

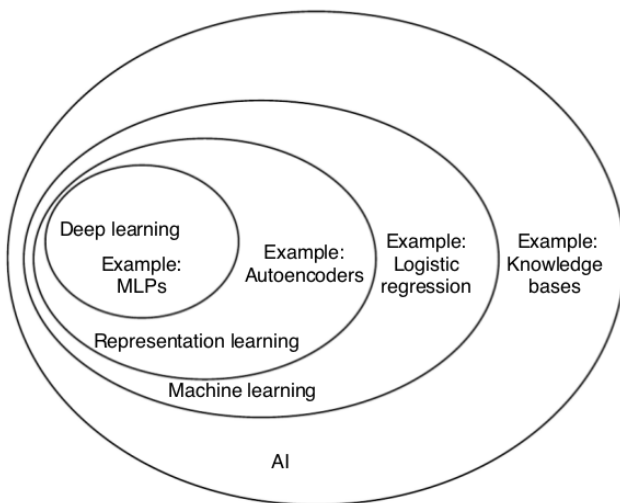
0.1 Dal machine learning al deep learning

Oggi l'intelligenza artificiale è un fiorente campo di ricerca, con l'obiettivo di risolvere una grande varietà di problemi, che per essere risolti attraverso la programmazione classica necessitano di una grande quantità di conoscenze pregresse, spesso non disponibili.

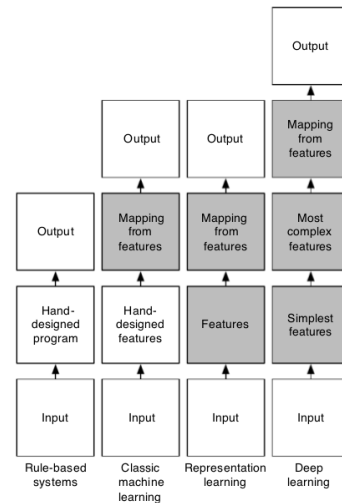
I Programmi basati sul paradigma dell'intelligenza artificiale si propongono di superare questi ostacoli acquisendo direttamente queste conoscenze dai dati grezzi, tale capacità è nota come machine learning. Sotto questa famiglia di algoritmi si trovano altri sottogruppi quali il representation learning e all'interno di quest'ultimo il deep learning.

Il deep learning rispetto ai metodi più classici, tipicamente in grado di riconoscere relazioni lineari (come ad esempio il noto SVM o support vector machine), si propone come alternativa per l'apprendimento di funzioni a molte variabili non lineari anche molto complesse. Le applicazioni più comuni comprendono il riconoscimento di oggetti nelle immagini, delle parole in tracce audio o anche per le traduzioni multilingue.

Il termine "deep learning" deriva proprio dalla capacità di questi modelli di riuscire a cogliere delle relazioni molto "profonde" tra i dati di ingresso e di uscita, approssimando con un certo errore, dipendente dal caso specifico, la funzione che dati gli input restituisce l'output desiderato, purtroppo però presentano il grande problema di non rendere disponibile in maniera chiara le relazioni apprese.



famiglie di algoritmi



differenze generali tra
le diverse tipologie di algoritmi

0.1.1 Le reti feed-forward (MLP)

Uno dei modelli della famiglia del deep learning più comuni e semplici da utilizzare sono le reti neurali feed-forward o multi layer perceptron, le quali sono assimilabili a modelli di regressione statistica.

Le reti neurali sono un modello matematico molto semplificato del cervello, e per parlarne è necessario qualche riferimento al modello biologico.

Il neurone biologico

L'unità base del cervello è rappresentata dai neuroni (Figura 2), i quali a loro volta sono composti dai dendriti, dal soma, dall'assone e dalle sinapsi.

I dendriti sono l'apparato di input del neurone, attraverso il quale riceve i segnali dall'ambiente o da altri neuroni, tali segnali caricano il neurone, che accumula un potenziale elettrico all'interno del soma. In determinate condizioni o al raggiungimento di una certa soglia di potenziale il neurone emette un impulso lungo l'assone giungendo alle sinapsi, poste alla fine dell'assone, le quali sono connesse con altri neuroni o con altri apparati del corpo, come i muscoli.

Questa particolare cellula oltre a scambiare impulsi è in grado di accumulare informazioni attraverso l'ispessimento o l'assottigliamento della guaina mielinica, la quale è presente lungo l'assone. Le variazioni dello spessore della guaina comportano un aumento o una diminuzione della resistenza elettrica dell'assone determinando una variazione dell'impulso percepito dai neuroni connessi alle sinapsi a parità di potenziale rilasciato.

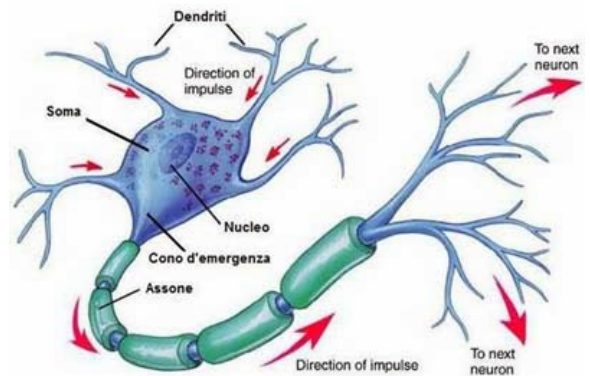


Figura 2: Il neurone biologico

Il perceptrone

Il neurone artificiale è composto da una serie di connessioni in ingresso, come i dendriti nel caso del neurone biologico, questa volta però il compito di immagazzinare informazione viene svolto da queste connessioni in input e non più dalle connessioni in output. Ad ogni input del neurone è associato un coefficiente che moltiplica il potenziale in ingresso, a questo punto gli ingressi pesati giungono al nodo sommatore, il corrispettivo del soma, che ne accumula la sommatoria. Lo step successivo è il blocco contenente la funzione di trasferimento, una funzione che prende in input il potenziale del neurone e rende disponibile il risultato come output finale o come input per i neuroni successivi. La funzione di trasferimento del neurone può essere di vari tipi, può variare infatti anche tra i neuroni della medesima rete, alcune delle più comuni sono:

- (a) la funzione gradino.
- (b) la funzione lineare con saturazione
- (c) la funzione sigmoide tra 0 e 1
- (d) la funzione sigmoide tra -1 e 1

Volendo esprimere il perceptrone in termini matematici si ottiene la seguente equazione:

$$y = f(P) = f(\sum_i W_i \cdot X_i)$$

y - output del perceptrone

f - funzione di trasferimento del perceptrone

P - potenziale del neurone

W_i - peso o coefficiente dell' i -esimo ingresso

X_i - potenziale d'ingresso dell' i -esimo neurone

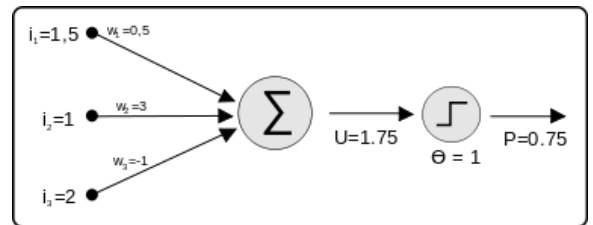


Figura 3: Il perceptrone

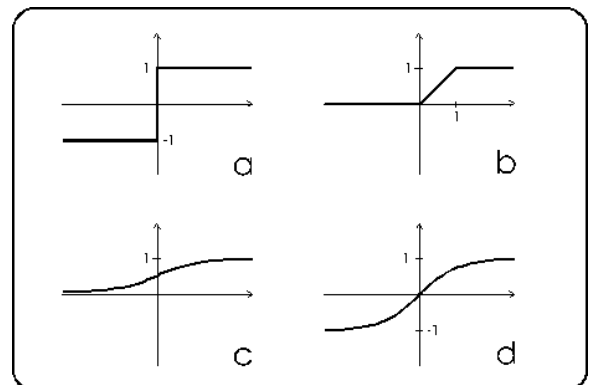
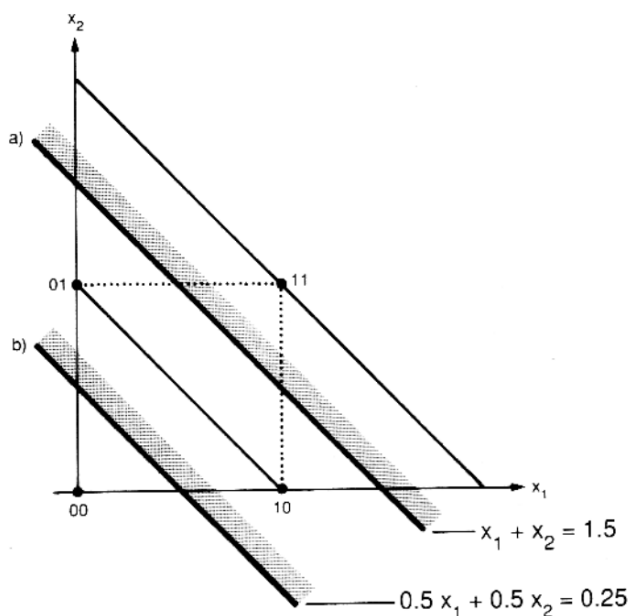
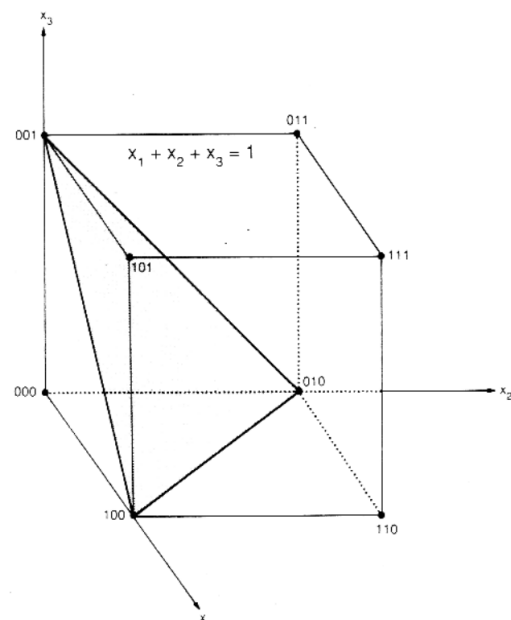


Figura 4: Le funzioni di trasferimento

Una delle caratteristiche più importanti del perceptrone è la sua capacità di compiere separazioni lineari. Tali separazioni avvengono in uno spazio a n dimensioni, dove n corrisponde al numero di variabili in ingresso al neurone, perciò si avrà ad esempio per un perceptrone a due input la capacità di separare i valori su \mathbb{R}^2 attraverso una retta, nel caso \mathbb{R}^3 si potranno separare i valori dello spazio tridimensionale attraverso un piano, e così via.



esempio di linee di separazione per neurone a 2 ingressi binari
(a - funzione OR)
(b - funzione AND)



esempio di piano di separazione lineare
a 3 ingressi per un neurone binario

Il problema dello XOR

Le reti multi-strato e la separazione non lineare

Le reti multi strato sono composte da strati di percettroni, interconnessi tra loro in modo che gli output degli strati precedenti siano connessi con gli input dei percettroni successivi.

0.1.2 La funzione di addestramento: Il back-propagation