



Universidade Eduardo Mondlane  
Faculdade de Engenharia  
Departamento de Engenharia Electrotécnica

Teste 2 de Inteligência Artificial

Curso de Licenciatura em Engenharia Informática

(120 min)

Nome: \_\_\_\_\_ Turno: \_\_\_\_\_ Classificação: \_\_\_\_\_

ASSINALE A ALTERNATIVA CORRECTA

1. Qual biblioteca Python é comumente usada para visualização de dados?  
a) NumPy. c) **Matplotlib.**  
b) SciPy. d) Pandas.
2. Qual é o tipo de gráfico mais adequado para visualizar a distribuição de uma variável numérica contínua?  
a) Gráfico de barras.  
b) Gráfico de dispersão.  
c) **Histograma.**  
d) Gráfico de linhas.
3. Como criar um gráfico de dispersão usando a biblioteca Matplotlib?  
a) Utilizando a função `plot()`.  
b) **Utilizando a função `scatter()`.**  
c) Utilizando a função `bar()`.  
d) Utilizando a função `hist()`.
4. Qual é a finalidade de um gráfico de barras?  
a) Visualizar a distribuição de uma variável contínua.  
b) Comparar a relação entre duas variáveis contínuas.  
c) **Comparar a frequência de categorias em uma variável categórica.**  
d) Visualizar a tendência temporal de uma variável.
5. Como personalizar um gráfico de barras com a biblioteca Matplotlib?  
a) Alterando as cores das barras.  
b) Adicionando um título ao gráfico.  
c) Definindo rótulos para os eixos x e y.  
d) **Todas as respostas anteriores estão corretas.**
6. Qual é o objetivo principal da aprendizagem de máquina?  
a) Automatizar tarefas repetitivas.  
b) Realizar análises estatísticas avançadas.  
c) **Extrair informações úteis a partir de dados e fazer previsões.**  
d) Desenvolver algoritmos mais eficientes.
7. Quais são os dois principais tipos de aprendizagem de máquina?  
a) **Aprendizado supervisionado e aprendizado não supervisionado.**  
b) Aprendizado ativo e aprendizado por reforço.  
c) Aprendizado online e aprendizado em lote.  
d) Aprendizado por transferência e aprendizado profundo.
8. O que é um modelo de aprendizado de máquina?  
a) Uma representação matemática de um conjunto de dados.  
b) Um conjunto de regras e instruções para treinar um algoritmo.  
c) **Uma função matemática que mapeia entradas para saídas.**  
d) Uma técnica estatística para analisar dados.
9. Como avaliar a qualidade de um modelo de aprendizado de máquina?  
a) Calculando a acurácia do modelo em dados de treinamento.  
b) **Dividindo o conjunto de dados em dados de treinamento e teste e calculando métricas de desempenho, como acurácia ou erro médio.**  
c) Comparando o modelo com outros modelos disponíveis na biblioteca Python.  
d) Realizando validação cruzada com diferentes conjuntos de hiperparâmetros.
10. Qual é a diferença entre overfitting e underfitting em modelos de aprendizado de máquina?  
a) Overfitting ocorre quando o modelo é muito simples e underfitting ocorre quando o modelo é muito complexo.  
b) **Overfitting ocorre quando o modelo se ajusta demasiadamente aos dados de treinamento e perde generalização, enquanto underfitting ocorre quando o modelo é muito simples e não consegue capturar padrões nos dados.**  
c) Overfitting e underfitting referem-se ao mesmo problema de um modelo sendo incapaz de se ajustar aos dados.  
d) Overfitting ocorre quando o modelo não é treinado corretamente e produz previsões imprecisas, enquanto underfitting ocorre quando o modelo é muito complexo e requer mais recursos computacionais.
11. Qual é a biblioteca Python mais comumente usada para processamento de texto?  
a) NumPy. c) Matplotlib  
b) SciPy. d) **NLTK (Natural Language Toolkit).**
12. O que é tokenização de texto?  
a) A remoção de palavras irrelevantes de um texto.  
b) A transformação de um texto em um formato numérico.  
c) **O processo de dividir um texto em unidades menores, como palavras ou frases.**  
d) A substituição de palavras em um texto por sinónimos.
13. Como realizar a lematização de palavras em um texto usando a biblioteca NLTK?  
a) Utilizando a função `stem()` da classe `WordNetLemmatizer`.  
b) **Utilizando a função `lemmatize()` da classe `WordNetLemmatizer`.**  
c) Utilizando a função `tokenize()` da classe `WordNetLemmatizer`.  
d) Utilizando a função `normalize()` da classe `WordNetLemmatizer`.
14. O que é um n-grama em processamento de texto?  
a) Um modelo de linguagem baseado em redes neurais.  
b) Um método para identificar entidades nomeadas em um texto.  
c) **Uma sequência contígua de n palavras em um texto.**  
d) A etapa de pré-processamento de texto em que as palavras são transformadas em vetores numéricos.
15. Qual é a técnica comumente usada para representar palavras como vetores numéricos em processamento de texto?  
a) One-Hot Encoding.  
b) Bag-of-Words.  
c) TF-IDF.  
d) **Word2Vec.**
16. O que é Big Data?  
a) Conjuntos de dados pequenos que podem ser armazenados e processados em um único computador.  
b) **Conjuntos de dados que são grandes demais para serem armazenados em um único computador e exigem técnicas e infraestrutura especiais para armazenamento e processamento.**  
c) Conjuntos de dados que contêm informações confidenciais e requerem medidas de segurança adicionais.  
d) Conjuntos de dados que são difíceis de analisar devido à alta dimensionalidade.
17. Qual é a finalidade da análise distribuída?  
a) Dividir um conjunto de dados em partes menores para facilitar o armazenamento.  
b) Realizar análises estatísticas avançadas em conjuntos de dados grandes.  
c) **Dividir um problema de análise em várias tarefas menores e executá-las em paralelo em um cluster de computadores.**  
d) Transformar um conjunto de dados em um formato adequado para análise.
18. Qual biblioteca Python é comumente usada para análise distribuída de Big Data?  
a) Pandas.  
b) NumPy.  
c) **Dask.**  
d) Matplotlib.
19. O que é um cluster de computadores?  
a) **Um conjunto de computadores conectados em rede que trabalham juntos para realizar tarefas complexas.**  
b) Um único computador com alta capacidade de armazenamento.  
c) Um computador especialmente projetado para processar grandes conjuntos de dados.  
d) Um servidor de banco de dados usado para armazenar Big Data.
20. Quais são os benefícios da análise distribuída em comparação com a análise em um único computador?  
a) Menor custo de infraestrutura.  
b) Maior velocidade de processamento.  
c) Capacidade de lidar com conjuntos de dados maiores.  
d) **Todas as respostas anteriores estão corretas.**