R22 Regulation *Subject code: 4E6FB*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech VI Semester Regular Examinations, May 2025

DATA WAREHOUSING AND DATA MINING

(IT)

Maximum Marks: 60 Date: 16.06.2025 Duration: 3 hours

Note: **1.This question paper contains two parts A and B.**
2. Part A is compulsory which carries 10 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 (10X1M=10 Marks)		Marks	CO	Bloom Tx
1.a)	List out the Data Warehouse Characteristics.	1M	1	BTL1
b)	Define Snow Flake Schema.	1M	1	BTL1
c)	What is Discretization and Binarization?	1M	2	BTL1
d)	Outline about the KDD challenges.	1M	2	BTL2
e)	Interpret the Frequent item set generation.	1M	3	BTL2
f)	Write the APRIOIRI algorithm.	1M	3	BTL2
g)	What is classification Techniques.	1M	4	BTL1
h)	List the KNN algorithm Characteristics.	1M	4	BTL1
i)	Identify the major elements in Agglomerative method.	1M	5	BTL1
j)	List the issues of Hierarchical Clustering Algorithm.	1M	5	BTL1

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	Bloom Tx
2	a) Define schema design and explain about any of their 3 types.	5M	1	BTL1
	b) Describe various OLAP operations performed on Multidimensional Data Model.	5M	1	BTL3
	OR			
3	a) Distinguish between designing a data warehouse and an OLAP cube.	5M	1	BTL2
	b) Explain clearly about the Dimension table characteristics.	5M	1	BTL2
4	a) What is Data Mining? And extend the data mining tasks.	5M	2	BTL1
	b) Examine the Data transformation techniques.	5M	2	BTL4
	OR			
5	a) Analyze clearly about KDD process.	5M	2	BTL4
	b) Build the Dimensionality Reduction techniques.	5M	2	BTL3
6	Assume 5 transactions and explain the two step approach to generate frequent item sets and to mine association rules using Apriori algorithm.	10M	3	BTL3
	OR			

7	a) Solve the Apriori algorithms for frequent item set mining in transactional databases. Apply these algorithms to the following data: TID LIST OF ITEMS 1 Bread, Milk, Sugar, Tea Powder, Cheese, Tomato 2 Onion, Tomato, Chillies, Sugar, Milk 3 Milk, Cake, Biscuits, Cheese, Onion 4 Chillies, Potato, Milk, Cake, Sugar, Bread	5M	3	BTL5
	b) What is compact representation of frequent itemset?	5M	3	BTL1
8	a) Estimate the general approaches to solving a classification problem.	5M	4	BTL5
	b) Label the KNN classification algorithm and list the characteristics.	5M	4	BTL1
	OR			
9	a) Demonstrate clearly about classification techniques.	5M	4	BTL2
	b) Identify the methods for Expressing attribute test conditions.	5M	4	BTL3
10	a) Define clustering and omit the Evolution of Clustering algorithm	5M	5	BTL1
	b) What is outlier detection? Explain distance-based outlier detection.	5M	5	BTL1
	OR			
11	What are the advantages and disadvantages of k-means clustering against model-based clustering? You are given a set of numbers {2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377}. Use the following techniques to find two clusters from this data set. (i) K-Means with initial centroids {1} and {378} (ii) K-Means with initial centroids {21} and {34} Explain the differences between K-means clustering and kernel K-means clustering.	10M	5	BTL3