

SMASCHERARE LE FAKE NEWS, ANCHE QUELLE SUL COVID-19: L'EVOLUZIONE DELL'INTELLIGENZA ARTIFICIALE

A cura di Davide Sgro, AI Project Manager di Indigo.ai

Non è sicuramente una breaking news, ma per fortuna neanche una **fake news**: è infatti ormai assodato che per combattere le notizie false un alleato fondamentale è l'**Intelligenza Artificiale**. Complotti, sospetti, accuse, colpe... le fake news **fanno affidamento sulla buona fede delle persone** per creare allarmismi, se non addirittura sentimenti di rabbia, contro una parte politica, un'azienda, il governo...

Dall'epoca in cui tutti potevano avere un'opinione siamo passati a quella in cui **ognuno sembra aver diritto alla propria realtà**. «È vero perché lo dico io» o anche «è vero perché l'ho letto in un articolo»... ma abbiamo imparato che non è così. Ecco perché per smascherare le notizie false non bastano più i cosiddetti **fact checker**, ovvero quelle persone che di lavoro verificano notizie che suonano poco o tanto strane, documentandosi, studiando e indagando. Non sempre è facile venirne a capo soprattutto per la mole di dati che bisogna controllare. Ecco allora che arrivano in soccorso le "macchine", o meglio, gli algoritmi.

"Come l'intelligenza artificiale può rilevare le fake news"

Titola così un articolo pubblicato sul [Telegraph](#) dedicato all'AI, che spiega come la ricerca volta a utilizzare sistemi di intelligenza artificiale contro il fenomeno della disinformazione online stia procedendo con buoni risultati. Non a caso, il riconoscimento dell'AI come arma per combattere le fake news arriva soprattutto dalle istituzioni. Tra i gruppi che guidano la ricerca su questi temi, c'è il **Computer Science and Artificial Intelligence Laboratory (CSAIL) del MIT** in collaborazione con il **Qatar Computing Research Institute (QCRI)**, in cui i ricercatori coinvolti stanno utilizzando vari metodi per analizzare i siti dei media, gli account Twitter associati, la reputazione della fonte, il traffico web e altri fattori, al fine di ipotizzare classifiche di alta, media e bassa veridicità. Altri esempi di intelligenza artificiale in grado di elaborare fake news al fine di individuare contenuti informativi non veri sono **Grover**, creato dall'**Università di Washington**, e **Fandango**, finanziato dall'**Unione Europea**. O anche quello dell'**Università di Harvard** e del **MIT-IBM Watson AI Lab**, che hanno creato **GLTR** (Giant Language Model Test Room), uno strumento per aiutare a combattere la diffusione della disinformazione che utilizza l'intelligenza artificiale per rilevare i modelli di testo statistici.

Anche **in Italia** qualcosa si sta muovendo e player come **Pagella Politica**, da anni impegnata in prima linea per combattere le fake news, proprio in questi giorni si sta preparando a lanciare un nuovo progetto, **in collaborazione con noi di Indigo.ai**, per aiutare le persone a fare chiarezza sulle tante cose che si sono dette e che ancora si dicono sul nuovo **Coronavirus**.

indigo.ai

Medicina, AI & fact-checking: i tre ingredienti di un progetto innovativo

L'intelligenza artificiale si sta dimostrando un'arma importante anche nella **lotta al Coronavirus**, dai robot intelligenti per le prime diagnosi ai software che prevedono i numeri del contagio con elevati gradi di certezza. Fino all'AI a servizio della **battaglia contro alla disinformazione rispetto al virus Covid-19**. Da domande usali come **"È vero che indossare la mascherina fa male?"** a interrogativi più sui generis come **"È vero che Bill Gates aveva creato il vaccino contro il Coronavirus nel 2014?"**, sono tante le questioni più o meno verosimili che ci siamo trovati ad affrontare in questi mesi, magari attraverso messaggi WhatsApp, articoli di giornale, post condivisi su Facebook da qualche nostro contatto. E quante volte le nostre certezze hanno vacillato, o ci abbiamo creduto senza alcun dubbio?

Se il fenomeno delle fake news è pericoloso sempre, ancora di più lo diventa se il protagonista è la salute e se la notizia riguarda tutti. **L'OMS ha parlato di infodemia**, ossia una diffusione eccessiva di notizie false, fuorvianti o dalle fonti poco attendibili. Proprio per questo, player come **Facebook** e **l'International Fact-Checking Network** hanno dato vita ad un bando internazionale dal nome **"Coronavirus Fact-Checking Grants"** per dare una risposta concreta all'emergenza di tutelare le persone su un **argomento così delicato** e che **coinvolge la nostra salute**. Nell'ambito di questo impegno per la collettività è nata **Vera**, un **chatbot** che aiuta a **verificare notizie e avere informazioni affidabili sul Covid-19**, grazie a una conversazione interattiva. Vera è un'assistente che a partire dalle domande degli utenti sul Coronavirus aiuta a trovare la risposta, ed è stata creata da Indigo.ai in collaborazione con **Pagella Politica, Facta, Irccs Ospedale San Raffaele ed Eurecom**.

Grazie alla tecnologia da noi sviluppata, Vera non solo dà una risposta alla domanda posta dall'utente, ma lo guida anche per avere più informazioni. Chatbot di questo tipo esistono in ambito pharma e assicurativo, ma è **la prima volta che viene implementato per una piattaforma di news e fact-checking**. Vera è uno strumento rivoluzionario, perché se alcune fake news suonano così false da smascherarsi da sole, altre invece sembrano (e lo sono) costruite con così tanta minuzia o suonano così tecniche da riuscire a **ingannare anche le menti più attente**. Se per esempio è facile immaginare che il 5G non possa essere la causa principale della diffusione del virus (una delle bufale diceva proprio così), più difficile è sapere se il tasso di mortalità riportato da un sito sia giusto o sbagliato. Il buon senso aiuta, ma **l'approccio scientifico** è l'unico che può davvero funzionare per combattere le notizie false.

Se l'intelligenza è artificiale... quanto serve l'uomo?

La ricerca sull'AI è in continua evoluzione, ed oggi l'avanguardia è incentrata sullo sviluppo di algoritmi di apprendimento profondo, meglio conosciuto come Deep Learning.

indigo.ai

In precedenza, i modelli di intelligenza artificiale si basavano fortemente su modelli statistico-matematici che miravano a *modellizzare* i dati con forti assunzioni iniziali, ovvero avevano bisogno di una serie di scelte che determinavano il modo in cui il modello cercava di rappresentare la realtà: il cosiddetto **bias induttivo**.

La nuova tendenza, ovvero **l'apprendimento profondo** (che è una sottocategoria del Machine Learning), si basa sul concetto di **reti neurali artificiali** per consentire all'intelligenza artificiale di fare previsioni e prendere decisioni in modo indipendente. Questo perché le reti neurali riescono a crearsi in autonomia una rappresentazione dei dati e quindi viene meno l'esigenza di forti assunzioni da parte dei *data scientist*.

Ad oggi, i migliori risultati sono ottenuti da una categoria molto specifica di rete neurale: i **transformers**, che **simulano la capacità di attenzione dell'essere umano**. Per esempio, quando leggiamo una frase ci sono alcune parole che sono più importanti rispetto ad altre per capirne il senso. Oppure in un'immagine sono alcuni particolari o dettagli che ci fanno distinguere un oggetto o riconoscere una persona. Questa innovazione ha portato ad una vera svolta in termini di risultati ottenuti, avvicinando sempre di più le intelligenze artificiali al livello umano: l'ultimo incredibile esempio è *GPT-3*, un modello di generazione di testo sviluppato da OpenAI e capace di scrivere articoli indistinguibili da quelli umani. Diventa quindi sempre più importante integrare la supervisione umana a ciò che viene generato dall'AI, come viene fatto nel caso del *fact checking*.

Ma l'AI **può eliminare autonomamente bias sociali e discriminazioni** all'interno dei suoi algoritmi? Questo è attualmente un argomento di dibattito, con la maggior parte del settore che concorda che la risposta dipende sia da come vengono raccolti i dati sia da come vengono sviluppati i modelli. Poiché ci si basa sull'addestramento umano e su dati reali, è facile che i pregiudizi e le percezioni di coloro che sviluppano l'AI si riflettano nella tecnologia o che si perda il controllo su ciò che viene imparato dalle macchine. Un esempio eclatante è stato Tay, un bot Twitter sviluppato da Microsoft nel 2016 che in poche ore ha iniziato a scrivere affermazioni razziste, a insultare e a negare l'Olocausto, obbligando i suoi ideatori a spegnerlo immediatamente. Bisogna quindi aumentare l'attenzione della comunità scientifica su questi temi per fare in modo che nei sistemi di AI vengano rispecchiati i concetti di equità e correttezza, la cosiddetta *AI Fairness*.

Informazioni su [Indigo.ai](https://indigo.ai)

Siamo uno studio di Conversational AI che progetta e costruisce assistenti virtuali, tecnologie di linguaggio ed esperienze conversazionali. Nati a Settembre 2016 tra i banchi del Politecnico di Milano da un'idea di cinque giovani (Gianluca Maruzzella, Enrico Bertino, Marco Falcone, Andrea Tangredi e Denis Peroni - ad oggi quasi tutti under30), abbiamo realizzato assistenti virtuali per alcune delle aziende più innovative al mondo, tra cui banche, assicurazioni, case farmaceutiche, etc. Abbiamo costruito un framework proprietario di Natural Language Processing che, sfruttando l'intelligenza artificiale, è in grado di comprendere le informazioni nel testo o nella voce in maniera completamente automatica: grazie a questo framework aiutiamo le aziende ad automatizzare conversazioni, efficientare processi, alleggerire il customer care ed ingaggiare i clienti in maniera super personalizzata. Il nostro team è formato da 20 persone e operiamo sia in Italia che all'estero. Tra il 2017 e il 2020 siamo stati scelti due volte

indigo.ai

come rappresentanti della delegazione delle start-up italiane al CES di Las Vegas e abbiamo vinto tre riconoscimenti del premio Gaetano Marzotto – tra i più importanti nel panorama dell’innovazione.

Ufficio Stampa Indigo.ai: ddl studio

indigo@ddlstudio.net

Mara Linda Degiovanni | +39 3496224812

Elisa Giuliana | +39 3386027361

Sabrina Barozzi | +39 3336158644