

## 1. Introdução

A segunda meta (M2) da avaliação prática, tem como objetivos principais:

- produzir uma abordagem sistemática à aplicação de modelos de *Supervised Learning (SL)*, tendo em consideração o *dataset* selecionado e preparado na meta M1.
- tendo por base estratégias baseadas em *Unsupervised Learning*, avaliar se é possível melhorar a classificação obtida com os modelos SL.

## 2. Meta M2

### 2.1 Supervised Learning

Nesta meta M2 pretende-se criar uma *short-list* de algoritmos de SL que possam ser utilizados no problema de classificação proposto.

- A seleção dos algoritmos deve ser efetuada tendo em consideração que se trata de um problema de classificação binária. Esta seleção deve ser devidamente justificada.
- Os dados devem ser preparados/transformados de forma a potenciar a performance dos classificadores.
- A avaliação dos modelos candidatos a integrar a *short-list* deve prever *cross-validation*.
- As métricas para avaliação do classificador devem ser escolhidas tendo por base as características do *dataset* a classificar.
- A comparação do desempenho dos classificadores deve ser devidamente justificada.

## 2.2. Unsupervised Learning

- Os algoritmos de *clustering* podem ser considerados de forma isolada ou, em alternativa como pré-processamento para tentar melhorar a performance dos classificadores anteriormente utilizados. As duas abordagens devem ser exploradas (anexo1\_M2).
- Devem ser considerados diversos algoritmos de *clustering* (<https://scikit-learn.org/stable/modules/clustering.html>; (anexo2\_M2).
- Problema de classificação binária.
- Os dados devem ser preparados/transformados de forma a potenciar a performance dos algoritmos.
- A avaliação do desempenho dos algoritmos deve ter em consideração que se trata de um problema de aprendizagem não supervisionada.

## 3. Avaliação

Na implementação desta meta M2 devem ser observadas as seguintes regras:

- os grupos formados na meta M1 permanecem naturalmente inalterados;
- deve ser elaborada uma pequena apresentação (ppt; **10 minutos**) a apresentar no dia **11/01/2023**. Os dois elementos do grupo terão de ter uma participação equivalente durante a apresentação.
- deve ser elaborado um pequeno documento em que apresente de forma totalmente objetiva e condensada (máx. 10 páginas) a abordagem seguida;
- implementação com base em Jupyter Notebook;
- esta meta de avaliação representa 5 valores (25%) da nota final da disciplina.
- atempadamente será disponibilizado um pequeno formulário com os horários possíveis para a realização da apresentação;
- todos os elementos de avaliação (\*.ipynb; \*.csv; \*.doc; \*.ppt) devem ser entregues na plataforma nónio, num ficheiro comprimido (.zip), **até às 23h59 de 09/01/2023**, com a seguinte nomenclatura:

*primeiroUltimoNomeAluno1\_nºAluno1\_primeiroUltimoNomeAluno2\_nºAluno2*

**Exemplo:** *FernandoPessoa\_1231234\_MiguelTorga\_1234123*