

Université Mohamed Boudiaf – M'sila
Département Informatique
Master Intelligence Artificielle
2020-2021

Chapitre 1
Introduction au Machine
Learning

Dr. Mehenni Tahar

Préambule

- De manière générale, un programme informatique tente de résoudre un problème pour lequel nous avons la solution. Par exemple : calculer la moyenne générale des étudiants, classer les étudiants selon leur moyenne. . .
- Pour certains problèmes, nous ne connaissons pas de solution exacte et donc nous ne pouvons pas écrire de programme informatique. Par exemple : reconnaître automatiquement des chiffres écrits à la main à partir d'une image scannée, déterminer automatiquement une typologie des clients d'une banque, jouer automatiquement aux échecs contre un humain ou un autre programme. . .
- En revanche, pour ces problèmes il est facile d'avoir une base de données regroupant de nombreuses instances du problème considéré.

Une définition

- L'apprentissage automatique consiste alors à programmer des algorithmes permettant d'apprendre automatiquement des données et d'expériences passées, un algorithme cherchant à résoudre au mieux un problème considéré.
- « **How** can we **build *computer systems*** that automatically improve with experience, and **what** are **the fundamental laws** that govern ***all learning processes?*** » *Tom Mitchell, 2006*

Apprentissage automatique (Machine Learning)

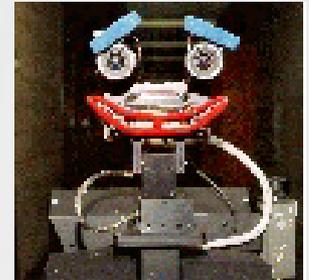
- Looking for a *model* of the world from **observations** in order to make *predictions* and to *understand*
- Chercher un modèle du monde réel à partir d'observations dans l'objectif de faire des **prévisions** ou de mieux **décrire**.

Apprentissage automatique (Machine Learning)

- **Changes** in a system that allows it to realize the *same type of tasks* than during training with a ever *better performance*.
- L'apprentissage est un **changment du système** dans le but de réaliser **les mêmes tâches** que celle de la période d'apprentissage mais avec une **meilleure performance**.

Motivations

- Concepts difficult to hand-code
 - Permissible moves for a robot
 - Person to recruit / or not
 - Predispositions for certain types of cancer
- Imitation
- Behavioral learning:
 - Learning to walk (Brooks's « insects »)
 - Learning to act on an unknown planet
- Learning to play
 - Adapt to the adversary
 - Learning to not repeat past faults
 - Learning to play within a team (teams of robots)
- Autonomous vehicle



Un domaine pluri-disciplinaire

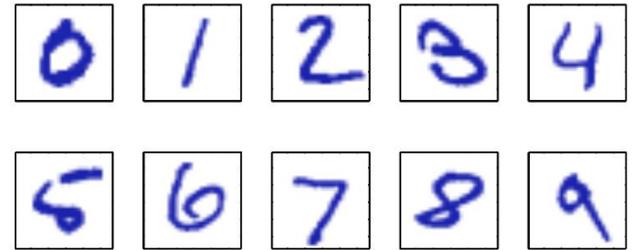
- L'apprentissage automatique (Machine Learning (ML)) est à la croisée de plusieurs disciplines :
- **Les statistiques** : pour l'inférence de modèles à partir de données.
- **Les probabilités** : pour modéliser l'aspect aléatoire inhérent aux données et au problème d'apprentissage.
- **L'intelligence artificielle** : pour étudier les tâches simples de reconnaissance de formes que font les humains (comme la reconnaissance de chiffres par exemple), et parce qu'elle fonde une branche du ML dite symbolique qui repose sur la logique et la représentation des connaissances.
- **L'optimisation** : pour optimiser un critère de performance afin, soit d'estimer des paramètres d'un modèle, soit de déterminer la meilleure décision à prendre étant donné une instance d'un problème.
- **L'informatique** : puisqu'il s'agit de programmer des algorithmes et qu'en ML ceux-ci peuvent être de grande complexité et gourmands en termes de ressources de calcul et de mémoire.

Exemple 1: prédiction d'achat

- Supposons qu'une entreprise veuille connaître le montant total dépensé par un client ou une cliente à partir de ses factures.
- Il suffit d'appliquer un algorithme classique, à savoir une simple addition : un algorithme d'apprentissage n'est pas nécessaire.
- Supposons maintenant que l'on veuille utiliser ces factures pour déterminer quels produits le client est le plus susceptible d'acheter dans un mois.
- Bien que cela soit vraisemblablement lié, nous n'avons manifestement pas toutes les informations nécessaires pour ce faire.
- Cependant, si nous disposons de l'historique d'achat d'un grand nombre d'individus, il devient possible d'utiliser un algorithme de machine learning pour qu'il en tire un modèle prédictif nous permettant d'apporter une réponse à notre question.

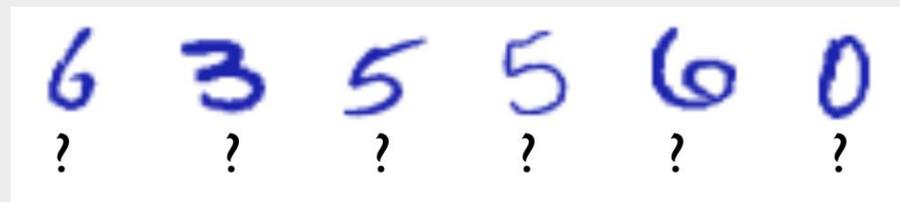
Exemple 2: Reconnaissance de caractère

- Comment reconnaître des Caractères manuscrits ?
- **par énumération de règles**
 - si intensité pixel à la position ...alors c'est un "3"
 - long et fastidieux, difficile de couvrir tous les cas
- **en demandant à la machine d'apprendre**
 - lui laisser faire des essais et apprendre de ses erreurs
 - apprentissage machine (machine-learning)



Exemple 2: Reconnaissance de caractère

- **Comment ça marche ?**
 - On donne à l'algorithme des données d'entraînement.
 - l'algorithme d'apprentissage machine apprend un modèle capable de généraliser à de nouvelles données.



Pourquoi utiliser le machine learning ?

Le machine learning peut servir à résoudre des problèmes

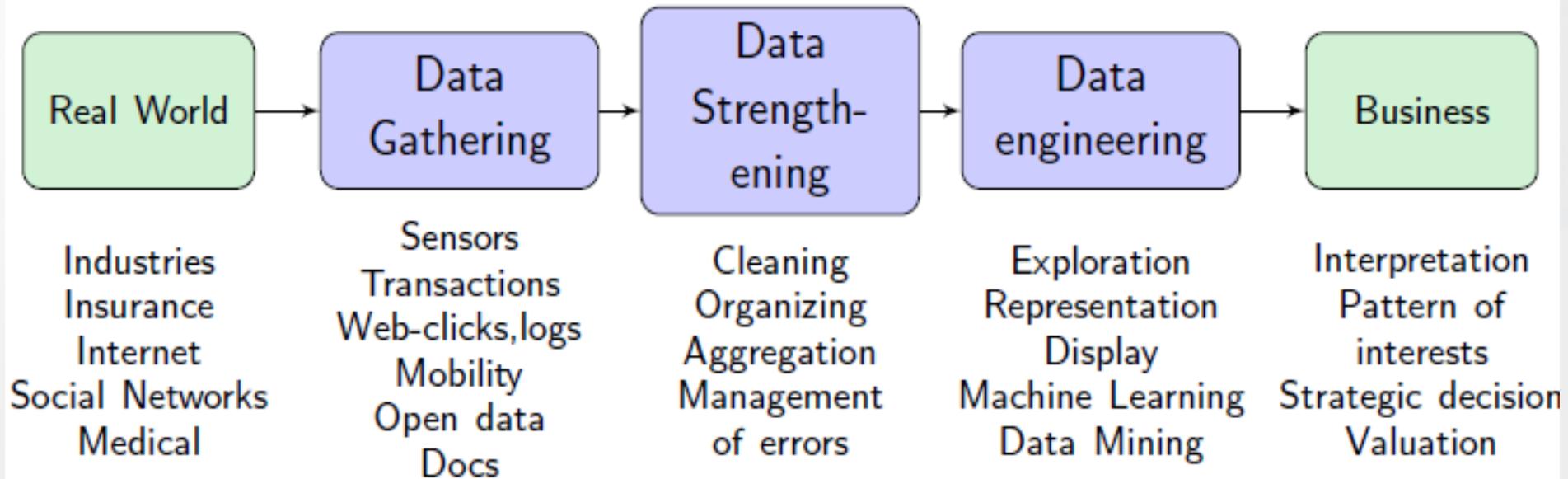
- que l'on ne sait pas résoudre (comme dans l'exemple de la prédiction d'achats ci-dessus) ;
- que l'on sait résoudre, mais dont on ne sait formaliser en termes algorithmiques comment nous les résolvons (c'est le cas par exemple de la reconnaissance d'images ou de la compréhension du langage naturel) ;
- que l'on sait résoudre, mais avec des procédures beaucoup trop gourmandes en ressources informatiques (c'est le cas par exemple de la prédiction d'interactions entre molécules de grande taille, pour lesquelles les simulations sont très lourdes).

Ingrédients du machine learning

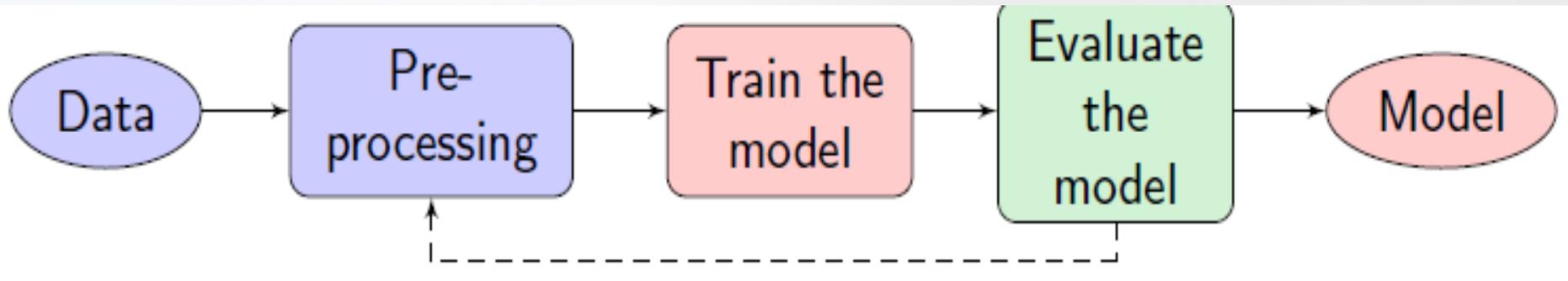
Le machine learning repose sur deux piliers fondamentaux :

- d'une part, **les données**, qui sont les exemples à partir duquel l'algorithme va apprendre ;
- d'autre part, **l'algorithme d'apprentissage**, qui est la procédure que l'on fait tourner sur ces données pour produire un modèle. On appelle *entraînement* le fait de faire tourner un algorithme d'apprentissage sur un jeu de données.

Implémenter un projet Machine Learning



Procéssus d'ingénierie des données



- 1 Understand and specify project goals
- 2 Pre-processing/visualize/analyze data
- 3 Which ML problem is it?
- 4 Design a solving approach
- 5 Evaluate its performance
- 6 Go to to 2) if needed

Approaches of Machine Learning

