

Aprendizado

Capítulo 5

309

O que é Aprendizado?

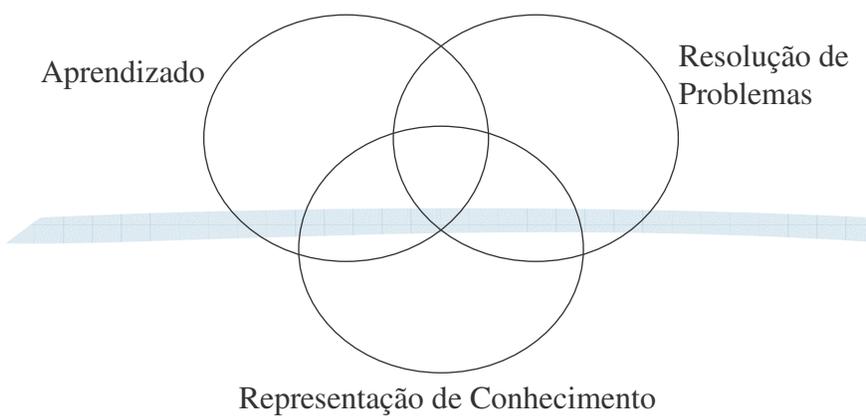
- “...mudanças que tornam sistemas capazes de adaptar-se e realizar tarefas semelhantes de maneira mais eficiente” [Simon]

310

Tipos de Aprendizado

- Refinamento de Habilidades *vs.* Aquisição do Conhecimento
- Aprendizagem por Memorização
- Aprendizagem através de Conselhos
- Aprendizagem através da Solução de Problemas
- Aprendizagem através de Exemplos

311



312

Aprendizagem por Memorização

- Armazenamento simples de valores
- Armazenamento e recuperação são mais baratos que o cálculo
- O armazenamento otimiza cálculos futuros
- Objetiva melhorar o desempenho em sistemas de IA

313

Aprendizagem por Aconselhamento

- Assim como pessoas, programas também podem aprender através de conselhos dados por seres humanos
- Os conselhos precisam ser codificados para que possam ser interpretados pelo sistema
- Novos conselhos podem ser utilizados pelo sistema com o objetivo de melhorar seu desempenho

314

Aprendizado por Exemplos: Indução

- Na estratégia de aprendizado por indução, o sistema adquire os conceitos através de inferências indutivas realizadas sobre fatos fornecidos ou observados.

315

Indução & Dedução

- Exemplo Dedução:
 - todos os homens são fortes
 - **Se** Pedro é homem **Então** Pedro é forte
- Exemplo de Indução:
 - A maioria dos homens é forte
 - **Se** Pedro é homem **Então** Pedro é forte

316

Aprendizado por Indução: Classificação

- O objetivo é associar a cada exemplo, ou observação, uma classe à qual ele pertence
- Os conceitos construídos estão representados na forma de um classificador

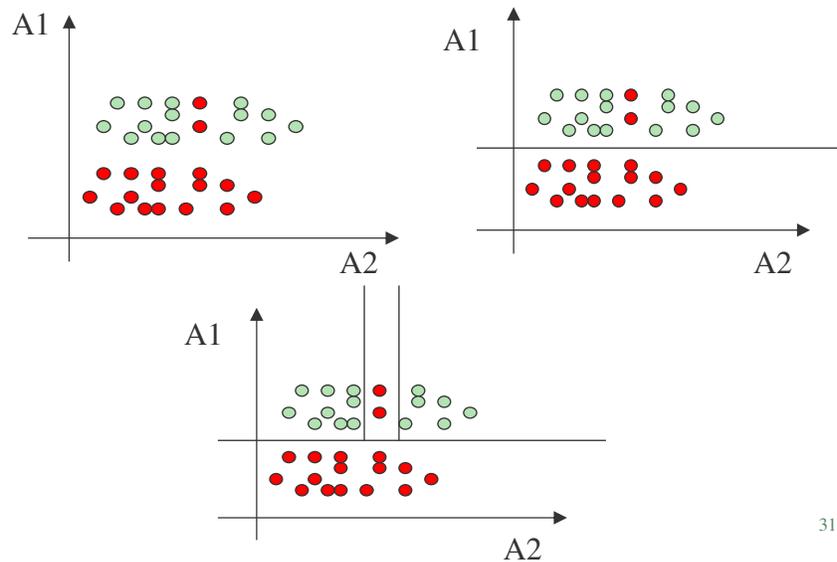


Aprendizado por Indução: Classificação (Cont.)

- O exemplo a ser classificado é submetido aos conceitos adquiridos, e uma decisão sobre a sua classe é devolvida pelo classificador



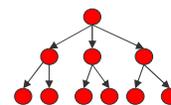
Problemas com Classificação



319

Aprendizado - Árvores de Decisão

• Representação

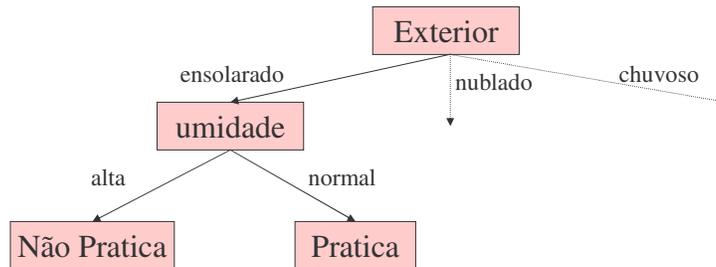


- formada por nós folha que representam as classes de cada exemplo contido na folha
- formada por nós internos que representam condições sobre os valores de um determinado atributo
- formada por ligações entre nós que representam os valores possíveis para um determinado atributo

320

•Aprendizado - Árvores de Decisão

(Cont.)



• Algoritmo de Construção

- Entrada: Conjunto de Exemplos E
- Saída: Árvore de Decisão T

321

Aprendizado - Árvores de Decisão

(Cont.)

$E = \{E_1, \dots, E_n\}$ = Conjunto de Exemplos

{ Inicialmente todos os exemplos pertencem a raiz }

Se todas as tuplas pertencem a uma mesma classe **então** parar

senão

escolha um atributo;

divida o conjunto de tuplas em subconjuntos, de acordo com cada valor do atributo escolhido;

repita o procedimento recursivamente em cada subconjunto criado;

fim_senão

fim_se

322

Aprendizado - Árvores de Decisão

(Cont.)

• Escolha do Atributo

• Entropia

- Entropia é uma grandeza da teoria da informação que mede a desordem de uma determinada coleção de exemplos
- Quanto maior a entropia, maior a desordem dos exemplos

323

Aprendizado - Árvores de Decisão

(Cont.)

• Fórmula da Entropia

$$E(S) = - \sum_{i=1}^k p(i) \log_2 p(i)$$

onde:

- S é um conjunto de exemplos
- i é uma determinada classe
- $p(i)$ é a probabilidade de ocorrência da classe i no conjunto S
- k é o número de classes

324

Aprendizado - Árvores de Decisão

(Cont.)

• Entropia de um atributo

$$E(S, A) = - \sum_{i=1}^v \frac{|S_i|}{|S|} E(S_i)$$

onde:

- A é um determinado atributo
- i é um determinado valor para o atributo A
- S_i é o subconjunto de exemplos que possuem valor i para o atributo A
- $|S_i|$ é o número de exemplos de S_i
- v é o número de valores para o atributo A

325

Aprendizado - Árvores de Decisão

(Cont.)

• Escolha baseada no ganho de informação

$$\text{ganho}(S, A) = E(S) - E(S, A)$$

• Informação da Divisão

$$\text{Inf}(S, A) = - \sum_{i=1}^v \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|}$$

326

Aprendizado - Árvores de Decisão

(Cont.)

- Escolha baseada no ganho proporcional de informação

$$\text{ganhoratio}(S, A) = \frac{\text{ganho}(S, A)}{\text{inf}(S, A)}$$

- Melhor atributo é aquele que maximiza o ganho proporcional

327

Aprendizado - Árvores de Decisão

(Cont.)

- Exemplo

- Dado um conjunto de exemplos, verificar se um determinado dia está próprio para prática de esportes

328

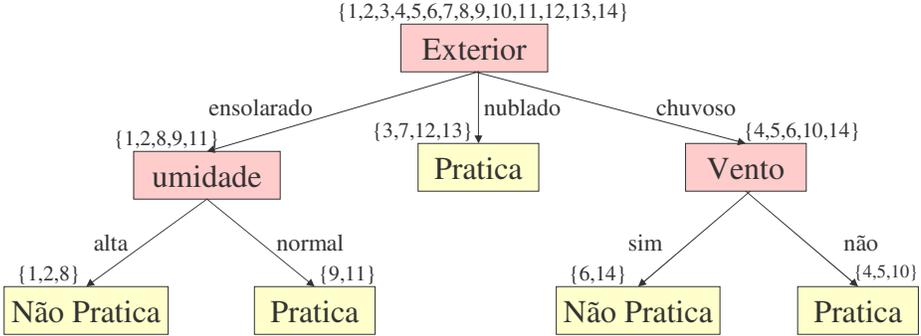
N	Atributos (A)				Classe (C)
	Exterior	Temperatura	Umidade	Vento	
1	ensolarado	alta	alta	não	np
2	ensolarado	alta	alta	sim	np
3	nublado	alta	alta	não	pe
4	chuvoso	média	alta	não	pe
5	chuvoso	baixa	normal	não	pe
6	chuvoso	baixa	normal	sim	np
7	nublado	baixa	normal	sim	pe
8	ensolarado	média	alta	não	np
9	ensolarado	baixa	normal	não	pe
10	chuvoso	média	normal	não	pe
11	ensolarado	média	normal	sim	pe
12	nublado	média	alta	sim	pe
13	nublado	alta	normal	não	pe
14	chuvoso	média	alta	sim	np

329

Aprendizado - Árvores de Decisão

(Cont.)

- Árvore de Decisão criada

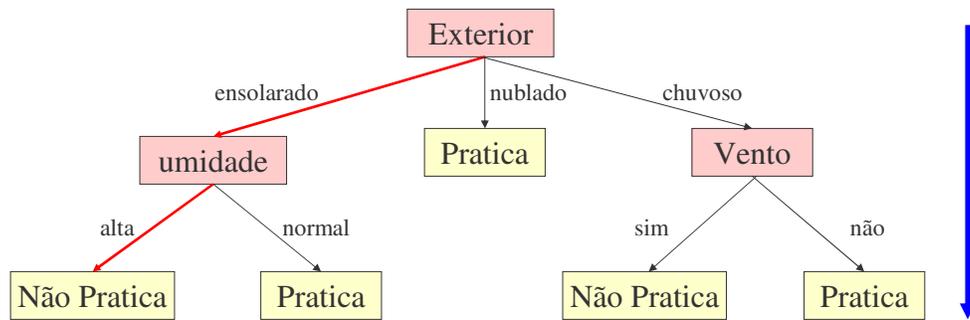


330

Aprendizado - Árvores de Decisão

(Cont.)

- Classificação do exemplo <ensolarado média alta não>



331

Aprendizado - Árvores de Decisão

(Cont.)

• Regras de produção geradas

1. **Se** exterior = ensolarado
e umidade = alta
Então classe = np
2. **Se** exterior = ensolarado
e umidade = normal
Então classe = pe

332

Aprendizado - Árvores de Decisão

(Cont.)

• Regras de produção geradas

3. **Se** exterior = nublado

Então classe = pe

4. **Se** exterior = chuvoso

e vento = sim

Então classe = np

333

Aprendizado - Árvores de Decisão

(Cont.)

• Regras de produção geradas

5. **Se** exterior = chuvoso

e vento = não

Então classe = pe

334

Aprendizado - Árvores de Decisão

(Cont.)

• Principais problemas encontrados

- tamanho da árvore
- superadequação aos dados
- presença de informações inconsistentes

• Possíveis estratégias para a solução dos problemas encontrados

- poda (pré-poda e pós-poda)
- utilização de conceitos da Lógica Paraconsistente para manipulação adequada de dados inconsistentes

335

Aprendizado - Árvores de Decisão (Cont.)

• Vantagens

- Podem adquirir e representar facilmente conceitos disjuntos
- Podem manipular dados com ruídos
- Generalizam conceitos a partir de poucos atributos
- Possuem representação simples e rápida classificação
- Têm obtido muito sucesso em diversas áreas de aplicação — vários sistemas estão disponíveis comercialmente

336

Aprendizado - Árvores de Decisão (Cont.)

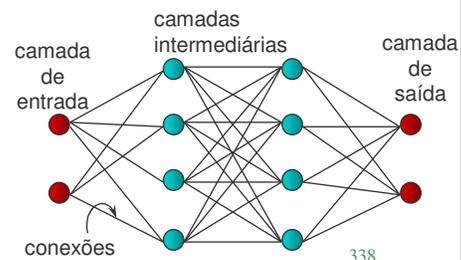
- Desvantagens
 - Vários conceitos para representar uma mesma classe podem ser criados
 - A maioria dos sistemas de árvore de decisão não implementam *backtraking* (retrocesso) na escolha de um atributo (sistemas determinísticos)
 - Poda não adequada pode generalizar conceitos de forma excessiva

337

Outros métodos de Aprendizado por Exemplos

• Redes Neurais

- também conhecidas como o modelos conexionistas
- são redes interconectadas formadas por elementos computacionais muito simples
- baseadas no modelo de funcionamento do cérebro



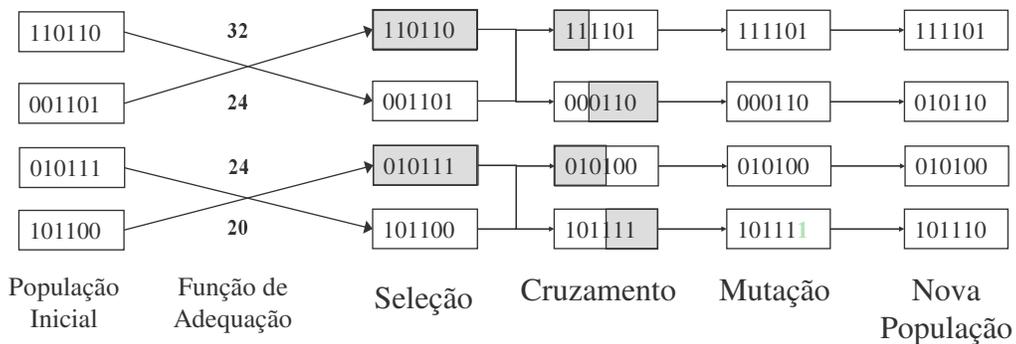
Outros métodos de Aprendizado por Exemplos

• Algoritmos Genéticos

- tiveram origem nos estudos realizados sobre autômatos celulares (Holland)
- baseado no mecanismo de seleção natural e não no processo de raciocínio simulado
- procura gerar novas descrições, a partir de descrições iniciais, que possuam maior qualidade, criando novas gerações de descrições

339

Algoritmos Genéticos



340