

SAE53

MACHINE LEARNING

Groupe: ER ROUASSE Ayoub, KONE Shyli, MOREL Robin et THIAW Alioune

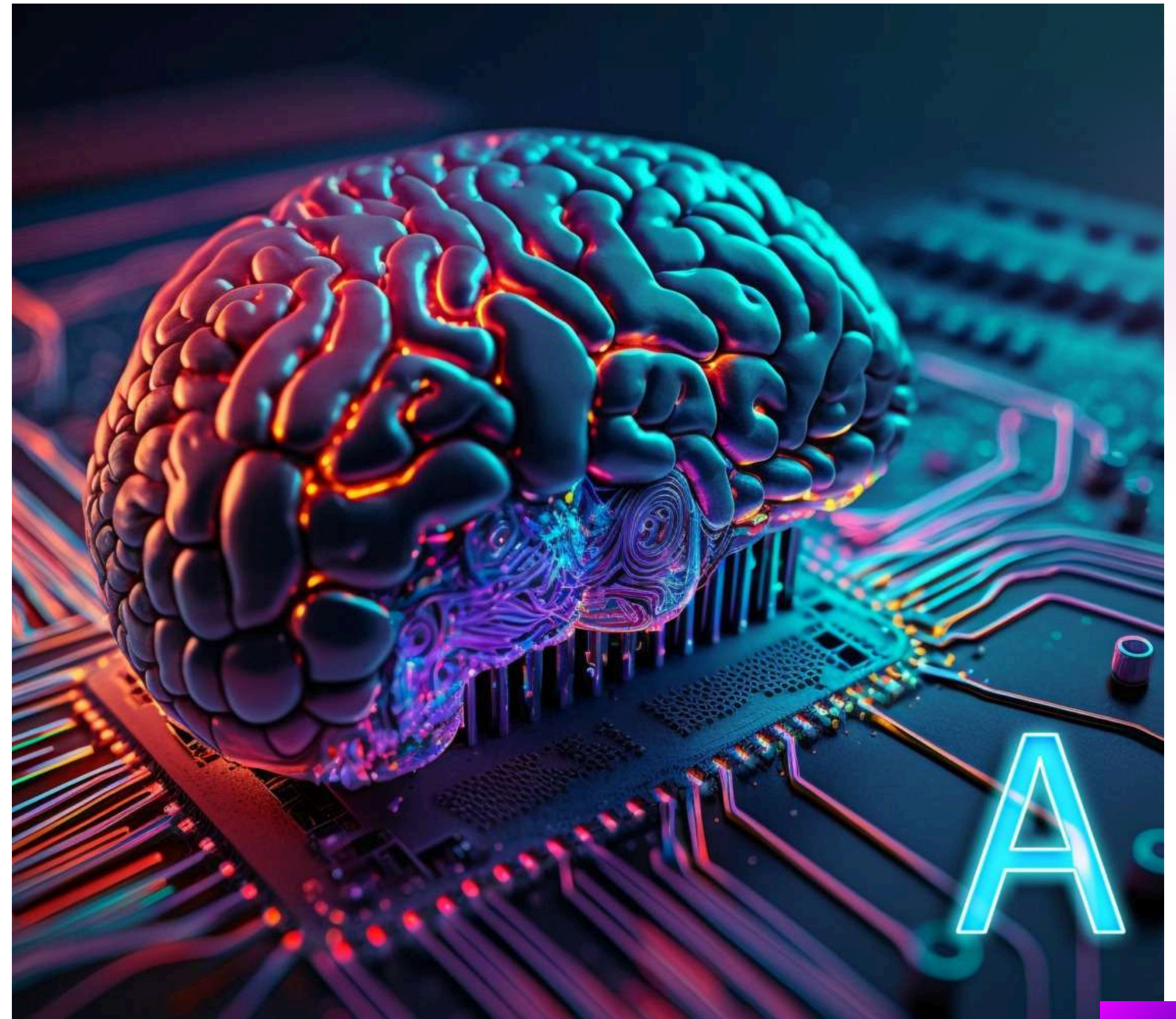


Présentation du projet



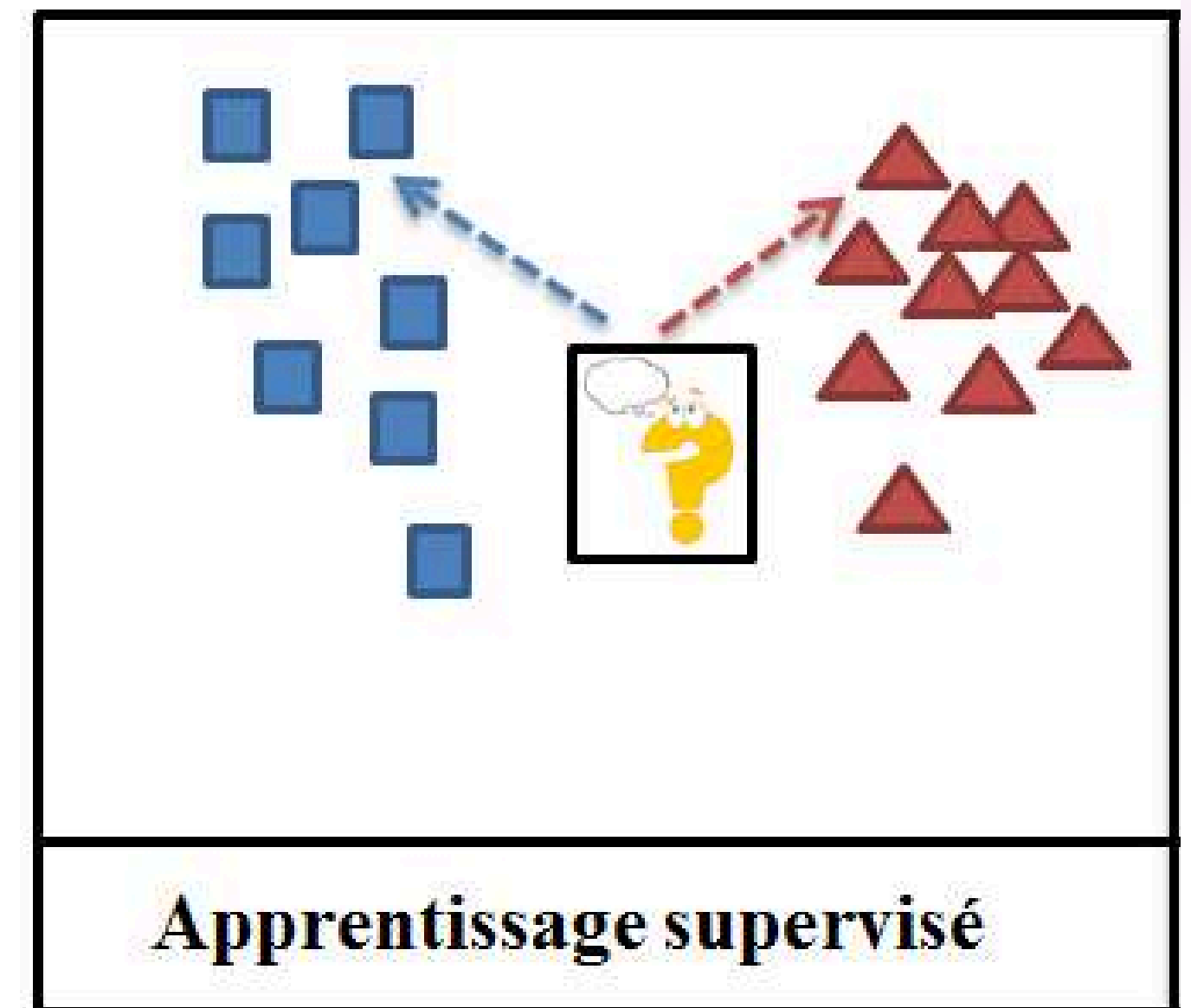
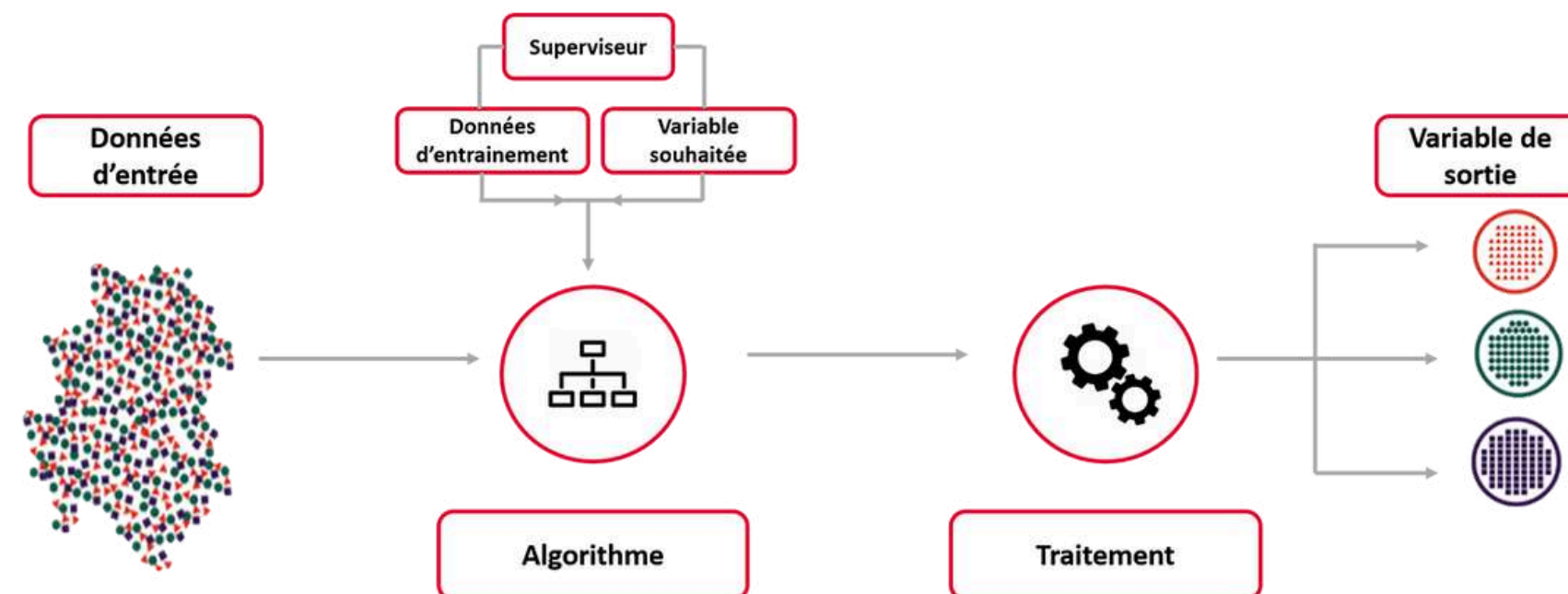
Sommaire

- **Apprentissage supervisé**
 - Kppv (k plus proches voisins)
 - MLP (Perceptron multicouche)
 - Discrimination fonctionnelle
 - SVM (Machines à vecteurs de support)
- **Apprentissage non supervisé**
 - Kmeans
 - DBScan
 - Carte de Kohonen





Apprentissage Supervisé





K-Plus Proche Voisin (KNN)

Principe

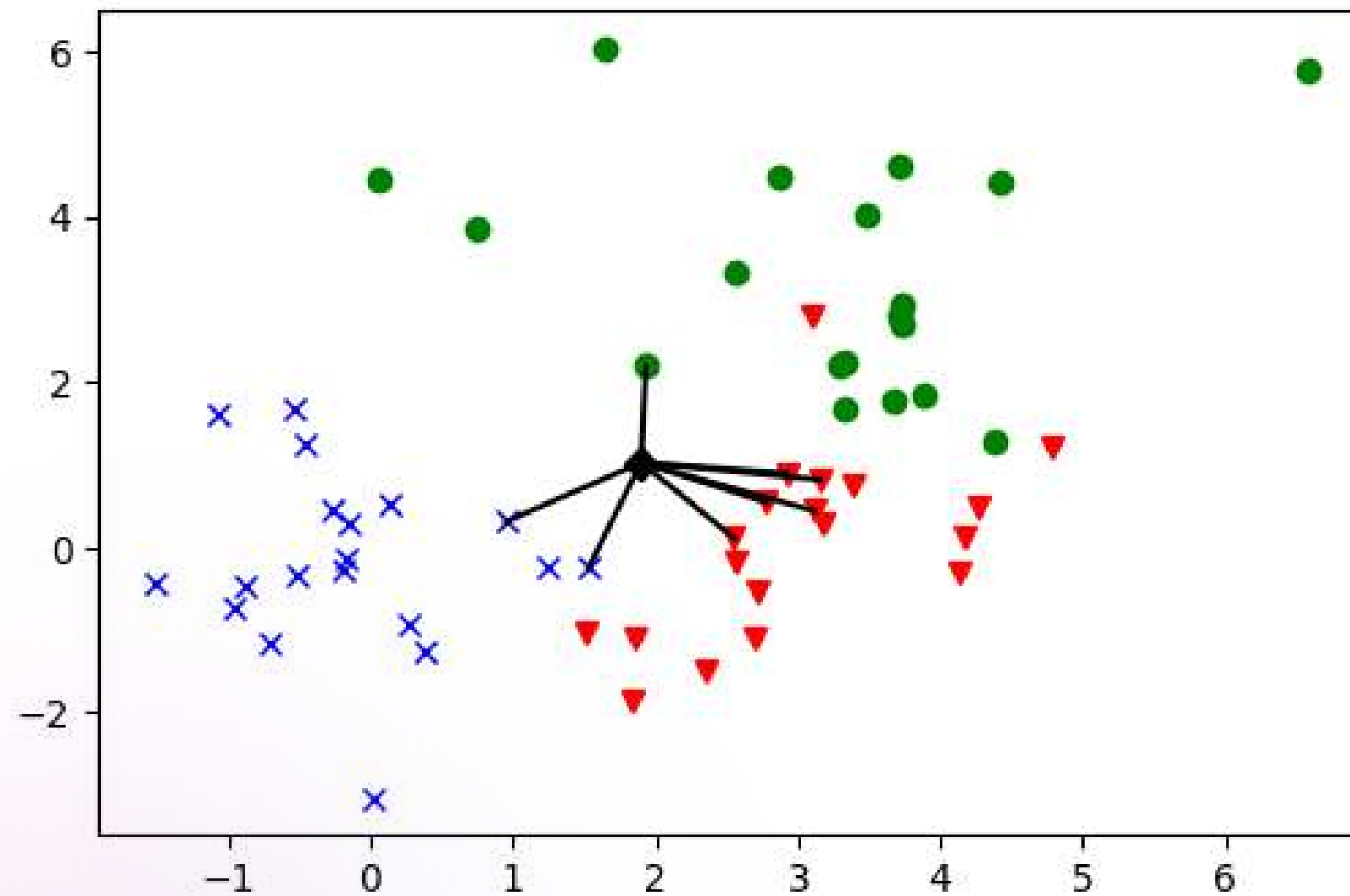
- Classe données sur la distance des k voisins les plus proches
- Plusieurs types de distances :
 - Euclidienne
 - Manhattan
 - Minkowski

Avantages

- Simple
- Efficace sur les petits échantillons

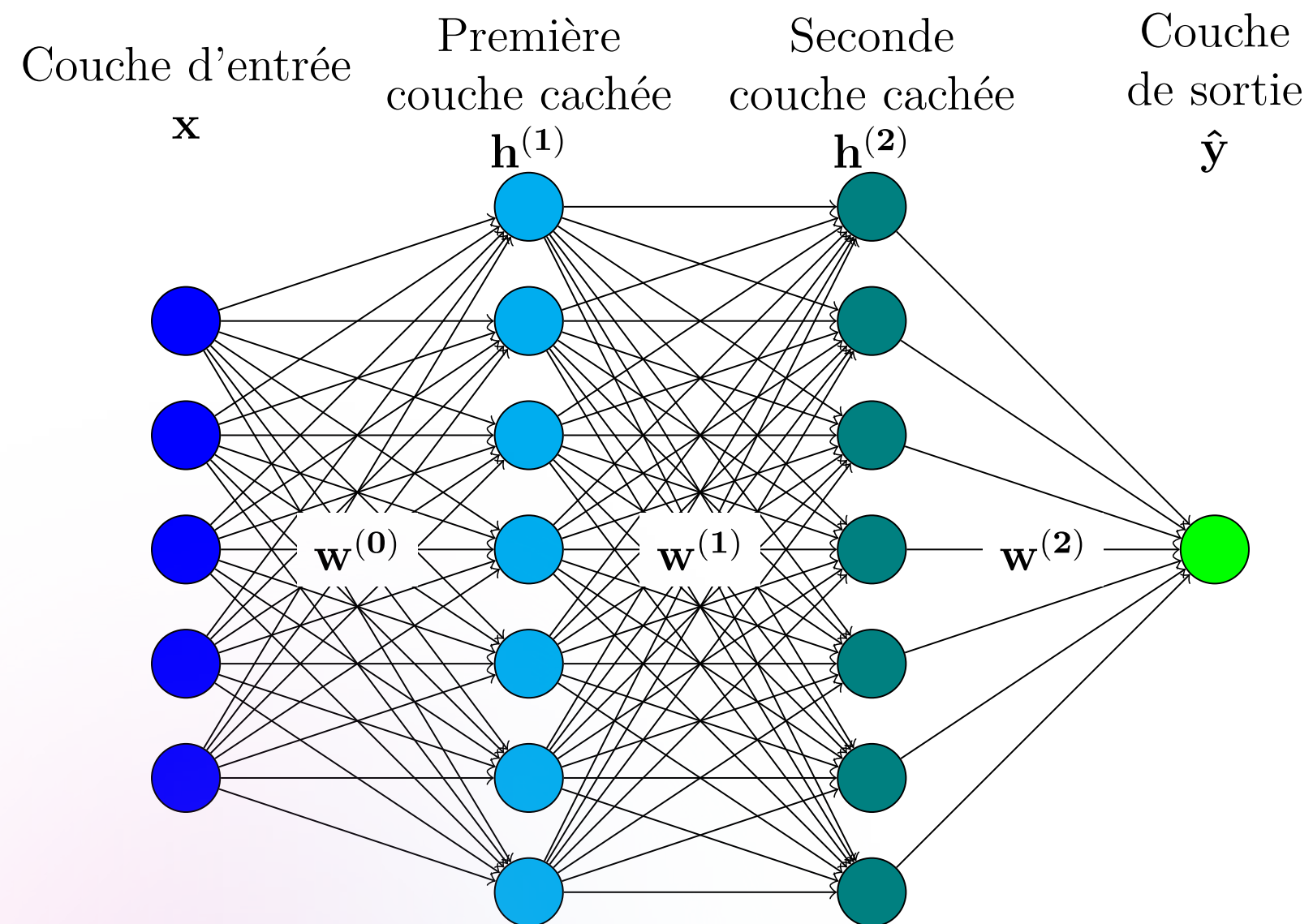
Inconvénients

- Problème de performance sur les grands échantillons





Perceptron Multicouche (MLP)



Principe

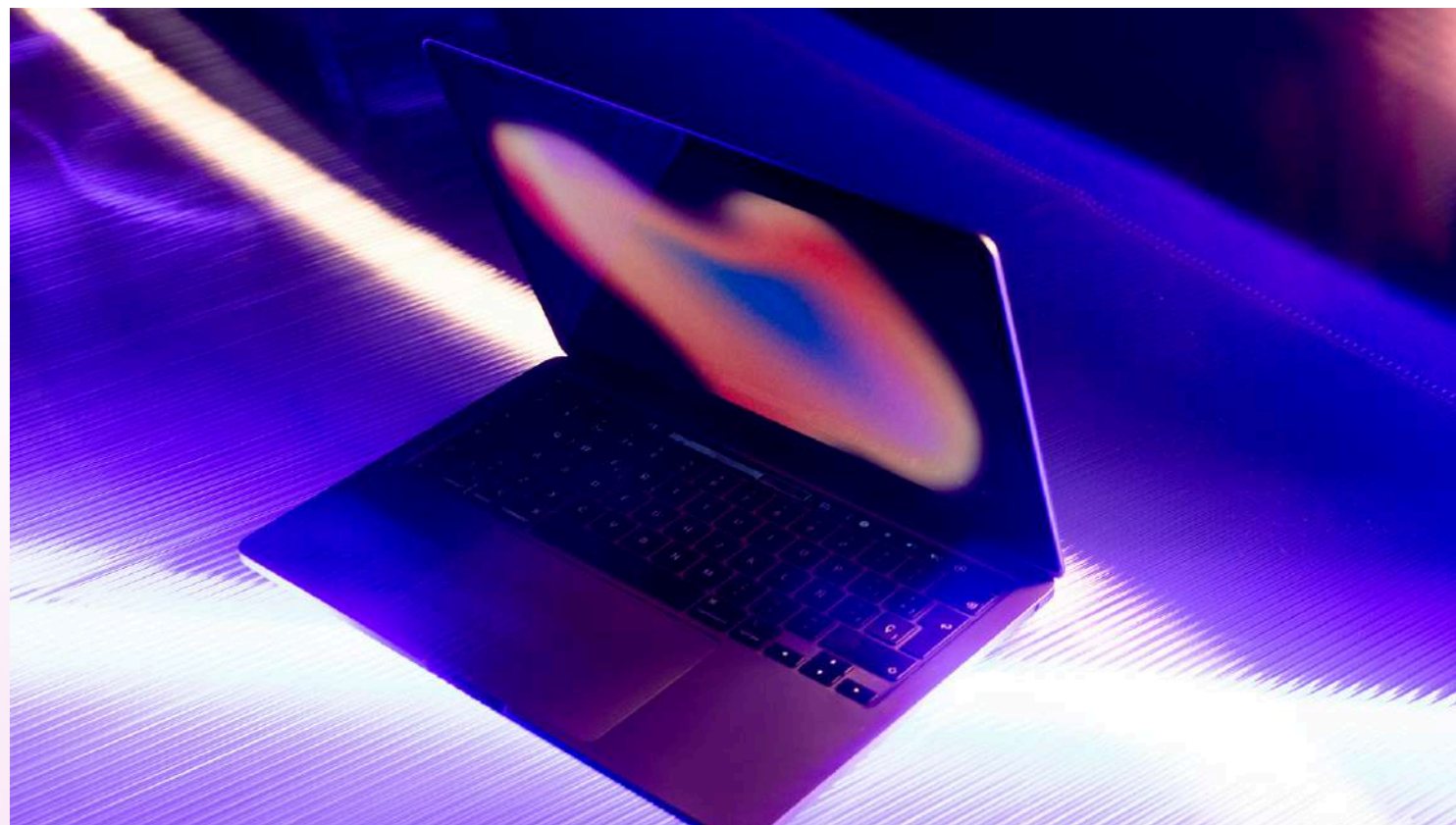
- Réseau de neurones artificiels sur plusieurs couches
- Transforme les données d'entrées
- Fonctions d'activation

Avantages

- Puissant pour modéliser des relations complexes

Inconvénients

- Nécessite beaucoup de données et de puissance de calcul



Discrimination Fonctionnelle

Principe

- Séparer des classes en trouvant une frontière optimale

Avantages

- Efficace sur des données bien séparées linéairement

Inconvénients

- Moins performant sur des structures complexes.



Machines à Vecteurs de Support (SVM)

Principe

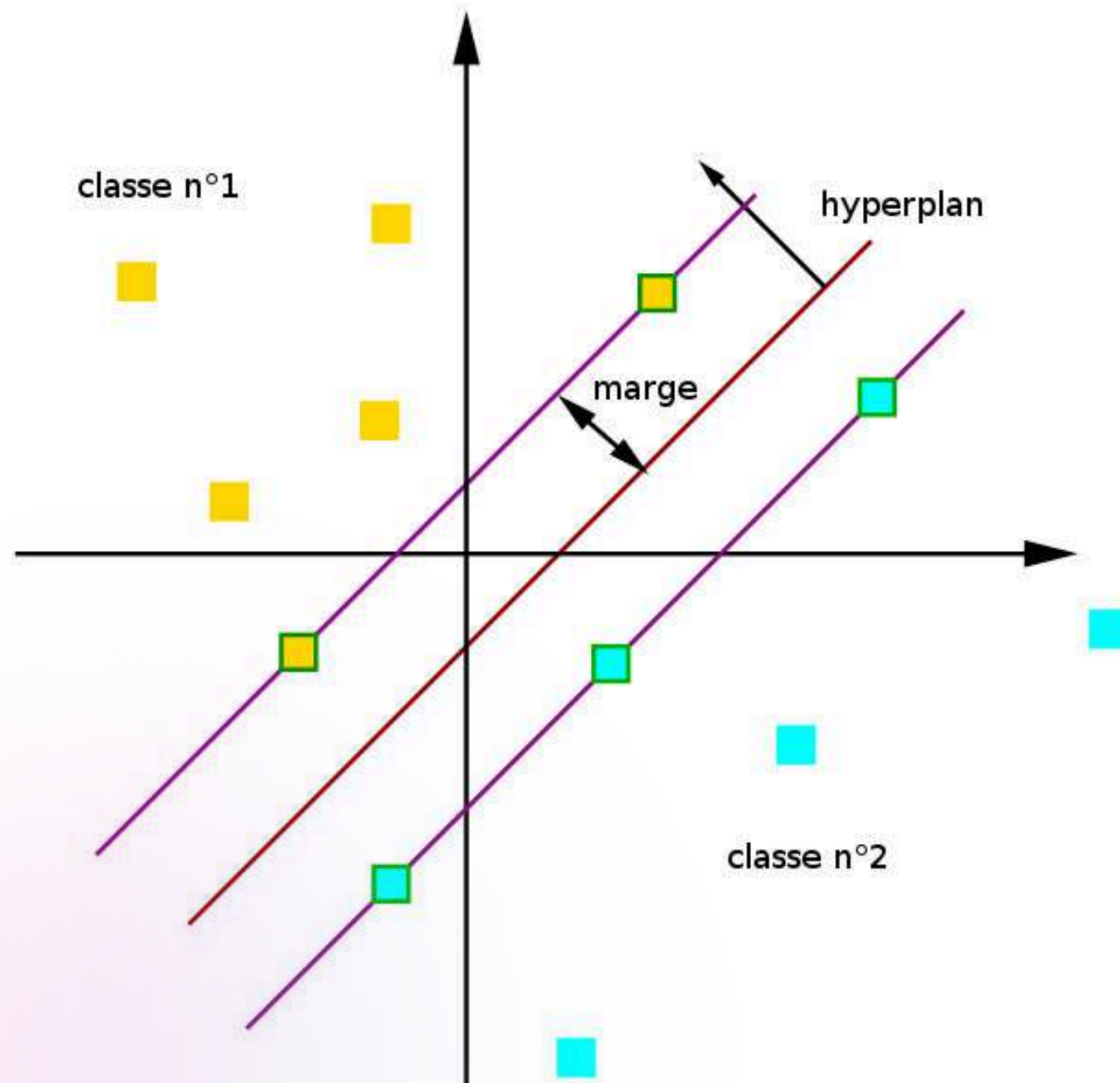
- Algorithme de classification trouvant une hyperplan optimal.
- Maximisation de la marge entre classes.
- Utilisation de noyaux (linéaire, gaussien, polynomial, etc.).

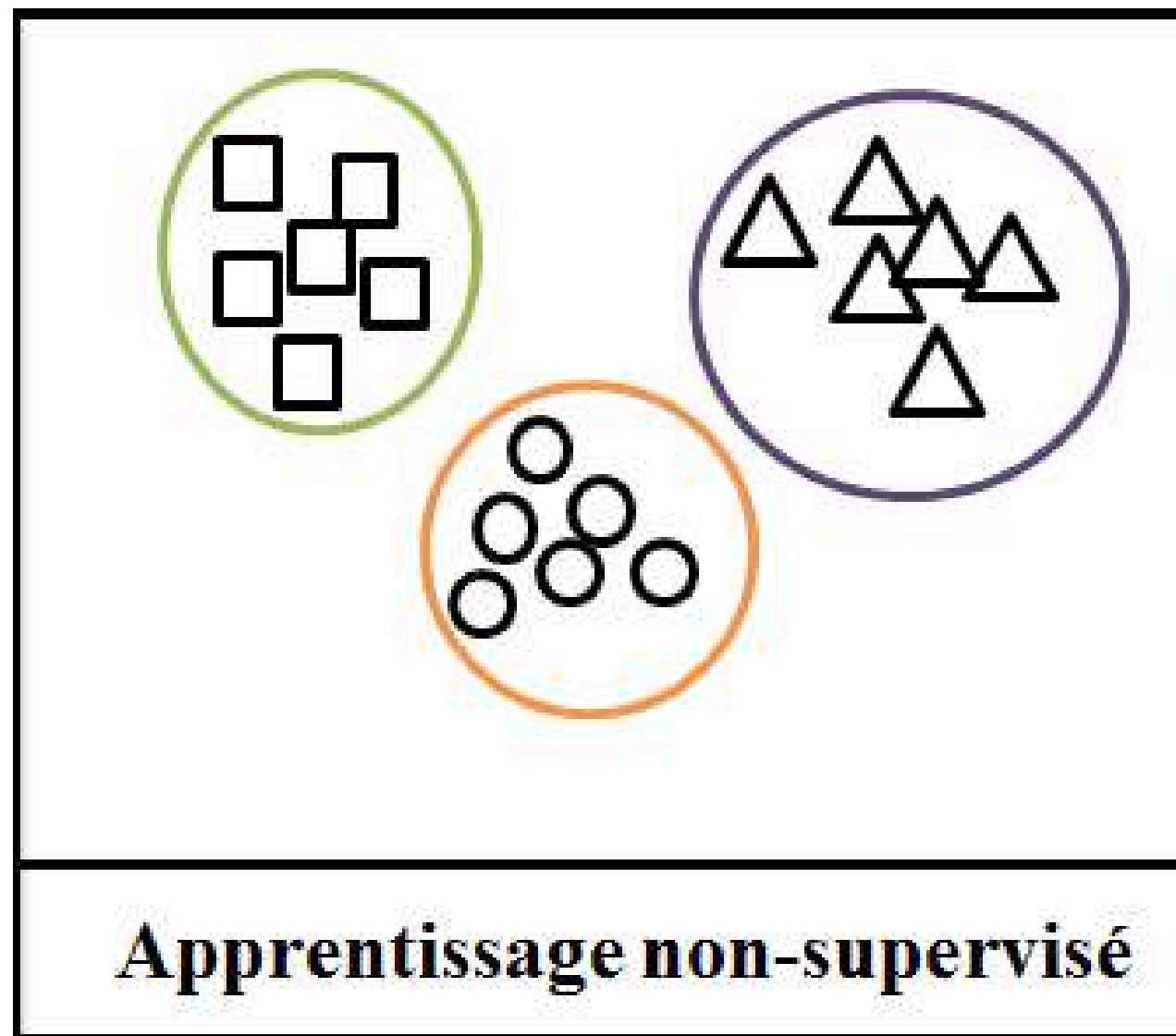
Avantages

- + Efficace pour petites dimensions

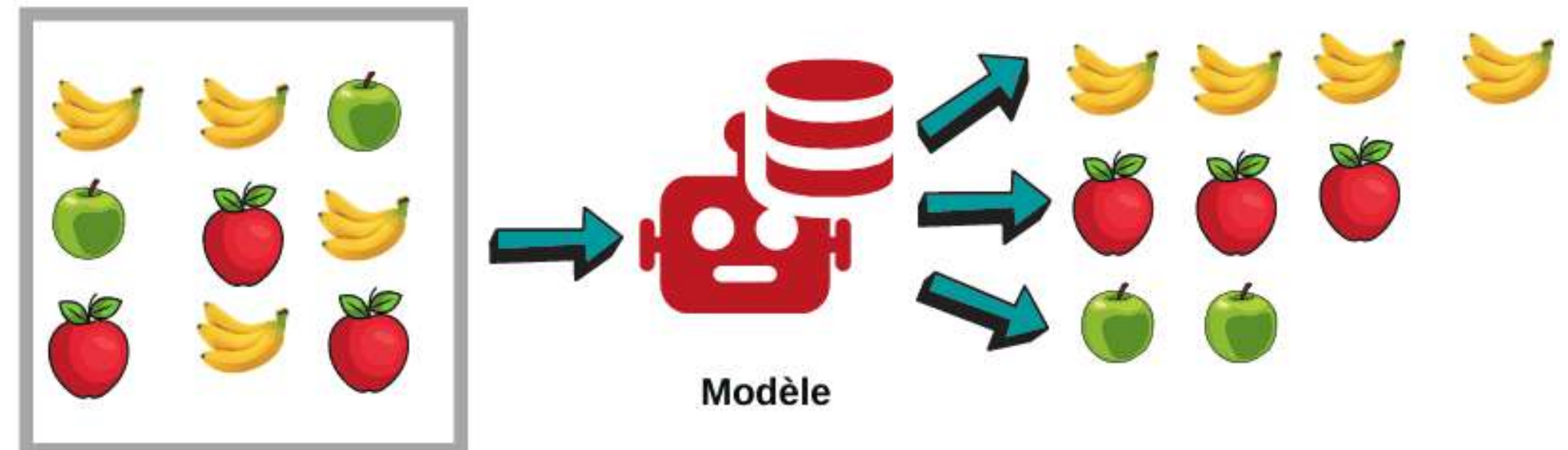
Inconvénients

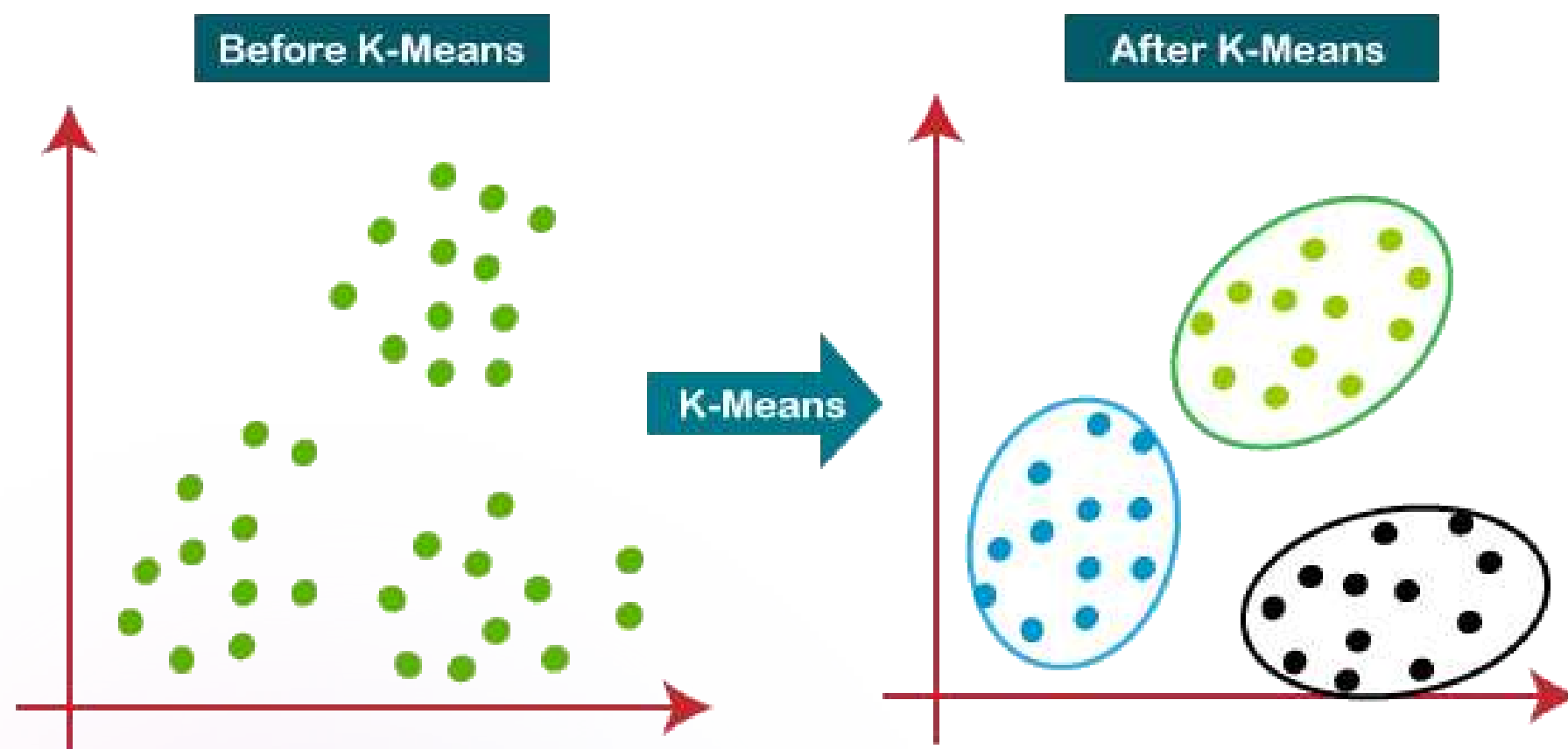
- - Performant sur grands jeux de données avec chevauchement





Apprentissage Non-Supervisé





K-means

Principe

- Algorithme de clustering basé sur la minimisation de la variance intra-classe.
- Sélection aléatoire de k centres.
- Attribution des points au centre le plus proche.
- Mise à jour des centres.

Avantages

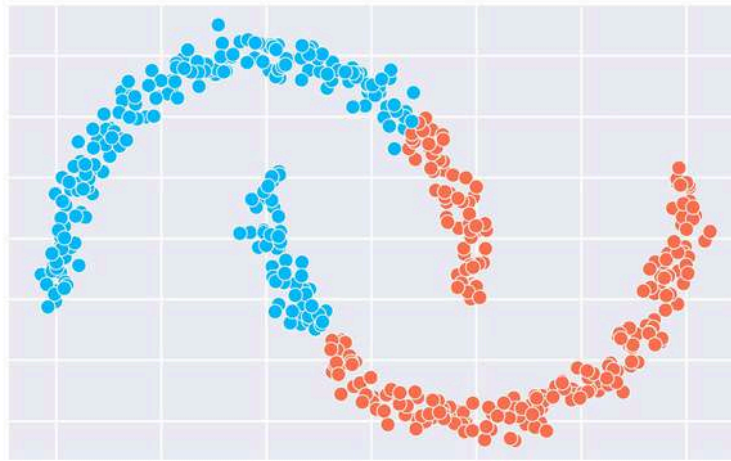
- Rapide
- Simple

Inconvénients

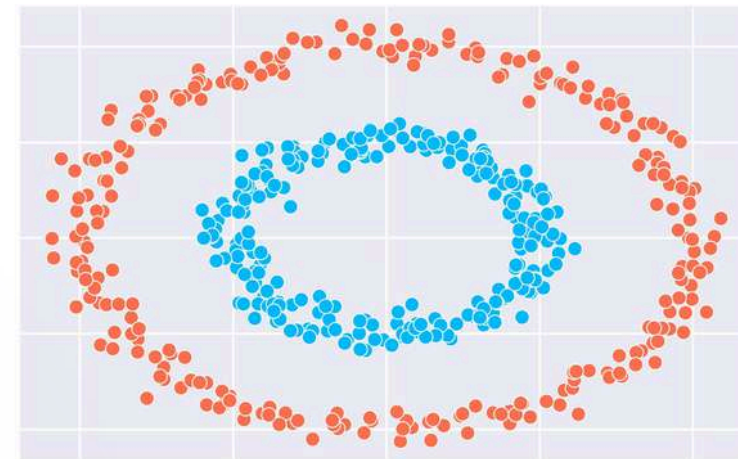
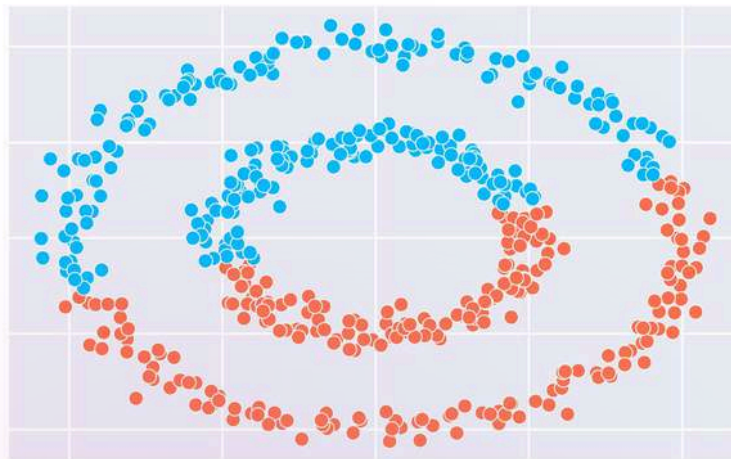
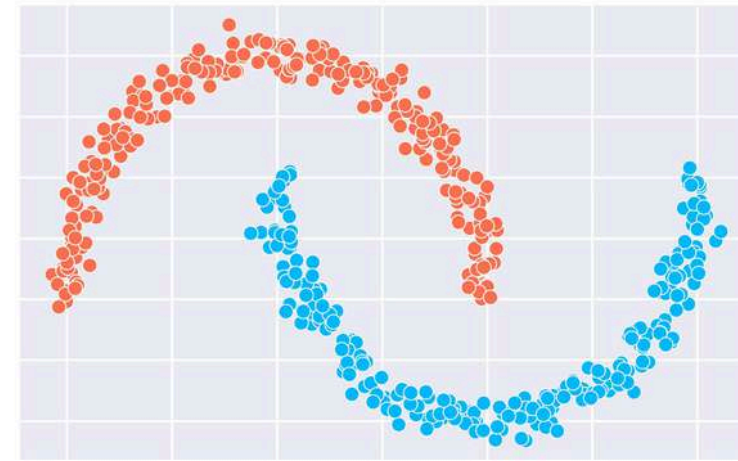
- K à choisir manuellement (délicat)
- Valeurs extrêmes influencent l'algo



KMeans



DBSCAN



DBScan

Principe

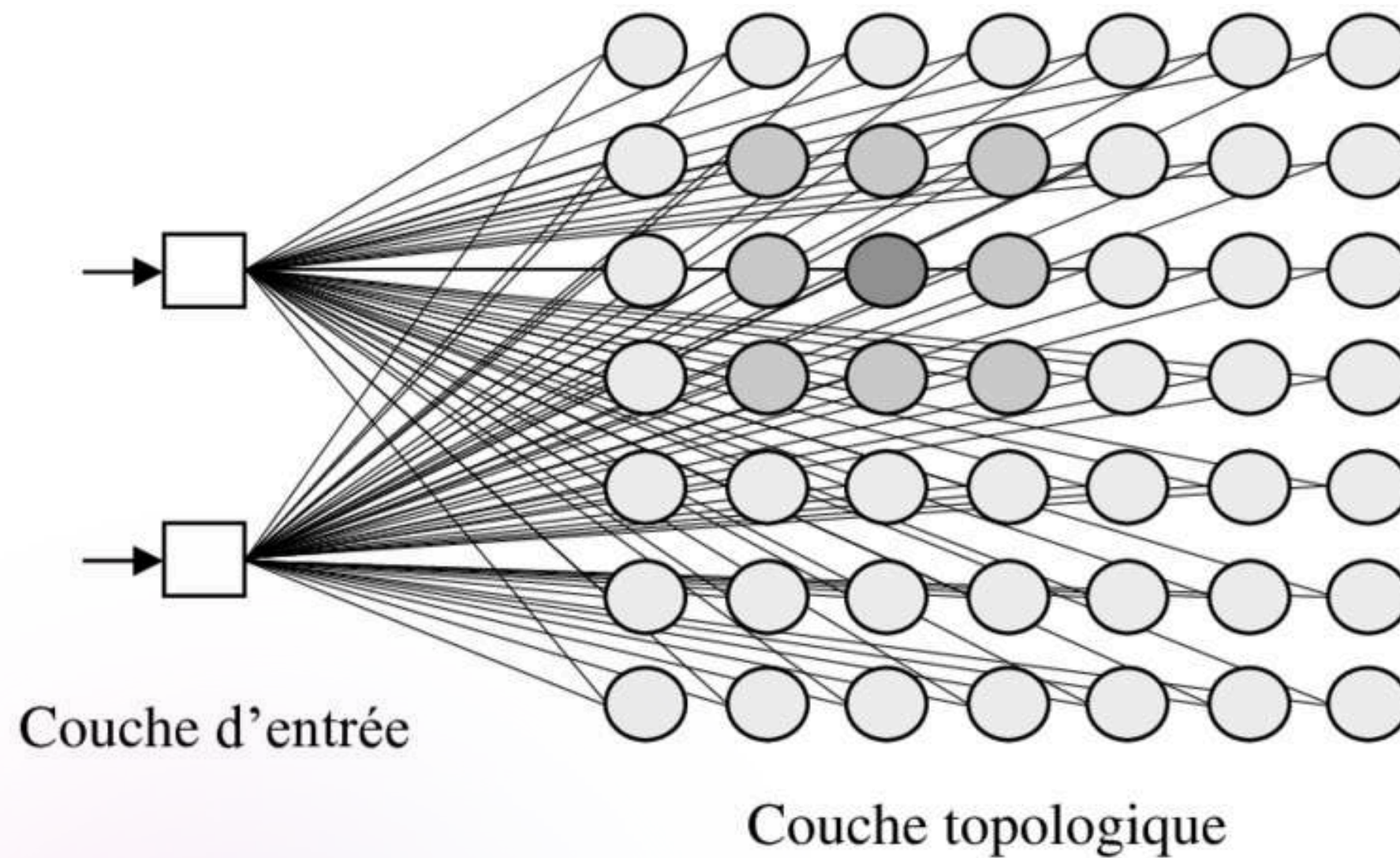
- Algorithme de clustering basé sur la densité.
- Définit des « noyaux » et regroupe les points densément connectés.
- Gère les valeurs extrêmes

Avantages

- efficace pour données bruitée
- efficace pour clusters de formes variées

Inconvénients

- Sensible aux paramètres (epsilon, minPts).



Carte de Kohonen

Principe

- Réseau de neurones non supervisé pour la classification.
- Projection des données en espace réduit.
- Auto-organisation des neurones selon la proximité des données.

Avantages

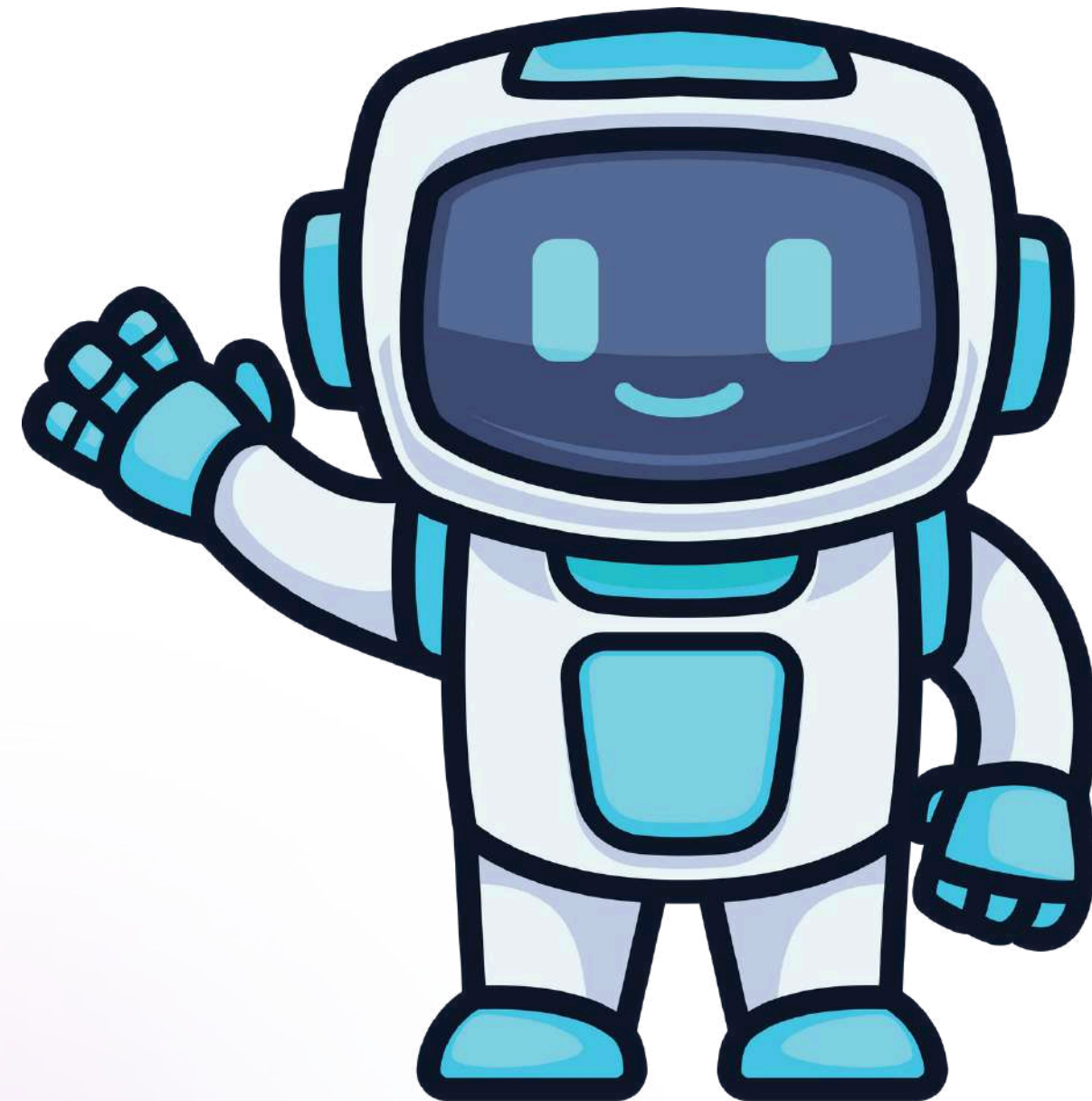
- Visualisation intuitive

Inconvénients

- Temps de calcul élevé



CONCLUSION



MERCI