

TECNOLOGIA

Un altro guru dell'AI avverte: «Questi sistemi fanno paura»

PIETRO SACC

Geoffrey Hinton, vincitore di un premio Turing per gli studi sulle reti neurali, lascia Google per poter parlare liberamente dei pericoli di ChatGPT e simili «Non sviluppateli finché non sarà chiaro se sarete capaci di controllarli» Milano Ancora una volta uno dei massimi esperti dell'AI, l'intelligenza artificiale, ha ammesso di essere spaventato da quello che potrebbero fare i sistemi di elaborazione del linguaggio naturale, a partire da ChatGPT. Lui è Geoffrey Hinton, scienziato britannico 75 anni, 40 dei quali passati a studiare le reti neurali, vincitore del premio Turing (il Nobel dell'informatica) nel 2018 assieme ai colleghi Yann LeCun e Yoshua Bengio. Lavorando con i suoi studenti all'Università di Toronto, Hinton nel 2012 ha creato un sistema capace di analizzare migliaia di fotografie e riconoscere con estrema precisione all'interno delle immagini oggetti come fiori, automobili, volti. Qualcosa considerato "normale" oggi ma ancora estremamente innovativo dieci anni fa. Tanto che attorno all'azienda fondata da Hinton con i due studenti - chiamata DNNresearch - si accese un'asta tra la cinese Baidu e Google, conclusa nel 2013 con il passaggio della startup a Google per 44 milioni di dollari. Hinton è rimasto a lavorare con Google (oggi Alphabet) fino a quest'oggi, quando ha comunicato al ceo Sundar Pichai la decisione di lasciare perché, come ha scritto lui stesso su Twitter, «voglio poter parlare dei pericoli dell'AI senza preoccuparmi dell'impatto che questo ha su Google». La sua, ha chiarito lo studioso, non è una critica ad Alphabet (che anzi, dice Hinton, ha agito fin qui «in modo molto responsabile») ma in generale all'evoluzione dell'AI. Hinton potrà spiegare pubblicamente le sue ragioni già oggi, quando parlerà al forum EmTech Digital a Boston.

Nell'attesa, in due interviste, una al New York Times e una alla MIT Technology Review, Hinton ha detto di avere paura di quello che sono già diventati i sistemi di linguaggio basati sull'intelligenza artificiale. «Sono cose che sono totalmente diverse da noi. A volte penso che è come se gli alieni fossero atterrati e la gente non se ne fosse accorta perché parlano un ottimo inglese» dice lo scienziato. Lo spaventano le possibilità aperte dagli sviluppi di questa tecnologia. Molti degli allarmi sollevati da Hinton coincidono con quelli già emersi nei primi mesi dall'apertura al pubblico di ChatGPT.

Come i timori per la facilità con cui il sistema può facilitare la creazione di disinformazione in enormi quantità o per il rischio di distruzione di centinaia di milioni di posti di lavoro.

Ma lo studioso non nasconde di essere spaventato per il futuro dell'umanità.

«L'idea che questa roba potesse effettivamente diventare più intelligente delle persone... alcune persone ci credevano - ha spiegato Hinton -. Ma la maggior parte delle persone pensava che fosse una prospettiva lontana. Anch'io ho pensato che fosse lontana. Pensavo che mancassero dai 30 ai 50 anni o anche di più. Ovviamente non lo penso più». Gli scenari disegnati dallo scienziato britannico sono



Avvenire

catastrofici. Dice ad esempio Hinton parlando degli obiettivi secondari che l'IA potrebbe auto-assegnarsi mentre è chiamata a completare qualche mansione per gli umani: «C'è un sotto-obiettivo che in biologia aiuta quasi sempre: ottenere più energia. La prima cosa che potrebbe accadere è che questi robot diranno: "Otteniamo più potenza. Reindirizziamo tutta l'elettricità ai nostri chip". Un altro grande obiettivo secondario sarebbe quello di fare più copie di se stesso. Tisogna bene?». Ipotesi che sarebbero divertenti in qualche film di fantascienza ma che spaventano quando ad avanzarle è un vincitore di un premio Turing. Hinton non era tra i firmatari della lettera dello scorso marzo con cui centinaia di esperti di intelligenza artificiale chiedevano di interrompere le sperimentazioni e riprenderle solo dopo avere definito regole globali sull'IA. La sua posizione, però, è la stessa: «Non credo dovrebbero aumentare ulteriormente la potenza di questi sistemi fino a quando non avranno capito se possono controllarli » ha spiegato. RIPRODUZIONE RISERVATA.