

Le reti neurali

Le reti neurali, dal nome affascinante quanto misterioso, sono una delle aree di ricerca più interessanti tanto da essere oggetto di numerosi studi. Spieghiamo con questo tutorial i principi che stanno alla base di questo sistema d'elaborazione dell'informazione

ROBERTO ACCOMANDO

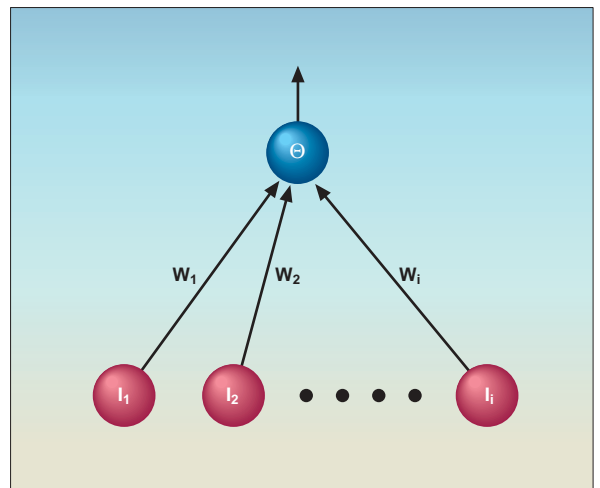
Le reti neurali (o neuronali) sono una delle aree più interessanti della ricerca in Intelligenza Artificiale. Nonostante la loro ideazione risalga ai primi anni '60, solo di recente si sono avuti notevoli risvolti applicativi e teorici. L'idea che sta alla base delle reti neurali è quella di 'imitare' il funzionamento del cervello, utilizzando un certo numero di semplici elementi di calcolo fortemente interconnessi tra loro per elaborare dati: i neuroni.

Perché funzionano le reti neurali?

Volendo spiegare i principi fondamentali che reggono il loro funzionamento dobbiamo necessariamente arrivare a definire le reti neurali. Queste sono a tutti gli effetti dei modelli matematici che tentano di emulare il sistema nervoso centrale umano. Lo scopo finale delle reti neurali artificiali è realizzare i meccanismi di apprendimento del cervello umano, facendo in modo che la rete interagisca con l'ambiente esterno senza l'intervento umano, oltre quello della creazione.

Lo scopo di questi complessi metodi matematici è il tentativo di penetrare le leggi della natura meglio delle tecniche di calcolo tradizionali.

Si basano sul concetto che, in determinate situazioni, è possibile far 'apprendere' a un apparato matematico determinate leggi che non si conoscono a priori, semplicemente facendogli analizzare un elevato numero di casi reali. In parole più semplici, le reti neurali rappresentano



Il neurone artificiale è molto simile per struttura a quello naturale. Infatti è composto da diverse sinapsi (che ricevono i dati di input) collegate a un unico centro, dal quale si diparte un assone che invia dati di output

un potente, e non ancora completamente compreso, mezzo statistico. Dopo aver processato una mole notevole di dati nella fase cosiddetta di apprendimento o assimilazione, queste reti propongono una linearizzazione del problema cui si vuole dare una soluzione per mezzo di una formula matematica che lega tra loro le diverse variabili, espressione del problema stesso.

Tale linearizzazione esprime la probabilità statistica che

ha un nuovo caso mai esaminato prima di collocarsi in un'area di 'soluzione' già verificatasi e validata nella fase di addestramento.

Nel dettaglio

Come accennato precedentemente, le reti neurali utilizzano per il loro funzionamento dei semplici elementi di calcolo chiamati (per motivi storici) neuroni. In una rete neurale alcuni neuroni sono utilizzati come ingressi, altri come uscite, altri ancora (non sempre presenti) non sono connessi con l'esterno e svolgono una funzione di 'calcolo' e vengono chiamati neuroni 'nascosti' (hidden). Una rete neurale è dunque costituita da diversi neuroni connessi tra di loro.

Ogni neurone artificiale valuta l'intensità di ogni input, somma i diversi input opportunamente moltiplicati con dei coefficienti di peso e confronta il risultato con una soglia opportuna. Se il valore ottenuto risulta maggiore della soglia il neurone produrrà un'uscita altrimenti rimarrà silente.

Estendendo questo principio di funzionamento a elaborazioni più complesse della somma, e combinando diversi neuroni tra di loro si viene così a realizzare una rete neurale.

La grande differenza rispetto ad altri sistemi di calcolo consiste nel fatto che la rete non è programmata per svolgere un determinato compito, ma viene 'addestrata' mostrandole degli esempi (sotto forma di coppie ingresso-uscita), la rete impara così ad associare ad ogni ingresso un'uscita e anche in una certa misura a generalizzare quello che ha imparato per ingressi che non le erano stati presentati in fase di addestramento.

Così per ottenere l'uscita voluta è sufficiente che l'ingresso 'somigli' a quello originario.

Questa proprietà rende le reti neurali adatte a tutti i compiti che hanno a che fare con il riconoscimento e classificazione di input (caratteri, immagini, suoni...).

Le capacità della rete sono determinate sia dal numero di neuroni, sia soprattutto dalla topologia (cioè da come sono collegati tra loro i neuroni).

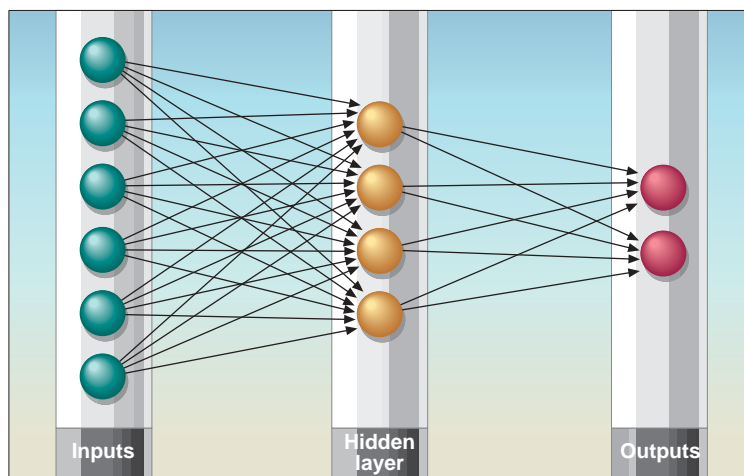
Esistono infinite topologie possibili, ma le più utilizzate si possono schematizzare in due classi: le reti alimentate in avanti (feed forward) e le reti ricorrenti (o ricorsive). Nelle prime non esiste nessun collegamento che riporti in qualche modo le uscite in ingresso, questo le rende abbastanza semplici da studiare e addestrare ma non gli permette di mantenere uno 'stato' interno, cioè una memoria di quello che è successo nel passato.

Questo tipo di rete è il più studiato e utilizzato. Le reti ricorrenti hanno delle capacità più estese però non sono state in generale ancora ben comprese e non si conosce sempre un metodo adatto per addestrarle.

Tra i numerosi tipi di reti neurali esistenti sono da ricordare inoltre le reti di Hopfield e le self organizing nets di Kohonen. Le prime sono delle reti ricorrenti formate da un solo strato di neuroni. Ciascuna uscita è collegata agli ingressi di tutti gli altri neuroni.

Questa rete viene addestrata utilizzando delle particolari formule lineari e può funzionare come rete auto associativa. Dopo la fase di allenamento la rete è in grado di riconoscere una configurazione che ha imparato e di presentare in uscita la versione originale.

Le reti di Kohonen invece costituiscono un esempio di reti che sono in grado di apprendere senza bisogno di



Una semplice rete neurale in cui sono presenti neuroni d'ingresso, d'uscita e nascosti specificatamente pensati per eseguire le funzioni di calcolo

supervisione, cioè hanno la capacità di ricevere diverse configurazioni in ingresso e di classificarle per somiglianza. Anche in questo caso le formule usate nell'addestramento sono diverse da quelle usate per i perceptron.

Le principali applicazioni

I principali campi d'applicazione delle reti neurali possono essere accumulati da una caratteristica peculiare. Infatti, sono tutti quelli dove l'analisi statistica di tutte le variabili di un problema risulta difficoltosa o dispendiosa in termini di calcolo, ma soprattutto dove non sia chiaro a priori quali relazioni deterministiche esistano tra le diverse variabili che caratterizzano il problema.

I principali campi in cui finora le reti neurali hanno dimostrato di poter dare validi contributi sono le applicazioni legate alle previsioni meteorologiche, i modelli di risposta, le indagini di rischio o le più comuni applicazioni OCR (Optical Character Recognition).

Tutti questi settori (si pensi ad esempio alle discipline legate alle previsioni meteorologiche) sono caratterizzati da una grossa quantità di dati sperimentali che è possibile elaborare cercando una qualche 'regola' che possa nascere da una continuativa attività d'osservazione di un fenomeno reale. ■