Seat No.: Enrolment No	
------------------------	--

## **GUJARAT TECHNOLOGICAL UNIVERSITY**

M. E. - SEMESTER - II • EXAMINATION - SUMMER • 2013

	•	code: 1722304 Date: 03-06-2013 Name: Data Warehouse and Data Mining	
		0.30 am – 01.00 pm Total Marks: 70	
Ins	struc	tions:	
		<ol> <li>Attempt all questions.</li> <li>Make suitable assumptions wherever necessary.</li> <li>Figures to the right indicate full marks.</li> </ol>	
Q.1	(a)	What is the difference between Data and knowledge? Explain all the steps involved in the process of knowledge discovery.	07
	(b)	What are the various components of a typical data mining system? Explain them in brief and also draw a figure of typical data mining system.	07
Q.2	(a)	Explain all data warehouse models. What are the advantages and disadvantages of the top-down and bottom up approaches to the data warehouse development.	07
	(b)	Why data warehouse systems use back-end tools and utilities? What are the different functions performed by these tools?  OR	07
	(b)	Explain the definition of data warehouse by using the following keywords: Subject oriented, Integrated, Time variant and Non-Volatile.	07
Q.3	(a)	Why we do not perform on-line analytical processing directly on operational databases instead of spending additional time and resources to construct a separate data warehouse? Justify your answer.	07
	(b)	What are the different schemas for multidimensional databases? Explain any one from them.	07
Q.3	(a) (b)	What do you mean by OLTP and OLAP? Compare OLAP and OLTP systems. What do you mean by DSS? What the different key features of Micro strategy DSS web? Explain.	07 07
Q.4	(a)	List the major clustering methods and discuss Balanced Iterative Reducing and Clustering Using Hierarchies method for clustering.	07
	<b>(b)</b>	Explain Agglomerative and Divisive Hierarchical Clustering OR	07
Q.4	(a)	Justify the efficiency of k-Medoids algorithm on large data set by discussing CLARA method.	07
	<b>(b)</b>	What are the different approaches of text mining on the basis of input taken. Explain	07
Q.5	(a)	Compare the web mining with keyword based web search. What are the different challenges that web poses for effective resource and web technology?	07
	<b>(b)</b>	Describe the ID3 algorithm of the decision tree construction. Why it is unsuitable for Data mining applications?	07
Q.5	(a)	OR Describe typical requirements of clustering in data mining?	07
ų.s	(a) (b)	What is the 'Apriori propertyø? How is it used by the APRIORI algorithm? What are the drawbacks of the Apriori algorithm? Discuss with suitable example.	07