

# Deep Learning

Paul Lehaut

October 1, 2025

# Contents

<b>1 Bases du Machine Learning</b>	<b>3</b>
1.1 Types d'apprentissage . . . . .	3
1.2 Apprentissage Supervisé Paramétrique . . . . .	3
<b>2 Backpropagation</b>	<b>3</b>

# 1 Bases du Machine Learning

Le Machine Learning est une discipline qui vise à concevoir des algorithmes généraux applicables à de nombreux problèmes et ce en partant uniquement des données.

Le data engineering consiste en l'étude approfondie des données et à les rendre exploitables. Il s'agit d'une étape cruciale qui concentre une grande partie du travail dans les problèmes réelles.

## 1.1 Types d'apprentissage

Selon les scénarios, il existe différents types d'apprentissage:

- L'apprentissage supervisé: les données sont composées à la fois des données d'entrées et des sorties attendues. Le but est d'implémenter un algorithme capable de généraliser proprement pour de nouvelles entrées.
- L'apprentissage non supervisé: les données sont fournies sans les sorties attendues, il s'agit alors de trouver des structures au sein des données (typiquement à l'aide de méthodes de clustering).
- Il en existe encore d'autres type comme le reinforcement learning qui permet à un algorithme d'interagir avec son environnement.

Il faut également faire la distinction entre deux types de méthodes:

- La méthode paramétrique: elle définit un ensemble de fonctions pour lesquelles il convient de trouver les meilleurs paramètres.
- La méthode non paramétrique: elle dépend directement des données.

## 1.2 Apprentissage Supervisé Paramétrique

Soit  $X$  un espace d'entrée et  $Y$  un espace de sortie, on va chercher à apprendre une fonction de prédiction:  $f : X \rightarrow Y$  paramétrée par un paramètre  $\theta$ . On va donc chercher les, voire les, paramètre  $\theta$  qui maximise la performance de la fonction  $f$ .

Formellement, on définit une fonction de coût  $l(y', y)$ , qui mesure l'erreur entre la prédiction  $y'$  et la valeur attendue  $y$ , ainsi qu'une fonction de risque:  $R(f) = \mathbb{E}_{(x,y)}(l(f(x), y))$ .

Malheureusement, on ne connaît jamais la distribution réelle des données mais plutôt simplement un échantillon  $Z$ . On définit alors le risque empirique:

$$R_Z(f) = \frac{1}{|Z|} \sum_{(x,y) \in Z} l(f(x), y).$$

L'apprentissage consiste alors à minimiser une fonction de perte définie à partir de ce risque et, éventuellement, d'autres termes.

Pour éviter qu'un modèle over/underfit les données, il est souvent judicieux de séparer les données en trois sets: un pour l'entraînement (train set), un pour la validation du modèle (validation set) et enfin un pour estimer la performance réelle (test set).

# 2 Backpropagation

La backpropagation se base essentiellement sur la règle de dérivation en chaîne. Cet algorithme considère chaque opération différentiable élémentaire comme un module capable de calculer;