JHARKHAND Rai University RANCHI

PRACTICE SET

End Semester Examination Dec 2025

Course: Data Mining and Warehousing

Program: MCA

Semester: III

Course Code: 3CIT202

Course Outcomes

CO1	Describe the fundamentals of data mining systems as well as issues related to access and retrieval of data at scale.
CO2	Explain the various data mining functionalities and data warehousing techniques.
CO3	Apply the various data mining techniques to solve classification, clustering and association rule mining problems.
CO4	Analyze and choose among different approaches of a data mining task.
CO5	Design and evaluate data mining models to be used in solving real life problems, keeping in view social impacts of data mining.

Section-A

Short Answer Type Questions (5 Marks each)

1. Define data mining. Explain how it differs from traditional data processing systems. [UNIT-1] [CO1] [Remember][LOT]
2. What are the major components of a data mining system? Briefly describe each. [UNIT-1] [CO1] [Understand][LOT]
3. List and explain any five major data mining functionalities. [UNIT-1] [CO1] [Understand][LOT]

4. What is metadata? Why is it important for data retrieval in large-scale systems? [UNIT-1] [CO1] [Understand][LOT]
5. Explain the challenges involved in handling large-scale data for mining and retrieval. [UNIT-1] [CO1] [Understand][LOT]
6. Define data mining and explain how it differs from traditional data analysis. [UNIT-2] [CO2] [Remember][LOT]
7. What are the major components of a data warehouse architecture? [UNIT-2] [CO2] [Understand][LOT]
8. List and explain any three data mining functionalities with examples. [UNIT-2] [CO2] [Understand][LOT]
9. What is OLAP? Describe its basic operations such as roll-up, drill-down, and slice. [UNIT-2] [CO2] [Understand][LOT]
10. Explain the term data preprocessing and its importance in data mining. [UNIT-2] [CO2] [Understand][LOT]
11. Define classification in data mining and explain its main objective. [UNIT-3] [CO3] [Understand][LOT]
12. Differentiate between supervised and unsupervised learning with suitable examples. [UNIT-3] [CO3] [Understand][LOT]
13. What is clustering? Describe any two clustering methods used in data mining. [UNIT-3] [CO3] [Understand][LOT]
14. Explain the concept of support and confidence in association rule mining. [UNIT-3] [CO3] [Remember][LOT]
15. Describe how decision tree induction works for classification problems. [UNIT-3] [CO3] [Apply][LOT]
16. Define data mining and explain the difference between predictive and descriptive data mining tasks.[UNIT-4] [CO4] [Understand][LOT]
17. What is the role of data preprocessing in improving data mining results? [UNIT-4] [CO4] [Understand][LOT]
18. Explain the concept of association rule mining with a suitable example. [UNIT-4] [CO4] [Apply][LOT]

19. List and briefly describe the major steps involved in a data mining process. [UNIT-4] [CO4] [Remember][LOT]

20. What are classification and clustering? How do they differ in their data mining approach? [UNIT-4] [CO4] [Understand][LOT]

21. Define data mining models and explain their role in solving real-life business problems. [UNIT-5] [CO5] [Understand][LOT]

22. Differentiate between classification and clustering techniques with suitable examples. [UNIT-5] [CO5] [Remember][LOT]

23. What are the major ethical and social issues associated with data mining applications? [UNIT-5] [CO5] [Understand][LOT]

24. Explain the concept of model evaluation in data mining. Why is it important? [UNIT-5] [CO5] [Apply][LOT]

25. Describe how association rule mining can help in decision-making processes. [UNIT-5] [CO5] [Apply][LOT]

Section- B

Descriptive Type Questions (10 Marks each)

26. Discuss the architecture of a data mining system with a neat diagram. [UNIT-1] [CO1] [Analyse][HOT]

27. Compare and contrast data cleaning, data integration, and data transformation processes in the context of data preprocessing. [UNIT-1] [CO1] [Evaluate][HOT]

28. Illustrate the architecture of a data warehouse and explain the role of ETL (Extract, Transform, Load) in it.[UNIT-2] [CO2] [Analyse][HOT]

29. Compare and contrast data cleaning, data integration, and data transformation in the context of data warehousing.[UNIT-2] [CO2] [Analyse][HOT]

30. Illustrate the steps involved in the Apriori algorithm with an example dataset. [UNIT-3] [CO3] [Analyse][HOT]*)*

31. Discuss the role of data preprocessing in improving classification and clustering accuracy. Provide suitable examples.[UNIT-3] [CO3] [Evaluate][HOT]

32. Analyze the different data mining approaches classification, clustering, and regression and discuss their applications in business intelligence. [UNIT-4] [CO4] [Analyse][HOT]

33. Compare and contrast supervised and unsupervised learning techniques in data mining. Highlight their advantages and limitations. [UNIT-4] [CO4] [Evaluate][HOT]

34. Analyze how data mining can be applied in healthcare for disease prediction and management. Discuss its benefits and limitations. [UNIT-5] [CO5] [Analyse][HOT]

35. Evaluate the impact of data privacy laws and ethical guidelines on the deployment of data mining models. [UNIT-5] [CO5] [Evaluate][HOT]

Section-C

Long Answer Type Questions (20 Marks each)

36. Elaborate on the KDD (Knowledge Discovery in Databases) process. Discuss each step in detail and explain how data mining fits into this process with a real-world example. [UNIT-1] [CO1] [Analyse][HOT]

37. Design and explain a complete data mining process model, including stages from data collection to knowledge presentation, and discuss how each stage supports decision-making.[UNIT-2] [CO2] [Analyse][HOT]

38. Given the following transaction data for a retail store, apply the Apriori algorithm to mine the frequent item sets of market basket analysis.

Transactions:

Transaction ID	Item sets
1	{Bread, Butter, Milk}
2	{Bread, Diaper, Butter, Eggs}
3	{Milk, Diaper, Butter, Cola}
4	{Bread, Milk, Diaper, Butter}
5	{Bread, Milk, Cola}

Perform the following tasks:

(a) Write the steps for solving problem.

(b) Mine the frequent item sets using the Apriori algorithm with a minimum support threshold of 50%.

(c) Generate the association rules from the frequent item sets with a minimum confidence threshold of 80%. [CO3][UNIT-3][Evaluate][HOT]

39. Evaluate different data mining techniques K-Means, and Apriori for a real-world dataset retail. Justify which approach is most effective and why. [UNIT-4] [CO4] [Evaluate][HOT]

40. Critically evaluate how data mining influences social behaviour and decision-making in areas marketing. Provide real-world examples to justify your points.[UNIT-5] [CO5] [Evaluate][HOT]

Summary Sheet

Course Outcomes (CO) Wise

CO	Question Numbers								Marks
CO1	1	2	3	4	5	26	27	36	65
CO2	6	7	8	9	10	28	29	37	65
CO3	11	12	13	14	15	30	31	38	65
CO4	16	17	18	19	20	32	33	39	65
CO5	21	22	23	24	25	34	35	40	65
Total									**325**

Unit Wise

Unit	Question Numbers								Marks
Unit-1	1	2	3	4	5	26	27	36	65
Unit -2	6	7	8	9	10	28	29	37	65
Unit -3	11	12	13	14	15	30	31	38	65
Unit -4	16	17	18	19	20	32	33	39	65
Unit -5	21	22	23	24	25	34	35	40	65
Total									**325**

Blooms Taxonomy Level (BTL) Wise

BTL	Question Numbers								Marks
LOT	1	2	3	4	5	6	7	8	**125**
	9	10	11	12	13	14	15	16	
	17	18	19	20	21	22	23	24	
	25								
HOT	26	27	28	29	30	31	32	33	**200**
	34	35	36	37	38	39	40		
Total									**325**

Prepared by: Chandray Soren

Disclaimer: - This is a Practice Set. The Question in End term examination will differ from the Practice set. This Practice set is meant for practice only.