

# Intelligence Artificielle et réseaux neuronaux

Franck Jeannot

Montréal, Canada, 2022, Z766, v1.2, Maj 17 Avril 2023

---

---

## 1. Introduction

En août 2022, le journaliste scientifique Matthieu Dugal (Radio Canada) à animé un documentaire en 3 épisodes<sup>1</sup> intitulé « **IA : être ou ne pas être** » [1] qui fait une introduction au sujet de l'intelligence artificielle. Dans ce documentaire, le journaliste se lance dans la création de son double numérique en puisant dans ses données personnelles et en utilisant les plus récentes technologies de l'IA, et par ce biais introduit plusieurs notions par des discussions avec quelques spécialistes du sujet.

Le sujet de l'IA est aussi traité par Matthieu Dugal dans un balado [2] « L'intelligence artificielle et nous » .

## 2. Sciences des données et réseaux neuronaux

Dans le documentaire présenté précédemment, une des intervenantes, Aurélie Jean<sup>2</sup>, met en contexte certaines notions d'intelligence artificielle. Dans l'ouvrage d'Aurélie Jean « *De l'autre côté de la machine. Voyage d'une scientifique au pays des algorithmes* » [3], elle introduit<sup>3</sup> **Stéphane Mallat** (professeur au Collège de France et la présentation inaugurale de celui-ci sur les « *Sciences des données* » [4] et celle-ci rappelle « *Ces réseaux neuronaux peuvent devenir très volumineux, et si complexes qu'il est difficile de les analyser afin de caractériser a posteriori les critères implicites de l'algorithme entraîné. Le mathématicien Stéphane Mallat [4] l'explique très bien dans sa conférence inaugurale<sup>4</sup> au Collège de France.* »

---

1. La série documentaire IA, être ou ne pas être est présentée à partir du samedi 23 juillet 2022 à 22 h 30 sur ICI Télé

2. [https://fr.wikipedia.org/wiki/Aur%C3%A9lie\\_Jean](https://fr.wikipedia.org/wiki/Aur%C3%A9lie_Jean)

3. p. 113/171

4. leçon inaugurale prononcée le 11 janvier 2018

### 3. Leçon inaugurale - Sciences des données, par Stéphane Mallard - Janv. 2018

En se référant à la synthèse de la « *leçon inaugurale* » rappelée dans la lettre 118 (2020) de l'annuaire du Collège de France [5] on retrouve que Stéphane Mallard « introduit les principes des algorithmes d'apprentissage et certains grands problèmes posés par cette discipline. Une question centrale est de comprendre les conditions dans lesquelles il est possible d'apprendre et donc de généraliser à partir d'exemples, suivant la complexité des données et de la fonction qu'il faut prédire. Cela met en jeu l'existence de régularités connues a priori, utilisées par des algorithmes comme l'apprentissage à noyaux, ou qui peuvent être découvertes par des algorithmes comme les réseaux de neurones. Ces notions de régularité font appel à diverses notions mathématiques comme la parcimonie, la séparation d'échelles par ondelettes<sup>5</sup>, ou des notions d'invariance relativement à l'action de groupes de symétries. Au cours des dernières années, les réseaux de neurones profonds ont obtenu des résultats spectaculaires dans de nombreux domaines d'applications, que les mathématiques sont loin de pouvoir expliquer. Ces succès posent aussi des questions sociétales et éthiques importantes, qui sont évoquées en lien avec les problèmes scientifiques sous-jacents. »

### 4. L'apprentissage face à la malédiction de la grande dimension

En présentant certains fondamentaux mathématiques à l'intelligence artificielle et au machine learning, dans sa leçon inaugurale [6], Stéphane Mallard introduit les **Réseaux Neuronaux**<sup>6</sup>.

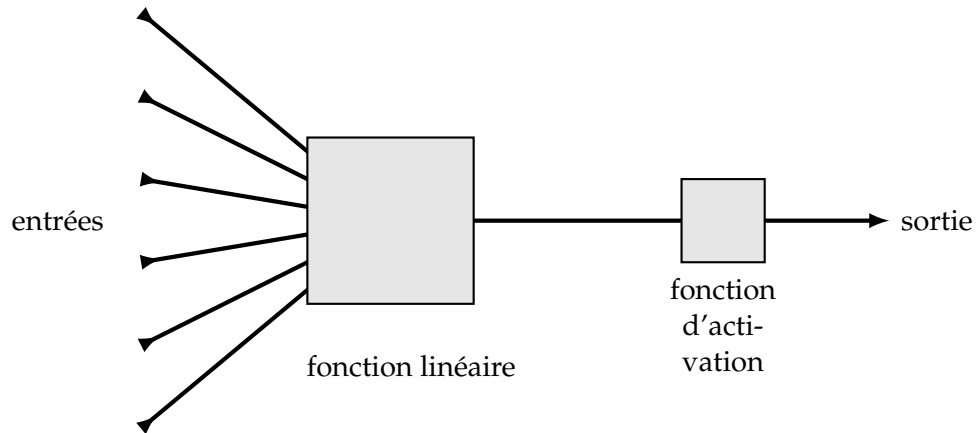
---

5. décomposition similaire à la transformée de Fourier à court terme, utilisée dans le traitement du signal

6. repris plus tard sous forme de notes par J.E Campagne [7]

## 5. Réseaux de neurones et perceptron

Le principe du **perceptron linéaire** est de prendre des valeurs en entrées, de faire un calcul simple et de renvoyer une valeur en sortie. Les calculs dépendent de paramètres propres à chaque perceptron <sup>7</sup>.



Les **réseaux de neurones** sont notés souvent **NN** pour "Neural Network", ou "MLP : Multi Layer Perceptron" pour « réseau **perceptron multicouches** <sup>8</sup> ».

*Un perceptron est l'unité de base du réseau neuronal, et en suivant l'analogie biologique, il joue le rôle d'un neurone qui reçoit de l'énergie à travers ses connexions synaptiques et la transmet à son tour selon une fonction d'activation [9].*

Concrètement, chaque perceptron calcule sa valeur de sortie comme :

$$v_{i,j} = f\left(\sum_{k=0}^n w_{i,j,k} \cdot x_k\right)$$

- $v_{i,j}$  représente la valeur de la variable de sortie ou de la fonction d'activation pour le neurone situé à la position  $i, j$  dans un réseau neuronal.
- $f(\cdot)$  est la fonction d'activation appliquée à la somme pondérée à l'intérieur des parenthèses.
- $\sum_{k=0}^n$  est le symbole de la sommation, indiquant que l'on effectue une somme sur la variable d'index  $k$  de 0 à  $n$ .

---

7. p. 65 de [8]

8. [https://fr.wikipedia.org/wiki/Perceptron\\_multicouche](https://fr.wikipedia.org/wiki/Perceptron_multicouche)

- $w_{i,j,k}$  est le poids associé à la connexion entre le neurone d'entrée  $x_k$  et le neurone de sortie  $v_{i,j}$ . Les poids sont appliqués de manière multiplicatrice aux valeurs d'entrée pour obtenir une somme pondérée.
- $x_k$  représente la valeur d'entrée provenant du neurone d'entrée situé à la position  $k$ .
- $n$  est le nombre total de neurones d'entrée dans le réseau neuronal.

où  $\{x_1 \dots x_n\}$  sont les entrées du neurone,  $x_0 = -1$  est une entrée de biais fixe et  $f$  est une fonction d'activation non linéaire<sup>9</sup>. L'entraînement du réseau consiste à calculer les poids synaptiques  $\{w_{i,j,0} \dots w_{i,j,n}\}$  pour chaque neurone  $(i, j)$  ( $i$  est la couche,  $j$  est la position du neurone dans la couche).

L'entraînement se réalise par rétropropagation<sup>10</sup> : pour la couche de sortie, à partir d'une sortie attendue  $\{y_1 \dots y_m\}$  on calcule (pour chaque neurone) une estimation de l'erreur :

$$\delta_{i,j} = f'(v_{i,j}) (y_j - v_{i,j})$$

Celle-ci est propagée aux couches précédentes par :

$$\delta_{i,j} = f'(v_{i,j}) \sum_k w_{i,j,k} \cdot \delta_{i+1,j}$$

où la somme est calculée sur tous les neurones connectés avec le neurone  $(i, j)$ .

Les nouveaux poids, au temps  $t + 1$ , sont :

$$w_{t+1;i,j,k} = w_{t;i,j,k} + \Delta w_{t;i,j,k}$$

où  $\Delta w_t$  dépend d'un facteur de correction et aussi de la valeur de  $\Delta w_{t-1}$  multipliée par un momentum  $\mu$  (cela donne aux modifications une sorte d'énergie cinétique) :

$$\Delta w_{t;i,j,k} = (\lambda \cdot \delta_{i+1,k} \cdot v_{i,j}) + \mu \cdot \Delta w_{t-1;i,j,k}$$

Le facteur de correction dépend des valeurs  $\delta$  calculées et aussi d'un facteur d'apprentissage  $\lambda$  qui peut être ajusté pour accélérer la convergence du réseau.

## 6. Exemple de réseaux de neurones simplifié

Dans l'exemple<sup>11</sup> plus bas, on commence par les observations  $x_1$ ,  $x_2$  et  $x_3$ . Ensuite, on a les résultats  $y_1$  and  $y_2$ . La couche cachée avec les neurones  $z_1$ ,  $z_2$ ,  $z_3$  et

9. utilisation de *tanh*, la tangente hyperbolique

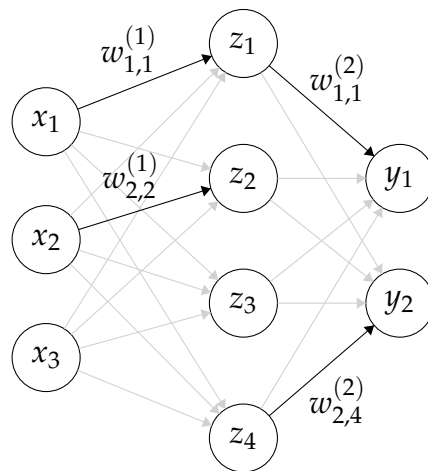
10. <https://fr.wiktionary.org/wiki/r%C3%A9tropropagation>

11. inspiré de "Une introduction à l'apprentissage automatique" Ralph Sarkis, 2018

$z_4$ . Finalement, les liens  $w$  entre les neurones.

Comment calcule-t-on les activations dans les autres couches ? Chaque flèche dans le graphique représente un coefficient que l'on dénotera par  $w$  :

$$w_{i,j}^{(\ell)} \begin{cases} \ell & \text{Numéro de la couche} \\ i & \text{Indice du neurone de la couche actuelle} \\ j & \text{Indice du neurone de la couche précédente} \end{cases}$$



*Entraînement du réseau.* Au début, les coefficients sont arbitraires. Ensuite, le réseau classe des observations déjà cataloguées et s'ajuste pour donner des meilleures réponses. Quand l'erreur  $J$  est petite, on se rapproche de la solution. On cherche donc à minimiser l'erreur en utilisant une technique d'optimisation simple.

Tel que l'énonce **Jean-Eric Campagne** (2018, École normale supérieure de Paris) « Les réseaux de neurones <sup>12</sup> (...) apprennent en même temps la représentation et le classificateur. La représentation est une **cascade de classificateurs** <sup>13</sup> **linéaires** que l'on passe dans des **filtres non-linéaires.** »

12. <https://avant-gare.on-train.com/2022/01/tout-savoir-sur-les-reseaux-de-neurones/>

13. p 108/112 de [7]

## Références

- [1] Dugal, Matthieu et al., *IA : être ou ne pas être* (08 2022).  
URL <https://ici.radio-canada.ca/tele/ia-etre-ou-ne-pas-etre/site>
- [2] De lancer, Alexis and Dugal, Matthieu and Côté-Légaré, Guillaume, *L'intelligence artificielle et nous (Ca s'explique)*, Alexis De Lancer (animation), Matthieu Dugal (invité) et Guillaume Côté-Légaré (réalisation).  
URL <https://ici.radio-canada.ca/ohdio/balados/6108/ca-sexplique-balado-info-alexis-de-lancer/647986/intelligence-artificielle-histoire-avancement-quotidien>
- [3] Jean, Aurélie, *De l'autre côté de la machine : voyage d'une scientifique au pays des algorithmes*, Éditions de l'Observatoire, 2019.
- [4] Mallat, Stéphane, *Sciences des données et apprentissage en grande dimension, La lettre du Collège de France [En ligne]*, 2017-2018; leçon inaugurale prononcée le 11 janvier 2018; mis en ligne le 21 mai 2019 (01 2018).  
URL <https://www.college-de-france.fr/site/stephane-mallat/inaugural-lecture-2018-01-11-18h00.htm>
- [5] Mallat, Stéphane, *Sciences des données*.  
URL <https://journals.openedition.org/annuaire-cdf/15441>
- [6] Mallat, Stéphane, *Sciences des données, L'annuaire du Collège de France. Cours et travaux (118) (2020) 31–42*, Paris; coédition Collège de France/Fayard coll. Leçons inaugurales; 2018; 80 p.
- [7] Campagne, Jean Eric, *Notes et commentaires au sujet des conférences de S. Mallat du Collège de France (2018): L'apprentissage face à la malédiction de la grande dimension*.  
URL <https://www.di.ens.fr/~mallat/College/Cours-2018-Mallat-Jean-Eric-Campagne.pdf>
- [8] Bodin, Arnaud and Recher, François, *Deepmath - Mathématiques (simples) des réseaux de neurones (pas trop compliqués): Algorithmes et mathématiques*.  
URL <https://github.com/exo7math/deepmath-exo7>
- [9] Burrioni, Javier and Sarraute, Carlos, *Outrepasser les limites des techniques classiques de prise d'empreintes grâce aux reseaux de neurones*, arXiv preprint arXiv :1006.2844 (2010).