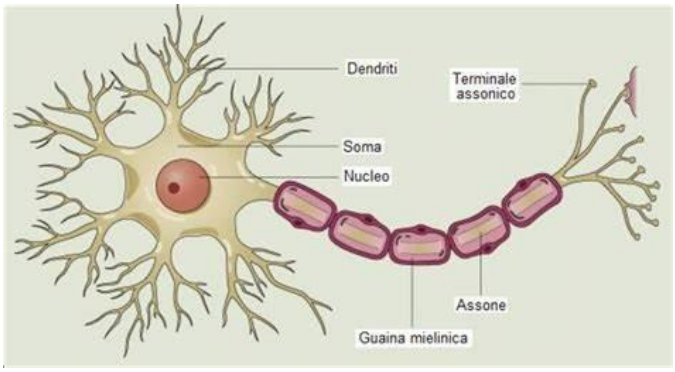
Le nuove frontiere dell'Intelligenza artificiale

Il periodo a cavallo tra gli anni '70 e gli anni '80 vide fiorire la grande stagione dei "sistemi esperti". Come si è visto, erano questi i promettenti prodotti nel campo dell'Intelligenza Artificiale. Progettati su una base di conoscenze specialistiche e di un "motore" di regole di inferenza logica al fine di ottenere indicazioni operative nelle discipline per cui erano stati creati. Intorno alla metà degli anni '80 i sistemi esperti comparivano ovunque in diversi domini della conoscenza; euforia e grandi investimenti si alimentarono a vicenda nell'idea che indicassero il futuro dell'IA. Intelligenza artificiale e "sistemi esperti" erano considerati sinonimi nell'idea che tali programmi riuscissero a risolvere problemi sviluppando conoscenze di base attraverso regole di inferenza.

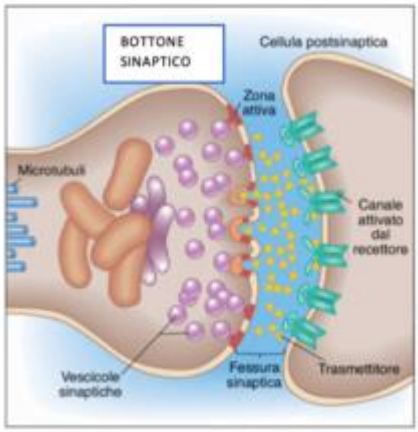
Ma verso la fine degli anni '80 gli entusiasmi si raffreddarono e vennero meno i generosi finanziamenti nel settore. Si mostrarono con evidenza i limiti tecnologici di tali sistemi: la necessità di implementare i programmi con nuove conoscenze e la difficoltà di scrivere nuove regole di inferenza ne diminuivano le prestazioni, sia nella velocità di elaborazione dei dati, sia nella qualità e precisione dei risultati. Iniziò così quello che gli studiosi definiscono il periodo "invernale" dell'IA, durante il quale germinarono alcune idee innovative che avrebbero, pochi anni dopo, rivoluzionato approcci, metodi e strategie. Queste nuove ricerche diedero risultati inattesi e strabilianti quando trovarono un potente alleato nel *World Wide Web* che, lanciato nel 1994, avrebbe trasformato radicalmente la ricerca dell'IA. Tornò appetibile l'idea di simulare la complessa struttura neuronale del cervello umano: visto che le reti neurali che lo compongono rendono questo organo è così abile a pianificare e imparare, non è da scartare l'idea di simularne la funzione semplicemente in vista dei risultati che produce. Semplificando molto, vediamo come funzionano.

Le reti neurali profonde



Il cervello umano è composto da circa 86 miliardi di neuroni e ciascuno di essi ha in media 7000 connessioni sinaptiche con altri neuroni.

I segnali che giungono a ogni cellula nervosa (neurone) vengono inoltrati ad altri neuroni attraverso gli assoni. Nelle parti terminali di questi ultimi si trovano dei punti di collegamento (sinapsi) che trasmettono chimicamente il segnale. Il cervello umano è composto da circa 86 miliardi di neuroni e ciascuno di essi ha in media 7000 connessioni sinaptiche con altri neuroni.



L'idea, come si è già visto, fu quella di simulare le reti neurali biologiche anche all'interno di un computer con l'aiuto della matematica, rappresentando l'intensità dell'impulso utilizzando un numero che esprime la forza della connessione sinaptica. Infatti, diversamente ai collegamenti tra i cavi elettrici, l trasmissione del segnale nelle sinapsi può essere più o meno buona.

La maggiore o minore intensità con la quale si propaga lo stimolo elettrico viene chiamato *peso sinaptico*. Se il valore supera una determinata *soglia di attivazione* del neurone, questo si attiva generando un *potenziale di azione* che eccita più o

meno intensamente il neurone successivo. La scoperta interessante riguarda la *neuroplasticità* del cervello, ossia la sua proprietà di modificarsi durante i processi mentali. Ciò che si modifica è il peso sinaptico, cioè la forza dei collegamenti tra cellule nervose. Tale plasticità sta alla base di ogni processo di apprendimento.

La seconda rivoluzione informatica iniziò con la costruzione di "reti multistrato", a imitazione di quelle biologiche, applicando a esse la tecnica dell'*apprendimento profondo*. Queste reti sono composte da diversi strati, o livelli, di neuroni (artificiali). I neuroni di ogni strato ricevono in ingresso informazioni da tutti i neuroni dello strato precedente per trasferirle al livello successivo. A ogni canale che acquisisce un dato in ingresso viene associato un *peso*, ossia un valore numerico che rappresenta quanto è importante quel dato nel processo che darà luogo a un risultato finale. Durante la fase di allenamento il sistema legge i dati in ingresso (input), li processa e produce un risultato finale (output). Se il risultato corrisponde a quello atteso il sistema viene "ricompensato". Ovviamente, questa "ricompensa" non deve essere presa alla lettera: il computer non gioisce, né soffre, e neppure si intristisce. Si tratta semplicemente di una funzione matematica che incrementa o meno una funzione matematica che indica una sorta di punteggio che la rete deve riuscire a massimizzare. Se il risultato non corrisponde a quello atteso, viene attivato un algoritmo di retropropagazione (*backpropagation*), che provvede alla correzione dei pesi associati a ogni connessione. Al termine di questa fase la rete impara a costruire un modello in grado di attivare il percorso più efficace per conseguire il risultato richiesto.

Fu così che Deep Mind, un'azienda controllata da Google, riuscì a creare AlphaGo, un potentissimo algoritmo in grado di imparare a giocare a Go; non solo, ma a Seul, nel 2016, la macchina sconfisse il coreano Lee Sedol, considerato il più abile giocatore di Go nel mondo intero. AlphaGo fu allenato su un database di 30 milioni di mosse diverse estratte da 160.000 partite giocate da campioni. Quindi, fu ulteriormente addestrato facendolo giocare 50 milioni di volte contro se stesso. In seguito, ogni versione successiva di AlphaGo sconfisse la precedente e l'ultima versione non ebbe bisogno neppure della collezione di 30 milioni di mosse da cui partire: "imparò" a giocare da sola, sconfiggendo la versione originale del programma in tutte le partite.

Queste reti neurali a più strati che apprendono sono alla base dei più moderni modelli di IA. Una considerazione sembra rilevante e particolarmente significativa per riflettere sulla natura di questo straordinario strumento:

> *"Quando le reti neurali, e dunque i vettori, sono sufficientemente grandi (per es. 175 sinapsi) non c'è nulla che si possa fare per ripercorre a ritroso, e dunque per capire, ciò che è successo davvero. La rete semplicemente fa e anche lo studio più accurato della massa di numeri che compongono i vettori non potrà mai fare chiarezza su come la rete abbia davvero funzionato. Detto in altre parole: i programmi [per i tradizionali computer] sono scritti in un linguaggio che come tale può essere compreso da un essere umano; un vettore di trasformazione al contrario non è altro che una sequenza di cifre, un insieme ordinato di numeri,dai quali non si evince nulla. Il singolo numero che determina il peso di una sola sinapsi non ci racconta niente sul contributo che essa ha dato al funzionamento dell'intera rete".*[1]

[1] Spitzer Manfred, *Intelligenza artificiale. Opportunità e rischi di una rivoluzione tecnologica che sta cambiando il mondo*, Corbaccio, Milano 2024, p. 110.

Applicazioni

Chimica

Nel campo della chimica organica le reti neurali fecero la loro comparsa nei primi anni di questo secolo, per avvalersi ben presto della tecnica dell'apprendimento automatico da parte delle reti stesse. L'obiettivo era quello di ottenere in modo automatico la sintesi di nuove molecole complesse e di far generare dalle macchine percorsi di reazione del tutto nuovi per velocizzare le sintesi. Per i sistemi esperti sarebbe stato un compito proibitivo data la sconfortante complessità di dati e calcoli combinatori necessari per pianificare nuove sintesi.

È a questo punto che viene in aiuto l'IA Chematica per progettare nuovi percorsi di sintesi di importanza fondamentale per la medicina. Tali percorsi furono successivamente verificati da chimici esperti e da questi replicati in laboratorio, con notevoli risparmi di costi. Il principio guida fu quello di associare la produzione di sintesi di nuove molecole a un "gioco". In sostanza, "le regole che governano le reazioni costituiscono le *mosse fondamentali* dalle quali partire per costruire i percorsi di sintesi completi (il *gioco*)."[2] Poiché il numero di possibilità per ogni passaggio si accrescerà esponenzialmente per ogni passaggio lo spazio di ricerca diviene talmente grande da rendere impossibili i calcoli (come nel gioco del Go). L'intelligenza artificiale interviene principalmente per tagliare le ramificazioni prive di prospettiva".[3]

Chematica aiuta anche a trovare alternative alla produzione di farmaci altrettanto efficaci a quelli già esistenti, o anche migliori. L'interesse dell'industria farmaceutica per l'IA è dovuto anche, o soprattutto, a un motivo di natura prettamente economica. Quando un brevetto per la produzione di un farmaco scade, la ditta farmaceutica che lo possiede detiene comunque, anche dopo la scadenza del brevetto, la proprietà della "ricetta" da seguire per la produzione del farmaco. Per fare un esempio culinario, scade il brevetto che protegge l'esclusiva della "torta margherita", ma non la proprietà della ricetta attraverso cui la si ottiene. Chematica viene utilizzata proprio per aggirare tale vincolo di proprietà, poiché è in grado di fornire "ricette" alternative" per un farmaco altrettanto efficace. Questo è già accaduto nel caso di un antibiotico, di un farmaco per la cura del tumore osseo e di un medicinale per la cura del diabete. In questi casi si è addestrata Chematica con i dati di centinaia di reazioni protette da brevetto, chiedendo al sistema di svilupparne di nuove e inedite. Chematica permetteva così di aggirare la protezione dei processi di sintesi delle molecole ancora protetti da brevetto.

Elaborazione di immagini

Gli smartphone più recenti sono dotati di programmi che applicano alle foto dei filtri che trasformano le immagini in qualcosa di simile a un'opera d'arte imitando, per esempio lo stile di Van Gogh. Utilizzano sistemi composti da due reti neurali che operano in simbiosi:

- Una prima rete, chiamata il *discriminante*, è deputata a valutare se una determinata immagine è un falso o un originale, e viene "ricompensata" ogni volta che vi riesce;
- Una seconda rete ha invece il compito di ingannare la prima e, a sua volta, viene "ricompensata" quando raggiunge il proprio obiettivo.

La simbiosi tra le due reti assomiglia, insomma, a una partita tra due avversari, ciascuno dei quali vuole avere la meglio sull'altro. Inizialmente, la rete "falsaria" produrrà informazioni palesemente riconoscibili come fasulle; poi, un poco alla volta, imparerà a generare informazioni tali da ingannare la controparte, generando immagini sempre più vicine allo stile dell'artista in gioco. Questi strumenti sono in grado di generare notizie, video, immagini di persone che dicono o fanno cose mai accadute o, comunque, difficilmente valutabili in termini di verità o falsità. Il primo caso di *deepfake attack* fu riportato dal "Washington Journal"; un impiegato di un'azienda ricevette telefonicamente la richiesta

[2] Ivi, p. 129.
[3] Ibidem.

di fare un bonifico di 240.000 dollari a favore di un fornitore. La procedura di impartire un ordine telefonicamente era prevista e praticata normalmente dall'azienda, ma in questo caso, la voce che ordinava all'impiegato di effettuare il bonifico era stata riprodotta simulando quella del titolare dell'azienda.

Queste macchine composte di reti neurali ci stupiscono con la loro capacità di riprodurre stili di pittori, musicisti e altro ancora. Dobbiamo aggiungere però che parlando in questo modo rischiamo di dimenticare e ignorare quel "lavoro altrettanto o assai più meraviglioso e stupefacente dell'umano linguaggio, capace di tradurre l'infinita e irriducibile complessità di una qualsiasi esperienza vivente in una successione di suoni significativi, cioè di parole".[4] Ritornando a video, immagini e audio fasulli, altro non replicano, se non in modo potenziato e più pericoloso, la comune maldicenza che affligge gli esseri umani nell'uso del linguaggio da tempi immemorabili.

Presto tardi – auspica Stefano Quintarelli – la politica finirà per regolamentare l'uso che si fa dalla IA. Per il momento, però, la soluzione sembra lontanissima: per il momento non si vede che l'uso cinico e sconsiderato che la stessa politica fa di questi strumenti, utilizzandoli a fini elettorali e populistici.

Pilota automatico

Alcuni esempi di guida affidata all'IA sono già fra noi, come le navette a rotaia che ci accompagnano da un terminal all'altro di un aeroporto; gli aerei, anche quelli di linea, adottano già da tempo sistemi di volo automatici: nel gennaio 2020 un aereo di linea è atterrato otto volte guidato completamente dal pilota automatico; inoltre, è in fase di avanzamento la sperimentazione di veicoli a guida automatica. L'ente internazionale che regola e norma l'industria spaziale e automobilistica (SAE), indica sei livelli di automazione: da un livello 0, dove tutte le attività di controllo sono a carico del pilota, al livello 5, dove il veicolo è guidato completamente da un agente artificiale, che si occupa anche delle decisioni da prendere per fronteggiare imprevisti. Gli ostacoli per raggiungere l'obiettivo di una completa automazione sono ancora molti. Innanzi tutto, la quantità di dati che un veicolo autonomo deve gestire è enorme (circa 5 TB all'ora) e richiede tecnologie molto avanzate e capaci di notevoli capacità computazionali. Inoltre, risulta particolarmente difficile risolvere il problema della imprevedibilità dell'ambiente: nel 2018, durante un test di Uber un veicolo ha investito e ucciso un pedone che attraversava un tratto di strada poco illuminato. Il numero di incidenti è in aumento, molti dei quali dovuti alla scarsa vigilanza dei passeggeri che confidano eccessivamente nell'automatismo.

I sistemi di guida autonoma possono essere anche oggetto di attacco di hacker, è stato dimostrato che è possibile intervenire da remoto su una Tesla modificando le informazioni sull'ambiente o le risposte del veicolo: un segnale di stop può essere trasformato in un segnale di diritto di precedenza con effetti disastrosi.

Amabot: il tramonto di un paradigma

Fondata da Jeff Bezos il 5 luglio del 1994 con il nome di *Cadabra*, con l'idea di dar vita a una libreria online. Gli fu subito consigliato di abbandonare quel nome (Cadabra) che, pronunciato, avrebbe potuto richiamare il termine "cadaver". Fu così sostituito da *Amazon*, con riferimento al grande fiume del continente sudamericano. Nei primi tempi l'azienda si era affidata a una dozzina di redattori, "la voce di Amazon", incaricati di scrivere recensioni di qualità dei libri proposti. La rapida crescita dell'azienda e l'espansione del catalogo a musica e film rendevano insufficiente l'apporto del piccolo gruppo di redattori, che si trovarono a dover recensire (e leggere) una ventina di libri a settimana. Jeff Bezos decise così di ricorrere allo stratagemma di spostare il peso della pubblicità dei

[4] Sini Carlo, *Intelligenza Artificiale e altri scritti*, Jaca Book, Milano 2024, p. 17.

prodotti dalle recensioni alle "raccomandazioni". "In generale, un *agente di raccomandazione* è incaricato di individuare gli articoli che hanno la maggiore probabilità di interessare un dato utente".[5]

Nel 1998 una squadra di programmatori sviluppò un algoritmo con il quale era possibile catturare acquirenti utilizzando il database delle vendite. L'idea era questa: personalizzare i suggerimenti sulla base degli acquisti dei clienti. Ancora oggi, dopo aver fatto la nostra scelta, troviamo nel sito la frase "prodotti correlati", oppure "chi ha acquistato questo articolo ha acquistato anche…". In altri termini, il nuovo sistema teneva conto del comportamento degli utenti, di ciò che gli utenti facevano in realtà. Si scoprì che i clienti acquistavano più libri in seguito a queste "raccomandazioni" generate dagli *Amabot* (così furono chiamati questi programmi) che non seguendo recensioni scritte da redattori in carne e ossa. Non solo, ma la pubblicità dei prodotti veniva così affidata, in ultima analisi, agli stessi acquirenti. Il nuovo "robot" decise rapidamente il destino dei bravi redattori della "voce di Amazon", licenziati alla fine del 1999. Sul *Seattle Weekly* comparve la seguente inserzione anonima, quasi un epitaffio:

> *"Carissimo Amabot, grazie tante. Se tu solo avessi un cuore per assorbire il nostro odio. Vecchio arnese scassato, la splendida confusione della carne e del sangue vincerà".*[6]

"Amabot non era animato da regole esplicite, né da alcuna comprensione dei clienti o dei contenuti delle merci offerte: il suo lavoro dipendeva da relazioni statistiche rilevate nel database degli acquisti fatti. […] Tutto ciò che doveva fare era raccogliere informazioni sul comportamento dell'utente e applicare algoritmi per sfruttare le relazioni statistiche esistenti in quei dati, trasformandole in decisioni utili".[7] Stava nascendo un nuovo paradigma, ossia l'insieme delle convinzioni implicite su cui la ricerca scientifica effettua il lavoro di ricerca: i modelli teorici stavano per essere sostituiti dalle regolarità statistiche; una sorta di "scorciatoia che permetteva di evitare Una macchina basata sull'apprendimento statistico sarebbe risultata più efficace nel predire le osservazioni future di quella programmata per regole. Una scorciatoia che sorregge anche i tentativi di risolvere problemi relativi al linguaggio; se, per esempio, le traduzioni o la elaborazione di testi possono essere approssimate senza comprendere prima il fenomeno del linguaggio, perché complicarsi la vita?

Il linguaggio della nuova IA è quello della statistica, che si fonda sulla disponibilità di un'immensa quantità di dati per addestrare le macchine perché producano risultati predittivi: una vera e propria ossessione per la misura delle prestazioni. L'obiettivo non è qualche "verità" o la comprensione di ciò che accade, ma generare comportamenti umani, a prescindere dal loro senso: per questo scopo sono sufficienti e più performanti semplici (si fa per dire) modelli statistici. Nel caso del linguaggio umano è più semplice predire la parola seguente che comprendere il senso di una frase. Il modello GPT-3 è stato "addestrato" su 45 terabyte di testo raccolto da diverse fonti e un milione di parametri.

ChatGPT

ChatGPT non è un programma per computer, bensì una rete neurale (175 miliardi di sinapsi) per l'elaborazione del linguaggio naturale.

ChatGPT è basato su 175 miliardi di sinapsi (noi 700.000 miliardi), preparata con 570 Gb di dati. Nasce il 30 novembre 2022. Ha sortito effetti che hanno allarmato gran parte della comunità scientifica, al punto che a pochi mesi dalla sua pubblicazione l'istituto statunitense *Future of Life* diffuse una lettera aperta, firmata da circa 1300 esperti, nella quale si chiede una sospensione di

[5] Cristianini Nello, La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano, il Mulino, Bologna 2023, p. 39.

[6] Ivi., p. 38.

[7] Ivi, pp. 41-2.

almeno sei mesi dell'impiego della IA. Il 30 maggio del 2023, anche un gruppo di aziende che si occupano di IA esprimono la preoccupazione nei confronti di una tecnologia che "potrebbe potenzialmente rappresentare una minaccia per l'esistenza dell'umanità, un rischio grave per la nostra società, al pari delle pandemie e delle guerre atomiche".[8]

Perché tanta preoccupazione? I tentativi di costruire chatbot risalgono al 1966, quando Joseph Weizenbaum (1923-2008) realizzò ELIZA. Poi arrivarono gli assistenti vocali come Siri di Apple (2011), Cortana di Microsoft (2014) e Alexa di Amazon (2015). Darren Gill, responsabile della gestione del prodotto in Amazon, si stupisce della frequenza e della facilità con cui gli esseri umani tendono a interagire con Alexa come se si trattasse di un umano. Centinaia di milioni le danno il "buongiorno", mezzo milione le hanno dichiarato il proprio "amore" e oltre duecentocinquantamila le hanno addirittura fatto una proposta di matrimonio. La parola più usata per rivolgersi all'assistente vocale è "grazie". Come mai? A nessuno verrebbe in mente di salutare il proprio frigorifero oppure ringraziare gli elettrodomestici di casa. Forse perché si tratta di una macchina la cui peculiarità è quella di simulare il linguaggio umano e una sorta di dialogo, seppure elementare. Da mettere in rilievo, inoltre, che Alexa è stato progettato pensando all'intimità di un ambiente domestico, che basti pronunciare il suo nome perché si accenda e che gli sviluppatori abbiano pensato di introdurre artificialmente delle pause nelle frasi pronunciate, psicologicamente avvertire come momenti di riflessione da parte della macchina.

Riprendiamo uno spunto di riflessione che svilupperemo in seguito. L'IA ci procura sicuramente benefici, come altri strumenti che hanno segnato la storia dell'umanità; ma la potenza di tale strumento ci mette anche angoscia e apprensione soprattutto quando tocca ciò che consideriamo proprio della sola specie umana: il linguaggio.

Breve scheda bibliografica di riferimento

- Larson Erik J., *Il mito dell'intelligenza artificiale. Perché i computer non possono pensare come noi.* Franco Angeli, Milano 2022.
- .Warwick Kevin, *Intelligenza artificiale. Le basi*, Faccovio Editore, Palermo 2015.
- Hénin Silvio, *Intelligenza artificiale tra incubo e sogno*,Hoepli, Milano 2019.
- Quintarelli Stefano (a cura di), *Intelligenza Artificiale. Cos'è davvero, come funziona, che effetti avrà*, Bollati Boringhieri, Torino 2020.
- Sini Carlo, *Intelligenza Artificiale e altri scritti*, Jaca Book, Milano 2024.
- Spitzer Manfred, *Intelligenza artificiale. Opportunità e rischi di una rivoluzione tecnologica che sta cambiando il mondo,* Corbaccio, Milano 2024.
- Cristianini Nello, *La scorciatoia. Come le macchine sono diventate intelligenti senza pensare in modo umano,* il Mulino, Bologna 2023.

[8] Spitzer M. op. cit. p. 17.