

**GUJARAT TECHNOLOGICAL UNIVERSITY****BE - SEMESTER-VI (NEW) EXAMINATION – SUMMER 2023****Subject Code:3161610****Date:14-07-2023****Subject Name:Data Warehousing and Mining****Time:10:30 AM TO 01:00 PM****Total Marks:70****Instructions:**

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

|   |  | <b>MARKS</b> |
|---|--|--------------|
| <b>Q.1</b>  | (a) Explain various OLAP operations.   | <b>03</b>    |
|   | (b) Compare Linear and nonlinear regression.   | <b>04</b>    |
|   | (c) Explain Star, Snowflake and Fact constellation” schemas of data warehouse with suitable example.         | <b>07</b>    |
| <b>Q.2</b>  | (a) Define the following terms:  | <b>03</b>    |
|   | 1. OLAP  |              |
|   | 2. OLTP  |              |
|   | 3. OLAM  |              |
| (b) What is data mining integration in data warehousing? Explain with an example  | <b>04</b>  |              |
| (c) Discuss data discretization and concept hierarchy generation.   | <b>07</b>  |              |
| <b>OR</b>   |  |              |
| (c) Explain Naïve Bayesian classification in detail with example.   | <b>07</b>  |              |
| <b>Q.3</b>  | (a) Define techniques to improve the efficiency of Apriori algorithm.  | <b>03</b>    |
|   | (b) Define nominal and ordinal variables   | <b>04</b>    |
|   | (c) What is data transformation? Explain the different data transformation approaches for transforming data. | <b>07</b>    |
| <b>OR</b>   |  |              |
| <b>Q.3</b>  | (a) What is feature selection in data mining?  | <b>03</b>    |
|   | (b) Define Fact Table and dimension table.   | <b>04</b>    |
|   | (c) What is the confusion matrix, and how is it used to evaluate a classifier?                               | <b>07</b>    |
| <b>Q.4</b>  | (a) Define Support & Confidence.   | <b>03</b>    |
|   | (b) Discuss Issues regarding Classification and prediction   | <b>04</b>    |
|   | (c) Describe and explain the different types of clustering methods.  | <b>07</b>    |
| <b>OR</b>   |  |              |
| (a) What is outlier? Discuss different methods for outlier detection.   | <b>03</b>  |              |
| (b) Explain the difference between a data warehouse and a data mart   | <b>04</b>  |              |
| (c) What are the reasons for the presence of ‘noise’ in data collected for mining?<br>Explain the methods to deal with noise. | <b>07</b>  |              |
| <b>Q.5</b>  | (a) Define data mart.  | <b>03</b>    |
|   | (b) What is association rule mining? Explain with an example.  | <b>04</b>    |
|   | (c) What is Decision Tree? Explain how classification is done using decision tree induction.                 | <b>07</b>    |
| <b>OR</b>   |  |              |
| <b>Q.5</b>  | (a) What is a data cube?   | <b>03</b>    |
|   | (b) Discuss the limitations and challenges of data mining  | <b>04</b>    |
|   | (c) Explain Web Mining in detail.  | <b>07</b>    |

\*\*\*\*\*