



NOTES
D'ÉTONNEMENT

MAI 2019

RECHERCHE EN IA @ GOOGLE

L'IA, C'EST QUOI ? De manière générique l'on pourrait dire que l'intelligence artificielle (IA) c'est l'art et la science de rendre les choses intelligentes. Une partie « historique » de cette science, la plus naturellement comprise par le public, vise à reproduire l'intelligence humaine mais, à côté de cela, un champ de recherche immense prend aujourd'hui un plein essor : il s'agit du *machine learning*, en français l'apprentissage automatique. La démarche marque une rupture puisque, au lieu de chercher à comprendre et interpréter un phénomène pour pouvoir ensuite le répliquer, en recourant à la modélisation, nous allons chercher à construire un modèle capable d'apprendre par lui-même et de répliquer un phénomène ou un comportement à partir de la connaissance empirique que nous lui donnerons, en l'occurrence des données. Ici déjà, le traitement de données structurées fournit des résultats prodigieux et, si voir une machine apprendre à distinguer un chat d'un chien peut paraître anodin, il s'agit en réalité d'une véritable prouesse.

Qu'il s'agisse d'*apprentissage profond* ou d'*apprentissage fédéré*, fondamentalement le schéma du machine learning est le suivant, toujours le même : celui d'entrer une donnée dans une fonction mathématique extrêmement complexe pour obtenir une sortie. En résumé donc, l'IA c'est : une entrée → un calcul complexe → une sortie... Cela fait finalement moins figure d'intelligence qu'il n'y paraissait. La seule intelligence qui vaille ici est bien celle des humains qui ont conçu le modèle mathématique qui permet à la machine de réaliser ces prouesses.

POURQUOI AUJOURD'HUI CET ENGOUEMENT POUR L'IA ? La dernière décennie a vu une explosion des ressources de calcul et cette évolution purement technologique a élargi l'accès à des modèles très complexes qu'il était jusqu'à présent difficile de mettre en œuvre et de déployer. Il est d'ailleurs tout à fait symbolique de voir que le Turing Award 2018 a récompensé trois pionniers du deep learning : les Canadiens Geoffrey Hinton et Yoshua Bengio et le Français Yann Le Cun. Yoshua Bengio a contribué à développer les réseaux GAN (réseaux antagonistes génératifs), Yann Le Cun a été au cœur d'avancées fulgurantes pour la reconnaissance d'images, et Geoffrey Hinton a prouvé que l'apprentissage profond pouvait résoudre des problèmes insolubles par d'autres approches. Yoshua Bengio dirige aujourd'hui l'institut d'intelligence artificielle Mila, basé à Montréal ; Yann Le Cun dirige la recherche scientifique de Facebook en matière d'IA et Geoffrey Hinton poursuit ses recherches au sein de Google Brain.¹

¹ Les trois récipiendaires se verront remettre le prix le 15 juin 2019 à San Francisco.
<https://amturing.acm.org/>

“ *La vraie rupture viendra quand les machines seront devenues réellement intelligentes, ce qui suppose de trouver quelque chose dont on ne dispose pas encore aujourd’hui.* ”

Bertrand RONDEPIERRE

Research Program Manager – Google AI Europe

En charge du volet santé de la mission Villani sur l’intelligence artificielle

QUELLES AVANCÉES PROCHAINES ATTENDRE DE L’IA ? Les cinq prochaines années verront un enrichissement des répertoires de mises en œuvre de l’IA. Toutefois la vraie rupture viendra quand les machines seront devenues réellement intelligentes, ce qui suppose de « trouver » quelque chose dont on ne dispose pas encore aujourd’hui. L’on verra alors des applications que l’on estime aujourd’hui inatteignables.

LA RECHERCHE EN IA @ GOOGLE

Google est un acteur avant tout industriel mais qui souhaite rester pleinement engagé dans la recherche fondamentale. Cette recherche est distribuée le plus largement possible entre l’Amérique du Nord (à New-York et Toronto), l’Europe (en France, où Google a ouvert son centre de recherche sur l’IA à Paris en septembre 2018, après Londres et Zurich, et concomitamment à Berlin et Amsterdam), Israël (Haïfa), l’Asie (Taiwan et Beijing) et même l’Afrique avec l’ouverture en avril dernier d’un centre de recherche, toujours sur l’IA, à Accra (Ghana). Les équipes de Google parcourent le monde pour échanger et s’enrichir. Leurs spécialités recouvrent l’intégralité du spectre des problématiques suscitées par le développement de l’IA. Certaines de ces équipes, dont notamment celle de Paris, effectuent de la recherche fondamentale sans avoir en ligne de mire aucun produit ou aucune application définie.

Pendant ce temps, d’autres équipes travaillent plus à la limite de la recherche appliquée sur la recherche de nouveaux produits et de nouvelles activités mais pour lesquelles les technologies ne sont pas encore matures. C’est le cas, par exemple, dans le domaine médical, lequel offre d’immenses perspectives et où des recherches sont donc engagées sur un spectre très large d’applications. Une troisième partie de la recherche de Google en matière d’IA vise directement la contribution aux produits de l’entreprise qu’il s’agit de maintenir à l’échelle pour des milliards d’utilisateurs. Enfin, un quatrième et dernier pilier de la recherche chez Google concerne les infrastructures logicielles. Ces dernières doivent permettre d’engager un grand nombre d’expérimentations de façon très rapide et réactive sur des ressources de calcul particulièrement volumineuses.

En pratique, les projets de recherche ne sont jamais compartimentés dans l’une ou l’autre de ces catégories mais émergent dans chacune d’elles. Une fois qu’un produit a été développé, les solutions trouvées aux problématiques posées sont transférées aux diverses équipes pour élaborer d’autres produits ou éléments d’infrastructures. En réalité donc, tous les domaines communiquent. Le transfert de connaissances est d’autant plus facilité que ce sont parfois aussi les mêmes ingénieurs qui font de la recherche et qui font des produits.

MIEUX : chez Google, les chercheurs sont libres de travailler sur les problématiques qu'ils souhaitent. Ceci fait évidemment rêver beaucoup de jeunes ingénieurs et chercheurs. Google reçoit ainsi chaque année 2 millions de candidatures.² L'aspect incertain de la démarche pour l'entreprise est plus que compensé par la facilité donnée au transfert de connaissances et par le large spectre des applications qu'elle met ou entend mettre en œuvre. Il y aura toujours un moment où les fruits de n'importe quelle recherche pourront être récoltés.

Ainsi, par exemple, posons-nous LA QUESTION : quelle est la différence entre interpréter une image de chat et interpréter une image de rétine ?

Rappelons que la rétinopathie diabétique, qui est en France la première cause de cécité avant 65 ans, peut cependant être décelée très en amont à partir d'un fond d'œil.

LA RÉPONSE À NOTRE QUESTION EST : AUCUNE. Il ne fait pour la machine aucune différence pour traiter l'image d'un chat et celle d'une rétine. L'algorithme développé pour interpréter la première servira à interpréter la seconde. Ceci est évidemment une très bonne nouvelle pour les 400 et quelques millions de diabétiques de type 2 (ceux menacés par la rétinopathie diabétique) de par le monde.

Comme nous l'avons évoqué lors d'une précédente session,³ les équipes de Google ont annoncé l'été dernier que l'intelligence DeepMind est dorénavant capable de recommander la décision de référence pour 50 maladies oculaires avec 94 % de précision, comme le feraient les ophtalmologues des deux prestigieux centres londoniens impliqués dans l'étude, l'Institut d'ophtalmologie de l'University College et le Moorfields Eye Hospital.⁴ Pour mettre au point ce logiciel, les chercheurs ont « nourri » leur système avec près de 15 000 scanners rétinien de patients précédemment diagnostiqués. Mais avant cela, ils ont eu recours à une première séance d'apprentissage réalisée à l'aide de 877 images dont la segmentation avait été réalisée manuellement. La machine a ainsi appris à segmenter les images elle-même, c'est-à-dire à repérer les différentes couches rétinienne et à les caractériser. C'est à tel point qu'elle est aujourd'hui capable de préciser au seul examen de l'iris si ce dernier appartient à un homme ou à une femme, ce qu'ignorait la science jusqu'à présent.

Mais, au-delà même des maladies oculaires, la rétine change au cours de la vie et se modifie en fonction de facteurs tels que l'âge, la pression artérielle, le tabagisme, le taux de cholestérol ou l'obésité... qui sont tous facteurs de risque cardiovasculaire. L'intelligence artificielle est donc capable aujourd'hui de d'évaluer ce risque cardiovasculaire à partir de seules photographies de rétine.⁵ C'est une avancée une fois de plus extrêmement intéressante car cet examen est délicat à réaliser pour les médecins, qui doivent généralement le compléter par des tests sanguins.

Ainsi, une même recherche finit par servir et accélérer le développement de multiples applications.

² "6 Reasons Working at Google Isn't Right for Most People", J.T. O'Donnell, WorkItDaily / Inc.com, 2018
<https://www.inc.com/jt-odonnell/6-reasons-working-at-google-isnt-right-for-most-people.html>

³ Note d'étonnement de la session du club Boussole Nouvelle Vague du 21 mars 2019 avec Alkéos Michail.

⁴ L'étude a été publiée dans la revue *Nature Medecine*, août 2018.

⁵ Publication scientifique dans *Nature Biomedical Engineering*, mars 2018.

ACCÉLÉRER LA RECHERCHE EN IA

Au cours des dernières années, Google a publié de nombreux outils d'IA, dont des jeux de données, du code et des logiciels. L'une des contributions les plus importantes a été TensorFlow, développé à l'origine pour un usage interne par Google Brain, l'équipe de deep learning, et rendu disponible en open source à la fin de 2015. TensorFlow est une bibliothèque de machine learning : une boîte à outils permettant de résoudre des problèmes mathématiques extrêmement complexes. Elle permet aux chercheurs de développer des architectures expérimentales d'apprentissage et de les transformer en logiciels. Récemment, un nouveau module a été ajouté au framework : TensorFlow Privacy, qui protège les données des utilisateurs sur lesquels les modèles d'intelligence artificielle sont entraînés. Une seconde bibliothèque a aussi récemment été mise en open source, nommée GPipe et conçue par Google AI pour créer des réseaux neuronaux profonds (deep learning) à grande échelle, plus rapidement, plus simplement et potentiellement avec moins de puissance de calcul.

En faisant ce choix de l'open source, Google entend accélérer « collectivement » la recherche en IA, aux côtés de ses partenaires qui en IA sont essentiellement académiques.

Pour faire face aux besoins d'un si grand nombre de chercheurs internes ou externes travaillant sur l'IA, Google a choisi de développer son processeur « maison », baptisé non pas CPU mais TPU pour Tensor Processor Unit. Les pods de TPU de 3^e génération (machines comprenant un grand nombre de TPU) ont une puissance de calcul de plus de 100 pétaflops (10¹⁵ opérations par seconde), soit plus de 6 fois supérieure à celle du supercalculateur HPC-IA attendu au centre de calcul du CNRS à Saclay, faisant 15 pétaflops.

POSER UN CADRE ÉTHIQUE

L'IA est une technologie émergente et doit être développée dans un cadre éthique. Mais cela vaut pour l'intelligence artificielle comme pour n'importe quelle autre technologie émergente... avec bon sens.

PARTIE REMISE. Fin mars 2019, Google annonçait la création d'un comité dédié à l'éthique dans le domaine de l'intelligence artificielle. L'objectif de ce comité était d'encadrer les travaux de l'entreprise dans le domaine et de s'assurer que ces travaux suivent les grands « Principes IA » fixés par le CEO Sundar Pichai en 2018. Ces Principes IA avaient été établis après qu'il ait été révélé que Google participait à un projet de drones militaires autonomes initié par le Pentagone. Ce comité, appelé Advanced Technology External Advisory Council (ATEAC), a donc réuni dans une première « mouture » des chercheurs de renom issus de domaines allant de l'intelligence artificielle à la philosophie en passant par la psychologie et la robotique. Cependant, il incluait aussi des politiciens. Parmi eux, le président de l'Heritage Foundation, Kay Coles James, connu pour ses idées conservatrices et anti-LGBTQ. Sa présence au sein de l'ATEAC a semé la discorde et obligé Google à dissoudre le comité une semaine après sa mise en place. La firme compte à présent poursuivre ses efforts dans le domaine de l'IA de façon responsable et cherchera d'autres façons de recevoir des opinions extérieures sur les questions liées à l'éthique.

Source : Le BigData.fr, 5 avril



“ Dans les cinq années qui viennent, l’IA ne va probablement pas apporter beaucoup de nouvelles choses révolutionnaires, elle va, en revanche à coup sûr, nous proposer beaucoup de nouvelles choses découlant des innovations qui ont déjà montré leur potentiel.

Bertrand RONDEPIERRE

DANS CINQ ANS, L’IA ?... Le travail sur les données structurées, images, paroles, etc. va être poursuivi et amplifié. Les progrès que l’on peut attendre ici sont tels que les chercheurs ne vont sur le court terme (cinq ans) que peu sortir de ce périmètre. L’IA ne va donc probablement pas apporter beaucoup de nouvelles choses révolutionnaires, elle va, en revanche à coup sûr, nous proposer beaucoup de nouvelles choses découlant de ces innovations qui ont déjà montré leur potentiel. Nous avons évoqué l’exemple d’applications dans le domaine médical, s’y ajouteront les domaines de la sûreté et de la sécurité, et plus globalement tout ce qui concerne l’interaction avec les humains (agents conversationnels et autres systèmes interactionnels). Tous ces sujets verront aussi à coup sûr une augmentation de l’efficacité, de l’efficience (avec, par exemple, une progression du temps réel) et surtout de l’interactivité : nous pourrons sûrement avoir à ce moment un vrai échange avec nos assistants vocaux, au-delà des quelques ordres que nous pouvons dès à présent leur passer.

« Google, peux-tu m’arrêter l’enregistrement et m’envoyer tout cela au format texte sur mon PC ? Attention aux fautes et fais, s’il te plaît, un effort de style, c’est tout de même une note d’étonnement ! Merci. »

Martine LE BEC
rédactrice en chef de la revue *Prospective Stratégique* – CEPS