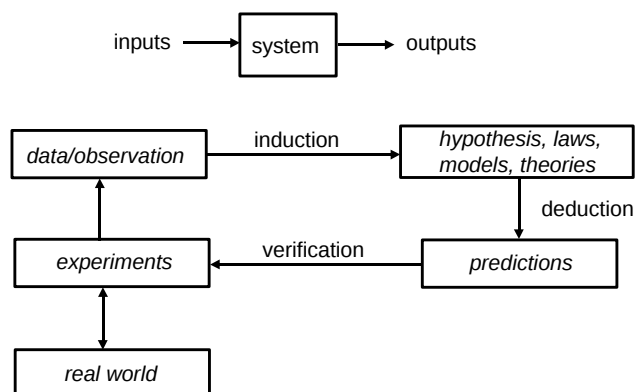


## Neural Computation

Two different perspectives:

- Artificial Intelligence and Cognitive Science
- Statistics, Mathematics

### **Models - the good, the bad, and the ugly:** the nature of the modelling enterprise



### Some historical background - 1

- McCulloch and Pitts (1943):
  - A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 9, 127-147;
- D.O. Hebb (1949):
  - concerned with adaptation laws involved in neural systems,
  - recurrent networks;
- Rosenblatt (1958):
  - coined the name *Perceptron*;
- B. Widrow (1962):
  - Adaline, Madaline,
  - LMS learning rule with multiple adaptive elements;
- M.L. Minsky and S.A. Papert (1969):
  - rigorous analysis of Perceptrons proving many properties and limitations;

### Some historical background - 2

- S. Grossberg (1976):
  - novel hypothesis about the underlying principles governing neural systems;
- J.J. Hopfield (1982):
  - considered the one that contributed to legitimize the field of neural networks,
  - recurrent networks (Hopfield networks or Amari-Hopfield models );
- T. Kohonen (1982):
  - self-organizing systems

## Redes Neurais Artificiais

- Uma definição:
  - uma rede neural artificial é um sistema composto por muitas unidades simples de processamento operando em paralelo, cuja função é determinada pela estrutura da rede, valor das conexões e o processamento realizado por cada unidade

## Inspiração em modelos biológicos

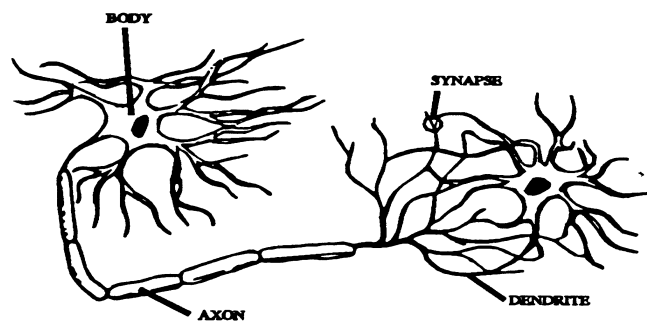
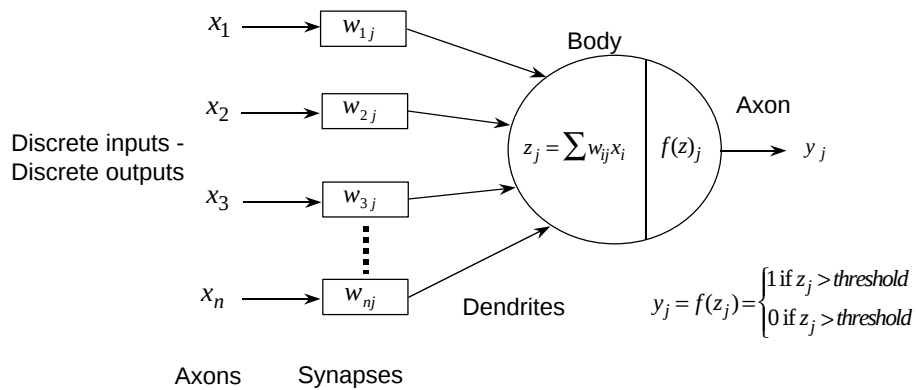


Figure 1: A Biological Neuron

The beginning: the McCulloch-Pitts neuron  
also known as: threshold neuron, threshold logic unit (TLU) and Perceptron



Para que servem ?

- A maior parte das Redes Neurais Artificiais mais utilizadas (como Multilayer Percetron Networks e Radial Basis Function Networks) estão diretamente relacionadas com métodos estatísticos e constituem modelos de regressão não lineares.

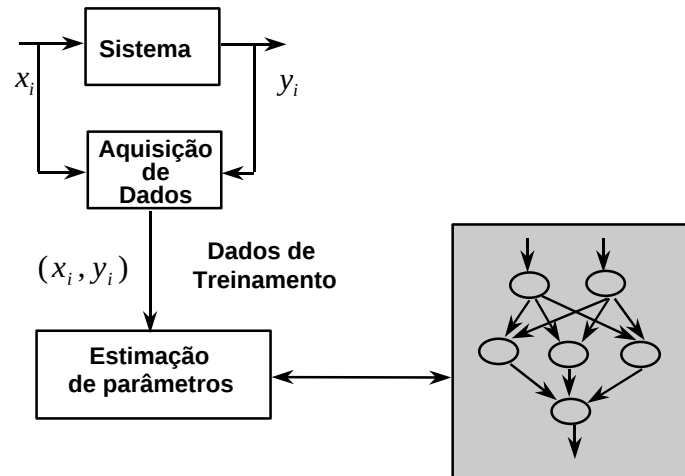
## Para que servem ?

- Podem ser utilizados para:
  - classificação:
    - reconhecimento de padrões,
    - deteção de falhas;
  - aproximação de funções não lineares:
    - sistemas estáticos,
    - sistemas dinâmicos.

## Características

- Vantagens
  - Não é necessário grande quantidade de informação a priori;
  - Alta flexibilidade (aproximadores universais): dada a complexidade adequada (número de parâmetros) pode-se aproximar qualquer função não linear;
- Desvantagens:
  - Alta dependência em relação aos dados de treinamento;
  - Poder de generalização limitado ???????.

## Aprendizagem: estimação de parâmetros



## Tipos de redes neurais artificiais

- As rede neurais podem ser classificadas em dois tipos distintos de acordo com o método de aprendizagem (estimação de parâmetros):
  - Aprendizado não supervisionado: aprendizagem autônoma sem a utilização de uma função de medida do erro;
  - Aprendizado supervisionado: utilização de uma função de medida do erro.

## Tipos de redes neurais artificiais

- De acordo com a estrutura da rede:

– Feedforward nets



– Feedback nets



## Tipos de redes neurais artificiais

- Aprendizado não supervisionado

– Feedback nets

- Continuous Hopfield (CH)
- Kohonen Self Organizing Maps (SOM)

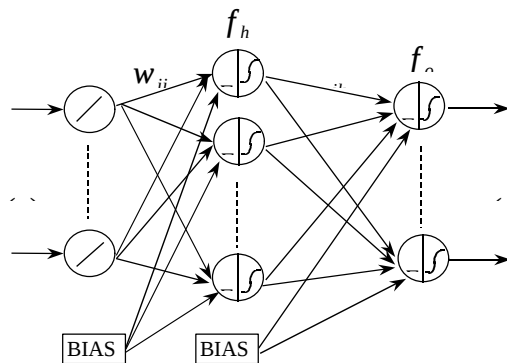
– Feedforward nets

- Fuzzy Associative Memories (FAM)
- Counterpropagation (CPN)

## Tipos de redes neurais artificiais

- Aprendizado supervisionado
  - Feedback nets
    - Boltzmann machines (BM)
    - Mean Field Annealing (MFT)
  - Feedforward nets
    - Multi-Layer Perceptron (MLP)
    - Radial Basis Function (RBF)

## Multi-layer Perceptrons



$$y^k = f_o \left( a^k + \nabla w^{jk} f_h \left( a^j + \nabla w^{ij} x^i \right) \right)$$

## Multi-layer Perceptrons

⇒  $f_h$  é usualmente a função logística

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

⇒  $f_o$  pode ser linear, logística ou limite

⇒ O número de unidades na hidden layer pode variar

## Associative memory networks

- Lattice networks
  - Neuro-fuzzy models
  - B-splines networks
- Non-lattice networks
  - Radial Basis Function (RBF) networks
  - ASMOD
  - ABBMOD

## Modelos neuro-fuzzy

conjunto de regras de inferência

$$\left\{ \begin{array}{l} \text{Rule}_q: \text{IF } x_1 \text{ is } A_1^q \text{ AND } x_2 \text{ is } A_2^q \text{ AND } \dots \text{ AND } x_n \text{ is } A_n^q \text{ THEN } y \text{ is } w_q \\ q = 1, \dots, Q \\ \tilde{x} = (x_1, \dots, x_l, \dots, x_n) \end{array} \right.$$

funções de pertinência

$$\left\{ \begin{array}{l} A_l^q \rightarrow f_{A_l^q}(x_l) \\ f_q(\tilde{x}) = \prod_{l=1}^n f_{A_l^q}(x_l) \end{array} \right.$$

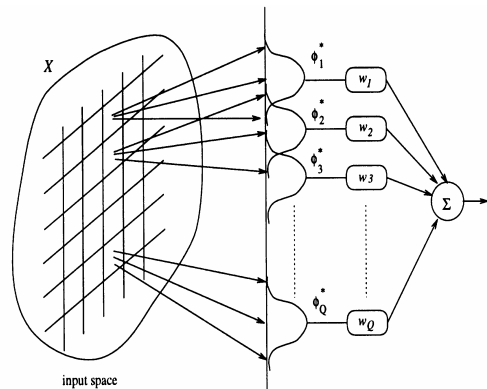
algoritmo de inferência

$$\left\{ \begin{array}{l} y = h(\tilde{x}) = \frac{\sum_{q=1}^Q w_q f_q(\tilde{x})}{\sum_{q=1}^Q f_q(\tilde{x})} \end{array} \right.$$

Modelos neuro-fuzzy são equivalentes a redes de memória associativa

$$h(\tilde{x}) = \sum_{q=1}^Q w_q f_q^*(\tilde{x})$$

$$f_q^*(\tilde{x}) = \frac{f_q(\tilde{x})}{\sum_{q=1}^Q f_q(\tilde{x})}$$



## CMAC - 1

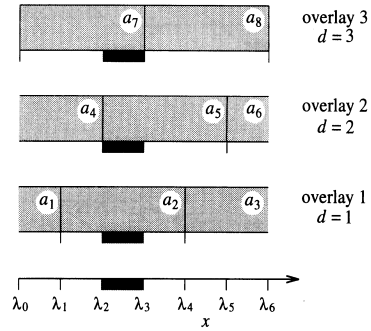


Figure 6.2 Univariate CMAC overlays with  $\rho = 3$ . Each overlay is composed of adjacent non-overlapping supports, and the input (which lies in the third interval) lies in the support of  $a_2, a_5$  and  $a_7$ .

## CMAC - 2

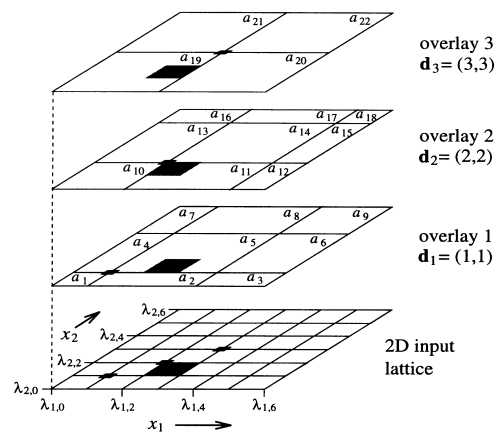


Figure 6.3 2-dimensional CMAC overlays with  $\rho = 3$ . The overlays are displaced relative to each other in *both* dimensions, and the input lies in the support of  $a_5, a_{11}$  and  $a_{19}$ .

## CMAC - 3

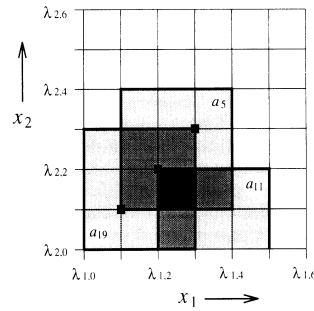


Figure 6.4 CMAC generalisation in a two-dimensional input space, where the generalisation parameter is 3. The black cell shows the input and is used to denote the cells which map to three common active basis functions. The dark grey, light grey and white areas show those cells which have two, one and zero active basis functions in common with the original input. The thick lines denote the support of the input's active basis functions.

## CMAC - 4

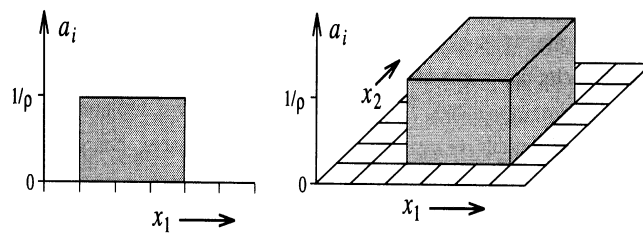


Figure 6.7 One and two-dimensional CMAC binary basis functions,  $\rho = 3$ .

### Algumas aplicações comerciais

- Controle de processos;
- Reconhecimento óptico de caracteres;
- Comunicações:
  - compressão de imagens;
  - roteamento de tráfego de informações em redes de computadores;
- Predições de séries temporais de mercados financeiros; etc.

### Exemplo 1: Estabilização de aeronaves

- Accurate Automation Corp.
  - Um simulador virtual de aeronaves foi projetado;
  - Diferentes situações foram simuladas e as ações dos pilotos que conseguiam estabilizar a aeronave foram armazenadas;
  - Dados armazenados são utilizados como dados de treinamento da RNA.

### Exemplo 2: Controle reconfigurável

- NASA Ames Research Center e Dryden Flight Research Center
  - Um controlador híbrido contendo uma RNA e um controlador tradicional foi capaz de aterrizsar um avião MD11 numa simulação de falha hidráulica total.

### Exemplo 3: Previsão de demanda de energia elétrica

- Companhia: Bayernwerke, Bavaria, Alemanha;
- Aplicação: Previsão da demanda de energia elétrica para otimização do sistema de gerenciamento de energia;
- Benefícios:
  - erros da ordem de 1-2% (regressão múltipla erros da ordem 2-3%). Obs: 1% = 10MW.

## Exemplo 4: Reconhecimento automático de escrita manual

- Helvetia Krankenkasse, Suíça.
  - Companhia de seguro saúde com 1.4 milhões de clientes;
  - Divisão da Swiss Care Organization com 2.4 milhões de clientes (30% da população Suíça).
  - Recebe cerca de 6 milhões de processos por ano (cerca de 25 mil por dia);
  - Possui 1300 escritórios regionais e 11 centros administrativos

## Exemplo 4

- Método de processamento antigo:
  - Centros de digitação transferiam as informações dos formulários a uma taxa de 150 por hora. Terceirização de cerca de 3 milhões de processos a um custo de \$0.75 por unidade;
- Novo método:
  - Formulários são lidos através de SCANNERS e caracteres são identificados através de um software de reconhecimento de caracteres implementado por redes neurais;

## Exemplo 4

### – Benefícios:

- apenas 3 em cada 10.000 caracteres são reconhecidos incorretamente;
- são processados 400 processos por hora;
- nenhum serviço é terceirizado;
- economia de \$2.300.000 por ano.

## Modelagem de Sistemas

### Sistemas bem definidos

- Modelagem baseada em modelos matemáticos bem definidos
- Modelos paramétricos
- Sistemas lineares

### Conexionismo

- Modelos black-box (Aprendizagem comandada pelo conjunto de dados)
- Sistemas não lineares

### Sistemas mal definidos

- Processamento simbólico
- Modelos não matemáticos

Quantitativo

Redes Neurais

Qualitativo

Lógica Fuzzy (Nebulosa)

## Modelling types according to the level of prior knowledge



**White box models:** model is perfectly known by prior knowledge and physical insight;



**Grey box models:** some physical insight is available but several parameters must be estimated from observed data.

- Physical modelling:** model structure can be built on physical grounds and parameters must be estimated from data;

- Semi-physical modelling:** physical insight is used to suggest certain nonlinear combinations of measured data signal. These new signals are then used to model black box structures.



**Black box models:** no physical insight is available.