



Intelligence Artificielle : dissection d'un 'buzzword'

Alban Jago - a.jago@cra.wallonie.be, en collaboration avec les unités U4, U6 et U12

Intelligence Artificielle : Quésako ?

De nos jours, l'Intelligence Artificielle (IA) est particulièrement à la mode et le mot est utilisé à tort et à travers. Selon les contextes, il cache des techniques très diverses, parfois banales et parfois très pointues. Difficile de ne pas être perdu et de savoir s'il est utilisé à bon escient ou dans un simple but de marketing. La question est immensément vaste et il n'existe pas de "Comité de l'IA" qui mettrait tout le monde d'accord mais voici quelques éléments de réflexion.

Au sens large, le concept désigne le fait qu'une technique ou une tâche :

- qui nécessite normalement une réflexion humaine (=> *intelligence*)
- puisse être effectuée par une machine (=> *artificielle*)

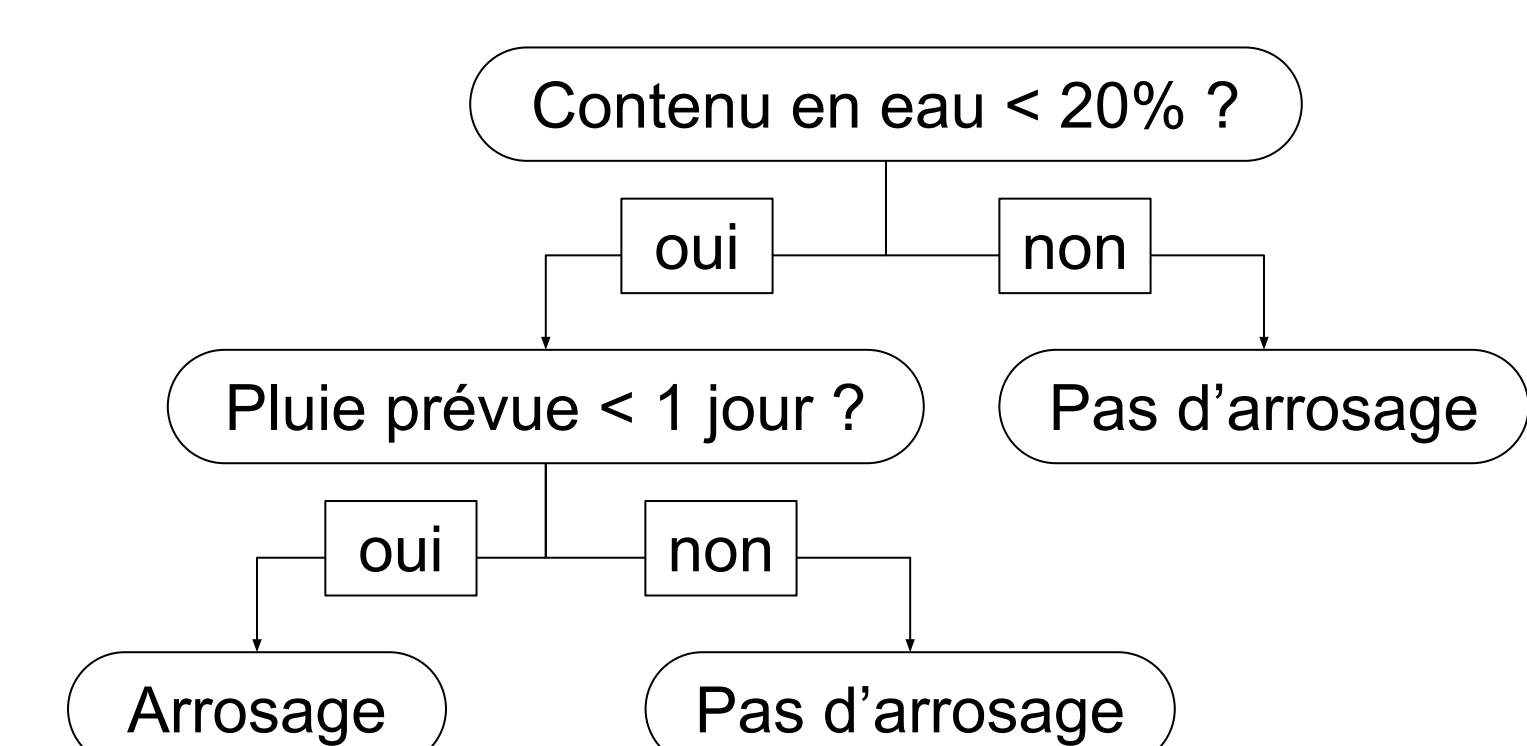
Aujourd'hui, dire qu'on "utilise des techniques d'IA" signifie souvent travailler avec une grande quantité de données dans sa totalité ("big data"). Mais cela ne dit pas grand chose sur les techniques utilisées pour leur traitement et leur analyse, qui peuvent aller de la simple régression linéaire à des réseaux de neurones très complexes.

Pourquoi tant de succès ?

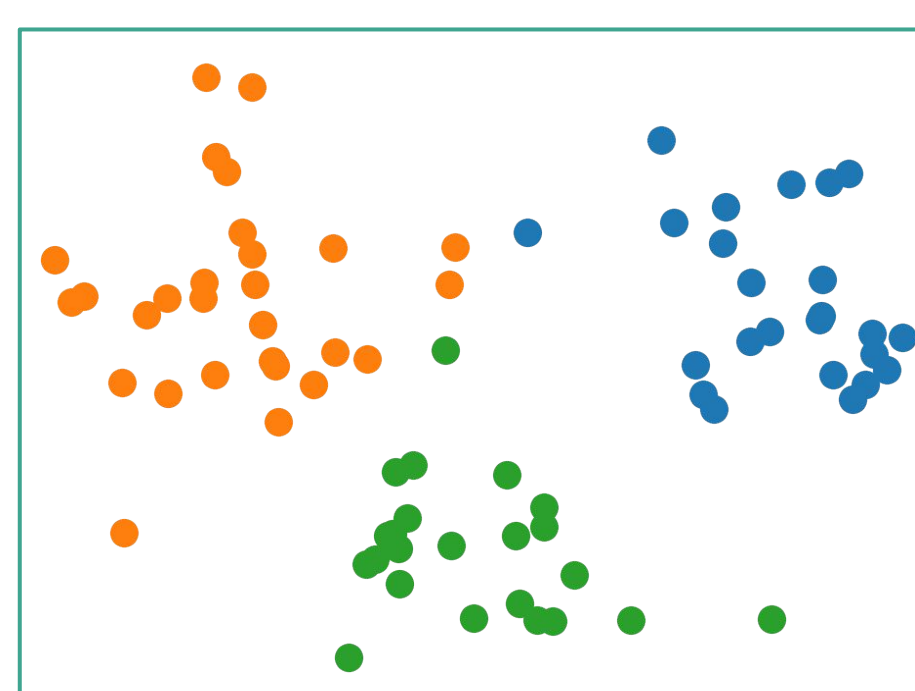
Le concept n'est pas nouveau, un premier boom ayant eu lieu depuis les années 50, suivis de hauts et de bas. Récemment, son explosion est due à une conjonction de plusieurs éléments :

- Augmentation drastique des données disponibles : capteurs embarqués, internet, photos, ...
- Puissance de calcul : datacenters, calcul parallèle, cloud computing, ...
- Nouveaux algorithmes (mais les anciens sont encore beaucoup utilisés)
- Facilité d'accès et d'utilisation : bibliothèques open-source, tutoriels, ...

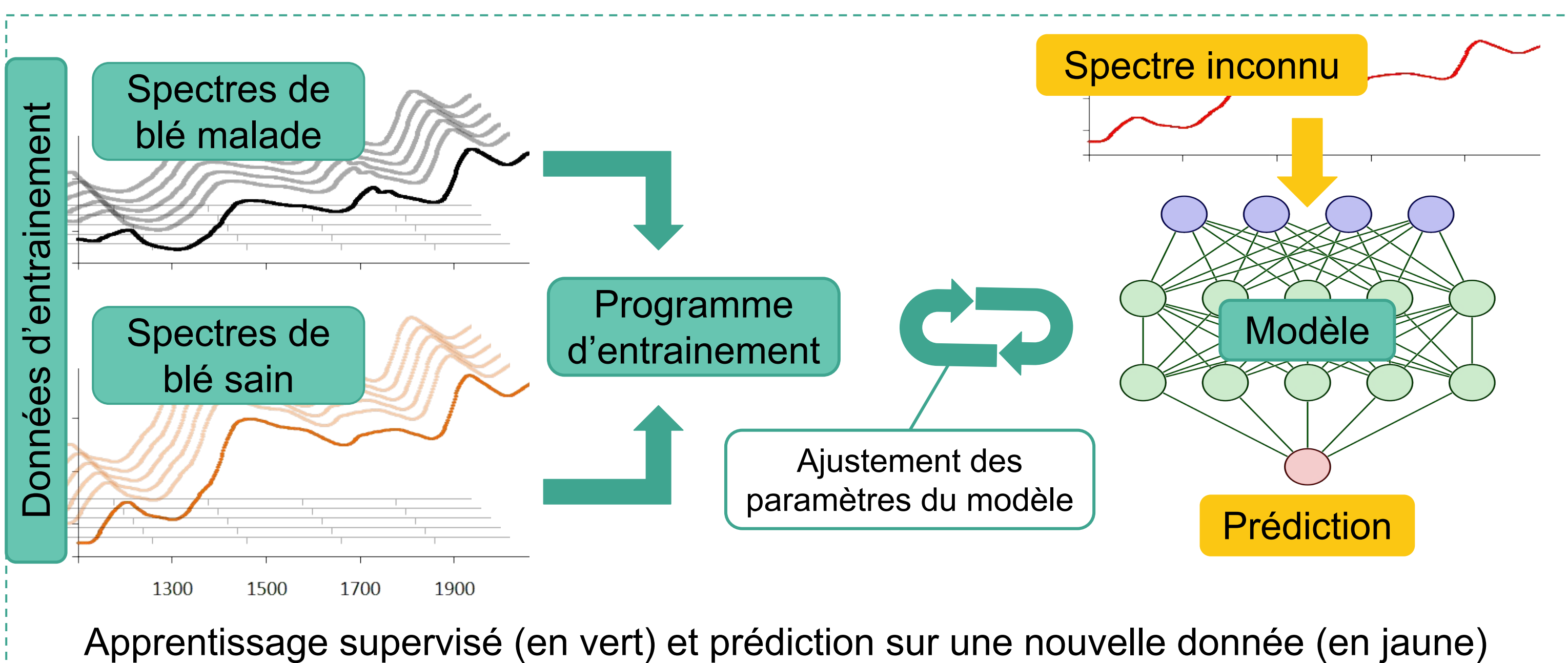
Ceci a permis des avancées majeures dans des domaines comme la reconnaissance d'image, la conduite autonome, le marketing, ...



Système de décision basé sur des règles explicites



Exemple de clustering non-supervisé



Apprentissage supervisé (en vert) et prédiction sur une nouvelle donnée (en jaune)

Et le Deep Learning ?

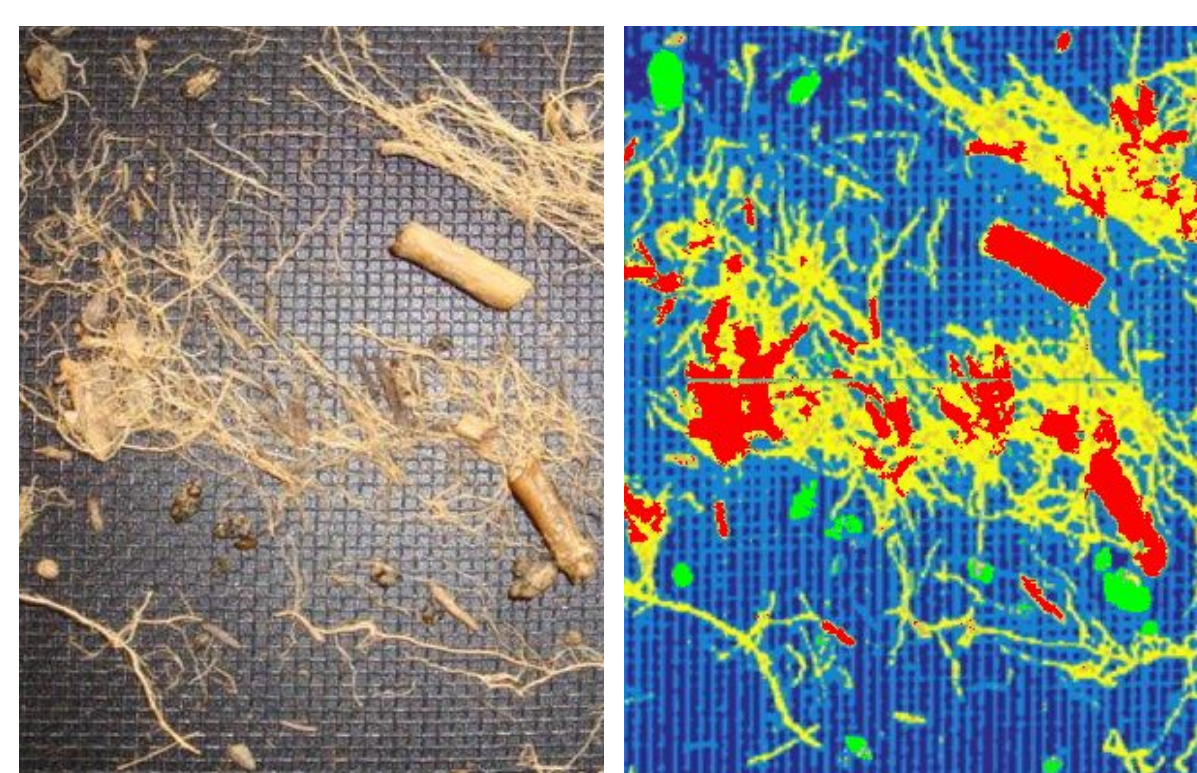
Les réseaux de neurones artificiels sont des algorithmes inspirés du fonctionnement du cerveau humain. Ils sont constitués de "neurones" connectés entre eux et organisés en couches successives. L'apprentissage profond (*deep learning*) est un sous-domaine du machine learning qui concerne les réseaux de neurones *profonds*, contenant un très grand nombre de couches. Il nécessite un très grand nombre de données et beaucoup de puissance de calcul.

Des exemples au CRA-W

Plusieurs unités du CRA-W travaillent avec des grands ensembles de données. Les techniques de machine learning, pas nécessairement les plus complexes, permettent de transformer ces données en applications agronomiques concrètes.

A l'U12, en spectroscopie, des SVM permettent de classer les spectres afin, par exemple, de détecter des éléments interdits dans des farines animales, ou de quantifier les résidus de cultures dans des échantillons.

Dans le projet InnovEau de l'U4, l'IA permet de repérer automatiquement les adventices.



Des approches différentes

Approche "classique" (*système expert*, ...)

Le comportement de la machine est dicté par des règles explicites prédéfinies par des humains. Exemples :

- aux échecs, tester un très grand nombre de positions et retenir la meilleure selon des règles prédéfinies par des experts
- système de décision basé sur des conditions if/else explicites

Apprentissage automatique (machine learning ; *conduit par les données*)

Les règles de la machine ne sont pas prédéfinies à l'avance par un humain : un programme d'apprentissage déduit ces règles sur base de données qui lui servent d'exemples

- Apprentissage supervisé : un programme générique est écrit avec des paramètres qui ne sont pas fixés. Un programme d'apprentissage déduit automatiquement, sur base de nombreux exemples de spectres de blé sain et de blé malade (p. ex.), la valeur des paramètres qui permet au programme de distinguer les deux avec le plus de précision.
- Apprentissage non-supervisé : par exemple, regroupement (*clustering*) de données en groupes semblables inconnus à l'avance
- Apprentissage par renforcement : un logiciel de jeu ajuste son comportement automatiquement en jouant à répétition contre lui-même

Algorithmes :

- | | |
|--|-------------------------------------|
| régression linéaire | machine à vecteurs de support (SVM) |
| forêt aléatoire (<i>random forest</i>) | méthode des k plus proches voisins |
| réseau de neurone artificiel | ... |

Quelques réflexions

- **Simplicité** : la tentation est grande d'utiliser des modèles complexes pour faire plus sérieux. Pourtant, des modèles simples suffisent souvent, avec l'avantage d'être plus faciles à expliquer et robustes.
- **Transparence** : bien qu'on sache *comment* ils fonctionnent, les modèles complexes donnent rarement le *pourquoi* d'une décision.
- **Pertinence** : si les techniques d'IA sont devenues très accessibles, la qualité des modèles dépend crucialement de la qualité des données et de l'entraînement.

