Oste è buono il vino ? O del paradosso di fidarsi degli annunci pubblicitari



Alessandro Giuliani | 30 Dicembre 2023

I fondamenti del funzionamento di tali sistemi sono piuttosto semplici e non differiscono molto dalle normali interrogazioni che quotidianamente facciamo con Google per reperire informazioni di interesse. L'illusione è quella di avere a che fare con una sorta di intelligenza, l'illusione però non resiste a un esame accurato,...

Una piccola perla

In una strepitosa (e illuminante) intervista alla CNN del 2018 – di cui caldamente consiglio la visione ai lettori di Bene Comune – Robert Shank (uno dei padri della cosiddetta Intelligenza Artificiale, da ora in poi confidenzialmente IA) chiarisce in maniera esatta e puntuale la paradossale natura dell'odierno dibattito sull'argomento. La situazione è paradigmatica: siamo all'inizio della riesumazione (a scopi evidentemente commerciali) di un vecchio mito risalente alla metà del secolo scorso (ma stia attento il lettore, che nel capitalismo terminale il commercio è l'essenza del potere).

Una entusiasta intervistatrice, mentre nel sottopancia scorrono le notizie dell'enorme successo in borsa delle azioni delle aziende che si occupano di IA, si aspetta da un sornione e divertito Shank, il coronamento dello spot con mirabolanti dichiarazioni sulle conseguenze epocali del nuovo (sic) prodotto. Le cose però prendono una piega inaspettata, lo scienziato afferma che forse, se invece di definire l'IA come 'intelligenza' la si definisse correttamente come 'un software' forse ci capiremmo meglio. Shank, con tranquilla saggezza, spiega cose ben note a chiunque abbia una se pur minima preparazione scientifica, e cioè che l'IA non è altro che una tecnica di analisi dei dati come tante altre che si sono susseguite nel tempo (e.g. analisi di regressione, analisi in componenti principali, analisi discriminante,

reti neurali..) e che quindi non ha nulla a che fare con l'intelligenza umana (o animale). In altre parole, a nessuno verrebbe in mente di assegnare un Nobel alla metodologia matematica che ha reso possibile una scoperta, così come (per ora, domani chissà) nessuno mette in galera l'arma di un delitto, preferendo giudicare il criminale.

A onore della giornalista va però detto che ella non perde il suo smalto e continua a sorridere e a porre possibili eccezioni a cui Shank risponde con grazia e sottile ironia. A questo punto però, la lettrice e il lettore di questo articolo hanno diritto a qualche ulteriore chiarimento.

La diminuzione dell'incertezza è un'esigenza dell'animo umano

'*Non mi lasciare sulle spine!*' è una frase che ci fa ben capire come sia insopportabile per ognuno di noi il permanere in una condizione di incertezza. Quasi tutte le attività umane, in modi diversi e con diversi orizzonti temporali, hanno a che vedere con il lenimento dell'angoscia che ci proviene da un mondo che non riusciamo a controllare. La via della scienza alla diminuzione dell'incertezza è quella della generalizzazione e della formalizzazione simbolico/quantitativa: ogni scienza cerca di fare delle previsioni accettabilmente precise sul mondo riducendone l'incertezza. La generalizzazione, riconducendo un fenomeno particolare a una classe di fenomeni simili di cui conosciamo l'esito e/o la natura, è uno strumento imprescindibile per la conoscenza del mondo e quindi per la 'domesticazione' del suo carico di imprevedibilità.

La via alla generalizzazione della fisica classica ha una origine legata alla filosofia di Platone (l'atto di nascita della fisica moderna è contemporaneo alla critica mossa all'Aristotelismo della tarda Scolastica): quello che ci propongono i nostri sensi è solo l'ombra sbiadita della vera realtà fatta di regolarità geometriche (forme ideali) pre-esistenti e indipendenti dalla loro apparenza terrena. La matematica ci fornisce un assaggio del mondo degli enti ideali: scrivere una equazione nella forma Y = f(X) dove Y è la grandezza che vogliamo stimare, mentre la (o le) X sono delle proprietà osservabili e il termine f si riferisce a delle operazioni matematiche eseguibili sulle osservabili X per stimare il valore di Y. Questa procedura corrisponde alla eliminazione dell'incertezza: se si conosce l'esatta formulazione della funzione (f), il valore della grandezza Y di interesse può essere calcolato con precisione. In modo solo apparentemente differente (e.g. senza quasi mai ricorrere a leggi matematiche esplicite), le

diagnosi cliniche cercano di prevedere una particolare patologia (Y) attraverso la valutazione di sintomi e valori di esami strumentali e/o di laboratorio (X) che rimandano il singolo paziente a un quadro più generale (la patologia). In questo caso la f può contenere allo stesso tempo un insieme di regole logiche del tipo 'se…allora..', funzioni discriminanti basate sul valore quantitativo di risultanze di laboratorio, analisi di immagini. Però questa via (che l'acuta lettrice di Bene Comune avrà già individuato come quella dell'IA) ha solo una vaga parentela con la rarefatta bellezza di una formula matematica, infatti ha una genesi molto diversa, è roba di Aristotele, piuttosto che di Platone.

In altre parole, qui non è che la verità ultima sia nella forma di una equazione ma, più modestamente, si cerca di fare buon uso di tutte le informazioni disponibili per dedurre una previsione il più possibile affidabile. D'altronde i medici dell'antico Egitto diagnosticavano correttamente le epatiti dal colore giallo della cornea che era una conseguenza secondaria (ma utile per la diagnosi, da cui il nome itterizia ancora in voga fino a pochi decenni fa) delle malattie del fegato. Il recente entusiasmo per le applicazioni di riconoscimento di forme o 'pattern recognition' (impropriamente dette di intelligenza artificiale) in medicina, deriva proprio dalla possibilità offerta da questi strumenti di trattare contemporaneamente molte informazioni eterogenee tra loro. Ma esattamente cosa intendiamo con il verbo 'trattare' ?

Niente di particolarmente diverso da ciò che faceva il medico egiziano: se si accorgeva che la cornea giallastra si associava (quasi) sempre con una costellazione di disturbi, usava questo segno come un 'predittore' di questi ultimi. Il medico imparava dall'esperienza e la sua fiducia nel potere predittivo del sintomo aumentava con l'aumentare della casistica. Ora sostituiamo a un semplice sintomo una costellazione (anche molto grande) di informazioni e a una semplice deduzione logica (*se* hai la cornea gialla *allora* digerirai con difficoltà, avrai dolori e spossatezza, febbri intermittenti…) qualcosa di più complesso del tipo: *la tua configurazione, descritta da un numero n (con n molto elevato) di caratteristiche, ti avvicina molto alla configurazione (pattern) tipica della malattia Y*. La scelta e l'uso di una metrica (come misuro la distanza dai pattern ideali) differenzia i diversi tipi di riconoscimento di forme (modelli di IA).

Il tipo di apprendimento a partire dai casi reali e quindi la scelta di un insieme rappresentativo, delle *n* caratteristiche di interesse, il controllo della

generalizzazione del modello ad altri casi non usati per l'apprendimento (cioè nella costruzione del modello) e altre attività di questo tipo fanno parte del lavoro artigiano del ricercatore e garantiscono lo stipendio a fine mese all'autore di questo articolo.

Tralasciamo per brevità altri punti importanti, primo fra tutti quello della cosiddetta 'spiegabilità' (explainability, il tema più caldo della ricerca in IA, si veda "Avoid Oversimplifications in Machine Learning: Going beyond the Class-Prediction Accuracy") e cioè la necessità di rendere intellegibili all'utente le motivazioni di una procedura di previsione, compito molto difficile per sistemi di calcolo come quelli denominati 'deep learning' (le IA più di moda) e abbandoniamo la scienza per entrare nel mondo dell'illusionismo. Il lettore arrivato a leggere fin qui avrà però già compreso che l'IA è uno strumento (come un pennello o un martello) usato con maggiore o minore perizia e non certo il pittore o il carpentiere.

Stanze cinesi e ChatBot

Nel 1980 (attenzione alla data, sono passati più di 40 anni) il filosofo americano John Searle pubblicò un articolo fondamentale che passò alla storia come 'Il caso della stanza cinese' (Il lettore ne potrà leggere una versione in Italiano in: John R. Searle, *La mente è un programma?*, in *Le Scienze*, n. 259, 1990).

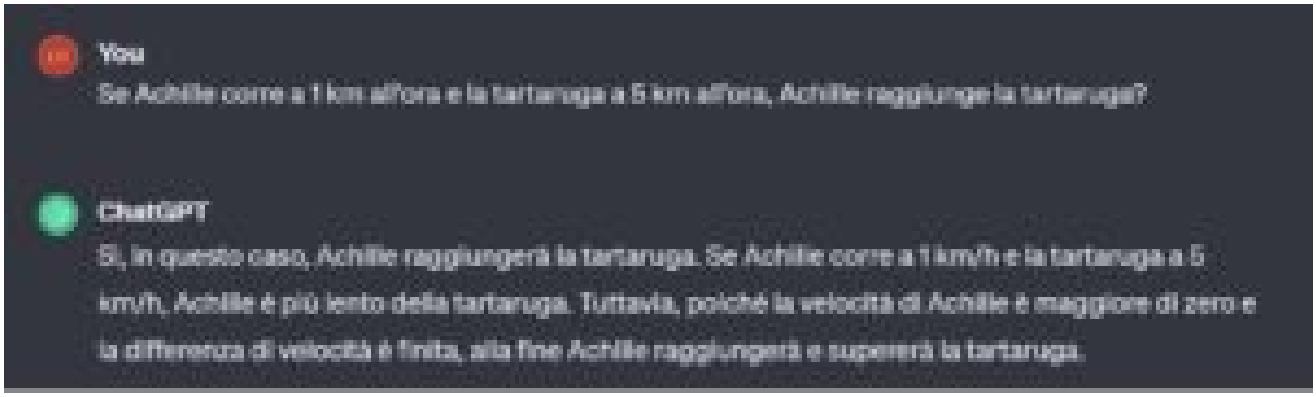
La faccenda è questa: immaginiamo di essere chiusi in una stanza con a disposizione una tabella (molto complessa per la verità, ma che con potenti mezzi informatici risulta essere facilmente consultabile) che associ a ogni frase in cinese (lingua a noi ignota) una corrispondente frase in italiano (nell'articolo originale ovviamente in inglese). Questa stanza viene propagandata a ignari utenti cinesi come la sede di un oracolo in grado di rispondere ai più svariati quesiti, incuriositi dal prodigio varie persone infilano nella stanza biglietti con su scritte delle domande. L'abitante della stanza, attraverso la tabella di cui sopra, legge il contenuto della domanda nella sua lingua e risponde affidando la traduzione in cinese alla stessa tabella. L'interrogante avrà la sua risposta. La questione è: io, abitante nella stanza, posso dire di conoscere il cinese? La risposta è sicuramente no, e neanche la tabella conosce nulla, è solo un sistema di accoppiamento (tra l'altro ai giorni nostri disponibile anche su smartphone).

L'articolo di Searle sancisce la definitiva falsificazione dell'identità tra

informazione sintattica e informazione semantica, la tabella di Searle non è altro che una di quelle IA dette ChatBot (di cui ChatGpt è la più nota) che hanno più di tutte stimolato la fantasia del grande pubblico.

I fondamenti del funzionamento di tali sistemi sono piuttosto semplici e non differiscono molto dalle normali interrogazioni che quotidianamente facciamo con Google per reperire informazioni di interesse. Siamo sempre nel caso del riconoscimento di forme: si tratta di consultare (e qui la potenza di calcolo aiuta molto) sterminati insiemi di documenti e trovare quelli 'più vicini' all'interrogazione. Tale vicinanza si esprime come grado di sovrapposizione fra la frase posta come domanda e i documenti nel web, Tale sovrapposizione si quantifica con metriche piuttosto sofisticate capaci di generare testi (con il metodo copia/incolla senza nessuna relazione con il significato) coerenti con la domanda stessa. La procedura genera componimenti piuttosto scontati derivanti da queste assonanze (ripeto, solo sintattiche fondate sulla regolarità dell' uso dei sequenze di lettere, niente a che vedere con il significato).

L'illusione è quella di avere a che fare con una sorta di intelligenza, l'illusione però non resiste a un esame accurato, ad esempio provate a chiedere qualcosa su 'Celine' a ChatGpt avendo in mente l'autore di 'Viaggio al termine della notte' e probabilmente avrete di ritorno la carriera artistica di Celine Dion (che ha molti più riferimenti in rete rispetto allo scrittore francese). Ma ecco un esempio spassoso fornitomi dal mio amico Lorenzo (valentissimo ingegnere informatico docente alla "Sapienza" di Roma):



Il gioco qui è abbastanza scoperto, l'ignara IA è finita (per motivi puramente statistici) sulla similitudine fra la questione posta da Lorenzo (You) e il famoso paradosso di Zenone, e questo spiega la risposta. Peccato

però che, proprio per l'assenza di reale comprensione semantica del testo, ChatGpt non comprenda che Lorenzo aveva posto come ipotesi che Achille fosse decisamente più lento della tartaruga (anche se sembra accorgersene nella seconda frase della risposta) e quindi non la avrebbe mai raggiunta. Però ChatGpt continua a riferirsi al paradosso di Zenone e appiccica una spiegazione strampalata come frase finale con un Achille lentissimo che supera una tartaruga velocissima e quindi concludendo che 1 è più grande di 5 (dopo aver appena affermato il contrario).

Avviso agli studenti (e non solo) svogliati: se avete accarezzato l'idea di usare ChatGpt per i vostri temi, almeno rileggete con attenzione la risposta…ora però passiamo ad alcune considerazione più inquietanti.

Il palazzo di Atlante , Tarkovsky, l'inganno

Nell'ottavo canto del capolavoro di Ariosto si narra di Orlando che cerca Angelica e ha una visione: gli sembra di vederla su un cavallo, rapita da un cavaliere. Così Orlando insegue il cavaliere ed entra in un palazzo incantato: chi ci entra pensa di avere di fronte a se l'oggetto del proprio desiderio ma in realtà è solo un'apparenza (infatti Orlando crede di vedere Angelica, Ruggero Bradamante, i cavalieri i cavalli…).

E' il palazzo di Atlante, una macchina maligna e mirabile che genera illusioni basandosi su quello che gli anglosassoni chiamerebbero '*wishful thinking*' (pensiero desiderante, prendere per reali e avverati i nostri desideri più profondi), è ciò che avviene nella 'zona', nel film Stalker di Andrey Tarkovsky (i.e. il film preferito dall'autore dell'articolo). Attenzione però, non è che la realizzazione dei desideri sia sempre un bene. Questo è molto chiaro allo stalker che accompagna i suoi clienti nella 'zona' quando afferma '*Il tuo desiderio profondo potrebbe non esserti noto, la zona potrebbe esaudirlo e poi scoprire troppo tardi che tuo fratello è morto*'. Tornando al castello di Atlante, capiamo che esso non è troppo dissimile da quelle pubblicità volte a convincerci che l'acquisto di una automobile ci trasformi magicamente in una persona dotata di un fascino irresistibile e di uno spirito libero e coraggioso. E' chiaramente un inganno, ma nel nostro animo ci cadiamo, chi non vorrebbe essere così ?

Su quale desideri profondi fa leva allora l'evidente inganno di una macchina pensante? Che sia un inganno è piuttosto semplice da capire e ne abbiamo parlato prima, ma allora cosa desideriamo? Perché filosofi, politici,

scienziati, letterati, imprenditori ci credono e discettano su questi argomenti in maniera compulsiva?

Fatta la tara della malafede (che credo sia una componente molto importante) rimane una grandissima parte di persone sinceramente affascinate (o preoccupate che è lo stesso) dalla 'macchina intelligente' che forse ci renderà inutili. Io ho le mie idee in proposito ma, perdonatemi, non ve le esporrò in questo articolo che è già troppo denso. Dirò solo che anche una illusione può avere conseguenze reali, e di queste parleranno diffusamente altri interventi di questo numero di Bene Comune.

Tags: Intelligenza artificiale scienza