

# Introdução a Aprendizagem de Máquina Utilizando a Linguagem Python

**Profº Gustavo Miranda**  
**[gustavo.oliveira@penedo.ufal.br](mailto:gustavo.oliveira@penedo.ufal.br)**



**UFAL**  
CAMPUS ARAPIRACA  
U. E. PENEDO

# Tópicos da Aula

— — — —

1º

**Introdução a Inteligência Artificial**

2º

**Definições e Conceitos da Inteligência Artificial**

3º

**Introdução a Aprendizagem de Máquina**

4º

**Exemplo de Classificação de Dados**

5º

**Aplicação de Árvore de Decisão na Classificação de Dados**



# Quem sou eu?



## Professor Gustavo Miranda:

- **Doutor** em Inteligência Computacional - **UFPE**
- **Mestre** em Inteligência Computacional - **UFPE**
- **Graduado** em Licenciatura em Computação - **UPE**
- **Experiência** nos seguintes campos:
  - Machine Learning
  - Ciência de Dados

## Informações Adicionais

- <https://github.com/GustavoHFM0>

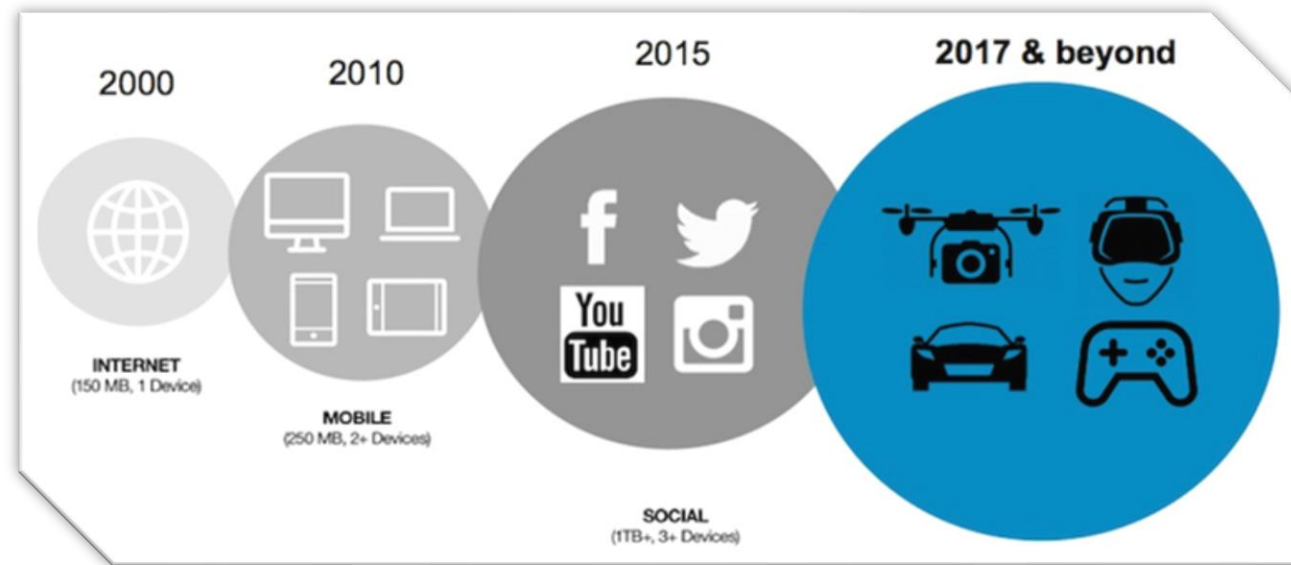
# Introdução a Inteligência Artificial



**UFAL**  
CAMPUS ARAPIRACA  
U. E. PENEDO

# Crescimento de Dados Digitais

- Nos últimos anos, aplicações do mundo real têm lidado com um grande **crescimento na quantidade de dados digitais**.
  - Uso de celulares, sensores, redes sociais, aplicações entre outros.



# Crescimento Contínuo de Dados

— — — —

NOTÍCIAS-CORPORATIVAS

**Compras Online: após pandemia, setor registra 40% de crescimento e vendas pela internet tornam-se a modalidade favorita dos consumidores**

🕒 29 de junho de 2020 👁 243 Visualizações 📖 4 minutos para ler

TECNOLOGIA

**Tráfego de internet no Brasil cresce 6 vezes nas redes da Angola Cables na pandemia**

*Dado foi apresentado pelo CEO da Angola Cables durante o Futurecom Digital Summit*

🕒 4 de julho de 2020 📖 4 minutos de leitura

**Claro estreia serviço de internet 5G no Brasil a partir da semana que vem**

Operadora implementará a novidade em São Paulo e no Rio de Janeiro inicialmente, regiões com maior demanda de tráfego de dados no País

Início > Análises > Brasil registra 33,2 milhões de contratos de Internet fixa em maio

Análise

**Brasil registra 33,2 milhões de contratos de Internet fixa em maio**

Por **Henrique Julião** - 09/07/20, 18:26 Atualizado em 13/07/20, 13:30



# Entendimento dos Dados


-----

## O que fazer com os dados armazenados?

- Os analistas de dados perceberam que **conhecimento pode ser extraído desses dados.**
  - Conhecimento vale dinheiro.



# Surgimento de Áreas de Pesquisa



IT Forum 365  
A VOZ DA TI

Eventos Colunas Notícias Podcasts

## Como a ciência de dados já está transformando a educação

Em palestra durante a Campus Party, cientista Priscilla Nascimento abordou vantagens tecnologia no ensino



Diário do Nordeste

HOME CORONAVÍRUS DIAS MELHORES METRO POLÍTICA JOGADA

## Com Ciência de Dados, solução otimiza atendimento ao cliente

Escrito por Meireles e Freitas Cobrança Digital, 04:00 / 01 de Julho de 2020.

Pacote de tecnologias Smart One, da Meireles e Freitas Cobrança Digital, auxilia empresas com alto número de inadimplentes.



JORNAL DA USP

USP

PÁGINA INICIAL CIÊNCIAS TECNOLOGIA EDUCAÇÃO CULTURA ATUALIDADES UNIVERSIDADE INSTITUCIONAL

Home > Universidade > Ações para comunidade > Pesquisadores aplicam ciência de dados e Inteligência Artificial no combate ao crime

Ações para comunidade - 27/07/2020

## Pesquisadores aplicam ciência de dados e Inteligência Artificial no combate ao crime

Identificar padrões de crimes e construir ferramentas para prever sua ocorrência são objetivos da iniciativa que terá como laboratório São Carlos, no interior de São Paulo

## Quase 60% das empresas admitem que pandemia ampliou demanda por soluções que aplicam inteligência artificial, aponta estudo global

Terça, 28 Julho 2020 18:02 Escrito ou enviado por Alan Mariasch Adicionar comentário SEGS.com.br - Categoria: Seguros Imprimir





# IA na Mídia Convencional

Why artificial intelligence is enjoying a renaissance - *Economist*<sup>1</sup>

How Cognitive Systems Could Redefine The Way Governments Work - *Forbes*<sup>2</sup>

## Artificial-Intelligence Experts Are in High Demand

- *The Wall Street Journal*<sup>3</sup>

## *Artificial Intelligence Swarms Silicon Valley on Wings and Wheels*

- *The New York Times*<sup>4</sup>

### March of the machines

What history tells us about the future of artificial intelligence—and how society should respond - *Economist*<sup>5</sup>

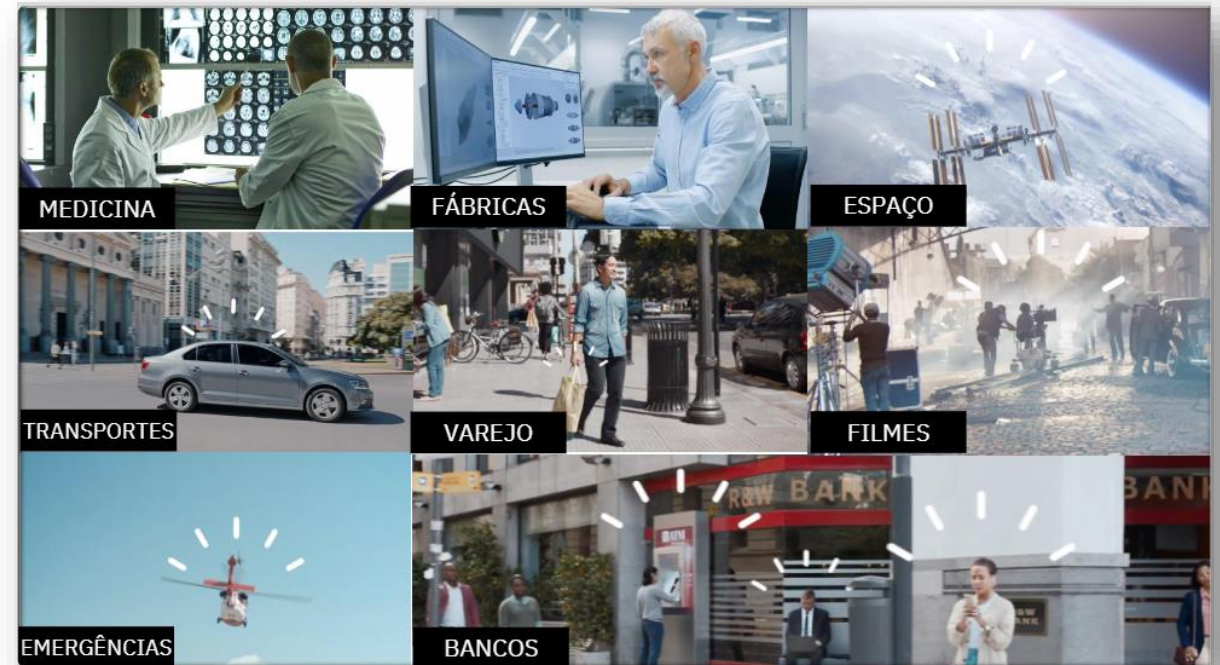
We have to upgrade our skills to match intelligent machines - *Business Insider*<sup>6</sup>



# 4º Revolução Industrial

-----

- A IA é uma das maiores facetas dessa revolução.
- Ela afetará quase todos os setores, assim como nas Revoluções Industriais anteriores.



# Aplicações da Inteligência Artificial

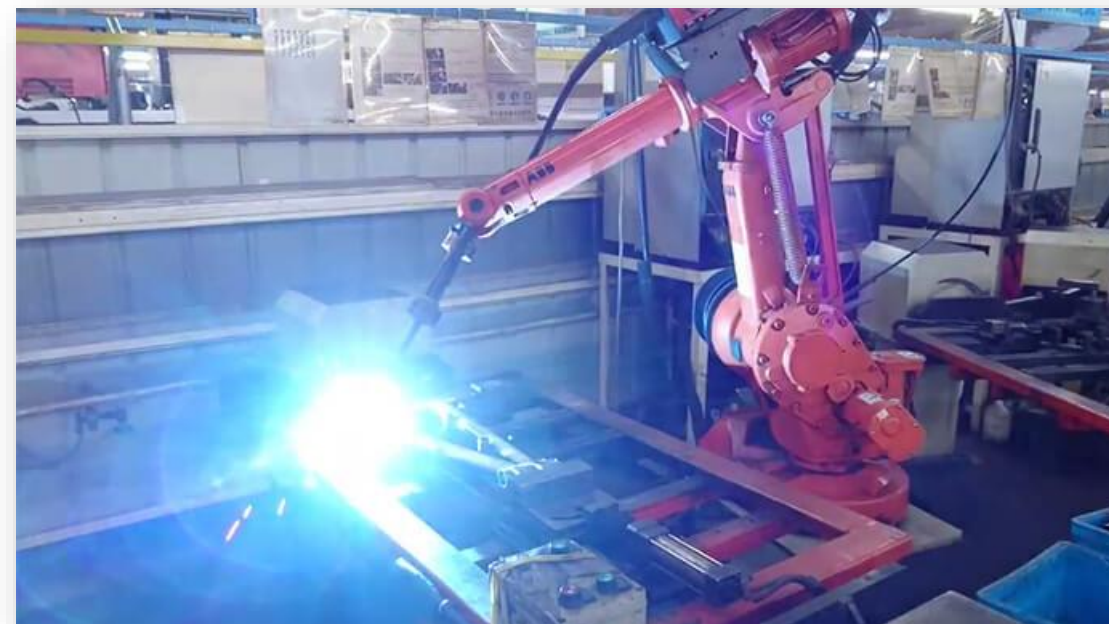


**UFAL**  
CAMPUS ARAPIRACA  
U. E. PENEDO

# Automação com IA

----

- **Detecção de anomalias em máquinas usando uma rede neural.**
  - Leituras anômalas de sensores preveem possíveis falhas em máquinas.
- **Robôs soldadores:**
  - Detecção preventiva de falhas e avaliação de manutenção.



# O Carro Autônomo

----

- Veículos automáticos têm se tornado **cada vez mais acessíveis**.
- Carros vão se **comunicar, socializar e colaborar com outras coisas**, incluindo outros veículos, semáforos, estacionamentos e revendedores.

estacionamentos e revendedores

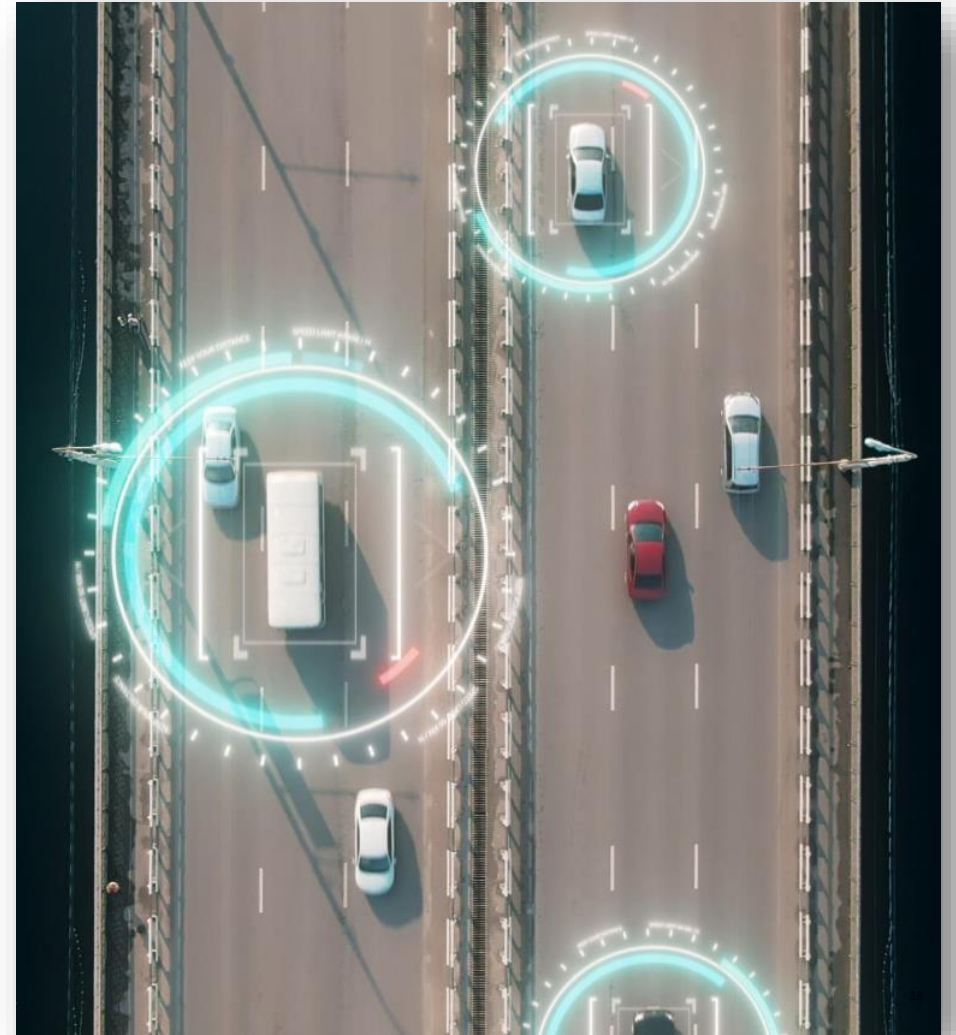




# Carro Conectado Hoje

----

- Produz mais de **25 gb** de dados toda hora.
- **Conectado com:**
  - O fabricante para compartilhar alertas e registros de performance.
  - Proporciona uma experiência única ao passageiro.



# Audi usa Aprendizado de Máquina

— — — —

- Dados são **coletados de sensores** para determinar a posição e as dimensões dos carros.
- O Aprendizado de máquina otimiza o movimento dos carros e **melhora o uso do espaço.**

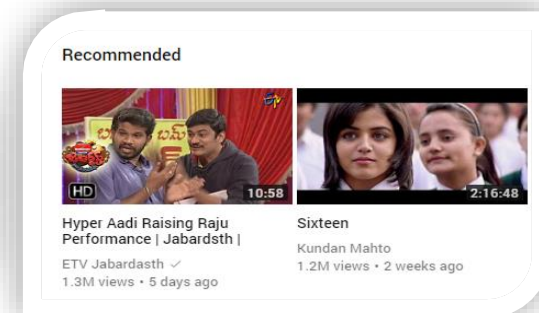


# Aplicações da Inteligência Artificial

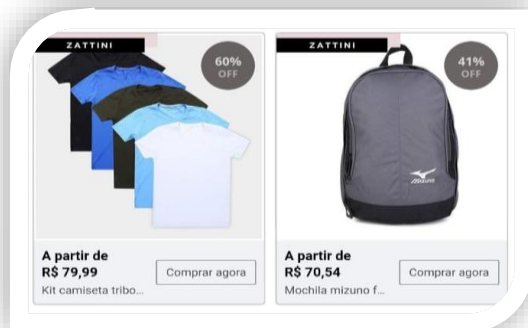
## Reconhecimento Facial



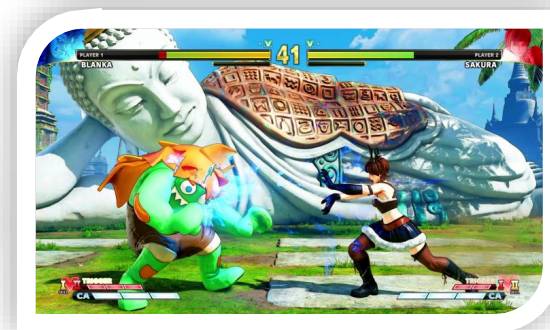
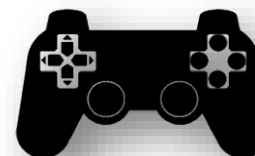
## Recomendação de Vídeos



## Recomendação de Produtos



## Níveis de Dificuldade em Jogos

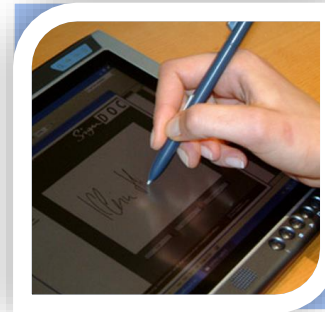




# Aplicações da Inteligência Artificial



**Reconhecimento de voz:** entendimento da linguagem natural.



**Reconhecimento de caligrafias:** reconhecer assinaturas.



**Tradução de texto:** por exemplo, do inglês para o português.



**Filtros de correios eletrônicos:** separando mensagens potencialmente perigosas das mensagens permitidas.

# Definições e Conceitos da Inteligência Artificial



# Inteligência Artificial (IA)

----

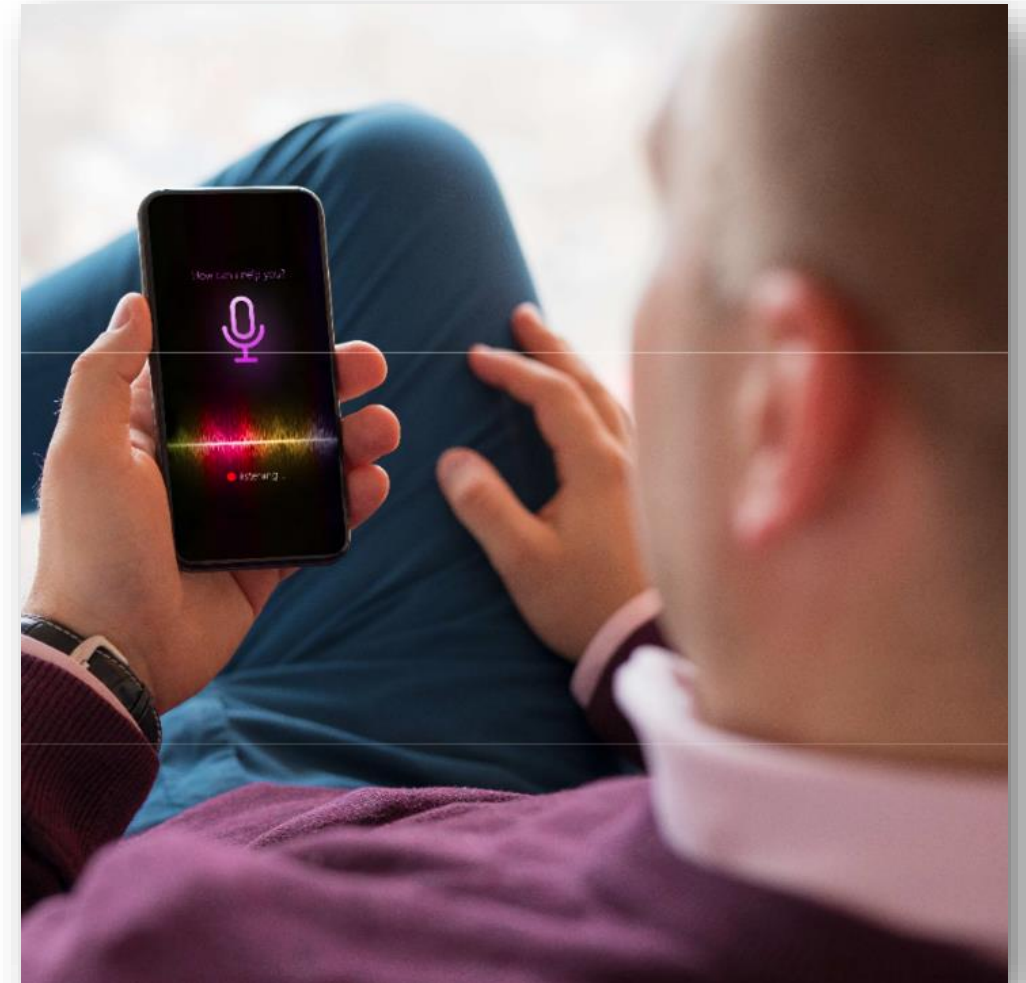
“Podemos definir **inteligência artificial**, no grosso modo, como a **capacidade das máquinas de pensarem como seres humanos**: aprender, perceber e decidir quais caminhos seguir, de forma racional, diante de determinadas situações.”



# IA Fraca

-----

- Focada em **realizar tarefas específicas.**
- **Exemplo:**
  - Comprar um livro através de um dispositivo de voz.



# IA Abrangente

-----

- **Utilizam conhecimento e dados** de uma indústria para **treinar** seus sistemas.
- **Exemplos:**
  - Carros autônomos.
  - Coleção de IA Fracas.
  - Tomam decisões.





# IA Forte

-----

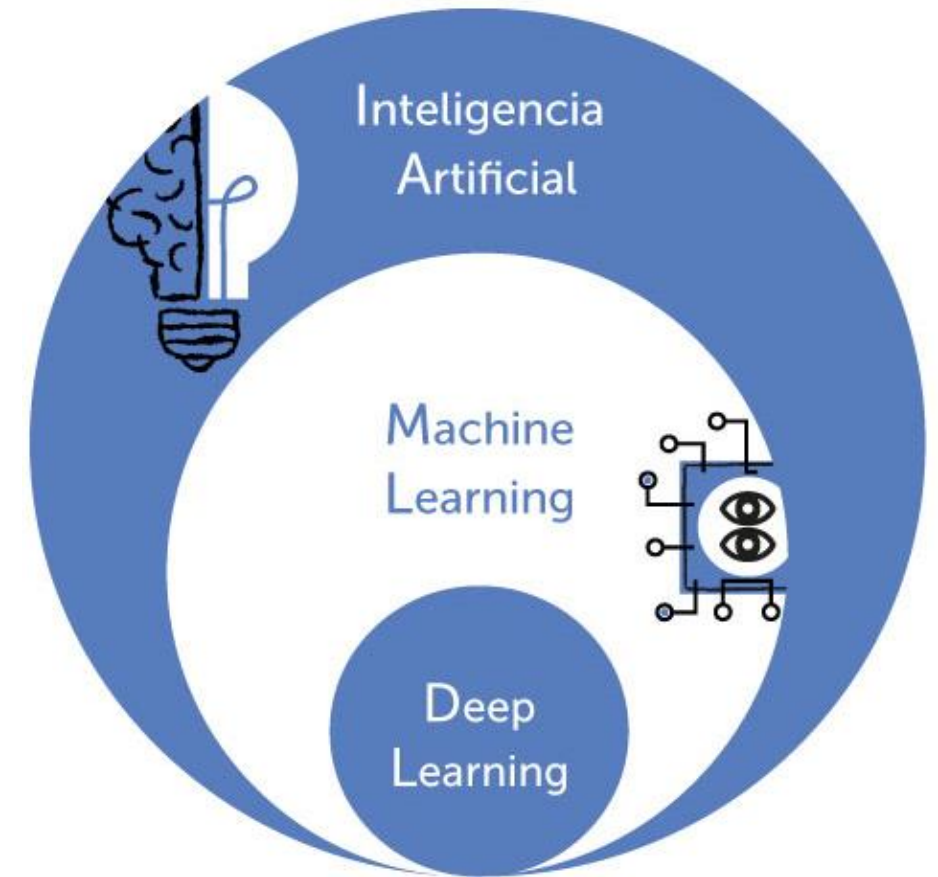
- Máquinas que podem **desempenhar trabalho intelectual como um humano.**
- **Exemplo:**
  - Atualmente a IA não consegue elaborar ideias criativas.



# Aprendizagem de Máquina

-----

- “Machine learning é uma área da ciência da computação que significa **Aprendizado de Máquina**. Faz parte do conceito de **Inteligência Artificial**, que busca meios de ensinar a máquina a realizar tarefas que seriam executadas por pessoas.”



# Introdução a Aprendizagem de Máquina



**UFAL**  
CAMPUS ARAPIRACA  
U. E. PENEDO



# Tipos de Aprendizado de Máquina



Aprendizado  
Supervisionado

A diagram showing the text 'Aprendizado Supervisionado' centered between two gray curved lines that form an open circle.



Aprendizado Não-  
supervisionado

A diagram showing the text 'Aprendizado Não-supervisionado' centered between two gray curved lines that form an open circle.



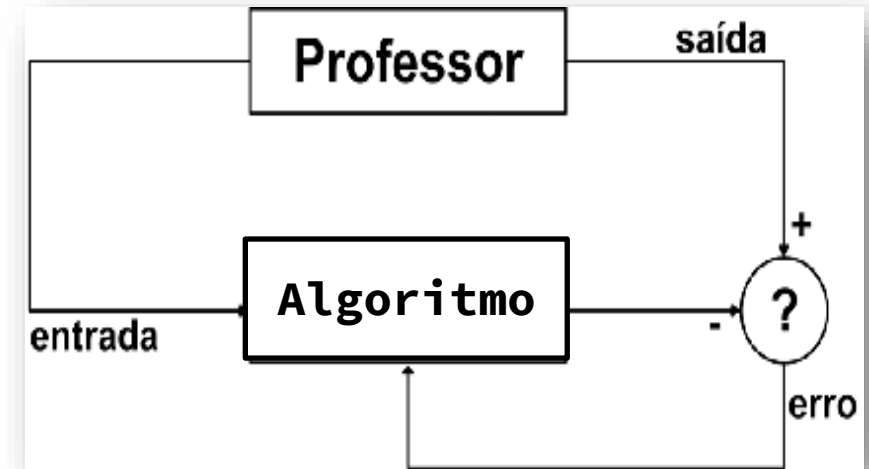
Aprendizado por  
Reforço

A diagram showing the text 'Aprendizado por Reforço' centered between two gray curved lines that form an open circle.

# Aprendizado Supervisionado

-----

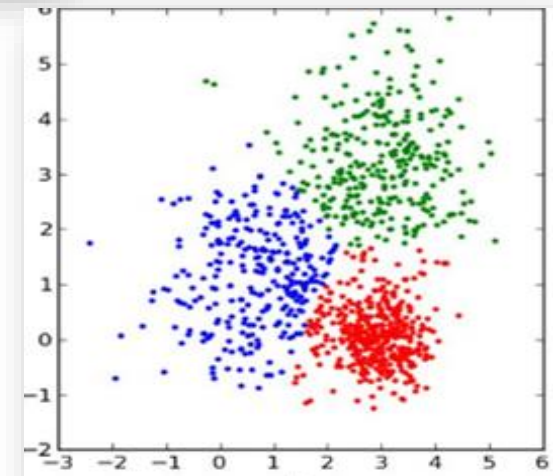
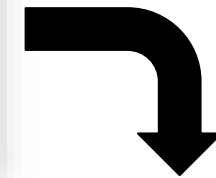
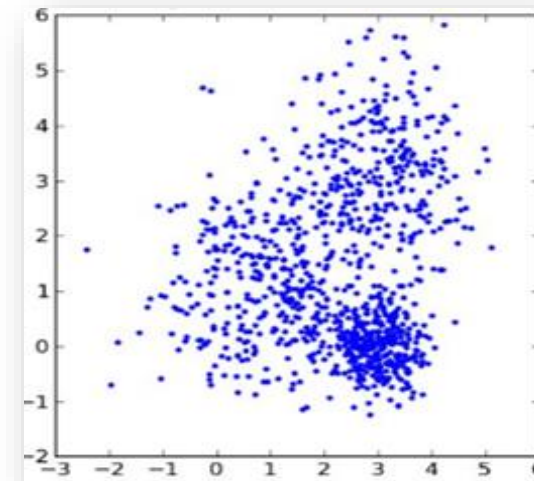
- São apresentados ao computador exemplos de **entradas e saídas desejadas**, fornecidas por um “professor”.
- O objetivo é aprender uma regra geral que **mapeia** as entradas e saídas.
- Exemplos de entrada e saída:
  - **Entrada:** foto; **Saída:** nome da pessoa.



# Aprendizado Não-supervisionado

-----

- **Nenhum tipo de saída é dada** ao algoritmo de aprendizado. O algoritmo **por conta própria** forma grupos com os dados mais semelhantes.
- Descobrir padrões nos dados:
  - **Identificar grupos:**
- Exemplos de entradas:
  - **Entrada:** várias fotos de três pessoas diferentes.

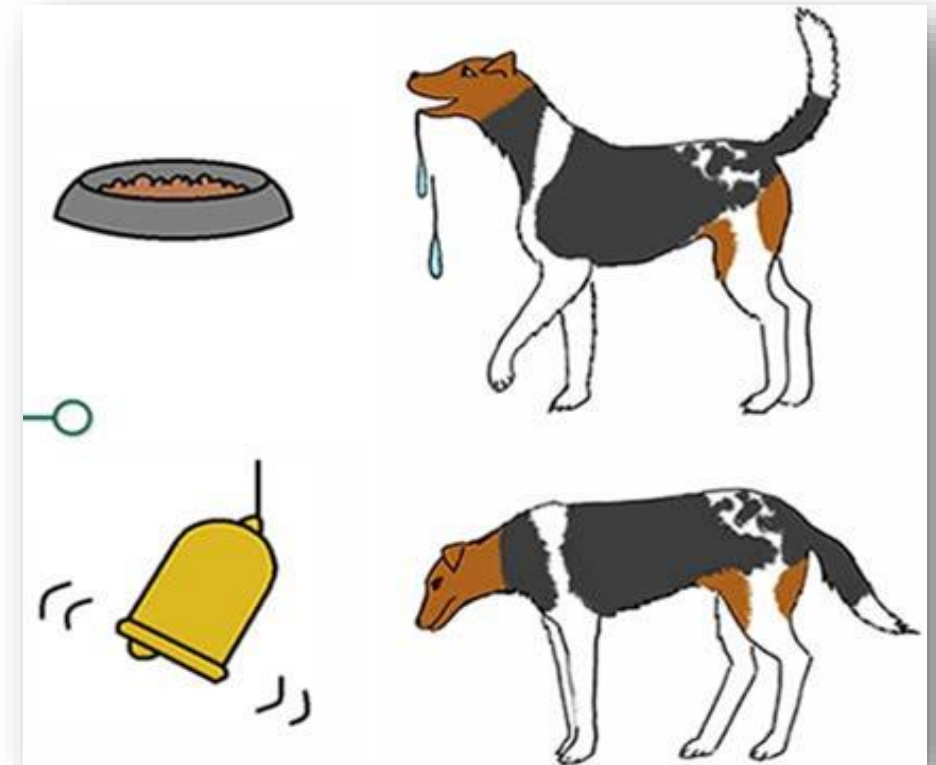


UFAL

# Aprendizado Por Reforço

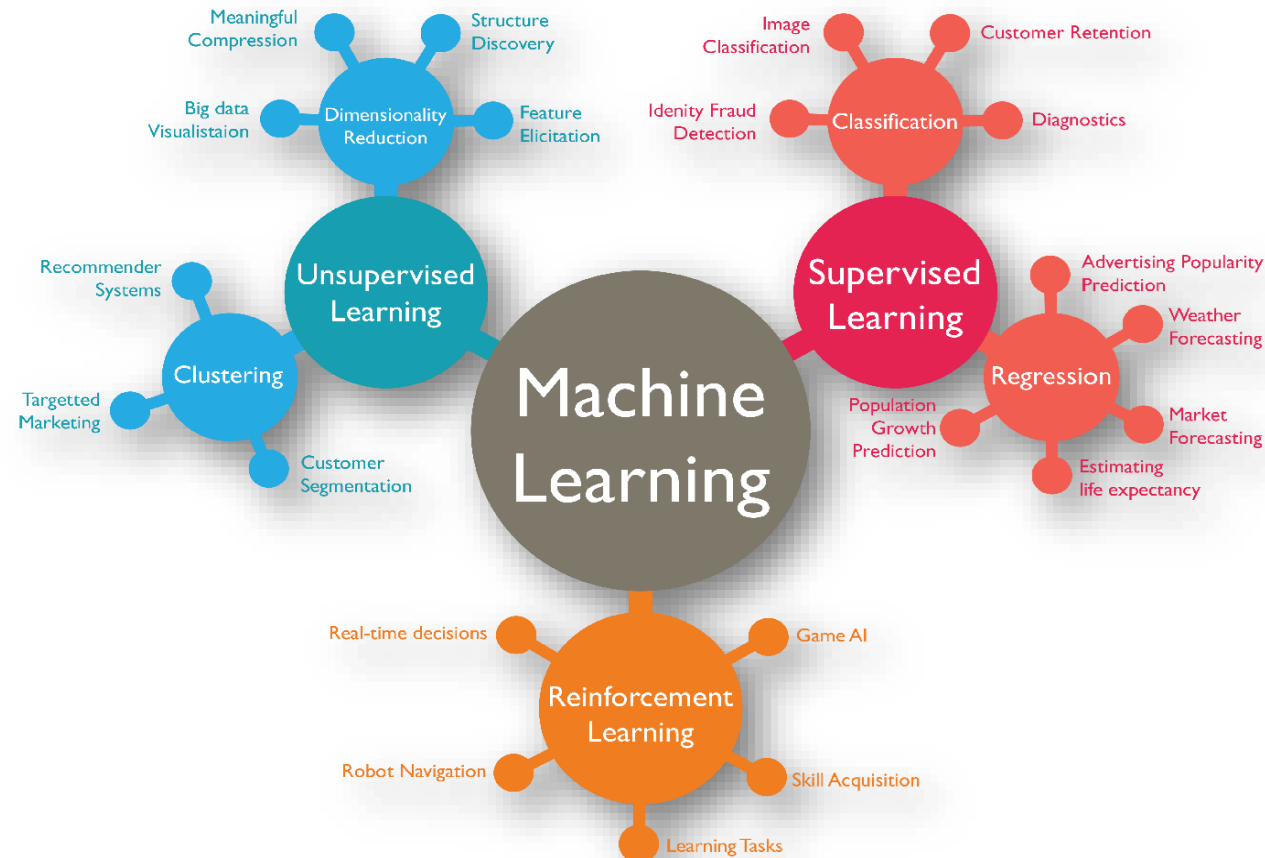
-----

- O programa desempenha uma atividade em um ambiente dinâmico.
- O ambiente dinâmico fornece ao programa **premiações** e **punições**, na medida que ele explora as opções.
- **Exemplo:** Melhorar a estratégia de jogo apenas jogando contra um oponente.



# Subáreas de Aprendizagem de Máquina

-----



# Aplicações do Aprendizado Supervisionado

— — — —

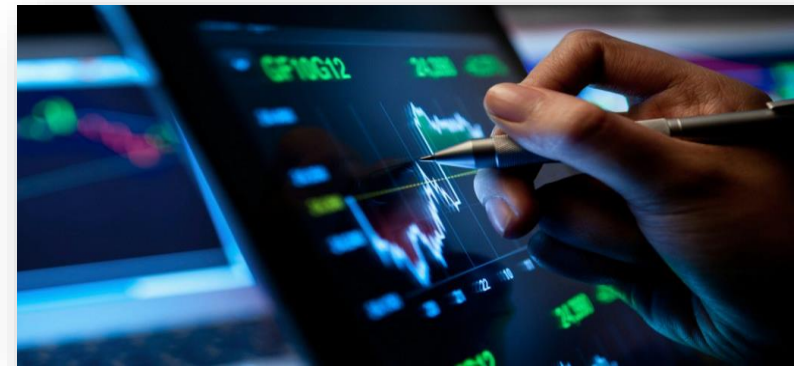
- **Classificação:**

- Classificar novos exemplos em grupos existentes.
- **Ex:** reconhecimento facial.



- **Regressão:**

- Estimar um valor flutuante baseando em informações passadas.
- **Ex:** previsão do mercado de ações.



# Aplicações do Aprendizado Não-Supervisionado

## • Agrupamento de dados:

- Verifica os dados mais semelhantes e os agrupa.
- **Ex:** agrupar fotos da mesma pessoa.



## • Sistemas de Recomendação:

- Baseado nas características do usuário, recomenda opções semelhantes.
- **Ex:** recomendação de produtos ou filmes.





# Aplicações do Aprendizado Por Reforço

-----

- **IA em Games:**
  - Verifica quais golpes geram mais efeitos e quais são mais punidos.
  - **Ex:** níveis de dificuldade em jogos de luta.
- **Navegação de Robôs:**
  - Aprende a navegar pelo terreno a partir da experiência.
  - **Ex:** robôs enviados para navegar em marte.





# Exemplo de Classificação de Dados



# Exemplo de Classificação de Dados

----

- Uma empresa de botânica precisa de um **sistema capaz de classificar plantas** do tipo íris nas suas **três subespécies** de maneira automática.



# Classificação de Espécies

— — — —

- As **três espécies da flor íris** são **muito semelhantes** e levam a confusão de não especialistas durante a classificação.



**Iris Versicolor**



**Iris Setosa**

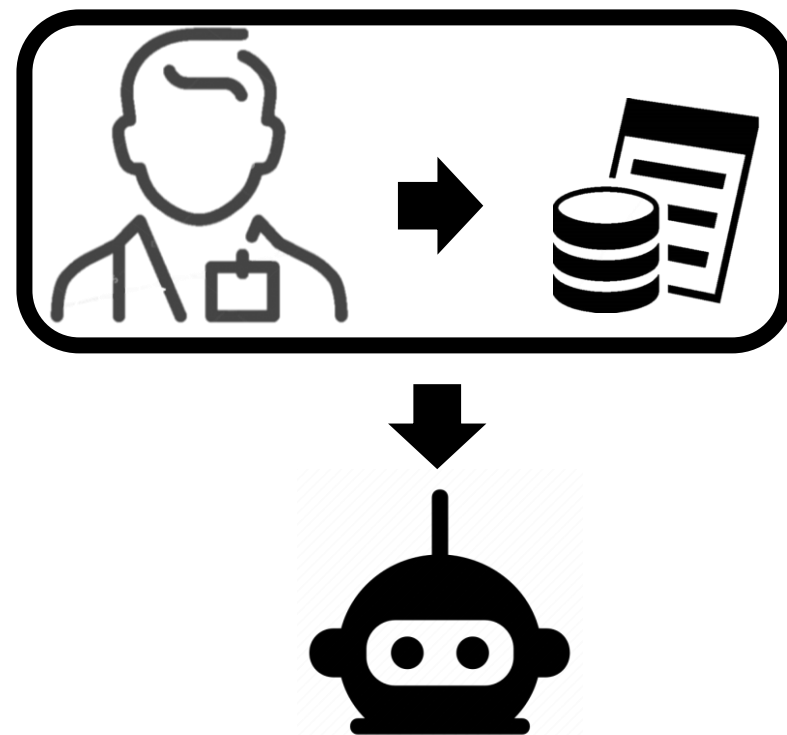


**Iris Virginica**

# Papel do Especialista

-----

- Um dos critérios a serem considerados é a **existência de profissionais especialistas** que tenham um bom domínio do conhecimento relacionado ao problema.
- Os dados na base de conhecimento são essencialmente **adicionados por** seres humanos **especialistas** em um domínio específico.



# Reconhecer Padrões nos Dados

----

- **Dados de Entrada (X):**

- Sépala:

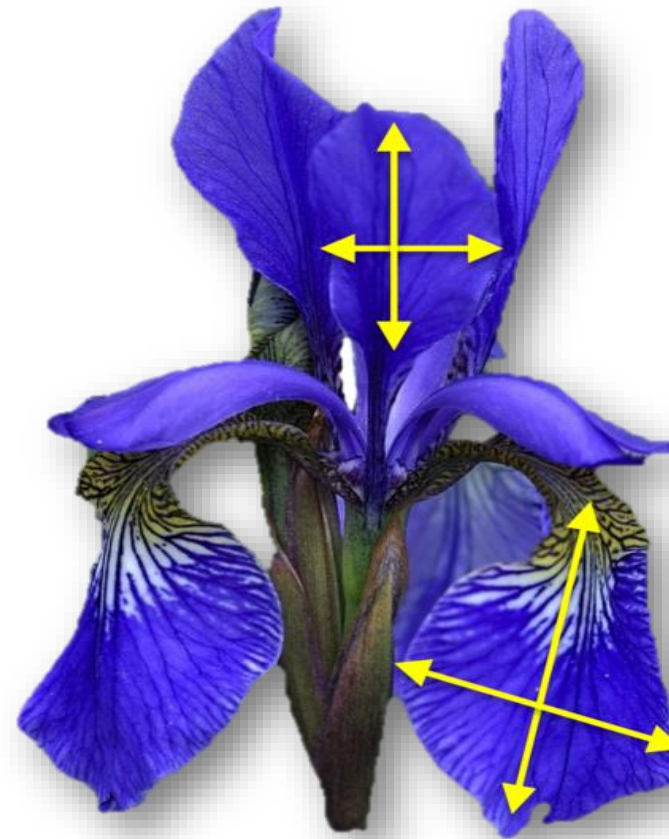
- Comprimento.
    - Largura.

- Pétala:

- Comprimento.
    - Largura.

- **Dados de Saída (Y):**

- Tipo de flor.



**Pétala**

**Sépala**

# Conjunto de Dados

-----

- Agrupamento de variáveis do mesmo interesse em um **arquivo organizado**.

Coletas	Sépala Largura	Sépala Cumprimento	Pétala Largura	Pétala Cumprimento	Classe
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
3	6.4	3.5	4.5	1.2	Versicolor
4	5.9	3.0	5.0	1.8	Versicolor
...	...	...	...	...	....
150	5.8	5.1	4.1	3	Virginica

# Características do Conjunto de Dados

Coletas	Sépala Largura	Sépala Cumprimento	Pétala Largura	Pétala Cumprimento	Classe
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
3	6.4	3.5	4.5	1.2	Versicolor
4	5.9	3.0	5.0	1.8	Versicolor
...	...	...	...	...	....
150	5.8	5.1	4.1	3	Virginica

Quantidade de exemplos para treinamento = **N**

# Características do Conjunto de Dados

Coletas	Sépala Largura	Sépala Cumprimento	Pétala Largura	Pétala Cumprimento	Classe
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
3	6.4	3.5	4.5	1.2	Versicolor
4	5.9	3.0	5.0	1.8	Versicolor
...	...	...	...	...	....
150	5.8	5.1	4.1	3	Virginica



Dados de Entrada = **X**



# Características do Conjunto de Dados

Coletas	Sépala Largura	Sépala Cumprimento	Pétala Largura	Pétala Cumprimento	Classe
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
3	6.4	3.5	4.5	1.2	Versicolor
4	5.9	3.0	5.0	1.8	Versicolor
...	...	...	...	...	....
150	5.8	5.1	4.1	3	Virginica

Dados de Saída = Y

# Características do Conjunto de Dados

-----

Coletas	Sépala Largura	Sépala Cumprimento	Pétala Largura	Pétala Cumprimento	Classe
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
3	6.4	3.5	4.5	1.2	Versicolor
4	5.9	3.0	5.0	1.8	Versicolor
...	...	...	...	...	....
150	5.8	5.1	4.1	3	Virginica

→ Observação  $x_t$

# Características do Conjunto de Dados

Coletas	Sépala Largura	Sépala Cumprimento	Pétala Largura	Pétala Cumprimento	Classe
1	5.1	3.5	1.4	0.2	Setosa
2	4.9	3.0	1.4	0.2	Setosa
3	6.4	3.5	4.5	1.2	Versicolor
4	5.9	3.0	5.0	1.8	Versicolor
...	...	...	...	...	....
150	5.8	5.1	4.1	3	Virginica

Classe  $y_t$  da respectiva observação  $x_t$



# Extração de Características dos Dados

- Os dados devem ser **separados por classe** para a **extração de características**.

Sépala Largura	Sépala Cumprimento	Pétala Largura	Pétala Cumprimento
5.1	3.5	1.4	0.2
4.9	3.0	1.4	0.2

Classe
Setosa
Setosa

**Classe Setosa**

6.4	3.5	4.5	1.2
5.9	3.0	5.0	1.8

Versicolor
Versicolor

**Classe Versicolor**

...	...	...	...
5.8	5.1	4.1	3

...
Virginica

**Classe Virginica**

# Aprendizado Sobre os Dados



**UFAL**  
CAMPUS ARAPIRACA  
U. E. PENEDO

# Aprendizado Sobre os Dados

----

- Para entender como funciona o processo de classificação, algumas etapas são necessárias:

**1) Visualização dos dados.**

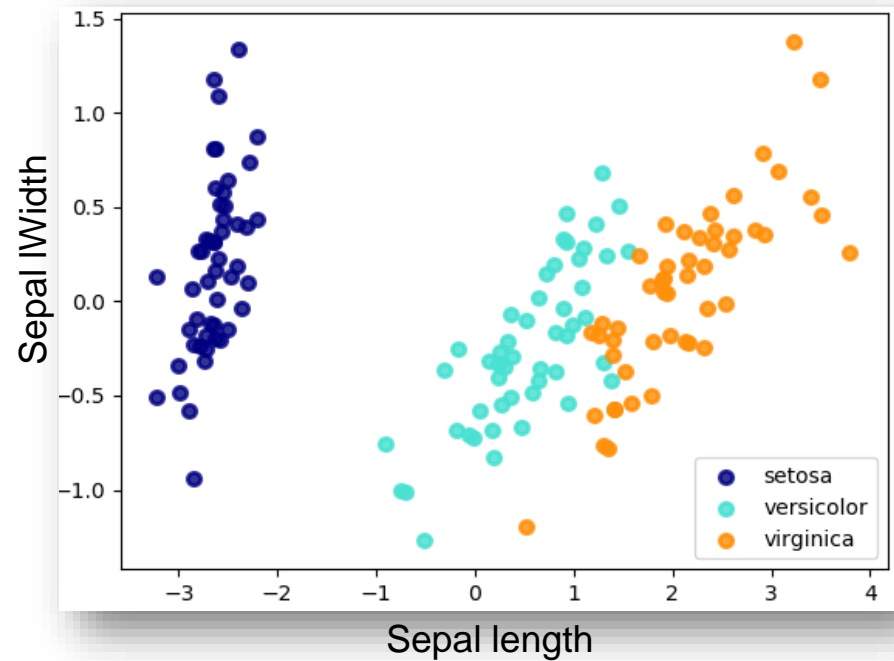
**3) Ajuste das fronteiras de decisão.**

**2) Criação das fronteiras de decisão.**

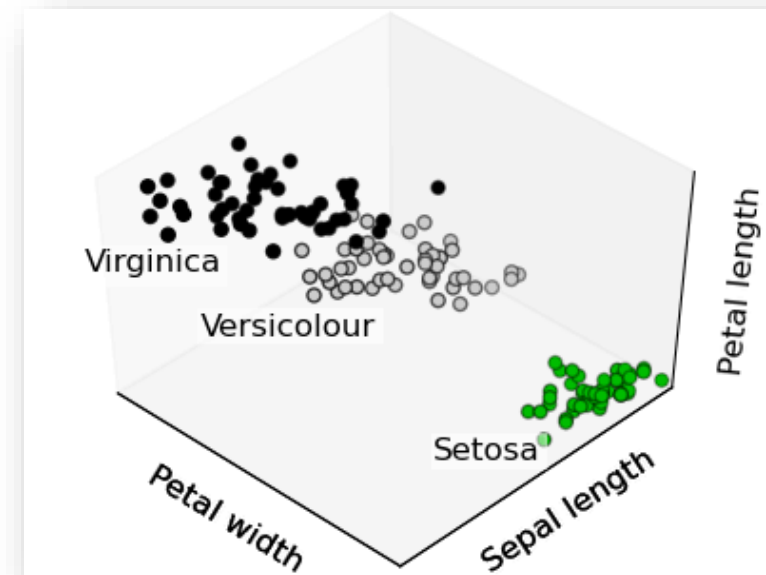
**4) Classificação de novos exemplos.**

# Visualização dos Dados

- A visualização do espaço de características ajuda a entender a distribuição dos dados e escolher o algoritmo adequado.



Dados com **dois** atributos.

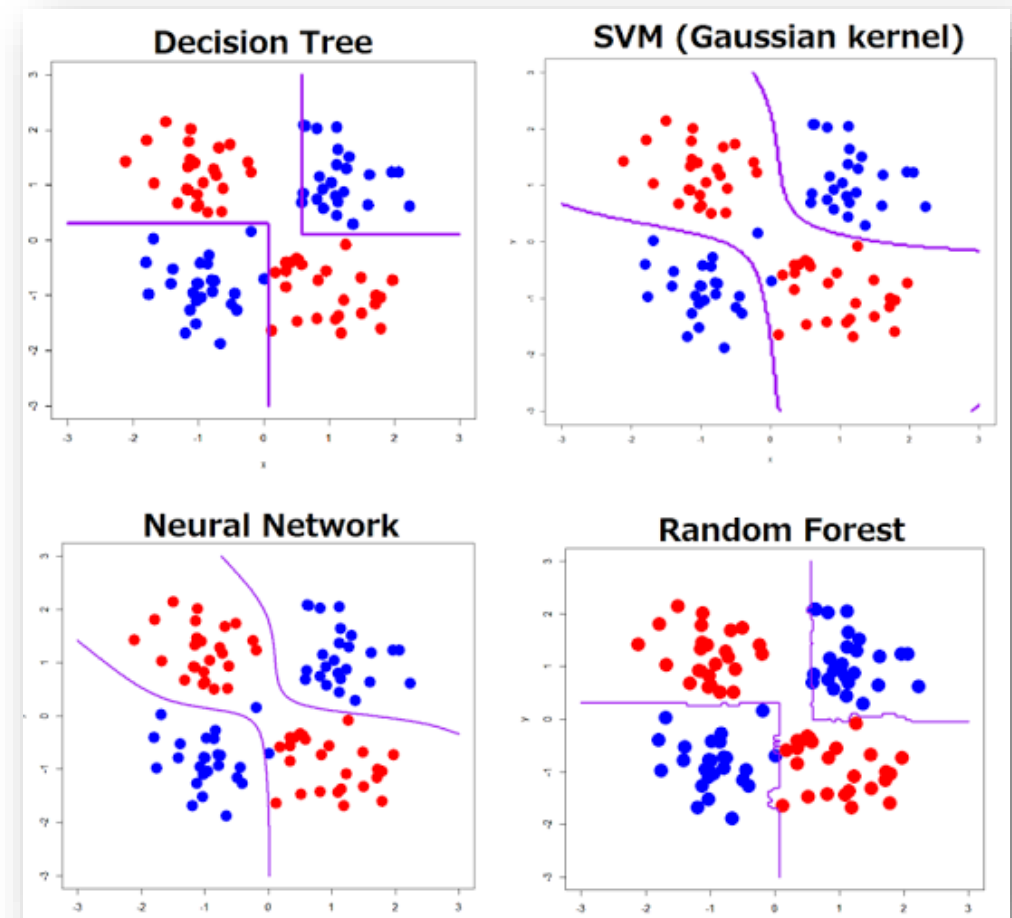


Dados com **três** atributos.

# Construção das Fronteiras de Decisão

## ■ Fronteiras de decisão:

- Modelagens computacionais que **separam** os dados de entrada (X) de acordo com suas respectivas classes (Y).
- As fronteiras de decisão são criadas durante **treinamento supervisionado** e representam o **conhecimento/aprendizado** sobre os dados.
- Cada algoritmo de aprendizado supervisionado cria suas fronteiras de decisão de maneira diferente.



*Fronteiras geradas por diferentes algoritmos.*



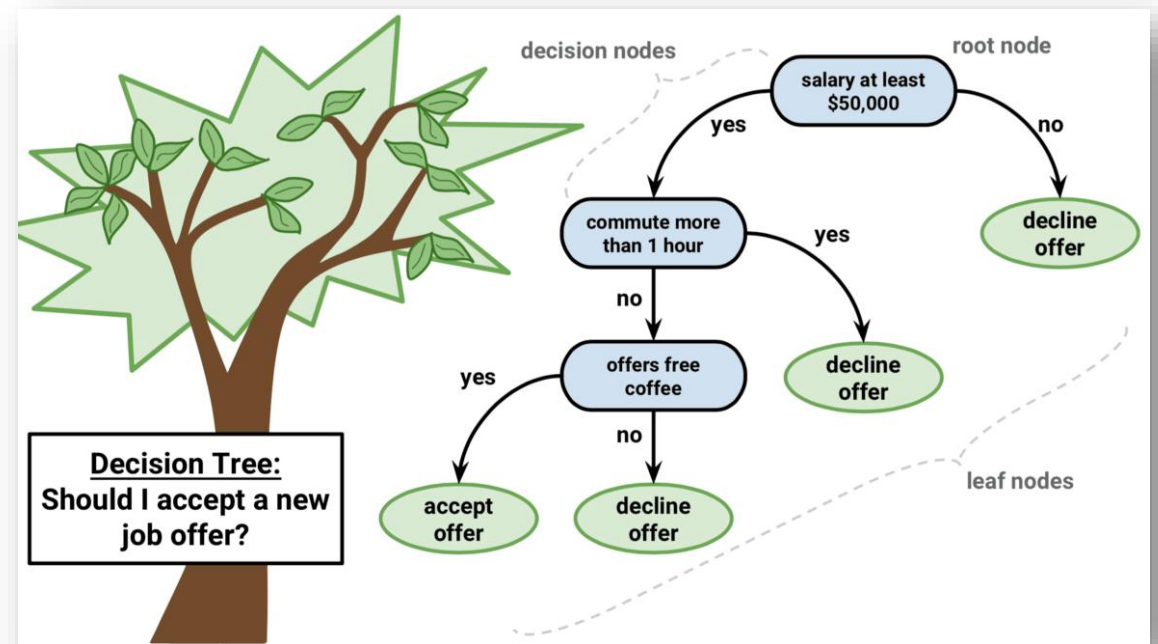
# Árvore de Decisão



**UFAL**  
CAMPUS ARAPIRACA  
U. E. PENEDO

# Árvore de Decisão

- Árvores são **estruturas de dados** formadas por um conjunto de elementos:
  - **Raiz:** característica com mais ligação para outros elementos.
  - **Nós:** informações que ligam a outros elementos.
  - **Folhas:** informações terminais que não possuem ligações.
- Uma árvore de decisão é um **algoritmo** que **armazena um conjunto de regras**, em que suas folhas representam a decisão a ser tomada.



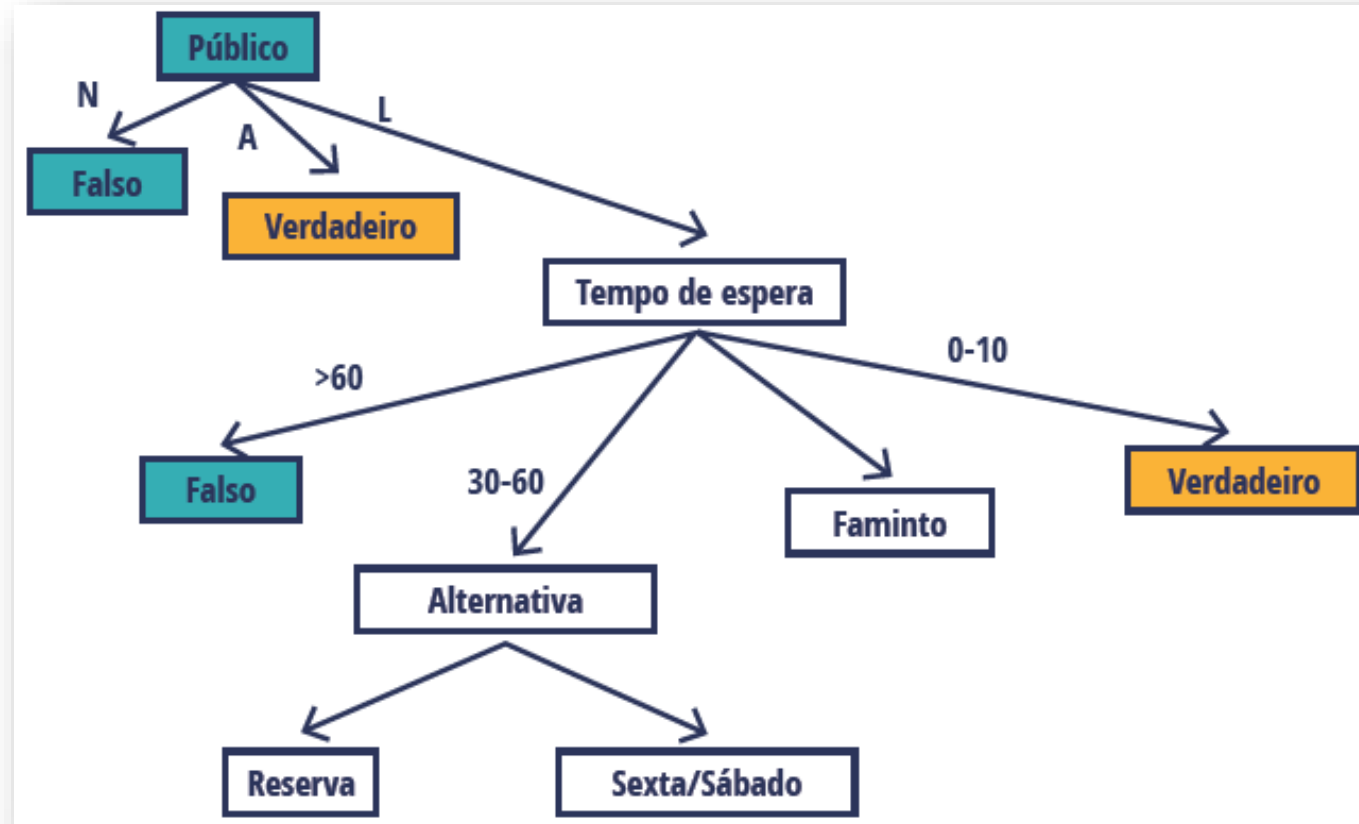
# Exemplo de Entrada de uma Árvore

— — — —

- **Decisão:**
  - devo esperar por uma mesa em um restaurante?
- **Entradas:**
  1. **Alternativa:** há um restaurante alternativo na redondeza?
  2. **Bar:** existe um bar confortável onde se esperar?
  3. **Sexta/Sábado:** hoje é sexta ou sábado?
  4. **Faminto:** estou com fome?
  5. **Público:** número de pessoas no restaurante (**Nenhuma (N)**, **Algumas (A)**, **Lotado (L)**).
  7. **Preço:** faixa de preços (R\$, R\$\$, R\$\$\$).
  8. **Chovendo:** está a chover?
  9. **Reserva:** temos reserva?
  10. **Tipo:** tipo do restaurante (francês, italiano, tailandês).
  11. **Tempo de espera:** tempo de espera estimado (0-10, 10-30, 30-60, >60).

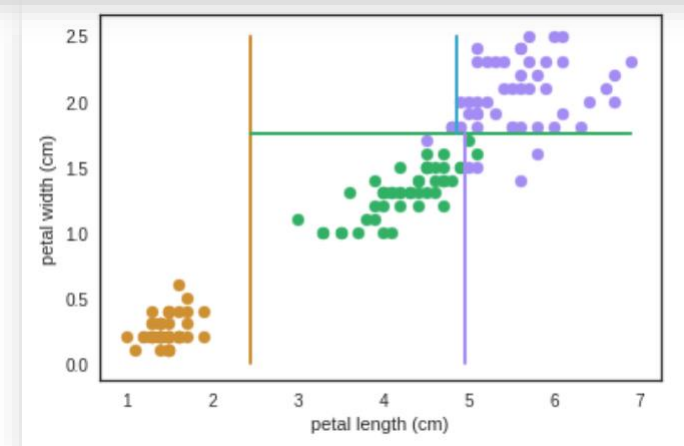
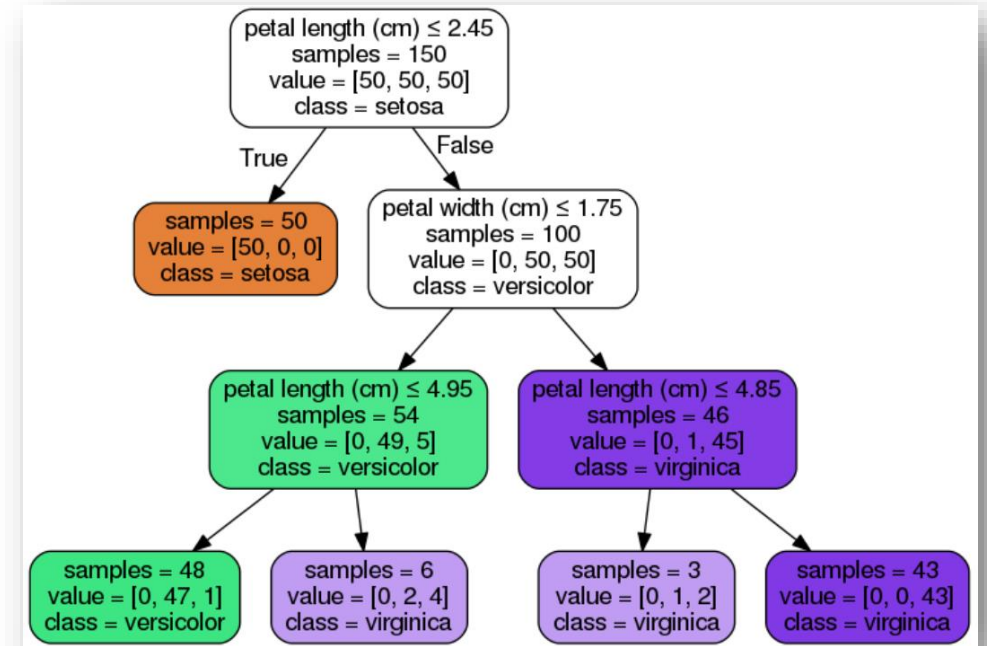


# Exemplo de Entrada de uma Árvore



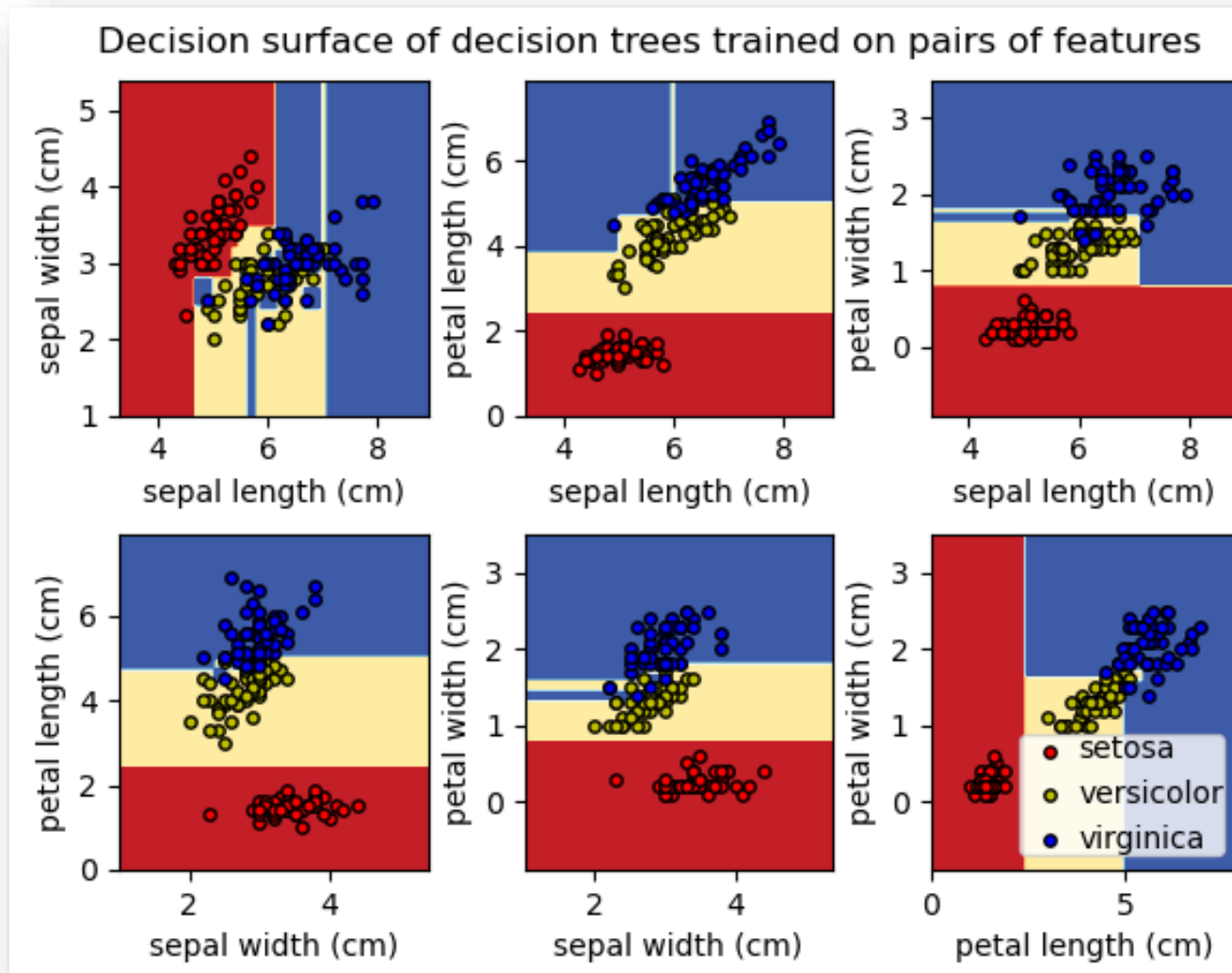
# Algoritmo ID3

- O algoritmo ID3 (**inductive decision tree**) é dos **mais utilizados** para a construção de árvores de decisão. Este algoritmo segue os seguintes passos:
  1. Recebe os dados de treinamento.
  2. Escolhe o atributo que melhor divide os exemplos.
  3. Cria nos filhos para cada valor possível ao atributo.
  4. Transporta os exemplos para cada nó filho.
  5. Repete o procedimento até que todos os exemplos sejam comportados.
- Como **saber qual o melhor atributo a escolher?**
  - **Entropia:** mede a aleatoriedade de uma variável.
  - **Ganho:** mede a redução da entropia.



# Fronteiras de Decisão de uma Árvore

-----



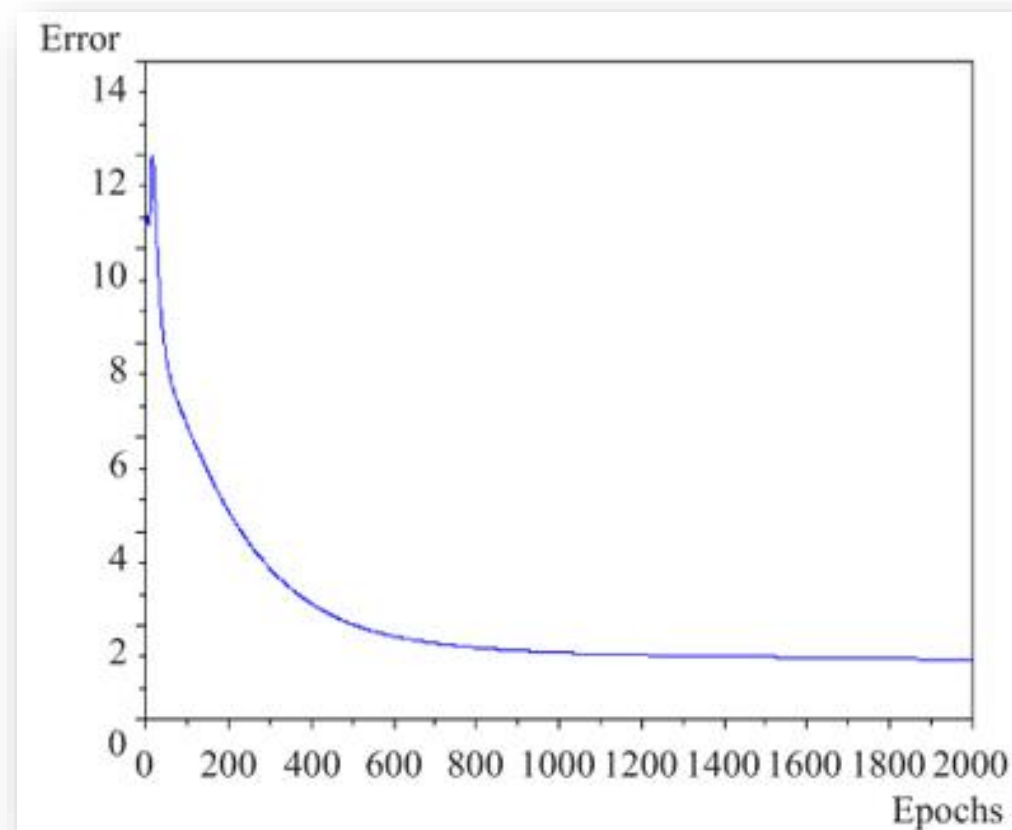
# Aprendizado Sobre os Dados



**UFAL**  
CAMPUS ARAPIRACA  
U. E. PENEDO

# Treinamento

- O treinamento consiste em realizar o ajuste de pesos das camadas intermediárias baseado nos tipos de entradas e saídas recebidas.
- Durante a etapa de treinamento, é realizado o **cálculo de erro**. O erro verifica o quanto próximo a rede está do resultado esperado. Logo, o objetivo do treinamento é **minimizar o erro**.

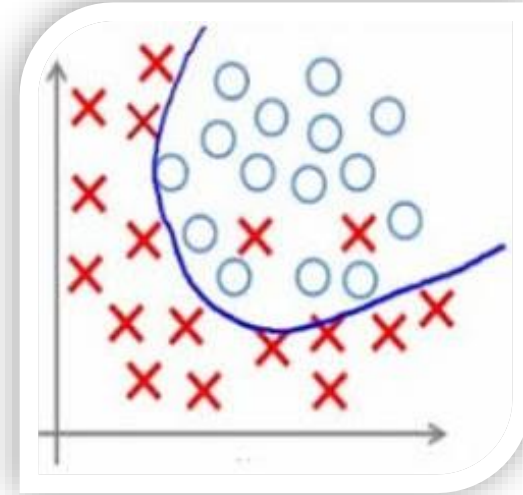




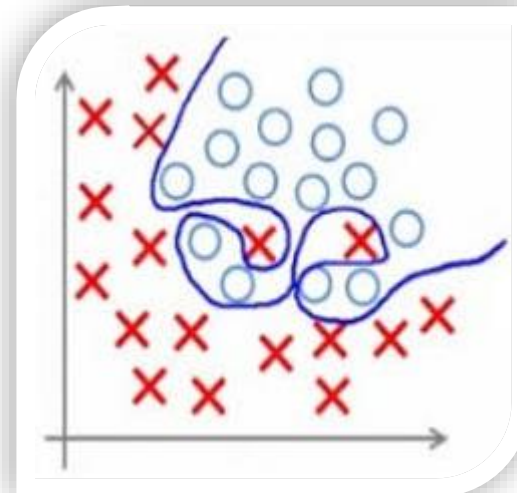
# Cuidados no Treinamento

- **Generalização:** consegue replicar bem o conhecimento aprendido.
- **Overfitting:** decora o treinamento e erra o teste.
- **Underfitting:** não aprende devidamente os padrões.

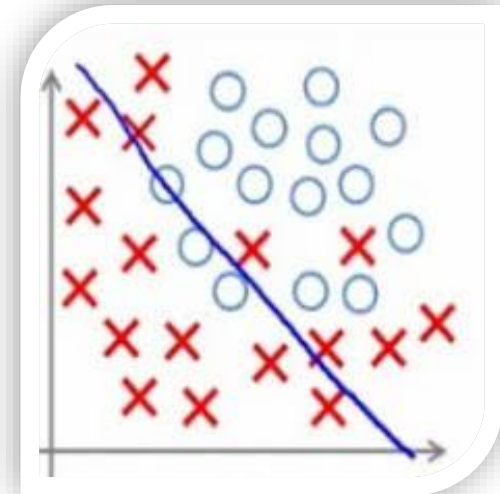
Generalização



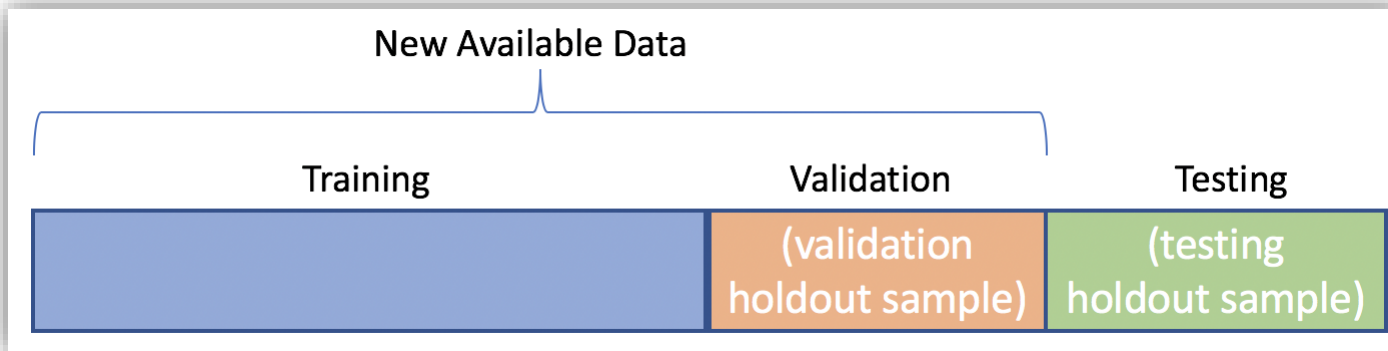
Overfitting



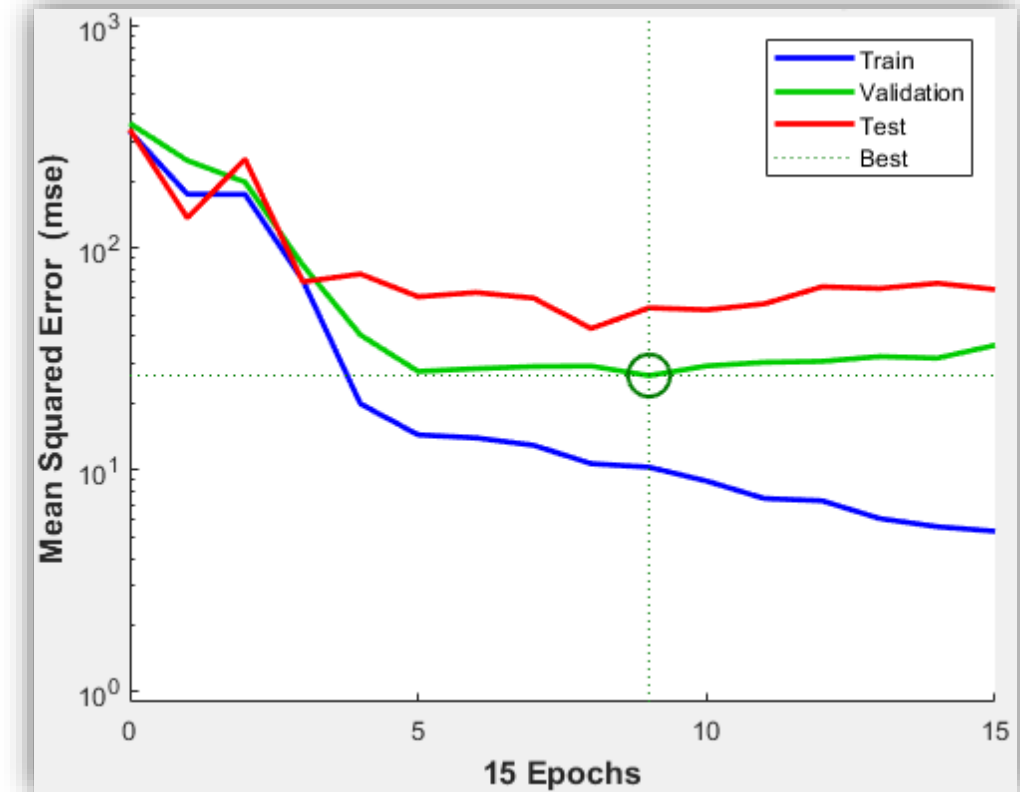
Underfitting



# Divisão dos Dados para Treinamento



- **Treinamento:** utiliza-se os dados para **ensinar** a rede os padrões existentes.
- **Validação:** utilizado **APENAS** para avaliar o desempenho da rede **DURANTE** o treinamento em dados não treinados (**Generalização**).
- **Teste:** **simula uma situação real**, em que a rede seria utilizada como uma ferramenta de tomada de decisão.

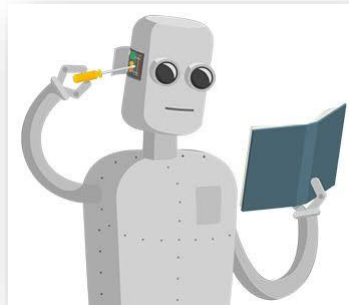


- **Critério de parada:** aumento do erro no conjunto de validação.

# Classificação de Novos Exemplos

----

- Após ser treinado, o sistema computacional **recebe uma observação sem classe** e estima a classificação para esse exemplo.



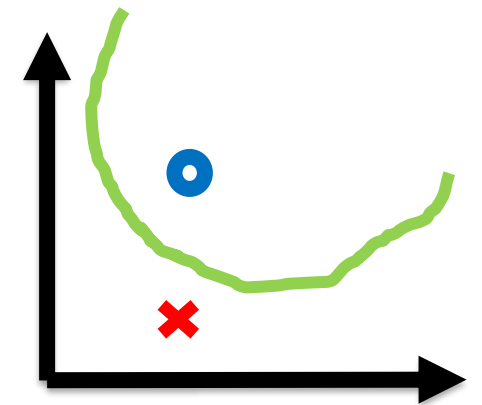
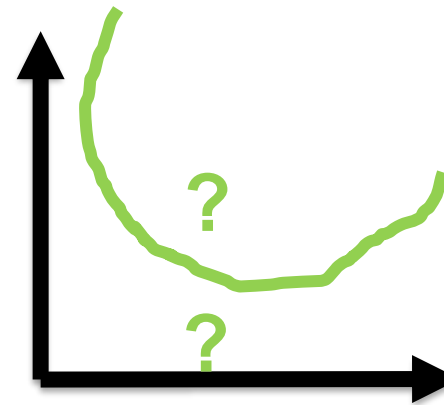
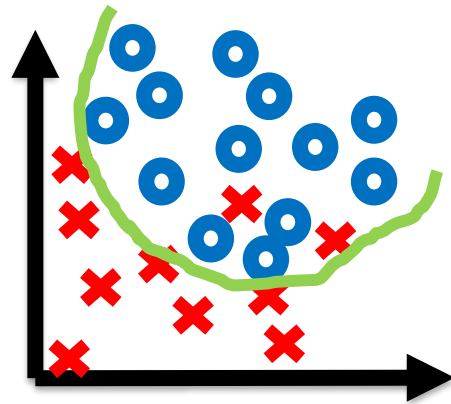
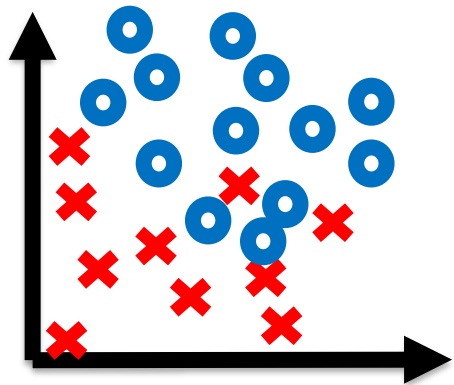
**Subespécie:  
Setosa**

- As classificações são utilizadas para as **tomadas de decisões** de acordo com os requisitos da empresa/usuário.

# Classificação de Novos Exemplos

-----

1) Dados de treinamento. 2) Aprendizado sobre os dados. 3) Recebimento de novos dados. 4) Classificação de novos dados.



# Avaliação do Resultado

-----

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$



# Introdução a Aprendizagem de Máquina Utilizando a Linguagem Python

**Profº Gustavo Miranda**  
**[gustavo.oliveira@penedo.ufal.br](mailto:gustavo.oliveira@penedo.ufal.br)**



**UFAL**  
CAMPUS ARAPIRACA  
U. E. PENEDO