

GUJARAT TECHNOLOGICAL UNIVERSITY
MBA – SEMESTER –III-EXAMINATION – SUMMER-2022

Subject Code: 4539251**Date: 15-07-2022****Subject Name: Data Warehousing and Data Mining****Time: 02:30 PM TO 05:30 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** Give definition of following terms: **14**
1. BCNF
 2. DBMS
 3. OLTP
 4. Data Lake
 5. Clustering
 6. Decision Tree
 7. Web Mining
- Q.2** (a) What is normalization? Explain 1NF and 2NF with examples. **07**
- (b) What is SQL? Explain various features of SQL. **07**
- OR**
- (b) What are subqueries? Explain various types of subqueries in detail. **07**
- Q.3** (a) Explain various data integration issues with suitable examples. **07**
- (b) Explain data warehouse process framework. **07**
- OR**
- Q.3** (a) Explain difference between Bayesian classification and rule-based classification. **07**
- (b) What is text mining? Explain role of text mining during the online purchase of products. **07**
- Q.4** (a) What is data mining? Explain various techniques of data mining. **07**
- (b) What is cluster analysis? Explain various applications of cluster analysis. **07**
- OR**
- Q.4** (a) Explain role of web mining in information search. **07**
- (b) Explain decision tree induction with suitable examples. **07**

Q.5 VC Mart is small online grocery marketer. VC markets provide all types of groceries which is available in small to large sizes. Recently VC mart launched a mobile app to find the various items which are available in their store. The success totally depends on the data mining tools and techniques. By making a v&c data mining program, they have clustered into various groups like bakery items, cereals, breakfast foods, soups, canned goods, frozen foods, dairy products and so on. Consumers can identify based on type of food, quantity, types, discounts, offers, then v&c finds the items in its database. They also provide like, dislike, and comments for each of the items and various offers. This again used by another data mining tool for refining their results. They also use various factoring methods for selection of songs.

- (a) Which are the variables data mining used to identify the require items? 07
- (b) Explain pros and cons of data mining methods used by the company. 07

OR

- Q.5**
- (a) Do you think data mining method used to identify the item is appropriate? Give proper justification of your answer. 07
 - (b) Can you suggest better data mining method? 07
