

Printed Page:-

Subject Code:- AMTCSE0114

Roll. No:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

NOIDA INSTITUTE OF ENGINEERING AND TECHNOLOGY, GREATER NOIDA

(An Autonomous Institute Affiliated to AKTU, Lucknow)

M.Tech

SEM: I - THEORY EXAMINATION (2022 - 2023)

Subject: Data Warehousing & Data Mining

Time: 3 Hours

Max. Marks: 70

General Instructions:

IMP: Verify that you have received the question paper with the correct course, code, branch etc.

1. This Question paper comprises of **three Sections -A, B, & C.** It consists of Multiple Choice Questions (MCQ's) & Subjective type questions.
2. Maximum marks for each question are indicated on right -hand side of each question.
3. Illustrate your answers with neat sketches wherever necessary.
4. Assume suitable data if necessary.
5. Preferably, write the answers in sequential order.
6. No sheet should be left blank. Any written material after a blank sheet will not be evaluated/checked.

SECTION A

15

1. Attempt all parts:-

- 1-a. A data warehouse is which of the following? (CO1) 1
- (a) Can be updated by end users.
 - (b) Contains numerous naming conventions and formats.
 - (c) Organized around important subject areas.
 - (d) Contains only current data
- 1-b. The generic two-level data warehouse architecture includes which of the following? (CO2) 1
- (a) At least one data mart
 - (b) Data that can extracted from numerous internal and external sources
 - (c) Near real-time updates
 - (d) ORDBMS
- 1-c. The basic idea of the Apriori algorithm is to generate _____ item sets of a particular size & scans.(CO3) 1

- (a) Candidate.
 - (b) Primary.
 - (c) Secondary.
 - (d) Superkey.
- 1-d. The human brain consists a network of _____. (CO4) 1
- (a) Neurons
 - (b) Cells
 - (c) Tissue
 - (d) Muscles
- 1-e. The granularity of the fact is the ____ of detail at which it is recorded. (CO5) 1
- (a) Transformation
 - (b) Summarization
 - (c) Level
 - (d) Transformation and summarization

2. Attempt all parts:-

- 2.a. Distinguish between Data Mart & Meta Data? (CO1) 2
- 2.b. What is the difference between operational data store and data warehouse? (CO2) 2
- 2.c. Why data cleaning routines are required? (CO3) 2
- 2.d. What is the need for a separate database for decision makers? (CO4) 2
- 2.e. Explain various data reduction techniques in the pre-processing steps of data mining. (CO5) 2

SECTION B

20

3. Answer any five of the following:-

- 3-a. Differentiate between Data warehouse and OLAP. (CO1) 4
- 3-b. Enlist some commands of DDL, DML and DCL? (CO1) 4
- 3-c. Differentiate between Knowledge Discovery and Data Mining. (CO2) 4
- 3-d. Draw the diagram for key steps of data mining. (CO2) 4
- 3.e. What is the decision tree algorithm? Write the steps involved in it. (CO3) 4
- 3.f. Explain about the three-tier Data Warehouse architecture with neat diagram. (CO4) 4
- 3.g. Discuss about Spatial Data Mining. (CO5) 4

SECTION C

35

4. Answer any one of the following:-

- 4-a. Illustrate with an example which schema is better while designing the data warehouse. (CO1) 7
- 4-b. What is OLAP? What are the functionalities of OLAP? List different types of OLAP servers. (CO1) 7

5. Answer any one of the following:-

- 5-a. Define data discretization. Explain the various approaches in data discretization. (CO2) 7
- 5-b. What are the issues in data integration? (CO2) 7

6. Answer any one of the following:-

- 6-a. Define Classification. How it different from Prediction? Explain the process of classification. (CO3) 7
- 6-b. Explain the concept of Bayesian classification. How it is helpful in datamining? (CO3) 7

7. Answer any one of the following:-

- 7-a. List out the types of clustering methods. Differentiate between clustering and classification. (CO4) 7
- 7-b. What is the role of Outlier Analysis in data mining? Explain in detail with example? (CO4) 7

8. Answer any one of the following:-

- 8-a. Discuss various sectors in present scenario where data mining is done and is used frequently? (CO5) 7
- 8-b. Explain data modeling life cycle in detail. (CO5) 7