

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

L'intelligenza artificiale salverà il mondo? I FOUNDATION MODEL

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1885291> since 2023-06-04T20:40:27Z

Publisher:

Egea spa

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

L'intelligenza artificiale salverà il mondo? I FOUNDATION MODEL

Il dibattito sull'IA oscilla da anni tra previsioni nefaste, che porteranno a rendere superfluo l'essere umano e far scomparire la maggior parte dei lavori che oggi conosciamo, e incredibili progressi che ci apriranno le porte di una vita migliore. Cureremo più malattie, diminuiranno le ricadute negative sull'ambiente, conquisteremo lo spazio. Gli scettici partecipano alla discussione, sicuri che questa tecnologia non arriverà mai a raggiungere gli effetti sperati se non, forse, tra molti anni. Questo approfondimento non ha la pretesa di indirizzare il lettore verso una posizione sul tema, ma semplicemente di far conoscere sugli sviluppi reali e tangibili che stanno alimentando l'intelligenza artificiale attraverso una serie di casi. E immaginare di poter trarre qualche vantaggio nelle attività della nostra vita.

I FOUNDATION MODEL

L'intelligenza artificiale (IA) che ha dato buoni risultati è uscita dai paper di ricerca e dai test dei laboratori ed è entrata poco alla volta nella nostra vita. I social media e le applicazioni che usiamo ogni giorno impiegano algoritmi di intelligenza artificiale per suggerire film, canzoni, articoli o percorsi stradali. Le banche usano l'IA per determinare quale forma di finanziamento o piano di investimento sia meglio per i loro clienti, le fabbriche fanno uso di robot guidati da intelligenza artificiale per la produzione, i centri di ricerca usano sofisticati algoritmi per analizzare gli eventi e immaginare nuovi scenari ai quali rispondere velocemente.

Negli ultimi 10 anni gli astri della IA hanno iniziato a brillare con una luce più intensa. *Gli investimenti privati nell'IA nel 2021 sono più che raddoppiati rispetto all'anno precedente sfiorando i 93,5 miliardi di dollari. Mentre gli Stati hanno continuato a investire per arrivare a raggiungere il primato del miglior ecosistema autonomo di Intelligenza Artificiale all'interno dei loro confini, le aziende si concentrano sullo sviluppo di una tecnologia che definisca la base di qualunque progetto di intelligenza artificiale. L'impegno dei privati ha dato i suoi frutti arrivando a sviluppare quelli che vengono definiti foundation model o modelli di base.*

La tecnologia

Secondo il centro dell'Università di Stanford Center for Research on Foundation Models (CRFM), un modello di base è "un qualsiasi modello che viene addestrato su dati ampi e che può essere adattato per un'ampia gamma di compiti a valle¹". I foundation model non sono nuovi: si basano sulle reti neurali profonde (deep neural networks) e sull'apprendimento auto-supervisionato (self-supervised learning) che esistono già da tempo. I modelli precedenti di IA erano associati a problemi specifici e richiedevano un notevole dispendio di tempo e risorse per essere sviluppati e utilizzati. Ogni nuova applicazione necessitava di individuare un set di dati ampio e ben etichettato per il compito specifico da affrontare. Se il dataset non esisteva, era compito dei ricercatori impiegare centinaia o migliaia di ore per trovare ed etichettare immagini, testi o grafici appropriati per il dataset.

Il risultato era che modelli specifici venivano sviluppati per progetti specifici. Il riconoscimento del linguaggio, quello delle immagini, o ancora lo studio delle molecole o dei farmaci, si basavano su approcci lunghi, costosi e personalizzati. Inoltre ciò che veniva sviluppato per un modello non poteva essere utilizzato per un altro.

I foundation model superano queste limitazioni grazie sia a caratteristiche tecniche sia all'evoluzione di contesto. A livello tecnico i modelli di base sono abilitati dall'apprendimento per trasferimento, dall'utilizzo del self supervised learning e dalla possibilità di essere scalabili. L' apprendimento per trasferimento (transfer learning), rappresenta la capacità di spostare la "conoscenza" appresa in un compito (ad esempio, il riconoscimento di oggetti nelle immagini) e applicarla a un altro compito (ad esempio, il riconoscimento di attività nei video) costituisce la base della scalabilità del modello. *Le tecniche di apprendimento auto-supervisionato, invece addestrano l'Intelligenza artificiale automaticamente dal sistema con dati non "etichettati" diminuendo di molto tempi e costi di sviluppo².* La scalabilità è ottenuta grazie al pre-addestrato su un compito surrogato e poi all'adattamento ai diversi casi. Per quanto riguarda le evoluzioni di contesto, grazie all'aumento della capacità di calcolo³ e alla disponibilità di una grande mole di dati per il pre-addestramento i

¹ <https://arxiv.org/pdf/2108.07258.pdf>

² la diminuzione dei costi di sviluppo e di addestramento del software si attesta intorno all 63,6% con un miglioramento dei tempi di "produzione" pari al 94,4%

³ la memoria delle GPU e il throughput (ossia la capacità) sono aumentate di dieci volte negli ultimi quattro anni

foundation model possono svilupparsi in modo adeguato. Più dati sono utilizzati, più grande è il modello, maggiore la generalizzazione che si può ottenere e di conseguenza le applicazioni che si potranno sviluppare. Esattamente come l'elettricità, le autostrade, o internet, l'intelligenza artificiale grazie ai foundation model diventa "general purpose" ossia utilizzata per finalità differenti.

I primi potenti esempi soprattutto nel campo del linguaggio naturale e dell'elaborazione delle immagini non mancano e hanno dato vita ad un complesso ecosistema di governi e attori privati. A seguire una breve descrizione.

L'ecosistema dei foundation model

Molti sono gli stati che ad oggi hanno definito strategie e investimenti tra cui Europa, America e Cina.

La strategia europea per l'IA mira non solo allo sviluppo di IA affidabile e incentrata sull'uomo ma a creare un polo di eccellenza mondiale nel vecchio continente.

Nell'aprile 2021, la Commissione ha presentato un pacchetto sull'IA, che comprende sia la promozione attraverso la comunicazione dell'approccio europeo all'intelligenza artificiale, sia un aggiornamento del Piano coordinato sull'intelligenza artificiale (con gli Stati membri dell'UE), sia una proposta di regolamento che stabilisca norme armonizzate sull'IA (European AI Act) e la relativa valutazione d'impatto.

La Commissione Europea e gli Stati membri hanno concordato nel promuovere l'eccellenza nell'IA unendo le forze in materia di politiche e investimenti. Attraverso i programmi Digital Europe e Horizon Europe, prevedendo di investire 1 miliardo di euro all'anno nell'IA. Ulteriori investimenti dal settore privato e dagli Stati membri saranno utili per raggiungere un volume di investimenti annuali di 20 miliardi di euro nel corso del decennio digitale.

Anche gli Stati Uniti d'America attraverso la National Artificial Intelligence Initiative (NAII) hanno istituito una legge bipartisan promulgata il 1° gennaio 2021: la National Artificial Intelligence Initiative Act of 2020 (NAIIA). *Oltre a garantire la continua leadership degli Stati Uniti nella ricerca e nello sviluppo dell'intelligenza artificiale, l'iniziativa si pone diversi obiettivi di carattere geopolitico e di autonomia strategica. Gli obiettivi riguardano la guida del mondo nello sviluppo e nell'uso di sistemi di intelligenza artificiale affidabili nei settori pubblico e privato; la preparazione della forza lavoro statunitense attuale e futura all'integrazione dei sistemi di intelligenza artificiale in tutti i settori dell'economia e della società e infine, il coordinamento delle*

attività di intelligenza artificiale in corso in tutte le agenzie federali per garantire che ciascuna informi il lavoro delle altre.

La Cina non è da meno e con l'obiettivo di diventare la superpotenza mondiale dell'intelligenza artificiale, il Consiglio di Stato della Repubblica Popolare Cinese ha dichiarato di voler investire 150 miliardi di dollari entro il 2030. Non c'è oggi in Cina una legislazione che disciplini in modo diretto l'IA, nonostante il paese si ponga tra i primi per i finanziamenti, la ricerca e l'utilizzazione di questo strumento ⁴.

Il principale punto di riferimento in materia resta però il “New Generation AI Development Plan” (Aidp) adottato dal Consiglio di Stato il 20 luglio 2017, **che indica le linee strategiche per fare della Cina entro il 2030 il paese più avanzato nel campo della IA. È un documento che non ha un'immediata valenza prescrittiva ma, nel modello cinese, condiziona fortemente il comportamento degli attori pubblici e privati del settore.**

Mentre a livello paese l'intelligenza artificiale è affrontata in modo generale, a livello industriale, l'ecosistema basato sui foundation model nel campo del linguaggio naturale inizia a tratteggiarsi. **Alla base del settore privato troviamo le aziende già focalizzate sui cloud e datacenter che forniscono la potenza necessaria per l'addestramento dell'Intelligenza artificiale. I protagonisti di questo primo layer sono Amazon, Google, Microsoft. Il secondo livello è formato dai centri di ricerca e big tech che sviluppano modelli di intelligenza artificiale e utilizzano le cloud factory per fare il training dei loro modelli e migliorarli. Il terzo livello è formato da aziende che costruiscono servizi di intelligenza artificiale. Alcune di particolare interesse AI21 e OpenAI, DeepMind. Infine abbiamo gli utilizzatori. Da Netflix a Spotify a Google, a chiunque abbia necessità di usare questa tipologia di modelli.**

AI21 Labs fondata nel 2017 è una start up israeliana che conta ad oggi 120 dipendenti, un round di Serie B da 64 milioni di dollari e una valutazione dell'azienda pari a 664 milioni di dollari. **Il team di AI21 Labs è guidato da Yoav Shoham, un professore di informatica alla Stanford University e membro del comitato scientifico di Google Brain, e da Yoav Artzi, un professore di informatica**

⁴ Y. Luo, Z. Yu “An interview with Covington & Burling discussing Artificial Intelligence in China”, in Lexology, 27 novembre 2020.

all'Università di Washington che in passato ha ricoperto il ruolo di direttore di ricerca a Facebook AI Research. Il primo prodotto dell'azienda è stato Wordtune, un aiuto alla scrittura alimentato dall'intelligenza artificiale destinato a competere con Grammarly, che suggerisce di riformulare il testo ovunque gli utenti scrivano. I modelli più evoluti dell'azienda sono anche in grado di riassumere contenuti, rispondere a operazioni matematiche "non banali" formulate in linguaggio naturale, di semplificare domande "complesse" e rispondere a operazioni matematiche "non banali" formulate in linguaggio naturale.

Nell'ottobre del 2015, Elon Musk fonda a San Francisco insieme ad altri importanti investitori, OpenAI una non-profit dedicata alla ricerca e allo sviluppo di intelligenza artificiale in modo da garantirne un utilizzo sicuro ed etico. I prodotti sviluppati sono open source e disponibili a tutti. L'azienda ha sviluppato tra gli altri prodotti DALL-E 2 capace di trasformare le descrizioni testuali in immagini così creative e uniche da aver permesso a Jason M. Allen di vincere il concorso artistico annuale della Colorado State Fair. Il sistema sta migliorando così rapidamente che le immagini di DALL-E 2 sono quattro volte più dettagliate di quelle generate dal DALL-E originale, presentato solo un anno prima. DALL-E 2 non rielabora immagini esistenti su Internet, ma ne crea di nuove attraverso un complesso processo di I.A. noto come "diffusion" (diffusione), che parte da una serie casuale di pixel, perfezionandola ripetutamente finché non corrisponde a una determinata descrizione testuale. Ma l'arte non è l'unico settore in cui l'intelligenza artificiale ha fatto passi da gigante. Il 18 dicembre 2019, OpenAI ha reso disponibile il proprio sistema di generazione di testi GPT-2, un modello di machine learning che è in grado di generare testi in base a un prompt iniziale. Il sistema è basato su un modello di machine learning chiamato Transformer, inizialmente sviluppato da Google nel 2017, e addestrato su un enorme corpus di dati di testi, raccolti da Internet. OpenAI ha dichiarato che GPT-2 è "in grado di generare testi coerenti e realistici in molti contesti", e che si tratta di un "passo avanti significativo nell'intelligenza artificiale". Il successore di GPT-2 è GPT-3, sviluppato per generare testi in molti stili e contesti diversi, imitando il modo in cui gli esseri umani scrivono. La rete neurale è in grado di produrre testi coerenti e ben scritti in molti contesti differenti, come racconti, articoli di giornale, elenchi di cose da fare, e-mail e altro ancora. Per esempio, allenato su un dataset di oltre dieci miliardi di

parole e frasi, è realizzato con lo scopo di migliorare l'auto-completamento di qualunque tipo di testo o codice appartenente a qualunque settore. *Dalla sanità, all'ambiente, all'educazione fino all'arte. Già oggi i programmatori di Microsoft utilizzano GPT3 per sviluppare codice e potersi concentrare su attività più importanti o per entrare nel dettaglio di processi più critici.* Kaving Scott il CTO di Microsoft in una recente intervista sull'Economist ha dichiarato in una recente intervista dell'Economist⁵ che “i programmatori dicono di amare molto questo programma”...“il programma non li ha sostituiti ma gli ha permesso di fare più cose e di focalizzarsi su attività che per loro erano più importanti entrando più nel dettaglio dei problemi”

DeepMind, nata nel 2010, con sede a Londra è stata acquistata da Google nel 2014 per un importo stimato di \$400 milioni. DeepMind è principalmente nota per i suoi successi nell'apprendimento automatico, in particolare per il suo programma di apprendimento profondo AlphaGo, che ha sconfitto il campione del mondo di Go, Lee Sedol, nel 2016. L'evoluzione di AphaGo è AlphaFold. che nel 2020 ha scoperto qualcosa di unico nel campo della ricerca scientifica risolvendo il problema “del ripiegamento delle proteine”. Un problema apparentemente irrisolvibile che riguarda il numero di modi in cui una proteina può teoricamente piegarsi prima di stabilizzarsi nella sua struttura tridimensionale finale e alla base di quasi tutte le malattie, compreso il cancro, la demenza e persino le malattie infettive come il Covid-19. Nel 2021 la rivista Science ha riconosciuto l'importanza di AlphaFold, nominandolo la più grande scoperta scientifica dell'anno. Ma DeepMind continua ad investire in AlphaFord e nella calda estate 2022 ha annunciato che il sistema ha fatto previsioni su quasi tutte le 200 milioni di proteine conosciute, producendo un tesoro di dati che aiuterà i ricercatori medici a sviluppare nuovi farmaci e vaccini per gli anni a venire.

LE IMPLICAZIONI FUTURE

Le implicazioni dell'utilizzo dell'intelligenza artificiale sono molteplici alcune sicuramente meritano attenzione. L'implicazione sociale e politica non è da sottovalutare. Sono pochi i soggetti che possono oggi costruire questo genere di modelli individuabili principalmente nelle bigtech. La concentrazione in poche mani è

⁵ The Economist “Artificial Intelligence enters its industrial age” June 7th 22

pericolosa non solo perché aumenta il ruolo politico di alcuni soggetti rendendo difficoltose le azioni di collaborazione con altre aziende o paesi a condizioni eque, ma anche perché non è possibile capire esattamente su quali logiche e con quali dati i modelli vengono costruiti. Sarà quindi necessario muoversi con cautela nell'utilizzo di questa tecnologia in aree sensibili come l'educazione, la formazione e la giustizia. Già oggi framework legislativi, in primis quello europeo, sono indirizzati verso richieste di trasparenza e valutazione alle aziende che sviluppano l'IA.

Per quanto riguarda il loro funzionamento, essendo addestrati su grandi quantità di dati provenienti da Internet, compresi i social media, questi modelli anche se usati correttamente possono generare testi tossici e distorti basati su un linguaggio simile a quello incontrato durante l'addestramento. E visto che su essi si basa lo sviluppo di altri modelli di intelligenza artificiale potrebbero queste problematiche potrebbero diffondersi velocemente.

Infine la tecnologia potrebbe essere usata in modo discordante per aumentare la disinformazione su vasta scala, la produzione di deepfake, la polarizzazione, l'odio on line. Non è uno scenario futuro. Già oggi abbiamo prove di usi distorti di questa tecnologia. Tutte queste problematiche hanno scatenato una miriade di panel, conferenza, think thank, comitati etici. Da Davos ai Ted talks passando per la pletora di eventi sulla tecnologia e il digitale, nessuno si esime dal delineare piani emergenziali per un futuro distopico. Rari sono invece i momenti di incontro in cui si parla in modo neutrale di ciò che i sistemi di I.A. di oggi sono effettivamente in grado di fare delineando una strategia per utilizzare questi sistemi, coglierne le opportunità e mitigare i rischi.

CONCLUSIONI

Per mitigare le ricadute sulla società, aggiornarsi sullo stato dell'arte di questa tecnologia e sviluppare un'esperienza diretta sulle potenzialità dello strumento, sono sfide che i governi non dovrebbero ignorare. Istituire un laboratorio per sperimentare l'uso dell'Intelligenza Artificiale nel settore pubblico potrebbe costituire, sul versante delle nuove tecnologie, un buon inizio anche per il governo italiano. La collaborazione tra le poche aziende proprietarie della tecnologia e i governi democratici è fondamentale.

E' ormai evidente quanto siano difficili da gestire per aziende e paesi l'interesse e la sicurezza delle nazioni come i diritti sulle persone, se considerati in ritardo rispetto allo sviluppo tecnologico. Gli esempi non mancano, innanzitutto sui social network.

Superare alcune barriere, aiutarci a guardare più lontano, indirizzando la nostra attenzione sulla parte di lavoro che è più piacevole, interessante e importante, saranno opportunità da cogliere per noi lavoratori.

Questo caso ne è la dimostrazione: scritto insieme con l'intelligenza artificiale GPT-3 e Wordtune di A21. Ho provveduto ad "istruire" il sistema attraverso gli articoli che ho scritto quest'anno sull'Intelligenza Artificiale . **Le parti in grassetto sono scritte interamente dall'Intelligenza Artificiale GPT-3, quelle in corsivo sono scritte da me ma revisionate dall'intelligenza artificiale Wordtune. le parti scritte in grassetto e corsivo sono scritte da GPT-3 e revisionate da Wordtune.** Le conclusioni costituiscono la parte più creativa caso, alla quale ho dedicato maggior impegno e attenzione. Senza il ricorso all'intelligenza artificiale.

Bibliografia

<https://www.ai.gov/strategy-documents/>

<https://aiindex.stanford.edu/report/>

https://www.wordtune.com/?utm_source=ai21_web

<https://beta.openai.com/>

https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_it

https://www.wordtune.com/?utm_source=ai21_web