



© Shutterstock.com/Dmitriy Rybin

LES DÉFIS DES BIG DATA

Comment faire face à ce déluge de données qui constitue les *big data* ? Comment les stocker, les exploiter, les analyser, les valoriser... ? Les algorithmes et les calculateurs à qui ces tâches incombent doivent être à la hauteur. La recherche sur ces deux aspects essentiels du traitement des *big data* est particulièrement active. Le développement des intelligences artificielles, « éduquées » par des quantités énormes de données, est une illustration de ces efforts. Ces systèmes sont parfois si performants que certains y voient une menace sur la façon de faire de la science.

L'ESSENTIEL

● L'intelligence artificielle telle qu'on l'imaginait dans les années 1950 a été plus difficile à développer que prévu.

● Depuis quelques années, le domaine a connu un grand renouveau avec les techniques de l'apprentissage profond, inspirées des réseaux de neurones du cerveau.

● Un réseau de neurones artificiels à apprentissage profond acquiert sans cesse de l'expertise à partir de nouvelles données.

● Un tel réseau a déjà battu un joueur de go professionnel, d'autres reconnaissent des images ou la parole...

L'AUTEUR



YOSHUA BENGIO est professeur d'informatique à l'université de Montréal. Il est l'un des pionniers du développement des méthodes d'apprentissage profond.



La révolution de l'APPRENTISSAGE PROFOND

Pour créer une intelligence artificielle, pourquoi ne pas s'inspirer d'une intelligence naturelle ? C'est le pari des réseaux multicouches de neurones artificiels, des « machines » qui apprennent à partir de grandes quantités de données. Leurs succès sont saisissants !

E

n 2016, le programme AlphaGo, mis au point par Google DeepMind, faisait sensation en battant au jeu de go, par le score de 4 à 1, le Sud-Coréen Lee Sedol, réputé le meilleur joueur du monde. Un an après, en octobre 2017, AlphaGo s'est fait humilier avec un score de 100 à 0 par... AlphaGo Zéro, le nouveau logiciel de Google DeepMind!

La différence entre les deux logiciels ? Là où AlphaGo a été nourri par des millions de parties humaines, AlphaGo Zéro s'est contenté des règles du jeu et des positions des pierres sur le plateau. Il a appris en jouant des millions de parties contre lui-même, d'abord au hasard, puis en affinant sa stratégie. En quarante jours, il est devenu le meilleur !

Les deux logiciels ont toutefois un point commun : ils fonctionnent grâce à l'apprentissage profond, aboutissement de décennies de recherche sur l'intelligence artificielle. Aujourd'hui, les machines apprennent grâce à une architecture inspirée de celle du cerveau et d'un nombre considérable de données (AlphaGo Zéro les a générées lui-même). Un retour historique s'impose.

Dans les années 1950, le programme élaboré par Arthur Samuel pour jouer aux dames commence à damer le pion de joueurs de bon niveau. Stupéfaction dans le monde de l'informatique naissante. Le programme en question ▶



est le premier à inclure un autoapprentissage. L'enthousiasme voire l'euphorie gagne les laboratoires. La conférence de Dartmouth de 1956 consacre la naissance de l'«intelligence artificielle». L'un des pionniers, Marvin Minsky – en 1951, il avait construit la première machine à réseau neuronal, le Snarc – n'hésite pas à déclarer en 1970, dans *Life*: «Dans trois à huit ans nous aurons une machine avec l'intelligence générale d'un être humain ordinaire.» Cet optimisme était un tantinet précipité...

LA CHASSE AU SNARC

Les logiciels de l'époque, par exemple pour aider les médecins dans leur diagnostic, n'ont pas tenu leurs promesses. Les algorithmes étaient trop simples et manquaient des données nécessaires pour parfaire leur apprentissage. La puissance de traitement informatique était également insuffisante pour mener à bien les calculs complexes requis pour imiter, même approximativement, des subtilités de la pensée humaine.

Au milieu des années 2000, le rêve de construire des machines aussi intelligentes que des humains avait presque été abandonné. À cette époque, même le terme d'«intelligence artificielle» semblait avoir déserté le domaine de la science sérieuse. Les chercheurs décrivent la période allant des années 1970 au milieu des années 2000 par l'expression «AI winters», une sorte de long hiver glaciaire de l'intelligence artificielle où tous les espoirs avaient été brisés.

Les choses ont commencé à changer en 2005. Les perspectives dans le domaine de l'intelligence artificielle ont radicalement changé avec l'apprentissage automatique et l'émergence de l'«apprentissage profond», qui revisite le connexionnisme des années 1960 et s'inspire des neurosciences. Les réalisations actuelles de l'apprentissage profond sont à la hauteur des promesses d'antan, et les grandes entreprises des technologies de l'information, y consacrent désormais des milliards de dollars.

Qu'entend-on par apprentissage profond? Il s'agit d'un traitement effectué par un grand nombre de neurones artificiels (imitant de façon très simplifiée les neurones biologiques) qui, par leurs interactions, permettent au système d'apprendre progressivement à partir d'images, de textes ou d'autres données. L'apprentissage repose sur des principes mathématiques généraux. Le résultat de l'apprentissage est une représentation (par exemple, «Cette image contient des éléments différents»), une décision (par exemple «Cette image représente Jeanne Dupont») ou une transformation (par exemple la traduction d'un texte dans une autre langue).

La technique de l'apprentissage profond a transformé la recherche en intelligence artificielle, ranimant des ambitions oubliées pour la vision par ordinateur, la reconnaissance automatique de la parole et la robotique. Les premières applications ont vu le jour en 2012 pour la compréhension de la parole (Siri, qui équipe les iPhone en une illustration). Et peu après sont arrivés des logiciels qui identifient le contenu d'une image, une fonctionnalité qu'intègre maintenant le moteur de recherche de Google.

Tous ceux qui sont frustrés par les menus malcommodes de leur téléphone peuvent apprécier les avantages d'utiliser Siri ou un autre «assistant personnel» sur leur appareil. Et pour ceux qui se rappellent combien la reconnaissance d'objets était mauvaise il y a quelques années seulement, les avancées dans le domaine de la vision artificielle ont été phénoménales: aujourd'hui, moyennant certaines conditions, les ordinateurs savent reconnaître sur des images un chat, une pierre ou des visages presque aussi bien que les humains. Les logiciels d'intelligence artificielle font désormais partie du quotidien de millions d'individus, qui par exemple, n'écrivent plus de messages, mais les dictent à leur téléphone!

Ces progrès ont ouvert la voie à de nouvelles réalisations, à des applications commercialisées, et l'intérêt ne cesse de croître. La concurrence fait rage entre les entreprises pour attirer les jeunes talents... et les moins jeunes. De nombreux professeurs d'université spécialistes du domaine (la majorité, d'après certaines estimations) sont passés du milieu académique à l'industrie, séduits par des équipements de pointe et par de généreux salaires.

Les efforts déployés pour répondre aux défis de l'apprentissage profond ont débouché sur des réussites époustouflantes, les performances d'AlphaGo Zéro en sont l'illustration. Les applications vont s'étendre à d'autres domaines de l'expertise humaine, et pas seulement aux jeux. Par exemple, un algorithme d'apprentissage profond serait capable de diagnostiquer des insuffisances cardiaques sur des images d'IRM aussi bien qu'un cardiologue.

Ce succès récent de l'apprentissage profond a de quoi surprendre quand on se penche sur l'histoire de l'intelligence artificielle. Pourquoi celle-ci a-t-elle buté sur tant d'obstacles dans les décennies précédentes? Parce qu'apprendre des choses nouvelles est, pour une machine, difficile. L'essentiel de la connaissance que nous avons du monde autour de nous n'est pas formalisé dans un langage écrit sous forme d'un ensemble de tâches explicites, ce qui est indispensable pour écrire un programme informatique. C'est pourquoi nous n'avons pas pu directement programmer un ordinateur pour effectuer la plupart des tâches qui nous sont

APPRENTISSAGE PROFOND ET ÉTUDE DE L'ADN

Dans le domaine de la génomique, l'accumulation de données nécessite des algorithmes puissants pour en extraire des informations. En 2015, Brendan Frey, de l'université de Toronto, et ses collègues ont développé DeepBind, un logiciel qui recourt à l'apprentissage profond pour analyser la façon dont des protéines se lient à l'ADN et à l'ARN. Les chercheurs détectent ainsi des mutations qui perturbent les processus cellulaires entraînant des maladies.