

Disciplina de Aprendizado de Máquina - 2024/02
Professor Frederico Gualberto Ferreira Coelho

Tema: Métodos de Reamostragem - Validação Cruzada (Hold-Out, Leave-One-Out e K-Fold), Bootstrap.

Integrantes: Áquila Oliveira, Daniel Nogueira, Felipe Augusto, Felipe Veloso, João Viola, Júlia Machado, Raissa Diniz, Rafael Camacho

Exercício Proposto: Avaliação de modelos com métodos de reamostragem

Objetivo: Avaliar o desempenho de um modelo de classificação usando diferentes métodos de reamostragem para validar os resultados.

Vamos utilizar o dataset "*Iris*" da biblioteca `sklearn` para construir um modelo de classificação que prevê a espécie de uma flor com base em características como comprimento e largura das pétalas e sépalas. A ideia é explorar como diferentes métodos de reamostragem afetam a avaliação do modelo.

Instruções:

- 1. Carregar o dataset Iris:**
Utilizar `sklearn.datasets.load_iris()` para carregar o dataset.
- 2. Pré-processamento dos dados:**
Dividir os dados em features (X) e target (y).
- 3. Criar um modelo de classificação:**
Utilizar um classificador simples como o `LogisticRegression` do `sklearn`.
- 4. Aplicar e comparar métodos de validação cruzada:**
 - **Hold-Out (Divisão Simples):** Separar os dados em treinamento e teste (80%-20%) usando `train_test_split`.
 - **K-Fold Cross-Validation (K=5):** Usar `KFold` do `sklearn.model_selection` com K=5.
 - **Leave-One-Out Cross-Validation (LOO):** Usar `LeaveOneOut` do `sklearn.model_selection`.
 - **Bootstrap:** Implementar uma validação com reamostragem bootstrap (bônus: podemos usar `resample` do `sklearn.utils` para criar as amostras).
- 5. Avaliação do Modelo:** Calcular e comparar as métricas de desempenho em cada método de validação.

A ideia é disponibilizar de antemão o código que faz tudo isso, semelhante aos exercícios propostos em sala.

Tarefa para a Turma:

1. Executar o código e comparar os resultados dos diferentes métodos de reamostragem.
2. Modificar o número de `splits` no K-Fold para 10 e analisar os impactos nos resultados.
3. Discutir os prós e contras de cada método de reamostragem no contexto do problema.