Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

M. E. - SEMESTER – II • EXAMINATION – WINTER • 2014

Subject code: 1720108 **Date: 05-12-2014**

Subject Name: Data Mining and Data warehousing

Time: 02:30 pm - 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)**	What kind of process a data mining application can perform? List and explain the functions of a data mining application in brief.	**07**
	(b)	Explain various data repositories on which data mining techniques are applied.	**07**
Q.2	**(a)**	What are the various OLAP operations that can be performed on multidimensional Data Model? Explain the three tier Data Warehouse architecture.	**07**
	(b)	Describe the process of mapping the data warehouse to multiprocessor architecture for a sample database.	**07**
		OR	
	(b)	What is the importance of visualization of discovered patterns? Explain the role of presentation in pattern visualization. Discuss various visualization techniques in KDD.	**07**
Q.3	**(a)**	What are the steps in the process of knowledge discovery? What are the various techniques for Data Cleaning?	**07**
	(b)	What is data transformation? Briefly explain various steps of data transformation.	**07**
		OR	
Q.3	**(a)**	Why is outlier mining important? Briefly describe the different approaches behind statistical-based outlier detection, distance-based outlier detection and deviation-based outlier detection.	**07**
	(b)	What is Decision Tree? Explain how classification is done using decision tree induction.	**07**
Q.4	**(a)**	Why naïve Bayesian classification is called ñnaïveò? Briefly outline the major idea of naïve Bayesian classification.	**07**
	(b)	Describe the ID3 algorithm of the decision tree construction. Why it is unsuitable for data mining applications?	**07**
		OR	
Q.4	**(a)**	Explain single dimensional Boolean association rules from Transactional databases and multidimensional association rules from relational database with suitable examples.	**07**
	(b)	What is market basket analysis? Explain association rule with confidence and support.	**07**
Q.5	**(a)**	What are the challenges for effective resource and knowledge discovery in mining the world wide web?	**07**
	(b)	What do you mean by text mining? Explain various issues involved in it.	**07**
		OR	

Q.5	**(a)**	Differentiate between partitioning and hierarchical methods for clustering. Explain any one hierarchical method in brief with suitable example.	**07**
	(b)	With a case study, explain the role of data mining in financial data analysis.	**07**

□□□□□□□□□□□□□