Sistemas Inteligentes

Incerteza

Incerteza

- Seja a ação A_t = sair para o aeroporto t minutos antes do vôo.
- A_t me levará ao aeroporto a tempo?
- Dificuldades de saber o resultado da ação:
  - Estados parcialmente observáveis
    - Estados das estradas, trânsito, etc.
  - Sensores ruidosos
    - Relatórios de trânsito
  - Incerteza quanto ao efeito das ações
    - Acidentes, pneu furado, etc.
  - Grande complexidade em prever e modelar o trânsito

Incerteza

- Um procedimento puramente lógico não é muito útil nesse caso, porque:

1. Arriscaria deduzir algo potencialmente falso
   - “A_{45} me levará a tempo ao aeroporto”
2. Levaria a conclusões fracas para tomada de decisões
   - “A_{45} me levará a tempo ao aeroporto, se nenhum acidente ocorrer na ponte, se não chover, se nenhum pneu furar, etc.”
3. Levaria a conclusões não práticas
   - “A_{1440} me levará a tempo ao aeroporto”

Lidando com a incerteza

- Probabilidade
  - Modela o **grau de crença** de um agente dadas as evidências disponíveis
    - "A_{25} chegará a tempo com probabilidade 0,04"
    - "A_{45} chegará a tempo com probabilidade 0,85"
    - "A_{60} chegará a tempo com probabilidade 0,95"

Probabilidade

- A probabilidade proporciona um meio para resumir a incerteza que vem de:
  - Preguiça = falha em enumerar todas as possíveis exceções à regra
  - Ignorância = falta de conhecimento sobre fatos relevantes, condições iniciais

Probabilidade

- Probabilidade subjetiva ou bayesiana
  - Estabelece o estado de crença do agente em uma sentença, dadas as evidências.
  - Muda quando novas evidências chegam
    - $P(A_{25}|\text{nenhum acidente}) = 0{,}06$
    - $P(A_{25}|\text{nenhum acidente, 5 a.m.}) = 0{,}15$

- As sentenças são verdadeiras ou falsas.
  - O que muda é o grau de crença do agente na sentença.
  - Atribuir probabilidade 0 a uma sentença significa acreditar que ela é falsa com certeza absoluta.
  - Atribuir probabilidade 1 a uma sentença significa acreditar que ela é verdadeira com certeza absoluta.

Decisões sob incerteza

- Suponha o seguinte conjunto de crenças:
  - $P(A_{25}$ chega a tempo | ...$) = 0{,}04$
  - $P(A_{90}$ chega a tempo | ...$) = 0{,}70$
  - $P(A_{120}$ chega a tempo | ...$) = 0{,}95$
  - $P(A_{1440}$ chega a tempo | ...$) = 0{,}9999$

- Que ação o agente deve tomar?
  - Depende de suas **preferências** sob perder o vôo versus o tempo esperando no aeroporto.
    - Teoria da utilidade = representação de preferências
    - Teoria da decisão = teoria da probabilidade + teoria da utilidade

Agente de teoria da decisão

- Representação não apenas de *possibilidades*, mas de *probabilidades* dos estados do mundo.
- Esboço da estrutura de um agente que usa a teoria da decisão para selecionar ações:

> **função** AGENTE-TD(*percepção*) **retorna** uma *ação*
>   **variáveis estáticas:** *estado_de_crença*, crenças probabilísticas sobre o estado atual do mundo
>   *ação*, a ação do agente
>
>   atualizar *estado_de_crença* com base em *ação* e *percepção*
>   calcular probabilidades de resultados de ações,
>   dadas descrições de ações e o *estado_de_crença* atual
>   selecionar *ação* com utilidade esperada mais alta
>   dadas as probabilidades de resultados e informações de utilidade
>   **retornar** *ação*

Introdução à probabilidade

- Elemento básico: variável aleatória
  - Análogo à lógica proposicional
    - Mundos possíveis são definidos pela atribuição de valores às variáveis.
  - Cada variável aleatória tem um **domínio** que determina seus valores possíveis.
    - Tipos de domínio
      - **Booleano**, ex.: *Cárie* possui valores em <verdadeiro,falso>
      - **Discreto**, ex.: *Clima* possui valores em <ensolarado, chuvoso, nublado, neve>
      - **Contínuo**, ex.: *Temperatura*

Introdução à probabilidade

- Proposições elementares
  - São construídas por meio da atribuição de valores a variáveis.
  - Ex.: *Clima* = ensolarado, *Cárie* = falso (abreviado como $\neg$*Cárie*)
- Proposições complexas
  - São formadas a partir de proposições elementares e conectivos lógicos padrão
  - Ex.: *Clima* = ensolarado $\vee$ *Cárie* = falso

Introdução à probabilidade

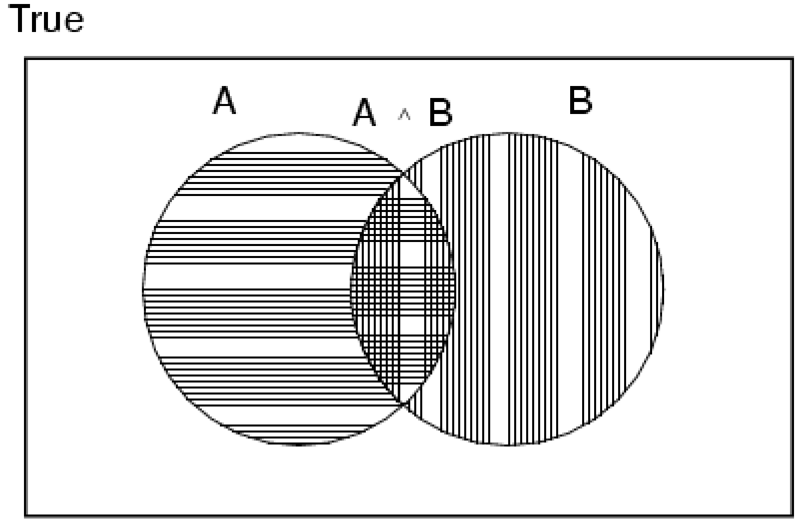
- Evento atômico
  - Especificação completa do estado do mundo sobre o qual o agente está incerto.
    - Uma atribuição de valores a TODAS as variáveis das quais o mundo é formado.
    - Eventos atômicos são mutuamente exclusivos e exaustivos.

Evento atômico: exemplo

- Se o mundo consistir somente de 2 variáveis booleanas (*Cárie* e *DorDeDente*), então há 4 eventos atômicos distintos:
  - *Cárie* = verdadeiro $\wedge$ *DorDeDente* = verdadeiro
  - *Cárie* = verdadeiro $\wedge$ *DorDeDente* = falso
  - *Cárie* = falso $\wedge$ *DorDeDente* = verdadeiro
  - *Cárie* = falso $\wedge$ *DorDeDente* = falso

Axiomas da Probabilidade

- Para quaisquer proposições *A, B*
  - $0 \leq P(A) \leq 1$
  - P(*verdade*) = 1 e P(*falso*) = 0
  - $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$



Probabilidade

- A probabilidade de uma proposição é igual à soma das probabilidades dos eventos atômicos em que ela é válida:

$$P(a) = \Sigma_{ei \in e(a)} P(e_i)$$

- Essa equação permite calcular a probabilidade de qualquer proposição dada uma distribuição conjunta total que especifique todos os eventos atômicos.

Probabilidade incondicional ou “a priori”

- É o grau de crença em uma proposição na ausência de outras informações.
  - Exemplos:
    - P(*Cárie* = verdadeiro) = 0,1
    - P(*Clima* = ensolarado) = 0,72
- Distribuição de probabilidades
  - Dá probabilidades a todos os valores possíveis de uma variável aleatória.
    - Climas: <ensolarado, chuvoso, nublado, neve>
    - P(*Clima*) = <0,72; 0,1; 0,08; 0,1> (normalizado, soma 1)

Distribuição de Probabilidade Conjunta

- Probabilidades de todas as combinações de valores de um conjunto de variáveis aleatórias.
  - P(*Clima*,*Cárie*) = tabela 4 x 2 de valores:

***Clima* =**	**ensolarado**	**chuvoso**	**nublado**	**neve**
Cárie = verdadeiro	0,144	0,02	0,016	0,02
Cárie = falso	0,576	0,08	0,064	0,08

- Uma distribuição conjunta total especifica a probabilidade de qualquer evento atômico.
  - Qualquer probabilidade nesse domínio pode ser calculada a partir da distribuição conjunta total.

Probabilidade condicional ou “a posteriori”

- É o grau de crença em uma proposição dada a presença de evidências (valores de variáveis aleatórias conhecidos).
- “|” é pronunciado como “considerando que”
  - Exemplos:
    - P(*Cárie* = verdadeiro | *DorDeDente* = verdadeiro) = 0,8
    - P(*Cárie* = verdadeiro | *DorDeDente* = verdadeiro, *Cárie* = verdadeiro) = 1
    - P(*Cárie* = verdadeiro | *DorDeDente* = verdadeiro, *Ensolarado* = verdadeiro) = P(*Cárie* = verdadeiro | *DorDeDente*) = 0,8

- Distribuição condicional
  - **P**(Y|X) fornece o valor de $P(Y=y_i \mid X=x_j)$ para cada valor de *i* e *j* possíveis.

Probabilidade Condicional

- Definição de probabilidade condicional:

  $P(a \mid b) = P(a \wedge b) / P(b)$ se $P(b) > 0$

- Regra do produto como formulação alternativa:

  $P(a \wedge b) = P(a \mid b)\, P(b) = P(b \mid a)\, P(a)$

- Pode ser generalizado para distribuições totais,

  $\mathbf{P}(\mathit{Clima}, \mathit{Cárie}) = \mathbf{P}(\mathit{Clima} \mid \mathit{Cárie})\, \mathbf{P}(\mathit{Cárie})$

  (um conjunto de 4 × 2 equações)

- Regra da cadeia é obtida a partir de aplicações sucessivas da regra do produto:

$$\begin{aligned}
\mathbf{P}(X_1, \ldots, X_n) &= \mathbf{P}(X_1, \ldots, X_{n-1})\, \mathbf{P}(X_n \mid X_1, \ldots, X_{n-1}) \\
&= \mathbf{P}(X_1, \ldots, X_{n-2})\, \mathbf{P}(X_{n-1} \mid X_1, \ldots, X_{n-2})\, \mathbf{P}(X_n \mid X_1, \ldots, X_{n-1}) \\
&= \ldots \\
&= \pi_{i=1}^{n}\, \mathbf{P}(X_i \mid X_1, \ldots, X_{i-1})
\end{aligned}$$

Inferência Probabilística

- Inferência probabilística: a computação a partir de evidências observadas de probabilidades posteriores para proposições de consulta.

- Inferência com o uso de distribuições conjuntas totais: base de conhecimento a partir da qual são derivadas respostas para todas as consultas.

Exemplo: Inferência Probabilística

- Suponha um domínio com a seguinte distribuição conjunta total: :

	dordedente		*¬dordedente*	
	boticão	*¬boticão*	*boticão*	*¬boticão*
cárie	.108	.012	.072	.008
¬cárie	.016	.064	.144	.576

- Para qualquer proposição φ, P(φ) é as soma dos eventos atômicos ω onde φ ocorre: $P(\varphi) = \Sigma_{\omega:\omega \models \varphi} P(\omega)$

Exemplo: Inferência Probabilística

- Suponha um domínio com a seguinte distribuição conjunta total:

	dordedente		*¬dordedente*	
	boticão	*¬boticão*	*boticão*	*¬boticão*
cárie	.108	.012	.072	.008
¬cárie	.016	.064	.144	.576

- Para qualquer proposição φ, P(φ) é as soma dos eventos atômicos ω onde φ ocorre: $P(\varphi) = \Sigma_{\omega:\omega \models \varphi} P(\omega)$

- P(*dordedente*) = 0,108 + 0,012 + 0,016 + 0,064 = 0,2

Exemplo: Inferência Probabilística

- Suponha um domínio com a seguinte distribuição conjunta total:

	dordedente		*¬dordedente*	
	boticão	*¬boticão*	*boticão*	*¬boticão*
cárie	.108	.012	.072	.008
¬cárie	.016	.064	.144	.576

- Para qualquer proposição φ, P(φ) é as soma dos eventos atômicos ω onde φ ocorre: $P(\varphi) = \Sigma_{\omega:\omega \models \varphi} P(\omega)$

P(*dordedente* ∨ *cárie*) = 0,108 + 0,012 + 0,016 + 0,064 + 0,072 + 0,008 = 0,28

Exemplo: Inferência Probabilística

- Podemos calcular probabilidades condicionais:

	dordedente		*¬dordedente*	
	boticão	*¬boticão*	*boticão*	*¬boticão*
cárie	.108	.012	.072	.008
¬cárie	.016	.064	.144	.576

$$
\begin{aligned}
P(\neg \textit{cárie} \mid \textit{dordedente}) &= \frac{P(\neg \textit{cárie} \wedge \textit{dordedente})}{P(\textit{dordedente})} \\
&= \frac{0{,}016+0{,}064}{0{,}108 + 0{,}012 + 0{,}016 + 0{,}064} \\
&= 0{,}4
\end{aligned}
$$

Normalização
Inferência Probabilística

	dordedente		*¬dordedente*	
	boticão	*¬boticão*	*boticão*	*¬boticão*
cárie	.108	.012	.072	.008
¬cárie	.016	.064	.144	.576

- O denominador pode ser visto como uma constante de normalização α.

P(*Cárie*|*dordedente*) = α **P**(*Cárie*,*dordedente*)
= α [**P**(*Cárie*, *dordedente*, *boticão*) + **P**(*Cárie*, *dordedente*, *¬boticão*)]
= α [<0,108; 0,016> + <0,012; 0,064>]
= α [<0,12; 0,08>]
= <0,6; 0,4>

Inferência Probabilística: Inferência por Enumeração

- Objetivo: calcular a distribuição de probabilidade das **variáveis de consulta** X (*Cárie* no exemplo), dados valores para as **variáveis de evidência** E (*DorDeDente* no ex.).
  – Sejam Y as variáveis restantes, não observadas (*Boticão* no ex.), temos

$$\mathbf{P}(X|e) = \alpha \sum_y \mathbf{P}(X, e, y)$$

Note que cada **P**(X,e,y) aparece na distribuição conjunta total

Problemas com a inferência por enumeração

- Complexidade de tempo (pior caso): $\mathbf{O}(d^n)$

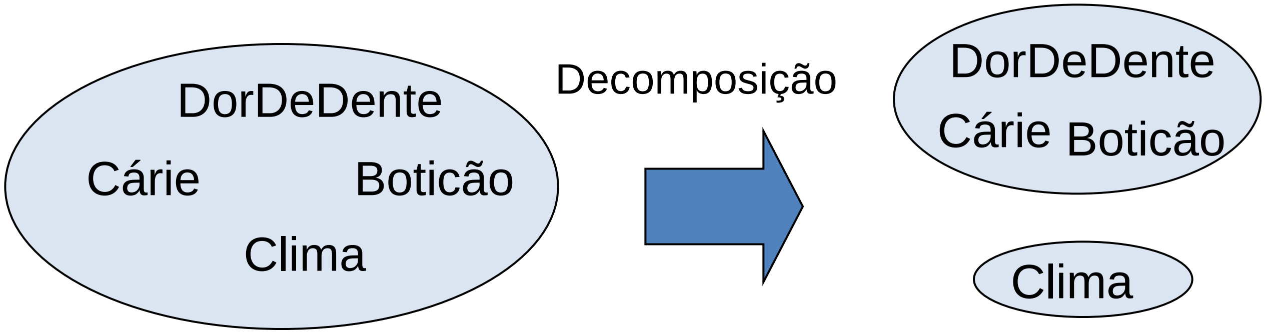
  onde *d* é a cardinalidade do maior domínio e *n* é o número de variáveis (se booleano, d=2).

- Complexidade de espaço: $\mathbf{O}(d^n)$ para armazenar a distribuição conjunta.
- Como encontrar as probabilidades para $\mathbf{O}(d^n)$ elementos?
- Em um problema realista, *n* pode ser maior que 100, e a forma tabular não é prática.

Independência

- A e B são independentes se e somente se

  $\mathbf{P}(A|B) = \mathbf{P}(A)$ ou $\mathbf{P}(B|A) = \mathbf{P}(B)$ ou $\mathbf{P}(\mathrm{A}, \mathrm{B}) = \mathbf{P}(A)\,\mathbf{P}(B)$



$\mathbf{P}(\text{DorDeDente,Cárie,Boticão,Clima})$
$= \mathbf{P}(\text{DorDeDente,Cárie,Boticão})\mathbf{P}(\text{Clima})$
(32 entradas reduzidas a 12).

- Porém, independência total é rara.

Independência Condicional

- Se eu tenho cárie, a probabilidade de uso do boticão não depende de eu ter ou não dor de dente.
  - **P**(Boticão|dordedente,cárie) = **P**(Boticão|cárie)

- A mesma independência ocorre se eu não tiver cárie.
  - **P**(Boticão|dordedente,¬cárie) = **P**(Boticão| ¬cárie)

- Logo Boticão é condicionalmente independente de DorDeDente dado cárie:
  - **P**(Boticão|DorDeDente,Cárie) = **P**(Boticão| Cárie)

Independência Condicional

- Escrevendo a distribuição total usando a regra da cadeia:
  P(DorDeDente,Boticão,Cárie)
  =P(DorDeDente|Boticão,Cárie)P(Boticão|Cárie)P(Cárie)
  =P(DorDeDente|Cárie)P(Boticão|Cárie)P(Cárie)

- Nesse exemplo, o número de valores para especificar a distribuição conjunta passa de 8 para 6.

- Na maioria dos casos, o uso da independência condicional reduz o tamanho da distribuição conjunta de exponencial em *n* para linear em *n*.

Regra de Bayes

- Da regra do produto $P(a \wedge b) =$

  $$P(a \mid b)\, P(b) = P(b \mid a)\, P(a)$$

  $\Rightarrow$ Regra de Bayes: $P(a \mid b) = P(b \mid a)\, P(a) / P(b)$

- Ou na forma da distribuição conjunta::

  – $\mathbf{P}(Y|X) = \mathbf{P}(X|Y)\, \mathbf{P}(Y) / \mathbf{P}(X) = \alpha \mathbf{P}(X|Y)\, \mathbf{P}(Y)$

- Útil para acessar regras probabilísticas de diagnóstico por meio de probabilidades causais:

  – P(Causa|Efeito) = P(Efeito|Causa) P(Causa) / P(Efeito)

Regra de Bayes

- P(Causa|Efeito) = P(Efeito|Causa) P(Causa) / P(Efeito)

- Exemplo:

  - O médico sabe que a meningite faz o paciente ter uma rigidez no pescoço em aproximadamente 70% do tempo

  - O médico conhece alguns fatos incondicionais:
    - Probabilidade *a priori* de um paciente com meningite: 1 / 50000
    - Probabilidade *a priori* de um paciente ter rigidez no pescoço: 1%

  - Seja *M* meningite, *S* rigidez no pescoço:

    $P(m|s) = P(s|m)\ P(m) / P(s) = 0{,}7 \times (1/50000) / 0{,}01 = 0{,}0014$

  - Nota: probabilidade posterior de meningite ainda é muito pequena (1 em 5000 pacientes com pescoço duro).

Regra de Bayes e Independência Condicional

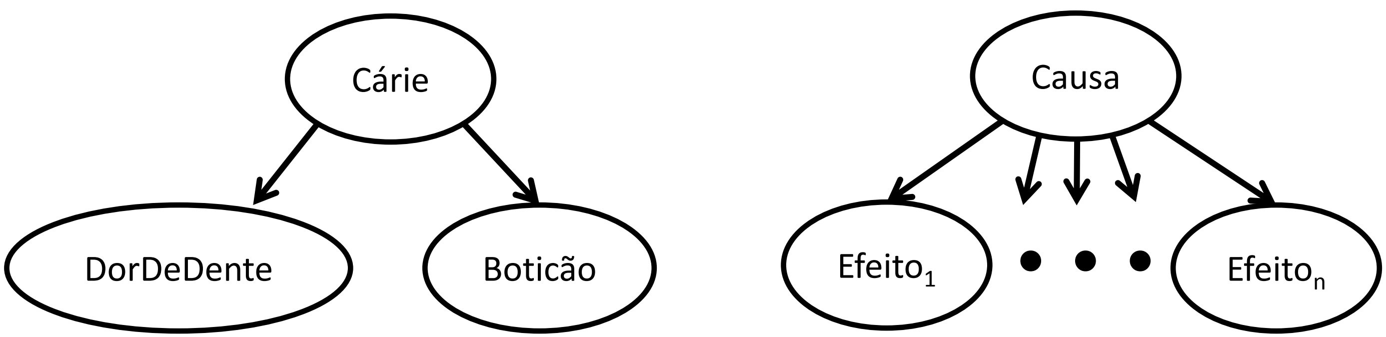
P(*Cárie* | *dordedente* $\wedge$ *boticão*)

= α **P**(*dordedente* $\wedge$ *boticão* | *Cárie*) **P**(*Cárie*)

= α **P**(*dordedente* | *Cárie*) **P**(*boticão* | *Cárie*) **P**(*Cárie*)

- Este é um exemplo de modelo de Bayes ingênuo (naive Bayes):

$$\mathbf{P}(\text{Causa}, \text{Efeito}_1, \ldots, \text{Efeito}_n) = \mathbf{P}(\text{Causa}) \prod_i \mathbf{P}(\text{Efeito}_i \mid \text{Causa})$$



- O número total de parâmetros é linear *n*

Classificação

Usando Naive Bayes

Apresentação baseada:

no livro Introduction to Data Mining (Tan, Steinbach, Kumar) e

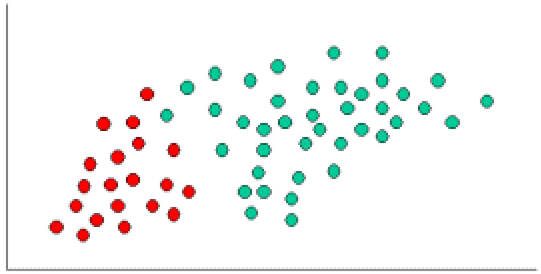
em apresentações dos Profs. José Todesco (UFSC) e Luis OtavioAlvares (UFSC)

Naive Bayes

- Abordagem estatística, baseada no teorema de Bayes. Naive (ingênuo) porque considera que os atributos são independentes.

Naive Bayes – visão geral

Seja o exemplo de dados:



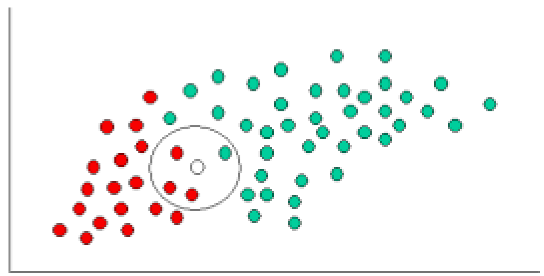
Os objetos podem ser classificados em vermelho ou verde

Como há mais objetos verdes que vermelhos, a probabilidade *a priori* é que um novo objeto seja verde

Probabilidade a priori de verde = número de objetos verdes/ número total de objetos = 40/60 = 4/6

Probabilidade a priori de vermelho = número de objetos vermelhos / número total de objetos = 20/60 = 2/6

Naive Bayes – visão geral



- Queremos classificar um novo objeto X (ponto branco)
- Como os objetos estão agrupados, é razoável considerar que quanto mais objetos de uma classe houver "parecidos" com X, maior a chance de X ser daquela classe.
- Vamos considerar o "parecido" pelo círculo na figura (estar dentro do círculo) e calcular a probabilidade:
- *Probabilidade de "parecido" dado que é verde = número de objetos verdes no círculo/ número total de verdes= 1/40*
- *Probabilidade de "parecido" dado que é vermelho = número de objetos vermelhos no círculo/ número total de vermelhos= 3/20*

Naive Bayes – visão geral

- Na análise Bayesiana, a classificação final é realizada considerando estas duas informações usando a probabilidade condicional do Teorema de Bayes:
- A probabilidade condicional de X ser verde dado que é “parecido” = probabilidade a priori de verde vezes *Probabilidade de “parecido” dado que é verde =*

  4/6 . 1/40 = 1/60
- *Analogamente,*
- A probabilidade condicional de X ser vermelho dado que é “parecido” = 2/6 . 3/20 = 1/20
- Portanto, a classe predita de X seria vermelho, pois é a maior probabilidade

Mais tecnicamente....

- Aprendizagem da classificação: qual é a probabilidade da classe dado um exemplo?
  - Evidência *E = exemplo (registro, com os valores dos atributos)*
  - Hipótese *H = valor da classe para o exemplo*
- Teorema de Bayes (1763):

$$P(H|E) = \frac{P(E|H).P(H)}{P(E)}$$

- Suposição do classificador bayesiano ingênuo: evidência pode ser separada em partes independentes (os atributos do exemplo)

$$P(E_1, E_2, ..., E_n | H) = P(E_1|H).P(E_2|H)... .P(E_n|H)$$

$$P(H|E) = \frac{P(E_1|H).P(E_2|H)... .P(E_n|H).P(H)}{P(E_1).P(E_2)... .P(E_n)}$$

Exemplo:Naive Bayes

Dia	Aspecto	Temperatura	Umidade	Vento	Decisão
1	Sol	Quente	Alta	Fraco	N
2	Sol	Quente	Alta	Forte	N
3	Nublado	Quente	Alta	Fraco	S
4	Chuva	Agradável	Alta	Fraco	S
5	Chuva	Fria	Normal	Fraco	S
6	Chuva	Fria	Normal	Forte	N
7	Nublado	Fria	Normal	Forte	S
8	Sol	Agradável	Alta	Fraco	N
9	Sol	Fria	Normal	Fraco	S
10	Chuva	Agradável	Normal	Fraco	S
11	Sol	Agradável	Normal	Forte	S
12	Nublado	Agradável	Alta	Forte	S
13	Nublado	Quente	Normal	Fraco	S
14	Chuva	Agradável	Alta	Forte	N

Exemplo: Naive Bayes

Qual será a *decisão* (valor da classe), se o dia estiver com sol, a temperatura fria, a umidade alta e o vento forte ?

P(Jogar = S | Aspecto = Sol, Temperatura = Fria, Umidade = Alta e Vento = Forte) = ?

P(Jogar = N | Aspecto = Sol, Temperatura = Fria, Umidade = Alta e Vento = Forte) = ?

Exemplo: Naive Bayes

$$P(H|E) = \frac{P(E_1|H).P(E_2|H)\ldots.P(E_n|H).P(H)}{P(E_1).P(E_2)\ldots.P(E_n)}$$

P(Jogar = S | Aspecto = Sol, Temperatura = Fria, Umidade = Alta e Vento = Forte) =

$$= \frac{P(Sol|S) * P(Fria|S) * P(Alta|S) * P(Forte|S) * P(S)}{P(Sol) * P(Fria) * P(Alta) * P(Forte)}$$

Exemplo: Naive Bayes

$P(Jogar = S) = 9/14$; $P(Jogar = N) = 5/14$;

$P(Aspecto = Sol \mid Jogar = S) = 2/9$;

$P(Aspecto = Sol \mid Jogar = N) = 3/5$;

$P(Temperatura = Fria \mid Jogar = S) = 3/9$;

$P(Temperatura = Fria \mid Jogar = N) = 1/5$;

$P(Umidade = Alta \mid Jogar = S) = 3/9$;

$P(Umidade = Alta \mid Jogar = N) = 4/5$;

$P(Vento = Forte \mid Jogar = S) = 3/9$;

$P(Vento = Forte \mid Jogar = N) = 3/5$;

Exemplo: Naive Bayes

P(Aspecto = Sol) = 5/14

P(Temperatura = Fria) = 4/14

P(Umidade = Alta) = 7/14

P(Vento = Forte) = 6/14

Exemplo: Naive Bayes

P(Jogar = S | Aspecto = Sol, Temperatura = Fria, Umidade = Alta e Vento = Forte) =

$$= \frac{P(\text{ Sol}|S) * P(\text{ Fria}|S) * P(\text{Alta}|S) * P(\text{Forte}|S) * P(S)}{P(\text{ Sol}) * P(\text{ Fria}) * P(\text{Alta}) * P(\text{Forte})} =$$

$= (2/9 * 3/9 * 3/9 * 3/9 * 9/14) / (5/14 * 4/14 * 7/14 * 6/14) =$

$= 0{,}0053 / 0{,}02186 =$ **0,242**

Exemplo: Naive Bayes

P(Jogar = N | Aspecto = Sol, Temperatura = Fria, Umidade = Alta e Vento = Forte) =

$$= \frac{P(\,Sol|N) * P(\,Fria|N) * P(Alta|N) * P(Forte|N) * P(N)}{P(\,Sol) * P(\,Fria) * P(Alta) * P(Forte)} =$$

= (3/5 * 1/5 * 4/5 * 3/5 * 5/14) / (5/14 * 4/14 * 7/14 * 6/14) =

= 0,0206 / 0,02186 = **0,942**

Como (J=N) 0,942 > (J=S) 0,242 **Então Jogar = Não**

O problema da *frequência zero*

- Se um valor de atributo nunca ocorrer para uma classe (como por exemplo Aspecto=nublado para a classe N)
  - A probabilidade será zero! P(nublado | N) = 0
  - A probabilidade a posteriori será zero, independentemente dos outros valores! P(N | E) = 0
- Solução: *Estimador de Laplace* ⇒ somar 1 à contagem de todas as combinações de classe e valor de atributo.
- Resultado: as probabilidades nunca serão zero!

Naive Bayes

- Vantagens:
  - Rápido
  - Bons resultados em dados reais
- Desvantagens:
  - Resultados não tão bons em problemas complexos

- Mozilla Thunderbird e Microsoft Outlook usam classificadores naive bayes para filtrar (marcar) emails que seriam spam

Redes Bayesianas: Cálculo da Probabilidade Conjunta

- Redes Bayesianas levam em consideração a **Independência Condicional** entre **subconjuntos** de variáveis

$$P(y_1 \wedge \ldots y_n) = P(y_n \mid y_1 \wedge \ldots y_{n-1}) \ldots P(y_2 \mid y_1)\, P(y_1)$$

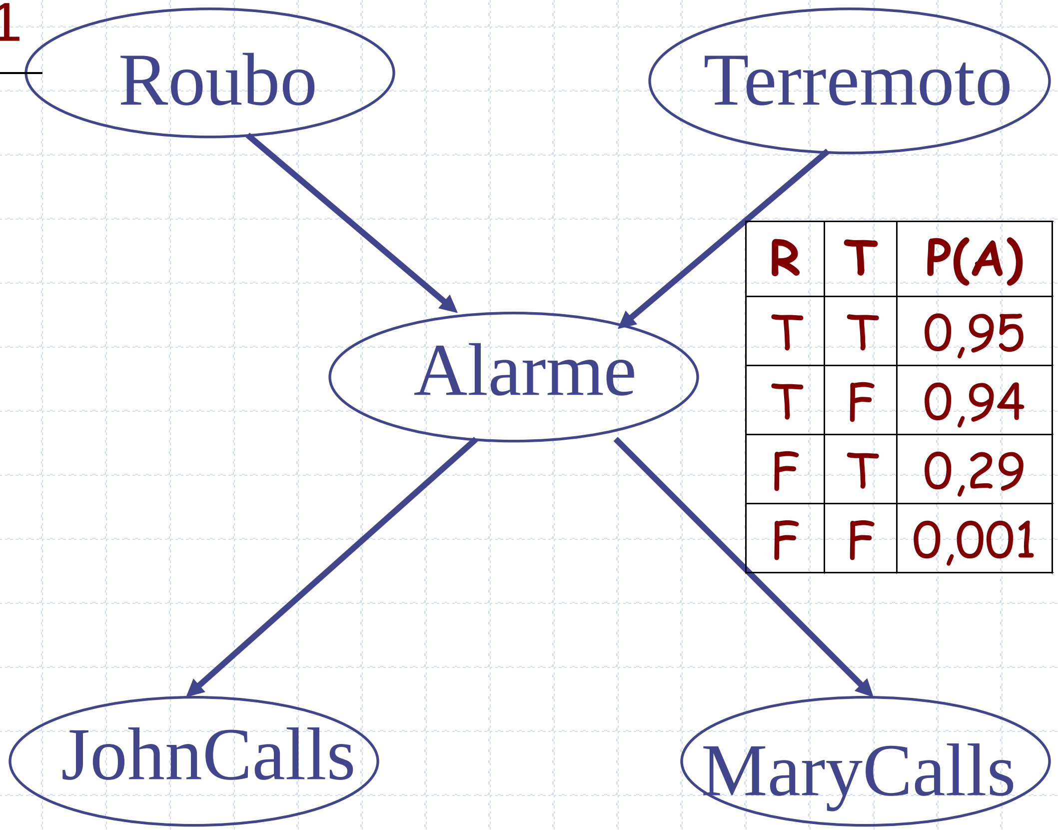
$P(y_n \mid \text{Predecessores}(Y_n))$ $P(y_2 \mid \text{Predecessores}(Y_2))$

$$P(y_1, \cdots, y_n) = \prod_{i=1}^{n} P(y_i \mid \text{Predecessores}(Y_i))$$

Exemplo do **alarme** para experimentar rede Baysiana no software Hugin

P(R)
0,001

P(T)
0,002



R	T	P(A)
T	T	0,95
T	F	0,94
F	T	0,29
F	F	0,001

A	P(J)
T	0,90
F	0,05

A	P(M)
T	0,70
F	0,01

Questão 2 da P2 (25%)

- Instalar o software Hugin
  - http://www.hugin.com/productsservices/demo/hugin-lite
  - (é preciso um pequeno registro; ao executar pela primeira vez, um código, recebido por e-mail, deve ser digitado)
- Testar uma rede pronta, alterando probabilidades para perceber a propagação. Há vários exemplos em:
  - <pasta de instalação>/Hugin Expert/Hugin Lite 8.1/Samples
- Criar uma rede qualquer, de preferência, sobre um assunto que tenha conhecimento ou interesse
- Entregar no Moodle o arquivo ".net" do Hugin e um arquivo PDF explicando a modelagem
  - Prazo: 19/11