(6 pages) **Reg. No. :** .....

Code No. : 30610 E **Sub. Code : SMCS 63**

B.Sc. (CBCS) DEGREE EXAMINATION,
APRIL 2020.

Sixth Semester

Computer Science – Main

DATA WAREHOUSING AND DATA MINING

(For those who joined in July 2017 owards)

Time : Three hours Maximum : 75 marks

PART A — (10 × 1 = 10 marks)

Answer ALL questions.

Choose the correct answer:

1. A ———————— could be a set of denormalised, summarised or aggregated data.
   (a) Metadata
   (b) data mart
   (c) EIS
   (d) Data query

2. The functionality of data transformation includes

(a) Removing unwanted data from operational databases

(b) Converting to common data names and definitions

(c) Calculating summaries and derived data

(d) All of the above

3. How many types of Database activity?

(a) 2 (b) 4

(c) 6 (d) 8

4. OLAP data servers can also go in the reverse direction and automatically display detail data which comprises consolidated data This is called

(a) Dicing (b) Consolidation

(c) Slicing (d) drill-downs

5. A ———————— query language can be designed to incorporate these primitives, allowing users to flexibly interact with data mining systems.

(a) Data mining

(b) Data Warehousing

(c) OLAP

(d) OLTP

6. Data mining often requires ———————— the merging of data from multiple data stores.

(a) Data cleaning

(b) Data integration

(c) Data Transformation

(d) Data Reduction

7. ———————— association rules can be mined efficiently using concept hierarchies under a support-confidence framework

(a) Multilevel (b) Multidimensional

(c) Rare patterns (d) Quantitative

8. The units in the hidden layers and output layer are sometimes referred to as ————————.

(a) neurons (b) Neurodes

(c) cell (d) None of these

9. An agglomerative hierarchical clustering method uses ———————— strategy.

(a) Top-down (b) Bottom-up

(c) Structural (d) Procedural

10. A tree structure called a clustering.

(a) Single-linkage

(b) dendrogram

(c) AGNES

(d) BIRCH

PART B — (5 × 5 = 25 marks)

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

Each answer should not exceed 250 words.

11. (a) Write short notes on Data warehouse Database.

Or

(b) Describe the Tangible benefits.

12. (a) Enumerate the OLAP Guidelines.

Or

(b) Write the need for Application OLAP.

13. (a) Give an account of Data Mining Concept.

Or

(b) Denote the Integration of a Data Mining system with a Data Warehouse.

14. (a) Criticize the Apriori Algorithm.

Or

(b) Explain about Tree pruning in decision tree.

15. (a) Write short notes on DBSCAN.

Or

(b) Describe the Grid based clustering method.

PART C — (5 × 8 = 40 marks)

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

Each answer should not exceed 600 words.

16. (a) Explain in detailed about Access Tools in Data Warehouse.

Or

(b) Summarize the design consideration for building Data Warehouse.

17. (a) Determine the importance of Multidimensional Data Model in OLAP.

Or

(b) Write brief notes on MOLAP.

18. (a) Clarify about what kind of data can be mined.

Or

(b) Illustrate the functionalities of Data Cleaning.

19. (a) Explain how to mining various kinds of Association Rules.

Or

(b) Write detail notes on Bayesian classification methods.

20. (a) Determine the concept of Cluster analysis.

Or

(b) Describe about Hierarchical Methods.

———————