

# Strumenti innovativi di IA e Big Data in sanità: opportunità e problematiche

Mauro Grigioni, Evaristo Cisbani, Antonella Rosi, Giuseppe D'Avenio

Centro Nazionale Tecnologie Innovative in Sanità Pubblica, Istituto Superiore di Sanità, Roma  
mauro.grigioni@iss.it, evaristo.cisbani@iss.it, antonella.rosi@iss.it, giuseppe.davenio@iss.it

**Abstract** La sempre maggiore disponibilità di grandi quantità di dati combinata a strumenti computazionali sempre più performanti ha dato un forte impulso allo sviluppo e diffusione di applicazioni basate su tecnologie di Intelligenza Artificiale (IA). Le opportunità offerte da questi strumenti nella sanità sono brevemente richiamate insieme alle problematiche che essi portano con sé, specificatamente dal punto di vista della *governance* e delle infrastrutture necessarie perché vengano colte le opportunità, nell'ambito delle caratteristiche di equità e universalità del SSN.

## 1 Introduzione

Negli ultimi anni si assiste a un crescente interesse verso le applicazioni di IA in medicina, date le potenzialità in vari ambiti, quali la patologia computazionale, la neurologia e le scienze radiologiche, in particolare per la diagnosi per immagini: gli algoritmi di analisi di immagine possono rivelare caratteristiche che l'occhio non è in grado di valutare: differenze di intensità non apprezzabili neanche da un clinico esperto sono quantificabili da molti strumenti software automatici, peraltro in grado di offrire performance costanti, indipendenti dal carico di lavoro.

L'imaging diagnostico è uno dei campi più in sviluppo per applicazioni di IA, ma non l'unico; un ambito altrettanto fertile è quello dell'analisi delle cartelle cliniche (e dei dati medici più in generale), che contengono una quantità enorme di informazioni (solo parzialmente strutturate), ancora sfruttate solo in minima parte. L'integrazione con i dati derivati dalla genomica e da altri ambiti clinici può essere di grande ricaduta medica (ad es., possibilità di effettuare studi longitudinali su molteplici categorie di pazienti, ricavando indici predittivi di importanti patologie).

Accanto a queste applicazioni in cui l'IA non interagisce direttamente con il paziente, si assiste oggi alla proliferazione di dispositivi robotici che partecipano direttamente alla cura del paziente, dalla compagnia alle persone anziane all'intervento chirurgico [Hamet, 2017].

Il paradigma più comune di IA è probabilmente quello che raggruppa le tecniche di Machine Learning (ML) il cui sviluppo più avanzato è rappresentato dal Deep Learning (DL) in cui si usa una rete neurale (NN) di decine o anche centinaia di strati di "neuroni", o unità elementari di elaborazio-

ne. La complessità delle architetture di DL fa sì che le stesse si comportano come "scatole nere" la cui descrizione teorica resta ad oggi largamente lacunosa: in pratica è impossibile individuare il meccanismo esatto per cui il sistema fornisce certe risposte: il meccanismo non-lineare è generato dalla stessa DL a differenza di quanto avviene nei sistemi esperti tradizionali.

Questo problema - in prospettiva sempre più critico, con il prevedibile sviluppo delle soluzioni DL - ha determinato, tra l'altro, la comparsa del concetto di *Explainable AI* (XAI): è importante poter individuare il percorso fatto da un sistema IA per arrivare a certe conclusioni ed essere sicuri che i "ragionamenti" eseguiti siano condivisibili. Nello specifico delle applicazioni alla salute, l'utilizzo di "scatole nere" certamente pone problematiche non banali nel definire il percorso di certificazione, di *quality assurance*, di standardizzazione, per potenziali dispositivi medici basati su DL.

## 2 I dati

Gli strumenti di ML (e DL) richiedono una o più fasi di apprendimento con dati quanto più comprensivi ed esaustivi dell'ambito di applicazione; seppur non un requisito indispensabile, la disponibilità di "Big Data" favorisce lo sviluppo di tecnologie di AI e per certi aspetti lo impone.

Con il termine "dati" si intendono oggetti (o categorie di oggetti) molto diversi fra loro. Nel campo dell'imaging radiologico, i dati sono costituiti da immagini registrate in determinate condizioni strumentali, dopo aver verificato che l'apparecchiatura (ad es. nella tomografia computerizzata, CT o ad emissione di positroni PET) fosse rispondente a certi requisiti. Evidentemente, ogni analisi che abbia come obiettivo il rigore scientifico deve partire da una base di informazioni coerenti: non sembra proponibile il sacrificio dei requisiti metodologici delle registrazioni diagnostiche a favore di una maggiore quantità di dati clinici da fornire ai tools ML/DL. Questo aspetto è rilevante rispetto alle tecniche di addestramento delle reti e al loro accrescimento: non sarebbero sufficienti i dati relativi ad un solo centro servizi sanitario. È auspicabile in tal senso la disponibilità dell'insieme dei dati clinici relativi a singole patologie, così da realizzare un addestramento continuo su base nazionale, laddove l'equità e l'universalità del supporto tecnologico in ambito diagnostico-terapeutico è una richiesta del SSN. Al contempo localmente le specificità genomiche e cliniche del

paziente permettono poi di utilizzare sistemi di supporto alla decisione in grado di personalizzare le terapie.

Occorre sottolineare la differenza fra i concetti di “dato” e “informazione”. Come affermato in [Rasetti, 2018] “Quello di trasformare i dati in informazione; l’informazione in conoscenza e, infine, la conoscenza in ‘sapere’ è uno dei più difficili problemi mai affrontato dalla scienza. È il *data mining*; estrarre valore dai dati, come un minatore estrae minerali preziosi dal terreno.” L’utilizzo dei dati in IA, in un settore delicato come quello della salute, amplifica le problematiche di privacy, tanto più quanto si combinano dati sensibili (condizioni di salute) con dati non sensibili (stile di vita) [Spina, 2018].

### 3 Qualità dei dati e dei processi in radiologia

L’imaging biomedico continua ad essere un pilastro del trattamento clinico, soprattutto in campo oncologico. I recenti sviluppi dei dispositivi di imaging - in termini di qualità, sensibilità e risoluzione - consentono di discriminare minime differenze di densità tissutali e di generare una mole di dati difficilmente utilizzabili anche da specialisti di grande esperienza se non attraverso un cambio di paradigma verso strumenti più potenti, tipici dell’IA. In contrasto con i metodi tradizionali basati su caratteristiche predefinite, gli algoritmi del DL risultano tanto più prestazionali quanto maggiore è il numero di dati generato. Questo importante sviluppo impatta fortemente sulla compliance a programmi di garanzia di qualità in termini di personalizzazione dei trattamenti, riduzione del numero di esposizioni (ogni immagine può essere sfruttata al massimo del suo contenuto informativo riducendo la necessità di immagini ripetute), riduzione del numero di attività di routine e conseguente riduzione del carico di lavoro. L’ISS produce linee di indirizzo in questo settore da oltre 20 anni, grazie all’attività di Assicurazione di Qualità nelle scienze radiologiche.

### 4 IA e Sistema Sanitario Nazionale

Il superamento delle disuguaglianze in sanità, già affermato in sede di istituzione del SSN, è un obiettivo ribadito dal [Patto, 2014]. A tale proposito, il processo di diffusione di strumenti IA in sanità, se non governato opportunamente, potrebbe aumentare le distanze fra regioni e la spesa, non ottenendo una coerenza infrastrutturale nel paese.

Data l’importanza dei dati in IA, si può infatti prevedere che le zone più organizzate dal punto di vista sanitario siano anche dotate di archivi informatici più efficienti e con maggiore numerosità e varietà di dati clinici e pertanto risultano più propense a raccogliere i maggiori benefici da questo processo, almeno nella fase iniziale, con una probabile perdita di equità del sistema. Appare quindi ragionevole pensare a forme di compensazione, nel senso di pianificare la fruizione delle conoscenze tratte dalle esperienze migliori a contesti meno avvantaggiati.

Inoltre, tenuto conto della continua e rapida evoluzione degli strumenti di IA, la possibilità di una scarsa sincronizza-

zione delle prestazioni delle varie piattaforme, a livello nazionale come anche regionale, potrebbe compromettere il requisito dell’accesso equo alle cure. D’altra parte, come già accennato, la gestione dei dati in sanità ha un forte impatto organizzativo-gestionale (necessità di progettare architetture informatiche efficienti ed espandibili), come anche dal punto di vista della sicurezza dei dati, la cui condivisione (anche al fine di uniformare l’offerta dello SSN) deve essere ovviamente in accordo con la normativa applicabile (es. [GDPR, 2016]).

### 5 Il quadro regolatorio

Negli USA è già in atto un’iniziativa della FDA per trattare la certificazione di software (SW) derivato da IA, come dispositivo medico (DM) [SPP, 2019]; in questo documento, si riconosce che le specificità della produzione di SW non sono bene inquadrabili nel tradizionale approccio alla verifica della sicurezza e delle prestazioni dei DM. Pertanto, la FDA propone un paradigma regolatorio agile per rispondere al ritmo accelerato di sviluppo e al potenziale innovativo dei prodotti software.

In Europa, l’approccio tuttora seguito – sia con la Direttiva 93/42/CEE sia con il nuovo Regolamento 2017/745 - è invece quello di trattare il SW come un DM vero e proprio, senza “sconti” per la sua natura peculiare. Evidentemente, questo impone al fabbricante un onere superiore rispetto a quanto preventivabile nel prossimo futuro per l’immissione dei SW sul mercato USA.

Sono quindi prevedibili delle criticità nella certificazione di prodotti DL in sanità in Europa, tenuto conto che buona parte di tali prodotti, per quanto ricordato sopra, si comportano essenzialmente come “scatole nere”, non trasparenti rispetto alle analisi eseguite, con stime non deterministiche. Metodi innovativi di progettazione e test e strumenti normativi adeguati saranno sicuramente dei requisiti necessari per la diffusione di strumenti IA in sanità.

### Riferimenti bibliografici

- [GDPR, 2016] General Data Protection Regulation, Regolamento UE 2016/679
- [Hamet, 2017] Hamet P. Tremblay Johanne, Artificial intelligence in medicine, *Metabolism* 69 (2017) 536-540
- [Patto, 2014] Intesa Stato Regioni 10 luglio 2014
- [Rasetti, 2017] <https://www.sicurezzanazionale.gov.it/sisr.nsf/comunicazione/decennale-intelligence/rasetti.html>
- [SPP, 2019] *Software Precertification Program: Working Model – Version 1.0* (January 2019) <https://www.fda.gov/downloads/MedicalDevices/DigitalHealth/DigitalHealthPreCertProgram/UCM629276.pdf>
- [Spina, 2018] A. Spina La medicina degli algoritmi: Intelligenza Artificiale, medicina digitale e regolazione dei dati personali, in F. Pizzetti (a cura di) *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino 2018