

Redes Bayesianas

TEORIA E IMPLEMENTAÇÃO



Prof. Dr. Thales Levi Azevedo Valente

Equipe: Euderlan Freire, Hissa Bárbara, Maria Clara, Lucas Costa

SÚMARIO



INTRODUÇÃO



**IINFERÊNCIA EM
REDES BAYESIANAS**



**RACIOCÍNIO
PROBABILÍSTICO**



APLICAÇÃO PRÁTICA



REDES BAYESIANAS



REFERÊNCIAS

— Introdução

Sistemas que agem racionalmente

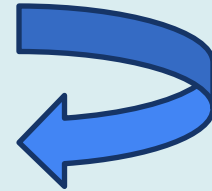


Raciocínio Lógico



Raciocínio Probabilístico

- ✓ Situações onde não se conhece todo o escopo
- ✓ Redes *Bayesianas*
 - Teoria de probabilidades
 - Teoria de grafos



— Raciocinando sobre incertezas

- *“A principal vantagem de raciocínio probabilístico sobre raciocínio lógico é fato de que agentes podem tomar decisões racionais mesmo quando não existe informação suficiente para se provar que uma ação funcionará”***[Russel]**
- Alguns fatores podem condicionar a falta de informação em uma base de conhecimento:
 - ✓ Ignorância Teórica
 - ✓ Impossibilidade



— Raciocinando sobre incertezas

- Utilizar **conectivos** que manipulem níveis de **certeza** e não apenas valores booleanos.
 - "Tenho **80% de chance** de fazer um bom trabalho"
 - "A **probabilidade** de um trabalho de IA ser bom é **50%**"
 - "A chance de um bom trabalho tirar nota máxima é **90%**"

Ao invés de dizer "sim" ou não,
dizemos:

- "Qual a **probabilidade** de eu tirar 10 neste trabalho?"
- Grafos podem representar relações causais entre eventos





— Raciocinando sobre incertezas

- **Considere o seguinte Domínio:** *Pela manhã meu **Del Rey** não irá funcionar. Eu posso ouvir a ignição, mas nada acontece. Podem existir várias razões para o problema. O rádio funciona, então a bateria está boa. A causa mais provável é que a gasolina tenha sido roubada durante a noite ou que a mangueira esteja entupida. Também pode ser que seja o carburador sujo, um vazamento na ignição ou algo mais sério. Para descobrir primeiro eu verifico o medidor de gasolina. Ele indica $\frac{1}{2}$ tanque, então eu decido limpar a mangueira da gasolina”*

• **Cenário: Carro
não funciona**

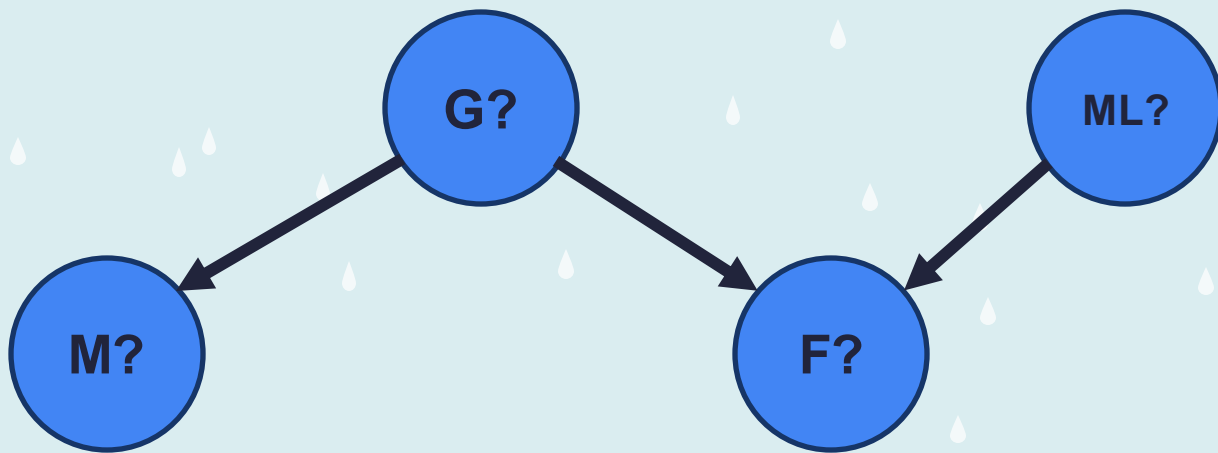
Gasolina? (Sim/Não)

**Mangueira Limpa?
(Sim/Não)**

**Funcionando?
(Sim/Não)**

**Medidor?
(Cheio/1/2/Vazio)**

— Raciocinando sobre incertezas



— Raciocínio Probabilístico

- Conceitos Importantes
 - Experimento Aleatório
 - Espaço Amostral
 - Evento
 - Probabilidade



— Raciocínio Probabilístico

- **Probabilidade**
 - Dado pelo intervalo $[0, 1]$
 - Se o espaço amostral consiste de N elementos igualmente prováveis e o evento A corresponde a um subconjunto de r elementos do espaço amostral, então a probabilidade é dada por: $P(A) = r/N$

— Raciocínio Probabilístico

- **Probabilidade Conjunta**

- Dois ou mais eventos acontecendo simultaneamente, dado que os eventos são independentes

$$P(A \wedge B) = P(A)P(B)$$

— Raciocínio Probabilístico

- **Probabilidade Condicional**
 - Dado um evento A, pode-se saber qual a probabilidade do evento B acontecer

$$P(A / B) = \frac{P(A \cap B)}{P(B)}$$

— Raciocínio Probabilístico

- **Probabilidade Condicional - Exemplo**
 - Qual a probabilidade de uma pessoa ter câncer, dado que ela fuma?

	Não Fuma	Fuma	Total
Sem Câncer	40	10	50
Câncer	7	3	10
Total	47	13	60

- $P(C="S" | F="S") = P(C="S" \wedge F="S") / P(F="S")$
 $= 3/60 / 13/60$
 $= 0,23$

— Raciocínio Probabilístico

- **Marginalização**
 - Também chamada de Totalização, é o processo de obter a probabilidade de uma variável sem considerar os valores das demais
 - Usa a distribuição conjunta para somar os eventos onde a variável desejada assume um determinado valor

— Raciocínio Probabilístico

- **Marginalização - Exemplo**

- Suponha duas variáveis: DorDeDente (D): {Verdadeiro, Falso}, Cárie (C): {Verdadeiro, Falso}. Qual a probabilidade de Cárie = Verdadeiro?

Cárie (C)	DorDeDente (D)	Probabilidade $P(C,D)P(C, D)$
Verdadeiro	Verdadeiro	0,108
Verdadeiro	Falso	0,012
Falso	Verdadeiro	0,072
Falso	Falso	0,008

- Solução: $P(C=V)=0,108+0,012=0,120$

— Raciocínio Probabilístico

- **Lei da Probabilidade Total**
 - Se um evento A ocorre em m condições distintas, todas mutuamente exclusivas, então a probabilidade de ocorrer o evento A é a soma das probabilidades dele ocorrer nas m condições diferentes

$$P(A) = \sum_{j=1}^m P(A \mid B_j) \cdot P(B_j)$$

— Raciocínio Probabilístico

- Teorema de Bayes

- Lida com a incertezas e atualiza a crença em um determinado evento à medida que novas informações chegam
- Ferramenta útil para inferir a probabilidade *a posteriori* de um evento baseado na evidencia e num conhecimento *a priori* de outros eventos

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$



— Raciocínio Probabilístico

- **Teorema de Bayes - Exemplo**

- Um médico sabe que meningite causa dor no pescoço em 50% dos casos. Ele sabe que a probabilidade a priori de um paciente ter meningite (M) é $1/50000$ e a possibilidade a priori de qualquer paciente ter uma dor no pescoço (S) é $1/20$
- Tem-se que: $P(S|M) = 1/2$, $P(M) = 1/50000$, $P(S) = 1/20$

— Raciocínio Probabilístico

- Teorema de Bayes – Exemplo
 - Um paciente chega ao consultório com dor no pescoço. Qual a probabilidade dele estar com meningite – $P(M|S)$?

$$P(M|S) = \frac{P(S|M) \cdot P(M)}{P(S)}$$



$$\begin{aligned}P(S|M) &= \frac{1}{2} \\ P(M) &= \frac{1}{50000} \\ P(S) &= \frac{1}{20}\end{aligned}$$

— Raciocínio Probabilístico

- **Teorema de Bayes – Exemplo**
- Um paciente chega ao consultório com dor no pescoço. Qual a probabilidade dele estar com meningite – $P(M|S)$?

$$P(M|S) = \frac{\frac{1}{2} \cdot \frac{1}{50000}}{\frac{1}{20}} = \frac{\frac{1}{100000}}{\frac{1}{20}}$$

$$P(M|S) = \frac{1}{100000} \cdot \frac{20}{1} = \frac{20}{100000} = \frac{1}{5000}$$



$$\frac{1}{5000} = 0,0002 = 0,02\%$$

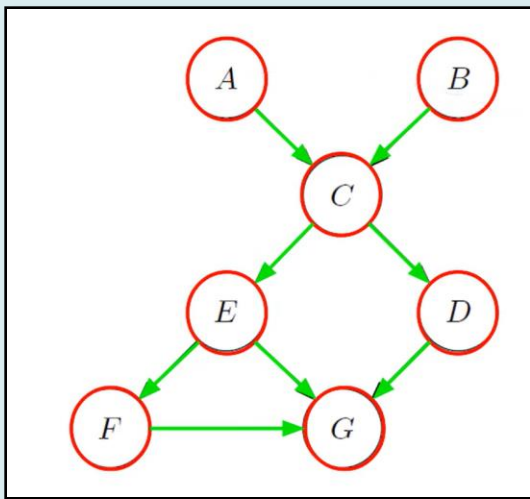
— Redes Bayesianas

- **Definição e Conceitos**
- **O que são:** Também chamadas de redes de crença, redes probabilísticas ou redes causais
- Teoria dos grafos
- Distribuição de probabilidades
- **Objetivo:** Representar situações, variáveis e estados para realizar inferências



— Redes Bayesianas

- Estrutura de uma Rede Bayesiana



- Um conjunto de **variáveis**
- Um conjunto de **arcos dirigidos**
- Cada variável possui um conjunto de finito estados mutuamente exclusivos
- Uma tabela de probabilidade condicionada para cada variável e seus pais
- O grafo é acíclico

— Redes Bayesianas



- **Exemplo de uma Rede Bayesiana**

"Um novo alarme contra assaltos é instalado, mesmo sendo muito confiável na detecção de assaltos ele pode disparar caso ocorra um terremoto. Os dois vizinhos João e Maria se disponibilizaram a telefonar caso o alarme dispare. João sempre liga quando ouve o alarme, entretanto algumas vezes ele confunde o alarme com o telefone e também liga nestes casos. já Maria gosta de ouvir música alta e as vezes não houve o alarme disparar, não ligando nestes casos"

— Redes Bayesianas

- Exemplo de uma Rede Bayesiana

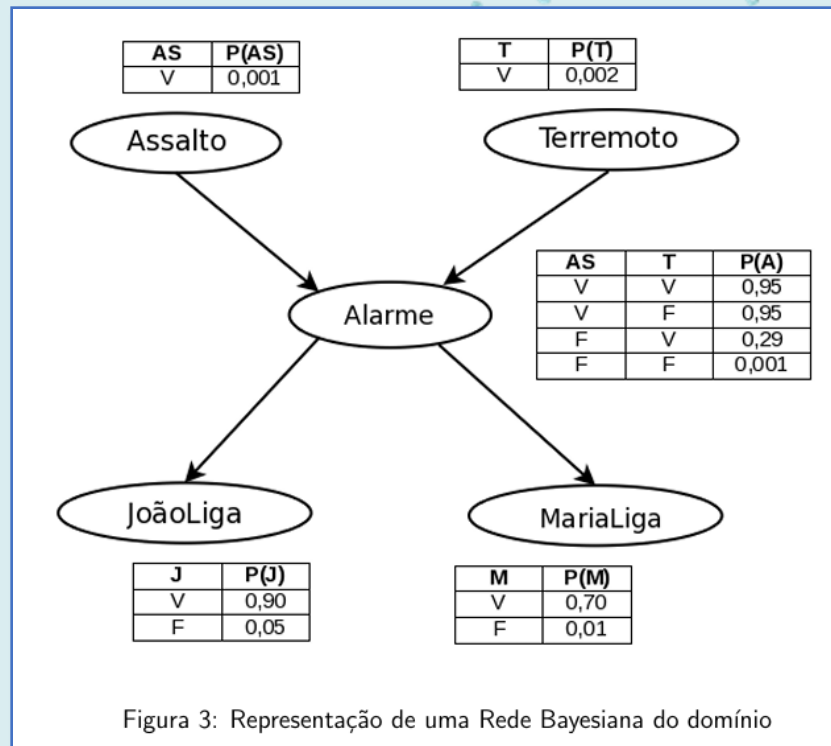


Figura 3: Representação de uma Rede Bayesiana do domínio

— Redes Bayesianas

- Exemplo de uma Rede Bayesiana
- *Considere que se deseja calcular a probabilidade do alarme ter tocado, mas nem um ladrão nem um terremoto aconteceram, e ambos, João e Maria ligaram, ou $P(J \wedge M \wedge A \wedge \neg L \wedge \neg T)$*
- $$\begin{aligned} &P(J \wedge M \wedge A \wedge \neg L \wedge \neg T) \\ &= P(J | A)P(M | A)P(A | \neg L \wedge \neg T)P(\neg L) P(\neg T) \\ &= 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998 \\ &= 0.00062 \end{aligned}$$

— Inferência em Redes Bayesianas

- **Definição:** Processo de extrair conhecimento representado na rede
- **Base:** Utilizar evidências observadas para calcular probabilidades de eventos não observados



— Inferência em Redes Bayesianas

- Tipos de Inferência

1. Inferência Diagnósticos: partindo dos efeitos para as causas;

Ex: Observar que o alarme disparou → inferir probabilidade de assalto

2. Inferência Preditiva: partindo das causas para os efeitos;

Ex: Saber que houve assalto → inferir probabilidade do alarme disparar



— Inferência em Redes Bayesianas

- Tipos de Inferência

3. **Inferência Intercasual:** entre causas de um efeito comum;

Ex: Saber que houve assalto → como isso afeta a probabilidade de terremoto?

4. **Inferência Mista:** combinação de dois ou mais tipos de inferência

Cenários mais realistas e complexos



— Aplicação

- Prever enchentes em São Luís



- Análise Exploratória dos Dados

❑ Utilização da IA para geração do dados para evitar conflitos

❑ **Fonte:** Dataset sao_luis_enchentes_limpo.csv

❑ **Estrutura Inicial dos Dados:**

- **Registros:** 52 linhas
- **Variáveis:** 8 colunas
- **Principais Variáveis:**
 - bairro (Localização)
 - topografia
 - enchente
 - drenagem

- Tabela para pré-visualização

Essa tabela mostra alguns dados retirados dos Datasets

	precipitacao_mm	intensidade_chuva	duracao_chuva	mare	bairro	topografia	drenagem	enchente
0	85.2	intensa	6h+	alta	Centro	baixa	inadequada	True
1	45.1	moderada	2-6h	média	Vinhais	média	adequada	False
2	92.7	intensa	6h+	alta	João Paulo	baixa	inadequada	True
3	23.5	fraca	<2h	baixa	Renascença	alta	adequada	False
4	67.8	intensa	2-6h	média	Cohama	baixa	regular	True
5	15.2	fraca	<2h	baixa	Alemanha	média	regular	False
6	76.3	intensa	6h+	alta	São Francisco	baixa	inadequada	True
7	38.9	moderada	2-6h	baixa	Calhau	baixa	regular	False
8	55.1	intensa	2-6h	alta	Centro	baixa	inadequada	True
9	29.7	fraca	2-6h	média	Vinhais	média	adequada	False

- Análise Exploratória dos Dados

❑ Limpeza dos dados

```
▶ # Conta o número de valores nulos por coluna  
print(df.isnull().sum())
```

```
⇒ precipitacao_mm      0  
   intensidade_chuva    0  
   duracao_chuva        0  
   mare                 0  
   bairro                0  
   topografia            0  
   drenagem              0  
   enchente              0  
   dtype: int64
```

```
[ ] # Conta o número de linhas duplicadas  
duplicados = df.duplicated().sum()  
print(f"Número de linhas duplicadas: {duplicados}")  
  
# Se houver, você pode removê-las  
if duplicados > 0:  
    df = df.drop_duplicates()  
    print("Linhas duplicadas foram removidas.")
```

```
⇒ Número de linhas duplicadas: 0
```

- Análise Exploratória dos Dados

☐ Análise Descritiva

```
▶ # Gera estatísticas descritivas para colunas numéricas  
print(df.describe())
```

```
⇒
```

	precipitacao_mm
count	52.000000
mean	51.494231
std	27.352911
min	8.300000
25%	28.000000
50%	48.750000
75%	75.700000
max	102.400000

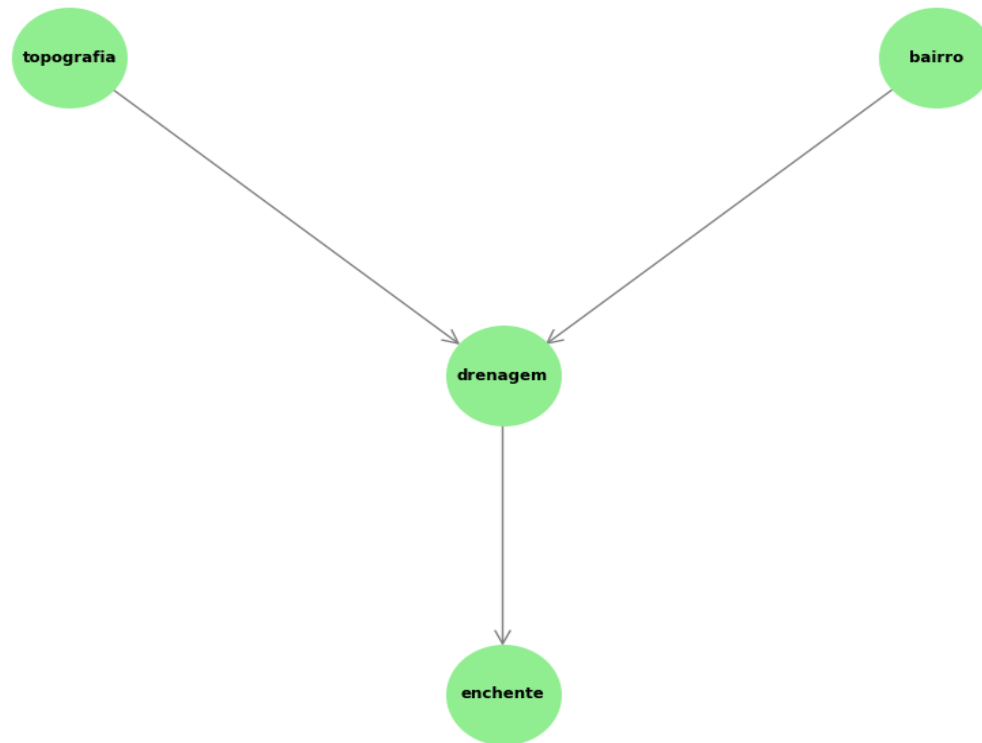
```
▶ # Conta quantas vezes cada bairro aparece  
if 'bairro' in df.columns:  
    print(df['bairro'].value_counts())  
  
# Conta quantas casas alagaram ou não  
if 'alagou_casa' in df.columns:  
    print(df['alagou_casa'].value_counts())
```

```
⇒
```

bairro	
Centro	8
Vinhais	7
João Paulo	7
Renascença	6
Cohama	6
Alemanha	6
São Francisco	6
Calhau	6

Name: count, dtype: int64

Rede Bayesiana: Fatores Geográficos que Afetam Enchentes em São Luís/MA



Estrutura descoberta automaticamente dos dados

Estimação das Tabelas de Probabilidades Condicionais (CPDs)

CPD para a variável: drenagem

bairro	... bairro(Vinhais)	
+-----+-----+		
topografia	... topografia(média)	
+-----+-----+		
drenagem(adequada)	... 1.0	
+-----+-----+		
drenagem(inadequada)	... 0.0	
+-----+-----+		
drenagem(regular)	... 0.0	
+-----+-----+		

CPD para a variável: enchente

drenagem	... drenagem(regular)	
+-----+-----+		
enchente(False)	... 0.7222222222222222	
+-----+-----+		
enchente(True)	... 0.2777777777777778	
+-----+-----+		

Estimação das Tabelas de Probabilidades Condicionais (CPDs)

CPD para a variável: bairro

bairro(Alemanha)	0.115385
bairro(Calhau)	0.115385
bairro(Centro)	0.153846
bairro(Cohama)	0.115385
bairro(João Paulo)	0.134615
bairro(Renasença)	0.115385
bairro(São Francisco)	0.115385
bairro(Vinhais)	0.134615

CPD para a variável: topografia

topografia(alta)	0.115385
topografia(baixa)	0.634615
topografia(média)	0.25

- Validação da Rede Bayesiana

☐ Método Hold-Out

- Separa parte dos dados para treino (70%) e parte para teste (30%)
- Treina a rede com os dados de treino
- Usa os dados de teste para ver se a rede acerta as previsões
- Calcula métricas para medir esse desempenho

- Validação da Rede Bayesiana

❑ Método Hold-Out

```
WARNING:pgmpy:Replacing existing CPD for drenagem
WARNING:pgmpy:Replacing existing CPD for enchente
WARNING:pgmpy:Replacing existing CPD for bairro
WARNING:pgmpy:Replacing existing CPD for topografia
WARNING:pgmpy:Replacing existing CPD for drenagem
WARNING:pgmpy:Replacing existing CPD for enchente
WARNING:pgmpy:Replacing existing CPD for bairro
WARNING:pgmpy:Replacing existing CPD for topografia
```

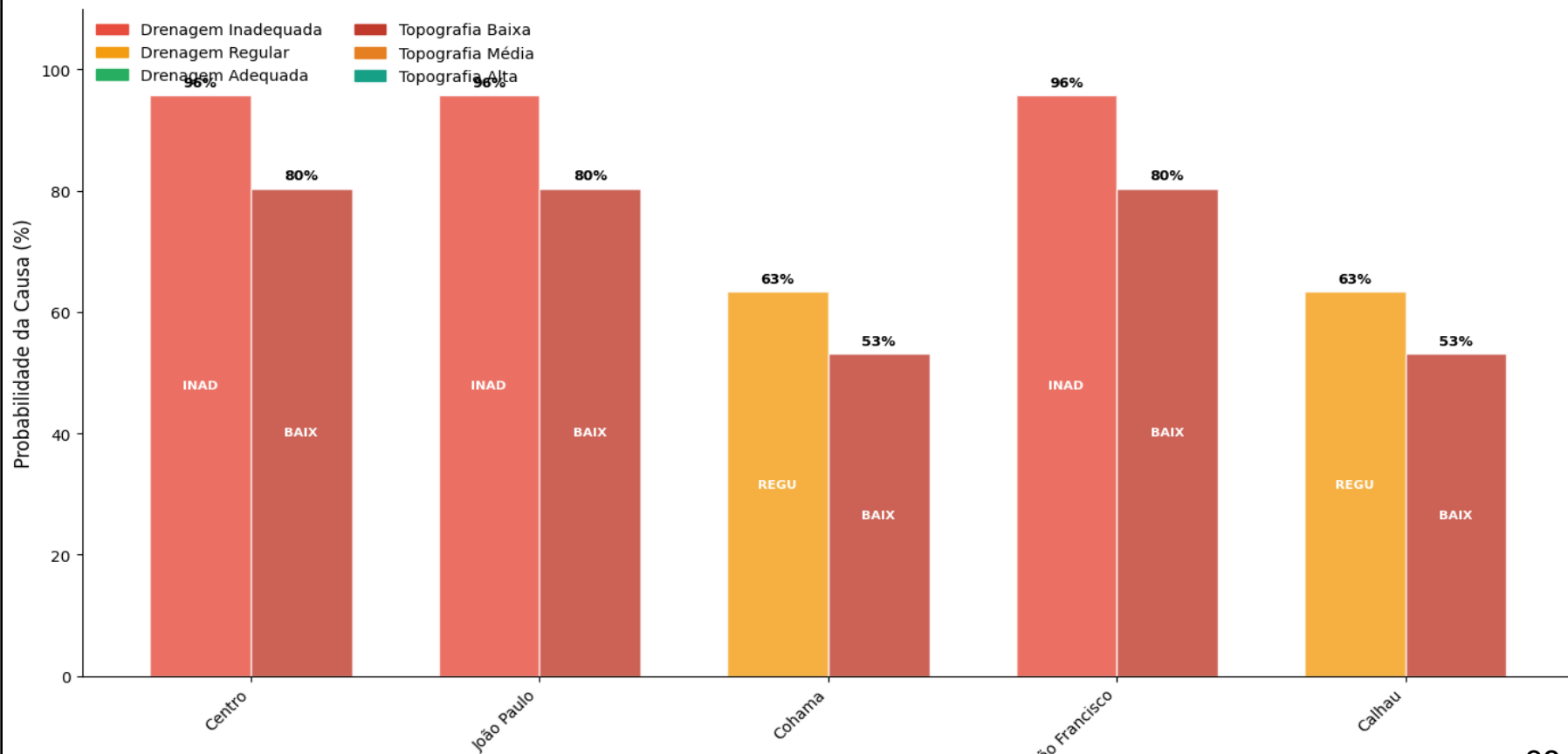
Desempenho do Modelo (Hold-out):

ROC AUC: 0.969

Matriz de Confusão:

```
[[8 0]
 [1 7]]
```

DIAGNÓSTICO: Por que houve enchente?



Bairros com Enchente Observada

INFERÊNCIAS:

Consulta 01: Probabilidade de enchente dado bairro='Centro', topografia='baixa', e drenagem 'inadequada'

+-----+	+-----+
enchente	phi(enchente)
+=====+	+=====+
enchente(False)	0.0000
+-----+	+-----+
enchente(True)	1.0000
+-----+	+-----+

INFERÊNCIAS:

Consulta 02: Probabilidade dos estados de drenagem dado bairro='Vinhais', topografia='média', e enchente=False (não houve enchente)

+-----+	+-----+
drenagem	phi(drenagem)
+=====+	+=====+
drenagem(adequada)	1.0000
+-----+	+-----+
drenagem(inadequada)	0.0000
+-----+	+-----+
drenagem(regular)	0.0000
+-----+	+-----+

INFERÊNCIAS:

Consulta 03: Probabilidade de enchente dado bairro='Calhau' e topografia='baixa' (o estado da drenagem não é fornecido como evidência e é inferido pelo modelo)

Situação 1: Drenagem Inferida

+-----+-----+	
enchente	phi(enchente)
+=====+=====+	
enchente(False)	0.7222
+-----+-----+	
enchente(True)	0.2778
+-----+-----+	

INFERÊNCIAS:

Consulta 04: Probabilidade de enchente dado bairro='Calhau', topografia='baixa', e drenagem='inadequada' (o estado da drenagem é fornecido diretamente)

Situação 2: Drenagem Observada como 'Inadequada'

+-----+-----+	
enchente	phi(enchente)
+=====+	
enchente(False)	0.0000
+-----+	
enchente(True)	1.0000
+-----+	

– Referências

"IA – Redes Baysianas", Site – YouTube. Disponível em:

<<https://www.youtube.com/watch?v=PGaxIKVP4ao&t=1155s>>. Acesso em 02 de jun de 2025.

MARQUES, Roberto Ligeiro; DUTRA, Inês. **Redes Bayesianas: o que são, para que servem, algoritmos e exemplos de aplicações**. Rio de Janeiro: COPPE/UFRJ, 2003. Disponível em: <https://www.cos.ufrj.br/~ines/courses/cos740/leila/cos740/Bayesianas.pdf>. Acesso em: 2 jun. 2025.

GONÇALVES, André Ricardo. **Redes Bayesianas**. Campinas: Universidade Estadual de Campinas, 2009. Disponível em: <https://andreric.github.io/files/pdfs/bayesianas.pdf>. Acesso em: 2 jun. 2025.

Obrigado!



CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**

Please keep this slide for attribution