



B.Tech V Semester Supplementary Examinations, July 2024

**DATA WAREHOUSING AND DATA MINING
(IT)**

Maximum Marks: 70

Date: 24.07.2024 Duration: 3 hours

- Note:**
1. This question paper contains two parts A and B.
 2. Part A is compulsory which carries 20 marks. Answer all questions in Part A.
 3. Part B consists of 5 Units. Answer any one full question from each unit which carries 10M.
 4. Each question carries 10 marks and may have a, b, c, d as sub questions.

Part-A

All the following questions carry equal marks		(10X2M=20 Marks)	CO	Bloom Tx
1	List the characteristics of a data ware house.		1	L1
2	Define OLAP.		1	L1
3	Define data mart.		2	L1
4	Define Data mining.		2	L1
5	what is data transformation?		3	L1
6	What is data warehouse metadata?		3	L1
7	Define frequent set and border set.		4	L1
8	What is classification?		4	L1
9	Explain the differences between star and snowflake schema.		5	L1
10	What are outliers?		5	L1

Part-B

Answer All the following questions.		(5X10M=50Marks)												
11	Define data warehouse and Differentiate between operational database systems and data warehouses? [10]		1	L2										
OR														
12	Draw the architecture of data warehouse and explain the three tier in detail. [10]		1	L2										
13	Describe the Architecture of a typical data mining system/Major Components. [10]		2	L2										
OR														
14	Consider the Data set D. Given the minimum support ₂ , apply Apriori algorithm on this dataset. [10]		2	L2										
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 50%;">Transaction ID</th> <th style="width: 50%;">Items</th> </tr> </thead> <tbody> <tr><td style="text-align: center;">100</td><td style="text-align: center;">A,C,D</td></tr> <tr><td style="text-align: center;">200</td><td style="text-align: center;">B,C,E</td></tr> <tr><td style="text-align: center;">300</td><td style="text-align: center;">A,B,C,E</td></tr> <tr><td style="text-align: center;">400</td><td style="text-align: center;">B,E</td></tr> </tbody> </table>		Transaction ID	Items	100	A,C,D	200	B,C,E	300	A,B,C,E	400	B,E			
Transaction ID	Items													
100	A,C,D													
200	B,C,E													
300	A,B,C,E													
400	B,E													

15	Briefly explain about the multidimensional data cube and it's operations. [10]	3	L2
	OR		
16	What is classification and prediction? Explain with an example. [10]	3	L2
17	Explain the algorithm for constructing a decision tree from training samples	4	L2
	OR		
18	Explain Outlier analysis with example. [10]	4	L2
19	Explain Hierarchical method clustering of classification with example. [10]	5	L2
	OR		
20	Explain Density-Based Clustering Methods of classification with example. [10]	5	L2