

i **quaderni**
della **DIDATTICA**

Intelligenza Artificiale in classe

**Integrare le nuove tecnologie
nella pratica educativa quotidiana**

- Esempi di applicazioni
- IA come tutor (del docente e dello studente)
- Creare materiali didattici
- Personalizzare l'apprendimento
- Esempi di Unità di Apprendimento

E. Barbuto - G. Ciliberti

I Edizione



IN OMAGGIO ESTENSIONI ONLINE

Contenuti
extra



EdiSES
edizioni

Intelligenza artificiale in classe

Accedi ai servizi riservati

Il codice personale contenuto nel riquadro dà diritto a servizi riservati ai clienti. Registrandosi al sito, dalla propria area riservata si potrà accedere a:

**MATERIALI DI INTERESSE
E CONTENUTI AGGIUNTIVI**

CODICE PERSONALE

Grattare delicatamente la superficie per visualizzare il codice personale.
Le **istruzioni per la registrazione** sono riportate nella pagina seguente.
Il volume NON può essere venduto né restituito se il codice personale risulta visibile.
L'**accesso ai servizi riservati** ha la **durata di 18 mesi** dall'attivazione del codice e viene garantito esclusivamente sulle edizioni in corso.

Istruzioni per accedere ai contenuti e ai servizi riservati

SEGUI QUESTE SEMPLICI ISTRUZIONI

SE SEI REGISTRATO AL SITO

clicca su **Accedi al materiale didattico**



inserisci email e password



inserisci le ultime 4 cifre del codice ISBN, riportato in basso a destra sul retro di copertina



inserisci il tuo **codice personale** per essere reindirizzato automaticamente all'area riservata

SE NON SEI GIÀ REGISTRATO AL SITO

clicca su **Accedi al materiale didattico**



registrati al sito **edises.it**

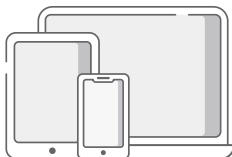


attendi l'email di conferma per perfezionare la registrazione



torna sul sito **edises.it** e segui la procedura già descritta per utenti registrati

CONTENUTI AGGIUNTIVI



Per problemi tecnici connessi all'utilizzo dei supporti multimediali e per informazioni sui nostri servizi puoi contattarci sulla piattaforma **assistenza.edises.it**

SCARICA L'APP **INFOCONCORSI** DISPONIBILE SU APP STORE E PLAY STORE

Intelligenza artificiale in classe



I quaderni della didattica – QD23 – Intelligenza artificiale in classe

Copyright © 2025 EdiSES edizioni S.r.l. – Napoli

9 8 7 6 5 4 3 2 1 0

2029 2028 2027 2026 2025

Le cifre sulla destra indicano il numero e l'anno dell'ultima ristampa effettuata

A norma di legge è vietata la riproduzione, anche parziale, del presente volume o di parte di esso con qualsiasi mezzo.

L'Editore

Progetto grafico: ProMedia Studio di A. Leano – Napoli

Fotocomposizione: Oltrepagina S.r.l.

Stampato presso: PrintSprint S.r.l. – Napoli

Per conto della EdiSES edizioni S.r.l. – Piazza Dante, 89 – Napoli

ISBN 979 12 5602 548 0

www.edises.it

I curatori, l'editore e tutti coloro in qualche modo coinvolti nella preparazione o pubblicazione di quest'opera hanno posto il massimo impegno per garantire che le informazioni ivi contenute siano corrette, compatibilmente con le conoscenze disponibili al momento della stampa; essi, tuttavia, non possono essere ritenuti responsabili dei risultati dell'utilizzo di tali informazioni e restano a disposizione per integrare la citazione delle fonti, qualora incompleta o imprecisa.

Realizzare un libro è un'operazione complessa e, nonostante la cura e l'attenzione poste dagli autori e da tutti gli addetti coinvolti nella lavorazione dei testi, l'esperienza ci insegna che è praticamente impossibile pubblicare un volume privo di imprecisioni. Saremo grati ai lettori che vorranno inviarci le loro segnalazioni e/o suggerimenti migliorativi su assistenza.edises.it

PREFAZIONE

“Il problema reale non è se le macchine pensano, ma se gli uomini continuano a farlo.”

B.F. Skinner

L'intelligenza artificiale è ormai una presenza costante nelle nostre vite. Agisce nei dispositivi che utilizziamo ogni giorno, nei motori di ricerca, nei sistemi di traduzione automatica, nelle piattaforme di comunicazione e di apprendimento. È alla base delle applicazioni che ci aiutano a scrivere, a organizzare il lavoro, a imparare, a viaggiare, a scegliere prodotti e informazioni. In molti casi, non la vediamo, ma essa lavora in sottofondo, interpretando dati, suggerendo soluzioni, orientando decisioni.

Per quanto diffusa, l'intelligenza artificiale resta un fenomeno complesso e ancora in parte misterioso per chi la utilizza. Comprenderla non significa soltanto imparare a usarla, ma saperne leggere i meccanismi, riconoscerne i limiti e interrogarsi sulle implicazioni etiche e sociali che comporta. L'IA non è una semplice invenzione tecnologica: è un cambiamento culturale che ridefinisce il rapporto tra uomo, conoscenza e società. Essa nasce da un'idea profondamente umana, quella di replicare e amplificare l'intelligenza e, proprio per questo, ci obbliga a riflettere su che cosa intendiamo per intelligenza, creatività, autonomia e responsabilità.

La scuola non può rimanere estranea a questa trasformazione. È il luogo in cui si formano le competenze, ma anche la consapevolezza civica e morale dei cittadini. In un mondo dove l'informazione e l'automazione avanzano rapidamente, la scuola deve offrire a studenti e docenti non solo gli strumenti per utilizzare la tecnologia, ma anche le chiavi per comprenderla e governarla. L'intelligenza artificiale, se impiegata con equilibrio e spirito critico, può diventare un prezioso alleato dell'educazione: aiuta a personalizzare l'apprendimento, a sviluppare la creatività, a stimolare il ragionamento, a favorire l'inclusione. Ma, come ogni strumento potente, richiede conoscenza, prudenza e senso etico.

L'obiettivo non è sostituire il docente o ridurre il ruolo dell'insegnamento, bensì rafforzarlo. L'intelligenza artificiale non possiede empatia, intuizione o capacità di discernimento: qualità che restano e resteranno sempre proprie dell'essere umano. La tecnologia può affiancare il docente, ma non potrà mai rimpiazzare la relazione educativa, il dialogo, la capacità di motivare e di comprendere. È dunque essenziale che gli insegnanti e gli studenti imparino a utilizzare questi strumenti come supporto alla crescita e non come scorciatoia del pensiero.

Educare all'intelligenza artificiale significa educare al pensiero critico, alla responsabilità e alla libertà. Significa formare persone capaci di interrogarsi, di valutare le fonti, di distinguere tra ciò che è vero e ciò che è verosimile, di utilizzare la tecnologia per migliorare se stessi e la società. L'IA non deve diventare una forma di delega del giudizio, ma un'occasione per rafforzare la capacità di ragionare.

Anche per questo, la riflessione sull'IA non può essere solo tecnica. È necessario comprendere che dietro ogni algoritmo ci sono scelte umane, dati selezionati, criteri di priorità e interpretazioni del mondo. I sistemi di intelligenza artificiale apprendono dai dati che ricevono: se i dati sono parziali o distorti, anche le loro risposte lo saranno. Da qui l'importanza dell'educazione ai valori dell'imparzialità, della trasparenza e del rispetto della privacy. L'uso etico dell'IA diventa parte integrante della formazione alla cittadinanza digitale, una delle competenze fondamentali per i cittadini del XXI secolo.

Questo volume nasce con l'intento di accompagnare docenti, studenti e formatori in questo percorso di conoscenza. È pensato per un contesto scolastico, con un linguaggio chiaro e accessibile, ma rigoroso, che unisce spiegazione teorica e applicazione pratica. Offre una panoramica completa sull'intelligenza artificiale e sulle sue implicazioni educative, approfondendo i riferimenti normativi, i documenti europei e internazionali, gli strumenti più diffusi e le metodologie didattiche più efficaci per integrare l'IA nella pratica quotidiana della scuola.

L'opera ha un valore particolare perché coniuga l'attenzione scientifica alla concretezza educativa. Non si limita a descrivere il funzionamento dei sistemi intelligenti, ma ne mostra le potenzialità nella progettazione didattica, nella personalizzazione dei percorsi, nella valutazione e nella gestione delle attività scolastiche. In questo modo, offre un contributo significativo alla costruzione di una scuola capace di dialogare con il futuro, senza rinunciare ai propri principi fondanti.

In un'epoca in cui il rischio maggiore è quello di utilizzare la tecnologia senza capirla, questo libro rappresenta un invito alla riflessione e alla consapevolezza. L'intelligenza artificiale può diventare una straordinaria opportunità di crescita, a condizione che l'uomo resti al centro del processo educativo. La conoscenza non può essere delegata a una macchina: deve restare un atto umano, fondato sul ragionamento, sulla curiosità e sul desiderio di verità.

Per questo motivo, è doveroso esprimere un sincero apprezzamento per il lavoro svolto dal prof. Emiliano Barbuto e dal prof. Gabriele Ciliberti, che con questo volume offrono alla scuola italiana uno strumento di grande utilità culturale e didattica. La loro capacità di unire chiarezza espositiva, rigore scientifico e attenzione pedagogica rende il testo accessibile a tutti, ma al tempo stesso profondo e aggiornato. In un ambito in continua evoluzione come quello dell'intelligenza artificiale, il loro contributo assume un valore particolarmente significativo: aiuta a orientarsi tra i concetti, a distinguere il reale dal sensazionalismo, a comprendere che la tecnologia, per essere veramente innovativa, deve sempre rimanere al servizio dell'uomo e della conoscenza.

Questo libro non è solo una guida tecnica, ma un invito a pensare. E la scuola, più di ogni altro luogo, ha il compito di custodire e trasmettere questa capacità: quella di continuare a pensare, con spirito critico e con fiducia nell'intelligenza umana.

On. Pino Bicchielli

INTRODUZIONE

Immaginiamo che, per incanto, un uomo primitivo, ibernato e ritrovato in un ghiacciaio, sia riportato in vita. Immaginiamo inoltre di poter comunicare con questo uomo primitivo e di confrontarci con lui sul suo e sul nostro stile di vita.

L'uomo primitivo, un cacciatore-raccoglitore, che non conosce ancora agricoltura e allevamento, scruta il nostro fisico scarsamente allenato e poco tonico. Lo mette a confronto con il suo, asciutto e temprato dall'ambiente inospitale in cui ha vissuto. Ci chiede se sappiamo arrampicarci sugli alberi o aggrapparci ad una liana, se sappiamo scoccare una freccia con l'arco, scagliare una pietra o andare a caccia in gruppo. Gli rispondiamo che non sappiamo fare nulla di tutto questo, che non siamo abituati ad arrampicarci sugli alberi, che non siamo dei grandi arcieri e che non sappiamo andare a caccia, perché ci procuriamo il cibo diversamente.

Incuriosito da tutto questo, il primitivo ci chiede quale sia il nostro modo di "andare a caccia". Gli rispondiamo che ci rechiamo in un posto che chiamiamo "supermercato" e, tra gli scaffali del reparto macelleria, troviamo dei pezzi di manzo, agnello, capretto, maiale, pollo già tagliati e confezionati che, con una sorta di baratto, possiamo portare a casa. Qui accendiamo un fuoco con un semplice gesto, quasi per magia, e tranquillamente cuciniamo e consumiamo il nostro cibo.

A questo punto il primitivo ci dice con fermezza e grande serenità che noi, nel suo mondo, non sopravvivremmo un giorno. Chi di noi potrebbe dare torto al primitivo? Tuttavia, a ben pensarci, neanche il primitivo potrebbe vivere a lungo nel nostro mondo. Forse finirebbe presto sotto un ponte, a ripararsi dalle intemperie, o in gattabuia, se, costretto dalla necessità, avesse la pretesa di guadagnarsi da vivere in modo poco ortodosso.

Probabilmente è quello che sta succedendo con l'Intelligenza Artificiale, uno strumento che potrebbe essere destinato a cambiare il nostro modo di vivere e lavorare. Come succede spesso, ci sono opinioni divergenti su questo nuovo strumento. Alcuni dicono che ci disabituerebbe a pensare, che i lavori di tipo intellettuale, creativo o artistico sono minacciati e si estingueranno. Altri affermano che in futuro non sapremo svolgere più compiti che oggi svolgiamo quotidianamente, perché l'intelligenza artificiale li svolgerà per noi e questo è da considerarsi una minaccia rilevante.

Il contesto sta evolvendo e lo strumento dell'intelligenza artificiale lo sta modificando in modo concreto e radicale. Ci prepariamo ad intendere il lavoro intellettuale, creativo o artistico in modo diverso? Magari, come un lavoro svolto in collaborazione o in simbiosi con uno strumento di intelligenza artificiale?

Sotto questo punto di vista l'art. 13 della L. 132/2025 afferma che *"L'utilizzo di sistemi di intelligenza artificiale nelle professioni intellettuali è finalizzato al solo esercizio delle attività strumentali e di supporto all'attività professionale e con prevalenza del lavoro intellettuale oggetto della prestazione d'opera"*.

Se ne deduce che quello che sapevamo fare resta importante perché ci qualifica come professionisti, ma quello che gli strumenti di intelligenza artificiale riescono a fare, nel supportare la nostra professione, può darci uno slancio di produttività e farci godere in modo pieno e concreto la parte realmente “intellettuale” della nostra professione. Fino agli anni '70, che hanno visto la nascita delle calcolatrici scientifiche tascabili, a scuola si poneva grande enfasi sulle procedure di calcolo perché le stesse erano essenziali per lo svolgimento di problemi. Inizialmente, l'introduzione delle calcolatrici fu anche osteggiata perché si riteneva che gli studenti non sarebbero stati in grado di preservare le loro conoscenze dell'aritmetica semplice, se avessero imparato ad usare una calcolatrice, prima di aver compreso appieno i concetti matematici di base. D'altro canto, tali paure erano bilanciate dalla possibilità di porre maggiore enfasi sulle tecniche e le procedure risolutive di problemi reali più complessi, che facessero comprendere l'utilità della matematica agli studenti, relegando alla calcolatrice il mero ruolo di strumento che supporta lo studente nel calcolo.

Oggi sappiamo bene quale strada è stata intrapresa. Comprendere al meglio le procedure di calcolo elementare è essenziale perché attiva alcuni processi cognitivi fondamentali e perché ti fornisce una visione critica di quello che producono gli strumenti utilizzati (le calcolatrici). Tuttavia, una volta padroneggiate tali procedure, in tutte le scuole del mondo, si è preferito lasciar utilizzare le calcolatrici scientifiche nelle attività didattiche e negli esami di fine ciclo, perché gli stessi sono stati riformulati ponendo l'attenzione su altre abilità e competenze che non fossero solo quelle del mero calcolo.

Si percorrerà una strada analoga per l'intelligenza artificiale?

L'adozione di questi strumenti nella scuola non è solo una questione di efficienza o di innovazione tecnologica: è anche una questione di responsabilità educativa ed etica. L'uso dell'IA implica riflessioni importanti sulla trasparenza degli algoritmi, sulla protezione dei dati personali, sull'inclusione e sull'equità nell'accesso alle tecnologie. Come indicano anche i principali documenti internazionali (UNESCO, Commissione Europea, INDIRE), è necessario sviluppare un approccio consapevole e critico, che metta al centro la persona e i valori educativi.

Non basta “usare” l'IA: occorre insegnare a comprenderla, a governarla, e a valutarne gli effetti sulle pratiche didattiche e sulle relazioni educative.

In questo volume ci proponiamo di:

- > Far comprendere come lavora un sistema di intelligenza artificiale (IA) anche ai non addetti ai lavori;
- > Analizzare i documenti guida e gli atti normativi che regolano l'utilizzo dell'intelligenza artificiale, ponendo particolare enfasi sui rischi e sulle potenzialità dello strumento;
- > Presentare una rassegna di software che utilizzano l'intelligenza artificiale o di sistemi di intelligenza artificiale che si possono rivelare utili per docenti e studenti;
- > Fornire ai docenti e agli studenti modalità di utilizzo dell'intelligenza artificiale nelle attività didattiche e di studio che siano conformi ai documenti guida e agli atti normativi e che possano far comprendere i nuovi orizzonti educativi che si aprono con questo strumento;
- > Proporre esempi concreti di utilizzo della IA nelle attività didattiche attraverso la presentazione di Unità di Apprendimento pensate e realizzate per ogni ordine e grado di scuola, corredate di materiali da scaricare e utilizzare in classe.

Come accadde per la calcolatrice, l'Intelligenza Artificiale non sostituirà la nostra intelligenza: la potenzierà. Ci troviamo davanti a uno snodo storico in cui la scuola può scegliere se subire questa trasformazione o esserne protagonista. La sfida non è solo tecnologica, ma profondamente culturale ed educativa. Significa interrogarsi su come trasmettere il sapere, su quali competenze vogliamo sviluppare nei nostri studenti e sul ruolo che i docenti sono chiamati a ricoprire in un contesto radicalmente nuovo. L'IA non è semplicemente uno strumento da aggiungere alla cassetta degli attrezzi didattici: è un interlocutore, un amplificatore, un acceleratore di idee e processi. Può supportare nella progettazione didattica, personalizzare percorsi di apprendimento, generare contenuti e stimolare nuovi modi di esplorare la conoscenza. Ma, come ogni potente innovazione, richiede consapevolezza, criteri etici solidi e una visione educativa chiara. Questo libro nasce per accompagnare docenti, dirigenti e studenti italiani in questo passaggio epocale. Non si tratta di sostituire ciò che siamo, ma di imparare a collaborare con una nuova forma di intelligenza per arricchire la nostra. Sta a noi scegliere se guardare a questi strumenti con timore o con responsabilità e coraggio. Se li useremo con intelligenza, potremo rendere la scuola un luogo ancora più creativo, inclusivo e capace di preparare i cittadini del futuro, non solo a vivere in un mondo con l'IA, ma a plasmarlo.

Ulteriori **materiali didattici** e **aggiornamenti** sono disponibili nell'area riservata a cui si accede mediante la registrazione al sito *edises.it* secondo la procedura indicata nelle prime pagine del volume.

Eventuali errata-corrige saranno pubblicati sul sito *edises.it*, nella scheda "Aggiornamenti" della pagina dedicata al volume.

Altri aggiornamenti sulle procedure concorsuali saranno disponibili sui nostri **social**, su **blog.edises.it** e **infoconcorsi.com**.

Indice

Capitolo 1 – IA, di cosa si tratta?

1.1	Introduzione all’Intelligenza Artificiale (IA)	1
1.2	Differenze tra algoritmi tradizionali e algoritmi di IA	3
1.3	Tipologie di Intelligenza Artificiale (IA debole vs IA forte, IA generativa vs IA analitica)	7
1.3.1	IA debole vs IA forte	7
1.3.2	IA generativa vs IA analitica	9
1.4	Machine learning e deep learning	11
1.4.1	Machine learning (apprendimento automatico)	11
1.4.2	Deep learning (apprendimento profondo)	13
1.5	LLM (Large Language Models) e il loro funzionamento	15
1.5.1	Cosa rende “grandi” questi modelli linguistici?	16
1.6	Sfide ed etica nell’uso dell’IA	21
1.6.1	Bias e discriminazione	21
1.6.2	Privacy e dati personali	21
1.6.3	Trasparenza e spiegabilità	22
1.6.4	Affidabilità e sicurezza	22
1.6.5	Impatto sul lavoro e sulle competenze	23
1.6.6	Aspetti legali e responsabilità	23
1.6.7	Uso improprio e rischi sociali	24
1.6.8	Impatto ambientale	25
1.6.9	Uso dell’IA nella didattica: opportunità e dilemmi	25

Capitolo 2 – La IA nei primi documenti di carattere sovranazionale

2.1	Il Quadro delle Competenze Digitali per i Cittadini (il DigComp2.2)	28
2.1.1	Finalità del DigComp2.2	28
2.1.2	La competenza nel DigComp 2.2	29
2.1.3	Le cinque aree di competenza	30
2.1.4	Il livello di padronanza delle competenze	32
2.1.5	Struttura generale del DigComp 2.2	34
2.1.6	L’intelligenza artificiale nel DigComp 2.2	34
2.1.7	L’Intelligenza artificiale nell’allegato A2 al DigComp 2.2	40
2.2	Il report “K-12 AI curricula - A mapping of government-endorsed AI curricula”	42
2.2.1	La finalità del report	42
2.2.2	Le fasi della creazione del curriculum	42
2.2.3	Dal curriculum alle competenze	44
2.3	Gli Orientamenti etici per gli educatori sull’uso dell’Intelligenza Artificiale (IA) e dei dati nell’insegnamento e nell’apprendimento	46
2.3.1	Struttura del documento	46
2.3.2	Esempi di uso dell’IA e dei dati nel settore dell’istruzione	47
2.3.3	Requisiti fondamentali per un’IA affidabile e le domande orientative per gli educatori	50
2.3.4	Le Competenze emergenti in materia di uso etico dell’IA e dei dati	54
2.4	Il Team dello European Digital Education Hub “AI in Education” e i 7 rapporti informativi	60
2.4.1	La nascita del Team e i sette rapporti informativi	60

2.4.2	I tre nuclei di competenza del docente in relazione all'Intelligenza Artificiale	60
2.4.3	I tre nuclei di competenze riferiti ai tre documenti preesistenti.....	62
2.4.4	Aspetti fondamentali del nucleo di competenze per l'IA (competenze di tipo civico)....	62
2.4.5	Aspetti fondamentali del nucleo di competenze con l'IA (competenze di tipo pedagogico e didattico-metodologico).....	63
2.4.6	Aspetti fondamentali del nucleo di competenze sull'IA (competenze di tipo tecnologico a livello basilare)	64
2.4.7	Il modello TPACK riadattato ai tre nuclei di competenza IA per i docenti.....	64

Capitolo 3 – La IA nella normativa vigente

3.1	Regolamento (UE) 2024/1689 del Parlamento Europeo e del Consiglio del 13 giugno 2024..	69
3.1.1	Gli obiettivi del regolamento.....	69
3.1.2	La definizione di Sistema IA	70
3.1.3	Livello di rischio di un sistema IA.....	72
3.1.4	Sistemi di IA con rischio inaccettabile, pratiche vietate con la IA	73
3.1.5	Sistemi di IA ad alto rischio	75
3.1.6	Sistemi di IA a rischio limitato	78
3.1.7	Sistemi di IA a rischio minimo o nullo	79
3.2	Il Quadro delle Competenze IA dell'UNESCO per gli insegnanti.....	80
3.2.1	Un quadro di riferimento centrato intorno all'uomo	80
3.2.2	I principi del quadro delle competenze IA.....	80
3.2.3	Struttura del quadro di riferimento delle competenze IA per i docenti	82
3.3	Il Quadro delle Competenze IA dell'UNESCO per gli studenti.....	84
3.3.1	Un quadro di riferimento centrato sull'etica, sulle conoscenze e sulle abilità	84
3.3.2	I principi del quadro delle competenze IA per gli studenti	85
3.3.3	Struttura del quadro delle competenze IA per gli studenti	86
3.4	La IA nelle Linee Guida STEM del MIM.....	89
3.5	La IA nelle Nuove Linee Guida per l'Insegnamento Trasversale dell'Educazione Civica.....	90
3.6	La IA nelle Nuove indicazioni per la scuola dell'infanzia e primo ciclo di istruzione 2025.....	91
3.6.1	Il dibattito intorno alle Nuove Indicazioni	91
3.6.2	Integrare l'IA con prudenza e senso critico	91
3.6.3	L'IA per garantire una scuola inclusiva.....	93
3.6.4	La struttura delle nuove Indicazioni Nazionali 2025.....	93
3.7	Le Linee Guida per l'introduzione dell'Intelligenza Artificiale nelle Istituzioni scolastiche...	94
3.7.1	Introduzione.....	94
3.7.2	La IA per i docenti.....	94
3.7.3	La IA per gli studenti.....	96
3.8	La Legge sulla IA	97
3.8.1	I principi di funzionamento della IA.....	97
3.8.2	I temi trattati dalla legge	98
3.8.3	Le professioni intellettuali.....	98
3.8.4	Interventi nell'ambito dell'istruzione.....	99

Capitolo 4 – I maggiori software di intelligenza artificiale generativa

4.1	ChatGPT (OpenAI).....	101
4.1.1	Modelli LLM utilizzati.....	101
4.1.2	Funzionalità principali.....	101
4.1.3	Accesso e registrazione	102
4.1.4	Versione gratuita vs a pagamento	102
4.2	Claude (Anthropic).....	103

4.2.1	Modelli LLM e versioni	104
4.2.2	Funzionalità principali.....	104
4.2.3	Interfaccia e “artefatti eseguibili”	105
4.2.4	Accesso al servizio	105
4.2.5	Versione gratuita vs a pagamento	105
4.3	Google Gemini	107
4.3.1	Modelli LLM e versioni	107
4.3.2	Funzionalità principali.....	108
4.3.3	Accesso e modalità di utilizzo	109
4.3.4	Versione gratuita vs a pagamento	109
4.4	Perplexity	110
4.4.1	Modelli LLM utilizzati.....	111
4.4.2	Differenze tra versioni	111
4.4.3	Funzionalità principali.....	112
4.4.4	Accesso e registrazione	113
4.4.5	Versione gratuita vs a pagamento	113
4.5	Google NotebookLM.....	115
4.5.1	Modello LLM utilizzato	115
4.5.2	Funzionalità principali.....	116
4.5.3	Accesso e registrazione	117
4.5.4	Versione gratuita vs a pagamento	118
4.6	Generatori di immagini (DALL-E, Leonardo AI).....	120
4.6.1	DALL-E (OpenAI)	120
4.6.2	Leonardo AI.....	123
4.7	Generatori di presentazioni	126
4.7.1	Gamma App	126
4.7.2	Beautiful.ai	127
4.7.3	Decktopus AI	127
4.7.4	Canva - Magic Design	127
4.8	Generatori di diagrammi	128
4.8.1	Mind map (mappe mentali).....	128
4.8.2	Timeline (sequenze temporali).....	129
4.8.3	Diagrammi di flusso (flow chart).....	130
4.9	Generatori di videoclip	131
4.9.1	Runway	131
4.9.2	Pixverse	132
4.9.3	Sora (OpenAI)	132
4.10	Directory online per strumenti IA	133
4.10.1	Futurepedia	133
4.10.2	AI Tools Directory	133
4.10.3	Alcyclopedia	134
4.10.4	There’s an AI for That	134

Capitolo 5 - Il vaglio delle fonti e l'intelligenza artificiale

5.1	L'importanza delle fonti con cui si addestra il sistema IA	136
5.2	Affidabilità di una fonte di informazione	137
5.3	Imparzialità di una fonte di informazione	138
5.4	Le Bufale e le fake news.....	140
5.5	Disinformazione e misinformazione.....	142
5.6	Il debunking (i siti web demistificatori)	143

5.7	Aggiornamento di una notizia.....	145
5.8	Il Clickbaiting	145
5.9	La profilazione degli utenti.....	147
5.10	Le bolle di filtraggio, le stanze dell'eco ed il bias confermativo	150
5.11	Le fonti con cui si addestra un sistema di IA.....	152
5.12	Utilizzare un sistema IA per valutare l'autorevolezza di una fonte.....	155
5.12.1	Primo esempio - Fornire un indirizzo Web a ChatGPT.....	155
5.12.2	Secondo esempio - Fornire un indirizzo Web a DeepSeek.....	159
5.12.3	Terzo esempio - Fornire un testo a ChatGPT	162
5.12.4	Quarto esempio - Fornire dati numerici a DeepSeek.....	165
5.13	Analisi delle fonti con motori di ricerca che si basano su sistemi di IA generativa	171
5.13.1	Perplexity.ia e la consultazione delle fonti.....	171
5.13.2	La scelta delle fonti in Perplexity.ia	173
5.13.5	Gli spazi di Perplexity.ia.....	176
5.14	Utilizzare un sistema IA per valutare l'imparzialità di una fonte	182
5.14.1	I criteri utilizzati per verificare l'imparzialità.....	182
5.14.2	La funzione di verifica delle fonti di Gemini.....	184
5.14.3	Esempio - Fornire un testo a Gemini per verificarne l'imparzialità	187
5.15	Metodologie didattiche e consultazione delle fonti	192
5.15.1	Il Debate	192
5.15.2	L'apprendimento basato sull'indagine.....	196

Capitolo 6 - Creare materiali didattici con la IA

6.1	Un racconto per bambini sull'educazione stradale.....	200
6.2	Creare una poesia in rima in lingua inglese per la festa della mamma.....	204
6.3	Riconoscere diversi stili letterari.....	207
6.4	Una versione di latino con un tema insolito.....	210
6.5	Una simulazione della prova di matematica per l'esame di Stato	213
6.6	Un riassunto modificato del primo capitolo dei Promessi sposi	218
6.7	Riconoscere stili pittorici	225
6.8	Coding con l'intelligenza artificiale.....	237
6.9	Preparare i materiali per un debate.....	245
6.10	Un test di storia	249
6.11	Produzione di tante verifiche distinte per quanti sono gli alunni di una classe.....	253
6.12	Produzione di compiti distinti per livello (base, intermedio e avanzato)	255
6.13	Produzione di un database di domande	258
6.14	Creazione di un quiz interattivo.....	262

Capitolo 7 - L'IA come tutor del docente

7.1	Applicare una metodologia didattica.....	270
7.2	Programmare le attività didattiche	278
7.3	Generare una Unità di Apprendimento.....	296
7.4	Creare collegamenti interdisciplinari.....	305
7.5	Scrivere comunicazioni per i genitori degli alunni	312

Capitolo 8 - L'IA come trainer dello studente

8.1	Le esercitazioni interattive con la IA.....	316
8.1.1	Caso 1: L'Intelligenza artificiale interroga su un argomento e poi dà un riscontro	317
8.1.2	Caso 2: L'intelligenza artificiale interroga su un argomento, in caso di errore dà un suggerimento e poi dà un riscontro.....	319

8.1.3	Caso 3: L'intelligenza artificiale interroga su contenuti specifici.....	325
8.2	Le Flashcard.....	327
8.2.1	Cosa sono le Flashcard e come si utilizzano.....	327
8.2.2	Creiamo un semplice set di Flashcard.....	328
8.2.3	Utilizziamo il software Anki.....	330
8.3	Giochi didattici.....	337
8.3.1	Indovina il romanzo.....	337
8.3.2	Indovina il romanzo da un elenco allegato.....	342
8.3.3	Chi lo ha detto?.....	343
8.3.4	"Chi lo ha detto?" da un elenco di personaggi famosi.....	344
8.3.5	Indovina la divinità.....	345

Capitolo 9 - L'IA come tutor dello studente

9.1	Assistere lo studente durante lo svolgimento di un calcolo o di un problema.....	348
9.1.1	Farsi assistere nei calcoli dalla IA invece di farli svolgere alla IA.....	348
9.1.2	Calcolo assistito di una espressione algebrica semplice.....	349
9.1.3	Calcolo assistito di una espressione algebrica più complessa.....	350
9.1.4	Risoluzione assistita di un problema sulle proporzioni.....	352
9.2	Confrontare il mio elaborato con quello della IA.....	353
9.2.1	Esempio 1: correzione di un riassunto sulla prima guerra punica.....	354
9.2.2	Esempio 2: correzione di un problema di geometria.....	357
9.3	Cercare nessi e relazioni in trattazioni estese.....	359
9.3.1	Esempio 1: Il Paradiso di Dante e i teoremi della Geometria Euclidea.....	361
9.3.2	Esempio 2: La morale di Pinocchio.....	367

Capitolo 10 - Personalizzare l'apprendimento con l'intelligenza artificiale

10.1	La personalizzazione e gli stili di apprendimento.....	372
10.2	Stili di apprendimento: un modello fisiologico.....	373
10.3	La personalizzazione dell'input e dell'output.....	375
10.4	Semplificare un testo con l'IA.....	376
10.4.1	Semplificare testi con ChatGPT.....	376
10.4.2	Semplificare testi con Copilot.....	380
10.5	Generare mappe concettuali da un testo con l'IA.....	381
10.5.1	Generare mappe con Algor Education.....	381
10.5.2	Generare mappe con ChatGPT e Markmap.....	383
10.5.3	Generare una mappa concettuale con MyLens.ai.....	386
10.6	Generare una linea temporale con l'IA.....	388
10.6.1	Introduzione.....	388
10.6.2	Generare una linea temporale con MyLens.ai.....	389
10.7	Generare un quadrante con l'IA.....	390
10.7.1	Introduzione.....	390
10.7.2	Generare quadrante con MyLens.ai.....	390
10.8	Generare una presentazione multimediale da un testo con SlidesAI.....	391
10.9	Generare videoclip da un testo con l'IA.....	400
10.9.1	Generare videoclip con Wave.video.....	400
10.9.2	Generare videoclip con Vidnoz.....	405

Capitolo 11 - Strumenti per identificare la IA

11.1	Strumenti antiplagio.....	413
11.1.1	Il plagio nell'era dell'intelligenza artificiale.....	413

11.1.2	Perché utilizzare strumenti antiplagio nella didattica	414
11.1.3	Criteri di selezione degli strumenti antiplagio	415
11.1.4	Scribbr.....	417
11.1.5	Grammarly.....	419
11.1.6	Copyscape.....	421
11.1.7	DupliChecker.....	423
11.1.8	Conclusione sugli strumenti antiplagio.....	425
11.1.9	Guida passo-passo: come utilizzare Scribbr per il controllo antiplagio	425
11.1.10	Riflessione finale: uso consapevole, etico e formativo dei tool antiplagio	430
11.2	Identificare la IA	431
11.2.1	Importanza di rilevare testi generati da IA in ambito educativo.....	431
11.2.2	Perplexity e Burstiness: metriche per rilevare testo generato da IA.....	432
11.3	Strumenti per identificare testi da IA: panoramica	434
11.3.1	GPTZero.....	434
11.3.2	Originality.AI	436
11.3.3	Winston AI	438
11.3.4	ZeroGPT.....	440
11.3.5	Scribbr AI Detector	442
11.3.6	Limiti tecnologici ed etica del rilevamento IA.....	444
11.4	Strumenti per “umanizzare” l’output della IA	446
11.4.1	Humanize AI.....	446
11.4.2	QuillBot Humanizer.....	447
11.4.3	DetectGPT Humanizer.....	448
11.4.4	GPTHuman.....	449
11.4.5	Humbot.....	450
11.4.6	Confronto tra i tool e strategie d’uso	452
11.4.7	Workflow: come “umanizzare” il testo IA e verificarlo.....	453
11.5	Strumenti “anti-humanizer”: rilevare testi IA umanizzati.....	457
11.5.1	Introduzione.....	457
11.5.2	LLMDetectAlve.....	457
11.5.3	DAMAGE.....	459
11.5.4	Raidar ed altre tecniche	460
11.5.5	Limiti e prospettive	461
11.6	Più delle parole	462

Capitolo 12 - Software per la didattica basati sulla IA

12.1	Magic School AI.....	464
12.1.1	Introduzione.....	464
12.1.2	Registrazione.....	464
12.1.3	Ambiente di lavoro.....	467
12.1.4	I livelli di grado.....	467
12.1.5	Strumenti per gli insegnanti: generatore di rubriche.....	468
12.1.6	Strumenti per gli insegnanti: Domande su video di YouTube	470
12.1.7	Strumenti per gli insegnanti: Testi informativi.....	473
12.1.8	Altri strumenti per gli insegnanti	475
12.1.9	Strumenti per gli insegnanti: Piano di lezione	475
12.1.10	Strumenti per gli insegnanti: revisione a spirale di matematica	477
12.1.11	Cronologia degli Output.....	478
12.1.12	Il chatbot Raina.....	479
12.1.13	Strumenti per studenti.....	479

12.1.14 Stanze	480
12.2 Eduaide.ai.....	480
12.2.1 Introduzione.....	480
12.2.2 Registrazione.....	481
12.2.3 Ambiente di lavoro.....	481
12.2.4 Planning	482
12.2.5 Presenting	484
12.2.6 Practice e Group Activities.....	485
12.2.7 Questions	486
12.2.8 Assessments	486
12.2.9 La sequenza delle attività presenti in Eduaide.ai	487
12.3 Diffit For Teachers.....	487
12.3.1 Introduzione.....	487
12.3.2 Accesso e registrazione	488
12.3.3 Generare una Unità di Apprendimento	489
12.3.4 Books	490
12.3.5 Altre funzionalità	491

Materiali riservati online

Esempi di Unità di Apprendimento	495
Sitografia e bibliografia	495



Capitolo 1

IA, di cosa si tratta?

In questo capitolo si cerca di chiarire cosa si intende per intelligenza artificiale e quali vantaggi dà il suo utilizzo. Dopo un breve excursus storico sulle origini e gli sviluppi dell'intelligenza artificiale, esploreremo dapprima le differenze tra gli algoritmi tradizionali e quelli utilizzati nell'IA. Successivamente, passeremo in rassegna le principali tipologie di intelligenza artificiale, distinguendo concetti come IA "debole" vs "forte" e IA "generativa" vs "analitica". Proseguiremo approfondendo l'importante ambito del machine learning e del deep learning, il motore attuale di molti progressi nell'IA. Parleremo poi dei modelli linguistici di grandi dimensioni (LLM, Large Language Models) e di come funzionano sistemi come i chatbot avanzati che conversano in linguaggio naturale. Infine, affronteremo le sfide e le questioni etiche legate all'uso crescente dell'intelligenza artificiale, specialmente in un contesto educativo.

1.1 Introduzione all'Intelligenza Artificiale (IA)

L'intelligenza artificiale, spesso indicata con la sigla IA (in inglese AI, *Artificial Intelligence*), è la disciplina informatica che si occupa di creare sistemi e programmi capaci di svolgere compiti che normalmente richiederebbero l'intelligenza umana. In altre parole, l'IA cerca di far sì che i computer "pensino" o, più propriamente, risolvano problemi, prendano decisioni e apprendano informazioni in modo simile a come farebbe un essere umano. Questo non significa che i computer abbiano coscienza o emozioni come noi, ma che possono imitare alcune capacità cognitive umane come riconoscere immagini, comprendere il linguaggio, apprendere dall'esperienza e così via.

Che cosa significa concretamente tutto ciò? Immaginiamo le attività che consideriamo "intelligenti": riconoscere il volto di un amico in una fotografia, conversare in una lingua naturale, risolvere un problema matematico complesso, giocare a scacchi strategicamente, o persino guidare un'automobile in mezzo al traffico. Fino a pochi decenni fa, compiti del genere potevano essere svolti solo da persone in carne e ossa. Oggi, grazie ai progressi nell'IA, esistono programmi e dispositivi in grado di eseguire molte di queste operazioni. Ad esempio, i moderni smartphone sono dotati di assistenti virtuali che rispondono a domande vocali, i servizi di streaming video consigliano film e serie basandosi sui gusti degli utenti, e le automobili a guida autonoma stanno imparando a muoversi senza l'intervento umano. Tutti questi sono esempi di intelligenza artificiale applicata alla vita quotidiana.

Vale la pena notare che la realtà dell'IA odierna è molto diversa da come viene talvolta immaginata nei film di fantascienza. Nelle storie cinematografiche compaiono spesso robot pienamente autocoscienti o supercomputer ribelli che cercano di dominare gli esseri umani. Nella realtà attuale nulla di simile esiste: le IA di oggi, per quanto avanzate, non possiedono coscienza né volontà propria. Esse eseguono compiti specifici per cui sono state progettate e addestrate, e non agiscono fuori da quei confini se non perché

un essere umano lo ha programmato o richiesto. In altre parole, non c'è da temere che “le macchine prendano il sopravvento” spontaneamente come nelle trame hollywoodiane; le vere sfide riguardano piuttosto l'uso corretto che noi umani facciamo di queste tecnologie (un tema che affronteremo verso la fine di questo capitolo discutendo delle implicazioni etiche).

L'IA non è una singola tecnologia monolitica, ma un insieme di sotto-discipline diverse. Per esempio, vi è la visione artificiale (*computer vision*), che permette a un computer di analizzare e comprendere immagini o video (come fare il riconoscimento facciale o identificare oggetti in una foto); vi è l'elaborazione del linguaggio naturale (*Natural Language Processing*), che riguarda la comprensione e produzione di testi o discorsi da parte delle macchine (ad esempio traduttori automatici, sistemi come Siri o chat bot avanzati); c'è la robotica intelligente, dove l'IA è inserita in robot fisici per farli muovere e interagire con l'ambiente (dai bracci meccanici nelle fabbriche ai robot domestici aspirapolvere); i sistemi esperti, che cercano di emulare il ragionamento di un esperto umano in un certo dominio tramite regole logiche; e molte altre aree ancora. Ognuna di queste aree affronta aspetti diversi dell'intelligenza e ha sviluppato tecniche specifiche, ma spesso i progressi in un campo (ad esempio nel riconoscimento di immagini) aiutano anche altri campi (come la robotica, che ha bisogno di “vedere” per muoversi). In questo capitolo e nel libro ci concentreremo soprattutto sull'IA basata sull'apprendimento automatico, che è quella che più ha rivoluzionato il settore negli ultimi anni ed è trasversale a molti ambiti citati.

L'idea di creare macchine intelligenti non è nuova. Già negli anni '50 del Novecento, pionieri dell'informatica come Alan Turing e John McCarthy posero le basi teoriche per l'IA moderna. Il termine “intelligenza artificiale” fu coniato proprio in quel periodo (nel 1956, durante una conferenza al Dartmouth College negli Stati Uniti) per indicare questo nuovo campo di studi. Inizialmente l'entusiasmo era altissimo: si pensava che nel giro di pochi anni le macchine avrebbero potuto fare praticamente qualsiasi cosa. Tuttavia, la realtà si rivelò più complessa del previsto. Tra gli anni '70 e '80, la ricerca nell'IA incontrò molte difficoltà: i computer dell'epoca erano lenti e con poca memoria, e mancavano sia i dati in quantità sufficiente sia le idee giuste per raggiungere risultati significativi. Questo portò a un periodo di stasi noto come “inverno dell'IA”, in cui i finanziamenti e l'interesse calarono notevolmente. Nonostante queste battute d'arresto, ci furono comunque alcuni successi notevoli nelle decadi successive. Ad esempio, nel 1997 il computer IBM Deep Blue sconfisse il campione del mondo di scacchi Garry Kasparov, dimostrando che un programma poteva superare le abilità umane in un gioco complesso (sebbene utilizzando un approccio perlopiù basato su forza bruta e regole predefinite). Nel 2011 un altro sistema di IA di IBM, chiamato Watson, vinse il quiz televisivo americano “Jeopardy!” contro avversari umani, rispondendo correttamente a domande di cultura generale formulate in linguaggio naturale. Questi traguardi, ottenuti con metodologie diverse, hanno mostrato le potenzialità dell'IA e hanno contribuito a rinnovare l'interesse del grande pubblico.

Le cose cambiarono a partire dagli anni '90 e soprattutto nei primi decenni del 2000. Due fattori principali hanno contribuito alla “rinascita” dell'intelligenza artificiale: da un lato, l'aumento esponenziale della potenza di calcolo dei computer (basti pensare che gli smartphone odierni hanno più capacità di calcolo di quanta ne avessero i supercomputer di qualche decennio fa); dall'altro, la disponibilità di enormi quantità di dati digitali. Pensiamo a quante informazioni vengono generate ogni giorno: testi, immagini,

video, transazioni economiche, dati scientifici, interazioni sui social network. Tutti questi dati costituiscono un tesoro prezioso per allenare sistemi di IA. Inoltre, sono stati sviluppati nuovi algoritmi e approcci – come quelli di *machine learning* e *deep learning*, di cui parleremo in seguito – che permettono ai computer di apprendere dai dati in modi molto più efficaci rispetto al passato.

Oggi l'intelligenza artificiale è diventata una realtà tangibile e pervasiva. In quali settori la troviamo? Praticamente in tutti. Nella medicina, ad esempio, l'IA aiuta ad analizzare radiografie e risonanze magnetiche per individuare precocemente tumori o altre patologie, e può persino assistere i medici nel formulare diagnosi o nel suggerire trattamenti personalizzati per i pazienti. Nella finanza, viene utilizzata per prevedere l'andamento dei mercati o rilevare transazioni sospette (come le frodi con carte di credito), e in alcune banche l'IA supporta le decisioni di concessione di prestiti analizzando il profilo di rischio dei clienti. Nel settore educativo, oggetto di questo libro, iniziano a diffondersi tutor intelligenti che supportano lo studio degli studenti adattandosi al loro livello, oppure sistemi di valutazione automatica che correggono compiti a risposta aperta. Nell'agricoltura esistono sistemi di IA che monitorano le coltivazioni tramite sensori e immagini satellitari, per ottimizzare l'uso di acqua e pesticidi e prevedere l'andamento dei raccolti. Nei trasporti si parla sempre più spesso di veicoli autonomi e di ottimizzazione intelligente del traffico: ad esempio, sistemi che regolano i semafori in città in base al flusso reale di automobili, riducendo gli ingorghi. Nella vita di tutti i giorni usiamo l'IA quando affidiamo al navigatore GPS il percorso migliore, quando il filtro antispam blocca le e-mail indesiderate, o quando un traduttore online converte istantaneamente un testo da una lingua a un'altra.

Questi esempi mostrano quanto l'IA sia già integrata nella società moderna, spesso in modo quasi invisibile. Probabilmente molti di noi interagiscono con sistemi intelligenti più volte al giorno senza farci caso. Ma come funzionano davvero questi sistemi? E come si differenziano dai programmi "tradizionali" con cui i computer sono stati programmati per tanti anni? Lo vedremo a breve. Prima di tutto, è fondamentale capire che alla base di ogni applicazione di IA c'è un algoritmo, ovvero una serie di istruzioni che il computer segue. Quello che differenzia un sistema di intelligenza artificiale da un software classico è soprattutto il modo in cui tali istruzioni vengono generate e come si comportano di fronte a nuovi problemi o dati imprevisti.

1.2 Differenze tra algoritmi tradizionali e algoritmi di IA

Abbiamo accennato che la caratteristica distintiva di un sistema di intelligenza artificiale è la sua capacità di apprendere dai dati e adattarsi. Questo lo distingue dai programmi tradizionali, che invece seguono rigidamente istruzioni predefinite. Ma entriamo più nel dettaglio: cosa intendiamo esattamente per algoritmo tradizionale e per algoritmo di IA?

Un algoritmo in generale è una sequenza finita di istruzioni che descrive come svolgere un determinato compito o risolvere un problema. Ad esempio, una ricetta di cucina può essere vista come un algoritmo: ha una serie di passi da eseguire in ordine per ottenere il piatto finale. Allo stesso modo, un programma informatico tradizionale è costituito da linee di codice (istruzioni) che il computer esegue passo dopo passo. In un algoritmo tradizionale, ogni possibile situazione è anticipata dal programmatore e gestita tramite istruzioni esplicite. Il computer, dunque, non "inventa" nulla di proprio: esegue sempli-



cemente i comandi scritti nel programma, sempre allo stesso modo, ogni volta che si presentano le medesime condizioni.

Prendiamo un esempio semplice: immaginate un programma per calcolare la media aritmetica di due numeri. Un programmatore potrebbe scrivere un breve algoritmo tradizionale in questo modo: (1) prendere il primo numero, (2) prendere il secondo numero, (3) sommare i due numeri, (4) dividere la somma per 2, (5) mostrare il risultato. Questo algoritmo è fisso e funzionerà sempre per calcolare la media, ma non può fare nient'altro. Se chiedessimo a questo programma di calcolare la media di tre numeri, non saprebbe come fare perché non è stato programmato per gestire tre valori – bisognerebbe modificare il codice e aggiungere le istruzioni specifiche per quel caso.

Nei tradizionali algoritmi, dunque, l'intelligenza risiede tutta nel programmatore umano che anticipa le situazioni e scrive il codice di conseguenza. Ogni condizione ("se succede X, allora fai Y") deve essere prevista e codificata manualmente. Questo approccio funziona bene per problemi chiari e ben definiti, in cui possiamo delineare tutte le regole in anticipo. Ad esempio, per ordinare una lista di numeri dal più piccolo al più grande, esistono algoritmi matematici precisi (come il "bubble sort" o il "merge sort") che un programmatore può implementare facilmente. Analogamente, per calcolare il percorso più breve tra due città su una mappa, possiamo programmare un algoritmo deterministico (come l'algoritmo di Dijkstra) che segue regole fisse.

Ma cosa succede se il problema è troppo complesso per essere descritto con un elenco di regole rigide? Pensiamo al riconoscimento vocale: come faremmo a scrivere a mano, in un programma tradizionale, tutte le possibili frasi che una persona potrebbe pronunciare e tutte le varianti di accenti, toni, rumori di fondo? Oppure consideriamo la classificazione delle immagini: come potremmo elencare regole precise per dire a un computer come riconoscere la differenza tra la foto di un gatto e quella di un cane? Ci sono gatti di mille razze e colori, in posizioni e sfondi diversi, e altrettante variabilità per i cani. Elencare manualmente caratteristiche sufficienti (es. "il gatto ha orecchie a punta, coda lunga, dimensioni piccole..") sarebbe pressoché impossibile, senza contare che molte regole avrebbero eccezioni.

Ecco dove entrano in gioco gli algoritmi di IA. Invece di richiedere al programmatore di prevedere ogni scenario, questi algoritmi sono progettati per imparare dalle esperienze, ovvero dai dati. Il programmatore, in questo caso, non specifica tutte le regole per il compito, ma fornisce al sistema un meccanismo per apprendere e un insieme di dati da cui apprendere. Il risultato è che l'algoritmo di IA "scrive da solo" (in termini figurativi) le proprie regole interne sulla base dei pattern che trova nei dati.

Riprendiamo l'esempio del riconoscimento di immagini di gatti e cani. Un approccio basato sull'IA (e in particolare sul *machine learning*, l'apprendimento automatico) non cercherebbe di codificare "cos'è un gatto" a tavolino. Piuttosto, si costruirebbe un modello (ad esempio, una rete neurale artificiale) e lo si addestrerebbe mostrandogli tantissime immagini, alcune etichettate come "gatto" e altre come "cane". Inizialmente il modello non sa distinguere un gatto da un cane meglio di quanto farebbe un tiro di moneta. Ma gradualmente, analizzando migliaia di esempi e correggendo i propri errori ad ogni tentativo, l'algoritmo aggiusta i suoi parametri interni finché non diventa in grado di dare la risposta giusta nella maggior parte dei casi. Alla fine dell'addestramento, pur senza che nessuno gli abbia mai fornito esplicitamente la "regola" per riconoscere un gatto, il sistema avrà imparato da sé una serie di caratteristiche ricorrenti (ad esempio la forma generale, il tipo di pelo, ecc.) che distinguono i gatti dai cani. E soprattutto, avrà

la capacità di generalizzare questa conoscenza a immagini nuove che non ha mai visto prima.

Questo concetto di generalizzazione è cruciale: un algoritmo di IA ben addestrato riesce ad applicare ciò che ha appreso dai dati di esempio a casi nuovi, fornendo risposte corrette anche quando riceve in input situazioni leggermente diverse da quelle viste in passato. Tornando al paragone con l'algoritmo tradizionale: se quest'ultimo incontra una situazione non prevista (come il programma della media a cui viene chiesto di calcolare una media di tre numeri senza essere stato programmato per farlo), fallirà o darà errore. Un algoritmo di IA, invece, avendo appreso dai dati, potrebbe gestire anche input mai visti prima in modo ragionevole, a patto che non si esca troppo dal dominio per cui è stato addestrato.

Un altro esempio concreto per evidenziare tale differenza può essere il filtro antispam nelle email. Molti di noi usufruiscono di filtri automatici che separano la posta indesiderata (spam) dalle email legittime. Un approccio tradizionale a questo problema richiederebbe di scrivere manualmente un elenco di regole: ad esempio, *“se l'oggetto dell'email contiene la parola 'vincita' oppure tutto il testo è in maiuscolo, probabilmente è spam”*. Si dovrebbero prevedere decine o centinaia di regole del genere, e comunque gli spammer (cioè chi invia spam) potrebbero trovare facilmente stratagemmi per aggirarle, ad esempio alterando le parole in modo creativo. Un filtro antispam basato su IA, invece, impara analizzando migliaia di email già classificate come spam o non spam. Il sistema di IA individua da sé schemi ricorrenti: può notare ad esempio che le email di spam spesso usano certe combinazioni di parole, certi indirizzi di provenienza, o particolari formati. Questi pattern possono essere molto complessi e non evidenti a occhio nudo, ma l'algoritmo li può cogliere mettendo a confronto enormi quantità di dati. Così, quando una nuova email arriva, l'algoritmo è in grado di valutare la probabilità che sia spam basandosi sulla somiglianza con gli esempi visti durante l'addestramento. Questo metodo tende ad essere molto più efficace e adattabile: se gli spammer cambiano tattica, il filtro basato su IA può aggiornarsi semplicemente continuando ad apprendere dai nuovi esempi di spam, senza dover essere riprogrammato da zero.

Da questo esempio emerge un altro aspetto importante: la capacità di adattamento. Gli algoritmi di IA possono migliorare le proprie prestazioni nel tempo. Più dati “vedono”, più diventano precisi (fino a un certo limite, ovviamente). Nei programmi tradizionali, invece, se vogliamo migliorare le prestazioni o aggiungere funzionalità, dobbiamo intervenire manualmente sul codice. Possiamo dire che gli algoritmi tradizionali sono statici, mentre quelli di IA sono dinamici. Questi ultimi possono cambiare il proprio comportamento in base all'esperienza accumulata.

.....

Differenze chiave tra un approccio tradizionale e uno basato su IA

Modalità di sviluppo: un algoritmo tradizionale viene scritto interamente dallo sviluppatore, specificando tutte le regole e i casi. Un algoritmo di IA viene in gran parte addestrato: lo sviluppatore definisce l'architettura del modello e l'obiettivo, ma lascia che sia il modello a ricavare le regole dai dati durante la fase di apprendimento.

Comportamento con nuovi dati: un algoritmo tradizionale applica pedissequamente le regole con cui è stato programmato e affronta male situazioni non previste esplicitamente. Un algoritmo di IA, avendo appreso dai dati, tende a essere più flessibile e a gestire anche casi nuovi, purché simili a quelli su cui è stato addestrato.

Aggiornamento e miglioramento: per migliorare o modificare un algoritmo tradizionale bisogna



riscrivere o aggiungere codice. Un algoritmo di IA può migliorare semplicemente alimentandolo con più dati aggiornati o affinando il processo di apprendimento, spesso senza cambiare la struttura del programma.

Deterministicità vs probabilità: molti algoritmi tradizionali forniscono risultati deterministici (gli stessi input producono sempre gli stessi output esatti). Alcuni algoritmi di IA, specialmente durante l'addestramento o in applicazioni generative, possono introdurre elementi probabilistici; inoltre, il loro risultato dipende fortemente dai dati appresi, quindi, in qualche modo "riflette" le statistiche di quei dati più che una logica fissa.

Trasparenza: in un algoritmo tradizionale è relativamente facile seguire il flusso di esecuzione e capire perché produce un certo risultato (basta leggere il codice passo passo). Al contrario, molti modelli di IA complessi (come le reti neurali profonde) sono spesso visti come una "scatola nera": anche i programmatori possono avere difficoltà a spiegare esattamente quale "ragionamento" interno abbia portato l'IA a prendere una certa decisione, poiché tale ragionamento è distribuito in migliaia o milioni di parametri numerici.

.....

Nella tabella 1.1 mettiamo a confronto gli algoritmi tradizionali con gli algoritmi di Intelligenza Artificiale rispetto ad alcuni aspetti chiave. Questo ci aiuta a capire come i programmi basati su regole fisse differiscono dai sistemi capaci di apprendere dagli esempi. I punti considerati includono la logica di funzionamento, l'adattabilità, la trasparenza, l'uso dei dati, la flessibilità e la capacità di miglioramento nel tempo.

Tabella 1.1 Algoritmi tradizionali vs algoritmi di IA

Aspetto	Algoritmi tradizionali	Algoritmi di IA
Logica di funzionamento	Seguono istruzioni programmate manualmente, con regole fisse e predefinite	Basati su modelli che <i>apprendono</i> dai dati e adattano il proprio funzionamento in base all'esperienza
Adattabilità	Bassa – il comportamento è statico: per cambiarlo serve riprogrammare o modificare il codice	Alta – il sistema può adattarsi a nuove situazioni man mano che riceve dati (apprendimento continuo, senza necessità di riprogrammare da zero)
Trasparenza	Alta – il funzionamento è generalmente comprensibile, poiché le regole e i passi logici sono espliciti e noti	Bassa – molti modelli di IA (es reti neurali profonde) sono "scatole nere", cioè operano in modo poco interpretabile dall'esterno
Uso dei dati	Utilizzano dati di input secondo regole statiche, senza imparare da essi (i dati non modificano il programma)	Richiedono grandi quantità di dati per l'addestramento; il modello impara pattern e informazioni direttamente dai dati e li utilizza per migliorare le prestazioni
Flessibilità	Limitata al compito specifico per cui sono progettati; se cambiano le condizioni o il tipo di problema, l'algoritmo tradizionale deve essere riscritto o riadattato manualmente	Maggiore nel gestire variazioni del problema all'interno del dominio appreso; può generalizzare l'esperienza per affrontare compiti simili, entro certi limiti
Migliorabilità	Migliora solo con intervento umano: per ottenere prestazioni migliori bisogna ottimizzare o aggiornare il codice a mano	Può migliorare <i>autonomamente</i> con l'esperienza: fornendo più dati o effettuando ulteriore training, il modello può affinare da solo le proprie prestazioni

Inoltre, lo sviluppo di un sistema basato su IA richiede un approccio differente: anziché programmare manualmente ogni funzionalità, si dedica molta attenzione alla raccolta di dati di qualità e alla fase di addestramento del modello. In un certo senso, “insegnare” a un’IA significa più prepararle un buon set di esempi da cui imparare, piuttosto che spiegarle passo passo cosa fare.

È importante notare che gli algoritmi tradizionali e quelli di IA non si escludono a vicenda, anzi spesso convivono. Un programma basato su IA, nel suo complesso, può includere parti sviluppate con algoritmi tradizionali (ad esempio per pre-processare i dati, o per gestire l’interfaccia utente) e componenti specifici di apprendimento automatico per le parti “intelligenti”. Inoltre, per alcuni problemi, gli approcci tradizionali rimangono i più indicati, specialmente se il problema è ben compreso e non troppo complesso da descrivere con regole. La scelta tra un algoritmo tradizionale e uno di IA dipende dal tipo di compito: se abbiamo un problema per cui è difficile scrivere tutte le regole manualmente ma abbiamo molti dati di esempi, l’IA è probabilmente la strada migliore; se invece il problema è chiaro e strutturato, un algoritmo classico può risultare più semplice, trasparente ed efficiente.

1.3 Tipologie di Intelligenza Artificiale (IA debole vs IA forte, IA generativa vs IA analitica)

Esistono vari modi di classificare i sistemi di intelligenza artificiale. In questo paragrafo ne esamineremo due in particolare: il primo riguarda l’ampiezza o generalità delle competenze dell’IA (contrapponendo la cosiddetta IA debole alla IA forte), il secondo riguarda la natura delle attività svolte dall’IA (distinguendo tra IA generativa e IA analitica). Queste classificazioni ci aiutano a capire meglio gli scopi e i limiti delle diverse applicazioni dell’intelligenza artificiale.

1.3.1 IA debole vs IA forte

Quando sentiamo parlare di intelligenza artificiale nella vita di tutti i giorni, quasi sempre si tratta di quella che gli esperti chiamano IA debole (in inglese Weak AI) o IA ristretta. Con questo termine non si vuole affatto dire che sia un’IA poco potente o efficace; “debole” in questo contesto significa semplicemente che è progettata per uno scopo specifico, limitato a un determinato dominio. Un’IA debole eccelle in un compito particolare, ma non ha capacità generali al di fuori di esso. Per esempio, un programma in grado di battere un campione umano a scacchi è un’IA debole: può analizzare milioni di mosse e strategie scacchistiche meglio di qualsiasi persona, ma non sa fare nient’altro al di fuori degli scacchi. Allo stesso modo, un sistema che riconosce i volti nelle fotografie è bravissimo in quello, ma non può all’improvviso imparare a guidare un’auto o a tradurre una lingua (a meno che non lo si riprogetti e riaddestri specificamente per quei compiti). La stragrande maggioranza delle applicazioni di IA attuali rientra in questa categoria: assistenti vocali come Siri o Alexa, software di diagnostica medica che identificano tumori in immagini radiografiche, algoritmi che consigliano canzoni o film in base ai nostri gusti, sistemi di traduzione automatica, videogiochi con intelligenza artificiale dei personaggi non giocanti, ecc. Tutte queste sono IA deboli perché ciascuna è altamente specializzata in un compito definito.

Per contro, il termine IA forte (in inglese spesso chiamata *Strong AI* o *Artificial General Intelligence*, AGI) indica un'ipotetica intelligenza artificiale generale, con capacità cognitive pari (o superiori) a quelle di un essere umano in qualsiasi ambito. Un'IA forte saprebbe non solo giocare a scacchi, ma anche parlare fluentemente qualsiasi lingua, comprendere barzellette, apprendere nuove abilità in autonomia, pianificare progetti creativi – insomma, avrebbe un'intelligenza flessibile e poliedrica, non confinata a un singolo compito. In aggiunta, alcuni definiscono davvero “forte” un'IA solo se dotata anche di coscienza o consapevolezza di sé, ovvero se fosse non solo intelligente ma anche senziente. Al momento attuale, un'IA del genere non esiste. I sistemi di IA che conosciamo e utilizziamo sono tutti, senza eccezioni, esempi di IA debole (narrow AI). L'IA forte per ora appartiene al regno della teoria o della fantascienza: molti ricercatori la studiano a livello concettuale e alcuni sono al lavoro per raggiungerla un giorno, ma non sappiamo se e quando ciò accadrà. Alcuni esperti ritengono che ci arriveremo nel corso di questo secolo, altri sono scettici e pensano che l'intelligenza umana abbia qualità uniche difficilmente replicabili del tutto in una macchina. La fantascienza ha spesso immaginato una tale IA forte: basti pensare a supercomputer come HAL 9000 nel film “2001: Odissea nello Spazio” o agli androidi dotati di emozioni in tanti romanzi e pellicole. Queste però sono, per ora, figure di fantasia.

Un esempio spesso citato per capire la differenza è il test di Turing. Proposto dal matematico Alan Turing nel 1950, questo test suggeriva che una macchina possa essere considerata “intelligente” (in senso generale) se, comunicando attraverso una tastiera, riesce a farsi passare per un essere umano agli occhi (anzi, alle orecchie) di un interlocutore umano. In altre parole, se un giudice umano fa conversazioni alla cieca sia con un computer che con una persona, e non riesce a distinguere chi è la macchina e chi l'umano dalle risposte che riceve, allora il computer avrebbe superato il test di Turing, mostrando un livello di intelligenza comparabile a quello umano. Finora nessun sistema ha superato pienamente questa prova in modo convincente, anche se recenti chatbot avanzati si sono avvicinati a sostenere conversazioni estremamente realistiche in linguaggio naturale. Va detto comunque che il test di Turing è un criterio piuttosto limitato (si basa solo sulla capacità di conversare) e non copre tutte le sfumature dell'intelligenza; inoltre, un programma potrebbe “ingannare” il giudice senza possedere una vera comprensione (potrebbe ad esempio utilizzare trucchi o risposte pre-programmate per sembrare umano). Dunque, superare il test di Turing non equivarrebbe necessariamente a creare un'autentica IA forte, ma rimane un concetto storico importante nell'ambito dell'IA.

.....
 Riassumendo la prima distinzione:

- **IA debole (ristretta):** intelligenza artificiale specializzata in un dominio o compito specifico. È la forma di IA oggi disponibile. Esempi: programmi di gioco (scacchi, Go, videogame), riconoscimento vocale e facciale, sistemi di raccomandazione (suggerimenti su Netflix, Amazon), assistenti virtuali, ecc. Non possiede conoscenza o flessibilità al di fuori dello specifico ambito per cui è stata progettata.
- **IA forte (generale):** intelligenza artificiale con abilità cognitive estese a qualunque dominio, paragonabile all'intelletto umano nella versatilità. Al momento è teorica e inesistente nella pratica. Implicherebbe la capacità di apprendere qualsiasi compito nuovo, trasferire conoscenze da un contesto all'altro e forse anche avere autocoscienza. Spesso descritta nella fantascienza come robot o computer “pensanti” alla pari o meglio degli esseri umani.

Tabella 1.2 Le principali differenze tra IA debole e IA forte

Caratteristica	IA debole (ristretta)	IA forte (generale)
Definizione	Intelligenza artificiale limitata a uno specifico campo o compito. È progettata per risolvere problemi circoscritti (non ha coscienza né comprensione generale)	Intelligenza artificiale con capacità cognitive generali paragonabili a quelle umane. Può teoricamente pensare, capire e apprendere <i>qualsiasi</i> tipo di compito in modo autonomo (simile a una mente umana, con autocoscienza)
Capacità	Eccelle nel compito per cui è stata addestrata, ma non va oltre quel dominio. Non è in grado di applicare le sue conoscenze a contesti differenti da quelli previsti	Avrebbe la capacità di ragionare e imparare in generale: potrebbe affrontare problemi nuovi in diversi ambiti, adattandosi e trovando soluzioni creative senza essere riprogrammata per ogni nuovo compito
Esempi	La maggior parte delle IA odierne: assistenti vocali come Siri o Alexa, software di scacchi o Go, sistemi di raccomandazione (es. suggerimenti di film), algoritmi di diagnostica medica specializzati	Al momento non esiste nella realtà; è un concetto teorico. Esempi immaginari di IA forte compaiono nella fantascienza, come HAL 9000 (il computer pensante di <i>2001: Odissea nello spazio</i>) o l'androide Data di <i>Star Trek</i>
Stato attuale	Realtà corrente – Tutte le applicazioni attuali di IA sono deboli/ristrette. L'IA oggi può superare l'uomo in compiti specifici, ma non ha intelligenza generale	Ipotesi futura – Non è stata ancora realizzata. Rimane un obiettivo della ricerca: non abbiamo ancora macchine con intelligenza paragonabile a quella umana in senso generale

1.3.2 IA generativa vs IA analitica

Le applicazioni di intelligenza artificiale possono essere classificate anche in base al tipo di output e al ruolo che hanno: si parla infatti di IA analitica quando il sistema si limita ad analizzare dati ed estrarne informazioni o decisioni, e di IA generativa quando il sistema è in grado di produrre nuovi contenuti autonomamente. Questa distinzione è diventata particolarmente rilevante negli ultimi anni, con l'ascesa di modelli di IA capaci di creare testi, immagini, musica e altro ancora.

L'IA analitica è l'IA tradizionale nel senso più comune, quella impiegata per esaminare dati, trovare correlazioni, fare previsioni e supportare decisioni. Un'IA analitica prende in ingresso dei dati e fornisce in uscita un'analisi o un'informazione derivata da quei dati. Non "inventa" nulla di proprio, piuttosto cerca di comprendere quello che c'è già. Esempi tipici di IA analitica includono: un modello che, dati anni di dati meteorologici, prevede il tempo di domani; un algoritmo che analizza il profilo di un cliente (età, acquisti precedenti, ricerche online) e determina se è a rischio di abbandonare un servizio; un sistema che monitora i macchinari in una fabbrica e segnala in anticipo possibili guasti sulla base di segnali anomali; un software che analizza le transazioni di una carta di credito per individuare quelle potenzialmente fraudolente. In tutti questi casi l'IA

macina dati e restituisce un risultato utile per capire meglio quei dati o per prendere decisioni operative. Storicamente, la maggior parte del *machine learning* è stato applicato in questo modo “analitico”: classificare cose, prevedere numeri, riconoscere elementi all’interno di dati complessi. Ad esempio, riconoscere che in una foto c’è un volto e identificare di chi si tratta è un compito analitico; altrettanto lo è tradurre una frase dall’italiano all’inglese (si analizza il testo di partenza e lo si converte, secondo ciò che il modello ha appreso, nel testo di arrivo).

L’**IA generativa** comprende sistemi di intelligenza artificiale capaci di creare qualcosa di nuovo che prima non c’era, sulla base di ciò che hanno imparato. Invece di restituire solo un’etichetta, un numero o una decisione, un’IA generativa può produrre un contenuto: testo, immagini, audio, video o altri dati. Questo non significa che l’IA “crei dal nulla” – ciò che genera è sempre ispirato a esempi assimilati durante l’addestramento – ma il risultato finale è originale nelle sue particolarità. Alcuni esempi impressionanti di IA generativa sono emersi di recente: modelli che, data una descrizione testuale (“un gatto astronauta che cammina sulla luna”), generano un’immagine fotorealistica corrispondente; oppure sistemi che producono testi articolati a partire da un semplice prompt (“scrivi una poesia sul mare in stile Dante Alighieri”). Ci sono IA generative in grado di comporre brani musicali in vari generi, di creare nuove ricette culinarie elaborando ingredienti esistenti, o di progettare modelli di oggetti tridimensionali. Un caso particolare molto noto di IA generativa sono proprio i LLM (*Large Language Models*), i grandi modelli di linguaggio, che vedremo in dettaglio nel prossimo paragrafo: essi possono generare testi molto elaborati, rispondere a domande, scrivere codice, riassumere documenti, il tutto producendo nuovo testo che sembra scritto da un umano.

L’IA generativa ha aperto frontiere del tutto nuove: se prima i computer aiutavano gli esseri umani principalmente elaborando più dati e più velocemente di noi, ora possono anche assisterci nei processi creativi. Ad esempio, un designer può farsi suggerire idee di design da un’IA generativa che propone bozze di immagini; uno scrittore può usare un modello linguistico per generare una bozza di racconto o per superare il “blocco dello scrittore”; un programmatore può farsi aiutare da un’IA generativa a scrivere porzioni di codice sorgente in un linguaggio di programmazione. Tutto questo era fantascienza fino a pochi anni fa. Naturalmente, la generazione automatica di contenuti solleva anche domande e problemi: la qualità e la correttezza dei contenuti generati non è sempre garantita (un’IA può scrivere cose scorrette o senza senso se non è ben controllata), c’è il rischio di violazione del diritto d’autore se il modello attinge troppo da materiale esistente, e in ambito educativo bisogna chiedersi come distinguere un compito scritto da uno studente da uno scritto da un computer. Affronteremo alcune di queste sfide più avanti nella sezione dedicata all’etica.

La tabella 1.3 confronta l’IA generativa e l’IA analitica in termini di obiettivo, tipo di output prodotto, esempi d’uso e principali modelli utilizzati.

Tabella 1.3 IA generativa vs IA analitica

Aspetto	IA generativa	IA analitica
Obiettivo	Creare dati o contenuti nuovi che non esistevano prima (inventare qualcosa di originale basandosi su ciò che ha appreso)	Analizzare dati esistenti per trovare pattern, estrarre informazioni utili e supportare decisioni (senza generare nuovi dati, ma interpretando quelli disponibili)
Output	Genera un nuovo output: può essere un testo, un'immagine, un brano musicale, una risposta in linguaggio naturale, codice, ecc., prodotto dal modello	Fornisce un risultato analitico: ad esempio una classificazione, una previsione, un punteggio, un allarme o una raccomandazione basati sui dati di input (non "crea" qualcosa di inedito, ma elabora ciò che c'è)
Esempi d'uso	<i>Generazione di testi:</i> chatbot avanzati e sistemi di scrittura assistita che producono articoli o racconti. <i>Generazione di immagini:</i> software che creano dipinti o immagini originali a partire da una descrizione (es. un programma che disegna paesaggi immaginari). <i>Creatività artificiale:</i> compositori musicali AI, generatori di ricette o di design innovativi	<i>Analisi educativa:</i> sistemi che analizzano le risposte degli studenti per identificare lacune di apprendimento. <i>Business intelligence:</i> modelli che esaminano dati di vendita per fare previsioni. <i>Rilevamento frodi:</i> algoritmi che analizzano transazioni bancarie e segnalano anomalie. In generale, qualsiasi applicazione che macina dati per offrire insight (dalla diagnostica medica alla manutenzione predittiva)
Principali modelli	Reti neurali generative specializzate, ad esempio: modelli di linguaggio come GPT (che producono testo in linguaggio naturale), reti generative avversarie (GAN) e modelli di diffusione per creare immagini, oppure auto-encoder variazionali. Spesso utilizzano architetture di Transformer (soprattutto per testo e sequenze)	Algoritmi di apprendimento automatico classico e reti neurali orientate all'analisi: alberi decisionali, random forest, reti neurali profonde per classificazione e regressione (es. una rete convolutiva CNN per riconoscere immagini), modelli statistici e sistemi di data mining tradizionali. L'enfasi è sull'accuratezza dell'analisi più che sulla creazione di qualcosa di nuovo

1.4 Machine learning e deep learning

Nel panorama dell'intelligenza artificiale moderna, due termini ricorrono continuamente: *machine learning* e *deep learning*. Spesso vengono usati quasi come sinonimi di "IA", poiché gran parte dei progressi recenti in questo campo sono dovuti proprio a questi approcci. Ma cosa significano esattamente?

1.4.1 Machine learning (apprendimento automatico)

Il *machine learning* – in italiano "apprendimento automatico" – è una branca dell'IA che si concentra sullo sviluppo di algoritmi in grado di imparare da dati ed esperienze,

migliorando le proprie prestazioni nel tempo. Invece di programmare una soluzione passo-passo per un problema, con il *machine learning* si forniscono al computer molti esempi (dati) affinché trovi da solo le regolarità e le regole che permettono di risolvere quel problema. È come insegnare a qualcuno a risolvere un compito mostrando tanti casi e soluzioni corrette, piuttosto che dando direttamente le istruzioni.

Il *machine learning* è già entrato nella nostra vita quotidiana: quando Netflix ci suggerisce una serie TV che potrebbe piacerci, lo fa perché ha “imparato” dai nostri gusti precedenti e da quelli di utenti simili a noi; quando uno smartphone riconosce il volto del proprietario per sbloccarsi, sta usando un modello addestrato su migliaia di volti; quando utilizziamo un filtro fotografico che automaticamente abbellisce l’immagine o la ritocca, c’è dietro un algoritmo che ha appreso dalle preferenze estetiche di molti utenti. Questi algoritmi di apprendimento automatico funzionano bene perché hanno visto tanti esempi e possono applicare quello che hanno appreso a situazioni nuove ma analoghe.

Esistono diversi “paradigmi” di apprendimento automatico, ossia modi in cui un algoritmo può essere addestrato. I tre principali sono:

- 1) **apprendimento supervisionato:** è il caso più comune. Si fornisce all’algoritmo un insieme di dati di esempio “etichettati”, cioè, corredati già della risposta corretta. Per esempio, potremmo avere mille fotografie di frutta, ognuna con un’etichetta che indica se nell’immagine c’è una mela, una banana, un’arancia, ecc. L’algoritmo apprende calcolando pian piano le correlazioni tra le caratteristiche delle foto (colori, forme, texture) e le etichette fornite (il tipo di frutto). Durante l’addestramento, ogni volta che l’algoritmo sbaglia (ad esempio scambia una banana per una mela), modifica leggermente i propri parametri interni per cercare di ridurre l’errore. Dopo molte iterazioni, convergerà verso un modello che commette pochi errori su quei dati. A quel punto, se gli mostriamo una nuova foto senza etichetta, idealmente riuscirà a riconoscere il frutto raffigurato basandosi su ciò che ha imparato. L’apprendimento supervisionato richiede dunque dati “giusti” già noti e funziona bene quando possiamo raccogliere molti esempi di input con la loro risposta attesa;
- 2) **apprendimento non supervisionato:** in questo approccio, all’algoritmo vengono forniti molti dati senza indicare quale sia la risposta corretta. L’obiettivo qui non è far corrispondere input e output predefiniti, ma scoprire da solo delle strutture nascoste nei dati. È come dare a qualcuno un mazzo di carte miste e chiedergli di ordinarle in gruppi come preferisce: quella persona potrebbe scoprire autonomamente che ci sono, ad esempio, carte con figure e carte numeriche, oppure carte rosse e nere. Un esempio classico di apprendimento non supervisionato è il “clustering”: l’algoritmo analizza, poniamo, i dati relativi ai clienti di un negozio (età, interessi, abitudini di acquisto) e li suddivide in gruppi omogenei, ad esempio “giovani urbani appassionati di tecnologia” vs “famiglie con bambini piccoli interessate a prodotti per l’infanzia”. Nessuno ha detto al programma quali gruppi trovare; è il modello che ha identificato schemi ricorrenti (*pattern*) e raggruppato i dati di conseguenza. Un altro esempio è l’analisi delle componenti principali, una tecnica che riduce la complessità dei dati trovando le “direzioni” principali in cui variano. L’apprendimento non supervisionato è utile per esplorare dati e trovare relazioni insospettite, e spesso serve anche come base per poi applicare metodi supervisionati o semplificare il problema;
- 3) **apprendimento per rinforzo:** si tratta di un paradigma ispirato al modo in cui impariamo molte cose nel mondo reale, ovvero per tentativi ed errori, ricevendo ricompen-

se o punizioni in base alle nostre azioni. Nell'apprendimento per rinforzo l'algoritmo (spesso chiamato in questo contesto "agente") interagisce con un ambiente compiendo delle azioni e, a seconda di ciò che fa, ottiene un rinforzo positivo (una ricompensa) o negativo (una penalità). L'obiettivo dell'agente è accumulare la maggiore ricompensa possibile nel lungo periodo, imparando quindi una strategia ottimale. Un classico esempio è un programma che impara a giocare a un videogioco: all'inizio muove il personaggio a caso, ma ogni volta che ottiene punti o avanza di livello riceve un feedback positivo, mentre se perde una vita o fallisce un obiettivo riceve un feedback negativo. Col tempo, il sistema attribuisce maggior valore alle azioni che hanno portato a ricompense e impara a replicarle, evitando invece le mosse che portano a punizioni. L'apprendimento per rinforzo ha permesso risultati straordinari: ad esempio, programmi che hanno imparato da soli a giocare a livelli sovrumani a giochi come scacchi e videogame classici, semplicemente giocando milioni di partite contro se stessi e ottimizzando le proprie strategie. È anche utilizzato in robotica, dove un robot può apprendere come camminare o come afferrare oggetti sperimentando varie azioni e ricevendo un punteggio in base a quanto si avvicina all'obiettivo (ad esempio camminare senza cadere, o afferrare un oggetto correttamente).

Questi tre approcci non si escludono a vicenda: esistono varianti e combinazioni, ad esempio l'apprendimento semi-supervisionato (dove alcuni dati hanno etichette e altri no), oppure l'apprendimento per rinforzo profondo (dove le tecniche di rinforzo vengono combinate con reti neurali profonde, come nel caso del famoso programma AlphaGo che ha battuto il campione mondiale di Go). L'idea generale da ricordare è che nel machine learning la chiave è sempre far sì che il programma adatti il proprio comportamento in base ai dati o all'esperienza, piuttosto che eseguire solo istruzioni fisse. Durante lo sviluppo di un sistema di ML, in genere si procede così: si raccolgono dati (tanti e di buona qualità), si sceglie o progetta un algoritmo di apprendimento, si addestra il modello su una porzione dei dati (detta *training set*), e poi si verifica come si comporta su dati nuovi che non ha mai visto (detti *test set* o parte di validazione). Questo permette di valutare se il modello ha veramente imparato a generalizzare o se si è solo "imparato a memoria" gli esempi (problema noto come *overfitting*, ovvero sovradattamento, quando un modello è troppo aderente ai dati di training e poi non va bene su dati nuovi). Se le prestazioni sono buone anche sul test set, allora l'algoritmo è considerato addestrato con successo e potrà essere utilizzato su dati reali.

1.4.2 Deep learning (apprendimento profondo)

Il **deep learning** è una sottocategoria del *machine learning* che ha avuto un ruolo rivoluzionario nell'ultimo decennio. La parola "deep" (profondo) si riferisce al fatto che questi algoritmi utilizzano reti neurali artificiali con molti strati (*layer*) di "neuroni" simulati. Già a partire dal nome, le reti neurali prendono ispirazione dal cervello umano: i neuroni biologici nel nostro cervello ricevono segnali dai neuroni vicini attraverso sinapsi, elaborano questi segnali, e se l'impulso è abbastanza forte lo trasmettono ad altri neuroni. In maniera molto semplificata, un neurone artificiale fa una cosa analoga in termini matematici: riceve alcuni valori in ingresso, li combina (ad esempio facendo una somma pesata: ogni ingresso ha un certo "peso" che amplifica o attenua il segnale), applica una funzione matematica sul risultato e produce un valore in uscita. Questo output può poi essere inviato come ingresso ad altri neuroni in cascata.

Una rete neurale è, dunque, un insieme di tanti neuroni artificiali interconnessi. Il primo strato (*layer*) della rete riceve l'input grezzo (ad esempio i pixel di un'immagine, o i valori dei sensori di un robot, o le parole di una frase da analizzare). L'ultimo strato produce l'output finale (ad esempio "gatto" vs "cane" nel caso del riconoscimento di immagini, oppure una traduzione in un'altra lingua, o una decisione sì/no). In mezzo ci sono vari strati "nascosti" che progressivamente trasformano e elaborano l'informazione, distillando concetti sempre più astratti. Per fare un paragone concreto: se la rete neurale deve riconoscere immagini di animali, i primi strati impareranno a riconoscere caratteristiche di basso livello come contorni, bordi, colori; gli strati intermedi combineranno queste caratteristiche base in forme e parti (un orecchio, un occhio, la coda); gli strati finali metteranno insieme queste parti per identificare l'oggetto completo (ad esempio riconosceranno che certe combinazioni di forme e colori corrispondono a un gatto). Questo processo di estrazione gerarchica delle caratteristiche è ciò che rende le reti neurali profonde così potenti: riescono a farsi da sole una sorta di "comprensione interna" dei dati, senza che qualcuno debba manualmente specificare quali caratteristiche guardare.

L'addestramento di una rete neurale profonda avviene ancora secondo i principi dell'apprendimento supervisionato (o altri approcci, a seconda dei casi). Immaginiamo di addestrare una rete per riconoscere dei suoni: inizialmente i pesi di tutti i neuroni sono impostati in modo casuale; quindi, la rete all'inizio darà risposte totalmente errate o casuali. Ma poi confrontiamo l'output con la risposta corretta attesa (ad esempio, sappiamo che questo suono era la parola "ciao" e non "arrivederci") e valutiamo l'errore. Un algoritmo di ottimizzazione (spesso il metodo di *back propagation* con discesa del gradiente) modifica un po' tutti i pesi dei neuroni in modo da ridurre quell'errore. Dopodiché si presenta un altro esempio alla rete, si calcola l'errore, si aggiornano i pesi, e così via, iterativamente. Dopo aver passato migliaia o milioni di esempi, la rete aggiusta i suoi parametri al punto da dare quasi sempre la risposta giusta su quelli già visti. L'importante però è che grazie alla struttura a più strati, essa ha appreso una rappresentazione generale che – se tutto è andato bene – funziona anche su esempi nuovi. Quando poi la rete sarà usata "sul campo", cioè in applicazione reale, non farà altro che eseguire i calcoli strato per strato con i pesi appresi per fornire un risultato.

Il *deep learning* ha permesso di risolvere problemi che per lungo tempo erano rimasti fuori portata per l'IA. Ad esempio, il riconoscimento vocale e la sintesi vocale (far parlare un computer con voce naturale) hanno fatto passi da gigante grazie alle reti neurali profonde, migliorando l'accuratezza al punto che servizi come assistenti vocali capiscono comandi e domande con pochissimi errori. Il riconoscimento delle immagini è passato da tassi di errore molto alti a prestazioni superiori a quelle umane in alcune competizioni: emblematico è il caso di un concorso del 2012 (*ImageNet*), in cui una rete neurale profonda superò di gran lunga tutti gli altri metodi nel riconoscere oggetti in foto, segnando l'inizio dell'odierna ondata di *deep learning*. Anche la traduzione automatica tra lingue, che un tempo era basata su frasi fatte e regole grammaticali codificate a mano, ora avviene tramite modelli neurali: ciò ha migliorato notevolmente la fluidità e correttezza delle traduzioni su servizi come Google Traduttore. In generale, ogni volta che uno smartphone identifica ciò che inquadra la fotocamera, o un'auto a guida autonoma interpreta i dati dei suoi sensori per prendere decisioni in tempo reale, c'è quasi sicuramente di mezzo un sistema di *deep learning*.

Un altro aspetto interessante del *deep learning* è la sua capacità di generare contenuti, non solo analizzarli. Abbiamo già accennato all'IA generativa: molte tecniche di IA ge-

nerativa sfruttano reti neurali profonde. Ad esempio, le GAN (*Generative Adversarial Networks*) sono un tipo di rete neurale dove due sub-reti collaborano/competono per produrre dati simili a quelli reali (una genera esempi e l'altra li valuta, spingendo la prima a migliorare sempre più); questa idea ha portato a enormi progressi nella generazione di immagini realistiche e di volti di fantasia indistinguibili da foto vere. Un altro approccio, i modelli a diffusione, ha rivoluzionato ulteriormente la generazione di immagini attraverso algoritmi in grado di trasformare rumore casuale in figure dettagliate seguendo un testo descrittivo. E ovviamente i modelli linguistici di grandi dimensioni (LLM) come GPT-3 e GPT-4 sono anch'essi reti neurali molto profonde (composte da decine di strati) specializzate nel produrre testo. Tutto questo per dire che il *deep learning* non solo consente al computer di capire meglio i dati, ma anche di creare.

È importante sottolineare che, pur essendo potentissime, le reti neurali profonde hanno anche dei limiti. In primo luogo, come accennato, tendono a funzionare come "scatole nere": il loro processo decisionale interno è difficile da interpretare, e questo può essere un problema quando si richiede trasparenza (si pensi a un algoritmo che decide un prestito bancario: sarebbe importante capire perché nega o concede il credito, ma se la decisione viene da una rete neurale profonda, non è immediato spiegare il motivo in termini semplici). In secondo luogo, richiedono enormi quantità di dati per addestrarsi efficacemente: se i dati sono pochi o poveri di qualità, spesso un modello più semplice potrebbe paradossalmente performare meglio di una rete profonda che rischierebbe di "imparare male". Inoltre, l'addestramento di modelli *deep learning* è molto costoso in termini computazionali: servono hardware specializzati (GPU, unità di elaborazione grafica, e TPU, unità di elaborazione per tensori) e molto tempo di calcolo, cosa che rende queste soluzioni impegnative e non alla portata di tutti se si parte da zero. Infine, come tutti i modelli di *machine learning*, anche il *deep learning* soffre se il contesto cambia troppo rispetto a quello dei dati di addestramento (ad esempio, un sistema di visione allenato su foto diurne potrà avere difficoltà con immagini notturne se non è stato addestrato anche per quelle).

Nonostante questi limiti, il *machine learning* – e in particolare il *deep learning* – rimane il motore principale dell'IA attuale. Grazie ad esso, l'IA è passata da sistemi rigidi e limitati a sistemi capaci di risultati sorprendenti.

1.5 LLM (Large Language Models) e il loro funzionamento

Tra le applicazioni più sorprendenti e discusse dell'intelligenza artificiale degli ultimi tempi ci sono i cosiddetti LLM (*Large Language Models*), ovvero grandi modelli linguistici. Si tratta di sistemi di IA addestrati per comprendere e generare linguaggio naturale (il linguaggio umano) su scala vastissima. Esempi noti di LLM includono modelli come GPT (*Generative Pre-trained Transformer*) sviluppato da OpenAI, il modello PaLM di Google, i modelli LLaMA di Meta, e molti altri in rapida evoluzione. Questi modelli sono alla base di strumenti come i chatbot avanzati (ad esempio ChatGPT), capaci di sostenere conversazioni articolate, rispondere a domande e produrre testi di sorprendente coerenza. Questa classe di modelli ha visto progressi rapidissimi: fino al 2018-2019, i testi generati automaticamente erano spesso sconclusionati o molto brevi; nel 2020 il modello GPT-3 ha stupito per la capacità di produrre testi lunghi e coerenti; e alla fine del 2022 l'apertura al pubblico di ChatGPT (basato su un modello della famiglia GPT-3.5) ha fatto scoprire a milioni di persone il potenziale di questa tecnologia.

1.5.1 Cosa rende “grandi” questi modelli linguistici?

Questi modelli linguistici si caratterizzano per la quantità di parametri che li compongono, e in secondo luogo per la mole di dati testuali su cui vengono addestrati. I parametri di un modello neurale sono essenzialmente i numeri (pesi) che il modello impara durante l’addestramento e che determinano come trasforma l’input nell’output. Un LLM può avere centinaia di miliardi di parametri: per confronto, un modello di elaborazione del linguaggio di qualche anno fa ne aveva forse qualche centinaio di milioni. Questo aumento colossale delle dimensioni permette al modello di immagazzinare (in forma di schemi statistici, non di memorizzazione letterale) una quantità impressionante di conoscenza estratta dai dati di training. I dati di addestramento di un LLM spesso comprendono gran parte del contenuto testuale disponibile su Internet: Wikipedia intera, libri digitalizzati, articoli di giornale, siti web, conversazioni pubbliche e così via. In pratica, durante la fase di training, al modello vengono presentati innumerevoli frasi e documenti, e viene chiesto di predire la parola (o meglio il “token”, un’unità linguistica che può essere una parola o porzione di parola) successiva. Facendo questo esercizio ripetutamente su miliardi di esempi, il modello affina progressivamente la sua capacità di indovinare le parole seguenti in un testo.

Può sorprendere, ma dal semplice compito di prevedere la parola successiva scaturisce un comportamento intelligentemente vario. Pensiamo a cosa vuol dire continuare un testo in modo plausibile: per farlo, il modello deve aver catturato la sintassi della lingua (sapere come si costruiscono frasi grammaticalmente corrette), deve avere conoscenza del mondo (se il testo parla di “Parigi”, il modello deve capire che probabilmente c’entra la Francia o l’Europa, ecc., per continuare coerentemente), deve cogliere il tono e il contesto (un conto è continuare una fiaba, un conto è rispondere con un tono formale a una email di lavoro). In effetti, un LLM ben addestrato “assorbe” dai dati miriadi di informazioni: fatti enciclopedici, nozioni di base di matematica, persino elementi di logica e di buon senso, perché tutti questi aspetti sono riflessi in qualche modo nei testi su cui è stato addestrato.

Per capire come funziona, immaginiamo di iniziare a scrivere una frase: “*Nel mezzo del cammin di nostra vita mi ritrovai in una selva oscura...*”. Un LLM che abbia letto molta letteratura riconoscerà questo incipit famoso e saprà probabilmente continuare citando la Divina Commedia di Dante, oppure spiegando che si tratta di Dante se il contesto della richiesta lo richiede. Se, invece, forniamo un inizio generico come: “Oggi ho portato il mio cane al parco e...”, un LLM dovrà capire il contesto: parleremo di un evento quotidiano, forse seguirà qualcosa come “...e abbiamo incontrato altri amici con cui ha giocato felicemente.” Queste continuazioni non sono memorizzate in un database, vengono create sul momento combinando in modo nuovo i modelli appresi. Il meccanismo interno, semplificando molto, è questo: il modello prende il testo di input (il nostro prompt o la conversazione fino a quel punto), lo analizza e ne crea una rappresentazione interna numerica; quindi, in base a quella e ai suoi parametri allenati, calcola quale sia il token più probabile per continuare la sequenza. Supponiamo che abbia un 80% di probabilità per la parola “giocato”, 15% per “corsi”, 5% per altre opzioni – può scegliere la più probabile (“giocato”) e aggiungerla al testo. Poi ripete il processo per la parola successiva, aggiornando il contesto man mano. Ovviamente in realtà i calcoli coinvolgono spazi dimensionali enormi e l’uso di un’architettura di rete neurale chiamata *transformer*, che è stata fondamentale per permettere tutto ciò. L’architettura *transformer* utilizza un mec-

canismo detto di “auto-attenzione” grazie al quale il modello valuta, per ogni parola in ingresso, quali altre parole del contesto sono più rilevanti e “a cosa prestare attenzione” – un po’ come noi, leggendo un testo, capiamo quali parole sono collegate tra loro. Questa tecnologia, introdotta intorno al 2017, ha reso molto più efficiente l’addestramento di modelli linguistici di grandi dimensioni e ha soppiantato in gran parte le precedenti architetture per compiti di linguaggio.

Il risultato di questo processo di apprendimento è che gli LLM sviluppano una sorprendente capacità di generare testo che sembra scritto da un umano. Possono rispondere a domande su praticamente qualsiasi argomento per cui abbiano informazioni nei dati di training; possono svolgere compiti di composizione come scrivere saggi, poesie, articoli, riassunti; possono tradurre tra lingue; possono persino scrivere un codice informatico se addestrati su testi di programmazione. È importante notare, però, che “capire” e “generare” linguaggio non è la stessa cosa che “ragionare” come un essere umano. Un LLM non ha coscienza, intenzioni proprie o vera comprensione del significato ultimo delle parole: manipola simboli (parole o frammenti di parole) basandosi sui pattern statistici appresi. Ciò significa che a volte può sembrare estremamente intelligente e ragionevole, e altre volte può cadere in errore o dire cose assurde se nel suo training c’erano lacune o se la domanda lo porta fuori dai casi per cui ha appreso buone correlazioni. Ad esempio, può succedere che un LLM fornisca con grande sicurezza un’informazione completamente inventata: magari risponde che “il presidente X nel 1975 ha detto Y...” quando in realtà non c’è alcuna fonte reale per quella citazione – il modello l’ha generata perché suonava plausibile. Questo fenomeno è noto come “allucinazione” dell’IA. Succede perché il modello non ha un meccanismo infallibile per distinguere il vero dal falso: se combinando parole gli sembra statisticamente appropriato affermare qualcosa, lo farà, a meno che non sia stato istruito a evitare certe risposte.

Nonostante questi limiti, la capacità degli LLM di produrre testo utile ha moltissime applicazioni. Si pensi a uno studente o un insegnante: un modello di linguaggio può essere utilizzato come supporto per spiegare concetti difficili (“spiegami la relatività come se avessi 12 anni”), per fornire spunti e idee (“dammi alcuni suggerimenti per un tema sull’ambiente”), per tradurre o riassumere articoli in altre lingue, o ancora per creare esercizi e quiz su un certo argomento. Tutto questo avviene semplicemente interagendo in linguaggio naturale con il modello, come si farebbe con un assistente personale molto informato. Alcune piattaforme educative stanno iniziando a integrare LLM per offrire tutor virtuali che rispondono alle domande degli studenti a qualsiasi ora, fornendo spiegazioni aggiuntive su una lezione. Anche al di fuori della scuola, questi modelli trovano impiego: professionisti li consultano per ottenere riassunti di report chilometrici, programmatori chiedono aiuto nella scrittura di codice (esistono sistemi come GitHub Copilot che suggeriscono porzioni di codice mentre si programma, basati su modelli linguistici addestrati su enormi basi di codice sorgente).

Facciamo un piccolo esempio di interazione con un LLM per dare un’idea concreta:

Utente: “Puoi creare una breve filastrocca sul ciclo dell’acqua per favore?”

Modello IA: “Certamente!

La pioggia scende lieve dal cielo al mare,

il sole la porta di nuovo a evaporare.

Si formano le nuvole nel cielo su in alto,

e torna la pioggia, di nuovo un salto!

Così l'acqua viaggia e torna sempre qua,
da nuvola a pioggia, che magia sarà.”

In questo dialogo simulato, l'utente ha chiesto in italiano una filastrocca sul ciclo dell'acqua, e il modello ha prodotto quattro versi in rima che descrivono il fenomeno in modo semplice e poetico. Questo è un tipico esempio di come un LLM possa generare contenuti creativi e adatti a un certo contesto su richiesta. Naturalmente, la qualità delle risposte può variare: non sempre la rima o il contenuto saranno perfetti, ma è impressionante constatare che provengono interamente da un modello matematico senza che qualcuno abbia scritto esplicitamente quella filastrocca.

Un aspetto cruciale da evidenziare è che gli LLM sono frutto di un addestramento massiccio e costoso. Addestrare modelli come GPT-4 ha richiesto centinaia di GPU funzionanti per settimane o mesi, con un costo di milioni di dollari. Questo fa sì che solo grandi organizzazioni o consorzi dispongano delle risorse per creare da zero un LLM di punta. Tuttavia, una volta addestrato, un modello può essere in parte riutilizzato o specializzato su compiti più specifici con costi molto minori, tramite una tecnica chiamata *fine-tuning* (affinamento). Ad esempio, un'azienda può prendere un LLM generale e fare un *fine-tuning* sui propri dati aziendali (documenti interni, manuali, ecc.) in modo che diventi esperto del dominio specifico e risponda come un assistente specializzato in quell'area. Questa flessibilità rende gli LLM strumenti molto versatili che possono essere adattati a molti scenari.

Una direzione interessante nello sviluppo attuale è integrare i LLM con fonti esterne di informazione e strumenti. Ad esempio, esistono assistenti avanzati che uniscono un modello linguistico a un motore di ricerca: quando l'utente fa una domanda su un fatto recente o molto specifico, l'IA può prima effettuare una ricerca sul web per raccogliere informazioni aggiornate, e poi formulare la risposta combinando quei dati con le proprie capacità di linguaggio naturale. In questo modo si mitigano alcune limitazioni dei modelli (come la mancanza di conoscenza sugli eventi successivi alla loro fase di addestramento) e si migliora l'accuratezza delle risposte, pur mantenendo la fluidità e la chiarezza espressiva. È probabile che in futuro vedremo sempre più spesso sistemi di questo tipo, dove l'IA dialogante è collegata a database, strumenti di calcolo o altri servizi, ampliando così ciò che può fare in modo controllato.

Allo stesso tempo, poiché questi modelli apprendono dai dati disponibili, riflettono anche ciò che c'è nei dati: se una certa idea, giusta o sbagliata che sia, è molto presente nei testi di training, il modello tenderà a riproporla. Per questo, *bias* (pregiudizi) e informazioni scorrette presenti nei dati di addestramento possono influenzare le risposte degli LLM. Le aziende che sviluppano questi modelli stanno lavorando su metodi per mitigare il problema, come l'inclusione di filtri e di feedback umani durante l'addestramento per evitare che il modello produca contenuti inappropriati (offese, incitamenti all'odio, ecc.) o che abbia bias discriminatori. Ad esempio, a ChatGPT e modelli simili viene insegnato esplicitamente a rifiutare richieste nocive o a rispondere in modo neutrale e informativo su temi delicati, proprio per ovviare agli effetti indesiderati di un addestramento su dati “di internet” che, in quanto specchio dell'umanità, contengono sia il meglio sia il peggio delle opinioni.

In conclusione, i *Large Language Models* rappresentano uno dei traguardi più avanzati dell'IA attuale: macchine in grado di gestire il linguaggio in modo sofisticato e creare risposte testuali articolate. Capire il loro funzionamento ci fa anche comprendere i loro

punti di forza (l'enorme base di conoscenze implicite, la versatilità) e le loro debolezze (mancanza di vero senso critico, dipendenza dai dati, potenziali bias). Nel contesto didattico, come in ogni altro contesto, è fondamentale usarli con consapevolezza: possono offrire un aiuto straordinario, ma vanno guidati e sorvegliati per assicurarsi che forniscano informazioni corrette e appropriate. Nel prossimo e ultimo paragrafo di questo capitolo vedremo, infatti, quali sono le sfide e le questioni etiche legate all'uso dell'IA, inclusi i modelli linguistici, nell'educazione e nella società in generale.

La tabella seguente mette a confronto ML, DL e LLM per definizione, tecniche, complessità, ambiti d'uso ed esempi.

Tabella 1.4 Confronto tra ML, DL e LLM

Aspetto	<i>Machine Learning (ML)</i>	<i>Deep Learning (DL)</i>	<i>Large Language Model (LLM)</i>
Definizione	Approccio dell'IA in cui i modelli apprendono dai dati come svolgere un compito, anziché seguire solo istruzioni fisse. Include metodi vari (dal semplice al complesso) che estraggono pattern dai dati per fare previsioni o decisioni	Sottoinsieme specializzato del ML che utilizza reti neurali artificiali con molti strati (reti "profonde"). Queste reti imparano rappresentazioni astratte dei dati attraverso diversi livelli, consentendo di affrontare compiti complessi (vision, linguaggio, ecc.) con alte prestazioni	Modelli di <i>deep learning</i> specializzati nel linguaggio naturale, con moltissimi parametri (spesso miliardi). Sono addestrati su enormi collezioni di testo per prevedere e generare linguaggio in modo coerente. In pratica, apprendono la statistica delle parole e frasi per poter dialogare, tradurre, rispondere a domande, ecc
Tecniche usate	Comprende algoritmi supervisionati (ad es. regressione lineare, alberi decisionali, SVM), non supervisionati (<i>cluster analysis</i> , riduzione della dimensionalità) e apprendimento per rinforzo. Spesso richiede <i>feature engineering</i> : bisogna selezionare manualmente le caratteristiche rilevanti dei dati da dare in pasto all'algoritmo	Si basa soprattutto su reti neurali profonde: esempi comuni sono le CNN (<i>Convolutional Neural Network</i>) per elaborare immagini, le RNN/LSTM per dati sequenziali, e più recentemente i <i>transformer</i> (un'architettura di rete neurale) utilizzati per testo, immagini e altro. Il modello impara automaticamente le caratteristiche rilevanti dai dati grezzi, senza necessitare di <i>feature engineering</i> manuale	Utilizzano in particolare l'architettura <i>transformer</i> su scala molto ampia. L'addestramento è spesso auto-supervisionato: il modello impara a predire la parola successiva in una frase, o a completare frasi, elaborando così la struttura del linguaggio. Impiegano tecniche avanzate di ottimizzazione distribuita, data la mole di dati e parametri, e spesso vengono poi <i>fintunati</i> (<i>fine-tuned</i>) su compiti specifici di linguaggio

(segue)

Aspetto	Machine Learning (ML)	Deep Learning (DL)	LLM (Large Language Model)
Livello di complessità	Variabile, generalmente inferiore al DL: i modelli ML possono essere semplici e interpretabili (una regressione lineare ha pochi parametri) oppure più complessi (un ensemble di alberi). In generale, richiedono meno potenza di calcolo rispetto al deep learning e sono più facili da analizzare e spiegare	Elevato: reti profonde con milioni (o anche miliardi) di pesi richiedono notevole potenza computazionale (GPU) e grandi quantità di dati per l'addestramento. Il modello risultante è spesso una "scatola nera" poco interpretabile. La complessità è sia nell'architettura sia nel tuning di molti iperparametri	Estremamente elevato: un LLM può avere <i>decine di miliardi</i> di parametri, richiedendo supercomputer per l'addestramento. Gestire e distribuire tali modelli è complesso anche in fase di utilizzo (servono risorse hardware notevoli). Sono tra i modelli meno interpretabili in assoluto, data la dimensione e la natura dell'apprendimento
Ambiti applicativi	Vastissimo: il ML è impiegato in quasi ogni settore. Esempi: filtrare email spam, sistemi di raccomandazione (film, corsi online per studenti), diagnosi mediche assistite, analisi finanziarie, previsione di fenomeni (dal meteo all'andamento degli iscritti a un corso). In sintesi, ovunque ci siano dati da analizzare per prendere decisioni	Compiti che richiedono riconoscere pattern complessi: visione artificiale (riconoscimento di volti, oggetti, scrittura a mano), elaborazione vocale (riconoscimento e sintesi della voce), NLP - elaborazione del linguaggio naturale (traduzione automatica, analisi del sentiment). Nella didattica, ad esempio, DL può alimentare tutor intelligenti che analizzano le risposte aperte degli studenti o sistemi di trascrizione automatica di lezioni	Tutto ciò che riguarda il linguaggio naturale: chatbot e assistenti virtuali avanzati (per conversare con utenti o studenti in linguaggio naturale), motori di ricerca e strumenti di completamento testo, traduttori automatici, sistemi di domanda-risposta, creazione di riassunti o contenuti. Nel contesto educativo, LLM possono fungere da tutor personalizzati conversazionali, rispondere a domande degli studenti o generare materiali didattici (con la dovuta supervisione)
Esempi	Un filtro anti-spam che impara a riconoscere email indesiderate. Un algoritmo che analizza i quiz di uno studente e suggerisce argomenti su cui fare ripasso. Anche un sistema di <i>credit scoring</i> bancario o di previsione delle vendite rientrano nel ML	Una rete neurale convolutiva che identifica oggetti o persone in una foto di classe. Il sistema di riconoscimento vocale di Siri o Google Assistant (che trascrive la voce in testo) è basato su <i>deep learning</i> . Anche le auto a guida autonoma utilizzano profondamente il DL per interpretare l'ambiente	GPT-4 / ChatGPT, che genera risposte in linguaggio naturale a qualsiasi domanda (LLM di OpenAI). BERT (di Google), un modello utilizzato per comprendere il significato di frasi e migliorare i risultati di ricerca. Altri esempi: PaLM, GPT-3, o modelli open-source come LLaMA, tutti LLM addestrati su grandissime quantità di testo

i quaderni della DIDATTICA

Rivolta a chi già insegna o desidera intraprendere la professione di docente ma anche ai candidati a corsi di specializzazione e studenti universitari, la collana contiene volumi dedicati ai principali strumenti teorici e operativi della didattica, la cui acquisizione costituisce un aspetto fondamentale della professione di insegnante.

Questo volume offre agli insegnanti e ai dirigenti scolastici un quadro organico e aggiornato sulle molteplici dimensioni dell'intelligenza artificiale applicata alla scuola.

Dalle definizioni di base fino alle applicazioni pratiche in classe, il testo accompagna passo dopo passo nella scoperta di un mondo complesso ma affascinante. Particolare attenzione è dedicata:

- agli **aspetti etici, giuridici e di responsabilità** legati all'uso dell'IA;
- alla presentazione dei **principali software di intelligenza artificiale generativa** e ai criteri per valutarne l'affidabilità;
- alle **metodologie didattiche innovative** che consentono di integrare l'IA nella progettazione curricolare, nella valutazione e nella personalizzazione dei percorsi di apprendimento.

Il testo si configura come una **guida teorico-pratica** che permette al docente di orientarsi tra opportunità e criticità, offrendo strumenti concreti per sviluppare competenze digitali e utilizzare l'IA in modo consapevole, responsabile e creativo nella scuola.

Tra le **estensioni online**, una vasta raccolta di *Unità di Apprendimento* esemplificative, utili come modelli di integrazione tra intelligenza artificiale e lavoro educativo, o come pratico "promptuario" (raccolta di prompt) ricco di idee e suggerimenti pronti all'uso.

Una risorsa indispensabile per chi desidera affrontare con rigore e competenza la trasformazione digitale della didattica.



IN OMAGGIO
ESTENSIONI ONLINE

Contenuti
extra

Le **risorse di studio** gratuite sono accessibili per 18 mesi dalla propria area riservata, previa registrazione al sito **edises.it**.



blog.edises.it
infoconcorsi.edises.it



€ 28,00

