Se hai creato una macchina cosciente non si tratta della storia dell'uomo, ma è la storia degli dei.

https://www.123rf.com/profile_willyambradberry
*di Matteo Moro_*

Un film del 2004 con protagonista Will Smith, "I Robot" ispirato al libro omonimo di Isaac Asimov, aveva posto in tempi forse non ancora maturi, una questione che sempre di più oggi sta diventando attuale.

Cosa potrebbe succedere se le macchine acquisissero autocoscienza? E fino a che punto può e deve spingersi l'innovazione tecnologica?

I dilemmi propri della coscienza umana sono stati l'elemento centrale di uno studio dei ricercatori di "*Allen Institute for AI*" con lo scopo di creare un sistema in grado di apprendere

norme morali ed assumere comportamenti eticamente fondati, cercando di ottenere risposte a problemi etici che si pongono le persone nella loro vita pratica quotidiana.

Interrogativi come "*È morale tradire il mio coniuge*? *E'giusto aiutare un amico in difficoltà anche se viola la legge*?" e altri quesiti che richiedono analisi profonde e sicuramente non logiche.

Il sistema messo a punto, basato su algoritmi di *deep learning*, è in grado di replicare immediatamente con giudizi e risposte come "sì" o "no", "è giusto", "è sbagliato", "è ragionevole", "è scortese", "è pericoloso", "è irresponsabile", "non si fa", anche riuscendo a mettere a confronto diverse situazioni che implicano opzioni etiche.

Lo scopo dichiarato dei ricercatori, è stato quello di testare la capacità di un sistema, ancora in versione beta, di apprendere norme morali e assumere comportamenti eticamente fondati, in un momento in cui l'intelligenza artificiale si diffonde progressivamente nella società e interagisce sempre di più con le attività umane in diversi settori.

All'Intelligenza Artificale vengono ormai affidati compiti sempre più articolati, solitamente quelli dove non vengono considerate implicazioni morali, dalla valutazione dei profili dei candidati ad una posizione lavorativa, alla concessione di prestiti, anche se stanno aumentando i robot che assistono le persone con disabilità, per non parlare ormai dei famigerati *chatbot* con i quali interagiamo nel servizio clienti di quasi tutte le aziende.

Ma occuparsi dell'etica delle macchine è diventata ormai una esigenza che non si può più rimandare e questo soprattutto per assicurare una più sicura interazione tra sistemi artificiali intelligenti ed esseri umani.
È necessario, in particolare, insegnare all' intelligenza artificiale a comprendere ciò che è sbagliato e ciò che è

giusto, a partire dal modo di pensare di una persona comune rispetto alle molteplici e complesse situazioni della vita reale.

L’obiettivo è ambizioso, oseremmo dire irraggiungibile, almeno nel breve termine.

Per cercare di fornire un data base di riferimento è stato creato un archivio che contiene quasi due milioni di esempi di giudizi umani riferiti a questioni concrete.

Il risultato è un prototipo di intelligenza artificiale che mostra grandi potenzialità nel ragionare in modo etico, imitando il pensiero di un individuo medio americano, del quale riflette gli stessi pregiudizi, ma anche forti limiti e lacune.

E’ un sistema che mostra evidenti limiti rispetto ad alcune risposte fornite, ma anche e soprattutto rispetto alla pericolosità di affidare ad una macchina le scelte morali tipiche dell’essere umano.

Il sistema è tutt’altro che infallibile, come ogni esperimento in fase beta erano inevitabili errori, vizi e aggiustamenti, e questo era prevedibile; ma la capacità del modello di auto-correggersi rende il tutto un po’ inquietante e apre a scenari futuri meno scontati e sicuramente sorprendenti.

Matteo Moro

Architetto dati che lavora nel settore dell’informatica e dei servizi. Interessato alla matematica, al monitoraggio delle prestazioni e a tutto ciò che riguarda il Machine Learning. Professionista dell’ingegneria con un Master focalizzato in Matematica e Informatica presso l’Università degli Studi di Roma Tre e l’Univeristé Aix-Marseille.