A–10186	Sub. Code
	4BCAA4

U.G. DEGREE EXAMINATION, APRIL 2021 &

Supplementary/Improvement/Arrear Examinations

Computer Applications

Allied — DATA MINING AND WAREHOUSING

(CBCS – 2014 onwards)

Time : 3 Hours Maximum : 75 Marks

Part A (10 × 2 = 20)

Answer **all the** questions.

1. What is data mining?

2. Define data warehouse.

3. What is data smoothing?

4. What is histogram? List out the various types of histogram.

5. Mention the methods used for data normalization.

6. Define support and confident for an association rule.

7. What do you mean by clustering?

8. What is Medoid?

9. What is a similarity search?

10. What is web usage mining?

Part B (5 × 5 = 25)

Answer **all** questions, choosing either (a) or (b).

11. (a) Define KDD. Identity and describe the phases in KDD process.

Or

(b) Describe a three-tier data warehousing architecture with neat diagram.

12. (a) What is data reduction? Discuss about the techniques used for data reduction.

Or

(b) Explain the various methods for the generation of concept hierarchies.

13. (a) Compare classification and prediction.

Or

(b) Explain Bayesian classification.

14. (a) List out the requirements for cluster analysis.

Or

(b) Explain the K-Means algorithm with example.

15. (a) Illustrate the trends in data mining.

Or

(b) Briefly describe the theoretical foundations of data mining.

Part C

(3 × 10 = 30)

Answer any **three** questions.

16. Explain the major issues in data mining in detail.

17. Discuss how to handle the missing values in data cleaning as a process.

18. Describe the back propagation algorithm in classification.

19. Explain agglomerative algorithm.

20. Describe the social impacts of data mining.

———————