

Compte rendu de la conférence de Sébastien Loustau Intelligence Artificielle et Réchauffement Climatique

au lycée Louis Barthou le 28/05/2021

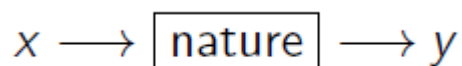
L'Intelligence Artificielle (IA)

On peut considérer qu'Alan Turing et Claude Shannon sont les pères fondateurs de l'intelligence artificielle. Alan Turing est l'un des premiers scientifiques à s'être intéressé à la question suivante dans les années 50: une machine peut-elle penser? Il a alors l'idée d'un test, appelé test de Turing. Dans cette expérience, un humain doit déterminer lors d'une conversation en aveugle, en mode texte, si il converse avec un homme ou une femme. Lors de la conversation, la femme dit la vérité et l'homme se fait passer pour une femme. L'idée de Turing est alors de remplacer l'homme par une machine, et de mesurer statistiquement sur un grand nombre d'expériences si la machine est capable de tromper l'humain mieux qu'un autre humain, réalisant ainsi une tâche complexe

Claude Shannon est le fondateur de la théorie de l'information. Il travaille sur la notion d'entropie (désordre présent dans un phénomène aléatoire), qui est toujours utilisée en IA aujourd'hui. Il est aussi à l'initiative de la conférence de Darmouth, considérée comme l'acte de naissance de l'IA.

Dans les années 70-80, les capacités de calcul des ordinateurs rendent possibles l'exécution des algorithmes de calcul que les chercheurs ont conçus dans les années 50. En effet, comme l'indique la loi de Moore (du nom du fondateur d'Intel Gordon Moore), la capacité d'un ordinateur double tous les deux ans. Cependant, dans les années 90, la limite physique des transistors est atteinte, et cette loi ne peut plus s'appliquer. On utilise alors pour les calculs en IA des processeurs graphiques, les GPU, optimisés pour les calculs matriciels (multiplication de matrices) et conçus pour les scènes en 3D.

Beaucoup de problèmes traités par l'IA peuvent se résumer au schéma suivant, où x représente l'entrée et y la sortie.



Prenons l'exemple de la voiture autonome. En entrée x il y a l'image de la route prise par une caméra fixée sur la voiture et en sortie y il y a l'action que doit faire le conducteur : tourner le volant à droite, à gauche, ne pas tourner le volant, accélérer, décélérer... La nature, dans cet exemple, c'est le conducteur. Avec l'IA, il n'y a plus de conducteur: à chaque image x d'entrée la machine doit prédire l'action y à faire, comme le ferai le conducteur. On remplace la nature par un algorithme que l'on entraîne en utilisant une base de donnée.

L'exemple d'Alexnet

Pour expliquer plus en détail, nous allons prendre l'exemple d'un réseau de neurone, appelé Alexnet, qui travaille sur des images pour reconnaître un animal ou un objet. Alexnet fait son apprentissage avec un million d'images qui viennent de la base de donnée IMAGENET (constituée de 20 millions

d'images). Chaque image de cette base est étiquetée, c'est à dire qu'on a associé une étiquette à chaque image: par exemple l'étiquette chat, l'étiquette chien, l'étiquette bateau... Il y a mille étiquettes différentes en tout. Depuis 2012 les chercheurs entraînent les réseaux de neurone à identifier ce qu'il y a dans les images, c'est à dire à associer à chaque image une étiquette. A cette époque le réseau de neurone Alexnet a fait de grand progrès grâce au deep learning et à l'utilisation de cartes graphiques (car les opérations réalisées sont des additions et des multiplications, opérations pour lesquelles les cartes graphiques sont optimisées).

Alexnet comporte 8 couches de neurones. Chaque couche est composée de deux fonctions. L_w est une fonction linéaire (qui correspond à une convolution) et σ est une fonction (appelée activation) non linéaire.

$$f_W(x) = \underbrace{\sigma \circ L_{W_8}}_{\text{couche 8}} \circ \underbrace{\sigma \circ L_{W_7}}_{\text{couche 7}} \circ \dots \circ \underbrace{\sigma \circ L_{W_1}(x)}_{\text{première couche}}$$

En sortie on a un vecteur de 1000 composantes et chaque composante est une probabilité d'appartenance à une étiquette (1000 étiquettes en tout).

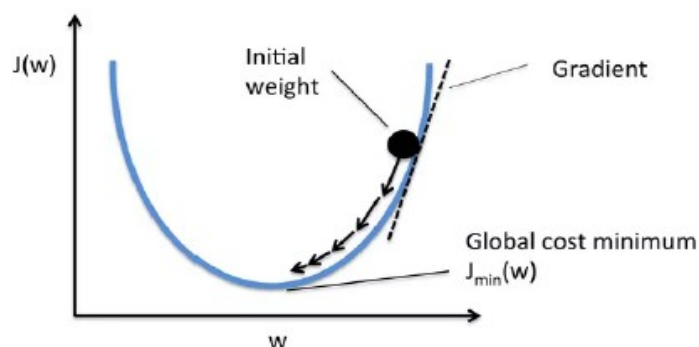
Les coefficients w sont des poids (weights). Toutes ses fonctions dépendent donc de poids, qui sont des paramètres à définir que l'algorithme fonctionne bien. Les paramètres vont être ajustés automatiquement dans la phase d'apprentissage avec la base de donnée. Dans le cas d'Alexnet il y a 60 millions de paramètres et chacun de ces paramètres est une valeur entre 0 et 255.

Comment fait-on pour trouver la bonne valeur pour chacun de ces paramètres? C'est un problème d'optimisation. Le machine learning est basé sur cette optimisation. On va chercher le meilleur ensemble de poids, de façon à minimiser cette somme:

$$\min_W \sum_{i=1}^n \ell(Y_i, f_W(X_i))$$

$\ell(Y_i, f_w(X_i))$ est l'erreur entre le résultat des calculs, $f_w(X_i)$, et la réponse attendue Y_i . L'optimisation revient à chercher l'ensemble de poids qui minimise cette somme des erreurs.

Pour rechercher un minimum, on utilise la dérivée de la fonction:



Sur cette représentation graphique, on voit que si l'on se trouve sur la partie croissante de la fonction, la dérivée sera positive. Il faudra donc se déplacer vers la gauche pour se rapprocher du minimum. De même si l'on se trouve sur la partie décroissante de la fonction, la dérivée est négative : il faudra donc se déplacer vers la droite pour se rapprocher du minimum. En procédant par itération, on peut donc trouver ainsi le minimum. C'est ce que l'on appelle la descente de gradient. Le poids initial w_0 est tiré au hasard. On calcul la dérivée et en fonction de son signe on se déplace soit vers la gauche soit vers la droite pour descendre et se rapprocher du minimum. On obtient ainsi une nouvelle valeur du poids, w_1 , qui est plus proche de l'abscisse du minimum que ne l'était w_0 . Ensuite on recommence l'opération pour avoir w_2 . Cette technique s'appelle la descente de gradient.

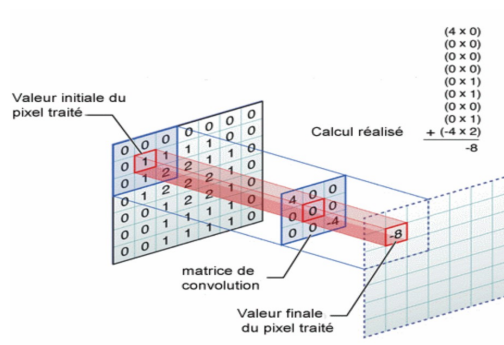
On utilise la formule suivante qui donne la valeur du poids suivant en fonction du poids précédent :

$$W^{(t+1)} = W^{(t)} - \alpha \nabla_W \left(\sum_{i=1}^n \ell(g \cdot (X_i), Y_i) \right) [W^{(t)}]$$

Pour calculer le poids W_{t+1} , on ajoute à W_t un terme proportionnel à l'opposé du gradient. Et on fait cela des millions de fois (le calcul d'une dérivée s'effectue en multipliant des matrices).

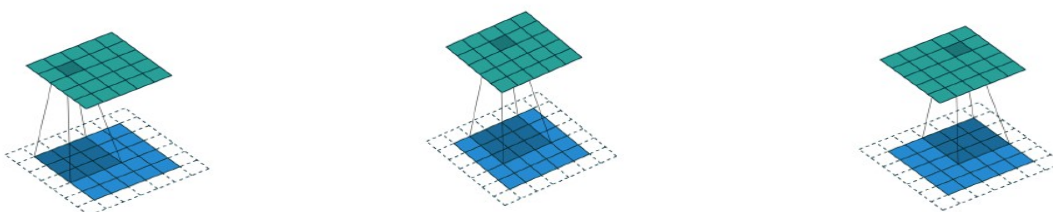
La convolution

Réaliser une convolution d'une partie de l'image (par exemple une fenêtre de 3×3 pixels au voisinage d'un pixel central) par un filtre (ou une matrice), revient à multiplier chaque coefficient du filtre par la valeur du pixel lui correspondant (même ligne et même colonne) comme l'indique le schéma suivant:



La convolution n'utilise que des multiplications et des additions. Si les valeurs finales obtenues sont élevées cela veut dire qu'il y a une corrélation entre le filtre et l'image. C'est un traitement local, autour d'un pixel.

Ensuite on déplace le filtre pour parcourir toute l'image d'origine pixel par pixel.



Cette opération va permettre d'extraire de l'image des informations intéressantes dans l'image, comme par exemple des contours.

Jusqu'au début des années 90 on déterminait les coefficients des filtres à l'avance, avec des mathématiques, en utilisant par exemple la théorie des ondelettes. Avec le deep learning, l'approche est différente. On choisit au hasard les coefficients du filtre au début. Puis, en utilisant la technique de la descente de gradient que l'on a vu précédemment, on va obtenir les bons coefficients. On dit que ces filtres sont appris, puisqu'au départ les coefficients sont pris au hasard puis, grâce à la descente de gradient, ils sont optimisés pour diminuer l'erreur. L'apprentissage est terminé quand l'erreur est la plus petite possible.

Il existe de nombreuses topologies différentes de réseaux. Pour Alexnet, il y a huit couches de neurones. On manipule des tenseurs (dimension trois) car chaque images en entrée est codée sous forme de trois matrices de 227×227 valeurs. Il y a trois matrices par images car les images sont en couleur (RVB) : une matrice pour la couleur Rouge, une autre pour le Vert et une autre pour le Bleu. En fin de traitement, on obtient un vecteur à mille composantes. Si le réseaux fonctionne bien, en fin de traitement on doit avoir une des mille composantes dont la valeur est proche de 1 et les 999 autres dont la valeur est proche de 0. Le réseau a donc associé à l'image une étiquette parmi les mille qu'il connaît (chat, chien , bateau, ...).

Pour améliorer les résultats, les chercheurs ont essayé un grand nombre de configurations possibles, en faisant varier le nombre de couches, les branchements entre les couches et la taille des filtres.

Quelques exemples d'application des réseaux de neurones

On peut réaliser de multiple traitement d'images avec les réseaux de neurones. Par exemple, on donne deux images au réseau de neurones: l'une représente l'image d'entrée et l'autre le style d'image que l'on souhaite. Après traitement, le réseau va proposer une image qui représente la même chose que l'image d'entrée mais avec le nouveau style. Par exemple si on donne au réseau l'image d'une maison n°1 en entrée, avec un style défini par l'image d'une maison n°2 au crépuscule, le réseau va fournir en sortie l'image de la maison n°1 au crépuscule.

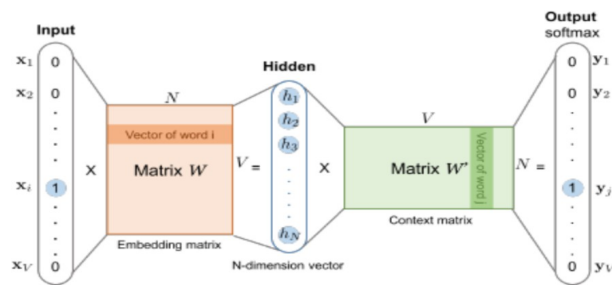


Le traitement par réseaux de neurones peut s'appliquer à tout type de données. On peut faire travailler un réseau de neurones sur du texte. L'idée est la suivante: pour trouver le sens d'un mot, on considère les mots qui se trouvent autours de lui.

Dans l'exemple suivant, on entre une citation en entrée du réseau, et le réseau va fournir en sortie une autre citation qui a le même sens mais dans un style différent.

La nuit porte conseil.
↓
La rosée tombe sur l'herbe, quand la nuit est la plus
silencieuse.
Nietzsche, Ainsi parlait Zarathoustra - Deuxième Partie

Ce réseau de neurones est capable d'associer à une citation en entrée « La nuit porte conseil » une citation en sortie dont le sens est similaire, mais dans un style « philosophique ». Cependant, il ne crée pas la citation de sortie: c'est l'une des citations, parmi un corpus plus ou moins important, qu'on a fourni au réseau durant la phase d'apprentissage.



On transforme chaque mot en un vecteur, avec une couche cachée (Hidden sur le schéma). Chaque mot peut être représenté par un vecteur de taille 50 par exemple.

Par exemple si l'algorithme constate que les mots « nuit », « lune » et « chouette », se trouvent souvent ensemble dans une phrase, il va en déduire qu'il y a un lien entre ces trois mots.

On peut citer aussi parmi les nombreuses applications des réseaux de neurones, les logiciels de traduction automatique qui vont associer à un mot en entrée dans une langue un mot en sortie dans une autre langue.

Mais tous ces algorithmes de plus en plus performants font appel à des capacités de calcul qui augmentent et donc qui consomment plus d'énergie. Et la consommation d'énergie est à l'origine du changement climatique auquel nous sommes confrontés...

Le réchauffement climatique.

Nous sommes habitués à consommer tous les jours des quantités d'énergie, puisées dans la nature, bien plus importante que notre seul corps humain serait capable de délivrer. Pour nous en faire prendre conscience, Sébastien Loustau nous propose une expérience de physique toute simple. Il fait appel à un élève volontaire qui va soulever puis reposer une masse de 1kg à 1m du sol pendant cinq minutes (en fait l'expérience n'a duré que quatre minutes pour réduire l'effort de l'élève). Un autre élève volontaire tient un mètre à côté pour mesurer l'amplitude du mouvement de son camarade.

En cinq minutes, on calcule que l'élève va fournir 3 000 Joules (10 joules par seconde pendant 5 minutes).

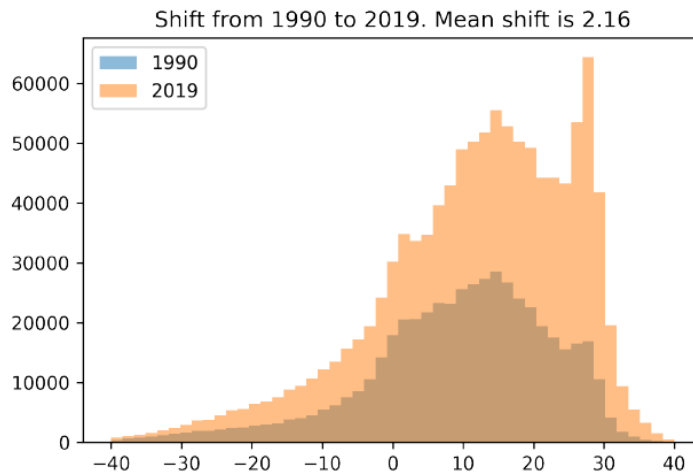
Un litre d'essence représente 10 kW/h, donc 36 000 000 Joules.

Sébastien montre 2,5 cl d'essence, c'est à dire 40 fois moins qu'un litre, dans une petite fiole de laboratoire. L'essence contenue dans cette fiole est capable de fournir 300 fois plus d'énergie que ce que l'élève va dépenser en cinq minutes.

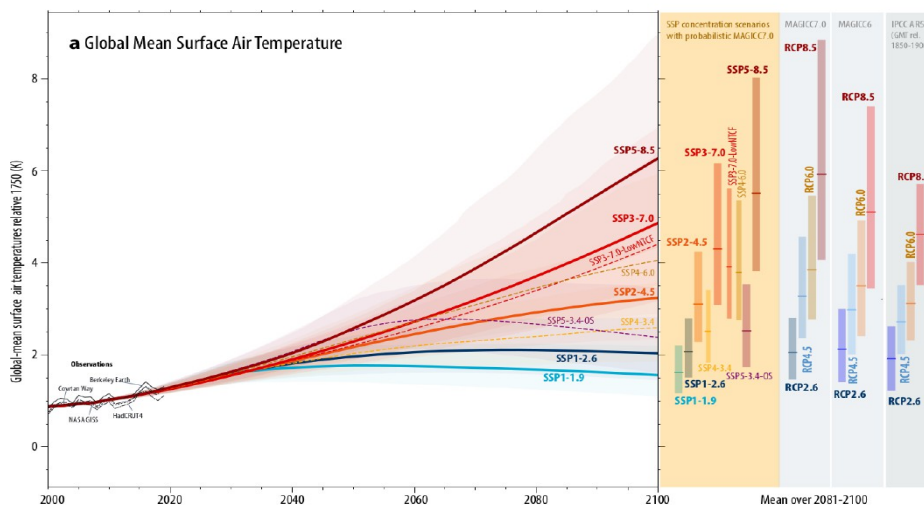
Chez l'homme, la puissance que développent les bras est d'environ 10W et les jambes 100W. Ce n'est rien comparé à un tracteur (60kW, soit 600 paires de jambes), un camion (400 kW, soit 4000 paires de jambes) ou un avion (100 MW, soit 1 million de paires de jambes).

Aujourd'hui un Français consomme en moyenne à peu près 4,3 litres de pétrole par jour. Un calcul rapide montre que c'est équivalent au travail musculaire que devraient fournir entre 30 et 60 randonneurs de 80kg avec 10 kg sur le dos pour faire 1 000 de dénivelé pendant un jour.

Plus on consomme de l'énergie fossile (charbon, gaz pétrole) plus cela rejette de gaz à effet de serre. On est passé d'une concentration de 280ppm de CO2 dans l'atmosphère en 1880, à l'aube de la révolution industrielle, à une concentration de 410 ppm en 2018. Cette concentration plus importante augmente l'effet de serre et donc la température moyenne de l'atmosphère. On ne réalise par forcément que la répartition des degrés supplémentaires n'est pas répartie de façon uniforme sur la plage des températures comme le montre la distribution des températures suivante:



Entre 1990 et 2019, il y a eu une augmentation moyenne des températures d'environ 2,16°. Mais on voit qu'il n'y a pas eu un simple décalage de l'histogramme foncé vers la droite de 2 degrés. Par exemple on voit qu'il y a beaucoup plus d'endroits où la température est proche des 30° en 2019 qu'en 1990. On voit aussi qu'il y a plus de températures extrêmes. Des endroits du globe deviennent invivables pour les populations (ce qui entraîne des déplacements de population). On estime à environ 800 millions le nombre de personnes qui pourraient se déplacer d'ici 2100. Jusqu'en 2030-2040, quoi que l'on fasse, le réchauffement moyen devrait continuer à augmenter par ce que le CO2, qui est très stable, disparaît lentement dans l'atmosphère. Ensuite plusieurs scénarios sont possibles en fonction des émissions nouvelles de CO2, du plus optimiste (SSP1) au moins optimiste (SSP5):



L'IA appliquée aux problèmes environnementaux

Sébastien Loustau fait parti d'une équipe de cinq chercheurs et ingénieurs sur Pau qui travaillent autour de la thématique du green IA, financée par l'UPPA et le label E2S.

L'équipe travaille sur deux axes de recherche :

- La diminution de la consommation de l'énergie utilisée par les algorithmes pour faire les descentes de gradient. L'objectif est de développer des méthodes utilisant les mathématiques pour consommer moins d'énergie pour les calculs. Des nouvelles métriques sont utilisées dans les algorithmes pour réduire la consommation d'énergie sans trop perdre en performances.
- Appliquer toutes les méthodes d'IA aux secteurs énergie ou environnement, pour contribuer à lutter contre le réchauffement climatique. Le travail sur cette thématique se fait en partenariat avec des acteurs locaux (entreprises, agriculteurs, collectivités).

Un projet en partenariat avec l'agglomération de Pau et les coopérative d'agricoles locales, vise à évaluer la capacité des exploitations agricoles à favoriser la photosynthèse dans les champs. L'objectif est de développer des outils à base d'IA qui vont permettre de mesurer la séquestrations de carbone dans les exploitations agricoles. C'est un enjeu actuel car les collectivités ou les entreprises souhaitent atteindre la neutralité en carbone, et si cette neutralité n'est pas possible en interne, les émissions de CO2 doivent être compensées. Une des stratégies utilisées est l'achat de puits de carbone. Certaines pratiques agricoles peuvent jouer ce rôle de séquestration du carbone: l'épandage avec du compost, le semi direct pour ne pas labourer. Les réseaux de neurones vont permettre de corréler les images aériennes avec les données de mesure de carbone, afin d'évaluer l'efficacité des pratiques agricoles en terme de séquestration de carbone.

Pour terminer, Sébastien Loustau cite un exemple de partenariat avec une entreprise Palloise (Altanové) qui conçoit des murs d'écrans pour monitorer des crises. Les préfectures et certains industriels possèdent des salles de crises pour réagir très rapidement lorsqu'une situation de crise survient (par exemple un incendie, une catastrophe naturelle ou industrielle). Les outils développés à base d'IA vont permettre de détecter des informations pertinentes dans tout le flux d'informations (réseaux sociaux, news, caméras de vidéo surveillance, capteurs) et de les afficher au bon moment en salle de crise pour aider à gérer ces situations critiques.