

GUJARAT TECHNOLOGICAL UNIVERSITY

MCA - SEMESTER-IV • EXAMINATION – SUMMER • 2014

Subject Code: 2640005

Date: 03-06-2014

Subject Name: Data Warehousing and Data Mining (DWDM)

Time: 10:30 am - 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)** Define and explain the following terms. **07**
- 1) Data Sources
 - 2) Query and Reporting
 - 3) MBR
 - 4) Virtual Warehouse
 - 5) Granularity
 - 6) Temporal Modeling
 - 7) OLTP
- (b)** What is data warehouse? Briefly explain components of data warehouse with figurative model. **07**
- Q.2 (a)** What is multidimensional data analysis? Explain its techniques. **07**
- (b)** Explain Concept Hierarchies with its advantages and example. **07**
- OR**
- (b)** How is class comparison performed? Explain discriminate between different classed. **07**
- Q.3 (a)** Why preprocess the Data? Discusses approaches to data cleaning as a process. **07**
- (b)** Discuss the various ways of handling missing value during data cleaning. **07**
- OR**
- Q.3 (a)** Briefly explain the Binning, Clustering and Regression techniques with suitable example. **07**
- (b)** What is data mining? Describe the steps involved in data mining when viewed as a process of knowledge discovery. **07**
- Q.4 (a)** What is the purpose of Apriori algorithm? Explain with suitable example. **07**
- (b)** What is decision tree? Describe tree pruning methods. **07**
- OR**
- Q.4 (a)** Describe the method of generating frequent item sets without candidate generation. **07**
- (b)** What is the difference between discrimination and classification? Difference between classification and prediction? For each of these pairs of tasks, how are they similar? **07**
- Q.5 (a)** What do you mean by cluster analysis? What are the requirements of cluster analysis? **07**
- (b)** How the data mining will be used in the telecommunication industry. **07**
- OR**
- Q.5 (a)** Briefly explain partition clustering algorithm. **07**
- (b)** Discuss the various methods which evaluate the accuracy of a classifier or a predictor. **07**
