

c. (CBCS) DEGREE EXAMINATION, APRIL 2022

Sixth Semester

Computer Science — Core

DATA WAREHOUSING AND DATA MINING

(For those who joined in July 2017 onwards)

Time : Three hours

Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL questions.

Choose the correct answer :

Data warehouse architecture is based on \_\_\_\_\_

- (a) DBMS (b) RDBMS  
(c) Sybase (d) SQL server

A crucial area of data warehousing is \_\_\_\_\_, which is a kind of data that describes the data warehouse itself.

- (a) Data Mart (b) Metadata  
(c) Data mining (d) Query

Classification rules are extracted from \_\_\_\_\_

- (a) Root node (b) Decision tree  
(c) Siblings (d) Branches

\_\_\_\_\_ are well suited for the mining of multidimensional association rules.

- (a) Data cubes (b) Clustering  
(c) Statistical (d) Nominal

A divisive hierarchical clustering method employs a \_\_\_\_\_ strategy.

- (a) bottom-up (b) top-down  
(c) structural (d) procedural

\_\_\_\_\_ is the process of grouping a set of data objects into multiple groups.

- (a) Clustering (b) Prediction  
(c) Association (d) Correlation

PART B — (5 × 5 = 25 marks)

Answer ALL questions, choosing either (a) or (b).

Each answer should not exceed 250 words.

- (a) Determine the data warehouse administration and management.

Or

- (b) Describe the intangible benefits of data warehouse.

3. The \_\_\_\_\_ is a virtually free resource, which provides a universal connectivity within and between companies.

- (a) Webpage (b) Internet  
(c) Message (d) Network

4. The process of viewing the cross-tab (single dimensional) with a fixed value of one attribute is

- (a) Slicing (b) Dicing  
(c) Pivoting (d) Both (a) and (b)

5. \_\_\_\_\_ is a summarization of the general characteristics of features of a target class of data.

- (a) Data characterization  
(b) Data discrimination  
(c) Constrasting classes  
(d) None of these

6. Which of the following is not involved in data mining?

- (a) Knowledge extraction  
(b) Data archaeology  
(c) Data exploration  
(d) Data transformation

12. (a) Write short on needs of OLAP.

Or

- (b) Explain the business analysis tool categories.

13. (a) Give an account of classification of data mining system.

Or

- (b) Write short notes on data smoothing techniques.

14. (a) Describe on IF-THEN rules for classification.

Or

- (b) Conclude the issues regarding to classification and prediction.

15. (a) Compare supervised vs unsupervised learning.

Or

- (b) Explain the various data mining applications.

PART C — (5 × 8 = 40 marks)

Answer ALL questions, choosing either (a) or (b).

Each answer should not exceed 600 words.

16. (a) Summarize about information Delivery System.

Or

- (b) Write a brief notes on metadata.

17. (a) Explain in detailed about managed query environment.

Or

- (b) Define OLAP. Categorize the various OLAP tools.

18. (a) Write brief notes on major issues in data mining.

Or

- (b) Illustrate about data transformation in data mining.

Page 5 Code No. : 30384 E

19. (a) Criticize the classification by back propagation.

Or

- (b) Elucidate the decision tree induction in classification.

20. (a) Discuss k-means-a centroid-based techniques.

Or

- (b) Enumerate the values of density-based clustering methods.

Page 6 Code No. : 30384 E