

Machine Learning

De l'importance des **données**

Données à dimension fixe ou dimension variable ?

- Image
- Texte (tf-idf : 50000 mots)
- ligne dans une table
- ...

Données à dimension variable

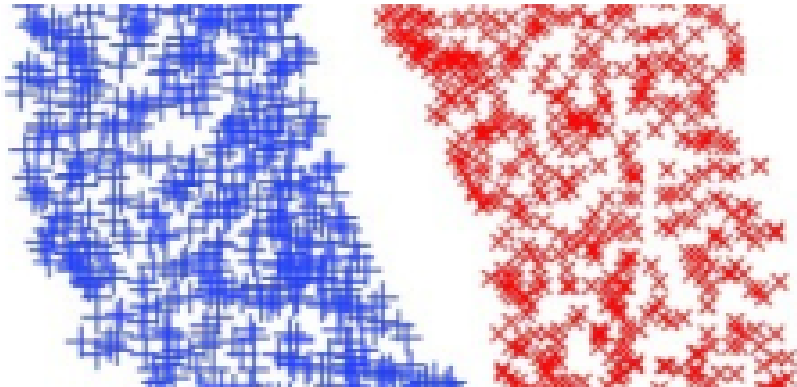
Données ancrées dans le temps.

Séquence de données à dimension fixe

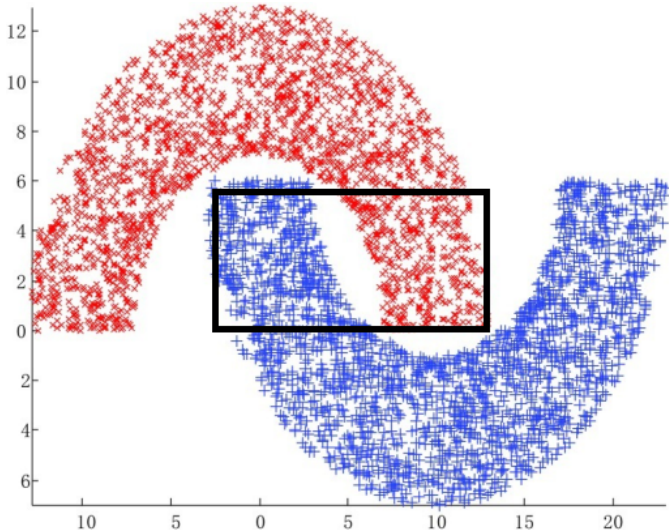
- Son
- Video
- Texte
- ...

La collection de **données** est l'élément central du machine learning, bien plus que l'algorithme d'apprentissage lui-même !

Quelle bonne méthode de séparation ?

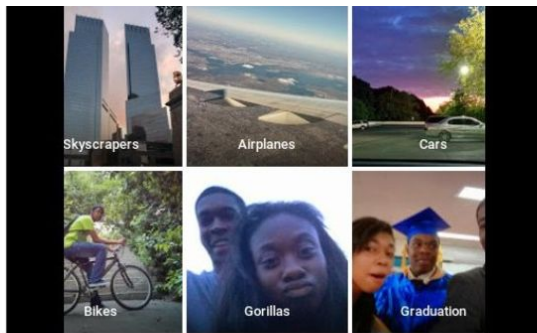


Quelle bonne méthode de séparation ?



Il faut porter une attention toute particulière aux données :

- Qualité Vs Quantité ?
- Représentation en vecteur ?
- Données trouées, équilibrées,... ?
- Éviter la fuite d'informations dégénérantes pour le modèle !



diri noir avec banan @jackyalcine · Jun 29

Google Photos, y'all [redacted] My friend's not a gorilla.



813



394



TWITTER