

Alfabetizzare le macchine: il problema della decisione

di Alfonso Musci

16-02-2022

«L'esplicabilità è fondamentale per creare e mantenere la fiducia degli utenti nei sistemi di "intelligenza artificiale" (IA). Non sempre è possibile spiegare, tuttavia, perché un modello matematico, statistico o algoritmico ha generato un particolare risultato o decisione (e quale combinazione di fattori di input vi ha contribuito). È il cosiddetto caso della "scatola nera" (black box algorithm) i cui algoritmi richiedono un'attenzione particolare ...». È quanto si legge nel documento Orientamenti etici per un'IA affidabile elaborato da un Gruppo di esperti ad alto livello sull'intelligenza artificiale incaricato dalla Commissione europea (2018).

Il grande quesito relativo all'alfabetizzazione delle macchine e alla loro autonomia decisionale non è una novità dei nostri tempi, ha una storia lunga. Inedita è però la presente dimensione del problema e delle sue conseguenze nell'orizzonte dell'apprendimento elettronico, (machine learning) nei vari processi di automazione scientifica, organizzativa, industriale e relazionale che investono le sfere sociali e biologiche. Questo ha spinto il nostro massimo decisore e regolatore politico (la Commissione europea) a occuparsi della questione in termini etici e inevitabilmente bioetici e politici.

A questo stato dell'arte si può aggiungere un ulteriore sospetto relativo alla domanda: chi istruisce le macchine? È il problema del "deep learning" (alfabetizzazione occulta delle macchine), una delle minacce del nostro tempo: black box (scatole nere) che macinano imponenti volumi di dati pronosticano, giudicano, votano e infine suggeriscono decisioni coerenti ma potenzialmente violente e ingiuste come adeguate e rispettose, senza spiegarne il perché (ne ha scritto Frank Pasquale, The Black Box Society, 2015).

Macchine istruite, geniali, mistiche, che operando su modelli risultanti ereditano in modo acritico giudizi, pregiudizi e inclinazioni (bias) nascosti nei dati di allenamento e apprendimento: tendenze destinate a comparire negli algoritmi di decisione. A queste condizioni, l'eventualità di un'estensione illimitata del machine learning e dell'automazione ai settori della pubblica amministrazione, della gestione delle risorse umane e dei servizi (sanità, trasporti, scuola, difesa, politica ecc...) potrebbe essere foriera di rischi imponderabili. Una dinamica fuori dal controllo democratico e tecnico, in cui gli algoritmi di decisione (i consulenti sciamani) potrebbero suggerire scelte di volta in volta coerenti (giuste, ingiuste, discriminatorie o sbagliate), ma imperscrutabili al loro interno e nel loro processo per il professionista, il dirigente, il medico, il politico o chi in generale delegasse loro la propria decisione o parte di essa (committente o apprendista stregone).

Occorre pertanto ribadire l'ovvio, che la macchina è per definizione actus e non agens, ma allo stesso tempo ammettere anche che la prerogativa della macchina di istruirsi, collezionare dati, confrontarli e fare valutazioni statistiche in una dimensione e in un tempo sempre più veloce renderà sempre più imperscrutabile (misterioso) il motivo del suo comportamento. Al punto da creare in chi la osserva all'opera l'effetto illusorio ma verosimile di una sua possibile autonomia decisionale (impostura). La macchina così da vaso d'elezione (vas electionis) diverrebbe non solo sibilla e veggente ma vaso di pandora.

Una frattura antropologica per la verità esiste ed è profonda. Nella relazione con le informazioni, macchine e società umane divergono radicalmente. Se il sovraccarico di informazioni produce uno strano e disinvolto effetto di leggerezza e protagonismo intellettuale delle macchine, al contempo schiaccia e aggrava la capacità di pensare autonomamente e in prospettiva delle società umane. A scavare questa divergenza e ad alimentare la dipendenza dalla macchina è proprio il medesimo

bisogno degli uomini di alleggerirsi, di riconquistare tempo, a patto però di saperlo poi vivere degnamente.

Credere tuttavia all'impostura di fondo presuppone un'idea sbagliata della scolarizzazione (senza maestri in carne e ossa e oltre i limiti ordinari del tempo) e un'idea sbagliata dell'automazione (illimitato potere e illimitata autonomia dell'artificiale). Wishful thinking di una tecnocrazia mantica e demiurgica al contempo. Risultato dell'estensione fuori confine di un matematismo senza pensiero. Ma della natura di questo falso mistero della decisione furono al contrario ben consapevoli proprio i padri del pensiero matematico contemporaneo.

Nel 1928 David Hilbert formulò l'Entscheidungsproblem (problema della decisione): esiste o no un metodo meccanico e rigoroso per decidere e stabilire sempre e in ogni caso se un'affermazione è vera oppure no? Due anni più tardi (1930) Kurt Gödel, schierandosi per il no, avanzò i due teoremi dell'incompletezza e della coerenza che si possono parafrasare come segue: 1) incompletezza: in ogni sistema formale coerente (vale a dire in grado di fornire una risposta univoca: cioè o vera o falsa e mai né vera né falsa o sia vera che falsa), vi sarà sempre una proposizione che, pur se sintatticamente corretta, sarà indimostrabile o inconfutabile all'interno del sistema stesso; 2) coerenza: nessun sistema formale coerente può dimostrare autonomamente il processo della sua coerenza.

Sei anni dopo (1936) toccò occuparsene, in una prospettiva meno astratta, anche ad Alan Turing (On Computable Numbers with an Application to the Entscheidungsproblem, «Proceedings of the London Mathematical Society», 1/1937). Per farlo immaginò l'esistenza di una macchina computazionale composta da un nastro infinito bidirezionale, diviso in caselle di memoria e da una testina in grado di leggere e scrivere elementi numerici a sistema binario (0 e 1) nelle singole caselle muovendosi, a seguito di un innesco (numeri scritti originariamente dall'uomo nelle caselle), lungo il nastro e in entrambe le sue direzioni. Turing dimostrò l'impossibilità per la macchina stessa di decidere autonomamente il suo innesco, il come procedere, il quando e il dove fermarsi. Era questo il grande dilemma dell'incompletezza dell'apprendimento elettronico (machine learning). Ed ecco l'epilogo, mai smentito, di quella temperie intellettuale: nessuna macchina avrà mai libero arbitrio; nessuna macchina sarà mai completa o avrà mai il privilegio di funzionare coerentemente ed essere contestualmente autocosciente e proprietaria del valore e della dimostrazione della sua coerenza.

Ovviamente questa premessa non ha impedito la crescita e l'espansione di macchine sempre più veloci e complesse, né tantomeno ha evitato il dilagare dell'automazione nei processi produttivi e organizzativi. L'uomo è stato in grado di creare macchine cui appaltare capacità di archiviazione, lettura, calcolo e simulazione infinitamente più veloci e potenti delle sue, ma tristemente incomplete. Su questa linea di produzione il creare dell'uomo non potrà mai essere accompagnato dalle parole indirizzate ad Adamo ed Eva (Genesi 3, 5-7) dal serpente (bruco della dea ragione: ricordarsi del culto di Hegel per il serpente in Ernst Bloch, *Atheismus in Christentum*, 1968) per convincerli a mangiare il frutto della conoscenza: eritis sicut deus, scientes bonum et malum, («Anzi, Dio sa che quando voi ne mangiaste, si aprirebbero i vostri occhi e diventereste come Dio, conoscendo il bene e il male»). Una conoscenza derivata dal morso della mela proibita. Celebre è l'aneddotica sull'ultima mela avvelenata morsa da Alan Turing, simbolo forse di una scienza impossibile, dell'impossibilità per le macchine di decidere e di mordere.

L'autonomia decisionale delle macchine, estremizzando, implicherebbe il loro potere creaturale. Non è escluso che possano in futuro sorgere interazioni sempre più stringenti tra vita naturale e vita artificiale; sono già molteplici le evidenze nella sfera della manipolazione biologica, chimica e corporea, ma questo livello di progresso non può essere equivocato col primato della sostanza inanimata.

L'assurdo è rappresentato dalla congettura che alla libera volontà creatrice dell'uomo possa un

giorno sostituirsi una libera volontà creatrice della macchina; vale a dire non solo in funzione ancillare o di complicità, ma in esclusiva e indipendente funzione generativa. Una tale congettura non è solo una bizzarria religiosa, è un delirio culturale che fraintende la storia naturale, la storia industriale e quella sociale e che potrebbe svelare la sua intera vuotezza solo al cospetto dei classici.

C'è un passo del primo libro del Capitale (studiato e commentato recentemente da Carlo Ginzburg, *La lettera uccide*, 2021 e Tom Rockmore, *Una riflessione su Vico e il materialismo marxista nel Capitale*, «Materialismo Storico», 1-2/2016) che è un concilio di classici: Vico, Darwin, Marx. Una nota al capitolo Macchine e grande industria dedicato alla nascita della «seconda macchina a vapore di Watt», un «motore primario» che «produce esso stesso la propria forza motrice» e il cui «potere energetico è interamente controllabile dall'uomo». Solo dopo questo evento gli «strumenti» passarono dall'essere «strumenti dell'organismo umano» all'essere «strumenti di un apparecchio meccanico» e la macchina assunse una «forma autonoma» completamente «emancipata dai limiti della forza umana» divenendo «agente generale della grande industria».

A ben considerare la rivoluzione industriale è qui presentata come un capitolo della lotta biologica dell'uomo per liberarsi dai limiti fisici energetici e temporali della sua forza naturale, dalle imperfezioni della sua capacità di calcolare e memorizzare. Ed è qui però che bisogna evidenziare quello che Marx chiama elemento storico, che non va rimosso neanche quando parliamo di machine learning o artificial intelligence. Trascurare l'elemento storico sarebbe come trasferire un bene irripetibile dell'uomo a un'entità dotata di autonomia parziale, che sprofonderebbe così nella storia naturale come un iriocervo nelle sabbie mobili. L'impossibile incubo del naturalismo delle macchine.

Questo assurdo conferimento sarebbe non solo l'acquisto di «forza motrice autonoma» ma di quella forza creatrice e decisionale che non ha bisogno di un innesco dall'esterno (l'impulso originario, il primo motor rappresentato dalla volontà umana) e che può evolvere senza eterodirezione.

Ed ecco le mirabili convergenze e divergenze tra Darwin e Vico, tra natura e storia, che si agitano nella mente di Marx: «Darwin ha richiamato l'interesse sulla storia della tecnologia naturale, cioè sulla formazione degli organi della pianta e dell'animale come strumenti di produzione della loro vita: non merita forse eguale attenzione la storia della formazione degli organi produttivi dell'uomo sociale, che costituiscono la base materiale di qualunque organizzazione della società? E non sarebbe più facile ricostruirla, dal momento che, come dice Vico, la storia umana si distingue dalla storia naturale perché noi non abbiamo fatto la seconda e abbiamo fatto la prima?». Un senso del limite che non andrebbe smarrito neanche a proposito della storia delle rivoluzioni industriali. Gli evoluzionisti hanno le idee chiare e sono molto scettici sulle possibilità di un'intelligenza artificiale che possa «evolvere per proprio conto» che possa trascendere i limiti dell'essere una «simulazione di intelligenza» con parametri preimpostati dagli uomini. Ci sono, è vero, piccole nicchie che sperimentano sui Mobot, i «robot formiche o sciame» la possibilità di un'evoluzione dell'IA verso «comportamenti emergenti» e larvate forme di autorganizzazione, ma non saranno certamente i cyborg a turbare i nostri sogni, prodotto di un'integrazione tra biologia e cultura che abbiamo deciso e condotto noi (L'innovazione tra evoluzione e editing genetico. Intervista a Telmo Pievani, a cura di Marco dal Pozzolo e Otello Palmi, in *Frontiere*, «Pandora Rivista», 1/2021).

Si potrebbe ora rievocare un'altra fumosa e più barocca intelligenza artificiale, da Edgar Allan Poe a Walter Benjamin: «È noto che sarebbe esistito un automa costruito in modo tale da reagire ad ogni mossa di un giocatore di scacchi con una contromossa che gli assicurava la vittoria [...] c'era seduto dentro un nano gobbo, maestro nel gioco degli scacchi, che guidava per mezzo di fili la mano del manichino» (Walter Benjamin, *Sul concetto di storia*, I, 1939-40). Al posto del nano e del gobbo Benjamin, intensificando e comprimendo ad alte densità immateriali l'elemento storico marxiano, vorrà vedervi come spiritus rector l'attesa messianica, la teologia «vecchia e rattrappita» che ama

nascondersi. Non solo fiducia per un avvenire ipertecnologico, ma principio di speranza politica (per la lettura adeguata delle tesi di Benjamin si tenga presente Michael Löwy, Segnalatore d'incendio, 2001).