



**Universidade do Minho**

Licenciatura em Engenharia Informática

Mestrado Integrado em Engenharia Informática

Aprendizagem e Decisão Inteligentes

3º/4º ano, 2º Semestre

Ano letivo 2022/2023

Enunciado Prático n.º 6

Março, 2023

<b>Tema</b>	Aplicação de técnicas de aprendizagem com KNIME: Regressão
<b>Objetivos de aprendizagem</b>	Com a realização desta ficha prática pretende-se que os alunos: <ul style="list-style-type: none"><li>• Apliquem nodos de aprendizagem e previsão;</li><li>• Usem nodos de avaliação de modelos;</li></ul>
<b>Enunciado</b>	<p>O problema descrito pelos dados do <i>dataset</i> «wine quality» é composto por 13 atributos (<i>features</i>), dos quais «quality1_10» e «quality_bin» são os atributos <i>target</i> do problema, representando, ambos, uma apreciação a qualidade do vinho feita por enólogos.</p> <p>Os dados descrevem dois problemas idênticos («winequality-red», qualidade de vinhos tintos; «winequality-white», qualidade de vinhos brancos), no sentido em que são caracterizados por atributos com o mesmo significado, sendo as avaliações sobre a sua qualidade resultado de uma apreciação individual e subjetiva de especialistas.</p>
<b>Tarefas</b>	<p>Descarregar o ficheiro disponível na plataforma de <i>e-learning</i> da UMinho, secção [Conteúdo] e realizar as tarefas seguintes:</p> <p>T1. Carregar o <i>dataset</i> «winequality-red» e fazer a exploração de dados;</p> <p>T2. Fazer a preparação e o tratamento de dados;</p> <p>T3. Utilizar o nodo LINEAR REGRESSION LEARNER para criar modelos de regressão e o nodo REGRESSION PREDICTOR para calcular as previsões dos modelos;</p> <p>    T3.1. Avaliar o desempenho dos modelos criados através de diferentes métricas (MAE, MSE, RMSE, entre outras);</p> <p>T4. Utilizar o nodo LOGISTIC REGRESSION LEARNER para criar modelos de classificação e o nodo LOGISTIC REGRESSION PREDICTOR para calcular as previsões dos modelos;</p> <p>    T4.1. Avaliar o desempenho dos modelos criados através de diferentes métricas (ACCURACY, PRECISION, entre outras);</p> <p>T5. Carregar o <i>dataset</i> «winequality-white» e desenvolver abordagem idêntica ao <i>dataset</i> «winequality-red»;</p> <p>T6. Discutir a possibilidade de fazer JOIN dos dois <i>datasets</i>:</p> <p>    a) É admissível proceder deste modo para realizar uma análise coerente do problema da qualidade dos vinhos, independentemente da sua cor?</p> <p>    b) É possível tratar os dois problemas deste modo? Se sim, como? Se não, porquê?</p>