

- (e) CART
- (f) Apriori Algorithm
- (g) Partitioning around medoids (PAM) algorithm

Total number of printed pages—4

52 (6) ITB-HE-6026/IT 6-2

2024

DATA MINING AND WAREHOUSING

Paper : ITB-HE-6026/IT 6-2

Full Marks : 80

Time : Three hours

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

1. Define the following terms (**any five**) : 3×5=15
- (a) Data Mining
 - (b) Multidimensional Data
 - (c) Frequent Itemset
 - (d) Data Cube
 - (e) Meta Data
 - (f) Support
 - (g) Pruning

2. Answer the following questions **(any six)** :

5×6=30

- (a) Define Data Warehouse. What are some of the important characteristics of data warehouse ?
- (b) What is association rule? Why do we need association rule? Give example.
- (c) What is clustering? What are the different types of clustering? Briefly explain.
- (d) What is KDD? What are the different steps in KDD? How is KDD different from data mining?
- (e) What is web content mining? Briefly explain.
- (f) Discuss few of the major application areas of data mining.
- (g) Explain CLARA, CLARANS.
- (h) What do you mean by Data Warehouse Deployment and Maintenance? Explain.

3. Differentiate between the following terms :

(any five)

3×5=15

- (a) Database and Data Warehousing
- (b) OLTP and OLAP
- (c) R-means and R-medoids
- (d) Clustering and Classification
- (e) Relational OLAP and Multidimensional OLAP
- (f) Hierarchical Clustering and Density-based Clustering
- (g) Partitional clustering and Hierarchical clustering

4. Write short notes on **(any five)** : 4×5=20

- (a) DBSCAN
- (b) Temporal Mining
- (c) Pincer-search Algorithm
- (d) Decision Tree