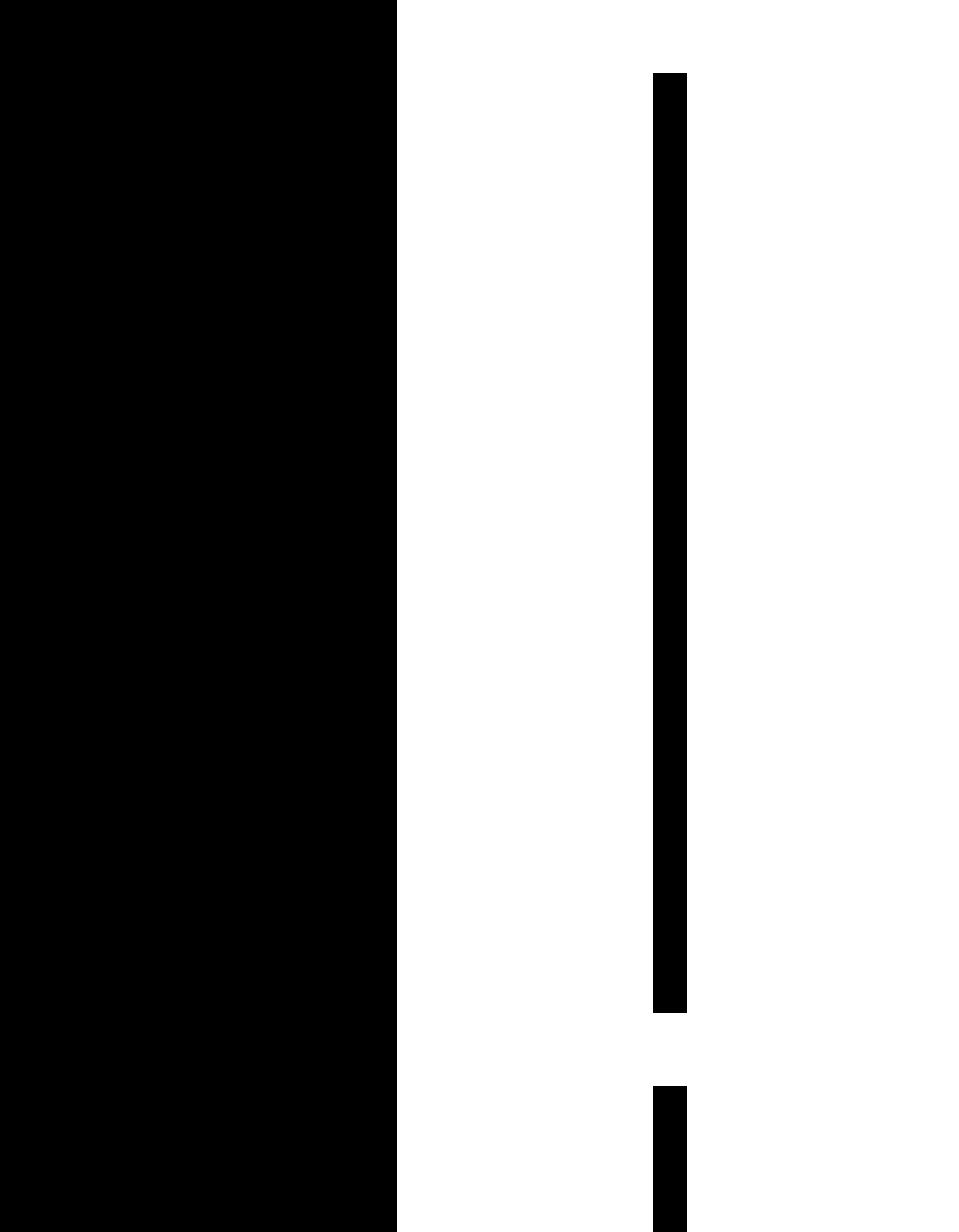


[IA]

anno 20 numero 39/ 10.25

asimmetrie

rivista semestrale dell'Istituto
Nazionale di Fisica Nucleare



asimmetrie

Care lettrici e cari lettori,

Da alcuni anni ormai, l'intelligenza artificiale (IA) è presente in molti momenti della vita quotidiana di tutti noi: nei nostri computer, sui nostri telefoni, sui social, nelle auto, negli elettrodomestici. Il boom dell'IA è avvenuto nel momento in cui si sono realizzate alcune condizioni tecnologiche abilitanti: reti veloci, grandi quantità di dati, sufficiente potenza di calcolo su architetture hardware dedicate, nuovi algoritmi.

Tuttavia, i primi modelli teorici di IA risalgono al secondo dopoguerra, mentre le prime applicazioni diventano possibili solo alla fine degli anni '80 del secolo scorso. Già in questo periodo il *machine learning* incontra la fisica delle particelle. I grandi esperimenti di quegli anni, LEP al CERN e Tevatron al Fermilab, iniziano a usare le reti neurali per analizzare i dati e riconoscere le tracce delle particelle pesanti, in particolare il quark *b*. In effetti, ripensandoci ora, quegli esperimenti possedevano, sia pure al livello possibile più di trent'anni fa, tutti gli elementi necessari per provare ad applicare l'IA con successo: potenza di calcolo e reti al top per l'epoca, grandi quantità di dati sperimentali, algoritmi di *machine learning*, non ancora *deep*, basati su perceptron e *back propagation* (che troverete discussi in questo numero). Pur con le limitazioni imposte dalla tecnologia di quegli anni, l'IA mostrò di essere uno strumento concorrenziale con gli altri approcci e si guadagnò un posto tra le tecniche per l'analisi dei dati utilizzate nella fisica delle particelle.

Con l'evoluzione tecnologica e algoritmica degli ultimi decenni, l'IA è diventata pervasiva in vari ambiti della ricerca dell'INFN: non c'è ormai analisi dati complessa che non preveda un approccio basato sull'IA, dalle particelle elementari, alle astroparticelle fino alle onde gravitazionali. In diversi casi, l'IA consente di abbreviare i tempi delle analisi e quindi ridurre i costi. Il *machine learning* viene utilizzato con successo anche nella fenomenologia delle particelle e ci sono prospettive per l'applicazione dell'IA generativa a problemi più teorici. Altrettanto diffusa ed efficace è la diffusione dell'IA in vari campi della fisica applicata. In fisica medica, ad esempio, l'IA trova ampia applicazione, dal *protein folding* alla diagnostica per immagini, dai gemelli digitali alla medicina di precisione.

Concludo ricordando che oggi, grazie al supporto del MUR tramite il PNRR, disponiamo del Centro Nazionale ICSC, che integra il supercomputer Leonardo del CINECA e l'infrastruttura di calcolo distribuita dell'INFN, mettendo a disposizione della comunità scientifica un centro adeguato ad affrontare con successo le sfide poste dai nuovi strumenti di ricerca, come l'IA.

Buona lettura.

Antonio Zoccoli
Presidente INFN



asimmetrie

as

39 / 10.25 [IA]

Ex machina di Daniele Bonacorsi	4	[as] riflessi Dove l'IA incontra la potenza di calcolo di Martina Galli	34
[as] IA in pillole	8	[as] intersezioni Bella senz'anima di Francesco Bianchini	36
Cogito ergo \sum di Elena Agliari e Adriano Barra	10	[as] illuminazioni Riconoscimento di immagini con "carta e penna" di Anna Greco	39
Rovescio vincente di Matteo Osella	14	[as] visioni Ma che domande! di Giuliana Galati	40
The imitation game di Antonio Celani	17	[as] traiettorie Dalle particelle all'IA di Francesca Mazzotta	42
Facile per la natura, difficile per l'uomo di Piero Fariselli	21	[as] spazi Il campus dell'INFN di Martina Bologna	44
Non ti avevo riconosciuto... di Tommaso Boccali	23	[as] convergenze Lacrime per il bosone di Higgs di Ottavia Fusco Squitieri	46
Cogli l'ottimo di Maurizio Pierini	26	[as] segni	48
Come ti leggo dentro di Stefano Forte	28		
Occhio clinico di Marianna La Rocca	31		

Ex machina

L'evoluzione dell'intelligenza artificiale

di Daniele Bonacorsi



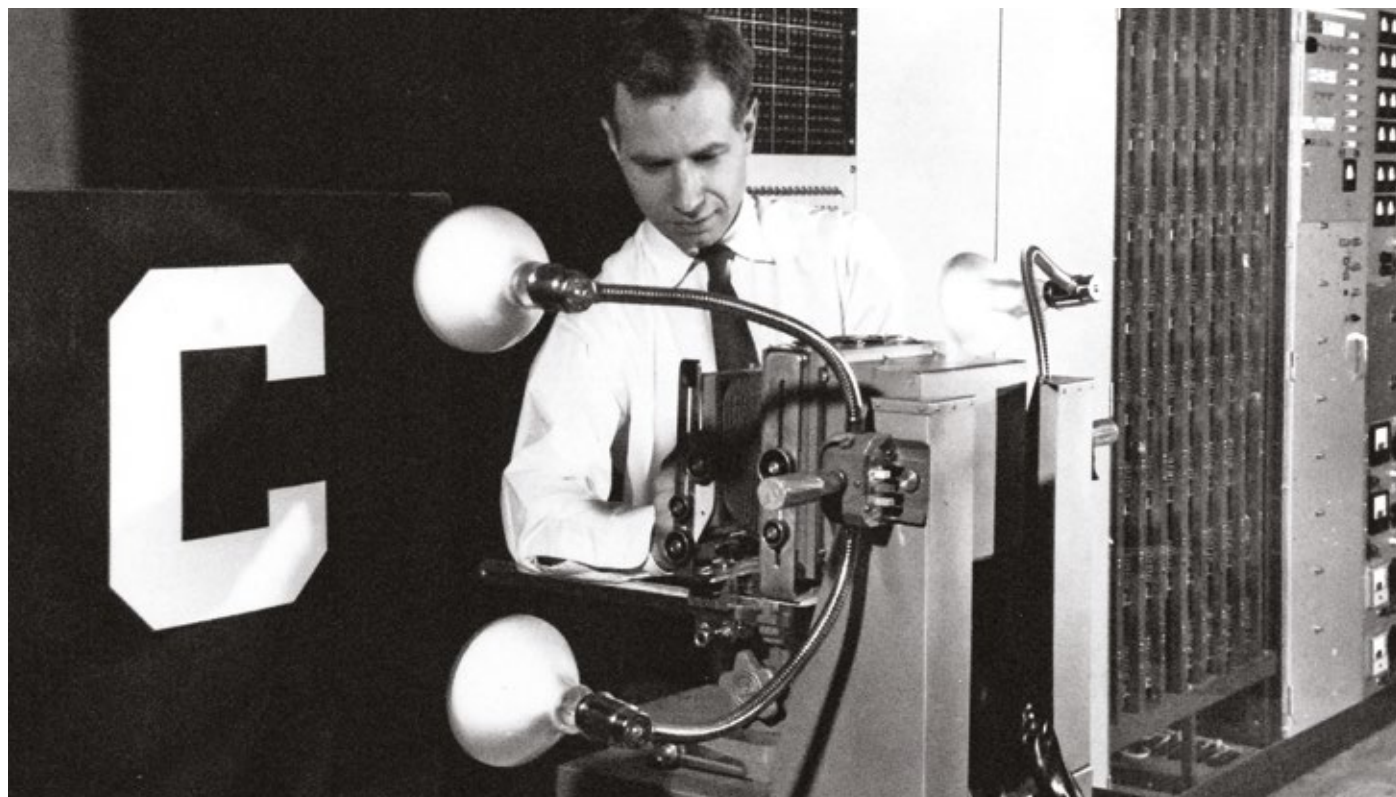
Negli ultimi anni, il termine “intelligenza artificiale” (IA) ha subito una rapida transizione da concetto puramente specialistico a espressione d'ampio uso quotidiano. Forte dei suoi successi e di una evidente ubiquità nella discussione pubblica, nonché di una pervasività senza precedenti nel mondo delle applicazioni software, dietro l'IA si nasconde tuttavia una lunga storia che alterna intuizioni geniali a progressi discontinui. Tale disciplina, dalle origini a oggi, ha seguito infatti un percorso complesso, che da alcuni anni sembra vivere un'esplosione senza fine, con evidenti impatti nella ricerca scientifica e nelle applicazioni tecnologiche e industriali. Oggi l'IA non è soltanto una delle sfide scientifiche e tecnologiche più ambiziose del nostro tempo, ma anche, e sempre più, un elemento profondamente trasformativo del tessuto della nostra società, che permea ambiti cruciali come il mondo del lavoro, le relazioni sociali, i sistemi di comunicazione, le modalità di accesso all'informazione, l'educazione, l'economia e molti processi decisionali, tanto individuali quanto collettivi.

Analizzandola fin dalle sue origini, quella dell'IA è una storia fatta di alti e bassi. Le sue radici affondano in un primo modello matematico semplificato del neurone, proposto nel 1943 da Warren McCulloch e Walter Pitts, che dimostrarono

come “reti” di unità binarie potessero, in linea di principio, realizzare qualsiasi computazione logica. Nel 1958, Frank Rosenblatt introdusse il “perceptrone”, un primo modello di strato di neuroni artificiali in grado di imparare da esempi mediante calcoli su cifre non più solo binarie, e passò poi anche alla disposizione su più strati (“layers”). Stavano nascendo le “reti neurali artificiali”, ovvero modelli computazionali ispirati alla struttura del cervello umano, composte da unità elementari (“neuroni” artificiali) interconnesse e organizzate in strati, in grado di elaborare informazioni complesse attraverso la propagazione e trasformazione di segnali numerici. Fu l'inizio della realizzazione di un sogno: quello di replicare, con metodi computazionali, alcune capacità cognitive tipiche degli esseri umani.

Tuttavia, il progresso fu irto di rallentamenti e ostacoli tecnici. Già negli anni '70 si verificarono i primi “inverni dell'IA”, lunghi periodi di stasi e disillusione, causati dai limiti teorici e computazionali dell'epoca: i modelli erano infatti troppo semplici per affrontare compiti complessi. Un punto di svolta cruciale nella storia dell'IA avvenne negli anni '80, con lo sviluppo di metodi efficaci per addestrare reti neurali su dati reali. Prima di approfondire tali metodi, occorre ricordare che in quegli anni tutti questi algoritmi

a.
Frank Rosenblatt con il perceptrone sviluppato nel 1958.



erano descritti nel cosiddetto “*machine learning*” (apprendimento automatico). Si tratta di una classe di algoritmi che consente ai computer di apprendere schemi e regole decisionali a partire dai dati, senza la necessità di una programmazione esplicita basata su regole. Questi sistemi sono in grado di generalizzare dagli esempi osservati per effettuare previsioni o prendere decisioni su input mai visti prima, e sono oggi alla base di moltissime applicazioni (vd. anche in Asimmetrie n. 27 p. 31, ndr). Solo nel secondo decennio del secolo corrente, invece, si è affermata una sotto-classe dell’apprendimento automatico, il cosiddetto “*deep learning*” (apprendimento profondo). Basate su reti neurali artificiali ma con molti “strati” (da cui il termine “*deep*”) di “neuroni”, le architetture di *deep learning* si sono dimostrate particolarmente efficaci nell’elaborazione sia di dati strutturati che non strutturati, come immagini, testo e audio, e sono oggi responsabili di molti dei risultati più spettacolari ottenuti negli ultimi anni nel campo dell’intelligenza artificiale.

Ma torniamo ai metodi per addestrare reti neurali su dati reali. L’idea di assegnare dei “pesi” ai collegamenti tra neuroni artificiali sorge dall’ispirazione di imitare il funzionamento del cervello biologico, dove la forza delle sinapsi influenza la trasmissione dell’informazione. Da qui, il passo successivo fu immaginare un meccanismo di apprendimento basato sulla modifica graduale di questi pesi, in risposta agli errori commessi durante il confronto tra le uscite prodotte dalla rete neurale e i risultati attesi. E questo divenne tecnicamente possibile con l’introduzione e la diffusione dell’algoritmo di “*backpropagation*”, che ha permesso l’addestramento di reti neurali anche complesse, a molti strati, aprendo la strada al *deep learning*. Al tempo stesso, un altro fattore decisivo nella rinascita dell’IA è stata la diffusione delle GP-GPU (“*general-purpose graphics processing unit*”): originariamente progettati per la grafica, questi dispositivi si sono rivelati ideali anche come acceleratori per svolgere in parallelo i numerosissimi calcoli necessari all’addestramento delle reti neurali, riducendo drasticamente i tempi necessari. Con la potenza di calcolo delle GP-GPU, la disponibilità di grandi volumi di dati digitali eterogenei (“*big data*”), i metodi di ottimizzazione e gli algoritmi di allenamento giunti a piena maturazione, tutti gli ingredienti erano finalmente sul tavolo: dopo oltre mezzo secolo, modelli teorici solidi hanno trovato terreno fertile per una loro applicazione efficace, e il *deep learning* è potuto diventare realtà, segnando l’inizio di una “primavera” dell’IA che, ad oggi, non sembra avere fine. Nel giro di pochi anni, concentrati nella seconda decade del XXI secolo, le reti neurali hanno vissuto un’esplosione di architetture innovative,

ciascuna progettata per un diverso tipo di dato o di compito da svolgere (vd. anche in Asimmetrie n. 27 p. 15, ndr). Le Convolutional Neural Network (CNN o ConvNet, reti neurali convoluzionali) hanno rivoluzionato la visione artificiale, diventando lo standard per il riconoscimento di immagini, volti e ogni tipo di schema o configurazione spaziale (in gergo “*pattern*”). Le Recurrent Neural Network (RNN, reti neurali ricorrenti) – insieme a varianti più stabili come le Long Short-Term Memory (LSTM) – hanno permesso alle reti neurali di gestire dati sequenziali e temporali, come ad esempio testi, suoni e in generale segnali di vario tipo. Gli Autoencoder (AE) hanno aperto la strada a tecniche di compressione, ricostruzione e riduzione dimensionale, fungendo anche da base per metodi di identificazione di anomalie (“*anomaly detection*”). L’introduzione delle Generative Adversarial Network (GAN, reti generative antagoniste) – in cui due reti in competizione tra loro riescono a creare immagini e contenuti sintetici di sorprendente realismo – hanno cambiato il modo di pensare la generazione di dati. Il passaggio dai modelli discriminativi di IA, progettati per distinguere e classificare dati esistenti, ai modelli generativi di IA, capaci di produrre autonomamente nuovi contenuti ad alto realismo, ha segnato un cambio di passo radicale: l’IA è passata dall’analizzare la realtà al saperne creare di nuove. In quest’ambito, più recentemente, i Diffusion Models (modelli di diffusione) hanno assunto un ruolo di primo piano nella generazione condizionata di dati guidata da istruzioni testuali (in gergo “*prompt*”), specialmente nel caso di immagini e audio. Infine, l’arrivo dei Transformer ha segnato una svolta epocale: progettati inizialmente per il “linguaggio naturale” (testo scritto o parlato), si sono rapidamente estesi a ogni tipo di dato – dal DNA alla musica – grazie al meccanismo di *self-attention*, diventando la spina dorsale dei Large Language Models (LLM), ovvero modelli linguistici di grande scala, e più in generale della nuova era dell’IA generativa.

Un riconoscimento importante per la storia dell’IA è arrivato nel 2024, con l’attribuzione di premi Nobel che toccano l’IA da molto vicino. Il Nobel per la chimica 2024 è stato assegnato a Demis Hassabis e John M. Jumper per lo sviluppo di un modello di IA capace di prevedere la struttura tridimensionale delle proteine: premiando la soluzione di un problema aperto da oltre 50 anni nella biologia strutturale, questo riconoscimento testimonia come lo sviluppo di tecniche basate su IA possa essere trasformativo nella progettazione di nuovi composti molecolari.

Il Nobel per la fisica del 2024 è stato conferito a John J. Hopfield e Geoffrey E. Hinton. Hopfield nel 1982 ha introdotto il modello di rete neurale che tuttora porta il suo nome, capace di



representare configurazioni di attività neuronale che corrispondono a *pattern* appresi, seguendo principi ispirati alla fisica statistica. Hinton ha successivamente sviluppato la macchina di Boltzmann, una rete neurale stocastica che modella la distribuzione di probabilità dei dati attraverso una funzione di energia, e che ha avuto un ruolo fondamentale nella diffusione dell'algoritmo di *backpropagation*, già menzionato, come elemento chiave nell'evoluzione del *deep learning*. In questo caso, il riconoscimento dell'Accademia svedese è rivoluzionario non solo per i contenuti premiati, ma anche per la scelta di sancire in modo autorevole che siamo di fronte a un punto di svolta sul piano epistemologico, in cui viene sfumata una linea di confine che tradizionalmente esisteva tra due discipline – la “*computer science*” e la fisica – quando esse si muovono su temi di IA, terreno in cui si influenzano e arricchiscono reciprocamente, procedendo insieme in un dialogo profondo e comune attorno a un tema di ricerca di assoluta frontiera. Con un cuore che batte nella fisica, quindi, la ricerca sull'IA viene ormai sempre meno percepita come una disciplina informatica o ingegneristica in senso stretto o esclusivo, e si tende invece a evidenziarne e rafforzarne sempre più i ricchi e fertili aspetti multiculturali e multidisciplinari. Le basi teoriche dell'IA affondano

le proprie radici in modo profondo nella fisica, in particolare nella fisica statistica e nella teoria dei sistemi complessi (vd. p. 10, ndr). Molti metodi che sono stati sviluppati per affrontare problemi in meccanica statistica, ad esempio, sono oggi strumenti chiave negli algoritmi di ottimizzazione e nell'addestramento delle architetture di reti neurali moderne. Le reti neurali stesse possono essere interpretate come sistemi complessi ad alta dimensionalità, in cui emergono proprietà collettive non previste dal comportamento dei singoli elementi, un ambito che la fisica conosce e studia da decenni.

Il legame tra fisica e intelligenza artificiale non è solo teorico: è diventato sempre più operativo e strategico. In campi come la fisica delle alte energie, l'astrofisica, la fisica dei materiali o la fisica medica, l'IA è ormai uno strumento indispensabile per analizzare ingenti volumi di dati, individuare *pattern* nascosti, identificare comportamenti anomali, simulare fenomeni complessi e – in ultima analisi – accelerare l'interpretazione sperimentale. Ma è sempre bene ricordare che la fisica non è soltanto un “cliente” dell'IA, ovvero uno dei domini applicativi di tecniche basate su reti neurali artificiali, ma è anche una delle sorgenti più profonde di ispirazione, innovazione e metodologia per la ricerca sull'IA.

d.
A sinistra John Hopfield e a destra Geoffrey Hinton, premi Nobel per la fisica del 2024 per le loro scoperte e invenzioni fondamentali che permettono l'apprendimento automatico con reti neurali artificiali.

Biografia

Daniele Bonacorsi è professore di fisica sperimentale all'Università di Bologna. Lavora da tempo in esperimenti al CERN con acceleratori di particelle. Ha lavorato alla progettazione, sviluppo e messa in opera della Grid per gli esperimenti di LHC ed è stato coordinatore internazionale del calcolo scientifico dell'esperimento CMS a LHC, dove ha promosso l'adozione di tecniche di *machine learning* e *deep learning*. Tiene corsi di *deep learning* e *quantum machine learning* e coordina il dottorato in *data science and computation* all'Università di Bologna.

[as] IA in pillole

MACHINE LEARNING (APPRENDIMENTO AUTOMATICO)

Una classe di algoritmi che consente ai computer di apprendere schemi e regole decisionali a partire dai dati, senza la necessità di una programmazione esplicita basata su regole. Questi sistemi sono in grado di generalizzare dagli esempi osservati per effettuare previsioni o prendere decisioni su input mai visti prima.

DEEP LEARNING (APPRENDIMENTO PROFONDO)

Un sottoinsieme dell'apprendimento automatico, basato su reti neurali artificiali con molti "strati" (da cui il termine "deep") di unità computazionali ("neuroni" artificiali). Le architetture di *deep learning* si sono dimostrate particolarmente efficaci nell'elaborazione sia di dati strutturati che non strutturati, come immagini, testo e audio.

SUPERVISED LEARNING (APPRENDIMENTO SUPERVISIONATO)

Una tecnica di apprendimento automatico, in cui un algoritmo viene addestrato su un insieme di dati etichettati, ossia per i quali è nota la risposta corretta ("label"). L'obiettivo è imparare a prevedere l'etichetta corretta per nuovi dati mai visti in precedenza. Si differenzia da "*unsupervised learning*", in cui i dati non hanno etichette e in fase di allenamento si cerca di individuare strutture o raggruppamenti nei dati stessi.

ALGORITMO DI BACKPROPAGATION

Un algoritmo di apprendimento supervisionato usato per addestrare reti neurali artificiali. Calcola il gradiente della funzione di costo rispetto a ciascun peso della rete applicando la cosiddetta "chain rule", propagando l'errore a ritroso dall'output verso gli strati precedenti per aggiornare i pesi tramite algoritmi di ottimizzazione come il "*gradient descent*" (discesa del gradiente).

CONVOLUTIONAL NEURAL NETWORK (CNN O CONVNET, RETE NEURALE CONVOLUZIONALE)

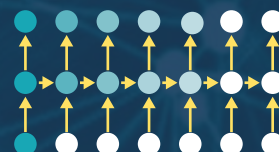
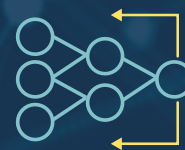
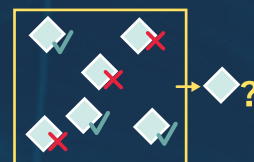
Un tipo di rete neurale profonda particolarmente efficace nell'elaborazione di dati strutturati spazialmente, come le immagini. Le CNN applicano filtri convoluzionali a regioni locali dell'input, sfruttando gerarchie spaziali e l'invarianza rispetto alla traslazione. Hanno avuto un enorme successo nella visione artificiale, ma mostrano anche notevole flessibilità se applicate a dati di natura diversa dalle immagini.

RECURRENT NEURAL NETWORK (RNN, RETE NEURALE RICORRENTE)

Una classe di reti neurali con memoria interna, in cui l'output a ogni passo dipende sia dall'input corrente sia dallo stato precedente, permettendo di catturare dipendenze temporali. Sono adatte all'elaborazione di sequenze di lunghezza variabile, alle previsioni su serie temporali o nella modellazione del linguaggio naturale (scritto o parlato).

LONG SHORT-TERM MEMORY (LSTM)

Una particolare architettura di RNN progettata per superare il cosiddetto "problema del gradiente evanescente" nell'addestramento delle reti ricorrenti. Le LSTM incorporano meccanismi di regolazione del flusso dell'informazione, permettendo loro di apprendere dipendenze in serie di dati sequenziali di lunghezza superiore rispetto alle RNN.



AUTOENCODER (AE)

Una rete neurale addestrata a ricostruire il proprio input dopo averlo compresso, in modo non supervisionato, in una rappresentazione latente a dimensionalità ridotta. Gli Autoencoder sono utilizzati per l'apprendimento auto-supervisionato di rappresentazioni, nella riduzione dimensionale e nella rilevazione di anomalie.



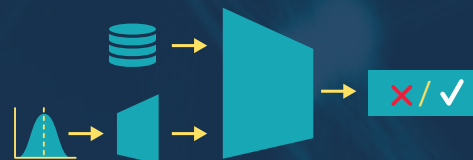
ANOMALY DETECTION (RILEVAZIONE DI ANOMALIE)

Si tratta dell'identificazione di dati che si discostano in modo significativo dal resto del *dataset*, indicando potenzialmente eventi rari e/o inaspettati. Ha ampie applicazioni in tutti i contesti in cui tali “outlier” (osservazioni che si discostano in modo statisticamente significativo dal resto dei dati) possono essere portatori di informazioni rilevanti.



GENERATIVE ADVERSARIAL NETWORK (GAN, RETE GENERATIVA AVVERSARIA)

Un modello generativo composto da due reti neurali contrapposte: un generatore che produce dati sintetici e un discriminatore che ne valuta l'autenticità in base a informazioni in suo possesso sui dati reali. Attraverso un processo di addestramento “avversario”, il generatore impara a produrre dati indistinguibili da quelli reali.



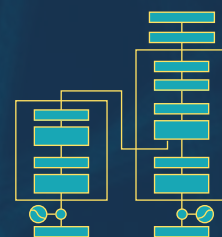
DIFFUSION MODEL (MODELLO DI DIFFUSIONE)

Una classe di modelli generativi che imparano a invertire un processo graduale di aggiunta di rumore ai dati. Modellando la traiettoria di “denoising” (rimozione del rumore, ovvero l'eliminazione delle componenti casuali, irregolari o indesiderate che mascherano l'informazione vera e utile), i modelli di diffusione sono in grado di generare campioni ad alta qualità, superando spesso le GAN in termini di diversità dei campioni e stabilità dell'addestramento.



TRANSFORMER

Un'architettura di rete neurale basata su meccanismi di “self-attention”, che elimina la necessità di ricorrenza e convoluzioni. Sviluppati inizialmente per l'elaborazione del linguaggio naturale, i Transformer sono oggi “foundation models” (modelli di IA generalisti, addestrati su scala estremamente ampia, e progettati per essere versatili) per molteplici modalità, tra cui video, audio e dati tabulari.



SELF-ATTENTION MECHANISM (MECCANISMO DI AUTO-ATTENZIONE)

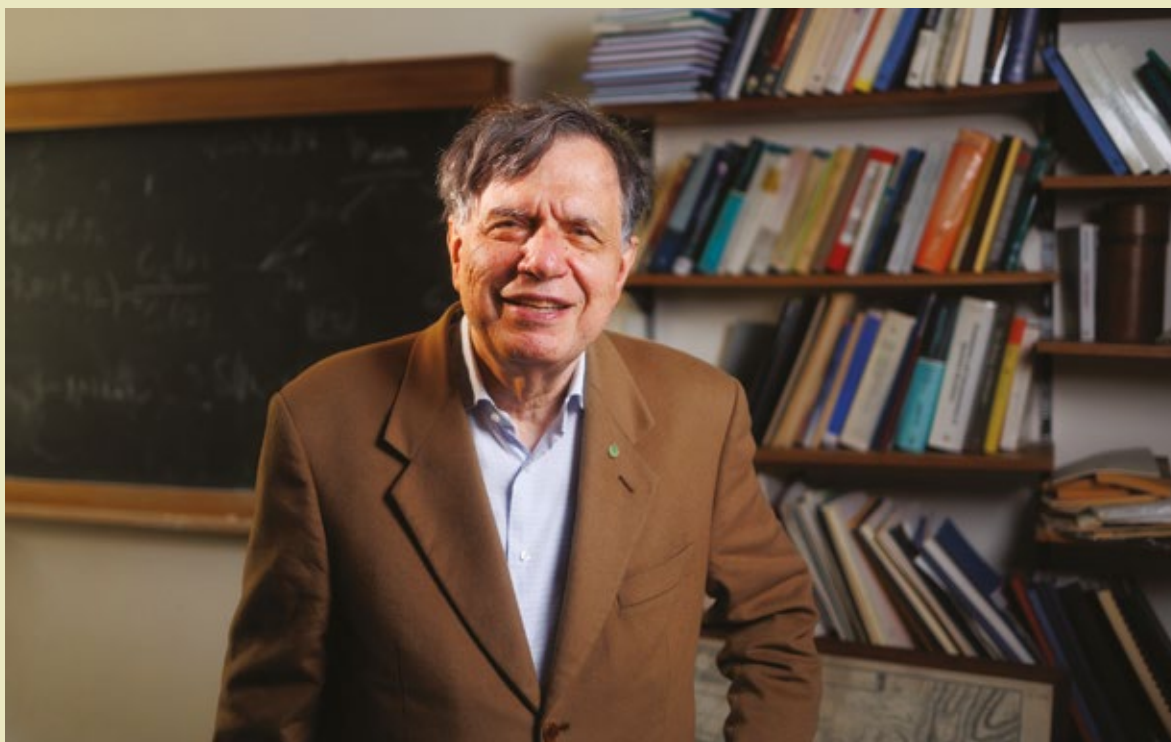
Una componente computazionale che consente a un modello di valutare l'importanza relativa delle diverse posizioni dei dati in una sequenza durante la codifica delle informazioni stesse. Permette la costruzione di rappresentazioni dinamiche e contestualizzate, in cui ciascun elemento dell'input può tenere conto, in modo esplicito e differenziato, di tutti gli altri elementi della sequenza.



Cogito ergo \sum

Il ruolo della meccanica statistica nelle reti neurali

di Elena Agliari e Adriano Barra



Gli anni '70 dello scorso secolo sono stati storiograficamente catalogati come l'inverno ("winter time") dell'IA. Proviamo a comprenderne le ragioni e ad analizzare come questa apparente stagnazione abbia in realtà favorito una profonda maturazione del campo, enfatizzando il ruolo che la fisica ha svolto nel condurre la ricerca verso una nuova primavera.

Nel 1970 erano ormai lontane le prime decadi del secolo in cui Ivan Pavlov costruiva le leggi teoriche della fisiologia, studiando il riflesso condizionato nei cani, mentre Santiago Ramon y Cajal e Camillo Golgi sviluppavano tecniche sperimentali per mettere in luce la morfologia del singolo neurone. Si erano anche già succeduti i primi modelli matematici di neurone, celebre quello di McCulloch e Pitts del '43, ed era già comparso il percettore di Rosenblatt

nel '58, una macchina che ne imita le caratteristiche computazionali. Semplificando, possiamo schematizzare un neurone come un interruttore, che alterna uno stato quiescente – quando la corrente ad esso afferente dai neuroni adiacenti non raggiunge una prestabilita soglia – e uno attivo, in cui tale soglia viene raggiunta e gli permette l'emissione di un impulso elettrico veicolato verso i neuroni ad esso connessi (vd. approfondimento). La natura elettrica della comunicazione neuronale e i primi modelli matematici si innestavano perfettamente nella logica di Boole (formalizzata a metà dell'800), portando a concepire, nella prima metà del secolo scorso, l'elaborazione dell'informazione nelle reti neurali come l'esecuzione di protocolli deterministici di operazioni logiche. Questa idea andò in frantumi nel '69,

a.
Il premio Nobel Giorgio Parisi, padre della meccanica statistica dei sistemi complessi (tra i cui esempi più celebri si annoverano le reti neurali).

quando Marvin Minsky e Seymour Papert mostrarono che il percettrone di Rosenblatt non era in grado di riprodurre tutte le funzioni logiche di base, congelando gli entusiasmi della comunità scientifica.

Il conseguente *winter time* segnò, in realtà, una profonda rivoluzione culturale che consentì, tacitamente, la transizione dal primo al secondo programma per l'IA. Infatti, si realizzò presto che, mentre un singolo percettrone non è in grado di simulare tutte le funzioni logiche fondamentali, un insieme di pochi percettroni opportunamente connessi sa farlo, e questo spostava l'attenzione dal soggetto (il neurone) alla rete delle sue interazioni (la rete neurale, vd. approfondimento), dando avvio a un significativo ammodernamento del bagaglio metodologico-scientifico sottostante alla loro trattazione.

La logica matematica e la neonata cibernetica venivano affiancate da discipline in grado di cogliere le proprietà emergenti in grandi "collezioni" di elementi semplici, come la meccanica statistica, l'inferenza statistica (un insieme di metodi che permettono di trarre conclusioni su una popolazione più ampia basandosi sui dati raccolti da un campione) e la teoria dei grafi. Per enfatizzare la portata del cambiamento di paradigma, per dirla con il filosofo Thomas Kuhn, notiamo che la (storicamente celebre) rivoluzione copernicana del passaggio dal sistema geocentrico al sistema eliocentrico, spostava l'attenzione da un soggetto (la Terra) a un altro (il Sole), mentre nel cambio che stiamo discutendo il focus si sposta dai soggetti alle loro interazioni: un significativo salto concettuale tanto che, ad oggi, quasi la totalità della Natura è stata ripensata in termini di reti di interazioni.

È importante ricordare che la meccanica statistica, nata sul chiudersi della Belle Époque e già ben consolidata negli anni '70, sapeva trattare esaurientemente reti di elementi binari (i.e. "spin a due stati") in interazione semplice (ad esempio, spin che tendono ad allinearsi se l'agitazione termica lo consente), mentre per reti "frustrate" (in cui coesistono interazioni "cooperative" che favoriscono l'allineamento tra due spin e interazioni "anti-cooperative", che ne favoriscono

l'anti-allineamento) non esistevano ancora gli strumenti teorici e matematici necessari per affrontarne lo studio. Il giusto riferimento modellistico nella meccanica statistica era però proprio quello delle reti frustrate poiché, interpretando i nodi della rete (i.e. gli spin) come neuroni a due stati (*on/off*) e le connessioni tra essi come sinapsi, era imprescindibile considerare modelli con interazioni sia cooperative che anti-cooperative, dato che nelle reti neurali esistono sia sinapsi eccitatorie che inibitorie. Da questo punto di vista, gli anni '70 furono una vera fucina di idee e di concetti preliminari all'edificazione di una meccanica statistica capace di descrivere quantitativamente il comportamento di tali sistemi complessi. Un decisivo punto di sintesi si raggiunse a cavallo tra gli anni '70 e '80, quando Giorgio Parisi sviluppò la sua celebre teoria che gli valse il premio Nobel per la fisica nel 2021, nella quale la compresenza di stimoli cooperativi e anti-cooperativi genera un "paesaggio" altamente non banale di minimi dell'energia. La meccanica statistica aveva così subito una trasformazione radicale, che si manifestava su più fronti. In primis, l'introduzione delle reti frustrate ne ampliò lo spettro applicativo ben oltre la fisica, suo ambito originario. In secondo luogo, venne codificato un nuovo vocabolario di osservabili e concetti (ad esempio "overlap" (sovrapposizione), "replica", "simmetria di replica" e sua "rottura") oggi di uso comune in informatica teorica e indispensabili per l'inferenza statistica in spazi ad alta dimensionalità, entrambi pilastri metodologici della moderna intelligenza artificiale.

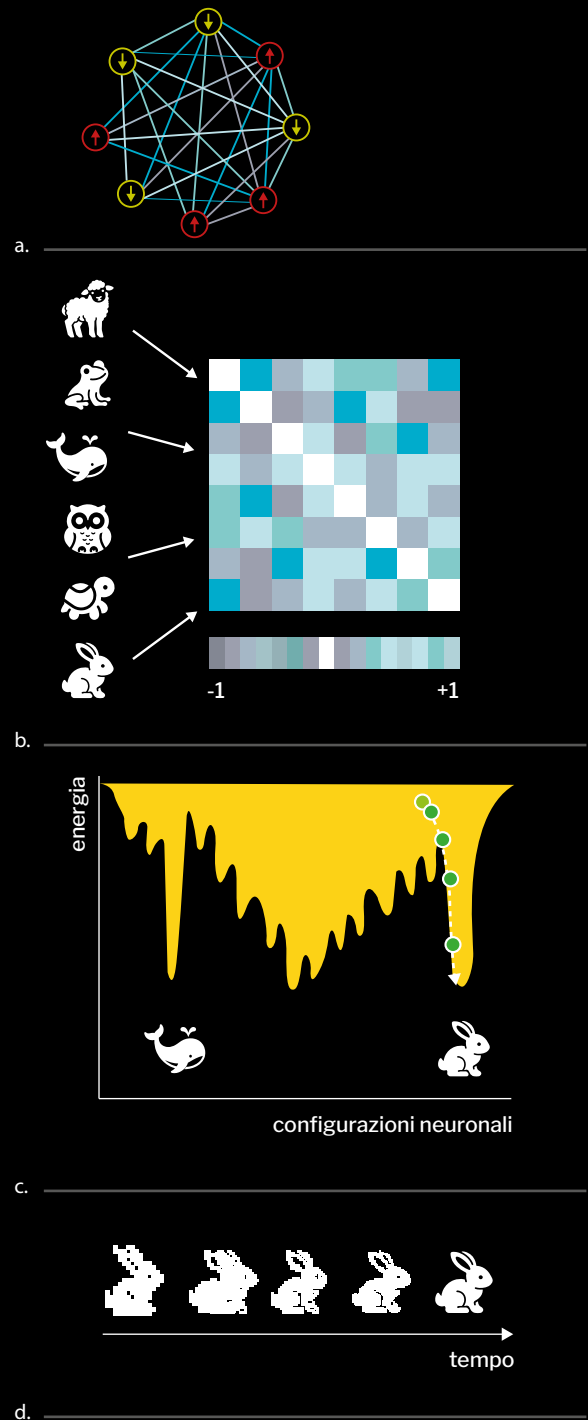
Non si dovette quindi aspettare molto per veder nascere i primi modelli di rete neurale, da quello proposto da Hopfield nel 1982 (che implementava le idee sull'apprendimento di Donald Hebb, a loro volta sintesi degli studi di Pavlov), alla "macchina di Boltzmann", il prototipo formulato da David Ackley, Geoffrey Hinton e Terrence Sejnowski nel 1985 (che sintetizza il ruolo egemone della meccanica statistica già nel suo nome, facendo riferimento a Ludwig Boltzmann). A cascata, tra il 1985 ed il 1987, Daniel Amit, Hanoch Gutfreund



b. Daniel Amit, pioniere delle reti neurali. Haim Sompolinsky e Hanoch Gutfreund e lo stesso Amit negli anni '80 furono i primi a usare la meccanica statistica delle reti frustrate per descrivere esaurientemente le proprietà di processazione d'informazione delle reti neurali, e in particolare nel modello di Hopfield. Amit è stato professore del corso sulle reti neurali alla Sapienza Università di Roma fino al 2007.

Il modello di Hopfield

Un sistema di unità interagenti può essere visualizzato attraverso un grafo, ovvero una rete in cui i nodi rappresentano le unità (per esempio neuroni) e le connessioni indicano l'esistenza di un'interazione (per esempio la presenza di una giunzione sinaptica). In fig. 1a è mostrata una rete neurale, in cui i neuroni evidenziano la loro modellizzazione mediante variabili binarie (i.e. "spin up/spin down"). Nel modello di Hopfield le intensità delle interazioni tra i neuroni sono determinate in maniera tale che la rete manifesti proprietà collettive che riproducono il funzionamento di una memoria associativa, ovvero la capacità di ricostruire un'informazione quando questa viene mostrata in maniera parziale o rumorosa. Le informazioni da memorizzare, per esempio una serie di immagini in bianco e nero, anche dette *pattern*, vengono opportunamente sommate (da cui è nato il titolo di questo articolo, dove "\sum" non è un errore ma indica l'operazione della "somma" in LaTeX, il linguaggio con cui gli scienziati scrivono i loro articoli), in modo da ottenere la "matrice sinaptica", ovvero una matrice i cui ingressi corrispondono alle intensità dell'interazione tra i nodi corrispondenti. Per esempio, in fig. 1b l'interazione tra il neurone 3 ed il neurone 5, corrispondente all'ingresso (3,5), è negativa e relativamente intensa (con un valore vicino a -1, secondo la scala di colore indicata nella parte bassa). Data una rete e specificate le interazioni tra i suoi elementi, è possibile associare ad ogni configurazione neuronale, che specifica quali neuroni sono attivi e quali neuroni sono a riposo, un'energia ovvero un "costo", ottenendo un "paesaggio energetico" come quello mostrato in fig. 1c. Il fatto che la matrice sinaptica abbia sia ingressi positivi che negativi definisce tale rete come frustrata e, di conseguenza, il paesaggio dell'energia ad essa associata appare corrugato, denso di minimi. Per far funzionare la rete è ora necessario introdurre una regola evolutiva per lo stato dei neuroni che assicuri la convergenza spontanea della rete da uno stato di input (p.es. l'immagine corrotta del coniglio) allo stato di output desiderato (l'immagine pulita del coniglio). Sotto opportune condizioni, la matrice sinaptica descritta prima realizza un panorama energetico tale che i minimi corrispondono a configurazioni neurali che riproducono i vari *pattern* incamerati. La regola evolutiva sarà allora particolarmente semplice: muoviti nella direzione che riduce il valore della funzione costo, cioè aggiorna lo stato di un neurone solo se ciò comporta una riduzione energetica. Un esempio di traiettoria è riportato in fig. 1c, insieme alla corrispondente ricostruzione del *pattern* "coniglio" in fig. 1d.



1. a) Grafo rappresentante una rete neurale composta da otto neuroni. b) *Pattern* da memorizzare e relativa matrice sinaptica con scala delle intensità. c) Panorama energetico della rete neurale. d) Evoluzione temporale della configurazione neurale, che realizza la ricostruzione di uno dei *pattern* immagazzinati nella matrice sinaptica.



C. Francesco Guerra, recentemente scomparso, è stato lo scienziato che, lavorando all'intersezione tra la fisica teorica e la matematica, ha sistematicamente gettato le basi metodologiche per rendere la teoria di Parisi matematicamente impeccabile, contribuendo in maniera cruciale alla sua riformulazione in termini di teoremi matematici.

e Haim Sompolsky pubblicavano la prima trattazione sistematica, interamente meccanico-statistica, di queste reti neurali, proprio sfruttando i concetti e le tecniche dei sistemi frustrati di Parisi: capito come associare ai minimi dell'energia le informazioni da immagazzinare, i.e. i *pattern* (vd. fig. 1a), queste reti potevano funzionare come memorie associative nelle quali richiamare un *pattern* (p.es. un'immagine) precedentemente appreso. Questo significava raggiungere spontaneamente uno degli stati di minimo (o di equilibrio) a seguito di un opportuno stimolo: tale cruciale capacità è diventata celebre con il nome di “*pattern recognition*” (vd. fig. 1b).

La meccanica statistica si consolidava così come disciplina cardine del settore, consacrata, oltre al Nobel a Parisi nel 2021, dal Nobel per la fisica conferito lo scorso anno a Hopfield e Hinton per i loro studi pionieristici su questi sistemi. Inoltre, vale la pena ricordare che, nel versante puramente matematico, l'importanza di questi studi è stata suggellata con l'assegnazione, sempre nel 2024, del premio Abel a Michel Talagrand, anche per aver dimostrato rigorosamente la

correttezza della formulazione parisiense dei sistemi complessi. Una curiosità storica riguarda la fisica teorica napoletana: precursore del modello di Hopfield è il modello di Shun'ichi Amari, un fisico giapponese sovente invitato da Eduardo Caianiello, fisico teorico napoletano, a spendere mesi in Costiera Amalfitana negli anni '60. Caianiello ed Amari si influenzarono reciprocamente e anticiparono i primi concetti sulle future teorie delle reti neurali. Tre decenni dopo, fu Francesco Guerra, di cui Caianiello fu mentore accademico, a costruire una serie di tecniche matematiche che permisero di dimostrare rigorosamente la teoria di Parisi, aprendo *de facto* la strada al premio Abel a Talagrand.

L'ampio intervallo di tempo che intercorre tra la formulazione di questi modelli e il conferimento del Nobel ai loro autori non deve sorprendere: questo è dipeso dal fatto che, quando la teoria venne edificata, non era sostanzialmente possibile metterla alla prova. Fino a un paio di decenni fa, non esisteva un mondo digitale in cui le reti neurali artificiali potessero apprendere.

Biografie

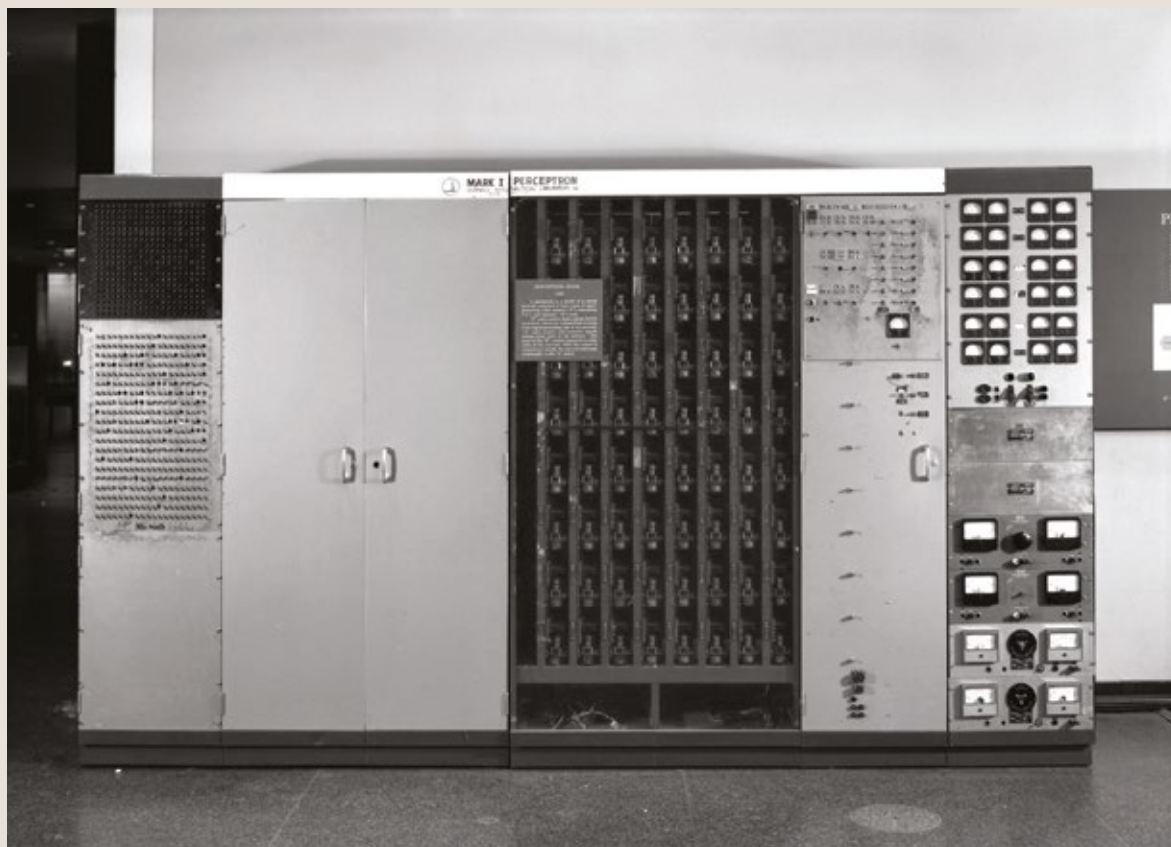
Elena Agliari, laureata e dottorata in fisica all'Università di Parma, è professoressa in fisica matematica presso il Dipartimento di Matematica dell'Università Sapienza di Roma, dove insegna modelli di reti neurali e matematica per il *machine learning*. Si è occupata di processi di diffusione su grafi, modelli matematici applicati alle reti immunitarie e di meccanica statistica dei sistemi disordinati.

Adriano Barra, laureato in fisica presso la Sapienza Università di Roma e dottorato in matematica presso la King's College University of London, è professore associato in fisica matematica presso il Dipartimento di Scienze di Base e Applicate per l'Ingegneria (SBAI) della Sapienza, dove tiene un corso di reti neurali e *machine learning* per il dottorato. La sua ricerca verte prevalentemente su meccanica statistica di reti neurali e sistemi disordinati e modellizzazione matematica di fenomeni biologici.

Rovescio vincente

Backpropagation e reti neurali

di Matteo Osella



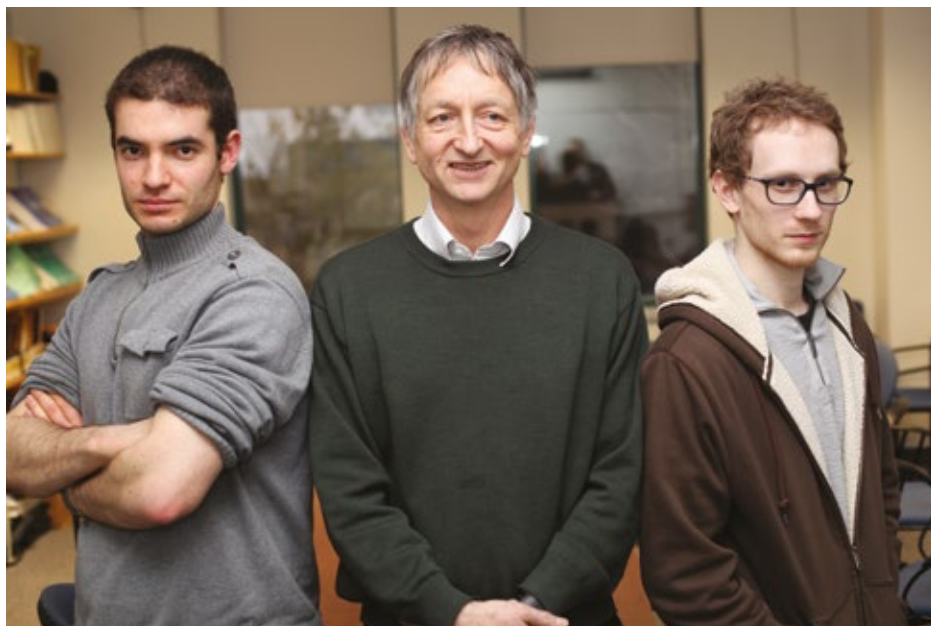
a.
Perceptron Mark I, la prima macchina costruita per mettere in pratica l'algoritmo del percettrone, il primo modello di neuroni artificiali. Era collegata a una fotocamera dotata di 400 sensori di luce (disposti in una griglia 20×20), in grado di catturare immagini semplici. Dietro le due ante nella parte sinistra della fotografia è alloggiato un pannello con cavetti, usato per collegare in modo diverso i segnali in ingresso e "insegnare" alla macchina a riconoscere modelli. Nella parte centrale, che assomiglia un po' a una griglia, si trovano dei "potenziometri" (piccoli componenti che permettono di regolare un potenziale elettrico), che servono a modificare i "pesi" dell'apprendimento, ovvero quanto ogni segnale influenzava la decisione finale.

Fin dagli anni '40, con la nascita del primo modello di neurone, il percettrone, si è fatta strada l'idea che funzioni cognitive complesse potessero emergere da reti di molte unità elementari interconnesse. In quest'ottica, le reti neurali rappresentano un esempio paradigmatico di sistema complesso: migliaia, milioni di unità semplici la cui dinamica collettiva dà luogo a comportamenti sofisticati, non riducibili alla somma delle parti.

Non sorprende quindi che la fisica statistica, con il suo bagaglio di concetti e tecniche sviluppate per modellizzare il comportamento collettivo di sistemi disordinati e interagenti, si sia presto interessata alle reti neurali. Esistono numerose connessioni formali e concettuali con la fisica statistica. Per esempio, nell'apprendimento supervisionato si conosce

l'output desiderato per un insieme di esempi, e su questi si cerca di minimizzare l'errore commesso dalla rete. Questo errore viene rappresentato da una "funzione costo", analoga all'energia in un sistema fisico. Infatti, come in fisica statistica si cercano le configurazioni di minima energia di un sistema, così durante l'apprendimento cerchiamo i minimi della funzione costo, corrispondenti a configurazioni di parametri che meglio spiegano i dati. Tuttavia, come in molti sistemi disordinati, il panorama energetico può presentare una geometria complessa e frastagliata, che l'algoritmo di apprendimento deve esplorare alla ricerca di "buoni" minimi, ovvero configurazioni in grado di fornire previsioni accurate anche su dati non visti.

Negli anni '80, fisici come Giorgio Parisi, Marc Mézard, Bernard



b.
Il professore Geoffrey Hinton al centro, con i suoi studenti Ilya Sutskever, a sinistra, e Alex Krizhevsky, a destra, all'Università di Toronto (Canada) nel 2013.

Derrida e Haim Sompolsky (per citarne alcuni), facendo leva su queste numerose analogie, iniziarono a costruire una “meccanica statistica” delle reti neurali. Negli stessi anni, nuovi modelli, esplicitamente ispirati alla fisica, per esempio dei vetri di spin (vd. in *Asimmetrie* n. 32 p. 12, ndr), vennero proposti. Un celebre esempio è il modello di memoria associativa di Hopfield (vd. approfondimento a p. 12, ndr).

Tuttavia, per lungo tempo, le reti neurali artificiali rimasero ai margini in termini di applicazioni pratiche. Erano, e sono tuttora, utilizzate nelle neuroscienze computazionali come possibile modello semplificato del funzionamento del cervello, anche se la maggior parte degli addetti ai lavori concordano sulla distanza abissale esistente tra queste due diverse incarnazioni di “intelligenza”. Analogamente, seppur le reti neurali fossero un attivo ambito di ricerca in *computer science*, avevano ancora uno scarso impatto pratico. L'ingegneria e l'industria preferivano soluzioni più robuste e controllabili. Il punto di svolta arrivò nel 2012, quando un lavoro di Geoffrey Hinton in collaborazione con Alex Krizhevsky e Ilya Sutskever (futuro cofondatore di OpenAI e figura chiave nello sviluppo di ChatGPT) mostrò che le reti profonde (o *deep*, da cui il termine *deep learning*), se addestrate con grandi quantità di dati e risorse computazionali, potevano superare ogni precedente metodo nel riconoscimento di immagini.

Gli ingredienti fondamentali erano già in circolazione: l'algoritmo di *backpropagation* per l'addestramento di reti profonde era stato sviluppato già negli anni '80, così come l'intuizione di costruire reti specializzate per l'analisi di immagini – le reti convoluzionali – che incorporavano il concetto di località. In analogia con la corteccia visiva del cervello, i neuroni nelle reti convoluzionali si specializzano nell'identificazione di *pattern* in regioni locali dell'immagine, contribuendo alla costruzione di una rappresentazione gerarchica che conserva le relazioni spaziali. Tuttavia, mancava ancora la giusta combinazione di questi elementi e l'hardware

adeguato (l'utilizzo delle GPU) per arrivare a un modello efficiente e addestrabile su grandi quantità di dati.

Da allora, le reti neurali sono entrate nella nostra vita quotidiana: nei nostri telefoni, negli algoritmi di raccomandazione, nella diagnostica medica, nei processi industriali. Negli ultimi anni abbiamo assistito anche alla rapida ascesa dei modelli generativi di linguaggio, addestrati semplicemente a predire la parola successiva in una frase basandosi su enormi raccolte di testi (circa 300 miliardi di parole nel caso di ChatGPT-3), ma capaci di sviluppare funzioni complesse e in parte inattese.

Eppure, mentre l'ingegneria dell'apprendimento automatico progredisce a ritmi straordinari, la teoria resta indietro. Oggi costruiamo modelli sempre più grandi, energivori e sofisticati, ma non possediamo ancora una vera comprensione del perché funzionino. Una teoria dell'intelligenza artificiale con la solidità predittiva della meccanica o della termodinamica ancora non esiste. E il concetto stesso di “intelligenza” resta sfuggente, privo di una definizione operativa condivisa e rigorosa. La situazione ricorda quella del XVIII secolo: l'utilizzo dei motori a vapore stava aprendo le porte alla rivoluzione industriale che avrebbe trasformato il mondo, ma la termodinamica non era ancora stata del tutto sviluppata. Solo con la formalizzazione teorica si poté comprendere il funzionamento di quei motori e guidare l'innovazione tecnologica verso forme più efficienti. Allo stesso modo, oggi ci troviamo ad applicare modelli potenti di apprendimento senza una teoria che ne spieghi a fondo il comportamento, che ne quantifichi i limiti, che ne guidi lo sviluppo futuro.

Esiste un ampio corpo di eleganti contributi teorici che arrivano principalmente dal mondo della matematica, ma spesso si concentrano su modelli fortemente semplificati e sui limiti delle loro performance in scenari “worst-case”, ovvero validi per qualsiasi possibile dato di input, ma poco rilevanti per i casi reali. I dati concreti che ci interessano hanno struttura,



regolarità, correlazioni complesse – proprietà che raramente entrano nelle trattazioni teoriche tradizionali. Inoltre, le architetture moderne sono così sofisticate da essere, almeno per ora, inaccessibili agli strumenti matematici formali di cui disponiamo.

Così si procede per tentativi: metodi empirici, ottimizzazioni *ad hoc*, intuizioni ingegneristiche. Un progresso rapido, ma in parte opaco, quasi “alchemico” come è stato definito in una discussa relazione a una delle più importanti conferenze del settore (NeurIPS).

È in questo scenario che i fisici potrebbero giocare un ruolo cruciale. Per colmare il *gap* attuale tra teoria e pratica serve precisamente ciò che la

fisica fa da sempre: costruire modelli fortemente semplificati, ma capaci di catturare gli ingredienti fondamentali e quindi di essere predittivi, anche a costo di rinunciare al rigore formale matematico. È il momento di cercare leggi generali domandandosi quali aspetti – la struttura dei dati, l’architettura della rete, la dinamica dell’addestramento, la funzione di errore – siano davvero cruciali e quali invece siano dettagli contingenti, frutto di scelte storiche o vincoli tecnologici.

In questo solco sarà forse possibile costruire una “teoria dell’intelligenza” che ci consenta non solo di costruire reti che funzionano, ma di capire anche perché funzionano, quando falliscono e come migliorarle.

c.

La situazione attuale dell’IA ricorda quella del XVIII secolo, quando l’utilizzo dei motori a vapore stava aprendo le porte alla rivoluzione industriale che avrebbe trasformato il mondo, ma la termodinamica non si era ancora sviluppata appieno.

Biografia

Matteo Osella si è laureato in fisica teorica all’Università degli Studi di Torino, dove ha proseguito gli studi con un dottorato di ricerca in sistemi complessi per la biologia. Ha lavorato alla Sorbonne Université di Parigi e ora è professore associato all’Università di Torino. È docente di reti neurali presso il Dipartimento di Fisica e i suoi temi di ricerca riguardano la fisica statistica, la fisica biologica e le reti neurali artificiali.

The imitation game

Apprendimento per rinforzo

di Antonio Celani

Un alieno sbarca sulla Terra. È arrivato in un bel posto di campagna, ma si sente a disagio. Sul suo pianeta tutto è leggero e immateriale. Il corpo umano che indossa per non dare nell'occhio invece è scomodo e ingombrante. Si sente anche spaesato: si era ripromesso di studiare tutto il possibile su questo pianeta durante il lungo viaggio, ma poi ha dormito tutto il tempo.

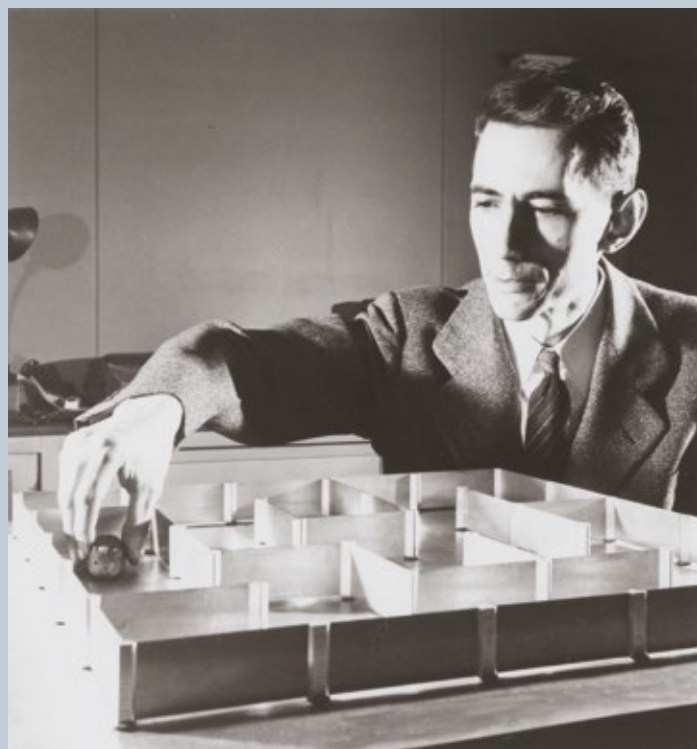
Ma ecco, un umano si avvicina, molto velocemente, in bicicletta. Poi si ferma, scende agilmente dalla bici che appoggia a un cancello, entra in una casa. Incuriosito l'alieno si avvicina. Anche lui vorrebbe muoversi con quella stessa destrezza. Ispeziona la bici e monta in sella. Mentre mette i piedi sui pedali, inizia a inclinarsi su un lato e cade rovinosamente a terra. Dopo averci pensato un po', decide di consultare il *database* collegato alla sua mente e in pochi secondi visualizza un milione di video di persone che vanno in bicicletta. "Ora so perfettamente cosa fare", si dice. Imitando i gesti che ormai gli sono familiari, si rimbecca i pantaloni, risale in sella e inizia a pedalare con entusiasmo. Ma dopo poche vigorose pedalate, purtroppo in senso sbagliato, la bici rimane sul posto e lui cade di nuovo.

"Devo chiedere aiuto a un esperto", pensa. Allora guarda un milione di *tutorial* su come si va in bicicletta, impara tutti i termini tecnici e studia l'equivalente di due dottorati in ingegneria e in fisica. Pensa: "Questa volta non posso sbagliare". E infatti sembra funzionare, la bici si muove, accelera, ma dopo pochi metri il novello ciclista perde l'equilibrio e finisce in un fosso.

Allertata dai rumori, la proprietaria della bicicletta esce di casa e si scaglia contro l'alieno urlando: "Ma cosa sta facendo?". Poi vedendo la sua espressione confusa e imbarazzata, si rende conto che non si tratta di un ladro. "Cosa pensava di fare?" chiede, più incuriosita che arrabbiata adesso. "Imparare ad andare in bicicletta", risponde. "Ho letto tutto quello che c'era da leggere, guardato tutto quello che c'era da guardare e studiato tutto quello che c'era da sapere, ma non riesco". Lei ride di cuore e gli spiega: "Ma signore, non lo sa che certe cose si imparano solo facendole?".

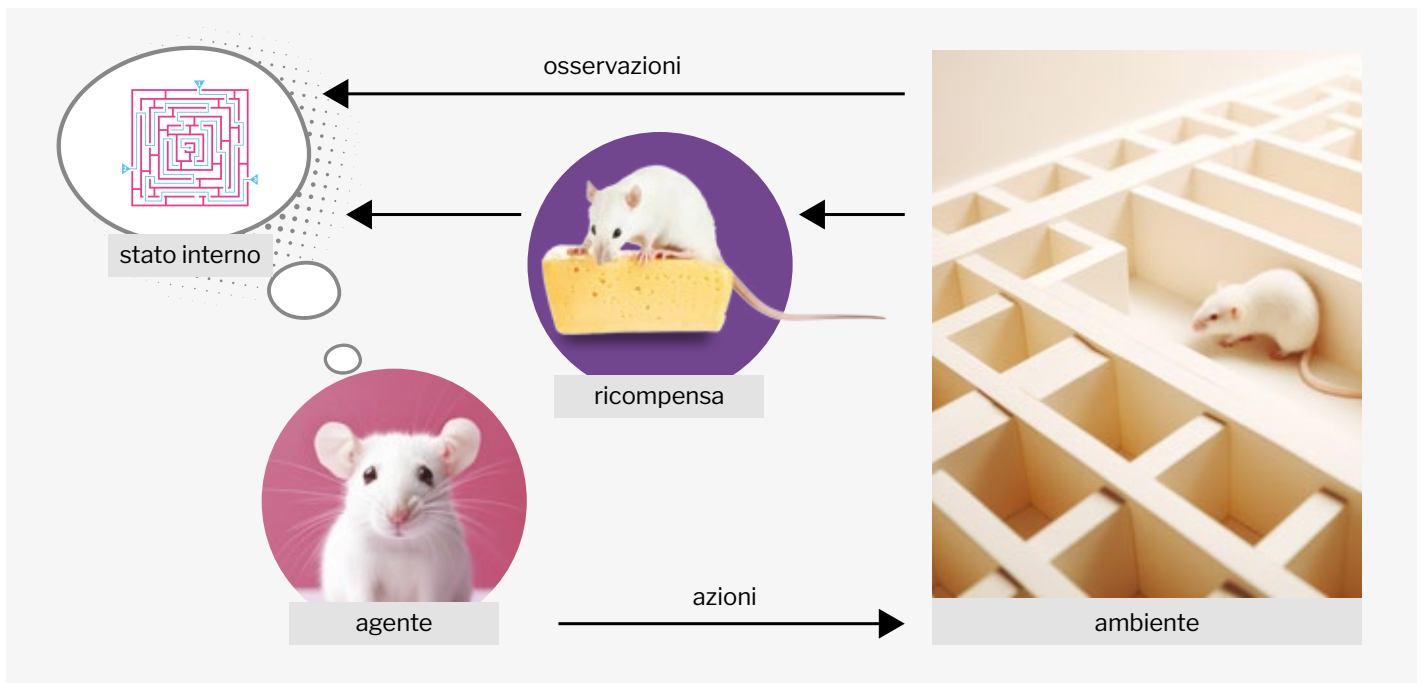
Quello che all'alieno sfugge è l'importanza dell'"apprendimento attivo".

Per noi umani, è un'esperienza talmente naturale che la diamo spesso per scontata. In pedagogia, l'apprendimento attivo considera i bambini protagonisti del proprio apprendimento,



a.

Claude Shannon, il fondatore della teoria dell'informazione con il topo meccanico "Theseus" in un labirinto, uno dei primi esperimenti di apprendimento in robotica. Shannon era anche noto per essere un abile giocoliere e ciclista di monocicli, quindi lui stesso dotato di una notevole "intelligenza incarnata".



b.

L'apprendimento per rinforzo descrive il processo di continua interazione tra agente e ambiente, qui esemplificata da un topolino in un labirinto. L'agente riceve informazioni sotto forma di osservazioni che possono essere in questo caso visive od olfattive. Le osservazioni contribuiscono alla formazione di una rappresentazione interna dell'ambiente, sotto forma di una mappa o anche più astratta. Lo stato interno guida il processo di decisione e determina le azioni, cioè i movimenti del topolino. Durante la navigazione, o anche solo al momento del raggiungimento del traguardo, l'agente riceve ricompense che rinforzano le decisioni favorevoli e scoraggiano quelle sfavorevoli. Dopo molte ripetizioni del compito, il topolino impara la strada più breve che gli permette di raccogliere la maggior quantità di formaggio.

non destinarli passivi di informazioni. La conoscenza viene costruita attivamente tramite l'esperienza diretta e pratica e non semplicemente assorbita. L'esplorazione, la manipolazione e la sperimentazione sono quindi fondamentali. I bambini utilizzano i sensi e le azioni materiali per comprendere il mondo: si tratta di una comprensione intuitiva e non formale della fisica e delle sue leggi, che possiamo chiamare "intelligenza incarnata".

L'apprendimento attraverso l'interazione fisica con l'ambiente non si limita ai bambini ma riveste un ruolo cruciale per tutti gli animali e in ogni stadio della loro vita. Per esempio, è fondamentale per imparare a ottimizzare l'acquisizione di risorse vitali come cibo e acqua, esplorando attivamente nuove fonti e apprendendo le strategie più efficaci per sfruttarle. Così anche per imparare a navigare con efficacia in spazi anche sconosciuti, creando mappe cognitive e scegliendo percorsi ottimali. In questo apprendimento per tentativi ed errori, un ruolo fondamentale è giocato dalla gratificazione ricevuta in risposta a un certo comportamento, che può consistere in una ricompensa materiale o anche in una soddisfazione intrinseca. Il processo attraverso il quale un animale impara comportamenti sempre più efficienti nel raccogliere delle ricompense si chiama "apprendimento per rinforzo" (*reinforcement learning*).

A differenza di altri tipi di apprendimento, i segnali di rinforzo non danno istruzioni esplicite su quale comportamento adottare, ma esprimono una forma di incoraggiamento o di dissuasione piuttosto generica. Inoltre, la gratificazione può arrivare solo in seguito a una lunga e complessa catena di decisioni, spesso caratterizzate dalla mancanza di ricompense intermedie, se non addirittura marcate da esperienze spiacevoli, proprio come quando si impara ad andare in bicicletta. Il *feedback* limitato e la distanza temporale tra le azioni prese e le loro conseguenze a lungo termine rendono difficile identificare le cause prime di successi e fallimenti, rallentando e frustrando il processo di apprendimento.

C'è anche un altro problema. Da una parte c'è la necessità di cercare attivamente informazioni che accrescono la conoscenza dell'ambiente, cioè il processo di esplorazione caratteristico dell'apprendimento attivo. Ma dall'altra parte, l'obiettivo è sempre quello di massimizzare la ricompensa che l'ambiente offre, cioè sfruttare le conoscenze che si sono già acquisite per il proprio scopo. Queste due richieste sono quasi sempre in tensione tra di loro. Se l'alieno si limitasse ad agire unicamente in base alla gratificazione ottenuta, probabilmente desisterebbe molto presto, visti i



risultati deludenti, se non proprio punitivi, dei primi tentativi. Viceversa, se una volta imparato a pedalare con una certa destrezza si abbandonasse al gusto per le più audaci sperimentazioni, magari guidare la bici con il manubrio invertito, la gratificazione ne risentirebbe, per usare un eufemismo. Trovare il giusto equilibrio tra curiosità e produttività è la chiave dell'apprendimento per rinforzo. Quindi, se vogliamo insegnare a una macchina a svolgere ogni tipo di compito che richiede una ripetuta interazione con un mondo esterno, come andare in bici, giocare a ping pong o a un videogioco, dobbiamo affrontare queste difficoltà in modo sistematico.

Qui entra in gioco la teoria dell'apprendimento per rinforzo. Si tratta di un insieme di metodi in cui confluiscono diverse discipline della matematica come la teoria dei processi di decisione, l'ottimizzazione, la statistica e la teoria dell'informazione, nutriti da una corrente sotterranea che attinge dalla psicologia, dalle neuroscienze e dall'economia, che sfociano in algoritmi messi in pratica dagli esperti di informatica e robotica, da ingegneri e da fisici. Questo approccio interdisciplinare permette di descrivere una grande varietà di situazioni di interesse pratico con un unico

linguaggio astratto, in cui nozioni come agente, ambiente, ricompense, osservazioni, strategie hanno un significato ben preciso che si declina poi caso per caso.

L'apprendimento per rinforzo ha diversi punti di contatto con la fisica. Innanzitutto, l'ottimizzazione dei processi di decisione è essenzialmente un problema di controllo di un sistema dinamico. Ad esempio, possiamo pensare alla robotica come una versione (molto) più complicata del problema di scoprire quando e quanta forza applicare a un sistema di punti materiali affinché raggiungano una certa configurazione desiderata in un dato tempo. Anche se non conosciamo tutti i parametri che definiscono questo sistema (masse, forze interne, attriti, perturbazioni, ecc.) possiamo comunque imparare il controllo ottimale per tentativi ed errori grazie all'apprendimento per rinforzo.

A un livello ancora maggiore di astrazione, la tensione tra ricompensa ed esplorazione ricorda il concetto di bilanciamento tra energia ed entropia in meccanica statistica. Questa analogia ha ispirato molti algoritmi di uso corrente in cui il processo di apprendimento è interpretato come un lento "raffreddamento" verso le strategie di decisione ottimali. Inoltre, dal punto di vista teorico, questa

c.
Un singolo algoritmo di apprendimento per rinforzo ha imparato a giocare a livelli sovrumani tre giochi diversi (scacchi, go e shōgi, quest'ultimo visibile in foto), battendo programmi specializzati e campioni del mondo.



connessione fornisce la chiave per comprendere la relazione tra apprendimento e teoria dell'informazione.

L'apprendimento per rinforzo ha già raggiunto traguardi importanti nelle applicazioni dell'intelligenza artificiale. Ha dimostrato abilità superiori a quelle umane in giochi complessi come *go* e *shōgi* e si è distinto nel controllo di robot per manipolazioni precise e compiti complessi, come risolvere il cubo di Rubik. Tra le maggiori aspettative per il futuro c'è sicuramente il suo uso in medicina per la personalizzazione delle cure e l'assistenza alla

chirurgia robotica minimamente invasiva. L'importanza dell'apprendimento per rinforzo nell'ambito dell'intelligenza artificiale è stata recentemente riconosciuta con il conferimento del "Turing award" a Richard Sutton e Andrew Barto, che hanno introdotto a partire dagli anni '80 la maggior parte dei concetti e degli algoritmi alla base di questa disciplina. Grazie ai recenti sviluppi, possiamo dire che la predizione di "macchine che imparano dall'esperienza", che fece Alan Turing nel 1947, sia adesso una realtà concreta. Per gli alieni in bicicletta invece, dovremo ancora aspettare.

d. Richard Sutton (a sinistra) e Andrew Barto (a destra), i pionieri della teoria dell'apprendimento per rinforzo. In mezzo a loro il collega John Moore.

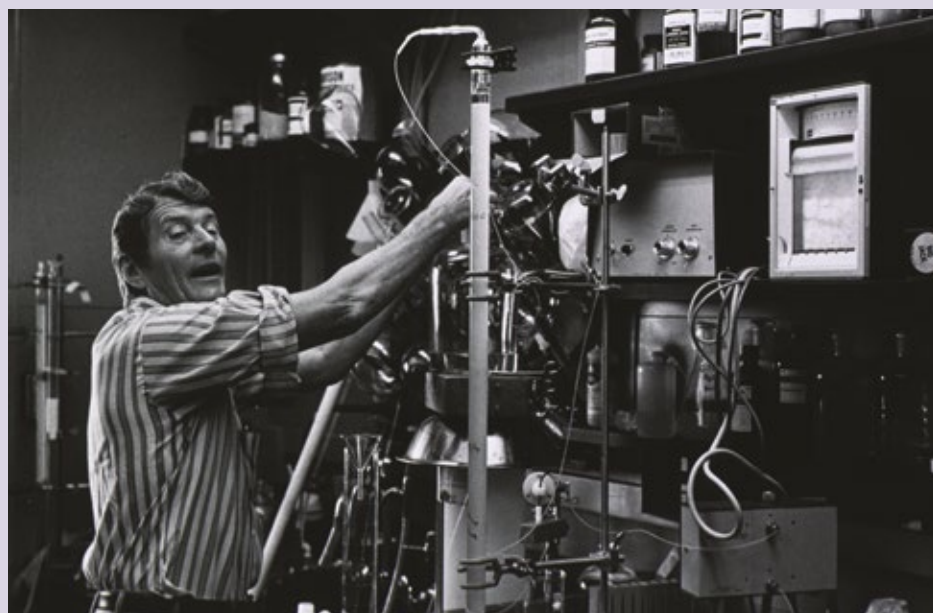
Biografia

Antonio Celani è *senior research scientist* presso l'International Center for Theoretical Physics di Trieste, dove coordina le attività della sezione di Quantitative Life Sciences. Si occupa di apprendimento per rinforzo e delle sue applicazioni al comportamento animale e artificiale.

Facile per la natura, difficile per l'uomo

Il problema del *protein folding*

di Piero Fariselli



a.
Christian Anfinsen, premio Nobel per la chimica del 1972. Dimostrò che l'informazione contenuta nella sequenza di amminoacidi è sufficiente a determinare la struttura tridimensionale di una proteina, e quindi la sua funzione biologica.

Se il DNA è la molecola che contiene le istruzioni per la costruzione e il funzionamento degli organismi viventi, sono le proteine a svolgere la maggior parte delle funzioni biologiche. Codificate nei geni, le proteine vengono tradotte come catene di amminoacidi, formate da una combinazione di 20 unità diverse. In ambiente fisiologico, queste catene si ripiegano spontaneamente in una forma tridimensionale specifica (la struttura "terziaria"), essenziale per la loro funzione. Nel 1972, Christian Anfinsen vinse il premio Nobel per la chimica per aver dimostrato che l'informazione contenuta nella sequenza degli amminoacidi è sufficiente a determinare la struttura tridimensionale di una proteina, la quale assume una geometria spaziale adatta a svolgere le sue funzioni, interagendo con le molecole dell'ambiente circostante. Da allora, ricercatori di tutto il mondo hanno cercato di risolvere il cosiddetto *protein folding problem*: predire la struttura tridimensionale di una proteina a partire dalla sua sequenza, facilmente ottenibile, quest'ultima, dal sequenziamento del DNA. I fisici hanno modellato il problema come una ricerca del minimo di energia libera ma, nonostante la teoria sia formalmente descrivibile

mediante la meccanica quantistica, le soluzioni pratiche sono impossibili da calcolare anche solo approssimativamente, dato l'enorme numero di atomi coinvolti.

Questa difficoltà ha portato alla formulazione del famoso "paradosso di Levinthal": se una proteina esplorasse casualmente tutte le possibili conformazioni, impiegherebbe un tempo astronomico per trovare quella corretta. Eppure, in natura, il *folding* avviene in modo rapido ed efficiente (nell'ordine dei micro- o millisecondi). Arthur Lesk ha riassunto perfettamente il dilemma con la frase: "Facile per la natura, difficile per l'uomo".

Per affrontare questo problema, la comunità scientifica ha sviluppato numerosi approcci computazionali e sperimentali. Nel 1994 è nato il CASP (Critical Assessment of protein Structure Prediction), una competizione biennale che valuta i progressi nella predizione delle strutture proteiche. Durante il CASP, vengono fornite le sequenze di proteine la cui struttura tridimensionale è stata determinata sperimentalmente, ma non ancora resa pubblica. I partecipanti devono predirne la struttura; i risultati vengono poi confrontati con le strutture



b.
David Baker, Demis Hassabis e John Jumper
all'Accademia Reale Svedese delle Scienze durante
la conferenza dei premi Nobel 2024.

reali, una volta rese note. Questa iniziativa ha permesso di monitorare nel tempo l'evoluzione delle tecniche di predizione, mostrando come l'integrazione tra dati evolutivi, modelli fisico-chimici e, più recentemente, intelligenza artificiale abbia progressivamente migliorato le prestazioni.

Il punto di svolta è arrivato nel 2020, quando l'azienda di intelligenza artificiale DeepMind ha presentato AlphaFold2, un sistema basato sul *deep learning*, che ha superato ogni aspettativa. I risultati presentati al CASP14 hanno mostrato una precisione paragonabile, e in molti casi equivalente, a quella di tecniche sperimentali (come la cristallografia a raggi X e la risonanza magnetica nucleare).

Il successo si basa su tre elementi fondamentali: l'utilizzo dell'informazione evolutiva, l'"autoconsistenza" geometrica e il *deep learning* come collante.

AlphaFold2 sfrutta enormi database di sequenze proteiche per costruire allineamenti multipli di sequenze (MSA). Le mutazioni naturali che si sono accumulate nel tempo conservano, attraverso la selezione, le strutture tridimensionali funzionali. Analizzando *pattern* di co-variazione tra residui conservati in famiglie evolutivamente correlate, il modello può dedurre quali amminoacidi tendono a essere vicini nello spazio, anche se lontani nella sequenza. Questa informazione è molto più potente di qualsiasi predizione basata solo sulla fisica.

Una delle intuizioni chiave è stata far "chiudere il cerchio" al modello: AlphaFold2 non predice solo distanze tra residui, ma genera una struttura tridimensionale completa, che viene poi reinterpretata internamente per verificare se è coerente con le informazioni iniziali. Questo processo, chiamato "auto-consistenza", consente al modello di correggersi autonomamente: la struttura generata viene reinserita nel

sistema per un nuovo ciclo di raffinamento, migliorando ogni volta la coerenza globale. Non serve alcun confronto con dati sperimentali durante la predizione. Infine, AlphaFold2 sfrutta un'architettura chiamata Evoformer, che consiste in una rete neurale trasformativa progettata per trattare in parallelo l'MSA e le relazioni tra residui. Questo modello è in grado di integrare contesto evolutivo e geometrico, imparando rappresentazioni complesse che legano la sequenza alla struttura.

Con AlphaFold2, DeepMind ha dimostrato che è possibile predire con elevata precisione la struttura tridimensionale di una proteina partendo dalla sola sequenza degli amminoacidi, sfruttando informazioni evolutive e un sofisticato meccanismo di raffinamento iterativo delle predizioni. Questo straordinario risultato non è passato inosservato: nel 2024 il premio Nobel per la chimica è stato assegnato a Demis Hassabis e John Jumper (AlphaFold) e David Baker (RoseTTAFold), proprio per aver rivoluzionato la predizione strutturale delle proteine attraverso l'intelligenza artificiale.

Abbiamo dunque ottenuto una soluzione efficace al problema, ma possiamo affermare di aver compreso realmente come avvenga il *folding*? Il modello predice accuratamente le strutture, eppure i meccanismi da esso appresi rimangono in gran parte oscuri. Il paradosso è che la comprensione scientifica umana sembra aver ceduto il passo alla straordinaria capacità predittiva delle reti neurali. Forse non è la natura a essere intrinsecamente complessa, ma piuttosto lo è il modo in cui la mente umana cerca di interpretarla. Possiamo dunque concludere che il problema del *folding*, così essenziale per la vita, oggi risulti semplice per la natura, calcolabile per l'intelligenza artificiale, ma ancora irrimediabilmente difficile per l'uomo.

Biografia

Piero Fariselli è professore all'Università di Torino e responsabile dell'unità di intelligenza artificiale e biomedicina computazionale presso il Dipartimento di Scienze Mediche. Da oltre trent'anni si occupa dello sviluppo di modelli di *machine learning* per applicazioni biofisiche e biomediche.

Non ti avevo riconosciuto...

Analisi dei dati e ricostruzione di segnali

di Tommaso Boccali

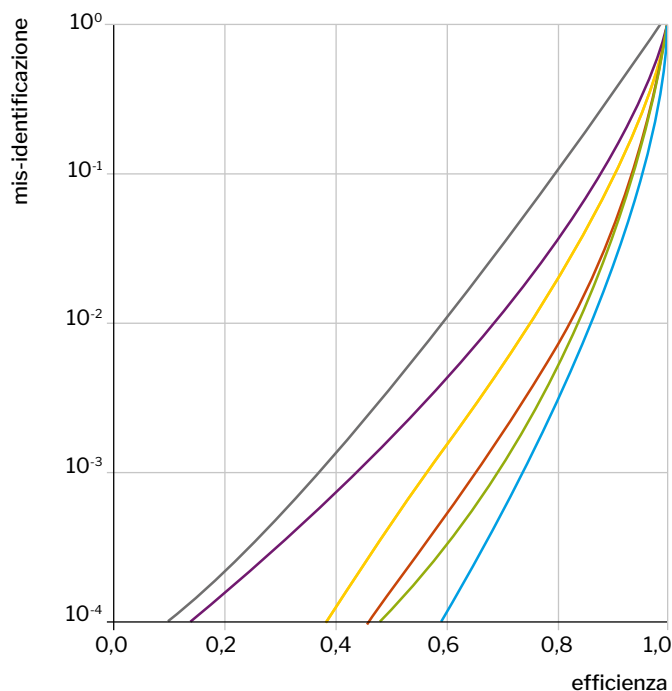


a.
L'esperimento ALEPH del Large Electron-Positron Collider (LEP), l'acceleratore di particelle del CERN di Ginevra negli anni '90 (precursore di LHC). L'utilizzo delle reti neurali per la ricerca di rari decadimenti del mesone B nell'esperimento ALEPH può essere considerato il primo esempio di uso dell'intelligenza artificiale nella fisica delle alte energie.

Il 30 novembre 2022 ChatGPT 3.5 è stato reso disponibile mediante ogni browser web, provocando una presa di coscienza da parte del grande pubblico delle potenzialità dell'intelligenza artificiale.

Nel mondo della ricerca, e della ricerca in fisica fondamentale in particolare, metodi ispirati al funzionamento del cervello umano sono utilizzati da un tempo molto maggiore. È ovviamente difficile risalire a un primo utilizzo, ma possiamo comunque considerare come uno dei primi nella fisica delle alte energie quello descritto nell'articolo di Philippe Rosnet e Pierre Henrard del 1997, sull'utilizzo delle reti neurali per la ricerca di rari decadimenti del mesone B nell'esperimento ALEPH del LEP, in cui viene proposta una struttura a 20 neuroni di input e due strati nascosti, con rispettivamente 15 e 10 neuroni. Il risultato viene condensato in un solo neurone nell'ultimo strato, per un

totale di 486 “pesi”, usati per calibrare la rete neurale. Una rete con tali caratteristiche era, a quei tempi, al limite della tecnologia informatica utilizzabile. Ai giorni nostri, l'utilizzo di grandi infrastrutture di calcolo e di processori dedicati come le GPU permette di addestrare e utilizzare reti con centinaia di miliardi di “pesi”. Insieme a una maggiore comprensione teorica delle reti neurali (in buona parte dovuta a studi che considerano tali reti come “sistemi complessi”, vd. p. 10, ndr) e alla disponibilità di modelli più avanzati, questa esplosione delle capacità tecnologiche è la maggior responsabile della “rivoluzione” IA nella fisica fondamentale. Oggi metodologie IA sono usate in quasi tutti gli ambiti della ricerca in tale campo: dalla simulazione degli eventi, capace di rivaleggiare con simulazioni dettagliate che hanno rappresentato lo standard fino ad oggi, alla ricostruzione delle particelle partendo dai segnali elettronici



— Likelihood ratio
 — Multi Layer Perceptron
 — Deep Neutral Network
 — Convolutional Neutral Network + Recursive Neutral Network
 — Dynamic Graph Convolutional Neutral Network
 — Transformer

b.

Un tipico grafico “mis-identificazione vs efficienza”, che mostra il miglioramento delle prestazioni introdotte dall’intelligenza artificiale nella fisica delle alte energie. La separazione tra eventi di segnale e rumore è tanto migliore quanto più è bassa la mis-identificazione (incorretta identificazione di eventi di fondo) a parità di efficienza nella corretta identificazione di eventi di segnale. Si nota come, man mano che si passa da algoritmi semplici e privi di IA (come Likelihood ratio) ad algoritmi che usano IA avanzata (p.es. i Transformer), la mis-identificazione si riduce sempre di più a parità di efficienza. L’obiettivo è avvicinarsi il più possibile all’angolo in basso a destra del diagramma.

grezzi, alle tecniche di analisi dei dati. Focalizzandoci su quest’ultima categoria, ci sono due esempi che possono essere utili per capire l’impatto dell’intelligenza artificiale nel settore. Uno dei segnali più interessanti nella fisica ai collisori ad alte energie degli ultimi decenni è rappresentato dagli eventi con la presenza di quark b (bottom, o beauty), il secondo quark più pesante dei sei del modello standard delle particelle elementari. La rivelazione di tali quark è un indicatore importante per moltissimi canali di interesse: la fisica dei sapori pesanti, dell’Higgs, del top e in moltissimi scenari oltre il modello standard. Tali quark, prodotti in misura minore rispetto a quark più leggeri, generano adroni con caratteristiche peculiari come una maggiore massa, una maggiore vita media e decadimenti con un numero elevato di particelle. Nessuna di queste caratteristiche è tale da permettere una separazione perfetta dagli eventi di rumore, per cui l’identificazione di eventi originati da quark b è fatta su base statistica. Le prestazioni degli algoritmi di IA si basa su due parametri: la loro efficienza nella corretta identificazione di eventi di segnale, in questo caso i quark b (detta “efficienza”), e alla percentuale di incorretta identificazione di eventi che non li contengono, il cosiddetto “fondo” (detta “mis-identificazione”). A parità di efficienza, l’algoritmo funziona meglio se la mis-identificazione è bassa.

Queste prestazioni possono essere riportate su un grafico bidimensionale come quello di fig. b, in cui l’algoritmo migliore è rappresentato dalla regione in basso a destra. Questo grafico ci mostra anche l’evoluzione di tali algoritmi di selezione per i quark b, che esistono da decenni. Tali algoritmi si sono evoluti passando da semplici selezioni basate su una sola delle caratteristiche sopra elencate a combinazioni sempre più complesse di diverse informazioni. Con l’arrivo dell’intelligenza artificiale, inizialmente con metodi come

Boosted Decision Trees e poi con semplici reti neurali dense (i cosiddetti Multi Layer Perceptron), che combinano un certo numero di caratteristiche in modo ottimizzato, si ottengono discriminazioni sempre migliori.

Arrivano poi tecniche avanzate di IA, come il *deep learning* e reti basate su grafi e i Transformer, che combinano un numero molto alto di caratteristiche, comprese alcune che a prima vista non paiono eccessivamente rilevanti, ma per le quali l’apprendimento automatico riesce a trovare comunque correlazioni tra le variabili utili, anche se piccole. Il risultato migliore ottenuto ad oggi, utilizzando metodi di intelligenza artificiale avanzati è, a parità di efficienza, più di 50 volte migliore del migliore degli algoritmi vecchio stile. Nella pratica questo vuol dire che un’analisi di fisica avrà 50 volte meno fondo e sarà in grado di selezionare segnali molto più puri. In aggiunta, in molti casi, c’è anche la richiesta di avere più quark b nell’evento, anche 2, 3 o 4. In questo caso il guadagno sulla mis-identificazione del fondo può scalare con le potenze di 50 (una per ogni quark b aggiuntivo), e quindi essere molto rilevante ai fini della precisione delle misure. Sono numeri ancora più significativi, con un impatto notevole che ha consentito la realizzazione di misure di fisica estremamente migliori di quanto inizialmente prospettato.

Il secondo esempio riguarda la rivelazione delle onde gravitazionali, uno dei campi di ricerca di più grande interesse, passato dalla fase di scoperta all’essere un efficace strumento per la ricerca astrofisica e cosmologica. I segnali raccolti in esperimenti come Virgo sono molto flebili, al limite della sensibilità sperimentale. In tale regime, la capacità di discriminazione fra segnali indotti da rumori locali legati all’ambiente esterno o ai sottosistemi del rivelatore e segnali reali è essenziale, ancor di più in un’ottica di ricerca multimessaggera, in cui l’allerta sulla



C.
Vista aerea
dell'interferometro
di onde gravitazionali
Virgo, a Cascina (nei
pressi di Pisa).

rivelazione dei segnali gravitazionali deve essere veloce per permettere ad altri esperimenti di orientarsi per una rivelazione congiunta. La pronta classificazione di segnali sperimentali (tipicamente mostrati in un diagramma tempo-frequenza detto “spettrogramma”) come appartenenti a classi di effetti indotti dal rivelatore o da eventi esterni (i cosiddetti “*glitch*”) è essenziale in questo processo ed è un’attività complessa, che non può essere svolta con metodi classici e algoritmici. E l’IA riesce a svolgere un ruolo cruciale anche in questa classificazione. Dal momento che i dati su cui si opera sono immagini (diagrammi tempo-frequenza), l’uso delle reti neurali convoluzionali (CNN) sulle immagini degli spettrogrammi è ideale. Perché l’algoritmo funzioni è necessario un set di immagini con la giusta etichetta che un umano esperto oppure un altro algoritmo possono fornire: evidentemente la qualità finale dell’algoritmo CNN dipende dalla qualità e dalla quantità di tale set. Una possibile soluzione per avere a disposizione

tale set di immagini è appoggiarsi a soluzioni di *citizen science*, nelle quali viene chiesto a una comunità di esperti ed entusiasti di aiutare in tale classificazione, ai fini della preparazione di un set da usare come allenamento per l’algoritmo. I risultati sono molto buoni e il tempo di esecuzione della classificazione è dell’ordine di pochi microsecondi, ottimale per caratterizzare i segnali di *glitch* in modo da studiarne l’origine per poterli rimuovere e poter quindi usare i segnali reali di onde gravitazionali per inviare rapidamente eventuali *alert* da sottoporre agli altri sistemi di rivelazione multimessaggera. Quelli descritti sono solo due esempi di come l’intelligenza artificiale possa essere utile nell’analisi dei dati della fisica fondamentale. In realtà, oggi è raro trovare ambiti di ricerca in cui l’IA non venga già usata o il cui uso non sia previsto per il prossimo futuro. I ricercatori dell’INFN, che lavorano in vari campi scientifici, stanno imparando a conoscere e usare queste tecnologie in modo efficace, così da far progredire la scienza.

Biografia

Tommaso Boccali è dirigente di ricerca presso l’INFN di Pisa. Già durante il suo dottorato presso la Scuola Normale Superiore, si è occupato di problematiche di calcolo scientifico, inizialmente per l’esperimento ALEPH e poi per l’esperimento CMS a LHC. Attualmente è supervisore del progetto CNAF Reloaded, che prevede l’integrazione delle risorse INFN al CNAF presso il Tecnopolo di Bologna, e *spoke 2 leader* del Centro Nazionale Big Data, HPC e Quantum Computing.

Cogli l'ottimo

La fisica delle particelle nell'era dell'IA

di Maurizio Pierini



a.
L'esperimento CMS al
CERN di Ginevra.

Una delle tecnologie alla base della rivoluzione dell'intelligenza artificiale è il *machine learning*, un approccio in cui un algoritmo apprende a svolgere un compito (come riconoscere oggetti o prendere decisioni) non attraverso regole fisse, ma tramite esempi. Più ampio e rappresentativo è il campione di dati usato per l'addestramento, più accurata sarà la soluzione proposta. In fisica delle particelle, il *machine learning* ha una lunga storia. Già dagli anni '80, i fisici hanno impiegato reti neurali e altri algoritmi per analizzare dati sperimentali. Tecniche di questo tipo hanno avuto, ad esempio, un ruolo importante nella determinazione degli elementi della matrice Cabibbo-Kobayashi-Maskawa, che spiega la simmetria tra materia e antimateria nei quark (premio Nobel per la fisica del 2008) e nella scoperta del bosone di Higgs, che spiega l'origine della massa delle particelle elementari (premio Nobel per la fisica del 2013).

Nel 2012, mentre i fisici del CERN annunciavano la scoperta del bosone di Higgs, Alex Krizhevsky, Ilya Sutskever e Geoffrey

Hinton presentavano AlexNet, una rete neurale profonda in grado di riconoscere immagini con una precisione senza precedenti. Questa rete diede il via alla nuova era del *deep learning*, che ha reso possibile l'attuale esplosione delle applicazioni dell'IA.

Durante i preparativi per il nuovo ciclo di dati ad alta energia di LHC previsto per il 2015, un piccolo gruppo di fisici cominciò a esplorare l'utilizzo di tecniche di *deep learning* per l'analisi dei dati sperimentali. I primi esperimenti si basarono su AlexNet per il riconoscimento di particelle, in particolare nei dati di LHC e di esperimenti sui neutrini. L'idea era trasformare i segnali raccolti dai rivelatori in immagini da fornire in *input* alla rete neurale. I risultati iniziali furono promettenti: le reti profonde superavano in prestazioni gli algoritmi tradizionali. Tuttavia, l'approccio "a immagine" mostrava dei limiti: i rivelatori di LHC hanno geometrie irregolari e complesse, spesso disposte in strati cilindrici, con sensori di forme e



b.
Rappresentazione artistica di una rete neurale che processa dati nell'esperimento ATLAS.

dimensioni diverse. I dati risultanti sono meglio descritti come una “nuvola di punti”, ciascuno con informazioni minime (posizione, carica elettrica, ecc.), che devono essere combinati per ricostruire le traiettorie delle particelle. In questo senso, i dati di LHC somigliano più a una rete sociale, dove ciascun “nodo” è un segnale e il significato globale emerge dalla relazione tra i segnali.

Dal 2017, gli esperimenti ATLAS e CMS al CERN hanno iniziato a utilizzare modelli neurali più adatti a strutture di dati complesse, come le *graph neural networks* (GNN) e, più recentemente, i Transformer – la stessa architettura usata per sviluppare ChatGPT. Queste reti sono in grado di cogliere relazioni complesse tra segnali, migliorando di un ordine di grandezza la capacità di identificare particelle o eventi rari nei dati.

Questo progresso ha portato alla formazione di una comunità all'interno degli esperimenti di LHC dedicata all'intelligenza artificiale. I ricercatori hanno iniziato ad applicare il *machine learning* a problemi fino ad allora risolti con metodi classici: simulazione del comportamento dei rivelatori, identificazione automatica di eventi di interesse, ricostruzione di particelle dai segnali registrati. L'uso dell'IA ha superato la fase di test: oggi è considerata uno strumento essenziale per la fisica delle particelle. Una pietra miliare si è raggiunta con l'inizio del *run 3* di LHC, quando l'esperimento CMS ha introdotto reti neurali profonde nel sistema di filtraggio in tempo reale dei dati (*trigger*). LHC produce circa 30 milioni di collisioni al secondo, ma solo poche migliaia possono essere salvate per l'analisi. La selezione iniziale

viene eseguita da algoritmi su FPGA (Field Programmable Gate Array), un tipo versatile di circuito integrato progettato per essere programmabile (vd. anche in *Asimmetrie* n. 27 p. 15, ndr). Sull'FPGA, i fisici degli esperimenti LHC eseguono gli algoritmi di selezione rappresentandoli come circuiti logici. Tali algoritmi devono decidere in appena 4 millisecondi se un evento merita di essere conservato. Una rete neurale inserita in questo sistema deve essere prima tradotta in un circuito logico. Poi, deve essere integrata sul FPGA. Il tutto va fatto in maniera tale che la rete possa prendere una decisione in 100 nanosecondi. Per confronto, una rete neurale su un'auto a guida autonoma ha tempi di reazione diecimila volte maggiori.

In questo contesto sono stati sviluppati due algoritmi, CICADA e AXOL1TL, progettati per identificare collisioni anomale, cioè eventi rari o inaspettati. Queste reti vengono addestrate solo su eventi noti, e valutano quanto una nuova collisione sia “tipica”. Se rilevano un'anomalia, salvano l'evento per ulteriori analisi. Dal 2024, CMS sta costruendo un archivio di eventi anomali, con l'obiettivo di individuare possibili segnali prodotti, ad esempio, da scenari di nuova fisica mai esplorati prima e quindi sfuggiti agli approcci tradizionali.

Questa nuova strategia mostra come l'intelligenza artificiale non sia solo uno strumento di supporto, ma una componente attiva della ricerca di punta. Se da questi eventi emergessero nuove scoperte, il ruolo dell'IA in fisica delle particelle sarebbe destinato a diventare ancora più centrale di quanto già lo sia oggi.

Biografia

Maurizio Pierini è un ricercatore del CERN e lavora all'esperimento CMS dal 2007. Si è specializzato in approcci innovativi per la rivelazione di nuovi fenomeni fisici all'LHC. Ha sviluppato sistemi di acquisizione dati e di rilevamento di anomalie basati sull'intelligenza artificiale.

Come ti leggo dentro

Applicazioni in fisica teorica

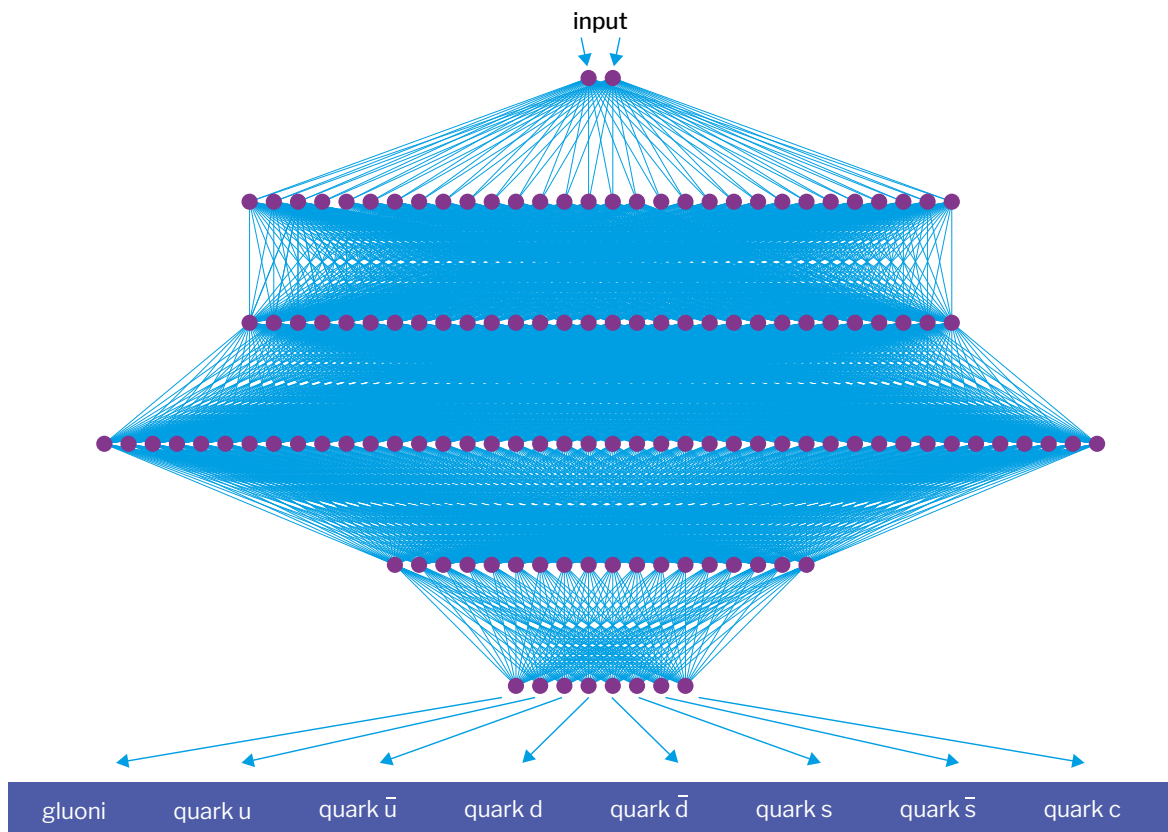
di Stefano Forte



L'intelligenza artificiale (IA) pervade il mondo della ricerca e sta entrando prepotentemente in molti aspetti della vita quotidiana. Nell'ambito della fisica teorica, il ruolo dell'IA è particolarmente significativo nella fisica delle particelle elementari. Compito della fisica teorica delle particelle è formulare le leggi che descrivono le forze

fondamentali della natura, che vengono messe alla prova confrontandole con i dati sperimentali, quali quelli provenienti dal CERN di Ginevra. La teoria permette di calcolare che cosa succede quando due protoni si scontrano frontalmente nel grande acceleratore di particelle LHC del CERN. I protoni sono fatti di costituenti elementari: si

a.
Vista aerea del CERN.



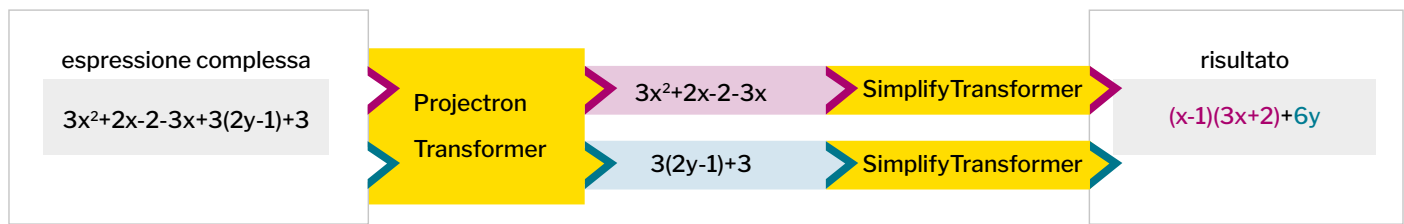
b. Schema di una rete neurale utilizzata per descrivere la struttura interna del protone. L'input (in alto) rappresenta la frazione longitudinale della quantità di moto dei costituenti del protone (quark e gluoni). Gli output (in basso) sono le probabilità che ciascun costituente porti proprio quella percentuale di quantità di moto. La rete neurale è costituita da vari livelli, ognuno composto da un diverso numero di neuroni (i pallini viola), collegati tra di loro da *link* (linee continue).

tratta dei quark, e dei gluoni che li legano assieme. In particolare, la teoria permette di combinare l'informazione sulla struttura del protone in termini di questi costituenti, con il calcolo di ciò che accade quando due costituenti si scontrano, producendo nuove particelle elementari: per esempio altri quark. Questi quindi si ricombinano fra loro, formando altre particelle dette "adroni", come ad esempio protoni o neutroni, che vengono visti mediante i rivelatori, giganteschi strumenti che ne intercettano il passaggio.

Diversi aspetti della procedura che porta dalla teoria fondamentale alla predizione di ciò che vede il rivelatore sono impossibili da affrontare con carta e penna. Innanzitutto, non si è attualmente in grado di calcolare come quark e gluoni sono legati dentro al protone: si tratta di un calcolo troppo difficile. L'interazione fra quark e gluoni invece si riesce a calcolare, ma al livello di precisione attuale è necessario manipolare espressioni algebriche fatte di migliaia o a volte anche milioni di termini. Infine, la ricombinazione dei quark in adroni che lasciano un segnale nel rivelatore è un processo troppo complicato per essere calcolabile.

Il modo tradizionale di affrontare questi problemi è sfruttando tecniche di simulazione numerica (vd.

in *Asimmetrie* n. 14 p. 25, ndr) e di manipolazione algebrica svolte da un computer. Ma l'intelligenza artificiale sta cambiando tutto, specialmente grazie alle tecniche sviluppate negli ultimi dieci anni: *deep learning* e modelli generativi. Vediamo come. Partiamo dalla struttura del protone: ci serve conoscerla per predire quello che succede quando due protoni si scontrano. Ma non la sappiamo calcolare: e dunque? Il modo tradizionale di procedere è attraverso la modellizzazione: si ipotizza un modello matematico con molti parametri indeterminati. Lo si usa per calcolare che succede quando due protoni si scontrano. Si confrontano i risultati con ciò che si osserva, e si scelgono i valori dei parametri che danno il migliore accordo. L'IA suggerisce di procedere diversamente. Supponiamo di vedere un'immagine, e dover decidere se si tratta di un cane o un gatto. Potremmo cercare di trovare un modello: un'equazione che distingue il cane dal gatto, ad esempio descrivendo la forma del muso. Ma con l'IA possiamo insegnare a una rete neurale a distinguere il cane dal gatto partendo da un addestramento supervisionato, cioè mostrando molti esempi finché la rete neurale non ne estrae le caratteristiche distintive. Analogamente, possiamo mostrare molti dati di collisioni fra protoni a una



c.
Una coppia di Transformer esegue la semplificazione di un'espressione algebrica: il primo raggruppa termini simili, il secondo li semplifica.

rete neurale, finché questa ne estrae la struttura del protone in termini di quark e gluoni. Grazie al *deep learning* si può non solo imparare come è fatto un protone, ma anche qual è il miglior modo di ottenere questo risultato. In una prima fase (iper-ottimizzazione) si determinano le caratteristiche ottimali della rete neurale e del metodo da usare per addestrarla. In una seconda fase la rete ottimale "impara" la struttura del protone. Passiamo alla relazione fra il calcolo dei prodotti della collisione e le tracce nel rivelatore. Tradizionalmente lo si affronta con il metodo Monte Carlo: ogni passo del cammino che porta dai quark prodotti nella collisione, alla loro ricombinazione con altri quark per formare gli adroni, e infine all'interazione di questi ultimi con il rivelatore, viene simulato numericamente al computer (vd. anche in *Asimmetrie* 27 p. 28, ndr). Confrontando i risultati simulati con quelli reali è possibile ricostruire, a partire da ciò che si osserva, che cosa ha prodotto la collisione. L'IA generativa affronta il problema usando la stessa tecnica che permette di scrivere un pezzo di musica nello stile di Beethoven: ne è un esempio l'uso delle Reti Generative Avversarie (GAN). Una rete neurale, il generatore, viene addestrata a generare dei dati artificiali. Nel nostro caso, a partire dalle tracce in un rivelatore lasciate da particelle prodotte in una collisione, il generatore produce un evento artificiale che si manifesta proprio con quelle tracce. Le tracce e l'evento generato vengono mostrate a un'altra rete neurale, il discriminatore, che deve capire se quello che sta vedendo è un vero evento, o un evento generato artificialmente. Il generatore e il discriminatore competono, e si migliorano a vicenda, finché il generatore non riesce più a distinguere gli eventi veri da quelli artificialmente prodotti

dal generatore. A questo punto, il generatore produce risultati indistinguibili da quelli reali, e dunque fornisce lo strumento desiderato, che dalle tracce ricostruisce l'evento soggiacente. Infine, la semplificazione di espressioni algebriche. Tradizionalmente si usa la *computer algebra*: l'elaboratore applica ripetutamente regole di trasformazione, un po' come i prodotti notevoli che si imparano alla scuola media. L'IA generativa affronta il problema con lo strumento usato dai Large Language Models (LLM) come ChatGPT: il Transformer. Questo prende in ingresso una sequenza e la completa: ad esempio le successive battute di un dialogo, in cui l'ultima è la domanda che facciamo a ChatGPT, e il completamento è la risposta di ChatGPT. Nel caso che ci interessa, un Transformer raggruppa i termini che possono essere semplificati, e un altro esegue la semplificazione. Come nei LLM, il Transformer procede in modo intuitivo: non applica delle regole prefissate, ma impara dall'esperienza sulla base di esempi. La forza del metodo è che se anche ogni tanto il Transformer sbaglia, non importa: si può sempre verificare se l'espressione iniziale e quella semplificata danno lo stesso risultato; altrimenti, il risultato si scarta e si ricomincia. Le tecniche che abbiamo visto sono di crescente complessità, e portano alla frontiera del progresso dell'IA in fisica teorica. Si tratta di un progresso bidirezionale. Da un lato, algoritmi sempre più avanzati forniscono strumenti di conoscenza sempre più potenti. Dall'altro, le esigenze della fisica teorica fanno da motore dello sviluppo di metodi sempre più sofisticati. L'intelligenza artificiale diventa così uno strumento per produrre conoscenza.

Biografia

Stefano Forte è professore ordinario di fisica teorica all'università di Milano dal 2003, dopo aver occupato posizioni di ricerca a Saclay (Parigi), al CERN e nell'INFN a Torino e Roma. Fin dai primi anni 2000 ha usato tecniche di intelligenza artificiale per la comprensione della struttura del protone.

Occhio clinico

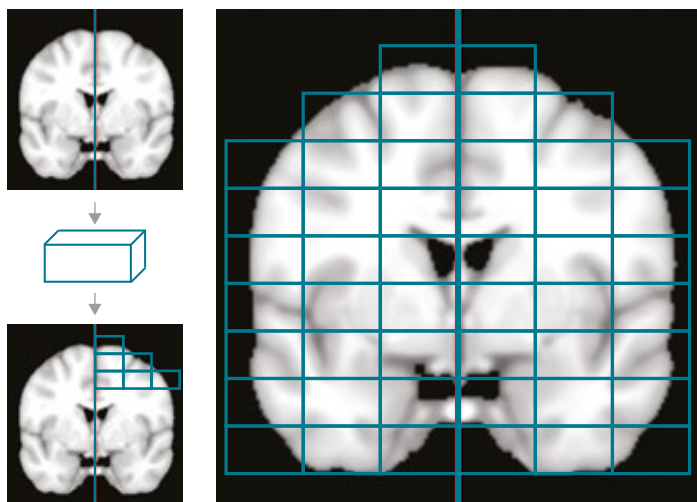
L'utilizzo in fisica medica

di Marianna La Rocca



Per cercare le debolezze del cervello umano si può usare uno strumento ispirato al cervello stesso? È proprio questa l'idea alla base di diversi studi e ricerche scientifiche che stanno portando avanti la sfida contro le malattie neurodegenerative come l'Alzheimer e il Parkinson, nell'ambito della fisica medica. Una sfida tanto ambiziosa quanto urgente, che mette in campo una combinazione potente: intelligenza artificiale e reti complesse. Le malattie neurodegenerative sono subdole. Agiscono lentamente e silenziosamente, spesso per anni, prima di manifestarsi con sintomi evidenti.

E a quel punto, spesso, è troppo tardi. Ecco perché è cruciale individuare indicatori precoci: serve a sperimentare terapie che possano rallentare, fermare o persino invertire il decorso della malattia. In questa missione, l'intelligenza artificiale si rivela un alleato formidabile. Ma come funziona? Pensiamo alle centinaia di connessioni che si formano ogni secondo nel nostro cervello durante l'apprendimento quotidiano, fin da quando siamo bambini. Le esperienze giocano un ruolo fondamentale nella formazione di una persona. Tutti abbiamo sentito dire almeno una volta: "si impara dai propri errori". E



b.

La figura mostra come le scansioni cerebrali ottenute con la risonanza magnetica (MRI) vengono suddivise in volumi rettangolari. Questi volumi, chiamati "patch", vengono poi usati per costruire una rete cerebrale: ogni patch rappresenta un nodo della rete, e i collegamenti (o link) tra i nodi indicano quanto questi volumi del cervello sono simili tra loro.

indovinate un po'? È proprio vero.

Immaginiamo un bambino che si brucia toccando una pentola sul fuoco, nonostante i genitori gli abbiano spiegato che il fuoco e l'acqua bollente possono far male. Dopo quell'esperienza spiacevole, il bambino capirà non solo che toccare qualsiasi oggetto rovente è una cattiva idea, ma anche che in presenza di fonti di calore eccessivo è meglio avvicinarsi con prudenza. Grazie a questa esperienza, sarà meno incline a commettere lo stesso errore e più preparato ad affrontare situazioni simili in futuro.

Allo stesso modo, per avere un sistema di intelligenza artificiale davvero efficace, non serve trasmettergli tutta la conoscenza possibile, ma bisogna insegnargli a imparare nel modo più autonomo ed efficiente. È importante anche che possa sbagliare ogni tanto, così da poter riorganizzare le sue strategie decisionali dopo ogni tentativo fallito, basandosi sui dati che gli abbiamo fornito. Solo così il sistema acquisirà la flessibilità necessaria per adattarsi a nuovi dati e potrà esplorare, tra un errore e l'altro, strade alternative che lo porteranno a ottenere risultati sempre migliori.

Nella lotta contro le malattie neurologiche, l'IA viene spesso allenata a distinguere immagini cerebrali di persone sane da quelle di pazienti con decadimento cognitivo lieve (*Mid Cognitive Impairment* - MCI), una condizione spesso preludio all'Alzheimer. Non è tutto, a questi sistemi di IA vengono richiesti compiti ancora più complessi: identificare quei soggetti MCI che svilupperanno la malattia negli anni successivi. Come si riesce in quest'impresa? Ogni immagine di risonanza magnetica cerebrale viene trasformata in una rete complessa. Proprio come una rete di aeroporti, dove ogni scalo (nodo della rete) rappresenta un piccolo volume del cervello, e ogni collegamento ("link") misura

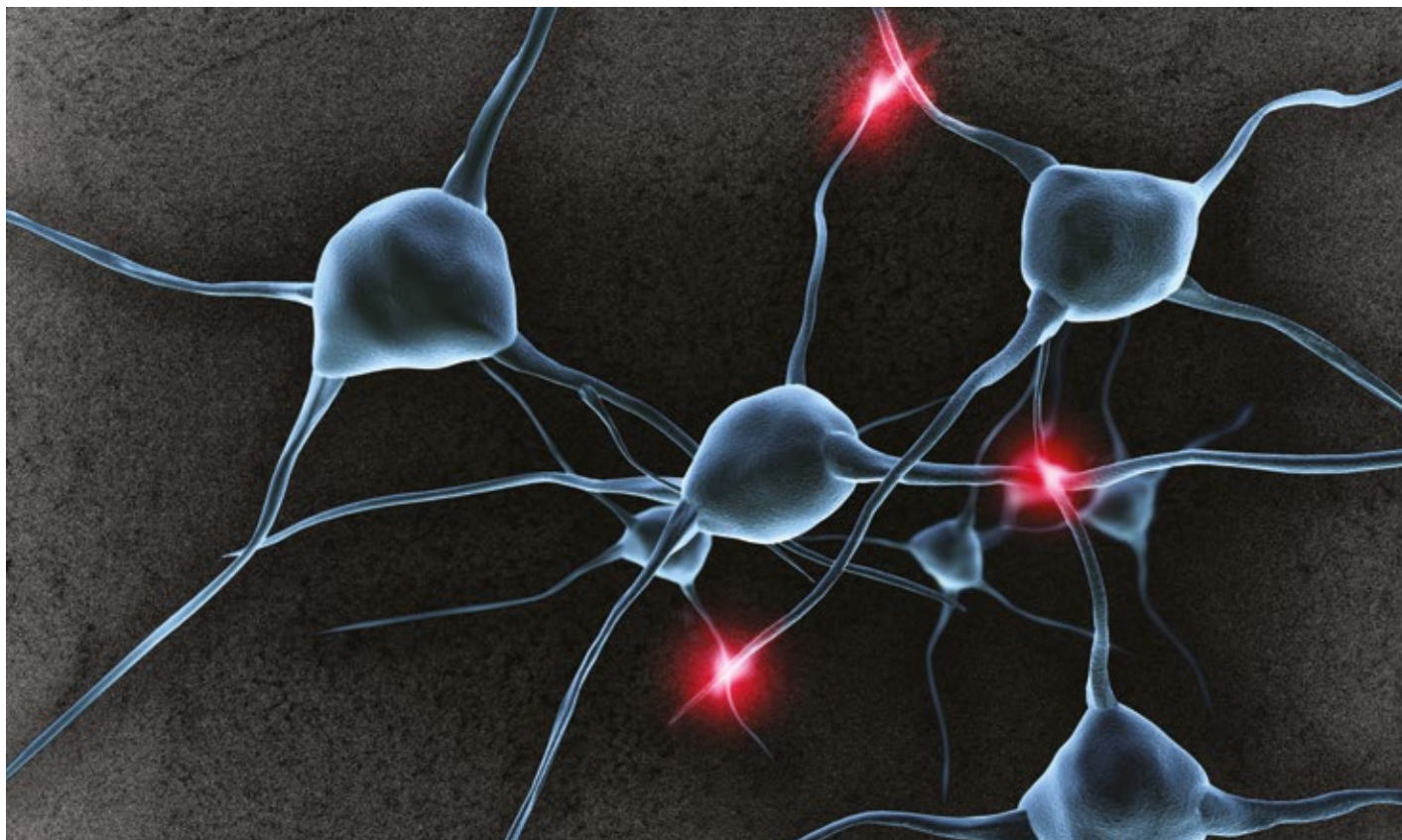
quanto queste aree siano simili tra loro. È così che si può quantificare come le connessioni cerebrali cambiano con l'avanzare della malattia.

L'IA analizza queste reti usando indicatori matematici, ad esempio il numero di connessioni di ogni nodo, per apprendere schemi tipici delle persone sane e di quelle malate. Una volta addestrata, la macchina è in grado di riconoscere questi schemi anche in nuovi soggetti mai visti prima.

I risultati sono promettenti: con una precisione superiore all'80%, si riesce a predire lo stato clinico anche nove anni prima della normale diagnosi. Un dato impressionante, che potrebbe rivoluzionare il modo in cui affrontiamo queste patologie.

Un aspetto cruciale nello studio delle malattie neurodegenerative è la scelta della "scala" con cui si analizzano le immagini cerebrali: se è troppo grande si rischia di perdere dettagli significativi, se è troppo piccola di non cogliere il quadro d'insieme e focalizzarsi su dettagli non importanti. La "scala ottimale" è quindi quella dimensione spaziale che permette di cogliere al meglio le alterazioni caratteristiche della malattia. Una scoperta sorprendente è che questa scala ottimale varia a seconda della patologia. Dividendo il cervello in piccoli blocchi, è emerso che per l'Alzheimer l'unità di analisi ideale è di circa 3000 mm³, una dimensione simile a quella dell'ippocampo, l'area cerebrale più colpita all'esordio della malattia. Per il Parkinson, invece, la scala ottimale è di soli 125 mm³, dimensione comparabile alla *substantia nigra*, una delle regioni più colpite all'esordio della malattia.

Questo approccio ha un vantaggio enorme: permette di analizzare l'intero cervello in modo sistematico, senza concentrarsi solo sulle aree già note. In questo modo si evita il rischio di introdurre errori legati ai metodi di segmentazione automatica,



c.
Immagine artistica di neuroni cerebrali.

che suddividono il cervello in aree predefinite senza un criterio specifico. Inoltre, non vincolandosi alle aree oggetto delle ipotesi preesistenti, è possibile scoprire nuove aree coinvolte nello sviluppo della patologia, aprendo la strada a scoperte fondamentali soprattutto per malattie ancora poco comprese. Un altro elemento innovativo è rappresentato dall'uso di tecniche note come "*explainable artificial intelligence*", che permettono di conoscere più a fondo le strategie applicate di caso in caso dagli algoritmi di IA. In questo modo è possibile conoscere quali sono le regioni cerebrali associate alle caratteristiche che influenzano maggiormente la discriminazione di due gruppi clinici. Questi approcci innovativi permettono una maggiore interpretabilità e affidabilità dell'IA, che è cruciale soprattutto nell'ambito medico e nell'ottica di una medicina personalizzata.

La combinazione delle reti complesse e dell'IA sta prendendo

sempre più piede nel mondo scientifico e si sta rivelando molto utile anche per la diagnosi precoce di altre malattie neurologiche, come l'epilessia post-traumatica o per la valutazione dell'efficacia neurologica di diverse terapie. Il punto di forza è che le reti complesse possono essere applicate e integrate in maniera versatile e quantitativa in qualsiasi tipo di dato medico: dati genetici, dati elettroencefalografici e immagini di risonanza magnetica funzionale e strutturale. L'IA nell'ambito della fisica medica e in particolare nella lotta contro patologie che distruggono completamente la vita delle persone colpite e delle loro famiglie, ci insegna a non lasciarci scoraggiare dai fallimenti. Anzi, è proprio da quei fallimenti che dobbiamo ripartire per riuscire a sconfiggere queste terribili malattie. Come diceva Albert Einstein, "*if you have never failed, you have never tried anything new*": se non hai mai fallito, non hai mai provato niente di nuovo.

Biografia

Marianna La Rocca è ricercatrice in fisica applicata presso il Dipartimento di Fisica dell'Università di Bari, dove ha anche conseguito il dottorato in fisica applicata. È stata per tre anni *postdoc* alla University of Southern California. Si occupa principalmente dello studio di malattie neurologiche usando reti complesse e tecniche di *machine learning*.

[as] riflessi

Dove l'IA incontra la potenza di calcolo

di Martina Galli

Giovane, frequentatissimo, all'avanguardia: il Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing (ICSC), inaugurato a Bologna nel 2022 su iniziativa dell'INFN e con finanziamenti del Ministero dell'Università e della Ricerca, ha rivoluzionato il calcolo a livello nazionale. Con l'obiettivo di mettere a sistema tutte le risorse disponibili nel Paese e potenziarle, ICSC è riuscito a realizzare una delle infrastrutture di calcolo più avanzate al mondo, a cui comunità scientifiche e aziende private riescono ad accedere in maniera dinamica. Tra le risorse messe a disposizione c'è Leonardo, un supercomputer per il calcolo parallelo potentissimo (tra i primi dieci a livello globale), il cui *upgrade* Lisa – in onore di Monna Lisa – sarà ottimizzato per l'addestramento di algoritmi di intelligenza artificiale (IA). In questa nuova epoca in cui la materia prima sono i dati, infatti, è importante non soltanto saperli processare con macchine estremamente performanti, ma anche riuscire a estrarre valore da essi. ICSC ha quindi affiancato alla sua infrastruttura distribuita lo sviluppo di applicazioni di IA, che con questa nuova materia prima sappiano creare opportunità. Abbiamo chiesto ad Antonio Zoccoli, presidente dell'INFN e della Fondazione ICSC, di raccontarcele.

[as]: Su quali settori si concentra ICSC per lo sviluppo di applicazioni di IA?

[Antonio Zoccoli]: L'intelligenza artificiale ha ormai un ruolo preponderante in tutte le linee scientifiche di ICSC. Ne

abbiamo individuate otto, otto “*spoke*”: scienza fondamentale e *space economy*, astrofisica, clima, ambiente e disastri naturali, ingegneria, scienza molecolare e dei materiali, medicina innovativa, *smart cities*. In tutte queste aree i dati possono fornire indicazioni preziose, e rivoluzionarle. Prendiamo per esempio il settore agricolo, in cui incrociando i dati satellitari e i dati climatici, si riesce a ottimizzare l'irrigazione di una determinata coltura; oppure l'ambito medico, in cui incrociando i dati genomici e quelli delle risonanze, si riescono a diagnosticare efficacemente le malattie e quindi a individuare le cure migliori. Nella scienza fondamentale gli algoritmi di IA possono guidarci nella ricerca di nuove particelle, nuovi sistemi solari, e indicarci nuove vie per studiare le leggi fondamentali dell'universo. E in ambito ingegneristico, i cosiddetti “gemelli digitali” possono offrirci innumerevoli vantaggi, permettendoci, per esempio, di testare la sicurezza degli edifici.

[as]: Come funzionano i gemelli digitali?

[AZ]: Un gemello digitale è sostanzialmente un codice che riproduce uno strumento, un impianto, un processo e ne simula il funzionamento ideale. Quando lo strumento, l'impianto, il processo assume un andamento diverso nella realtà – per esempio il traffico in una determinata città – allora il gemello digitale fornisce delle informazioni aggiuntive, diciamo dei suggerimenti di intervento. Questo è particolarmente rilevante nel caso di sistemi complessi, difficilmente simulabili con carta e matita, perché il gemello digitale permette di visualizzare



a.
Il presidente dell'INFN e della Fondazione ICSC, Antonio Zoccoli.



b.
Il Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing (ICSC) è ospitato a Bologna in un polo di edifici della Regione.

gli effetti di una eventuale rottura, deformazione, determinata azione, e quindi di mettere in campo delle azioni correttive, utili a prevenire effetti indesiderati o a ottimizzare i processi.

[as]: Quale sarà l'impatto di queste nuove tecnologie sul tessuto produttivo italiano?

[AZ]: L'impatto è potenzialmente trasformativo, ma è molto importante che siano proprio le aziende, conscie delle esigenze del mercato, a fare da traino. L'idea è quella di aiutarle con delle soluzioni operative, sviluppate *ad hoc* dalle accademie e dagli enti di ricerca che fanno parte del consorzio di ICSC, per rispondere precisamente alle loro esigenze. Con Unipol, per esempio, abbiamo lavorato sul monitoraggio delle automobili dotate della loro scatola nera; con Intesa Sanpaolo abbiamo sviluppato algoritmi di IA per ottimizzare i loro processi bancari; con Humanitas ci siamo concentrati sullo sviluppo di nuove terapie. Sono tantissime le applicazioni che il mondo produttivo può ottenere sfruttando le nostre competenze. La sfida con le aziende medio-piccole è piuttosto chiarire loro quale sia il reale vantaggio competitivo delle soluzioni che proponiamo, perché spesso non ne vedono il potenziale. Il loro coinvolgimento è fondamentale per innescare una transizione verso un sistema produttivo che sfrutti l'intelligenza artificiale, in cui chiunque abbia un po' di conoscenza, inventiva e coraggio per proporre soluzioni innovative troverà spazio.

[as]: Quali sfide computazionali si troverebbe ad affrontare?

[AZ]: Per addestrare un algoritmo di intelligenza artificiale si fa affidamento al calcolo parallelo, ovvero una serie di CPU che lavorano simultaneamente. A seconda di quanto è complicato l'algoritmo, e quindi del numero di parametri che lo determinano, il processo è più o meno complicato, più o meno lungo. Quello che conta è avere, per algoritmi molto complicati, un numero molto grande di CPU, quindi almeno un grande supercomputer. Avere un supercomputer da 100 CPU e avere due supercomputer da 50 CPU, infatti, non è la stessa cosa, perché non conta il numero assoluto di CPU, ma il numero di CPU utilizzabili contemporaneamente. Un supercomputer da 100 CPU in questo senso è più adatto; consente di addestrare il modello di IA in un tempo breve, e dunque di rimanere competitivi sul mercato. Per avere un'idea dei tempi di elaborazione: se io volessi fare il *training* di un Large Language Model – un algoritmo che sfrutta l'IA generativa come ChatGPT – su Leonardo, impiegherei tre mesi, su un computer normale circa 50 anni. Naturalmente a un supercomputer come Leonardo è utile affiancare computer a scalare più piccoli, che permettano di sviluppare algoritmi più semplici. A Bologna, per esempio, abbiamo creato dei sistemi di calcolo parallelo chiamati HPC *bubbles*, più piccoli e dinamici, finalizzati proprio a questo scopo.

[as] intersezioni

Bella senz'anima

di Francesco Bianchini

filosofo della scienza



Non è facile oggi parlare di etica dell'intelligenza artificiale (IA). Non è facile come lo sarebbe stato qualche anno fa o, ancora di più, qualche decennio fa. Le ragioni di questa difficoltà sono varie. Innanzitutto, l'IA è una disciplina scientifico-tecnologica e l'etica è una materia filosofica. E fin qui le cose sembrano essere piuttosto chiare. Dopodiché, la congiunzione fra queste due discipline può avvenire, anche qui senza troppe complicazioni, come una riflessione filosofica sulla scienza che chiama in causa aspetti morali e valoriali. Come per molte altre discipline scientifiche, è possibile analizzare gli aspetti etici riguardanti i principi di una disciplina scientifico-tecnologica, le implicazioni morali delle sue metodologie e dei suoi risultati, e le conseguenze che il suo impatto può avere sulla società in cui essa viene sviluppata e diffusa. Si pensi alla fisica e alle

riflessioni sulla visione dello spaziotempo portate dalla relatività; alla genetica e agli interrogativi che suscita in quanto disciplina in grado di influenzare e modificare la costituzione del mondo vivente, compreso quello umano; alla fisica delle particelle e alle considerazioni etiche sull'uso, bellico o meno, dell'energia nucleare; alla medicina e ai tanti risvolti etici che le sue pratiche implicano. Fin qui il discorso non presenta difficoltà: una disciplina scientifica, in genere nei suoi aspetti applicativi, presenta quesiti morali che la filosofia della scienza e la filosofia morale sono abituate ad affrontare. Il discorso per l'IA si complica se la consideriamo nelle sue varie sfaccettature. Si possono individuare due macroaree entro cui si sviluppa la riflessione etica relativa all'IA.

La prima riguarda, nel senso più classico relativo all'etica di una disciplina scientifica, questioni legate

a.
L'intelligenza artificiale alle prese con le scelte etiche.



b.
L'IA consente nuove opportunità nel mondo del lavoro, semplificando i processi e liberando tempo per attività a maggiore valore umano.

al suo sviluppo, alla sua applicazione e ai suoi risultati. Questa area è molto vasta e va da aspetti sociali e politici dell'impiego dell'IA a questioni più strettamente relative all'uso, diretto o indiretto, dell'IA da parte del singolo o in relazione all'individuo e alle sue attività. Questioni rilevanti da tale punto di vista riguardano l'impatto che l'IA ha sulla nostra società, sul mondo del lavoro, sulle *policy* da prevedere per un corretto uso degli strumenti IA disponibili.

L'IA consente nuove opportunità nel mondo del lavoro, semplificando i processi e liberando tempo per attività a maggiore valore umano.

Alcune delle domande cui si vuole rispondere sono: quale peso può avere l'IA nel determinare le scelte politiche e sociali? Come può influenzare i sistemi politici democratici, ad esempio per quanto riguarda i loro meccanismi elettivi e di scelta? Quale impatto ha l'IA sulle dinamiche occupazionali? Porterà a cambiamenti che cancelleranno posti di lavoro con terribili conseguenze sociali? O ne creerà di nuovi, dimostrandosi un'opportunità per le evoluzioni future delle società in cui si diffonderà capillarmente? O entrambe le cose? E in tal caso, con che tempistiche? Quali linee guida e sistemi di leggi devono essere previsti e strutturati per prevenire usi pericolosi dell'IA, sia per le società che per gli individui, e quali agenzie nazionali e transnazionali sono titolate e proporli? A tale gamma di preoccupazioni si sommano, infine, le questioni legate alla sostenibilità ambientale ed energetica dell'IA con il loro conseguente impatto sociale.

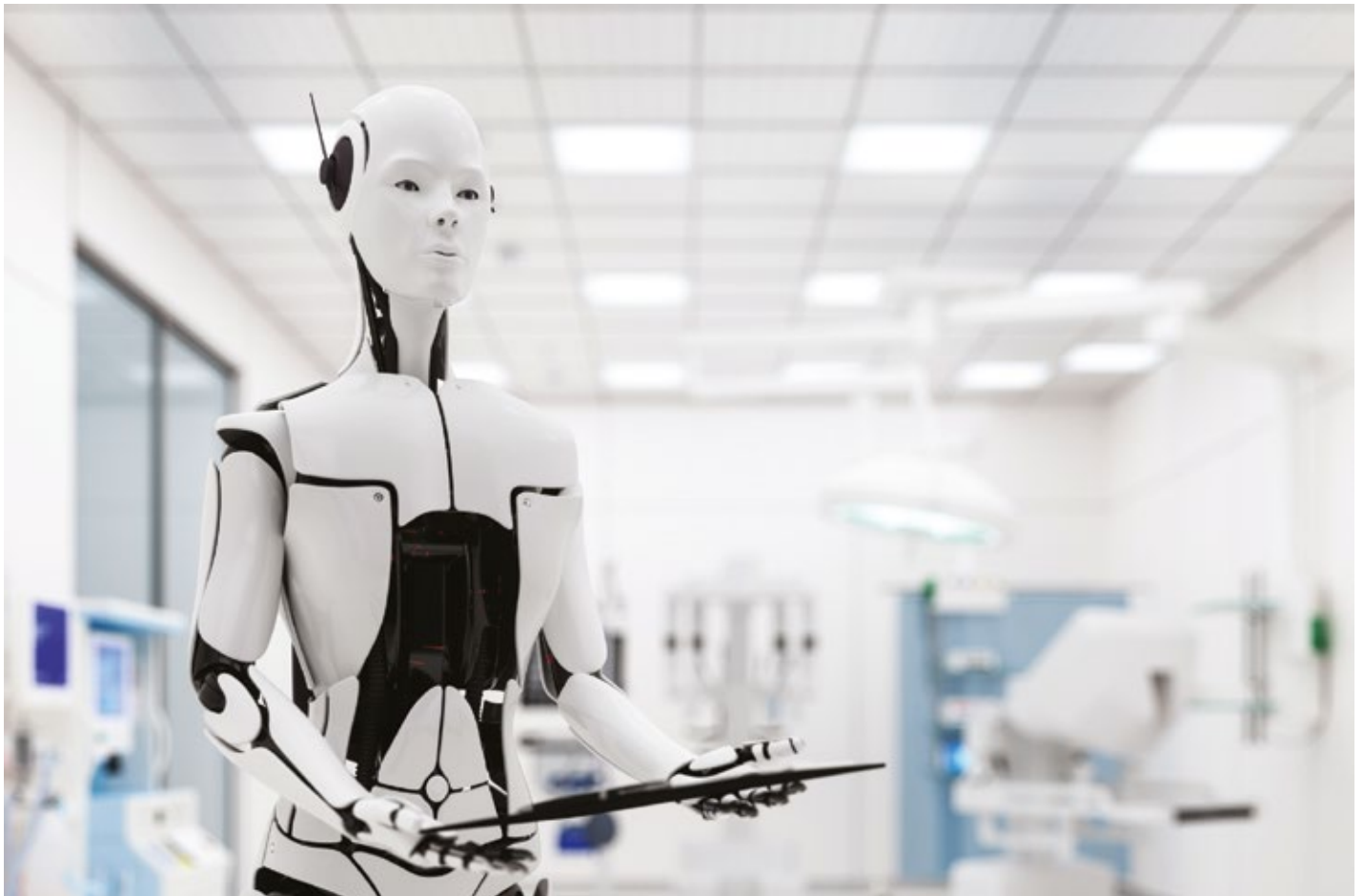
Inoltre, come si diceva, i problemi etici sociali si intrecciano con quelli relativi agli individui. Due fattori hanno reso pressante il problema etico dell'IA nei confronti delle singole persone: la tipologia algoritmica largamente prevalente dell'IA contemporanea e la sua recente diffusione su ampia scala. È vero infatti che oggi, anche più che soltanto dieci anni fa, l'IA è accessibile a un largo pubblico e pertanto è utilizzata, in senso diretto o indiretto, da utenti non specializzati. Ciò può generare usi incontrollati o poco consapevoli, come ad esempio la produzione di testi o immagini attraverso IA generative senza le opportune cautele e la necessaria competenza; o anche l'utilizzo dell'IA per manipolare/controllare gli individui, ad esempio attraverso algoritmi di profilazione in campo commerciale, finanziario, sociale, ecc. Inoltre, tali algoritmi sono in genere basati su reti neurali, in particolare sulle loro evoluzioni degli ultimi quindici anni, cioè vale a dire su metodi statistico-predittivi in grado di produrre risultati con un elevato grado di accuratezza, ma processualmente opachi. La loro elaborazione consiste in complesse funzioni matematiche che associano un

output a un input senza che i processi interni dell'esecuzione dell'algoritmo siano spiegabili in termini comprensibili, senza che sia possibile, cioè, assegnare loro un significato. La riflessione etica in questo caso si concentra, a partire da vari tentativi per raggiungerne la spiegabilità, su come proteggere la *privacy* degli utenti, sulla trasparenza dell'operato di tali sistemi, sul corretto modo di addestrarli partendo da dati che non siano condizionati da assunzioni implicite ingiuste o scorrette da un punto di vista etico e sociale (i cosiddetti *bias*). Negli ultimi anni sono state approntate numerose linee guida etiche, che condividono in genere principi quali trasparenza, equità, giustizia, beneficenza, non maleficenza. La seconda macroarea etica relativa all'IA, che la rende più originale rispetto ad altri campi scientifico-tecnologici, riguarda l'etica che i sistemi stessi di IA devono possedere. Si tratta della questione delle "macchine morali" o dell'"etica delle macchine". Poiché una delle caratteristiche che accomuna i sistemi dell'IA è la loro autonomia, essa deve essere in qualche modo normata anche dall'interno. Quali regole devono

guidare il comportamento delle IA? Quali comportamenti sono permessi e quali no? Ciò vale per i sistemi software, ma è ancora più evidente per i sistemi robotici; e vale sia per gli ambienti strutturati (si pensi ad esempio alla robotica industriale o a quella chirurgica) sia per quelli più indeterminati, dove è maggiore il grado di autonomia e di azione "libera" che i sistemi possono mettere in atto, specialmente quando sono presenti esseri umani (si pensi ad esempio al caso dei veicoli autonomi o a quello dei robot assistivi). È proprio l'interazione con gli esseri umani che fa nascere la richiesta di regole e norme per un comportamento adeguato e non dannoso, di una sorta di morale del sistema IA, anche nei casi estremi come quelli delle armi autonome. Seppure non sempre gli aspetti etici dell'IA vengano perseguiti o tenuti in debita considerazione, essi sono pervasivi e richiedono, oggi più che mai in misura crescente, competenze giuridiche, etiche ed epistemologiche con l'obiettivo morale effettivo di una IA non tanto migliore quanto più sostenibile per l'essere umano.

C.

Un "medico robot" umanoide che prepara la cartella medica in ospedale.



Riconoscimento di immagini con “carta e penna”

di Anna Greco

Le tecniche di *image recognition* hanno l'obiettivo di rendere un sistema artificiale in grado di riconoscere gli oggetti all'interno di un'immagine: un esempio fra tutti è la funzione che classifica le foto nella galleria di alcuni smartphone, riconoscendone i protagonisti. Attraverso un'attività in classe, si possono individuare i punti salienti del processo di riconoscimento e provare a simularlo con “carta e penna”. L'attività descritta è presente nei *lesson plan* di DIG4Future (Digital competencies, Inclusion and Growth for Future Generations), un progetto Erasmus+ dedicato a migliorare le competenze digitali dei giovani tra 11 e 14 anni che vivono in aree urbane e rurali svantaggiate in Italia, Bulgaria, Grecia e Romania. Inizialmente, il docente chiede alla classe se qualcuno ha mai sperimentato tecnologie di riconoscimento immagini e propone qualche esempio. Invita poi studenti e studentesse a ragionare su come gli umani riconoscono un'immagine: elencando i passaggi che ciascuno compie per riconoscere un soggetto, si introduce il concetto di algoritmo e quello di estrazione delle caratteristiche. Si può descrivere a questo punto il procedimento con cui un computer riconosce l'immagine, spiegando che la tecnica utilizzata per l'*image recognition* è quella delle reti neurali convoluzionali (vd. p. 8), che si ispirano al funzionamento della corteccia visiva, la parte del nostro cervello che elabora le informazioni che riceviamo dagli occhi e ci permette di vedere. Il giusto algoritmo permette di “allenare”

le reti neurali a riconoscere sempre meglio le immagini. A questo punto, la classe viene suddivisa in 4 o 5 gruppi. Tutti i gruppi meno uno ricevono la fotografia di un cane, diverso per ogni gruppo, e una “tabella delle razze”, che elenca le caratteristiche di ciascuna razza di cane. Questi gruppi si occupano dell'estrazione delle caratteristiche, compilandone un elenco relativo all'immagine ricevuta secondo la tabella. Il gruppo rimanente è quello degli “investigatori dell'IA”: ognuno dei componenti riceve una delle foto di cani fornite agli altri e compila la tabella della razza. Si leggono poi le risposte degli “investigatori” e si comparano con quelle degli altri gruppi, di cui non si conosce ancora l'immagine. Si attribuisce un punteggio di somiglianza ai gruppi, in base a quanto sono simili le proprie indicazioni con quelle di ciascuno degli “investigatori”. Il gruppo con il maggior punteggio di somiglianza rivelerà a questo punto la foto per verificare se coincida con quella descritta dagli “investigatori”. Se il procedimento ha funzionato, il cane dello studente “investigatore” dovrebbe corrispondere a quello del gruppo con il punteggio di somiglianza più alto. Gli studenti possono discutere i risultati, ragionando su come rendere più preciso ed efficiente il processo. Il docente può introdurre infine l'idea dei *database di training*, spiegando come sia utile aumentare il numero di immagini da cui trarre le caratteristiche per addestrare in modo più efficiente l'IA.

						
	chihuahua	spitz americano	bulldog francese	pastore tedesco	barboncino	komondor
dimensioni	piccolo	grande	piccolo	grande	piccolo	grande
colore del manto						
lunghezza del pelo	corto	medio	corto	medio	medio	lungo
dimensione degli occhi	grande	normale	normale	normale	piccola	?
colore degli occhi	scuri	azzurro	scuri	scuri	scuri	?
dimensione delle orecchie	grande	media	grande	media	media	?
orecchie all'insù	sì	sì	sì	sì	no	no
dimensione del naso	piccolo	medio	grande	medio	medio	grande

a.
Un esempio di tabella delle razze canine, compilata con le caratteristiche delle razze.

[as] visioni

Ma che domande!

di Giuliana Galati

La folla, in trepida attesa fino a un minuto prima, ammutolisce. Qualcuno mormora il proprio disappunto. Pensiero Profondo, il secondo più grande supercomputer di tutti i tempi e di tutti gli spazi, si è pronunciato: la risposta alla “Domanda fondamentale sulla vita, sull’universo e tutto quanto” è “42”. Loonquawl e Phouchg, dopo aver ascoltato la risposta, si allontanano sommessamente borbottando.

[Loonquawl & Phouchg]: Sette milioni e mezzo di anni abbiamo aspettato! Se solo i nostri avi Lunkwill e Fook avessero posto meglio la Domanda forse oggi avremmo avuto la risposta! Eppure deve esserci un senso... cosa può significare “42”?

[as]: Scusate se mi intrometto, ma 42 è un numero molto interessante! Per esempio, è la somma delle prime tre potenze di due con esponente dispari: $2^1 + 2^3 + 2^5 = 42$. Se consideriamo la successione il cui termine generale è la somma delle prime n (con $n > 0$) potenze di 2 con esponente dispari scopriamo che, in base 2, l' n -esimo elemento si può esprimere ripetendo la stringa “10” per n volte. Infatti, 42 in base 2 si scrive 101010. I numeri che appartengono a questa successione sono estremamente rari!

[L&P]: Grazie per la lezione matematica, ma noi esseri pandimensionali iperintelligenti conosciamo benissimo tutte le proprietà dei numeri. Se è per questo 42 fa parte anche delle successioni A105281, A000108, A005153 dell’OEIS, la “On-Line Encyclopedia of Integer Sequences”, oltre che della già da lei citata A020988, e poi è un numero sfenico, semiperfetto, colombiano, idoneo, pratico, scarsamente totiente, oblungo...

[as]: Va bene, va bene ho capito! Certo che ricordate a memoria un sacco di cose... Mi fa piacere notare che l’uso di supercomputer e intelligenza artificiale non ha compromesso le facoltà mentali! Fate largo uso di questi sistemi?

[L&P]: Certamente, ma non sono poi così utili per sopravvivere...

[as]: Ah no? E cosa è veramente utile?

[L&P]: Beh, forse l’oggetto più utile che uno possa avere è l’asciugamano. Lo dice anche la guida galattica per gli autostoppisti. E la guida galattica è infallibile... È la realtà, spesso, a essere inesatta.

[as]: E i robot? Ne avete tanti? Sono tutti depressi come Marvin?



a.
Non tutti i robot sono allegri! Marvin l’androide paranoico è uno dei personaggi principali della “Guida galattica per gli autostoppisti”, serie ideata da Douglas Adams.

[as] traiettorie

Dalle particelle all'IA

di Francesca Mazzotta

“Del mio lavoro apprezzo molto la flessibilità nel poter fare ricerca su quello che preferisco e la possibilità di collaborare con ambiti molto diversi tra loro, dalla fisica alle scienze sociali, fino alla linguistica”. Si racconta così Marco Dalla, oggi ricercatore all'Università di Cork, in Irlanda, dove si occupa di intelligenza artificiale e analisi dati.

[as]: In che cosa consiste il tuo lavoro?

[Marco]: Lavoro su un tema dell'informatica teorica che si chiama *boolean satisfiability*, o in italiano “soddisfacibilità booleana”. Questo problema consiste nel determinare se, per una formula booleana composta da più variabili, il cui valore può essere vero o falso, esista una soluzione che renda la formula sempre vera. Spesso questo problema viene risolto cercando di volta in volta soluzioni appropriate. La mia ricerca consiste, quindi, nell'adoperare tecniche di intelligenza artificiale, in particolare *machine* e *deep learning* per analizzare la struttura logica delle formule, stimarne la difficoltà e suggerire strategie di ragionamento che orientino la ricerca di una soluzione.

[as]: Come sei arrivato qui?

[M]: Abbastanza per caso. Mi sono avvicinato all'intelligenza artificiale applicando questi metodi nel campo della fisica delle particelle e, trattandosi di un tema molto in voga, ho pensato che potesse aprire valide opportunità sia nel campo accademico sia nel settore privato.

Ho una laurea triennale in fisica e una laurea magistrale in fisica nucleare e subnucleare, che ho conseguito presso l'Università di Bologna. Durante la tesi magistrale mi sono occupato della caratterizzazione di alcuni sensori dell'esperimento ATLAS al CERN. Non ero però sicuro di voler continuare con lo studio della fisica delle particelle e, su suggerimento del mio relatore, ho iniziato a lavorare su un progetto di ricerca sull'intelligenza artificiale legato a metodi di *machine learning* e *deep learning*. Questo progetto è stato il punto di svolta: mi ha permesso di vincere un dottorato all'Università di Cork, dove lavoro tuttora.

[as]: Hai incontrato difficoltà nella transizione tra fisica e intelligenza artificiale?

[M]: Quando ho iniziato il dottorato, non ero l'unico studente con una formazione diversa dall'informatica. Tra i miei colleghi c'erano statistici, matematici, ingegneri; un collega era perfino laureato in legge.

Nonostante questo, soprattutto all'inizio, ho sentito di avere delle lacune abbastanza grandi, che chi proveniva dall'ingegneria informatica non aveva. Tuttavia, durante i miei studi avevo appreso un metodo di studio efficace: sapevo come studiare e dove cercare le fonti, dovevo solo cambiare argomento. Così sono riuscito a colmare le mie lacune.

[as]: Ci sono stati altri vantaggi legati alla tua formazione?

[M]: Sicuramente ho avuto dei vantaggi legati ad alcune materie che ho approfondito nello studiare la fisica: l'analisi matematica ma anche l'algebra e la geometria servono



molto per studiare proprio le basi teoriche degli algoritmi di intelligenza artificiale. E anche la statistica, che in fisica delle particelle si utilizza sia per l'analisi dati sia per capire alcuni algoritmi di intelligenza artificiale, è stata utilissima.

[as]: Quali sono le differenze che hai notato tra la ricerca in fisica delle particelle e quella in intelligenza artificiale?

[M]: Nella mia breve parentesi di ricerca in fisica delle particelle ho notato che la ricerca si basa sul lavoro di un gruppo: ognuno fa una parte di lavoro e tutto questo contribuisce a un progetto più grande. Per quanto riguarda l'intelligenza artificiale ho svolto la maggior parte delle mie ricerche in solitaria: è molto più facile fare da soli un esperimento dalla A alla Z, basta avere accesso a risorse di calcolo neanche troppo potenti e una buona idea.

Un'altra differenza sostanziale è che spesso, nel campo dell'intelligenza artificiale, il fine pratico della ricerca è molto importante: per trovare fondi e finanziamenti è necessario sottolineare il lato pratico.

Inoltre, nella ricerca in intelligenza artificiale, c'è una maggiore collaborazione con il privato. In tutte le conferenze dove presento i miei lavori, infatti, ci sono molte aziende che partecipano. Ci sono anche molte aziende che investono su

ricerca non applicata e che contribuiscono a finanziare borse di dottorato. Durante il mio percorso di dottorato, ad esempio, avevamo l'opportunità di svolgere dei tirocini all'interno di aziende che avevano contribuito a finanziare le nostre borse. Io personalmente non l'ho fatto e ho scelto di insegnare all'interno dell'università, ma molti miei colleghi hanno fatto dei tirocini curriculari di alcuni mesi in queste aziende per riuscire a creare un contatto diretto e andare a lavorare con loro alla fine degli studi: un valore aggiunto anche per le aziende, che contribuiscono alla formazione di persone con un profilo elevato e possono assumere alla fine del percorso di dottorato candidati molto validi.

Tuttavia, ci sono argomenti molto in voga, come i Large Language Model, dove di fianco a comunità di ricerca aperte esistono grosse aziende come OpenAI o Google, i cui prodotti vengono gestiti con maggiore discrezione: la gran parte del lavoro resta interna, della fase di sviluppo trapela poco e di solito si rende pubblico solo il prodotto finito.

[as]: Che cosa ti auguri per il tuo futuro?

[M]: Mi piace fare ricerca, mi piace insegnare, mi piace l'ambiente universitario. Mi piace il mio lavoro e spero di poter continuare a farlo il più a lungo possibile.

b.

Marco Dalla durante uno dei suoi soggiorni al CERN di Ginevra.



[as] spazi

Il campus dell'INFN

di Martina Bologna



La scienza è fatta di formule, esperimenti e dati, ma soprattutto di curiosità, passione e incontri. È questo lo spirito che ha animato la prima edizione dell'INFN STEAM Summer Camp, un campo estivo nazionale ideato dall'Istituto Nazionale di Fisica Nucleare, in partnership con Cineca e ICSC (vd. p. 34, ndr), con l'obiettivo di avvicinare ragazze e ragazzi al mondo della ricerca scientifica, con uno sguardo aperto verso il futuro.

Dal 13 al 17 luglio, 17 docenti e 97 studenti e studentesse delle classi III, IV e V delle scuole secondarie di secondo grado italiane o degli istituti italiani all'estero si sono ritrovati a Bertinoro (FC) per vivere un'esperienza immersiva all'insegna di scienza, tecnologia, scoperta... e amicizia. Ospitati presso il Centro Universitario Residenziale di Bertinoro e selezionati attraverso progetti INFN per le scuole oppure tramite bando, i partecipanti hanno avuto la possibilità di esplorare alcuni dei temi più affascinanti e complessi della fisica moderna.

Il Summer Camp si è aperto con un'immersione nella fisica delle particelle, grazie agli interventi di alcuni dei protagonisti dell'INFN. Il presidente Antonio Zoccoli ha ripercorso la storia dell'ente, dai ragazzi di via Panisperna fino ai progetti più innovativi, mentre Diego Bettoni ha introdotto i concetti fondamentali della fisica delle particelle, dal bosone di Higgs alle onde gravitazionali. Pierluigi Paolucci ha raccontato l'impegno dell'INFN nel mondo della scuola e Sascha Schmeling ha presentato i progetti educativi promossi dal CERN e dedicati alle scuole, evidenziando il ruolo fondamentale che la formazione ha all'interno della comunità scientifica internazionale. "L'esperienza al Summer Camp INFN è stata molto formativa e stimolante, proprio perché l'anno prossimo studierò fisica all'università e questo mi ha permesso di avere una prima infarinatura dei prossimi anni di studio", commenta lo studente Paolo De Angelis, dell'I.I.S. A. Pacinotti di Scafati (SA), "questi giorni mi hanno lasciato molto emozionato e contento di intraprendere la mia prossima carriera accademica!".

a.
Foto di gruppo del
STEAM INFN Summer
Camp 2025.



b.
Giovani studenti e studentesse impegnati nel laboratorio di gaming dell'INFN Summer Camp 2025.

Il secondo giorno si è svolto, invece, a Bologna, presso il Tecnopolo, cuore pulsante dell'innovazione tecnologica italiana. Accompagnati da Francesco Ubertini, direttore del CINECA, Luca dell'Agnello, direttore CNAF-INFN, Daniele Bonacorsi, Università di Bologna e Daniele Gregori, Chief Scientific Officer dell'azienda E4 Computer Engineering, le studentesse e gli studenti hanno scoperto le potenzialità del supercalcolo, dell'intelligenza artificiale e del *quantum computing*. Hanno inoltre partecipato a una visita guidata al CNAF, il centro di calcolo dell'INFN, dove hanno potuto scoprire da vicino le infrastrutture che supportano la ricerca scientifica. Il pomeriggio è poi proseguito con un laboratorio pratico dove, tra mani sulla tastiera e occhi sul codice, hanno esplorato linguaggi di programmazione, software e strumenti di calcolo avanzato, sperimentando in prima persona la potenza dell'informatica applicata alla scienza.

Oltre agli approfondimenti sulla ricerca scientifica e il supercalcolo e le attività sperimentali, non sono mancati momenti dedicati al gioco e all'immaginazione. Gli studenti e le studentesse sono stati, infatti, protagonisti di laboratori di realtà aumentata e virtuale che hanno permesso alle ragazze e ai ragazzi di entrare, letteralmente, nell'esperimento giapponese Belle II del KEK di Tsukuba e vederne le collisioni, o di provare a "costruire l'universo", ripercorrendo le principali tappe dei primi 20 minuti dopo il Big Bang con un gioco da tavolo sviluppato da INFN Game. L'ultimo giorno è stato poi dedicato alla comunicazione scientifica. Con la guida dell'Ufficio Comunicazione dell'INFN e del fumettista Luca Ralli, le ragazze e i ragazzi hanno partecipato a un laboratorio

di *storytelling* scientifico, provando a trasformare concetti complessi in narrazioni accessibili e coinvolgenti. Tra vignette, metafore e invenzioni narrative, tutti i partecipanti si sono divisi in 15 squadre per elaborare e presentare i propri progetti, esercitando al contempo la creatività e la capacità di lavorare in gruppo.

A chiudere il Summer Camp, la sera stessa, nella cornice della Rocca di Bertinoro, è stata messa in scena la conferenza-spettacolo "Bit e particelle. Un'escursione tra arte e scienza, dal Big Bang al supercalcolo". Seguendo come filo conduttore il supercalcolo e il suo ruolo cruciale nell'analisi e gestione dei *big data* nella ricerca scientifica, il pubblico è stato accompagnato in un viaggio tra le frontiere della fisica, dagli esperimenti del CERN, alle profondità marine di Km3Net, fino a guardare al futuro con Einstein Telescope, due degli esperimenti di punta in cui l'INFN è impegnato attualmente. Ad arricchire la serata, musiche, letture teatrali e illustrazioni per intrecciare i diversi linguaggi in un racconto avvincente e accessibile a tutti.

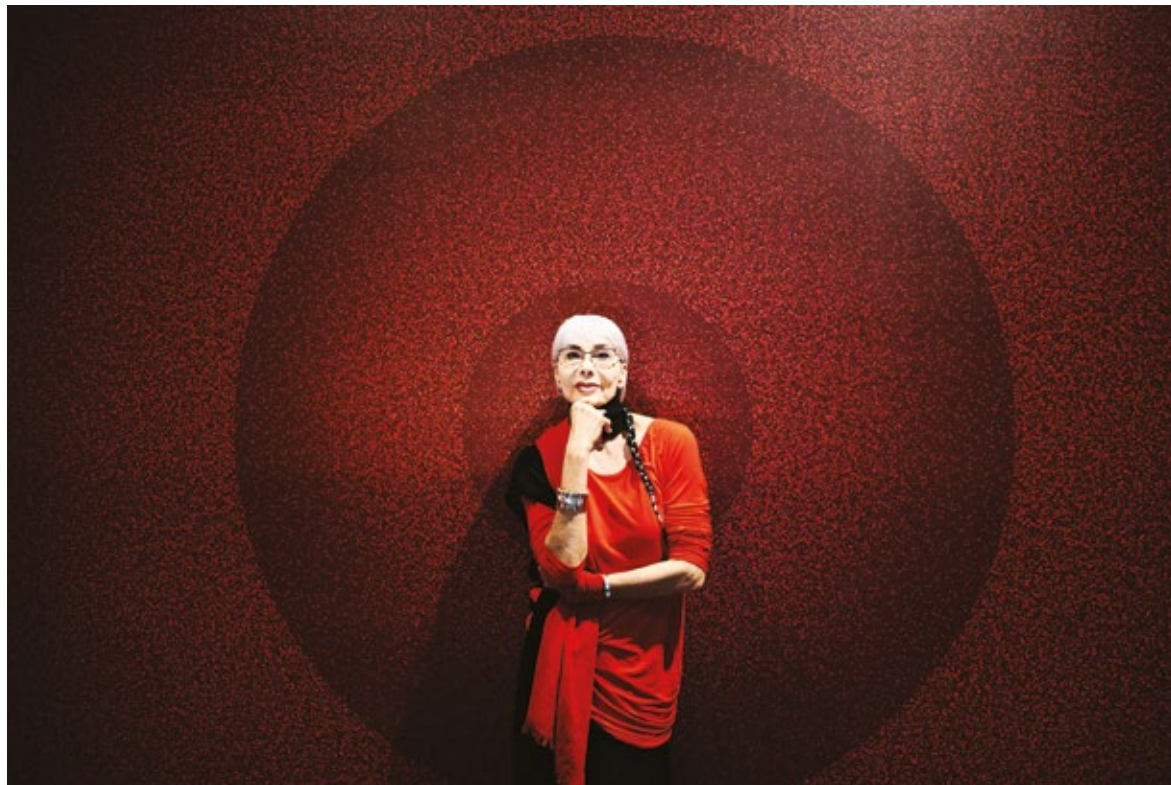
L'INFN STEAM Summer Camp, dunque, non è stato soltanto un'occasione di formazione, ma anche un momento unico di incontro tra ragazze e ragazzi uniti dalla stessa passione: la voglia di capire il mondo attraverso la scienza. Come racconta la studentessa Beatrice Massa Parodi, del Liceo Scientifico Santa Maria Immacolata di Genova, "il Summer Camp è stato un momento di incontro con tantissime nuove persone, nuovi amici, che porterò con me, assieme a tante nuove conoscenze e nozioni, ma soprattutto a un sacco di domande che spero di approfondire nel mio percorso di studi futuri".

[as] convergenze

Lacrime per il bosone di Higgs

di Ottavia Fusco Squitieri

attrice teatrale e cantante



Da sempre la mia passione si è rivolta all'infinitamente piccolo e all'infinitamente grande. Avevo appena sette anni quando presi un foglio dal mio quaderno a quadretti e scrissi (in italiano): "Gentili signori della NASA, sono una bambina di sette anni appassionata di voli Apollo. Se possibile vorrei ricevere la relazione di volo dell'ultima missione e una foto degli astronauti con autografo. Cordiali saluti". Chiusi il foglio in una busta con l'indirizzo NASA-Houston-Texas e imbucai ... Passarono due mesi quando un martedì, tornando da scuola (ricordo che era proprio di martedì, perché all'epoca la RAI trasmetteva all'ora di pranzo "Super Gulp, fumetti in TV"), trovai nella buca da lettere una grande busta gialla con l'aquilotto della NASA e, all'interno, ciò che avevo chiesto. Mi avevano risposto!!! La gioia fu incontenibile, anche perché i miei sogni di bambina erano diventare attrice o

astronauta. Peccato però che in quegli anni, almeno in Italia, le donne non fossero ancora contemplate nei progetti spaziali di volo. Fu così che, senza rimpianti, seguii la strada dell'arte.

Il sogno dell'Infinito, però, non si placava e il mio percorso esistenziale passava attraverso lo studio della filosofia, della mistica e della scienza: parallelismi più che plausibili e che Einstein ci ha ben segnalato. Più che altro sono sempre stata convinta che le grandi intuizioni mistiche delle filosofie orientali (più che di quelle occidentali, data la loro impostazione teologica, oggi a mio avviso del tutto anacronistica) siano sempre più confermate dai progressi della fisica quantistica.

Passando attraverso curiosità e approfondimenti, approdo così a un primo, grande progetto teatrale, che metteva insieme le mie passioni: lo spettacolo "Infinity", scritto e diretto da Ruggero Cappuccio,

con musiche originali di Franco Battiato, ispirato a Valentina Tereshkova, la prima donna nello Spazio. Anche l'incontro con il M° Battiato ha contribuito a confermare le mie convinzioni sul legame stretto tra mistica e scienza. In quell'occasione ebbi modo di conoscere il nostro astronauta Paolo Nespoli, che mi mise in contatto con Mosca, in particolare la Città delle Stelle (il centro di *training*, in cui sono stati formati i cosmonauti russi, ndr), e la grande Valentina attraverso la figlia Elena.

E, sempre grazie a Paolo, approdo poi al CERN di Ginevra. Siamo nel dicembre 2022 e per la prima volta nella mia vita mi ritrovo completamente immersa nell'intelligenza e nello stupore: una sensazione di commozione incommensurabile di fronte al grande mistero dell'esistenza. Respiro l'atmosfera della mitica caffetteria del CERN, dove tante idee sono emerse dai cervelli più brillanti del mondo davanti a una birra o a una tazza di caffè. Incontro Fabiola Gianotti e sotto la sapiente ed entusiasmante guida della comunicatrice scientifica del CERN, Paola Catapano, entro nel tunnel dell'LHC. Da quella visita, mi viene proposto di collaborare – con la mia voce in italiano che racconta la storia dell'universo – al progetto del Science Gateway di Renzo Piano, allora in allestimento e poi inaugurato il 7 ottobre 2023 (vd. in *Asimmetrie* n. 38 p. 46, ndr). Insomma, avevo raggiunto la “quadratura del cerchio” tra il mio essere attrice e le mie passioni scientifiche. Ma un episodio in particolare, che risale al febbraio 2022, credo sintetizzi al meglio le emozioni che ho cercato di comunicare fin qui, ma per raccontarlo sono necessarie un paio di premesse: la prima è che ho il raro privilegio di un'amicizia che va oltre il

senso comune dell'amicizia – di quelle cosiddette “eterne” – con Paola di Genova, detta Paolettina. La seconda è che, al momento dell'episodio che intendo narrare, la mia adorata e novantenne mamma Ester stava malissimo.

Cosa c'entrano queste premesse? C'entrano, c'entrano ... Ero a casa e mi apprestavo a seguire il documentario in onda su RAI3 dedicato a Fabiola Gianotti.

La trasmissione comincia e io la seguo con coinvolgimento ed emozione crescenti, fino a quando arriva il fatidico momento dell'annuncio della scoperta del bosone di Higgs, dall'Auditorium principale del CERN: l'emozione di Fabiola e poi tutti a esultare e poi il prof. Higgs in lacrime e poi e poi e poi ... Io, sempre da sola sul divano di casa, comincio a piangere come una fontana e a singhiozzare come un motore ingolfato. Devo assolutamente, ASSOLUTAMENTE, condividere con qualcuno quell'emozione unica, incontenibile! Acchiappo il cellulare e chiamo Paolettina ... Se quanto segue fosse un copione teatrale, la scena si sarebbe svolta così:

Paolettina: Sì?

Ottavia: Paolettinaaaaaaaa!!! (singhiozzi inarticolati)

Paolettina (spaventatissima): Oh Dio mio Ottavia, la mamma?!?

Ottavia: Il bosone di Hiiiiiiigggggsssss!!! (altri singhiozzi)

(PAUSA)

Paolettina: MA VAFFA*****!!!!!!!!!!!!!!

(CLIC)

Quel pomeriggio ho seriamente rischiato di giocarmi l'amicizia più importante della mia vita.

Cosa non si fa per la scienza...



b.
Ottavia Fusco nel
Science Gateway, al
CERN di Ginevra.

[as] segni



Intelligenza nata da un artificio tecnologico, sintesi del tutto e del nulla, frammenti d'informazione tradotti in conoscenza. Vivida, conserva lo spirito del suo creatore. Algida utilità senza tempo o usbergo di sentimenti?

"IA", pennarello, acquerello e matita su carta, illustrazione e testo di Alessandro Nagar



Istituto Nazionale di Fisica Nucleare

Per contattare
i laboratori dell'Infn:

Laboratori Nazionali di Frascati (Lnf)

T + 39 06 94032423

/ 2552 / 2643 / 2942

comedu@lnf.infn.it

www.lnf.infn.it

Laboratori Nazionali del Gran Sasso (Lngs)

T +39 0862 437265/450

visits@lngs.infn.it

www.lngs.infn.it

Laboratori Nazionali di Legnaro (Lnl)

T + 39 049 8068342 356

stage@lnl.infn.it

www.lnl.infn.it

Laboratori Nazionali del Sud (Lns)

T + 39 095 542296

visiteguidate@lists.lns.infn.it

www.lns.infn.it



Raccontaci che cosa pensi di Asimmetrie su Instagram.
Pubblica storie, post e reel con l'hashtag
#AsimmetriInfn e taggando l'account @infn_insights.
Condivideremo i tuoi contenuti.

www.infn.it



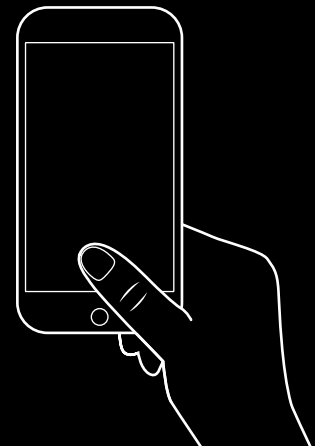
Sul sito **www.asimmetrie.it**
vengono pubblicate periodicamente
notizie di attualità scientifica.

Per **abbonarti gratuitamente** ad Asimmetrie
o per **modificare** il tuo abbonamento vai su:
<http://www.asimmetrie.it/index.php/abbonamento>

Si prega di tenere sempre aggiornato il proprio
indirizzo mail per ricevere le nostre comunicazioni.

Leggi anche le nostre **faq** su:
<http://www.asimmetrie.it/index.php/faq>

Asimmetrie è anche una app,
ricca di nuovi contenuti multimediali.





www.infn.it

rivista online
www.asimmetrie.it