

# Aula 6 – Aprendizado Supervisionado: Parte 2 (Classificação)

## Desvendando a Classificação: Como Máquinas Tomam Decisões Categorias

Bem-vindo à Aula 6 do nosso Curso de Inteligência Artificial Aplicada! Se você já se perguntou como seu e-mail sabe o que é spam, ou como um aplicativo de fotos consegue identificar um cachorro de um gato, você está prestes a descobrir os segredos por trás dessas "decisões" inteligentes. Na aula anterior, exploramos o Aprendizado Supervisionado através da Regressão, onde o objetivo era prever valores contínuos. Agora, vamos mergulhar na outra face dessa moeda: a **Classificação**.

A classificação é uma das tarefas mais fundamentais e ubíquas da Inteligência Artificial, presente em inúmeras aplicações que moldam nosso dia a dia. Ela permite que as máquinas organizem informações, identifiquem padrões e, em última instância, tomem decisões baseadas em categorias distintas. É como ter um assistente super inteligente que sabe exatamente onde cada coisa pertence.

Ao final desta aula, você não apenas entenderá os conceitos por trás da classificação, mas também será capaz de identificar diferentes tipos de problemas de classificação, compreender o funcionamento de algoritmos essenciais como Regressão Logística, K-Nearest Neighbors (KNN) e Árvores de Decisão, e, crucialmente, avaliar a performance desses modelos usando métricas como Acurácia, Precisão, Recall e a Matriz de Confusão. Prepare-se para ver o mundo através das lentes da IA, onde cada dado pode ser uma pista para uma categoria.

# Problemas de Classificação: Prevendo Categorias

Imagine por um momento que você é um bibliotecário. Sua tarefa não é apenas organizar livros, mas também prever a qual gênero um novo livro pertence antes mesmo de lê-lo por completo, apenas olhando para a capa, o título e a sinopse. Você não está prevendo o número de páginas ou o preço (isso seria regressão), mas sim a **categoria**: ficção científica, romance, biografia, etc. Essa é a essência dos problemas de classificação na Inteligência Artificial.

- ❏ No universo da IA, um problema de classificação surge quando nosso objetivo é prever uma **variável de saída categórica**. Isso significa que a resposta que buscamos não é um número contínuo, mas sim uma de um conjunto finito de categorias ou rótulos.

Pense em cenários como identificar se um e-mail é "spam" ou "não spam", se uma imagem contém um "cachorro" ou um "gato", ou se um cliente irá "comprar" ou "não comprar" um produto. Cada uma dessas respostas é uma categoria distinta.

A grande sacada aqui é que, para que um algoritmo de classificação funcione, ele precisa ser "treinado" com dados que já possuem essas categorias definidas. É como se o bibliotecário tivesse acesso a milhares de livros já classificados por gênero, aprendendo com eles as características que definem cada categoria. Com esse conhecimento, ele pode então aplicar o que aprendeu para classificar novos livros. Essa é a base do **Aprendizado Supervisionado** que exploramos na aula anterior, agora aplicada a um novo tipo de problema.

# Regressão Logística: A Porta de Entrada para a Classificação

Quando você ouve o termo "Regressão Logística", é natural que a palavra "regressão" o leve a pensar em prever valores numéricos contínuos, como fizemos na aula anterior. No entanto, o nome pode ser um pouco enganoso! A **Regressão Logística** é, na verdade, um dos algoritmos mais fundamentais e amplamente utilizados para resolver problemas de **classificação**, especialmente aqueles com apenas duas categorias (classificação binária).

## Probabilidades

Estima a probabilidade de uma instância pertencer a uma determinada classe (valor entre 0 e 1)

## Função Sigmoide

Transforma qualquer valor real em um valor entre 0 e 1 usando a função logística

## Decisão

Se probabilidade  $> 0.5$ , classifica como "sim"; caso contrário, como "não"

A magia da Regressão Logística reside em sua capacidade de estimar a **probabilidade** de uma instância pertencer a uma determinada classe. Em vez de prever diretamente "sim" ou "não", ela prevê a probabilidade de "sim" (um valor entre 0 e 1). Se essa probabilidade for maior que um certo limite (geralmente 0.5), o modelo classifica a instância como "sim"; caso contrário, como "não". Pense nisso como um "dimmer switch" para decisões: em vez de um interruptor de luz que é apenas "ligado" ou "desligado", o dimmer permite uma gradação, e você decide o ponto em que a luz está "acesa o suficiente".

Por exemplo, em um cenário de detecção de fraude, a Regressão Logística pode nos dizer que uma transação tem 0.85 (85%) de chance de ser fraudulenta. Com base nesse valor, podemos decidir se a transação deve ser bloqueada ou não. É uma ferramenta poderosa para transformar uma previsão numérica em uma decisão categórica clara.

# K-Nearest Neighbors (KNN): Vizinhança que Decide

Imagine que você acabou de se mudar para uma nova cidade e está tentando descobrir quais restaurantes são bons. Uma das maneiras mais intuitivas de fazer isso é perguntar aos seus vizinhos mais próximos. Se a maioria dos seus vizinhos recomenda um certo restaurante, é provável que ele seja bom, certo? Essa ideia simples está no cerne do algoritmo **K-Nearest Neighbors (KNN)**.

01

---

## Calcular Distâncias

Para cada novo ponto, calcula a distância (geralmente euclidiana) para todos os pontos do conjunto de treinamento

02

---

## Selecionar K Vizinhos

Identifica os K pontos mais próximos (vizinhos) do novo ponto de dados

03

---

## Votação por Maioria

A classe do novo ponto é determinada pela maioria dos votos entre os K vizinhos selecionados

O KNN é um algoritmo de classificação que se baseia na premissa de que "objetos semelhantes estão próximos uns dos outros". Em outras palavras, para classificar um novo ponto de dados, o KNN olha para os **K** pontos de dados mais próximos (seus "vizinhos") no conjunto de treinamento. A classe do novo ponto é então determinada pela **maioria dos votos** entre esses K vizinhos. Se  $K=3$  e dois dos seus três vizinhos mais próximos são da classe "A" e um é da classe "B", o novo ponto será classificado como "A".

A beleza do KNN reside em sua simplicidade e intuição. Ele não constrói um modelo complexo durante o treinamento; em vez disso, ele "memoriza" todo o conjunto de dados. Quando uma nova instância chega, ele calcula a distância entre essa nova instância e todas as instâncias existentes, seleciona os K mais próximos e então "vota". É uma abordagem preguiçosa, mas muitas vezes eficaz!

# Comparando Regressão Logística e KNN

Agora que exploramos a Regressão Logística e o KNN individualmente, é importante entender quando cada um brilha e quais são suas características distintivas. Ambos são poderosas ferramentas de classificação, mas operam sob princípios muito diferentes e são mais adequados para cenários distintos.

## Regressão Logística

### Modelo Paramétrico

- Assume uma forma funcional específica
- Eficiente computacionalmente
- Mais interpretável
- Limitada por relações não lineares

Como um *"especialista"* que aprendeu regras gerais para tomar decisões

## K-Nearest Neighbors

### Modelo Não Paramétrico

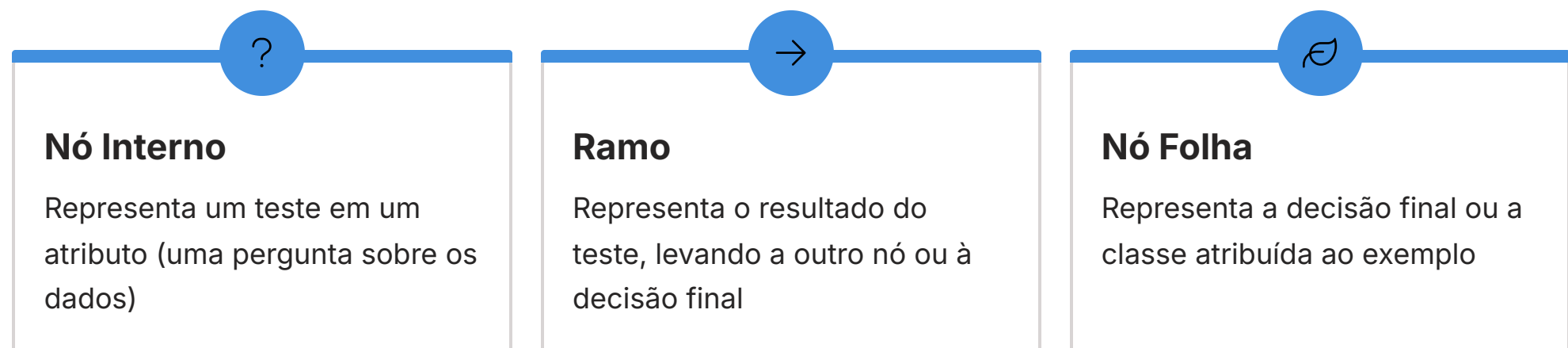
- Não faz suposições sobre distribuição
- Muito flexível
- Captura relações complexas
- Computacionalmente caro

Como um *"comitê de vizinhos"* que decide caso a caso

Conceito	Abordagem Principal	Vantagens	Desvantagens
<b>Regressão Logística</b>	Modelo paramétrico, estima probabilidades via sigmoide	Rápida, interpretável, bom para dados lineares	Pode ser limitada por relações não lineares
<b>K-Nearest Neighbors (KNN)</b>	Modelo não paramétrico, baseado em vizinhança	Flexível, simples, bom para relações complexas	Lento com muitos dados, sensível a ruído e escala

# Árvores de Decisão: O Caminho Lógico para a Decisão

Você já seguiu um fluxograma para resolver um problema ou tomar uma decisão? Se sim, você já tem uma boa intuição sobre como funcionam as **Árvores de Decisão**. Este algoritmo de classificação é incrivelmente intuitivo porque imita o processo de tomada de decisão humana, dividindo um problema complexo em uma série de perguntas mais simples.



Uma Árvore de Decisão é, literalmente, uma estrutura em forma de árvore, onde cada **nó interno** representa um teste em um atributo (uma pergunta), cada **ramo** representa o resultado desse teste, e cada **nó folha** (o final do caminho) representa a decisão ou a classe final. O algoritmo aprende a sequência de perguntas que levam à melhor classificação.

**Exemplo prático:** Para decidir se um cliente vai comprar um produto, a árvore pode primeiro perguntar: "O cliente tem mais de 30 anos?". Se sim, "Ele já comprou algo antes?". E assim por diante, até chegar a uma conclusão.

A grande vantagem das Árvores de Decisão é sua **interpretabilidade**. Ao contrário de alguns modelos de "caixa preta" na IA, você pode literalmente seguir o caminho da decisão e entender por que uma determinada classificação foi feita. Isso é extremamente valioso em áreas onde a transparência e a explicabilidade são cruciais, como medicina ou finanças. Elas são como um mapa claro que mostra exatamente como o modelo chegou à sua resposta.

# A Interpretabilidade das Árvores de Decisão e Seus Desafios

A capacidade de uma Árvore de Decisão de ser "lida" e compreendida por humanos é uma de suas maiores forças, especialmente no contexto da **IA Explicável (XAI - Explainable AI)**. Em um mundo onde os modelos de IA estão se tornando cada vez mais complexos e impactam decisões críticas, a interpretabilidade não é apenas uma conveniência, mas uma necessidade.

## Vantagens da Interpretabilidade

- Transparência total no processo de decisão
- Facilita auditoria e conformidade regulatória
- Constrói confiança entre usuários e stakeholders
- Permite identificação de vieses e erros lógicos

## Desafios e Limitações

- Árvores muito profundas se tornam complexas demais
- Propensão ao overfitting (memorização excessiva)
- Perda de interpretabilidade com o aumento da complexidade
- Necessidade de técnicas de poda para controle


No entanto, a história não termina aqui. Embora árvores pequenas e simples sejam fáceis de interpretar, elas podem se tornar excessivamente complexas e difíceis de visualizar à medida que o número de atributos e a profundidade da árvore aumentam. Uma árvore muito "profunda" pode ser tão confusa quanto um labirinto, perdendo sua vantagem de interpretabilidade.

Além disso, Árvores de Decisão são propensas a um problema chamado **overfitting**, onde o modelo se ajusta tão perfeitamente aos dados de treinamento que perde a capacidade de generalizar bem para novos dados. É como um estudante que memoriza todas as respostas de um livro, mas não consegue aplicar o conhecimento em uma nova situação.

Para mitigar o overfitting, técnicas como a **poda** (remover ramos desnecessários) ou a limitação da profundidade da árvore são comumente usadas. Apesar desses desafios, a interpretabilidade inerente das Árvores de Decisão as torna uma ferramenta valiosa, especialmente como base para algoritmos mais avançados, como as Florestas Aleatórias (Random Forests) e o Gradient Boosting.

# Métricas de Avaliação: Além do "Certo ou Errado"

Quando um modelo de classificação faz uma previsão, ele pode acertar ou errar. Intuitivamente, poderíamos pensar que a **Acurácia** (a proporção de previsões corretas) é a única métrica que importa. No entanto, a realidade é muito mais complexa.

 **Exemplo ilustrativo:** Imagine um modelo que detecta uma doença rara que afeta apenas 1% da população. Se o modelo simplesmente disser que "ninguém tem a doença" o tempo todo, ele terá uma acurácia de 99%! Mas ele falhou completamente em seu objetivo principal de identificar os doentes.

Isso nos leva à necessidade de um conjunto mais sofisticado de métricas para avaliar o desempenho de um modelo de classificação. Precisamos entender não apenas "quantos acertos", mas também "que tipo de acertos e erros" o modelo está cometendo. Um erro em um contexto (como detectar spam) pode ter consequências muito diferentes de um erro em outro (como diagnosticar câncer).



## Verdadeiro Positivo (VP)

Modelo previu "Positivo" e a classe real era "Positivo" - **Acerto**



## Falso Positivo (FP)

Modelo previu "Positivo" e a classe real era "Negativo" - **Falso Alarme**



## Falso Negativo (FN)

Modelo previu "Negativo" e a classe real era "Positivo" - **Falha em Detectar**



## Verdadeiro Negativo (VN)

Modelo previu "Negativo" e a classe real era "Negativo" - **Acerto**

Para isso, introduzimos os conceitos de **Verdadeiro Positivo (VP)**, **Verdadeiro Negativo (VN)**, **Falso Positivo (FP)** e **Falso Negativo (FN)**. Essas quatro categorias formam a base para todas as métricas de avaliação de classificação e nos permitem ter uma visão muito mais granular do desempenho do nosso modelo.

# Matriz de Confusão: O Mapa Detalhado dos Acertos e Erros

Para organizar e visualizar os Verdadeiros Positivos, Verdadeiros Negativos, Falsos Positivos e Falsos Negativos, utilizamos uma ferramenta essencial chamada **Matriz de Confusão**. Ela é, de fato, um mapa detalhado que nos mostra onde o modelo acertou e onde errou, e, mais importante, o tipo de erro que ele cometeu.

A Matriz de Confusão é uma tabela que compara as classes reais (o que realmente aconteceu) com as classes previstas pelo modelo. Para um problema de classificação binária (duas classes, por exemplo, "Positivo" e "Negativo"), a matriz terá 2x2 células:

## Verdadeiro Positivo (VP)

O modelo previu "Positivo" e a classe real era "Positivo". **(Acerto)**

## Falso Positivo (FP)

O modelo previu "Positivo" e a classe real era "Negativo". **(Erro tipo I - Falso Alarme)**

## Falso Negativo (FN)

O modelo previu "Negativo" e a classe real era "Positivo". **(Erro tipo II - Falha em Detectar)**

## Verdadeiro Negativo (VN)

O modelo previu "Negativo" e a classe real era "Negativo". **(Acerto)**

### Exemplo prático - Detecção de Spam:

- VP: E-mail que é spam e foi classificado como spam
- FP: E-mail legítimo classificado como spam (falso alarme)
- FN: Spam que passou como e-mail legítimo (falha em detectar)
- VN: E-mail legítimo classificado corretamente como legítimo

A Matriz de Confusão nos dá essa visão clara e concisa, sendo o ponto de partida para calcular todas as outras métricas importantes.

# Acurácia, Precisão e Recall: As Lentes para Avaliar Modelos

Com a Matriz de Confusão em mãos, podemos agora calcular as métricas que nos darão uma visão mais completa do desempenho do nosso modelo.



## Acurácia (Accuracy)

**Fórmula:**  $(VP + VN) / (VP + VN + FP + FN)$

É a métrica mais simples e intuitiva. Representa a proporção de previsões corretas sobre o total de previsões.

**Quando usar:** Útil quando as classes são balanceadas e os custos dos erros são simétricos.



## Precisão (Precision)

**Fórmula:**  $VP / (VP + FP)$

Responde: "Das vezes que o modelo previu 'Positivo', quantos estavam realmente corretos?"

**Quando usar:** Importante quando o custo de um Falso Positivo é alto. Ex: detecção de spam.



## Recall (Sensibilidade)

**Fórmula:**  $VP / (VP + FN)$

Responde: "Dos casos que eram realmente 'Positivos', quantos o modelo conseguiu detectar?"

**Quando usar:** Importante quando o custo de um Falso Negativo é alto. Ex: diagnóstico de doenças.

## Analogia do Pescador

**Precisão:** "Qualidade" das previsões positivas

*Um pescador com alta precisão só pega peixes comestíveis, mas pode perder muitos.*

## Analogia do Pescador

**Recall:** "Quantidade" de positivos capturados

Um pescador com alto recall pega quase todos os peixes, mas pode trazer muitos que não servem.

# O Dilema Precisão vs. Recall e a Importância do F1-Score

Em muitos cenários do mundo real, a Precisão e o Recall apresentam um **dilema**: melhorar um geralmente significa piorar o outro. É um trade-off. Por exemplo, se um modelo de detecção de fraude for ajustado para ter um Recall muito alto (para não perder nenhuma fraude), ele pode acabar gerando muitos Falsos Positivos (alertas de fraude para transações legítimas), o que pode irritar os clientes.



## Priorizar Precisão

### Quando Falsos Positivos são muito caros


- Sistema de recomendação de produtos
- Detecção de spam
- Aprovação de crédito



## Priorizar Recall

### Quando Falsos Negativos são muito caros

- Diagnóstico de câncer
- Detecção de intrusão em segurança
- Detecção de fraude financeira

 **F1-Score:** Para situações onde precisamos de um equilíbrio entre Precisão e Recall, utilizamos o F1-Score. É a média harmônica da Precisão e do Recall, fornecendo uma única métrica que penaliza modelos que têm um desempenho muito bom em uma métrica e muito ruim na outra.

A escolha de qual métrica priorizar depende criticamente do **contexto do problema** e dos **custos associados a cada tipo de erro**. O F1-Score é particularmente útil quando você busca um balanço entre os dois ou quando as classes são desbalanceadas.

# Estudo de Caso: Detecção de Spam em E-mails – Uma Aplicação Clássica

A detecção de spam é um dos exemplos mais clássicos e compreensíveis de um problema de classificação binária. Todos nós já lidamos com e-mails indesejados, e a capacidade de um sistema de e-mail de filtrar a maioria deles é um testemunho do poder dos algoritmos de classificação.

01

---

## Coleta e Rotulação

Milhares de e-mails são coletados e rotulados manualmente como "spam" ou "não spam" (ham)

03

---

## Treinamento do Modelo

O modelo aprende a distinguir entre spam e e-mails legítimos baseado nos padrões identificados

## Tipos de Erro e Consequências

**Falso Positivo (FP):** E-mail legítimo classificado como spam

*Consequência: Usuário pode perder mensagem importante*

---

**Falso Negativo (FN):** Spam classificado como legítimo

*Consequência: Spam polui a caixa de entrada*

02

---

## Extração de Características

O algoritmo aprende padrões: frequência de palavras suspeitas, estrutura do remetente, presença de links

04

---

## Classificação em Tempo Real

Quando um novo e-mail chega, o modelo aplica o conhecimento adquirido para fazer sua previsão

## Estratégia de Otimização

Nesse cenário, o **Falso Positivo** (perder um e-mail importante) é geralmente considerado mais grave do que um **Falso Negativo** (receber um spam).

Por isso, modelos de detecção de spam frequentemente buscam um bom equilíbrio entre Precisão e Recall, ou priorizam ligeiramente a **Precisão**.

# Desafios e Tendências em Classificação: Viés e Explicabilidade (XAI)

À medida que os modelos de classificação se tornam mais poderosos e são aplicados em domínios cada vez mais críticos – como saúde, justiça e finanças – surgem desafios importantes que vão além da simples acurácia. Duas das tendências mais discutidas e cruciais em 2025 são o **viés algorítmico** e a necessidade de **explicabilidade (XAI)**.

## Viés Algorítmico

Ocorre quando um modelo de IA, sem intenção, reproduz ou amplifica preconceitos presentes nos dados de treinamento.

**Exemplo:** Um modelo de classificação de risco de crédito treinado predominantemente com dados de um grupo demográfico específico pode classificar injustamente indivíduos de outros grupos como de maior risco.

**Impacto:** Decisões discriminatórias e injustas que perpetuam desigualdades sociais.

## IA Explicável (XAI)

Busca tornar os modelos de IA mais transparentes e compreensíveis para os humanos.

**Objetivo:** Em vez de aceitar uma classificação como uma "caixa preta", a XAI nos permite entender *por que* o modelo tomou aquela decisão.

**Importância:** Vital para auditar modelos, identificar vieses, garantir conformidade regulatória e construir confiança.

- ❑ **Regulamentação:** O AI Act da União Europeia exige transparência em sistemas de alto risco, tornando a explicabilidade não apenas uma boa prática, mas uma exigência legal em muitos contextos.

A detecção e mitigação de viés são áreas de pesquisa e desenvolvimento intensas, essenciais para uma IA ética e responsável. Modelos como Árvores de Decisão são inerentemente mais explicáveis, mas para modelos mais complexos (como redes neurais), técnicas de XAI são desenvolvidas para "abrir a caixa preta" e revelar os fatores mais influentes na decisão.

# O Futuro da Classificação e a IA Generativa

O campo da classificação está em constante evolução, e as inovações em **IA Generativa** estão começando a impactar significativamente como abordamos esses problemas. Modelos como GPT-4 (para texto), DALL-E 3 e Midjourney (para imagens) não são diretamente algoritmos de classificação, mas sua capacidade de gerar dados realistas e coerentes abre novas fronteiras.



## Dados Sintéticos

A IA Generativa pode criar dados sintéticos adicionais, aumentando a diversidade e o volume do conjunto de treinamento. Isso leva a modelos classificadores mais robustos e generalizáveis.

**Exemplo:** Para treinar um classificador de imagens de doenças raras, DALL-E 3 poderia gerar variações de imagens de casos existentes.



## Engenharia de Características

A IA Generativa auxilia na identificação ou criação de novas características a partir de dados brutos que podem melhorar o desempenho dos classificadores.

**Exemplo:** GPT-4 pode extrair informações relevantes de textos não estruturados, transformando-as em características úteis para classificação de sentimentos.



## Modelos Híbridos

A sinergia entre IA Generativa e classificação promete modelos mais eficientes, adaptáveis e capazes de lidar com cenários complexos e dados limitados.

**Futuro:** Sistemas que combinam geração de dados, extração de características e classificação em pipelines integrados.

# Consolidação: Classificando o Conhecimento

Chegamos ao fim da nossa jornada pela classificação no Aprendizado Supervisionado. Vimos que classificar é muito mais do que apenas "sim" ou "não"; é sobre entender categorias, probabilidades e as nuances dos erros. Exploramos algoritmos fundamentais como a Regressão Logística, que nos dá probabilidades, o KNN, que se baseia na sabedoria da vizinhança, e as Árvores de Decisão, que nos oferecem um caminho lógico e interpretável para a tomada de decisão.

## Definição Clara

Sempre defina claramente as categorias que seu modelo deve prever

## Escolha do Algoritmo

Escolha o algoritmo de classificação com base na natureza dos seus dados e na interpretabilidade desejada

## Avaliação Completa

Avalie seus modelos usando a Matriz de Confusão, Precisão, Recall e F1-Score, considerando os custos de cada tipo de erro

## Ética e Transparência

Esteja atento aos vieses nos seus dados e modelos, buscando sempre a explicabilidade

## Autoavaliação

- Qual das seguintes afirmações melhor descreve um problema de classificação em Aprendizado Supervisionado?**
  - a) Prever um valor numérico contínuo, como o preço de uma casa.
  - b) Agrupar dados sem rótulos pré-definidos em clusters.
  - c) Prever uma categoria ou rótulo discreto, como "spam" ou "não spam".
  - d) Reduzir a dimensionalidade de um conjunto de dados complexo.
- Um modelo de Regressão Logística prevê a probabilidade de um e-mail ser spam como 0.75. Se o limiar de decisão for 0.5, como o e-mail será classificado?**
  - a) Como "não spam", pois 0.75 é menor que 1.
  - b) Como "spam", pois 0.75 é maior que 0.5.
  - c) O modelo não pode classificar, apenas fornece probabilidades.
  - d) Depende do número de vizinhos mais próximos.
- Em um cenário de diagnóstico médico para uma doença rara, qual métrica de avaliação é geralmente mais crítica para minimizar o risco de não detectar um caso real da doença?**
  - a) Acurácia
  - b) Precisão
  - c) Recall
  - d) F1-Score
- Qual característica principal torna as Árvores de Decisão particularmente valiosas em contextos onde a transparência e a compreensão do "porquê" de uma decisão são cruciais?**
  - a) Sua alta velocidade de treinamento.
  - b) Sua capacidade de lidar com dados não lineares.
  - c) Sua inerente interpretabilidade e visualização como fluxogramas.
  - d) Sua resistência natural ao overfitting.
- Explique brevemente a importância da IA Explicável (XAI) no contexto dos modelos de classificação modernos, especialmente considerando as tendências regulatórias como o AI Act da União Europeia.

# Gabarito

**1** c) Prever uma categoria ou rótulo discreto, como "spam" ou "não spam".

**2** b) Como "spam", pois 0.75 é maior que 0.5.

**3** c) Recall

**4** c) Sua inerente interpretabilidade e visualização como fluxogramas.

**5** **Resposta Sugerida:**

A IA Explicável (XAI) é crucial para modelos de classificação modernos porque permite que humanos compreendam o raciocínio por trás das decisões da IA, transformando-os de "caixas pretas" em sistemas transparentes. Isso é vital para identificar e mitigar vieses algorítmicos, construir confiança nos sistemas de IA e garantir a conformidade com novas regulamentações, como o AI Act da União Europeia, que exige maior transparência e responsabilidade para sistemas de IA de alto risco.

# Próximos Passos e Recursos

## Próxima Aula: Aula 7 – Aprendizado Não Supervisionado: Descobrendo Padrões Ocultos

Prepare-se para explorar como a IA pode encontrar estruturas e grupos em dados sem a necessidade de rótulos pré-definidos!

### Recursos Adicionais

#### Livro

"Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow" por Aurélien Géron

*Para aprofundar nos algoritmos e implementações práticas*

#### Artigo


"Why is AI ethics important?"

*Para entender mais sobre viés algorítmico e XAI*

#### Documentação

Scikit-learn

*Para exemplos práticos de implementação em Python*

 **NOTA IMPORTANTE:** As informações regulatórias/legais/técnicas desta aula estão atualizadas até 2025. Consulte sempre fontes oficiais para verificar alterações.