

Seat No. **Summer Examination March - 2023**

Subject Name: B.C.A. (CBCS)_88233_Data Warehousing and Data Mining_26.05.2023_02.30 PM To 05.30 PM

Subject Code: 88233

Day and Date: Friday, 26-05-2023

Time: 02:30 pm to 05:30 pm

Total
Mark
s: 70**Instructions.:**

1) Figures to the right indicate full marks

Special Instruction.:

Que.1 and Que. 6 are compulsory and attempt any three Questions from Que. No.2 to Que. No. 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.refers to the overall process of discovering useful knowledge from data.

- A. KCC
- B. KLL
- C. KDD
- D. None of above

4. On what is data warehouse based?

- A. 1D model
- B. 2D model
- C. 3D model
- D. Multidimensional model

5. From where are classification rules extracted?

- A. Branches
- B. Decision tree
- C. Siblings
- D. Root node

6. Which one of the following can be considered as the final output of the hierarchal type of clustering?

- A. A tree which displays how the close thing are to each other
- B. Assignment of each point to clusters
- C. Finalize estimation of cluster centroids
- D. None of the above

7. Why is the snowflake schema applied?

- A. Transformation
- B. Aggregation
- C. Normalization
- D. Generalization

8. DSS in data warehouse stands for _____.

- A. Decision single system
- B. Decision support system
- C. Data support system
- D. Data storable system

9. Which one of the following statements about the K-means clustering is incorrect?

- A. The goal of the k-means clustering is to partition (n) observation into (k) clusters
- B. K-means clustering can be defined as the method of quantization
- C. The nearest neighbor is the same as the K-means
- D. All of the above

10. Which one of the following can be considered as the correct application of the data mining?

- A. Fraud detection
- B. Corporate Analysis & Risk management
- C. Management and market analysis
- D. All of the above

Q. 1 B Write short answers (Any two)

10

- 1. Discuss KDD process.
- 2. Explain 3 tier data warehouse architecture.
- 3. Explain accuracy and error measures.

- Q.2. What is multi-dimensional data model? Explain snowflakes and star model in detail. [10]
- Q.3. What is data mining? Discuss different applications of data mining. [10]
- Q.4. What is graph sampling? Explain frequent sub graph mining and tree mining. [10]
- Q.5. Discuss different clustering methods in detail. [10]

Q.6. Write short note on (Any four)

[20]

- 1. Components of data warehousing**
- 2. Data mining tasks**
- 3. Tree algorithm**
- 4. Rule based classification**
- 5. Apriori algorithm**
- 6. K means Cluster.**

38