




# Árvore de Decisão: **CART**

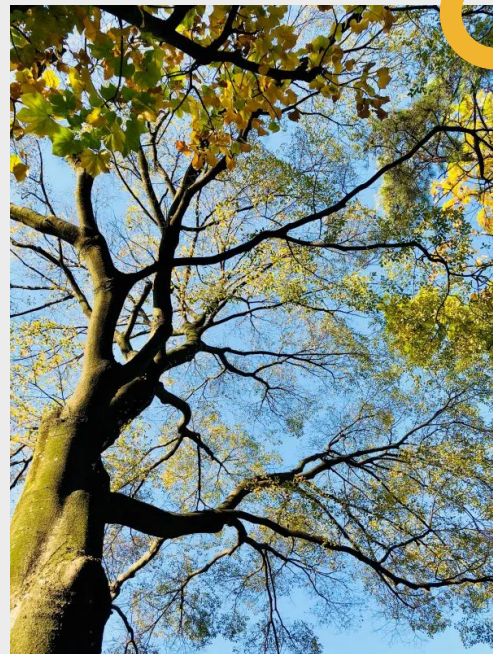
---



**Alunos:** Fernanda Mendes, Hugo Portes, Luiz Vitor, Pedro Olyntho

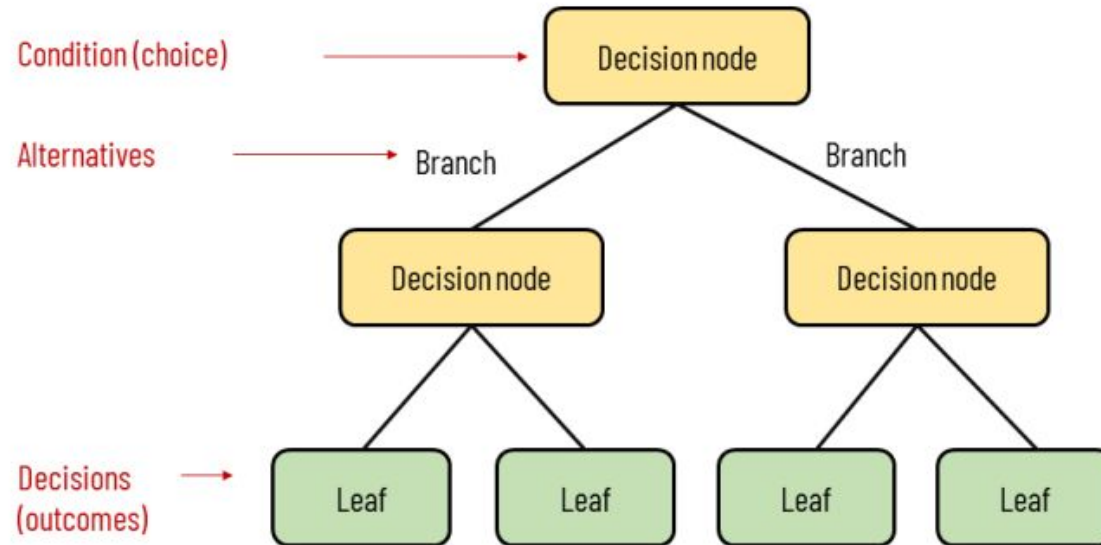
# Recapitulando... Árvores de Decisão

- Conjunto de **regras de decisão**
- Nós folha representam as classes a serem atribuídas
- Construção: seleção dos **atributos mais relevantes** para a divisão dos dados em subconjuntos de maneira eficiente
- **Regras de decisão hierárquicas** que permite classificar ou prever com base nas características dos dados



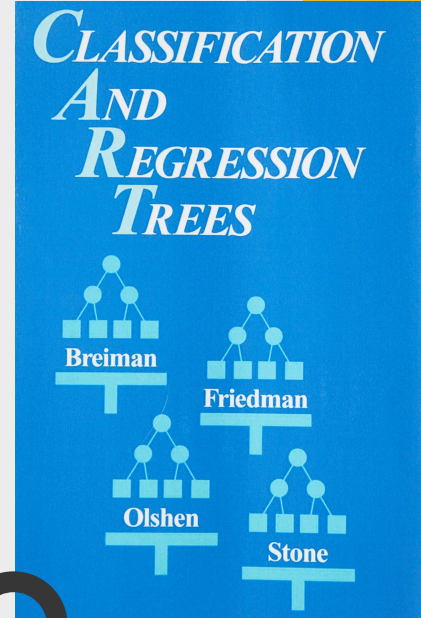
# Recapitulando... Árvores de Decisão

## Elements of a decision tree



# CART: Introdução

- Introduzido por Breiman em **1984**
- CART → Classification and Regression Tree
- Capaz de realizar:
  - **Classificação:** variável alvo é categórica
  - **Regressão:** variável alvo é contínua



# CART: Funcionamento

- Resultado é sempre uma árvore **binária**
- Para a escolha de atributos utiliza o **Critério Gini**
  - Quando este índice é igual a zero, o nó é puro.
  - Valor máximo é 0.5 → nó é impuro
- **Particionamento Recursivo** - divisão sequencial até que o último subconjunto puro seja encontrado ou seja alcançado o nº máximo de folhas possível.
- Realiza **pós poda**

$$Gini = 1 - \sum_{i=1}^n (p_i)^2$$

where  $p_i$  is the probability of an object being classified to a particular class.



# CART

## Vantagens

- Fácil de entender e interpretar
- Lida com dados ausentes
- Flexibilidade: problemas de regressão e de classificação
- Atributos categóricos e numéricos

## Desvantagens

- Tendência ao overfitting
- Pode gerar árvores instáveis
- Sensibilidade a pequenas variações nos dados

# Algoritmos de Árvore de Decisão: Comparativo

	<b>ID3</b>	<b>C4.5</b>	<b>CART</b>
Tipos de atributos	Discretos	Contínuos e Discretos	Contínuos e Discretos
Lida com dados ausentes	Não	Sim	Sim
Escolha de atributos	Ganho de informação	Gain Ratio	Índice de Gini
Tipo de partição	Sem restrições	Sem restrições	Binária
Realiza pós poda	Não	Sim	Sim
Tipo de problema	Classificação	Classificação	Classificação e Regressão

# Referências

- COMPARATIVE STUDY ID3, CART AND C4.5 DECISION TREE ALGORITHM: A SURVEY
- <https://www.geeksforgeeks.org/cart-classification-and-regression-tree-in-machine-learning>
- A comparative study of decision tree ID3 and C4.5
- Study and Analysis of Decision Tree Based Classification Algorithms
- Analysis of Various Decision Tree Algorithms for Classification in Data Mining