

PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc DEGREE EXAMINATION MAY 2022
(Sixth Semester)

Branch – SOFTWARE SYSTEMS
(Five year Integrated)

DATA WAREHOUSING AND MINING

Time: Three Hours

Maximum: 75 Marks

SECTION-A (10 Marks)

Answer ALL questions

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

1. A Data cube is a _____.
(i) Cuboids (ii) Cube
(iii) Constant (iv) Values
2. _____ is a visualization operation that rotates the data in view to provide an alternative data presentation.
(i) Slice (ii) Dice
(iii) Pivot (iv) Roll-up
3. _____ is the process of finding a model that describes and distinguishes data classes Or concepts.
(i) Data Characterization (ii) Data Classification
(iii) Data discrimination (iv) Data selection
4. Binning is a _____ splitting technique.
(i) Bottom up (ii) Top down
(iii) Top up (iv) Bottom down
5. Rules that satisfy both _____ and _____ are called strong.
(i) maximum support threshold (ii) minimum confidence threshold
(iii) both of these (iv) none of these
6. In the Rule-Based Classification the Learning model is represented as a set of _____.
(i) IF_THEN (ii) IF_ELSE
(iii) IF (iv) none of the above
7. _____ clustering techniques starts with all records in one cluster and then try to spilt that cluster into small pieces.
(i) Agglomerative (ii) Divisive
(iii) Partition (iv) Numeric
8. Clustering algorithm initially assumes that each data instance represents a single cluster.
(i) agglomerative clustering (ii) conceptual clustering
(iii) K-Means clustering (iv) expectation maximization
9. An Interesting mining Optimization method is called _____ can be adopted in spatial association analysis.
(i) Aggregate Refinement (ii) Analysis Management
(iii) Progressive Refinement (iv) None
10. Web page can be segmented by _____ tages.
(i) XML Tags (ii) User Defined Tags
(iii) Predefined Tags (iv) None

SECTION - B (25 Marks)

Answer ALL questions

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

11. a) Explain the types of OLAP servers.
[OR]
b) Illustrate the concept hierarchies.
12. a) Explain the classification of data mining systems.
[OR]
b) Discuss the attribute subset selection.
13. a) Explain the classification and prediction.
[OR]
b) Illustrate multi-level association rules mining.
14. a) Explain k-means algorithm with suitable illustration.
[OR]
b) Discuss k-medoid algorithm with suitable illustration.
15. a) Explain the spatial data cube construction and spatial OLAP.
[OR]
b) Discuss about generalization of structured data.

SECTION -C (40 Marks)

Answer ALL questions

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

Question no 16 is compulsory

16. Categorize the schemas for multidimensional databases.
17. a) Analyse the data mining functionalities.
[OR]
b) Analyse the methods for discretization and concept hierarchy for numeric data.
18. a) Apply the Apriori algorithm for a dataset.
[OR]
b) Elucidate classification by decision tree induction.
19. a) Categorize the hierarchical clustering algorithms.
[OR]
b) Analyse the density-based clustering algorithms.
20. a) Explain about mining the WWW.
[OR]
b) Classify the text mining approaches.

Z-Z-Z END