



R20 Regulation

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

Subject code:3P6EB

B.Tech VI Semester Regular Examinations, June/July 2023

DATA WAREHOUSING AND DATA MINING
(CSE)

Maximum Marks: 70

Date:24.06.2023 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

- 1 Write two differences between OLAP and OLTP . (10x2M=20 Marks)
- 2 Brief on Fact-Less-Facts
- 3 What is meant by KDD?
- 4 List the limitations of data mining.
- 5 When we can say the association rules are interesting?
- 6 Differentiate multiple and multilevel association rule mining
- 7 Define nominal and ordinal variables
- 8 Write the need for tree pruning in decision tree induction?
- 9 Differentiate between clustering and classification
- 10 What are the two data structures in cluster analysis

Part-B

Answer All the following questions.

- 11 A. Explain the fact constellation schema with an example. [5]
B. Give a note on OLAP Operations. [5] (5X10M=50Marks)
- 12 What is a Data Warehouse? Explain three types of schemas that are used for modeling data warehouse with examples. State the applications of Data mining. What is its need in Business? [10]
- 13 A. Briefly describe the four stages of knowledge discovery (KDD)? [5]
B. List and define the measures of Similarity and Dissimilarity. [5]
- 14 Illustrate the various data reduction techniques for data preprocessing [10]
- 15 State and explain Apriori Algorithm with an illustration [10]

OR

- 16 Explain FP-Growth algorithm and generate effective association rules with minimum support count =33% [10]

Transaction	LIST OF ITEMS
001	MILK, DAL, SUGAR, BREAD
002	DAL, SUGAR, WHEAT, JAM
003	MILK, BREAD, CURD, PANEER
004	WHEAT, PANEER, DAL, SUGAR
005	MILK, PANEER, BREAD
006	WHEAT, DAL, PANEER, BREAD

- 17 A. Discuss the algorithm for K- Nearest neighbor classification. [5]
 B. Explain in detail the Naive-Bayes Classifier. [5]

OR

- 18 Construct a decision tree with root node Type from the data in the table below. The first row contains attribute names. Each row after the first represents the values for one data instance. The output attribute is Class. [10]

Scale	Type	Shade	Texture	Class
One	One	Light	Thin	A
Two	One	Light	Thin	A
Two	Two	Light	Thin	B
Two	Two	Dark	Thin	B
Two	One	Dark	Thin	C
One	One	Dark	Thin	C
One	Two	Light	Thin	C

- 19 A. Differentiate Agglomerative and divisive Hierarchical Clustering. [5]
 B. Discuss about classification of outlier detection techniques. [5]

OR

- 20 How Density based clustering algorithms are different from partitioning based cluster algorithms. Compare both. Explain DBSCAN algorithm with suitable example. [10]