GUJARAT TECHNOLOGICAL UNIVERSITY
MCA - SEMESTER– IV• EXAMINATION – WINTER 2020

Subject Code:4649307   **Date:05/01/2021**

Subject Name:Data Mining and Data Visualization

Time:02:00 PM to 04:00 PM   **Total Marks: 56**

Instructions:
1. Attempt any FOUR questions out of EIGHT questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.

Q.1	**(a)**	Explain following terms.
		      1) Data cube
		      2) Outliers
		      3) Ordinal attribute
		      4) Strong rules
		      5) Supervised learning
		      6) InfoGraphics
		      7) Scatter plot
			07

	(b)	List the various Data Pre-processing methods. Discuss various Data Cleaning techniques for Missing value and Noisy data.	**07**
Q.2	**(a)**	What is data mining? Describe the steps involved in data mining process with diagram.	**07**
	(b)	What is ensemble method? List the ensemble methods and explain any one.	**07**
Q.3	**(a)**	Explain the confusion matrix by giving appropriate example.	**07**
	(b)	What is Bayes theorem? Explain the working of Naïve Bayesian Classifier.	**07**
Q.4	**(a)**	Describe Decision Tree Induction algorithm. You can describe it with the help of an example. How are the Rules induced from the Decision Tree?	**07**

	(b)	Discuss following in brief:
		      1) Information Gain
		      2) Gain Ratio
			07

Q.5	**(a)**	What is cluster analysis? What are the requirements of cluster analysis?	**07**
	(b)	Discuss classical partitioning methods of clustering: k-means and k-medoids.	**07**
Q.6	**(a)**	Explain the different charting primitives.	**07**
	(b)	Write a short note on the categorization of clustering methods.	**07**
Q.7	**(a)**	Discuss about SVG and HTML canvas.	**07**
	(b)	What is Google chart? How to adding animations in Google chart?	**07**
Q.8	**(a)**	What is data visualization? Explain the applications of data visualization.	**07**
	(b)	Explain the charts used for advanced visualization.	**07**

****