MAPPE DEL LABIRINTO
UNA SERIE DI PHANTOM MAZE — AI & LANGUAGE LAB

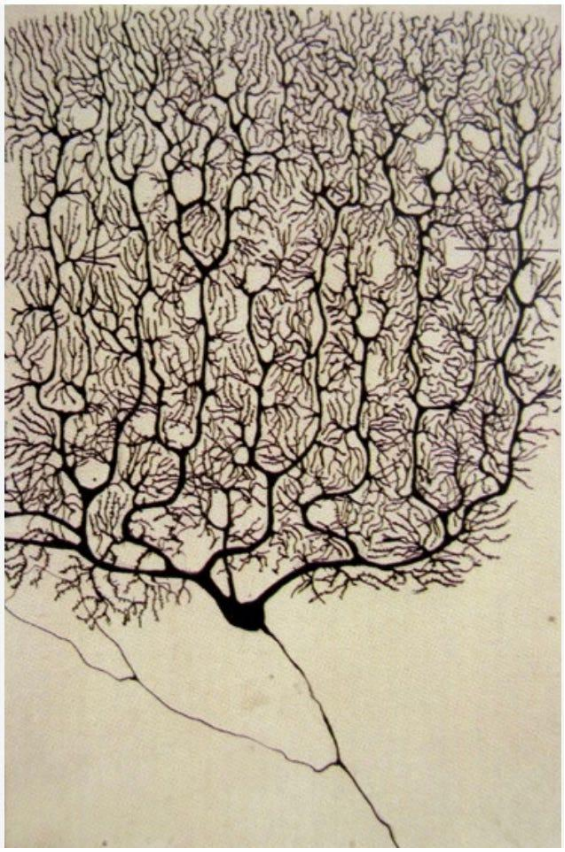
SAGGIO 17

Reti neurali: la ragnatela senza ragno

Claudia Marsico · Hernán Inverso

DOI 10.5281/zenodo.23575286

Che cosa fa davvero un neurone artificiale



La rete neurale è la struttura matematica che sta sotto gli attuali modelli linguistici, e il suo nome porta con sé la metafora più carica di tutta l'intelligenza artificiale. Ripetiamo così tanto "rete neurale", "neuronale", "neurale" che suona come qualcosa di ovvio, e la parte più strana ci sfugge. In che senso è costruita come un cervello di silicio? Conviene guardare con attenzione, perché è un terreno in cui tecnica e licenza poetica si intrecciano in modo diabolico.

Il termine "rete" si porta dietro la sua storia di pescatori e cacciatori, con la loro maglia di fili annodati. Da lì viene "reticolo", la "piccola rete", e anche "retina", perché in fondo all'occhio si vedeva una

specie di rete. Anche vedere è, almeno nell'etimologia, catturare il mondo in una rete. "Neurone" non è lontano. Il greco *neuron* non aveva nulla a che fare con il cervello, che chiamavano *enképhalos*, e ancora meno con le cellule del cervello, di cui nessuno sapeva l'esistenza. *Neuron* era il tendine, la corda tesa del muscolo e più tardi il nervo. Quando nel 1891 Wilhelm von Waldeyer fece la sintesi dei lavori del settore, scelse "neurone" per questa unità del sistema nervoso. Santiago Ramón y Cajal avrebbe suggerito anni dopo che il nome orecchiabile fu tutto il contributo di Waldeyer. Non è poco, dopotutto.

Le maglie di nodi ci parlano da molto lontano. Nella tradizione formatasi attorno al *Sutra dell'ornamento fiorito*, l'*Avatamsaka Sutra*, comparve l'immagine della rete del dio Indra, che il patriarca Dushun, nella Cina dei secoli VI-VII, trasformò in un insegnamento capitale sulla compenetrazione di tutte le cose. Immaginiamo una rete infinita tesa sul cosmo che in ogni nodo porta una gemma. Ogni gemma riflette tutte le altre, e in ciascuno di quei riflessi sono a loro volta riflesse tutte le altre, all'infinito. Nessun nodo ha senso da solo. Ciascuno è ciò che è per l'intera trama che si rispecchia in lui. È un'intuizione bellissima sulla vita del significato nelle relazioni, nella rete che ripete la rete in ogni punto. Una rete neurale artificiale funziona un po' così, perché nessun nodo isolato vuol dire nulla fuori dalla trama di riflessi e di punti. Allo stesso tempo è molto diversa, dato che la rete di Indra è un mondo di gemme che si contemplano, mentre questa è una maglia di somme e moltiplicazioni che nessuno contempla.

Alla fine del XIX secolo la teoria dominante, quella di Camillo Golgi, intendeva il sistema nervoso come un reticolo continuo, un'unica maglia di fibre fuse senza soluzione di continuità in cui tutto toccava tutto. Contro quell'idea lavorò, con un microscopio e un talento da disegnatore, lo spagnolo Santiago Ramón y Cajal, che abbiamo già nominato. Verso il 1888, dopo aver studiato tessuti colorati, respinse l'idea della maglia continua e affermò che il cervello è una moltitudine di cellule separate, ciascuna con il suo corpo, i suoi rami e il suo terminale, che si avvicinano senza fondersi, comunicando attraverso spazi che più tardi si sarebbero chiamati sinapsi. Cajal vedeva in quel folto, secondo le sue stesse parole, una foresta di alberi dai rami protesi. Lo disegnò migliaia di volte con una bellezza che adorna i musei. La discordia fu così tenace che, quando nel 1906 diedero a entrambi lo stesso premio Nobel, Golgi approfittò del suo discorso per continuare a difendere la rete continua che stava perdendo la partita, e Cajal commentò, con amarezza, che il destino era stato crudele a legarlo come un gemello siamese a un avversario scientifico di carattere tanto opposto.

L'episodio non finì lì. Cajal vinse la battaglia anatomica, ma la "rete" sopravvisse. Il cervello non era il reticolo continuo immaginato da Golgi, ma quando mezzo secolo dopo comparvero le reti neurali artificiali, conservarono entrambe le immagini insieme: unità separate, come i neuroni di Cajal, organizzate in una trama di connessioni, solo che quei "neuroni" e quelle "connessioni" diventarono astrazioni matematiche.

Nel 1943 il neurofisiologo Warren McCulloch e il logico Walter Pitts proposero un modello matematico estremamente semplificato di una cellula nervosa. L'idea di fondo era ridurre la cellula nervosa a un'unità logica: riceveva segnali da altre unità e, se quelli eccitatori raggiungevano una certa

soglia e non interveniva un segnale inibitorio, produceva un'uscita; altrimenti no. Della complessità chimica ed elettrica di un neurone reale restava appena un piccolo dispositivo logico.

Le reti attuali hanno conservato la struttura generale, ma hanno cambiato molto il funzionamento. Un'unità riceve numeri, li moltiplica per i suoi pesi, li somma, può spostare il risultato verso l'alto o verso il basso con un parametro chiamato bias, e infine fa passare quel totale attraverso una regola che si chiama funzione di attivazione. Quest'ultima operazione decide se la rete servirà a qualcosa. Una delle attivazioni più semplici e famose, ReLU, fa questo: se il numero che arriva è negativo, lo sostituisce con zero; se è positivo, lo lascia passare tale e quale. Niente di più. I modelli attuali usano varianti più morbide, ma l'idea è sempre quella di inserire nella catena di moltiplicazioni e somme una regola che non sia proporzionale, per rompere ciò che i matematici chiamano linearità. Se raddoppio ciò che entra, raddoppia ciò che esce, e l'effetto di due cose insieme è la somma dei loro effetti separati. Comprare il doppio delle mele costa il doppio. Questo è lineare. Il mondo, però, è pieno di cose che non funzionano così. Raddoppiare la dose di un farmaco non raddoppia il miglioramento e può, al contrario, uccidere il paziente. Un po' di sale migliora lo stufato e il doppio lo rovina. Soglie, saturazioni, cambiamenti di stato sfuggono alla proporzionalità.

Da qui nasce il problema di impilare strati senza attivazioni. Pensiamo a cambiare valuta senza banche né commissioni di mezzo: pesos in dollari, dollari in euro, euro in yen. Ogni passaggio è una moltiplicazione per un numero, e la catena completa, per quanti anelli abbia, equivale sempre a un'unica moltiplicazione, quella che converte direttamente i pesos in yen. Aggiungere intermediari non apre né chiude nulla. Con gli strati succede lo stesso: se ciascuno si limitasse a trasformazioni lineari o affini, senza attivazioni, cento strati si potrebbero far collassare in uno e avremmo una rete di cento piani con la capacità di uno solo. L'attivazione impedisce quel collasso e fa sì che il secondo piano possa vedere qualcosa che il primo non poteva, e che la profondità serva a qualcosa.

Il linguaggio ci chiede di farlo perché il senso di una frase non è la somma dei sensi delle sue parole. Un "non" davanti rovescia tutto ciò che segue, "non è male" non è la sottrazione di "è" e "male". Per non parlare dell'ironia, che significa il contrario di ciò che dice. "Salvo che", "a meno che", "ma" sono cerniere, non addendi. Per tutto questo una catena strettamente lineare non può rappresentare quel genere di interazioni e ha bisogno di soglie e di ribaltamenti non lineari.

Immaginiamo punti rossi e blu disegnati su un foglio, con il compito di separarli con una sola linea retta. Se i rossi stanno a sinistra e i blu a destra è facile, ma se i rossi formano un anello attorno ai blu non esiste nessuna retta che li separi. Servono una o più curve. Gli strati con attivazioni permettono alla rete di piegare quel confine tante volte quante le occorrono. Una rete "profonda" è semplicemente una rete in cui molte di quelle trasformazioni si succedono l'una all'altra, e ciascuna può piegare un po' di più ciò che ha ricevuto dalla precedente. La profondità del *deep learning* conta quegli strati, privi di misteri insondabili.

Che cos'è, allora, una rete neurale? Ricordiamo che una funzione è una regola che prende un ingresso e restituisce un'uscita. Per esempio, una tabella di conversione delle temperature riceve gradi Celsius e restituisce Fahrenheit. Una rete neurale fa qualcosa del genere, ma su scala smisurata. Riceve molti

numeri, li fa attraversare strati successivi di operazioni e restituisce altri numeri. Questa funzione è inoltre speciale perché contiene un'enorme quantità di parametri regolabili, sicché cambiare quei valori cambia la relazione tra ingresso e uscita.

La portata di quella flessibilità fu formalizzata alla fine degli anni Ottanta nei cosiddetti teoremi di approssimazione universale. Pensiamo a una circonferenza fatta con tre segmenti. Somiglia poco a un cerchio, con venti gli somiglia parecchio e con mille quasi non distinguiamo la differenza. Nel caso delle reti, combinando abbastanza pezzi semplici, possono approssimare con precisione relazioni continue molto più complicate. Ora, il teorema garantisce soltanto che, sotto certe condizioni, quella configurazione esiste, non che sia facile da trovare. Sapere che una qualche combinazione di milioni o miliardi di parametri produce il comportamento che cerchiamo non ci dice quale sia. Come conoscere la combinazione della cassaforte? A questo serve l'addestramento, che consiste nel modificare via via i parametri fino a trovare una configurazione che funzioni bene.

La storia di questa idea è un giro sulle montagne russe. Dopo il neurone di McCulloch e Pitts venne, nel 1958, il percettrone di Rosenblatt, quell'apparecchio dalle manopole motorizzate che abbiamo nominato a proposito di parametri e pesi. Interessa qui perché aggiunse un modo di regolare quei pesi a partire da esempi, qualcosa di nuovo rispetto al modello del 1943. Sul momento ci fu euforia, ma nel 1969 Marvin Minsky e Seymour Papert mostrarono con rigore i limiti dei percettroni a un solo strato. L'esempio che finì per diventare celebre fu XOR, l'"o esclusivo", che è vero quando accade una cosa o l'altra ma non entrambe insieme. Le sue quattro combinazioni di valori di verità non si possono separare con un unico confine lineare. Il risultato alimentò l'idea che le reti neurali fossero in un vicolo cieco, proprio ai tempi del primo inverno dell'intelligenza artificiale, che ebbe molte altre cause. Fatto sta che questa linea di ricerca ne uscì malconcia.

La primavera delle reti tornò a metà degli anni Ottanta, quando David Rumelhart, Geoffrey Hinton e Ronald Williams resero popolare la retropropagazione, un procedimento che permetteva di usare l'errore osservato in uscita per calcolare come convenisse modificare i pesi situati negli strati interni. Le reti multistrato potevano finalmente essere addestrate in modo pratico su problemi che il percettrone semplice non poteva rappresentare. Mancavano ancora i dati e le macchine, che arrivarono insieme nel 2012, quando una rete profonda chiamata AlexNet, opera di Alex Krizhevsky, Ilya Sutskever e dello stesso Hinton, vinse con un distacco schiacciante una grande competizione di riconoscimento di immagini e insegnò al mondo intero che la vecchia idea, con scala sufficiente, finalmente funzionava. Da lì ai modelli linguistici di oggi la strada si fa riconoscibile.

Soffermiamoci un momento su ciò che la macchina non è. Non è un cervello. Il "neurone" artificiale sta a un neurone vero come un omino stilizzato sta a un corpo umano. Ne condivide il nome e poco altro. Una cellula nervosa reale è un sistema elettrochimico di una complessità che non abbiamo ancora finito di capire; l'unità artificiale è una somma con una curva sopra, un'astrazione matematica ispirata alla lontana da alcuni tratti della prima. Le reti artificiali non riproducono l'anatomia, la dinamica né i meccanismi di apprendimento del cervello. La parentela è soltanto storica e metaforica. Prenderla alla lettera è la fonte di moltissime confusioni che circolano sull'intelligenza artificiale, come

l'idea che regolare pesi equivalga al modo in cui apprende un cervello o che da qualche parte nella rete debba esserci un neurone che custodisce una determinata idea. Quelle inferenze non discendono dal nome.

Non regge nemmeno l'abusato confronto numerico. Un cervello umano ha circa ottantasei miliardi di neuroni e dell'ordine di centomila miliardi di sinapsi; i grandi modelli linguistici si contano in centinaia di miliardi di parametri. La vicinanza delle cifre non significa nulla, perché un parametro non equivale a una sinapsi né un'unità artificiale a una cellula nervosa. L'analogia non migliora perché i numeri si avvicinano, anche se è interessante la nostra compulsione a confrontarli.

La rete di Indra era fatta di gemme che si riflettevano l'una nell'altra. Questa è fatta di operazioni numeriche la cui efficacia dipende anch'essa da relazioni distribuite, e finisce per riconoscere immagini, tradurre lingue o sostenere una conversazione senza bisogno di un cervello nascosto nei nodi. Una ragnatela naturale presuppone un ragno che l'ha tessuta e che occupa un posto rispetto a essa. In una rete neurale ciò che la rete sa non vive in nessun pezzo isolato, né c'è un centro. È, in questo senso, una ragnatela senza ragno.

COME CITARE

Marsico, Claudia e Hernán Inverso. 2026. "Reti neurali: la ragnatela senza ragno". *Mappe del labirinto*, saggio 17. Phantom Maze — AI & Language Lab. https://doi.org/10.5281/zenodo.23575286

Mappe del labirinto è una serie di Phantom Maze, AI & Language Lab. La collezione vive su **phantommaze.cloud**. L'iscrizione è gratuita.

Alcune delle idee che inquadrano questo testo sono sviluppate in *Il fantasma senza ego. Una mappa nel labirinto dell'intelligenza artificiale*, di Hernán Inverso e Claudia Marsico (Phantom Maze Press, 2026; ISBN 978-987-3729-19-5).

PHANTOM MAZE

Leggi questo saggio online · Tutti i saggi · Il libro

Fonti. Santiago Ramón y Cajal, *Textura del sistema nervioso del hombre y de los vertebrados: estudios sobre el plan estructural y composición histológica de los centros nerviosos adicionados de consideraciones fisiológicas fundadas en los nuevos descubrimientos*, Madrid, Imprenta y Librería de Nicolás Moya, 2 tomi in 3 voll., 1899–1904; Wilhelm von Waldeyer-Hartz, "Ueber einige neuere Forschungen im Gebiete der Anatomie des Centralnervensystems", *Deutsche Medizinische Wochenschrift* 17 (1891), pp. 1213–1218, 1244–1246, 1267–1270, 1287–1289, 1331–1332 e 1352–1356; Santiago Ramón y Cajal, *Recuerdos de mi vida*, 3ª ed., Madrid, Imprenta de Juan Pueyo, 1923, in particolare la seconda parte, *Historia de mi labor científica*; Camillo Golgi, "The Neuron Doctrine-Theory and Facts", Nobel Lecture, 11 dicembre 1906; Francis H. Cook, *Hua-yen Buddhism: The Jewel Net of Indra*, University Park, Pennsylvania State University Press, 1977; Warren S. McCulloch e Walter Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity", *The Bulletin of Mathematical Biophysics* 5, n. 4 (1943), pp. 115–133; Frank Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", *Psychological Review* 65, n. 6 (1958), pp. 386–408; Marvin Minsky e Seymour A. Papert, *Perceptrons: An Introduction to Computational Geometry*, Cambridge, MA, MIT Press, 1969; George Cybenko, "Approximation by Superpositions of a Sigmoidal Function", *Mathematics of Control, Signals, and Systems* 2, n. 4 (1989), pp. 303–314; Kurt Hornik, "Approxima-

tion Capabilities of Multilayer Feedforward Networks", *Neural Networks* 4, n. 2 (1991), pp. 251–257; David E. Rumelhart, Geoffrey E. Hinton e Ronald J. Williams, "Learning Representations by Back-Propagating Errors", *Nature* 323 (1986), pp. 533–536; Alex Krizhevsky, Ilya Sutskever e Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Advances in Neural Information Processing Systems* 25 (2012), pp. 1097–1105.