

APRENDIZADO DE MÁQUINAS

Redes Bayesianas para classificação

TÓPICOS

- 1. Conceito de Redes Bayesianas**
- 2. Exemplo de aplicação de Redes Bayesianas**
- 3. Métodos de construção**
- 4. Tipos de conhecimentos em Redes Bayesianas**
- 5. Tipos de inferência em Redes Bayesianas**
- 6. Aprendizagem em Redes Bayesianas**
- 7. Classificadores em Redes Bayesianas**
- 8. Considerações Finais**

REDES BAYESIANAS - CONCEITO

Estrutura de dados para representar as dependências entre variáveis e fornecer uma especificação concisa de qualquer distribuição de probabilidade conjunta total.

Composição:

- um conjunto de nós, um para cada variável aleatória
- grafo direcionado e acíclico (seta = "influência direta")
- cada nó tem uma distribuição condicional $P(X_i | \text{Pais}(X_i))$ que quantifica o efeito dos pais sobre o nó

No caso mais simples, a distribuição condicional é representada como uma tabela de probabilidade condicional (TPC), dada uma distribuição sobre X_i para cada combinação de valores dos pais.



REDES BAYESIANAS - EXEMPLO

“Você possui um novo alarme contra ladrões em casa. Este alarme é muito confiável na detecção de ladrões, entretanto, ele também pode disparar caso ocorra um terremoto. Você tem dois vizinhos, João e Maria, os quais prometeram telefonar-lhe no trabalho caso o alarme dispare. João sempre liga quando ouve o alarme, entretanto, algumas vezes confunde o alarme com o telefone e também liga nesses casos. Maria, por outro lado, gosta de ouvir música alta e às vezes não escuta o alarme.”

Estados



Ladrão



Terremoto



João



Fonte: br.freepik

Alarme



Maria

A topologia da rede reflete conhecimento "causal":

- Um roubo pode ligar o alarme
- Um terremoto pode ligar o alarme
- O alarme faz Maria telefonar
- O alarme faz João telefonar

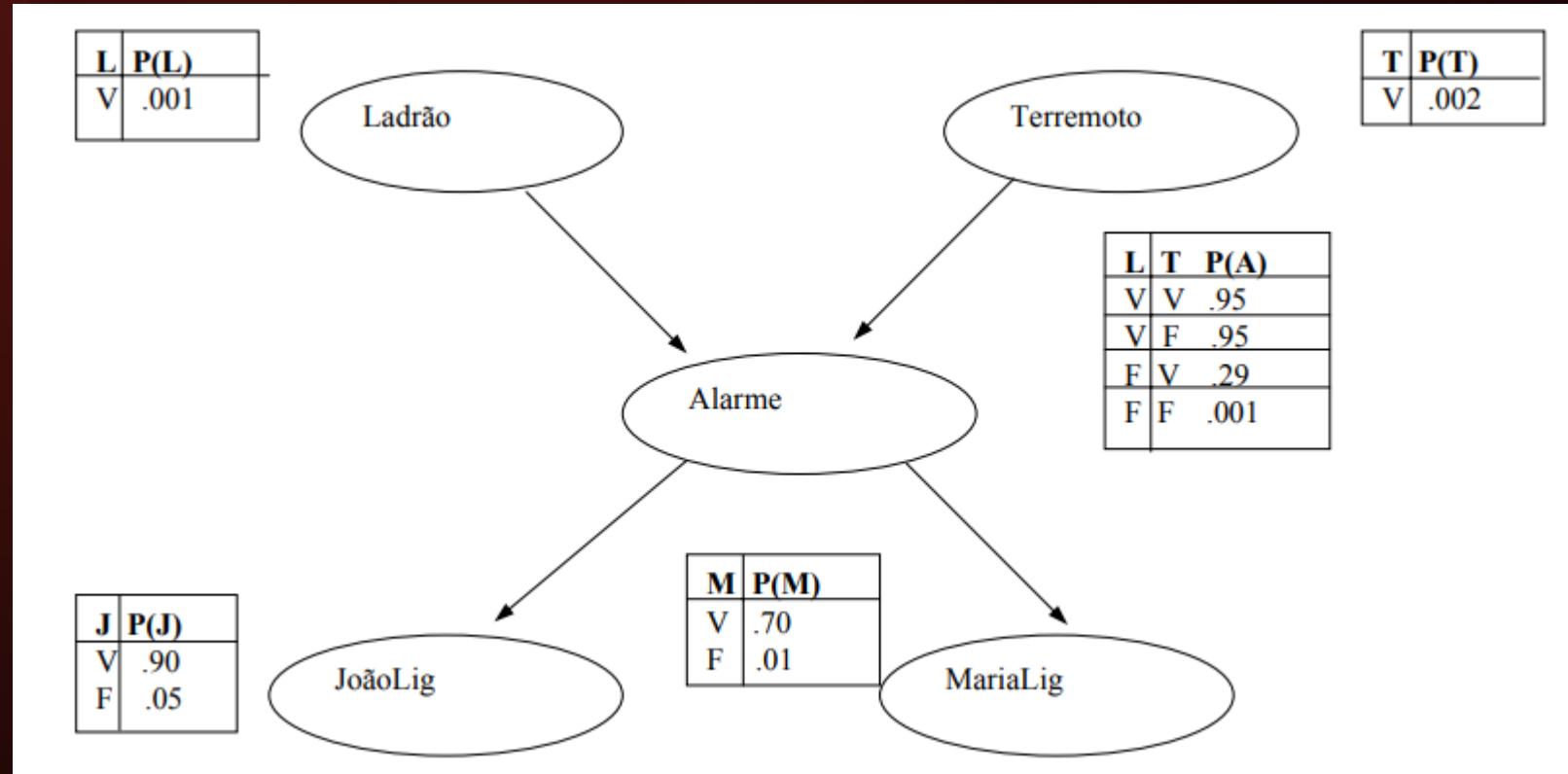
REDES BAYESIANAS - EXEMPLO

Se conhecemos a probabilidade da ocorrência de um ladrão e de um terremoto e ainda, a probabilidade de João e Maria telefonarem.

Podemos Calcular $P(\text{Alarme}|\text{Ladrão}, \text{Terremoto})$:

Ladrão	Terremoto	$P(\text{Alarme} \text{Ladrão}, \text{Terremoto})$	
		Verdadeiro	Falso
Verdadeiro	Verdadeiro	0.95	0.050
Verdadeiro	Falso	0.95	0.050
Falso	Verdadeiro	0.29	0.71
Falso	Falso	0.001	0.999

REDES BAYESIANAS - EXEMPLO



Roubos e terremotos afetam diretamente a probabilidade do alarme tocar; mas o fato de João e Maria telefonarem só depende do alarme. Desse modo, a rede representa nossas suposições de que eles não percebem quaisquer roubos diretamente, não notam os terremotos e não verificam antes de ligar!

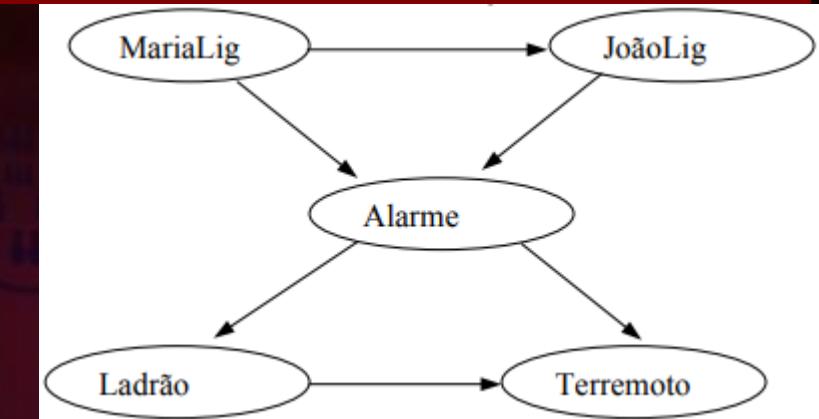
REDES BAYESIANAS - EXEMPLO

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 $P(J \wedge M \wedge A \wedge \neg L \wedge \neg T)$.

$$P(J \wedge M \wedge A \wedge \neg L \wedge \neg T) = P(J|A)P(M|A)P(A|\neg L \wedge \neg T)P(\neg L)P(\neg T) = 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998 = 0.00062$$

REDES BAYESIANAS – MÉTODO DE CONSTRUÇÃO

- Escolha um conjunto de variáveis X_i que descrevam o domínio.
- Escolha uma ordem para as variáveis.
- Enquanto existirem variáveis:
 - Escolha uma variável X_i e adicione um nó n a rede.
 - Determine os nós $Pais(X_i)$ dentre os nós que já estejam na rede e que tenham influência direta em X_i .
 - Defina a tabela de probabilidades condicionais para X_i



- MariaLig: raiz.
- JoãoLig: Se Maria ligou, então, provavelmente, o alarme tocou. Neste caso, MariaLig influencia JoaoLig. Portanto, MariaLig é pai de JoãoLig.
- Alarme: Claramente, se ambos ligaram, provavelmente o alarme tocou. Portanto, Alarme é influenciado por JoãoLig e MariaLig.
- Ladrão: Influenciado apenas por Alarme.
- Terremoto: Se o alarme tocou, provavelmente, um terremoto pode ter acontecido. Entretanto, se existe um Ladrão, então as chances de um terremoto diminuem. Neste caso, Terremoto é influenciado por Ladrão e Alarme.

REDES BAYESIANAS – TIPOS DE CONHECIMENTO

- Causal

Refletem a direção conhecida de causalidade no mundo: para algumas propriedades do mundo percepções são geradas.

Exemplo: $P(\text{DorDeDente}|\text{Cárie})$, $P(\text{MariaLiga}|\text{Alarme})$

- Diagnóstico

Infere a presença de propriedades escondidas diretamente da percepção.

Produzem conclusões fracas.

Exemplo: $P(\text{Cárie}|\text{DorDeDente})$, $P(\text{Alarme}|\text{MariaLiga})$

REDES BAYESIANAS – TIPOS DE INFERÊNCIA

Causal (da causa para o efeito)

- $P(\text{JoãoLiga/Roubo}) = 0,86$



Diagnóstico (do efeito para a causa)

- $P(\text{Roubo/JoãoLiga}) = 0,016$



REDES BAYESIANAS – TIPOS DE INFERÊNCIA

Intercausal (entre causas com um efeito comum)

$$P(\text{Roubo}/\text{Alarme}) = 0,376$$

$$P(\text{Roubo}/\text{Alarme} \wedge \text{Terremoto}) = 0,373$$



Mista (combinando duas ou mais das de cima)

$$P(\text{Alarme}/\text{JoãoLiga} \wedge \neg \text{Terremoto}) = 0,03$$

Este é um uso simultâneo de inferência causal e diagnóstico



REDES BAYESIANAS – APRENDIZAGEM

A aprendizagem Bayesiana pode ser vista como uma forma de obter a representação interna da rede que define um dado domínio, de modo a facilitar a extração do conhecimento.

O processo de aprendizagem é dividido em duas partes:

aprendizagem da estrutura (e relações entre as variáveis);

aprendizagem dos parâmetros numéricos (distribuição de probabilidade);

Formas de aprendizagem:

- Especialista
- Indutiva

Aprendizagem dos parâmetros:

- podem ser estimados através das frequências relativas, caso exista uma quantidade de dados significativa de uma amostra aleatória

Aprendizagem de estrutura:

- Métodos de Verossimilhança Máxima;
- Métodos de Teste de Hipóteses;
- Métodos de Verossimilhança Estendidos;
- Métodos "Minimum Information Complexity";
- Métodos "Resampling";
- Métodos Bayesianos, destacando o clássico algoritmo K2 (COOPER; HERSKOVITS, 1992).

REDES BAYESIANAS – CLASSIFICADORES

Classificadores bayesianos com k-dependências são aqueles em que cada atributo tem no máximo k nós atributos como pais.

■ Naive Bayes:

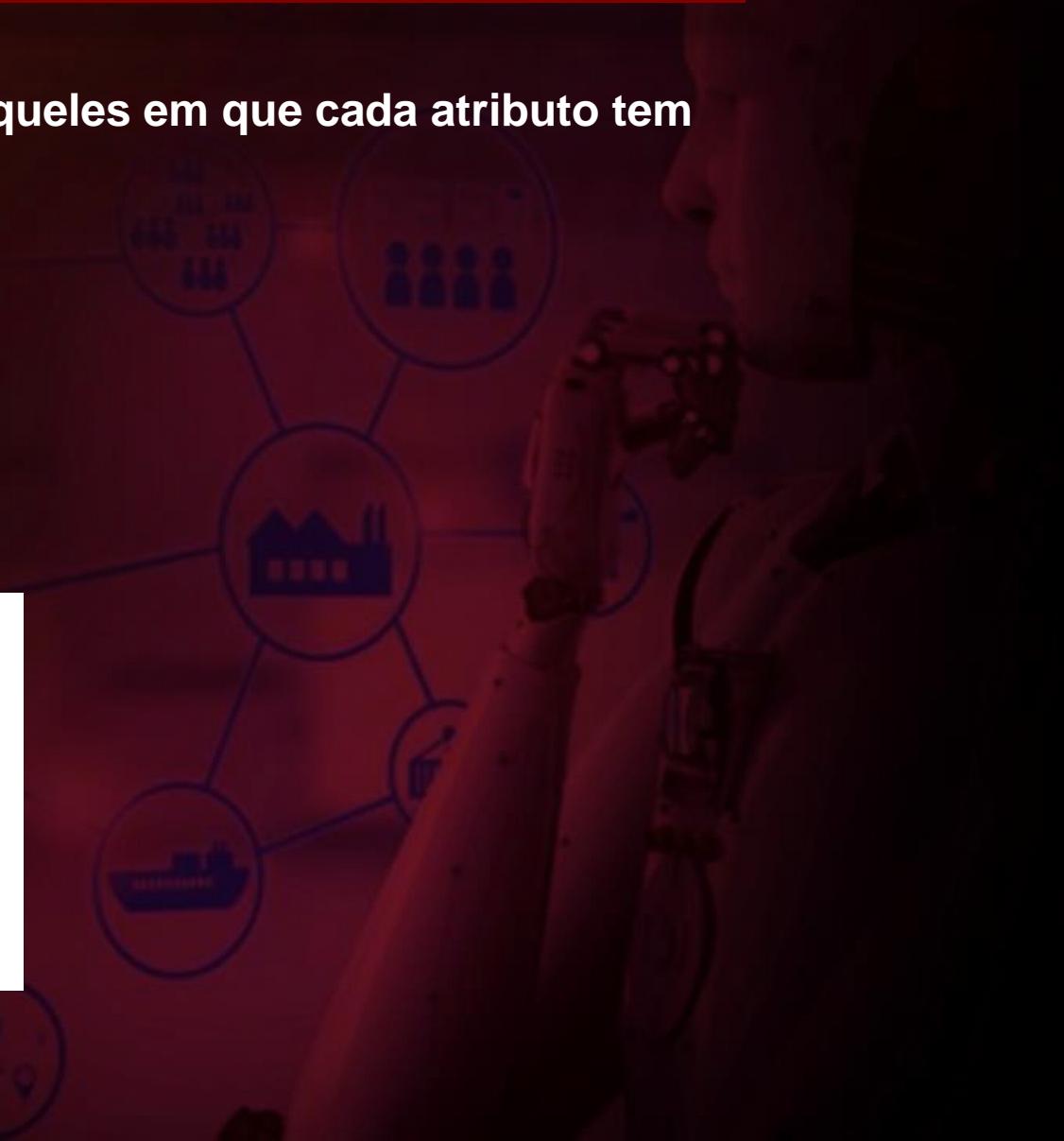
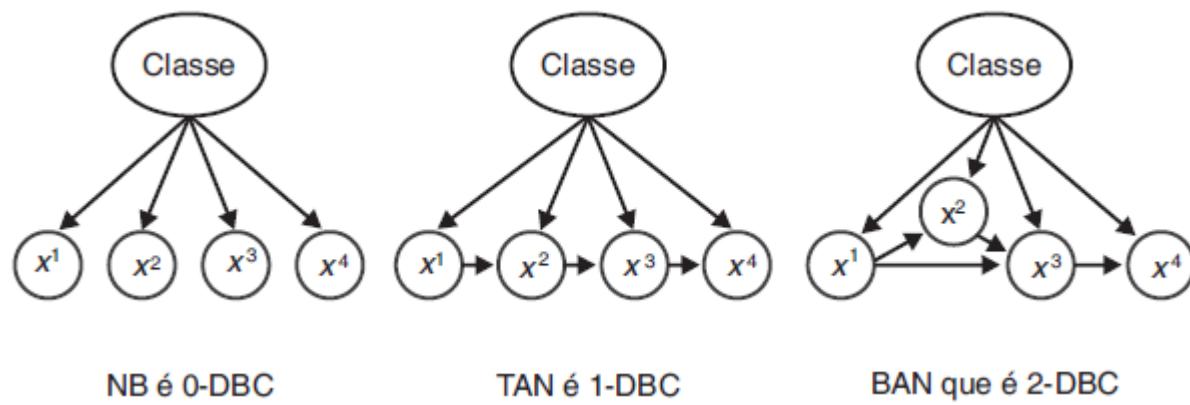
$$P(\text{Classe})P(x^1|\text{Classe})P(x^2|\text{Classe})P(x^3|\text{Classe})P(x^4|\text{Classe})$$

■ TAN:

$$P(\text{Classe})P(x^1|\text{Classe})P(x^2|x^1, \text{Classe})P(x^3|x^2, \text{Classe})P(x^4|x^3, \text{Classe})$$

■ BAN:

$$P(\text{Classe})P(x^1|\text{Classe})P(x^2|x^1, \text{Classe})P(x^3|x^1, x^2, \text{Classe})P(x^4|x^3, \text{Classe})$$



REDES BAYESIANAS – CONSIDERAÇÕES

- Redes bayesianas constituem uma forma natural para representação de informações condicionalmente independentes
- Boa solução para problemas onde conclusões não podem ser obtidas apenas do domínio do problema
- Combina conhecimento *a priori* com dados observados
- O impacto do conhecimento *a priori* (quando correto) é a redução da amostra de dados necessários
- Inferências sobre redes bayesianas
 - Podem ser executadas em tempo linear
 - NP-hard para maioria dos casos
 - Aplicação de técnicas



APRENDIZADO DE MÁQUINAS

Redes Bayesianas para classificação