



# Data mining: chaque entreprise renferme une mine d'or

Un constat clair apparaît si l'on se penche sur dix ans d'expérience dans des projets de data mining menés en Suisse: grâce aux logiciels spécialisés toujours plus conviviaux et complets, et grâce au savoir-faire en matière de conseil, ce qui était pendant longtemps pénible est devenu une promenade de santé. Daniel Schloeth

Qui ne désirerait pas voir l'avenir au lieu de ne s'occuper toujours que du passé? Les outils de reporting et les bases de données multidimensionnelles (OLAP) sont indispensables à la compréhension de ce qui a déjà eu lieu. Les requêtes aux bases de données permettent par exemple de connaître le chiffre d'affaires du mois précédent. Les cubes OLAP, tels qu'ils sont typiquement utilisés dans les applications de business intelligence, vont plus loin en permettant par exemple d'obtenir les revenus détaillés par produit ou par région commerciale. Si précis soient-ils, tous ces outils ne font cependant qu'offrir un point de vue dans le rétroviseur.

## Seul le data mining regarde dans le futur

À l'inverse, le data mining montre quelle personne va vraisemblablement acheter tel produit le mois prochain. Et c'est grâce à la reconnaissance algorithmique de *patterns* (modèles récurrents) dans le passé que des prévisions étonnamment précises sont possibles.

Le data mining basé sur de l'intelligence artificielle est adapté à presque toutes les questions sur les données et il est typiquement utilisé dans les domaines suivants:

- Comprendre ce que les clientes et clients veulent et les segmenter dans des groupes homogènes en fonction de leurs besoins.
- Identifier les clients les plus profitables et en gagner de nouveaux répondant à ce profil.
- Améliorer les ventes croisées (*cross-selling*) et l'augmentation des ventes aux clients existants (*up-selling*).
- Réduire les pertes de clients et augmenter leur loyauté.
- Diminuer les coûts des campagnes de marketing direct sans nuire à leurs résultats.

Dr. Daniel Schloeth est membre de la direction de SPSS Suisse et responsable des relations avec la clientèle.



L'exploitation de campagnes marketing est une entrée idéale dans le data mining. Source: SPSS Suisse

- Elever les revenus de la grande distribution au moyen d'analyse des paniers d'achat et d'optimisation des parcours et actions dans la grande surface.
- Déterminer les risques liés aux crédits.
- Reconnaître des fraudes ou du lavage d'argent.
- Identifier des électeurs potentiels et les motiver à voter.

Au-delà de ces usages commerciaux ou électoraux, le data mining ouvre un large champs d'applications pour les administrations, dans les sciences, la médecine et la recherche – en prédisant par exemple les phénomènes climatiques ou le développement d'une maladie.

Nettement en retard pas rapport aux pays anglo-saxons, l'utilisation professionnelle du data mining ne s'est établie en Suisse qu'à la fin des années 90 et au début de la décennie suivante. Tout d'abord dans des secteurs hautement compétitifs avec des entreprises proposant des offres relativement similaires à un grand nombre de clients – concrètement dans les banques et les télécoms. En effet, plus les

produits et services sont proches les uns des autres et donc aisément substituables, plus il est important de se différencier de la concurrence au moyen de campagnes marketing ciblées et adaptées aux clients individuels.

## Pourquoi le data mining ne fait-il pas encore partie du quotidien?

Le data mining répond à un paradoxe connu, qui veut que plus les données client sont nombreuses, plus il est difficile et laborieux de les analyser à l'aide de méthodes traditionnelles et d'en tirer une quelconque utilité. Ce qui devrait être une mine d'or de connaissances reste souvent inutilisé par manque de personnel, de temps et d'expertise.

Bien que le data mining des données client fasse sens pour la plupart des grandes et moyennes entreprises et bien qu'il apporte des retours rapides, beaucoup de sociétés suisses s'offrent encore le luxe d'y renoncer. En premier lieu parce que beaucoup d'entreprises savent que la qualité de leurs données n'est pas parfaite – elles préfèrent ainsi continuer de les tester ou investir encore plus dans

des bases de données. Ce faisant, elles s'empêchent encore plus profondément dans les données clients précieuses déjà accumulées au lieu de commencer à en retirer une utilité. D'autres entreprises encore plus raffinées remplacent un data warehouse fonctionnel existant par celui d'un autre fournisseur avant de planifier l'exploitation sérieuse de leurs données existantes. Il faudrait en fait prendre le chemin inverse: même s'ils s'avèrent nécessaires, les énormes investissements dans des systèmes de data warehouse et autre CRM ne deviennent véritablement profitables que grâce au data mining.

D'autres sociétés disposent d'immenses quantités de données et, grâce aux énormes investissements dans un ERP, connaissent l'historique complet de chacun de leurs clients. Cependant, elles n'en usent que pour traiter les commandes aussi rapidement que possible, puis les archiver. Enfin, certaines entreprises sont conseillées par des consultants pendant des périodes si longues qu'à la fin il ne reste plus d'argent pour la réalisation des projets de data mining.

Mais le tableau n'est pas si sombre: le data mining s'impose dans de nombreuses branches en Suisse, même si cela se fait par des chemins sinueux. Et il s'agit là de véritable data mining, et non pas de données de business intelligence certes raffinées mais orientées vers le passé, que certains fournisseurs aimeraient continuer à vendre sous l'appellation de data mining.

### Comment démarrer?

Les campagnes marketing constituent une porte d'entrée idéale dans le domaine du data mining. Les vertus de la démarche apparaissent en effet facilement en sélectionnant la moitié des adresses d'un mailing de manière habituelle et l'autre moitié au moyen d'algorithmes de data mining. Des campagnes marketing plus ciblées permettent de réduire les coûts jusqu'à 80 %, si une offre n'est envoyée qu'aux personnes ayant une grande probabilité de s'y intéresser. Le retour sur investissement en résultant est ainsi nettement plus important que dans des campagnes diffusées à large échelle.

Que faire pourtant lorsqu'une entreprise ne souhaite pas déployer une infrastructure analytique ou que les collaborateurs en place ne peuvent être assignés à ces nouvelles tâches, si importantes soient-elles? A une époque où les sociétés sont peu disposées à supporter de nouveaux travaux ou à engager du personnel supplémentaire, l'outsourcing du data mining à des spécialistes peut s'avérer utile à l'obtention de gains rapides.

### Dix ans de progrès...

Beaucoup de progrès ont été accomplis depuis l'arrivée des premiers logiciels de data mining en Suisse. Chaque nouvelle version des outils les plus connus apporte son lot de nouveaux algorithmes, parfois exotiques. Après des années de concurrence fonctionnelle entre les fournisseurs, les usagers disposent aujourd'hui d'un vaste choix de fonctionnalités telles que CARMA, machines à vecteurs de support, régression de Cox, *self learning response model* - avec pour effet que les utilisateurs occasionnels peinent parfois à s'y retrouver.

Les éditeurs ne s'y trompent pas et leurs initiatives récentes cherchent à alléger le travail de l'utilisateur plutôt que de le laisser seul devant une pléthore d'options. Tout d'abord en automatisant la préparation des données, ce qui est souvent le secret d'un bon résultat ; on pense ici en particulier au choix des variables, à la catégorisation et à la dichotomisation. Ces processus prennent en général beaucoup de temps et exigent une grande expertise. Les logiciels de data mining les plus modernes résolvent cette problématique de manière élégante: ils proposent une préparation adaptée pour chacune des particularités présentes, se basant tant sur des caractéristiques intrinsèques des données elles-mêmes que sur la connaissance de l'ensemble. Au final, les données sont préparées correctement et l'on peut passer directement à la modélisation. Un second apport consiste à augmenter la productivité au moyen de méta-instructions qui lancent plusieurs algorithmes en parallèle et en représentent immédiatement les résultats comparés, afin de pouvoir choisir le modèle optimal.

La vie des analystes s'est considérablement simplifiée ces dix dernières années. Leurs logiciels de data mining leur permettent d'analyser des quantités de données beaucoup plus grandes, ils disposent d'un vaste choix d'algorithmes, et de nombreuses étapes fastidieuses sont désormais exécutées automatiquement en arrière-plan. Sans oublier bien sûr que ces logiciels sont bien moins chers que par le passé.

### ... et ça continue

Le data mining est par ailleurs sorti ces dernières années du carcan de l'analyse des bases de données:

- Le *text mining* permet la reconnaissance automatique d'assertions au sein de textes libres et améliore les pronostics de comportement grâce à la prise en compte d'informations non-structurées à disposition.
- Le *web mining* étudie les logs sur internet et permet de prédire le comportement des

clients en ligne en temps réel. Les internautes se voient ainsi proposés des offres adaptées individuellement, augmentant par la même le nombre d'accès et la profitabilité du site.

- En complétant les données dures des clients par des informations *soft* issues de questionnaires clients, il est possible de prendre en compte leurs opinions et préférences personnelles. Cela permet de jeter un regard sur les motivations d'achat de la clientèle.
- Avec l'arrivée d'applications de *speech mining*, vient la promesse d'analyses en temps réel et de prédictions à partir des conversations téléphoniques.
- À l'aide des méthodes évoquées précédemment, l'analyse prédictive à l'échelle de l'entreprise lie toutes les informations à disposition d'une société et fournit comme résultat des suggestions pour répondre et réagir aux désirs des clients de façon optimale au moyen des systèmes opérationnels existants.

Une seule chose n'a pas changé durant la dernière décennie: il est toujours aussi crucial de connaître ses propres données, de posséder un savoir-faire sectoriel et de bien évaluer les résultats obtenus. Se borner à acheter un logiciel et en espérer qu'il fournisse la solution à tous les problèmes ne fonctionnera pas non plus à l'avenir. <



Le data mining aide à comprendre ce que les clients veulent et à les segmenter dans des groupes homogènes en fonction de leurs besoins.

Source: SPSS Suisse