

**GUJARAT TECHNOLOGICAL UNIVERSITY****BE - SEMESTER-VII EXAMINATION – SUMMER 2016****Subject Code:171601****Date:16/05/2016****Subject Name:Data warehousing and Data Mining****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.

- |            |   |           |
|------------|---|-----------|
| <b>Q.1</b> | (a) Draw and explain three tier architecture of data warehouse.   | <b>07</b> |
|            | (b) What is the need of data preprocessing? Write a note on data cleaning process for missing value and noisy data treatment.           | <b>07</b> |
| <b>Q.2</b> | (a) Write a note on following:<br>1. Association Rule Mining<br>2. Support and Confidence   | <b>07</b> |
|            | (b) Write a note on FP growth algorithm to generate frequent itemset.   | <b>07</b> |
|            | <b>OR</b>   |           |
|            | (b) Discuss various issues of Apriori algorithm. Also list and discuss various alternatives to improve efficiency of Apriori algorithm. | <b>07</b> |
| <b>Q.3</b> | (a) Write the differences between classification and prediction. Also discuss various issues of classification and prediction.          | <b>07</b> |
|            | (b) Write a note on Rule based classification.  | <b>07</b> |
|            | <b>OR</b>   |           |
| <b>Q.3</b> | (a) Discuss Information gain. Also discuss steps to measure classification accuracy.  | <b>07</b> |
|            | (b) Discuss Linear and Non-Linear Regression methods of prediction.   | <b>07</b> |
| <b>Q.4</b> | (a) Write steps of k-Means clustering algorithm with its pros and cons.   | <b>07</b> |
|            | (b) Compare hierarchical and partitioned clustering techniques.   | <b>07</b> |
|            | <b>OR</b>   |           |
| <b>Q.4</b> | (a) Define clustering. List and discuss various clustering techniques.  | <b>07</b> |
|            | (b) Give differences between k-Means and k-Medoids clustering algorithms.   | <b>07</b> |
| <b>Q.5</b> | (a) List and explain various OLAP operations.   | <b>07</b> |
|            | (b) Define Data Mining. List and discuss various issues of data mining.   | <b>07</b> |
|            | <b>OR</b>   |           |
| <b>Q.5</b> | (a) Write a note on Fuzzy logic with its applications.  | <b>07</b> |
|            | (b) Discuss the following:<br>1. Mining on time-series and sequence data<br>2. Types of Web mining                                      | <b>07</b> |

\*\*\*\*\*

