R18 Regulation *Subject code:2P6EC*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech VI Semester Supplementary Examinations, May 2025

DATA WAREHOUSING AND DATA MINING

(CSE)

Maximum Marks: 70 Date: 25.06.2025 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.
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)		Marks	CO	BTL
1	Define Data Warehousing	2M	1	L1
2	List out the operations of OLAP	2M	1	L1
3	What is Data Mining?	2M	2	L1
4	Define Binaryzation?	2M	2	L1
5	Define APRIORI Principle	2M	3	L1
6	What are Maximal Frequent Item Set	2M	3	L1
7	How prediction is different from classification?	2M	4	L1
8	How effective are Bayesian classifiers?	2M	4	L1
9	Define Clustering.	2M	5	L1
10	Discuss on Agglomerative and divisive clustering techniques	2M	5	L1

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	BTL
11	What is the Significance of OLAP in Data warehouse.? Describe OLAP operations with neat diagram/example.	10M	1	L2
	OR			
12	Explain with suitable diagrams Star, Snow-Flake and Fact constellation Schema.	10M	1	L2
13	Explain different Data Mining tasks for knowledge discovery database (KDD).	10M	2	L2
	OR			
14	Explain the various Data pre-processing techniques. How data reduction helps in data pre-processing.	10M	2	L2
15	A database has five transactions. Let *min_sup* = 60% and *min_conf* = 80%	10M	3	L2

TID	*Items_bought*
T100	{M, O, N, K, E, Y}
T200	{D, O, N, K, E, Y}
T300	{M, A, K, E}
T400	{M, U, C, K, Y}
T500	{C, O, O, K, I, E}

Find all Frequent item sets using Apriori algorithm. List out all the strong association rules.

	OR			
16	a) Explain how can you improve the performance of Apriori algorithm. b) What are the advantages of FP-Growth algorithm?	5M 5M	3	L2

17	Highlight the general approach to solve a classification problem.	10M	4	L2
	OR			
18	a) State and explain k- nearest neighbor algorithm. b) List the characteristics of Nearest Neighbour Classifier.	5M 5M	4	L2
19	Discuss Basic Agglomerative Hierarchical Clustering Algorithm.	10M	5	L2
	OR			
20	What are the different clustering methods? Explain in detail.	10M	5	L2