

25093

M.Sc. DEGREE EXAMINATION, APRIL 2015.

SECOND SEMESTER

Computer Science

Paper III — DATA MINING

Time : Three hours

Maximum : 75 marks

(No additional sheet will be supplied)

PART A — (5 × 3 = 15 marks)

Answer any FIVE questions.

Each question carries 3 marks.

Each answer should not exceed 1 page.

1. Write the difference between database and knowledge base.
2. Specify the advantages of using concept hierarchies in the data mining applications.
3. Define Support and Confidence in Association rule mining.
4. Compare and contrast classification and prediction.
5. Give the three advantages of Web mining.
6. What is spatial database? Mention its features.
7. Differentiate between parallel and distributed algorithms.
8. Define Fuzzy sets, neural networks and Genetic algorithm.

PART B — (4 × 15 = 60 marks)

Answer ALL questions.

Each question carries 15 marks.

Each answer should not exceed 6 pages.

9. Discuss various data mining task primitives in detail.

Or

10. (a) Differentiate operational database systems and data warehousing.
(b) Explain Pattern matching with example.

11. State Bayes theorem and explain how Bayesian classifier works.

Or

12. (a) Explain about basic decision tree induction algorithm.

(b) Explain multilevel association rule mining with example.

13. (a) Explain Statistical-based outlier detection and Deviation-based outlier detection with an example.

(b) With an example explain Clustering with categorical attributes.

Or

14. Discuss the Categorization of Major Clustering Methods.

15. What is Web mining? Explain the techniques in web mining.

Or

16. (a) Differentiate between Generalization and Specialization.

(b) Explain Classification in Spatial data mining.

