

Exercício Prático 01 - Primeiro contato com o Scikit-Learn

1. Como elaborar

Utilize o Google Colab. Se preferir, use o notebook que foi disponibilizado como ponto de partida.

2. Contextualização

Na primeira aula, vimos um exemplo de como carregar datasets de exemplo ("toy datasets") do Scikit-learn. O dataset visto foi o Iris. Vimos também como funciona o fluxo básico de aprendizagem supervisionada scikit-learn: preparação de dados, treinamento e predição. Para isto, utilizamos dados produzidos artificialmente.

3. O que fazer neste exercício

Nesta atividade, você deve implementar o fluxo descrito abaixo, utilizando o scikit-learn, adaptando os conceitos vistos nos exemplos em sala.

1. Carregar o dataset Iris, utilizando o scikit-learn.
2. Usando matplotlib / pyplot, gere algumas visualizações para ter uma ideia como certos pares de features se relacionam para dar algum tipo de insight sobre sua utilidade. Se estiver com pressa, pule essa etapa.
3. Instancie um objeto *model* da classe *GaussianNB* do módulo *navie_bayes*. Não precisa especificar parâmetros.
4. Treine um modelo para classificação das flores com todos os exemplos da base.
5. Faça 6 predições diferentes especificando dados elaborados manualmente. Antes, veja alguns exemplos contidos na base de treino para ter ideia (grosseira) da faixa de valores das features.

