

UTFPR - Departamento de Computação
Inteligência Artificial
Projeto Final - Aprendizado de Máquina

Prof. Danilo Sipoli Sanches
danilo@utfpr.edu.br

31 de Outubro de 2023

1. Considere as seguintes informações:

- Base de dados **Adult Census Income** (<https://www.kaggle.com/datasets/uciml/adult-census-income>), pede-se:
- Utilizar os classificadores: Poderá ser utilizado os classificadores Decision Tree, KNN, Redes Neurais e K-Means.

2. A notas de projeto serão classificadas de acordo com as seguintes etapas:

- Limpeza e pré-processamento dos dados(Remoção de dados faltantes NaN, transformação de dados categóricos e normalização);
- Balanceamento das classes;
- Comparação com todos os classificadores descritos anteriormente;
- Ajuste dos hiperparâmetros com GridSearchCV e RandonSearchCV;
- Amostragem por validação cruzada estratificada (10 folds);
- Considerar uma versão da base de dados sem o atributo classe **income**; Faça o agrupamento dos dados baseado no algoritmo K-Means. Para isso, utilize as colunas **Age** e **education** (para education, considerar etapa de transformação dos dados categóricos com LabelEncoder).

3. Informações sobre o envio:

- Projeto individual ou grupo (máximo 3 alunos);

- Data de envio: 30 de Novembro;
- Enviar os arquivos: .ipynb e .pdf do código fonte(imprimir o notebook no formato pdf com a saída de todas as células)
- Enviar o link do vídeo gravado com uma breve explicação do projeto (máximo 10 minutos) .