GUJARAT TECHNOLOGICAL UNIVERSITY
BE - SEMESTER– VII (NEW) EXAMINATION – WINTER 2021

Subject Code:2170715   **Date:01/12/2021**
Subject Name:Data Mining and Business Intelligence
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.**

			MARKS
Q.1	**(a)**	Define supervised, unsupervised and active learning.	03
	(b)	Differentiate between OLTP and OLAP.	04
	(c)	Explain KDD process in detail.	07
Q.2	**(a)**	Differentiate between data warehouse and data mart.	03
	(b)	Explain mean, median and mode with example.	04
	(c)	How to improve efficiency of Apriori algorithm. Discuss in detail.	07

OR

Q.2 (c) Generate strong association rules for given dataset. (Min_supp = 60% and   **07**
Min_conf = 80%)

TID	Item_bought
1	{M,O,N,K,E,Y}
2	{D,O,N,K,E,Y}
3	{M,A,K,E}
4	{M,U,C,K,Y}
5	{C,O,O,K,I,E}

Q.3	**(a)**	Explain ROLAP and MOLAP.	03
	(b)	Discuss multidimensional data models of data warehouse.	04
	(c)	Explain various data cleaning methods.	07

OR

Q.3	**(a)**	Explain *Support* and *Confidence* measures by taking appropriate example.	03
	(b)	What is classification? Discuss phases of classification.	04
	(c)	Explain Naïve Bayesian classifier with example. Why it is called Naïve?	07

Q.4	**(a)**	Give a list of data transformation strategies and explain any one in brief.	03
	(b)	Define True Positive, True Negative, False Positive and False Negative.	04
	(c)	Explain Tree Pruning techniques.	07

OR

Q.4	**(a)**	Differentiate clustering and classification.	03
	(b)	Explain partitioning and hierarchical methods of clustering.	04
	(c)	Explain decision tree algorithm with suitable example.	07

Q.5	**(a)**	Differentiate traditional database and Big data.	03
	(b)	What is business intelligence? Give its advantages.	04
	(c)	What is distributed file system? Explain HDFS architecture in detail.	07

OR

Q.5	**(a)**	What is MapReduce?	03
	(b)	Explain Hadoop, Pig, Hive and NoSql.	04
	(c)	Write a short note on data mining application.	07