

Department of CSE

B. Tech. Mid Question Bank (R20 Regulation)

Academic Year: 2024-25

Semester: VII

Subject Name: DATA MINING

Faculty Name: R Sai Krishna, G Shankar

PART-A

Q.No	Questions	Marks	BL	CO	Unit No
1	Define data mining.	2	BL1	CO1	1
2	What is meant by outlier analysis?	2	BL1	CO1	1
3	Define maximal frequent item set.	2	BL1	CO1	1
4	What types of data to be mined?	2	BL1	CO1	1
5	Why patterns are interested?	2	BL1	CO1	1
6	Explain types of integration schemas.	2	BL2	CO1	1
7	How to compute confidence of an association rule? Give example.	2	BL1	CO2	2
8	What is meant by test data?	2	BL1	CO2	2
9	Describe constraint -based association mining.	2	BL4	CO2	2
10	Explain types of association rules in data mining.	2	BL2	CO2	2
11	What is SPM?	2	BL1	CO2	2
12	Determine State classification problem.				
13	Explain the major steps of decision tree classification.	2	BL1	CO3	3
14	What is rule based classification?	2	BL5	CO3	3
15	Define Bayes theorem?	2	BL2	CO3	3
16	What is the significance of information gain?	2	BL1	CO3	3
17	What is lazy learner?	2	BL1	CO3	3
18	Explain basic concepts of classification.	2	BL2	CO3	3
19	What is cluster analysis?	2	BL1	CO4	4
20	Explain types of Data Cluster Analysis.	2	BL2	CO4	4
21	Define Outlier Analysis.	2	BL1	CO4	4
22	What are the requirements of clustering?	2	BL1	CO4	4
23	Predict major tasks of clustering evaluation?	2	BL6	CO4	4
24	Explain applications of clustering.	2	BL2	CO4	4
25	List the applications of web usage	2	BL1	CO5	5

	mining.				
26	Give examples for unstructured text.	2	BL2	CO5	5
27	Explain time series data.	2	BL2	CO5	5
28	Define Mining object.	2	BL1	CO5	5
29	Explain mining sequence patterns in transactional database.	2	BL2	CO5	5
30	What is recommender system?	2	BL1	CO5	5

PART-B

Q.No	Questions	Marks	BL	CO	Unit No
1	Discuss data mining as a step in knowledge discovery process and various challenges associated.	6	BL6	CO1	1
2	Explain Various Data Mining Functionalities with an example.	6	BL2	CO1	1
3	Explain major issues in data mining.	6	BL2	CO1	1
4	Illustrate about Data Mining Task Primitives.	3	BL2	CO1	1
5	What is Data Cleaning? Describe various methods of Data Cleaning.	3	BL1	CO1	1
6	Discuss about the Issues to be considered during Data Integration.	3	BL6	CO1	1
7	What is preprocessing? Explain about Data Transformation techniques.	3	BL1	CO1	1
8	Discuss with few words about binning methods	3	BL6	CO1	1
9	How data mining is differed from data warehousing?	3	BL2	CO1	1

EXPLORE TO INVENT

10	<p>How to find all the frequent item sets using Apriori algorithm for the given data where min-sup = 2.</p> <p>Transactional Data for an <i>AllElectronics</i> Branch</p> <table border="1"> <thead> <tr> <th><i>TID</i></th> <th><i>List of item IDs</i></th> </tr> </thead> <tbody> <tr> <td>T100</td> <td>I1, I2, I5</td> </tr> <tr> <td>T200</td> <td>I2, I4</td> </tr> <tr> <td>T300</td> <td>I2, I3</td> </tr> <tr> <td>T400</td> <td>I1, I2, I4</td> </tr> <tr> <td>T500</td> <td>I1, I3</td> </tr> <tr> <td>T600</td> <td>I2, I3</td> </tr> <tr> <td>T700</td> <td>I1, I3</td> </tr> <tr> <td>T800</td> <td>I1, I2, I3, I5</td> </tr> <tr> <td>T900</td> <td>I1, I2, I3</td> </tr> </tbody> </table>	<i>TID</i>	<i>List of item IDs</i>	T100	I1, I2, I5	T200	I2, I4	T300	I2, I3	T400	I1, I2, I4	T500	I1, I3	T600	I2, I3	T700	I1, I3	T800	I1, I2, I3, I5	T900	I1, I2, I3	6	BL1	CO2	2
<i>TID</i>	<i>List of item IDs</i>																								
T100	I1, I2, I5																								
T200	I2, I4																								
T300	I2, I3																								
T400	I1, I2, I4																								
T500	I1, I3																								
T600	I2, I3																								
T700	I1, I3																								
T800	I1, I2, I3, I5																								
T900	I1, I2, I3																								
11	List out different kinds of Association Rules with an example for each.	6	BL1	CO2	2																				
12	Describe in detail about constraint based association mining.	6	BL2	CO2	2																				
13	Explain about maximal frequent Item set and closed frequent Item set.	3	BL2	CO2	2																				
14	<p>Apply FP –Growth algorithm to the following transactional database to find frequent item sets. List of items</p> <table border="1"> <tbody> <tr> <td>001</td> <td>I1,I3,I5,I7</td> </tr> <tr> <td>002</td> <td>I1,I5,I6,I7</td> </tr> <tr> <td>003</td> <td>I6,I7</td> </tr> <tr> <td>004</td> <td>I2,I3,I6,I7</td> </tr> <tr> <td>005</td> <td>I8,I1,I6</td> </tr> <tr> <td>006</td> <td>I2,I5,I8</td> </tr> </tbody> </table>	001	I1,I3,I5,I7	002	I1,I5,I6,I7	003	I6,I7	004	I2,I3,I6,I7	005	I8,I1,I6	006	I2,I5,I8	3	BL3	CO2	2								
001	I1,I3,I5,I7																								
002	I1,I5,I6,I7																								
003	I6,I7																								
004	I2,I3,I6,I7																								
005	I8,I1,I6																								
006	I2,I5,I8																								
15	How to improve the efficiency of Apriori?	3	BL1	CO2	2																				
16	Explain comparison of pattern evaluation Measures.	3	BL2	CO2	2																				
17	What is graph pattern mining and write its types?	3	BL1	CO2	2																				
18	Explain correlation analysis.	3	BL2	CO2	2																				
19	State classification problem and briefly explain general approaches to solve it.	3	BL3	CO3	3																				

20	Discuss about Decision tree induction algorithm with an example.	3	BL6	CO3	3
21	Discuss about Naïve-Bayes classification algorithm with an example.	3	BL6	CO3	3
22	Justify the selection of k value for KNN classifier.	3	BL5	CO3	3
23	Explain about tree pruning.	3	BL2	CO3	3
24	Explain rule extraction from decision tree with an example.	3	BL2	CO3	3
25	Discuss hierarchical methods for clustering and contrast agglomerative and divisive approaches.	6	BL6	CO4	4
26	Explain statistical based outlier detection.	6	BL2	CO4	4
27	Describe how categorization of major clustering methods is being done	6	BL2	CO4	4
28	Criticize the evaluation metrics used for clusters.	3	BL5	CO4	4
29	List out various clustering methods.	3	BL1	CO4	4
30	How to cluster the data sets using k-means clustering algorithm?	3	BL1	CO4	4
31	How to determine the number of clusters?	3	BL1	CO4	4
32	Write and explain k means clustering algorithm.	3	BL1	CO4	4
33	Explain hierarchical clustering with example.	3	BL2	CO4	4
34	Discuss various kinds of patterns to be mined from web/server logs in web usage mining.	6	BL6	CO5	5
35	Compare and contrast text mining with web content mining using lucid examples.	6	BL4	CO5	5
36	Demonstrate the application of data mining on time series data.	6	BL2	CO5	5
37	What is web content mining? Discuss in detail.	3	BL1	CO5	5
38	Explain briefly about Web mining.	3	BL2	CO5	5

39	How to convert unstructured text into features in text mining?	3	BL1	CO5	5
40	Explain about a. Spatial data mining b. Mining world wide web.	3	BL2	CO5	5
41	Explain the process for mining in time series data.	3	BL2	CO5	5
42	How will you measure precision and recall for text retrieval?	3	BL3	CO5	5

