Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

ME - SEMESTER–I(New course)• EXAMINATION – WINTER- 2015

Subject Code: 2712309 **Date: 04/01/2016**

Subject Name: Data Mining

Time:2:30 pm to 5:00 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)** Describe the steps involved in KDD process? **07**

(b) Compare OLAP and OLTP. **07**

Q.2 **(a)** Explain different techniques for data reduction. **07**

(b) Group of 12 sales price records has been sorted as follows. **07**

5, 10, 11, 13, 15, 35, 50, 55, 72, 92, 204, 215.

Partition them into three bins by each of the following methods.

1. Equal-frequency partitioning 2. Equal-width partitioning 3. Clustering.

OR

(b) 1. What are the value ranges of the following normalization methods? **03**

(a) min-max normalization (b) z-score normalization (c) normalization by decimal scaling.

2. Normalize the following group of data using min-max normalization by setting *min* = 0 and *max* = 1 **04**

200, 300, 400, 600, 1000.

Q.3 **(a)** 1. What is the purpose of Apriori Algorithm? Define support and confidence. **04**

2. How to generate association rules from frequent item sets? **03**

(b) What is class comparison? How it is performed? **07**

OR

Q.3 **(a)** Explain: Multidimensional association rule, Joining and Pruning in apriory. **07**

(b) Discuss any two techniques to improve Apriory algorithm. **07**

Q.4 **(a)** Explain data cube and fact table with example. **07**

(b) Explain multilayer feed forward neural network in detail. **07**

OR

Q.4 **(a)** The following table shows the paired data where x is number of years of work experience of a college graduate and y is salary . **07**

x years experience	3	8	9	13	3	6	11	21	1	16
y salary	30	57	64	72	36	43	59	90	20	83

1. Using least squares method, find equation for prediction of salary of college graduate. 2. Predict the salary of a college graduate with 10 years of experience.

(b) Explain preprocessing steps for data for classification and prediction. Based on which criteria classification and prediction methods can be compared? **07**

Q.5 **(a)** Explain OLAP operations with example. **07**

(b) Write a short note on click stream mining. **07**

OR

Q.5 **(a)** What is clustering? Explain with example. **07**

(b) Explain how data mining is applicable for telecommunication industry? **07**

****