

## Data Mining : Règles d'association

### TD N : 4

#### Exercice 1

On veut appliquer le modèle des "Règles d'association" à un problème de TextMining.

Le tableau suivant représente les mots-clés (les mots les plus importants) extraits à partir de 7 textes.

Appliquez l'algorithme a priori pour trouver toutes les règles d'association qui vérifient  $\text{minsup} \geq 40\%$  et donnez leur confiance.

N° Texte	Mots clés
1	Finance, Marché, Budget, Economie
2	Ouverture, Finance, Economie
3	Ouverture, Assemblée, Handball, Sport
4	Directeur, Budget, Finance, Economie
5	Directeur, Assemblée, Handball, Sport
6	Ouverture, Marché, Economie
7	Ouverture, Assemblée, Directeur, Handball, Sport

#### Exercice 2

Dans un magasin de produits informatiques, on a enregistré les achats suivants faits par les clients (transactions) dans le tableau ci-dessous.

On veut appliquer le modèle des "Règles d'association" pour extraire des règles de types :  $\{X \rightarrow Y\}$  ou  $\{X, Y \rightarrow Z\}$  ou  $\{X \rightarrow Y, Z\}$ .

Pour faciliter la suite de l'exercice, on associe à chaque item un code qui sera utilisé pour répondre aux différentes questions. Les codes associés sont les suivants :

- P1 = CD
- P2 = DVD
- P3 = Jeu
- P4 = Antivirus
- P5 = Souris
- P6 = Caméra
- P7 = Tablette
- P8 = FlashDisk

Transactions	Items
1	CD, DVD, Jeu, Antivirus
2	CD, Jeu, Souris, Caméra, Antivirus
3	Antivirus, FlashDisk, CD, Jeu, Caméra
4	Jeu, DVD, Tablette
5	FlashDisk, Antivirus, Tablette, Souris

**Travail à faire :**

1. Trouver tous les items et les ensembles d'items de taille 2 et de taille 3 fréquents sachant que **minsup = 60%**. Détaillez toutes les étapes.
2. Dédurre toutes les règles d'associations à partir des ensembles d'items fréquents.
1. Quelles sont les règles ayant une **confiance  $\geq 80\%$** . Montrez tous les calculs.