

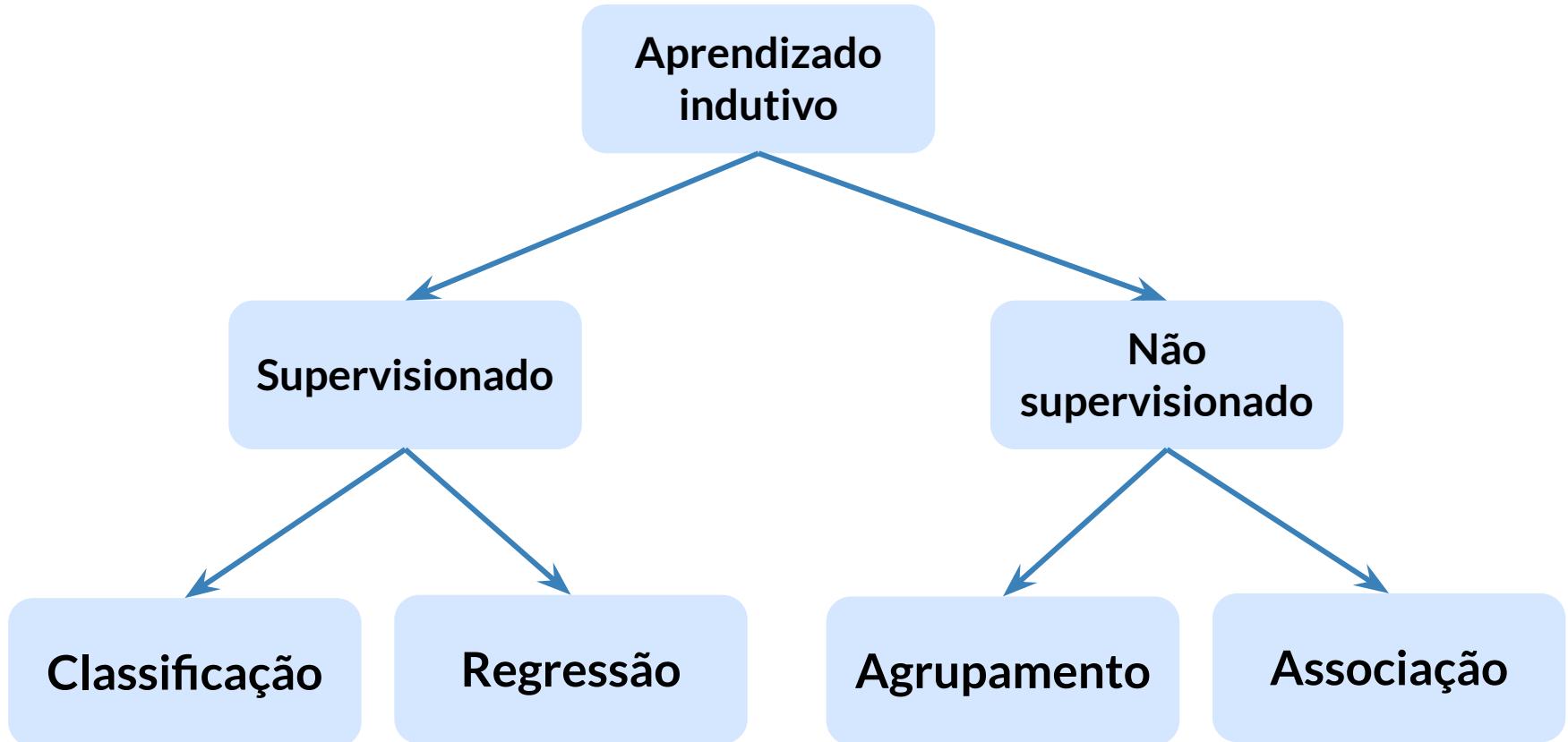
Árvores de Decisão e Regressão



Gabriel P. Oliveira

Inteligência Artificial
2026/1

Tarefas de Aprendizado



Tarefas de Aprendizado

Preditivas → Encontrar uma função (hipótese) a partir dos dados de treinamento para prever um rótulo ou valor com base nos atributos de entrada

Algoritmos seguem o paradigma do **aprendizado supervisionado**

- Existe um “supervisor externo” que conhece a saída desejada para cada exemplo – o atributo de saída

Tarefas de Aprendizado

Descritivas → Explorar ou descrever um conjunto de dados

Algoritmos seguem o paradigma do **aprendizado não supervisionado**

- *Não fazem uso o atributo de saída*

1.

O que são Árvores de Decisão?

Definição e funcionamento

Árvore de Decisão



Árvores de Decisão/Regressão
são algoritmos de aprendizado
preditivo

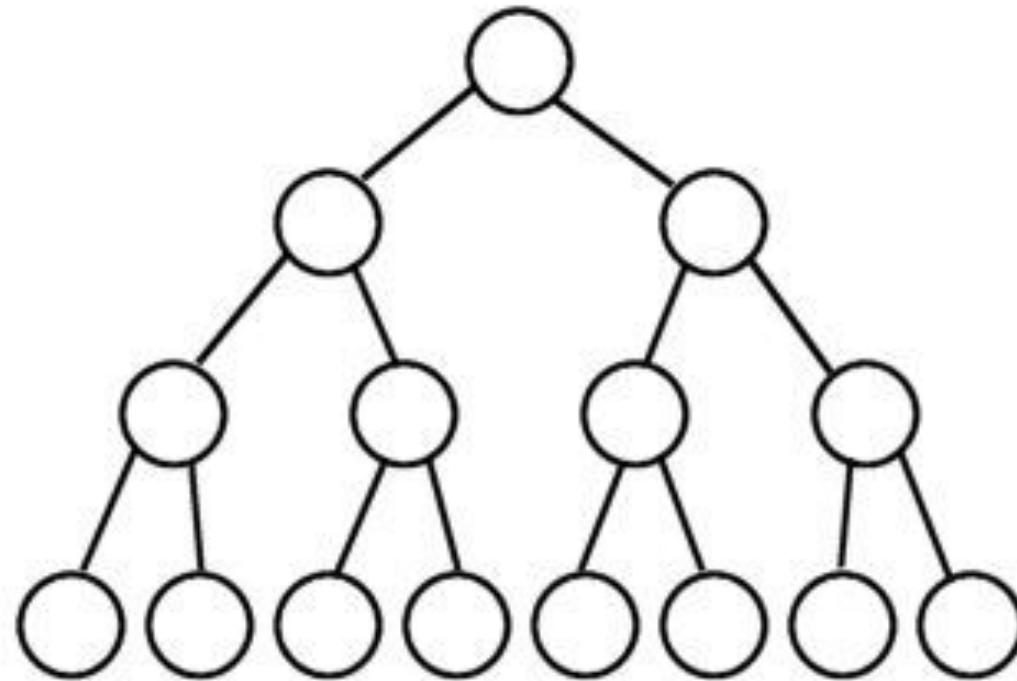
Árvore de Decisão

“Função que toma como entrada um vetor de valores de atributos e retorna uma “decisão” – um valor de saída único.” ([Definição do livro de Russell e Norvig](#))

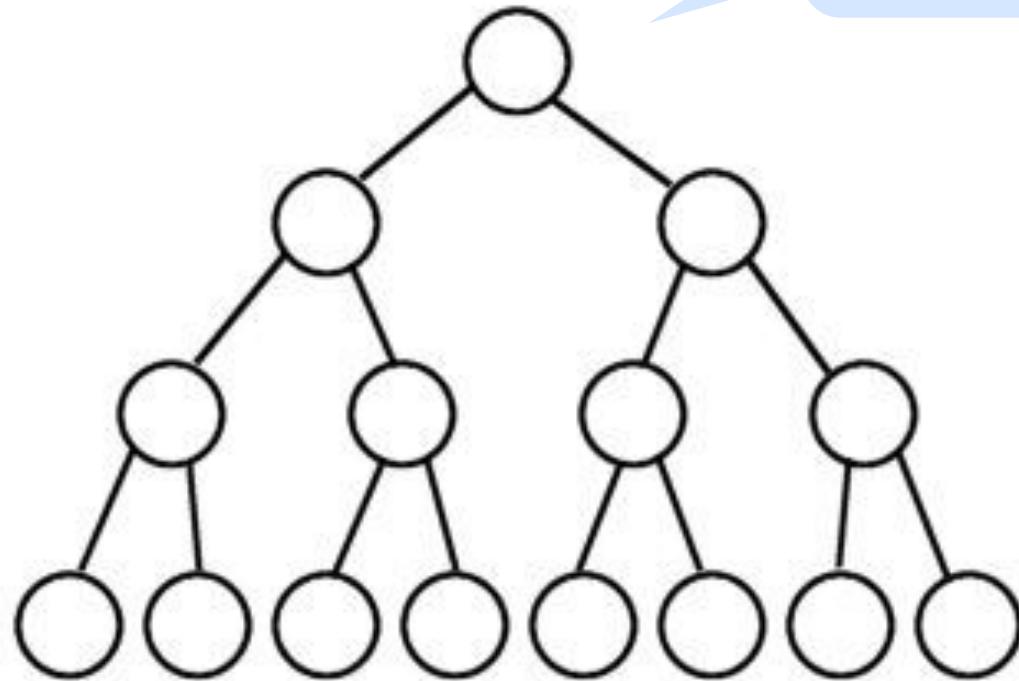
Os valores de entrada/saída podem ser **discretos** ou **contínuos**

- ▷ Árvores de Decisão podem lidar com problemas de **classificação e regressão**

Árvores

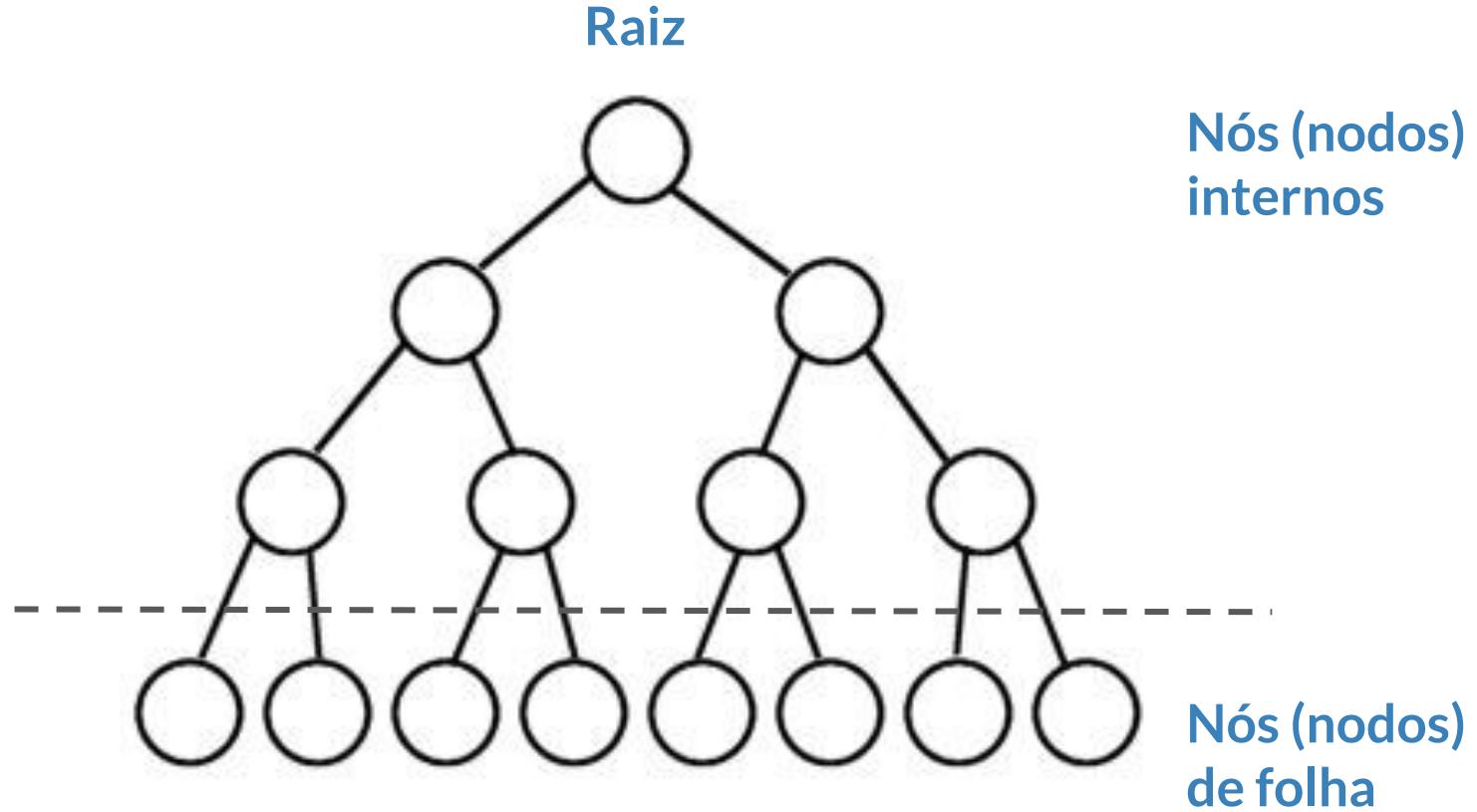


Árvores

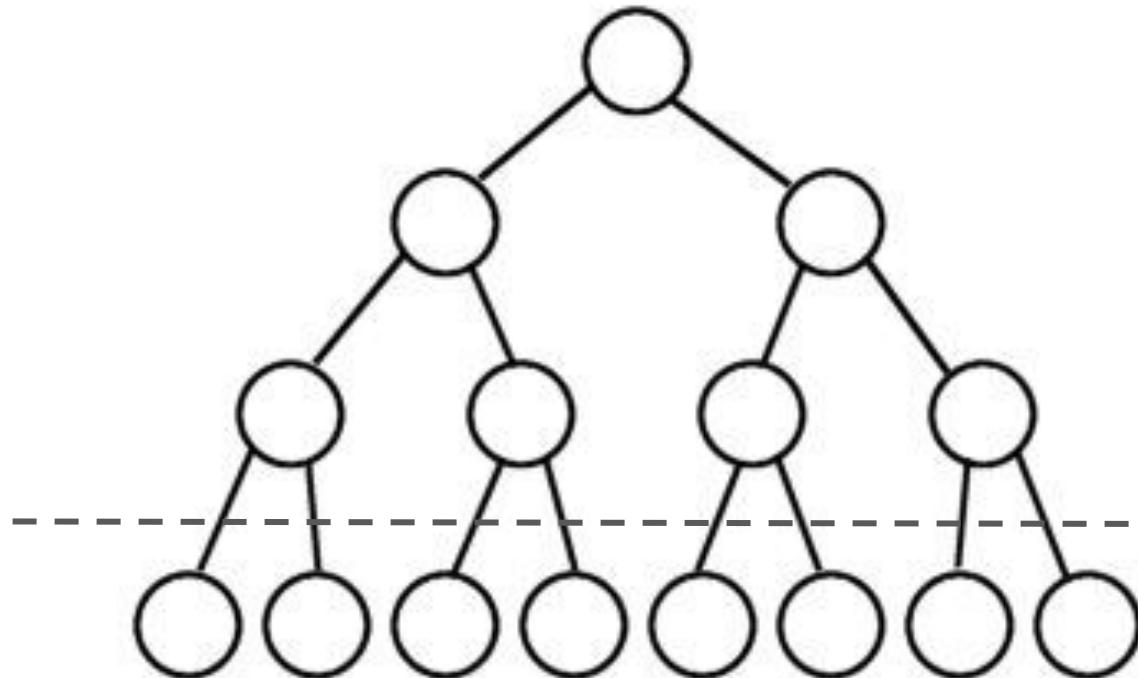


Oi, eu sou um tipo
específico de **grafo** :-)

Árvores



Árvores



Nós (nodos)
internos

Teste do valor
de um dos
atributos de
entrada

Nós (nodos)
de folha

Valor a ser
retornado

Árvore de Decisão

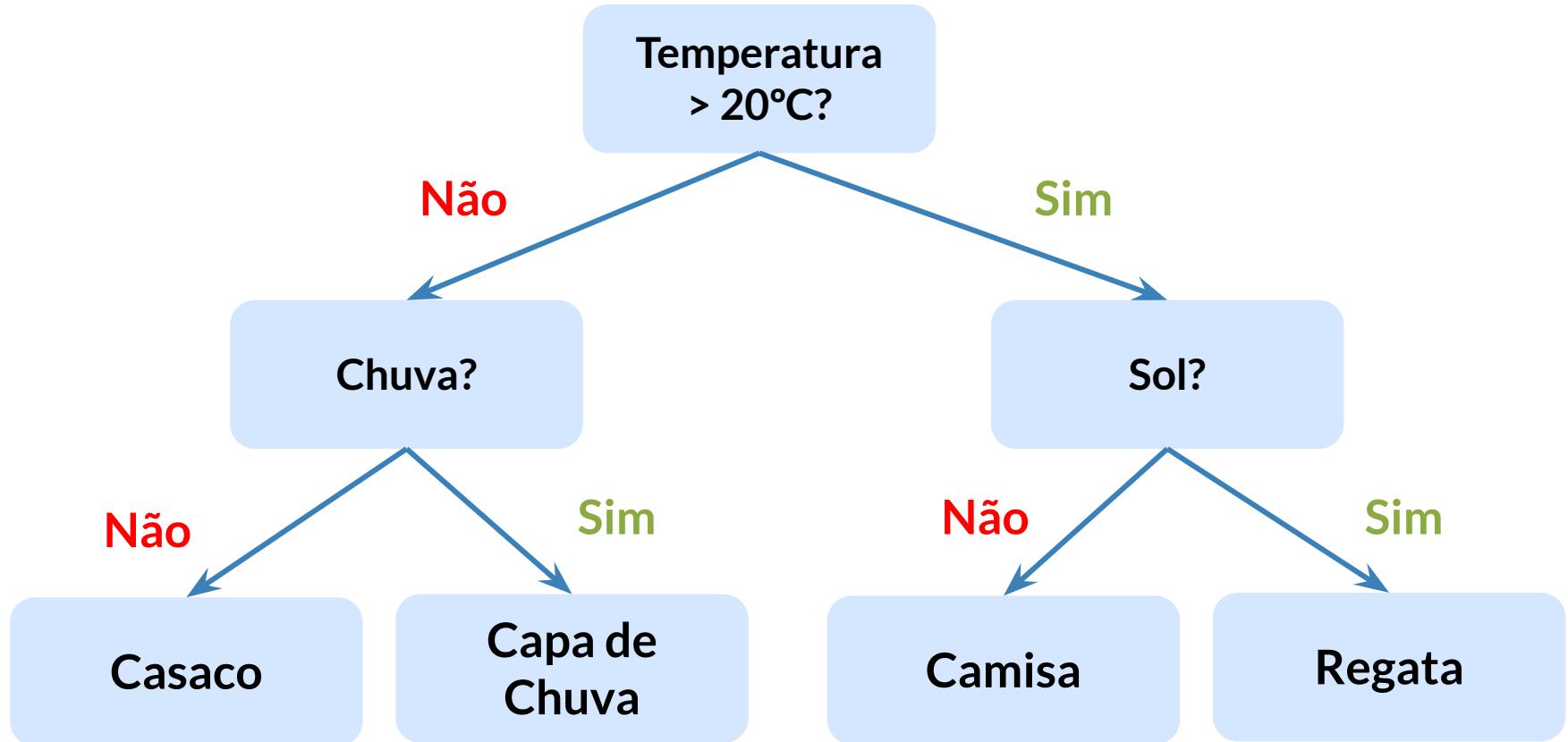
Decisão é tomada a partir de **sequência de testes**

- ▷ Várias perguntas **sim/não**
- ▷ Hierarquicamente arranjadas
- ▷ Estrutura da árvore determinada por meio de aprendizado

Muitas das decisões que tomamos no dia-a-dia podem ser compreendidas através de árvores de decisão

- ▷ Considerar **vários fatores** e seguir uma **lógica**

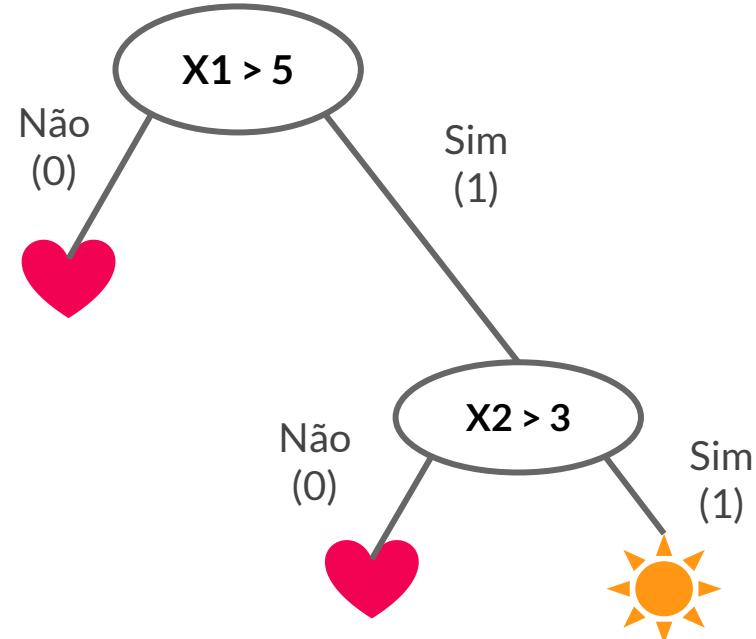
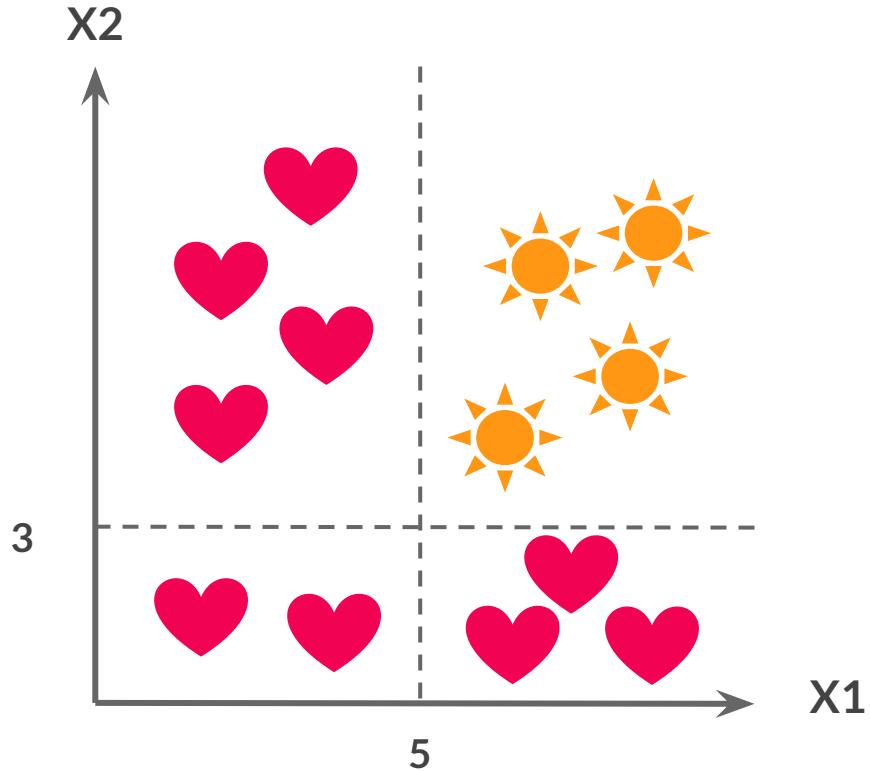
Com que roupa eu vou?



Árvore de Decisão

Abrangem todo o espaço de instâncias, ou seja, árvores de decisão podem fazer previsões para qualquer exemplo de entrada.

Árvore de Decisão



2. Indução de Árvores de Decisão

Como gerar?

Exemplo – Esperar uma mesa?

Problema: Estou em um restaurante e preciso decidir se vou esperar uma mesa.

Exemplo – Esperar uma mesa?

Problema: Estou em um restaurante e preciso decidir se vou esperar uma mesa.

É um problema de **classificação!**

Objetivo: Construir uma árvore de decisão para decidir a espera ou não de uma mesa em um restaurante. Aqui, o objetivo é aprender uma definição para o predicado *VaiEsperar* (sim/não).

Quais seriam possíveis atributos?

Exemplo – Esperar uma mesa?

Atributos disponíveis:

1. **Alternativa:** Há outro restaurante por perto?
2. **Bar:** Existe uma área de bar confortável onde se possa esperar?
3. **Sex/Sáb:** Hoje é sexta ou sábado?
4. **Faminto:** Estamos com fome?
5. **Clientes:** Quantas pessoas estão no restaurante? (Nenhum, Alguns e Cheio).
6. **Preço:** Faixa de preços do restaurante (\$, \$\$, \$\$\$).
7. **Chovendo:** Está chovendo do lado de fora?
8. **Reserva:** Fizemos uma reserva?
9. **Tipo:** Francês, italiano, tailandês, hambúrguer, etc.
10. **EsperaEstimada:** 0-10 minutos, 10-30, 30-60, >60

Exemplo – Esperar uma mesa?

Exemplo	Alt.	Bar	Sex/Sab	Fome	Cientes	Pr.	Chuva	Res.	Tipo	Estim.	VaiEsperar
x1	Sim	Não	Não	Sim	Alguns	\$\$\$	Não	Sim	Francês	0-10	Sim
x2	Sim	Não	Não	Sim	Cheio	\$	Não	Não	Tailandês	30-60	Não
x3	Não	Sim	Não	Não	Alguns	\$	Não	Não	Burguer	0-10	Sim
x4	Sim	Não	Sim	Sim	Cheio	\$	Sim	Não	Tailandês	10-30	Sim
x5	Sim	Não	Sim	Não	Cheio	\$\$\$	Não	Sim	Francês	>60	Não
x6	Não	Sim	Não	Sim	Alguns	\$\$	Sim	Sim	Italiano	0-10	Sim
x7	Não	Sim	Não	Não	Nenhum	\$	Sim	Não	Burguer	0-10	Não
x8	Não	Não	Não	Sim	Alguns	\$\$	Sim	Sim	Tailandês	0-10	Sim
x9	Não	Sim	Sim	Não	Cheio	\$	Sim	Não	Burguer	>60	Não
x10	Sim	Sim	Sim	Sim	Cheio	\$\$\$	Não	Sim	Italiano	10-30	Não
x11	Não	Não	Não	Não	Nenhum	\$	Não	Não	Tailandês	0-10	Não
x12	Sim	Sim	Sim	Sim	Cheio	\$	Não	Não	Burguer	30-60	Sim

Exemplo – Esperar uma mesa?

Exemplo	Alt.	Bar	Sex/Sab	Fome	Cientes	Pr.	Chuva	Res.	Tipo	Estim.	VaiEsperar
x1	Sim	Não	Não	Sim	Alguns	\$\$\$	Não	Sim	Francês	0-10	Sim
x2	Sim	Não	Não	Sim	Cheio	\$	Não	Não	Tailandês	30-60	Não
x3	Não	Sim	Não	Não	Alguns	\$	Não	Não	Burguer	0-10	Sim
x4	Sim	Não	Sim	Sim	Cheio	\$	Sim	Não	Tailandês	10-30	Sim
x5	Sim	Não	Sim	Não	Cheio	\$\$\$	Não	Sim	Francês	>60	Não
x6	Não	Sim	Não	Sim	Alguns	\$\$	Sim	Sim	Italiano	0-10	Sim
x7	Não	Sim	Não	Não	Nenhum	\$	Sim	Não	Burguer	0-10	Não
x8	Não	Não	Não	Sim	Alguns	\$\$	Sim	Sim	Tailandês	0-10	Sim
x9	Não	Sim	Sim	Não	Cheio	\$	Sim	Não	Burguer	>60	Não
x10	Sim	Sim	Sim	Sim	Cheio	\$\$\$	Não	Sim	Italiano	10-30	Não
x11	Não	Não	Não	Não	Nenhum	\$	Não	Não	Tailandês	0-10	Não
x12	Sim	Sim	Sim	Sim	Cheio	\$	Não	Não	Burguer	30-60	Sim

Atributos

Exemplo – Esperar uma mesa?

Exemplo	Alt.	Bar	Sex/Sab	Fome	Cientes	Pr.	Chuva	Res.	Tipo	Estim.	VaiEsperar
x1	Sim	Não	Não	Sim	Alguns	\$\$\$	Não	Sim	Francês	0-10	Sim
x2	Sim	Não	Não	Sim	Cheio	\$	Não	Não	Tailandês	30-60	Não
x3	Não	Sim	Não	Não	Alguns	\$	Não	Não	Burguer	0-10	Sim
x4	Sim	Não	Sim	Sim	Cheio	\$	Sim	Não	Tailandês	10-30	Sim
x5	Sim	Não	Sim	Não	Cheio	\$\$\$	Não	Sim	Francês	>60	Não
x6	Não	Sim	Não	Sim	Alguns	\$\$	Sim	Sim	Italiano	0-10	Sim
x7	Não	Sim	Não	Não	Nenhum	\$	Sim	Não	Burguer	0-10	Não
x8	Não	Não	Não	Sim	Alguns	\$\$	Sim	Sim	Tailandês	0-10	Sim
x9	Não	Sim	Sim	Não	Cheio	\$	Sim	Não	Burguer	>60	Não
x10	Sim	Sim	Sim	Sim	Cheio	\$\$\$	Não	Sim	Italiano	10-30	Não
x11	Não	Não	Não	Não	Nenhum	\$	Não	Não	Tailandês	0-10	Não
x12	Sim	Sim	Sim	Sim	Cheio	\$	Não	Não	Burguer	30-60	Sim

Exemplo – Esperar uma mesa?

Queremos uma árvore que seja **consistente com os exemplos** e seja a **menor possível**.

Exemplo – Esperar uma mesa?

Queremos uma árvore que seja **consistente com os exemplos** e seja a **menor possível**.

O algoritmo de aprendizagem em árvore de decisão adota uma estratégia gulosa de **dividir para conquistar**

- ▷ Sempre testar o atributo mais importante primeiro (o que faz mais diferença para classificação)

Exemplo – Esperar uma mesa?

Queremos uma árvore que seja **consistente com os exemplos** e seja a **menor possível**.

O algoritmo de aprendizagem em árvore de decisão adota uma estratégia gulosa de **dividir para conquistar**

- ▷ Sempre testar o atributo mais importante primeiro (o que faz mais diferença para classificação)

Qual atributo podemos testar?

Exemplo – Esperar uma mesa?

Vamos testar o atributo “Tipo”?

1	3	4	6	8	12
2	5	7	9	10	11

Exemplo – Esperar uma mesa?

Vamos testar o atributo “Tipo”?

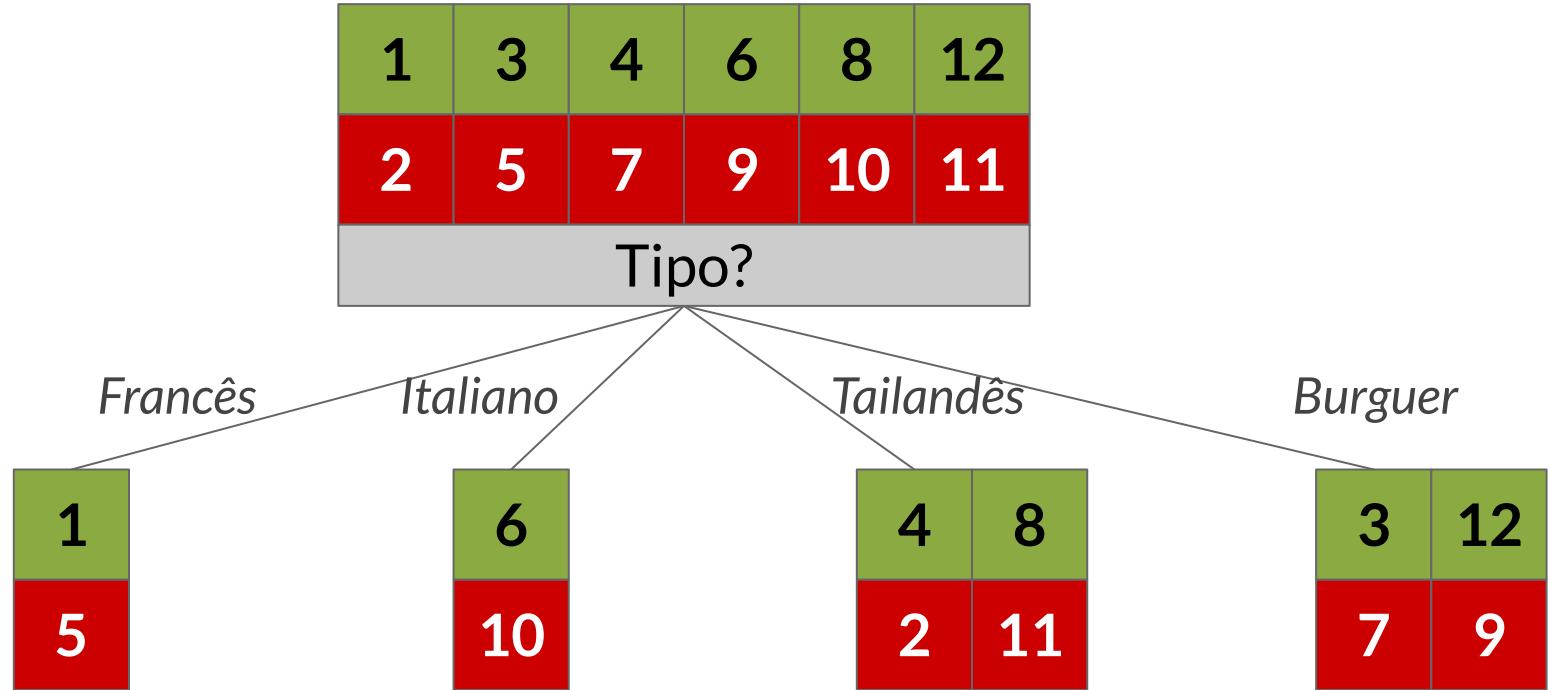
1	3	4	6	8	12
2	5	7	9	10	11

Decidi esperar

Decidi não esperar

Exemplo – Esperar uma mesa?

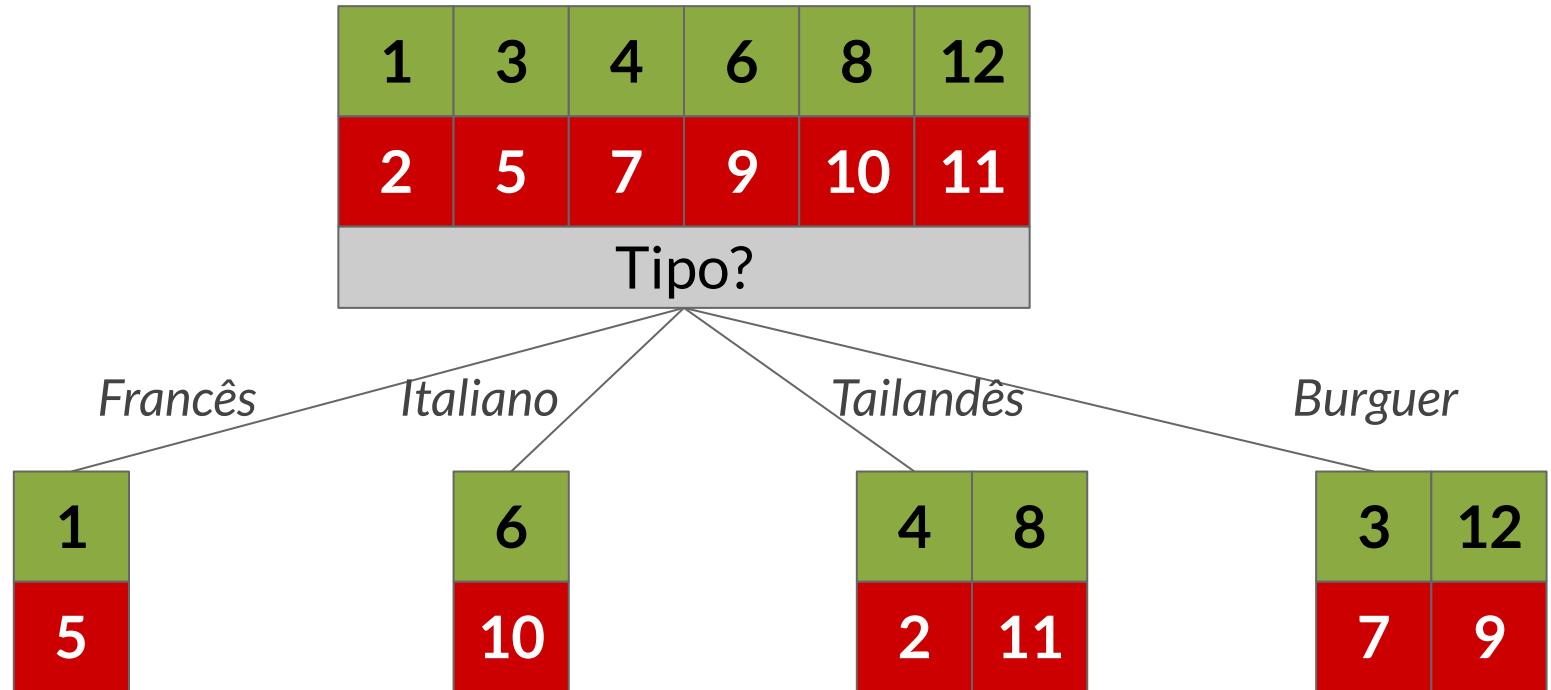
Vamos testar o atributo “Tipo”?



É um bom atributo?

Exemplo – Esperar uma mesa?

Vamos testar o atributo “Tipo”?



“Tipo” é um **atributo fraco** porque cada resultado tem o mesmo número de exemplos positivos e negativos.

Exemplo – Esperar uma mesa?

Algum outro atributo?

Exemplo	Alt.	Bar	Sex/Sab	Fome	Cientes	Pr.	Chuva	Res.	Tipo	Estim.	VaiEsperar
x1	Sim	Não	Não	Sim	Alguns	\$\$\$	Não	Sim	Francês	0-10	Sim
x2	Sim	Não	Não	Sim	Cheio	\$	Não	Não	Tailandês	30-60	Não
x3	Não	Sim	Não	Não	Alguns	\$	Não	Não	Burguer	0-10	Sim
x4	Sim	Não	Sim	Sim	Cheio	\$	Sim	Não	Tailandês	10-30	Sim
x5	Sim	Não	Sim	Não	Cheio	\$\$\$	Não	Sim	Francês	>60	Não
x6	Não	Sim	Não	Sim	Alguns	\$\$	Sim	Sim	Italiano	0-10	Sim
x7	Não	Sim	Não	Não	Nenhum	\$	Sim	Não	Burguer	0-10	Não
x8	Não	Não	Não	Sim	Alguns	\$\$	Sim	Sim	Tailandês	0-10	Sim
x9	Não	Sim	Sim	Não	Cheio	\$	Sim	Não	Burguer	>60	Não
x10	Sim	Sim	Sim	Sim	Cheio	\$\$\$	Não	Sim	Italiano	10-30	Não
x11	Não	Não	Não	Não	Nenhum	\$	Não	Não	Tailandês	0-10	Não
x12	Sim	Sim	Sim	Sim	Cheio	\$	Não	Não	Burguer	30-60	Sim

Exemplo – Esperar uma mesa?

Vamos testar o atributo “Clientes”?

1	3	4	6	8	12
2	5	7	9	10	11

Decidi esperar

Decidi não esperar

Exemplo – Esperar uma mesa?

Vamos testar o atributo “**Cientes**”?

1	3	4	6	8	12
2	5	7	9	10	11
Clientes?					

Nenhum

7	11
---	----

Alguns

1	3	6	8
---	---	---	---

Cheio

4	12
2	5

9 10

“Cliente” é um atributo bastante importante.

Exemplo – Esperar uma mesa?

Vamos testar o atributo “**Cientes**”?

1	3	4	6	8	12
2	5	7	9	10	11
Clientes?					

Nenhum

7	11
---	----

Alguns

1	3	6	8
---	---	---	---

Cheio

4	12
2	5
9	10

**Novo problema de aprendizagem
de árvore de decisão com menos
exemplos e um atributo a menos.**

Exemplo – Esperar uma mesa?

1	3	4	6	8	12
2	5	7	9	10	11
Clientes?					

Nenhum

7	11
---	----

Alguns

1	3	6	8
---	---	---	---

Cheio

4	12		
2	5	9	10
Faminto?			

Não

5	9
---	---

Sim

4	12
2	10

Árvores de Decisão

função APRENDIZAGEM-EM-ÁRVORE-DE-DECISÃO(*exemplos, atributos, exemplos-pais*)

retorna uma árvore de decisão

se *exemplos* é vazio **então retornar** VALOR-DA-MAIORIA (*exemplos_pais*)

senão se todos os *exemplos* têm a mesma classificação **então retornar** a classificação

senão se *atributos* é vazio **então retornar** VALOR-DA-MAIORIA(*exemplos*)

senão

$A \leftarrow \operatorname{argmax}_{a \in \text{atributes}} \text{IMPORTÂNCIA}(a, \text{exemplos})$

árvore \leftarrow uma nova árvore de decisão com teste de raiz *A*

para cada valor v_k de *A* **faça**

exs $\leftarrow \{e : e \in \text{exemplos} \text{ e } e.A = v_k\}$

subárvore \leftarrow APRENDIZAGEM-EM-ÁRVORE-DE-DECISÃO (*exs, atributos — A, exemplos*)

adicionar uma ramiificação à árvore com rótulo ($A = v_k$) e subárvore *subárvore*

retornar *árvore*

Árvores de Decisão

função APRENDIZAGEM-EM-ÁRVORE-DE-DECISÃO(*exemplos, atributos, exemplos-pais*)

retorna uma árvore de decisão

se *exemplos* é vazio **então retornar** VALOR-DA-MAIORIA (*exemplos_pais*)

senão se todos os *exemplos* têm a mesma classificação **então retornar** a classificação

senão se *atributos* é vazio **então retornar** VALOR-DA-MAIORIA(*exemplos*)

senão

A $\leftarrow \operatorname{argmax}_{a \in \text{atributes}} \text{IMPORTÂNCIA}(a, \text{exemplos})$

árvore \leftarrow uma nova árvore de decisão com teste de raiz *A*

para cada valor v_k de *A* faça

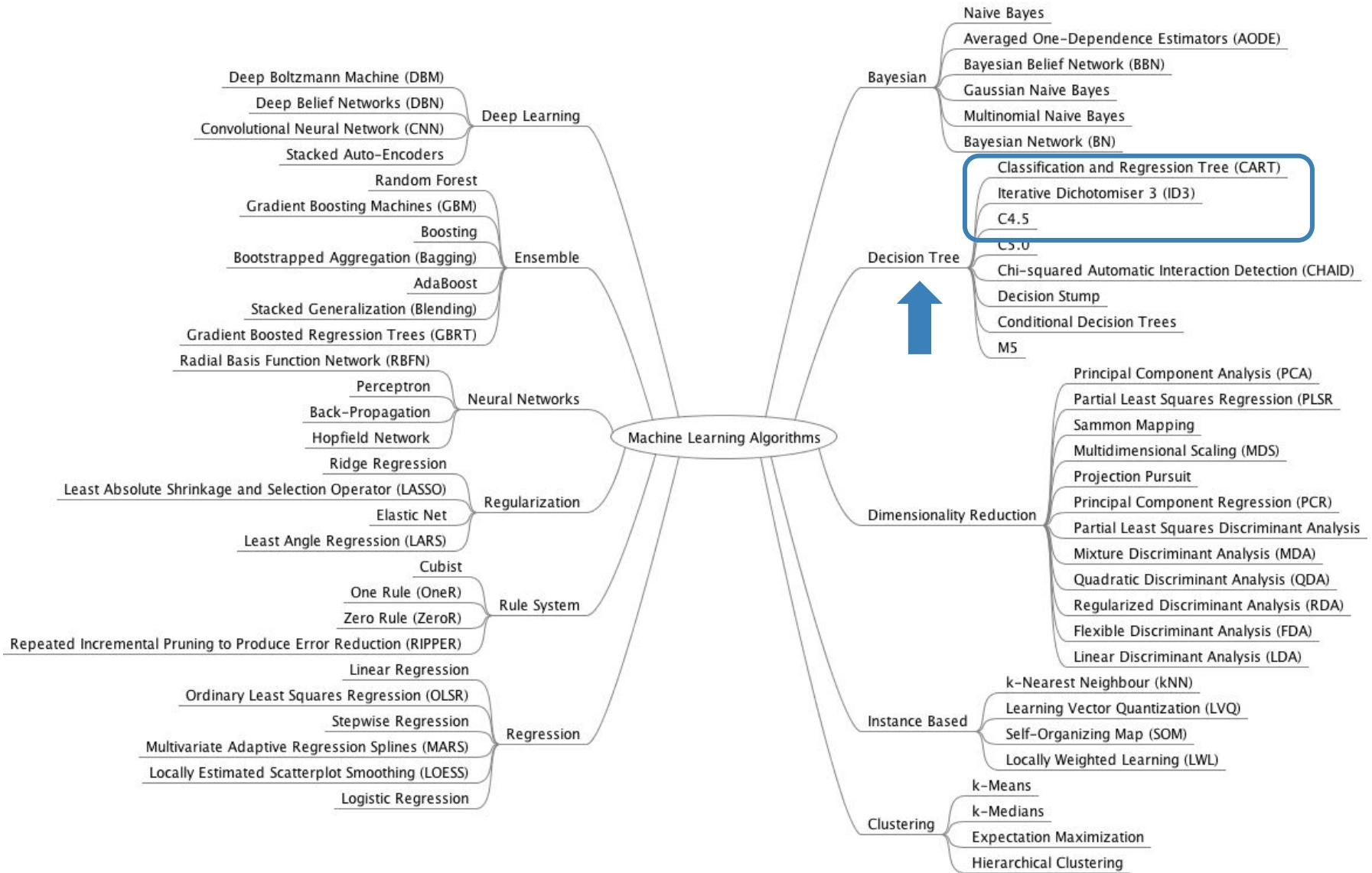
exs $\leftarrow \{e : e \in \text{exemplos} \text{ e } e.A = v_k\}$

subárvore \leftarrow APRENDIZAGEM-EM-ÁRVORE-DE-DECISÃO (*exs, atributos — A, exemplos*)

adicionar uma ramiificação à árvore com rótulo ($A = v_k$) e subárvore *subárvore*

retornar árvore

Árvores de Decisão



Árvores de Decisão

Como avaliar **matematicamente** o ganho de cada atributo?

- ▷ Ou seja, como decidir o atributo mais importante?

Árvores de Decisão

Como avaliar matematicamente o ganho de cada atributo?

- ▷ Ou seja, como decidir o atributo mais importante?

Próxima aula!

3. Vantagens e Desvantagens

Quais são?

ADs – Vantagens

Flexibilidade

- ▷ Não assumem nenhuma distribuição para os dados
- ▷ Elas são métodos não paramétricos
- ▷ O espaço de objetos é dividido em subespaços, e a cada subespaço é ajustado com diferentes modelos
- ▷ Cobertura exaustiva do espaço de instâncias

ADs – Vantagens

Seleção de atributos

- ▶ Produz modelos que tendem a ser bastante **robustos contra a adição de atributos irrelevantes e redundantes**

ADs – Vantagens

Interpretabilidade

- ▷ Decisões complexas e globais podem ser aproximadas por uma **série de decisões mais simples e locais** (subproblemas)
- ▷ Todas as **decisões são baseadas nos valores dos atributos** usados para descrever o problema

ADs – Vantagens

Eficiência

- ▷ O algoritmo para aprendizado de árvore de decisão é um **algoritmo guloso** construído de cima para baixo (*top-down*)
- ▷ Utiliza estratégia **dividir para conquistar** sem *backtracking* (voltar quando existe uma falha)

ADs – Desvantagens

Valores ausentes

- ▷ Algoritmos devem empregar mecanismos especiais para abordar **falta de valores**

Atributos contínuos

- ▷ Alguns autores estimam que a operação de ordenação consuma 70% do tempo necessário para induzir uma árvore de decisão em grandes conjuntos de dados com muitos atributos contínuos

ADs – Desvantagens

Instabilidade

- ▷ Pequenas variações no conjunto de treinamento podem produzir **grandes variações na árvore final**
- ▷ Forte tendência a inferência feitas próximo das folhas serem menos confiáveis que aquelas feitas próximas da raiz

Referências

RUSSEL, Stuart J. e NORVIG, James. Artificial Intelligence: a modern approach. (Cap. 18 - 3^a ed. ou Cap. 19 - 4^a ed.)

Material de aula da Profa. Cristiane Nobre.