

19. Gennaio

Il ruolo del “overfitting “ nell’ utilizzo di algoritmi nella medicina di precisione

*Tutto è vago al punto che non te ne rendi conto fino a che
non hai tentato di renderlo preciso.*

Bertrand Russel

Il mistero non esiste che nelle cose precise.

Jean Cocteau

La **medicina di precisione** promette trattamenti su misura per i profili individuali dei pazienti. I modelli di **machine learning** sono stati annunciati come gli strumenti per accelerare la medicina di precisione vagliando grandi quantità di dati complessi per individuare **marcatori genetici, sociodemografici o biologici** che predicono il trattamento giusto per la persona giusta al momento giusto.

Tuttavia, l’entusiasmo iniziale per questi strumenti predittivi avanzati si trova ora ad affrontare un **confronto con la realtà** che fa riflettere.

I modelli computerizzati non riescono a prevedere i risultati del trattamento della schizofrenia. Nonostante abbiano la stessa diagnosi, i pazienti affetti da schizofrenia presentano un’ampia gamma di sintomi e i trattamenti che funzionano per alcune persone potrebbero non funzionare per altri. I modelli computerizzati, che possono rapidamente ordinare gruppi di dati clinici e identificare le tendenze, potrebbero potenzialmente prevedere quali farmaci saranno i più efficaci, consentendo ai ricercatori di creare piani di trattamento su misura per ciascun individuo. Ma quanto sono accurate queste previsioni?



Il team di **Adam Chekroud** del Dipartimento di Psichiatria, Yale University nel report

Chekroud AM et al.

Illusory generalizability of clinical prediction models.

Science. 2024 Jan 12;383(6679):164-167.

mostrano che i modelli di apprendimento automatico che prevedono la risposta al trattamento ai farmaci antipsicotici tra gli individui con schizofrenia in uno studio clinico non sono riusciti a generalizzare i dati provenienti da studi clinici nuovi e mai visti.

I risultati non solo evidenziano la necessità di standard metodologici più rigorosi per gli approcci di apprendimento automatico, ma richiedono anche un riesame delle sfide pratiche che la medicina di precisione si trova ad affrontare.

Il team di Yale Utilizzando i dati di cinque studi clinici, i ricercatori hanno valutato la capacità di un algoritmo di apprendimento automatico di prevedere il successo di diversi farmaci antipsicotici. Quando il modello è stato addestrato utilizzando i dati campionati da un singolo studio, ha previsto con precisione i risultati dei pazienti all'interno di quello specifico studio.

Tuttavia, il modello ha fallito quando i ricercatori lo hanno applicato a un set di dati diverso. Anche l'utilizzo di tutti e cinque i set di dati per addestrare il modello non è riuscito a risolvere il problema: le sue previsioni non potevano essere generalizzate a un nuovo studio clinico indipendente.



Il problema, secondo la neuroscienziata **Frederike Petzschner** del Dipartimento di Psichiatria e Comportamento Umano e il **Carney Institute for Brain Science presso la Brown University**, direttrice del programma Carney Brainstorm che accelera la traduzione della scienza del cervello computazionale in applicazioni cliniche e commercializzazione

in una prospettiva correlata, probabilmente ha a che fare con un fenomeno noto come “overfitting”, in cui un algoritmo informatico si conforma troppo strettamente ai dati utilizzati per addestrarlo.

L'overfitting è un fenomeno della data science che si verifica quando un modello statistico si adatta esattamente ai suoi dati di addestramento. Quando questo accade, l'algoritmo purtroppo non può funzionare correttamente in presenza di dati non osservati in precedenza, vanificando il suo scopo.

La generalizzazione di un modello sui nuovi dati è in definitiva ciò che ci consente di utilizzare gli algoritmi di *machine learning* ogni giorno per fare delle previsioni e classificare i dati.

Quando vengono creati degli algoritmi di apprendimento automatico, essi sfruttano un set di dati campione per addestrare il modello.

Tuttavia, quando il modello si addestra troppo a lungo su dati campione o quando il modello è troppo complesso, può iniziare ad apprendere il "rumore", o le informazioni irrilevanti, all'interno del set di dati. Quando il modello memorizza il rumore e si adatta troppo strettamente al set di addestramento, il modello diventa soggetto a "overfitting" e non è in grado di generalizzare bene sui nuovi dati.

Se un modello non è in grado di generalizzare bene sui nuovi dati, allora non sarà in grado di eseguire le attività di classificazione o previsione per le quali è stato concepito. Bassi tassi di errore e una varianza elevata sono buoni indicatori di overfitting. Per prevenire questo tipo di comportamento, una parte del set di dati di addestramento è tipicamente messa da parte come "set di test" per controllare l'eventuale presenza di overfitting. dati di addestramento con un basso tasso di errore e dati di test con un alto tasso di errore sono indice di overfitting.

Complessivamente **Chekroud** che Il risultato da "*un quadro deludente ma realistico della nostra capacità di sviluppare modelli veramente utili per i risultati del trattamento della schizofrenia*".

Petzschner concorda, sottolineando che i risultati *"non solo evidenziano la necessità di standard metodologici più rigorosi per gli approcci di apprendimento automatico, ma richiedono anche un riesame delle sfide pratiche che la medicina di precisione si trova ad affrontare"*.

I modelli computerizzati non riescono a prevedere i risultati del trattamento della schizofrenia. Nonostante abbiano la stessa diagnosi, i pazienti affetti da schizofrenia presentano un'ampia gamma di sintomi e i trattamenti che funzionano per alcune persone potrebbero non funzionare per altri. I modelli computerizzati, che possono rapidamente ordinare gruppi di dati clinici e identificare le tendenze, potrebbero potenzialmente prevedere quali farmaci saranno i più efficaci, consentendo ai ricercatori di creare piani di trattamento su misura per ciascun individuo.

In sintesi

I modelli di apprendimento automatico che prevedono la risposta al trattamento ai farmaci antipsicotici tra gli individui con schizofrenia non sono riusciti a generalizzare i dati provenienti da studi clinici nuovi e mai visti.

I risultati non solo evidenziano la necessità di *standard metodologici più rigorosi* per gli approcci di apprendimento automatico, ma richiedono anche un riesame delle sfide pratiche che la medicina di precisione si trova ad affrontare.

VITTORIA

Vittoria, 68 anni, con insufficienza cardiaca, malattia renale cronica e ipertensione è ricoverata nel reparto di terapia intensiva per insufficienza respiratoria.



Negli ultimi 5 giorni ha manifestato una progressiva mancanza di respiro associata a dolore toracico pleurítico sul lato destro e tosse che occasionalmente produce espettorato schiumoso giallo striato di sangue.

Oggi, la sua famiglia l'ha portata al pronto soccorso dove è stata notata una marcata dispnoia con una frequenza respiratoria di 36 respiri al minuto e una saturazione di ossigeno dell'81% mentre respirava aria ambiente.

È stata immediatamente intubata e ricoverata nel reparto di terapia intensiva medica.

All'esame, è sedata ma apre gli occhi a un vigoroso stimolo tattile.

La sua temperatura è 38,3°C.

Il suo polso venoso giugulare non è visibile. Ci sono crepitii grossolani in entrambi i campi polmonari e nessun soffio o galoppo; non c'è edema periferico. Il resto dell'esame è normale.

I test di laboratorio danno i seguenti risultati:

	Valore del paziente	Intervallo di riferimento
Conta dei leucociti (per mm ³)	15.700	4500–11.000
Emoglobina (g/dl)	10.2	12.0–16.0
Conta piastrinica (per mm ³)	126.000	150.000–350.000
Creatinina (mg/dl)	2.5	0,6–1,1
Troponina T (ng/mL) (test di 4a generazione)	0,02	0–0,03
Emogasanalisi arteriosa (ottenuta su una frazione di ossigeno inspirato pari a 0,40)		
pH	7.32	7.38–7.44
Pressione parziale dell'ossigeno (mm Hg)	76	80-100
Pressione parziale dell'anidride carbonica (mm Hg)	30	35–45

Si ottiene una radiografia del torace portatile



Qual è l'esame diagnostico più utile nella valutazione e nella gestione dell'insufficienza respiratoria acuta di questo paziente?

- 1- Scansione di perfusione ventilatoria
- 2- Livello di proteina C-reattiva
- 3- Cateterizzazione dell'arteria polmonare
- 4- Livello del peptide natriuretico cerebrale
- 5- Soltura del tratto respiratorio inferiore