

Exercício prático - PCA-kNN

Analise o programa em anexo ao material da aula que compara diferentes estratégias para a inicializações do algoritmo k-means e responda as seguintes questões:

1. (Téorica) Com base no arquivo `kmeans.py`, descreva as métricas utilizadas para avaliar a qualidade dos clusters produzidos pelo algoritmo k-means. Segundo as métricas, qual o efeito ao se utilizar o PCA?
2. (Prática) Considerando as notebooks (e a base de dados) disponibilizados no arquivo `leaf.zip`:
 - (a) procure pela base no kaggle, para mais detalhes.
 - (b) analise os códigos, procurando relacionar o que foi visto na disciplina com o que está implementado.
 - (c)
3. (Prática) Acrescente no programa disponibilizado no arquivo `kmeans.py` o “método de **Elbow**”, que procura determinar a quantidade ideal de clusters.