Seat No.: _____   Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY
M. E. - SEMESTER – II • EXAMINATION – WINTER • 2013

Subject code: 1722304   **Date: 31-12-2013**
Subject Name: Data Warehouse and Data Mining
Time: 10.30 am – 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.**

Q.1	**(a)**	Define the following : characterization, discrimination, association and classification, Give example of each, using a real-life database.	**07**
	(b)	In a real-world data, Tuples with missing values for some attributes are a common thing. Describe various methods for handling this problem.	**07**
Q.2	**(a)**	Give the difference between OLAP & OLTP	**07**
	(b)	Explain Data ware house architecture. How to implement Data Warehouse?	**07**
		OR	
	(b)	Explain any two methods of Data Mining for Complex type of data.	**07**
Q.3	**(a)**	Explain:-basic issues regarding Classification & Bayesian classification.	**07**
	(b)	Discuss the various techniques to remove the noise from data.	**07**
Q.3	**(a)**	Explain the advantages of Data integration and data reduction in Data mining. Explain with example.	**07**
	(b)	Explain:- Clustering Partition methods & Density based Methods.	**07**
Q.4	**(a)**	Suppose that the data for analysis include the attribute **age.** The age values for the data tuples are (in increasing order): 12, 13, 14, 16, 17, 17, 19, 28, 28, 28, 32, 34, 35, 42, 42, 42, 47, 50, 52, 52, 56, 57, 57, 57, 59, 63, 63, 66, 68, 68, 69, 71, 72, 72, 75.	**07**
		a) Use smoothing by bin means to smooth the above data, using a bit depth of 5. Illustrate your steps. Comment on the effect of this technique for the given data.	
		b) How might you determine outliers in the data?	
	(b)	What is the use of correlation analysis? Explain in detail	**07**
		OR	
Q.4	**(a)**	Suppose that the data for analysis include the attribute **age.** The age values for the data tuples are (in increasing order): 12, 13, 14, 16, 17, 17, 19, 28, 28, 28, 32, 34, 35, 42, 42, 42, 47, 50, 52, 52, 56, 57, 57, 57, 59, 63, 63, 66, 68, 68, 69, 71, 72, 72, 75.	**07**
		a) Use smoothing by bin means to smooth the above data, using a bit depth of 5. Illustrate your steps. Comment on the effect of this technique for the given data.	
		b) How might you determine outliers in the data?	
	(b)	Explain:- associative classification, Temporal mining	**07**
Q.5	**(a)**	Explain Text and web mining related algorithms.	**07**
	(b)	Explain Classification by Back Propagation.	**07**
		OR	
Q.5	**(a)**	Explain any one algorithm for outlier analysis.	**07**
	(b)	List and describe the five primitives for specifying a data mining task.	**07**

****