

E se i dati fossero tutto ciò che abbiamo?

2 Marzo 2017

Tecniche di Machine Learning e Deep Learning per sviluppare un approccio nuovo e originale per l'integrazione di dati fenotipici e genotipici nel disturbo dello spettro autistico.

Quando uso espressioni come modello predittivo, analisi di dati e statisticamente significativo, i miei colleghi matematici solitamente storcono il naso. Per un matematico, la statistica applicata e il data mining non possono essere considerate materie logiche e razionali alla stregua di algebra, geometria, analisi... Se il tuo lavoro consiste nel trovare modelli in grandi moli di dati, non sei un matematico, sei un mago. Tanto varrebbe estrarre un coniglio da un cappello.

L'incertezza e la mancanza di dimostrazioni rigorose spaventano. Come matematica tendo a condividere lo stesso pensiero. Ma poi mi domando: "E se i dati fossero tutto quello che abbiamo?" La natura è una struttura imperfetta e gli esseri umani, come manifestazione di essa, sono anche loro imperfetti. Pertanto risulta difficile applicare con successo ai fenomeni sociali e naturali un sistema rigoroso come la matematica. Tuttavia, alcuni eventi possono essere compresi attraverso procedimenti matematici (per esempio la diffusione di una epidemia). Inoltre, informazioni raccolte nel passato possono risultare utili nel predire avvenimenti futuri (tutti sanno che le previsioni meteorologiche non sono [quasi mai] previsioni magiche). La matematica applicata offre quindi la possibilità di indagare i fenomeni naturali, individuando regole e modelli dove non sembrano essercene. Si tratta indubbiamente di un lavoro duro e, purtroppo, gli errori sono all'ordine del giorno. Sono però fermamente convinta che, impiegando gli strumenti giusti, usati con cautela e cognizione di causa, sia possibile giungere ad importanti scoperte scientifiche.

Prediamo come esempio il Disturbo dello Spettro Autistico (ASD). Alla stregua di altri disturbi del neurosviluppo, come la schizofrenia, la sua eziologia non è ancora nota. Il Manuale Diagnostico e Statistico dei Disturbi Mentali, quinta edizione, (DSM-5) definisce l'ASD come un disturbo caratterizzato da (1) deficit nella comunicazione sociale e nell'interazione sociale, (2) modelli di comportamento ristretti e ripetitivi, mentre lo sviluppo atipico del linguaggio è considerato in comorbidità. La diagnosi viene effettuata solitamente entro i tre anni di età. La caratterizzazione del disturbo dello spettro autistico ha subito diverse modifiche nel tempo e attualmente ASD indica un disordine complesso ed eterogeneo. Studi genetici hanno individuato diverse mutazioni che interferiscono con il neurosviluppo prenatale e infantile normotipico, mentre studi di neuroimaging hanno portato a importanti scoperte riguardanti i cambiamenti patologici in alcune zone cerebrali dei soggetti con ASD.

Il disturbo dello spettro autistico è quindi caratterizzato da eterogeneità fenotipica e genotipica, inoltre le caratteristiche diagnostiche non sono fisse nel corso dello sviluppo.

Risulta quindi problematico collegare i dati psicometrici cross-sectional ai dati di espressione genica, poiché le possibili relazioni dipendono dall'istante temporale nel quale i dati comportamentali sono raccolti. Una possibile soluzione a questo problema è fornita dalle traiettorie di sviluppo, che meglio caratterizzano il fenotipo dei soggetti con ASD.

E qui incominciano i problemi... Innanzitutto è necessario avere databases longitudinali, in grado di fornire una testimonianza dei diversi profili fenotipici. Questi profili psicometrici devono poi essere integrati con i dati genetici. Infine è importante sviluppare metodi analitici in grado di individuare i diversi profili e metterli in collegamento con le mutazioni genetiche. Per affrontare un approccio così complesso **diventa quindi necessario un tipo di ricerca interdisciplinare**.

Dato il mio interesse per la matematica applicata (soprattutto in ambito medico e biologico) che mi accompagna fino dai primi anni di studio all'Università di Pisa, ho iniziato a collaborare con il Dipartimento di Psicologia e Scienze Cognitive (Università di Trento) in ambito di tesi magistrale. Questa collaborazione è proseguita anche in seguito alla laurea ed è stato in quel periodo che ho iniziato a studiare le difficoltà collegate al disturbo dello spettro autistico.

Attualmente sono una studentessa di dottorato in Scienze Psicologiche e della Formazione (Università di Trento), con borsa finanziata dalla Fondazione Bruno Kessler (FBK). In questo modo ho la possibilità di unire le due discipline che devono necessariamente "comunicare" per fare chiarezza sull'eziologia dell'ASD: psicologia e matematica. Da una parte sto imparando come i dati psicometrici sono collezionati e interpretati dai colleghi del Laboratorio di Osservazione, Diagnosi e Formazione (ODFLab) a Rovereto. Quali strumenti di valutazione vengono utilizzati per la diagnosi di ASD? Quali interventi producono miglioramenti comportamentali e dello sviluppo nei soggetti con ASD? Dall'altra, in FBK (Modelli Predittivi per la Biomedicina e l'Ambiente), sto studiando le tecniche di machine learning e deep learning per lo sviluppo di un approccio nuovo e originale per l'integrazione di dati fenotipici e genotipici. Gli aspetti relativi alla biologia computazionale saranno supervisionati dal Centro di Biologia Integrata (CIBIO, Università di Trento).

Gli ultimi due anni spesi a lavorare con gli psicologi sono stati molto stimolanti. Abbiamo un background diverso, usiamo parole diverse e ci avviciniamo ad uno stesso problema con metodi diversi. Passo dopo passo però le differenze possono essere superate, anzi possono diventare un punto di forza nell'affrontare difficili domande di ricerca. Mi è ormai chiaro che, più mi applico in questa materia per me nuova, più sarò in grado di avvicinarmi allo studio dell'ASD da una nuova prospettiva, che include la conoscenza clinica del disordine. Sfortunatamente però una formazione interdisciplinare non è sufficiente, come ho già sottolineato, è fondamentale avere accesso a grandi database (in particolare database longitudinali) e collaborare con ricercatori che condividono lo stesso obiettivo. Solo così si può aumentare la probabilità di individuare le cause genetiche associate all'ASD, riducendo gli errori e i falsi positivi.

Date le scoperte recenti, penso che un approccio matematico possa aiutare nell'individuare i fattori genetici responsabili dei diversi fenotipi associati al disturbo dello spettro autistico.

Spero nel futuro sia possibile creare trattamenti individualizzati per i soggetti con ASD, trovando le opzioni terapeutiche migliori. Per fare questo è necessario un approccio interdisciplinare rigoroso, prestando particolarmente attenzione alla riproducibilità dei risultati.

<https://magazine.fbk.eu/it/news/e-se-i-dati-fossero-tutto-cio-che-abbiamo/>

AUTORI

- Isotta Landi