

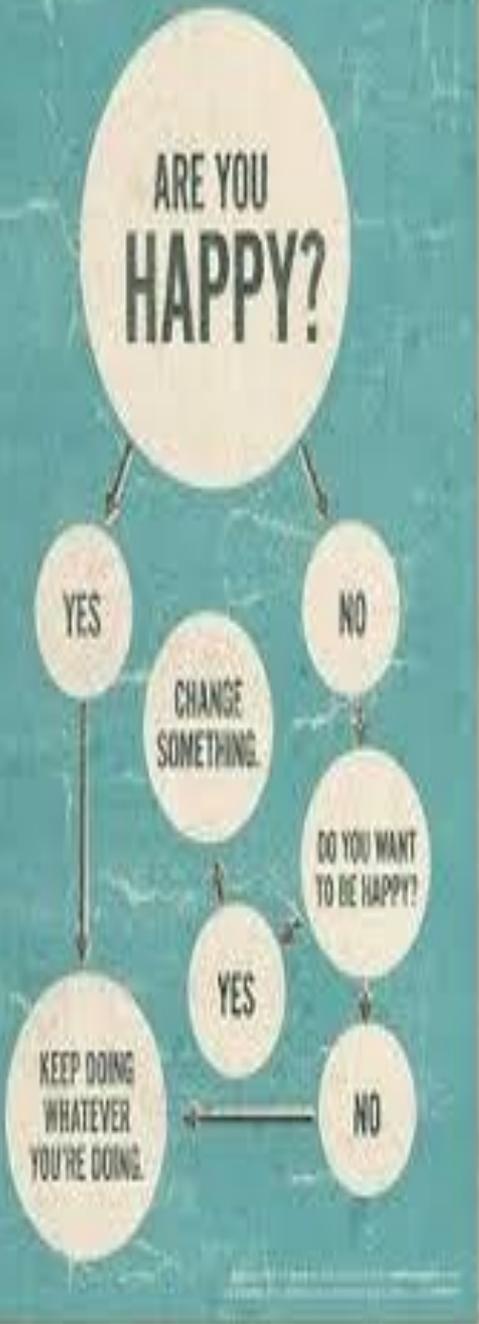
Universidade Federal de Campina Grande
Departamento de Sistemas e Computação
Pós-Graduação em Ciência da Computação

Inteligência Artificial

Aprendizagem – Parte II **(Árvores de Decisão - Exemplo)**

Prof.^a Joseana Macêdo Fachine Régis de Araújo
joseana@computacao.ufcg.edu.br

<https://www.pinterest.com/carlymundo/decision-tree-infographics/>



Aprendizagem

Tópico

- Aprendizagem
 - Árvores de Decisão – Exemplo



Aprendizagem - Exemplo

- Construa uma Árvore de decisão (Ad) para o seguinte problema: Decidir se vou Jogar Tênis.
 - Parâmetros do ambiente:
 - o Aspecto do Céu,
 - a Temperatura,
 - a Umidade e
 - o Vento.



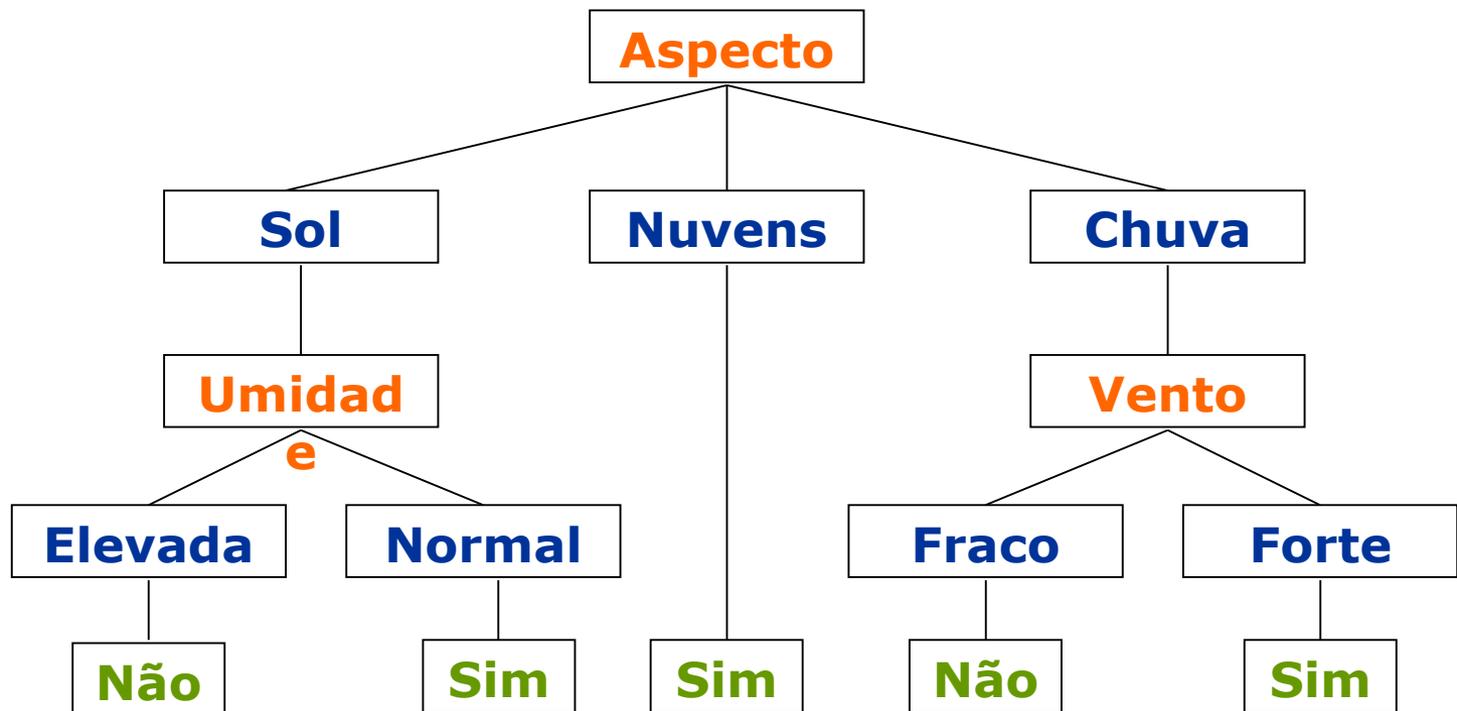
Aprendizagem - Exemplo

Exemplos de Treinamento

Dia	Aspecto	Temp.	Umidade	Vento	Jogar Tênis
D1	Sol	Quente	Elevada	Fraco	Não
D2	Sol	Quente	Elevada	Forte	Não
D3	Nuvens	Quente	Elevada	Fraco	Sim
D4	Chuva	Ameno	Elevada	Fraco	Sim
D5	Chuva	Fresco	Normal	Fraco	Sim
D6	Chuva	Fresco	Normal	Forte	Não
D7	Nuvens	Fresco	Normal	Fraco	Sim
D8	Sol	Ameno	Elevada	Fraco	Não
D9	Sol	Fresco	Normal	Fraco	Sim
D10	Chuva	Ameno	Normal	Forte	Sim
D11	Sol	Ameno	Normal	Forte	Sim
D12	Nuvens	Ameno	Elevada	Forte	Sim
D13	Nuvens	Quente	Normal	Fraco	Sim
D14	Chuva	Ameno	Elevada	Forte	Não

Aprendizagem - Exemplo

Solução:



Aprendizagem - Exemplo

Como construir a Ad?

Algoritmo ID3 (*inductive decision tree*) - um dos mais utilizados para a construção de Ad.

Passos do algoritmo:

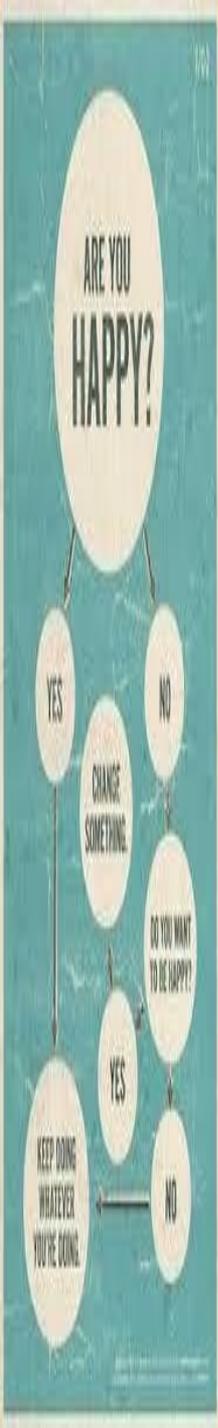
1. Começar com todos os exemplos de treino;
2. Escolher o teste (atributo) que melhor divide os exemplos, ou seja agrupar exemplos da mesma classe ou exemplos semelhantes;
3. Para o atributo escolhido, criar um nó filho para cada valor possível do atributo;
4. Transportar os exemplos para cada filho tendo em conta o valor do filho;
5. Repetir o procedimento para cada filho não "puro". Um filho é puro quando cada atributo X tem o mesmo valor em todos os exemplos.



Aprendizagem - Exemplo

Como escolher o melhor atributo?

- ❑ Entropia
- ❑ Ganho



Aprendizagem - Exemplo

Entropia

- ❑ A entropia de um conjunto pode ser definida como sendo uma medida do grau de impureza do conjunto.
- ❑ Este conceito define a medida de "falta de informação", mais precisamente o número de bits necessários, em média, para representar a informação em falta, usando codificação ótima.
- ❑ Entropia é uma medida da aleatoriedade de uma variável.



Aprendizagem - Exemplo

Entropia

Dado um conjunto S , com instâncias pertencentes à classe i , com probabilidade p_i , tem-se:

$$Entropia(S) = \sum p_i \log_2 p_i$$

Pode-se ter:

$$Entropia(S) = -p_+ \log_2 p_+ - p_- \log_2 p_-$$

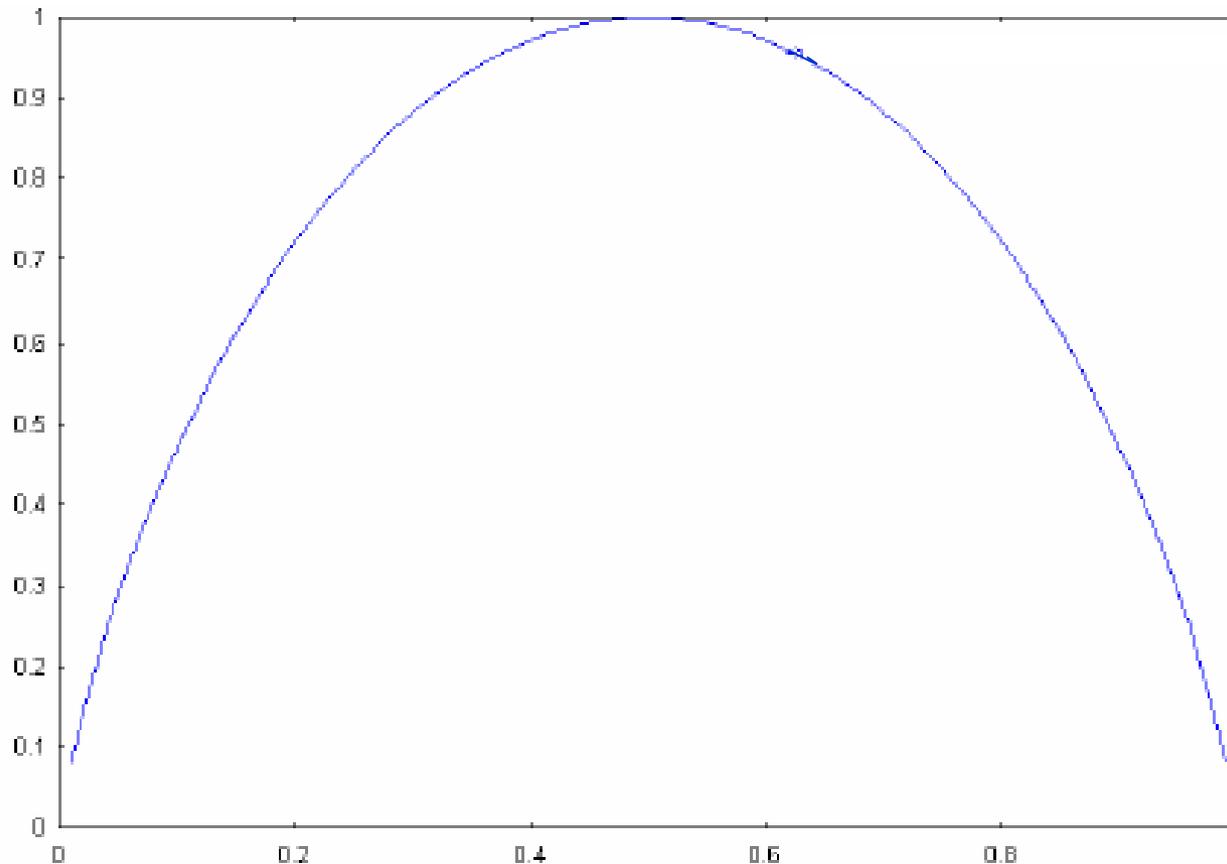
- S é o conjunto de exemplo de treino;
- p_+ é a porção de exemplos positivos;
- p_- é a porção de exemplos negativos.
- A Entropia(S) = 0 se existe um i tal que $p_i = 1$
- É assumido que $0 * \log_2 0 = 0$.

se $p_+ = 1$, por exemplo, o receptor sabe que o exemplo tirado será positivo, assim não há necessidade de enviar mensagem, e a entropia é zero.

Aprendizagem - Exemplo

Entropia

$$Entropia(S) = -p_+ \log_2 p_+ - p_- \log_2 p_-$$



Varição da Entropia a medida que a proporção de positivos e negativos se altera.



Aprendizagem - Exemplo

- Se p_{\oplus} é 1, o destinatário sabe que o exemplo selecionado será positivo.
 - Nenhuma mensagem precisa ser enviada
 - Entropia é 0 (mínima)
- Se p_{\oplus} é 0.5, um bit é necessário para indicar se o exemplo selecionado é \oplus ou \ominus .
 - Entropia é 1 (máxima)
- Se p_{\oplus} é 0.8, então uma coleção de mensagens podem ser codificadas usando-se - em média menos de um bit - códigos mais curtos para \oplus e mais longos para \ominus .

Entropia(S) => especifica o número mínimo de bits de informação necessário para codificar uma classificação de um membro arbitrário de S.



Aprendizagem - Exemplo

Ganho

- ❑ Define a redução na entropia.
- ❑ $Ganho(S, A)$ significa a redução esperada na entropia de S , ordenando pelo atributo A .

$$Ganho(S, A) = Entropia(S) - \sum_{v \in \text{valores}(A)} \frac{|S_v|}{|S|} \cdot Entropia(S_v)$$

- ❑ Para responder à pergunta anterior, "**Como escolher o melhor atributo?**" é usado o ganho.
 - Em cada iteração do algoritmo é escolhido o atributo que apresenta uma maior ganho.

Obs.: $\text{valores}(A)$ é o conjunto de todos possíveis valores para o atributo A , e S_v é o subconjunto de S para qual o atributo A tem valor v .



Aprendizagem - Exemplo

- ❑ **Entropia** - medida da impureza do conjunto de treino. Assumindo o valor máximo (1) quando existem tantos elementos positivos como negativos, e o valor mínimo (0) quando todos os elementos são da mesma classe.
- ❑ **Ganho de informação** - redução esperada no valor da Entropia, devido à ordenação do conjunto de treino segundo os valores do atributo A.

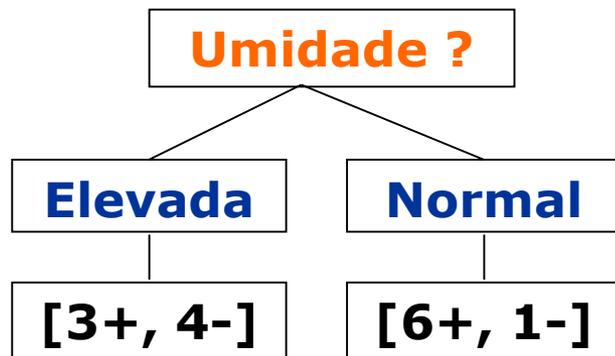


Aprendizagem - Exemplo

- Primeiro passo: são analisados todos os atributos, começando pela **Umidade**.

$$S = [9+, 5-]$$

$$E(\text{Entropia}) = -\left(\frac{9}{14}\right)\log_2\left(\frac{9}{14}\right) - \left(\frac{5}{14}\right)\log_2\left(\frac{5}{14}\right) = 0,940$$



$$E=0,985$$

$$E=0,592$$

$$\text{Ganho}(S, \text{Umidade}) = 0,940 - \left(\frac{7}{14}\right) \times 0,985 - \left(\frac{7}{14}\right) \times 0,592$$

$$\text{Ganho}(S, \text{Umidade}) = 0,151$$

Dia	Aspecto	Temp.	Umidade	Vento	Jogar Tênis
D1	Sol	Quente	Elevada	Fraco	Não
D2	Sol	Quente	Elevada	Forte	Não
D3	Nuvens	Quente	Elevada	Fraco	Sim
D4	Chuva	Ameno	Elevada	Fraco	Sim
D5	Chuva	Fresco	Normal	Fraco	Sim
D6	Chuva	Fresco	Normal	Forte	Não
D7	Nuvens	Fresco	Normal	Fraco	Sim
D8	Sol	Ameno	Elevada	Fraco	Não
D9	Sol	Fresco	Normal	Fraco	Sim
D10	Chuva	Ameno	Normal	Forte	Sim
D11	Sol	Ameno	Normal	Forte	Sim
D12	Nuvens	Ameno	Elevada	Forte	Sim
D13	Nuvens	Quente	Normal	Fraco	Sim
D14	Chuva	Ameno	Elevada	Forte	Não

$$\log_b x = (\log_m x) / (\log_m b).$$

Aprendizagem - Exemplo

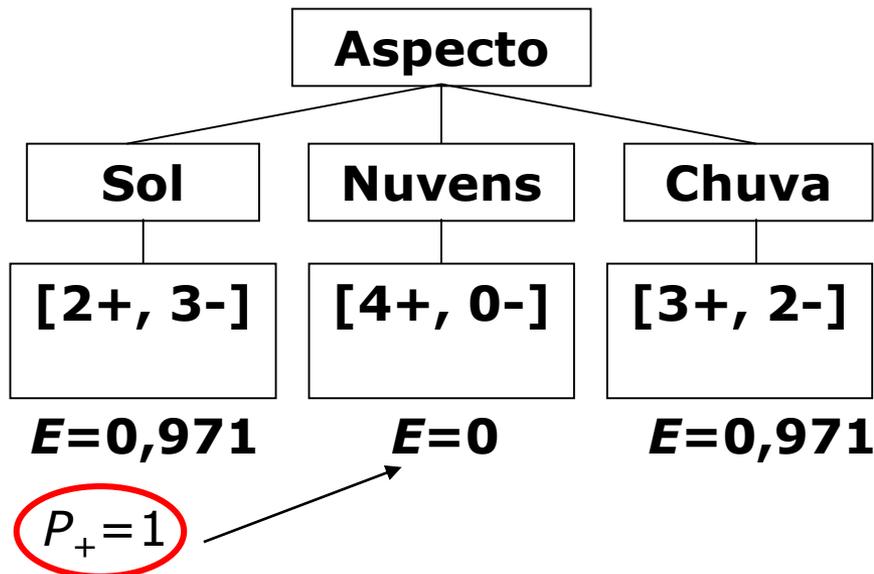
Calculando o ganho para todos os atributos \Rightarrow o que tem maior ganho é o **Aspecto**.

$$S = [9+, 5-]$$

$$E = 0,940$$

$$\text{MAX} \left\{ \begin{array}{l} \text{Ganho}(S, \text{Umidade}) = 0,151 \\ \text{Ganho}(S, \text{Vento}) = 0,048 \\ \text{Ganho}(S, \text{Aspecto}) = 0,247 \\ \text{Ganho}(S, \text{Temp.}) = 0,029 \end{array} \right.$$

$= \text{Ganho}(S, \text{Aspecto})$ 



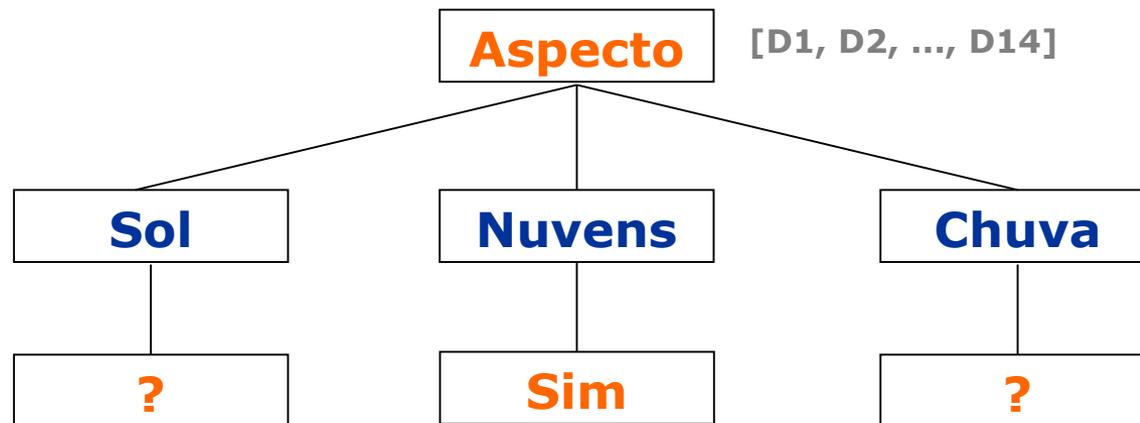
Dia	Aspecto	Temp.	Umidade	Vento	Jogar Tênis
D1	Sol	Quente	Elevada	Fraco	Não
D2	Sol	Quente	Elevada	Forte	Não
D3	Nuvens	Quente	Elevada	Fraco	Sim
D4	Chuva	Ameno	Elevada	Fraco	Sim
D5	Chuva	Fresco	Normal	Fraco	Sim
D6	Chuva	Fresco	Normal	Forte	Não
D7	Nuvens	Fresco	Normal	Fraco	Sim
D8	Sol	Ameno	Elevada	Fraco	Não
D9	Sol	Fresco	Normal	Fraco	Sim
D10	Chuva	Ameno	Normal	Forte	Sim
D11	Sol	Ameno	Normal	Forte	Sim
D12	Nuvens	Ameno	Elevada	Forte	Sim
D13	Nuvens	Quente	Normal	Fraco	Sim
D14	Chuva	Ameno	Elevada	Forte	Não

$$\text{Ganho}(S, \text{Aspecto}) = 0,940 - \left(\frac{5}{14}\right) \times 0,971 - \left(\frac{4}{14}\right) \times 0,0 - \left(\frac{5}{14}\right) \times 0,971$$

$$\text{Ganho}(S, \text{Aspecto}) = 0,247$$

Aprendizagem - Exemplo

No próximo passo o atributo **Aspecto** já não é considerado.



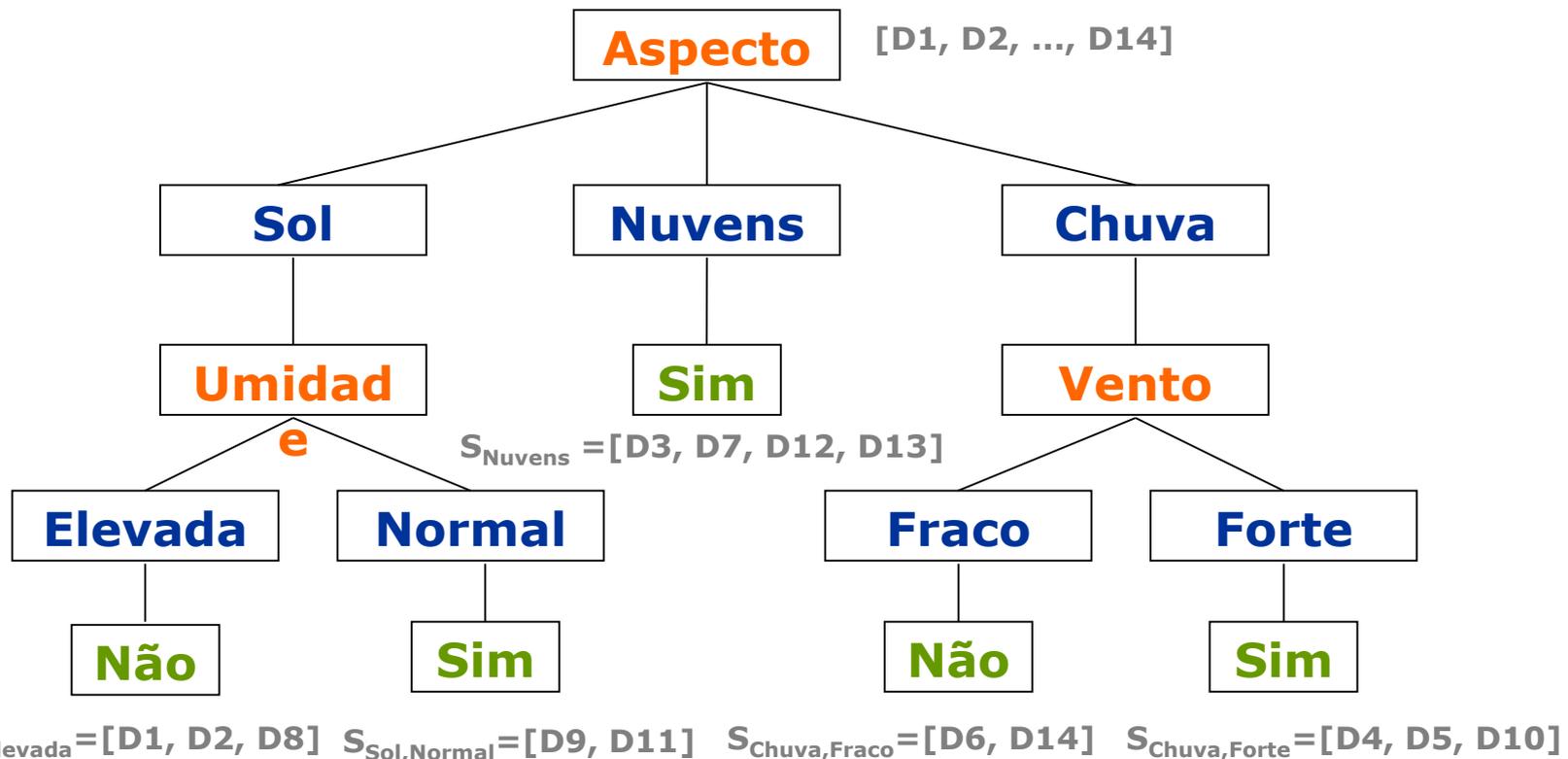
$S_{Sol} = [D1, D2, D8, D9, D11]$ $S_{Nuvens} = [D3, D7, D12, D13]$ $S_{Chuva} = [D4, D5, D6, D10, D14]$

$$\text{MAX} \left(\begin{array}{l}
 \text{Ganho}(S_{Sol}, \text{Umidade}) = 0,971 - \left(\frac{3}{5}\right) \times 0,0 - \left(\frac{2}{5}\right) \times 0,0 = 0,971 \\
 \text{Ganho}(S_{Sol}, \text{Temp.}) = 0,971 - \left(\frac{2}{5}\right) \times 0,0 - \left(\frac{2}{5}\right) \times 1,0 - \left(\frac{1}{5}\right) \times 0,0 = 0,570 \\
 \text{Ganho}(S_{Sol}, \text{Vento}) = 0,971 - \left(\frac{2}{5}\right) \times 1,0 - \left(\frac{3}{5}\right) \times 0,918 = 0,019
 \end{array} \right)$$

$= \text{Ganho}(S_{Sol}, \text{Umidade})$ ←

Aprendizagem - Exemplo

Quando em todos os nós a entropia for nula, o algoritmo para e obtêm-se a seguinte Árvore de decisão:



Aprendizagem - Exemplo

Aplicações

- Diagnóstico médico e de equipamentos
- Análise de crédito
- Recuperação de Informação

Exemplo de implementação:

- <http://www.cs.ualberta.ca/%7Eaixplore/learning/DecisionTrees/Applet/DecisionTreeApplet.html>

