

# **Introduction au Machine Learning**

# Présentation

- Le machine learning ???
- Comment ça marche ?
- La mise en oeuvre
- Quelques modèles

# **Le machine learning ???**

# Vocabulaire

- artificial intelligence (AI)
- machine learning (ML)
- deep learning (DL)

# Définition

Le but du ML est de **prendre des décisions** ou **faire des prédictions** en se basant sur les **données**.

Les **données** sont **primordiales**.

# Prendre des décisions

# Faire des prédictions



# Applications

- La recommandation de musique
- Le filtrage des mails
- Les feeds sur les réseaux sociaux
- Les assistants vocaux
- Les assistants type ChatGPT
- Les voitures autonomes
- ...

# Les principaux types

- apprentissage supervisé
- apprentissage non-supervisé
- apprentissage par renforcement

# L'apprentissage supervisé

Trouver une fonction  $f(X)$  telle que:

$$f(X) = Y$$

avec  $X$  : les exemples

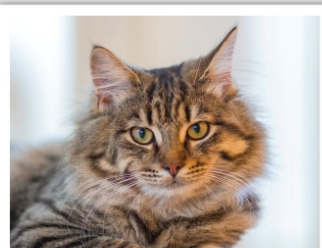
avec  $Y$  : les étiquettes

Image

Label



Cat



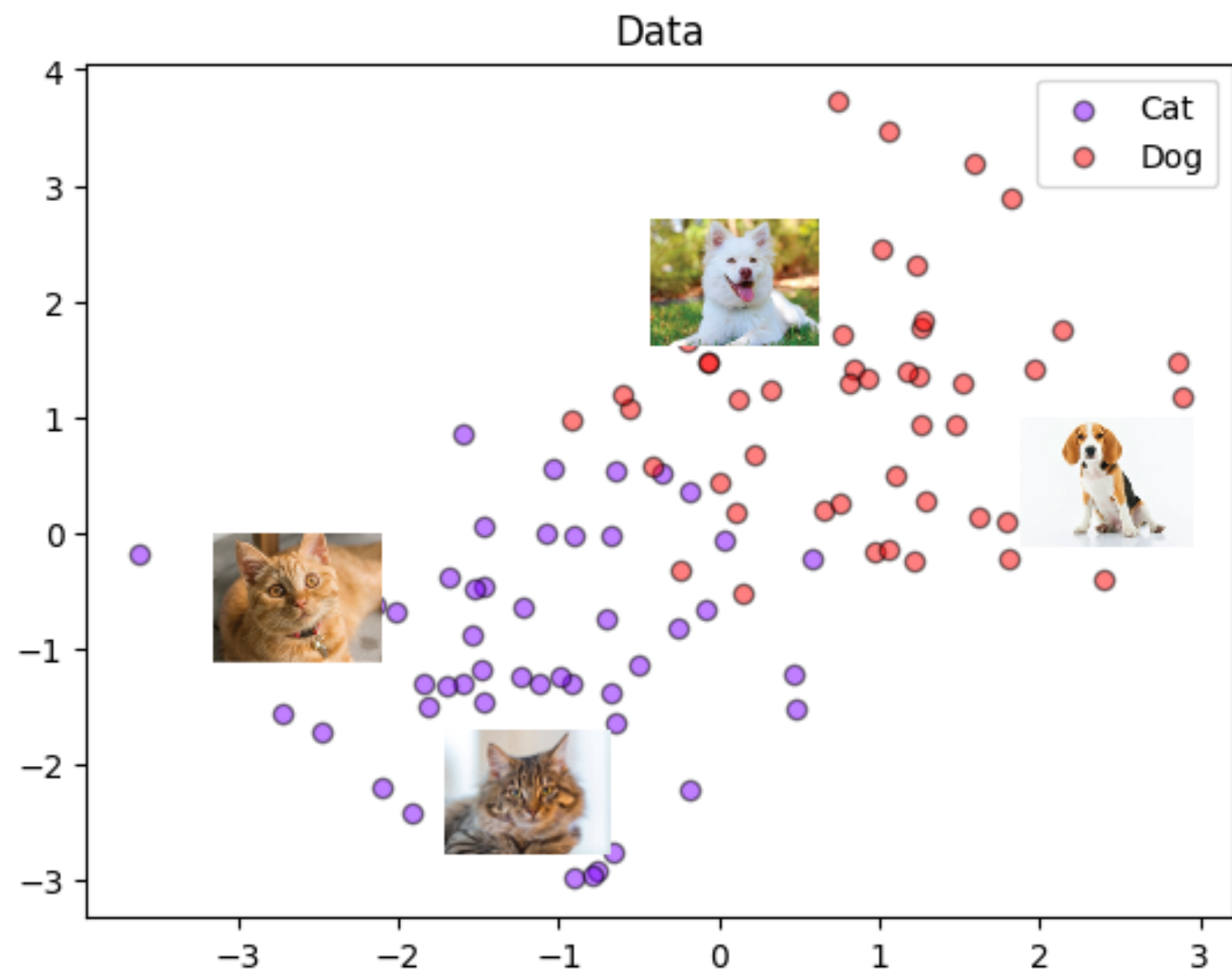
Cat

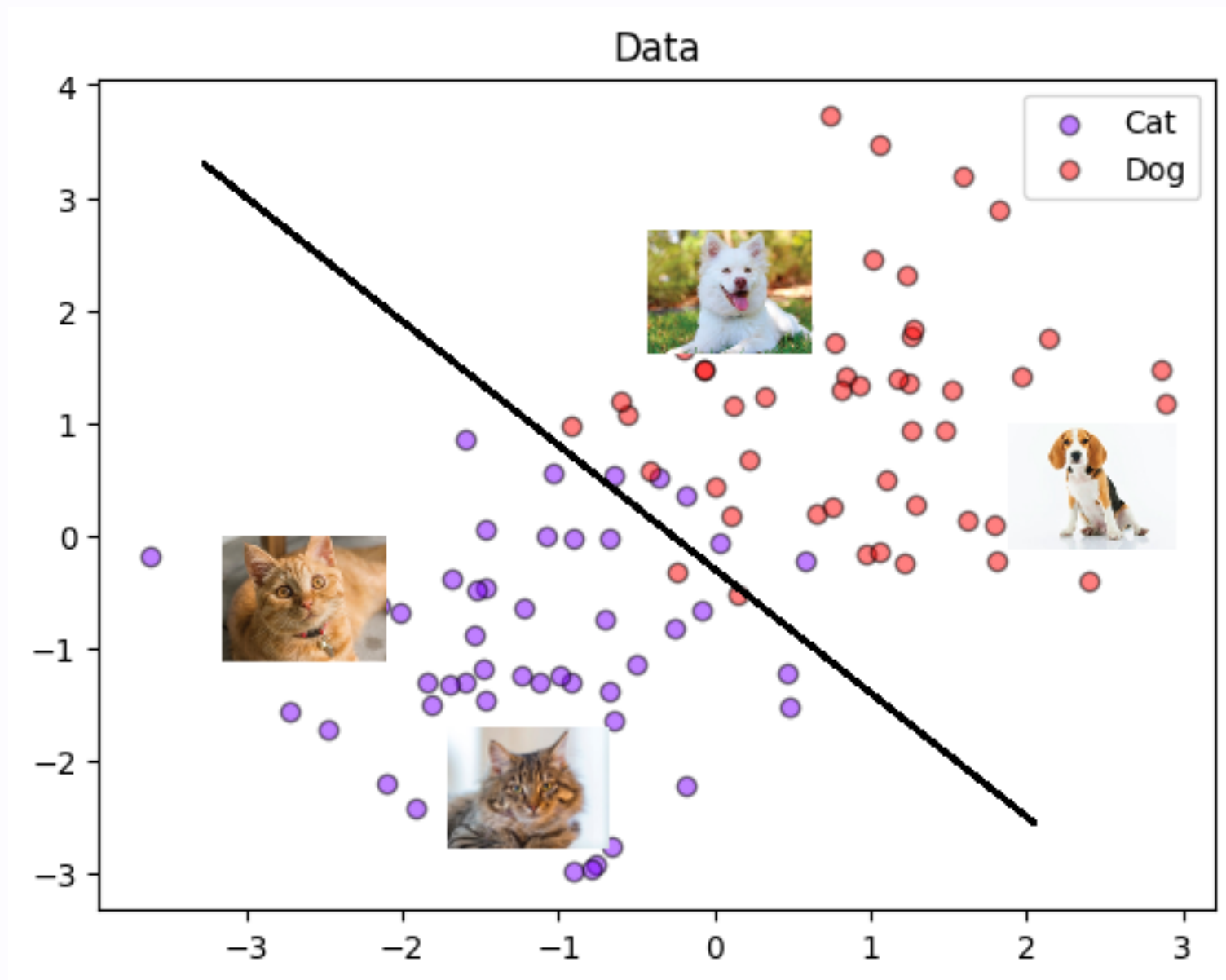


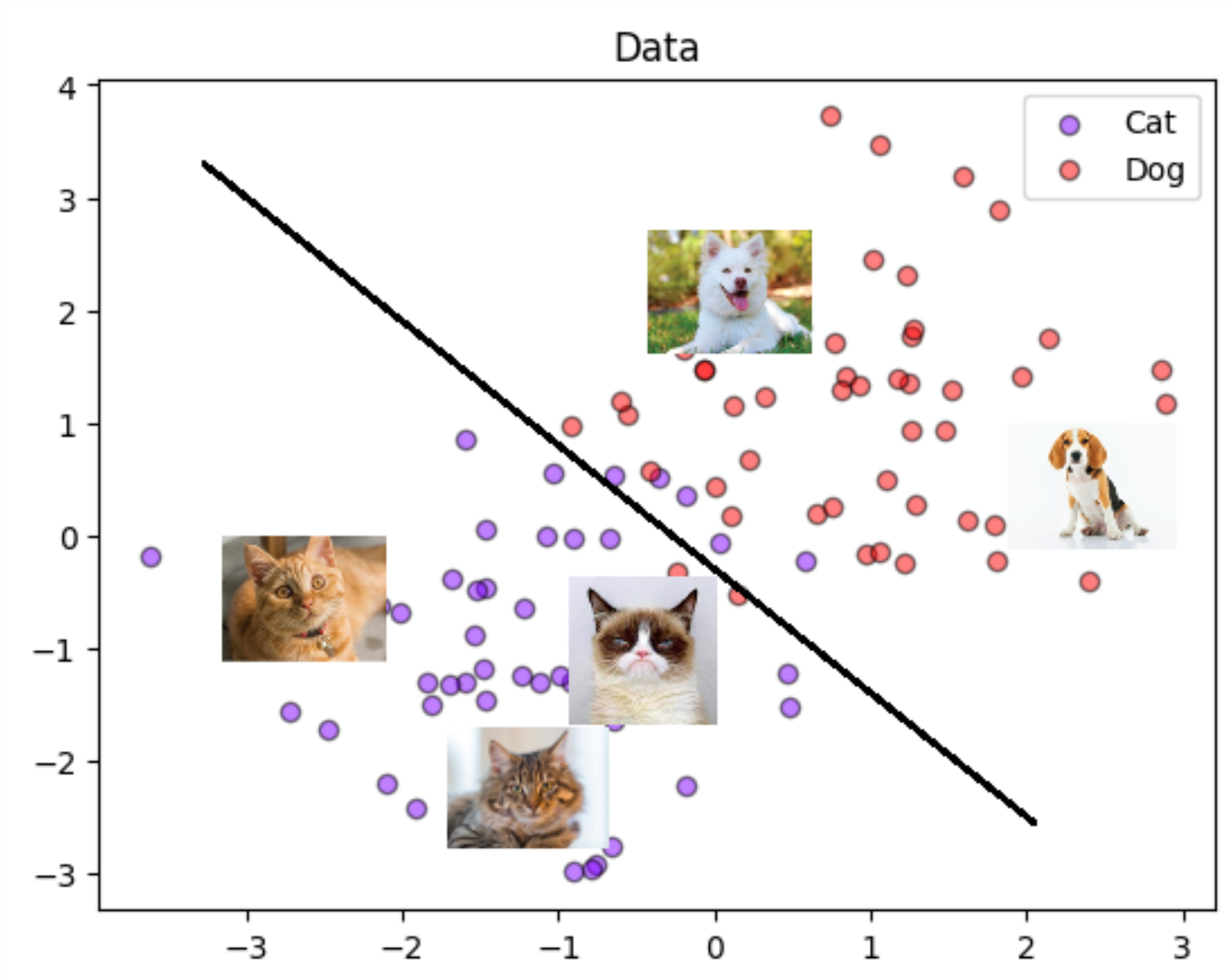
Dog




Dog







$$f(\text{img}) = \text{cat}$$
A close-up photograph of Grumpy Cat, a brown and white cat with a grumpy expression. The cat has a white patch on its face and chest, and its eyes are squinted and slanted downwards, giving it a perpetually grumpy look. The background is a plain, light gray.



$f(\text{img}) = \text{cat}$



f(



)

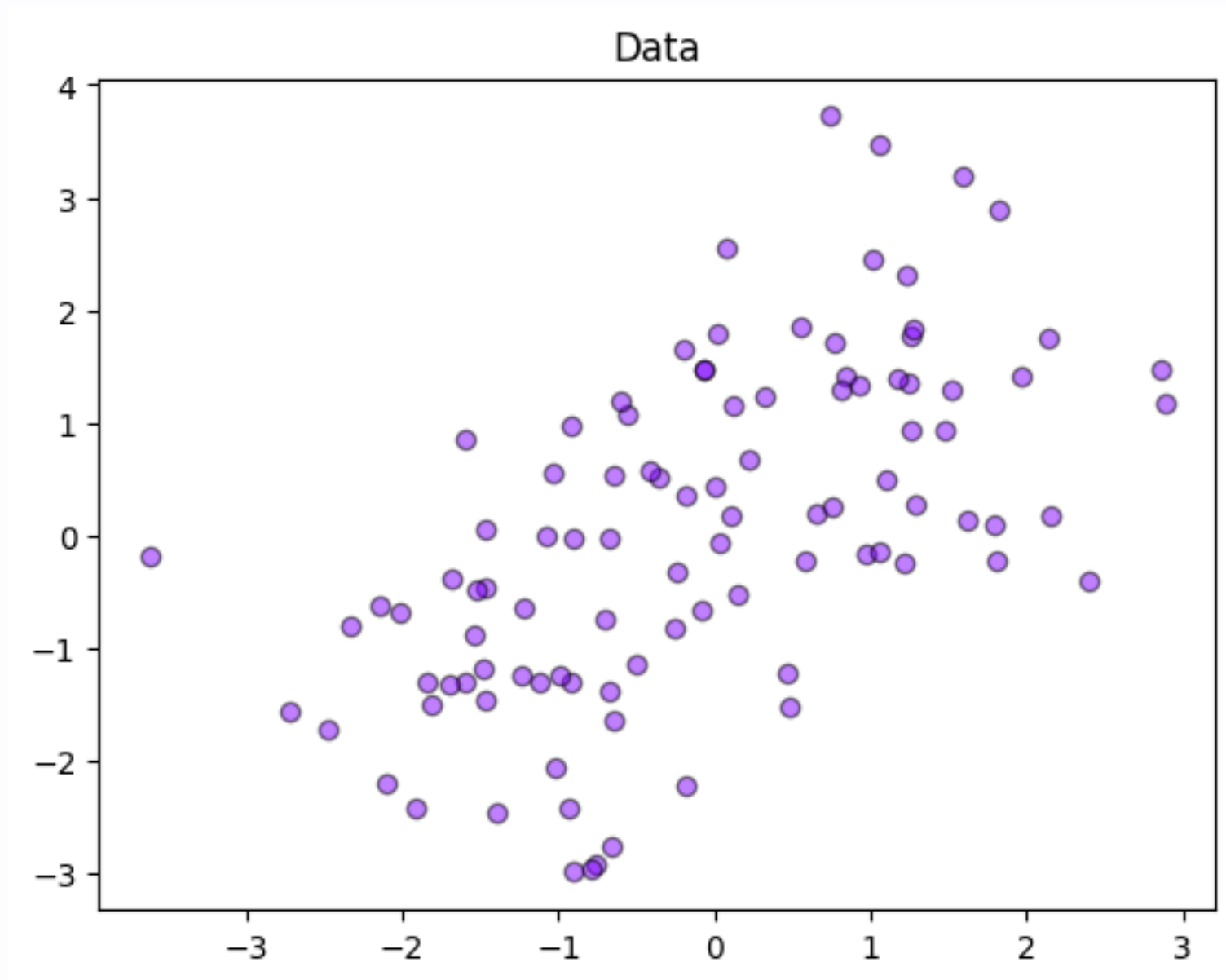
= dog

# **L'apprentissage non-supervisé**

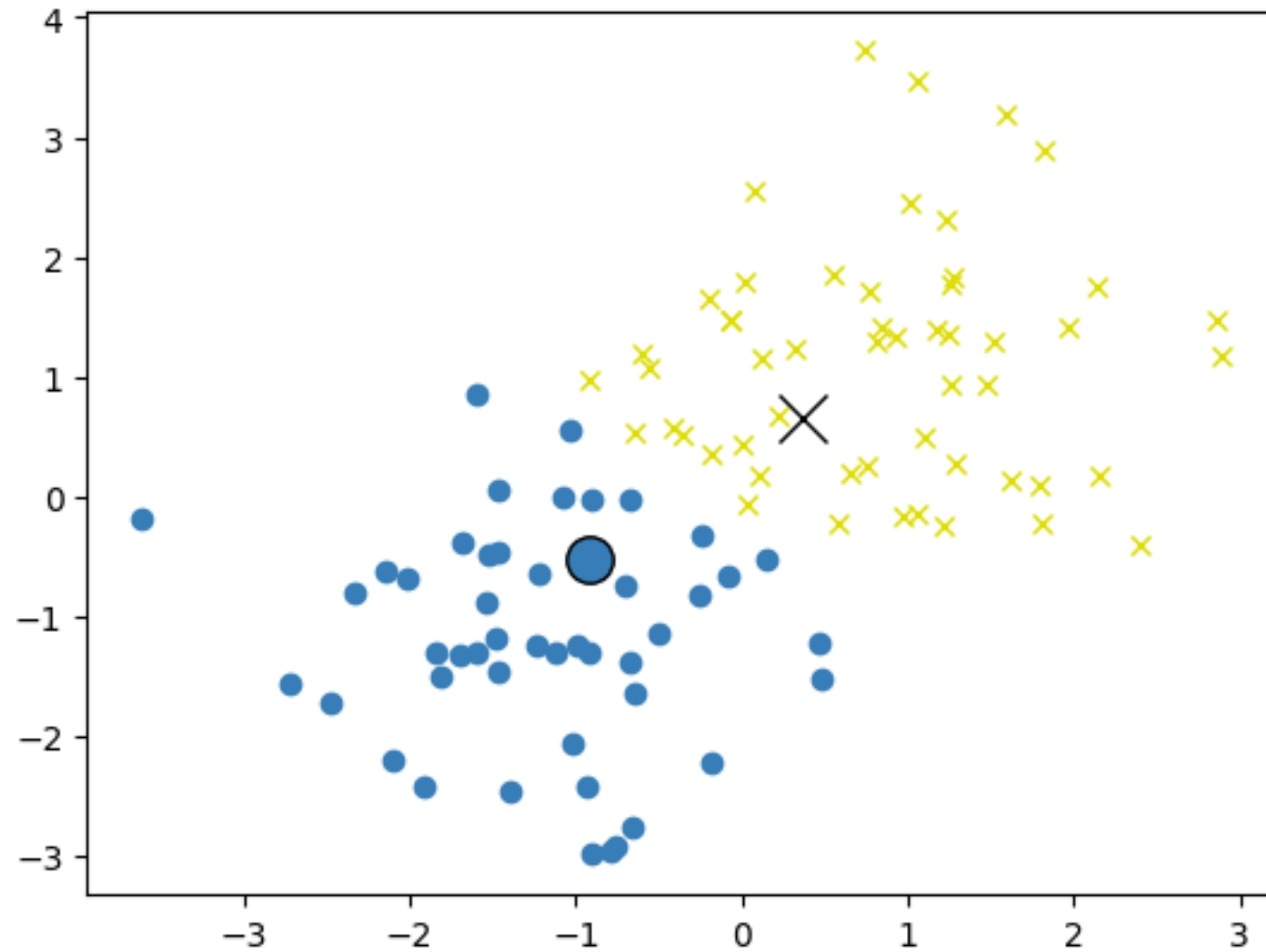
Trouver des patterns ou des structures cachés dans des données.

## Image

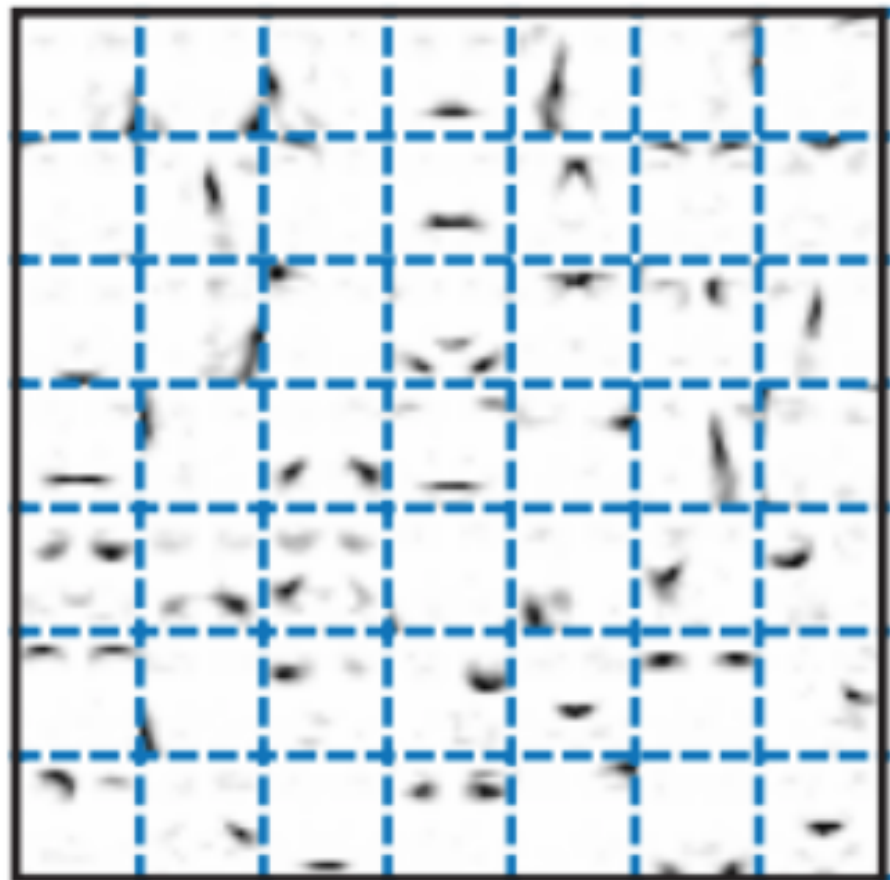




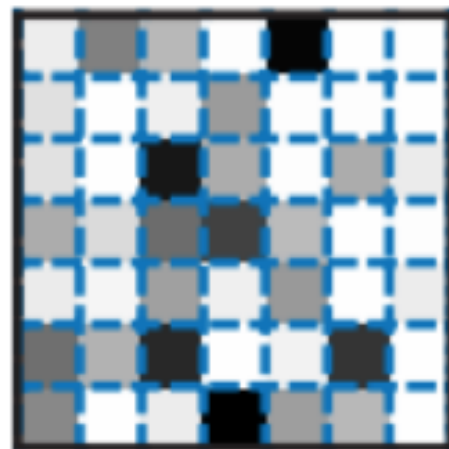
Estimated number of clusters: 2



NMF



$\times$



$=$

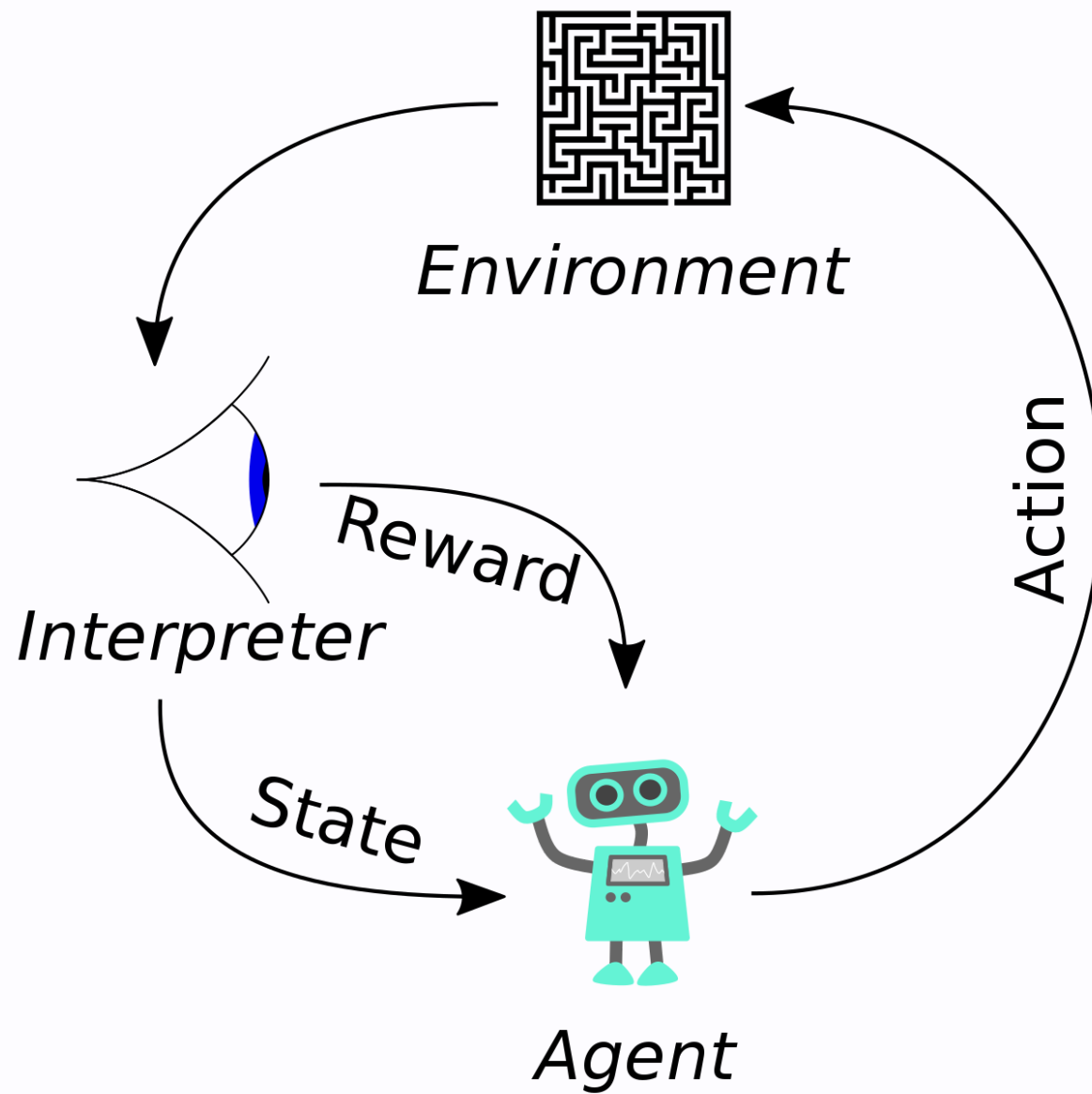
Original

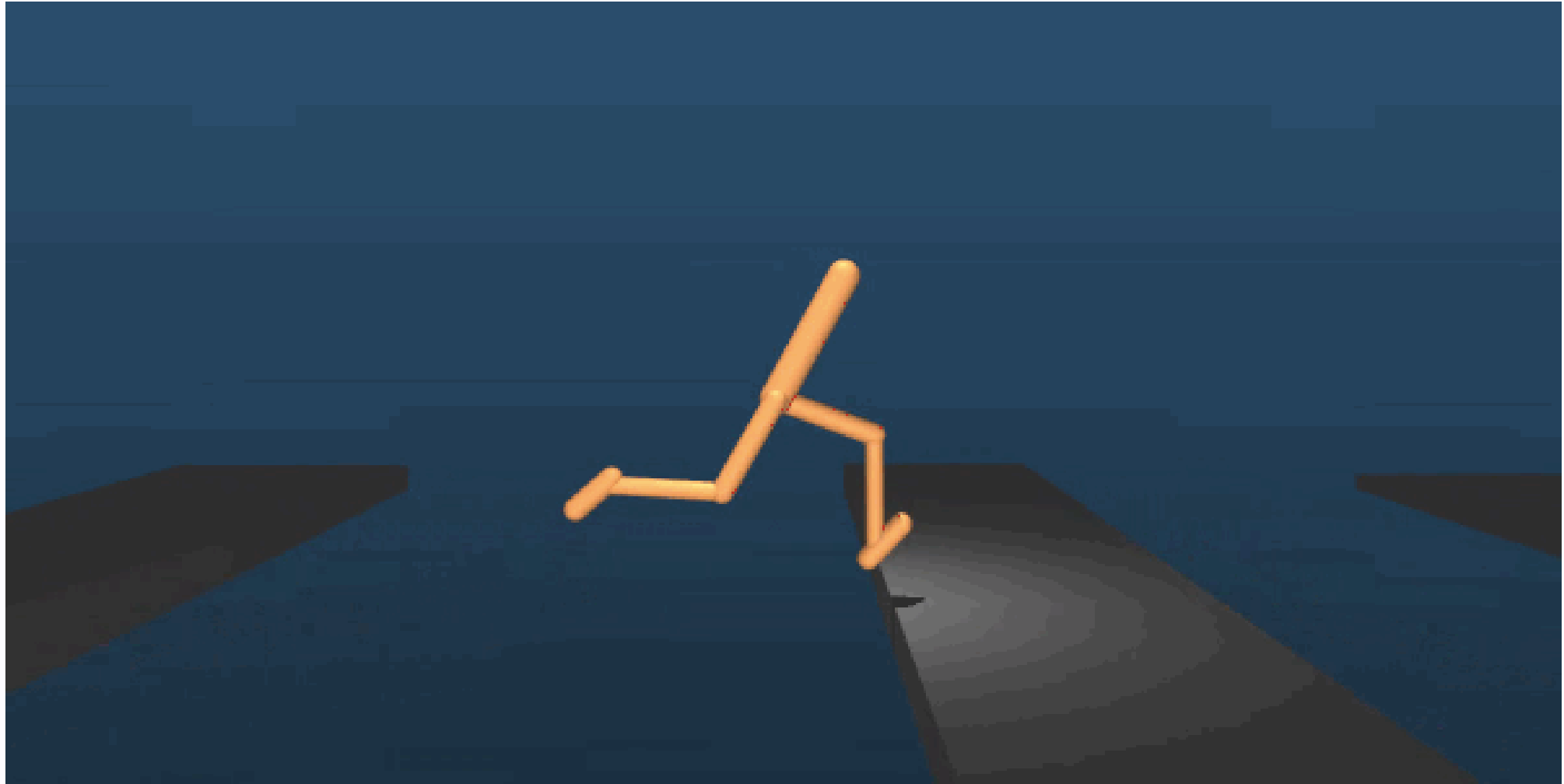


# **L'apprentissage par renforcement**

Faire évoluer un agent dans un environnement afin qu'il apprenne à réaliser des actions qui le récompensent.





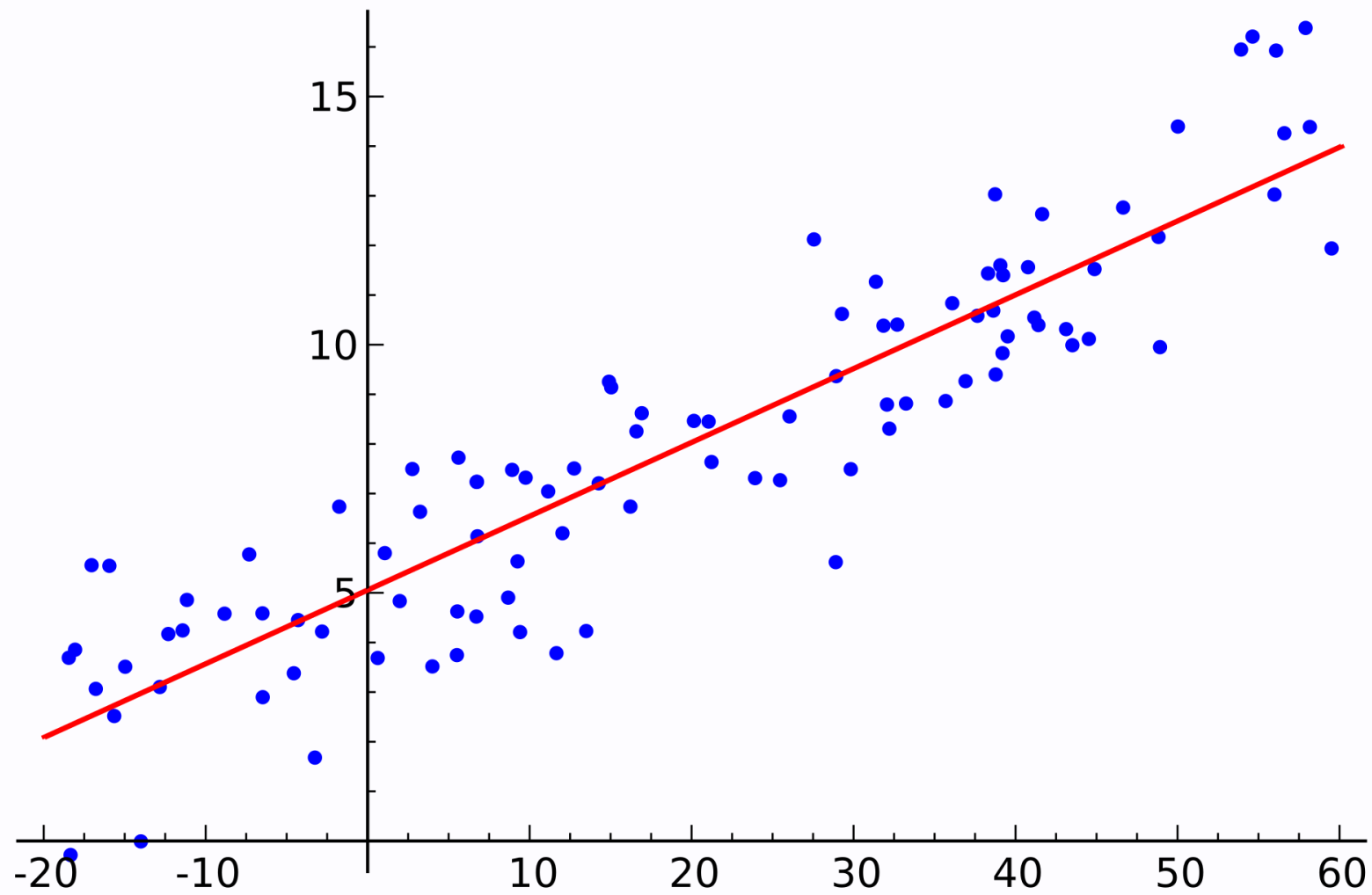




**Comment ça marche ?**

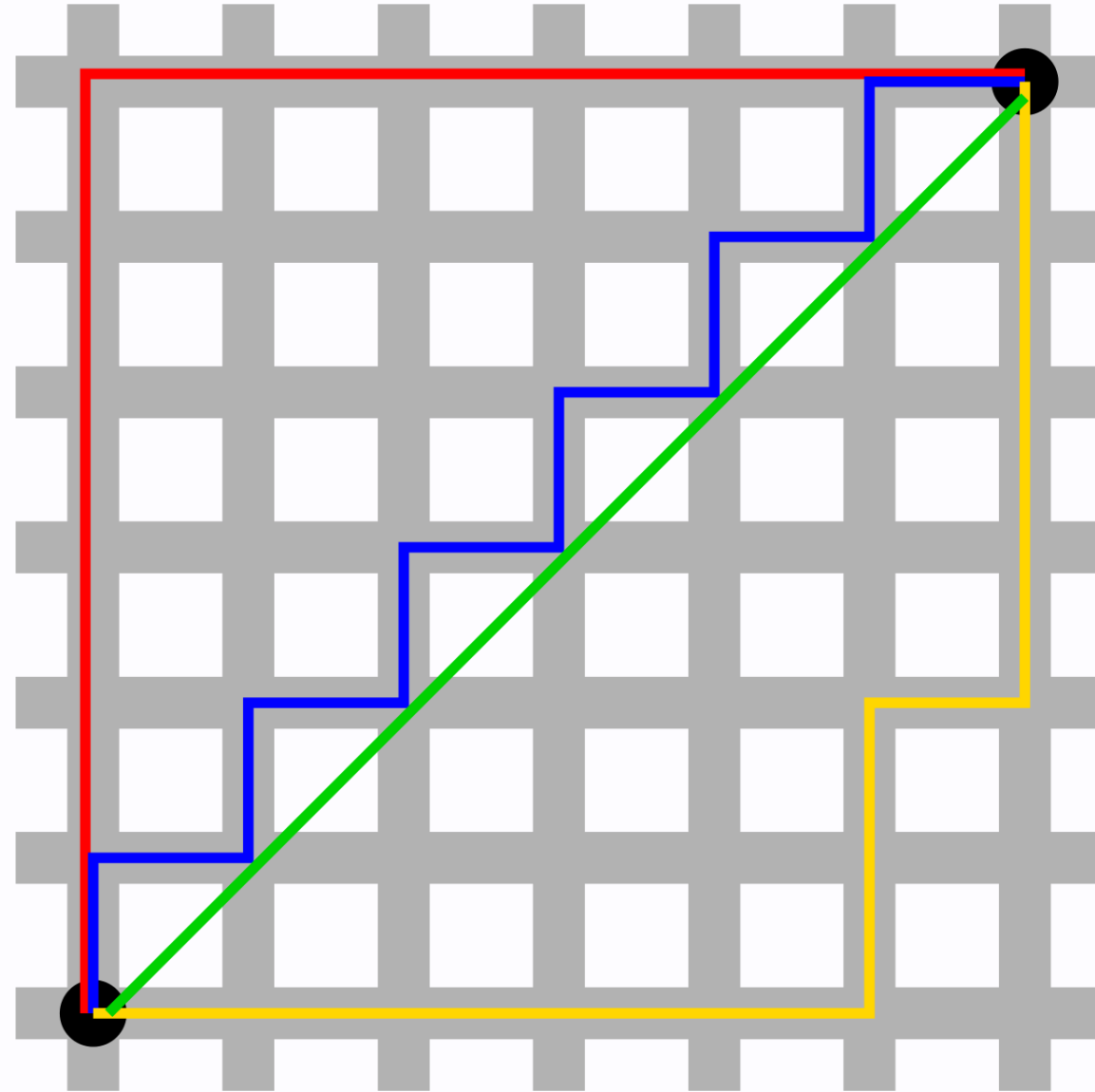
# L'entrainement d'un modèle

0. On initialise un modèle
1. On utilise le modèle pour faire une prédiction
2. On compare la prédiction à ce qu'on attend
3. On corrige le modèle
4. On recommence à partir de 1 jusqu'à être satisfait



# Définir un critère d'évaluation de l'erreur

- Distance L1, L2
- Précision, rappel, score F1
- Distance de Levenshtein
- ...



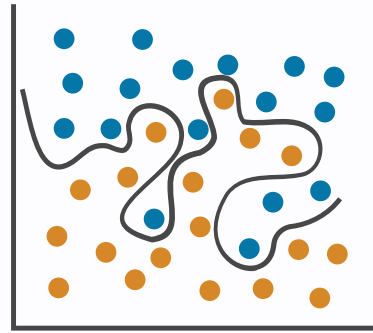


# Éviter le sous-apprentissage et le sur-apprentissage

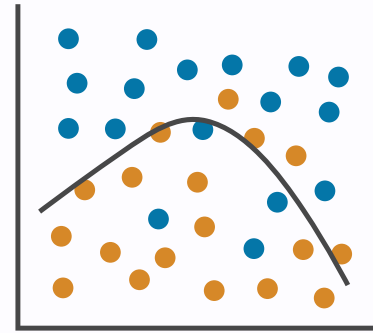
Un modèle de ML doit être capable d'**estimer** et de **généraliser**.

Classification

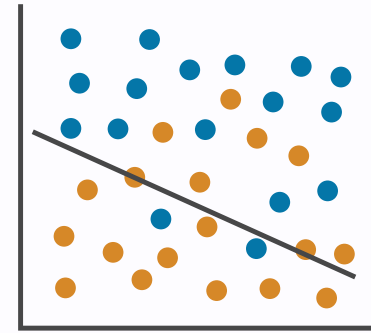
Overfitting



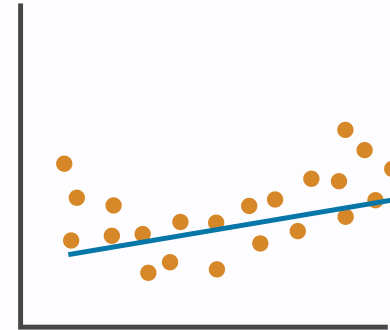
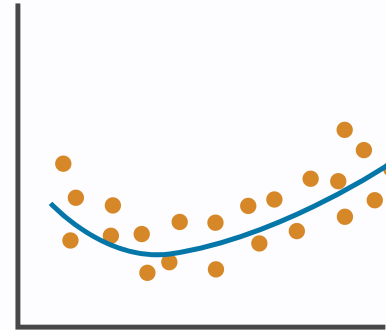
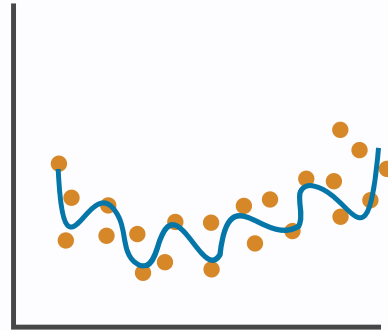
Right Fit



Underfitting

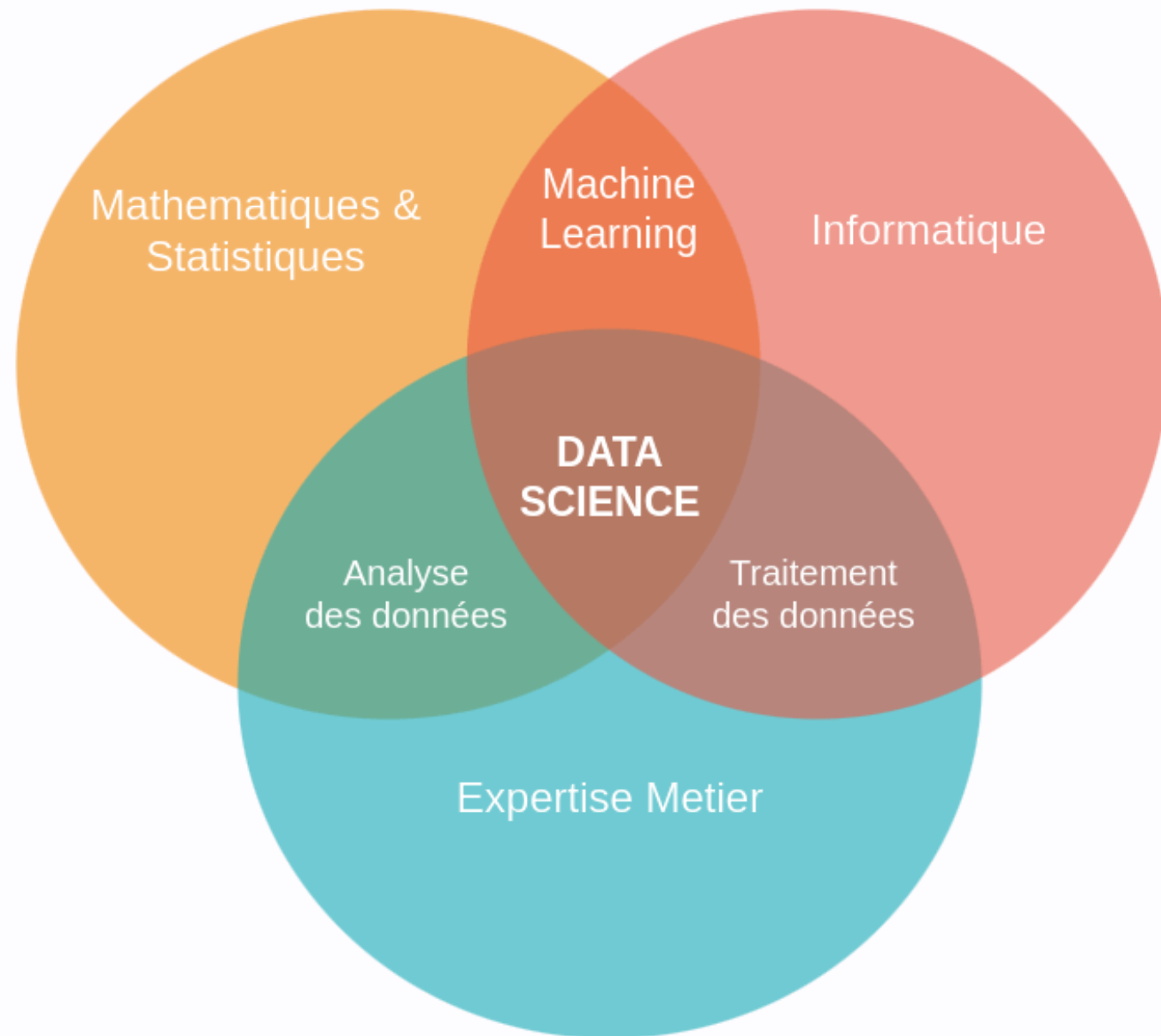


Regression



# **La mise en oeuvre**



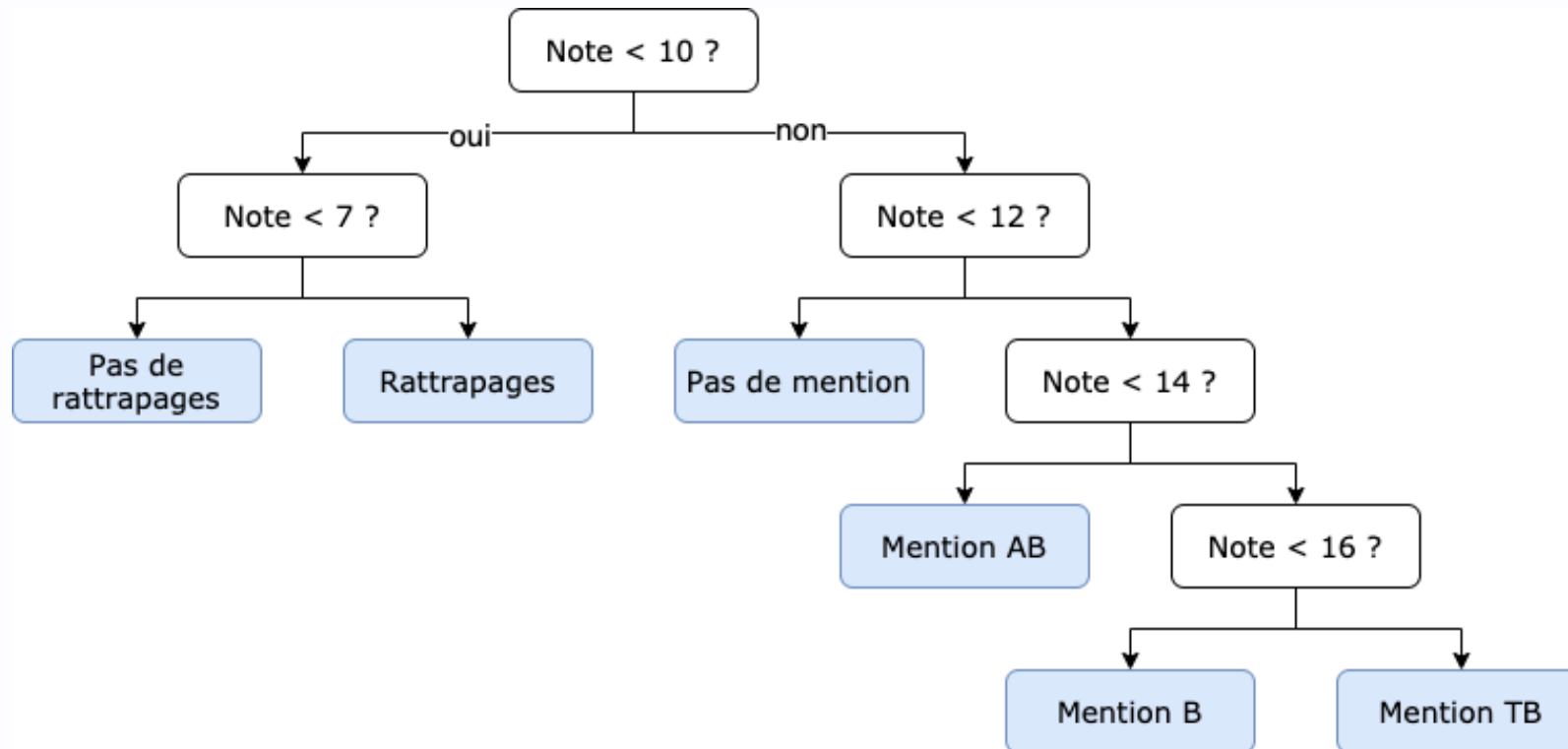


# Cadrer le problème

- identifier le type du problème
- poser des hypothèses sur les données
- choisir un critère d'évaluation
- choisir un modèle
- choisir un algorithme pour l'apprentissage
- choisir un algorithme pour l'inférence

# Quelques modèles

# Les arbres de décision

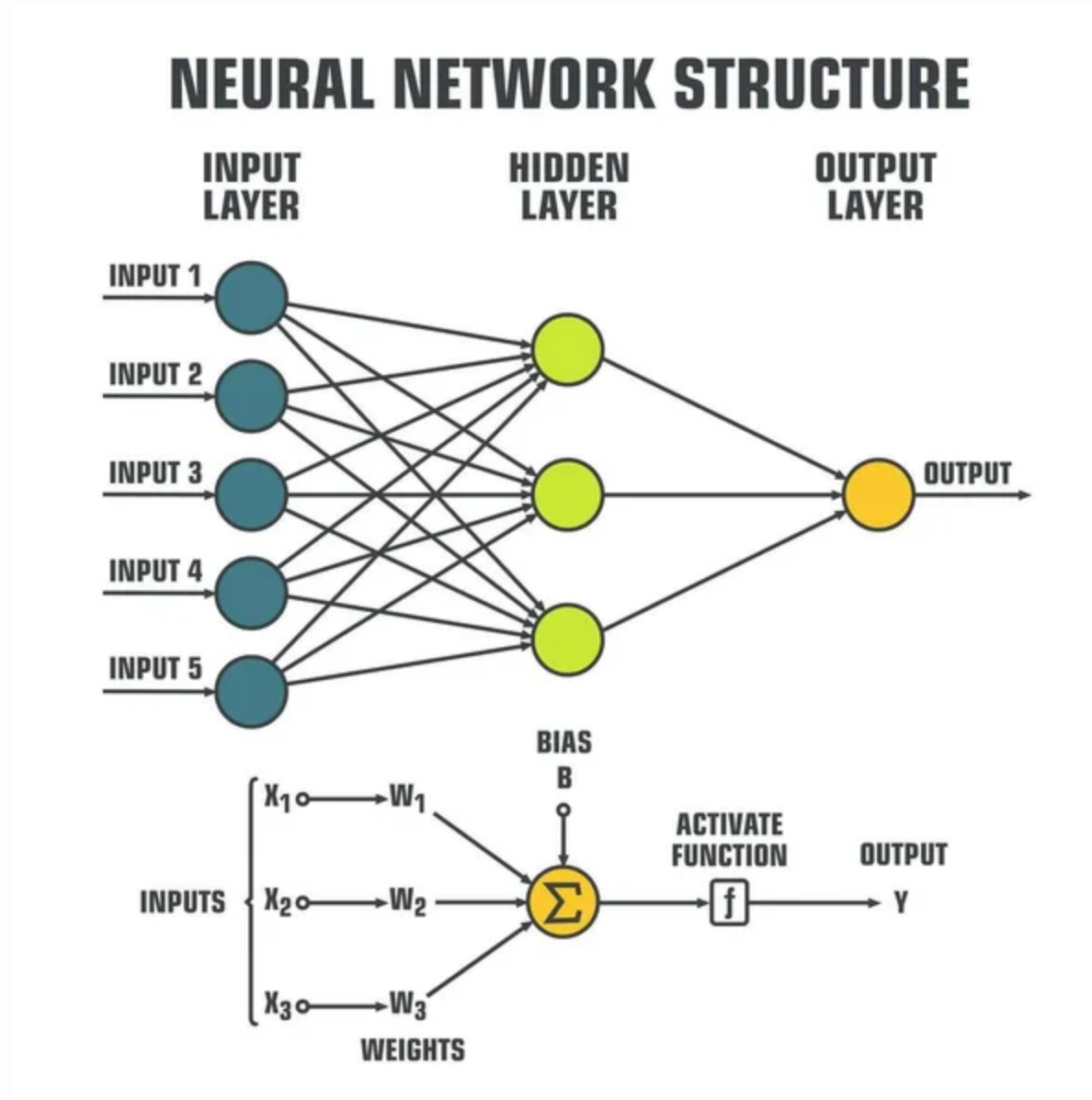




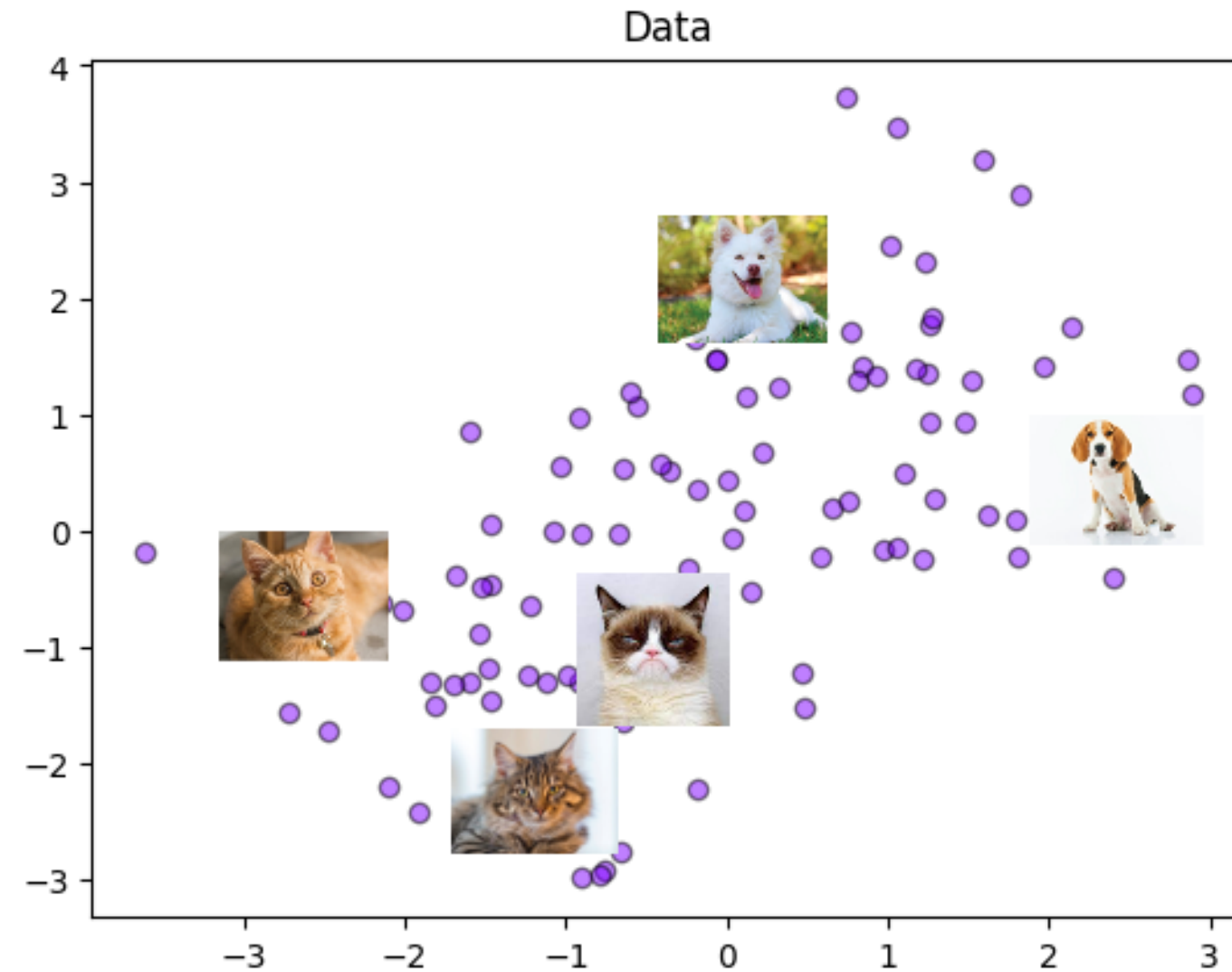
# **Les forêts aléatoires**

Un ensemble d'arbres de décision

# Les réseaux de neurones



# Les KNN



# Conclusion

Le Machine Learning vise à faire des prédictions ou prendre des décisions en se basant sur des données.

Il faut :

- identifier le type du problème
- comprendre les données

# Ressources pour aller plus loin :

- [https://fr.wikipedia.org/wiki/Apprentissage\\_automatique](https://fr.wikipedia.org/wiki/Apprentissage_automatique)
- [https://scikit-learn.org/stable/user\\_guide.html](https://scikit-learn.org/stable/user_guide.html)
- cours en accès libre :
  - ★ <https://www.kaggle.com/learn>
  - <https://www.coursera.org/specializations/machine-learning-introduction>
  - [https://openlearninglibrary.mit.edu/courses/course-v1:MITx+6.036+1T2019/courseware/Week1/intro\\_ml](https://openlearninglibrary.mit.edu/courses/course-v1:MITx+6.036+1T2019/courseware/Week1/intro_ml)