R22 Regulation *Subject code:* **4E4HA**

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech IV Semester Supplementary Examinations, December 2024

DATA WAREHOUSING & DATA MINING
(CSE(DS))

Maximum Marks: 60 Date: 03.12.2024 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	CO	Bloom Tx
All the following questions carry equal marks	(10x1M=10 Marks)		
1.a)	Differentiate OLAP, ROLAP	CO1	L1
b)	Give a brief note on Fact-Less-Facts.	CO1	L1
c)	Define data mining.	CO2	L1
d)	Enumerate the applications of data mining	CO2	L1
e)	What are the challenges in data mining that motivate the mining tasks?	CO3	L1
f)	Mention different characteristics to construct a decision tree.	CO3	L1
g)	What is classification?	CO4	L1
h)	h) Write the need for tree pruning in decision tree induction.	CO4	L1
i)	What are the applications of clustering?	CO5	L1
j)	Mention the Key issues in hierarchical clustering.	CO5	L1
	Part-B		
Answer All the following questions.	(5X10M=50Marks)		Bloom Tx level
2	a) Draw the Data warehouse Architecture and explain its Components. [5M] b) Explain Star and Snow-Flake Schemas. [5M]	CO1	L2 L2
	OR		
3	a) Write the difference between designing a data warehouse and an OLAP cube. [5M] b) Give a brief note on ROLAP. [5M]	CO1	L2 L2
4	a) List the steps of the Knowledge Discovery in Databases (KDD) and describe each of them. [5M] b) Explain the following with examples. [5M] i) Aggregation ii) Dimensionality reduction	CO2	L3 L2
	OR		
5	a) Describe the Feature Subset Selection. [5M] b) Illustrate the Data Transformation by Normalization. [5M]	CO2	L2 L2

6	a) How are association rules generated from frequent itemsets? Illustrate. [5M] b) Explain the procedure to mining closed Frequent Item Set. [5M]	CO3	L2 L2
	OR		
7	Discuss the FP-Growth algorithm with an example. [10M]	CO3	L3
8	a) Describe the data classification process with a neat diagram. How does the Naïve Bayesian classification works? Explain. [5M] b) Briefly explain the Evaluation of Classifiers. [5M]	CO4	L3 L2
	OR		
9	a) Discuss on Measures for Selecting the Best Split for the various types of attributes. [5M] b) Explain k-NN Algorithm for data classification with an example. [5M]	CO4	L3 L3
10	a) Give a brief note on PAM Algorithm. [5M] b) What is the drawback of k-means algorithm? How can we modify the algorithm to diminish that problem? [5M]	CO5	L3 L2
	OR		
11	a) Differentiate Agglomerative and divisive Hierarchical Clustering. [5M] b) Explain Partitioning Clustering-K-Means Algorithm with an example. [5M]	CO5	L3 L2