U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

André Sousa (up202005277@fe.up.pt)
Gonçalo Pinto (up202004907@fe.up.pt)
Pedro Fonseca (up202008307@fe.up.pt)

Supervised Learning

AI Second Project

Specification of the work

- **Goal:** Analyze and study the process of classification (data preprocessing, feature selection, model selection, model training and model evaluation). In this case, the goal is to use data collected from past records of an insurance company to predict whether a given client will file a claim in the next 6 months.

- **Project Objectives:**
  - Given a data set, predict the target feature
  - Compare multiple machine learning models and feature selection to achieve the best results

Software used and related work

- Stuart Russell, Peter Norvig; Artificial intelligence. ISBN: 978-0-13-207148-2
- Richard S. Sutton; Reinforcement learning. ISBN: 978-0-262-03924-6
- Stuart Russel, Peter Norvig; Artificial Intelligence: A modern Approach.
- Exercise5 IART SupervisedLearning
- Other approaches to the problem

Tools and algorithms

- PyCharm Professional
- Python 3.9
- Git
- Python packages:
  - scikit-learn
  - Pandas
  - Seaborn
  - Matplotlib
- Machine Learning models:
  - Decision Tree
  - Neural Network (MLP)
  - k-nearest neighbors (KNN)
  - Support Vector Machine (SVM)

Implementation work already carried out

- Data preprocess
- Correlation analysis
- Feature selection (based on correlation results)
- Initial algorithm comparison