R20 Regulation Subject code: 3P6EB

TKR COLLEGE OF ENGINEERING AND TECHNOLOGY

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

B.Tech VI Semester Supplementary Examinations, February 2024

Data Warehousing and Data Mining

(CSE)

Maximum Marks:70M Date:17.02.2024 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)		Bloom's Tx	CO
1	List and define the characteristics of data warehouse	L1	CO1
2	Define metadata	L1	CO1
3	List the methods for filling missing values	L1	CO2
4	Define KDD process	L1	CO2
5	State apriori property	L1	CO3
6	Define closed item set and maximal item set	L1	CO3
7	What is rule classification	L1	CO4
8	Define Decision tree	L1	CO4
9	Define clustering	L1	CO5
10	What are the requirements for cluster analysis	L1	CO5
	Part-B		
Answer All the following questions. (5X10M=50Marks)			
11	Draw and explain 3-tier architecture of a data warehouse. (10M)	L2	CO1
	OR		
12	List and explain OLAP operations with an example. (10M)	L3	CO1
13	What is data mining? List and explain the motivating challenges of data mining. (10M)	L2	CO2
	OR		
14	With the help of a neat diagram Explain data mining as a step process of knowledge discovery. (10M)	L2	CO2
15	Consider the following transactional data base , generate association rules and identify strong association rules using APRIORI algorithm with support=2 confidence=60%. (10M)	L3	CO3

<table>
<tr><td></td><td><table>
<tr><td>Tid</td><td>items</td></tr>
<tr><td>T1</td><td>Hotdogs, bun, ketchup</td></tr>
<tr><td>T2</td><td>Hotdogs, bun</td></tr>
<tr><td>T3</td><td>Hotdogs, coke, chips</td></tr>
<tr><td>T4</td><td>coke, chips</td></tr>
<tr><td>T5</td><td>Chips,ketchup</td></tr>
<tr><td>T6</td><td>Hotdogs, coke, chips</td></tr>
</table></td><td></td><td></td></tr>
<tr><td></td><td colspan="2">OR</td><td></td></tr>
<tr><td>16</td><td>Write and explain FP-growth algorithm with an example. (10M)</td><td>L3</td><td>CO3</td></tr>
<tr><td>17</td><td>What is Bayesian belief network? Explain with an example. (10M)</td><td>L3</td><td>CO4</td></tr>
<tr><td></td><td colspan="2">OR</td><td></td></tr>
<tr><td>18</td><td>Write KNN classification algorithm with an example. (10M)</td><td>L3</td><td>CO4</td></tr>
<tr><td>19</td><td>a) Write about evaluation of clustering algorithms. (5M)
b) Write the key issues in hierarchical clustering algorithm. (5M)</td><td>L2</td><td>CO5</td></tr>
<tr><td></td><td colspan="2">OR</td><td></td></tr>
<tr><td>20</td><td>Explain agglomerative methods and construct a dendrogram for below distance matrix using single linkage. (10M)
<table>
<tr><td></td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr>
<tr><td>1</td><td>0</td><td></td><td></td><td></td><td></td></tr>
<tr><td>2</td><td>9</td><td>0</td><td></td><td></td><td></td></tr>
<tr><td>3</td><td>3</td><td>7</td><td>0</td><td></td><td></td></tr>
<tr><td>4</td><td>6</td><td>5</td><td>9</td><td>0</td><td></td></tr>
<tr><td>5</td><td>11</td><td>10</td><td>2</td><td>8</td><td>0</td></tr>
</table></td><td>L2</td><td>CO5</td></tr>
</table>