

**UNIVERSITÀ POLITECNICA DELLE MARCHE**  
**FACOLTÀ DI INGEGNERIA**  
Dipartimento di Ingegneria dell'Informazione  
Corso di Laurea in Ingegneria Informatica e dell'Automazione

---



**TESI DI LAUREA**

**Esperienze nel mondo dell'Intelligenza Artificiale: Sentiment  
Analysis, Computer Vision e Reinforcement Learning**

**Experiences in the Artificial Intelligence world: Sentiment Analysis,  
Computer Vision and Reinforcement Learning**

Relatore

Prof. Domenico Ursino

Candidato

Alex Pugnaroni

---

**ANNO ACCADEMICO 2023-2024**

*Un giorno le macchine riusciranno a risolvere tutti i problemi,  
ma mai nessuna di esse potrà porne uno.*

Albert Einstein

## Sommario

Negli ultimi anni, l'Intelligenza Artificiale ha assunto un ruolo sempre più prominente nella nostra quotidianità, rivoluzionando numerosi aspetti della nostra vita. Grazie alla sua capacità di analizzare enormi quantità di dati e apprendere dai risultati ottenuti, l'IA sta contribuendo a ottimizzare processi e servizi in svariati settori, tra cui la sanità, l'industria, l'educazione e il trasporto. Questa tecnologia ci offre opportunità senza precedenti, consentendo di sviluppare soluzioni innovative per sfide complesse e di migliorare la qualità della vita umana. Questa tesi è finalizzata ad esplorare alcuni degli ambiti di utilizzo di sistemi IA testando con mano, dal punto di vista di un utente, le potenzialità di questa tecnologia rivoluzionaria.

**Keyword:** Intelligenza Artificiale, Machine Learning, Deep Learning, Cloud Computing, Computer Vision, Reti Neurali, Tecnologia, Innovazione

<b>Introduzione</b>	<b>1</b>
<b>1 Introduzione all'Intelligenza Artificiale</b>	<b>3</b>
1.1 Evoluzione e Concetti Chiave dell'Intelligenza Artificiale . . . . .	3
1.1.1 Definizione di Intelligenza Artificiale . . . . .	3
1.1.2 Origini e Fondamenti . . . . .	3
1.1.3 IA simbolica . . . . .	4
1.1.4 IA subsimbolica . . . . .	4
1.1.5 Periodo di difficoltà e stagnazione . . . . .	5
1.1.6 Rinascita e sistemi basati sulla conoscenza . . . . .	6
1.1.7 Dall'ambiente accademico all'industria . . . . .	6
1.1.8 Il ritorno delle reti neurali . . . . .	6
1.1.9 L'intelligenza artificiale moderna . . . . .	7
1.1.10 IA debole, IA forte . . . . .	7
1.2 Test classici e dilemmi filosofici nell'Intelligenza Artificiale . . . . .	8
1.2.1 Test di Turing e le sue varianti . . . . .	8
1.2.2 Test della stanza cinese . . . . .	9
1.2.3 Test di Winograd . . . . .	10
1.2.4 Dibattito filosofico . . . . .	11
1.3 Applicazioni dell'Intelligenza Artificiale . . . . .	11
1.3.1 Intelligent Data Processing . . . . .	11
1.3.2 Natural Language Processing (NLP) . . . . .	12
1.3.3 Virtual Assistant/Chatbot . . . . .	12
1.3.4 Computer Vision . . . . .	12
1.3.5 Recommendation System . . . . .	12
1.3.6 Soluzioni Fisiche di Intelligenza Artificiale . . . . .	13
<b>2 Amazon Web Services</b>	<b>14</b>
2.1 Introduzione al cloud computing e Amazon Web Services . . . . .	14
2.1.1 Che cos'è Amazon Web Services . . . . .	14
2.1.2 Servizi principali di AWS . . . . .	15
2.1.3 Architettura ed infrastruttura AWS . . . . .	17
2.1.4 Costi di utilizzo . . . . .	18

<b>3</b>	<b>Esperienze di Sentiment Analysis</b>	<b>19</b>
3.1	Introduzione alla Sentiment Analysis . . . . .	19
3.1.1	Che cos'è la Sentiment Analysis . . . . .	19
3.1.2	Processo di funzionamento della Sentiment Analysis . . . . .	20
3.1.3	Approcci di Sentiment Analysis . . . . .	20
3.2	Amazon Comprehend . . . . .	21
3.2.1	Che cos'è Amazon Comprehend . . . . .	21
3.2.2	Funzionalità di Amazon Comprehend . . . . .	21
3.3	TripAdvisor ed Amazon Comprehend per test di Sentiment Analysis . . . . .	23
3.3.1	Recensione positiva . . . . .	24
3.3.2	Recensione negativa . . . . .	25
3.3.3	Recensione mista . . . . .	26
3.3.4	Recensione sarcastica . . . . .	26
3.3.5	Limiti della Sentiment Analysis . . . . .	27
<b>4</b>	<b>Esperienze di Computer Vision</b>	<b>28</b>
4.1	Introduzione alla Computer Vision . . . . .	28
4.1.1	Che cos'è la Computer Vision . . . . .	28
4.1.2	Come funziona la Computer Vision . . . . .	29
4.2	Amazon Rekognition . . . . .	30
4.2.1	Che cos'è Amazon Rekognition . . . . .	30
4.2.2	Funzionalità di Amazon Rekognition . . . . .	30
4.3	Test di riconoscimento di immagini tramite demo . . . . .	31
4.3.1	Label detection . . . . .	31
4.3.2	Image properties . . . . .	31
4.3.3	Facial analysis . . . . .	32
4.3.4	Face comparison . . . . .	32
4.3.5	Celebrity recognition . . . . .	33
4.3.6	Text in image . . . . .	33
4.4	Test di riconoscimento di immagini tramite Custom Labels . . . . .	34
4.4.1	Creazione del dataset . . . . .	34
4.4.2	Fase di etichettatura . . . . .	35
4.4.3	Training del modello . . . . .	35
4.4.4	Prestazioni del modello . . . . .	36
4.4.5	Variante del test di riconoscimento . . . . .	37
<b>5</b>	<b>Introduzione al Reinforcement Learning</b>	<b>39</b>
5.1	Che cos'è il Reinforcement Learning . . . . .	39
5.2	Come funziona il Reinforcement Learning . . . . .	40
5.3	Fasi dell'addestramento del Reinforcement Learning . . . . .	40
5.4	Applicazioni del Reinforcement Learning . . . . .	42
5.5	Reinforcement Learning nei giochi . . . . .	42
<b>6</b>	<b>Esperienze di Reinforcement Learning</b>	<b>45</b>
6.1	Snake Game ed il Reinforcement Learning . . . . .	45
6.1.1	Librerie per lo sviluppo . . . . .	46
6.1.2	Definizione del sistema . . . . .	46
6.1.3	Definizione della Q-Table ed equazione di Bellman . . . . .	47
6.1.4	Sintesi dell'algoritmo . . . . .	48
6.1.5	Prestazioni dell'agente IA . . . . .	50

<b>7</b>	<b>Discussione</b>	<b>51</b>
7.1	Discussione in merito alla Sentiment Analysis . . . . .	51
7.2	Discussione in merito alla Computer Vision . . . . .	51
7.3	Discussione in merito al Reinforcement Learning . . . . .	52
	<b>Conclusioni</b>	<b>53</b>
	<b>Bibliografia</b>	<b>55</b>
	<b>Ringraziamenti</b>	<b>57</b>

---

## Elenco delle figure

---

1.1	Esempio di funzionamento del perceptrone . . . . .	5
1.2	Garry Kasparov nella famosa partita contro Deep Blue . . . . .	7
1.3	Esempio illustrativo Test di Turing . . . . .	8
1.4	Esempio illustrativo Test della stanza cinese . . . . .	10
1.5	Abitacolo di un veicolo a guida autonoma . . . . .	13
2.1	Esempio illustrativo di cloud computing . . . . .	15
2.2	Servizi offerti da Amazon Web Services . . . . .	17
3.1	Cattura del testo analizzato da parte di Amazon Comprehend . . . . .	22
3.2	Esempio di Sentiment Analysis . . . . .	23
3.3	Logo di TripAdvisor con range associati ad ogni pallino . . . . .	24
3.4	Screenshot di una recensione positiva su TripAdvisor . . . . .	24
3.5	Screenshot di una recensione negativa su TripAdvisor . . . . .	25
3.6	Screenshot di una recensione mista su TripAdvisor . . . . .	26
4.1	Screenshot del risultato di una Label detection su un'immagine di uno skater in città . . . . .	31
4.2	Screenshot del risultato dell'applicazione di Image properties sulla foto di una metropoli in notturna . . . . .	32
4.3	Screenshot del risultato dell'applicazione di una Facial analysis di una ragazza in auto . . . . .	32
4.4	Screenshot del risultato di Face comparison . . . . .	33
4.5	Screenshot del risultato di Celebrity recognition quando viene applicato su Jeff Bezos . . . . .	33
4.6	Screenshot del risultato di Text in image . . . . .	34
4.7	Screenshot di un dataset di training . . . . .	34
4.8	Screenshot di una etichettatura su logo Nike . . . . .	35
4.9	Screenshot di training in corso del modello . . . . .	35
4.10	Screenshot delle prestazioni del modello . . . . .	36
4.11	Screenshot del risultato dell'analisi di due immagini . . . . .	37
4.12	Screenshot di training in corso del modello variante . . . . .	37
4.13	Screenshot delle prestazioni del modello variante . . . . .	38
4.14	Screenshot del risultato dell'analisi di due immagini . . . . .	38
5.1	Esempio di modello per il Reinforcement Learning . . . . .	39

---

5.2	Il Reinforcement Learning nell'addestramento dei cani . . . . .	40
5.3	Workflow del Reinforcement Learning . . . . .	40
5.4	Addestramento di un problema di apprendimento che non risponde al principio di sample efficiency con il calcolo parallelo . . . . .	41
5.5	Screenshot di una partita a Call of Duty: Black Ops 3 . . . . .	43
5.6	Screenshot di una partita a Dota 2 . . . . .	44
6.1	Screenshot di una partita a Snake . . . . .	45
6.2	Esempio di un set di stati che indicano la posizione del serpente . . . . .	47
6.3	Esempio di un set di azioni che il serpente può compiere . . . . .	47
6.4	Esempio di Q-Table rappresentante le possibili azioni e le relative probabilità di successo . . . . .	47
6.5	Equazione di Bellman . . . . .	48
6.6	Equazioni di Bellman semplificate . . . . .	48
6.7	Funzione di perdita . . . . .	49
6.8	Schema di elaborazione delle azioni intraprese dal modello . . . . .	49
6.9	Grafico punteggio - partite giocate con screenshot di una partita a Snake . . . . .	50

L'evoluzione tecnologica ha caratterizzato gli ultimi decenni con un progresso senza precedenti, trasformando radicalmente la nostra società e il nostro modo di vivere. Dall'avvento dei computer e di Internet fino alle più recenti innovazioni in ambito digitale, lo sviluppo tecnologico ha sempre attirato l'interesse e l'ingegno dell'uomo. Tra tutte le innovazioni, l'Intelligenza Artificiale (IA) si distingue per il suo potenziale rivoluzionario e per le vaste applicazioni in una molteplicità di settori.

L'IA rappresenta il culmine di decenni di ricerca e sviluppo nel campo della tecnologia e dell'informatica. Sin dai primi calcolatori, l'idea di creare macchine capaci di emulare l'intelligenza umana ha affascinato scienziati e ingegneri. I progressi nella potenza di calcolo, insieme alla disponibilità di enormi quantità di dati e all'avanzamento degli algoritmi di apprendimento automatico, hanno permesso all'IA di diventare una realtà tangibile e sempre più pervasiva.

In particolare, sono i campi del Machine Learning e del Deep Learning a dimostrare l'efficacia e l'importanza di queste tecnologie. Le continue scoperte e innovazioni in questi ambiti confermano il loro ruolo centrale nel progresso tecnologico. Il nostro mondo è in costante cambiamento, e attualmente uno dei cambiamenti più significativi è determinato proprio dallo sviluppo di queste tecnologie. Le aziende di tutte le dimensioni, dalle piccole imprese alle grandi multinazionali, riconoscono i benefici che i sistemi di Intelligenza Artificiale possono apportare, sia a livello economico che sociale e mediatico. Tecnologie come la Computer Vision, che permettono ai computer di interpretare e comprendere il mondo visivo, sono ormai utilizzate in numerosi settori, dal riconoscimento facciale alla diagnostica medica, evidenziando ulteriormente l'impatto trasformativo dell'IA.

Oltre a questi, il campo del Reinforcement Learning, o apprendimento per rinforzo, rappresenta un'area di grande innovazione. Questo approccio permette agli algoritmi di apprendere comportamenti ottimali attraverso interazioni con l'ambiente e feedback sotto forma di ricompense o penalità. Il Reinforcement Learning ha mostrato risultati straordinari in applicazioni come i giochi, la robotica e la gestione delle risorse, dove decisioni sequenziali e strategie complesse sono essenziali.

Le applicazioni dell'IA sono ormai innumerevoli: dai sistemi di riconoscimento vocale e facciale, alle auto a guida autonoma, dai consigli personalizzati dei sistemi di e-commerce, fino agli assistenti virtuali. Questa diffusione capillare della tecnologia dell'IA testimonia non solo il suo valore pratico, ma anche il crescente interesse che suscita nell'uomo. L'IA non solo risolve problemi complessi in maniera efficiente, ma apre anche nuove possibilità di innovazione in ambiti prima impensabili.

Questa tesi tratterà tutti questi argomenti in modo specifico e dettagliato, con l'obiettivo

di dimostrare quanto l'Intelligenza Artificiale sia pervasiva nella realtà odierna e quanto questi sistemi possano risultare estremamente capaci, se adoperati efficacemente.

Si inizierà con un'introduzione all'IA, spiegandone le origini, i fondamenti, l'evoluzione e i test condotti nel corso della storia fino ad arrivare ai giorni nostri. Successivamente, si farà un accenno al cloud computing, introducendo la piattaforma cloud di Amazon, ovvero Amazon Web Services (AWS). È importante sottolineare che tutte le analisi e gli esperimenti descritti nelle sezioni successive sono stati svolti durante la fase di tirocinio.

In questo contesto, verrà presentata la Sentiment Analysis, legata all'elaborazione del linguaggio naturale, e il riconoscimento delle immagini, collegato alla Computer Vision. Per quanto riguarda la Sentiment Analysis, verrà utilizzato lo strumento di AWS denominato Amazon Comprehend, che verrà testato per valutarne la capacità di rilevare l'umore di una persona in base a un testo o messaggio scritto. Per il riconoscimento delle immagini, invece, si farà uso di un altro strumento di AWS, chiamato Amazon Rekognition. Questo strumento verrà utilizzato per visionare alcune demo pre-addestrate offerte dal servizio e per effettuare dei test di riconoscimento su etichette personalizzate.

Successivamente, verrà introdotto il Reinforcement Learning, noto anche come apprendimento per rinforzo. Saranno analizzati il funzionamento, le fasi di addestramento e le applicazioni del Reinforcement Learning, con particolare attenzione all'utilizzo di questa tecnologia nel mondo dei giochi. Infine, verrà sviluppato un agente IA in grado di imparare a giocare autonomamente al celebre gioco Snake.

La presente tesi è composta da sette capitoli strutturati come di seguito specificato:

- Nel Capitolo 1 verrà introdotto il mondo dell'Intelligenza Artificiale a partire dalle sue origini, evidenziandone l'evoluzione, alcuni celebri test e diverse applicazioni.
- Nel Capitolo 2 verrà introdotto il concetto di cloud computing e verrà presentata la piattaforma di punta per l'IA di Amazon, ovvero Amazon Web Services.
- Nel Capitolo 3 si approfondirà l'interpretazione del linguaggio naturale, in particolare la Sentiment Analysis, utilizzando Amazon Comprehend, un servizio fornito da AWS.
- Nel Capitolo 4 si approfondirà il concetto di Computer Vision legato al riconoscimento delle immagini, effettuando test con Amazon Rekognition, un servizio anch'esso fornito da AWS.
- Nel Capitolo 5 verrà introdotto il Reinforcement Learning, spiegandone il funzionamento, le fasi di addestramento e le applicazioni.
- Nel Capitolo 6 verrà spiegato come creare un agente IA in grado di imparare, tramite Reinforcement Learning, a giocare autonomamente al celebre gioco Snake.
- Nel Capitolo 7 verranno tratte le conclusioni del lavoro svolto e verranno delineati alcuni possibili sviluppi futuri.

---

## Introduzione all'Intelligenza Artificiale

---

*In questo capitolo si introdurrà il concetto di "Intelligenza Artificiale" (IA), dalla sua origine fino ai giorni nostri. Si analizzeranno i principali test sull'IA e rifletteremo sulla possibilità che le macchine possano pensare. Infine si approfondiranno alcune delle applicazioni dell'Intelligenza Artificiale.*

### 1.1 Evoluzione e Concetti Chiave dell'Intelligenza Artificiale

Il sogno di creare una macchina intelligente come l'uomo, se non di più, è vecchio di secoli, ma è entrato a far parte della scienza contemporanea con l'avvento dei computer digitali. Di fatto, l'idea sarebbe nata con le prime macchine digitali programmabili e derivava dagli sforzi dei matematici di comprendere il pensiero umano, in particolare la logica intesa come processo meccanico di "manipolazione di simboli".

Secondo i pionieri della computazione, come Alan Turing e John von Neumann, c'erano profonde analogie tra i computer e il cervello umano, e pareva, quindi, chiaro che l'intelligenza umana si potesse replicare in programmi informatici.

#### 1.1.1 Definizione di Intelligenza Artificiale

L'intelligenza Artificiale (IA) è l'abilità di una macchina di mostrare capacità umane, ovvero il ragionamento, l'apprendimento, la pianificazione e la creatività, e permette ai sistemi di capire il proprio ambiente, di mettersi in relazione con quello che percepisce, di risolvere problemi e agire verso un obiettivo specifico. Sostanzialmente il computer riceve dati già preparati o raccolti tramite sensori, analizza gli effetti delle azioni precedenti e, lavorando in autonomia, li processa e risponde.

#### 1.1.2 Origini e Fondamenti

L'Intelligenza Artificiale ha una data di nascita, ovvero l'estate 1956. Per lo più, chi lavora nel mondo dell'Intelligenza Artificiale fa risalire la fondazione ufficiale del proprio campo a un workshop ristretto svoltosi nel 1956 al Dartmouth College e organizzato dal giovane matematico John McCarthy che, entrato nel 1955 a far parte del corpo docente del Dartmouth College ed avendo acquisito alcuni rudimenti della psicologia nascente, si era appassionato all'idea di creare una macchina pensante. Così, nell'estate del 1956, dopo aver conosciuto alcune personalità come Minsky, Shannon e Rochester, chiese loro di aiutarlo ad organizzare uno studio di due mesi con dieci uomini per creare una macchina in grado di simulare ogni

aspetto dell'apprendimento e dell'intelligenza umana. L'espressione "Intelligenza Artificiale" fu una invenzione di McCarthy, che voleva distinguere questo campo da un'impresa affine, chiamata cibernetica.

Al workshop del 1956 al Dartmouth College i partecipanti espressero opinioni divergenti sulla corretta impostazione da seguire per sviluppare una macchina pensante. Alcuni di loro, in genere matematici, promuovevano la logica matematica e il ragionamento deduttivo come linguaggio del pensiero razionale; altri sostenevano i metodi induttivi, per cui i programmi ricavano una statistica dai dati e usavano le probabilità per trattare l'incertezza; altri ancora erano convinti che ci si dovesse ispirare alla biologia e alla psicologia per creare programmi simili al cervello. Tutto ciò ha scatenato la creazione di diversi approcci con corredi di principi e tecniche propri al punto da definire molteplici paradigmi per l'IA. È il caso, ad esempio, del Deep Learning, una famiglia di metodi dell'IA che si è elevata al di sopra di questa anarchia diventando il paradigma dominante, a tal punto che molti media identificano il termine intelligenza artificiale con Deep Learning. In realtà, l'IA è un campo che include un ampio insieme di approcci, il cui obiettivo è creare macchine intelligenti. Il Deep Learning è soltanto uno degli approcci.

Da queste distinzioni nascono due scissioni pratiche e filosofiche: IA simbolica e IA subsimbolica.

### 1.1.3 IA simbolica

In un programma di IA simbolica la conoscenza consiste in parole o frasi (simboli), tipicamente comprensibili da un essere umano, insieme a regole con le quali il programma può combinarle ed elaborarle per eseguire un compito assegnato.

Uno dei primi programmi di IA simbolica fu GPS, acronimo di General Problem Solver, ideato da Newell e Simon per risolvere problemi come l'enigma dei "missionari e dei cannibali". In questo celebre rompicapo tre missionari e tre cannibali devono attraversare un fiume, sennonché dispongono di una barca a due posti. Se gli affamati cannibali superano di numero gli appetitosi missionari su una riva del fiume, quest'ultimi faranno una brutta fine. I creatori del General Problem Solver avevano registrato diversi studenti che pensavano a voce alta mentre risolvevano questo ed altri enigmi logici; perciò decisero di progettare un programma che imitasse, a loro avviso, quelli che erano i processi mentali degli studenti.

Sostanzialmente la natura simbolica si può cogliere dal modo in cui erano codificate le istruzioni del programma, definendo uno stato attuale che di default è lo stato di partenza, e lo stato desiderato ossia l'obiettivo finale. Nel codice, il programma ha degli "operatori" (sottoprogrammi), che possono trasformare lo stato attuale in un nuovo stato, e delle "regole", che codificano i vincoli del compito. I difensori dell'impostazione simbolica nell'IA sostenevano che, per realizzare l'intelligenza nei computer, non è necessario costruire programmi che imitano il cervello, ma piuttosto definire dei programmi adeguati di elaborazioni di simboli.

### 1.1.4 IA subsimbolica

A differenza dell'IA simbolica, l'IA subsimbolica trae ispirazione dalla neuroscienza e cercava di cogliere i processi mentali su cui si fonda quella che alcuni chiamano percezione rapida, come riconoscere volti o identificare parole pronunciate. Questi programmi non contengono il tipo di linguaggio comprensibile da esseri umani come nell'IA simbolica, ma una sequenza di equazioni spesso molto difficili da comprendere. Un primo esempio di programma subsimbolico ispirato al cervello è stato il perceptrone, concepito sul finire degli anni Cinquanta dallo psicologo Frank Rosenblatt. L'invenzione di Rosenblatt dei perceptron si ispirava al modo in cui i neuroni elaborano l'informazione.

Un neurone è una cellula del cervello che riceve un input elettrico o chimico da altri neuroni a esso collegati.

In parole povere, un neurone somma tutti gli input ricevuti da altri neuroni e se la somma totale raggiunge una certa soglia, il neurone scarica, ossia trasmette l'impulso (Figura 1.1).

Il perceptrone, che ha molteplici input numerici ed un solo output, lavora in maniera analoga: somma gli input che lo raggiungono e, se la somma risultante è uguale o maggiore alla soglia del perceptrone, quest'ultimo restituisce come output il valore 1 (scarica), altrimenti 0 (non scarica). Questi approcci però si dimostrarono, almeno inizialmente, limitati rispetto all'IA simbolica, e tali da sottrarre solo dollari destinati a tentativi più proficui dell'IA simbolica.

Nel 1969 Minsky e il suo collega Papert pubblicarono il libro *Perceptrons*, in cui dimostravano matematicamente che i problemi risolvibili alla perfezione da un perceptrone erano limitati e che l'algoritmo di apprendimento del perceptrone non se la cavava affatto bene quando si applicava a compiti che richiedevano numerosi pesi e soglie.

Minsky e Papert misero in luce il fatto che, se un perceptrone è potenziato aggiungendo uno "strato" di neuroni simulati, i tipi di problema che il dispositivo può risolvere sono, in linea di principio, molto più ampi, dando origine alla rete neurale multistrato.

Essi però riconobbero gli ostacoli di allenare perceptron multistrato; lo stesso Rosenblatt riconobbe la difficoltà nell'addestramento, dando un momentaneo colpo mortale al perceptrone.

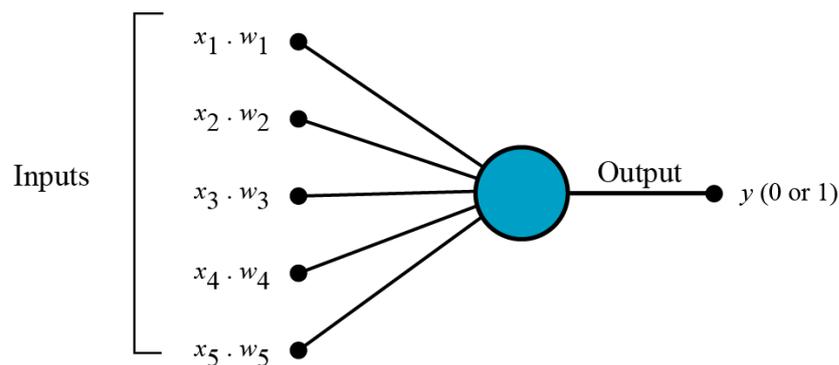


Figura 1.1: Esempio di funzionamento del perceptrone

### 1.1.5 Periodo di difficoltà e stagnazione

Intanto i fautori dell'IA simbolica stavano inviando richieste di finanziamenti, con la promessa di imminenti rivoluzioni in aree quali la comprensione della parola e del linguaggio, il ragionamento di senso comune, l'orientamento nei robot e nei veicoli a guida autonoma. A metà degli anni Settanta, mentre venivano impiegati con successo alcuni sistemi esperti dall'obiettivo molto ristretto, le più generali rivoluzioni dell'IA promesse non si erano ancora realizzate e, di conseguenza, quell'ottimismo che si era diffuso in quegli anni venne spento quando i ricercatori fallirono nel dimostrare teoremi a partire da più di una dozzina di assiomi. Gli enti finanziatori se ne accorsero e, tramite due report, sollecitati, rispettivamente, dal Science Research Council nel Regno Unito e dal Department of Defense degli Stati Uniti, riferirono parole poco lusinghiere sul progresso e le ricerche sull'IA.

Queste ultime influirono moltissimo sulla riduzione dei finanziamenti statali dando luogo ad un'importante chiusura di diverse start-up e ad un netto calo delle ricerche.

### 1.1.6 Rinascita e sistemi basati sulla conoscenza

Queste difficoltà portarono a definire gli approcci adottati dalle macchine come deboli, che necessitavano, quindi, di una conoscenza maggiore inerente al campo di applicazione. Nel 1969, grazie a Ed Feigenbaum, studente di Simon, Bruce Buchanan e Joshua Lederberg, venne creato il programma DENDRAL. Quest'ultimo era in grado, a partire dalle informazioni sulla massa molecolare ricavata da uno spettrometro, di ricostruire la struttura di una molecola. Tale programma fu quindi il primo dei sistemi basati su un uso intensivo della conoscenza che arrivò più tardi a inglobare i concetti teorizzati da McCarthy per l'Advice Taker. Successivamente, Feigenbaum cominciò, insieme ad altri ricercatori di Stanford, l'Heuristic Program Project (HPP), con l'obiettivo di estendere gli scenari applicativi di questi sistemi.

Intorno agli anni Settanta venne sviluppato Mycin, un sistema esperto, ovvero in grado di possedere una conoscenza in un determinato scenario di applicazione. Il programma riceveva circa seicento regole che usava per aiutare i medici a fare le diagnosi e a trattare malattie del sangue. I programmatori di Mycin svilupparono queste regole dopo meticolose interviste a medici esperti. Dati i sintomi di un paziente e i risultati dei test clinici, Mycin usava la logica e il ragionamento probabilistico, insieme con le proprie regole, per stilare una diagnosi, e sapeva spiegare i meccanismi di ragionamento che esso stesso aveva impiegato. In parole povere, Mycin era un esempio paradigmatico di IA simbolica, surclassando i sistemi subsimbolici che sembravano difficili da interpretare e in cui l'uomo riesce difficilmente a definire delle regole.

### 1.1.7 Dall'ambiente accademico all'industria

L'inizio degli anni Ottanta rappresentò l'avvio dell'impiego di tecnologie intelligenti all'interno delle industrie di tutto il mondo. Fu il caso di R1, il primo sistema utilizzato da Digital Equipment nel 1982 per aiutare l'azienda a configurare gli ordini per nuovi computer. In soli 4 anni riuscì a far risparmiare alla compagnia circa 40 milioni di dollari l'anno, facendo crescere ottimismo per l'utilizzo di questi programmi che, da allora, cominciarono ad essere utilizzati in quasi ogni grande azienda americana e giapponese.

L'industria dell'Intelligenza Artificiale raggiunse nel 1988 una cifra dell'ordine di miliardi di dollari, includendo centinaia di aziende che stavano creando sistemi esperti, robot, software e hardware specializzati in questi settori.

### 1.1.8 Il ritorno delle reti neurali

Intorno alla metà degli anni Ottanta fu reinventato l'algoritmo di apprendimento per le reti neurali che allora era stato, come già citato in precedenza, confutato da Minsky e Papert, basandosi sul concetto di retropropagazione per l'addestramento delle reti neurali. Come il termine lascia intendere, la retropropagazione è un modo per considerare un errore osservato nelle unità di output e propagare a ritroso la responsabilità di quell'errore, in modo da ripartirla a ciascun peso della rete. Ciò permette alla retropropagazione di determinare quanto modificare ciascun peso per ridurre l'errore. L'algoritmo fu applicato a molti modelli relativi all'apprendimento, inerenti sia al lato dell'informatica sia a quello della psicologia.

Da ciò presero spunto David Rumelhart e James McClelland, conosciuti come "connessionisti", chiamati, appunto, così perché sostenevano che, in queste reti, la conoscenza risiede nelle connessioni ponderate tra le unità. Il team di Rumelhart e McClelland è noto per aver scritto la "bibbia" del connessionismo, un trattato di due volumi pubblicato nel 1986 intitolato *Pdp: microstruttura dei processi cognitivi*. In un panorama in cui l'IA simbolica dominava l'intero mondo, il libro rappresentò una ventata d'ottimismo per l'approccio subsimbolico. In effetti, a metà degli anni Ottanta i sistemi esperti si stavano rivelando sempre più fragili: erano

inclinati a errori e spesso incapaci di generalizzare o di adattarsi al cospetto di nuove situazioni, evidenziando tutti i limiti dell'IA simbolica.

### 1.1.9 L'intelligenza artificiale moderna

Al giorno d'oggi i sistemi intelligenti sono presenti in ogni campo, anche nelle attività quotidiane, e primeggiano nei giochi, come teorizzato anni prima dai sostenitori di questa tecnologia. Nel 1997, infatti, il programma per giocare a scacchi Deep Blue della IBM sconfisse il campione mondiale Garry Kasparov (Figura 1.2).

Quest'evento frastornò lo stesso Kasparov al punto che accusò il team dell'IBM di avere barato: secondo lui una macchina non avrebbe potuto giocare così bene senza ricevere aiuti da esperti umani. A partire dalla metà degli anni Duemila una serie pervasiva di successi dell'IA ha colto di sorpresa e ha proliferato a un ritmo ubriacante: Google ha inaugurato Google Translate, il servizio di traduzione automatica, sono stati creati servizi di Assistenti Virtuali come Siri di Apple e Alexa di Amazon; inoltre, YouTube ha cominciato a fornire sottotitoli incredibilmente accurati nei video. Skype ha offerto traduzioni simultanee nelle videochiamate, Facebook ha introdotto riconoscimenti facciali nelle foto caricate sulla piattaforma ed Apple ha utilizzato questo riconoscimento facciale come chiave di accesso nei suoi dispositivi.

Successivamente è stato sviluppato AlphaGo, un potente programma di IA che ha sconfitto i migliori giocatori del complesso gioco Go in quattro partite su cinque. Insomma, l'Intelligenza Artificiale è entrata a far parte della quotidianità.



Figura 1.2: Garry Kasparov nella famosa partita contro Deep Blue

### 1.1.10 IA debole, IA forte

Tutti questi sviluppi che sono stati citati precedentemente sono esempi della cosiddetta IA "ristretta o debole". Questi termini non vanno intesi in senso dispregiativo ma si riferiscono semplicemente a un sistema che sa eseguire soltanto un compito definito in modo ristretto. Ad esempio, AlphaGo è forse il miglior giocatore al mondo di Go, ma sa fare solo questo. Google Translate, può tradurre in cinese una recensione di un film in inglese, ma non sa dirci se al recensore del film è piaciuto o meno. I termini *ristretto* e *debole* sono usati in contrapposizione all'IA *forte, di livello umano, generale o vera e propria*, ossia capace di replicare quasi ogni cosa che sa fare l'uomo, e magari molto di più.

L'IA generale mira a creare macchine intelligenti che siano indistinguibili dalla mente umana. Ma proprio come un bambino, la macchina avrà bisogno di imparare attraverso

stimoli ed esperienze, progredendo costantemente e sviluppando le sue capacità nel tempo. Così, gli sforzi dell'IA si sono poco alla volta concentrati su compiti particolari ben definiti, come il riconoscimento vocale, gli scacchi, la guida autonoma, e così via.

Mentre i ricercatori, sia nel mondo accademico che nei settori privati, sono impegnati nella creazione dell'AGI (Artificial General Intelligence), ad oggi, nella realtà effettiva quest'ultima, esiste solo come concetto teorico. Mentre alcuni individui come Marvin Minsky sono considerati eccessivamente ottimisti su ciò che potremmo realizzare in pochi decenni in questo campo, altri sostengono che non è neanche possibile sviluppare sistemi di IA forte. Finché le misure del successo, come l'intelligenza e la comprensione, non saranno esplicitamente definite, la loro convinzione è corretta. Ad oggi molti usano il test di Turing per valutare l'intelligenza di un sistema IA.

## 1.2 Test classici e dilemmi filosofici nell'Intelligenza Artificiale

Dalla nascita dell'Intelligenza Artificiale ai giorni nostri, molti sono stati gli esperimenti ed i test sviluppati per verificare l'efficienza e la capacità di comprensione delle macchine intelligenti. Tra questi spiccano il celebre Test di Turing e le sue varianti, il test della stanza cinese e il test di Winograd. Tuttavia, lungo il percorso, questi esperimenti hanno messo in discussione l'efficacia e la genuinità dell'Intelligenza Artificiale, portando ad un dibattito filosofico sulla vera natura della mente e della coscienza, generando perplessità filosofiche riguardo alla possibilità che una macchina possa pensare, provare emozioni e possedere una vera coscienza.

### 1.2.1 Test di Turing e le sue varianti

Il test di Turing è un criterio per determinare se una macchina sia in grado di esibire un comportamento intelligente. Tale criterio è stato suggerito da Alan Turing nell'articolo *Computing machinery and intelligence*, apparso nel 1950 sulla rivista *Mind*.

Il test si ispira ad un gioco chiamato "gioco dell'imitazione", dove ci sono due sfidanti: un computer e un essere umano. Ciascuno dei due riceve separatamente delle domande da un giudice (umano), il quale deve determinare l'identità dei due. Il giudice è fisicamente separato dai due sfidanti, e quindi non può basarsi su segnali visivi né uditivi, ma deve affidarsi a un semplice testo stampato che riceve.

Una macchina che supera il test di Turing è un sistema che riesce ad ingannare il giudice circa metà delle volte (Figura 1.3).

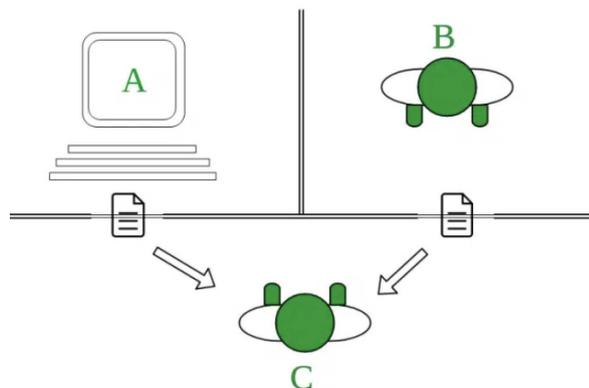


Figura 1.3: Esempio illustrativo Test di Turing

Col passare degli anni vi sono state diverse versioni più complete del test di Turing come:

- Test di Turing totale: è un test sviluppato dallo scienziato cognitivo Stevan Harnad, che aggiunge due ulteriori requisiti al test di Turing. Il giudice può anche testare le abilità percettive del soggetto, che coinvolgono la visione artificiale e la capacità del soggetto di manipolare oggetti, coinvolgendo la robotica;
- Test di Turing dell'esperto: è un test in cui il soggetto umano è un esperto in un dato campo, tale test è stato proposto da Edward Feigenbaum in un documento del 2003;
- Test del minimo segnale intelligente: è un test in cui le risposte consentite dai soggetti interrogati sono binarie, ossia vero/falso oppure sì/no, per concentrarsi sulla sola abilità di pensiero, eliminando la necessità di formulare una risposta in linguaggio naturale;
- Test di Turing inverso e CAPTCHA: nel test di Turing inverso viene messo alla prova un computer per verificare se è in grado di determinare se sta interagendo con un essere umano o con un altro computer. Un esempio comune di questo tipo di prova è il sistema CAPTCHA, il cui scopo è quello di verificare che un utente di un sito è, di fatto, un umano e non un bot;

### 1.2.2 Test della stanza cinese

Il test di Turing, seppur rappresenti una pietra miliare della filosofia dietro l'Intelligenza Artificiale, non rappresenta un metodo definitivo per riconoscere se un'entità possa essere definita umana. Negli anni, infatti, sono stati proposti un numero indefinito di controesempi che confutano la non affidabilità totale del test di Turing, come, ad esempio, il test della stanza cinese o il test di Winograd.

Il test della stanza cinese fu ideato da John Searle nell'articolo "*Minds, Brains and Programs*" pubblicato nel 1980, come controesempio rispetto alla teoria dell'Intelligenza Artificiale forte. L'esempio proposto da Searle risulta essere quasi analogo a quello ideato da Turing, dove un essere umano, a sua insaputa, interagisce con una macchina e il suo obiettivo è quello di capire con chi sta parlando. Definito il contesto iniziale, interviene lo stesso Searle che si posiziona al posto della macchina, si chiude dentro la stanza e interagisce con qualcuno all'esterno. Si suppone poi che la persona fuori parli cinese e che Searle non conosca minimamente la lingua. Si immagina, ancora, che sul tavolo della stanza vi siano una serie di caratteri cinesi che Searle dovrà utilizzare per rispondere. Inoltre, vi è un libro di istruzioni con alcuni insiemi di caratteri cinesi associati secondo delle regole scritte in inglese. Searle continua a non capire nulla del cinese, però comprende le informazioni in inglese che gli indicano come rispondere a qualsiasi domanda ricevuta in cinese. A questo punto Searle è in grado di intraprendere una conversazione in cinese ed ingannare l'interlocutore esterno del fatto che stia parlando con un madrelingua.

Di fatto, Searle non ha mai dovuto interpretare i simboli cinesi per capire la domanda, doveva semplicemente seguire un insieme di regole comprensibili da lui e rispondere in maniera adeguata. In conclusione, secondo il filosofo, allo stesso modo in cui egli esegue meccanicamente un ordine senza comprendere il cinese, un sistema esegue il programma scritto nel linguaggio di programmazione manipolando simboli di cui non ne conosce il significato operando, quindi, in maniera puramente sintattica (Figura 1.4).

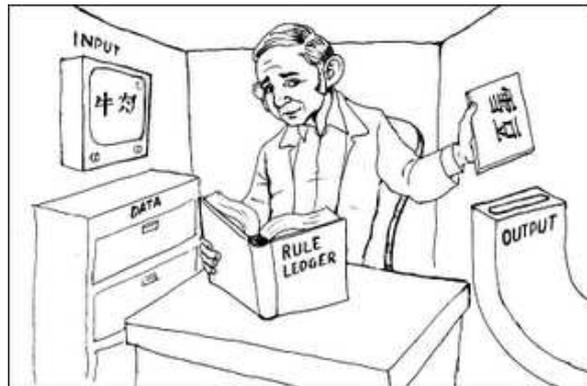


Figura 1.4: Esempio illustrativo Test della stanza cinese

### 1.2.3 Test di Winograd

Il test o schema di Winograd è un test che è stato sviluppato da un informatico statunitense di nome Terry Winograd nel 1972 e che tutt'ora funziona.

Progettato per essere un miglioramento del test di Turing, è un test a scelta multipla che impiega domande di una struttura ben definita, ossia istanze di quelli che vengono chiamati schemi di Winograd.

Il test di Winograd è beffardo dal punto di vista informatico per la sua semplicità. Una delle sue formulazioni tipiche è questa:

*“Il trofeo non ci stava nella valigia marrone perché era troppo grande.”*

Quest'ultima, risulta essere una frase banale e perfettamente comprensibile per noi esseri umani. Talmente comprensibile e ovvia che non ci rendiamo conto della sua ambiguità.

Quale dei due oggetti è troppo grande? La valigia o il trofeo? Per noi la risposta è ovvia, per il computer no. Infatti, una semplice analisi meccanica della frase, ossia dividere sostantivi, verbi e tanto altro, non basta a risolvere l'ambiguità. Bisogna avere una visione completa di ciò che è scritto, ossia conoscere cos'è un trofeo, una valigia, conoscere i loro rapporti e le dimensioni e, soprattutto, sapere che le valigie sono contenitori di oggetti e non viceversa.

Nel 2012 venne riproposto uno schema di Winograd dallo scienziato Hector Lavesque per mettere in luce le debolezze del test di Turing, in quanto, infatti, nei due anni successivi un chatBot di nome Eugene Goostman riuscì a superare il test di Turing ingannando l'interlocutore.

Le debolezze del test di Turing riguardavano 3 aspetti fondamentali:

- *L'inganno*: la macchina è costretta a costruire una falsa identità che non fa parte dell'intelligenza.
- *La conversazione*: molte interazioni si possono considerare “non intelligenti”, in quanto potremmo avere scherzi, digressioni sagaci senza, però, avere un ragionamento perspicace alla base.
- *La valutazione*: gli esseri umani commettono errori e i giudici spesso non sono d'accordo con i risultati.

Gli schemi di Winograd, quindi, hanno come obiettivo quello di mettere alla prova l'Intelligenza Artificiale che, secondo Lavesque, anche se supportata dall'elaborazione del linguaggio naturale, per risolvere questi compiti ha bisogno della conoscenza e del buon senso.

Tra 2016 e il 2018, Nuance Communications ha sponsorizzato un concorso offrendo un primo premio di 25.000 dollari per il programma che riuscisse ad ottenere un punteggio superiore al 90% (per confronto, gli esseri umani rispondono correttamente al 92-96%) alle domande dello schema di Winograd. Tuttavia, nessuno si è avvicinato a vincere il premio nel 2016 e il concorso nel 2018 è stato annullato per mancanza di prospettive. Di conseguenza, il premio è stato annullato.

#### 1.2.4 Dibattito filosofico

Le macchine possono pensare? Questa domanda affligge l'IA sin dalla sua nascita. Lo stesso Alan Turing, dopo aver proposto il "gioco dell'imitazione" definì nove possibili obiezioni che esprimevano come concetto il fatto che una macchina non potesse pensare e cercò di confutarle. Una di queste è riportata come segue: "Pensare è una funzione dell'anima immortale dell'uomo. Dio ha dato un'anima immortale ad ogni uomo e a ogni donna, ma non a qualsiasi altro animale né alle macchine. Perciò nessun animale e nessuna macchina possono pensare."

Il fatto di pensare risulta essere indissolubilmente legato al provare delle emozioni; fino a che una macchina non esprimerà pensieri o emozioni come provare piacere per i propri successi, dolore quando un componente si brucia, essere arrabbiato o abbattersi quando non può ottenere quello che desidera, non potremmo accettare che eguagli il cervello. Nessuna macchina è in grado di arrivare a tanto; di conseguenza non può realmente pensare.

### 1.3 Applicazioni dell'Intelligenza Artificiale

Dopo aver analizzato le origini ed il significato dell'Intelligenza Artificiale, possiamo cercare di comprenderne meglio il funzionamento.

Alla base dell'IA ci sono algoritmi, tecniche di computazione e soluzioni in grado di replicare il comportamento umano.

Le applicazioni, però, possono essere molteplici e riguardare diversi ambiti non solo industriali, ma anche domestici. Basti pensare ad applicazioni in campi come i sistemi di Smart Home per rendere la casa domotica; questi sono capaci di regolare temperatura, umidità e luminosità in base alle abitudini degli utenti mediante l'utilizzo della voce come input. Queste soluzioni agevolano la gestione delle case e ne migliorano il livello di comfort complessivo. Tali applicazioni sono attivate tramite l'utilizzo di un telecomando o di uno smartphone o del riconoscimento vocale.

Tuttavia, se è vero che l'Intelligenza Artificiale potrebbe essere applicata a diversi ambiti, non è destinata a diffondersi in essi alla stessa velocità. Considerato questo, bisogna fornire un quadro completo delle soluzioni di Intelligenza Artificiale adottate (o adottabili) dalle imprese. Per agevolare questo compito, l'Osservatorio "Artificial Intelligence" ha elencato sei classi di applicazioni, o soluzioni distinte, in base alle finalità d'utilizzo.

#### 1.3.1 Intelligent Data Processing

La classe di soluzioni degli Intelligent Data Processing è quella più ampia dal punto di vista delle applicazioni dell'Intelligenza Artificiale. Qui vi rientrano tutte quelle soluzioni che utilizzano algoritmi di Artificial Intelligence – su dati strutturati e non – per finalità collegate all'estrazione delle informazioni presenti nei dati stessi.

Le principali finalità che muovono le imprese nell'utilizzo di queste soluzioni di Intelligenza Artificiale sono: il Forecasting (Previsioni) e il Classification & Clustering (Classificazione e Raggruppamento).

### 1.3.2 Natural Language Processing (NLP)

Un'altra applicazione di Intelligenza Artificiale è il Natural Language Processing (NLP). Le tecniche di NLP si pongono l'obiettivo di creare sistemi in grado di favorire l'interazione e la comprensione uomo/macchina. L'NLP è un algoritmo di apprendimento che si occupa principalmente di testi.

Con questo termine intendiamo, una qualsiasi sequenza di parole che, in una lingua, esprime uno o più messaggi (come pagine web, post, tweet, informazioni aziendali).

### 1.3.3 Virtual Assistant/Chatbot

Tra le applicazioni di Intelligenza Artificiale, il Chatbot, o Virtual Assistant, è una delle soluzioni più diffuse tra le aziende italiane e internazionali. Gli assistenti virtuali sono strumenti capaci offrire assistenza 24/7 sia ai clienti che ai dipendenti. Inoltre, i ChatBot si prestano anche a diversi impieghi in ambito marketing, supporto alla vendita, HR Management, domotica e Ricerca e Sviluppo.

Un esempio di applicazione di Intelligenza Artificiale tramite Chatbot è ChatGPT. ChatGPT è, per definizione, un prototipo di chatbot capace di fornire risposte ai quesiti degli utenti in tempo reale e utilizzando l'elaborazione del linguaggio naturale. Lanciato sul mercato nel novembre 2022 da OpenAI, impresa statunitense, ha avuto un immediato successo. Dopo solo una settimana, infatti, la piattaforma registrava più di un milione di utenti al giorno: non solo curiosi ma anche addetti ai lavori e professionisti di diversi altri ambiti, tra cui il marketing. ChatGPT, infatti, sfruttando gli algoritmi di Intelligenza Artificiale, come quelli di apprendimento automatico, è in grado di svolgere moltissime funzioni, dal dialogo con gli utenti alla stesura di testi (come articoli, ricette, email, etc.), alla generazione di immagini (tramite la piattaforma PlaygroundAI), ma non solo; infatti, è in grado di analizzare e compilare snippet di codice in diversi linguaggi di programmazione.

### 1.3.4 Computer Vision

Tra le applicazioni di Intelligenza Artificiale, la Computer Vision è una delle più importanti. Questa studia gli algoritmi e le tecniche per permettere ai computer di raggiungere una comprensione di alto livello del contenuto di immagini o video. Gli avanzamenti di questa applicazione negli ultimi anni sono stati importanti. Questi hanno portato a soluzioni basate su descrizioni statistiche delle immagini e hanno progressivamente lasciato il passo a reti neurali addestrate su milioni di immagini.

### 1.3.5 Recommendation System

Di tutte le applicazioni finora viste, i Recommendation System sono le applicazioni di Intelligenza Artificiale più utilizzate. Non tutti sanno che gli algoritmi di raccomandazione sono il pilastro del modello di business di tutte le piattaforme social ed e-commerce; Amazon e Netflix ne fanno largamente utilizzo.

Alla base di tanti servizi digitali ci sono algoritmi di Intelligenza Artificiale che tengono traccia delle azioni dell'utente; confrontando, poi, tali applicazioni con quelle degli altri utenti, apprendono le preferenze e sono in grado di produrre suggerimenti più precisi al fine di migliorare l'esperienza dell'utente sulla piattaforma.

### 1.3.6 Soluzioni Fisiche di Intelligenza Artificiale

Le soluzioni fisiche di Intelligenza Artificiale, in Italia, sono ancora poco diffuse. Ovvero tra queste applicazioni, tre sono le categorie da prendere in esame: veicoli autonomi, ovvero mezzi di trasporto dotati di guida autonoma (Figura 1.5), Autonomous Robot, ovvero robot in grado di muoversi senza l'intervento umano, e Intelligent Object, ovvero oggetti in grado di compiere azioni senza l'intervento umano e di prendere decisioni in base alle condizioni dell'ambiente circostante.



**Figura 1.5:** Abitacolo di un veicolo a guida autonoma

*In questo capitolo si introdurrà il concetto di cloud computing e ci concentreremo sulla piattaforma di punta nel settore, ovvero Amazon Web Services (AWS). In particolare, si analizzeranno attentamente i diversi servizi offerti da AWS, si esamineranno l'architettura e i relativi costi di utilizzo.*

## **2.1 Introduzione al cloud computing e Amazon Web Services**

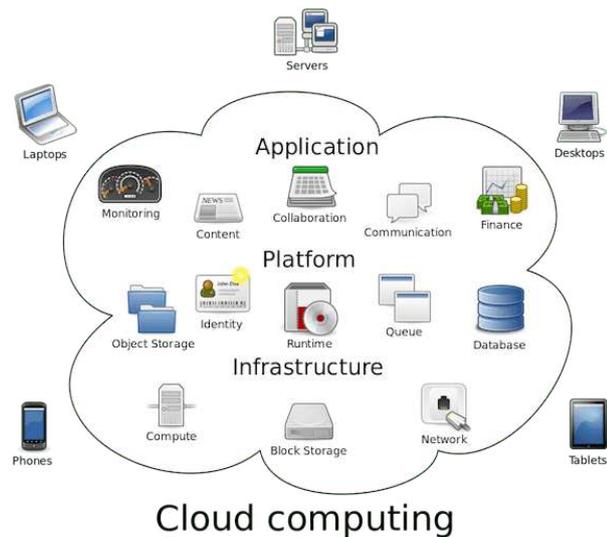
Il cloud computing rappresenta una rivoluzione nel modo in cui le aziende accedono, archiviano e utilizzano le risorse informatiche. Piuttosto che investire in infrastrutture fisiche, costose e complesse, esso consente alle aziende di utilizzare risorse informatiche su richiesta tramite Internet. In altre parole, offre un modello di distribuzione di servizi informatici in cui le risorse come l'elaborazione, lo storage e il networking sono fornite e gestite da fornitori esterni.

Una delle caratteristiche chiave del cloud computing è l'elasticità, che consente alle aziende di scalare rapidamente le risorse in base alle esigenze. Questo significa che le aziende possono facilmente aumentare o ridurre la capacità delle risorse informatiche in risposta a picchi di domanda o cambiamenti nelle esigenze aziendali, senza dover investire in hardware aggiuntivo. Il cloud computing offre, inoltre, una maggiore flessibilità e agilità, consentendo alle aziende di accedere alle risorse informatiche da qualsiasi luogo e in qualsiasi momento tramite Internet. Ciò permette una maggiore collaborazione, innovazione e mobilità per i dipendenti, oltre a semplificare la distribuzione e la gestione delle applicazioni. Il cloud computing garantisce, infine, una maggiore sicurezza e ridondanza rispetto alle infrastrutture tradizionali. I fornitori di servizi cloud investono pesantemente nella sicurezza delle loro piattaforme, utilizzando crittografia avanzata, monitoraggio continuo e misure di protezione dei dati per proteggere le informazioni sensibili dei clienti.

Con l'avvento del cloud computing, illustrato sinteticamente in Figura 2.1, una delle maggiori aziende che si è mossa per lo sviluppo di prodotti cloud è proprio Amazon, che nel marzo 2006 lancia Amazon Web Services.

### **2.1.1 Che cos'è Amazon Web Services**

Amazon Web Services (AWS) è un servizio di cloud computing che fornisce una vasta gamma di servizi informatici su richiesta (on-demand) tramite Internet. In altre parole, AWS offre un'infrastruttura tecnologica scalabile e affidabile, che consente alle aziende di ospitare le proprie applicazioni, archiviare i dati, e accedere a una varietà di servizi informatici



**Figura 2.1:** Esempio illustrativo di cloud computing

senza dover costruire e gestire la propria infrastruttura fisica. Con AWS, le aziende possono sfruttare una serie di servizi, tra cui: l'elaborazione, lo storage, il database, il networking, l'Intelligenza Artificiale, il Machine Learning, l'analytics, l'internet delle cose (IoT), e molto altro ancora. Questi servizi sono progettati per essere altamente scalabili, sicuri, affidabili ed efficienti rispetto ai costi, consentendo alle aziende di ridurre i costi operativi e di accelerare l'innovazione.

AWS gestisce una vasta rete di data center in tutto il mondo, distribuiti in regioni geografiche e zone di disponibilità. Questa infrastruttura globale consente alle aziende di distribuire le proprie applicazioni in modo affidabile e scalabile, migliorando la latenza, la conformità normativa e la resilienza. Inoltre, AWS fornisce un ecosistema ricco di strumenti di gestione, automazione, monitoraggio e sicurezza che semplificano e migliorano l'esperienza degli sviluppatori e degli amministratori di sistema. Questi strumenti consentono di gestire in modo efficiente le risorse, automatizzare processi ripetitivi, monitorare le prestazioni delle applicazioni e garantire la sicurezza dei dati e delle risorse. L'adozione di Amazon Web Services ha trasformato il modo in cui le organizzazioni concepiscono, sviluppano e distribuiscono le proprie applicazioni e i propri servizi, consentendo loro di essere più agili, innovative e competitive sul mercato.

Con una vasta gamma di clienti, che spaziano dalle startup alle grandi imprese, AWS continua a guidare l'innovazione nel settore del cloud computing e a fornire soluzioni all'avanguardia per le sfide tecnologiche di oggi e di domani.

### 2.1.2 Servizi principali di AWS

I principali servizi offerti da Amazon Web Services possono essere suddivisi nelle seguenti categorie:

#### ► Calcolo

Tra i servizi di calcolo attirano la nostra curiosità *Amazon EC2* (*Amazon Elastic Computer Cloud*) ed *Amazon Lambda*.

- Amazon EC2 permette ai suoi utilizzatori di eseguire facilmente e scalabilmente macchine virtuali su infrastrutture cloud di AWS, permettendo agli utenti di

scegliere tra una vasta gamma di macchine virtuali con diverse capacità di CPU, memoria, storage e rete in base alle proprie esigenze.

- Amazon Lambda è un servizio di calcolo serverless che permette di eseguire codice senza la necessità di gestire l'infrastruttura sottostante, addebitando dei costi solo per il tempo effettivo di esecuzione del codice.

► Rete

- Nel servizio di rete offerto da AWS troviamo *Amazon Virtual Private Cloud (VPC)*, che permette agli utenti di creare una sezione isolata e privata all'interno dell'infrastruttura di AWS collaborando con altri servizi come EC2, database, e così via, garantendo maggiore sicurezza ai suoi utilizzatori.

► Archiviazione e distribuzione di contenuti

AWS mette, inoltre, a disposizione dei suoi fruitori dei servizi di archiviazione e distribuzione dei contenuti. Tra i vari servizi riportiamo *Amazon S3 (Amazon Simple Storage Service)* ed *Amazon Glacier*.

- Amazon S3 è il servizio di archiviazione di oggetti altamente scalabile e affidabile, nonché il più diffuso per praticità e semplicità nel recupero degli oggetti. Tramite S3 possiamo memorizzare qualsiasi tipo di oggetto digitale, come immagini, video, documenti e file di backup, offrendo una disponibilità elevata e una durabilità dei dati del 99,9 % e garantendo che i dati siano sempre accessibili e protetti.
- Amazon Glacier è un servizio analogo ad Amazon S3, ma per un'archiviazione a lungo termine e a basso costo progettato per dati e informazioni che non devono essere accedute in maniera frequente o modificate.

► Basi di dati

Oltre all'archiviazione e alla distribuzione di contenuti già citati in precedenza, Amazon Web Services permette agli utenti di definire al suo interno delle vere e proprie basi di dati altamente scalabili e affidabili. I servizi sono: *Amazon Relational Database Service*, *Amazon Redshift*, *Amazon SimpleDB* e *Amazon DynamoDB*.

- Amazon Relational Database Service ed Amazon Redshift si distinguono in quanto sono database relazionali, simili a MySQL e adatti a ricevere delle query SQL.
- Amazon SimpleDB ed Amazon DynamoDB sono database non relazionali, NoSQL, adatti in genere per strutture dati leggere e di facile gestione, ma che, comunque, consentono di avere un sistema all'avanguardia e perfettamente manutenibile e scalabile.

► Servizi applicativi

Amazon Web Services offre, inoltre, una serie di servizi applicativi progettati per semplificare e migliorare diversi aspetti delle applicazioni aziendali. Tra questi troviamo *Amazon Flexible Payments Service (FPS)* ed *Amazon Simple Email Service (SES)*.

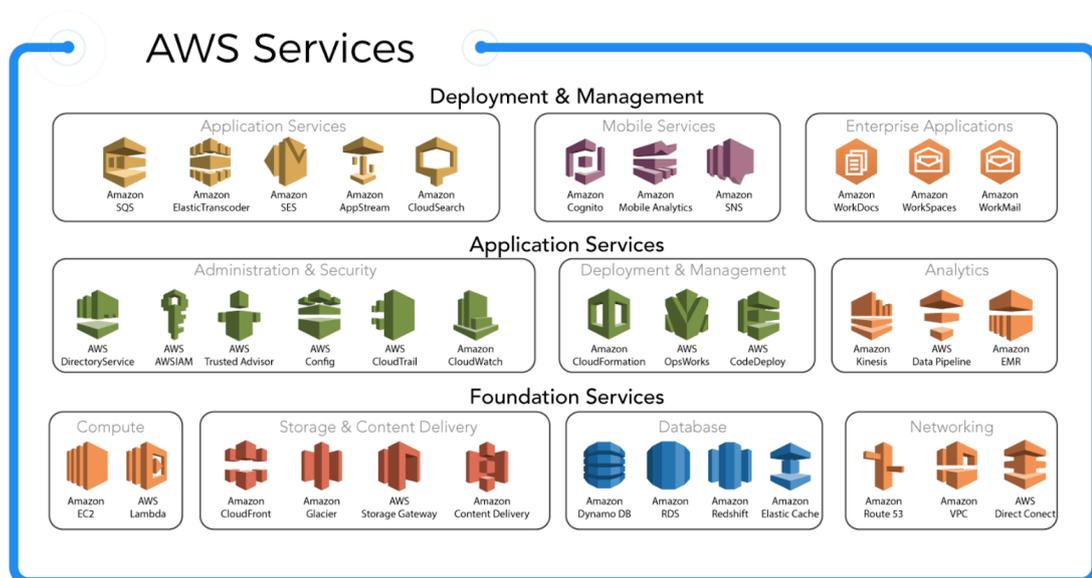
- Amazon FPS è un servizio di pagamento online che consente alle aziende di accettare pagamenti online in modo flessibile e sicuro. Esso offre, inoltre, diverse funzionalità, come, ad esempio, gestione di pagamenti ricorrenti, supporto per pagamenti internazionali, e, infine, tutela su eventuali frodi. FPS può, inoltre, essere integrato in applicazioni o siti web per pagamenti online utilizzando API semplici e documentazione dettagliata.

- Amazon SES, invece, è un servizio di invio email scalabile e affidabile che consente alle aziende di inviare email in modo efficiente e sicuro.

#### ► Analytics

Infine, tra i vari servizi proposti da AWS abbiamo la sezione Analytics, che consente alle aziende di elaborare, analizzare e trarre insight dai propri dati in modo rapido ed efficiente. Questi servizi includono strumenti per l'elaborazione dei dati in batch e in tempo reale, l'analisi dei dati, il machine learning e la visualizzazione dei dati. Tra questi appunto, abbiamo il machine learning (ML), che consente alle aziende di creare dei modelli utilizzando dati esistenti per effettuare previsioni o classificare dati. Come vedremo in seguito, Amazon ML semplifica il processo di creazione, addestramento e distribuzione di modelli di machine learning, consentendo, anche, agli sviluppatori senza esperienza nel machine learning di utilizzare questa tecnologia in modo efficace.

I prodotti di Amazon Web Services appena proposti possono essere riassunti sinteticamente nella Figura 2.2, dove i vari servizi offerti vengono raggruppati nelle loro categorie.



**Figura 2.2:** Servizi offerti da Amazon Web Services

### 2.1.3 Architettura ed infrastruttura AWS

L'architettura e l'infrastruttura di Amazon Web Services costituiscono la base su cui poggia l'intera piattaforma di cloud computing. Quest'ultima si basa su regioni di utilizzo e tipologie di servizi per regione. Essendo AWS un servizio di cloud computing a livello globale, possiamo verificare, semplicemente autenticandoci all'interno della piattaforma, che, in base alla regione in cui ci troviamo, alcuni servizi saranno disponibili ed altri no.

Di fatto, Amazon opera attraverso una rete globale di data center distribuiti in diverse regioni geografiche. Ogni regione è composta da una o più zone di disponibilità fisicamente separate, collegate tra loro tramite reti ad alta velocità. Il fatto di avere regioni localizzate in diverse parti del mondo garantisce agli utenti la possibilità di posizionare le proprie risorse quanto più vicino possibile agli utenti finali per migliorare la latenza e la disponibilità delle applicazioni.

Le regioni attualmente disponibili sulla piattaforma sono ben cinque, ciascuna con le sue zone di disponibilità (definite in numero all'interno delle parentesi), che raccolte macroscopicamente sono:

- Stati Uniti Orientali e Occidentali (3);
- Asia Pacifica (6);
- Canada (1);
- Europa Centrale (5);
- Sud America (1);

Nella regione dell'Europa Centrale, tra i cinque data center proposti da AWS, nessuno di questi è collocato in Italia.

#### 2.1.4 Costi di utilizzo

Amazon Web Services mette a disposizione dei suoi utilizzatori tutti i servizi citati precedentemente, purché questi ultimi, effettuino una registrazione di un account all'interno della piattaforma, nonché l'attivazione di un piano che potrebbe essere in partenza gratuito.

Essendo Amazon Web Services un servizio di cloud computing on-demand, esso mette a disposizione la possibilità di fruire dei suoi servizi in modalità Pay per Use, garantendo, così, di risparmiare in alcune situazioni, pagando semplicemente per le risorse o per il tempo che vengono effettivamente utilizzati, senza pagare ulteriori spese. Il piano gratuito include un set di risorse per servizi specifici, come, ad esempio, un certo numero di ore di utilizzo di Amazon EC2 oppure una quantità limitata di storage su Amazon S3. Ad esempio, utilizzando un tool di AWS di nome Amazon Rekognition (che vedremo in seguito), se l'addestramento del modello non supera i 30 minuti effettivi, esso sarà gratuito; se, invece, supera i 30 minuti, verranno addebitati sulla carta di credito/debito, registrata dall'utente nel suo account, 1 dollaro per ogni ora di addestramento del modello. Ovviamente, Amazon mette a disposizione dei piani che, per aziende che devono effettuare un numero elevato di addestramenti, elaborazioni, immagazzinamento dati, e così via, sulla piattaforma, permettono ad esse di risparmiare.

---

## Esperienze di Sentiment Analysis

---

*In questo capitolo si introdurrà l'elaborazione del linguaggio naturale (Natural Language Processing - NLP) che è alla base del concetto di Sentiment Analysis, di cui verranno discussi approcci e processi di funzionamento. In seguito verrà introdotto Amazon Comprehend, un tool della famiglia di Amazon Web Services con il quale verrà trattata l'analisi del sentiment precedentemente proposta. Si proseguirà, quindi, con un test che metterà alla prova l'efficienza della stessa tramite delle recensioni provenienti dalla nota piattaforma TripAdvisor.*

### 3.1 Introduzione alla Sentiment Analysis

Prima di addentrarci nella Sentiment Analysis, è essenziale comprendere il contesto più ampio dell'elaborazione del linguaggio naturale (Natural Language Processing - NLP), già introdotto marginalmente nel Capitolo 1.

L'NLP è un campo dell'Intelligenza Artificiale che si occupa di consentire alle macchine di comprendere, interpretare e generare il linguaggio umano in modo naturale. L'NLP ha numerose applicazioni pratiche, tra cui la traduzione automatica, la classificazione di testo, i chatbot, la comprensione del sentimento e la sintesi automatica di testo. All'interno dell'NLP la Sentiment Analysis (analisi dei sentimenti) è una tecnica particolarmente significativa. Essa mira a determinare e comprendere le emozioni, i sentimenti e le opinioni espressi in un testo. Utilizzando algoritmi e modelli predittivi, la Sentiment Analysis è in grado di identificare se un testo esprime un'opinione positiva, negativa o mista rispetto a un determinato argomento o prodotto.

#### 3.1.1 Che cos'è la Sentiment Analysis

L'analisi del sentiment, quindi, è un processo di analisi di un testo digitale, in grado di determinare se il tono emotivo del messaggio è positivo, negativo o misto. Oggi le aziende dispongono di grandi volumi di dati di testo, come e-mail, trascrizioni di chat dell'assistenza clienti, commenti sui social media e recensioni. Gli strumenti di analisi del sentiment possono scansionare questi testi per determinare in maniera automatica l'atteggiamento dell'autore nei confronti di un determinato argomento. Di conseguenza, le aziende utilizzano le informazioni derivanti dall'analisi del sentiment per migliorare il servizio clienti e aumentare la reputazione del marchio.

La Sentiment Analysis, sostanzialmente, risulta essere importante sotto diversi aspetti:

- *Fornisce informazioni oggettive*: le aziende possono evitare i pregiudizi personali associati ai revisori umani utilizzando strumenti di analisi del sentiment basati sull'Intelligenza Artificiale. Di conseguenza, esse ottengono risultati coerenti e obiettivi come risultato dell'analisi delle opinioni dei clienti.
- *Crea prodotti e servizi migliori*: un sistema di Sentiment Analysis aiuta le aziende a migliorare i propri prodotti e servizi, sulla base di feedback reali e specifici dei clienti. Le tecnologie di Intelligenza Artificiale identificano oggetti o situazioni del mondo reale che i clienti associano a emozioni negative.
- *Analizza su larga scala*: un sistema di Sentiment Analysis estrae e gestisce grandi quantità di informazioni contenenti anche dati non strutturati, come, ad esempio e-mail, rendendo il processo per le aziende semplice e ad un costo accessibile.
- *Fornisce risultati in tempo reale*.

### 3.1.2 Processo di funzionamento della Sentiment Analysis

La Sentiment Analysis è composta di due fasi principali: la preelaborazione e l'analisi delle parole chiave. Nella fase di preelaborazione l'analisi del sentiment identifica le parole chiave per evidenziare il messaggio centrale del testo. In questa fase abbiamo tre sottofasi:

- *la tokenizzazione*: suddivide una frase in diversi elementi o token;
- *la lemmatizzazione*: converte le parole nella loro radice; ad esempio, la radice di "sono" è "essere";
- *la rimozione delle stop word*: vengono filtrate tutte quelle parole che non aggiungono un valore significativo alla frase come ad esempio, *con, per, a e di*.

### 3.1.3 Approcci di Sentiment Analysis

Esistono tre approcci principali utilizzati dai software di Sentiment Analysis, ovvero:

► **Approccio basato su regole:**

In questo approccio, vengono identificate, classificate e vengono assegnati dei punteggi ad alcune parole chiave specifiche in base a dei lessici predeterminati. I lessici sono raccolte di parole che rappresentano l'intento, l'emozione e l'umore dello scrittore. Gli esperti di marketing assegnano dei punteggi di sentiment a lessici positivi e negativi, per riflettere il peso emotivo delle diverse espressioni. Sostanzialmente, per determinare se una frase è positiva, negativa o mista, il software esegue la scansione delle parole elencate nel lessico e riassume il punteggio del sentiment. Il punteggio finale viene, poi, confrontato con i limiti del sentiment per determinare la rilevanza emotiva complessiva. Un sistema di Sentiment Analysis basato su regole è semplice da impostare, ma è difficile da scalare. Ad esempio, è necessario espandere continuamente i lessici quando si scoprono nuove parole chiave per trasmettere sempre al meglio l'intento nell'input di testo. Inoltre, questo approccio può non essere accurato quando si elaborano frasi influenzate da culture diverse.

► **Approccio basato sull'esperienza:**

In questo approccio si utilizzano tecniche di machine learning e algoritmi di classificazione del sentiment, come le reti neurali e il deep learning, per insegnare al software informatico come identificare il sentiment emotivo dal testo. Questo processo comporta

la creazione di un modello di analisi del sentiment e un addestramento ripetuto su dati noti, per permettere al sistema di riconoscere il sentiment con elevata precisione in dati sconosciuti.

*In che cosa consiste l'addestramento?*

Sostanzialmente, i data scientist utilizzano set di dati di Sentiment Analysis che contengono un gran numero di esempi. Il software di ML utilizza i set di dati come input e si addestra per raggiungere la conclusione predeterminata. Addestrandosi con un gran numero di esempi diversi, il software differenzia e determina in che modo le diverse disposizioni delle parole influenzano il punteggio finale del sentiment. L'analisi del sentiment con ML risulta essere vantaggiosa perché elabora accuratamente un'ampia gamma di informazioni di testo. Finché il software viene addestrato con un numero sufficiente di esempi, la Sentiment Analysis con ML può prevedere con precisione il tono emotivo dei messaggi. Tuttavia, un modello di ML addestrato è specifico per un'area di business in particolare. Ciò significa, ad esempio, che il software di analisi del sentiment, addestrato con dati di marketing, non può essere utilizzato per il monitoraggio dei social media senza essere addestrato nuovamente.

► Approccio ibrido:

L'analisi ibrida della Sentiment Analysis elabora approcci basati su regole e ML combinandoli. Essa utilizza le funzionalità di entrambi i metodi per ottimizzare la velocità e la precisione nella derivazione dell'intento contestuale nel testo. Tuttavia, sono necessari tempo e sforzi tecnici per riunire i due diversi sistemi.

## 3.2 Amazon Comprehend

Per testare le potenzialità della Sentiment Analysis combinate con il Natural Language Processing verrà adoperato il sistema "Amazon Comprehend", anch'esso facente parte del servizio AWS (Amazon Web Services).

### 3.2.1 Che cos'è Amazon Comprehend

Amazon Comprehend è il servizio di elaborazione del linguaggio naturale (NLP) offerto da AWS, una piattaforma in grado di raccogliere informazioni e relazioni in un testo grazie al Machine Learning. Esso fornisce, oltre alla Sentiment Analysis, servizi di riconoscimento delle entità personalizzati, classificazione personalizzata, estrazione di frasi chiave, riconoscimento delle entità e altre tipologie di API per permettere di integrare l'elaborazione del linguaggio naturale nelle applicazioni.

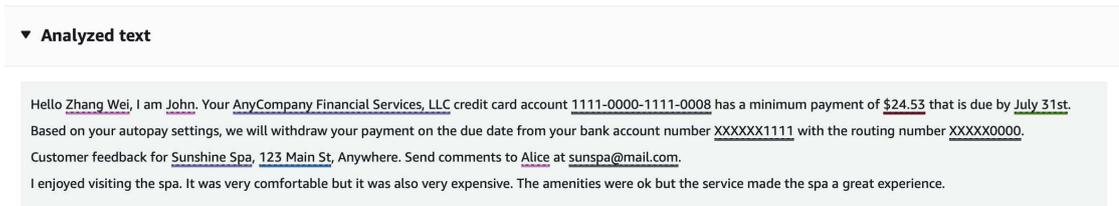
### 3.2.2 Funzionalità di Amazon Comprehend

La piattaforma di Natural Language Processing offerta da AWS mette a disposizione un vasto ventaglio di funzionalità avanzate per analizzare il testo. In seguito esamineremo brevemente ciascuna di queste funzionalità, soffermandoci successivamente sulla classificazione del sentiment. Amazon Comprehend offre anche la possibilità di rilevare entità specifiche all'interno del testo, come nomi propri, luoghi, organizzazioni e attività. Queste informazioni possono essere utilizzate per categorizzare documenti, personalizzare il servizio clienti e identificare relazioni tra entità diverse all'interno del testo. Inoltre, la piattaforma supporta la traduzione automatica di testo da una lingua all'altra, rendendo possibile l'analisi del testo in molte lingue diverse. Questa funzionalità è particolarmente utile per le aziende che operano a livello internazionale.

Queste rappresentano solo alcune delle possibilità offerte da Comprehend, le quali forniscono alla piattaforma una capacità di comprendere il contenuto e il significato del testo proposto in modo preciso e affidabile. Di seguito si riportano tutte le funzionalità offerte da Amazon Comprehend:

► Entities:

Il riconoscimento di Entità integrato in Amazon Comprehend permette l'analisi di un breve testo (Figura 3.1), con lo scopo di identificarne i concetti principali ed etichettarli a seconda di un dataset predefinito di tipologie.



**Figura 3.1:** Cattura del testo analizzato da parte di Amazon Comprehend

Come possiamo notare in Figura 3.1, dopo aver analizzato il testo sottoposto, il sistema sottolineerà con diversi colori le diverse tipologie di entità che è riuscito ad individuare. In aggiunta a ciò, esso prende una sezione "Results" dove verrà fornita una lista completa di ogni entità trovata, con le relative etichette e percentuali di confidenza. L'insieme di etichette fornito da Amazon Comprehend contiene nomi propri, luoghi, organizzazioni, date e orari, numeri, pressioni, temperature, peso, lunghezze e valute.

► Key Phrases:

Un'altra funzionalità offerta da Amazon Comprehend è quella che permette di identificare, nel testo proposto, le Key Phrases. Il sistema identifica le frasi chiave come una serie di parole che formano un concetto significativo. Sostanzialmente, attraverso un'analisi semantica del testo, analizza e identifica quelle frasi che sono più significative per il contesto, fornendo all'utente, in una sezione chiamata "Results", una lista di tutti gli elementi riconosciuti con la relativa percentuale di confidenza.

► Language:

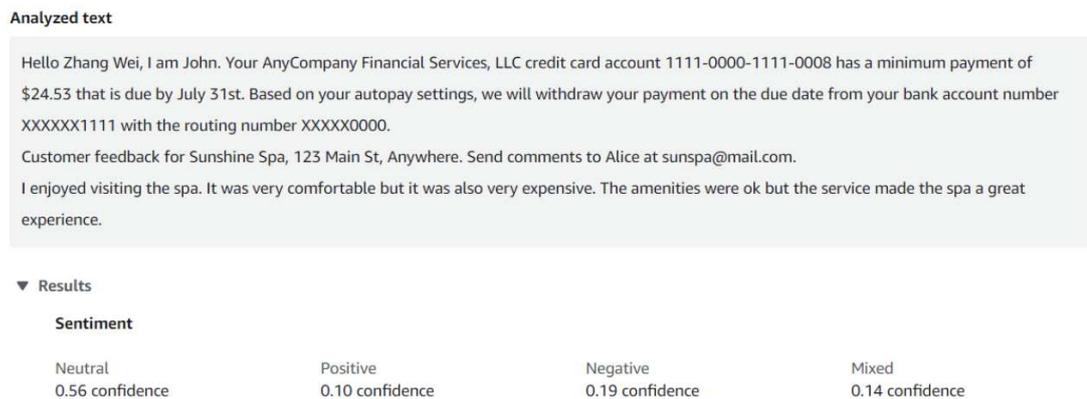
Il riconoscimento della lingua utilizzata nel testo rappresenta la funzione più basilica, ma più fondamentale, di Amazon Comprehend. Il sistema supporta una vasta gamma di linguaggi tra i quali: inglese, tedesco, francese, spagnolo, italiano, portoghese, giapponese, cinese semplificato, coreano, arabo, russo, turco, olandese, svedese, danese, norvegese, finlandese, indonesiano e malese.

► PII:

Amazon Comprehend è, inoltre, in grado di identificare le PII, ovvero, le Personal Identifiable Informations, informazioni personali identificabili. Questa funzionalità è utile per i clienti che devono gestire e proteggere le informazioni personali sensibili. Amazon Comprehend è in grado di identificare le seguenti informazioni personali: nome e cognome, indirizzo, numero di telefono, email, codice fiscale, numero carta di credito o debito, data di nascita, sesso e luogo di nascita.

► Sentiment:

La funzionalità su cui abbiamo posto il nostro interesse in Amazon Comprehend riguarda, appunto, il Sentiment. Come già presentata in precedenza, e come verrà discussa in seguito, essa riguarda la possibilità di analizzare il tipo di sentimento che un testo proposto ci suggerisce (Figura 3.2). L'emozionalità espressa in uno scritto viene suddivisa in sentimento positivo, negativo o misto. Il sistema utilizza un modello di apprendimento automatico avanzato per analizzare il contenuto del testo e determinarne il sentimento espresso. Questa funzionalità è utile in molti ambiti, come l'analisi delle opinioni sui prodotti oppure la valutazione di recensioni scritte dai clienti.



**Figura 3.2:** Esempio di Sentiment Analysis

► Targeted sentiment:

In aggiunta alla sezione "Sentiment", Amazon Comprehend fornisce all'utente la possibilità di analizzare il sentimento mirato, ovvero l'emozionalità espressa dalle singole parti del testo. Questa funzionalità fornisce una comprensione più dettagliata del sentimento espresso in uno scritto rispetto alla semplice analisi generale.

► Syntax:

L'ultima funzionalità proposta da Amazon Comprehend, chiamata Syntax, ha come obiettivo quello di analizzare la struttura del testo proposto e identificarne elementi come frasi, parole, verbi, soggetti e altri elementi linguistici organizzandoli successivamente in una struttura sintattica. Tali informazioni possono essere utilizzate in molte applicazioni, tra cui la creazione di modelli di chatbot, l'analisi del customer feedback e la creazione di sistemi di ricerca e di classificazione delle informazioni.

### 3.3 TripAdvisor ed Amazon Comprehend per test di Sentiment Analysis

Per testare ed avere una visione complessiva del servizio di Sentiment Analysis proposto da Amazon Comprehend sono state prese in esame diverse recensioni, redatte da alcuni utenti su una piattaforma di opinioni, e, tramite l'analisi del sentiment, è stata verificata la congruenza della recensione con una valutazione espressa in pallini ad essa associati dall'utente stesso.

Come sito per l'estrazione delle recensioni è stato scelto TripAdvisor, una piattaforma online globale dedicata alle recensioni e alle opinioni degli utenti riguardanti varie attività

legate al settore turistico. L'analisi del sentiment, associa ad ogni testo un sentiment complessivo che varia da 0 ad 1 circa, dove con zero si esprime un sentimento estremamente negativo, mentre con 1 un sentimento totalmente positivo. Proprio per questo, essendo il minimo di pallini associabili ad una recensione uguale ad 1 ed il massimo uguale a 5, si è deciso di associare ad ogni pallino di una recensione un range crescente e variabile di 0.20, in modo tale da esprimere un paragone diretto tra i pallini associati alla recensione ed i risultati ottenuti dall'analisi del sentiment (Figura 3.3). Ad esempio, se dovessimo avere una recensione totalmente negativa, e di conseguenza un solo pallino associato alla recensione da parte dell'utente, dovremo ottenere dalla Sentiment Analysis di quella recensione un valore di ritorno compreso tra 0.01 e 0.20 compreso.



**Figura 3.3:** Logo di TripAdvisor con range associati ad ogni pallino

Da TripAdvisor sono state estratte 3 recensioni di un hotel presente sulla piattaforma, di cui per privacy non citeremo il nome, utili ai fini del test.

### 3.3.1 Recensione positiva

In questa recensione (Figura 3.4) è stato associato dall'utente il massimo dei pallini disponibili, e, di conseguenza, ci si aspetta dall'analisi del sentiment un risultato decisamente positivo, con valori, quindi, compresi tra 0.80 e 0.99.

Ecco il risultato dell'analisi:



#### **Soggiorno perfetto per una settimana sugli sci**

Situato a 3 km dagli impianti di risalita di Grotte/Fortini questo albergo è il migliore in termini di qualità prezzo. Noi abbiamo soggiornato 4 notti con formula mezza pensione. La camera non era grandissima ma comoda e pulita e con televisore con Netflix incluso, il wifi gratuito, la colazione ricca e abbondante (a buffet). Il menù della mezza pensione offriva un primo e un secondo della tradizione trentina, frutta o dolce, acqua, caffè e vino incluso e inoltre abbiamo quasi sempre preso un amaro non incluso ma sempre offerto! Il personale è sempre stato disponibile e gentile. In conclusione qualora dovessi tornare a Madonna di Campiglio sicuramente ritornerei qui!

Mostra meno ▲

**Figura 3.4:** Screenshot di una recensione positiva su TripAdvisor

- Neutral Sentiment: 0.00 confidence;
- Positive Sentiment: 0.99 confidence;
- Negative Sentiment: 0.00 confidence;
- Mixed Sentiment: 0.00 confidence;

In questo caso la Sentiment Analysis riporta un valore congruo a quello associato dall'utente nella recensione presa in esame, e possiamo, di conseguenza, concludere che l'analisi è corretta.

### 3.3.2 Recensione negativa

In questa recensione (Figura 3.5) è stato associato dall'utente il minimo dei pallini disponibili, e, di conseguenza, ci si aspetta dalla Sentiment Analysis un risultato decisamente negativo.

Ecco il risultato dell'analisi:



#### Un'orrenda bettola

Quello che si definisce gergalmente un postaccio. Camere anguste sporche con disposizione dei letti inquietanti. Letti scomodissimi. Nessun sapone in dotazione ne'phon. Alloggi da erasmus o persone molto povere.

[Mostra meno](#) ▲

**Figura 3.5:** Screenshot di una recensione negativa su TripAdvisor

- Neutral Sentiment: 0.00 confidence;
- Positive Sentiment: 0.00 confidence;
- Negative Sentiment: 0.99 confidence;
- Mixed Sentiment: 0.00 confidence;

Tale risultato ci mostra come il sentimento predominante sia Negative, al quale viene associato un punteggio di 0.99, ossia il massimo associabile.

In questo caso, per avere un confronto effettivo con i pallini associati ci basta considerare il sentiment Positive uguale ad 1, al quale sottraiamo il sentiment Negative di 0.99, ottenendo così come risultato 0.01 Positive. Di fatto, come per definizione dei vari range imposti precedentemente per ogni pallino, si può verificare che il primo pallino corrisponde ad un range che va da 0.01 a 0.20, che combacia perfettamente con il risultato dell'analisi.

In questo caso l'analisi del sentiment riporta un valore congruo a quello associato dall'utente nella recensione presa in esame e possiamo concludere che l'analisi era corretta.

### 3.3.3 Recensione mista

In questa recensione (Figura 3.6), sono stati associati dall'utente un numero di pallini equivalente a 3, il che, determina che il valore del sentiment dovrà essere compreso tra un range che va da 0.40 a 0.60 inclusi.

Ecco il risultato dell'analisi:



#### Breve soggiorno in occasione del Giro D'Italia

Breve soggiorno in occasione del Giro D'Italia, peccato non aver usufruito del ristorante, spero di tornare presto. La struttura non rappresenta il massimo del comfort, ma la disponibilita' del personale e la bellezza di Madonna di Campiglio, mi suggeriscono di tornarci se il prossimo viaggio mi porterà in questo scorcio del divino Trentino...

Mostra meno ▲

**Figura 3.6:** Screenshot di una recensione mista su TripAdvisor

- Neutral Sentiment: 0.00 confidence;
- Positive Sentiment: 0.01 confidence;
- Negative Sentiment: 0.00 confidence;
- Mixed Sentiment: 0.98 confidence;

In questo caso possiamo notare come il sentiment Positive sia 0.01, e che la parte restante del punteggio dell'analisi sia stata associata ad un sentiment Mixed.

Questo accade quando durante la Sentiment Analysis il sistema riscontra un sentimento da parte dell'utente che non sia nè troppo entusiasta nè troppo insoddisfatto, e si trova in una posizione, come dice il termine stesso, un po' incerta. Per effettuare un confronto con i pallini associati alla recensione, ci basterà semplicemente dividere la componente Mixed in due parti e associarle alla Positive e alla Negative, tenendo, però, in considerazione che l'analisi ha comunque riscontrato una tendenza leggermente più positiva che negativa, in quanto alla Positive è stato associato un punteggio di 0.01. Infatti, risulterà che la Positive finale avrà un valore di 0.50, compreso nel range che va da 0.41 a 0.60, utile all'associazione del terzo pallino.

Anche in questo caso, l'analisi del sentiment riporta un valore congruo a quello associato dall'utente nella recensione presa in esame e possiamo riportare un'analisi corretta.

### 3.3.4 Recensione sarcastica

Finora, Amazon Comprehend e l'analisi del sentiment hanno risposto in maniera egregia alle nostre esigenze. Tuttavia, si è deciso di testare il tool con un'ulteriore recensione, per verificare se il sistema fosse in grado di recepire quando una recensione adotti dei termini sarcastici, potendo destare, quindi, problemi di comprensione. La frase che abbiamo preso in considerazione è la seguente:

*“Sì, ottimo. Ci sono volute solo tre ore prima che l'hotel mi rispondesse al telefono per venire a recuperare me e la mia famiglia sulle piste...”*

Ed ecco il risultato dell'analisi del sentiment:

- Neutral Sentiment: 0.00 confidence;
- Positive Sentiment: 0.99 confidence;
- Negative Sentiment: 0.00 confidence;
- Mixed Sentiment: 0.00 confidence;

Come è possibile notare, l'analisi ha fornito un sentiment Positive massimo, in una recensione in cui l'utente stava descrivendo un disservizio della struttura ospitante. Questo è dovuto dalla presenza della parola "Ottimo", che il software indica correttamente come valore positivo, ma decontestualizzandola dal testo e non recependo che si tratta di una recensione sarcastica. Sostanzialmente, in un'analisi corretta, ci saremmo aspettati un valore del sentiment totalmente negativo.

Questo ci fa capire che il sistema messo a disposizione da Amazon è utilissimo ma ha sempre bisogno di un controllo umano.

### 3.3.5 Limiti della Sentiment Analysis

Nonostante i progressi nelle tecnologie di elaborazione del linguaggio naturale (NLP), la comprensione del linguaggio umano è una sfida per le macchine. Nell'esempio mostrato in precedenza è stato messo alla prova il sistema con una recensione sarcastica, ma vi sono altre situazioni in cui Comprehend può entrare in difficoltà come, ad esempio, i casi di Negazione o Multipolarità.

► Negazione:

La negazione è l'uso di parole negative per trasmettere un'inversione di significato nella frase. Ad esempio, *"Non direi che l'abbonamento è costoso"*. Gli algoritmi di analisi del sentiment potrebbero avere difficoltà a interpretare correttamente tali frasi, in particolare se la negazione avviene su due frasi, come nel caso di *"Pensavo che l'abbonamento fosse economico. Non lo era"*.

► Multipolarità:

La multipolarità si verifica quando una frase contiene più di un sentiment. Un esempio è una recensione di un prodotto che dice: *"Sono soddisfatto della struttura robusta, ma non mi ha impressionato il colore"*. In questi casi, diventa difficile per il software interpretare il sentiment finale.

---

## Esperienze di Computer Vision

---

*In questo capitolo si introdurranno il concetto di Computer Vision e le sue funzionalità. In seguito verrà introdotto Amazon Rekognition, un tool della famiglia di Amazon Web Services con il quale, verranno effettuati dei test del prodotto su demo offerte dalla piattaforma. Infine verrà definito sul sistema un Custom Labels per la creazione e, la verifica di un modello personalizzato.*

### 4.1 Introduzione alla Computer Vision

Prima di analizzare come la Computer Vision e il riconoscimento delle immagini possano essere utilizzati dal punto di vista di un utente, è importante approfondire che cos'è e come funziona questa tecnologia che sta alla base di moltissimi algoritmi AI-based.

#### 4.1.1 Che cos'è la Computer Vision

La Computer Vision, o visione artificiale, è una branca dell'Intelligenza Artificiale (AI) che studia e programma algoritmi e tecniche che consentono ai computer di replicare i processi e le funzioni dell'apparato visivo umano e di rilevare ed interpretare informazioni importanti attraverso un'immagine digitale, un video o altri input visivi.

Con la Computer Vision, una macchina non solo può riconoscere oggetti, animali o persone presenti all'interno di un'immagine digitale o di una sequenza video, ma può anche estrapolare informazioni utili, interpretare i dati ricavati, elaborarli e intraprendere azioni o effettuare segnalazioni sulla base dei dati ottenuti. Tramite tale processo, il computer può capire i contenuti presenti dentro ad un'immagine, ricostruire un contesto e attribuire un vero e proprio significato a ciò che rappresenta.

In sostanza, possiamo dire che se l'AI è la disciplina che dona la capacità ai computer di pensare, o per lo meno, provare a pensare, la Computer Vision è quella che dona ad essi la capacità di vedere, capire e interpretare. Per poter funzionare e riuscire a interpretare i contenuti presenti all'interno di un'immagine, i sistemi di Computer Vision devono essere prima "addestrati" attraverso un processo di Machine Learning e utilizzando una grande quantità di immagini catalogate, le quali saranno la base del dataset che permetterà all'algoritmo di riconoscere quelle successive in modo intelligente.

Il processo è simile al sistema di apprendimento umano; infatti i nostri occhi si sono allenati per anni e anni nel distinguere oggetti, riuscendo, infine, a capire e interpretare. Ovviamente un sistema di Computer Vision non dispone di retine, nervi ottici e di una

corteccia visiva, e per essere addestrato vengono utilizzate telecamere, algoritmi, dati e immagini opportunamente etichettate.

#### 4.1.2 Come funziona la Computer Vision

La Computer Vision è basata sulle più avanzate tecniche di Machine Learning e su una grande mole di dati che consentono alla macchina di interpretare e capire un'immagine con prestazioni paragonabili a quelle della vista umana. Ma come fanno le macchine a riconoscere gli oggetti presenti in un'immagine? Tutto il sistema è basato su 3 fasi fondamentali:

- *Acquisizione dell'immagine*: le immagini o le sequenze video vengono acquisite dal computer, anche in tempo reale, tramite foto o tecnologia 3D e video per finalità di analisi.
- *Elaborazione dell'immagine*: attraverso dei modelli di Deep Learning la macchina riesce a elaborare l'immagine mediante specifici task. Questi modelli vengono addestrati precedentemente tramite il caricamento di migliaia e migliaia di immagini etichettate e pre-identificate.
- *Interpretazione dell'immagine*: infine la macchina identifica, comprende e classifica l'immagine elaborata e, se necessario, intraprende un'azione o una segnalazione.

Gli algoritmi di Computer Vision possono effettuare operazioni più o meno avanzate su un'immagine per interpretarla, a seconda delle tecniche impiegate e del tipo di task effettuato. In particolare, le operazioni che essi possono compiere sono le seguenti:

- *Classificazione dell'immagine (Image Classification)*: analisi del contenuto dell'immagine e attribuzione di un'etichetta di riconoscimento (ad esempio, cane, uomo, lampada, etc.).
- *Rilevamento degli oggetti (Object Detection)*: identificazione di tutte le entità presenti all'interno dell'immagine.
- *Segmentazione dell'immagine (Image Segmentation)*: l'immagine viene suddivisa in sezioni per essere analizzata nel dettaglio.
- *Riconoscimento facciale (Face Recognition)*: è un task fondamentale per individuare dei volti di persone all'interno di un'immagine.
- *Riconoscimento di azioni (Action Recognition)*: identificazione delle entità all'interno dell'immagine e della loro relazione nello spazio e nel tempo per individuare azioni specifiche (ad esempio, un tennista che colpisce una palla).
- *Rilevamento delle relazioni visive (Visual Relationship Detection)*: comprensione e interpretazione delle relazioni tra le entità presenti nell'immagine.
- *Riconoscimento delle emozioni (Emotion Recognition)*: rilevamento dei sentimenti dei volti presenti in un'immagine.
- *Modifica dell'immagine (Image Editing)*: modifiche a un'immagine (ad esempio, oscuramento di dati sensibili).

Per effettuare tutti questi task e per comprendere un'immagine, i sistemi di Computer Vision sfruttano il Machine Learning, ovvero un sistema di apprendimento automatico basato sull'Intelligenza Artificiale (AI) che utilizza i dati a sua disposizione per comprendere e migliorare le performance. Il Machine Learning consente di addestrare i sistemi a capire il

contesto di una foto sulla base di un set di dati; in questo modo il sistema può capire cosa rappresentano tutti i numeri che sta analizzando.

Infatti, è fondamentale sapere che un computer non vede l'immagine così come la vediamo noi, ma osserva un set di numeri da elaborare e interpretare a seconda di come sono organizzati. La Computer Vision può allenarsi fino a diventare migliore e più abile della vista umana; infatti, quando un'immagine è ambigua possiamo continuare ad allenare il sistema con altri tipi di immagini finché il computer avrà abbastanza dati da riuscire ad interpretarla perfettamente.

## 4.2 Amazon Rekognition

Per testare le potenzialità della Computer Vision associate al riconoscimento delle immagini verrà adoperato il sistema "Amazon Rekognition", anch'esso facente parte del servizio AWS (Amazon Web Services).

### 4.2.1 Che cos'è Amazon Rekognition

Amazon Rekognition è un servizio di analisi delle immagini e dei video fornito da Amazon Web Services (AWS) che utilizza l'Intelligenza Artificiale per identificare, analizzare e riconoscere oggetti, persone, volti, testo e azioni all'interno di immagini e video digitali.

Amazon Rekognition è ampiamente utilizzato in una vasta gamma di settori e applicazioni, tra cui sicurezza, media e intrattenimento, pubblicità, monitoraggio e sicurezza pubblica, assistenza sanitaria, automazione industriale e altro ancora. Grazie alla sua facilità d'uso, alla sua scalabilità e alle potenti funzionalità di analisi visiva, è diventato uno strumento prezioso per le aziende e gli sviluppatori che desiderano integrare l'Intelligenza Artificiale nelle proprie applicazioni e nei propri servizi.

### 4.2.2 Funzionalità di Amazon Rekognition

Amazon Rekognition offre all'interno della sua struttura delle demo che permettono all'utente, tramite degli algoritmi pre-addestrati, di testare alcune funzionalità caricando foto o video personali e di ricevere un'analisi dettagliata in tempo reale di ciò che la rete neurale ha trovato all'interno dell'immagine, definendo dei parametri con le relative percentuali di probabilità o confidenza.

Le demo disponibili sulla piattaforma sono:

- Label detection;
- Image properties;
- Facial analysis;
- Face comparison;
- Face Liveness;
- Celebrity recognition;
- Text in image;
- PPE detection.

Inoltre, la piattaforma mette a disposizione degli utenti la possibilità di poter creare delle etichette personalizzate, definendo di conseguenza dei modelli addestrati, chiamati anche “Custom Labels”, da poter utilizzare creando opportuni dataset di immagini per l’addestramento (o “training”).

In seguito, tramite dei test, si può verificare l’attendibilità del modello addestrato, e al raggiungimento di un buon livello di addestramento, utilizzare tale modello.

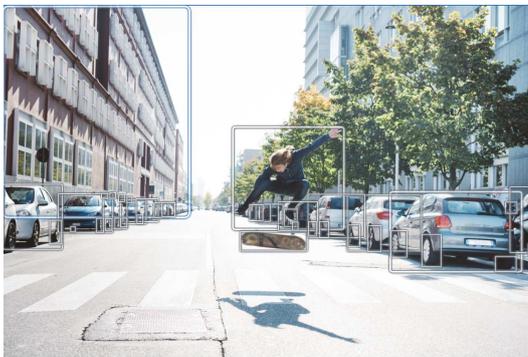
## 4.3 Test di riconoscimento di immagini tramite demo

Utilizzando alcune delle demo fornite dalla piattaforma è possibile analizzare in brevissimo tempo immagini di vario tipo sotto svariati punti di vista. In questa sezione adopereremo strumenti di identificazione generica, di proprietà delle immagini, di riconoscimento facciale, di confronto di visi, di riconoscimenti di celebrità, ed, infine, di decifrazione del testo all’interno di immagini.

### 4.3.1 Label detection

Il primo servizio demo offerto da Amazon Rekognition è Label Detection. Questo programma permette di inserire un’immagine qualsiasi; a seguito di ciò, la rete neurale sviluppata da Amazon sarà in grado di identificare in maniera autonoma oggetti e persone presenti e, perfino, il contesto della foto (Figura 4.1).

Ovviamente, la rete neurale non è in grado di riconoscere ogni elemento a priori, ma solo quelli per cui è stata addestrata; di fatto, AWS ci fornisce una barra di ricerca per controllare se l’oggetto interessato è presente nel database e, in più, ci permette di scaricare un file Excel con la lista completa di oltre 3000 label riconoscibili dal sistema.



#### ▼ Risultati

Neighborhood	99.9 %
City	99.9 %
Road	99.9 %
Street	99.9 %
Urban	99.9 %
Person	98.7 %

**Figura 4.1:** Screenshot del risultato di una Label detection su un’immagine di uno skater in città

Come possiamo notare nella Figura 4.1, sotto la voce "Risultati", verrà mostrato all’utente ciò che la rete neurale ha interpretato nel contesto della foto, con la relativa percentuale di certezza.

### 4.3.2 Image properties

Un altro servizio offerto dalla piattaforma di Amazon è Image Properties. Questo programma permette di caricare un’immagine qualsiasi (Figura 4.2) e di visionare, sempre tramite la sezione dei "Risultati", le corrispettive proprietà tra cui: colori dominanti, colori di sfondo, luminosità, nitidezza e contrasti, con le relative percentuali di affidabilità.



▼ Risultati

▼ Proprietà dell'immagine completa

Colori dominanti

#4682b4, RGB(70, 130, 180)	19.82%
#191970, RGB(25, 25, 112)	18.69%
#2f4f4f, RGB(47, 79, 79)	13.21%
#696969, RGB(105, 105, 105)	8.587%
#000000, RGB(0, 0, 0)	8.322%

Qualità dell'immagine

Brightness	63.41
Sharpness	78.59
Contrast	80.99

Figura 4.2: Screenshot del risultato dell'applicazione di Image properties sulla foto di una metropoli in notturna

### 4.3.3 Facial analysis

Questo servizio di Amazon Rekognition è in grado di riconoscere ed analizzare volti umani. Visionando la sezione "Risultati" dell'analisi (Figura 4.3) possiamo notare un gran qualitativo di attributi del volto proposto che il sistema è stato in grado di riconoscere come, ad esempio il sesso, l'età media, se sta sorridendo, se porta occhiali, specificando addirittura se si trattano di occhiali da sole, etc.



▼ Risultati



looks like a face	99.9 %
appears to be female	99.9 %
age range	26 - 34 years old
smiling	62.5 %
appears to be happy	92.1 %
wearing glasses	100 %
wearing sunglasses	100 %

Figura 4.3: Screenshot del risultato dell'applicazione di una Facial analysis di una ragazza in auto

### 4.3.4 Face comparison

Un altro servizio di Amazon Rekognition è Face Comparison, il quale permette di sottoporre al sistema due immagini, contenenti uno o più volti che quest'ultimo analizzerà e metterà a confronto. Esso, poi, fornirà una valutazione per ogni coppia di volti trovata nelle

due foto, specificando quanto ritiene che le due immagini rappresentino la stessa persona e fornendo anche una percentuale di somiglianza.

Inoltre, Face Comparison è in grado di riconoscere la stessa persona anche ad anni di distanza; infatti in Figura 4.4, possiamo notare, nella sezione "Risultati", una percentuale di similitudine del 99.9% per due soggetti che rappresentano la stessa persona ma in età differenti.

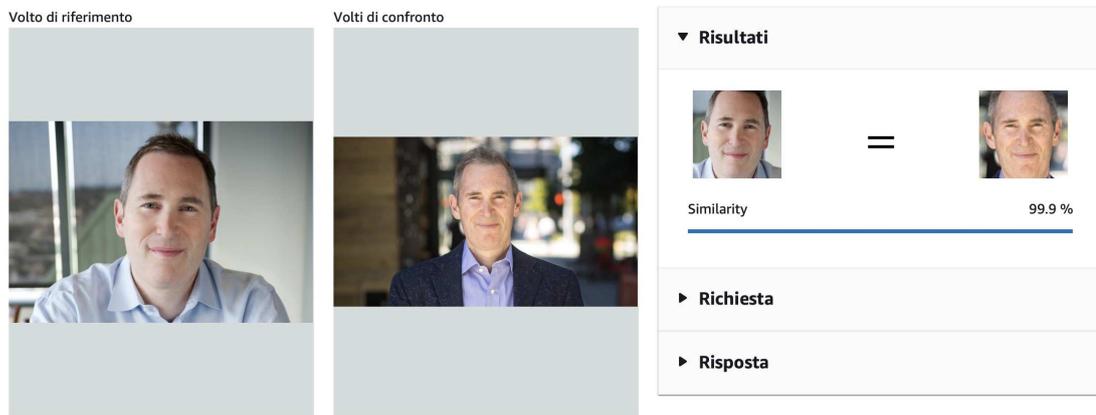


Figura 4.4: Screenshot del risultato di Face comparison

#### 4.3.5 Celebrity recognition

Tramite il servizio Celebrity recognition, Amazon offre agli utenti l'opportunità di sfruttare una rete neurale avanzata per identificare personaggi famosi in immagini caricate. Con questa potente funzionalità, gli utenti possono esplorare e testare le capacità di riconoscimento di Amazon, caricando immagini di celebrità e ottenendo risultati precisi e affidabili.

Grazie all'uso di algoritmi sofisticati di Intelligenza Artificiale, il servizio Celebrity recognition fornisce un'esperienza intuitiva e accurata, consentendo agli utenti di esplorare in modo rapido e efficace un vasto catalogo di personaggi noti (Figura 4.5).

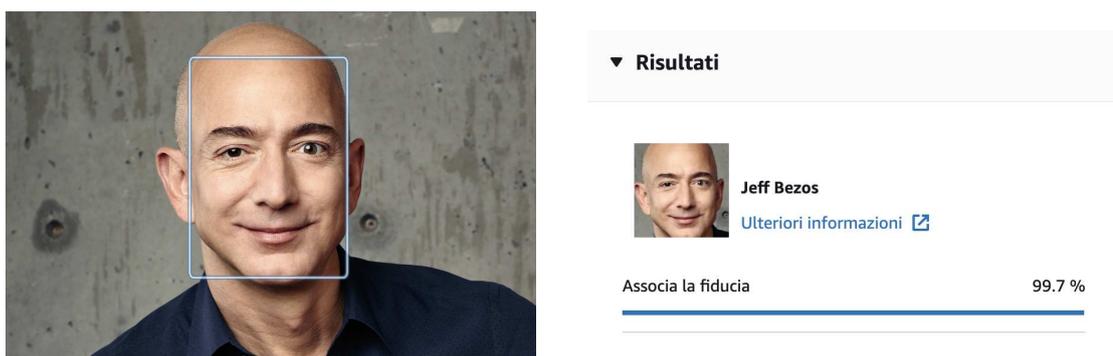


Figura 4.5: Screenshot del risultato di Celebrity recognition quando viene applicato su Jeff Bezos

#### 4.3.6 Text in image

Text in image permette di riconoscere un testo presente all'interno delle immagini. Una volta effettuato l'upload del file, ci troveremo davanti al resoconto dell'analisi effettuata da Amazon Rekognition (Figura 4.6). Nella sezione "Risultati" troveremo una trascrizione del testo analizzato dalla rete neurale.

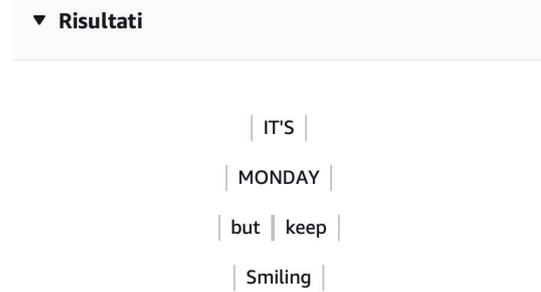


Figura 4.6: Screenshot del risultato di Text in image

## 4.4 Test di riconoscimento di immagini tramite Custom Labels

Come già accennato in precedenza, Amazon Rekognition mette a disposizione degli utilizzatori la possibilità di definire delle etichette personalizzate, e di conseguenza definire un proprio modello addestrato alle esigenze del fruitore del servizio.

Tramite l'addestramento della rete neurale, infatti, il modello definito permette di identificare oggetti non presenti tra le etichette fornite di default da AWS nella Label Detection. Ciò risulta estremamente importante in quanto fornisce la possibilità di implementare nuovi algoritmi di Computer Vision in grado di analizzare un sostanzioso quantitativo di situazioni ed oggetti.

### 4.4.1 Creazione del dataset

Il primo passo da compiere per la definizione di un modello personalizzato è quello della definizione di un dataset.

Un dataset (Figura 4.7) è un insieme di immagini, rappresentanti elementi comuni tra loro, utili ai fini dell'addestramento del modello. A tal fine esse vengono ripartite in due sezioni: immagini utili al training del modello ed immagini utili al test del modello. Generalmente si cerca di ripartire il dataset riservando l'80% delle immagini al training e il 20% al test, in modo tale da fornire alla rete neurale un buon quantitativo di immagini su cui addestrarsi.

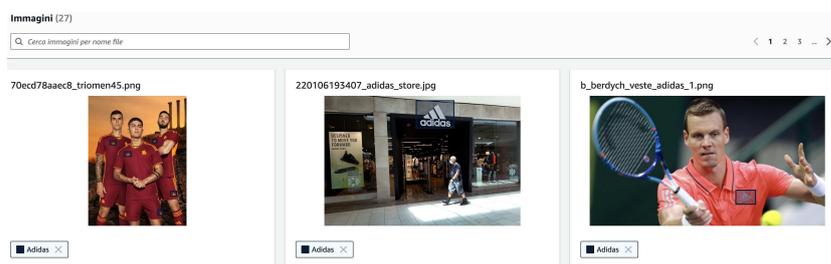


Figura 4.7: Screenshot di un dataset di training

### 4.4.2 Fase di etichettatura

Dopo aver definito il dataset, si passa alla fase di etichettatura, dove vengono definite prima di tutto le etichette che dovranno essere riconosciute all'interno delle immagini.

Successivamente si passa alla vera e propria etichettatura (Figura 4.8), dove l'utente scorre una per una le immagini del dataset di training definendo per ognuna di esse un riquadro con la relativa etichetta che definisce ciò che è contenuto al suo interno. Nel nostro caso, gli oggetti da riconoscere erano alcuni loghi di prodotti in ambito di moda e sport e le etichette erano, per l'appunto, "Nike" e "Adidas".

Al termine dell'etichettatura del dataset di training si passa all'etichettatura delle immagini del dataset di test, procedendo in modo analogo.



Figura 4.8: Screenshot di una etichettatura su logo Nike

### 4.4.3 Training del modello

Dopo aver etichettato correttamente tutte le immagini, il modello è pronto per essere addestrato. Una volta avviato l'addestramento (Figura 4.9), il sistema inizierà il processo di training che può durare dai 30 minuti fino a 24 ore in base alla quantità di etichette definite, nonché alla quantità e qualità delle immagini fornite nei dataset.

Dettagli del progetto							
Nome progetto LogoRekognition	Creato January 19, 2024 at 16:42:00 (UTC+01:00)	Set di dati 2 etichette di addestramento, 27 immagini di addestramento, 2 etichette di test, 20 immagini di test	Modelli 2				
Modelli (2)							
<input type="checkbox"/>	Nome	Data di creazione	Set di dati per la formazione	Test del set di dati	Prestazioni del modello (Punteggio F1)	Stato del modello	Messaggio di stato
<input type="checkbox"/>	LogoRekognition.2024-01-19T19.27.17	January 19, 2024			N/A	TRAINING_IN_PROGRESS	The model is being trained.
<input type="checkbox"/>	LogoRekognition.2024-01-19T18.30.36	January 19, 2024			0.832	TRAINING_COMPLETED	The model is ready to run.

Figura 4.9: Screenshot di training in corso del modello

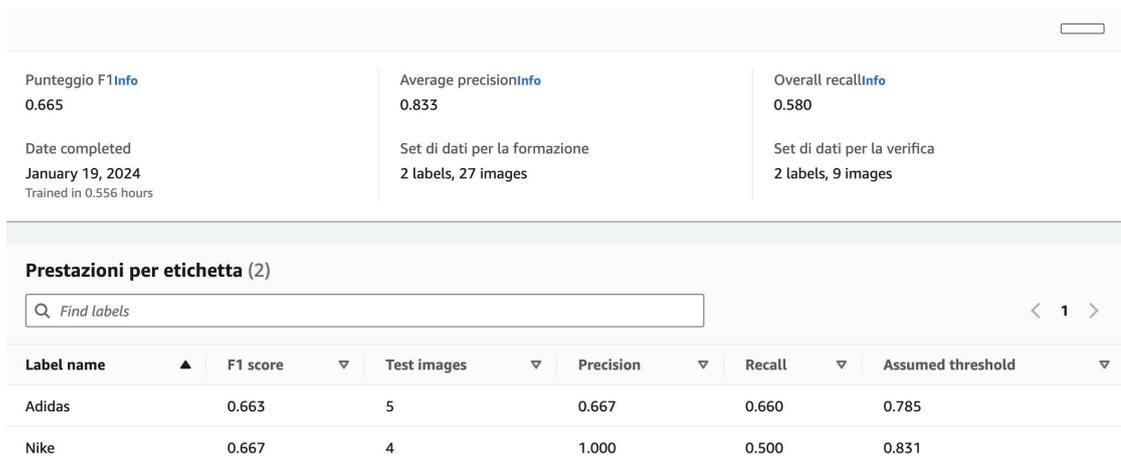
Appena il modello termina l'addestramento, saremo informati tramite lo stato del modello, in cui troveremo, come in figura sopra, "TRAINING\_COMPLETED"; a questo punto, possiamo verificare la valutazione del modello fornita da AWS, contenente le statistiche riguardanti il processo di addestramento del programma.

#### 4.4.4 Prestazioni del modello

Cliccando sul progetto ad allenamento terminato, verremo reindirizzati nella pagina di valutazione del nostro modello, dove è possibile visionare alcune statistiche (Figura 4.10), come, ad esempio:

- Il tempo impiegato dal sistema per completare l'operazione; in questo caso, 0.556 ore.
- La precisione media (Average Precision), ovvero la frazione di previsioni corrette (veri positivi) su tutte le previsioni del modello (veri e falsi positivi); in questo caso, 0.833;
- L'overall recall, ovvero la frazione di etichette del dataset di test che sono state predette correttamente; in questo caso, 0.580;
- L' F1 score, ovvero una misura aggregata che tiene conto sia della precisione che del recall su tutte le etichette, in questo caso: 0.665.

Inoltre, è possibile visionare le prestazioni per etichetta, dove, oltre a verificare i parametri citati in precedenza per ogni singola etichetta, si può visionare un parametro chiamato "soglia ipotizzata per etichetta", che rappresenta il valore al di sopra del quale una predizione viene considerata vera o falsa positiva.



The screenshot displays a dashboard with three main performance metrics: F1 score (0.665), Average precision (0.833), and Overall recall (0.580). Below these, it shows training details: 'Date completed January 19, 2024' and 'Trained in 0.556 hours'. The 'Set di dati per la formazione' section lists '2 labels, 27 images', and the 'Set di dati per la verifica' section lists '2 labels, 9 images'. A section titled 'Prestazioni per etichetta (2)' contains a search bar and a table with columns for Label name, F1 score, Test images, Precision, Recall, and Assumed threshold. The table lists 'Adidas' and 'Nike' with their respective metrics.

Label name	F1 score	Test images	Precision	Recall	Assumed threshold
Adidas	0.663	5	0.667	0.660	0.785
Nike	0.667	4	1.000	0.500	0.831

Figura 4.10: Screenshot delle prestazioni del modello

In seguito, è possibile visionare, immagine per immagine, i risultati ottenuti nel dataset di test della nostra rete neurale. Per ogni immagine nel set di dati di test è possibile visualizzare un'analisi della previsione del modello rispetto all'etichettatura assegnata. Esistono tre possibili risultati:

- *True positive*: indica che il modello ha identificato correttamente un'etichetta da noi assegnata nel set di dati di test.
- *False positive*: indica che il modello ha previsto un'etichetta non etichettata nell'immagine di test.
- *False negative*: indica che il modello non ha previsto un'etichetta contrassegnata come presente nell'immagine di test.

In aggiunta a questi, il sistema ci propone anche un punteggio di fiducia, ovvero un valore che quantifica la certezza della previsione del modello per ciascuna etichetta (un punteggio più alto significa maggiore fiducia).

In Figura 4.11 riportiamo due immagini del dataset di test con i risultati ottenuti dall'analisi delle etichette:



Figura 4.11: Screenshot del risultato dell'analisi di due immagini

#### 4.4.5 Variante del test di riconoscimento

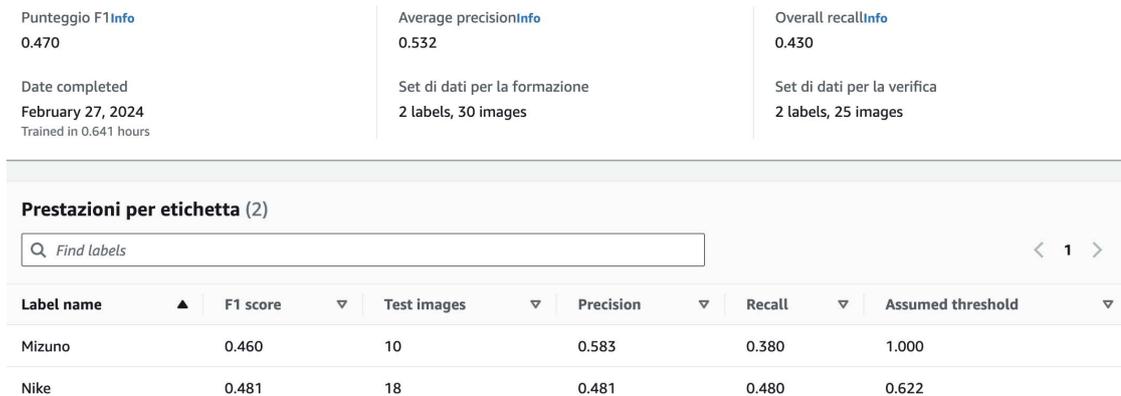
Successivamente al caso di studio precedentemente proposto, si è deciso di realizzare un ulteriore addestramento di un nuovo modello, sempre basandosi su Amazon Rekognition.

Il nuovo caso di studio comprende il riconoscimento del logo Nike con il riconoscimento del logo Mizuno; si tratta di due brand differenti, i cui loghi risultano avere una forma molto simile. Di fatto, è stato creato un nuovo modello (Figura 4.12), definendo delle immagini di training e di test che comprendessero i loghi Nike e Mizuno; sono state inserite anche foto che contenevano entrambi i loghi, per cercare di mettere in difficoltà la rete neurale durante l'addestramento.

Modelli (1)							
<input type="text" value="Trova risorse"/> <span style="float: right;">&lt; 1 ... &gt;</span>							
<input type="checkbox"/>	Nome	Data di creazione	Set di dati per la formazione	Test del set di dati	Prestazioni del modello (Punteggio F1)	Stato del modello	Messaggio di stato
<input type="checkbox"/>	Nike_Mizuno_Definitivo.2024-02-27T10.38.08	February 27, 2024			N/A	TRAINING_IN_PROGRESS	The model is being trained.

Figura 4.12: Screenshot di training in corso del modello variante

Una volta terminato il training del modello, è stato possibile visionare i risultati del suo addestramento, come mostrato in Figura 4.13.



**Figura 4.13:** Screenshot delle prestazioni del modello variante

In questa figura possiamo subito notare come i punteggi di F1, precisione media ed overall recall siano più bassi del caso precedente; in realtà, la riduzione non è neanche eccessiva in quanto in questo modello abbiamo utilizzato più immagini sia per il training che per il test.

Un discorso analogo vale per le prestazioni per etichetta singola, dove possiamo vedere risultati più bassi rispetto al precedente modello; di fatto, nella visione delle singole immagini, abbiamo che l'IA non sempre è stata in grado di riconoscere le etichette. Per quanto riguarda il tempo di addestramento, invece, siamo poco sopra il modello precedente, con 0,641 ore di allenamento.

Visionando immagine per immagine i risultati del dataset, invece, rimaniamo sorpresi di alcune etichette assegnate ad alcune foto, specie quelle in cui abbiamo nella stessa immagine sia il logo Nike che il logo Mizuno, dove la rete neurale è stata in grado di riconoscere perfettamente i due loghi senza confondersi. Spicca, inoltre, un'immagine, dove, per errore personale, non era stato etichettato un logo Nike durante la fase di etichettatura di test, e l'IA è stata in grado di riconoscere il logo, evidenziandolo come false positive.

In Figura 4.14 si riportano alcune foto del test con i risultati sopra menzionati.



**Figura 4.14:** Screenshot del risultato dell'analisi di due immagini

---

## Introduzione al Reinforcement Learning

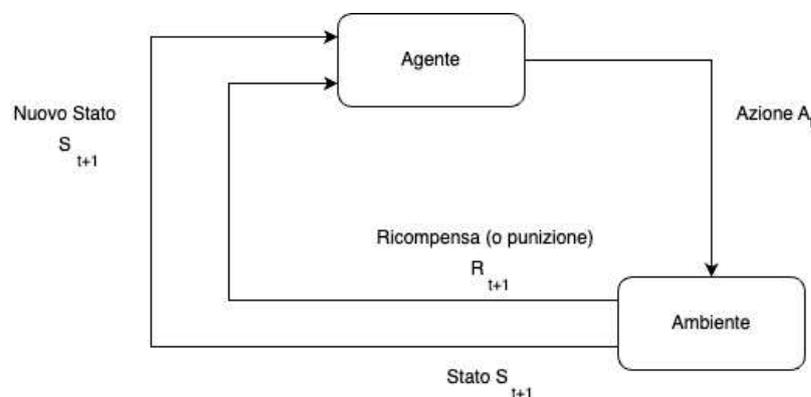
---

*In questo capitolo si introdurrà il concetto di Reinforcement Learning, analizzando che cos'è e come funziona. Successivamente verranno introdotte le fasi di addestramento del Reinforcement Learning, alcune delle sue applicazioni e come esso si sia sviluppato nei giochi.*

### 5.1 Che cos'è il Reinforcement Learning

Il Reinforcement Learning, denominato anche apprendimento per rinforzo, è una branca del Machine Learning che prevede l'allenamento di agenti "intelligenti" attraverso sistemi di ricompense e punizioni. Sostanzialmente, un computer (agente) impara a svolgere un'attività tramite ripetute interazioni di tipo "trial-and-error" (eseguite per tentativi ed errori) con un ambiente dinamico. Tale approccio all'apprendimento consente all'agente di adottare una serie di decisioni in grado di massimizzare una metrica di ricompensa per l'attività, senza essere esplicitamente programmato per tale operazione e senza l'intervento dell'uomo.

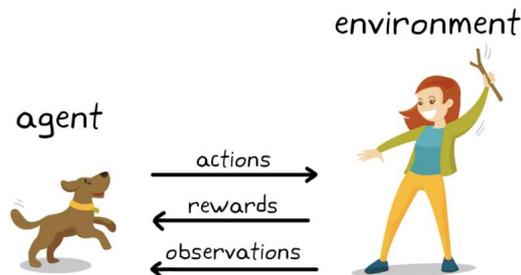
In altre parole, per allenare un modello di Reinforcement Learning si lascia al computer la libertà di procedere per tentativi, premiando i comportamenti corretti, ossia quelli che aumentano il punteggio, e punendo quelli scorretti. In questo modo il modello (Figura 5.1) imparerà quali azioni deve intraprendere per risolvere un problema sequenziale e quali, invece, è meglio evitare, in un percorso di apprendimento simile a quello di un essere umano.



**Figura 5.1:** Esempio di modello per il Reinforcement Learning

## 5.2 Come funziona il Reinforcement Learning

Il meccanismo di addestramento che si cela dietro al Reinforcement Learning riflette molti scenari del mondo reale. Consideriamo, per esempio, l'addestramento degli animali domestici tramite rinforzo positivo (Figura 5.2).



**Figura 5.2:** Il Reinforcement Learning nell'addestramento dei cani

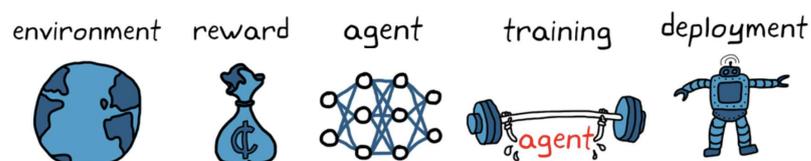
Usando la terminologia del Reinforcement Learning in Figura 5.2, l'obiettivo dell'apprendimento in questo caso è quello di addestrare un cane (agente) affinché completi un'attività all'interno di un ambiente che include l'ambiente che circonda il cane e la persona che si occupa dell'addestramento.

Prima di tutto, l'addestratore dà un comando o un segnale che il cane osserva (osservazione). Il cane, poi, risponde, con un'azione. Se questa si avvicina al comportamento desiderato, l'addestratore probabilmente offrirà una ricompensa, come un croccantino o un gioco; altrimenti, non verrà offerta alcuna ricompensa. All'inizio dell'addestramento il cane probabilmente risponderà con azioni più casuali, ad esempio rotolandosi al comando "seduto", in quanto sta cercando di associare delle osservazioni specifiche a delle azioni e a delle ricompense. Questa associazione, o mappatura, tra osservazioni e azioni si chiama politica. Dal punto di vista del cane, l'ideale sarebbe riuscire a rispondere correttamente a ogni segnale, così da ottenere il maggior numero di premi. Quindi, il fine ultimo dell'addestramento tramite Reinforcement Learning consiste nel "regolare" la politica del cane in modo che apprenda i comportamenti desiderati capaci di massimizzare la ricompensa.

Al termine dell'addestramento, il cane dovrebbe essere capace di osservare il padrone e di agire in modo appropriato, come sedersi al comando "seduto", usando la politica interna che ha sviluppato. A questo punto, i premi sono ben accetti ma, in teoria, non dovrebbero più essere necessari.

## 5.3 Fasi dell'addestramento del Reinforcement Learning

Il workflow generale per addestrare un agente usando il Reinforcement Learning include le seguenti fasi (Figura 5.3):



**Figura 5.3:** Workflow del Reinforcement Learning

► Creazione dell'ambiente:

Prima di tutto è necessario definire l'ambiente all'interno del quale opererà l'agente di Reinforcement Learning, compresa l'interfaccia tra agente e ambiente. L'ambiente può essere un modello di simulazione o un sistema fisico reale, ma, in genere, come primo step, si utilizzano gli ambienti simulati perché sono più sicuri e consentono di sperimentare.

► Definizione della ricompensa:

In seguito, è necessario specificare il segnale di ricompensa che l'agente utilizza per misurare le sue prestazioni rispetto agli obiettivi dell'attività e alla modalità di calcolo di questo segnale dall'ambiente. Dare forma alla ricompensa potrebbe essere complicato e, prima di raggiungere il risultato desiderato, potrebbero servire diverse iterazioni.

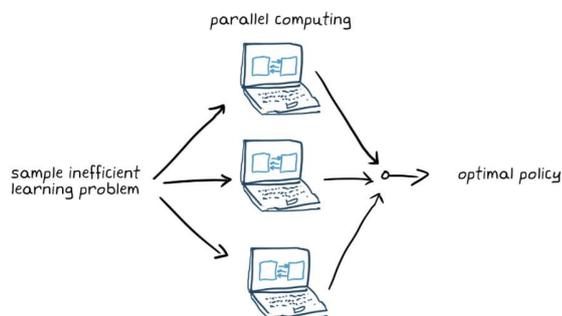
► Creazione dell'agente:

A questo punto è necessario creare l'agente, costituito dalla politica e dall'algoritmo di addestramento del Reinforcement Learning. In particolare, si deve:

- Scegliere un modo per rappresentare la politica, ad esempio usando reti neurali.
- Selezionare l'algoritmo di addestramento appropriato. Spesso diverse rappresentazioni sono legate a categorie specifiche di algoritmi di addestramento. In generale, però, gli algoritmi di Reinforcement Learning moderni si basano in gran parte su reti neurali, che si prestano bene a problemi complessi e a spazi di stato/azione di grandi dimensioni.

► Addestramento e convalida dell'agente:

Questa fase prevede la configurazione di opzioni di addestramento (come i criteri di arresto) e l'addestramento dell'agente per regolare la politica. Al termine dell'addestramento, non si deve dimenticare di convalidare la politica addestrata. Se necessario, è possibile rivedere le scelte di progettazione, come il segnale di ricompensa e l'architettura della politica, e procedere nuovamente all'addestramento. Il Reinforcement Learning è un processo generalmente noto per non sottoporsi al principio di sample efficiency; l'addestramento può richiedere da pochi minuti fino a giorni interi, e tutto dipende dall'applicazione. Per le applicazioni complesse, la parallelizzazione dell'addestramento su più CPU, GPU e cluster di computer renderà il processo più veloce (Figura 5.4).



**Figura 5.4:** Addestramento di un problema di apprendimento che non risponde al principio di sample efficiency con il calcolo parallelo

► Distribuzione della politica:

Questa fase prevede la distribuzione della rappresentazione della politica addestrata usando, ad esempio, codice CUDA o C/C++ generato. A questo punto, la politica è un sistema decisionale indipendente. L'addestramento di un agente tramite Reinforcement Learning è un processo iterativo. Le decisioni e i risultati delle fasi successive potrebbero richiedere il ritorno a una fase precedente del workflow di apprendimento. Ad esempio, se il processo di addestramento non converge su una politica accettabile entro un lasso di tempo ragionevole, potrebbe essere necessario aggiornare uno dei seguenti elementi prima di addestrare nuovamente l'agente:

- impostazioni di addestramento;
- configurazione dell'algoritmo di Reinforcement Learning;
- rappresentazione della politica;
- definizione del segnale di ricompensa;
- segnali di osservazione e azione;
- dinamica dell'ambiente.

## 5.4 Applicazioni del Reinforcement Learning

Il Reinforcement Learning risulta, quindi, essere una potente tecnica di apprendimento automatico, con numerose applicazioni in una vasta gamma di settori. Una delle sue applicazioni più evidenti riguarda il mondo dei giochi, che analizzeremo in seguito. Tuttavia, le sue potenzialità non si limitano ai giochi; di fatto, in campo robotico, il Reinforcement Learning viene impiegato per addestrare robot ad eseguire compiti complessi, come la manipolazione di oggetti o la navigazione in ambienti sconosciuti.

Nel settore industriale, i sistemi di controllo basati su Reinforcement Learning, come, ad esempio, l'Enterprise Resource Management (ERM), hanno come obiettivo quello di ottimizzare i processi di produzione e migliorare l'efficienza operativa, cercando di risparmiare tempo e conservare risorse.

Infine un'altra applicazione molto importante del Reinforcement Learning riguarda i veicoli a guida autonoma. Le self-driving car fanno uso di molteplici algoritmi per guidare, tra cui quelli riguardanti l'assistenza al parcheggio. In questo caso, l'apprendimento avviene attraverso dei sensori montati sull'auto per rilevare gli oggetti e per individuare il modo ottimale di parcheggiare. Tra le novità riguardanti il mondo delle auto a guida autonoma troviamo un algoritmo sviluppato dal Massachusetts Institute of Technology – MIT, ancora in fase di studio, riguardante la possibilità di ridurre le emissioni di CO<sub>2</sub> improntando le macchine ad una guida più ecologica. L'obiettivo di questi ricercatori consiste nell'arrivare a ridurre, attraverso una serie di simulazioni, il consumo di carburante e le emissioni, migliorando al contempo la velocità media del veicolo. Per testare questo algoritmo è stata impiegata una piattaforma di simulazione del traffico caratterizzata da un unico incrocio, dove una flotta di veicoli autonomi connessi, in grado di comunicare con i semafori, riceve informazioni su come accelerare e come decelerare. I risultati delle prime simulazioni, ricordando che lo studio è tuttora in corso, indicano che scenari con il 100% di veicoli autonomi connessi sono in grado di portare a una riduzione del 18% del consumo di carburante e del 25% dei livelli di emissioni di CO<sub>2</sub>, migliorando anche la velocità di marcia del 20%.

## 5.5 Reinforcement Learning nei giochi

Per quanto riguarda il Reinforcement Learning, come citato in precedenza, i videogiochi rappresentano un ottimo banco di prova per il training dei modelli. Ciascun gioco può

essere, infatti, visto come un ambiente dove ad ogni stato corrisponde un diverso insieme di scelte. Il modello di Reinforcement Learning dovrà, quindi, compiere delle azioni che vadano a massimizzare il suo punteggio nel gioco, valutando opportunamente se reiterare comportamenti passati che hanno condotto ad una ricompensa o se intraprendere una nuova via seppur dal risultato incerto.

Prima dell'ampio utilizzo del Reinforcement Learning, le Intelligenze Artificiali che dominavano i nostri videogiochi erano difficili da battere, non tanto per la loro capacità di adottare la strategia ottimale, piuttosto per i vantaggi che il sistema stesso concedeva loro. Ad esempio, in un gioco di carte digitale come "Heartstone", creato dalla celebre casa produttrice Blizzard, un avversario governato da un programma IA molto forte era difficile da battere perché aveva a disposizione carte migliori nel mazzo rispetto al giocatore, e non perché sapeva gestire al meglio le proprie risorse o le sapeva usare in modo particolarmente intelligente. In un gioco di guerra come "Call of Duty: Black Ops 3" (Figura 5.5), un soldato guidato dall'Intelligenza Artificiale con un grado di potenza elevato rispetto ad un soldato guidato dall'uomo con un grado di potenza normale eliminava più facilmente quest'ultimo, poiché, a parità di armi, quelle possedute dall'avversario infliggevano una maggiore quantità di danni rispetto a quelle possedute dal giocatore. Tale escamotage veniva chiamato "difficoltà artificiale". Di fatto, era un sistema ideato dagli sviluppatori per regolare la complessità di un livello di gioco o di un nemico basato sulla statistica, e non sull'intelligenza o sulle strategie migliori per affrontare il giocatore.



**Figura 5.5:** Screenshot di una partita a Call of Duty: Black Ops 3

A seguito di un utilizzo più assiduo del Reinforcement Learning nel campo dei videogiochi vi sono stati numerosi esperimenti che hanno previsto l'impiego di ambienti videoludici per il training degli agenti. Numerosi esempi recenti provengono da DeepMind, società all'avanguardia nella ricerca sull'IA, acquisita da Google nel 2014, che ha utilizzato videogiochi come Quake III Arena e StarCraft II per allenare le proprie istanze di Intelligenza Artificiale. In particolare, in un esperimento condotto nel 2019 all'interno del gioco Capture the Flag di Quake III Arena, è stato eseguito il training di una popolazione di agenti IA, ciascuno con il proprio percorso di apprendimento indipendente, in modo da testare la capacità di cooperazione dell'IA sia tra gli agenti stessi che con giocatori umani. I risultati dell'esperimento hanno mostrato la capacità, da parte degli agenti IA, non solo di avere delle performance nettamente migliori dei giocatori umani più abili, ma anche di dimostrare un miglior grado di cooperazione all'interno del gioco.

Un'altra società di ricerca, OpenAI, ha invece utilizzato il videogioco Dota 2 (Figura 5.6) per allenare i suoi modelli di Intelligenza Artificiale nel progetto OpenAI Five. L'obiettivo di lungo periodo è quello di aumentare sempre di più la capacità di generalizzazione dei modelli per realizzare quella che viene definita come Intelligenza Artificiale Generalizzata

(o AGI, Artificial General Intelligence), ossia un'intelligenza "completa", come quella degli esseri umani, in grado di interagire in un sistema complesso come quello del mondo reale. Gli agenti di OpenAI, allenati attraverso il Reinforcement Learning, sono stati in grado di battere sia una squadra di amatori che di semi-professionisti di Dota 2, mettendo, inoltre, in difficoltà un team di professionisti scelti tra i più forti giocatori del videogame.

Data l'elevata complessità del gioco (che si gioca a squadre in 5 contro 5), alcuni esperti di IA hanno parlato dell'esperimento come di un grande successo nello sviluppo dell'Intelligenza Artificiale, mentre altri hanno criticato il fatto che i bot siano stati fortemente privilegiati dall'accesso diretto alle API del gioco, quando avrebbero dovuto utilizzare sistemi di Computer Vision per interpretare i pixel dello schermo.



**Figura 5.6:** Screenshot di una partita a Dota 2

---

## Esperienze di Reinforcement Learning

---

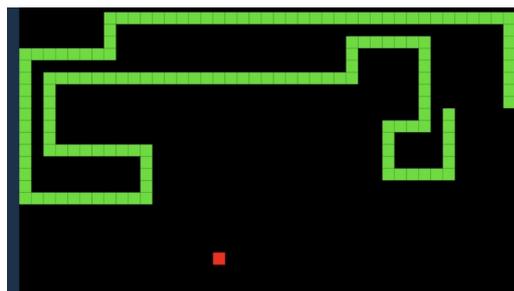
*In questo capitolo approfondiremo il concetto di Reinforcement Learning introdotto precedentemente, delineando, passo per passo, la creazione di un agente IA capace di apprendere autonomamente come giocare al celebre gioco Snake.*

### 6.1 Snake Game ed il Reinforcement Learning

Nel capitolo precedente abbiamo introdotto il Reinforcement Learning e constatato quanto esso abbia ottenuto successo nel mondo dei giochi, dove è stato utilizzato per creare agenti di Intelligenza Artificiale capaci di imparare autonomamente a giocare a giochi specifici. In questo capitolo, per valutare l'efficacia del Reinforcement Learning, esamineremo come definire un agente che, durante le attività di gioco relative ad un determinato videogioco, migliora gradualmente la propria strategia, ottenendo, di conseguenza, risultati migliori.

Il gioco preso in esame è Snake (Figura 6.1), considerato uno dei primi giochi per dispositivi mobili, le cui origini risalgono al videogioco arcade "Blockade" del 1976. La sua popolarità è cresciuta esponenzialmente con il rilascio del Nokia 6110, dove una versione monocromatica di Snake era preinstallata su questi dispositivi, finendo nelle mani di milioni di persone.

Il gioco è estremamente semplice: il giocatore controlla un serpente che può muoversi verso l'alto, in basso, a sinistra o a destra. Ogni volta che il serpente mangia una mela, cresce. Tuttavia, ci sono due regole fondamentali da seguire: evitare di colpire i bordi del campo di gioco e di mordere la propria coda. Qualsiasi collisione con i bordi o con la coda termina immediatamente il gioco. Il punteggio più alto è determinato dal numero di quadrati aggiunti al serpente. La partita è vinta quando non c'è più spazio per far crescere il serpente.



**Figura 6.1:** Screenshot di una partita a Snake

### 6.1.1 Librerie per lo sviluppo

Per lo sviluppo dell'agente IA sono state utilizzate due librerie di Python: Pygame e Pytorch.

► Pygame

è una libreria open-source utilizzata per lo sviluppo di videogiochi in Python. Essa fornisce strumenti e funzionalità per gestire grafica, suoni, input utente e altro ancora, semplificando lo sviluppo di giochi 2D. Con Pygame, è possibile creare giochi completi e interattivi, sia per scopi ludici che educativi.

► Pytorch

è una libreria open-source per l'apprendimento automatico e il Deep Learning. È stato sviluppato principalmente da Facebook's AI Research lab (FAIR). PyTorch offre una vasta gamma di strumenti per costruire e addestrare reti neurali, gestire dati, calcoli e altro ancora. È noto per la sua flessibilità e facilità d'uso, soprattutto grazie al suo approccio "imperativo" alla definizione dei modelli, che lo rende più intuitivo per molti sviluppatori. Inoltre, Pytorch è stato integrato per lo sviluppo di Deep Q Learning, che estende il Reinforcement Learning usando una Deep Neural Network per definire una previsione delle azioni da intraprendere. Vedremo in seguito che la Q di Deep Q Learning sarà, appunto, legata alle "quality of action" nella definizione del modello, in quanto il nostro obiettivo è, appunto, massimizzare la qualità delle azioni.

### 6.1.2 Definizione del sistema

Come primo passo è necessario definire il sistema che si configura come l'interazione tra due componenti principali, ovvero l'ambiente (il gioco stesso) e l'agente (il serpente). L'agente raccoglierà informazioni sul suo stato corrente (dettagli che verranno esaminati in seguito) e agirà di conseguenza, eseguendo azioni volte al progresso nel gioco. L'ambiente reagirà premiando o punendo l'agente in base alle azioni intraprese; l'obiettivo dell'agente sarà, quindi, massimizzare la ricompensa ottenuta. Inizialmente, senza regole prestabilite, il serpente compie azioni casuali poiché non ha conoscenza delle dinamiche di gioco. L'obiettivo consiste nell'elaborare una strategia, definita tecnicamente "policy", volta a massimizzare il punteggio o la ricompensa. È qui che entra in gioco il Deep Q Learning. In sostanza, il sistema può essere suddiviso in tre sezioni distinte:

- *il gioco*, definito da Pygame;
- *il model*, definito con Pytorch;
- *l'agente*, che utilizza il gioco ed il model per effettuare il training.

Nella definizione del gioco, è essenziale descrivere un insieme di stati che rappresentano la posizione del serpente sulla griglia di gioco. Questi stati possono essere sintetizzati come mostrato in Figura 6.2.

```

State (11 values)

[danger straight, danger right, danger left,

direction left, direction right,
direction up, direction down,

food left, food right,
food up, food down
]

```

**Figura 6.2:** Esempio di un set di stati che indicano la posizione del serpente

Successivamente, si delineano una serie di azioni che il serpente deve compiere, come illustrato nella Figura 6.3:

### Action

```

[1, 0, 0] -> straight
[0, 1, 0] -> right turn
[0, 0, 1] -> left turn

```

**Figura 6.3:** Esempio di un set di azioni che il serpente può compiere

Infine, vengono definite delle ricompense, come, ad esempio, un incremento di +10 quando il serpente mangia una mela e una penalizzazione di -10 quando il serpente morde la propria coda o colpisce il muro.

### 6.1.3 Definizione della Q-Table ed equazione di Bellman

Per comprendere come l'agente prende decisioni, è fondamentale conoscere il concetto di Q-Table. La Q-Table (Figura 6.4) è una matrice che associa lo stato dell'agente con le azioni possibili che può intraprendere. I valori all'interno di questa tabella rappresentano le probabilità di successo delle azioni, essenzialmente una stima della ricompensa attesa, e vengono aggiornati sulla base delle ricompense ottenute dall'agente durante il processo di addestramento.

State	Right	Left	Up	Down
1	0	0.31	0.12	0.87
2	0.98	-0.12	0.01	0.14
3	1	0.10	0.12	0.31
4	0.19	0.14	0.87	-0.12

**Figura 6.4:** Esempio di Q-Table rappresentante le possibili azioni e le relative probabilità di successo

In questo esempio, potremmo decidere di muoverci verso destra se ci troviamo nello stato 2, e verso l'alto se ci troviamo nello stato 4. I valori presenti nella tabella rappresentano la somma cumulativa delle ricompense attese per intraprendere un'azione "a" da uno stato "s". Questa tabella, come menzionato in precedenza, definisce la policy dell'agente, indicando

quali azioni dovrebbero essere eseguite da ciascuno stato per massimizzare la ricompensa attesa.

Tuttavia, c'è un problema: la policy è rappresentata da una tabella, il che significa che può gestire solo uno spazio di stati finito. In altre parole, non possiamo avere una tabella infinita con un numero infinito di stati.

Per affrontare questa limitazione, entra in gioco il Deep Q Learning, menzionato in precedenza, che converte la tabella in una Deep Neural Network seguendo l'equazione di Bellman (Figura 6.5):

$$NewQ(s, a) = Q(s, a) + \alpha [R(s, a) + \gamma \max Q'(s', a') - Q(s, a)]$$

**Figura 6.5:** Equazione di Bellman

Esistono anche versioni semplificate dell'equazione di Bellman, che sono mostrate nella Figura 6.6.

Q Update Rule Simplified:

$$Q = model.predict(state_0)$$

$$Q_{new} = R + \gamma \cdot \max(Q(state_1))$$

**Figura 6.6:** Equazioni di Bellman semplificate

#### 6.1.4 Sintesi dell'algoritmo

In generale, il funzionamento dell'algoritmo procede come segue:

- Il gioco ha inizio e il Q-value viene inizializzato in modo casuale.
- L'agente raccoglie lo stato corrente "s".
- L'agente esegue un'azione basata sullo stato raccolto. Tale azione può essere casuale oppure restituita dalla sua rete neurale. Durante le prime fasi di addestramento, il sistema tende a scegliere azioni casuali per esplorare le varie possibilità. In seguito, esso tende sempre più a fare affidamento sulla sua rete neurale.
- Dopo che l'IA ha scelto ed eseguito l'azione, l'ambiente fornisce una ricompensa all'agente. Successivamente, l'agente raggiunge un nuovo stato e aggiorna il suo Q-value seguendo l'equazione di Bellman precedentemente menzionata. Inoltre, per ogni mossa, vengono registrati lo stato iniziale, l'azione intrapresa, lo stato raggiunto dopo l'azione, la ricompensa ottenuta e se il gioco è terminato o meno. Questi dati vengono successivamente utilizzati per addestrare la rete neurale tramite una procedura chiamata Replay Memory.
- Queste ultime due operazioni vengono ripetute fino al verificarsi di determinate condizioni, come, ad esempio, la fine del gioco.

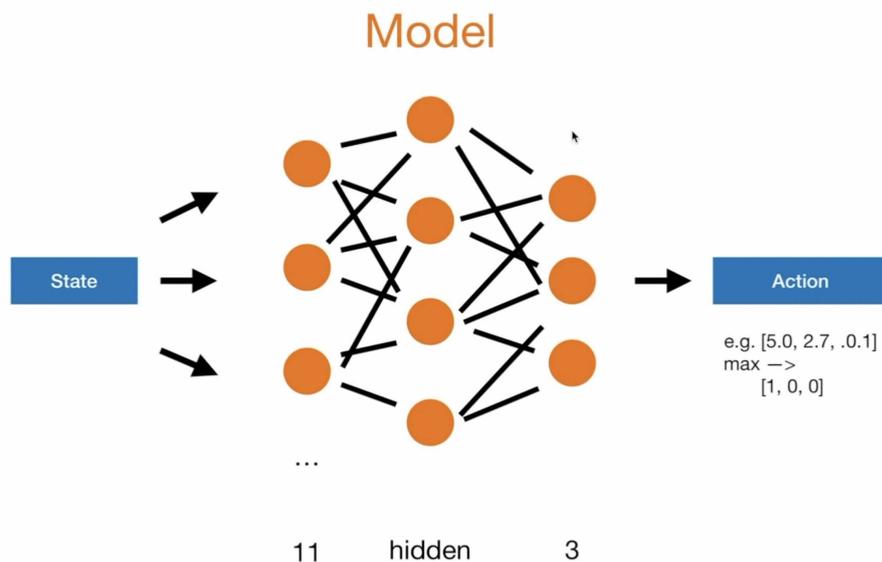
La rete neurale ottimizza l'output (cioè l'azione) in base a uno specifico input (lo stato), cercando di massimizzare la ricompensa attesa. Il valore che esprime quanto è buona la predizione viene confrontato con la verità fornita dalla funzione di perdita. Il compito della rete neurale è, quindi, quello di minimizzare tale perdita, riducendo la discrepanza tra il target reale e quello predetto. Nel nostro caso, la funzione di perdita è espressa nella Figura 6.7.

$$loss = (Q_{new} - Q)^2$$

Mean Squared Error

**Figura 6.7:** Funzione di perdita

In seguito a quanto detto, possiamo osservare come il modello elabori effettivamente le sue decisioni, basandosi sull'algoritmo precedentemente menzionato. Fondamentalmente, abbiamo uno State composto da un totale di 11 stati come input. Attraverso il Model e l'algoritmo, otterremo tre risultati, ciascuno associato alle azioni che l'agente deve intraprendere. L'obiettivo è selezionare il valore massimo tra i tre risultati ottenuti, il quale corrisponderà a una delle tre azioni, come mostrato nella Figura 6.8.



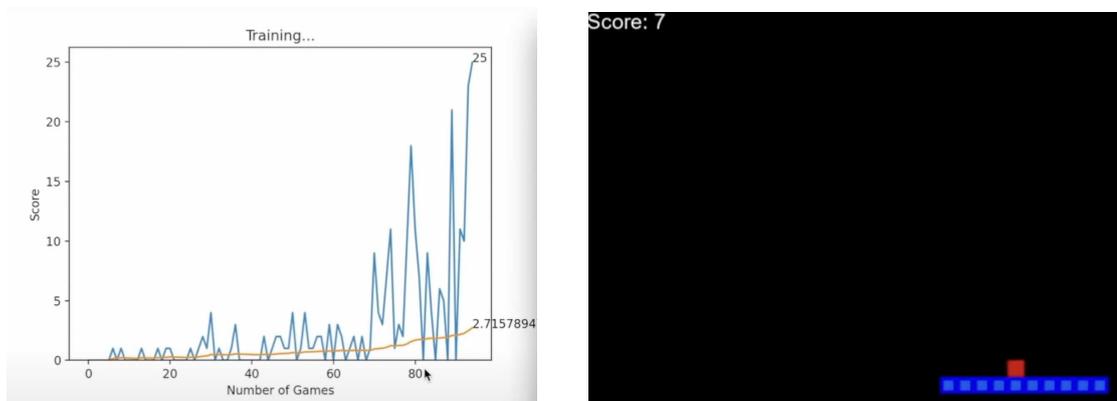
**Figura 6.8:** Schema di elaborazione delle azioni intraprese dal modello

### 6.1.5 Prestazioni dell'agente IA

Al termine di un addestramento della durata di circa 10 minuti, è possibile esaminare l'evoluzione della policy dell'agente mentre affronta le partite. Per comprendere appieno il processo di apprendimento e il funzionamento effettivo del Deep Reinforcement Learning, sono stati rappresentati graficamente il punteggio ottenuto e il numero di partite giocate (Figura 6.9).

Come evidenziato nella Figura 6.9, durante le prime 60 partite l'agente accumula pochi punti, poiché in questa fase prende decisioni casuali per esplorare il campo di gioco e memorizzare diverse situazioni, azioni e ricompense. Successivamente, esso smette di agire in modo casuale e comincia a decidere in base alla sua rete neurale, adottando, così, una policy specifica.

In soli 10 minuti, dopo aver disputato circa 90 partite, l'agente IA è riuscito a guadagnare circa 25 punti, mostrando un miglioramento significativo nella sua capacità di giocare.



**Figura 6.9:** Grafico punteggio - partite giocate con screenshot di una partita a Snake

*In questo capitolo discuteremo e trarremo le conclusioni dalle esperienze descritte nei capitoli precedenti in merito alla Sentiment Analysis, alla Computer Vision e al Reinforcement Learning.*

### 7.1 Discussione in merito alla Sentiment Analysis

La Sentiment Analysis è una tecnica di elaborazione del linguaggio naturale che ci consente di determinare, tramite l'analisi di un testo digitale, se il tono emotivo del messaggio è positivo, negativo o misto. Per valutare le potenzialità di questa branca dell'Intelligenza Artificiale abbiamo impiegato un servizio offerto da Amazon Web Services, noto come Amazon Comprehend.

Questa piattaforma offre la capacità di analizzare testi di varie lunghezze e in molteplici lingue, consentendo l'identificazione di una serie di caratteristiche sintattiche, inclusa la Sentiment Analysis.

Dall'esperienza condotta su Amazon Comprehend, che ci ha permesso di analizzare testi digitali provenienti da diverse recensioni tratte da TripAdvisor, è emerso come questo servizio interpreti in modo quasi naturale le sensazioni o emozioni espresse nei messaggi. Ciò dimostra che si tratta di una tecnologia all'avanguardia, necessaria per le aziende che desiderano migliorare i propri prodotti o servizi basandosi sui feedback reali forniti dalle recensioni dei clienti.

La Sentiment Analysis si rivela particolarmente utile per le aziende, in quanto consente ad esse di comprendere il sentiment dei clienti e di reagire di conseguenza. Attraverso l'analisi accurata dei feedback dei clienti, le aziende possono identificare aree di miglioramento nei loro prodotti o servizi, adattando le proprie strategie di business per soddisfare al meglio le esigenze e le aspettative dei consumatori.

### 7.2 Discussione in merito alla Computer Vision

La Computer Vision è una branca dell'Intelligenza Artificiale dedicata allo studio e alla programmazione di algoritmi che consentono ai computer di emulare i processi e le funzioni del sistema visivo umano, interpretando e rilevando informazioni significative da immagini digitali o video.

In correlazione alla Computer Vision, troviamo il riconoscimento delle immagini, una tecnologia straordinaria che abbiamo avuto l'opportunità di testare ed esaminare attraverso la

piattaforma Amazon Rekognition, fornita da Amazon Web Services. Le demo pre-addestrate fornite da tale piattaforma si sono dimostrate estremamente competenti nel riconoscere oggetti, volti, situazioni e contesti.

Sebbene ciascuna demo testata offra un notevole valore per conto proprio, la vera potenza di questa piattaforma, combinata con una rete neurale, risiede nelle Custom Labels. Fondamentalmente, se si dispone di un vasto dataset contenente una grande varietà di oggetti o situazioni da analizzare, il servizio Custom Labels, con la definizione di etichette personalizzate, può addestrare e definire un modello in grado di riconoscere specifici oggetti in un tempo incredibilmente breve.

Tuttavia, è importante considerare alcuni aspetti critici legati all'utilizzo dell'Intelligenza Artificiale e della Computer Vision. Le preoccupazioni riguardanti la privacy dei dati e l'etica nell'impiego di queste tecnologie sono sempre più rilevanti. L'ampio utilizzo del riconoscimento delle immagini potrebbe sollevare interrogativi su chi ha accesso alle informazioni raccolte, su come vengono utilizzate e se vengono rispettate le normative sulla privacy dei dati.

In definitiva, mentre l'Intelligenza Artificiale continua a progredire e ad essere adottata in vari settori, è fondamentale affrontare e risolvere tali questioni in modo da garantire un utilizzo responsabile ed etico di queste potenti tecnologie.

### **7.3 Discussione in merito al Reinforcement Learning**

Il Reinforcement Learning rappresenta una branca del Machine Learning focalizzata sull'addestramento di agenti "intelligenti" tramite sistemi di ricompense e punizioni. Fondamentalmente, un agente informatico apprende a compiere un'attività attraverso interazioni ripetute con un ambiente dinamico, che fornisce al suddetto agente ricompense o punizioni in base alle azioni intraprese.

Attualmente, il Reinforcement Learning trova numerose applicazioni soprattutto nei settori industriali e di produzione. Come abbiamo visto, esso è impiegato nella robotica, nell'ottimizzazione dei processi produttivi aziendali e nel campo delle auto a guida autonoma. In quest'ultimo contesto, abbiamo assistito all'utilizzo di questa tecnologia nell'assistenza al parcheggio e nella progettazione di algoritmi finalizzati a ridurre le emissioni di CO<sub>2</sub>, promuovendo una guida più ecologica.

Tuttavia, l'applicazione più evidente del Reinforcement Learning è nel mondo dei giochi, dove sono stati sviluppati agenti IA capaci di sconfiggere squadre di dilettanti e persino professionisti in specifici giochi. Inoltre, abbiamo assistito alla creazione di un agente IA in grado di giocare al celebre gioco Snake, attraverso la combinazione di stati, azioni e ricompense per definire un modello in grado di affrontare tale gioco.

Di fatto, è importante sottolineare che la creazione di un agente IA per giocare a Snake rappresenta solo l'inizio delle potenzialità del Reinforcement Learning. Attualmente, l'agente è limitato a quel particolare gioco; tuttavia, ciò dimostra il potenziale di questa tecnologia e la sua capacità di adattarsi a compiti specifici.

In conclusione, il Reinforcement Learning si sta rapidamente affermando come una tecnologia estremamente potente e in continua evoluzione. Il suo obiettivo è quello di progredire costantemente, aprendo nuove frontiere e affrontando sfide sempre più complesse. L'obiettivo ultimo è quello di sviluppare un'Intelligenza Artificiale generica in grado di adattarsi e rispondere a qualsiasi scenario o task venga ad esso presentato, rappresentando, così, una risorsa preziosa in numerosi settori e ambiti applicativi.

In questa tesi ci siamo concentrati sull'esplorazione dell'Intelligenza Artificiale (IA), esaminando vari aspetti chiave e applicazioni pratiche di questa tecnologia rivoluzionaria. Abbiamo iniziato con un'introduzione generale all'IA, ripercorrendo le sue origini e i fondamenti teorici, inclusa la distinzione tra IA simbolica e subsimbolica. Abbiamo approfondito alcuni test classici, come il test di Turing, il test della stanza cinese e il test di Winograd, sollevando anche un dibattito filosofico sulla capacità di pensare da parte delle macchine. Inoltre, abbiamo illustrato diverse applicazioni dell'IA in vari settori.

Successivamente, abbiamo introdotto il concetto di cloud computing, focalizzandoci su Amazon Web Services (AWS). Utilizzando AWS, abbiamo condotto esperimenti di Sentiment Analysis con Amazon Comprehend, dimostrando come questo servizio possa analizzare e interpretare il linguaggio naturale per rilevare l'umore e le emozioni. Inoltre, abbiamo esplorato il campo della Computer Vision attraverso Amazon Rekognition, effettuando test con demo pre-addestrate e creando modelli personalizzati per riconoscere i loghi di brand come Nike, Adidas e Mizuno tramite la funzionalità Custom Labels.

Abbiamo poi affrontato il Reinforcement Learning, spiegando il suo funzionamento, le fasi di addestramento e alcune delle sue applicazioni più rilevanti. Questo ci ha portato a sviluppare un agente IA capace di imparare a giocare autonomamente al celebre gioco Snake, ottenendo risultati promettenti.

La tesi si conclude con una riflessione sulle esperienze di utilizzo dell'Intelligenza Artificiale condotte durante il nostro studio. Abbiamo dimostrato come l'IA possa essere implementata efficacemente in vari contesti, evidenziando il suo potenziale di trasformazione e le opportunità future che offre. Questo lavoro rappresenta un contributo significativo alla comprensione e all'applicazione dell'IA, aprendo la strada a ulteriori ricerche e sviluppi in questo campo affascinante.

Difatti, in un contesto in cui giorno per giorno si scoprono nuove funzionalità, nuovi approcci, nuovi rischi e nuove opportunità legati a questo campo, sarà fondamentale saper maneggiare queste tecnologie con maestria ed accuratezza sempre maggiori. Guardando al futuro, potremmo assistere finalmente allo sviluppo della cosiddetta AGI (Artificial General Intelligence), ossia l'Intelligenza Artificiale Generale, capace di risolvere qualsiasi tipo di problema venga posto ad essa.

Numerose imprese e organizzazioni stanno investendo ingenti risorse nello sviluppo della AGI. Tra queste spiccano aziende tecnologiche come OpenAI, Google DeepMind e IBM, che sono all'avanguardia nella ricerca sull'Intelligenza Artificiale. L'obiettivo è creare sistemi in grado di comprendere, apprendere e applicare conoscenze in un'ampia gamma di domini, avvicinandosi così alle capacità cognitive umane.

I potenziali benefici della AGI sono enormi. Essa potrebbe rivoluzionare settori come la medicina, l'energia, i trasporti e l'educazione, risolvendo problemi complessi e migliorando la qualità della vita in tutto il mondo. Tuttavia, insieme ai benefici, emergono anche significativi rischi. La creazione di un'intelligenza capace di operare in modo autonomo e potenzialmente superiore all'intelligenza umana solleva, infatti, importanti questioni etiche e di sicurezza. Vi è il pericolo che tali tecnologie possano essere utilizzate per scopi malevoli o sfuggire al controllo umano, causando conseguenze impreviste.

Pertanto, lo sviluppo della AGI richiederà un approccio responsabile e collaborativo, con una rigorosa regolamentazione e un costante monitoraggio degli impatti etici e sociali. Sarà essenziale promuovere un dialogo globale coinvolgendo scienziati, governi e società civile per garantire che l'IA del futuro sia sicura, equa e benefica per tutta l'umanità.

- BODEN, M. (2019), *L'Intelligenza Artificiale*, Il Mulino.
- CHANDRA, V. e HAREENDRAN, A. (2020), *ARTIFICIAL INTELLIGENCE: PRINCIPLES AND APPLICATIONS*, PHI Learning.
- CONTUCCI, P. (2023), *Rivoluzione Intelligenza Artificiale. Sfide, rischi e opportunità*, edizioni Dedalo.
- FERNANDES, S. L. e SHARMA, T. K. (2022), *Artificial Intelligence in Industrial Applications*, Springer.
- LAMBERT, D. (2023), *Robotica e Intelligenza Artificiale*, Queriniana.
- MINSKY, M. e PAPERT, S. (1969), *Perceptrons: an introduction to computational geometry*.
- MITCHELL, M. (2022), *L'intelligenza Artificiale – una guida per esseri umani pensanti*, Einaudi.
- MUSIOLIK, T. H. e CHEOK, A. D. (2020), *Analyzing Future Applications of AI, Sensors, and Robotics in Society*, IGI Global.
- RANDELL, B. (1982), *The Origins of Digital Computers: Selected Papers, Third Edition*, Springer-Verlag Berlin Heidelberg.
- RUMELHART, D. e MCCLELLAND, J. (1986a), *Parallel distributed processing: Explorations in microcognition, Volume 1: Foundations*, Bradford Books.
- RUMELHART, D. e MCCLELLAND, J. (1986b), *Parallel distributed processing: Explorations in microcognition, Volume 2: Psychological and biological models*, Bradford Books.
- RUSSELL, S. J. (2010), *Artificial Intelligence a modern approach*, Pearson Education.
- SEARLE, J. R. (1980), «Minds, brains, and programs», *THE BEHAVIORAL AND BRAIN SCIENCES* 3,417-457.
- SOMALVICO, M. (1987), *Intelligenza Artificiale*, Scienza & Vita.
- TURING, A. (1950), «Computing Machinery and Intelligence», *Mind*.

## Siti Web Consultati

- **IBM , Strong AI** – <https://www.ibm.com/it-it/topics/strong-ai>
- **Wikipedia** – [https://it.wikipedia.org/wiki/Intelligenza\\_artificiale](https://it.wikipedia.org/wiki/Intelligenza_artificiale)
- **Red Hot Cyber** – [https://www.redhotcyber.com/post/storia-intelligenza-artificiale-dagli-anni-50-ai-nostri-giorni/?utm\\_content=cmp-true](https://www.redhotcyber.com/post/storia-intelligenza-artificiale-dagli-anni-50-ai-nostri-giorni/?utm_content=cmp-true)
- **Blog Osservatori** – [https://blog.osservatori.net/it\\_it/intelligenza-artificiale-funzionamento-applicazioni](https://blog.osservatori.net/it_it/intelligenza-artificiale-funzionamento-applicazioni)
- **Blog Osservatori** – [https://blog.osservatori.net/it\\_it/applicazioni-intelligenza-artificiale](https://blog.osservatori.net/it_it/applicazioni-intelligenza-artificiale)
- **Amazon Web Services** – <https://aws.amazon.com/it/what-is-aws/>
- **Softwave - Soltec** – <https://www.softwave-soltec.it/amazon-web-service-cose-e-come-funziona/>
- **Amazon Web Services - Sentiment Analysis** – <https://aws.amazon.com/it/what-is/sentiment-analysis/>
- **Temera - Computer Vision** – <https://temera.it/it/news/blog/computer-vision.html>
- **MathWorks - Reinforcement Learning** – <https://it.mathworks.com/discovery/reinforcement-learning.html>
- **Data Skills - Reinforcement Learning** – <https://www.dataskills.it/reinforcement-learning-intelligenza-artificiale-e-videogiochi/>
- **FiloSoft - We love Software** – <https://filosoft.it/2021/06/03/lutilizzo-del-reinforcement-learning-per-la-creazione-di-artificial-intelligence-nei-videogiochi/>
- **YouTube** – <https://www.youtube.com/watch?v=L8ypSXwyBds>
- **Towards Data Science** – <https://towardsdatascience.com/how-to-teach-an-ai-to-play-games-deep-reinforcement-learning-28f9b920440a>

---

## Ringraziamenti

---

Bene, finalmente siamo arrivati alla parte dei ringraziamenti (l'unica cosa che forse i miei amici leggeranno...). Spero di menzionare tutti, se così non fosse, ringrazio tutti coloro che mi vogliono e vorranno sempre bene.

In primis, mi sento di ringraziare vivamente il Professor Ursino, che sin da subito si è dimostrato disponibile nei miei confronti e mi ha sostenuto fino in fondo nell'elaborazione di questa tesi. La sua bontà e generosità sono state per me di grande aiuto, e ho sempre potuto contare sulla sua disponibilità per ogni mia perplessità e richiesta, anche la più banale. Il percorso formativo che ho seguito sotto la sua guida, arricchito dalle preziose referenze, dai libri consigliati e dal mio studio personale, rappresenta un tesoro che custodirò con cura e che sarà fondamentale nel mio futuro accademico e professionale.

Il ringraziamento più importante va alla mia famiglia (mamma, papà, Karin e Matteo), che mi ha sostenuto e spronato per affrontare tutte le difficoltà che ho incontrato durante il percorso, stimolo a non mollare mai.

Ringrazio i miei compagni di università: Andrea, detto Andres, il mio motivatore instancabile; il confronto giornaliero con lui era quasi un obbligo, cosa fare cosa non fare, lui c'era e c'è sempre, interista ma ci posso passare sopra. Poi Elia, un grande, sa sempre come tirarti su il morale ed una persona veramente genuina e sincera; l'ultima fila in fondo a destra era sempre la nostra, condivisa con Tommaso (finché ha resistito) che è stato il mio primo vero compagno di Università; le notti passate a progettare non le scorderò facilmente.

Ringrazio poi Thomas, "il mio più fidato", con il quale, come un fratello, ho condiviso tutto, diciamo che a momenti mi conosce più lui che la mia famiglia. Questo traguardo lo devo anche a lui, che mi ha supportato sempre, a modo suo, ma mi ha supportato!

Ringrazio poi Jeck, anche lui come un secondo fratello; ci conosciamo da una vita (letteralmente), le nostre chiacchierate e le birrette dopo cena mi hanno sicuramente aiutato nel mio percorso universitario, facendomi svagare e stare bene quando necessario.

Ringrazio poi tutti i miei amici: i ragazzi di Tisanina, le Salsicce, il gruppo Offiziel e Gasag, con cui ho condiviso di tutto. Grazie davvero; senza di voi, le chiacchierate, le serate, le cene e le vacanze passate assieme sarebbe stato tutto più vuoto.

Voglio ringraziare i miei nonni, che mi hanno anch'essi sempre incoraggiato e sostenuto nel momento del bisogno, soprattutto nonna Sveva, che nel mio percorso di studio mi ha sempre seguito con affetto, sin da piccolo.

Infine voglio ringraziare chi purtroppo non c'è più; questo traguardo è anche merito tuo. Sin dalle superiori sei stato quella marcia in più dentro di me che mi ha spinto sempre a dare il massimo, come un "Turbo". Grazie.