



RACIOCÍNIO E REPRESENTAÇÃO DO CONHECIMENTO | 21097 | 2021/22

Instruções:

- Nesta prova considere os 3 dígitos menos significativos do seu número de estudante. Exemplo: no número de estudante 2012345, os três dígitos menos significativos são o número **345**. No enunciado é utilizado d_2 para referir o terceiro dígito menos significativo (aqui 3), ao d_1 o segundo dígito menos significativo (aqui 4) e ao d_0 o dígito menos significativo (aqui 5). Existem também questões que utilizam valores binários com base na paridade destes dígitos. Neste caso as variáveis utilizadas são b_2 a b_0 , ficando com 1 para os dígitos par e com 0 para os dígitos ímpar. No caso deste exemplo, apenas d_1 é par, pelo que b_2 e b_0 são 0, e b_1 é 1. Deve preencher na folha de resolução a seguinte tabela, aqui preenchida com o exemplo.

Número: (exemplo: 2012**345**)

Dígito	Valor	Binário	Valor
d_2	(exemplo: 3)	b_2	(exemplo: 0)
d_1	(exemplo: 4)	b_1	(exemplo: 1)
d_0	(exemplo: 5)	b_0	(exemplo: 0)

Nota: as variáveis binárias podem conter expressões, que têm de ser avaliadas. É utilizada a barra superior para negação e operador + para a disjunção, distinguindo-se dos operadores utilizados nas variáveis

Trabalho a desenvolver:

Pergunta 1 - [1 valor]

No âmbito de um estudo sobre as causas do cancro, foi recolhida informação sobre um conjunto de pessoas com e sem cancro, nos quais foram classificados como estando ou não em cada situação (binário), exceto a idade com três níveis. Pretende-se determinar quais as principais causas que são responsáveis pelo cancro.

Causas:

1. Excesso consumo de **Tabaco** e/ou **Álcool**;
2. **Obesidade** e/ou **Dieta** (rica em gordura);
3. Exposição a radiação **Solar**;
4. **Idade** (jovem, adulto, idoso);
5. Falta de **Exercício Físico**.

Construa uma árvore de decisão para o seguinte conjunto de treino, utilizando o algoritmo Decision-Tree-Learning, indicando **todos os passos**. Opte pela variável que permitir resolver mais casos em cada passo.

Conjunto de Treino: *consultar folha de Excel e colocar número de estudante*

Pergunta 2 - [1 valor]

Considere a situação da pergunta anterior, mas desta vez pretende treinar uma rede neuronal. Treine a rede fazendo uma passagem pelo conjunto de treino (da pergunta anterior), utilizando um valor de alfa de $(3 + d_0 + d_1 + d_2)\%$.

Rede neuronal (*threshold* function) - 1 nível, 1 neurónio com valores iniciais dos pesos pela ordem do conjunto de treino (constante seguido de C1 a C5). Apresente **todos os passos**.

w (nível 1)	11
0	<i>consultar folha de Excel e colocar número de estudante</i>
1	
2	
3	
4	
5	

Pergunta 3 - [2 valores]

Considere a situação das perguntas anteriores. Pretende-se que utilize o conjunto de treino da pergunta 1 para indicar qual a classificação para cada um dos casos do conjunto de teste, para o algoritmo dos 3-vizinhos mais próximos, e para a árvore de decisão da pergunta 1 e rede neuronal da pergunta 2. Atendendo a que a maior parte das variáveis são binárias, considere a distância como sendo o número de elementos distintos.

Conjunto de teste: *consultar folha de Excel e colocar número de estudante*

Critérios de avaliação e cotação

Cotação: O e-fólio tem 3 perguntas com a cotação junto de cada pergunta.

Descontos: Detecção de fraude (total ou parcial): 100%

Normas a respeitar

Forma de entrega: Um ficheiro em formato pdf de até 10 páginas A4, com todos os cálculos e todas as opções tomadas na resolução das perguntas.

Não são aceites entregas fora da plataforma Moodle.