A–8994

Sub. Code
4BCAA4

U.G. DEGREE EXAMINATION, NOVEMBER 2019

Computer Applications

Allied – DATA MINING AND WAREHOUSING

(CBCS – 2014 onwards)

Time : 3 Hours Maximum : 75 Marks

Part A (10 × 2 = 20)

Answer **all** questions.

1. What is data mining?

2. What is the major difference between the snowflake and start schema models?

3. What is meant by Dimensionality reduction?

4. Define clustering.

5. Define support and confidence.

6. What are Bayesian classifiers?

7. What are outliers?

8. What is the difference between centroid and medoid?

9. Define spatial data mining.

10. What is regression?

Part B $(5 \times 5 = 25)$

Answer **all** questions.

11. (a) How does data mining impact society? Explain.

Or

(b) Describe the architecture of a data warehousing system.

12. (a) Explain the necessity of preprocessing the data.

Or

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

13. (a) Explain the issues regarding classification and prediction.

Or

(b) How does back propagation work? Explain.

14. (a) Explain about the distance measures in algorithmic methods.

Or

(b) What advantages does STING offer over other clustering methods? Explain.

15. (a) Briefly describe the trends in data mining.

Or

(b) Explain the important issues related to web mining.

Part C (3 × 10 = 30)

Answer any **three** questions.

16. Write a note on data mining functionalities.

17. Discuss about the basic methods for data cleaning.

18. Describe a method for generating association rules from transactional databases with examples.

19. How does DBSCAN find clusters? Explain with an example.

20. Discuss the social impacts of data mining.

———————