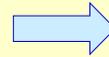


PSI 2533 – Módulo de Redes Redes Neurais Artificiais e Aprendizado

Prof. Emilio Del Moral Hernandez
Escola Politécnica da Universidade de São Paulo
Departamento de Eng. De Sistemas Eletrônicos



emilio_del_moral@ieee.org
www.lsi.usp.br/~emilio

Inteligência Computacional e Neurocomputação

- Neurocomputação
- Inteligência Artificial Simbólica
- Fuzzy Logic
- Algoritmos Genéticos
- Classificação Estatística
- Filtros Adaptativos
- Otimização
- Identificação de modelos paramétricos
- Aprendizado de máquina
- Sistemas para o apoio à decisão
- Processamento de informações não estruturadas
- Processamento / reconhecimento / classificação de sinais e imagens
- Interfaces homem-máquina naturais
- Modelagem de sistemas não lineares multivariáveis
- Bifurcação e dinâmica caótica no contexto de neurocomputação

Grupo **ICONE** - Inteligência Computacional, Modelagem e Neurocomputação Eletrônica

www.lsi.usp.br/ICONE



PSI-2533-2013– Prof. Emilio Del Moral

3

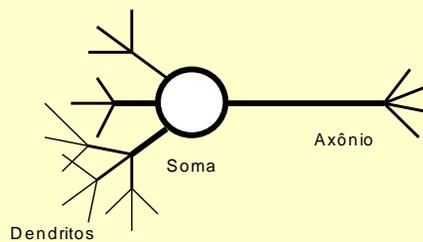
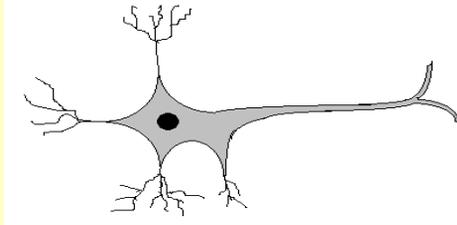
Redes Neurais Artificiais

- São sistemas computacionais, de implementação em hardware ou software, que *imitam* as habilidades computacionais do sistema nervoso biológico, usando um grande número de neurônios artificiais simples e interconectados entre si.
- Algumas aplicações típicas ...
 - Reconhecimento de caracteres
 - Reconhecimento e Síntese de Voz
 - Visão artificial
 - Riscos de inadimplência / detecção de padrões de risco
 - Previsão de vendas / previsão de séries temporais

PSI-2533-2013– Prof. Emilio Del Moral

5

Dos Neurônios Biológicos aos Artificiais



**Neurônio Biológico
e
Neurônio Artificial**

PSI-2533-2013– Prof. Emilio Del Moral

6

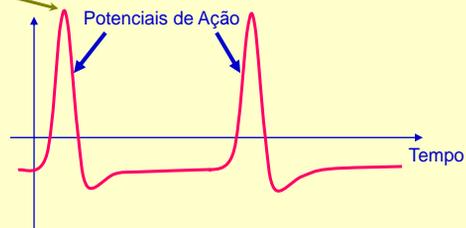
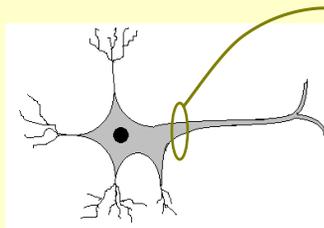
O Neurônio Biológico

- **Décadas de 50 e 60: neurônio biológico passou a ser visto como um dispositivo computacional elementar do sistema nervoso, com muitas entradas e uma saída.**

Entrada - Conexões Sinápticas

Informação - Impulsos Nervosos

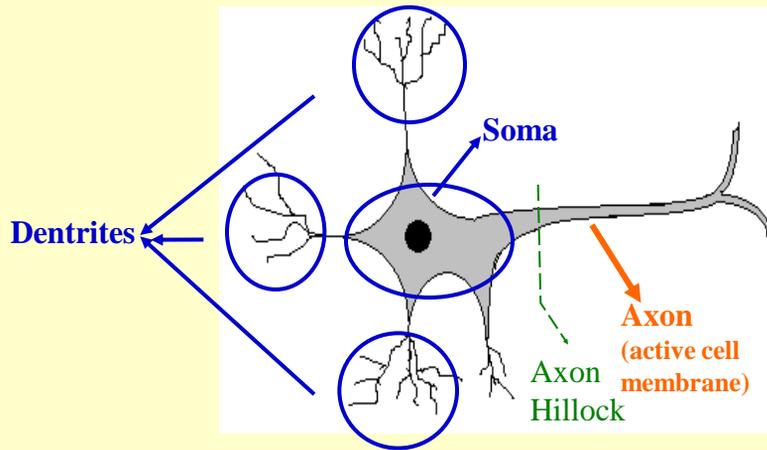
- **Impulso Nervoso: depolarização da membrana celular**



PSI-2533-2013– Prof. Emilio Del Moral

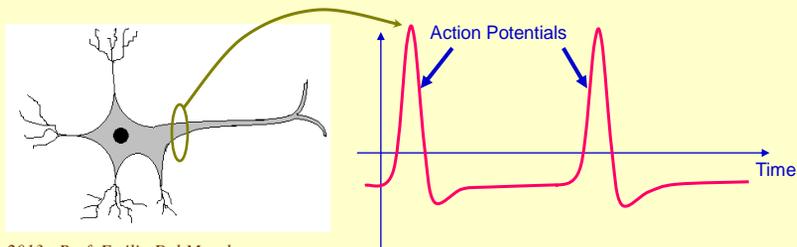
7

The biological neuron

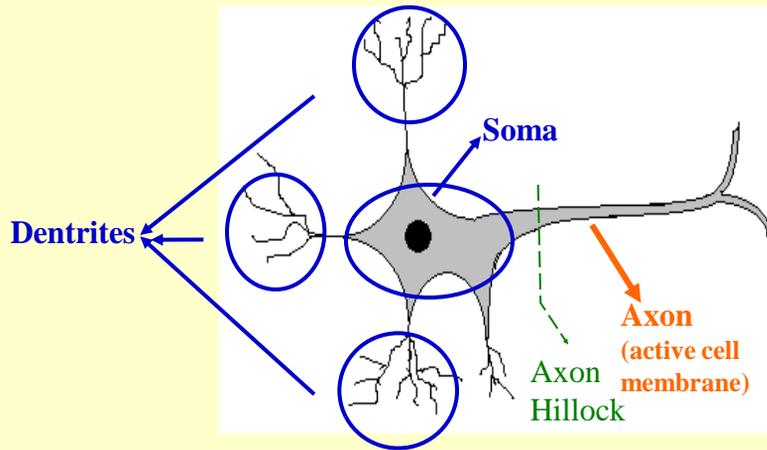


The biological neuron

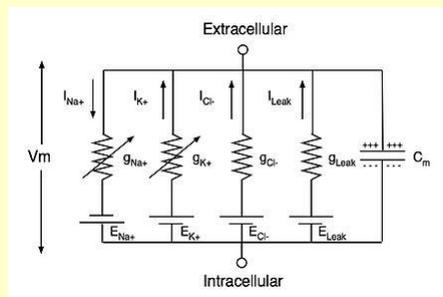
... Information is represented in electrical pulses generated by the nervous cell, named action potentials



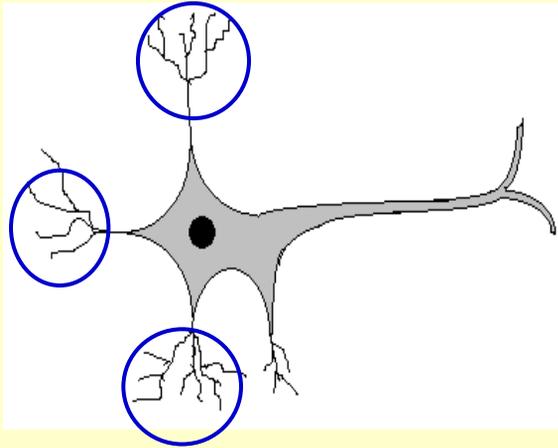
The biological neuron



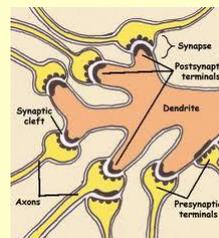
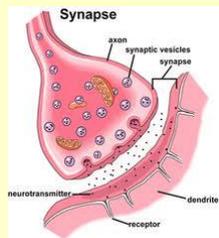
Hodgkin and Huxley model (explica tensão de membrana e os potenciais de ação)



Árvore Dendrítica

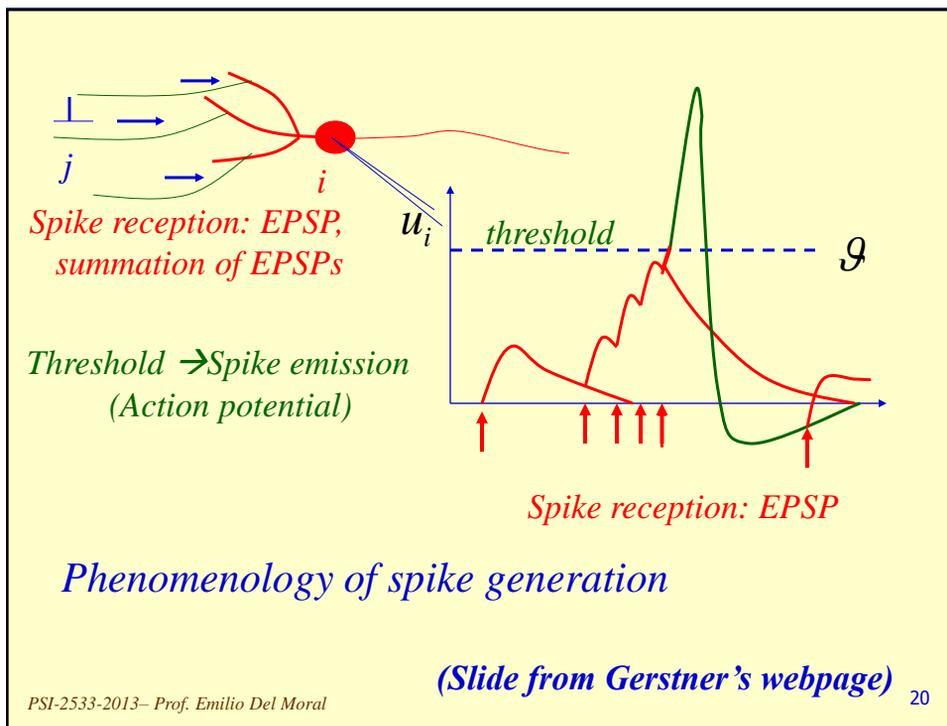


Sinapses



Understanding features and role of the Action Potentials

- Membrane potential at rest: approximately -50 mV
- Depolarization as a result of stimulation
- Once a threshold of depolarization is reached ... a pulse with characteristic dynamics is produced
- Amplitude and shape “does not code”
- Frequency of pulses “codes” (we have non repeatability of timing of spiking)



Understanding features and role of the Action Potentials

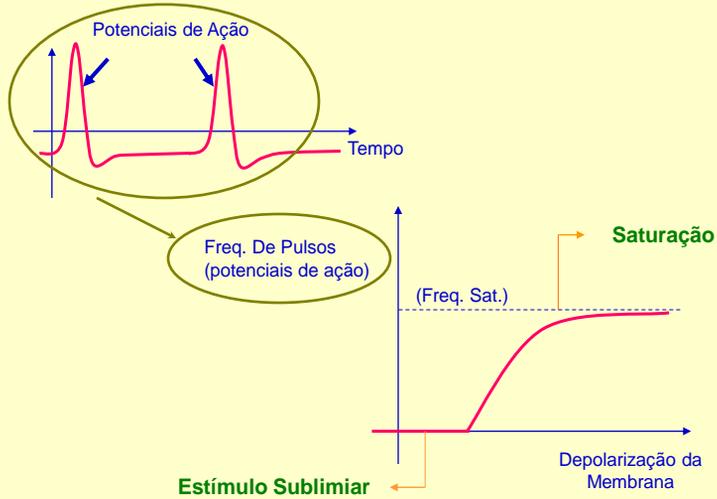
- Membrane potential at rest: approximately – 50 millivolts
- Depolarization as a result of stimulation
- Once a threshold of depolarization is reached ... a pulse with characteristic dynamics is produced
- **Amplitude and shape “does not code”**
- **Frequency of pulses “codes” (we have non repeatability of timing of spiking)**

Codificação de informação em neurônios biológicos

- **Frequencial**
- **Phase**
- **Sincronização**
- **??? Outros ???**

– **Os modelos neurais mais clássicos >>>
Predominantemente codificação FREQUENCIAL!!!!**

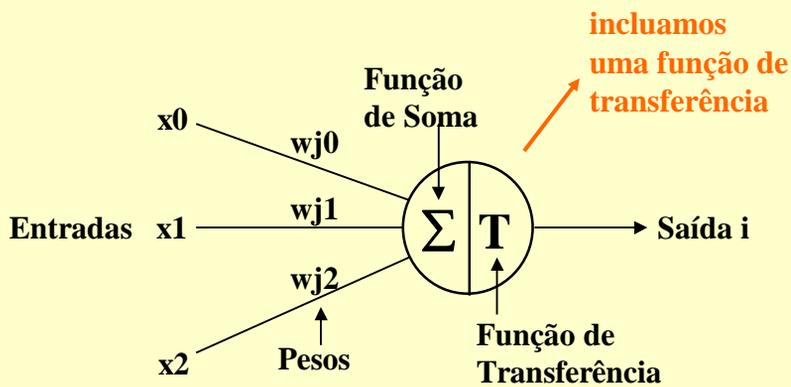
... Estudando a relação não linear entre volume de estímulo e volume de atividade de saída



PSI-2533-2013– Prof. Emilio Del Moral

23

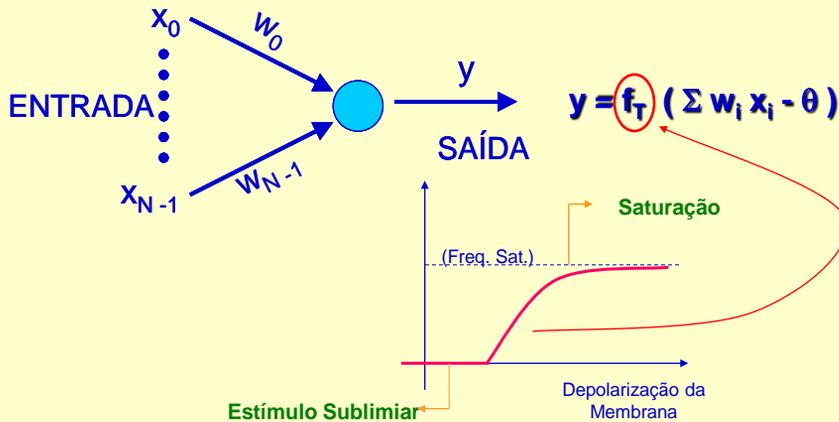
Modelando a Relação Entrada / Saída



PSI-2533-2013– Prof. Emilio Del Moral

24

Resumo ... um nó (neural) realiza a seguinte computação analógica



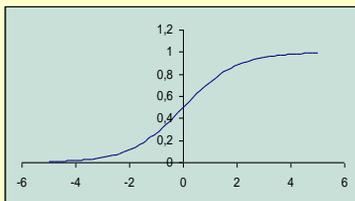
PSI-2533-2013– Prof. Emilio Del Moral

25

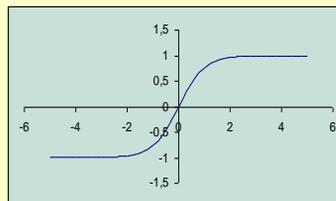
Função de transferência não linear do tipo sigmóide ou tangente hiperbólica

- Preserva - se o fenômeno da saturação
- Preserva - se a monotonicidade na faixa dinâmica
- Temos funções matematicamente amigáveis

SIGMÓIDE



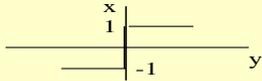
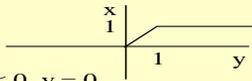
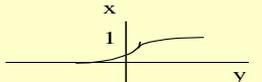
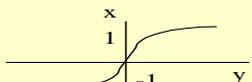
TANGENTE HIPERBÓLICA



PSI-2533-2013– Prof. Emilio Del Moral

26

Outras funções de transferência não linear

<p>Hard Limiter (limite rápido)</p>  <p> $s < 0, y = -1$ $s > 0, y = 1$ </p>	<p>Ramping Function (função de rampa)</p>  <p> $s < 0, y = 0$ $0 \leq s \leq 1, y = s$ $s > 1, y = 1$ </p>
<p>Sigmoide Function (função sigmóide)</p>  <p> $y = 1/(1 + e^{-s})$ </p>	<p>Sigmoide Function (função sigmóide)</p>  <p> $x \geq 0, y = 1 - 1/(1+s)$ $x < 0, y = -1 + 1/(1-s)$ </p>

- Com escalamento do argumento, pode-se abarcar os universos digital e analógico / linear e não linear simultaneamente

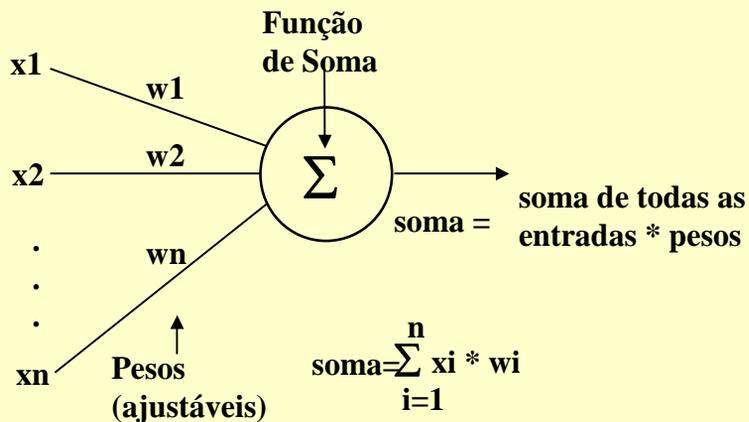
Temos pois duas componentes de cálculo complementares possíveis:

1) Uma linear (soma ponderada das entradas)

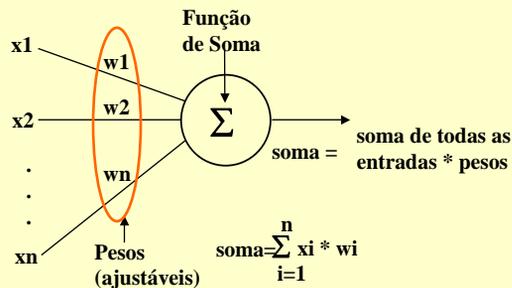
2) Outra não linear (Função de transferência da classe sigmoide)

A seguir: alguns comentários sobre a componente linear apenas

Adaline: Passos Iniciais ... um Modelo Linear



Adaline como Previsor Linear / Plasticidade sináptica



Filtros adaptativos / previsores lineares
 $x(n+1) = a.x(n) + b.x(n-1) + c.x(n-2) + \dots$

a, b, c ... São adaptados para otimizar a previsão
Face a um conjunto de treinamento

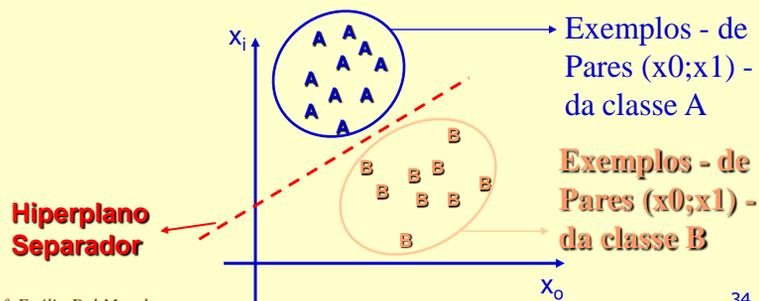
Does Adaline model well the processes in biological neurons?

- It does not represent well the mapping from stimulus to output in situations of very LOW LEVELS OF STIMULUS
- It does not represent well the mapping from stimulus to output in situations of very HIGH LEVELS OF STIMULUS

A seguir: alguns comentários sobre a componente NÃO linear apenas

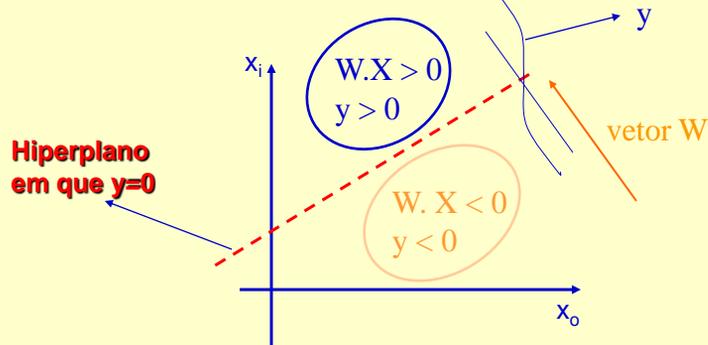
O Perceptron: $y = \text{signal}(\sum w_i x_i - \theta)$ (função de transferência tipo “degrau”)

- Viabiliza a classificação de padrões com separabilidade linear
- O algoritmo de aprendizado adapta os W_s de forma a encontrar o hiperplano de separação adequado
- Aprendizado por conjunto de treinamento

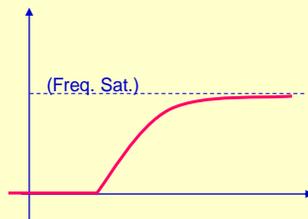


E se a saída do nosso problema não for digital? O “Perceptron Contínuo”: $y = \text{tgh}(\sum w_i x_i - \theta)$

- Que problemas de entradas contínuas conseguimos atacar usando uma função de transferência tangente hiperbólica)

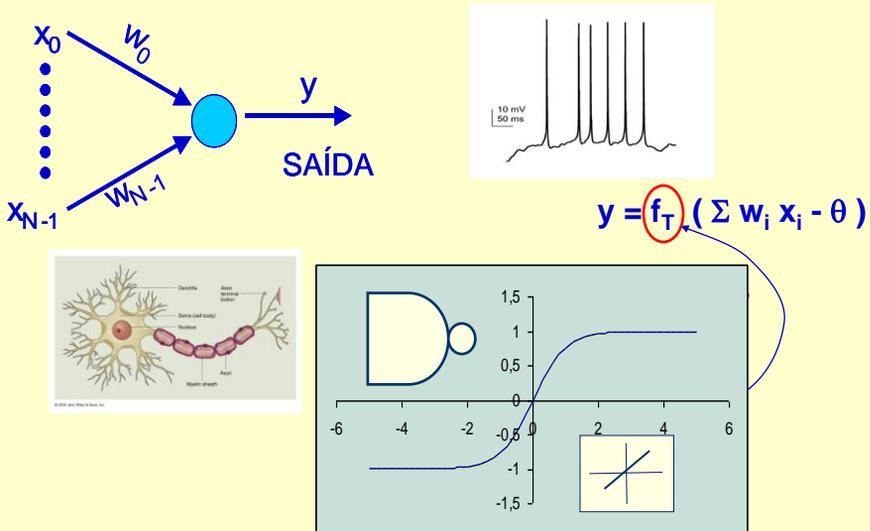


The same model can be useful for different tasks



- Linear input-output mapping
- Digital output applications
- Non-linear input-output mapping

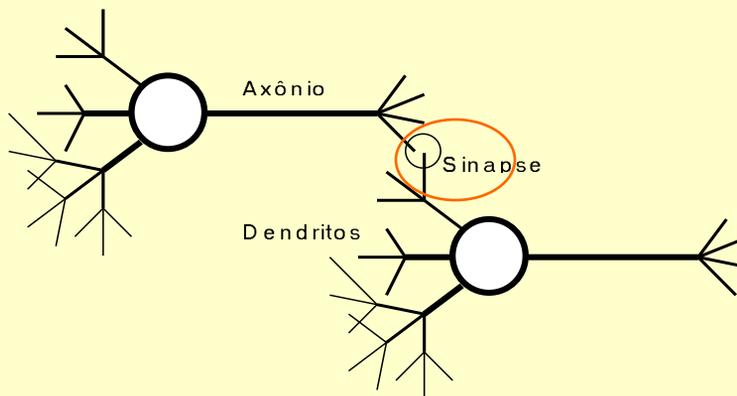
Computação e adaptabilidade em um nó neural



PSI-2533-2013– Prof. Emilio Del Moral

37

Cômputos mais complexos ... são realizados pelo encadeamento de vários nós



A conexão entre um axônio de um neurônio e um dendrito de outro é denominada **Sinapse**

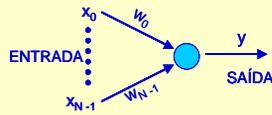
PSI-2533-2013– Prof. Emilio Del Moral

39

Aqui começamos a falar de REDES neurais

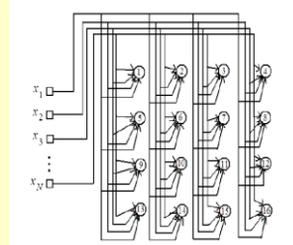
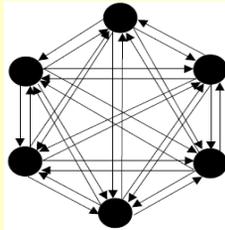
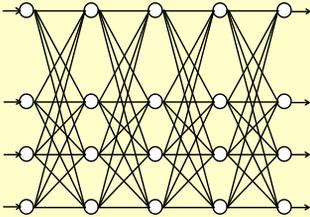
- **Mais de um neurônio artificial envolvido**
- **Formas específicas de organização e interconexão**
- **Formas específicas de representação de informação de acordo com as grandezas do problema**
- **Formas específicas de programação dos pesos sinápticos**
- **Formas específicas de “entrar” informação na rede e de “retirar” o resultado fornecido por ela**

***Três RNAs Clássicas,
com grande número de
ferramentas,
metodologias e nichos
importantes de
aplicações***



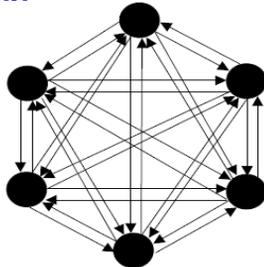
3 arquiteturas clássicas:

MLP / Hopfield / SOMs de Kohonem



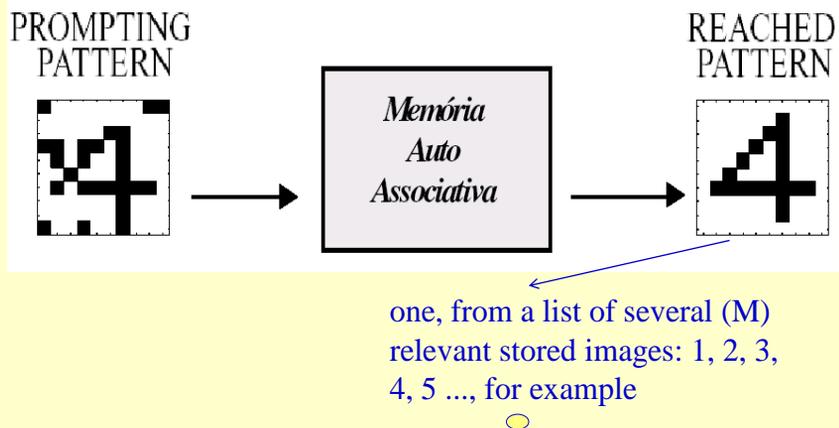
Associative (Neuro) Memory proposed by John Hopfield

- In 1982, John Hopfield presented an artificial neural system which was able to memorize images in the interconnections between the model neurons



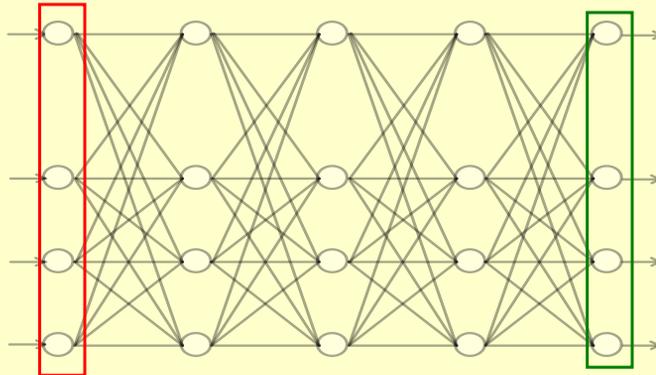
- Concept of attractors and energy landscape

Example of auto-association in the recovery of corrupted images ... each pixel = one node



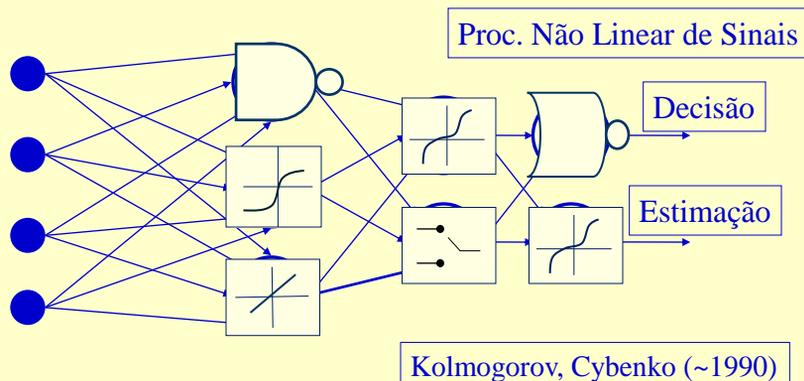
O Multi Layer Perceptron (MLP)

- Múltiplas **entradas** / Múltiplas **saídas**
- Ambas podem ser analógicas ou digitais
- Não há mais a restrição de separabilidade linear



O Multi Layer Perceptron (MLP)

- Múltiplas entradas / Múltiplas saídas / Múltiplas camadas
- Variáveis (internas e externas) analógicas ou digitais
- Relações lineares ou não lineares entre elas



PSI-2533-2013– Prof. Emilio Del Moral

46

Resumindo os aspectos conceituais principais

- 1) Não linearidade com a função neural sigmoidal
- 2) Possibilidade de conjugar na mesma estrutura ...
 - Cálculos digitais
 - Cálculos lineares multivariáveis
 - Funções genéricas não lineares multivariáveis
- 3) Comportamento adaptativo com aprendizado através de exemplos
 - Problemas complexos, multidimensionais, não lineares e mesmo aqueles sem teoria conhecida
 - Decisão automática, estimação, reconhecimento de padrões, classificação, processamento não linear de sinais, *clustering* multidimensional ...

PSI-2533-2013– Prof. Emilio Del Moral

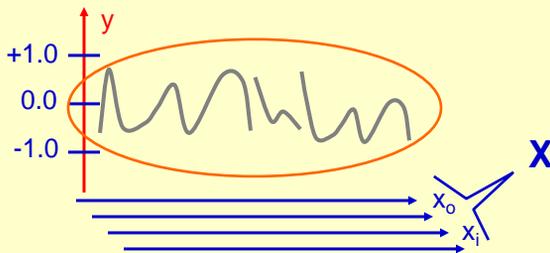
47

Aprendizado em RNAs do tipo MLP – Multi Layer Percetron – através do algoritmo Error Back Propagation

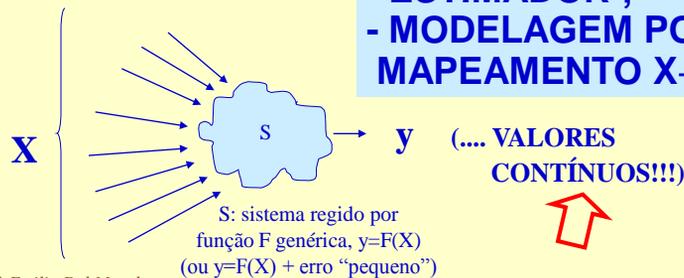
Contexto: Redes neurais do tipo MLPs aplicadas para a aproximação / representação de funções de múltiplas variáveis

$$y=F(X)$$

A função $y(X)$ “a descobrir”, num caso geral de função contínua $y(X)$



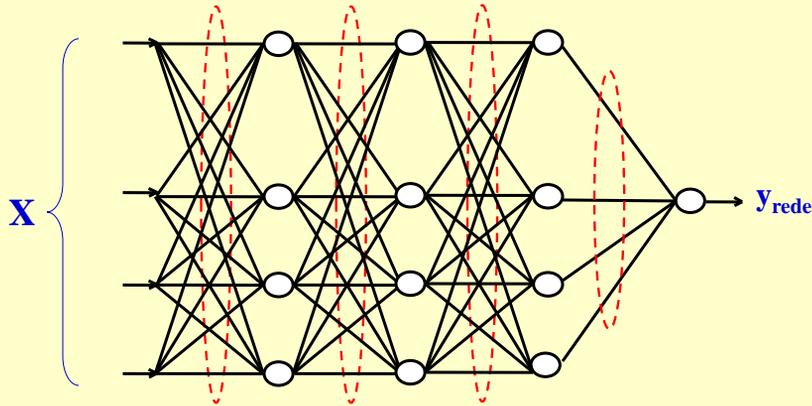
- ESTIMADOR ;
- MODELAGEM POR
MAPEAMENTO $X \rightarrow y$



Um exemplo de modelagem por funções $y(X)$ de múltiplas variáveis: previsão de Ibovespa assumindo que o próximo Ibovespa é função dos anteriores

(nota: se preferir, expanda o modelo com mais entradas X – valores do dólar por ex. – ou com elementos de dinâmica externos à função $y(X)$)

Como sabemos no E.B.P. se a rede está sendo boa ou ruim (como modelo de um conjt. de treino), para cada configuração específica de pesos w_{ij} sendo explorada?



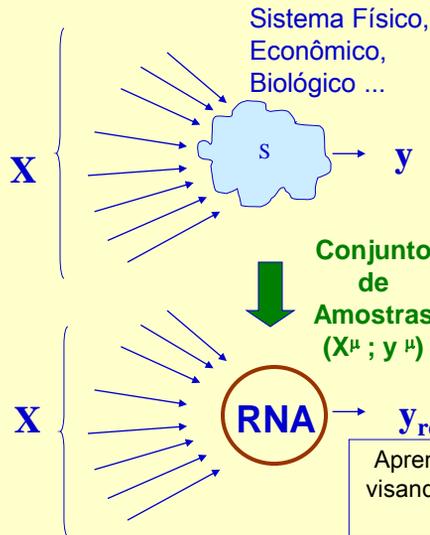
(na lousa) ... erro da rede com relação ao conjunto de treinamento como um todo; simbologia (\bar{X}^μ ; y^μ); Erro quadrático de exemplar (Eq^μ); Erro quadrático médio (Eqm)

Resumo de principais resultados em lousa ...

μ identifica um de M exemplos de treinamento

$$Eqm = \frac{1}{M} \sum_{\mu=1}^M (y_{rede}(\bar{X}^\mu) - y^\mu)^2$$

Conjunto de treino em arquiteturas supervisionadas (ex. clássico: MLP com Error Back Propagation)



A computação desejada da rede pode ser definida simplesmente através de amostras / exemplos do comportamento requerido

$$Eqm = \frac{1}{M} \sum_{\mu=1}^M (y_{rede}(\vec{X}^\mu, W) - y^\mu)^2$$

$$\vec{\Delta W} = -\eta \cdot \vec{\nabla} Eqm$$

Aprendizado: Espaço de pesos W é explorado visando aproximar ao máximo a computação da rede da computação desejada