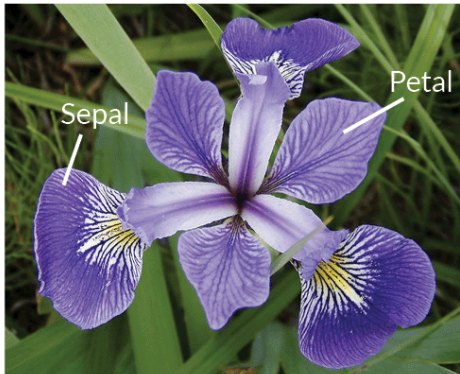


## 1. Objetivo

O objetivo principal deste trabalho é projetar e analisar os modelos de aprendizado de máquina criados para identificar corretamente a espécie de flor na base das características observadas da sépala e da pétala. Durante o tempo de execução do projeto, você terá que projetar o pré- processamento dos dados, experimentar com diferentes métodos de classificação e avaliar o desempenho do modelo gerado.

## 2. Dataset



**Iris Versicolor**



**Iris Setosa**



**Iris Virginica**

O dataset Iris é um dos conjuntos de dados mais conhecidos e amplamente utilizados em machine learning e estatística, especialmente para problemas de classificação. Ele foi coletado por Ronald A. Fisher em 1936 e contém informações sobre três espécies de flores do gênero *Iris*:

- Iris setosa
- Iris versicolor
- Iris virginica

### Estrutura do Dataset

O dataset contém 150 amostras, divididas igualmente entre as três espécies de flores (50 amostras por espécie). Cada amostra possui 4 características medidas em centímetros, que representam aspectos morfológicos das flores. Essas características são:

- **SepalLengthCm**: Comprimento da sépala
- **SepalWidthCm**: Largura da sépala
- **PetalLengthCm**: Comprimento da pétala
- **PetalWidthCm**: Largura da pétala

Além dessas características, o dataset contém uma coluna chamada Species, que é o alvo de classificação e indica a espécie de cada flor.

### 3. Código

Neste trabalho, o objetivo é completar o código fornecido no arquivo **main.py**, utilizando o dataset **iris**. Esse exercício visa praticar o desenvolvimento de modelos de machine learning, desde o carregamento e pré-processamento de dados até o treinamento e teste dos modelos.

#### Estrutura do Projeto

O projeto é composto por duas partes principais:

1. **main.py** - Arquivo que contém o código base para o trabalho. Esse é o código que você deve completar.
2. **iris.data** - Arquivo com o dataset Iris. Este dataset é amplamente utilizado para classificação de espécies de flores e está em formato CSV.

#### Instruções Detalhadas

##### 1. Classe **Modelo**

A classe **Modelo** no **main.py** já possui uma estrutura básica que você deve completar. Ela contém os métodos:

- **CarregarDataset:** Esta função deve carregar o dataset **iris**. Utilize o argumento `path` para especificar o caminho do arquivo.
- **TratamentoDeDados:** Aqui, você deve realizar o pré-processamento dos dados. Sugestões:
  - Use **`self.df.head()`** para visualizar o início dos dados e compreender sua estrutura.
  - Verifique se há dados faltantes e, caso necessário, trate-os.
  - Analise os dados e decida quais colunas podem ser mais relevantes ou devem ser removidas.
- **Treinamento:** Este método deve realizar o treinamento do modelo. Sugestões:
  - Divida os dados em conjuntos de treino e teste usando **`train_test_split`**.
  - Experimente diferentes modelos (como SVM ou regressão linear) para ver qual deles apresenta melhores resultados.
  - Para melhorar a acurácia, considere usar técnicas como validação cruzada (cross-validation).
- **Teste:** Este método é responsável por testar o modelo treinado.

##### 2. Treinamento e Comparação de Modelos

- É solicitado que você implemente ao menos dois modelos diferentes e compare os resultados. No arquivo, já estão importados os modelos de SVM e Regressão Linear para que você possa começar, mas você está livre para escolher outros modelos, caso prefira.

### 3. Método **Train**

- Esse método está quase completo, e sua função é orquestrar o processo, chamando os métodos de carregamento de dados, tratamento, treinamento e teste.

### 4. **Requisitos e Entrega**

- Mínimo de Dois Modelos: Complete o projeto implementando pelo menos dois modelos de machine learning e compare-os.
- Análise e Documentação: Para cada modelo, analise a acurácia obtida e documente as diferenças entre eles.
- Testes e Conclusão: Após o treinamento, utilize o método **Teste** para avaliar o desempenho dos modelos.
- O código deve ser colocado no github até dia 11/11/2024
  - No mesmo dia iremos disponibilizar um formulário para vocês nos passarem o link do projeto no github de vocês