R20 Regulation *Subject code:* **3P4HA**

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech IV Semester Regular Examinations, July 2022

DATA WAREHOUSING & DATA MINING

CSE (DATA SCIENCE)

Maximum Marks: 70 Date:22.07.2022 ***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)

1 What are the characteristics of a data warehouse?
2 List out the models of data warehousing.
3 What is data mining explain?
4 Give the process of data cleaning.
5 What is an association rule in data mining?
6 Is APRIORI supervised or unsupervised? Justify
7 Which technique helps solve classification problems?
8 What is KNN algorithm?
9 Define clustering and give example.
10 What is outlier with example in data mining?

Part-B

Answer All the following questions. (10MX 5=50Marks)

11 A. Explain about the components of data warehouse architecture? (5 Marks)
B. What are the characteristics of fact table and dimension table? (5 Marks)

OR

12 A. Explain in detail about data modeling with example. (5 Marks)
B. Explain OLAP architecture with suitable diagram. (5 Marks)

13 A. What is difference between KDD and data mining? (5 Marks)
B. Explain data transformation technique with relevant diagrams. (5 Marks)

OR

14 A. Explain data mining tasks with suitable diagrams. (5 Marks)
B. Briefly explain about Data discretization and binarization. (5 Marks)

15 A. What is the APRIORI property and how is it used in the APRIORI algorithm? (5 Marks)
B. How many phases are there in FP growth algorithm? Explain. (5 Marks)

OR

16 A. How do you calculate support and confidence in APRIORI algorithm? (5 Marks)
B. What are frequent item sets and closed item sets? Explain. (5 Marks)

17 A. How is decision tree constructed? Explain in detail. (5 Marks)
B. What is K nearest neighbor classification technique? Explain. (5 Marks)

OR

18 A. What is classification problem list different classification methods and explain any one in detail? (5 Marks)
B. Explain decision tree induction in data mining with example (5 Marks)

19 A. How do you evaluate a clustering algorithm? Give example. (5 Marks)
B. What are the strengths and weaknesses of hierarchical methods? (5 Marks)

OR

20 A. Explain hierarchical clustering methods in data mining with example. (5 marks)
B. What is PAM algorithm? How does its clustering work? (5 marks)