

TD 3 en Fouille de données

Exercice 1 : Soit l'exemple de cours des données transactionnelles d'une épicerie, matérialisées par des tickets de caisse.

ID du ticket	Pain	Lait	Jus	Beurre	Œufs	Cola
1	1	1	0	0	0	0
2	1	0	1	1	1	0
3	0	1	1	1	0	1
4	1	1	1	1	0	0
5	1	1	1	0	0	0

1. Dessiner le treillis des itemsets de ces données.
2. En prenant comme support minimum la valeur 40%, donner les tableaux successifs des k-itemsets fréquents ($k \geq 2$) en illustrant les itemsets éliminés à chaque étape. Le passage entre deux tableaux consécutifs doit se faire par fusion des itemsets (algorithme Apriori).
3. En prenant comme confiance minimale la valeur 60%, générer les règles d'associations pour chaque itemset fréquent.

Exercice 2 : Un fabricant de téléphones portables a collecté des données d'activation des options dans la barre supérieure pendant une période donnée. Les options sont les suivantes : Données mobiles (D), Silence (S), Bluetooth (B), Localisation (L), Rotation Automatique (R). Le tableau ci-dessous représente l'ensemble des options activées en même temps pour différents téléphones pendant une fenêtre de temps. On veut analyser les associations entre ces options

DLR	DSBLR	SBR	DBL	DSBLR
DLR	SBR	DSLRL	DSLRL	DB
SBLR	DSR	SLR	DBL	DSR
BLR	DSBLR	SBLR	DLR	DSBLR

1. En utilisant l'algorithme Apriori, donner les tableaux successifs des itemsets en indiquant à chaque fois les itemsets fréquents pour un support de 50%.
2. Quelle est la meilleure règle en termes de support/confiance ? Cette règle est-elle utile ? Justifier la réponse.
3. Comment peut-on utiliser cette règle concrètement (donner un exemple)
4. Sachant que les options "Données mobiles", "Bluetooth" et "Localisation" sont consommatrices d'énergie et les deux autres sont de bien-être numérique, expliquer comment il est possible de connaître à quel degré
 - a. Le choix d'une seule option consommatrice d'énergie influence le choix d'au moins une option de bien-être numérique.
 - b. Le choix d'une seule option de bien-être entraîne la décharge maximale de la batterie du téléphone