E 1253

(Pages : 2)

Reg. No.....

Name.....

B.Sc./B.C.A. DEGREE (C.B.C.S.S.) EXAMINATION, MARCH 2018

Sixth Semester

Choice Based Core Course—DATA MINING

(Common for B.Sc. Computer Science, B.C.A. and B.Sc. Computer Applications [Triple Main])

Time : Three Hours

Maximum Marks : 80

Part A

Answer **all** *questions.*
Each question carries 1 *mark.*

1. Define data cleaning.
2. Which association rules contain single predicate ?
3. Which techniques consider data tuples as objects ?
4. Define data warehouse.
5. Define data mart.
6. Which are used to build heterogeneous database integration ?
7. What are two steps of data classification ?
8. What is mean by Clustering ?
9. What are the two states of binary variable ?
10. Define spatial data warehouse.

(10 × 1 = 10)

Part B

Answer any **eight** *questions.*
Each question carries 2 *marks.*

11. Define outlier analysis.
12. What do you mean by frequent patterns ?
13. Define Transactional Data.
14. Define OLAP.
15. List three tiers of data warehousing architecture.
16. Define virtual data warehouse.

Turn over

17. What is relevance analysis ?
18. List the criteria of association rule mining.
19. What are the categorization of major clustering methods ?
20. What is meant by standardized measurement or z-score ?
21. Define basic measures for text retrieval.
22. What are the choices of spatial measures in spatial data cube construction ?

(8 × 2 = 16)

Part C

Answer any **six** *questions.*
Each question carries 4 *marks.*

23. Compare and contrast data integration and data reduction.
24. Define cluster analysis.
25. Compare and contrast enterprise warehouse and data mart.
26. What are the choices of data materialization ?
27. How can extract classification rules from decision trees ?
28. Explain about classification based on concepts from association rule mining.
29. Explain about CLARANS and discuss its advantages
30. Compare and contrast latent semantic indexing and inverted indexing.
31. Discuss the categories of mining associations in multimedia data.

(6 × 4 = 24)

Part D

Answer any **two** *questions.*
Each question carries 15 *marks.*

32. Explain about data mining and knowledge discovery from data with neat diagrams.
33. Describe about multidimensional data model with neat diagrams.
34. Briefly describe about different partitioning methods.
35. Explain about multidimensional analysis and descriptive mining of complex data objects.

(2 × 15 = 30)