

US  
N

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



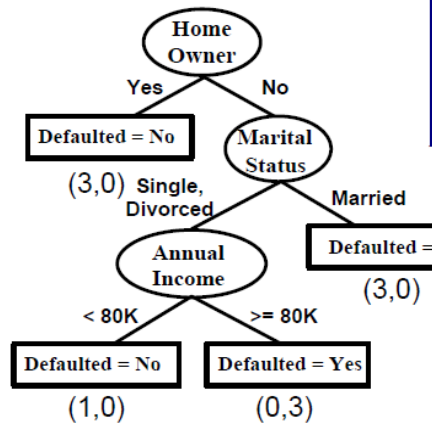
**Internal Assessment Test II- July-2024---Scheme**

Sub :	<b>DATA MINING AND DATA WAREHOUSING</b>					Sub Code:	<b>21IS643</b>	Branch:	<b>ISE</b>
Date:	<b>11/07/2024</b>	Duration:	<b>90 min</b>	Max Marks:	<b>50</b>	Sem/Sec:	<b>VI / C</b>		<b>OBE</b>

**Answer any FIVE FULL Questions**

		MAR KS	CO	RB T												
1	<p>Apply the Apriori algorithm for generating the frequent Item Set for the following set of five transactions, Give algorithm for the same. Let min sup=60% and min conf =80%</p> <table border="1"> <thead> <tr> <th>TID</th> <th>T100</th> <th>T200</th> <th>T300</th> <th>T400</th> <th>T500</th> </tr> </thead> <tbody> <tr> <td>ITEM S</td> <td>{M, O, N, K, E, Y}</td> <td>{D, O, N, K, E, Y}</td> <td>{M, A, K, E}</td> <td>{M, U, C, K, Y}</td> <td>{C, O, O, K, I, E}</td> </tr> </tbody> </table> <p><b>Solution :</b> Given: Support = 60% = 60/100 = 3/5 Confidence = 70% ITEMS : E,K,O : 3 <b>ASSOCIATION RULE:</b></p> <ol style="list-style-type: none"> <li>[E,K] →→ O = 3/4 = 75%</li> <li>[K,O] →→ E = 3/3 = 100%</li> <li>[E,O] →→ K = 3/3 = 100%</li> <li>E →→ [K,O] = 3/4 = 75%</li> <li>K →→ [E,O] = 3/5 = 60%</li> <li>O →→ [E,K] = 3/3 = 100%</li> </ol> <p>∴ Rule no. 5 is discarded because confidence <math>\geq 70\%</math> ,So, Rule 1,2,3,4,6 are selected</p>	TID	T100	T200	T300	T400	T500	ITEM S	{M, O, N, K, E, Y}	{D, O, N, K, E, Y}	{M, A, K, E}	{M, U, C, K, Y}	{C, O, O, K, I, E}	[10]	CO3	L3
TID	T100	T200	T300	T400	T500											
ITEM S	{M, O, N, K, E, Y}	{D, O, N, K, E, Y}	{M, A, K, E}	{M, U, C, K, Y}	{C, O, O, K, I, E}											
2	<p>Construct the FP tree and generate the frequent item set using FP growth algorithm. Let Min Support =3.</p> <table border="1"> <thead> <tr> <th>TID</th> <th>T1</th> <th>T2</th> <th>T3</th> <th>T4</th> <th>T5</th> </tr> </thead> <tbody> <tr> <td>ITEM S</td> <td>I1, I3, I4</td> <td>I2, I3, I5, I6</td> <td>I1, I2, I3, I5</td> <td>I2, I5</td> <td>I1, I3, I5</td> </tr> </tbody> </table> <div style="text-align: center;"> <pre> graph TD     Null[Null] --&gt; I3_4["{I3:4}"]     Null --&gt; I5_1["{I5:1}"]     I3_4 --&gt; I1_1["{I1:1}"]     I3_4 --&gt; I5_3["{I5:3}"]     I5_1 --&gt; I2_1["{I2:1}"]     I1_1 --&gt; I4_1["{I4:1}"]     I5_3 --&gt; I2_1_2["{I2:1}"]     I5_3 --&gt; I1_2["{I1:2}"]     I2_1_2 --&gt; I6_1["{I6:1}"]     I1_2 --&gt; I2_1_3["{I2:1}"]                     </pre> </div> <p><b>Solution:</b></p>	TID	T1	T2	T3	T4	T5	ITEM S	I1, I3, I4	I2, I3, I5, I6	I1, I2, I3, I5	I2, I5	I1, I3, I5	[10]	CO3	L3
TID	T1	T2	T3	T4	T5											
ITEM S	I1, I3, I4	I2, I3, I5, I6	I1, I2, I3, I5	I2, I5	I1, I3, I5											
3	<p>Explain the various approaches for candidate generation with examples. <b>Solution : Brute Force ,Fk-1 X Fk-1, Fk-1 x F1 Item sets</b></p>	[10]	CO3	L2												
4	<p>Explain how to build Decision tree using Hunt's algorithm with Example.</p>															

Tid	Home Owner	Marital Status	Annual Income	Defaulted Borrower
1	Yes	Single	125 K	No
2	No	Married	100 K	No
3	No	Single	70 K	No
4	Yes	Married	120 K	No
5	No	Divorced	95 K	Yes
6	No	Married	60 K	No
7	Yes	Divorced	220 K	No
8	No	Single	85 K	Yes
9	No	Married	75 K	No
10	No	Single	90 K	Yes



**Solution:**

(1,0) (0,3)

5 Explain about various methods for attribute test conditions with example.

**Solution :** With figure

Binary

◦Nominal

◦Ordinal

◦Continuous

6 Illustrate the measures used to select best split in Decision tree classifier. Give example for each. (Write about Gini Index, Information Gain with example.) along with decision tree induction algorithm.

**Solution :**

Gini Index 
$$GINI(t) = 1 - \sum_j [p(j|t)]^2$$

Entropy 
$$Entropy(t) = -\sum_j p(j|t) \log p(j|t)$$

Misclassification error

$$Error(t) = 1 - \max P(i|t)$$

Faculty Signature

CCI Signature

HOD Signature

