

# Machine Learning

Travaux Pratiques

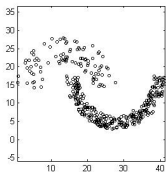
---

Il y a comme d'habitude des fichiers ipynb provenant de la documentation de sklearn pour apprendre à manipuler des algorithmes de clustering

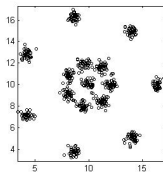
Pour la prochaine fois :

- Vous pouvez vous amuser avec les données dans `./data/toyproblem/`
- Choisissez un dataset dans `./data/real-data/` et
  - Afficher et présenter un résultat avec une méthode non-supervisée
  - Idem avec une méthode supervisée

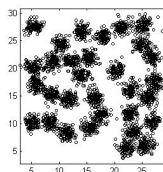
Des problèmes jouets pour comprendre les limitations des algorithmes de clustering :



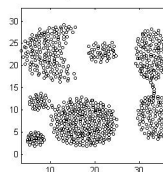
A.K. Jain's Toy problem



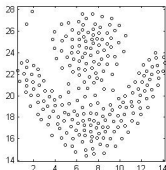
R15



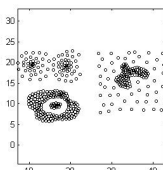
D31



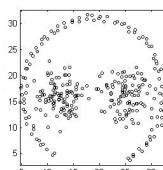
Aggregation



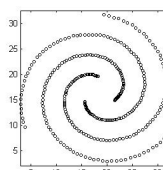
Flame



Zahn's Compound



Path-based1



path-based2: spiral

Les caractéristiques détaillées de chaque joueur dans FIFA 19

<https://www.kaggle.com/karangadiya/fifa19>

Des données sur toutes les apps du Play Store de Google

<https://www.kaggle.com/lava18/google-play-store-apps>

Envie d'aider à combattre les maladies du coeur ?

<https://www.kaggle.com/ronitf/heart-disease-uci>

2.0GB de données socio-économiques agrégée par quartier en direct des USA.

[https://www.kaggle.com/safegraph/  
census-block-group-american-community-survey-data](https://www.kaggle.com/safegraph/census-block-group-american-community-survey-data)