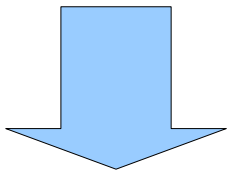


Reti Neurali

Modelli ad ispirazione biologica

Le reti neurali si ispirano al modo in cui i neuroni agiscono ed interagiscono

NB: i neuroni artificiali non sono modelli fedeli dei neuroni biologici, ne catturano solo alcuni principi



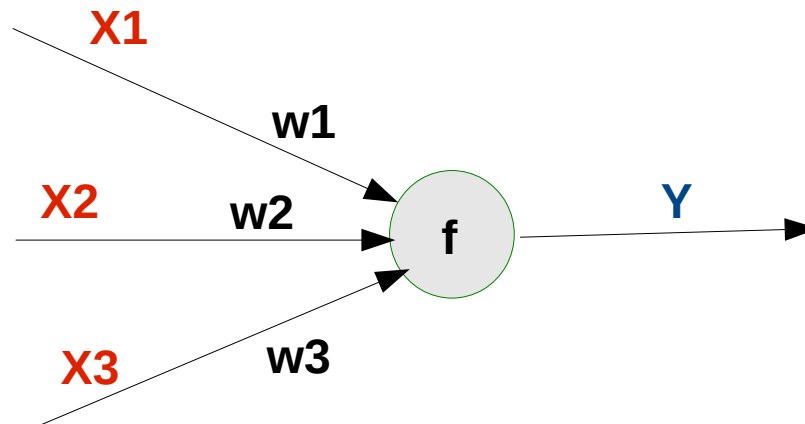
Perceptron (Rosenblatt)

B-machine di Turing

(http://www.alanturing.net/turing_archive/pages/Reference%20Articles/connectionism/Turing%27santicipation.html)

Fra i primi modelli proposti

Perceptron



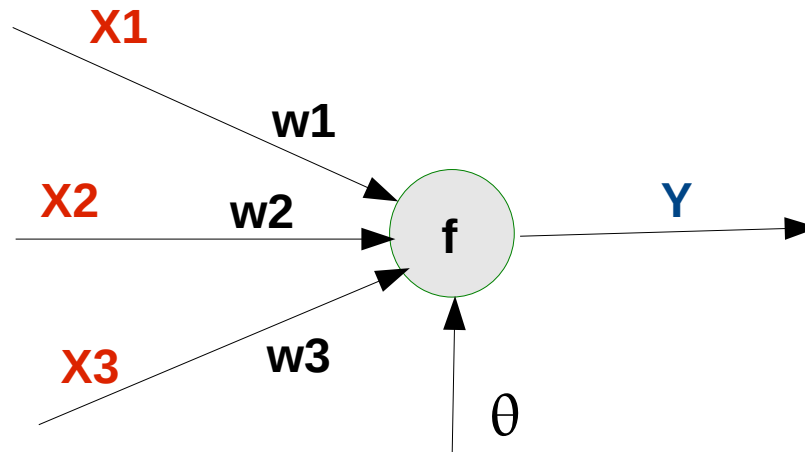
Un perceptron è un elemento computazionale, dotato di una piccola memoria in grado di calcolare una **funzione di attivazione** in esso strettamente codificata:

$$Y = f(\text{net})$$

tale funzione è calcolata su una composizione dei valori in ingresso, opportunamente pesati ...

Originariamente f era la **funzione gradino** (valori possibili per y: 0 oppure 1)

Net e sigmoide



$$net = \sum_{i=1}^n w_i X_i$$

Combinazione lineare

$$Y = f(net) = \frac{1}{1 + e^{-\alpha * (net - \theta)}}$$

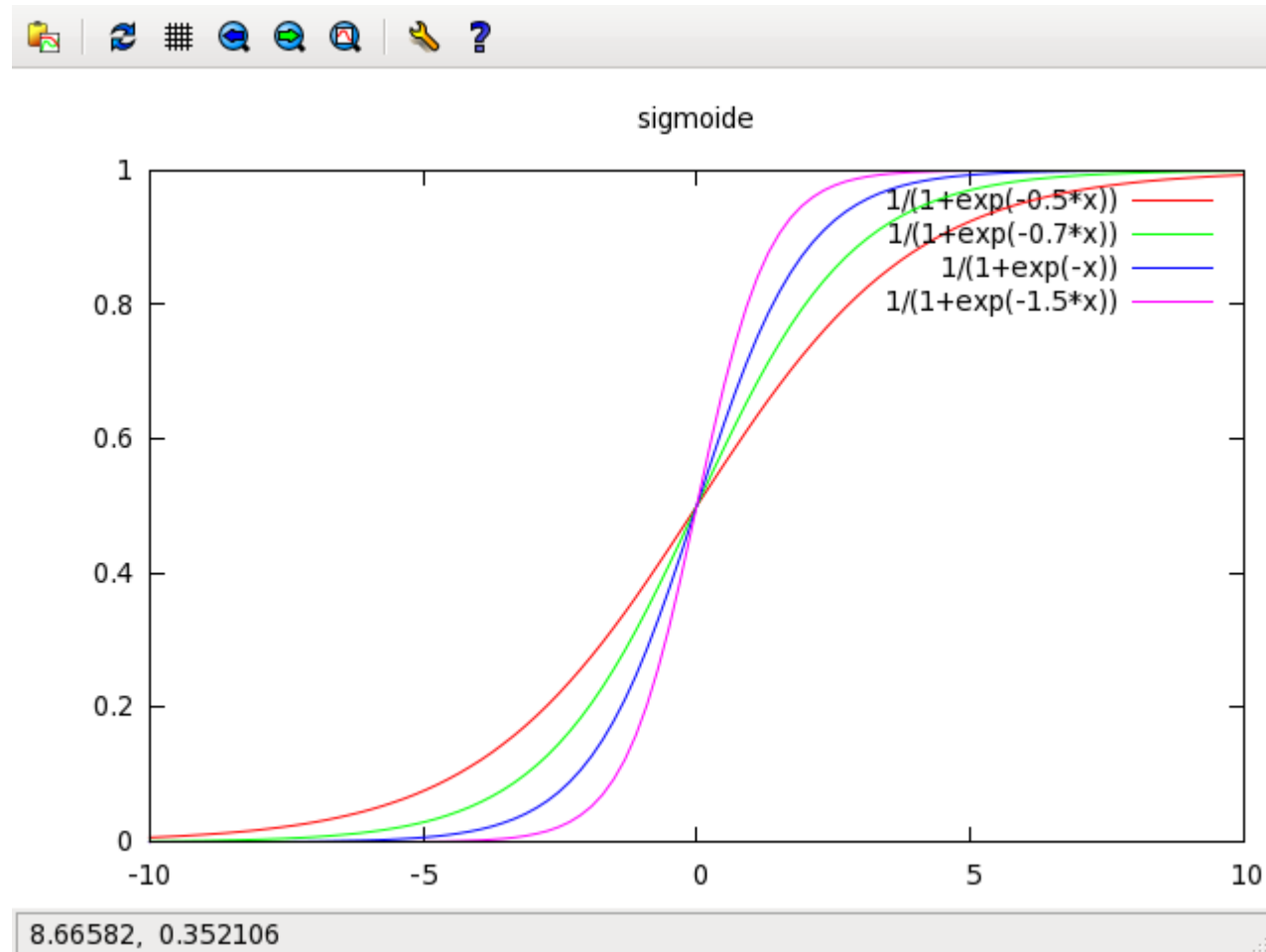
Funzione sigmoide

θ = soglia o bias
(è un parametro preimpostato)

α : controlla la pendenza

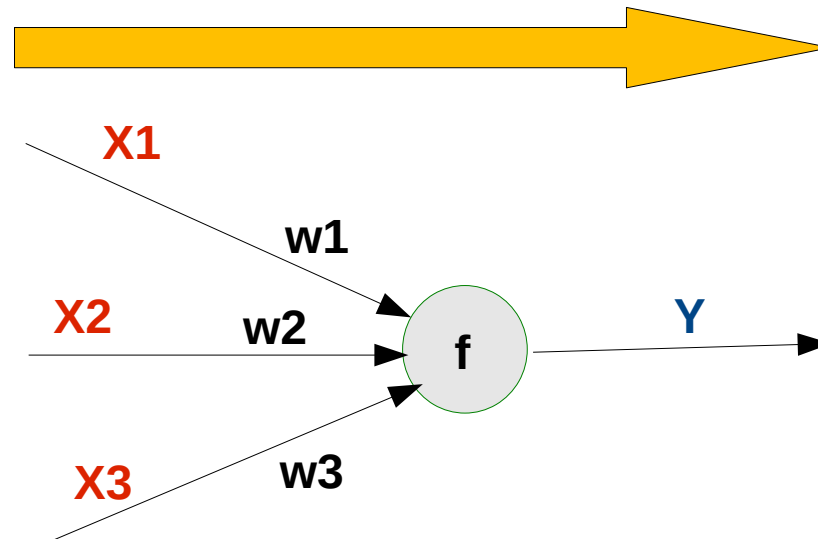
Sostituisce la funzione gradino, ha il vantaggio di essere derivabile

sigmoide



Nota: al crescere del parametro α la sigmoide tende alla funzione gradino

Passata forward

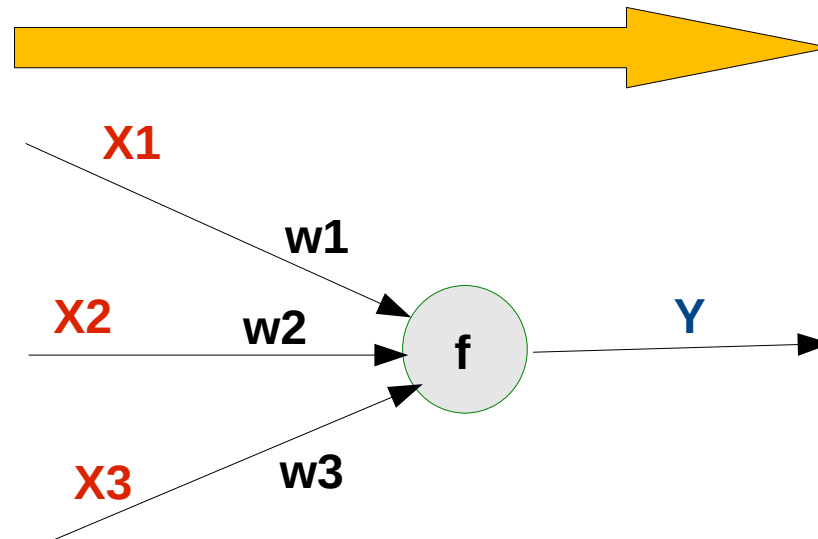


Il perceptron elabora dati che sono visti come punti in uno spazio N-dimensionale (lo spazio dei dati o spazio degli input)

N = numero degli attributi descrittivi (anche detti feature descrittive), sono variabili Indipendenti; nella figura $N = 3$

Data un'istanza questa viene propagata attraverso al neurone che calcola il valore di **Y** tramite la sua funzione di attivazione

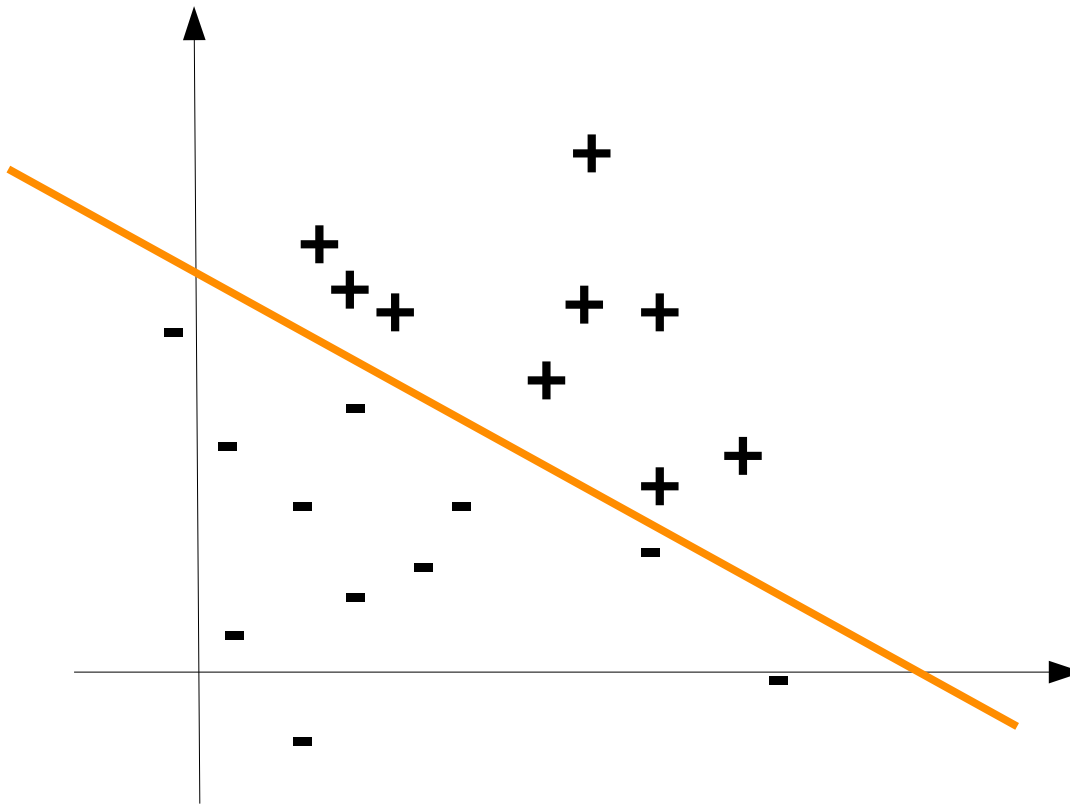
Passata forward



Possiamo vedere questo processo come una forma di **classificazione**:

Il dato ricade nella categoria 1 oppure nella categoria 0?
Dipende da cosa restituisce il perceptron

Cosa codifica un perceptron?



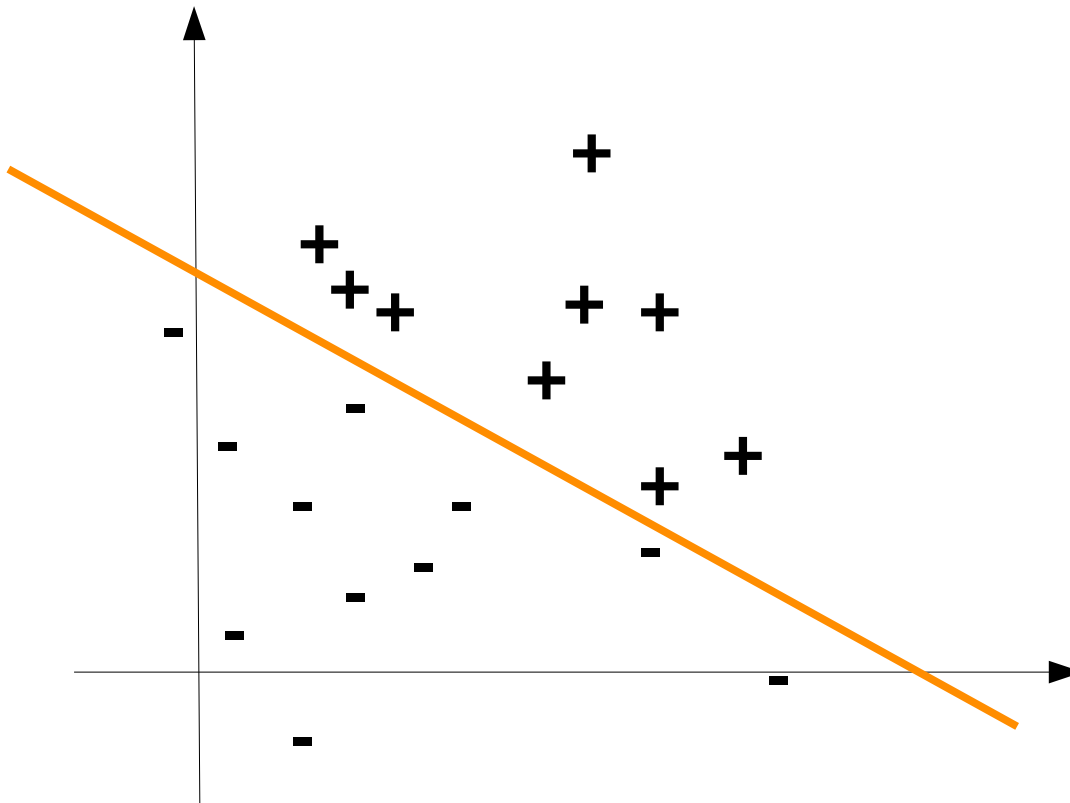
Un perceptron codifica un test lineare

Ciò che cade al di sopra dell'iperpiano codificato dai suoi pesi fa attivare il neurone, ciò che è sotto non fa attivare il neurone

Classificazione:

ciò che è sopra all'iperpiano è riconosciuto come appartenente alla classe obiettivo, sotto all'iperpiano le istanze negative

Cosa codifica un perceptron?



Un perceptron codifica un test lineare

Ciò che cade al di sopra dell'iperpiano codificato dai suoi pesi fa attivare il neurone (restituisce 1), ciò che è sotto non fa attivare il neurone (restituisce 0)

I **pesi** sulle connessioni in ingresso definiscono la posizione e la pendenza dell'iperpiano nello spazio in cui sono definiti gli input. I **pesi** caratterizzano i neuroni e **costituiscono la conoscenza del neurone**


SONO PERSISTENTI

Caratteristiche del perceptron

- Adatto a svolgere compiti di tipo numerico
- Consente di risolvere problemi separabili linearmente
- Conoscenza data dai pesi
- I pesi sono persistenti
- **Apprendimento da esempi, supervisionato**
- **Imparare = individuare la posizione corretta dell'iperpiano nello spazio**

quale segnale può essere usato per guidare l'apprendimento?

Apprendimento

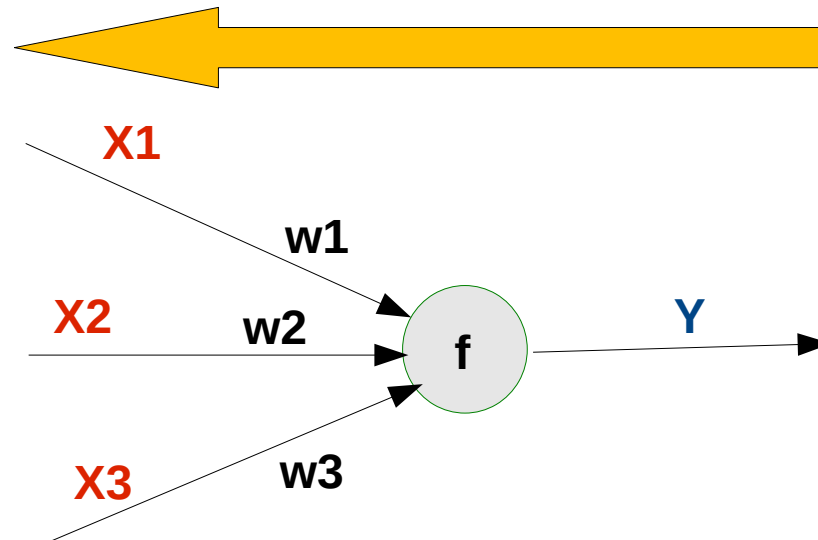
$$w_j^{(k+1)} = w_j^k + \alpha * (d - o) x_j$$


Poiché il problema affrontato è numerico possiamo sfruttare l'errore inteso come differenza fra valore desiderato e ottenuto

Il peso sulla connessione fra l'input j-mo e il neurone viene aggiornato sommandogli l'errore moltiplicato per un fattore di scalamento α e per il valore della componente j-ma dell'input precedentemente elaborato

Novikoff ha dimostrato che quando il problema è linearmente separabile l'apprendimento converge alla soluzione, quando non è linearmente separabile invece non converge

Passata backward



Fatta una passata forward, il perceptron ha prodotto un valore (o = output ottenuto). L'apprendimento è supervisionato, per cui sappiamo qual è l'output che avrebbe dovuto restituire (d = output desiderato)

L'errore ($d-o$) diventa l'informazione che permette di modificare i pesi in modo da rendere, in futuro, il comportamento del perceptron più vicino a quello desiderato

Epoca di apprendimento

Learning set (supervisionato)

l1		c1
l2		c2
l3		c3
...		...

L'apprendimento percorre tutte le istanze di learning: l'istanza viene passata forward nel perceptron, che produce un output

L'errore viene usato in una passata backward che modifica i pesi

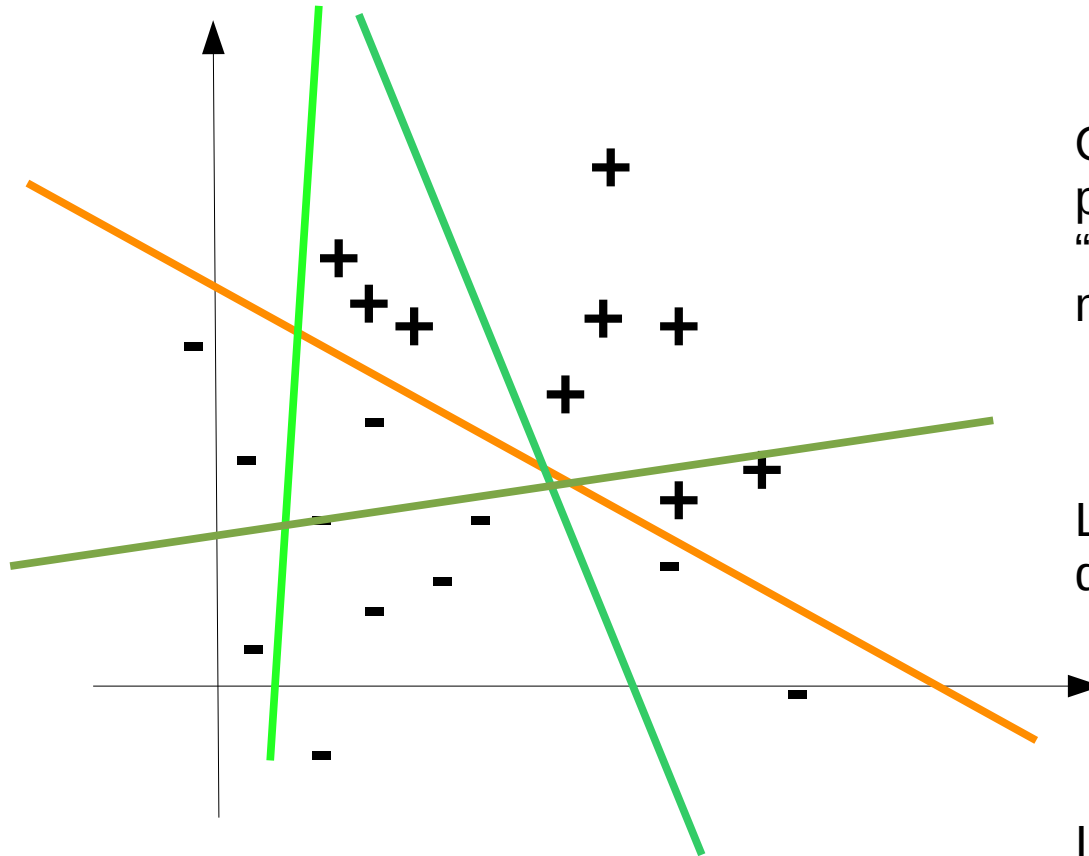
Poi si passa all'istanza successiva

EPOCA DI APPRENDIMENTO:

elaborazione di tutte le istanze contenute nel learning set

L'addestramento può richiedere molte epoche di apprendimento

Convergenza alla soluzione



Ogni esempio "tira" l'iperpiano per fare in modo che il suo "caso" venga trattato correttamente dal perceptron

L'iperpiano percorre lo spazio delle istanze

Il fattore di scalamento evita la rincorsa all'ultimo esempio e forza una generalizzazione

Limiti dei perceptron

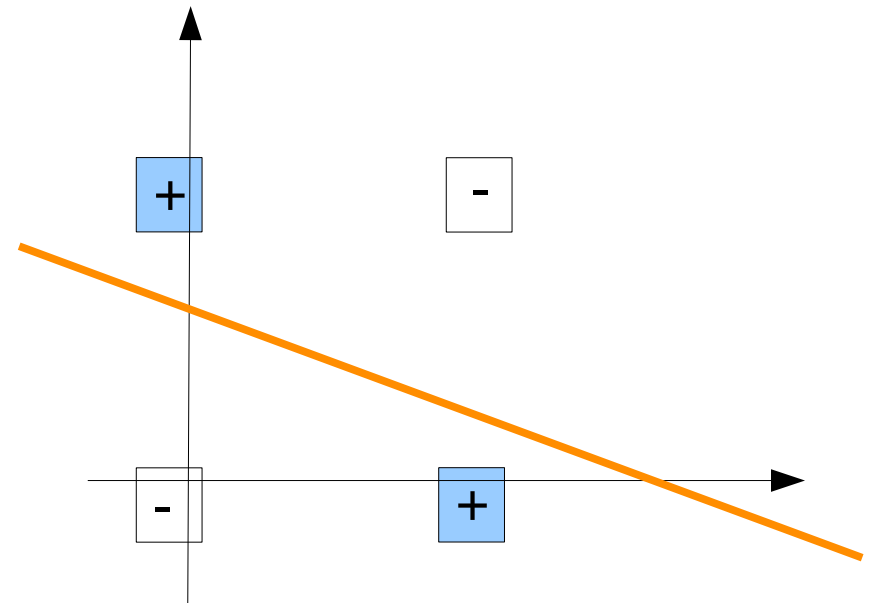
I limiti del perceptron furono evidenziati tramite un esempio semplicissimo: l'or esclusivo (XOR)

A1	A2	XOR
0	0	0
1	0	1
0	1	1
1	1	0



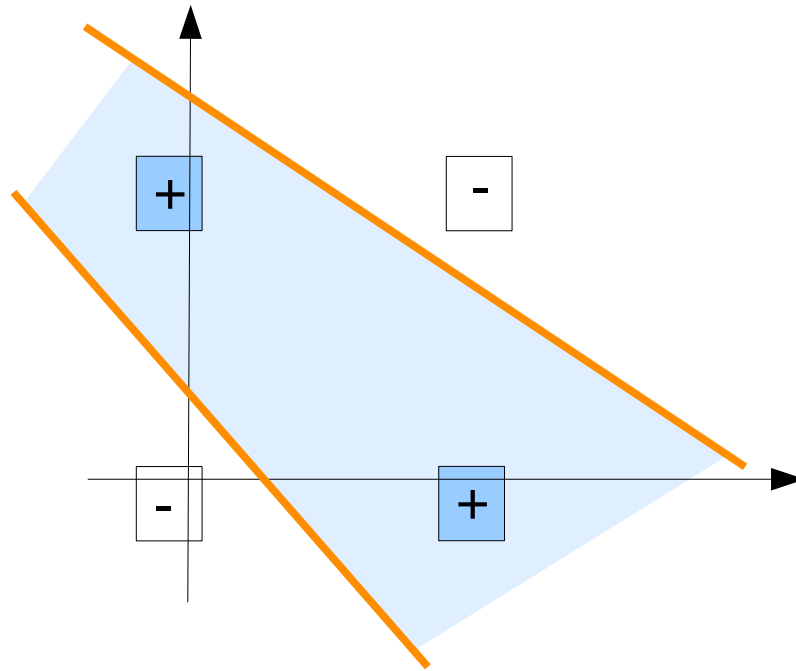
Valori degli
attributi

Output desiderato



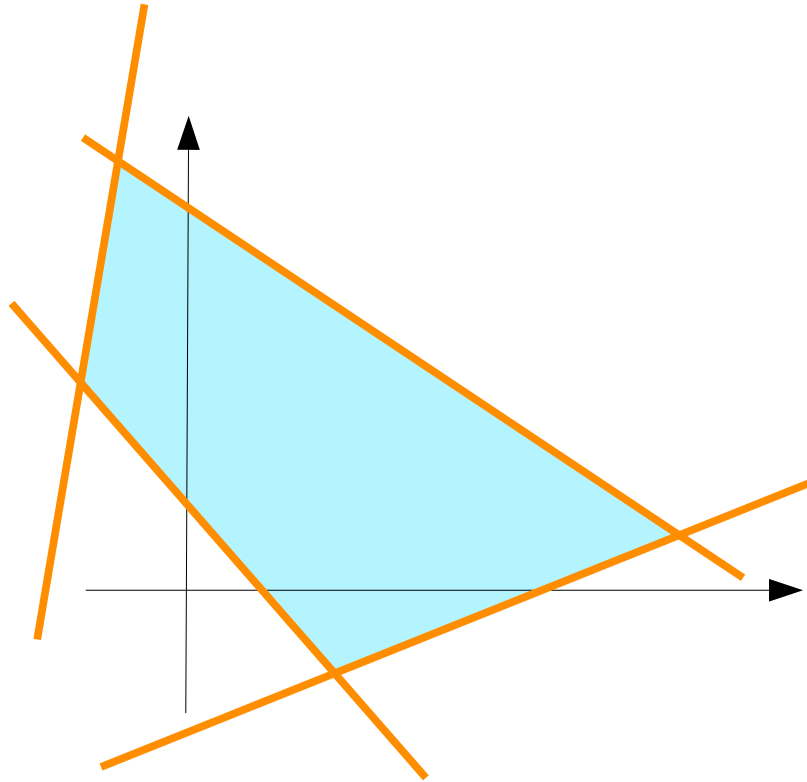
Interpreto le coppie come coordinate di punti e il risultato dello XOR come il fatto che il punto debba stare sopra o sotto all'iperpiano (non c'è verso)

Intuizione ...



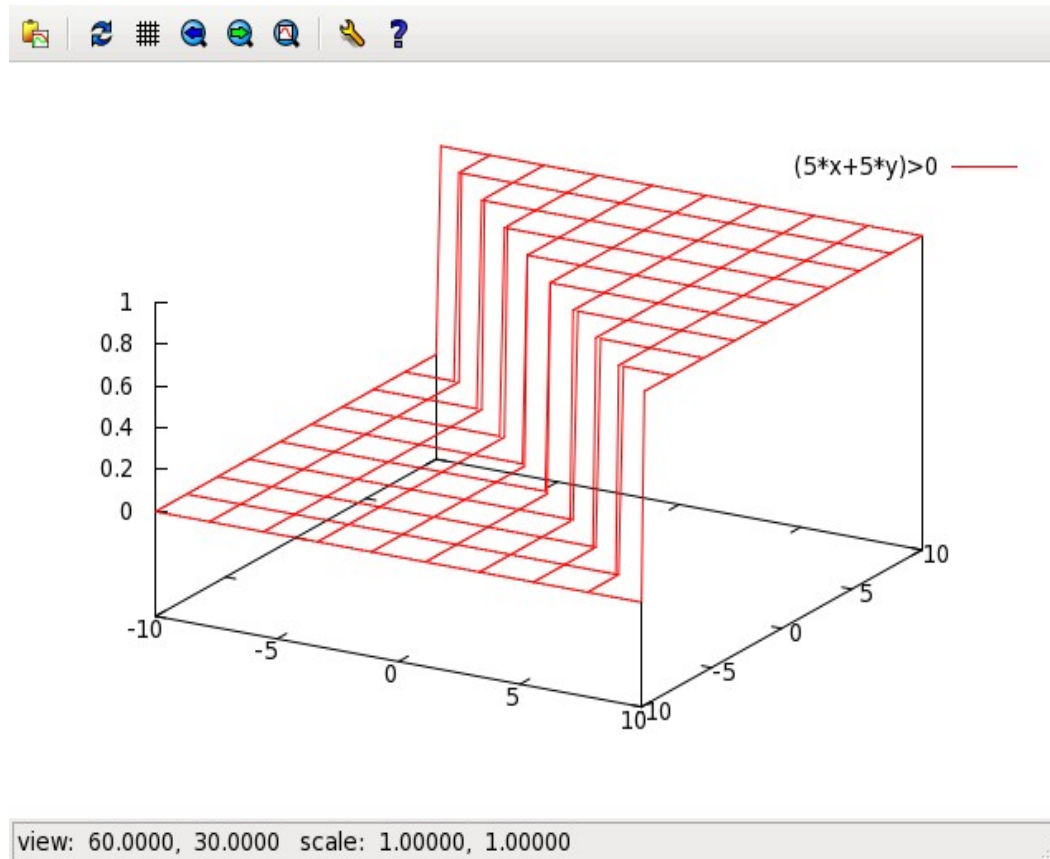
Un'iperpiano da solo non ce la fa ma se ne avessi a disposizione 2 e avessi la capacità di combinare le loro risposte?

Intuizione ...



Se ne avessi a disposizione tanti potrei addirittura circoscrivere delle regioni chiuse!

Come decide il perceptron?



Si parte dalla
combinazione lineare
degli input pesati (net)

La retta tracciata è disegnata
da $w_1 * x_1 + w_2 * x_2 = 0$

$$x_1 = \frac{-w_2}{w_1} * x_2$$

Il perceptron si attiva per i
punti che stanno sopra alla
retta cioè tali che

$$x_1 > \frac{-w_2}{w_1} * x_2$$

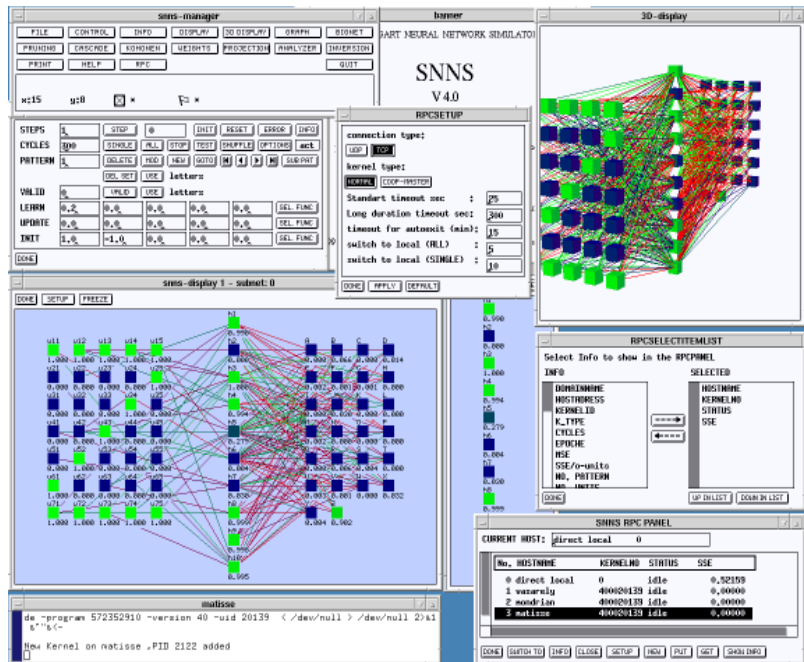
e cioè:

$$w_1 * x_1 + w_2 * x_2 > 0$$

Reti neurali

Una rete neurale è un approssimatore universale di funzioni, avente natura distribuita. È costituita da un insieme di neuroni, collegati fra di loro secondo una topologia, che dipende dal modello di rete realizzato. I neuroni possono implementare funzioni diverse. Possono essere software oppure hardware.

<http://cortex.cs.nuim.ie/tools/spikeNNS/index.html>



SNNS: Stuttgart Neural Network simulator



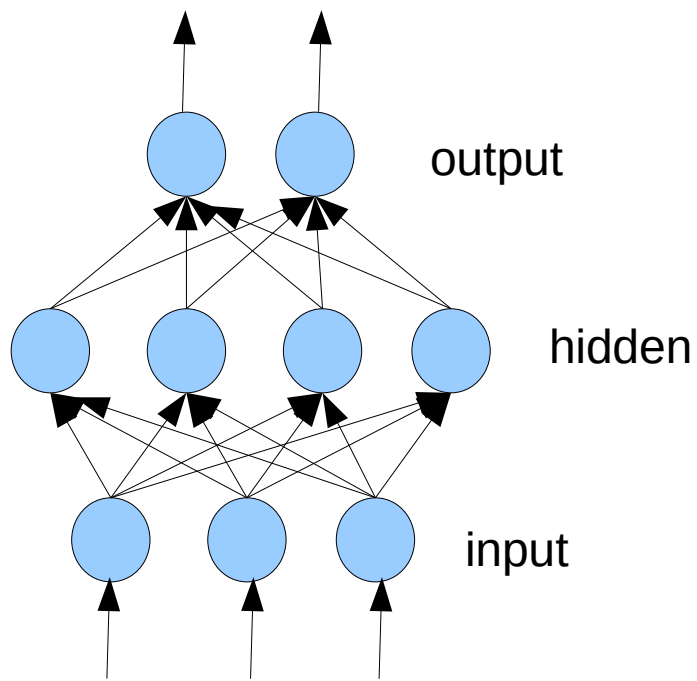
<http://steim.org/projectblog/?p=114>

Topologia di una rete neurale

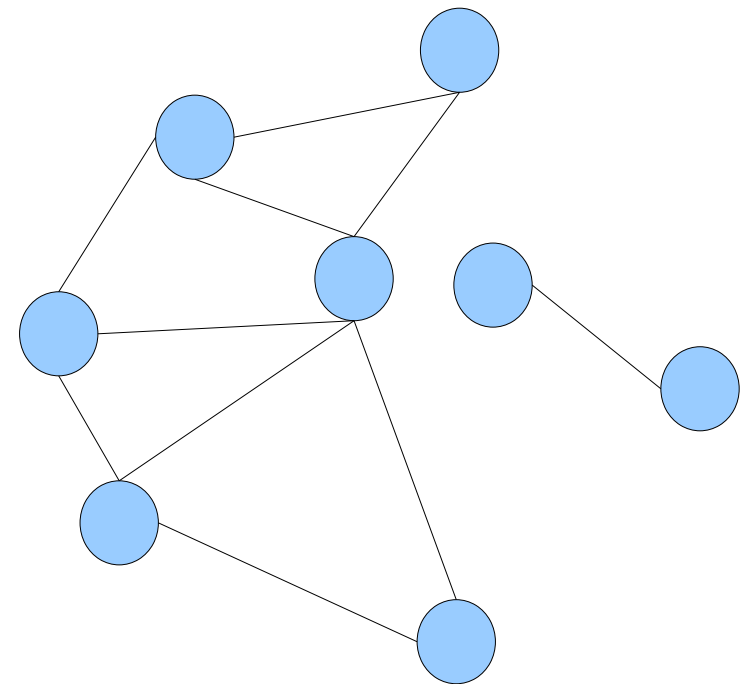
È il modo in cui sono organizzati i neuroni (o nodi) e le loro connessioni

Esempi

A strati (layered)



A vicinato



Multi-layer perceptron

Il multi-layer perceptron (MLP) è un modello di NN con topologia a strati, Feed-forward (flusso di calcolo in una sola direzione). Di solito i neuroni di input implementano la funzione identità. I neuroni hidden sono perceptron, che usano la sigmoide (o altre funzioni tipo a scalino, derivabili), i neuroni di output combinano i risultati dei neuroni hidden. Si possono avere più layer hidden

