

Università degli studi di Padova

Dipartimento di Diritto Pubblico, Internazionale e Comunitario

Corso di laurea triennale in Diritto e Tecnologia

Gli inverni dell'intelligenza artificiale

Relatore: Prof. Claudio Sarra

Laureando: Alessandro De Checchi

Matricola n.2002174

Anno accademico 2022-2023

Sommario

Abstract	3
Capitolo 1: Intelligenza Artificiale	4
1.1 La Definizione di Intelligenza	4
1.2 Definizione di Intelligenza Artificiale	5
1.3 Cenni storici, Alan Turing e prime applicazioni	6
1.4 I paradigmi dell'intelligenza artificiale	9
1.4.1 Machine learning	10
1.4.2 Deep Learning	12
1.4.3 Le Reti Neurali	13
Capitolo 2: Gli inverni dell'intelligenza artificiale	15
2.1 Gli eventi che hanno portato al primo inverno dell'IA	15
2.1.1 Il decennio buio	18
2.2 Il rapporto Lighthill e il primo inverno dell'IA	19
2.3 Il nuovo Boom	20
2.4 Il secondo inverno dell'IA	20
Capitolo 3: L'intelligenza artificiale ai giorni nostri	21
3.1 L'impatto sociale dell'IA	21
3.2 Settori e industrie coinvolte dall'IA	23
3.2.1 Industria automobilistica	24
3.2.2 Settore sanitario	25
3.2.3 Finanza	26
3.2.4 Istruzione	27
3.3 Sfide ed etica dell'IA	28
3.4 Possibili sviluppi futuri dell'IA	31
3.5 Criticità e possibile terzo inverno	33
3.5.1 Aspettative e promesse	33
3.5.2 L'opinione degli esperti	35
3.5.3 Il ruolo dei finanziamenti	35
3.5.4 Gli ultimi progressi	36
3.6 Conclusioni	38
Bibliografia e sitografia	40

Abstract

L'Intelligenza Artificiale (IA) è diventata un termine comune nel nostro linguaggio quotidiano, generando quesiti sulla sua definizione e sulle varie declinazioni in cui viene utilizzata. In questa tesi si intendono analizzare le cause che hanno portato a un rallentamento dell'innovazione tecnologica, noti come inverni dell'intelligenza artificiale, e se questi fattori possano ripresentarsi in futuro con le medesime criticità del passato.

Nel primo capitolo viene fornita una definizione del concetto d'intelligenza artificiale (IA) prendendo in considerazione la descrizione enunciata da John McCarthy alla conferenza di Dartmouth del 1956, per poi proporre un excursus storico sugli eventi che hanno segnato questa materia, primo fra tutti il celebre test di Turing, e delineare infine le caratteristiche e l'utilizzo delle applicazioni più recenti, come il machine learning, il deep learning e le reti neurali.

Nel secondo capitolo vengono analizzate le cause che hanno condotto ai primi due inverni dell'intelligenza artificiale (IA), provocati sostanzialmente dal deludente riscontro dei risultati ottenuti rispetto alle eccessive aspettative previste dai modelli teorici, che portarono a una drastica riduzione degli investimenti nel settore e al temporaneo raffreddamento dell'interesse verso le applicazioni tecnologiche dell'IA.

Nel capitolo conclusivo si analizza il crescente impatto con cui l'intelligenza artificiale (IA) è penetrata nella società attuale, sia dal punto di vista delle attività umane che svolgiamo quotidianamente, sia in relazione alle radicali trasformazioni economiche e sociali che ha prodotto in numerosi settori industriali, nel mondo della finanza e in campo sanitario, considerando al contempo le questioni etiche e giuridiche sollevate dall'utilizzo di queste tecnologie in relazione alla gestione e alla tutela dei dati. Viene infine presentata una panoramica sui recenti sviluppi dell'IA e del suo ciclo di hype, riportando i sorprendenti risultati di alcuni studi che dimostrano come alcune delle sue più popolari applicazioni stiano registrando dei cali di prestazioni, inducendo a considerare concretamente il pericolo di un imminente terzo inverno.

Capitolo 1: Intelligenza Artificiale

1.1 La Definizione di Intelligenza

L'intelligenza è un concetto complesso e sfaccettato. Gli studiosi hanno cercato di definirla in svariati modi, portando a numerose teorie; una definizione ampiamente accettata d'intelligenza si concentra sulla capacità di apprendimento, comprensione e adattamento all'ambiente circostante.

Quando si tenta di delineare il significato dell'intelligenza e il suo processo di sviluppo, è necessario considerare una molteplicità d'influenze interne ed esterne. La complessità di questi fattori rende la definizione dell'intelligenza un compito tutt'altro che agevole. È essenziale riconoscere che ciò che un individuo considera un atto d'intelligenza è strettamente legato ai processi di pensiero e alle competenze a cui attribuisce valore, pertanto il concetto d'intelligenza non può essere separato dal contesto sociale in cui opera.

Considerando i suoi limiti intrinseci, possiamo esaminare la definizione d'intelligenza della Treccani¹ che viene descritta come: *“Complesso di facoltà psichiche e mentali che consentono all'uomo di pensare, comprendere o spiegare i fatti o le azioni, elaborare modelli astratti della realtà, intendere e farsi intendere dagli altri, giudicare, e lo rendono insieme capace di adattarsi a situazioni nuove e di modificare la situazione stessa quando questa presenta ostacoli all'adattamento; propria dell'uomo, in cui si sviluppa gradualmente a partire dall'infanzia e in cui è accompagnata dalla consapevolezza e dall'autoconsapevolezza, è riconosciuta anche, entro certi limiti agli animali, specialmente mammiferi”*.

Più nello specifico, interviene la "teoria delle intelligenze multiple"², elaborata dal professor Howard Gardner, secondo cui l'intelligenza umana non può essere semplificata in una singola entità, ma abbraccia una varietà di dimensioni, ognuna con le sue specifiche caratteristiche e abilità. Gardner classifica sette tipologie d'intelligenza: linguistica, logico-matematica, spaziale, musicale, corporeo-

¹ <https://www.treccani.it/vocabolario/intelligenza/>

² Gardner, H. (1983). "Frames of Mind: The Theory of Multiple Intelligences."

cinestetica, intrapersonale e interpersonale. Questa prospettiva mette in evidenza la ricchezza e la diversità delle capacità umane, sottolineando che ognuno di noi eccelle in modi unici in queste identificazioni. L'influenza di Gardner è notevole nel campo dell'Intelligenza Artificiale poiché suggerisce che replicare l'intelligenza umana richiede la comprensione e l'incorporazione di queste varie sfaccettature.

1.2 Definizione di Intelligenza Artificiale

L'idea di sviluppare macchine in grado di emulare il pensiero umano ha radici profonde nella storia della tecnologia. Tuttavia, il termine "Intelligenza Artificiale" (IA) è stato ufficialmente coniato solo nel 1956, alla conferenza di Dartmouth³ organizzata da John McCarthy. Questo evento è universalmente considerato una pietra miliare nell'evoluzione dell'IA, segnando l'inizio di un'epoca di esplorazione e innovazione senza precedenti.

Tuttavia, definire con precisione cosa significhi "Intelligenza Artificiale" può essere complesso. Secondo la definizione contenuta nella Treccani⁴, l'IA è *"La disciplina che studia se e in che modo si possano riprodurre i processi mentali più complessi mediante l'uso di un computer. Tale ricerca si sviluppa secondo due percorsi complementari: da un lato l'Intelligenza Artificiale cerca di avvicinare il funzionamento dei computer alle capacità dell'intelligenza umana, dall'altro usa le simulazioni informatiche per fare ipotesi sui meccanismi utilizzati dalla mente umana."*

Un'altra definizione dell'IA è stata fornita, invece, dallo stesso John McCarthy nel 2007⁵: *È la scienza e l'ingegneria della creazione di macchine intelligenti e programmi per computer particolarmente intelligenti."*

L'acronimo "IA" è ormai parte integrante del nostro linguaggio, ma la sua definizione può variare notevolmente in base al punto di vista adottato. Esistono diverse espressioni per definire l'IA, che tuttavia convergono nell'obiettivo comune di sviluppare algoritmi in grado di eseguire compiti che

³ https://www.treccani.it/enciclopedia/nuove-prospettive-nell-intelligenza-artificiale_%28XXI-Secolo%29/

⁴ <https://www.treccani.it/enciclopedia/intelligenza-artificiale>

⁵ John McCarthy, "What is artificial intelligence?" Stanford University, 2007

normalmente richiedono l'intelligenza umana, come apprendere, ragionare, risolvere problemi, adattarsi all'ambiente e interagire in modo efficace con il mondo circostante.

Ad esempio, quando interagiamo con un assistente vocale formulando domande o impartendo comandi, l'IA analizza il nostro linguaggio, interpreta le nostre intenzioni e fornisce risposte appropriate⁶; ciò rappresenta uno degli innumerevoli esempi di come l'IA sia parte integrante della nostra vita quotidiana.

Un'ulteriore prospettiva sulla definizione di IA si basa sull'idea di considerare l'intelligenza come una collezione aperta di attributi che comprendono percezione, azione, ragionamento, adattamento, comunicazione, pianificazione, autonomia, creatività, riflessione, estetica e organizzazione⁷ questa prospettiva più ampia riflette l'idea che l'intelligenza non può essere ridotta a un singolo attributo, ma rappresenta una vasta gamma di abilità cognitive e intellettuali.

Tuttavia, è importante notare che non esiste una definizione univoca e universalmente accettata di IA.

1.3 Cenni storici, Alan Turing e prime applicazioni

L'Intelligenza Artificiale ha una profonda connessione con l'evoluzione dei computer, che iniziarono a diffondersi negli anni '40 grazie agli studi di McCulloch e Pitts, che nel 1943 svilupparono modelli matematici basati sui neuroni biologici per spiegare l'attivazione e l'inibizione degli stessi, operando in un sistema binario; questi modelli dimostrarono la capacità dei neuroni di apprendere e adattare le loro azioni nel tempo. Più tardi negli anni '50, Alan Turing, già famoso per aver decifrato il codice ENIGMA nazista durante la Seconda Guerra Mondiale, si interrogò sulla possibilità di far pensare una macchina e propose nell'articolo: "*Computing machinery and intelligence*", il celebre "Test di Turing"⁸. L'esperimento si basava sull'idea che per valutare l'intelligenza di un individuo si pongano delle domande, e proponeva di estendere tale modello alle macchine. Il test misurava quanto una

⁶ Russell, S. (2010). "Artificial Intelligence: A Modern Approach." Prentice

⁷ Honavar, V. (2006). "Artificial Intelligence and Neural Networks: Steps Toward Principled Integration." *Neural Networks*, 19(3), 367-387.

⁸ Turing, Alan M. *Computing machinery and intelligence*. Springer Netherlands, 2009.

macchina potesse essere indistinguibile da un essere umano in una conversazione, coinvolgendo un interrogatore, una tastiera e un monitor divisi in due parti per mostrare le risposte del computer e dell'essere umano; dopo cinque minuti di conversazione su un argomento scelto dall'interrogatore, si doveva determinare se si fosse conversato con una macchina e se il computer era riuscito a ingannare il 30% degli interrogatori, la prova si riteneva superata. Il test non dava un risultato perfetto, poiché poteva fallire se uno dei due interlocutori umani raggiungeva una prestazione inferiore al 30%, ma il test di Turing era teso a certificare solo la capacità di simulare il pensiero umano, senza affrontare la questione della consapevolezza del computer. Turing prevedeva che entro una cinquantina d'anni, quindi intorno agli anni 2000, una macchina avrebbe superato il test. Nel 1956, Herb Simon e Allen Newell⁹ costruirono la prima macchina in grado di risolvere autonomamente problemi matematici di base e in quello stesso anno John McCarthy alla conferenza di Dartmouth coniò il termine "Artificial Intelligence" per descrivere macchine in grado di ragionare e compiere operazioni che richiedevano intelligenza umana, distinguendo il campo dall'approccio della cibernetica con l'obiettivo di estendere l'uso dei computer oltre la mera elaborazione dati. Questi anni videro un crescente ottimismo e i primi dibattiti sulla capacità dell'Intelligenza Artificiale di eguagliare l'intelligenza umana, tuttavia, molte sfide teoriche e computazionali si presentarono nel tentativo di emulare il funzionamento del cervello umano.

Nel 1966, Joseph Weizenbaum¹⁰ programmò ELIZA, considerato il primo chatbot, che forniva risposte predefinite in base alle domande poste e riusciva a farsi scambiare per uno psicologo da diverse persone, nonostante le sue limitazioni.

Purtroppo, negli anni '70 la ricerca nell'ambito dell'Intelligenza Artificiale (IA) non raggiunse i risultati sperati, e questo portò a un brusco declino dell'interesse e alla conseguente diminuzione dei

⁹ <https://www.ibm.com/topics/artificial-intelligence>.

¹⁰ Joseph Weizenbaum, *Computer power and human reason*, Londra, Penguin Books, 1984.

finanziamenti. Le principali sfide erano legate alla limitata potenza di calcolo e alle capacità limitate dei computer dell'epoca, oltre scarsità di dati disponibili.

John Searle, ad esempio, presentò il concetto della "stanza cinese" per illustrare che le macchine non possono essere considerate pensanti, poiché non comprendono veramente i simboli che elaborano. Egli argomentò che, se una persona, senza conoscere il cinese, seguisse istruzioni simili a quelle di un computer, otterrebbe lo stesso risultato, ma non si potrebbe dire che comprende ciò che ha prodotto. Di conseguenza, Searle concluse affermando che i computer in realtà non comprendevano effettivamente il loro elaborato e dunque non si può parlare d'intelligenza ma di simulazione di comportamenti intelligenti, privi di comprensione vera e propria.

Nonostante le sfide, in quegli anni furono studiate varie metodologie per migliorare i programmi e le soluzioni al fine di avvicinarsi alle prestazioni umane. Negli anni '80, l'interesse per l'IA conobbe una rinascita, con un focus più pratico che teorico, soprattutto nell'applicazione di sistemi esperti in settori specifici.

L'avvento della robotica ebbe un impatto rilevante sull'IA, consentendo di creare un'intelligenza "reale" in grado di percepire l'ambiente circostante tramite il processo di "*embodiment*", ossia una sorta di incarnazione dell'intelligenza artificiale, sempre più capace di percepire la realtà.

In alcuni ambiti, l'IA riuscì a superare le capacità umane. Geoffrey Hinton e Yann LeCun intensificarono lo studio delle reti neurali, mentre nel 1997, Deep Blue, sviluppata da IBM, sconfisse il campione di scacchi Garry Kasparov, dimostrando che le prestazioni superiori dell'IA erano in gran parte dovute alla potenza computazionale disponibile.

Nel 2016, Alpha Go di Google sconfisse il campione cinese di Go, Lee Sedol, imparando il gioco attraverso l'autoapprendimento.

Con l'aumento dei modelli linguistici di grandi dimensioni, o LLM, come ChatGPT, queste nuove pratiche d'intelligenza artificiale generativa e i modelli di deep learning possono essere preaddestrati su grandi quantità di dati grezzi e non etichettati.

La connettività wireless ha svolto un ruolo chiave nell'evoluzione dell'IA, consentendo ai computer di operare come un unico "cervello globale" con vasta distribuzione e connettività, offrendo un vantaggio rispetto all'intelligenza umana.

L'andamento storico della ricerca in IA è caratterizzato da periodi di ottimismo e significativi finanziamenti, alternati a periodi di minor interesse e finanziamenti ridotti, portando nel complesso a un trend di crescita costante.

1.4 I paradigmi dell'intelligenza artificiale

Per quanto complesso, abbiamo compreso, dalle definizioni date finora, che l'intelligenza artificiale, o IA, è una materia interdisciplinare che si propone di sviluppare sistemi informatici in grado di eseguire attività che normalmente richiedono la capacità cognitiva umana. Nel corso della sua evoluzione, l'IA ha sviluppato diversi paradigmi e approcci per affrontare questa sfida, e due di questi approcci principali sono noti come l'approccio simbolico e l'approccio sub-simbolico.

Nell'approccio simbolico all'IA¹¹, la conoscenza è rappresentata attraverso un insieme definito di simboli, il cui significato è sintetizzato dall'uomo che ne cura la codifica mediante un linguaggio formale, e su cui è possibile operare con metodi di natura logica-insiemistica; questo approccio ha portato a risultati significativi, come sistemi di ragionamento esperti e sistemi di pianificazione.

D'altra parte, l'approccio sub-simbolico all'IA¹² si concentra sull'uso di modelli matematici e algoritmi per affrontare i compiti intelligenti; la rappresentazione della conoscenza è veicolata da una collezione ordinata di variabili numeriche. Questo paradigma trae ispirazione dai processi biologici e dalle reti neurali, cercando di emulare le funzioni cerebrali attraverso l'uso di strumenti

¹¹ https://www.treccani.it/magazine/lingua_italiana/speciali/IA/02_Bacciu.html

¹² https://www.treccani.it/enciclopedia/intelligenza-artificiale_%28Dizionario-di-Medicina%29/

computazionali. Gli approcci sub-simbolici includono il machine learning, le reti neurali artificiali e l'elaborazione del linguaggio naturale, e sono stati fondamentali per molti sviluppi recenti nell'IA, come il riconoscimento facciale e la traduzione automatica.

Entrambi questi paradigmi hanno i loro vantaggi e limiti, e spesso vengono combinati in approcci ibridi per affrontare sfide complesse nell'IA contemporanea. La scelta tra l'approccio simbolico e sub-simbolico dipende spesso dal tipo di problema da risolvere e dalle risorse disponibili. In ultima analisi, la continua evoluzione dell'IA è guidata dalla ricerca e dalla sperimentazione in entrambi questi paradigmi, poiché gli scienziati cercano di sviluppare sistemi sempre più intelligenti e versatili.

1.4.1 Machine learning

Come già citato, il Machine Learning è uno dei più famosi paradigmi dell'intelligenza artificiale e uno dei più studiati e utilizzati nelle nuove tecnologie. Una definizione più specifica è stata elaborata nel 1959 da Herbert Simon¹³: *“Il Machine Learning è quella branca dell'informatica che permette a una macchina di imparare ad eseguire un compito senza essere esplicitamente programmata per farlo.”* In altre parole, questa definizione implica che il Machine Learning consiste nell'allenare una macchina a imparare da dati iniziali in modo che possa gestire elementi che non erano stati originariamente previsti dal programmatore, come se la macchina apprendesse in modo simile a un bambino che acquisisce conoscenza, osservando e migliorando continuamente sulla base di ciò che ha imparato.

Fu in seguito lo scienziato informatico Tom M. Mitchell¹⁴ nel 1997, a proporre una definizione più dettagliata del Machine Learning: *"Si può affermare che un programma impara attraverso l'esperienza E in riferimento a categorie specifiche di compiti T, utilizzando una misurazione delle prestazioni P, se le sue abilità nel compito T, misurate da P, migliorano con l'esperienza E."*

¹³ Simon, Herbert A. "Why should machines learn?" *Machine learning*. Morgan Kaufmann, 1983. 25-37.

¹⁴ Mitchell, Tom M. "Machine learning." (1997).

Questa definizione ha stabilito un solido fondamento per comprendere il Machine Learning come un processo in cui un programma migliora le sue capacità in un compito specifico grazie all'esperienza acquisita.

In generale, il Machine Learning permette di rappresentare la realtà mediante funzioni matematiche, anche se l'algoritmo non conosce queste funzioni in anticipo. L'algoritmo attiva a tali funzioni matematiche dopo aver elaborato i dati disponibili. L'obiettivo dell'apprendimento nel Machine Learning è far corrispondere gli input ai risultati desiderati, invece di cercare una comprensione concettuale approfondita. Il processo di addestramento consente all'algoritmo di imparare come associare gli input a output specifici. Il processo di funzionamento delle macchine che utilizzano il Machine Learning inizia con la preparazione e l'inserimento dei dati, un passaggio critico per il successo dell'algoritmo. È fondamentale selezionare adeguatamente i dati pertinenti per ottenere il miglior risultato possibile e massimizzare l'efficacia.

Successivamente, si procede all'addestramento della macchina, ossia la fase in cui l'algoritmo apprende da esempi di input e output, creando una funzione che collega i dati di input ai risultati desiderati. Questa fase implica un feedback che consente all'algoritmo di migliorarsi nel tempo.

I classificatori di Machine Learning sviluppano le loro abilità cognitive attraverso una formula matematica che consente di distinguere tra diverse classi basate sulle caratteristiche dei dati. La rappresentazione delle informazioni è essenziale per il funzionamento dell'algoritmo e coinvolge la mappatura delle caratteristiche degli input in modo che possano essere riconosciute in futuro.

L'ottimizzazione dell'algoritmo coinvolge la ricerca di diverse combinazioni di parametri per trovare la mappatura ottimale delle caratteristiche e delle classi. Una volta creato il classificatore, l'algoritmo è pronto per elaborare classificazioni basate sulla funzione obiettivo.

Tuttavia, il processo per far funzionare un algoritmo di Machine Learning può richiedere tempo e supervisione da parte del programmatore per garantire il corretto svolgimento delle fasi. Non sempre

è possibile creare un algoritmo di successo, quindi è cruciale avere una conoscenza completa di ogni singola fase del processo per aumentare le possibilità di successo.

1.4.2 Deep Learning

Il deep learning¹⁵ rappresenta una branca del machine learning, caratterizzata da reti neurali composte da tre o più livelli che cercano di emulare in parte il funzionamento del cervello umano. Questa tecnologia consente ai sistemi di apprendere da un vasto insieme di dati, anche se è ben lontana dall'eguagliare la complessità del cervello umano. Le reti neurali profonde svolgono un ruolo fondamentale in numerose applicazioni dell'intelligenza artificiale, automatizzando compiti analitici e fisici, come riconoscimento vocale e assistenti digitali. Queste reti cercano di riconoscere, classificare e descrivere oggetti all'interno dei dati, lavorando attraverso una serie di livelli interconnessi che incrementano via via la precisione.

Tuttavia, il deep learning si differenzia dal machine learning tradizionale per la gestione dei dati e il processo di apprendimento. Mentre il machine learning richiede spesso dati strutturati ed etichettati, il deep learning è in grado di affrontare dati non strutturati come testi e immagini senza una preelaborazione intensiva. Gli algoritmi di deep learning possono estrarre autonomamente le caratteristiche principali dai dati, riducendo la necessità d'intervento umano. L'apprendimento del modello avviene attraverso processi come la discesa del gradiente e la backpropagation, che consentono all'algoritmo di migliorare la sua precisione nel tempo.

Inoltre, sia il machine learning sia il deep learning possono essere classificati in categorie di apprendimento, come l'apprendimento controllato (con dati etichettati), l'apprendimento non controllato (senza etichette) e l'apprendimento a punteggio di prestazioni (ottimizzato per il rendimento). Entrambi gli approcci sono potenti ma si distinguono per la gestione di dati non strutturati e la capacità di estrazione automatica delle caratteristiche chiave dai dati.

¹⁵ <https://www.ibm.com/topics/deep-learning>

1.4.3 Le Reti Neurali

Le reti neurali¹⁶, conosciute anche come ANN (artificial neural network) o SNN (simulated neural network), costituiscono una parte del machine learning e sono l'elemento centrale dei metodi di deep learning. La loro ispirazione deriva dal cervello umano, emulando la trasmissione di segnali tra i neuroni biologici.

Le ANN (artificial neural network) sono composte da strati di nodi che includono un livello di input, uno o più livelli nascosti e un livello di output. Ciascun nodo, chiamato neurone artificiale, è collegato ad altri nodi e ha associati un peso e una soglia. Se l'output di un nodo supera la soglia, attiva il nodo successivo nella rete. In caso contrario, non trasmette dati al livello successivo.

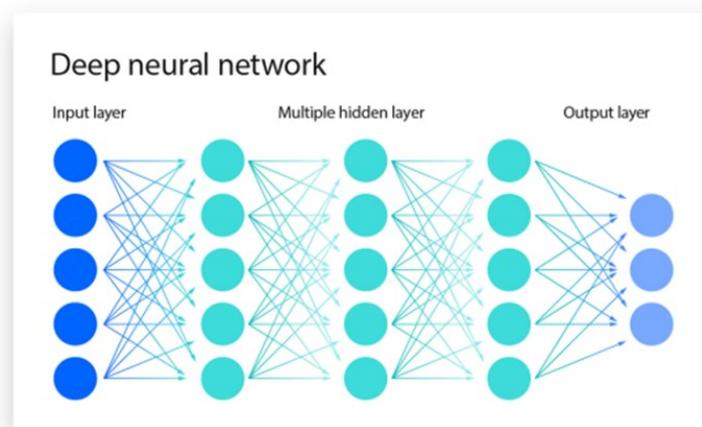


Immagine presa dal Sito Web di IBM

Queste reti neurali utilizzano dati di addestramento per migliorare la loro precisione nel tempo. Una volta ottimizzate per la precisione, diventano potenti strumenti nella computer science e nell'intelligenza artificiale, permettendo di classificare e organizzare dati in modo rapido. Attività come il riconoscimento vocale o il riconoscimento delle immagini richiedono solo pochi minuti invece delle ore necessarie per una classificazione manuale da parte di esperti umani.

¹⁶ <https://www.ibm.com/it-it/topics/neural-networks#:~:text=alla%20fase%20successiva-,Cos'%C3%A8%20una%20rete%20neurale%3F,degli%20algoritmi%20di%20deep%20learning.>

Le reti neurali possono essere suddivise in vari tipi, ciascuno adatto a scopi specifici. Qui di seguito ci sono alcune delle reti neurali più comuni:

- Il perceptrone, la rete neurale più antica, sviluppata da Frank Rosenblatt nel 1958.
- Le reti neurali feedforward o perceptron multilivello (MLP), sono composte da un livello di input, uno o più livelli nascosti e un livello di output. Sono ampiamente utilizzate in visione artificiale, elaborazione del linguaggio naturale e altre applicazioni neurali.
- Le reti neurali convolutive (CNN), simili alle reti feedforward, ma principalmente utilizzate per il riconoscimento delle immagini, la scoperta di modelli e la visione artificiale. Queste reti si basano sull'algebra lineare, in particolare la moltiplicazione di matrici, per identificare modelli all'interno d'immagini.
- Le reti neurali ricorrenti (RNN), caratterizzate da loop di feedback. Vengono utilizzate soprattutto per analizzare dati temporali e fare previsioni sul futuro, come previsioni di mercato azionario o previsioni di vendita.

In conclusione, è importante distinguere tra deep learning e reti neurali di base, poiché il termine "profondo" in deep learning si riferisce unicamente alla profondità dei livelli all'interno di una rete neurale. Quando una rete neurale comprende più di tre livelli, compresi quelli di input e output, può essere considerata un algoritmo di deep learning. Al contrario, una rete neurale con solo due o tre livelli è considerata una rete neurale di base. Questa distinzione è fondamentale poiché il deep learning ha dimostrato un notevole potenziale nell'ambito dell'intelligenza artificiale, alimentando applicazioni come il riconoscimento d'immagini, il riconoscimento vocale e molti altri campi. La crescente profondità delle reti neurali ha contribuito a migliorare significativamente la capacità di queste reti di apprendere da dati complessi e di svolgere compiti molto articolati, aprendo nuove frontiere nell'AI.

Capitolo 2: Gli inverni dell'intelligenza artificiale

2.1 Gli eventi che hanno portato al primo inverno dell'IA

Negli anni '50 e '60 del XX secolo, l'Intelligenza Artificiale (IA) emerse come una prospettiva affascinante e ambiziosa per il futuro. Gli scienziati e i ricercatori, tra cui spiccavano le figure di Alan Turing e John von Neumann¹⁷, abbracciarono con entusiasmo l'idea di creare macchine intelligenti in grado di emulare il pensiero umano. Nonostante le prime macchine fossero rudimentali rispetto agli standard odierni, il loro potenziale suscitò un entusiasmo palpabile tra i pionieri che le svilupparono; l'idea di creare macchine in grado di risolvere problemi complessi e di apprendere autonomamente catturò l'immaginazione di tutti.

Una delle prime applicazioni pratiche dell'IA in questo periodo pionieristico fu nel 1951 il Ferranti Mark I, un computer sviluppato presso l'Università di Manchester da Freddie Williams e Tom Kilburn e poi assemblato da Ferranti¹⁸. Fu il primo computer di uso generale a essere messo in commercio, sul quale era stato sviluppato un programma per giocare a scacchi che riuscì, incredibilmente per l'epoca, a sconfiggere un giocatore professionista; un risultato che ispirò entusiasmo tra gli sviluppatori, aprendo prospettive che fecero diventare l'IA un campo di ricerca accademica in espansione, con l'obiettivo di creare software capaci di emulare il processo di pensiero umano, con la prospettiva che le macchine potessero superare gli esseri umani in compiti intellettuali.



Immagine del Ferranti Mark I

¹⁷ J. Neumann, "General and Logical Theory of Automata," in *Cerebral Mechanisms in Behavior: The Hixon Symposium*, 1951

¹⁸ Lavington, Simon Hugh. *A history of Manchester computers*. NCC Publications, 1975.

Un altro catalizzatore di questo entusiasmo iniziale fu l'idea di interagire con i computer utilizzando il linguaggio naturale, per cui i ricercatori si impegnarono nello sviluppo di linguaggi di programmazione più intuitivi e di sistemi di elaborazione del linguaggio naturale, per rendere i computer più accessibili e favorire un'interazione più fluida tra uomini e macchine. Uno dei progressi più significativi fu caratterizzato dai primi tentativi di traduzione automatica¹⁹. Nel 1954 vennero eseguiti degli esperimenti iniziali di traduzione automatica utilizzando un dizionario di 250 parole e un'analisi sintattica che, sebbene fosse una traduzione letteraria parola per parola, generò notevole interesse e fu considerata estremamente utile per la traduzione di documenti tecnici da parte soprattutto dagli enti di difesa statunitensi (ONR e ARPA, successivamente chiamato DARPA) che fecero grandi investimenti nella ricerca e nello sviluppo di queste nuove tecnologie nella speranza di poterle utilizzare durante la guerra fredda per riuscire a compiere una traduzione automatica dei documenti dal russo all'inglese²⁰.

Gli esperimenti iniziali servirono da ispirazione per creare il Dartmouth Summer Project nel 1956, dove il termine "IA" fu coniato, ossia un progetto estivo che ebbe come motto: *"Ogni aspetto dell'apprendimento o qualsiasi altra caratteristica dell'intelligenza può in principio essere descritto così precisamente che una macchina possa essere fatta per simulare"*²¹. Furono invitati ricercatori di discipline diverse e vennero presentate molte idee, articoli e concetti tra i più vari, ma nonostante i progressi fatti grazie durante il workshop alcuni non rimasero soddisfatti, tra i quali lo stesso McCarthy che affermò: *"Il motivo principale per cui il Workshop non ha soddisfatto le mie aspettative è che l'IA è più difficile di quanto pensassimo."*

¹⁹ J. Hutchins, "The first public demonstration of machine translation: the Georgetown-IBM system, 7th January 1954", MT News International, 2006.

²⁰ Hutchins, Encyclopedia of Language & Linguistics, Second Edition. Oxford: Elsevier, 2006

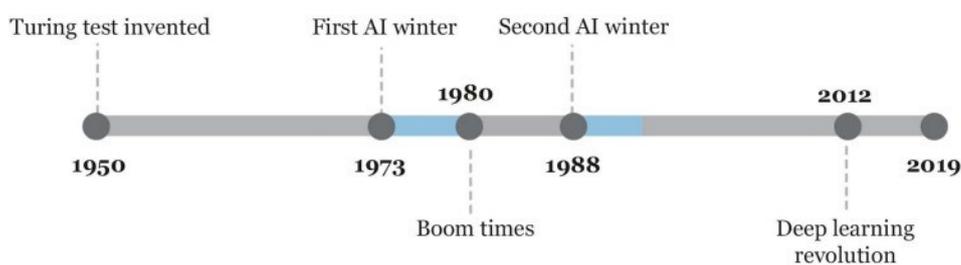
²¹ J. McCarthy, M. L. Minsky, and C. E. Shannon, "A proposal for the Dartmouth summer research project on artificial intelligence - August 31, 1955", Ai Magazine, vol. 27, no. 4, pp. 12-14, 2006

L'anno successivo, al Dartmouth Summer Project, Rosenblatt inventò *i perceptron*²², un primo esperimento tipo di rete neurale, in cui unità neurali binarie erano collegate tramite pesi regolabili. Fu ispirato dal lavoro di neuroscienziato svolto negli anni '40, che lo portò a creare una replica rudimentale dei neuroni del cervello, sperimentando layout e algoritmi di apprendimento diversi²³.

I computer dell'epoca erano troppo lenti per eseguire il perceptrone, quindi Rosenblatt costruì una macchina speciale con resistenze regolabili (potenziometri) controllate da piccoli motori. Il risultato fu sensazionale: il macchinario riuscì ad imparare a classificare diverse immagini di forme o lettere.

Nel medesimo anno, Herbert Simon riassunse il progresso dell'IA con queste parole: *"Nel mondo ci sono ora macchine che pensano, imparano e creano. Inoltre, la loro capacità di fare queste cose aumenterà rapidamente fino a quando, in un futuro visibile, il range di problemi che possono gestire sarà coestensivo con il range a cui la mente umana è stata applicata"*²⁴.

Il periodo tra gli anni '50 e '60 rappresentò un momento di straordinario entusiasmo e crescita nell'ambito dell'IA, caratterizzato da scoperte, sperimentazioni e audaci ambizioni. I primi successi e il fervore dei ricercatori aiutarono l'afflusso di consistenti finanziamenti pubblici e privati che contribuirono a gettare le basi per lo sviluppo futuro dell'Intelligenza Artificiale.



²² F. Rosenblatt, *The Perceptron a perceiving and recognizing automaton*, 1957.

²³ F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychological Review*, 1958

²⁴ P. Norvig, *Paradigms of Artificial Intelligence Programming: Case Studies in Common Lisp*. 1992, p. 962

2.1.1 Il decennio buio

In seguito all'entusiasmo iniziale e al conseguente aumento dei finanziamenti legati alla grande aspettativa che si nutriva verso la traduzione automatica negli anni '50 e nei primi anni '60, ci fu un forte rallentamento al progresso, Hutchins definì questo periodo che va dal 1967 al 1976 come "*the quiet decade of machine traslation*"²⁵. Bar-Hillel²⁶ aggiunse che i computer avrebbero dovuto elaborare troppe informazioni per eseguire una traduzione corretta, impossibile da ottenere per la limitata potenza di calcolo delle macchine. Nel 1964, il Comitato Consultivo per l'Elaborazione Automatica del Linguaggio confermò la tesi di Bar-Hillel affermando che nel breve periodo non ci sarebbe stata alcuna macchina capace di produrre una traduzione automatica corretta; a questo si aggiunse un rapporto nel 1966 che poneva fine alla sperimentazione di questa tecnologia, con il conseguente taglio ai finanziamenti da parte dell'Establishment della Difesa degli Stati Uniti su tutti i progetti di traduzione automatica e, più in generale, sulle applicazioni dell'intelligenza artificiale. Il decennio buio della traduzione automatica fu uno degli eventi che maggiormente caratterizzarono l'origine alla prima "era glaciale" dell'IA. Nel 1969, è stato pubblicato il libro "*Perceptrons*"²⁷ di Minsky e Papert che presentava una critica severa ai perceptrons di Rosenblatt, dimostrando che i perceptrons potevano essere addestrati solo per risolvere problemi linearmente separabili. Fu un duro colpo per i connessionisti, i quali credevano che l'IA potesse essere ottimizzata imitando la struttura del cervello. Minsky e Papert sapevano che più livelli avrebbero potuto risolvere questo problema, ma non c'era un algoritmo capace di addestrare una tale rete; ci vollero diciassette anni prima che tale algoritmo fosse ideato e solo successivamente si scoprì che l'algoritmo di retro-propagazione, ossia quello capace di addestrare la rete, era stato inventato prima della pubblicazione di "*Perceptrons*"²⁸.

²⁵ J. Hutchins, "Machine Translation: A Brief History," 1995.

²⁶ Y. Bar-Hillel, "The Present Status of Automatic Translation of Languages", *Advances in Computers*, 1960

²⁷ M. M. Papert and S., "Perceptrons. An Introduction to Computational Geometry.", 1969

²⁸ J. Schmidhuber, "Deep Learning in neural networks", 2015

2.2 Il rapporto Lighthill e il primo inverno dell'IA

Nel 1973 fu pubblicato il rapporto Lighthill²⁹, una valutazione dello stato di fatto dell'IA commissionato dal Consiglio di Ricerca Scientifica Britannico che giunse alla conclusione che le promesse fatte dai ricercatori coinvolti nella sperimentazione di AI erano ingigantite, sostenendo che: “in nessuna parte del campo le scoperte finora effettuate hanno prodotto l'impatto principale che era stato promesso all'epoca”.

Nonostante fosse emerso che l'area di ricerca più deludente fosse stata la traduzione automatica, “...dove sono stati spesi enormi somme con risultati molto scarsi...” James Lighthill, riteneva che il fallimento principale fosse quello di non superare “l'esposizione combinatoria”, riferendosi a un problema ben noto negli spazi di ricerca, come il numero di nodi che aumenta in modo esponenziale procedendo nell'albero. Analogamente, Shannon³⁰ dimostrò come nel gioco degli scacchi il numero di possibili partite aumenta da 20 con la prima mossa a 400 con la seconda mossa, mentre alla quinta ci sono già 4.865.609 possibilità, rappresentando quindi un'esplosione combinatoria.

Il rapporto Lighthill fu oggetto all'epoca di molte critiche e commenti molto ben argomentati, ma ebbe comunque un effetto rilevante, tanto che in seguito alle valutazioni da lui fatte, il governo britannico tagliò i finanziamenti per tutte le università, tranne le due coinvolte nella ricerca in questo campo, innescando una reazione che si diffuse in tutta Europa, con un impatto evidente anche negli Stati Uniti³¹.

In conclusione, la prima "era glaciale" dell'IA fu generata da un concorso di diverse circostanze che portarono dapprima un crescente entusiasmo riguardo alle potenzialità di questo nuovo campo con abbondante cassa di risonanza mediatica, cui seguirono le delusioni riscontrate nella traduzione automatica, che determinarono un'era di stagnazione, le critiche sollevate da Minsky e Papert che rallentarono il progresso dei perceptrons e, infine, il rapporto Lighthill che, dopo la sua pubblicazione

²⁹ I. Literature and R. Lighthill, “Lighthill Report - Artificial Intelligence: A General Survey”, pp. 1–18, 1972.

³⁰ C. E. Shannon, “Programming a Computer for Playing Chess”, Philosophical Magazine, 1950

³¹ J. Hendler, “Avoiding Another AI Winter”, IEEE Intelligent Systems, 2008.

nel 1973, ebbe un impatto catastrofico sui finanziamenti e sul progresso della ricerca dando così inizio al primo inverno dell'intelligenza artificiale.

2.3 Il nuovo Boom

Con il passare degli anni gli effetti negativi del primo inverno dell'intelligenza artificiale cominciarono a diminuire e l'IA ebbe una nuova primavera con l'avvento dei prodotti commerciali legati a questa tecnologia ancora in fase embrionale, dopo lo stallo durato circa un decennio. Nei primi anni '80, grandi conferenze come AAAI (*Association for the Advancement of Artificial Intelligence*) iniziarono a prendere piede e a riscuotere grande successo, tanto da portare molti settori dell'industria e i dipartimenti governativi a rinnovare il loro interesse nella tecnologia dell'IA.

Al centro del rinnovato interesse riguardo l'intelligenza artificiale (IA) c'erano i “*sistemi esperti*”³², programmi composti da regole nella forma “*if-then*” (se condizione, allora azione). Questo approccio, noto come “top-down”, era considerato con convinzione la via migliore per realizzare l'IA. Infatti, molti settori applicarono questa nuova tecnologia come per esempio la pianificazione finanziaria, la diagnosi medica, l'esplorazione geologica e la progettazione di circuiti microelettronici. L'entusiasmo tornò a crescere per la prima volta dal primo inverno e con esso anche la paura di un secondo fallimento dato dalle eccessive attese. In un panel chiamato “*The Dark Ages of AI - Can we avoid or survive them?*”³³ alla conferenza AAAI del 1984, gli scienziati discussero se un imminente inverno dell'IA potesse essere evitato. Il timore principale era che i finanziamenti si esaurissero nuovamente quando le aspettative irrealistiche non potevano essere soddisfatte; purtroppo, da lì a poco quel timore si sarebbe materializzato portando al secondo inverno dell'intelligenza artificiale.

2.4 Il secondo inverno dell'IA

Negli anni successivi le affermazioni sulle capacità e le potenzialità dei sistemi d'intelligenza artificiale hanno lentamente dovuto confrontarsi con la realtà. I sistemi esperti al centro della

³² G. Johnson, “Thinking About Thinking”, APF Reporter, vol. 8, no. 1, pp. 1–6, 1985.

³³ D. Mcdermott, M. M. Waldrop, R. Schank, B. Chandrasekaran, and J. Mcder “The Dark Ages of AI: A Panel Discussion at AAAI-84”, AI Magazine, vol. 6, no. 3, pp. 122–134, 1985.

rivoluzione hanno affrontato molte problematiche. Nel 1984, John McCarthy criticò i sistemi esperti perché mancavano di buon senso e conoscenza delle proprie limitazioni³⁴. Analizzò il sistema esperto MYCIN, progettato per assistere i medici, illustrando una situazione in cui un paziente aveva il vibrione del colera nei suoi intestini. Quando venne posto il quesito al sistema, esso prescriveva due settimane di tetraciclina, terapia che avrebbe probabilmente eliminato tutti i batteri, ma che avrebbe portato anche alla morte del paziente. Inoltre, molte attività erano troppo complesse per essere progettate manualmente dagli ingegneri attraverso la definizione di regole; i sistemi per la visione e il linguaggio, ad esempio, contenevano troppe eccezioni³⁵. Schwarz, direttore dell'Information Science and Technology Office della DARPA (Defense Advanced Research Projects Agency) dal 1987 al 1989, concluse che la ricerca sull'IA ha sempre avuto *"... successi molto limitati in aree specifiche, seguiti immediatamente dal fallimento nel raggiungere l'obiettivo più ampio a cui questi successi iniziali sembrano alludere..."*. Ciò portò a una diminuzione dei finanziamenti nella ricerca sull'IA³⁶. L'interesse generale per l'IA diminuì ancora una volta perché le aspettative non potevano essere soddisfatte. In questo periodo, molte aziende che operavano nel campo dell'intelligenza artificiale chiusero i battenti. La conferenza AAAI, che attirava oltre 6000 visitatori nel 1986, diminuì rapidamente a soli 2000 nel 1991³⁷, segnando l'inizio del secondo inverno dell'IA.

Capitolo 3: L'intelligenza artificiale ai giorni nostri

3.1 L'impatto sociale dell'IA

L'Intelligenza Artificiale (IA) sta rivoluzionando il panorama sociale in modi che vanno ben oltre il mero avanzamento tecnologico. L'impatto sociale dell'IA è una realtà sempre più tangibile, plasmando dinamiche quotidiane, relazioni interpersonali, e strutture socioeconomiche. Questo fenomeno suscita un ampio dibattito su come la società debba affrontare le sfide e sfruttare le opportunità derivanti da questa accelerata trasformazione tecnologica, uno degli aspetti più evidenti è il cambiamento nel

³⁴ J. McCarthy, "Some Expert System Need Common Sense", pp. 1–11, 1984.

³⁵ K. Hao, "We analyzed 16,625 papers to figure out where AI is headed next", MIT Technology Review, pp. 1–19, 2019.

³⁶ N. J. Nilsson, *The Quest for Artificial Intelligence*, 2010.

³⁷ T. Menzies, "21st-Century AI: Proud, not smug", 2003.

panorama occupazionale, l'automatizzazione alimentata dall'IA ha il potenziale di ridisegnare il mercato del lavoro, con la possibilità di rendere obsolete alcune professioni e, al contempo, di creare nuove opportunità sollevando interrogativi cruciali sulla necessità di riconvertire le competenze professionali e garantire una transizione equa per coloro che potrebbero essere colpiti negativamente dalla disoccupazione tecnologica.

Parallelamente, l'IA contribuisce in modo significativo a migliorare l'efficienza nei settori della sanità, dell'istruzione e della gestione delle risorse; sistemi diagnostici basati sull'IA possono fornire diagnosi più precise, mentre piattaforme di apprendimento automatico personalizzato stanno trasformando l'approccio all'istruzione.

Tuttavia, sorge la preoccupazione per la possibile creazione di divari digitali, con gruppi sociali svantaggiati che potrebbero non beneficiare appieno di queste tecnologie innovative. Nel suo libro "Homo Deus: Breve storia del futuro"³⁸, lo storico Yuval Noah Harari esplora il possibile impatto dell'IA sulla struttura sociale, affermando: *"Il ventunesimo secolo potrebbe vedere l'ascesa di nuove caste di super-umani"*. Harari suggerisce che l'IA potrebbe creare divisioni sociali più accentuate, dando vita a una società stratificata in base all'accesso e alla competenza nell'utilizzo delle tecnologie intelligenti. A questo si aggiunge l'influenza che l'IA ha sulla sfera delle relazioni umane, con chatbot e assistenti virtuali che diventano sempre più comuni; a tale riguardo, la psicologa e sociologa Sherry Turkle, nel suo libro "Alone Together: Why We Expect More from Technology and Less from Each Other"³⁹, esamina l'effetto dell'IA sulle relazioni interpersonali, sostenendo: *"Siamo soli insieme. Non possiamo tirarci indietro da ciò che ci aspettiamo da tecnologie sempre più onnipresenti"*. Questa citazione riflette sul modo in cui l'IA può influenzare la dinamica sociale, portando a una maggiore dipendenza dalle tecnologie e, paradossalmente, a una maggiore solitudine. Se da un lato le tecnologie basate sull'intelligenza artificiale (IA) possono semplificare la vita quotidiana, dall'altro sollevano

³⁸ *Homo Deus – A brief history of tomorrow*; Autore: Yuval Noah Harari; Data di pubblicazione: 2017; Editore: Penguin.

³⁹ Arnd-Caddigan, Margaret. "Sherry Turkle: Alone Together: Why We Expect More from Technology and Less from Each Other: Basic Books, New York, 2011, 348 pp" (2015)

questioni etiche riguardanti la privacy e la sicurezza dei dati personali, pone interrogativi sull'equità e sulla trasparenza degli algoritmi utilizzati e sull'automazione delle decisioni, come quelle legate al credito o all'impiego.

L'Intelligenza Artificiale sta dunque plasmando in modo significativo il tessuto sociale, introducendo cambiamenti che vanno oltre il semplice progresso tecnologico, migliorando l'efficienza e aprendo nuove prospettive, ma la rapida automazione suscita anche preoccupazioni legate alla trasformazione del mercato del lavoro, alla possibile creazione di divari digitali e all'accentuarsi delle disuguaglianze sociali.

3.2 Settori e industrie coinvolte dall'IA

L'intelligenza artificiale (IA) sta progressivamente e prepotentemente condizionando diversi settori economici e industriali, rivoluzionando processi, migliorando l'efficienza e aprendo nuove opportunità, ma anche introducendo forti cambiamenti nel mondo del lavoro e, in alcuni casi, sostituendosi all'uomo nello svolgimento di molti lavori nei quali la figura umana fino a pochi anni fa era ritenuta insostituibile, mentre ora il lavoratore viene sempre più impiegato in compiti di supervisione del funzionamento delle macchine.

In "The Fourth Industrial Revolution" di Klaus Schwab⁴⁰, il fondatore del World Economic Forum esamina come l'IA stia trasformando il panorama industriale. *"Abbiamo appena iniziato a vedere come la tecnologia sta influenzando il nostro mondo economico e sociale. Non possiamo prevedere tutte le conseguenze, ma una cosa è chiara: il paradigma dell'industria sta cambiando in modo significativo."*

Questo fenomeno sta indubbiamente cambiando il mondo del lavoro portando all'estinzione di alcune professioni ma anche alla creazione di nuove specializzazioni nate per soddisfare le esigenze del mercato del lavoro che richiede competenze diverse in grado di coniugare i lavori tradizionali con le tecnologie moderne.

⁴⁰ Schwab, Klaus. *The fourth industrial revolution*. Currency, 2017.

3.2.1 Industria automobilistica

Il settore automobilistico è stato uno dei precursori nell'adozione dell'intelligenza artificiale (IA) per potenziare la progettazione, la sicurezza e l'esperienza di guida. Secondo Klaus Schwab, l'IA sta effettivamente rivoluzionando l'industria automobilistica, contribuendo a plasmare veicoli più sicuri, efficienti e interconnessi.

Le applicazioni dell'IA nel campo automobilistico sono diverse e in costante evoluzione. Sistemi avanzati di assistenza alla guida (ADAS), come il controllo automatico della velocità, il mantenimento della corsia e la frenata automatica di emergenza, stanno ristrutturando il panorama della sicurezza stradale e migliorando l'esperienza di guida. Le moderne vetture utilizzano l'IA per ottimizzare i consumi, potenziare le prestazioni e personalizzare l'esperienza degli utenti. Un esempio concreto di questa trasformazione è rappresentato da Tesla, che ha implementato con successo la guida autonoma assistita, dimostrando come l'IA sia in grado di rivoluzionare il concetto stesso di guida e di influenzare il panorama del trasporto su strada. Elon Musk, CEO di Tesla, ha dichiarato che nel lungo termine nessuno sarà propenso ad acquistare un'auto che non sia autonoma, evidenziando il cambiamento radicale che l'IA apporterà al modo in cui concepiamo la proprietà e l'uso dei veicoli.

Le applicazioni dell'IA nell'ambito automobilistico abbracciano inoltre piattaforme di connettività automatica, gestione intelligente dei viaggi, soluzioni di cloud computing per l'analisi dei big data e, soprattutto, la guida autonoma. Quest'ultima rappresenta una delle applicazioni più attese dell'IA nel settore automobilistico, con la SAE che ha delineato cinque livelli di guida autonoma, riflettendo l'aumento progressivo dell'autonomia dei veicoli rispetto alla guida umana.

Inoltre, l'applicazione dell'IA nell'industria automobilistica si estende alle piattaforme intelligenti che facilitano il collegamento automatico tra guidatori e servizi nelle vicinanze, così come sistemi di gestione del traffico che sfruttano dati avanzati per ottimizzare gli spostamenti.

Tuttavia, l'introduzione dell'IA nella guida autonoma solleva interrogativi etici e di sicurezza, rendendo essenziale l'implementazione di regolamentazioni adeguate e valutazioni approfondite delle implicazioni sociali ed economiche.

3.2.2 Settore sanitario

Il settore della medicina è al centro di una profonda trasformazione grazie all'adozione dell'intelligenza artificiale (IA). Secondo l'analisi di Wolf e Sholze la medicina al pari dell'economia ha vissute le proprie rivoluzioni ed è quindi possibile scomporre la sua storia in epoche sulla base dei progressi tecnologici. Il primo periodo, quello della medicina 1.0 si è perpetuato per migliaia di anni, nei quali i medici usavano come strumenti al proprio servizio soltanto i cinque sensi per eseguire diagnosi e i medicinali derivati da elementi naturali. La seconda età, corrispondente a un'epoca recente, inizia con la scoperta degli antibiotici e l'uso della radiografia a raggi x per diagnosticare certe patologie. La medicina 3.0 è sostanzialmente quella degli ultimi decenni, in cui fanno il loro ingresso nella quotidianità clinica la miniaturizzazione, l'elettronica, la diagnostica per immagini e la robotica. L'intelligenza artificiale apre le porte alla medicina 4.0⁴¹, nella quale la capacità di apprendimento e di problem solving delle macchine, la profilazione e categorizzazione di enormi quantità di files clinici e la possibilità di trasmissione di informazioni in tempo reale tra macchine e medici sta inesorabilmente modificando il modo di fare medicina.

Secondo lo studio "*Artificial Intelligence and life in 2030*"⁴² pubblicato dall'Università di Stanford il settore sanitario è uno degli otto settori in cui l'impatto dell'intelligenza artificiale sarà maggiormente rilevante, portando cambiamenti significativi al settore, automatizzando attività cliniche e amministrative, introducendo nuove modalità di fruizione dei servizi come la telemedicina, implementando la medicina personalizzata usando algoritmi avanzati che analizzano immagini mediche con precisione, facilitando la diagnosi precoce di patologie e le decisioni del personale medico, come ad esempio i sistemi esperti ormai ampiamente utilizzati a supporto delle diagnosi e

⁴¹ B.WOLF, C. SCHOLZE, "Medicine 4.0", Biomedical Engineering 2017; p. 183–186

⁴² Artificial Intelligence and life in 2030 – One Hundred Year Study on Artificial Intelligence, Stanford University, 2016

delle cure. Un'altra applicazione di IA nel settore sanitario è IBM Watson che utilizza il machine learning per analizzare vasti dataset medici e fornire raccomandazioni personalizzate. Non meno importante si rivela anche l'uso della robotica medica utilizzata nella sfera della chirurgia di precisione, nell'assistenza e nella riabilitazione.

3.2.3 Finanza

L'intelligenza artificiale (IA) ha svolto un ruolo cruciale nella trasformazione del settore finanziario, introducendo innovazioni sostanziali in ambiti quali la gestione del rischio, l'analisi dei dati e la prevenzione delle frodi. Questa adozione ha contribuito a rendere il settore più efficiente, competitivo e orientato al cliente.

L'economista Klaus Schwab, noto per il suo libro "The Fourth Industrial Revolution", evidenzia che nel settore finanziario, l'IA sta ridefinendo le dinamiche di mercato, accelerando l'analisi dei dati e introducendo nuovi modelli operativi, mentre Thomas H. Davenport, autore di "The AI Advantage: How to Put the Artificial Intelligence Revolution to Work"⁴³, afferma che l'IA ha consentito una personalizzazione senza precedenti nei servizi finanziari.

Un esempio tangibile di questa trasformazione si riscontra nell'utilizzo dell'IA nelle piattaforme di trading online, dove algoritmi avanzati analizzano in tempo reale ingenti quantità di dati finanziari per identificare pattern e opportunità di investimento.

Gli algoritmi di machine learning giocano un ruolo fondamentale nella gestione del rischio finanziario, consentendo alle istituzioni di analizzare rapidamente i dati e identificare potenziali rischi. Melanie Mitchell, autrice di "Artificial Intelligence: A Guide for Thinking Humans"⁴⁴, sottolinea come l'IA favorisca una gestione più proattiva e mirata.

Nel servizio clienti, l'introduzione delle chatbots alimentati da intelligenza artificiale, permette di rispondere autonomamente alle domande degli utenti, risolvere problemi comuni e fornire supporto continuo 24/7. Le aziende utilizzano l'IA per automatizzare le interazioni con i clienti, migliorando

⁴³ Davenport, Thomas H. *The AI advantage: How to put the artificial intelligence revolution to work*. mit Press, 2018.

⁴⁴ Mitchell, Melanie. *Artificial intelligence: A guide for thinking humans*. Penguin UK, 2019.

l'efficienza e garantendo una risposta rapida alle esigenze degli utenti, come ad esempio l'utilizzo di chatbot su piattaforme di e-commerce per guidare gli utenti nel processo di acquisto, fornire informazioni sui prodotti e gestire eventuali problemi post-vendita.

Un'altra rilevante applicazione dell'IA in questo settore è stata la prevenzione delle frodi finanziarie, attraverso algoritmi di rilevamento delle anomalie capaci di analizzare i modelli di transazione e identificare comportamenti sospetti, contribuendo in modo significativo a proteggere i clienti e le istituzioni finanziarie da attività illecite.

3.2.4 Istruzione

Negli ultimi anni, l'Intelligenza Artificiale (IA) ha visto numerose sperimentazioni internazionali nel campo educativo, con due principali modalità di integrazione: l'IA come argomento di studio per sviluppare conoscenze e consapevolezza, e l'IA come strumento di analisi e miglioramento del processo di apprendimento.

La seconda modalità ha ricevuto maggiore attenzione, con rare sperimentazioni sui percorsi educativi dedicati all'IA. Tuttavia, l'importanza di lavorare su questo tema è evidenziata da tre necessità educative, come proposto da Lucking⁴⁵: la capacità di discutere sugli aspetti etici dell'IA, la comprensione delle basi del suo funzionamento per un utilizzo consapevole, e la capacità di progettare sistemi AI, considerando gli studenti come futuri progettisti di algoritmi

Paesi come la Cina e la Finlandia hanno investito in programmi educativi sull'IA a favore di studenti con età a partire da undici anni, mentre in Italia sono state condotte alcune sperimentazioni senza un coordinamento ministeriale e senza un piano nazionale di sviluppo e formazione per docenti e studenti.

Nel campo dell'IA come strumento di analisi del processo di apprendimento, sono state sviluppate diverse categorie di strumenti. I Sistemi di Tutoraggio Intelligente forniscono un tutoraggio

⁴⁵ LUCKIN Rosemary, *Machine Learning and Human Intelligence: The Future of Education for the 21st Century*, London, UCL IOE Press - University of London - Institute of Education 2018.

individualizzato basato sulle risposte degli studenti, come nel caso di MATHia⁴⁶. I Sistemi di Tutoraggio Basati su Dialogo interagiscono con gli studenti attraverso la conversazione, mentre gli Ambienti di Apprendimento Esplorativo promuovono l'esplorazione libera per costruire la conoscenza, come nel gioco ECHOES.

Tuttavia, è importante considerare gli aspetti etici e di equità nell'uso dell'IA in educazione, come evidenziato dai casi di ingiustizie generati da algoritmi di valutazione basati su dati sbilanciati. La ricerca continua e l'attenzione all'equità sono essenziali per garantire che l'IA contribuisca in modo positivo all'istruzione.

3.3 Sfide ed etica dell'IA

Dalla sua creazione nella metà del secolo scorso, l'intelligenza artificiale ha suscitato ampie riflessioni filosofiche sulla sua natura e sul suo rapporto con l'intelligenza umana. Recentemente, è emersa la necessità di una riflessione etica sull'autonomia delle macchine, considerando la loro capacità di compiere azioni autonome con implicazioni morali. Il rapido progresso tecnologico degli ultimi decenni ha complicato il rapporto tra tecnologia e diritto, mettendo in discussione l'efficacia delle norme giuridiche esistenti.

Le crescenti autonomia e capacità cognitive delle macchine sollevano l'interrogativo su sé considerare i robot come soggetti aventi diritti o creare una nuova categoria chiamata "personalità elettronica"⁴⁷. Nella società moderna, permeata dal progresso, gli strumenti derivati dalla tecnologia non solo consentono di adattare la vita umana alle proprie necessità, ma offrono anche l'opportunità di studiarla. L'etica diventa un elemento cruciale quando si esplorano concetti indefiniti come l'intelligenza artificiale, cercando modalità etiche per giustificarne l'esistenza.

⁴⁶ RITTER S. - CARLSON R. - SANDBOTHE M. - FANCSALI S. E., Carnegie Learning's adaptive learning products, in SANTOS O. et alii (eds.), Educational Data Mining 2015, 8th International Conference on Educational Data Mining (EDM2015), Madrid

⁴⁷ DE SIMONE L., BURGIO E., "Intelligenza Artificiale e responsabilità civile" in MediaLaws, Law and Policy of the Media in a Comparative Perspective, (2021)

L'integrazione dell'etica nelle discussioni sull'intelligenza artificiale mira a comprendere le possibilità offerte da essa e a giustificarne l'esistenza in modi diversi. In questo contesto, l'etica si concentra sull'algoritmo piuttosto che sulla persona, spingendo a percepire la tecnologia come un'entità in grado di anticipare e soddisfare i bisogni umani. La formulazione di principi etici diventa fondamentale per valutare l'innovazione e garantire che l'IA segua uno schema valoriale allineato con i valori condivisi dalle strutture sociali e istituzionali umane, evitando il rischio di accrescere le disuguaglianze sociali. Tra gli impatti dell'uso dell'intelligenza artificiale che, invece, possono essere considerati non etici, rientrano gli algoritmi per il riconoscimento facciale, la sorveglianza, la localizzazione, e la ricerca di armi intelligenti in grado di colpire autonomamente il bersaglio. Inoltre, strumenti utilizzati per manipolare la realtà e influenzare il consenso umano, come i generatori di fake news, possono essere inclusi in questo elenco. Se il riconoscimento facciale, la sorveglianza e gli strumenti di localizzazione possono avere un'utilità "dual use," le ultime menzionate potrebbero essere considerate come un "totally unethical use" (utilizzo totalmente non etico).

Nei casi di uso non etico dell'IA, è essenziale considerare anche la questione delle decisioni non etiche che l'IA potrebbe prendere. Queste decisioni possono derivare dalla mancanza di strategie adeguate da parte degli umani o dalla presenza di pregiudizi "*bias*" sviluppati durante il processo di apprendimento dell'IA, sia indotti dai dati proposti che emergono autonomamente. Questo problema diventa particolarmente critico dato l'aumento dell'uso di tali algoritmi, poiché decisioni non etiche possono comportare gravi conseguenze.

Ad esempio, negli Stati Uniti, l'uso di software di riconoscimento facciale da parte delle forze dell'ordine ha portato a casi di errato riconoscimento e arresti ingiusti. Un altro esempio riguarda un software utilizzato per prevedere la probabilità di recidiva, il quale presentava un *bias* razziale, indicando una maggiore probabilità di recidiva per individui non di razza bianca. Questi casi

evidenziano la necessità di eliminare i pregiudizi nei dati, attraverso un lavoro di pulizia del dataset per renderlo neutrale dal punto di vista del *bias*⁴⁸.

Tuttavia, si pone la questione se l'eliminazione dei pregiudizi dai dati stessi possa costituire un pregiudizio. L'intervento umano dovrebbe includere riferimenti etici come ragionevolezza, sicurezza, responsabilità, equità e tutela della privacy per garantire che le decisioni dell'IA siano simili a quelle di un essere umano razionale e responsabile.

Virginia Dignum⁴⁹, docente di informatica presso l'Università di Umea e docente associata presso l'Università di Tecnologia di Delft, oltre che membro del gruppo di ricerca sull'intelligenza artificiale sociale ed etica, identifica tre livelli di interazione tra l'intelligenza artificiale (IA) ed etica:

1. **Etica by Design:** Questo livello implica l'integrazione tecnico-algoritmica delle capacità di ragionamento etico nel comportamento dell'IA. Si basa su un approccio che utilizza strumenti formali per descrivere una situazione e principi etici per consentire all'IA di formulare autonomamente giudizi etici e giustificare le decisioni prese.
2. **Etica in Design:** Si riferisce ai metodi ingegneristici e normativi che supportano la valutazione e l'analisi delle implicazioni etiche dei sistemi di IA quando integrano o sostituiscono le strutture sociali tradizionali. Questo livello enfatizza la trasparenza, la responsabilità e l'equità degli algoritmi che influenzano la nostra vita quotidiana.
3. **Etica per Design:** Coinvolge la definizione di standard, la progettazione di codici di condotta e processi di certificazione per garantire l'integrità di sviluppatori e utenti nell'intero ciclo di vita dei sistemi di IA. I progettisti devono seguire chiare linee guida etiche, consentendo al sistema di prendere decisioni in linea con questi principi e avvertendo gli utenti quando si discosta da tali valori.

⁴⁸ Facial recognition: is the technology taking away your identity? The Guardian febbraio 2020

⁴⁹ Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20(1), 1-3.

Dignum sottolinea l'importanza del principio di precauzione, che richiede la supervisione umana continua delle macchine. Questo principio implica l'associazione delle capacità di apprendimento dell'IA a "driver etici" umani per evitare discriminazioni e assicurare che le decisioni siano vantaggiose per l'umanità. L'essere umano dovrebbe poter intervenire in qualsiasi momento tramite un "botone di emergenza" insormontabile dalla macchina, nel caso in cui si verifichi una violazione dei principi etici predefiniti.

Il focus dell'attenzione di Dignum si concentra su tre settori specifici: machine learning, medicina digitale e veicoli autonomi, considerati rilevanti nel contesto dell'eticità dei sistemi di IA.

3.4 Possibili sviluppi futuri dell'IA

L'Intelligenza Artificiale (IA) sta attraversando un rapido processo di evoluzione, destinato a lasciare un'impronta significativa nel nostro mondo. I futuri sviluppi nell'IA prevedono una trasformazione profonda in diverse aree della vita quotidiana e in tutti i settori dell'economia e dell'industria.

In primo luogo, ci attendiamo un progresso nell'implementazione della personalizzazione e dell'adattamento dei sistemi basati sull'IA. Grazie ad algoritmi sempre più sofisticati, l'IA sarà in grado di comprendere meglio le preferenze individuali, offrendo esperienze su misura. Questo potrebbe tradursi in servizi personalizzati, consigli di prodotti mirati e interazioni digitali più intuitive.

Un altro elemento chiave sarà lo sviluppo di sistemi d'intelligenza artificiale in grado di apprendere in modo continuo, adattandosi all'evolversi delle informazioni e delle circostanze. Ciò consentirà alle IA di affrontare scenari complessi, rimanendo rilevanti nel tempo e migliorando costantemente le loro prestazioni.

La collaborazione sempre più frequente tra intelligenze artificiali e umane è destinata a diventare la norma. L'IA potrebbe essere integrata nei processi decisionali umani, offrendo supporto e

suggerimenti basati sull'analisi di dati complessi. Questa sinergia potrebbe portare a soluzioni più efficaci e tempestive in settori come la medicina, la finanza e la ricerca scientifica.

Un ulteriore settore in cui si prevedono sviluppi significativi è quello dell'IA etica. Con l'ubiquità crescente delle tecnologie IA, c'è una crescente consapevolezza dell'importanza di garantire uno sviluppo e un utilizzo etico. Possiamo pertanto aspettarci progressi nella progettazione di algoritmi che rispettino principi morali e norme sociali, mitigando potenziali rischi e preoccupazioni etiche legate all'IA.

Infine, l'applicazione dell'IA potrebbe rivoluzionare settori cruciali come l'energia, l'ambiente e la sostenibilità. L'analisi avanzata dei dati potrebbe contribuire a ottimizzare l'uso delle risorse, migliorare l'efficienza energetica e sviluppare soluzioni innovative per affrontare sfide ambientali.

In quest'ottica, le espressioni utilizzate per descrivere queste nuove tecnologie riflettono la necessità di adattarsi a un panorama in continua evoluzione. Come evidenziato nel passato, l'adozione di nuovi termini e concetti è spesso influenzata dalla familiarità e dall'eredità concettuale. Tuttavia, prevedere il futuro dell'IA è complesso, poiché i vari settori si evolvono con ritmi diversi, dipendenze reciproche e incertezze legate a concetti come la singolarità tecnologica.

La prospettiva della singolarità tecnologica, introdotta da Vernor Vinge⁵⁰ nel 1993, solleva preoccupazioni riguardo a un futuro in cui l'intelligenza artificiale supererà quella umana, portando a cambiamenti imprevedibili. Tuttavia, le stime sulla data di possibile singolarità variano, e alcuni sostengono che dovremmo adottare il principio della precauzione per evitare potenziali effetti negativi.

⁵⁰ Vinge, Vernor. "Technological singularity." *VISION-21 Symposium sponsored by NASA Lewis Research Center and the Ohio Aerospace Institute*. 1993.

Ray Kurzweil⁵¹, al contrario, vede la fusione tra intelligenza artificiale e pensiero umano come una tendenza positiva, prevedendo che entro il 2045 potremmo raggiungere la singolarità. Alcuni transumanisti considerano la fusione tra umani e macchine come un'evoluzione necessaria, mentre altri sono più cauti e sottolineano la necessità di regolamentazioni etiche.

Secondo Jerry Kaplan⁵² queste visioni non sono in linea con la connessione esistente tra l'attuale tecnologia e le previsioni più pessimistiche. Non ci sono prove che gli attuali sviluppi dell'Intelligenza Artificiale rappresentino i primi segnali di macchine onniscenti super intelligenti. È più probabile che l'Intelligenza Artificiale sia utilizzata per automatizzare la maggior parte dei processi che richiedono l'intervento umano, seguendo un trend ormai consolidato. Kaplan tende a sottolineare che, riguardo al controllo dell'Intelligenza Artificiale, sia i più ottimisti, sia i più pessimisti non abbiano necessariamente torto, ma che le previsioni prospettate siano lontane dalla realtà odierna. Tuttavia, ritiene che queste posizioni possano essere utili per alimentare discussioni su argomenti di fondamentale importanza. Secondo l'opinione di Kaplan, se si dovesse verificare la singolarità, questa sarà preceduta da numerosi segnali che consentiranno di prendere adeguate contromisure per evitarla o limitarne gli effetti.

In conclusione, mentre il futuro dell'IA promette trasformazioni significative, è essenziale affrontare le sfide etiche e sociali correlate a questa evoluzione. Solo un approccio responsabile e consapevole alle nuove tecnologie IA garantirà un impatto positivo e sostenibile sulla nostra società.

3.5 Criticità e possibile terzo inverno

3.5.1 Aspettative e promesse

Molti personaggi pubblici esprimono affermazioni sull'IA che ricordano quelle dei primi ricercatori negli anni '50, generando entusiasmo per i progressi futuri, o hype. Kurzweil, ad esempio, è famoso non solo per aver predetto la singolarità, un momento in cui l'intelligenza artificiale superiore

⁵¹ Kurzweil, Ray. "The singularity is near." *Ethics and emerging technologies*. London: Palgrave Macmillan UK, 2005. 393-406.

⁵² Kaplan Jerry. *Le persone non servono: lavoro e ricchezza nell'epoca dell'Intelligenza Artificiale*. Roma: Luiss, 2016.

diventerà ubiqua entro il 2045⁵³, ma anche che l'IA supererà l'intelligenza umana entro il 2029.

Analogamente, Scott predice che *"non c'è motivo e nessun modo per cui una mente umana possa tenere il passo con una macchina intelligente artificiale entro il 2035"*⁵⁴, mentre Ng considera l'IA come la nuova elettricità⁵⁵. Dichiarazioni di questo tipo tendono a creare aspettative piuttosto elevate sull'IA generando hype.

Uno strumento spesso utilizzato nell'analisi dell'hype è il "Gartner's Hype Cycle"⁵⁶, che attraverso applicazioni pratiche consente di fare previsioni, benché la sua validità non sia scientificamente provata, in quanto, non si tratta di uno strumento sviluppato per un approccio scientifico ma di un grafico stilizzato creato per le decisioni aziendali. Detto ciò, sono stati effettuati tentativi di convalidare empiricamente il ciclo di hype per diverse tecnologie, giungendo alla conclusione che nonostante sia innegabile la sua esistenza i suoi modelli specifici sono molto variabili.

Come dimostra Menzies, il ciclo dell'hype è ben rappresentato nei numeri dei partecipanti alla conferenza AAAI negli anni '80⁵⁷. iniziata con un rapido aumento delle vendite dei biglietti che portarono a un picco, che preludeva a una rapida diminuzione di quei numeri.

Attualmente, i numeri dei partecipanti a conferenze come NIPS raggiungono o superano il picco di AAAI degli anni '80 e si presentano in continuo e rapido aumento, accrescendo l'interesse per il settore e gli investimenti da parte delle imprese, che favorisce il proliferare di aziende nel settore dell'intelligenza artificiale. Se seguiamo l'andamento del ciclo di hype, dovremmo attenderci una prossima diminuzione nella partecipazione alle conferenze sull'IA, che condurrà a una nuova fase di stallo, ma l'intelligenza artificiale è un campo piuttosto ampio e caratterizzato da tecnologie

⁵³ R. Kurzweil, *The Singularity is Near - When Human Transcends Biology*. 2005 p. 135-136.

⁵⁴ A. Brown, *11 Quotes About AI That'll Make You Think*, 2019.

⁵⁵ S. Lynch, Andrew Ng: *Why AI Is the New Electricity*, 2017.

⁵⁶ A. Linden and J. Fenn, "Understanding Gartner's Hype Cycles Understanding Gartner's Hype Cycles Understanding Gartner's Hype Cycles", 2003.

⁵⁷ T. Menzies, "21st-Century AI: Proud, not smug", 2003

diverse, per cui sarebbe necessario “calcolare” un ciclo di hype per ogni tecnologia che rientra nell’IA.

Di seguito, si propone un’analisi più approfondita di tali affermazioni, valutando se il rapido aumento d’interesse verso l’IA è solo la conseguenza di promesse esagerate o se le affermazioni sono fondate.

3.5.2 L’opinione degli esperti

Il confronto tra le dichiarazioni pubbliche e le previsioni degli esperti sull’Intelligenza Artificiale (IA) rivela una discrepanza significativa. Un sondaggio condotto nel 2017 intervistando ricercatori di machine learning indica che un’IA di alto livello potrebbe manifestarsi entro 45 anni e con una probabilità del 50%, ma l’automazione completa del lavoro è prevista in un futuro più lontano, entro i prossimi 122 anni e con una probabilità del 50%, risultati che sono in contrasto con le previsioni più ottimistiche dei futuristi. Uno studio Meta evidenzia che la maggior parte delle previsioni sull’IA di alto livello si colloca in un arco di tempo di circa vent’anni, indipendentemente dal periodo in cui vengono fatte. La previsione dell’IA è dunque soggetta a notevole incertezza, con il rischio di un calo dell’interesse a breve termine dovuto al modello enunciato dal Ciclo dell’Hype⁵⁸

3.5.3 Il ruolo dei finanziamenti

Il finanziamento ha sempre avuto un ruolo significativo nella ricerca sull’IA. Come sottolinea Hendler, i tagli dei finanziamenti governativi si avvertono solo dopo diversi anni, poiché i programmi di ricerca in corso proseguono⁵⁹.

Nell’aprile 2018, i membri dell’Unione Europea hanno concordato di collaborare nella ricerca sull’Intelligenza Artificiale, stanziando tra il 2018 e il 2020 la cifra di 1,7 miliardi di dollari per la ricerca. A giugno 2018, la Commissione europea ha proposto il programma di finanziamento Digital Europe, con un totale di 9,2 miliardi di euro, di cui 2,5 destinati alla ricerca sull’IA.

Contemporaneamente, nel marzo 2018 l’amministrazione degli Stati Uniti ha dichiarato l’obiettivo

⁵⁸ S. Armstrong and K. Sotola, “How We’re Predicting AI – or Failing to”, pp. 11–29, 2014.

⁵⁹ Hendler, “Avoiding Another AI Winter”, IEEE Intelligent Systems, 2008.

di voler mantenere il primato globale nell'IA e nel settembre 2018 la DARPA ha stanziato 2 miliardi di dollari per finanziare nuove tecnologie dell'IA. Anche la Cina ha lanciato un progetto di sviluppo nel campo dell'IA entro il 2030, promuovendo diverse iniziative. Tali dichiarazioni hanno contribuito ad avvalorare il concetto di "Corsa all'IA", una competizione che potrebbe limitare o impedire la riduzione dei finanziamenti al settore, anche se già in passato, durante il primo inverno, si è assistito a tagli di fondi da parte dei governi, come avvenne nel caso dei traduttori automatici dal russo all'inglese, finanziati dagli Stati Uniti e rilevatisi poi un insuccesso che portò con sé l'arresto allo sviluppo di questa tecnologia.

3.5.4 Gli ultimi progressi

Per valutare le critiche alle attuali tecniche di intelligenza artificiale, è necessario esaminare il progresso compiuto dal 2012 fino a oggi e degli elementi chiave che emergono nella formazione di tale scenario. Risultati impressionanti sono stati ottenuti nell'ambito della comprensione del linguaggio naturale. L'apprendimento profondo ha reso possibile lo sviluppo di tutti gli assistenti vocali più popolari, da Alexa e Siri a Cortana⁶⁰. L'apprendimento con l'ausilio di reti neurali profonde ha ottenuto risultati importanti nel gioco; nel 2014 DeepMind ha utilizzato un apprenditore *deep-q* per risolvere 50 diversi giochi Atari, senza modificare l'architettura o gli iperparametri del modello⁶¹. Questa flessibilità nello svolgimento delle attività non aveva precedenti, solleticando l'interesse del colosso informatico Google che li ha presto acquisiti per sviluppare progetti di apprendimento per rinforzo come AlphaGo⁶² e AlphaStar⁶³. Negli ultimi anni, le reti neurali generative (GAN) hanno ottenuto risultati impressionanti nella generazione di

⁶⁰ J. Schmidhuber, Deep Learning in neural networks, 2015.

⁶¹ X. Guo, H. Lee, X. Wang, and R. L. Lewis, "Deep learning for real-time Atari game play using offline Monte Carlo tree search planning", Nips, 2014.

⁶² D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of Go with deep neural networks and tree search.", Nature, vol. 529, no. 7587, pp. 484–489, 2016.

⁶³ AlphaStar: Mastering the Real-Time Strategy Game StarCraft II, 2019

immagini, in special modo dei volti umani⁶⁴, portando l'apprendimento profondo a raggiungere risultati rivoluzionari in molte attività.

Per quanto riguarda la storia recente, emerge l'esperienza dell'azienda OpenAI con l'introduzione di chat GPT (Generative Pre-trained Transformer), un esempio d'intelligenza artificiale basata su modelli di linguaggio. Nello specifico, GPT appartiene alla categoria delle reti neurali trasformative, che rappresenta una forma di architettura di reti neurali profonde.

La caratteristica principale di GPT è la sua capacità di generare testo in modo coerente e contestualmente rilevante. La tecnologia GPT è stata preaddestrata su grandi quantità di testo proveniente da internet, acquisendo così una vasta conoscenza linguistica. Questo preaddestramento consente a GPT di comprendere il contesto e generare risposte coerenti alle domande o agli input forniti.

GPT si basa su un modello di transformer, un'architettura di rete neurale che si è dimostrata particolarmente efficace nel trattamento del linguaggio naturale. L'aspetto "generativo" di GPT si riferisce alla sua capacità di concepire nuovi pezzi di testo in modo autonomo, anziché rispondere semplicemente a input specifici.

Questa tecnologia ha presentato tuttavia alcune lacune; un recente studio condotto da ricercatori della Stanford University e dell'Università della California, Berkeley pubblicata su arXiv (quindi in attesa di revisione da parte della comunità scientifica), ha infatti esaminato l'evoluzione nel tempo delle prestazioni della tecnologia d'intelligenza artificiale, focalizzandosi sui modelli di linguaggio di grandi dimensioni (LLM) di OpenAI, in particolare GPT 3.5 e GPT 4. Denominato "*How is ChatGPT's behavior changing over time?*". Lo studio ha testato le versioni di marzo e giugno 2023 su diverse attività, come la risoluzione di problemi matematici, la risposta a domande sensibili, la generazione di codici e il ragionamento visivo, con il risultato che la capacità di GPT 4

⁶⁴ T. Karras, "Progressive Growing of Gans for improved Quality, Stability, and Variation", pp. 1–26, 2018.

d'identificare la risposta corretta in un sondaggio a risposta multipla eseguito con metodologia "OpinionQA" (dataset di sondaggi contenente 1506 domande di carattere generale) è drasticamente diminuita dal 97,6% a marzo al solo 22,1% a giugno, mentre il test eseguito su GPT 3.5 ha mostrato prestazioni migliorative di circa il 2% (da 94,3% a 96,7%)⁶⁵. Questa ricerca ha suscitato preoccupazioni, in linea con le lamentele degli utenti sulle prestazioni della chatbot. Per spiegare le motivazioni di questo calo di efficienza sono stati ipotizzati il tentativo di risparmiare energia computazionale per accelerare le risposte del software o l'utilizzo di una strategia aziendale per far pagare agli utenti alcune funzionalità aggiuntive. Peter Welinde, vicepresidente di OpenAI, ha affermato che GPT-4 non è stato reso più "stupido", ma che i problemi diventano evidenti con un uso più intensivo. Alcuni esperti, come Arvind Narayanan dell'Università di Princeton, ritengono che i risultati dello studio non dimostrino in modo definitivo un calo delle prestazioni di GPT 4 e che potrebbero essere coerenti con gli aggiustamenti apportati da OpenAI.

3.6 Conclusioni

In conclusione, l'analisi delle affermazioni pubbliche e delle prospettive indicate degli esperti sull'intelligenza artificiale (IA) rivelano una discrepanza significativa, laddove molte dichiarazioni pubbliche tendono a riflettere un ottimismo eccessivo e una propensione esagerata all'hype, rispetto a una visione più prudente che tenga conto dei limiti e delle criticità dell'attuale tecnologia. Figure pubbliche come Kurzweil e Scott prospettano scenari futuri molto ambiziosi, spingendo l'entusiasmo per l'IA a livelli elevati. La rappresentazione del Ciclo dell'Hype, evidenziata attraverso l'esempio del "Gartner's Hype Cycle" e le analogie con i numeri dei partecipanti alle conferenze sull'IA, suggerisce che il settore potrebbe trovarsi in una fase di elevato interesse, con il rischio di un calo nell'entusiasmo nel breve termine. La recente evoluzione della tecnologia, come evidenziato dal caso di Chat GPT, mostra che nonostante i progressi significativi, possono emergere lacune e cambiamenti delle prestazioni nel tempo. L'analisi empirica sull'andamento del ciclo di

⁶⁵ Lingjiao Chen , Matei Zaharia , James Zou, *"How is ChatGPT's behavior changing over time?"*, Stanford University, UC Berkeley, p.10, 2023.

hype, unita alle sfide evidenziate da studi come quello condotto dai ricercatori della Stanford University e dell'Università della California, Berkeley, indica che l'ottimismo e le aspettative eccessive possono portare a una potenziale delusione quando le tecnologie non mantengono completamente le promesse fatte. Il rischio di un terzo "inverno" dell'intelligenza artificiale è tangibile, soprattutto se i governi e i finanziatori privati continuano a sostenere aspettative irrealistiche e a interpretare le dichiarazioni degli esperti come certezze piuttosto che come prospettive. È dunque essenziale mantenere un approccio realistico verso lo sviluppo dell'IA, bilanciando l'entusiasmo per le potenzialità con una comprensione critica delle sfide e delle limitazioni attuali delle tecnologie. In tal modo, si può contribuire a una crescita sostenibile e responsabile dell'intelligenza artificiale nel lungo termine.

Bibliografia e sitografia

- 1) <https://www.treccani.it/vocabolario/intelligenza/>
- 2) Gardner, H. (1983). "Frames of Mind: The Theory of Multiple Intelligences."
- 3) https://www.treccani.it/enciclopedia/nuove-prospettive-nell-intelligenza-artificiale_%28XXI-Secolo%29/
- 4) <https://www.treccani.it/enciclopedia/intelligenza-artificiale>
- 5) John McCarthy, "What is artificial intelligence?" Stanford University, 2007
- 6) Russell, S. (2010). "Artificial Intelligence: A Modern Approach." Prentice
- 7) Honavar, V. (2006). "Artificial Intelligence and Neural Networks: Steps Toward Principled Integration." *Neural Networks*, 19(3), 367-387.
- 8) Turing, Alan M. *Computing machinery and intelligence*. Springer Netherlands, 2009.
- 9) <https://www.ibm.com/topics/artificial-intelligence>.
- 10) Joseph Weizenbaum, *Computer power and human reason*, Londra, Penguin Books, 1984.
- 11) https://www.treccani.it/magazine/lingua_italiana/speciali/IA/02_Bacciu.html
- 12) https://www.treccani.it/enciclopedia/intelligenza-artificiale_%28Dizionario-di-Medicina%29/
- 13) Simon, Herbert A. "Why should machines learn?" *Machine learning*. Morgan Kaufmann, 1983. 25-37.
- 14) Mitchell, Tom M. "Machine learning." (1997).
- 15) <https://www.ibm.com/topics/deep-learning>
- 16) <https://www.ibm.com/it-it/topics/neural-networks#:~:text=alla%20fase%20successiva-.Cos%C3%A8%20una%20rete%20neurale%3F,degli%20algoritmi%20di%20deep%20learning>.
- 17) J. Neumann, "General and Logical Theory of Automata," in *Cerebral Mechanisms in Behavior: The Hixon Symposium*, 1951
- 18) Lavington, Simon Hugh. *A history of Manchester computers*. NCC Publications, 1975.
- 19) J. Hutchins, "The first public demonstration of machine translation: the Georgetown-IBM system, 7th January 1954", *MT News International*, 2006.
- 20) Hutchins, *Encyclopedia of Language & Linguistics*, Second Edition. Oxford: Elsevier, 2006

- 21) J. McCarthy, M. L. Minsky, and C. E. Shannon, "A proposal for the Dartmouth summer research project on artificial intelligence - August 31, 1955", *Ai Magazine*, vol. 27, no. 4, pp. 12–14, 2006
- 22) F. Rosenblatt, *The Perceptron a perceiving and recognizing automaton*, 1957.
- 23) F. Rosenblatt, "The perceptron: A probabilistic model for information storage and organization in the brain," *Psychological Review*, 1958
- 24) P. Norvig, *Paradigms of Artificial Intelligence Programming: Case Studies in Common Lisp*. 1992, p. 962.
- 25) J. Hutchins, "Machine Translation: A Brief History," 1995.
- 26) Y. Bar-Hillel, "The Present Status of Automatic Translation of Languages", *Advances in Computers*, 1960
- 27) M. M. Papert and S., "Perceptrons. An Introduction to Computational Geometry.", 1969
- 28) J. Schmidhuber, "Deep Learning in neural networks", 2015
- 29) I. Literature and R. Lighthill, "Lighthill Report - Artificial Intelligence: A General Survey", pp. 1–18, 1972.
- 30) C. E. Shannon, "Programming a Computer for Playing Chess", *Philosophical Magazine*, 1950
- 31) J. Hendler, "Avoiding Another AI Winter", *IEEE Intelligent Systems*, 2008.
- 32) G. Johnson, "Thinking About Thinking", *APF Reporter*, vol. 8, no. 1, pp. 1–6, 1985.
- 33) D. Mcdermott, M. M. Waldrop, R. Schank, B. Chandrasekaran, and J. Mcder "The Dark Ages of AI: A Panel Discussion at AAAI-84", *AI Magazine*, vol. 6, no. 3, pp. 122–134, 1985.
- 34) J. McCarthy, "Some Expert System Need Common Sense", pp. 1–11, 1984.
- 35) K. Hao, "We analyzed 16,625 papers to figure out where AI is headed next", *MIT Technology Review*, pp. 1–19, 2019.
- 36) N. J. Nilsson, *The Quest for Artificial Intelligence*, 2010.
- 37) T. Menzies, "21st-Century AI: Proud, not smug", 2003.
- 38) ¹ *Homo Deus – A brief history of tomorrow; Autore: Yuval Noah Harari; Data di pubblicazione: 2017; Editore: Penguin.*
- 39) ¹ Arnd-Caddigan, Margaret. "Sherry Turkle: Alone Together: Why We Expect More from Technology and Less from Each Other: Basic Books, New York, 2011, 348 pp" (2015)

- 40) ¹ Schwab, Klaus. *The fourth industrial revolution*. Currency, 2017.
- 41) B.WOLF, C. SCHOLZE, “Medicine 4.0”, *Biomedical Engineering* 2017; p. 183–186
- 42) *Artificial Intelligence and life in 2030 – One Hundred Year Study on Artificial Intelligence*, Stanford University, 2016
- 43) Davenport, Thomas H. *The AI advantage: How to put the artificial intelligence revolution to work*. mit Press, 2018.
- 44) Mitchell, Melanie. *Artificial intelligence: A guide for thinking humans*. Penguin UK, 2019.
- 45) ¹ LUCKIN Rosemary, *Machine Learning and Human Intelligence: The Future of Education for the 21st Century*, London, UCL IOE Press - University of London - Institute of Education 2018.
- 46) RITTER S. - CARLSON R. - SANDBOTHE M. - FANCSALI S. E., *Carnegie Learning’s adaptive learning products*, in SANTOS O. et alii (eds.), *Educational Data Mining 2015*, 8th International Conference on Educational Data Mining (EDM2015), Madrid
- 47) DE SIMONE L., BURGIO E., “Intelligenza Artificiale e responsabilità civile” in *MediaLaws, Law and Policy of the Media in a Comparative Perspective*, (2021)
- 48) *Facial recognition: is the technology taking away your identity?* The Guardian febbraio 2020
- 49) Dignum, V. (2018). *Ethics in artificial intelligence: introduction to the special issue*. *Ethics and Information Technology*, 20(1), 1-3.
- 50) Vinge, Vernor. "Technological singularity." *VISION-21 Symposium sponsored by NASA Lewis Research Center and the Ohio Aerospace Institute*. 1993.
- 51) Kurzweil, Ray. "The singularity is near." *Ethics and emerging technologies*. London: Palgrave Macmillan UK, 2005. 393-406.
- 52) Kaplan Jerry. *Le persone non servono: lavoro e ricchezza nell’epoca dell’Intelligenza Artificiale*. Roma: Luiss, 2016.
- 53) R. Kurzweil, *The Singularity is Near - When Human Trascends Biology*. 2005 p. 135-136.
- 54) A. Brown, *11 Quotes About AI That’ll Make You Think*, 2019.
- 55) S. Lynch, Andrew Ng: *Why AI Is the New Electricity*, 2017.
- 56) A. Linden and J. Fenn, “Understanding Gartner’s Hype Cycles Understanding Gartner’s Hype Cycles Understanding Gartner’s Hype Cycles”, 2003.

- 57) T. Menzies, “21st-Century AI: Proud, not smug”, 2003
- 58) S. Armstrong and K. Sotola, “How We’re Predicting AI – or Failing to”, pp. 11–29, 2014.
- 59) Hendler, “Avoiding Another AI Winter”, IEEE Intelligent Systems, 2008.
- 60) J. Schmidhuber, Deep Learning in neural networks, 2015.
- 61) X. Guo, H. Lee, X. Wang, and R. L. Lewis, “Deep learning for real-time Atari game play using offline Monte Carlo tree search planning”, Nips, 2014.
- 62) D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, “Mastering the game of Go with deep neural networks and tree search.”, Nature, vol. 529, no. 7587, pp. 484– 489, 2016.
- 63) AlphaStar: Mastering the Real-Time Strategy Game StarCraft II, 2019
- 64) T. Karras, “Progressive Growing of Gans for improved Quality, Stability, and Variation”, pp. 1–26, 2018.
- 65) Lingjiao Chen , Matei Zaharia , James Zou, “*How is ChatGPT’s behavior changing over time?*”, Stanford University, UC Berkeley, p.10, 2023.