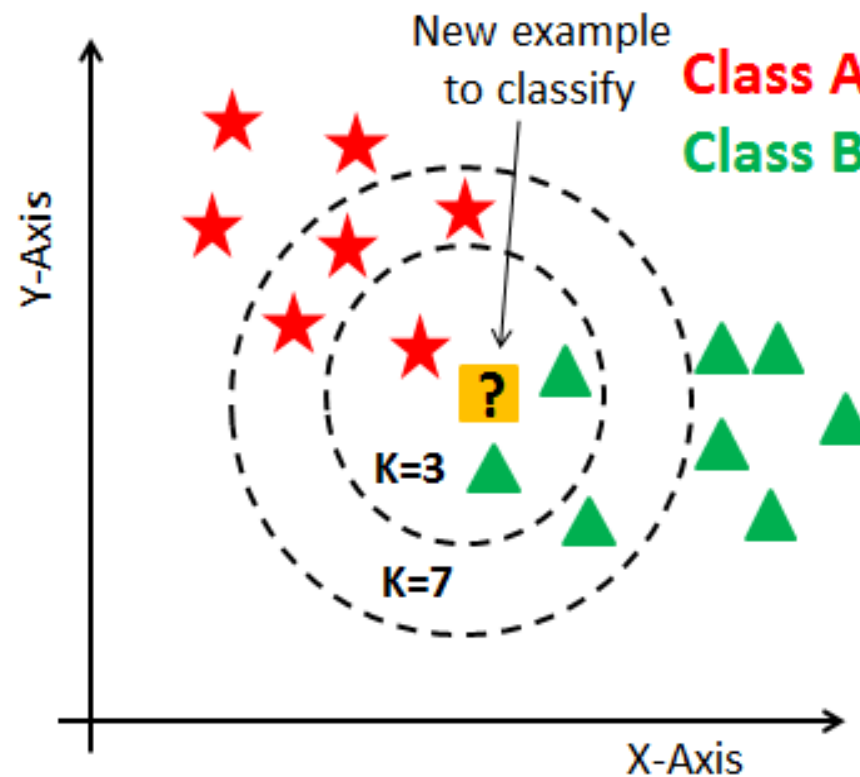


Algorithme des k plus proche voisins



Sommaire

- Introduction
- L'objectif de l'algorithme
- Le fonctionnement de celui-ci
- Qu'est-ce qu'un cluster ?
- Exemples où l'algorithme est utile

Introduction

L'algorithme des k plus proche voisins, ou knn (k Nearest Neighbors), est un algorithme d'apprentissage supervisé. L'apprentissage supervisé ou machine learning est le fait de donner des données à étiquetées(que l'on définit) à l'algorithme afin qui les assimile.

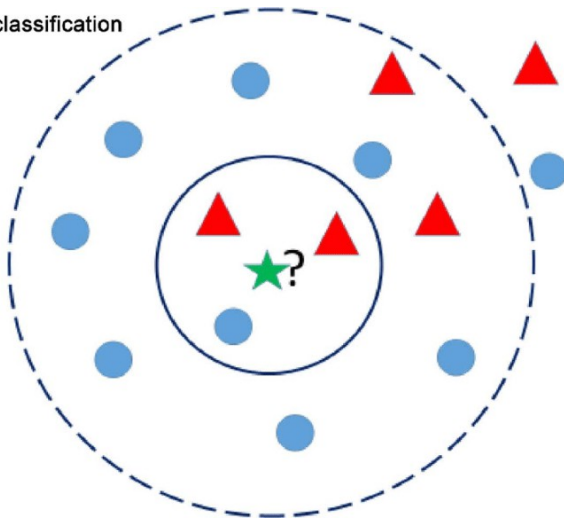
L'objectif de l'algorithme

L'algorithme des k plus proche voisins a pour principe « *dis moi qui sont tes voisins, je te dirais qui tu es* » . L'algorithme fonctionne grâce à des données et en fonction de ces données il va chercher les instances les plus proche de notre individu à étudier en particulier leur output pour « prédire » le output de notre individu.

Le fonctionnement de celui-ci

- Pour expliquer le fonctionnement de l'algorithme, rien de mieux qu'un exemple :

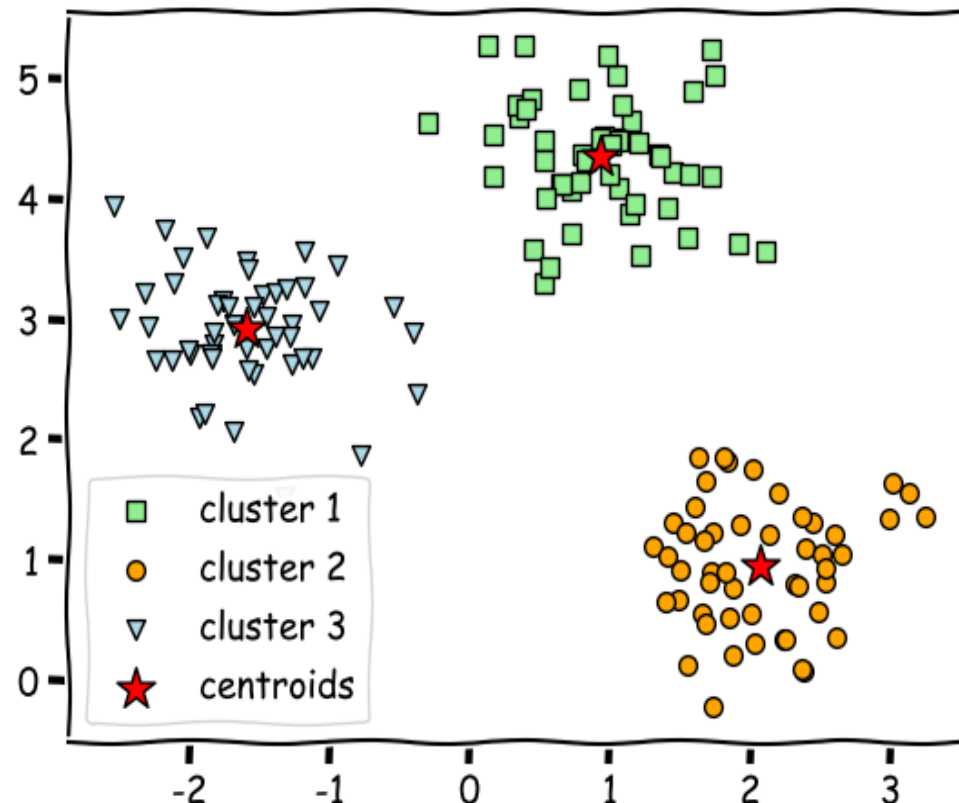
Example of KNN classification



Nous avons un individu, dont on veut déterminer la classe, ici ça sera l'étoile, pour ce faire on prend un nombre k de voisins. Si on prend $k=3$, on remarque que les 3 plus proches voisins sont 2 triangles et un cercle. L'individu étudié appartient donc à la classe des triangles. Cependant si l'on prend un nombre de voisins supérieur, $k=11$ on remarque que les 11 plus proches voisins sont 7 cercles et 4 triangles. On en déduit donc la classe de l'individu, la classe des cercles.

Qu'est-ce qu'un cluster ?

Un cluster un regroupement d'objet ayant des caractéristique équivalente.



Sur l'image,
nous pouvons
voir 3 cluster

Exemples où l'algorithme est utile

- L'algorithme des k plus proches voisins peut être utile dans plusieurs domaines tels que :
 - La reconnaissance faciale (ex : Face Id)
 - La filtration des mails(ex : les spams)
 - L'épidémiologie (ex : avec le coronavirus)
 - Le marketing avec le ciblage de personne potentiel (target)

Merci de votre attention