

GUJARAT TECHNOLOGICAL UNIVERSITY
MBA-SEMESTER-III-EXAMINATION-WINTER-2024

Subject Code: 1539603**Date: 13/12/2024****Subject Name: Data Warehousing and Data Mining****Time: 10:30 AM TO 01:30 PM****Total Marks: 70****Instructions:**

1. Attempt all questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.
4. Use of simple calculators and non-programmable scientific calculators are permitted.

Q. No.	Question Text and Description	Marks
Q.1	(a) 1NF in database normalization. (b) DefineDW Process framework. (c) Meta Data Management (d) Associative Classification. (e) Support Vector Machines (f) DefineK-Means. (g) Define Outlier Analysis.	14
Q.2	(a) Provide a detailed example of an SQL query that uses multiple clauses (e.g., SELECT, WHERE, GROUP BY, HAVING, and ORDER BY). (b) Describe the key features of Structured Query Language (SQL). How is SQL used in database management?	07 07
	OR	
	(b) Discuss the role of Reporting and Query tools in a Data Warehouse environment.	07
Q.3	(a) What is Metadata Management in Data Warehousing? How do Data Marts fit into the Data Warehouse architecture? (b) Discuss the different types of data and their attributes commonly used in Data Mining.	07 07
	OR	
Q.3	(a) What are the various methods used for Mining Frequent Patterns? (b) What are the basic concepts of Classification and Prediction in data mining? How are they different?	07 07
Q.4	(a) What is Bayesian Classification? Explain its principles and provide an example of how it can be applied. (b) What is Cluster Analysis in Data Mining? Discuss the importance of clustering in uncovering patterns in large datasets.	07 07
	OR	
Q.4	(a) Explain the K-Means Clustering algorithm. How does it work, and what are its limitations? (b) What is Constraint-Based Cluster Analysis, and how does it differ from traditional clustering methods?	07 07

Q.5

CASE STUDY:

A large hospital network is looking to enhance patient care and streamline its operations by utilizing the vast amount of data generated from patient records, diagnostic tests, treatment histories, and medical devices. The hospital is struggling to make sense of this data due to its dispersed nature, stored in various departments such as radiology, pharmacy, and electronic health records (EHR).

To solve this problem, the hospital builds a Data Warehouse to centralize and integrate the data from all its departments. Additionally, it uses Data Mining techniques such as classification, clustering, and predictive modeling to discover patterns that can help predict disease outbreaks, optimize treatment plans, and identify high-risk patients for early intervention.

The implementation of data warehousing and data mining not only improves the hospital's ability to manage patient information effectively but also provides valuable insights that lead to better patient outcomes, more efficient use of resources, and cost reduction in operations.

- | | | |
|-----|---|-----------|
| (a) | How can the use of Data Mining techniques in conjunction with a Data Warehouse help the hospital network improve patient outcomes and operational efficiency? | 07 |
| (b) | What are the potential ethical and privacy concerns when implementing Data Mining and Data Warehousing in a healthcare setting, and how can these be mitigated? | 07 |

OR

Q.5

- | | | |
|-----|---|-----------|
| (a) | What challenges might the hospital network face when integrating data from multiple departments into a centralized Data Warehouse, and how can these challenges be addressed? | 07 |
| (b) | How can predictive analytics derived from Data Mining enhance decision-making for healthcare administrators and clinicians? | 07 |
