



# K-Nearest Neighbors - Exemplo de Uso

☰ Ciclo	Ciclo 02: Aprendizado supervisionado - Classificação
# Aula	13
🕒 Created	@January 14, 2023 5:53 PM
☑ Done	☑
☑ Ready	☑

## Objetivo da Aula:

- ☐ Exemplos de uso
- ☐ Vantagens e Desvantagens
- ☐ Resumo
- ☐ Próxima aula

## Conteúdo:

### ▼ 1. Exemplos de uso

O KNN consegue encontrar um rótulo a partir de seus K vizinhos mais próximos.

#### ▼ 1.1 Sistema de recomendação

**Livros que você pode gostar?**

Dado um filme, encontre os  $k$  filmes mais similares.

## ▼ 1.2 Classificação de notícias

Dado uma nova notícia, em qual categoria ela mais se encaixa?

## ▼ 1.3 Agrupamento de clientes

Dado uma nova cliente, qual o tipo de outras clientes parecidas com ela?

## ▼ 1.4 Classificação de imagem

Dado uma nova imagem sem rótulo e grupos de imagens com fotos de pandas, ursos, cachorros, cavalos e etc, em qual grupo a nova imagem pode ser classificada?

## ▼ 1.5 Sistema de Busca

Dado uma nova consulta, quais as notícias que podem ser retornadas como resultado?

# ▼ 2. Vantagens e Desvantagens

## ▼ 2.1 Vantagens

1. O KNN é fácil de entender e simples de explicar o seu funcionamento.
2. Não há a necessidade de construir um modelo, ajustar vários parâmetros ou assumir premissas adicionais.
3. O KNN é versátil. Ele pode ser usado para resolver problemas de classificação, regressão e busca.

## ▼ 2.2 Desvantagens

- ▼ 1. O KNN se tornar significativamente devagar quanto mais o número de exemplos e/ou preditores aumenta.

Com o aumento da dimensionalidade dos dados através das colunas, o cálculo da distância fica mais demorado.

▼ 2. Em alta dimensionalidade, a distância entre os pontos podem ser distorcidas.

Os mesmos pontos podem se aproximar ou se distanciar dependendo da quantidade de colunas na tabela e isso pode alterar os rótulos. O que foi classificado como A, pode ser reclassificado como B.

### ▼ 3. Resumo

1. A principal vantagem do KNN é a versatilidade do algoritmo. Ele pode ser usado para solucionar problemas de classificação, regressão e busca.
2. A principal desvantagem é a perda de performance em velocidade de classificação a medida que a quantidade de dados ou o número de características aumentam.

### ▼ 4. Próxima aula

Métricas de avaliação I: Matriz de confusão e acurácia