vii) BIRCH

———————

800