

Rapporti di lavoro

Intelligenza artificiale e processi decisionali in azienda

Evidenziato da uno studio in ambiente sanitario che la qualità del modello di IA influisce sul valore della prestazione professionale

di Andrea Rossetti e Luca Barbieri

15 Febbraio 2024

Il ricorso a sistemi di intelligenza artificiale (IA) che possano contribuire a svolgere un'analisi approfondita ed estesa delle dinamiche e trasformazioni proprie di un'organizzazione o di processi complessi può ovviamente influire, anche in misura decisiva, sulle decisioni adottate dagli organi direttivi di un'impresa e sulla definizione di strategie e politiche aziendali.

Stampa

Un recente studio condotto con riguardo agli impatti che in ambito sanitario possono derivare dall'impiego di un sistema di intelligenza artificiale ha rivelato, con il conforto di misurazioni ed evidenze anche di natura quantitativa, che la qualità del modello di IA influisce sul valore della prestazione professionale, incidendo significativamente sull'accuratezza diagnostica del personale medico. In *Measuring the Impact of AI in the Diagnosis of Hospitalized Patients*, studio pubblicato recentemente sul *Journal of the American Medical Association (JAMA)*, è dimostrato come l'impiego di un modello di IA che sia preceduto da un'opportuna formazione circa il suo funzionamento e che offra spiegazioni in relazione al risultato frutto dell'elaborazione del sistema stesso incrementi l'accuratezza delle diagnosi effettuate nella misura di 4,4 punti percentuali. Quando invece il modello di IA sia caratterizzato da numerose distorsioni e incontrollati pregiudizi e non siano stati resi noti agli utenti i principi del suo funzionamento, l'accuratezza della diagnosi è drasticamente ridotta (11,3 punti percentuali).

In altri termini, lo studio evidenzia come l'utilizzo di un modello di IA conforme alle finalità perseguite migliori le competenze professionali del medico utente e incrementi il grado d'accuratezza dell'attività di diagnosi, mentre l'impiego di un modello di IA il cui funzionamento sia caratterizzato da distorsioni sistematiche incide negativamente sulla qualità della prestazione professionale, anche quando sia stata condotta un'attività di formazione volta a rilevare i profili critici e distorsivi del sistema, esponendo così al rischio di perpetuazione e intensificazione di pregiudizi (bias) ed effetti distorsivi che potrebbero causare danni ai pazienti.

Lo studio ha altresì portato in luce l'evidente difficoltà del personale medico nel rilevare la presenza di distorsioni sistematiche nel funzionamento del modello di IA. È per tale ragione che la Food and Drug Administration (FDA) attribuisce cruciale rilievo alla possibilità per il personale medico di valutare autonomamente le raccomandazioni fornite dal modello di IA posto a supporto alla decisione clinica e che la proposta di una Carta dei diritti per l'IA avanzata dal White House Office of Science and Technology Policy sottolinea l'importanza di introdurre efficaci strumenti esplicativi che consentano d'individuare gli errori di funzionamento di un modello di IA, garantendo una protezione affidabile dalle distorsioni sistematiche, ancor più rilevanti quando il modello di IA elabori immagini. La mitigazione dei rischi a cui si è accennato, e che non sono propri del settore sanitario ma caratterizzano qualsivoglia settore economico nel quale s'impieghi un modello di IA, può essere realizzata: attraverso una formazione specifica circa i limiti e le potenziali distorsioni che è possibile siano prodotte da un sistema di IA, in modo che l'utente abbia la capacità di rilevare eventuali anomalie e distorsioni; sviluppando, in collaborazione con l'utente, strumenti esplicativi che permettano gli opportuni adattamenti del modello di IA rispetto alle esigenze pratiche e cognitive del contesto; mediante la produzione di flussi informativi standardizzati da parte del modello di IA, espressi in un linguaggio chiaro e accessibile e accompagnati da una verifica empirica; tramite una rigorosa validazione del modello di IA, anche sulla base di test specifici, affinché eventuali distorsioni possano essere identificate e corrette prima che il modello di IA sia in uso.

L'adozione di tali misure collettive è peraltro promossa dalle norme contenute nel Titolo III dell'Artificial intelligence act nella formulazione da ultimo concordata in sede di trilogò e che contempla, ad esempio, l'obbligo per il fornitore di un sistema di IA ad alto rischio di i) adottare un sistema di gestione del rischio (articolo 9) e della qualità (articolo 17) ed una specifica governance per la gestione dei dati in sede di addestramento del modello (articolo 10), ii) consegnare all'utente la documentazione tecnica afferente al sistema di IA (articolo 11), il cui

funzionamento dev'essere peraltro incardinato sul principio di trasparenza in modo che l'utente sia posto nelle condizioni di poter interpretare i dati (articolo 13), iii) progettare un sistema che consenta di svolgere un'attività di sorveglianza e verifica del funzionamento del sistema (articolo 14) e di iv) garantire un funzionamento coerente con le finalità perseguite nonché un adeguato livello di accuratezza del sistema (articolo 15). L'adozione di tali misure e l'osservanza degli obblighi imposti sia al fornitore che all'utente (articolo 29) sono dunque elementi decisivi per consolidare l'affidabilità di un modello di IA e consentire un utilizzo consapevole e prudente del modello stesso, affinando al tempo stesso le capacità e le competenze professionali dell'utente e, quindi, la fondatezza e la fluidità dei processi decisionali.

Il Sole 24 ORE aderisce a  **The Trust Project**

P.I. 00777910159 © Copyright Il Sole 24 Ore Tutti i diritti riservati

Il Sole
24 ORE