

Apprentissage Automatique

- Comprendre le positionnement du Deep Learning dans le champ plus large de l'apprentissage automatique
- Savoir quand le Deep Learning n'est pas approprié

- Pas de définition précise (champ très vaste)
- Idée transversale : éviter la programmation **explicite**.
- Création de programmes qui utilisent des données ou des algorithmes généraux pour apprendre à réaliser leurs tâches

- Désigne l'ensemble des méthodes qui utilisent des réseaux de neurones avec de nombreuses couches
- Nombreuses couches = hiérarchisation = grande expressivité

- Symbolique (grammaire, automate, ...)
- Probabiliste (modèle bayésien)
- Statistique (SVM, arbre de décision, réseaux de neurones, ...)

- Première branche explorée historiquement
- Analyse stricte de symboles
- Tolère souvent mal les erreurs
- Difficultés à généraliser à l'inconnu

Exemple : apprentissage de grammaires.

- Branche popularisée par Judea Pearl
- Aussi appelée approche probabiliste
- Tolère mieux les erreurs dans les données
- Interprétable

Exemple : modélisation d'un score de crédit par un graphe probabiliste.

- Ne cherche pas à modéliser exactement comme l'apprentissage symbolique
- Dégage des tendances statistiques
- Tolère bien les données bruitées
- Peut être interprétable

Exemple : Classification de documents avec des SVM.

Par rapport à l'apprentissage automatique statistique « classique » :

- Nécessite plus de données
- Plus dur à entraîner
- Moins interprétable
- Plus expressif

Exemple : Reconnaissance vocale.

Un modèle **facilement calculable** est souvent **peu expressif**.

Inversement, un modèle **peu calculable** est souvent **expressif**.

Critères pour savoir si l'on doit choisir un modèle calculable ou expressif :

- Quantité de données à disposition
- Qualité du signal d'apprentissage dans les données
- Difficulté du problème à résoudre

Tableau récapitulatif

Famille	Expressivité	Calculabilité	Quantité de données	Interprétabilité
Symbolique	-/+	++/-	--	+/-
Probabiliste	+	+/-	+	++
Statistique	++	+/-	-/++	-/+
Deep Learning	+++	--	+++	---

Le Deep Learning rassemble des modèles **expressifs**, **difficiles à entraîner**, **peu interprétables** et nécessitant **beaucoup de données**.

Il faut donc majoritairement l'appliquer à des **problèmes difficiles** où **beaucoup de données** sont disponibles.

Les approches Big Data :

- Permettent de traiter des données plus volumineuses
- Sont **beaucoup** plus lourdes à mettre en place

→ Les utiliser avant des volumes de données vraiment importants est une perte de temps.

- Faire l'inventaire de vos problèmes d'apprentissage récents
- Les discuter en termes d'expressivité requise et de quantité de données