

REDES NEURAIS

Marley Maria B.R. Vellasco

ICA: Núcleo de Pesquisa em Inteligência
Computacional Aplicada
PUC-Rio



CONTEÚDO

- **Introdução**
 - Motivação, Objetivo, Definição, Características Básicas e Histórico
- **Conceitos Básicos**
 - Neurônio Artificial, Modos de Interconexão
- **Processamento Neural**
 - *Recall e Learning*
- **Regras de Aprendizado**
 - Regra de Hebb, Perceptron, Back Propagation, Hopfield e Competitive Learning



MOTIVAÇÃO

Constatação que o cérebro processa informações de forma diferente dos computadores convencionais

CÉREBRO

velocidade 1 milhão de vezes mais lenta que qualquer gate digital



COMPUTADOR

processamento extremamente rápido e preciso na execução de sequência de instruções

**Processamento altamente paralelo
(10^{11} neurônios com 10^4 conexões cada)**



Problema dos 100 Passos

Neurônio: 2ms

Processador: 2ns



Processador é 10^6 mais rápido que o neurônio



Cérebro reage a um estímulo entre 0,2 e 1 seg.



O cérebro responde em **100 passos**



MOTIVAÇÃO

- Observações:

- O cérebro tem ~ **10 bilhões** de neurônios.
 - Cada neurônio tem ~ **1.000** a **10.000** conexões
- ↓
- **60 trilhões** de conexões - **10^{14} sinapses!**
- ↓
- Cada pessoa pode dedicar **100.00 conexões** para armazenar cada segundo de experiência
(**65 anos** ⇒ **2.000.000.000 de segundos!**)
 - Durante os 2 primeiros anos de vida, **1.000.000 de sinapses** são formadas por segundo!!



OBJETIVO

Estudar a teoria e a implementação de sistemas **massivamente paralelos**, que possam processar informação com **eficiência comparável** ao cérebro.



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.



Aquisição de Conhecimento: Aprendizado

Treinamento efetuado através da apresentação de exemplos



Existe uma variedade de algoritmos que estabelecem **QUANDO** e **COMO** os parâmetros da Rede Neural devem ser atualizados



Substitui a programação necessária para a execução das tarefas nos computadores



APLICAÇÕES GERAIS

- ☑ Reconhecimento de Padrões
- ☑ Classificação de Padrões
- ☑ Correção de Padrões
- ☑ Previsão de Séries Temporais
- ☑ Aproximação de Funções
- ☑ Suporte à Decisão
- ☑ Extração de Informações



Características Básicas

Devido à similaridade com a estrutura do cérebro, as Redes Neurais exibem características similares ao do comportamento humano, tais como:



Características Básicas

- Procura Paralela e Endereçamento pelo Conteúdo:

O cérebro **não** possui **endereço de memória** e **não procura** a informação **sequencialmente**



Características Básicas

- Aprendizado:

A rede **aprende por experiência**, não necessitando explicitar os algoritmos para executar uma determinada tarefa



Características Básicas

- Associação:

A rede é capaz de fazer associações entre padrões diferentes

Ex: Pessoa → Nome
Perfume → Pessoa



Características Básicas

- 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.



Características Básicas

- Abstração:

Capacidade de **abstrair a essência de um conjunto de entradas**, isto é, a partir de padrões ruidosos, extrair a informação do padrão sem ruído.



Características Básicas

- Robustez e Degradação Gradual:

A perda de um conjunto de elementos processadores não causa o mal funcionamento da rede neural.



EVOLUÇÃO

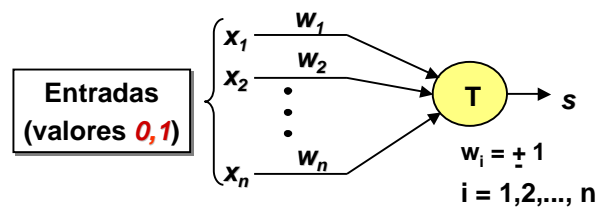
McCulloch & Pitts (Mathematical Bio-Physics, Vol. 5, 1943),
 "A Logical Calculus of Ideas Immanent in Nervous Activity"

	Von Neumann <i>Machine Intelligence</i>	Marvin Minsky <i>Macroscopic Intelligence</i>	Frank Rosenblatt <i>Microscopic Intelligence</i>
1945	Digital Computers		
1950		Black-Box AI (LISP)	Perceptron, Adaline
1960	Mainframes	Theorem Solver	
1970	Vax 780 (Time Sharing)	Expert Systems	
1980	Workstations, PCs	Commercialization of E.S.	Rumelhart, Hopfield
1990	Desktop Supercomputers		Commercialization of N.N.



EVOLUÇÃO

- **Modelo de McCulloch-Pitts:**

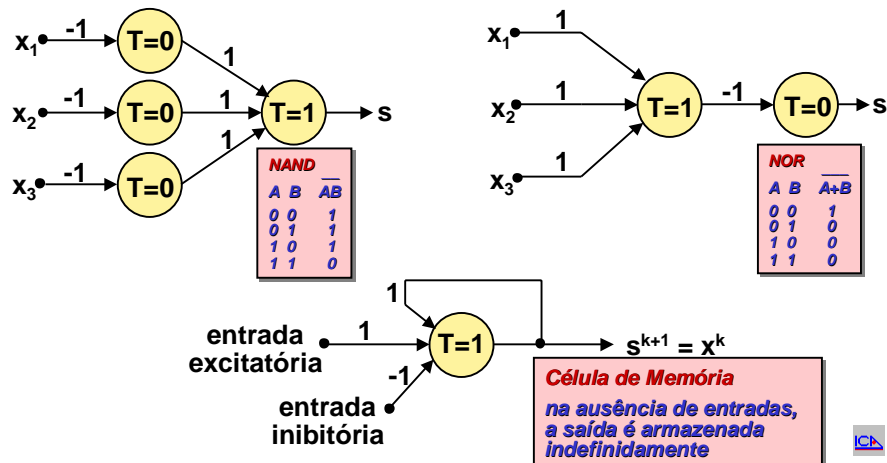


$$s^{k+1} = \begin{cases} 1 & \text{se } \sum_{i=1}^n w_i x_i^k \geq T \\ 0 & \text{se } \sum_{i=1}^n w_i x_i^k < T \end{cases}$$



EVOLUÇÃO

• Modelo de McCulloch-Pitts:



HISTÓRICO

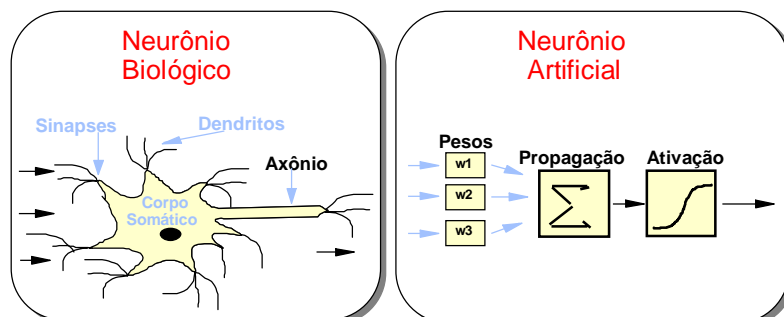
- McCulloch & Pitts (1943):
 - modelo computacional para o neurônio artificial. Não possuía capacidade de aprendizado
- Hebb (1949):
 - modelo de aprendizado (*Hebbian Learning Rule*)
- Rosenblatt (1957):
 - Perceptron, com grande sucesso em certas aplicações e problemas em outras aparentemente similares
- Minsky & Papert (*Perceptrons* 1969):
 - prova matemática de que as redes Perceptron são incapazes de solucionar problemas simples tipo OU-EXCLUSIVO
- Rumelhart (início da década de 80):
 - novos modelos que superaram os problemas dos Perceptrons.

CONCEITOS BÁSICOS

- **Neurônio Artificial**
 - (Elemento Processador)
- **Estruturas de Interconexão**
 - FeedForward de 1 camada
 - FeedForward de Múltiplas Camadas
 - Recorrente (com realimentação)



Elemento Processador



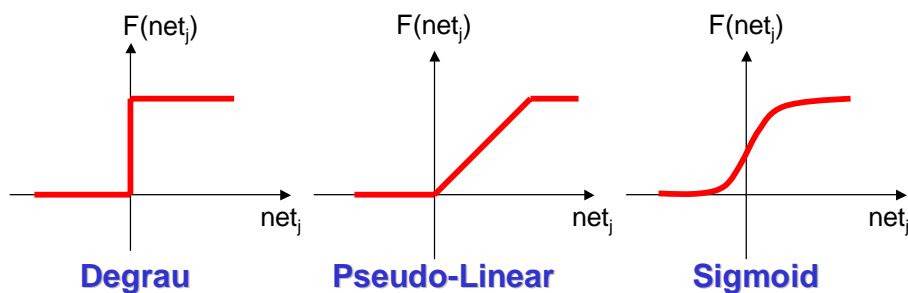
Elementos Básicos

- **Estado de Ativação** $\rightarrow s_j$
- **Conexões entre Processadores**
 - a cada conexão existe um **peso sináptico** que determina o efeito da entrada sobre o processador $\rightarrow w_{ij}$
- **Função de Ativação**
 - determina o novo valor do **Estado de Ativação** do processador $\rightarrow s_j = F(\text{net}_j)$



Funções de Ativação

É a função que determina o nível de ativação do Neurônio Artificial - $s_j = F(\text{net}_j)$



Tipos de Processadores

Input → Recebe os dados de entrada

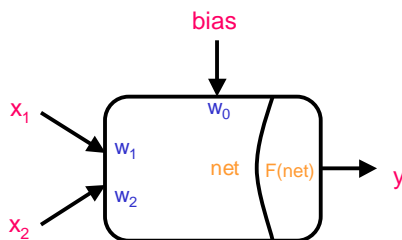
Output → Apresenta os dados de saída

Hidden → As suas entradas e saídas permanecem dentro do sistema

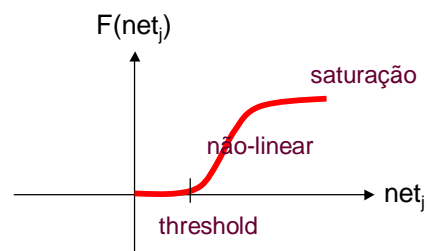


Neurônio Artificial

3 pontos importantes: ↻ **Thresholding**
↻ **Não-Linearidade**
↻ **Saturação**



$$\text{net} = w_0 + w_1x_1 + w_2x_2$$
$$F(\text{net}) = \frac{1}{1 + e^{-\text{net}}} \text{ (sigmoid)}$$



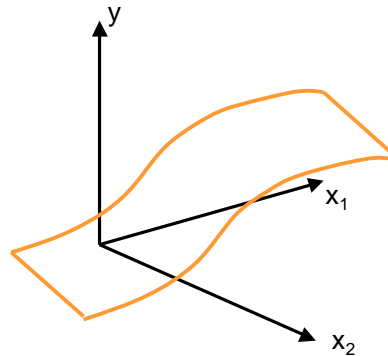
Neurônio Artificial

Em função das equações de **net** e **F(net)**:

$$y = F(\text{net}) = \frac{1}{1 + e^{-(w_0 + x_1 w_1 + x_2 w_2)}}$$



Fórmula matemática representada pelo neurônio artificial



Exemplos

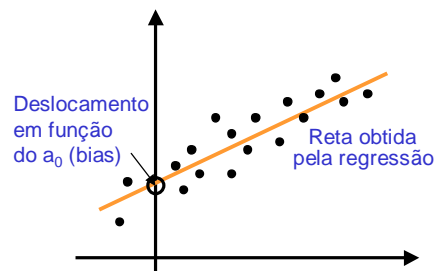
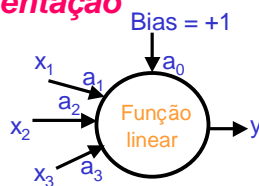
Regressão Linear:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$$

Variáveis explicativas

Acha a **reta** com **erro mínimo** que passe pelos **pontos existentes** (padrões de treinamento)

Representação Neural

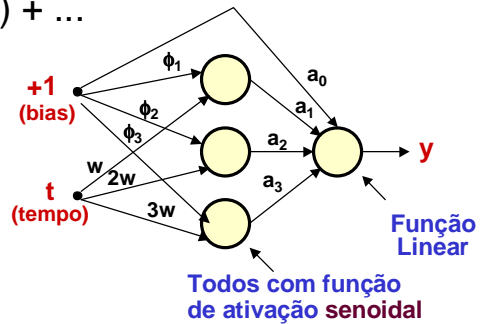


Exemplos

Transformada de Fourier:

$$y = a_0 + a_1 \text{sen}(wt + \phi_1) + a_2 \text{sen}(2wt + \phi_2) + a_3 \text{sen}(3wt + \phi_3) + \dots$$

**Representação
Neural**



CONCEITOS BÁSICOS

- **Neurônio Artificial**
 - (Elemento Processador)
- **Estruturas de Interconexão**
 - FeedForward de 1 camada
 - FeedForward de Múltiplas Camadas
 - Recorrente (com realimentação)

Topologias das Redes Neurais

- **Redes Feed-Forward:**

- redes de uma ou mais camadas de processadores, cujo **fluxo de dados** é sempre em **uma única direção**, isto é, não existe realimentação.

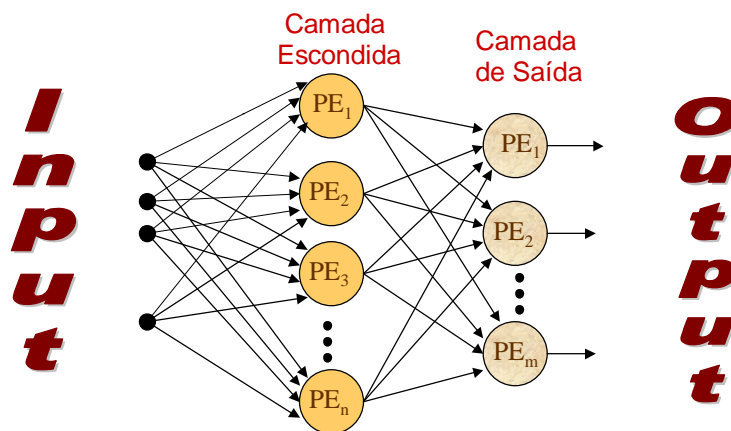
- **Redes Recorrentes:**

- redes com conexões entre processadores da mesma camada e/ou com processadores das camadas anteriores (**realimentação**).

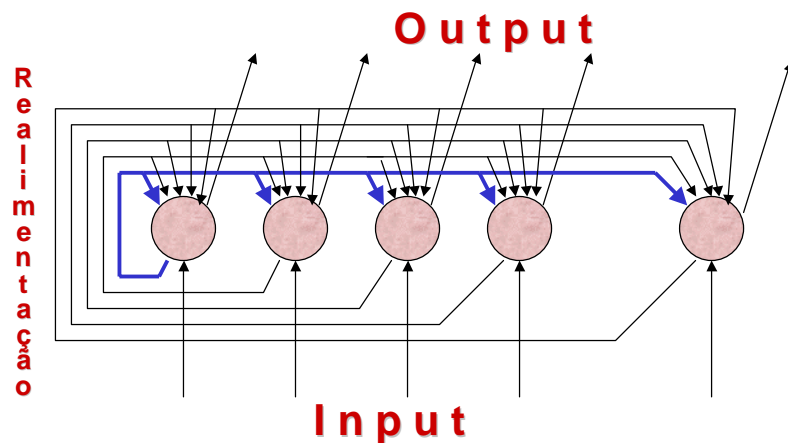


Redes Feed-Forward

Rede de Múltiplas Camadas



Redes Recorrentes



Processamento Neural

O processamento de uma Rede Neural pode ser dividido em duas fases:

Recall



Processo de **cálculo da saída da rede**, dado um certo padrão de entrada -
Recuperação da Informação

Learning



Processo de **atualização dos pesos sinápticos** para a aquisição do conhecimento -
Aquisição da Informação

Recuperação de Dados



Recuperação de Dados

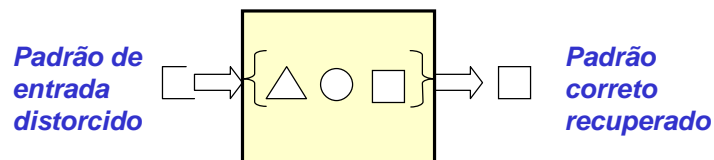
Assumindo que um conjunto de padrões tenha sido **armazenado**, a Rede Neural pode executar as seguintes tarefas:



Recuperação de Dados

Autoassociação:

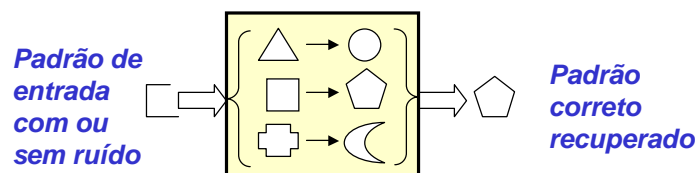
_A Rede Neural recupera o padrão armazenado mais semelhante ao padrão de entrada apresentado. ➡ **Recuperação de padrões ruidosos**



Recuperação de Dados

Heteroassociação:

_A Rede Neural armazena a associação entre um par de padrões. ➡ **Recuperação de um padrão diferente do da entrada.**

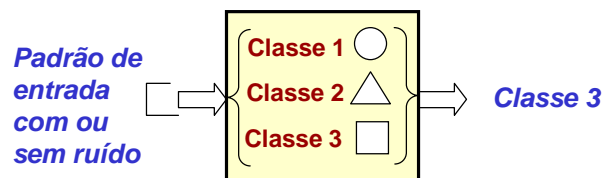


Recuperação de Dados

Classificação:

— A Rede Neural responde com a informação relativa à **classe** a qual o padrão de entrada pertence. ➡ **Caso especial de Heteroassociação**

Ex: Padrões de entrada divididos em 3 classes distintas.



GENERALIZAÇÃO

— A Rede Neural responde corretamente a um padrão de entrada fora do conjunto de treinamento. ➡ **Interpola corretamente os novos pontos apresentados**



Aprendizado



Aprendizado

- Processo pelo qual os **parâmetros livres** - *pesos sinápticos* - de uma rede neural são adaptados através de um **processo contínuo** de estimulação pelo ambiente.
- Existem 3 tipos básicos de aprendizado:
 - ☑ Treinamento Supervisionado;
 - ☑ Treinamento Não-Supervisionado;
 - ☑ Treinamento em “*Batch*”.



Treinamento Supervisionado

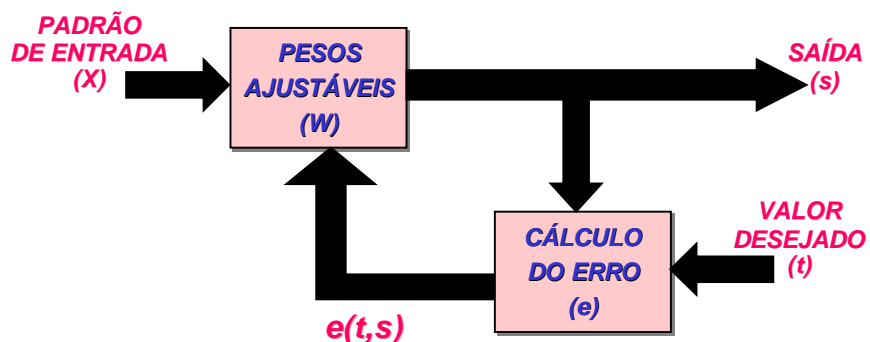
A rede é treinada através do fornecimento dos valores de **entrada** e seus respectivos valores da **saída desejada** (“*training pair*”).



Geralmente efetuado através do processo de **minimização do erro** calculado na saída.



Treinamento Supervisionado



Treinamento Não-Supervisionado

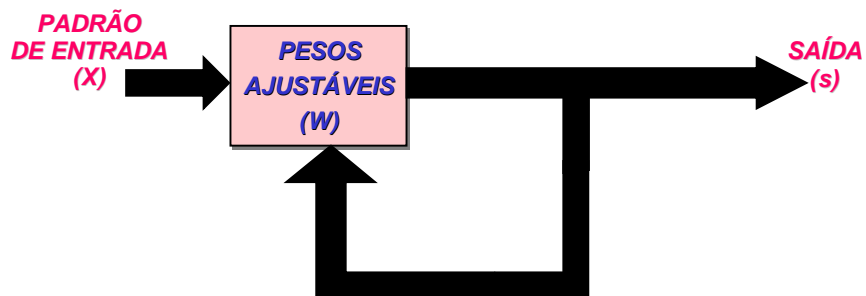
“*Self-Organization*” → Não requer o valor desejado de saída da rede. O sistema extrai as características do conjunto de padrões, agrupando-os em classes inerentes aos dados.



Aplicado a problemas de *Clusterização*



Treinamento Não-Supervisionado



Treinamento em “Batch”

Os valores dos pesos sinápticos são estabelecidos a priori, em um único passo. → Também chamado de **Gravação** - “*Recording*”



Exemplos de Treinamento Supervisionado

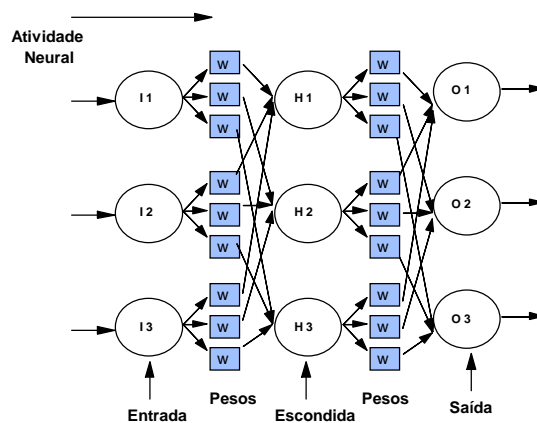
- ① Reconhecimento de Caracteres
- ② Previsão de Séries Temporais



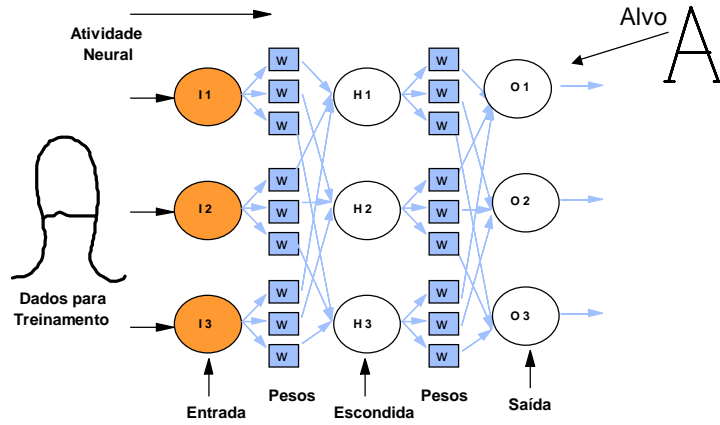
Reconhecimento de Caracteres



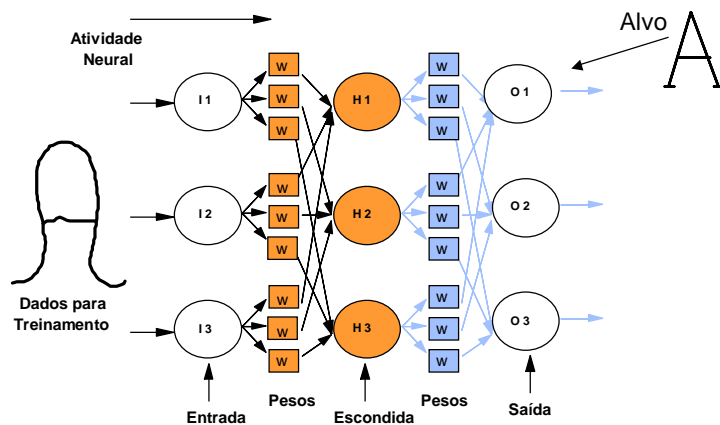
Estrutura da Rede Neural



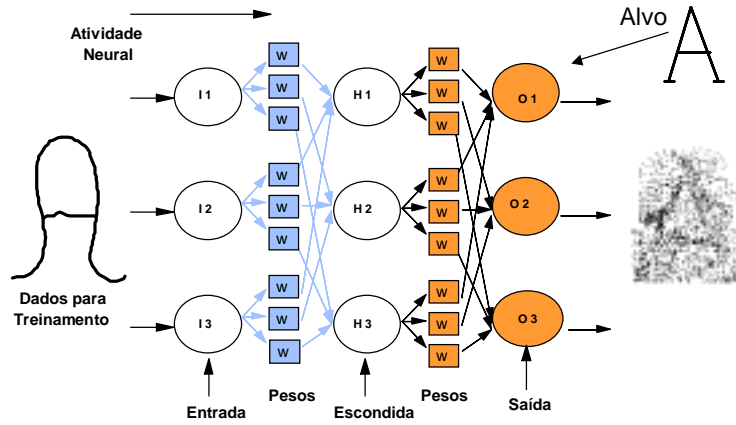
Processo de Aprendizizado



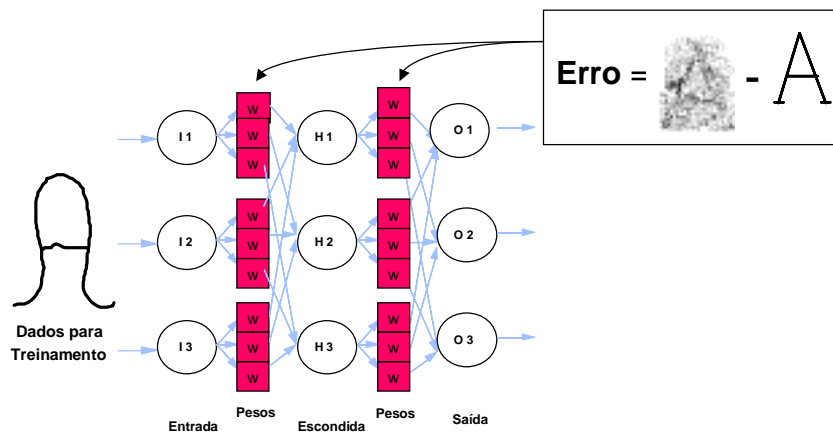
Processo de Aprendizizado



Processo de Aprendizado



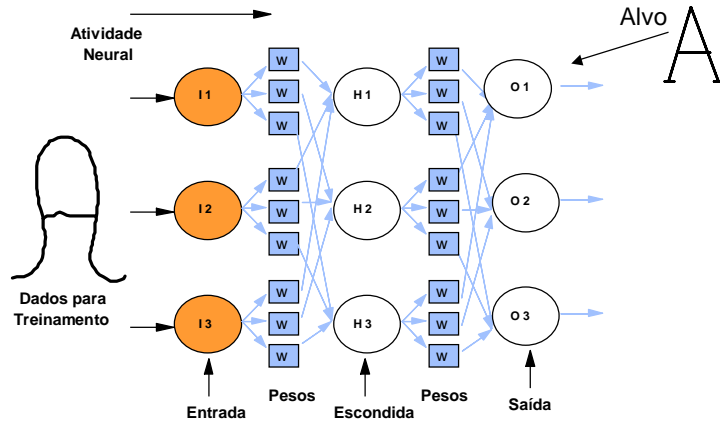
Processo de Aprendizado



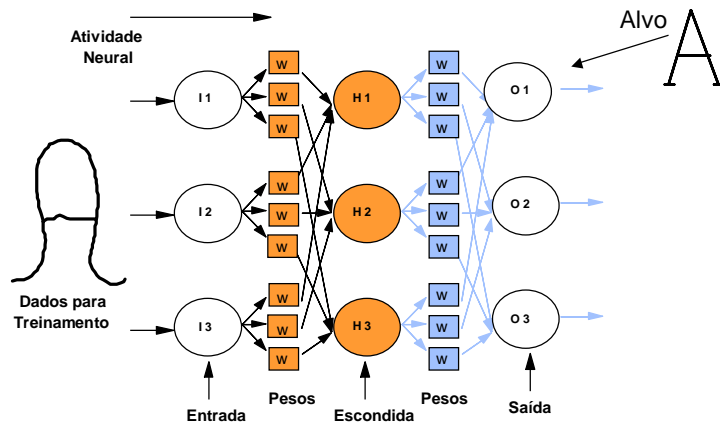
Atualização dos pesos em função do erro



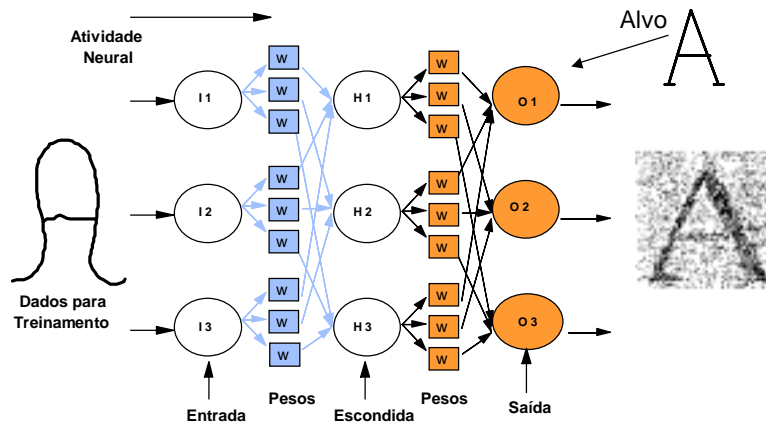
Processo de Aprendizado



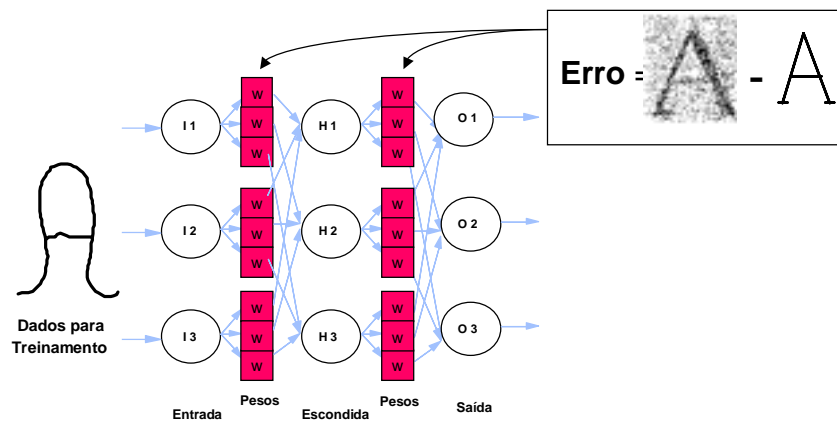
Processo de Aprendizado



Processo de Aprendizado



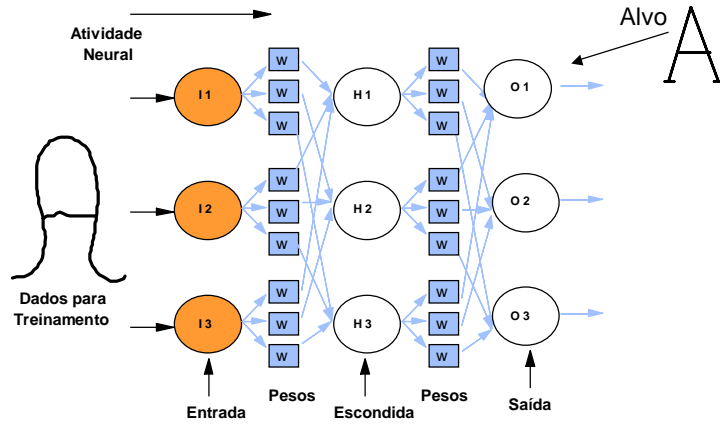
Processo de Aprendizado



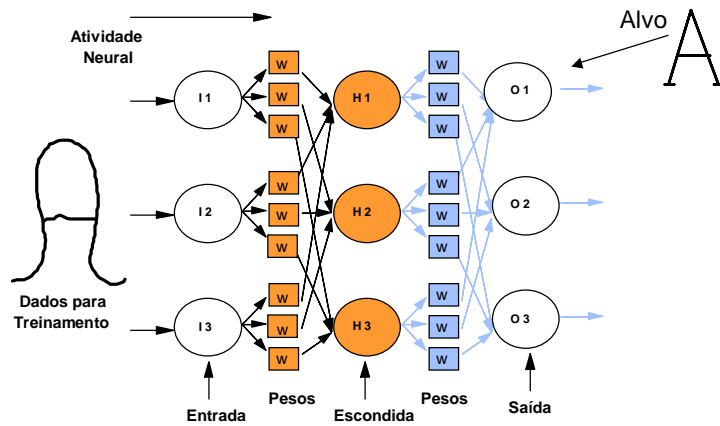
Atualização dos pesos em função do erro



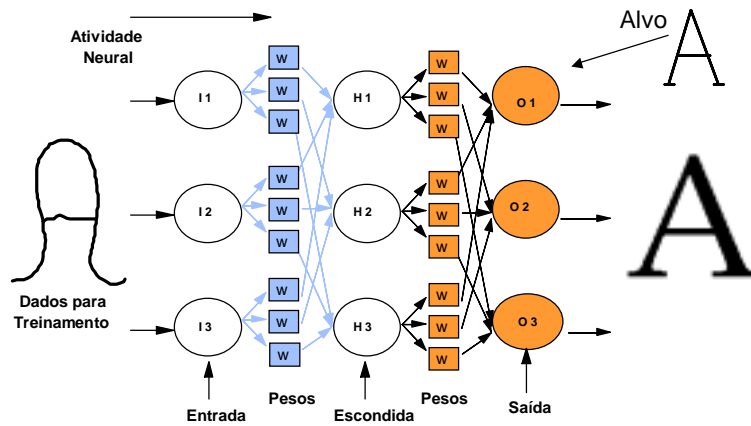
Processo de Aprendizado



Processo de Aprendizado



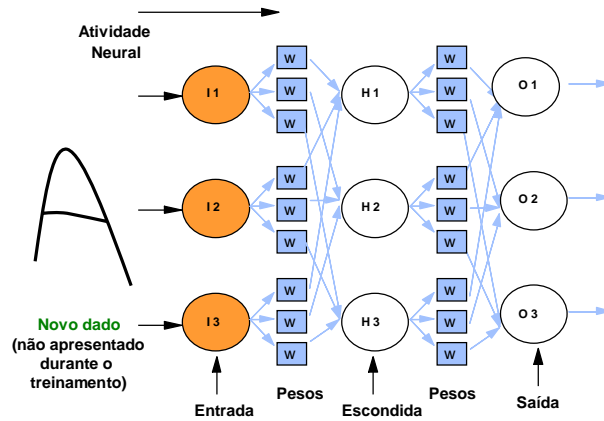
Processo de Aprendizado



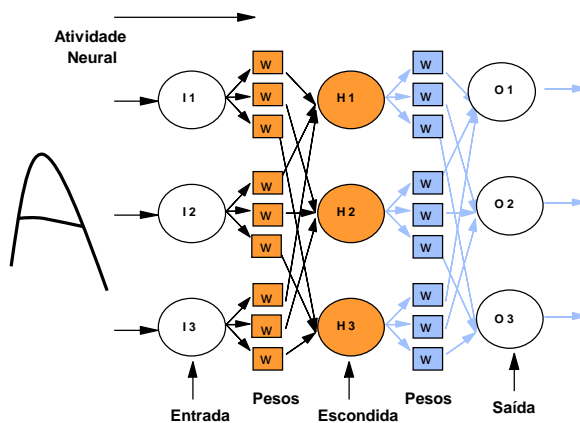
Processo de Generalização

Recuperação da Informação
Aprendida

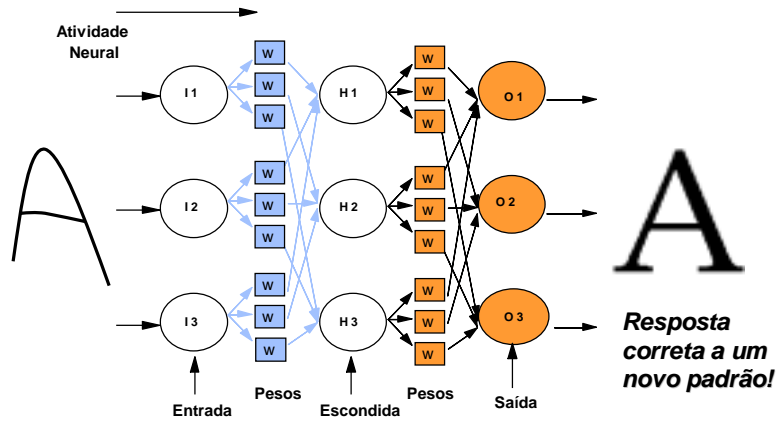
Processo de Generalização



Processo de Generalização



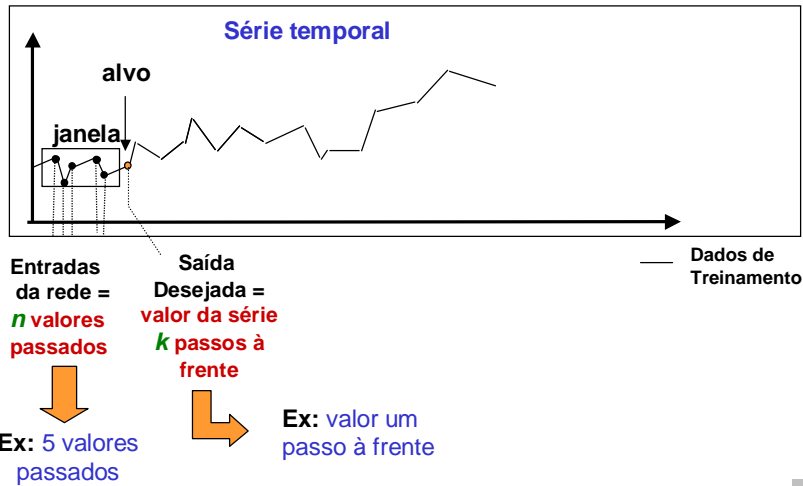
Processo de Generalização



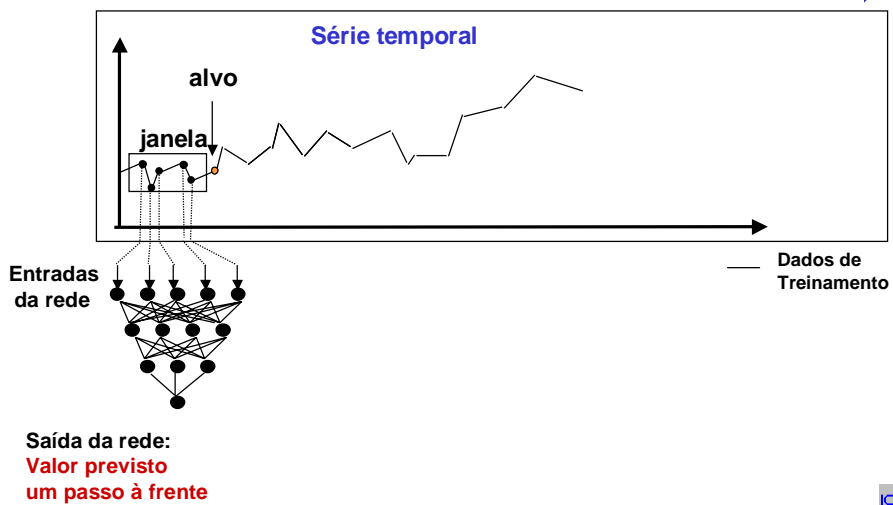
Previsão de Séries Temporais



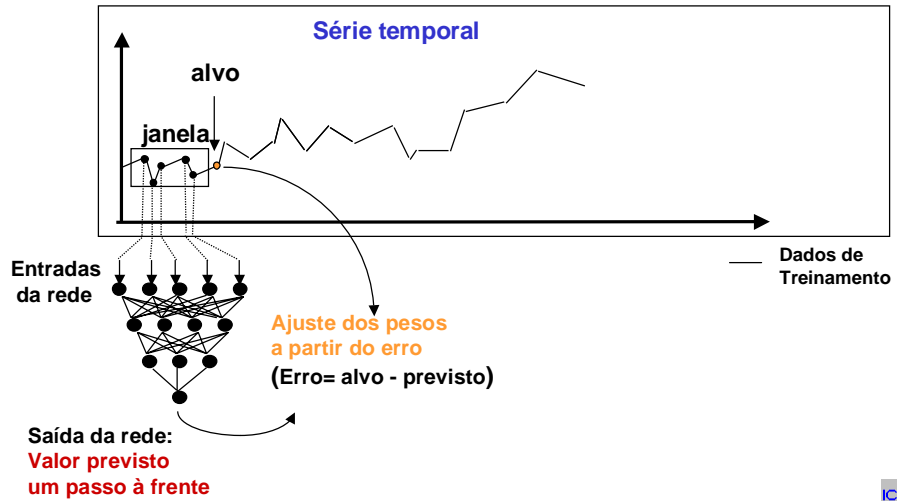
Estrutura da Rede Neural



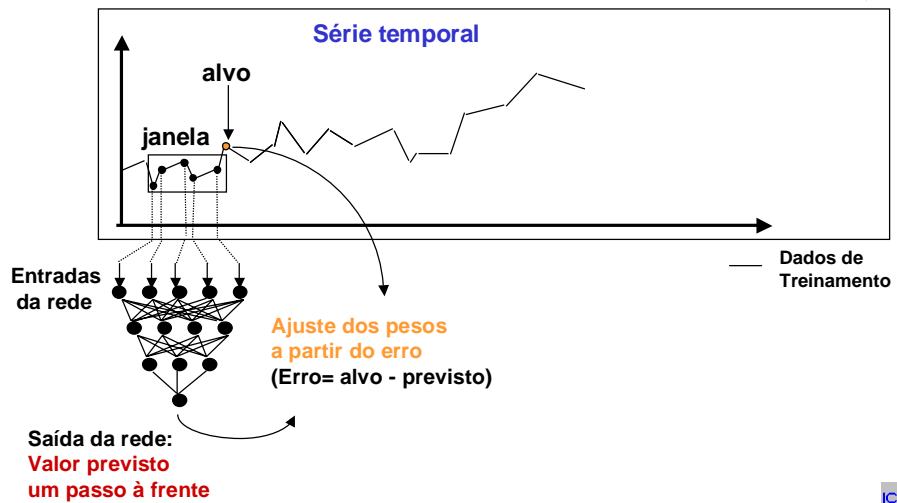
Processo de Aprendizado



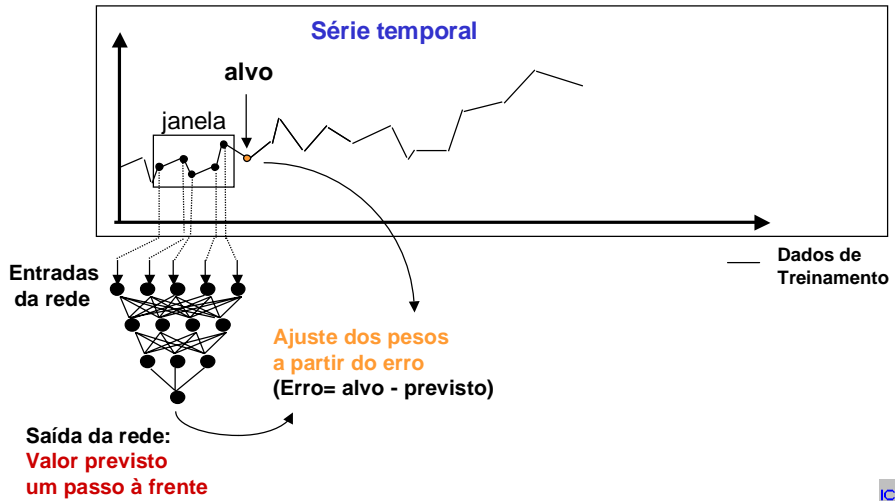
Processo de Aprendizado



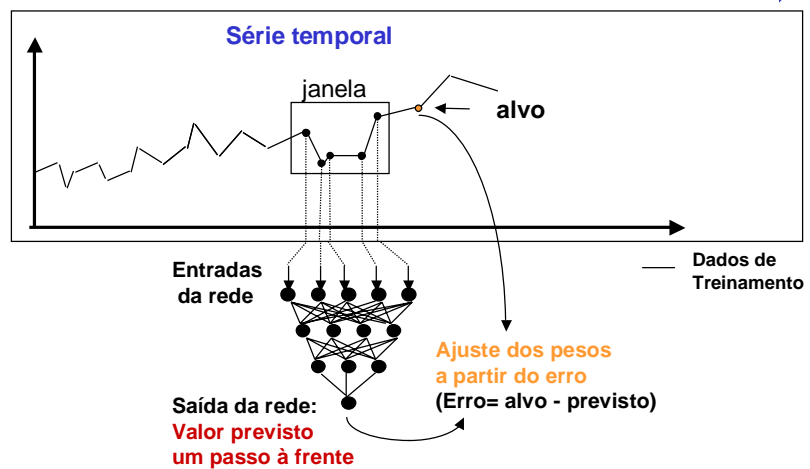
Processo de Aprendizado



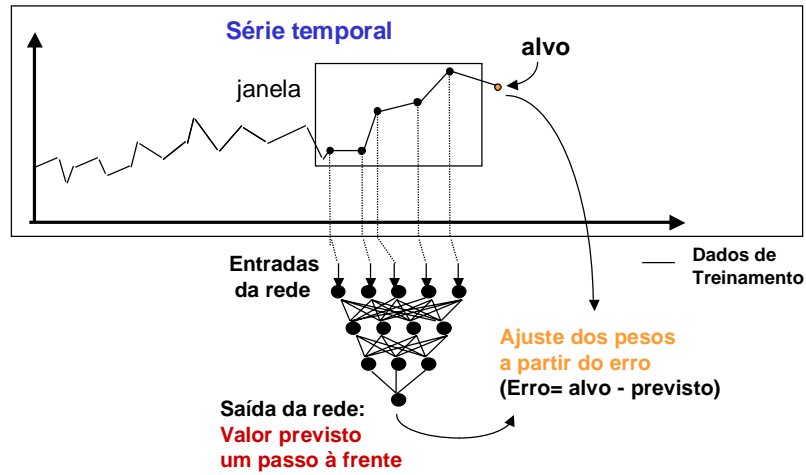
Processo de Aprendizado



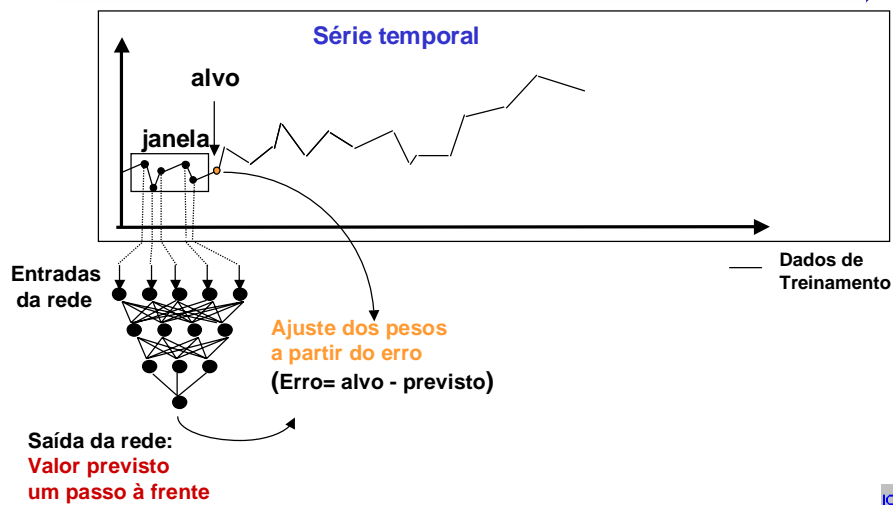
Processo de Aprendizado



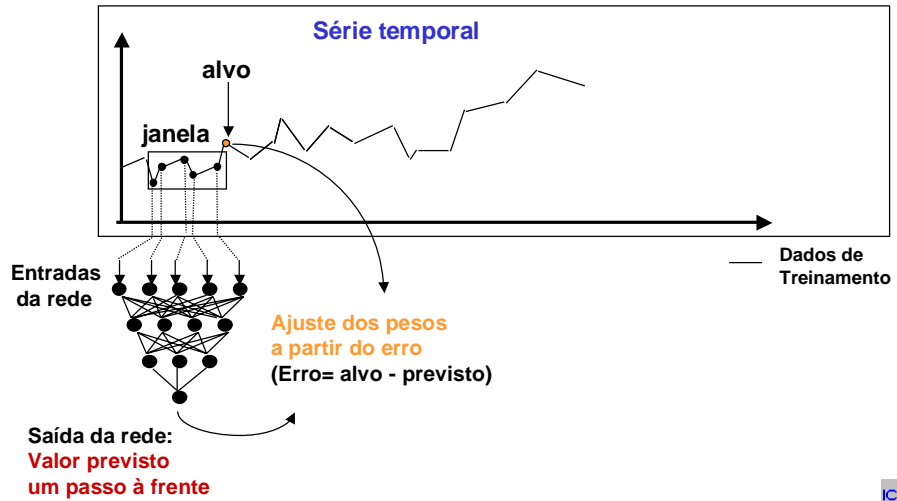
Processo de Aprendizado



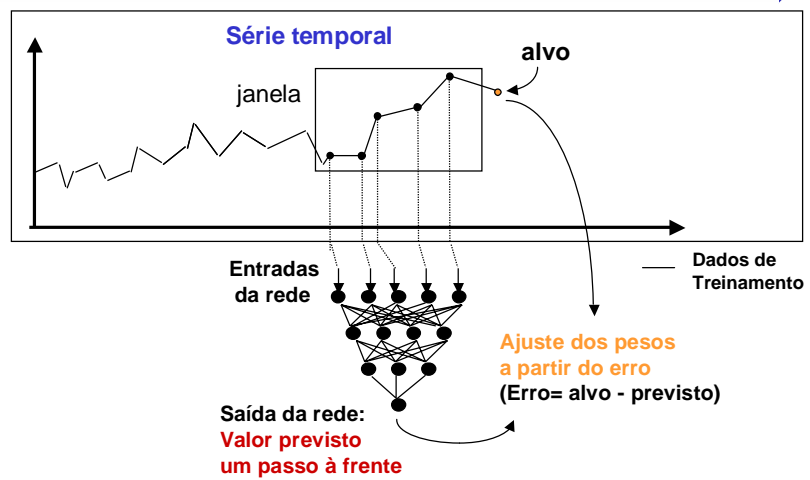
Processo de Aprendizado



Processo de Aprendizado



Processo de Aprendizado

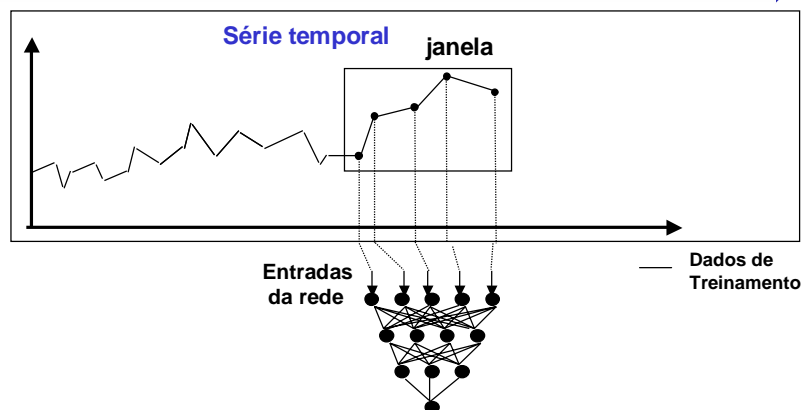


Processo de Generalização

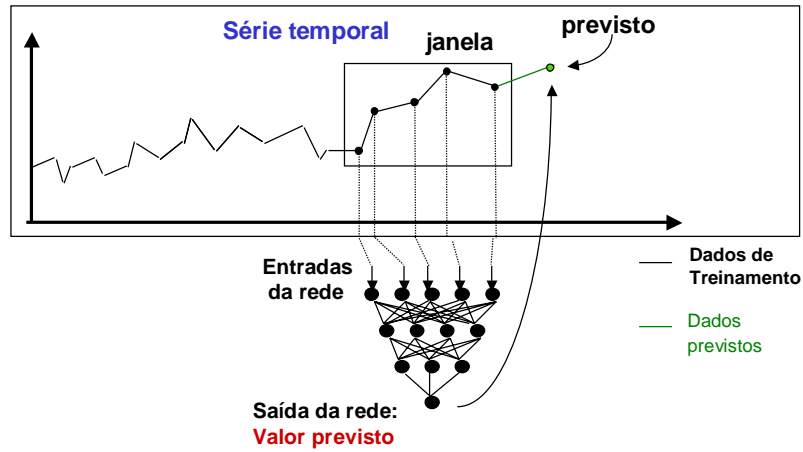
Recuperação da Informação
Aprendida



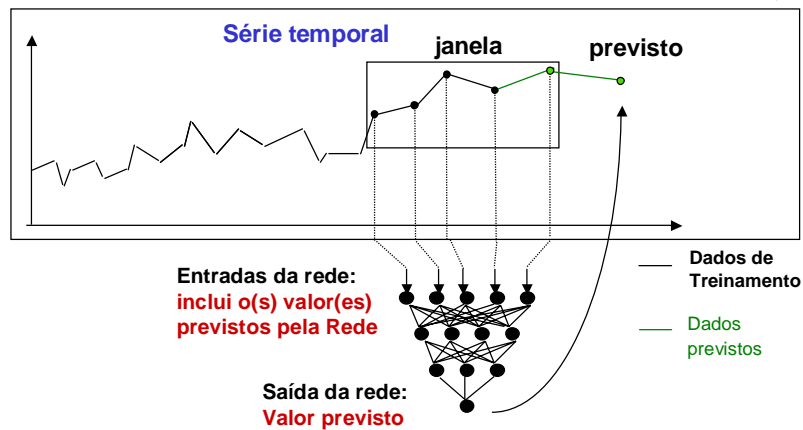
Processo de Generalização



Processo de Generalização



Processo de Generalização



Processo de Generalização

