

ALESSANDRO AMOROSO

RELAZIONE TRA L'INTELLIGENZA ARTIFICIALE E I SUOI DATI

SOMMARIO: 1. Introduzione. — 2. Sviluppo dell'intelligenza artificiale. — 3. Reti neurali e machine learning. — 3.1. Come è fatta una rete neurale. — 3.2. Come funziona una rete neurale. — 3.3. Machine learning di una rete neurale. — 4. Conclusioni.

1. INTRODUZIONE.

L'obiettivo di questo lavoro è quello di presentare alcune riflessioni sui sistemi di Intelligenza Artificiale, d'ora in poi IA, quali oggetto di indagine giuridica che prende spunto anche da una recente attività effettuata in qualità di CTU. In questa sede si considera che le IA siano basate su *reti neurali*, come nella stragrande maggioranza degli attuali sistemi. Il presente contributo vuole evidenziare come le differenze sostanziali tra un tradizionale programma informatico e uno di IA richiedano necessariamente di essere valutate con un approccio nuovo.

In questo contributo non si intende affrontare argomenti estremamente affascinanti, quali ad esempio: il significato di intelligenza, le automobili a guida autonoma, i robot umanoidi, oppure l'utilizzo della IA ai fini decisionali in un processo, che richiederebbero una trattazione di tipo enciclopedico. Verrà, invece fornita una descrizione di carattere generale di come siano articolati gli attuali sistemi di IA, senza la quale le argomentazioni espresse risulterebbero assiomatiche. In altri termini, si vuole fornire una descrizione funzionale di cosa sia una IA, in modo da illustrarne gli aspetti fondamentali.

In generale, un sistema informatico tradizionale viene sviluppato dal produttore e spesso viene personalizzato per lo specifico cliente. I risultati generati dipendono dai seguenti fattori principali: la base di dati utilizzata, gli input ricevuti e il codice sorgente usato per svilupparlo. I risultati prodotti da una IA, invece, dipendono oltre che dagli input, anche da due elementi ulteriori, che non sono presenti nei sistemi tradizionali: dalla *rete neurale*, così come è stata definita dal produttore, e dai *dati* con cui questa è stata addestrata. Prima di diventare operativa una IA ha bisogno di essere addestrata con enormi quantità di dati, che la plasmano e che determineranno le risposte fornite agli input

successivi. I dati con cui viene addestrata una IA possono essere forniti dall'utente o provenire da terzi e naturalmente devono essere di qualità per evitare un effetto ben noto in informatica: il cosiddetto *GIGO*, dall'inglese *Garbage In - Garbage Out* ¹.

Di conseguenza, la valutazione di un sistema di IA non può limitarsi alla sola valutazione del cosiddetto *codice sorgente* della rete neurale, ma deve necessariamente considerare anche i *dati* che sono stati utilizzati per il suo addestramento, e gli effetti che questi hanno prodotto sulla rete neurale originale. Questo aspetto ha una particolare rilevanza, ad esempio si consideri un contratto in cui non sia stato stabilito chi sia il titolare dei molteplici diritti del sistema di IA ² in uso presso un cliente. Questa IA generalmente è il risultato dalla fornitura della rete neurale al cliente che la addestra con i propri dati. La sola rete neurale originale, così come fornita dal produttore, per quanto raffinata e sofisticata non è in grado di elaborare alcun risultato se prima non è stata addestrata. Parimenti, i dati che sono utilizzati per l'addestramento non hanno di per sé alcuna capacità elaborativa e sono spesso complessi da analizzare. Inoltre, spesso i sistemi di IA "imparano" anche nel corso del loro utilizzo; quindi, evolvono durante il loro utilizzo.

Il caso in cui i dati di addestramento non siano di titolarità dell'utente ma di soggetti terzi introduce un ulteriore livello di complessità contrattuale, ma non cambia i principi sostanziali esposti in questa sede.

In estrema sintesi, in questo contributo si intende evidenziare come la classica identificazione del *codice sorgente* quale oggetto da tutelare nel software tradizionale sia ampiamente insufficiente nel caso di sistemi di AI, dove i *dati* utilizzati per l'addestramento della IA sono altrettanto importanti. Si vedrà di seguito come nell'ambito delle IA sia necessario superare la mera tutela del *codice sorgente* per estenderla, almeno, anche ai dati di addestramento.

¹ Letteralmente *spazzatura in ingresso - spazzatura in uscita*. Espressione coniata alla fine degli anni '50 dello scorso secolo, spesso riassunto con l'acronimo *GIGO*, esprime il concetto che se i dati in ingresso ad un sistema di calcolo sono di bassa qualità, ossia *garbage*, allora anche i risultati saranno di bassa qualità, ossia ancora *garbage*. Nel tempo *GIGO* si è rivelata applicabile per le basi di dati, dove, ovviamente, se i dati memorizzati sono di bassa qualità, allora anche i risultati che si otterranno interrogando i dati memorizzati sa-

ranno di bassa qualità. Nel caso delle IA indica che i risultati ottenuti sono errati se quelli con cui è stata addestrata la IA sono a loro volta errati.

² Come è ampiamente noto, molte controversie legali relative all'informatica traggono origine da contratti spesso lacunosi o decisamente oscuri e ambigui. Ad esempio, nella mia esperienza di consulente ho visto contratti in cui l'oggetto stesso era estremamente vago, oppure contratti che erano semplicemente le slide di una presentazione firmate dalle parti.

2. SVILUPPO DELL'INTELLIGENZA ARTIFICIALE.

Da sempre si è subito il fascino di macchine in grado di esibire una qualche forma di intelligenza o di comportamento autonomo. Risalgono già al mondo ellenistico descrizioni di sofisticati dispositivi meccanici in grado di svolgere calcoli specifici. Nella antica letteratura cinese sono descritti diversi tipi di automi meccanici ³. La cosiddetta *macchina di Anticitera* — che risale ad un periodo tra il 150 e il 100 a.C. e che fu rinvenuta nel relitto di una nave romana nel 1900 — era un complesso planetario meccanico che per mezzo di ruote dentate calcolava, tra l'altro, il sorgere e il tramonto del sole e le fasi lunari ⁴. Lo studioso islamico *Al Jazari* nel suo libro più famoso ⁵ già nel 1206 descriveva in dettaglio 50 macchine, tra cui dieci fontane che cambiavano forma. Sempre *Al Jazari* mostrava un fattivo interesse nel realizzare macchine con sembianze umane per lo svolgimento di compiti specifici ⁶. Anche Leonardo Da Vinci progettò un cavaliere meccanico in armatura apparentemente in grado di compiere alcuni movimenti, tra cui alzarsi in piedi e muovere braccia e mascella ⁷. Più di recente, nella seconda metà del diciottesimo secolo, ebbe molto successo un automa meccanico che giocava a scacchi, chiamato *Il turco*, che veniva esibito nelle corti europee. Quasi un secolo dopo si scoprì che l'automa nascondeva al suo interno una persona che lo muoveva; quindi, l'intero meccanismo serviva essenzialmente per suscitare stupore e per celare la presenza dello scacchista umano ⁸. Naturalmente, a nessuno di questi esempi si associa l'idea di comportamento intelligente della macchina.

Una trentina di anni fa, nel 1996, accadde un evento considerato straordinario: per la prima volta un computer vinse una partita di scacchi contro il campione del mondo in carica, il russo Garry Kasparov. Il computer era il famosissimo *Deep Blue* ⁹ che non aveva affatto sembianze umane ed aveva sostanzialmente l'aspetto di un armadio metallico grande come un frigorifero.

³ Ad esempio, in NEEDHAM, JOSEPH, *Science and Civilisation in China, Part 2, Mechanical Engineering*, Cambridge University Press, 1965.

⁴ Questa macchina viene considerata, tra l'altro, il più antico calcolatore meccanico conosciuto. Una più completa descrizione può essere trovata in WRIGHT, MICHAEL T., "The Antikythera mechanism reconsidered", *Interdisciplinary science reviews*, 32.1 (2007): 27-43.

⁵ KING, DAVID A., "The Book of Knowledge of Ingenious Mechanical Devices: Kitab fi ma'arifat al-hiyal al-handasiya

by Ibn al-Razzaz al-Jazari" by DR Hill", *History of Science* 13.4 (1975): 284-289.

⁶ Si facciariferimento a ROSHEIM, MARK E., *Robot evolution: the development of robotics*, John Wiley & Sons, 1994.

⁷ Si trovano alcuni disegni progettuali nel *Codice Atlantico*.

⁸ STANDAGE, TOM, *The Turk: The life and times of the famous eighteenth-century chess-playing machine*. Walker & Company, 2002.

⁹ HSU, FENG-HSIUNG, *Behind Deep Blue: Building the computer that defeated the world chess champion*, Princeton University Press, 2002.

Deep Blue perse comunque l'incontro, articolato in più partite ¹⁰. L'anno successivo, invece, una versione più evoluta di Deep Blue si aggiudicò l'intero incontro, sempre contro il campione mondiale Kasparov ¹¹.

Questi eventi vennero considerati da molti come primi casi di computer "intelligente", pur se estremamente specializzato e non ancora generico. In realtà si trattava di applicazioni di tecniche computazionali ben note e molto raffinate, rese possibili dalla enorme, per i tempi, capacità di calcolo di Deep Blue.

L'idea che fosse possibile realizzare macchine "intelligenti" risale ad almeno la metà dello scorso secolo, quando Alan Turing, uno dei padri fondatori dell'informatica, la ipotizzò nel suo lavoro seminale intitolato « Computing Machinery and Intelligence » ¹². In quello scritto Turing proponeva anche il suo famoso test per definire il comportamento intelligente di un computer. In estrema sintesi, Turing proponeva la seguente metodologia: se un computer interrogato da una persona tramite una interfaccia testuale fosse stato capace di essere scambiato per una persona, dissimulando la sua natura di macchina, allora avrebbe mostrato un comportamento intelligente ¹³.

Tutti i passati tentavi di realizzare delle IA si sono scontrati con la insufficiente potenza di calcolo e la capacità di memorizzazione dei computer allora disponibili. La potenza di calcolo nel tempo ha seguito una crescita letteralmente esponenziale, raddoppiando ogni due anni ¹⁴. Anche la capacità di memorizzazione è enormemente cresciuta nel tempo.

Fino alla fine del 2022 l'IA era una tecnologia limitata a poche specifiche applicazioni, molto specialistiche, ed era considerata dai non specialisti quasi esclusivamente per le sue apparizioni nel mondo della fantascienza ¹⁵. La pubblicazione sul web di *ChatGPT* ha completamente cambiato questa situazione. Chiunque poteva interagire con una IA tramite linguaggio naturale,

¹⁰ In un incontro di sei partite Deep Blue vinse una partita, ne pareggiò due e ne perse tre.

¹¹ Deep Blue vinse due partite, ne pareggiò tre e ne perse solo una. Un dettagliato resoconto di entrambi gli incontri può essere consultato al sito <https://www.chess.com/article/view/deep-blue-kasparov-chess> (ultimo accesso marzo 2024).

¹² Pubblicato nel 1950 e liberamente disponibile su numerosi siti online, ad esempio al seguente indirizzo (ultima consultazione febbraio 2024): <https://redirect.cs.umbc.edu/courses/471/papers/turing.pdf>.

¹³ Sostanzialmente l'utilizzo di una interfaccia testuale, come ad esempio il terminale di un computer, semplifica il problema, escludendo la necessità anche di eventuali

simulazioni visive e vocali, focalizzandosi solo su aspetti di "intelligenza".

¹⁴ Restando ancora oggi aderente alla cosiddetta *Congettura o Legge di Moore* formulata nel 1965 e leggermente aggiornata successivamente; Gordon Moore fu uno dei fondatori e uno dei CEO della società di microprocessori Intel. Per ulteriori approfondimenti si veda, ad esempio, <https://www.britannica.com/technology/Moores-law> (ultima consultazione luglio 2024).

¹⁵ Si ricordano tra i più famosi: il computer *Hall 9000* del film *2001 Odissea nello Spazio* di Stanley Kubrick, il robot *R. Daneel Olivaw* che compare in molti racconti di Isaac Asimov, il robot *Andrew* del film *L'uomo del Bicentenario* di Chris Co-

anche se solo in forma di inglese scritto, ossia si potevano formulare domande e ricevere risposte su moltissimi argomenti. Dopo cinque giorni dalla pubblicazione *ChatGPT* aveva già più di un milione di utenti registrati, che hanno superato i 100 milioni in meno di due mesi ¹⁶. Questo tasso di crescita del numero di utenti registrati soverchia tutti quelli precedentemente noti. Per avere una idea, basta considerare che per arrivare al loro primo milione di utenti registrati *Facebook* ha impiegato dieci mesi e *Twitter* due anni ¹⁷.

L'eccezionale successo di *ChatGPT* ha suscitato al contempo grandi entusiasmi e profonde paure. Da un alto, l'attuale nuova generazione di IA, di cui *ChatGPT* è stato l'alfiere, prospetta futuristici scenari in cui l'IA verrà utilizzata per svolgere molti lavori che adesso si ritiene possano essere svolti solo da persone; questi lavori potrebbero essere eseguiti con maggiore accuratezza e velocità, ad un costo estremamente basso, permettendo all'intera umanità sviluppi adesso neanche immaginabili. Dall'altro lato ¹⁸, si teme che una larga e capillare diffusione di sistemi basati su IA possa portare a catastrofi epocali: si potrebbe arrivare allo sviluppo di una imprevista autocoscienza delle macchine, che potrebbero addirittura prendere il sopravvento sull'intera umanità ¹⁹.

Siamo solo agli inizi degli sviluppi tecnologici delle IA, che inevitabilmente presenteranno nuove sfide concettuali e che adesso propongono nuove prospettive. Di seguito si presenteranno gli aspetti salienti delle attuali AI per proporre migliori strumenti concettuali necessari a interpretarle.

Le intelligenze artificiali hanno in un certo senso superato la classica idea di algoritmo. Gli algoritmi sono tuttora esistenti nell'ambito delle IA, ma rappresentano solo una parte del tutto. Le IA attuali sono essenzialmente il risultato di una sorta di motore per l'apprendimento che è stato addestrato su dati specifici. Gli algoritmi svolgono un ruolo primario nelle tecniche di apprendimento, ossia nella progettazione e nello sviluppo del motore di apprendimento. La IA risultante però dipende forte-

lumbus e l'intelligenza artificiale *Samantha* del film *Her* di Spike Jonze.

¹⁶ Si veda il The Guardian all'indirizzo <https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app>.

¹⁷ Interrogazione nel luglio 2024 della IA *Perplexity* (<https://www.perplexity.ai>) con il seguente prompt « how long it took to facebook to reach 1 million registered users? »

¹⁸ L'idea che la creatura possa ribellarsi al creatore è spesso indicata come *Sindrome di Frankenstein*, dal mito del ro-

manzo di Mary Shelley. In tempi più recenti, i film *2001 Odissea nello Spazio* e *Blade Runner*, riproponevano lo stesso mito ambientato in un futuro non troppo remoto.

¹⁹ Alcuni ipotizzano che si raggiungerà la cosiddetta *Singolarità*, che sarà quando le IA supereranno l'intelligenza complessiva dell'intera specie umana e le macchine inizieranno autonomamente a progettare e costruire altre macchine sempre più intelligenti. Si faccia riferimento, tra gli altri a <https://www.britannica.com/technology/singularity-technology> (ultima consultazione agosto 2024).

mente anche dai dati che sono stati utilizzati per addestrare il motore. Il solo motore per l'apprendimento non è in grado di produrre alcun risultato, per quanto possa essere estremamente raffinato e sofisticato. Parimenti, anche i soli dati non posso produrre alcun risultato; oltretutto i dati sono un insieme passivo che non è in grado di elaborare alcunché. È fondamentale evidenziare che la qualità di una IA dipende fortemente anche dalla qualità dei dati su cui è stata addestrata. Solo dopo che il motore di ricerca è stato addestrato su quegli specifici dati diventa una IA.

Allo stato attuale, la tecnologia basata sulle *reti neurali* è alla base della gran parte dei motori per l'apprendimento delle IA. Attraverso un processo detto di *machine learning*, una rete neurale viene addestrata, per mezzo di un insieme cospicuo di dati, a riconoscere gli scenari che questi dati rappresentano. Dopo la fase di *machine learning*, una rete neurale diventa capace di riconoscere gli scenari che ha "imparato" nei nuovi dati che riceve in input. Inoltre, una rete neurale potrebbe continuare ad "imparare" nuovi scenari dai dati ricevuti come input.

Molti dei termini appena utilizzati possono sembrare eccessivamente antropomorfi e suggerire una sorta di personalità delle macchine e nello specifico delle IA. I termini tecnici in uso richiamano il meccanismo tecnico a cui si riferiscono suggerendo una antropomorfizzazione della macchina, ma questi spesso ben identificano concetti che altrimenti richiederebbero dei neologismi per essere indicati.

2.1. Come è fatta una rete neurale.

Una rete neurale è un insieme di unità elaborative piuttosto semplici, dette *nodi*, *perceptroni*, o *neuroni*, collegate tra loro per mezzo di canali di trasmissione detti *link*. La Figura 1 illustra una semplicissima rete neurale. I nodi sono rappresentati con dei cerchi e i link con delle frecce. Ogni link parte da un nodo, all'inizio della freccia, e arriva in un altro nodo, alla punta della freccia.

Ogni link rappresenta un canale di comunicazione tra il nodo di partenza e quello di arrivo; in altri termini, il flusso dei dati segue il verso delle frecce. Ad ogni link è associato un *peso numerico*, non rappresentato in figura, che può cambiare nel corso dell'apprendimento. Ogni nodo calcola un risultato in base ai valori di input trasmessi dai link in ingresso, correlati ai pesi numerici dei rispettivi link. I pesi hanno un ruolo fondamentale sia nell'*apprendimento* sia nel successivo funzionamento della rete neurale.

I nodi sono organizzati in *strati*, che nella Figura 1 sono rappresentati come nodi incolonnati. Nello specifico, la figura mostra una rete neurale a quattro strati, in cui i dati fluiscono da sinistra verso destra. Il primo strato, quello più a sinistra,

rappresenta lo *strato di input*, ossia lo strato che riceve i dati che dovranno essere elaborati dalla IA. Il numero di link che arrivano allo strato di input definisce il numero di *feature*, ossia *caratteristiche*, che costituiscono ogni singolo input. La Figura 1 mostra una rete neurale che accetta tre caratteristiche per ogni input, i.e. ogni singolo input è costituito da una terna di valori.

I due strati intermedi mostrati nella Figura 1 sono detti *strati nascosti*, perché sono invisibili dall'esterno del sistema, che mostra solo gli strati di input e di output. Gli strati nascosti possono essere anche più di due e contenere molti più nodi di quelli mostrati. I risultati sono prodotti dallo strato di output.

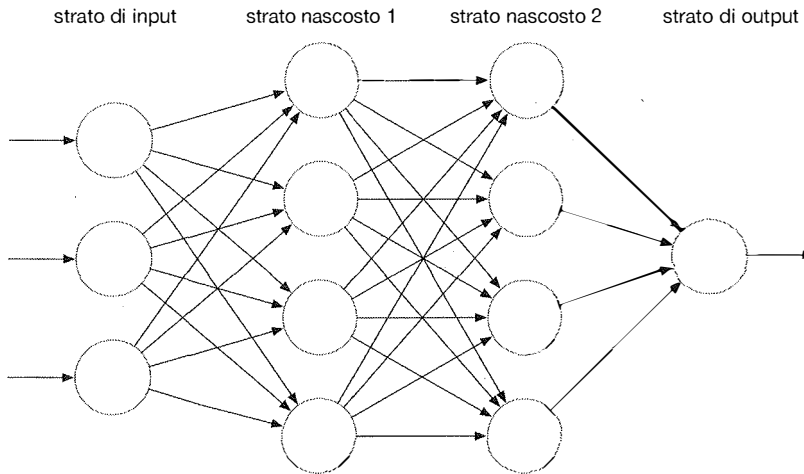


Figura 1. Schema di una semplice rete neurale

● Ogni link che collega un nodo ad uno dei nodi dello strato successivo viene denominato *parametro* della rete neurale. Il numero di parametri è una metrica per misurare la complessità di una rete neurale. La Figura 1 mostra una rete neurale con 35 parametri²⁰. Per avere un'idea della dimensione della rete neurale di ChatGPT 3, ossia la versione meno potente, non a pagamento, si consideri che la stima è di 175 *miliardi* di parametri. Per ChatGPT 4, ossia la successiva versione, a pagamento, la stima

²⁰ Per spirito di precisione, ogni nodo ha un ulteriore link in ingresso, detto *bias*, che in questo contributo non viene considerato, per evitare un eccessivo appesantimento della trattazione, senza perdere significativamente in completezza. Si può considerare il *bias* come una sorta di input ulteriore per ogni nodo, che è utilizzato da questo per calcolare il proprio risultato assieme ai valori dei link e dei loro pesi. In effetti, quindi, la rete neurale mostrata in Figura 1 ha 47 parametri. In questo caso il

termine *bias* non ha nessuna accezione negativa, ma sta solo ad indicare un elemento che altera leggermente gli input. Questo *bias* è altro rispetto agli eventuali pregiudizi, *bias* appunto, che una IA potrebbe sviluppare a livello operativo, ad esempio come descritto in <https://www.technologyreview.com/2020/12/10/1013617/racism-data-science-artificial-intelligence-ai-opinion/> e anche in <https://time.com/5520558/artificial-intelligence-racial-gender-bias/> (ultima consultazione agosto 2024).

del numero di parametri è di 1,76 *trillioni*, ossia *mille miliardi*, quindi circa dieci volte più complessa del predecessore ²¹.

2.2. Come funziona una rete neurale.

I nodi dello strato di input, che nella Figura 1 sono mostrati più a sinistra, ricevono l'insieme di input, che è costituito da un certo numero di caratteristiche che lo definiscono. Ogni nodo riceve una singola caratteristica. A esempio, un insieme di terne di valori in input per la rete mostrata in figura potrebbe essere composto da tre caratteristiche quali il nome, il peso, l'altezza di una persona. In questo caso, il genere di quella persona potrebbe essere ignoto e il risultato della computazione sarebbe il genere più probabile per quella persona. Il ruolo dei nodi di input è molto semplice: si limitano a trasferire il loro input ad ognuno dei nodi a cui sono collegati.

Ognuno dei nodi di uno strato nascosto, i due strati centrali mostrati nella figura, calcola il proprio risultato in base alla somma dei valori in ingresso da ogni link, ognuno di questi moltiplicato per il peso numerico del relativo link, e ad una *funzione di attivazione*, calcolata su questa somma ²². Il risultato viene trasmesso allo strato nascosto successivo, i cui nodi ripetono lo stesso tipo di computazione con gli opportuni valori numerici. Gli strati nascosti possono essere più di due, e composti da un differente numero di nodi, ma comunque continuano a ripetere lo schema computazionale appena descritto.

Quando i dati giungono ai nodi di output, lo strato più a destra nella Figura 1, questi restituiscono il risultato calcolato dalla rete neurale. Restando aderenti all'esempio, l'output della rete neurale mostrata in Figura 1 sarebbe il genere presunto della persona corrispondente alla terna in input.

2.3. Machine learning di una rete neurale.

Come suggerito da quanto esposto in precedenza, i pesi dei link che trasmettono i dati in una rete neurale giocano un ruolo fondamentale nella computazione della rete, ossia nelle risposte della IA. Il valore di questi pesi rappresenta quanto "imparato" dalla rete neurale. In altri termini, la cosiddetta fase di apprendimento della rete neurale ha lo scopo di stabilire i pesi numerici

²¹ OpenAI, la società che produce ChatGPT, non ha comunicato questi valori; i numeri indicati sono da considerarsi delle stime, che però si suppone siano decisamente verosimili e sono condivise da molti studiosi. Si faccia riferimento, ad esempio, all'indirizzo <https://medium.com/@francesc>

[o.casalegno/chatgpt-explained-whats-inside-the-ml-model-from-gpt-1-to-gpt-4-631d325368ef](https://medium.com/@francesc) (ultimo accesso marzo 2024).

²² I valori di *bias*, citati in una nota precedente e non sono mostrati in figura, contribuiscono al calcolo della funzione di attivazione.

dei link in modo che successivamente la rete possa produrre il risultato più probabile in base ai dati di tutti gli input utilizzati per l'addestramento.

L'addestramento di una rete neurale si svolge in turni successivi sottoponendo alla rete neurale una larghissima quantità di insiemi di dati in input, uno per volta. Per ogni insieme si valuta la risposta e si aggiorna il sistema di conseguenza. È necessario evidenziare che i pesi numerici dei parametri hanno un valore iniziale generico prima della fase di addestramento della rete e sono messi a punto nel corso dell'addestramento. Risulta impossibile ricostruire la sequenza di passaggi successivi in base alla quale ogni singolo peso numerico abbia acquisito il valore corrente.

L'addestramento di una rete neurale è algoritmico, nel senso che i meccanismi utilizzati sono stabiliti con specifici algoritmi. In estrema sintesi si possono identificare due principali famiglie di addestramento: con o senza supervisione.

Nel cosiddetto addestramento *con supervisione* di una rete neurale si utilizzano dati completi, ossia che hanno delle "etichette" che li identificano; in altri termini, si utilizzano dati di cui si conosce a priori la risposta. In base al confronto fra il risultato calcolato e quello corretto, la rete ricalcola i pesi numerici di ogni link con un meccanismo detto di *back propagation*. Questo aggiustamento dei pesi dei link va nella direzione di ridurre il più possibile la differenza fra i due risultati calcolato e corretto.

Nell'esempio già presentato, un insieme di dati per l'addestramento potrebbe essere composto da quaterne complete, ossia con il nome, il peso, l'altezza e il genere di ogni persona. Della intera quaterna solo il nome, il peso e l'altezza costituirebbero l'input, mentre il genere costituirebbe l'etichetta, ossia il valore corretto con cui confrontare la risposta.

Il cosiddetto addestramento *senza supervisione* di una rete neurale prevede che i dati non abbiano etichette. In questo caso la rete neurale cerca di identificare le affinità tra gli input raggruppandoli in base a determinate caratteristiche.

È importante considerare anche l'ordine di grandezza dell'insieme di dati necessario per addestrare una IA. Secondo stime condivise, si ritiene che l'intero dataset utilizzato per addestrare ChatGPT 3 sia composto da più di 500 miliardi di *token*, ossia di parole e loro frammenti²³. Per dare un'idea di questa misura si consideri che l'intera Wikipedia in inglese, ossia una delle più

²³ In informatica il termine *token* assume significati differenti in base al contesto in cui viene utilizzato. Nel caso della linguistica computazionale un *token* verbale è una parola o un suo frammento significa-

tivo, che non viene ulteriormente scomposto, definendo quindi una sorta di unità atomica. Questo significato è altro, ad esempio, da quelli di *token* utilizzati nelle

grandi basi di conoscenza organizzate del modo è composta da circa 2,5 miliardi di token. In altri termini, si stima che per addestrare ChatCPT 3 sia stato necessario un insieme di dati circa duecento volte più esteso della intera Wikipedia in inglese ²⁴.

A livello macroscopico, nell'esempio descritto in precedenza, le caratteristiche che determinano il risultato, ossia il genere della persona, potrebbero essere sia la lettera finale del nome — che in italiano segue delle regole rigide con poche eccezioni e se le persone dell'insieme di addestramento hanno nomi italiani — sia le due caratteristiche fisiche di peso e altezza, sia una combinazione di questi. Sempre a livello macroscopico, senza sostanziali modifiche dei meccanismi di funzionamento e apprendimento della rete neurale non c'è modo di sapere quale siano le motivazioni che hanno portato la rete a dare le risposte che ha dato. Spesso è molto complesso, o addirittura impossibile, effettuare queste modifiche, che comportano un incremento notevole del grado di complessità del sistema.

In base a quanto su esposto, le attuali IA più diffuse sono il risultato dell'apprendimento dai dati eseguito da una rete neurale.

Una rete neurale può essere estremamente sofisticata e complessa oltre che composta da un numero enorme di nodi e link. Anche solo basandosi sulla sommaria esposizione precedente, si deduce che la complessità di una rete neurale deriva, tra altre cose, dal numero di nodi, dai collegamenti fra di essi, dalla loro disposizione in strati, delle funzioni di attuazione, dalle regole di *back propagation* e molto altro.

Parimenti, un insieme di dati, necessariamente cospicuo e ricco di informazioni, ha valore solo se analizzato attraverso sistemi di elaborazione. Di per sé, un insieme di dati risulta sostanzialmente impossibile da sfruttare, ad esempio per estrarne risultati aggregati; a tale scopo occorrono sofisticati strumenti di analisi, spesso basati su metodologie statistiche.

Nessuna delle due entità precedenti, rete neurale e insieme di dati per l'addestramento, ha un valore rilevante senza la necessaria relazione con l'altro. In altri termini, si può affermare che l'insieme è decisamente superiore alla somma delle parti.

Stanno diventando di realizzazione piuttosto comune le cosiddette personalizzazioni delle IA. Ad esempio, una azienda che realizza una personalizzazione di ChatGPT partendo dalla versione corrente e addestrandola ulteriormente anche su propri

blockchain oppure nei meccanismi di autenticazione e sicurezza.

²⁴ Si faccia riferimento, ad esempio, a <https://expertbeacon.com/chatgpts-training-dataset-a-treasure-trove-for-scientific-research/> e anche a <https://medium.com/@dlaytonj2/chatgpt-show-me-the-data-sources-11e9433d57e8> (ultima consultazione agosto 2024).

earch/ e anche a <https://medium.com/@dlaytonj2/chatgpt-show-me-the-data-sources-11e9433d57e8> (ultima consultazione agosto 2024).

dati. In questo caso otterrebbe una AI che saprebbe sfruttare le capacità linguistiche della IA originale e che è in grado di interagire anche su argomenti propri dell'azienda, realizzando ad esempio un *chatbot* aziendale.

È necessario, a giudizio dello scrivente, che tutte le parti di una fornitura di una IA definiscano con attenzione, già in sede contrattuale, quali siano i diritti di tutte le parti, anche per i futuri sviluppi.

Nel corso delle attività di una analisi di una IA occorre sempre tenere presente che una IA è il frutto almeno del contributo sia dei dati sia della rete neurale. Per meglio definire il ruolo delle due parti occorre considerare che la stessa rete neurale di partenza può evolvere in modi anche molto diversi tra loro a seconda dell'insieme dei dati di addestramento e del suo uso sino a quel momento. Allo stesso modo, utilizzando lo stesso insieme di dati di addestramento per due differenti reti neurali si otterranno differenti IA.

Queste prime riflessioni hanno lo scopo di evidenziare che, a differenza di quanto accade nel codice tradizionale, nell'ambito delle IA ha un ruolo fondamentale l'insieme della rete neurale, che costituisce il codice sorgente, e dei dati su cui quel codice è stato addestrato. Quindi, mentre per un sistema informatico tradizionale tutte le considerazioni sulle funzionalità e sui diritti si possono sostanzialmente risolvere analizzando il codice sorgente, questo non accade più nei sistemi di IA, dove bisogna considerare anche l'interrelazione tra codice e dati di addestramento.

Abstract

This work considers Artificial Intelligence as object of legal investigations. For sake of simplicity, we may consider that AI is implemented by neural networks, as it happens in most of the current cases.

An AI system is made of both the neural network and the data set used to train it. Neither of them is enough to define the AI, in other words, we can say that the whole is bigger than the sum of its parts. The mere analysis of the source code, ignoring the effects of the training data set, it is not adequate to understand an AI.

It results that AI requires a new approach with respect to traditional software.