

Intelligenza artificiale: una chiave per il futuro di Domenico Marino (UNIRC)

1. Introduzione

L'intelligenza artificiale (IA) sta avendo un impatto sempre più rilevante non solo sull'economia, ma anche sulla vita quotidiana delle persone, creando nuove opportunità, ma anche introducendo potenziali rischi. L'IA non può essere considerata una soluzione universale ai problemi, ma deve essere utilizzata con cautela, tenendo conto dei rischi che derivano dalla sua crescente potenza e autonomia. Ad esempio, un sistema di IA progettato per ridurre il consumo di carburante di un aereo potrebbe prendere decisioni non sicure in situazioni impreviste, come fermare completamente l'aereo per raggiungere il suo obiettivo primario.

Le questioni etiche legate all'IA sono complesse e stanno diventando centrali nei dibattiti scientifici. Asimov, noto scrittore di fantascienza, aveva anticipato il problema cercando di codificare tre leggi della robotica, ma queste sono troppo semplici per il mondo reale. Gli sviluppi dell'IA moderna richiedono una riflessione più approfondita su come bilanciare le opportunità offerte dalla tecnologia con le sue implicazioni morali e legali. In particolare, il concetto di "personalità non umana" solleva nuove domande sulla possibilità di attribuire diritti e responsabilità agli esseri non umani e non biologici, come le IA avanzate.

2. IA e Etica: una panoramica

Negli ultimi anni, il dibattito sull'etica dell'IA ha prodotto numerose riflessioni accademiche. Uno dei principali problemi evidenziati è la separazione tra moralità e responsabilità. Floridi (2019) sottolinea come un agente artificiale possa essere programmato per seguire norme morali, ma senza che ciò implichi una vera e propria responsabilità. Allen et al. (2006) hanno sottolineato come, in caso di danni causati dall'IA, sia complesso individuare chi dovrebbe essere ritenuto responsabile. Questo perché, anche se accettiamo l'idea che un'IA possa seguire regole morali, manca ancora un collegamento chiaro tra moralità e responsabilità.

Numerosi altri studiosi hanno affrontato la questione. Susan Anderson ha esplorato il problema dell'IA che crea altre IA, e come queste nuove creazioni possano rappresentare nuove sfide etiche. Floridi (2019) ha individuato cinque principi fondamentali per l'IA benevola, tra cui beneficenza, non maleficenza, autonomia, giustizia ed esplicabilità. Quest'ultimo principio riguarda sia la comprensibilità delle decisioni prese dall'IA sia la capacità di individuare chi è responsabile del suo funzionamento.

A livello normativo, l'Unione Europea ha introdotto diverse linee guida, come le Guidelines for Trustworthy Artificial Intelligence (2019), che delineano sette principi per garantire un'IA etica e affidabile. Questi includono la supervisione umana, la robustezza tecnica, la privacy, la trasparenza e l'equità. Il White Paper on Artificial Intelligence (2020) dell'UE ha tracciato un percorso per uno

sviluppo etico dell'IA, culminato nell'approvazione del Artificial Intelligence Act (2023), il primo atto legislativo globale per regolamentare l'IA in modo da proteggerne l'uso sicuro ed etico.

3. I rischi di un'IA non etica

Un importante rischio è rappresentato dall'uso non etico dell'IA, come nel caso della sorveglianza, del riconoscimento facciale, delle armi autonome e della manipolazione delle informazioni. I sistemi di IA che gestiscono il riconoscimento facciale potrebbero essere usati da regimi totalitari per reprimere la libertà individuale, mentre le armi autonome possono decidere da sole di colpire obiettivi civili.

Oltre all'uso non etico, vi è anche il rischio che un'IA possa prendere decisioni non etiche a causa di errori nei dati di apprendimento o di pregiudizi. Gli algoritmi di apprendimento automatico possono sviluppare bias autonomamente o attraverso i dataset forniti dagli esseri umani. Questo è accaduto, ad esempio, con un software utilizzato negli Stati Uniti per prevedere la probabilità di recidiva dei criminali, che discriminava le persone di colore. Un altro esempio è quello delle decisioni prese da algoritmi nel settore delle consegne, dove alcuni lavoratori possono essere penalizzati ingiustamente rispetto ad altri.

4. Verso una "Machine Ethics"

Per affrontare questi problemi, è necessario sviluppare una "Machine Ethics", ossia una struttura etica integrata nel funzionamento delle IA. Virginia Dignum ha proposto tre approcci all'interazione tra etica e IA:

- **Ethics by Design:** consiste nell'integrazione di principi etici all'interno degli algoritmi, in modo che l'IA possa prendere decisioni eticamente giuste.
- **Ethics in Design:** prevede l'adozione di normative e metodi ingegneristici che valutino le implicazioni etiche dei sistemi di IA.
- **Ethics per Design:** si occupa della creazione di codici di condotta e standard che guidino lo sviluppo e l'utilizzo dell'IA in modo etico.

In tutti questi casi, è fondamentale garantire che l'IA sia sempre supervisionata dall'uomo, soprattutto quando le sue decisioni potrebbero avere ripercussioni su larga scala. L'interazione tra essere umano e macchina deve essere strutturata in modo tale che l'uomo mantenga il controllo finale e possa intervenire qualora l'IA prenda decisioni dannose. Questo concetto, noto come "intelligenza artificiale estesa", mira a integrare i driver etici umani nelle decisioni delle IA, per evitare il rischio che un'IA ottimizzi una singola parte di un problema, trascurando l'interesse generale.

L'IA deve dunque essere addestrata seguendo valori umani condivisi, basati sul rispetto della dignità umana e della solidarietà. Per farlo, le IA devono essere "educate" a riconoscere e rispettare i principi etici universali, attraverso un processo che Floridi ha definito "costruttivismo etico". L'obiettivo è che le IA facciano sempre scelte a beneficio dell'umanità, evitando comportamenti dannosi o discriminatori.

5. Implicazioni per la business ethics

L'introduzione dell'IA nel contesto aziendale solleva questioni etiche rilevanti. Le imprese sono sempre più responsabili per i danni causati dalle scelte autonome dell'IA, che siano danni civili o, in alcuni casi, penali. Ad esempio, un robot sanitario potrebbe essere chiamato a rispondere di un errore medico, oppure un algoritmo di selezione del personale potrebbe essere accusato di discriminazione durante il reclutamento. In tutti questi casi, l'azienda che utilizza l'IA potrebbe essere ritenuta legalmente responsabile.

Per minimizzare questi rischi, le aziende devono introdurre meccanismi che garantiscano la supervisione umana sulle decisioni prese dall'IA. Ad esempio, nel caso di una diagnosi medica, un medico dovrebbe sempre convalidare l'output fornito dall'algoritmo, così come nella guida autonoma ci dovrebbe essere un controllo umano in grado di intervenire in qualsiasi momento. Anche se ciò potrebbe comportare costi maggiori, è necessario per evitare danni peggiori.

Un altro elemento fondamentale per ridurre i rischi è garantire la trasparenza degli algoritmi. I soggetti coinvolti nelle decisioni dell'IA devono avere accesso a tutte le informazioni rilevanti, in modo da poter esercitare un controllo sulle decisioni prese. Questo significa che l'algoritmo deve essere progettato in modo da consentire una comprensione chiara del suo funzionamento e dei dati utilizzati.

Infine, le aziende devono prestare grande attenzione alla selezione dei dati utilizzati per l'addestramento dell'IA. Durante la pandemia di COVID-19, ad esempio, i ricercatori hanno frettolosamente aggregato dati provenienti da fonti diverse, alcuni dei quali includevano scansioni di bambini sani usate come archetipi di casi non COVID. Di conseguenza, l'algoritmo ha imparato a identificare i bambini invece che i casi di COVID-19. È quindi cruciale utilizzare dati certificati e verificati, e garantire che il processo di apprendimento continui ad essere accurato anche quando vengono introdotti nuovi dati.

Per affrontare il problema delle decisioni eticamente corrette, sarebbe utile creare un sistema di certificazione delle IA basato sulla loro capacità di prendere decisioni etiche. Solo le IA che superano determinati test di etica dovrebbero essere certificate, limitando così la responsabilità legale delle aziende che le utilizzano.

6. IA, dati e privacy

Uno dei temi più delicati riguardo all'IA è la gestione dei dati e la privacy, soprattutto nel settore sanitario. L'IA sta rivoluzionando la diagnosi medica, la pianificazione dei trattamenti e la gestione dei pazienti, ma al contempo solleva preoccupazioni sulla privacy e l'utilizzo dei dati sanitari da parte di aziende farmaceutiche o tecnologiche per fini commerciali.

L'AI Act dell'Unione Europea introduce regole rigorose per la tutela dei dati personali, in particolare quelli sanitari. Le aziende devono minimizzare la quantità di dati raccolti e garantire che questi siano utilizzati solo per scopi specifici, come la diagnosi o la ricerca medica, nel pieno rispetto delle normative sulla privacy. Inoltre, è richiesto il consenso esplicito dei pazienti per il trattamento dei loro dati.

Il trasferimento dei dati sanitari nel cloud è un processo complesso che richiede attenzione per evitare violazioni della privacy e della sicurezza. I vantaggi del cloud includono la riduzione dei costi e la flessibilità nella gestione dei dati, ma è necessario implementare rigorosi protocolli di sicurezza per proteggere le informazioni sensibili.

L'IA ha dunque un enorme potenziale per migliorare molti settori, ma per garantire un uso etico e responsabile è fondamentale che vengano stabilite regole chiare. Solo attraverso una supervisione umana e una regolamentazione accurata sarà possibile evitare i rischi e massimizzare i benefici dell'IA a favore della società nel suo complesso.

Alcune considerazioni conclusive

L'intelligenza artificiale è uno strumento potente che sta già rivoluzionando numerosi settori, ma comporta anche significativi rischi etici e legali. Per garantire che l'IA operi in modo sicuro e responsabile, è necessario sviluppare un sistema di etica della macchina basato su valori condivisi e implementare meccanismi di supervisione umana. Le aziende devono assumere un ruolo attivo nell'introdurre trasparenza, selezione dei dati accurata e responsabilità nel loro utilizzo dell'IA. Allo stesso tempo, la regolamentazione deve continuare a evolversi per mantenere il passo con i rapidi sviluppi tecnologici, assicurando che l'IA continui a servire l'umanità rispettando i diritti e la dignità di tutti.