

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

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

3 de novembro de 2023

1. Considere as seguintes informações:

- Base de dados:
 - **Travel Insurance Prediction Data** (<https://www.kaggle.com/datasets/tejashvi14/travel-insurance-prediction-data>);
 - **Student Stress Factors: A Comprehensive Analysis** (<https://www.kaggle.com/dataset/stress-factors-a-comprehensive-analysis>);
- Utilizar os classificadores: Decision Tree, KNN e K-Means.

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

- Pré-processamento dos dados(transformação de dados categóricos e normalização);
- Apresentação dos resultados de acurácia e matriz confusão;
- **Travel Insurance Prediction**
 - Considerar uma versão da base de dados sem o atributo classe **TravelInsurance**; Faça o agrupamento dos dados baseado no algoritmo K-Means. Para isso, utilize as colunas **AnnualIncome** e **FrequentFlyer**. Realizar a plotagem da distribuição dos dados por cluster;
- **Student Stress Factors: A Comprehensive Analysis**
 - Considerar uma versão da base de dados sem o atributo classe **stress_level**; Faça o agrupamento dos dados baseado no algoritmo K-Means. Para isso, utilize as colunas **anxiety_level** e **sleep_quality**. Realizar a plotagem da distribuição dos dados por cluster;

3. Informações sobre o envio:

- Projeto individual;
- Data de envio: 26 de Novembro;
- Enviar os arquivos: .ipynb e .pdf do código fonte(imprimir o notebook no formato pdf com a saída de todas as células)