



R20 Regulation

Subject code: 3P4HA

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech IV Semester Supplementary Examinations, December 2025

DATA WAREHOUSING AND DATA MINING (CSE(DS))

Maximum Marks: 70

Date: 18.12.2025

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	Define Data Extraction.	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	Mention the importance of association rule mining.	2M	3	L1
7	How prediction is different from classification.	2M	4	L1
8	Define Decision tree.	2M	4	L1
9	What is the importance of outlier analysis?	2M	5	L1
10	Write about K-Means algorithm.	2M	5	L1

Part-B

Answer All the following questions. (5X10M=50Marks)		Marks	CO	BTL
11	Explain fully additive, semi additive and non-additive measures with an example.	10M	1	L2
OR				
12	a) Explain different types of OLAP servers. b) Draw and explain star, snowflake and fact constellation schema for Hospital management system.	5M 5M	1	L2
13	Given the following measurements for the variable age 18,22,25,42,28,43,33,35,56,28 Standardize the variable by using Z-score.	10M	2	L2
OR				
14	a) Discuss the data smoothing methods with an example. b) Explain how to handle missing values with an example.	5M 5M	2	L2
15	a) Explain the procedure to mining closed data item set b) Explain fp-growth algorithm. Find frequent itemsets using fp growth algorithm and construct fp tree for the below dataset.	5M 5M	3	L2

	Tid	Items brought																																						
	T1	K,A,D,B																																						
	T2	D,A,C,E,B																																						
	T3	A,C,B,E																																						
	T4	B,D,E																																						
	T5	A,D,B																																						
OR																																								
16	Write and explain Apriori algorithm with an example.		10M	3	L2																																			
17	What is Bayesian belief network? Explain with an example.		10M	4	L2																																			
OR																																								
18	Write KNN classification algorithm with an example.		10M	4	L2																																			
19	a) Write about evaluation of clustering algorithms. b) Write the key issues in hierarchical clustering algorithm.		5M 5M	5	L2																																			
OR																																								
20	Explain agglomerative methods and construct a dendrogram for below distance matrix using single linkage.		10M	5	L2																																			
		<table border="1"> <thead> <tr> <th></th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> </tr> </thead> <tbody> <tr> <th>1</th> <td>0</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <th>2</th> <td>9</td> <td>0</td> <td></td> <td></td> <td></td> </tr> <tr> <th>3</th> <td>3</td> <td>7</td> <td>0</td> <td></td> <td></td> </tr> <tr> <th>4</th> <td>6</td> <td>5</td> <td>9</td> <td>0</td> <td></td> </tr> <tr> <th>5</th> <td>11</td> <td>10</td> <td>2</td> <td>8</td> <td>0</td> </tr> </tbody> </table>		1	2	3	4	5	1	0					2	9	0				3	3	7	0			4	6	5	9	0		5	11	10	2	8	0		
	1	2	3	4	5																																			
1	0																																							
2	9	0																																						
3	3	7	0																																					
4	6	5	9	0																																				
5	11	10	2	8	0																																			