Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

ME - SEMESTER– II(New course) • EXAMINATION (Remedial) – WINTER- 2015

Subject Code: 2720209 **Date: 10/12/2015**

Subject Name: Data Mining and Data Warehousing

Time:2:30 pm to 5:00 pm **Total Marks: 70**

Instructions:

1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**

Q.1	**(a)**	**1.** Compare following Design Methods for Multidimensional Database: Star schema and fact-constellation schema.	**04**
		2. Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45,46, 52, 70. (a) What is the mean of the data? What is the median? (b) What is the mode of the data? Comment on the dataôs modality (i.e., bimodal). (c) What is the midrange of the data?	**03**
	(b)	Describe Major issue in Data Mining.	**07**
Q.2	**(a)**	Draw and Explain Data Mining Architecture.	**07**
	(b)	Explain Apriori Algorithm for finding Frequent Item-sets.	**07**
		OR	
	(b)	Explain in detail correlation analysis.	**07**
Q.3	**(a)**	Explain Bayeôs Theorem and Bayesian Classification with suitable example.	**07**
	(b)	What do you mean by text mining? Explain various issues involved in it.	**07**
		OR	
Q.3	**(a)**	What is classification and prediction? Explain issues regarding classification and prediction.	**07**
	(b)	Write a short note on Web Mining.	**07**
Q.4	**(a)**	What is data transformation? Briefly explain various steps of data transformation.	**07**
	(b)	Explain mining in following Databases with example. i. Temporal Databases ii. Sequence Databases iii. Spatial Databases and Spatiotemporal Databases.	**07**
		OR	
Q.4	**(a)**	Describe any two hierarchical methods of cluster analysis in detail.	**07**
	(b)	Describe any two density based methods of cluster analysis in detail.	**07**
Q.5	**(a)**	Differentiate Data Warehouse and Data Mart.	**07**
	(b)	Explain outlier analysis with suitable example and algorithm	**07**
		OR	
Q.5	**(a)**	Define multilevel association rules. Give suitable example and explain various methods to mine them.	**07**
	(b)	Draw and Explain the three ï tier data warehousing architecture.	**07**

****