

DEPARTMENT OF CSE / CSE(AI&ML) CSE(DS)

Course Code & Name: 22CSPC53/22CAPC53/22CDPC53 & Data Mining and Data Analytics

Academic Year: 2025-2026

Year : III B.Tech Semester I

Mid 1 - Assignment Questions

ANSWER ALL THE QUESTIONS

SET-1

S.No	Question	BTL	COs	POs & PSOs
1	Explain the major issues in data mining related to performance, scalability, and data quality.	2	1	1,2,3,12,1
2	Analyze the steps involved in the data preprocessing stage with an example dataset.	4	1	1,2,3,12,1
3	Summarize the working principles of the Apriori algorithm with a suitable transactional dataset.	2	2	1,2,3,12,1
4	Illustrate the FP-Growth algorithm with a sample dataset and explain how it differs from Apriori.	2	2	1,2,3,12,1
5	Explain the process of decision tree induction with a suitable example.	2	3	1,2,3,12,1

SET-2

S.No	Question	BTL	COs	POs & PSOs
1	Identify and analyze the role of task primitives in defining a data mining query.	3	1	1,2,3,12,1
2	Define data summarization and visualization in data mining, and discuss their importance in knowledge presentation.	1	1	1,2,3,12,1
3	Apply Apriori algorithm on a small transactional database and identify the frequent itemsets.	3	2	1,2,3,12,1
4	Explain the method of generating association rules from frequent itemsets using confidence threshold.	2	2	1,2,3,12,1
5	Apply the ID3 algorithm on a small training dataset and draw the decision tree.	3	3	1,2,3,12,1

SET-3

S.No	Question	BTL	COs	POs & PSOs
1	Explain the differences between descriptive and predictive data mining functionalities with examples.	2	1	1,2,3,12,1
2	Describe output representation primitives and how they aid in result interpretation.	2	1	1,2,3,12,1
3	Analyze the impact of reducing candidate generation in Apriori using hash-based techniques.	4	2	1,2,3,12,1
4	Use FP-Growth to mine frequent patterns from a sample dataset and compare the results with Apriori.	3	2	1,2,3,12,1

5	Discuss the architecture of a multilayer feedforward neural network and explain how backpropagation is used in training.	2	3	1,2,3,12,1
---	--	---	---	------------

SET-4

S.No	Question	BTL	COs	POs & PSOs
1	Analyze the impact of missing data and outliers in a dataset and outline how preprocessing can address them.	4	1	1,2,3,12,1
2	How can domain knowledge or user belief systems influence the selection and interpretation of “interesting” patterns?	1	1	1,2,3,12,1
3	Compare and contrast frequent itemset mining using Apriori vs. FP-Growth in terms of time and space complexity.	4	2	1,2,3,12,1
4	Describe the Apriori algorithm and illustrate its level-by-level approach with an example.	2	2	1,2,3,12,1
5	Train a k-Nearest Neighbor classifier on a dataset and evaluate its performance using accuracy and confusion matrix.	3	3	1,2,3,12,1
