



UNIVASF

UNIVASF

FUNDAÇÃO UNIVERSIDADE FEDERAL  
DO VALE DO SÃO FRANCISCO



# Aprendizagem de Máquina

**Professor:** Rosalvo Ferreira de Oliveira Neto

**Disciplina:** Inteligência Artificial



# Tópicos

1. Definições
2. Tipos de aprendizagem
3. Paradigmas de aprendizagem
4. Modos de aprendizagem
5. Medida de desempenho
6. Conceitos Importantes
7. Exemplo

Um processo de raciocínio indutivo de maneira geral, é um raciocínio que parte do particular para o geral, da parte para o todo

Por exemplo, se notamos que:

Todos os pacientes com **Déficit de Atenção** atendidos em **2006** sofriam de **Ansiedade**

Todos os pacientes com **Déficit de Atenção** atendidos em **2007** sofriam de **Ansiedade**

...

Podemos inferir logicamente, que Todos os pacientes que sofrem de Déficit de Atenção também sofrem de Ansiedade

Isto pode **ser ou não verdade**, mas propicia uma boa generalização

# Aprendizado Indutivo

O objetivo do aprendizado indutivo é encontrar uma hipótese ( $h$ ) que concorde com um conjunto de exemplos.

Uma hipótese é considerada satisfatória quando possui uma boa capacidade de **generalização**.

# Aprendizado de Máquina - uma definição

"Um programa aprende a partir da experiência **E**, em relação a uma classe de tarefas **T**, com medida de desempenho **P**, se seu desempenho em **T**, medido por **P**, melhora com **E**".

Mitchell, 1997

# Aprendizado de Máquina - Exemplo

Ex.: Detecção de bons clientes para um cartão de crédito

**Tarefa T:** classificar potenciais novos clientes como bons ou maus pagadores

**Medida de Desempenho P:** porcentagem de clientes classificados corretamente

**Experiência de Treinamento E:** uma base de dados histórica em que os clientes já conhecidos são previamente classificados como bons ou maus pagadores

# Aprendizado de Máquina - Exemplo

Ex.: Diagnóstico de gravidez de risco

**Tarefa T:** classificar novas gestantes com potenciais riscos na gravidez

**Medida de Desempenho P:** porcentagem de pacientes classificadas corretamente

**Experiência de Treinamento E:** base de dados histórica contendo exemplos de gestantes com ou sem gravidez de risco

- ✓ **Supervisionado**
- ✓ **Não Supervisionado**

## Aprendizado Supervisionado

- O algoritmo de aprendizado (indutor) recebe um conjunto de exemplos de treinamento para os quais os rótulos da classe associada são conhecidos
- Cada exemplo (instância ou padrão) é descrito por um vetor de valores (atributos) e pelo rótulo da classe associada
- O objetivo do indutor é construir um classificador que possa determinar corretamente a classe de novos exemplos ainda não rotulados

## Aprendizado Não Supervisionado

- O indutor analisa os exemplos fornecidos e tenta determinar se alguns deles podem ser agrupados de alguma maneira, formando agrupamentos ou clusters
- Após a determinação dos agrupamentos, em geral, é necessário uma análise para determinar o que cada agrupamento significa no contexto do problema que está sendo analisado

# Paradigmas de Aprendizado de Máquina

- ✓ **Simbólico**
- ✓ **Estatístico**
- ✓ **Baseado em Exemplos**
- ✓ **Conexionista**
- ✓ **Evolutivo**

# Paradigmas de Aprendizado de Máquina

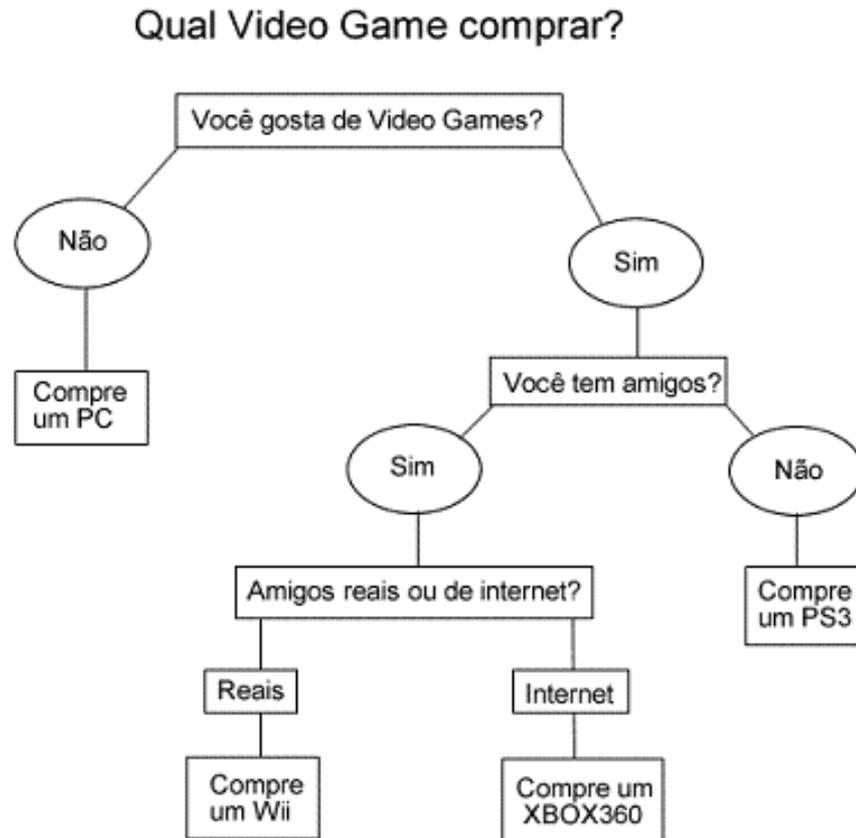
## ✓ Simbólico

Neste paradigma um conceito é representado em uma estrutura simbólica, e o aprendizado é realizado através da apresentação de exemplos e contra-exemplos deste conceito. As estruturas simbólicas estão tipicamente representadas em alguma expressão lógica, como por exemplo: regras de produção.

Exemplos de técnicas que utilizam este paradigma são: Agentes Inteligentes e Árvores de Decisão.

# Paradigmas de Aprendizado de Máquina

## ✓ Simbólico



# Paradigmas de Aprendizado de Máquina

## ✓ Estatístico

No paradigma estatístico é utilizado um modelo estatístico que encontre uma hipótese que possua uma boa aproximação do conceito a ser induzido. O aprendizado consiste em encontrar os melhores parâmetros para o modelo. Estes modelos podem ser paramétricos (quando fazem alguma suposição sobre a distribuição dos dados, ou podem ser não paramétricos, quando não fazem suposição sobre a distribuição dos dados.

Dentre os modelos estatísticos utilizados em aprendizagem de máquina podemos destacar os modelos Bayesianos.

# Paradigmas de Aprendizado de Máquina

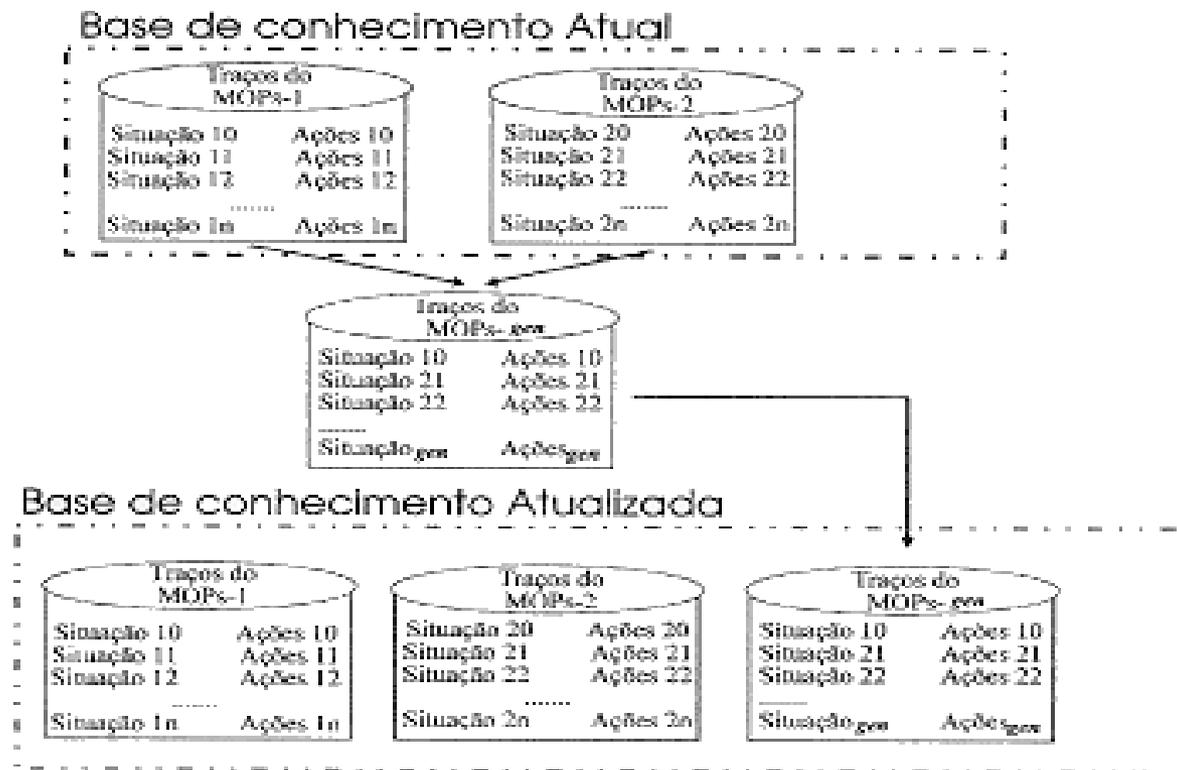
## ✓ Baseado em Exemplos

Uma forma de classificar um novo padrão é lembrar-se de exemplos parecidos classificados anteriormente, e assim atribuir ao novo exemplo uma classe de um padrão parecido. Esta é a idéia central deste paradigma.

A técnica de raciocínio baseado em casos é um exemplo de técnica que utiliza este paradigma.

# Paradigmas de Aprendizado de Máquina

## ✓ Baseado em Exemplos



# Paradigmas de Aprendizado de Máquina

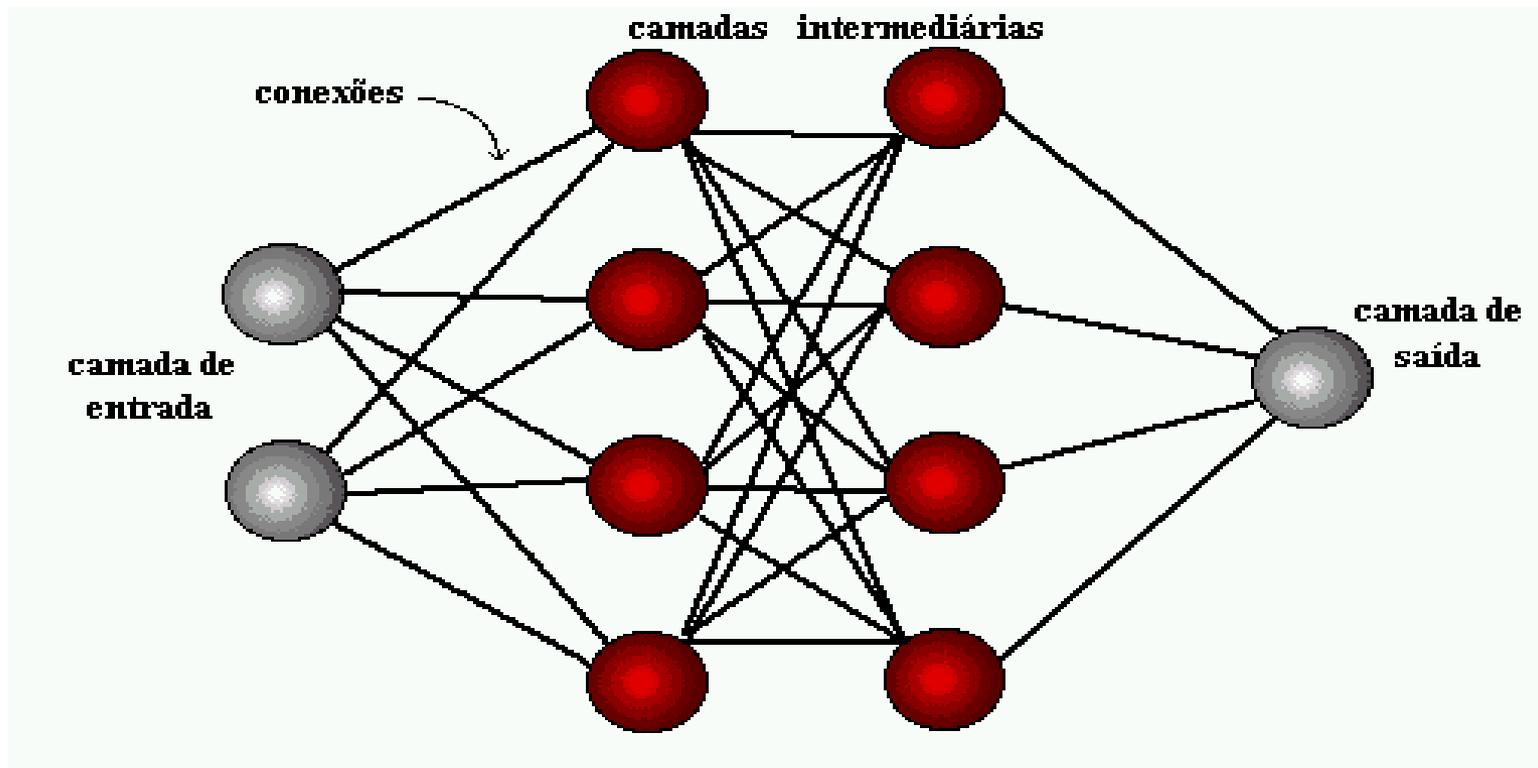
## ✓ **Conexionista**

O nome conexionista vem da área de pesquisa de Redes Neurais Artificiais (RNA). Uma rede neural artificial é um modelo computacional inspirado no funcionamento do cérebro humano. Uma RNA possui três componentes principais: unidade de processamento “os neurônios”, conexões “sinapses” e uma topologia. As redes neurais possuem como principal característica aprender através de exemplos e poder de generalização.

As redes Multi Layer Perceptron (MLP) e Self Organing Map (SOM) são exemplos de técnicas que utilizam este paradigma de aprendizado.

# Paradigmas de Aprendizado de Máquina

## ✓ Conexionista



# Paradigmas de Aprendizado de Máquina

## ✓ Evolutivo

Este paradigma foi inspirado na teoria da evolução das espécies de Charles Darwin. O algoritmo inicia com uma população de indivíduos, onde cada indivíduo representa uma possível solução. Os indivíduos competem entre si, os indivíduos com menor desempenho são descartados, e os indivíduos com melhores desempenho são selecionados para reprodução (Crossover), os novos indivíduos gerados podem ou não sofrer mutação. A população evolui através de várias gerações, até que uma solução ótima seja encontrada.

Algoritmos genéticos e Programação genética são exemplo de técnicas que utilizam este paradigma.

# Modos de Aprendizado de Máquina

## ✓ Incremental (On-Line)

- A hipótese é atualizada a medida que cada exemplo é apresentado ao indutor.

## ✓ Não Incremental (Batch)

- A hipótese só é atualizada após a apresentação de todos os exemplos ao indutor.

# Medida de Desempenho

A medida de desempenho natural de um indutor é sua taxa de erro global, porém é **otimista!**

## Solução:

- Treinamento
- Teste

# Conceitos Importantes

- ✓ Exemplo (caso, registro ou padrão) é uma tupla de valores de atributos
  - Um cliente, uma paciente
  
- ✓ Atributo: descreve uma característica ou um aspecto de um exemplo.
  - Nominal: sexo, estado civil
  - Contínuo: peso, altura, idade
  
- ✓ Classe: atributo especial (aprendizado supervisionado), denominado rótulo ou classe.

## Conceitos Importantes

- ✓ Conjunto de exemplos: Um conjunto de exemplos é composto por exemplos contendo valores de atributos bem como a classe associada.

Sexo	Idade	Renda	Estado Civil	Classe
Masculino	18	R\$ 1.000,00	Solteira	MAU
Feminino	21	R\$ 800,00	Solteira	MAU
Masculino	50	R\$ 2.000,00	Casado	BOM
Feminino	60	R\$ 8.000,00	Casado	BOM

# Conceitos Importantes

## ✓ Missing Values

Em geral, indicados por valores ausentes;

### **Tipos:**

- Desconhecidos
- Não registrados

### **Razões**

- Erro de funcionamento
- Mudanças na definição do escopo

# Conceitos Importantes

- ✓ Outliers (Pontos aberrantes)

Em geral, indicados por valores fora do padrão.

Exemplo rendas com valores R\$ 1,00 ou R\$ 1.000.000,00

# Exemplo

**RUSSEL, S.; NOVIG, P. Inteligência Artificial. Elsevier, 2004**

- Capítulo 18

18.1 Formas de aprendizagem

18.2 Aprendizagem indutiva

18.4 Aprendizagem por agrupamento

**Inteligência Artificial. Estruturas e Estratégias para a Solução de Problemas Complexos**

- capítulo 9

- capítulo 10

**Witten, I. H. and Frank, E. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann Series in Data Management Sys. Morgan Kaufmann, 2005.**

Capitulo 1

1.4 Machine learning and statistics

1.5 Generalization

- Capitulo 2- Input: Concepts, instances, and attributes



UNIVASF

FUNDAÇÃO UNIVERSIDADE FEDERAL  
DO VALE DO SÃO FRANCISCO