

(FOR THE CANDIDATES ADMITTED

22UDA413

DURING THE ACADEMIC YEAR 20

ONLY)

REG.NO. :

N.G.M.COLLEGE (AUTONOMOUS) : POLLACHI

END-OF-SEMESTER EXAMINATIONS : MAY 2024

BSC CS WITH DA(SF)

MAXIMUM MARKS: 50

SEMESTER- IV

TIME : 3 HOURS

PART - III

22UDA413-DATA MINING

SECTION – A

(10 X 1 = 10 MARKS)

ANSWER THE FOLLOWING QUESTIONS. (K1)

1. What is the primary goal of data mining?
(a) Data storage (b) Data retrieval (c) Knowledge discovery (d) Data visualization
2. What does the "support" of an association rule measure?
(a) The probability of the rule being true (b) The percentage of transactions containing the rule's items (c) The confidence in the rule's accuracy (d) The statistical significance of the rule
3. In hierarchical clustering, what does the dendrogram represent?
(a) Centroids of clusters (b) Connectivity between data points (c) The hierarchy of clusters (d) The density of data points
4. Which type of Decision Tree can handle both categorical and numerical features naturally?
(a) ID3 (b) CART (c) C4.5 (d) Random Forest
5. In spatial data mining, what does "clustering" refer to?
(a) Finding patterns in sequential data (b) Grouping similar spatial entities together (c) Identifying trends over time (d) Predicting future spatial patterns

ANSWER THE FOLLOWING IN ONE (OR) TWO SENTENCES (K2)

6. Define. Knowledge Discovery in Databases
7. What is the purpose of the "lift" metric in association rule evaluation?
8. Which clustering algorithm is based on density rather than distance?
9. In Decision Trees, what is the criterion used to determine the best split at each node?
10. What is the primary focus of Temporal Data Mining?

SECTION – B

(5 X 3 = 15 MARKS)

ANSWER EITHER (a) OR (b) IN EACH OF THE FOLLOWING QUESTIONS. (K3)

11. a) Explain the fundamental concepts of Knowledge Discovery in Databases (KDD) and Data Mining. How are they related, and what distinguishes one from the other?
(OR)
b) Discuss the challenges associated with data quality in data mining. How does preprocessing address issues such as missing values, outliers, and noise are to be rectified?

(CONT....2)

12. a) Discuss the Apriori algorithm for mining association rules.
(OR)
b) Explain the concept of incremental algorithm in the context of association rule mining.
13. a) Compare and contrast hierarchical and partitional clustering approaches. What are the strengths and weaknesses of each?
(OR)
b) Explain the underlying principle behind the CLARA algorithm.
14. a) Discuss tree construction principles.
(OR)
b) Discuss the different splitting criteria used in Decision Trees.
15. a) Discuss the concept of spatial sequence patterns.
(OR)
b) Explain the techniques used for detecting spatial trends in data mining.

SECTION – C**(5 X 5 = 25 MARKS)****ANSWER EITHER (a) OR (b) IN EACH OF THE FOLLOWING QUESTIONS.(K4 (Or) K5)**

16. a) Explain data mining techniques in detail.
(OR)
b) Explore real-world applications of data mining across various industries.
17. a) Explain the concept of Association Rule Mining in data mining. How does it contribute for discovering patterns and relationships within large datasets?
(OR)
b) Explain the fundamental concepts of the FP tree -Growth algorithm in association rule mining.
18. a) Explain density-based clustering algorithms, such as DBSCAN. How do they identify clusters based on density variations in the data?
(OR)
b) Explain the concept of ROCK algorithm, and how does it differ from traditional clustering methods of data mining?
19. a) Compare and contrast popular Decision Tree algorithms such as ID3, C4.5, and CART. What are the key differences in their methodologies and applications?
(OR)
b) Explain the key characteristics of the CHAID algorithm. How does it approach the task of decision tree construction?
20. a) Explain the concept of temporal data mining.
(OR)
b) Explain the concept of spatial clustering.