

REDES NEURAIS

Aprendizagem em redes neurais e
seus principais paradigmas



TÓPICOS

1. O que é aprendizagem?

2. Principais Técnicas

- 1. Correção de erros
- 2. Memória
- 3. Competitiva
- 4. Boltzman
- 5. Hebbiana

3. Paradigmas de aprendizagem



O QUE É APRENDIZAGEM?

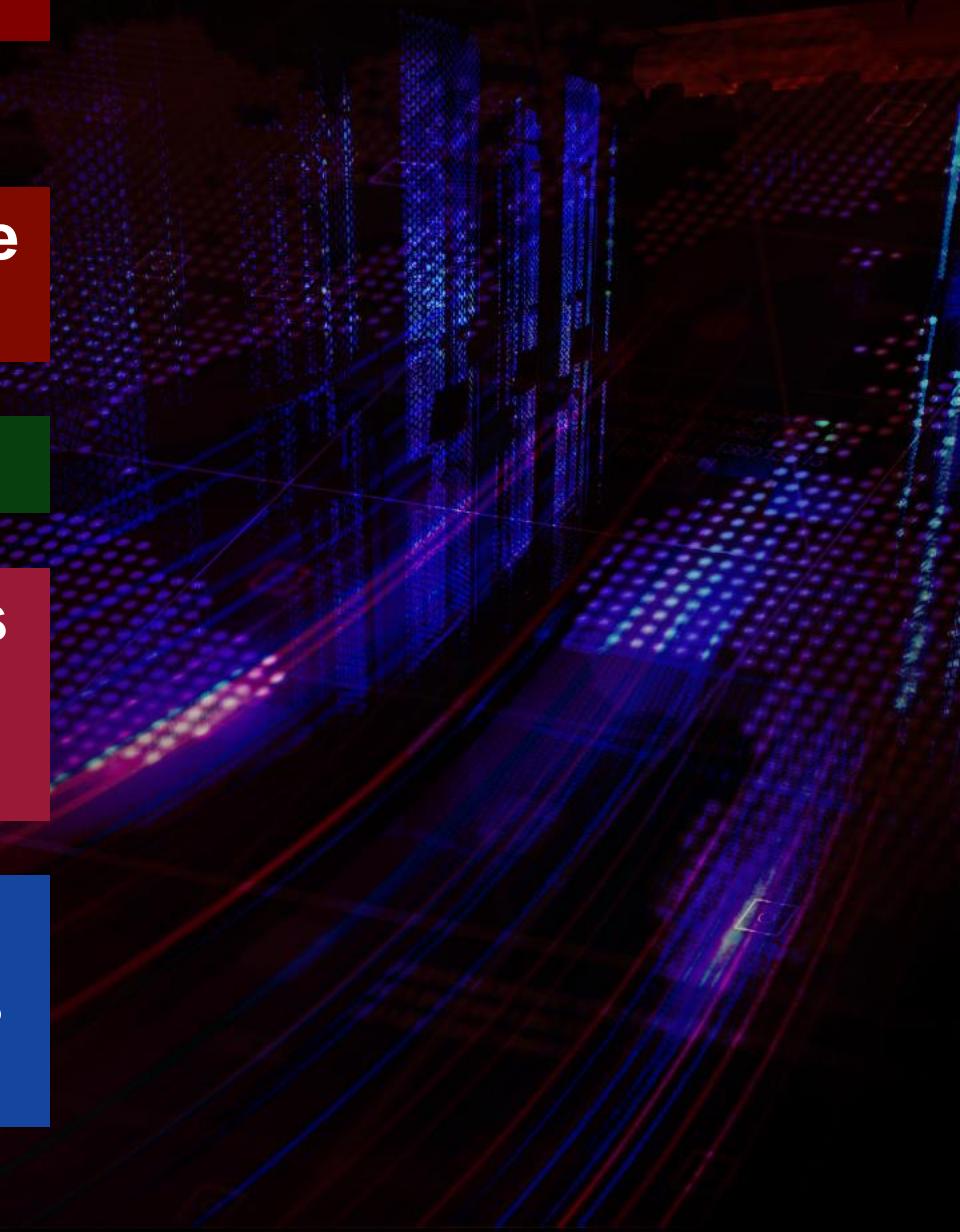
- Aprendizado é o processo pelo qual os **parâmetros** (os pesos das conexões entre os neurônios) e **hiperparâmetros** (taxas, topologia, etc.) de uma rede neural são ajustados através de uma forma continuada de estímulo pelo ambiente no qual a rede está operando
- Também denominado **Treinamento**



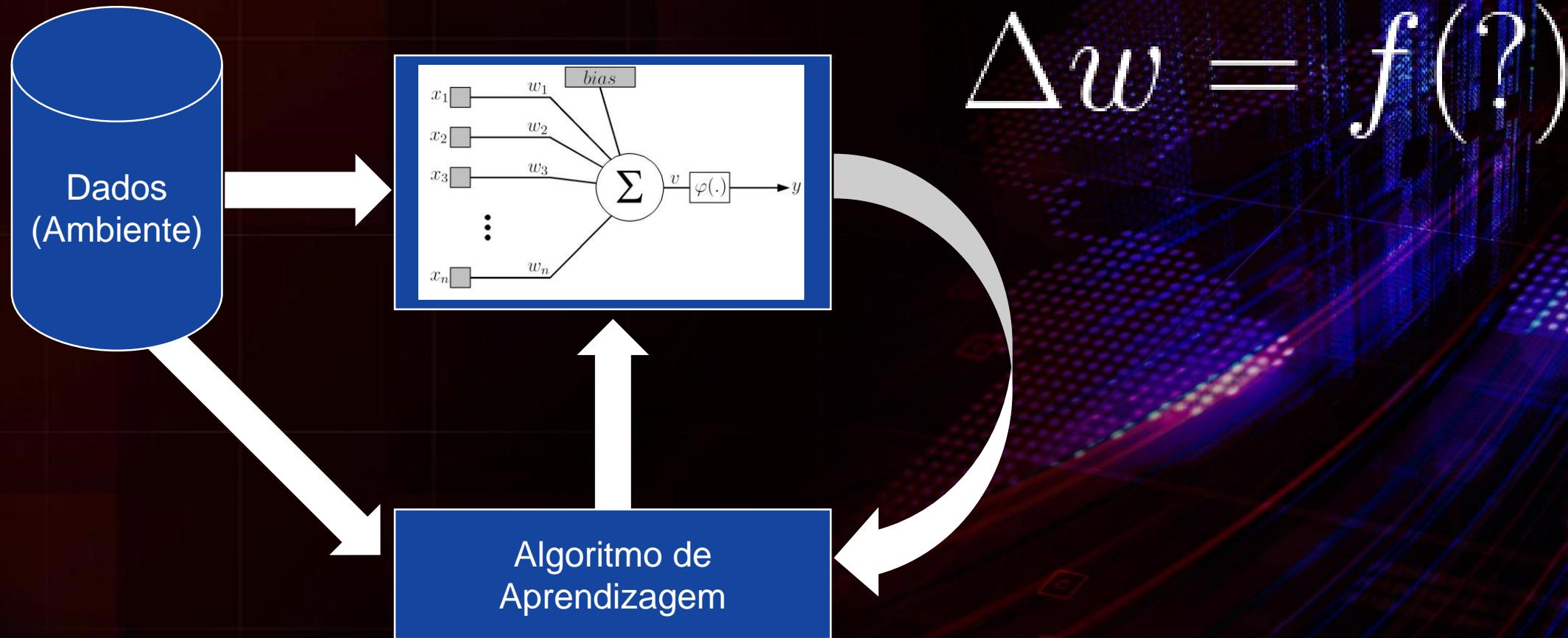
O QUE É APRENDIZAGEM?

O processo de aprendizado implica a seguinte sequência de eventos:

1. A rede neural é estimulada pelo ambiente.
2. A rede neural sofre modificações nos seus parâmetros livres como resultado deste estímulo.
3. A rede neural responde de uma maneira nova ao ambiente, devido às modificações ocorridas na sua estrutura interna.

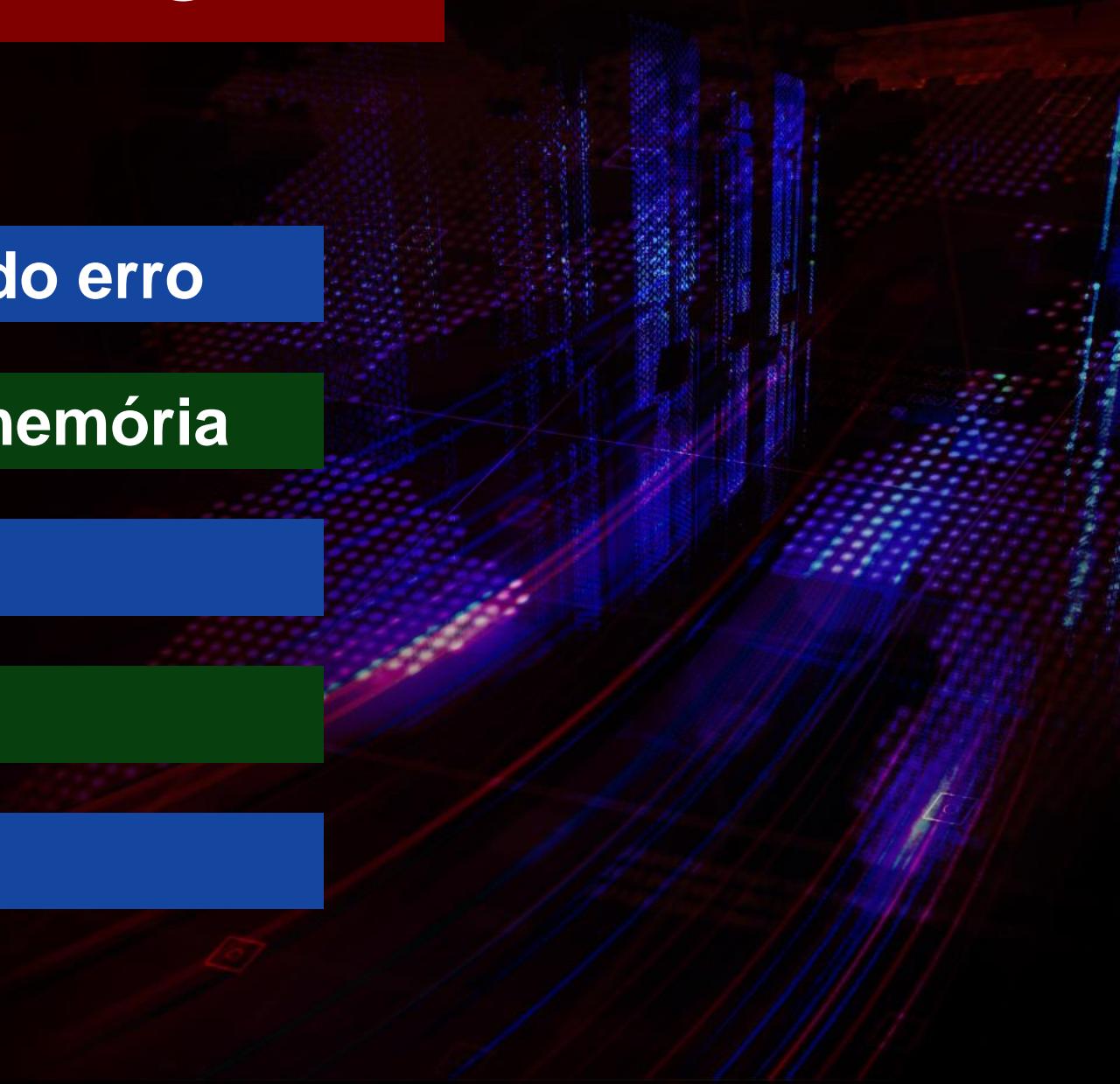


O QUE É APRENDIZAGEM?

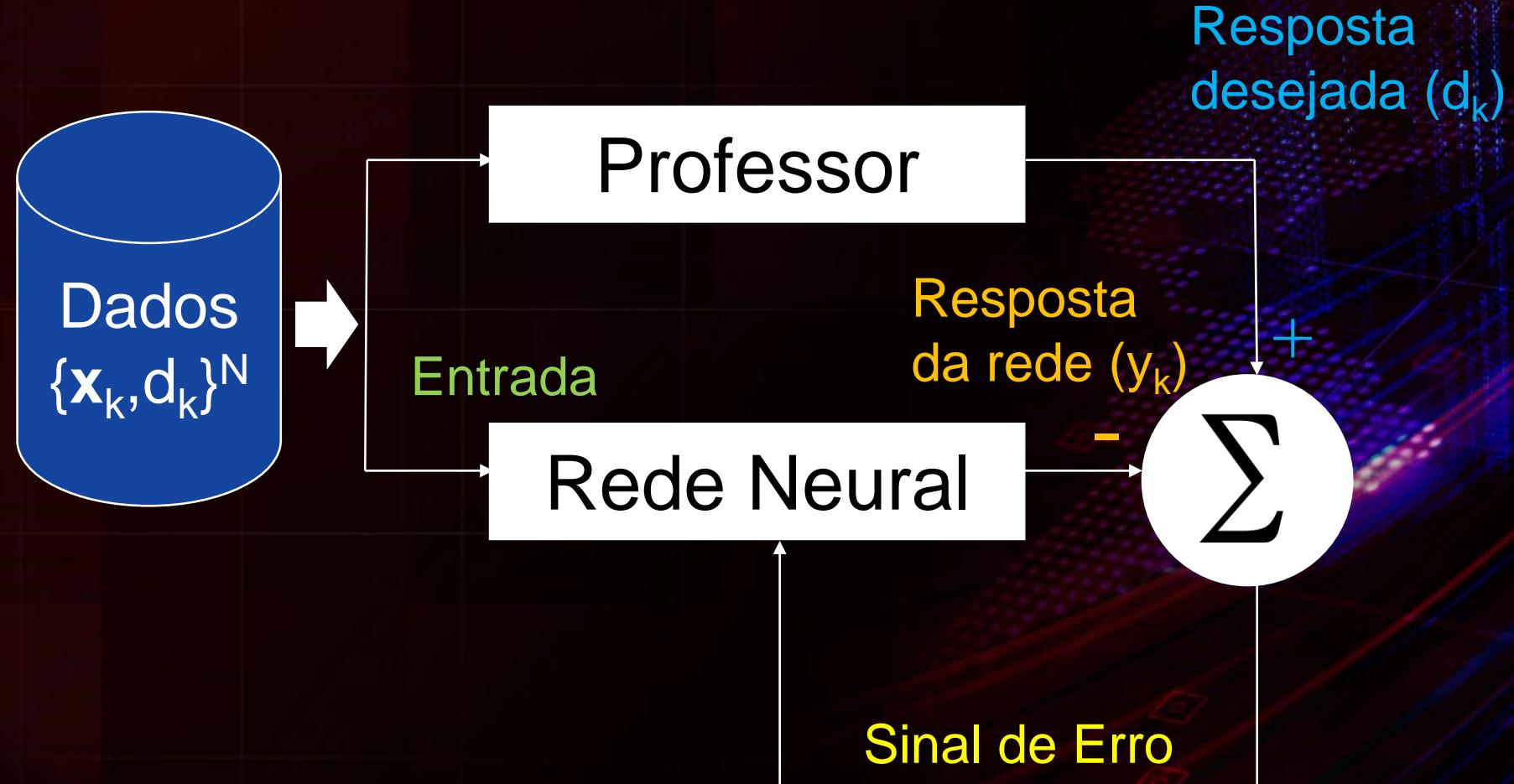


REGRAS DE APRENDIZAGEM?

- Aprendizado por correção do erro
- Aprendizado baseado em memória
- Aprendizado Hebbiano
- Aprendizado competitivo
- Aprendizado de Boltzmann



CORREÇÃO DE ERRO



CORREÇÃO DE ERRO

- A Resposta do modelo (y_k) é comparada à resposta desejada (d_k) e um sinal de erro é gerado:

$$e_k(n) = d_k(n) - y_k(n)$$

- A técnica tem por objetivo reduzir o erro ao longo do processo de treinamento



CORREÇÃO DE ERRO

- Diversos métodos podem ser usados para reduzir o erro. Comumente, utiliza-se o método do gradiente descendente (regra delta)

$$E(n) = \frac{1}{2} \sum (d_k(n) - y_k(n))^2$$

$$\Delta w_{kj} = -\eta \frac{\partial E}{w_{kj}} \quad \Delta w_{kj} = \eta e_k x_j$$

APRENDIZADO BASEADO EM MEMÓRIA

- Armazena parte das experiências existentes em uma memória:

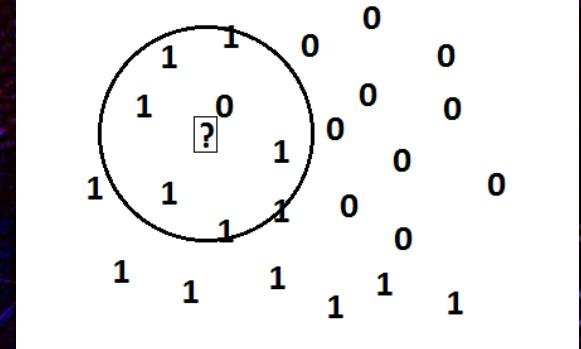
$$\{(\mathbf{x}_k, d_k)\}_{i=1}^N$$

- Para se classificar um novo dado $\mathbf{x}_{\text{teste}}$, consulta-se a memória disponível

- O exemplo é classificado segundo os exemplos contidos na vizinhança de $\mathbf{x}_{\text{teste}}$

APRENDIZADO BASEADO EM MEMÓRIA

- **Dois passos fundamentais:**
 1. Critério para definir a vizinhança de x_{teste}
 2. Regra de aprendizagem aplicada aos exemplos pertencentes à vizinhança



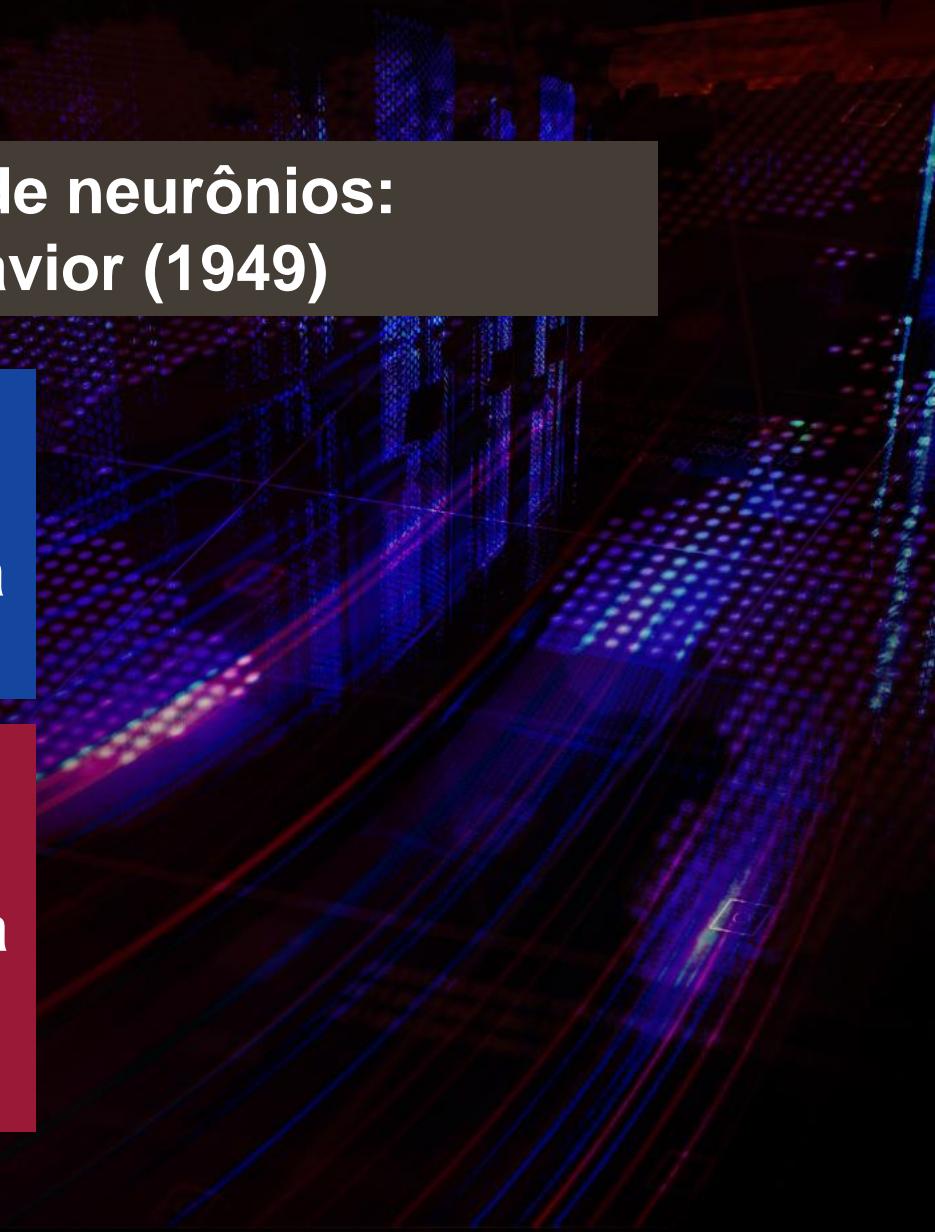
- **No caso mais simples: regra do vizinho mais próximo**

$$\min_i d(\mathbf{x}_i, \mathbf{x}_{teste}) = d(\mathbf{x}', \mathbf{x}_{teste}) \quad y_{teste} = y'$$

APRENDIZADO HEBBIANO

Primeiro estudo sobre aprendizado em redes de neurônios:
Postulado de Hebb – The Organization of Behavior (1949)

- Se dois neurônios em ambos os lados de uma sinapse (conexão) são ativados simultaneamente, então a força daquela sinapse é seletivamente aumentada
- Se dois neurônios em ambos os lados de uma sinapse são ativados assincronamente, então a força daquela sinapse é seletivamente enfraquecida ou eliminada

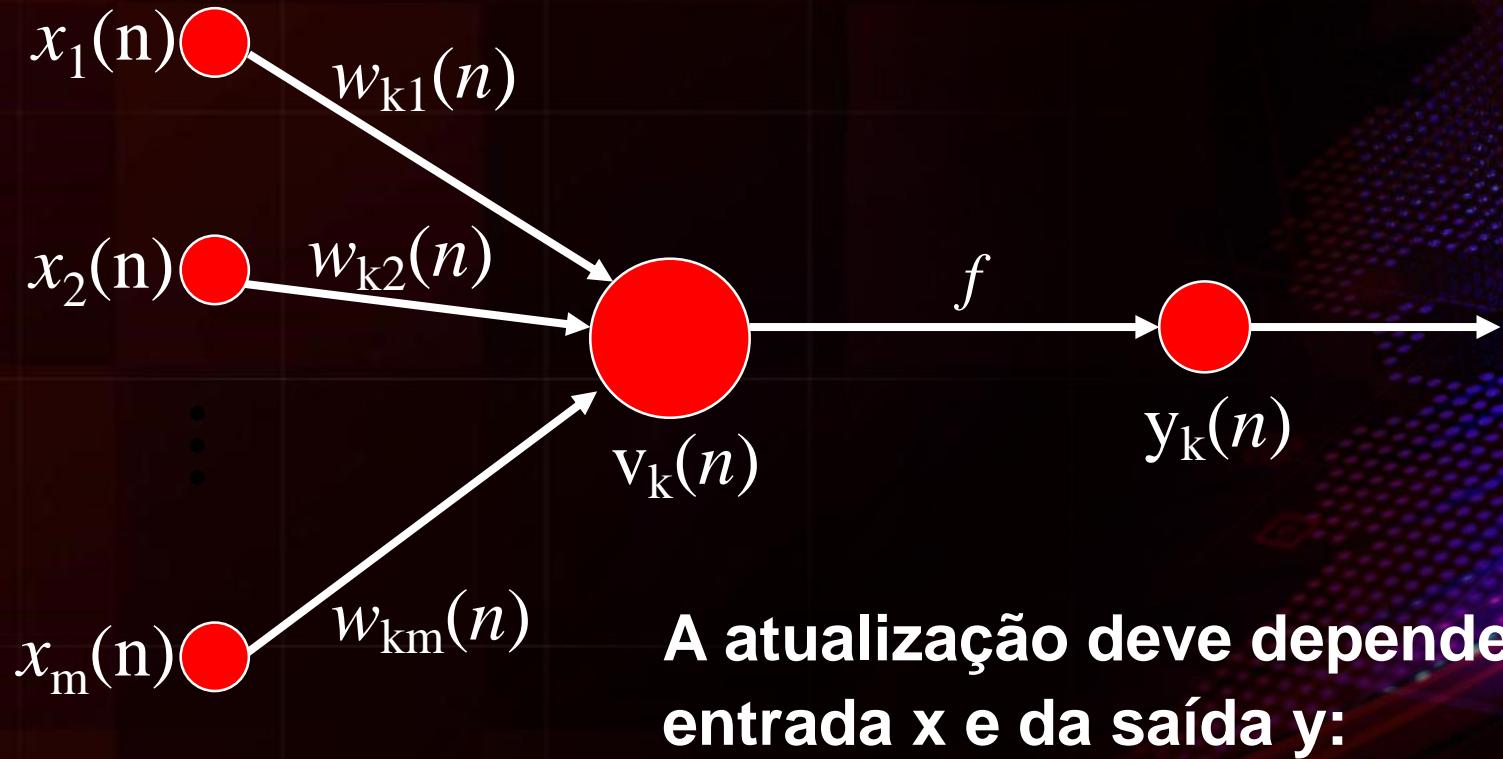


APRENDIZADO HEBBIANO

PROPRIEDADES DA SINAPSE HEBBIANA:

- **Mecanismo dependente do tempo:** dependem do tempo exato de ocorrência dos sinais pré e pós-sinápticos
- **Mecanismo local:** as modificações são locais (sinapses)
- **Mecanismo interativo:** depende explicitamente da interação entre os neurônios pré e pós-sinápticos de uma sinapse
- **Mecanismo correlativo:** a correlação temporal entre a atividade de ambos os lados da sinapse define a modificação da eficiência sináptica

APRENDIZADO HEBBIANO



A atualização deve depender da
entrada x e da saída y :

$$\Delta w_{kj}(n) = F(y_k(n), x_j(n))$$

APRENDIZADO HEBBIANO

$$\Delta w_{kj}(n) = F(y_k(n), x_j(n))$$

- **Forma Correlativa:**

$$\Delta w_{kj}(n) = \eta y_k(n)x_j(n)$$

- **Forma da Covariância:**

$$\Delta w_{kj}(n) = \eta(y_k(n) - \bar{y})(x_j(n) - \bar{x})$$

- **Forma com Saturação (regularização)**

$$\Delta w_{kj}(n) = \eta y_k(n) [x_j(n) - y_k(n)w_{kj}]$$

APRENDIZADO COMPETITIVO

- Neurônios de saída competem entre si para se tornarem ativos (disparar)
- Um neurônio (vencedor) ganha o direito de disparar enquanto os demais permanecem em repouso
- Apenas o vencedor (e seus vizinhos) são atualizados (aprendem)

APRENDIZADO COMPETITIVO

TEMOS TRÊS ELEMENTOS BÁSICOS:

1. Um conjunto de neurônios equivalentes, com pequenas modificações na distribuição dos pesos
2. Um limite de ativação é imposto sobre a força dos neurônios
3. Um mecanismo de competição para que apenas um neurônio se torne vencedor a um dado padrão: “O vencedor leva tudo”

APRENDIZADO COMPETITIVO

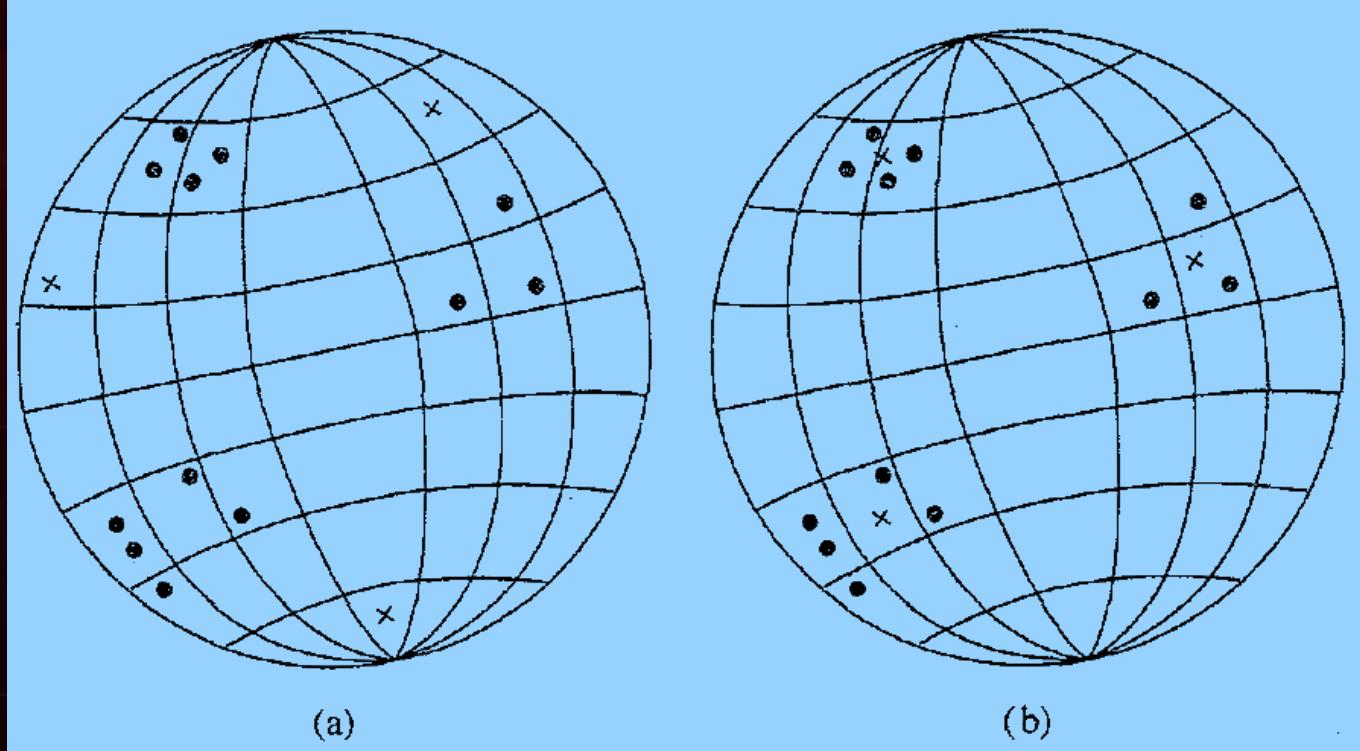
- A ativação do neurônio é consolidada a partir do campo local induzido:

$$y_k = \begin{cases} 1 & \text{se } v_k > v_j \forall j, j \neq k \\ 0 & \text{caso contrario} \end{cases}$$

- Regra de atualização:

$$\Delta w_{kj} = \begin{cases} \eta(x_j - w_{kj}) & \text{se } k \text{ e o neurônio vencedor} \\ 0 & \text{caso contrario} \end{cases}$$

APRENDIZADO COMPETITIVO



Fonte: Haykin (2001)

Os neurônios se especializam e se tornam protótipos dos dados existentes

APRENDIZADO DE BOLTZMANN

- Regra de aprendizagem enraizada nas ideias da mecânica estatística. Homenagem a Ludwig Boltzmann
- Máquina de Boltzmann – Proposta por Geoffrey Hinton e Terry Sejnowski em 1985
- Neurônios representam células binárias
- A máquina é caracterizada por uma função de energia

$$E = - \sum_j \sum_k w_{kj} x_k x_j$$

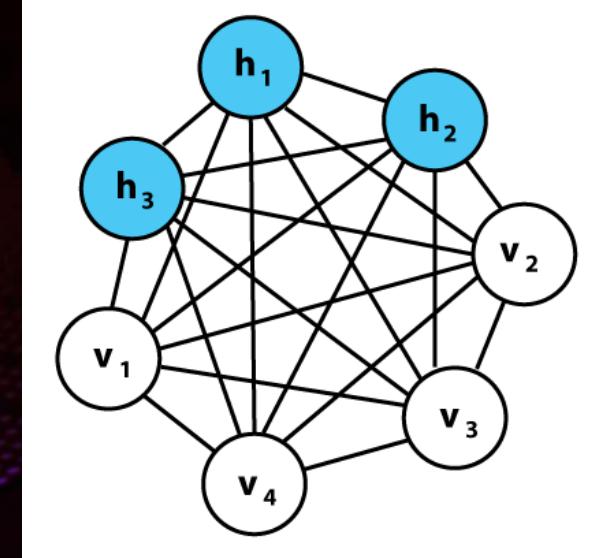
BOLTZMANN

- A máquina opera de forma estocástica, alterando o estado de um dado neurônio com probabilidade:

$$P(x_k) = \frac{1}{1 + \exp(-\Delta E_k/T)}$$

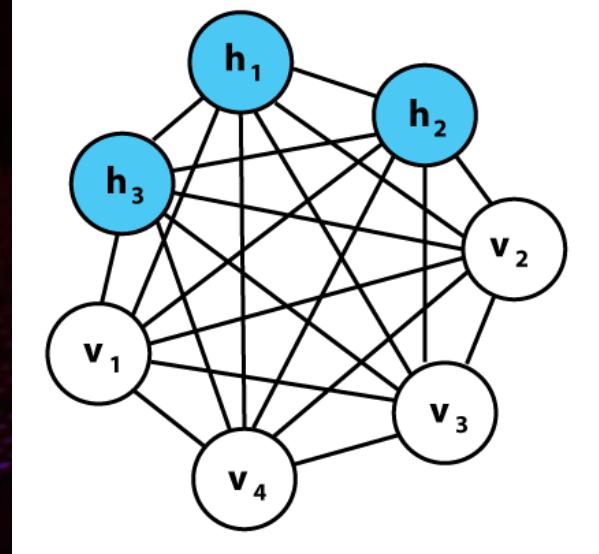
- A aplicação contínua dessa regra levará a máquina a um estado de equilíbrio “térmico”

- Há dois tipos de neurônios na máquina de Boltzmann: **visíveis** e **ocultos**



BOLTZMANN

- A máquina opera em duas condições:
 - Presa: os neurônios visíveis permanecem com valores fixos
 - Livre: todos os neurônios podem evoluir

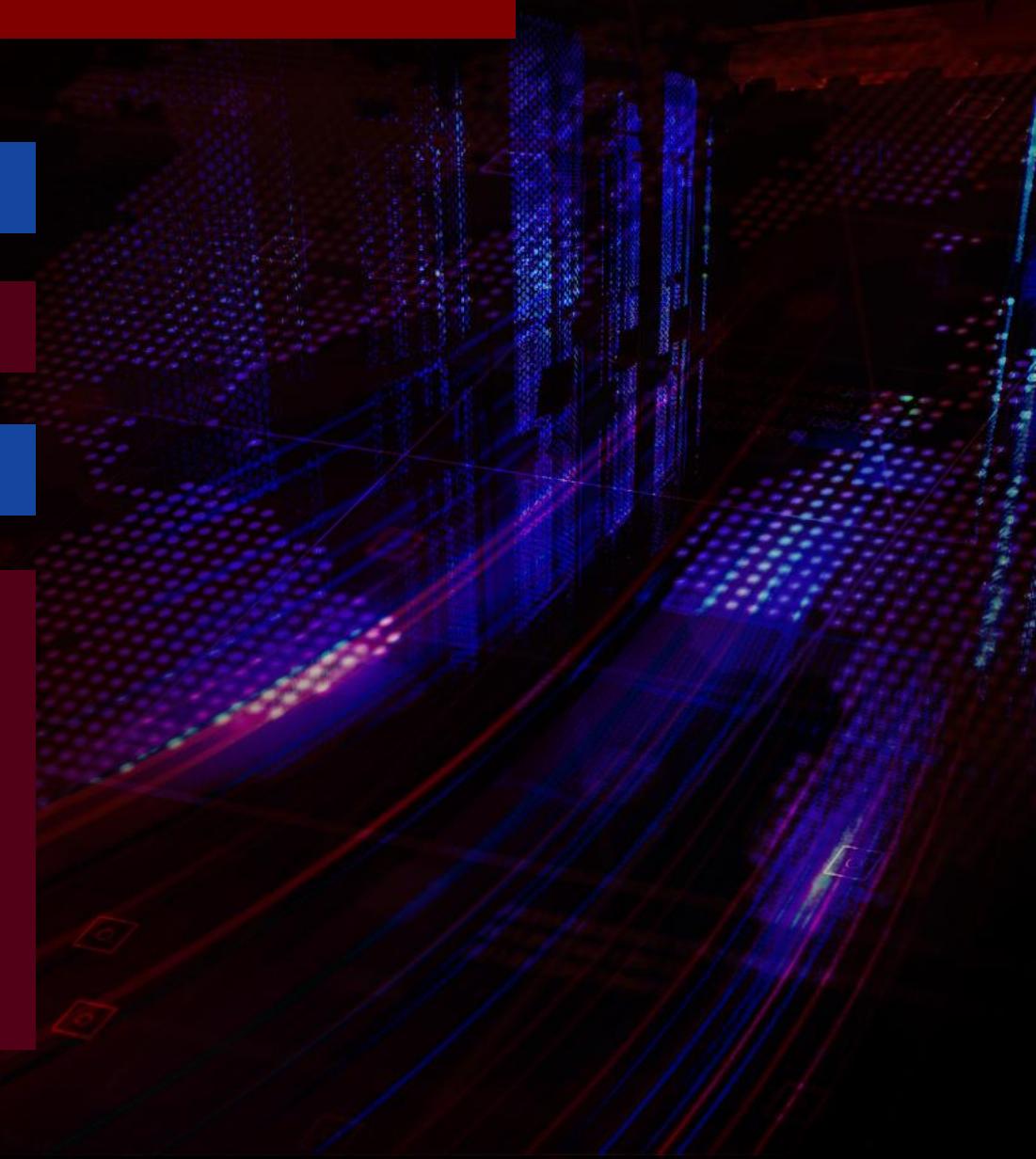


- A regra de atualização da máquina é definida por:

$$\Delta w_{kj} = \eta (\rho_{kj}^+ - \rho_{kj}^-)$$

PARADIGMAS DE APRENDIZAGEM

- Aprendizado Supervisionado
- Aprendizado Não-Supervisionado
- Aprendizado por Reforço
- Outras formas:
 - Aprendizado Auto-supervisionado
 - Aprendizado Semissupervisionado
 - Aprendizado Ativo



SUPERVISIONADO

- Todas as saídas desejadas são fornecidas pelo professor

$$\{(\mathbf{x}_k, d_k)\}_{i=1}^N$$

- Tem por objetivo criar um mapeamento entre as entradas e saídas do problema

- Forma mais comum: aprendizagem por correção de erros

SUPERVISIONADO

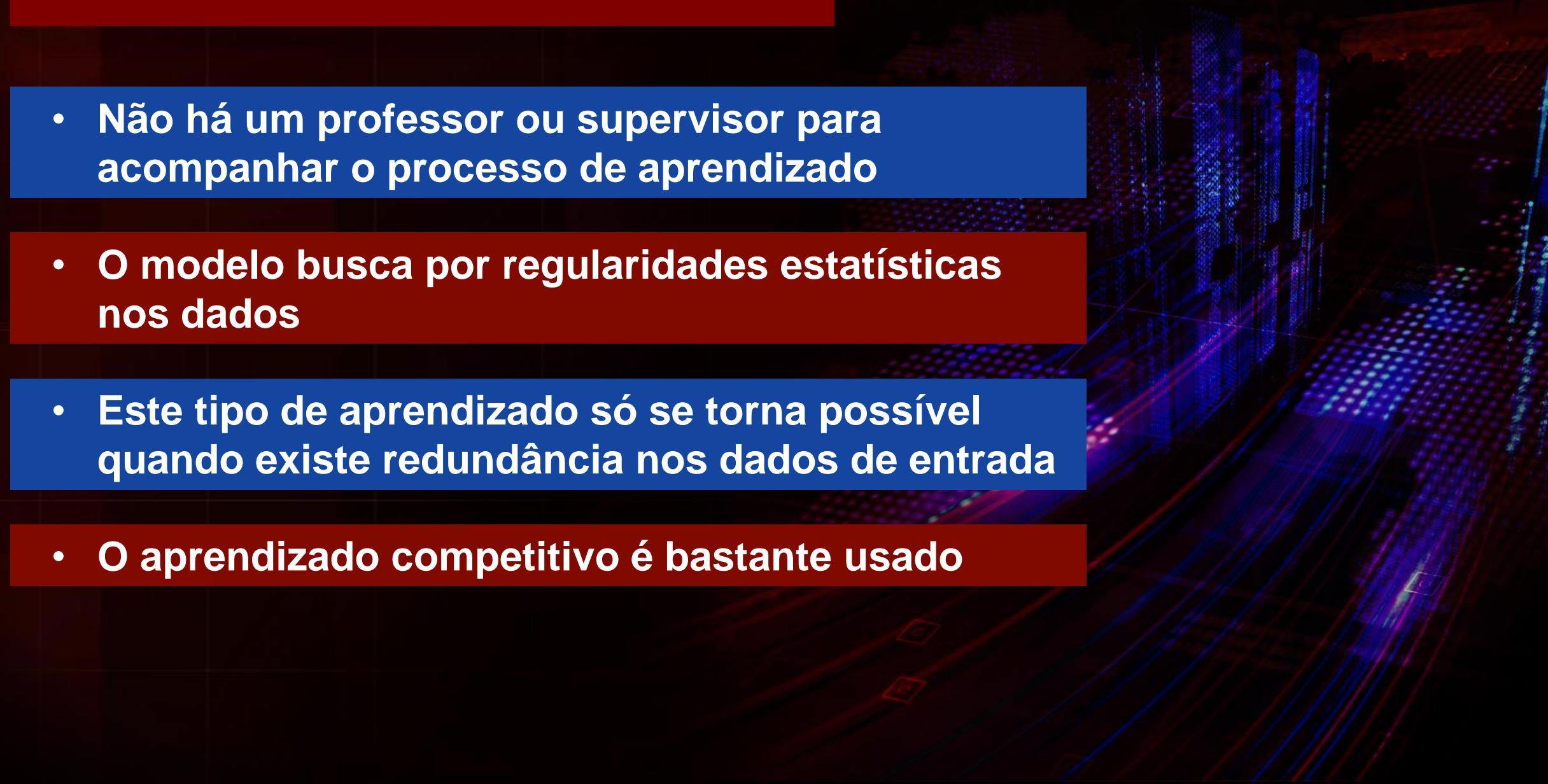
PRINCIPAIS APLICAÇÕES:

- Classificação de dados
- Regressão
- Tradução
- Reconhecimento
- Seleção de atributos



NÃO-SUPERVISIONADO

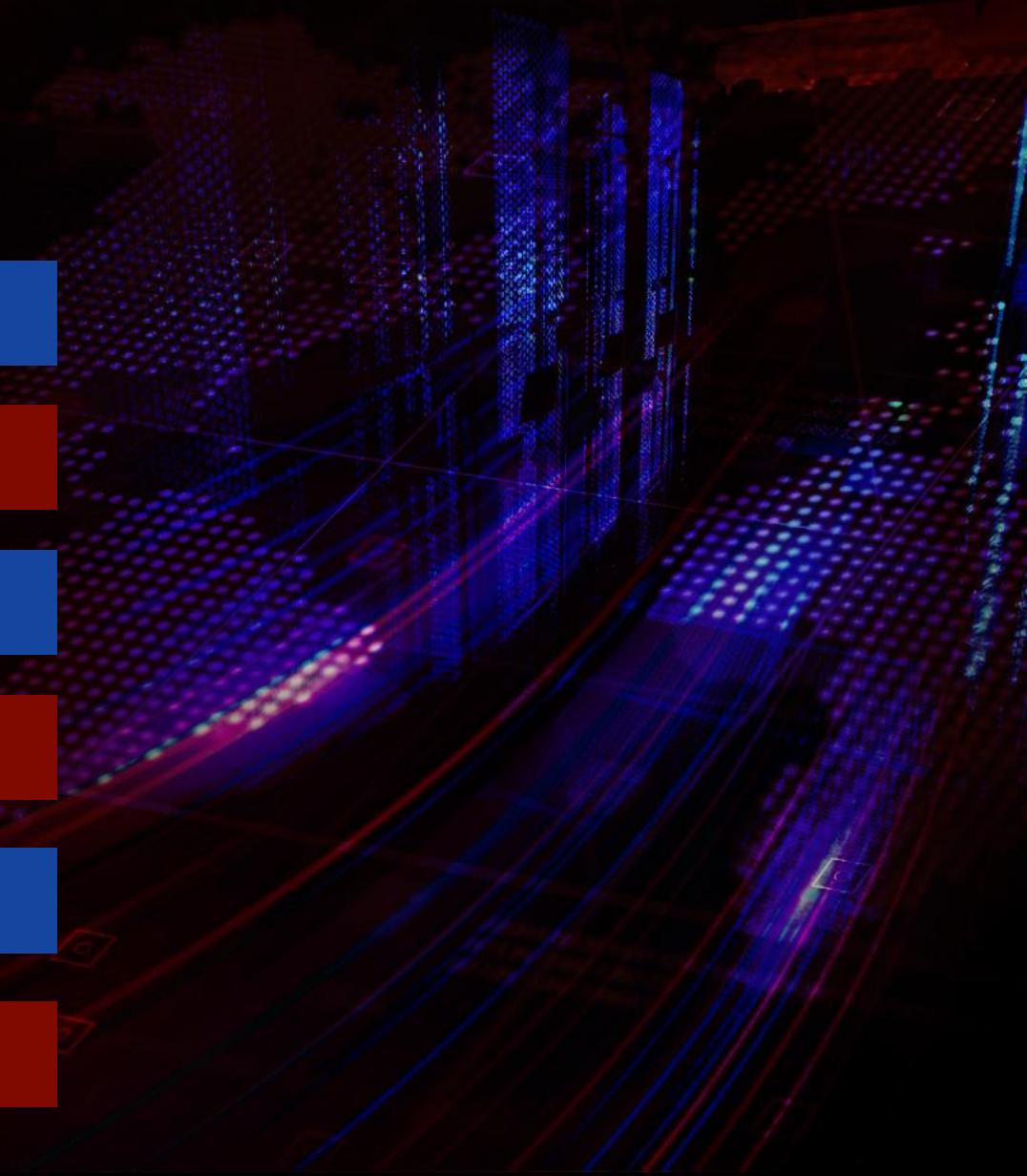
- Não há um professor ou supervisor para acompanhar o processo de aprendizado
- O modelo busca por regularidades estatísticas nos dados
- Este tipo de aprendizado só se torna possível quando existe redundância nos dados de entrada
- O aprendizado competitivo é bastante usado



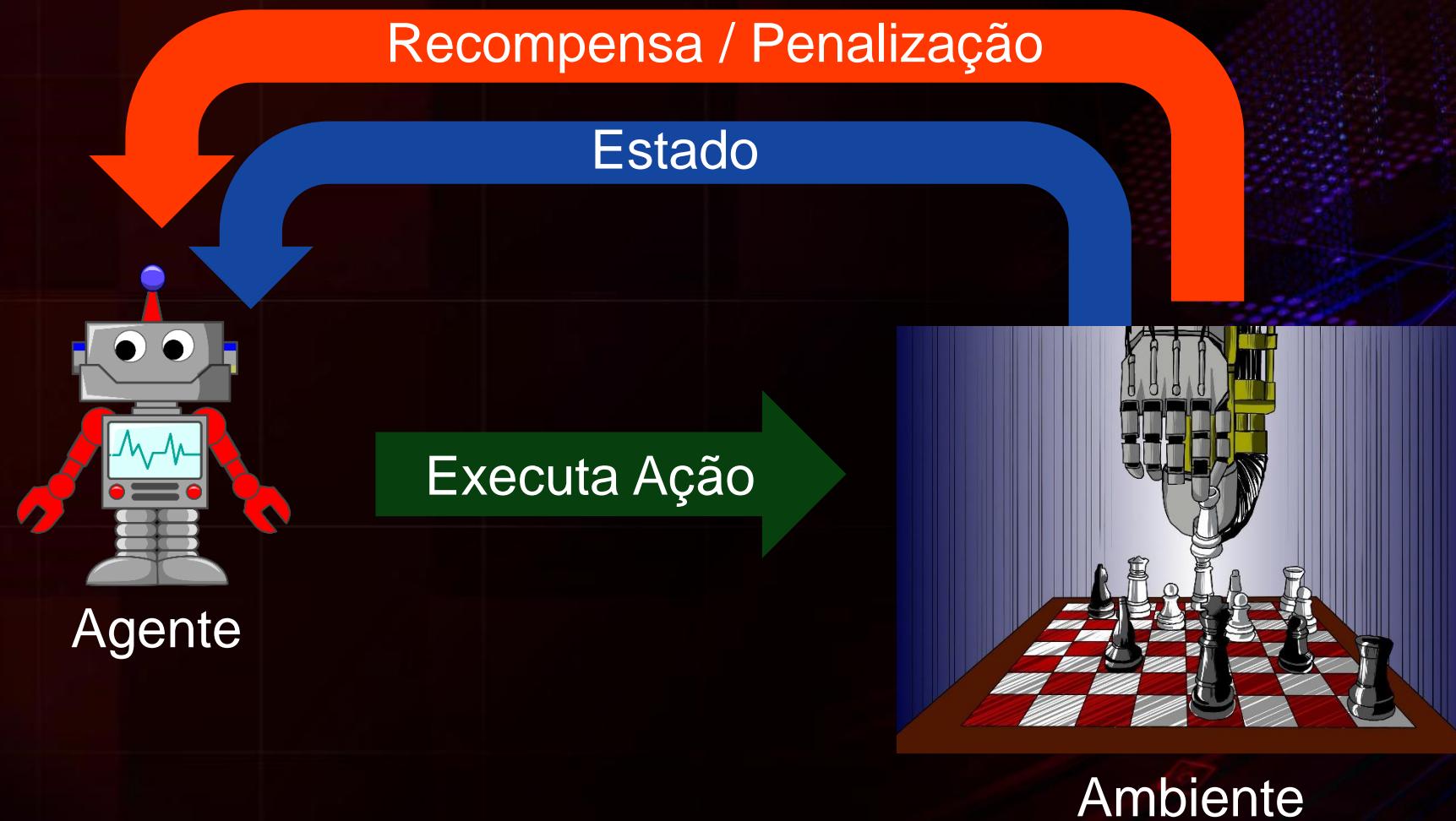
NÃO-SUPERVISIONADO

PRINCIPAIS APLICAÇÕES:

- Agrupamento de dados
- Redução de dimensionalidade
- Seleção de atributos
- Associação de padrões
- Detecção de anomalias
- Aprendizado de representações



APRENDIZADO POR REFORÇO

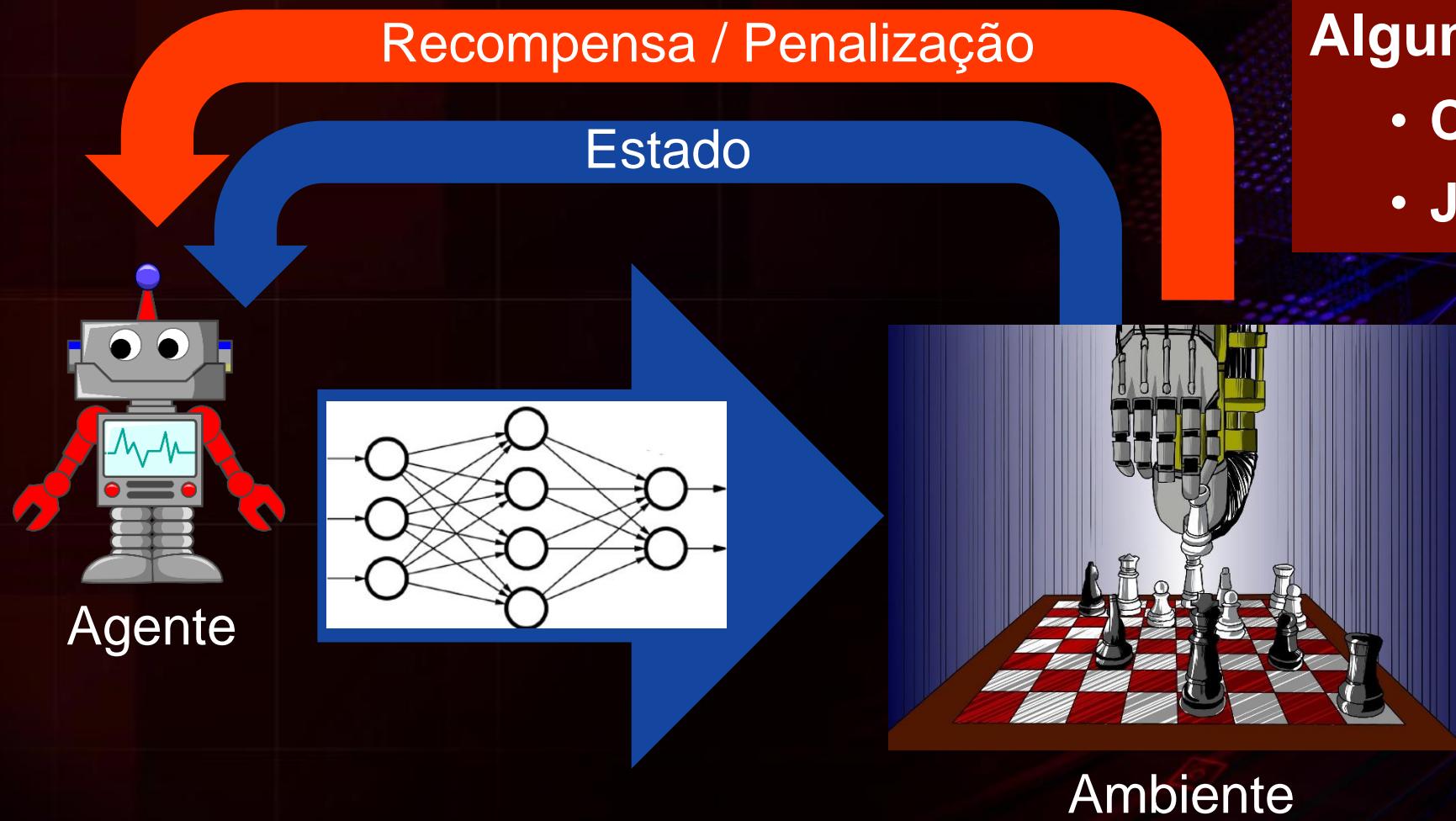


APRENDIZADO POR REFORÇO

- **Objetivo do algoritmo de aprendizagem:** Aprender as regras a partir da resposta do ambiente
- Normalmente usamos a Q-Table

Estado	Ação 1	Ação 2	Ação 3	Ação 4
1	0,1	0,2	0,6	0,1
2	0,9	0,0	0,1	0,0
3	0,0	0,2	0,0	0,8

APRENDIZADO POR REFORÇO



Algumas Aplicações

- Controle
- Jogos

SEMISSUPERVISIONADO

- Considera tanto dados **rotulados** quanto dados **não-rotulados**
- Combina métodos supervisionados e não-supervisionados
- Objetivo central:
 - Construir modelos melhores que os obtidos via aprendizado supervisionado
 - Reduzir o custo de desenvolvimento de um modelo de aprendizado de máquina



OUTRAS FORMAS DE APRENDIZAGEM

➤ APRENDIZADO AUTO-SUPERVISIONADO

- Extração automática de representações

➤ APRENDIZADO ATIVO

- Uso de um oráculo para avaliar novos dados
- Apenas dados representativos são selecionados

➤ TRANSFER LEARNING

- Aproveitar o conhecimento obtido em outros domínios (relacionados) e generalizá-lo



O QUE VIMOS?

- Conhecemos o que é o aprendizado em redes neurais
- Conhecemos as principais técnicas de aprendizagem
- Revisitamos os principais paradigmas de aprendizagem e algumas das suas aplicações

PRÓXIMA VIDEOAULA

- Aprenderemos a modelar a arquitetura de redes neurais via grafos direcionados
- Conheceremos as principais arquiteturas de redes neurais
 - Redes Alimentadas adiante
 - Redes Recorrentes
 - Redes Convolucionais
 - Redes Construtivas



REDES NEURAIS

Aprendizagem em redes neurais e
seus principais paradigmas

