GUJARAT TECHNOLOGICAL UNIVERSITY
MCA INTEGRATED – SEMESTER VI- EXAMINATION –WINTER-2025

Subject Code: 2668603   **Date: 18/11/2025**
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)**	1) Association rules		**07**
		2) Cuboids		
		3) ETL		
		4) Data Mining		
		5) HOLAP		
		6) Frequent itemset		
		7) Confusion matrix		
	(b)	Explain KDD process with its steps.		**07**
Q.2	**(a)**	(i) Agglomerative versus Divisive Hierarchical Clustering		**04**
		(ii) Difference between k-means and k-Medoids.		**03**
	(b)	What is Outlier? Explain types of Outliers.		**07**
		OR		
	(b)	Explain Rule Induction Using a Sequential Covering Algorithm.		**07**
Q.3	**(a)**	What are the techniques used in data reduction? Explain two of them in detail.		**07**
	(b)	Explain Market Basket Analysis in detail.		**07**
		OR		
Q.3	**(a)**	1. Define Data Warehousing. And explain the key Features of Data Warehousing.		**04**
		2. Discuss Supervised learning and unsupervised learning.		**03**
	(b)	Compare between OLTP & OLAP.		**07**
Q.4	**(a)**	Explain Attribute selection measures with example.		**07**
	(b)	List and explain describe major issues in data mining.		**07**
		OR		
Q.4	**(a)**	Explain Clustering Methods in detail		**07**
	(b)	What is decision tree? Describe Decision Tree Induction algorithm.		**07**
Q.5	**(a)**	Discuss Generating association rules from frequent itemsets and also methods to Improve Apriori Efficiency.		**07**

(b) Explain with figure: 3-Tire Data Warehouse Architecture. 07

OR

Q.5 **(a)** Explain Typical OLAP Operations with appropriate figure. 07

(b) State and explain the Bayesian classification. 07

****