MAY 2023 **P/ID 16307/PITE2**

Time : Three hours Maximum : 80 marks

PART A — (10 × 2 = 20 marks)

Answer any TEN questions.

1. State the idea behind relational databases.

2. What are all the data mining tasks?

3. List out the importance of data pre-processing.

4. What is mean by noisy data?

5. Define the uses of metarules.

6. State the three common binning strategies.

7. What is the use of multidimensional association rules?

8. What is mean by regression analysis?

9. State the two approaches of tree pruning.

10. List out the types of data in cluster analysis.

11. Define density based cluster.

12. What is mean by Murkowski distance?

PART B — (5 × 6 = 30 marks)

Answer any FIVE questions.

13. State neat algorithm for neural network learning for classification using back propagation algorithm.

14. Explain in brief about strategies of data reduction.

15. What are all the important properties of Apriori property?

16. Explain in detail about Naive Bayesian classification working nature.

17. How data mining techniques are used in social implications?

18. List out the requirements of clustering in data mining.

19. State few examples and explain about multilayer feed-forward network.

PART C — (3 × 10 = 30 marks)

Answer any THREE questions.

20. Elaborate discretization and concept hierarchy generation for numerical data.

21. Write the importance of object oriented databases.

22. Explain in detail about mining guided by rule constraints and its example.

23. List out the types of hierarchical clustering methods and explain with an example.

24. State the uses of Bayes' theorem uses in the naive Bayesian classifier.

———————