

w

Cho cơ sở dữ liệu giao dịch sau:

TID	Nội dung giao dịch
100	A C D I
200	A C I
300	C E I
400	ABDE
500	BDI
600	ABDE

Minsup = 0.3, minconf = 1

- Hãy sử dụng thuật toán Apriori để tìm tất cả các tập phổ biến
- Tìm các luật kết hợp

Bài làm

Minsup = 0.3 $\Rightarrow$ count = 2 $\Rightarrow$ Tập phổ biến

Tập ứng cử C1		Tập phổ biến L1	
items	Count	items	Count
A	4	A	4
B	3	B	3
C	3	C	3
D	4	D	4
E	3	E	3
I	4	I	4

Tập ứng cử C2		Tập phổ biến L2	
Items	Count	items	Count
AB	2	AB	2
AC	2	AC	2
AD	3	AD	3
AE	2	AE	2
AI	2	AI	2
BC	0	BD	3
BD	3	BE	2
BE	2	CI	3
BI	1	DE	2
CD	1	DI	2
CE	1		
CI	3		
DE	2		
DI	2		
EI	1		

Tập ứng cử C3		Tập phổ biến L3	
items	Count	items	Count
ABD	2	ABD	2
ABE	2	ABE	2
ACI	2	ACI	2
ADE	2	ADE	2
ADI	1	BDE	2
BDE	2		

Tập ứng cử C4		Tập phổ biến L4	
items	Count	items	Count
ABDE	2	ABDE	2

FI = {L4 hợp l3 hợp l2 hợp l1}

FI = {{ABDE}:2, {BDE}:2 }

ABDE

- Xét tập phổ biến ADBE, tìm luật kết hợp từ tập phổ biến này ($\text{minconf} = 1$)
- Xét loại bỏ 1 tập mục:
 - Xét A:
 - R1: DBE $\square$ A
 - $\text{conf}(R1) = \text{supp}(\text{ADBE}) / \text{supp}(\text{DBE}) = 2/2 = 1 \square$ luật kết hợp
 - Xét D:
 - R2: ABE $\square$ D
 - $\text{conf}(R2) = 1 \square$ R2 là luật kết hợp
 - Xét B:
 - R3: ADB $\square$ B
 - $\text{conf}(R3) = 1 \square$ R3 là luật kết hợp
 - Xét E:
 - R4: ADB $\square$ E.
 - $\text{conf}(R4) = 1 \square$ R4 là luật kết hợp
- Xét loại bỏ 2 tập mục:
 - Xét AD:
 - R5: BE $\square$ AD.
 - $\text{Conf}(R5) = 1 \square$ R5 là luật kết hợp
 - Xét AB:
 - R6: DE $\square$ AB.
 - $\text{Conf}(R6) = 1 \square$ R6 là luật kết hợp
 - Xét AE:
 - R7: DB $\square$ AE.
 - $\text{Conf}(R7) = \frac{2}{3} < 1 \square$ R7 không phải luật kết hợp
 - Xét DB:
 - R8: AE $\square$ DB.
 - $\text{Conf}(R8) = 1 == 1 \square$ R8 là luật kết hợp
 - Xét DE:

- R9: $AB \sqsubseteq DE$.
- $\text{Conf}(R9) = 1 \sqsubseteq R9$ là luật kết hợp
- Xét BE:
 - R10: $AD \sqsubseteq BE$.
 - $\text{Conf}(R10) = \frac{2}{3} < 1 \sqsubseteq R10$ không là luật kết hợp
- Xét loại bỏ 3 tập mục:
 - Xét ADB:
 - R11: $E \sqsubseteq ADB$.
 - $\text{Conf}(R11) = \frac{\text{supp}(ADBE)}{\text{supp}(E)} = \frac{2}{3} < 1 \sqsubseteq R11$ không là luật kết hợp
 - Xét ADE:
 - R12: $B \sqsubseteq ADE$
 - $\text{Conf}(R12) = \frac{\text{supp}(ADBE)}{\text{supp}(B)} = \frac{2}{3} < 1 \sqsubseteq R12$ Không là luật kết hợp
 - Xét ABE:
 - R13: $D \sqsubseteq ABE$
 - $\text{Conf}(R13) = \frac{\text{supp}(ADBE)}{\text{supp}(D)} = \frac{2}{4} = 0.5 < 1 \sqsubseteq R13$ Không là luật kết hợp
 - Xét DBE:
 - R14: $A \sqsubseteq DBE = \frac{\text{supp}(ADBE)}{\text{supp}(A)} = \frac{2}{4} = 0.5 < 1 \sqsubseteq R14$ Không là luật kết hợp
- Xét loại bỏ 4 thuộc tính $\sqsubseteq 1$ bên là rỗng $\sqsubseteq$ dừng thuật toán

ABDE