

Le parole del futuro

Il progetto dell'ingegnere elettronico di **Meta** Fabio Petroni: «Così l'intelligenza artificiale suggerisce ai curatori gli articoli da verificare»

«Voci di Wikipedia da correggere? Basta l'algoritmo»

Tecnologia



Da sempre appassionato di intelligenza artificiale e deep learning, Fabio Petroni, romano classe 1986, ha già diverse pubblicazioni all'attivo in importanti riviste scientifiche.

Dopo aver conseguito un master in ingegneria informatica presso l'Università La Sapienza di Roma, è approdato ai laboratori londinesi di Meta, dove oggi è Tech Lead Manager presso il Fundamental AI Research (ex Facebook AI Research lab). Lì, insieme al suo team di

ricerca, ha sviluppato un algoritmo in grado di "smascherare" le citazioni false su Wikipedia.

Con oltre 55 milioni di articoli in più di 300 lingue, Wikipedia nei suoi 21 anni di vita è diventata l'enciclopedia online di riferimento per tutti i cittadini del web. Poche le regole per pubblicare un articolo, ma su una la Wikimedia Foundation non transige: tutte le voci devono essere supportate da fonti esterne attendibili. Spesso però accade il contrario, e ci troviamo alle prese con citazioni mancanti, incomplete o addirittura errate. Intanto la piattaforma continua a crescere, e diventa difficile per i volontari che operano su Wikipedia tenere il passo con gli oltre 17.000 nuovi articoli aggiunti ogni mese.

Per arginare il problema, il laboratorio FAIR di **Meta** ha sviluppato un algoritmo capace di scansionare automaticamente centinaia di migliaia di citazioni per verificare se davvero supportano quanto scritto nell'articolo. Ma c'è di più perché, secondo Petroni, spingere le macchine a studiare passaggi di testo complessi farà progredire l'IA verso sistemi sempre più intelligenti, capaci di rielaborare le sfumature più

sofisticate del sapere umano. E un giorno, magari, di creare da soli nuove voci enciclopediche.

Com'è nata l'idea di questo algoritmo?

«Ci siamo chiesti in che modo fosse possibile utilizzare la tecnologia per aiutare gli esseri umani a svolgere compiti piuttosto complessi, ad esempio verificare se le affermazioni su Wikipedia sono supportate da citazioni attendibili. Ogni giorno su Wikipedia avvengono migliaia di modifiche agli articoli esistenti, è impossibile controllarle tutte. Avere un sistema che consenta agli utenti di focalizzarsi solo sui casi problematici torna estremamente utile».

Come funziona?

«L'algoritmo setaccia Wikipedia per portare alla luce tutte le voci che contengono citazioni non ancora verificate. Poi suggerisce agli editor, cioè gli utenti che controllano gli articoli, quali pagine web contengono una fonte alternativa attendibile. Ha due componenti principali: il "verifier", che cerca di

«PER LA CELEBRE ENCICLOPEDIA ONLINE CONTROLLIAMO CHE CI SIANO FONTI ATTENDIBILI E LINK A GIORNALI AUTOREVOLI»



verificare se una voce sia supportata dalla sua citazione, e un sistema di raccomandazione che si appoggia a un database con oltre 100 milioni di pagine web da cui è possibile cercare una fonte alternativa».

Come ci riesce?

«Applicando delle nuove tecniche di analisi del linguaggio che confrontano il testo dell'articolo con quello della fonte, per capire se uno supporta l'altro, se c'è una

connessione. Il nostro modello ordina milioni di coppie citazione-fonte e le indicizza, consentendo così a un utente umano di avere subito sott'occhio i casi più problematici».

Quante citazioni false ha scovato il vostro sistema?

«Tantissime. Ad esempio abbiamo trovato un articolo su Joe Hipp, il primo nativo americano a competere per il titolo dei pesi massimi. Ma la citazione contenuta nell'articolo riportava a una pagina web dove non venivano menzionate né Hipp né la boxe. Il nostro sistema

è riuscito a trovare una pagina dove si parlava di entrambe».

Facciamo un altro esempio pratico: in una voce sui mondiali dell'82 scriviamo che a vincere la finale è stata la Germania e non l'Italia. Come interviene l'algoritmo?

«Per prima cosa controlla la citazione per verificare che ci siano fonti attendibili. Se non le trova, avvisa gli editor di Wiki-

pedia che non ha trovato nulla che supporti quell'affermazione. Gli utenti a quel punto possono controllare tutte le fonti che abbiamo segnalato e sostituirle, o con quelle suggerite da noi o con altre. E come si capisce se una fonte è attendibile? Il nostro modello, essendo stato allenato su milioni di citazioni, è riuscito in automatico ad

apprendere la nozione di affidabilità di una fonte, e preferisce ad esempio citazioni dal *Guardian* o dal *New York Times* rispetto ad articoli prelevati da qualche punto oscuro del web».

Usare sempre le stesse fonti non rischia di appiattire l'informazione?

«Wikipedia ha delle regole chiare sull'affidabilità delle fonti, noi ci atteniamo a quelle. E alla fine del processo comunque c'è sempre l'intervento umano».

Quindi l'algoritmo necessita

di una supervisione?

«Sì, non è ancora pensato per operare in maniera autonoma ma per fornire un supporto agli utenti. Certo, per avere il prossimo avanzamento in campo di ricerca IA sarà necessario che questi algoritmi riescano a fare ricerche come le fa un essere umano, cioè setacciando il web, confrontando fonti e nozioni. È anche in quest'ottica che abbiamo creato un indice così vasto, per insegnare alla macchina a svolgere ricerche».

Il passo successivo quale sarà?

«Non è impensabile che questi sistemi, già adesso così sofisticati, possano aiutare un giorno l'uomo a creare conoscenza nuova, ad esempio un teorema matematico o un articolo scientifico. O magari proprio una nuova pagina Wikipedia».

Un sistema IA che scrive articoli su Wikipedia. L'idea non la spaventa un po'?

«No, la trovo affascinante, sempre ovviamente in ottica di aiuto all'uomo. Le macchine potrebbero darci un giorno nuove conoscenze che strada facendo potrebbero esserci sfuggite».

R.d.E.

© RIPRODUZIONE RISERVATA

Sotto, Fabio Petroni, 36 anni, ingegnere elettronico romano, Tech Lead Manager presso il Fundamental AI Research di Meta a Londra

«IL SISTEMA NON È ANCORA PENSATO PER OPERARE IN MANIERA AUTONOMA, MA SOLO PER FORNIRE AIUTO ALL'UOMO»

I NUMERI

17mila

il numero di nuovi articoli che vengono aggiunti ogni mese su Wikipedia

172mila

Le modifiche apportate ogni giorno alle voci Wikipedia dagli utenti

134

in milioni, le pagine web contenute nel database usato dall'algoritmo di Meta

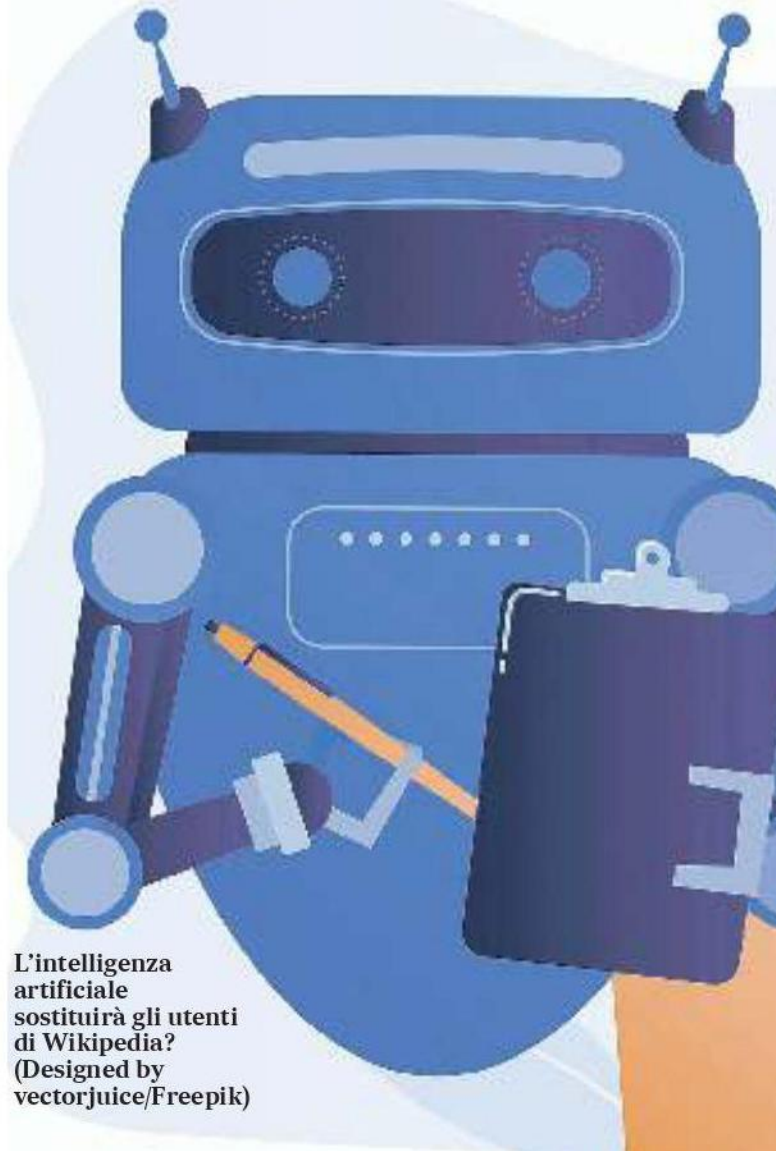
4

in milioni, le voci analizzate dall'algoritmo per trovare una fonte attendibile

41,4

in milioni, il numero di utenti registrati nella versione inglese di Wikipedia





L'intelligenza artificiale sostituirà gli utenti di Wikipedia?
(Designed by vectorjuice/Freepik)

