

GUJARAT TECHNOLOGICAL UNIVERSITY
MCA INTEGRATED– SEMESTER VI- EXAMINATION –WINTER-2023

Subject Code:2668603**Date: 12/12/2023****Subject Name: Data Mining****Time: 02:30 PM TO 05:00 PM****Total Marks: 70****Instructions:**

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

- Q-1 (a) Explain Following (Any Seven) : 07**
- (i) Data discrimination
 - (ii) KDD
 - (iii) ETL
 - (iv) KDD
 - (v) OLAP
 - (vi) Bagging and Boosting
 - (vii) Decision Tree
 - (Viii)Data Integration
- (b) What is Data Mining? Explain Data Mining Process in detail. 07**
- Q.2 (a) Explain Differences between Operational Database Systems and Data Warehouses. 07**
- (b) Explain Major Issues in Data Mining. 07**
- OR**
- (b) Explain data warehousing architecture. 07**
- Q.3 (a) Explain Stars, Snowflakes, and Fact Constellations Schemas. 07**
- (b) Explain Starnet Query model. 07**
- OR**
- Q.3 (a) Explain Information Gain, Gain Ratio and Gini Index 07**
- (b) Explain Classification and Regression for Predictive Analysis. 07**
- Q.4 (a) Write down Apriori algorithm for discovering frequent itemsets for mining Boolean Association rules. 07**
- (b) Explain Major Tasks of Data Preprocessing. 07**
- OR**
- Q.4 (a) Explain the major steps of decision tree classification. 07**
- (b) What are Outliers? Explain types of outliers. 07**

- Q.5** (a) Explain Market Basket Analysis in detail **07**
(b) Explain Density Based Cluster Methods. **07**
Clustering Methods in detail.

OR

- Q.5** (a) Explain Partitioning Method. **07**
(b) What is Cluster Analysis? Explain Requirements for Cluster Analysis. **07**
