R20 Regulation *Subject code:3P6GB*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech VI Semester Supplementary Examinations, May 2025

DATA WAREHOUSING AND DATA MINING

(CSE(AI&ML))

Maximum Marks: 70 Date: 18.06.2025 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.**
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)		Marks	CO	BTL
1	Differentiate between ROLAP and HOLAP.	2M	1	L1
2	What is the use of metadata in data warehouse?	2M	1	L1
3	List the methods of filling missing values.	2M	2	L1
4	State the Apriori property.	2M	2	L1
5	List the applications of data mining	2M	3	L1
6	What is the need of confidence measure in association rule mining?	2M	3	L1
7	Mention types of classifier techniques.	2M	4	L1
8	Define Pre pruning and post pruning.	2M	4	L1
9	Give the different types of data in cluster analysis	2M	5	L1
10	Define clustering.	2M	5	L1

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	BTL
11	Compare and contrast online transaction processing with online analytical processing.	10M	1	L2
	OR			
12	Draw and explain the three tier data warehouse architecture.	10M	1	L2
13	What is data preprocessing and Explain Data preprocessing techniques?	10M	2	L2
	OR			
14	a)Write short notes on Measures Similarity and Dissimilarity? b)Explain dimensionality reduction.	5M 5M	2	L2
15	Explain Apriori algorithm with an example Support threshold=50%, Confidence= 60% TABLE-1 Transaction — List of items T1 — I1,I2,I3 T2 — I2,I3,I4 T3 — I4,I5 T4 — I1,I2,I4 T5 — I1,I2,I3,I5 T6 — I1,I2,I3,I4	10M	3	L3

	OR			
16	Apply FP-Growth algorithm to the following transactional data to find frequentitemsets. List all frequent itemsets with their support count. TID | List of Item IDs 1 | I1,i3,i5,i7 2 | I2,i4,i6,i8 3 | I1,i3,i5,i7 4 | I9,i7,i5,i1 5 | I2,i4,i6,i7 6 | I1,i2,i3,i4 7 | I3,i4,i5,i6 8 | I7,i8,i6,i1 9 | I8,i5,i3,i2 10 | I1,i3,i4,i6	10M	3	L2
17	Describe the data classification process with a neat diagram. How does the NaiveBayesian classification works? Explain.	10M	4	L2
	OR			
18	a) State Bayes theorem How can it be applied for data classification? b) With example explain Bayesian belief network.	5M 5M	4	L2
19	a)What are the advantages of PAM algorithm over k-means algorithm? b)Explain hierarchical clustering.	5M 5M	5	L2
	OR			
20	Appraise the importance of outlier detection and its application. Explain any one approachfor outlier detection.	10M	5	L2