R20 Regulation *Subject code: 3P6EB*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech VI Semester Supplementary Examinations, April 2026

DATA WAREHOUSING AND DATA MINING

(CSE)

Maximum Marks: 70 Date: 16.04.2026 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	List and define the characteristics of data warehouse.	2M	1	L1
2	Define metadata.	2M	1	L1
3	List the methods for filling missing values.	2M	2	L1
4	Define KDD process.	2M	2	L1
5	State Apriori property.	2M	3	L1
6	Define closed item set and maximal item set.	2M	3	L1
7	What is rule classification?	2M	4	L1
8	Define Decision tree.	2M	4	L1
9	Define Clustering.	2M	5	L1
10	What are the requirements for cluster analysis?	2M	5	L1

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	BTL
11	Draw and explain 3-tier architecture of a data warehouse.	10M	1	L2
	OR			
12	List and explain OLAP operations with an example.	10M	1	L2
13	What is data mining? List and explain the motivating challenges of data mining.	10M	2	L2
	OR			
14	With the help of a neat diagram Explain data mining as a step process of knowledge discovery.	10M	2	L2
15	Consider the following transactional data base , generate association rules and identify strong association rules using **APRIORI** algorithm with **support=2 confidence=60%**	10M	3	L2

Tid	**items**
T1	Hotdogs, bun, ketchup

	T2 Hotdogs, bun T3 Hotdogs, coke, chips T4 coke, chips T5 Chips,ketchup T6 Hotdogs, coke, chips			
	OR			
16	Discuss compact representation of frequent itemset (closed and maximal) with min-sup=3 **Tid** **items** T1 {A,B,C,D} T2 {A,B,C,D} T3 {A,B,C} T4 {B,C,D} T5 {C,D}	5M 5M	3	L2
17	Explain Naïve-Bayers classifiers and predict instance to the class (weather data set) if outlook=rainy,temp=cool, humidity=high, windy=true, play =?	10M	4	L2

No.	1: outlook	2: temperature	3: humidity	4: windy	5: play
	Nominal	Nominal	Nominal	Nominal	Nominal
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mild	high	FALSE	yes
5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool	normal	TRUE	yes
8	sunny	mild	high	FALSE	no
9	sunny	cool	normal	FALSE	yes
10	rainy	mild	normal	FALSE	yes
11	sunny	mild	normal	TRUE	yes
12	overcast	mild	high	TRUE	yes
13	overcast	hot	normal	FALSE	yes
14	rainy	mild	high	TRUE	no

	OR			
18	Explain how evaluation of classifiers will be done using confusion matrix.	10M	4	L2
19	Explain the outlier detection methods.	10M	5	L2
	OR			
20	Discuss key issues involved in hierarchical and K-Means clustering.	10M	5	L2