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

Subject Code: 2668603   **Date: 21/06/2023**
Subject Name: Data 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. Use of simple calculators and non-programmable scientific calculators are permitted.

Q.1	**(a)**	Explain Stars, Snowflakes, and Fact Constellations Schemas.	**07**
	(b)	Draw and explain the Data warehouse architecture.	**07**
Q.2	**(a)**	Describe the steps involved in data mining when viewed as a process of knowledge discovery.	**07**
	(b)	Explain research issues in Data Mining.	**07**
		OR	
	(b)	What is Data Mining? Explain Data Mining Process in detail.	**07**
Q.3	**(a)**	Consider a transactional database where 1, 2, 3, 4, 5, 6, 7 are items. Suppose the minimum support is 60%. Find all frequent item sets using Apriori algorithm.	**07**

ID	ITEMS
T_1	1, 2, 3, 5
T_2	1, 2, 3, 4, 5
T_3	1, 2, 3, 7
T_4	1, 3, 6
T_5	1, 2, 4, 5, 6

	(b)	Define following terms & differentiate them: Data Mart, Enterprise Warehouse & Virtual Warehouse.	**07**
		OR	
Q.3	**(a)**	Discuss the variations of the Apriori algorithm to improve the efficiency.	**07**
	(b)	Explain Classification and Regression for Predictive Analysis.	**07**
Q.4	**(a)**	Why naïve Bayesian classification is called "naïve"? Briefly outline the major ideas of naïve Bayesian classification.	**07**
	(b)	What do you mean by Random Forest Algorithm? Why is Random Forest Algorithm popular?	**07**
		OR	
Q.4	**(a)**	What are Outliers? Describe types of outliers also explain challenges of outliers detection.	**07**
	(b)	List down the features of Bagged Trees. What are the Limitations of Bagging Trees?	**07**
Q.5	**(a)**	Enlist the steps of K-Mean clustering algorithm. Explain it with suitable example.	**07**
	(b)	Enlist the steps of ID3 decision tree generation algorithm. Explain it with suitable example and generate the tree.	**07**
		OR	
Q.5	**(a)**	Explain Clustering Methods in detail.	**07**
	(b)	What is Decision Tree? Explain how classification is done using decision tree induction.	**07**

****