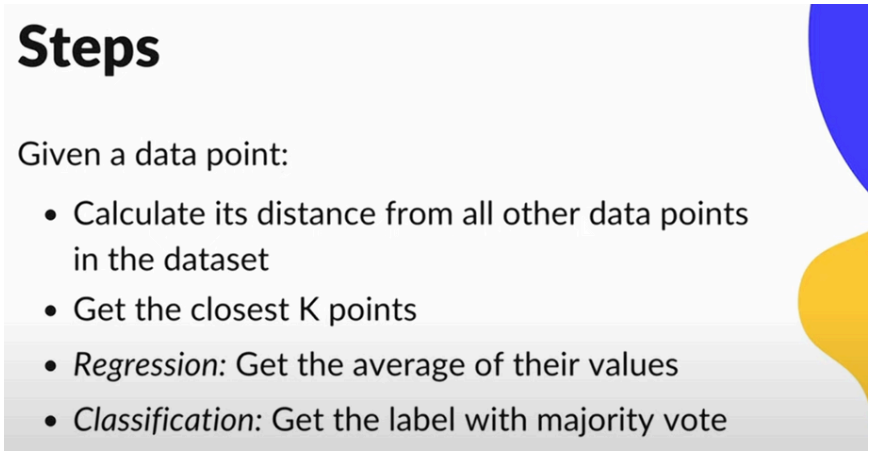
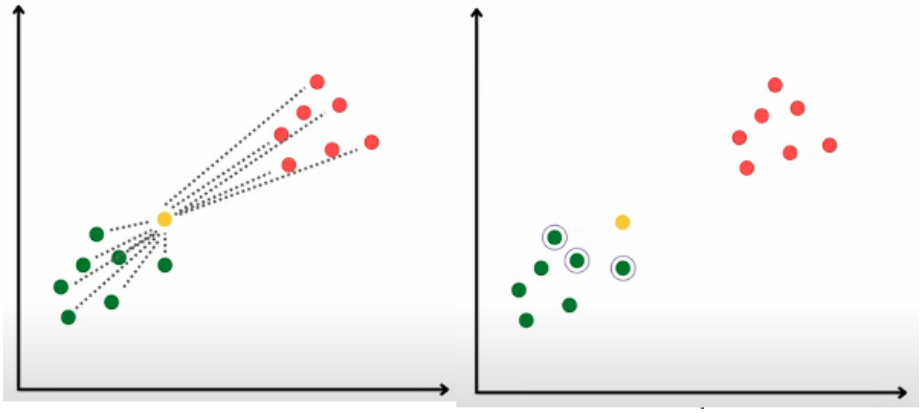
# KNN

Salman Hanif - 13523056

## Cara Kerja Algoritma

Algoritma K-Nearest Neighbors (KNN) bekerja dengan cara memprediksi kelas sebuah titik data baru berdasarkan mayoritas kelas dari tetangga terdekatnya dalam dataset pelatihan.



Pada Algoritma class KNN yang dibuat :

1. Method fit, ini digunakan hanya untuk menyimpan X_train dan y_train yang dimasukkan sebagai argumen
2. Method predict, saat digunakan akan memanggil method subpredict untuk setiap titik data sementara predict lah yang menampung hasil dari subPredict. SubPredict melakukan :

- Menghitung semua jarak, untuk satu titik data x yang akan diprediksi dengan metode [erhitungan jarak yang dipilih : euclidean etc
- Mencari tetangga terdekatnya sejumlah K sesuai input

- Dari sejumlah K tetangganya, dicari modus dari beliau-beliau ini dan itulah yang menjadi nilai y_predict nya

## Hasil Evaluasi model dari hasil scratch dan dari library. Jelaskan perbedaan

Dalam evaluasi, digunakan metode Hold-out dan K-Fold pada library scikit-learn dan KNN buatan saya.

Saya melakukan 2 pengujian, dengan method euclidean dan nilai K ganjil yaitu 5, maka hasil yang diberikan sama persis dalam confussion matrix. semua elemen (TP, FN, FP, TN) tidak ada yang beda satu pun.

Sedangkan pada pengujian dengan nilai K genap yaitu 4 dengan jarak manhattan, terdapat sedikit perbedaan di mana model saya performanya di bawah scikit-learn. Hal ini saya yakini pada klasifikasi binary yang hanya 1 dan 0, sangat rawan terjadi keseimbangan antara 1 dan 0, misal 4 terdekatnya [1,1,0,0], algoritma saya langsung mengambil urutan terdepan di dictionary. Sedangkan KNN scikit-learn mungkin memiliki logika yang berbeda.

## Improvement yang bisa dilakukan

Improvement yang bisa saya lakukan adalah mengembangkan algoritma pengambilan keputusan ketika hasil dalam K jumlahnya seimbang untuk performa model agar lebih baik lagi.