GUJARAT TECHNOLOGICAL UNIVERSITY
BE - SEMESTER–VI (NEW) EXAMINATION – SUMMER 2024

Subject Code:3161610   **Date:24-05-2024**
Subject Name:Data Warehousing and Mining
Time:10:30 AM TO 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.**
4. **Simple and non-programmable scientific calculators are allowed.**

			MARKS
Q.1	**(a)**	Explain various features of data warehouse.	**03**
	(b)	Explain why data warehouses are needed for developing business solutions from today's perspective.	**04**
	(c)	Explain three-tier data warehouse architecture.	**07**
Q.2	**(a)**	Give the comparisons of OLAP and OLTP system.	**03**
	(b)	List challenges of data mining regarding data mining methodology and user interaction issues.	**04**
	(c)	What is cuboid? Explain OLAP Operations with examples.	**07**

OR

	(c)	Explain star, snowflake, fact constellation schema for multidimensional database.	**07**
Q.3	**(a)**	Explain various application area of data mining.	**03**
	(b)	How to handle missing values in data set?	**04**
	(c)	With the help of a suitable diagram explain the process of knowledge discovery from database.	**07**

OR

Q.3	**(a)**	Why preprocessing is required?	**03**
	(b)	Explain min-max and Z-score normalization.	**04**
	(c)	Explain Mean, Median, Mode, Variance and Standard Deviation with suitable database example.	**07**
Q.4	**(a)**	Suppose a group of sales price records has been sorted as follow 6,9,12,13,15,25,50,70,72,92,204,232 partition them into three bins by equal frequency. Perform data smoothing by bin mean.	**03**
	(b)	Explain association rules with confidence and support.	**04**

(c) State the Apriori Property. Find frequent item-sets and association rules using Apriori algorithm on the following data set with minimum support count is 2 and minimum confidence=75%. **07**

Sr.No	TID	List of items_IDs
1	T100	I1,I2,I5
2	T200	I2,I4
3	T300	I2,I3
4	T400	I1,I2,I4
5	T500	I1,I3
6	T600	I2,I3
7	T700	I1,I3
8	T800	I1,I2,I3,I5
9	T900	I1,I2,I3

OR

Q.4	**(a)**	Write the difference between classification and predication.	**03**
	(b)	Explain rule based classification.	**04**
	(c)	Explain Naïve Bayesian classifier with example. Why it is called Naïve?	**07**

Q.5	**(a)**	List strength and weakness of neural network as classifier.	**03**
	(b)	Define True Positive, True Negative, False Positive and False Negative.	**04**
	(c)	Explain decision tree algorithm with suitable example.	**07**

OR

Q.5	**(a)**	Write short note on web mining.	**03**
	(b)	Explain text mining.	**04**
	(c)	Explain k-means clustering algorithm.	**07**

****

.