

Seat No.: \_\_\_\_\_

Enrolment No. \_\_\_\_\_

**GUJARAT TECHNOLOGICAL UNIVERSITY**  
**MCA– SEMESTER –IV EXAMINATION –SUMMER-2019**

**Subject Code:2640005**

**Date: 28-05-2019**

**Subject Name: Data Warehousing & Data Mining**

**Time:10.30 am to 1.00 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)** Give definitions of following terms **07**
- 1) Noisy data
  - 2) Fact Table
  - 3) Clustering
  - 4) Outlier
  - 5) Classification
  - 6) Association
  - 7) Sequence Discovery
- (b)** Explain data mart architecture with neat diagram **07**
- Q.2 (a)** What is star schema? Compare it with snowflake schema **07**
- (b)** What is data cleaning? Explain any 2 data cleaning techniques **07**
- OR**
- (b)** What is ETL? Discuss it with neat diagram **07**
- Q.3 (a)** Differentiate classification with clustering **07**
- (b)** Write short note on 4 keywords of data warehousing definition **07**
- OR**
- Q.3 (a)** Write short note on native bayes classification **07**
- (b)** Explain any one partitioning methods of clustering with example **07**
- Q.4 (a)** Write a short note on various operations of OLAP **07**
- (b)** What is classification? Compare it with prediction **07**
- OR**
- Q.4 (a)** Write apriori algorithm with suitable example **07**
- (b)** What is data mart? Compare depended & independent data marts **07**
- Q.5 (a)** Write a short note on Bagging & Boosting **07**
- (b)** Explain the application of data mining with telecommunication industry **07**
- OR**
- Q.5 (a)** Discuss major issues of data mining **07**
- (b)** Explain various primitives of data mining **07**

\*\*\*\*\*