

PARUL UNIVERSITY
FACULTY OF ENGINEERING & TECHNOLOGY
M. Tech. Summer 2018 - 19 Examination

Semester: 2
Subject Code: 03205180
Subject Name: Data Mining & Warehousing

Date: 01/05/2019
Time: 02:00pm to 04:30pm
Total Marks: 60

Instructions:

1. All questions are compulsory.
2. Figures to the right indicate full marks.
3. Make suitable assumptions wherever necessary.
4. Start new question on new page.

Q.1 A) Explain KDD process with neat diagram. **(05)**

B) Explain measures for finding association rule interestingness with example. **(05)**

C) Use min-max normalization method to normalize the following group of data by setting min = 0 and max = 1 **(05)**

Data : 200, 300, 400, 600, 1000

Q.2 Answer the following questions. (Attempt any three) (Each five mark) **(15)**

A) Explain Distributive, Holistic and Algebraic measures.

B) Describe various methods for handling missing data values.

C) Define the following terms: i) Business Intelligence,
 ii) Data Mart,
 iii) Closed frequent itemset,
 iv) Correlation analysis,
 v) Outlier Analysis

D) Suppose a group of sales price records has been sorted as follows: 6, 9, 12, 13, 15, 25, 50, 70, 72, 92, 204, 232 Partition them into three bins by equal-frequency (equi-depth) partitioning method. Perform data smoothing by bin mean.

Q.3 A) Using Apriori algorithm, generate frequent item sets (min_sup = 33.3%) for the following transaction database. **(07)**

Trans_id	Itemlist
T1	{A, B, D, K}
T2	{A, B, C, D, E}
T3	{A, B, C, E}
T4	{B, D}
T5	{A, C}
T6	{B, D}

B) Explain Data warehouse architecture. **(08)**

OR

B) Write short note on web mining. **(08)**

Q.4 A) What is cluster Analysis? Explain any one partitioning algorithm for clustering. **(07)**

OR

A) Explain Algorithm to find Frequent itemset without Candidate generation. **(07)**

B) What is purpose of attribute selection measure? Explain different attribute selection measures. **(08)**