

Predire il futuro

Gli algoritmi aiuteranno il medico. Cambieranno il suo modo di lavorare. Soprattutto, i big data potranno rivelarsi una buona notizia per i malati

“ Sono gli algoritmi e non i dataset che dimostreranno di essere un agente di cambiamento. ”

Questa è la frase a effetto che apre l'articolo "Predicting the future: big data, machine learning, and clinical medicine", di Ziad Obermeyer e Ezekiel J. Emanuel, uscito sul *New England of Medicine* a fine settembre 2016.

La ragione della convinzione dei due autori è collegata all'importanza che avrà in futuro l'intelligenza computazionale (machine learning). Ne abbiamo parlato con uno dei due autori, Ziad Obermeyer, che studia da tempo gli algoritmi a supporto delle decisioni cliniche.

[...] Stiamo iniziando a intravedere un uso nuovo dei computer in medicina. Oggi, usiamo i computer per applicare delle regole: ci avvertono se proviamo a prescrivere la ciprofloxacina insieme al coumadin. Sono regole che già conosciamo e che talvolta semplicemente dimentichiamo, perché è notte fonda o perché è troppo presto al mattino oppure perché non sappiamo che quel paziente assume il coumadin.

In un futuro neanche troppo distante, degli algoritmi ci diranno del nostro paziente cose che non sappiamo: ci comunicheranno la probabilità che il malato possa morire avviando un determinato regime chemioterapico o ci avvertiranno del rischio di infarto miocardico di un altro paziente nei sette giorni successivi al momento della nostra interrogazione, sulla base della sua storia clinica. Sarà entusiasmante, non solo per la pratica clinica, ma anche come guida per nuovi modi di comprendere la salute e la malattia utilizzando dati molto complessi.

Perché ritenete che lasciar parlare i dati da soli possa essere problematico? Come possiamo affrontare un problema del genere?

Quando sviluppiamo un algoritmo, vediamo molto il fenomeno del cosiddetto *overfitting* (che si verifica quando il modello stesso ha troppi parametri relativi al numero di osservazioni, condizione che porta a sovrastimarne l'utilità, ndr). Possiamo predire molto bene un determinato esito sulla base di un dataset ma a causa di alcune bizzarrie nei dati possiamo scoprire che le conclusioni alle quali siamo arrivati non funzionano una volta che sono applicate al mondo reale. In molte delle gare per scienziati che vengono organizzate – ma non relative alla medicina



Intervista a
Ziad Obermeyer

Assistant Professor of Emergency medicine e Assistant Professor of Health care policy Harvard Medical School

– questo problema viene gestito rilasciando un set di dati a chi è in gara per lo sviluppo degli algoritmi, e una volta che gli algoritmi finali sono uploadati vengono testati per la valutazione su un dataset completamente separato. Questo principio di separare i dati utili allo sviluppo del modello da quelli per la valutazione di quest'ultimo è estremamente importante, come sa già chi lavora alla produzione di algoritmi non basati sull'analisi computazionale.

La qualità dei dati è collegata alla loro quantità?

Ehm, probabilmente è inversamente proporzionale...

La correlazione non implica la causalità: è un bel problema in epidemiologia clinica. Dovremmo aspettarci una soluzione dal machine learning?

Ne dubito. Ci sono alcuni ricercatori che stanno svolgendo un lavoro molto interessante in questo ambito, utilizzando l'analisi computazionale per costruire dei metodi migliori di aggiustamento del rischio, propensity score e variabili strumentali, ma alla fine questi algoritmi trovano delle correlazioni, e questa è la loro forza. È improbabile che possano risolvere anche qualcuno dei problemi fondamentali della inferenza causale nei dataset sperimentali.

Quali sono le aree della medicina che lei e Ezekiel J. Emanuel ritenete possano essere davvero radicalmente cambiate dalla trasformazione del dato in conoscenza?

Certamente, questa trasformazione aiuterà il medico nella formulazione della prognosi: comprendere quando un paziente morirà o se un cancro metastatizzerà. Predire il futuro

è qualcosa per cui questi algoritmi sono estremamente adatti.

Migliorerà anche l'accuratezza diagnostica: saranno suggeriti esami ad alto valore aggiunto e potrà ridursi la richiesta di test inutili. Questo però è più complicato, e dovremmo attendercelo in un futuro meno prossimo, perché persino il medico spesso non ha chiaro quale sia il "gold standard" da indicare alla macchina per farle apprendere come formulare la diagnosi corretta! È un aspetto che rende più difficile l'addestramento degli algoritmi.

Un punto collegato è che gli algoritmi sottrarranno molto lavoro ai radiologi e agli anatomopatologi, che lavorano su dati digitali... che potrebbero essere invece inviati a una macchina. E infine interpreteranno un flusso di dati dalla anestesiologia e dalla terapia intensiva.

Così, in 20 anni, i radiologi non esisteranno da nessuna parte del mondo nella forma attuale: dovranno adattarsi al cambiamento, un po' come i lavoratori delle costruzioni che fanno oggi un lavoro del tutto differente da quello che compivano prima della meccanizzazione avvenuta cento anni fa. I cassieri in banca non maneggiano più contante ma devono essere capaci di eseguire transazioni molto più complesse di una volta. La tecnologia non sempre elimina il lavoro; talvolta lo cambia e coloro che devono adattarsi possono finire col diventare i veri vincitori. F

1. Obermeyer Z, Emanuel EJ. Predicting the Future: Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med* 2016;375:1216-9.
2. Parikh RB, Obermeyer Z, Bates DW. Making Predictive Analytics a Routine Part of Patient Care. *Harvard Business Review* 2016; 21 aprile. hbr.org/2016/04/making-predictive-analytics-a-routine-part-of-patient-care - Ultimo accesso 14 ottobre 2016.

“Gli algoritmi sono tanto affidabili quanto lo sono i dati su cui sono basati.”