

## Especialização em Inteligência Artificial – IFMG

### Trabalho da disciplina de Aprendizado de Máquina

Escolher **duas** bases de dados para a tarefa de aprendizado supervisionado do tipo classificação. As bases de dados podem ser relacionadas ao tema de pesquisa do aluno ou podem ser retiradas de repositórios públicos tais como:

- UC Irvine Machine Learning Repository
  - <https://archive.ics.uci.edu/>
- Kaggle Machine Learning and Data Science community
  - <https://www.kaggle.com/>
- Penn Machine Learning Benchmarks (PMLB)
  - *Olson, Randal S., William La Cava, Patryk Orzechowski, Ryan J. Urbanowicz, and Jason H. Moore. PMLB: a large benchmark suite for machine learning evaluation and comparison. BioData mining 10, no. 1 (2017): 1-13. BioData Mining 10, page 36.*

Para cada base de dados, efetuar o pipeline completo de aprendizado de máquina para os seguintes métodos: K-NN, Naive Bayes, Árvore de decisão.

Lembre-se de separar seus dados em conjuntos de treinamento e teste, utilizando alguma metodologia vista em aula para a partição: *hold-out*, *k-fold cross validation*, etc.

Discutir os resultados obtidos quanto ao desempenho estimado e comparar os diferentes métodos.

**Entregável: relatório individual com o desenvolvimento feito, resultados obtidos, discussões e conclusões.**

**Data de entrega: 28 de julho de 2023.**

Professor Frederico Gadelha Guimarães  
DCC – UFMG

[fredericoguimaraes@ufmg.br](mailto:fredericoguimaraes@ufmg.br)  
<https://minds.eng.ufmg.br/>