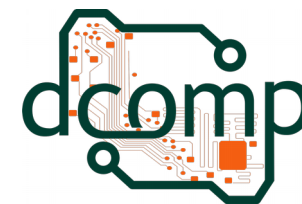




Universidade Federal do Espírito Santo
Centro de Ciências Agrárias – CCENS UFES
Departamento de Computação



Aprendizado de Máquina

Inteligência Artificial

Site: <http://jeiks.net>

E-mail: jacsonrcsilva@gmail.com

Aprendizado

- Qual sua importância para as espécies?
- O que ele permite aos animais?
- Qual sua importância em sistemas inteligentes?
- O que ele permite aos sistemas inteligentes?
- Onde pode ser utilizado?

O que é Aprendizado?

- “Memorizar/Abrigar/Armazenar” informações;
- Aprender fatos por meio de **observação e exploração**;
- Melhorar habilidades motoras/cognitivas por meio de **prática**;
- Organizar novo conhecimento em **representações** efetivas e gerais;

Aprendizado de Máquina

- Um sistema deve conseguir aprender sobre um determinado conjunto de informações.
- Para isso, é realizado:
 - **O Treinamento:**
 - Trabalho realizado sobre um conjunto inicial de entrada (e geralmente, das saídas desejadas), com intuito de *produzir* resultados corretos.
 - Aplicação do conhecimento aprendido:
 - Capacidade de **produzir** a *saída esperada* sobre os dados treinados; e
 - Capacidade de **generalizar** os dados do treinamento, tornando-se capaz de produzir uma *resposta esperada* para valores ainda não observados.

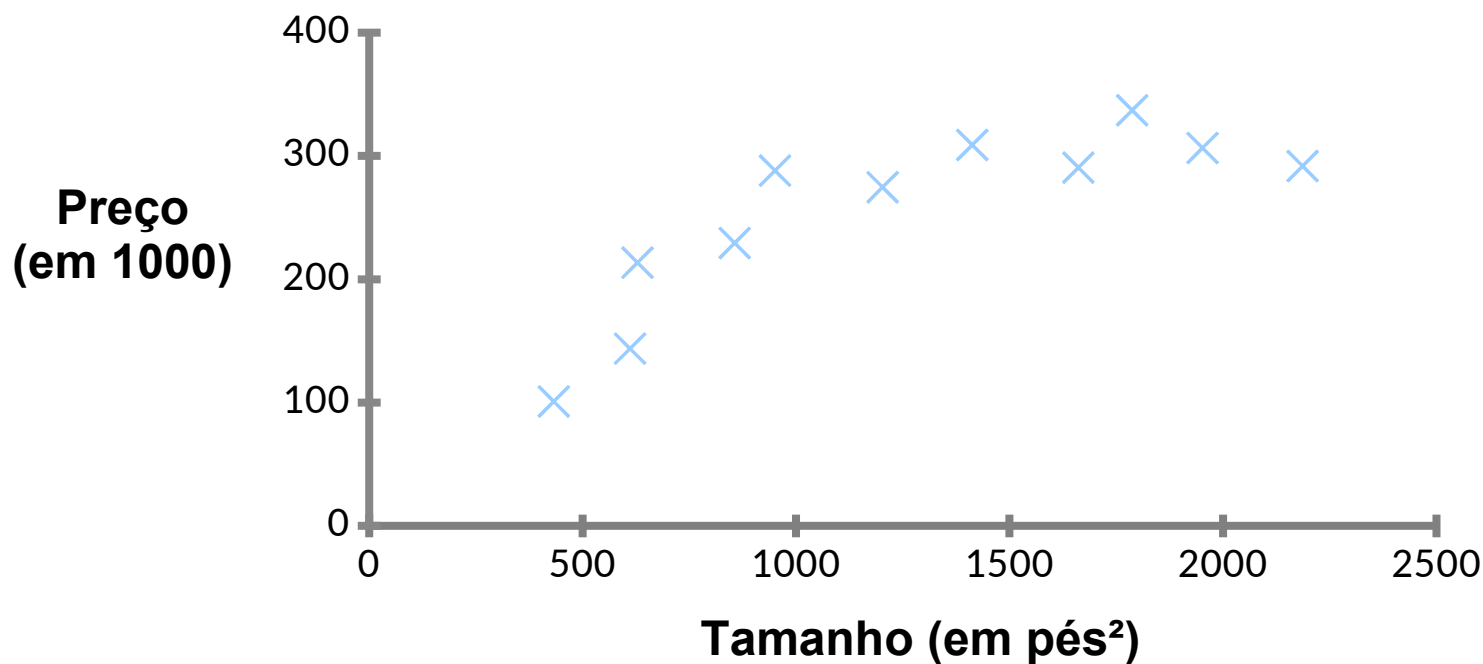
Treinamento

- Para maioria dos problemas de aprendizado:
 - A tarefa é aprender a classificar entradas de acordo com um conjunto finito (às vezes, infinito) de dados.
- Tipicamente, o sistema possui inicialmente um conjunto de dados classificados manualmente.
- O sistema tenta então aprender a partir desses dados de treinamento:
 - Fornecendo o resultado certo para estes dados; e também
 - Fornecendo resultados à novos dados ainda não observados (através de generalização).

Treinamento

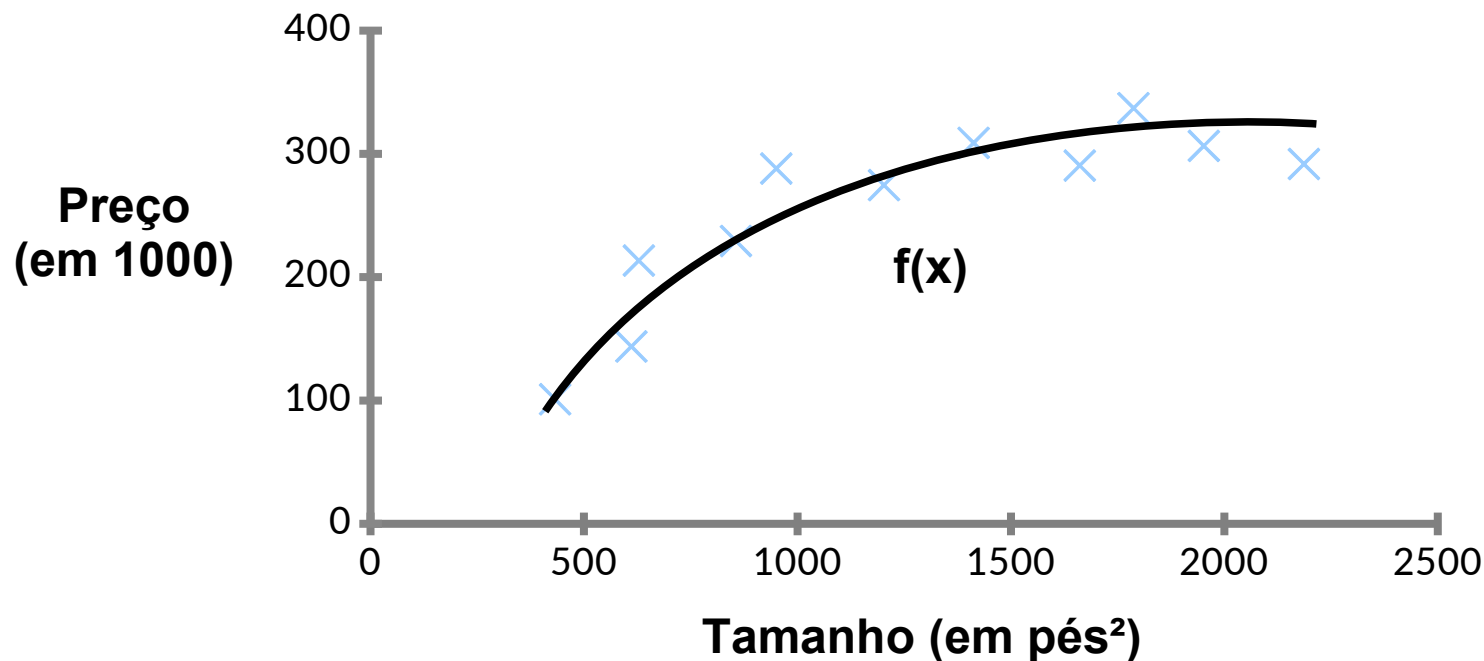
- Espera-se que haja relação entre os dados de entrada e suas classificações.
- Assim, existirá uma função que poderá ser gerada dos dados (x) para pertencer a classificação (y):

$$f(x) = y$$



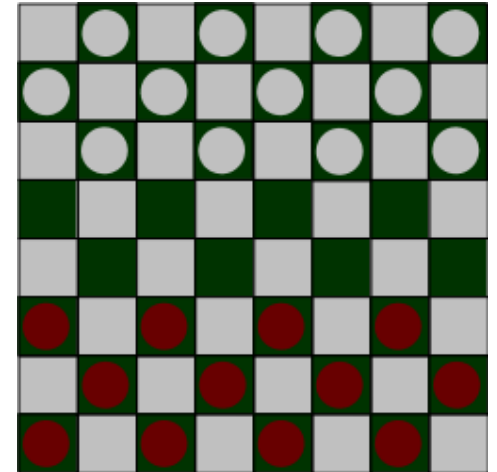
Regressão

- A regressão consiste no resultado da função $f(x)$, que forneça valores próximos aos resultados já conhecidos.



Definição de Aprendizado de Máquina

- Arthur Samuel (1959):
 - Campo de estudo que dá aos computadores a capacidade de aprender sem ser explicitamente programado.
 - Tornou-se famoso nos anos 50 quando escreveu um programa para jogar damas.
- Tom Mitchell (1998):
 - Um problema de aprendizagem bem colocado é:
Um programa aprende a partir da experiência **E**,
em relação a uma tarefa **T**,
com alguma medida de desempenho **P**,
se seu desempenho em **T**, medido por **P**, melhora com **E**.



Tipos de Aprendizado de Máquina

- Principais classificações:
 - Supervisionado:
 - Apresenta-se um conjunto de entradas, onde cada padrão de entrada possui sua saída. A saída corresponde às classes desejadas.
 - Exemplos de técnicas utilizadas: Redes Neurais Artificiais supervisionadas; Algoritmos Genéticos; Árvores de Decisão.
 - Não Supervisionado:
 - Apenas as entradas são fornecidas. O padrão de saída deve ser deduzido pelo próprio sistema de aprendizado.
 - Exemplos de técnicas utilizadas: Redes Neurais Artificiais não supervisionadas; Algoritmos de Agrupamento; Regras de Associação.
 - Por reforço:
 - O aprendizado ocorre com a interação com o ambiente.
 - Os algoritmos devem transformar as situações do mundo em ações, de maneira que maximize sua medida de desempenho.
 - O agente deve explorar o ambiente e descobrir quais ações oferecem mais recompensas ou punições.

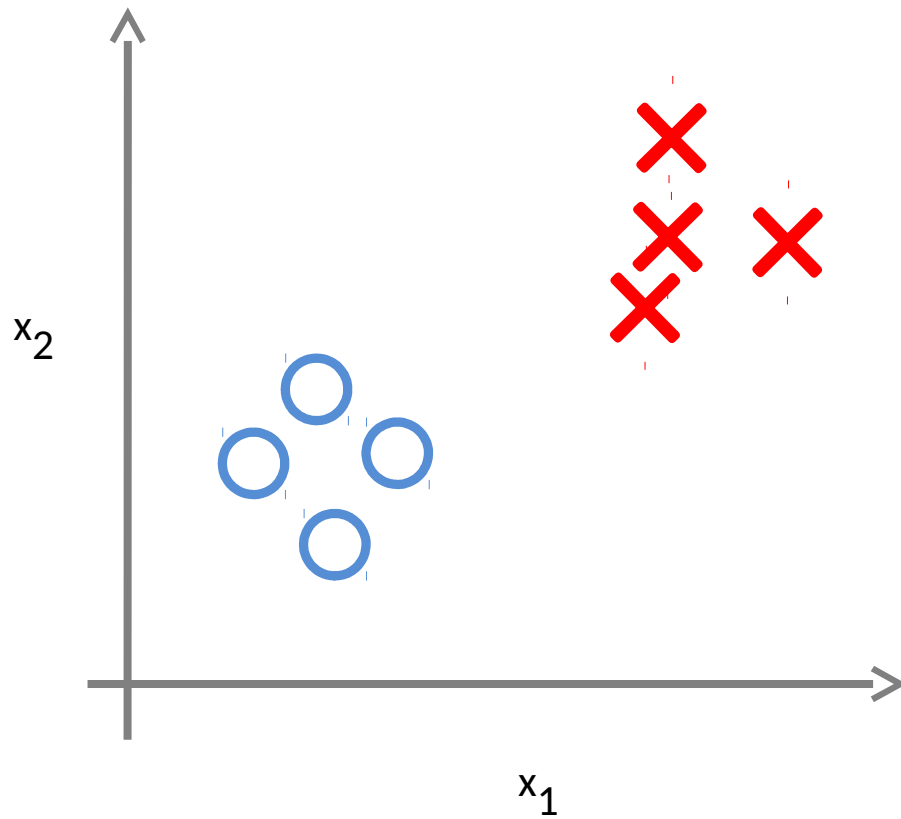
Aprendizado Supervisionado

- O *algoritmo de aprendizado* recebe um **conjunto** de exemplos de **treinamento**.
- Cada exemplo é descrito por um **padrão**, composto por um vetor de valores e pelo rótulo da classe associada. Ex.:
 - Entradas: [0,5 ; 0,7 ; 1,0] Rótulo (Saída): 1 (classe A)
- O objetivo é *construir* um **classificador** que possa determinar corretamente a classe de novos exemplos ainda não rotulados.
- Para rótulos de classe *discretos*:
 - Ocorre uma **classificação**; e
- Para valores *contínuos*,
 - Ocorre uma **regressão**.

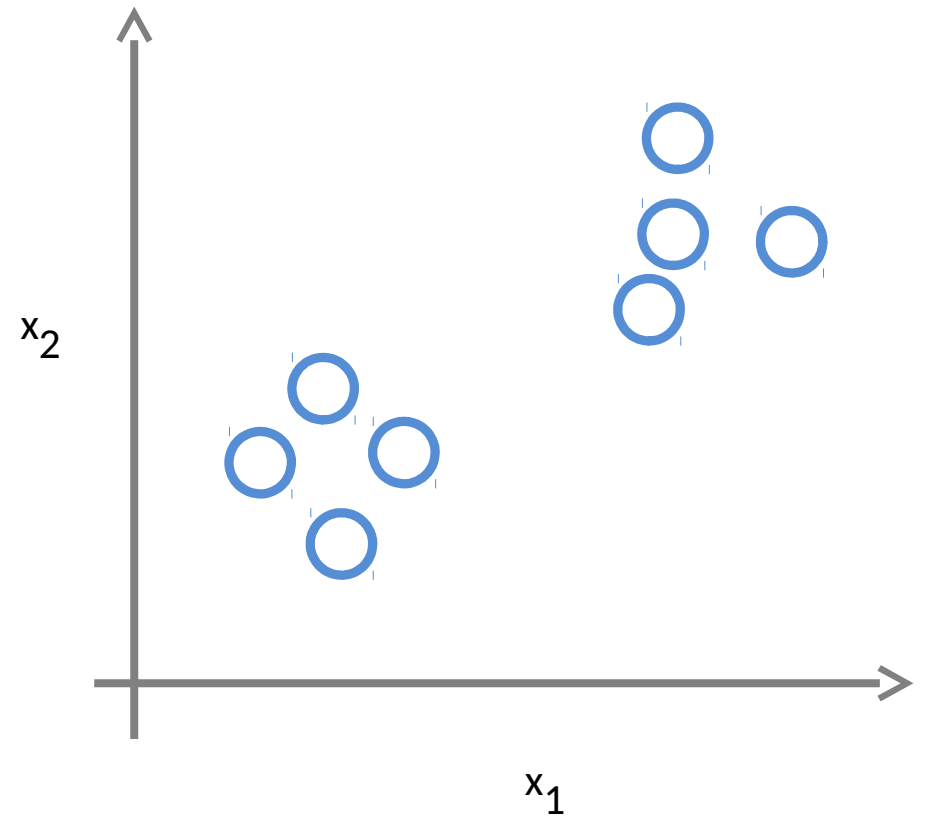
Aprendizado Não Supervisionado

- O *algoritmo de aprendizado* também recebe um **conjunto** de exemplos de **treinamento**.
- Cada exemplo é composto por um vetor de valores de entrada, porém **sem** o rótulo da classe associada. Ex.:
 - Entradas: [0,5 ; 0,7 ; 1,0] [0,5 ; 0,4 ; 1,2] [0,6 ; -1,0 ; 0,7]
- O objetivo é *construir* um **classificador** que possa criar agrupamentos, ou seja, **rotular** as entradas baseando-se nas similaridades de seus valores.
- O classificador analisa os exemplos fornecidos e tenta determinar se alguns deles podem ser agrupados de alguma maneira, formando agrupamentos ou *clusters*.
- Geralmente, após a determinação dos agrupamentos é necessário uma análise para determinar o que cada agrupamento significa no contexto problema sendo analisado.

Aprendizados

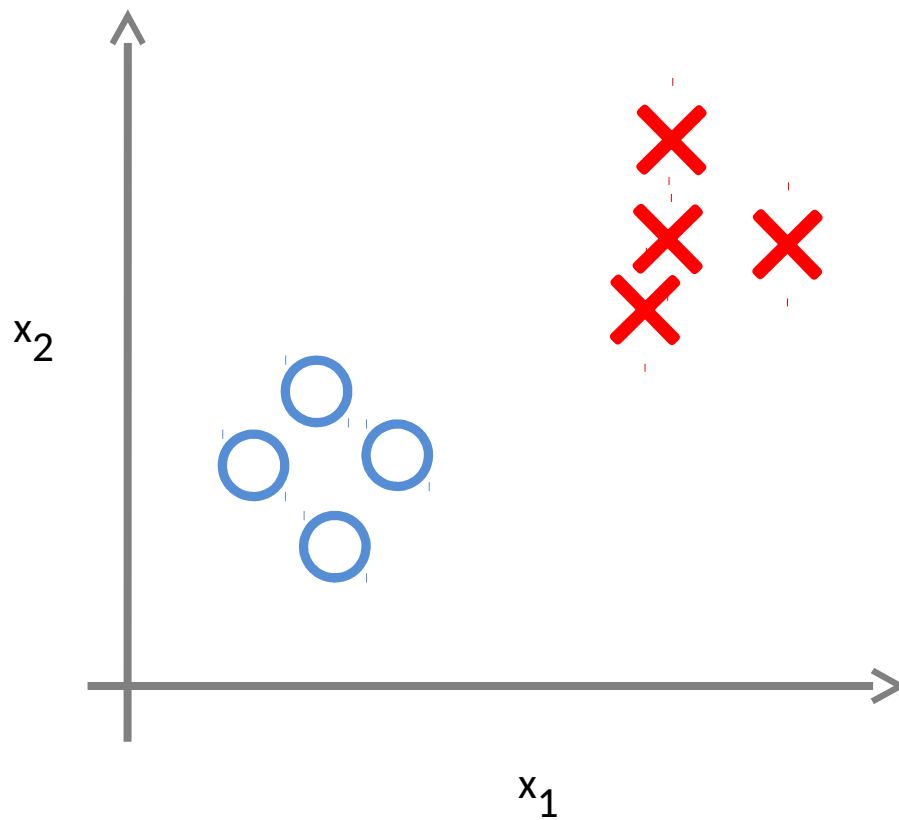


Supervisionado

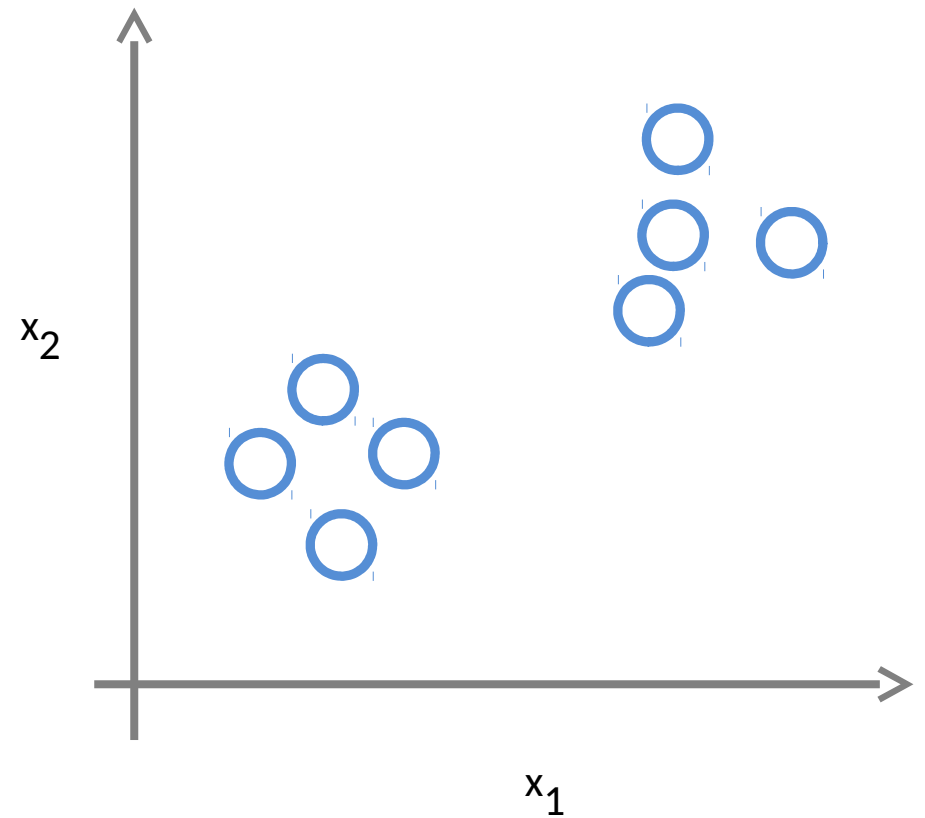


Não Supervisionado

Aprendizados



Supervisionado



Não Supervisionado

Esboce o resultado da função de classificação nos gráficos acima.

Grau de compreensibilidade

- Sistemas tipo caixa-preta
 - Sua representação interna não pode ser facilmente interpretada por humanos;
 - Não fornecem esclarecimento ou explicação do processo de reconhecimento.
- Sistemas orientados a conhecimento
 - Objetivam a criação de estruturas simbólicas que sejam compreensíveis por humanos.

Paradigmas de Aprendizado

- Simbólico;
- Estatístico;
- Baseado em Exemplos;
- Conexionista;
- Evolutivo.

Paradigmas de Aprendizado

- Simbólico
 - Buscam aprender construindo representações simbólicas de um conceito
 - Necessita de exemplos e contra-exemplos desse conceito para analisar
 - Estão tipicamente na forma de alguma expressão lógica, como:
 - Árvores de decisão;
 - Regras; ou
 - Rede semântica.

Paradigmas de Aprendizado

- Estatístico:
 - Utilizar métodos estatísticos para encontrar uma boa aproximação do conceito induzido;
- Baseados em exemplos:
 - Classificar exemplos nunca vistos por meio de exemplos similares conhecidos;
 - Sistema denominado lazy (preguiçoso);
 - Necessitam manter exemplos na memória para classificar novos exemplos.

Paradigmas de Aprendizado

- **Conexionista:**
 - Diretamente ligado às Redes Neurais;
 - Envolve unidades altamente interconectadas.
- **Evolutivo:**
 - Derivado do modelo biológico de aprendizado;
 - Consiste em:
 - Uma população de elementos de classificação que competem para fazer a predição;
 - Elementos com performance fraca são descartados;
 - Os elementos mais fortes proliferam, produzindo variações de si mesmos.

Exercício

- Sabendo que:
 - Temos os seguintes dados de entrada ($\Phi_1 = \{ x_1, x_2 \}$):

| | |
|---------------------------|----------------------------|
| $\Phi_1 = \{ 1 ; 0,8 \};$ | $\Phi_3 = \{ 1 ; -0,5 \};$ |
| $\Phi_2 = \{ 1 ; 1,4 \};$ | $\Phi_4 = \{ 1 ; -1,0 \};$ |
 - Temos os grupos:
 - Grupo 1 (Φ_1 e Φ_2): resultado da função (t) = -1;
 - Grupo 2 (Φ_3 e Φ_4): resultado da função (t) = +1;
 - Temos como função de classificação:

$$\mathbf{x}_1\omega_1 + \mathbf{x}_2\omega_2$$
 , onde $\omega_1 = 0,5$, e $\omega_2 = 1,0$.
 - Temos como correção:

$$\omega_{\text{novo}} = \omega_{\text{velho}} + \Phi_{\text{errado}} * t_{\text{desejado}}$$
- Faça os passos necessários para que os dados sejam corretamente classificados.