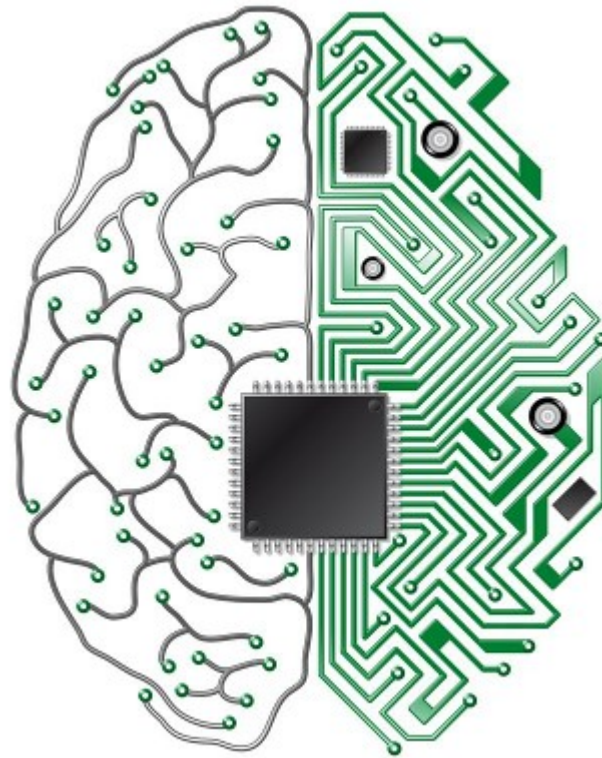


# Le reti neurali artificiali



Giacomo Trudu aka *Wicker25*

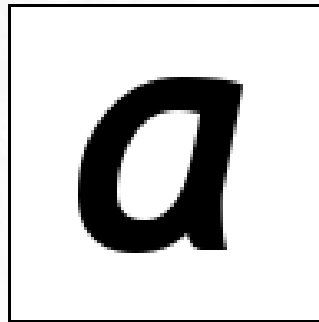
# Sommario

- L'approccio algoritmico
- Le reti neurali artificiali
- Apprendimento delle reti neurali
- Il neurone biologico
- Il percettrone
- L'apprendimento del percettrone
- Il percettrone multistrato
- La retropropagazione dell'errore

# L'approccio algoritmico

- Processo risolutivo definito, sequenziale e deterministico
- Linguaggio naturale della Macchina di Turing
- Richiede la piena comprensione del problema preso in esame (problemi definiti)
- Non necessita di dati sperimentali
- Modellistica di tipo “white box”

# L'approccio algoritmico



Quale carattere è rappresentato nell'immagine? Possiamo formalizzare il problema in un algoritmo risolutivo?

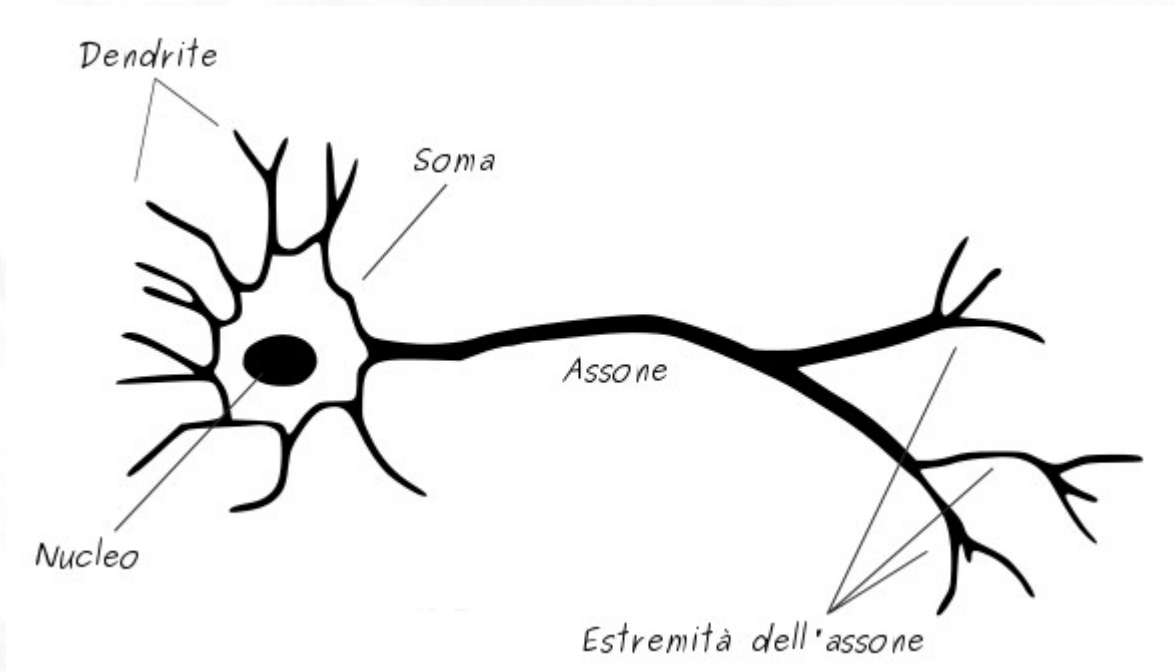
# Le reti neurali artificiali

- Sistemi per il trattamento dell'informazione
- Basate sulla teoria del connessionismo
- Riproducono alcuni processi del cervello umano (paradigma di ispirazione biologica)
- Sono modelli di regressione non lineare
- Modellistica di tipo “black box”

# Le reti neurali artificiali

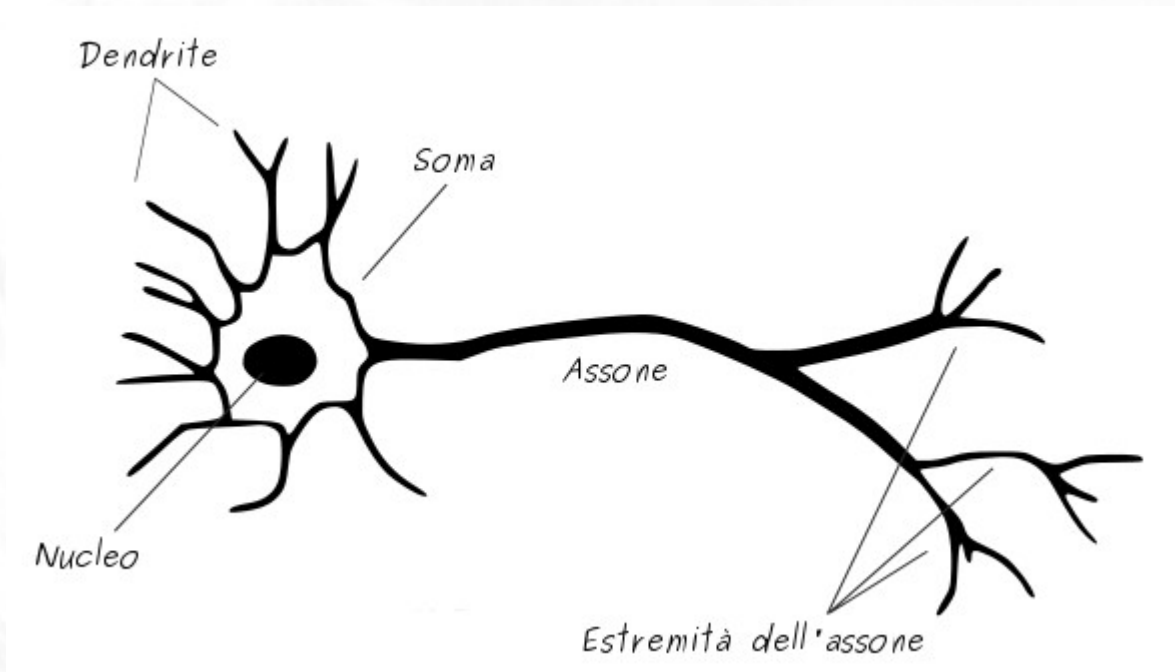
- Apprendimento basato su dati empirici
- Elaborano i dati in modo parallelo e distribuito
- Possono generalizzare le soluzioni apprese
- Possono astrarre nuove soluzioni
- Sono in grado di trattare dati “rumorosi”

# Il neurone biologico



- Nel cervello umano sono circa 100 miliardi delle dimensioni di 10 micron ciascuno

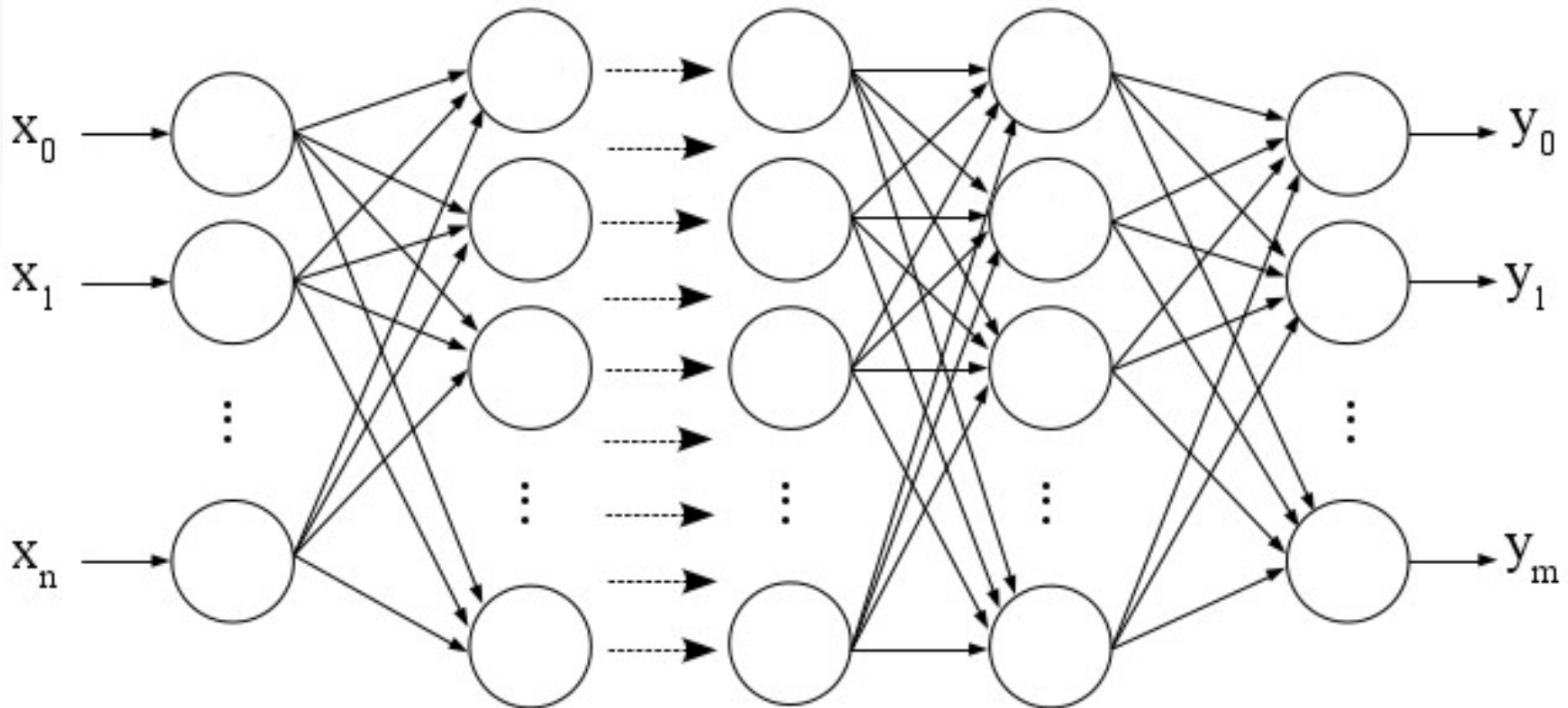
# Il neurone biologico



- Neuroni sensori
- Neuroni motori (o motoneuroni)
- Neuroni intermedi (o interneuroni)



# Le reti neurali artificiali



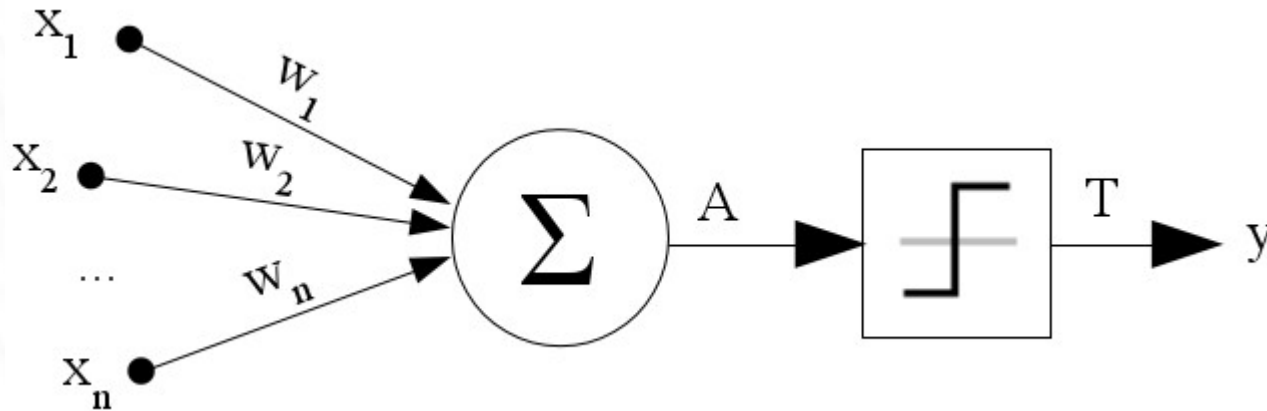
# Le reti neurali artificiali

- Reti totalmente connesse
- Reti parzialmente connesse
  
- Reti feedforward
- Reti feedback

# Apprendimento delle reti neurali

- Apprendimento supervisionato
- Apprendimento non supervisionato
- Apprendimento per rinforzo

# Il percettrone



- Funzione di attivazione  $A(w_1, x_1, w_2, x_2, \dots, w_n, x_n)$
- Funzione di trasferimento  $T(a)$

# Il perceptrone

Formula di uscita:

$$y = T \left( \sum_{i=1}^n w_i x_i - \theta \right)$$

Con l'aggiunta del *Bias*:

$$y = T \left( \sum_{i=1}^{n+1} w_i x_i \right)$$

# Il percettrone

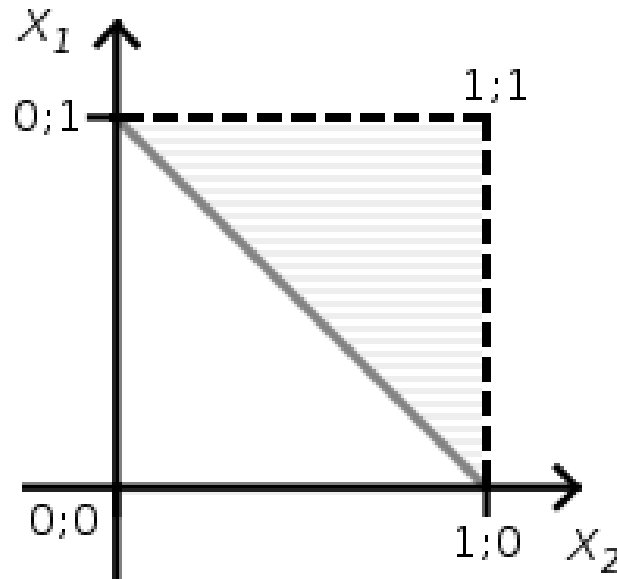
Con

$$T(x) = \begin{cases} 1, & \text{se } x > 0 \\ 0, & \text{se } x \leq 0 \end{cases}$$

Diventa

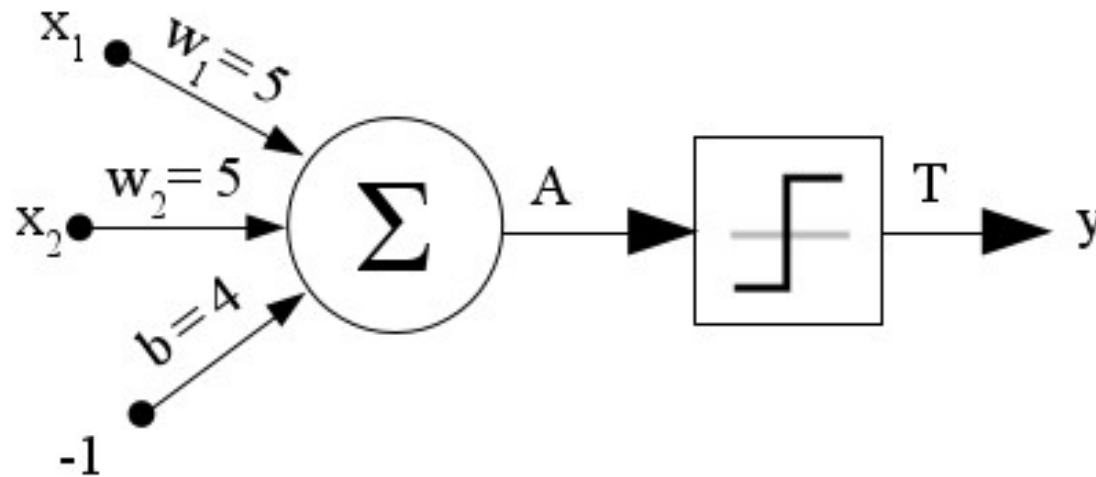
$$\sum_{i=1}^{n+1} w_i x_i = w_1 x_1 + w_2 x_2 + \dots + w_n x_n - \theta > 0$$

# Il percettrone: forma logica OR



$x_1$	$x_2$	$y$
0	0	0
1	0	1
0	1	1
1	1	1

# Il percettrone: forma logica OR



$x_1$	$x_2$	$5x_1 + 5x_2 - 4 > 0$	$y$
0	0	$5 \cdot 0 + 5 \cdot 0 - 4 > 0$	0
1	0	$5 \cdot 1 + 5 \cdot 0 - 4 > 0$	1
0	1	$5 \cdot 0 + 5 \cdot 1 - 4 > 0$	1
1	1	$5 \cdot 1 + 5 \cdot 1 - 4 > 0$	1



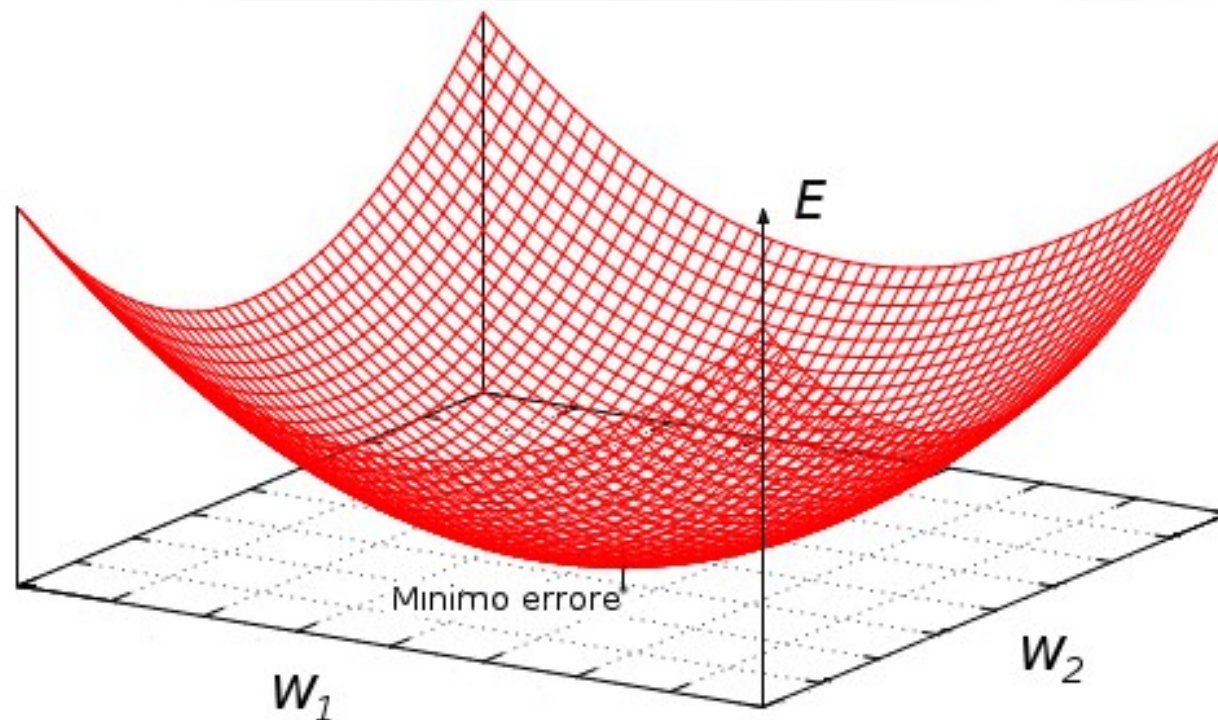
# Apprendimento del percettrone

Dato un insieme di addestramento non vuoto formato da “S” coppie ingressi-uscite, si definisce una funzione di errore (o funzione costo)

$$E = \sum_{c=1}^s \frac{1}{2} (d_c - y_c)^2$$

# Apprendimento del percettrone

In una rete a due ingressi



# Apprendimento del perceptrone

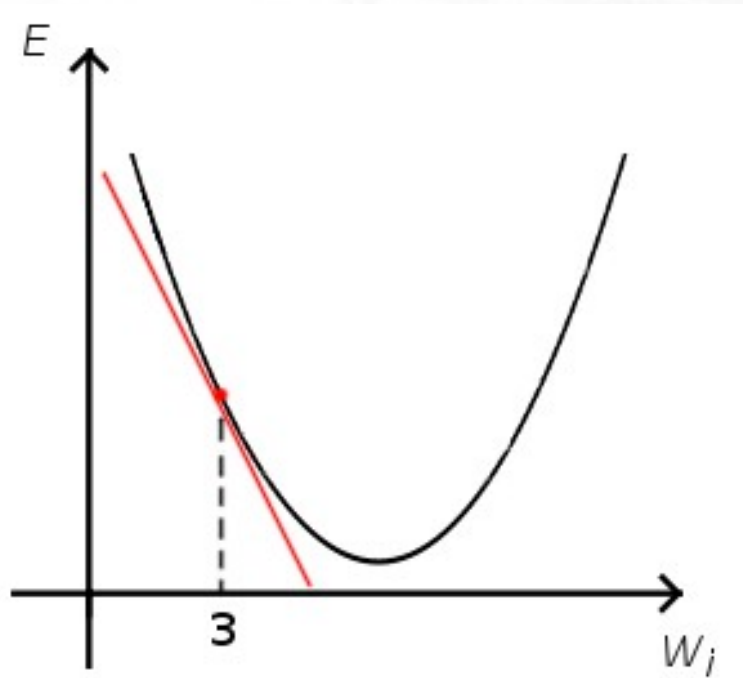
Calcoliamo il gradiente della funzione costo:

$$\nabla E(\mathbf{w}) = \left[ \frac{\partial E}{\partial w_0}, \frac{\partial E}{\partial w_1}, \dots, \frac{\partial E}{\partial w_n} \right]$$

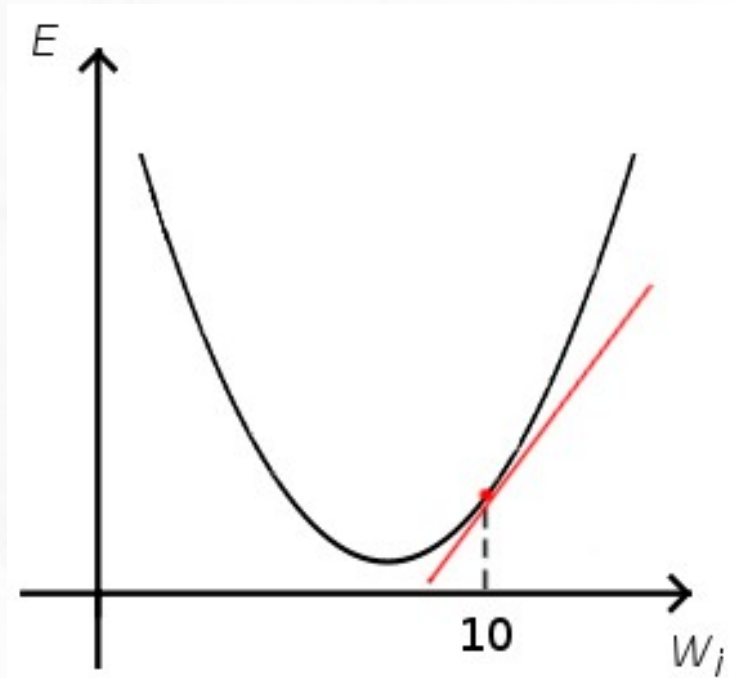
Correggiamo i pesi sinaptici sulla base della derivata parziale del peso:

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i}; \quad w_i^t = w_i^{t-1} + \Delta w_i$$

# Apprendimento del percettrone



$$\frac{\partial E}{\partial w_i} = -2$$



$$\frac{\partial E}{\partial w_i} = 1$$

# Apprendimento del perceptrone

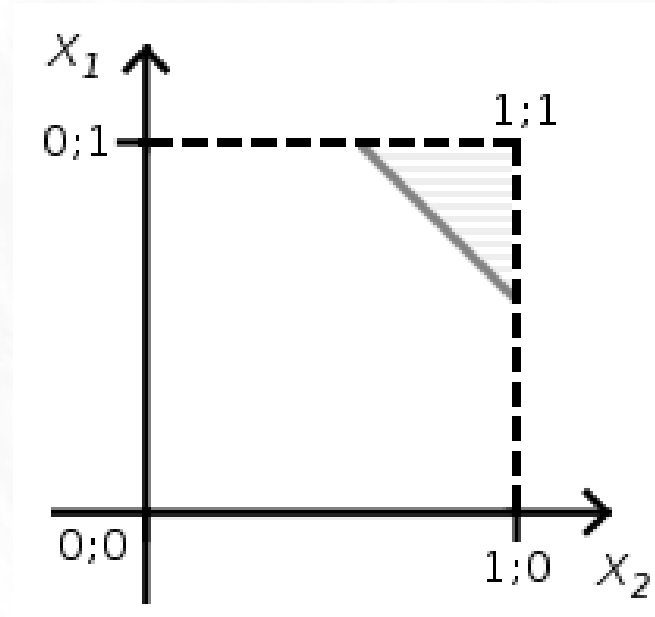
Applicando *la regola di derivazione delle funzioni composte* otteniamo:

$$\frac{\partial E}{\partial w_i} = -(d - y) x_i$$

Considerando tutto l'insieme di addestramento diventa:

$$\frac{\partial E}{\partial w_i} = \sum_{c=1}^S [-(d_c - y_c) x_{ci}]$$

# Il percettrone: forma logica AND



$x_1$	$x_2$	$y$
0	0	0
1	0	0
0	1	0
1	1	1

# Il percettrone: forma logica AND

Apprendimento:

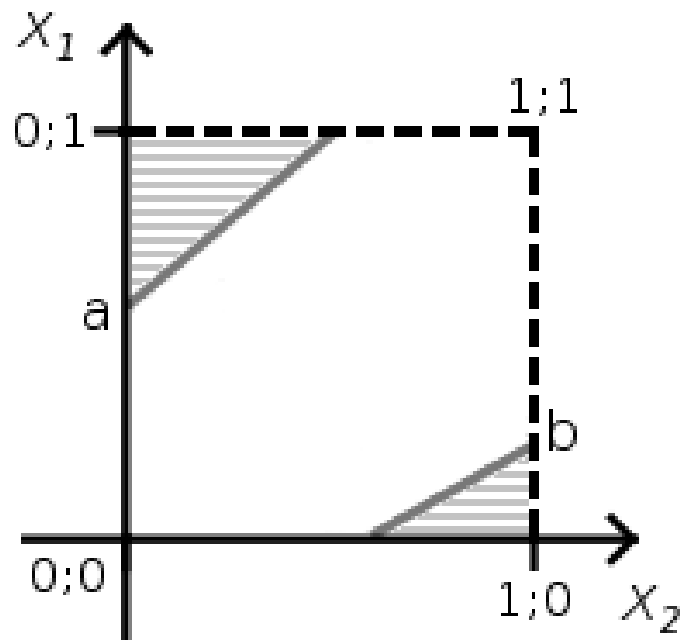
Epoche	$w_1$	$\partial E/\partial w_1$	$w_2$	$\partial E/\partial w_2$	$w_b$	$\partial E/\partial w_b$	E
0	7,00	-1,00	2,00	-1,00	15,00	1,00	0,50
1	7,50	-1,00	2,50	-1,00	14,50	1,00	0,50
2	8,00	-1,00	3,00	-1,00	14,00	1,00	0,50
3	8,50	-1,00	3,50	-1,00	13,50	1,00	0,50
4	9,00	-1,00	4,00	-1,00	13,00	1,00	0,50
5	9,50	-1,00	4,50	-1,00	12,50	1,00	0,00

Esecuzione:

$x_1$	$x_2$	$9,5 x_1 + 4,5 x_2 - 12,5 > 0$	$y$
0	0	$9,5 \cdot 0 + 4,5 \cdot 0 - 12,5 > 0$	0
1	0	$9,5 \cdot 1 + 4,5 \cdot 0 - 12,5 > 0$	0
0	1	$9,5 \cdot 0 + 4,5 \cdot 1 - 12,5 > 0$	0
1	1	$9,5 \cdot 1 + 4,5 \cdot 1 - 12,5 > 0$	1

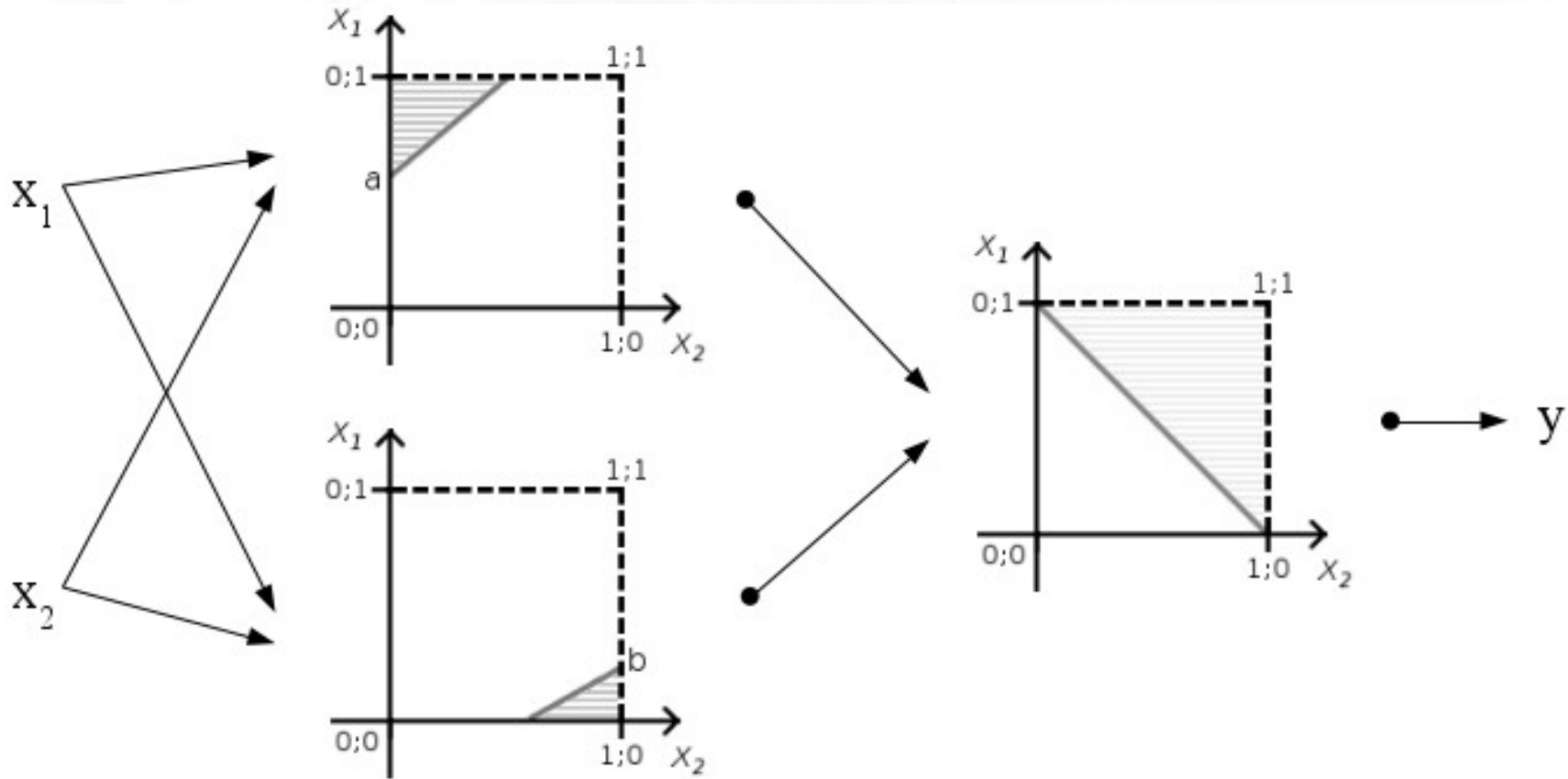
# Il percettrone multistrato

Problemi di separazione non lineare:

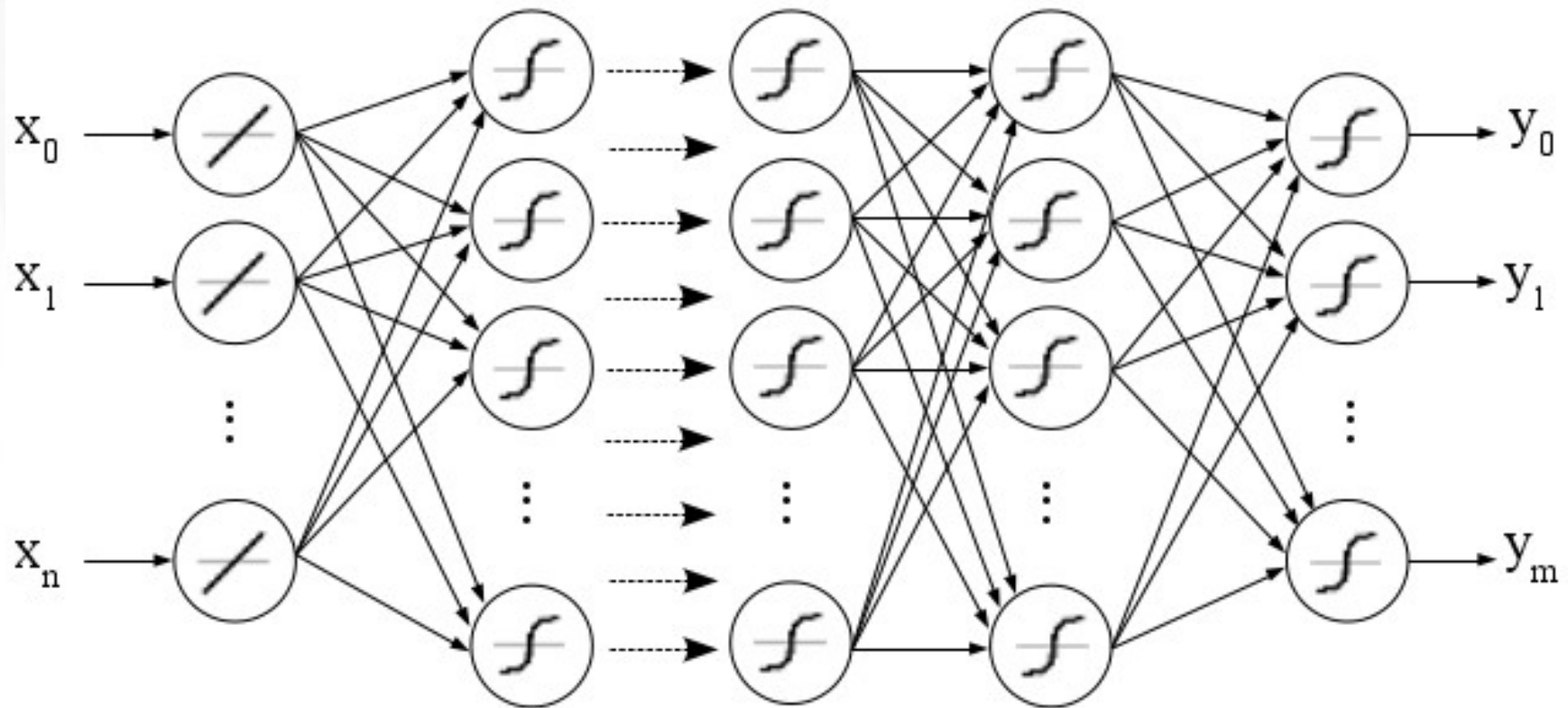




# Il percettrone multistrato



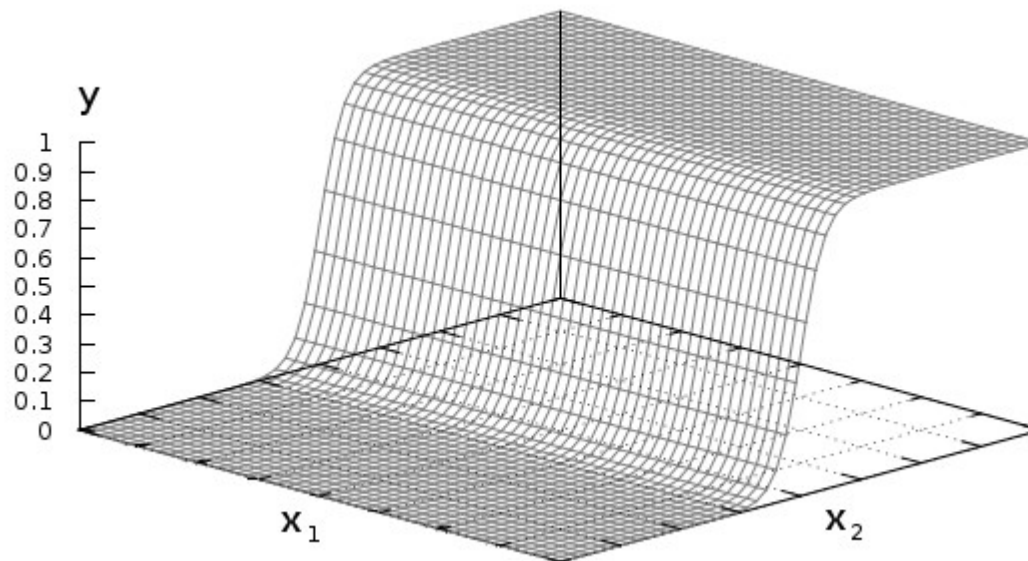
# Il perceptrone multistrato



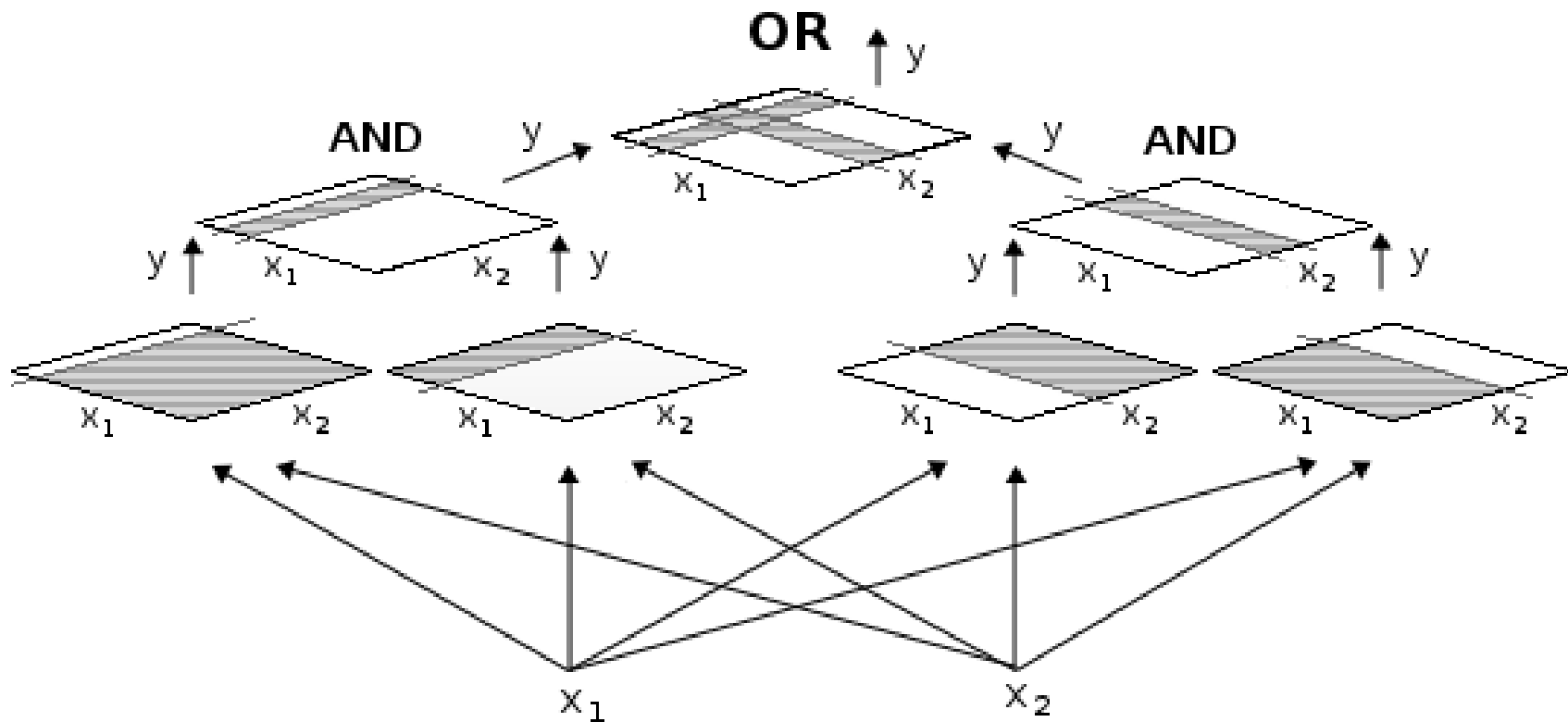
# Il perceptrone multistrato

Formula di uscita:

$$\text{sigmoid} \left( \sum_{i=1}^{n+1} w_i x_i = w_1 x_1 + w_2 x_2 + \dots + w_n x_n - \theta \right)$$



# Il percettore multistrato



# Retropropagazione dell'errore

Dato un insieme di addestramento non vuoto formato da “S” coppie ingressi-uscite, si definisce una funzione di errore (o funzione costo)

$$E = \sum_{c=1}^S \sum_{j=1}^n \frac{1}{2} (d_{cj} - y_{cj})^2$$

# Retropropagazione dell'errore

Calcoliamo il gradiente della funzione costo:

$$\nabla E(\mathbf{w}) = \left[ \dots, \frac{\partial E}{\partial w_{jk}}, \dots \right]$$

Correggiamo i pesi sinaptici sulla base della derivata parziale del peso:

$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}}; \quad w_{jk}^t = w_{jk}^{t-1} + \Delta w_{jk}$$

# Retropropagazione dell'errore

Applicando *la regola di derivazione delle funzioni composte* otteniamo:

$$\Delta w_{ki} = -\eta \delta_k z_i; \quad w_{ki}^t = w_{ki}^{t-1} + \Delta w_{ki}$$

$$\delta_k = \begin{cases} -(d_k - y_k) T'(P_k) & \text{se è un nodo di uscita} \\ \sum_{j=1}^{p+1} (\delta_j w_{jk}) T'(P_k) & \text{se è un nodo interno} \end{cases}$$

# Retropropagazione dell'errore

Considerando tutto l'insieme di addestramento diventa:

$$\Delta w_{ki} = -\eta \sum_{c=1}^S (\delta_{ck} z_{ci}); \quad w_{ki}^t = w_{ki}^{t-1} + \Delta w_{ki}$$



# Retropropagazione dell'errore

Considerando tutto l'insieme di addestramento diventa:

$$\Delta w_{ki} = -\eta \sum_{c=1}^S (\delta_{ck} z_{ci}); \quad w_{ki}^t = w_{ki}^{t-1} + \Delta w_{ki}$$

# Le reti neurali artificiali

- Riconoscimento di immagini e suoni (*face recognition, object detection, etc*)
- Diagnosi precoce di malattie (*riconoscimento di tessuti tumorali, di malattie neurodegenerative, etc*)
- Sistemi di supporto (*eye tracking, interfacce vocali, riconoscimento di cartelli stradali, etc*)
- Sistemi di sorveglianza e di tracciamento (*riconoscimento di targhe, di situazioni pericolose, etc*)

# Links

- Logo della copertina: <http://spinewave.co.nz/>
- Tesi: <http://www.hackyourmind.org/>