Serie superconvergenza e rivoluzione delle biotecnologie 2024: premonizioni di AlphaFold 3

Pubblicato il 27 agosto 2024

Christopher Gannatti, CFA
Global Head of Research

Principali insegnamenti

- ChatGPT ha richiamato l'attenzione sull'idea di utilizzare l'IA per predire il testo, che poi porta naturalmente a predire immagini o addirittura video in base a diversi prompt. Possiamo utilizzare questi concetti per prevedere i diversi modi in cui proteine e molecole potrebbero interagire all'interno delle cellule?
- Il ripiegamento e la struttura delle proteine pongono problemi affascinanti da analizzare: una singola proteina potrebbe potenzialmente ripiegarsi in qualcosa come 10^300 modi diversi, un numero troppo elevato per qualsiasi sistema che tenti di trovare una soluzione alla volta.
- È straordinario notare che il quesito sul ripiegamento delle proteine si protrae da circa 50 anni e che abbiamo da poco iniziato a vedere progressi costanti, come AlphaFold, AlphaFold 2, AlphaFold 3... La rapidità con cui la convergenza tra IA, capacità di calcolo e comprensione delle biotecnologie sta facendo progredire tali quesiti è davvero sorprendente.
- Prodotti correlati WisdomTree BioRevolution UCITS ETF – USD Acc, WisdomTree Artificial Intelligence UCITS ETF – USD Acc Scopri di più

Nel costruire la nostra strategia dedicata alla rivoluzione delle biotecnologie, abbiamo collaborato con il futurista dott. Jamie Metzl, membro del comitato di esperti sull'editing del genoma umano dell'Organizzazione mondiale della sanità. Riteniamo di essere alla vigilia di un periodo straordinario, che potrebbe durare qualche decennio, in cui metteremo in discussione e infine trasformeremo il nostro modo di fare le cose, ad esempio:

- come gestiamo l'assistenza sanitaria per gli esseri umani;
- come coltiviamo il cibo per una popolazione globale in espansione;
- come generiamo materiali, prodotti chimici ed energia innovativi da fonti biologiche;

- come pensiamo di memorizzare enormi quantità di dati con una densità e una fedeltà superiori a quelle del passato.

Il dott. Metzl ha recentemente pubblicato il libro *Superconvergence: How the Genetics, Biotech, and AI Revolutions will Transform our Lives, Work and World*. (Superconvergenza: in che modo la rivoluzione della genetica, delle biotecnologie e dell'IA trasformerà le nostre vite, il nostro lavoro e il nostro mondo). Nel corso dell'estate, pubblicheremo una serie di post che illustreranno alcune delle idee presentate nella sua opera.

I risultati

In un certo senso, l'investimento tematico consiste nel raccontare una storia e Superconvergence è riuscito a trasmettere in modo eccellente quella alla base del WisdomTree BioRevolution ESG Screened Index.

Grazie ChatGPT

A mio parere, l'aspetto migliore di ChatGPT risiede nel fatto che ha reso il concetto di intelligenza artificiale (IA) generativa accessibile a quasi tutti. Nelle sue fasi iniziali, la popolarità dell'IA generativa risiedeva nella generazione di testi, ma da allora si è evoluta fino alla generazione di immagini, suoni e video. È logico pensare che tutto ciò che ha un sistema e una struttura, in grado di essere formato in base a diverse regole e relazioni, possa prestarsi a essere predetto.

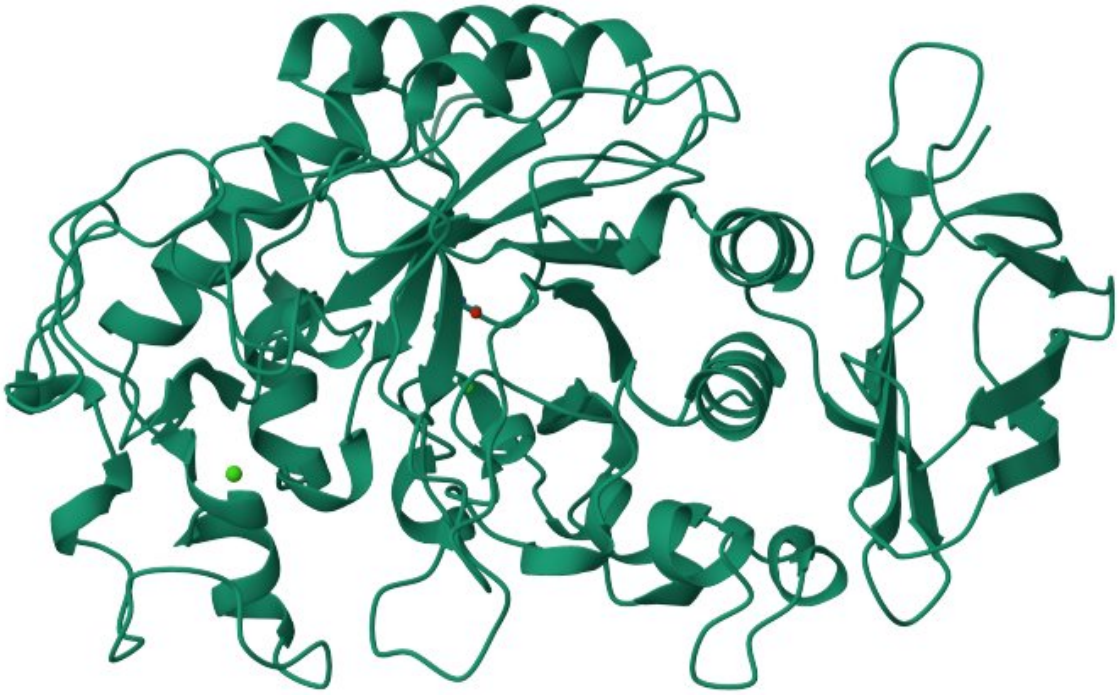
La potenza di questi strumenti risiede nella loro capacità di formulare continuamente previsioni su cui poi riflettere, non nel fatto che tali previsioni siano sempre accurate al 100%. Passando alle biotecnologie, non si tratta di "prevedere una cura", ma piuttosto di "prevedere un nuovo percorso di ricerca" che possa condurre a un risultato terapeutico interessante.

I sistemi di IA generativa sono in grado di prevedere strutture proteiche utili?

Prevedere diverse serie di parole che abbiano senso, o definire la posizione di pixel all'interno di un'immagine sulla base di un prompt è una cosa, ma prevedere le interrelazioni tra le molecole in un sistema biologico è completamente diverso.

Riportiamo la Figura 1 per far comprendere a chi non ha una formazione in biologia molecolare (me compreso) la complessità di una singola proteina da un punto di vista visivo; in questo caso, si tratta dell'amilasi salivare (saliva) umana proveniente dal Protein Data Bank in Europa. È impossibile non rimanere impressionati da come gli scienziati che eseguono fisicamente la cristallografia a raggi X abbiano gettato le basi per determinare la struttura delle proteine.

Figura 1: Struttura delle proteine della saliva umana



Fonte: https://www.ebi.ac.uk/pdbe/entry/pdb/1smd

Annuncio di AlphaFold nel 2020

Il professor John Moult, co-fondatore e presidente del Critical Assessment of Protein Structure Prediction (CASP) dell'Università del Maryland, ha dichiarato[1]:

Sono circa 50 anni che siamo bloccati su questo problema: come si ripiegano le proteine. Vedere DeepMind produrre una soluzione allo stesso, dopo averci lavorato personalmente per così tanto tempo e dopo tante battute d'arresto e ripartenze, chiedendoci se ci saremmo mai arrivati, è un momento davvero speciale.

Gli esperimenti del CASP mirano a fare il punto della situazione nel campo della previsione della struttura delle proteine, a identificare i progressi e a mettere in evidenza le aree su cui concentrare gli sforzi futuri. Questi hanno cadenza biennale e il primo si è svolto nel 1994[2].

Nel 1969, quando la sfida era prevedere la struttura tridimensionale di una proteina, si è stimato che una proteina tipica potesse avere qualcosa come 10.000 possibili conformazioni. Questo ci dice che un approccio di tipo “forza bruta”, in cui un sistema esamina tutte le possibilità, non è fattibile[3].

Simulare una cellula

Possiamo imparare molto da simulazioni migliori e a più alta risoluzione. Il 18 marzo 2024, Nvidia ha annunciato un gemello digitale del clima terrestre: il concetto è che se riuscissimo a simulare il clima e il tempo a una risoluzione sempre maggiore, potremmo usare tale simulazione per capire meglio e infine prevedere i cambiamenti del tempo e del clima[4]. Sebbene l'idea sembri semplice, raccogliere ed elaborare

un numero sufficiente di dati per avere anche solo una ragionevole possibilità di essere abbastanza precisi da risultare affidabili non è facile.

Considerando la difficoltà e il costo legati all'organizzazione di studi clinici per testare le diverse terapie possibili, è facile pensare che se potessimo simulare il corpo umano e i suoi sistemi sottostanti, potremmo imparare molto e dipendere meno dall'organizzazione di sperimentazioni cliniche sull'uomo.

È interessante fare un passo indietro e riconoscere che nel corso della storia abbiamo analizzato DNA, rRNA, aminoacidi e proteine. Stiamo edificando la nostra comprensione pezzo dopo pezzo. La complessità di ciascuna di queste fasi è sbalorditiva.

Da *Superconvergence*:

Nel 2022, il fondatore di DeepMind, Demis Hassabis, ha dichiarato:

Uno dei miei sogni nei prossimi 10 anni è quello di produrre una cellula virtuale. Per "cellula virtuale" intendo la modellazione dell'intero funzionamento della cellula con un sistema di IA. Si potrebbero fare esperimenti virtuali su quella cellula e le previsioni che ne deriverebbero sarebbero confermate dalla verifica in laboratorio. Riuscite a immaginare quanto sarebbe più veloce ed efficiente l'intero processo di scoperta dei farmaci e di sperimentazione clinica se avessimo a disposizione una cosa del genere? . . . Possiamo considerare quello che abbiamo fatto con AlphaFold come il primo gradino della scala. . . . Poi poco a poco si progredisce, magari arrivando ai pathway e alle cellule e poi, forse, all'intero organismo. Questo è il sogno[5].

Presentiamo AlphaFold 3

Un aspetto divertente della pubblicazione di scritti su internet o degli interventi in diversi podcast è che, se sai come cercarli, puoi trovare le previsioni più disparate che le persone hanno formulato o le aspirazioni che hanno espresso in un particolare momento. È interessante leggere una tale affermazione da parte di Demis Hassabis a luglio del 2024, quando siamo ormai a conoscenza del recente lancio di AlphaFold 3.

L'8 maggio 2024 è stato pubblicato quanto segu[6]:

Presentiamo AlphaFold 3, un nuovo modello di IA sviluppato da Google DeepMind e Isomorphic Labs. Prevedendo con precisione la struttura di proteine, DNA, RNA, ligandi e altro ancora, e il modo in cui interagiscono, speriamo che trasformi la nostra comprensione del mondo della biologia e della scoperta dei farmaci.

Secondo la mia lettura, dichiaratamente amatoriale, se AlphaFold 3 simulasse un'intera cellula, l'avrebbero già scritto, quindi ritengo che questo rappresenti un passo importante lungo il percorso e che dovremmo aspettarci sempre più versioni in grado di applicare una risoluzione sempre maggiore alle cellule e, in ultima analisi, agli organismi viventi. È affascinante che, anche durante la pubblicazione del libro Superconvergence, i progressi nel percorso verso la capacità di predire la struttura delle proteine e, quindi, di simulare le cellule, siano continuati. Abbiamo sequenziato il genoma, previsto come si ripiegano le diverse proteine e ora stiamo iniziando a prevedere come interagiscono le varie molecole e proteine.

Sembra che questi progressi avvengano a un ritmo sempre più rapido. Se immaginiamo una linea temporale, questo appare davvero evidente: il Progetto genoma umano è stato completato nel 2003[7], AlphaFold è stato annunciato nel 2020[8] e AlphaFold 3 nel 2024[9]. Con l'avanzare dell'hardware di calcolo e dell'IA, il settore delle biotecnologie dovrebbe potenzialmente essere in grado di progredire allo stesso ritmo. La convergenza di questi diversi megatrend potrebbe rendere i prossimi anni davvero interessanti.

[1] Fonte: https://deepmind.google/discover/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology/
[2] Fonte: https://predictioncenter.org/index.cgi
[3] Fonte: Levinthal, Cyrus (note di A. Rawitch). 1969. How to fold graciously. In Mössbauer Spectroscopy in Biological Systems, rivisto da P. Debrunner, J. C. M. Tsibris, e E. M?nck. Atti di una riunione tenutasi presso la Allerton House, 17-18 marzo 1969, Monticello, Illinois. Urbana, Ill.: University of Illinois Press.
[4] Fonte: https://nvidianews.nvidia.com/news/nvidia-announces-earth-climate-digital-twin#:~:text=GTC%E2%80%94To%20accelerate%20efforts%20to,and%20climate%20at%20unprecedented%20scale.
[5] Fonte: Metzl, Jamie. Superconvergence: How the Genetics, Biotech, and AI Revolutions will Transform our Lives, Work and World. Timber Press: 2024.
[6] Fonte: https://blog.google/technology/ai/google-deepmind-isomorphic-alphafold-3-ai-model/#life-molecules
[7] Fonte: https://www.genome.gov/human-genome-project/timeline#:~:text=More%20%2B-,2003,two%20years%20ahead%20of%20schedule.
[8] Fonte: https://deepmind.google/discover/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology/
[9] Fonte: https://blog.google/technology/ai/google-deepmind-isomorphic-alphafold-3-ai-model/#life-molecules

Important Risks Related to this Article

Informazioni importanti

Comunicazioni di marketing emesse all'interno dello Spazio economico europeo ("SEE"): Il presente documento è stato emesso e approvato da WisdomTree Ireland Limited, società autorizzata e regolamentata dalla Central Bank of Ireland.

Comunicazioni di marketing emesse in giurisdizioni non appartenenti al SEE: Il presente documento è stato emesso e approvato da WisdomTree UK Limited, società autorizzata e regolamentata dalla Financial Conduct Authority del Regno Unito.

Per fare riferimento a WisdomTree Ireland Limited e a WisdomTree UK Limited si utilizza per entrambe la denominazione "WisdomTree" (come applicabile). La nostra politica sui conflitti d'interesse e il nostro inventario sono disponibili su richiesta.

Solo per clienti professionali. Le informazioni contenute nel presente documento sono fornite a titolo meramente informativo e non costituiscono né un'offerta di vendita né una sollecitazione di un'offerta di acquisto di titoli o azioni. Il presente documento non deve essere utilizzato come base per una qualsiasi decisione d'investimento. Gli investimenti possono aumentare o diminuire di valore e si può perdere una parte o la totalità dell'importo investito. Le performance passate non sono necessariamente indicative di performance future. Qualsiasi decisione d'investimento deve essere basata sulle informazioni contenute nel Prospetto informativo di riferimento e deve essere presa dopo aver richiesto il parere di un consulente d'investimento, fiscale e legale indipendente.

L'applicazione di regolamenti e leggi fiscali può spesso portare a una serie di interpretazioni diverse. Eventuali punti di vista o opinioni espresse in questa comunicazione rappresentano le opinioni di WisdomTree e non devono essere interpretate come consulenza normativa, fiscale o legale. WisdomTree non fornisce alcuna garanzia o dichiarazione circa l'accuratezza di qualsiasi punto di vista o opinione espressa in questa comunicazione. Qualsiasi decisione di investimento dovrebbe essere basata sulle informazioni contenute nel prospetto appropriato e dopo aver richiesto una consulenza finanziaria, fiscale e legale indipendente.

Il presente documento non è, e in nessun caso deve essere interpretato come, una pubblicità o qualsiasi altro strumento di promozione di un'offerta pubblica di azioni o titoli negli Stati Uniti o in qualsiasi provincia o territorio degli Stati Uniti. Né il presente documento né alcuna copia dello stesso devono essere acquisiti, trasmessi o distribuiti (direttamente o indirettamente) negli Stati Uniti.

Benché WisdomTree si adoperi per garantire l'esattezza del contenuto del presente documento, WisdomTree non garantisce né assicura la sua esattezza o correttezza. Qualsiasi terzo fornitore di dati di cui ci si avvalga per reperire le informazioni contenute nel presente documento non rilascia alcuna garanzia o dichiarazione di sorta in relazione ai suddetti dati. Laddove WisdomTree abbia espresso dei pareri relativamente al prodotto o all'attività di mercato, si ricorda che tali pareri possono cambiare. Né WisdomTree, né alcuna consociata, né alcuno dei rispettivi funzionari, amministratori, partner o dipendenti,

accetta alcuna responsabilità per qualsiasi perdita, diretta o indiretta, derivante dall'utilizzo del presente documento o del suo contenuto.