

**GUJARAT TECHNOLOGICAL UNIVERSITY****MCA - SEMESTER– IV EXAMINATION – WINTER 2019****Subject Code: 3640006****Date: 18-11-19****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 marks.

- Q.1** (a) (i) Define the following terms: **03**
- a) Data warehouse
  - b) Data Discretization
  - c) Frequent patterns
- (ii) Explain the differences between Knowledge discovery and data mining. **04**
- (b) What is the relation between data warehousing and data mining? List out different sources of information. **07**

- Q.2** (a) Explain following clustering algorithm in details: **07**
- i) CLARA
  - ii) BIRCH
- (b) What do you mean by Data Processing? Describe different data cleaning approaches. **07**

**OR**

- (b) Explain the different methods of handling the missing values. **07**

- Q.3** (a) (i) Define the terms support and confidence. **03**
- (ii) How is association rule mined from large databases? **04**
- (b) Explain in brief about apriori: candidate generation and test approach with an example. **07**

**OR**

- Q.3** (a) (i) When we can say the association rules are interesting? **03**
- (ii) Explain whether association rule mining is supervised or unsupervised type of learning. **04**
- (b) Consider the transaction dataset: **07**

Tid	Items	Tid	Items
1	{a,b}	6	{a,b,c,d}
2	{b,c,d}	7	{a}
3	{a,c,d,e}	8	{a,b,c}
4	{a,d,e}	9	{a,b,d}
5	{a,b,c}	10	{b,c,e}

Construct the FP tree by showing the trees separately after reading each transaction.

- Q.4** (a) What is rule based classifier? Explain how a rule classifier works. **07**

- (b) What is Bayes Theorem? Show how it is used for classification **07**
- OR**
- Q.4** (a) Write the algorithm for k-nearest neighbor classification. **07**
- (b) Discuss methods for estimating predictive accuracy of classification. **07**
- Q.5** (a) Explain the typical requirements of cluster analysis. **07**
- (b) Write a short note on: Spatial Mining. **07**
- OR**
- Q.5** (a) Explain cardinal, ordinal and ratio – scaled variables. **07**
- (b) Write a short note on: Web Mining. **07**

\*\*\*\*\*