Redes Neurais

Tópico

- Aprendizagem
  - *Conceitos Básicos*
  - *Redes Neurais*

Aprendizagem

Aprendizagem (realimentação - *feedback*):

- Aprendizagem **supervisionada**
- Aprendizagem **não supervisionada**
- Aprendizagem **por reforço**

Aprendizagem

Aprendizagem Supervisionada:

- Ocorre nas situações em que é possível perceber tanto as entradas como as saídas;
- Freqüentemente as saídas são fornecidas por um supervisor(especialista) humano;
- Envolve aprendizagem de uma **função** a partir de exemplos de suas entradas e saídas.

Aprendizagem Não-Supervisionada:

- Envolve a aprendizagem de padrões na entrada, quando não são fornecidos valores de saída específicos.

Aprendizagem por Reforço:

- O agente deve aprender a partir do reforço (recompensa).

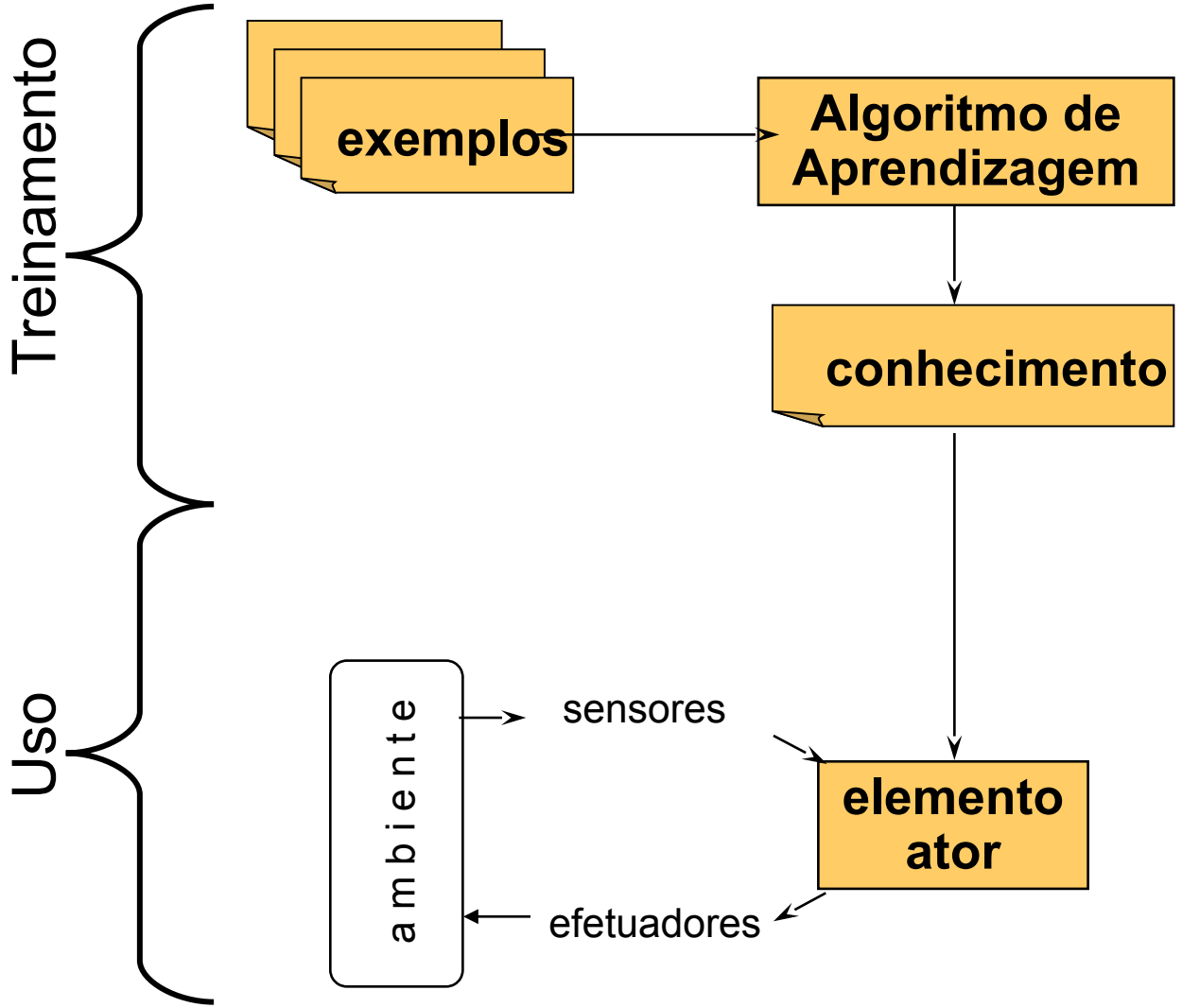
Aprendizagem

- **Tipos**

  Segundo Ginsberg(1993), aprendizagem pode ser dividida em:

  - Aprendizagem por Descoberta;
  - Aprendizagem por Generalização;
    - Aprendizagem Dedutiva;
    - **Aprendizagem Indutiva**.

Modelo do Agente Aprendiz (*off-line*)



Por que a aprendizagem funciona?

- *“É quase certo que qualquer hipótese que esteja seriamente errada será "desmascarada" com alta probabilidade após um pequeno número de exemplos, por que fará uma previsão incorreta. Desse modo, qualquer hipótese que seja consistente com um conjunto bastante grande de exemplos de treinamento terá pouca probabilidade de estar seriamente errada: isto é, ela deve estar* ***provavelmente aproximadamente correta*** *(aprendizagem PAC).*

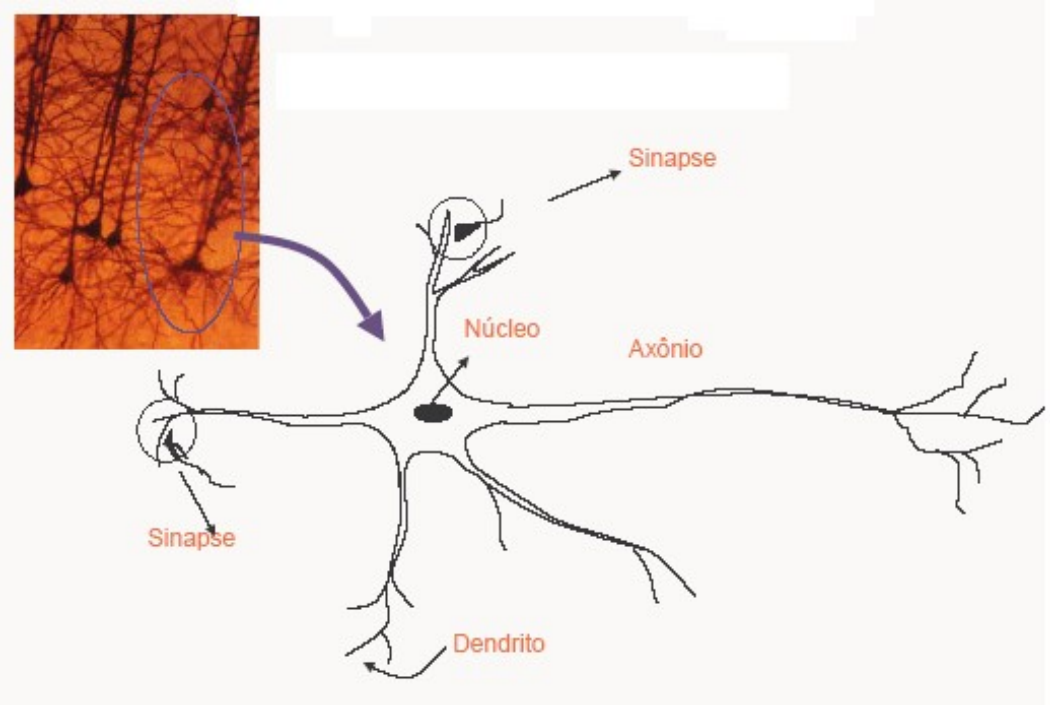
- *Para tanto, os conjuntos de treinamento e teste devem ser extraídos ao acaso e de forma independente da população de exemplos com a mesma distribuição de probabilidade - suposição de* ***hipótese de estacionariedade****.”*

(Russel e Norvig)

Redes Neurais

- É uma célula do cérebro cuja principal função é coletar, processar e disseminar sinais elétricos.
- Acredita-se que a capacidade de processamento de informações no cérebro vem de redes de neurônios.

Componentes de um neurônio



Redes Neurais

Componentes de um Neurônio

- **Axônio** – transmissão de sinais a partir do corpo celular; poucas ramificações e compridos;
- **Dendritos** – conduzem sinais para a célula; têm muitas ramificações (zonas receptivas);
- **Sinapses** – estruturas funcionais elementares que mediam as conexões entre os neurônios

Redes Neurais

- **Plasticidade de um neurônio** – capacidade de adaptação ao ambiente.

- Mecanismos de Plasticidade (cérebro de um adulto)
  - Criação de novas conexões sinápticas
  - Modificação das sinapses existentes

- Plasticidade - essencial para as Redes Neurais Artificiais

Neurociência computacional

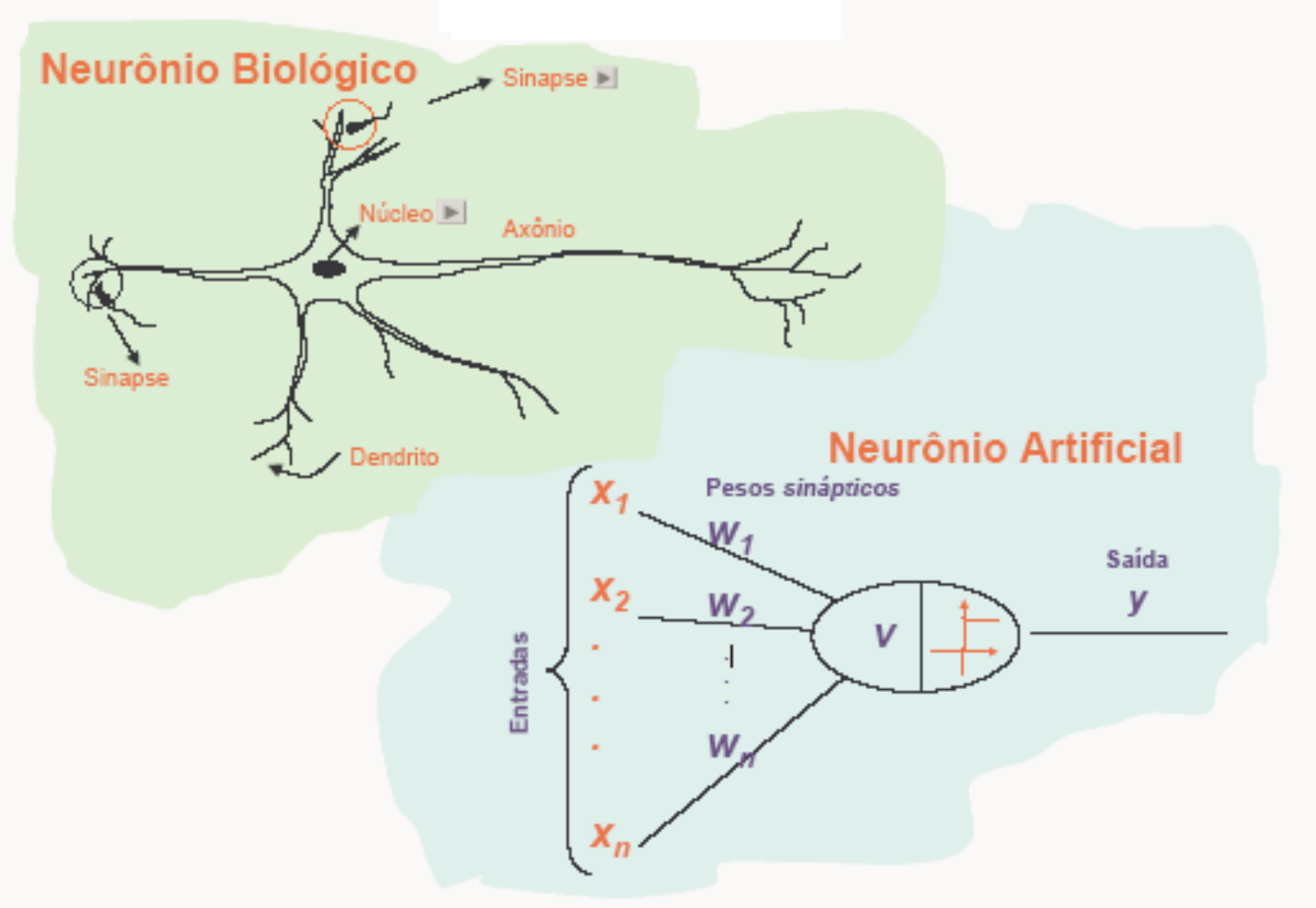
- Modelo matemático do neurônio criado por McCulloch e Pitts (1943)

- Interesse nas propriedades mais abstratas da RNs
  - Habilidade para executar computação distribuída
  - Habilidade para tolerar entradas ruidosas
  - **Habilidade para aprender**

- **Uma das formas mais populares e efetivas de sistemas de aprendizagem**

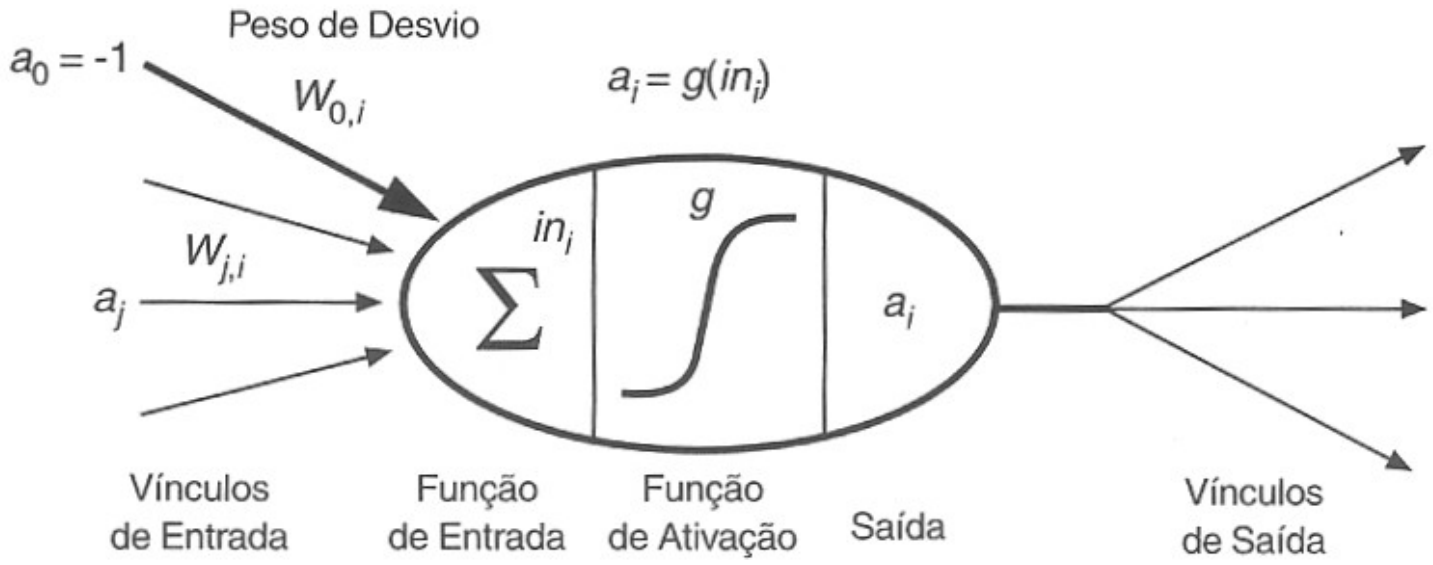
Redes Neurais Artificiais

Inspiração Biológica



Neurônio Artificial

- "Dispara" quando uma combinação linear de suas entradas excede algum limiar.



Redes Neurais Artificiais

- Cada unidade *i* calcula
  - Uma soma ponderada de suas entradas

$$in_i = \sum_{j=0}^{n} w_{j,i} a_j$$

  - Depois, aplica uma **<u>função de ativação</u>** g a essa soma para derivar a saída

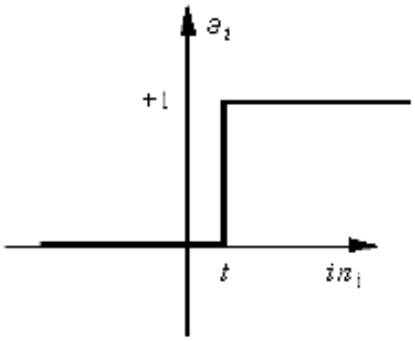
$$a_i = g(in_i) = g\left(\sum_{j=0}^{n} w_{j,i} a_j\right)$$

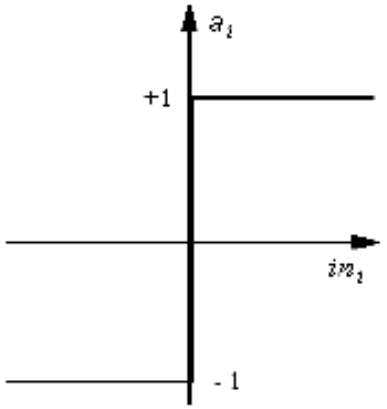
A função de ativação

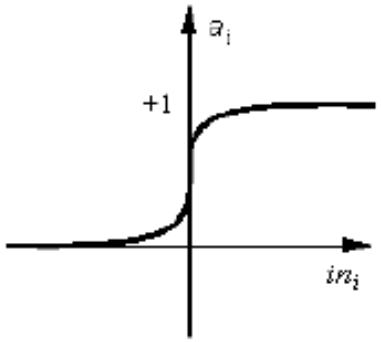
- g é projetada para atender a dois propósitos
  - Deseja-se a unidade “ativa” (próxima de 1) quando as entradas “corretas” forem recebidas e “inativa” (próxima de 0) quando as entradas “erradas” forem recebidas

Aprendizagem

Funções de Ativação:







$$step_t(x) = \begin{cases} 1, x \geq t \\ 0, x < t \end{cases}$$

t = limiar

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

$$sign(x) = \begin{cases} +1, x \geq 0 \\ -1, x < 0 \end{cases}$$

Aprendizagem

Rede de memória associativa com pesos fixos

- Aprende a associar valores de entrada com valores de saída
- A matriz de pesos é determinada pela multiplicação do vetor de entrada com o de saída

$$w_{i,j} = \sum_{m} a_i^{(m)} . b_j^{(m)}$$

a = vetores de saída no conjunto de treinamento
b= vetores de entrada
j = índice dos links de entrada
i = índice do neurônios de saída
m = índice do conjunto de treinamento

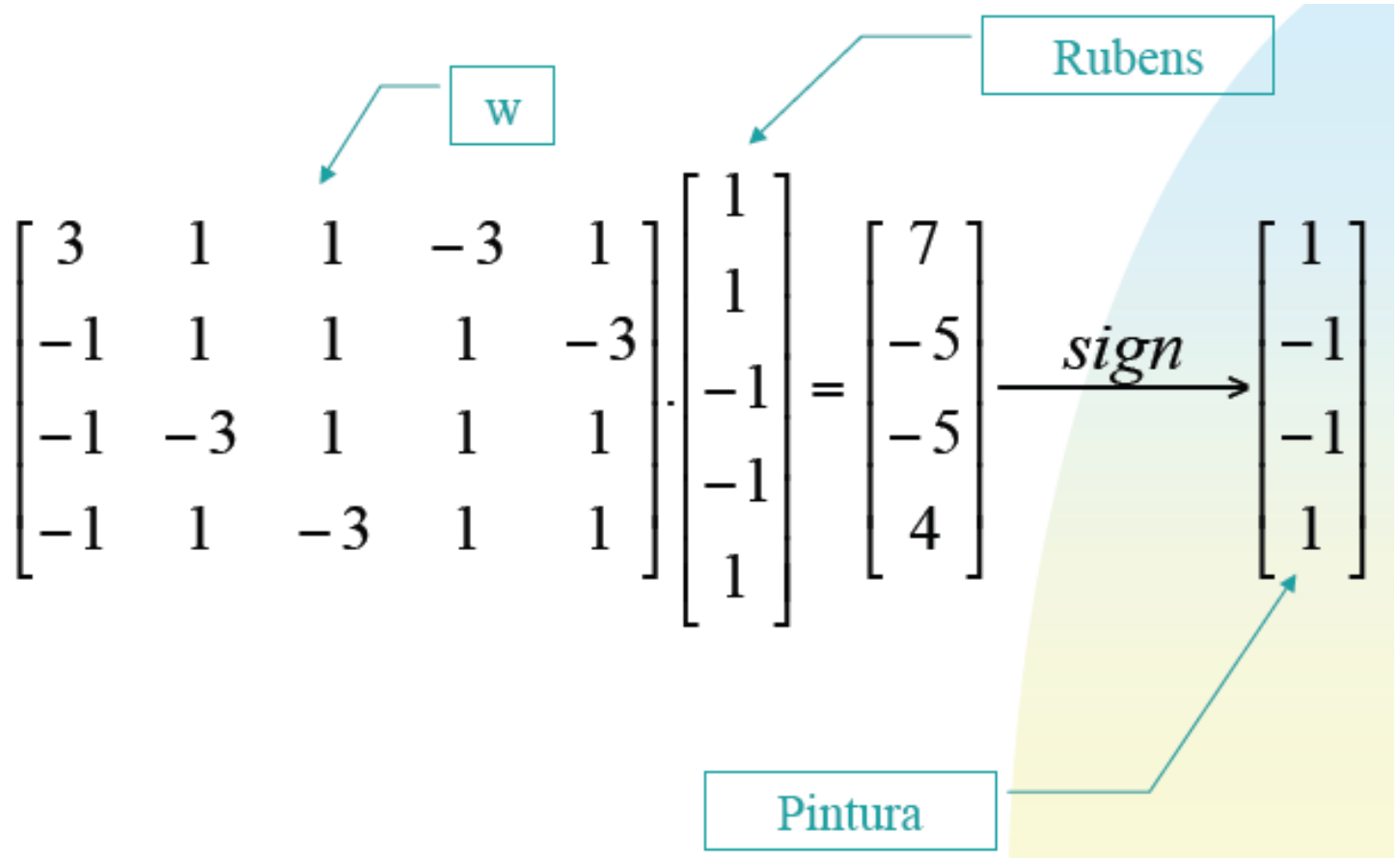
Aprendizagem - Ilustração

- A RN deve aprender as seguintes associações
  - Homero - Literatura
  - Bethoven - Música
  - Rubens - Pintura
- Representação binária dos valores
  - Homero [1,-1,1,-1,1] = b(1)
  - Literatura [1, -1, 1, -1] = a(1)
  - Bethoven [1,1,1,-1,-1] = b(2)
  - Música [1,1,-1,-1] = a(2)
  - Rubens [1,1,-1,-1,1] = b(3)
  - Pintura [1,-1,-1,1] = a(3)

- Variáveis: m = 1,..,3; j=1,..,5; i=1,..,4
- Usar como função de transferência: ***sign***
- Calcular W - alguns exemplos:

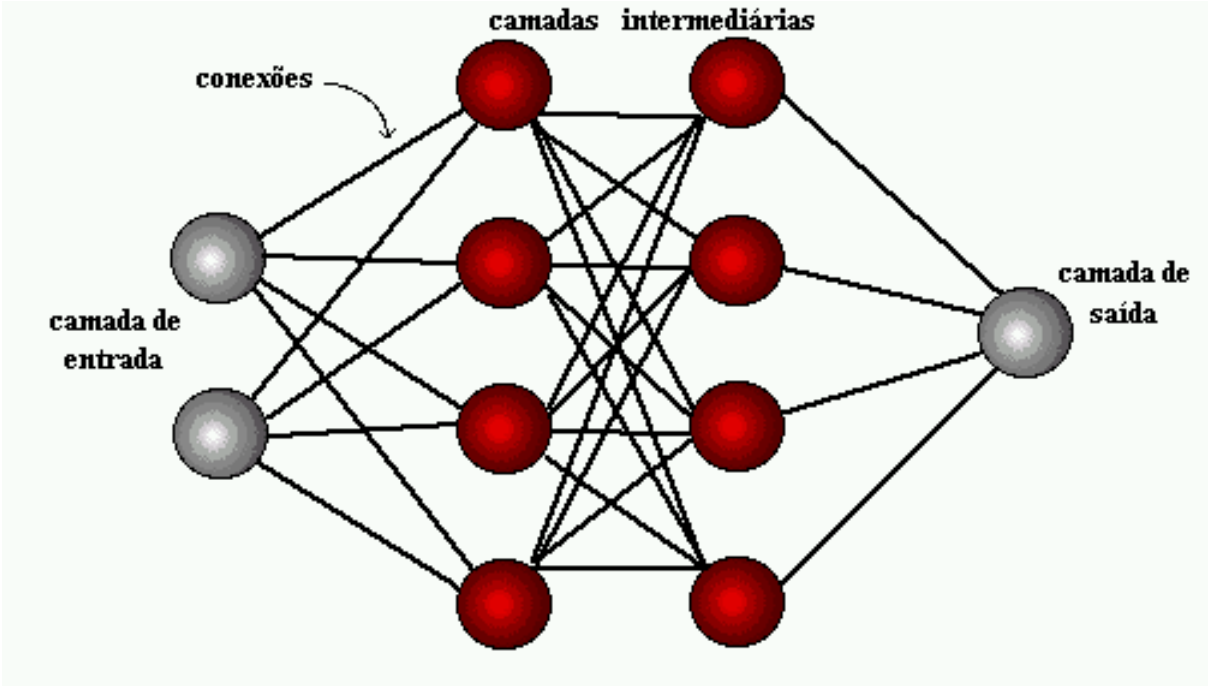
$$w_{1,1} = a_1^{(1)} b_1^{(1)} + a_1^{(2)} b_1^{(2)} + a_1^{(3)} b_1^{(3)}$$
$$= 1.1 + 1.1 + 1.1 = 3$$
$$w_{1,2} = a_1^{(1)} b_2^{(1)} + a_1^{(2)} b_2^{(2)} + a_1^{(3)} b_2^{(3)}$$
$$= 1.-1 + 1.1 + 1.1 = 1$$

Aprendizagem - Ilustração



Redes Neurais Artificiais

Rede Neural - Organização em Camadas (Exemplo)



Redes Neurais Artificiais

- As camadas são classificadas em três grupos (usualmente):
  - **Camada de Entrada**: onde os padrões são apresentados à rede;
  - **Camadas Intermediárias ou Escondidas**: onde é feita a maior parte do processamento, através das conexões ponderadas; podem ser consideradas como extratoras de características;
  - **Camada de Saída**: onde o resultado final é concluído e apresentado.

Redes Neurais Artificiais

Redes Neurais: Classificação dos Modelos Conexionistas

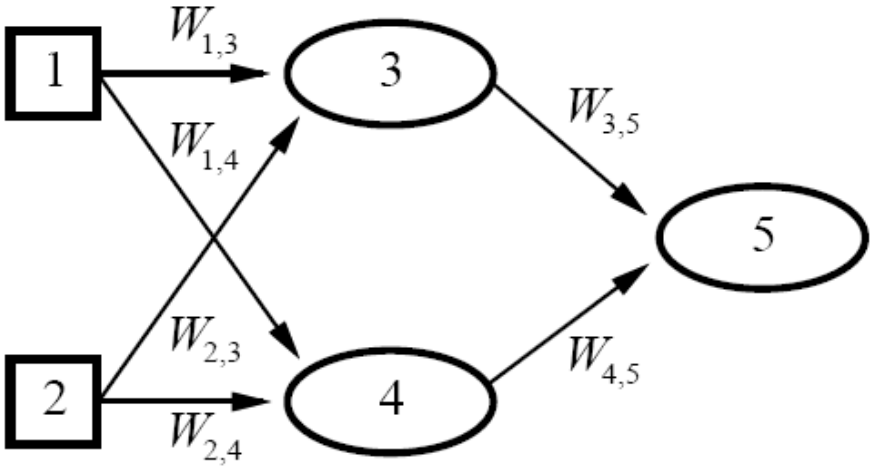
- Em relação à estrutura da rede:
  - Redes de uma única camada
  - Redes de múltiplas camadas
  - Redes do tipo uni-direcional
  - Redes do tipo recorrentes
  - Redes com estrutura estática (não altera a sua estrutura)
  - Redes com estrutura dinâmica (altera a estrutura)
  - Redes com conexões de ordem superior

Estruturas de rede

- Redes acíclicas ou redes de alimentação direta:
  - Representam uma função de sua entrada atual;
  - NÃO têm nenhum estado interno além dos pesos.
- Redes cíclicas ou redes recorrentes:
  - Utilizam suas saídas para realimentar suas próprias entradas;
  - Níveis de ativação da rede formam um sistema dinâmico
    - Pode atingir um estado estável ou exibir oscilações;
  - A resposta da rede a uma determinada entrada pode depender de entradas anteriores (como um flip-flop).

Redes de alimentação direta

- Representa uma função de suas entradas



- Dados os valores de entrada a_1 e a_2 a rede calcula:

$$a_5 = g(W_{3,5}a_3 + W_{4,5}a_4)$$
$$= g(W_{3,5}g(W_{1,3}a_1 + W_{2,3}a_2) + W_{4,5}g(W_{1,4}a_1 + W_{2,4}a_2))$$

Redes de alimentação direta

- Expressando a saída de cada unidade oculta como uma função de suas entradas percebe-se que:
  - a_5 é uma função das entradas da rede;
  - Os pesos da rede atuam como **parâmetros** dessa função;
  - A rede calcula h_w**(x);**
  - Ajustando os pesos muda-se a função que a rede representa;
  - **Aprendizagem!**

Redes Neurais Artificiais

Processos de Aprendizado

- A propriedade mais importante das redes neurais é a habilidade de aprender com seu ambiente e com isso melhorar seu desempenho.
- Isso é feito a partir de um processo iterativo de ajustes aplicado a seus pesos, o **treinamento**.
- O aprendizado ocorre quando a rede neural atinge uma solução generalizada para uma classe de problemas.

Redes Neurais Artificiais

Algoritmo de Aprendizado

- Conjunto de regras bem definidas para a solução de um problema de aprendizado.
- Existem muitos tipos de algoritmos de aprendizado específicos para determinados modelos de redes neurais.
- Os algoritmos diferem entre si principalmente pelo modo como os pesos são modificados.

Redes Neurais Artificiais

Redes Neurais: Classificação - Aprendizado

- Em relação ao aprendizado:
  - Aprendizado supervisionado
  - Aprendizado semi-supervisionado
  - Aprendizado não supervisionado
  - Aprendizado instantâneo
  - Aprendizado por pacotes
  - Aprendizado contínuo
  - Aprendizado ativo
  - Aprendizado: aproximação de funções
  - Aprendizado: classificação
  - Usar apenas uma base de exemplos de aprendizado
  - Usar uma base de aprendizado e uma base de teste de generalização

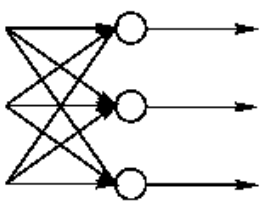
Redes Neurais Artificiais

Redes Neurais: Classificação dos Modelos Conexionistas

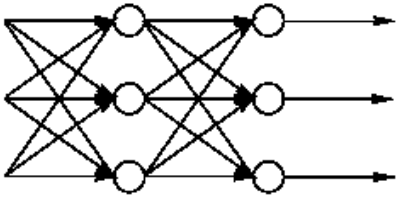
- Em relação as unidades da rede:
  - Redes baseadas em Perceptrons (MLP - *Multi-Layer Perceptron*)
  - Redes baseadas em Protótipos (RBF - *Radial Basis Function*)

Redes Neurais Artificiais

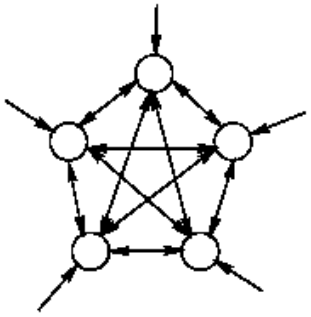
Exemplos de Redes Neurais



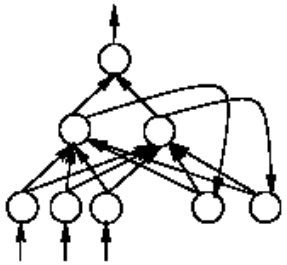
a) Single-Layer Perceptron



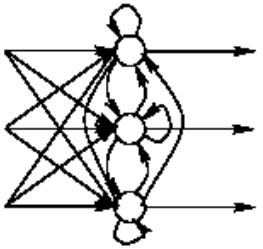
b) Multi-Layer Perceptron



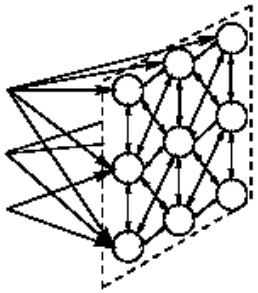
c) Hopfield network



d)Elman recurrent network



e) Competitive networks



f) Self-Organizing Maps

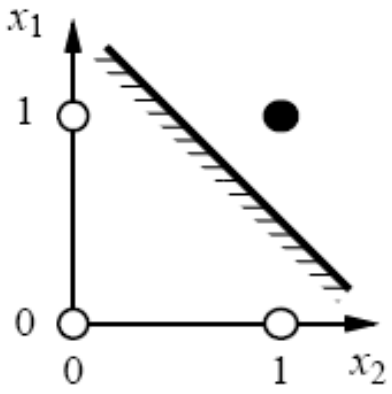
Redes Neurais Artificiais

Aplicações Práticas

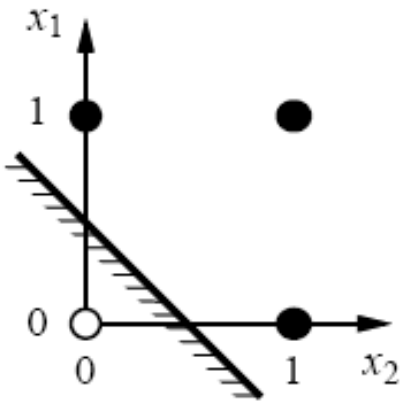
- Reconhecimento de Padrões: Caracteres, Imagens, Voz, etc
- Sistemas de Auxílio ao Diagnóstico: Médico, Falhas Mecânicas, etc
- Robótica Inteligente
- Previsão: Tempo, Cotações da Bolsa de Valores, etc
- Sistemas de Controle
- Processamento de Sinais
  - Processamento de Linguagem Natural
- *Data Mining*

Exemplos de Aplicações

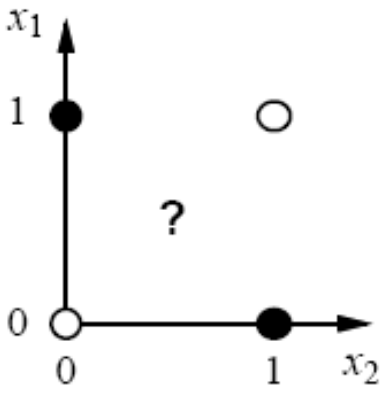
- O perceptron de limiar é chamado **separador linear**
  - Porque traça um plano entre os pontos de entrada em que a saída é zero ou um



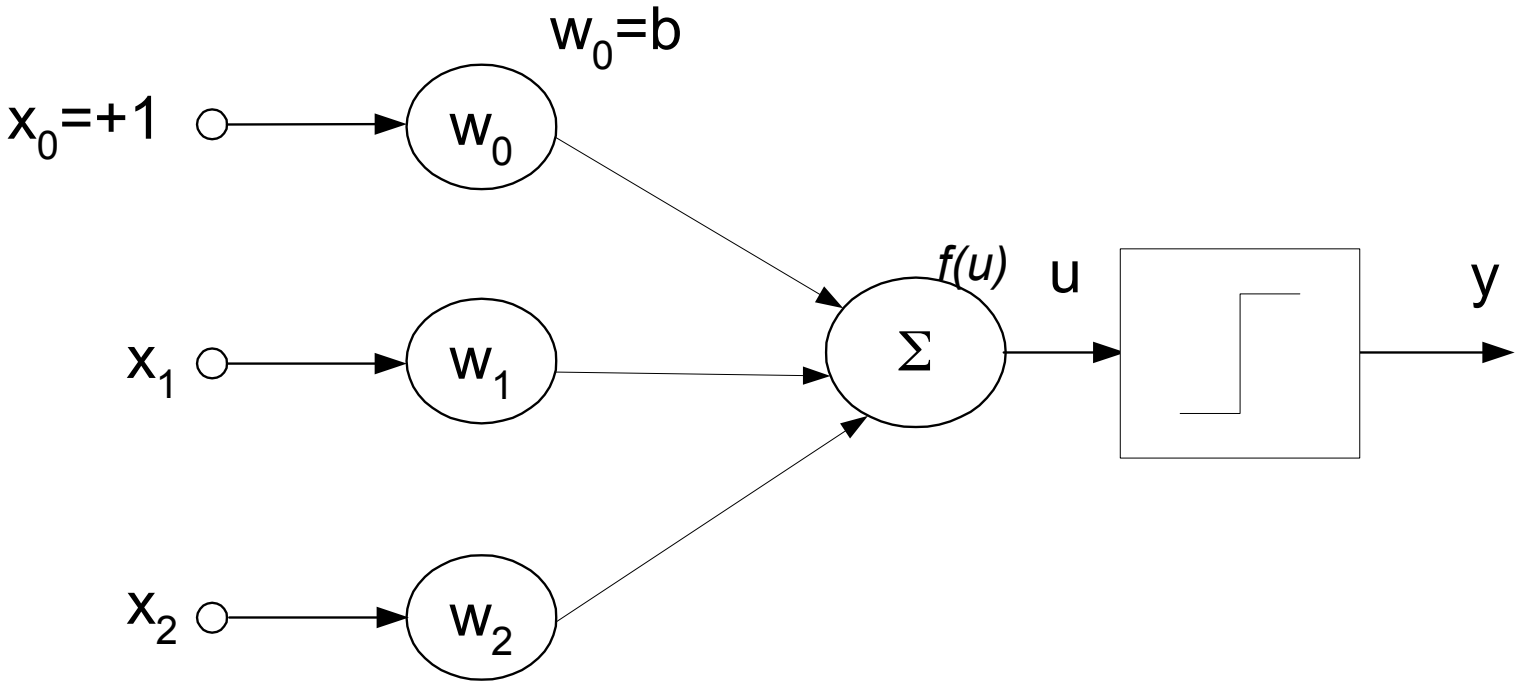
(a) x_1 **and** x_2



(b) x_1 **or** x_2



(c) x_1 **xor** x_2



$$y = f\left(w_1 x_1 + w_2 x_2 + w_0\right), \text{ sendo } \begin{cases} f(u) = 1 & \text{se } u \geq 0 \\ f(u) = 0 & \text{se } u < 0 \end{cases}$$

Com os parâmetros w_0, w_1 e w_2, a função *f(u)* separa o espaço de entradas em duas regiões, usando uma linha reta dada por:

$$w_1 x_1 + w_2 x_2 + w_0 = 0$$

Exemplos de Aplicações

Um neurônio para a função AND de 2 entradas

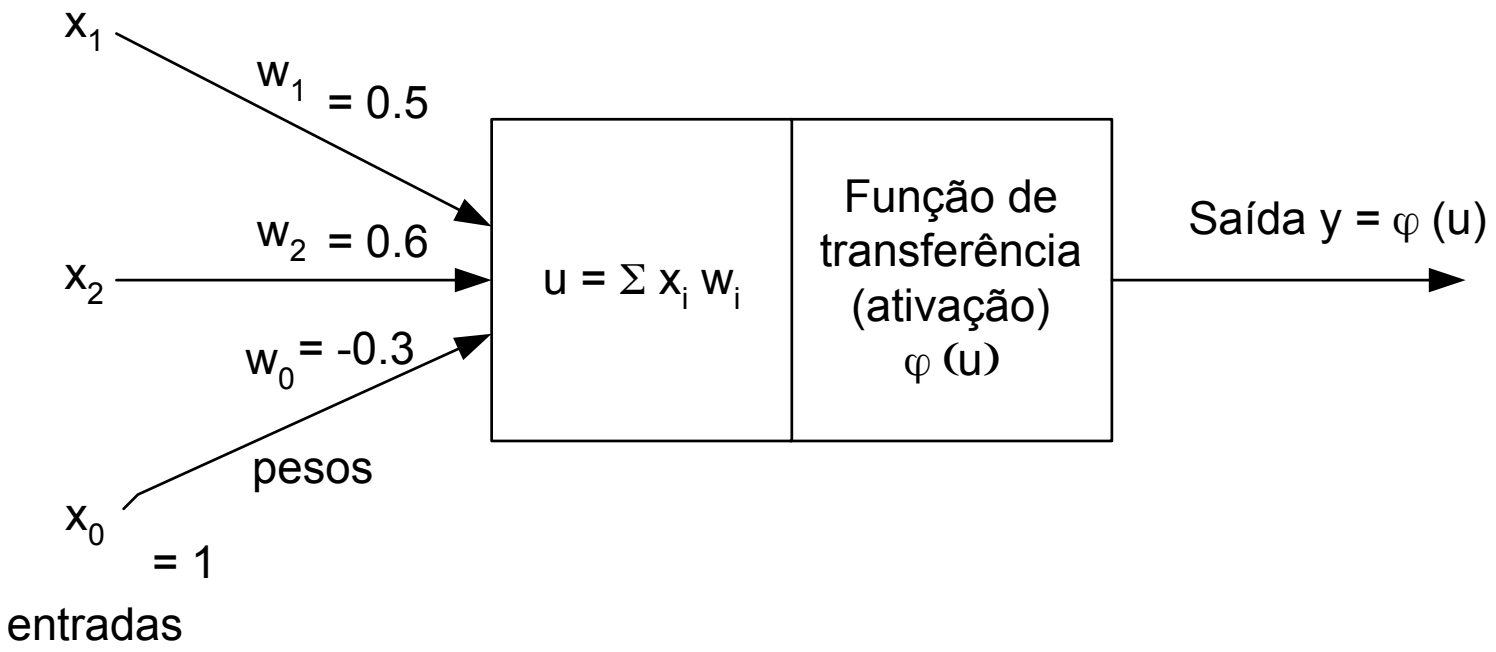
Iniciamos o neurônio com os pesos 0.5 e 0.6 para as duas entradas, e -0.3 para o limiar (w_0). Isso é equivalente à equação:

$$u = 0.5\ x_1 + 0.6\ x_2 - 0.3\ x_0$$

x_0 é a entrada estável sempre igual a 1.

Assim, a saída y deve disparar quando u >= 0).

Neurônio para a função AND de 2 entradas



x1	x2	u	y
0	0	-0.3	0
0	1	0.3	1
1	0	0.2	1
1	1	0.8	1

a saída é 1 para os padrões 01, 10 e 11, enquanto desejamos que a saída seja 1 somente para o padrão 11.

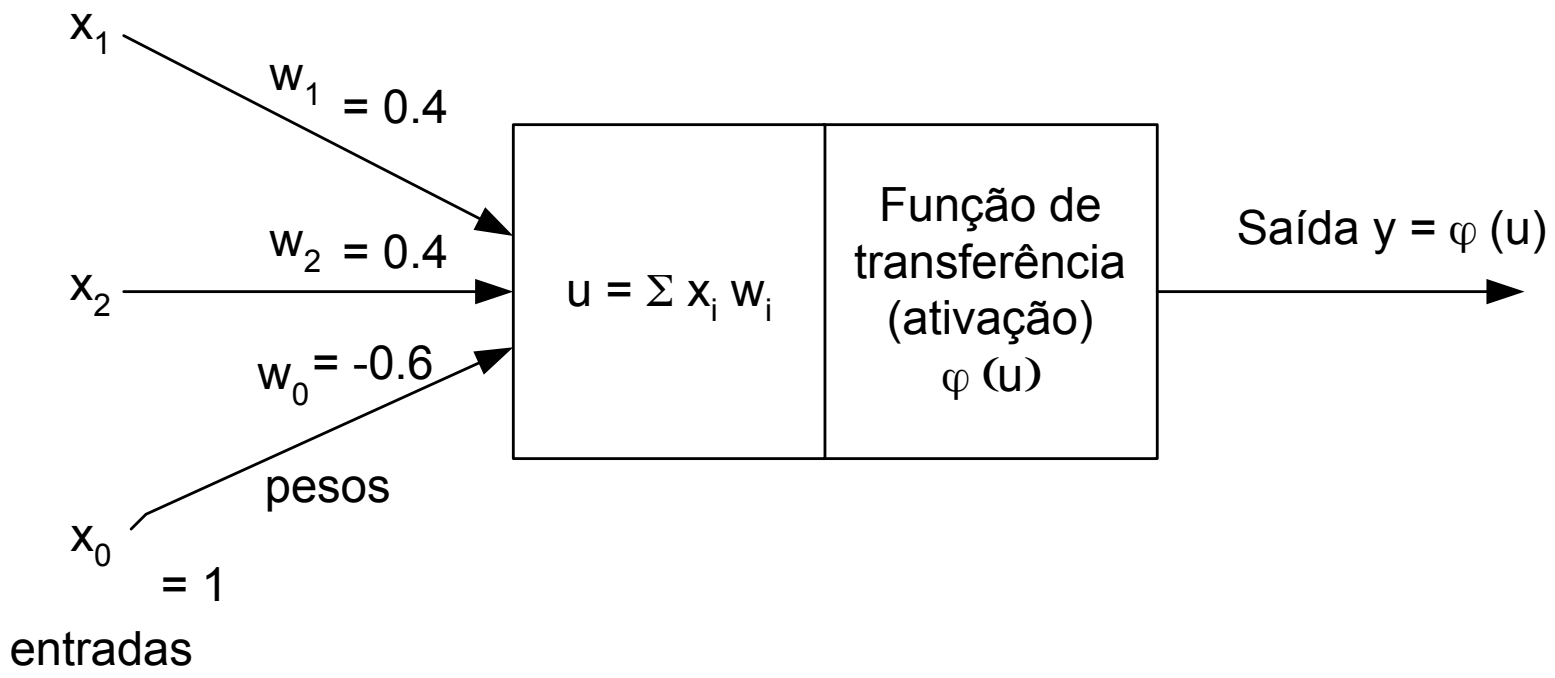
Neurônio para a função AND de 2 entradas

Seqüência de passos na aplicação do algoritmo

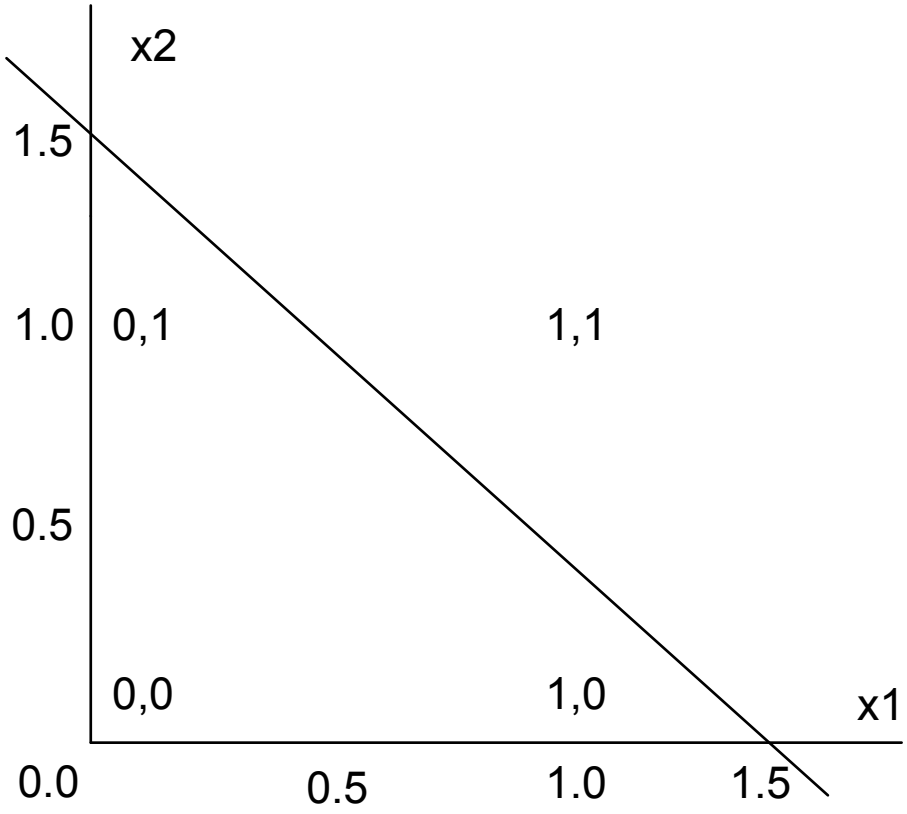
Inicio			w1 = 0.5 w2 = 0.6 w0 = -0.3
Entrada 0 0	u = - 0.3	y = 0 correta	
Entrada 0 1	u = 0.3	y = 1 incorreta	
Correção dos pesos de 0.1 para baixo →			w1 = 0.5 w2 = 0.5 w0 = -0.4
Entrada 1 0	u = 0.1	y = 1 incorreta	
Correção dos pesos de 0.1 para baixo →			w1 = 0.4 w2 = 0.5 w0 = -0.5
Entrada 1 1	u = 0.4	y = 1 correta	
Entrada 0 0	u = - 0.5	y = 0 correta	
Entrada 0 1	u = 0	y = 1 incorreta	
Correção dos pesos de 0.1 para baixo →			w1 = 0.4 w2 = 0.4 w0 = -0.6
Entrada 1 0	u = - 0.2	y = 0 correta	
Entrada 1 1	u = 0.2	y = 1 correta	
Entrada 0 0	u = -0.6	y = 0 correta	
Entrada 0 1	u = - 0.2	y = 0 correta	
Entrada 1 0	u = - 0.2	y = 0 correta	
Entrada 1 1	u = 0.2	y = 1 correta	
Fim			w1 = 0.4 w2 = 0.4 w0 = -0.6

Neurônio para a função AND de 2 entradas

Resultado do aprendizado



Neurônio para a função AND de 2 entradas



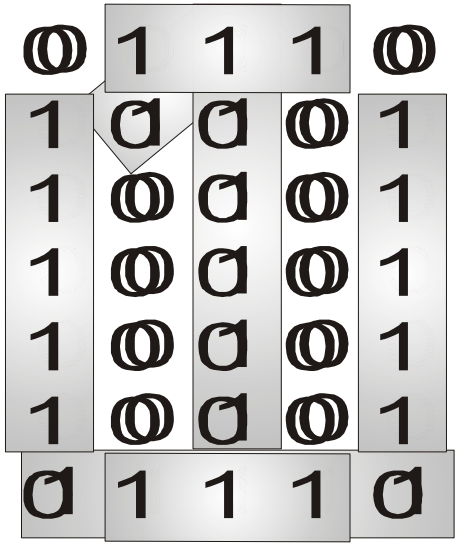
A reta 0.4 x1 + 0.4 x2 - 0.6 = 0 separa os pontos 00, 01 e 10, do ponto 11.

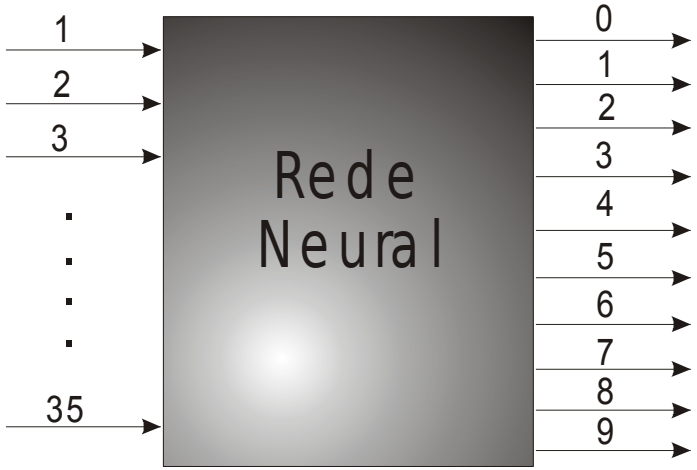
Exemplos de Aplicações

Reconhecimento de Dígitos

Parâmetros

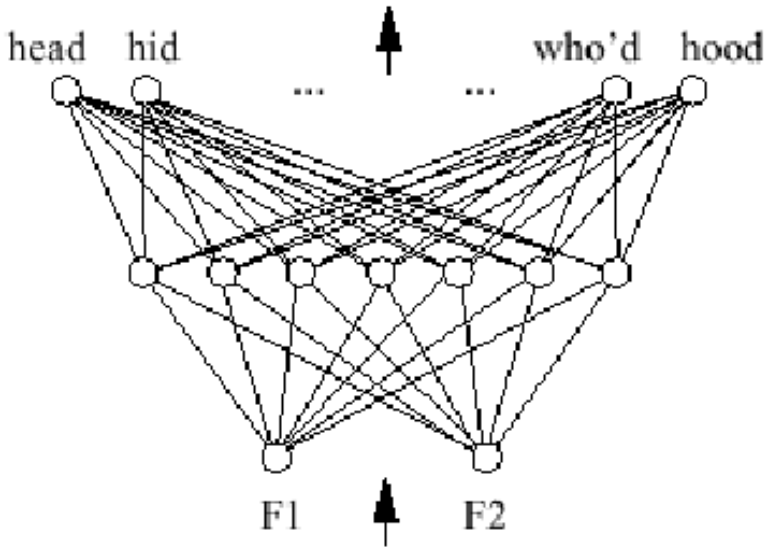
- rede 35 x 4 x 10
- entrada grid 5x7
- saída 10 bits

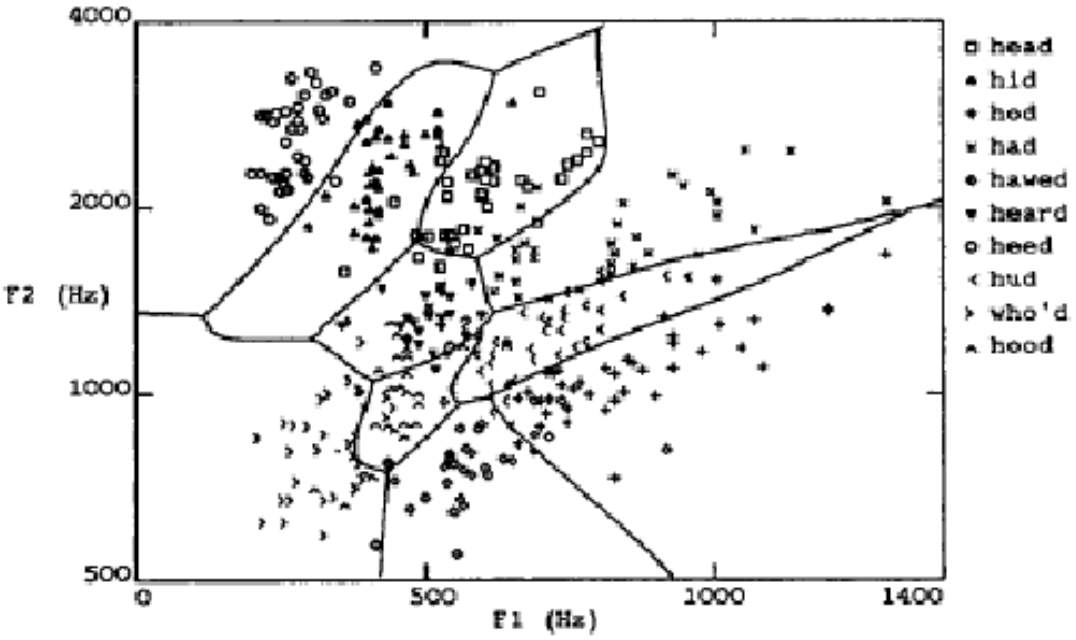




Exemplos de Aplicações

- **Exemplo**: Reconhecimento de fala – reconhecer 1 de 10 vogais entre h_d (em inglês)
  - – F1 e F2: parâmetros retirados da análise espectral do som da palavra.





Redes Neurais Artificiais

Aplicações Práticas

- http://fbim.fh-regensburg.de/~saj39122/begrolu/kohonen.html
- http://www.nd.com/neurosolutions/products/ns/nnandnsvideo.html

Exemplos de Aplicações

- Neural Networks Java Applets:
  - Perceptron Learning Rule (CNNL)
  - Image Compression Using Backprop (CNNL)
  - Generalizations of the Hamming Associative Memory (CNNL)

- Joone - Java Object Oriented Neural Engine
  - Porta XOR

- Simulador de Redes Neurais: SNNS (*Stuttgart Neural Network Simulator*)

Fonte: Java Demonstrations of Neural Net Concepts, http://neuron.eng.wayne.edu/software.html

Redes Neurais Artificiais

Vantagens

- Aquisição automática de conhecimentos empíricos a partir de uma base de exemplos de aprendizado referente a um problema;
- Manipulação de dados quantitativos, aproximados e mesmo incorretos com uma degradação gradual das respostas;
- Grande poder de representação de conhecimentos através da criação de relações ponderadas entre as entradas do sistema;

Redes Neurais Artificiais

Desvantagens

- Dificuldade de configuração das redes em relação à sua estrutura inicial e também no que se refere aos parâmetros dos algoritmos de aprendizado;
- Dificuldade de explicitar os conhecimentos adquiridos pela rede a partir de uma linguagem compreensível para um ser humano;
- Dificuldade de convergência (bloqueios) e instabilidade, inerentes aos algoritmos de otimização empregados;
- “Lentidão” do processo de aprendizado / adaptação.