R22 Regulation *Subject code:4E6FB*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech VI Semester Regular/Supplementary Examinations, April 2026

DATA WAREHOUSING AND DATA MINING

(IT)

Maximum Marks: 60 Date:13.04.2026 Duration: 3 hours

Note: 1.This question paper contains two parts A and B.
2. Part A is compulsory which carries 10 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 (10X1M=10 Marks)		Marks	CO	Bloom Tx
1.a)	What are the characteristics Data warehouse?	1M	1	1
b)	Define ROLAP?	1M	1	1
c)	What is Data Mining?	1M	2	1
d)	Define discretization.	1M	2	1
e)	What is Apriori principle?	1M	3	1
f)	Define Support.	1M	3	1
g)	Discuss evaluation of classifiers.	1M	4	1
h)	What is Bayesian Belief Network?	1M	4	2
i)	What is clustering?	1M	5	1
j)	Define dendrogram.	1M	5	1

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	Bloom Tx
2	a) Explain any two of the schemas for multidimensional databases. b) Describe Fully Addictive, Semi-Addictive, Non Addictive Measures.	5M 5M	1	2
	OR			
3	Illustrate the difference between Snow – Flake Schema and Fact Constellation with a real time example.	10M	1	3
4	a) Discuss steps in KDD process with a diagram. b) Explain methods to handle missing data.	5M 5M	2	2
	OR			
5	a) Write about the methods used for dimensionality reduction. b) Explain a method that mines the complete set of frequent item sets without candidate generation for the table below. Explain the procedure in detail with minimum support = 3.	5M 5M	2	2

TID	Items
100	F, A, C, D, G, I, M, P
200	A, B, C, F, L, M, O
300	B, F, H, J, O, W
400	B, C, K, S, P
500	A, F, C, E, L, P, M, N

<table>
<tr><td>6</td><td>Write the the Apriori algorithm for discovering frequent item sets for mining Boolean association rules.</td><td>10M</td><td>3</td><td>3</td></tr>
<tr><td></td><td colspan="4">OR</td></tr>
<tr><td>7</td><td>a) Write the FP-growth algorithm.
b) Discuss the difference between Maximal and Closed frequent itemsets.</td><td>5M
5M</td><td>3</td><td>3</td></tr>
<tr><td>8</td><td>a) Discuss about general approaches to solving a classification problem.
b) Explain Decision Tree Induction algorithm.</td><td>5M
5M</td><td>4</td><td>2</td></tr>
<tr><td></td><td colspan="4">OR</td></tr>
<tr><td>9</td><td>a) Discuss Bayes theorem with an example.
b) Demonstrate KNN Classification algorithm with an example.</td><td>5M
5M</td><td>4</td><td>2</td></tr>
<tr><td>10</td><td>State K-means algorithm. Apply k-means algorithm with two iterations to form two clusters by taking the initial cluster centers as subjects 1 and 4.
<table>
<tr><th>Subject</th><th>A</th><th>B</th></tr>
<tr><td>1</td><td>1.0</td><td>1.0</td></tr>
<tr><td>2</td><td>1.5</td><td>2.0</td></tr>
<tr><td>3</td><td>3.0</td><td>4.0</td></tr>
<tr><td>4</td><td>5.0</td><td>7.0</td></tr>
<tr><td>5</td><td>3.5</td><td>5.0</td></tr>
<tr><td>6</td><td>4.5</td><td>5.0</td></tr>
<tr><td>7</td><td>3.5</td><td>4.5</td></tr>
</table></td><td>10M</td><td>5</td><td>4</td></tr>
<tr><td></td><td colspan="4">OR</td></tr>
<tr><td>11</td><td>a) Explain how a dendrogram is used in Hierarchical clustering.
b) Explain basic Agglomerative Hierarchical clustering algorithm.</td><td>5M
5M</td><td>5</td><td>3</td></tr>
</table>