R20 Regulation *Subject code: 3P6EB*

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech VI Semester Supplementary Examinations, November 2025

DATA WAREHOUSING AND DATA MINING

(CSE)

Maximum Marks: 70 Date: 21.11.2025 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	What are the characteristics of star schema.	2M	1	L1
2	List any four differences between OLTP and OLAP.	2M	1	L1
3	What are the different forms of data preprocessing?	2M	2	L1
4	Define the process of KDD.	2M	2	L1
5	Mention the importance of association rule mining.	2M	3	L1
6	State the property of Apriori.	2M	3	L1
7	How prediction is different from classification?	2M	4	L1
8	Write a short note on confusion matrix?	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	Perform Decimal scaling normalization on sal attribute in below dataset	10M	2	L2

ID	GENDER	AGE	SAL
1	F	27	19000
2	M	51	64
3	M	52	10000
4	F	33	55000
5	M	45	45000

	OR			
14	a) Briefly 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.	5M	3	L2

<table>
<tr><td></td><td>b) Explain fp-growth algorithm. Find frequent itemsets using fp growth algorithm and construct fp tree for the below dataset.
<table>
<tr><td>Tid</td><td>Items brought</td></tr>
<tr><td>T1</td><td>K,A,D,B</td></tr>
<tr><td>T2</td><td>D,A,C,E,B</td></tr>
<tr><td>T3</td><td>A,C,B,E</td></tr>
<tr><td>T4</td><td>B,D,E</td></tr>
<tr><td>T5</td><td>A,D,B</td></tr>
</table></td><td>5M</td><td></td><td></td></tr>
<tr><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td>16</td><td>a) Explain Apriori algorithm for mining association rules.
b) Generate frequent itemset in below transactional dataset using Apriori algorithm (min-sup=40% min-conf=70%)
<table>
<tr><td>Tid</td><td>Items brought</td></tr>
<tr><td>T1</td><td>Bread, butter, milk</td></tr>
<tr><td>T2</td><td>Bread, butter</td></tr>
<tr><td>T3</td><td>Biscuits, cookies,diapers</td></tr>
<tr><td>T4</td><td>Bread, butter, milk, diapers</td></tr>
<tr><td>T5</td><td>Biscuits,diapers</td></tr>
</table></td><td>5M
5M</td><td>3</td><td>L2</td></tr>
<tr><td>17</td><td>What is Bayesian belief network and explain with an example.</td><td>10M</td><td>4</td><td>L2</td></tr>
<tr><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td>18</td><td>Write KNN classification algorithm with an example.</td><td>10M</td><td>4</td><td>L2</td></tr>
<tr><td>19</td><td>Explain the outlier detection methods.</td><td>10M</td><td>5</td><td>L2</td></tr>
<tr><td></td><td>OR</td><td></td><td></td><td></td></tr>
<tr><td>20</td><td>Discuss key issues involved in hierarchical and K-Means clustering.</td><td>10M</td><td>5</td><td>L2</td></tr>
</table>