

Bernard Ourghanlian, Microsoft : « Pourquoi le Machine Learning va gagner l'entreprise »

Silicon.fr : Pourquoi les grands fournisseurs de l'IT, dont Microsoft, mettent-ils autant l'accent aujourd'hui sur le Machine Learning, un concept qui est loin d'être nouveau ?

Bernard Ourghanlian : Le Machine Learning est une branche de l'Intelligence Artificielle qui existe, c'est vrai, depuis longtemps. Mais plusieurs éléments viennent changer la donne. Primo, on dispose aujourd'hui de données massives, or le lien entre Machine Learning et Big Data est assez naturel. Secundo, le domaine bénéficie de progrès très importants réalisés au milieu des années 90 par le mathématicien russe [Vladimir Vapnik](#). Ce dernier a formalisé la théorie statistique de l'apprentissage, ce qui permet de dépasser les approches bayésiennes (dérivées du théorème de Bayes, NDLR) utilisées jusqu'alors, surtout dans la lutte contre le spam. C'est donc ce double phénomène – des progrès techniques à la confluence de l'informatique et des maths, et l'océan numérique qui nous environne – qui explique que cette thématique passe aujourd'hui au premier plan.

Sans oublier le fait qu'on est face à des réalisations de plus en plus incontestables. Il faut ainsi souligner les progrès dans la traduction en temps réel, une technologie déjà ancienne puisqu'elle émane du département de la Défense américain au temps de la guerre froide, et qu'elle a ensuite été reprise par IBM pour les besoins de traduction du gouvernement canadien. Plus près de nous, Google a exploité le patrimoine du Web pour descendre à un taux d'erreur sur les mots (word error rate) d'environ 20 %. En travaillant sur des réseaux de neurones profonds (NDLR : modèle de calcul inspiré du fonctionnement des neurones biologiques, selon [Wikipedia](#)), Microsoft Research a récemment ramené ce taux à quelque 5 %. C'est cette technologie qui est exploitée dans Skype Translator (traducteur temps réel de l'outil de visioconférence, aujourd'hui en bêta, NDLR). On peut aussi parler de cette démonstration d'hélicoptère autonome, capable de s'adapter aux conditions de vol les plus impossibles. Dans ces deux cas, il serait impossible de programmer toutes les situations auxquelles le système va être confronté.

Pour l'entreprise, à quoi le Machine Learning peut-il servir ?

L'usage le plus naturel concerne les prévisions. Citons par exemple la start-up Lokad, qui vend des prévisions aux hypermarchés, sur n'importe quel produit, n'importe quelle implantation à n'importe quel moment. L'objectif étant de réduire le niveau des stocks sans entraîner de rupture de produits. Le but du Machine Learning consiste à augmenter l'explication des phénomènes que sous-tendent les données collectées. Les travaux de Vapnick permettent d'utiliser des réseaux de neurones à 10, 20 ou 30 couches, aboutissant à des modèles plus efficaces. Le problème consiste souvent à trouver le juste équilibre entre complexité du modèle et capacité à la calculer dans des temps raisonnables. C'est là où le Cloud joue un rôle déterminant.

Le Machine Learning est intimement lié à des théories mathématiques pointues. Or les

entreprises n'ont que rarement ces compétences en interne...

D'où notre choix visant à démocratiser ces technologies. Sur Azure ML ([l'offre de Machine Learning du premier éditeur mondial](#) sur le Cloud, NDLR), les entreprises ne sont pas obligées de disposer tout de suite d'un statisticien pour coder un algorithme sophistiqué. L'outil leur laisse la possibilité de démarrer, de créer des modèles et de les tester sur des sets de données pour s'approcher le plus possible de la réalité d'un phénomène. Contrairement à des outils pour statisticiens, comme SPSS ou SAS, les utilisateurs n'ont pas à coder en Java ou en C les modèles qu'ils ont créés. Dans Azure, le cycle se réalise en un clic : les modèles sont immédiatement mis en production et appliqués à de nouveaux jeux de données. L'autre avantage, c'est que le service peut s'appuyer sur les autres fonctions d'Azure, comme le stockage dans des tables, des blocs, sur des technologies de bases de données classiques ou NoSQL. Le service est par ailleurs associé à une place de marché d'algorithmes prêts à l'emploi, qu'on peut intégrer tels quels dans des applications. Ces algorithmes sont commercialisés sur Azure et peuvent être écrits en à peu près tous les langages, dont celui de R (logiciel Open Source de traitement de données et d'analyse statistique). Déjà, plus de 1 000 algorithmes sont disponibles sur le service.

Si la technologie est plus simple à aborder, affirmer qu'on peut la maîtriser sans compétences en statistiques serait exagéré. En fonction du cas d'usage, il faut travailler sur des algorithmes très différents. Dans certains cas, l'objectif est d'écartier les données aberrantes, alors que dans d'autres - la détection d'intrusion sur les réseaux par exemple -, il va s'agir au contraire de travailler sur les signaux faibles. Azure ML sert donc davantage à instaurer un pont entre statisticiens et informaticiens.

Combien de chercheurs travaillent sur le sujet chez Microsoft ?

Pas loin des deux-tiers des chercheurs de Microsoft Research travaillent de près ou de loin sur le Machine Learning, soit environ 600 personnes. On peut y ajouter une cinquantaine de personnes travaillant au sein de [notre laboratoire commun monté avec l'Inria](#). De facto, on retrouve des fonctions de Machine Learning dans de nombreux produits de Microsoft : dans notre assistant vocal Cortana, dans la recherche Bing mais aussi dans Excel pour des fonctions de complétion automatique de colonnes par exemple. Nos équipes de recherche travaillent aussi sur des sujets plus expérimentaux comme l'interprétation des images cardiaques, la reconnaissance automatique des images, l'écriture automatique de scripts de films pour leur indexation ou l'indexation des documents et connaissances sur Pompéi à partir d'un modèle 3D de la ville romaine ensevelie et dont les vestiges se dégradent rapidement avec l'afflux des touristes.

Récemment, le scientifique Stephen Hawking et l'entrepreneur Elon Musk (PayPal, SpaceX, Tesla) se sont inquiétés du développement rapide de l'intelligence artificielle, y voyant une menace pour l'Humanité. Qu'en pensez-vous ?

Ces craintes sont naturelles. Mais dire qu'elles vont se matérialiser demain matin me paraît exagéré. Nous n'avons toujours fondamentalement pas compris ce qu'était l'intelligence humaine, même si on parvient aujourd'hui à l'émuler dans un certain domaine de façon remarquable. A plus court terme, le développement du Machine Learning pose une autre question, celle de la disparition des classes moyennes. Car, avec des machines de plus en plus intelligentes, un grand nombre d'emplois risquent de disparaître.

A lire aussi :

[Un tiers des emplois remplacés par des robots d'ici 2025, selon Gartner](#)

[Machine Learning : IBM Watson fait son entrée dans les applications d'entreprise](#)

[Delve : Microsoft injecte du machine learning dans Office 365](#)