

6/H-73 (viii) (c) (Syllabus-2015)

2018

(April)

COMPUTER SCIENCE

(Honours)

(Data Mining)

(CS-602 CT)

Marks : 75

Time : 3 hours

The figures in the margin indicate full marks for the questions

Answer **one** question from each Unit

UNIT—I

1. What is data mining? Why is it important? 2+3=5
2. Define five data mining functionalities. 5

UNIT—II

3. Explain five techniques to deal with missing values. Write a short note on 'noisy' data. 10+5=15

(Turn Over)

(2)

(3)

4. Write a short note on dimensionality reduction. Explain, with an example, the star schema of a multidimensional data model.

5+10=15

UNIT—III

5. (a) Define itemset, support and confidence with respect to association rule mining. Write a priori algorithm.

6+6=12

- (b) Explain two disadvantages of a priori algorithm. Give two modifications of a priori algorithm.

4+4=8

6. Write short notes on multilevel association rule, multi-dimensional association rule, constraint-based association rule and metarule-guided association rule.

5×4=20

UNIT—IV

7. (a) What are classification and prediction? Explain, with an example, the decision tree algorithm to find the classification rules.

5+15=20

- (b) Write short notes on linear regression and nonlinear regression.

5+5=10

8D/1896

(Continued)

8. (a) What is cluster analysis? Explain, with an example, the DBSCAN algorithm.

4+12=16

- (b) What is outlier? Explain two methods for outlier detection.

4+10=14

UNIT—V

9. Describe how clustering is used for fraud detection in telecommunication industry.

5

10. Explain one new data mining trend.

5

8D—400/1896

6/H-73 (viii) (c) (Syllabus-2015)