2018

MCA

MCA-5.3

DATA MINING AND WAREHOUSING

Full Marks : 75

Time : 3 hour

The figures in the margin indicates full marks for the questions

1. (a) What is data mining? 2
   (b) What is the difference between KDD and datamining? Explain in detail about KDD process. 2+6

2. (a) Describe candidate generation and pruning process of a apriori algorithmwith proper example. 6
   (b) Write down the apriori algorithm. 6

3. Explain the algoritm for mining frequent item sets without candidate generation for the bgiven set(min-sup =30%). 10

TID	Items brought
100	a,c,d,f,g,i,m,p
200	a,b,c,f,l,m,o
300	b,f,h,j,o,w
400	b,c,k,s,p
500	a,f,c,e,l,p,m,n
600	c,d,h,j,o,w

4. Write down the PAM algorithm. What is the advantage of CLARANS over PAM? 6+4=10

P.T.O

5. (a) Define data warehouse. State general stages involved in data warehouse. 2+4=6

(b) What is OLAP? Describ four basic OLAP operations. 2+6=8

6. Give brief description about meta data. 10

Or

What do you mean by data cube in multidimentional data model? Describe the different schemas for multidimensional databases in data warehouse? 2+8=10

7. Describe about requirements of clustering analysis in data mining. 9

***