

Quando l'algoritmo non dice la verità

- Teresa Numerico, 20.01.2019

SCIENZA. Quando l'algoritmo non dice l'«Outnumbered», del matematico David Sumpter edito da Bloomsbury. Il volume analizza i meccanismi in relazione a fake news e social network attraversando l'attualità. L'autore è convinto che le raffinate tecniche di descrizione delle persone usate dagli algoritmi di profilazione siano molto più efficienti e corrette del giudizio umano

Secondo Bernard Stiegler, un noto filosofo francese, la costruzione dell'identità degli esseri umani è l'esito di un processo di adattamento alle tecnologie che hanno impiegato nel tempo, e dalle quali sono sempre stati dipendenti. La capacità di invenzione tecnica e quella di pensare e agire attraverso i dispositivi hanno rappresentato per l'umano la condizione di possibilità della sopravvivenza in un ambiente ben più brutale e forte di loro. Un posto speciale è riservato per il filosofo alle tecniche per la conservazione e costruzione della memoria, da cui discenderebbe la formazione stessa dell'interiorità umana: dal linguaggio al disegno, dalla scrittura alla stampa, dal computer alla rete e alla digitalizzazione. Eppure la struttura psichica umana potrebbe anche scomparire, in conseguenza dell'adozione di nuovi strumenti di esternalizzazione della memoria, con lei incompatibili.

DAVID SUMPTER professore di matematica applicata a Uppsala contesta questa idea che gli algoritmi avrebbero conseguenze trasformative sulla psicologia e le scelte collettive. Nel suo ultimo libro *Outnumbered* che significa *In inferiorità numerica* (Bloomsbury, pp. 272), discute di meccanismi algoritmici in relazione a fake news, social network, filter bubble, attraversando l'attualità, con l'analisi dei processi di categorizzazione della personalità nel caso Cambridge Analytica, il funzionamento dei sondaggi sulle elezioni e l'addestramento dei bot, piccoli agenti software capaci di interagire con gli utenti reali dei social network o come assistenti personali.

L'autore è convinto che le raffinate tecniche di descrizione delle persone usate dagli algoritmi di profilazione siano molto più efficienti e corrette del giudizio umano. Da questa prospettiva gli strumenti per trovare l'anima gemella online o i criteri con cui gli algoritmi classificano i tipi di consumatori funzionano molto meglio della limitata capacità umana di interpretare e giudicare la psicologia di amici, sconosciuti e conoscenti, carica di pregiudizi e ignorante di informazioni cruciali. Le sue opinioni sono circostanziate e costruite su solidi argomenti matematici come nel caso in cui afferma che gli esiti deteriori delle fake news e della costruzione di messaggi mirati per elettori incerti, instabili o fragili sarebbero solo un'altra informazione fallace. Le persone, infatti, secondo Sumpter, si lascerebbero influenzare dai messaggi falsificatori degli algoritmi di profilazione esclusivamente nel caso fossero già convinte dell'opinione che la fake news sostiene. In azione ci sarebbe, quindi, il bias di conferma, quella fallacia epistemica che ci spinge a essere più indulgenti nel valutare informazioni che rafforzano le nostre convinzioni.

LA SOGGETTIVITÀ che ha in mente Sumpter si riferisce a esperti consapevoli, capaci di attivare attenzione e coscienza per comprendere ciò che sta succedendo intorno a loro e che quindi sono in grado di smascherare un eventuale tentativo di condizionamento epistemico o una gentle nudge (spinta gentile), come ora è di moda chiamarla. Sarebbe bello che avesse ragione, ma le cose sono molto diverse da come può apparire a un matematico esperto e abile come l'autore del volume, che sembra riflettere se stesso nell'utente tipo dei social network. Tuttavia è giusto rilevare che i successi delle destre al potere non possano essere derubricati come effetti collaterali della credulità popolare.

Nella seconda parte del volume, l'autore si cimenta con temi controversi e prende una posizione molto più critica nei confronti delle velleità degli algoritmi di dare senso alla realtà che ci circonda,

di quanto abbia fatto in prima battuta. Il primo caso riguarda gli algoritmi che si usano per comprendere i testi, come quello noto come GloVe (global vectors for word representation) sviluppato dallo Stanford Natural Language Processing Group. L'algoritmo viene addestrato su grandi corpora testuali come Wikipedia e Gigaword (un database di 4.3 miliardi di parole prese dai siti di informazione nel mondo).

GloVe tende a essere razzista e sessista, cioè incorpora nel suo database gli impliciti pregiudizi che sono inclusi nei testi sui quali è stato addestrato. Insomma quando si tenta di far comprendere il linguaggio a un algoritmo quello restituisce una visione orientata dalle credenze e dai vincoli epistemici che sono seminati implicitamente nel suo dataset di addestramento. La difficoltà consiste nel definire le condizioni in cui sarebbe corretto esercitare la capacità di astrazione o generalizzazione. Si tratta del problema, mai risolto, dell'induzione, per dirla con Hume. Quando è possibile universalizzare, attribuendo a un soggetto una caratteristica in conseguenza di un numero di casi individuali di quella soggettività che la presentano? Se tante donne sono associate ai lavori e alla vita domestica, vorrà allora dire che è più probabile e più adeguato attribuir loro quelle qualità e non quelle connesse a professioni come ingegnere o scienziata? Com'è costruita la serie dei casi da cui si parte per generalizzare? È qui che bisogna cercare l'equivoco e l'origine del pregiudizio.

LA QUESTIONE vale anche per gli esseri umani che hanno prodotto quella testualità orientata rispetto al genere e alle sue prerogative. L'algoritmo, però, ha altri impedimenti: l'incapacità di fare differenze, di distinguere tra le affermazioni che archivia, di scegliere gli interlocutori, di valutare le fonti criticamente, di districarsi nell'ambivalenza del linguaggio, perché è governato dalla sua rigida ragione calcolante. L'intelligenza umana si riconosce dalla sua versatilità e dalla propensione a cambiare e a rivedere le precedenti convinzioni. Questa capacità è il frutto probabilmente della «neotenia», la tendenza biologica a restare a lungo bambini e soprattutto a conservare parte delle caratteristiche dell'infanzia nell'età adulta, che spinge gli umani a un costante processo di apprendimento e trasformazione. È questo che consente loro di modificarsi, di esercitare la critica, di variare la propria visione e riorganizzare le proprie credenze, avendo il coraggio di tenerle continuamente sotto osservazione e di adattarsi a condizioni sempre nuove, senza paura di tradire il passato. L'età dell'apprendimento negli umani è sempre, perché la loro vita è dominata da una costante relazione dinamica con l'ambiente.

OUTNUMBERED affronta infine l'intelligenza artificiale. Anche in questo caso Sumpter sembra essere poco fiducioso sui progressi dell'intelligenza degli algoritmi. Il vincolo più importante riguarda la loro incapacità di usare la soluzione di un problema come base per risolvere un nuovo compito non ancora affrontato. Gli umani, invece, sono molto bravi a farlo. Sumpter ritiene, quindi, che discutere della singularity, cioè di un'intelligenza artificiale generale sia troppo prematuro per avere carattere scientifico.

La difficoltà che l'autore non vede, ma che contribuisce ad alimentare con alcune argomentazioni, non è se le macchine potranno manifestare un'intelligenza artificiale paragonabile a quella degli umani, ma se questi ultimi saranno disponibili a cedere la presa di decisione e la conseguente responsabilità a chi programma le macchine, intelligenti o meno che siano. Non a caso il nodo della responsabilità decisionale, che deve restare umana, anima le Draft Ethics Guidelines for Trustworthy AI (Bozza di linee guida etiche per un'IA affidabile) pubblicate a dicembre scorso dall'High-Level Expert Group on Artificial Intelligence della Commissione Europea. Non solo l'Europa ma tutto il mondo già si sta misurando con la sfida di costruire un'intelligenza artificiale affidabile che permetta agli umani di restare autonomi, e ci riuscirà solo comprendendo i limiti e le prerogative di una tecnologia capace principalmente di eseguire calcoli.