

6/H-80 (xv) (b) (Syllabus-2015)

(2)

2 0 2 1

(July)

COMPUTER APPLICATION

(Honours)

(Data Warehousing and Data Mining)

[BCA-602 (b)]

Marks : 75

Time : 3 hours

*The figures in the margin indicate full marks
for the questions*

Answer any **two** questions from Unit—I and **one**
question each from the remaining Units

UNIT—I

1. What is meant by metadata? Describe the different types of metadata in a data warehouse. 3+7=10
2. What is a data warehouse? What are the various data sources for the data warehouse? 2+8=10

20D/1344

(Turn Over)

3. What is the STAR schema? What are fact table and dimension table? 3+7=10

UNIT—II

4. Define data mining. Write short notes on any two data mining techniques. 2+4+4=10
5. Describe the different stages of KDD process with the help of a diagram. 5+5=10

UNIT—III

6. (a) Differentiate the following : 2½+2½=5
 - (i) Clustering and classification
 - (ii) Numerical and categorical data
 - (b) Describe the working of DBSCAN algorithm. 10
7. (a) Differentiate between partitional and hierarchical clustering. 5
 - (b) Describe the working of CLARA algorithm. 10

20D/1344

(Continued)

(3)

UNIT—IV

8. Use a priori algorithm to find the frequent itemsets from the following data set. Show the result of every step. Assume minimum support count as 2 : 15

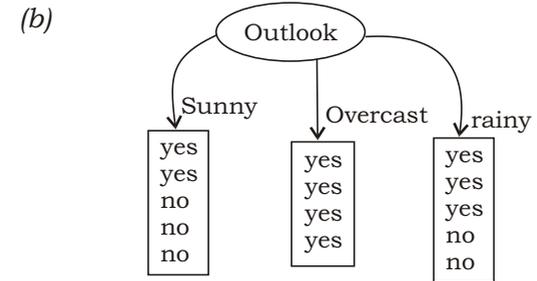
Transaction	Items
1	A, B, C
2	B, C
3	B, D
4	A, B, C
5	A, D
6	B, D
7	A, D
8	A, B, D, C
9	A, B, D

9. (a) Write short notes on the following : 5+5=10
- (i) Generalized association rule
 - (ii) Quantitative association rule
- (b) The Pincer-search algorithm is a bi-directional search. Comment. 5

UNIT—V

10. (a) What is a splitting attribute? What is a splitting criterion? Are the splitting criteria different for numeric attributes and categorical attributes? 1+1+6=8

(4)



Calculate the entropy for the (3) three different outlooks in the above diagram. 7

11. Describe the ID3 algorithm of the decision tree construction. 15
