R18 Regulation *Subject code:2P6EC*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech VI Semester Supplementary Examinations, April 2026

DATA WAREHOUSING AND DATA MINING

(CSE)

Maximum Marks: 70 Date: 23.04.2026 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	List the characteristics of data warehouse.	2M	1	L1
2	Define metadata.	2M	1	L1
3	List the methods for filling missing values.	2M	2	L1
4	Define KDD process.	2M	2	L1
5	State Apriori property.	2M	3	L1
6	Define closed item set and maximal item set.	2M	3	L1
7	What is rule classification?	2M	4	L1
8	Define Decision tree.	2M	4	L1
9	Define clustering.	2M	5	L1
10	What are the requirements for cluster analysis?	2M	5	L1

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	BTL
11	Draw and explain 3-tier architecture of a data warehouse.	10M	1	L2
	OR			
12	List and explain OLAP operations with an example.	10M	1	L2
13	What is data mining? List and explain the motivating challenges of data mining.	10M	2	L2
	OR			
14	With the help of a neat diagram Explain data mining as a step process of knowledge discovery.	10M	2	L2
15	Consider the following transactional data base, generate association rules and identify strong association rules using **APRIORI** algorithm with **support=2 confidence=60%.**	10M	3	L2

Tid	**items**
T1	Hotdogs, bun, ketchup

		T2	Hotdogs, bun			
		T3	Hotdogs, coke, chips			
		T4	coke, chips			
		T5	Chips,ketchup			
		T6	Hotdogs, coke, chips			
	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	Describe the general approaches to solve a classification problem.			10M	4	L2
	OR					
18	Consider the following data, apply KNN algorithm to find class when acid durability=3 and strength=7.			10M	4	L2

name	acid durability	strength	class
Type1	7	7	Bad
Type2	7	4	Bad
Type3	3	4	Good
Type4	1	4	Good

19	Explain the outlier detection methods.			10M	5	L2
	OR					
20	Discuss key issues involved in hierarchical and K-Means clustering.			10M	5	L2