GUJARAT TECHNOLOGICAL UNIVERSITY
ME - SEMESTER II (NEW) – • EXAMINATION – SUMMER – 2017

Subject Code:2720209   **Date:29/05/2017**
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)**	Explain Data Mining functionalities in detail.	**07**
	(b)	What is data cleaning? Discuss various ways of handling missing values during data cleaning.	**07**
Q.2	**(a)**	Explain following Data Transformation Stretegies in detail : i) Normalization ii) Discretization	**07**
	(b)	Explain following in Detail: i) Information Gain ii) Gini Index	**07**
		OR	
	(b)	Define multilevel association rules. Give suitable example and explain various methods to mine them.	**07**
Q.3	**(a)**	What are classification and prediction? Explain issues regarding classification and prediction.	**07**
	(b)	Explain Classification by Backpropagation.	**07**
		OR	
Q.3	**(a)**	Explain Baye's Theorem and Naïve Bayesian Classification.	**07**
	(b)	Explain prepruning and postpruining in Decision tree Induction. Why is it Require?	**07**
Q.4	**(a)**	Explain requirements of clustering in Data Mining.	**07**
	(b)	Explain following methods for cluster analysis in detail: i) k-Means ii) k-Medoids	**07**
		OR	
Q.4	**(a)**	Explain Grid-based methods for cluster analysis in detail.	**07**
	(b)	Explain following methods for cluster analysis in detail: i) BIRCH ii) Chameleon	**07**
Q.5	**(a)**	Explain outlier analysis with suitable example and algorithm.	**07**
	(b)	Write a short note on Multimedia Mining.	**07**
		OR	
Q.5	**(a)**	Write a short note on Web Mining.	**07**
	(b)	What do you mean by text mining? Explain various issues involved in it.	**07**

****