



Universidade do Minho

Licenciatura em Engenharia Informática

Aprendizagem e Decisão Inteligentes

3º Ano, 2º Semestre

Ano letivo 2024/2025

Trabalho de Grupo

Fevereiro, 2025

Tema

Conceção de modelos de aprendizagem e decisão.

Objetivos de aprendizagem

Com a realização deste Trabalho de Grupo pretende-se sensibilizar e motivar os estudantes para a conceção e desenvolvimento de um projeto prático, no âmbito da unidade curricular de Aprendizagem e Decisão Inteligentes, utilizando diferentes paradigmas de aprendizagem e recorrendo a técnicas de *machine learning* abordados no decurso da unidade curricular.

Enunciado

Este enunciado é o ponto de partida para a implementação de um projeto de análise e extração de conhecimento, atendendo à distribuição dos *datasets* disponíveis e o respetivo desenvolvimento de modelos de aprendizagem. Os problemas propostos deverão ser resolvidos em grupo e com recurso à plataforma [KNIME](#).

O Trabalho de Grupo engloba duas tarefas e nenhuma tem precedência sobre a outra, pelo que devem ser abordadas em simultâneo.

Tarefa **Dataset Grupo**

- Consultar, analisar e selecionar um *dataset* de entre os que estão acessíveis a partir de fontes como, por exemplo, o [Kaggle](#) ou o [UCI Machine Learning Repository](#);
- Explorar, analisar e preparar os dados do *dataset* selecionado, procurando extrair conhecimento relevante no contexto do problema em análise;
- Conceber e otimizar diversos modelos de *Machine Learning*;
- Realizar uma análise crítica de resultados.

Tarefa **Dataset Atribuído**

- Os grupos com número ímpar trabalharão com o *dataset* disponível no Blackboard em «Dataset grupos ímpar»;
- Os grupos com número par trabalharão com o *dataset* disponibilizado no Blackboard em «Dataset grupos par»;
- O ficheiro zip de cada *dataset* disponibiliza informação de contexto, que pode ser entendida complementar aos dados fornecidos sobre cada problema;
- Estes *datasets* não poderão ser utilizados para a **Tarefa Dataset Grupo**;
- Explorar, analisar e preparar os dados do caso atribuído, procurando extrair conhecimento relevante no contexto do problema em análise;
- Conceber modelos de *Machine Learning* de classificação e de regressão;
- Realizar uma análise crítica de resultados.

Nesta tarefa, cada conjunto de dados é composto pelos seguintes ficheiros:

- [*_description.txt*] breve descrição dos dados;
- [*_train.csv/xlsx*] *dataset* com dados para treino;
- [*_test.csv/xlsx*] *dataset* com dados para teste;
- [*_new_problems.csv/xlsx*] dados com problemas novos (sem solução);

Resultados

Os resultados deverão ser objeto de um único relatório, limitado a 30 páginas, que apresente, entre outros:

- Os domínios a tratar, os objetivos a alcançar e como os atingir;
- A metodologia seguida e como foi aplicada;
- A descrição e exploração detalhada dos *datasets* de ambas as tarefas e do tratamento de dados efetuado;
- A descrição dos modelos desenvolvidos e quais as suas características, parâmetros de treino, entre outros detalhes que enriqueçam a explicação;
- O sumário dos resultados obtidos e respetiva análise crítica;
- A apresentação de sugestões e recomendações após análise dos resultados obtidos e dos modelos desenvolvidos.

A descrição do trabalho desenvolvido deverá ser acompanhada de exemplos e indicações que permitam reproduzir os passos realizados e os resultados alcançados.

Entrega e Avaliação

Durante o período de aulas na semana de 24 a 29 de março de 2025 realizar-se-á um *checkpoint* do trabalho desenvolvido pelos grupos. Cada grupo disporá de 10 minutos.

Na semana de 2 a 5 de junho de 2025 decorrerão as apresentações finais do trabalho desenvolvido. Cada grupo disporá de 15 minutos para realizar a apresentação, utilizando os meios que considerar adequados.

A organização do *checkpoint* e da apresentação final será divulgada oportunamente.

O relatório e os restantes elementos produzidos deverão ser compactados num único ficheiro que será submetido na plataforma de *e-learning* da Universidade do Minho, por apenas um membro de cada grupo.

A data limite para a submissão dos trabalhos é o dia 19 de maio de 2025.

Avaliação por pares

Cada grupo deverá realizar uma análise coletiva sobre o contributo que cada membro apresentou no desenvolvimento do trabalho, definindo a avaliação por pares (APP). Para o efeito, dispõe de tantos pontos quantos os membros do grupo (p.ex., grupo de 3 pessoas = 3 pontos; grupo de 4 pessoas = 4 pontos).

Na APP o grupo deve distribuir os seus pontos proporcionalmente ao contributo relativo de cada um dos seus membros. A APP deve ser consensualizada no seio do grupo de trabalho.

Por exemplo, A=1, B=1 e C=1, significa que todas as pessoas A, B e C desempenharam as suas tarefas em condições de equidade. Noutro caso, X=1, Y=1,5 e Z=0,5, significa que X participou no trabalho na medida do esperado, que Y se envolveu com mais esforço e que Z teve um empenho inferior ao esperado.

A soma final de todas as parcelas deve ser igual ao número de pontos atribuído ao grupo: $X_1 + X_2 + \dots + X_N = N$, sendo N o número de elementos do grupo de trabalho e X_i os pontos atribuídos a cada um dos seus membros "i", podendo ser utilizados valores com até 1 casa decimal.

Cada grupo deverá enviar a APP numa mensagem de correio eletrónico para a equipa docente (Cesar Analide analide@di.uminho.pt, Inês Alves d12090@di.uminho.pt, Pedro Oliveira pedro.jose.oliveira@algoritmi.uminho.pt, Afonso Bessa d13959@di.uminho.pt), colocando em cc: todos os membros do grupo de trabalho, com assunto/ *subject* «[ADI^3] APP Grupo XX».

As classificações do instrumento de avaliação Trabalho de Grupo não são calculadas na falta desta informação.

Código de conduta

Os intervenientes neste trabalho académico declararão ter atuado com integridade e confirmarão não ter recorrido a práticas de plágio nem a qualquer forma de utilização indevida de informação ou falsificação de resultados em nenhuma etapa decorrente da sua elaboração.

Mais declararão conhecer e respeitar o [Código de Conduta Ética da Universidade do Minho](#).

Referências bibliográficas

Além do material disponibilizado nas aulas, aconselha-se a consulta de fontes como:

- Machine Learning. T. Michell, McGraw Hill, ISBN ISBN: 978-1259096952, 2017.
- Introduction to Machine Learning. Alpaydin, E. ISBN: 978-0-262-02818-9. Published by The MIT Press, 2014.
- Computational Intelligence: An Introduction. Engelbrecht A., Wiley & Sons. 2nd Edition, ISBN: 978-0470035610, 2007.
- The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Hastie, T., R. Tibshirani, J. Friedman, 12th Edition, Springer, ISBN: 978-0387848570, 2016.
- Machine Learning: A Probabilistic Perspective. K.P. Murphy, 4th Edition, The MIT Press, ISBN: 978-0262018029, 2012.