TI MAPEADA · 300 MAPAS MENTAIS DE TI

Módulo 10

Inteligência Artificial e Tecnologias Emergentes

30 mapas · mapas 272 a 301



Neste módulo:

272 Inteligência Artificial: Conceitos
273 Machine Learning: Tipos
274 Classificação e Regressão
275 Clustering
276 Aprendizado por Reforço
277 Overfitting e Underfitting
278 Métricas de Classificação
279 Matriz de Confusão
280 Validação de Modelos
281 Árvores de Decisão e Ensembles
282 KNN, SVM e Naive Bayes
283 Regressão Linear e Logística
284 Redes Neurais
285 Deep Learning e CNN
286 RNN e LSTM
287 Transformers
288 LLMs (Grandes Modelos de Linguagem)
289 Engenharia de Prompt
290 RAG
291 Embeddings e Bancos Vetoriais
292 Processamento de Linguagem Natural
293 Visão Computacional
294 IA Generativa
295 IA Responsável
296 Pré-processamento de Dados
297 Redução de Dimensionalidade
298 Internet das Coisas
299 Blockchain
300 Computação Quântica
301 Python para Dados

Inteligência Artificial: Conceitos

O QUE É

Área que cria sistemas capazes de executar tarefas que exigem inteligência.

1 IA fraca (estreita)

- tarefas específicas

2 IA forte (geral)

- inteligência ampla como a humana
- hipotética

3 Abordagens

- simbólica (regras)
- conexionista (redes neurais)

4 Subáreas

- aprendizado de máquina
- PLN
- visão
- robótica

5 Teste de Turing

- imitar um humano em conversa

⚠ CAI EM PROVA

Todos os sistemas atuais são IA estreita.

Material de estudo e revisão · uso pessoal

Machine Learning: Tipos

O QUE É

Algoritmos que aprendem padrões a partir de dados.

1 Supervisionado

- dados rotulados

2 Não supervisionado

- sem rótulos
- acha estruturas

3 Por reforço

- agente aprende com recompensas

4 Semissupervisionado

- poucos rótulos

5 Autossupervisionado

- rótulos gerados dos próprios dados

⚠ CAI EM PROVA

Agrupamento (clustering) é aprendizado não supervisionado.

Material de estudo e revisão · uso pessoal

Classificação e Regressão

O QUE É

As duas grandes tarefas do aprendizado supervisionado.

1 Classificação

- prevê categoria
- ex. spam ou não

2 Regressão

- prevê valor numérico
- ex. preço

3 Binária e multiclasse

- duas ou várias classes

4 Algoritmos

- árvores
- SVM
- regressão logística

5 Saída

- rótulo x número contínuo

⚠ CAI EM PROVA

Regressão logística é usada para classificação, apesar do nome.

Clustering

O QUE É

Agrupa dados parecidos sem rótulos prévios.

1 K-means

- k grupos
- centróides

2 Hierárquico

- dendrograma
- aglomerativo ou divisivo

3 DBSCAN

- densidade
- detecta ruído

4 Escolha de k

- método do cotovelo

5 Uso

- segmentação de clientes

⚠ CAI EM PROVA

No K-means o número de grupos precisa ser definido antes.

Material de estudo e revisão · uso pessoal

Aprendizado por Reforço

O QUE É

Agente aprende por tentativa e erro, interagindo com um ambiente.

1 Elementos

- agente
- ambiente
- estado
- ação
- recompensa

2 Política

- estratégia de escolha das ações

3 Exploração x exploração

- testar novo x usar o conhecido

4 Algoritmos

- Q-learning
- SARSA

5 Uso

- jogos
- robótica
- ajuste de LLMs (RLHF)

⚠ CAI EM PROVA

RLHF usa feedback humano como recompensa para ajustar modelos de linguagem.

Overfitting e Underfitting

O QUE É

Problemas de generalização de modelos.

1 Overfitting

- decora o treino
- vai mal em dados novos

2 Underfitting

- modelo simples demais
- vai mal em tudo

3 Viés x variância

- equilíbrio necessário

4 Combate ao overfitting

- regularização
- mais dados
- dropout

5 Mais

- validação
- early stopping

⚠ CAI EM PROVA

Erro baixo no treino e alto no teste indica overfitting.

Material de estudo e revisão · uso pessoal

Métricas de Classificação

O QUE É

Medem a qualidade de um classificador.

1 Acurácia

- acertos sobre o total

2 Precisão

- VP / (VP + FP)

3 Recall (sensibilidade)

- VP / (VP + FN)

4 F1

- média harmônica de precisão e recall

5 AUC-ROC

- separação entre as classes

⚠ CAI EM PROVA

Com classes desbalanceadas, acurácia alta pode enganar.

Matriz de Confusão

O QUE É

Tabela que cruza classes reais e previstas.

1 VP

- positivo previsto como positivo

2 VN

- negativo previsto como negativo

3 FP

- negativo previsto como positivo (erro tipo I)

4 FN

- positivo previsto como negativo (erro tipo II)

5 Especificidade

- VN / (VN + FP)

⚠ CAI EM PROVA

Falso positivo é o erro tipo I; falso negativo, o tipo II.

Validação de Modelos

O QUE É

Avaliar o modelo em dados que ele não viu.

1 Divisão

- treino
- validação
- teste

2 K-fold

- k partes
- cada uma vira teste uma vez

3 Estratificada

- mantém a proporção das classes

4 Vazamento de dados

- informação do teste no treino

5 Hiperparâmetros

- ajustados na validação

⚠ CAI EM PROVA

O conjunto de teste não deve ser usado para ajustar o modelo.

Material de estudo e revisão · uso pessoal

Árvores de Decisão e Ensembles

O QUE É

Modelos que decidem por regras em forma de árvore e combinações delas.

1 Árvore

- nós de decisão
- folhas com a resposta

2 Critérios

- ganho de informação (entropia)
- Gini

3 Random Forest

- muitas árvores
- bagging

4 Boosting

- árvores em sequência
- XGBoost

5 Vantagem

- árvores são interpretáveis

⚠ CAI EM PROVA

Random Forest usa bagging; XGBoost usa boosting.

Material de estudo e revisão · uso pessoal

KNN, SVM e Naive Bayes

O QUE É

Algoritmos clássicos de classificação.

1 KNN

- classe dos k vizinhos mais próximos

2 KNN

- preguiçoso
- não treina modelo

3 SVM

- hiperplano de margem máxima

4 SVM kernel

- separa dados não lineares

5 Naive Bayes

- probabilidade
- atributos independentes

⚠ CAI EM PROVA

O Naive Bayes assume independência entre os atributos.

Material de estudo e revisão · uso pessoal

Regressão Linear e Logística

O QUE É

Modelos estatísticos de base para previsão.

1 Linear

- reta que minimiza o erro quadrático

2 Métricas

- MAE
- MSE
- RMSE
- R^2

3 Logística

- função sigmoide
- probabilidade

4 Regularização

- Lasso (L1)
- Ridge (L2)

5 Lasso

- pode zerar coeficientes

⚠ CAI EM PROVA

R^2 mede quanto da variação o modelo explica.

Material de estudo e revisão · uso pessoal

Redes Neurais

O QUE É

Modelos inspirados nos neurônios, com camadas de unidades conectadas.

1 Perceptron

- neurônio artificial simples

2 Camadas

- entrada
- ocultas
- saída

3 Ativação

- ReLU
- sigmoide
- tanh
- softmax

4 Treino

- backpropagation e gradiente descendente

5 Hiperparâmetros

- taxa de aprendizado
- épocas
- lote

⚠ CAI EM PROVA

Perceptron simples não resolve o XOR, que não é linearmente separável.

Material de estudo e revisão · uso pessoal

Deep Learning e CNN

O QUE É

Redes profundas; CNNs são especializadas em imagens.

1 Deep learning

- muitas camadas ocultas

2 CNN

- convolução extrai características

3 Pooling

- reduz a dimensão

4 Uso de CNN

- classificação de imagens
- detecção de objetos

5 Transfer learning

- reaproveita modelo treinado

⚠ CAI EM PROVA

Camadas convolucionais compartilham pesos, reduzindo parâmetros.

Material de estudo e revisão · uso pessoal

RNN e LSTM

O QUE É

Redes para dados sequenciais.

1 RNN

- memória do passo anterior

2 Problema

- gradiente que desaparece

3 LSTM

- portas de entrada, esquecimento e saída

4 GRU

- versão simplificada da LSTM

5 Uso

- séries temporais
- texto
- voz

⚠ CAI EM PROVA

A LSTM foi criada para reduzir o problema do gradiente que desaparece.

Material de estudo e revisão · uso pessoal

Transformers

O QUE É

Arquitetura baseada em atenção, base dos modelos de linguagem atuais.

1 Artigo

- Attention Is All You Need (2017)

2 Autoatenção

- cada token olha para os outros

3 Partes

- encoder
- decoder

4 Paralelismo

- processa a sequência inteira de uma vez

5 Exemplos

- BERT (encoder)
- GPT (decoder)

⚠ CAI EM PROVA

Transformers dispensam recorrência e usam atenção.

Material de estudo e revisão · uso pessoal

LLMs (Grandes Modelos de Linguagem)

O QUE É

Modelos com bilhões de parâmetros treinados em grandes volumes de texto.

1 Treino

- pré-treino prevê o próximo token

2 Ajuste

- fine-tuning
- RLHF

3 Token

- pedaço de texto processado

4 Janela de contexto

- tokens considerados de uma vez

5 Alucinação

- resposta plausível, mas falsa

⚠ CAI EM PROVA

LLM gera texto prevendo tokens; pode alucinar fatos.

Material de estudo e revisão · uso pessoal

Engenharia de Prompt

O QUE É

Técnicas para escrever instruções que melhoram as respostas do modelo.

1 Zero-shot

- só a instrução

2 Few-shot

- instrução com exemplos

3 Cadeia de pensamento

- raciocínio passo a passo

4 Papel e contexto

- define persona e formato

5 Temperatura

- maior = mais criativo e variado

⚠ CAI EM PROVA

Few-shot inclui exemplos no prompt sem retreinar o modelo.

RAG

O QUE É

Geração aumentada por recuperação: busca documentos para embasar a resposta.

1 Etapa 1

- indexar documentos em pedaços (chunks)

2 Etapa 2

- gerar embeddings e guardar em banco vetorial

3 Etapa 3

- recuperar trechos parecidos com a pergunta

4 Etapa 4

- o LLM responde usando os trechos

5 Benefícios

- menos alucinação
- conhecimento atualizado

⚠ CAI EM PROVA

RAG não altera os pesos do modelo, diferente do fine-tuning.

Embeddings e Bancos Vetoriais

O QUE É

Representações numéricas do significado e onde guardá-las.

1 Embedding

- vetor que representa texto ou imagem

2 Similaridade

- cosseno
- distância euclidiana

3 Bancos vetoriais

- FAISS
- Pinecone
- pgvector

4 Busca

- vizinhos mais próximos aproximados (ANN)

5 Uso

- busca semântica
- recomendação

⚠ CAI EM PROVA

Textos de significado parecido geram vetores próximos.

Material de estudo e revisão · uso pessoal

Processamento de Linguagem Natural

O QUE É

Faz computadores entenderem e gerarem linguagem humana.

1 Pré-processamento

- tokenização
- stopwords
- stemming
- lematização

2 Representação

- bag of words
- TF-IDF
- embeddings

3 Tarefas

- análise de sentimento
- NER
- tradução

4 Modelos

- Word2Vec
- BERT
- GPT

5 Avaliação

- BLEU
- ROUGE

⚠ CAI EM PROVA

Lematização volta à forma de dicionário; stemming só corta o radical.

Visão Computacional

O QUE É

Extrai informação de imagens e vídeos.

1 Classificação

- o que tem na imagem

2 Detecção

- onde estão os objetos
- YOLO

3 Segmentação

- classe por pixel

4 OCR

- reconhece texto em imagem

5 Biblioteca

- OpenCV

⚠ CAI EM PROVA

Detecção localiza objetos com caixas; segmentação classifica cada pixel.

IA Generativa

O QUE É

Modelos que criam conteúdo novo: texto, imagem, áudio e código.

1 Texto

- LLMs

2 Imagem

- modelos de difusão

3 GANs

- gerador contra discriminador

4 VAEs

- codificam e geram variações

5 Riscos

- deepfakes
- direitos autorais

⚠ CAI EM PROVA

Na GAN, o discriminador tenta distinguir dados reais dos gerados.

Material de estudo e revisão · uso pessoal

IA Responsável

O QUE É

Uso ético, seguro e transparente da inteligência artificial.

1 Viés

- dados de treino desbalanceados

2 Explicabilidade

- entender as decisões (XAI)

3 Privacidade

- proteção dos dados usados

4 Supervisão humana

- decisões de alto risco

5 Regulação

- AI Act europeu por nível de risco

⚠ CAI EM PROVA

O AI Act da União Europeia classifica sistemas por nível de risco.

Pré-processamento de Dados

O QUE É

Preparar dados para o treino dos modelos.

1 Limpeza

- valores ausentes
- duplicados
- outliers

2 Normalização

- escala entre 0 e 1

3 Padronização

- média 0 e desvio 1

4 Codificação

- one-hot para categorias

5 Balanceamento

- oversampling
- undersampling
- SMOTE

⚠ CAI EM PROVA

Ajuste o normalizador só no treino para evitar vazamento.

Redução de Dimensionalidade

O QUE É

Diminui o número de variáveis mantendo a informação.

1 PCA

- componentes principais
- maior variância

2 t-SNE e UMAP

- visualização em 2D ou 3D

3 Seleção de atributos

- escolhe variáveis

4 Extração de atributos

- cria novas variáveis

5 Maldição da dimensionalidade

- dados esparsos

⚠ CAI EM PROVA

PCA é não supervisionado e cria componentes não correlacionados.

Material de estudo e revisão · uso pessoal

Internet das Coisas

O QUE É

Objetos físicos conectados que coletam e trocam dados.

1 Camadas

- percepção
- rede
- aplicação

2 Protocolos

- MQTT
- CoAP
- Zigbee
- LoRaWAN

3 MQTT

- publish/subscribe leve
- broker

4 Desafios

- segurança
- energia
- escala

5 Uso

- cidades inteligentes
- indústria 4.0

⚠ CAI EM PROVA

MQTT usa o modelo publicar e assinar com um broker.

Blockchain

O QUE É

Registro distribuído e imutável de transações em blocos encadeados.

1 Bloco

- dados
- hash próprio
- hash do anterior

2 Consenso

- prova de trabalho
- prova de participação

3 Tipos

- pública
- privada
- consórcio

4 Contratos inteligentes

- código executado na rede

5 Características

- descentralização
- rastreabilidade

⚠ CAI EM PROVA

Alterar um bloco muda seu hash e quebra o encadeamento.

Material de estudo e revisão · uso pessoal

Computação Quântica

O QUE É

Usa fenômenos quânticos para processar informação.

1 Qubit

- pode estar em superposição de 0 e 1

2 Emaranhamento

- qubits correlacionados

3 Algoritmo de Shor

- fatora números
- ameaça o RSA

4 Algoritmo de Grover

- acelera buscas

5 Resposta

- criptografia pós-quântica

⚠ CAI EM PROVA

O algoritmo de Shor ameaça a criptografia assimétrica como o RSA.

Material de estudo e revisão · uso pessoal

Python para Dados

O QUE É

Bibliotecas Python mais usadas em ciência de dados.

1 NumPy

- arrays e cálculo numérico

2 Pandas

- DataFrames
- leitura de CSV

3 Matplotlib e Seaborn

- gráficos

4 scikit-learn

- modelos clássicos de ML

5 TensorFlow e PyTorch

- deep learning

⚠ CAI EM PROVA

O DataFrame é a estrutura principal do Pandas.