# Reconhecimento de Padrões (RPD0041): Introdução, Regras e Datas

André E. Lazzaretti

UTFPR/CPGEI

# Ementa

- Introdução à análise de padrões
- Modelos Lineares Generativos (Métodos estatísticos)
- Métodos de Otimização
- Modelos Lineares Discriminativos
- Modelos Não-Lineares
- Aprendizagem não-supervisionada
- Detecção de Novidades
- Extração de Características e Redução de Dimensionalidade
- Redes Neurais, Deep Learning e Convolutional Neural Networks
- Análise e Comparação de Desempenho
- Modelos de Markov (*Hidden Markov Models*) (*)

# Calendário

- Semana 1 (hoje): Introdução à análise de padrões
- Semana 2: Métodos estatísticos parte 1
- Semana 3: Métodos estatísticos parte 2
- Semana 4: Métodos de otimização
- Semana 5: Modelos lineares
- Semana 6: Modelos não-lineares
- Semana 7: Análise e Comparação de Desempenho
- Semana 8: Aprendizagem não-supervisionada
- Semana 9: Detecção de Novidades
- Semana 10: Extração de Características e Redução de Dimensionalidade
- Semana 11: Redes Neurais e Deep Learning
- Semana 12: Modelos de Markov, apresentação das propostas e entrega dos exercícios

# Pré-requisitos

- Álgebra Linear;
- Probabilidade e Estatística;
- Cálculo Diferencial e Integral;
- Programação (Matlab, Python).

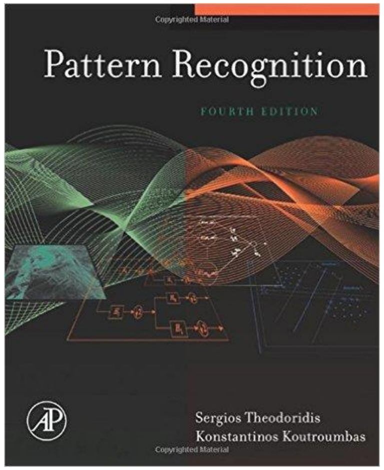
# Aulas

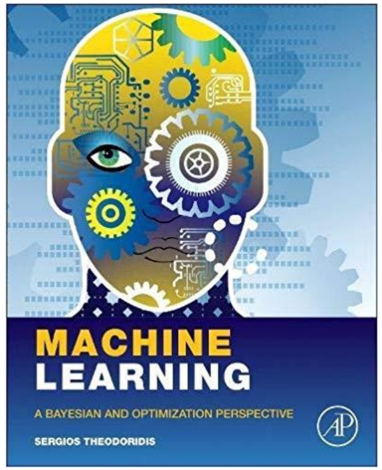
- Uso de slides para apresentação dos conceitos;
- Detalhamento de equações/deduções e resoluções de exercícios – Quadro ou com auxílio PC;
- Matlab/Octave!
- **Listas de exercícios para entrega** para facilitar compreensão – mais práticas, porém são colocados alguns pontos para deduções e soluções analíticas;

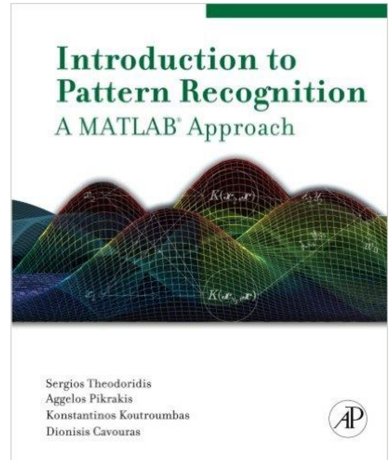
# Avaliação

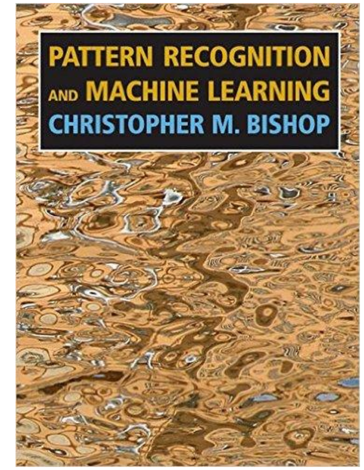
- Entrega dos Exercícios propostos.
- Trabalho Final:
  – Reproduzir integralmente ou parcialmente um artigo;
  – Propor uma técnica de RP para um determinado problema;
  – Procurar algo relacionado com sua dissertação/tese.
  – Regras do Trabalho Final:
    - Individual;
    - Apresentar a proposta antes do final do trimestre para validação (por escrito);
    - Trabalho escrito no formato de artigo (Padrão IEEE – Conferência – mín. de quatro páginas) + apresentação a ser agendada durante o trimestre;
- Nota Final:
  – 0,3*Notas dos Exercícios + 0,1*Apresentação da proposta + 0,3*Artigo do Trabalho Final + 0,3*Apresentação do Trabalho Final

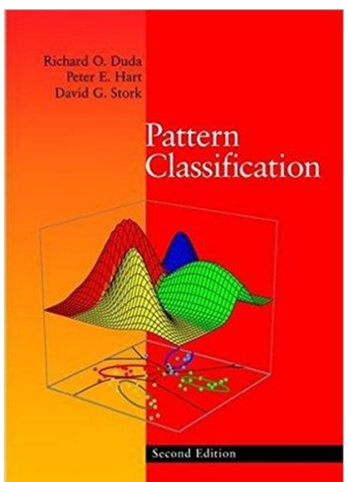
# Bibliografia

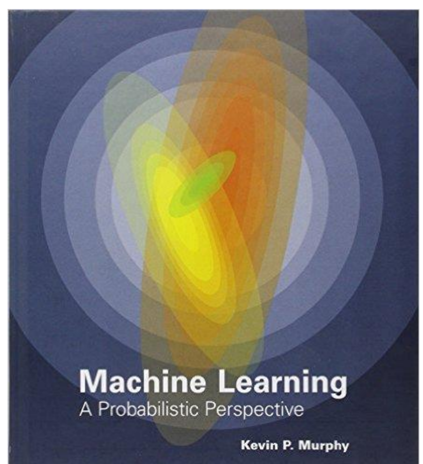


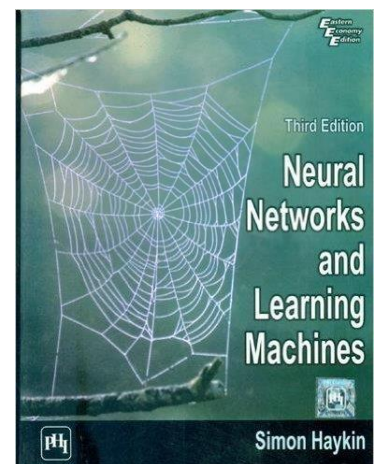



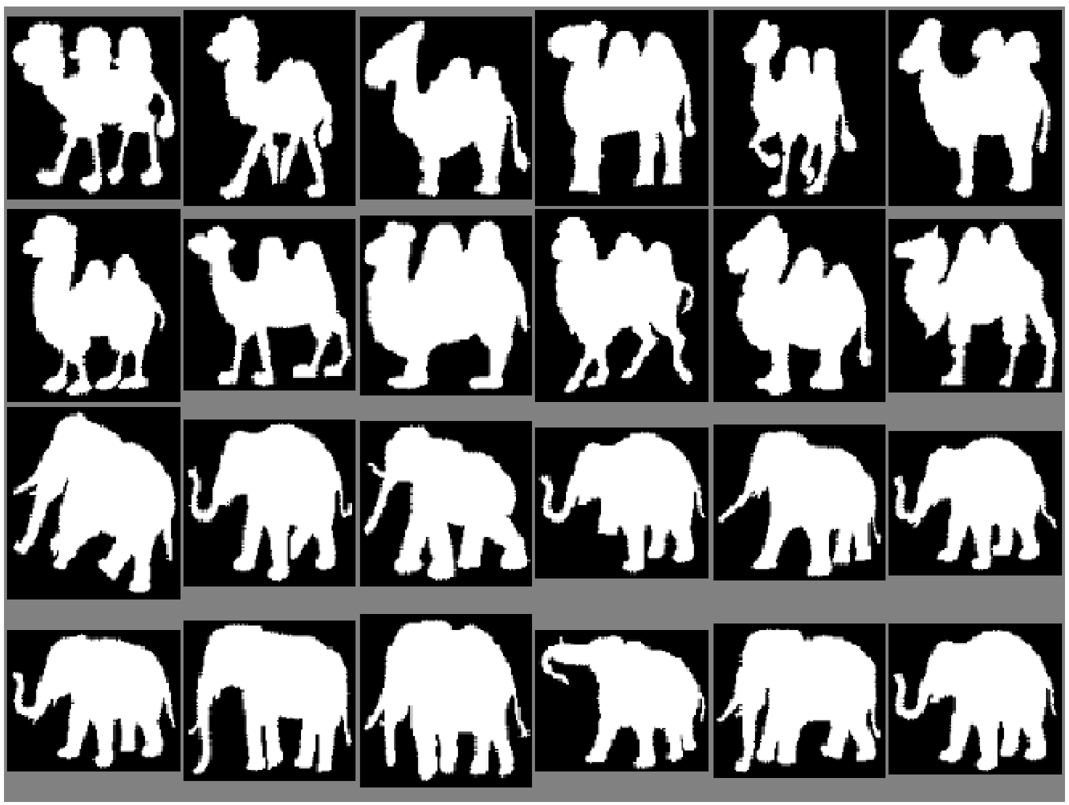



# Aulas de Outras Universidades

- Andrew Ng (Stanford):
  - http://www.andrewng.org/courses/
  - https://www.youtube.com/playlist?list=PLJ_CMbwA6bT-n1W0mgOlYwccZ-j6gBXqE
- Nando de Freitas (UBC-Oxford):
  - http://www.cs.ubc.ca/~nando/540-2013/lectures.html

# Introdução

- Exemplo: diferenciar camelos de elefantes



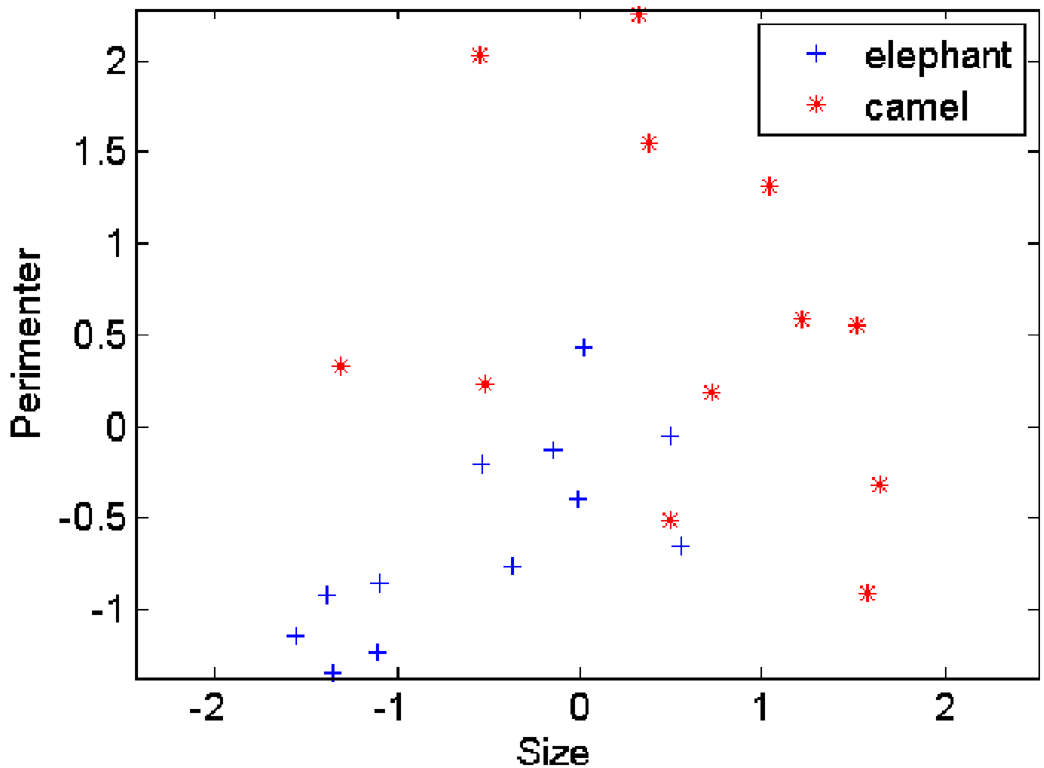
KIMIA dataset

Para alguém que nunca viu esses animais antes (p.ex. criança de três-quatro anos). Ele(a) seria capaz de agrupar as figuras similares?

Se fossem apresentadas novas figuras e fossem dados nomes às classes, ele(a) conseguiria classificá-las?

# Introdução

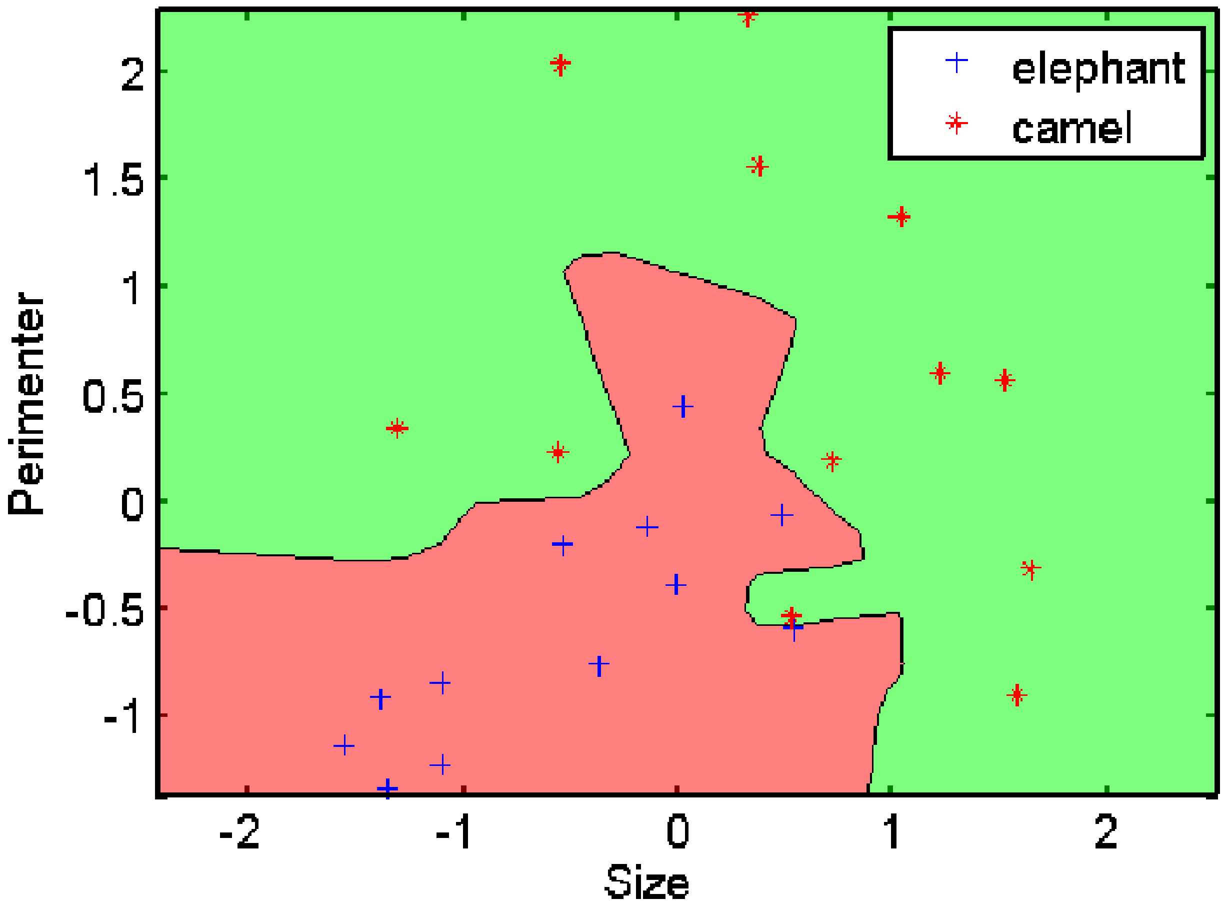
- Como poderíamos fazer o mesmo com um computador?
- Abordagem numérica → calcular similaridades;
- Area e Perímetro dos animais:



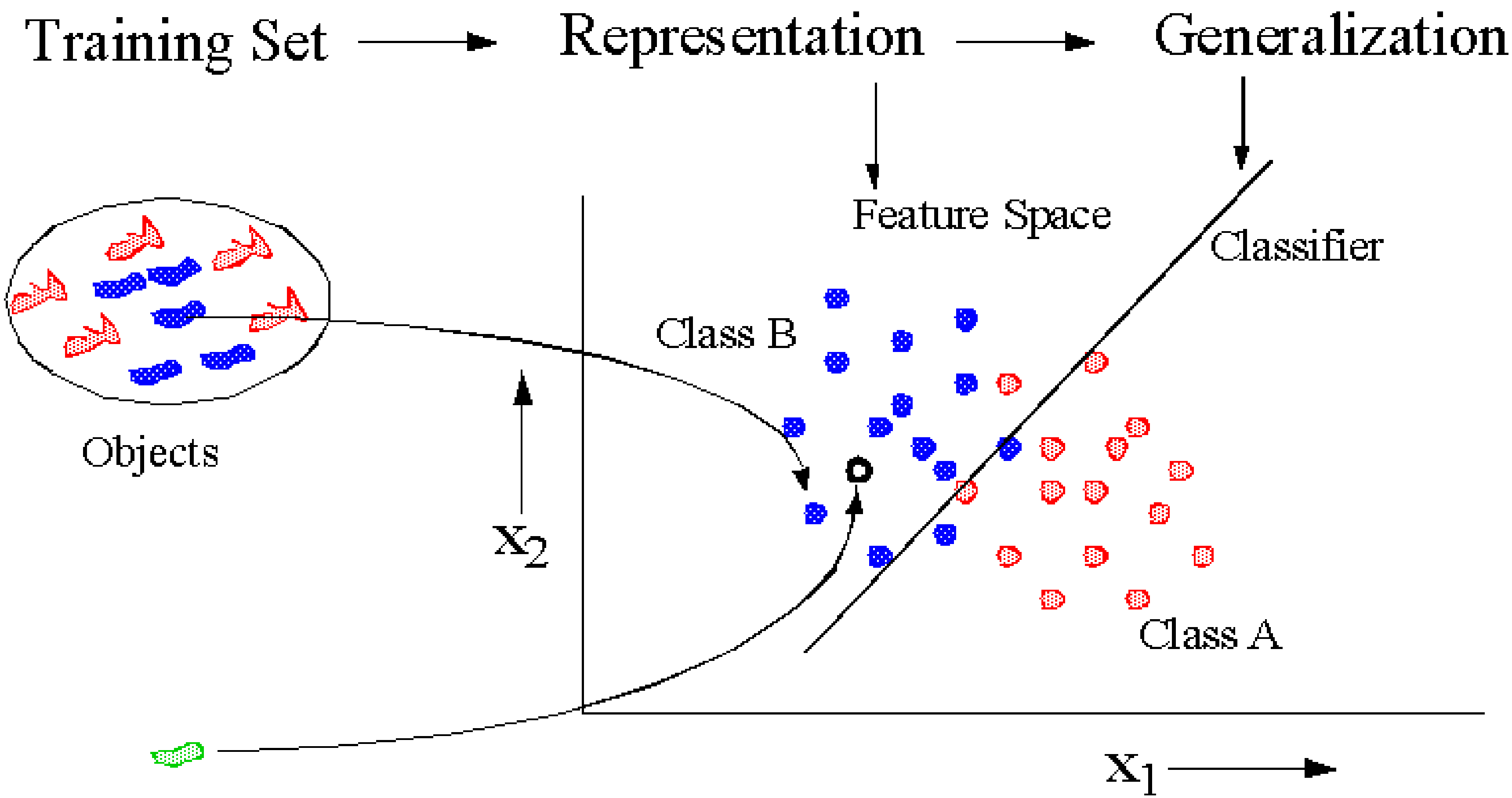

**Espaço de Características (Feature Space)**

# Introdução

- Com isso, pode-se criar um classificador (processo de aprendizagem por exemplos):

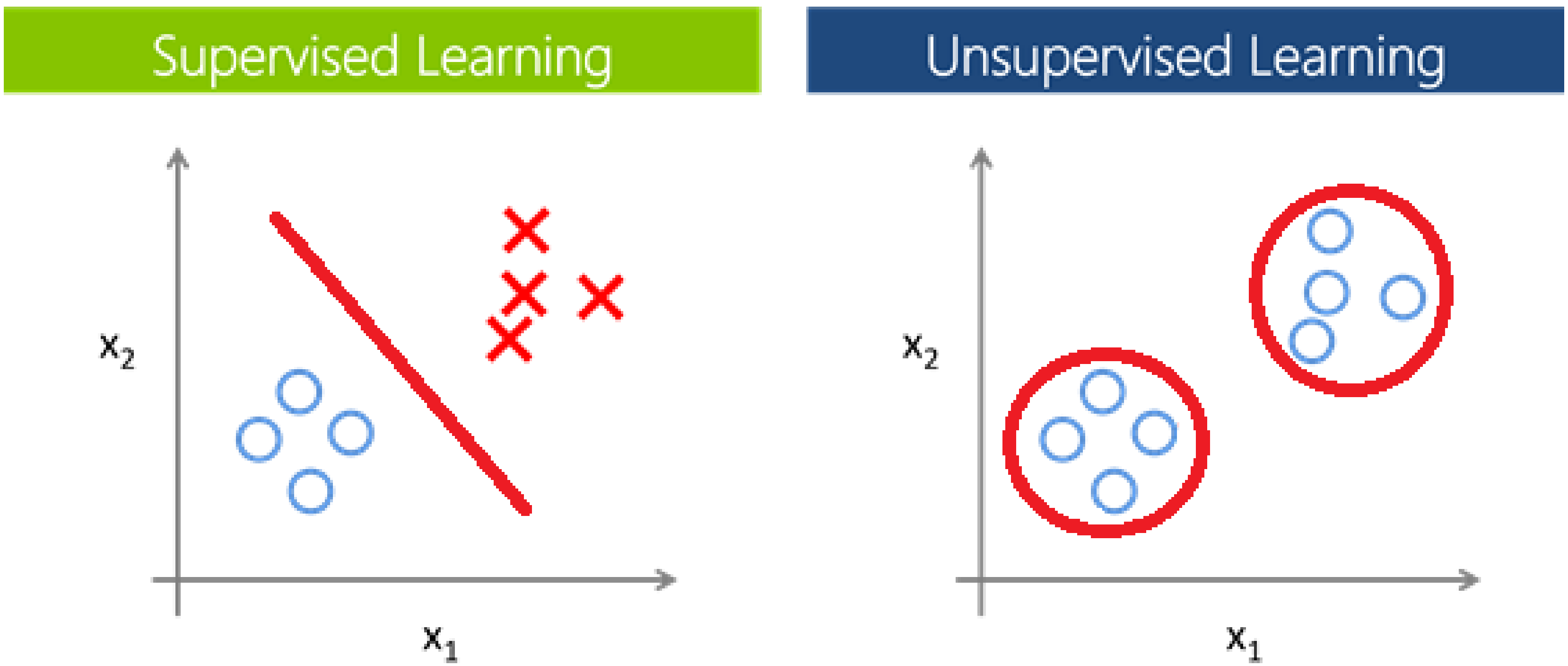


# Visão Geral

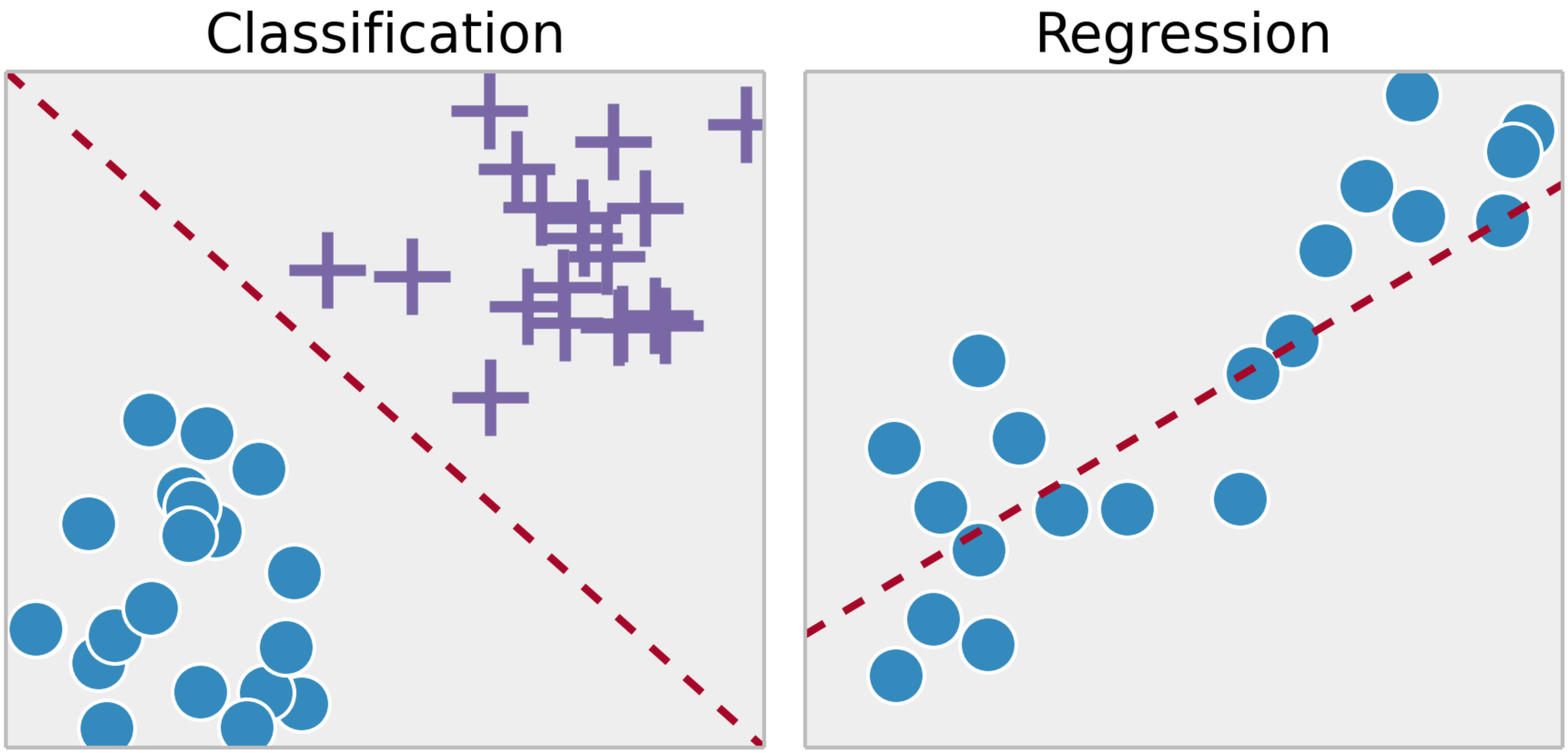


Training Set
Representation
Generalization
Feature Space
Classifier
Class B
Class A
Objects
$x_2$
$x_1$

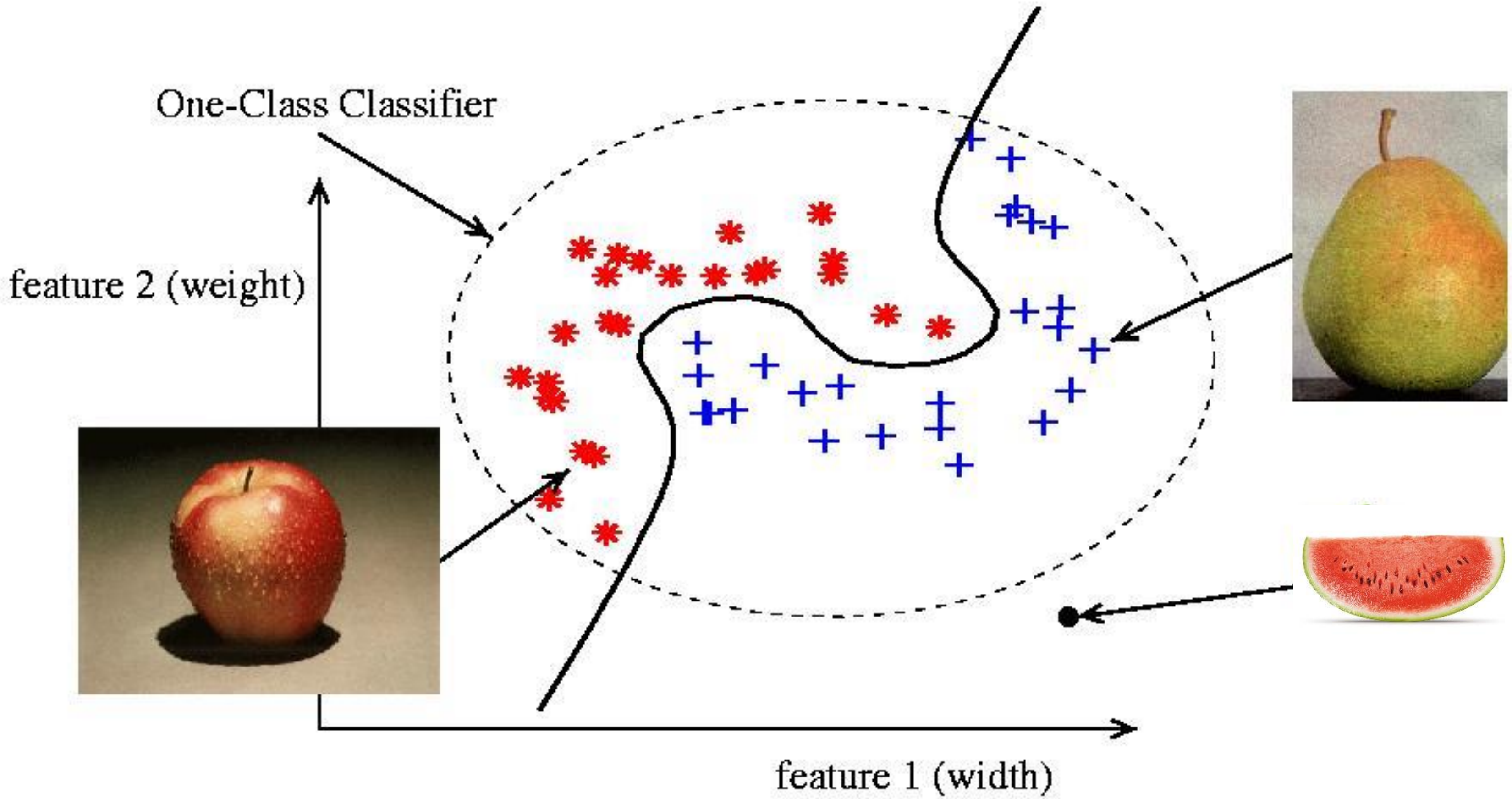
# Aprendizado Supervisionado x Não-Supervisionado



# Regressão

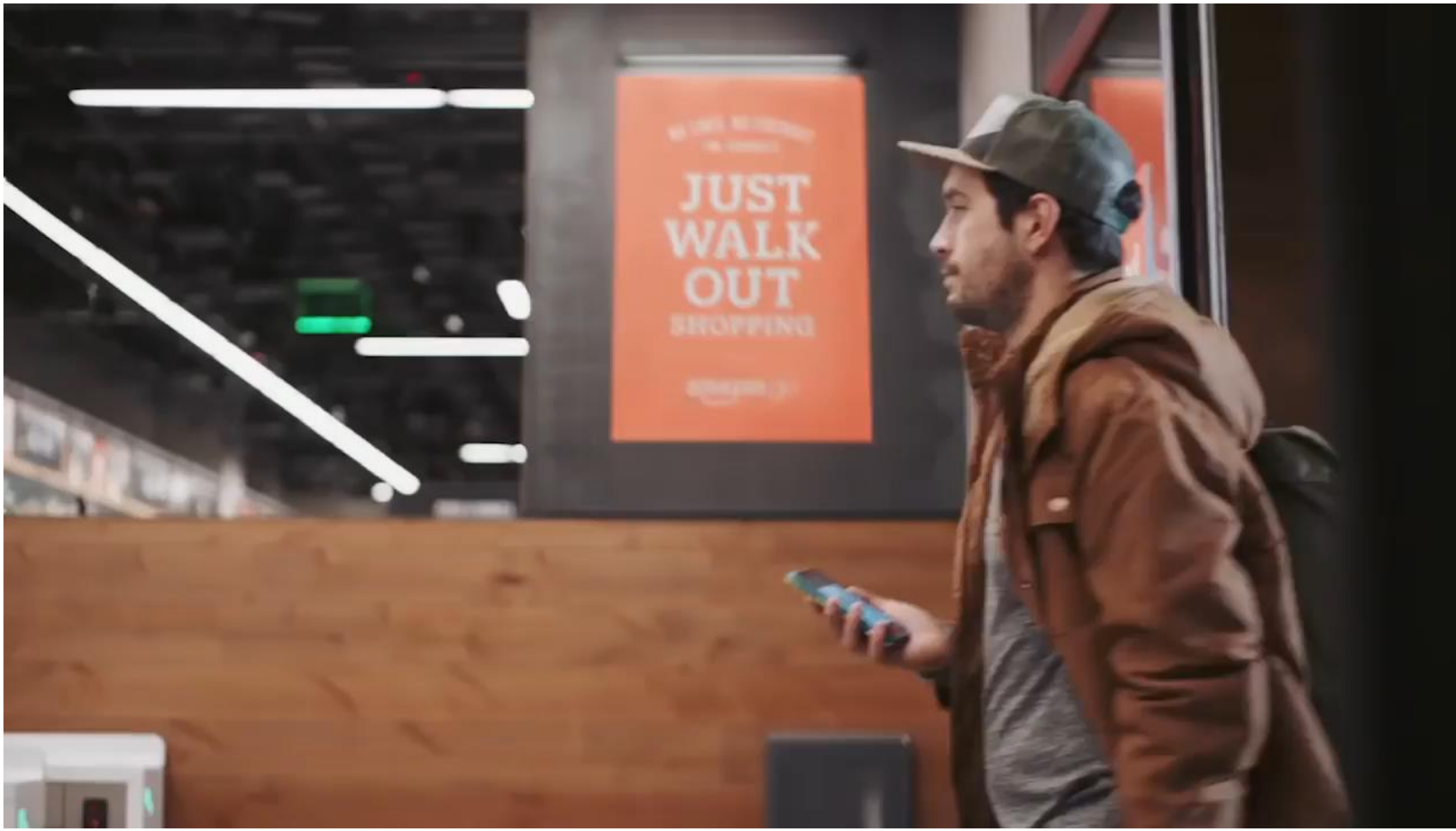


# Detecção de Novidades (one-class)



# Tecnologias já Disponíveis

# Amazon GO



# Conceitos Iniciais

- Probabilidade:
  - Incerteza é um conceito chave na área de reconhecimento de padrões/aprendizado de máquina.
  - **Aprendizado → redução de incerteza**.
  - Probabilidade de um evento é a fração de vezes que o evento ocorre em relação a um número total de tentativas, no limite em que o número de tentativas tende a infinito.
- Regras Básicas de Probabilidade:
  - Soma.
  - Produto.

# Leis Básicas de Probabilidade

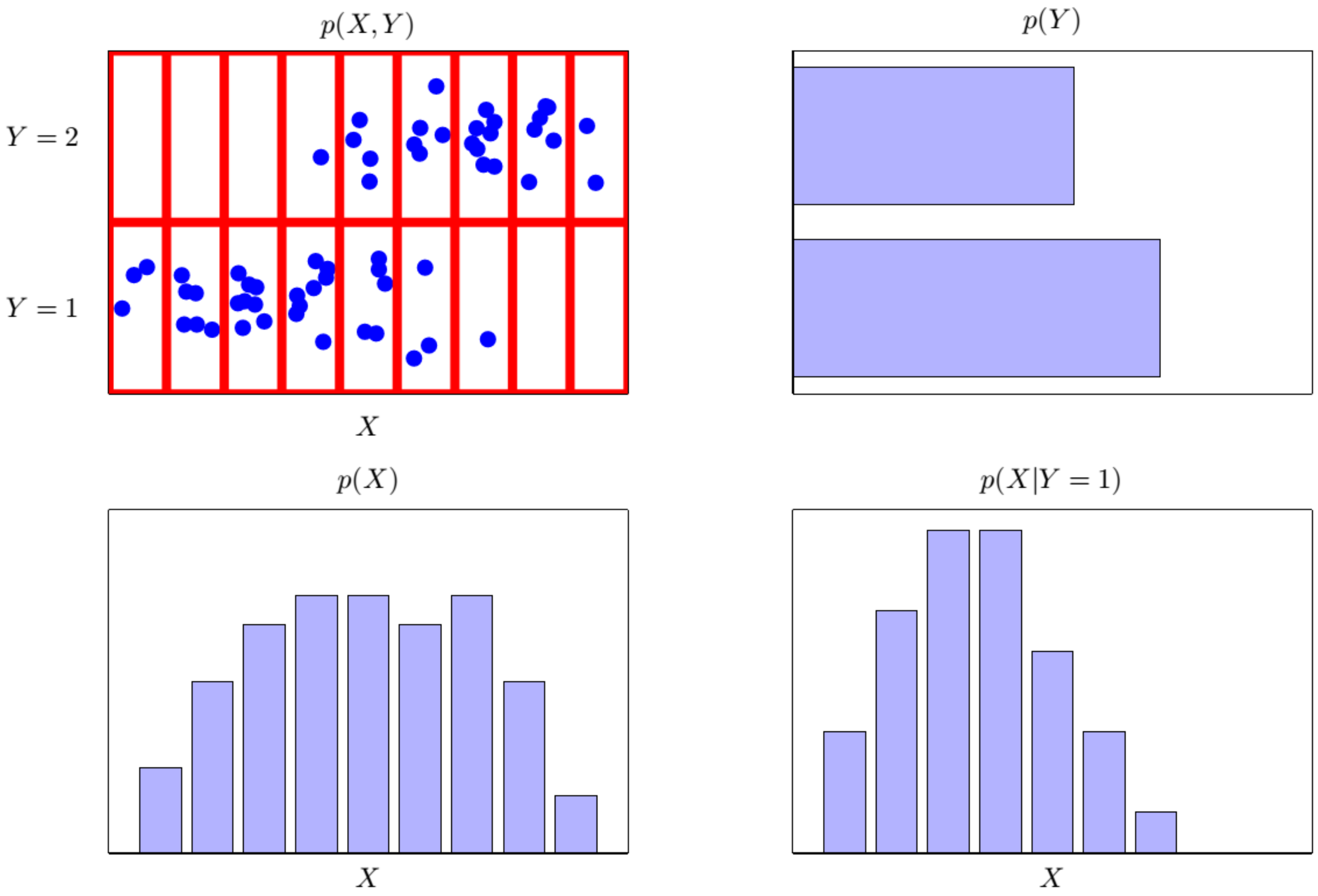
**sum rule** $$p(X) = \sum_Y p(X,Y)$$

**product rule** $$p(X,Y) = p(Y|X)p(X).$$

*Bayes' theorem* $$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

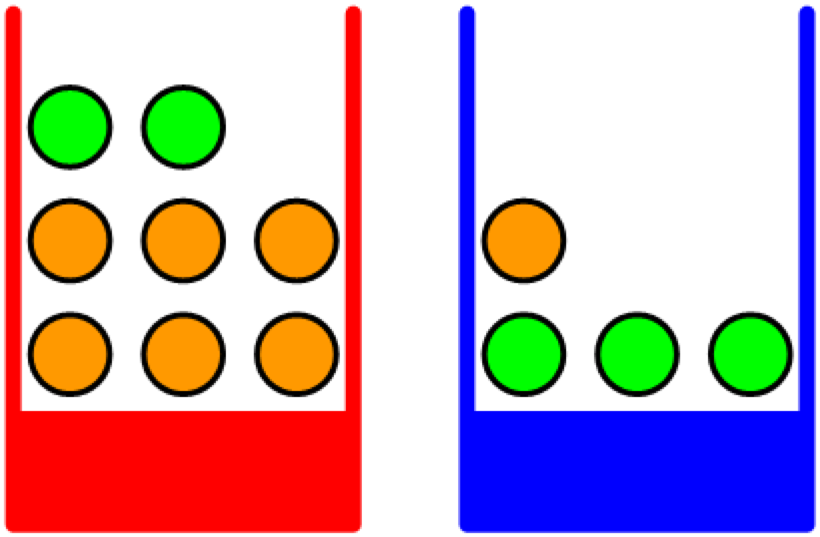
p(X,Y) é a probabilidade conjunta (probabilidade de X **e** Y)
p(Y|X) é a probabilidade condicional (probabilidade de Y **dado** X)
p(X) é a probabilidade marginal (probabilidade de X)

# Leis Básicas de Probabilidade



# Leis Básicas de Probabilidade

- Laranjas e Maçãs em duas caixas:



**Perguntas:**

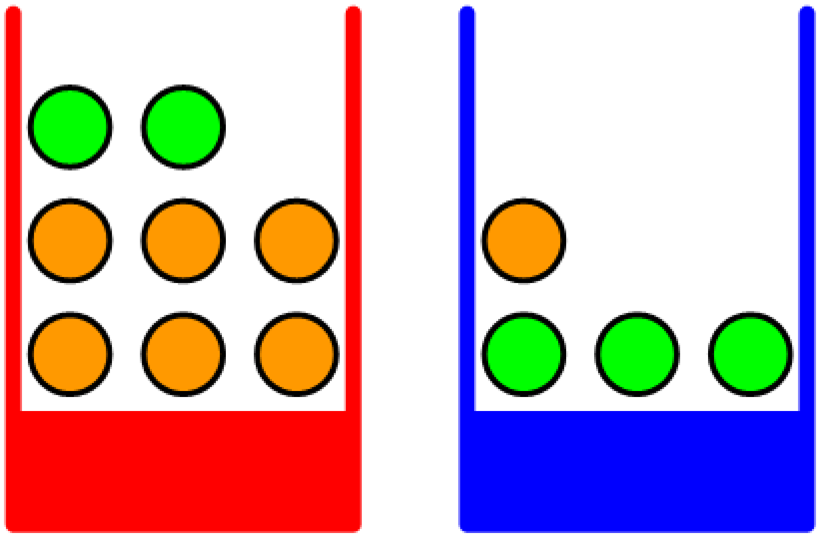
1) Qual a probabilidade de selecionar uma maçã?

2) **Dado que foi observada uma laranja, qual a probabilidade da caixa escolhida ser a azul?**

**Leis Básicas da Probabilidade!**

# Leis Básicas de Probabilidade

- Problema das Laranjas e Maçãs:



$$p(B = r) = 4/10$$
$$p(B = b) = 6/10$$
$$\begin{aligned} p(F = a|B = r) &= 1/4 \\ p(F = o|B = r) &= 3/4 \\ p(F = a|B = b) &= 3/4 \\ p(F = o|B = b) &= 1/4. \end{aligned}$$

$$p(B = r) + p(B = b) = 1$$

$$p(F = a|B = r) + p(F = o|B = r) = 1$$

$$p(F = a|B = b) + p(F = o|B = b) = 1$$

# Leis Básicas de Probabilidade

$$p(B=r) = 4/10$$

$$p(B=b) = 6/10$$

$$\begin{aligned} p(F=a|B=r) &= 1/4 \\ p(F=o|B=r) &= 3/4 \\ p(F=a|B=b) &= 3/4 \\ p(F=o|B=b) &= 1/4. \end{aligned}$$

- Qual a probabilidade de escolher maçã?

$$p(X) = \sum_Y p(X|Y)p(Y)$$

$$\begin{aligned} p(F=a) &= p(F=a|B=r)p(B=r) + p(F=a|B=b)p(B=b) \\ &= \frac{1}{4} \times \frac{4}{10} + \frac{3}{4} \times \frac{6}{10} = \frac{11}{20} \end{aligned}$$

$$p(F=o) = 1 - 11/20 = 9/20$$

# Leis Básicas de Probabilidade

$$p(B = r) = 4/10$$

$$p(B = b) = 6/10$$

$$p(F = o) = 1 - 11/20 = 9/20$$

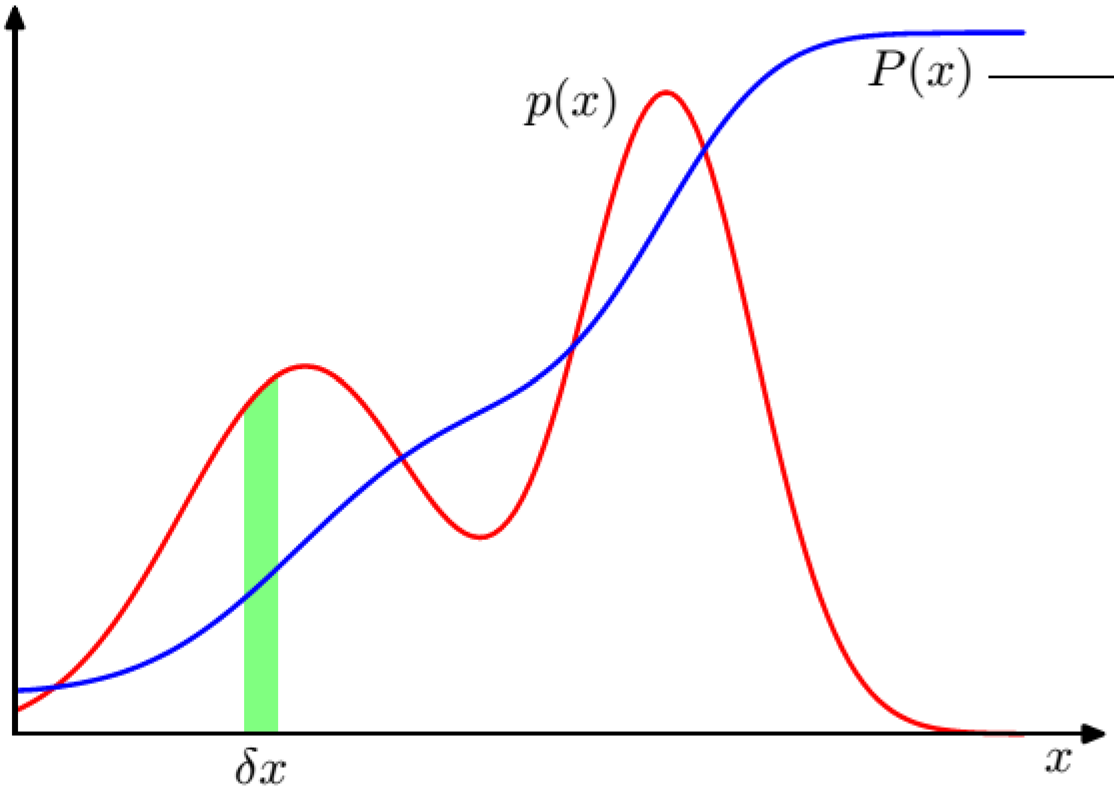
$$\begin{aligned} p(F = a|B = r) &= 1/4 \\ p(F = o|B = r) &= 3/4 \\ p(F = a|B = b) &= 3/4 \\ p(F = o|B = b) &= 1/4. \end{aligned}$$

- Qual a probabilidade da caixa escolhida ser a vermelha, dado que a fruta selecionada é laranja?

$$p(Y|X) = \frac{p(X|Y)p(Y)}{p(X)}$$

$$p(B = r|F = o) = \frac{p(F = o|B = r)p(B = r)}{p(F = o)} = \frac{3}{4} \times \frac{4}{10} \times \frac{20}{9} = \frac{2}{3}.$$

# Função Densidade de Probabilidade




Cumulative distribution function

$$P(z) = \int_{-\infty}^{z} p(x)\,\mathrm{d}x$$

$$P'(x) = p(x)$$

$$\boxed{\begin{aligned} p(x) &\geqslant 0 \\ \int_{-\infty}^{\infty} p(x)\,\mathrm{d}x &= 1. \end{aligned}}$$

Probabilidade que x esteja no intervalo (x, x + δx)
é dado por p(x)δx para δx → 0.

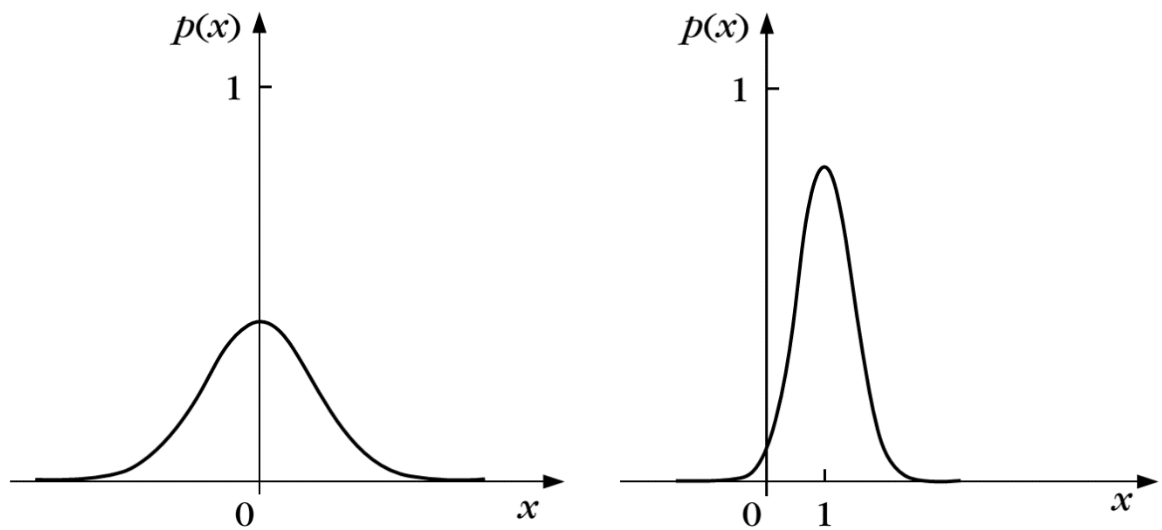
# Função Densidade de Probabilidade Gaussiana

- Caso unidimensional:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$\mu = E[x] \equiv \int_{-\infty}^{+\infty} x p(x) dx$$

$$\sigma^2 = E[(x-\mu)^2] \equiv \int_{-\infty}^{+\infty} (x-\mu)^2 p(x) dx$$



# Função Densidade de Probabilidade Gaussiana

- Caso multidimensional:

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{l/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x}-\boldsymbol{\mu})\right)$$

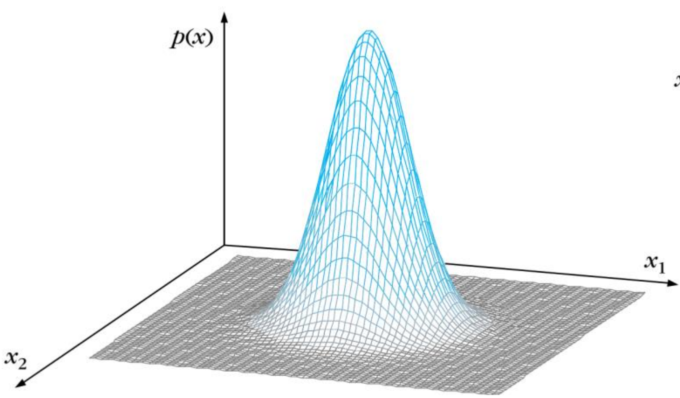
$$\boldsymbol{\mu} = E[\boldsymbol{x}]$$

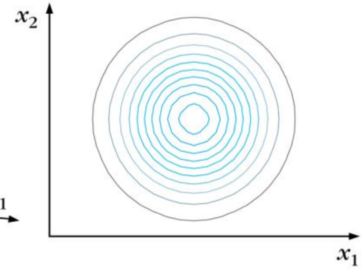
$$\Sigma = E[(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T]$$

**2-D:**

$$\Sigma = E\left[\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}\begin{bmatrix} x_1-\mu_1, & x_2-\mu_2 \end{bmatrix}\right]$$

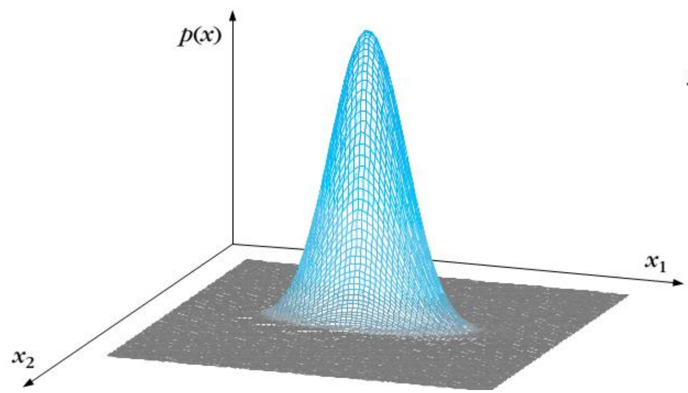
$$= \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

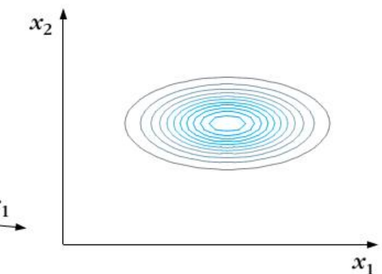


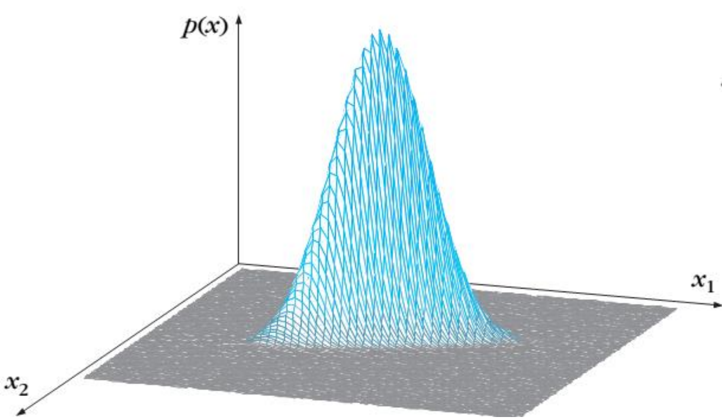

$\sigma_1^2 = \sigma_2^2.$

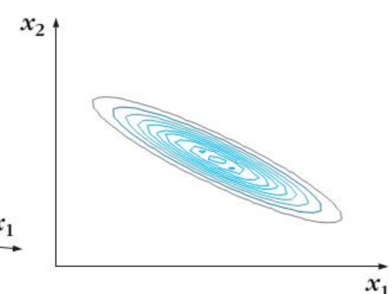
$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$




$\sigma_1^2 >> \sigma_2^2.$

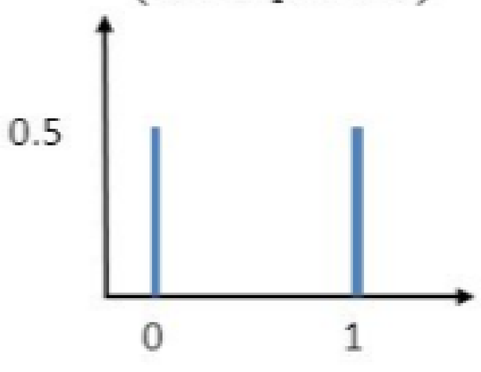



nondiagonal $\Sigma$

# Bernoulli distribution

- Bernoulli (pmf):




$$f(k;p) = \begin{cases} p & \text{if } k = 1, \\ 1-p & \text{if } k = 0. \end{cases}$$

$$f(k;p) = p^k (1-p)^{1-k} \quad \text{for } k \in \{0, 1\}.$$

# Referências

- Livro Pattern Classification – Duda/Hart;
- Livro Pattern Recognition and Machine Learning - Bishop;
- Aulas Prof. David M. J. Tax e Prof. Bob Duin – TU Delft.