(6 Pages)

Reg. No. : ....

Code No. : 30384 E   Sub. Code : SMCS 63

B.Sc. (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 :

1. Data warehouse architecture is based on ———

    (a) DBMS   (b) RDBMS

    (c) Sybase   (d) SQL server

2. 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

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

Page 2   Code No. : 30384 E

7. Classification rules are extracted from ——————

   (a) Root node   (b) Decision tree

   (c) Siblings   (d) Branches

8. —————— are well suited for the mining of multidimensional association rules.

   (a) Data cubes   (b) Clustering

   (c) Statistical   (d) Nominal

9. A divisive hierarchical clustering method employs a —————— strategy.

   (a) bottom-up   (b) top-down

   (c) structural   (d) procedural

10. —————— 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.

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

Or

   (b) Describe the intangible benefits of data warehouse.

Page 3   **Code No. : 30384 E**

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.

Page 4   **Code No. : 30384 E**

[P.T.O]

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.

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.

_____