



Politecnico  
di Torino

Dipartimento di Scienze  
Matematiche "G. L. Lagrange"

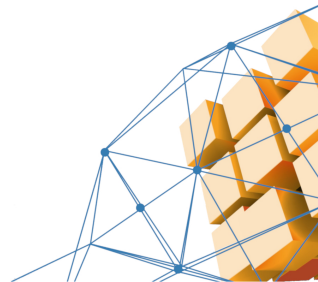


## Reti Neurali

Candidato:  
Riccardo Kiefer

Relatrice:  
Prof.ssa V. Chiadò Piat

06/07/2021  
Laurea in Matematica per l'Ingegneria





# Reti Neurali

## Neuroni artificiali

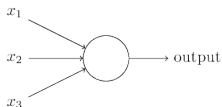
Unità logica della rete neurale

Tipi di neuroni artificiali

- Perceptrons
- Sigmoid Neurons

## Perceptrons

### Modello neurone artificiale



$$output = \begin{cases} 0, & \text{if } w \cdot x + b \leq 0 \\ 1, & \text{if } w \cdot x + b > 0 \end{cases} \quad (1)$$

- input ed output binari.
- $w$  sono i pesi.
- $b$  è il bias.

Modificare pesi e bias per settare risposta desiderata dal neurone.

## Sigmoid Neurons

$$\sigma(z) \doteq \frac{1}{1 + e^{-z}} \quad (2)$$

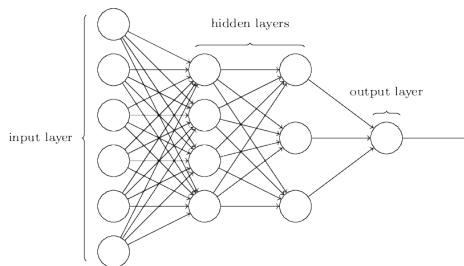
- input ed output reali compreso tra 0 ed 1.
- $\sigma$  regolarizza l'input.
- piccole variazioni di pesi e bias comportano piccole variazioni dell'output.

Possibilità di avere un comportamento più prevedibile modificando pesi e bias.

$$\Delta output \approx \sum_j \frac{\partial output}{\partial w_j} \Delta w_j + \frac{\partial output}{\partial b} \Delta b \quad (3)$$

# Reti Neurali

## Architettura reti neurali



- input layer: pesano informazione in input.
- hidden layer: elaborano l'informazione.
- output layer: emettono l'output.



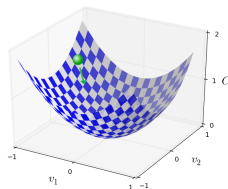
# Reti Neurali

## Apprendimento autonomo

- training and test data.
- metodo discesa del gradiente.
- metodo stocastico discesa del gradiente.
- algoritmo backpropagation.

## Metodo discesa del gradiente

Metodo numerico utilizzato per minimizzare una funzione sufficientemente regolare.



Consiste nel calcolo ripetuto del gradiente per poi utilizzare questo valore per muoversi in direzione opposto nel dominio della funzione.

$$\nu \rightarrow \nu' = \nu - \eta \nabla C \quad (4)$$



## Metodo stocastico discesa del gradiente

Riduce il campione di apprendimento selezionando un piccolo sottoinsieme di elementi scelti casualmente per poi stimare il gradiente tramite la sua media campionaria. Vantaggio:

- 1 Ridurre il costo computazionale e quindi il tempo di calcolo.

## Algoritmo di backpropagation

Metodo numerico per il calcolo rapido del gradiente. Si basa su quattro equazioni che vengono usate per calcolare le componenti del gradiente:

$$\delta^L = \nabla_a C \odot \sigma'(z^L) \quad (5)$$

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l) \quad (6)$$

$$\frac{\partial C}{\partial b_j^l} = \delta_j^l \quad (7)$$

$$\frac{\partial C}{\partial w_{jk}^l} = a_k^{l-1} \delta_j^l \quad (8)$$

$\odot$  rappresenta il prodotto di Hadamard: dati due vettori  $s$  e  $t$ :  $(s \odot t)_j = s_j t_j$ .

- 1 equazione:** Descrive l'errore  $\delta^L$  nello strato di output.
- 2 equazione:** Mette in relazione l'errore  $\delta^l$  con l'errore nello strato successivo  $\delta^{l+1}$ . In particolare quando applichiamo la matrice trasposta a  $\delta^{l+1}$  è come se spostassimo l'errore indietro di uno strato all'interno del network.
- 3 equazione:** Afferma che l'errore  $\delta_j^l$  è la derivata parziale di  $C$  rispetto  $b_j^l$  ( $\partial C / \partial b_j^l$ ).
- 4 equazione:** Fornisce un modo semplice per calcolare la derivata parziale  $C$  rispetto il peso  $w_{jk}^l$ .

Il costo di tale algoritmo è lo stesso di quello richiesto per il passaggio in avanti dell'informazione dallo strato in input a quello di output.

Questi metodi applicati a:

$$C(w, b) \equiv \frac{1}{2n} \sum_x \|y(x) - a\|^2 \quad (9)$$

- $w$  l'insieme dei pesi del network;
- $b$  l'insieme dei bias;
- $n$  il numero totale degli input;
- $a$  il vettore dell'output dato  $x$ ;
- $y(x)$  il valore esatto dell'output dato  $x$ ;

Regolano il processo di apprendimento automatico.

## Un'applicazione: riconoscimento di numeri scritti a mano

504192

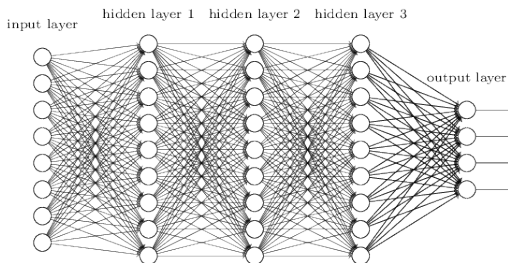
Rete neurale con architettura a tre strati: l'input layer, un hidden layer e l'output layer.

- **Input layer:** Un vettore di dimensione  $28 \times 28 = 784$ .
- **Hidden Layer:** Riconoscano un'immagine vedendo da che parti è composta.
- **Output layer:** composto da 10 neuroni ognuno dei quali indica una cifra da 0 a 9. I neuroni che si attivano indicano il valore finale dell'output della rete.

# Deep Learning

## Network Neurali Profondi

Tipo di network dall'architettura multistrato.

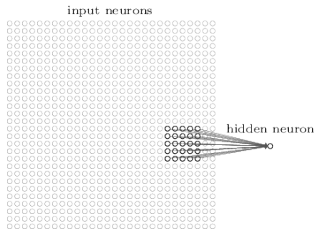




## Fenomeno scomparsa del gradiente

La scomparsa del gradiente è il fenomeno per cui durante la fase di apprendimento le componenti del gradiente della funzione di costo negli strati più interni della rete risultano progressivamente minori nel tempo rispetto a quelle degli strati esterni, questo comporta velocità di apprendimento differenti, compromettendo l'apprendimento omogeneo della rete e rendendo il processo instabile.

## Network convolutivi



- 1 Campi ricettivi locali:** Ogni neurone dell'hidden layer sarà connesso delle piccole aree circoscritte di 5x5 pixel nella griglia di input
- 2 Pesi e bias condivisi:** Tutti i neuroni di uno stesso hidden layer hanno pesi e bias uguali.
- 3 Raggruppamento:** Il raggruppamento (*pooling*) consiste in un ulteriore layer, intermedio tra quello convolutivo e l'output, utilizzato per semplificare le informazioni in output dallo strato precedente.



## Intelligenza Artificiale e Deep Learning

- Algoritmi di riconoscimento vocale utilizzati negli assistenti virtuali di smartphone come Siri o Google Assistant.
- Algoritmi per la guida autonoma di veicoli.
- Algoritmi sofisticati per il riconoscimento di volti come DeepFace, una rete neurale sviluppata da Facebook che ha un grado di accuratezza del 97%.
- Un particolare tipo di rete neurale chiamato GAN in grado di generare immagini realistiche di persone mai esistite.
- Algoritmi utilizzati per i giochi di strategia come DeepBlue della IBM oppure AlphaGo di Google.



Grazie per l'attenzione



Politecnico  
di Torino