

**GUJARAT TECHNOLOGICAL UNIVERSITY**

**BE- SEMESTER-VI (NEW) EXAMINATION – WINTER 2024**

**Subject Code:3161610**

**Date:05-12-2024**

**Subject Name:Data Warehousing and Mining**

**Time:02:30 PM TO 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.
4. Simple and non-programmable scientific calculators are allowed.

		<b>Marks</b>
<b>Q.1</b>	(a) What is Data Warehousing? Explain its features.	<b>03</b>
	(b) List out issues in data mining and describe any two.	<b>04</b>
	(c) What is Data Mining? Explain KDD process.	<b>07</b>
<b>Q.2</b>	(a) Compare OLAP and OLTP systems.	<b>03</b>
	(b) Draw the three tier architecture of data warehouses.	<b>04</b>
	(c) With the help of a suitable example, illustrate the OLAP operations: 'drill-down', 'roll-up', 'slice' and 'dice'.	<b>07</b>
<b>OR</b>		
	(c) What is an 'Outlier'? How do Outliers impact the results of Mining? List out various methods to detect Outliers and explain any one in brief.	<b>07</b>
<b>Q.3</b>	(a) Why pre-processing is required in data mining process?	<b>03</b>
	(b) List out the various noise removing methods and explain any one in detail.	<b>04</b>
	(c) Explain the steps of the Apriori algorithm for mining frequent item sets with candidate generation.	<b>07</b>
<b>OR</b>		
<b>Q.3</b>	(a) List out various Concept Hierarchy and explain any one.	<b>03</b>
	(b) List out various methods for data discretization and explain any one in detail.	<b>04</b>
	(c) What is data transformation? Explain the different data transformation approaches for transforming data.	<b>07</b>
<b>Q.4</b>	(a) Why tree pruning useful in decision tree induction?	<b>03</b>
	(b) Write a short note on text mining.	<b>04</b>
	(c) Explain Linear and Non-linear Regression.	<b>07</b>
<b>OR</b>		
<b>Q.4</b>	(a) Discuss dimensionality reduction in brief.	<b>03</b>
	(b) What is feature selection and explain methods for feature selection.	<b>04</b>
	(c) Explain Naïve Bayesian classification in detail with example.	<b>07</b>
<b>Q.5</b>	(a) Discuss k-NN classification algorithm.	<b>03</b>
	(b) How k-Mean clustering method differs from k-Medoids clustering Method?	<b>04</b>
	(c) What is the confusion matrix, and how is it used to evaluate a classifier?	<b>07</b>
<b>OR</b>		

- Q.5** (a) List out various issues regarding Classification and prediction and describe any one in brief. **03**
- (b) Define Fact Table and dimension table. **04**
- (c) List out various types of Web Mining and explain any two in detail. **07**

\*\*\*\*\*