



MARC DURANTON CEA-List

L'APPRENTISSAGE PROFOND RENOUVELLE L'INTELLIGENCE ARTIFICIELLE

Ce n'est qu'à partir de la seconde moitié du ^{XX}e siècle que les machines ont vraiment amplifié les capacités mentales de l'homme. Mais l'histoire de l'intelligence artificielle a été assez chaotique et ce n'est qu'avec les succès récents des techniques d'apprentissage profond qu'elle connaît un essor aussi fort que prometteur.

REPÈRES

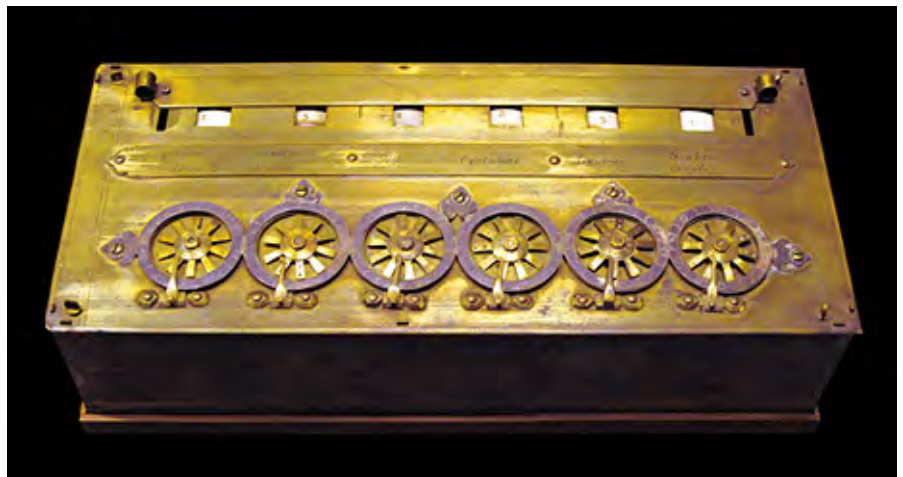
De tout temps, les hommes ont cherché à faire des machines amplifiant leurs capacités physiques, puis mentales. Le cerveau n'a pas toujours été le centre de l'intelligence : pour Aristote, il ne servait qu'à refroidir le cœur. Mais l'approche préconisée par Platon, Hippocrate, Démocrite pour lequel le cerveau est le centre de la prise de conscience des sensations et le gardien de l'intelligence a finalement prévalu et de nombreuses générations de chercheurs ont cherché, et cherchent encore à analyser son fonctionnement. L'idée de l'imiter pour faire des systèmes « intelligents » n'est pas neuve, mais ce sont les découvertes du ^{XX}e siècle qui ont déclenché les premiers résultats.

EN S'INSPIRANT du savoir des biologistes de leur époque, Warren Sturgis McCulloch, neurologue américain et Walter Pitts, mathématicien et psychologue ont proposé en 1943 un modèle mathématique du fonctionnement simplifié des neurones biologiques, cellules étant une des composantes du cerveau. Leur papier, *A Logical Calculus of Ideas Immanent in Nervous Activity*, a été publié en 1943 dans le *Bulletin of Mathematical Biophysics* (5:115-133) et reste la base des réseaux de neurones formels. Leur modèle est simple : un neurone effectue une fonction binaire qui compare à un seuil la somme pondérée de ses entrées (connectées aux autres neurones).

« Un réseau suffisamment complexe permet de "calculer" n'importe quelle fonction »

UNE PREMIÈRE DESCRIPTION DE L'ORDINATEUR MODERNE DÈS 1945

Ils ont montré qu'un réseau suffisamment complexe permet de « calculer » n'importe quelle fonction. John von Neumann, dont on considère que son *First Draft of a Report on the EDVAC* est la première description d'un ordinateur moderne (la machine de J. von Neumann), ne cite que ce papier de McCulloch et Pitts dans ce rapport de 1945 et déduit de l'article de McCulloch et Pitts que « tout ce qui peut être décrit de manière exhaustive et sans ambiguïté [...] peut être conçu comme un réseau neuronal approprié ». Il entérine qu'un réseau de neurones peut représenter une machine universelle de



De tout temps, les hommes ont cherché à faire des machines amplifiant leurs capacités.



© 2011 LOS ALAMOS NATIONAL SECURITY, LLC TOUS DROITS RÉSERVÉS

On considère que John von Neumann fut le premier à décrire un ordinateur moderne.

Turing, et donc un calculateur universel. Malheureusement, les limitations de la technologie de l'époque ne lui permirent pas de développer l'approche hautement parallèle des réseaux de neurones, et donc il aboutit à l'architecture *mémoire, unité de contrôle, unité arithmétique* et *unités d'entrée et de sortie*, que l'on retrouve dans les ordinateurs actuels.

1957 : UN ALGORITHME INSPIRÉ DES NEURONES

C'est en 1957 que la psychologue Frank Rosenblatt invente un algorithme baptisé Perceptron. Pour ce classifieur, la pondération entre neurones est inspirée par la règle de Hebb qui considère que lorsque deux neurones sont excités conjointement, leur lien se renforce. La règle du Perceptron tient compte de l'erreur observée lorsqu'on propage une entrée dont le Perceptron calcule la fonction de sortie. Le premier hiver des réseaux de neurones a été provoqué par le livre de Marvin Minsky et Seymour Papert, *Perceptrons: an introduction to computational geometry*, qui montre des limitations des Perceptrons. Le livre, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* de David Everett Rumelhart et James McClelland, relance le domaine avec une approche testable des réseaux multicouches (essentiellement avec une couche intermédiaire,

UN NOUVEL HIVER POUR LES RÉSEAUX DE NEURONES

Les machines à vecteurs de support (en anglais *Support Vector Machine*, SVM) ont sonné le début d'un nouvel hiver des réseaux de neurones en offrant de meilleures performances que les MLP (*Multi-Layer-Perceptrons*) pour la classification d'images. Les principes furent explorés entre 1963 et 1970 par Vladimir Vapnik, mais ce n'est qu'en 1992 que l'article de Boser, Guyon et Vapnik synthétise les résultats et permet un large développement des SVM pour la classification.

dite cachée) dits *Multi-Layer-Perceptrons* (MLP).

UN REGAIN D'INTÉRÊT APRÈS 1985

Une règle d'apprentissage, appelée rétropropagation du gradient, permettant de calculer les poids des couches intermédiaires a été publiée dans sa thèse en 1985 par Yann LeCun (maintenant chez Facebook), puis largement diffusée par David Rumelhart, Geoffrey Hinton (maintenant chez Google Brain) et Ronald Williams en 1986. Cela a abouti à une première explosion des utilisations des réseaux de neurones dans les années 1990. Ils furent utilisés pour la reconnaissance de caractères manuscrits (pour reconnaître les codes postaux), pour de l'analyse d'images, etc. Une première ère de développements de

circuits spécialisés a suivi, mais les techniques de l'époque ne permettent qu'un parallélisme limité, et l'avancée rapide des processeurs généralistes a limité leur expansion.

RÉTROPROPAGATION ET RÉSEAUX CONVOLUTIONNELS

Pendant ce temps, les réseaux de neurones deviennent plus profonds (avec plus de couches), grâce à des méthodes permettant d'utiliser les approches de la rétropropagation du gradient à des réseaux avec plus d'une couche cachée. Les réseaux se complexifèrent, en spécialisant les couches comme dans le cortex visuel. Les résultats du neuroscientifique David Marr et de David Hubel et Torsten Wiesel (tous deux prix Nobel en 1981 pour leurs découvertes concernant le traitement de l'informa-

« Les techniques de l'époque ne permettent qu'un parallélisme limité »



Torsten Wiesel a reçu le prix Nobel 1981 avec David Hubel pour leurs découvertes sur le traitement de l'information visuelle.

DOSSIER

tion dans le système visuel) inspirèrent les chercheurs pour faire des réseaux plus adaptés à la reconnaissance d'objets. Leur ancêtre est le *Neocognitron* inventé dans les années 1980 par Kunihiko Fukushima. Les réseaux convolutionnels profonds tels qu'utilisés actuellement ont plus de vingt ans, mais « le récent succès pratique de l'apprentissage profond [...] est dû en grande partie à l'augmentation spectaculaire de la taille des ensembles de données et de la puissance des ordinateurs [...], ce qui nous a permis de former des réseaux gigantesques [...] ».

2012, ANNÉE DU RENOUVEAU

Le renouveau a été provoqué par Alex Krizhevsky, Ilya Sutskever et Geoffrey Hinton en 2012 qui utilisèrent des réseaux de neurones convolutionnels profonds pour le challenge ImageNet, qui consiste à classer des images dans la base de données d'images ImageNet. Le réseau *Supervision* de Hinton bat les autres approches avec un taux d'erreur de 15,3 % contre 26,1 % pour le second. Dès 2013, les 8 premiers du challenge sont à base de réseaux de neurones profonds.

« *Les réseaux profonds sont meilleurs qu'un humain sur certains challenges* »

Depuis, les réseaux profonds sont meilleurs qu'un humain sur ce challenge, avec moins de 3,5 % d'erreurs. Le tableau suivant montre l'amélioration très rapide des algorithmes d'apprentissage profond, jusqu'à être meilleurs que les humains.

UN CHAMP D'APPLICATION DE PLUS EN PLUS LARGE

Grâce au fait qu'un réseau profond est formé par apprentissage et non explicitement programmé, il est appliqué dans de nombreuses applications où il est difficile de définir un algorithme, comme la reconnaissance d'image (indispensable pour les véhicules autonomes), la compréhension de la parole (tous les assistants personnels, de Siri à Alexa ou Google Now, utilisent des réseaux profonds souvent récurrents), la lecture sur les lèvres et la participation à divers jeux. Une grande base de données « labellisée » (indexée) est tout ce qui est nécessaire; celles-ci sont souvent disponibles auprès des grands acteurs d'Internet (Google, Baidu, Facebook, Microsoft, Apple, etc.), expliquant pourquoi ils mènent les recherches et les utilisations de l'apprentissage en profondeur. Par



© SERGEY

AlphaGo peut battre les meilleurs joueurs de go.

exemple, plus de 2 milliards de photos passent chaque jour à travers deux types de réseaux profonds chez Facebook, Instagram, Messenger, WhatsApp pour la reconnaissance d'image/l'indexation et pour la reconnaissance de visages (mais pas en Europe).

Les réseaux et les techniques se complexifient, en combinant plusieurs approches. Par exemple pour le programme AlphaGo développé par Google DeepMind qui a battu Lee Sedol (un professionnel 9-dan dans le jeu de go) en mars 2016, générant beaucoup de publicité pour l'apprentissage en profondeur et les techniques d'IA.

UNE UTILISATION EN DEUX PHASES

En général, il existe deux phases dans l'utilisation des réseaux profonds : la phase d'apprentissage, dans laquelle les paramètres du réseau (topologie et poids des connexions) sont déterminés par la règle d'apprentissage et la phase d'inférence dans laquelle le réseau est utilisé pour classer les données. La phase d'apprentissage est la plus exigeante, avec des millions ou des milliards de présentations d'exemples et des modifications des paramètres du réseau. Il est maintenant généralement fait sur les GPU en virgule flottante 16 bits ou sur ces circuits spé-

Nom de l'algorithme	Date	Erreur sur le jeu de test
Supervision	2012	15,3 %
Clarifai	2013	11,7 %
GoogleNet	2014	6,66 %
Niveau humain		5 %
Microsoft	05/02/2015	4,94 %
Google	02/03/2015	4,82 %
Baidu/Deep Image	10/05/2015	4,58 %
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences CNN à 152 couches	10/12/2015	3,57 %
Google Inception-v3	2015	3,5 %
	Maintenant	?

DES CENTAINES DE MILLIERS DE NEURONES ARTIFICIELS

À titre d'exemple, *Supervision* (le réseau de G. Hinton) est composée de 650 000 neurones artificiels connectés par 630 000 000 connexions partagées (synapses). Un apprentissage d'un réseau actuel demande quelques exaflops (plus d'un milliard de milliards d'opérations).

cialisés comme les *Tensor Processing Units* (TPU) de Google. La phase d'inférence est moins exigeante et peut être effectuée avec moins de précision (en nombre entier, même réduit à 8 bits). C'est généralement cette phase qui est implémentée dans des dispositifs embarqués pour la reconnaissance d'image, etc. Les poids synaptiques sont téléchargés après apprentissage et peuvent être mis à jour après un nouvel apprentissage, étendant le nombre d'objets reconnus.

PLUSIEURS MODES D'APPRENTISSAGE

Il existe un grand nombre d'approches pour la phase d'apprentissage, mais elles peuvent être classées en 3 grandes classes : l'apprentissage supervisé (présentation lors de l'apprentissage des entrées ET des résultats souhaités correspondant à la classe particulière de l'apport présenté) ;

l'apprentissage non supervisé (le réseau détermine sa sortie à partir de différentes entrées qui n'ont alors pas besoin d'être labellisées et essaie de discriminer les entrées dans différentes classes) ; l'apprentissage par renforcement qui se concentre sur la prédiction d'une récompense. C'est ce type d'apprentissage qui a été utilisé pour former le programme AlphaGo et ses suc-

cesseurs, comme Alpha Zero, qui, en quelques heures, et sans connaissance du domaine sauf les règles, bat tous ses prédécesseurs au jeu de go, mais aussi aux échecs. D'autres approches sont en développement, comme les *Generative adversarial networks* (GAN) qui mettent des réseaux en compétition. On commence même à voir apparaître des recherches utilisant des approches d'apprentissage profond pour créer d'autres réseaux d'apprentissage profond plus optimisés.

« *Les investissements relatifs à l'intelligence artificielle sont supposés atteindre près de 11 milliards de dollars en 2025* »

UNE OFFRE D'OUTILS QUI S'ÉTOFFE

Les grands acteurs du domaine fournissent leurs outils de développement de réseau profond comme logiciel libre, comme TensorFlow (Google), CNTK (Microsoft), DSSTNE (Amazon), Theano, Caffe (Berkeley), Torch (Facebook avec *open source*), N2D2 (CEA), les modules d'apprentissage Torchnet, OpenAI Gym (de Open AI), MXNet, etc. En fait, le logiciel est un élément non crucial pour créer un système efficace d'apprentissage en profondeur. Une grande base de données et la topologie des réseaux neuronaux sont les principaux ingrédients : la valeur réside dans la topologie du réseau neuronal et ses poids, déterminés après l'apprentissage sur une base de données particulière.

DES INVESTISSEMENTS MASSIFS POUR DES TRANSFORMATIONS PROFONDES

Dans le domaine économique, les investissements relatifs à l'intelligence artificielle sont supposés atteindre près de 110 milliards d'euros en 2025. Beaucoup de start-up travaillant dans le domaine de l'IA ont récemment été acquises par de grandes entreprises. Par exemple, en 2014, Google a acheté DeepMind au Royaume-Uni (la société qui a créé AlphaGo et Alpha Zero), tandis qu'en 2016 Intel a acheté Movidius en Irlande et aux États-Unis (spécialisé dans les systèmes de vision basse consommation, utilisés par exemple dans les drones) et Nervana. Au total, Google, IBM, Yahoo, Intel, Apple et Salesforce ont acquis au cours des cinq dernières années plus de 30 entreprises travaillant sur l'IA.

Des scientifiques bien connus et les grandes entreprises investissent énormément en IA et sur l'apprentissage profond, et des pays comme les États-Unis, la Chine et le Japon lancent de grands projets d'IA et ont la conviction que de nouvelles percées se produiront et que cela aura certainement un impact profond sur notre société dans les années à venir. Le président Obama a dit : « Mon successeur gouvernera un pays transformé par l'IA », montrant l'impact que pourrait avoir l'IA dans le futur. ■



© ALEXTYPE

Intel a acheté Movidius en Irlande et aux États-Unis (spécialisé dans les systèmes de vision basse consommation, utilisés par exemple dans les drones).