

Nome:

1. **(1,5 pt)** Quais as vantagens do Planejamento de Ordem Parcial (POP) em relação ao Planejamento de Ordem Total? Explique o funcionamento de um planejador POP.
2. **(1,0 pts)** Por que o pré-processamento dos dados é muito importante para os algoritmos de aprendizagem de máquina? Comente algumas técnicas usadas para isso.
3. **(1,0 pts)** Relacione as colunas, associando as tarefas em aprendizagem de máquina às suas características.

(1) Agrupamento	()	Tarefa supervisionada cuja meta é encontrar um modelo a partir dos dados de treino que possa ser usado para prever um rótulo de um novo exemplo.
(2) Associação	()	Tarefa descritiva cuja meta é encontrar um subconjunto dos dados que descreva todo o conjunto de dados.
(3) Classificação	()	Tarefa preditiva cuja meta é encontrar uma função a partir dos dados de treino que possa ser usada para definir o valor (contínuo) de um novo exemplo.
(4) Regressão	()	Tarefa não supervisionada que permite descobrir relações entre os dados.
(5) Sumarização	()	Tarefa descritiva capaz de organizar dados (sem rótulos) em subgrupos de acordo com alguma medida de similaridade.
4. **(0,5 pt)** Analise as afirmativas abaixo sobre algoritmos de agrupamento e assinale V para afirmativa verdadeira e F para Falsa.
 - a) () Os algoritmos particionais dividem o conjunto de dados em subconjuntos (grupos) disjuntos.
 - b) () Nos algoritmos hierárquicos, os subconjuntos (grupos) são aninhados na forma de uma árvore.
 - c) () Os algoritmos hierárquicos exigem que seja informado o número de grupos desejados previamente.
 - d) () Os algoritmos particionais podem ser construídos de forma aglomerativa ou por divisão.
 - e) () Average-link é um algoritmo hierárquico que se baseia na distância média entre dados dos grupos para organizá-los .
5. **(2,0 pts)** Os dados abaixo precisam ser agrupados, simule a execução de uma iteração do algoritmo $k - Means$, usando, para isso, a distância de Manhattan $d(x_i, x_j) = \sum_{l=1}^d |x_i^l - x_j^l|$. Ao final de sua simulação, indique como os dados foram agrupados e os valores dos centróides. Considere

$k = 2$ e que os dados 1 e 5 foram usados para inicializar os centróides (use apenas uma casa decimal em seus cálculos).

N.	Atributo1	Atributo2	Atributo3
1	2	3	5
2	1	2	9
3	0	1	1
4	2	1	2
5	1	0	1

6. **(2,0 pts)** Explique a lógica do algoritmo k -NN e indique a classe da nova amostra considerando $k = 1$ e $k = 3$. Também use a distância de Manhattan para definir os vizinhos.

N.	Atributo1	Atributo2	Classe
1	1	1	A
2	3	3	A
3	0	2	A
4	4	9	B
5	5	6	B
Nova Amostra	3	5	?

7. **(2,0 pts)** Considerando que existe uma coleção com 1000 textos jornalísticos em Língua Portuguesa. Desses textos, 500 são da Seção de Esporte, 170 são da Seção de Saúde, 30 são da Seção de Turismo e os 300 restantes, da Seção Economia. Sabendo que a tarefa a ser realizada é de categorização de textos, analise as afirmações abaixo e assinale V para afirmação verdadeira e F para falsa. Justifique as afirmativas falsas.
- ☐ O desbalanceamento dos dados pode afetar o desempenho do algoritmo usado para categorização.
 - ☐ Tanto uma rede Multi-Layer Perceptron (MLP) quanto o algoritmo k -NN poderiam ser usados para essa tarefa.
 - ☐ O algoritmo usado teria 1000 atributos de entrada, pois são 1000 textos jornalísticos.
 - ☐ O algoritmo usado teria 4 atributos de entrada, pois os textos são de Esporte, Saúde, Turismo e Economia.
 - ☐ A rede MLP, se usada, poderia ter 2 ou 4 neurônios na camada de saída.
 - ☐ Se 80% dos textos fossem usados como conjunto de treino, uma época em uma MLP, corresponderia a 800 iterações visto que o conjunto de dados é formado por 1000 textos.
 - ☐ O modelo indutivo aprendido para esta tarefa por uma MLP, através do algoritmo error-backpropagation, será armazenado nos pesos sinápticos dessa rede.
 - ☐ Uma das maneiras de estruturar os textos para viabilizar o uso dos algoritmos de Aprendizagem de Máquina é a técnica de bag-of-words.
 - ☐ Nesta tarefa, os atributos usados para representar o conjunto de treino podem ser diferentes daqueles usados para o conjunto de teste.
 - ☐ A rede neural Perceptron ou o algoritmo Support Vector Machine podem ser usados como classificadores nessa tarefa.