

INTELLIGENCE ARTIFICIELLE, LA RÉALITÉ DERRIÈRE L'EMBALLEMENT

Face à l'engouement pour l'intelligence artificielle, ses applications révolutionnaires tardent à venir. Qu'en est-il réellement ?

Visite guidée des coulisses et des enjeux avec **Francis Bach**, tenant d'une IA « responsable et efficace ».

Tout le monde parle d'intelligence artificielle (IA) depuis une dizaine d'années, dans les sphères scientifiques, économiques et personnelles. Ce renouveau vient, comme les vagues d'engouement précédentes, avec ses promesses (la voiture autonome d'ici dix ans), ses peurs (mon emploi va-t-il être remplacé par une machine ?), voire ses fantasmes (les robots vont-ils prendre le contrôle ?).

APPRENTISSAGE AUTOMATIQUE

Dans le cadre de notre partenariat avec l'Académie des sciences, les académiciens nouvellement élus fin 2019 présentent un éclairage sur leur discipline et ses enjeux scientifiques, éthiques, politiques et sociétaux, à travers leur expérience personnelle.

notre appareil photo se charge de la mise au point sur les visages ou quand les meilleurs joueurs d'échecs et de go se font battre par les machines.

L'IA est aussi présente de manière plus cachée, voire insidieuse, quand elle est utilisée pour les campagnes de marketing, pour nous suivre et nous proposer des produits à acheter lors de notre navigation sur Internet ou pour nous proposer un contenu personnalisé sur les différents moteurs de recherche, agrégateurs d'articles de presse ou réseaux sociaux.

Au-delà de cette surcommunication et de toutes ces attentes, l'IA est déjà parmi nous de manière très visible et impressionnante, lorsque nous dictons un message à notre téléphone, quand nous utilisons un traducteur automatique, quand

Les fantasmes de prise de pouvoir hostile des robots ou de transhumanisme ont été rapidement évacués. Les craintes de menaces sur l'emploi sont légitimes, mais l'histoire économique nous apprend que, pour les « technologies d'usage général », comme la machine à vapeur, l'électricité, l'ordinateur ou Internet, et maintenant l'intelligence artificielle, les effets sur l'emploi sont multiples et difficiles à prévoir, car, si certains emplois disparaissent, de nouveaux apparaissent.

Qu'en est-il des grandes promesses, comme la voiture autonome pour tous ou la médecine personnalisée ? Pourquoi les applications révolutionnaires promises par l'IA mettent-elles du temps à arriver ? Majoritairement parce que les connaissances actuelles en informatique et en mathématiques ne sont pas suffisantes. Des efforts de recherche sont donc cruciaux pour aboutir à une IA responsable et efficace.

UNE IA VRAIMENT INTELLIGENTE ?

Plongeons-nous d'abord dans les facteurs du renouveau de l'IA. Ces progrès récents sont en grande partie le fruit de l'utilisation systématique d'algorithmes dits d'« apprentissage automatique » (« machine learning » en anglais), qui utilisent des masses de données où de nombreux exemples du comportement final désiré sont disponibles (on parle de données « étiquetées »). L'apprentissage est en général « supervisé » : la machine apprend à partir d'exemples,



PROFIL

Directeur de recherche à Inria, institut national de recherche en sciences et technologies du numérique, responsable de l'équipe Sierra d'apprentissage statistique – une des facettes de l'intelligence artificielle –, **Francis Bach** est membre de l'Académie des sciences. Pionnier et spécialiste mondialement connu de l'apprentissage automatique, il s'intéresse en particulier à ses aspects algorithmiques et théoriques.

qui eux-mêmes sont fournis par des humains. Par exemple, pour la reconnaissance d'objets dans des images, il a fallu fournir « à la main » une description du contenu de millions d'images, et pour la traduction automatique, des millions de phrases traduites par des humains. Apprendre ainsi à partir de données massives nécessite une puissance de calcul importante.

En caricaturant, les technologies actuelles de l'IA ne peuvent se comporter intelligemment – ou presque – que dans les situations déjà observées précédemment et suffisamment souvent. Pour la reconnaissance d'objets dans les images ou la traduction automatique, les données étiquetées existantes ou créées pour l'occasion suffisent, dans des contextes non critiques où les erreurs commises par l'IA ne prêtent pas à conséquence. Qu'en est-il pour les applications futures ?

L'IA est déjà parmi nous de manière très visible et impressionnante, lorsque nous dictons un message à notre téléphone, quand notre appareil photo se charge de la mise au point sur les visages ou quand les meilleurs joueurs d'échecs et de go se font battre par les machines...

AU-DELÀ DE L'APPRENTISSAGE SUPERVISÉ

Prenons l'exemple de la voiture autonome. Les capacités de généralisation au-delà de ce qui a été observé étant actuellement limitées, il est nécessaire d'avoir proposé aux algorithmes d'apprentissage des observations venant de paysages (campagne, ville ou banlieue), de saisons différentes (neige, pluie, canicule), d'heures de la journée différentes, de pays différents, voire de styles de comportement des différents usagers de la route (observation différenciée du Code de la route), etc. À chacun des critères précédents correspond la nécessité de multiplier le nombre de données par un grand facteur (disons arbitrairement 10, pour fixer les esprits). À cause de cet effet multiplicatif, il sera fastidieux et coûteux d'acquérir et d'étiqueter tant de données.

C'est pour cela que, dans la plupart des applications de l'IA, de nouveaux algorithmes sont nécessaires pour réduire la quantité de données étiquetées, et utiliser les masses disponibles de données sans supervision ou avec une supervision faible (comme un bébé observe son environnement) pour « comprendre comment marche le monde » ou, plus précisément, pour créer une représentation du monde permettant d'apprendre avec une quantité de données étiquetées inférieure.

PROTECTION DE LA VIE PRIVÉE ET ÉQUITÉ

La collecte de données nécessaire aux algorithmes d'apprentissage, comme toute collecte de données personnelles, pose d'épineux problèmes de protection de la vie privée. Dois-je donner mon accord explicite pour que mes données de localisation et de navigation sur Internet ou le contenu de mes e-mails soient utilisés par de grandes entreprises technologiques pour me proposer de nouvelles publicités ? Est-ce réellement dangereux de partager de manière contrôlée, protégée et anonyme mes contacts physiques récents pour participer à lutter contre une pandémie ? Si j'autorise l'utilisation de mes données médicales pour des études épidémiologiques, vais-je bénéficier de nouveaux traitements potentiels ? Ces questions concernent tout traitement de données numériques et vont au-delà de l'IA, mais leur prise en compte en amont par les chercheurs en IA est nécessaire pour une utilisation généralisée et responsable d'algorithmes personnalisés.

L'utilisation massive de données pose un autre problème, d'ordre éthique : celui des biais dans les données utilisées et de la difficulté d'aboutir à des prédictions équitables. Prenons l'exemple d'un algorithme cherchant à proposer des offres d'emploi à partir de CV : s'il a appris à partir de bases de données du passé, il ne va proposer qu'aux hommes des emplois dans le monde de l'informatique. Il y avait un biais dans »

Prenons l'exemple d'un algorithme cherchant à proposer des offres d'emploi. S'il a appris à partir de bases de données du passé, il ne va proposer qu'aux hommes des emplois dans l'informatique.

» les données, que l'algorithme a reproduit. Dans une application médicale, si le panel de patients inclus dans l'étude initiale n'est pas équilibré, le diagnostic ne sera pas aussi précis pour les patients aux caractéristiques sous-représentées. Ces biais se retrouvent à de multiples niveaux, en particulier social ou ethnique. Leur prise en compte est un enjeu actuel important de l'apprentissage automatique, aussi bien au niveau du développement de nouveaux algorithmes dits « équitables » que dans leur mise en œuvre pratique.

ROBUSTESSE, FIABILITÉ ET CERTIFICATION

L'utilisation de l'informatique dans des applications critiques, comme le transport ou la médecine, nécessite depuis longtemps l'établissement de garanties de bon fonctionnement des logiciels utilisés. Dans le cadre d'algorithmes d'apprentissage, il faut aller au-delà des bugs informatiques communs à tout logiciel, que les technologies actuelles de certification permettent en principe d'éviter. En effet, les méthodes d'apprentissage sont avant tout des méthodes statistiques pour lesquelles les résultats comportent des marges d'erreur (comme pour les sondages). Quantifier cette incertitude est indispensable dans les systèmes critiques, les grandes difficultés étant que les taux d'erreur finaux qui seront admis par les utilisateurs seront très faibles,

et que l'estimation de ces incertitudes est pour l'instant hors d'atteinte pour les derniers algorithmes qui ont permis les avancées récentes, comme les réseaux de neurones profonds.

Au-delà de la variabilité naturelle des données, les algorithmes doivent aussi être robustes aux attaques malicieuses (par exemple s'assurer que la photo d'une personne ne suffit pas pour débloquer son téléphone par reconnaissance faciale ou éviter de pouvoir changer quelques pixels bien choisis d'une image afin de tromper la machine dans sa prédiction). Là aussi, le développement de méthodes robustes et fiables est un domaine de recherche particulièrement actif.

FRUGALITÉ

Traiter des masses de données demande une forte puissance de calcul, ce qui a un impact sur l'environnement, à travers notamment la construction des ordinateurs nécessaires (énergie dépensée, utilisation de matériaux rares, transport et recyclage), et la consommation d'électricité requise pour faire fonctionner ces ordinateurs et les refroidir.

Comme l'ensemble des technologies numériques, l'IA permettra de remplacer des activités existantes coûteuses pour l'environnement. On pense naturellement au télétravail rendu possible par les outils de vidéoconférence qui, au-delà des considérations sanitaires, réduit les transports quotidiens, avec, comme pour toute technologie, des effets rebonds parfois inattendus (pour le télétravail, un déménagement dans un logement plus éloigné du lieu de travail). Au-delà de la mesure précise de l'impact de l'IA au niveau environnemental, celui-ci ne pourra être positif que si une IA plus frugale voit le jour : ceci nécessite un effort de recherche que la communauté de l'apprentissage automatique commence tout juste à prendre en compte.

RENDEZ-VOUS DANS DIX ANS ?

Ainsi, pour que l'IA puisse bénéficier à toutes et tous, de nombreux verrous scientifiques devront être levés. Il est difficile et sans doute illusoire de vouloir faire des prédictions sur les performances des algorithmes dans dix ans. En revanche, il est désormais certain qu'un formidable effort de recherche est nécessaire, pour lequel une recherche académique bien financée est indispensable pour relever les défis techniques, mais aussi pour informer les citoyens sur les forces et les faiblesses de l'intelligence artificielle. ★

EN SAVOIR PLUS

Le site de l'Académie des sciences : WWW.ACADEMIE-SCIENCES.FR



« **Quand la machine apprend. La révolution des neurones artificiels et de l'apprentissage profond** », de Yann Le Cun. Odile Jacob, 2019.

En ligne, le dossier « **Comment l'intelligence artificielle va changer nos vies** » (2018), du « CNRS, le journal », regroupe tous ses contenus sur cette technologie : <https://lejournal.cnrs.fr/dossiers/comment-lintelligence-artificielle-va-changer-nos-vies>

« **Introduction au machine learning** », de Chloé-Agathe Azencott. Dunod, 2019.

Pourquoi les applications révolutionnaires promises, telle la voiture autonome, mettent-elles du temps à arriver ? Majoritairement parce que les connaissances actuelles en informatique et en mathématiques ne sont pas suffisantes.



GETTY IMAGES / ISTOCKPHOTO