

Algorithmes d'apprentissage

non supervisé

01-01

Introduction au partitionnement de données

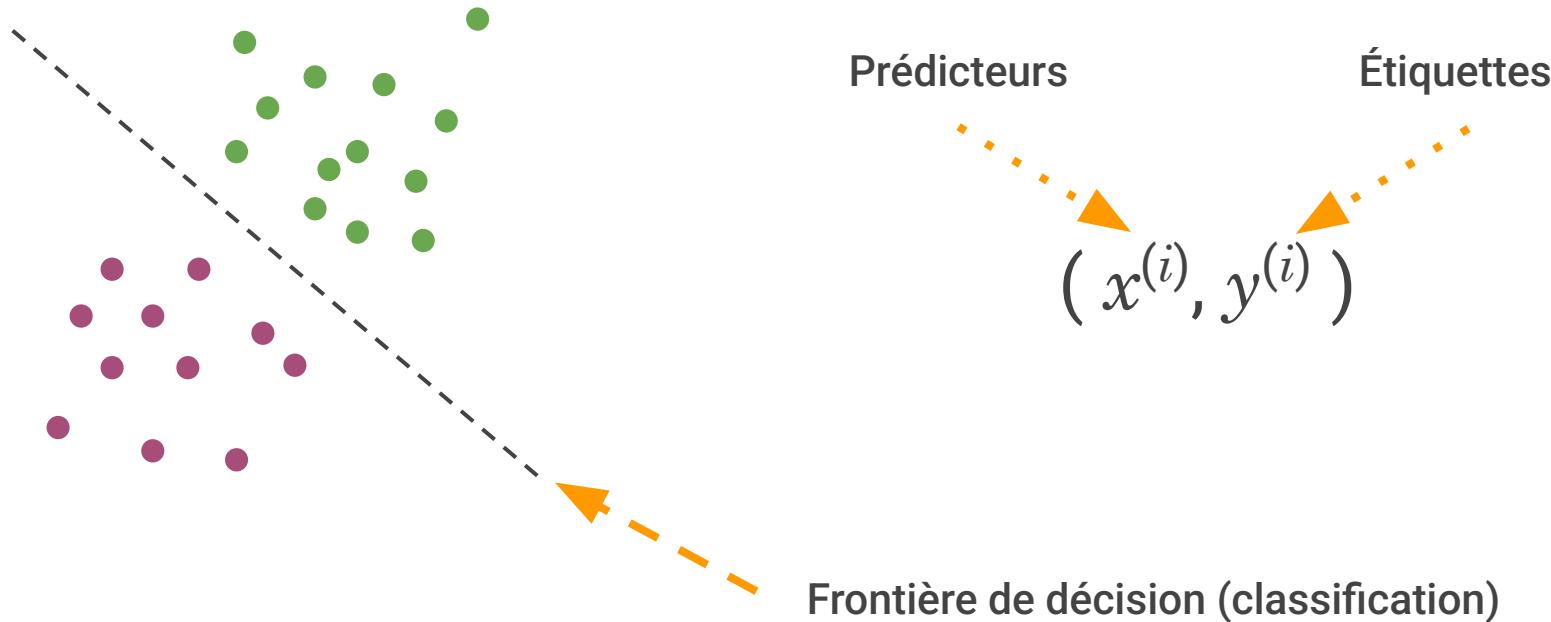
Au programme

- Introduction
- Les principaux types de partitionnement
- Applications
- Lecture obligatoire
- Références

Introduction

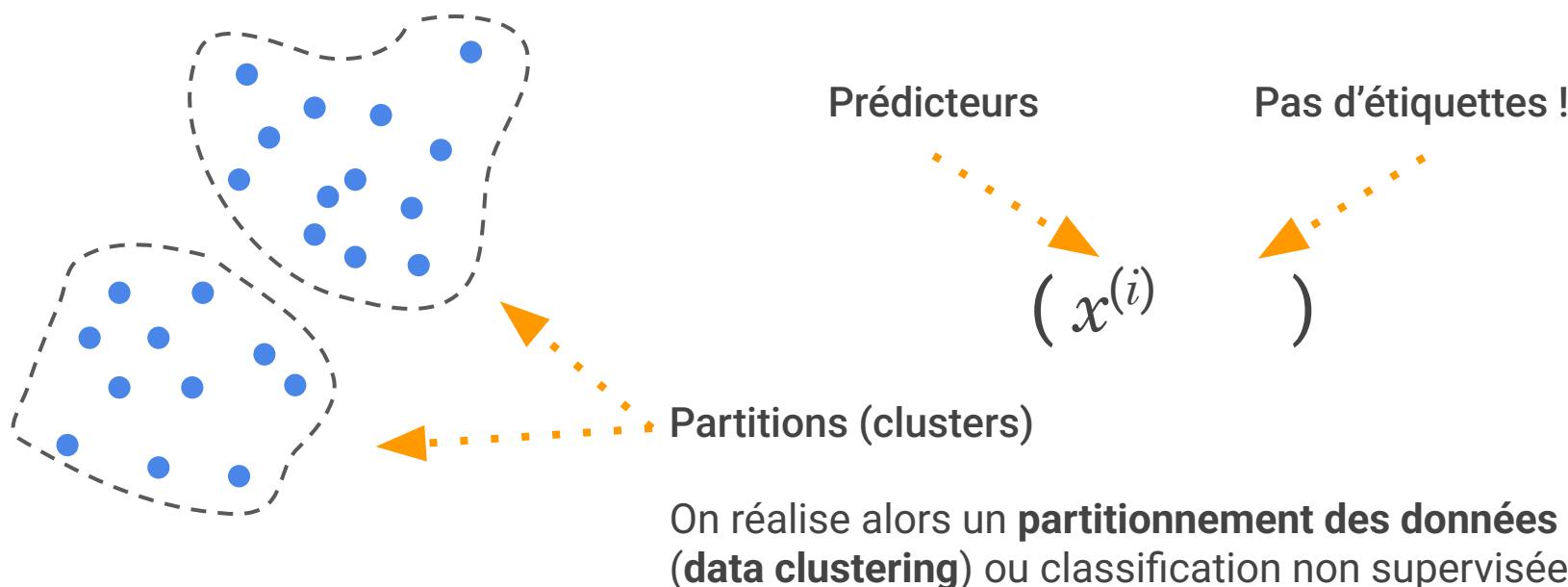
Introduction

Rappelez-vous, en apprentissage supervisé ...



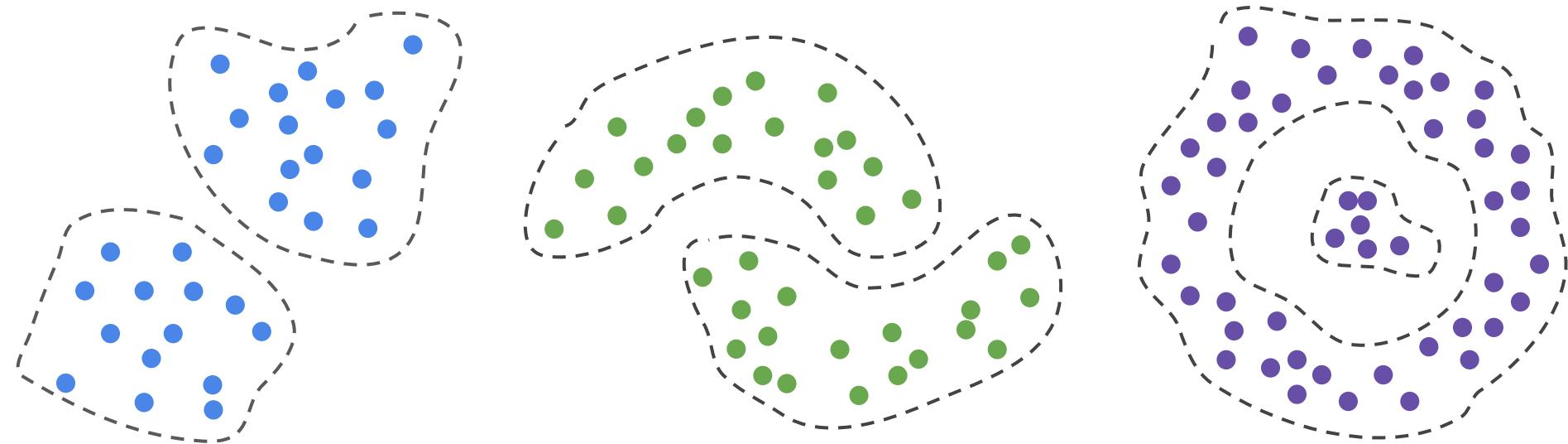
Introduction

Maintenant, en apprentissage non supervisé



Introduction

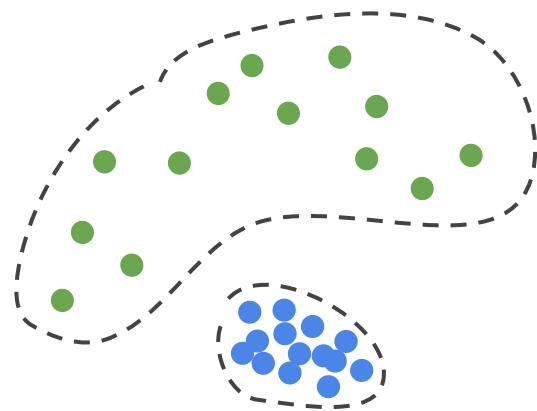
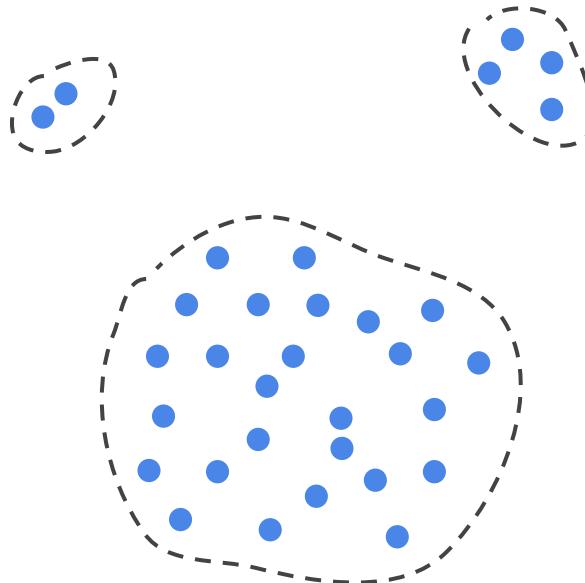
Les clusters peuvent avoir différentes morphologies



Introduction

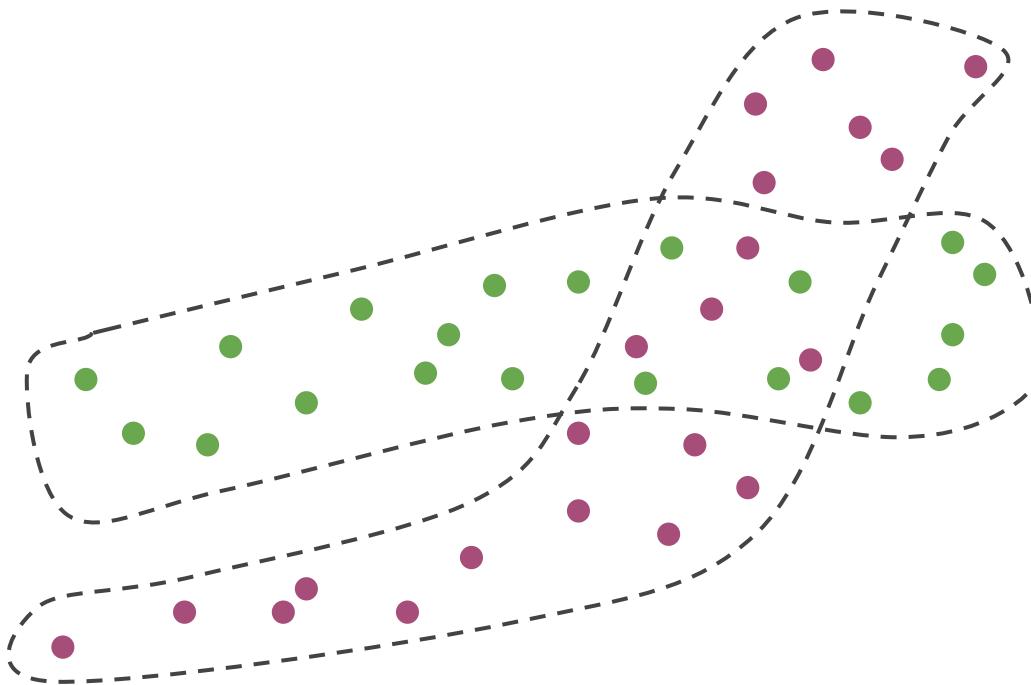
Ils peuvent être de différentes tailles ...

... et de différentes densités.



Introduction

Dans certains cas, il peut même y avoir chevauchement



Les principaux types de partitionnement

Les principaux types de partitionnement

- Partitionnement basé sur
 - les centroïdes (K-moyennes, CURE, ...)
 - la connectivité (hiérarchique, ...)
 - la distribution (BFR, ...)
 - la densité (DBSCAN, OPTICS, ...)
 - les grilles
- Et d'autres

Les principaux types de partitionnement

- Partitionnement basé sur
 - les centroïdes (**K-moyennes**, CURE, ...) ➡️
 - la connectivité (**hiérarchique**, ...) ➡️
 - la distribution (BFR, ...)
 - la densité (**DBSCAN**, OPTICS, ...) ➡️
 - les grilles
- Et d'autres

Applications

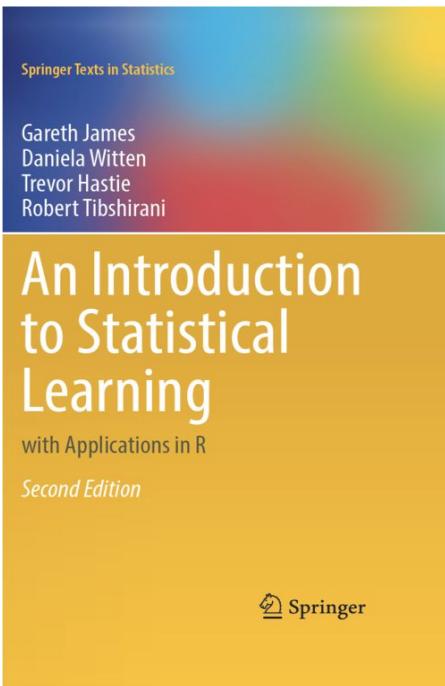


Applications du partitionnement

- Optimisation de portfolio en finance
- Centres de données
- Astrophysique
- Cybersécurité
- Marketing
- Génétique
- Télédétection
- ...

Lecture obligatoire

Lecture obligatoire



- Introduction to Statistical Learning with Applications in R
Second edition (2021)

Lire le chapitre 12.4 Clustering Methods
Pages 516-532

Références

Références

- [1] CS229: Machine Learning - Stanford University
- [2] Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, “Introduction to Statistical Learning with Applications in R - Second edition”
- [3] [Popular Unsupervised Clustering Algorithms \(Kaggle\)](#)