



TD 4 Extraction de connaissances (Partie 1)

Exercice 1 :

1. Citer deux synonymes de datamining ?
2. Citer deux disciplines liées au datamining autres que les statistiques ?
3. Quelles sont les caractéristiques des BDD utilisées en datamining ?
4. Citer deux types de datamining selon la nature des données ?
5. La décision est-elle toujours automatique en datamining ?
6. Citer deux domaines d'application de datamining ?
7. Citer deux techniques simples de datamining ?

Exercice 2 : Soit la base suivante issue des courriers signalés comme spam ou non.

Ami ,spam, transfert
banque, spam, aide
Ami, banque, spam, aide
banque, aide



1. Pour le 1-item « banque » quel est le support ?
2. Pour le 2-item « ami, aide » quel est le support ?
3. Pour le 3-item « ami,aide, banque » quel est le support ?

Un item fréquent est un item dont le support dépasse le seuil ($\text{minSupp} = 2/4$).

4. Citer un 1-item non fréquent ?
5. Citer un 2 item fréquent ?
6. Quelle est la confiance de la règle d'association « ami,aide \rightarrow spam »

La confiance minimale = 100%.

7. Citer une règle d'association sûre ?