



QUESTION BANK

Sub.Code : CS1011

Semester : VII

Sub.Title : DATA WAREHOUSING AND MINING

UNIT I

PART - A

1. Define Data warehouse?
2. What are operational databases?
3. Define OLTP?
4. Define OLAP?
5. How a database design is represented in OLAP systems?
6. Write short notes on multidimensional data model?
7. Define data cube?
8. What are facts?
9. What are dimensions?
10. Define dimension table?
11. Define fact table?
12. What are lattice of cuboids?
13. What is apex cuboid?
14. List out the components of star schema?
15. What is snowflake schema?
16. List out the components of fact constellation schema?
17. Point out the major difference between the star schema and the snowflake schema?
18. Which is popular in the data warehouse design, star schema model (or) snowflake schema model?
19. Define concept hierarchy?
20. Define Total Order?
21. Define partial order?
22. Define schema hierarchy?
23. List out the OLAP operations in multidimensional data model?
24. What is roll-up operation?
25. What is drill-down operation?
26. What is slice operation?
27. What is dice operation?
28. What is pivot operation?
29. List out the views in the design of a data warehouse?
30. List out the steps of the data warehouse design process?
31. Define ROLAP?
32. Define MOLAP?
33. Define HOLAP?
34. What is data mart?
35. What are dependent and independent data marts?
36. What is virtual warehouse?
37. Define indexing?

38. What are the types of indexing?
39. Define metadata? 40. Define VLDB?

PART – B

1. Discuss the components of data warehouse. (16)
2. List out the differences between OLTP and OLAP. (16)
3. Discuss the various schematic representations in multidimensional model. (16)
4. Explain the OLAP operations multidimensional model. (16)
5. Explain the design and construction of a data warehouse. (16)
6. Explain the three-tier data warehouse architecture. (16)
7. (a) Explain indexing. (8)
(b) Write short notes on VLDB. (8)

UNIT II

PART - A

1. Define Data cleaning.
2. Define Data mining.
3. What is Data Generalization?
4. Define Data mining Query Language.
5. What are the different types of coupling?
6. Define Data Integration.
7. Define Data Transformation.
8. What is GIS?
9. What is the Graphical User Interface?
10. What is Concept Description?
11. What are the ways to control a generalization process?
12. What are the approaches categorized in flexible generalization of large dataset?
13. What are the methods of class comparison?
14. What is the difference between semi-tight coupling and tight coupling?
15. What are the uses of statistics in data mining?
16. What is the dispersion?
17. What is a q-q plot?

PART B

1. What are the types of data pre-processing techniques? Explain in detail about them? (16)
2. (a) Explain the smoothing Techniques? (8)
(b) Explain Data transformation in detail? (8)
3. Explain data reduction? (16)
4. Explain parametric methods and non-parametric methods of reduction? (16)
5. Explain Data Discretization and Concept Hierarchy Generation? (16)
6. Explain Data mining Primitives? (16)
7. Explain Attribute Oriented Induction? (16)
8. Explain Statistical measures in data bases? (16)

UNIT III

PART – A

1. Define Association Rule Mining.
2. When we can say the association rules are interesting?
3. Explain Association rule in mathematical notations.
4. Define support and confidence in Association rule mining.
5. How is association rules mined from large databases?
6. Describe the different classifications of Association rule mining.
7. What is the purpose of Apriori Algorithm?
8. Define anti-monotone property.
9. How to generate association rules from frequent item sets?
10. Give few techniques to improve the efficiency of Apriori algorithm.
11. What are the things suffering the performance of Apriori candidate generation technique.
12. Describe the method of generating frequent item sets without candidate generation.
13. Define Iceberg query.
14. Mention few approaches to mining Multilevel Association Rules
15. What are multidimensional association rules?
16. Define constraint-Based Association Mining.

PART - B

1. (a) Explain multilevel association rule? (8)
(b) Explain Apriori algorithm. (8)
2. Explain the apriori algorithm for finding frequent item sets without candidate generation? (16)
3. (a) Explain the apriori algorithm for finding frequent item sets? (8)
(b) Explain Apriori algorithm. (8)
4. Explain mining single – dimensional Boolean associated rules from transaction Database? (16)
5. Explain mining Multi – dimensional Boolean associated rules from transaction Database? (16)
6. Explain constraint-based association mining? (16)

UNIT IV

PART – A

1. Define the concept of classification.
2. What is Decision tree?
3. What is Attribute Selection Measure?
4. Describe Tree pruning methods.
5. Define Pre Pruning
6. Define Post Pruning.
7. What is meant by Pattern?
8. Define the concept of prediction.
9. Define Clustering?

10. What do you mean by Cluster Analysis?
11. What are the fields in which clustering techniques are used?
12. What are the requirements of cluster analysis?
13. What is interval scaled variables?
14. Define Binary variables? And what are the two types of binary variables?
15. Define nominal, ordinal and ratio scaled variables?
16. What do u mean by partitioning method?
17. Define CLARA and CLARANS?
18. What is Hierarchical method?
19. Differentiate Agglomerative and Divisive Hierarchical Clustering?
20. Define Density based method?
21. What is a DBSCAN?
22. What are Bayesian Classifiers?
23. Where is decision trees mainly used?
24. How will you solve a classification problem using decision trees?
25. What is decision tree pruning?
26. Explain ID3
27. What is Association rule?
28. What is Association rule?
29. Define support.
30. Define Confidence.
31. How is association rules mined from large databases?
32. What is the classification of association rules based on various criteria?
33. What are the different types of data used for cluster analysis?

PART – B

1. Explain various classification methods? (16)
2. Explain Bayesian classification? (16)
3. Explain Classifier accuracy with examples? (16)
4. Explain Hierarchical method of classifications? (16)
5. Explain Partition Methods of clustering? (16)
6. Explain classification by Decision tree induction? (16)
7. Explain Outlier analysis? (16)
8. Explain the types of data in cluster analysis. (16)
9. (a) Discuss the requirements of clustering in data mining. (8)
(b) Write sort notes on clustering methods (8)

UNIT V

PART - A

1. Define Spatial Databases.
2. What are Time-Series databases?
3. What is Time Series Analysis?
4. What is Smoothing?
5. Name some of the data mining applications
6. Define Multimedia Database

PART B

1. Explain mining complex types of data? (16)
2. (a) Discuss about WWW. (8)
(b) Write short notes on Text Database. (8)
3. Describe mining for Spatial Database? (16)
4. Describe mining for Multimedia Database. (16)
5. Explain data mining application? (16)

