

# Data mining et marketing bancaire

Alain Lot  
Business intelligence  
solutions director  
IBM Europe



Le *data mining* permet des segmentations et des croisements de plus en plus fins dans les bases de données. Il constitue un instrument marketing particulièrement adapté aux grandes banques possédant plusieurs millions de clients.

L'informatique est un outil essentiel pour le traitement des transactions quotidiennes des établissements financiers. Non seulement elle a permis des gains de productivité importants et l'arrivée de services supplémentaires tels que guichets automatiques, banque à domicile... mais elle autorise aujourd'hui des développements nouveaux, notamment en matière de marketing.

Le *data mining*, c'est-à-dire l'exploitation des grands ensembles historiques de transactions, constitue l'un d'eux. Il a pour objectif de découvrir des informations cachées dans l'historique des transactions concernant placements, clients, etc. La grande différence avec les systèmes classiques de recherche dans les bases de données historiques, tient à ce qu'il n'implique

pas la génération de questions pertinentes en amont.

Par exemple, un responsable de clientèle constatant que la moyenne des placements baisse dans certaines agences n'aura pas besoin d'émettre des hypothèses en vérifiant les investissements de tous les clients de la tranche d'âge et de la catégorie socio-professionnelle dont il suppose, à tort ou à raison, que ce segment est en cause. Non seulement le *data mining* le fera automatiquement à sa place mais il pourra mettre en évidence, sans intervention extérieure, l'existence d'autres segments de clientèle intervenant d'une manière beaucoup plus grande qu'on aurait pu le soupçonner.

Autre exemple ❶, une banque peut avoir déterminé, par des méthodes classiques, que ses clients les

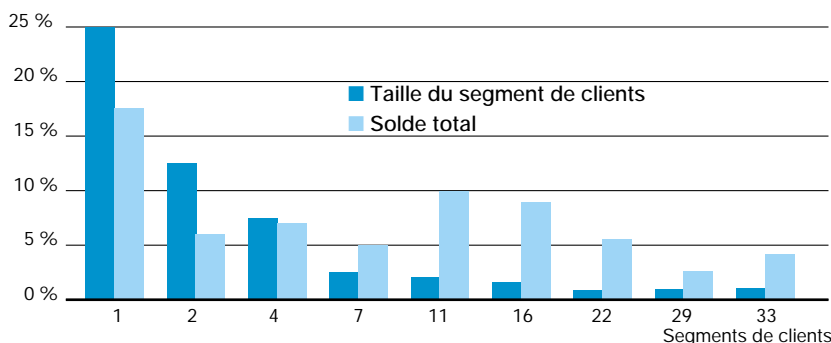
plus rentables se situent dans les segments 1, 2 et 4. Une opération de *data mining*, plus pointue, fera apparaître qu'il existe en réalité six autres segments de clientèle cachés dans les trois premiers, que les études précédentes (statistiques, analyse multidimensionnelle...) n'ont pas permis d'identifier. Si ces six segments 7, 11, 16, 22, 29 et 33 représentent un nombre faible de clients (raison pour laquelle une interrogation directe ne les a pas fait ressortir), leur contribution se révèle très importante en termes de rentabilité.

## Une identification automatique des paramètres

Le *data mining* diffère aussi des techniques classiques par le fait qu'il traite, avec des temps de calcul très rapides, des ensembles de transactions pouvant concerner jusqu'à plusieurs millions d'acteurs eux-mêmes caractérisés par plusieurs centaines de paramètres, identifiant de lui-même les paramètres de chaque segment, par ordre d'importance.

Ainsi pour reprendre l'exemple précédent, chacun des dix segments de clientèle, mentionné par sa taille en pourcentage de la population et sa contribution au revenu, est caractérisé par ses principaux paramètres : existence d'un compte courant, d'une carte de crédit, avoir moyen, nombre

### ❶ Segments de clients les plus rentables



Le *data mining* fait appel à des algorithmes mathématiques complexes : analyse relationnelle, réseaux neuronaux, arbres de décision... qui demandent des puissances de calcul informatique importantes, sur postes de travail ou sur ordinateurs parallèles, dans la mesure où l'on veut exploiter des bases de données historiques de plusieurs centaines de milliers, voire de plusieurs millions de transactions, chacune étant caractérisée par plusieurs dizaines, voire plusieurs centaines de paramètres. Sur les ordinateurs modernes, les algorithmes de *data mining*

demandent des temps de calcul raisonnables. Les résultats peuvent être présentés sous forme graphique (logiciel *Intelligent Miner* d'IBM), où la description des segments de clients est représentée par ses paramètres, par ordre d'importance.

Le *data mining* s'appuie sur les données historiques des transactions, pouvant provenir de plusieurs systèmes opérationnels : comptes courants, cartes bancaires, investissements, etc. Les données sont en général entreposées dans un système différent des

systèmes de production, appelé entrepôt de données (*data warehouse*) qui a pour fonction d'extraire ces données des systèmes opérationnels, de les rassembler sous un format commun après nettoyage, dédoublonnage, etc., de réaliser les catalogues d'information ainsi que d'automatiser les opérations de mises à jour répétitives. La mise en œuvre d'un *data warehouse* demande des compétences d'«intégration de systèmes» pour extraire et faire fonctionner ensemble des données de systèmes informatiques multiples. L'état de l'art actuel consiste

à mettre en place l'entrepôt de données par étapes, en créant tout d'abord des magasins de données (*data warehouse*), entrepôts de données spécialisés par fonction, par exemple pour le marketing, que l'on intègre ensuite.

Il n'est pas indispensable d'avoir un *data warehouse* ou un *data mart* pour effectuer une opération de *data mining*, dans la mesure où l'on peut mettre en forme les données spécialement pour cette occasion ; néanmoins ; des opérations répétées tireront profit de son existence.

de transactions, nombre moyen de débits, âge du titulaire, accord de découvert, assurance vie, etc. ②.

Les applications du *data mining* sont nombreuses. Dans le domaine du marketing, en particulier, par une connaissance fine de la clientèle, il

permet un ciblage beaucoup plus pertinent. Ainsi, dans les opérations de publipostage dont le rendement est traditionnellement faible, il fournit une précision nettement améliorée. Toujours dans le même exemple, la banque concernée a réussi à faire pas-

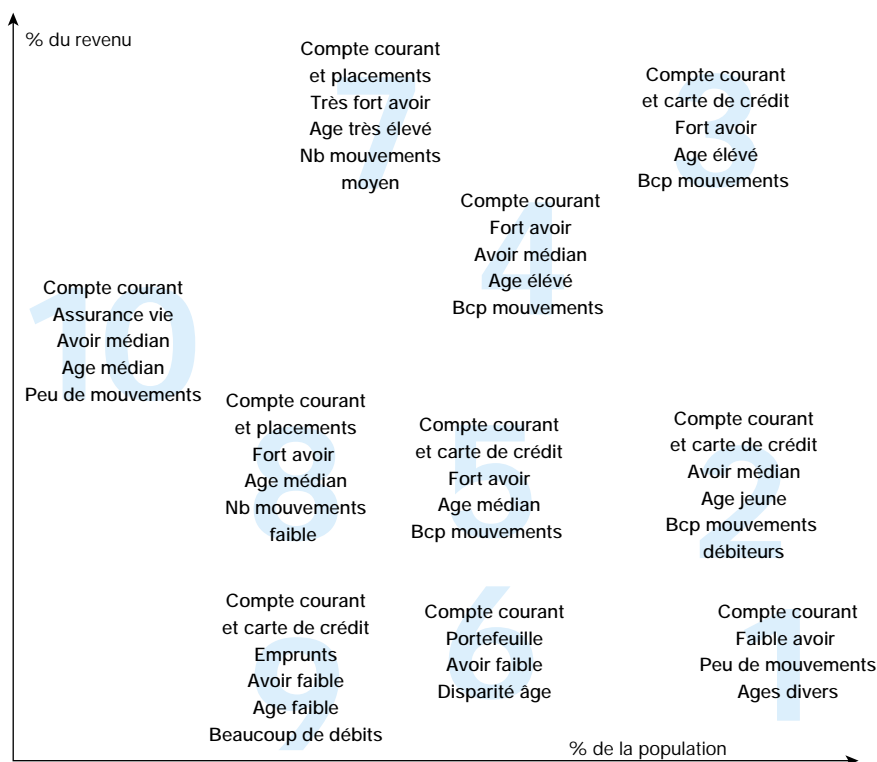
ser le taux de réponse à son publipostage de 2,2 à 3,7 %, en ciblant de manière beaucoup plus personnalisée les différents segments de clients visés.

### Traiter chaque client comme un client unique

La stratégie marketing d'«intimité avec le client» (*customer intimacy*) consiste à traiter chacun d'eux et de ses proches comme s'il était unique, grâce à la connaissance par l'analyse de ses transactions passées du segment type de clientèle auquel il appartient. La segmentation de clientèle est donc un outil indispensable pour y parvenir et le *data mining* est le moyen de la réaliser avec des segments pouvant atteindre plusieurs centaines, sur un nombre de paramètres importants. L'analyse va jusqu'au calcul de la rentabilité de chaque segment type, ainsi que de l'évaluation du risque que ce type de client a de quitter l'établissement financier.

Or, on admet généralement que conquérir un nouveau client coûte environ cinq fois plus cher que conserver un client existant. Sachant que les segments types présentant un risque d'attrition sont représentés par des paramètres tels que nombre d'erreurs d'imputation, nombre de réclamations, diminution du nombre de transactions, diminution des investissements, etc., il est très important de les identifier afin de monter des actions marketing adaptées.

## ② Paramètres des 10 segments les plus importants



Le *data mining* permet également des regroupements de données indiquant l'ensemble des produits et services utilisés par un client et ses proches, d'où, là encore, un marketing spécifique.

Le dédoublement, opération indispensable au nettoyage des fichiers, et dont les résultats vont bien au-delà de la seule économie réalisée, est obtenu par des fonctions de reconnaissance capables d'identifier que deux personnes ayant la même adresse, dont les noms ont été épelés différemment par erreur, peuvent être un individu unique ou posséder un lien de parenté. De la même façon, deux personnes ayant le même nom, mais dont les adresses ont été par erreur enregistrées différemment, peuvent être la même, ou parentes.

Cette reconnaissance est notamment réalisée grâce à des techniques linguistiques (phonétique...) identifiant les différentes manières d'épeler un mot avec une prononciation identique. La même technique est utilisée pour la reconnaissance de la parole par les ordinateurs (dictée automatique).

### **Déterminer les opportunités de ventes croisées**

L'«intimité avec le client» va jusqu'à prendre en compte dans les paramètres le contenu de sa correspondance avec sa banque. Cette fonction d'analyse de texte (*text mining*), identifie les mots importants de cette correspondance : recherche d'information, demande de transaction, insatisfaction, erreur... ; elle est obtenue, au moyen d'outils informatiques nouveaux qui se développent aujourd'hui et font aussi appel aux techniques linguistiques ainsi qu'à la segmentation, pour reconnaître les mots traitant de sujets semblables.

Le *data mining* est utilisé également pour déterminer les associations de produits acquis au cours des différentes transactions ; en partant de l'étude de ces associations de produits (placements, assurance vie, etc.), l'analyse des grandes bases de données indique quels sont ceux qui ont été acquis simultanément. Ces différentes informations permettent d'identifier le profil type du client devant posséder simultanément plu-

sieurs types de produits, et peuvent ainsi être utilisées par les établissements pour des opérations de ventes croisées (*cross selling*).

La détection de fraudes est rendue plus facile par le *data mining*. Par exemple, connaissant le profil type des transactions carte bancaire de chaque client après l'analyse historique de ses transactions, la banque est à même de détecter toute anomalie importante. Dans un autre domaine, il peut servir pour déterminer le meilleur emplacement d'un guichet automatique, en analysant les transactions du parc de guichets installés, en fonction des multiples paramètres : proximité de commerces (bureau de tabac, librairie, café...), de transports (gares, autobus, métro...), de la concurrence ; facilité d'accès ; visibilité, etc.

Dans la mesure où les coûts informatiques ont baissé de manière significative et où, de ce fait, les données historiques des transactions sont accessibles en direct, les perspectives du *data mining* sont considérables et vont bien au-delà du seul aspect marketing. ■