

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BSc DEGREE EXAMINATION MAY 2022
(Fourth Semester)

Branch – COMPUTER SCIENCE WITH DATA ANALYTICS

DATA MINING

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer ALL questions

ALL questions carry EQUAL marks (10 x 1 = 10)

1. The _____ operation defines a subcube by performing a selection on two or more dimensions.
(i) drill (ii) point (iii) slice (iv) dice
2. _____ is a visualization operation which rotates the data axes in order to provide an alternative presentation of the same data.
(i) Out Order (ii) Pivot (iii) In order (iv) Drill
3. Apriori algorithm is also called as _____ algorithm.
(i) Border (ii) Partition (iii) Level Wise (iv) Incremental
4. An item set that was a border set before update and becomes a frequent set after update is called a _____ border item set.
(i) Promoted (ii) Dominated (iii) Under (iv) Upper
5. The goal of _____ is to discover both the dense and the sparse regions in a data set.
(i) Clubbing (ii) Clustering (iii) Sinking (iv) Leveling
6. PAM uses a _____ method to identify the clusters.
(i) k-medoid (ii) k-cluster (iii) k-data (iv) k-analysis
7. ROCK stands for _____
(i) Robust Hierarchical Clustering with Links
(ii) Random Hierarchical Clustering with Links
(iii) Robust Hierarchical Clustering with Kings
(iv) Random Hierarchical Clustering with Kings
8. Decision trees are built from available _____ data.
(i) Slicing (ii) task (iii) data (iv) frequency
9. A sequence is an ordered list of _____.
(i) Frequency (ii) Temporal (iii) Catalogue (iv) Events
10. The algorithm for solving sequence mining problems is mostly based on _____ algorithm.
(i) A priori (ii) Hans (iii) Mining (iv) Decision

SECTION - B (25 Marks)

Answer ALL questions

ALL questions carry EQUAL Marks (5 x 5 = 25)

11. (a) Explain the importance of Data Mining.
(OR)
(b) Discuss on Data Mining Applications.
12. (a) Give an example for Apriori Algorithm.
(OR)
(b) Evaluate the methods to discover Association rule.
13. (a) Analyze on DBSAN Algorithm with an example.
(OR)

Cont...

(b) Give the importance of Clustering algorithms with an example.

14. (a) Expand the concept of Splitting criteria with example.

(OR)

(b) What are the disadvantages of the decision tree algorithm?

15. (a) What is Temporal Data Mining? Explain the types of Temporal Data.

(OR)

(b) Give the importance of Spatial Clustering.

SECTION -C (40 Marks)

Answer **ALL** questions

ALL questions carry **EQUAL** Marks (5 x 8 = 40)

16. (a) Elucidate on the Issues and Challenges in Data Mining.

(OR)

(b) Discuss the Stages of KDD with examples.

17. (a) Explain in detail on FP-Tree Growth Algorithm with an example.

(OR)

(b) Evaluate Incremental algorithm with an example.

18. (a) Discuss on ROCK Clustering Algorithm with an example.

(OR)

(b) Elucidate on CLARANS Algorithm with an example.

19. (a) Comment on the concept of Splitting Indices with examples.

(OR)

(b) Explain the ID3 algorithm for Decision Tree Construction. Give examples.

20. (a) Elucidate in detail on Sequence mining with example.

(OR)

(b) Illustrate with examples the concept of Temporal Association Rule.

Z-Z-Z END