



**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° 5

Março, 2023

<b>Tema</b>	Aplicação de técnicas de aprendizagem com KNIME: Árvores de Decisão
<b>Objetivos de aprendizagem</b>	<p>Com a realização desta ficha prática pretende-se que os alunos:</p> <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> «bcw» (<i>breast cancer wisconsin</i>) é composto por 10 atributos (<i>features</i>), dos quais «class» é o atributo classificador com o significado descrito nos ficheiros disponíveis na plataforma e e-learning.</p> <p>Em síntese, os dados referem-se a diversas características que oncologistas especialistas em cancro de mama avaliaram em termos da sua relevância.</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> «bcw» e usar nodos de exploração de dados, analisando:</p> <ul style="list-style-type: none"><li>a) Tendências centrais</li><li>b) Dispersão estatística</li><li>c) Correlação entre <i>features</i></li></ul> <p>T2. Crie gráficos para visualizar as características dos dados</p> <p>T3. Tratar os dados em falta (<i>missing values</i>).</p> <p>T4. Utilize um nodo DECISION TREE LEARNER para treinar uma árvore de decisão e um nodo DECISION TREE PREDICTOR para calcular as previsões do modelo obtido;</p> <p>T5. Avalie o desempenho dos modelos criados através de diferentes métricas (ACCURARY, PRECISION, entre outras);</p> <p>T6. Experimente diversas parametrizações do nodo DECISION TREE LEARNER, analise e documente os desempenhos alcançados;</p> <p>T7. Experimente outras formas de preparação dos dados e analise as consequências:</p> <ul style="list-style-type: none"><li>a) Selecionar <i>features</i> mais relevantes</li><li>b) Fazer <i>pruning</i></li></ul>