R20 Regulation *Subject code: 3P4HA*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech IV Semester Supplementary Examinations, February 2024

Data Warehousing & Data Mining

(CSE (DATA SCIENCE))

Maximum Marks: 70 Date:19.02.2024 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		
All the following questions carry equal marks (10X2M=20 Marks)		CO	Bloom Tx
1	Write the three tiers in Data warehouse Architecture.	CO1	L1
2	What is Data Extraction?	CO1	L1
3	Define Data mining.	CO2	L1
4	What are the benefits of Dimensionality reduction?	CO2	L1
5	Define Association Rule Mining.	CO3	L1
6	Give few techniques to improve the efficiency of APRIOIRI algorithm.	CO3	L2
7	Define Decision Tree induction.	CO4	L1
8	What is K- Nearest neighbor classification-Algorithm?	CO4	L1
9	Illustrate the meaning of cluster analysis.	CO5	L2
10	mention the fields in which clustering techniques are used.	CO5	L1
	Part-B		
Answer All the following questions. (5X10M=50Marks)			
11	Discuss the following A. Star schema. (3M) B. Snow Flake schema. (3M) C. Fact constellation schema. (4M)	CO1	L2
	OR		
12	Explain OLAP and its types. Also explain the advantages of OLAP. (10M)	CO1	L2
13	A. What is KDD? Explain about data mining as a step in the process of knowledge discovery. (8M) B. Why do we preprocess the data? Discuss. (2M)	CO2	L1 L2
	OR		
14	What is data cleaning? What are the different techniques for handling missing values? (10M)	CO2	L2

15	Write and explain the APRIORI algorithm with an example. (10M)	CO3	L2
	OR		
16	Differentiate between Maximal Frequent Item Set and Closed Frequent Item Set. (10M)	CO3	L3
17	Explain about the classification and discuss with an example. (10M)	CO4	L2
	OR		
18	How a Naive Bayes classifier works? Explain with an example. (10M)	CO4	L2
19	Explain the categories of major clustering methods. (10M)	CO5	L2
	OR		
20	Describe the different types of hierarchical methods. (10M)	CO5	L2