



Regulation R18

**TKR COLLEGE OF ENGINEERING AND TECHNOLOGY**

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

Subject code: 2P5FC

**B.Tech V Semester Supplementary Examination, January 2023**  
**DATA WAREHOUSING AND DATA MINING**  
(INFORMATION TECHNOLOGY)

Maximum Marks: 70

11/01/23

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

Answer all the following questions

(10x2M=20 Marks)

- 1 What is a fact table?
- 2 What is a data mart?
- 3 Define OLTP.
- 4 What is data mining?
- 5 What is frequent itemset?
- 6 Define Slice and Dice operation.
- 7 What is classification?
- 8 Define information gain.
- 9 What is an outlier?
- 10 List the demerits of k-means algorithm

Part-B

Answer All the following questions.

(5X10M=50Marks)

- 11 Explain data mining as a step process of knowledge discovery. Mention the Functionalities of Data mining. [10]
- OR
- 12 Differentiate Operational database systems and data warehousing. Explain the star schema and fact constellation schemas. [10]
  - 13 a) Explain various data mining functionalities. [5]  
b) Discuss briefly about data cleaning techniques. [5]
- OR
- 14 Briefly explain Frequent Item Set generation using Apriori algorithm? [10]
  - 15 Describe the data classification process with a neat diagram. How does the Naive Bayesian classification work? Explain. [10]
- OR
- 16 What is prediction? Explain the various prediction techniques. Explain about Decision tree Induction classification technique. [10]
  - 17 Explain about the partitioning methods. [10]
- OR
- 18 Differentiate between AGNES and DIANA algorithms. [10]
  - 19 Explain about DBSCAN Algorithm? [10]
- OR
- 20 Explain Outlier analysis with example? [10]

