

# Saídas

---

Carlos Diego Rodrigues

5 de outubro de 2021

Universidade Federal do Ceará

## Saídas: representando o conhecimento

- Tabelas
- Modelos Lineares
- Árvores de decisão
- Regras
- Regras de classificação

## Saídas representando padrões estruturais

- Várias formas de representar padrões.
- Árvores de decisão, regras, baseados na instância, ...
- Também chamados representação do "conhecimento".
- A representação está ligada ao método de inferência.
- Entender a forma como o conhecimento é representado é fundamental para entender os métodos de aprendizado.
- Diferentes formas de saídas para diferentes problemas de aprendizado (classificação, regressão, associação, ...)

# Tabelas de decisão

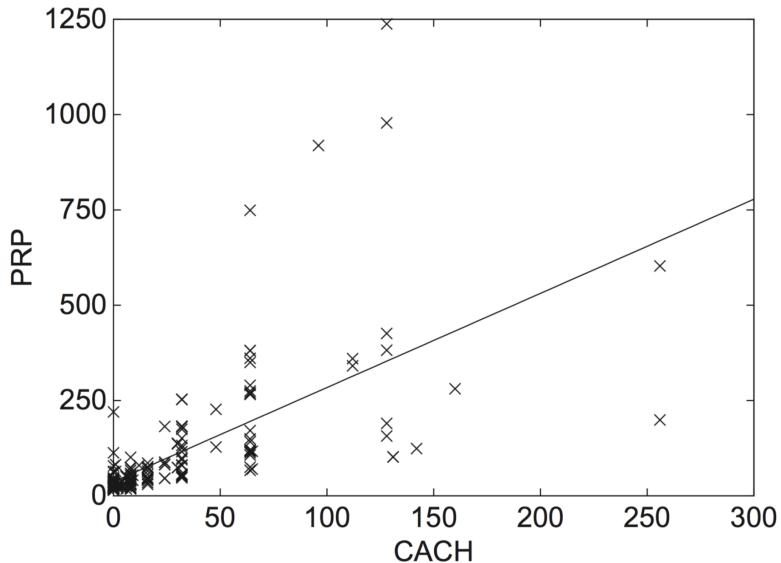
- Forma mais simples de representar a saída

No.	1: outlook	2: temperature	3: humidity	4: windy	5: <b>play</b>
	Nominal	Nominal	Nominal	Nominal	Nominal
1	sunny	hot	high	FALSE	no
2	sunny	hot	high	TRUE	no
3	overcast	hot	high	FALSE	yes
4	rainy	mild	high	FALSE	yes
5	rainy	cool	normal	FALSE	yes
6	rainy	cool	normal	TRUE	no
7	overcast	cool	normal	TRUE	yes
8	sunny	mild	high	FALSE	no
9	sunny	cool	normal	FALSE	yes
10	rainy	mild	normal	FALSE	yes
11	sunny	mild	normal	TRUE	yes
12	overcast	mild	high	TRUE	yes
13	overcast	hot	normal	FALSE	yes
14	rainy	mild	high	TRUE	no

- Principal problema: selecionar os atributos certos.

- Forma de representação simplificada
- Tradicionalmente conhecida para regressão:
  - Entradas (atributos) e saídas são todos numéricos.
  - Saída é uma soma ponderada das entradas.
- Objetivo é achar bons valores para os pesos.
- Muitas formas de fazer, a mais famosa é o método de minimizar o desvio quadrático.

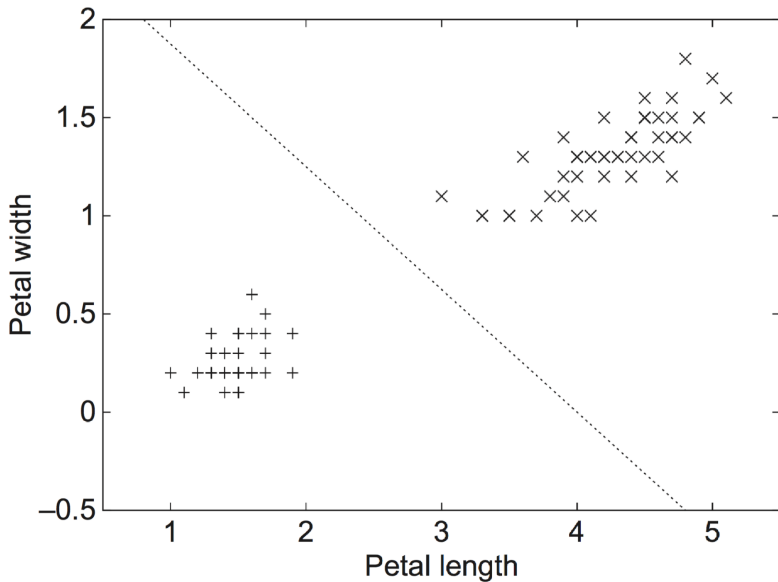
## Regressão linear - exemplo



# Modelos lineares para classificação

- Usado para classificação binária
- Uma linha separa duas classes: borda de decisão.
- Predição é realizada a partir de uma função cuja entrada são os atributos e a saída é uma classe (tipicamente  $\{0, 1\}$ ,  $\{-1, 1\}$  ou  $< 0$  e  $\geq 0$ ).
- A borda se torna um hiperplano em dimensões maiores (mais atributos).

## Classificação setosas vs. versicolor





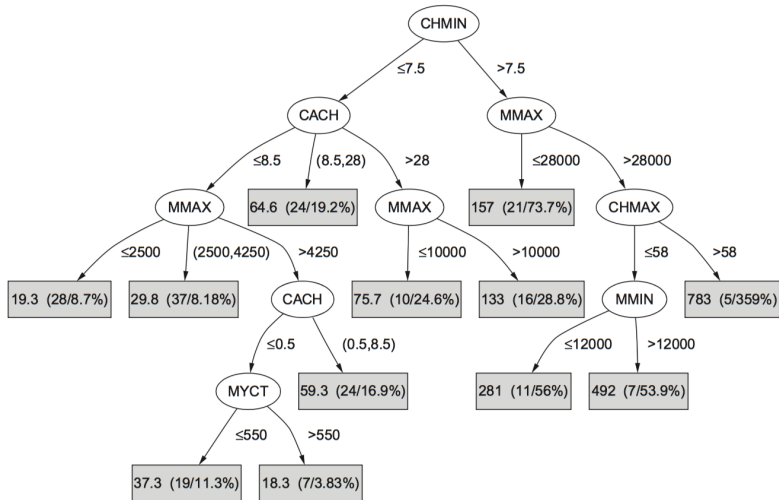
# Árvores de decisão

- Uma abordagem "dividir para conquistar" produz uma árvore.
- Nós da árvore consistem em um teste sobre um atributo em particular.
- Tipicamente um atributo é comparado a uma constante.
- Outras possibilidades:
  - Comparar valores de dois atributos.
  - Usar uma função sobre um ou mais atributos.
- Folhas representam uma ou mais classes ou distribuições de probabilidade entre as classes.
- Dada uma nova instância, faz-se a rota dela a partir da raiz da árvore.

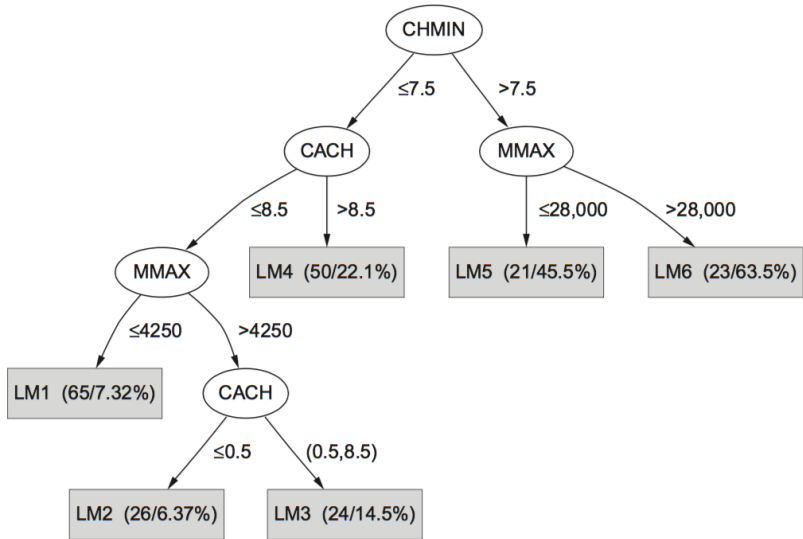
# Árvores de decisão para regressão

- Regressão é o processo de computar uma expressão que prediz uma quantidade numérica.
- Árvore de regressão é uma árvore de decisão onde cada folha é uma regressão
  - Média dos valores cobertos por aquele ramo.
  - Um modelo de regressão associado a cada folha.
  - Corretivos lineares aproximam a árvore de regressão a uma função contínua.

# Árvore de regressão



# Árvore de modelos regressão



# Regras de classificação

- Alternativa popular às árvores de decisão
- *Antecedente*: uma série de testes (similares aos da árvore de decisão).
- Em geral eles são uma conjunção de testes (**e**).
- *Consequente*: classes, conjuntos de classes ou uma distribuição de probabilidade entre as classes.
- Mais de uma regra pode ser aplicada em um formato de disjunção (**ou**), o que pode gerar conflitos caso resultados diferentes surjam.

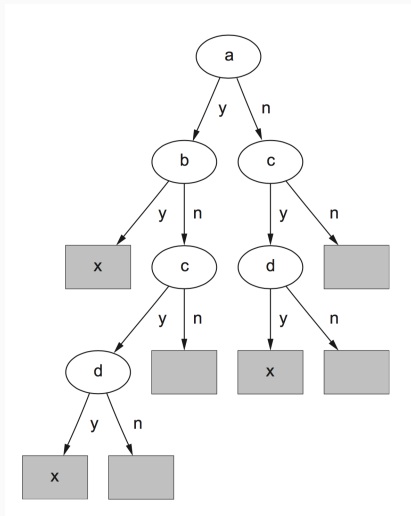
# Conversão de árvores em regras

- Conversão simples: cada folha se torna uma regra.
- Produz regras não-ambíguas
- Pode produzir regras desnecessariamente complexas e em grande número.

## Conversão de regras em árvores

- Nem sempre é possível pois as árvores não conseguem representar disjunções.
- É preciso desfazer a simetria.
- Pode exigir réplicas de subárvores.

# Árvore de modelos regressão





# Árvore de modelos regressão

If  $x=1$  and  $y=1$  then class = a  
If  $z=1$  and  $w=1$  then class = a  
Otherwise class = b

