

ÉTUDE TECHNIQUE

# Analyse prédictive

## L'essor et la valeur de l'analyse prédictive dans la prise de décisions

« *Donnez-moi un point d'appui et un levier et je soulèverai le monde.* » — Archimède, 250 av. J.-C.

Au cours des dernières années, l'analyse prédictive est passée de technique avantgardiste peu répandue à une arme concurrentielle dont la portée se développe rapidement. L'adoption croissante de l'analyse prédictive est alimentée par des tendances convergentes : le phénomène des données massives, l'amélioration des outils d'analyse de données et un afflux constant de réussites démontrées dans le cadre de nouvelles applications. Aujourd'hui, l'analyste moderne dirait sûrement « Donnez-moi suffisamment de données et je prédirai n'importe quoi. »

La façon dont les modèles prédictifs apportent de la valeur ajoutée constitue un concept très simple : ils permettent de prendre des décisions plus éclairées, de façon plus rapide et à moindre coût. Ils améliorent les décisions humaines en augmentant leur efficacité et leur efficacité. Dans certains cas, ils permettent même d'automatiser un processus complet de prise de décision.

L'évaluation du crédit constitue un exemple classique d'analyse prédictive au travail. Les modèles de risque de crédit, utilisant l'information de chaque demande de prêt pour prédire les risques de perte, ont été conçus et remaniés au fil des ans, à un point tel qu'ils jouent désormais un rôle indispensable en matière de décisions de crédit. Le secteur du crédit à la consommation tel que nous le connaissons aujourd'hui ne pourrait pas opérer sans modèles de prévision des risques de crédit.

Les banques l'ont rapidement adoptée, mais l'éventail d'applications et d'organisations employant l'analyse prédictive avec succès s'est rapidement étendu :

- **Ventes et marketing direct.** Les pistes provenant du site web d'une entreprise peuvent être évaluées afin de déterminer la probabilité de conclusion d'une entente et d'établir son niveau de priorité. Les campagnes peuvent cibler les candidats les plus susceptibles de répondre.
- **Relations avec la clientèle.** Les caractéristiques et le comportement du client permettent de prévoir le niveau d'attrition (p. ex. : contrats de téléphonie mobile et cartes de crédit). Les modèles d'attrition ou de « désabonnement » aident les entreprises à établir des stratégies pour réduire les variations de taux d'attrition au moyen de communiqués et d'offres spéciales.
- **Optimisation des prix.** À l'aide d'une quantité suffisante de données, la relation entre la demande et le prix peut être modélisée pour n'importe quel produit puis utilisée pour établir la meilleure stratégie en matière de prix. L'analyse de prix et la gestion des revenus sont fortement employées dans les secteurs du transport aérien, du loisir et de l'amusement, des biens de consommation emballés et du secteur bancaire de détail et commencent à se répandre dans de nouveaux domaines comme les routes à péage et le commerce électronique de détail.
- **Résultats en matière de santé.** Les modèles liant les symptômes aux traitements sont de plus en plus répandus auprès des fournisseurs. Par exemple, un modèle peut prédire la probabilité qu'un patient présentant certains symptômes soit en fait victime d'une crise cardiaque, aidant ainsi le personnel des urgences à déterminer le traitement requis et le niveau d'urgence.
- **Fraude à l'assurance.** Plusieurs types de fraudes ont une approche prévisible et peuvent être identifiés à l'aide de modèles statistiques, ce qui aide la prévention, les enquêtes après fraude et le recouvrement.
- **Paiements indus d'assistance sociale et fraude.** Des paiements Indus d'assistance sociale en matière de santé, de bien-être, de chômage, de logement et autres sont parfois effectués, gaspillant l'argent des contribuables et réduisant la disponibilité des ressources pour les gens qui en ont réellement besoin. Des modèles semblables à ceux employés pour contrer la fraude à l'assurance peuvent aider à prévenir et à récupérer ces pertes.
- **Perception d'impôt.** Plusieurs cas de perception d'impôt additionnel (pour des non-déclarants, des sous-déclarations et des remboursements gonflés) peuvent être décelés. Le *Internal Revenue Service* (IRS) et plusieurs gouvernements d'État emploient des modèles de perception d'impôt qu'ils améliorent constamment.
- **Prévision et prévention de la criminalité, de la violence familiale et du terrorisme.** En plus de lier les techniques d'analyse aux enquêtes de crimes, les modèles prédictifs permettent de déterminer les situations à haut risque et les mesures préventives à employer.

## CONCEVOIR DES MODÈLES PRÉDICTIFS EFFICACES

Les modèles prédictifs requièrent des données. L'élaboration, l'évaluation et la mise au point de ces modèles requièrent des données décrivant : 1) les renseignements connus au moment de la prévision, et 2) les résultats potentiels. Par exemple, pour développer un modèle de risques de crise cardiaque propres aux patients se présentant à l'urgence, nous devons disposer de données décrivant les symptômes des patients à leur arrivée, et les résultats après examen (souffraient-ils réellement d'une crise cardiaque?). La capacité de générer des données à l'aide de ces caractéristiques constitue un facteur essentiel à la réussite d'une application de modèle prédictif.

## À PROPOS DE CETTE ÉTUDE

Comme ses applications concrètes ne cessent d'augmenter, l'analyse prédictive est à l'heure actuelle en plein essor. Les modèles prédictifs permettent de prendre de meilleures décisions, offrent une cohérence accrue et contribuent à réduire les coûts.

Les modèles prédictifs contribuent à apporter une valeur ajoutée considérable dans une foule de domaines, notamment le marketing, le maintien de la clientèle, l'optimisation des prix et la prévention de la fraude, et la liste ne cesse de s'allonger.

Ce document présente la méthode de conception des modèles, les situations propices à leur application, le calcul du rendement du capital investi, les tendances en matière de modèles prédictifs et plus encore. Grâce à l'analyse prédictive, les organismes gouvernementaux et les entreprises peuvent tirer une valeur accrue de leurs données, faciliter leur prise de décision et obtenir un avantage concurrentiel accru.

Les techniques statistiques comme la régression linéaire et les réseaux RNA, sont ensuite appliquées pour cerner les signes précurseurs et calculer les modèles actuels. Des logiciels de l'institut SAS, du SPSS d'IBM et de la trousse d'outils en code source libre R sont souvent employés à cette étape de l'analyse de modélisation.

Après l'assemblage des données, les analystes peuvent trouver jusqu'à 20 facteurs prédictifs connus pour chaque patient (dans notre exemple de salle d'urgence) et leur assigner des niveaux d'incidence à l'aide de logiciels statistiques (p. ex. : 50 points pour une pression sanguine anormalement basse). Le logiciel statistique utilise des algorithmes pour optimiser le modèle de facteurs d'incidence pour que leur combinaison offre les prédictions les plus précises possible à l'aide des données disponibles.

Les résultats obtenus, combinés à l'ensemble des facteurs et leurs niveaux d'incidence, constitueront un indicateur de risques efficace pouvant être utilisé comme critère de décision avec d'autres règles pour le traitement du patient. Les résultats seront non seulement mis en corrélation avec le risque de maladie cardiaque, mais pourront également être calibrés pour établir une relation mathématique. Cet avantage crucial permet de transformer un risque de type « inconnu-inconnu » en un risque « connu-inconnu », ou en d'autres termes, en risque calculé.

### APPLICATIONS DE QUALITÉ SUPÉRIEURE POUR LA MODÉLISATION PRÉDICTIVE

Les processus d'affaires de la plupart des organisations peuvent tirer parti de la modélisation prédictive. Cela dit, dans les situations suivantes, les modèles prédictifs peuvent offrir une valeur ajoutée considérable :

- Processus qui requièrent un grand nombre de décisions semblables.
- Lorsque les résultats ont une incidence considérable en terme d'argent et en vies.
- Lorsqu'une grande quantité de données informatiques est disponible pour la prise de décision et l'analyse des résultats.
- Lorsqu'il est possible d'insérer un modèle de calcul dans le processus d'entreprise actuel, soit pour automatiser la prise de décision ou soutenir les décideurs.

L'évaluation de crédit est une application de qualité supérieure :

- Présence d'un flot régulier de demandes de prêts contenant de l'information à propos des emprunteurs.
- Le résultat de chaque décision a une incidence financière considérable et est saisi dans des systèmes de comptabilisation des prêts.
- Les modèles s'intègrent aux processus d'octroi du prêt de façon logique pour soutenir la prise de décision humaine.

Tout comme l'évaluation des risques cardiaques :

- Les symptômes des patients sont maintenant saisis dans son dossier médical électronique, tout comme les diagnostics et les résultats ultérieurs.
- Les décisions en matière de traitement ont des répercussions sur la vie ou la mort.
- Il est possible d'utiliser un modèle pour guider les décisions immédiates en matière de traitement à partir du moment où l'information du patient est inscrite lorsqu'il est aux urgences.

*Pour les banques et les prêteurs, l'évaluation du crédit constitue l'un des nombreux domaines où l'analyse prédictive permet d'apporter une valeur ajoutée.*

*En utilisant un modèle pour prédire un enjeu d'affaires essentiel, une organisation peut transformer un risque de type « inconnu-inconnu » en un risque « connu-inconnu », ou en d'autres termes, un risque calculé.*

*Les processus d'affaires de la plupart des organisations peuvent tirer parti de la modélisation prédictive.*

## CALCUL DU RENDEMENT DU CAPITAL INVESTI DE L'ANALYSE PRÉDICTIVE

Dans plusieurs cas, il est possible de mesurer les avantages potentiels et même d'estimer le rendement du capital investi d'un modèle prédictif à l'aide d'une méthode simple : l'ensemble swap. Comme le démontre le tableau ci-dessous, l'ensemble swap est une série de décisions qui ont pu être améliorées grâce à un modèle prédictif.

Les pistes de vente provenant du site web d'une entreprise constituent un bon exemple d'application. Une fois reçues, les pistes sont assignées à l'un des représentants commerciaux internes pour un suivi immédiat, ou une réponse automatique est acheminée par courriel à ce dernier.

L'entreprise souhaite accroître ses ventes en conservant le même effectif. Elle développe alors un modèle prédictif pour mesurer la probabilité que la piste mène à la conclusion d'une vente. Les règles d'évaluation sont modifiées pour intégrer le modèle en assignant les pistes à forte probabilité aux représentants commerciaux. L'ensemble swap compare alors les résultats de l'ancienne méthode d'évaluation axée sur l'intuition à ceux du nouveau modèle. Il en résulte une augmentation mensuelle de 200 000 \$.

Processus existant (Processus compatible au modèle prédictif)	Assigné au rep.	Courriel seulement	Pistes totales
<b>Assigné au représentant</b>	1 000 pistes	1 000 pistes	2 000 pistes, ventes de 1 000 000 \$
<b>Courriel seulement</b>	1 000 pistes	7 000 pistes	8 000 pistes, ventes de 500 000 \$
<b>Pistes totales</b>	2 000 pistes ventes de 700 000 \$	8 000 pistes ventes de 600 000 \$	10 000 pistes ventes de 1 500 000 \$ (modèle) ventes de 1 300 000 \$ (ancien) Augmentation de 200 000 \$

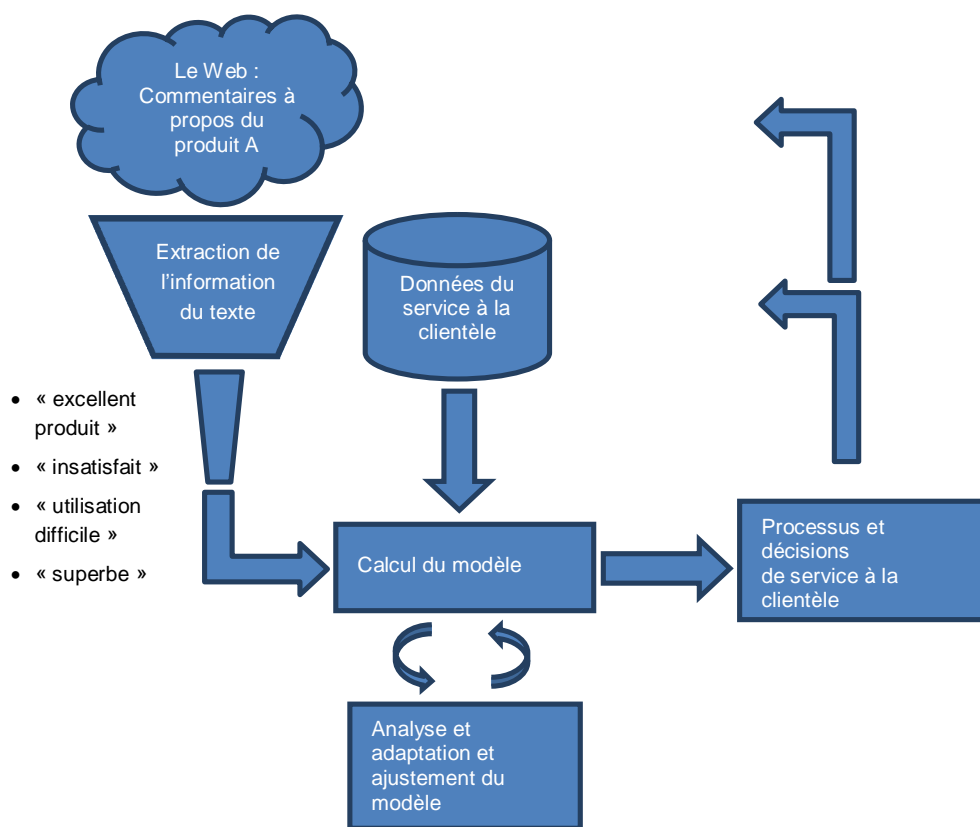
Démonstration de l'ensemble swap avec l'évaluation des pistes de vente. L'utilisation d'un modèle d'assignation des pistes amène des décisions de traitement différentes des décisions précédentes et génère des résultats accrus avec les mêmes coûts de vente.

## TENDANCES EN MATIÈRE DE MODÉLISATION PRÉDICTIVE

Les principaux changements et avancées en modélisation prédictive résultent de l'explosion de données non structurées, documents écrits, vidéos et images, et des techniques d'analyse qui évoluent à un rythme effréné. En bref, la modélisation prédictive requiert de l'information structurée comme on en trouve dans les bases de données par exemple l'analyse de sentiments provenant de publications en ligne.

L'information peut être extraite à partir de publications de clients sur des forums, des blogues ou de toutes autres sources afin de prédire la satisfaction de la clientèle et les tendances de ventes de nouveaux produits. Il serait toutefois pratiquement impossible de concevoir un modèle prédictif à l'aide du texte des publications. Une étape d'extraction est requise pour obtenir de l'information exploitable sous forme de mots-clés, d'expressions et de significations provenant du texte publié, comme le démontre le graphique ci-dessous. Il est alors possible d'établir une corrélation entre les occurrences de l'expression « problèmes avec le produit » et l'augmentation du volume d'appels de service par exemple.

*Dans plusieurs cas, il est possible de mesurer les avantages potentiels et même d'estimer le rendement du capital investi d'un modèle prédictif à l'aide d'une méthode simple – l'ensemble swap.*



Démonstration du flux d'information et des processus d'une application d'analyse de sentiments

Toutes formes de données non structurées (p. ex. : texte, images, vidéos et son) sont accompagnées d'un ensemble de techniques permettant d'extraire l'information utile des données brutes et de l'analyser, notamment au moyen de la modélisation prédictive. Une fois cette étape terminée, les méthodes visant à trouver des applications de modélisation prédictive supérieures sont les mêmes pour les données non structurées que pour les données structurées : décisions stratégiques, données numérisées tant pour les résultats que pour les indicateurs prévisionnels et façon d'insérer un modèle dans le processus actuel de prise de décision.

## CONCLUSION

Selon CGI, le modèle prédictif constitue une façon extrêmement efficace d'obtenir une valeur ajoutée des données et offre une foule d'applications qui n'ont pas encore été pleinement exploitées. Le phénomène de données massives viendra augmenter le nombre d'utilisations de la modélisation prédictive au sein du gouvernement et de l'industrie.

Parallèlement, l'analyse prédictive ne convient pas à tous les types d'applications. Comme pour n'importe quelle technologie puissante, une attention particulière doit être portée pour implanter des modèles prédictifs de façon pragmatique afin de créer une réelle valeur ajoutée.

L'adoption à grande échelle de l'analyse prédictive est freinée par le manque d'expertise pour analyser les données, pour créer des modèles et pour les implanter de manière efficace. Les entreprises souhaitant valoriser leurs données au moyen de l'analyse prédictive devraient investir dans la formation en analyse, dans les programmes de mentorat et profiter de l'expertise des firmes partenaires.

## À PROPOS DE CGI

Chez CGI, nous nous engageons à contribuer au succès de tous nos partenaires. Nos 69 000 professionnels présents dans plus de 40 pays fournissent des services complets en technologies de l'information (TI) et en gestion des processus d'affaires qui facilitent l'évolution constante des activités de nos clients.

Nous sommes déterminés à aider nos clients à atteindre leurs objectifs d'affaires, à offrir des carrières gratifiantes à nos professionnels et à procurer à nos actionnaires des rendements supérieurs au fil des années. La raison d'être de CGI est de respecter ses engagements en produisant des résultats.

Pour en savoir plus, visitez [www.cgi.com](http://www.cgi.com) ou communiquez avec nous à [info@cgi.com](mailto:info@cgi.com).