



NARRAZIONE E ROBOTICA PER LA COMUNICAZIONE CON LABORATORIO APPLICATIVO

Laurea Triennale
Scienze e Tecnologie della Comunicazione
AA 2022-2023

Prof. Giorgio Poletti
giorgio.poletti@unife.it





Concetti e storia per
iniziare

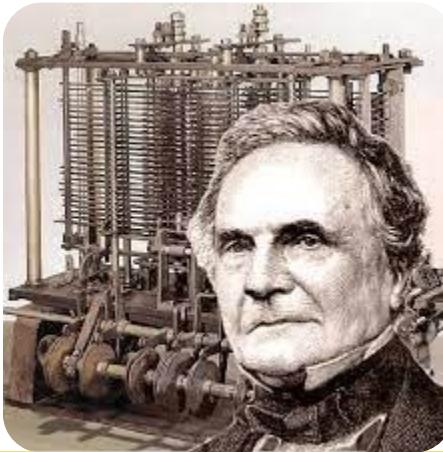
Machine Learning

Nessuno la esprime in questo modo, ma penso che l'intelligenza artificiale sia quasi una disciplina umanistica. È davvero un tentativo di comprendere l'intelligenza umana e la cognizione umana

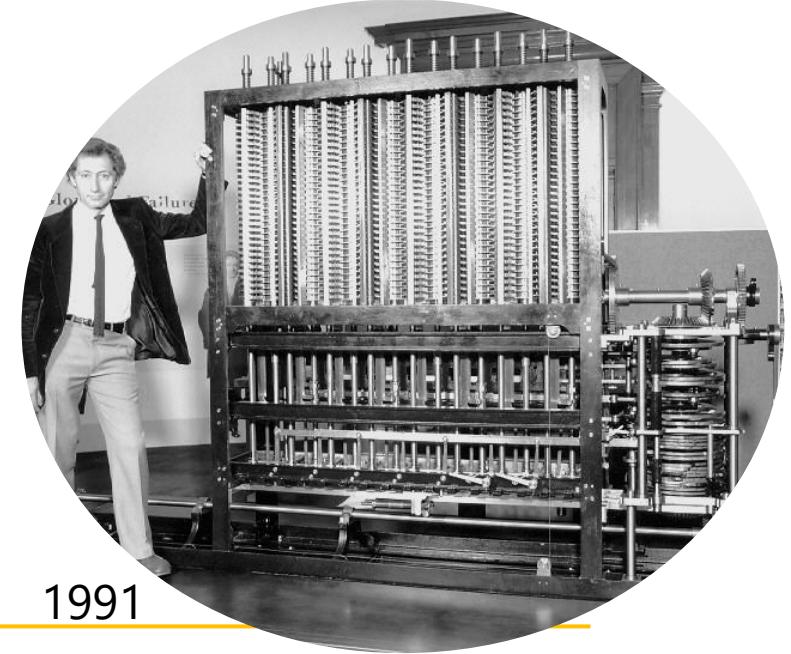
Sebastian Thrun



1 AI e Machine Learning



1821



1991

- L'AI è la disciplina che studia la progettazione, lo sviluppo e la realizzazione di sistemi in grado di simulare le abilità, il ragionamento e il comportamento umani.
- È nel **XVII** secolo che vengono costruite le prime macchine in grado di effettuare calcoli automatici (Blaise Pascal, Gottfried Wilhelm von Leibniz).
- Nella prima metà dell'Ottocento **Charles Babbage** con la sua "macchina analitica" anticipava le caratteristiche dei moderni calcolatori. A partire dalla seconda metà degli anni '30 vengono pubblicati i lavori di **Alan Touring**, padre dell'informatica.

2

AI e Machine Learning

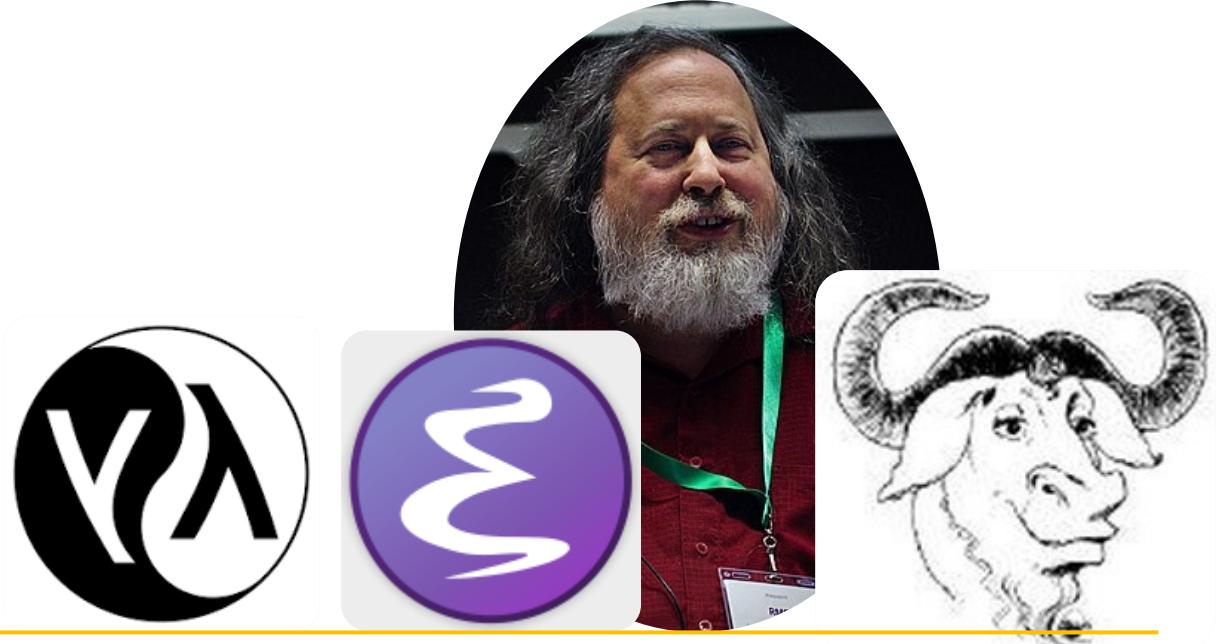


Intelligence Project, il Massachusetts Institute
of Technology di Cambridge

- Nel 1943, con il lavoro del neurofisiologo Warren Sturgis McCulloch e del matematico Walter Harry Pitts, mostrano, con un modello matematico, come dei semplici neuroni possano essere combinati per calcolare le tre operazioni logiche elementari NOT, AND, OR.
- È da questi assunti che nasceranno le reti *neurali artificiali*: vengono così avviati gli studi *sull'intelligenza artificiale*.
- “INTELLIGENZA ARTIFICIALE” ha una data di nascita precisa: viene utilizzata per la prima volta dai matematici e informatici John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon in un documento del 1955.

3

Machine Learning

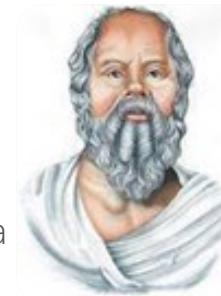


- Nel 1958 McCarthy sviluppa il linguaggio Lisp (List Processor) per studiare la computabilità di funzioni ricorsive su espressioni simboliche che per lungo tempo è stato il linguaggio di riferimento nei progetti di intelligenza artificiale.
- Nel 1986, Jay McClelland e David Rumelhart in una pubblicazione sotengono che una rete neurale è un grafo diretto non lineare, nel quale ogni elemento di elaborazione (ogni nodo della rete) riceve segnali da altri nodi ed emette a sua volta un segnale verso altri nodi.
- Si sviluppa il modello di rete neurale multistrato (MLP- Multilayer Perceptron) dove ogni strato di nodi è completamente connesso con quello successivo e utilizza una tecnica di apprendimento supervisionato chiamata retropropagazione dell'errore.

4

AI e Machine Learning

Legge universale partendo da singoli casi particolari.



Procedimento inferenziale che permette di derivare una conclusione a partire da alcune premesse

- L'AI debole non comprende totalmente i processi cognitivi umani, ma si occupa sostanzialmente di PROBLEM SOLVING (*risposte a problemi sulla base di regole conosciute*).
- La caratteristica distintiva dell' AI forte è l'analisi del linguaggio per comprenderne il significato, anche attraverso metodi induttivi (*maieutica*) o deduttivi.
- AI → MACHINE LEARNING (*apprendimento automatico*): la capacità di imparare ed eseguire compiti da parte della macchina sulla base di algoritmi che imparano dai dati in modo iterativo.

5

Machine Learning

Il computer apprende la nuova conoscenza in vari modi:

- Dall'osservazione dell'ambiente esterno: l'agente osserva il mondo esterno e impara dai feedback delle sue azioni e dai suoi errori.
- Dalla base di conoscenza: l'esperienza e la conoscenza dell'ambiente operativo sono conservate in un database detto base di conoscenza (*knowledge base o KB*).
- La base di conoscenza non è mai vuota ma contiene alcune informazioni iniziali inserite manualmente dal progettista (conoscenza pre-esistente o pregressa). Successivamente, durante il processo di apprendimento, la **base cognitiva (KB)** viene modificata tramite i feedback con l'ambiente, in base alla esperienza diretta maturata nel corso del tempo.

6

Machine Learning

La **teoria della calcolabilità**, della computabilità, e della ricorsione cerca di comprendere quali funzioni possono essere calcolate tramite un procedimento automatico

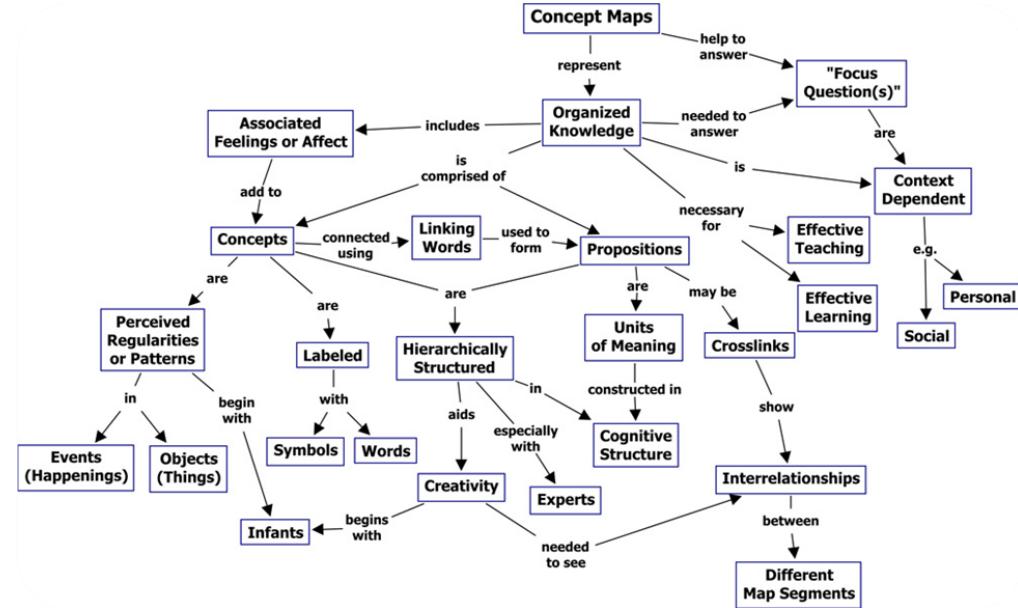
Nel machine learning la pre-elaborazione (o pre-processing) è la fase in cui si organizzano i dati, prima di avviare l'algoritmo di apprendimento. In particolar modo, nella fase di pre-processing si analizza il dataset per individuare eventuali correlazioni nei dati.

Gli attributi fortemente correlati possono essere indicati come unica variabile nel dataset, senza alcuna perdita di informazione → riduce la dimensione dei dati nel dataset e la complessità computazionale dell'algoritmo.

I rumori (dati irrilevanti) sono eliminati dal dataset.

La teoria della **complessità computazionale** parte della teoria della computabilità che studia le risorse minime necessarie per la risoluzione di un problema

1 Alberi decisionali

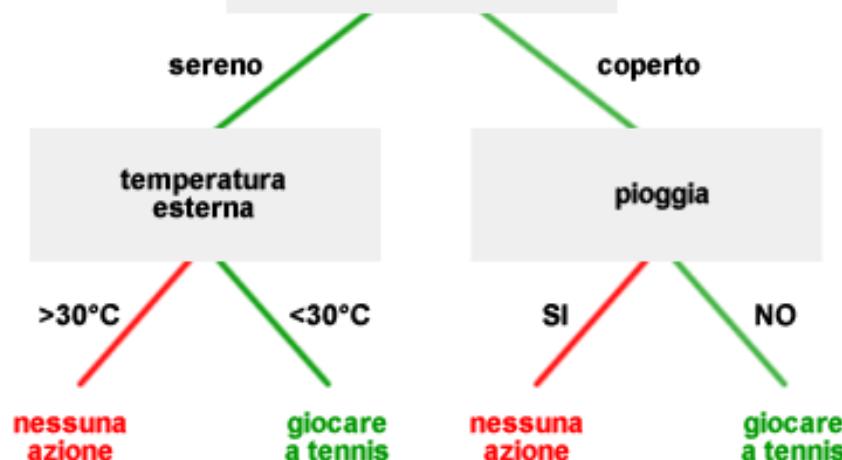


- Alberi di decisione sono utilizzati nei processi di apprendimento induttivo, quelli basati sull'osservazione dell'ambiente circostante.
- Il processo decisionale è rappresentato con un albero logico rovesciato, che parte da un nodo radice (root) e ogni nodo è una funzione condizionale. Azione o un calcolo basati su un test di dati
- Ogni nodo interno rappresenta un test su una proprietà o attributo,
- Un arco verso un nodo figlio rappresenta un possibile valore per quella proprietà o attributo,
- Una foglia il valore finale previsto a partire dai valori delle altre proprietà, che nell'albero è rappresentato dal cammino (*path*) dal nodo radice al nodo foglia.

2

Alberi decisionali

condizioni del cielo



*Regola 1: If (Outlook=Sunny) •
(Humidity=High) Then PlayTennis=No*

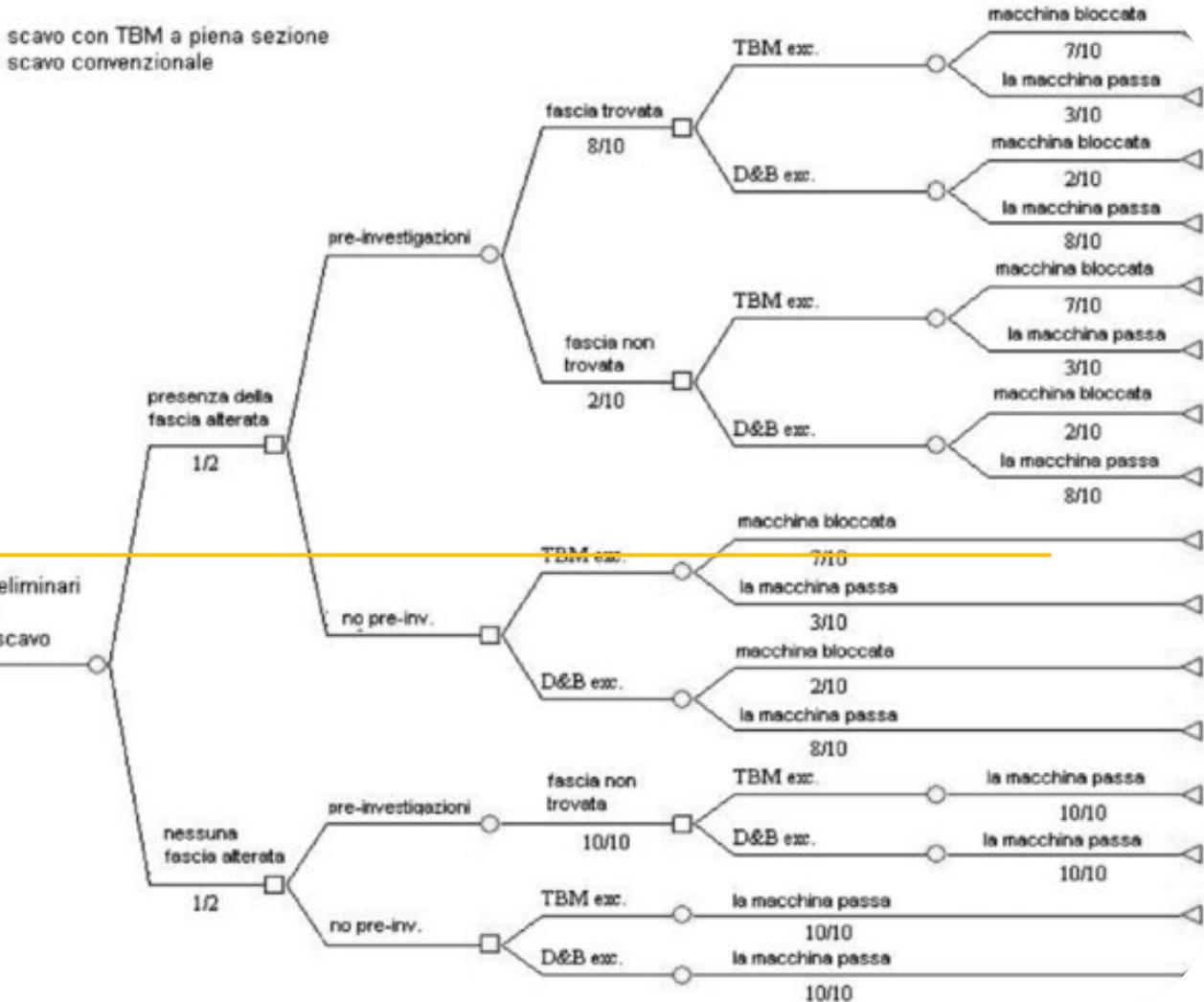
*Regola 2: If (Outlook=Sunny) •
(Humidity=Normal) Then PlayTennis=Yes*

*Regola 3: If (Outlook=Overcast)
Then PlayTennis=Yes*

*Regola 4: If (Outlook=Rain) • (Wind=Strong)
Then PlayTennis=No*

*Regola 5: If (Outlook=Rain) • (Wind=Weak)
Then PlayTennis=Yes*

ΔM exc. : scavo con TBM a piena sezione
D&B exc. : scavo convenzionale



3

Alberi decisionali

- Gli alberi logici, **vantaggi**
 - ✓ semplicità
 - ✓ facili da capire e da eseguire
 - ✓ facilmente comprensibile dagli esseri umani; l'uomo può verificare come la macchina giunge alla decisione. Eventualmente dissentire.
- Esempio. Un albero decisionale applicato alla medicina fornisce delle diagnosi. Essendo una decisione importante per il paziente, è sempre opportuno che un medico verifichi il processo di classificazione che ha portato la macchina a prendere quella decisione.
- Per un uomo è più facile farlo leggendo un albero decisionale che una rete neurale

La rappresentazione ad albero decisionale è poco adatta per i problemi complessi: lo spazio delle ipotesi diventa troppo grande si pone il problema della «complessità spaziale» dell'algoritmo.



Introduzione alle reti neurali

**Non c'è alcun modello del cervello più semplice del cervello stesso.
Per simulare alcuni comportamenti della struttura cerebrale degli esseri
viventi attraverso un insieme di regole di calcolo, occorre conoscere la
struttura del cervello umano e cercare di
riprodurlo con un modello matematico.**

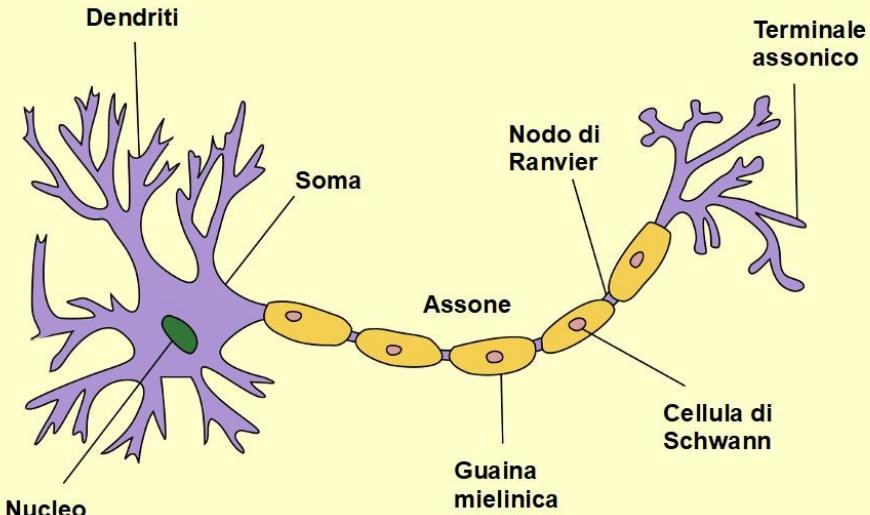
John von Neumann

1

Il neurone

Unità fondamentale del cervello umano. Ogni neurone riceve come input i segnali elettrici da tutti i dendriti, e se la somma pesata supera il valore di attivazione emette un impulso elettrico in uscita verso l'assone.

Schema della struttura di un Neurone

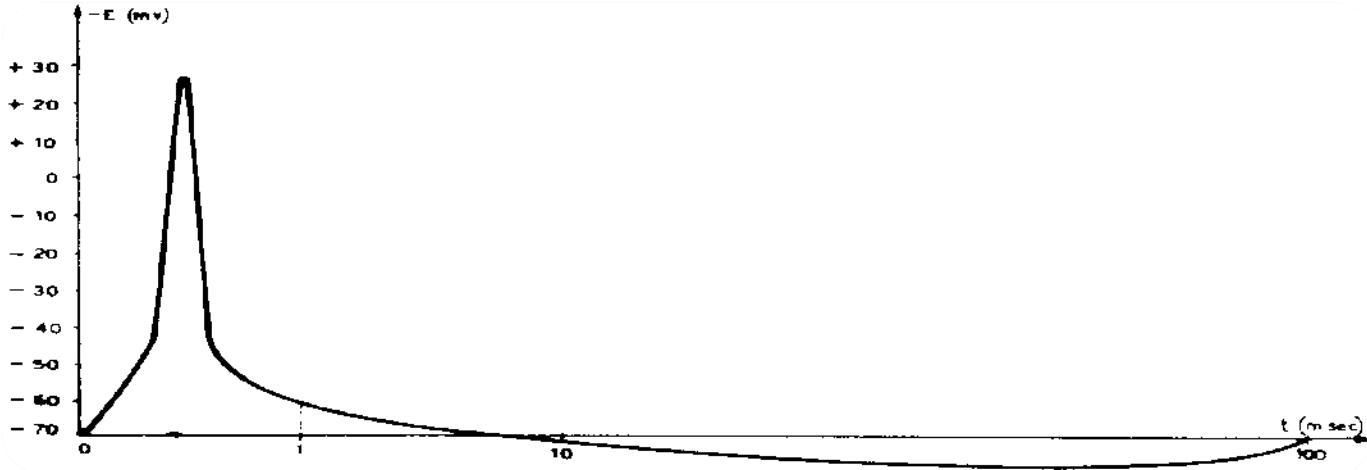


- ✓ In input i *dendriti* per la raccolta dei segnali provenienti da altri neuroni. Ogni neurone ha una *soglia di attivazione*.
- ✓ In output gli *assoni* per la distribuzione dei segnali ad altri neuroni.
- ✓ Un neurone attivo è attivo in misura variabile: il *livello di attività* è influenzato dai livelli di attività degli assoni provenienti da altri neuroni congiunti ai dendriti del neurone attraverso *sinapsi*.
- ✓ Una sinapsi ha una forza da positiva (eccitatoria), a zero (nessun effetto), a negativa (inibitoria). Le sinapsi *"pesano"* il segnale.

1

Il neurone

FUNZIONE DI SPARO



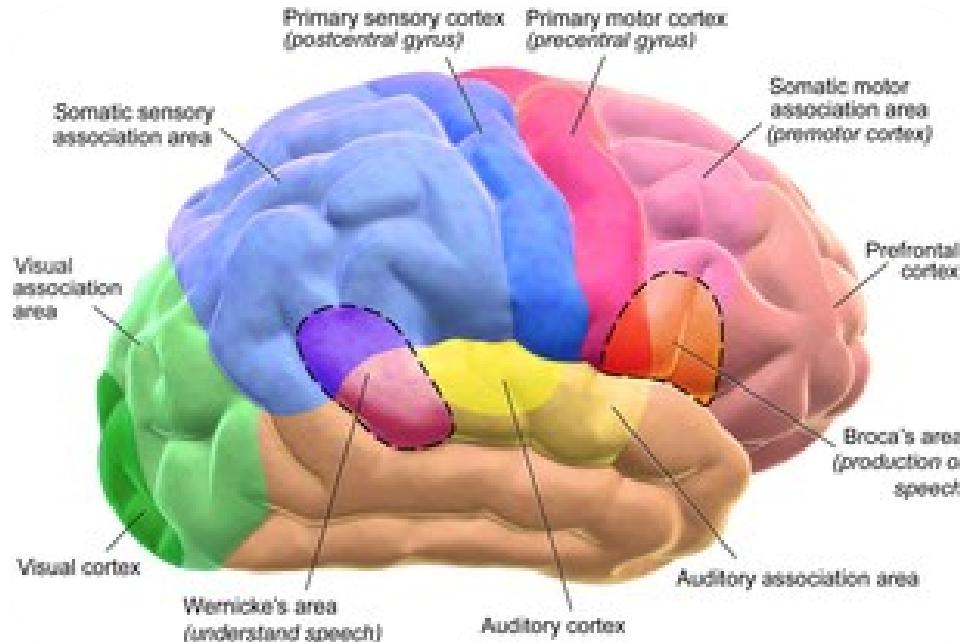
Andamento di $-E$, differenza di potenziale fra interno ed esterno della membrana cellulare del neurone.

Per uno stimolo con adeguata intensità, il neurone risponde con uno spike o non risponde: non c'è nessun tipo di risposta intermedia.

2

Il neurone

IL CERVELLO UMANO



- ✓ Elaborazione dell'informazione in parallelo e diffusa tra migliaia di neuroni porta all'emergere di processi cognitivi
- ✓ Elaborazione risiede in due emisferi separati da un corpo calloso che sono, a loro volta, suddivisi in regioni ben localizzate negli emisferi
- ✓ La struttura cerebrale è sempre in evoluzione: ogni giorno centinaia di neuroni muoiono, ma non c'è declino mentale perché aumentano le sinapsi per compensare le perdite

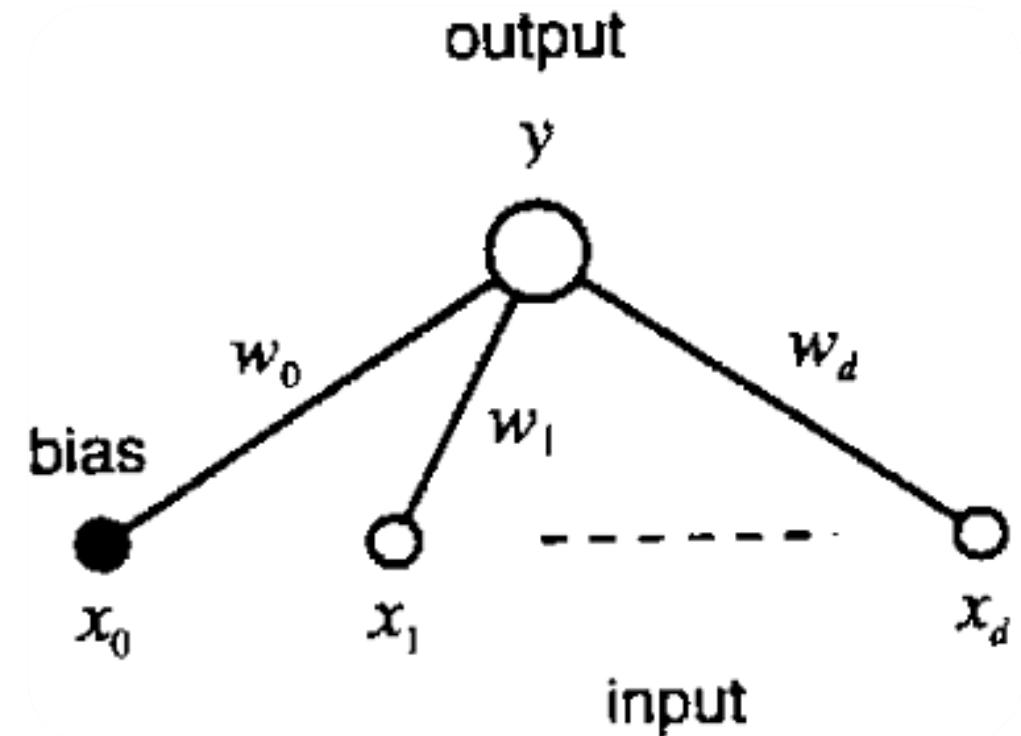
3

Il neurone

NEURONE ARTIFICIALE

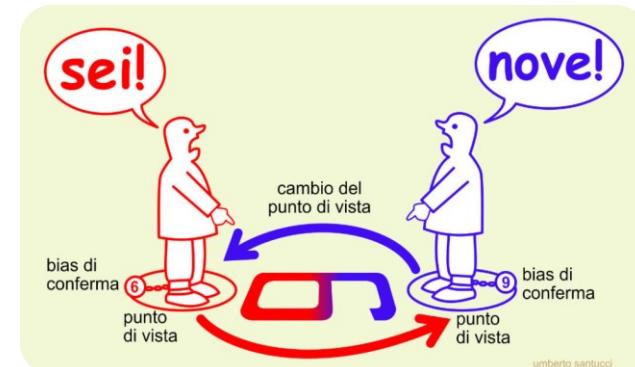
Modello matematico semplificato del neurone biologico

➤ Ad ogni input x_i è associato un peso w_i con valore positivo o negativo per eccitare o inibire il neurone. Il BIAS varia secondo la propensione del neurone ad attivarsi, per variare la soglia di attivazione del neurone.



BIAS (provenz. ant. biais «obliquo»: tendenza, inclinazione, distorsione).

- STATISTICA: *la tendenza a deviare dal valore medio.*
- ELETTRONICA: *la polarizzazione.*



4

Il neurone

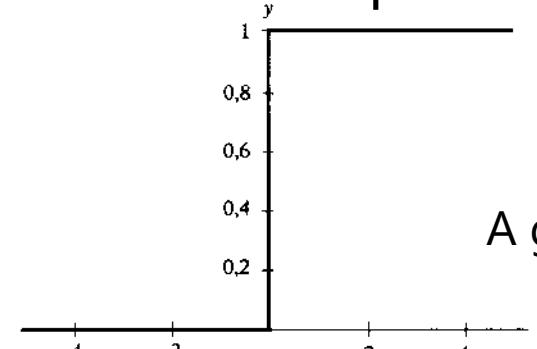
ALGORITMO E FUNZIONE DI ATTIVAZIONE

- Caricare i valori degli input x_i e dei pesi relativi w_i
- Calcolare la somma dei valori input pesata con i relativi pesi
- Calcolare il valore della funzione di attivazione g con il risultato della somma pesata
- y output del neurone è il risultato della funzione di attivazione

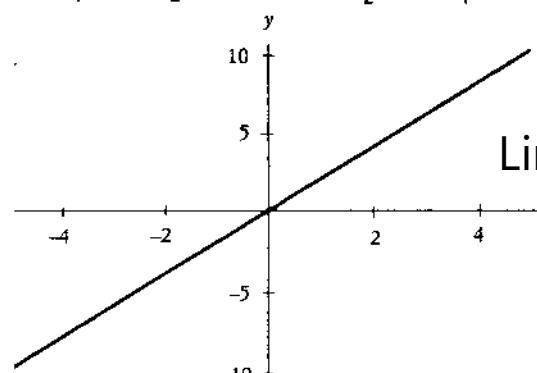
$$y(x) = g\left(\sum_{i=1}^d w_i x_i + w_0\right) = \bar{w}^T \bar{x}$$

$$g(A) = \frac{1}{1 + e^{-A}}$$

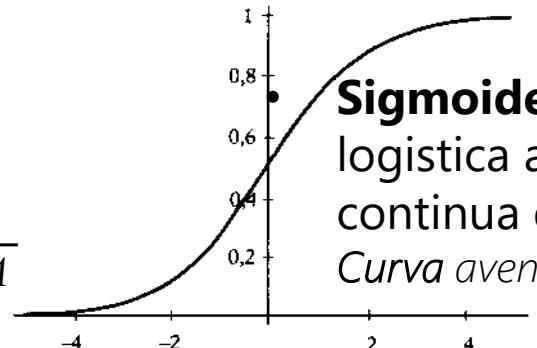
Determina la risposta del neurone



A gradino



Lineare continua



Sigmoide o funzione logistica a valori positivi, continua e derivabile
Curva avente un andamento ad "S"

La derivata descrive come varia una funzione $f(x)$ quando varia il suo argomento x .

5

Il neurone

LA RETE NEURONALE

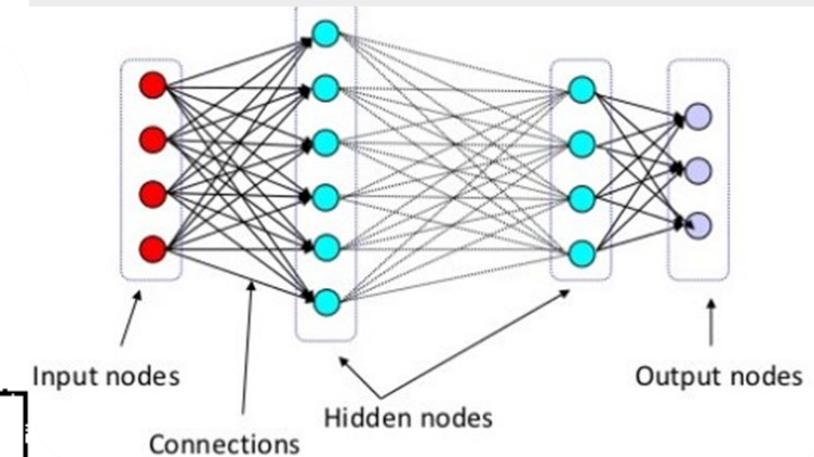
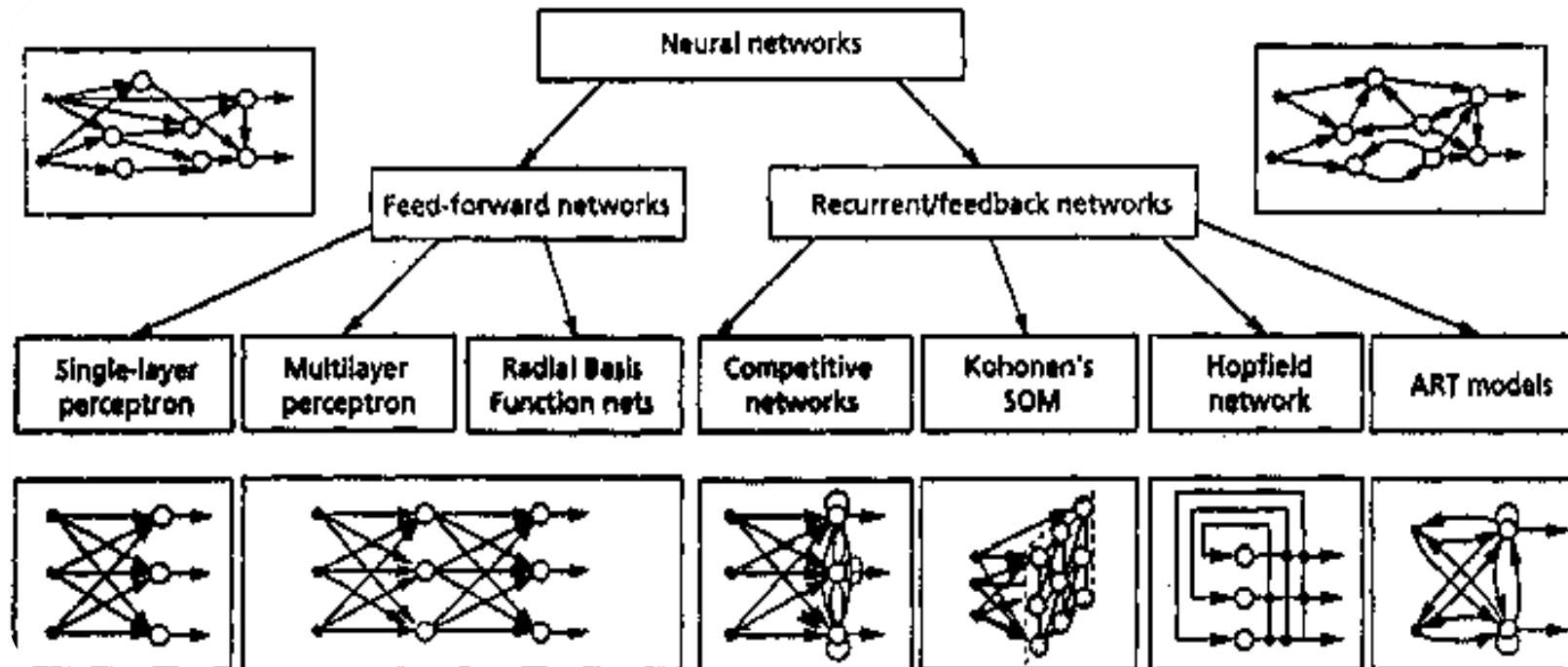
- Sistema dinamico avente la topologia di un grafo orientato con nodi, i neuroni artificiali, ed archi, i pesi sinaptici. Il termine rete è riferito alla topologia dei collegamenti tra i neuroni.
- Scatola nera, di cui si può ignorare il funzionamento, che associa un input a un output e le associazioni possono essere create con l'apprendimento
- Modello matematico che calcola la funzione $output=f(\text{input},\text{pesi})$ al variare dei pesi e senza specificare la forma della funzione f
- La rete neurale è un modello matematico che usa neuroni matematici

In media in un cervello ci sono 10 miliardi di cellule nervose dette **neuroni**. In media, ogni neurone `e collegato ad altri neuroni attraverso circa 10.000 sinapsi

6

Il neurone

RETI NEURONALI - ESEMPI



7

Il neurone

RETI NEURONALI - VANTAGGI



- Adatte per problemi che *non chiedono risposte accurate*, ma risposte approssimate con un grado di errore o di variazione
- **Generalizzazione**: producono buone risposte anche con input non considerati durante la creazione e l'addestramento
- **Facili da implementare**, basta definire il neurone e poi crearne delle copie e creare i collegamenti tra i neuroni
- **Funzionamento veloce** perché parallelo: ogni neurone usa solo il suo input
- **Stabilità dell'output** rispetto a valori di input: incompleti, con rumore, non ben noti, che accettano un grado di errore o di variazione
- Determinano il risultato tenendo conto contemporaneamente di tutti gli input

8

Il neurone

RETI NEURONALI - SVANTAGGI



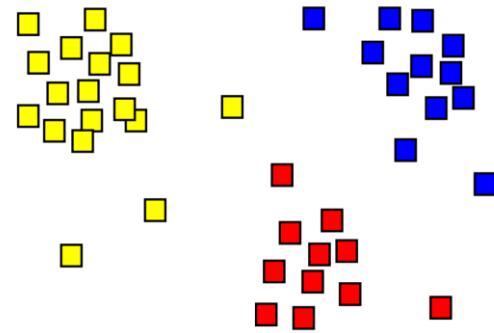
- Incapacità di rendere conto dell'elaborazione: non si può capire perché ha dato quel risultato specifico
- Non si può descrivere e localizzare la conoscenza memorizzata nella rete
- Carenza di hardware con cui implementare, si usano su computer seriali
- Tecniche di addestramento sofisticate che richiedono molto tempo di calcolo
- Non sempre esiste una rete che risolve il problema, perché non sempre esiste un algoritmo di apprendimento che converge dando un output della rete con basso errore
- Valori di output non sono precisi, ma hanno un margine in cui possono variare
- Serve una casistica di esempi molto ampia per ottenere un buon apprendimento e un basso errore di output

9

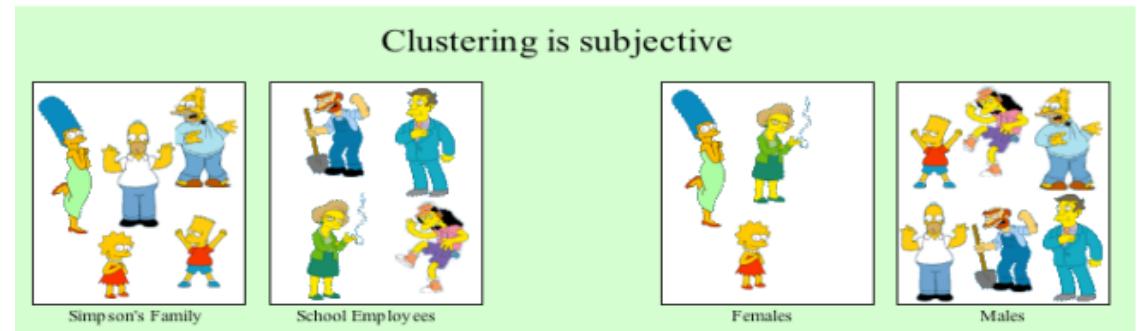
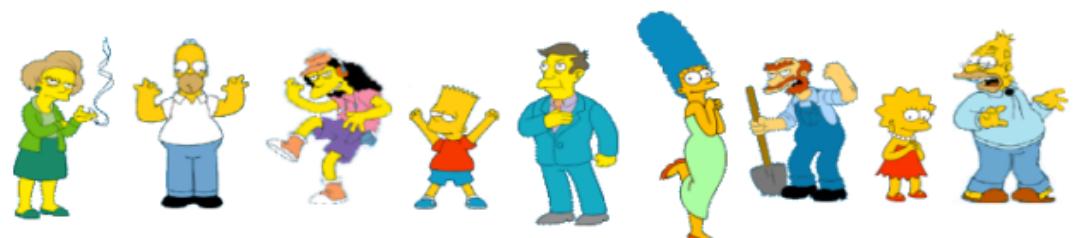
Il neurone

RETI NEURONALI - CAMPI DI APPLICAZIONE

- Pattern classificazione
- Clustering
- Approssimazione di funzioni
- Predizioni in serie temporali
- Ottimizzazione
- Memorie associative (**CAM**, Content Addressable Memory - *memoria a contenuto indirizzabile*)
- Controllo di apparati
- Elaborazione di segnali ed immagini



Kruskal (1977): " ... We call **clusters natural** if the membership is determined fairly well in a natural way by the data, and we call the clusters arbitrary if there is a substantial arbitrary element in the assignment process".



CLUSTERING: tecniche di analisi multivariata dei dati volte alla selezione e raggruppamento di elementi omogenei in un insieme di dati. Le tecniche di clustering si basano su misure relative alla somiglianza tra gli elementi

10

Il neurone

RETI NEURONALI E ALGORITMI GENETICI

Gli **algoritmi genetici** possono essere applicati alle reti neurali per risolvere problemi come:

- Scegliere la struttura della rete
- Scegliere i valori dei pesi per ridurre l'errore di output

Reti neurali e logica fuzzy

*Logica polivalente, ossia un'estensione della logica booleana. Legata alla teoria degli **insiemi sfocati** Funzione di grado di appartenenza, continuo [0,1]*

- Le reti neurali non calcolano con le variabili linguistiche e non usano regole qualitative. Esistono le reti neuro-fuzzy, strumento matematico molto efficiente e complesso che unisce i vantaggi delle due tecniche.

Cromosoma: una delle soluzioni ad un problema considerato e generalmente è codificata con un vettore di bit o di caratteri.

Popolazione: insieme di soluzioni relative al problema considerato.

Gene: parte di un cromosoma. Generalmente consiste in una o più parti del vettore di bit o caratteri che codificano il cromosoma.

Fitness: grado di valutazione associato ad una soluzione. La valutazione avviene in base ad una funzione appositamente progettata detta *funzione di fitness*.

Crossover: generazione di una nuova soluzione mescolando delle soluzioni esistenti.

Mutazione: alterazione casuale di una soluzione.

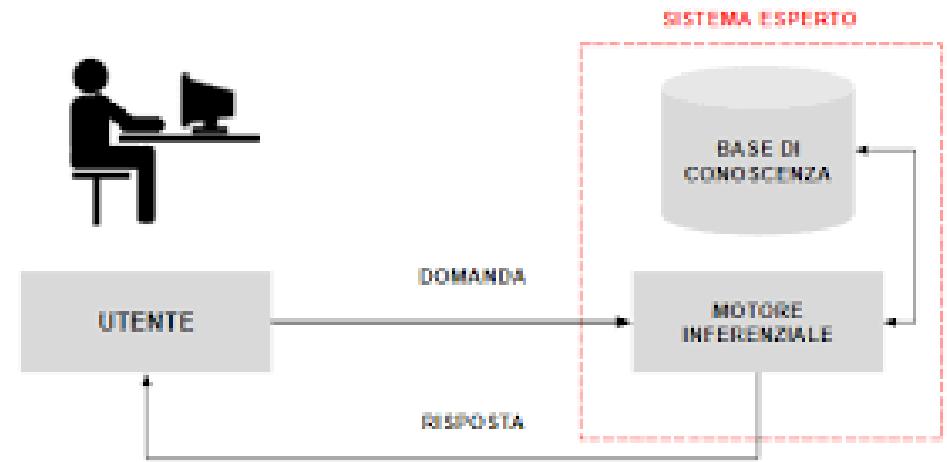
11

Il neurone

RETI NEURONALI E SISTEMI ESPERTI

Le reti neurali si differenziano molto dai **sistemi esperti** perci...

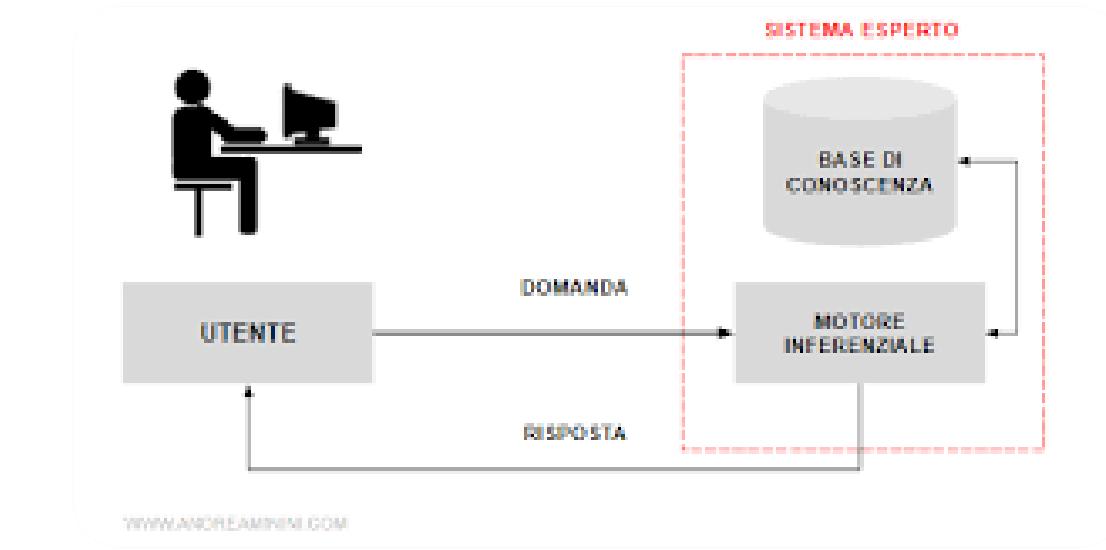
- Non usano conoscenze esplicite, ma conoscenze implicite contenute in una casistica molto ampia di esempi formata da vettori di numeri
- Usano numeri invece di simboli e regole
- Non usano regole tipo IF...THEN..
- Tutti i neuroni collaborano insieme
- Non vengono programmate, vengono addestrate con l'apprendimento
- Si accettano dati parziali e con rumore
- Il sistema esperto può spiegare perché e come ha ottenuto una conclusione, le reti neurali NON può giustificare perché e come ha ottenuto dei valori specifici in uscita



12

Il neurone

RETI NEURONALI E SISTEMI ESPERTI



Un sistema esperto può fare ragionamenti come "Socrate è un uomo, gli uomini sono mortali, Socrate è mortale", la rete neurale no ma può facilmente riconoscere una faccia

Un sistema esperto NON tiene conto del supporto materiale che ragiona → si dedica solo alla funzione del ragionamento e cerca di imitare la mente umana.

Il connessionismo è l'approccio allo studio della mente con le reti neurali; considera fondamentale la struttura del supporto materiale che ragiona, per cui cerca di imitare il cervello umano

A detailed illustration of a futuristic control room. In the foreground, several people in dark suits are seated at desks, working on computer monitors that display abstract blue patterns. The room is dominated by a massive, organic-looking brain with a large, bulbous head and many tentacle-like tentacles branching out across the ceiling and walls. The architecture is a mix of steampunk and modern elements, with pipes, ladders, and metal structures. The lighting is warm and golden, coming from various sources including chandeliers and built-in lights in the ceiling.

*A neural network and an expert system
discussing the solution to a social problem*

Apprendimento della rete

La sola cosa che interferisce con il mio apprendimento è la mia istruzione

Albert Einstein

1

Apprendimento in rete

LA PROGRAMMAZIONE

Esempio: per riconoscere un computer bisogna scrivere un programma con struttura IF...THEN che include tutti i casi possibili, quindi bisogna prevederli tutti

IF ha microprocessore THEN

IF ha memoria THEN

IF ha tastiera THEN

.....

PRINT "è un computer"

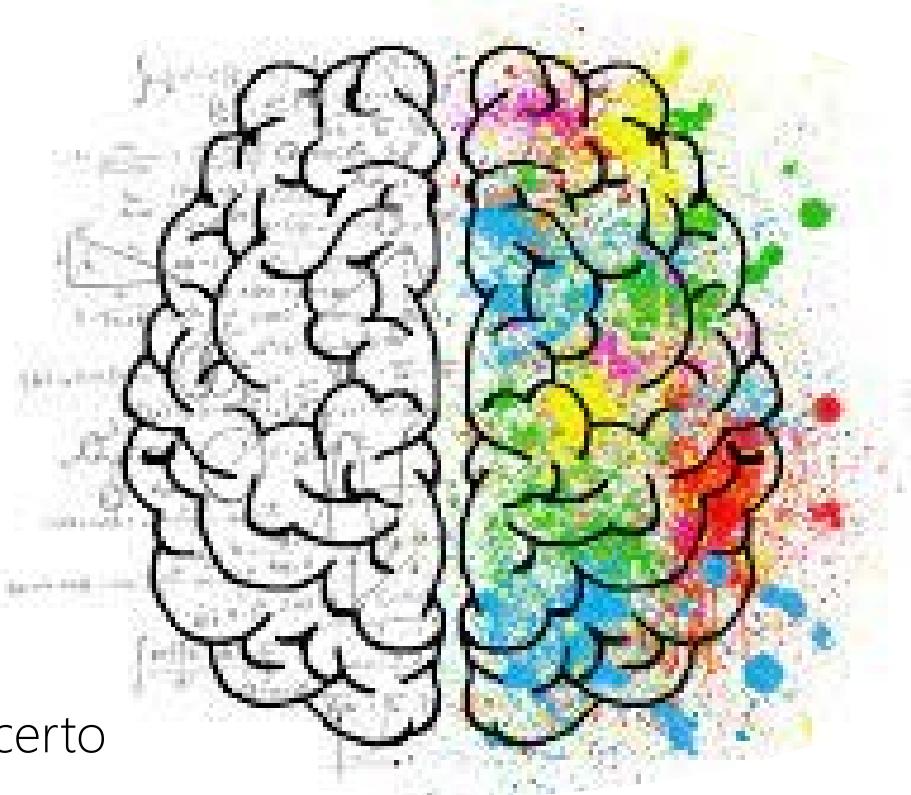


2

Apprendimento in rete

L'APPRENDIMENTO

- Apprendere «migliorare la capacità di esecuzione di un certo compito attraverso l'esperienza».
- Esempio: *all'inizio la rete neurale non conosce il concetto di computer. Si creano tanti esempi di computer e si fornisce ogni esempio alla rete neurale dicendogli che è un esempio di computer. La rete neurale impara e si crea una esperienza sul riconoscimento dei computer, così la prossima volta che ha in input una descrizione diversa di un computer lo riconosce subito.*



3

Apprendimento in rete

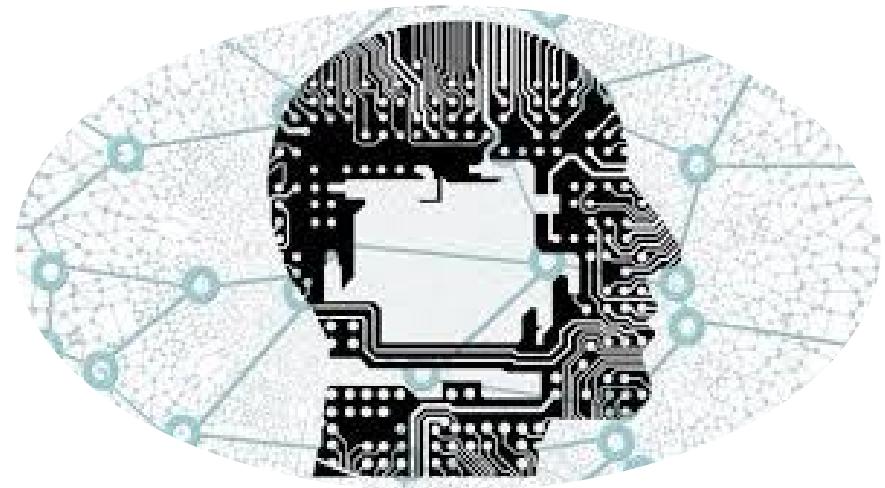
NON PROGRAMMARE MA APPRENDERE

- La PROGRAMMAZIONE serve solo per creare il software che crea la rete e l'algoritmo di apprendimento. Per insegnare alla rete a risolvere un problema, occorre un periodo di apprendimento in cui insegnare alla rete come comportarsi con l'input che riceve, perché all'inizio la rete non ha nessuna forma di conoscenza.
- L'operatore crea la struttura della rete e quindi i pesi sono gli unici parametri che possono essere modificati. Infatti la conoscenza è memorizzata nei suoi pesi e la rete apprende usando tecniche di ottimizzazione per variare i valori dei pesi, cercando di minimizzare una funzione di errore.

Tipi di apprendimento

1. *SUPERVISIONATO*
2. *NON SUPERVISIONATO*
3. *HEBBIANO*

Ogni modello di rete neurale ha il suo specifico tipo di apprendimento



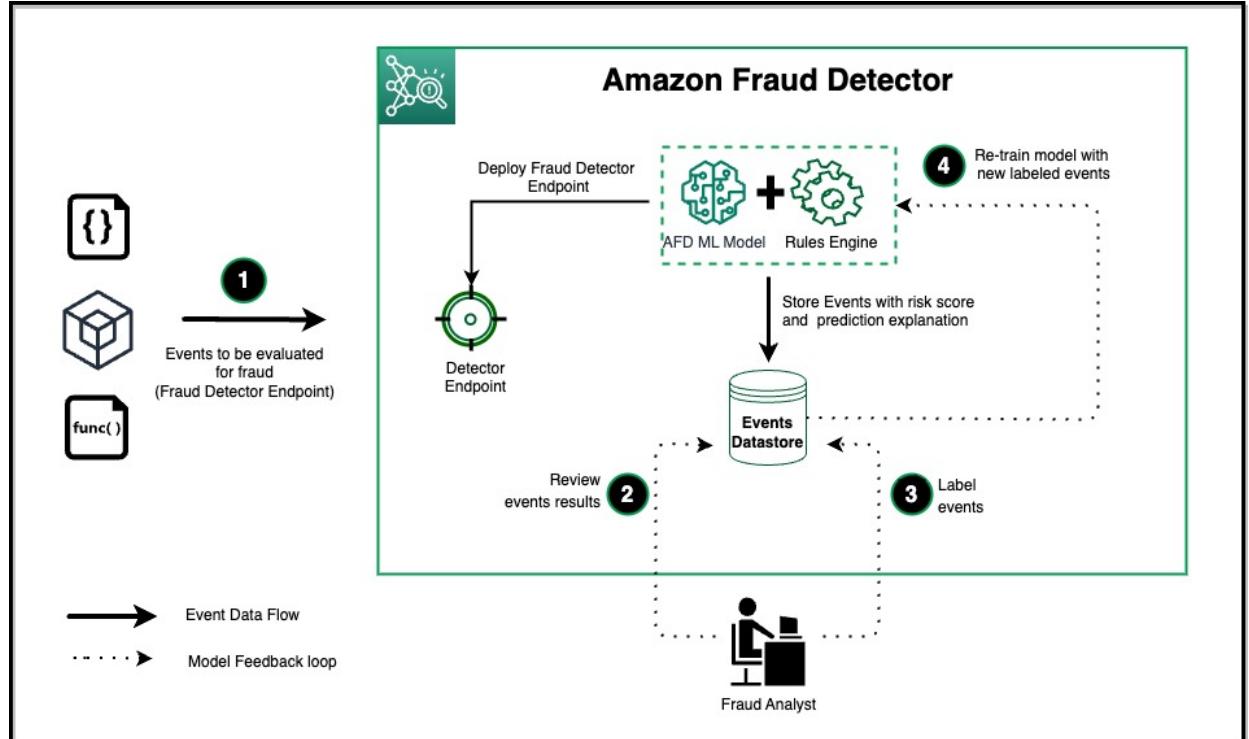
4

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO

Ipotesi induttiva

- Alla rete viene presentato un training set preparato da un supervisore esterno, e composto da molte coppie significative di valori (input,output atteso);
 - la rete riceve l'input e calcola il suo corrispondente output
 - per un certo input, l'errore è dato dalla differenza tra l'output della rete e l'output atteso; serve a supervisionare l'apprendimento per far capire alla rete quanto si sbaglia nel calcolare quell'output
 - la rete modifica i pesi in base all'errore cercando di minimizzarlo e commetterà sempre meno errori.
- Occorre quindi preparare alcuni esempi di funzionamento studiati appositamente, e la rete impara da questi esempi.
- Modelli che usano questo apprendimento: *perceptron (un tipo di classificatore binario)*, *multi layer perceptron*, *radial basis function*

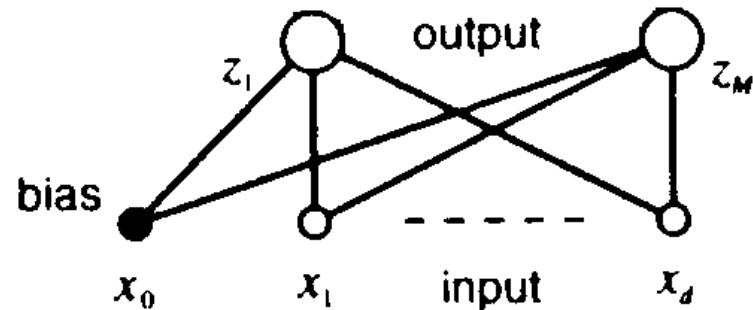


5

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: PERCEPTRON

- Modello semplice creato da *Frank Rosenblatt* nel 1957 per riconoscere immagini simulando la percezione umana, da cui il nome del modello. Ogni neurone ha funzione di attivazione a gradino ed uscita con valori binari: 1 indica la presenza di un oggetto, 0 l'assenza. Questo modello è composto da una fila di neuroni affiancati.



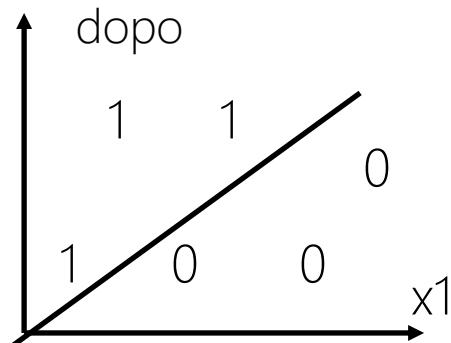
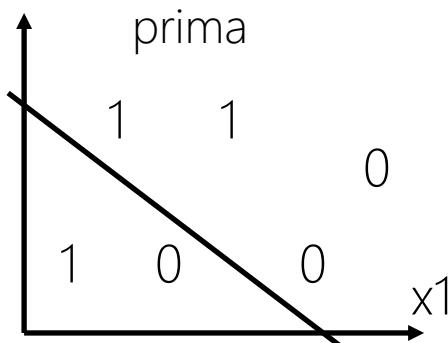
- L'apprendimento consiste nel variare i pesi e nel muovere una retta di separazione del piano fino a dividere correttamente i valori input in due gruppi.

6

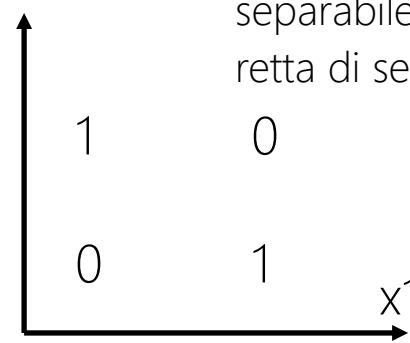
Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: PERCEPTRON

- Esempio: un oggetto è caratterizzato dalla sua altezza x_1 e larghezza x_2 , e ci sono due tipi di oggetti. Si riportano sul piano degli input le misure di alcuni oggetti e si rappresentano i due tipi di oggetto con 1 e 0. Si vuole stabilire una semplice regola per decidere a quale classe appartengono.
- Spazio di 2 valori input linearmente separabile e Perceptron con apprendimento trova sempre una retta di separazione del piano nei due semipiani con i gruppi di oggetti



➤ Spazio di 2 valori input non linearmente separabile e Perceptron non trova una retta di separazione



7

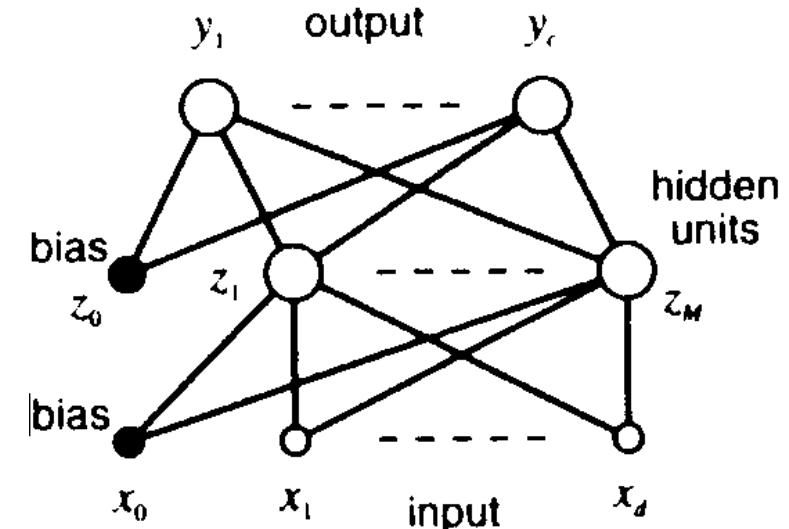
Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: MULTI LAYER PERCEPTRON (MLP)

- Formato dalla sovrapposizione di vari Perceptron
- Livello di input di contiene neuroni, variabili esplicative (1) , in figura sono 2
- Ogni neurone è collegato con tutti i neuroni dello strato precedente e successivo, i neuroni sullo stesso strato non sono collegati, non esistono cicli dallo strato di output verso lo strato di input
- La conoscenza viene elaborata dal livello input verso il livello output, cioè si calcola l'output di tutti i nodi di un livello che diventano poi gli input dei nodi del livello successivo

(1) Variabili di input che viene fornita alla rete neurale per eseguire una predizione o una classificazione.

Ad esempio, se stiamo addestrando una rete neurale per riconoscere immagini di animali, le variabili esplicative potrebbero essere i pixel dell'immagine stessa, o le caratteristiche estratte dall'immagine come contorni, colori, ecc.



8

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: MULTI LAYER PERCEPTRON (MLP)

- MLP con d neuroni input, M neuroni hidden, c neuroni output, g funzione di attivazione dei neuroni, w_{kj} il peso dal neurone k al neurone j

$$y_k(x) = g\left(\sum_{j=0}^M w_{kj}^{(2)} g\left(\sum_{i=0}^c w_{ji}^{(1)} x_i\right)\right)$$

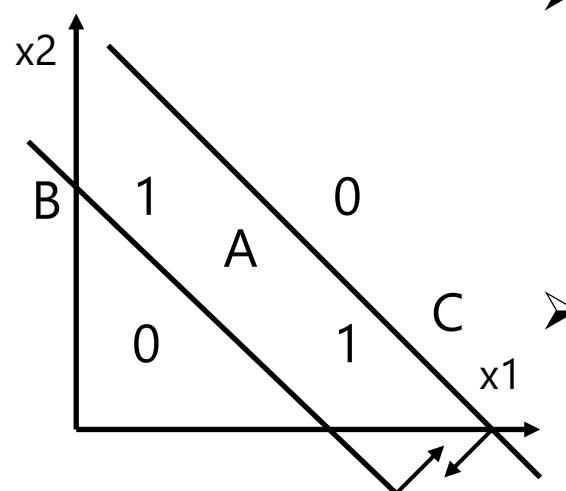
- tutti i neuroni allo stesso livello hanno la stessa funzione di attivazione
- la funzione di attivazione **softmax** crea ogni output in $[0,1]$ e somma di tutti gli output pari a 1, in modo da interpretare la risposta della rete come stime di probabilità

$$g(a_i) = \frac{e^{a_i}}{\sum_{i=1}^n e^{a_i}}$$

9

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: REGIONE DELIMITATA
MULTI LAYER PERCEPTRON (MLP)



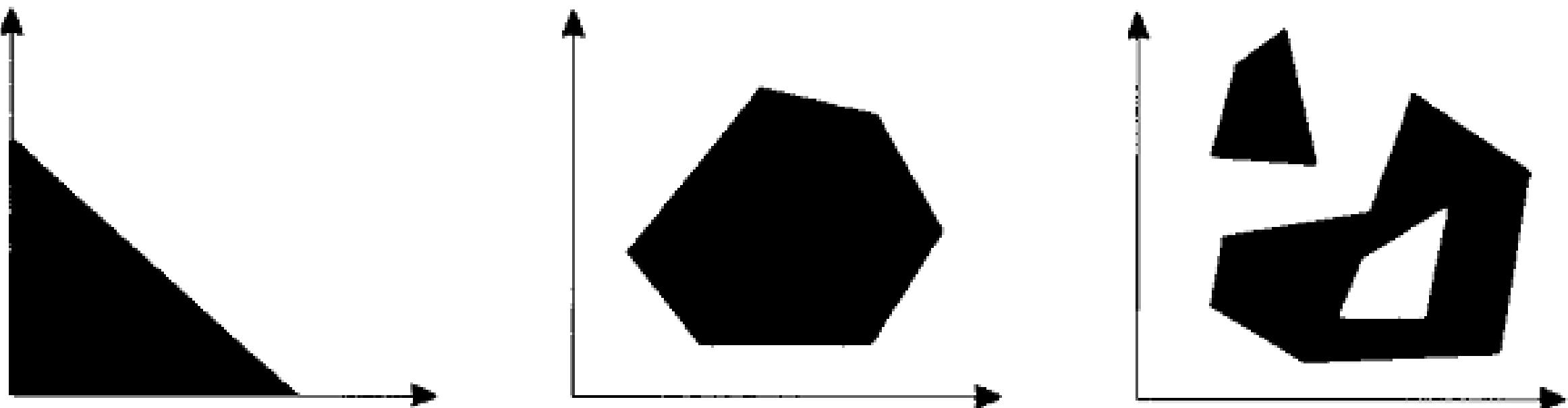
- Supera i limiti del Perceptron usando strati di neuroni *hidden* (interni) che realizzano una rappresentazione interna dell'input più complessa, perché individua regioni arbitrarie intersecando iperpiani (1) nel iperspazio dei valori input.
- Nell'esempio i Perceptron B e C creano ciascuno un semipiano individuato dalle rette e il Perceptron A interseca i due semipiani individuando la parte di piano tra le rette.

(1) IPERPIANO: uno spazio di dimensione $n-1$ in uno spazio euclideo di dimensione n . Spazio euclideo spazio geometrico in cui è possibile definire concetti come distanza, angoli, e operazioni di vettori.

10

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: REGIONE DELIMITATA
MULTI LAYER PERCEPTRON (MLP)

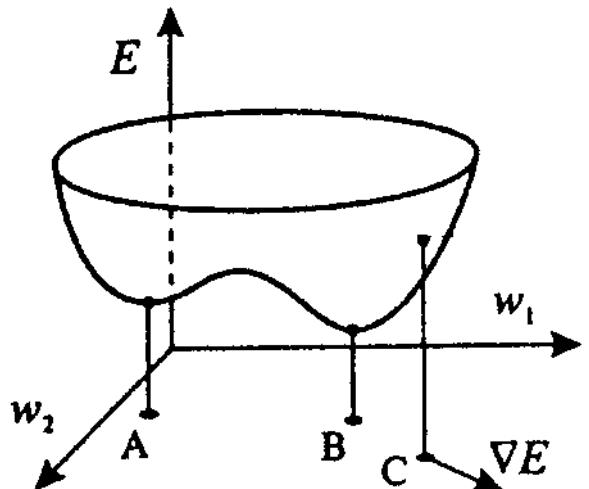


Usando la funzione di attivazione a gradino, a sinistra la regione delimitata da un Perceptron, al centro la regione delimitata da MLP a 1 livello hidden, a destra la regione delimitata da MLP con 2 livelli hidden che può delimitare aree arbitrarie.

11

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: FUNZIONE DI ERRORE

- Esprime la differenza fra l'output della rete y e l'output desiderato y' nell'apprendimento. $E(w)$ perché l'output è ottenuto in base al valore dei pesi; quindi E è derivabile nei pesi w e occorre trovare il vettore w che rende minimo l'errore. Esistono molte formulazioni di E .
- $E(w) = \sum_{i=1}^c ((y_i - y'_i)^2)$
- E è non lineare: gli algoritmi cercano un minimo nella sua superficie con modifiche di w in base al gradiente di E .
- A sinistra c'è una superficie di E quadratica nei pesi w_1 e w_2 ; A e B sono minimi e C punto di calcolo del gradiente locale della superficie, si va nella direzione opposta al gradiente.
- Può non convergere verso il minimo assoluto B ma verso un minimo locale A; in alcuni punti il gradiente è nullo.

«Gradiente» si riferisce a un vettore che rappresenta la direzione e l'entità della massima variazione di una funzione scalare in un punto specifico dello spazio. In altre parole, indica la direzione in cui la funzione cresce più rapidamente.

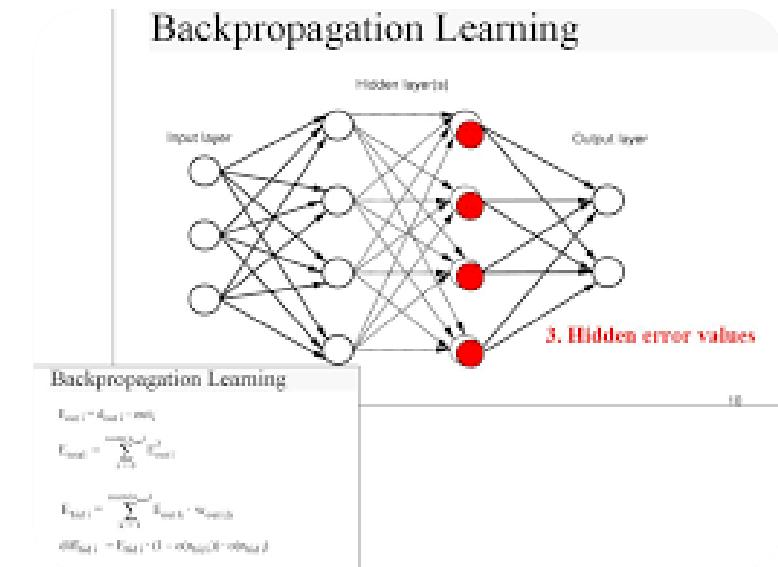
12

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: FUNZIONI DI BACK-PROPAGATION

- Ogni neurone computa con z_j l'attivazione di unità che manda connessione a unità j e w_{ji} il peso tra unità j e i , la sommatoria è su tutte le unità connesse all'unità j .
$$a_j = \sum_i w_{ji} z_i$$
- Usando funzione attivazione g si ha l'attivazione z_j dell'unità j
Se le z_j indicano unità input allora sono $z_j = x_j$, se le unità j sono di output l'attivazione è denotata $y_k = z_j$. È l'errore per n pattern input.

Tecnica di ottimizzazione utilizzata per regolare i pesi delle connessioni tra i neuroni in modo da minimizzare l'errore di predizione del modello.



13

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: ALGORITMO DI BACK-PROPAGATION

1. si esegue un'epoca di apprendimento, cioè si attuano per tutte le coppie (input, output) significative del training set i passi:
2. prende una coppia e calcola la risposta della rete per quell'input; il calcolo procede dal livello input verso il livello output calcolando l'attivazione di tutte le unità, quindi propaga in avanti l'errore
3. calcola l'errore E tra l'output della rete e output della coppia e calcola δ_i delle unità output
4. propaga all'indietro l'errore verso il livello di input, calcolando δ_i per ogni unità hidden
5. variazione dei pesi
6. ripete dal passo 1 fino a terminare le coppie
7. calcola l'errore globale e se è ancora alto si ripete l'epoca di apprendimento.

14

Apprendimento in rete

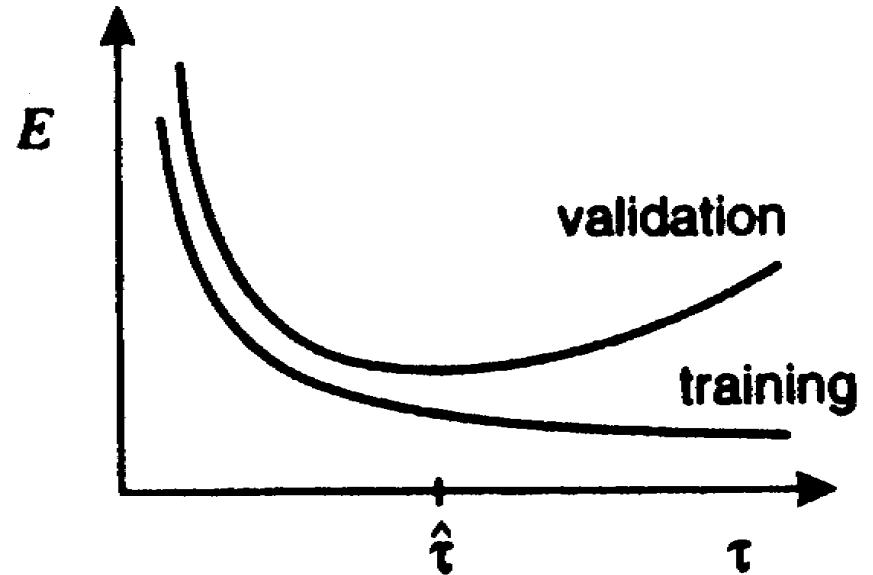
APPRENDIMENTO SUPERVISIONATO: PROBLEMI DEL BACK-PROPAGATION

- ✓ Algoritmo **lento** che può finire intrappolato in un minimo di E credendo di aver trovato il valore ottimo dei pesi che rende minima la E
- ✓ Dipendenza dal valore iniziale dei pesi w ; può capitare di cominciare da un punto della superficie di errore già molto vicino a un minimo locale
- ✓ scelta critica di η : troppo piccolo crea apprendimento lento, troppo grande crea oscillazioni
- ✓ esistono tecniche che aumentano la velocità di convergenza verso il minimo assoluto di E e che smorzano le oscillazioni aggiungendo alla formula di variazione del peso un altro parametro detto **momentum** (*inerzia nel cambiamento dei pesi*).

15

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: OVERFITTING



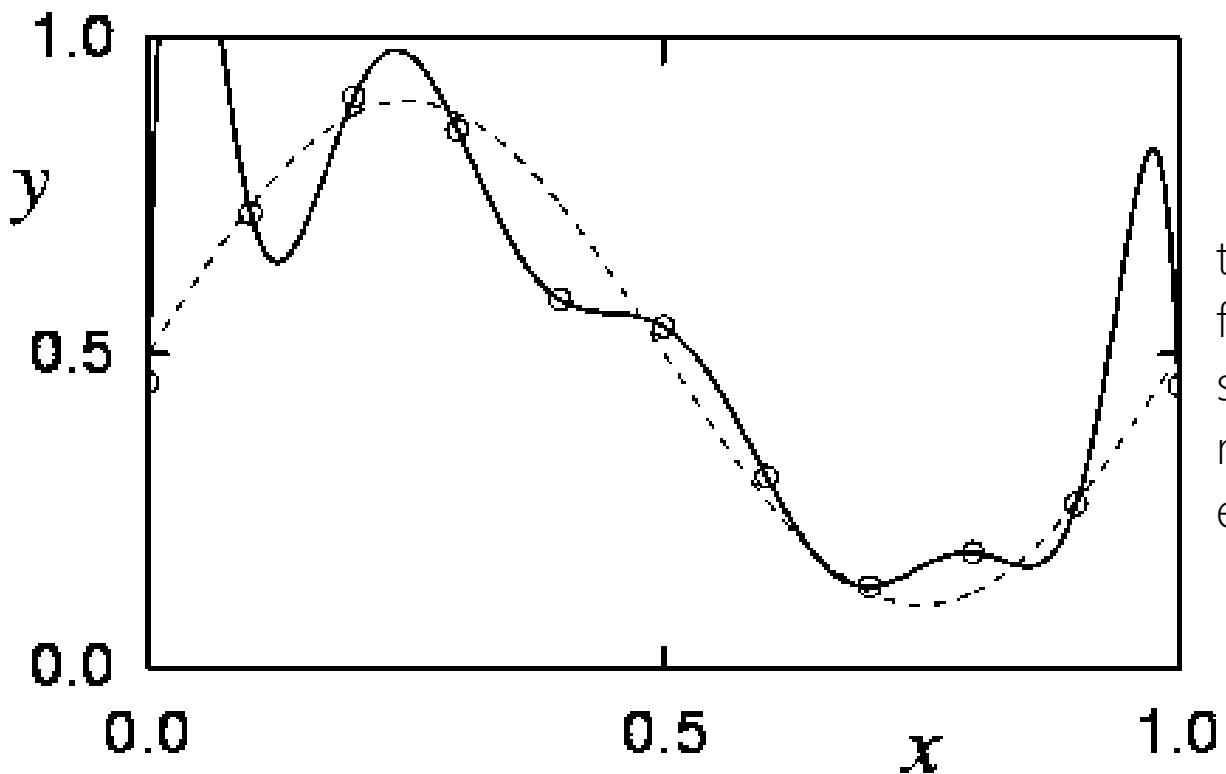
Si adatta troppo bene ai dati di addestramento, al punto da perdere la capacità di generalizzare correttamente a nuovi dati o a situazioni diverse.

- La rete neurale deve avere capacità di comprensione del modello statistico dei dati, non memorizzare i soli dati del training set.
- Così può generalizzare, cioè rispondere esattamente a input non in training set. La tecnica migliore per evitare ciò è *l'early stopping*. Nell'apprendimento l'errore E tende a 0 perché la rete sta imparando meglio. Usando un validation set di coppie non usate nel training set, si misura la E e si crea la seconda curva che tende a un minimo e poi τ ricresce: da quel punto la rete sta imparando il training set e non il suo $\hat{\tau}$ modello statistico. Si ferma lo apprendimento al minimo di E rispetto al validation set.

16

Apprendimento in rete

APPRENDIMENTO SUPERVISIONATO: OVERFITTING



I cerchi indicano i punti che formano il training set, la curva in tratteggio è la funzione da apprendere, la curva continua è la funzione appresa dalla rete neurale con errore nullo di training senza uso di *early stopping*. Evidente la differenza tra le curve nei punti diversi dai cerchi, perché la rete neurale ha imparato esattamente solo i cerchi e fa grossi errori negli altri punti.

17

Apprendimento in rete

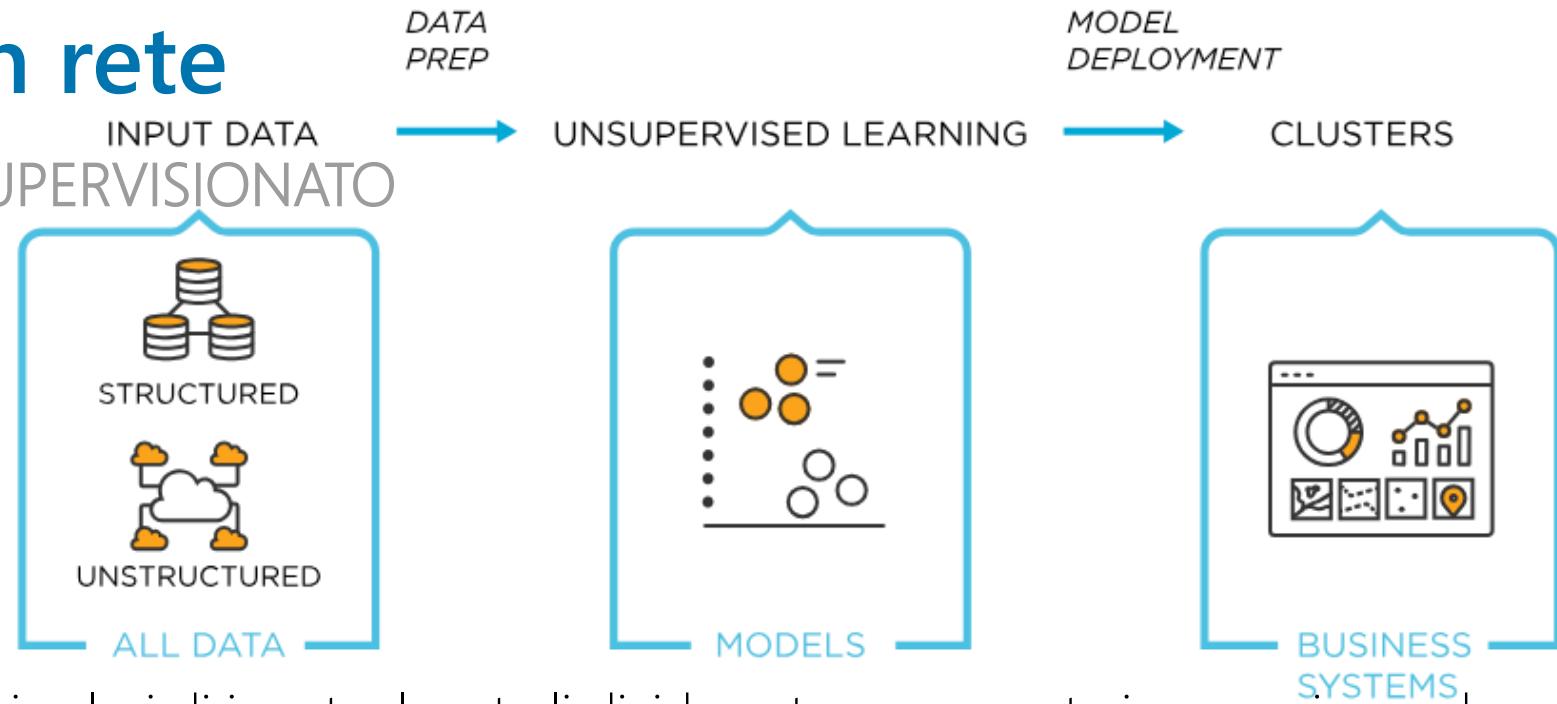
APPRENDIMENTO SUPERVISIONATO: VALUTARE LE PRESTAZIONI

- Per misurare le prestazioni di una rete neurale dopo l'apprendimento, si crea il test set formato da coppie non usate per i *training* e *validation* set. In genere il test set è un terzo del training set ed è composto da input critici su cui la risposta della rete deve essere buona, altrimenti si butta via la rete.
- Per ogni coppia del test set:
 - calcolare la risposta della rete all'input
 - calcolare l'errore dato dalla differenza tra output rete e output di coppia
- L'errore totale è dato dalla somma degli errori avuti per ogni coppia.
- *Secondo il tipo di problema, si usano varie tecniche statistiche per decidere se usare o meno la rete creata; in genere si accetta una rete se sul test set ha mostrato un errore inferiore al 20-25% delle coppie.*

18

Apprendimento in rete

APPRENDIMENTO NON SUPERVISIONATO



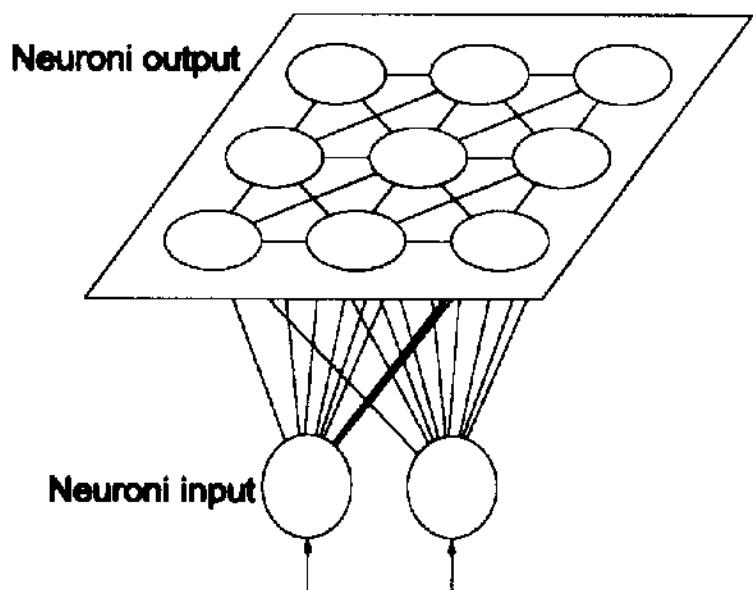
- Alla rete vengono presentati solo i valori di input e la rete li divide autonomamente in gruppi usando misure di similarità, senza usare confronti con output noti, e cercando di mettere input simili nello stesso gruppo. E' un apprendimento autonomo e non c'è controllo esterno sull'errore. Adatto per ottimizzare risorse e se non si conoscono a priori i gruppi in cui dividere gli input.
- Modelli che usano questo apprendimento:
 1. Kohonen
 2. Hopfield

18

Apprendimento in rete

APPRENDIMENTO NON SUPERVISIONATO: SELF ORGANIZING MAPS (SOM)

- Ideate da Tuevo Kohonen nel 1982 ispirandosi alla topologia della corteccia del cervello. Tengono conto delle connessioni tra neuroni e dell'influenza che può avere un neurone sui suoi vicini: i neuroni vicini a neuroni attivi rinforzano i legami, mentre a quelli che si trovano ad una data distanza vengono indeboliti i legami.
- Una rete **SOM** è composta da un livello input ed un livello di competizione che sono localizzati su una griglia ad una o due dimensioni e con diverse forme. Ciascun neurone di input è connesso a tutti i neuroni della griglia; ogni neurone di output ha il vettore dei pesi con le stesse dimensioni del vettore di input.

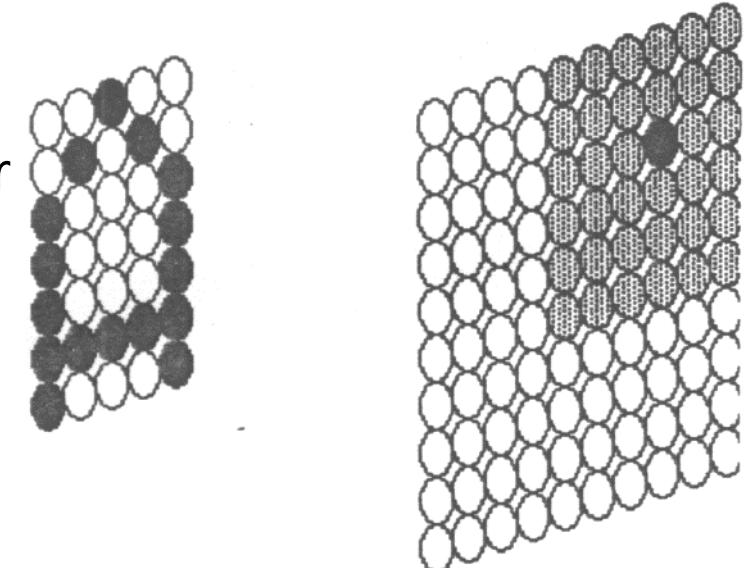


19

Apprendimento in rete

APPRENDIMENTO NON SUPERVISIONATO: INTERPRETAZIONE DELLA MAPPA

- La configurazione finale dei pesi dei singoli neuroni permette di suddividere gli elementi forniti in ingresso in cluster (raggruppamenti di oggetti simili) che di quegli elementi rappresentano una classificazione.
- Occorre localizzare sulla mappa i neuroni attivi ed associarli con gli input presentati. Alla fine dell'addestramento, ogni lettera attiverà una differente combinazione di neuroni contenente il neurone vincitore per quella lettera.
- Il neurone vincente sarà attivato anche da sottomissione della lettera parziale o con rumore.



20

Apprendimento in rete

APPRENDIMENTO NON SUPERVISIONATO: APPRENDIMENTO IN SOM

1. Creare un training set con molti vettori di input significativi; stavolta non ci sono output con cui paragonare la risposta della rete.
2. Riepilogando per ogni vettore di input:
 - ogni neurone di output riceve le componenti del vettore input e ne calcola la distanza euclidea dal suo vettore di pesi
 - il neurone di output, che ha la minima distanza euclidea dall'input, si attiva e dà una risposta maggiore, modifica poi i suoi pesi e quelli dei neuroni vicini per avvicinarli all'input.
3. Così input simili attiveranno neuroni vicini.
4. Occorre stabilire quali sono i neuroni vicini e come modificare i loro pesi.

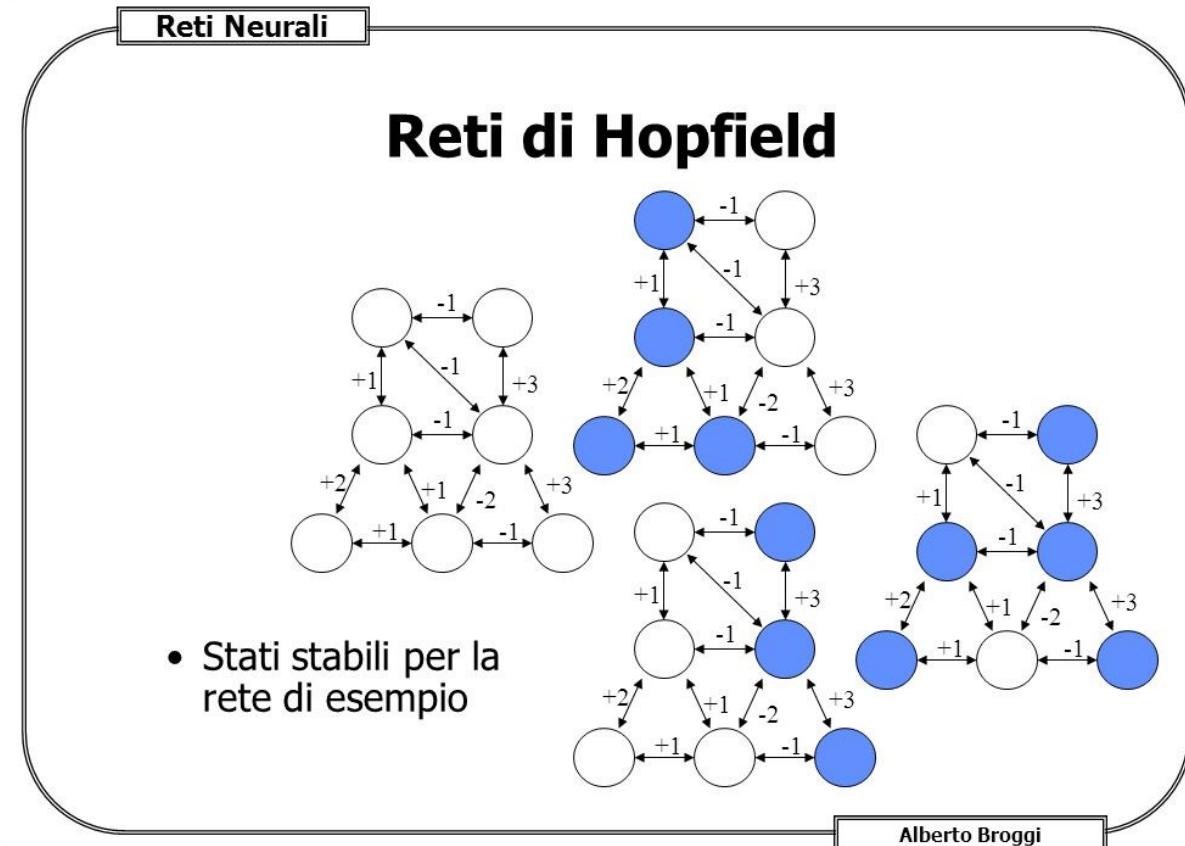
20

Apprendimento in rete

APPRENDIMENTO NON SUPERVISIONATO: RETI DI HOPFIELD

- John Hopfield nel 1982 propone una rete per memorizzare informazioni. E' composta da neuroni completamente connessi con funzione di attivazione. Ogni neurone è nodo di ingresso e di uscita, ha stato attivo o disattivo, ha funzione di attivazione a gradino
- Alla rete è associata una funzione energia da minimizzare durante la sua evoluzione con una successione di stati, fino a raggiungere uno stato finale stabile corrispondente al minimo della funzione energia.

John Hopfield (fisico) pubblica un articolo di un modello matematico: una rete che permette "l'emergere spontaneo di nuove capacità computazionali dal comportamento collettivo di un gran numero di semplici elementi d'elaborazione".



Una rete di Hopfield, è modello di rete che simula le capacità del cervello umano di ricordare le cose o di ricostruire le immagini distorte.

21

Apprendimento in rete

APPRENDIMENTO NON SUPERVISIONATO: RETI DI HOPFIELD

PROPRIETÀ

- ✓ rappresentazione distribuita su tutti i neuroni
- ✓ adatta per realizzare una memoria indirizzabile per contenuti: per recuperare un'informazione basta produrne una versione parziale o qualche indizio
- ✓ adatta per compiti di ottimizzazione vincolata, quando il problema può essere espresso in termini di una funzione obiettivo da trasformare in una funzione energia
- ✓ recupero efficiente di informazioni anche con danni alla rete

22

Apprendimento in rete

APPRENDIMENTO HEBBIANO



Donald Hebb (Psicologo 1904-1985)

L'attivazione simultanea di due neuroni connessi determina la modifica dei pesi delle connessioni che li uniscono, in modo da aumentare la probabilità che uno dei due neuroni si attivi quando l'altro si attiva. Così si rafforza la connessione tra due neuroni molto usati e si indebolisce la connessione tra due neuroni poco attivati contemporaneamente. (anni '40)

«Due neuroni che scaricano assieme si potenziano reciprocamente» → introdusse inoltre il concetto di assemblamento cellulare, per definire un gruppo di neuroni che costituiscono un'unità di elaborazione, e ipotizzò che la combinazione delle connessioni tra gruppi di neuroni costituisse l'algoritmo, in continua mutazione, che dettava al cervello la risposta ai diversi stimoli.

La teoria del campo (matrice gestaltica), nella quale l'attenzione ai processi percettivi, viene estesa da Kurt Lewin fino ai processi sociali e di gruppo.

Intelligenza Artificiale

Cos'è davvero, come funziona, che effetti avrà



Intelligenza Artificiale

Cos'è davvero, come funziona, che effetti avrà

«Timeline» del libro



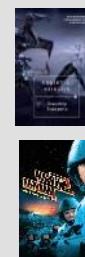
Punto di svolta
AI Life Wide

“When one teaches two learn” (“Quando uno
insegna, due imparano”)

Robert Heinlein

(Scrittore di fantascienza... «Fanteria nello spazio»
Starship Troopers)

Tecnocrazia (“il governo dei cervelloni”),



Roboetica

Max Tegmark, Vita 3.0. Essere umani
nell'era dell'intelligenza artificiale

