

Priyadarshini College of Engineering, Nagpur
Sessional Examination (2022-23) Even Semester
B.Tech. Sixth Semester (Computer Technology) (C.B.C.S.)
Data Warehousing and Mining

P. Pages: 2

PCE/KS/23/BTCT602T

Time: Three Hours

Max. Marks: 70

Notes:

1. All questions carry marks as indicated.
2. Solve Question 1 or Question 2.
3. Solve Question 3 or Question 4.
4. Solve Question 5 or Question 6.
5. Solve Question 7 or Question 8.
6. Solve Question 9 or Question 10.
7. Due credit will be given to neatness and adequate dimensions.
8. Assume suitable data wherever necessary.
9. Illustrate your answers whenever necessary with the help of neat sketches.

Q. No.	Question	CO	BT	Marks
1.	a) Explain various OLAP operations with example.	CO1	2	7
	b) Discuss: 1) ROLAP 2) MOLAP 3) HOLAP	CO1	2	7
	OR			
2.	a) With the help of suitable example, explain multidimensional data model for data warehouse.	CO1	2	7
	b) Differentiate between Data Warehouse and Operational DBMS	CO1	2	7
	OR			
3.	a) List and explain the major issues in Data Mining? Explain each in detail.	CO2	2	7
	b) Discuss data mining functionalities in detail.	CO2	2	7
	OR			
4.	a) List the major components of a typical data mining? Draw architecture of data mining system and explain it.	CO2	2	7
	b) Explain the need of Data Preprocessing? Explain Data Cleaning and Data Transformation in short.	CO2	2	7
	OR			
5.	a) Describe k-means algorithm. Also illustrate the strength and weakness of k-means in comparison with the k-medoids algorithm.	CO3	2	7
	b) Discuss Bayesian classification with suitable example.	CO3	2	7
	OR			
6.	a) Define the hierarchical clustering approach? Compare agglomerative and divisive hierarchical clustering.	CO3	2	6
	b) Explain tree induction algorithm for building decision tree.	CO3	3	8

Q. No.	Question	CO	BT	Marks
7.	a) What is Frequent pattern mining and Association Rules? What is the use of both? Explain.	CO4	2	6
	b) Consider Transactional data for an <i>AllElectronics</i> branch. <div style="display: flex; justify-content: space-between;"> <div>TID</div> <div>List of item IDs</div> </div> 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, There are nine transactions in this database, that is, $ D = 9$. Apply the Apriori algorithm for finding frequent itemsets in D . consider $\text{min_support}=2$ and $\text{confidence}=50\%$	CO4	3	8
OR				
8.	a) A database has five transactions. Let $\text{min sup} = 2$. <div style="display: flex; justify-content: space-between;"> <div>TID</div> <div>items bought</div> </div> T1 {A, B, C, D, E,} T2 {B, C, D} T3 {B, C, D, E} T4 {A, B, C, D, E} T5 {B, C, D, E} Find all frequent itemsets using FP-growth Algorithm.	CO4	3	8
	b) Explain how efficiency of Apriori algorithm can be improved?	CO4	2	6
9.	a) What are the challenges involved in web data mining? Explain	CO5	2	4
	b) How to access accuracy of text retrieval in text mining system?	CO5	2	6
	c) What are the applications of WEB data mining?	CO5	2	4
OR				
10	a) Discuss i) Web Content Mining ii) Web Usage Mining iii) Web Structure Mining iv) Visual Web Data Mining	CO5	2	14

*****All The Best*****