

Il lato oscuro dell'intelligenza

Will Knight, Mit Technology Review, Stati Uniti

Foto di Max Aguilera-Hellweg

L'intelligenza artificiale si basa su una tecnologia che permette ai computer di imparare da soli. Ma c'è un problema: gli stessi ricercatori che l'hanno progettata non riescono a capire in che modo le macchine prendono le decisioni

Un anno fa una strana automobile senza conducente è stata mandata in giro per le tranquille strade di Monmouth County nel New Jersey, negli Stati Uniti. Il prototipo, sviluppato da un gruppo di ricercatori del produttore di processori grafici Nvidia, aveva un aspetto simile a quello di altre auto che si guidano da sole. Ma aveva qualcosa di completamente diverso rispetto ai modelli di Google, Tesla e General Motors, che dimostrava il potere crescente dell'intelligenza artificiale. L'auto della Nvidia non seguiva le istruzioni di un ingegnere o di un programmatore, ma si affidava interamente a un algoritmo che aveva imparato da solo a guidare osservando un essere umano.

Progettare un'auto di questo tipo è un'impresa notevole. Ma anche un po' inquietante, perché non è del tutto chiaro come la macchina prenda le sue decisioni. Le informazioni registrate dai sensori del veicolo finiscono in un'enorme rete di neuroni artificiali che elaborano i dati e danno le istruzioni necessarie per usare il volante, i freni e gli altri sistemi. I risultati sembrano simili al comportamento che ci si aspetterebbe da un guidatore in carne e ossa. Ma cosa succederebbe se un giorno l'auto facesse una mossa inaspettata, per esempio andasse a sbattere contro un albero o si fermasse con il semaforo verde? Allo stato attuale sarebbe molto difficile capire perché è successo. Il sistema è così complicato che

perfino gli ingegneri che l'hanno progettato hanno difficoltà a individuare i motivi all'origine di ogni decisione. Ed è impossibile chiedere una spiegazione alle macchine: non c'è un criterio assodato di progettare il sistema in modo che sia sempre capace di spiegare perché fa quello che fa.

La mente misteriosa di quest'automobile rimanda a una questione aperta sull'intelligenza artificiale. La tecnologia alla base della macchina, nota come *deep learning*, o apprendimento profondo, negli ultimi anni si è dimostrata molto efficace nella soluzione dei problemi ed è stata spesso usata per scopi come la traduzione, il riconoscimento vocale e l'*image captioning* (la descrizione di immagini attraverso le didascalie). Ora ci si aspetta che le stesse tecniche siano in grado di diagnosticare malattie mortali, fare investimenti milionari e rivoluzionare interi

settori industriali. Ma tutto questo non succederà (o almeno non dovrebbe) se non si troverà il modo di rendere le tecniche come il *deep learning* più comprensibili ai loro creatori e responsabili di fronte a chi le usa. Altrimenti sarà difficile prevedere quando ci saranno degli inconvenienti, che ci saranno inevitabilmente. Non a caso l'auto della Nvidia è ancora in fase sperimentale.

Oggi negli Stati Uniti si usano già dei modelli matematici per decidere, per esempio, chi può ottenere la libertà vigilata, chi può ricevere un prestito e chi deve essere assunto per un impiego. Accedendo a questi modelli si può capire il loro modo di ragionare. Negli ultimi tempi, tuttavia, le banche, l'esercito, le imprese e altri soggetti stanno rivolgendo l'attenzione a nuovi sistemi complessi di apprendimento automatico che rischiano di rendere del tutto imperscrutabili i processi decisionali automatizzati. Il *deep learning*, il sistema più diffuso, è un modo totalmente nuovo di programmare i computer. "Già ora è un problema rilevante, è lo sarà molto di più in futuro", osserva Tommi Jaakkola, un professore del Massachusetts Institute of Technology (Mit) che lavora sulle applicazioni dell'apprendimento automatico. "Per qualsiasi decisione - in campo finanziario, medico o militare - non ci si può affidare semplicemente a una scatola nera", cioè a un sistema per capire gli eventi solo a cose fatte.

Qualcuno già sostiene che la possibilità di chiedere a un'intelligenza artificiale com'è arrivata a determinate conclusioni

Da sapere

Le foto di questo articolo

◆ **Max Aguilera-Hellweg** ha fotografato per anni i robot sviluppati nei principali laboratori di ricerca del mondo. Il risultato è un libro intitolato *Humanoid* (Blast Books 2017), che raccoglie i ritratti di robot con sembianze umane. In queste pagine ci sono due ritratti - scattati tra il 2010 e il 2011 - di **Cb2**, un robot umanoide realizzato dall'Asada Laboratory dell'università di Osaka, in Giappone. Cb2 è un robot "bambino" creato per studiare i meccanismi attraverso cui le macchine intelligenti acquisiscono nuove capacità e imparano a svolgere alcune funzioni.