

Bacharelado em Ciência da Computação

Disciplina: DCE6634 - Inteligência Artificial

Valor: 10 pontos (peso 0.25)

Data: 15/07/2025

Professor: Rodrigo Martins Pagliares



Gabarito: Agradeço se você puder repassar suas respostas para quadro abaixo. Isso facilita durante minha correção.

1	2	3	4	5	6	7	8	9	10
A	B	A	B	C	B	B	B	A	A
11	12	13	14	15	16	17	18	19	20
C	B	C	A	C	A	B	B	B	A

01. Assinale a alternativa CORRETA sobre aprendizado *online* e sua aplicação prática:

- a) Permite que o modelo aprenda continuamente à medida que novos dados chegam, sem necessidade de reprocessar todo o conjunto de dados.
- b) Requer que todos os dados sejam carregados na memória para atualização incremental. ×
- c) É incompatível com modelos supervisionados, pois depende exclusivamente de dados não rotulados. ×
- d) É utilizado apenas em problemas de classificação binária em tempo real. ×
- e) Não permite a atualização de parâmetros já aprendidos pelo modelo. ×

02. Assinale a alternativa CORRETA sobre aprendizado fora do núcleo (*out-of-core learning*):

- a) É utilizado quando o conjunto de dados é pequeno e cabe facilmente na memória RAM. ×
- b) Divide o conjunto de dados em lotes menores para processar sequencialmente, evitando sobrecarga de memória.
- c) Exige armazenamento dos dados exclusivamente em memória volátil. ×
- d) É restrito a modelos não supervisionados devido à limitação de espaço. ×
- e) Faz uso de redes neurais convolucionais. ×

03. Assinale a alternativa CORRETA sobre tarefas comuns de aprendizado supervisionado e não supervisionado:

- a) Classificação e regressão são exemplos de aprendizado *supervisionado*.
- b) Agrupamento e detecção de anomalias são exemplos de aprendizado *supervisionado*.
- c) Redução de dimensionalidade é uma tarefa supervisionada típica.
- d) Classificação hierárquica e análise discriminante são tarefas não supervisionadas.
- e) Aprendizado supervisionado não pode ser usado para prever valores numéricos.

04. Assinale a alternativa CORRETA sobre o impacto da escolha do algoritmo nos resultados do modelo:

- a) Algoritmos baseados em instâncias normalmente superam algoritmos baseados em modelos em termos de generalização.
- b) Algoritmos baseados em modelos criam representações matemáticas dos padrões nos dados e usam essas representações para prever novas instâncias.
- c) Algoritmos baseados em instâncias descartam todos os exemplos após o treinamento.
- d) Algoritmos baseados em modelos não podem ser aplicados a dados contínuos.
- e) A escolha do algoritmo é irrelevante quando os dados são rotulados.

**05. Assinale a alternativa CORRETA sobre o papel da validação na escolha de hiperparâmetros:**

- a) O conjunto de validação é usado para avaliar o modelo final antes da implantação.
- b) Ajustar hiperparâmetros usando o conjunto de teste garante melhor generalização.
- c) A validação cruzada pode ser usada para estimar o desempenho com diferentes configurações de hiperparâmetros.
- d) A escolha de hiperparâmetros é feita exclusivamente durante o pré-processamento dos dados.
- e) Não é necessário conjunto de validação se o conjunto de treinamento for grande o suficiente.

**06. Assinale a alternativa CORRETA sobre o conceito de aprendizado no contexto de inteligência artificial:**

- a) É o processo de armazenar dados em um banco de dados sem realizar nenhuma operação sobre eles.
- b) É a capacidade de um sistema adquirir conhecimento e reproduzi-lo em situações futuras para auxiliar na tomada de decisão.
- c) É a simples execução de instruções pré-programadas sem adaptação ao ambiente.
- d) É a transferência direta do conhecimento humano para máquinas por meio de programação procedural.
- e) É o uso de funções determinísticas para garantir que a saída seja sempre idêntica à entrada.

**07. Assinale a alternativa CORRETA sobre o papel do “supervisor” em aprendizado supervisionado:**

- a) Fornecer apenas os dados de entrada, deixando o modelo criar suas próprias classes sem referência.
- b) Avaliar as saídas do modelo, comparar com os valores esperados e indicar ajustes para reduzir o erro.
- c) Criar novos atributos para os dados sempre que houver baixa precisão.
- d) Controlar exclusivamente a taxa de aprendizado do modelo durante o treino.
- e) Interromper o treinamento assim que o modelo atingir 100% de acurácia nos dados de teste.

**08. Assinale a alternativa CORRETA sobre modelos baseados em árvores:**

- a) São compostos exclusivamente por uma única árvore de decisão.
- b) Incluem abordagens como ID3, C4.5 e ensembles de árvores, como Random Forest.
- c) Não podem ser combinados com outros modelos supervisionados.
- d) São sempre mais precisos do que redes neurais.
- e) São aplicáveis apenas a problemas não supervisionados.

**09. Assinale a alternativa CORRETA sobre as vantagens e desvantagens do KNN:**

- a) É simples, flexível e fácil de implementar, mas pode ter custo computacional alto e ser sensível a ruídos e ao número de vizinhos escolhidos.
- b) É complexo de implementar e exige ajustes matemáticos avançados para qualquer aplicação.
- c) Não é afetado pelo aumento do número de instâncias na base de dados.
- d) Sua principal desvantagem é que não pode ser usado para classificação.
- e) É exclusivamente aplicável a problemas de regressão linear.

**10.** No exemplo prático apresentado na vídeo aula do professor Ricardo, cobrada como exercício para casa, qual foi a principal função da biblioteca `scikit-learn` no contexto do algoritmo K-means?

- a) Fornecer ferramentas para visualização tridimensional dos dados
- b) Implementar uma versão otimizada do algoritmo K-means
- c) Gerar dados sintéticos para o experimento
- d) Normalizar e padronizar os dados de entrada
- e) Fornecer funções para leitura de arquivos CSV

**11.** Assinale a alternativa CORRETA sobre a relação entre tipos de algoritmos de aprendizado e os conjuntos de dados utilizados:

- a) Algoritmos supervisionados requerem apenas um conjunto de teste para aprender padrões nos dados.  $\times$
- b) Algoritmos não supervisionados utilizam rótulos no conjunto de treinamento para ajustar hiperparâmetros.  $\times$
- c) Algoritmos supervisionados precisam de dados rotulados para aprender a mapear entradas para saídas.
- d) Algoritmos online são incompatíveis com aprendizado supervisionado devido à natureza sequencial dos dados.  $\times$
- e) O aprendizado fora do núcleo exige que todos os dados estejam disponíveis na memória principal.  $\times$

**12.** Assinale a alternativa CORRETA sobre estratégias para lidar com problemas de generalização ruim em aprendizado de máquina:

- a) Reduzir o tamanho do conjunto de teste para aumentar a precisão medida do modelo.
- b) Coletar mais dados relevantes ou aplicar regularização para reduzir o overfitting.
- c) Aumentar o número de parâmetros do modelo para memorizar os exemplos de treinamento.
- d) Ajustar hiperparâmetros com base exclusivamente no desempenho no conjunto de teste.
- e) Utilizar algoritmos não supervisionados para melhorar modelos supervisionados sem alterar os dados.

**13.** Assinale a alternativa CORRETA sobre a diferença entre parâmetros e hiperparâmetros de um modelo:

- a) Parâmetros são definidos antes do treinamento e hiperparâmetros são ajustados pelo algoritmo durante o treino.
- b) Hiperparâmetros são aprendidos automaticamente a partir dos dados de treinamento.
- c) Parâmetros são aprendidos pelo algoritmo durante o treino, enquanto hiperparâmetros são definidos antes do treinamento.
- d) Ambos são ajustados automaticamente pelo algoritmo durante o processo de validação cruzada.
- e) Apenas modelos não supervisionados possuem hiperparâmetros.

**14.** Assinale a alternativa CORRETA sobre redução de dimensionalidade:

- a) É usada para simplificar os dados mantendo informações relevantes.
- b) É utilizada para aumentar a quantidade de características irrelevantes.
- c) É aplicável apenas a dados de imagem.
- d) Não pode ser combinada com algoritmos de classificação.
- e) É exclusiva do aprendizado supervisionado.

**15. Assinale a alternativa CORRETA sobre aplicações de aprendizado de máquina e escolha de algoritmos:**

- a) Para detectar *spam* em *e-mails*, é mais adequado utilizar um algoritmo de agrupamento não supervisionado.
- b) A segmentação de clientes em grupos homogêneos é tipicamente realizada por algoritmos supervisionados.
- c) A navegação autônoma de um robô em terrenos desconhecidos pode ser resolvida por aprendizado por reforço.
- d) A regressão linear é a abordagem mais indicada para reconhecimento de padrões visuais complexos.
- e) Para prever categorias de texto, utiliza-se exclusivamente redes neurais convolucionais.

**16. Assinale a alternativa CORRETA sobre impacto de dados ruidosos:**

- a) Dados ruidosos podem comprometer a capacidade do modelo de identificar padrões reais.
- b) Dados ruidosos sempre aumentam a precisão no conjunto de teste.
- c) Ruído nos dados de treinamento é irrelevante para modelos baseados em instâncias.
- d) Apenas modelos não supervisionados são afetados por ruído.
- e) Ruído afeta exclusivamente dados categóricos.

**17. Assinale a alternativa CORRETA sobre medidas de similaridade:**

- a) São usadas apenas em aprendizado supervisionado para comparar previsões com rótulos reais.
- b) Permitem que algoritmos baseados em instâncias classifiquem novos exemplos com base nos mais próximos no espaço de características.
- c) São exclusivas de algoritmos não supervisionados de agrupamento hierárquico.
- d) São irrelevantes para métodos de regressão.
- e) São aplicáveis somente a dados textuais.

**18. Assinale a alternativa CORRETA sobre desafios comuns em aprendizado de máquina:**

- a) O *overfitting* pode ser resolvido apenas aumentando a complexidade do modelo.
- b) Dados desbalanceados podem prejudicar o desempenho de classificadores.
- c) O *underfitting* ocorre quando o modelo se ajusta excessivamente aos dados de treinamento.
- d) A ausência de ruído nos dados compromete a generalização.
- e) Dados redundantes nunca afetam o desempenho do modelo.

**19. Assinale a alternativa CORRETA sobre a deteção de spam como problema de aprendizado:**

- a) É um problema de aprendizado não supervisionado por natureza.
- b) Pode ser resolvido com classificação supervisionada usando *e-mails* rotulados como *spam* ou não *spam*.
- c) Depende exclusivamente de agrupamento para identificar padrões nos textos.
- d) Não requer pré-processamento dos dados textuais.
- e) É inviável utilizar aprendizado supervisionado nesse contexto.

20. Assinale a alternativa CORRETA sobre aprendizado não supervisionado:

- a) Trabalha com dados não rotulados para encontrar padrões e estruturas ocultas.
- b) Necessita de pares entrada-saída para aprender.
- c) É usado exclusivamente para prever valores numéricicos.
- d) Não pode ser aplicado a dados de alta dimensionalidade.
- e) Depende unicamente de redes neurais convolucionais.

Boa prova!

Rodrigo Martins Pagliares