

Aula Prática KNN

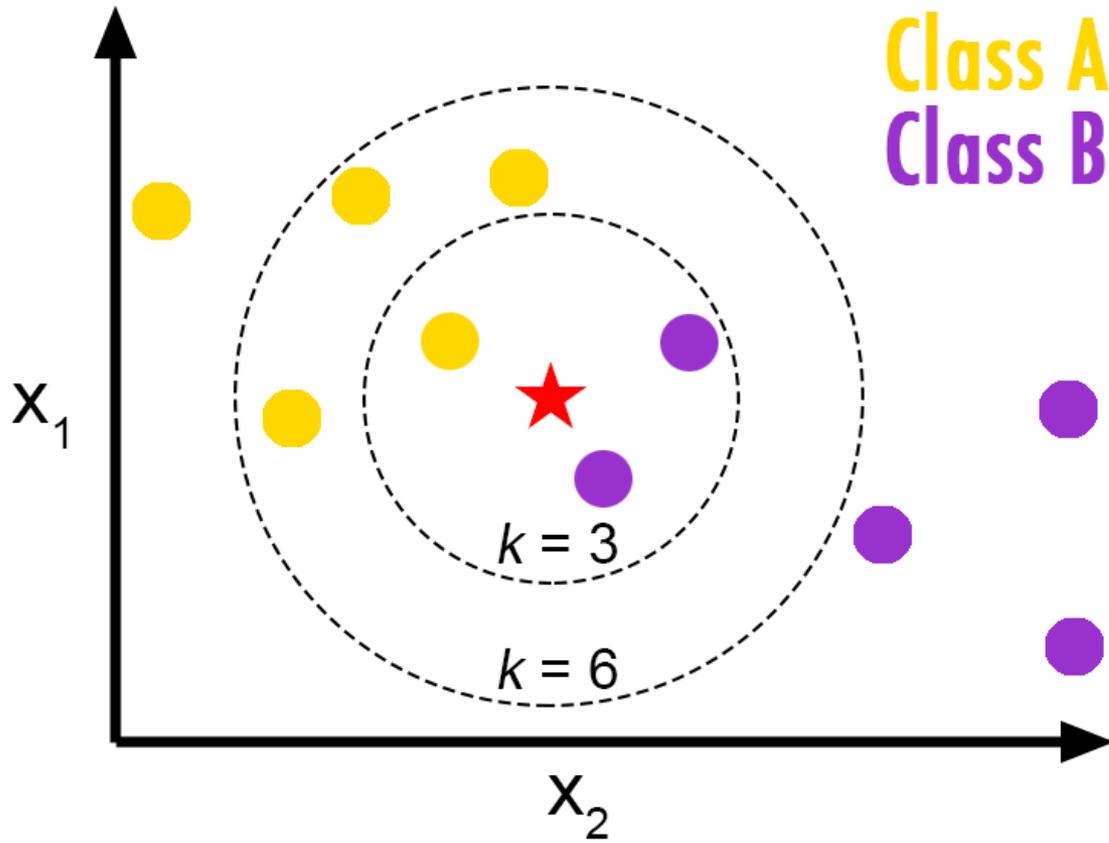
Eraylson Galdino

egs@cin.ufpe.br

Roteiro

- Relembrar sobre o KNN
- Jupyter Notebook
- Mão na massa:
 - Carregar base de dados
 - Dividir para treinamento e teste
 - Treinar o modelo
 - Avaliar o modelo
- Aplicação
- Pontos negativos do KNN

KNN



KNN

- Definir um valor para K;
- Para cada novo padrão:
 - Calcular a distância entre o novo padrão e os padrões já armazenados;
 - Retornar as classes dos k padrões mais próximos;
 - Classificar o novo padrão com a classe majoritária;



Jupyter

Mão na Massa: Carregar os dados

```
In [1]: from sklearn.datasets import load_iris
```

```
In [2]: dados_iris = load_iris()
```

```
In [3]: x_entrada = dados_iris.data  
        y_saida = dados_iris.target
```

Dividir em Treinamento e Teste

```
#dividir os dados em treinamento e teste  
from sklearn.cross_validation import train_test_split
```

```
x_treinamento, x_teste, y_treinamento, y_teste = train_test_split(x_entrada, y_saida, test_size=0.3)
```

Treinar o modelo

```
#realiza a criação do modelo  
from sklearn.neighbors import KNeighborsClassifier
```

```
knn_1 = KNeighborsClassifier(n_neighbors=3)
```

```
#treinamento  
knn_1.fit(x_treinamento, y_treinamento)
```

Avaliar o modelo

```
: #avaliando o modelo no conjunto de treinamento  
  from sklearn.metrics import accuracy_score
```

```
: pred_treinamento = knn_1.predict(x_treinamento)
```

```
: print(accuracy_score(y_treinamento, pred_treinamento))
```

Classificando Câncer

- Através do conjunto de dados breast cancer, crie um modelo KNN que possa classificar os novos dados com uma precisão maior que 90%.

Pontos Negativos do KNN

- Custoso para produção;
- Dimensão dos dados;
- Interpretação dos resultados;