LUCI ED OMBRE DELL'INTELLIGENZA ARTIFICIALE

L'intelligenza artificiale promette progressi incredibili, ma c'è un problema spinoso: **per noi umani non sempre è possibile spiegare come gli algoritmi di apprendimento giungano alle loro decisioni.**

È un fenomeno responsabile di una nuova ondata di radicali cambiamenti: l'automazione cognitiva, che può essere riassunta con la parola chiave "AI" (Artificial Intelligence).

L'intelligenza artificiale è un territorio dove le opinioni si dividono. Per alcuni, raffigurati in modo preminente dall'imprenditore tecnologico Elon Musk, rappresenta un'altra, se non la minaccia decisiva, per l'umanità. Per altri, tra i quali il fondatore di Facebook Mark Zuckerberg, al contrario, saranno le macchine per l'apprendimento a rendere la nostra vita migliore e più sicura in futuro. Da qualunque parte ci si possa schierare in questa fondamentale disputa tra ottimisti della tecnologia e pessimisti, è indiscutibile che gli algoritmi di apprendimento saranno sempre più presenti in molti settori, che si tratti di medicina, di guida autonoma o della valutazione automatica di gigantesche quantità di dati generati da esperimenti scientifici. L'intelligenza artificiale sta però causando una serie di conseguenze indesiderate come violazioni della privacy, discriminazioni, incidenti e manipolazione dei sistemi politici; poiché è un fenomeno relativamente nuovo, spesso si trascurano i potenziali pericoli.

Le persone temono di essere sostituite dalla tecnologia. La maggior parte di noi reagisce ai cambiamenti tecnologici nel migliore dei casi con disagio, nel peggiore con il panico e, come molte volte in passato, siamo preoccupati che questo nuovo set di tecnologie ci possa danneggiare, che l'AI porterà alla disoccupazione di massa o che diventerà sovrumana e sceglierà di distruggerci.

L'apprendimento automatico è strettamente legato al concetto d'intelligenza artificiale, se si forniscono sufficienti esempi ad un computer, questo può riconoscere i modelli sottostanti e incorporare informazioni per generare un algoritmo.

L'intelligenza si basa sulla capacità di apprendere e di adattarsi ai cambiamenti delle condizioni. Le **reti neurali** artificiali sono particolarmente efficaci e sono costituite da diversi strati di neuroni artificiali interconnessi, il loro uso è possibile vista la potenza di calcolo oggi disponibile.

Queste reti sono "addestrate" sulla base di dati, in modo che il risultato del calcolo possa essere confrontato con i valori conosciuti, se i due non concordano l'algoritmo si adatta di conseguenza per migliorare le sue prestazioni. **L'algoritmo di apprendimento** costruisce da solo tutte le strutture necessarie per classificare i dati in modo indipendente.

Gli scienziati hanno dovuto sviluppare nuovi metodi matematici perché le trasformazioni non lineari eseguite dalle reti neurali artificiali non sono molto trasparenti e di difficile interpretazione. Il problema delle "scatole nere", gli algoritmi opachi del computer, è una delle principali preoccupazioni dei ricercatori.

Ad esempio, alcuni scienziati hanno esaminato i risultati di due diversi algoritmi di classificazione delle immagini, entrambi hanno raggiunto un'accuratezza simile nel riconoscimento delle immagini nonostante i diversi metodi di base. Il primo algoritmo si basava su una rete neurale profonda, mentre il secondo utilizzava un altro metodo di apprendimento, quindi vi sono varie strade per lo stesso risultato, cosa che complica ulteriormente la ricerca.

Un'altra volta sono stati esaminati algoritmi per classificare i testi in base al loro contenuto. Anche in questo caso entrambi gli algoritmi hanno raggiunto una precisione simile, anche se i meccanismi di valutazione che hanno sviluppato erano molto diversi. Le indagini dimostrano quanto sia pericoloso affidarsi all'apprendimento di algoritmi che sono trattati come scatole nere. Un aspetto del problema è che l'analisi dei dati può rivelare correlazioni che non hanno causalità. In questo senso si può concordare con Elon Musk nella richiesta di un controllo dell'intelligenza artificiale.

Le reti neurali sono sempre più utilizzate per costruire programmi in grado di prevedere e classificare in una molteplicità di situazioni, tutto ciò ha prodotto modelli che possono migliorare la produttività e l'efficienza, tuttavia non sappiamo davvero come funzionano.

Perché la necessità di avere modelli spiegabili? Le Reti Neurali non sono infallibili, non sappiamo davvero i motivi delle scelte che fanno. I model-

li diventano sempre più complessi, il compito di produrne una versione interpretabile diventa sempre più difficile inoltre la relativa facilità di ingannare queste reti è preoccupante.
L'intelligenza artificiale non deve essere seguita ciecamente ma non possiamo dimenticare i vantaggi che può portare.
Le aziende si stanno muovendo rapidamente per applicare l'apprendimento automatico ai processi decisionali. Nuovi programmi vengono costantemente lanciati, impostando algoritmi complessi per lavorare su grandi insiemi di dati.
L'AI è utilizzata in molte decisioni con implicazioni commerciali e personali, come l'approvazione di prestiti nel settore bancario.
Le macchine di apprendimento sono state progettate per emulare il funzionamento del cervello umano. **Se i pregiudizi colpiscono l'intelligenza umana, che dire di quella artificiale? Le macchine sono di parte?** La risposta, ovviamente, è sì, per alcune ragioni di base. Gli algoritmi di apprendimento automatico sono inclini a incorporare i pregiudizi dei loro creatori umani. Gli algoritmi possono formalizzare parametri distorti; dove l'apprendimento automatico predice i risultati comportamentali, la dipendenza da criteri storici rafforzerà gli errori del passato.
La gravità di questi pregiudizi può essere amplificata da algoritmi che devono presupporre per funzionare che le cose continueranno in maniera simile al passato. Un altro fattore di base che genera distorsioni è costituito dall'incompletezza dei dati, ogni algoritmo opera interamente all'interno del mondo definito dai dati che sono stati utilizzati per calibrarlo. Le limitazioni nell'insieme dei dati influenzano i risultati, a volte in modo grave.
L'apprendimento automatico può rivelare preziose intuizioni in insiemi di dati complessi, ma le anomalie e gli errori dei dati possono portare a risultati fuorvianti.
Occorre quindi affrontare i pregiudizi contenuti negli algoritmi di apprendimento automatico; è importante sensibilizzare l'opinione pubblica su ciò che veramente preoccupa quando si tratta di Al: la manipolazione molto efficace e scalabile del comportamento umano e il suo possibile uso malevolo da parte delle aziende e dei governi. Naturalmente questo non è l'unico rischio tangibile che deriva dallo sviluppo delle tecnologie cognitive, ce ne sono molte altre potenziali.
Sulla manipolazione della popolazione di massa, il rischio è pressante e sottovalutato. Le aziende e i governi stanno ora raccogliendo quantità impressionanti di dati su di noi, in particolare attraverso i social network.
Quasi la metà del traffico online è falsa, generata per lo più da bot, cioè da software automatici studiati per dare l'impressione di essere navigatori umani.
Un grande sottoinsieme dell'Al - in particolare il ***"reinforcement learning"*** - riguarda lo sviluppo di algoritmi per risolvere problemi di ottimizzazione nel modo più efficiente possibile e raggiungere il pieno controllo dell'obiettivo, in questo caso, noi. Spostando le nostre vite verso il regno digitale, diventiamo vulnerabili agli algoritmi quindi anche a semplici modelli di manipolazione sociale, è però importante sottolineare che l'esistenza di questa minaccia non significa che tutti i dati algoritmici o che tutta la pubblicità mirata sia cattiva.
La gestione algoritmica dell'informazione ha un enorme potenziale per aiutarci, per consentire agli individui di realizzare il loro potenziale e per aiutare la società a gestire meglio se stessa, il problema non è l'Al in sé, il problema è il controllo.
Gli algoritmi di gestione delle informazioni non dovrebbero essere una forza misteriosa imposta per servire fini che vanno contro il nostro interesse ma dovrebbero invece essere uno strumento da usare per i nostri scopi, ad esempio, per l'educazione e l'intrattenimento personale.
L'intelligenza artificiale sta diventando sempre più una cosa scontata nella nostra vita. **Che cosa ci aspettiamo dalle macchine intelligenti, in quale modo la loro presenza nella nostra vita quotidiana cambia la nostra immagine e la nostra interazione con le altre persone? E quali libertà vogliamo dare alle macchine?**
Particolare attenzione etica è rivolta in questo momento a quei sistemi tecnici destinati a compensare o sostituire le capacità umane, questa è la novità della quarta rivoluzione industriale.
Per molto tempo l'uomo è stato considerato l'unico essere che ha avuto la libertà per quanto riguarda la sua volontà, le sue decisioni e le sue azioni. Questo assunto molto semplificato è ora rimesso in discussione dalle conquiste dell'intelligenza artificiale. Alcune forme di moralità e razionalità

sembrano essere soddisfatte ancora meglio dalle macchine che dagli esseri umani, quelle dotate d'intelligenza artificiale possono "decidere" alcuni compiti in modo più rapido, preciso e incorruttibile.

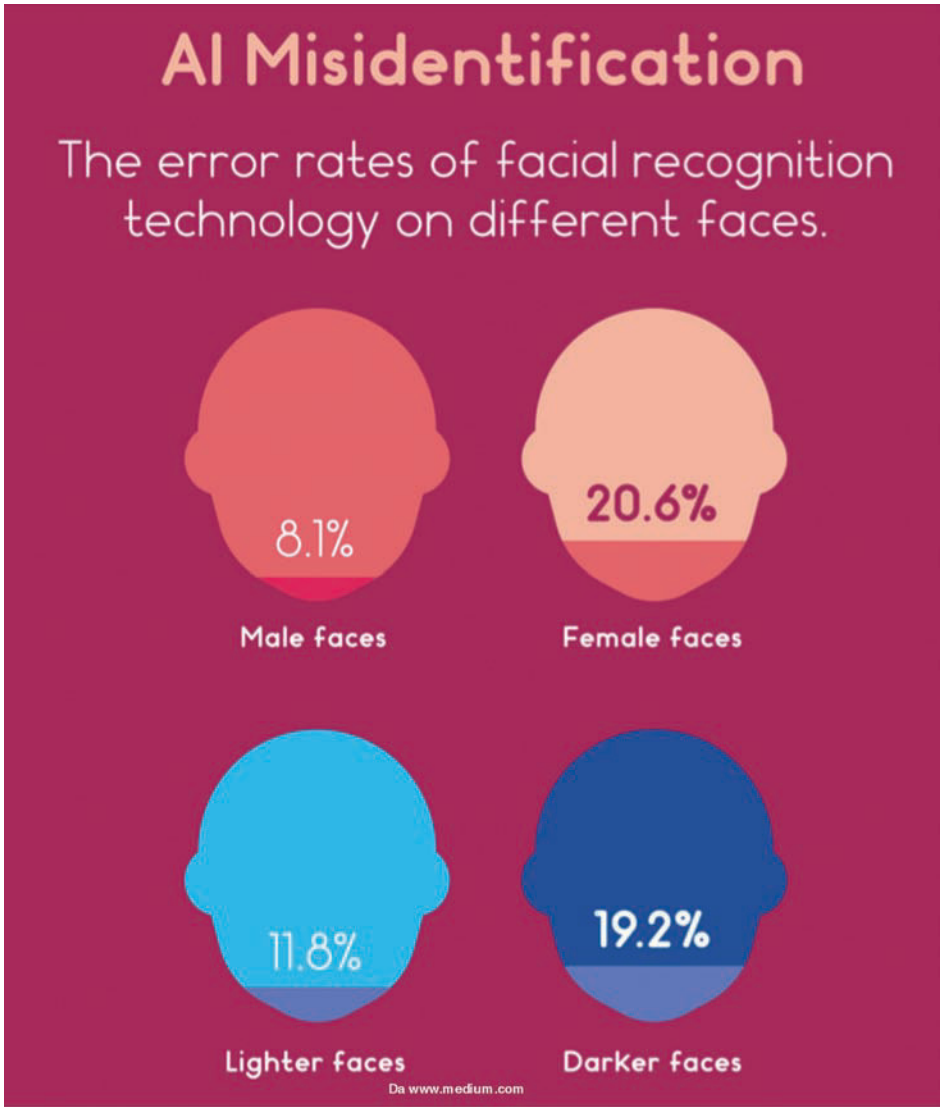
In questo periodo si discute criticamente se le persone rischino di perdere la supremazia in un numero sempre maggiore di campi d'azione perché le macchine sono semplicemente migliori nelle loro prestazioni e stanno fissando sempre più gli standard.

Questioni di giustizia sorgono, ad esempio, nel campo del lavoro, dove i robot e l'intelligenza artificiale possono fare sempre più e costringere le persone ad abbandonare alcune attività o perdere delle competenze, ad esempio le capacità diagnostiche dei medici.

Il fatto che la tecnologia sostituisca il lavoro umano non è un problema nuovo. In questo caso, tuttavia, la situazione è che i nuovi posti di lavoro creati richiedono qualifiche completamente nuove e più elevate, non accessibili a parte dei lavoratori.

Si ritiene comunque possibile per le imprese realizzare utili e al tempo stesso seguire un **approccio etico** e che possano esistere società sostenibili che siano in grado di operare con successo, conducendo i consumatori in un percorso virtuoso, usando il proprio potere per qualcosa che vada oltre la semplice realizzazione di profitti.

Nel caso della ricerca sono state stabilite adeguate protezioni solo dopo che si sono verificate gravi trasgressioni etiche. Si potrebbero creare enti di supervisione che valutino la ricerca sull'AI, tali comitati dovrebbero essere composti da un mix di scienziati e non, con il compito di identificare e valutare i rischi etici delle nuove forme dell'AI.



C'è ancora molto da imparare sui rischi potenziali che le organizzazioni, gli individui e la società devono affrontare, sul giusto equilibrio tra innovazione e rischio e sull'introduzione di controlli per gestire l'inimmaginabile.

Un altro imperativo è di impegnarsi in un serio dibattito sull'etica dell'applicazione dell'AI e su dove tracciare linee che ne limitino l'uso. Anche l'azione collettiva, che potrebbe comportare un dibattito a livello industriale sull'autodisciplina e l'impegno con le autorità di regolamentazione, è destinata a crescere d'importanza. Le organizzazioni che s'impegneranno in questo settore saranno meglio posizionate per servire efficacemente la società quindi per evitare le difficoltà etiche, commerciali e normative.