



# Competição do Kaggle - 1

## Aprendizado de Máquina

Prof. Me. Otávio Parraga

---

## Introdução

A plataforma **Kaggle** é um ambiente online que reúne cientistas de dados para resolver problemas reais por meio de competições. Nesta atividade, você utilizará ferramentas como **Pandas** (para manipulação de dados) e **Scikit-Learn** (para modelos de machine learning) para desenvolver um classificador de problemas de coluna vertebral. Ambientes interativos como Jupyter Notebooks, Google Colab ou Kaggle Notebooks são recomendados para implementação e teste do código.

---

## Objetivo e Tarefas

O objetivo é desenvolver um modelo de classificação para identificar se um paciente possui **Hérnia**, **Espondilolistese** ou **nenhum problema** (Normal). Para isso, você deverá:

1. Implementar os modelos **K-Nearest Neighbors (KNN)**, **Naïve Bayes** e **Árvores de Decisão** usando o Scikit-Learn.
2. Realizar validação cruzada, ajustar hiperparâmetros e treinar os modelos com o conjunto de dados anotados.
3. Gerar previsões para o conjunto de teste (sem anotações) e submeter o arquivo formatado no Kaggle, seguindo o exemplo fornecido.

A atividade pode ser feita individualmente ou em grupos de até 3 pessoas. Utilize os notebooks de exemplo disponíveis no Moodle como referência.

---

## Pontos Extra (0.5 na Primeira Prova)

Para ganhar a bonificação, é necessário:

- Atingir **acurácia mínima de 81%** no desafio.
- Postar o notebook completo que gerou o resultado no fórum do Moodle, garantindo que os resultados sejam reproduzíveis (ex.: seeds fixos, pré-processamento claro).
- Apenas um integrante do grupo deve enviar a submissão final no Kaggle.