

Aprendizado de Máquina e Deep Learning

Implementação K-means

Prof. Dr. Thiago Meirelles Ventura

Criação de bases para experimentos

```
from sklearn.datasets import make_blobs
```

```
seed = 24
```

```
n = 1500
```

```
X, y = make_blobs(n_samples=n,  
                  random_state=seed)
```

K-means

- Importação

```
from sklearn.cluster import KMeans
```

K-means

- Modelo e treinamento

```
kmeans = KMeans(n_clusters=3, random_state=seed)
```

```
kmeans.fit(X)
```

K-means

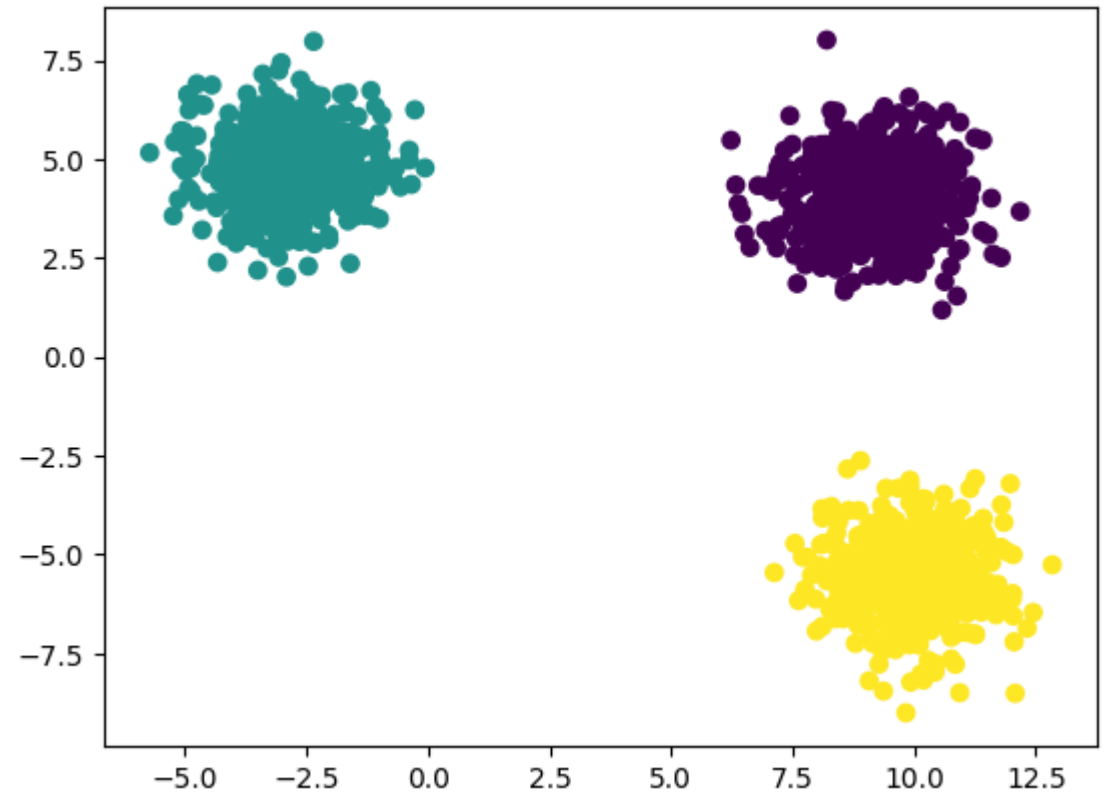
- Predição

```
predicao = kmeans.predict(X)
```

K-means

- Visualização

```
plt.scatter(X[:, 0],  
            X[:, 1],  
            c=predicao)
```



K-means

script kmeans.ipynb

Exercício 4

- Use o método "predict" com novos valores e verifique se faz sentido os grupos indicados a estes dados

Exercício 5

- Modifique o número de clusters e verifique como ficam os resultados
 - Teste com valores menores que 3
 - Teste com valores maiores que 3

Exercício 6

- Dados: happiness
 - Contém características que podem representar o nível de felicidade da população de um país
 - Além do nome do país, há atributos de índices que representam o produto interno bruto, suporte social, expectativa de vida, liberdade, generosidade e corrupção
- Agrupe os países de acordo com o nível de felicidade da população
 - Utilize o método K-means para criar 10 clusters com os dados de treinamento
 - Avalie em qual grupo pertence cada país

Exercício 7

- Avalie o resultado dos clusters criados com dados de teste

Outros métodos

- O Scikit learn possui diversos métodos para executar um aprendizado não supervisionado
- K-means é apenas um deles
- Pode ser visto outros métodos de clustering neste link
 - <https://scikit-learn.org/stable/modules/clustering.html>