



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 nº 3

Março, 2023

Tema	Preparação e exploração de dados com KNIME.
Objetivos de aprendizagem	<p>Com a realização desta ficha prática pretende-se que os alunos:</p> <ul style="list-style-type: none">• Explore a plataforma de análise de dados KNIME;• Experimentem tarefas de preparação e exploração avançada de dados;
Enunciado	<p>O enunciado diz respeito a dados de alunos de 2 escolas distintas. Na última página deste enunciado podem encontrar os atributos presentes neste <i>dataset</i> e respetiva descrição.</p>
Tarefas	<p>Deverão começar por descarregar o ficheiro disponível na plataforma de <i>e-learning</i> da UMinho, secção [Conteúdo]. O <i>dataset</i> «student_data» contém dados relativos ao desempenho escolar dos alunos destas escolas.</p> <p>Realize as tarefas seguintes:</p> <p>T1. Carregar, no <i>Knime</i>, o <i>dataset</i>.</p> <p>T2. Aplicar nodos para exploração de dados de forma a:</p> <ol style="list-style-type: none">a. Analisar os dados em relação às suas características e padrões;b. Verificar, graficamente, a distribuição de alunos nas diferentes escolas;c. Verificar, graficamente, a distribuição do emprego das mães dos alunos;d. Agrupar por escola e obter a idade média dos alunos que a frequentam. <p>T3. Aplicar nodos de exploração de dados que respondam às seguintes questões:</p> <ol style="list-style-type: none">a. A coabitação (ou não) dos alunos com os pais afeta a sua classificação final (G3)?b. O sexo do estudante afeta a classificação final? E o tempo que demora a chegar à escola?c. Como se comporta a classificação final em relação à idade?d. Como se comporta a classificação final em relação aos tipos de empregos que ambos os pais dos alunos desempenham? E em relação ao nível de educação de ambos?e. Há alguma relação entre a classificação final e o facto de os alunos quererem ou não frequentar o ensino superior?f. Qual a relação entre o número de chumbos e a classificação final? E entre o número de chumbos e o tempo semanal de estudo? <p>T4. Explorar as outras <i>features</i> e retirar informações relevantes sobre o <i>dataset</i> (ex.: fazer e responder a perguntas como em T3).</p>

Atributos do Dataset

school	escola do aluno
sex	género do aluno
age	idade
address	morada do aluno ('U' - urbano, 'R' - rural)
famsize	tamanho da família ('LE3' - menor ou igual que 3, 'GT3' - mais do que 3)
Pstatus	situação de coabitação dos pais ('T' - living together; 'A' - apart)
Medu	nível de educação da mãe (0 - nenhum, 1 - primária, 2 - 3.º ciclo, 3 - ensino secundário, 4 - ensino superior)
Fedu	nível de educação do pai (0 - nenhum, 1 - primária, 2 - 3.º ciclo, 3 - ensino secundário, 4 - ensino superior)
Mjob	tipo de emprego da mãe
Fjob	tipo de emprego do pai
reason	razão de escolha da escolha
guardian	quem detém a guarda do aluno
traveltime	tempo que demora a chegar à escola (1 - <15min., 2 - 15 a 30min., 3 - 30min. a 1h, 4 - >1h)
studytime	tempo semanal de estudo (1 - <2h, 2 - 2 a 5h, 3 - 5 a 10h, 4 - >10h)
failures	número de chumbos
schoolsup	suporte extra educacional
famsup	suporte familiar
paid	aulas extra pagas
activities	actividades extra-curriculares
nursery	frequentou a creche
higher	pretende frequentar o ensino superior
internet	acesso à internet em casa
romantic	numa relação amorosa
famrel	qualidade das relações familiares (de 1 - muito mau a 5 - muito bom)
freetime	tempo livre depois da escola (de 1 - muito baixo a 5 - muito alto)
goout	saídas com amigos (de 1 - muito baixo a 5 - muito alto)
Dalc	consumo de álcool à semana (de 1 - muito baixo a 5 - muito alto)
Walc	consumo de álcool ao fim de semana (de 1 - muito baixo a 5 - muito alto)
health	estado de saúde (de 1 - muito mau a 5 - muito bom)
absences	número de faltas
G1	classificação do 1.º período
G2	classificação do 2.º período
G3	classificação final