



Ministero della Salute

Consiglio Superiore di Sanità

Sessione LII (2019-2022)

Presidente: Prof. Franco Locatelli

Sezione V*

*Presidente: Prof. Giuseppe Remuzzi
Segretario tecnico: Dr. Franco Abbenda*

***“I sistemi di intelligenza artificiale come strumento
di supporto alla diagnostica”***

Coordinatore: Prof. A. Laghi

9 novembre 2021

* Il lavoro è stato condotto in collaborazione con gli Esperti designati dalla Sezione I del CSS.

INDICE

| | |
|--|----|
| EXECUTIVE SUMMARY | 3 |
| 1. INTRODUZIONE | 5 |
| 2. DEFINIZIONE E ASPETTI TECNOLOGICI | 6 |
| 2.1 AI in Medicina..... | 7 |
| 2.2 Classificazione..... | 8 |
| 2.3 Metodologie per la diagnostica..... | 9 |
| 2.4 Passi per la realizzazione di sistemi AI basati su Machine Learning (ML) | 12 |
| 2.5 Metodologie e tecnologie abilitanti..... | 15 |
| 3. STATO DELL'ARTE DEI SISTEMI DI AI IN DIAGNOSTICA | 18 |
| 3.1 Principali applicazioni dell'AI in sanità a supporto della diagnostica..... | 18 |
| 3.2 Considerazioni metodologiche e cliniche..... | 20 |
| 3.3 Considerazioni tecniche..... | 22 |
| 3.4 Criticità e sfide future..... | 24 |
| 4. ASPETTI ETICI E LEGALI | 27 |
| 4.1 Il quadro normativo..... | 27 |
| 4.2 Le diverse tipologie di sistemi tra autonomia e automazione..... | 29 |
| 4.3 Le implicazioni nella relazione medico-paziente. Il consenso informato..... | 30 |
| 4.4 Raccolta, trattamento e condivisione dei dati personali sanitari..... | 32 |
| 4.5 La disciplina della responsabilità nei rapporti tra produttore e operatore/utilizzatore...33 | |
| 5. CONCLUSIONI E PROPOSTE OPERATIVE | 36 |
| 6. APPENDICE | 38 |
| 6.1 Tabella 1. Criteri di inclusione della metodologia di ricerca..... | 38 |
| 6.2 Tabella 2. Tipologia di modellizzazione con esempi pratici di utilizzo del ML in ambito diagnostico dalla letteratura..... | 38 |
| 7. GRUPPO DI LAVORO | 41 |

EXECUTIVE SUMMARY

L'Intelligenza Artificiale (*Artificial Intelligence, AI* nella dizione anglosassone), in questi ultimi anni, ha rivoluzionato la quotidianità dei cittadini e, considerando la velocità dello sviluppo tecnologico, ne modificherà i comportamenti e le abitudini anche nel prossimo futuro.

Il mondo sanitario in generale, e quello della diagnostica per immagini in particolare, seppur con comprensibile ritardo data la maggior complessità dei processi, sta subendo il medesimo sconvolgimento. La prospettiva, più o meno a breve termine, sarà di doversi confrontare con sistemi esperti in grado di modificare significativamente i percorsi diagnostici e terapeutici, le modalità decisionali del Medico e, in ultimo, anche il rapporto Medico-Paziente.

Uno sviluppo incontrollato e non governato dell'AI non è scevro da potenziali rischi, derivanti, ad esempio: 1) dall'uso di sistemi di AI privi di una rigorosa validazione scientifica; 2) dalla mancanza di controllo sulla modalità di processazione dei dati da parte dei sistemi esperti; 3) da possibili violazioni della privacy degli utenti; 4) da discriminazioni (ad esempio di razza e/o di genere) introdotte dalla programmazione degli algoritmi; 5) dall'assenza di informazioni circa la sicurezza e la riproducibilità nell'uso dei sistemi di AI; 6) dalla mancanza di norme circa la responsabilità professionale del Medico nell'interazione con gli algoritmi; 7) dalla impreparazione del personale Medico e sanitario al corretto utilizzo dei sistemi di AI e alla appropriata modalità di comunicazione del loro utilizzo ai pazienti; 8) dall'incomprensione da parte dell'utente/cittadino dei reali benefici e limitazioni dei sistemi di AI.

Per poter introdurre in modo sicuro nella pratica clinica i sistemi di AI e per competere in ambito internazionale nella programmazione e nello sviluppo degli stessi, è auspicabile che, nel nostro Paese, siano attuati i seguenti interventi:

- realizzazione di infrastrutture organizzative, informatizzate, a livello locale, regionale o nazionale, di *data stewardship* e *data governance*;
- creazione di una struttura di *governance* dei sistemi di AI da parte delle agenzie regolatorie italiane, in particolare il Ministero della Salute per ciò che riguarda i dispositivi medici e AIFA per gli eventuali aspetti terapeutici, con lo scopo di stabilire delle regole rigorose per l'approvazione e la registrazione di tali sistemi;
- predisposizione di Linee Guida nazionali riguardanti le modalità di integrazione e il corretto utilizzo dei sistemi di AI nella diagnostica, in accordo con le società scientifiche di riferimento;
- creazione di un osservatorio nazionale permanente presso il Ministero della Salute, per il monitoraggio delle performance dei sistemi di AI immessi sul mercato (analisi *post-market*);
- predisposizione di moduli formativi universitari e post-universitari per migliorare le conoscenze e competenze in materia di AI del personale medico e delle professioni sanitarie;

- integrazione di elementi metodologici in tema di AI all'interno dei programmi della scuola secondaria superiore e creazione di contenuti informativi, anche tramite canali informatici, al servizio del cittadino.

1. INTRODUZIONE

L'Intelligenza Artificiale (*Artificial Intelligence*, *AI* nella dizione anglosassone) si sta rapidamente diffondendo nella nostra Società, a un ritmo probabilmente più veloce di ogni ottimistica previsione. Ogni giorno, anche inconsapevolmente, utilizziamo tecnologie basate sull'AI, che ci facilitano compiti, più o meno complessi, della vita quotidiana: dalla navigazione in automobile, alla ricerca di un ristorante o al suggerimento all'acquisto di un determinato bene di consumo.

Il mondo sanitario in generale, e quello della diagnostica in particolare, subisce, allo stesso modo, questo processo di sviluppo. Ad esempio, tecnologie basate sull'AI controllano già oggi le grandi apparecchiature di diagnostica per immagini (Tomografia Computerizzata, TC o Risonanza Magnetica, RM), standardizzando i protocolli di acquisizione e riducendo drasticamente i tempi di acquisizione degli esami, al fine di migliorare la *compliance* e il *comfort* dei Pazienti. Esistono anche algoritmi in grado di supportare il Radiologo nell'identificazione di patologie (ad esempio, noduli mammari o polmonari, fratture ossee, pneumotorace, ecc), riducendo al minimo gli errori di percezione, oppure aiutare lo specialista nella caratterizzazione di lesioni (ad esempio, il Dermatologo con il melanoma cutaneo), con lo scopo di migliorare la diagnosi. Le ultime evidenze scientifiche riportano, inoltre, studi sull'uso di sistemi esperti "intelligenti", in grado di categorizzare il rischio del Paziente, determinare una prognosi e supportare il medico nella successiva decisione terapeutica.

Questo tumultuoso sviluppo rischia, d'altra parte, di essere incontrollato e fonte di potenziali rischi, derivanti, ad esempio, dall'uso di sistemi di AI privi di una rigorosa validazione scientifica, dalla mancanza di controllo sui dati processati dai sistemi esperti, da possibili violazioni della privacy degli utenti e da discriminazioni introdotte dalla programmazione degli algoritmi; senza dimenticare le aspettative illusorie e fuorvianti per sanitari e pazienti derivanti da un utilizzo improprio dei sistemi di AI¹. Si aprono, quindi, nuove problematiche etiche² e legali, che coinvolgono la responsabilità professionale e il ruolo del Medico nell'interazione con i sistemi di AI, arrivando anche a modificare potenzialmente il rapporto Medico-Paziente.

La Commissione Europea ha pubblicato un Libro bianco sulla strategia europea per l'AI, presentato nel 2018³, e significativamente intitolato "*Libro bianco sull'Intelligenza Artificiale. Un approccio europeo all'eccellenza e alla fiducia*"⁴, che ribadisce l'importanza strategica della tecnologia dell'AI, in grado di offrire indubbi benefici ai cittadini, anche, ma non solo, nel campo della salute, purché si "*segua un approccio antropocentrico, etico, sostenibile e rispettoso dei valori e dei diritti*

¹ Laghi A. Cautions about radiologic diagnosis of COVID-19 infection driven by artificial intelligence. *Lancet Digit Health*. 2020 May

² Intelligenza Artificiale e Medicina: aspetti etici, 29 maggio 2020. <http://bioetica.governo.it/it/pareri/pareri-gruppo-misto-cnbcnbbsv/intelligenza-artificiale-e-medicina-aspetti-etici/>

³ AI for Europe, COM/2018/237 final

⁴ WHITE PAPER. On Artificial Intelligence - A European approach to excellence and trust. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

fondamentali". A seguito della pubblicazione del Libro Bianco, la Commissione Europea, anche su esplicita richiesta del Parlamento Europeo, ha lanciato un'ampia consultazione con le parti interessate, al fine di armonizzare le leggi europee in tema di AI, che ha prodotto un documento reso pubblico nel mese di aprile 2021⁵. L'obiettivo è quello di assicurare un mercato interno europeo dell'AI ben funzionante e nel quale siano adeguatamente valutati i benefici e i rischi dei software di AI, che devono essere sviluppati seguendo i principi di sicurezza, affidabilità ed etica propri dell'Unione Europea.

Anche il nostro Paese, nelle "Proposte per una strategia italiana per l'intelligenza artificiale"⁶, elaborate dal gruppo di esperti del Ministero dello Sviluppo Economico (MISE) sull'AI e pubblicate nel luglio del 2020, ribadisce gli stessi concetti, senza entrare, però, in dettaglio nell'area della Medicina.

Il Consiglio Superiore di Sanità ha ritenuto, pertanto, opportuno approfondire il tema dei sistemi di AI in diagnostica medica, partendo dall'analisi delle attuali evidenze scientifiche, regolatorie ed etiche, con lo scopo di formulare delle proposte operative per introdurre in modo sicuro nella pratica clinica i sistemi di AI e per competere in ambito internazionale nella programmazione e nello sviluppo degli stessi.

2. DEFINIZIONE E ASPETTI TECNOLOGICI

Due semplici definizioni di AI, tratte dal noto *Merriam Webster Dictionary* e discusse in Combi⁷, sono:

- l'AI è la branca dell'informatica che studia la simulazione di comportamenti intelligenti nei computer;
- con il termine AI si intende la capacità di una macchina di imitare il comportamento umano intelligente.

Tali definizioni, pur nella loro semplicità, permettono di avvicinarci ad alcune caratteristiche essenziali dell'AI, che, al centro, ha:

- 1) l'ottenimento per via algoritmica di risultati, come l'estrazione di contenuto informativo e la produzione di nuova conoscenza, da insiemi di dati di grande mole e complessità, spesso al di là delle possibilità del singolo essere umano;
- 2) le attività, sia astratte sia concrete, che implicano la produzione di giudizi e/o predizioni

⁵ Commissione europea, Proposal for a Regulation of the European Parliament and of the Council. Laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts, 21 aprile 2021, COM (2021) 206.

⁶ "Proposte per una strategia italiana per l'intelligenza artificiale", Elaborata dal Gruppo di Esperti MISE sull'intelligenza artificiale, 2020

⁷ Combi C. Editorial from the new Editor-in-Chief: Artificial Intelligence in Medicine and the forthcoming challenges. *Artif Intell Med.* 2017; 76:37-39. doi: 10.1016/j.artmed.2017.01.003

2.1 AI in medicina

L'AI in Medicina (AI M) può dunque essere definita attraverso la declinazione dei comportamenti intelligenti che caratterizzano la Medicina e le attività cliniche correlate.

L'AI M comprende, dunque, tutti gli ambiti dove la conoscenza medica richiede di essere rappresentata ed estesa attraverso differenti tipi di ragionamento. D'altro canto, l'AI M considera i comportamenti intelligenti, che sono alla base delle numerose attività decisionali in Medicina, quali la diagnosi, la terapia, la prognosi e la gestione del monitoraggio del paziente. Tali attività, oltre a essere caratterizzanti della pratica clinica, prevedono la capacità di fondere e usare conoscenze di base, conoscenze specifiche sul paziente, sul contesto ambientale e altro, al fine di prendere, entro un lasso di tempo accettabile, la migliore decisione possibile rispetto all'evoluzione dello stato di salute del paziente (o di interi gruppi di pazienti).

I primi studi, e relativi sistemi software, inerenti l'AI M sono comparsi nei primi anni '70. In quel periodo, e negli anni successivi, l'attenzione si è concentrata su tecniche e software in grado di simulare il ragionamento umano rispetto a una conoscenza medico-clinica opportunamente rappresentata in modo formale. Gli ambiti di interesse riguardavano dunque il supporto alle decisioni cliniche in specifici ambiti di applicazione (ad esempio, MYCIN, un sistema esperto sviluppato a Stanford negli anni '70, si proponeva di supportare i clinici nella definizione di terapie per pazienti con infezioni). Con l'avvento di computer sempre più potenti, l'evoluzione dell'informatica, e la disponibilità di enormi moli di dati grazie a internet e alle piattaforme "social" ad esso legate, l'interesse dell'AI M si è spostato da applicazioni basate su grandi moli di conoscenza ad applicazioni basate su grandi moli di dati⁸. In modo corrispondente, le applicazioni (chiamate sistemi esperti) che si proponevano di supportare i medici con ragionamenti e relativi suggerimenti basati su conoscenza specialistica sono state sostituite da applicazioni che si propongono di supportare le attività decisionali dei medici attraverso l'individuazione veloce e fruibile delle informazioni rilevanti. Con il termine "data mining" si intende un insieme di attività il cui scopo è estrarre dai dati conoscenza e informazioni in essi nascoste o, in altre parole, implicite⁹. Fra queste attività ricordiamo fra le principali: la *classificazione*, che consiste nella collocazione di oggetti all'interno di un sistema di classi predefinito, sulla base delle informazioni disponibili; la *clusterizzazione*, che permette di individuare gruppi omogenei di oggetti; la derivazione di *regole di associazione*, per estrarre associazioni ricorrenti nei dati; la *predizione*, che permette di stimare eventi futuri sulla base dei dati disponibili, descrittivi di quanto accaduto nel passato; *l'analisi what-if*, che permette di supportare ragionamenti ipotetici rispetto

⁸ Combi C, Keravnou-Papailiou E, Shahar Y. Temporal information systems in medicine. Springer Science & Business Media; 2010

⁹ Patel VL, Shortliffe EH, Stefanelli M, Szolovits P, Berthold MR, Bellazzi R, Abu-Hanna A. The coming of age of artificial intelligence in medicine. *Artif Intell Med* 2009;46(1):5-17, <http://dx.doi.org/10.1016/j.artmed.2008.07.017>

a differenti situazioni prospettate; l'estrazione di *pattern*, o configurazioni ricorrenti, da grandi moli di dati, spesso caratterizzate da dimensioni temporali.

2.2 Classificazione

Differenti tecniche e formalismi sono stati proposti nell'ambito dell'AI per implementare le varie attività di *data mining*, di rappresentazione della conoscenza e ragionamento, con importanti applicazioni in Medicina. Senza la pretesa di essere completi, nel seguito si vuole distinguere le tecniche studiate rispetto ad una classificazione comunemente usata.

Approcci simbolici

Tali approcci si propongono di rappresentare i problemi di volta in volta considerati attraverso formule matematiche e/o logiche, stringhe di caratteri opportunamente strutturati, reti ed altre strutture di dati, e di applicare ad essi algoritmi in grado di estrarre contenuto informativo, ossia di derivare nuova conoscenza, di analizzare le informazioni a disposizione e di individuare anomalie, pattern ricorrenti e così via. Fra i sistemi di AI M fondati su un approccio simbolico, vi sono i sistemi basati su regole, i sistemi di inferenza - basati su logiche descrittive, logiche multivalore, logiche *fuzzy*, logiche temporali e altro - i sistemi basati su ontologie, i sistemi di pianificazione e gli alberi decisionali. Tali sistemi sono stati studiati in Medicina, sin dagli anni '70, con applicazioni nella rappresentazione di conoscenza medica, nel supporto all'applicazione di linee guida, percorsi di cura, diagnosi e terapia, nell'estrazione di conoscenza medico-clinica da testo in linguaggio naturale, e, in generale, in ambiti nei quali erano richieste capacità di ragionamento su concetti astratti e strutturalmente complessi. Con tali sistemi è relativamente semplice estrarre una spiegazione di quanto ottenuto o proposto dal sistema.

Approcci sub-simbolici

In questi ultimi anni, mentre l'interesse per i sistemi simbolici di AI M ha continuato a consolidarsi, gli approcci sub-simbolici si sono affermati e hanno riscosso un livello di attenzione estremamente elevato in un breve lasso di tempo. Tali sistemi non si fondano sulla rappresentazione simbolica del problema affrontato, ma si rifanno a paradigmi di elaborazione che cercano di simulare il funzionamento di neuroni e delle loro connessioni, le dinamiche di lavoro cooperativo di colonie di insetti, o che usano metodi numerici che propongono predizioni/classificazioni apprendendo dai dati via via forniti. I vari approcci di questa categoria vengono spesso raccolti sotto il termine "*Machine Learning*" (*apprendimento automatico*), per la loro capacità di apprendere dai dati. Tali approcci hanno riscosso un interesse ed uno sviluppo rapidissimo negli ultimi dieci anni, e sono stati spesso applicati in Medicina per la predizione e classificazione di patologie/rischi basati su grandi moli di dati, per stimare l'evoluzione di fenomeni pandemici, per la diagnosi su immagini e segnali biomedici.

2.3 Metodologie per la diagnostica

La realizzazione di sistemi computerizzati di supporto alla diagnostica clinica è stato uno dei primi temi affrontati dall'AI M, in particolare negli anni '80 e '90 del secolo scorso, mediante la realizzazione di "Sistemi Esperti"¹⁰. Un aspetto particolarmente significativo di questi sistemi riguarda la separazione fra la rappresentazione della conoscenza medica, spesso descritta mediante regole di produzione di tipo IF - THEN, e gli algoritmi per utilizzare questa conoscenza di fronte a un caso clinico. Le implicazioni di questo approccio sono duplici. Da una parte viene richiesto di rappresentare in modo formalizzato ciò che si conosce su uno *specifico dominio clinico*, dall'altra viene richiesto di definire una strategia per usare questa conoscenza al fine di formulare una diagnosi sullo *specifico caso clinico*. Sebbene l'effettiva adozione di questi sistemi nella pratica sia stata molto limitata, l'approccio impiegato dai sistemi esperti ha generato una serie di risultati molto importanti. Le tecniche di rappresentazione della conoscenza vengono impiegate nella realizzazione di sistemi di classificazione, terminologie e ontologie, di cui è possibile trovare una vasta raccolta curata dalla *Open Biological and Biomedical Ontology (OBO) Foundry*¹¹. A titolo di esempio, si può ricordare SNOMED, una complessa terminologia e ontologia di tutti i concetti medici largamente impiegata in anatomia patologica, e la *Gene Ontology*, una raccolta di tre ontologie di concetti legati ai geni e ai loro prodotti tra le più utilizzate nel campo della biologia molecolare. Inoltre, un settore di notevole interesse è costituito dall'implementazione di linee guida informatizzate, incluse almeno in parte nelle cartelle cliniche elettroniche o anche in *app*¹². Da ultimo, ma non per importanza, è necessario ricordare la rappresentazione formale e informatizzata dei flussi di cura dei pazienti di specifiche strutture ospedaliere. Questo tipo di conoscenza clinica, che rappresenta l'operatività di un reparto o di un servizio, è di fondamentale importanza nella realizzazione di sistemi di supporto alle decisioni¹³.

I più recenti successi nel campo dell'AI appartengono all'ambito del *Machine Learning*, il settore che si occupa di creare sistemi che apprendono, ovvero migliorano le proprie prestazioni, alla luce dei dati che sono messi a loro disposizione¹⁴. In particolare, le strategie di maggiore impiego nella diagnostica possono essere ricondotte ai metodi di apprendimento induttivo, in cui vengono perseguiti degli obiettivi macroscopicamente classificabili in tre ambiti:

- i) l'apprendimento supervisionato;

¹⁰ Perry CA. Knowledge bases in medicine: a review. Bull Med Libr Assoc. 1990 Jul;78(3):271-82. PMID: 2203499; PMCID: PMC225405

¹¹ Smith, B., Ashburner, M., Rosse, C. et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nat Biotechnol 25, 1251–1255 (2007). <https://doi.org/10.1038/nbt1346>

¹² Riaño D, Peleg M, Ten Teije A. Ten years of knowledge representation for health care (2009-2018): Topics, trends, and challenges. Artif Intell Med. 2019 Sep;100:101713. doi: 10.1016/j.artmed.2019.101713. Epub 2019 Sep 7. PMID: 31607346

¹³ Adlassnig KP, Combi C, Das AK, Keravnou ET, Pozzi G. Temporal representation and reasoning in medicine: Research directions and challenges. Artif Intell Med. 2006 Oct;38(2):101-13. doi: 10.1016/j.artmed.2006.10.001. Epub 2006 Nov 1. PMID: 17081736.

¹⁴ Mitchell, T. Machine Learning, Mc Graw Hill, 1997

- ii) l'apprendimento non supervisionato;
- iii) l'apprendimento con rinforzo¹⁴.

Nel primo ambito l'obiettivo è mettere in relazione un insieme di variabili misurate, come i dati clinici di un paziente, con una variabile di interesse, come la diagnosi. L'apprendimento è supervisionato perché si suppone che un "supervisore" abbia fornito una base dati che contiene un insieme di casi in cui sono presenti al contempo le variabili misurate e la loro corrispondente variabile di interesse. Se questa è discreta o categorica, ovvero può assumere uno in un insieme di possibili valori, come in un problema diagnostico, il problema è detto di classificazione. Viceversa, se la variabile di interesse ha valori continui, come un parametro clinico quale la glicemia, il problema è detto di regressione. A queste due classi di problemi può essere aggiunta l'analisi di sopravvivenza. Queste tipologie di problemi sono ben note nel campo della statistica; gli sviluppi recenti nel campo del *Machine Learning* hanno tuttavia permesso di ampliare grandemente la natura e la numerosità delle variabili in ingresso rispetto agli approcci noti nella statistica classica, e di ricavare regole di classificazione e di predizione in grado di cogliere aspetti complessi nei dati a disposizione.

Nel campo dell'apprendimento non supervisionato, viceversa, l'obiettivo è trovare delle regolarità nei dati in ingresso, senza che sia stata definita una variabile di interesse a priori. Una classe di approcci molto impiegata nella ricerca biomedica è il *clustering*, che identifica gruppi di casi con caratteristiche simili fra loro (o *cluster*) e che siano sufficientemente distinti dai casi degli altri gruppi. Il *clustering* è stato molto impiegato nel campo della diagnostica molecolare per scoprire sottogruppi in una determinata patologia, ad esempio nei tumori, sulla base dei dati provenienti da metodi di misura *high-throughput*, come il *Next Generation Sequencing*.

Infine, l'apprendimento con rinforzo, utilizzato con successo nell'ambito della robotica, si occupa di sviluppare un agente artificiale che deve raggiungere un obiettivo, e, sulla base degli stimoli forniti dall'ambiente e dall'esito delle sue azioni, apprende la miglior strategia per arrivare all'obiettivo stesso.

Nel campo della diagnostica clinica, i metodi di classificazione supervisionata sono senza dubbio quelli che hanno maggior interesse per la realizzazione di sistemi basati su AI. In questo ambito le metodologie impiegate spaziano da strategie derivanti dalla tradizione più prettamente statistica, come la regressione logistica e i metodi Bayesiani, ad approcci che fanno parte sia della letteratura statistica che di quella informatica, come gli alberi decisionali e le *random forest*. Infine, vi sono strategie maturate nell'ambito del *Machine Learning* stesso, come le *support vector Machines*, i metodi di *boosting* e le reti neurali artificiali¹⁵. Un aspetto decisivo nel successo dell'applicazione di qualunque procedura di tipo diagnostico, e in generale di qualunque metodo di predizione, è legato alla scelta delle

¹⁵ Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, 2008

variabili da misurare sul caso in questione. Negli approcci “tradizionali” del *Machine Learning*, le variabili sono ricavate sia direttamente dai dati clinici (come i dati di laboratorio), sia mediante misure ed estrapolazioni effettuate dagli operatori, come nel caso di osservazioni effettuate nell’ambito dell’*imaging* diagnostico. Negli ultimi anni, uno dei fattori di successo più significativi dell’AI è stata la semplificazione di questo passo di estrapolazione dei dati, mediante l’utilizzo di nuove architetture di reti neurali artificiali denominate collettivamente metodi di *Deep Learning*¹⁶. Questi approcci svolgono congiuntamente due compiti: trasformano le variabili in ingresso e contestualmente effettuano la classificazione (e in generale qualunque compito previsionale). Nella fase di trasformazione, i valori delle variabili sono combinati matematicamente al fine di creare un insieme di variabili latenti. Queste trasformazioni sono in grado di tenere conto di complesse relazioni interne tra le variabili misurate, anche analizzando i dati a diversi livelli di astrazione. Per questo motivo, il *Deep Learning* si è dimostrato estremamente efficace nel trattare dati che tradizionalmente necessitavano di essere prima analizzati manualmente, come immagini, video, segnali e testi. Le componenti chiave delle architetture di *Deep Learning* sono una serie di trasformazioni, o *layers*, progettate per affrontare compiti specifici. Ad esempio, un tipo di architettura ampiamente sfruttata nell’elaborazione delle immagini, è nota come *Convolutional Neural Network*¹⁷. Altre architetture di successo sono quelle impiegate per analizzare sequenze, come le sequenze di caratteri o le serie temporali, tra cui le *Recurrent Neural Networks* (RNNs) e i loro sottotipi, *Gated Recurrent Units* (GRUs) e reti *Long short-term memory* (LSTM)¹⁸. Recentemente, nel campo dell’elaborazione del linguaggio naturale, sono state introdotte con grande successo architetture note come *transformers*^{19, 20}. I modelli di *Deep Learning*, tuttavia, sono caratterizzati da un altissimo numero di parametri liberi che devono essere stimati statisticamente (ossia appresi) dai dati disponibili. Ad esempio, uno dei modelli di *Deep Learning* più popolari, Inception V.3, pubblicato nel 2015 per affrontare la sfida di riconoscimento visivo su larga scala nota come ImageNet, ha 23 milioni di parametri. BERT, attualmente la migliore architettura per l’analisi del linguaggio naturale, ha più di 110 milioni di parametri²¹. Questa caratteristica delle reti di *Deep Learning* ha tre importanti implicazioni. Innanzitutto, queste architetture sono molto flessibili e possono memorizzare un numero estremamente elevato di relazioni tra le variabili. In secondo luogo, l’apprendimento di questi tipi di modelli richiede una quantità molto elevata di dati e grandi risorse

¹⁶ Hinton G. Deep Learning-A Technology With the Potential to Transform Health Care. JAMA. 2018 Sep 18;320(11):1101-1102. doi: 10.1001/jama.2018.11100. PMID: 30178065

¹⁷ Krizhevsky A, Sutskever I, Hinton G.E., ImageNet classification with Deep convolutional neural networks, Nips'12, Curran Associates Inc., 3 dicembre 2012, pp. 1097–1105

¹⁸ Si Y, Du J, Li Z, Jiang X, Miller T, Wang F, Jim Zheng W, Roberts K. Deep representation Learning of patient data from Electronic Health Records (EHR): A systematic review. J Biomed Inform. 2021 Mar;115:103671. doi: 10.1016/j.jbi.2020.103671. Epub 2020 Dec 31. PMID: 33387683

¹⁹ Devlin J, Chang M.W., Lee K, Toutanova K, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, 2019, arXiv, 1810.04805

²⁰ Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. 2017. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 6000–6010.

²¹ Vedi n.19

computazionali. In terzo luogo, alcuni dei compiti eseguiti da questi modelli di apprendimento possono essere considerati compiti cognitivi di base, come il riconoscimento delle forme nelle immagini. Per questo motivo, non è sempre necessario riaddestrare i modelli da zero in presenza di nuovi dati, ma è possibile sfruttare modelli pre-allenati, come meccanismi di trasformazione delle caratteristiche in ingresso. Le nuove funzionalità possono quindi essere utilizzate come componenti di base in una *pipeline* di *Machine Learning* più complessa, che preveda una fase di apprendimento dei parametri solo per la parte del modello che si deve occupare del ricavare la regola di classificazione a partire dalle variabili latenti. Questo è l'obiettivo del *transfer Learning*, una tecnica molto promettente per sfruttare convenientemente i modelli di *Deep Learning*, senza dover affrontare la difficoltà di apprendere molti parametri avendo a disposizione un numero insufficiente di esempi^{22,23,24}.

L'entusiasmo generato dall'elevata capacità dei sistemi di AI di riconoscere *pattern* complessi nei dati, in particolare nel campo della diagnostica per immagini, ha creato una forte aspettativa riguardo la possibilità concreta di realizzare applicazioni in grado di apprendere da qualsiasi tipo di esperienza clinica. Tuttavia, è importante ricordare come i successi dell'AI, con particolare riferimento al *Machine* e *Deep Learning*, sono legati alla capacità di svolgere compiti molto specifici, come la classificazione di immagini o testi, a valle di una procedura di apprendimento su grandissime quantità di dati. Quando queste capacità devono essere collocate in un sistema più ampio, come la realizzazione di un sistema di supporto alle decisioni cliniche che debba integrare più sorgenti informative e soprattutto continuamente adattarsi al contesto operativo di riferimento, il problema diventa molto più complesso e il numero di casi di successo riportati in letteratura scende vertiginosamente. I sistemi di AI dovranno quindi essere necessariamente progettati in modo da modellizzare esplicitamente il processo di cura del paziente, le sorgenti informative e le loro ontologie di riferimento, essere in grado di impiegare dati di elevata qualità per l'adattamento delle regole decisionali, ed essere dotati di strumenti di autodiagnosi e monitoraggio delle prestazioni²⁵.

2.4 Passi per la realizzazione di sistemi AI basati su Machine Learning (ML)

La realizzazione di sistemi AI basati su metodi di *Machine Learning* e *Deep Learning* presenta numerose criticità, in larga parte legate alla grande capacità di questi modelli di riprodurre i dati che

²²McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, Back T, Chesus M, Corrado GS, Darzi A, Etemadi M, Garcia-Vicente F, Gilbert FJ, Halling-Brown M, Hassabis D, Jansen S, Karthikesalingam A, Kelly CJ, King D, Ledsam JR, Melnick D, Mostofi H, Peng L, Reicher JJ, Romera-Paredes B, Sidebottom R, Suleyman M, Tse D, Young KC, De Fauw J, Shetty S. International evaluation of an AI system for breast cancer screening. *Nature*. 2020 Jan;577(7788):89-94. doi: 10.1038/s41586-019-1799-6. Epub 2020 Jan 1. PMID: 31894144

²³Lee J, Yoon W, Kim S, Kim D, Kim S, So CH, Kang J. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*. 2020 Feb 15;36(4):1234-1240. doi: 10.1093/bioinformatics/btz682. PMID: 31501885; PMCID: PMC7703786

²⁴Song B, Li Z, Lin X, Wang J, Wang T, Fu X. Pretraining model for biological sequence data. *Brief Funct Genomics*. 2021 Jun 9;20(3):181-195. doi: 10.1093/bfgp/elab025. PMID: 34050350; PMCID: PMC8194843

²⁵Stead WW. Clinical Implications and Challenges of Artificial Intelligence and Deep Learning. *JAMA*. 2018;320(11):1107-1108. doi:10.1001/jama.2018.11029

vengono utilizzati nella fase di apprendimento, dovuta all'elevato numero di parametri liberi. In altre parole, queste architetture possono “copiare” internamente tutti i dati disponibili, senza alcuna capacità di astrarre relazioni generali dai dati stessi. Ciò comporta almeno due conseguenze importanti: da una parte, durante la realizzazione di questi sistemi è fondamentale utilizzare delle strategie che siano in grado di assicurare buone prestazioni su casi non visti nella fase di apprendimento (ossia assicurarsi che l'approccio addestrato abbia una buona capacità di *generalizzazione*); dall'altra è inevitabile che il sistema che viene appreso dai dati sia indissolubilmente legato alle procedure di raccolta e alle eventuali distorsioni presenti nei dati stessi.

Le criticità che si sono verificate nell'utilizzo delle tecniche di AI per la diagnosi e la prognosi del COVID-19 sono un monito a procedere in modo molto rigoroso prima di adottare questi sistemi nella pratica clinica²⁶. Per affrontare queste criticità, la comunità internazionale ha definito delle procedure standard per il *reporting* degli studi che coinvolgono l'AI in medicina. Fra questi, lo standard MINIMAR (*MINimum Information for Medical AI Reporting*)²⁷ richiede che vengano sempre rispettati i seguenti principi: (1) includere informazioni sulla popolazione che fornisce i dati di addestramento, in termini di fonti di dati e principi di selezione della coorte; (2) includere i dati demografici della popolazione da cui è stato appreso il modello, in un modo che consenta un confronto con il proprio gruppo di pazienti o la nuova popolazione a cui si vuole applicare il modello stesso; (3) fornire dettagli sull'architettura computazionale e sullo sviluppo del modello in modo che sia chiara la finalità del modello di AI, sia possibile confrontarlo con modelli simili e sia possibile la replica dello studio; e (4) venga riportata in modo trasparente la procedura di valutazione statistica del modello e le procedure impiegate per la stima dei parametri liberi, sempre al fine di poter replicare i risultati. La valutazione statistica del modello deve a sua volta seguire delle procedure ormai standardizzate nella letteratura scientifica²⁸. Il principio fondamentale è la divisione dei dati a disposizione per la costruzione del modello in un sottoinsieme dedicato all'apprendimento del modello stesso (*training set*) e in un sottoinsieme su cui valutarne le prestazioni (*test set*). I due *data set* devono essere sempre mantenuti separati per poter ricavare delle valutazioni affidabili delle prestazioni del modello di AI²⁹. La letteratura riporta sia varie strategie che possono essere impiegate allo scopo sia diverse misure di performance che possono essere impiegate per valutare i classificatori. Infine, nella valutazione delle pubblicazioni scientifiche che riportano i risultati di algoritmi di AI è possibile impiegare delle *check-*

²⁶ Roberts, M., Driggs, D., Thorpe, M. et al. Common pitfalls and recommendations for using Machine Learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat Mach Intell* 3, 199–217 (2021). <https://doi.org/10.1038/s42256-021-00307-0>

²⁷ Hernandez-Boussard T, Bozkurt S, Ioannidis JPA, Shah NH. MINIMAR (MINimum Information for Medical AI Reporting): Developing reporting standards for artificial intelligence in health care. *J Am Med Inform Assoc.* 2020 Dec 9;27(12):2011-2015. doi: 10.1093/jamia/ocaa088. PMID: 32594179; PMCID: PMC7727333

²⁸ Witten I, Frank E, Hall M, Pal, C. *Data Mining, Practical Machine Learning Tools and Techniques*, 4th Edition, Morgan Kaufmann, 2016

²⁹ Cabitza F, Campagner A. The need to separate the wheat from the chaff in medical informatics. *Int J Med Inform.* 2021 Jun 2;104510. doi: 10.1016/j.ijmedinf.2021.104510. Epub ahead of print. PMID: 34108105

list che verifichino se siano stati rispettati dei criteri minimi qualitativi in tutte le fasi principali della costruzione di un modello di AI, ossia: i) comprensione del problema; ii) comprensione e analisi dei dati disponibili; iii) costruzione del modello; iv) validazione del modello; v) utilizzo nella pratica clinica^{30,31}. Nel caso vengano utilizzati degli algoritmi di AI in trial clinici, è fortemente raccomandato che siano sempre completate le *check-list* degli *standard* CONSORT-AI (*Consolidated Standards of Reporting Trials–Artificial Intelligence*)³² e SPIRIT-AI (*Standard Protocol Items: Recommendations for Interventional Trials–Artificial Intelligence*)³³, che permettono di aumentare la qualità dell’evidenza prodotta negli studi.

Come detto, i modelli di AI più recenti, tra cui quelli basati su metodi di *Deep Learning*, sono caratterizzati da architetture computazionali molto complesse. In questo caso il comportamento di un sistema di AI viene visto come una scatola nera, o *black-box*: il sistema è in grado di compiere azioni “intelligenti”, come riconoscere immagini o tradurre testi in molte lingue in maniera accurata, ma l’utente non è in grado di comprendere il funzionamento del sistema stesso. L’esigenza di avere un certo livello di spiegazione dei risultati, anche in relazione al contesto critico quale quello biomedico e al diritto alla spiegazione degli utenti, come previsto dalla *General Data Protection Regulation*, ha fortemente spinto la ricerca a definire strategie di *eXplainable AI* (XAI)³⁴. La spiegazione dei risultati di un modello di AI si può riferire al risultato ottenuto su un singolo paziente (*local explainability*) oppure al comportamento generale dell’algoritmo (*global explainability*). La ricerca sta attualmente ponendo grande attenzione a queste tematiche e i modelli di AI dovrebbero, quindi, essere sempre dotati di meccanismi che consentano un certo livello di spiegazione dei risultati.

La dipendenza dei modelli di *Machine Learning* dai dati di apprendimento richiede infine una particolare attenzione nel loro impiego durante la pratica clinica, in particolare in centri che potrebbero avere pazienti con caratteristiche diverse da quelle dei dati di partenza. Per questo motivo un ulteriore requisito che dovrebbe essere incluso nei sistemi di AI basati su *Machine Learning* è l’analisi dell’affidabilità dei risultati, composta da due aspetti: una valutazione della confidenza della

³⁰ Bellazzi R, Zupan B. Predictive data mining in clinical medicine: current issues and guidelines. *Int J Med Inform.* 2008 Feb;77(2):81-97. doi: 10.1016/j.ijmedinf.2006.11.006. Epub 2006 Dec 26. PMID: 17188928.

³¹ Cabitza F, Campagner A, Soares F, García de Gadiana-Romualdo L, Challa F, Sulejmani A, Seghezzi M, Carobene A. The importance of being external. methodological insights for the external validation of Machine Learning models in medicine. *Comput Methods Programs Biomed.* 2021 Sep; 208:106288. doi: 10.1016/j.cmpb.2021.106288. Epub 2021 Jul 22. PMID: 34352688.

³² Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Lancet Digit Health.* 2020 Oct;2(10):e537-e548. doi: 10.1016/S2589-7500(20)30218-1. Epub 2020 Sep 9. PMID: 33328048; PMCID: PMC8183333

³³ Cruz Rivera S, Liu X, Chan AW, Denniston AK, Calvert MJ; SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Digit Health.* 2020 Oct;2(10):e549-e560. doi: 10.1016/S2589-7500(20)30219-3. Epub 2020 Sep 9. PMID: 33328049; PMCID: PMC8212701

³⁴ Kundu S. AI in medicine must be explainable. *Nat Med.* 2021 Aug; 27(8):1328. doi: 10.1038/s41591-021-01461-z. PMID: 34326551

classificazione effettuata³⁵ e la stima di quanto i dati del paziente di interesse siano simili o diversi da quelli impiegati nella fase di apprendimento³⁶.

Il successo delle tecnologie di AI nel compiere *task* diagnostici sta promuovendo la rapida commercializzazione di nuove soluzioni software in campo biomedico (<https://www.fda.gov/media/145022/download>). Anche grazie alla recente spinta data dalla bozza di regolamento sull'Intelligenza Artificiale, che classifica tutti i sistemi biomedici come ad alto rischio, è evidente che prodotti di AI in Italia e in Europa dovranno seguire il percorso di marcatura CE, così come previsto dalla regolamentazione per i dispositivi medici entrata recentemente in vigore (<https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>). I sistemi di supporto alle decisioni, peraltro, rientrano già nella prassi attuale nei dispositivi medici di classe II (a/b) (https://ec.europa.eu/health/sites/default/files/md_topics-interest/docs/md_mdcg_2019_11_guidance_en.pdf). La combinazione delle procedure di marcatura e l'auspicabile aumento di attenzione sulla qualità degli studi pubblicati e sull'aderenza alle linee guida internazionali di validazione dei sistemi di AI dovrebbero portare alla realizzazione di strumenti affidabili, e far emergere le eventuali criticità precedentemente riportate.

2.5 Metodologie e tecnologie abilitanti

Se le metodologie di AI basate su *Machine Learning* trovano il loro fondamento nella disponibilità di dati, e in particolare di grandi quantità di dati, è evidente che il più importante insieme di tecnologie che favoriscono l'implementazione dell'AI è rappresentato dalla realizzazione di sistemi di raccolta, integrazione e condivisione delle informazioni, tali da includere dati strutturati, segnali, immagini e testi. A questo scopo è necessario definire e applicare strategie che permettano l'interoperabilità sintattica e semantica fra i sistemi, sia all'interno delle realtà ospedaliere e cliniche, sia fra tutti gli attori coinvolti nella sanità. L'utilizzo di standard per la rappresentazione dei dati, come i sistemi di codifica ICD, LOINC e le terminologie ed ontologie come SNOMED e UMLS, rappresenta un passo ineludibile per la creazione di coorti sufficientemente grandi e condivisibili di dati tali da poter costituire la base per la realizzazione di nuovi modelli di AI e per impiegare utilmente strategie di *transfer Learning*³⁷.

Tuttavia, come già descritto precedentemente, la realizzazione di una sanità dotata di sistemi di supporto alle decisioni “che apprendono” dinamicamente dai dati raccolti durante la pratica clinica necessita di una profonda trasformazione dei meccanismi di raccolta e messa a disposizione dei dati,

³⁵ Hüllermeier E, Waegeman W. Aleatoric and Epistemic Uncertainty in Machine Learning: An Introduction to Concepts and Methods. *Machine Learning* 110 (3): 457–506, 2021. <https://doi.org/10.1007/s10994-021-05946-3>

³⁶ Davis SE, Greevy RA Jr, Lasko TA, Walsh CG, Matheny ME. Detection of calibration drift in clinical prediction models to inform model updating. *J Biomed Inform.* 2020 Dec;112:103611. doi: 10.1016/j.jbi.2020.103611. Epub 2020 Nov 4. PMID: 33157313

³⁷ Coiera E. *Guide to Health Informatics*. CRC Press. 2015

che preveda la realizzazione di una infrastruttura organizzativa di *data stewardship* e *governance*³⁸. La *data stewardship* si occupa della gestione dei metadati, o dati sui dati raccolti, e ne verifica la qualità e la coerenza, “così che chi ne fruisce sia certo di utilizzare dati conformi al loro significato apparente, completi e affidabili”³⁹. La *data governance* affronta, invece, tutte le problematiche legate all’accesso ai dati, ai profili autorizzativi e al rispetto della normativa sulla *privacy*. Un’organizzazione adeguata deve anche richiedere un aumento di consapevolezza nella raccolta delle informazioni, che deve essere attenta a partire dal singolo operatore, ma resa snella e, quando possibile, automatizzata. Un aspetto spesso trascurato riguarda la necessità di rappresentare esplicitamente i processi di cura impiegati, in modo da poter riconoscere le caratteristiche delle diverse sorgenti informative, tenendo conto anche del loro specifico contesto⁴⁰.

A livello nazionale e internazionale sono attualmente attivi diversi progetti di ricerca che hanno lo scopo di creare delle infrastrutture multicentriche in grado di condividere dati per favorire lo sviluppo di strumenti di AI (es.: <https://www.alleanzacontroilcancro.it/progetti/health-big-data/>). Molte di queste progettualità sono nate prendendo a modello iniziative per la condivisione di specifiche tipologie di dati, come i dati molecolari, i dati di *imaging* o i dati clinici di patologie di interesse^{41,42}. La notevole varietà di informazioni raccolte rende difficile ottenere una visione organica, in particolare in assenza di quello strato organizzativo di base di cui si parlava nel paragrafo precedente. È comunque interessante notare la presenza di piattaforme e di standard per la raccolta e per la condivisione dei dati, che sono largamente diffuse a livello internazionale e che potrebbero rappresentare una base per la realizzazione di infrastrutture condivise, come HL7 FHIR (<https://www.hl7.org/fhir/>), OMOP/ODHSI (<https://ohdsi.org/>), XNAT (<https://www.xnat.org/>) e i2b2/transmart (<https://i2b2transmart.org>).

Dal punto di vista delle tecnologie abilitanti ICT, le strategie a supporto della condivisione dei dati sembrano convergere in questo periodo verso la realizzazione di *data lakes*, ovvero dei *repository* di dati archiviati nel loro formato originale, includendo sia dati raccolti in database relazionali, sia in formati semistrutturati come file XML o JSON, sia, infine, dati non strutturati come testi, immagini e segnali. Le diverse sorgenti di dati contenute nei *data lakes* devono essere opportunamente indicizzate e i dati in esse contenuti dovranno essere estratti e attentamente preparati per poter essere analizzati con

³⁸ Geissbuhler A, Safran C, Buchan I, Bellazzi R, Labkoff S, Eilenberg K, Leese A, Richardson C, Mantas J, Murray P, De Moor G. Trustworthy reuse of health data: a transnational perspective. *Int J Med Inform.* 2013 Jan;82(1):1-9. doi: 10.1016/j.ijmedinf.2012.11.003. Epub 2012 Nov 20. PMID: 23182430

³⁹ Giacchino A. <https://www.itware.com/blog-itware/big-data-analytics-data-management/item/1311-data-steward-per-mestiere-piu-semplce-che-data-analyst-piu-di-valore-per-l-impresa>, 2017

⁴⁰ Hripcsak G, Albers DJ, Perotte A. Parameterizing time in electronic health record studies. *J Am Med Inform Assoc.* 2015 Jul;22(4):794-804. doi: 10.1093/jamia/ocu051. Epub 2015 Feb 26. PMID: 25725004; PMCID: PMC6169471

⁴¹ Harrison PW, Lopez R, Rahman N, Allen SG, Aslam R, Buso N, Cummins C, Fathy Y, Felix E, Glont M, Jayathilaka S, Kadam S, Kumar M, Lauer KB, Malhotra G, Mosaku A, Edbali O, Park YM, Parton A, Pearce M, Estrada Pena JF, Rossetto J, Russell C, Selvakumar S, Sitjà XP, Sokolov A, Thorne R, Ventouratou M, Walter P, Yordanova G, Zadissa A, Cochrane G, Blomberg N, Apweiler R. The COVID-19 Data Portal: accelerating SARS-CoV-2 and COVID-19 research through rapid open access data sharing. *Nucleic Acids Res.* 2021 Jul 2;49(W1):W619-W623. doi: 10.1093/nar/gkab417. PMID: 34048576; PMCID: PMC8218199

⁴² Lenert LA, Ilatovskiy AV, Agnew J, Rudisill P, Jacobs J, Weatherston D, Deans KR Jr. Automated production of research data marts from a canonical fast healthcare interoperability resource data repository: applications to COVID-19 research. *J Am Med Inform Assoc.* 2021 Jul 30;28(8):1605-1611. doi: 10.1093/jamia/ocab108. PMID: 33993254; PMCID: PMC8243354

algoritmi di AI. Le esigenze di interoperabilità semantica e sintattica descritte precedentemente sono, quindi, fondamentali per utilizzare in modo efficiente queste soluzioni per la raccolta dati.

Allo stato dell'arte esistono tre possibili configurazioni che consentano di effettuare analisi su grandi quantità di dati, potenzialmente condivisi fra più centri clinici e di ricerca.

In primo luogo, è possibile creare dei *data warehouse* o *data mart* centralizzati, in cui le informazioni siano fisicamente trasferite e che possano utilizzare delle risorse computazionali dedicate. Ciò può essere fatto sia mediante un sistema di calcolo ad alte prestazioni opportunamente realizzato, sia grazie all'utilizzo di servizi *cloud*, che consentono una configurazione “*on demand*” della soluzione computazionale, senza chiedere preventivamente un investimento infrastrutturale. La centralizzazione dei dati ha, tuttavia, diversi inconvenienti, primo fra tutti la necessità di tener conto dei vincoli legati alla *privacy* e alla protezione delle informazioni sensibili.

Una seconda strategia consiste nella realizzazione di analisi distribuite e nella successiva integrazione dei risultati con strategie di meta-analisi. Questo tipo di approccio ha il vantaggio di aumentare le tutele nella gestione della *privacy* e di ridurre le esigenze di sincronizzazione della banca dati centralizzata con la periferia. L'approccio, però, non consente di sfruttare pienamente la mole dei dati disponibili, che potrebbero, viceversa, consentire la realizzazione di modelli di AI diagnostici più performanti grazie all'accesso alle informazioni individuali.

Una terza strategia, particolarmente promettente, riguarda l'utilizzo di metodi di AI che eseguono l'apprendimento su architetture distribuite. In questo caso, l'algoritmo di AI è in grado di “muoversi” da un centro partecipante a un altro, accedendo localmente ai dati e costruendo poi delle regole di classificazione con lo stesso livello di dettaglio della soluzione centralizzata. Nell'ambito di questa strategia alcuni consorzi internazionali stanno sperimentando delle soluzioni con elevato livello di sicurezza; in questo caso gli algoritmi svolgono le analisi in modo distribuito, avendo accesso soltanto a dati crittografati⁴³. Nell'ambito delle strategie distribuite, bisogna, inoltre, segnalare la tecnologia dell'*edge computing*, che consente di svolgere calcoli distribuiti utilizzando in parte le infrastrutture di rete⁴⁴. Infine, è necessario ricordare che ognuna delle soluzioni sopra descritte può essere realizzata con una notevole varietà e disponibilità di architetture di calcolo, che spaziano dai

⁴³ Scheibner J, Raisaro JL, Troncoso-Pastoriza JR, Ienca M, Fellay J, Vayena E, Hubaux JP. Revolutionizing Medical Data Sharing Using Advanced Privacy-Enhancing Technologies: Technical, Legal, and Ethical Synthesis. *J Med Internet Res*. 2021 Feb 25;23(2):e25120. doi: 10.2196/25120. PMID: 33629963; PMCID: PMC7952236

⁴⁴ Warnat-Herresthal S, Schultze H, Shastry KL, Manamohan S, Mukherjee S, Garg V, Sarveswara R, Händler K, Pickkers P, Aziz NA, Ktena S, Tran F, Bitzer M, Ossowski S, Casadei N, Herr C, Petersheim D, Behrends U, Kern F, Fehlmann T, Schommers P, Lehmann C, Augustin M, Rybniker J, Altmüller J, Mishra N, Bernardes JP, Krämer B, Bonaguro L, Schulte-Schrepping J, De Domenico E, Siever C, Kraut M, Desai M, Monnet B, Saridaki M, Siegel CM, Drews A, Nuesch-Germano M, Theis H, Heyckendorf J, Schreiber S, Kim-Hellmuth S; COVID-19 Aachen Study (COVAS), Nattermann J, Skowasch D, Kurth I, Keller A, Bals R, Nürnberg P, Rieß O, Rosenstiel P, Netea MG, Theis F, Mukherjee S, Backes M, Aschenbrenner AC, Ulas T; Deutsche COVID-19 Omics Initiative (DeCOI), Breteler MMB, Giamarellos-Bourboulis EJ, Kox M, Becker M, Cheran S, Woodacre MS, Goh EL, Schultze JL. Swarm Learning for decentralized and confidential clinical Machine Learning. *Nature*. 2021 Jun;594(7862):265-270. doi: 10.1038/s41586-021-03583-3. Epub 2021 May 26. PMID: 34040261; PMCID: PMC8189907

cluster di processori *multi-core* alle GPU, e che in futuro potranno impiegare soluzioni avveniristiche come il *quantum computing*⁴⁵.

3. STATO DELL'ARTE DEI SISTEMI DI AI IN DIAGNOSTICA

Il tema dell'AI e quello della diagnostica sono strettamente connessi in virtù del concetto chiave di “predizione”. Se da una parte una buona predizione consente di raggiungere una diagnosi accurata, dall'altra l'AI è in grado di implementare dei veri e propri “modelli predittivi” matematici capaci di identificare precocemente gli esordi di una malattia anticipandone l'insorgenza. Pertanto, si è ritenuto opportuno effettuare un'indagine preliminare (v. Tabella 1, in appendice) su *Medline* che prendesse in considerazione entrambi i concetti di diagnosi e di predizione.

La ricerca, che è stata limitata agli ultimi 5 anni e ha riguardato le revisioni sistematiche e le meta-analisi, ha prodotto oltre 500 risultati. Tali articoli sono il frutto di una ricerca più ampia che, estesa a tutte le tipologie di articoli presenti su *Medline*, avrebbe fornito oltre 150.000 articoli, con una media di 35.000 pubblicazioni sin dal 2013. A partire dalle principali (e più recenti) revisioni sistematiche abbiamo evidenziato quali sono le tendenze principali attorno ai temi scelti, analizzando *abstract* e *full text* per avere una panoramica olistica sul fenomeno preso in esame.

3.1 Principali applicazioni dell'AI in sanità a supporto della diagnostica

I contesti in cui vi è più numerosa letteratura sul tema sono: la radiologia, la radioterapia, l'oftalmologia, e la dermatologia⁴⁶, ma anche la gastroenterologia, la ginecologia oncologica e la senologia⁴⁷, l'ematologia⁴⁸ e l'infettivologia⁴⁹ (fra cui, non da ultimo, anche il COVID-19⁵⁰). Per esempio, in uno studio in ambito radiologico, l'accuratezza di un algoritmo sviluppato con l'obiettivo di rilevare la polmonite partendo da oltre 112.000 immagini radiografiche del torace è stata confrontata con quella di quattro radiologi, risultando superiore⁵¹. Evidenze esistono, anche, nel campo della identificazione di tumori polmonari, dove algoritmi di *machine learning*, istruiti attraverso la scansione di oltre 34.000 radiografie toraciche, hanno raggiunto un livello di accuratezza superiore a 17 su 18

⁴⁵ Huang HY, Broughton M, Mohseni M, Babbush R, Boixo S, Neven H, McClean JR. Power of data in quantum Machine Learning. *Nat Commun.* 2021 May 11;12(1):2631. doi: 10.1038/s41467-021-22539-9. PMID: 33976136; PMCID: PMC8113501

⁴⁶ Naylor C.D. (2018), “On the prospects for a (deep) learning health care system”, *JAMA*, 320, 1099–1100

⁴⁷ Sangaiah A.K., et al. (2020), “An intelligent learning approach for improving ECG signal classification and arrhythmia analysis”, *Artif Intell Med.* 103:101788

⁴⁸ Radakovich N, et al. (2020), “Artificial Intelligence in Hematology: Current Challenges and Opportunities”, *Curr Hematol Malig Rep.*, 15(3):203-210

⁴⁹ Peiffer-Smadja N, et al. (2020), “Machine learning for clinical decision support in infectious diseases: a narrative review of current applications”, *Clin Microbiol Infect.*, 26(5):584-595

⁵⁰ Ozsahin I, et al. (2020), “Review on Diagnosis of COVID-19 from Chest CT Images Using Artificial Intelligence”, *Comput Math Methods Med.*, 2020:9756518

⁵¹ Wang X., et al. (2017), “ChestX-ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases”, Preprint at <https://arxiv.org/abs/1705.02315>

radiologi usati come confronto⁵². Risultati simili si sono ottenuti nell'identificazione di tumori della mammella dove un sistema di AI adeguatamente istruito ha portato a una riduzione assoluta del 5,7% e 1,2% (rispettivamente negli Stati Uniti e nel Regno Unito) nei falsi positivi e del 9,4% e 2,7% nei falsi negativi, e, nel confronto con l'operato di 6 radiologi, a un aumento del 11,5% della sensibilità⁵³. In campo radiologico, passi importanti sono stati fatti nello sviluppo di sistemi di AI in grado di diagnosticare con maggiore precisione le fratture del polso, aumentando la sensibilità dall'81% registrata nella diagnosi da parte del personale di pronto soccorso al 92% ottenuta da un sistema di AI, che ha consentito, anche, di ridurre le interpretazioni errate del 47%⁵⁴.

In area dermatologica, sono diversi gli studi condotti sul campo per dimostrare l'accuratezza di sistemi di AI finalizzati all'identificazione di neoplasie. Uno studio, che utilizzava un ampio set di dati di addestramento di quasi 130.000 immagini, ha dimostrato l'affidabilità di un sistema di *machine learning* nell'identificazione di carcinoma e di melanoma con una sensibilità rispettivamente del 96% e del 94%, assolutamente sovrapponibili a quelle di 21 dermatologi americani certificati⁵⁵. Tuttavia, è l'area oftalmologica quella nella quale più numerose sono le prove di affidabilità dei sistemi di AI. Il primo sistema di AI approvato dalla *Food and Drug Administration* come *medical device* è proprio un algoritmo di *machine learning* (adeguatamente istruito con fotografie del fondo retinico e OCT) che in uno studio prospettico condotto su 900 pazienti diabetici è stato in grado di identificare retinopatie con una sensibilità dell'87% e una specificità del 91%⁵⁶.

Non mancano strumenti di AI basati su tecniche di *Natural Language Processing* e addestrati sui dati contenuti nei sistemi di cartelle cliniche elettroniche contenenti la storia clinica dei pazienti composta da dati sia strutturati sia non strutturati. Un sistema sviluppato in ambito pediatrico, addestrato con 100 milioni di dati riferiti a circa 1,4 milioni di visite condotte da oltre 500.000 pazienti, ha permesso di ottenere elevate percentuali di successo nel diagnosticare le malattie pediatriche più comuni (95% per sinusiti e altre infezioni respiratorie, 94% per l'influenza, 97% per l'infezione mani-piedi-bocca) e quelle più pericolose o complicate (97% per gli attacchi di asma, 93% per la meningite batterica e per la varicella, 93% per la rosolia, 90% per la mononucleosi), con percentuali di sensibilità e specificità simili a quelle osservate tra i medici più esperti⁵⁷.

Esistono anche sistemi di AI sviluppati per supportare la diagnosi di COVID-19. Uno studio ha valutato le prestazioni di un sistema di AI nel rilevare pazienti con COVID-19 analizzando le

⁵² Nam J.G., et al. (2018), "Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs", *Radiology* <https://doi.org/10.1148/radiol.2018180237>

⁵³ McKinney SM, et al. (2020), "International evaluation of an AI system for breast cancer screening", *Nature*, 577(7788):89-94. doi: 10.1038/s41586-019-1799-6

⁵⁴ Lindsey R., et al. (2018), "Deep neural network improves fracture detection by clinicians", *Proc. Natl. Acad. Sci. USA* 115, 11591-6

⁵⁵ Esteva A., et al. (2017), "Dermatologist-level classification of skin cancer with deep neural networks", *Nature* 542, 115-8

⁵⁶ Abramoff M., et al. (2018), "Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices", *N PJ Digit. Med.* 1, 39

⁵⁷ Liang H, et al. (2019), "Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence", *Nat Med.*, 25(3):433-8. doi: 10.1038/s41591-018-0335-9

radiografie del torace e dimostrando la sua affidabilità rispetto all'operato di sei radiologi con una sensibilità dell'85% e una specificità del 61%⁵⁸; mentre un altro ha dimostrato che un algoritmo di *deep learning* è in grado di riconoscere il COVID-19 rispetto ad altre malattie polmonari analizzando le TAC del torace dei pazienti⁵⁹. Anche se l'*American College of Radiology* e i *Centers for Disease & Control* sconsigliano di utilizzare la TAC e le radiografie del torace per lo *screening* o come test di prima linea per diagnosticare il COVID-19⁶⁰, gli studi suggeriscono che l'approccio basato sul *deep learning* può essere promettente.

Altri strumenti di AI sono stati sviluppati e studiati in ambito cardiologico per diagnosticare infarti e aritmie sulla base dell'analisi di elettrocardiogrammi e per classificare adeguatamente i pazienti sulla base dei loro ecocardiogrammi⁶¹. Dal punto di vista tecnologico e modellistico, sono diverse gli approcci usati per implementare sistemi di AI. La tabella 2 (v. Appendice) illustra i principali desunti dalla analisi della letteratura scientifica.

3.2 Considerazioni metodologiche e cliniche

Sebbene siano molti gli studi che sembrano fornire prove di affidabilità a favore dei sistemi di AI usati in un contesto diagnostico, esistono alcune revisioni sistematiche e meta-analisi che ne mettono in discussione la validità scientifica e la metodologia impiegata per raggiungere tali risultati^{62,63}.

Pochi, infatti, sono gli studi clinici che confrontano la diagnosi raggiunta da un sistema di AI o di *machine learning* rispetto a quella raggiunta da un operatore sanitario o un esperto della materia, e la maggior parte di questi ha un elevato rischio di *bias* perché non correttamente condotti dal punto di vista metodologico. A ciò occorre aggiungere il fatto che molti studi clinici sono retrospettivi e basati su set di dati precedentemente assemblati, mentre pochi sono quelli prospettici condotti in ambienti clinici reali, in un contesto di "*real world*" e basati sul modello delle sperimentazioni cliniche controllate randomizzate. A complicare le cose e a mettere ulteriormente in discussione i risultati raggiunti dagli studi clinici è il fatto che molti di questi adottano una validazione interna del sistema di AI a scapito di una validazione esterna, più accurata nel raggiungere gli obiettivi, mentre il confronto

⁵⁸ Murphy K., et al. (2020), "COVID-19 on Chest Radiographs: A Multireader Evaluation of an Artificial Intelligence System", *Radiology*, 296(3):E166-E172. doi: 10.1148/radiol.2020201874

⁵⁹ Li L., et al. (2020), "Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy", *Radiology*, 296(2):E65-71. doi: 10.1148/radiol.2020200905

⁶⁰ American College of Radiology (ACR) (2020), "Recommendations for the use of Chest Radiography and Computed Tomography (CT) for Suspected COVID-19 Infection", 11 Marzo 2020. Fonte: <https://www.acr.org/Advocacy-and-Economics/ACR-Position-Statements/Recommendations-for-Chest-Radiography-and-CT-for-Suspected-COVID19-Infection>

⁶¹ Topol E.J., (2019), "High-performance medicine: the convergence of human and artificial intelligence", *Nat Med.*, 25(1):44-56. doi: 10.1038/s41591-018-0300-7

⁶² Liu X., et al. (2019), "A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis", *Lancet Digit Health*. 1(6):e271-e297. doi: 10.1016/S2589-7500(19)30123-2. Erratum in: *Lancet Digit Health*. 2019 Nov;1(7):e334

⁶³ Nagendran M., et al. (2020), "Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies", *BMJ*, 368:m689. doi: 10.1136/bmj.m689

tra le performance dei modelli di *machine learning* e quelle degli esperti (laddove esiste) solo in pochi casi è condotto utilizzando lo stesso set di dati.

Altri problemi metodologici risiedono nell'ampio *range* di metriche usate per misurare le *performance* diagnostiche dei sistemi di AI che ne rende difficile il confronto e l'interpretazione, il ridotto numero di operatori sanitari usati per il confronto delle prestazioni diagnostiche (quando ciò avviene è stimato in un numero medio di 5, di cui 4 esperti¹¹⁰), la limitata generalizzazione dei risultati, la ridotta disponibilità di data-set e codici che rendono difficile la riproducibilità degli studi pubblicati. Per esempio, una recente revisione sistematica accompagnata da una meta-analisi⁶⁴, partendo da circa 32.000 studi ne ha identificati 82 che confrontavano le *performance* diagnostiche di sistemi di *deep learning* basate sull'analisi delle immagini rispetto a quelle fornite da esperti del settore (tra gli studi più rappresentativi, 18 erano quelli condotti in ambito oftalmologico, 17 in ambito radiologico e 9 in ambito dermatologico). I ricercatori hanno osservato che una validazione esterna era stata, però, effettuata solo in 25 studi, di cui solo 14 avevano effettuato il confronto tra modelli di *deep learning* e operatori sanitari sullo stesso *data-set*, rilevando una sensibilità complessiva dell'87,0% (IC 95% 83,0–90,2) per i modelli di *deep learning* e dell'86,4% (79,9–91,0) per gli operatori sanitari e una specificità del 92,5% (IC 95% 85,1–96,4) per i modelli di *deep learning* e del 90,5% (80,6–95,7) per gli operatori sanitari. Gli autori suggeriscono che la presunta parità in termini di affidabilità diagnostica tra i sistemi di *deep learning* e gli esperti debba essere vista alla luce dei limiti metodologici e scientifici identificati nel corso della revisione e precedentemente indicati. Risultati simili sono stati riportati da un'altra revisione sistematica, condotta sui registri di sperimentazioni cliniche, che ha analizzato gli studi in corso e quelli completati che mettevano a confronto le prestazioni diagnostiche di sistemi di *deep learning* con l'operato di esperti⁶⁵.

Non bisogna, inoltre, dimenticare altre limitazioni che pregiudicano in parte, a oggi, l'impiego di sistemi di AI in medicina e, in particolare, lo sviluppo di strumenti diagnostici: i *bias* di selezione nella scelta dei campioni su cui tali sistemi vengono istruiti e che rendono l'algoritmo non applicabile a tutta la popolazione, la mancanza di trasparenza dell'algoritmo (nota con il termine di *black box*) che impedisce al medico che impiega il sistema di conoscere il ragionamento applicato per arrivare alla soluzione proposta, i problemi di sicurezza dei dati dei pazienti che possono mettere a rischio la loro *privacy*⁶⁶.

Infine, esiste il problema della scarsa collaborazione tra clinici e sviluppatori di sistemi di AI. Per esempio, una revisione sistematica⁶⁷ degli studi di *imaging* polmonare assistito dall'AI per COVID-19

⁶⁴ vedi nota 62

⁶⁵ Vedi nota 63

⁶⁶ Santoro E. (2017), "L'intelligenza artificiale in medicina: quali limiti, quali ostacoli, quali domande", *Recenti Prog Med.*, 108(12):500-502. doi: 10.1701/2829.28580

⁶⁷ Abbasi J. (2021), "Artificial Intelligence in COVID-19 Imaging Mismatched to the Clinic", *JAMA*. 13;326(2):124. doi: 10.1001/jama.2021.10888

ha rivelato una discrepanza tra ciò su cui si concentrano gli sviluppatori e ciò di cui i medici hanno bisogno. La revisione ha riguardato articoli scientifici apparsi su riviste mediche o su server *pre-print* che avevano come argomento l'*imaging* polmonare in pazienti COVID-19 per i quali era specificata la tipologia di immagine usata (tra TAC, radiografia e ecografia). I ricercatori hanno identificato 463 articoli che sperimentavano sistemi di AI e 2496 articoli clinici che non la utilizzavano. Confrontando i due gruppi di articoli si sono accorti che mentre l'84% degli studi clinici utilizzava la TAC, solo il 39% degli studi sull'AI usava questo tipo di immagini. La maggior parte della ricerca di *imaging* polmonare assistita dall'AI si è basata su radiografie, che invece costituivano solo il 10% degli studi clinici senza l'ausilio di AI. Inoltre, il 72% degli articoli sull'AI era focalizzato sulla diagnosi di COVID-19 anziché sulla gravità o sulla prognosi della malattia, anche se l'immagine non era consigliata o utilizzata di routine per questo obiettivo. Infine, solo 12 dei manoscritti sull'AI erano di buona qualità e vedevano la collaborazione tra gli esperti di tecnologia e gli esperti medici.

3.3 Considerazioni tecniche

Gli algoritmi di AI mostrano la loro più grande utilità nelle analisi multifattoriali di dati per consentire la creazione di modelli matematici predittivi sempre più precisi sulla base delle informazioni che si possono reperire dal paziente, da sottopopolazioni con caratteristiche affini, da popolazioni con la medesima patologia o da popolazioni apparentemente distinte. I dati possono essere clinici, genomici, metabolomici, di *imaging*, di laboratorio, demografici, nutrizionali, psicologici, biologici e comportamentali. Di seguito, si riportano i principali *task* assegnati agli algoritmi per fornire supporto alla diagnostica e le loro principali applicazioni in sanità. Considerando le limitazioni descritte nella metodologia, l'obiettivo di questa revisione non è onnicomprensivo, ma esplorativo.

Basando le predizioni e le modellizzazioni su grandi set di dati classificati sulla base di analogie, regole, connessioni, reti neurali, statistiche o probabilità, i *task* principali degli algoritmi a supporto della diagnostica sono i seguenti:

- Utilizzo di ontologie (modelli condivisi, concordati e dettagliati di un determinato problema o dominio) per creare dei modelli semantici di dati collegandosi con istanze di base^{68,69}
- Predizione del rischio e diagnosi di svariate patologie, specialmente oncologiche, nelle loro tipologie, *features* e livelli di complessità⁷⁰;

⁶⁸ Zouri M, et al. (2019), "An Ontology Approach for Knowledge Representation of ECG Data", *Stud Health Technol Inform.*, 257:520-525

⁶⁹ Questa modellizzazione è definita knowledge-base, ed è quella che viene utilizzata dal motore di ricerca PubMed: le ontologie sono le parole chiave [definite e consolidate nel MeSH Browser e associate ad ogni articolo], le istanze di base sono le *query*, il risultato è un modello semantico di dati, frutto di un data mining all'interno del database attraverso le parole chiave.

⁷⁰ Kourou K., et al. (2014), "Machine learning applications in cancer prognosis and prediction", *Comput. Struct. Biotechnol. J.*, 13, 8-17

- Classificazione delle tipologie di cancro attraverso la valutazione visiva delle cellule tumorali^{71,72};
- Integrazione di dati omici e multi-omici per la personalizzazione diagnostica di patologie complesse^{73,74};
- Identificazione di potenziali *cluster*, biomarcatori o fenotipi clinici quali predittori di rischio^{75,76};
- Identificazione di *pathway* biologici e clusterizzazione dei dati per l'analisi dei *pathway*^{77,78};
- Identificazione di elementi genomici e molecolari sensibili a trattamenti esistenti o innovativi per predire eventi avversi⁷⁹;
- Identificazione di nuove associazioni fra le patologie e le loro cause scatenanti, inclusi i casi di comorbidità e multimorbidità^{80,81};
- Identificazione di relazioni non-lineari in cartella elettronica (es. *text mining*)^{82,83};
- Identificazione di nuovi percorsi causali e causativi di patologia generando nuove ipotesi⁸⁴.
- Visualizzazione e analisi di strutture di conoscenza (estratte dall'*Internet-of-Things* - IoT) per identificare *pattern* espliciti e nascosti⁸⁵;
- Osservazione di *outcome* terapeutici patologia-specifici differenti, per identificare nuove istanze (es. mutazioni o alterazioni genetiche) che originano nuovi tumori⁸⁶.

⁷¹ Sun R, et al. (2018), "A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study", *Lancet Oncol.*, Sep;19(9):1180-1191

⁷² Koelzer V.H., et al. (2019), "Precision immunoprofiling by image analysis and artificial intelligence", *Virchows Arch.*, 474(4):511-522

⁷³ Patel S.K., et al. (2020), "Artificial Intelligence to Decode Cancer Mechanism: Beyond Patient Stratification for Precision Oncology", *Front Pharmacol.*, 11:1177

⁷⁴ Cesario, A., et al. (2021), "Personalized Clinical Phenotyping through Systems Medicine and Artificial Intelligence", *J. Pers. Med.*, 2021, 11, 265

⁷⁵ Hedman Å.K., et al. (2020), "Identification of novel pheno-groups in heart failure with preserved ejection fraction using machine learning", *Heart.*, 106(5):342-349

⁷⁶ Sun X, et al. (2020), "Identification of dynamic signatures associated with smoking-related squamous cell lung cancer and chronic obstructive pulmonary disease", *J Cell Mol Med.*, 24(2):1614-1625

⁷⁷ Zhu Z., et al. (2017), "Identification of pivotal genes and pathways for spinal cord injury via bioinformatics analysis", *Mol Med Rep.*, 16(4):3929-3937

⁷⁸ Zhang C., et al. (2017), "The identification of key genes and pathways in hepatocellular carcinoma by bioinformatics analysis of high-throughput data", *Med Oncol.* 34(6):101

⁷⁹ Fornecker LM, et al. (2019), "Multi-omics dataset to decipher the complexity of drug resistance in diffuse large B-cell lymphoma", *Sci Rep.*, Jan 29; 9(1):895

⁸⁰ Zamborlini V., et al. (2017), "Analyzing interactions on combining multiple clinical guidelines", *Artif Intell Med.*, 81:78-93

⁸¹ DeParis S., et al. (2018), "An Analytical Method for Multimorbidity Management Using Bayesian Networks", *Stud Health Technol Inform.*, 247:820-824. X

⁸² Brunekreef T.E., et al. (2021), "Text Mining of Electronic Health Records Can Accurately Identify and Characterize Patients With Systemic Lupus Erythematosus", *ACR Open Rheumatol.* 3(2):65-71

⁸³ van Dijk W.B., et al. (2021), "Text-mining in electronic healthcare records can be used as efficient tool for screening and data collection in cardiovascular trials: a multicenter validation study", *J Clin Epidemiol.*, 132:97-105

⁸⁴ Naylor C.D. (2018), "On the prospects for a (deep) learning health care system", *JAMA*, 320, 1099–1100

⁸⁵ Sangaiah A.K., et al. (2020), "An intelligent learning approach for improving ECG signal classification and arrhythmia analysis", *Artif Intell Med.* 103:101788

⁸⁶ Vedi n.126

3.4 Criticità e sfide future

La validazione scientifica dei sistemi di AI in medicina, e in particolare per sviluppare strumenti diagnostici, è condizione assolutamente indispensabile per il loro impiego in un contesto assistenziale. Studi clinici metodologicamente solidi (che siano prospettici, randomizzati, che prevedano un confronto con medici davvero esperti sugli stessi dati, che prevedano una validazione esterna del sistema o che siano basati sulla metodologia di ricerca corrente o su nuove modalità che, pur preservando il rigore scientifico, potrebbero essere più affini ai nuovi strumenti) sono necessari per dimostrare, per esempio, che una diagnosi fatta da un sistema di AI sia affidabile tanto quanto quella fatta da un medico specializzato, che il software non contenga *bias* di selezione (per esempio evitando che algoritmi di *machine learning* imparino solo sui dati di alcune categorie di pazienti, ignorandone altre), che le soluzioni proposte siano generalizzabili e clinicamente sicure ed efficaci.

Potrebbero essere adottati anche altri modelli di ricerca clinica. Per esempio, si potrebbero condurre studi clinici randomizzati che confrontano l'operato di un medico rispetto a quello del medico che usa un sistema di AI. Nel campo dell'anatomia patologica questo modello è già stato sperimentato con successo (l'algoritmo ha notevolmente accelerato la revisione dei vetrini)⁸⁷ enfatizzando così la sinergia tra il patologo e il sistema di AI.

Tutto ciò presuppone la necessità di una *governance* di questi sistemi da parte delle agenzie regolatorie italiane. Negli Stati Uniti la *Food and Drug Administration* ha affrontato la questione già da diversi anni con l'adozione di una regolamentazione che prevede la necessità di approvazione di sistemi di AI in medicina come dispositivi medici (a oggi 343 sono gli strumenti approvati come tali)⁸⁸. Nonostante l'aggiornamento a cui la regolamentazione è costantemente sottoposta, il fatto che ai fini registrativi e approvativi siano accettati anche studi retrospettivi o che forniscono prove di efficacia poco solide la rende però potenzialmente vulnerabile (<https://www.nature.com/articles/s41591-021-01312-x>). In Europa trovano, invece, spazio il Regolamento (UE) 2017/745⁸⁹ che, a partire dal 26 maggio 2021, disciplina i dispositivi medici (tra i quali rientrano i sistemi di AI in ambito diagnostico) e la nuova proposta di Regolamento europeo sull'intelligenza artificiale⁹⁰ adottata dalla Commissione europea il 21 aprile 2021, che dedica uno spazio proprio all'uso di questi sistemi in ambito medico. A tali modelli le istituzioni regolatorie italiane potrebbero ispirarsi per trovare soluzioni che possano superare i limiti della regolamentazione americana.

⁸⁷ Steiner D.F., et al. (2018), "Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer", *Am. J. Surg. Pathol.* 42, 1636–46

⁸⁸ <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices?fbclid=IwAR3LstJEfxkw0Xl3GmFUKUm6-v83ay4X1g0fPWYJ6BassG3krtDEpk7u4>

⁸⁹ https://www.salute.gov.it/portale/news/p3_2_1_1_1.jsp?lingua=italiano&menu=notizie&p=dalministero&id=5499

⁹⁰ <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206>

Sebbene l'AI venga utilizzata per ridurre gli errori medici, diagnosticare precocemente le malattie e prevederne l'insorgenza per ridurre il rischio di peggioramenti e complicanze, è opportuno inoltre considerare la presenza di alcune sfide estremamente rilevanti. Alcune sono note in letteratura e includono le seguenti (ma non si limitano a):

- Irrobustire il *framework* etico e algoretico, normativo e regolatorio in materia di protezione e condivisione dei dati, rispetto della *privacy* e consenso al trattamento^{91,92};
- Ampliare la collaborazione fra diversi centri di ricerca e unità sanitarie per condividere i dati⁹³;
- Superare le barriere interdisciplinari creando linguaggi e competenze comuni fra clinici, informatici e altre figure professionali⁹⁴;
- Certificare/validare l'affidabilità delle fonti di provenienza dei dati^{95,96};
- Gestire e coordinare molteplici standard, strutture, tipi e formati di dati⁹⁷;
- Verificare l'adeguatezza metodologica e strumentale per l'analisi dei dati⁹⁸;
- Governare e monitorare la rapida crescita dei dati eterogenei⁹⁹;
- Implementare una filiera trasparente e sicura per la raccolta, semplificazione e conversione dei dati da grezzi a raffinati, per una loro gestione e distribuzione^{100,101};
- Ripulire gli errori presenti nei dati al fine di non generare *bias* al momento della modellizzazione¹⁰²;
- Implementare una filiera *evidence-based* sull'addestramento, verifica e validazione degli algoritmi^{103,104};
- Identificare correttamente le informazioni presenti nelle prescrizioni mediche (dipende da come viene addestrato e supervisionato l'algoritmo)¹⁰⁵;

⁹¹ Paterson M. and Witzleb N. (2018), "The Privacy-Related Challenges Facing Medical Research in an Era of Big Data Analytics: A Critical Analysis of Australian Legal and Regulatory Frameworks", *J Law Med.* 26(1):188-203

⁹² Balthazar P., et al. (2018), "Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics", *J Am Coll Radiol.*, 15(3 Pt B):580-586

⁹³ Dewey M., et al. (2019), "Clinical Imaging Research: Higher Evidence, Global Collaboration, Improved Reporting, and Data Sharing Are the Grand Challenges", *Radiology*, 291(3):547-552

⁹⁴ Kusters R., et al. (2020), "Interdisciplinary Research in Artificial Intelligence: Challenges and Opportunities", *Front Big Data.*, 3:577974

⁹⁵ Househ M. and Aldosari B. (2017), "The Hazards of Data Mining in Healthcare", *Stud Health Technol Inform.* 238:80-83

⁹⁶ Markus AF, et al. (2021), "The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies", *J Biomed Inform.*, 113:103655

⁹⁷ Bukowski M, et al. (2020), "Implementation of eHealth and AI integrated diagnostics with multidisciplinary digitized data: are we ready from an international perspective?", *Eur. Radiol.*, 30(10):5510-5524

⁹⁸ Yanagihara RT, et al. (2020), "Methodological Challenges of Deep Learning in Optical Coherence Tomography for Retinal Diseases: A Review", *Transl Vis Sci Technol.*, 9(2):11

⁹⁹ Xu J., et al. (2019), "Translating cancer genomics into precision medicine with artificial intelligence: applications, challenges and future perspectives. *Hum Genet.*, 138(2):109-124

¹⁰⁰ Cutiongco MFA, et al. (2020), "Predicting gene expression using morphological cell responses to nanotopography", *Nat Commun.*, 11(1):1384

¹⁰¹ vedi nota n.143

¹⁰² Hueso M, et al. (2020), "Leveraging Data Science for a Personalized Haemodialysis", *Kidney Dis (Basel)*, 6(6):385-394

¹⁰³ Bi W.L., et al. (2019), "Artificial intelligence in cancer imaging: Clinical challenges and applications", *CA Cancer J Clin.*, 69(2):127-157

¹⁰⁴ vedi nota n.143

¹⁰⁵ Negro-Calduch E., et al. (2021), "Technological progress in electronic health record system optimization: Systematic review of systematic literature reviews", *Int J Med Inform.*, 152:104507

- Comprendere l'analisi effettuata dagli algoritmi per una corretta interpretazione e inferenza clinica¹⁰⁶;
- Confrontarsi con la carenza di strumenti tecnologici in alcuni ambiti che integrano dati clinici e analitici su più livelli, così come di sistemi che combinano vari approcci alle interazioni biologiche¹⁰⁷;
- Confrontarsi con i costi computazionali che lo sviluppo e la manutenzione di queste tecnologie richiedono¹⁰⁸.

Altre sfide riguardano, più nello specifico, il ML ad uso clinico, come per esempio considerare che la quantità non è sempre sinonimo di qualità, utilizzare dati prospettici e valutare l'interpretabilità intrinseca del modello rispetto ai potenziali *bias* di ricerca. Vi sono poi delle sfide riguardanti la necessità di implementare metriche di valutazione standard (*Model Performance Evaluation*) per osservare come un modello sia adeguato nel diagnosticare una patologia, nonché la sensibilità e la specificità dell'algoritmo stesso. A queste ultime sfide è possibile rispondere con alcune strategie che mirino alla regolarizzazione statistica di modelli “*underfitting*” o “*overfitting*” riducendo i *bias* per una maggiore precisione degli algoritmi. Altre sfide, invece, costituiscono dei veri e propri esiti non intenzionali, come il rischio di dare una fiducia eccessiva a fronte della presunta affidabilità dell'algoritmo, rinunciando a mettere in atto (o riducendone l'applicazione) meccanismi di vigilanza e supervisione^{109,110}.

Non da ultimo, è necessario creare delle linee guida sull'integrazione e il corretto utilizzo dell'AI nella diagnostica¹¹¹. A questo proposito oltre alle già citate linee guida sviluppate dalla *Food and Drug Administration*¹¹², è interessante segnalare le linee guida e i suggerimenti proposti dalla *National Academy of Medicine*¹¹³. Altrettanto importante è garantire la formazione dei professionisti (e degli utenti beneficiari, quali i pazienti)¹¹⁴ sul significato del dato modellizzato, così come raffinare gli studi sul gradimento dell'interazione fra l'uomo e la macchina (*Human-Computer Interaction – HCI, User Experience – UX, Technology Acceptance, TA*)^{115,116}.

¹⁰⁶ Schwendicke F., et al. (2020), “Artificial Intelligence in Dentistry: Chances and Challenges”, *J Dent Res.*, 99(7):769-774

¹⁰⁷ Kipp W., et al. (2018), “Artificial Intelligence in Cardiology”, *J Am Coll Card.*, 71, 2668-2679

¹⁰⁸ Gilvary C., et al. (2019), “The Missing Pieces of Artificial Intelligence in Medicine”, *Trends Pharmacol Sci.*, 40(8):555-564

¹⁰⁹ vedi nota n.78

¹¹⁰ Cabitza F., et al. (2017), “Unintended consequences of machine learning in medicine”, *JAMA*, 318,517-518

¹¹¹ Ploug T. and Holm S. (2020), “The four dimensions of contestable AI diagnostics - A patient-centric approach to explainable AI”, *Artif Intell Med.*, 107:101901

¹¹² vedi nota n.135

¹¹³ Matheny M.E., et al. (2020), “Artificial Intelligence in Health Care: A Report From the National Academy of Medicine”, *JAMA*, 323(6):509-510. doi: 10.1001/jama.2019.21579

¹¹⁴ Amann J., et al. (2020), “Explainability for artificial intelligence in healthcare: a multidisciplinary perspective”, *BMC Med Inform Decis Mak.*, 20(1):310

¹¹⁵ Rundo L., et al. (2020), “Recent advances of HCI in decision-making tasks for optimized clinical workflows and precision medicine”, *J Biomed Inform.*, 108:103479

¹¹⁶ Lee L.I.T., et al. (2019), “The Current State of Artificial Intelligence in Medical Imaging and Nuclear Medicine”, *BJR Open*, 1(1):20190037

Per concludere, occorre essere consapevoli che la convalida delle prestazioni di un algoritmo in termini di accuratezza non equivale a dimostrarne l'efficacia clinica. Come Keane e Topol osservano, “*un algoritmo con un AUC di 0,99 non vale molto se non è dimostrato che migliora i risultati clinici*”¹¹⁷. Non sappiamo ancora quale impatto possa avere l'AI applicata in contesti sanitari e in particolare in quello diagnostico e non lo sapremo fino a quando non saranno disponibili solide prove scientifiche raccolte in maniera prospettica in contesti clinici reali, frutto di una rigorosa metodologia di ricerca e di analisi statistica. Avviene così per qualunque genere di intervento farmacologico o non farmacologico prima del loro inserimento in un contesto assistenziale: l'AI non può essere un'eccezione.

4. ASPETTI ETICI E LEGALI

4.1 Il quadro normativo

Il settore dell'AI è stato oggetto, in questi ultimi anni, di una serie di interventi da parte del legislatore comunitario indirizzati a delineare, attraverso la prescritta procedimentalizzazione regolativa, un possibile quadro comune di riferimento, proprio al fine di allineare le singole discipline degli Stati membri, almeno in termini generali. Ovviamente la regolamentazione euro-unitaria *in itinere* deve tenere conto di quanto già eventualmente presente in termini di disciplina (sia comunitaria sia interna). Per quanto riguarda l'ordinamento italiano il riferimento è, anzitutto, all'attuale regolamentazione della responsabilità per danno da prodotti difettosi, a partire dalla direttiva 85/374/CEE e di cui all'originario decreto del Presidente della Repubblica 24 maggio 1988, n. 224, oggi trasfuso negli artt. 114-127 del decreto legislativo 6 settembre 2005, n. 206. Per quanto attiene invece alla posizione comunitaria, l'inizio del percorso redazionale può farsi risalire al progetto di relazione preliminare alla Risoluzione del Parlamento europeo del 16 febbraio 2017, recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica; con questa Risoluzione si è infatti voluta rappresentare l'opportunità (se non, addirittura, la necessità) di approntare una serie di norme in materia (una sorta, per l'appunto, di diritto civile della robotica, pur non essendo estranea la preoccupazione etica). Una correlativa prima analisi consente di avvertire, anzitutto, la stretta connessione risultante dalla prospettata corrispondenza (quali ordini di connessi problemi) tra ambiti di richiesta regolazione (segnatamente in tema di responsabilità) e proposta istituzione di uno specifico *status* giuridico per i robot (segnatamente in tema di personalità digitale o elettronica). Allo stesso tempo deve segnalarsi l'intervento del legislatore euro-unitario, nello specifico settore sanitario,

¹¹⁷ vedi nota n.61

con il Regolamento del Parlamento europeo e del Consiglio del 5 aprile 2017, n. 745, relativo ai dispositivi medici.

I due settori indicati (responsabilità e soggettività) hanno poi formato oggetto del riscontro da parte del (consultivo) Comitato economico e sociale europeo (CESE) che, nel documento pubblicato il 31 agosto 2017 ha preso espressa posizione (in verità parzialmente differente rispetto a quanto proposto dal Parlamento europeo), trovando poi conferma nella successiva Risoluzione del Parlamento europeo del 20 ottobre 2020, recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale. In tale direzione si colloca anche la più recente Risoluzione del Parlamento europeo del 20 gennaio 2021 sull'intelligenza artificiale, in ordine alle questioni relative all'interpretazione e applicazione del diritto internazionale nella misura in cui l'UE è interessata relativamente agli impieghi civili e militari e all'autorità dello Stato al di fuori dell'ambito della giustizia penale.

Ultimo in ordine di tempo il documento della Commissione europea, che ha pubblicato il 21 aprile 2021 la Proposta di regolamento sull'approccio europeo all'AI, espressione del primo quadro giuridico europeo in questo specifico ambito¹¹⁸. Le soluzioni normative ivi contenute dovranno formare oggetto di approvazione da parte degli Stati membri e del Parlamento europeo e la relativa entrata in vigore non avverrà prima del 2023. In tale documento i rischi dell'AI formano oggetto di puntuale valutazione, con l'obiettivo di salvaguardare i valori e i diritti fondamentali dell'Unione europea e la sicurezza degli utenti¹¹⁹. In questa direzione si comprende la previsione di una serie di pratiche vietate di intelligenza artificiale¹²⁰, in quanto in grado (in termini assolutamente generali) di incidere su comportamenti individuali o di gruppi specifici di persone con correlativi pregiudizi (effettivi o potenziali) di ordine fisico o psicologico; divieto previsto anche rispetto all'utilizzo di sistemi identificazione biometrica remota "in tempo reale" in spazi accessibili al pubblico ai fini dell'applicazione della legge, senza che ricorra alcuna delle motivazioni specificamente previste (in particolare per contrastare il crimine), volendosi così scongiurare il gravissimo pericolo del ricorso a tecniche di sorveglianza di massa¹²¹.

Si tratta, in buona sostanza, di una articolata risposta su un duplice versante, etico e giuridico, deputata a fissare, per un verso, criteri di valutazione e di impiego dei nuovi strumenti di intelligenza artificiale e diretta a individuare, per altro verso, regole di disciplina rispetto al risalente interrogativo

¹¹⁸ v. nota 5

¹¹⁹ A ciò fa riferimento quanto contenuto nel titolo I della proposta, rubricato "General Provisions"

¹²⁰ Secondo quanto espressamente previsto nel titolo II della proposta, rubricato "Prohibited Artificial Intelligence Practices", contenente peraltro il solo articolo 5

¹²¹ Le questioni correlate alle (sempre più sofisticate) tecniche di acquisizione e raccolta di dati (compresa ovviamente l'area medico-sanitaria) sono ormai oggetto di attenta analisi all'interno del complesso fenomeno dei cosiddetti Big Data. Per taluni recenti approfondimenti A.C. Di Landro, *Big Data. Rischi e tutele nel trattamento dei dati personali*, Napoli, 2020; specificamente, per quel che in questa sede rileva, A. Marchese, *Profili civilistici dell'information technology in ambito sanitario*, Napoli, 2020

«se le attuali norme giuridiche siano sufficienti per regolamentare la robotica in ambito civile o, al contrario, se sia necessario creare regole specifiche»¹²².

Di questo ordine di questioni, in punto di metodo ma con utili indicazioni nel merito, si sono altresì recentemente fatti carico, a livello interno, con particolare attenzione al settore sanitario, il Comitato Nazionale per la Bioetica e il Comitato Nazionale per la Biosicurezza, le Biotecnologie e le Scienze della Vita, che hanno esitato il 29 maggio 2020, su espressa richiesta della Presidenza del Consiglio, il parere “Intelligenza artificiale e medicina: aspetti etici”, che costituisce uno strumento di riflessione (certo non normativo ma) a supporto delle prossime iniziative legislative.

4.2 Le diverse tipologie di sistemi tra autonomia e automazione

Il tipo di problemi sollevati dall'introduzione dei sistemi robotici va attentamente rapportato alla relativa tipologia e alle corrispondenti differenziazioni strutturali e funzionali; ciò vale in termini generali rispetto agli svariati settori di utilizzo e, con tratti di chiara peculiarità, con specifico riferimento all'ambito medico. In tal senso, infatti, può essere utile richiamare la prospettata distinzione tra robot tele-operati, robot autonomi e robot cognitivi¹²³. I primi, infatti, sono sostanzialmente semplici strumenti nelle mani di un operatore che dunque li controlla dal punto di vista delle azioni da intraprendere e intraprese; i secondi, invece, sono in grado di svolgere il compito affidato e programmato in assenza di intervento umano durante la fase attuativa ed esecutiva, con una (differenziabile) autonomia (di comprensione e di azione) che consente loro di rapportarsi con l'ambiente in cui operano, sia pure con qualche distinguo tra capacità di limitata reazione e capacità di integrale percezione; i terzi, infine, sono ritenuti capaci di comportamenti cosiddetti intelligenti, nella misura in cui è ravvisabile una maggiore autonomia, aggiungendosi alla rappresentazione interna del contesto esterno il possibile adattamento alla mutevolezza dell'ambiente medesimo. A questo tipo di differenziazione corrisponde, in buona sostanza, quella (ancora più specifica dal punto di vista dell'evoluzione tecnologica) tra *calculating machine*, *computing machine* e *learning machine*, che tratteggia il passaggio dagli originari modelli “deterministici” a quelli più recenti dai tratti “intelligenti”. Al riguardo, nel rapporto tra la struttura e il funzionamento dei sistemi algoritmici, sono stati ravvisati possibili criteri distintivi deputati ad assicurare catalogazioni di senso per potere poi operare scelte di indirizzo quanto alla soluzione dei problemi rappresentati. Si è così operato un distinguo, in ragione del modo attraverso cui, sulla base del tipo di macchina utilizzata, si ottiene il risultato decisionale, tra:

¹²² A. ZORNOZA-M. LAUKYTE, *Robotica e diritto: riflessioni critiche sull'ultima iniziativa di regolamentazione in Europa*, in *Contr. e impr./Europa*, 2016, p. 808; in termini similari, in precedenza, A. SANTOSUOSSO-C. BOSCARATO-F. CAROLEO, *Robot e diritto: una prima ricognizione*, in *Nuova giur. civ. comm.*, 2012, II, p. 497; più di recente U. SALANITRO, *Intelligenza artificiale e responsabilità. La strategia della Commissione Europea*, in *Riv. dir. civ.*, 2020, p. 1248; da ultimo Ar. FUSARO, *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in *Nuova giur. civ. comm.*, 2020, p. 1344

¹²³ Il riferimento è ad A. SANTOSUOSSO-C. BOSCARATO-F. CAROLEO, *Robot e diritto*, cit., p. 498 s.

- a) le decisioni basate-su-algoritmi (*algorithm-based decisions*): sono decisioni integralmente umane, anche se basate in tutto o in parte su informazioni ottenute mediante calcoli algoritmici;
- b) le decisioni guidate-da-algoritmi (*algorithm-driven decisions*): queste decisioni sono ancora prevalentemente umane, ma in termini molto ristretti rispetto a quanto determinato dall'esito delle procedure informatiche;
- c) le decisioni determinate-da-algoritmi (*algorithm-determined decisions*): sono gli esiti dei processi digitali ad assumere le decisioni, senza alcun intervenuto umano»¹²⁴.

Proprio il salto di qualità nell'ambito del processo decisionale fornisce la misura del passaggio di grado, nel senso che al sistema automatizzato governato da algoritmi non si chiede l'attuazione di quanto deciso da un altro soggetto, quanto piuttosto di decidere in autonomia¹²⁵. Si realizza, in altri termini, una vera e propria modifica sistemica (con possibili, decisive conseguenze), transitandosi dall'utilizzo dell'algoritmo in termini di supporto a decisioni comunque assunte da parte di agenti umani all'utilizzo dell'algoritmo quale vero e proprio (autonomo) agente decisionale¹²⁶, salvo, però, verificare come il passaggio dal percorso determinista (quello del primo tipo di uso) al percorso decisionista (quello del secondo tipo di uso) rifletta effettivamente una diversità che giustifichi il parallelismo (per equivalenza) con l'agente umano.

In tal senso, infatti, il pur individuato *discrimen* dell'autonomia, rapportato alla tipologia di macchine e al correlativo grado di "intelligenza", viene in buona sostanza misurato in relazione alla provenienza del controllo operativo, se cioè totalmente esterno (e, dunque, estraneo al sistema) piuttosto che interno (ma limitato) o autoprogrammato (quindi marcatamente autonomo)¹²⁷.

4.3 Le implicazioni nella relazione medico-paziente. Il consenso informato

La particolare caratterizzazione del tipo di sistema, per come sopra differenziato, reca con sé comprensibili tratti di differenziazione in ordine alle varie problematiche che concernono l'ambito sanitario, a partire dalla incidenza nella relazione tra il medico e il paziente. Quest'ultima, infatti, si connota tradizionalmente, e trova oggi pieno riconoscimento a livello normativo all'art. 1, comma 2 legge 22 dicembre 2017, n. 219, recante norme in materia di consenso informato e di disposizioni anticipate di trattamento, quale rapporto che traduce per l'appunto una dimensione essenzialmente

¹²⁴ S. AMATO, *Biodiritto 4.0*, cit., p. 87, richiamando peraltro sul punto la sintesi operata da un documento tedesco (dal quale è tratto quanto riportato in testo)

¹²⁵ Tale è il rilievo di A. SIMONCINI, *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in *BioLaw Journal*, n. 1, 2019, p. 69.

¹²⁶ Affida alla diversità dell'uso che verrà fatto dell'algoritmo la possibilità di ricorrere alle regole vigenti o a nuove elaborabili discipline E. PELLECCIA, *Profilazione e decisioni automatizzate al tempo della black box society: qualità dei dati e leggibilità dell'algoritmo nella cornice della responsible research and innovation*, in *Le nuove leggi civ. comm.*, 2018, p. 1209 ss.; proprio rispetto al concetto chiave di autonomia, quale elemento di svolta nel rapporto uomo-cosa, rispetto al tipo di controllo esercitabile dall'uomo sul robot "intelligente", v., in punto di individuazione delle regole applicabili, N.F. FRATTARI, *Robotica e responsabilità da algoritmo. Il processo di produzione dell'intelligenza artificiale*, in *Contr. e impr.*, 2020, p. 458 ss.

¹²⁷ Indicazioni in tal senso sono rinvenibili in N.F. FRATTARI, *Robotica e responsabilità da algoritmo*, cit., p. 468 s., che corrispondentemente distingue tra assenza di autonomia, autonomia debole e autonomia forte.

relazionale, caratterizzata dai tratti contenutistici della cura e della fiducia e che il legislatore stesso ha assunto il compito di promuovere e valorizzare. Del resto, soprattutto con riferimento alla cura delle persone anziane, è stato evidenziato che tra i fattori che influiscono sulla sensazione di sicurezza di questo tipo di pazienti è proprio l'aver relazioni sicure con il personale sanitario, avere il controllo delle routine quotidiane ed esserne informati.

Una particolare rilevanza assume, pertanto, l'aspetto comunicativo, deputato a rendere effettivo ed efficace il rapporto tra medico e paziente. Comunicare significa mettere qualcuno a parte di qualcosa e, più in particolare per quanto qui di interesse, essere in rapporto scambievole con qualcuno¹²⁸; si tratta, allora, di un modello che presenta una necessaria connotazione etica, importando ciò, in positivo, l'adesione al corrispondente carattere interpersonale di tale processo (che impone di riconoscere l'altro come persona) e, in negativo, l'impossibilità di limitare la comunicazione a una mera informazione (il che tradirebbe la stessa valenza etimologica della terminologia utilizzata)¹²⁹.

L'elemento distintivo e fondamentale di una comunicazione significativa, in grado cioè di non rendere estraneo e non appropriato il processo comunicativo, viene individuato in uno svolgimento caratterizzato, attraverso l'utilizzo di un linguaggio adeguato, da un doppio binario; quest'ultimo, infatti, traduce alcune delle coordinate fondamentali della comunicazione tra medico e paziente, quello cioè del comprendere (in sintonia, per il tramite dell'ascolto, con l'interlocutore) e quello dello spiegare (sia per far capire sia per indurre un cambiamento nell'interlocutore)¹³⁰. Solo un'adeguata attenzione a una delle esigenze fondamentali dell'essere umano, per l'appunto il comunicare, può dunque consentire a tale processo appropriata efficacia e potere curativo (rappresentando ciò, propriamente, l'obiettivo legislativo, pur indicato in termini descrittivi), in quanto diretto ad assicurare la necessaria piena umanità della relazione terapeutica¹³¹, nella quale, come consapevolmente avvertito, «la personalizzazione della cura implica l'incontro tra il sapere su di sé del paziente e il sapere professionale del medico»¹³². Quando allora si mettono in relazione tali coordinate con il tema dell'intelligenza artificiale e, più specificamente, con la possibilità che attraverso un sistema di assistenza cognitiva automatizzata nell'attività diagnostica si provveda alla raccolta di dati clinici e documentali, determinando poi la elaborazione degli stessi per realizzare il richiesto processo di analisi, si avvertono subito taluni, non agevoli punti di snodo. In primo luogo, la necessità di mantenere a livello umano e personale (non, dunque, delegabile al sistema intelligente) il rapporto con il paziente. Correlativamente,

¹²⁸ In tal senso I. CARRASCO DE PAULA, *La comunicazione medico-paziente: elementi per una fondazione etica*, in *Medicina e Morale*, 2002, p. 609.

¹²⁹ Il rischio dell'impoverimento del processo comunicativo (come se fosse solo una tecnica quantunque appaia tecnicabile) è segnalato, già, da CARRASCO DE PAULA, *La comunicazione medico-paziente*, cit., p. 610 ss., il quale indica un radicale mutamento di prospettiva rispetto alla medicina classica, in cui predominava il medico.

¹³⁰ Indicazioni in tal senso in CAGLI, *La comunicazione tra medico e paziente*, in *L'arco di Giano*, 2004, n. 41, p. 69 ss.

¹³¹ È l'auspicio di ZATTI, «Parole tra di noi così diverse». *Per una ecologia del rapporto terapeutico*, in *Nuova giur. civ. comm.*, 2012, II, p. 150; di parallela evoluzione di una relazione umana e comunicativa discute VALLINI, *Pianificazione delle cure, medicina palliativa. I nuovi paradigmi del "fine vita"*, in *Riv. it. med. leg.*, 2016, p. 1150 ss.

¹³² NICOLUSSI, *Al limite della vita: rifiuto e rinuncia ai trattamenti sanitari*, in *Quad. cost.*, 2010, p. 276

quale riflesso di questo criterio ordinante e permanente, va preservato il percorso destinato ad acquisire il consenso (autonomo, libero e consapevole) da parte del paziente attraverso un'opera di informazione che presenta, tuttavia, talune difficoltà. Non è infatti revocabile in dubbio che la comunicazione da parte del medico al paziente dell'utilizzo di nuovi strumenti comporta anche la spiegazione, in termini comprensibili e chiari, delle relative modalità operative. Ciò indubbiamente può non essere agevole sia rispetto alla capacità comunicativa da parte del medico sia rispetto alla capacità di comprensione da parte di un paziente non sempre in possesso di conoscenze generali ben più elementari. In altri termini l'efficacia dell'azione comunicativa, rispetto a un oggetto altamente complesso, va parametrata sempre e dipende dal tipo di paziente sottoposto a un determinato trattamento.

Proprio per rendere la comunicazione più efficace, vanno preventivamente messe in atto misure dirette a migliorare le conoscenze e competenze in materia di AI sia del personale medico, all'interno del percorso universitario e post-universitario, sia del cittadino, agendo al livello della formazione secondaria superiore e predisponendo un'informazione neutrale, scientificamente corretta ma chiara e comprensibile, strutturata per livelli, da offrire sui canali Web mediante contenuti testuali, illustrativi, filmati, sulle due direttrici della metodologia (come vengono gestiti e usati i dati, come giungono a un risultato gli algoritmi di AI) e della consapevolezza civile ed etica (quali sono i diritti del paziente rispetto ai suoi dati, come esercitarli, che cosa è un consenso informato e perché è importante concederlo conoscendone gli effetti).

Infatti, proprio la EU insiste sul concetto di “trust”¹³³, “fiducia”, da costruire nei cittadini mediante una comunicazione schietta e comprensibile, accessibile a tutti. Ciò principalmente per prevenire la naturale diffidenza che potrebbe rallentare l'adozione di soluzioni basate su AI anche in casi di comprovata efficacia e sicurezza.

Inoltre, come da più parti segnalato, la possibile opacità da parte del sistema di intelligenza artificiale, ossia la non effettiva ricostruibilità del percorso seguito per giungere a una determinata soluzione, non consente di spiegare al paziente come si sia pervenuti a una certa conclusione o si siano prospettate determinate soluzioni.

4.4 Raccolta, trattamento e condivisione dei dati personali sanitari

Quanto poi al tema che in termini generali viene ricondotto alla protezione della *privacy* dell'interessato rispetto ai propri dati personali sanitari occorre preliminarmente mettere in evidenza la diversità delle possibili azioni che qui vengono in considerazione. Anzitutto la relativa acquisizione ai fini del successivo trattamento presuppone la necessità del previo consenso (informato ed esplicito), alla luce del generale principio di trasparenza di cui all'art. 5 Regolamento UE 2016/679 che deve

¹³³ https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence_en

regolare il profilo informativo¹³⁴, proprio laddove, come nel caso, di specie, raccolta e trattamento dei dati siano affidati, già a monte e per la successiva elaborazione, al sistema automatizzato che, a propria volta, si rapporta con un contesto sempre più ricco di dati da incrociare. Peraltro, la tecnica oggi ritenuta più sicura e adeguata all'archiviazione dei dati personali sanitari è quella connessa all'utilizzo della particolare infrastruttura *web* nota come *cloud computing* (il cui sistema ibrido di tipo *community* consente di superare anche talune, possibili rilevate criticità).

In corrispondenza con il dato tecnico (della maggiore e più estesa condivisione dei dati) il dato normativo utilizza, nella distinzione giuridica tra interessato, titolare e responsabile del trattamento, il concetto della co-titolarità nel trattamento dei dati di cui all'art. 26 Regolamento UE 2016/679, in presenza di una pluralità di soggetti deputati a condividere la gestione dei dati medesimi, delineandone così i correlativi profili di responsabilità. Questo criterio può consentire, rispetto a un sistema caratterizzato dalla più ampia possibilità tecnica di condivisione dei dati e in riscontro alla corrispondente sempre più pressante richiesta scientifica in tal senso, di introdurre efficaci azioni nell'ottica della migliore tutela.

4.5 La disciplina della responsabilità nei rapporti tra produttore e operatore/utilizzatore

Messo in disparte il tema della possibile soggettività degli agenti *software*, il versante della responsabilità derivante dal relativo utilizzo costituisce, ad oggi, il maggior problema, a cui è stata peraltro rivolta particolare attenzione da parte del legislatore sovranazionale.

Con riguardo allo specifico tema della responsabilità civile per i danni causati dai robot lo stesso Parlamento europeo nella Risoluzione del 16 febbraio 2017, ha già ritenuto che «almeno nella fase attuale, la responsabilità deve essere imputata a un essere umano e non a un robot» (punto 56). Come qui allo stesso tempo rilevato «la tendenza all'automazione esige che i soggetti coinvolti nello sviluppo e nella commercializzazione di applicazioni dell'intelligenza artificiale integrino gli aspetti relativi alla sicurezza e all'etica fin dal principio, riconoscendo pertanto che devono essere preparati ad accettare di essere legalmente responsabili della qualità della tecnologia prodotta». Dall'altro, e parimenti, si rileva che «alcuni Stati esteri quali Stati Uniti, Giappone, Cina e Corea del Sud stanno prendendo in considerazione, e in una certa misura hanno già adottato, atti normativi in materia di robotica e intelligenza artificiale, e che alcuni Stati membri hanno iniziato a riflettere sulla possibile elaborazione di norme giuridiche o sull'introduzione di cambiamenti legislativi per tenere conto delle applicazioni emergenti di tali tecnologie».

¹³⁴ In proposito v. Garante per la protezione dei dati personali, 7 marzo 2019, *Chiarimenti sull'applicazione della disciplina per il trattamento dei dati relativi alla salute in ambito sanitario*, in www.garanteprivacy.it.

Nella prospettiva delineata la riparazione del danno prodotto viene dunque ancora affidata, nella tendenzialmente comune rappresentazione dei sistemi intelligenti in termini di strumenti attraverso quel modello di responsabilità, che muove dalla produzione del bene, ma può non escludere, in relazione allo specifico caso concreto, la posizione dell'utilizzatore-proprietario. In realtà è proprio la già sussistente via della responsabilità del produttore a essere ritenuta preferibile, in considerazione della agevolata posizione dell'attore danneggiato che deve dimostrare l'esistenza del danno e del difetto, unitamente al nesso causale tra difetto e danno¹³⁵. Da questo punto di vista il richiamo all'assicurazione obbligatoria viene ritenuto «certamente utile per consentire di affrontare, almeno temporaneamente, gli aspetti connessi alla responsabilità dal punto di vista economico»¹³⁶.

Peraltro la generale e ordinante prospettiva non sembra mutare anche con riferimento all'ipotesi di sistemi sempre più complessi e sofisticati, la cui condizione di autonomia e imprevedibilità viene ritenuta piuttosto ragione necessaria per evitare percorsi giuridici di deresponsabilizzazione dell'uomo nella costruzione e nell'uso di robot, rilevandosi che «la responsabilità umana, ancorché in forme magari non completamente sovrapponibili all'esistente, deve aumentare progressivamente, a misura della accresciuta capacità tecnica di introdurre nell'ambiente potenziali nuove fonti di rischio»¹³⁷. Attualmente il punto di riferimento è costituito dalle già riferite Risoluzione del Parlamento europeo del 20 ottobre 2020 e conseguente Proposta di regolamento della Commissione europea del 21 aprile 2021, proprio sul tema dei rapporti tra responsabilità e intelligenza artificiale, sulla base della precedente Relazione al riguardo¹³⁸ e del Libro bianco sull'intelligenza artificiale¹³⁹. In merito, infatti, ritenendosi nella Risoluzione opportuno concentrarsi «sulle azioni per responsabilità civile nei confronti dell'operatore di un sistema di AI» (punto 10) attraverso correlative differenti previsioni per i diversi tipi di rischi, si prospetta l'introduzione di un regime comune di responsabilità oggettiva per sistemi di intelligenza artificiale (autonomi) a rischio intrinseco elevato e un regime di responsabilità per colpa per tutti gli altri sistemi di intelligenza artificiale (dunque non autonomi e in assenza di un rischio intrinseco elevato) (punti 14 ss.).

Sulla base di tali indicazioni e riprendendo quanto sostanzialmente contenuto nell'Allegato alla Risoluzione, in ordine al possibile testo da predisporre, la Commissione europea ha elaborato una Proposta di regolamento (non a caso in termini di *artificial intelligence act*) che si caratterizza, nei limiti di quel che in questa sede può rilevare, per una calibrata prospettazione delle questioni che interessano

¹³⁵In tal senso E. PALMERINI, *Robotica e diritto*, cit., p. 1839, al termine di un'attenta disamina delle plurime proposte delineate per affrontare il cosiddetto *responsability gap*.

¹³⁶A. ZORNOZA-M. LAUKYTE, *Robotica e diritto*, cit., p. 812 s., peraltro critici in ordine all'idea di creare un registro specifico dei robot cosiddetti avanzati, ritenuta inquisitoria, anche in ragione del fatto che tutti i prodotti fabbricati e commercializzati nell'Unione Europea hanno (già) un numero di registro; riferimenti anche in E. PALMERINI, *Robotica e diritto*, cit., p. 1842.

¹³⁷E. Palmerini, *Robotica e diritto*, cit., p. 1840.

¹³⁸Si fa qui riferimento alla Relazione della Commissione al Parlamento europeo, al Consiglio e al Comitato economico e sociale europeo, *Relazione sulle implicazioni dell'intelligenza artificiale, dell'Internet delle cose e della robotica in materia di sicurezza e di responsabilità*, 19 febbraio 2020, COM (2020) 64 final

¹³⁹v. nota 4.

i quadri giuridici della sicurezza e della responsabilità; ciò, anche sulla base della disamina compiuta dal gruppo di esperti sulla responsabilità e le nuove tecnologie, istituito per fornire consulenza alla Commissione in merito all'applicabilità della direttiva sulla responsabilità per danno da prodotti difettosi e delle norme nazionali in materia di responsabilità civile, rispetto altresì alla elaborazione di orientamenti per possibili adeguamenti della normativa applicabile rispetto alle nuove tecnologie. Deve, peraltro, allo stesso tempo rilevarsi che lo specifico regime di responsabilità sembra sostanzialmente ricavabile da una serie di previsioni che traducono un certo tipo di orientamento, con evidente spazio per un intervento successivo, quantunque allo stesso tempo si tenda a fare chiarezza su taluni profili oggetto del dibattito in corso.

Risulta così recepita la differenziazione - fermo il già richiamato divieto di utilizzo di sistemi di intelligenza artificiale dal rischio inaccettabile - tra sistemi di intelligenza artificiale a rischio elevato o meno; in particolare, ai primi è dedicata la puntuale regolazione contenuta nel titolo III, dedicato ai "High-Risk AI Systems" (artt. 6-51). In merito, infatti, nella strutturazione dell'approntato schema normativo appare evidente la correlazione tra tipologie di sistemi, attraverso la classificazione contenuta nel relativo capitolo I quanto ai sistemi a rischio elevato (artt. 6 e 7) e di cui all'allegato III; requisiti degli stessi, secondo le indicazioni contenute nel relativo capitolo II (artt. 8-15); corrispondenti obbligazioni dei soggetti coinvolti, con consequenziali assetti di responsabilità, secondo il quadro delineato nel capitolo III (artt. 16-29).

Il modello prospettato mette pertanto in stretta correlazione tipologie di sistemi di intelligenza artificiale e corrispondenti modalità di immissione nel circuito di utilizzo, con la connessa individuazione dei soggetti che vi possono avere a che fare, sia pure con ruoli differenti, rispetto agli obblighi gravanti su ciascuno, con l'obiettivo di garantire adeguati standard di tutela e, soprattutto, assicurare la predeterminazione dei diversi livelli di responsabilità¹⁴⁰. Da un lato, pertanto, la previa immissione sul mercato di un sistema di intelligenza artificiale a rischio elevato implica la sottoposizione a una serie di controlli preordinati a garantire la sicurezza, attraverso una valutazione di conformità; dall'altro, il tema di un'appropriata progettazione e del consequenziale sviluppo del sistema di intelligenza artificiale risulta correlato alla possibilità di un'effettiva ed efficace sorveglianza da parte di persone fisiche allorquando lo stesso è in uso.

Vengono così poi fissati gli obblighi in capo ai vari soggetti interessati, ossia fornitori, importatori, distributori, utilizzatori o altre terze parti, diverse dunque da tutte le altre in precedenza indicate. Per ciascuna di queste figure sono previsti specifici obblighi in ragione della corrispondente posizione rispetto al sistema di intelligenza artificiale. Così, per esemplificare, in capo al fornitore è posto,

¹⁴⁰ Si tratta, peraltro, di un tipo di approccio (unitamente a quanto allo stesso più puntualmente correlato) che era stato indicato da parte del gruppo di esperti e di cui dà conto, da ultimo, R. MONTINARO, *Responsabilità del prodotto difettoso e tecnologie digitali tra soft law e hard law*, in *Persona e mercato*, 2020, 4, p. 375 ss. e spec. p. 378 s., rilevando altresì che ciò era accaduto indipendentemente dalla considerazione di uno specifico regime di responsabilità civile.

anzitutto (tra le altre previsioni contenute nell'art. 16), un obbligo di garanzia in ordine alla conformità del sistema di intelligenza artificiale rispetto a quanto richiesto; mentre, proprio in ragione del differente ruolo, sarà compito del distributore (ai sensi dell'art. 27) verificare la sussistenza della documentazione deputata a certificarne la conformità (di competenza, come visto, del fornitore); infine, l'utilizzatore (che nel settore qui di interesse corrisponde al medico) è tenuto all'uso conforme alle istruzioni, risultando viceversa assimilabile, dal punto di vista della responsabilità, al fornitore ove abbia apportato una modifica sostanziale al sistema di intelligenza artificiale (al pari, secondo quanto previsto dall'art. 28, di distributori, importatori o altre terze parti a cui sia riferibile un comportamento tale da incidere, nei diversi modi indicati, sul sistema immesso sul mercato). In questo contesto appare chiaro che condotte tenute in violazione degli obblighi normativamente stabiliti non possono non generare responsabilità.

Si tratta di un modello che sembra riproporre, per un verso, quanto già suggerito a proposito della sostanziale tenuta della disciplina in tema di danno da prodotto difettoso, richiamandosi cioè il generale principio (da applicare anche nel campo dell'intelligenza artificiale) secondo cui i prodotti che non soddisfano le norme di sicurezza obbligatorie sono considerati difettosi (indipendentemente dalla colpa del produttore)¹⁴¹; che non sembra escludere, per altro verso, con particolare riferimento ai sistemi di intelligenza artificiale con un profilo di rischio specifico in termini di responsabilità, la già prospettata apertura a regimi di responsabilità oggettiva¹⁴².

5. CONCLUSIONI E PROPOSTE OPERATIVE

I sistemi di AI stanno entrando rapidamente nel mondo medico, con la prospettiva, più o meno a breve termine, di modificare significativamente i percorsi diagnostici e terapeutici, le modalità decisionali del Medico e, in ultimo, anche il rapporto Medico-Paziente. Prima di poter essere impiegati in ambito assistenziale, i sistemi di AI necessitano di una rigorosa validazione scientifica, basata su studi metodologicamente solidi (prospettici, possibilmente randomizzati e condotti in ambienti clinici

¹⁴¹ Sul punto la Relazione della Commissione al Parlamento europeo, al Consiglio e al Comitato economico e sociale europeo, *Relazione sulle implicazioni dell'intelligenza artificiale*, cit., p. 16, riprendendo la posizione del gruppo di esperti, rileva che «si potrebbero prevedere adeguamenti delle leggi nazionali per agevolare l'onere della prova per le vittime di danni dovuti all'intelligenza artificiale. Ad esempio l'onere della prova potrebbe essere collegato al rispetto (da parte dell'operatore interessato) di specifici obblighi in materia di cibersicurezza o di altri obblighi di sicurezza stabiliti dalla legge: in caso di mancato rispetto di tali obblighi, potrebbe applicarsi una modalità dell'onere della prova per quanto riguarda la colpa e il nesso di causalità».

¹⁴² Anche in tal caso la Relazione della Commissione al Parlamento europeo, al Consiglio e al Comitato economico e sociale europeo, *Relazione sulle implicazioni dell'intelligenza artificiale*, cit., p. 18, alla stregua della posizione del gruppo di esperti, «sollecita osservazioni e commenti sulla necessità di e sulla misura in cui possa essere necessaria la responsabilità oggettiva, quale prevista dalle legislazioni nazionali per rischi analoghi ai quali è esposto il pubblico (ad esempio per il funzionamento di veicoli a motore, aerei o centrali nucleari), per consentire il risarcimento efficace delle possibili vittime. La Commissione sollecita inoltre osservazioni e commenti sulla necessità di abbinare alla responsabilità oggettiva l'eventuale obbligo di sottoscrivere un'assicurazione disponibile, seguendo l'esempio della direttiva sull'assicurazione auto, al fine di garantire il risarcimento indipendentemente dalla solvibilità del responsabile e di contribuire a ridurre i costi dei danni».

reali) che dimostrino la non inferiorità, o la superiorità, oltreché la costo-efficacia, rispetto al percorso diagnostico e decisionale convenzionale. È, inoltre, importante dimostrare la sicurezza e la riproducibilità nell'uso del software, nonché considerare le emergenti problematiche etiche e legali inerenti la responsabilità professionale del Medico nell'interazione con gli algoritmi.

Per poter introdurre in modo sicuro nella pratica clinica i sistemi di AI e per competere in ambito internazionale nella realizzazione e nello sviluppo degli stessi, è auspicabile che, nel nostro Paese, siano attuati i seguenti interventi:

- realizzazione di una infrastruttura organizzativa, informatizzata, a livello locale, regionale o nazionale, di *data stewardship* e *data governance*;
- creazione di una struttura di *governance* dei sistemi di AI da parte delle agenzie regolatorie italiane, in particolare il Ministero della Salute per ciò che riguarda i dispositivi medici e AIFA per gli eventuali aspetti terapeutici, con lo scopo di stabilire delle regole rigorose per l'approvazione e la registrazione di tali sistemi;
- predisposizione di Linee Guida nazionali riguardanti le modalità di integrazione e il corretto utilizzo dei sistemi di AI nella diagnostica, in accordo con le società scientifiche di riferimento;
- creazione di un osservatorio nazionale permanente presso il Ministero della Salute, per il monitoraggio delle performance dei sistemi di AI immessi sul mercato (*analisi post-market*);
- predisposizione di moduli formativi universitari e post-universitari per migliorare le conoscenze e competenze in materia di AI del personale medico e delle professioni sanitarie;
- integrazione di elementi metodologici in tema di AI all'interno dei programmi della scuola secondaria superiore e creazione di contenuti informativi, anche tramite canali informatici, al servizio del cittadino.

6. APPENDICE

Tabella 1. Criteri di inclusione della metodologia di ricerca.

| Search engine | Mediline |
|-------------------|---|
| Date | 25/08/2021 |
| Query | "artificial intelligence" AND ["diagnosis" OR "prediction"] |
| Field | All |
| Text availability | Abstract |
| Publication date | 5 years (2016-2022) |
| Language | English |
| Article type | meta-analysis; systematic review |
| Total results | 563 |

Tabella 2. Tipologia di modellizzazione con esempi pratici di utilizzo del ML in ambito diagnostico dalla letteratura.

| Tipologia | Esempi pratici di utilizzo del ML in ambito diagnostico |
|---|---|
| Regressione lineare (<i>linear regression</i>) | Viene applicata in moltissime analisi e predizioni computazionali, dall'identificazione di fattori di rischio prognostici rilevanti ¹⁴³ alla comprensione dei pattern prevalenti di immunodeficienza da HIV umana ¹⁴⁴ . |
| Regressione logistica (<i>logistic regression</i>) | Viene applicata per ottenere modelli predittivi ed esplicativi ottenendo profili di rischio nel caso di patologie complesse (es. tubercolosi ¹⁴⁵ , tumore della mammella ¹⁴⁶), la diagnosi di patologie cardiache coronariche ¹⁴⁷ o il linfoma non-Hodgkin ¹⁴⁸ . |
| Analisi discriminante (<i>discriminant analysis</i>) | Il suo utilizzo consente di predire precocemente la diagnosi clinica di immunodeficienze primarie ¹⁴⁹ , identificare i sintomi del morbo di Parkinson ¹⁵⁰ , valutare il rischio di malattie croniche ¹⁵¹ o diagnosticare l'ipercalcemia ¹⁵² . |
| Support Vector Machine (SVM) | Viene utilizzato in una grande varietà di contributi che vanno dalla classificazione dei sintomi all'analisi per migliorare l'accuratezza diagnostica, all'identificazione dei biomarcatori di <i>imaging</i> per le patologie neurologiche e psichiatriche, alla validazione per il tumore della mammella e della cervice analizzando i dati |

¹⁴³ Schneider A., et al. (2010), "Linear regression analysis: part 14 of a series on evaluation of scientific publications", *Deutsches Arzteblatt International*, 107, 776–782X

¹⁴⁴ Madigan E.A., et al. (2008), "Workforce analysis using data mining and linear regression to understand HIV/AIDS prevalence patterns", *Hum. Res. Health*, 6, 2

¹⁴⁵ Xu W., et al. (2018), "Differential analysis of disease risk assessment using binary logistic regression with different analysis strategies", *J. Int. Med. Res.*, 46, 3656–3664

¹⁴⁶ García-Laencina P.J., et al. (2015), "Missing data imputation on the 5-year survival prediction of breast cancer patients with unknown discrete values", *Comput. Biol. Med.*, 59, 125–133

¹⁴⁷ Xu W., et al. (2018), "Differential analysis of disease risk assessment using binary logistic regression with different analysis strategies", *J. Int. Med. Res.*, 46, 3656–3664

¹⁴⁸ Nick T.G., et al. (2007), "Topics in biostatistics", *Methods Mol. Biol.*, 404

¹⁴⁹ Murata C., et al. (2019), "Discriminant analysis to predict the clinical diagnosis of primary immunodeficiencies: a preliminary report", *Rev. Alerg. Mex.*, 62, 125–133

¹⁵⁰ Armañanzas R., et al. (2013), "Unveiling relevant non-motor Parkinson's disease severity symptoms using a machine learning approach", *Artif. Intell. Med.*, 58, 195–202

¹⁵¹ Jen X C.-H., et al. (2012), "Application of classification techniques on development an early-warning system for chronic illnesses", *Expert Syst. Appl.* 39, 8852–8858

¹⁵² Johnson K.R., et al. (1984), "Differential laboratory diagnosis of hypercalcemia", *Crit. Rev. Clin. Lab. Sci.*, 21, 51–97.

| | |
|--|--|
| | clinici, patologici ed epidemiologici, alla validazione del cancro orale attraverso i dati clinici, genomici e molecolari ^{153,154,155,156,157} . |
| Naïve Bayes | Viene utilizzato per la modellizzazione di diverse patologie oncologiche (cervello, mammella, prostata) e non (asma ¹⁵⁸ ; Alzheimer ¹⁵⁹), per classificare le cartelle elettroniche ¹⁶⁰ , per il supporto della decisione clinica per le patologie cardiache ¹⁶¹ . |
| K-nearest neighbor (K-NN) | Viene utilizzato per modellizzare la performance diagnostica ¹⁶² , per classificare i set di dati sanitari ¹⁶³ , per classificare le metastasi dei linfonodi nel tumore gastrico ¹⁶⁴ e per la predizione del tumore pancreatico combinato con la letteratura pubblicata e le cartelle elettroniche ¹⁶⁵ . |
| Deep Learning | Viene utilizzato per la diagnostica in moltissimi ambiti clinici, soprattutto oncologici ^{166,167,168,169} , dermatologici ¹⁷⁰ , per la retinopatia diabetica ¹⁷¹ e le anomalie da insufficienza cardiaca ¹⁷² , analizzando immagini provenienti da radiografie e marcatori genomici. |
| Modello di Markov nascosto (Hidden Markov Model - HMM) | Inizialmente utilizzato per risolvere i problemi legati al linguaggio, viene utilizzato per modellizzare stati “salutari” e “non salutari” ¹⁷³ sulla base dei dati dei pazienti, minimizzare le reazioni avverse ai farmaci dall’analisi di forum sanitari online ¹⁷⁴ o monitorare l’attività circadiani ¹⁷⁵ . |

¹⁵³ Huang S., et al. (2018), “Applications of support vector machine (SVM) learning in cancer genomics”, *Cancer Genom. Proteom.* 15, 41–51

¹⁵⁴ Cho G., et al. (2019), “Review of machine learning algorithms for diagnosing mental illness”, *Psychiatry Investig.*, 16, 262–269

¹⁵⁵ Cruz J.A. and Wishart D.S. (2007), “Applications of machine learning in cancer prediction and prognosis”, *Cancer Inform.*, 2, 59–77

¹⁵⁶ Huang S, et al. (2018), “Applications of Support Vector Machine (SVM) Learning in Cancer Genomics”, *Cancer Genomics Proteomics*, 15(1):41-51

¹⁵⁷ Ahmed K., et al. (2020). “Artificial Intelligence with multi-functional machine learning platform development for better healthcare and precision medicine”, *Database*, 1-35

¹⁵⁸ Langarizadeh M. and Moghbeli F. (2016), “Applying naive Bayesian networks to disease prediction: a systematic review”, *AIM: journal of the Society for Medical Informatics of Bosnia & Herzegovina*, 24, 364–369

¹⁵⁹ Wei W., et al. (2011), “The application of naive Bayes model averaging to predict Alzheimer’s disease from genome-wide data”, *JAMIA*, 18, 370–375

¹⁶⁰ Al-Aidaros K.M, et al. (2012), “Medical data classification with naïve Bayes approach”, *Inform. Technol. J.*, 11, 1166–1174

¹⁶¹ Srinivas K., et al. (2010), “Applications of data mining techniques in healthcare and prediction of heart attacks”, *Int. J. Comput. Sci. Eng.*, 2, 250–255

¹⁶² Zhang Z. (2016), “Introduction to machine learning: k-nearest neighbors”, *Ann. Translat. Med.*, 4, 218

¹⁶³ Hu L.Y., et al. (2016), “The distance function effect on k-nearest neighbor classification for medical datasets”, *SpringerPlus*, 5, 1304

¹⁶⁴ Sarkar M. and Leong T.Y. (2000), “Application of K-nearest neighbors algorithm on breast cancer diagnosis problem”, *AMIA Symposium*, 2000, 759–763

¹⁶⁵ Zhao D. and Weng C. (2011), “Combining PubMed knowledge and EHR data to develop a weighted bayesian network for pancreatic cancer prediction”, *J. Biomed. Inform.*, 44, 859–868

¹⁶⁶ Hosny A., et al. (2018), “Artificial intelligence in radiology”, *Nat. Rev. Cancer*, 18, 500–510

¹⁶⁷ Langlotz C.P., et al. (2019), “A roadmap for foundational research on artificial intelligence in medical imaging from the 2018 NIH/RSNA/ACR/the Academy Workshop”, *Radiology*, 291, 781–791

¹⁶⁸ Haenssle H.A., et al. (2018), “Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists”, *Ann. Oncol.*, 29, 1836–1842

¹⁶⁹ Esteva A., et al. (2017), “Dermatologist-level classification of skin cancer with deep neural networks”, *Nature*, 542, 115–118

¹⁷⁰ Olsen T.G., et al. (2018), “Diagnostic performance of deep learning algorithms applied to three common diagnoses in dermatopathology”, *J. Pathol. Inform.*, 9, 32

¹⁷¹ Gulshan V., et al. (2016), “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs”, *JAMA*, 316, 2402–2410

¹⁷² Rajkomar A., et al. (2018), “Scalable and accurate deep learning with electronic health records”, *Nat. Digital Med.*, 1, 18

¹⁷³ Wall M. M. and Li R. (2009), “Multipl eindicator hidden Markov model with an application to medical utilization data”, *Stat. Med.*, 28, 293-310

¹⁷⁴ Sampathkumar H., et al. (2014), “Mining adverse drug reactions from online healthcare forums using hidden Markov model”, *BMC Med. Inform. Dec. Mak.*, 14, 91

¹⁷⁵ Huang Q., et al. (2018), “Hidden Markov models for monitoring circadian rhythmicity in telemetric activity data”, *J. R. Soc. Interface*, 15, 20170885

| | |
|--|--|
| Albero di decisione (<i>decision tree</i>) | Viene utilizzato in ambito medico a supporto della diagnostica e del <i>decision-making</i> ¹⁷⁶ , per individuare i fattori legati all'ipertensione ¹⁷⁷ o scoprire fattori associati alla pressione ulcerosa nei soggetti anziani ¹⁷⁸ . |
| Foresta casuale (<i>random forest</i>) | Fra le applicazioni di questo modello rientra la diagnosi della malattia di Alzheimer ¹⁷⁹ e la sua classificazione ¹⁸⁰ , la predizione della mortalità dei pazienti in terapia intensiva ¹⁸¹ e la predizione di <i>pathway</i> metabolici ¹⁸² . |
| Algoritmo genetico (<i>genetic algorithm</i>) | Viene utilizzato specialmente nel campo della radiologia, oncologia, cardiologia, endocrinologia, pediatria, chirurgia, pneumologia, malattie infettive, radioterapia, riabilitazione, ortopedia, neurologia, farmacoterapia, ginecologia e ostetricia ¹⁸³ . Alcuni esempi pratici riguardano l'identificazione delle microcalcificazioni per il tumore alla mammella ^{184,185} , lo sviluppo di tecniche non invasive per la diagnosi di tumore alla cervice ¹⁸⁶ , analizzare i microarray dalle linee di cellule cancerose ¹⁸⁷ , predizione degli outcome di pazienti con tumore polmonare non a piccole cellule (<i>non-small cell lung cancer</i> – NSCLC ¹⁸⁸) o mutazioni della malattia di Parkinson ¹⁸⁹ . |

¹⁷⁶ Bae J.M.T. (2014), "Clinical decision analysis using decision tree", *Epidemiol. Health*, 36, e2014025

¹⁷⁷ Tayefi M., et al. (2017), "The application of a decision tree to establish the parameters associated with hypertension", *Comput. Methods Programs Biomed.*, 139, 83–91

¹⁷⁸ Moon M. and Lee S.K. (2017), "Applying of decision tree analysis to risk factors associated with pressure ulcers in long-term care facilities", *Health. Inform. Res.*, 23, 43–52

¹⁷⁹ O'Bryant S.E., et al. (2010), "A serum protein-based algorithm for the detection of Alzheimer disease", *Arch. Neurol.*, 67, 1077–1081

¹⁸⁰ Byeon H. (2019), "Developing a random forest classifier for predicting the depression and managing the health of caregivers supporting patients with Alzheimer's disease", *Technol. Health Care*, 1–14

¹⁸¹ Lee J. (2017), "Patient-specific predictive modeling using random forests: An observational study for the critically ill", *JMIR Med. Inform.*, 5, e3

¹⁸² Toubiana D., et al. (2019), "Combined network analysis and machine learning allows the prediction of metabolic pathways from tomato metabolomics data", *Commun. Biol.*, 2, 214

¹⁸³ Ghaheri A., et al. (2015), "The applications of genetic algorithms in medicine", *Oman Med. J.*, 30, 406–416

¹⁸⁴ Karnan M. and Thangavel K. (2007), "Automatic detection of the breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications", *Comput. Methods Programs Biomed.*, 87, 12–20

¹⁸⁵ Pereira D.C., et al. (2014), "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm", *Comput. Methods Programs Biomed.*, 114, 88–101

¹⁸⁶ Baum K.G., et al. (2011), "Evaluation of novel genetic algorithm generated schemes for positron emission tomography (PET)/magnetic resonance imaging (MRI) image fusion", *J. Digit Imaging*, 24, 1031–1043

¹⁸⁷ Wold S., et al. (2001), "PLS-regression: a basic tool of chemometrics", *Chemom. Intell. Lab. Syst.*, 58, 109–130

¹⁸⁸ Jefferson M.F., et al. (1997), "Comparison of a genetic algorithm neural network with logistic regression for predicting outcome after surgery for patients with nonsmall cell lung carcinoma", *Cancer*, 79, 1338–1342

¹⁸⁹ Smigrodzki R., et al. (2005), "Genetic algorithm for analysis of mutations in Parkinson's disease", *Artif. Intell. Med.*, 35, 227–241

Consiglio Superiore di Sanità

Sessione LII (2019-2022)

Sezione V

Gruppo di lavoro "Intelligenza Artificiale – A.I."

Prof. Andrea Laghi

Coordinatore Gdl - Consigliere Sezione V CSS

Professore Ordinario di Radiologia, Dipartimento di Scienze medico-chirurgiche e medicina traslazionale, Università degli Studi di Roma La Sapienza. Direttore UOC Radiologia, AOU Sant'Andrea, Roma

Dr. Franco Abbenda

Segretario tecnico GdL

Dirigente medico, Coordinatore e Direttore della Struttura tecnica di Segreteria della Sezione V del Consiglio Superiore di Sanità, Ministero della salute

Prof. Giuseppe Remuzzi

Presidente Sezione V CSS

Direttore Istituto di Ricerche farmacologiche IRCS "Mario Negri", Milano - Direttore del Dipartimento di Medicina dell'Azienda Ospedaliera Papa Giovanni XXIII di Bergamo.

Prof. Mario Barbagallo

Vice presidente Sezione V CSS

Professore Ordinario di Medicina interna e Geriatria – Direttore Dipartimento Patologie emergenti e continuità assistenziale Direttore della Scuola di specializzazione in geriatria - Università degli Studi di Palermo

Prof. Giovanni Scambia

Vice presidente Sezione I CSS

Professore Ordinario di Ginecologia e Ostetricia – Direttore U.O. Ginecologia Università Cattolica del Sacro Cuore, Policlinico A. Gemelli - Direttore Dipartimento Scienze salute della donna e del bambino e Direttore Scientifico Fondazione Policlinico Universitario A. Gemelli di Roma

Prof. Renato Bernardini

Consigliere Sezione V CSS

Professore Ordinario di Farmacologia – Dipartimento Scienze biomediche e biotecnologiche (BIOMETEC) - Direttore Scuola di Specializzazione in Farmacologia e Tossicologia Clinica – Università degli Studi di Catania

Prof. Giuseppe Curigliano

Consigliere Sezione V CSS

Professore Associato di Oncologia medica - Dipartimento di Oncologia ed Emato-oncologia, Università degli Studi di Milano Direttore Struttura complessa della Divisione e sviluppo di nuovi farmaci per terapie innovative, Istituto Europeo di Oncologia (IEO)

Prof. Andrea Giustina

Consigliere Sezione V CSS

Professore Ordinario di Endocrinologia e malattie del metabolismo, Università Vita-Salute San Raffaele di Milano Direttore U.O. Endocrinologia, IRCCS Ospedale San Raffaele Milano

Prof. Riccardo Bellazzi

Professore Ordinario di Bioingegneria Elettronica e Informatica - Direttore, Dipartimento di Ingegneria Industriale e dell'Informazione, Università di Pavia - Direttore Laboratorio di Informatica e Sistemistica per la Ricerca Clinica, I.R.C.C.S. Istituti Clinici Scientifici (ICS) Maugeri, Pavia

Prof. Carlo Combi

Professore Ordinario di Informatica, Dipartimento di Informatica, Università degli Studi di Verona

Prof. Giovanni Di Rosa

Professore Ordinario di Diritto civile e Biodiritto – Dipartimento Giurisprudenza, Università degli Studi di Catania

Dott. Eugenio Santoro

Responsabile del laboratorio di Informatica Medica dell'Istituto di Ricerche Farmacologiche Mario Negri IRCCS

Prof. Vincenzo Valentini

Professore Ordinario di Radioterapia Oncologica, Università Cattolica del Sacro Cuore di Roma - Direttore Dipartimento di Diagnostica per Immagini, Radioterapia Oncologica ed Ematologia, Fondazione Policlinico Universitario A. Gemelli, Roma

IL COORDINATORE

Prof. Andrea Laghi

IL SEGRETARIO DELLA SEZIONE V

Dr. Franco Abbenda

IL PRESIDENTE DELLA SEZIONE V

Prof. Giuseppe Remuzzi