



Chapitre 1

Introduction

Savoirs et compétences :

□ *A voir*

1 Définitions

- Data scientist
- machine learning
- deep learning
- réseaux de neurons
- régression
- data mining
- bigdata
- données continues, données discrètes, données nominales, données ordinales, données (semi-)structurées et non structurées
- algorithmes supervisés, non supervisés
- algorithmes de régression et de classification

Définition — Machine Learning – Apprentissage automatique – Wikipedia. Champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'« apprendre » à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune.

La première phase de l'apprentissage consiste à estimer un modèle à partir de données, appelées observations, qui sont disponibles et en nombre fini, lors de la phase de conception du système.

La seconde phase correspond à la mise en production : le modèle étant déterminé, de nouvelles données peuvent alors être soumises afin d'obtenir le résultat correspondant à la tâche souhaitée.

Définition — Apprentissage supervisé – Apprentissage non supervisé – Apprentissage par renforcement – Wikipedia. Si les données sont étiquetées (c'est-à-dire que la réponse à la tâche est connue pour ces données), il s'agit d'un apprentissage supervisé. On parle de :

- classification ou de classement si les étiquettes sont discrètes ;
- régression si elles sont continues.

Si le modèle est appris de manière incrémentale en fonction d'une récompense reçue par le programme pour chacune des actions entreprises, on parle d'apprentissage par renforcement.

Dans le cas le plus général, sans étiquette, on cherche à déterminer la structure sous-jacente des données (qui peuvent être une densité de probabilité) et il s'agit alors d'apprentissage non supervisé.

Exemples : <https://makina-corpus.com/blog/metier/2017/initiation-au-machine-learning-avec-python>

Références

[1] Éric Biernat et Michel Lutz. *Data science : fondamentaux et études de cas*. Eyrolles.