Seat No.

Winter Examination Oct/Nov 2023

Subject Name: B.C.A. (CBCS)_88233_88659 - Data Warehousing and Data Mining_26.10.2023_02.30 PM To 05.30 PM

Subject Code: 88233

Day and Date: - Thursday, 26-10-2023
Time: - 02:30 pm to 05:30 pm

Total Marks: 70

Instructions.:
  1) Figures to the right indicate full marks
Special Instruction.:
  1. Que. no. 1 and 6 are compulsory. 2. Attempt any three questions from Que. 2 to 5

Q.1. Q. 1. A) Choose the correct options   **[20]**
                            10

1. Where is data warehousing used?
A. Transaction system
B. Logical system
C. Decision support system
D. None of above

2. ....is defined as the removal of noisy and irrelevant data from the collection.
A. Data extraction
B. Data cleaning
C. Data Modeling
D. None of above

3. ....is defined as the heterogeneous data from multiple sources combined into a common source.
A. Data extraction
B. Data integration
C. Data Modeling
D. None of above

4. KDD stands for ....
A. Knowledge discovery and data
B. Knowledge discovery from data
C. Knowledge did from data
D. None of above

5. On what is data warehouse based?
A. 1D model
B. 2D model
C. 3D model
D. Multidimensional model

6. From where are classification rules extracted? [20]
A. Branches
B. Decision tree
C. Siblings
D. Root node

7. Which of the following also used as the first step in the knowledge discovery process?
A. Data selection
B. Data cleaning
C. Data transformation
D. Data integration

8. Why is the snowflake schema applied?
A. Transformation
B. Aggregation
C. Normalization
D. Generalization

9. Choose incorrect property of data warehouse
A. Volatile
B. Subject-oriented
C. Time variant
D. Collection from heterogeneous sources

10. Which one of the following is not a clustering method?
A. K- nearest neighbor method
B. K-means method
C. Agglomerative method
D. All of the above

Q. 1 B Write short answers (Any two)
10
1. Explain 3 tier data warehouse architecture.
2. Discuss KDD process.
3. Explain Rule based classification

Q.2. What is graph sampling? Explain frequent sub graph mining and tree mining. [10]

Q.3. What is data mining? Discuss different applications of data mining. [10]

Q.4. What is multi-dimensional data model? Explain data cube and star model in detail. [10]

Q.5. Explain in detail partitioning methods like K-mean and K-medoids. [10]

Q.6.   **Write short note on (Any four)**   **[20]**

1. Components of data warehousing
2. Data mining tasks
3. Tree algorithm
4. Accuracy and error measures
5. Types of data in cluster analysis
6. Apriori algorithm