# CLINICI BASATI SULLA INTELLIGENZA ARTIFICIALE(MACHINE LEARNING): COME INTERPRET

**Data** 23 maggio 2021
**Categoria** Medicinadigitale

**BREVE GUIDA PER NON ESPERTI**

[b]Cosa è il Machine Learning[/b]
Il "Machine Learning" o apprendimento automatico è una branca importantissima ed in continuo sviluppo della Intelligenza Artificiale:in buona sostanza si tratta di hardware e software che partendo dai dati inseriti dai ricercatori rilevano correlazioni od incompatibilità e propongono ipotesi di approfondimento o addirittura soluzioni ai quesiti posti.
Negli ultimi anni si sono affermati i dispositivi di "Deep Learning" nei quali potenti computer, basati su reti neurali organizzate su molteplici strati interattivi, lavorano senza supervisione e propongono risultati ricavati da milioni o miliardi di computazioni, impossibili da controllare per noi umani: i risultati di questa processazione dei dati dovrebbero quindi essere valutati con grande attenzione anche mediante proiezioni e previsioni a lungo termine

Machine Learning e Deep Learning producono risultati (output) che sono strettamente correlati ai dati introdotti (input), alle loro caratteristiche strutturali e alle modalità di addestramento utilizzato per le macchine. La velocità di elaborazione, l'enorme quantità dei dati prodotti, e talora la sorprendente originalità dei risultati possono indurre facilmente a una sopravvalutazione di questi dispositivi, erroneamente ritenuti infallibili. Inoltre, la naturale tendenza della mente umana a ottenere il miglior risultato possibile con il minimo sforzo, ci porta ad accettare acriticamente, come "sapienza oracolare" le conclusioni delle intelligenze artificiali.
L'autorevole rivista americana JAMA e il prestigioso BMJ inglese ci hanno proposto preziose indicazioni metodologiche che ci aiutano a valutare gli studi in medicina effettuati con le metodiche della intelligenza artificiale: **qualunque siano le conclusioni tratte dai dispositivi di Intelligenza Artificiale noi dovremmo sottoporle ad una analisi critica basata su questi criteri:**

**Validità dello studio**
**Cosa valutare?**
1) Sono chiaramente descritte le metodologie le procedure utilizzate nello studio? Sono corrette?
2) Le procedure e i dati sono stati controllati anche da esperti umani o solo dai dispositivi di intelligenza artificiale?
3) Ci sono riferimenti e confronti con studi già effettuati sul tema e considerati affidabili?
4) L'algoritmo di addestramento delle macchine è controllabile ed è coerente con gli obiettivi di salute che si propone lo studio?

**Risultati dello Studio**
1) Lo studio modifica e arricchisce in qualche modo le nostre conoscenze su quell'argomento?
2) Lo studio è riproducibile nelle medesime condizioni ed è trasferibile in realtà diverse e su popolazioni di pazienti diversi da quelli studiati?
3) Nel caso in cui apporti nuove conoscenze, lo studio include criteri di verifica e di falsificabilità ovvero vi sono eventi che verificandosi possono smentirlo?
4) Nel caso in cui i risultati dello studio siano verificabili, affidabili e riproducibili, va ulteriormente valutata la sua applicabilità nella specifica realtà in cui operiamo e nei confronti della popolazione di pazienti che seguiamo…

**RiflessioniFinali**
Il criterio di falsificabilità, proposto da Karl Popper quale criterio di validazione scientifica di tutte le nuove acquisizioni, è un importante strumento di controllo anche nei confronti dei risultati forniti dai dispositivi di deep learning: se, ad esempio, il sistema di intelligenza artificiale stima che un nuovo farmaco attivo sul coronavirus possa eradicarlo in una percentuale compresa tra il 60% e l'80% dei casi, sarà certamente possibile in un arco di tempo ragionevole confermare o smentire la ipotesi di partenza.
Se invece lo stesso sistema di intelligenza artificiale stima la efficacia di un farmaco sulla base di endpoint surrogati (ad esempio, riduzione della glicemia e/o della colesterolemia quali sostituti di una diminuzione delle complicanze cardiovascolari), dovremmo accettare con molta prudenza queste conclusioni, in quanto non basate su prove inconfutabili.
A chi avesse dubbi in merito ricordiamo le tristi esperienze dei glitazioni per la cura del diabete e della cerivastatina come ipocolesterolizzante.
E' quindi di grandissima importanza, di fronte al moltiplicarsi delle ricerche basate sulla intelligenza artificiale, applicare un metodo di lettura e di valutazione ispirato ai medesimi criteri di critica rigorosa che la nostra imperfetta mente da secoli ha utilizzato per migliaia di studi e sperimentazioni, che hanno finora garantito alla umanità un mondo imperfetto ma comunque migliore.

**Riccardo De Gobbi e Giampaolo Collecchia**

**Bibliografia**
1) Angus DC. Randomized clinical trials of artificial intelligence. JAMA 2020; 323(11): 1043-5.
2) Matheny ME, Whicher D, Thadaney Israni S. Artificial intelligence in health care: a report from the National Academy of Medicine. JAMA 2020; 323(6): 509-10.
3) Liu Y, Chen PC, Krause J, Peng L. How to read articles that use machine learning: users' guides to the medical literature. JAMA 2019; 322(18): 1806-16.
4) Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ 2020; 368: m689.
5) Popper K. Logica della scoperta scientifica. Torino: Einaudi, 2010; 66-83.
6) Vollmer S, Mateen BA, Bohner G, et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. BMJ 2020; 368: l6927.

Perapprofondire:
**Collecchia G. De Gobbi R.: Intelligenza Artificiale e Medicina Digitale. Una guida critica. Il Pensiero Scientifico Ed. Roma 2020**
**pensiero.it/catalogo/libri/pubblico/intelligenza-artificiale-e-medicina-digitale**