

	Terceira Avaliação		Nota:
	Curso:	Ciência da Computação	9,5
	Disciplina:	Introdução a Mineração de Dados	
	Aluno(a):	[Redacted]	
			Data: 22/07/2025

1) Marque V (verdadeira) ou F (falsa) para cada uma das afirmativas dadas: (1 pt)

- ☒ a) (F) Selecionar atributos sempre irá melhorar a performance de um algoritmo.
- ☒ b) (V) A seleção de atributos pode melhorar a performance de todo o processo de descoberta de conhecimento, por exemplo, na etapa de extração de dados.
- ☒ c) (F) A seleção de atributos deve ser executada apenas quando temos problemas de performance.
- ☒ d) (V) A seleção de atributos gera um subconjunto de atributos que são mais relevantes.

2) Marque V (verdadeira) ou F (falsa) para cada uma das afirmativas dadas: (1 pt)

- ☒ a) (V) Filtros supervisionados levam em consideração as classes da relação.
- ☒ b) (F) Filtros são mecanismos que podem aplicar outros algoritmos de classificação.
- ☒ c) (V) Filtros aplicados a partir das colunas são os filtros de atributos.

3) Qual a palavra que melhor define a diferença entre Naive Bayes e Redes Bayesianas. Selecione apenas uma alternativa. (1 pt)

- ☒ a) (X) Independência.
- ☐ b) ( ) Probabilidade.
- ☐ c) ( ) Sequência.

4) Marque V (verdadeira) ou F (falsa) para cada uma das afirmativas dadas: (3 pts)

- ☒ a) (V) A seleção de atributos pode atuar de duas formas: selecionando um subconjunto dos atributos mais relevantes de uma relação ou criando novos atributos gerados a partir de atributos originais da relação.
- ☒ b) (V) O método "Random Forest" induz várias árvores de decisão e realiza um processo de avaliação de desempenho nestas.
- ☒ c) (V) As técnicas de seleção de atributos são úteis para reduzir a dimensionalidade aumentando as chances de se obter um modelo genérico e não superajustado.

5) Discorra sobre a avaliação de desempenho incluindo os tópicos: relevância e, caso seja possível, como pode ser realizada no Weka. (2 pts)

6) Discorra sobre a ferramenta Knowledgeflow do Weka. (2 pts)



5. A avaliação de desempenho tem sua relevância dada por poder medir a eficácia de modelos preditivos. É importante para comparar modelos, evitar overfitting e garantir uma boa generalização. Sendo feita no Weka utilizando a aba classify e escolhendo o algoritmo que será usado e aplicar a validação cruzada. 2+ (1,5)

6. Knowledgeflow é uma interface gráfica do Weka que permite criar fluxos de processamento de dados por meio do drag'n drop. Funcionando como alternativa na aba Explorer do programa. Sendo útil para automatizar experimentos. C