

D'un gratte-ciel, nous attendons qu'il respecte un cahier des charges, avec pour *minimum minimorum* de supporter son propre poids et de résister à un séisme. Rien de tel avec l'une des technologies les plus importantes du monde moderne. Avec elle, nous avançons à l'aveuglette, bricolons des configurations, mais tant qu'elle n'a pas été testée, nous ne savons pas vraiment ce qu'elle peut faire, ni où elle échouera. Cette technologie est le réseau de neurones artificiels ; c'est lui qui soutient les systèmes d'intelligence artificielle les plus avancés à ce jour. De plus en plus profondément intégré dans notre société, il filtre le flux des réseaux sociaux que nous recevons, aide les médecins à diagnostiquer des maladies, et influence même la durée d'emprisonnement d'un justiciable condamné pour crime.

En réalité, selon le mathématicien Boris Hanin, «la meilleure approximation de ce que nous savons est que... nous ne savons presque rien à propos du fonctionnement réel des réseaux neuronaux, ni ce que dirait une théorie vraiment pertinente sur le sujet !» Cet enseignant à l'université de Princeton aime l'analogie avec une autre technologie révolutionnaire : la machine à vapeur. De prime abord, ces machines n'étaient bonnes qu'à pomper de l'eau. Puis, elles ont actionné des locomotives, ce qui correspond peut-être au degré de sophistication atteint aujourd'hui par les réseaux neuronaux. Enfin, les scientifiques et les mathématiciens, grâce à la théorie de la thermodynamique, ont pu comprendre exactement ce qu'il se passe à l'intérieur de n'importe quelle machine. Au final, cette technologie nous a conduits sur la

### — En bref —

> Les réseaux neuronaux ont une architecture inspirée du cerveau sous forme de couches successives de « neurones » reliées entre elles.

> Adapter la profondeur et la largeur des couches est encore très empirique. Il faut beaucoup d'essais et erreurs pour réussir.

> L'un des objectifs majeurs de la discipline est d'établir une théorie sur le lien entre telle forme et tel résultat. Cela nécessite, par exemple, de déterminer la largeur minimale d'une couche.

Lune. «D'abord, une belle ingénierie, puis d'excellents trains, puis un peu de compréhension théorique pour en arriver aux fusées», résume Boris Hanin.

Au sein de la communauté tentaculaire des chercheurs en intelligence artificielle, pourtant, un petit groupe d'individus à l'esprit bouillonnant de maths essaye de mettre au point une théorie capable d'expliquer leur fonctionnement et de garantir que, si vous construisez tel réseau neuronal en suivant telle méthode prescrite, il exécutera telles tâches précises. Ce travail n'en est qu'à ses balbutiements, mais, en 2018, plusieurs études ont détaillé la relation entre forme et fonction dans les réseaux neuronaux. L'objectif est d'en définir les fondations mêmes. Il montre que, bien avant de pouvoir certifier qu'ils peuvent conduire des voitures, vous devrez prouver que les réseaux neuronaux sont aptes à réussir une simple multiplication. Voilà qui demande quelques explications.

### UN CHIEN, C'EST QUATRE PATTES POILUES

Les réseaux neuronaux visent à imiter le cerveau humain, organe que l'on peut représenter comme une structure qui avance au fil d'abstractions de plus en plus larges. Dès lors, la complexité de la pensée est mesurée par la palette d'abstractions de base sur lesquelles vous pouvez vous appuyer, et le nombre de fois où vous pouvez les combiner en de nouvelles de plus haut niveau – c'est ainsi que nous apprenons à distinguer les chiens des oiseaux.

«Comme être humain, apprendre à reconnaître un chien revient à savoir identifier quatre pattes poilues, analyse Maithra Raghu, chercheuse chez Google Brain. Idéalement, nous aimerions que nos réseaux neuronaux procèdent de la même manière.» L'abstraction vient naturellement au cerveau humain. Pas aux réseaux neuronaux. Ils sont faits de blocs de construction appelés «neurones», connectés de diverses manières. Chacun peut représenter un attribut, ou une combinaison d'attributs, que le réseau prend en compte à chaque niveau d'abstraction. Quand ils relient ces neurones, les ingénieurs doivent faire de nombreux choix, en particulier décider de combien de couches de neurones le réseau doit disposer, autrement dit, quelle est sa «profondeur». Imaginez que l'enjeu soit de reconnaître des objets dans des images. L'image entre dans le système à la première couche. À la couche suivante, le réseau peut avoir des neurones qui détectent simplement les contours dans l'image. Puis la suivante combine les courbes en formes et textures, et la dernière traite formes et textures pour conclure sur la nature de l'objet : c'est un mammouth laineux !

Chaque couche combine plusieurs aspects de la précédente. Ainsi, un cercle peut être vu comme une combinaison de courbes, une courbe comme une combinaison de segments. Les ingénieurs doivent aussi décider de la «largeur» de chaque couche, laquelle correspond au nombre de caractéristiques différentes que le réseau prend en compte à chaque niveau d'abstraction. Dans le cas de la reconnaissance d'image, cette largeur peut être le nombre de types de segments, de courbes et de formes considérés.



*Cet article a d'abord été publié en anglais par **Quanta Magazine**, une publication en ligne indépendante soutenue par la Simons Foundation afin de favoriser la diffusion des sciences : [bit.ly/36dMtRh](http://bit.ly/36dMtRh)*

Au-delà de la profondeur et de la largeur d'un réseau, il faut aussi choisir comment câbler les neurones dans et entre les couches, et quel poids donner à chaque connexion. Dans ces conditions, une fois l'objectif fixé, comment savoir quelle architecture de réseau neuronal l'atteindra au mieux ? Quelques règles générales se dégagent. Pour les tâches impliquant des images, les ingénieurs utilisent des réseaux neuronaux «convolutifs», qui offrent le même motif de connexions entre les couches, répété encore et encore. Pour le traitement du langage naturel – reconnaissance automatique de la parole ou génération de textes –, les réseaux neuronaux «récurrents» semblent mieux fonctionner : les neurones peuvent y être connectés à des couches non adjacentes.

## LA RÉUSSITE COMME PRODUIT D'ESSAIS ET D'ERREURS

Au-delà de quelques lignes directrices, les ingénieurs n'ont d'autre choix que d'obtenir des preuves expérimentales : ils lancent 1 000 réseaux neuronaux différents et observent simplement lequel fait le boulot. «Leurs choix sont le produit d'essais et d'erreurs, constate Boris Hanin. C'est dur, car il y a un nombre infini de choix possibles et nul ne sait quel est le meilleur.»

Il conviendrait de réduire la quantité d'essais et d'erreurs et de mieux comprendre au préalable ce que l'architecture d'un réseau neuronal peut apporter. Quelques études récentes font avancer dans cette direction. Selon David Rolnick, aujourd'hui professeur de sciences