

**MASTER OF COMPUTER
APPLICATIONS (MCA-NEW)
Term-End Examination
December, 2025**

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

Time : 3 Hours

Maximum Marks : 100

(Weightage : 70%)

Note : *Question No. 1 is compulsory. Answer any **three** questions from the rest.*

1. (a) With the help of a block diagram, explain Extract, Transform and Loading (ETL) process of a Data Warehouse. Also, discuss how it is different from ELT process. 10

- (b) With the help of an example, explain the Binning method for solving the problem of Noisy data. 10
 - (c) Discuss the following text transformation technique that help text sentences into numeric vectors :
“Bag-of-Words (BoW)”
Also give an example. Mention the drawbacks of using BoW. 10
 - (d) Write and explain the K-Nearest Neighbour’s Algorithm. Also mention its advantages and disadvantages. 10
2. (a) What is clustering technique in Data Mining ? Give an example. Explain the following clustering methods briefly : 10
- (i) Density-based method
 - (ii) Constraint-based method
- (b) Explain briefly the following text preprocessing stages used in typical NLP system : 10
- (i) Segmentation

- (ii) Tokenization
 - (iii) Removal of stop words
 - (iv) Stemming
3. (a) Explain snowflake and fact constellation schemas of dimensional modeling in a Data Warehouse with the help of an example for each. 10
- (b) Describe briefly the following types of OLAP architectures along with a block diagram for each : 10
- (i) ROLAP architecture
 - (ii) MOLAP architecture
4. (a) Explain Real-Time Data Warehouse (RTDW) architecture along with a block diagram. Give examples where these type of data warehouses are useful. 10
- (b) What are Data Marts ? How are they different from a Data Warehouse ? Discuss the design of the Data Marts. 10

5. Write short notes on the following : $4 \times 5 = 20$

- (a) Data Lake
- (b) Star Schema
- (c) Market Basket Analysis
- (d) Cloud Data Warehouse

× × × × ×