

Décryptage

Le décisionnel se projette dans l'avenir **P 18**

Pratique

Orange décortique sa base de clients **P 22**

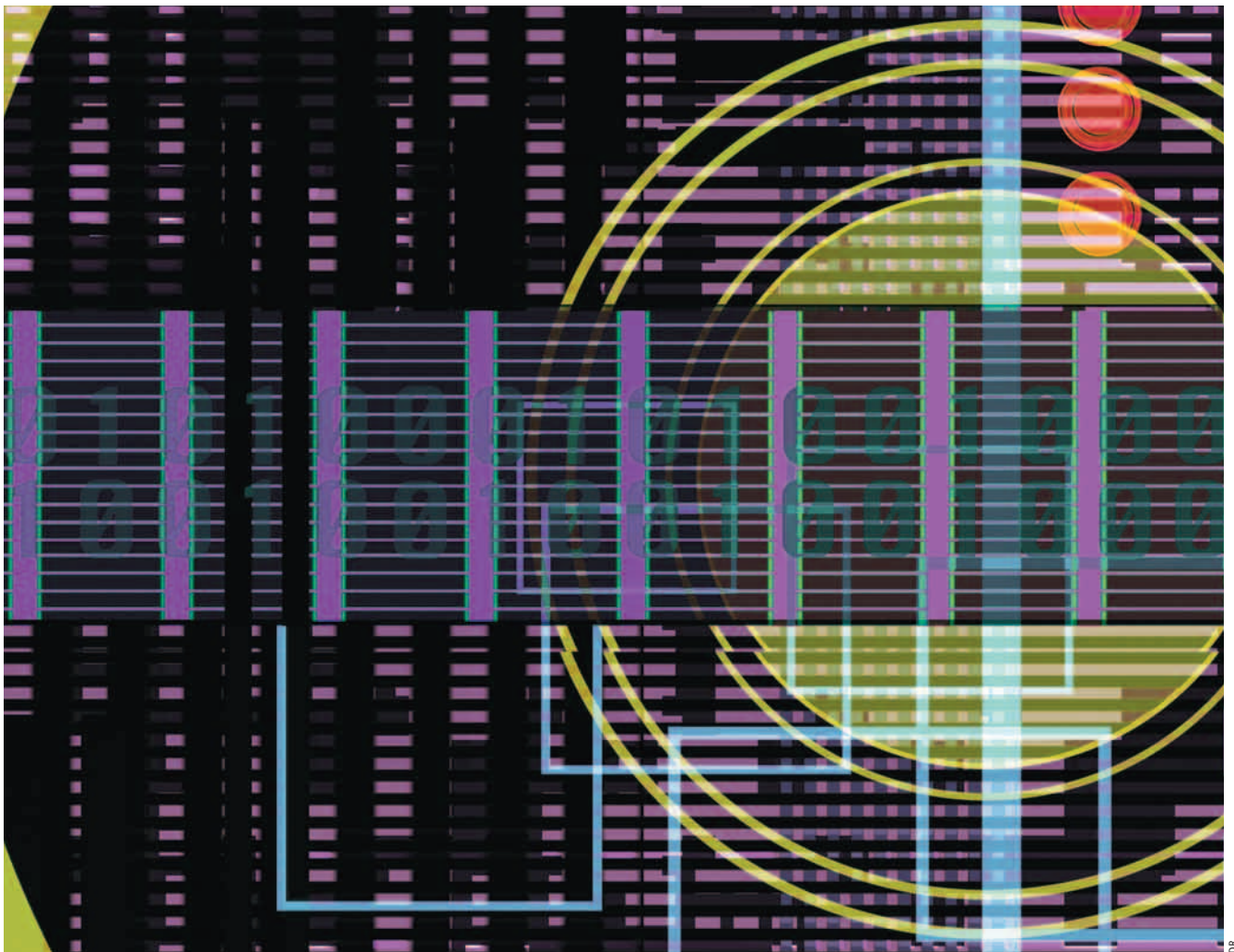
Zoom

La distribution : au-delà du marketing de masse **P 24**

Prospective

Le décisionnel frappe à la porte des DRH **P 25**

Décisionnel : détour vers le futur



Après avoir décortiqué les événements passés, le décisionnel se tourne vers l'avenir avec les outils d'analyse prédictive. Basées sur les données issues de la GRC et surtout des PGI, ces applications tentent de prévoir le comportement des clients, afin d'apporter les réponses les plus à même de les fidéliser. Des projets hautement stratégiques, mais aussi difficiles à mener. Car l'analyse prédictive passe par de lourdes manipulations de données.

ORANGE FRANCE EN FAITS ET EN CHIFFRES

Activité : téléphonie mobile.
Chiffre d'affaires : près de 8 milliards d'euros en 2003.
Effectif : 6 645 personnes en 2003.
Abonnés : 20,3 millions en 2003.

PROBLÉMATIQUE :

Mieux connaître ses abonnés et suivre son moteur de croissance.

En 2000, Orange se rend compte que sa croissance sera désormais tirée par la fidélisation de ses d'abonnés, et non plus par la conquête de nouveaux clients. La connaissance des abonnés (côté prédictif) et le suivi du moteur de croissance (base de clients) deviennent donc stratégiques. Pour répondre à cette adaptation, Orange dispose déjà d'une plate-forme de GRC analytique, qui consiste en un entrepôt de données séparé en deux datamarts.

LE PROJET :

Historique :

Fin du second semestre 2000

Démarrage du projet. Orange constate les limites de l'outil décisionnel dont il dispose et passe un appel d'offres pour une solution d'analyse prédictive.

Mi-2001 Installation de l'outil K2R de l'éditeur Xken en une semaine, sur des postes autonomes.

Septembre 2001 Déploiement de l'outil SAS Enterprise Miner (tous modules confondus). Trois mois seront nécessaires pour rendre les bases de données opérationnelles et corriger les bogues.

Fin mars 2002 Démarrage progressif de l'utilisation des modules d'analyse prédictive de SAS. Une trentaine de personnes effectuent de l'analyse prédictive sur la plate-forme SAS et Xken.

Solution technique :

- **SAS Enterprise Miner** sert à la compréhension des phénomènes complexes, comme l'étude du mode de vie des clients, facteur qui conditionne leur consommation. L'outil se connecte à la plate-forme GRC d'Orange, un entrepôt de données séparé en deux datamarts.
- **Le logiciel K2R**, de Xken, prend en charge les mesures de potentiel des échantillons de clients (*scoring*).

Investissement :

n.c. (plusieurs millions d'euros).

Pratique Orange : connaître les clients pour les fidéliser

Les opérateurs de téléphonie mobile sont à la pointe de l'analyse prédictive. Pour mieux connaître ses clients et éviter leur départ à la concurrence, Orange a retenu deux outils informatiques différents, pour des usages complémentaires.

Le développement de l'analyse prédictive, ou GRC (gestion de la relation client) analytique, est né d'un virage stratégique de l'opérateur mobile Orange. "Courant 2000, nous nous sommes aperçus que la croissance des revenus n'allait plus dépendre de l'acquisition de nouveaux clients, mais de ceux que nous avions déjà", précise Arnaud Muller, responsable du datamining décisionnel chez Orange. Les limites des outils décisionnels dont nous disposions sont alors apparues clairement. "Un constat qui vaut pour l'ensemble des opérateurs mobiles. "Dans le domaine de l'analyse prédictive, les opérateurs de téléphonie mobile sont partis de zéro et possèdent maintenant quelques années d'expérience", note ainsi Michel de Bollivier, chef de projet chez Business & Decision. Avec plus de 20 millions de clients, Orange France a rencontré les problèmes les plus importants posés par cette nouvelle discipline : gros volumes, qualité et évolution des données.

Si Orange a opté pour deux outils informatiques très différents, des éditeurs SAS et Xken, ses efforts en matière d'analyse prédictive portent essentiellement sur une seule et même cible, sa base de clients. De façon générale, les opérateurs cherchent à mieux cerner leurs clients, à réduire le taux de *churn* (c'est-à-dire le



AXEL SANE/FRANCE TÉLÉCOM

La croissance d'Orange dépend aujourd'hui avant tout de la consommation de ses abonnés, et moins de la conquête de nouveaux clients. Un constat qui rend la plate-forme d'analyse prédictive stratégique.



nombre d'abonnés qui partent à la concurrence une fois leur abonnement terminé) et à prévoir les consommations saisonnières à travers la détermination de profils d'utilisateurs. Ce qui permet de s'apercevoir, par exemple, que les gros clients se servent moins de leur portable l'été, tandis que le phénomène inverse se produit pour les autres.

Sur le plan technique, les problématiques proviennent en grande partie de la gestion des millions de clients. "La complexité de ce genre de projet augmente exponentiellement avec le volume des données à traiter", reprend Ar-

naud Muller. Avec de telles masses, les temps de réponse des requêtes dépassent fréquemment des normes acceptables. "Il faut dix minutes, par exemple, pour récupérer avec SAS Enterprise Miner des échantillons sur des bases de 20 millions de personnes." La qualité des données en entrée (celles qui sont recueillies dans les points de vente ou par les conseillers clients en particulier) s'avère également primordiale pour obtenir de bonnes analyses en sortie.

Dans la pratique, les deux outils sont complémentaires. "Au départ logiciel de statistique, SAS Enterprise Miner s'est

AVIS D'EXPERT

"Prévoir le comportement face à une action donnée"



Michel de Bollivier, chef de projet GRC chez Business & Decision

Dans une entreprise, les moteurs de l'analyse prédictive sont souvent les équipes du marketing stratégique et opérationnel. Il s'agit de se poser les bonnes questions. Quelle est l'utilité d'un projet de GRC analytique ? Recherche-t-on des scores *[outils de mesure du potentiel d'un échantillon de clients, NDLR]*, de la prédiction des réactions du client ? Il faut aussi savoir à quel stade du décisionnel l'entreprise se situe, de la gestion d'études sur la clientèle à une exigence de

retour sur investissement. De manière générale, avant de se lancer dans l'analyse statistique, il faut savoir ce que l'on veut en faire, afin de ne pas tourner en rond. En travaillant uniquement sur une base de données des clients, on ne peut pas tout comprendre d'eux. Mais l'on peut prévoir leur comportement face à une action donnée. Il faut cibler, c'est-à-dire se focaliser sur un aspect particulier du client, plutôt que de rester trop généraliste.

L'expérience de la téléphonie

mobile montre un des intérêts du datamining : il peut aider à trouver les caractéristiques qui décrivent le mieux un client. En outre, le datamining est un critère de plus pour les campagnes de marketing. Mais chaque secteur présente ses propres spécificités. Ainsi, la gestion de campagne des opérateurs diffère de celle des banques (intervention des agences régionales) ou des assurances (dans le cas notamment de la vente d'assurance décès).

MÉTHODOLOGIE : QUATRE CONSEILS CLÉS

1 Faire une cartographie des gisements de données

Objectif : éviter de se noyer dans le flot des informations de l'entreprise.

2 S'assurer de la qualité des données

Ceux qui se lancent dans un projet de GRC analytique croient dur comme fer en leurs données... qui s'avèrent souvent de piètre qualité.

3 Prendre patience

Ce n'est pas un petit projet. Mieux vaut ne pas promettre des résultats immédiats à la direction générale et aux directions opérationnelles.

4 Acquérir les compétences

Le datamining requiert une double compétence, à la fois informatique (pratique de la programmation) et généraliste (capacité d'analyser les résultats, de les rendre compréhensibles et de les résumer).

étendu au datamining. Il nécessite une certaine expertise et des connaissances en programmation, contrairement à l'outil de Xxen. Ce dernier est un module très spécialisé et non propriétaire, utile pour commencer à faire de l'analyse prédictive", évalue Michel de Bollivier. Complet, mais aussi complexe, Enterprise Miner sert "à la compréhension des phénomènes pour les problématiques approfondies", selon Arnaud Muller. Car, en toile de fond, "l'usage du mobile dépend du mode de vie et non des catégories socioprofessionnelles, alors que les achats dans un hypermarché sont liés à la structure familiale". Ces facteurs clés sont étudiés au travers du logiciel Enterprise Miner notamment. "Cet outil est en fait un gestionnaire de base de données relationnelle de gros volumes, parce que l'outil de dépouillement de statistiques a été bâti sur un système de base de données qui se greffe sur le système d'information de l'entreprise." Très présent dans les grands comptes français, SAS dispose d'un autre avantage : "Tous ceux qui font des études de statistiques en France apprennent sur des outils de cet éditeur", remarque Arnaud Muller. Ainsi, le temps consacré à la formation des équipes de l'informatique décisionnelle est quasiment nul.

Une autre application de la GRC analytique réside dans la construction de scores, pour laquelle l'opérateur se sert de l'outil K2R de Xxen. Le scoring consiste en un système de notation bâti

sur une importante quantité de données différentes (type de produits achetés, fréquence et montant des achats...). En téléphonie mobile, les opérateurs établissent ainsi des "scores de churn" (dont l'objectif est de déterminer la probabilité, à moyen ou long terme, qu'un client ne renouvelle pas son abonnement) ou de PCM, pour "programme changer de mobile" (probabilité qu'un client change d'appareil).

Une solution tout intégrée ?

Les différences des deux outils se retrouvent dans leur mise en œuvre. Si l'installation de K2R sur des postes autonomes a pris une semaine et a pu être réalisée en interne, il en va tout autrement pour l'outil de SAS. Le premier n'est d'ailleurs pas, contrairement au second, intégré au système d'information. L'analyse prédictive avec SAS Enterprise Miner est venue se greffer sur l'existant en GRC (un entrepôt de données séparé en deux datamarts). Lancé à la fin du deuxième semestre 2000, le déploiement d'Enterprise Miner s'est achevé fin 2001. Pour mener à bien le projet, Orange a fait appel à TMIS, une société de conseil externe qui l'a accompagné jusqu'à deux mois avant la livraison. Elle a participé notamment à la définition de l'architecture, ainsi qu'à la construction et à l'élaboration de la solution. L'intégration a été réalisée par SAS et supervisée par la société de conseil. Depuis la mise en production, Accentures s'occupe de la maintenance et des évolutions.

Quant aux perspectives du projet, elles s'organisent selon deux grands axes. A court terme, Orange vise "l'amélioration des performances grâce à l'installation d'un nouveau serveur". A plus long terme, une réflexion sur l'adoption des solutions tout intégrées de Tera-data est en cours. "Ce n'est pour l'instant qu'une piste de développement, nous ne sommes pas sûrs de nous engager", remarque Arnaud Muller. L'idée revient à mettre toutes les informations sur le même plan : toutes les bases de données seraient dans le même environnement et communiqueraient. "On saisit l'intérêt dans le domaine du décisionnel : récupérer des informations de multiples provenances.

Reste la question du retour sur investissement de l'analyse prédictive. "Les opérateurs nous assurent qu'ils obtiennent des résultats en la matière", confie Michel de Bollivier. Orange semble confirmer cette tendance. Même si Arnaud Muller tempère : "Le retour sur investissement est difficilement mesurable, mais l'analyse prédictive est une activité d'avenir chez Orange. Si l'on veut atteindre des objectifs de croissance, il nous faut travailler dans cette voie." ●

MARC DI ROSA

Zoom Grande distribution : sortir du marketing de masse

Devant une concurrence de plus en plus forte, notamment celle du hard discount, la grande distribution s'intéresse aux outils d'analyse prédictive.

C omble de l'ironie, la grande distribution ne connaît pas, ou très peu, ses clients. Bien qu'en contact direct avec eux, elle n'exploite pas les informations dont elle dispose. Alors qu'il s'agit d'une véritable mine d'or. De plus en plus consciente de cet enjeu, la grande distribution commence seulement à s'équiper d'outils d'analyse prédictive. Ceux-ci ayant pour vocation d'analyser les futurs comportements des clients. Car, aujourd'hui, rester compétitif réclame quelques efforts, notamment à l'heure où les chaînes traditionnelles font face à l'agressivité des hard discounters.

Reste à combler le retard pris par le secteur sur d'autres activités, comme la banque ou les opérateurs de télécommunications. "La grande distribution en est encore au stade de la mise en place des solutions d'analyse prédictive, alors que d'autres secteurs en sont à l'optimisation des processus utilisant ces outils", fait remarquer Hervé Mignot, directeur des opérations France de SPSS. Si pendant longtemps, la grande distribution ne s'est pas intéressée à l'analyse prédictive, ce n'est pas faute d'offres matures. "Seulement, pour la grande distribution, la priorité était ailleurs, remarque Renaud Smaghe, consultant chez Pierre Audoin Consultants (PAC). Elle se concentrait principalement sur la logistique et la traçabilité des denrées alimentaires."

Le ticket de caisse pour collecter des informations

Insuffisant toutefois pour se distinguer des concurrents. La seule arme efficace : connaître ses clients afin de les fidéliser. Ce que proposent en somme les outils d'analyse prédictive. "Néanmoins, leur efficacité dépend essentiellement de la richesse des informations collectées", estime Hervé Mignot. Malheureusement, la plupart du temps, les enseignes n'ont pas encore la capacité d'identifier chaque client passant en caisse. L'une des façons d'y parvenir consiste à offrir une carte de fidélité. "Au-delà de la relation avec le client qu'établit ce type d'outil, c'est aussi un bon moyen de suivre l'acheteur et de déterminer ses habitudes", avoue Hervé Mignot. "Néanmoins, la carte de fidélité a ses limites et n'est pas



Après avoir bombardé leurs clients de messages non ciblés, les distributeurs tentent de personnaliser leur marketing, par exemple avec des cartes de fidélité.

fiable à 100 %, car elle n'est pas utilisée par tous les acheteurs", constate Renaud Smaghe. D'autres façons de procéder peuvent être employées. Par exemple, en s'appuyant sur le ticket de caisse. Dans ce cadre, il ne s'agit plus d'identifier les profils de chaque client,

mais plutôt de concevoir des panels de marchandises vendues. Et, par la même occasion, de mieux négocier le prix et le choix des produits. Mais les bénéfices de tels outils ne s'arrêtent pas là. Ils peuvent aussi être employés pour optimiser la disposition des produits

AVIS D'UTILISATEUR

Stéphane Marchet,
directeur marketing de Conforama

"Un volume considérable de données à gérer"



En 2003, en France, un lit sur trois et un meuble sur cinq ont été achetés dans nos magasins. Mais rester compétitif dans le secteur de la distribution demande des services de plus en plus professionnels. Concrètement, cela suppose de connaître son client pour mieux anticiper ses besoins. C'est aussi indispensable lorsqu'il s'agit de négocier avec les industriels. Le projet d'analyse prédictive n'a pu être mené à bien qu'au prix de longs efforts et de beaucoup de persévérance. Nombre de difficultés viennent du secteur lui-même. D'abord, parce que le volume des données à gérer est considérable.

Chez Conforama, pas moins de 20 000 références produits sont traitées par magasin, soit environ 50 000 au niveau de l'enseigne. D'autre part, le nombre de clients est important. Autre difficulté, l'éclatement des points de vente qui conduit à des problèmes de régionalisme. Pour y remédier, il faut mettre en place des processus d'exploitation des informations adaptés à la fois au niveau régional et au niveau national. Enfin, facteur déterminant si l'on souhaite des analyses fiables, les données doivent être d'une qualité irréprochable et homogène.

QUELQUES PRÉCURSEURS

Botanic Ce magasin spécialisé dans le jardinage a choisi SPSS Base pour analyser le comportement des 200 000 adhérents du Club Botanic. Le projet a porté sur trois axes principaux : identification des profils de clients, détermination des associations de produits et évaluation des performances des opérations commerciales.

Sears Cette chaîne américaine de grands magasins a choisi Kxen Analytic Framework pour mieux cibler ses clients lors de ses campagnes de marketing direct. Lancé à la fin de l'été 2004, ce projet permet déjà à l'équipe marketing de développer des modèles qui identifient les clients les plus à même de réaliser des achats.

Conforama Le spécialiste de l'ameublement a choisi Enterprise Miner de SAS Institute pour mieux identifier les achats des clients. Lancé en 1999, le projet a permis à des enseignes du groupe de segmenter leur portefeuille de clients. Le projet va être étendu à l'ensemble des magasins.

dans les rayons. Par exemple, en étudiant sur un échantillon donné les linéaires les plus représentés dans le panier d'achats. En fonction du résultat obtenu, les rayons ou les produits seront agencés différemment dans le magasin. Autre usage possible : l'optimisation des prix. Par exemple, en positionnant l'enseigne par rapport à ses concurrents. Ou bien en mesurant l'impact que provoque l'affichage de tel prix sur tel type d'article.

Malgré les nombreux avantages qu'elles présentent, les solutions d'analyse prédictive tardent à percer. "Parmi les principales raisons, des difficultés d'ordre légal", signale Renaud Smaghe. Notamment en ce qui concerne les bases de données client. Celles-ci comportent des informations pouvant relever du domaine privé. "Acelas'ajoutent des problèmes organisationnels", poursuit le consultant. Et des difficultés techniques. En effet, pas de projet d'analyse prédictive sans assurance sur la qualité des données présentes dans la base. Concrètement, cela passe par l'homogénéisation du système d'information. Une tâche ardue dans la distribution, dont les implantations sont le plus souvent très dispersées, une enseigne comprenant généralement un siège, plusieurs magasins et une centrale d'achats. Et le coût élevé de ce type de projets n'est pas là pour accélérer les investissements. ●

SANDRINE CHILOTTI

Prospective Le multidimensionnel frappe à la porte des DRH

Les DRH ont des comptes à rendre. Au-delà du reporting légal, elles doivent aussi analyser, prévoir et simuler. Des solutions décisionnelles émergent, mais certaines entreprises ont déjà mis en place leurs propres outils.

Mes salariés sont-ils équitablement rémunérés par rapport au niveau moyen de rémunération sur le marché ? Les coûts salariaux augmentent-ils trop vite et pourquoi ? Quel est le taux de mobilité interne de l'entreprise ? Quelles ont été les démarches de recrutement les plus rentables ? Dans leur version la plus classique, les systèmes d'information exploités par les directions des ressources humaines (SIRH) ne délivrent pas de réponse directe à ces questions. Pourtant, "ces directions sont de plus en plus soumises à une logique de résultat, reconnaît Stanislas de la Foye, enseignant au Ciffop (Centre interdisciplinaire de formation à la fonction personnel) et spécialiste des SIRH. On leur demande de chiffrer, quantifier, prouver. Le DRH doit redoubler d'imagination pour trouver le bon indicateur, le communiquer aux bonnes personnes et se placer dans une perspective d'action et de prospective qui va bien au-delà de ses obligations légales de reporting."

Les entreprises se sont équipées d'outils pour satisfaire à ces obligations légales. Chaque année, elles publient un bilan social avec des rapports détaillés sur leurs effectifs, la formation professionnelle, les rémunérations et les accidents du travail. Certaines ont déjà bâti des infocentres pour rassembler des données gérées par des applications RH distinctes (paie, gestion des temps, de la formation, etc.).

Des informations qualitatives aux données d'analyse

Mais la plupart n'ont pas encore d'outils pour des investigations plus poussées. Des applications qui, par exemple, permettraient au DRH d'un groupe chapeautant une centaine de filiales, et autant de DRH départementales, d'effectuer des analyses multidimensionnelles sur les données agrégées de l'ensemble des salariés. Ou de réaliser du pilotage à l'aide de tableaux de bord et d'indicateurs clés de performances.

"Après le reporting légal, l'analyse de la masse salariale et la gestion prévisionnelle, l'enjeu consiste maintenant à transformer les informations qualitatives en données d'analyse, commente Charles Parat, directeur conseil pour le décisionnel chez Mi-



La solution de Cognos et HR Access. Les grands du décisionnel et des RH ont développé des produits spécifiques pour répondre aux besoins des directions des ressources humaines.

cropole-Univers. Par exemple parvenir à effectuer des mesures sur les entretiens d'évaluation des salariés, qui sont des éléments essentiels dans la conservation du capital humain."

Des solutions répondent déjà à ce besoin. "Les acteurs traditionnels du do-

maine ont développé pour les DRH des applications spécifiques, comme ils l'avaient fait pour les services financiers et les relations avec les clients", indique Stanislas de la Foye. Il cite des éditeurs comme Cognos, Business Objects, Hyperion ou SAS Institute qui,

souvent, s'associent à des cabinets de conseil, experts dans les RH (comme EasyWay) ou dans les outils d'analyse (Business & Decision, Décisionnel, Micropole-Univers, Keyrus, Aston...).

Les éditeurs de progiciels intégrés, spécialisés dans les RH (ADP-GSI, Sopra, HR Access...) ou plus généralistes (PeopleSoft, SAP, Oracle, Arès...), ont leurs propres offres. Celles-ci s'appuient aussi sur les acteurs du décisionnel. HR Access, par exemple, a intégré les outils de pilotage Metrics Manager de Cognos. Et Sopra travaille avec Business Objects. "Nous proposons des univers BO paramétrés par notre progiciel Pléiades, avec des exemples de requêtes sur les données RH", confirme Véronique Montamat, directrice marketing de Sopra. SAP et PeopleSoft exploitent aussi les offres les plus courantes (BO, Cognos...). Sur ce marché plutôt encombré, un acteur de niche, le Franco-Néerlandais Allshare, dispose d'une solution reconnue pour la simulation de la masse salariale et les budgets des frais de personnel. Le Français Enablon, lui, propose du pilotage et du reporting opérationnel sur des données sociales et environnementales, dans une perspective de développement durable.

La concurrence des solutions maison

L'offre est fournie. L'ampleur de la demande est moins évidente. Même si les entreprises se sont aperçues que leur SIRH ne répondait pas forcément à leurs attentes de pilotage. Mais pour passer à la vitesse supérieure (analyse, prévision, simulation), elles ne sont pas toujours tentées par une solution standard. En particulier celles qui exploitent un infocentre bâti avec des équipes internes. Celles-là ont développé une couche de reporting qui fonctionne parfaitement et ne sont pas prêtes à investir lourdement dans un nouvel outil et un nouveau modèle de données.

"Les secteurs d'activité qui investissent sont ceux qui disposent traditionnellement de budgets informatiques importants, le secteur bancaire, notamment, remarque Stanislas de la Foye. Mais les besoins ne sont pas moindres dans l'industrie ou la fonction publique où l'on voit en ce moment des projets émerger. Il s'agit surtout d'une question de budget."

MARYSE GROS

MISE EN ŒUVRE

Une mine d'or à exploiter pour le Club Med



Dans les villages de vacances, une simple connexion Internet donne accès de façon sécurisée à des données salariales ciblées. Les analyses plus stratégiques se font au siège où la DRH du Club Med peut, par exemple, calculer son turnover managérial et opérationnel suivant une dizaine d'axes. "Nous n'en sommes qu'au début, nous avons de nombreux sujets d'analyse à mettre en place, une mine d'or à exploiter", se réjouit Jean-Yves Corre, responsable informatique back-office.

Ce projet de reporting et de pilotage s'est inscrit dans une refonte plus globale du SIRH, avec la mise en production en 2002 de l'offre de PeopleSoft (la paie reste gérée par GXP d'ADP-GSI). Fin 2002, l'entreprise retient la solution HCM (Human Capital Management) de SAS Institute, pour analyser l'ensemble de ses données RH.

"De nombreuses structures d'analyse multidimensionnelles sont

prédéfinies dans HCM, mais notre DRH a préféré conserver ses propres définitions, précise Jean-Yves Corre. Nous avons donc élaboré avec elle les indicateurs de pilotage qui lui convenaient." Pour les requêtes ad hoc, la DSI utilise SAS Enterprise Guide.

Club Med exploitait déjà les outils de SAS pour le reporting commercial. La DSI a voulu mutualiser les compétences acquises. Elle a également apprécié le déploiement en mode Web qui réduit les coûts de maintenance.

M. G.

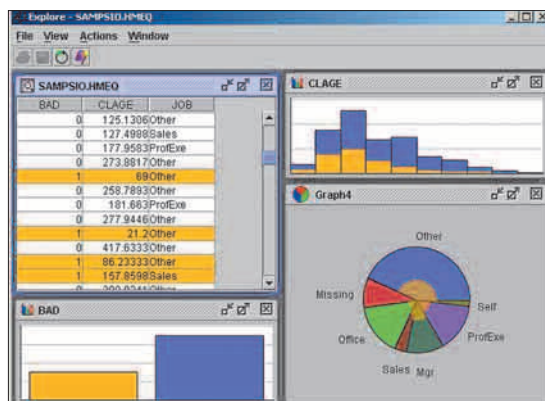
LA SOLUTION ADOPTÉE

SAS HCM, environnement IBM AIX, serveur d'applications BEA WebLogic 7.0 ; SAS HCM s'appuie sur des datamarts (recélant les données sur les postes, les affectations, etc.) qui puisent leurs données dans un entrepôt de données lui-même alimenté par le SIRH PeopleSoft et la gestion de la paie de ADP-GSI.

Décryptage Quand le décisionnel se projette dans l'avenir

Grâce en partie à la mise en œuvre des projets de gestion de la relation client, et surtout des progiciels de gestion intégrés, qui ont rendu les données plus facilement exploitables, les entreprises s'exercent à l'analyse prédictive à la demande des métiers. Objectif : anticiper les comportements de leurs clients et répondre en conséquence.

Les dix dernières années ont vu la mise en place progressive des outils décisionnels basés sur des systèmes dits de "business intelligence". L'objectif de ces applications : fournir quantité de rapports et tableaux de bord construits et modélisés, basés sur une valeur essentielle de l'entreprise et dont le potentiel est resté longtemps insoupçonné, ses données. Isabelle Carcassonne, directrice marketing et communication chez Cognos Europe du Sud, explique : "Jusqu'à présent, et c'est ce que ces outils permettaient de faire, les sociétés regardaient dans le passé et tentaient d'analyser ce qui s'était produit pour prendre leurs décisions. C'est l'effet rétroviseur." Avec l'analyse prédictive, c'est une étape supplémentaire qui est franchie : les phares s'allument. On tente désormais de prévoir le comportement des clients dans tel ou tel cas de figure. On est passé de la prise de décision à partir des informations tirées de l'analyse de l'historique (comprendre) à une modélisation prévisionnelle anticipant sur les comportements futurs et permettant donc une meilleure réactivité (prédire). "Ce type de projet d'analyse prédictive est néanmoins réservé à des entreprises qui ont déjà mis en œuvre des outils de reporting et



L'interface du logiciel Enterprise Miner de SAS Institute. Quand ils mettent en œuvre de gros volumes de données, l'usage des outils d'analyse prédictive reste un domaine de spécialistes.

qui possèdent une bonne maîtrise de leurs flux de données", souligne Hervé Chapron, directeur business intelligence chez Keyrus.

Une demande des directions fonctionnelles

Ces projets émanent essentiellement de la demande des directions fonctionnelles, qui disposent de l'expertise technique, de la connaissance métier, et qui sont de plus "poussées par la concurrence sous l'effet moutons de

Panurge", s'amuse Charles Parat, directeur du conseil autour de la business intelligence chez Micropole Univers. Benoît Contamine, directeur du pôle décisionnel chez Valoris, souligne : "Par exemple, dans la grande distribution, la tendance est de passer du marketing de masse au marketing individualisé. Ce choix stratégique a été enclenché par les directions fonctionnelles." Le projet fonctionne ensuite en tandem avec les directions informatiques. Marie-Claude Santon, directrice marketing et communication chez SAS, pré-

cise : "Parce que la gestion des données est là et que leur préparation est essentielle. Et puis, pour les DSI, ce type de projet, avec un retour sur investissement mesurable, est plutôt valorisant." Même s'ils ne le considèrent pas encore comme un enjeu majeur et disposent encore de peu de compétences sur le sujet. Sans oublier dans la boucle les équipes de statisticiens, dépendant des directions fonctionnelles et très en ligne avec les métiers, qui travaillent sur les modèles.

Fidéliser les clients

Mais si elle vient des métiers, l'analyse prédictive est-elle directement utilisable par ses demandeurs ? Deux écoles s'affrontent. "S'il y a de gros volumes de données et un jeu complexe, l'analyse prédictive reste un domaine de spécialiste", affirme Roger Haddad, président de l'éditeur spécialisé Kxen. Ce que confirme Isabelle Carcassonne : "L'analyse prédictive n'est plus un travail d'informaticien, mais ce ne sont pas des outils à mettre entre toutes les mains." Le son de cloche est différent chez Ilog. "Depuis quatre ans, des efforts ont été accomplis pour rendre ces produits accessibles aux utilisateurs non techniques", argumente Nicolas Robbe, directeur de l'équipe business solution chez Ilog. Même tonalité chez SAS, où des équipes d'ergonomes travaillent sur les produits pour qu'ils soient utilisables par le plus grand nombre.

Parallèlement à la demande des métiers, deux leviers essentiels ont permis à ce type d'applications de s'imposer dans les entreprises. Côté marché tout d'abord : la pression concurrentielle s'accroissant, notamment sur les secteurs liés aux particuliers, une réactivité accrue en termes d'offre est attendue. Les clients, fréquemment sollicités par la concurrence, se montrant de plus en plus infidèles, il faut créer et mettre en œuvre de nouveaux programmes ou produits pour les fidéliser ou en attirer de nouveaux. Hervé Chapron ajoute : "Conquérir un nouveau client coûte sept fois plus cher que d'en garder un. Les entreprises cherchent donc par tous les moyens à conserver ceux qu'elles ont déjà et à les développer." Il devient donc indispensable de proposer à ses clients "au plus juste prix" le type

Des projets par îlots pour des livrables rapides

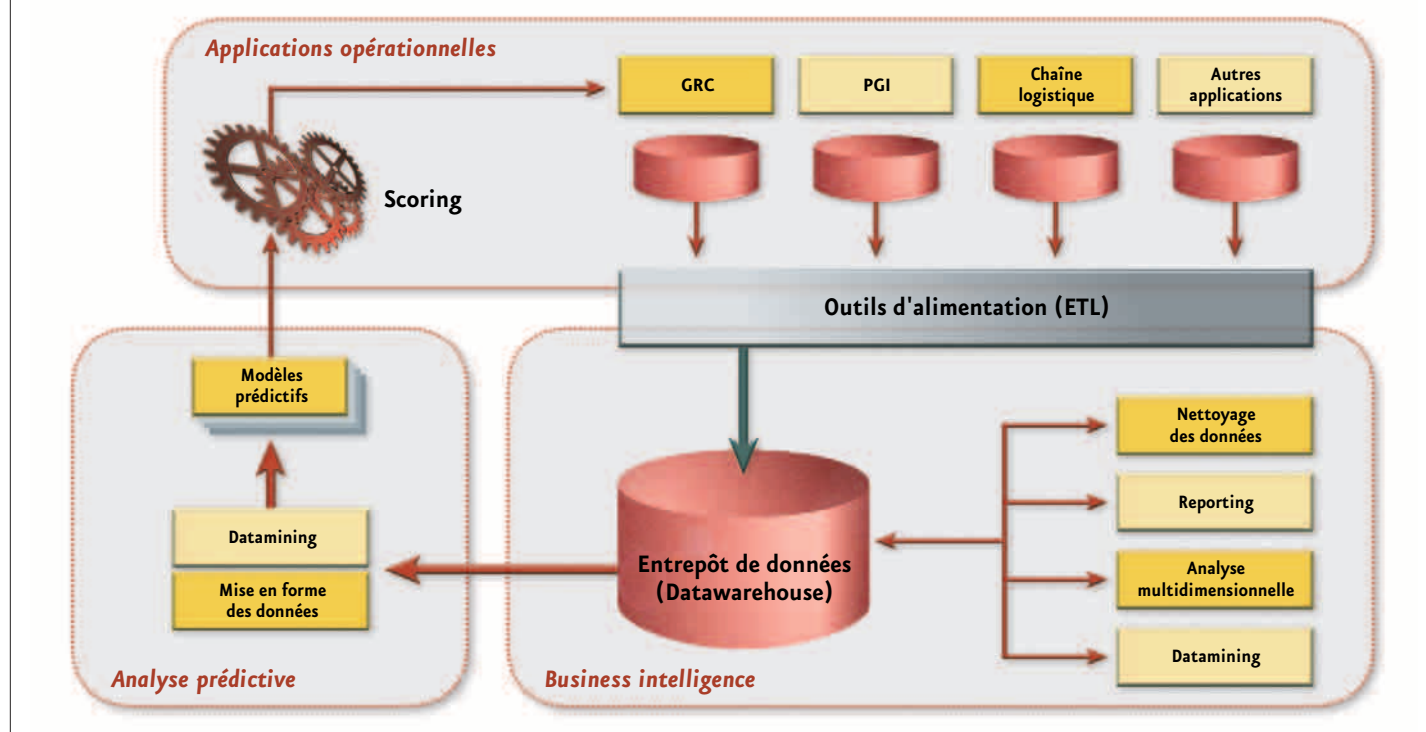
Si les projets d'analyse prédictive exigent que l'on dispose d'un volume important de données et d'un historique, leur mise en œuvre est le plus souvent découpée en îlots. Car sur cette masse d'informations à disposition, tout n'est pas directement exploitable. Le périmètre de l'application doit également être soigneusement cadré. Pour une application donnée, il faut ainsi identifier les critères intéressants et déterminer des modèles qui soient robustes et reproductibles. Tout l'art consiste à passer d'une problématique statistique à une problématique métier.

"Il faut faire attention à ne pas avoir trop d'ambition

au démarrage, affirme Charles Parat, de Micropole Univers. Il n'y a pas de projet décisionnel big bang qui marche. Il est nécessaire de cadencer et de sérier le projet pour éviter l'effet tunnel et de commencer petit, même si l'on voit grand et que les investissements sont importants." Comme les demandes viennent des directions fonctionnelles, ces dernières exigent des livrables rapides, car c'est le cœur de leur métier. Et puis, le concept de base étant d'améliorer la réactivité de l'entreprise, c'est le projet lui-même qui doit rester réactif, car, fréquemment, les demandes changent durant son déroulement. "Il ne faut laisser personne d'autre que la direction fonctionnelle

piloter le projet, ajoute Hervé Chapron, directeur business intelligence chez Keyrus. Et c'est aux métiers de travailler sur la recette." D'où l'importance des prototypes et de l'utilisation d'une technologie permettant de démontrer le bien-fondé de la conception. Nicolas Robbe, de la société Ilog, explique : "Les utilisateurs métier n'ont pas l'approche de la technologie, mais ils veulent voir un résultat tangible. Voir pour croire." "La plupart du temps, afin de démontrer la valeur du projet, on travaille sur une partie des données pendant trois mois. Il faut absolument être capable de montrer rapidement des résultats", conclut Marie-Claude Santon, de SAS. S. H.

Suite page 20 ➤



Source : Ilog/LMI

900 millions de dollars

C'est le chiffre d'affaires du marché mondial de l'analyse prédictive en 2003, selon le cabinet d'études IDC.

+8 %

C'est le taux de croissance enregistré par la GRC prédictive (ou analytique) en 2003, selon le cabinet d'études PAC. Par opposition, la GRC opérationnelle enregistre une baisse de 6,5 %.

LES SECTEURS CONSOMMATEURS

La banque Un secteur pionnier dans la mise en place de systèmes d'analyse prédictive, notamment pour la délivrance des crédits à la consommation. Aujourd'hui, ces outils sont orientés vers la relation client au sens large (vente croisée ou vente incitative) et vers la détection des fraudes.

Les assurances L'analyse prédictive est utilisée pour l'estimation des risques et la proposition de contrats au plus juste prix. Le secteur des mutuelles devient également consommateur.

Les opérateurs de télécommunications Un des domaines les plus actifs, car la concurrence y fait rage. L'analyse prédictive vise à y limiter le risque de départ des clients (*churn*).

La grande distribution Ce secteur analyse bien sûr ses données clients, mais c'est moins vital que dans d'autres secteurs. Le marketing de masse produit encore des résultats.

Les secteurs qui démarrent L'industrie pharmaceutique, le service public, l'industrie (notamment pour le contrôle qualité), la santé.

CES ENTREPRISES FRANÇAISES QUI SE SONT LANCÉES

SOFINCO

Plate-forme : SAS
Nombre d'utilisateurs : une vingtaine de personnes
Calendrier : six mois en 1995 pour mettre en œuvre l'analyse prédictive sur les risques, un mois en 1998 pour son extension au marketing.
Objectif : sécuriser le risque de crédit et améliorer le taux d'acceptation des dossiers. Côté marketing, lancer un plus grand nombre de campagnes.
Montant : environ 200 000 euros (construction du système analytique).

TEREOS

Plate-forme : Easy Decision
Nombre d'utilisateurs : 5
Calendrier : début du projet fin 2001. Durée d'un an environ (de la phase de réflexion à la mise en production).
Objectif : exploiter la totalité des données de l'entrepôt de données (auparavant 15 % seulement des données étaient utilisées).
Montant : environ 58 000 euros.

LOUIS VUITTON

Plate-forme : SPSS
Nombre d'utilisateurs : 3
Calendrier : un mois au total, janvier 2002 (analyse des besoins et installation spécifique).
Objectif : modéliser le comportement du client et analyser des données sous de multiples dimensions.
Montant : environ 12 000 euros (conseil et développement sur mesure compris).

NATEXIS ASSURANCE

Plate-forme : SPSS
Nombre d'utilisateurs : 2
Calendrier : 40 jours-hommes d'avril à septembre 2002.
Objectif : gains en productivité et en rentabilité dans le domaine du marketing.
Montant : entre 150 000 et 180 000 euros.

► d'offre qui les incitera à rester alors qu'ils n'ont pas encore réfléchi à leur départ. C'est ce qu'on appelle le *churn* dans le secteur des opérateurs de télécoms ou "l'attrition" dans le domaine bancaire.

Les données, un capital à exploiter

Côté technique ensuite. Contrairement à ce que l'on pourrait penser, le développement de l'analyse prédictive n'est pas fondamentalement lié à celui des applications de GRC (gestion de la relation client), qui permettent de collecter nombre d'informations sur les clients. "L'analyse prédictive en tant que technique de décision existe depuis longtemps dans les entreprises, historiquement sur le secteur du crédit à la consommation et à travers les solutions de scoring pour la gestion du risque", précise Benoît Contamine. "C'est plutôt l'arrivée à maturité des PGI qui, en offrant une vision intégrée et centrale des données clients, en lieu et place de modules indépendants, a participé à son développement", ajoute Xavier Gazay, associé en charge des lignes de services GRC pour Accenture. Finalement, les applications de GRC ne représentent qu'un canal d'accès de plus à ces précieuses données, stockées depuis longtemps par l'entreprise par différents canaux et sous de multiples formes, mais dont la valeur en tant que "capital" prend tout son sens aujourd'hui avec leur exploitation par ces applications décisionnelles.

Les données, quantitatives autant que qualitatives, restent d'ailleurs tech-



Marie-Claude Santon, directrice marketing et communication chez SAS : "Pour les DSI, ce type de projet, avec un retour sur investissement mesurable, est plutôt valorisant."

Xavier Gazay, associé en charge des lignes de services GRC pour Accenture : "La grosse difficulté réside dans l'industrialisation de l'utilisation des données. Aujourd'hui c'est très souvent du bricolage."



iquement au cœur de la problématique liée à la mise en œuvre de ce type de projet. Pour envisager l'utilisation d'un système d'analyse prédictive, l'entreprise doit impérativement disposer d'un volume et d'un historique important, qui garantissent l'efficacité de la mise en œuvre de modèles. Ensuite, il faut qu'elles soient exploitables. "La grosse difficulté, commente Xavier Gazay, réside dans l'industrialisation

ILS ONT DIT



Isabelle Carcassonne, directrice marketing et communication de Cognos Europe du Sud : "L'analyse prédictive n'est plus un travail d'informaticien, mais ce ne sont pas des outils à mettre entre toutes les mains."

de la manière dont on utilise ces données pour faire de l'analyse prédictive. Aujourd'hui c'est très souvent du bricolage. De plus, ce type d'application travaille sur des indicateurs et pas sur des données normalisées comme le font les systèmes transactionnels." La partie planification, qui consiste à repérer et récupérer les données nécessaires, fait ainsi figure d'étape la plus cruciale et la plus longue d'un projet d'analyse prédictive. "Il faut récupérer les valeurs manquantes, homogénéiser les données, créer un dictionnaire d'information, le datamart, ou un référentiel unique, ce qui est souvent difficile", ajoute Marie-Claude Santon. Et puis donner une définition, du chiffre d'affaires par exemple, qui peut être différente entre le directeur financier et le directeur commercial. C'est la plus grosse partie du projet." Constatation confirmée par Xavier Gazay : "Ce type de projet est constitué d'environ 80 % de manipulation de données et de 20 % de modélisation à proprement parler, cette dernière partie n'étant pas la plus compliquée."

Si, en amont, il est nécessaire de disposer de données exploitables, et même, parallèlement, de remettre à plat et contrôler les processus afin d'encadrer ce cercle vertueux (sensibilisation des opérateurs à la qualité des données par exemple), il faut aussi en aval définir un périmètre d'exploitation. Et s'assurer que les modèles statistiques sont déployables réellement dans un contexte opérationnel. "Indépendamment de la qualité du modèle, il faut industrialiser la propagation de l'information, et ne pas en sous-estimer les difficultés techniques", conclut Xavier Gazay. ●

SOPHIE HUET, AVEC SANDRINE CHILOTTI

CHOISIR LA PLATE-FORME EN FONCTION DU VOLUME DES DONNÉES

On peut répartir les logiciels de datamining en trois familles, suivant le volume de données traitées :

• **Bases de données de gros volume.** C'est-à-dire des millions de données. Il s'agit de produits comme Insightful Miner de Insightful, Statistica de Statsoft ou Intelligent Miner d'IBM. Ces solutions couvrent en général plusieurs techniques d'analyse pour traiter les informations. Elles reposent sur l'association de plusieurs algorithmes (réseaux de neurones, arbres de décision, règles d'association ou réseaux bayésiens). Ces solutions peuvent également fonctionner en mode client-serveur.

• **Bases de données de petit volume.** Soit quelques milliers de données. Cela concerne des logiciels comme Spad de Decisia, Neuralware de Predict ou Alice d'Isoft. Ces outils sont en général plus faciles à installer et moins coûteux que les précédents. Mais ils ne combinent pas plusieurs types d'algorithmes.

• **Bases de données de volume moyen.** Environ une centaine de milliers de données. Ces solutions peuvent faire appel à un ou à plusieurs algorithmes. Par exemple, Answer Tree de SPSS n'a recours qu'aux arbres de décision alors que Darwin d'Oracle utilise à la fois les réseaux de neurones, les arbres de décision et la typologie.

Glossaire

Analyses prédictives : elles déterminent statistiquement, à partir des données historiques du client, les comportements futurs de celui-ci.

Datamining : technique de recherche et d'analyse de données qui permet d'établir des tendances ou des corrélations cachées parmi des masses d'informations.

Entrepôt de données (datawarehouse) : base dans laquelle les données sont centralisées. Elle est conçue de manière à ce que les personnes intéressées aient accès rapidement à l'information stratégique dont elles ont besoin.

Segmentation : découpage d'une population ou d'une série de données pour identifier des sous-classes pertinentes, à partir d'un ou de plusieurs critères (géographiques, comportementaux...).

Datamart : base de données orientée métier et propre au décisionnel.

Echantillon : ensemble de données tirées, aléatoirement ou non, de l'entrepôt de données et permettant d'effectuer des tests et des recherches sur des volumes restreints.

L'OFFRE D'ANALYSE PRÉDICTIONNELLE

Editeur	Produit	Méthodes de modélisation
PETITS VOLUMES (ENVIRON 10 000 LIGNES DANS LA BASE)		
Elseware	Previas	Réseaux de neurones, arbres de décision, réseaux bayésiens
Grimmersoft	StartMiner	Réseaux de neurones, arbres de décision
Isoft	Alice	Arbres de décision
Netral	NeuroOne	Réseaux de neurones
Neuralware	Predict	Réseaux de neurones
Wizwhy	Wizsoft	Détection d'associations
VOLUMES MOYENS (ENVIRON 100 000 LIGNES)		
Angoss	KnowledgeSeeker	Arbres de décision
Bayesia	BayesiaLab	Réseaux bayésiens
Decisia	Spad	Arbres de décision
Oracle	Darwin	Typologie, réseaux de neurones, arbres de décision, réseaux bayésiens
Salford Systems Cart		n.c.
SPSS	Answer Tree	Arbres de décision
GROS VOLUMES (QUELQUES MILLIONS DE LIGNES)		
IBM	Intelligent Miner	Réseaux de neurones, arbres de décision
Insightful	Insightful Miner	Arbres de décision, segmentation, régression...
Kxen	Kxen Analytic Framework	Règles d'association, régression, K-means supervisé (segmentation)
SAS Institute	Enterprise Miner	Règles d'association, régression, K-means supervisé (segmentation)
SPSS	SPSS Base	Régression
SPSS	PredictiveMarketing	Réduction de règles, arbres de décision
SPSS	Clementine	Réseaux de neurones, régression, arbres de décision, induction de règles
StatSoft	Statistica Data Miner	Réseaux de neurones