

Algoritmi, strumenti opachi

- Teresa Numerico, 30.06.2019

TEMPI PRESENTI. Una riflessione nell'ambito della intelligenza artificiale, tra ricadute etiche e politiche che ci riguardano. La fiducia nel potere di anticipazione, di carattere prescrittivo, rischia di produrre conseguenze sgradite sul rispetto dei diritti sociali e lavorativi dei soggetti più deboli

La prevedibilità statistica degli esseri umani misurata dalle procedure automatiche suggerisce un nuovo equilibrio di potere tra esseri umani e agenti artificiali, governati a loro volta da un gruppo di ingegneri. Le domande sul futuro degli eventi e dei comportamenti sono difficili da affrontare per gli umani perché riguardano l'imprevedibilità del futuro e la formulazione di giudizi che attribuiscono profili e classificazioni. L'organizzazione di classi con cui rappresentare gli esseri umani è sempre un processo sociale che ha vincitori e vinti, anche se tendiamo naturalmente a definire categorizzazioni. Il *pattern recognition* è la principale attività che gli algoritmi eseguono sui dati. Si tratta di un metodo per astrarre regolarità e modelli da una grande quantità di informazioni. Come umani, siamo molto bravi a riconoscere pattern (schemi) anche in assenza di troppi dettagli. Per esempio sappiamo riconoscere la scrittura a mano meglio di qualsiasi sistema automatico. Tuttavia è chiaro che quando estraiamo modelli dall'analisi dei dati stiamo esercitando la nostra capacità di interpretazione, stiamo esprimendo le nostre opinioni. Interpretazioni e opinioni sono inserite nei sistemi di potere.

QUANDO GLI ALGORITMI emettono i loro giudizi, vengono invece considerati neutrali e affidabili di per sé solo per il fatto che i loro metodi sono rappresentati attraverso sistemi matematici e logici. Ci affidiamo agli algoritmi sempre di più, sebbene sappiamo che sono agenti artificiali opachi e che aziende private, con i loro scopi di guadagno, li hanno prodotti. L'intelligenza artificiale è intessuta di questi strumenti opachi e sta diventando sempre più importante anche nel supportare la presa di decisione umana. Tuttavia dobbiamo fare attenzione perché l'intelligenza non è una nozione astratta che può essere definita obiettivamente e universalmente e quello che siamo preparati a considerare intelligente dipende non solo dalle capacità del sistema, ma soprattutto dall'ignoranza o dalla mancanza di esperienza dell'osservatore. È cruciale, quindi, stabilire in anticipo quali caratteristiche dell'intelligenza siamo pronti ad attribuire alla macchina.

Gli algoritmi lavorano secondo due assunzioni implicite. La prima è il principio di induzione che stabilisce che quello che è già successo tenderà a ripetersi in futuro, assume cioè un certo livello di regolarità negli eventi da prevedere. Il rischio è che, anche se la previsione è di successo, contribuisca a convalidare e replicare lo status quo. La seconda assunzione riguarda il meccanismo di attribuzione della somiglianza a due oggetti o persone, che stabilisce un criterio secondo cui entità simili si comporteranno in modo analogo in certe circostanze. Tale ipotesi è fondamentale per l'efficienza della «clusterizzazione», uno dei più importanti strumenti per interpretare e anticipare comportamenti umani.

I due presupposti sono incerti dal punto di vista epistemologico, mentre sono dei potenti strumenti interpretativi capaci di implementare ogni tipo di pregiudizi e stereotipi, accettati come conclusione solo per il carattere di presunta neutralità del dispositivo che si esibisce senza trasparenza sulla spiegazione del suo operato. La fiducia cieca nel potere di anticipazione degli algoritmi rischia di produrre conseguenze sgradite sul rispetto dei diritti sociali e lavorativi dei soggetti più deboli, anche per il carattere prescrittivo delle previsioni algoritmiche.

UN ESEMPIO dell'esito pregiudiziale di questi strumenti è fornito dall'articolo di Joy Buolamwini, *Gender shades*, illustrato fra l'altro nella mostra sull'intelligenza artificiale che è al Barbican Centre di Londra fino al 26 agosto, dal titolo *AI: More than Human*. Il progetto analizza strumenti

commerciali di riconoscimento facciale per misurare la loro capacità di distinguere maschi e femmine dal volto. Lo studio della ricercatrice ha mostrato che tutti i software analizzati - lo stato dell'arte della tecnica di face recognition - faticano a riconoscere i volti delle donne con la pelle scura. La studiosa ha fondato la Algorithmic justice league per combattere ingiustizie sociali e pratiche discriminatorie esercitate attraverso l'uso inadeguato di dati e algoritmi.

A proposito di software per il face recognition, la città di San Francisco, dove questi strumenti sono stati inventati, ne ha proibito l'uso da parte della polizia e delle altre agenzie di intelligence. Amazon Rekognition rileva che il suo sistema fa previsioni, non prende decisioni e che il livello di affidabilità del suo servizio di riconoscimento facciale deve essere incorporato in un sistema di revisione umana.

POSSIAMO ATTRIBUIRE all'intelligenza artificiale delle attività di pattern recognition basate sui Big Data, ma dobbiamo tener presente che si tratta di opinioni incorporate nei sistemi matematici per la previsione statistica e dobbiamo negoziare politicamente chi è in controllo di queste valutazioni, come sono espresse, cosa è codificato nei dati di training. È stato dimostrato tecnicamente che non esiste l'obiettività algoritmica.

È necessario che gli algoritmi usati per supportare la presa di decisione umana, in ambiti che riguardano la giustizia sociale e la lotta alla discriminazione siano valutati e controllati pubblicamente e periodicamente. I giudizi - perché di questo si tratta - espressi dagli algoritmi hanno a che fare con il potere come è sempre il caso quando si tratta di valutazioni. Il principale elemento da esplicitare è cosa sia considerato un successo per l'algoritmo.

Dobbiamo tenere a mente che la tecnologia non è una scienza che ha a che fare, sia pure con molte limitazioni, con come stanno le cose nel mondo. La tecnologia riguarda come il mondo deve funzionare perché i dispositivi siano di successo. Ha quindi un carattere prescrittivo e non descrittivo rispetto all'ambito cui si applica.

Se costruisco un ponte, decido quanto deve essere solido per non cadere, giudizio che non sempre si rivela azzeccato come mostra tristemente il crollo del ponte Morandi a Genova. Ma decido anche altre specifiche, per esempio quanto devono essere basse le sue arcate per impedire ai bus di passare sotto i suoi cavalcavia, come nel caso dei ponti che, all'inizio del Novecento Robert Moses progettò per la Southern State Parkway, che impedivano agli afroamericani, senza auto privata, di arrivare sulle spiagge di Long Island.

I DISPOSITIVI tecnologici si propongono come forza trasformativa del mondo che li circonda secondo i loro metodi. Sono sistemi sociotecnici con diretto impatto normativo sulla società, se non sono regolati e governati dalla politica. Non tutto ciò che è fattibile, è legittimo. La stessa cosa vale per gli algoritmi. Hanno un impatto regolativo perché con le loro predizioni (accurate o meno) creano un contesto obbligatorio per le persone giudicate secondo i loro criteri.

Il sistema di valutazione, la definizione di successo, il modello di astrazione dei dati, il modo in cui sono etichettati i dati di addestramento (training set) sono tutti passaggi che includono decisioni etiche che riguardano gli esseri umani e sono prese dai programmatori, non importa se lo vogliano o non ne siano consapevoli. L'effetto di queste valutazioni e opinioni è istituire e definire il futuro che stanno anticipando.

Non tutte le tecnologie disponibili vengono adoperate, ci sono i limiti di velocità sulle strade, la bomba atomica non è stata ancora lanciata dopo la prima volta, abbiamo smesso di usare l'amianto e il freon e stiamo cercando di ridurre l'uso della plastica. Meglio pianificare in anticipo invece di adottare ciecamente soluzioni che possono causare conseguenze incontrollabili.

DOBBIAMO MONITORARE con attenzione le attività degli algoritmi di intelligenza artificiale, in particolare quelli usati per prendere decisioni che impattano sulla vita delle persone. Dobbiamo controllare se legalmente e, soprattutto, politicamente le procedure sono in grado di preservare il «diritto alla spiegazione», sancito anche dal Regolamento europeo per il trattamento dei dati personali (Gdpr). È necessario preservare il diritto a una negoziazione sociale degli algoritmi e non

consentire la segretezza di metodi che impattano sulla possibilità di discriminazione e sul rispetto dei diritti sociali. La lotta per l'affidabilità e l'equità degli algoritmi è politica e va combattuta su questo terreno. Non si possono sollevare questioni di proprietà intellettuale o di protezione del vantaggio competitivo, o peggio di impossibilità di comprendere il funzionamento degli algoritmi di nuova generazione (Deep Learning) quando questi dispositivi hanno il potere di danneggiare le persone senza spiegare perché accada.

© 2019 IL NUOVO MANIFESTO SOCIETÀ COOP. EDITRICE