

Série TD 1 (2 séances)

Motifs fréquents et règles d'association.

Exercice 1

Soit la table des transactions suivante :

TID	Items
T001	P1, P2, P3
T002	P1, P3
T003	P1, P2, P3
T004	P1, P3
T005	P2, P3
T006	P4

- 1) Représenter cette table par une table de transactions formelle.
- 2) Trouver, en utilisant l'algorithme Apriori, l'ensemble L des motifs fréquents sachant que le seuil de support (minsup) est égal à 2.
- 3) Dessiner le treillis des itemsets puis déterminer les motifs fréquents fermés ainsi que les motifs fréquents maximaux.
- 4) Lister toutes les règles solides d'association sachant que le seuil de confiance (minconf) est de 55 %.

Exercice 2

Une base de données possède cinq transactions. On suppose que le support minimum est de 60% et la confiance minimale est de 80%.

TID	Items
T100	M, O, N, K, E, Y
T200	D, O, N, K, E, Y
T300	M, A, K, E
T400	M, U, C, K, Y
T500	C, O, O, K, I, E

- 1) Représenter cette table par une table formelle.
- 2) Trouver, en utilisant l'algorithme Apriori, l'ensemble L des motifs fréquents.

- 3) Lister toutes les règles solides d'association correspondant à la métarègle suivante, où X représente les clients et les item i dénotent les articles achetés :

$$\text{Acheter}(X, \text{item1}) \wedge \text{Acheter}(X, \text{item2}) \Rightarrow \text{Acheter}(X, \text{item3})$$

Exercice 3

Dans un supermarché, on dispose de la base de transactions suivante :

TID	Items	TID	Items
T1	Lait, Jus, Fromage	T6	Lait, Fromage, Pain, Beurre
T2	Pain, Beurre, Lait	T7	Pain, Beurre, Fromage
T3	Lait, Fromage, Sucre	T8	Jus, Fromage
T4	Pain, Beurre, Sucre	T9	Lait, Fromage, Pain, Beurre
T5	Jus, Sucre, Fromage	T10	Jus, Sucre

En utilisant l'algorithme Apriori avec un support minimum de 20% et une confiance minimale de 75%, trouver :

- Les motifs fréquents,
- Les motifs fréquents fermés,
- Les motifs fréquents maximaux,
- Les règles solides d'association de type $A, B \Rightarrow C$.

Exercice 4

Considérons la base des séquences (sequence database) suivante :

SID	Séquence
1	$\langle \{a, b\}, \{c\}, \{f, g\}, \{g\}, \{e\} \rangle$
2	$\langle \{a, d\}, \{c\}, \{b\}, \{a, b, e, f\} \rangle$
3	$\langle \{a\}, \{b\}, \{f, g\}, \{e\} \rangle$
4	$\langle \{b\}, \{f, g\} \rangle$

Déterminer si les séquences ci-dessous sont fréquentes ou non :

- ✓ $\langle \{a\}, \{g\} \rangle$
- ✓ $\langle \{a\}, \{f\}, \{e\} \rangle$
- ✓ $\langle \{b\}, \{f\} \rangle$
- ✓ $\langle \{b\}, \{f, g\} \rangle$
- ✓ $\langle \{f\} \rangle$
- ✓ $\langle \{b\}, \{e\} \rangle$

Avec $\text{minsup} = 3$.