

SCIENZA

* Il «trattamento di dati» alla ricerca di fenomeni ricorrenti non può essere considerato un paradigma

ANGELO VULPIANI

■ Da anni ci sentiamo ripetere che siamo nell'era dei «Big Data» e che quest'abbondanza di informazioni, insieme alla disponibilità di intelligenza artificiale e algoritmi, non potrà che essere una risorsa fondamentale. Sicuramente avere a disposizione una grande mole di informazioni può essere utile: più dati sull'influenza permettono una migliore prevenzione, ad esempio. Ma Big Data e algoritmi fanno nascere una serie di delicati problemi: chi è il proprietario dei nostri dati sanitari? Quale uso possono farne società di assicurazioni, datori di lavoro, tribunali? Possiamo controllare la procedura usata dagli algoritmi che decidono di aspetti importanti della mia vita?

Nel libro di Cathy O'Neil *Armi di distruzione matematica* (Bompiani, 2016) si trova un'attenta analisi dell'uso degli algoritmi in ambito sociale. Un esempio: nella contea di Broward, in Florida, un algoritmo aiuta a decidere se una persona accusata di un reato debba essere rilasciata su cauzione prima del processo. Tra i neri, l'algoritmo classifica «ad alto rischio» un numero sproporzionato di individui che però successivamente non commettono nuovi reati. La società che gestisce l'algoritmo sostiene che la metodologia non abbia pregiudizi e che classifica tutti con la stessa accuratezza. Il problema è che ci sono almeno 20 plausibili definizioni di imparzialità, in molti casi mutualmente esclusive. Inoltre, gli algoritmi istruiti dai dati raggiungono una complessità tale che ormai sfugge agli stessi creatori.

SIAMO TANTO SICURI che una descrizione molto dettagliata sia necessariamente positiva? Borges, nel breve racconto «Funes o della memoria» descrive un personaggio che, in seguito a un incidente, ricordava tutto sin nei minimi dettagli della più comune delle situazioni. Questo, ben lungi dall'essere un vantaggio, comportava la quasi totale incapacità di un pensiero astratto: «infastidiva Funes che il cane delle tre e quattordici (visto di profilo) avesse lo stesso nome del cane delle tre e un quarto (visto di fronte)».

Naturalmente i Big Data hanno un'utilità per la ricerca scientifica e anche la politica ha scoperto le loro potenzialità. Il governo italiano, nell'area milanese dell'ex-Expo, intende avviare il progetto «Human Technopole», centrato in gran parte proprio sul trattamento di grandi quantità di dati. È prevista una



Perché non possiamo scavalcare le teorie

Big Data e algoritmi ci illudono di poter capire tutto, ma non è così facile

spesa pari a una percentuale consistente dell'investimento complessivo in ricerca pubblica, come mostra l'accurata analisi della senatrice Elena Cattaneo.

Per alcuni, grazie alla disponibilità di informazioni, saremmo di fronte ad una nuova rivoluzione scientifica, con la creazione di un quarto paradigma accanto alle tre metodologie già esistenti: al metodo sperimentale, a quello teorico-matematico e a quello computazionale, si aggiungerebbe ora un nuovo approccio, consistente nel trattamento di dati alla ricerca di fenomeni ricorrenti. Nel 2008 il guru informatico Chris Anderson è arrivato a sostenere, in un articolo provocatoriamente intitolato «La fine della teoria: il diluvio di dati rende obsoleto il metodo



La visione pragmatica è ormai imperante negli atenei italiani e stranieri, in cui si privilegiano aspetti con finalità solo pratiche a scapito del senso critico e di una solida formazione

scientifico», che ormai «la correlazione è sufficiente» e possiamo smettere di cercare modelli.

SECONDO ALCUNI, come l'autore di *L'algoritmo definitivo* (Bollati Boringhieri, 2016) Pedro Domingos, in breve tempo avremo un super-algoritmo che governerà la politica in modo da far declinare la povertà e farci diventare più longevi, felici e produttivi. Ma che una correlazione tra due quantità non dica granché è cosa nota: si potrebbe citare la correlazione tra il numero di pirati e la temperatura media sulla terra, o tra il numero dei divorzi nel Maine e il consumo di margarina negli Usa per accorgersene. Quando si analizzano le cose più in profondità diventa chiaro che un approccio puramente induttivo, basato solo

sui Big Data, non può che fallire se si cerca di fare previsioni.

Infatti, per prevedere l'evoluzione futura di un sistema si potrebbe cercare nel passato una situazione «vicina» a quella di oggi. Se la si trovasse al giorno 25 gennaio 1923, ad esempio, ci si potrebbe aspettare che domani il sistema si comporti in modo simile a quanto fece il 26 gennaio 1923. Sarebbe tutto facile, in particolare nell'era dei Big Data, e potremmo non perdere tempo con la teoria. Ma quanto indietro si deve andare per trovare un simile giorno? Secondo un risultato matematico dovuto a Henri Poincaré, il tempo dipende esponenzialmente dal numero minimo N di variabili necessarie per descrivere il problema. Se con $N=1$ per trovare uno

stato simile con una certa precisione bastano cento dati, allora per $N=2$ ne servono diecimila e già per $N=6$ ce ne vogliono mille miliardi. Questa «maledizione esponenziale» non permette di farsi illusioni se il sistema non è troppo semplice. La filosofia su cui si basa l'approccio Big Data sembra dimenticare il fatto che quasi mai la scienza avanza per accumulo di dati, bensì per la capacità di eliminare gli aspetti secondari (quel «difalcare gli impedimenti» di Galileo). E ovviamente non è così facile.

DUNQUE, non convince la retorica su una presunta rivoluzione di cui finora, nonostante le dichiarazioni enfatiche, almeno nella ricerca di base non c'è traccia. L'idea di usare dati e algoritmi per fondare una scienza senza basi teoriche sembra in linea, in maniera allarmante, con la visione pragmatica ormai imperante negli atenei italiani e stranieri, in cui si privilegiano aspetti con finalità immediatamente pratiche a scapito di una solida formazione di base e soprattutto della possibilità di sviluppare senso critico. Ben vengano i «Big Data» e l'intelligenza artificiale, ma non illudiamoci troppo di fare scienza, almeno quella interessante, in maniera automatica: dobbiamo rassegnarci a studiare e farci venire qualche buona idea.

Baby scienziati



I libri del fisico Chris Ferrie hanno titoli da manuale universitario come «Fisica Newtoniana», «Chimica organica», «Fisica Statistica» o «Chimica organica». Nonostante gli argomenti siano oggettivamente ostici, nel mondo anglosassone sono diventati bestseller. Qual è il segreto di Ferrie, ricercatore all'università di Sidney? I suoi sono libri per

bambini e spiegano argomenti difficilissimi con poche e colorate figure, apprezzabili ancor prima di saper leggere e sfogliabili nel giro di pochi minuti. Negli Usa ha venduto centinaia di migliaia di copie e ora l'editrice Il Castoro ha deciso di tradurli in italiano nella collana «Baby scienziati». I primi quattro titoli appena usciti sono «Ingegneria Spaziale», «Relatività Generale», «Fisica Quantistica» e «Optica Fisica». Bastano due disegni, ad esempio, per spiegare come si formi l'arcobaleno. Noi grandi lo sappiamo già, non è vero? (An. Cap.)

Prime passeggiate spaziali tra donne



Le astronave statunitensi Jessica Meir e Christina Koch giovedì sono rimaste sospese a 400 km di quota per quasi 6 ore per sostituire una batteria difettosa di una delle apparecchiature montate sulla Stazione Spaziale Internazionale. La prima passeggiata spaziale femminile è stata effettuata dalla sovietica Svetlana

Savitskaya nel 1984. Dopo di lei, nella storia dell'astronautica, le donne protagoniste di «attività extra-veicolari» (questo è il nome ufficiale delle «passeggiate») sono state 14, ma nessuna di loro si era avventurata nello spazio senza l'accompagnamento di un uomo. Secondo i piani, in questa missione Koch dovrebbe battere un altro record, quello della permanenza di una donna nello spazio. La sua missione finirà infatti a febbraio 2020, dopo 328 giorni. Sono solo 12 in meno del recordman Scott Kelly, che è rimasto in orbita per 340 giorni. (An. Cap.)

Piantare alberi non è sempre una buona idea



Una ricerca recente pubblicata su «Science» ha quantificato il potenziale beneficio della riforestazione nella lotta al cambiamento climatico. Secondo lo studio, guidato da Jean-François Bastin del Politecnico di Zurigo, la riforestazione di 900 milioni di ettari potrebbe assorbire 205 giga-tonnellate di anidride

carbonica. Ora lo studio è contestato sulla stessa rivista da un altro gruppo di scienziati, guidati Joseph Veldman dell'università A&M del Texas; il beneficio attribuito alla riforestazione sarebbe sovrastimato: una superficie alberata ha un minore potere riflettente della radiazione solare e può persino aumentare la temperatura dell'atmosfera. Inoltre, la ricerca di Bastin includeva nelle aree da riforestare anche savane tropicali come il Cerrado brasiliano, ecosistemi che andrebbero perduti a causa della riforestazione, con danni ambientali ancora maggiori. (An. Cap.)