

FICHE SAVOIR MARKETING –méthode d'analyse et de traitement : Le Data Mining

Définition

Le Data Mining est en fait un terme générique englobant toute une famille d'outils facilitant l'exploration et l'analyse des données contenues au sein d'une base décisionnelle de type Data Warehouse ou DataMart.

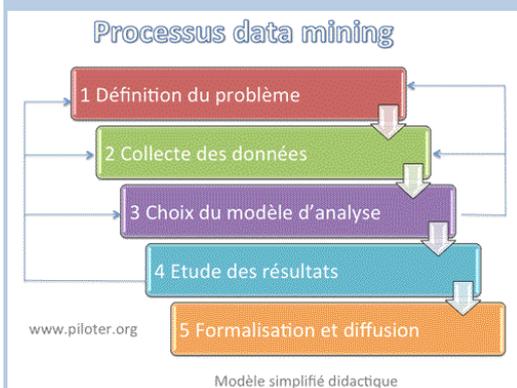
Les techniques mises en œuvre

Différentes techniques sont proposées. Elles sont à choisir en fonction de la nature des données et du type d'étude que l'on souhaite entreprendre

-Les méthodes utilisant les techniques de classification et de segmentation

- Les méthodes utilisant des principes d'arbres de décision assez proches des techniques de classification
- Les méthodes fondées sur des principes et des règles d'associations ou d'analogies
- Les méthodes exploitant les capacités d'apprentissage des réseaux de neurones
- Et pour les études d'évolution de populations, les algorithmes génétiques
- *Algorithmes Naïve Bayes, séries chronologiques, régression linéaire...*

Le principe : une démarche en 5 temps majeurs



Définition du problème

Quel est le but de l'analyse, que recherche-t-on ? Quels sont les objectifs ? Comment traduire le problème en une question pouvant servir de sujet d'enquête pour cet outil d'analyse bien spécifique ? A ce sujet, se souvenir que l'on travaille à partir des données existantes, la question doit être ciblée selon les données disponibles.

Collecte des données

Une phase absolument essentielle. On n'analyse que des données utilisables, c'est à dire "propres" et consolidées.

Construire le modèle d'analyse

Ne pas hésiter à valider vos choix d'analyse sur plusieurs jeux d'essais en variant les échantillons. Une première évaluation peut nous conduire à reprendre les points 1 ou 2.

Etude des résultats

Il est temps d'exploiter les résultats. Pour affiner l'analyse on n'hésitera pas à reprendre les points 1, 2 ou 3 si les résultats s'avéraient insatisfaisants.

Formalisation et diffusion

Les résultats ne seront utiles qu'une fois devenus une connaissance partagée. C'est bien là l'aboutissement de la démarche. C'est aussi là que réside la difficulté d'interprétation et de généralisation.