GUJARAT TECHNOLOGICAL UNIVERSITY
BE - SEMESTER–VIII EXAMINATION – WINTER 2015

Subject Code:183103   **Date:12/12/2015**
Subject Name: Business Intelligence & Data Mining
Time: 2:30pm to 5:00pm   **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 the KDD process in details.	**07**
	(b)	Define Data warehouse. Explain terms Data mart and Virtual warehouse.	**07**
Q.2	**(a)**	Explain generalization using attribute oriented induction with a complete example. Show necessary steps.	**07**
	(b)	Explain Data Mining Issues Related to Performance & Diversity of Data types.	**07**

OR

	(b)	Discuss various OLAP operations in the multidimensional data model with example.	**07**
Q.3	**(a)**	Explain prepruning and postpruining in Decision tree Induction. Why it is Require?	**07**
	(b)	All electronics company have sales department. Sales consider 4 dimensions namely time, branch & location. The schema contain a central fact tables sales with two measures dollars_sold and unit_sold. Design star schema, snowflake schema & fact constellation for same.	**07**

OR

Q.3	**(a)**	Explain Bayesian classification and Rule based Classification.	**07**
	(b)	Write Short note on ARCS .	**07**
Q.4	**(a)**	Use the two methods below to *normalize* the following group of data: 200, 300, 400, 600, 1000 (a) min-max normalization by setting *min* = 0 and *max* = 1 (b) z-score normalization	**07**
	(b)	Brief outline the major steps of decision tree classification.	**07**

OR

Q.4	**(a)**	Find all frequent itemsets & generate strong association rules in the following database using Apriori Algorithm. Take minimum support count = 2 and min_conf = 70%	**07**

TID	List of item_ids
T1	I1,I2,I3
T2	I1,I2
T3	I2,I5
T4	I1,I2,I5

	(b)	Explain following classification method: i. K-nearest neighbor classifiers. ii. Case-based reasoning	**07**
Q.5	**(a)**	What is Business Intelligence and How it is related to Data Mining.	**07**

	(b)	Explain three tier Data Warehousing Architecture.	**07**

<div align="center">**OR**</div>

Q.5	**(a)**	Describe the life cycle of data in detail.	**07**
	(b)	Explain FP growth to generate frequent item set using specific example.	**07**

<div align="center">************</div>