

**« Du bon usage des algorithmes dans un monde digital »**

Intervention de Didier Warzée du pôle Fintech-Innovation de l'ACPR

Colloque organisé conjointement par l'Institut des actuaires et le groupe SCOR

le 14 décembre 2017

Bonjour,

David Dubois, Président de l'Institut des Actuaires, m'a sollicité pour intervenir sur le sujet du « bon usage des algorithmes ». Je remarque avec soulagement que nous n'en sommes pas encore arrivés au stade où la question serait posée par des algorithmes, qui eux s'interrogeraient sur « le bon usage des actuaires ».

Si nous n'en sommes pas encore là, il ne fait toutefois aucun doute que le *big data*, l'intelligence artificielle – dans toutes ses acceptions - sont des technologies clés pour la finance numérique, à condition que les enjeux de sécurité, de transparence et de protection des consommateurs soient correctement pris en compte.

Je souhaite en préalable à mon propos appeler l'attention sur le caractère parfois assez prospectif de cet exposé, étant par exemple précisé que l'on peut envisager des opportunités et des risques liés à l'usage de l'intelligence artificielle (IA), sans toutefois que ceux-ci ne se soient déjà pleinement matérialisés. Nous resterons notamment dans le champ de l'intelligence artificielle dite faible, le développement d'une IA dotée d'une forme de conscience restant à ce stade hypothétique.

L'intelligence artificielle s'appuie sur des quantités de données en croissance exponentielle et sur la sophistication des algorithmes utilisés, ce qui va en général de pair avec le caractère disruptif de leurs applications. Ces données sont rarement accessibles dans une quantité et une qualité adaptées, et les technologies d'intelligence artificielle de type *deep learning* semblent pour le

moment essentiellement se destiner à des cas impliquant de l'analyse d'images. Des technologies de *machine learning* moins sophistiquées présentent dès à présent des potentiels d'applications très variés, même s'ils sont *a priori* moins disruptifs que les technologies plus avancées.

La question de la bonne utilisation de ces algorithmes est en réalité multiple :

- Elle recèle tout d'abord une question qui a trait à l'efficacité et à la pertinence des algorithmes. Est-ce qu'un algorithme est efficace et pertinent pour tarifer ou provisionner un risque, ou même être utilisé dans une démarche commerciale ?
- Elle pose ensuite la question du cadre réglementaire s'appliquant à l'usage d'une telle technologie. Sur ce sujet, ce n'est pas tant la technologie sous-jacente qui est en question – il n'existe pas à ma connaissance une liste noire de méthodes algorithmiques défendues – mais le respect *in fine* d'obligations réglementaires concernant : la possibilité ou non d'utiliser les données qui alimentent cet algorithme, le respect des contraintes de gestion de ces données et les obligations vis-à-vis des consommateurs.
- Mais elle soulève la question, que se pose probablement un public d'actuaire, de l'utilisation de méthodes algorithmiques complexes dans le cadre de travaux soumis à la réglementation prudentielle.

### **I – Sur la question de l'efficacité des algorithmes**

Si les algorithmes de type *machine learning* semblent posséder des capacités intéressantes, ils posent tout de même, aux organismes comme aux superviseurs, un certain nombre de questions légitimes, notamment par rapport à des méthodes plus traditionnelles.

a/ En premier lieu, ces outils utilisent des techniques qui leur permettent d'apprendre et d'améliorer leurs performances en optimisant l'extraction et l'analyse de volumes importants de données. Cela pose donc le problème immédiat de la qualité et de la pertinence de ces données.

- Ce sujet, plutôt relatif au *Big Data*, est fréquemment associé aux 5V (Volume, Vitesse, Variété, Véracité et Valeur). Au-delà de la quantité extrêmement importante de données, on retrouvera donc également la variété de ces données et la possibilité de les exploiter de plus en plus rapidement. L'exactitude des données et leur valeur ajoutée sont des critères qualitatifs également fondamentaux.
- Même si ces algorithmes identifieront généralement des données non prédictives, ils pourraient tout de même être dégradés par des informations biaisées présentant un apparent pouvoir prédictif. C'est une responsabilité qui incombe à l'établissement que de s'assurer de l'exactitude des données mais également de justifier que leur utilisation n'introduise pas de biais.
  - Par rapport à un travail traditionnel de modélisation, la plus grande automatisation à l'œuvre dans la sélection des variables laissera moins de champ d'action au « sens critique » du modélisateur, dont le jugement peut reposer sur des critères extra-statistique (connaissance métier, éthique, sens moral).
  - Cela rejoint d'ailleurs un deuxième risque qui est « le renforcement de biais » : si une population a fait par le passé l'objet d'une tarification élevée en raison d'un facteur affectant temporairement leur risque (quel qu'il soit), il est vraisemblable que de plus mauvais risques ont été recrutés au sein de cette population et qu'un algorithme automatique les pondérera plus négativement qu'il ne le devrait dans une vision totalement neutre.

## b / Il conviendra par ailleurs d'identifier les méthodologies appropriées

- En avant-propos sur cette question, bien évidemment ces outils ne sont pas magiques – même si certains le croient ou le prétendent – et ne pourront extraire des données que l'information qui y est présente et utilisable.
  - Il faudra ainsi être prudents en face des vendeurs d'outils présentant des performances formidables par rapport aux techniques traditionnelles, et bien contrôler que ces performances sont robustes, stables et fondées sur des données opérationnellement utilisables.
- La première tâche sera déjà d'identifier la méthode d'apprentissage la plus pertinente dans un contexte donné – ce qui impliquera une mise à jour régulière des connaissances des actuaires travaillant sur ces sujets très évolutifs.
  - Notons toutefois que cette question n'a de sens que si elle est placée dans un contexte, à savoir une problématique donnée. C'est pourquoi déterminer le problème et par la suite la méthode d'apprentissage la plus adéquate semble la démarche la plus appropriée.
  - Par conséquent, la bonne réponse pourra parfois être que ce n'est pas un algorithme de *machine learning* qui est la solution la plus adaptée à la question posée.

## c / Enfin, il sera impératif d'évaluer l'adéquation de l'algorithme aux conditions concrètes d'utilisation

- Par exemple, l'intérêt de certains de ces algorithmes réside dans leur faculté de procéder à des mises à jour en continu des formules en fonction des données reçues : si cela peut s'avérer théoriquement intéressant, l'utilisation pratique peut en revanche s'avérer plus compliquée :

- On peut en effet subir des difficultés d'implantation concrète dans des systèmes d'informations qui ne sont pas conçus dans cette optique.
- Cela peut aussi être délicat du point de vue de l'efficacité commerciale : comment expliquer à un client qu'entre hier et aujourd'hui la proposition qui lui est faite change du tout au tout alors que rien n'a changé dans son profil par exemple ?

Sur tous ces points l'actuaire verra évidemment ses compétences en *data science* de plus en plus sollicitées, mais deux aspects moins techniques sont également très importants : le sens critique et la connaissance métier. En effet, ce sont ces deux dimensions qui feront le plus défaut aux algorithmes, et sur lesquelles la valeur ajoutée de l'actuaire sera fondamentale.

## **II – Même efficace, un algorithme doit être utilisé dans le respect des obligations réglementaires**

En effet, l'un des intérêts majeurs de ces algorithmes est de permettre l'exploitation automatique de volumes extrêmement importants d'informations, d'une façon généralement plus rapide et efficace que les techniques traditionnelles.

Toutefois, il convient au préalable de s'assurer que l'utilisation de ces données respecte bien la réglementation sur les données personnelles (notamment dans l'optique du RGPD qui entre en application en mai 2018<sup>1</sup>) et dans certains cas potentiellement d'autres réglementations (données de santé/génétiques, données sur les Organismes d'Importance Vitale...).

Au-delà de la simple question de la capacité légale d'utilisation d'une donnée, il sera impératif que la sécurité de cette donnée soit assurée, et ce dans un environnement informatique qui est de plus en plus ouvert.

---

<sup>1</sup> Règlement (UE) n° 2016/679, dit règlement général sur la protection des données (RGPD)

Il sera également important d'être attentif aux comportements déviants que pourraient adopter des algorithmes : dès lors qu'ils n'ont pas de moralité intrinsèque, leurs décisions pourraient ne pas respecter de critères éthiques.

- Ainsi, nous avons évoqué précédemment le renforcement de biais comme une problématique d'efficacité, mais elle peut aussi être une problématique éthique si un comportement discriminatoire se trouvait renforcé par l'apprentissage automatique.
- Prenons l'exemple d'un organisme de crédit qui a été confronté à une fraude organisée par des individus se prétendant tous plombiers, avec pour conséquence une défiance excessive ultérieure des services d'octroi face aux demandes issues de cette profession. Il en résulte que, dans la base d'apprentissage de l'algorithme, être plombier ressort comme un critère de risque élevé, qui appliqué sans sens critique peut conduire à un refus systématique évidemment discriminatoire de cette profession.
- Cela illustre le fait qu'un algorithme optimisant des critères statistiques sans être doté du sens critique dont est normalement doté un être humain, pourrait conduire à un comportement discriminatoire. Ce cas est assez caricatural car il focalise simplement sur une variable, mais on peut imaginer des algorithmes reproduisant des comportements discriminatoires via des combinaisons de variables moins faciles à identifier.
- Avant d'envisager l'utilisation opérationnelle d'un algorithme, il est donc nécessaire de prévoir des tests suffisants pour être certains de ne pas déployer un outil plus discriminatoire que statistiquement discriminant.

Au-delà de ces questions, qui tombent pour celles ayant trait aux données personnelles dans le champ de compétences de la CNIL, les règles de protection des consommateurs affectent également les possibilités d'utilisation d'algorithmes. Ainsi ces technologies doivent respecter la bonne information

du client et la loyauté à la fois des explications, mais également des conseils et préconisations qui lui sont donnés.

- Tout d'abord, et avant même de parler des conditions réglementaires de leur utilisation, il faut rappeler la problématique de l'acceptation par les consommateurs de ce type d'outils. Ceux-ci ne sont pas forcément tous prêts à gérer leurs opérations les plus sensibles via des algorithmes. Le bon usage des algorithmes pourra donc évoluer en fonction de la maturité des clientèles. Un dosage entre le conseil automatisé et le conseil humain méritera aussi d'être défini.
- En tout état de cause, la loyauté et la transparence vis-à-vis du client imposent qu'il soit informé de la nature de son interlocuteur : en effet, la réglementation prévoit qu'il puisse accéder au besoin à un conseiller humain.
- Les algorithmes utilisés pour des décisions concernant un client, ou même une proposition qui lui est faite, doivent donc pouvoir expliquer les motivations du résultat proposé, et démontrer qu'ils ne participent pas d'un éventuel conflit d'intérêt. Cette tâche pourrait être assez malaisée pour des algorithmes très complexes. Elle sera incontournable et nous avons connaissance de travaux en cours sur des méthodes permettant de limiter l'effet « *black box* » que peuvent avoir ces outils (au moins pour certains d'entre eux).
- Aujourd'hui, il semble d'ailleurs que l'un des apports les plus intéressants des algorithmes pourrait résider dans la plus-value apportée aux conseillers humains, qui deviendront grâce à ces outils ce que l'on pourrait appeler des « conseillers augmentés », mieux outillés pour fournir un conseil de qualité et adapté aux besoins du client.
- Cela étant, il est certain que l'usage d'un algorithme ne dispensera pas des prérequis d'un conseil humain. Il est donc impératif que ces algorithmes ne conduisent pas à « manipuler » un consommateur : des

algorithmes qui orienteraient un client non pas en fonction de ses besoins et de ses intérêts mais de ce qu'il serait prêt à payer, serait totalement hors-jeu : le fait de fournir un conseil via un algorithme ne pourra en aucune façon excuser le défaut de conseil.

### **III – Même pour l'optimisation d'outils internes sans lien direct avec le client, comme pour le calcul de provisions techniques, les établissements financiers doivent tenir compte de certaines règles**

Il s'agit en réalité des règles qui s'appliquent également aux méthodologies « classiques », notamment les règles issues de Solvabilité 2, mais qui doivent s'adapter aux spécificités des algorithmes. On trouvera donc parmi ces exigences :

- La qualité des données en entrée, car un algorithme auto-apprenant qui se fonderait sur des données de mauvaises qualités sera nécessairement moins performant.
  - o Certes des méthodes existent permettant de gérer des données tronquées, et on peut imaginer que des algorithmes intégrerai-ent des fonctionnalités de correction avancée sur ce type de problématique.
  - o Mais il faudra en tout état de cause que la qualité des données soit précisément évaluée et que les méthodes employées soient en conséquence justifiées.
- Par ailleurs, si l'on considère le calcul de *Best Estimate*, tout l'enjeu sera d'adapter les requis habituels aux algorithmes envisagés :
  - o Pour ce qui est de la documentation technique, elle restera bien sûr indispensable et devra clairement justifier de la pertinence technique de l'algorithme envisagé. La difficulté pourra notamment résider, en particulier dans le cas de recours à des prestataires, dans la capacité à fournir le détail méthodologique qui sous-tend l'outil utilisé.

- De même, l'utilisation de ces outils appelle la présentation d'un dispositif de contrôle interne, incluant une piste d'audit et des processus de validation et de *back-testing* de ces outils, qui sont par nature très évolutifs.
- La compréhension du comportement des algorithmes sera impérative pour évaluer s'ils maintiennent une réponse appropriée en situation de *stress* et voir si leur optimisation en période plus calme ne se fait pas au prix de comportements inefficaces quand les conditions deviennent tendues.

\*\*\*

En conclusion, je voudrais donc dire que le bon usage de l'intelligence artificielle dans un monde digital implique de préserver ou de développer une grande quantité d'intelligence humaine, à la fois pour concevoir, alimenter, déployer et contrôler ces outils.

À cet égard, permettez au superviseur que je suis d'appeler l'attention sur les risques qui peuvent naître des dépendances à des prestataires – fournisseurs de données ou prestataires technologies : tout d'abord le risque de continuité de service voire le risque de réputation; ensuite les risques financiers, par exemple en cas de fuites de données causée par des mesures de cybersécurité insuffisantes générées par des systèmes plus ouverts. Notons donc que la dépendance à des fournisseurs communs pourrait même constituer un nouveau risque systémique.

Ensuite se posera la question stratégique de la pertinence d'utilisation d'un tel algorithme par une entreprise : la réponse prendra naturellement en compte les bénéfices attendus mais aussi les différents risques générés. Ces risques peuvent être issus de l'inadéquation avec les objectifs de l'entreprise, en matière d'appétit pour le risque ou de positionnement par rapport aux clients.

En tout état de cause, aux côtés des enjeux de sécurité liés à la collecte, au stockage et l'utilisation de masses de données de plus en plus étendues, les

enjeux d'auditabilité et d'explicabilité des algorithmes seront significatifs pour les entreprises comme pour les superviseurs. Ces défis, qui ne doivent pas occulter les bénéfices attendus de l'intelligence artificielle, doivent faire l'objet d'échanges constructifs, parfois assez prospectifs, entre les superviseurs et les entreprises financières. C'est ce que nous faisons ce soir et c'est ce que nous continuerons de faire pour que les acteurs financiers et leurs clients tirent le meilleur parti de cette nouvelle révolution technologique.

Je vous remercie de cette invitation.