

Regras

Regras Rudimentarias

- ▶ 1 R: aprende uma regra por atributo
 - atribuí a classe mais freqüente
 - taxa de erro: proporção de instâncias que não pertence a classe majoritária
 - escolhe o atributo com menor taxa de erro

Pseudo-código para 1R

- ▶ Para cada atributo
 - para cada valor do atributo, faça uma regra como:
 - conte a frequência de cada classe
 - encontre a classe mais freqüente
 - atribua a classe mais freqüente a esta regra
 - calcule a taxa de erro da regra
 - escolha a regra com taxa de erro menor

Avaliando os atributos da base Weather

Outlook	Temp	Humidity	Windy	Play				
Sunny	Hot	High	False	No	Attribute	Rules	Errors	Total errors
Sunny	Hot	High	True	No	Outlook	Sunny → No	2/5	4/14
Overcast	Hot	High	False	Yes		Overcast → Yes	0/4	
Rainy	Mild	High	False	Yes		Rainy → Yes	2/5	
Rainy	Cool	Normal	False	Yes				
Rainy	Cool	Normal	True	No	Temp	Hot → No*	2/4	5/14
Overcast	Cool	Normal	True	Yes		Mild → Yes	2/6	
Sunny	Mild	High	False	No		Cool → Yes	1/4	
Sunny	Cool	Normal	False	Yes				
Rainy	Mild	Normal	False	Yes	Humidity	High → No	3/7	4/14
Sunny	Mild	Normal	True	Yes		Normal → Yes	1/7	
Overcast	Mild	High	True	Yes	Windy	False → Yes	2/8	5/14
Overcast	Hot	Normal	False	Yes		True → No*	3/6	
Rainy	Mild	High	True	No				

Atributos Numéricos

- Discretize os atributos numéricos
- Divida cada atributo em intervalos
 - ◆ Ordene as instâncias de acordo com os valores
 - ◆ Coloque um ponto de corte na alteração de classe
 - ◆ Isto minimiza o erro total
- Example: *a temperatura da base Weather*

64 65 68 69 70 71 72 72 75 75 80 81 83 85
Yes | No | Yes Yes Yes | No No Yes | Yes Yes | No | Yes Yes | No

Outlook	Temperature	Humidity	Windy	Play
Sunny	85	85	False	No
Sunny	80	90	True	No
Overcast	83	86	False	Yes
Rainy	75	80	False	Yes
...	...	...	...	...

O problema de sobre especialização

- Este procedimento é muito sensível a ruído
 - ♦ Uma instância com uma classe incorreta produzira um intervalo separado
 - ♦ Solução simples:
 - ♦ *Forçar um numero mínimo de instâncias no intervalo*
- Exemplo (com $\text{min} = 3$):

64	65	68	69	70	71	72	72	75	75	80	81	83	85
Yes	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	No
64	65	68	69	70	71	72	72	75	75	80	81	83	85
Yes	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes	No	Yes	Yes	No

Evitando Sobre especialização

- Conjunto resultante:

Attribute	Rules	Errors	Total errors
Outlook	Sunny $\rightarrow$ No	2/5	4/14
	Overcast $\rightarrow$ Yes	0/4	
	Rainy $\rightarrow$ Yes	2/5	
Temperature	$\leq 77.5 \rightarrow$ Yes	3/10	5/14
	$> 77.5 \rightarrow$ No*	2/4	
Humidity	$\leq 82.5 \rightarrow$ Yes	1/7	3/14
	> 82.5 and $\leq 95.5 \rightarrow$ No	2/6	
	$> 95.5 \rightarrow$ Yes	0/1	
Windy	False $\rightarrow$ Yes	2/8	5/14
	True $\rightarrow$ No*	3/6	

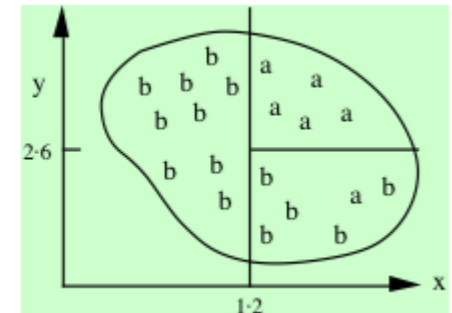
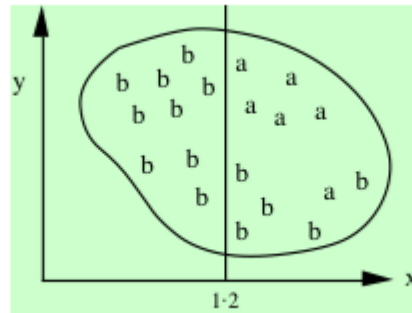
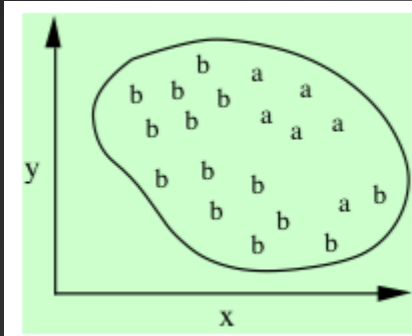
Discussão

Apesar de simples apresenta resultados similares a arvores de decisão!!!!

Algoritmos 'Covering'

- ▶ Para cada classe são geradas um conjunto de regras.
- ▶ Abordagem 'Covering'
 - a cada passo são identificadas regras que cobrem algumas instâncias

Exemplo



Se True então a

Se $x > 1.2$ então a

Se $x > 1.2$ e $y > 2.6$
então a

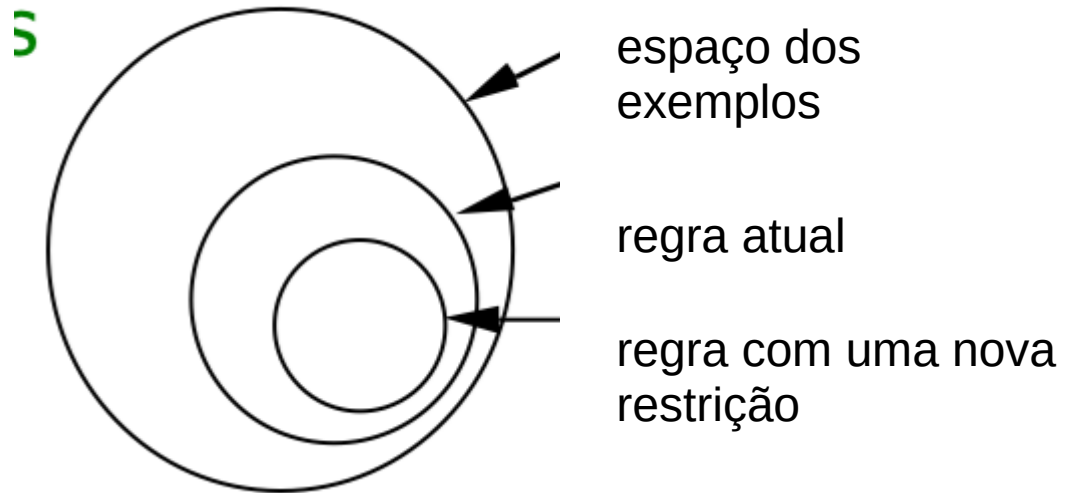
Possível conjunto de regras para a classe b

Se $x < 1.2$ então classe b

se $x > 1.2$ and $y < 2.6$ então classe b

Algorithm 'Covering'

- ▶ Gere uma regra adicionando restrições que maximizem a precisão da regra.
- ▶ Similar a situação das árvores



Selecionando a Restrição

- ▶ Objetivo: Melhorar a precisão
 - t numero total de instâncias cobertas pela regra
 - p exemplos positivos da classe cobertos pela regra
 - $t-p$ numero de erros cometidos
 - Maximizar p/t ??

Exemplo: Receita de óculos ???

Age	Spectacle prescription	Astigmatism	Tear production rate	Recommended lenses
Young	Myope	No	Reduced	None
Young	Myope	No	Normal	Soft
Young	Myope	Yes	Reduced	None
Young	Myope	Yes	Normal	Hard
Young	Hypermetrope	No	Reduced	None
Young	Hypermetrope	No	Normal	Soft
Young	Hypermetrope	Yes	Reduced	None
Young	Hypermetrope	Yes	Normal	hard
Pre-presbyopic	Myope	No	Reduced	None
Pre-presbyopic	Myope	No	Normal	Soft
Pre-presbyopic	Myope	Yes	Reduced	None
Pre-presbyopic	Myope	Yes	Normal	Hard
Pre-presbyopic	Hypermetrope	No	Reduced	None
Pre-presbyopic	Hypermetrope	No	Normal	Soft
Pre-presbyopic	Hypermetrope	Yes	Reduced	None
Pre-presbyopic	Hypermetrope	Yes	Normal	None
Presbyopic	Myope	No	Reduced	None
Presbyopic	Myope	No	Normal	None
Presbyopic	Myope	Yes	Reduced	None
Presbyopic	Myope	Yes	Normal	Hard
Presbyopic	Hypermetrope	No	Reduced	None
Presbyopic	Hypermetrope	No	Normal	Soft
Presbyopic	Hypermetrope	Yes	Reduced	None
Presbyopic	Hypermetrope	Yes	Normal	None

Primeira Restrição

```
If ?  
    then recommendation = hard
```

Age = Young	2/8
Age = Pre-presbyopic	1/8
Age = Presbyopic	1/8
Spectacle prescription = Myope	3/12
Spectacle prescription = Hypermetrope	1/12
Astigmatism = no	0/12
Astigmatism = yes	4/12
Tear production rate = Reduced	0/12
Tear production rate = Normal	4/12

Regra

```
If astigmatism = yes  
    then recommendation = hard
```

Instâncias cobertas pela regra

Age	Spectacle prescription	Astigmatism	Tear production rate	Recommended lenses
Young	Myope	Yes	Reduced	None
Young	Myope	Yes	Normal	Hard
Young	Hypermetrope	Yes	Reduced	None
Young	Hypermetrope	Yes	Normal	hard
Pre-presbyopic	Myope	Yes	Reduced	None
Pre-presbyopic	Myope	Yes	Normal	Hard
Pre-presbyopic	Hypermetrope	Yes	Reduced	None
Pre-presbyopic	Hypermetrope	Yes	Normal	None
Presbyopic	Myope	Yes	Reduced	None
Presbyopic	Myope	Yes	Normal	Hard
Presbyopic	Hypermetrope	Yes	Reduced	None
Presbyopic	Hypermetrope	Yes	Normal	None

Refinando

estado Atual

```
If astigmatism = yes  
    and ?  
    then recommendation = hard
```

possíveis atributos

Age = Young	2/4
Age = Pre-presbyopic	1/4
Age = Presbyopic	1/4
Spectacle prescription = Myope	3/6
Spectacle prescription = Hypermetrope	1/6
Tear production rate = Reduced	0/6
Tear production rate = Normal	4/6

Regra Modificada e dados

regra com o melhor atributo adicionado

```
If astigmatism = yes  
    and tear production rate = normal  
then recommendation = hard
```

Instâncias cobertas pela regra modificada

Age	Spectacle prescription	Astigmatism	Tear production rate	Recommended lenses
Young	Myope	Yes	Normal	Hard
Young	Hypermetrope	Yes	Normal	hard
Pre-presbyopic	Myope	Yes	Normal	Hard
Pre-presbyopic	Hypermetrope	Yes	Normal	None
Presbyopic	Myope	Yes	Normal	Hard
Presbyopic	Hypermetrope	Yes	Normal	None

Continuando

Estado atual

```
If astigmatism = yes  
    and tear production rate = normal  
    and ?  
    then recommendation = hard
```

Possíveis testes

Age = Young	2/2
Age = Pre-presbyopic	1/2
Age = Presbyopic	1/2
Spectacle prescription = Myope	3/3
Spectacle prescription = Hypermetrope	1/3

Resultado

- ▶ Regra final

```
If astigmatism = yes  
and tear production rate = normal  
and spectacle prescription = myope  
then recommendation = hard
```

- ▶ Segunda regra (construída a partir dos exemplos que não são cobertos pela regra anterior)

```
If age = young and astigmatism = yes  
and tear production rate = normal  
then recommendation = hard
```