



K-NEAREST NEIGHBOR

KNN

Alunos :

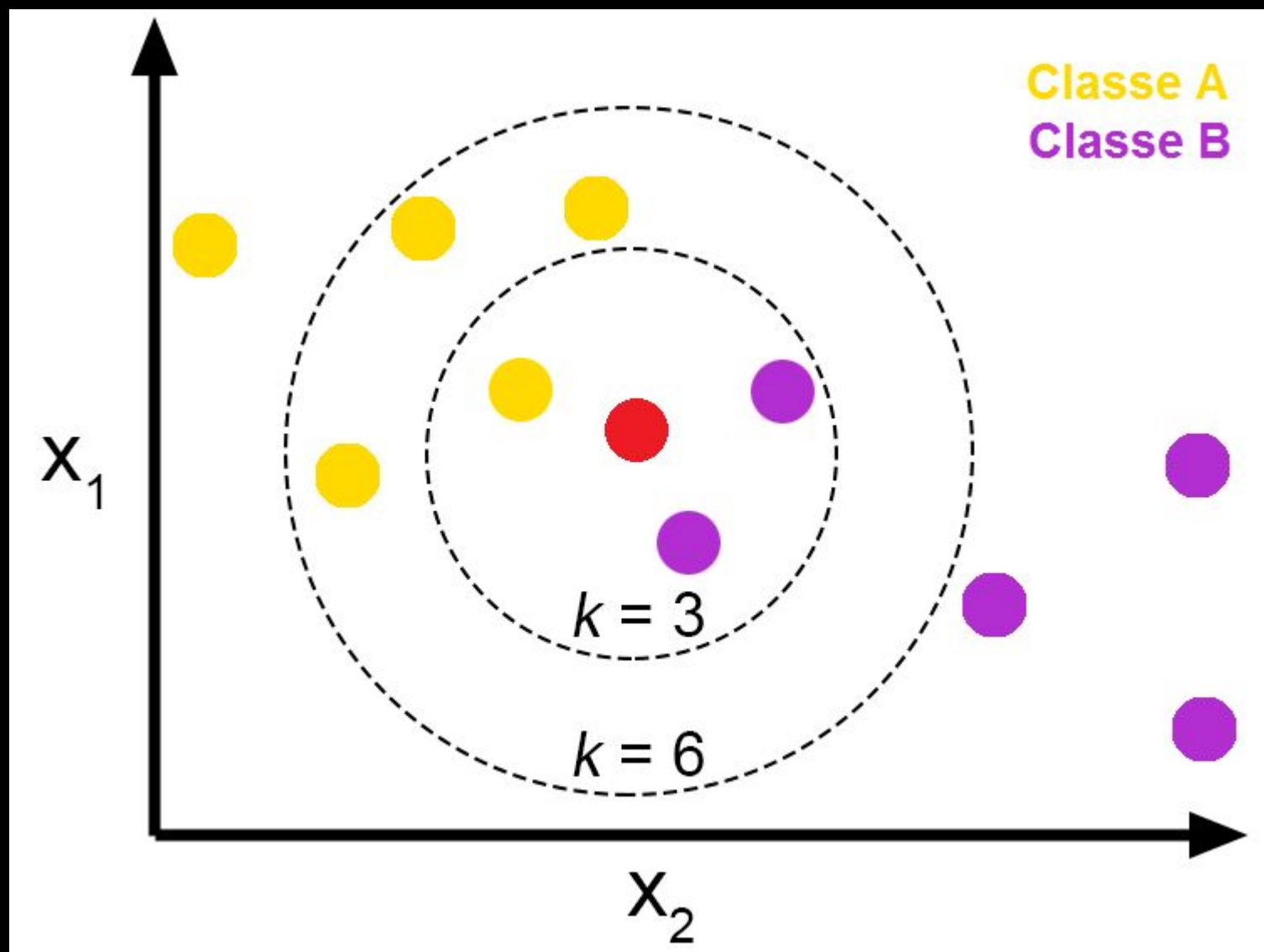
- ▶ **João Gondim**
- ▶ **Yndyra Kavendyche de Araujo
Leite Pinheiro**

Conceito

- ▶ **Aprendizado supervisionado**
- ▶ **Classificação simples**
- ▶ **Proximidade no espaço com a base de treinamento**

O que é necessário ?

- ▶ Base de treinamento
- ▶ Métrica para cálculo de distância
- ▶ Valor de K



Cálculo da distância entre 2 pontos

Distance functions

Euclidean

$$\sqrt{\sum_{i=1}^k (x_i - y_i)^2}$$

Manhattan

$$\sum_{i=1}^k |x_i - y_i|$$

Minkowski

$$\left(\sum_{i=1}^k (|x_i - y_i|)^q \right)^{1/q}$$

Vantagens

- ▶ **Técnica simples , facilmente implementada**
- ▶ **Flexível**
- ▶ **Ótimos resultados**

Desvantagens

- ▶ Base de treinamento muito grande
- ▶ Ruído ou características irrelevantes

Código para classificação da localização de proteínas– Ecoli Dataset

