Artificial Neural Networks:
an introduction

Prof Sébastien Marcel
Senior researcher
www.idiap.ch/~marcel

Idiap Research Institute
Martigny, Switzerland
www.idiap.ch

January 25, 2010

Outline

Outline

Learning and Learning

XXX

- Learning by heart:

$$\begin{array}{ll} 1+0=1 & 1\times 0=0 \\ 1+1=2 & 1\times 1=1 \\ 1+2=3 & 1\times 2=2 \\ \ldots & \ldots \end{array}$$

any computer can do that !

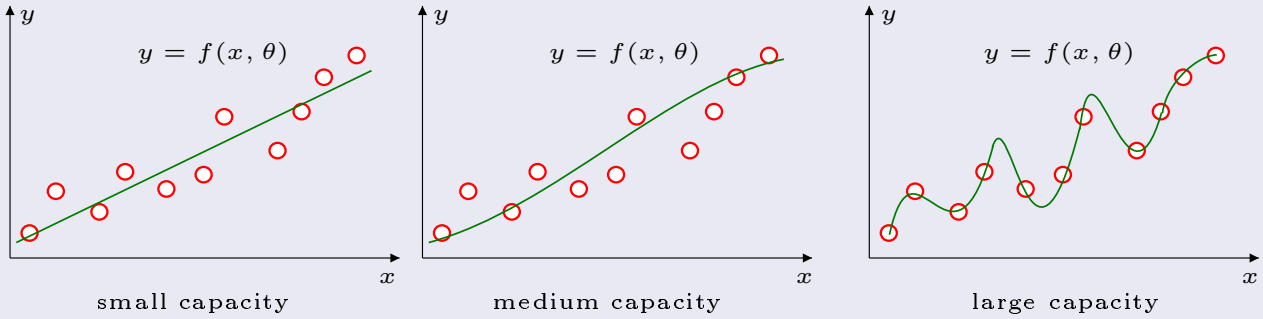
- Learning by heart is not learning:
  - learning is "to gain knowledge or understanding of or skill in by study, instruction, or experience" (Merriam-Webster),
  - the difficulty of learning is to be able to **generalize**.

Outline

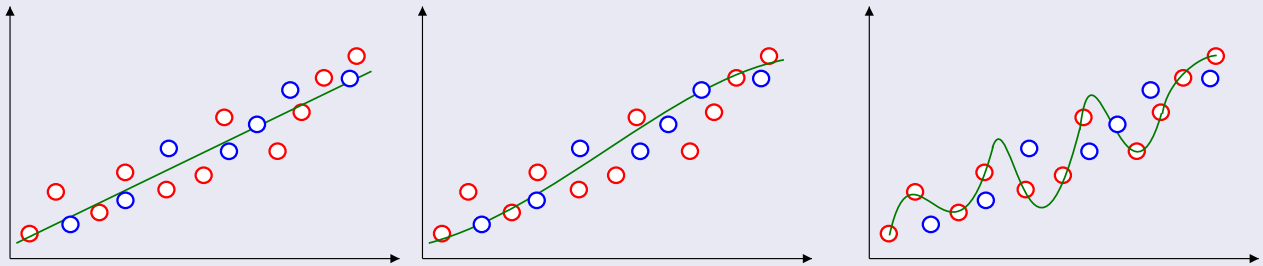
Capacity and Generalization

XXX

- Capacity: # of parameters (θ) required to fit the data with a function



- Generalization: the performance of the above function on unseen data

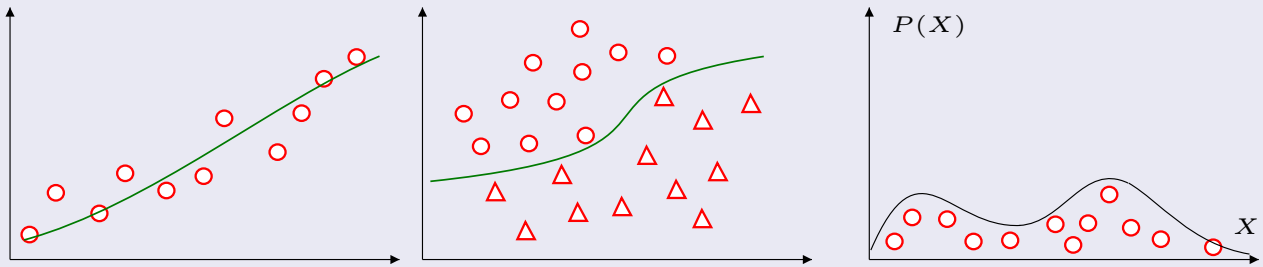


Outline

Regression, Classification and Density Estimation

XXX

- there are 3 kinds of problems:



regression, classification and density estimation

- Machine Learning Algorithms address the above problems using various tools:

XXX

- Artificial Neural Networks,
- Support Vector Machines,
- Gaussian Mixture Models,
- Hidden Markov Models

Outline

Applications

XXX

- in Computer Vision:
  - Face detection, face recognition, face orientation estimation,
  - Gesture recognition,
  - Optical character recognition,
  - Handwritten recognition.
- in Speech Processing:
  - Speech recognition,
  - Speaker recognition.
- but also in Finance, Telecoms, Games, Robotic and more ...

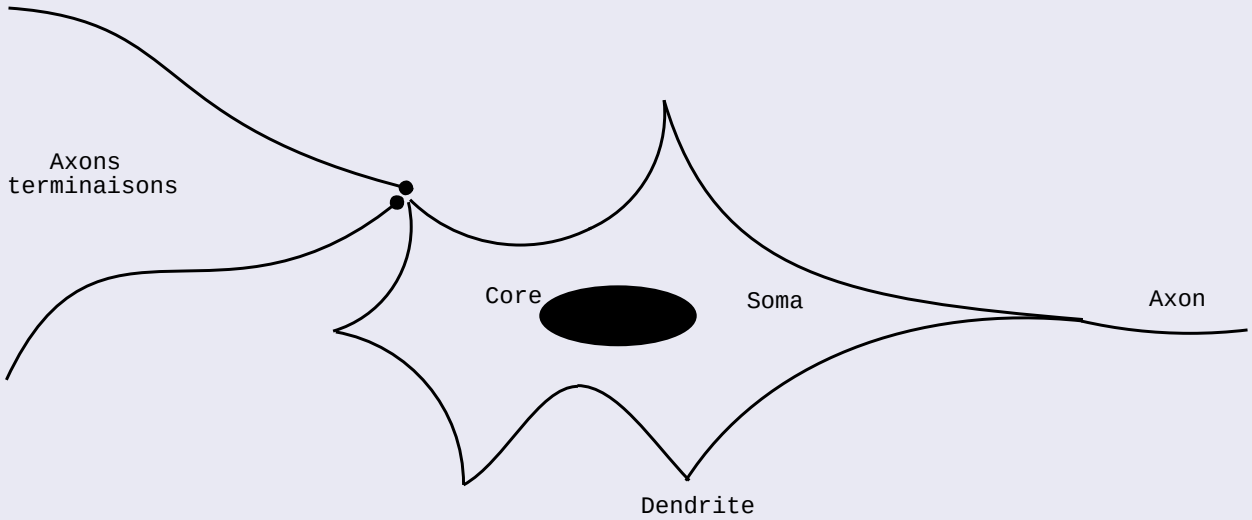
Outline

Outline

ANN: Biological Bases

XXX

- the brain:
  - 10^{12} neurons massively connected,
  - 1 neuron is connected with 10^3 others in average,
- the neuron:



  - core: where DNA belongs (approx. 30000 genes)
  - dendrite: receives a signal from other neurons (via their axon),
  - axon: propagates the signal to other neurons.

Outline

History of ANN

XXX

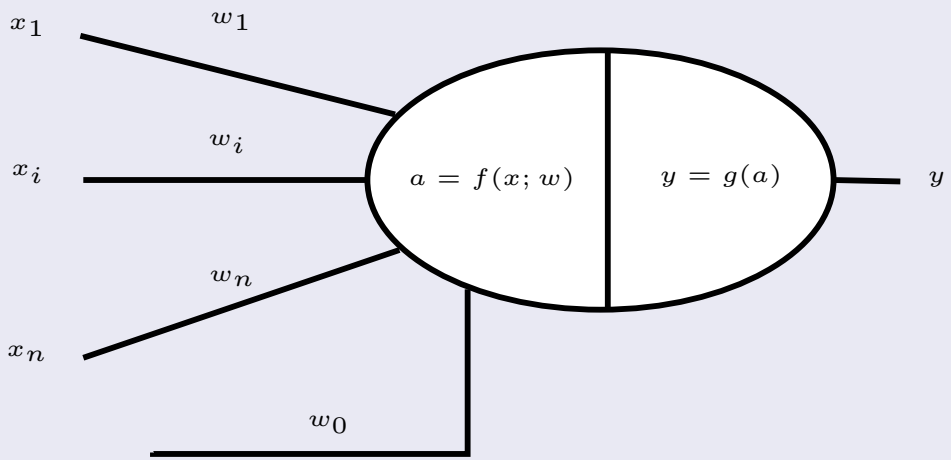
- Mc Culloch and Pitts (1943): formal neuron inspired from biology
- Hebb (1949): the first training rule
- Rosenblatt (1962): the Perceptron
- Minsky and Papert (1969): limitations of Perceptron
- then, nothing during 16 years, research goes for symbolic AI
- Hopfield (1982): auto-associative memory
- Werbos (1974), Rumelhart (1986), Parker (1985), Le Cun (1985): Multi Layer Perceptron and Back-propagation algorithm
- Kohonen (1995): auto-organizing maps

Outline

The Formal Neuron

XXX

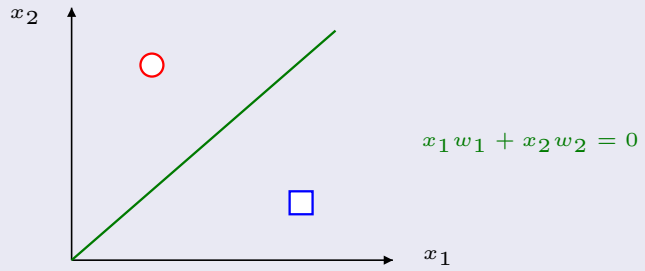
- Mc Culloch and Pitts (1943):
  - x is the input $\in \mathbb{R}^n$,
  - $w_1 ... w_n$ are the weights,
  - w_0 is the bias,
  - a is the result of the integration function $f(x; w) = \sum_{i=1}^{n} w_i x_i + w_0$,
  - y is the output of the transfer function $g(a) = tanh(a)$.



The Role of the Bias

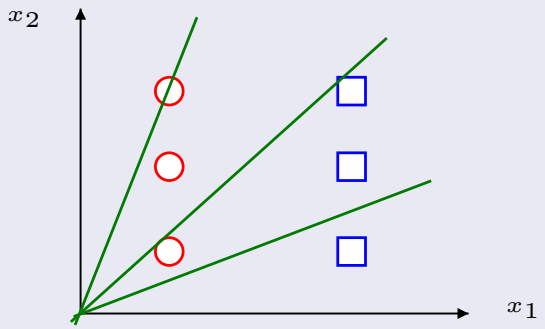
XXX

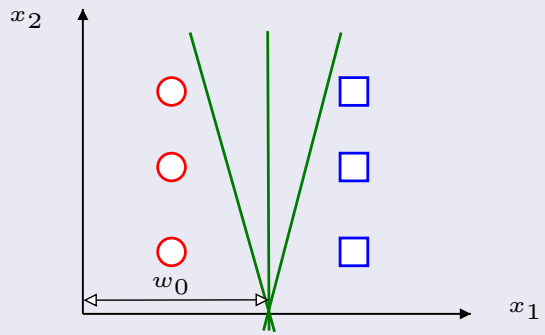
- the formal neuron is a linear separator:



$$x_1 w_1 + x_2 w_2 = 0 \quad \Leftrightarrow \quad x_2 = -\frac{w_1}{w_2} x_1$$

- without the bias the linear separation is not always possible:



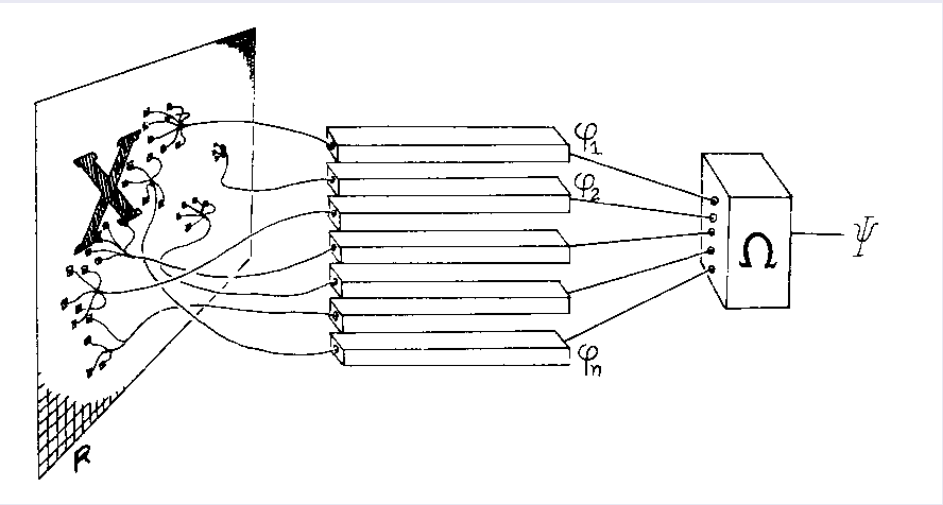


Outline

The Perceptron

XXX

- Rosenblatt (1962):

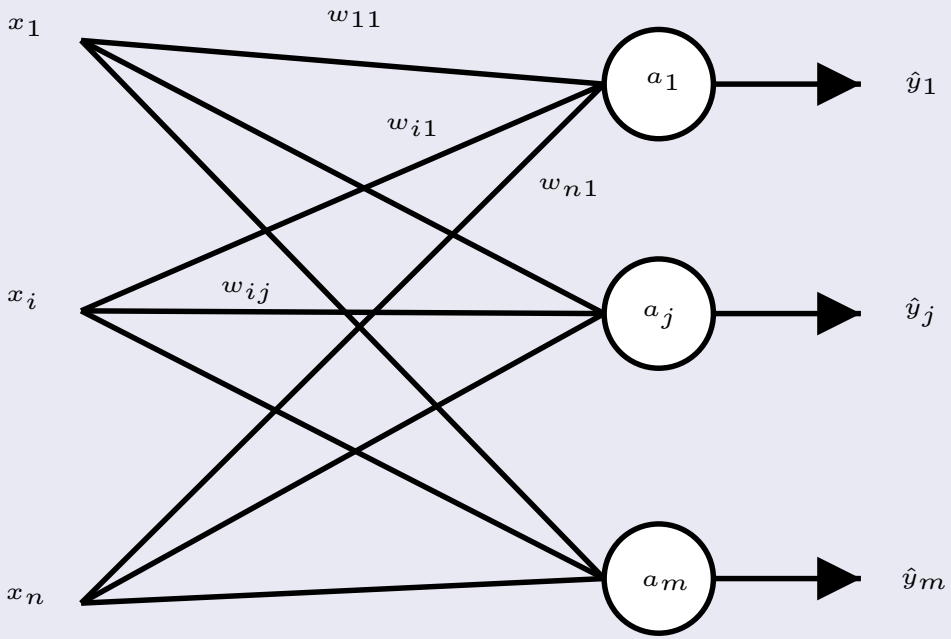


  - a retina: binary input of the perceptron,
  - association cells: "pre-processing",
  - decision cells: linear units.

The Perceptron

XXX

- Rosenblatt (1962):



- Training rules:

XXX

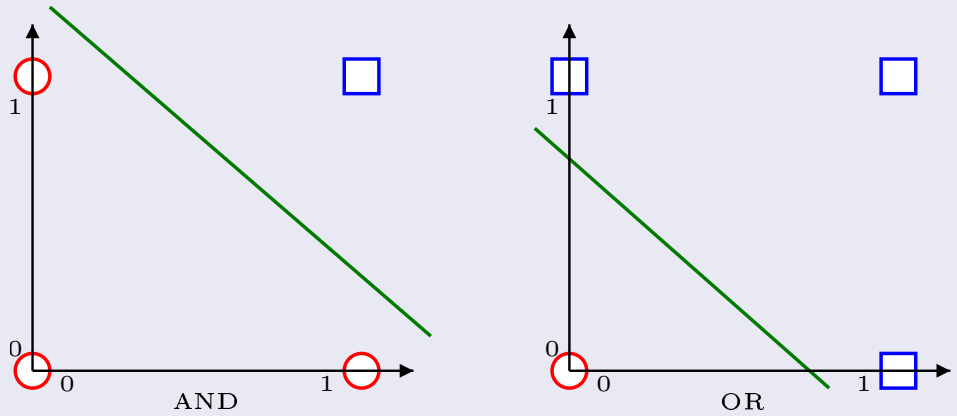
- $w_{ij}^{t+1} = w_{ij}^{t} + \alpha(\hat{y}_j - y_j)x_i$ (Rosenblatt)

Outline

Linearly Separable

XXX

- OR and AND: are linearly separable:



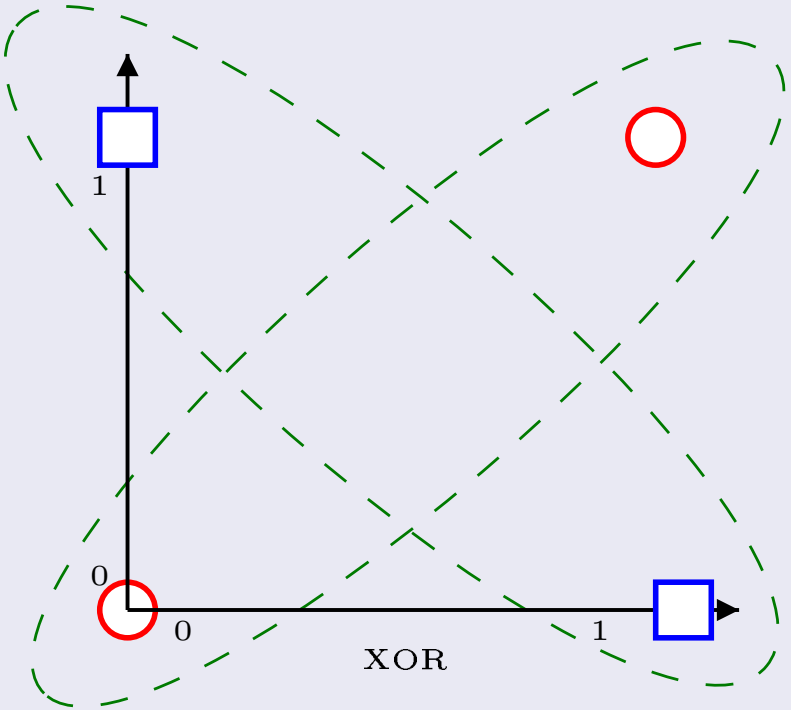
- one solution to AND: $w_1 = 1$, $w_2 = 1$ and $w_0 = 1.5$
- one solution to OR: $w_1 = 1$, $w_2 = 1$ and $w_0 = -0.5$

Outline

The problem of XOR

XXX

- XOR is not linearly separable:



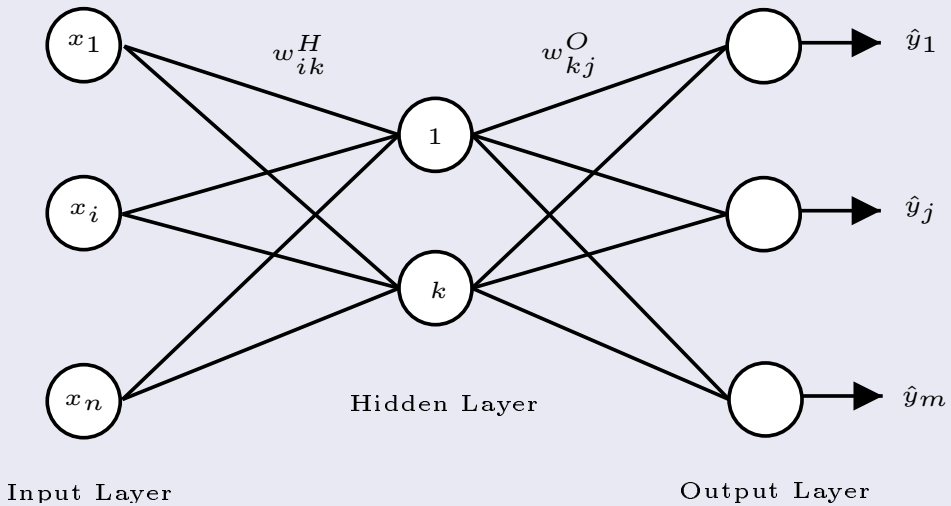
- impossible to solve $x_1 w_1 + x_2 w_2 + w_0 = 0$, but what about multiple equations $x_1 w_{11} + x_2 w_{21} + w_{01} = 0$, $x_1 w_{12} + x_2 w_{22} + w_{02} = 0$,...

Outline

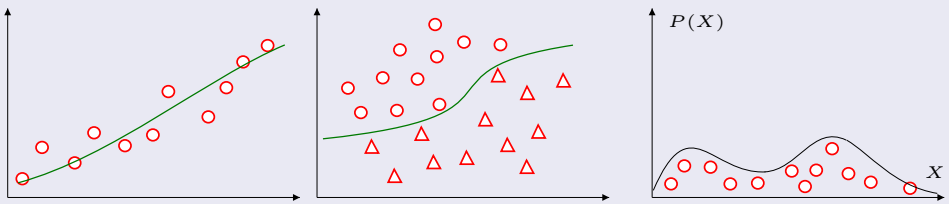
The Multi Layer Perceptron

XXX

- It contains 1 input layer, 1 or several hidden layer and 1 output layer:



- It can approximate any continuous functions,



regression, classification and density estimation