

L'ultimo miglio: dove l'intelligenza artificiale incontra la realtà

Sviluppare e valutare il nuovo nel contesto che lo vedrà protagonista

Sembra che la priorità sia migliorare le performance dell'intelligenza artificiale quando invece dovremmo concentrarci maggiormente sull'implementazione delle tecnologie basate sui dati di questo ramo dell'informatica. Sul *Journal of Medical Internet Research* Enrico Coiera invita a non sottovalutare il "terzo miglio" di quel viaggio che parte dalla raccolta dei dati e termina con l'ingresso dell'innovazione nel mondo reale. Serve un approccio più agile: sviluppare e testare in modo interattivo l'intelligenza artificiale laddove verrà utilizzata.

Lo sviluppo di tecnologie basate sui dati, come il machine learning, richiede a grandi linee tre fasi. Nel primo miglio una volta selezionato il compito da svolgere, come per esempio la diagnosi, i dati vengono acquisiti, possibilmente etichettati, e pre-elaborati o "puliti". Il miglio successivo, quello di mezzo, si focalizza sullo sviluppo e sulla valutazione delle performance tecniche dei diversi algoritmi creati a partire da questi dati. Solo nell'ultimo miglio, gli algoritmi vengono incorporati nei processi del real world e testati sulla base dei loro effetti sugli outcomes in contesti reali.

Ciascun miglio ha le sue sfide. Quelle del primo miglio riguardano la raccolta e il trattamento di dati di alta qualità. Per tecnologie come il machine learning, che dipendono spesso da grandi quantità di dati di alta qualità, un collo di bottiglia nell'acquisizione dei dati si traduce in un ostacolo all'applicazione della tecnologia. Il miglio di mezzo contiene invece tutte le sfide dello sviluppo di algoritmi basati sui dati, tra cui la gestione dei bias, la replicabilità, l'inferenza causale, l'eventuale eccesso di adattamento ai dati usati per l'addestramento e il miglioramento della generalizzazione di tutti i modelli e algoritmi sviluppati.

Nell'ultimo miglio, si affronta la realtà che l'intelligenza artificiale non fa nulla da sola. In qualche modo, l'intelligenza artificiale deve essere collegata ai processi del real world e il suo impatto su tali processi deve essere consequenziale. È a questo punto che una tecnologia sviluppata per sé stessa è destinata a fallire. Per esempio, una cosa è dimostrare che il machine learning riesca a diagnosticare il tumore a partire dalle scansioni della tiroide, con la stessa accuratezza dell'uomo, se non superiore – una prodezza tecnica; altra cosa è l'utilità di questa prodezza. Nell'attuale contesto, in cui il carcinoma tiroideo viene diagnosticato e anche trattato oltremisura, non è detto che abbiamo bisogno di diagnosi migliori. Piuttosto abbiamo bisogno di un approccio più sfumato e meno aggressivo².

Le sfide dell'ultimo miglio sono quindi quelle dell'implementazione e, per i ricercatori, della scienza dell'implementazione. Queste sfide esistono a diversi livelli e includono quanto segue.

Misurazione. Le metriche standard delle performance dell'intelligenza artificiale si riferiscono al modo in cui viene completato il task assegnato. Tradizionalmente si adottano misure di performance come sensibilità, specificità e area al di sotto della curva caratteristica operativa del ricevitore. Vi è, tuttavia, una lunga serie di eventi che devono verificarsi tra elevate performance tecniche e impatto effettivo sull'esito clinico. Il successo in ciascuna delle fasi iniziali di questa *information value chain* è necessario ma non sufficiente per garantire l'impatto dell'applicazione nel mondo reale³. La valutazione dei risultati del mondo reale richiede quindi un passaggio dalla misurazione dell'accuratezza tecnica alla valutazione dell'impatto sui processi e sulle persone. Per esempio la curva caratteristica operativa del ricevitore può essere fuorviante in contesti clinici⁴ in quanto misura le performance complessive di un algoritmo attraverso l'intera curva caratteristica operativa del ricevitore, mentre il funzionamento nel mondo reale ideale può al meglio essere limitato a un segmento di esso⁵.

Le tecnologie dovrebbero nascere nel profondo della rete in cui vivranno. ”

Generalizzazione e calibrazione. Nel momento in cui l'intelligenza artificiale viene addestrata su dati storici, le sue performance future dipendono dal modo in cui i nuovi dati corrispondono a quelli storici. Una sfida comune nell'implementazione è scoprire che un algoritmo ad alte prestazioni, sviluppato a partire dai dati di una popolazione, si deteriora quando applicato su un'altra popolazione, riflettendo le differenze sottostanti nella frequenza e nella natura degli eventi all'interno dei set di dati. Per questo motivo, l'intelligenza artificiale potrebbe dover essere sintonizzata specificamente sulla popolazione finale. In molti contesti, anche questa popolazione finale sarà dinamica, variando a causa di eventi ricorrenti come per esempio modifiche stagionali delle malattie, cambiamenti nelle caratteristiche della popolazione e nuovi eventi imprevedibili quali focolai di malattie. Ciò significa che potrebbe essere necessario ricalibrare l'intelligenza artificiale, periodicamente o dinamicamente, per rispecchiare i cambiamenti della popolazione. Dovremo inoltre monitorare attentamente le performance dell'intelligenza artificiale per rilevare cambiamenti nel suo comportamento che indicano la necessità di ricalibrare il software⁶. Maggiori sono le difficoltà quando più efficacemente l'intelligenza artificiale migliora i risultati e più velocemente le sue performance sembrano peggiorare, in quanto il suo stesso successo può alterare l'associazione tra predittori nel suo modello e outcomes raggiunti⁷.



Enrico Coiera

Direttore
Centre for health
informatics
Macquarie university
Sydney

Contesto locale. Un principio fondamentale delle scienze dell'implementazione è che alle differenze nel contesto in cui la tecnologia viene inserita si associno cambiamenti nelle performance. Se consideriamo un'organizzazione come una rete di persone, processi e tecnologie, è chiaro che i network alla base di due organizzazioni distinte saranno diversi. L'implementazione può essere interpretata come l'atto di adattare una nuova tecnologia o un processo a una rete organizzativa preesistente e la bontà di adattamento della tecnologia alla rete determinerà qualsiasi impatto sulle prestazioni organizzative⁸. Ciò vale tanto per l'intelligenza artificiale quanto per la salute digitale, in generale, o per qualsiasi nuovo processo o tecnologia. A complicare ulteriormente le cose è la dinamicità delle reti organizzative stesse. L'impatto di una tecnologia cambierà quindi con il tempo, poiché cambia il modo in cui "si adatta" a una rete organizzativa e probabilmente a causa della sua stessa presenza: le vecchie connessioni scompariranno o saranno sostituite da nuove connessioni.

Il mondo del software ha compiuto un cambiamento importante dal considerare lo sviluppo del software come un processo lineare che inizia con le esigenze degli utenti e termina con prove utente finale, a un processo flessibile in cui gli utenti sono integrati in un processo rapido e iterativo che adatta il software agli utenti. La scienza dell'implementazione deve percorrere lo stesso percorso e in particolar modo con l'intelligenza artificiale che cambia dinamicamente.

Le tecnologie dovrebbero nascere nel profondo della rete in cui vivranno.

Lo sviluppo dell'intelligenza artificiale non dovrebbe essere interpretato come un viaggio lineare che si estende dal primo all'ultimo miglio. In tal modo si rischierebbe di arrivare a un prodotto finale che non soddisfa le esigenze del mondo reale, proprio come accade con i software. Al contrario l'implementazione dovrebbe essere vista come un processo agile, iterativo e leggero per ottenere dati da usare per l'addestramento e per sviluppare algoritmi e trasformarli in strumenti e flussi di lavoro. Sarà essenziale trovare il giusto equilibrio tra il riutilizzo di una qualsiasi tecnologia e la necessità di rispondere alle esigenze locali⁹. In ogni caso, le applicazioni dell'intelligenza artificiale non dovrebbero essere create lontane dal contesto in cui verranno impiegate. Idealmente, dovrebbero nascere nel profondo della rete in cui vivranno. ■

Questo articolo riprende parte dell'articolo di Enrico Coiera "The Last Mile: Where Artificial Intelligence Meets Reality" pubblicato sul *Journal of Medical Internet Research* www.jmir.org/2019/11/e16323.

Ringraziamo gli autori e il *Journal of Medical Internet Research* per l'autorizzazione.

- Li X, et al. *Lancet Oncol* 2019;20:193-201.
- Coiera E. *Lancet Oncol* 2019;20:166-7.
- Coiera E. *Stud Health Technol Inform* 2019;263:35-48.
- Cook NR. *Circulation* 2007;115:928-35.
- Thomas G, et al. *Diagn Progn Res* 2017;1:17.
- Ong M, et al. *J Am Med Inform Assoc* 2013;20:506-12.
- Lenert M, et al. *J Am Med Inform Assoc* 2019;26:1645-50.
- Coiera E. Chapter 12: Implementation, in *Guide to Health Informatics*. Boca Raton: CRC Press, 2016.
- Johannessen L, et al. *Comput Supported Coop Work* 2009;18:607-34.