



QP CODE: 26019128



Reg No :

Name :

M.Sc DEGREE (CSS) EXAMINATION, MARCH 2026

Fourth Semester

M Sc COMPUTER SCIENCE

Core Courses - CA010401 - DATA MINING

2019 Admission Onwards

C9CB6820

Time: 3 Hours

Weightage: 30

Part A (Short Answer Questions)

*Answer any **eight** questions.*

Weight 1 each.

1. Define data mining.
2. What is meant by interactive data mining?
3. What do you mean by quality of data?
4. How data reduction strategies work?
5. What is the basic concept behind the FP growth algorithm?
6. What are rule constraints?
7. What are the different Naïve Bayesian classifiers?
8. What is cross validation?
9. Write the the applications of Cluster Analysis?
10. Define AGNES and DIANA.

(8×1=8 weightage)

Part B (Short Essay/Problems)

*Answer any **six** questions.*

Weight 2 each.

11. Explain the association rule with an example.
12. What is a cuboid? Explain the different types of cuboids.





13. Describe the different methods for data cleaning.
14. How correlation coefficient and covariance can be used in detecting data redundancies.
15. Explain various correlation measures used for mining large data sets.
16. What are the different types of nodes in decision trees? Explain.
17. What are the disadvantages of Information Gain? How these disadvantages can be rectified?
18. What are statistical approaches to detect outliers?

(6×2=12 weightage)

Part C (Essay Type Questions)

*Answer any **two** questions.*

Weight 5 each.

19. Explain various data transformation strategies.
20. Consider an itemset $I=\{a,b,c,d\}$. How many frequent itemsets can be generated? Draw a lattice structure for the given itemset.
21. What is backpropagation? Explain the backpropagation training algorithm of a one hidden layer feed forward network.
22. Explain various Hierarchical clustering methods.

(2×5=10 weightage)

