

GUJARAT TECHNOLOGICAL UNIVERSITY**BE- SEMESTER-VII (NEW) EXAMINATION – WINTER 2024****Subject Code:2170715****Date:30-11-2024****Subject Name: Data Mining and Business Intelligence****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.

- Q.1** (a) Define the following terms: (i) Data Mining (ii) Data mart **03**
 (b) Explain the role of “metadata repository” in a data warehouse and discuss its contents. **04**
 (c) What is KDD? With the help of a neat diagram discuss the steps in the KDD process. **07**
- Q.2** (a) Explain min-max normalization using suitable example. **03**
 (b) With the help of a suitable example explain the following OLAP operations: **04**
 (i) roll-up (ii) drill-down (iii) slice (iv) dice.
 (c) What are the reasons for noise in real world data? Discuss in detail the different ways to deal with noisy data. **07**
- OR**
- (c) Discuss major issues in Data Mining. **07**
- Q.3** (a) Define “Data Discretization”. **03**
 (b) Briefly discuss the following Data Warehouse Schemas: (i) Star (ii) Fact Constellation **04**
 (c) With the help of a suitable example explain the Apriori algorithm for mining frequent itemsets. **07**
- OR**
- Q.3** (a) Define “Numerosity Reduction”. **03**
 (b) Differentiate between OLAP and OLTP. **04**
 (c) What are the limitations of the Apriori algorithm? Discuss any two methods to improve its efficiency. **07**
- Q.4** (a) Define “strong association rule”. **03**
 (b) Discuss main features of Bayesian classification. **04**
 (c) Explain linear and non-linear regression. **07**
- OR**
- Q.4** (a) Explain the following OLAP servers: (i) ROLAP (ii) MOLAP **03**
 (b) Briefly discuss Rule-based classification. **04**
 (c) What is Information Gain? Explain classification based on Decision Tree induction with the help of a suitable example. **07**
- Q.5** (a) Define “clickstream mining”. **03**
 (b) Discuss the role of data scientists. **04**
 (c) Write a detailed note on Text Mining. **07**
- OR**
- Q.5** (a) Define the term “Clustering”. **03**
 (b) What is Big Data? Discuss the 4 V’s of Big Data. **04**
 (c) Discuss applications of data mining for Banking and Finance sector. **07**
