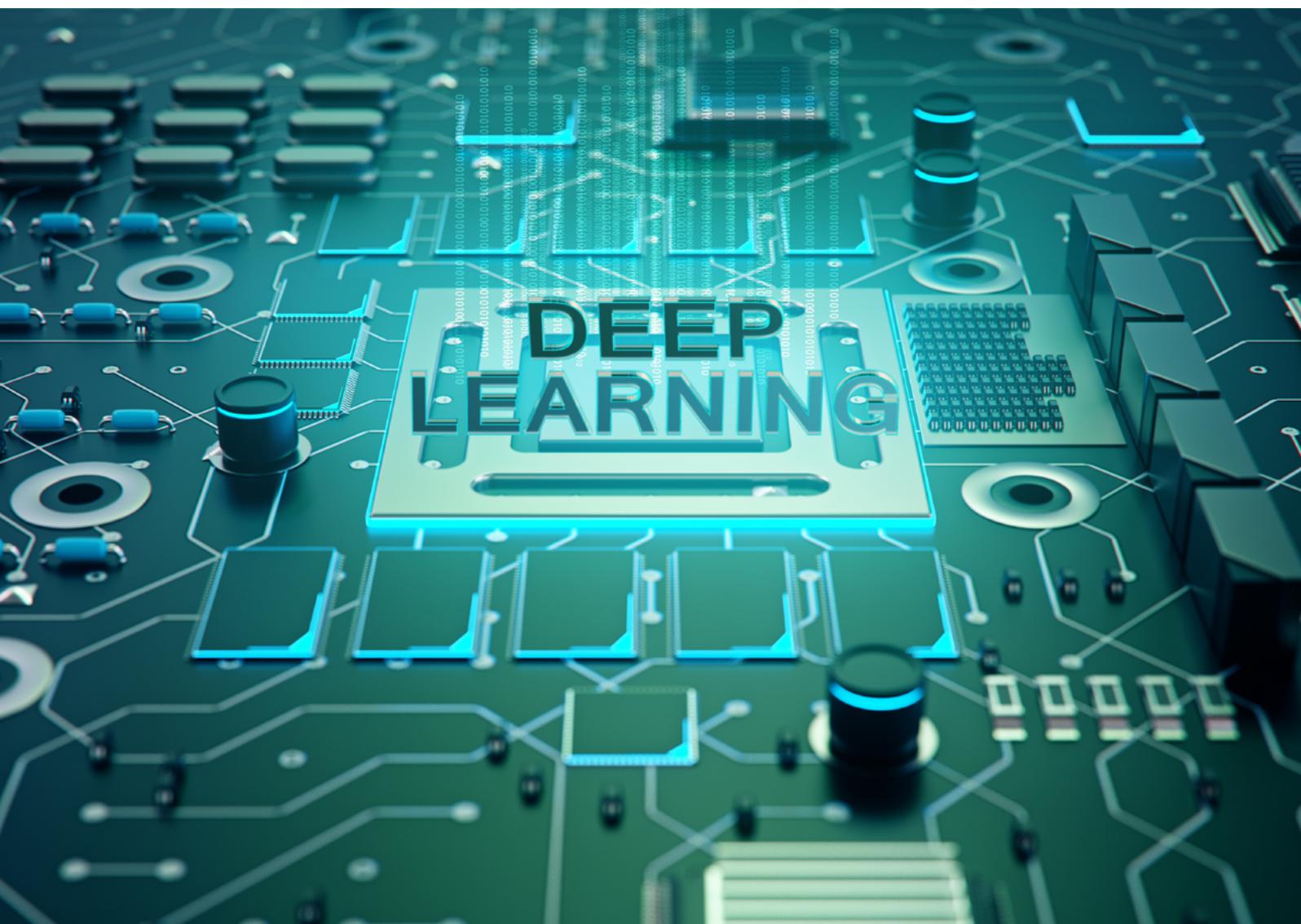


Deep Learning, cos'è l'apprendimento profondo, come funziona e quali sono i casi di applicazione

Testi a cura della redazione AI4Business

[LEGGI SUL SITO](#)





INDICE DEGLI ARGOMENTI

1. Cos'è il Deep Learning	5
2. Le reti neurali artificiali, la base del Deep Learning	
3. Come funziona il Deep Learning	6
4. Come si addestra un sistema di Deep Learning	10
5. Deep Learning Framework: perché quello di Facebook sta attirando l'attenzione	12
6. I casi d'uso e i tipi di applicazioni del Deep Learning	18
6.1 Driverless car	
6.2 Cinema	
6.3 Quantum Intelligence	
6.4 Ologrammi	

*La traduzione letterale è **apprendimento profondo** ma il Deep Learning, sotto categoria del [Machine Learning](#) e del più ampio mondo dell'[Intelligenza Artificiale](#), sottende a qualcosa di molto più ampio del “semplice” apprendimento su più livelli delle macchine. Cerchiamo allora di capire cos'è il Deep Learning, come funziona e che tipo di applicazioni può avere.*

**360ONPODCAST****Intelligenza artificiale: cos'è e com'è nata**



COS'È IL DEEP LEARNING

Il Deep Learning, la cui traduzione letterale significa **apprendimento profondo**, è una sottocategoria del Machine Learning (che letteralmente viene tradotto come apprendimento automatico) e indica quella branca dell'intelligenza artificiale che fa riferimento agli **algoritmi ispirati alla struttura e alla funzione del cervello** chiamate **reti neurali artificiali**.

Da un punto di vista scientifico, potremmo dire che il Deep Learning è l'apprendimento da parte delle "macchine" attraverso dati appresi grazie all'utilizzo di algoritmi (prevalentemente di calcolo statistico).

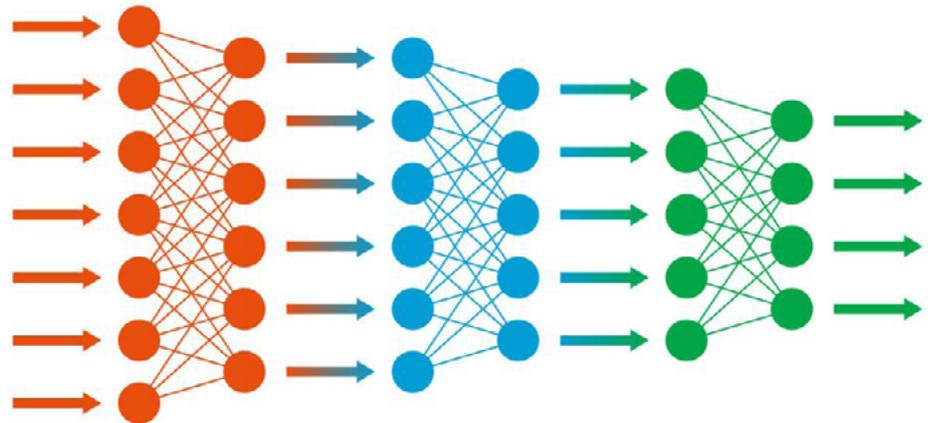
Il Deep Learning (noto anche come apprendimento strutturato profondo o apprendimento gerarchico), infatti, fa parte di una più ampia famiglia di metodi di Machine Learning basati sull'assimilazione di rappresentazioni di dati, al contrario di algoritmi per l'esecuzione di task specifici.

Le architetture di Deep Learning (con le quali oggi si riporta all'attenzione anche del grande pubblico il concetto di rete neurale artificiale) sono per esempio state applicate nella **computer vision**, nel **riconoscimento automatico della lingua parlata**, nell'**elaborazione del linguaggio naturale**, nel riconoscimento audio e nella **bioinformatica** (l'utilizzo di strumenti informatici per descrivere dal punto di vista numerico e statistico determinati fenomeni biologici come le sequenze di geni, la composizione e la struttura delle proteine, i processi biochimici nelle cellule, ecc.).

Raccogliendo le diverse interpretazioni di alcuni tra i più noti ricercatori e scienziati nel campo dell'apprendimento profondo (quali: **Andrew Yan-Tak Ng**, fondatore di Google

Brain, Chief Scientist di Baidu e professore e Director dell'AI Lab alla Stanford University; **Ian J. Goodfellow**, ricercatore riconosciuto come uno dei migliori innovatori del mondo – under 35 – dal MIT di Boston; **Yoshua Bengio**, uno degli scienziati più riconosciuti proprio nel campo del Deep Learning; **Ilya Sutskever**, Research Director di OpenAI; **Geoffrey Everest Hinton**, una delle figure chiave del Deep Learning e dell'Intelligenza artificiale, primo ricercatore ad aver dimostrato l'uso di un algoritmo di backpropagation generalizzato per l'addestramento di reti neurali multistrato...), potremmo definire il Deep Learning come un sistema che sfrutta una classe di algoritmi di apprendimento automatico che:

- 1) usano **vari livelli di unità non lineari a cascata per svolgere compiti di estrazione di caratteristiche e di trasformazione**. Ciascun livello successivo utilizza l'uscita del livello precedente come input. Gli algoritmi possono essere sia di tipo supervisionato sia non supervisionato e le applicazioni includono l'analisi di pattern (apprendimento non supervisionato) e classificazione (apprendimento supervisionato);
- 2) sono basati sull'**apprendimento non supervisionato di livelli gerarchici multipli di caratteristiche (e di rappresentazioni) dei dati**. Le caratteristiche di più alto livello vengono derivate da quelle di livello più basso per creare una rappresentazione gerarchica;
- 3) fanno parte della più ampia classe di algoritmi di apprendimento della rappresentazione dei dati all'interno dell'apprendimento automatico (**Machine Learning**);
- 4) apprendono multipli livelli di rappresentazione che corrispondono a differenti livelli di astrazione; questi livelli formano una **gerarchia di concetti**.



Applicando il Deep Learning, avremo quindi una “macchina” che riesce autonomamente a classificare i dati ed a strutturarli gerarchicamente, trovando quelli più rilevanti e utili alla risoluzione di un problema (esattamente come fa la mente umana), migliorando le proprie prestazioni con l’apprendimento continuo.

LE RETI NEURALI ARTIFICIALI, LA BASE DEL DEEP LEARNING

Come accennato nel precedente paragrafo, l’apprendimento profondo (deep learning), basa il suo funzionamento sulla classificazione e “selezione” dei dati più rilevanti per giungere ad una conclusione, esattamente come fa il nostro cervello biologico che per formulare una risposta ad un quesito, dedurre un’ipotesi logica, arrivare alla risoluzione di un problema, mette in moto i propri neuroni biologici e le connessioni neurali (Neuroni biologici interconnessi formano le nostre **reti neurali cerebrali**, quelle che permettono a ciascun individuo di ragionare, fare calcoli in parallelo, riconoscere suoni, immagini, volti, imparare e agire).

Il Deep Learning si comporta allo stesso modo e sfrutta le reti neurali artificiali, **modelli di calcolo matematico-informatici basati sul funzionamento delle reti neurali biologiche, ossia modelli costituiti da interconnessioni di informazioni.**

Una rete neurale di fatto si presenta come un **sistema “adattivo” in grado di modificare la sua struttura (i nodi e le interconnessioni) basandosi sia su dati esterni sia su informazioni interne** che si connettono e passano attraverso la rete neurale durante la fase di apprendimento e ragionamento.

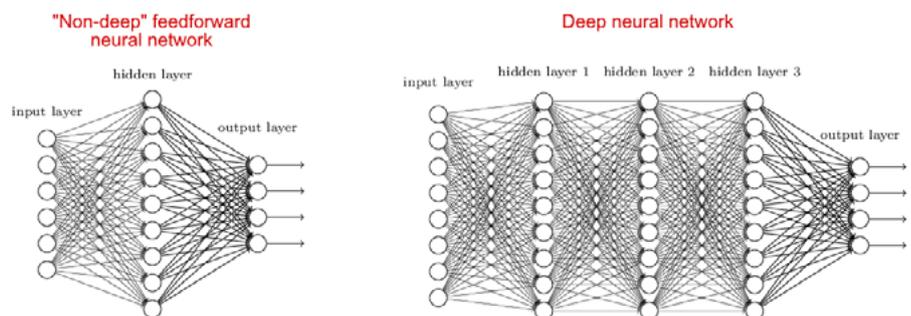
PER APPROFONDIMENTI SULLE RETI NEURALI RIMANDIAMO A: [“Reti neurali: cosa sono e a cosa servono”](#)

COME FUNZIONA IL DEEP LEARNING

Con il Deep Learning vengono simulati i processi di apprendimento del cervello biologico attraverso sistemi artificiali (le reti neurali artificiali, appunto) per insegnare alle macchine non solo ad apprendere autonomamente ma a farlo in modo più “profondo” come sa fare il cervello umano dove profondo significa “su più livelli” (vale a dire sul numero di layer nascosti nella rete neurale – chiamati **hydden layer**: quelle “tradizionali” contengono 2-3 layer, mentre le reti neurali profonde possono contenerne oltre 150).

L'immagine di seguito (tratta dall'[eBook disponibile online gratuitamente](#) “Neural Networks and Deep Learning”)

può aiutare a comprendere meglio la “struttura” delle reti neurali profonde.



Le reti neurali profonde sfruttano un numero maggiore di strati intermedi (hidden layer) per costruire più livelli di astrazione, proprio come si fa nei circuiti booleani (modello matematico di computazione usato nello studio della teoria della complessità computazionale che, in informatica, afferisce alla teoria della computabilità ossia studia le risorse minime necessarie – principalmente tempo di calcolo e memoria – per la risoluzione di un problema).

Proviamo a fare un esempio concreto di funzionamento di una rete neurale profonda con il riconoscimento visivo dei pattern: i neuroni del primo strato potrebbero imparare a riconoscere i bordi, i neuroni nel secondo strato potrebbero imparare a riconoscere forme più complesse, ad esempio triangoli o rettangoli, create dai bordi. Il terzo strato riconoscerebbe forme ancora più complesse, il quarto riconosce ulteriori dettagli e così via... **i molteplici livelli di astrazione possano dare alle reti neurali profonde un vantaggio enorme nell'imparare a risolvere complessi problemi di riconoscimento di schemi** proprio perché ad ogni livello intermedio aggiungono informazioni e analisi utili a fornire un output affidabile.

E' abbastanza facile intuire che quanti più livelli intermedi ci sono in una rete neurale profonda (e quindi quanto più è grande la rete neurale stessa) tanto più efficace è il risultato (il compito che è "chiamata" a svolgere) ma, di contro, la scalabilità della rete neurale è strettamente correlata ai data set, ai modelli matematici e alle risorse computazionali.

Seppur la richiesta di capacità computazionali enormi possa rappresentare un limite, la scalabilità del **Deep Learning grazie all'aumento dei dati disponibili e degli algoritmi è ciò che lo differenzia dal Machine Learning**: i sistemi di Deep Learning, infatti, migliorano le proprie prestazioni all'aumentare dei dati mentre le applicazioni di Machine Learning (o meglio, i cosiddetti sistemi di apprendimento superficiale) una volta raggiunto un certo livello di performance non sono più scalabili nemmeno aggiungendo esempi e dati di training alla rete neurale.

Questo perché nei sistemi di Machine Learning le caratteristiche di un determinato oggetto (nel caso di sistemi di riconoscimento visivo) vengono estratte e selezionate manualmente e servono per creare un modello in grado di categorizzare gli oggetti (in base alla classificazione e al riconoscimento di quelle caratteristiche); nei sistemi di Deep Learning, invece, l'estrazione delle caratteristiche avviene in modo automatico: **la rete neurale apprende in modo autonomo come analizzare dati grezzi e come svolgere un compito** (per esempio classificare un oggetto riconoscendone, autonomamente, le caratteristiche).

Se dal punto di vista delle potenzialità il Deep Learning può sembrare più "affascinante" e utile del Machine Learning, va precisato che **il calcolo computazionale richiesto per il loro funzionamento è davvero impattante, anche dal punto di vista economico**: le CPU più avanzate e le GPU top di gamma utili a "reggere" i workload di un sistema

di Deep Learning costano ancora migliaia di dollari; il ricorso a capacità computazionali via Cloud attenuano solo in parte il problema perché la formazione di una rete neurale profonda richiede spesso l'elaborazione di grandi quantità di dati utilizzando cluster di GPU di fascia alta per molte, molte ore (non è detto quindi che acquistare "as a service" la capacità di calcolo necessaria risulti economico).

4

COME SI ADDESTRA UN SISTEMA DI DEEP LEARNING

Un semplicissimo quanto efficace esempio per capire il reale funzionamento di un sistema di Machine Learning (e la differenza con un sistema di Deep Learning) ci viene fornito da [Tech Target](#):

«Mentre gli algoritmi di apprendimento automatico tradizionali sono lineari, gli algoritmi di apprendimento profondo sono impilati in una gerarchia di crescente complessità e astrazione. Per capire l'apprendimento profondo, immaginiamo un bambino la cui prima parola è "cane". Il bambino impara cos'è un cane (e cosa non lo è) indicando oggetti e dicendo la parola cane. Il genitore dice "Sì, quello è un cane" o "No, non è un cane". Mentre il bambino continua a puntare agli oggetti, diventa più consapevole delle caratteristiche che tutti i cani possiedono. Ciò che il bambino fa, senza saperlo, è chiarire un'astrazione complessa (il concetto di cane) costruendo una gerarchia in cui ogni livello di astrazione viene creato con la conoscenza che è stata acquisita dallo strato precedente della gerarchia».

A differenza del bambino, che impiegherà settimane o addirittura mesi per comprendere il concetto di “cane” e lo farà con l’aiuto del genitore (quello che viene definito apprendimento supervisionato), **una applicazione che utilizza algoritmi di Deep Learning può mostrare e ordinare milioni di immagini, identificando con precisione quali immagini contengono i set di dati, in pochi minuti pur non avendo avuto alcun tipo di indirizzamento sulla correttezza o meno dell’identificazione di determinate immagini nel corso del training.**

Solitamente, nei sistemi di Deep Learning, l’unica accortezza degli scienziati è “etichettare” i dati (con i **meta tag**), per esempio inserendo il meta tag “cane” all’interno delle immagini che contengono un cane ma senza spiegare al sistema come riconoscerlo: è il sistema stesso, attraverso livelli gerarchici multipli, che intuisce cosa caratterizza un cane (le zampe, la coda, il pelo, ecc.) e quindi come riconoscerlo.

Questi sistemi si basano, in sostanza, su un **processo di apprendimento “trial-and-error”** ma perché l’output finale sia affidabile sono necessarie enormi quantità di dati. Pensare subito ai Big Data e alla facilità con cui oggi si producono e distribuiscono dati di qualsiasi forma e da qualsiasi fonte come facile risoluzione sarebbe però un errore: l’accuratezza dell’output richiede, almeno nella prima fase di addestramento, l’utilizzo di dati “etichettati” (contenenti dei meta tag) che significa che l’utilizzo di dati non strutturati potrebbero rappresentare un problema. I dati non strutturati possono essere analizzati da un modello di apprendimento profondo una volta formato e raggiunto un livello accettabile di accuratezza, ma non per la fase di training del sistema.

Non solo, i sistemi basati su Deep Learning sono difficili da addestrare a causa del numero stesso di strati nella rete neurale. Il numero di strati e collegamenti

tra i neuroni nella rete è tale che può diventare difficile calcolare le “regolazioni” che devono essere apportate in ogni fase del processo di addestramento (un problema indicato come problema della [scomparsa del gradiente](#)); questo perché per il training comunemente si usano i cosiddetti **algoritmi di retropropagazione dell'errore (backpropagation)** attraverso il quale si rivedono i pesi della rete neurale (le connessioni tra i neuroni) in caso di errori (la rete propaga all'indietro l'errore in modo che i pesi delle connessioni vengano aggiornati in modo più appropriato). Un processo che continua in modo iterativo finché il gradiente (l'elemento che dà la direzione verso cui l'algoritmo deve muoversi) è nullo.

DEEP LEARNING FRAMEWORK: perché quello di Facebook sta attirando l'attenzione

Uno dei framework specifici per il Deep Learning più in uso tra ricercatori, sviluppatori e [data scientist](#) è **TensorFlow**, nota libreria software open source (progetto supportato da **Google**) per l'apprendimento automatico che fornisce moduli testati e ottimizzati per la realizzazione di algoritmi da impiegare in diversi tipi di software e con diversi tipi di linguaggi di programmazione, da Python, C/C++, Java, Go, RUST, R, ... (in particolare per “compiti percettivi” e per la comprensione del linguaggio naturale).

Nel 2019 però ha iniziato a prendere piede un altro framework che, a detta di alcuni analisti ([come riportato dall'analista Janakiram MSV in un suo articolo pubblicato su Forbes](#)), sta rapidamente diventando il preferito di sviluppatori e data scientist. Si tratta di

PyTorch, un progetto open source di **Facebook** già ampiamente utilizzato all'interno dell'azienda.

Inizialmente (e per diversi anni) gli sviluppatori di Facebook hanno utilizzato un framework noto come **Caffe2**, che è stato adottato anche da molte università e ricercatori. Già nel 2018 però Facebook annunciò che stava lavorando ad un altro tipo di framework capitalizzando gli sforzi dello sviluppo di Caffe2 con l'obiettivo di creazione un nuovo framework accessibile alla comunità open source.

In realtà, ciò che sta facendo Facebook è combinare il meglio di Caffe2 e ONNX in un nuovo framework (PyTorch); ONNX sta per Open Neural Network Exchange ed è un framework interoperabile cui anche Microsoft e AWS contribuiscono attivamente fornendo supporto per Microsoft CNTK e Apache MXNet.

PyTorch 1.0, di fatto, combina il meglio di Caffe2 e ONNX (è uno dei primi framework con supporto nativo per i modelli ONNX).

Ciò su cui si stanno concentrano gli sviluppatori di Facebook (ma non solo) è la creazione di un framework molto più semplice ed accessibile rispetto a TensorFlow. PyTorch, per esempio, utilizza una tecnica nota come calcolo dinamico che semplifica l'addestramento delle reti neurali. Non solo, "il modello di esecuzione di PyTorch imita il modello di programmazione convenzionale noto a uno sviluppatore medio di [Python](#). Offre inoltre formazione distribuita, profonda integrazione in Python e un vivace ecosistema di strumenti e librerie (come [Keras](#)), che lo rende popolare tra ricercatori e ingegneri", scrive nel suo articolo l'analista Janakiram MSV.

I CASI D'USO E I TIPI DI APPLICAZIONI DEL DEEP LEARNING

Nonostante le problematiche che abbiamo illustrato, i sistemi di Deep Learning hanno compiuto enormi passi evolutivi e sono migliorati moltissimo negli ultimi cinque anni, soprattutto per la grandissima quantità di dati a disposizione ma soprattutto per la disponibilità di infrastrutture ultra performanti (CPU e GPU in particolare).

Nell'ambito della ricerca sull'Artificial Intelligence, l'apprendimento automatico ha riscosso un notevole successo negli ultimi anni, consentendo ai computer di superare o avvicinarsi alle prestazioni umane corrispondenti in aree che vanno dal riconoscimento facciale al riconoscimento vocale e linguistico. L'apprendimento profondo invece consente ai computer di fare un passo in avanti, in particolare di risolvere una serie di problemi complessi.

Già oggi ci sono casi d'uso ed ambiti di applicazione che possiamo notare anche come "comuni cittadini" non esperti di tecnologia. Dalla computer vision per le [auto senza conducente](#), fino [robot droni](#) impiegati per la consegna di pacchi o anche per l'assistenza in casi di emergenza (per esempio per la consegna di cibo o sangue per trasfusioni in zone terremotate, alluvionate o in zone che devono affrontare crisi epidemiologiche, ecc.); riconoscimento e sintesi vocale e linguistica per [chatbot](#) e **robot di servizio**; riconoscimento facciale per **sorveglianza** in paesi come la Cina; riconoscimento immagini per aiutare i radiologi a **individuare i tumori nei raggi X**, oppure per aiutare i ricercatori a **individuare le sequenze genetiche correlate alle malattie** e identificare

le molecole che potrebbero portare a farmaci più efficaci o addirittura personalizzati; sistemi di analisi per la manutenzione predittiva su una infrastruttura o un impianto analizzando i dati dei sensori dell'IoT; e ancora, la visione del computer che rende possibile il supermercato Amazon Go senza cassa.

Guardando invece ai **tipi di applicazione** (intesi come compiti che una macchina può svolgere grazie al Deep Learning), quelli di seguito sono quelli ad oggi più maturi:

1) **colorazione automatica delle immagini in bianco e nero** (per la rete neurale significa riconoscere bordi, sfondi, dettagli e conoscere i colori tipici di una farfalla, per esempio, sapendo esattamente dove collocare il colore corretto);

2) **aggiunta automatica di suoni a filmati silenziosi** (per il sistema di Deep Learning significa sintetizzare suoni e collocarli correttamente all'interno di una situazione particolare riconoscendo immagini ed azioni, per esempio inserendo il suono del martello pneumatico, della rottura dell'asfalto e i sottofondi di una strada cittadina trafficata in un video in cui si vedono dei lavoratori che stanno rompendo l'asfalto con il martello pneumatico);

3) **traduzione simultanea** (per il sistema di Deep Learning significa ascoltare e riconoscere il linguaggio naturale, riconoscere la lingua parlata e tradurre il significato in un'altra lingua);

4) **classificazione degli oggetti all'interno di una fotografia** (il sistema in questo caso è in grado di riconoscere e classificare tutto ciò che vede in un'immagine, anche molto complessa dove per esempio c'è un paesaggio di sfondo, per esempio delle montagne, persone che camminano lungo un sentieri, degli animali al pascolo, ecc.);

5) **generazione automatica della grafia** (ci sono già oggi sistemi di Deep Learning capaci di utilizzare la grafia umana per scrivere addirittura apprendendo gli stimoli della scrittura a mano degli esseri umani ed imitandola);

6) **generazione automatica di testo** (sono sistemi che hanno imparato a scrivere correttamente in una determinata lingua rispettando ortografia, punteggiatura, grammatica e persino imparando ad usare stili di scrittura differenti a seconda dell'output da produrre, per esempio un articolo giornalistico o una novella);

7) **generazione automatica di didascalie** (in questo caso il riconoscimento delle immagini, l'analisi del contesto e la capacità di scrittura consentono ad un sistema di scrivere in automatico le didascalie di una immagine descrivendone perfettamente la scena);

8) **gioco automatico** (abbiamo imparato a capire le potenzialità di un sistema in grado di imparare autonomamente come giocare a un determinato gioco grazie a DeepMind – oggi parte di Google – che ha sviluppato un sistema di Deep Learning – AlphaGo – che non solo ha imparato a giocare al complessissimo gioco Go ma è riuscito anche a battere il campione mondiale, umano).

Driveless car

[Traffic Sign Detection](#) (TDS) è una funzione presente su molte auto di nuova fabbricazione che permette di riconoscere i segnali stradali. Si tratta di una applicazione di machine learning che utilizza le reti neurali convoluzionali e framework come Tensorflow.

Cinema

All'inizio del 2021 è stata presentata una nuova metodologia di AI applicabile nella produzione cinematografica. Il nuovo approccio si avvale di diversi strumenti di **deep learning** in parallelo e in serie (VGG16, MLP, transfer learning) e adopera differenti tipologie di dataset (immagini con feature diverse messe in risalto) al fine di poter classificare in modo accurato le inquadrature cinematografiche e permetterne così un utilizzo professionale e operativo nel processo di film making o nell'attività di indicizzazione dei contenuti streaming. Parliamo di [AI applicata](#) all'**image processing**.

Quantum Intelligence

Sul [quantum computing](#) si concentrano grandi aspettative. Un paradigma di elaborazione che richiede nuovo hardware, nuovi algoritmi e nuove soluzioni. La caratteristica principale è la capacità di semplificare molto la soluzione di alcuni problemi, riducendone la complessità esponenziale. Un esempio: la determinazione dei fattori di un numero, alla base di tanti problemi di crittografia e su cui si fondano moltissime applicazioni di sicurezza informatica.

Ologrammi

Uno studio condotto sugli [ologrammi](#) nel 2021 dai ricercatori del Massachusetts Institute of Technology (MIT) ha provato come, grazie a una nuova tecnica di deep learning denominata "olografia tensoriale", è possibile generare un **video olografico** in modo istantaneo, sfruttando la capacità computazionale presente su un semplice computer. La peculiarità è quella di essere composta da tensori addestrabili, che possono apprendere come elaborare le informazioni visive e di profondità in modo simile a quanto fa il cervello umano.

NETWORK **DIGITAL** 360

Network Digital360 è il più grande network in Italia di testate e portali B2b dedicati ai temi della Trasformazione Digitale e dell'Innovazione Imprenditoriale, con oltre 50 fra portali, canali e newsletter.

Ha la missione di diffondere la cultura digitale e imprenditoriale nelle imprese e pubbliche amministrazioni italiane e di fornire a tutti i decisori che devono valutare investimenti tecnologici informazioni aggiornate e approfondite.

Il Network è parte integrante di [Digital360HUB](#), il polo di Demand Generation di Digital360, che mette a disposizione delle tech company un'ampia gamma di servizi di comunicazione, storytelling, pr, content marketing, marketing automation, inbound marketing, lead generation, eventi e webinar.

VIA COPERNICO, 38

20125 - MILANO

TEL. 02 92852785

MAIL: MARKETING@DIGITAL4.BIZ

©ICT & Strategy

