



**Universidade do Minho**

Licenciatura em Engenharia Informática

Mestrado integrado em Engenharia Informática

Aprendizagem e Decisão Inteligentes

3º/4º ano, 2º semestre

Ano letivo 2021/2022

Trabalho Prático de Grupo

Março, 2022

<b>Tema</b>	Conceção de modelos de aprendizagem.
<b>Objetivos de Aprendizagem</b>	Com a realização deste Trabalho Prático de Grupo pretende-se sensibilizar e motivar os alunos para a conceção e desenvolvimento de um projeto, no âmbito da unidade curricular de Aprendizagem e Decisão Inteligentes, utilizando os modelos de aprendizagem abordados ao longo do semestre.
<b>Enunciado</b>	<p>Este enunciado pretende ser o ponto de partida para a análise e extração de conhecimento atendendo à distribuição de dados disponíveis, e o respetivo desenvolvimento de modelos de aprendizagem. Para a resolução do problema proposto, a utilização da plataforma <a href="#">KNIME</a> será obrigatória.</p> <p>Este enunciado prático engloba duas fases.</p> <p>A <u>1ª fase da componente prática de avaliação em grupo</u> incidirá sobre as seguintes tarefas:</p> <ol style="list-style-type: none"><li>1. Consultar, analisar e selecionar um <i>dataset</i> de entre os que estão acessíveis a partir de fontes como, por exemplo, o <a href="#">Google Dataset Search</a>, o <a href="#">Kaggle</a>, ou a <a href="#">UCI Machine Learning Repository</a>;</li><li>2. Para além do <i>dataset</i> selecionado no ponto 1., os grupos deverão também trabalhar com um <i>dataset</i> proposto pela equipa docente, nos seguintes termos:<ul style="list-style-type: none"><li>• Os grupos 1, 4, 7 ... 58 trabalharão com o <i>dataset</i> disponível em (wine_quality_classification), de modo a determinar a qualidade de um vinho tinto produzido em Portugal;</li><li>• Os grupos 2, 5, 8 ... 59 trabalharão com o <i>dataset</i> disponível em (USA_housing_regression), com o objetivo de estimar o valor da habitação numa região dos Estados Unidos da América;</li><li>• Os grupos 3, 6, 9 ... 60 trabalharão com o <i>dataset</i> disponível em (salary_classification), com o intuito de prever o salário de um trabalhador, atendendo a um conjunto de características;</li><li>• O <i>link</i> disponibilizado para cada <i>dataset</i> proposto pela equipa docente, permite o acesso aos dados e à sua descrição;</li><li>• Estes <i>datasets</i> propostos não poderão ser utilizados para a tarefa dada no ponto 1.</li></ul></li><li>3. Explorar, analisar e preparar ambos os <i>datasets</i>, procurando extrair conhecimento relevante no contexto dos problemas em questão;</li></ol> <p>A <u>2ª fase da componente prática de avaliação em grupo</u> incidirá sobre as seguintes tarefas:</p> <ol style="list-style-type: none"><li>4. Conceção de modelos de aprendizagem para ambos os problemas, i.e., para o <i>dataset</i> selecionado pelo grupo de trabalho e para o <i>dataset</i> proposto pela equipa docente;</li><li>5. Obtenção e análise crítica de resultados;</li><li>6. Interpretação dos resultados adquiridos e definição da sua utilidade no contexto dos problemas subjacentes aos <i>datasets</i> trabalhados. Determinar e explicitar os resultados mais relevantes.</li></ol>

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

7. Quais os domínios a tratar, quais os objetivos a alcançar e como os atingir;
8. Qual a metodologia seguida e como foi aplicada;
9. Descrição e exploração detalhada de ambos os *datasets* e do tratamento de dados efetuado;
10. Descrição dos modelos desenvolvidos e quais as suas características, parâmetros de treino, entre outros detalhes que enriqueçam a explicação;
11. Sumário dos resultados obtidos e respetiva análise crítica;
12. Apresentação de sugestões e recomendações após análise dos resultados obtidos e dos modelos desenvolvidos.

Todo o processo deverá ser acompanhado de exemplos e indicações que permitam reproduzir todos os passos realizados bem como os resultados obtidos.

### **Entrega e Avaliação**

Durante o período de aulas na semana de 4 a 9 de abril de 2022 realizar-se-á um breve *checkpoint* do trabalho desenvolvido pelos grupos na 1ª fase. Este *checkpoint* decorrerá de forma remota, num período de 10 minutos por grupo, sendo utilizada a ferramenta *Blackboard Collaborate Ultra* da plataforma de *e-learning* da Universidade do Minho.

Na semana de 9 a 14 de maio de 2022 decorrerão as apresentações finais do trabalho desenvolvido. Cada grupo disporá de 15 minutos para realizar a apresentação, utilizando os meios que considerar mais adequados para o efeito. A apresentação decorrerá de forma remota, através da ferramenta *Blackboard Collaborate Ultra* da plataforma de *e-learning* da Universidade do Minho.

A organização das apresentações de ambas as fases será divulgada oportunamente.

O relatório, assim como os restantes elementos produzidos, deverão ser compactados num único ficheiro que deverá ser submetido na plataforma de *e-learning* da Universidade do Minho (em [Conteúdo/Instrumentos de Avaliação em Grupo/Submissão TPG] ), por apenas um membro de cada grupo.

A data limite para a submissão dos trabalhos é o dia 6 de maio de 2022.

### **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: 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, 12nd Edition, Springer, ISBN: 978-0387848570, 2016.
- Machine Learning: A Probabilistic Perspective. K.P. Murphy, 4th Edition, The MIT Press, ISBN: 978-0262018029, 2012.