

GUJARAT TECHNOLOGICAL UNIVERSITY**MCA INTEGRATED– SEMESTER VI- EXAMINATION –SUMMER-2022****Subject Code: 2668603****Date: 06/06/2022****Subject Name: Data Mining****Time: 10:30 AM to 01:00 PM****Total Marks: 70****Instructions:**

1. Attempt all questions.
2. Make Suitable assumptions wherever necessary.
3. Figures to the right indicate full mark

- Q-1 (a) Explain Following :** **07**
- (i) Data Mart
(ii) OLTP
(iii) ETL
(iv) KDD
(v) Data discrimination
(vi) Frequent itemsets
(vii) Confusion matrix
- (b) What is Data Mining? Explain Data Mining Process in detail.** **07**
- Q.2 (a) Explain Differences between Operational Database Systems and DataWarehouses.** **07**
- (b) (i) Explain Tree pruning? Explain scalability and decision tree induction.** **04**
ii) Explain frequent item sets and closed item sets **03**
- OR**
- (b) Explain data warehousing architecture.** **07**
- Q.3 (a) Explain Stars, Snowflakes, and Fact Constellations Schemas.** **07**
- (b) (i) Explain methods for attribute subset selection .** **04**
(ii) Explain PCA. **03**
- OR**
- Q.3 (a) Explain Major Tasks of Data Preprocessing.** **07**
- (b) Explain Classification and Regression for Predictive Analysis.** **07**
- Q.4 (a) Explain Apriori Algorithm with example.** **07**
- (b) Explain Information Gain, Gain Ratio and Gini Index.** **07**
- OR**
- Q.4 (a) Explain Market Basket Analysis in detail.** **07**
- (b) (i) Agglomerative versus Divisive Hierarchical Clustering** **04**
(ii) Difference between k-means and k-Medoids. **03**
- Q.5 (a) Explain decision tree algorithm.** **07**
- (b) Explain Clustering Methods in detail.** **07**
- OR**
- Q.5 (a) What are Outliers? Describe types of outliers also explain challenges of outliers detection .** **07**
- (b) What is Cluster Analysis? Explain Requirements for Cluster Analysis.** **07**
