

APRIL 2022

51326/SEE6H/SEZ6H

Time : Three hours

Maximum : 75 marks

PART A — (10 × 2 = 20 marks)

Answer any TEN questions.

1. What is meant by Data Mining?
2. Define : Cluster Analysis.
3. What defines a data mining task?
4. What is data generalization?
5. What is the use of Apriori algorithm?
6. What is decision tree?
7. What is Data cleaning Routines?
8. State Bayes Theorem.
9. Define : Cluster.
10. Expand : DBSCAN.
11. What is sampling?
12. Write the purpose of 'Cross over' Operator in genetic Algorithm.

14 Bsc CS → Data Mining

PART B — (5 × 5 = 25 marks)

Answer any FIVE questions.

13. Discuss briefly on : Classification of Data Mining Systems.
14. Write short notes on : Data Integration.
15. How to perform attribute relevance analysis for concept description?
16. Write about Association Rule Mining.
17. Mention the basic strategy for Decision tree Induction.
18. Write briefly on : Class comparison methods and Implementation.
19. Explain how the k-means algorithm works.
22. Describe the steps in Basic Algorithm for attribute oriented Induction.
23. How to estimate and increase classifier Accuracy? Explain.
24. Describe the statistical approach for model based clustering methods.

PART C — (3 × 10 = 30 marks)

Answer any THREE questions.

20. Describe in detail about Data Mining Functionalities.
21. Explain the concept of Data Reduction.