

# **INTRODUÇÃO AO ESTUDO DE REDES NEURAIS ARTIFICIAIS**

## **MÓDULO INTRODUTÓRIO**

**Laboratório de Conexionismo e Ciências Cognitivas**

**L3C**

**Grupo SICRES**

**INE - UFSC**



# Objetivo

- Oferecer ao aluno uma introdução à abordagem da IA conexionista, descrevendo características de funcionamento, formas de aprendizado, modelos de redes neurais e aplicações típicas.



# Introdução

- As abordagens da Inteligência Artificial
- IA Simbólica:
  - Em um primeiro momento, a inteligência era geralmente associada a uma característica unicamente humana, de representação de conhecimentos e resolução de problemas, refletindo um ponto de vista altamente antropocêntrico.
  - Um sistema simbólico é capaz de manifestar um comportamento inteligente.
  - HIPÓTESE DO SISTEMA DE SÍMBOLOS FÍSICOS (Newell & Simon)



# Introdução

- As abordagens da Inteligência Artificial
- IA Conexionista:
  - Se for construído um modelo suficientemente preciso do cérebro, este modelo apresentará um comportamento inteligente. Se apenas uma pequena parte do cérebro for reproduzida, a função exercida por esta parte emergirá do modelo.
  - Existem uma série de problemas que os seres vivos, e os seres humanos em particular, parecem resolver de maneira inata: o processamento de imagens, o reconhecimento da fala, a recuperação de informações de maneira associativa, a filtragem adaptativa de sinais, o aprendizado de novos fatos e idéias, etc.

# Introdução

## ■ IA Simbólica

- Conhecimento representado por regras (ou outra estrutura similar) que podem ser facilmente tratadas e analisadas.
- Permite a explicação do processo que levou a uma determinada resposta.
- Fácil inserção de novos conhecimentos obtidos a partir do especialista ou através de métodos automáticos de aquisição de conhecimento.

## ■ IA Conexionista

- Conhecimento representado por exemplos (casos resolvidos).
- Dificuldade de explicar todos os conhecimentos relativos ao problema através de regras simbólicas.
- A inserção de novos casos é trabalhosa, pois a rede deve ser re-treinada.
- São capazes de generalizar o seu conhecimento a partir de exemplos anteriores.

# Introdução

## ■ IA Simbólica X IA Conexionista

- Conhecimento Teórico / Empírico

- $AND(A,B) =$

- if  $A=0$

- then  $AND=0$

- else if  $B=0$

- then  $AND=0$

- else  $AND=1$

A	B	AND
0	0	0
0	1	0
1	0	0
1	1	1

# Introdução

- IA Simbólica X IA Conexionista

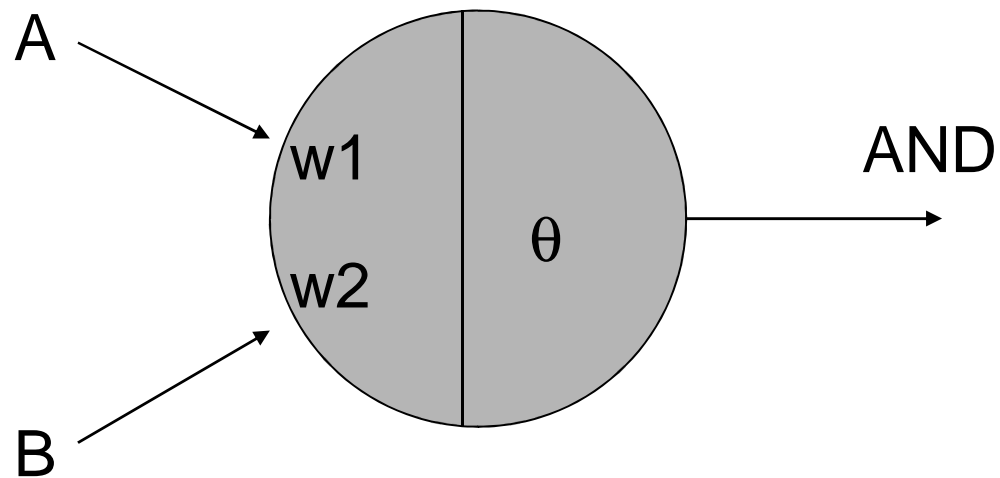
A	B	AND
0	0	0
0	1	0
1	0	0
1	1	1

$$w1=1$$

$$w2=1$$

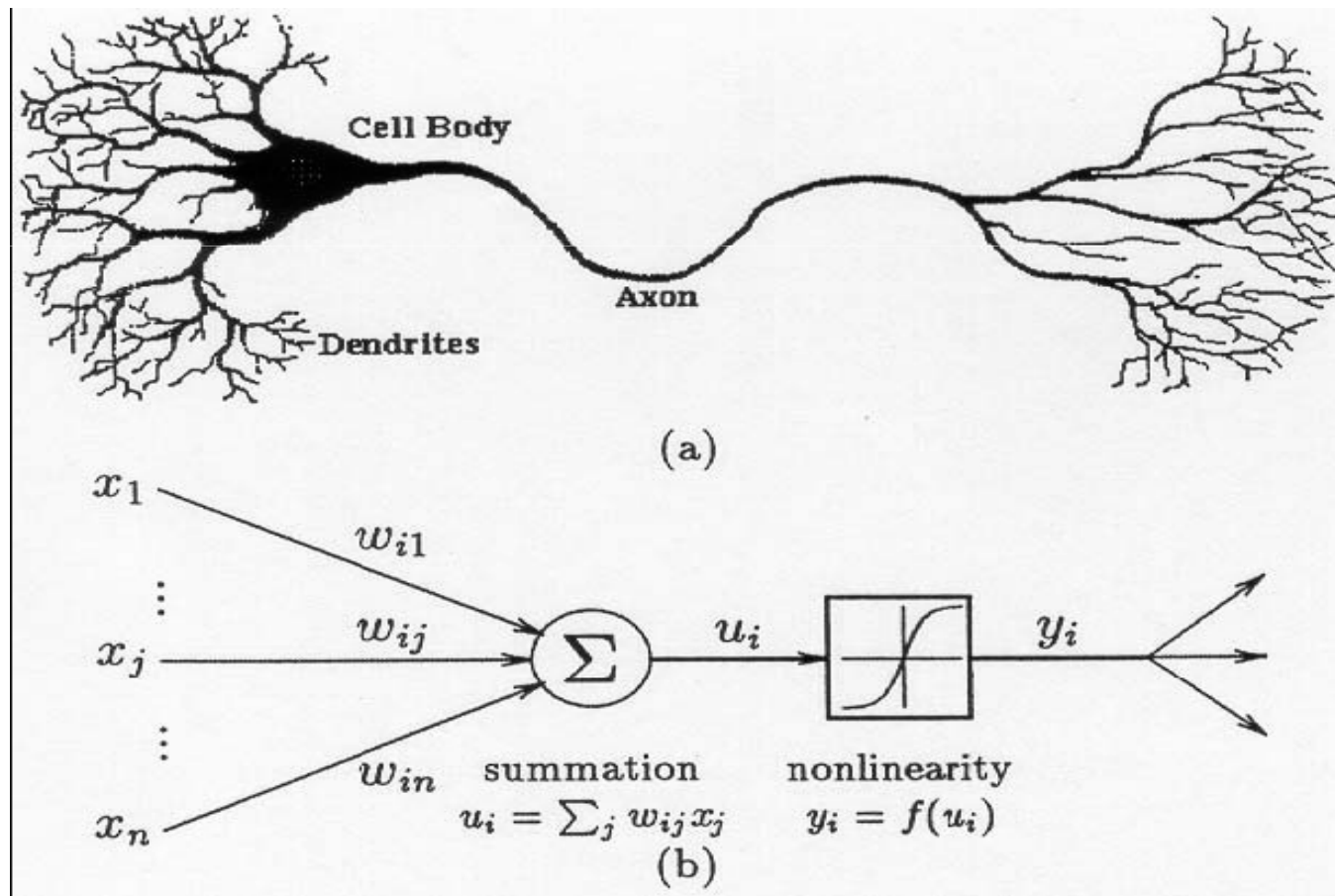
$$\theta = 2$$

$$A.w1+B.w2 \geq \theta$$



# Introdução

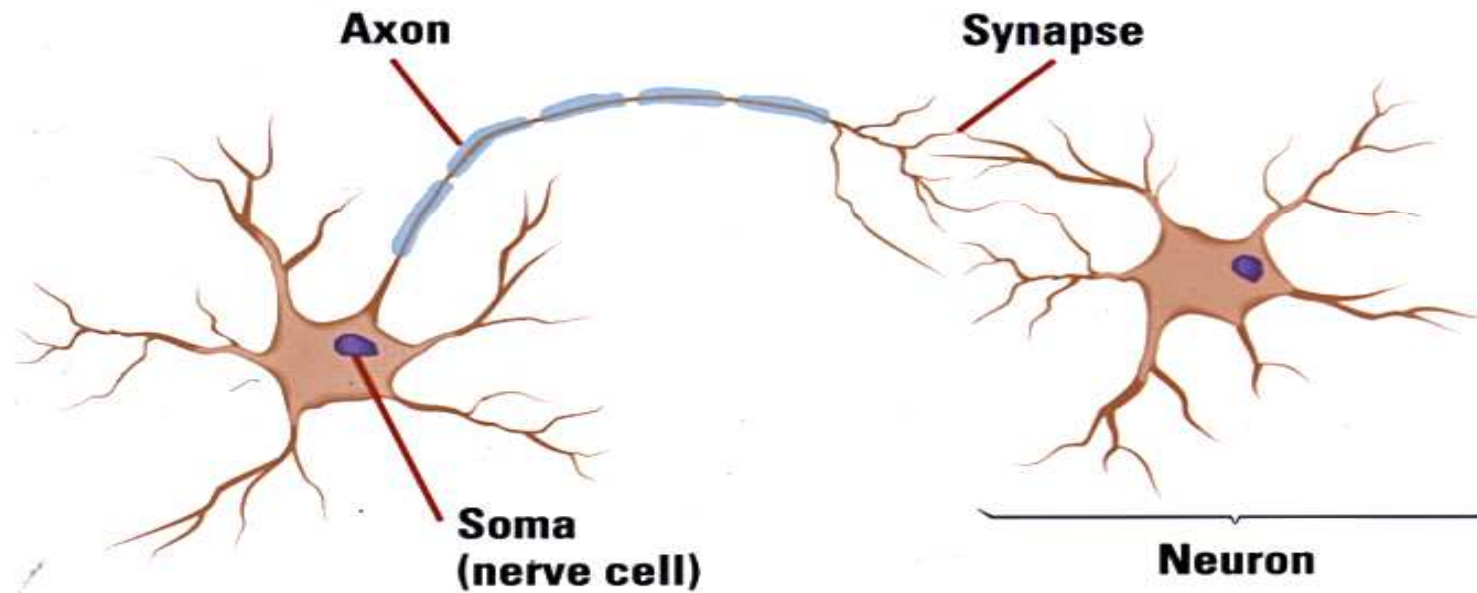
- Fundamentos Biológicos





# Introdução

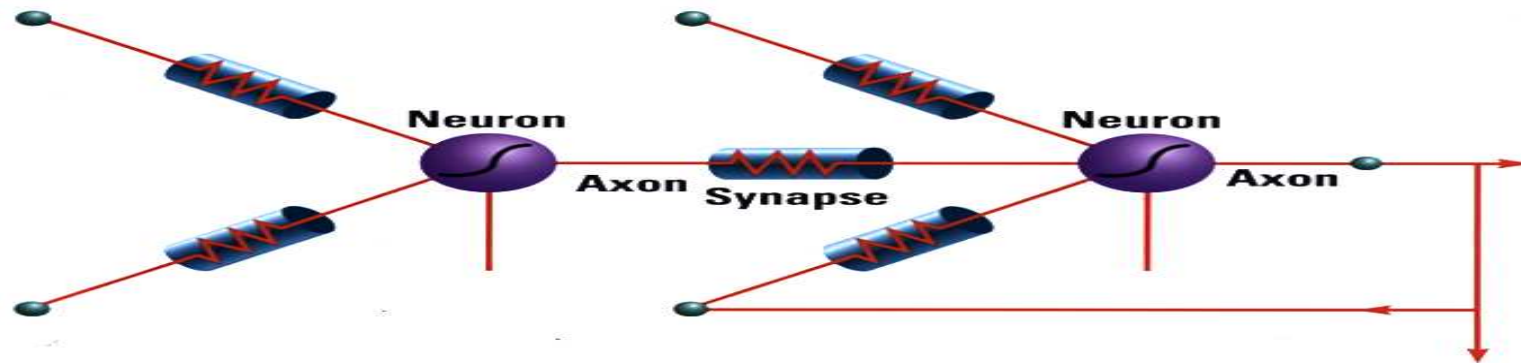
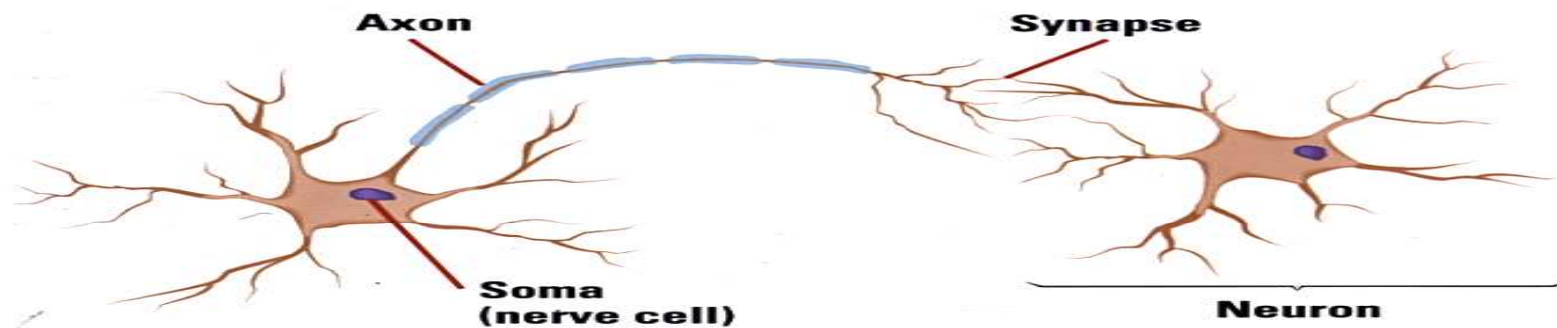
- Fundamentos Biológicos



*Natural Neurons*

# Introdução


- Fundamentos Biológicos



**Artificial Neurons**



# Introdução

- O Cérebro
    - Composto por Neurônios  
( $10^5$  na mosca da fruta,  $5 \times 10^6$  no rato e  $10^{11}$  no homem)
  - O Cérebro Humano
    - O cérebro humano tem ~ 10 bilhões de neurônios.
    - Cada neurônio tem ~ 1.000 a 10.000 conexões
    - 60 trilhões de conexões -  $10^{14}$  sinapses!
    - Massa: 1-2kg no adulto - 2% do peso  
20% do peso do recém-nascido.
    - Usa 20% do oxigênio, 25% da glucose, 15% do fluxo de sangue.
- 



# Introdução

## ▪ DEFINIÇÃO

- Redes Neurais Artificiais são sistemas inspirados nos neurônios biológicos e na estrutura massivamente paralela do cérebro, com capacidade de adquirir, armazenar e utilizar conhecimento experimental.

## ▪ IDÉIA BÁSICA

- Sistemas compostos de diversas unidades simples (neurônios artificiais) ligadas de maneira apropriada, podem gerar comportamentos interessantes e complexos.
- Comportamento é determinado pela estrutura das ligações (topologia) e pelos valores das conexões (pesos sinápticos)



# Introdução

- Devido à similaridade com a estrutura do cérebro, as Redes Neurais exibem características similares ao do comportamento dos seres vivos, tais como:
  - Procura Paralela e Endereçamento pelo Conteúdo:
    - O cérebro não possui endereço de memória e não procura a informação sequencialmente.
  - Aprendizado:
    - A rede aprende por experiência, não necessitando explicitar os algoritmos para executar uma determinada tarefa.



# Introdução

- Devido à similaridade com a estrutura do cérebro, as Redes Neurais exibem características similares ao do comportamento dos seres vivos, tais como:
  - Associação:
    - A rede é capaz de fazer associações entre padrões diferentes.
      - Cidade → Pessoa
      - Perfume → Pessoa
      - Pessoa → Nome



# Introdução

- Devido à similaridade com a estrutura do cérebro, as Redes Neurais exibem características similares ao do comportamento dos seres vivos, tais como:
  - Generalização:
    - Redes Neurais são capazes de generalizar o seu conhecimento a partir de exemplos anteriores.
    - Habilidade de lidar com ruídos e distorções, respondendo corretamente a padrões novos.
  - Robustez e Degradação Gradual:
    - A perda de um conjunto de elementos processadores e/ou conexões sinápticas não causa o mal funcionamento da rede neural.



# **Modelos de Neurônios Artificiais**

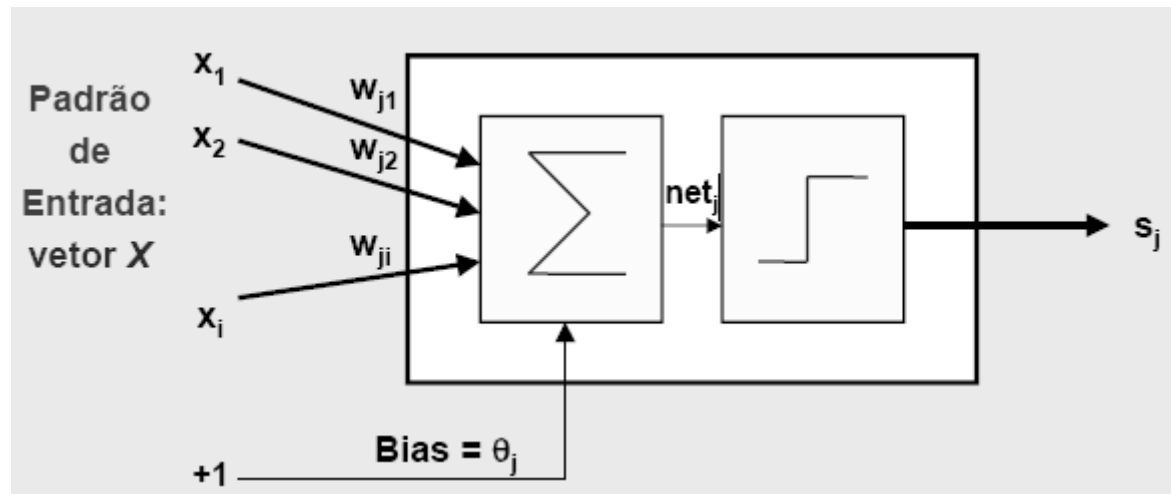
## **▪ Evolução Histórica**

- McCulloch & Pitts (1943):
  - modelo computacional para o neurônio artificial. Não possuía capacidade de aprendizado
- Hebb (1949):
  - modelo de aprendizado (Hebbian Learning Rule)
- Rosenblatt (1957):
  - modelo Perceptron, com grande sucesso em certas aplicações e problemas em outras aparentemente similares



# Modelos de Neurônios Artificiais

- O PERCEPTRON (Rosenblatt (1957))
  - Na sua forma mais simples o modelo perceptron consiste de:



# Modelos de Neurônios Artificiais

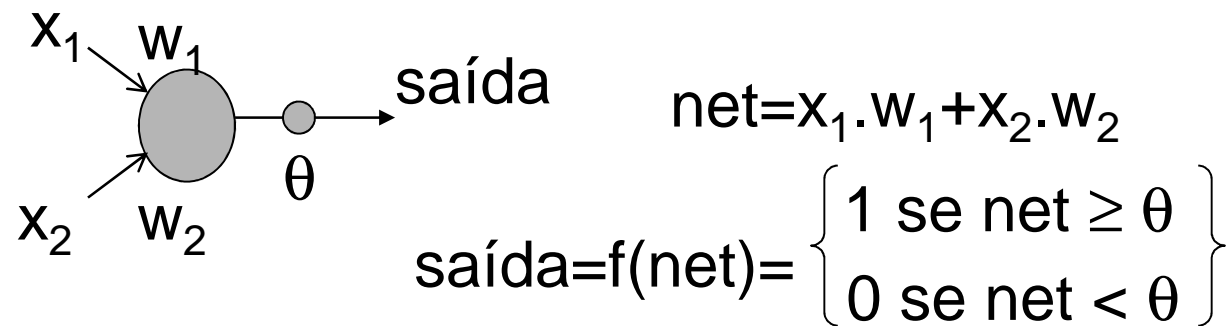
- **O PERCEPTRON (Rosenblatt (1957))**

- Características Básicas:

- Regra de Propagação:  $net_j = \sum x_i \cdot w_{ji} + \theta_j$
    - Função de Ativação: Degrau
    - Topologia: Uma única camada de processadores.
    - Algoritmo de Aprendizado: Supervisionado
    - Valores de Entrada/Saída: Binários

# Modelos de Neurônios Artificiais

## ▪ O PERCEPTRON (Rosenblatt (1957))

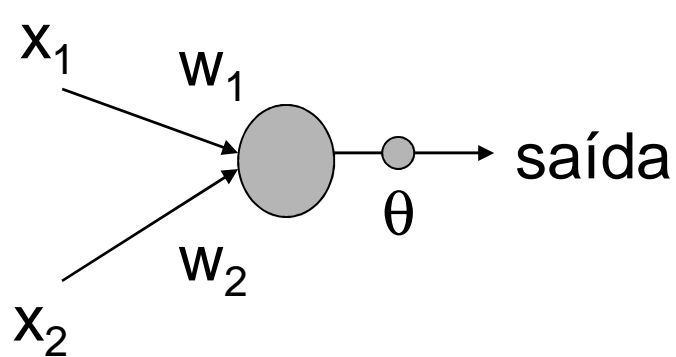


EQUAÇÃO FUNDAMENTAL DO PERCEPTRON

$$w_1 \cdot x_1 + w_2 \cdot x_2 = \theta \quad \longleftarrow \text{EQUAÇÃO DE UMA RETA}$$

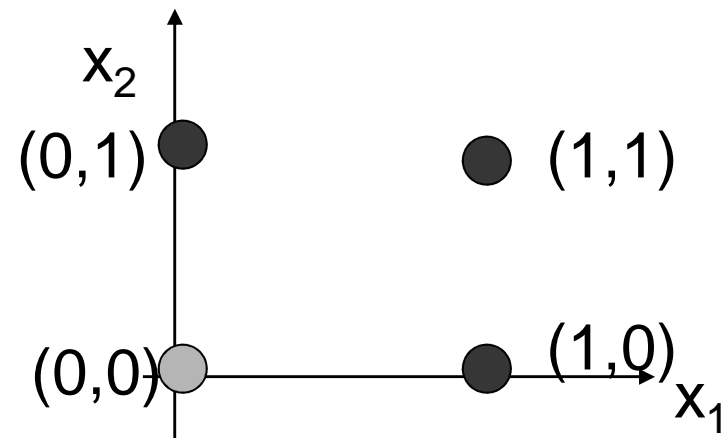
# Modelos de Neurônios Artificiais

## ▪ O PERCEPTRON (Rosenblatt (1957))



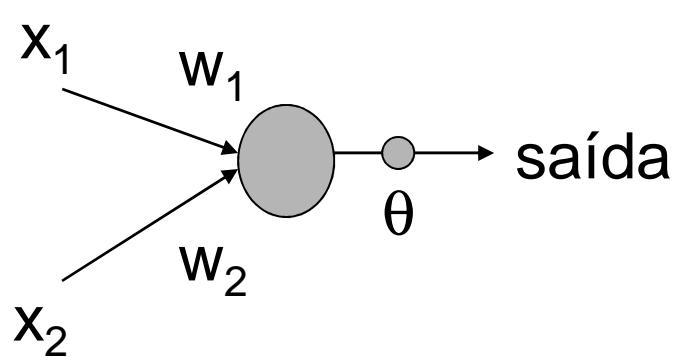
$$\text{saída} = f(\text{net}) = \begin{cases} 1 & \text{se } \text{net} \geq \theta \\ 0 & \text{se } \text{net} < \theta \end{cases}$$

Determinar  $w_1$ ,  $w_2$  e  $\theta$   
para que o Perceptron  
aprenda a função OU  
Lógico (OR)



# Modelos de Neurônios Artificiais

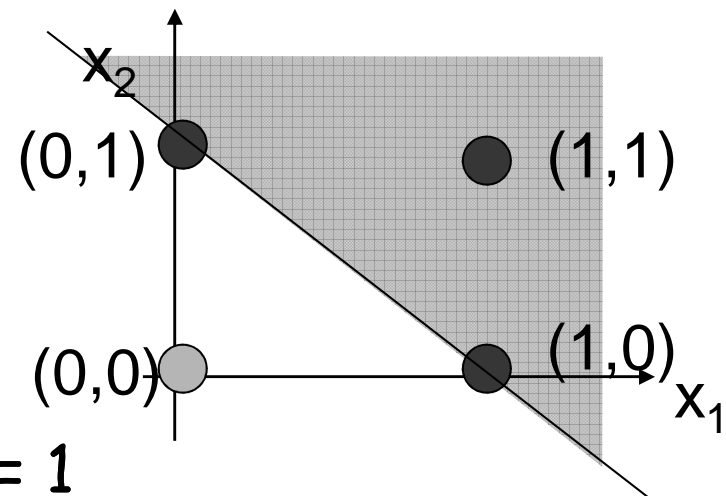
## ▪ O PERCEPTRON (Rosenblatt (1957))



$$\text{saída} = f(\text{net}) = \begin{cases} 1 & \text{se } \text{net} \geq \theta \\ 0 & \text{se } \text{net} < \theta \end{cases}$$

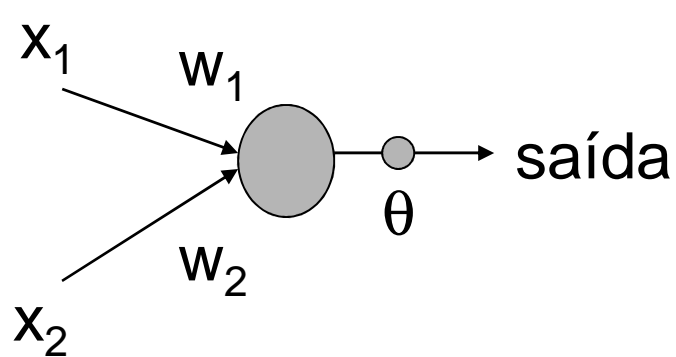
Determinar  $w_1$ ,  $w_2$  e  $\theta$   
para que o Perceptron  
aprenda a função OU  
Lógico (OR)

Possível solução:  $w_1 = w_2 = \theta = 1$



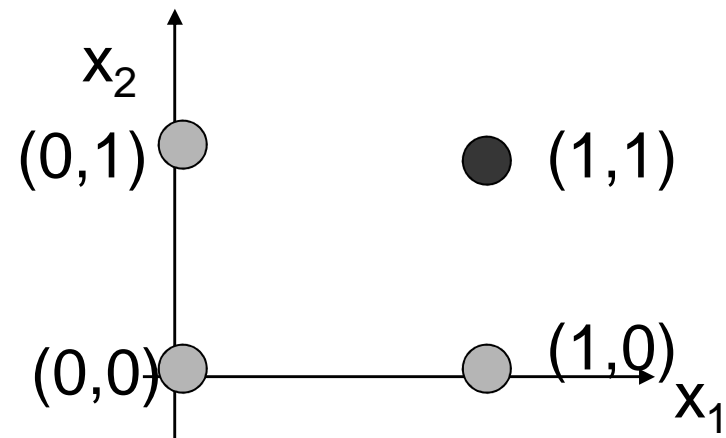
# Modelos de Neurônios Artificiais

## ▪ O PERCEPTRON (Rosenblatt (1957))



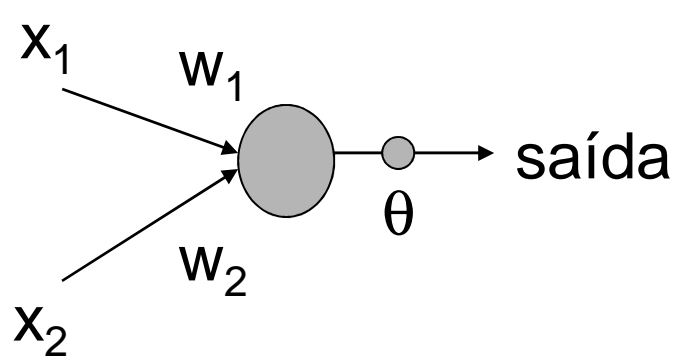
$$\text{saída} = f(\text{net}) = \begin{cases} 1 & \text{se } \text{net} \geq \theta \\ 0 & \text{se } \text{net} < \theta \end{cases}$$

Determinar  $w_1$ ,  $w_2$  e  $\theta$   
para que o Perceptron  
aprenda a função E Lógico  
(AND)



# Modelos de Neurônios Artificiais

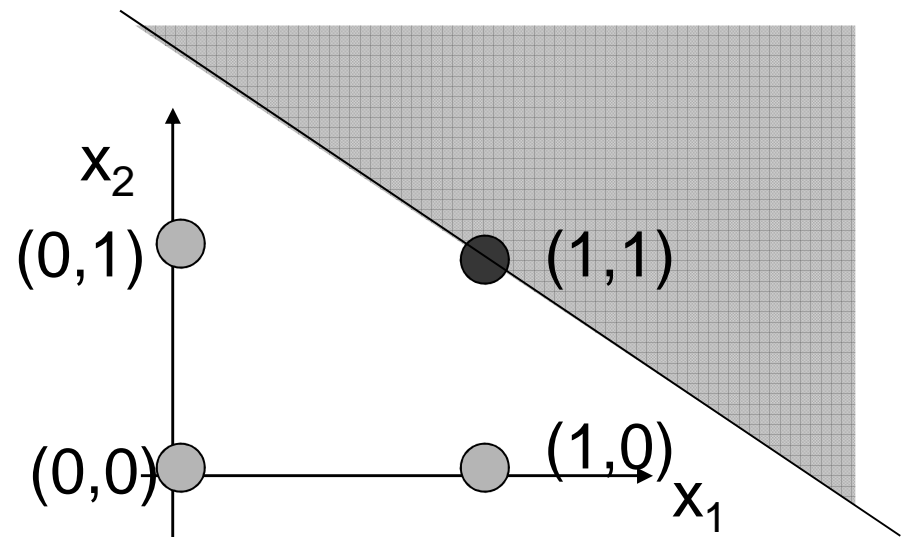
## ▪ O PERCEPTRON (Rosenblatt (1957))



$$\text{saída} = f(\text{net}) = \begin{cases} 1 & \text{se } \text{net} \geq \theta \\ 0 & \text{se } \text{net} < \theta \end{cases}$$

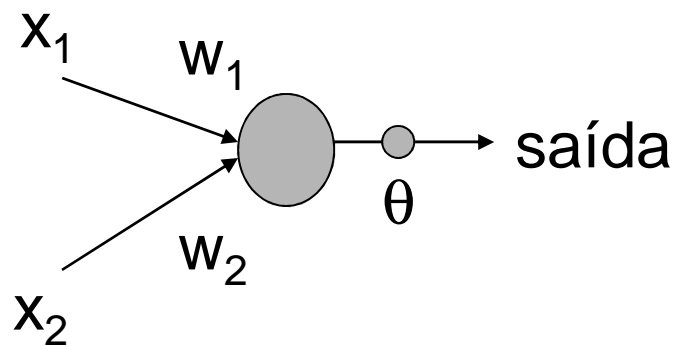
Determinar  $w_1$ ,  $w_2$  e  $\theta$   
para que o Perceptron  
aprenda a função E Lógico  
(AND)

Possível solução:  $w_1 = w_2 = 1$   
 $\theta = 2$



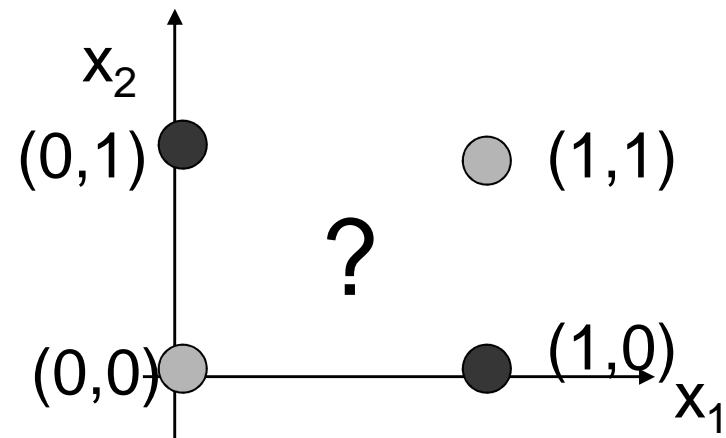
# Modelos de Neurônios Artificiais

## ▪ O PERCEPTRON (Rosenblatt (1957))



$$\text{saída} = f(\text{net}) = \begin{cases} 1 & \text{se } \text{net} \geq \theta \\ 0 & \text{se } \text{net} < \theta \end{cases}$$

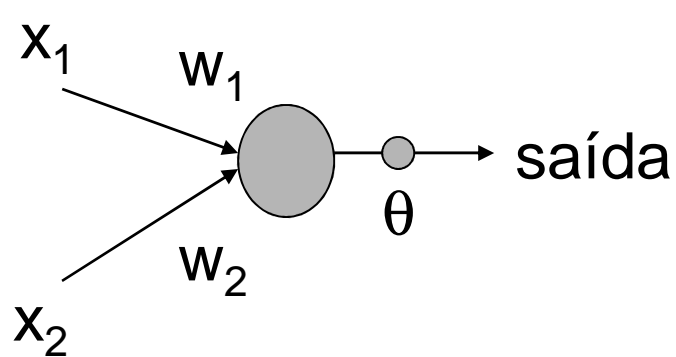
Determinar  $w_1$ ,  $w_2$  e  $\theta$   
para que o Perceptron  
aprenda a função OU-  
EXCLUSIVO (XOR)





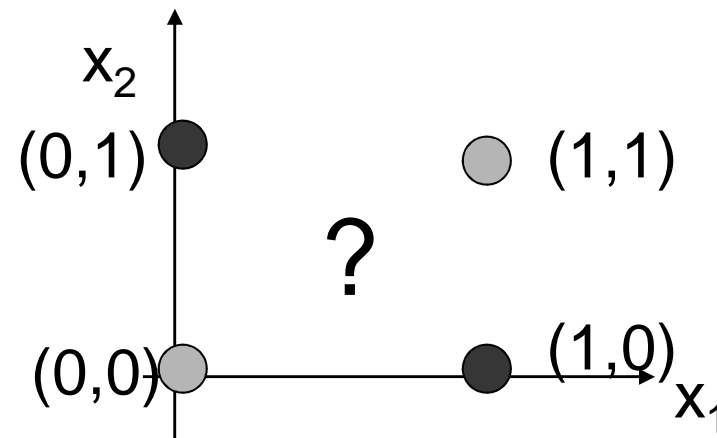
# Modelos de Neurônios Artificiais

## ▪ O PERCEPTRON (Rosenblatt (1957))



$$\text{saída} = f(\text{net}) = \begin{cases} 1 & \text{se } \text{net} \geq \theta \\ 0 & \text{se } \text{net} < \theta \end{cases}$$

Determinar  $w_1$ ,  $w_2$  e  $\theta$   
para que o Perceptron  
aprenda a função OU-  
EXCLUSIVO (XOR)

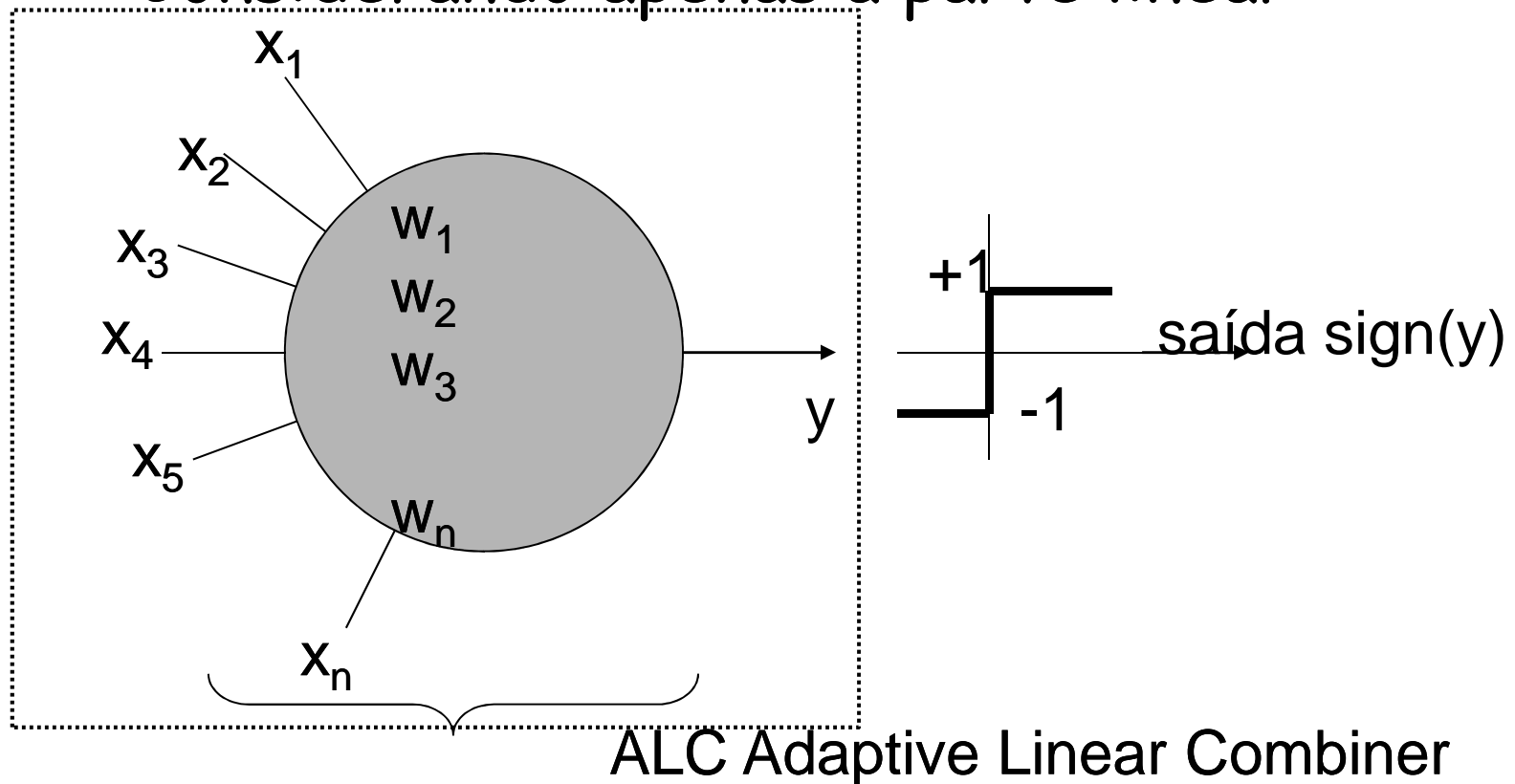


**CONCLUSÃO: O PERCEPTRON É CAPAZ DE DISTINGUIR APENAS PADRÕES LINEARMENTE SEPARÁVEIS!!!**

# Modelos de Neurônios Artificiais

- O PERCEPTRON (Aprendizado)

- Considerando apenas a parte linear:



# Modelos de Neurônios Artificiais

## ▪ O PERCEPTRON (Aprendizado)

### ▫ Formulação Vetorial

$$W = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ \dots \\ w_n \end{pmatrix} \quad X = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \quad y = W^T \cdot X$$

- Suponha que temos um conjunto de vetores de entrada  $\{X_1, X_2, \dots, X_L\}$ , cada um com seu valor de saída correto (desejado)  $\{d_1, d_2, \dots, d_L\}$ .
- O Método de Aprendizado procura achar os pesos de forma a minimizar a diferença entre a saída desejada e a saída obtida com o vetor de entrada. **(REGRESSÃO LINEAR)**

# Modelos de Neurônios Artificiais

- O PERCEPTRON (Aprendizado)
  - Como Minimizar o Erro Médio Quadrático
    - Erro para o k-ésimo vetor de entrada:
    - $E_k = d_k - y_k$  desejado - obtido

$$\langle E_k^2 \rangle = \frac{1}{L} \sum_{k=1}^L E_k^2$$

$$\langle E_k^2 \rangle = \frac{1}{L} \sum_{k=1}^L (d_k - y_k)^2 = \langle (d_k - y_k)^2 \rangle$$

$$\text{mas } y_k = W^T X_k$$

$$\text{logo } \langle E_k^2 \rangle = \langle (d_k - W^T X_k)^2 \rangle =$$

$$\langle d_k^2 \rangle + W^T \langle X_k X_k^T \rangle W - 2 \langle d_k X_k^T \rangle W$$

# Modelos de Neurônios Artificiais

- O PERCEPTRON (Aprendizado)
  - Como Minimizar o Erro Médio Quadrático
  - (método analítico - resolução de um sistema de equações lineares)

$$\langle E_k^2 \rangle = \langle d_k^2 \rangle + W^T \langle X_k X_k^T \rangle W - 2 \langle d_k X_k^T \rangle W$$

$$\frac{\partial \langle E_k^2 \rangle}{\partial W} = 0$$

$$2 \langle X_k X_k^T \rangle W - 2 \langle d_k X_k \rangle = 0$$

# Modelos de Neurônios Artificiais

- O PERCEPTRON (Aprendizado)
  - Como Minimizar o Erro Médio Quadrático (método analítico - resolução de um sistema de equações lineares)

$$2 \langle X_k X_k^T \rangle W - 2 \langle d_k X_k \rangle = 0$$

$$\langle X_k X_k^T \rangle = XX^T$$

$$\langle d_k X_k \rangle = XD^T$$

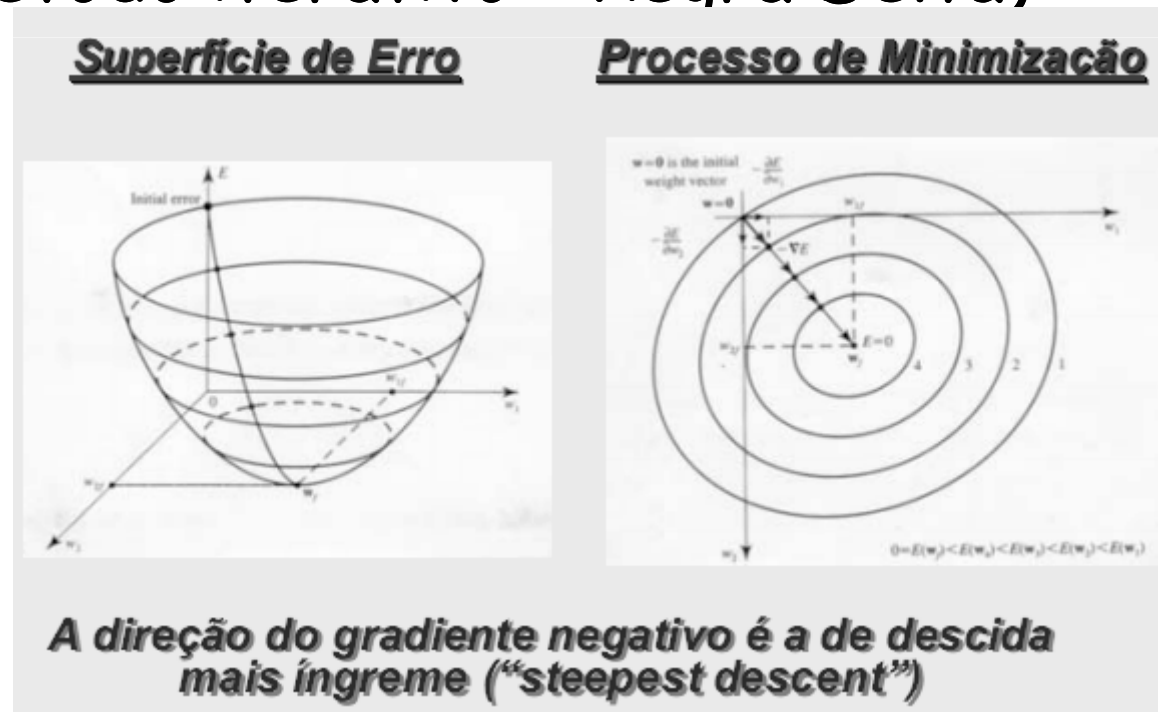
$$W = (XX^T)^{-1} (XD^T)$$

# Modelos de Neurônios Artificiais

- **O PERCEPTRON (Aprendizado)**
  - Como Minimizar o Erro Médio Quadrático (método iterativo - Regra Delta)
    - 1-Começar especificando valores aleatórios para os pesos.
    - 2-Aplicar um vetor de entrada  $X_k$ .
    - 3-Determinar o erro  $E_k(t)$  utilizando  $W(t)$ .
$$E_k^2(t) = (d_k - W^T(t)X_k)^2$$
    - 4-Supor que  $E_k^2(t)$  é uma aproximação razoável para o Erro Médio Quadrático  $\langle E_k^2(t) \rangle$

# Modelos de Neurônios Artificiais

- O PERCEPTRON (Aprendizado)
  - Como Minimizar o Erro Médio Quadrático (método iterativo - Regra Delta)





# Modelos de Neurônios Artificiais

- O PERCEPTRON (Aprendizado)
  - Como Minimizar o Erro Médio Quadrático (método iterativo - Regra Delta)

5-Calcular o gradiente do erro, isto é a “direção” em que derivada é maior.

$$E_k^2(t) = (d_k - W^T(t)X_k)^2$$

$$\nabla E_k^2(t) = 2(d_k - W^T(t)X_k) \frac{\partial E_k}{\partial W} =$$

$$2E_k(t).(-X_k) = -2E_k X_k$$

# Modelos de Neurônios Artificiais

- O PERCEPTRON (Aprendizado)
  - Como Minimizar o Erro Médio Quadrático (método iterativo - Regra Delta)

6-Atualizar o Vetor de Pesos.

$$W(t+1) = W(t) - \mu \nabla E_k^2(t)$$

$$W(t+1) = W(t) + 2\mu E_k X_k$$

com

$$E_k = d_k - W^T(t) X_k$$



# **Modelos de Neurônios Artificiais**

- **O PERCEPTRON (Aprendizado)**
  - Como Minimizar o Erro Médio Quadrático (método iterativo - Regra Delta)
    - 7-Repetir os passos de 2 a 6 até o erro alcançar um valor suficientemente pequeno.



# **Modelos de Neurônios Artificiais**

- **O PERCEPTRON (O Período Negro)**
  - Minsky e Papert - 1969 - Livro "Perceptrons". O Perceptron é incapaz de classificar corretamente padrões não linearmente separáveis.
  - A maioria dos problemas são não linearmente separáveis.
  - Se colocar mais uma camada de neurônios é possível contornar esta limitação, mas como achar os pesos para estes neurônios intermediários?

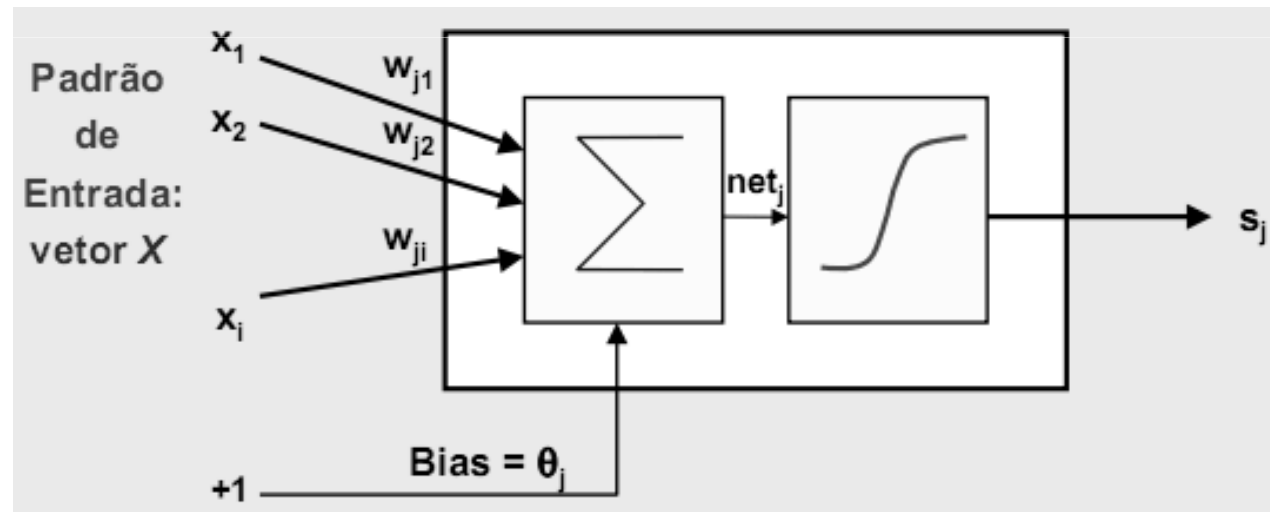


# **Modelos de Neurônios Artificiais**

- **O PERCEPTRON (O Renascimento)**
  - Apesar do descrédito gerado sobre a área da neurocomputação, entre 1969 e 1982 os estudos neste campo continuaram, ainda que englobadas em outras linhas de pesquisa
  - Os primeiros resultados da retomada do desenvolvimento sobre redes neurais foram publicados em 1986 e 1987, através dos trabalhos do grupo PDP (Parallel and Distributed Processing), onde ficou consagrada a técnica de treinamento por backpropagation.

# Modelos de Neurônios Artificiais

- O Modelo Moderno de NEURÔNIO ARTIFICIAL



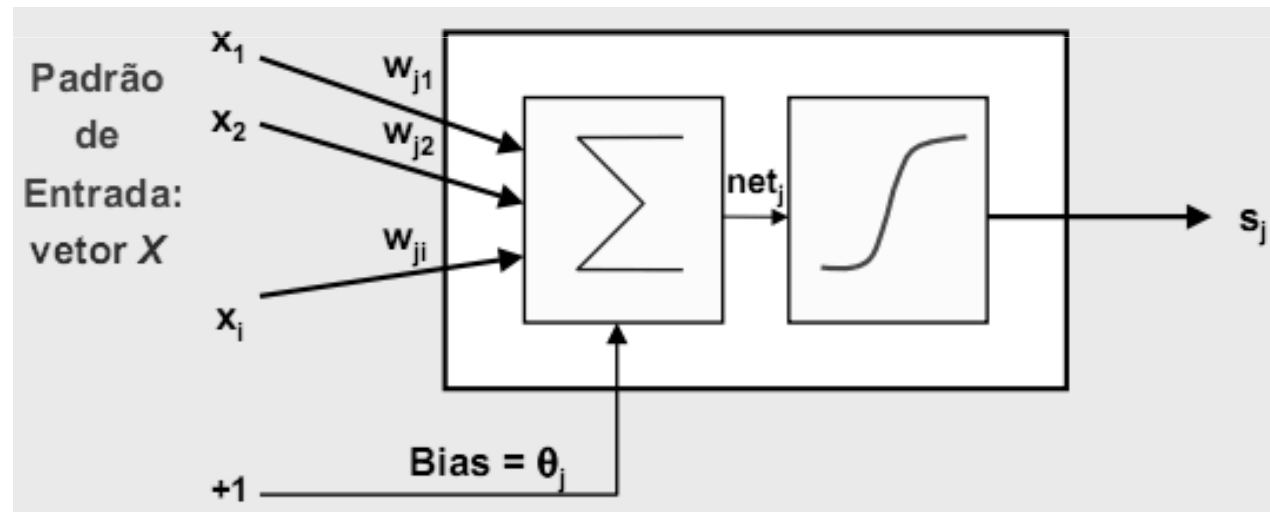


# **Modelos de Neurônios Artificiais**

- **O Modelo Moderno de NEURÔNIO ARTIFICIAL**
  - **As Entradas**
    - As entradas de um neurônio podem ser as saídas de outros neurônios, entradas externas, um bias ou qualquer combinação destes elementos.

# Modelos de Neurônios Artificiais

- O Modelo Moderno de NEURÔNIO ARTIFICIAL







# **Modelos de Neurônios Artificiais**

- **O Modelo Moderno de NEURÔNIO ARTIFICIAL**
  - **A Combinação das Entradas - O "Net"**
    - O somatório de todas estas entradas, multiplicadas por seus pesos, dá origem ao chamado "net" de um neurônio.
    - $w_{ji}$  é um número real que representa a conexão sináptica da entrada do  $j$ -ésimo neurônio com a saída do  $i$ -ésimo neurônio.
    - Após a determinação do net, o valor de saída do neurônio é produzido através da função de ativação.



# **Modelos de Neurônios Artificiais**

- **O Modelo Moderno de NEURÔNIO ARTIFICIAL**
  - **A Função de Ativação ou de Saída**
    - Essencialmente, qualquer função contínua e derivável pode ser utilizada como função de saída na modelagem neural.
    - As funções mais comumente utilizadas como funções de saída são:
      - A Função Linear
      - A Função Sigmoidal ou Logística
      - A Função Tangente Hiperbólica

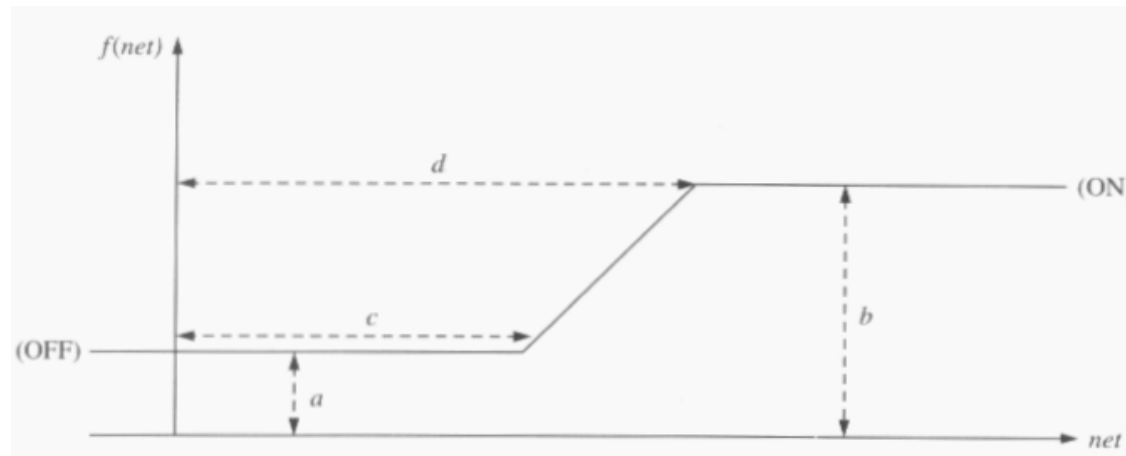
# Modelos de Neurônios Artificiais

- O Modelo Moderno de NEURÔNIO ARTIFICIAL

- A Função de Ativação ou de Saída

- A Função Linear

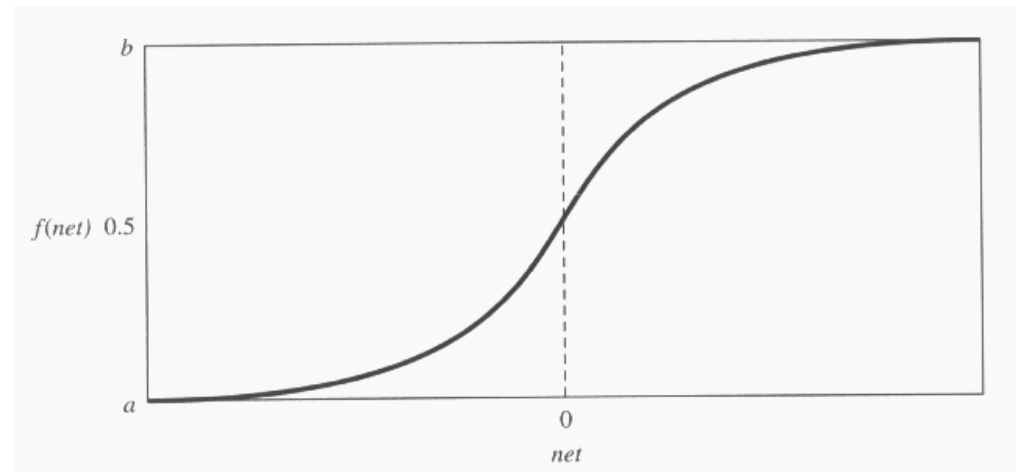
$$y(x) = ax$$



# Modelos de Neurônios Artificiais

- O Modelo Moderno de NEURÔNIO ARTIFICIAL
  - A Função de Ativação ou de Saída
    - A Função Sigmoidal

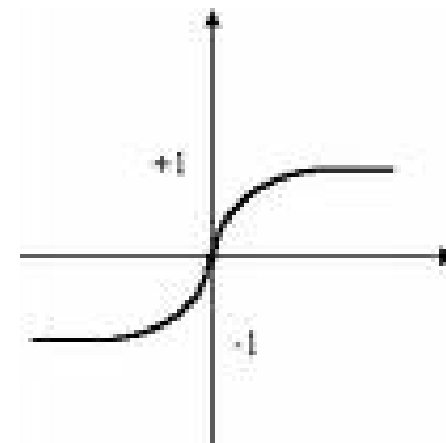
$$y(x) = \frac{1}{1 + e^{-kx}}$$



# Modelos de Neurônios Artificiais

- O Modelo Moderno de NEURÔNIO ARTIFICIAL
  - A Função de Ativação ou de Saída
    - A Função Tangente Hiperbólica

$$y(x) = \tanh(kx) = \frac{1 - e^{-kx}}{1 + e^{-kx}}$$





# **Topologias de Redes Neurais**

## **■ Como os neurônios se Conectam**

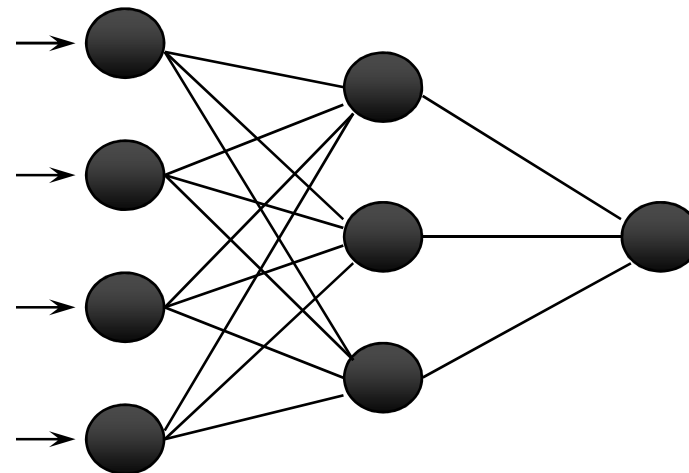
- Um único neurônio ou elemento processador é insuficiente para a maioria dos problemas práticos, e, redes com grande número de neurônios interconectados são freqüentemente utilizadas.
- A maneira como os neurônios são interconectados determina como as capacidades computacionais da rede e se constitui em uma das primeiras e mais importantes decisões no desenvolvimento de uma rede neural.

# **Topologias de Redes Neurais**

- **Como os neurônios se Conectam**
  - Redes Diretas - "Feedforward" - MLP (Multi-Layer Perceptrons)
    - Não existem ciclos.
    - Conexões para a "frente".
  - Redes Recorrentes ou Cíclicas
    - Existem ciclos.
    - Conexões para trás e/ou para os lados.

# Topologias de Redes Neurais

- **Redes Neurais Diretas-Feedforward**
  - **MLP (Multi-Layer Perceptrons)**
    - As redes diretas são aquelas cujo grafo não tem ciclos.
    - Frequentemente é comum representar estas redes em camadas e neste caso são chamadas redes de camadas.





# Topologias de Redes Neurais

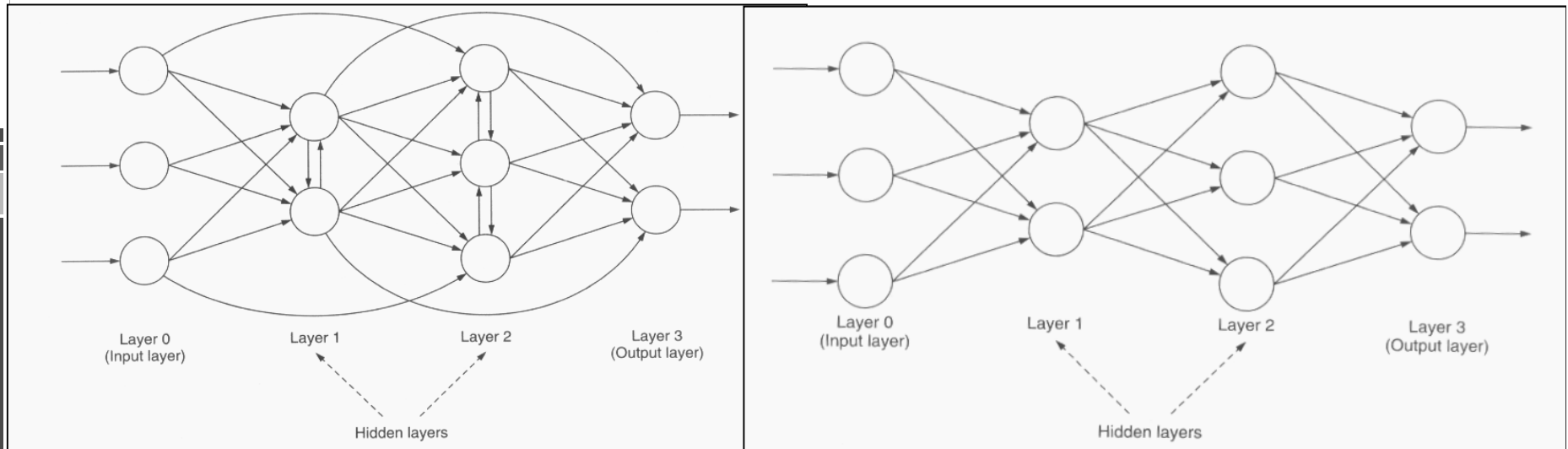
## ▪ Redes Neurais Diretas em Camadas

- Neurônios que recebem sinais de excitação do ambiente são chamados de camada de entrada ou primeira camada.
- Neurônios que tem sua saída como saída da rede pertencem a camada de saída ou última camada.
- Neurônios que recebem sinais apenas de outros neurônios e enviam suas saídas para outros neurônios, pertencem a camadas intermediárias ou escondidas (hidden).

# Topologias de Redes Neurais

## ■ Redes Neurais Recorrentes

- Redes recorrentes são aquelas no qual um neurônio de uma camada  $i$  pode possuir conexão com um ou mais neurônios de uma camada  $j$ , em que  $j \leq i$ .

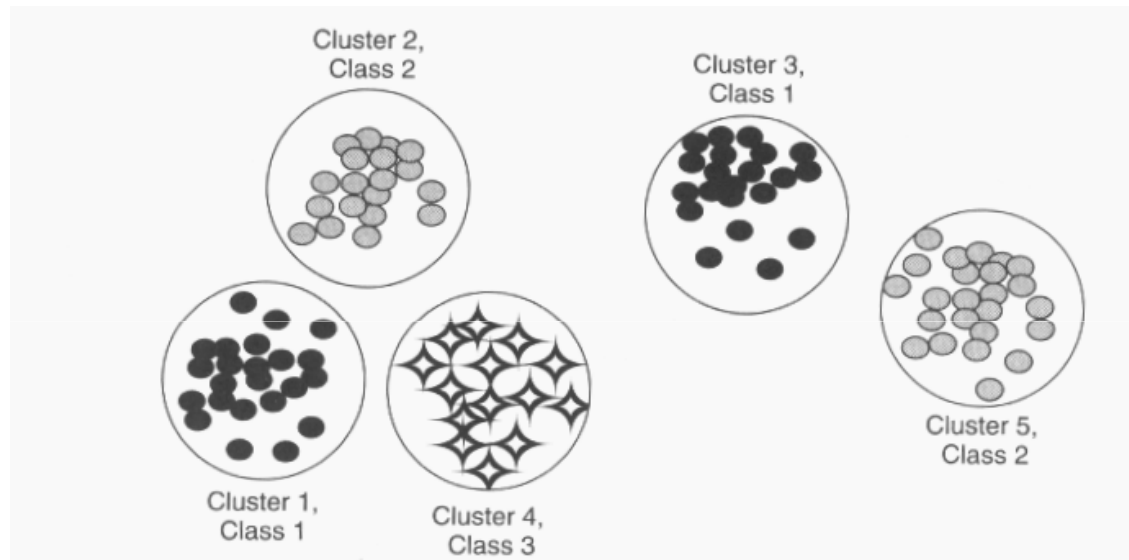


# Aprendizado de Redes Neurais

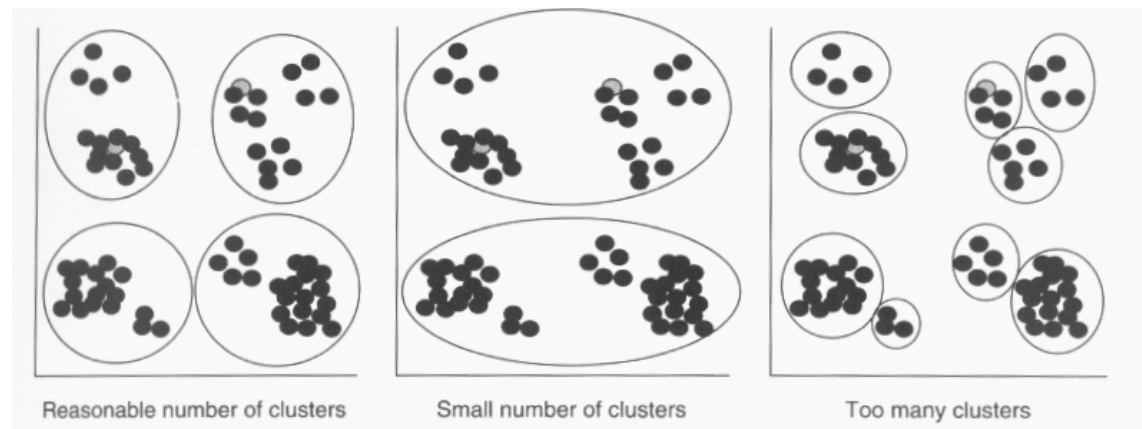
- Um neurônio é considerado ser um elemento adaptativo. Seus pesos sinápticos são modificáveis dependendo do algoritmo de aprendizado.
- Por exemplo, alguns, dependendo do sinal de entrada que recebem, tem seus valores de saída associados a uma resposta diante de um **aprendizado supervisionado** por uma espécie de "professor".
- Em alguns casos o sinal do "professor" não está disponível e não há informação de erro que possa ser utilizada para correção dos pesos sinápticos, assim o neurônio modificará seus pesos baseado somente no sinal de entrada e/ou saída, sendo o caso do **aprendizado não-supervisionado**.

# Aplicações de Redes Neurais

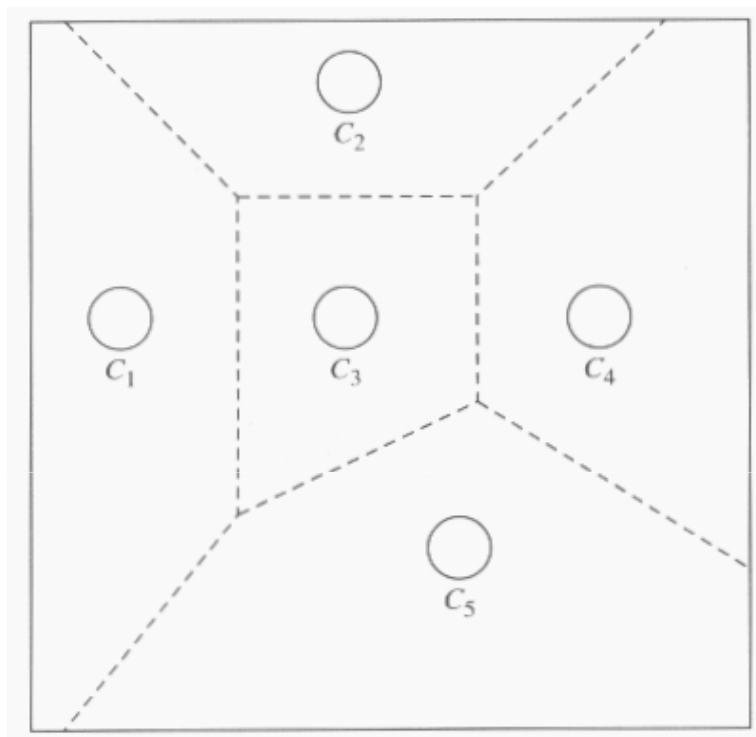
## CLASSIFICAÇÃO



## CLUSTERIZAÇÃO

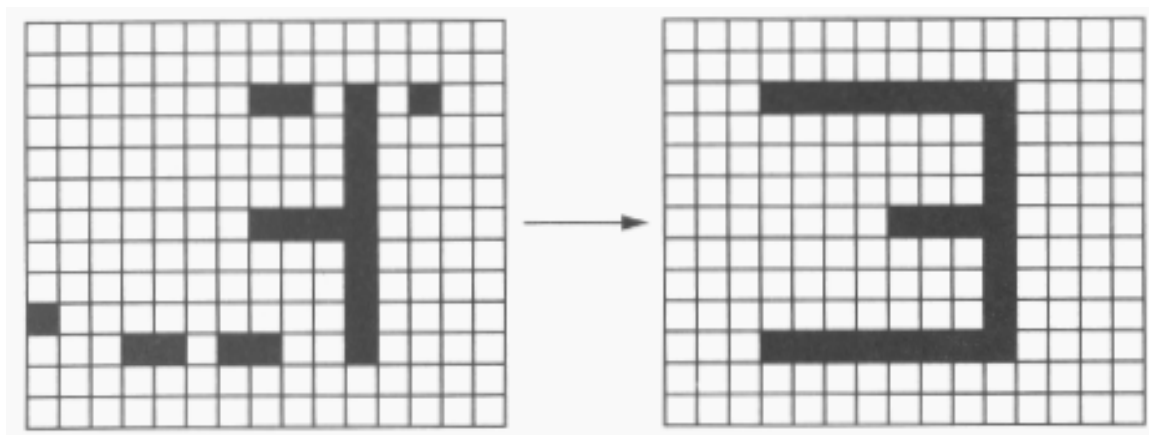


# Aplicações de Redes Neurais



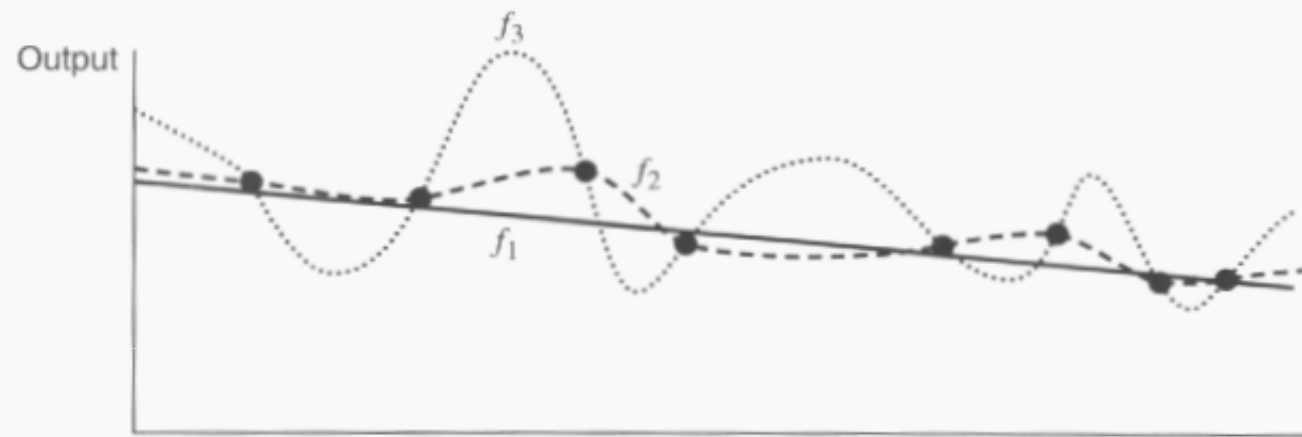
QUANTIZAÇÃO  
VETORIAL

ASSOCIAÇÃO  
DE PADRÕES

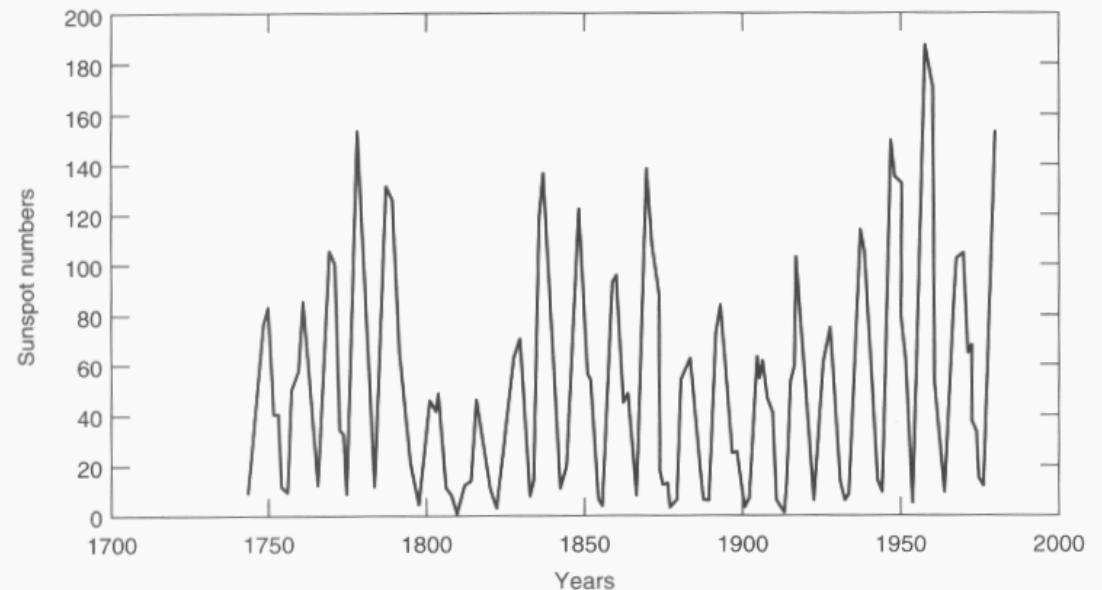


# Aplicações de Redes Neurais

## APROXIMAÇÃO DE FUNÇÕES

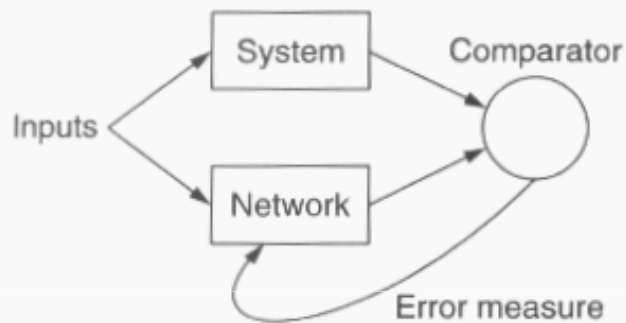


## PREVISÕES DE SÉRIES TEMPORAIS

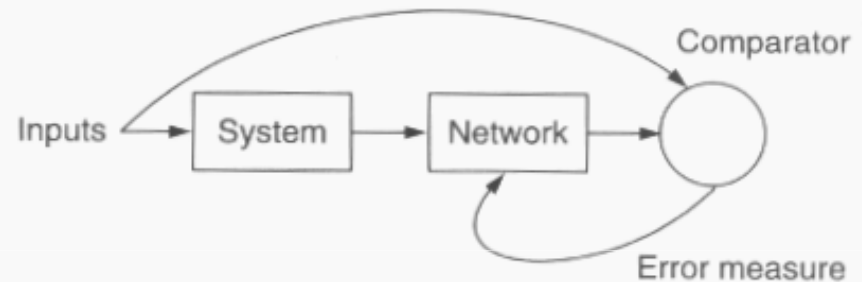


# Aplicações de Redes Neurais

## APLICAÇÕES EM CONTROLE



(a) Forward system identification



(b) Inverse system identification



**FIM DO MÓDULO  
INTRODUTÓRIO**

