

Introduction au machine learning

Module 1

Objectifs

- cerner ce qu'est le machine learning
- appréhender les différentes facettes du domaine

Machine Learning

Qu'est-il pour vous ?



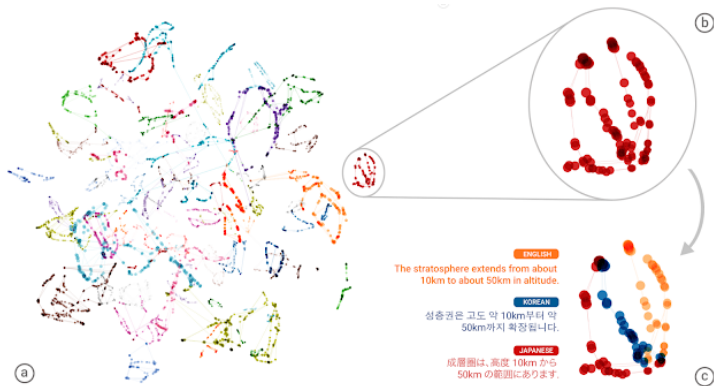
- pas de définition exacte
- idée transversale : éviter la programmation **explicite**.
- création de programmes qui utilisent des données ou des algorithmes généraux pour apprendre à réaliser leurs tâches

Beaucoup de façons de voir le machine learning. Basées sur :

- les paradigmes (supervisé, non supervisé, renforcement, en ligne, ...)
- les modèles (arbres, grammaires, automates, réseaux de neurones ...)
- les données (tabulaire, image, texte, vidéo, graphe, ...)
- les techniques (statistiques, symboliques, probabilistes, ...)
- les contraintes (real time, embarqué, big data, multilingue, ...)

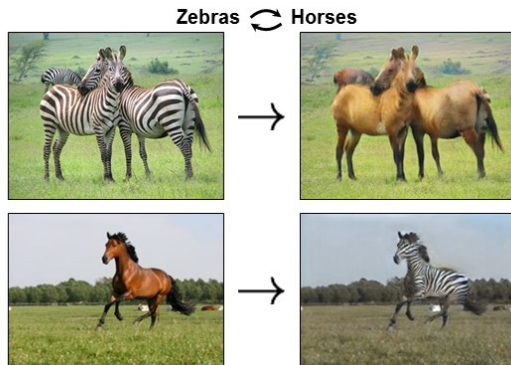
→ Domaine **extrêmement** vaste.

Facettes du machine learning — supervised learning



Demande beaucoup de données, parfois coûteuses. Modèles performants en sortie.

Facettes du machine learning — unsupervised learning



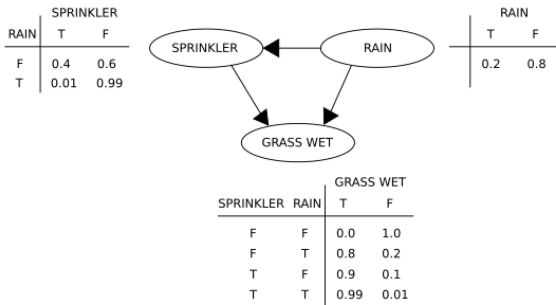
Pas besoin d'annotation → données moins couteuses. Limite les possibilités des modèles.

Facettes du machine learning — reinforcement learning



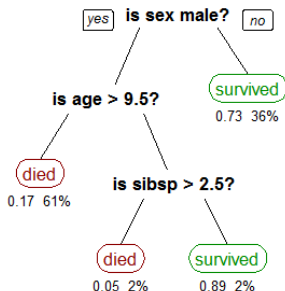
Paradigme de d'acquisition des données différent. Modèles potentiellement extrêmement performants.

Facettes du machine learning — bayesian networks



Très interprétable, requêtable.

Facettes du machine learning — decision trees



Assez interprétable, robuste, couteau-suisse du machine learning tabulaire.

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

Inversement, un modèle **peu calculable** est souvent **expressif** (sinon mauvais modèle).

Critères pour s'orienter dans les approches de machine learning :

- quantité de données à disposition
- qualité du signal d'apprentissage dans les données
- difficulté du problème à résoudre
- besoin d'interprétabilité
- contraintes techniques
- contraintes de délai
- ... et d'autres en fonction des domaines métiers

- le machine learning est un champ vaste.
- il existe sûrement un modèle/paradigme pour vos besoins
- l'important est de définir les bons critères

- à quelles données allez-vous appliquer le machine learning ? À quels besoins ?
- aurez-vous besoin de modèles interprétables ou simplement très performant en prédiction ?
- quelles sont vos contraintes ?