Code No: 814BD **R13**

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD
MCA IV Semester Examinations, June/July - 2018
DATA WAREHOUSING AND DATA MINING

Time: 3 Hours **Max. Marks: 60**

Note: This question paper contains two parts A and B.
Part A is compulsory which carries 20 marks. Answer all questions in Part A. Part B consists of 5 Units. Answer any one full question from each unit. Each question carries 8 marks and may have a, b, c as sub questions.

PART - A

5 × 4 Marks = 20

1.a) What are the characteristics of an interesting pattern? [4]
b) What is meant by multi dimensional data model? [4]
c) Give examples for a single dimensional association rule and a quantitative multidimensional association rules. [4]
d) What are the accuracy measures for a classifier? [4]
e) List the merits and demerits of hierarchical agglomerative clustering. [4]

PART - B

5 × 8 Marks = 40

2. What is data mining? Explain it as a step in knowledge discovery process. [8]

OR

3. Demonstrate attribute subset selection as a preprocessing technique. [8]

4. Define data warehouse. Compare it with database management systems. [8]

OR

5. Explain BUC algorithm for data cube computation. [8]

6. Using FP Growth algorithm find frequent item sets(support threshold 30%) for the following data: [8]

TID	List of Items
1	Pen, eraser, marker, calculator, drafter
2	Pencil, marker, eraser, cutter
3	Pen, Pencil, eraser, A4 papers
4	A4 papers, CD, marker
5	Pencil, eraser, stapler, marker
6	Pen, eraser, sharpener, calculator
7	A4 papers, Pencil, eraser
8	Calculator, drafter, Pen
9	Pen, Pencil, CD, A4 papers.

OR

7. What is correlation analysis? Explain the significance of lift measure for correlation analysis. [8]

8. How to prepare data for classification? Explain with suitable data set. [8]

OR

9. What are the characteristics of neural network that make a good classifier? Describe back propagation algorithm. [8]

10. Explain k-means algorithm and contrast it with k-medoid algorithm. [8]

OR

11. What is an outlier? What is the need of outlier detection? Explain any one technique for outlier analysis. [8]

---oo0oo---