

La questione delle macchine

Tom Standage, *The Economist*, Regno Unito. Foto di James Ball

L'intelligenza artificiale si perfeziona sempre più rapidamente. Qualcuno teme che diventi così efficiente e autonoma da poter fare a meno dell'umanità. Ma la storia dell'innovazione può aiutarci a non aver paura del futuro

La paura che le macchine tolgano il lavoro a tutti sconvolgendo la società e favorendo solo pochi privilegiati sembra uguale in ogni epoca. Due secoli fa, in piena rivoluzione industriale, scatenò furiosi dibattiti nel Regno Unito. All'epoca, in realtà, non si parlava di "rivoluzione industriale" ma di "questione delle macchine". Posto per la prima volta dall'economista David Ricardo nel 1821, il problema riguardava "l'influenza delle macchine sugli interessi delle varie classi sociali" e, in particolare, "l'opinione diffusa nella classe operaia che il loro uso spesso danneggiasse i suoi interessi". Nel 1839 Thomas Carlyle inveiva contro il "demone della meccanica" che con la sua forza dirompente stava "sconvolgendo la vita di intere moltitudini di lavoratori".

Oggi la questione delle macchine si sta riproponendo, ma in una nuova forma. Gli esperti di tecnologia, gli economisti e i filosofi dibattono sulle possibili conseguenze dell'intelligenza artificiale, una tecnologia in rapido sviluppo che permette alle macchine di svolgere compiti finora riservati agli esseri umani. La sua influenza potrebbe essere enorme e diventare un pericolo anche per persone il cui lavoro sembrava impossibile da automatizzare: dai radiologi ai praticanti degli studi legali. Da una ricerca molto citata, condotta da Carl Benedikt Frey e Michael Osborne dell'università di Oxford e pubblicata nel 2013, è emerso che

presto negli Stati Uniti il 47 per cento dei posti di lavoro correrà il rischio di essere "occupato dai computer". La Bank of America Merrill Lynch prevede che entro il 2025 il valore annuale delle trasformazioni imposte dall'intelligenza artificiale potrebbe oscillare tra i 14mila e i 33mila miliardi di dollari, compresi novemila miliardi legati alla riduzione dei salari. Il McKinsey global institute sostiene che l'intelligenza artificiale sta contribuendo a una trasformazione della società "dieci volte più veloce e trecento volte più grande, e quindi con conseguenze tremila volte maggiori", rispetto alla rivoluzione industriale. Come due secoli fa, quindi, molti temono oggi che le macchine produrranno milioni di disoccupati, facendo crescere le disuguaglianze e i disordini.

Ma dove alcuni vedono pericoli, altri individuano opportunità. Gli investitori sono molto interessati al settore. I colossi della tecnologia stanno comprando aziende di intelligenza artificiale e si fanno concorrenza per assicurarsi i migliori ricercatori. Secondo Quid, una società di analisi dei dati,

Da sapere Storia illustrata

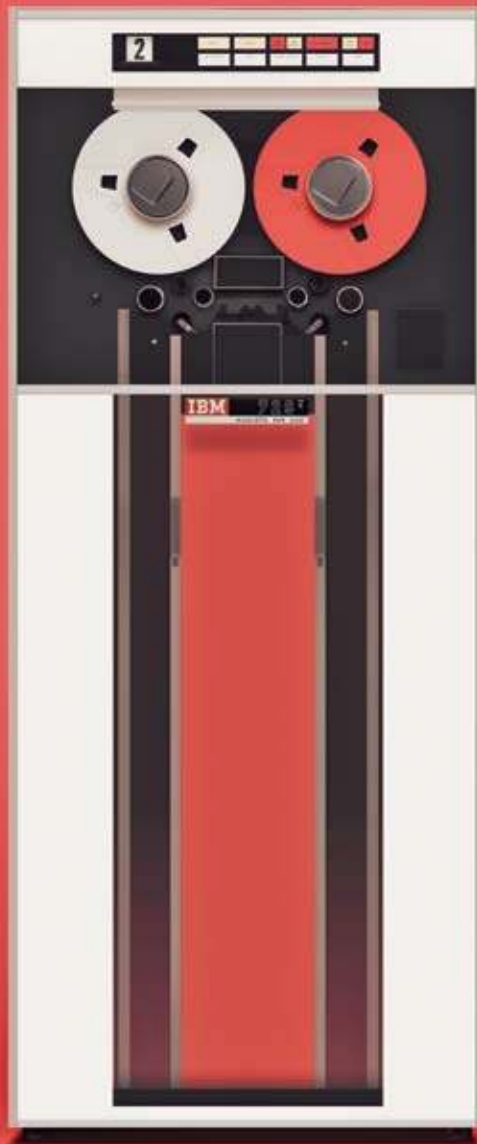
◆ Le foto di questo articolo sono tratte da *Guide to computing*, un lavoro sulla storia dei computer realizzato dal fotografo britannico James Ball insieme allo studio grafico Ink. Ball ha ritoccato delle foto d'epoca per mostrare i computer nel loro aspetto originale.

nel 2015 sono stati investiti 8,5 miliardi di dollari nelle aziende che si occupano di intelligenza artificiale, quasi il quadruplo del 2010. Google, Facebook, Ibm, Amazon e Microsoft stanno cercando di creare ecosistemi intorno ai servizi di intelligenza artificiale forniti attraverso il *cloud computing*. "Questa tecnologia sarà applicata a quasi tutti i settori che usano un qualche tipo di dati, dalla genetica alle immagini e al linguaggio", dice Richard Socher, il fondatore della MetaMind, un'azienda comprata di recente dal colosso del *cloud computing* Salesforce. "L'intelligenza artificiale sarà ovunque".

Proposta di ricerca

Ma come ha fatto l'intelligenza artificiale, che all'inizio sembrava così deludente, a diventare all'improvviso il settore più avanzato della tecnologia? Il termine fu coniato nell'ambito di una proposta di ricerca presentata nel 1956. Nel documento si suggeriva che sarebbe stato possibile far risolvere alle macchine "il tipo di problemi che oggi sono riservati agli esseri umani. Basta che un gruppo selezionato di scienziati lavori insieme per un'estate". Quest'ipotesi si sarebbe rivelata eccessivamente ottimistica. La riabilitazione dell'intelligenza artificiale si può far risalire al 2012, in particolare al concorso online ImageNet Challenge.

ImageNet è un database online che raccoglie milioni di immagini, tutte taggate manualmente. Per ogni tag, cioè per ogni termine etichetta, come "palloncino" o

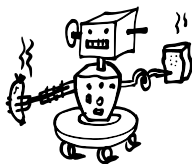


IBM 729 Magnetic Tape Unit

“fragola”, ImageNet contiene diverse centinaia di immagini. L’ImageNet Challenge invita ogni anno le persone che lavorano in questo campo a misurare i progressi fatti nell’insegnare ai computer a riconoscere e a taggare le immagini automaticamente. I sistemi sono prima collaudati usando una serie di foto già taggate correttamente, poi devono etichettarne altre mai viste. Durante un workshop che si svolge dopo il concorso, i vincitori condividono e discutono le loro tecniche. Nel 2010 il sistema che vinse era in grado di taggare correttamente il 72 per cento delle immagini (per gli esseri umani la media era del 95 per cento). Nel 2012 un’equipe guidata da Geoffrey Hinton, dell’università di Toronto, ha raggiunto un livello di precisione dell’85 per cento grazie a una nuova tecnica chiamata *deep learning* (apprendimento profondo). Quest’innovazione ha prodotto rapidi miglioramenti e un’accuratezza del 96 per cento nell’ImageNet Challenge del 2015, l’edizione in cui le macchine hanno superato per la prima volta gli esseri umani. I risultati del 2012 sono stati giustamente considerati una svolta decisiva, ma erano stati ottenuti “combinando cose che già esistevano”, dice Yoshua Bengio, un informatico dell’università di Montréal che, con Hinton e pochi altri, è ritenuto uno dei pionieri del *deep learning*. In pratica questa tecnica usa enormi quantità di dati e di capacità di calcolo per sviluppare una vecchia idea che risale all’alba dell’intelligenza artificiale: quella delle cosiddette reti neurali artificiali (*artificial neural networks*), sistemi che s’ispirano alle reti di cellule cerebrali.

Nel cervello umano ogni neurone può essere stimolato e attivato da altri neuroni e a sua volta stimolarne e attivarne altri. Una rete neurale artificiale semplice ha uno strato di neuroni d’ingresso, attraverso il quale i dati possono essere inseriti, e uno strato di uscita, che produce i risultati. Poi ci possono essere un paio di strati intermedi dove sono elaborate le informazioni. Ogni neurone ha una serie di “pesi” e una “funzione di attivazione” che controlla la produzione dei risultati. Per addestrare una rete neurale è necessario correggere i pesi dei neuroni in modo che i dati in ingresso producano i dati in uscita desiderati. All’inizio degli anni novanta le reti neurali artificiali stavano cominciando a dare risultati, per esempio nel riconoscimento dei numeri scritti a mano. Ma i tentativi di usarle per compiti più complessi non hanno funzionato.

Negli ultimi dieci anni alcune nuove tecniche e una semplice modifica della funzione di attivazione hanno reso possibile l’addestramento di reti di *deep learning*. Allo stesso tempo la nascita di internet ha reso disponibili per l’addestramento miliardi di documenti, immagini e filmati. Tutto questo richiede una grande capacità di calcolo, che è stata raggiunta intorno al 2009 quando diverse équipes di ricerca si sono rese conto che i processori grafici, i microchip usati nei pc e nelle console per videogiochi per generare la grafica, erano



adatti anche a gestire gli algoritmi del *deep learning*. Un gruppo di ricercatori dell’università di Stanford, diretta da Andrew Ng, che poi sarebbe passato a Google e ora lavora per il colosso di internet cinese Baidu, ha scoperto che i processori grafici potevano accelerare il suo sistema di apprendimento fino a cento volte. All’improvviso per addestrare una rete neurale a quattro strati, che fino a quel momento aveva richiesto settimane, bastava meno di un giorno. È una piacevole coincidenza, dice Jen-Hsun Huang, l’amministratore delegato della Nvidia, un produttore di processori grafici, che gli stessi microchip usati per creare mondi immaginari nei videogiochi possano aiutare i computer a capire il mondo reale con il *deep learning*.

I risultati di ImageNet hanno dimostrato chiaramente quello che è in grado di fare il *deep learning*. All’improvviso tutti hanno cominciato a prestare attenzione alle nuove scoperte. Da allora i sistemi di *deep learning* sono diventati più potenti: ormai esistono reti a venti o trenta strati, i ricercatori della Microsoft ne hanno costruita una che

ne ha 152. Le reti più profonde raggiungono un più alto livello di astrazione e producono risultati migliori, dimostrando di poter risolvere una vasta gamma di problemi.

“Quello che ha entusiasmato chi lavora nel settore è che un’unica tecnica di apprendimento è applicabile a tanti campi”, dice John Giannandrea, direttore dell’unità di intelligenza artificiale di Google e oggi responsabile anche del motore di ricerca. Google sta usando il *deep learning* per migliorare la qualità dei suoi risultati, per capire i comandi vocali degli smartphone, facilitare la ricerca di immagini, suggerire risposte automatiche alle email, migliorare il servizio di traduzione delle pagine web e aiutare le automobili che si guidano da sole a capire cosa succede intorno a loro.

Una serie di esempi

Esistono diversi tipi di *deep learning*. Quello più usato è il *supervised learning* (apprendimento guidato), una tecnica che può essere usata per addestrare un sistema con l’aiuto di una serie di esempi. Per filtrare lo spam della posta elettronica, si può usare un enorme database di messaggi etichettati come “spam” e “non spam”. Un sistema di *deep learning* può essere addestrato passando in rassegna varie volte gli esempi e correggendo i pesi all’interno della rete neurale per aumentare la sua precisione nel valutare lo spam. Il grande merito di questo metodo è che non c’è bisogno di una lista di regole compilata da un esperto umano né di un programmatore che le codifichi, perché il sistema impara direttamente dai dati etichettati.

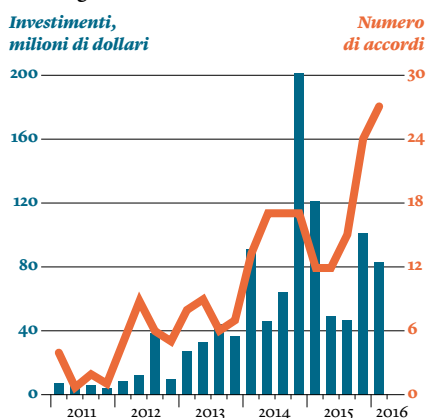
I sistemi di questo tipo sono usati anche per classificare immagini, riconoscere il parlato, individuare i tentativi di frode con le carte di credito, i programmi dannosi e le pubblicità, tutto quello per cui esiste una serie di risposte note in base a casi precedenti. Quando carichiamo una fotografia, Facebook è in grado di riconoscere e taggare i nostri amici e parenti, e di recente ha introdotto un sistema capace di descrivere il contenuto delle foto per i non vedenti (“due persone che sorridono, occhiali da sole, esterno, acqua”). Esiste un enorme serbatoio di dati a cui si può applicare l’apprendimento guidato, dice Ng.

Un’altra tecnica, quella dell’*unsupervised learning* (apprendimento non guidato), addestra la rete presentandole un enorme numero di esempi, ma senza dirle cosa cercare. La rete impara da sola a riconoscere certi tratti e a raggruppare esempi simili, rivelando così collegamenti e schemi nascosti nei dati. L’apprendimento non guida-

Da sapere

Interesse crescente

Finanziamenti alle aziende che si occupano d’intelligenza artificiale. Fonte: *The Economist*





Endim 2000, computer sviluppato nell'ex Germania Est

to può essere usato per individuare cose che non sappiamo come sono: per esempio, per monitorare gli schemi di traffico in rete alla ricerca di anomalie che potrebbero indicare un attacco informatico oppure esaminare un gran numero di richieste di risarcimento ricevute dalle assicurazioni per individuare nuovi tipi di frodi. Un esempio famoso è quello che è capitato a Ng quando lavorava a Google nel 2011. Dirigeva un progetto chiamato Google Brain, in cui un gigantesco sistema di apprendimento non guidato doveva cercare schemi ricorrenti in migliaia di video di YouTube non taggati. Un giorno, ricorda Ng, uno dei suoi collaboratori gli chiese di dare un'occhiata al suo computer. Sullo schermo c'era un muso peloso ottenuto filtrando migliaia di esempi. Il sistema aveva scoperto i gatti.

Poi c'è il *reinforcement learning* (apprendimento rinforzato), una via di mezzo tra i due precedenti. La rete neurale è addestra-

ta a interagire con un ambiente e riceve solo ogni tanto un feedback sotto forma di premio. In pratica l'addestramento consiste nel correggere i pesi della rete per cercare una strategia che generi regolarmente ricompense maggiori. Un'azienda specializzata in questo settore è la DeepMind, finita su tutti i giornali quando a Seoul il suo sistema AlphaGo ha battuto 4 a 1 Lee Sedol, uno dei migliori giocatori del mondo di *go* (un gioco tradizionale cinese). AlphaGo è un sistema di apprendimento rinforzato con alcune caratteristiche insolite. È costituito da diversi moduli interconnessi, tra cui due reti neurali profonde, ognuno dei quali è specializzato in un campo diverso, come i moduli del cervello umano. Una delle reti è stata addestrata ad analizzare milioni di partite per suggerire possibili mosse vincenti, che poi sono valutate dall'altra con l'aiuto di una tecnica basata sul campionamento casuale. I ricercatori del settore di-

scutono da decenni quale sia il metodo migliore, ma AlphaGo li usa entrambi. "È un sistema ibrido perché pensiamo che avremo bisogno di qualcosa di più del *deep learning* per imitare l'intelligenza", dice Demis Hassabis, della DeepMind.

Hassabis e altri ricercatori stanno già pensando al passo successivo, che chiamano *transfer learning* (apprendimento trasferito). Questa tecnica consentirebbe a un sistema di apprendimento rinforzato di costruire su quanto ha appreso in precedenza invece di essere addestrato daccapo ogni volta. Gli esseri umani lo fanno senza nessuno sforzo, osserva Hassabis. Giannandrea ricorda che quando aveva quattro anni sua figlia era in grado di capire che un biciclo era un tipo di bicicletta anche se non ne aveva mai visto uno. "I computer non sono capaci di farlo".

La maggior parte delle persone scoprirà i progressi dell'intelligenza artificiale quan-

In copertina

do miglioreranno i servizi online che già usa. I motori di ricerca produrranno risultati più rilevanti, i consigli saranno più precisi. Tra qualche anno i sistemi intelligenti saranno quasi ovunque, prevede Hassabis. La tecnologia permetterà ai computer di conversare e fare previsioni, non solo di funzionare in base a icone e menù. I computer diventeranno accessibili anche a persone che non sanno né leggere né scrivere, aggiunge Bengio.

Ma quando le macchine raggiungeranno una certa soglia e saranno in grado di svolgere compiti fino a quel momento riservati agli esseri umani, potrebbero verificarsi dei cambiamenti improvvisi. Le auto che si guidano da sole stanno migliorando rapidamente e presto potrebbero sostituire i taxi, almeno nelle zone centrali delle città. I droni per le consegne potrebbero fare concorrenza ai corrieri. Con i nuovi sistemi di riconoscimento e i progressi della robotica le macchine potrebbero riempire gli scaffali dei supermercati e spostare le merci nei magazzini.

Qualcuno è preoccupato, perché teme che l'intelligenza artificiale estenderà la gamma delle attività automatizzate e toglierà lavoro alle persone, come fecero le macchine a vapore duecento anni fa. "La scoperta della forza vapore ha accelerato un processo già in atto, ma lo ha reso troppo veloce", osservava all'epoca il poeta inglese Robert Southey. Temeva che quella scoperta fosse arrivata "prima che sapessimo come usarla nel modo giusto". E oggi molte persone pensano la stessa cosa dell'intelligenza artificiale.

Fabbriche di graffette

Non mancano neanche le preoccupazioni di carattere etico. Il *paperclip maximizer* (ottimizzatore di graffette) è un esperimento mentale proposto da Nick Bostrom, filosofo dell'università britannica di Oxford. Immaginate un'intelligenza artificiale programmata per fabbricare graffette che decide di produrne il maggior numero possibile, dice Bostrom. Dedica tutte le sue energie alla fabbricazione e resiste a ogni tentativo di distrarla dal compito. Alla fine "comincerà a trasformare la Terra e porzioni sempre più grandi dello spazio in fabbriche di graffette". Quest'ipotesi apparentemente demenziale serve a dimostrare che le intelligenze artificiali non devono necessariamente avere motivazioni o psicologie simili a quelle umane. Possono evitare alcuni tipi di errori o pregiudizi, ma commetterne altri, come fissarsi sulle graffette. E anche se all'inizio il loro obiettivo può sem-



Hdr75, computer sviluppato nell'ex Germania Est, a Dresda

brare innocuo, se fossero in grado di progettare nuove macchine e migliorarsi continuamente, potrebbero diventare pericolose. Perfino una "superintelligenza controllata", che gira in un computer isolato, potrebbe convincere i suoi controllori umani a liberarla. L'intelligenza artificiale avanzata non è una tecnologia qualunque, sostiene Bostrom, costituisce una minaccia esistenziale per l'umanità.

L'idea delle macchine che si ribellano ai loro creatori non è una novità, risale a prima del *Frankenstein* di Mary Shelley (1818). Non è nuova neanche l'ipotesi, avanzata nel 1965, di un'intelligenza artificiale che subisce "un'esplosione d'intelligenza" continuando a migliorarsi. Ma gli ultimi progressi del settore hanno suscitato nuovi timori, e Bostrom è diventato il più noto sostenitore dei pericoli legati all'intelligenza artificiale avanzata o, come preferisce chiamarla lui, la "superintelligenza".

L'interesse di Bostrom per l'intelligenza artificiale nasce dall'analisi delle possibili minacce esistenziali per l'umanità. A differenza di una pandemia, dell'impatto con un

asteroide o dell'eruzione di un grande vulcano, l'emergere di una superintelligenza è qualcosa che il genere umano può controllare. Il libro di Bostrom, *Superintelligence*, ha spinto Elon Musk, il fondatore della Tesla, a dichiarare che l'intelligenza artificiale è "potenzialmente più pericolosa del nucleare". Timori sulla sua pericolosità sono stati espressi anche dal fisico Stephen Hawking e da Martin Rees, ex presidente della Royal society, la più prestigiosa istituzione scientifica britannica. Queste tre persone, e molte altre che lavorano nel settore, hanno firmato una lettera aperta in cui chiedono ai ricercatori di garantire che i sistemi d'intelligenza artificiale siano "robusti e benefici". Quasi tutti concordano nel dire che l'intelligenza artificiale deve essere sviluppata in modo da portare benefici all'umanità, ma c'è molta meno sintonia su come raggiungere l'obiettivo.

Musk pensa che la chiave sia la trasparenza. Nel dicembre del 2015 è stato uno dei fondatori di OpenAI, un istituto di ricerca con fondi per più di un miliardo di dollari che condurrà studi sull'intelligenza artifi-

ciale e renderà pubblici i risultati. “Pensiamo che l’intelligenza artificiale avrà un enorme effetto sul futuro della civiltà, e vorremmo fare in modo che sia un buon futuro”, dice Musk. A questo scopo l’intelligenza artificiale dovrebbe essere decentralizzata il più possibile. I computer devianti della fantascienza, come l’Hal 9000 di *2001 Odissea nello spazio* e lo Skynet della serie *Terminator*, sono grandi macchine centralizzate, ed è questo che le rende pericolose quando diventano cattive. Una migliore distribuzione garantirebbe i servizi a tutti, e se qualcosa andasse storto le conseguenze sarebbero meno gravi, dice Musk.

Non tutti sono d’accordo con lui. Qualcuno sostiene che la vera preoccupazione di Musk sia la concentrazione del mercato, per esempio il monopolio di Facebook o quello di Google. Musk liquida queste accuse definendole “meschine”. Per il momento Google, Facebook e le altre grandi aziende tecnologiche stanno sviluppando buona parte dei loro programmi d’intelligenza artificiale da sole e in ogni caso i risultati delle loro ricerche sono a disposizione di tutti. Bostrom, da parte sua, non è convinto che mettere l’intelligenza artificiale a disposizione di tutti sia necessariamente una cosa positiva. In un recente saggio ha osservato che l’esistenza di diverse intelligenze artificiali “non garantisce che agiranno nell’interesse degli esseri umani né che rimarranno sotto il loro controllo”. Il filosofo aggiunge che la proliferazione potrebbe rendere la tecnologia più difficile da controllare e da regolamentare.

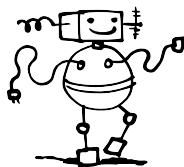
Il timore che l’intelligenza artificiale diventi maligna non è condiviso da tutti quelli che sono all’avanguardia della ricerca nel settore. “Gran parte dell’allarmismo è alimentato da persone che non lavorano direttamente in questo campo, quindi immaginano scenari da fantascienza”, dice Hassabis della DeepMind. “Non credo sia utile parlarne in termini emotivi, si provoca solo isterismo”. Hassabis considera l’idea di un computer impazzito che produce miliardi di graffette “poco realistica”, ma pensa che Bostrom faccia bene a sottolineare l’importanza degli obiettivi dell’intelligenza artificiale. La ricerca si deve porre il problema di come imporre obiettivi e valori giusti all’intelligenza artificiale e di come garantire che restino invariati nel tempo, dice Hassabis (la DeepMind ha appena pubblicato un articolo con il Future of humanity institute di Bostrom sulla necessità di aggiungere ai sistemi d’intelligenza artificiale una sorta di “interruttore per poterli spegnere”). Anche al convegno che si è tenuto nel 2009 ad

Asilomar, in California, gli esperti del settore sono arrivati alla conclusione che la sicurezza dell’intelligenza artificiale dev’essere oggetto di ricerca, ma che per ora non è un problema. La scelta della sede del convegno è stata significativa, perché nel 1975 i biologi si erano riuniti proprio lì per stabilire delle linee guida volontarie sull’ingegneria genetica.

Scenari da fantascienza

Diversi ricercatori del settore condividono le preoccupazioni di Bostrom, anche se lui afferma di voler solo mettere in evidenza i potenziali rischi posti dall’intelligenza artificiale. Non dice che sia pericolosa in questo momento. Andrew Ng di Baidu si rivolge a Musk quando sostiene che preoccuparsi delle superintelligenze ora “è come preoccuparsi della sovrappopolazione di Marte prima ancora di metterci piede”. L’intelligenza artificiale spaventa la gente, osserva Marc Andreessen, uno dei pionieri di internet, perché combina due paure profonde: quella che le macchine portino via tutti i posti di lavoro e quella che un giorno un’intelligenza artificiale “si svegli” e faccia qualcosa di testa sua come la creatura di Frankenstein. Entrambe queste paure “saltano fuori continuamente”. E decenni di fantascienza le hanno rese più concrete di cose come il cambiamento climatico, che è un pericolo molto maggiore.

I ricercatori del settore indicano vari motivi per cui certi timori sono esagerati, almeno per il momento. Prima di tutto, essere intelligenti non è la stessa cosa che essere senzienti o coscienti, dice Ng, anche se i tre concetti sono spesso confusi tra



loro. L’ipotesi che le macchine “un giorno si sveglino e cambino idea su quello che devono fare” non è realistica, aggiunge Francesca Rossi, che si occupa di etica dell’intelligenza artificiale per l’Ibm. In secondo luogo, l’eventualità di “un’esplosione di intelligenza” è ritenuta improbabile, perché significherebbe che, diventando più intelligente, un’intelligenza artificiale sia in grado di creare una nuova versione di se stessa in meno tempo della precedente. Invece sappiamo che quasi tutti i processi informatici, anche quelli più semplici, quando diventano più complessi richiedono più tempo.

Inoltre, anche se le macchine possono apprendere dalle loro esperienze passate, non imparano continuamente. Un’automobile che si guida da sola, per esempio, non si riprogramma a ogni percorso. I sistemi di deep learning, invece, hanno una fase di addestramento in cui i parametri della rete neurale sono regolati per costruire un modello in grado di svolgere un particolare compito: si

tratta di un processo che può richiedere giorni. Il modello ottenuto è poi inserito in un sistema attivo, dove può funzionare usando meno potenza di calcolo, permettendo così il suo impiego nelle auto, nei droni, nelle applicazioni e in altri prodotti. Ma quelle macchine e quei droni non imparano nel vuoto. I dati che raccolgono quando sono in missione sono rimandati indietro e servono a perfezionare il modello, che poi sarà reinstallato. Quindi un singolo sistema non può imparare un comportamento negativo in un particolare ambiente e “diventare cattivo”, perché in quel momento non sta davvero imparando.

Tra i timori suscitati dall’intelligenza artificiale c’è il rischio che l’urgenza di stabilire norme etiche su questa tecnologia sia sottovalutata. I sistemi di riconoscimento facciale basati sul deep learning, per esempio, potrebbero rendere gli strumenti di sorveglianza molto più potenti. FaceNet di Google stabilisce con il 99,6 per cento di precisione se in due fotografie compare la stessa persona (gli esseri umani ci riescono al 98 per cento). DeepFace di Facebook è quasi allo stesso livello. Quando qualche tempo fa il social network ha lanciato Moments, un’applicazione che raggruppa automaticamente le foto di una stessa persona, in Europa ha dovuto disabilitare alcune funzioni per non violare la legge sulla privacy irlandese.

Intanto in Russia sono scattate le proteste contro FindFace, un’applicazione che

Da sapere Meglio degli esseri umani

Accuratezza degli algoritmi di riconoscimento facciale, percentuale

	%
Baidu	99,77
Tencent BestImage	99,65
Google FaceNet	99,63
DeepID3 (CUKH)	99,53
Megvii Face++	99,50
Livello umano	97,64
Facebook DeepFace	97,35

Fonte: *The Economist*

permette di scattare foto di estranei e scoprire chi sono attraverso le immagini usate nei loro profili sui social network. I creatori dell'applicazione dicono che è solo un modo per contattare persone viste per strada o in un bar. Ma la polizia ha cominciato a usarla per identificare sospettati e testimoni. Il rischio che si corre è chiaro: sarebbe la fine dell'anonimato. Le immagini di una folla riprese a centinaia di metri di distanza potrebbero essere usate per sapere chi ha partecipato a una manifestazione di protesta, anche a distanza di anni. In realtà il deep learning ha reso impossibile partecipare a un evento pubblico senza essere riconosciuti, a meno che non s'indossi una maschera (un'azienda giapponese ha lanciato PrivacyVisor, una serie di buffi occhiali che servono a impedire il riconoscimento facciale).

Il deep learning, con la sua capacità di riconoscere schemi ricorrenti e individuare gruppi di esempi simili, può sicuramente servire a combattere la criminalità, ma può anche permettere ai governi autoritari di spiare i cittadini. Le autorità cinesi stanno analizzando i profili delle persone sui social network per individuare possibili dissidenti, dice Patrick Lin, uno specialista di etica dell'intelligenza artificiale dell'università di Stanford. Negli Stati Uniti, invece, la polizia di Fresno, in California, sta testando Beware, un sistema che valuta quanto può essere pericolosa una persona sospetta in base all'analisi degli archivi della polizia, dei registri immobiliari e dei post sui social network. Un altro sistema, Compas, aiuta i giudici a stabilire le pene, calcolando le probabilità che l'imputato commetta di nuovo lo stesso reato. Questi sistemi che prima o poi faranno uso del deep learning, invalidano "il concetto stesso di giusto processo", dice Lin.

Un'altra preoccupazione è che quando i sistemi di apprendimento automatico saranno inseriti in un numero sempre maggiore di attività commerciali, potrebbero involontariamente diventare discriminatori nei confronti di alcuni gruppi di persone. Un esempio famoso è quello di Google, che ha dovuto scusarsi quando il sistema di identificazione automatica della sua applicazione Photos ha cominciato a etichettare i neri come "gorilla". Anche Compas è stato accusato di discriminare i neri.

L'intelligenza artificiale "sta toccando la vita delle persone comuni, perciò è importante che non inglobi pregiudizi", dice Richard Socher di MetaMind. Nessuno decide di creare un sistema razzista, aggiunge, ma "se il sistema è addestrato con dati

scorretti farà previsioni scorrette". Ormai non si tratta più solo di calcoli, ma anche di decisioni di tipo etico, dice Lin, affidate a "quelle che in pratica sono solo scatole nere". Per fortuna ci sono vari modi di guardare dentro quelle scatole nere e capire come arrivano a certe conclusioni. Per esempio, si può chiedere a una rete neurale che elabora immagini di evidenziare le zone di una certa immagine che hanno influito di più sulla sua decisione. Molti ricercatori, inoltre, stanno lavorando sulle varianti di una tecnica chiamata "estrazione di regole", che consente alle reti di spiegare come hanno ragionato. Il campo in cui questo problema è stato più studiato è quello delle auto che si guidano da sole.

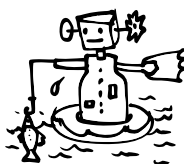
Se un sistema è addestrato con dati scorretti farà previsioni scorrette

Questi veicoli sollevano altri tipi di problemi etici, soprattutto per quanto riguarda il comportamento nei casi d'emergenza. Per esempio, un'auto dovrebbe rischiare di fare male ai suoi passeggeri per evitare d'investire un bambino che sbuca all'improvviso per strada? Non sono più questioni teoriche. Questioni come la responsabilità in caso di incidente, le verifiche necessarie prima di mettere in circolazione queste auto e le regole da stabilire vanno discusse ora, dice Hassabis. Ng affronta la questione da una prospettiva diversa, sostenendo che i ricercatori hanno l'obbligo morale di costruire le auto con il pilota automatico al più presto per salvare vite umane: la maggior parte degli incidenti stradali mortali è dovuta a un errore umano. Ma anche se le auto che si guidano da sole sono più sicure, osserva Daniel Suskind, un economista dell'università di Oxford, dovremo prima cambiare la mentalità delle persone. La gente sembra ritenere tollerabili gli incidenti mortali causati dagli esseri umani, ma dalle macchine pretendono molto di più. "Da loro ci aspettiamo la perfezione, dagli esseri umani che svolgono gli stessi compiti no", dice.

Molte persone sono preoccupate anche dalle applicazioni militari dell'intelligenza artificiale, in particolare dalla possibilità che siano costruite armi autonome in grado di agire senza alcun intervento umano. Yoshua Bengio, dell'università di Montréal, vorrebbe un "divieto assoluto" per i

militari di usare l'intelligenza artificiale. Le decisioni che implicano la vita o la morte di qualcuno dovrebbero essere prese da esseri umani, non da macchine, anche perché queste non possono assumersene la responsabilità. Hassabis è d'accordo. Quando Google ha comprato la sua azienda, la DeepMind, Hassabis ha insistito per avere la garanzia che la tecnologia non sarebbe stata usata a scopi militari. Lui e Bengio hanno firmato una lettera aperta in cui chiedono il divieto di costruire "armi offensive autonome". Ronald Arkin, del Georgia Institute of Technology, invece sostiene che i robot militari basati sull'intelligenza artificiale potrebbero essere moralmente superiori ai soldati: non violenterebbero, non saccheggerebbero e non prenderebbero decisioni sbagliate nei momenti di tensione.

Un'altra idea di Hassabis è stata quella di istituire un comitato etico, composto anche da alcuni osservatori indipendenti (la DeepMind, però, è stata criticata perché si è rifiutata di rendere noti i nomi dei commissari). Anche se le aziende che si occupano di intelligenza artificiale non sono d'accordo con gli allarmisti, devono ammettere che esistono alcune cose di cui è giusto preoccuparsi, e stabilire delle regole prima che gli siano imposte dall'esterno. È improbabile, però, che l'intelligenza artificiale finisca per avere un ente di vigilanza, come quello per l'aviazione civile, perché può essere applicata in molti campi. Sem-



bra più probabile che richieda un aggiornamento delle leggi esistenti. Le regole più famose che governano il comportamento dei sistemi d'intelligenza artificiale sono le "tre leggi della robotica"

di Isaac Asimov. Quello che rendeva interessanti i suoi racconti era che i robot si comportavano in modo inaspettato, perché in pratica le leggi non funzionavano. Quindi bisognerà sbrigarci a trovarne di nuove e più efficaci.

Cambiamenti positivi

È bene ricordare, però, che nell'ottocento la questione delle macchine alla fine si risolse da sola. Nonostante i timori espressi da David Ricardo e da altri, l'effetto complessivo della meccanizzazione fu quello di creare un numero senza precedenti di posti di lavoro. Le macchine permisero agli operai di produrre di più, fecero scendere il prezzo di molte merci e aumentarono la domanda, generando la necessità di altri operai. Nacquero lavori completamente nuovi per controllare le macchine. Le aziende eb-



Cdc 6600, il primo supercomputer della storia

bero bisogno di più dirigenti, contabili e altro personale. Con l'arrivo delle ferrovie, del telegrafo e dell'elettricità sorsero anche industrie che fino a quel momento nessuno aveva immaginato. Il timore della disoccupazione lasciò il posto a un discorso molto più ampio sulle condizioni di lavoro, favorendo la nascita delle teorie socialiste e comuniste e del movimento dei lavoratori. Alla fine dell'ottocento il problema delle macchine non si poneva più, perché la risposta era evidente.

Cosa ci dice questo sull'intelligenza artificiale di oggi? Finora il dibattito è stato dominato dal timore di un'enorme perdita di posti di lavoro e dalla possibilità che i computer prendano il sopravvento. Delle previsioni positive, secondo cui i computer cambieranno il mondo in meglio, in genere si parla meno. Facciamo tre esempi. Prima di tutto, l'intelligenza artificiale potrebbe modificare radicalmente l'uso dei traspor-

ti e la vita delle città, a partire dalle automobili senza conducente. Potendone chiamare una in qualsiasi momento, non sarebbe più necessario possedere un'auto, e questo ridurrebbe notevolmente il numero di veicoli in circolazione sulle strade e gli incidenti mortali. Le città rinascerebbero grazie al calo dell'inquinamento e gli spazi riservati ai parcheggi potrebbero essere trasformati in parchi, zone residenziali e piste ciclabili.

In secondo luogo, l'intelligenza artificiale potrebbe permettere alle persone di parlare con un'ampia gamma di oggetti: ovviamente con la loro casa e la loro automobile, come i personaggi di *Star Trek*, ma anche con avatar di aziende e altre organizzazioni, servizi d'informazione e consulenti. Questo potrebbe diventare un nuovo modo per interagire con i computer e accedere alle informazioni. L'intelligenza artificiale potrebbe anche aiutare gli esseri

umani a comunicare tra loro, traducendo in tempo reale i discorsi di persone che parlano lingue diverse.

Infine, l'intelligenza artificiale potrebbe fare un'enorme differenza nella ricerca scientifica e medica. Potrebbe diventare un'assistente infaticabile in campi come la ricerca sul cancro e il cambiamento climatico, analizzando i dati, leggendo migliaia di saggi scientifici e avanzando ipotesi o indicando correlazioni. Dopo anni di frustrazione per la lentezza dei progressi dell'intelligenza artificiale, ora è paradossale pensare che questa tecnologia stia avanzando troppo rapidamente. Negli anni quaranta dell'ottocento John Stuart Mill scriveva che "le prove del vantaggio delle invenzioni meccaniche per i lavoratori in futuro saranno evidenti". Un futuro economista potrebbe dire la stessa cosa dei vantaggi dell'intelligenza artificiale. Non solo per i lavoratori, ma per tutti. ♦ *bt*