Internal Assessment Test II— July-2024Scheme   Sub   DATA MINING AND DATA   Sub   Code:   2118643   Branch:   ISE	US														į.	STATE OF THE PARTY	R		
Sub WAREHOUSING       Sub WAREHOUSING       Sub Code: Code: Code: 21IS643       Branch: ISE         Branch: ISE         Answer any FIVE FULL Questions       IMAR       CO RE         Answer any FIVE FULL Questions       IMAR       CO RE         Any United 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 conference Solution:       TID TIDO TIDO TIDO TIDO TIDO TIDO TIDO T	N		CMR HISTIT AGENEEN								JTE OF TECHNO	DLOGY, BENG GRADE BY N	ALURU.						
Sub WAREHOUSING       Sub WAREHOUSING       Sub Code: Code: Code: 21IS643       Branch: ISE         Branch: ISE         Answer any FIVE FULL Questions       IMAR       CO RE         Answer any FIVE FULL Questions       IMAR       CO RE         Any United 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 conference Solution:       TID TIDO TIDO TIDO TIDO TIDO TIDO TIDO T	Internal Assessment Test II—July-2024Scheme																		
Date: 11/07/2024 Duration: 90 min Max Marks: 50 Sem/Sec: VI/C OBE  Answer any FIVE FULL Questions  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%.  TID 1100 T200 T300 T400 T500 T60, K, E, Suy 1 T100 T100 T00, K, E, Y  M, A, K, E  M, U, C, K, C, G, G, K, L, E  Solution:  Given: Support = 60% = 60100x560100x5 = 3 Confidence = 70%  TTEMS: E, K, O: 3  ASSOCIATION RULE:  1. [E, K] $\rightarrow$ 0 = 3/4 = 75%  2. [K, O] $\rightarrow$ E = 3/3 = 100%  4. E $\rightarrow$ [E, O] = 3/5 = 60%  6. O $\rightarrow$ [E, K] = 3/3 = 100% $\therefore$ Rule no. 5 is discarded because confidence $\geq$ 70%, So, Rule 1,2,3,4,6 are selected  2. Construct the FP tree and generate the frequent item set using FP growth algorithm.  Let Min Support = 3.  TID 11 T2 T3 T4 T5  T1EM 11, 13, 14 12, 13, 15, 16 11, 12, 13, 15 12, 15 11, 13, 13 15  Solution: 3 Explain the various approaches for candidate generation with examples. [10] CO3 L2 Solution: Brute Force, Fk-1 X Fk-1, Fk-1 x F1 Item sets	Sub	DATA MINING AND DATA Sub 211S643 B								В	ranch:	ISE							
Answer any FIVE FULL Questions    Mark   CO   RB   CO   RB   Set	Date:					90	) min				50				V	I/C		OBE	
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%    TID					Ans	swer	any	FIVE			uest	ions	Į.					СО	
ITEM       (M, O, N, K, E, S)       {D, O, N, K, E, Y}       {M, A, K, E}       {M, U, C, K, Y}       {C, O, O, K, I, E}         Solution:         Given:         Support = 60% = 60100x560100x5 = 3         Confidence = 70%         ITEMS: E, K, O: 3         ASSOCIATION RULE:         1. [E, K] → D = 3/4 = 75%         2. [K, O] = $M$		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%										[10]	CO3						
Solution: Given: Support = $60\% = 60100x560100x5 = 3$ Confidence = $70\%$ ITEMS: E, K, O: 3  ASSOCIATION RULE:  1. [E, K] → 0 = $3/4 = 75\%$ 2. [K, O] → E = $3/3 = 100\%$ 3. [E, O] → K = $3/3 = 100\%$ 4. E → [K, O] = $3/4 = 75\%$ 5. K → [E, O] = $3/5 = 60\%$ 6. O → [E, K] = $3/3 = 100\%$ ∴ Rule no. 5 is discarded because confidence ≥> $70\%$ , So, Rule 1,2,3,4,6 are selected  2 Construct the FP tree and generate the frequent item set using FP growth algorithm.  Let Min Support = 3.  TID T1 T2 T3 T4 T5 T1EM I1, 13, 14 12, 13, 15, 16 11, 12, 13, 15 12, 15 11, 13, 15  Solution:  3 Explain the various approaches for candidate generation with examples.  [10] CO3 L2  Solution: Brute Force ,Fk-1 X Fk-1, Fk-1 x F1 Item sets				, N, K,	, E,			, K, E,			K, E}					), K, I, E			
algorithm.  Let Min Support = 3.  TID T1 T2 T3 T4 T5  ITEM I1, I3, I4 I2, I3, I5, I6 I1, I2, I3, I5 I2, I5 I1, I3, I5  Solution:  Solution:  3 Explain the various approaches for candidate generation with examples.  Solution: Brute Force ,Fk-1 X Fk-1, Fk-1 x F1 Item sets	Support = $60\% = 60100x560100x5 = 3$ Confidence = $70\%$ ITEMS: E,K,O: 3 <b>ASSOCIATION RULE:</b> 1. [E,K] $\rightarrow \rightarrow 0 = 3/4 = 75\%$ 2. [K,O] $\rightarrow \rightarrow E = 3/3 = 100\%$ 3. [E,O] $\rightarrow \rightarrow K = 3/3 = 100\%$ 4. E $\rightarrow \rightarrow$ [K,O] = $3/4 = 75\%$ 5. K $\rightarrow \rightarrow$ [E,O] = $3/5 = 60\%$ 6. O $\rightarrow \rightarrow$ [E,K] = $3/3 = 100\%$ $\Rightarrow$ Rule no. 5 is discarded because confidence $\geq \geq 70\%$ , So, Rule 1,2,3,4,6 are											F1.01							
ITEM   I1, I3, I4   I2, I3, I5, I6   I1, I2, I3, I5   I2, I5   I1, I3, I5   I5   I1, I3, I5   I2, I5   I1, I3, I5   I1, I3, I5   I2, I5   I1, I2, I3, I5   I2, I5   I1, I3, I5   I2, I5   I1, I3, I1		algorithm.										[10]	COS	L3					
Solution:    Solution:   Solut		-		T •			12		16	T1				5 1					
3 Explain the various approaches for candidate generation with examples.  Solution: Brute Force, Fk-1 X Fk-1, Fk-1 x F1 Item sets  [10] CO3 L2			S				12		{ 3:4}   		(15:3)	Null [17	{15:1}						
	3	3 Explain the various approaches for candidate generation with examples.									[10]	CO3	L2						
	4													Exan	nple.				

Tid Home Marital Annual Defaulted	l l							
Owner Status Income 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		ı						
Defaulted = No  (3,0) Single, Divorced  Annual Income  (3,0)  Solution:  Marrial Status  Married  Oefaulted = (3,0)  Defaulted = Yes  (1,0)  (0,3)								
5 Explain about various methods for attribute test conditions with example.		1						
Solution: With figure								
Binary								
Nominal								
o Ordinal								
		ı						
o Continuous								
6 Illustrate the measures used to select best split in Decision tree classifier. Give								
example foreach. (Write about Gini Index, Information Gain with example.).along		1						
with decision tree induction algorithm.								
Solution:								
Gini Index $GINI(t) = 1 - \sum_{j} [p(j t)]^{2}$		l						
Entropy								
Entropy $Entropy(t) \vdash \neg \nabla n(i t) \log n(i t)$	Entropy							
$Entropy(t) \models -\sum_{j} p(j \mid t) \log p(j \mid t)$		ı						
Misclassification error								
Misclassification cirol								
$Error(t) = 1 - \max P(i \mid t)$	1							