

INTRODUCTION AU MACHINE LEARNING

MOTIVATION ET FORMALISATION

Théo Lopès-Quintas

BPCE Payment Services,
Université Paris Dauphine

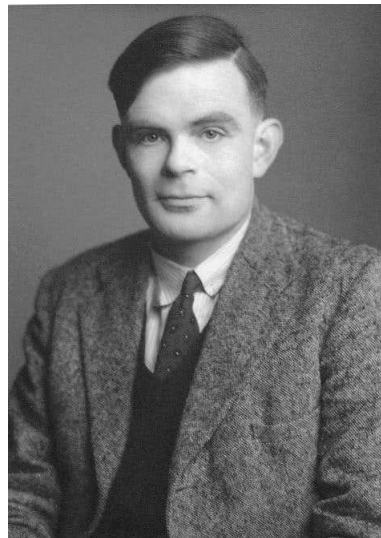
2022-2024

MOTIVATION

FUTUR

Nous ne pouvons qu'avoir un aperçu du futur, mais cela suffit pour comprendre qu'il y a beaucoup à faire.

— Alan Turing (1950)



MOTIVATION

EXERCICE

En règle générale, je dirais que l'on n'apprend que dans les cours où l'on travaille sur des problèmes. Il est essentiel que les étudiants tentent de résoudre des problèmes. [...] Se borner à écouter ne sert pas à grand chose.

— Werner Heisenberg (1963)



MOTIVATION

MATHÉMATIQUES

La logique ne fait que sanctionner les conquêtes de l'intuition.

— Jacques Hadamard (1972)



LES DIFFÉRENTES APPROCHES DU MACHINE LEARNING

SUPERVISÉ OU NON-SUPERVISÉ

LES DIFFÉRENTES APPROCHES DU MACHINE LEARNING

SUPERVISÉ OU NON-SUPERVISÉ

Exercice 1

Nous travaillons dans une concession automobile et avons à notre disposition une base de données avec l'ensemble des caractéristiques de chaque voiture, chaque ligne de cette base de données étant un modèle de voiture que l'on vend.

Donner pour chaque demande le type d'approche que l'on peut suivre.

- 1. Prédire le type de voiture*
- 2. Visualiser en deux dimensions la base de données*
- 3. Prédire le prix d'une voiture*
- 4. Recommander des voitures à un client se rapprochant de sa voiture de rêve*

APPRENTISSAGE SUPERVISÉ

DATASET

$$\mathcal{D} = \left\{ (x^{(i)}, y_i) \mid \forall i \leq n, x^{(i)} \in \mathbb{R}^d, y_i \in \mathcal{Y} \right\} \quad (1)$$

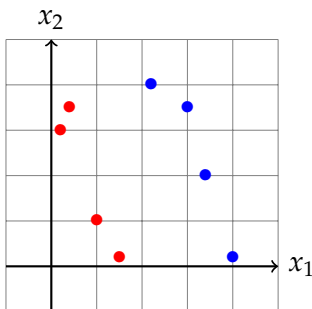
Nombre d'observations (pointing to n)

Nombre d'informations (pointing to d)

APPRENTISSAGE SUPERVISÉ

FONCTION DE PERTE ET FONCTION DE COÛT

Exercice 2



Trouver la meilleure fonction f_θ qui renvoie 0 pour les ronds bleus et 1 pour les ronds rouges parmi les propositions suivantes :

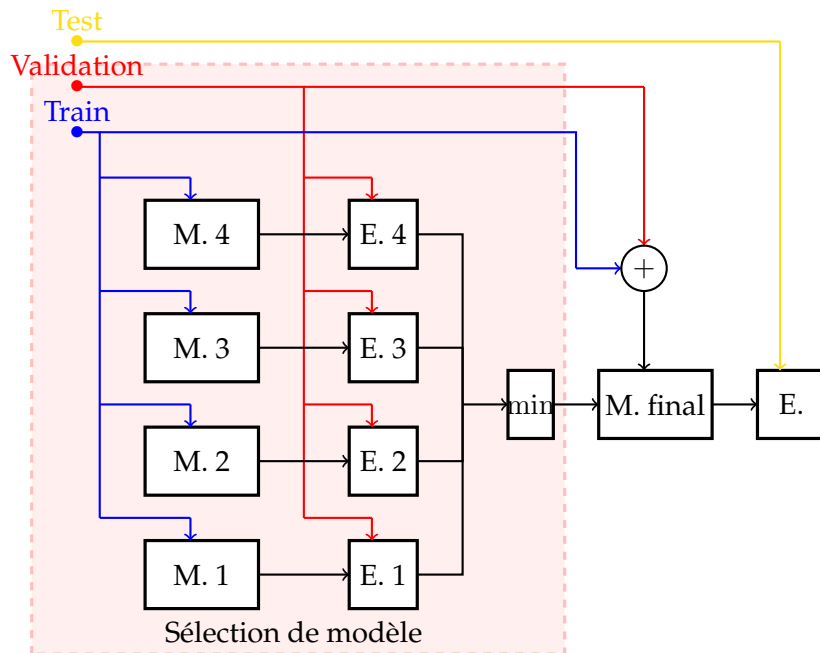
1. $f_\theta(x_1, x_2) = \mathbb{1}_{x_1 \leq \theta}$

2. $f_\theta(x_1, x_2) = \mathbb{1}_{x_2 \leq \theta}$

Avec $\mathcal{L}(\theta; x, y) = \mathbb{1}_{y \neq f_\theta(x_1, x_2)}$ donc $\mathcal{C}(\theta; X, y) = \sum_{i=1}^n \mathcal{L}(\theta, x, y)$.

SÉLECTION DE MODÈLE

TRAIN, VALIDATION, TEST



SÉLECTION DE MODÈLE

VALIDATION CROISÉE

