MAY 2023 **P/ID 17621/PCAE5**

Time : Three hours Maximum : 80 marks

PART A — (10 × 2 = 20 marks)

Answer any TEN questions.

1. What are the basic units of analysis?

2. What is meant by missing values in data analysis?

3. Define the terms LGD and CLV.

4. What is called hyperplane in Support Vector Machine?

5. What is meant by descriptive analysis?

6. What is the aim of segmentation?

7. What is called Self-Organizing Maps?

8. List out the types of Social Networks.

9. What is meant by binary-link in relational logistic regression?

10. What is the idea of Benchmarking in data analytics?

11. Why privacy is an important concern in data analysis?

12. Why documentation should be transparent and comprehensive?

PART B — (5 × 6 = 30 marks)

Answer any FIVE questions.

13. Write short notes on sampling with suitable example.

14. Describe the processing of standardizing data.

15. Write the differences between logistic and linear regression.

16. Explain the performance measurement for classification models.

17. Write short notes on processing of Association Rule Mining.

18. Explain about the components of a social network learner.

19. Write short notes on the types of software and tools for doing analytics.

PART C — (3 × 10 = 30 marks)

Answer any THREE questions.

20. Explain in details about data sources and data elements.

21. Determine the evaluation of predictive models.

22. Discuss about architecture of Artificial Neural Networks with suitable diagram.

23. Explain in detail about processing steps in K-means clustering with suitable diagram.

24. Discuss about significance of model design and documentation.

———————