

Apprendimento automatico, salute mentale ed eugenetica

Dan Mcquillan

Introduzione

L'apprendimento automatico è destinato a essere irreversibilmente coinvolto nella diagnosi della salute mentale. Il cosiddetto fenotipo digitale è vertiginosamente ampio, tanto che i cambiamenti nell'uso degli smartphone, i dati dei fitness tracker o il tono delle parole che usiamo su Twitter diventeranno fattori di diagnosi predittiva. Ma si tratta di una collisione tra l'esattezza computazionale e le incoerenze nascoste dall'etichettatura psichiatrica. Mentre l'idea della 'psichiatria di precisione' promette di individuare l'insorgenza precoce della psicosi prima che i servizi o l'individuo stesso siano in grado di farlo, i movimenti degli utenti sostengono che molto di ciò che viene medicalizzato è l'espressione di un trauma o di un disagio. I volumi di dati necessari per l'IA richiederanno una sorveglianza pervasiva che amplificherà l'ansia già instillata da un sistema di prestazioni, con una natura diffidente e punitiva. L'IA non è un attore neutrale, ma peserà pesantemente sul lato del riduzionismo biologico, rafforzando la comprensione dei problemi di salute mentale come disturbi dei circuiti cerebrali piuttosto che come conseguenza di eventi avversi della vita. Ciò è parallelo al campo emergente della sociogenomica e all'uso di studi di associazione genomica (GWAS) per correlare lo status sociale a fattori genetici distribuiti, riaprendo narrazioni eugenetiche che erano state considerate come sepolte nella storia. Sia la psichiatria di precisione che i GWAS classificano i problemi di salute mentale come tendenze innate e agiscono come cortine fumogene per oscurare le condizioni sociali e politiche. Per superare la coscienza di spettatore dell'IA, abbiamo bisogno di pratiche tecniche critiche che possano svincolare le distanze vettoriali dalle differenze sociali. Ciò può avvenire attra-

verso un'IA femminista che attinga alla teoria degli standpoint e agli approcci femministi alla scienza, combinati con strutture collettive di ricerca che includano coloro che sono più coinvolti nel processo di indagine. Abbiamo bisogno di una psicopolitica alternativa dell'apprendimento automatico.

Fenotipo digitale

La generazione di dati da parte dei nostri numerosi dispositivi e l'analisi di tali dati da parte dell'apprendimento automatico è destinata a diventare così profondamente coinvolta nella diagnosi delle malattie, da offuscare i confini tra il digitale e il biologico e cambiare ciò che è considerato conoscibile sui nostri corpi e cervelli.

Un'idea chiave di questi sviluppi è nota come fenotipizzazione digitale¹. Essa descrive i nostri tratti osservabili, come l'aspetto fisico, la biochimica e il comportamento. A differenza del genotipo, l'insieme dei geni che codificano le caratteristiche, il fenotipo è il prodotto dell'espressione genica e dei fattori ambientali. Sappiamo che il colore degli occhi è determinato geneticamente e che l'alimentazione infantile influisce sulla statura da adulti. Ma ora si propone che i tempi di reazione alla tastiera, l'attenzione allo schermo e le caratteristiche catturate digitalmente della nostra voce costituiscano i nostri fenotipi digitali². In altre parole, si tratta di corollari sufficientemente solidi alle osservazioni tradizionali rispetto ai mutamenti delle malattie in modo da poter essere sia diagnostici sia prognostici, identificando i sintomi prima che siano altrimenti osservabili e fornendo potenziali percorsi di intervento precoce.

La portata del fenotipo digitale è vertiginosamente ampia e, in alcuni casi, piuttosto superficiale. L'attività dello smartphone durante la notte è vista come un tracker per l'insonnia³, e i dati dei monitor

¹ S.H. Jain, B.W. Powers, J.B. Hawkins, J.S. Brownstein, *The Digital Phenotype. Comments and Opinion*, «Nature Biotechnology», 5/2015, pp. 462-463.

² T.R. Insel, *The NIMH Research Domain Criteria (RDoC) Project: Precision Medicine for Psychiatry*, «American Journal of Psychiatry», 4/2014, pp. 395-397.

³ D.J. McIver, J.B. Hawkins, R. Chunara, A.K. Chatterjee, A. Bhandari, T.P. Fitzgerald, S.H. Jain, J.S. Brownstein, *Characterizing Sleep Issues Using Twitter*, «Journal of Medical Internet Research», 6/2015, pp. 227-233.

per il fitness sono usati per caratterizzare le malattie cardiometaboliche⁴. Ma la fenotipizzazione digitale ha implicazioni più profonde sia dal punto di vista medico sia politico. Queste stanno diventando visibili nella sua proposta di applicazione alla salute mentale, dove si fornirà «una valutazione longitudinale passiva, oggettiva, continua e ubiquitaria dell'umore e della cognizione (e) firme per la previsione e la prevenzione»⁵. La fenotipizzazione digitale è proposta come modalità di accesso diretto ai sintomi della salute mentale; i comportamenti bipolari catturati dai social media, i segnali sensoriali di entropia correlati alle valutazioni dell'umore, la depressione e la psicosi individuate dall'incoerenza semantica nei campioni di parlato.

Non si tratta solo di una sostituzione con il monitoraggio digitale, ma della capacità del fenotipo digitale di rivelare nuovi segnali clinicamente rilevanti e di modificare la percezione della manifestazione di una patologia. L'idea è che il potere predittivo dei dati e l'apprendimento automatico raggiungeranno il graal della medicina preventiva, oltre a diventare un canale diretto per il cambiamento dei comportamenti, evitando la fatica della formazione alla salute pubblica. Il termine biomarcatore si riferiva ai risultati di una biopsia o di un esame del sangue, ma ora, secondo uno dei maggiori ricercatori statunitensi del settore «the biomarker is some combination of GPS and, from the watch, seeing your heart rate spike», così che un qualsiasi direttore di un Centro per la salute mentale digitale di una qualsiasi università può dire, riferendosi ai dati dei sensori degli smartphone come l'accelerometro e il microfono, «this is going to be our MRI scan in behavior»⁶. La fenotipizzazione digitale non consiste solo nel disporre di un'enorme serie di dati, ma nell'utilizzare un software per trasformarli in qualcosa di riconoscibile, il che richiederà spesso l'apprendimento automatico anche a causa della complessità dell'adattamento dei

⁴ M.A. Kirk, M. Amiri, M. Pirbaglou, P. Ritvo, *Wearable Technology and Physical Activity Behavior Change in Adults With Chronic Cardiometabolic Disease: A Systematic Review and Meta-Analysis*, «American Journal of Health Promotion», 5/2019, pp. 778-791.

⁵ P. Dagum, *Digital Phenotyping in Mental Health*, interview presented at the National Alliance in Mental Health Convention in December 2017.

⁶ P. Biegler, *Tech Support: How Our Phones Could Save Our Lives by Detecting Mood Shifts*, «The Sydney Morning Herald», 10 novembre 2017.

modelli. Un fatto spiacevole riguardo alla diagnosi della salute mentale è l'incoerenza delle diagnosi tra i singoli psichiatri. Le ricerche rivelano che l'affidabilità delle valutazioni psichiatriche è statisticamente debole⁷; cioè, la diagnosi varia così tanto tra gli psichiatri da risultare empiricamente inaffidabile⁸. Gli studi dimostrano anche che i criteri per il disturbo depressivo sono scarsamente correlati e incoerenti tra loro, mentre altri ricercatori mettono in dubbio proprio le basi concettuali per la diagnosi di schizofrenia⁹.

Le categorizzazioni apparentemente empiriche della psichiatria sono impantanate in varie controversie e sono fortemente contestate, non da ultimo da molti utenti dei servizi di salute mentale. Per capire perché l'apprendimento automatico si schiera così fortemente in questo dibattito, dobbiamo esaminare il modo in cui esso immagazzina i dati. L'apprendimento automatico impara adattando interattivamente i dati a un insieme di risultati etichettati, o target. Le operazioni matematiche ottimizzano una funzione di perdita a cui si somma il divario statistico tra le previsioni e gli obiettivi, in modo che l'algoritmo sia in grado di classificare i nuovi dati in arrivo con un errore minimo. Per la salute mentale, i dati sono casi noti e gli obiettivi dell'addestramento sono le diagnosi, definite dal Manuale diagnostico e statistico dei disturbi mentali, quinta edizione (DSM-5) (American Psychiatric Association (APA) 2018) o dall'equivalente dell'Organizzazione mondiale della sanità, la Classificazione internazionale delle malattie versione 11 (ICD-11) (Organizzazione mondiale della sanità 2018). Alcune forme di apprendimento automatico sono concettualmente semplici, come il K-nearest neighbors, che classifica un punto nello spazio delle caratteristiche in base ai punti di formazione a cui è più vicino. Altre, come la retropropagazione utilizzata nelle reti neurali, trattano dati molto più complessi utilizzando una serie di calcoli paralleli. In ogni caso, la matematica che permette di adattare i punti ai risultati tar-

⁷ R. Freedman, D.A. Lewis, R. Michels, D.S. Pine, S.K. Schultz, C.A. Tamminga, G.O. Gabbard, *The Initial Field Trials of DSM-5: New Blooms and Old Thorns*, «American Journal of Psychiatry», 1/2013, pp. 1-5.

⁸ J. Carney, *Unstable Reliability Ratings - Boycott DSM5*, «Mad In America blog», 26 marzo 2013.

⁹ M.J. Owen, *Is There a Schizophrenia to Diagnose?*, «World Psychiatry», 1/2011, pp. 34-35.

get è ben definita. Tuttavia, le conseguenze dell'applicazione dell'apprendimento automatico alla salute mentale sono sismiche, a causa della collisione tra questa precisione matematica e le complessità nascoste dalle etichette diagnostiche. Alcune di queste complessità possono essere viste nel modo in cui l'aggiornamento dal DSM-4 al DSM-5 è stato criticato da tutte le professioni della salute mentale perché ha medicalizzato condizioni che in precedenza erano considerate naturali, come la depressione che può accompagnare un lutto, gli scatti d'ira nei bambini (Disruptive Mood Dysregulation Disorder) o le lievi dimenticanze in età avanzata (Minor Neurocognitive Disorder) attirando così l'attenzione sull'influenza dietro le quinte delle case farmaceutiche e portando un importante psichiatra ad avvertire che una «medicalizzazione della normalità sarebbe stata un bene per l'industria farmaceutica»¹⁰. Sottolineare che le diagnosi di salute mentale possono mancare di solide basi empiriche o essere soggette al pensiero di gruppo dei comitati professionali non significa affermare che la diagnosi non sia mai utile per guidare il trattamento o per convalidare le esperienze di una persona. Ma il potere dell'etichettatura psichiatrica può alterare radicalmente sia l'identità di sé, sia il modo in cui una persona viene trattata dai servizi e dalla società in generale, mentre i movimenti degli utenti sostengono che ciò che viene medicalizzato è spesso l'espressione di un disagio profondo o degli effetti di un trauma precoce come l'abuso¹¹.

In queste condizioni è fondamentale comprendere gli effetti della classificazione proposta dalla 'macchina'; come per esempio che le operazioni computazionali che trattano le diagnosi come dati di fatto non automatizzino semplicemente un processo di classificazione, ma costituiscano un intervento distorsivo che pone l'implacabile matematica ai ferri corti con la testimonianza dell'esperienza vissuta e la sua irriducibile pluralità¹². Per i sostenitori dell'analisi predittiva, l'instabilità dell'etichettatura psichiatrica non è un problema, ma un'oppor-

¹⁰ A. Frances, *A Warning Sign on the Road to DSM-V: Beware of Its Unintended Consequences*, «Psychiatric Times», 27 June 2009.

¹¹ S. Timimi, *No More Psychiatric Labels*, «Asylum Magazine», 24 May 2012.

¹² D. Mcquillan, *Mental Health and Artificial Intelligence: Losing Your Voice*, «Open Democracy», 12 November 2018.

tunità: la missione dell'apprendimento automatico è quella di fornire una psichiatria di precisione, spazzando via le dolorose contraddizioni che circondano il disagio mentale. Frustrati dal fatto che non esiste una misura di laboratorio oggettiva per i problemi di salute mentale, i sostenitori dell'analisi predittiva cercano una combinazione di fattori biologici, comportamentali e sociali, un fenotipo digitale che catturi definitivamente i diversi tipi di disfunzioni mentali, portando una precisione non solo nella classificazione ma anche una risoluzione temporale superiore all'attuale pratica clinica, identificando con una precisione simile a quella di un laser i primi segni di psicosi o di altre condizioni ben prima che vengano notati dall'individuo stesso o rilevati dai servizi sanitari.

Sfortunatamente, il sogno di precisione sembra sempre portare danni collaterali nella fase pratica, allo stesso modo in cui le armi intelligenti utilizzate negli attacchi aerei causano ancora un numero impressionante di vittime civili. Un tipo di danno collaterale che deriva dall'applicazione sociale dell'IA è la sorveglianza, la necessità di enormi quantità di dati di addestramento per ottenere una qualche precisione dagli algoritmi. Nel Regno Unito, le persone con problemi di salute mentale sperimentano già un sistema di sussidi che dà priorità al ritorno al lavoro produttivo rispetto a qualsiasi nozione di rete di sicurezza sociale. Nell'ambito di questa sfiducia istituzionalizzata, le riprese delle telecamere a circuito chiuso del supermercato Sainsbury's vengono statisticamente combinate con quelle degli abbonamenti in palestra, dei filmati degli aeroporti e dei post sui social media per suggerire che le persone mentono sulle proprie disabilità, in una politica deliberata di sorveglianza come deterrente volta a instillare ansia e interiorizzare il sospetto morale (Big Brother Watch 2018). L'ironia, naturalmente, è che il senso di iper-visibilità, di essere presi di mira, di non poter uscire per paura di essere visti come se stessero facendo qualcosa di normale, dove la retorica dello scroccone significa che essere visti giocare con i propri figli potrebbe essere usato contro di noi, può portare ulteriori livelli invalidanti di paura, ansia e paranoia in aggiunta ad altri problemi di salute mentale esistenti. E ancora, i trattamenti mirati forniti dall'IA richiederanno un monitoraggio ancora più pervasivo. Enormi volumi di dati sono indispensabili per qualsiasi apprendimento automatico, per quanto benefica sia la sua missione, per

non parlare dell'applicazione dell'IA sotto l'austerità, dove è vista come la quadratura del cerchio tra la crescente domanda e le risorse politicamente ridotte.

Riduzionismo dell'IA

Oltre a richiedere una sorveglianza sempre più pervasiva, l'applicazione dell'IA alle diagnosi psichiatriche avrà un peso notevole a favore del riduzionismo biologico.

Nella psichiatria e nella psicologia clinica i tradizionali disturbi mentali sono comunemente definiti come disturbi cerebrali. Per quanto complessi siano i fenomeni osservati e per quanto sembrano invocare un senso, la sensazione è che possano essere ricondotti a disfunzioni della materia grigia. Tuttavia, questa convinzione che i problemi di salute mentale siano disturbi cerebrali non è stata accompagnata dalla capacità delle neuroscienze di individuare i problemi. Nonostante decenni di sforzi, non esistono marcatori biologici definitivi per le condizioni psichiatriche comuni¹³. È difficile trovare definizioni dettagliate di come i disturbi mentali siano esattamente disturbi cerebrali e, sebbene gli studi che utilizzano le neuroimmagini come la risonanza magnetica funzionale siano pieni di correlazioni tra l'attività cerebrale e le diagnosi del DSM, non rivelano se si tratti di cause, effetti o epifenomeni. Almeno alcune delle meraviglie di rendere visibile il funzionamento del cervello tramite la risonanza magnetica sono illusorie; una meta-analisi condotta nel 2017 su centinaia di studi di risonanza magnetica che hanno coinvolto migliaia di partecipanti non ha mostrato differenze cerebrali consistenti tra schizofrenia, disturbo bipolare, disturbo depressivo maggiore, disturbi d'ansia e disturbo ossessivo compulsivo. Di conseguenza, mette in guardia dall'attribuire un'eccessiva specificità ai cambiamenti funzionali del cervello quando si formano modelli esplicativi dei disturbi psichiatrici¹⁴. Sebbene al-

¹³ E. Fried, *All Mental Disorders Are Brain Disorders ... Not*, «Eiko Fried blog», 6 January 2018.

¹⁴ E. Sprooten, A. Rasgon, M. Goodman, A. Carlin, E. Leibu, W. Hee Lee, S. Frangou, *Addressing Reverse Inference in Psychiatric Neuroimaging: Meta-Analyses*

cuni di questi problemi possano derivare da incongruenze diagnostiche, la fMRI ha anche i suoi problemi di affidabilità dei biomarcatori a livello individuale¹⁵ e fa parte del più ampio fenomeno di 'crisi della riproducibilità' nella scienza a causa di problemi di p-hacking (in cui i campioni vengono selezionati per produrre risultati statistici apparentemente significativi)¹⁶. La contro-narrazione suggerisce che i problemi di salute mentale non hanno origine all'interno del cervello, ma sono la risposta della mente a traumi, abusi o angosce¹⁷. Le ricerche dimostrano, ad esempio, che le avversità infantili aumentano il rischio di psicosi¹⁸. Per quanto riguarda la depressione e il PTSD, ad esempio, i risultati sono che gli eventi avversi della vita spiegano circa due ordini di grandezza in più della varianza rispetto ai fattori biologici.

In questa ambiguità di fattori biologici e sociali si inserisce la cosiddetta 'psichiatria di precisione', che promuove concetti di disturbi mentali non come lesioni, ma come disturbi dei circuiti cerebrali che possono essere identificati attraverso gli strumenti della genetica, delle neuroscienze e dei dati digitali¹⁹. La psichiatria intelligente analizza i dati digitali di tutti i giorni che presumibilmente forniscono indizi sul comportamento, sulla cognizione e sull'umore, che si tratti di cambiamenti nella velocità di battitura, nell'attività fisica, nel tono di voce o nella forma delle parole che usiamo su Twitter, utilizzando le possibilità dello smartphone, mediate dalla lente gravitazionale dell'intel-

of Task-Related Brain Activation in Common Mental Disorders, «Human Brain Mapping», 4/2017, pp. 1846-1864.

¹⁵ J.H. Fröhner, V. Teckentrup, M.N. Smolka, N.B. Kroemer, *Addressing the Reliability Fallacy: Similar Group Effects May Arise from Unreliable Individual Effects*, «Neuroimage», 195/2017, pp. 174-189.

¹⁶ C. Pernet, J.B. Poline, *Improving Functional Magnetic Resonance Imaging Reproducibility*, «GigaScience», 4/2015, pp. 1-8.

¹⁷ P. Kinderman, *Why We Need to Abandon the Disease-Model of Mental Health Care*, «Scientific American Blog Network», 17 November 2014.

¹⁸ F. Varese, F. Smeets, M. Drukker, R. Lievers, T. Lataster, W. Viechtbauer, J. Read, J. Van Os, R.P. Bentall, *Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective- and cross-sectional cohort studies*, «Schizophrenia Bull», 4/2012, pp. 661-671.

¹⁹ P. Dagum, *Digital Phenotyping in Mental Health*, intervista presentata alla National Alliance in Mental Health Convention 2017, Washington DC, 30 giugno 2017.

ligenza artificiale, per rivelare le perturbazioni più profonde del nostro cervello.

La capacità dell'apprendimento automatico di lavorare con dati eterogenei, che spaziano dai grafi sociali ai sensori di inclinazione, non viene utilizzata per ampliare la nostra prospettiva, ma per separare le cause sociali dal disagio e, come un GPS neurale, localizzarne le origini nelle nostre strutture corticali.

L'effetto di questo intervento digitale nella questione mente-corpo non è semplicemente quello di sostenere un'epistemologia riduttiva, ma di partecipare al divenire dell'umano, i cui confini sono sempre co-costruiti dagli strumenti del tempo. La posta in gioco non è solo la colonizzazione algoritmica della mente, ma il modo in cui l'epoca dell'IA disegna i confini del naturale.

Le domande su ciò che costituisce il corpo e la mente umana, sulla loro composizione e sui loro limiti, sono sempre attenuate dalle pratiche tecniche a nostra disposizione. I geni, come oggetti contemporanei di conoscenza di sé, dipendono per la loro concretezza dai meccanismi della tecnoscienza distribuita: sequenziatori, biobanche, banche dati, algoritmi, istituzioni e internet. Allo stesso modo, i fenotipi digitali del disagio mentale verranno ad esistere attraverso l'impilamento verticale della diagnosi psichiatrica con gli smartphone e gli algoritmi delle reti neurali artificiali, e il substrato dei microprocessori GPU in grado di eseguire calcoli massivamente paralleli in tempi fattibili²⁰. Invece delle fantasie dei transumanisti della Silicon Valley, secondo cui le nostre menti saranno caricate sul cloud e quindi sfuggiranno ai vincoli della fisicità, i circuiti dell'analisi dei dati si stanno invece fondendo con i nostri cervelli incarnati.

Il complesso computazionale-psichiatrico costruisce spiegazioni autorevoli dei nostri comportamenti incatenando ipotesi riduttive sulla mente alle correlazioni dell'apprendimento automatico. Questo tentativo di costruire un complesso computazionale-psichiatrico non avviene nel vuoto, ma è parallelo al campo emergente della sociogenomica che, cercando di salvare il potere esplicativo della genetica sposandolo con i big data, ha fatto riemergere alcuni pregiudizi

²⁰ F. Shaikh, *Why Are GPUs Necessary for Training Deep Learning Models?*, «Analytics Vidhya blog», 18 May 2017.

suprematisti della scienza che si pensava fossero tranquillamente sepolti nella storia.

Nonostante le promesse che hanno accompagnato la corsa al sequenziamento del genoma umano, il fatto scomodo è che poche malattie comuni possono essere attribuite a singoli geni. Per questo motivo, sono in corso studi sull'intero genoma di grandi gruppi di persone alla ricerca di modelli di variazioni in un milione di polimorfismi a singolo nucleotide, o SNP, in un processo noto come studi di associazione genome-wide (GWAS)²¹. L'idea è quella di trovare correlazioni tra questi modelli e tratti fenotipici osservabili. Se si riesce a trovare un modello statistico, il genoma di qualsiasi individuo successivo può essere confrontato con esso per generare una gradazione di rischio poligenico²²; la probabilità predittiva che quella persona abbia o sviluppi il tratto. Utilizzando dati provenienti da archivi pubblici e da aziende di genomica di consumo come 23andMe, gli studi GWAS hanno trovato modelli di SNP associati a una serie vertiginosa di tratti. L'elenco della Biobanca del Regno Unito, per esempio, comprende risultati relativi al cancro e ad altre malattie, ma anche correlazioni genetiche legate ad altri aspetti della vita sociale come i turni di lavoro in azienda, ma anche il tempo trascorso alla guida ecc., e naturalmente modelli apparentemente associati a depressione, ansia e schizofrenia (The Neale Lab 2018).

Come l'apprendimento automatico, la sociogenomica si basa sulla correlazione piuttosto che sulla causalità. Sebbene la presenza di queste correlazioni suggerisca che questi tratti possano essere ereditabili, i professionisti più sobri sottolineano che anche se qualcosa è in qualche modo ereditabile, le differenze ambientali e culturali sono spesso i principali fattori di differenza²³. Anche se una parte della differenza fenotipica può essere attribuita a punteggi poligenici, non significa che un tratto sia immutabile o naturale, poiché molti fenotipi sono modificabili. Inoltre, la cosa più critica è che le variazioni identifica-

²¹ J.S. Witte, *Genome-Wide Association Studies and Beyond*, «Annual Review of Public Health», 31/2010, pp. 9-20.

²² F. Dudbridge, *Power and Predictive Accuracy of Polygenic Risk Scores*, «PLOS Genetics», 3/2013.

²³ G. Coop, *Polygenic Scores and Tea Drinking*, «Gcbias blog», 14 March 2018.

te sono spesso piccole, una questione di pochi punti percentuali, così che i fattori ambientali e culturali le sovrastano in ordine di importanza. Eppure alcuni scienziati sono propensi a citare i GWAS come basi genetiche di fenomeni sociali complessi, promuovendo l'idea di pagelle che predicono i rischi non solo di varie malattie, ma anche le propensioni a comportamenti futuri come la fedeltà coniugale o la prudenza finanziaria²⁴.

Alcuni studiosi come Robert Plomin, insistono sul fatto che «individual differences in income are, like everything else, substantially heritable, about 40 percent. Income correlates with intelligence, and genetics drives this correlation»²⁵. La giustificazione del diritto, che è vecchia come il cucco, viene trasformata dalla sua nozione di 'ricchezza genetica' in una sorta di neoliberalismo genomico, in cui gli SNP sono gli elementi di un meccanismo di libero mercato che produrrà necessariamente il risultato ottimale. Nel frattempo, il cosiddetto movimento per la biodiversità umana, che fa parte della più ampia famiglia della destra politica, sfrutta gli studi GWAS per affermare una base scientifica per le disuguaglianze razziali e utilizza la nozione di 'realismo razziale' per giustificare il razzismo come ragionevole ed empirico. Ricordiamo che tali visioni distorte possono diventare parte dell'apparato sociale. All'inizio del XX secolo, le idee eugenetiche portarono ad alcune leggi negli Stati Uniti contro la *miscegenation* (la mescolanza di gruppi razziali diversi attraverso il matrimonio o le relazioni sessuali) e furono alla base della politica di sterilizzazione dei 'non idonei'.

Francis Galton, il cugino di Darwin che coniò il termine eugenetica, credeva che il progresso della civiltà dipendesse dall'eliminazione dei deboli. Nell'ambito di una ricerca su una popolazione sedentaria, sviluppò il concetto di correlazione matematica, che divenne un pilastro centrale della statistica. Il suo seguace eugenista Karl Pearson creò il concetto di coefficiente di correlazione²⁶. Naturalmente, il con-

²⁴ N. Comfort, *Nature still battles nurture in the haunting world of social genomics*, «Nature», 553/2018, pp. 278-280.

²⁵ R. Plomin, *Blueprint: Come il DNA ci rende ciò che siamo*, Allen Lane ed., Londra 2018.

²⁶ N. Comfort, *Sociogenomics Is Opening a New Door to Eugenics*, «MIT Technology Review», 23 October 2018.

retto di correlazione e il coefficiente di correlazione hanno trascorso la loro origine per diventare parte dell'analisi numerica generale, con la stessa probabilità di essere applicati a favore della giustizia sociale o contro di essa. Ma in una congiunzione di suprema ironia storica, questi metodi statistici sono il nucleo matematico dell'apprendimento automatico e rischiano di essere iscritti in forme distorte di selezione sociale, in modo tale che la matematica nata accanto all'eugenetica si ricongiunga ad essa attraverso l'IA.

La scienza da sola non è il baluardo che potremmo sperare contro questi sviluppi. Come sottolineano Sandra Harding e altri, il suo modello di oggettività è efficace nel controllo incrociato tra gli esperimenti, ma è miope nell'identificare i pregiudizi della cultura²⁷. Tuttavia, la scienza rivendica per sé il potere di squalificare altre modalità di spiegazione e, inserita in asimmetrie istituzionali di potere, i giudizi algoritmici ereditano questa autorità.

Le decisioni istituzionali si baseranno su metriche predittive come quelle del sistema di protezione dei minori della contea di Allegheny²⁸ o delle classifiche di valutazione degli insegnanti di Washington D.C.²⁹. Sia che si basino sulla sociogenomica o sui fenotipi digitali, le analisi predittive acquisteranno forza di legge anche quando non saranno esplicitamente previste, soprattutto quando faranno parte di cicli chiusi come il ciclo di previsione-intervento-monitoraggio della psichiatria di precisione. Assumere che le persone non siano dei fogli bianchi non equivale ad affermare un ordine naturale, né l'esistenza di differenze tra i cervelli equivale a dire che queste sono il fattore più importante per la salute mentale, ma sedimentare queste domande sotto l'opacità algoritmica porterà alla loro promozione attraverso stati di eccezione³⁰. I presupposti riduzionistici incorporati nei sistemi di apprendimento automatico, che a loro volta sono incor-

²⁷ S. Harding, *Rethinking Standpoint Epistemology: What Is "Strong Objectivity"?*, Routledge, Londra 2004.

²⁸ V. Eubanks, *A Child Abuse Prediction Model Fails Poor Families*, «Wired», 15 gennaio 2018.

²⁹ W. Rinehart, *Book Review: Cathy O'Neil's 'Weapons of Math Destruction'*, «Technology Liberation Front», 7 November 2018.

³⁰ D. McQuillan, *Algorithmic States of Exception*, «European Journal of Cultural Studies», 4-5/2015, pp. 564-576.

porati nelle urgenze dei servizi di prima linea, renderanno operativi la prelazione, l'esclusione e la detenzione.

Psicopolitica dell'apprendimento automatico

Il target delle persone selezionate attraverso biomarcatori digitali, sostenuto da ipotesi 'biologizzanti' sui disturbi cerebrali, rafforza una visione riduzionista secondo cui i problemi di salute mentale sono una propensione innata, una maledizione individuale. La cosiddetta 'psichiatria dell'IA' diventa una cortina di fumo high-tech che oscura i fattori sociali, come i traumi emotivi e psicologici e gli effetti devastanti della precarietà, della povertà e della mancanza di una casa. Le scosse di assestamento dell'austerità spingono un numero sempre maggiore di persone in condizioni disperate, aggravate da un sistema di sussidi punitivo che incolpa anche l'individuo, producendo indebitamento e disperazione. Il tributo alla salute mentale di questo sistema è visibile nei suicidi di persone che si sono viste negare i sussidi di invalidità³¹. Nel suo rapporto sulla Gran Bretagna, il relatore delle Nazioni Unite sulla povertà estrema e i diritti umani ha concluso che l'austerità non è la causa diretta, ma la copertura di una riduzione rivoluzionaria dei livelli di equità e giustizia sociale, che equivale a uno smantellamento del contratto sociale:

Compassion for those who are suffering has been replaced by a punitive, mean-spirited, and often callous approach apparently designed to instill discipline where it is least useful, to impose a rigid order on the lives of those least capable of coping with today's world, and elevating the goal of enforcing blind compliance over a genuine concern to improve the well-being of those at the lowest levels of British society³².

³¹ B. Barr, D. Taylor-Robinson, D. Stuckler, R. Loopstra, A. Reeves, M. Whitehead, *'First, Do No Harm': Are Disability Assessments Associated with Adverse Trends in Mental Health? A Longitudinal Ecological Study*, «Journal of Epidemiol Community Health», 4/2016, pp. 339-345.

³² P. Alston, *Statement on Visit to the United Kingdom*, United Nations Special Rapporteur on Extreme Poverty and Human Rights, 16 November 2018.

Egli individua nell'automazione e nell'IA il fulcro di questa intensificazione della strategia governativa, attraverso l'analisi completamente automatizzata dei rischi e il ricalcolo in tempo reale di sussidi e sanzioni; una mobilitazione algoritmica del mantra di Margaret Thatcher secondo cui 'la società non esiste'³³. Eppure sappiamo già che, quando si parla di benessere, i determinanti sociali superano quelli biologici. La ricerca suggerisce che l'impatto maggiore sui risultati della salute mentale deriva da fattori esterni al trattamento, come le circostanze sociali e i livelli di supporto di una persona. Invece di permettere all'IA di definire come unica opzione gli interventi miranti a obiettivi individualizzati, dovremmo chiederci se possiamo ripensare i nostri metodi algoritmici in modo che riducano i fattori di rischio sociale e creino beni pubblici che aumentino il benessere delle comunità, perché la prevenzione piuttosto che la prelazione implica la crescita dell'intera popolazione.

La tecnologia, come la scienza, è allo stesso tempo il prodotto della matrice sociale e ne è costitutiva. Non solo agisce nel mondo, ma riflette una visione collettiva di ciò che è possibile e desiderabile, che Sheila Jasanoff ha definito immaginario sociotecnico³⁴. Le speranze e le paure di una cultura sono legate ai suoi mezzi di conoscenza, un processo in costante evoluzione, poiché tecnologie come la genetica e la computazione ereditano una storia, ma indicano modi diversi per raggiungere futuri promettenti. L'immaginario sociotecnico lega insieme il nostro senso di ciò che è e di ciò che dovrebbe essere, co-producendo significati sulla natura del mondo e sul nostro posto nella società che durano perché sono eseguiti collettivamente, ma che mutano quando emergono nuove possibilità.

Contestando la visione psichiatrica della salute mentale, il movimento degli utenti e dei sopravvissuti ha già mostrato la possibilità di un'articolazione dal basso verso l'alto del sapere e di una diversa visione di ciò che è desiderabile, mobilitandosi intorno ai poli del trauma-

³³ Si veda M. Thatcher, *Aids, Education and the Year 2000*, interwove for Woman's Own, 23 September 1987.

³⁴ Sull'argomento, cfr. S. Jasanoff, K. Sang-Hyun, *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*, University of Chicago Press, Chicago-London 2015.

ma, dell'abuso e del disagio piuttosto che della malattia e della patologia, e sviluppando strategie di self-help come la “minimizzazione del danno”³⁵. La necessità di una psicopolitica alternativa della salute mentale è intensificata dalle nuove forme di conoscenza neurale e dal nascente divenire di un complesso psichiatrico-computazionale. C'è un bisogno sociale e politico di pratiche alternative, basate sui valori e sui principi della solidarietà e dell'aiuto reciproco, che spingano una controcultura dell'IA attraverso tutti i suoi strati di infrastrutture, dati, algoritmi e statistiche. Non si tratta di una questione urgente solo per la salute mentale, poiché lo schema del riduzionismo sostenuto dall'IA si ripeterà in tutta la sanità, l'assistenza sociale e il sistema giudiziario. Il vero pregiudizio dell'apprendimento automatico è il modo in cui distorcerà l'immaginario sociotecnico attraverso la coscienza degli spettatori abilitata dalle GPU. Piuttosto che affidarci alla rappresentazione di esperienze incommensurabili tramite vettori computazionali, abbiamo bisogno di nuove pratiche algoritmiche che mettano in discussione le risonanze generate tra le distanze vettoriali e le differenze sociali, e che svincolino l'ottimizzazione della rete neurale dall'ottimizzazione neoliberale dei mercati. Queste pratiche tecniche critiche possono essere realizzate solo con l'inclusione dei diretti interessati.

Ripensare la psicopolitica dell'apprendimento automatico significa lavorare contemporaneamente con la programmazione e la politica, risolvendo problemi tecnici e mantenendo l'attenzione sull'impatto sociale, un processo che non è né ingegneria né politica sociale. Richiede sia precisione a livello matematico sia apertura verso le diverse realtà possibili. Un approccio è offerto dal modello di scienza femminista elaborato da Roy, Spanier e Harding³⁶, per esempio, che amplia la metodologia scientifica come modalità di indagine per includere (1) la localizzazione delle origini della problematica, (2) la scoperta degli scopi dell'indagine e (3) l'instaurazione di una relazione tra il ricercatore e il suo oggetto di indagine. Coloro che de-

³⁵ M. Cresswell, H. Spandler, *Psicopolitica: Peter Sedgwick's Legacy for the Politics of Mental Health*, «Social Theory & Health», 2/2009, pp. 129-147.

³⁶ D. Roy, *Feminist Theory in Science: Working Toward a Practical Transformation*, «Hypatia», 1/2004, pp. 255-279.

siderano sviluppare un apprendimento automatico non oppressivo non devono accettare un problema come dato di fatto, ma devono iniziare a localizzarne le origini; in altre parole, devono diventare consapevoli delle forze strutturali che lo hanno reso prioritario e di ciò che la sua formulazione rivela sull'algebra sociale del potere. Scoprire gli scopi dell'apprendimento automatico significa andare oltre la previsione accurata dei dati di validazione attraverso l'ottimizzazione degli iper-parametri. Significa considerare questo ristretto scopo tecnico come parte di un insieme più ampio di impatti, chiedendosi a chi servirà, chi potrebbe escludere e come influirebbe sul benessere più ampio della società. Forse l'aspetto più radicale per l'apprendimento automatico è che questa metodologia femminista stabilisce una relazione tra chi indaga e il suo oggetto d'indagine, richiedendo di mettere volutamente da parte la coscienza di spettatore che alimenta l'arroganza dell'IA. Il modo più diretto per mettere in pratica questo metodo femminista con l'apprendimento automatico è attraverso strutture collettive di ricerca che includano il 'target group' nel processo di indagine. Il nuovo apprendimento automatico deve combinarsi con una pedagogia critica collegata a sua volta a pratiche tecniche critiche.

Conclusioni

Le questioni sollevate dall'IA nel periodo dell'austerità non sono fatte di filosofie astratte su come vivere con le macchine pensanti, ma si concentrano sui modi in cui le operazioni concrete saranno iscritte dalla politica punitiva e sulla necessità di immaginare alternative che risuonino con la liberazione. Il complesso computazionale-psichiatrico sarà uno dei tanti che nasconde accomodamenti repressi a ogni livello, dalle categorie sociali all'analisi dei dati all'elaborazione dei segnali, piegando le dimensioni del politico fuori dalla vista. Sappiamo già che il modernismo dualistico del nostro immaginario sociotecnico ha un cuore oscuro; che alcune iterazioni del suo dispiegamento tecnico-sociale sostengono le malefatte su scala, sia attraverso la craniometria del colonialismo sia attraverso le macchine Hollerith del nazional-

socialismo³⁷. È particolarmente pericoloso “biologizzare” l’angoscia in un periodo di ascesa della politica di estrema destra; una politica che esprime ciò che Roger Griffin chiama ultranazionalismo palingenetico, cercando di porre fine alla degenerazione della nazione e di provocare la sua imminente rinascita dalla decadenza attraverso la xenofobia razionalizzata e l’etnocentrismo biologicamente determinista³⁸.

Ma anche quando avremo rifiutato la rivalutazione dell’eugenetica e ucciso di nuovo il fascismo, dovremo continuare a contestare i confini del politico e del naturale. Come ci ha ricordato Donna Haraway, «the point is to learn to remember that we might have been otherwise and might yet be, as a matter of embodied fact»³⁹. Forme calcolatrici, come l’IA e la sociogenomica, si stanno sedimentando nelle nostre società; dato che l’apprendimento automatico incanala questi cambiamenti, abbiamo urgentemente bisogno di un rinnovamento strutturale dell’apprendimento automatico stesso attraverso un’agenzia collettiva di empatia e solidarietà, l’applicazione di metodi femministi e un impegno al mutuo soccorso antifascista.

³⁷ E. Black, *IBM e l’Olocausto: The Strategic Alliance Between Nazi Germany and America’s Most Powerful Corporation-Expanded Edition*, Dialog press, Washington DC. 2012.

³⁸ R. Griffin, *The Palingenic Core of Fascist Ideology*, in A. Campi (a cura di), *Che cos’è il Fascismo? Interpretazioni e prospettive di ricerca*, Ideazione editrice, Roma 2003, pp. 97-122.

³⁹ D.J. Haraway, *Modest_Witness@Second_Millennium.FemaleMan_Meets_OncoMouse: Feminism and Technoscience*, Routledge, Londra 1997.