

Examen final

Règles d'association

le 18 décembre 2014, durée 1h

L'ordre de résolution des sujets n'est pas imposé. Les sujets sont indépendants.

1. Comment peut-on traiter un aspect temporel¹ à des transactions de type "achat" (panier de la ménagère) ?
2. Souvent les campagnes marketing visent à mettre en évidence des produits ciblés (nouveaux sur le marché ou qu'on envisage de remplacer ou changer de présentation). Il serait donc intéressant de pouvoir obtenir et étudier exclusivement les règles de type :

$$X \Rightarrow \alpha$$

avec X un ensemble (quelconque) d'items, $Sink$ l'ensemble des items à étudier et $\alpha \subset Sink$ un sous-ensemble non-vide de $Sink$. On appellera ces règles "partie droite imposée"

Proposer une méthode (originale ou adaptation de l'algorithme *A priori*) qui permet d'obtenir les règles d'association à partie droite imposée.

3. La base suivante traduit une série de transactions de type "panier".

1. exemples d'aspect temporel : tard le soir, pause déjeuner, veille de fête, mercredi

Id	Transaction
T100	M, O, N, K, E, Y
T200	D, O, N, K, E, Y
T300	K, E, Y
T400	M, O, N, E, Y
T500	D, O, N, E
T600	M, O, D, E
T700	O, N, E
T800	E, Y, D

Soient les limites du support ($min_support$) à 25% et de la confiance ($min_confidence$) à 80%.

- Appliquant l'algorithme **A priori** calculer tous les itemsets fréquents par rapport à $min_support$ (*Attention : la limite est donnée en pourcentage !*)
- A partir du résultat calculé au point précédent calculer des règles d'association de type $x, y \Rightarrow E|Y|(E, Y)$ (la notation x, y indique que la partie gauche a deux éléments, pour la partie droite la notation $E|Y|(E, Y)$ indique au moins un des éléments E et Y) qui sont au-delà de $min_confidence$.