

COLETÂNEA CONCURSOS PÚBLICOS



GLOSSÁRIO DO EDITAL DO CONCURSO DO BACEN

**CONHECIMENTOS ESPECÍFICOS PARA O CARGO 1:
ANALISTA – ÁREA 1 – TECNOLOGIA DA INFORMAÇÃO**



**COM QUESTÕES INÉDITAS COMENTADAS
PÓS-EDITAL 2024**

APRESENTAÇÃO

A ideia central deste ebook preparatório para o concurso do BACEN é que o estudante possa aprender de forma rápida todo o conteúdo do edital, portanto, para um maior aprofundamento em cada assunto é necessário utilizar materiais complementares.

Este ebook foi elaborado com auxílio de inteligência artificial e revisado e complementado pelo professor Izaias Batista dos Santos.

Siga a Kuasarnex nas redes sociais para receber notificações sobre conteúdos e oportunidades.



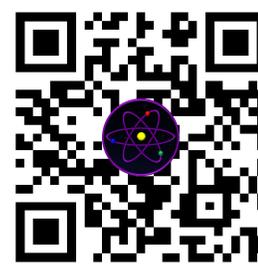
[Seguir no Instagram](#)



[Inscrição no Canal](#)



[Seguir no LinkedIn](#)



[Site Oficial](#)

AUTOR: IZAIAS BATISTA DOS SANTOS

Olá, seja muito bem-vindo(a)!

Sou Izaias Batista dos Santos, autor deste material preparatório especialmente desenvolvido para auxiliar você em sua jornada rumo à aprovação no concurso do BACEN. Tenho orgulho de compartilhar um pouco da minha trajetória com você, pois acredito que conhecer quem está por trás dos conteúdos reforça a confiança no estudo e nos resultados.

Sou mestre em **Tecnologias Computacionais para o Agronegócio** pela **Universidade Tecnológica Federal do Paraná (UTFPR)**, especialista em **Engenharia de Software** pela **Pontifícia Universidade Católica de Minas Gerais (PUC-Minas)**, possuo **MBA em Gerenciamento de Projetos de TI** pelo **Instituto de Gestão em Tecnologia da Informação (IGTI)** e sou bacharel em **Sistemas de Informação** pelo **Centro Universitário Dinâmica das Cataratas (UDC)**.

Minha trajetória profissional foi construída com dedicação e propósito em diversas áreas da Tecnologia da Informação. Atuei como **coordenador de TIC no Exército Brasileiro**, **analista de sistemas sênior** na **Fundação Parque Tecnológico Itaipu**, **coordenador de projetos de software** na empresa **Eits Prognus Software Livre** e também fui **professor substituto de informática e suas tecnologias** e **técnico em tecnologia da informação e comunicações** no **Instituto Federal do Paraná**.

Atualmente, sou **analista de sistemas e processos** no **Conselho Federal de Química**, onde atuo com **gestão por processos** e exerço a função de **gestor técnico de sistemas**, sempre buscando alinhar boas práticas, inovação e eficiência para garantir resultados sustentáveis e soluções de alto impacto.

Com base em minha experiência acadêmica e profissional, preparei este e-book para oferecer a você um conteúdo claro, atualizado e voltado aos temas mais cobrados nos concursos públicos da área de Tecnologia da Informação. Meu objetivo é ajudá-lo(a) a dominar os assuntos e encurtar o caminho até a aprovação.

E, nos momentos em que bater a dúvida sobre sua capacidade de aprender e vencer esse desafio, lembre-se sempre da poderosa frase:

“Tudo posso naquele que me fortalece!”

Ah, e caso queira acompanhar mais conteúdos, dicas e novidades sobre concursos públicos e Tecnologia da Informação, será um prazer ter você comigo nas redes sociais!

Siga-me e vamos juntos nessa jornada:



[Seguir no Instagram](#)



[Inscrição no Canal](#)



[Seguir no LinkedIn](#)

Estarei sempre compartilhando conhecimento, informações atualizadas e motivação para que você se mantenha firme no seu propósito. Será uma honra contar com sua presença por lá!

Bons estudos e sucesso em sua caminhada!

OUTROS MATERIAIS

Conheça outros materiais acessando o seguinte



[Acessar Conteúdos](#)

Conheça também os materiais gratuitos acessando o seguinte



[Material Gratuito](#)

CARGO 1: ANALISTA - ÁREA 1 - TECNOLOGIA DA INFORMAÇÃO

CAPÍTULO I - CIÊNCIA DE DADOS (14 Questões):

- 1 Aprendizado de Máquina.
- 2 Deep learning.
- 3 Processamento de linguagem natural.
- 4 Big data.
- 5 Qualidade de Dados.
- 6 Tipos de Aprendizado: Supervisionado, Não Supervisionado, Semi Supervisionado, Por Reforço, Por Transferência.
- 7 Grandes Modelos de Linguagem (LLM), IA Generativa.
- 8 Redes Neurais.
- 9 MLOps: Gestão de código, treinamento, implantação, monitoramento e versionamento de modelos, automação do ciclo de produção.
- 10 Governança e Ética na IA: Transparência, Responsabilidade, Explicabilidade, Privacidade, Segurança, Viés.

CAPÍTULO II - SEGURANÇA DA INFORMAÇÃO (7 Questões):

- 1 Gestão de Identidades e Acesso: Autenticação e Autorização, Single Sign-On (SSO), Security Assertion Markup Language (SAML), OAuth2 e OpenId Connect.
- 2 Privacidade e segurança por padrão.
- 3 Principais tipos de ataques e vulnerabilidades.
- 4 Controles e testes de segurança para aplicações Web e Web Services.
- 5 Múltiplos Fatores de Autenticação (MFA).
- 6 Soluções para Segurança da Informação: Firewall, Intrusion Detection System (IDS), Intrusion Prevention System (IPS), Security Information and Event Management (SIEM), Proxy, Identity Access Management (IAM), Privileged Access Management (PAM), Antivírus, Antispam.
- 7 Frameworks de segurança da informação e segurança cibernética, como, por exemplo, MITRE ATT&CK, CIS Controls e NIST CyberSecurity Framework (NIST CSF).
- 8 Tratamento de Incidentes Cibernéticos.
- 9 Assinatura e certificação digital, criptografia e proteção de dados em trânsito e em repouso.
- 10 Segurança em nuvens e de contêineres.

CAPÍTULO III - ENGENHARIA DE SOFTWARE (24 Questões):

- 1 Arquitetura de sistemas web: protocolo HTTP, HTTP/2, gRPC, WebSockets, TLS, servidores proxy, cache, DNS, balanceamento de carga, tolerância a falhas e escalabilidade em sistemas web.
- 2 Princípios e práticas de DevOps e DevSecOps, integração e entrega contínuas (CI/CD).
- 3 Técnicas de desenvolvimento seguro.
- 4 Testes de software: Testes unitários, Testes de Integração, TDD, BDD.
- 5 Arquiteturas em camadas, baseada em serviços, microserviços (orquestração de serviços e API gateway), orientação a eventos, cliente-servidor, serverless.
- 6 Práticas de UX e UI design.
- 7 Programação assíncrona.

- 8 RESTful e GraphQL.
- 9 Web services.
- 10 Padrões: GoF e GRASP.
- 11 Git.
- 12 Python e Java.
- 13 Transações distribuídas.
- 14 Distributed Ledger Technology (DLT).

CAPÍTULO IV - INFRAESTRUTURA EM TI (17 Questões):

- 1 Conceitos de infraestrutura como código e automação de infraestrutura de TI.
- 2 Docker, Kubernetes: boas práticas para infraestrutura e orquestração de containers.
- 3 Serviços de Rede Microsoft Windows Server: DNS, DHCP, Radius, Autenticação, Certificados, Active Directory (AD).
- 4 Monitoração, observabilidade e análise de sistemas em produção por meio do uso de ferramentas de monitoramento e logging, como o Nagios, Prometheus, Grafana, Elasticsearch, Kibana, Application Performance Monitoring (APM), entre outras.
- 5 Protocolos da camada de aplicação, como Simple Mail Transfer Protocol (SMTP), Hypertext Transfer Protocol (HTTP), Hypertext Transfer Protocol (HTTPS), SSL/TLS, Lightweight Directory Access Protocol (LDAP), Network File System (NFS), Server Message Block (SMB).
- 6 Tolerância a falhas e continuidade de operação.
- 7 Implantação e administração de serviços de nuvem (IaaS, PaaS e SaaS).
- 8 Administração e gerenciamento de ambientes de virtualização.
- 9 Administração e operação de sistemas operacionais Microsoft Windows Server e Linux.
- 10 Conceitos de LAN, WAN e SDN.
- 11 Conceitos e ferramentas de orquestração e automação de infraestrutura: Puppet, Ansible.

CAPÍTULO V - BANCOS DE DADOS (4 Questões):

- 1 SGBDs SQL e NOSQL.
- 2 Modelagens de dados: relacional, multidimensional, nosql.
- 3 SQL (Procedural Language / Structured Query Language).
- 4 Arquitetura de Inteligência de Negócio: DataWarehouse, DataMart, DataLake, DataMesh.

CAPÍTULO VI - GESTÃO EM TI (4 Questões):

- 1 Kanban.
- 2 Scrum.
- 3 Governança de Dados.
- 4 ITIL 4.

CAPÍTULO I - CIÊNCIA DE DADOS

Para este capítulo serão cobradas 14 Questões relacionadas aos tópicos seguintes

- 1 Aprendizado de Máquina.
- 2 Deep learning.
- 3 Processamento de linguagem natural.
- 4 Big data.
- 5 Qualidade de Dados.
- 6 Tipos de Aprendizado: Supervisionado, Não Supervisionado, Semi Supervisionado, Por Reforço, Por Transferência.
- 7 Grandes Modelos de Linguagem (LLM), IA Generativa.
- 8 Redes Neurais.
- 9 MLOps: Gestão de código, treinamento, implantação, monitoramento e versionamento de modelos, automação do ciclo de produção.
- 10 Governança e Ética na IA: Transparência, Responsabilidade, Explicabilidade, Privacidade, Segurança, Viés.

1 Aprendizado de Máquina

1. Definição do Tema

Aprendizado de Máquina (Machine Learning - ML) é um subcampo da inteligência artificial que envolve o desenvolvimento de algoritmos e modelos estatísticos que permitem que os sistemas de computador aprendam e melhorem suas performances em tarefas específicas com base em dados. Em vez de serem explicitamente programados para executar uma tarefa, os modelos de aprendizado de máquina usam padrões e inferências para fazer previsões ou tomar decisões.

2. Exemplos de Aplicação

- **Reconhecimento de Imagens e Visão Computacional:** Algoritmos de aprendizado de máquina são usados em sistemas de reconhecimento facial, diagnósticos médicos a partir de imagens (como a detecção de câncer em radiografias), e em veículos autônomos para interpretar o ambiente ao redor.
- **Processamento de Linguagem Natural (PLN):** Técnicas de ML são utilizadas em chatbots, assistentes virtuais (como Alexa e Siri), tradução automática e análise de sentimentos em textos.
- **Sistemas de Recomendação:** Plataformas como Netflix, Amazon e Spotify usam modelos de aprendizado de máquina para sugerir filmes, produtos ou músicas com base no histórico e nas preferências dos usuários.
- **Finanças e Detecção de Fraudes:** Bancos e empresas financeiras utilizam algoritmos de ML para detectar atividades fraudulentas, avaliar riscos de crédito e otimizar carteiras de investimentos.
- **Manutenção Preditiva:** Na indústria, modelos de aprendizado de máquina são empregados para prever falhas em máquinas e equipamentos, ajudando a realizar manutenção preventiva e reduzir o tempo de inatividade.
- **Saúde Personalizada:** O ML é utilizado para criar planos de tratamento personalizados com base nos dados médicos do paciente, prever surtos de doenças e desenvolver novos medicamentos.

3. Tecnologias Relacionadas

- **Linguagens de Programação:** Python e R são amplamente usadas devido às suas bibliotecas e frameworks robustos para aprendizado de máquina, como TensorFlow, PyTorch, Scikit-learn e Keras.
- **Bibliotecas e Frameworks:**
 - **TensorFlow:** Uma biblioteca open-source de ML desenvolvida pelo Google que é amplamente utilizada para criar e treinar modelos de aprendizado profundo (Deep Learning).
 - **PyTorch:** Desenvolvida pelo Facebook, esta biblioteca também é usada para aprendizado profundo e é conhecida por sua flexibilidade e facilidade de uso.
 - **Scikit-learn:** Uma biblioteca em Python para aprendizado de máquina tradicional que oferece uma vasta gama de algoritmos e ferramentas para modelagem estatística.

- **Keras:** Uma API de alto nível para redes neurais, que pode ser executada sobre TensorFlow e é conhecida por sua simplicidade e rapidez no desenvolvimento de protótipos.
- **Plataformas e Ferramentas de Implementação:**
 - **Google Cloud ML Engine:** Serviço gerenciado de aprendizado de máquina que facilita a construção e o deployment de modelos.
 - **AWS SageMaker:** Plataforma da Amazon Web Services para desenvolvimento, treinamento e implantação de modelos de ML.
 - **Microsoft Azure Machine Learning:** Serviço de nuvem da Microsoft que oferece ferramentas para criar, treinar e implantar modelos de ML.
- **Ambientes de Desenvolvimento Integrado (IDEs):**
 - **Jupyter Notebook:** Um ambiente amplamente utilizado para programação em Python e R, permitindo a criação de documentos que contêm código, visualizações e explicações textuais.
 - **Google Colab:** Uma versão online do Jupyter Notebook que permite a execução de código em GPUs gratuitamente.

O aprendizado de máquina está em constante evolução, impulsionado pelo crescimento exponencial de dados e avanços em poder de computação. Ele está transformando indústrias e criando novas oportunidades para inovação e eficiência.

Estilo Cebraspe - Questões Certo ou Errado sobre Aprendizado de Máquina

Questão 1

Estilo Cebraspe - Questão inédita

O aprendizado supervisionado é um tipo de aprendizado de máquina em que o modelo é treinado com um conjunto de dados rotulados, ou seja, cada exemplo de treinamento é composto por uma entrada e a saída desejada correspondente.

Comentário:

No aprendizado supervisionado, o objetivo é aprender uma função que mapeie uma entrada para uma saída com base em exemplos de entrada-saída fornecidos. Exemplos de algoritmos supervisionados incluem regressão linear, regressão logística, máquinas de vetor de suporte (SVM) e redes neurais.

Resposta: Certo

Questão 2

Estilo Cebraspe - Questão inédita

No aprendizado não supervisionado, os algoritmos de aprendizado de máquina são usados para encontrar padrões ou estruturas nos dados que não possuem rótulos ou categorias predefinidas.

Comentário:

O aprendizado não supervisionado se concentra em explorar a estrutura dos dados sem a necessidade de rótulos de saída. Algoritmos comuns incluem clustering (como K-means) e métodos de redução de dimensionalidade (como PCA - Análise de Componentes Principais).

Resposta: Certo

Questão 3

Estilo Cebraspe - Questão inédita

Em uma rede neural, a função de ativação ReLU (Rectified Linear Unit) é frequentemente utilizada devido à sua capacidade de introduzir não-linearidades no modelo, o que é essencial para a aprendizagem de padrões complexos.

Comentário:

A função de ativação ReLU é definida como $f(x) = \max(0, x)$, e é popular em redes neurais profundas devido à sua eficiência computacional e capacidade de mitigar o problema de gradientes desaparecendo, comum em outras funções de ativação como sigmoid ou tanh.

Resposta: Certo

Questão 4

Estilo Cebraspe - Questão inédita

A validação cruzada (cross-validation) é uma técnica utilizada para avaliar a performance de um modelo de aprendizado de máquina dividindo o conjunto de dados em várias partes, onde cada parte é utilizada para treinamento e validação em diferentes iterações.

Comentário:

A validação cruzada, especialmente a K-fold cross-validation, ajuda a garantir que o modelo generalize bem para dados não vistos, dividindo o conjunto de dados em K partes (folds) e usando cada uma delas para validação enquanto as demais são usadas para treinamento.

Resposta: Certo

Questão 5

Estilo Cebraspe - Questão inédita

O overfitting ocorre quando um modelo de aprendizado de máquina se ajusta muito bem aos dados de treinamento, mas falha em generalizar para novos dados, indicando que o modelo é demasiadamente complexo para o problema em questão.

Comentário:

Overfitting é um problema comum em aprendizado de máquina, onde o modelo captura ruído e peculiaridades dos dados de treinamento em vez de aprender padrões subjacentes. Técnicas como regularização, poda de árvores de decisão e aumento do conjunto de dados podem ajudar a mitigar o overfitting.

Resposta: Certo

Questão 6

Estilo Cebraspe - Questão inédita

Algoritmos de aprendizado por reforço (reinforcement learning) são utilizados para resolver problemas onde um agente toma ações em um ambiente com o objetivo de maximizar uma recompensa cumulativa, sem a necessidade de um conjunto de dados rotulados previamente.

Comentário:

No aprendizado por reforço, o agente aprende interagindo com o ambiente e recebendo recompensas ou punições com base em suas ações. Este tipo de aprendizado é amplamente utilizado em problemas de controle e jogos, como AlphaGo e controle de robôs.

Resposta: Certo

Questão 7

Estilo Cebraspe - Questão inédita

O aprendizado supervisionado é um tipo de aprendizado de máquina em que o modelo é treinado com um conjunto de dados rotulados, ou seja, cada exemplo de treinamento é composto por uma entrada e a saída desejada correspondente.

Comentário:

No aprendizado supervisionado, o objetivo é aprender uma função que mapeie uma entrada para uma saída com base em exemplos de entrada-saída fornecidos. Exemplos de algoritmos supervisionados incluem regressão linear, regressão logística, máquinas de vetor de suporte (SVM) e redes neurais.

Resposta: Certo

Questão 8

Estilo Cebraspe - Questão inédita

No aprendizado não supervisionado, os algoritmos de aprendizado de máquina são usados para encontrar padrões ou estruturas nos dados que não possuem rótulos ou categorias predefinidas.

Comentário:

O aprendizado não supervisionado se concentra em explorar a estrutura dos dados sem a necessidade de rótulos de saída. Algoritmos comuns incluem clustering (como K-means) e métodos de redução de dimensionalidade (como PCA - Análise de Componentes Principais).

Resposta: Certo

Questão 9

Estilo Cebraspe - Questão inédita

Em uma rede neural, a função de ativação ReLU (Rectified Linear Unit) é frequentemente utilizada devido à sua capacidade de introduzir não-linearidades no modelo, o que é essencial para a aprendizagem de padrões complexos.

Comentário:

A função de ativação ReLU é definida como $f(x) = \max(0, x)$, e é popular em redes neurais profundas devido à sua eficiência computacional e capacidade de mitigar o problema de gradientes desaparecendo, comum em outras funções de ativação como sigmoid ou tanh.

Resposta: Certo

Questão 10

Estilo Cebraspe - Questão inédita

A validação cruzada (cross-validation) é uma técnica utilizada para avaliar a performance de um modelo de aprendizado de máquina dividindo o conjunto de dados em várias partes, onde cada parte é utilizada para treinamento e validação em diferentes iterações.

Comentário:

A validação cruzada, especialmente a K-fold cross-validation, ajuda a garantir que o modelo generalize bem para dados não vistos, dividindo o conjunto de dados em K partes (folds) e usando cada uma delas para validação enquanto as demais são usadas para treinamento.

Resposta: Certo

Questão 11

Estilo Cebraspe - Questão inédita

O overfitting ocorre quando um modelo de aprendizado de máquina se ajusta muito bem aos dados de treinamento, mas falha em generalizar para novos dados, indicando que o modelo é demasiadamente complexo para o problema em questão.

Comentário:

Overfitting é um problema comum em aprendizado de máquina, onde o modelo captura ruído e peculiaridades dos dados de treinamento em vez de aprender padrões subjacentes. Técnicas como regularização, poda de árvores de decisão e aumento do conjunto de dados podem ajudar a mitigar o overfitting.

Resposta: Certo

Questão 12

Estilo Cebraspe - Questão inédita

Algoritmos de aprendizado de reforço (reinforcement learning) são utilizados para resolver problemas onde um agente toma ações em um ambiente com o objetivo de maximizar uma recompensa cumulativa, sem a necessidade de um conjunto de dados rotulados previamente.

Comentário:

No aprendizado de reforço, o agente aprende interagindo com o ambiente e recebendo recompensas ou punições com base em suas ações. Este tipo de aprendizado é amplamente utilizado em problemas de controle e jogos, como AlphaGo e controle de robôs.

Resposta: Certo

Questão 13

Estilo Cebraspe - Questão inédita

Redes neurais convolucionais (CNNs) são particularmente eficazes para tarefas de processamento de linguagem natural devido à sua capacidade de capturar a hierarquia espacial das características.

Comentário:

As CNNs são principalmente usadas em tarefas de visão computacional devido à sua capacidade de capturar padrões espaciais e hierárquicos em imagens. Para processamento de linguagem natural, arquiteturas como RNNs (Redes Neurais Recorrentes) e Transformers são mais adequadas.

Resposta: Errado

Questão 14

Estilo Cebraspe - Questão inédita

Redes neurais recorrentes (RNNs) são bem-sucedidas em tarefas sequenciais, como previsão de séries temporais e processamento de linguagem natural, devido à sua capacidade de manter informações sobre estados anteriores.

Comentário:

RNNs são projetadas para lidar com dados sequenciais, mantendo uma memória de estados anteriores, o que as torna eficazes em tarefas como tradução de idiomas, reconhecimento de fala e previsão de séries temporais. No entanto, elas podem sofrer de problemas como gradientes desaparecendo.

Resposta: Certo

Questão 15

Estilo Cebraspe - Questão inédita

As redes generativas adversariais (GANs) consistem em dois componentes principais: um gerador que tenta criar dados falsos e um discriminador que tenta distinguir entre dados reais e falsos.

Comentário:

GANs são compostas por duas redes neurais que competem entre si: o gerador cria amostras falsas na tentativa de enganar o discriminador, enquanto o discriminador tenta distinguir entre amostras reais e falsas. Este processo resulta em geradores que produzem dados altamente realistas.

Resposta: Certo

Questão 16

Estilo Cebraspe - Questão inédita

A regularização L2, também conhecida como regularização de Tikhonov, adiciona a soma dos quadrados dos pesos ao erro de treinamento, ajudando a reduzir o overfitting em redes neurais profundas.

Comentário:

A regularização L2 adiciona um termo de penalização proporcional ao quadrado da magnitude dos pesos ao objetivo de minimização, encorajando a rede a manter pesos menores e reduzindo a complexidade do modelo, o que ajuda a prevenir overfitting.

Resposta: Certo

Questão 17

Estilo Cebraspe - Questão inédita

A função de ativação sigmoide é amplamente utilizada em camadas ocultas de redes neurais profundas devido à sua eficiência na mitigação do problema de gradientes desaparecendo.

Comentário:

A função de ativação sigmoide pode exacerbar o problema de gradientes desaparecendo,

especialmente em redes profundas, pois ela tende a saturar e levar gradientes a valores muito próximos de zero. Funções como ReLU (Rectified Linear Unit) são mais eficazes em mitigar esse problema.

Resposta: Errado

Questão 18

Estilo Cebraspe - Questão inédita

O treinamento de redes neurais profundas geralmente requer grandes quantidades de dados e recursos computacionais significativos devido à complexidade dos modelos e à quantidade de parâmetros a serem ajustados.

Comentário:

Treinar redes neurais profundas é intensivo em dados e computação. Grandes volumes de dados são necessários para evitar overfitting e para que a rede aprenda representações robustas. Além disso, a quantidade de parâmetros nas redes profundas demanda considerável poder de processamento e memória.

Resposta: Certo

Questão 19

Estilo Cebraspe - Questão inédita

O uso de dropout durante o treinamento de uma rede neural é uma técnica que consiste em adicionar um ruído gaussiano aos pesos da rede para aumentar sua robustez e prevenir overfitting.

Comentário:

O dropout é uma técnica de regularização que funciona desligando aleatoriamente uma fração de neurônios durante o treinamento, forçando a rede a não depender excessivamente de unidades específicas e, assim, melhorando sua capacidade de generalização. Não se trata de adicionar ruído gaussiano aos pesos.

Resposta: Errado

Questão 20

Estilo Cebraspe - Questão inédita

A tokenização é uma etapa fundamental no processamento de linguagem natural, onde o texto é dividido em unidades menores como palavras, frases ou sub-palavras, facilitando o processamento subsequente.

Comentário:

A tokenização é essencial no PLN, pois transforma o texto em unidades menores (tokens) que podem ser palavras, sub-palavras ou até mesmo caracteres, permitindo que os algoritmos de PLN processem e analisem o texto de maneira eficiente.

Resposta: Certo

Questão 21

Estilo Cebraspe - Questão inédita

A lematização e a stemming são técnicas de normalização de texto que sempre retornam a mesma forma base para uma palavra, garantindo resultados consistentes.

Comentário:

Embora a lematização e a stemming sejam técnicas de normalização de texto, elas não garantem os mesmos resultados. A lematização considera o contexto e retorna a forma base correta (dicionário), enquanto o stemming simplesmente corta sufixos, podendo gerar formas não reais da palavra.

Resposta: Errado

Questão 22

Estilo Cebraspe - Questão inédita

O modelo Transformer revolucionou o campo do PLN ao introduzir a atenção multi-cabeça, permitindo que o modelo aprenda relações contextuais entre palavras em uma sequência de forma mais eficiente do que modelos baseados em RNN.

Comentário:

O modelo Transformer, introduzido por Vaswani et al. em 2017, utiliza mecanismos de atenção multi-cabeça para capturar dependências contextuais em todo o texto, melhorando significativamente a eficiência e o desempenho em várias tarefas de PLN em comparação com modelos baseados em RNN.

Resposta: Certo

Questão 23

Estilo Cebraspe - Questão inédita

A codificação one-hot é a técnica mais eficiente para representar palavras em um espaço vetorial, sendo amplamente utilizada em modelos modernos de PLN.

Comentário:

A codificação one-hot é uma técnica simples e direta para representar palavras, mas não é a mais

eficiente, especialmente para grandes vocabulários. Técnicas mais avançadas como word embeddings (Word2Vec, GloVe) e representações contextuais (BERT) são preferidas em modelos modernos de PLN.

Resposta: Errado

Questão 24

Estilo Cebraspe - Questão inédita

O modelo BERT (Bidirectional Encoder Representations from Transformers) é projetado para entender o contexto bidirecional de uma palavra em uma frase, o que melhora seu desempenho em várias tarefas de PLN.

Comentário:

BERT é um modelo baseado em Transformers que captura o contexto bidirecional de palavras em uma frase, treinando em grandes volumes de texto para melhorar a compreensão contextual e o desempenho em uma ampla gama de tarefas de PLN, como classificação de texto, resposta a perguntas e NER.

Resposta: Certo

Questão 25

Estilo Cebraspe - Questão inédita

No processamento de linguagem natural, a análise de sentimento é a tarefa de identificar entidades nomeadas em um texto, como pessoas, organizações e locais.

Comentário:

A análise de sentimento é a tarefa de determinar a polaridade das emoções expressas em um texto (positiva, negativa ou neutra). A tarefa de identificar entidades nomeadas é chamada de Reconhecimento de Entidades Nomeadas (NER).

Resposta: Errado

Questão 26

Estilo Cebraspe - Questão inédita

A técnica de bag-of-words (BoW) representa um texto ignorando a ordem das palavras e considerando apenas a frequência de ocorrência de cada palavra, o que pode levar à perda de informações contextuais importantes.

Comentário:

A técnica de bag-of-words (BoW) representa um texto como uma coleção de palavras sem considerar

a ordem ou contexto, apenas a frequência das palavras. Isso pode resultar em perda de informações contextuais, o que é uma limitação significativa dessa abordagem.

Resposta: Certo

Questão 27

Estilo Cebraspe - Questão inédita

O aprendizado supervisionado é caracterizado pela utilização de um conjunto de dados rotulados para treinar o modelo, onde cada entrada é associada a uma saída correspondente.

Comentário:

No aprendizado supervisionado, o modelo é treinado usando um conjunto de dados onde cada exemplo de entrada possui uma saída correspondente (rótulo). Exemplos incluem a regressão linear e a classificação de imagens.

Resposta: Certo

Questão 28

Estilo Cebraspe - Questão inédita

O aprendizado não supervisionado é utilizado quando o objetivo é identificar padrões ou estruturas nos dados sem a necessidade de rótulos pré-definidos, como é o caso dos algoritmos de clustering.

Comentário:

No aprendizado não supervisionado, o modelo busca encontrar padrões ou estruturas intrínsecas nos dados sem rótulos. Exemplos incluem o clustering (agrupamento), como K-means, e a análise de componentes principais (PCA).

Resposta: Certo

Questão 29

Estilo Cebraspe - Questão inédita

O aprendizado semi-supervisionado utiliza uma combinação de dados rotulados e não rotulados para treinar o modelo, aproveitando uma pequena quantidade de dados rotulados para melhorar a precisão do aprendizado.

Comentário:

O aprendizado semi-supervisionado combina uma pequena quantidade de dados rotulados com uma grande quantidade de dados não rotulados, aproveitando ambos para melhorar o desempenho do modelo. Este método é útil quando a rotulação de dados é custosa ou demorada.

Resposta: Certo

Questão 30

Estilo Cebraspe - Questão inédita

No aprendizado por reforço, um agente aprende a tomar decisões sequenciais em um ambiente interativo, recebendo recompensas ou punições com base nas ações executadas.

Comentário:

O aprendizado por reforço envolve um agente que aprende interagindo com o ambiente, tomando ações e recebendo feedback na forma de recompensas ou punições. O objetivo é maximizar a recompensa acumulada ao longo do tempo, como em jogos ou controle de robôs.

Resposta: Certo

Questão 31

Estilo Cebraspe - Questão inédita

O aprendizado por transferência permite que um modelo treinado em uma tarefa seja reutilizado e adaptado para outra tarefa relacionada, muitas vezes reduzindo o tempo e os recursos necessários para o treinamento.

Comentário:

O aprendizado por transferência envolve usar um modelo pré-treinado em uma tarefa e adaptá-lo para uma nova tarefa relacionada. Isso é eficiente, pois o modelo já aprendeu características relevantes que podem ser úteis para a nova tarefa, reduzindo o tempo e recursos necessários.

Resposta: Certo

Questão 32

Estilo Cebraspe - Questão inédita

Um dos principais desafios do aprendizado supervisionado é a necessidade de uma grande quantidade de dados rotulados, o que pode ser custoso e demorado de obter.

Comentário:

O aprendizado supervisionado depende de grandes volumes de dados rotulados para treinar modelos de forma eficaz. A obtenção e rotulação desses dados podem ser caras e demoradas, constituindo um desafio significativo.

Resposta: Certo



www.kuasarnex.com

 [@kuasarnex](https://www.instagram.com/kuasarnex)