



Regulation R18

Subject code:2P5FC

**TKR COLLEGE OF ENGINEERING AND TECHNOLOGY**

(Autonomous, Accredited by NAAC with 'A+' Grade)

Index in Chapter | International in Excellence

**B.Tech V Semester Supplementary Examinations, February 2024**

**DATA WAREHOUSING AND DATA MINING  
(INFORMATION TECHNOLOGY)**

Maximum Marks: 70

Date:20.02.2024 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

Answer all the following questions (10X2M=20 Marks)		CO	Bloom Tx
1	Define data warehouse.	1	L1
2	List the Data warehouse Characteristics.	1	L1
3	What is data cleaning?	2	L1
4	Define OLAP Data cube.	2	L1
5	Define data mart.	3	L1
6	Define Slice and Dice operation.	3	L1
7	Explain the differences between star and snowflake schema.	4	L2
8	What is decision tree?	4	L1
9	What is an outlier?	5	L1
10	Define cluster Analysis	5	L1

Part-B

Answer All the following questions (5X10M=50Marks)													
11	Define data warehouse. Draw the architecture of data warehouse and explain the three tiers in detail. [10]	1	L2										
OR													
12	Explain the data preprocessing techniques in detail. [10]	1	L2										
13	Consider the Data set D. Given the minimum support <sub>2</sub> , apply apriori algorithm on this dataset and find maximal frequent item set. [10]	2	L3										
<table border="1" style="margin-left: 40px;"> <thead> <tr> <th>Transaction ID</th> <th>Items</th> </tr> </thead> <tbody> <tr> <td>100</td> <td>A,C,D</td> </tr> <tr> <td>200</td> <td>B,C,E</td> </tr> <tr> <td>300</td> <td>A,B,C,E</td> </tr> <tr> <td>400</td> <td>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												
OR													
14	Explain Naïve Bayesian classification in detail with example. [10]	2	L2										

15	Explain about Basic Agglomerative Hierarchical Clustering Algorithm with an example. [10]	3	L2
	OR		
16	What is prediction? Explain the various prediction techniques. Explain about Decision tree Induction classification technique. [10]	3	L2
17	Discuss about model based clustering methods. [10]	4	L2
	OR		
18	Explain about the partitioning methods. [10]	4	L2
19	Explain Outlier analysis with example. [10]	5	L2
	OR		
20	Explain about DBSCAN Algorithm. [10]	5	L2