



# L'apprentissage non-supervisé et ses contradictions

Jérémie Sublime<sup>1, 2</sup>

L'apprentissage machine, également appelé apprentissage artificiel [2], ou encore *machine learning* en anglais, est un domaine scientifique lié à l'intelligence artificielle et qui se trouve à la frontière entre l'informatique et les statistiques. L'objectif de ce domaine est la conception et l'étude d'algorithmes capables « d'apprendre » à partir de données. Ces algorithmes d'apprentissage sont ensuite utilisés pour des tâches diverses et variées : analyse et traitement d'images (i.e. reconnaissance de visages), détection d'anomalies (i.e. connexions frauduleuses), systèmes de recommandation (Google, Youtube, Amazon, Netflix, etc.), tri et catégorisation d'information, et même pour apprendre à des programmes à jouer aux échecs, au Go<sup>3</sup>, ou aux jeux vidéos<sup>4</sup>.

On divise généralement l'apprentissage machine en trois grandes catégories : l'apprentissage supervisé, non-supervisé et par renforcement.

Dans l'apprentissage supervisé, les algorithmes sont entraînés à reconnaître et catégoriser des données à partir d'exemples étiquetés. On apprendra, par exemple, à un réseau de neurones à faire la différence entre des photos de chats ou de chiens en lui montrant des centaines d'images et en lui précisant à chaque fois de quel animal il s'agit de sorte qu'il finira par apprendre ce qui caractérise un chat ou un chien, et saura reconnaître ces caractéristiques sur une photo qu'il n'aura jamais vue.

1. ISEP, École d'ingénieurs du numérique.

2. LIPN - CNRS UMR 7030, LaMSN - Université Sorbonne Paris Nord.

3. <https://www.bbc.com/news/technology-35785875>.

4. DeepMind/AlphaStar.

Dans l'apprentissage non-supervisé, conçu à la base comme une tâche exploratoire, les algorithmes sont amenés à traiter des données non-étiquetées pour y trouver des structures, des groupes d'objets similaires, ou des anomalies. Si on reprend l'exemple précédent avec les photos de chats et de chiens, on fournira cette fois-ci toutes les photos à l'algorithme sans lui préciser de quel animal il s'agit. Avec un peu de chance, la méthode non-supervisée s'apercevra toute seule qu'il y a deux animaux différents selon les photos, et fera un tri pour séparer les chiens et les chats. Mais il se peut aussi qu'elle décide de les trier par couleur (en mélangeant chiens et chats), qu'elle fasse de multiples catégories par race, ou encore qu'elle fasse un tri selon des critères difficilement compréhensibles pour un humain (la couleur du papier peint par exemple) en ne se préoccupant nullement des animaux présents sur les images. On parle de *clustering* lorsque l'apprentissage non-supervisé regroupe ensemble des données considérées comme similaires selon des critères plus ou moins transparents. Les groupes ainsi formés sont appelés *clusters*.

L'apprentissage par renforcement est un domaine un peu différent des précédents puisqu'on cherche ici à influencer le comportement d'un algorithme d'intelligence artificielle à partir de récompenses si l'algorithme se comporte comme on le souhaite, et de pénalités dans le cas contraire. Ce type d'apprentissage est davantage utilisé pour apprendre aux algorithmes à jouer à des jeux (en choisissant les bons coups à jouer) plutôt que pour traiter des données.

L'apprentissage supervisé est sans doute la catégorie la plus connue des chercheurs tous domaines confondus, avec ses réseaux de neurones aux performances extraordinaires en analyse d'image. L'apprentissage par renforcement, quant à lui, est le plus médiatisé auprès du grand public pour ses succès lorsque les machines battent les meilleurs humains sur des jeux ou des tâches que l'on estimait il y a quelques années encore trop complexes pour qu'un algorithme puisse vaincre un jour un humain. Quant à moi, je travaille depuis le début de ma carrière de chercheur sur l'apprentissage non-supervisé, la branche moins connue de l'apprentissage, la plus mal posée aussi, et qui sera le sujet de cet article. J'aborderai un certain nombre de contradictions et de questions que l'on peut voir émerger dans ce domaine un peu particulier.

## Un apprentissage non-supervisé détourné de son but exploratoire initial ?

Lorsqu'il est présenté à des étudiants dans l'enseignement supérieur, l'apprentissage non-supervisé est généralement abordé sous l'angle du *clustering* : cette tâche exploratoire que j'ai évoquée plus haut et qui consiste à regrouper ensemble des données similaires dans des *clusters*. Ce cours sur le *clustering* fait généralement partie d'un module plus large d'apprentissage machine et est l'occasion de présenter les principales familles d'algorithmes de *clustering* : on présentera les K-moyennes [11]