APRIL 2021 **51326/SEE6H/SEZ6H**

Time : Three hours Maximum : 75 marks

SECTION A — ($10 \times 2 = 20$ marks)

Answer any TEN questions.

1. What is Noise?

2. Define Data Cleaning.

3. What defines a data mining task?

4. Give the syntax for characterization of the kind of knowledge lobe mined.

5. What is concept description?

6. How to specify the task relevant data?

7. Give the purpose of using the Apriority property.

8. Define interdimension Association Rules.

9. What is Apriori?

10. Name any two applications of classification and prediction.

11. Define Clustering

12. How hierarchical clustering methods are classified?

SECTION B — (5 × 5 = 25 marks)

Answer any FIVE questions

13. How data mining helps in the process of knowledge discovery?

14. How the information from data warehouse help the organizations?

15. Discuss the primitives for specifying a data mining task.

16. How mining can be done by partitioning the data?

17. Explain the techniques of improving the efficiency of Apriori algorithm.

18. Mention the basic strategy for Decision tree induction.

19. What are different types of Data in Cluster Analysis? Discuss about them.

SECTION C — (3 × 10 = 30 marks)

Answer any THREE questions.

20. Explain the concept of Data Reduction.

21. Discuss on concept description.

22. Discuss the methods for mining multilevel association rules involving items at different levels of abstraction.

23. Describe an algorithm for inducing a decision tree from training samples.

24. Describe the statistical approach for model based clustering methods.

———————