

No. of Printed Pages : 3

MCS-221

**MASTER OF COMPUTER
APPLICATIONS (MCA-NEW)**

Term-End Examination

June, 2023

**MCS-221 : DATA WAREHOUSING AND DATA
MINING**

Time : 3 Hours

Maximum Marks : 100

Weightage : 70%

Note : (i) *Question No. 1 is compulsory.*

(ii) *Answer any **three** questions from the
rest.*

1. (a) Discuss the following data preprocessing stages briefly : 10
- (i) Data Cleaning
 - (ii) Data Integration
 - (iii) Data Reduction
 - (iv) Data Transformation

P. T. O.

- (b) Define Aggregate fact tables and derived dimension tables. What are their significance ? Give an example. Also mention their advantages and disadvantages. 10
- (c) Enumerate the key challenges in data warehouse design. 5
- (d) Differentiate between a data lake and a data warehouse. 5
- (e) Write and explain the Apriori algorithm to identify the most frequently occurring elements and meaningful associations in a dataset. 10
2. (a) What is Cluster Analysis ? How is this used in Data Mining ? Give an example. Also mention few applications of cluster analysis in data mining. 10
- (b) List and discuss various types of Web-mining. 10
3. (a) Explain the following techniques for Dimensionality Reduction : 10
- (i) Feature Selection
- (ii) Feature Extraction
- (b) Discuss the layered implementation of ETL in a Data Warehouse. 10

[3]

4. (a) Define OLAP. Differentiate between Multicube and Hypercube. Mention the applications of OLAP reporting system. 10
- (b) List and explain the following types of Data Warehouses : 10
- (i) Enterprise Data Warehouse (EDW)
 - (ii) Operational Data Store
5. Write short notes on the following : 4×5=20
- (a) ELT *vs.* ETL
 - (b) Data Marts
 - (c) Applications of Data Mining
 - (d) OLAP data cube operations