

Analisi delle Reti Neurali (cenni)

- L'IA funziona combinando grandi quantità di dati con un'elaborazione veloce e iterativa e algoritmi intelligenti, consentendo al software di imparare automaticamente dai modelli o dalle caratteristiche dei dati.
- L'IA è un campo di studio molto ampio che comprende diverse teorie, metodi e tecnologie.

- La **rete neurale** artificiale è una analisi di tipo machine learning costituito da unità interconnesse (come i neuroni) che elabora le informazioni rispondendo agli input esterni e trasmettendo le informazioni tra ogni unità. Il processo richiede ai dati più passaggi per trovare le connessioni e ricavare significato dai dati non definiti.

- Il **Deep Learning** utilizza enormi reti neurali con molti livelli di unità di elaborazione. Sfrutta i progressi nella potenza di calcolo e le migliorate tecniche di apprendimento per imparare i modelli complessi presenti nella grandi quantità di dati. Le applicazioni più comuni includono il riconoscimento di immagini e voce.

- Il Cognitive Computing è una branca dell'intelligenza artificiale che vuole ottenere un'interazione naturale con le macchine, simile a quella umana. Utilizzando l'intelligenza artificiale e il calcolo cognitivo, l'obiettivo finale è una macchina che simuli i processi umani attraverso la capacità di interpretare immagini e parlato e che sia poi in grado di rispondere in modo coerente. La **computer vision** si basa sul riconoscimento dei modelli e sul deep learning per riconoscere ciò che c'è in un'immagine o in un video. Quando le macchine sono in grado di elaborare, analizzare e comprendere il contenuto, possono catturare immagini o video in tempo reale e interpretare l'ambiente circostante.

- Il Natural Language Processing (elaborazione del linguaggio naturale) è la capacità dei computer di analizzare, comprendere e generare il linguaggio umano, compreso il parlato. La fase successiva dell'NLP è l'interazione linguistica naturale, che consente agli esseri umani di comunicare con i computer utilizzando un linguaggio normale e quotidiano per svolgere le loro attività.

- Le unità di elaborazione grafica (GPU) sono fondamentali per l'intelligenza artificiale perché forniscono la potenza di calcolo necessaria per l'elaborazione iterativa. L'apprendimento delle reti neurali richiede grandi dati e potenza di calcolo elevata.
- L'Internet of Things genera enormi quantità di dati dai dispositivi connessi, la maggior parte dei quali non analizzati. L'automazione dei modelli con l'IA ci permetterà di utilizzarne di più.

- Le [API](#), le interfacce di programmazione di un'applicazione, sono pacchetti di codice che consentono di aggiungere funzionalità di IA a prodotti e pacchetti software esistenti. Possono aggiungere funzionalità di riconoscimento delle immagini ai sistemi di sicurezza domestica e funzionalità di Q&A che descrivono dati, creano didascalie e titoli o richiamano modelli e informazioni sui dati.

Il cervello: un sistema di elaborazione delle informazioni

- Le reti neurali artificiali cercano di simulare all'interno di un sistema informatico il funzionamento di sistemi nervosi biologici del cervello.
- Il cervello umano contiene circa 10 miliardi di cellule nervose dette neuroni. In media, ogni neurone è collegato ad altri neuroni attraverso circa 10000 sinapsi (i dati effettivi variano notevolmente a seconda della neuroanatomia locale).
- La rete cerebrale di neuroni forma un massiccio sistema
- parallelo di elaborazione delle informazioni.

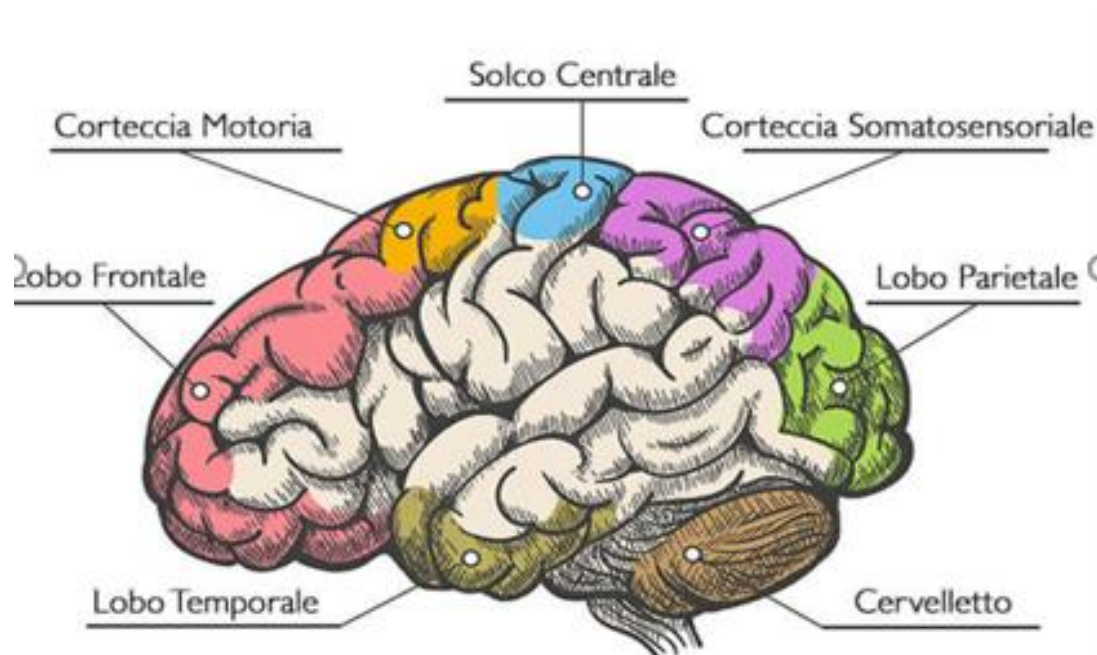
Nota: (Gallo C., De Bonis M .)

Il cervello: un sistema di elaborazione delle informazioni

- Il cervello quantunque sia costituito da un hardware molto lento, rispetto ai processori artificiali, ha notevoli capacità:
- le sue prestazioni tendono a peggiorare “dolcemente” in caso di danni parziali. Al contrario, la maggior parte dei programmi e dei sistemi artificiali sono molto fragili: se si rimuovono alcune parti arbitrarie,
- può imparare (riorganizzarsi) attraverso l’esperienza;
- il recupero parziale dei danni `e possibile se le unità sane possono imparare ad assumere le funzioni precedentemente svolte dalle aree danneggiate;
- esegue i calcoli in modo altamente parallelo ed estremamente efficiente.
- Nota: (Gallo C., De Bonis M .)

Le reti neurali nel cervello

- Il cervello non è una entità omogenea: si distinguono corteccia, mesencefalo, tronco cerebrale e cervelletto. Ognuna di queste parti può essere gerarchicamente suddivisa in molte altre regioni e all'interno di ciascuna regione in aree, sia in base alla struttura anatomica delle reti neurali sia in base alla funzione svolta. Nota: (Gallo C., De Bonis M.)



Le reti neurali nel cervello

Il modello generale di proiezioni (fasci di connessioni neurali) tra le aree è complesso. Si distinguono proiezioni feedforward che vanno dalla prima fase di trasformazione (stimoli sensoriali) all'ultima trasformazione (visualizzazione dell'immagine), e proiezioni di tipo feedback, cioè connessioni che vanno nella direzione opposta. Oltre a questi collegamenti ad ampio raggio, i neuroni possono mettersi in contatto con molte migliaia di neuroni vicini. In questo modo si forma una serie di reti locali dense e complesse

(Gallo C., De Bonis M.)

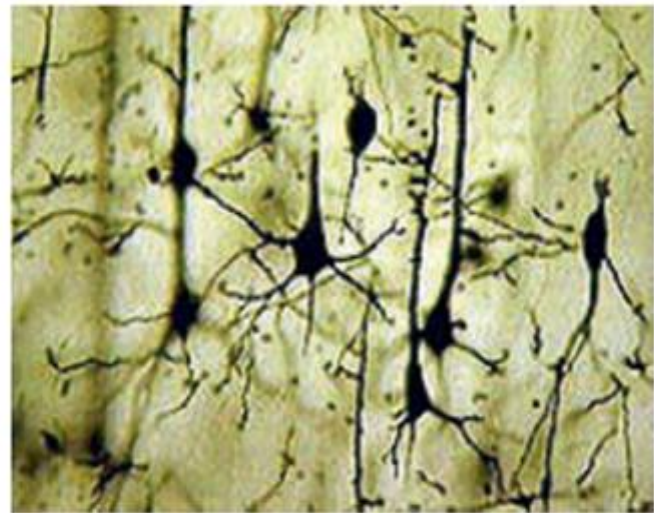


Figura 1. *Rete neurale biologica*

Le reti neurali nel cervello

Neuroni e sinapsi

L'unità di base di calcolo nel sistema nervoso è la cellula nervosa, o neurone (Figura 2 e Figura 3). Un neurone è costituito dai seguenti elementi: Dendriti (ingressi), Soma (corpo della cellula), Assone (uscite)

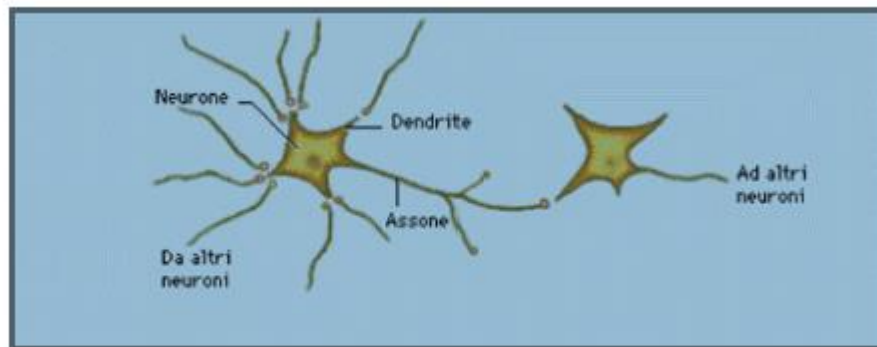


Figura 3. *Struttura di un neurone biologico*

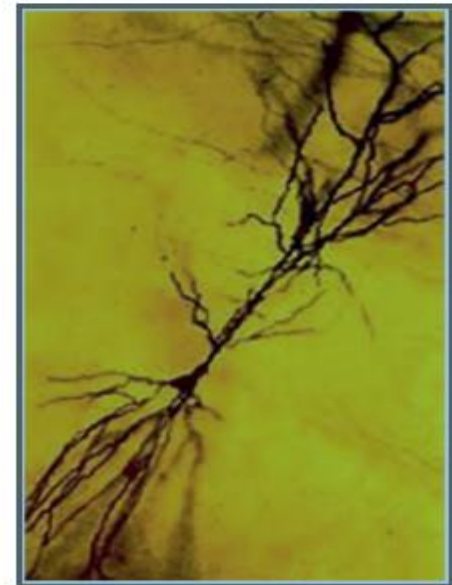


Figura 2. *Singolo neurone biologico*

Nota: (Gallo C., De Bonis M.)

- Un neurone riceve l'input da altri neuroni (molte migliaia in genere). Gli ingressi tendono a sommarsi, a dare un contributo cumulativo al neurone destinatario.
- Una volta che l'ingresso supera un valore di soglia detto livello critico, si genera un impulso detto spike: un impulso elettrico che viaggia dal corpo, giù per l'assone, verso il neurone successivo o a tutti gli altri recettori a cui è connesso. Questo evento è anche chiamato depolarizzazione, ed è seguito da un periodo refrattario, durante il quale il neurone non è in grado di essere recettivo.

Nota: (Gallo C., De Bonis M.)

- Le terminazioni degli assoni sono quasi a contatto con i dendriti o con il corpo cellulare del neurone successivo.
- La trasmissione di un segnale elettrico da un neurone all'altro avviene tramite i neurotrasmettitori, una serie di informazioni di tipo chimico che vengono rilasciati dal neurone trasmettitore e si legano ai recettori nel neurone destinatario. Questo collegamento `e chiamato **sinapsi**.
- La misura in cui il segnale da un neurone passa a quello successivo dipende da molti fattori.

Nota: (Gallo C., De Bonis M .)

Apprendimento sinaptico

Il cervello impara, naturalmente. Da quello che conosciamo delle strutture neuronali, un modo che il cervello ha per imparare `e di modificare i punti di forza delle connessioni tra neuroni: aggiungere o eliminare le connessioni tra i neuroni stessi. Inoltre si impara sulla base dell'esperienza.

L'efficacia di una sinapsi pu` variare in funzione del risultato di una determinata esperienza, fornendo sia la memoria sia l'apprendimento attraverso il potenziamento a lungo termine (LTP - Long Term Potentiation). Un modo in cui ci` avviene `e attraverso il rilascio di altri neurotrasmettitori.

- Il potenziamento a lungo termine `e un aumento duraturo dell'efficacia sinaptica (> 1 ora) come il risultato di una alta frequenza di stimolazione di vari ingressi attraverso quello che `e comunemente denominato il Postulato di Hebb:
- “Quando un assone della cellula A eccita la cellula B ripetutamente o prende parte in modo persistente a questo processo di stimolazione, un processo di crescita, o cambiamento metabolico, avviene in una o entrambe le cellule in modo che l'efficienza di A come cellula di trasmissione di B `e aumentata.”

Alcuni ricercatori scoprirono che nel cervello il processo denominato LTP (LTP - Long Term Potentiato) denota le seguenti conseguenze:

- le sinapsi diventano più o meno importanti nel tempo (fenomeno della plasticità)
- LTP è basato sull'esperienza
- LTP è basato solo su fenomeni locali di scambio delle informazioni (Postulato di Hebb)

Il cervello: un sistema di elaborazione delle informazioni

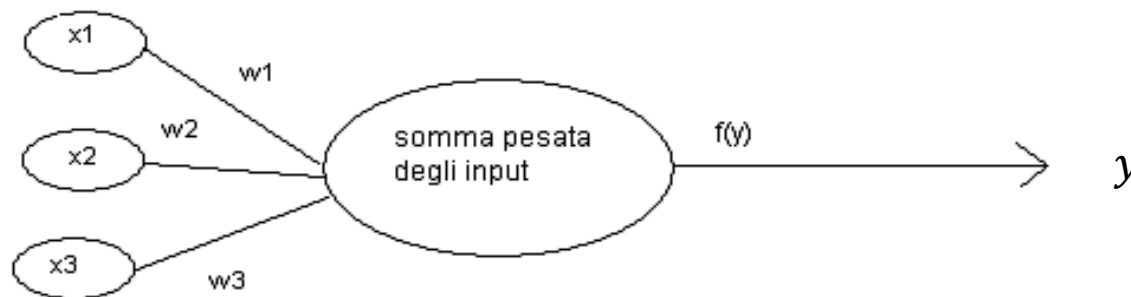
- Come branca dell'Intelligenza Artificiale, le reti neurali artificiali sono, quindi, il tentativo di portare i computer un po' più vicino alle capacità del cervello imitandone alcuni aspetti di elaborazione delle informazioni, in un modo altamente semplificato.

Le Reti Neurali

- Le reti neurali artificiali costituiscono un settore di ricerca scientifico e tecnologico in rapida espansione: sono propriamente dei sistemi di elaborazione dell'informazione descritti nel linguaggio matematico e che traggono ispirazione dai sistemi nervosi biologici.
- Il neurone artificiale rappresenta quello biologico; in questo caso gli stati di attivazione e le relazioni risultano da formule matematiche e non da fenomeni biologici.

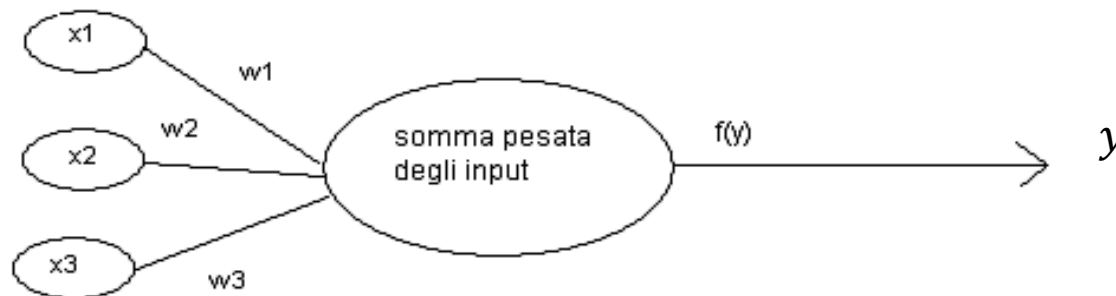
Un neurone artificiale è caratterizzato:

- dagli x gli *input layer* segnali d'entrata,
- dai pesi sinaptici w , che sono i punti di connessione tra un neurone e l'altro (positiva/negativa, debole/forte)
- da una soglia di attivazione A
- da una funzione di attivazione (trasferimento) $f(y)$



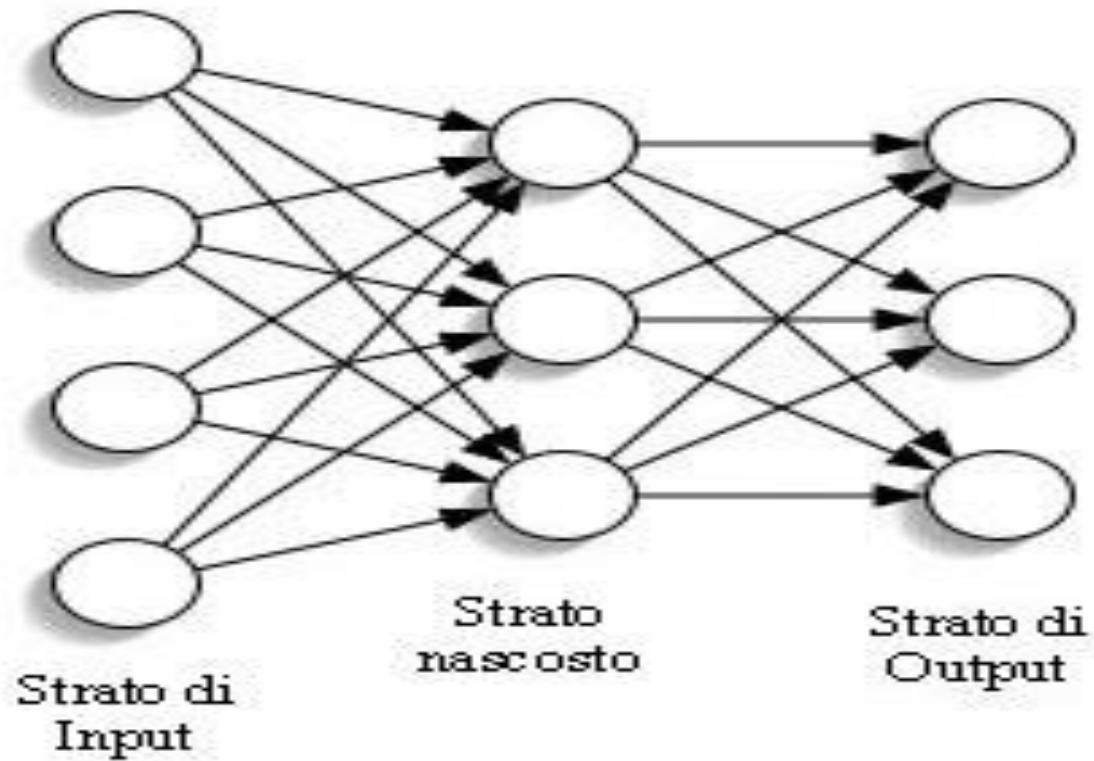
Un neurone artificiale è caratterizzato:

- dagli x gli *input layer*
- dai pesi sinaptici w , che sono i punti di connessione tra un neurone e l'altro,
- da una soglia di attivazione A ,
- da una funzione di attivazione $f(y)$
- y risposta del neurone



Struttura della rete neurale

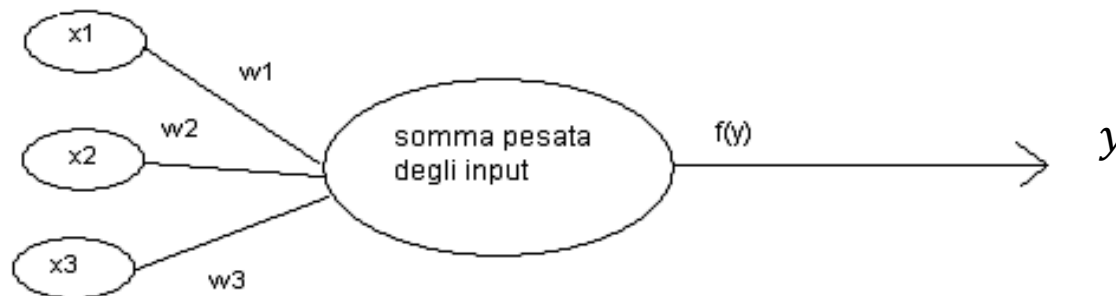
- La rete neurale artificiale è suddivisa in strati e ciascuno strato è costituito da unità elementari connesse tra loro, chiamate neuroni o nodi.
- Le tipologie di strati della rete neurale sono tre:
- Strato di input detto *input layer*: costituito da neuroni che ricevono segnali dall'ambiente esterno;
- Strato/i nascosto/i chiamato *hidden layer*: costituito da neuroni che comunicano solamente con le unità all'interno della rete;
- Strato di output detto *output layer*: costituito da neuroni che emettono informazioni all'ambiente esterno.



Schema generale della rete neurale; «Reti neurali», <http://www.ildiogene.it/EncyPages/Ency=retineurali.html>.

Un neurone artificiale è caratterizzato:

- dagli x gli *input layer*
- dai pesi sinaptici w , che sono i punti di connessione tra un neurone e l'altro,
- da una soglia di attivazione A ,
- da una funzione di attivazione $f(y)$
- y risposta del neurone



Quando un neurone riceve un segnale (dall'ambiente esterno se si tratta di una unità di input o da un altro neurone altrimenti), la cui intensità supera una certa soglia, esso si attiva ed emette a sua volta un impulso attraverso i canali di comunicazione fino alle altre unità con le quali è connesso.

I punti di connessione tra un neurone ed un altro hanno il compito di filtrare i messaggi e di trasformarli in segnali inibitori o eccitatori modificandone l'intensità a seconda delle caratteristiche individuali.

- I punti di connessione vengono chiamati *pesi sinaptici* o *sinapsi* w perché ricordano le sinapsi biologiche e hanno il compito di pesare l'intensità dei segnali trasmessi.
- L'input A_i di un neurone i -esimo è dato da:

$$A_i = \sum_{j=1}^N w_{ij} x_j - \vartheta_i \quad i=1, \dots, M \quad j=1, \dots, N$$

dove x_j è il segnale d'ingresso, w_{ij} sono i valori delle sinapsi corrispondenti al neurone j -esimo e a quello i -esimo e ϑ_i il valore soglia del neurone i -esimo.

- La risposta del neurone viene calcolata trasformando l'input attraverso la funzione di attivazione f
- $y_i = f(A_i) = f(\sum_{j=1}^N w_{ij}x_j - \vartheta_i)$
- Le funzioni di attivazione possono essere di varia tipologia e determinano il tipo di risposta che un neurone è in grado di emettere.

La Funzione Di Attivazione

- La funzione di attivazione può essere una funzione non lineare o lineare, la scelta di una soluzione, piuttosto che un'altra, dipende dal tipo di problema analizzato.
- Il suo compito è di convertire i segnali in input nel valore finale di output della rete, accettando come argomento la funzione somma (valore di attivazione).
- Solitamente vengono scelte funzioni semplici cioè che siano continue e derivabili su tutto il campo di definizione.
- Non esiste una regola generale per la scelta della funzione di attivazione e neanche per la scelta se utilizzare una funzione di attivazione in ogni strato della rete.

Esempi di funzioni di attivazione

Logistica	$f(x) = \frac{1}{1 + e^{-x}}$
Logistica simmetrica	$f(x) = \left(\frac{2}{1 + e^{-x}} \right)^{-1}$
Tangente iperbolica	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$
Gaussiana	$f(x) = e^{-x^2}$
Gaussiana inversa	$f(x) = 1 - e^{-x^2}$
Sinusoide	$f(x) = \sin(x)$
Tangente corretta	$f(x) = \tan(c \cdot x)$
Lineare	$f(x) = x$

Figura1-Le principali funzioni di attivazione

Caratteristiche delle reti neurali

- **Robustezza:** le reti neurali sono costituite da numerosi neuroni e sono in grado di svolgere la propria attività anche se una o più unità smette di funzionare. Inoltre, in alcuni casi, le unità lesionate possono essere riaddestrate affinché riacquistino la loro abilità.
- **Flessibilità:** una rete neurale non viene programmata, ma ha la capacità di apprendere dall'ambiente esterno; perciò può essere adoperata in diversi campi applicativi senza che vi sia una conoscenza analitica del problema analizzato.

- **Generalizzazione:** una volta che viene conclusa la fase di apprendimento, la rete neurale è in grado di fornire risposte su nuove sequenze di dati. Questa caratteristica è dovuta al fatto che il modello neurale non memorizza le caratteristiche del singolo, ma estrae le peculiarità invarianti dell'input.
- **Recupero in base al contenuto:** le reti neurali sono capaci di recuperare informazioni memorizzate anche partendo da dati parziali o corrotti.

Campi applicativi

- **Informatica:** le reti neurali artificiali sono utilizzate nella scienza dell'informazione per:
 - Compressione di dati: sono stati sviluppati codici utilizzati per la compressione e decompressione di immagini.
 - Eliminazione del rumore: sono state utilizzate reti neurali capaci di riconoscere ed eliminare disturbi di segnale dai canali di trasmissione.
 - Riconoscimento di segnali sonar: i modelli neurali hanno trovato grande applicazione nel riconoscimento e classificazione di oggetti marini perché in grado di separare il segnale di interesse da disturbi provocati da altri oggetti.

Campi applicativi

- **Informatica:** le reti neurali artificiali sono utilizzate nella scienza dell'informazione per:
 - Macchina da scrivere fonetica: è stata utilizzata una rete neurale che impara a segmentare il discorso di un parlante e ad associare in tempo reale ciascun segmento alle lettere dell'alfabeto.
 - Riconoscimento di caratteri: attraverso l'uso di reti neurali è stato possibile riconoscere caratteri scritti a mano con bassissimo indice di errore.

Campi applicativi

- **Sistemi di controllo:** le reti neurali artificiali hanno trovato applicazione nel campo dei sistemi di controllo per robot e veicoli intelligenti.
- **Analisi finanziarie:** le caratteristiche del modello neurale ha permesso l'applicazione di questo meccanismo per compiere predizioni finanziarie, valutare rischi di alcuni investimenti e analizzare il mercato borsistico

Campi applicativi

- **Medicina:** una rete neurale può anche essere d'aiuto per la formulazione di diagnosi e prognosi di patologie, in quanto può riconoscere immagini e può operare su dati simbolici che rappresentano una descrizione dei sintomi o dei parametri del paziente.
- **Neuroscienza:** la neuroscienza computazionale è la disciplina che utilizza le reti neurali artificiali per poter comprendere ulteriormente il funzionamento del sistema nervoso.

Campi applicativi

- **Psicologia:** lo studio delle reti neurali in psicologia è chiamato Connessionismo; i modelli neurali artificiali, affiancati ai modelli seriali tradizionali, permettono di dare una spiegazione a fenomeni macrostrutturali riguardanti i processi cognitivi.

Architetture

Definire l'architettura di una rete neurale significa:

- Distinguere tra neuroni di ingresso e neuroni di uscita:
- le reti *etero-associative* hanno nodi di ingresso che sono distinti dai nodi di uscita;
- quelle *auto-associative* possiedono un unico strato di unità connesse tra loro che ricevono sia input esterni sia input provenienti dagli altri nodi. La funzione principale delle reti auto-associative consiste nell'apprendere, immagazzinare e ricostruire modelli.

Architetture

Definire l'architettura di una rete neurale significa:

- Scegliere il numero di strati di sinapsi (o di neuroni): esistono reti con un unico strato di neuroni e invece altre che possiedono uno o più di uno strato intermedio o nascosto (es. *Multy-Layer Perceptron*).

Architetture

Definire l'architettura di una rete neurale significa considerare connessioni di feedback:

- le reti *feedforward* sono architetture in cui ciascun nodo riceve segnali unicamente dagli strati inferiori e quindi il flusso di informazione procede in un'unica direzione. Queste reti non riescono a cogliere eventuali strutture temporali dell'informazione e per ovviare a questo problema è possibile inserire le cosiddette *connessioni feedback* o *connessioni ricorrenti*. Una connessione di questo tipo permette la ricezione di segnali provenienti da altri nodi dello stesso strato o da nodi degli strati superiori.

Apprendimento

- La rete neurale risponde all'ingresso di input modificando i pesi delle connessioni sinaptiche in base ad un algoritmo di apprendimento, quindi non è programmata, ma si addestra attraverso tali procedure che possono essere di due tipi:
 - apprendimento supervisionato
 - apprendimento non supervisionato.

Apprendimento

- supervisionato (supervised learning): in questo caso il training set contiene sia i dati di ingresso che i dati di uscita reali;
- non supervisionato (unsupervised learning): i valori del modello della rete neurale artificiale vengono modificati solo in funzione dei dati di ingresso.

Apprendimento

Apprendimento supervisionato:

consiste nel modificare i pesi sinaptici in seguito ad un confronto tra la risposta fornita dalla rete e la risposta desiderata, cioè la variazione dei pesi è una funzione opportuna dell'errore commesso dalla rete.

L'addestramento viene eseguito su coppie del tipo (X, D) dove $X = (x_1, x_2, \dots, x_n)$ è il vettore degli input e $D = (d_1, d_2, \dots, d_n)$ è il vettore degli output attesi; l'insieme delle coppie coinvolte viene chiamato *training set*.

Apprendimento

Apprendimento supervisionato:

- Una tipologia particolare di apprendimento supervisionato è l'*apprendimento per rinforzo* secondo il quale l'ambiente esterno risponde all'output della rete con un segnale positivo o negativo che rappresenta un indice di bontà dell'operazione.

Apprendimento

Apprendimento non supervisionato:

è un processo durante il quale la rete si auto-organizza durante il suo funzionamento in base ad alcune semplici regole. In questo modo la rete è in grado di estrarre informazioni dall'esterno e di classificare autonomamente gli input.

Osservazioni

- Nonostante la sostanziale differenza tra le due tipologie di apprendimento, non è facile delineare confini precisi tra le due procedure. In particolare esiste un gruppo di algoritmi di apprendimento, gli algoritmi genetici, che rientra in entrambi i processi. Esistono, tuttavia, delle caratteristiche generali comuni ad entrambe le procedure:

Osservazioni

- I valori iniziali dei pesi sinaptici della rete vengono assegnati in modo casuale oppure viene assegnato a tutti i pesi il valore nullo;
- L'apprendimento consiste nell'inserimento ripetuto di campioni appartenenti al training set all'interno della rete (nell'addestramento supervisionato viene considerata la coppia (X,D) , invece nell'addestramento non supervisionato si utilizza il solo vettore degli input X).

Osservazioni

- Data la variabile tempo t , la nuova configurazione dei pesi sinaptici viene calcolata nel modo seguente:

$$w_{ij}^t = w_{ij}^{t-1} + \eta \Delta w_{ij}^t$$

- L'apprendimento si occupa di integrare le nuove conoscenze a quelle precedenti già consolidate e tale integrazione è graduale e avviene in modo ricorsivo (ogni campione viene inserito più volte nella rete e solo una frazione della modifica sinaptica viene effettivamente addizionata ai valori consolidati). La velocità di apprendimento è regolata dal *tasso di apprendimento* η ($0 \leq \eta \leq 1$).

Osservazioni

- Nel caso dell'addestramento supervisionato, la fase di apprendimento si conclude una volta che l'errore commesso tra valori di output desiderati e valori di output restituiti dalla rete raggiunge un livello minimo prefissato; raggiunta tale soglia, i valori sinaptici vengono registrati ed è possibile testare la rete neurale con nuovi campioni chiamati *test*.

Osservazioni

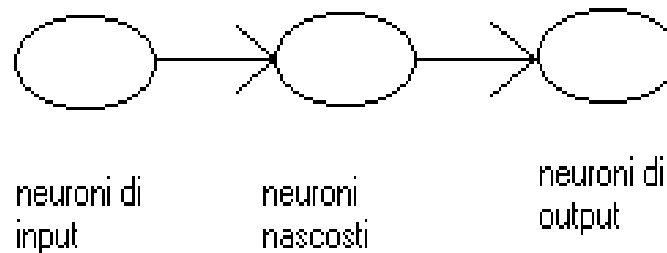
- Nel caso dell'addestramento non supervisionato, invece, non esiste una misura di errore, quindi si utilizzano altri indici, come, ad esempio, la misura della variazione dei pesi sinaptici. Quando la variazione dei pesi tende ad essere piccola, allora la rete ha raggiunto una soluzione.

Algoritmi di apprendimento

- Esistono diversi algoritmi di apprendimento tra cui:
- Algoritmo di apprendimento Hebbiano; esso è alla base di tutti gli algoritmi di apprendimento ed è basato sulla regola di Hebb che afferma che se due neuroni connessi tra loro sono contemporaneamente attivi, il peso sinaptico deve essere rafforzato.
- Algoritmo di apprendimento Back-Propagation; in questo tipo di procedura, l'errore che viene valutato una volta confrontato l'output ottenuto con quello desiderato, viene propagato all'indietro in modo tale da correggere i pesi sinaptici e procedere con un nuovo ingresso di input

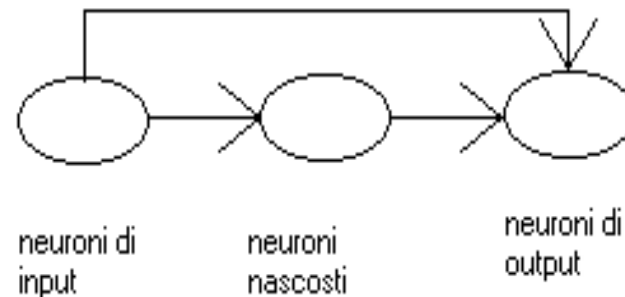
Esistono tre varianti principali delle reti *Back Propagation*:

- **Connessioni sequenziali.** Sono il tipo di connessione più semplice che si può instaurare fra i diversi strati di neuroni, gli strati possono essere di qualsivoglia numero, la particolarità di questa connessione è che non prevede salti e ritorni.



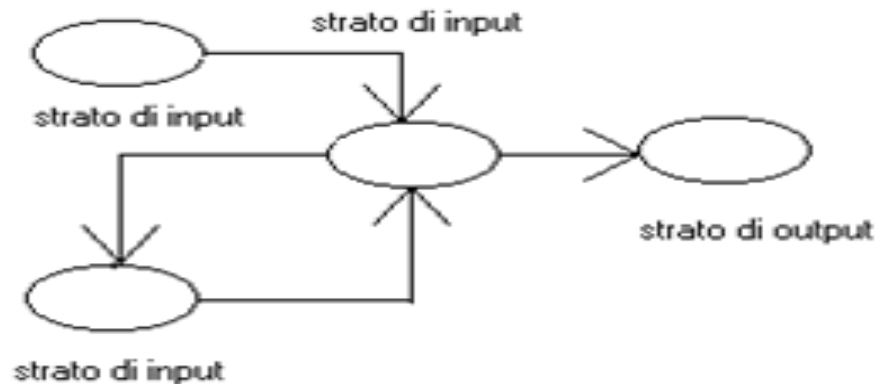
Esistono tre varianti principali delle reti *Back Propagation*:

- **Connessioni a salto.** Questo tipo di connessione è leggermente più complicata della precedente e prevede che si instaurino dei legami oltre che fra strati adiacenti anche fra strati non consecutivi. Anche questo tipo di connessione prevede che il numero di strati nascosti sia di qualsivoglia grandezza



Esistono tre varianti principali delle reti *Back Propagation*:

- **Connessioni ripetute.** Questo è il tipo di connessione più complesso delle tre, e sovente utilizzata nello studio delle serie storiche finanziarie in quanto possiede al suo interno una sorta di memoria storica. La particolarità di questo tipo di connessione è che i neuroni appartenenti agli strati nascosti possono ritornare indietro e presentarsi alla rete come neuroni di input. Anche in questo caso il numero di strati è lasciato alla scelta del ricercatore



Algoritmi di apprendimento

Algoritmi Genetici

- L'algoritmo genetico fa parte della grande famiglia dei modelli evolutivi e si basa sulla teoria evolutiva di Darwin.
- L'idea di base è quella di tradurre i principi dell'evoluzione naturale in un algoritmo in grado di risolvere i problemi in modo simile a quello utilizzato nei processi naturali.
- L'algoritmo genetico viene utilizzato come metodo di ottimizzazione e si basa sull'ipotesi che esista una qualche legge deterministica alla base, di cui non si conosce la funzione che la esprime; obiettivo è quello di determinare la combinazione di argomenti che rende massimo il valore di quest' ultima.

Algoritmi di apprendimento

Algoritmi Genetici

- Il principio base è quello di considerare ogni argomento come una combinazione di diversi elementi i quali sono, da un lato, ereditari, e quindi trasmissibili, e dall'altro sono sottoposti a mutazioni casuali.
- Il processo di riproduzione si effettua attraverso la tecnica dell'incrocio (*crossover*), il processo di mutazione avviene attraverso il cambiamento del valore di un gene. Ogni soluzione trovata viene tradotta in una stringa di numeri. Questo modello computazionale fu sviluppato da Holland negli anni settanta dello scorso secolo.

Modelli di reti neurali

Il modello di McCulloch-Pitts, 1943:

- il primo modello di rete neurale rappresenta la struttura fondamentale sulla quale si basano altri modelli neurali artificiali molto complessi. La rete è composta da un solo neurone e la funzione di attivazione, che in questo modello corrisponde ad una funzione a gradino, può assumere due soli valori:

- $$y_i = f(A_i) = f\left(\sum_{i=1}^N w_i x_i - \vartheta_i\right) = \begin{cases} 1 & \text{se } \sum_{i=1}^N w_i x_i \geq \vartheta_i \\ 0 & \text{se } \sum_{i=1}^N w_i x_i < \vartheta_i \end{cases}$$
- dove gli x_i rappresentano i segnali di ingresso, w_i sono i pesi, y rappresenta il segnale di output fornito dal modello, $f(\)$ è la funzione di attivazione che riceve come argomento il valore di attivazione $A_i = \sum_{i=1}^N w_i x_i - \vartheta_i$.

Modelli di reti neurali

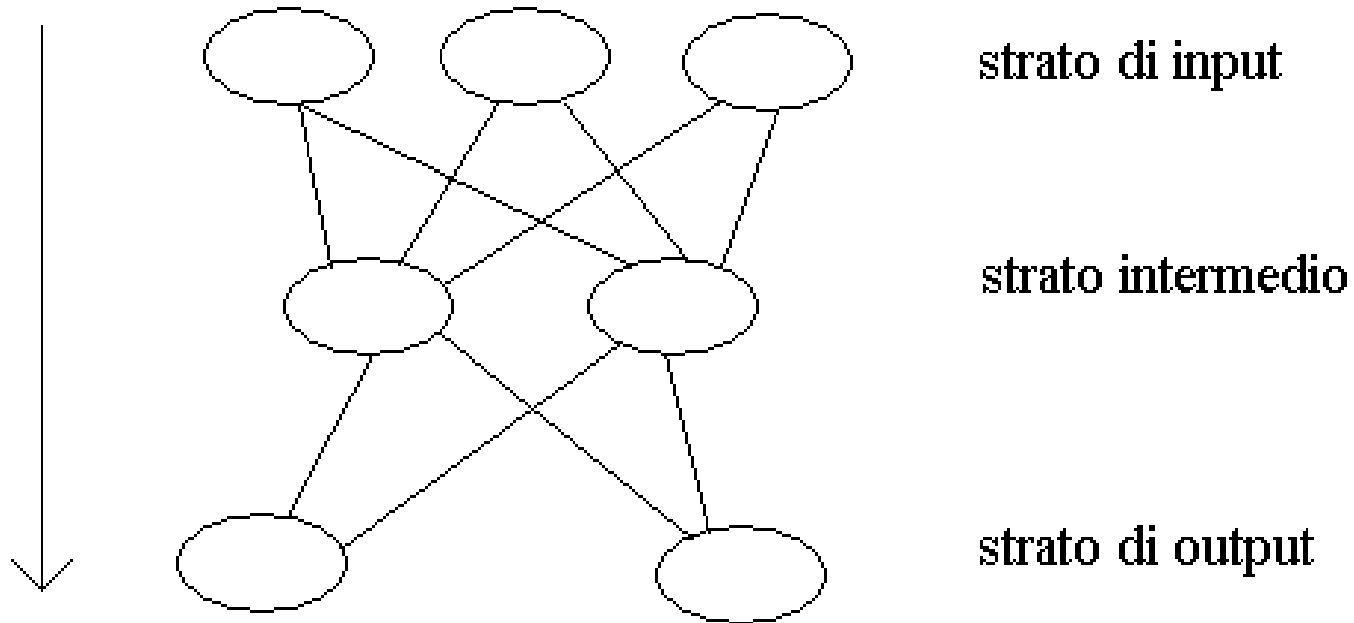
Multi Layer Perceptron:

- una rete Multi Layer Perceptron è caratterizzata solitamente da tre strati: strato di input, uno strato intermedio o nascosto e uno strato di output.
- Le unità appartenenti allo strato nascosto si chiamano *unità nascoste* in quanto il loro valore di attivazione non è misurabile così come non lo è il loro output. Reti neurali senza strato nascosto sono chiamate anche perceptron semplici.
- Una rete *Multi Layer Perceptron* è una rete che nasce con l'obiettivo di risolvere problemi non lineari.
- Una rete di questo tipo è detta *feed forward* poiché i segnali si muovono dallo strato di input allo strato intermedio prima, e di output dopo, senza avere connessioni trasversali (neuroni dello strato non sono collegati) o *back forward* (il segnale si muove in un senso e non nel senso opposto).

Modelli di reti neurali

- Le reti Multi Layer Perceptron sono di tipo feed-foward e sono generalmente più potenti in quanto hanno la capacità di risolvere qualsiasi tipo di problema. Tuttavia, questo tipo di rete richiede un algoritmo di apprendimento più complesso chiamato *Algoritmo di Back-Propagation* o altri algoritmi come IL METODO BFGS utilizzato nel pacchetto nnet di R.
- Questa libreria può essere considerata un pacchetto standard per l'utilizzo delle reti neurali con R; essa implementa le reti feed-foward e in particolare le Multi Layer Perceptron con un singolo strato nascosto. L'apprendimento non viene eseguito con l'algoritmo di backpropagation, ma con l'algoritmo BFGS che appartiene all'insieme dei metodi quasi-Newton.

Modelli di reti: Il Multi Layer Perceptron



Modelli di reti neurali

- Le reti di Ward e le Reti Jump Connection: sono tipologie di reti neurali che utilizzano l'architettura del tipo Back-Propagation;
 - Le reti di Ward vantano una maggior capacità previsiva e di generalizzazione in quanto sono composte da gruppi separati di neuroni che ricevono le informazioni contemporaneamente.
 - Le reti Jump Connection possiedono almeno due strati nascosti ciascuno dei quali è collegato con il precedente.

Modelli di reti neurali

- Le reti General Regression: si tratta di modelli neurali che si occupano di approssimare funzioni continue sulla base di un solo parametro che indica la qualità dell'approssimazione.
- Recurrent Networks: sono reti neurali in grado di considerare la componente temporale dei dati e per questo motivo sono quelle maggiormente utilizzate nelle previsioni finanziarie.

Modelli di reti neurali

- Reti di Hopfield: sono reti che si basano sulla struttura Back-Propagation e sono presentate come memorie auto-associative, cioè sono in grado di recuperare un'informazione sulla base di una componente parziale della stessa.
- Reti Radial Basis Function: sono reti computazionalmente meno pesanti perché si basano sull'approssimazione di funzioni complesse con combinazioni lineari di funzioni relativamente più semplici.

Modelli di reti neurali

Reti di Kohonen:

- particolari architetture che hanno lo scopo principale di scoprire regolarità tra i dati in modo autonomo considerando il fatto che l'attività di un neurone influenza l'attività dell'unità vicina. Le reti di Kohonen permettono di classificare oggetti senza la supervisione esterna. Esse sono in grado, visto il loro particolare tipo di addestramento, di correggere i pesi sinapsici in modo autonomo senza riferirsi ad un particolare output.
- In pratica questo tipo di rete è in grado di apprendere eventuali relazioni che si instaurano tra i vari input in modo da poterli organizzare in categorie. Queste reti tengono conto non solo delle connessioni sinottiche tra neuroni ma anche dell'influenza che può avere un neurone sul vicino.

Modelli di reti neurali

Reti di Kohonen:

- E' stato osservato che, nel caso biologico, i neuroni che sono fisicamente vicini a neuroni attivi hanno i legami più forti mentre quelli ad una particolare distanza hanno legami inibitori. A questa caratteristica Kohonen (Kohonen 1982) associa un algoritmo detto " mappa di Kohonen " il quale si basa sull'idea che esista una configurazione all'interno dei dati ma questa non è esplicita. L'obiettivo della rete è quello di scoprire questa regolarità tra i dati

Vantaggi e svantaggi di un modello a rete neurale

- le reti neurali lavorano in parallelo;
- l'elaborazione dei sistemi nervosi è distribuita su molti elementi, quindi ci sono molti elementi, neuroni, che lavorano alla stessa operazione;
- si accede al dato non attraverso i mezzi "tradizionali", cioè attraverso l'indirizzo di memoria a cui far riferimento, ma attraverso le informazioni, anche parziali, riferite al "contenuto";
- si impara attraverso un apprendimento on line, cioè in funzione dell'esperienza, con o senza l'aiuto di un istruttore esterno.

Vantaggi e svantaggi di un modello a rete neurale

- Le reti neurali sono tecniche che mirano a dotare una macchina della possibilità di “imparare” da set di dati, autoistruirsi al fine di risolvere problemi. Esse oggi appartengono ad un ramo dell’intelligenza artificiale conosciuto come machine learning.
- L’aspetto più rilevante del machine learning consiste nella possibilità di risolvere problemi senza scrivere un codice che ne implementi la soluzione; è la macchina a farlo partendo dai dati in input.
- Fondamentali per la riuscita dell’analisi sono:
 - l’algoritmo impiegato per la fase di autoapprendimento (detto *training*)
 - l’insieme dei dati utilizzati.

Vantaggi e svantaggi di un modello a rete neurale

- effetto black-box: i modelli prodotti dalle reti neurali, anche se molto efficienti, non sono spiegabili in linguaggio simbolico umano per cui i risultati vanno accettati “così come sono”;
- la realizzazione di una rete neurale dipende molto dall’esperienza del creatore e dalla sua sensibilità e capacità di adottare e capire la rete in funzione del dataset.

PROC HPNEURAL (SAS STUDIO)

- PROC HPNEURAL
-
- Parola chiave: [HPNEURAL](#)
- **Contesto:** [DEFINIZIONE DELLA PROCEDURA] PROC HPNEURAL
-
- Syntax: PROC HPNEURAL <DATA=SAS-data-set > < DISTR=ALL | SPLIT > < NOPRINT > ;
- PERFORMANCE performance-options ;
- ARCHITECTURE architecture-option ;
- ID variables ;
- INPUT variables < / LEVEL=INT | LEVEL=NOM < MISSING=MAP > > ;
- WEIGHT variable | _INVERSE_PRIORS_ ;
- HIDDEN number ;
- TARGET variables < / LEVEL=INT | LEVEL=NOM > ;
- PARTITION ROLEVAR=variable(TRAIN=value | VALIDATE=value) ;
- PARTITION FRACTION(TRAIN=number | VALIDATE=number) ;
- TRAIN < NUMTRIES=number > < MAXITER=number >
- <VALID=_NONE_ > <OUTMODEL=SAS-data-set > ;
- SCORE OUT=SAS-data-set <MODEL=SAS-data-set > ;
- CODE FILE='external-file' | fileref ;

PROC HPNEURAL (SAS STUDIO)

- The HPNEURAL procedure is a high-performance procedure that trains a multilayer perceptron neural network. For more information about multilayer perceptron neural networks, see Bishop (1995).
- PROC HPNEURAL can also use the trained network to score the input data set.
- PROC HPNEURAL reads and writes data in distributed form and makes full use of multicore computers and distributed computing environments to perform training and scoring.
- Training a multilayer perceptron neural network requires the unconstrained minimization of a nonlinear objective function.
- Because there are currently no practical methods to guarantee finding a global minimum of that objective function, one way to be reasonably sure of finding a good solution is to train the network multiple times using different sets of initial values for the weights.
- Thus, even problems with smaller numbers of variables and training observations can benefit from the use of multicore computers and distributed