GUJARAT TECHNOLOGICAL UNIVERSITY
MCA – SEMESTER V - EXAMINATION –SUMMER-2022

Subject Code: 4659306   **Date:07/06/2022**
Subject Name: Big Data Analytics
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.**

Q.1	**(a)**	**(I.) Answer the following.**	**03**
		1. Define BI.	
		2. List the applications of big data analytics.	
		3. What is Text mining?	
		(**II**) Explain types of big data.	**04**
	(b)	Which are the phases of data analytics life cycle? Explain any three phases in detail.	**07**
Q.2	**(a)**	**(I.)** List and explain drivers of big data in brief.	**04**
		(II) Explain Euclidean distance and Manhattan distance.	**03**
	(b)	Discuss User based collaborative filtering with example.	**07**
		OR	
	(b)	Explain Item based collaborative filtering in detail.	**07**
Q.3	**(a)**	Explain Descriptive Vs Predictive Modeling.	**07**
	(b)	What is Apriori Algorithm? Explain in brief how Apriori algorithm is used to find Frequent item sets.	**07**
		OR	
Q.3	**(a)**	Discuss autocorrelation and ARIMA model in brief.	**07**
	(b)	Explain simple exponential smoothing in detail.	**07**
Q.4	**(a)**	Explain Direct Vs Undirected Networks in brief.	**07**
	(b)	Write short note on hierarchical clustering. Also list out the limitations of hierarchical clustering.	**07**
		OR	
Q.4	**(a)**	Discuss emergence of big data ecosystem with reference to new approach to analytics.	**07**
	(b)	What is k-means clustering? Explain in brief.	**07**
Q.5	**(a)**	Explain bag of words approach with suitable example.	**07**
	(b)	Briefly explain advantages and disadvantages of social media analytics.	**07**
		OR	
Q.5	**(a)**	Discuss visualizing and analyzing networks in details.	**07**
	(b)	Explain Data Partitioning and Performance Evaluation in brief.	**07**

****