**Le responsable de l’Intelligence artificielle chez Facebook, Yann LeCun, veut débrider le « deep learning » et rendre les machines encore plus intelligentes.**

L’intelligence artificielle est passé à travers une période sombre, qualifiée d’  «  hiver de l’IA ». Ce n’est plus le cas aujourd’hui. L’IA a atteint un tel niveau que des géants tels que Google, Facebook, Apple, Baido ou Microsoft se disputent les chercheurs les plus compétents dans ce domaine. L’attrait actuel pour l’IA tient principalement aux avancées révolutionnaires apportées par les réseaux neuronaux à convolution. Cette technique d’apprentissage automatique devrait apporter des progrés majeurs dans des domaines tels que la reconnaissance optique, la reconnaissance vocal et le traitement du langage naturel. Vous en avez sans doute entendu parler sous le terme de « deep learning ».

Peu de gens ont été associé d’aussi près au « deep learning » que Yann LeCun. En travaillant comme chercheur au Bell Labs à la fin des années 1980, LeCun a développé la technique des réseaux neuronaux à convolution, et montré comment elle pouvait être utilisé pour améliorer de façon significative la reconnaissance de l’écriture manuscrite. La plupart des chèques émis aujourd’hui aux États-Unis sont traités en utilisant cette technique. Entre le milieu des années 1990 et la fin des années 2000, quand les réseaux neuronaux n’avaient plus la faveur des scientifiques, LeCun a fait partie de la poignée de chercheurs qui ont persévéré dans ce domaine. Devenu professeur à l’Université de New-York en 2003, il a eu un rôle majeur dans les avancées du « deep learning ».

Le « deep learning » et les techniques associés sont maintenant un des domaines les plus actifs de la recherche en informatique. C’est la raison pour laquelle en 2013, LeCun a été engagé pour diriger le tout nouveau Laboratoire de Recherche en Intelligence Artificielle de Facebook, tout en poursuivant son enseignement à l’Université de New-York.

Né en France, LeCun garde de ses origines la sensibilisation au rôle des « intellectuels ». Il publie et donne des conférences dans son domaine principal, mais ne rechigne pas à intervenir également dans d’autres domaines.

1. **Expliquer le « deep learning » en huit mots**

***IEEE Spectrum*:** **Les mots « Deep Learning » apparaissent ces jours dans de nombreuses publications. Quelle serait la définition proposée récemment dans les journaux qui aurait le moins votre accord ?**

**Yann LeCun:**  La définition avec laquelle je suis le plus en désaccord est « ça fonctionne tout juste comme le cerveau ». Je n’apprécie pas cette façon de présenter ce concept, car si le « deep learning » tire son inspiration de la biologie, cette technologie est très éloignée du fonctionnement réel du cerveau. Et la décrire comme un « cerveau » lui donne une aura un peu magique, ce qui n’est pas sans danger, et ressemble un peu trop à du battage médiatique. Les gens énoncent des choses qui ne sont pas vrai. L’Intelligence Artificielle vient de connaitre une période sombre parce que des gens ont annoncé des choses qu’ils n’ont pas pu réaliser.

***Spectrum*: Si vous étiez un reporter chargé de présenter le « deep learning », et n’aviez que huit mots pour cela, comme c’est souvent le cas en journalisme, lesquels choisiriez vous**

**LeCun:**  Laissez moi le temps d’y réfléchir … Je pense que l’on pourrait utiliser cette formule : « Machines qui apprennent à représenter le monde ». Cela fait juste huit mots. On pourrait aussi dire « Apprentissage global pour les machines », mais cela ne fait que cinq mots, et il faudrait un peu développer … L’idée étant que chaque composant d’une « machine apprenante » peut être éduqué.

 ***Spectrum*: Votre éditeur ne va pas trop aimer ça ….**

**LeCun:**  Je pense surtout que le grand public ne comprendrait pas cette formulation. Essyons autre chose : on pourrait décrire le « deep learning » comme la construction d’une machines apprenante, disons un système de reconnaissance de structure ou quelque chose de semblable, par l’assemblage de nombreux modules ou éléments qui sont tous éduqués de la même façon. Il y aurait donc un principe unique d’apprentissage pour tous les modules. Mais je crois que nous avons dépassé les huit mots …

***Spectrum*: Que peut faire le « deep learning » que les autres machines apprenantes ne peuvent pas faire ?**

**LeCun:** Voilà une façon plus intéressante de poser la question. Les systèmes précédents, que l’on pourrait appeler « systèmes d’apprentissage superficiel » ( NdT : par opposition au deep learning, qui signifie « apprentissage en profondeur » ), étaient limités dans leurs capacités à calculer des fonctions complexes. Imaginons que l’on code un algorithme de reconnaissance d’image en utilisant la classification linéaire dont sont capables les « systèmes d’apprentissage superficiels ». Il faudra pour cela l’alimenter avec une liste de particularités extraites de l’image. Mais construire une liste de particularité « à la main » est extrêmement difficile, et coûteux en temps.

Une autre possibilité serait d’utiliser une méthode de classification plus souple, telle qu’une machine vectorielle, ou un réseau neuronal à deux couches alimenté directement par les pixels de l’image. Mais un tel système ne serait pas capable de reconnaitre les formes avec précisions, à moins de le construire si gigantesque qu’il en deviendrait inutilisable.

***Spectrum*: Ces explications ne sont pas très faciles à comprendre. C’est sans doute pour cela que les journalistes qui doivent décrirent le « deep learning » préférent dire …**

**LeCun:** ( NdT : terminant la phrase de l’interviewer ) Ca fonctionne comme le cerveau !

1. **Une boite noire avec 500 millions de boutons.**

***Spectrum*:** **Une partie du problème est que la notion d’apprentissage automatique est de façon suprenante inaccessible aux personnes qui ne travaillent pas dans ce domaine. Des tas de gens arrivent à plus ou moins comprendre des termes techniques utilisés en informatiques, par exemple l’algorithme de classement des pages utilisé par Google. Mais je pense que seuls les professionnels comprendront un article détaillé sur les classificateurs linéaires ou les machines vectorielles. Serait-ce dû à une complexité ontologique dont serait affecté ce domaine ?**

**LeCun:** En fait, je crois que les bases de l’apprentissage automatique sont tout à fait simple à comprendre. J’ai expliqué ce concept à des étudiants de grandes écoles et à leur professeur, et aucun ne s’est endormi !

Imaginons une boîte avec 500 millions de boutons, mille ampoules électriques, et 10 millions d’images à lui faire ingurgiter. Voilà un bon exemple de ce que peut être le « deep learning ».

Un système de reconnaissance de formes est comme une boîte noire équipée d’une caméra, d’un voyant vert et d’un voyant rouge, et de pas mal de boutons.

L’algoritme d’apprentissage tâche d’ajuster les boutons pour avoir le voyant rouge allumé si c’est un chien qui se trouve devant la caméra, et le voyant vert si c’est une voiture.

Après avoir montré un chien à la machine, vérifiez l’état des voyants. Si le voyant rouge est brillant, n’intervenez pas. Si l’ampoule ne brille que faiblement, tournez les boutons pour obtenir une lumière plus vive. Si le voyant vert s’allume, tournez les boutons pour affaiblir sa luminosité. Ensuite, montrez une voiture à la machine, puis tournez les boutons pour affaiblir la lumière rouge et aviver la verte. Si vous répétez l’expérience de nombreuses fois, en ajustant à chaque fois un peu moins les boutons, la machine finira par fournir à chaque fois la bonne réponse.

La chose intéressante est qu’elle sera également capable de reconnaître comme « chien » ou « voiture » un objet qui ne lui a pas encore été présenté. L’astuce est de calculer dans quel sens il faut tourner chaque bouton et avec quelle finesse, sans les manipuler dans tous les sens. Cela implique de calculer pour chaque bouton un gradient indiquant les variations induites dans les voyants quand le bouton est tourné.

Maintenant, imaginons une boîte avec cinq cent millions de boutons, mille voyants et dix millions d’images à lui faire classifier. Cela permet de se faire une idée de ce qu’est un système de « deep learning ».

***Spectrum*: Vous utilisez sans doute le terme « apprentissage superficiel » d’une façon ironique. Je ne pense pas que les personnes qui travaillent avec des classificateurs linéaires considèrent leur travail « superficiel ». Est-ce que l’expression « deep learning » ne contiendrait pas une part de « communication grand public », en montrant le mot « profond ( deep ) » comme qualifiant ce qui est appris, alors qu’en réalité « profond » décrit le nombre d’étapes parcourues par le système ?**

**LeCun:** Oui, c’est un peu ironique, mais cela reflète une réalité. Les « systèmes d’apprentissage superficiel » ont une ou au maximum deux couches, quand un système de « deep learning » est couramment constitué de cinq et jusqu’à vingt couches. Ce n’est pas l’apprentissage qui est superficielle ou profond, mais l’architecture utilisée.

1. **A la poursuite des belles idées ( un peu de passion pour l’informatique est recommandé )**

 ***Spectrum*: Votre biographie officielle dit que vous avez exploré de nouvelles approches pour les réseaux neuronaux à une période où ceux-ci étaient tombés en défaveur. Qu’est ce qui vous a poussé, au mépris de toute sagesse, à persévérer ?**

**LeCun:** J’ai toujours été attiré par l’idée d’ « éduquer  » un système de bout en bout. Vous informez un système avec des données brutes, et parce que ce système est multi-couches, chaque couche va comprendre comment transformer le résultat du travail de la couche précédente, ce qui fait que la couche finale va fournir la réponse. Cette idée, qu’il soit possible de programmer un apprentissage global qui permette à une machine d’obtenir une représentation correcte des données, est ce qui m’a obsédé durant trente ans.

***Spectrum*: Votre travail est-il juste une passion pour l’informatique, ou est-ce une science ? Est-ce que vous essayez les choses de façon empirique jusqu’à obtenir un résultat, ou y a-t-il une théorie qui sous-tend vos travaux ?**

**LeCun:**  On pourrait plutôt parler d’interaction entre intuitions, modélisations théoriques, mise en œuvre pratique, empirisme, et analyses scientifiques. L’intuition est de la pensée créative, la modélisation est de la mathématique, la mise en œuvre est de l’ingénierie et le tout ensemble devient de la science. Ce que j’aime surtout est transformer des idées théoriques, dans leur beauté, en un système qui fonctionne.

J’ai peu de sympathie pour les gens qui théorisent tout, par facilité, en ignorant d’autres méthodes de recherche qui fonctionnent également, y compris l’empirisme, parce qu’ils sont incapables de les théoriser. On retrouve un peu de tout ça dans la communauté qui gravite autour de l’apprentissage automatique. D’une certaine façon, la période sombre qu’a connu le concept de réseaux neuronaux vers les années 2000 est une conséquence de cette philosophie, consistant à révérer la « théorie », au mépris des résultats obtenus empiriquement. Ce n’est pas de cette façon que l’on résoud les problèmes d’ingénierie.

Je disais précédemment que e que j’aime surtout est transformer des idées théoriques, dans leur beauté, en un système qui fonctionne. Mais il y a aussi des dangers dans une approche purement théorique. La reconnaissance vocale, par exemple, a toujours été traitée de façon empirique, en ce sens que seuls ont retenu l’attention les résultats de tests. Et cela étouffe la créativité, parce que pour parvenir au même résultat que des équipes qui travaillent depuis longtemps sur un domaine, vous devez reparcourir tout leur cheminement en utilisant une durée équivalente, et reconstruire votre propre infrastructure. C’est à la fois difficile et dangereux, et personne ne s’y hasarde. C’est pour cette raison que dans le domaine de la reconnaissance vocale, les progrès, pour avoir été continus, n’ont été que très lents, du moins jusqu’à l’émergence du « deep learning ».

***Spectrum*: Vous semblez vouloir prendre une certaine distance avec le monde des neuro-sciences et de la biologie. Par exemple vous parlez de « réseaux à convolution » et pas de « réseaus neuronaux à convolution ». Ou bien, pour décrire vos algorithmes, vous parlez d’“unités” et pas de “neurones.”**

**LeCun:** C’est exact. Certains aspects de nos moèles sont inspirés des neuro-sciences, mais beaucoup de composant ne le sont pas du tout, et tirent au contraire leur inspiration de la théorie, de l’intuition, ou d’une approche empirique. Nos modèles ne prétendent pas se calquer sur le modèle du cerveau, et nous ne prétendons pas imiter le comportement des neurones. Mais d’un autre côté, je ne rejette pas l’idée que l’architecture des réseaux à convolution est inspirée par une connaissance de base du fonctionnement du cortex visuel. Certaines personnes utilisent des concepts qui viennent des neuro-sciences, mais refusent de l’admettre. Pour ce qui est de moi, je n’ai pas de problèmes à le reconnaitre. L’apport des neuro-sciences a été très utile.

Mais je suis attentif à ne pas utiliser des mots qui donnent lieu à du battage médiatique. Parce que dans ce domaine, il y a déjà trop de battage médiatique, et que c’est dangereux.

1. **Publicité et pseudo-science.**

 ***Spectrum*:** **Le battage médiatique n’est pas une bonne chose, mais pourquoi dites vous qu’il est dangereux ?**

**LeCun:** Cela crée pour les donateurs, pour le public, pour les clients potentiels, pour les investisseurs, des espérances qui ne sont pas fondées, comme de croire que nous sommes sur le point de construire des systèmes d’une puissance équivalente à celle du cerveau, alors que c’est loin d’être le cas. Et cela risquerait aisément d’amener une autre « période noire » pour notre activité.

De plus, pour reprendre l’expression de Richard Feynman, Cela fait penser à « la science culte de l’avion cargo ». Richard Feynman emploie cette expression pour décrire des choses qui ressemblent à de la science, mais n’en sont pas.

***Spectrum*: Pouvez vous donner un exemple ?**

**LeCun:** Dans le culte du cargo, on copie l’apparence d’une machine sans comprendre les principes qui font fonctionner cette même machine. C’est comme construire une station de radio avec des bottes de paille. Une telle approche de l’aéronautique même à copier très préciséement un oiseau, plumes, ailes battantes, et tout le reste. C’est ce que les gens ont fait pour tenter de construire des machines volantes, jusqu’à la fin du XIXème siècle, mais sans trop de succés.

L’équivalent en intelligence artificielle est de copier chaque détail connu du fonctionnement des neurones et des synapses, puis de mettre le résultat en œuvre dans un « réseau neuronal » géant à l’aide d’un super-calculateur, et d’espérer que de l’intelligence en sorte. Il y a des gens très sérieux et très fortunés qui sont prêts à croire ce genre de choses.

***Spectrum*: Pensez vous que le projet IBM « True North » soit de la « science culte de l’avion cargo » ?**

**LeCun:** Ce serait peut-être un peu exagéré de dire ça. Mais je pense que quelques annonces d’IBM ont été un peu trop loin et peuvent facilement être mal interprétées.

Leurs annonces peuvent paraitre impressionnantes, mais ils n’implémentent en réalité rien de concret. Avant le projet True North, IMB avait utilisé un super-calculateur pour simuler un cerveau de rat. Mais c’était juste un réseau neuronal aléatoire qui n’a rien produit d’utile.

Une puce basée sur un réseau à convolution pourrait être utilisée dans pas mal de dispositif. Mais ce qu’a tenté IBM ne conduit à rien d’utilisable.

Ce que je trouve attristant dans le projet Blue North est qu’il aurait pu être beaucoup plus efficace si on n’avait pas tenté d’imiter de si près la biologie. Créer une puce informatique est très coûteux. J’ai aussi été un concepteur de puces, ce qui me permet de penser qu’on ne devrait créer une nouvelle puce que si on est VRAIMENT sûr qu’elle va apporter quelque chose de neuf. Une puce basée sur un réseau à convolution – et la technologie pour la fabriquer est maintenant connue – pourrait dès à présent rendre service dans de nombreux domaines. Mais ce que vient de faire IBM ne mène à rien.

***Spectrum*: D’autres exemples ?**

**LeCun:** Je vais me faire des ennemis en disant ça, mais une grosse partie du projet Européen « Human Brain Project » est basée sur l’idée qu’en construisant des puces informatiques qui reproduisent au plus près le fonctionnement des neurones, puis en les mettant en œuvre dans un super-calculateur, après avoir mis celui-ci en route et lui avoir fourni un algorithme d’apprentissage automatique, l’intelligence devrait en émerger … Foutaises !

Bon, je reconnais que ma description est un peu caricaturale. Et je ne voudrais pas critiquer toute personne qui serait impliquée dans ce projet. Après tout, beaucoup de gens qui y travaillent le font parce que c’est une source de revenus dont ils ne peuvent pas se passer …

1. **L’apprentissage non dirigé : celui qu’il faut aux machines.**

***Spectrum*:** Reste-t-il encore beaucoup à découvrir sur l’apprentissage automatique **?**

**LeCun:** Énormément. Le genre d’apprentissage que nous utilisons dans les systèmes actuels de type « Deep learning » est très restreint. Ce qui fonctionne dans ces systèmes est l’apprentissage « dirigé ». Nous montrons une image au système, nous lui enseignons que c’est une voiture, et il ajuste ses paramètres pour être capable de reconnaître une voiture la fois suivante. Puis nous faisons la même chose avec une chaise, puis avec un être humain. Et avec quelques millions d’exemples, et de nombreux jours ou de nombreuses semaines de calcul, selon la taille du système, il est capable de fournir une réponse correcte.

Mais les êtres humains ou les animaux n’apprennent pas de cette façon. On ne vous a pas enseigné le nom de chaque objet que vous avez vu depuis que êtes un bébé. Et même la notion d’un objet, le fait que le monde soit tri-dimensionnel, la compréhension que si un objet est caché derrière un autre, il continue quand même à exister, vous l’apprenez. Ce ne sont pas des concepts innés. Ils sont appris. Nous appelons ce type d’apprentissage « apprentissage non dirigé ».

Beaucoup de ceux qui sont impliqués dans la renaissance du « deep learning » depuis le milieu des années 200, parmi lesquels [Geoff Hinton](http://www.cs.toronto.edu/~hinton/), [Yoshua Bengio](http://www.iro.umontreal.ca/~bengioy/yoshua_en/index.html), et moi-même, ceux que l’on appelle « la conspiration du deep learning », comme Andrew Ng, partent de l’idée qu’il est préférable d’utiliser l’apprentissage non dirigé plutôt que l’apprentissage dirigé. L’apprentissage non dirigé peut aider à « pré-éduquer » des réseaux comportant un grand nombre de couches. Nous avons obtenu quelques résultats avec cette méthode, mais malgré tout, ce qui fonctionne réellement est la méthode éprouvée de l’apprentissage dirigé, néanmoins combiné avec des réseaux à convolutions, sur lesquels nous travaillons depuis vingt ans.

Mais du point de vue de la recherche, ce qui nous a intéressé est de trouver comment utiliser efficacement l’apprentissage non dirigé. Nous possédons maintenant des techniques de ce genre, qui fonctionnes. Le souci est que ces techniques sont surpassées si l’on nourrit le système avec un grand nombre de données, puis que l’on utilise l’apprentissage dirigé. C’est pourquoi dans l’industrie, les applications de « deep learning » utilisent toutes l’apprentissage dirigé. Mais il n’en sera pas toujours ainsi

Le point important à retenir est que le cerveau fonctionne bien mieux, du point de vue de l’apprentissage non dirigé, que nos systèmes. Cela signifie qu’il y a encore des choses qui nous échappent dans les principes de base de l’apprentissage « biologique », et que nous ne savons pas encore les incorporer à nos systèmes d’apprentissage automatique.