



Article from

Risk Management

December 2018

Issue 43

Dangers et perspectives d'avenir de l'analyse prédictive

par Mark Griffin

Aujourd'hui, de nombreux assureurs de personnes utilisent l'analyse prédictive (AP) à diverses fins. Le présent article traite des défis particuliers et des bouleversements que l'AP est susceptible d'engendrer pour les gestionnaires du risque, de même que des possibilités qu'elle laisse entrevoir.

Voyons tout d'abord les défis généraux :

- D'ordinaire, les modèles d'AP sont construits par un groupe de scientifiques de données récemment embauchés. Même si ceux-ci sont généralement très doués en analyse, il se peut qu'ils connaissent peu le secteur des assurances et les protocoles et techniques de gouvernance de modèles au sein des institutions financières.
- Les modèles d'AP sont souvent construits à partir d'un code source ouvert. Il sera facile et tentant pour les constructeurs de modèles d'importer des groupes de codes des bibliothèques ou d'autres sources pour effectuer certaines tâches. Il devrait y avoir un protocole précis à ce sujet.
- Les gestionnaires du risque d'assurance ont habituellement une bonne connaissance, sinon une expérience pratique, des modèles traditionnellement en usage pour prévoir l'actif, le passif et les autres valeurs financières en assurance. En ce qui concerne les modèles d'AP, il est peu probable que les gestionnaires du risque aient une expérience pratique ou quelque connaissance que ce soit en la matière. En fait, il se peut même que personne au sein de la société (à part ceux qui ont construit les modèles) ne puisse effectuer un examen par les pairs ou procéder à une évaluation critique des modèles. Même à l'extérieur de la société, il se peut que seul un nombre restreint de personnes soient en mesure de bien le faire.

En outre, en matière de gestion du risque, il ne faut pas oublier les craintes habituelles que fait naître tout grand projet de développement de systèmes. Par exemple, des modèles de gestion de projet et de sécurité des données devraient être en place, les avantages et les inconvénients de l'utilisation du nuage devraient être pris en compte, etc.

Les gestionnaires du risque doivent d'abord reconnaître que l'analyse prédictive est un domaine très dynamique. Kaggle est une plateforme très utilisée lors des concours d'analyse prédictive où sont réunis des groupes aux prises avec des problèmes et des données, et des scientifiques de données bien disposés. Plusieurs sociétés d'assurances ont parrainé des projets

en lien avec Kaggle. En décembre dernier, la Society of Actuaries (SOA) a publié des études de cas de son programme Kaggle Involvement¹. Se familiariser avec Kaggle est un premier pas assez simple à franchir. En outre, Google a annoncé dernièrement qu'il mettrait à la disposition de tous son savoir-faire en matière d'apprentissage automatique et d'intelligence artificielle.

Voici d'autres considérations et suggestions particulières.

En ce qui concerne les données :

- Le gestionnaire du risque devrait être au courant des décisions de haut niveau et des choix de modélisation. A-t-on utilisé un modèle linéaire généralisé ou une méthode d'apprentissage automatique? Avec un modèle linéaire généralisé, le processus de modélisation commence souvent par une hypothèse. D'ordinaire, une variable cible est établie. Ce sont deux étapes que le gestionnaire du risque devrait comprendre. Si beaucoup de données sont disponibles, les constructeurs de modèles pourraient tenter de déterminer un ensemble réduit d'« attributs » effectifs, qui sont des combinaisons de variables qui semblent influencer sur la variable cible. Les méthodes d'apprentissage automatique permettent à l'ordinateur de rechercher les liens de dépendance entre les données avec une interface humain-machine beaucoup moins grande. Bien que ces méthodes soient plus efficaces pour trouver des relations qui n'étaient pas reconnues ou pas bien comprises auparavant (et donc plus « indépendantes » des approches et des schèmes de pensée actuels), elles peuvent être utilisées pour détecter des relations peu étayées par des données.
- De nombreuses techniques d'AP équivalent à utiliser la puissance informatique pour vérifier si les résultats du modèle sont stables lorsque différents sous-ensembles de données sont utilisés. Pour ce faire, les données sont habituellement divisées en trois groupes : entraînement, validation et test, souvent dans des proportions de 70 %, 20 % et 10 %. Il se peut qu'il n'y ait pas assez de données pour se permettre ce luxe. De plus, toute dimension temporelle des données limitera ce processus. Les actuaires pourraient utiliser la signification statistique comme critère de test afin d'avoir un autre point de vue sur cette même question. L'article « La crédibilité est-elle encore crédible? », paru dans le numéro d'août 2017 de *Gestion du risque*, s'intéresse à cette question.
- Quelle que soit la méthode choisie, les spécialistes concernés ont-ils participé?
- Il existe une longue liste de considérations relatives à la gouvernance de modèles. Le contrôle des versions est-il effectué automatiquement par le logiciel ou par des développeurs? Existe-t-il un guide de rédaction pour le codage et la documentation? Une « piste de vérification » sera-t-elle produite?
- Le modèle peut-il être expliqué aux utilisateurs? À la direction? Les utilisateurs comprennent-ils le but et les résultats, et utiliseront-ils le modèle de façon appropriée, compte tenu de ses limites?

- Comment l'impact et la réussite du modèle seront-ils mesurés? Certains pourraient vouloir mesurer la rapidité du modèle; d'autres voudront peut-être mesurer sa fréquence d'utilisation, tandis que d'autres pourraient vouloir mesurer la variation des résultats financiers.

La modélisation AP offre l'occasion de bouleverser en profondeur les méthodes traditionnelles. Un bon exemple est celui de la simplification de la procédure de souscription de l'assurance temporaire. Les sociétés ont cherché à faciliter et à accélérer le processus d'achat d'assurance temporaire. L'élimination du prélèvement de liquides organiques de la procédure de souscription est la façon la plus rapide de le faire. La modélisation AP est donc employée pour construire des modèles de souscription qui utilisent des sources de données autres que les liquides organiques qui ont historiquement été prélevés, en plus d'autres sources qui peuvent avoir un pouvoir prédictif en matière de mortalité. Il pourrait s'agir des cotes de crédit, des dossiers de la pharmacie, etc. Les modèles obtenus sont ensuite testés par rapport aux données historiques pour voir dans quelle mesure ils reproduisent les résultats des processus classiques de souscription qui prévoient le prélèvement de liquides organiques.

Comme c'est le cas de tout bouleversement, il peut procurer de très gros avantages aux sociétés qui font le pari d'une adoption rapide et efficace, mais être très préjudiciable à celles qui restent à la traîne.

Il vaut la peine de prendre le temps d'imaginer la suite des choses. Au fur et à mesure que les sociétés modifieront leur procédure de souscription pour utiliser de nouvelles sources de données ou même simplement pour revoir leur pondération entre les nouvelles sources de données et les anciennes sources, le comportement des conseillers en sécurité financière et celui des consommateurs vont aussi changer et entraîneront un changement dans la composition du portefeuille de polices. À terme, les sociétés qui conserveront une approche traditionnelle pourraient perdre la majeure partie ou la totalité de leur part de marché – voire, en fin de compte, faire l'objet d'antisélection – au profit des sociétés qui font une meilleure sélection des risques avec les nouveaux modèles.

Bon nombre d'assureurs ont consacré leurs efforts AP au domaine de la mise en marché, tentant de faire face au bouleversement qui s'opère. À tout le moins, ils devraient rendre plus efficaces les méthodes et les canaux traditionnels de mise en marché et de vente.

Dans le secteur des assurances IARD, où l'on utilise couramment l'analyse prédictive depuis des décennies, la procédure de souscription a été transformée en une qui, dans bien des cas, ne prévoit aucune catégorie, juste un tarif unique pour chaque client. Avec des produits tels que l'assurance temporaire, pour lesquels la boucle de rétroaction entre l'activité de souscription et les résultats est plus lente, l'évolution peut être plus lente, mais se dirigerait naturellement dans la même direction.

Outre ces difficultés, l'analyse prédictive fait naître des possibilités du point de vue du gestionnaire du risque. Voici deux exemples :

1. Les techniques d'AP permettent de mieux comprendre les liens de dépendance entre les variables, en particulier celles qui ne se voient pas au premier coup d'œil. Existe-t-il un meilleur outil pour aider le gestionnaire du risque à concevoir des systèmes d'alerte avancée ou des indicateurs de risque clés?
2. Les sociétés d'assurances ont commencé à prendre conscience de la valeur des données. Par exemple, dans l'exemple précédent où il est question de modèles de souscription simplifiés, non seulement les nouvelles sources de données achetées, mais aussi les anciennes données peuvent être numérisées, ce qui donne au modèle plus de points de données et lui permet de mieux reproduire les données de souscription historiques. Avec l'information numérique de souscription, l'efficacité du modèle historique peut être bien évaluée et sans aucun doute améliorée.

Les spécialistes de l'analyse avancée (évoqués en début d'article) peuvent faire partie de l'équipe de gestion du risque. Qu'ils le soient ou non, leur réussite profite à l'activité de gestion du risque, de sorte qu'il est essentiel d'entretenir avec eux des rapports étroits. La diversité des idées et des approches est toujours l'ami du gestionnaire du risque!

L'émergence et l'évolution des méthodes d'AP déstabiliseront les activités de mise en marché, de souscription et sans aucun doute de nombreux autres aspects du monde actuel de l'assurance de personnes. Comme c'est le cas de tout bouleversement, il peut procurer de très gros avantages aux sociétés qui font le pari d'une adoption rapide et efficace, mais être très préjudiciable à celles qui restent à la traîne. En plus des nombreuses considérations exposées dans cet article, il serait sage pour les gestionnaires du risque d'évaluer périodiquement l'usage que leur société fait de l'analyse avancée par rapport à leurs concurrents. □



Mark Griffin, FSA, CERA, est membre de la haute direction et chef de la gestion des risques chez TIAA. On peut le joindre à mark.griffin@tiaa.org.

NOTE

- 1 <https://www.soa.org/predictive-analytics/kaggle-program/>