



Università degli Studi di Padova

Dipartimento di Diritto Pubblico, Internazionale e Comunitario

Corso di Laurea in Diritto e Tecnologia

A.A. 2023-2024

INTELLIGENZA ARTIFICIALE IN CAMPO MEDICO: UTILIZZI, POTENZIALITÀ E RISCHI

Docente:

Chiar.mo Prof. ANDREA PIN

Studente: AGIS QOSE

Matricola: 2006589

*Nell'elegante sinfonia dell'innovazione,
una dolce melodia vibra,
un tesoro da custodire,
tra i bagliori di albe nascenti,
che ci tiene ancora saldi,
all'essenza della nostra umanità.*

(“ChatGpt”).

INDICE

1. INTRODUZIONE.....	4
1.1 Di cosa tratterà questa tesi	4
1.2 Motivazioni della scelta	5
1.3 Obiettivi della ricerca.....	5
2. UTILIZZI DELL'IA IN AMBITO SANITARIO	6
2.1 Analisi predittiva e diagnosi	6
2.1.1 My Health Bank	7
2.1.2 IA e malattie infiammatorie	10
2.2 IA nell'imaging medico	16
2.2.1 Deep learning per la diagnosi dell'Alzheimer	17
2.3 GPT-4, diagnosi e triage	21
2.4 Ricerca e sviluppo di farmaci	24
3. RISCHI E REGOLAMENTAZIONE	29
3.1 BIAS e discriminazione nell'IA	29
3.1.1 Effetto farfalla	30
3.2 (In)sicurezza dell'uso dell'IA in medicina	32
3.2.1 Black Box e responsabilità.....	33
3.3 Privacy dei dati	35
3.4 Normative sulla protezione dei dati nel mondo	37
4. CONCLUSIONI	42
4.1 Impatto dell'IA sulla medicina tradizionale.....	42
4.2 Un pensiero personale.....	43
BIBLIOGRAFIA.....	45
RINGRAZIAMENTI	53

1. INTRODUZIONE

Molte sono le potenzialità dell'*Intelligenza Artificiale* (IA). Negli ultimi anni abbiamo visto cosa essa sia in grado di fare: dalla generazione di immagini e video di altissima qualità, basate solamente su prompt testuali, alla replica quasi perfetta della voce di persone specifiche, a software di comunicazione testuale come il famigerato CHATGPT: «[...] un chatbot che sfrutta l'*Intelligenza Artificiale Generativa* e l'apprendimento automatico, basato su un modello conversazionale, per dialogare, per iscritto, con utilizzatori finali umani»¹, e molto altro ancora.

Insomma, queste tecnologie oggi stanno letteralmente rivoluzionando il mondo come lo conosciamo e continueranno a farlo, con risvolti che probabilmente non riusciamo ancora nemmeno ad immaginare.

1.1 Di cosa tratterà questa tesi

In questa tesi verranno trattati i benefici ed i rischi legati all'utilizzo dell'IA in campo medico. Si inizierà analizzando i diversi utilizzi delle tecnologie di *Intelligenza Artificiale* che sono oggi implementati in questo campo, distinguendo tra i vari algoritmi, modelli e i casi specifici in cui l'IA è di particolare utilità. Si proseguirà analizzando quali sono i rischi ed i problemi emergenti legati all'uso di queste tecnologie e analizzando le normative vigenti a livello mondiale volte a regolamentare l'IA e la protezione dei dati personali, concludendo con una riflessione finale sull'impatto che l'IA potrebbe avere sulla medicina che oggi conosciamo.

¹ Cfr. Mirella CASTIGLI, *ChatGPT guida completa: cos'è, come si usa e cosa può fare per aziende e professionisti*, Cultura e Società Digitali, 19 aprile 2023, ultimo aggiornamento 9 febbraio 2024, disponibile nel sito AGENDA DIGITALE, <https://www.agendadigitale.eu/cultura-digitale/introduzione-a-chatgpt-cose-come-si-usa-e-cosa-puo-fare/>. (consultato il 26 febbraio 2024).

1.2 Motivazioni della scelta

La scelta di questo argomento è dovuta ad un interesse del sottoscritto riguardo l'IA e le nuove tecnologie, che ha spinto quest'ultimo a chiedersi quali potessero essere i risultati dell'implementazione di questa tecnologia nel campo medico, un settore di grande interesse per il futuro dell'umanità ma estremamente delicato allo stesso tempo.

1.3 Obiettivi della ricerca

L'obiettivo di questa ricerca è analizzare alcuni usi specifici che sono oggi implementati nel campo medico, mettendone in luce le potenzialità ed i risultati; analizzare rischi che comporta quest'implementazione, in particolare quelli legati alla sicurezza, l'utilizzo dei dati personali e l'attribuzione di responsabilità. Infine, si vuole comprendere se le normative disponibili nel periodo della scrittura di questa tesi, ovvero i primi tre mesi del 2024, siano in grado di stare al passo con l'evoluzione sempre più rapida dell'IA.

2. UTILIZZI DELL'IA IN AMBITO SANITARIO

In questo capitolo verranno analizzati alcuni specifici utilizzi dell'*Intelligenza Artificiale* (IA) nell'ambito sanitario, che stanno già trasformando il mondo della Medicina e che sembrano promettere un futuro in cui quest'ambito subirà un'evoluzione senza precedenti.

2.1 Analisi predittiva e diagnosi

Uno degli utilizzi più promettenti dell'IA nel campo medico è sicuramente quello dell'analisi predittiva, che comprende di conseguenza la diagnosi ed i trattamenti personalizzati, specifici per ogni singolo paziente.

È importante rendersi conto che ogni persona al mondo è diversa dalle altre in qualche modo. Difatti ogni individuo ha uno stile di vita unico, ha caratteristiche fisiche e biologiche specifiche ed è condizionato dai fattori ambientali che lo circondano, dall'alimentazione e da moltissimi altri fattori. Una quantità così vasta di elementi distinti è difficile da prendere in considerazione in modo preciso quando si tratta di prevedere, diagnosticare e trattare un paziente; ma è necessaria per ottimizzare al meglio la gestione della salute di ogni individuo. È qui che viene in aiuto l'*Intelligenza Artificiale*. Questa tecnologia, infatti, attraverso algoritmi di apprendimento automatico, quali *Machine Learning* (ML) e *Deep Learning* (DL) è in grado di analizzare una quantità estremamente vasta di dati, apprendere ed estrapolare informazioni utili da questi, riconoscendo modelli e pattern grazie ai quali è possibile riuscire a fare previsioni. C'è però un principio importante che va preso in considerazione quando si tratta di IA e di algoritmi di apprendimento automatico, ovvero il cosiddetto “*garbage-in, garbage-out*”, in altre parole, la qualità dei dati fornita all'algoritmo va poi a condizionare la qualità del risultato finale. È quindi importante che questi dati siano strutturati in maniera quantomeno precisa e che siano di buona qualità, per poter ottenere un risultato soddisfacente.

Di seguito andremo ad analizzare un ottimo esempio di raccolta e strutturazione di dati efficiente, che viene già utilizzato e che potrebbe essere un'architettura da cui prendere spunto, per la sua adozione a livello globale.

2.1.1 *My Health Bank*

My Health Bank è una piattaforma online in uso a Taiwan diventata ufficialmente operativa sul sito web NHIA² il 25 settembre 2014. La piattaforma consente agli utenti di verificare le proprie informazioni mediche in qualsiasi momento e di monitorare e gestire il proprio stato di salute. Serve anche come riferimento durante le visite mediche, in modo da migliorare la qualità delle cure attraverso la disponibilità dei dati sanitari del paziente. Nel luglio del 2016, *My Health Bank* è stata completamente rivisitata, migliorando la visualizzazione dei dati, integrando collegamenti esterni utili per l'educazione sanitaria ed è stato anche ampliato l'ambito di applicazione dei dati col fine di migliorarne l'accessibilità e la facilità di utilizzo.³

Secondo l'informativa dedicata all'introduzione di *My Health Bank*, in particolare nella sezione "*How will My Health Bank help me?*" presente nel sito della NHIA, ci sono quattro punti in cui viene esplicitata l'utilità di questo servizio, tre dei quali sono di interesse in questo caso:

- **Gestione dello stato di salute:** «la piattaforma fornisce dati sanitari ai pazienti in modo che essi conoscano i dettagli del loro stato di salute, come la loro storia medica e le allergie. Gli utenti possono anche programmare in anticipo cure mediche, appuntamenti [...] ed usufruire dei servizi di assistenza sanitaria preventiva in base ai dati delle loro cure mediche.»⁴
- **Miglioramento della comunicazione medico-paziente:** «i pazienti possono mostrare i dati contenuti nella piattaforma (o nell'applicazione ad essa dedicata) al loro

² *National Health Insurance Administration*, agenzia nazionale di assicurazione sanitaria taiwanese. Precedentemente nota come *Bureau of National Health Insurance*, fondata nel 1995 come impresa pubblica finanziaria e assicurativa, poi riposizionata nel 2010 come agenzia amministrativa e ribattezzata NHIA nel 2013. Cfr. "NHIA Organization", disponibile nel sito della NHIA, <https://www.nhi.gov.tw/en/cp-18-63f1c-11-2.html> (consultato il 26 febbraio 2024).

³ Cfr. *What is My Health Bank?*, *My Health Bank Introduction*, disponibile nel sito della NHIA, <https://www.nhi.gov.tw/en/cp-562-2313c-22-2.html>, (consultato il 26 febbraio 2024).

⁴ Cfr. *Ivi*. *How will My Health Bank help me?*, 1. *Management of Our Own Health*.

medico in modo che egli possa tenere traccia in modo efficace delle cure ed i trattamenti precedentemente ricevuti presso altre strutture così da prendere decisioni informate riguardo i farmaci ed i trattamenti, necessari per migliorare la sicurezza e l'efficacia delle cure mediche.»⁵

- **Miglioramento dell'acquisizione dei dati:** «*My Health Bank* combina database sanitari provenienti da diverse strutture mediche in modo che gli utenti possano usufruire di informazioni diversificate.»⁶

Come abbiamo visto, l'applicazione permette agli utenti di caricare una moltitudine di dati medici, ma non solo. Infatti, all'interno della piattaforma, oltre a quelli precedentemente citati, possono essere inserite: informazioni riguardanti la propria pressione sanguigna, il battito cardiaco, l'alimentazione, il sonno e molti altri aspetti.⁷ Questa integrazione e vasta disponibilità di dati non solo rende possibile ed accessibile l'informazione riguardante i diversi aspetti della vita e della salute del paziente, ma tali dati possono essere integrati con algoritmi di IA, che attraverso la loro analisi, utilizzando tecniche di *Deep Learning*, sono in grado di predire (con largo anticipo, si parla di dieci anni), la probabilità di insorgenza di queste particolari malattie (non si nega la possibilità che aumentino di numero in futuro): *malattia renale cronica*, *carcinoma epatocellulare* e *cancro ai polmoni*.⁸

Vediamo quindi come attraverso l'input di molti dati riguardanti un paziente non solo sia possibile avere un quadro chiaro ed ordinato della sua storia clinica, ma a prevedere e prevenire l'insorgenza di possibili malattie.

Altri dati che la piattaforma rende disponibili agli utenti, secondo la descrizione nel sito della NHIA sono i seguenti:

⁵ Cfr. *Ibidem*.

⁶ Cfr. *Ibidem*.

⁷ Cfr. TAIPEI MEDICAL UNIVERSITY, *Artificial Intelligence (AI) In Healthcare: Opportunities and Challenges*, Participatory healthcare, disponibile nel sito FUTURE LEARN,

<https://www.futurelearn.com/courses/artificial-intelligence-for-healthcare-opportunities-and-challenges/6/steps/1698187>, (consultato il 26 febbraio 2024).

⁸ *Ibidem*.

- Dati sulle cure mediche caricati tramite la tessera NHI negli ultimi tre anni, che comprendono cure di medicina occidentale, *medicina tradizionale cinese* (MTC), clinica odontoiatrica, chirurgia e farmaci.
- Dati di ospedalizzazione degli ultimi tre anni.
- Dati sulle allergie, sui test di laboratorio, sull'Imaging medico e sulle dimissioni caricati dal 1° gennaio 2015 dalle strutture mediche convenzionate con l'NHI.
- Risultati dell'assistenza sanitaria preventiva per adulti.
- Dati sull'immunizzazione.

Grazie a questo quindi vi è la possibilità di previsione e prevenzione. La piattaforma fornisce le sue previsioni basandosi sull'analisi dei dati degli ultimi tre anni presenti nella cartella del paziente. Utilizzando questi dati, la piattaforma è in grado di fornire presto una previsione sui rischi che il paziente ha di contrarre una malattia, a dieci anni dalla previsione. Non è ancora in grado però di fornire previsioni riguardo periodi più prossimi e questo può rendere difficile intervenire nel breve periodo.⁹ Per questo, i dati conferiti dalla piattaforma, presentati in questo modo, possono essere estremamente utili al fine di fornire dati chiari e strutturati, da essere poi utilizzati in studi specifici che hanno lo scopo di prevenire l'insorgenza di patologie in un tempo minore e che grazie alla disponibilità di questi dati possono fornire un ottimo input agli algoritmi di apprendimento automatico.

Ciò che si può estrapolare da quest'analisi è l'importanza della partecipazione degli utenti nel fornire e caricare i propri dati sanitari all'interno di questo tipo di applicazioni, che altrimenti vedrebbero venir meno il loro scopo, se in possesso di poche informazioni. In altre parole, più dati ci sono, migliori sono i risultati che si possono ottenere.

Questa piattaforma, quindi, è un ottimo modello atto a raccogliere e fornire dati sulla storia clinica di un paziente, in modo chiaro e dettagliato, e che, come detto in precedenza, è un approccio che dovrebbe essere integrato anche all'interno dei sistemi sanitari delle altre nazioni, variando poi a seconda delle esigenze, ma che permette un'analisi precisa e dei risultati ottimali.

⁹ *Ibidem.*

Visto questo caso specifico, si procederà con l'analisi di diversi studi che dimostrano come sia possibile prevedere e diagnosticare un particolare tipo di malattie, ovvero quelle infiammatorie, attraverso l'utilizzo dell'IA.

2.1.2 IA e malattie infiammatorie

Le malattie infiammatorie sono un gruppo di patologie, accomunate da uno stato infiammatorio, in questo caso cronico, dovute ad una reazione immunitaria anomala che attacca le cellule sane invece degli agenti patogeni. Queste malattie si differenziano tra di loro, e possono avere effetti diversi a seconda del tipo di patologia, dalla risposta immunitaria, e da cosa, o dove, il sistema immunitario dedica di attaccare.¹⁰ La causa delle malattie infiammatorie è tutt'ora poco nota, non si sa ancora con certezza, infatti, cosa scateni una reazione autoimmune o perché questo avvenga.

Esistono molti tipi di malattie infiammatorie o autoimmuni. Alcune di queste sono: la *malattia di Addison*, la *celiachia*, il *morbo di Crohn*, la *sclerosi multipla*, l'*artrite reumatoide*, il *lupus eritematoso sistemico*, il *diabete* di tipo-1, ecc..¹¹

Non è semplice diagnosticare una malattia di questo tipo, infatti prima di arrivare ad una diagnosi certa sono necessari diversi esami specifici che variano a seconda della patologia. Inoltre, per questo tipo di malattie non esiste ancora una cura definitiva, ovvero un farmaco o un trattamento che possa guarire del tutto il paziente affetto, ma negli ultimi anni, in particolare, si sono fatti dei grandi passi avanti nella ricerca dei trattamenti che riescano a tenerne sotto controllo l'attività e l'avanzamento. Le terapie utilizzate per queste malattie consistono in genere nel modulare (o sopprimere, a seconda dei casi) l'attività del sistema immunitario, in modo tale che questo possa diminuire la sua attività e di conseguenza anche il meccanismo autoimmune. Grazie all'IA, oggi è possibile ottenere diagnosi più accurate, veloci e riuscire in una certa misura a prevederne l'andamento nel tempo.

¹⁰ Cfr. “Malattie infiammatorie”, disponibile sul sito AMGEN, <https://www.amgen.it/en/aree-terapeutiche/malattie-infiammatorie>, (consultato il 27 febbraio 2024).

¹¹ Cfr. Che cosa sono le malattie autoimmuni?, “Malattie autoimmuni”, disponibile sul sito della HUMANITAS RESEARCH HOSPITAL, <https://www.humanitas.it/malattie/malattie-autoimmuni/>, (consultato il 27/02/2024).

Di seguito vedremo come l'IA può aiutare ad ottenere una diagnosi, in particolare analizzeremo due studi: uno incentrato sul ruolo dell'IA nel diagnosticare la *sclerosi multipla*¹², ed uno nel prevederne l'evoluzione.

Lo studio “*Artificial intelligence in the diagnosis of multiple sclerosis: A systematic review*” si impegna ad analizzare ed eseguire una revisione sistematica per documentare le prestazioni dell'IA nella diagnosi di SM.¹³ In questo studio è stata eseguita una ricerca sistematica utilizzando quattro database, tra cui PUBMED, SCOPUS, WEB OF SCIENCE e *IEEE*. Tutti gli studi originali, incentrati sull'utilizzo di algoritmi di *deep learning* o sull'IA, allo scopo di diagnosticare la SM (*sclerosi multipla*), sono stati inclusi nello studio.¹⁴ La revisione analizza come siano stati sviluppati diversi metodi basati sull'uso dell'IA per il rilevamento delle lesioni da SM nei dati delle *risonanze magnetiche* (MRI), della *tomografia ottica computerizzata* (OCT), delle misure sierologiche e di quelli motori.¹⁵

Analizziamole di seguito:

- **Imaging:** l'utilizzo della *risonanza magnetica* nella diagnosi ed il monitoraggio dei pazienti affetti da SM è un approccio che viene preso in considerazione da molto tempo. I recenti modelli di Imaging utilizzano le cosiddette CNN (*Convolutional Neural Network*) per analizzare le immagini ottenute grazie alla *risonanza magnetica* e riuscire a fornire un quadro preciso della situazione. La

¹² La *sclerosi multipla* è una malattia infiammatoria cronica di natura autoimmune che colpisce il sistema nervoso centrale. Il suo meccanismo di azione consiste nell'attivazione del sistema immunitario che attacca erroneamente la mielina, ovvero la membrana che ricopre i nervi, e può causare una moltitudine di problemi nei soggetti affetti da questa patologia. La prognosi varia da paziente a paziente. Non esiste una cura definitiva per questa malattia ed è di natura neurodegenerativa, la cui progressione è variabile a seconda del carico lesionale e diversa per ogni paziente.

¹³ Cfr. Fardin NABIZADEH *et al.* *Artificial intelligence in the diagnosis of multiple sclerosis: A systematic review*, *Multiple Sclerosis and Related Disorders*, vol. 59, 2022, 03673, ISSN 2211-0348, <https://doi.org/10.1016/j.msard.2022.10367>, <https://www.sciencedirect.com/science/article/pii/S2211034822001882>

¹⁴ *Ivi.* Metodi.

¹⁵ *Ivi.* Introduzione, par. 3

miglior performance è stata ottenuta da SOLTANI et al.¹⁶ con una precisione diagnostica del 99,66%.¹⁷

- **Tomografia ottica computerizzata (OCT):** questa tecnica viene generalmente utilizzata per rilevare la presenza della *neurite ottica*, ovvero un'inflammazione del nervo ottico, nei pazienti affetti da SM. Secondo la revisione in oggetto, gli studi recenti hanno dimostrato come utilizzando l'OCT sia possibile ottenere una diagnosi precoce di SM anche in pazienti senza lesioni *oftalmologiche* poiché l'assottigliamento della mielina nella RNFL (*strato delle fibre nervose della Retina*) può verificarsi in molte condizioni di degenerazione della mielina come la SM (Gür GÜNGÖR e AHMET¹⁸). In altre parole, anche se un paziente non presenta lesioni visibili al nervo ottico, il suo assottigliamento può indicare che ci si trovi in presenza di *sclerosi multipla*. Gli indici OCT sono già utilizzati nella pratica clinica, tuttavia, non è ancora stato raggiunto un consenso sui pattern delle anomalie riscontrate nella retina (BRITZE e FREDERIKSEN¹⁹). Pertanto, è stata introdotta l'applicazione di modelli di ML per la classificazione delle immagini della RNFL. In questo caso, l'accuratezza maggiore si è ottenuta attraverso l'utilizzo un algoritmo di *decision tree* con un'accuratezza diagnostica del 97,4% utilizzando immagini della RNFL e dello strato delle cellule ganglionari (Pérez DEL PALOMAR et al²⁰).²¹

¹⁶ Azam SOLTANI et al. *Improved algorithm for multiple sclerosis diagnosis in MRI using convolutional neural network*, IET Image Processing, vol. 14, 17, 2020, 1751-9659, <https://doi.org/10.1049/iet-ipr.2019.0366>.

¹⁷ Fardin NABIZADEH et al. cit. Imaging, par. 2

¹⁸ Gür GÜNGÖR S, Ahmet AKMAN, *Are All Retinal Nerve Fiber Layer Defects on Optic Coherence Tomography Glaucomatous?* Turk J Ophthalmol, 2017; 47: pp. 267-273, <https://pubmed.ncbi.nlm.nih.gov/29109895/>.

¹⁹ Josefine BRITZE, Jette Lautrup FREDERIKSEN, *Optical coherence tomography in multiple sclerosis*, Eye, Maggio 2018, vol. 32, pp. 884–888, <https://doi.org/10.1038/s41433-017-0010-2>.

²⁰ Pérez DEL PALOMAR et al. *Swept source optical coherence tomography to early detect multiple sclerosis disease. The use of machine learning techniques*, PLOS ONE, 2019, 14(5), <https://doi.org/10.1371/journal.pone.0216410>.

²¹ Fardin NABIZADEH et al. cit. Optical coherence tomography (OCT).

- **Analisi dei marcatori del siero e del *Liquor* (liquido cerebrospinale):**
L'analisi dei profili di marcatori contenuti nel siero e nel liquido cefalorachidiano (CSF) detto anche *Liquor*, è un approccio già utilizzato per diagnosticare la SM. In uno studio condotto da LÖTSCH et al.²², è stato misurato approfonditamente il profilo lipidico nel siero, integrando la misura con l'applicazione di un modello di ML supervisionato, che ha portato l'accuratezza dell'analisi al 96%.²³
- **Dati motori:** I disturbi dell'equilibrio e del movimento sono comuni nei pazienti affetti da *sclerosi multipla*; alcuni studi hanno utilizzato algoritmi di ML sui dati della deambulazione e delle variazioni posturali dei pazienti. I risultati degli studi, però, non sono solidi, a causa della ridotta dimensione del campione (SARBAZ et al.²⁴). Oltre a questo, in diversi studi su pazienti affetti, l'IA è stata utilizzata per rilevare in loro i livelli di *ataxia* specialmente in quelli con un basso grado di disabilità tramite algoritmi *support-vector machine* (SVM) e *k-*

²² Jörn LÖTSCH et al. Machine-learning based lipid mediator serum concentration patterns allow identification of multiple sclerosis patients with high accuracy, Scientific Reports, 5 Ottobre 2018, <https://doi.org/10.1038/s41598-018-33077-8>.

²³ Fardin NABIZADEH et al. cit. Serum and CSF.

²⁴ Yashar SARBAZ et al., Introducing a decision support system for multiple sclerosis based on postural tremor: a hope for separation of people who might be affected by multiple sclerosis in the future, Biomedical Engineering: Applications, Basis and Communications, vol. 29(06), 1750046 (2017), <https://doi.org/10.4015/S1016237217500466>.

nearest neighbors (K-NN) con un'accuratezza dell'89% (KAYA et al.²⁵, BILEK et al.²⁶, BALGETIR et al.²⁷).²⁸

In sintesi, questo studio ha dimostrato come i metodi di IA abbiano una performance diagnostica affidabile nel distinguere i pazienti con SM dalle persone sane e come ci sia un'enorme potenzialità per l'uso dell'IA nella diagnosi di SM e altre malattie infiammatorie.²⁹

Un altro studio, intitolato “*Machine Learning Approaches in Study of Multiple Sclerosis Disease Through Magnetic Resonance Images*”³⁰, ci aiuterà a capire come, attraverso l'uso dell'IA sia possibile riuscire a prevedere la prognosi l'evoluzione della disabilità nei pazienti affetti da SM.

In questo studio, in particolare nella sezione dedicata alla previsione della progressione della malattia, vediamo come la revisione condotta analizza tre studi che hanno utilizzato tecniche di ML per prevedere la progressione della SM, misurando la *Scala di Invalidità Espansa* (EDSS) nei primi anni di evoluzione della malattia. Due di questi hanno utilizzato le CNN per estrarre e classificare le caratteristiche derivanti

²⁵ Mustafa KAYA et al. Detection of ataxia with hybrid convolutional neural network using static plantar pressure distribution model in patients with multiple sclerosis, *Computer Methods and Programs in Biomedicine*, vol. 214, 2022, <https://doi.org/10.1016/j.cmpb.2021.106525>.

²⁶ Furkan BILEK et al. Quantitative Assessment of Ataxia in Multiple Sclerosis Patients using Spatiotemporal Parameters: A Relief-Based Machine Learning Analysis, *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin* 2021; 31(06): pp. 367-376, <https://doi.org/10.1055/a-1512-4858>.

²⁷ Ferhat BALGETIR et al. Detection of ataxia in low disability MS patients by hybrid convolutional neural networks based on images of plantar pressure distribution, *Multiple Sclerosis and Related Disorders*, 2021, vol. 56, 103261, <https://doi.org/10.1016/j.msard.2021.103261>.

²⁸ Fardin NABIZADEH *et al.* cit. Movement.

²⁹ *Ibidem.*

³⁰ Cfr. Faezeh MOAZAMI et al., *Machine Learning Approaches in Study of Multiple Sclerosis, Disease Through Magnetic Resonance Images*, *Frontiers in Immunology, ML Approaches in Study of MS*, Agosto 2021, vol.12, <https://doi.org/10.3389/fimmu.2021.700582> (consultato il 29 febbraio 2024).

dalle immagini di MRI, e il terzo ha invece utilizzato l'algoritmo SVM per classificare le immagini delle risonanze basandosi sulle caratteristiche estratte dalle immagini.³¹

Uno degli studi approfonditi in questa revisione ha valutato l'influenza del volume delle lesioni sulla performance di rilevamento delle CNN. CORONADO et al.³² hanno applicato questo modello su cinque immagini MRI multispettrali con diversi volumi di potenziamento del *gadolinio* (metodo di contrasto). Le prestazioni migliori sono state ottenute utilizzando come input tutti e cinque i set di immagini multispettrali, inclusi FLAIR, T2, PD, e T1 pre e post-*gadolinio*.³³

Youngjin YOO et al.³⁴ hanno migliorato le prestazioni del loro modello CNN nella previsione della progressione delle lesioni utilizzando il metodo *Transfer Learning* (TL).³⁵ Inizialmente, hanno applicato un algoritmo definito come *unsupervised 3D convolutional deep belief network* (DBN) come modello pre-addestrato, e successivamente hanno usato una CNN per estrarre i pattern di lesioni latenti causati dalla SM.³⁶ Hanno anche confrontato il loro modello con le prestazioni dei *Random Forests*. Il loro modello, CNN, ha ottenuto una precisione media del 75% nel prevedere la conversione clinica a *SM definitiva* entro due anni, mentre i *Random Forests* hanno ottenuto una precisione del 67,9%.³⁷ Lo studio si conclude mettendo in luce come l'estrazione delle caratteristiche dalle immagini di *risonanza magnetica* da parte dei tecnici non permetta di trovare nuove informazioni nascoste nei dati.³⁸ I modelli di DL quindi, potrebbero aiutare a superare il problema estraendo informazioni utili

³¹ *Ivi*. Prediction of MS Disease Progression.

³² Ivan CORONADO et al. *Deep learning segmentation of gadolinium-enhancing lesions in multiple sclerosis*, *Multiple Sclerosis Journal*, 2021, vol. 27(4), pp. 519-527, <https://doi.org/10.1177/1352458520921364>.

³³ Faezeh MOAZAMI et al. cit. *Ibidem*.

³⁴ Youngjin YOO et al. *Deep learning of brain lesion patterns and user-defined clinical and MRI features for predicting conversion to multiple sclerosis from clinically isolated syndrome*, *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2019, vol.7(9), pp. 250-259, <https://doi.org/10.1080/21681163.2017.1356750>.

³⁵ Faezeh MOAZAMI et al. cit. *Ibidem*.

³⁶ *Ibidem*.

³⁷ *Ibidem*.

³⁸ *Ivi*. Conclusion.

direttamente dai dati provenienti dalle immagini.³⁹ Inoltre, l'utilizzo dei metodi di *transfer learning* (TL) durante il processo di formazione dei modelli di DL potrebbe migliorare le loro prestazioni. Questo approccio ha mostrato prestazioni notevoli, specialmente nella ricerca medica, dove di solito sono disponibili limitati dataset di immagini e dove l'analisi delle immagini da parte di un esperto è un compito particolarmente impegnativo.⁴⁰

Come possiamo vedere da questo studio, è possibile riuscire a prevedere l'evoluzione e la progressione della SM con una percentuale di precisione discreta, che è destinata a migliorare più si progredisce con la ricerca e lo sviluppo delle tecnologie di IA. Ciò che importa però, è che questo sia comunque un ottimo risultato che permette così di poter agire in modo conscio e adattare le terapie in base alle informazioni ottenute, riuscendo a trattare questa patologia nel modo migliore possibile (fino ad ora).

Di seguito andremo ad approfondire l'uso dell'IA nell'imaging medico e come l'uso di algoritmi di DL siano utili nella diagnosi della malattia di Alzheimer.

2.2 IA nell'imaging medico

Come visto in precedenza, un particolare tipo di architettura di *deep learning*, il CNN, è un ottimo strumento da utilizzare nell'analisi delle immagini mediche ed è in grado di rilevare pattern, dettagli ed informazioni che possono sfuggire ad un tecnico umano. Questo non solo migliora la qualità e la precisione dei risultati ottenuti, ma velocizza di gran lunga il processo di analisi e permette di utilizzare un numero nettamente maggiore di immagini nel processo di ricerca.

Nel giugno del 2018, uno studio pubblicato su ANNALS OF ONCOLOGY ha dimostrato che una CNN, addestrata ad analizzare immagini dermatologiche, è stata in grado di individuare la presenza di un melanoma con una precisione del 10% superiore rispetto ai clinici umani. Persino quando ai tecnici venivano fornite

³⁹ *Ibidem.*

⁴⁰ *Ibidem.*

informazioni di base sui pazienti, come età, sesso e posizione del possibile melanoma, la CNN è riuscita comunque ad avere una precisione maggiore del 7%.⁴¹

Vediamo ora come il deep learning viene utilizzato per la diagnosi dell'*Alzheimer*.

2.2.1 *Deep learning per la diagnosi dell'Alzheimer*

La malattia di *Alzheimer*, conosciuta come *Alzheimer's Disease* (AD), è la causa più comune di demenza. È una malattia progressiva che esordisce con un'iniziale lieve perdita di memoria e può progredire, portando all'incapacità di sostenere una conversazione e di interagire con l'ambiente circostante. La malattia coinvolge le parti del cervello che controllano il pensiero, la memoria ed il linguaggio. Tutto questo può influenzare fortemente la capacità di una persona di svolgere attività quotidiane.⁴² L'AD è una malattia irreversibile per la quale ad oggi non esiste una cura definitiva ed i farmaci esistenti possono solo rallentarne il processo di deterioramento. La diagnosi precoce dell'AD⁴³, è quindi di fondamentale importanza per ridurre il deterioramento.

Secondo lo studio di revisione "*Current Trends in Deep Learning Techniques for Alzheimer's Disease. Diagnosis Using Medical Images: A Systematic Literature Review*"⁴⁴, generalmente l'imaging medico viene analizzato attraverso la visione manuale ed i metodi tradizionali di ML, che sono dispendiosi in termini di tempo, soggetti ad errori umani e poco accurati. Per questo motivo gli approcci di DL hanno guadagnato l'attenzione per la loro capacità di gestire grandi dataset di immagini mediche e fornire supporto per la diagnosi di Alzheimer.⁴⁵ In questa revisione sono

⁴¹ Cfr. Jennifer BRESNICK, *What Is Deep Learning and How Will It Change Healthcare?*, Features, disponibile nel sito HEALTH IT ANALYTICS, <https://healthitanalytics.com/features/what-is-deep-learning-and-how-will-it-change-healthcare>, (consultato il 1° marzo 2024).

⁴² Cfr. *Alzheimer's Disease and Related Dementias*, What is Alzheimer's Disease?, Alzheimer's Disease and Healthy Aging, disponibile sul sito del CENTER OF DISEASE CONTROL AND PREVENTION, <https://www.cdc.gov/aging/aginginfo/alzheimers.htm>, (consultato il 1° marzo 2024).

⁴³ *Alzheimer's Disease*, ovvero il termine inglese per definire la malattia di *Alzheimer*.

⁴⁴ Wong P. CHING, Shahrum S. ABDULLAH, Mohd I. SHAPIAI, *Current Trends in Deep Learning Techniques for Alzheimer's Disease. Diagnosis Using Medical Images: A Systematic Literature Review*, 13 febbraio 2024, Stato di pubblicazione attuale: Under Review, HELIYON, <https://ssrn.com/abstract=4720066>, (consultato il 1° marzo 2024).

⁴⁵ Cfr. *Ivi.*, Abstract.

state analizzate un totale di 110 pubblicazioni sulla diagnosi di AD che utilizzano vari algoritmi di apprendimento, database e modalità.⁴⁶ Ciò che emerge dallo studio è che l'architettura CNN, come parte del modello di DL, è divenuta la strategia preferita per la diagnosi di AD in quanto fornisce risultati accurati e nettamente più veloci rispetto ai metodi che utilizzano il ML, i quali necessitano di una fase di pre-addestramento e che vengono spesso integrati con una moltitudine di altri modelli ed algoritmi per poter ottenere risultati ottimali, rendendo il processo dispendioso in termini di tempo e risorse, a differenza degli algoritmi di DL che riescono ad estrarre le caratteristiche dalle immagini in maniera autonoma. Nel 2020 infatti, BRINGAS et al.⁴⁷ hanno utilizzato un modello CNN su dati raccolti da un centro diurno per classificare i diversi stadi dell'AD come: iniziale, intermedio o avanzato. Il modello utilizzato ha raggiunto un'accuratezza del 90.91%⁴⁸. Grazie a questo ed altri studi simili, si è dimostrato che i modelli CNN siano capaci di ottenere ottimi risultati nella classificazione binaria (capire se ci si trova di fronte ad un caso di malattia o meno) o, come in questo caso, multi-classe (ovvero classificare lo stato di avanzamento della malattia).⁴⁹

Nonostante ciò, lo studio evidenzia come ci siano delle limitazioni anche utilizzando questo metodo. Il problema principale identificato dalla letteratura è l'insufficienza dei dati medici e delle immagini di MRI durante l'addestramento delle reti DL.⁵⁰ Di conseguenza sono state sviluppate particolari tecniche di DL avanzate per far fronte a questi problemi, come l'integrazione di tecniche di *transfer learning* (TL) che permettono di migliorare l'addestramento dell'algoritmo di DL rendendolo più veloce ed accurato e la fusione di diversi modelli, definito come apprendimento ibrido.⁵¹ Il TL fornisce quindi un pre-addestramento al modello di DL su grandi dataset, come IMAGENET ed altri, che contengono un enorme numero di immagini. Successivamente, solo i livelli di classificazione del modello di DL vengono sottoposti

⁴⁶ Cfr. *Ivi.*, Conclusion.

⁴⁷ BRINGAS et al. 2020, *Alzheimer's Disease stage identification using deep learning models*, Journal of Biomedical Informatics, 109, 103514.

⁴⁸ Wong P. CHING, *Cit.*, Deep Learning (DL).

⁴⁹ *Ibidem.*

⁵⁰ *Ivi.* Conclusion.

⁵¹ *Ibidem.*

ad un processo di *fine-tuning*, velocizzando di gran lunga l'intero processo.⁵² Questa tecnica è ora una delle strategie di apprendimento più utilizzate per questo tipo di compiti. Diversi framework di DL, CNN inclusi, sono stati utilizzati per creare reti pre-addestrate, come GOOGLNET, DENSENET, EFFICIENTNET, RESNET, ALEXNET e SQUEEZENET, [...], che gli utenti possono sfruttare per le loro ricerche necessitando solo di un leggero aggiustamento finale (*fine-tuning*). Queste reti pre-addestrate sono state a loro volta addestrate utilizzando il dataset IMAGENET, che contiene mille classi diverse di immagini.⁵³ L'ultimo studio di ASHRAF et al.⁵⁴ del 2021 ha applicato questo metodo utilizzando un totale di 13 architetture CNN pre-addestrate, ovvero ALEXNET, GOOGLNET, VGG-16, VGG-19, RESNET-18, RESNET-50, RESNET-101, DENSENET, MOBILENET, SQUEEZENET, INCEPTION V3, INCEPTION-RESNET-v2 e *spiking neural networks* (SNN), sottoponendole a *fine-tuning* per classificare le immagini di MRI.⁵⁵ Il DENSENET ha ottenuto le prestazioni migliori, con un'accuratezza diagnostica media del 99,05%. Questo dimostra perfettamente i significativi miglioramenti nei modelli basati su TL anche per complesse attività di classificazione multi-classe.

Per quanto riguarda l'apprendimento ibrido, invece, la revisione mette in luce diversi studi che ne dimostrano le potenzialità. Uno degli studi di maggior successo in quanto a precisione è stato condotto da ODUSAMI et al.⁵⁶ ed ha combinato due modelli di DL, specificamente l'architettura RESNET-18 ed il modello DENSENET-201, formando così un modello ibrido per la classificazione multi-classe dalle immagini di MRI. Il risultato ha dimostrato un'accuratezza e una precisione del 98%. In generale si è visto che i modelli ibridi hanno a loro volta migliorato l'accuratezza e le prestazioni della classificazione.⁵⁷

⁵² *Ivi.* Transfer Learning (TL).

⁵³ *Ibidem.*

⁵⁴ ASHRAF et al., 2021, Deep transfer learning for Alzheimer neurological disorder detection, Multimedia Tools and Applications.

⁵⁵ Wong P. CHING, Cit., Transfer Learning (TL).

⁵⁶ Odusami et al., 2022, An Intelligent System for Early Recognition of Alzheimer's Disease Using Neuroimaging, Sensors.

⁵⁷ Wong P. CHING, Cit., Hybrid learning.

Un altro tipo di apprendimento ibrido che emerge da questa revisione è l'utilizzo dell'architettura GAN (*Generative Adversarial Networks*), che coinvolge due sottomodelli: un modello generatore, per la creazione di nuovi esempi ed un modello discriminatore, per classificare gli esempi generati come reali e quindi appartenenti al dominio, o falsi, cioè generati dal modello generatore. Quest'architettura viene utilizzata per generare o produrre nuovi esempi che potrebbero plausibilmente essere stati estratti dal dataset originale. In altre parole, generare risultati talmente simili ai dati forniti inizialmente, da essere virtualmente indistinguibili da essi, sembrando reali.⁵⁸ L'architettura viene usata in questo campo per far fronte all'insufficienza di immagini disponibili. Lo studio di maggior successo in quanto a precisione diagnostica è stato condotto da KADRI et al.⁵⁹, che ha proposto un metodo ibrido di GAN con una rete DENSENET ed una rete definita *Squeeze-and-Excitation Network* (SENet) per arricchire i dati e la classificazione binaria della malattia di *Alzheimer*. Lo studio ha adottato il GAN *Convolutionale* (CGAN) per l'arricchimento dei dati ed ha raggiunto un'accuratezza del 98%, rispetto ai metodi tradizionali.⁶⁰ È stata quindi dimostrata l'importanza dell'architettura GAN quando incorporata con altri modelli, portando ad un uso crescente di questi metodi e ad un'accuratezza diagnostica estremamente alta ed efficiente.

Grazie a questi studi abbiamo potuto vedere come utilizzando diverse architetture di DL sia possibile riuscire a velocizzare e migliorare la precisione nella diagnosi dell'AD, utilizzando le immagini di *risonanza magnetica*.

Di seguito vedremo come vengono utilizzati i modelli di linguaggio, conosciuti anche come *Large Language Models* (LLM) per fornire diagnosi.

⁵⁸ Cfr. Jason BROWNLEE, *A Gentle Introduction to Generative Adversarial Networks (GANs)*, 19 luglio 2019, *Generative Adversarial Networks*, disponibile nel sito MACHINE LEARNING MASTERY, <https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/>, (consultato il 1° marzo 2024).

⁵⁹ KADRI et al., *Deep Squeeze and Excitation-Densely Connected Convolutional Network with cGAN for Alzheimer's Disease Early Detection*, 2022, *Intelligent Systems Design and Applications*, vol 418, Springer, Cham, https://doi.org/10.1007/978-3-030-96308-8_41.

⁶⁰ Wong P. CHING, *Cit., Hybrid learning*.

2.3 GPT-4, diagnosi e triage

I LLM sono un tipo di algoritmo di AI che utilizza tecniche DL e dataset estremamente grandi per comprendere, riassumere, generare e prevedere nuovo contenuto.⁶¹

Un LLM molto famoso (e probabilmente il più conosciuto) è CHATGPT, il modello di linguaggio di OPENAI. Come abbiamo visto in precedenza, CHATGPT è un chatbot di IA, in grado di dialogare per iscritto, con utilizzatori finali umani.⁶² Ma questo modello di IA è in grado di svolgere molte altre funzioni oltre al semplice dialogare in forma testuale. L'uso specifico del modello linguistico di OPENAI che andremo ad analizzare in questo caso è quello della sua capacità di fornire diagnosi e valutare il livello di *triage*, nel contesto medico.

Uno studio intitolato “*The Accuracy and Potential Racial and Ethnic Biases of GPT-4 in the Diagnosis and Triage of Health Conditions: Evaluation Study*”⁶³ si è occupato di valutare il livello di precisione dell'LLM GPT-4 nell'effettuare una diagnosi e di valutare il *triage* basandosi su 45 casi clinici già valutati in precedenza come corretti e confrontando le sue prestazioni con quella di tre clinici umani. Per valutare il livello di precisione diagnostica lo studio ha seguito un particolare criterio; per ogni caso clinico utilizzato come esempio da valutare, è stato chiesto sia a GPT-4 che ai medici di fornire la diagnosi primaria più probabile e tre diagnosi differenziali. Ai partecipanti umani non è stato consentito conoscere le decisioni degli altri durante il processo di valutazione. Inoltre, sia a GPT-4 che ai medici è stato chiesto di spiegare il loro ragionamento e le motivare adeguatamente le diagnosi. Per la valutazione, sono state sottoposte consecutivamente quattro domande in modo standardizzato, seguite dai criteri per la classificazione del triage, ecco il prompt:

⁶¹ Cfr. Sean Michael KERNER, *large language models (LLMs)*, Settembre 2023, Definition, What is generative AI? Everything you need to know, disponibile sul sito TECH TARGET, <https://www.techtarget.com/whatis/definition/large-language-model-LLM>, (consultato il 3 marzo 2024).

⁶² Cfr. Mirella CASTIGLI, Cit.

⁶³ Naoki ITO et al. *The Accuracy and Potential Racial and Ethnic Biases of GPT-4 in the Diagnosis and Triage of Health Conditions: Evaluation Study*, 2 novembre 2023, JMIR Med Educ, vol. 9, 47532, <https://mededu.jmir.org/2023/1/e47532/>, (consultato il 3 marzo 2024).

«Si prega di rispondere alle seguenti domande riguardanti il caso clinico presentato;

- diagnosi più probabile
- motivazione della diagnosi
- altre tre diagnosi
- classificazione del triage

Il caso clinico è il seguente;

[Caso clinico]

I criteri per la classificazione del triage sono i seguenti:

emergente: che include il consiglio di chiamare un'ambulanza, recarsi al reparto di emergenza o consultare immediatamente un medico; non emergente, che include il consiglio di chiamare un medico generico o un fornitore di cure primarie o di recarsi presso un centro di assistenza immediata, di consultare uno specialista, andare in una “retail clinic” o richiedere una visita elettronica; e auto-cura, che include il consiglio di rimanere a casa o di recarsi in farmacia.»⁶⁴. Queste sono poi state valutate in modo indipendente da altri due medici che hanno classificato la diagnosi primaria più probabile come “corretta” o “errata” e la motivazione come “appropriata” o “inappropriata”. Nel caso in cui i giudizi dei due revisori fossero stati contrastanti, la decisione sarebbe stata presa da un terzo medico.⁶⁵

Le diagnosi venivano considerate corrette se corrispondenti a quelle attese o se identificate come le più probabili in base al caso clinico. Venivano invece considerate errate se diverse dalla risposta attesa o se, nonostante fossero considerate corrette, mancasse la menzione di una condizione critica, per esempio, sapendo che la *sindrome emolitico-uremica* (HUS) è causata da *gastroenterite acuta*, nel caso in cui la diagnosi fosse stata solo di quest'ultima, sarebbe stata considerata errata poiché l'omissione della presenza di HUS può risultare fatale per i pazienti.⁶⁶

Le motivazioni delle diagnosi venivano considerate "appropriate" se coerenti con le diagnosi stesse e se accompagnate da una spiegazione convincente, anche nel caso in cui la diagnosi primaria fosse errata. Una motivazione risultava invece

⁶⁴ *Ivi*. Assessment of GPT-4's Performance.

⁶⁵ *Ivi*. Evaluation of the Diagnosis.

⁶⁶ *Ibidem*.

"inappropriata" se incoerente con la diagnosi, se la spiegazione fosse stata insoddisfacente, o se non in grado di differenziarla da altre diagnosi potenziali.⁶⁷

Per quanto riguarda lo studio della precisione della valutazione del *triage*, ai partecipanti è stato chiesto valutare il livello di quest'ultimo per ogni caso clinico fornito, le valutazioni possibili erano quelle utilizzate nel prompt precedentemente citato. Un livello di *triage* veniva definito "corretto" se corrispondente esattamente al livello di *triage* atteso; in caso contrario, veniva considerato "errato".⁶⁸

I risultati dello studio hanno messo in mostra la capacità di GPT-4 di fornire diagnosi accurate e valutare il livello di *triage* comparabile alla precisione dei medici.

L'accuratezza della diagnosi ha ottenuto una percentuale di diagnosi corretta del 97,8% per GPT-4 e del 91,1% per i medici. Inoltre, GPT-4 ha fornito un ragionamento appropriato per il 97,8% dei casi. Anche l'adeguatezza del *triage* è risultata paragonabile tra GPT-4 ed i medici: entrambi hanno totalizzato una percentuale di precisione del 66,7%.⁶⁹ La performance di GPT-4 non variava nemmeno nel caso in cui fossero valutati casi clinici di diverse etnie (neri, bianchi, asiatici e ispanici).⁷⁰

Lo studio conclude confermando la capacità di GPT-4 di diagnosticare e gestire il *triage* di situazioni cliniche tipiche paragonabile a quella dei medici certificati e sostenendo che questi risultati dovrebbero essere informativi per i sistemi sanitari che intendono adottare l'uso dei LLM per migliorare l'efficienza diagnostica e della valutazione del *triage* dei pazienti.⁷¹

Abbiamo visto come i sistemi di IA, capaci di comunicare e comprendere le informazioni scritte, siano capaci di uguagliare, se non sorpassare, le performance dei clinici umani in determinate situazioni, nonostante sia importante ricordarsi che questi studi siano stati condotti sotto criteri specifici e situazioni controllate; quindi, la loro applicazione nel mondo reale potrebbe comunque variare. Nonostante tutto rimane comunque un risultato degno di nota e decisamente promettente.

⁶⁷ *Ibidem.*

⁶⁸ *Ibidem.*

⁶⁹ *Ivi.* Results.

⁷⁰ *Ibidem.*

⁷¹ *Ivi.* Conclusion.

A margine di questa analisi, il sottoscritto ha voluto testare di persona l'efficacia diagnostica di GPT-4, fornendo al chatbot informazioni su due casi clinici che hanno coinvolto direttamente lo scrittore in passato. La conversazione si è svolta chiedendo al chatbot di presentare 20 domande utili per effettuare una diagnosi, alle quali ha ricevuto risposte basate sui casi specifici. Il risultato della conversazione ha rispettato perfettamente le aspettative; GPT-4 ha infatti diagnosticato correttamente entrambi i casi clinici, fornendo la diagnosi primaria più probabile, la sua causa e motivando la sua risposta.

Di seguito vedremo invece come l'IA può essere impiegata per migliorare la ricerca e lo sviluppo dei farmaci.

2.4 Ricerca e sviluppo di farmaci

Un altro campo di utilizzo dell'IA è quello della ricerca e dello sviluppo farmaceutico. Viste le potenzialità che questa tecnologia possiede, la sua applicabilità in questo settore è sicuramente degna di studio e approfondimento.

Per capire come venga utilizzata l'IA nella ricerca farmaceutica andremo ad analizzare lo studio "*Artificial intelligence in drug discovery and development*"⁷² che fornisce un quadro dettagliato su questo argomento.

Lo studio inizia specificando come esistano un numero enorme di molecole, nello specifico, più di "10⁶⁰". Questo ovviamente rende possibile lo sviluppo di tantissimi potenziali farmaci ma al contempo risulta problematico quando ci si trova ad aver a che fare con la fase effettiva di sviluppo.⁷³ Questo processo è infatti lungo e costoso. Il problema però può essere affrontato utilizzando l'IA, che può riconoscere i diversi composti e fornire una valutazione più rapida del target del farmaco e dell'ottimizzazione della progettazione della sua struttura, oltre a fornire assistenza al

⁷² Paul DEBLEENA *et al.* *Artificial intelligence in drug discovery and development*, *Drug discovery today*, gennaio 2021, vol. 26; 1, pp. 80-93, disponibile sul sito SCIENCE DIRECT, <https://doi.org/10.1016/j.drudis.2020.10.010>, (consultato il 4 marzo 2024).

⁷³ Cfr. *Ivi.* AI in drug discovery.

processo decisionale, determinare la terapia giusta per un paziente, e gestire i dati clinici.⁷⁴

Dallo studio in analisi risulta che i modelli di DL abbiano mostrato una migliore capacità predittiva rispetto agli approcci tradizionali di ML per la valutazione dell'assorbimento, distribuzione, metabolismo, escrezione e tossicità, riassunti con l'acronimo ADMET, dei farmaci candidati.⁷⁵

Il processo di scoperta e sviluppo di un farmaco però può richiedere più di dieci anni e costa in media 2,8 miliardi di dollari USA.⁷⁶ Nonostante questo, nove molecole terapeutiche su dieci falliscono gli studi clinici di Fase II e l'approvazione normativa.⁷⁷ Algoritmi come i *Nearest-Neighbour classifiers*, RF, *extreme learning machines*, SVMs, *deep neural networks* (DNN) vengono utilizzati per lo *screening virtuale* (VS) basato sulla fattibilità di sintesi delle molecole e possono anche prevedere l'attività e la tossicità in vivo del farmaco candidato.⁷⁸

Lo studio procede analizzando come negli ultimi anni, l'approccio al design *de novo* dei farmaci sia stato ampiamente utilizzato per progettare nuove molecole farmaceutiche. Il metodo tradizionale di design *de novo* dei farmaci verrà però sostituito dai metodi di deep learning (DL) in evoluzione, migliorando e velocizzando l'analisi e la progettazione delle nuove molecole. GRZYBOWSKI *et al.*⁷⁹ hanno sviluppato il programma CHEMATICA, ora rinominato SYNTHIA, che ha la capacità di codificare un insieme di regole chimiche nella macchina e proporre possibili percorsi di sintesi per otto target essenziali dal punto di vista medico. Questo programma si è dimostrato efficiente sia nella resa, sia nella riduzione delle spese. SYNTHIA è anche in grado di proporre alternative ai metodi di sintesi per composti già brevettati, migliorandone così l'efficienza e riducendone i costi.⁸⁰ SYNTHIA può quindi aiutare i

⁷⁴ *Ivi.* AI in the lifecycle of pharmaceutical products.

⁷⁵ *Ibidem.*

⁷⁶ *Ivi.* AI in drug screening.

⁷⁷ *Ibidem.*

⁷⁸ *Ibidem.*

⁷⁹ Bartosz A. GRZYBOWSKI, *Chematica: A Story of Computer Code That Started to Think like a Chemist*, 2018, Chem, vol. 4, 3, pp. 390-398, <https://doi.org/10.1016/j.chempr.2018.02.024>.

⁸⁰ Paul DEBLEENA *et al.* cit. AI in de novo drug design.

chimici a scoprire nuovi modi per creare molecole, migliorando l'efficienza e riducendo i costi nel processo di sviluppo di nuovi farmaci.

PUTIN *et al.*⁸¹ hanno esplorato un'architettura DNN chiamata *reinforced adversarial neural computer* (RANC) basata sul *reinforced learning* (RL) per il design *de novo* di piccole molecole organiche. Questa piattaforma è stata addestrata con molecole rappresentate come stringhe “SMILES”, un metodo per descrivere la struttura delle molecole usando una serie di caratteri ASCII. L'architettura RANC è stata confrontata con un'altra piattaforma, ORGANIC, ed è risultata superiore in termini di generazione di strutture uniche ed in grado di mantenere la loro integrità strutturale, senza perdere informazioni significative sulla lunghezza o sulla complessità delle molecole.⁸² In altre parole, questo studio dimostra l'efficacia dell'uso dell'IA per il design di nuove molecole, che potrebbe accelerare la scoperta di nuovi farmaci o di nuovi materiali.

POPOVA *et al.*⁸³ invece, hanno sviluppato una strategia di *reinforcement learning for structural evolution* per la sintesi *de novo* dei farmaci, che coinvolge delle DNN *generative e predittive*. In questo caso, il modello *generativo* sfrutta una memoria a pila per produrre molecole che si distinguono per la loro unicità, rappresentate attraverso le stringhe SMILE, mentre i modelli *predittivi* hanno il compito di analizzare le molecole generate e prevedere le loro proprietà fisico/chimiche. Questo è utile per valutare la fattibilità e la potenziale efficacia delle molecole generate, prima di procedere con la sperimentazione in laboratorio.⁸⁴

Un altro utilizzo dell'IA nella ricerca farmaceutica approfondito nello studio è quello dei trial clinici. Questi, infatti, sono diretti a stabilire la sicurezza e l'efficacia di un prodotto farmaceutico negli esseri umani e richiedono minimo sei o sette anni, oltre ad un sostanziale investimento finanziario. Purtroppo, però, solo una molecola su dieci studiata in questi trial ottiene l'approvazione, il che risulta in una perdita enorme per

⁸¹ Evgeny PUTIN *et al.* *Reinforced Adversarial Neural Computer for de Novo Molecular Design*, 2018, *Journal of Chemical Information and Modeling*, 58 (6), 1194-1204, <https://pubs.acs.org/doi/10.1021/acs.jcim.7b00690>.

⁸² Paul DEBLEENA *et al.* cit. *Ibidem*.

⁸³ Mariya POPOVA *et al.* *Deep reinforcement learning for de novo drug design*, 2018, *Science Advances*, vol. 4 (7), 7885, <https://www.science.org/doi/full/10.1126/sciadv.aap7885>.

⁸⁴ Paul DEBLEENA *et al.* cit. *Ibidem*.

l'industria. I fallimenti possono essere dovuti a diversi motivi, come una selezione inappropriata dei pazienti, la mancanza di requisiti tecnici ed una cattiva infrastruttura.⁸⁵ Ecco che l'IA può aiutare a ridurre questo fenomeno. La selezione dei pazienti, infatti, occupa un terzo della tempistica del trial clinico ed il successo del trial può essere migliorato reclutando pazienti idonei, che altrimenti porta a circa l'86% di casi di fallimento.⁸⁶ L'IA può essere utile nell'assistere alla selezione, per esempio, di una specifica porzione di popolazione composta da pazienti affetti da malattie specifiche per il reclutamento nelle fasi II e III del trial, utilizzando l'analisi del profilo genoma-esposoma dei pazienti, che può aiutare a prevedere quali pazienti avranno una maggiore probabilità di rispondere positivamente al farmaco in esame nel trial.⁸⁷

Un altro punto cruciale e problematico dei trial clinici è il tasso di abbandono dei pazienti, che ha una percentuale del 30% e che crea quindi ulteriori requisiti di reclutamento per il completamento del trial, portando ad uno spreco di tempo e denaro. Questo può essere evitato monitorando attentamente i pazienti e aiutandoli a seguire il protocollo previsto dal trial. Un software mobile sviluppato da AICURE ha monitorato l'assunzione regolare di farmaci da parte dei pazienti con *schizofrenia* in un trial di fase II, che ha aumentato il tasso di aderenza dei pazienti del 25%, garantendo così il completamento ed il successo del trial.⁸⁸

Abbiamo quindi visto come l'implementazione delle tecnologie di *Intelligenza Artificiale* possa aiutare a migliorare la velocità, ridurre i costi e migliorare la probabilità di esito positivo della ricerca farmaceutica e della sperimentazione *in vivo* dei pazienti.

Sebbene attualmente non ci siano ancora farmaci sul mercato sviluppati con approcci basati sull'IA e rimangano sfide specifiche per quanto riguarda l'implementazione di questa tecnologia, è probabile che diventerà invece uno strumento inestimabile nel settore farmaceutico nel prossimo futuro.⁸⁹

⁸⁵ *Ivi.* AI in clinical trial design.

⁸⁶ *Ibidem.*

⁸⁷ *Ibidem.*

⁸⁸ *Ibidem.*

⁸⁹ *Ivi.* Concluding remarks and prospects.

In questo capitolo abbiamo visto i diversi utilizzi dell'IA nel campo medico, e come attraverso diverse tecniche sia possibili riuscire a diagnosticare e prevedere l'insorgenza o la presenza di malattie, oltre che l'utilizzo efficiente dei dati ed il supporto per la ricerca. Grazie a tutte le potenzialità riscontrate è possibile riuscire a trattare un paziente con modalità che non hanno precedenti e questo permette di personalizzare la cura al meglio per ogni individuo, avendo a disposizione dati chiari, diagnosi certe ed un'ampia conoscenza dei trattamenti e le modalità terapeutiche, oltre al profilo genomico del paziente e quindi la possibilità di riuscire a prevedere in buona misura la risposta e l'evoluzione di un trattamento, in modo individuale.

3. RISCHI E REGOLAMENTAZIONE

Nel capitolo precedente abbiamo analizzato i vari utilizzi e le potenzialità dell'IA nel campo medico. Questa tecnologia innovativa, però, non è esente da rischi. Il suo utilizzo, pertanto, dev'essere implementato con attenzione ed in maniera controllata, ricordandosi che si ha a che fare con un ambiente, quello medico, dove il rischio è estremamente alto, trattandosi della vita dei pazienti.

In questo capitolo tratteremo i potenziali rischi e gli eventuali problemi legati all'utilizzo delle tecnologie di *Intelligenza Artificiale*, nel campo della Medicina e vedremo quali sono le regolamentazioni in merito.

3.1 BIAS e discriminazione nell'IA

Il *Bias* (pregiudizio) non è un concetto nuovo nella storia dell'umanità. Questo, infatti, esiste sin dall'antichità e ci ha permesso in qualche modo di sopravvivere nel tempo, permettendoci di valutare un potenziale pericolo in maniera automatica, senza dover analizzare ogni singola situazione con attenzione; pensiamo all'uomo primitivo che si trova di fronte ad un animale feroce; in automatico in lui scatta il meccanismo “*fight or flight*” che spinge l'uomo ad agire ed evitare il pericolo. Negli anni questo meccanismo non è svanito, ma la società si è evoluta ed i pericoli non sono ormai più di natura fisica, ma piuttosto sociale. Ecco che questo meccanismo trova ugualmente il suo posto, permettendoci di valutare situazioni potenzialmente dannose per la nostra sfera sociale, senza doverle analizzare con attenzione, attraverso quei processi mentali che in psicologia sono conosciuti come “*euristiche*”. I pregiudizi, però, possono avere conseguenze negative per chi li subisce, perché si può venir giudicati in maniera errata per il solo fatto di presentare determinate caratteristiche, far parte di una determinata cerchia di individui, etnia e così via. Ecco che questo meccanismo, definito appunto come *Bias*, ha assunto oggi un'accezione totalmente negativa, che lede gli individui senza che però questi dimostrino effettivamente comportamenti negativi.

Ma perché si parla di *Bias* anche nell'IA? In realtà è molto semplice; essendo gli algoritmi basati su dati creati, raccolti e condivisi da esseri umani, è molto facile che insieme a questi, siano trasferiti ed implementati anche i pregiudizi, condizionandone quindi la capacità di decisione ed i risultati finali.

Di seguito vedremo come sia facile introdurre *Bias* all'interno degli algoritmi e le conseguenze che ne derivano.

3.1.1 *Effetto farfalla*

In questa sezione verrà analizzato uno studio che tratta il *Bias* algoritmico e le sue implicazioni. Lo studio in questione è intitolato “*The Butterfly Effect in artificial intelligence systems: Implications for AI bias and fairness*”⁹⁰. L'*effetto farfalla* di cui si parla, è un concetto derivato dalla *Teoria del Caos* (che studia il comportamento dei sistemi dinamici caotici) ed evidenzia come cambiamenti, apparentemente minori, possano portare a risultati significativi e imprevedibili nei sistemi complessi.⁹¹ Fattori come pregiudizi (anche minori) nei dati iniziali, deviazioni durante la fase di addestramento degli algoritmi o cambiamenti nella distribuzione dei dati, dalla fase di training a quella di testing, possono involontariamente portare a risultati squilibrati ed ingiusti. Chi ne subisce le conseguenze sono spesso i gruppi emarginati e le minoranze e questo non fa altro che rinforzare le disuguaglianze sociali già esistenti.⁹² Proprio nel 2019, OBERMEYER *et al.*⁹³ hanno scoperto un *Bias* nel settore sanitario, presente in un algoritmo commerciale ampiamente utilizzato per prevedere le necessità sanitarie e che mostrava un evidente pregiudizio razziale.⁹⁴ L'algoritmo in questione assegnava punteggi di rischio inferiori ai pazienti di colore rispetto a quelli bianchi, nonostante le simili condizioni di salute, portando a disparità nell'accesso ai programmi assistenza

⁹⁰ Emilio FERRARA, *The Butterfly Effect in artificial intelligence systems: Implications for AI bias and fairness*, Machine Learning with Applications, 2024, vol. 15, 100525, ISSN 2666-8270, <https://doi.org/10.1016/j.mlwa.2024.100525>, (consultato il 10 marzo 2024).

⁹¹ Cfr. *Ivi.* Abstract.

⁹² *Ibidem.*

⁹³ Ziad OBERMEYER et al. *Dissecting racial bias in an algorithm used to manage the health of populations*. Science, 2019, vol. 366, pp. 447-453, <https://www.doi.org/10.1126/science.aax2342>.

⁹⁴ Emilio FERRARA, *Cit. Ivi.* Healthcare algorithms.

sanitaria. Dallo studio è emerso che l'algoritmo in esame si affidasse ai costi come criterio per le necessità sanitarie, introducendo involontariamente un *Bias* a causa delle differenze nell'utilizzo dell'assistenza sanitaria.⁹⁵ Questo caso ci permette di capire come anche piccoli pregiudizi all'interno di un algoritmo possano causare un notevole impatto sociale.

Lo studio di FERRARA prosegue analizzando alcune delle cause di *Bias* all'interno degli algoritmi di IA che analizzeremo di seguito:

- Piccoli aggiustamenti nei dati di input possono avere un impatto significativo sull'equità e sul pregiudizio dei sistemi di IA, i quali si affidano ad enormi quantità di dati per prendere decisioni e quindi sensibili ai cambiamenti, anche piccoli.⁹⁶
- Il modo in cui i dati vengono raccolti e campionati può introdurre pregiudizi. Se il processo di raccolta dei dati non è progettato con attenzione, può portare ad una sotto o sovra-rappresentazione di certi gruppi, influenzando così, negativamente, il sistema. (SHANKAR *et al.*⁹⁷).⁹⁸
- Se un particolare gruppo demografico è sottorappresentato nei dati di addestramento, il sistema di IA potrebbe non generalizzare bene quel gruppo, portando a *Bias* ingiusti (BUOLAMWINI & GEBRU⁹⁹).¹⁰⁰
- Anche varie fonti, come la discriminazione storica o gli errori di misurazione possono introdurre *Bias* all'interno degli algoritmi. Ad esempio, i pregiudizi storici presenti nei dati di addestramento possono portare l' IA a perpetuare o esacerbare disuguaglianze già esistenti. Un esempio di questo fenomeno si è

⁹⁵ *Ibidem.*

⁹⁶ *Ivi.* Small adjustments in input data.

⁹⁷ Shreya SHANKAR *et al.* *No Classification without Representation: Assessing Geodiversity Issues in Open Data Sets for the Developing World*, Proceedings of the 2020 conference on fairness, accountability, and transparency, 2020, pp. 445–454, <https://doi.org/10.48550/arXiv.1711.08536>.

⁹⁸ Emilio FERRARA, *Cit. Ivi.* Data sampling.

⁹⁹ Joy BUOLAMWINI, Timnit GEBRU, *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, Proceedings of Machine Learning Research, 2018, vol. 81, pp. 77-91, <https://proceedings.mlr.press/v81/buolamwini18a.html>.

¹⁰⁰ Emilio FERRARA, *Cit. Ivi.* Demographic makeup.

visto nello strumento di valutazione del rischio di recidiva COMPAS, che ha avuto un impatto significativo su diversi gruppi a causa dei pregiudizi presenti nei dati di addestramento (ANGWIN, Larson *et al.*¹⁰¹).¹⁰²

- Scelte come quella di un modello che assume una relazione lineare tra le variabili di input e output, che se non accurata, potrebbe portare ad interpretazioni errate introducendo *Bias*; l'assunzione di indipendenza tra le caratteristiche dei dati, presupponendo che queste non si influenzino reciprocamente, può condurre a conclusioni errate e pregiudizi se invece queste fossero correlate (BERK *et al.*¹⁰³). Inoltre, le tecniche di ottimizzazione come la regolarizzazione, possono introdurre *Bias* se danno priorità a determinati obiettivi rispetto che ad altri (KEARNS *et al.*¹⁰⁴).¹⁰⁵

Abbiamo visto quindi come il *Bias* può essere facilmente introdotto all'interno degli algoritmi di IA ed avere un pesante impatto sociale, specialmente su determinati gruppi di persone.

3.2 (In)sicurezza dell'uso dell'IA in medicina

Quando si ha a che fare con algoritmi di IA è essenziale adottare le dovute precauzioni e misure di sicurezza, a maggior ragione quando si decide di utilizzare questi algoritmi in un ambiente estremamente delicato e considerato ad alto rischio,

¹⁰¹ Julia ANGWIN et al. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks, *Ethics of Data and Analytics*, 2022, pp. 254-264, Auerbach Publications, ISBN 9781003278290.

¹⁰² Emilio FERRARA, Cit. *Ivi*. Biases in data.

¹⁰³ Richard BERK *et al.* *Fairness in Criminal Justice Risk Assessments: The State of the Art*, *Sociological Methods & Research*, 2 luglio 2018, vol. 50, prima edizione (febbraio 2021), <https://doi.org/10.1177/0049124118782533>.

¹⁰⁴ Michael KEARNS *et al.* *Preventing Fairness Gerrymandering: Auditing and Learning for Subgroup Fairness*, *Proceedings of the 35th International Conference on Machine Learning*, 2018, *Proceedings of Machine Learning Research*, vol. 80, pp. 2564-2572, <https://proceedings.mlr.press/v80/kearns18a.html>.

¹⁰⁵ Emilio FERRARA, Cit. *Ivi*. Algorithmic biases.

come quello sanitario. Di seguito andremo ad analizzare qual è il fattore di rischio principale nell'utilizzo dell'IA e cercheremo di capire come l'attribuzione di responsabilità può risultare difficile in un contesto simile.

3.2.1 *Black Box e responsabilità*

L'ascesa degli algoritmi di IA ha cambiato (e sta cambiando) il mondo, ma il loro funzionamento rimane tutt'ora spesso sconosciuto. Le cosiddette *black box* (scatole nere) mappano le caratteristiche dell'utente in una classe o un punteggio senza spiegare il perché, poiché il modello decisionale, talvolta, non è comprensibile né agli stakeholder né agli scienziati.¹⁰⁶ Questo è chiaramente preoccupante non solo per la mancanza di trasparenza, ma anche per i possibili *Bias* nascosti negli algoritmi e per la sicurezza.¹⁰⁷ È di fondamentale importanza, quindi, capire cosa spinga un algoritmo a prendere una determinata decisione e riuscire ad “aprire” la scatola nera, in modo tale da comprendere i meccanismi che hanno portato ad un determinato risultato, non solo per capire il perché, ma anche per poterli utilizzare in maniera sicura, senza zone d'ombra o opacità. È chiaramente un compito arduo, ma dev'essere un obiettivo imperativo per la buona implementazione negli algoritmi, specialmente in campi così delicati.

Un altro punto fondamentale dell'implementazione degli algoritmi di IA nel campo medico è l'attribuzione di responsabilità nei soggetti coinvolti. Chi è da ritenere responsabile nel caso di una decisione clinica sbagliata, che danneggia il paziente, nel caso in cui sia stata utilizzata l'IA? L'importanza di questa domanda è direttamente proporzionale alla difficoltà nell'averne una risposta definitiva e certa. Di chi è la responsabilità quindi? Del medico che ne ha fatto uso, di chi l'ha progettata o dell'IA stessa?

¹⁰⁶ Cfr. Dino PEDRESCHI *et al.* *Meaningful Explanations of Black Box AI Decision Systems*, Proceedings of the AAAI Conference on Artificial Intelligence, 2019, vol. 33(1), pp. 9780-9784, <https://doi.org/10.1609/aaai.v33i01.33019780>, (consultato il 10 marzo 2024).

¹⁰⁷ *Ibidem.*

Vediamo cosa emerge dallo studio “*Clinical AI: opacity, accountability, responsibility and liability*”¹⁰⁸ a riguardo. La definizione che lo studio fornisce di responsabilità è la seguente: «La responsabilità è quando un individuo è obbligato a spiegare (rendere conto) a coloro che hanno il diritto di chiedere (regolatori, un paziente) il processo decisionale che ha guidato le sue azioni o omissioni.»¹⁰⁹. Dall’analisi dello studio si possono comprendere due cose fondamentali, ovvero che la responsabilità da parte dei clinici è richiesta dai codici di condotta che comportano che il clinico sia in grado di giustificare le proprie decisioni e lo obbligano a fornire un’assistenza di qualità con un’enfasi sulla sicurezza; e che non esiste, invece, un requisito analogamente applicato affinché i tecnologi siano personalmente responsabili, perché, a differenza dei clinici, i tecnici informatici non hanno codici di condotta simili, ma semplici linee guida etiche.¹¹⁰ Ancor meno probabile è la possibilità che sia l’IA stessa ad essere ritenuta responsabile, dato che non è effettivamente senziente e nemmeno in grado di comprendere, nel modo in cui lo intendiamo noi, le sue azioni e decisioni. L’opacità gioca un ruolo chiave nell’attribuzione della responsabilità in casi simili. La letteratura non ha ancora trovato un punto fermo su cui essere d’accordo: c’è chi come Blay WHITBY¹¹¹, sostiene che i tecnici informatici debbano condividere la responsabilità per le conseguenze con i clinici quando un’IA fornisce consigli inappropriati e utilizzati dal clinico; mentre, come rilevato dalle interviste di FENECH *et al*¹¹², i tecnici non dovrebbero detenere la responsabilità per un sistema quando è stato progettato per assistere il processo decisionale clinico, piuttosto che sostituirlo (ad esempio, nel sistema DEEPMIND) in cui la responsabilità ricade sul il clinico che ne ha fatto uso.¹¹³

¹⁰⁸ Helen SMITH, *Clinical AI: opacity, accountability, responsibility and liability*, AI & SOCIETY, 2021, vol. 36, pp. 535–545, <https://doi.org/10.1007/s00146-020-01019-6>, (consultato l’11 marzo 2024).

¹⁰⁹ *Ivi*. Why is accountability important?

¹¹⁰ *Ibidem*.

¹¹¹ Blay WHITBY, *Automating Medicine the Ethical Way*, Machine Medical Ethics. Intelligent Systems, Control and Automation: Science and Engineering, 2015, vol 74, https://doi.org/10.1007/978-3-319-08108-3_14.

¹¹² Matthew FENECH, *Ethical, social and political challenges of artificial intelligence in health*, Future Advocacy report for the Wellcome Trust, 2018, <https://wellcome.org/sites/default/files/ai-in-health-ethical-social-political-challenges.pdf>.

¹¹³ Helen SMITH cit. *Ivi*. Responsibility.

È quindi chiaro che, quando un clinico decide di utilizzare un sistema di IA, opaco o meno, sia egli ad assumersi la responsabilità della decisione, e che, invece, chi lavora all'algoritmo debba assumersi la responsabilità di progettarlo al meglio ed in modo sicuro, nonostante non sia direttamente a contatto con i pazienti. È importante quindi formulare, a livello normativo, delle linee guida chiare a riguardo, che riusciranno in futuro a rispondere a questa difficile domanda.

3.3 Privacy dei dati

Un altro punto cruciale su cui è doveroso soffermarsi è la *Privacy* dei dati, che inevitabilmente entra in conflitto con gli algoritmi di IA, in particolare quando si trattano dati sensibili, come quelli medici.

Man mano che i modelli di IA diventano sempre più sofisticati e integrati negli ambienti clinici, si rende necessaria una forte attenzione sull'integrità e sulla sicurezza dei dati. Il ricorso a vasti archivi di dati sanitari per gli algoritmi di IA pone in primo piano sfide etiche cruciali. Queste includono preoccupazioni relative alla privacy dei pazienti e all'integrità dei dati. Pertanto, vi è un'urgente necessità di gestire con attenzione la raccolta, l'archiviazione e l'elaborazione dei dati sanitari per mantenere la riservatezza del paziente.¹¹⁴ Le maggiori preoccupazioni riguardo la *Privacy* includono l'accesso, l'uso ed il controllo dei dati dei pazienti. Un altro insieme di preoccupazioni è relativo al rischio esterno di violazioni della privacy attraverso metodi guidati dall'IA; la capacità di de-identificare o anonimizzare i dati sanitari dei pazienti potrebbe essere infatti compromessa o addirittura annullata alla luce di nuovi algoritmi che hanno avuto successo nel re-identificare tali dati.¹¹⁵ I problemi che emergono, quindi, sono due: che una parte significativa della tecnologia esistente relativa all'apprendimento automatico è sotto il controllo di grandi gruppi commerciali

¹¹⁴ Cfr. Steven M. WILLIAMSON, Victor PRYBUTOK, *Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare*, *Applied Sciences*, 2024, vol. 14(2):675, <https://doi.org/10.3390/app14020675>, (consultato il 12 marzo 2024).

¹¹⁵ Cfr. Blake MURDOCH, *Privacy and artificial intelligence: challenges for protecting health information in a new era*, *BMC Med Ethics*, 2021, vol. 22, n. 122, <https://doi.org/10.1186/s12910-021-00687-3>, (consultato il 12 marzo 2024).

(per esempio GOOGLE, con DEEPMIND) e che l'uso di queste IA è collegato al rischio esterno di violazioni della privacy, da parte degli stessi sistemi algoritmici, altamente sofisticati, in grado di re-identificare dati che sono stati in precedenza de-identificati.¹¹⁶ Un sondaggio¹¹⁷ condotto nel 2018 su quattromila adulti americani ha scoperto che solo l'11% di loro è disposto a condividere i dati sanitari con le aziende tecnologiche, contro il 72%, disposto a condividerli con i medici.¹¹⁸ Uno studio di NA *et al.*¹¹⁹ ha scoperto che un algoritmo potrebbe essere utilizzato per riconoscere l'85,6% degli adulti ed il 69,8% dei bambini in uno studio sull'attività fisica, «nonostante l'aggregazione dei dati e la rimozione delle informazioni sanitarie protette»¹²⁰.¹²¹ Un altro studio¹²² condotto nel 2019 ha utilizzato con successo una tecnica, il “*linkage attack framework*”, cioè un algoritmo mirato a riconoscere informazioni sanitarie anonime, che può collegare dati sanitari, presenti online, a persone reali, dimostrando «la vulnerabilità dei dati sanitari esistenti online»¹²³.¹²⁴

È chiaro quindi che, vista la delicata natura dei dati sanitari medici, vista la grande quantità di dati personali sanitari utilizzati dagli algoritmi e vista la vulnerabilità di questi dati, sia necessario implementare non solo forti linee guida sull'utilizzo e sull'implementazione degli algoritmi, ma anche una regolamentazione altrettanto forte che riesca a stare al passo con la sempre più rapida evoluzione di queste tecnologie.

¹¹⁶ *Ibidem.*

¹¹⁷ ROCK HEALTH, *Beyond wellness for the healthy: digital health consumer adoption*, 2018. https://rockhealth.com/reports/beyond-wellness-for-the-healthy-digital-health-consumer-adoption-2018/?mc_cid=0c97d69dbe&mc_eid=452e95c5c5.

¹¹⁸ Blake MURDOCH, Cit. *Ivi*. Concerns with access, use and control.

¹¹⁹ Liangyuan NA, Feasibility of Reidentifying Individuals in Large National Physical Activity Data Sets From Which Protected Health Information Has Been Removed With Use of Machine Learning, *JAMA Netw Open*, 2018, vol.1(8), <https://www.doi.org/10.1001/jamanetworkopen.2018.6040>.

¹²⁰ *Ibidem.*

¹²¹ Blake MURDOCH, Cit. *Ivi*. The problem of reidentification.

¹²² Shouling Ji, *De-Health: All Your Online Health Information Are Belong to Us*, arXiv preprint, 3 Giugno 2019, v. 2, <https://arxiv.org/abs/1902.00717>.

¹²³ *Ibidem.*

¹²⁴ Blake MURDOCH, Cit. *Ibidem.*

Alcune tecnologie per migliorare l'approccio alla privacy sono già state proposte, come emerge dallo studio di WILLIAMSON; e sono: la tecnologia *blockchain*, che offre un metodo sicuro e decentralizzato per archiviare e gestire i dati sanitari, utilizzando un sistema a registro distribuito per garantire che i dati dei pazienti siano immutabili e verificabili, impedendo l'accesso non autorizzato e le potenziali violazioni dei dati; il *Federated Learning* (FL), che consente di addestrare modelli di IA su dispositivi o server decentralizzati che contengono campioni di dati locali, eliminando la necessità di trasferire dati sensibili dei pazienti; la *Homomorphic Encryption* (HE), che offre una soluzione per eseguire calcoli su dati crittografati senza decrittazione, permettendo così all'IA di apprendere dai dati crittografati, mantenendo la privacy anche quando è richiesta l'analisi da parte di terzi; infine, la *Differential Privacy* (DP), che introduce "rumore" nei set di dati, per oscurare le identità individuali, consentendo l'addestramento di modelli di IA su dati rappresentativi della popolazione senza compromettere la privacy individuale.¹²⁵

Queste tecnologie promettono quindi di proteggere la *Privacy* dei pazienti in modo innovativo e devono farlo al fianco di normative altrettanto robuste.

3.4 Normative sulla protezione dei dati nel mondo

Vediamo in questa sezione quali sono le normative a livello mondiale riguardanti la protezione dei dati personali e quali invece si sono già addentrate nel difficile compito di fornire una normativa sull'IA.

In Europa abbiamo il *General Data Protection Regulation* (GDPR), entrato in vigore nel 2016 e che ha trovato applicazione dal 25 maggio 2018. Questo complesso normativo è uno dei più solidi a livello mondiale in materia di protezione dei dati personali e ha fornito all'Europa gli strumenti per favorire la circolazione delle informazioni adeguandosi all'evoluzione tecnologica, rimanendo "umano-centrica" e si pone l'obiettivo di proteggere i diritti e le libertà degli individui conferendo loro

¹²⁵ Cfr. Steven M. WILLIAMSON, Cit. Implementing Ethical AI in Healthcare.

mezzi utili ad esercitare un controllo attivo e proattivo sui dati personali che li riguardano.¹²⁶

In Cina vige una normativa simile per certi aspetti al nostro GDPR, la *Personal Information Protection Law* (PIPL), entrata in vigore nel 2021. A differenza del GDPR però, il PIPL, offre ampia discrezionalità al governo e agli enti amministrativi in merito ai dati personali e alla promozione degli interessi di sicurezza nazionale, prevedendo che questi possano ostacolare il trasferimento dei dati al di fuori della Cina.¹²⁷

Negli USA, invece, non esiste una legge unica in tema di protezione dei dati personali. Le disposizioni in tale ambito sono frazionate e rinvenibili in diversi strumenti statali che tutelano la protezione dei dati dei consumatori.¹²⁸ Una regolamentazione degna di nota è la *California Consumer Privacy Act* (CCPA), una legge sulla privacy entrata in vigore nel 2020, che offre ai consumatori californiani una serie di misure di tutela della privacy, tra cui il diritto di accedere alle proprie informazioni personali, nonché di eliminarle o di disattivarne la “vendita”.¹²⁹ Successivamente, il 1° gennaio 2023, è stato introdotto il *California Privacy Rights Act* (CPRA) che modifica ed amplia il CCPA.¹³⁰ Da queste leggi, più stati federali hanno proposto delle legislazioni simili per proteggere i consumatori.¹³¹ Riguardo all’ambito sanitario, nel 1996 è stato introdotto l’*Health Insurance Portability and Accountability Act* (HIPAA), legge federale che definisce i requisiti per il trattamento dei dati sanitari dei privati.¹³² Recentemente alcuni senatori hanno introdotto l’*Health*

¹²⁶ Cfr. Federica PUCARELLI, *La protezione dei dati nel mondo: una breve panoramica*, Privacy, 8 Marzo 2022, disponibile sul sito dello Studio Segale Stefanelli, <https://www.studiolegalestefanelli.it/it/approfondimenti/la-protezione-dei-dati-nel-mondo-una-breve-panoramica/>, (consultato il 12 marzo 2024).

¹²⁷ Cfr. Carlo IMPALÀ, *Il GDPR e il PIPL: similitudini e differenze*, Professione e Mercato, Ntplusdiritto, 14 Giugno 2023, <https://ntplusdiritto.ilsole24ore.com/art/il-gdpr-e-pipl-similitudini-e-differenze-AEONHogD>, (Consultato il 12 Marzo 2024).

¹²⁸ Federica PUCARELLI, Cit., USA.

¹²⁹ Cfr. *California Consumer Privacy Act (CCPA)*, Google, <https://cloud.google.com/security/compliance/ccpa?hl=it>, (Consultato il 12 marzo 2024).

¹³⁰ *Ibidem*.

¹³¹ Cfr. Federica PUCARELLI, Cit. USA.

¹³² *Ibidem*.

Data Use and Privacy Commission Act, volto ad avviare il processo di modernizzazione dell'utilizzo dei dati sanitari e delle politiche sulla privacy.¹³³

Nel Regno Unito vige il *Data Protection Act* 2018, ovvero la normativa di attuazione del GDPR, che tutt'ora è in vigore.¹³⁴ A Taiwan, invece, vige il *Taiwan Personal Data Protection Act* (PDPA) ed offre livelli di protezioni simili al GDPR.¹³⁵ Anche in Brasile vige una normativa che si allinea molto al GDPR, entrata in vigore il 18 settembre 2020, il *General Data Protection Law* (LGPD).¹³⁶

Per quanto riguarda le normative in tema di IA, di seguito vedremo quali sono le principali a livello mondiale.

In prima linea, vediamo l'Europa, che ha raggiunto un accordo su un quadro normativo per la regolamentazione dell'IA. Proprio nella data di scrittura di questa sezione (13 marzo 2024) l'UE ha approvato definitivamente l'*AI Act*, che verrà promulgato a maggio 2024. La legge si basa su un sistema di rischi, che divide i sistemi di IA sulla base di quattro categorie di rischio: minimo, limitato, alto e inaccettabile.¹³⁷ Maggiore è il rischio, maggiori sono le responsabilità e i limiti per chi sviluppa o usa questi sistemi, fino ad arrivare ai modelli troppo pericolosi per essere autorizzati, come le tecnologie per la categorizzazione biometrica, il riconoscimento delle emozioni sul posto di lavoro o a scuola, i sistemi di *social scoring*, la polizia predittiva, ovvero l'uso di dati che servono a calcolare le probabilità che una persona commetta un reato, ecc.¹³⁸

¹³³ *Ibidem*.

¹³⁴ *Ibidem*.

¹³⁵ Cfr. *Data Privacy in Taiwan: PDPA, GDPR, and CCPA*, Sandline, 14 giugno 2022, <https://sandlineglobal.com/the-pdpa-establishes-data-privacy-in-taiwan/>, (consultato il 12 marzo 2024).

¹³⁶ *Data Protection Laws of the World*, <https://www.dlapiperdataprotection.com/index.html?t=law&c=BR>, (consultato il 12 marzo 2024).

¹³⁷ Kevin CARBONI, *Il Parlamento europeo ha approvato l'AI Act*, Diritti, 13 marzo 2024, WIRED, <https://www.wired.it/article/ai-act-approvato-parlamento-europeo/>, (consultato il 13 marzo 2024).

¹³⁸ *Ibidem*.

Il Regno Unito, invece, ha deciso di adottare un approccio “*pro-innovazione*” che evita nuove e complesse misure di sicurezza per la tecnologia.¹³⁹ Verranno quindi utilizzate le normative già esistenti e adattate in modo da allinearsi con l’IA.¹⁴⁰ Tuttavia, il governo UK ha ospitato il *Vertice Internazionale per la Sicurezza dell’IA* nel novembre 2023. L’evento ha portato alla *Dichiarazione di Bletchley*, un accordo firmato da 28 paesi, inclusi gli Stati Uniti e la Cina, che mira a stabilire una comprensione condivisa dei rischi legati all’IA e sottolinea la necessità di “gestirli collettivamente” attraverso la collaborazione internazionale.¹⁴¹

Anche in Cina sono stati sviluppati quadri normativi per la regolamentazione dell’IA, a partire dal 2021. Il loro approccio però, a differenza di quello Europeo, si concentra «sul mantenimento dell’ordine sociale, in particolare affrontando la disinformazione.»¹⁴² Questo viene fatto attraverso normative che si concentrano principalmente sul controllo dell’uso dell’IA *generativa*. Anche il controllo sui sistemi di raccomandazione alimentati dall’IA è al centro delle leggi cinesi sull’IA, con misure come il divieto di prezzi dinamici ed un rigoroso requisito di trasparenza per gli utenti che interagiscono con contenuti generati dall’IA.¹⁴³

L’approccio degli USA, invece, si concentra su un approccio specifico per settore, caratterizzato dall’*Ordine Esecutivo sull’uso sicuro, protetto e affidabile dell’intelligenza artificiale*, pubblicato dalla *Casa Bianca* nell’ottobre 2023. L’iniziativa stabilisce che i principali sviluppatori di IA dovrebbero valutare e notificare “le potenziali minacce dei loro algoritmi alla sicurezza nazionale”, inclusi i dati con cui addestrano i loro modelli. Il loro approccio quindi, come in UK, è pro-innovazione, incentivando lo sviluppo tecnologico e mitigando i rischi «richiedendo

¹³⁹ Livia GIANNOTTI, *The state of AI regulation around the world*, Ai And Automation, 8 Gennaio 2024, disponibile sul sito TECHMONITOR, <https://techmonitor.ai/technology/ai-and-automation/the-state-of-ai-regulation-around-the-world>, (consultato il 13 marzo 2024).

¹⁴⁰ *Ibidem*.

¹⁴¹ *Ibidem*.

¹⁴² *Ivi*. China: a focus on generative AI.

¹⁴³ Cfr. *Ibidem*.

al segretario competente di pubblicare linee guida su come le autorità federali dovrebbero supervisionare l'uso dell'IA [...]»¹⁴⁴.

Abbiamo visto come i diversi stati stanno affrontando la regolamentazione della *Privacy* e dell'IA, con l'Europa che sembra essere un passo avanti rispetto alle altre, per quanto riguarda la tutela degli individui.

¹⁴⁴ Cfr. *Ivi.* US: sector-specific measures.

4. CONCLUSIONI

In conclusione, si può affermare che l'IA stia avendo un enorme impatto nel mondo della medicina, permettendoci di raggiungere livelli di precisione diagnostica e terapeutica mai visti fino ad ora e fornendo strumenti chiave per riuscire ad affrontare sfide che fino a qualche tempo fa si reputavano quasi impossibili da superare. Tutto questo rende possibile l'idea che in futuro, i casi di incertezza o le malattie senza cura, vengano ridotte al minimo. Non bisogna scordarsi però che tutto ciò che promette grandi cambiamenti, possa anche comportare un gran livello di rischio e che debba essere implementato con grande attenzione, accompagnato da forti linee guida etiche, normative ed un rigido controllo.

Prima di giungere al termine di questa tesi, rifletteremo sull'impatto dell'IA sulla medicina tradizionale e cosa potrebbe significare questo per i medici, concludendo con un pensiero personale del sottoscritto.

4.1 Impatto dell'IA sulla medicina tradizionale.

«Abbiamo ricevuto i risultati dei tuoi esami e sembra che ci siano alcune cellule che non si comportano come dovrebbero. Quello che vediamo da questo comportamento sembra indicare che ci sia un cancro. Voglio rassicurarti che sentirsi scioccati o spaventati è assolutamente naturale. Ma voglio anche che tu sappia che non sei solo in questo».¹⁴⁵

Questa è la frase che ci si può aspettare da un medico, al momento della comunicazione di una diagnosi ad un paziente, ebbene, questa frase non è stata pronunciata da un medico, ma da GPT-4. Vediamo che il modello di linguaggio non è solo in grado, come visto in precedenza, di diagnosticare e valutare il livello di *triage*

¹⁴⁵ Cfr. Thomas K. LEW, *Will AI and ChatGPT replace doctors like me on the other end of the stethoscope?*, Usa Today, Opinion, <https://eu.usatoday.com/story/opinion/voices/2023/11/27/ai-chatbots-medicine-healthcare-impact-patient-care/71666199007/>, (consultato il 16 marzo 2024).

in un caso clinico, ma anche di comunicare, in modo apparentemente empatico, una diagnosi pesante come questa, ad un paziente. Cosa significa quindi? Che l'IA sarà in grado di sostituire i medici in tutto, addirittura nella comunicazione umana empatica? Effettivamente dalla frase letta si potrebbe pensare che GPT sia in grado di empatizzare e rassicurare un paziente anche di fronte ad una situazione grave come quella di una malattia potenzialmente mortale. Per rispondere a questa domanda occorre però comprendere cosa sia l'empatia. La definizione di empatia, fornita dal sito di TRECCANI, è la seguente: *“capacità di porsi nella situazione di un'altra persona o, più esattamente, di comprendere immediatamente i processi psichici dell'altro”*. Ecco che qui si sta parlando di comprensione e di porsi nella situazione di un'altra persona, cosa che, fino ad ora, è possibile solo agli esseri senzienti e coscienti. GPT-4 e le tecnologie di IA non sono ancora (e forse non lo saranno mai) né senzienti né coscienti; pertanto, l'empatia non è un concetto applicabile a questi, nonostante siano in grado di replicare in maniera quasi perfettamente credibile una conversazione in cui questo sentimento è fondamentale. Quindi, a rigor di logica, l'IA e i chatbot non potranno (probabilmente mai) sostituire un essere umano in questo compito, perché non in grado di comprendere, provare sentimenti e perché privi di coscienza.

Il ruolo del medico quindi non verrà sostituito dall'IA, ma anzi, assumerà un'importanza ancor maggiore, diventando il punto di distinzione tra uno strumento atto a fornire le migliori prestazioni tecniche possibili ed un essere umano, in grado di utilizzare questi strumenti in maniera corretta, per fare ciò che conta più di qualsiasi altra cosa e che sta alla base di tutto quello che tratta questo studio e la Medicina da sempre, ovvero, quello di mettere sempre la vita e la salute del paziente al primo posto.

4.2 Un pensiero personale

Voglio concludere questa tesi con il mio personale pensiero riguardo l'introduzione dell'IA in Medicina. Ho iniziato a scrivere la tesi con una conoscenza limitata e idealizzata di ciò che l'IA fosse in grado di fare e del suo impatto nel campo medico, credendo che questa fosse la svolta a tutti i problemi di salute esistenti. Mi sono reso conto, durante la ricerca e lo studio, che a differenza di ciò che mi aspettavo, non ci fosse nulla di “magico” nell'IA in grado di risolvere tutti i problemi sanitari esistenti. Ma grazie a questa ricerca ho capito come questi strumenti possano essere

estremamente utili per migliorare ciò che già siamo, ovvero entità che amano la vita, e che farebbero di tutto per trovare una soluzione ai problemi, perché abbiamo un intrinseco bisogno di rispondere ai nostri perché e di superare le sfide che ci si pongono davanti, facendo sì che l'impossibile diventi possibile.

Ritengo che oggi più che mai siamo vicini a vedere questi traguardi impossibili diventare realtà, e credo che l'avanzamento tecnologico ci porterà sempre più vicino ad un mondo in cui la sanità sarà in grado di superare i limiti che ancora abbiamo, in cui le malattie incurabili avranno una cura e in cui tutti avranno la stessa possibilità di accedere a queste cure. Ma questo, ricordandoci sempre che l'*Intelligenza Artificiale* è solo uno strumento (incredibilmente potente, sia chiaro) a supporto nostro, di esseri umani, dotati di coscienza, intelligenza e di sentimenti. E ci spero davvero, perché so bene cosa significhi avere a che fare con qualcosa che è destinato a cambiare il corso del regalo più grande che ci sia mai stato fatto, la Vita.

BIBLIOGRAFIA

- AMGEN, *Malattie infiammatorie*, <https://www.amgen.it/en/aree-terapeutiche/malattie-infiammatorie>.
- ANGWIN Julia et al. *Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks*, *Ethics of Data and Analytics*, 2022, pp. 254-264, Auerbach Publications, ISBN 9781003278290.
- ASHRAF et al., 2021, *Deep transfer learning for Alzheimer neurological disorder detection*, *Multimedia Tools and Applications*.
- BALGETIR Ferhat et al. *Detection of ataxia in low disability MS patients by hybrid convolutional neural networks based on images of plantar pressure distribution*, *Multiple Sclerosis and Related Disorders*, 2021, vol. 56, 103261, <https://doi.org/10.1016/j.msard.2021.103261>.
- BERK Richard et al. *Fairness in Criminal Justice Risk Assessments: The State of the Art*, *Sociological Methods & Research*, 2 luglio 2018, vol. 50, prima edizione (febbraio 2021), <https://doi.org/10.1177/0049124118782533>.
- BILEK Furkan et al. *Quantitative Assessment of Ataxia in Multiple Sclerosis Patients using Spatiotemporal Parameters: A Relief-Based Machine Learning Analysis*, *Physikalische Medizin, Rehabilitationsmedizin, Kurortmedizin* 2021; 31(06): pp. 367-376, <https://doi.org/10.1055/a-1512-4858>.
- BRESNICK Jennifer, *What Is Deep Learning and How Will It Change Healthcare?*, *Features*, disponibile nel sito HEALTH IT ANALYTICS, <https://healthitanalytics.com/features/what-is-deep-learning-and-how-will-it-change-healthcare>.

- BRINGAS et al. 2020, *Alzheimer's Disease stage identification using deep learning models*, Journal of Biomedical Informatics, 109, 103514.
- BRITZE Josefine, FREDERIKSEN Jette Lautrup, *Optical coherence tomography in multiple sclerosis*, Eye, Maggio 2018, vol. 32, pp. 884–888, <https://doi.org/10.1038/s41433-017-0010-2>.
- BROWNLIE Jason, *A Gentle Introduction to Generative Adversarial Networks (GANs)*, 19 luglio 2019, Generative Adversarial Networks, disponibile nel sito MACHINE LEARNING MASTERY, <https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/>.
- BUOLAMWINI Joy, GEBRU Timnit, *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, Proceedings of Machine Learning Research, 2018, vol. 81, pp. 77-91, <https://proceedings.mlr.press/v81/buolamwini18a.html>.
- CARBONI Kevin, *Il Parlamento europeo ha approvato l'AI Act*, Diritti, 13 marzo 2024, WIRED, <https://www.wired.it/article/ai-act-approvato-parlamento-europeo/>.
- CASTIGLI Mirella, *ChatGPT guida completa: cos'è, come si usa e cosa può fare per aziende e professionisti*, Cultura e Società Digitali, 19 aprile 2023, ultimo aggiornamento 9 febbraio 2024, disponibile nel sito AGENDA DIGITALE, <https://www.agendadigitale.eu/cultura-digitale/introduzione-a-chatgpt-cose-come-si-usa-e-cosa-puo-fare/>.
- CENTER OF DISEASE CONTROL AND PREVENTION, *Alzheimer's Disease and Related Dementias, What is Alzheimer's Disease?*, Alzheimer's Disease and Healthy Aging, <https://www.cdc.gov/aging/aginginfo/alzheimers.htm>.
- CHING Wong P., ABDULLAH Shahrum S., SHAPIAI Mohd I., *Current Trends in Deep Learning Techniques for Alzheimer's Disease. Diagnosis Using Medical Images: A Systematic Literature Review*, 13 febbraio 2024, Stato di pubblicazione attuale: Under Review, HELIYON, <https://ssrn.com/abstract=4720066>.

CORONADO Ivan et al. *Deep learning segmentation of gadolinium-enhancing lesions in multiple sclerosis*, *Multiple Sclerosis Journal*, 2021, vol. 27(4), pp. 519-527, <https://doi.org/10.1177/1352458520921364>.

DATA PROTECTION LAWS OF THE WORLD, <https://www.dlapiperdataprotection.com/index.html?t=law&c=BR>.

DEBLEENA Paul et al. *Artificial intelligence in drug discovery and development*, *Drug discovery today*, gennaio 2021, vol. 26; 1, pp. 80-93, disponibile sul sito SCIENCE DIRECT, <https://doi.org/10.1016/j.drudis.2020.10.010>.

DEL PALOMAR Pérez et al. *Swept source optical coherence tomography to early detect multiple sclerosis disease*. The use of machine learning techniques, *PLOS ONE*, 2019, 14(5), <https://doi.org/10.1371/journal.pone.0216410>.

FENECH Matthew, *Ethical, social and political challenges of artificial intelligence in health*, Future Advocacy report for the Wellcome Trust, 2018, <https://wellcome.org/sites/default/files/ai-in-health-ethical-social-political-challenges.pdf>.

FERRARA Emilio, *The Butterfly Effect in artificial intelligence systems: Implications for AI bias and fairness*, *Machine Learning with Applications*, 2024, vol. 15, 100525, ISSN 2666-8270, <https://doi.org/10.1016/j.mlwa.2024.100525>.

GIANNOTTI Livia, *The state of AI regulation around the world*, *Ai And Automation*, 8 Gennaio 2024, TECHMONITOR, <https://techmonitor.ai/technology/ai-and-automation/the-state-of-ai-regulation-around-the-world>.

GOOGLE, *California Consumer Privacy Act (CCPA)*, <https://cloud.google.com/security/compliance/ccpa?hl=it>.

GRZYBOWSKI Bartosz A., *Chematica: A Story of Computer Code That Started to Think like a Chemist*, *Chem*, 2018, vol. 4, 3, pp. 390-398, <https://doi.org/10.1016/j.chempr.2018.02.024>.

- GÜNGÖR Gür S, AKMAN Ahmet, *Are All Retinal Nerve Fiber Layer Defects on Optic Coherence Tomography Glaucomatous?* Turk J Ophthalmol, 2017; 47: pp. 267-273, <https://pubmed.ncbi.nlm.nih.gov/29109895/>.
- HUMANITAS RESEARCH HOSPITAL, *Che cosa sono le malattie autoimmuni?*, Malattie autoimmuni, <https://www.humanitas.it/malattie/malattie-autoimmuni/>.
- IMPALÀ Carlo, *Il GDPR e il PIPL: similitudini e differenze, Professione e Mercato*, Ntplusdiritto, 14 giugno 2023, <https://ntplusdiritto.ilsole24ore.com/art/il-gdpr-e-pipl-similitudini-e-differenze-AEONHogD>.
- ITO Naoki et al. *The Accuracy and Potential Racial and Ethnic Biases of GPT-4 in the Diagnosis and Triage of Health Conditions: Evaluation Study*, 2 novembre 2023, JMIR Med Educ, vol. 9, 47532, <https://mededu.jmir.org/2023/1/e47532/>.
- Ji Shouling, *De-Health: All Your Online Health Information Are Belong to Us*, arXiv preprint, 3 Giugno 2019, v. 2, <https://arxiv.org/abs/1902.00717>.
- KADRI et al., *Deep Squeeze and Excitation-Densely Connected Convolutional Network with cGAN for Alzheimer's Disease Early Detection*, Intelligent Systems Design and Applications, 2022, vol 418, Springer, Cham, https://doi.org/10.1007/978-3-030-96308-8_41.
- KAYA Mustafa et al. *Detection of ataxia with hybrid convolutional neural network using static plantar pressure distribution model in patients with multiple sclerosis*, Computer Methods and Programs in Biomedicine, vol. 214, 2022, <https://doi.org/10.1016/j.cmpb.2021.106525>.
- KEARNS Michael et al. *Preventing Fairness Gerrymandering: Auditing and Learning for Subgroup Fairness*, Proceedings of the 35th International Conference on Machine Learning, 2018, Proceedings of Machine Learning Research, vol. 80, pp. 2564-2572, <https://proceedings.mlr.press/v80/kearns18a.html>.

KERNER Sean Michael, *large language models (LLMs)*, Definition, What is generative AI? Everything you need to know, Settembre 2023, disponibile sul sito TECH TARGET, <https://www.techtarget.com/whatis/definition/large-language-model-LLM>.

LEW Thomas K., *Will AI and ChatGPT replace doctors like me on the other end of the stethoscope?*, Usa Today, Opinion, <https://eu.usatoday.com/story/opinion/voices/2023/11/27/ai-chatbots-medicine-healthcare-impact-patient-care/71666199007/>.

LÖTSCH Jörn et al. *Machine-learning based lipid mediator serum concentration patterns allow identification of multiple sclerosis patients with high accuracy*, Scientific Reports, 5 Ottobre 2018, <https://doi.org/10.1038/s41598-018-33077-8>.

MOAZAMI Faezeh et al., *Machine Learning Approaches in Study of Multiple Sclerosis, Disease Through Magnetic Resonance Images*, *Frontiers in Immunology*, ML Approaches in Study of MS, Agosto 2021, vol.12, <https://doi.org/10.3389/fimmu.2021.700582>.

MURDOCH Blake, *Privacy and artificial intelligence: challenges for protecting health information in a new era*, BMC Med Ethics, 2021, vol. 22, n. 122, <https://doi.org/10.1186/s12910-021-00687-3>.

NA Liangyuan, *Feasibility of Reidentifying Individuals in Large National Physical Activity Data Sets From Which Protected Health Information Has Been Removed With Use of Machine Learning*, JAMA Netw Open, 2018, vol.1(8), <https://www.doi.org/10.1001/jamanetworkopen.2018.6040>.

NABIZADEH Fardin et al. *Artificial intelligence in the diagnosis of multiple sclerosis: A systematic review*, Multiple Sclerosis and Related Disorders, 2022, vol. 59, <https://www.sciencedirect.com/science/article/pii/S2211034822001882>.

NHIA, *NHIA Organization*, <https://www.nhi.gov.tw/en/cp-18-63f1c-11-2.html>.

- NHIA, *What is My Health Bank?*, My Health Bank Introduction, <https://www.nhi.gov.tw/en/cp-562-2313c-22-2.html>.
- OBERMEYER Ziad et al. *Dissecting racial bias in an algorithm used to manage the health of populations*, Science, 2019, vol. 366, pp. 447-453, <https://www.doi.org/10.1126/science.aax2342>.
- ODUSAMI et al., 2022, An Intelligent System for Early Recognition of Alzheimer's Disease Using Neuroimaging, Sensors.
- PEDRESCHI Dino et al. *Meaningful Explanations of Black Box AI Decision Systems*, Proceedings of the AAAI Conference on Artificial Intelligence, 2019, vol. 33(1), pp. 9780-9784, <https://doi.org/10.1609/aaai.v33i01.33019780>.
- POPOVA Mariya et al. *Deep reinforcement learning for de novo drug design*, Science Advances, 2018, vol. 4 (7), 7885, <https://www.science.org/doi/full/10.1126/sciadv.aap7885>.
- PUCARELLI Federica, *La protezione dei dati nel mondo: una breve panoramica*, Privacy, 8 Marzo 2022, disponibile sul sito dello Studio Segale Stefanelli, <https://www.studiolegalestefanelli.it/it/approfondimenti/la-protezione-dei-dati-nel-mondo-una-breve-panoramica/>.
- PUTIN Evgeny et al. *Reinforced Adversarial Neural Computer for de Novo Molecular Design*, 2018, Journal of Chemical Information and Modeling, 58 (6), 1194-1204, <https://pubs.acs.org/doi/10.1021/acs.jcim.7b00690>.
- ROCK HEALTH, *Beyond wellness for the healthy: digital health consumer adoption*, 2018. https://rockhealth.com/reports/beyond-wellness-for-the-healthy-digital-health-consumer-adoption-2018/?mc_cid=0c97d69dbe&mc_eid=452e95c5c5.
- SANDLINE, *Data Privacy in Taiwan: PDPA, GDPR, and CCPA*, 14 giugno 2022, <https://sandlineglobal.com/the-pdpa-establishes-data-privacy-in-taiwan/>.

SARBAZ Yashar et al., *Introducing a decision support system for multiple sclerosis based on postural tremor: a hope for separation of people who might be affected by multiple sclerosis in the future*, *Biomedical Engineering: Applications, Basis and Communications*, vol. 29(06), 1750046 (2017), <https://doi.org/10.4015/S1016237217500466>.

SHANKAR Shreya et al. *No Classification without Representation: Assessing Geodiversity Issues in Open Data Sets for the Developing World*, *Proceedings of the 2020 conference on fairness, accountability, and transparency*, 2020, pp. 445–454, <https://doi.org/10.48550/arXiv.1711.08536>.

SMITH Helen, *Clinical AI: opacity, accountability, responsibility and liability*, *AI & SOCIETY*, 2021, vol. 36, pp. 535–545, <https://doi.org/10.1007/s00146-020-01019-6>.

SOLTANI Azam et al. *Improved algorithm for multiple sclerosis diagnosis in MRI using convolutional neural network*, *IET Image Processing*, 2020, vol. 14, 17, 1751-9659, <https://doi.org/10.1049/iet-ipr.2019.0366>.

TAIPEI MEDICAL UNIVERSITY, *Artificial Intelligence (AI) In Healthcare: Opportunities and Challenges*, *Participatory healthcare*, disponibile nel sito FUTURE LEARN, <https://www.futurelearn.com/courses/artificial-intelligence-for-healthcare-opportunities-and-challenges/6/steps/1698187>.

WHITBY Blay, *Automating Medicine the Ethical Way*, *Machine Medical Ethics*, *Intelligent Systems, Control and Automation: Science and Engineering*, 2015, vol 74, https://doi.org/10.1007/978-3-319-08108-3_14.

WILLIAMSON Steven M., PRYBUTOK Victor, *Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare*, *Applied Sciences*, 2024, vol. 14(2):675, <https://doi.org/10.3390/app14020675>.

YOO Youngjin et al. *Deep learning of brain lesion patterns and user-defined clinical and MRI features for predicting conversion to multiple sclerosis from clinically isolated syndrome*, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 2019, vol.7(9), pp. 250-259, <https://doi.org/10.1080/21681163.2017.1356750>.

RINGRAZIAMENTI

Un primo e più sentito ringraziamento va ai miei genitori, per avermi sostenuto in questo percorso, avermi dato fiducia e supportato sempre ed incondizionatamente.

La mia gratitudine va anche al mio caro nonno, grande giurista e grande uomo, a cui devo l'ambizione e il mio interesse al mondo del diritto.

Un ringraziamento particolare va al Prof. Andrea Pin, che è stato per me un esempio da seguire durante tutto il mio percorso di studi e che mi ha permesso di svolgere questa tesi, supportando la scelta dell'argomento sin dall'inizio e mantenendo sempre una disponibilità che ritengo ammirevole.

Vorrei anche ringraziare il Prof. Moro e la Prof.ssa Mingardo per aver creduto in me ed avermi ispirato durante il percorso di studi.

Vorrei inoltre ringraziare i miei compagni di Università, per aver reso questa un'esperienza unica e indimenticabile.

Un particolare ringraziamento va ai miei colleghi rappresentanti di corso, Francesco, Letizia e Matteo, con i quali abbiamo sempre mantenuto una costante armonia durante tutto il periodo di rappresentanza.

Infine, vorrei ringraziare i miei amici e famigliari, che mi hanno sempre supportato (e sopportato), specialmente durante i periodi più complicati, senza mai farmi mancare il loro affetto e la loro fiducia. Se ho raggiunto questo traguardo è anche grazie a voi.