Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE - SEMESTER–VII EXAMINATION – SUMMER 2016

Subject Code:171601 **Date:16/05/2016**

Subject Name:Data warehousing and 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.**

Q.1	**(a)**	Draw and explain three tier architecture of data warehouse.	**07**
	(b)	What is the need of data preprocessing? Write a note on data cleaning process for missing value and noisy data treatment.	**07**
Q.2	**(a)**	Write a note on following: 1. Association Rule Mining 2. Support and Confidence	**07**
	(b)	Write a note on FP growth algorithm to generate frequent itemset.	**07**
		OR	
	(b)	Discuss various issues of Apriori algorithm. Also list and discuss various alternatives to improve efficiency of Apriori algorithm.	**07**
Q.3	**(a)**	Write the differences between classification and prediction. Also discuss various issues of classification and prediction.	**07**
	(b)	Write a note on Rule based classification.	**07**
		OR	
Q.3	**(a)**	Discuss Information gain. Also discuss steps to measure classification accuracy.	**07**
	(b)	Discuss Linear and Non-Linear Regression methods of prediction.	**07**
Q.4	**(a)**	Write steps of k-Means clustering algorithm with its pros and cons.	**07**
	(b)	Compare hierarchical and partitioned clustering techniques.	**07**
		OR	
Q.4	**(a)**	Define clustering. List and discuss various clustering techniques.	**07**
	(b)	Give differences between k-Means and k-Medoids clustering algorithms.	**07**
Q.5	**(a)**	List and explain various OLAP operations.	**07**
	(b)	Define Data Mining. List and discuss various issues of data mining.	**07**
		OR	
Q.5	**(a)**	Write a note on Fuzzy logic with its applications.	**07**
	(b)	Discuss the following: 1. Mining on time-series and sequence data 2. Types of Web mining	**07**

****