Roll No. | | | | | | | | | | | | Total No. of Pages : 02

Total No. of Questions : 09

M.Sc.(IT) (Sem.–4)

DATA WAREHOUSING AND DATA MINING

Subject Code : PGCA1941

M.Code : 79564

Date of Examination : 24-05-2023

Time : 3 Hrs. Max. Marks : 70

INSTRUCTIONS TO CANDIDATES :

1. **SECTION-A is COMPULSORY consisting of TEN questions carrying TWO marks each.**
2. **SECTION - B & C have FOUR questions each.**
3. **Attempt any FIVE questions from SECTION B & C carrying TEN marks each.**
4. **Select atleast TWO questions from SECTION - B & C.**

SECTION-A

1. Write briefly:

a) Differentiate between OLAP and OLTP.

b) Briefly explain the ETL process.

c) What is the need for data warehousing?

d) What are the various techniques available for data transformation?

e) Discuss snowflake schema.

f) What is Bayes theorem?

g) What is the difference between Nominal and Ordinal Data?

h) How to choose the value of K in the K-Nearest Neighbour method?

i) Differentiate between classification and clustering.

j) What is the use of DBSCAN method?

SECTION-B

2. Explain the process of data preprocessing with all its types in detail.

3. Draw and explain the OLAP Three-tier architecture.

4. Discuss mapping of data warehouse to a multiprocessor architecture.

5. a. Discuss five ways of handling missing values.

   b. Discuss three ways of handling noisy data.

SECTION-C

6. Discuss the comparison between the following pairs :

a) Predictive and Descriptive Data

b) Linear and Multiple Regression

c) K-Mean and K-Mediods

d) Hierarchical and Density-based Clustering Techniques

e) KNN and SVM

7. Write the algorithm for the K-Means Clustering and demonstrate the same with an example.

8. Explain the following in detail using examples :

a) Associating Rule Mining

b) Apriori Algorithm.

9. a) What do the terms "Naïve" and "Bayes" mean in the Naive Bayes classification algorithm?

   b) Explain the mathematics behind the Naive Bayes model.

NOTE : Disclosure of Identity by writing Mobile No. or Marking of passing request on any paper of Answer Sheet will lead to UMC against the Student.