



Human-Centered Data & AI

# Vinicius Caridá, Ph.D.



Manager of Digital Customer Service  
Platforms, Data and AI - Itaú Unibanco

MBA Professor - FIAP

Google Developer Expert – Machine Learning



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida

“

# Zero to Hero Machine Learning na AWS

Parte 2/5

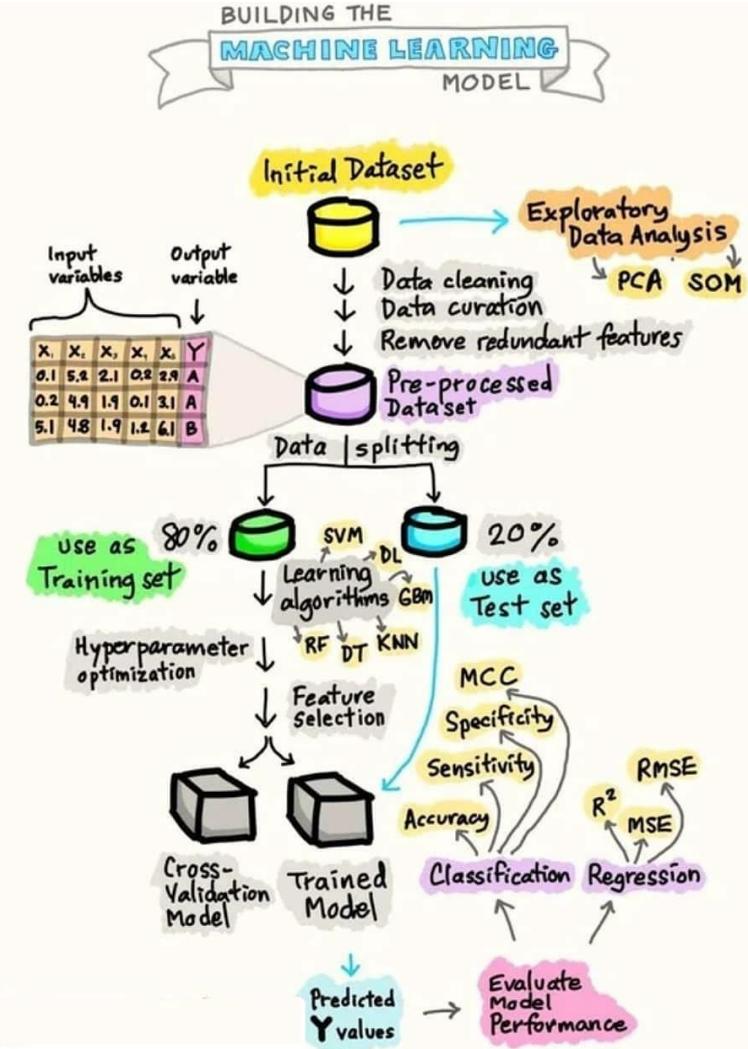
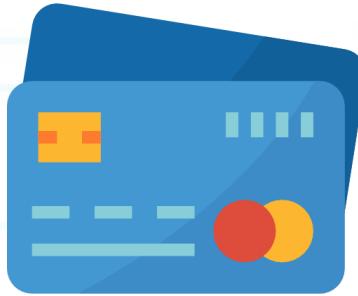
“

# Classificação

# Classificação



# Classificação



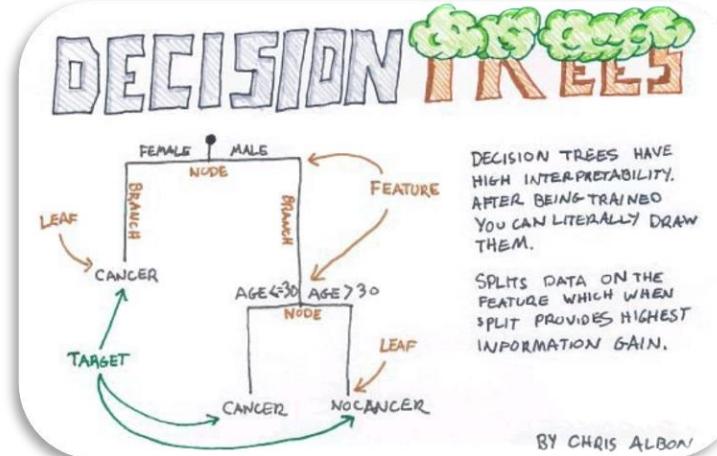
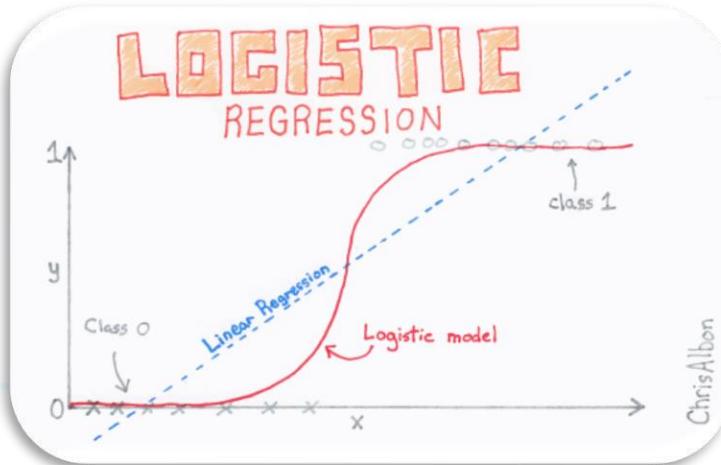
# Classificação



97%



# Classificação



## GAUSSIAN NAIVE BAYES CLASSIFIER

"Gaussian" because this is a normal distribution

$$P(\text{class} | \text{data}) = \frac{P(\text{data} | \text{class}) \times p(\text{class})}{P(\text{data})}$$

We don't calculate this in naive bayes classifiers

This is our prior belief

Chris Albon

## KNN NEIGHBORHOOD SIZE

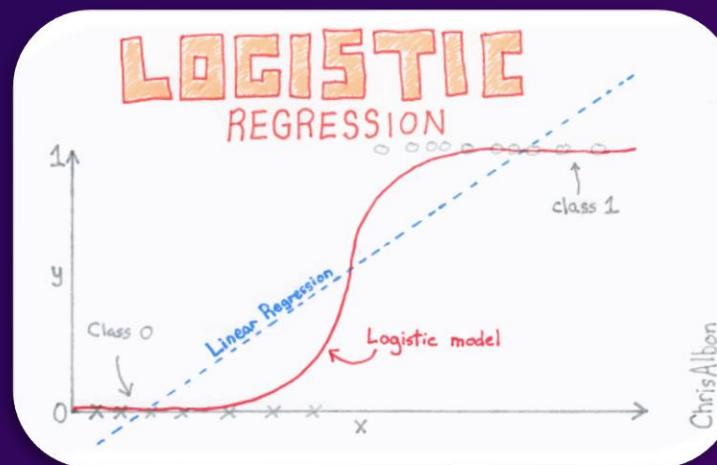
Small  $\downarrow$  K = Low Bias, High Variance

LARGE  $\uparrow$  K = High Bias, Low Variance

BY CHRIS ALBON

“

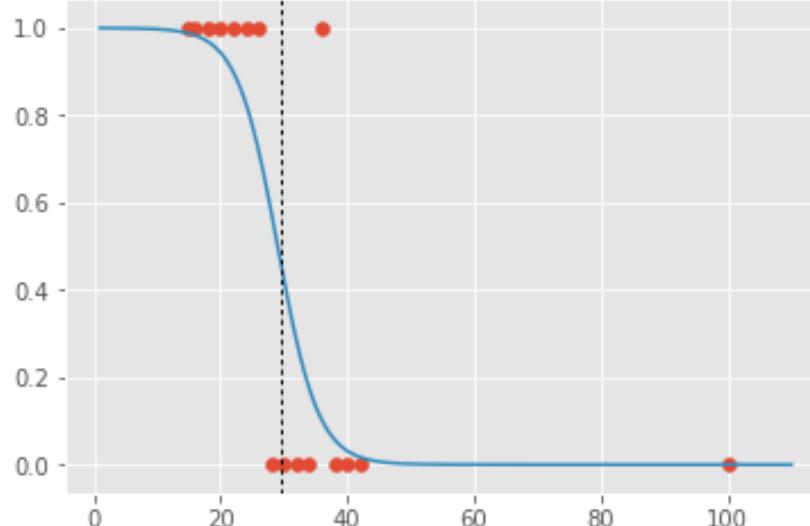
# Regressão Logística



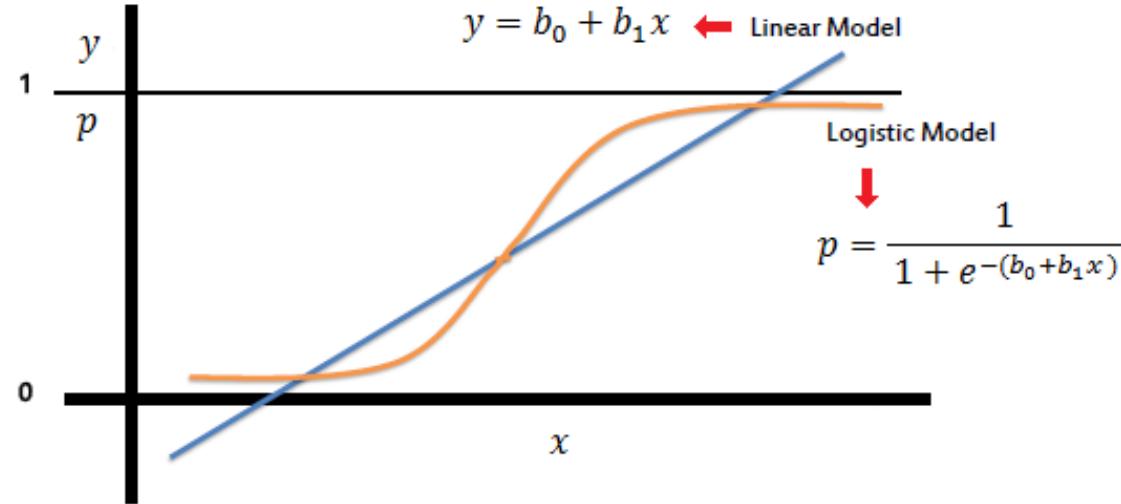
ChrisAlbon

# Regressão Logística

- A intuição por trás de regressão logística é bastante simples: em vez de acharmos a reta que melhor se ajusta aos dados, vamos achar uma curva em formato de 'S' que melhor se ajusta aos dados:



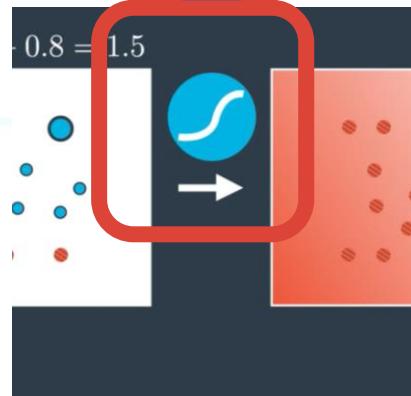
# Regressão Logística



A regressão logística é semelhante a uma regressão linear, mas a curva é construída usando o logaritmo natural das “probabilidades” da variável-alvo.

# Regressão Logística

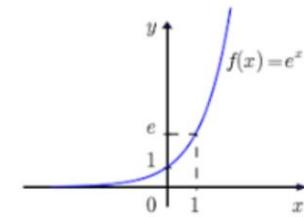
Redes Neurais



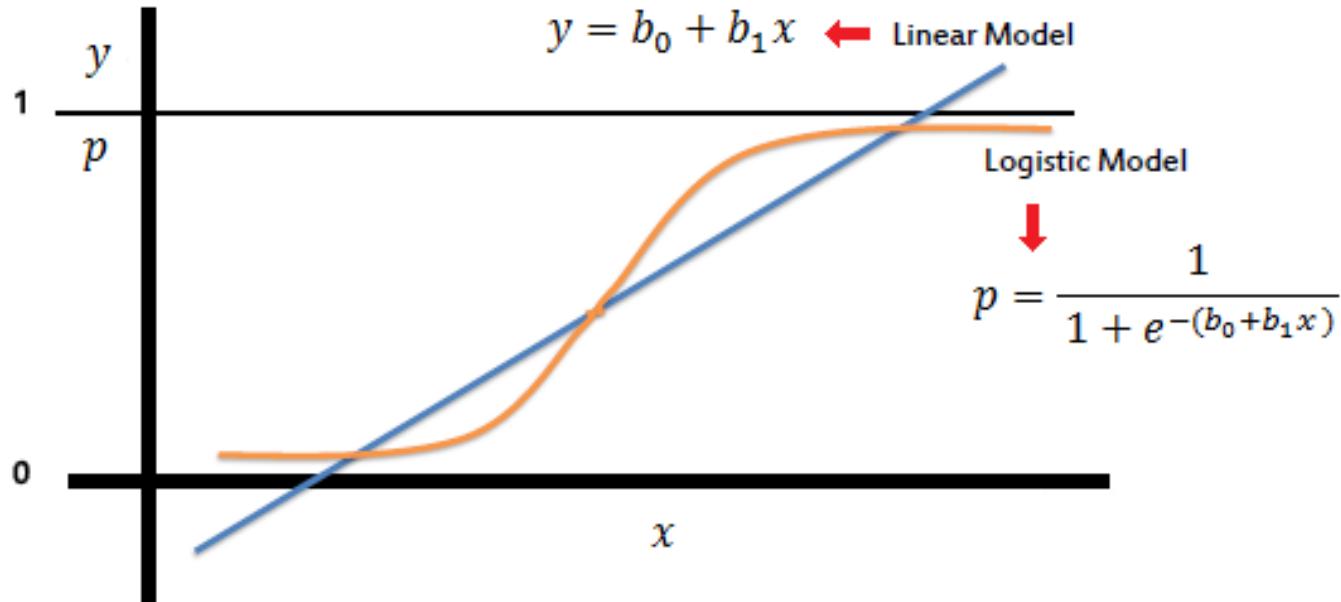
$$f(x) = \frac{1}{1 + e^{-x}} \text{ para todo } x \text{ real.}$$



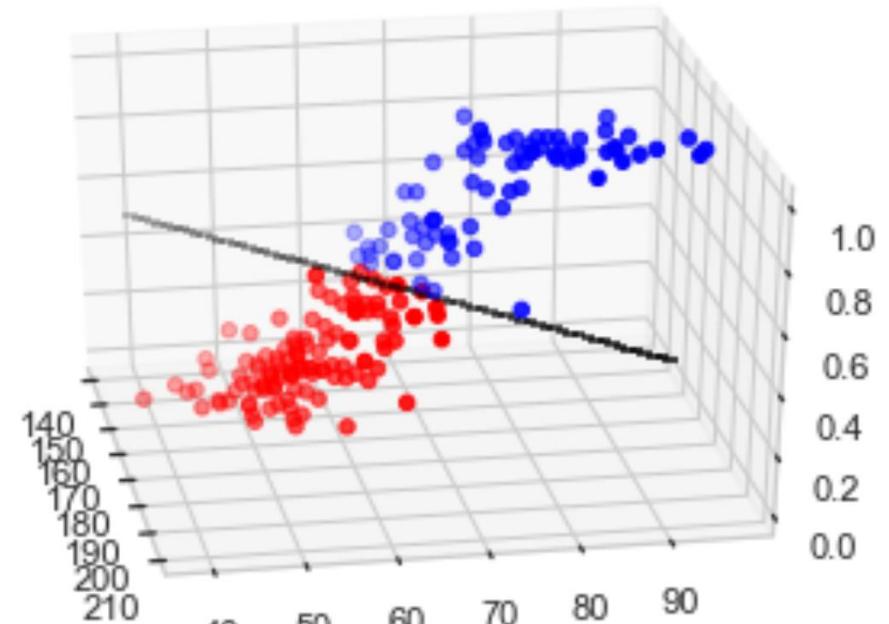
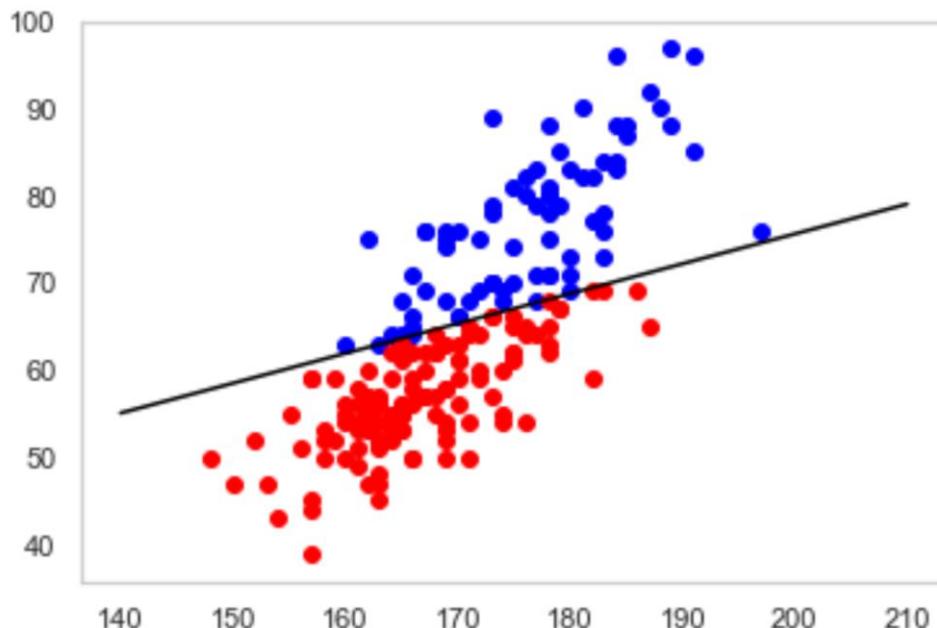
A função **exponencial** natural, denotada  $e^x$  ou  $\exp(x)$  é a função **exponencial** cuja base é o número de Euler (um número irracional que **vale** aproximadamente 2,718281828).



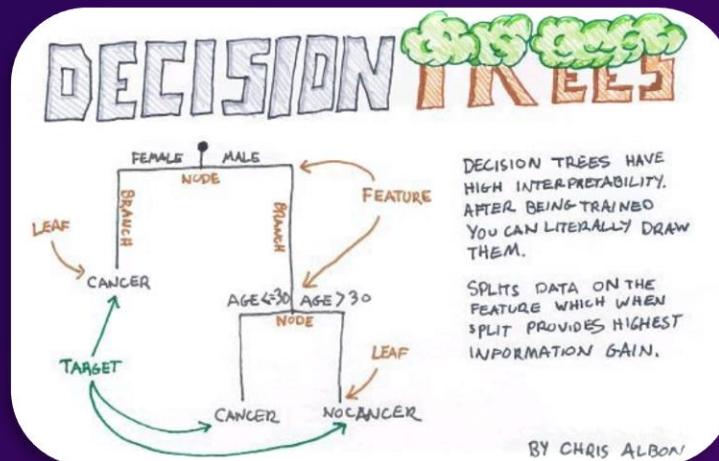
# Regressão Logística



# Regressão Logística



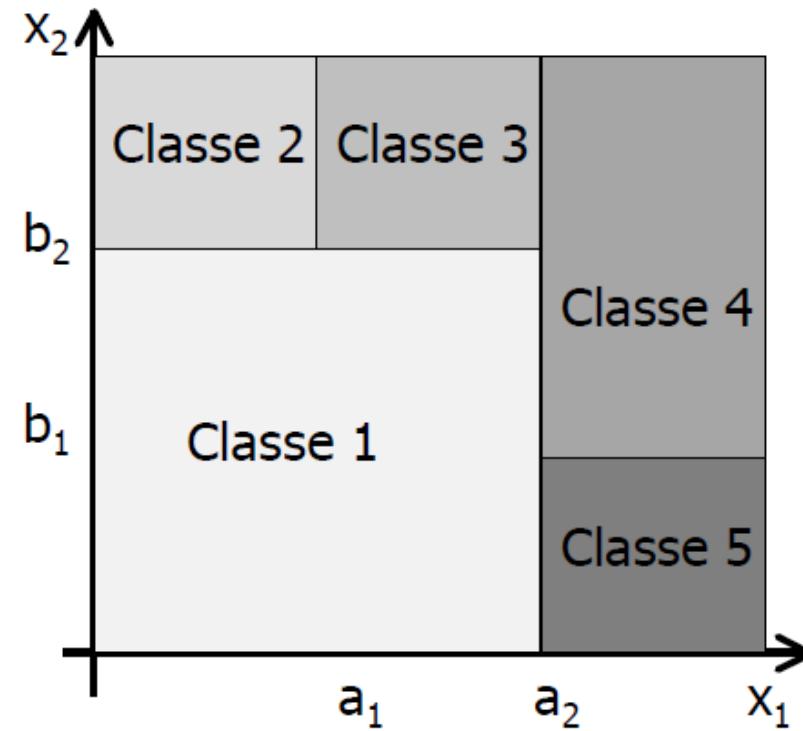
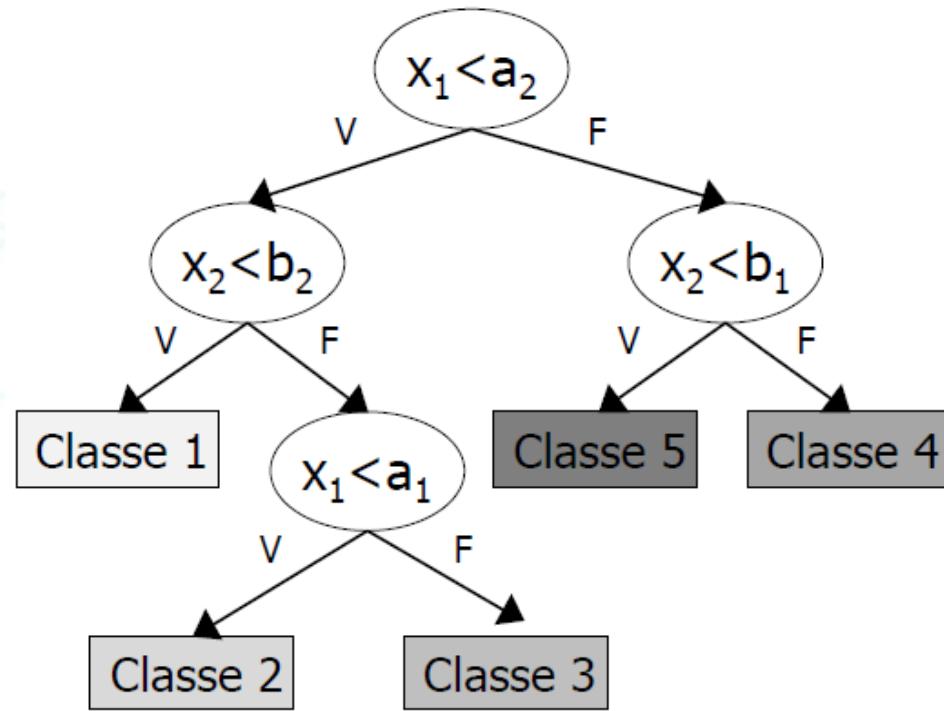
# Árvore de Decisão



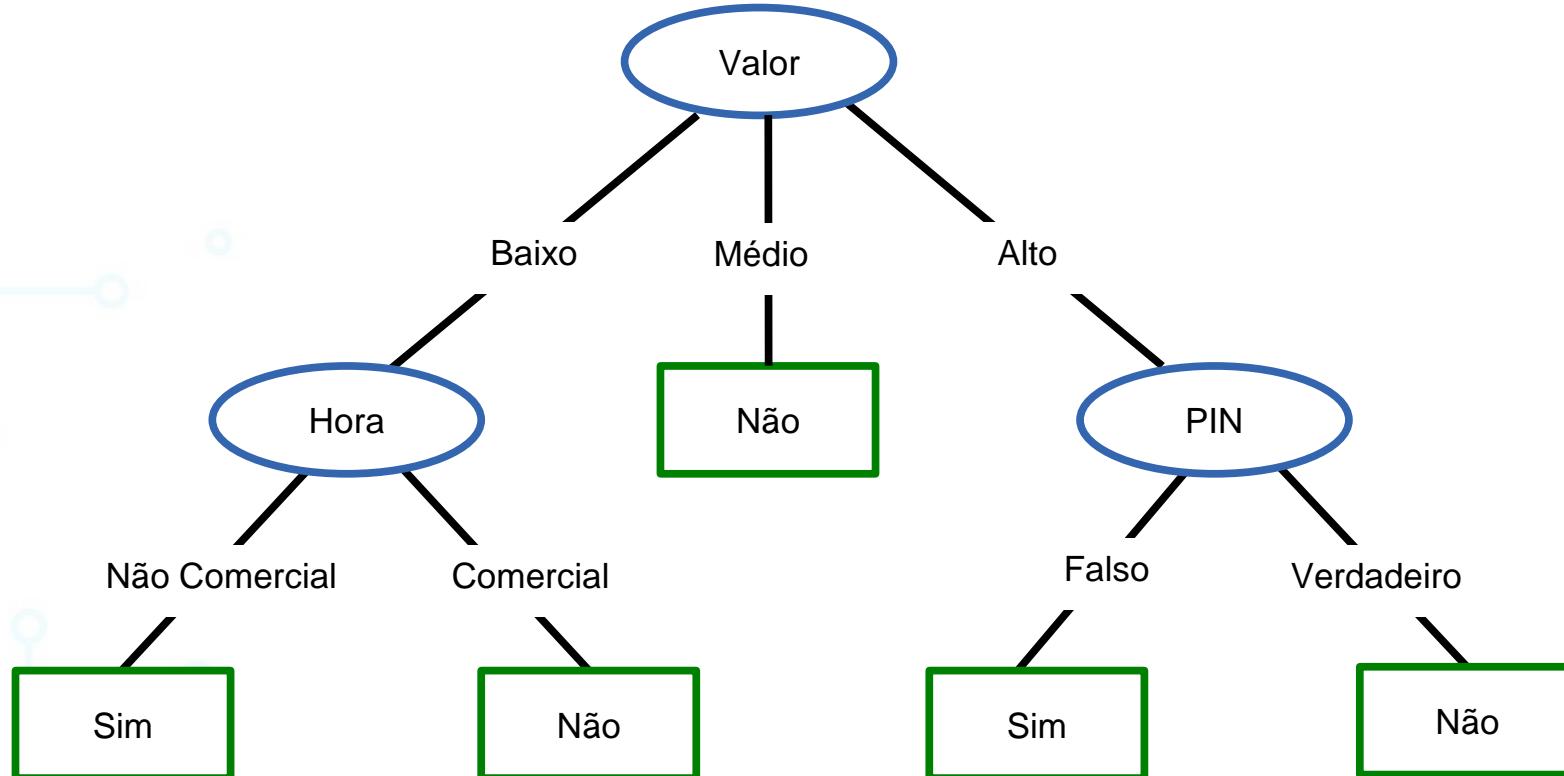
# Árvore de Decisão

- Uma das formas de aprendizado indutivo mais utilizadas
  - Aplicações práticas
  - Tarefas de Classificação (Aprendizado supervisionado)
- Aproximação de funções discretas robusta a ruídos
- Aprendizado de expressões disjuntivas

# Representação de uma AD



# Árvore de Decisão



# Árvore de Decisão

- ID3
  - Cria um ranking dos atributos mais adequados
  - Insere (com base no ranking) os atributos na árvore
  - Utiliza o ganho de informação (information gain) para criar o ranking
    - Método com base na entropia

# Árvore de Decisão

- Entropia
    - Medida definida na teoria da informação
    - Define a pureza de um conjunto de instâncias
    - Precisa de um conjunto de instâncias positivas e negativas
- $$\text{Entropia}(S) = - (p_+ \log_2 p_+) - (p_- \log_2 p_-)$$
- $p_+$ : proporção de instâncias positivas
- $p_-$ : proporção de instâncias negativas

# Árvore de Decisão

- O ganho de informação de um atributo é a redução esperada na entropia, causada pelo particionamento das instâncias de acordo com este atributo. Ou seja, o quanto se espera que a entropia seja reduzida quando se sabe o valor do atributo A.

$$\text{Ganho}(S, A) = \text{Entropia}(S) - \sum_{v \in \text{valores}(A)} \frac{|S_v|}{|S|} \text{Entropia}(S_v)$$

Valores(A): conjunto de todos os valores possíveis do atributo A

S<sub>v</sub>: subconjunto de S, no qual o atributo A possui valor v.

# Árvore de Decisão

Compra	Valor	Local	Hora	PIN	Aprovada
1	Baixo	Internacional	Não Comercial	Verdadeiro	Não
2	Baixo	Internacional	Não Comercial	Falso	Não
3	Médio	Internacional	Não Comercial	Verdadeiro	Sim
4	Alto	Mesmo	Não Comercial	Verdadeiro	Sim
5	Alto	BR	Comercial	Verdadeiro	Sim
6	Alto	BR	Comercial	Falso	Não
7	Médio	BR	Comercial	Falso	Sim
8	Baixo	Mesmo	Não Comercial	Verdadeiro	Não
9	Baixo	BR	Comercial	Verdadeiro	Sim
10	Alto	Mesmo	Comercial	Verdadeiro	Sim
11	Baixo	Mesmo	Comercial	Falso	Sim
12	Médio	Mesmo	Não Comercial	Falso	Sim
13	Médio	Internacional	Comercial	Verdadeiro	Sim
14	Alto	Mesmo	Não Comercial	Falso	Não

# Árvore de Decisão

Compra	Valor	Local	Hora	PIN	Aprovada
1	Baixo	Internacional	Não Comercial	Verdadeiro	Não
2	Baixo	Internacional	Não Comercial	Falso	Não
3	Médio	Internacional	Não Comercial	Verdadeiro	Sim
4	Alto	Mesmo	Não Comercial	Verdadeiro	Sim
5	Alto	BR	Comercial	Verdadeiro	Sim
6	Alto	BR	Comercial	Falso	Não
7	Médio	BR	Comercial	Falso	Sim
8	Baixo	Mesmo	Não Comercial	Verdadeiro	Não
9	Baixo	BR	Comercial	Verdadeiro	Sim
10	Alto	Mesmo	Comercial	Verdadeiro	Sim
11	Baixo	Mesmo	Comercial	Falso	Sim
12	Médio	Mesmo	Não Comercial	Falso	Sim
13	Médio	Internacional	Comercial	Verdadeiro	Sim
14	Alto	Mesmo	Não Comercial	Falso	Não

`values(PIN) = verdadeiro, Falso`

$$S = [9+, 5-]$$

# Árvore de Decisão

Compra	Valor	Local	Hora	PIN	Aprovada
1	Baixo	Internacional	Não Comercial	Verdadeiro	Não
2	Baixo	Internacional	Não Comercial	Falso	Não
3	Médio	Internacional	Não Comercial	Verdadeiro	Sim
4	Alto	Mesmo	Não Comercial	Verdadeiro	Sim
5	Alto	BR	Comercial	Verdadeiro	Sim
6	Alto	BR	Comercial	Falso	Não
7	Médio	BR	Comercial	Falso	Sim
8	Baixo	Mesmo	Não Comercial	Verdadeiro	Não
9	Baixo	BR	Comercial	Verdadeiro	Sim
10	Alto	Mesmo	Comercial	Verdadeiro	Sim
11	Baixo	Mesmo	Comercial	Falso	Sim
12	Médio	Mesmo	Não Comercial	Falso	Sim
13	Médio	Internacional	Comercial	Verdadeiro	Sim
14	Alto	Mesmo	Não Comercial	Falso	Não

$\text{values(PIN)} = \text{verdadeiro, Falso}$

$$S = [9+, 5-]$$

$$S_{\text{verdadeiro}} = [6+, 2-]$$

# Árvore de Decisão

Compra	Valor	Local	Hora	PIN	Aprovada
1	Baixo	Internacional	Não Comercial	Verdadeiro	Não
2	Baixo	Internacional	Não Comercial	Falso	Não
3	Médio	Internacional	Não Comercial	Verdadeiro	Sim
4	Alto	Mesmo	Não Comercial	Verdadeiro	Sim
5	Alto	BR	Comercial	Verdadeiro	Sim
6	Alto	BR	Comercial	Falso	Não
7	Médio	BR	Comercial	Falso	Sim
8	Baixo	Mesmo	Não Comercial	Verdadeiro	Não
9	Baixo	BR	Comercial	Verdadeiro	Sim
10	Alto	Mesmo	Comercial	Verdadeiro	Sim
11	Baixo	Mesmo	Comercial	Falso	Sim
12	Médio	Mesmo	Não Comercial	Falso	Sim
13	Médio	Internacional	Comercial	Verdadeiro	Sim
14	Alto	Mesmo	Não Comercial	Falso	Não

$\text{values(PIN)} = \text{verdadeiro, Falso}$

$$S = [9+, 5-]$$

$$S_{\text{verdadeiro}} = [6+, 2-]$$

$$S_{\text{Falso}} = [3+, 3-]$$

# Árvore de Decisão

Compra	Valor	Local	Hora	PIN	Aprovada
1	Baixo	Internacional	Não Comercial	Verdadeiro	Não
2	Baixo	Internacional	Não Comercial	Falso	Não
3	Médio	Internacional	Não Comercial	Verdadeiro	Sim
4	Alto	Mesmo	Não Comercial	Verdadeiro	Sim
5	Alto	BR	Comercial	Verdadeiro	Sim
6	Alto	BR	Comercial	Falso	Não
7	Médio	BR	Comercial	Falso	Sim
8	Baixo	Mesmo	Não Comercial	Verdadeiro	Não
9	Baixo	BR	Comercial	Verdadeiro	Sim
10	Alto	Mesmo	Comercial	Verdadeiro	Sim
11	Baixo	Mesmo	Comercial	Falso	Sim
12	Médio	Mesmo	Não Comercial	Falso	Sim
13	Médio	Internacional	Comercial	Verdadeiro	Sim
14	Alto	Mesmo	Não Comercial	Falso	Não

$\text{values(PIN)} = \text{verdadeiro, Falso}$

$$S = [9+, 5-]$$

$$S_{\text{verdadeiro}} = [6+, 2-]$$

$$S_{\text{Falso}} = [3+, 3-]$$

$$\text{Gain}(S, \text{PIN}) = \text{Entropia}(S) - \sum_F^V \left( \frac{Sv}{S} \right) \text{Entropia}(Sv)$$

$$= \text{Entropia}(S) - (8/14)\text{Entropia}(S_{\text{verdadeiro}}) - (6/14)\text{Entropia}(S_{\text{Falso}})$$

# Árvore de Decisão

Compra	Valor	Local	Hora	PIN	Aprovada
1	Baixo	Internacional	Não Comercial	Verdadeiro	Não
2	Baixo	Internacional	Não Comercial	Falso	Não
3	Médio	Internacional	Não Comercial	Verdadeiro	Sim
4	Alto	Mesmo	Não Comercial	Verdadeiro	Sim
5	Alto	BR	Comercial	Verdadeiro	Sim
6	Alto	BR	Comercial	Falso	Não
7	Médio	BR	Comercial	Falso	Sim
8	Baixo	Mesmo	Não Comercial	Verdadeiro	Não
9	Baixo	BR	Comercial	Verdadeiro	Sim
10	Alto	Mesmo	Comercial	Verdadeiro	Sim
11	Baixo	Mesmo	Comercial	Falso	Sim
12	Médio	Mesmo	Não Comercial	Falso	Sim
13	Médio	Internacional	Comercial	Verdadeiro	Sim
14	Alto	Mesmo	Não Comercial	Falso	Não

$\text{values(PIN)} = \text{verdadeiro, Falso}$

$$S = [9+, 5-]$$

$$S_{\text{verdadeiro}} = [6+, 2-]$$

$$S_{\text{Falso}} = [3+, 3-]$$

$$\text{Gain}(S, \text{PIN}) = \text{Entropia}(S) - \sum_F^V \left( \frac{Sv}{S} \right) \text{Entropia}(Sv)$$

$$= \text{Entropia}(S) - (8/14)\text{Entropia}(S_{\text{verdadeiro}}) - (6/14)\text{Entropia}(S_{\text{Falso}})$$

$$\text{Entropia}(S) = - (9 \log_2 9) - (5 \log_2 5)$$

$$\text{Entropia}(S_{\text{verdadeiro}}) = - (6 \log_2 6) - (2 \log_2 2)$$

$$\text{Entropia}(S_{\text{Falso}}) = - (3 \log_2 3) - (3 \log_2 3)$$

# Árvore de Decisão

Compra	Valor	Local	Hora	PIN	Aprovada
1	Baixo	Internacional	Não Comercial	Verdadeiro	Não
2	Baixo	Internacional	Não Comercial	Falso	Não
3	Médio	Internacional	Não Comercial	Verdadeiro	Sim
4	Alto	Mesmo	Não Comercial	Verdadeiro	Sim
5	Alto	BR	Comercial	Verdadeiro	Sim
6	Alto	BR	Comercial	Falso	Não
7	Médio	BR	Comercial	Falso	Sim
8	Baixo	Mesmo	Não Comercial	Verdadeiro	Não
9	Baixo	BR	Comercial	Verdadeiro	Sim
10	Alto	Mesmo	Comercial	Verdadeiro	Sim
11	Baixo	Mesmo	Comercial	Falso	Sim
12	Médio	Mesmo	Não Comercial	Falso	Sim
13	Médio	Internacional	Comercial	Verdadeiro	Sim
14	Alto	Mesmo	Não Comercial	Falso	Não

$\text{values(PIN)} = \text{verdadeiro, Falso}$

$$S = [9+, 5-]$$

$$S_{\text{verdadeiro}} = [6+, 2-]$$

$$S_{\text{Falso}} = [3+, 3-]$$

$$\text{Gain}(S, \text{PIN}) = \text{Entropia}(S) - \sum_F^V \left( \frac{Sv}{S} \right) \text{Entropia}(Sv)$$

$$= \text{Entropia}(S) - (8/14)\text{Entropia}(S_{\text{verdadeiro}}) - (6/14)\text{Entropia}(S_{\text{Falso}})$$

$$\text{Entropia}(S) = - (9 \log_2 9) - (5 \log_2 5)$$

$$\text{Entropia}(S_{\text{verdadeiro}}) = - (6 \log_2 6) - (2 \log_2 2)$$

$$\text{Entropia}(S_{\text{Falso}}) = - (3 \log_2 3) - (3 \log_2 3)$$

$$= 0.048$$

# Árvore de Decisão

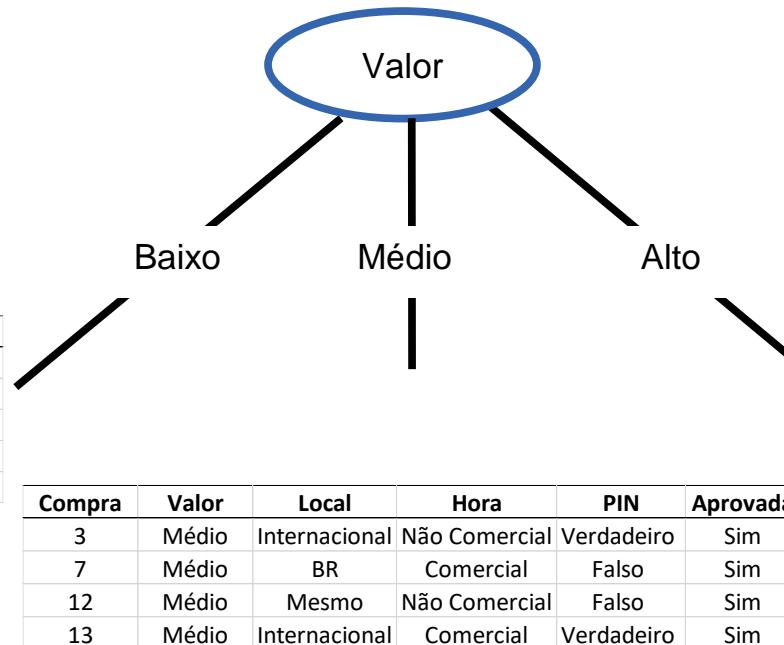
$\text{Gain}(S, \text{ valor}) = 0.246$

$\text{Gain}(S, \text{ Hora}) = 0.151$

$\text{Gain}(S, \text{ PIN}) = 0.048$

$\text{Gain}(S, \text{ Local}) = 0.029$

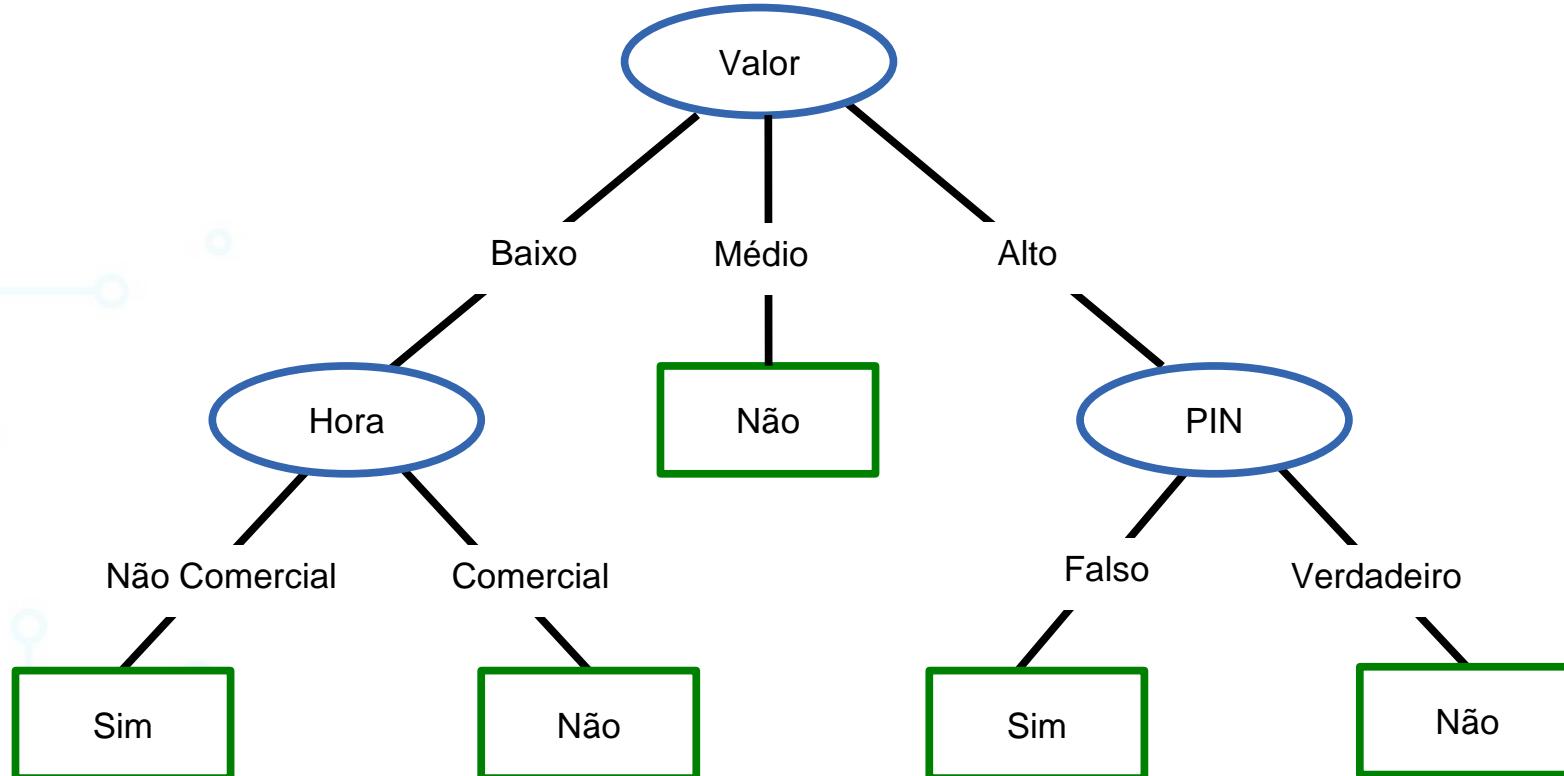
# Árvore de Decisão



Compra	Valor	Local	Hora	PIN	Aprovada
1	Baixo	Internacional	Não Comercial	Verdadeiro	Não
2	Baixo	Internacional	Não Comercial	Falso	Não
8	Baixo	Mesmo	Não Comercial	Verdadeiro	Não
9	Baixo	BR	Comercial	Verdadeiro	Sim
11	Baixo	Mesmo	Comercial	Falso	Sim

Compra	Valor	Local	Hora	PIN	Aprovada
3	Médio	Internacional	Não Comercial	Verdadeiro	Sim
7	Médio	BR	Comercial	Falso	Sim
12	Médio	Mesmo	Não Comercial	Falso	Sim
13	Médio	Internacional	Comercial	Verdadeiro	Sim

# Árvore de Decisão

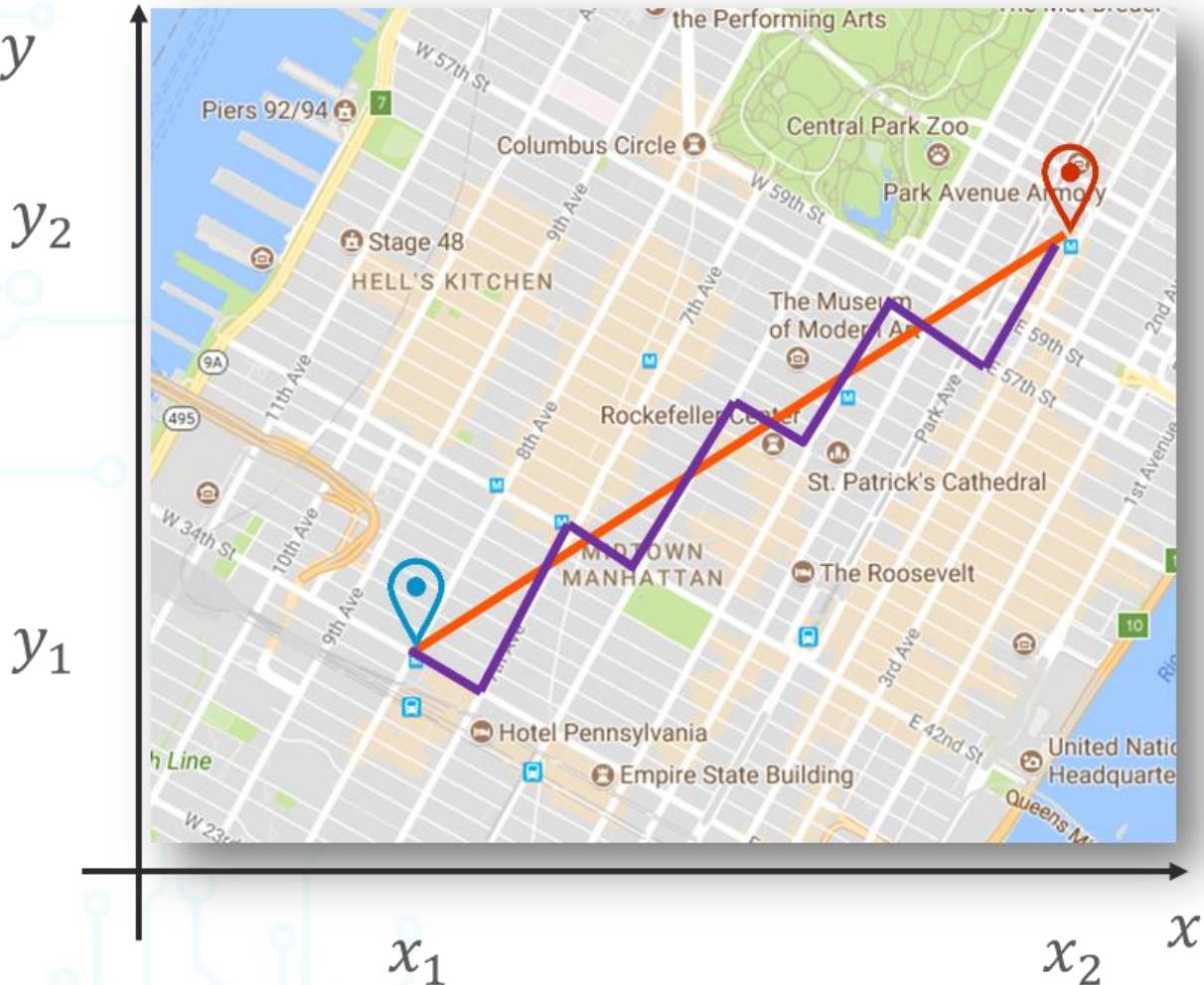


“

# K-Nearest Neighbors (KNN)



# Medidas de Distância



**Manhattan**

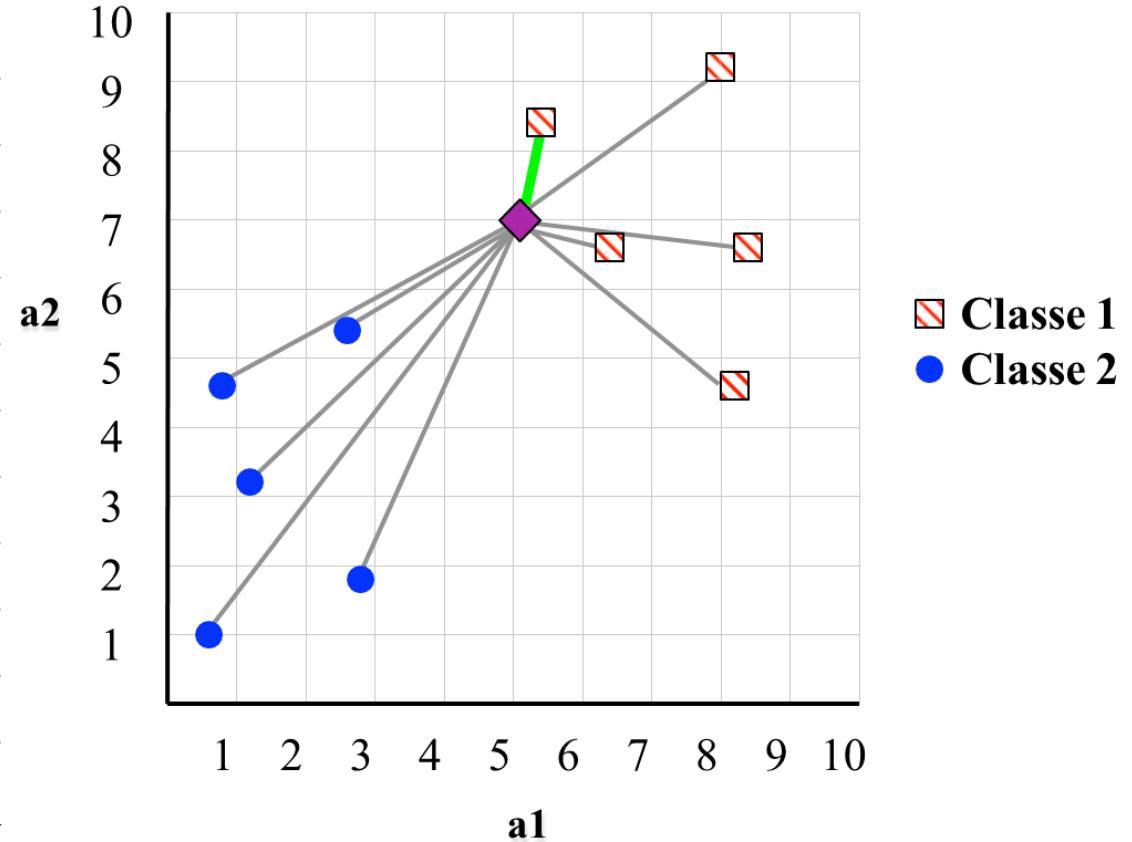
$$d_1 = |x_1 - x_2| + |y_1 - y_2|$$

**Euclidiana**

$$d_2 = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?



Fonte: Keogh, E. AGentle Introduction to Machine Learning and Data Mining for the Database Community, SBBD 2003, Manaus.

# KNN

- *K-Nearest Neighbors (KNN)*;
- Exemplos correspondem a pontos no espaço n-dimensional ( $\mathbb{R}^n$ );
- Em geral os vizinhos são definidos em função de uma medida de distância. Por exemplo, considerando-se dois vetores  $\mathbf{x}=[x_1, x_2, \dots, x_n]$  e  $\mathbf{y}=[y_1, y_2, \dots, y_n]$ , a distância Euclidiana entre estes dois vetores é dada por:

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#1, \#11) = \sqrt{(0.5 - 5)^2 + (1 - 7)^2}$$

$$d_E(\#1, \#11) = \sqrt{(-4.5)^2 + (-6)^2}$$

$$d_E(\#1, \#11) = \sqrt{20.25 + 36}$$

$$d_E(\#1, \#11) = \sqrt{56.25}$$

$$d_E(\#1, \#11) = 7.5$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#2, \#11) = \sqrt{(2.9 - 5)^2 + (1.9 - 7)^2}$$

$$d_E(\#2, \#11) = \sqrt{(-2.1)^2 + (-5.1)^2}$$

$$d_E(\#2, \#11) = \sqrt{4.41 + 26.1}$$

$$d_E(\#2, \#11) = \sqrt{30.51}$$

$$d_E(\#2, \#11) = 5.5$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#3, \#11) = 5.4$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#4, \#11) = 4.7$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#5, \#11) = 2.8$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#6, \#11) = 3.8$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#7, \#11) = 3.3$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#8, \#11) = 1.3$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#9, \#11) = 3.6$$

# KNN

#	a1	a2	Classe
1	0.5	1	2
2	2.9	1.9	2
3	1.2	3.1	2
4	0.8	4.7	2
5	2.7	5.4	2
6	8.1	4.7	1
7	8.3	6.6	1
8	6.3	6.7	1
9	8	9.1	1
10	5.4	8.4	1
11	5	7	?

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

$$d_E(\#10, \#11) = 1.4$$

# KNN

#	a1	a2	Classe	Distância para a instância #11
1	0.5	1	2	7.5
2	2.9	1.9	2	5.5
3	1.2	3.1	2	5.4
4	0.8	4.7	2	4.7
5	2.7	5.4	2	2.8
6	8.1	4.7	1	3.8
7	8.3	6.6	1	3.3
8	6.3	6.7	1	1.3
9	8	9.1	1	3.6
10	5.4	8.4	1	1.4
11	5	7	?	

# KNN

Compra	Valor	Local	Hora	PIN	Aprovada
1	Baixo	Internacional	Não Comercial	Verdadeiro	Não
2	Baixo	Internacional	Não Comercial	Falso	Não
3	Médio	Internacional	Não Comercial	Verdadeiro	Sim
4	Alto	Mesmo	Não Comercial	Verdadeiro	Sim
5	Alto	BR	Comercial	Verdadeiro	Sim
6	Alto	BR	Comercial	Falso	Não
7	Médio	BR	Comercial	Falso	Sim
8	Baixo	Mesmo	Não Comercial	Verdadeiro	Não
9	Baixo	BR	Comercial	Verdadeiro	Sim
10	Alto	Mesmo	Comercial	Verdadeiro	Sim
11	Baixo	Mesmo	Comercial	Falso	Sim
12	Médio	Mesmo	Não Comercial	Falso	Sim
13	Médio	Internacional	Comercial	Verdadeiro	Sim
14	Alto	Mesmo	Não Comercial	Falso	Não

# Naive Bayes

“GAUSSIAN NAIVE BAYES CLASSIFIER”

“Gaussian” because this is a normal distribution

This is our prior belief

$$P(\text{class} | \text{data}) = \frac{P(\text{data} | \text{class}) \times p(\text{class})}{P(\text{data})}$$

We don't calculate this in naive bayes classifiers

Chris Albon

# Naive Bayes

- O teorema de Bayes permite o cálculo da posteriori de  $h$
- MAP (Maximum a Posteriori) Hipótese Máxima a Posteriori
  - $P(h|D) = (P(D|h) P(h)) / P(D)$

# Naive Bayes

- O NB, simplifica o cálculo de  $P(a_1, a_2, \dots, a_n | v_j)$  assumindo que todos os atributos são condicionalmente independentes entre si dado o valor da classe.
- Independência dos atributos condicionada à classe:

$$P(a_1, a_2 \dots a_n | v_j) = \prod_i P(a_i | v_j)$$

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11)

$$P(\text{Classe}=0|A_1=0, A_2=1, A_3=0, A_4=1) = \\ P(A_1=0, A_2=1, A_3=0, A_4=1|\text{Classe}=0) * P(\text{Classe}=0)$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(A \cap B) = \frac{P(B|A)P(A)}{P(B)}$$

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11)

$$P(\text{Classe}=0|A_1=0, A_2=1, A_3=0, A_4=1) =$$

$$P(A_1=0, A_2=1, A_3=0, A_4=1|\text{Classe}=0) * P(\text{Classe}=0)$$

Como assumimos a independência condicional, temos que:

$$P(A_1=0, A_2=1, A_3=0, A_4=1|\text{Classe}=0) = \prod P(A_i=x_i | \text{Classe}=0)$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$P(A \cap B) = \frac{P(B|A)P(A)}{P(B)}$$

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11)

$$P(\text{Classe}=0|A_1=0, A_2=1, A_3=0, A_4=1) =$$

$$P(A_1=0, A_2=1, A_3=0, A_4=1|\text{Classe}=0) * P(\text{Classe}=0)$$

Como assumimos a independência condicional, temos que:

$$P(A_1=0, A_2=1, A_3=0, A_4=1|\text{Classe}=0) = \prod P(A_i=x_i | \text{Classe}=0)$$

Assim, temos:

$$P(\text{Classe}=0|A_1=0, A_2=1, A_3=0, A_4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0)$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$$P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A_1=0 | C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe} = 0)$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A_1=0 | C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe} = 0)$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) = & \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = & \\
 (P(A_1=0 | C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe} = 0) & \\
 (1/5) *
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) \\
 (1/5) *
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) =$$

$$(\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) =$$

$$(P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0)$$

$$(1/5) *$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) \\
 (1/5) * (2/5)
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) \\
 (1/5) * (2/5)
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) \\
 (1/5) * (2/5)
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) &= \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) &= \\
 (P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * \textcolor{blue}{P(A4=1|C=0)}) * P(\text{Classe}=0) & \\
 (1/5) * (2/5) * (5/5) * &
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) \\
 (1/5) * (2/5) * (5/5) *
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) = & \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = & \\
 (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) & \\
 (1/5) * (2/5) * (5/5) * (1/5) &
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado , deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) \\
 (1/5) * (2/5) * (5/5) * (1/5) * (5/10)
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

$$\begin{aligned}
 P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe}=0) \\
 (1/5) * (2/5) * (5/5) * (1/5) * (5/10) =
 \end{aligned}$$

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) &= \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) &= \\
 (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) & \\
 (1/5) * (2/5) * (5/5) * (1/5) * (5/10) &= \\
 0.2 * 0.4 * 1 * 0.2 * 0.5 & = 0.008
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) \\
 (1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

Agora fazemos o mesmo processo para a classe=1:

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 & P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = \\
 & (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 & (P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe}=0) \\
 & (1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 & 0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 & P(\text{Classe}=1 | A1=0, A2=1, A3=0, A4=1) = \\
 & (\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1) =
 \end{aligned}$$

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) &= \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) &= \\
 (P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe}=0) & \\
 (1/5) * (2/5) * (5/5) * (1/5) * (5/10) &= \\
 0.2 * 0.4 * 1 * 0.2 * 0.5 &= 0.008
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 P(\text{Classe}=1 | A1=0, A2=1, A3=0, A4=1) &= \\
 (\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1) &= \\
 (P(A1=0|C=1) * P(A2=1|C=1) * P(A3=0|C=1) * P(A4=1|C=1)) * P(\text{Classe}=1) &
 \end{aligned}$$

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A_1=0, A_2=1, A_3=0, A_4=1) &= (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 &= (P(A_1=0|C=0) * P(A_2=1|C=0) * P(A_3=0|C=0) * P(A_4=1|C=0)) * P(\text{Classe}=0) \\
 &= (1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 &= 0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 P(\text{Classe}=1 | A_1=0, A_2=1, A_3=0, A_4=1) &= (\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1) = \\
 &= (P(A_1=0|C=1) * P(A_2=1|C=1) * P(A_3=0|C=1) * P(A_4=1|C=1)) * P(\text{Classe}=1) \\
 &= (5/5) * (5/5) * (0/5) * (5/5) * (5/10) =
 \end{aligned}$$

# Naive Bayes

Considere por exemplo, que a partir do conjunto de instâncias de treinamento dado, deseja-se classificar uma nova instância (11):

$$\begin{aligned}
 P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe}=0) \\
 (1/5) * (2/5) * (5/5) * (1/5) * (5/10) = \\
 0.2 * 0.4 * 1 * 0.2 * 0.5 = 0.008
 \end{aligned}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 P(\text{Classe}=1 | A1=0, A2=1, A3=0, A4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1) = \\
 (P(A1=0|C=1) * P(A2=1|C=1) * P(A3=0|C=1) * P(A4=1|C=1)) * P(\text{Classe}=1) \\
 (5/5) * (5/5) * (0/5) * (5/5) * (5/10) = \\
 1 * 1 * 0 * 1 * 0.5 = 0
 \end{aligned}$$

# Naive Bayes

## Solução

$$\begin{aligned}
 P(\text{Classe}=0 | A1=0, A2=1, A3=0, A4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=0)) * P(\text{Classe}=0) = \\
 (P(A1=0|C=0) * P(A2=1|C=0) * P(A3=0|C=0) * P(A4=1|C=0)) * P(\text{Classe}=0) \\
 (1+0.5)/(5+1) * (2+0.5)/(5+1) * (5+0.5)/(5+1) * (1+0.5)/(5+1) * \\
 (5/10) = 0.25 * 0.41 * 0.91 * 0.25 * 0.5 = 0.011
 \end{aligned}$$

Agora fazemos o mesmo processo para a classe=1:

$$\begin{aligned}
 P(\text{Classe}=1 | A1=0, A2=1, A3=0, A4=1) = \\
 (\prod P(A_i=x_i | \text{Classe}=1)) * P(\text{Classe}=1) = \\
 (P(A1=0|C=1) * P(A2=1|C=1) * P(A3=0|C=1) * P(A4=1|C=1)) * P(\text{Classe}=1) \\
 (5+0.5)/(5+1) * (5+0.5)/(5+1) * (0+0.5)/(5+1) * (5+0.5)/(5+1) * \\
 (5/10) = 0.91 * 0.91 * 0.09 * 0.91 * 0.5 = 0.033
 \end{aligned}$$

# Naive Bayes

## Solução

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = 0.011$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \mathbf{0.033}$$

Instâncias	Classe	A1	A2	A3	A4
1	1	0	1	1	1
2	1	0	1	1	1
3	0	1	0	0	0
4	1	0	1	1	1
5	1	0	1	1	1
6	0	0	0	0	1
7	0	1	1	0	0
8	1	0	1	1	1
9	0	1	0	0	0
10	0	1	1	0	0
11	?	0	1	0	1

Normalizando, temos:

$$P(\text{Classe}=0|A1=0, A2=1, A3=0, A4=1) = 0.011/(0.011+\mathbf{0.033}) = 0.25$$

$$P(\text{Classe}=1|A1=0, A2=1, A3=0, A4=1) = \mathbf{0.033}/(0.011+\mathbf{0.033}) = 0.75$$

A classe mais provável é a classe 1.

# Thanks !



Vinicius Fernandes Caridá

[vfcarida@gmail.com](mailto:vfcarida@gmail.com)



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida



@vinicius caridá



@vfcarida