

Errori e rischi, il lato oscuro delle macchine: quando l'intelligenza artificiale dà segni di stupidità

di **Martina Gianecchini**



Le applicazioni dell'intelligenza artificiale alle attività lavorative sono ormai diffuse: robot che guidano automobili e camion, che coltivano orti e curano animali, che decidono chi assumere e promuovere in azienda, che lavorano in catena di montaggio e gestiscono magazzini. Accanto a queste esperienze di successo, si vanno però moltiplicando notizie di fallimenti e resistenze che suggeriscono come l'adozione dei sistemi di *machine learning* all'interno delle nostre case e dei nostri luoghi di lavoro sia ancora ricca di difficoltà.

Solo per fare alcuni esempi, è di qualche mese fa la notizia di un albergo giapponese che ha «licenziato» circa metà dei suoi 243 androidi addetti ai servizi ai clienti perché soggetti a guasti, costosi da mantenere e incapaci di adattarsi alle richieste dei clienti tanto da diventare quasi molesti. Un destino simile era toccato a Fabio, un robot impiegato in un supermercato con il compito

di dare indicazioni sulla collocazione dei prodotti sugli scaffali. All'inizio del 2018, si era poi sollevata una discussione intorno ai *robot trader*, sistemi di intelligenza artificiale che operano in Borsa, che per effetto di alcune fake news avevano compiuto una serie di investimenti errati. Oppure, più di recente, il software utilizzato da Amazon per analizzare i curriculum dei candidati è stato accusato di discriminare negativamente le donne, in quanto programmato a partire da informazioni sulle persone assunte in azienda nel corso degli ultimi dieci anni. Infine, in un interessante esperimento, il sistema di intelligenza artificiale Miss Debater di IBM, che si esprime con voce femminile ed è stato alimentato con oltre 10 miliardi di frasi prese da giornali e riviste scientifiche, è stato sconfitto da un essere umano in un dibattito relativo ai benefici del finanziamento pubblico alla scuola d'infanzia. Questa breve rassegna di esempi mostra come l'applicazione dei sistemi di intelligenza artificiale

non sia priva di rischi e di conseguenze inattese per le persone, per le imprese e per l'intera società. Con riferimento alle persone, si possono generare rischi per la loro salute e incolumità fisica (si pensi ad esempio al malfunzionamento di sistemi di guida autonoma o a una diagnosi errata di un modello di analisi medico), per la loro privacy e reputazione nel caso in cui vengano rubati e diffusi dati sensibili, per le loro opportunità occupazionali se ad esempio venissero scelti dalle aziende in base alle loro relazioni sui social network. Con riferimento alle aziende, possono soffrire di rischi economici collegati all'incapacità dei sistemi informatizzati di prevedere improvvisi cambiamenti nei mercati finanziari, rischi legali dovuti alla protezione dei dati dei loro collaboratori, rischi di mercato dovuti ad esempio all'uso di sistemi che generano in automatico avvisi pubblicitari, che potrebbero essere percepiti dai consumatori come fraudolenti o invadenti. Infine, con riferimento

all'intera società, numerosi casi di attualità stanno mostrando come i nostri sistemi politici ed elettorali siano facile preda di sistemi automatizzati in grado di generare informazioni che influenzano le opinioni degli elettori, e alcuni recenti falle e blocchi dei maggiori social media hanno mostrato la fragilità di un sistema di infrastrutture virtuali e fisiche fortemente interconnesso. Le cause di fallibilità di sistemi complessi come quelli di intelligenza artificiale sono numerose e possono verificarsi in diversi momenti della loro realizzazione: da quando vengono progettati a quando vengono alimentati con le informazioni che devono apprendere, a quando infine entrano in contatto con gli esseri umani. Seppure molti di questi rischi appaiano ancora remoti o controllabili, è opportuno che se ne studino le conseguenze sin da subito per evitare che sistemi intelligenti esprimano tutto il loro potenziale di stupidità.

© RIPRODUZIONE RISERVATA