

**G-3/374/22**

Roll No. ....

**III Semester Examination, January 2022****M.Sc.****INFORMATION TECHNOLOGY****Paper IV****(Data Mining and Warehousing)**

Time : 3 Hours ]

[ Max. Marks : 100

**Note :** All questions are compulsory. Question Paper comprises of 3 Sections. Section A is objective type/multiple choice questions with no internal choice. Section B is short answer type with internal choice. Section C is long answer type with internal choice.

**SECTION A****1×10=10****(Objective Type/Multiple Choice Questions)**

Choose the correct answer :

1. The @active data warehouse architecture includes which of the following ?
  - (a) At least one data mart
  - (b) Data that can extracted from numerous internal and external sources
  - (c) Near real time updates
  - (d) All of the above

P.T.O.

2. A goal of data mining includes which of the following ?
  - (a) To explain some observed event or condition
  - (b) To confirm that data exists
  - (c) To analyze data for expected relationship
  - (d) To create a new data warehouse
3. What do data warehouses support ?
  - (a) OLAP
  - (b) OLTP
  - (c) OLAP and OLTP
  - (d) Operational Database
4. What is KDD in data mining ?
  - (a) Knowledge Discovery Database
  - (b) Knowledge Discovery Data
  - (c) Knowledge Data Definition
  - (d) Knowledge Data House
5. Which one of the following can be considered as the final output of the hierarchical type of clustering ?
  - (a) A tree which display hoke the close things are to each other
  - (b) Assignment of each point to clusters
  - (c) Finalize estimation of cluster centroids
  - (d) None of the above

**G-3/374/22**

6. Which one of the clustering technique needs the merging approach ?  
 (a) Partitioned (b) Naive Bayes  
 (c) Hierarchical (d) Both (a) and (c)
7. The issues like efficiency, scalability of data mining algorithm comes under :  
 (a) Performance issues  
 (b) Diverse data type issues  
 (c) Mining methodology and user interaction  
 (d) All of the above
8. The term 'DMQL' stands for :  
 (a) Data Marts Query Language  
 (b) DB Miner Query Language  
 (c) Data Mining Query Language  
 (d) None of the above
9. What is the use of data cleaning ?  
 (a) To remove the noisy data  
 (b) Correct the inconsistencies in data  
 (c) Transformations to correct wrong data  
 (d) All of the above
10. Which of the following is true for classification ?  
 (a) Subdivision of a set  
 (b) A measure of the accuracy

- (c) The task of assigning a classification  
 (d) All of the above

**SECTION B****6×5=30****(Short Answer Type Questions)****Note :** Answer the following questions in **250** words.**Unit-I**

1. Explain the data mining functionality.

*Or*

Explain major requirements and challenges in data mining.

**Unit-II**

2. Explain the data preprocessing techniques in detail.

*Or*

Discuss data discretization and concept hierarchy generation.

**Unit-III**

3. Explain various classification methods.

*Or*

Discussion about prediction in detail.

**Unit-IV**

4. Explain density based clustering method in detail.

[ 5 ]

Or

Discuss about the grid based method.

### Unit-V

5. Explain the process of mining the World Wide Web.

Or

Explain spatial mining and time series mining.

### SECTION C

12×5=60

(Long Answer Type Questions)

**Note :** Answer the following questions in **500** words.

### Unit-I

1. Explain about data mining tools widely used in industry.

Or

Write in detail about data mining implementation process.

### Unit-II

2. Explain parametric and non-parametric method of data reduction.

Or

Discuss data integration and transformation in detail.

G-3/374/22

P.T.O.

[ 6 ]

### Unit-III

3. Explain mining multidimensional data from transactional databases and relational database.

Or

Explain Naive Bayesian classification in detail with example.

### Unit-IV

4. Elaborate the various partitioning method in detail.

Or

Explain major categories of clustering methods.

### Unit-V

5. Discuss about some or the case studies in data mining application.

Or

Define the following :

- (i) Web mining
- (ii) Web content mining
- (iii) Text mining
- (iv) Multimedia mining

★ ★ ★ ★ ★ c ★ ★ ★ ★ ★

G-3/374/22

6/50