

GUJARAT TECHNOLOGICAL UNIVERSITY**M. E. - SEMESTER – II • EXAMINATION – SUMMER • 2014****Subject code: 1722304****Date: 20-06-2014****Subject Name: Data Warehouse and Data Mining****Time: 02:30 pm - 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.

- Q.1** (a) Explain the three-tier data warehousing architecture with diagram. **07**
 (b) Draw and explain the diagram of KDD process. List and elaborate the major challenges of data mining. **07**

- Q.2** (a) Give differences between data warehouse and OLTP systems with the help of suitable diagram. **07**
 (b) What are the different types of multidimensional schemas? Explain any one schema in detail with proper diagram. **07**

OR

- (b) Discuss architectural strategies and organizational issues of data warehousing. **07**
- Q.3** (a) Describe the working of candidate generating steps of level wise search algorithm with its advantages, disadvantages and complexity. **07**
 (b) For following set of transactions, generate frequent item sets using FP-Growth Tree. Assume the Minimum support (min_sup) is 2. **07**

Transaction ID	List of Item sets
T1	I1, I2, I5
T2	I2, I4
T3	I2, I3
T4	I1, I2, I4
T5	I1, I3
T6	I2, I3
T7	I1, I3
T8	I1, I2, I3, I5
T9	I1, I2, I3

OR

- Q.3** (a) Define tree pruning. Discuss tree pruning methods with suitable example. **07**
 (b) Briefly discuss supervised learning process. Discuss information gain, entropy and gain ratio with proper example. **07**

- Q.4** (a) Define clustering. Categorize and explain various clustering techniques. **07**
 (b) Describe any two OLAP operations with example in the multidimensional data models. **07**

OR

- Q.4** (a) Explain the working of the DBSCAN algorithm with its pros and cons. **07**
 (b) Compare k-means clustering algorithm with k-medoid clustering algorithm. **07**
- Q.5** (a) Describe web mining briefly. List and describe various categories of web mining. **07**
 (b) Describe Time series data with its various tasks. **07**

OR

- Q.5** (a) How text mining is related to web mining? Which are the techniques of text mining? **07**
- (b) Explain the essential features of temporal data and temporal inferences. **07**
