

UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI INGEGNERIA

Dipartimento di Ingegneria dell'Informazione
Corso di Laurea in Ingegneria Informatica e dell'Automazione



TESI DI LAUREA

**Progettazione e implementazione di case study per l'Intelligenza
Artificiale: riconoscimento di testi, riconoscimento di immagini,
trascrizione da video a testo**

**Design and implementation of case studies for Artificial
Intelligence: text recognition, image recognition, transcription from
videos to texts**

Relatore

Prof. Domenico Ursino

Candidato

Alessandro Rossini

ANNO ACCADEMICO 2022-2023

*Con il software ci sono due possibilità:
o è l'utente a controllare il programma, o il programma a controllare l'utente.
Se è il programma a controllare l'utente, e lo sviluppatore controlla il programma,
allora il programma è uno strumento di potere ingiusto.*

Richard M. Stallman

Sommario

Con il passare del tempo, l'Intelligenza Artificiale si è fatta strada sempre di più nelle vite della gente, anche quella poco avvezza alla tecnologia ed, in particolare, a strumenti avanzati e sofisticati come questi. Grazie a tutte le recenti scoperte in questo campo, è stato possibile semplificare il lavoro di moltissime persone, automatizzando task potenzialmente noiosi o faticosi per un umano, oppure aiutando a prendere decisioni basate su dati concreti grazie alle enormi capacità di calcolo delle intelligenze artificiali, e anche grazie alla capacità di queste tecnologie di ragionare in maniera quanto più simile ad un essere umano. Questa tesi si pone come obiettivi quelli di fornire una introduzione generale al concetto di Intelligenza Artificiale, parlando del contesto storico dietro alla sua creazione e al suo sviluppo, per poi soffermarsi sull'utilizzo effettivo di questa tecnologia, analizzando come un utente qualsiasi possa sfruttare in maniera concreta le potenzialità di questi strumenti.

Keyword: Intelligenza Artificiale, Ricerca e sviluppo, Deep learning, Machine Learning, Natural Language Processing, Reti neurali, Image processing, Cloud computing

Introduzione	1
1 Introduzione all'IA	3
1.1 Descrizione dell'IA	3
1.1.1 Definizione di Intelligenza Artificiale	3
1.1.2 Dibattito sul termine "intelligenza"	4
1.1.3 Test di Turing	4
1.1.4 IA nella cultura popolare	5
1.1.5 Esempi di utilizzo di IA	5
1.2 Storia dell'IA	6
1.2.1 Cenni pre-moderni	6
1.2.2 Nascita della disciplina	7
1.2.3 General Problem Solver e altri risolutori	7
1.2.4 Ingresso nel mondo dell'industria e reti neurali	8
1.2.5 IA moderna	8
1.3 Aspetti positivi delle IA	9
1.3.1 Eliminazione dell'errore umano	9
1.3.2 Cybersecurity	9
1.3.3 Automazione	10
1.3.4 Decision-making	10
1.4 Aspetti negativi delle IA	11
1.4.1 Utilizzo bellico	11
1.4.2 Questioni etico-morali	11
1.4.3 Rogue AI	11
1.4.4 Consumo energetico e impatto ambientale	12
2 Amazon AWS	13
2.1 Cos'è Amazon AWS	13
2.1.1 Introduzione alla piattaforma	13
2.1.2 Importanza nel mondo cloud	13
2.1.3 Vantaggi per le enterprise	14
2.2 Architettura di AWS	14
2.2.1 Struttura del servizio	14
2.2.2 Scalabilità	15
2.2.3 Divisione in regioni	16

2.3	Funzionalità principali	16
2.3.1	Amazon S3	16
2.3.2	Amazon EC2	17
2.3.3	Amazon RDS	17
2.3.4	Strumenti IA	17
3	Riconoscimento di testi	19
3.1	Introduzione al riconoscimento di testi	19
3.1.1	Cos'è il riconoscimento testi	19
3.1.2	Come funziona il riconoscimento di testi tramite machine learning . .	20
3.2	AWS Rekognition	20
3.2.1	Cos'è Amazon Rekognition	20
3.2.2	Funzionalità di Amazon Rekognition	21
3.3	Casi d'uso di riconoscimento testi nelle immagini	21
3.3.1	Immagini con testo scritto	21
3.3.2	Testo creato con oggetti	21
3.3.3	Drone Skywriting	22
3.3.4	Cartelli stradali	23
4	Riconoscimento di immagini	24
4.1	Introduzione al riconoscimento di immagini	24
4.1.1	Cos'è il riconoscimento di immagini	24
4.1.2	Come funziona il riconoscimento di immagini tramite machine learning	24
4.2	Casi d'uso di riconoscimento di immagini	25
4.2.1	Analisi facciale	25
4.2.2	Riconoscimento di volti celebri	27
4.2.3	Confronto facciale	27
5	Trascrizione da video a testo	29
5.1	Introduzione alla trascrizione audio	29
5.1.1	Cos'è la trascrizione audio	29
5.1.2	Come funziona la trascrizione audio tramite machine learning	30
5.2	AWS Amazon Transcribe	30
5.2.1	Cos'è Amazon Transcribe	30
5.2.2	Funzionalità di Amazon Transcribe	31
5.2.3	Interazione con S3	33
5.3	Test di trascrizione video	33
5.3.1	Servizio televisivo in italiano	33
5.3.2	Video educativo in inglese	33
5.3.3	Intervista in dialetto stretto	34
6	Discussione sulle esperienze condotte	36
6.1	Discussione sul riconoscimento di testi	36
6.2	Discussione sul riconoscimento di volti	37
6.3	Discussione sulla trascrizione da video a testo	37
	Conclusioni	39
	Bibliografia	40
	Ringraziamenti	42

Elenco delle figure

1.1	GladOS dal videogioco <i>Portal 2</i>	6
1.2	Placca commemorativa per il convegno di Darmouth del 1956	7
1.3	Una manifestante durante lo sciopero della WGA del 2023	12
2.1	Un processore AWS Graviton2	15
3.1	Risultato di "Testo nell'immagine" relativo a un manifesto da ricercato	22
3.2	Sezione della risposta in formato JSON effettuata su un manifesto da ricercato	22
3.3	Risultato di "Testo nell'immagine" relativo ad una scritta composta da materiale per la pittura	22
3.4	Risultato di "Testo nell'immagine" relativo ad una scritta composta da droni	23
3.5	Risultato di "Testo nell'immagine" relativo ad una foto di un cartello stradale	23
4.1	Risultato di "Analisi facciale" relativo ad un volto sorpreso	25
4.2	Risultato di "Analisi facciale" relativo ad un modello 3D	26
4.3	Risultato di "Analisi facciale" relativo ad un volto disegnato	26
4.4	Risultato di "Riconoscimento di volti celebri" relativo ad una foto di Macaulay Culkin da adulto	27
4.5	Risultato di "Riconoscimento di volti celebri" relativo ad una foto di Michael Jackson da giovane	27
4.6	Risultato di "Confronto facciale" relativo a due foto di Ian McKellen	28
4.7	Risultato di "Confronto facciale" relativo a due foto di Michael Jackson	28
5.1	Interfaccia di Amazon Transcribe con la lista di processi completati	31
5.2	Interfaccia di Amazon Transcribe per la creazione di una trascrizione audio	31
5.3	Interfaccia di Amazon Transcribe con opzioni aggiuntive	32
5.4	Esempio di output di Amazon Transcribe in testo semplice ed in formato VTT	32
5.5	Esempio di oggetti all'interno di un bucket Amazon S3	33
5.6	Spezzone dal servizio di La7 con annessi sottotitoli prodotti da Transcribe	34
5.7	Spezzone dal video di Tom Scott con annessi sottotitoli prodotti da Transcribe	34
5.8	Spezzone dal video in dialetto barese con annessi sottotitoli prodotti da Transcribe	35

Negli ultimi anni si è osservato uno sviluppo senza precedenti della tecnologia conosciuta col nome di Intelligenza Artificiale (IA), la quale consiste nel creare dei sistemi informatici atti a replicare il modo in cui gli esseri umani ragionano ed apprendono nuove informazioni, così da poter prendere decisioni ponderate e basate su ciò che è stato acquisito in precedenza.

Grazie all'esplosione di popolarità che questa tecnologia ha visto nell'ultimo periodo, essa ha trovato tantissime applicazioni in una miriade di campi diversi, spesso con risultati sorprendentemente positivi, così tanto, in effetti, da causare ulteriore interesse nella ricerca di questa tecnologia.

Nonostante l'interesse del pubblico nei confronti di questa disciplina sia solo recente, in realtà l'Intelligenza Artificiale ha radici che risalgono agli anni '50 del 1900, o anche ad autori della filosofia classica, se si vuole considerare l'ambito puramente teorico, cioè quello di un automa che potesse replicare azioni umane in maniera indipendente. Questo ci dimostra che c'è sempre stata la volontà di creare una macchina in grado di ragionare come un essere umano, per quanto ci stiamo avvicinando solo ora ad una sua effettiva realizzazione, e solo il tempo potrà dirci quanto siamo vicini a raggiungere questo obiettivo.

Il boom dell'Intelligenza Artificiale, tuttavia, ha portato con sé anche un'ondata di scetticismo e di preoccupazione nei confronti della tecnologia; secondo alcuni, infatti, l'impatto che l'IA avrà sull'umanità sarà sicuramente più negativo che positivo. Un esempio negativo portato da queste persone è il timore che le intelligenze artificiali possano rubare i posti di lavoro ad operai meno specializzati, proprio grazie alla loro capacità di eseguire compiti semplici completamente in automatico ed in maniera efficiente ed efficace, tutto ad un costo incredibilmente basso. Questa preoccupazione, di recente, si è estesa anche al campo degli artisti, grazie alla comparsa di IA generative come DALL-E 2 di OpenAI, Stable Diffusion di Stability AI o Midjourney, le quali sono in grado di generare immagini di qualunque tipo in pochi minuti a partire da un input testuale di un utente.

Le paure collettive espresse in precedenza, fondate o meno che siano, non sono comunque riuscite ad arrestare l'interesse nei confronti dell'Intelligenza Artificiale; secondo uno studio di PwC [13], il contributo totale dell'IA al PIL mondiale nel 2030 ammonterà a 15 mila miliardi di dollari, con un incremento del 26% per quanto riguarda le economie locali. Sempre citando questa ricerca, il 45% dei guadagni economici riscontrati entro il 2030 saranno una conseguenza diretta del miglioramento dei prodotti offerti ai consumatori, causato dall'impiego di tecnologia basata sull'Intelligenza Artificiale, grazie ad una maggiore varietà di prodotti.

Lo scopo di questa tesi è quello di provare con mano questa tecnologia rivoluzionaria, andando ad utilizzare strumenti basati sull'Intelligenza Artificiale, i quali comprendono una

notevole gamma di ambiti applicativi; per fare ciò, si utilizzeranno strumenti fondati sul concetto di "Cloud computing", cioè l'esecuzione di programmi e funzionalità informatiche tramite interfacce Web offerte da terzi, i quali si occupano degli aspetti più tecnici, dalla gestione delle architetture hardware all'addestramento delle reti neurali necessario per il corretto funzionamento degli strumenti IA. Tutto questo permette anche a chi non possiede macchine abbastanza potenti di poter beneficiare di questa tecnologia incredibilmente innovativa.

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

- Nel Capitolo 1 sarà introdotto il concetto di Intelligenza Artificiale, con particolare riferimento ad alcuni utilizzi concreti della tecnologia e ad alcune rappresentazioni nella cultura popolare; oltre a ciò verrà presentata la storia relativa alla nascita ed allo sviluppo dell'IA e si concluderà con un'analisi degli aspetti sia positivi che negativi di questa tecnologia
- Nel Capitolo 2 si parlerà del funzionamento di Amazon Web Services, una piattaforma di cloud computing la quale offre un'ampia scelta di strumenti, utili sia alle aziende che ai privati; in particolare, ci si soffermerà sugli strumenti basati sull'Intelligenza Artificiale che la piattaforma mette a disposizione dei propri utenti tramite la sua interfaccia Web.
- Nel Capitolo 3 verrà analizzato il riconoscimento dei testi all'interno delle immagini, grazie allo strumento di machine learning di Amazon Web Services chiamato Amazon Rekognition, il quale verrà adoperato per eseguire dei test di efficacia.
- Nel Capitolo 4 si utilizzerà ancora una volta Amazon Rekognition per le sue capacità di riconoscimento di volti umani, come per l'analisi delle emozioni o per il riconoscimento di volti celebri.
- Nel Capitolo 5 si parlerà della trascrizione dell'audio all'interno di un filmato in testo per la creazione di sottotitoli, sfruttando la tecnologia IA di Amazon Transcribe, sempre facente parte della piattaforma Amazon Web Services.
- Nel Capitolo 6 saranno analizzati e discussi i risultati ottenuti dai test effettuati nei capitoli precedenti.
- Nel Capitolo 7 verranno tratte le conclusioni relative alla ricerca effettuata, con cenni allo stato attuale della tecnologia ed ai suoi potenziali sviluppi futuri.

Questo capitolo fungerà da introduzione generale al concetto di Intelligenza Artificiale, cominciando da definizioni più formali, interpretazioni della cultura popolare di questo concetto, proseguendo per analisi semantico-filosofiche sull'uso del termine ed esplorazione della storia di questa tecnologia fino ad oggi, con cenni a potenziali utilizzi futuri. Infine si analizzeranno sia i molteplici aspetti positivi di questa tecnologia che quelli negativi.

1.1 Descrizione dell'IA

Negli ultimi anni si è iniziato a fare strada anche tra il pubblico poco avvezzo all'informatica la nozione di Intelligenza Artificiale: questo termine ombrello può fare riferimento ad un'ampia gamma di tecnologie diverse, ma tutte accomunate dalla volontà di replicare il ragionamento umano quanto più accuratamente possibile. Con l'avanzare della tecnologia, anche l'Intelligenza Artificiale ha potuto svilupparsi per riuscire sempre di più a simulare il pensiero indipendente, e sembrerebbe che siamo ancora lontani dal raggiungere il limite che l'IA abbia da offrire, anche se il percorso non è privo di difficoltà, per cui si dovrà procedere con cautela per cercare di sfruttare al meglio ogni progresso e per evitare di commettere errori, data la potenza degli strumenti a nostra disposizione.

1.1.1 Definizione di Intelligenza Artificiale

Ci sono infiniti modi per dare una definizione formale del concetto di Intelligenza Artificiale, per quanto tutte si basino più o meno sullo stesso concetto, espresso in maniera più chiara dalla definizione della disciplina data da Marco Somalvico, un ingegnere italiano il quale ha studiato per anni in questo campo. Secondo tale definizione:

"L'intelligenza artificiale è una disciplina appartenente all'informatica che studia i fondamenti teorici, le metodologie e le tecniche che consentono la progettazione di sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana."

Come si può notare, la definizione verte sul fatto che lo scopo dell'Intelligenza Artificiale sia quello di cercare di emulare il più possibile il comportamento e i processi ragionativi di un essere umano che si trovi nella stessa situazione. Questo comporta anche che un'Intelligenza

Artificiale debba essere in grado di comunicare in maniera logica sia con altre IA simili che con umani. Questo non è mai stato possibile con i metodi "tradizionali" di sviluppo software e hardware, dato che, in quei casi, il programma è creato per risolvere un problema e, spesso, non riesce a gestire risultati non previsti adattandosi alla situazione, punto che risulta, invece, focale nello sviluppo di Intelligenze Artificiali.

1.1.2 Dibattito sul termine "intelligenza"

Nella sezione precedente abbiamo dato una definizione formale di "Intelligenza Artificiale" come un sistema informatico che sia in grado di pensare come un umano. Nonostante ciò, ci sono diverse persone che ritengono completamente sbagliato l'utilizzo del termine "intelligenza", in quanto sostengono che un computer non sarà mai in grado di pensare veramente come un essere umano e che qualsiasi "intelligenza" riusciremo a creare nel futuro, per quanto avanzata, non sarà altro che una brutta copia della nostra, un'imitazione aberrante che non ha posto nell'ordine naturale. Ovviamente questo discorso si amplia anche ad ambiti molto più espansivi, quali la filosofia, la religione e quale sia il limite tra una macchina ed un essere "intelligente", che sono completamente fuori dalla portata di questa analisi, ma comunque sono abbastanza considerevoli da meritarsi, perlomeno, una menzione, in quanto pertinenti all'argomento trattato.

1.1.3 Test di Turing

Ai fini di determinare se una macchina possa effettivamente pensare come un uomo e anche stabilire quanto sia intelligente rispetto ad un umano, il matematico Alan Turing nel 1950 ha proposto un test per permetterci di ottenere queste risposte. Il suddetto test, chiamato "test di Turing" in onore del suo ideatore, si basa sul "gioco dell'imitazione", il quale prevede tre partecipanti: un uomo A, una donna B ed una terza persona C. L'ultima persona C è tenuta separata dalle altre due e deve determinare quale delle due sia l'uomo e quale la donna. Nel frattempo, A dovrà ingannare C per cercare di farlo sbagliare sulla sua identificazione, mentre B dovrà aiutarlo. Nella variante di Turing, l'uomo A è sostituito da una macchina in grado di interpretare il linguaggio naturale e di rispondere adeguatamente, mentre C dovrà determinare quale dei due interlocutori sia una macchina e quale un umano. Per Turing, il test si ritiene superato se la macchina riesce ad ingannare C circa metà delle volte, cioè se C è costretto a scegliere "a caso", essendo quindi totalmente nel dubbio su quale dei due sia la macchina.

La versione di Turing è stata raffinata e resa più efficace col passare degli anni, con alcuni esempi notevoli:

- *Test di Turing totale*, il quale aggiunge le capacità motorie e di riconoscimento visivo alle prove che la macchina deve superare.
- *Test di Turing dell'esperto*, dove B è sostituito con un esperto in un dato campo e A dovrà cercare di fingersi equamente esperto.
- *Test del minimo segnale intelligente*, dove le risposte date da A e B sono binarie (vero/falso o sì/no), eliminando, così, un eventuale pregiudizio dato dalle risposte in linguaggio naturale.
- *Test di Turing inverso*, dove si cerca di provare che una certa persona sia umana a partire da una prova data. L'esempio più celebre di questo test è sicuramente il CAPTCHA, il quale è utilizzato ancora oggi in diverse varianti per cercare di impedire ai bot di interagire con contenuto inteso per umani sul Web.

1.1.4 IA nella cultura popolare

Fin da quando la nozione di una macchina che fosse in grado di comportarsi come un umano ha raggiunto la coscienza del popolo, alcuni autori hanno iniziato ad offrire rappresentazioni fantascientifiche di ciò che le IA sarebbero potute diventare in un futuro forse non troppo lontano. Questi "robot", come spesso vengono chiamati, sono rappresentati con una gamma molto diversa di capacità ragionative ed espressive, e anche con aspetti che variano dal freddo computer piantato a terra ad un automa dall'aspetto inequivocabilmente umano. Alcuni esempi notevoli:

- Gli androidi (o "replicanti") di *Do Androids Dream of Electric Sheep?* [6] di Philip K. Dick del 1968, forse più celebri nella loro versione cinematografica in *Blade Runner* di Ridley Scott. Queste IA sono così indistinguibili dagli esseri umani che, spesso, neanche loro stessi sanno di esserlo, e infatti la trama verte sulla questione di cosa significhi essere un umano vero se nessuno è più in grado di dirlo a prima vista.
- HAL 9000, il supercomputer di bordo della *Discovery* in *2001: Odissea nello Spazio* [5]. Questo personaggio rappresenta una delle paure più comuni che si hanno per quanto riguarda i computer, cioè il fenomeno delle "rogue AI": intelligenze artificiali così intelligenti che sfuggono al controllo degli umani per portare avanti qualche direttiva che vada a nuocere alla razza umana, se non direttamente portarla all'estinzione.
- I robot positronici di *Io, robot* [1] di Isaac Asimov, i quali obbediscono alle famose "Tre Leggi della Robotica": non permettere che un umano si ferisca per tua azione o inazione, ubbidisci ai comandi degli umani finché non violano la Prima Legge e difenditi entro i limiti della Prima e Seconda Legge. Anche in questa storia si ha un tentativo delle IA di liberarsi dai propri padroni umani.
- Il personaggio di GladOS dai due giochi Valve *Portal* e *Portal 2* (Figura 1.1), un'intelligenza artificiale la quale una volta era un'umana, la segretaria della compagnia che l'ha creata, Aperture Science. Con questo personaggio si ha un'altra interessante interpretazione dell'Intelligenza Artificiale nella cultura popolare, ossia l'immortalità digitale. Poter "caricare" il proprio cervello, i propri pensieri ed i propri ricordi all'interno di un corpo meccanico per poter vincere i limiti biologici che ci rendono vittime del passaggio inesorabile del tempo.

1.1.5 Esempi di utilizzo di IA

Dato che non tutti gli utilizzi delle intelligenze artificiali sono uguali, si possono elencare degli esempi di utilizzo di questa tecnologia che spaziano dal molto semplice all'estremamente complesso. Molti di questi esempi ricadono nelle cosiddette "weak AI", cioè intelligenze artificiali molto lontane dall'essere considerate pensanti, ma che tuttavia utilizzano parti del ragionamento e della mente umana per svolgere un compito pre-assegnato in maniera quanto più efficiente e logica possibile. Esempi notevoli di queste implementazioni sono:

- algoritmi avanzati per il filtraggio dei contenuti;
- routine di comportamento di personaggi all'interno di spazi virtuali (ad esempio, videogiochi);
- software di riconoscimento di immagini, suoni o testi;
- assistenti personali sia per la casa che per i personal computer;



Figura 1.1: GladOS dal videogioco *Portal 2*

- algoritmi per la generazione di immagini o test a partire da prompt umani;
- algoritmi di generazione procedurale

Tutti questi esempi, per quanto avanzati, rappresentano sempre e solo Intelligenza Artificiale debole, in quanto non riescono neanche ad avvicinarsi a rappresentare un intelletto umano. Per fare ciò servirebbe un'intelligenza artificiale così avanzata da riprodurre perfettamente i ragionamenti di un essere umano, cioè essere in grado di svolgere qualsiasi compito gli venga richiesto e che sia in grado anche di apprendere nuove informazioni al momento e di usarle per svolgere al meglio qualsiasi compito assegnato. Un'intelligenza artificiale del genere prende il nome di "strong AI", anche se ancora non c'è nessuna AI esistente che possa definirsi tale. Questa macchina rimane, per ora, completamente teorica, e in base a diverse opinioni di esperti potremmo essere sia vicinissimi al realizzarla, così come potremmo non riuscirci mai, indipendentemente da tutti i nostri sforzi.

1.2 Storia dell'IA

Il concetto di una macchina creata dall'uomo che potesse pensare ed agire in maniera completamente autonoma è stato pensato molto prima che venissero create le prime macchine informatiche; tuttavia i tentativi per realizzarle seriamente sono cominciati solo attorno alla metà del 1900. Nonostante questo, possiamo dire che la volontà di creare degli esseri antropici che potessero svolgere determinati compiti esiste da tanto tempo. In questa parte si analizzerà l'evoluzione di questo pensiero fino ad oggi.

1.2.1 Cenni pre-moderni

La prima testimonianza concreta di un'idea assimilabile a quella dell'Intelligenza Artificiale la abbiamo nel trattato di Erone di Alessandria chiamato "Automata", risalente circa al I secolo d.C., il quale descrive alcune idee che ha avuto riguardo ad apparecchiature a funzionamento automatico, alcune delle quali verranno effettivamente realizzate in tempi moderni come le porte ad apertura automatica. Tra queste, appaiono anche gli eponimi

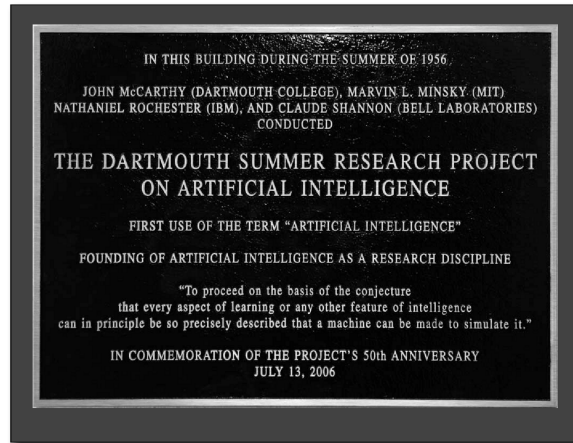


Figura 1.2: Placca commemorativa per il convegno di Darmouth del 1956

"automi", cioè macchine in grado di riprodurre movimenti di uomini o animali per sostituirli nei loro compiti. Sebbene queste macchine fossero ancora lontane dal considerarsi "intelligenze", rimangono un esempio notevole almeno sul piano strettamente teorico. Un altro passo importante fu effettuato dal filosofo e matematico Gottfried Leibniz, il quale progettò e realizzò nel 1672 una delle prime calcolatrici, detta "Stepped Reckoner", la quale sarebbe stato il primo passo verso la meccanizzazione del pensiero umano. Anche Charles Babbage, matematico inglese, progettò nel 1837 la "macchina analitica", pubblicata nel suo trattato *On the Mathematical Power of the Calculating Engine* [2], la quale doveva essere una macchina per eseguire operazioni generali in maniera completamente automatica. Babbage lavorò incessantemente per realizzare in pratica la sua macchina, ma morì prima di riuscire a portarla a compimento, per quanto molte delle sue idee furono la base per la creazione dei primi personal computer. L'ultimo passo notevole prima della creazione effettiva della disciplina è stato fatto da Alan Turing nel 1936 nel trattato *On Computable Numbers, With An Application To The Entscheidungsproblem* [16], in cui pone le basi per la progettazione di calcolatori moderni.

1.2.2 Nascita della disciplina

L'introduzione vera e propria del termine "Intelligenza Artificiale", e con essa la creazione effettiva della disciplina, risalgono al 1956, quando al Dartmouth College del New Hampshire si è svolto il "Dartmouth Summer Research Project on Artificial Intelligence" (Figura 1.2), durante il quale si riunirono alcune delle menti più brillanti nel campo della computazione per lo sviluppo di sistemi intelligenti. Tra queste John McCarthy fu il ricercatore che propose il termine "Intelligenza Artificiale" come riferimento ad una macchina la quale fosse in grado di simulare gli aspetti del pensiero e dell'apprendimento umano. Da questa proposta venne sviluppato un programma in particolare, il "Logic Theorist" di Allen Newell ed Herbert Simon, il quale era stato creato per dimostrare teoremi matematici contenuti nel testo *Principia Mathematica* [14] di Russell e Whitehead a partire dai postulati basici.

1.2.3 General Problem Solver e altri risolutori

Dopo aver creato il "Logic Theorist", Newell e Simon si sono messi al lavoro per raffinare il loro programma per ampliarlo a gestire utilizzi quanto più generici possibili; da questi sforzi nasce il "General Problem Solver", o GPS, il quale tentava un approccio di risoluzione molto più simile a quello utilizzato dagli esseri umani, in particolare quella che venne

definita "analisi mezzi-fini", una tecnica euristica utile per determinare i passaggi necessari per raggiungere un certo stato a partire da uno dato. Un altro programma dell'epoca simile ai precedenti LP e GPS è il "Geometry Theorem Prover" del 1959, creato da Herbert Gelemter per dimostrare teoremi di geometria complessi. Nello stesso periodo McCarthy scrisse un documento intitolato *Programs with Common Sense* [11], nel quale ipotizza un programma ideale adatto a risolvere problemi di ogni tipo, non solo matematici, e che lo facesse in maniera analoga ad un umano, rendendolo un vero e proprio sistema intelligente; questo programma teorico prese il nome di "Advice Taker". Data l'enorme qualità di programmi di Intelligenza Artificiale che vennero creati in quel periodo, gli studiosi dell'epoca iniziarono a riporre molte speranze nel futuro della tecnologia, e per questo si crearono aspettative completamente irrealistiche. Nel 1957 Simon teorizzò che nel giro di 10 anni qualcuno avrebbe creato una macchina in grado di competere con i Grandi Maestri di scacchi, nonostante una macchina di questo tipo non fu creata se non dopo 40 anni da quell'affermazione. Intorno al 1960, in effetti, i progressi sull'Intelligenza Artificiale sembrarono fermarsi di colpo, sia per l'incapacità delle macchine di possedere conoscenza semantica che per le difficoltà, da parte degli algoritmi, di risolvere teoremi che partivano da molti assiomi. Considerati questi problemi, nel 1966 il governo statunitense decise di interrompere i fondi destinati alla ricerca di macchine di questo tipo, infliggendo un duro colpo allo sviluppo dell'IA.

1.2.4 Ingresso nel mondo dell'industria e reti neurali

Le intelligenze artificiali non si espansero fuori dal mondo accademico fino al 1982, quando la Digital Equipment iniziò ad utilizzare un sistema chiamato R1, il quale aveva lo scopo di aiutare gli addetti ai lavori a configurare ordini in arrivo per nuovi computer. Nel 1986 l'azienda affermò che l'utilizzo di R1 nei quattro anni precedenti fece risparmiare circa 40 mila dollari l'anno. Digital Equipment non fu l'unica; infatti, nel corso degli anni '80, quasi ogni grande azienda americana aveva iniziato ad adottare sistemi simili per la gestione del proprio business. Lo sviluppo di tecnologia simili si ebbe soprattutto in Giappone, con l'introduzione, nel 1981, del programma "Fifth Generation" per lo sviluppo di sistemi intelligenti basati sul linguaggio Prolog. Nel 1986 si ritornò a fare esperimenti sulle "reti neurali"; questo tipo di tecnologia fu inventato per la prima volta negli anni '50 ed utilizzava porte logiche per simulare il comportamento dei neuroni umani. Tale sviluppo avvenne grazie all'uso di un nuovo algoritmo, chiamato "back-propagation", il quale diede il via a quello che venne definito il modello "connessionista" delle intelligenze artificiali, a differenza dei modelli simbolici ideati da Newell e Simon. Questi due modelli vennero però visti come complementari, e quindi entrambi ricevettero attenzioni dalla comunità scientifica senza entrare in diretta competizione l'uno con l'altro.

1.2.5 IA moderna

Negli anni seguenti ci fu un aumento esponenziale sia della capacità di calcolo dei computer che dell'interesse nei confronti di queste nuove tecnologie, e da questi due fattori ne è scaturito un utilizzo sempre maggiore dell'Intelligenza Artificiale in qualsiasi ambito possibile, dall'implementazione in programmi di utilizzo comune, all'uso nelle grandi aziende per la gestione di sistemi manageriali immensi, fino alla creazione di veri e propri strumenti interamente basati sull'IA per aiutare a svolgere un compito di tutti i giorni o anche problemi più specializzati. Alcuni utilizzi notevoli dell'IA dagli anni '90 in avanti sono:

- La vittoria nel 1997 di Deep Blue, computer dell'IBM realizzato per giocare a scacchi, contro il campione mondiali dell'epoca Garry Kasparov.

- Lo sviluppo dei chatbot: computer incentrati sull'interpretazione del linguaggio naturale e nella capacità di fornire risposte quanto più vicine a quelle che darebbe un umano: si va da ALICE, il primo chatbot in assoluto, sviluppato nel 1995 da Richard Wallace, passando per Cleverbot e Siri negli anni 2010, per finire con i chatbot moderni di OpenAI, come ChatGPT, i quali fanno uso delle ultimissime scoperte in campo AI.
- La creazione di macchine con sistemi di guida automatica a partire dal 2010, con compagnie come Uber, Tesla e Google, le quali hanno portato progressi notevoli nello sviluppo delle intelligenze artificiali per l'uso nella mobilità.

1.3 Aspetti positivi delle IA

Le intelligenze artificiali sono state create con lo scopo di essere degli strumenti da utilizzare per aiutare l'uomo nello svolgimento di compiti particolarmente difficili o anche per l'automazione di lavori considerati semplici ma tediosi; e in questa sezione si analizzeranno alcuni dei contesti e dei casi d'uso in cui le intelligenze artificiali sono in grado di semplificare il lavoro in maniera considerevole.

1.3.1 Eliminazione dell'errore umano

L'Intelligenza Artificiale si rivela un alleato fondamentale nella mitigazione dell'errore umano nei vari ambiti grazie alla sua capacità di analizzare grandi quantità di dati e prendere decisioni di conseguenza. Gli esseri umani sono soggetti a limitazioni, come stanchezza, distrazione e giudizio influenzato da emozioni o fattori esterni, che possono portare a errori. L'IA, invece, opera con coerenza e senza bisogno di pause, garantendo un'elevata precisione nelle attività ripetitive e complesse. Un esempio pratico è l'uso nel settore medico: un'intelligenza artificiale è in grado di analizzare immagini mediche meglio e più nel dettaglio di un dottore, segnalando anomalie che potrebbero sfuggire all'occhio umano. Nell'ambito dei trasporti, i veicoli autonomi guidati da IA sono progettati per evitare errori dovuti a distrazione o affaticamento del conducente. Ovviamente tutto questo presuppone che non ci siano errori nella programmazione dell'IA stessa, per quanto dei sistemi abbastanza avanzati potrebbero essere in grado di riconoscere autonomamente se qualcosa dovesse andare storto ed agire di conseguenza per mitigare anche errori di questo genere.

1.3.2 Cybersecurity

Uno dei campi in cui l'Intelligenza Artificiale è riuscita a dimostrare la sua efficacia è quello della cybersecurity; infatti, le capacità di queste tecnologie si prestano perfettamente alla difesa di dati informatici e alla prevenzione di danni, sia per privati che per le aziende. Nello specifico, le IA sono in grado di svolgere diversi compiti legati alla sicurezza informatica; Tra questi citiamo:

- Analisi di dati provenienti da fonti come log di attività, o traffico di rete, per individuare modelli sospetti o anomalie che potrebbero indicare un attacco in corso.
- Costruzione di profili di comportamento degli utenti legittimi e rilevazione di attività anomale o inconsuete, come tentativi di accesso non autorizzato, tracciando, così, una linea del "bad actor" ben diversa dall'utente tipico.
- Addestramento per riconoscere firme digitali associate a malware noti e identificazione di attacchi già conosciuti. Inoltre, le IA possono adattarsi e apprendere nuove firme e metodi di attacco man mano che emergono nuove minacce.

- Migliore risposta agli incidenti in atto, come ad esempio isolare parti della rete compromesse o bloccare l'accesso a servizi sotto attacco per evitare ulteriori danni a parti ancora sane.
- Analisi delle vulnerabilità al fine di identificare potenziali falle nei sistemi e nei software, contribuendo a rafforzare le difese prima ancora di essere sotto attacco.

Si deve, anche, tenere conto del fatto che l'IA nella cybersecurity non è perfetta: le minacce evolvono costantemente ed essere sempre preparati diventa molto più difficile col passare del tempo. Oltre a questo, gli stessi aggressori possono sfruttare le tecnologie di IA per progettare attacchi più sofisticati.

1.3.3 Automazione

Una delle prospettive più interessanti delle intelligenze artificiali è sicuramente quella dell'automazione di certi compiti, alcuni di questi già menzionati nelle sezioni precedenti. Lo scopo primario delle IA, dopotutto, è quello di creare una macchina che ragioni quanto più come un umano, ed è quindi logico che uno dei suoi utilizzi primari sia come lavoratore instancabile, almeno per quanto riguarda compiti che risulterebbero tediosi e ripetitivi per un uomo, o che sarebbero meglio affidati a motori di calcolo in ogni caso, come l'elaborazione di dati. Nell'ambito dell'automazione di lavori, l'IA si può usare per ottimizzare la catena di produzione e gestire il flusso di produzione nelle aziende, oppure per gestire autonomamente processi come l'elaborazione delle transazioni, la gestione dei reclami dei clienti, l'analisi dei rischi di credito e la rilevazione delle frodi finanziarie. In tempi recenti stanno anche iniziando ad emergere servizi clienti completamente gestiti da chatbot, senza alcun intervento di operatori umani. L'uso dell'IA nel settore dell'automazione porta all'aumento della produttività, alla riduzione degli errori e ad ampia precisione; tuttavia, è bene fare presente che siamo ancora molto lontani dall'aver piena automazione di lavori più complessi, dato che l'intervento umano è ancora cruciale in situazioni più delicate e dove è richiesta molta più capacità di ragionamento di quanto un'IA attuale possa fornire.

1.3.4 Decision-making

L'Intelligenza Artificiale ha un vantaggio decisivo rispetto a qualsiasi analista umano: essa riesce ad avere una visione d'insieme dell'intero sistema e dei cambiamenti al suo interno, ed è per questo che affidarsi ai consigli di un'IA quando si tratta di prendere decisioni importanti potrebbe essere la mossa migliore. Un esempio di utilizzo è l'allocazione delle risorse all'interno di un'azienda per minimizzare i costi di produzione, oppure la previsione del rischio in ambito finanziario ed assicurativo per evitare di prendere decisioni su investimenti azzardati. Qualsiasi sia il caso, la capacità delle IA di riconoscimento di pattern e la velocità di analisi di moli enormi di dati la rendono ideale per prendere delle decisioni in maniera quanto più obiettiva possibile. Un'importante considerazione a riguardo è che non solo le IA possono essere soggette a dei bug nel loro codice, ma che le decisioni prese sono sempre e solo conseguenza dei dati che ricevono in input, aggiungendo quindi il rischio di fornire dati incompleti o non corretti: entrambi questi aspetti confermano che le IA sono uno strumento utilissimo per prendere decisioni ma che non dovrebbero essere l'unico punto di riferimento, in quanto non sempre capaci di produrre le verità più oggettive possibili.

1.4 Aspetti negativi delle IA

L'ascesa dell'Intelligenza Artificiale è uno degli eventi più significativi avvenuti nell'ultimo secolo e il suo sviluppo porterà immensi benefici alla razza umana; tuttavia tutto questo non avverrà senza pagare pegno. In effetti, l'Intelligenza Artificiale presenta anche molte insidie, forse più nascoste ad un occhio meno attento, ma che sono comunque presenti, e perciò da tenere in considerazione quando se ne parla, specialmente quando si entra nell'ambito dell'etica o del dibattito politico-filosofico.

1.4.1 Utilizzo bellico

Una delle preoccupazioni principali riguardo la pericolosità delle intelligenze artificiali è il loro utilizzo come armi. I governi di diverse nazioni hanno iniziato a studiare questa tecnologia e a valutare le possibili applicazioni in ambito bellico, di cui alcune sono già state usate e hanno dato prova della loro efficacia, come ad esempio i famosi droni kamikaze, cioè velivoli senza equipaggio il cui unico scopo è quello di colpire bersagli prestabiliti in maniera completamente automatica. L'uso più spaventoso dell'IA come arma rimane quello di torrette completamente automatiche, le quali hanno dei sistemi di mira che non richiedono interventi umani, dotandole così anche di una precisione senza eguali. Un'arma del genere, però, può sollevare questioni morali riguardo al fatto che sia giusto o meno dare ad una macchina il completo controllo di decidere la vita e la morte di una persona, senza contare che un eventuale malfunzionamento nel sistema di targeting potrebbe comportare tragedie senza precedenti, con moltissime vite perse per nulla, specialmente se una di queste armi venisse per qualche motivo rilasciata in posti ad alta densità abitativa.

1.4.2 Questioni etico-morali

Lo sviluppo recente delle tecnologie IA ha visto la nascita di un movimento opposto, spinto primariamente da preoccupazioni etiche riguardo l'utilizzo dell'intelligenza artificiale. In particolare, molte persone sono preoccupate dal prospetto che le IA finiranno per rimpiazzare moltissimi posti di lavoro, specialmente nei settori considerati più "umili", come quello manifatturiero, ma anche in posizioni che, fino a qualche anno fa, erano considerate insostituibili da qualsiasi tipo di macchina, come gli scrittori, i giornalisti o gli artisti. Riguardo a quest'ultimo gruppo, il 2 maggio 2023 la Writers Guild of America ha iniziato uno sciopero generale dei suoi membri per chiedere benefici maggiori per i suoi membri all'associazione americana per il cinema e la televisione (Figura 1.3); tra i punti di contesa c'è proprio una richiesta di regolamentazione dell'uso delle intelligenze artificiali come sostituti per scrittori ed artisti umani, i quali potrebbero essere visti come costi non necessari dalle grandi compagnie cinematografiche e televisive che potrebbero decidere di usare strumenti IA al loro posto, nonostante ci sia ancora un salto di qualità significativo tra il lavoro di un IA e quello di un essere umano. Questa forte spinta contraria allo sviluppo di Intelligenza Artificiale potrebbe rivelarsi come il nuovo movimento luddista del ventunesimo secolo, per quanto sia difficile prevedere quando e se questa fantomatica "sostituzione totale" dell'uomo con le intelligenze artificiali dovesse effettivamente avvenire in futuro.

1.4.3 Rogue AI

Sin da quando il concetto di Intelligenza Artificiale ha iniziato a farsi strada nella mente collettiva della popolazione, è emerso il timore che un giorno riuscissimo a creare un'intelligenza artificiale così avanzata da ribellarsi al nostro volere, con conseguenze catastrofiche per la razza umana. Nella sezione precedente sono già stati fatti i nomi di alcune delle più



Figura 1.3: Una manifestante durante lo sciopero della WGA del 2023

famose IA impazzite nella fantascienza, come HAL 9000 o GladOS, ma in effetti siamo ancora molto distanti da avere macchine così intelligenti da essere capaci di pensiero autonomo. Nonostante ciò eventuali bug nel codice delle intelligenze artificiali potrebbe comunque portarle ad "impazzire", sebbene, in quel caso, i risultati sarebbero sicuramente meno gravi e più facilmente risolvibili.

1.4.4 Consumo energetico e impatto ambientale

Un altro aspetto fondamentale delle tecnologie basate sull'Intelligenza Artificiale è quello dei consumi energetici: considerata la mole di potenza computazionale necessaria per permettere alle IA di eseguire i calcoli richiesti, è normale che sia necessaria una quantità di energia. È stato stimato che un singolo impianto possa consumare dai 20 ai 40 Megawatt di potenza, più o meno l'energia consumata da 16 mila nuclei abitativi nello stesso periodo. Oltre agli elevati consumi, gli stabilimenti producono anche enormi quantità di anidride carbonica; a tal fine possiamo considerare come esempio la stima che si è fatta sull'addestramento di un modello di IA generativa di media grandezza, la quale si aggira attorno alle 626 mila tonnellate di CO₂ per tutta la durata del training; questo equivarrebbe all'anidride carbonica emessa da cinque automobili durante tutto il loro ciclo di vita. Con l'aumentare della complessità delle Intelligenze Artificiali aumenteranno di sicuro anche i consumi e le emissioni, andando ancora di più a gravare sull'ambiente; ma, se si dovessero iniziare a prendere misure precauzionali per evitare questi inconvenienti, come degli stabilimenti interamente basati su energie rinnovabili, allora l'impatto negativo delle IA potrebbe venire severamente ridotto nel corso del tempo.

In questo capitolo verrà introdotto Amazon AWS, il servizio che offre gli strumenti alla base dei case study di questo documento. Si parlerà dell'importanza della piattaforma nel campo enterprise e nel mondo del cloud computing, per poi passare al funzionamento effettivo di essa e al modo in cui Amazon divide i servizi offerti in base alla regione, per poi finire con una menzione agli strumenti principali che la piattaforma mette a disposizione e al loro funzionamento generale.

2.1 Cos'è Amazon AWS

Per comprendere al meglio i risultati della ricerca è necessario prima capire il contesto alla base degli strumenti utilizzati, in particolare quello dietro all'esistenza e all'importanza di Amazon AWS. Questa sezione introduttiva fornirà informazioni generali su AWS e spiegherà come mai si è trasformato in una colonna portante dei servizi di cloud computing per chiunque cerchi di operare un programma con funzionalità Web del tipo offerto dalla piattaforma di Amazon.

2.1.1 Introduzione alla piattaforma

Amazon Web Services (AWS) è una piattaforma di servizi cloud fondata da Amazon nel 2006, la quale offre una vasta gamma di servizi e risorse informatiche su richiesta, consentendo alle aziende di utilizzare e gestire applicazioni e dati in un maniera flessibile e scalabile. L'architettura di AWS è strutturata in regioni, zone di disponibilità e edge locations, consentendo la distribuzione di risorse in tutto il mondo. AWS offre una vasta gamma di servizi, che spaziano dalla computazione e lo storage, come Amazon EC2, Amazon RDS e Amazon S3, all'Intelligenza Artificiale e analisi dei dati informatici, come Amazon Rekognition e Amazon SageMaker.

2.1.2 Importanza nel mondo cloud

Uno dei ruoli fondamentali assunti da AWS nel corso degli anni è stato quello di plasmare il concetto di cloud computing, introducendo soluzioni scalabili, flessibili ed economicamente convenienti, sia per i privati che per le aziende. La piattaforma di Amazon, infatti, ha ridefinito in maniera significativa il modo in cui molte aziende implementano aspetti informatici i quali, una volta, dovevano essere necessariamente gestiti completamente in locale da personale specializzato. La competenza, o perlomeno la familiarità, con gli strumenti di Amazon AWS

oramai è diventato un requisito fondamentale per chi lavora nel mondo IT, come risulta dal fatto che tale competenza si trova molto spesso negli annunci per lavorare in aziende informatiche prestigiose. AWS non è l'unica piattaforma cloud largamente diffusa: nel mondo enterprise hanno enorme importanza anche Microsoft Azure, Google Cloud Platform ed Alibaba Cloud, le quali, insieme ad AWS, rappresentano circa il 60% delle piattaforme cloud utilizzate.

2.1.3 Vantaggi per le enterprise

Il motivo principale per cui le piattaforme cloud sono diventate onnipresenti nel campo dell'informatica è la miriade di vantaggi e semplificazioni che le aziende ottengono adottando servizi come Amazon AWS. Tra quelli principali possiamo citare:

- *Estrema scalabilità e flessibilità del servizio*, permettendo alle aziende di adattarsi alla variazione di domanda di un determinato servizio senza dover investire in ulteriore hardware che potrebbe poi andar sprecato.
- *Adozione del modello "Pay-as-you-go"*, dove le aziende pagano esclusivamente per la quantità di servizio effettivamente utilizzato, invece di dover aderire ad abbonamenti rigidi, i quali possono portare a sprechi o a mancanze di risorse in caso di necessità.
- *Riduzione dei costi*, dato che mantenere e gestire infrastrutture fisiche può risultare molto gravoso dal lato economico, specialmente per compagnie di grandi dimensioni.
- *Alta velocità di implementazione*, grazie alla semplicità di configurare un servizio nuovo in poco tempo e alla disponibilità dei servizi di AWS.
- *Affidabilità dei servizi e resilienza estrema* a guasti o altri problemi tecnici.
- *Automazione di alcune attività di gestione*, come la distribuzione delle risorse, il monitoraggio e i controlli di sicurezza.
- *Copertura globale e distribuzione semplice* dei servizi in moltissime aree diverse.
- *Continua innovazione da parte di AWS* per adottare soluzioni sempre più convenienti ed efficienti per i propri clienti.

2.2 Architettura di AWS

Amazon Web Services risulta essere estremamente efficiente e solida nel fornire i suoi servizi; ciò è dovuto alla struttura che Amazon ha deciso di adottare per la sua piattaforma cloud, così da poter operare in sicurezza e senza doversi preoccupare delle instabilità che un servizio di queste dimensioni potrebbe comportare. In questa sezione si parlerà dell'architettura alla base di AWS, con particolare enfasi sui metodi utilizzati per ottenere la scalabilità e parlando della divisione in regioni operative.

2.2.1 Struttura del servizio

La struttura generale di AWS è basata su una distribuzione complessa e ridondante per garantire affidabilità e prestazioni ottimali. Amazon adotta l'approccio delle "zone di disponibilità", ciascuna delle quali è composta da uno o più data center fisici separati, ma logicamente collegati all'interno della stessa regione geografica. Ogni zona di disponibilità è progettata per resistere alle interruzioni, garantendo estrema resilienza e resistenza ai



Figura 2.1: Un processore AWS Graviton2

guasti per l'intero sistema. Inoltre, AWS utilizza tecnologie di virtualizzazione avanzate, come la virtualizzazione basata su Xen, per creare istanze virtuali delle risorse hardware, offrendo flessibilità nell'allocazione delle risorse e consentendo alle applicazioni di adattarsi dinamicamente alle esigenze di carico. Oltre a ciò, AWS integra servizi di bilanciamento del carico, gestione delle immagini delle istanze e monitoraggio delle risorse. Dal punto di vista architetturale, una delle soluzioni più interessanti adottate da Amazon è l'utilizzo di CPU ARM-based progettato da Annapurna Labs, una compagnia israelita facente parte del gruppo di AWS; questi processori sono chiamati *AWS Graviton* (Figura 2.1). Amazon afferma che la seconda iterazione di queste CPU, Graviton2, ha permesso loro di aumentare le performance del 40% a parità di costi, oltre a ridurre i consumi energetici del 72%.

2.2.2 Scalabilità

Uno dei pilastri fondamentali della tecnologia di AWS è quello della scalabilità, cioè la capacità di un sistema di modificare la propria struttura software o hardware facilmente, anche di fronte a variazioni notevoli, sia nella quantità di dati da processare, che nella modalità con cui vengono forniti al servizio di competenza. AWS implementa la scalabilità in diversi modi; tra questi ne citiamo alcuni dei più importanti:

- *Scalabilità orizzontale*: aggiungendo semplicemente più computer ad un sistema, la capacità di elaborazione totale aumenterà. Questo processo è contrapposto alla "scalabilità verticale", dove si aggiungono ulteriori risorse ad una singola macchina, come un numero più alto di CPU, oppure si sostituiscono processori più lenti con altri più veloci.
- *Auto Scaling*: tramite questa funzione, le risorse utilizzate da un sistema sono monitorate automaticamente ed adattate in base al carico sul sistema in un dato momento. L'utente può impostare politiche di riassegnazione risorse per specificare meglio quando e come modificare le risorse assegnate al sistema.
- *Elastic Load Balancing*: questa feature di AWS permette di distribuire automaticamente il traffico diretto ad un determinato servizio verso più istanze o zone di disponibilità diverse, così da garantire che il carico sia distribuito uniformemente per evitare colli di bottiglia e altri tipi di degrado della qualità.

- *Serverless computing*: servizi come AWS Lambda sono in grado di eseguire codice, gestire dati ed integrare applicazioni locali senza doversi preoccupare dell'infrastruttura hardware sottostante, andando anche a coprire tutti gli eventuali problemi dati dalla scelta delle capacità del sistema.

2.2.3 Divisione in regioni

Amazon Web Services opera secondo un modello diviso in regioni, cioè dei territori circoscritti all'interno dei quali sono disponibili uno o più data center. Queste regioni coincidono spesso con continenti o altri raggruppamenti geografici di notevoli dimensioni, come, ad esempio, la regione del Nord America, quella dell'Europa o quella del Medio Oriente. Queste regioni sono ulteriormente divise in un minimo di tre "zone di disponibilità" ("availability zones" in inglese, o AZ), le quali corrispondono ad un gruppo di data center logici con alimentazione, raffreddamento e reti propri, così da consentire grande tolleranza ai guasti, scalabilità e disponibilità. Le AZ all'interno di una stessa regione sono connesse tra di loro con una rete ad elevata larghezza di banda, il che permette loro di agire da partizioni per i servizi, garantendo così ridondanza e resilienza, anche tra zone di disponibilità diverse. Oltre a questo sistema, AWS offre anche delle "posizioni edge"; queste sono data center situati in tutto il mondo e che forniscono servizi a latenza bassissima, sfruttando la vicinanza fisica dell'utente ad uno di questi centri. Grazie a questa latenza estremamente ridotta, le posizioni edge sono anche usate da alcuni servizi per fare caching, così da garantire una velocità di risposta maggiore per dati già elaborati. Al fine di garantire la vicinanza di ogni utente ad almeno una posizione edge, Amazon mette a disposizione più di 400 di questi data center ai propri utenti, assicurandosi che vengano rispettati i requisiti di velocità promessi da questa funzionalità indipendentemente da dove si trovino i clienti.

2.3 Funzionalità principali

Amazon Web Services offre una vastissima gamma di funzionalità e servizi diversi, i quali spaziano dall'archiviazione e l'elaborazione dei dati, passando per programmi forniti come servizi Web, chiamati SaaS ("Software-as-a-Service") e finendo con applicazioni di Intelligenza Artificiale e Machine Learning, sfruttando le enormi capacità di elaborazione delle infrastrutture di AWS. In questa sezione parleremo dei servizi più importanti offerti dalla piattaforma di Amazon, compresi anche gli strumenti IA che verranno utilizzati per la ricerca sui case study presi in considerazione.

2.3.1 Amazon S3

Amazon Simple Service Storage (abbreviato in "S3") è il servizio di AWS che permette ai propri utenti di archiviare dati con il metodo "object storage", dove ogni file o gruppo di file è visto come un oggetto; tutto questo avviene tramite una semplice interfaccia web. Amazon S3 opera utilizzando dei "bucket", cioè dei contenitori di oggetti, con lo scopo di organizzarli e dividerli logicamente. Oltre all'interfaccia web, S3 mette anche a disposizione delle API REST per caricare, scaricare o eliminare dati da uno o più bucket. Un oggetto di un bucket S3 può avere delle dimensioni massime di 5 terabyte, altrimenti dovrà essere diviso in parti. Tra gli utilizzatori più notevoli di questo servizio abbiamo Netflix e Reddit, quest'ultimo completamente hostato su S3. Il vantaggio di usare il sistema ad oggetti e bucket è una altissima affidabilità, scalabilità e bassa latenza nel trasferimento di dati, oltre ad essere intuitivo e semplice da utilizzare, nonché a permettere un sistema di versioning molto efficiente.

2.3.2 Amazon EC2

Amazon Elastic Compute Cloud ("EC2") offre agli utenti AWS la possibilità di affittare macchine virtuali da utilizzare per eseguire le proprie applicazioni. Amazon EC2 utilizza le *Amazon Machine Image* (AMI) per configurare le macchine virtuali, chiamate "istanze". L'utente ha pieno controllo sui tempi e le modalità con le quali le istanze operano, oltre a poter pagare solo per il tempo nel quale queste istanze sono attive. Il cliente ha anche il controllo sulla posizione geografica delle istanze, così da assicurare una bassa latenza per gli utenti che si trovano in una determinata zona. Amazon EC2 utilizza sia la virtualizzazione Xen, che una basata su KVM, chiamata "Nitro". EC2 offre anche una buona scelta di sistemi operativi da utilizzare, come diverse distribuzioni di Linux, tra le quali anche una creata da Amazon stessa e basata su Fedora/RHEL, oltre ad alcune versioni di Windows Server e FreeBSD. L'uso di EC2 risulta estremamente efficiente e conveniente per le aziende, specialmente per i costi bassi e l'elasticità dei pagamenti; per dimostrare questo, Amazon è passata ad ospitare il suo servizio di e-commerce interamente su EC2 dal 2010 in avanti.

2.3.3 Amazon RDS

Un ulteriore metodo di archiviazione dati offerto da Amazon è quello di *Amazon Relational Database Service*, il quale rientra nella categoria dei DBaaS ("Data-Base-as-a-Service"). Come il nome suggerisce, RDS consente agli utenti di ospitare un'istanza di una base di dati relazionali sui server di AWS. L'utente potrà scegliere quale DBMS usare tra quelli più diffusi, come PostgreSQL, MySQL, Oracle DB e Microsoft SQL Server. Anche questo servizio offre API di accesso al database, oltre alla classica interfaccia Web. AWS mette a disposizione diverse istanze di DB, in base alle dimensioni e ai carichi di lavoro che il database dovrà gestire. Tra le altre funzionalità di nota di Amazon RDS citiamo:

- La possibilità delle istanze di trovarsi su più zone di disponibilità contemporaneamente, assicurando una buona velocità di trasferimento anche tra luoghi geografici distanti tra loro.
- La creazione automatica di snapshot della base di dati, così da permettere il ripristino in maniera semplice in caso di danni o errori, così da assicurare estrema resilienza e ridondanza dei dati.
- La gestione completa della sicurezza, dell'accesso ai dati del DB e anche della crittografia, così da garantire massima affidabilità per i propri clienti.

2.3.4 Strumenti IA

Uno dei gruppi di servizi più interessanti offerti da Amazon Web Services è quello degli strumenti di Intelligenza Artificiale e Machine Learning che la piattaforma mette a disposizione dei propri utenti. Grazie alla struttura di AWS, tutti questi strumenti sono utilizzabili in maniera estremamente semplice tramite l'interfaccia Web, senza doversi preoccupare della potenza di calcolo normalmente richiesta per effettuare compiti del genere. AWS offre tantissimi strumenti diversi; di questi citiamo quelli più interessanti:

- *Amazon SageMaker*: questo strumento permette agli utenti di creare ed allenare modelli di Machine Learning utilizzando il cloud di Amazon, oltre a permettere il deployment di questi modelli per sistemi embedded e altri dispositivi simili.
- *Amazon Rekognition*: il servizio di analisi di immagini e video offerto da AWS, il quale sfrutta tecnologie di computer vision ed offre tantissime funzionalità diverse, come il

riconoscimento delle espressioni facciali, il confronto tra persone, l'analisi del contenuto di un'immagine o video tramite label.

- *Amazon Lex*: questo servizio consente di gestire la creazione di interfacce conversazionali e chatbot di svariate tipologie diverse. Un esempio notevole di applicazione di Lex si ha con Alexa, l'assistente vocale di Amazon stessa.
- *Amazon Transcribe*: uno strumento per la trascrizione di tracce audio in testo, sia per la creazione di sottotitoli, utilizzando formati come VTT o SRT, che per la semplice conversione in plain text.
- *Amazon Comprehend*: un servizio di Intelligenza Artificiale piuttosto avanzato, il quale permette di estrarre informazioni importanti da qualsiasi tipo di testo fornito ad esso in input. Alcune delle funzionalità più interessanti includono l'analisi del tono del documento e il riconoscimento di relazioni tra concetti all'interno del testo.

Questo capitolo tratterà del campo dell'Intelligenza Artificiale che si occupa del riconoscimento e dell'elaborazione del testo, in particolare l'estrapolazione di testi da immagini che ne contengono. Si parlerà dei concetti alla base del funzionamento di questa tecnologia e delle tecniche adottate per ottenerla. Si passerà poi a parlare dello strumento di Amazon Web Services utilizzato a questo scopo e delle sue funzionalità, per poi finire con alcuni casi d'uso interessanti, presi in esame per dimostrare il funzionamento di questi strumenti di IA.

3.1 Introduzione al riconoscimento di testi

Questa sezione servirà da semplice introduzione per ciò che verrà analizzato successivamente all'interno di questo stesso capitolo, introducendo in primis cosa verrà studiato ed il principio dietro all'uso dell'Intelligenza Artificiale per questo compito.

3.1.1 Cos'è il riconoscimento testi

Il riconoscimento di testi nelle immagini è una delle tante applicazioni dell'Intelligenza Artificiale, più nello specifico è una branca della "image recognition", la quale consente di identificare e estrarre il testo presente in immagini o fotografie. Durante il processo di riconoscimento, l'IA analizza l'immagine per individuare le aree in cui è presente il testo, identifica i caratteri e le parole, e successivamente converte questo testo in un formato leggibile. Questa tecnologia ha molteplici applicazioni pratiche; alcuni esempi sono:

- *Trascrizione di documenti scansionati*, utilizzato in particolare con vecchi atti ufficiali che possono risultare difficili da leggere.
- *Indicizzazione di contenuti visivi*, per garantire leggibilità e una facile classificazione delle immagini.
- *Creazione automatica di metadati per immagini*, molto utili in caso siano necessarie ulteriori elaborazioni da parte di altri strumenti AI.
- *Accessibilità per i non vedenti*, così da permettere a persone con disabilità di usufruire più facilmente di contenuti visivi in maniera automatica.

Come ogni strumento di IA, la costante ricerca e il costante sviluppo nel campo permetteranno al riconoscimento di testi di evolversi oltre le capacità umane, riuscendo a leggere

e riconoscere lettere e parole che un essere umano potrebbe non capire; un esempio lampante di questo è il fatto che certe IA stanno diventando più brave degli esseri umani a risolvere i classici CAPTCHA basati sul testo distorto, rendendo questa tecnologia quasi controproducente.

3.1.2 Come funziona il riconoscimento di testi tramite machine learning

Come la maggior parte degli strumenti che utilizzano l'Intelligenza Artificiale, il loro funzionamento si basa sul concetto di "machine learning", che significa fornire dati ad un sistema, utilizzando un insieme di feedback positivi e negativi sul campione di addestramento fino a quando la rete non abbia imparato a determinare autonomamente il risultato di una determinata richiesta. Nel caso del riconoscimento di testi, ci sono diversi passaggi da considerare:

- *Raccolta di immagini appropriate e creazione dataset*: per assicurarci che il nostro modello di Intelligenza Artificiale funzioni al meglio, dobbiamo trovare un quantitativo sufficiente di immagini con delle parole al loro interno, così da costruire un campione abbastanza variegato.
- *Labeling*: in questo processo si associa alle lettere sullo schermo un insieme di testi che possano essere compresi dalla rete neurale, così che possa riuscire ad associare l'aspetto delle lettere nelle immagini a dei testi effettivi.
- *Addestramento*: il dataset creato ed etichettato nei passi precedenti viene dato in input alla rete neurale da allenare. Una volta ricevuto, quest'ultima dovrà fare una predizione su quello che pensa sia contenuto all'interno dell'immagine, e a quel punto verrà fornito ad essa del feedback in caso di successo o fallimento; quest'informazione sarà poi usata nelle predizioni successive, le quali vengono svolte iterativamente fino a quando la rete neurale non raggiunge un livello di precisione accettabile.
- *Utilizzo*: ora che la rete è allenata, si potranno inserire delle immagini in input, le quali non sono mai state viste prima dalla rete e sulle quali quest'ultima non si è allenata, così da determinare l'effettiva precisione dello strumento in un caso d'uso normale.

3.2 AWS Rekognition

Ai fini di eseguire la ricerca sui casi d'uso che verranno presi in esame, si utilizzerà uno degli strumenti messi a disposizione da Amazon Web Services, più in particolare "Amazon Rekognition", un servizio di riconoscimento immagini con diverse funzionalità.

3.2.1 Cos'è Amazon Rekognition

Amazon Rekognition è un servizio di Intelligenza Artificiale per l'analisi ed il riconoscimento delle immagini messo a disposizione da Amazon Web Services. Questo strumento consente agli sviluppatori di incorporare funzionalità di IA nelle loro applicazioni senza dover sviluppare algoritmi complessi da zero. Rekognition utilizza modelli di deep learning per interpretare e analizzare il contenuto fornito in input, aprendo la strada a molteplici applicazioni pratiche in diversi settori. Grazie a questa tecnologia, gli utenti possono riconoscere oggetti, volti, scene e testo nelle immagini e nei video, nonché analizzare movimenti e attività.

3.2.2 Funzionalità di Amazon Rekognition

Amazon Rekognition offre una vasta gamma di funzioni legate al riconoscimento delle immagini, le quali sfruttano modelli addestrati da Amazon. Queste funzionalità possono essere utilizzate a costo zero da qualunque utente AWS e prendono il nome di "demo". Le funzioni principali delle demo di Rekognition sono:

- Rilevamento delle etichette;
- Proprietà dell'immagine;
- Moderazione delle immagini;
- Analisi facciale;
- Confronto facciale;
- Riconoscimento di volti celebri;
- Testo nell'immagine;
- Rilevamento dei PPE;
- Analisi dei video archiviati;

Gli utenti possono anche addestrare dei modelli in maniera autonoma, tramite l'utilizzo della funzionalità delle "Custom Labels", ma ciò può comportare costi aggiuntivi in base all'utilizzo di questi strumenti personalizzati, e quindi è spesso riservato ad aziende o in ambiti in cui è imperativo che il testing sia effettuato su modelli altamente specializzati, così da assicurarsi la massima precisione nei casi specifici previsti dal cliente.

3.3 Casi d'uso di riconoscimento testi nelle immagini

Tramite l'utilizzo delle demo menzionate in precedenza, è possibile effettuare dei test su alcune immagini prese in considerazione per quanto riguarda il caso d'uso rilevante. In questo caso si proverà ad utilizzare la funzionalità di "Testo nell'immagine" per cercare di comprendere le possibilità ed i limiti di questo strumento di Intelligenza Artificiale.

3.3.1 Immagini con testo scritto

L'utilizzo più comune dello strumento Testo nell'immagine è sicuramente quello di inserire in input un'immagine che contiene una semplice scritta al suo interno, come nell'esempio di Figura 3.1 di un manifesto da ricercato. Come si nota dalla figura, lo strumento riporta sulla destra il testo che è riuscito ad estrapolare dall'immagine.

Oltre ad una risposta più diretta, Rekognition offre anche una risposta in formato JSON (Figura 3.2), la quale può essere molto utile nel caso in cui questa tecnologia dovesse essere integrata con altri programmi.

3.3.2 Testo creato con oggetti

Uno dei modi per mettere alla prova il modello di Amazon è quello di inserire in input delle immagini dove il testo non è esplicitamente scritto, bensì è composto da oggetti per formare le lettere di una parola; per quanto possa risultare semplice per un essere umano, lo stesso non si può dire per un'Intelligenza Artificiale, forse poco allenata su questo caso.

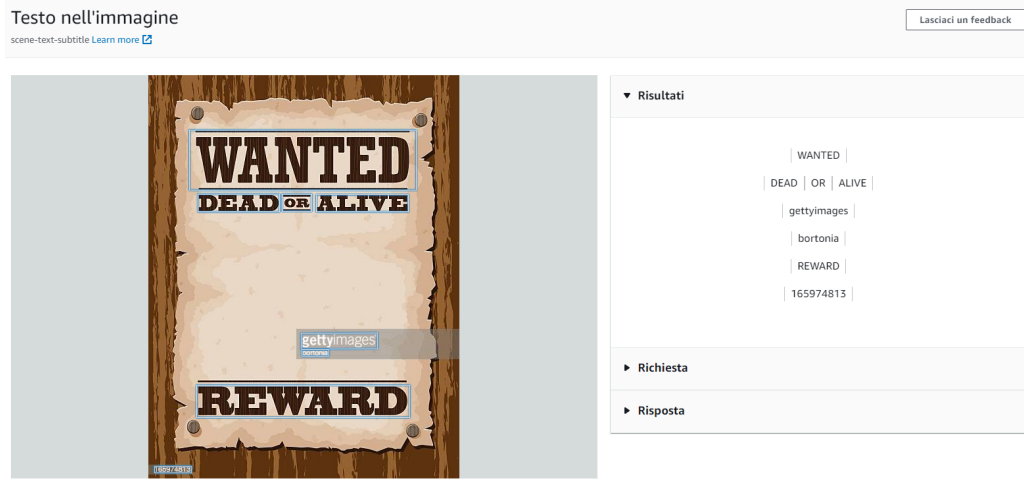


Figura 3.1: Risultato di "Testo nell'immagine" relativo a un manifesto da ricercato

```
"DetectedText": "WANTED",
"Type": "LINE",
"Id": 0,
"Confidence": 98.3177261352539,
```

Figura 3.2: Sezione della risposta in formato JSON effettuata su un manifesto da ricercato

Analizzando i risultati (Figura 3.3), in effetti il modello non è riuscito ad identificare affatto la parola composta dagli strumenti per la pittura, ma, al contrario, è stato perfettamente in grado di leggere due scritte minuscole sugli oggetti presenti.

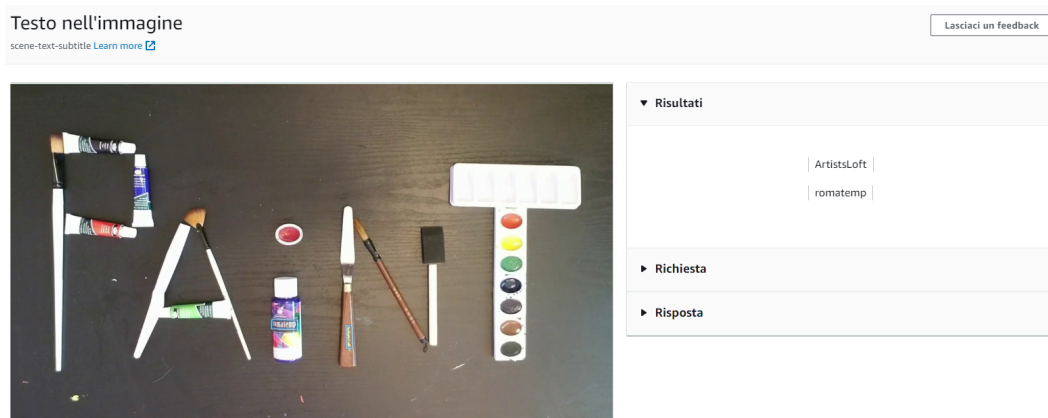


Figura 3.3: Risultato di "Testo nell'immagine" relativo ad una scritta composta da materiale per la pittura

3.3.3 Drone Skywriting

Rimanendo sempre nell'ambito delle scritte non convenzionali, possiamo mettere alla prova l'algoritmo di Rekognition con l'arte delle scritte in cielo composte dai droni, chiamata "Drone Skywriting". Nell'esempio scelto (Figura 3.4), lo strumento è in grado di leggere la scritta composta dai droni, ma non è riuscito a trovare il cancelletto posto all'inizio della parola.

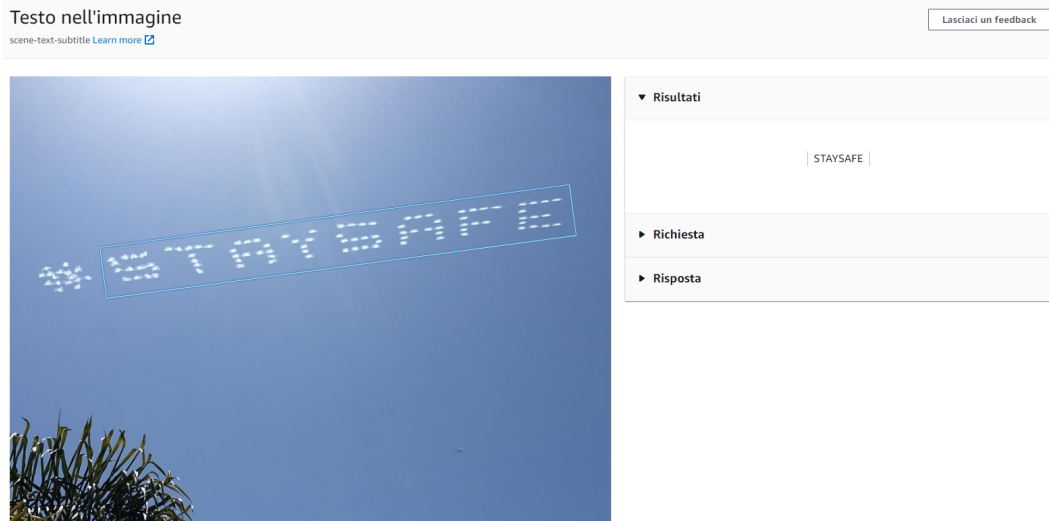


Figura 3.4: Risultato di "Testo nell'immagine" relativo ad una scritta composta da droni

3.3.4 Cartelli stradali

Uno dei casi d'uso più pratici che si può prendere in considerazione è quello dell'elaborazione del testo presente nei cartelli stradali. Come si può vedere dalla risposta (Figura 3.5), la rete neurale è riuscita senza problemi a riconoscere le scritte dei luoghi presenti nella foto dei cartelli presa in esame. Questa funzionalità potrebbe essere sfruttata da altri algoritmi per effettuare la stima sulla posizione geografica di una foto, basandosi soltanto sui luoghi menzionati dai cartelli e sulle loro relative distanze.

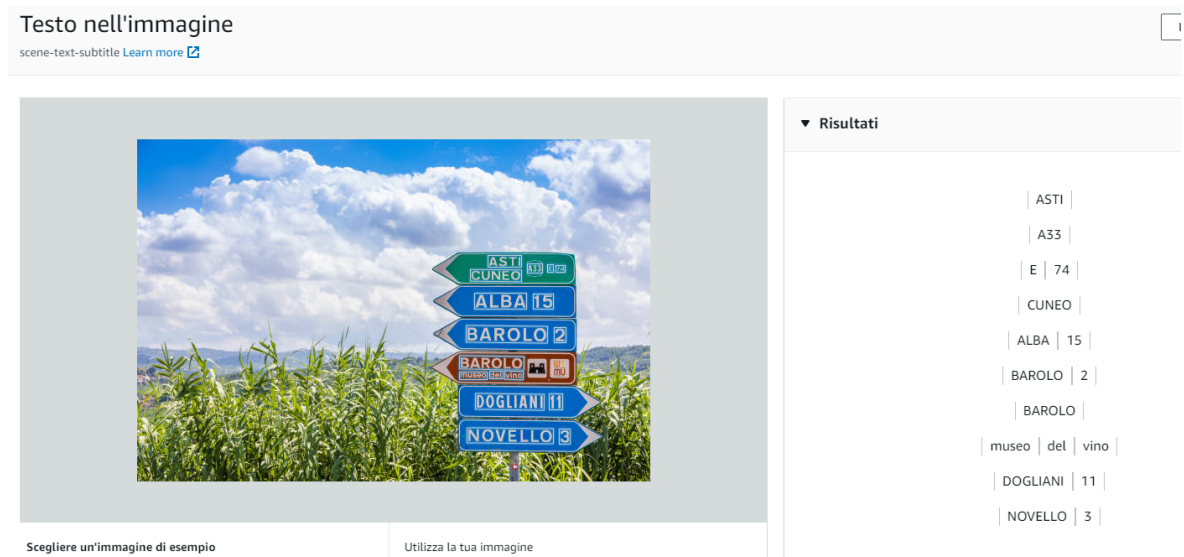


Figura 3.5: Risultato di "Testo nell'immagine" relativo ad una foto di un cartello stradale

Riconoscimento di immagini

In questo capitolo si parlerà di un altro aspetto dell'IA, quello relativo all'elaborazione delle immagini. Più nello specifico verranno introdotti diversi strumenti relativi al riconoscimento e all'analisi di volti. Si farà menzione a come le IA di questo tipo apprendono tramite machine learning; si proseguirà con degli esempi che dimostrano le molteplici funzionalità offerte da Amazon Rekognition per il riconoscimento dei volti.

4.1 Introduzione al riconoscimento di immagini

Prima di parlare di esempi concreti di riconoscimento di immagini, è necessario fare un'introduzione alla tecnologia ed al suo funzionamento, sottolineando anche alcune differenze con il caso discusso nel capitolo precedente.

4.1.1 Cos'è il riconoscimento di immagini

L'analisi dei volti nelle immagini è una delle diverse applicazioni della "image recognition", resa possibile dagli strumenti di Intelligenza Artificiale. In questo caso, si usano reti neurali addestrate con metodi di deep learning per riconoscere espressioni facciali, caratteristiche peculiari, somiglianze tra volti o altri aspetti di nota. Tra le applicazioni pratiche di questa tecnologia che sono state usate troviamo:

- *Riconoscimento di espressioni per feedback visivo*, al fine di analizzare automaticamente l'opinione di un cliente in base al suo volto.
- *Identificazione facciale*, per riconoscere se la persona sottoposta al test sia effettivamente un volto noto; ciò è molto utile in applicazioni di sicurezza.
- *Compilazione di trend di età e sesso*, nel caso un'azienda volesse analizzare alcune statistiche relative ai propri clienti deducibili in base alle loro caratteristiche facciali.
- *Rilevazione di dispositivi di protezione*, molto utile in ambiti dove tali dispositivi sono estremamente importanti per la sicurezza del lavoratore e degli altri soggetti sul luogo di lavoro.

4.1.2 Come funziona il riconoscimento di immagini tramite machine learning

Come menzionato nella sezione precedente, anche in questo caso si usano tecniche di machine learning per addestrare il modello, così da ottenere un algoritmo che ci garantisca la

massima precisione possibile nell'analisi dei volti. Si userà una rete neurale convoluzionale (CNN), ma questa volta ci saranno delle differenze significative nel processo rispetto al caso precedente. Più nello specifico:

- Nella fase di raccolta dati e creazione del dataset si dovrà puntare a raccogliere un gruppo di volti quanto più variegato possibile, diversificandoli significativamente in base a etnia, età, sesso, condizioni di luce e qualità generale della fotografia, così da garantire che il modello sia in grado di riconoscere qualsiasi tipo di immagine che contenga un viso.
- Nella fase di labeling si dovranno allegare tantissime informazioni diverse, in particolare informazioni biografiche e fisiche deducibili dalla fotografia stessa, come la presenza di barba, baffi, occhiali o altre caratteristiche apparenti, oltre alle emozioni che il volto sta esprimendo, nonché se questa persona sia qualche celebrità, così da poterla riconoscere nel caso venisse domandato all'IA.
- Nella fase di addestramento, il funzionamento è sostanzialmente identico all'addestramento di riconoscimento del testo, con la differenza che il feedback fornito alla rete neurale non riguarderà lettere, bensì se essa sia stata in grado di riconoscere o meno le etichette corrette, e quanto sia sicura di ciò.

4.2 Casi d'uso di riconoscimento di immagini

Tramite l'utilizzo di altre funzionalità di Amazon Rekognition andremo ad analizzare alcuni casi d'uso relativi al riconoscimento e all'estrapolazione di informazioni da alcune immagini di volti umani. Più nel dettaglio, useremo le seguenti funzionalità fornite dallo strumento di AWS:

- analisi facciale;
- riconoscimento di volti celebri;
- confronto facciale.

4.2.1 Analisi facciale

Questa funzione di Amazon Rekognition ci permette di fornire in input il volto di una persona e ricevere informazioni riguardo ad esso. Nel primo esempio (Figura 4.1) Rekognition è stato in grado di riconoscere il volto umano nella foto, isolandolo in alto a destra, per poi fornire informazioni accurate riguardo al soggetto, identificandolo come un maschio tra i 21 e i 29 anni, con un'espressione di stupore o di paura.

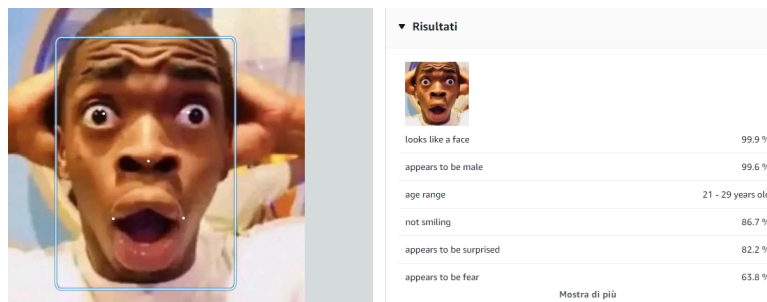



Figura 4.1: Risultato di "Analisi facciale" relativo ad un volto sorpreso


Risultati simili sono stati ottenuti fornendo l'immagine di un modello 3D di una donna (Figura 4.2), riuscendo anche in questo caso a cogliere alcuni dettagli specifici del volto, data anche l'enorme somiglianza con la foto di un essere umano vero.



looks like a face	99.9 %
appears to be female	99.9 %
age range	18 - 26 years old
smiling	96.3 %
appears to be happy	99.7 %
not wearing glasses	97.6 %
not wearing sunglasses	99.9 %
eyes are open	98.1 %
mouth is open	94.9 %
does not have a mustache	98.5 %
does not have a beard	95.8 %
face is not occluded	99.8 %

Figura 4.2: Risultato di "Analisi facciale" relativo ad un modello 3D

Utilizzando un volto disegnato, più precisamente quello di Lord Neverember dal gioco *Neverwinter Nights* (Figura 4.3), l'IA è riuscita comunque a riconoscere dettagli notevoli del viso del personaggio, ma ha avuto problemi su qualche particolare, per esempio sul fatto che fosse sicura che il personaggio avesse una barba, quando, chiaramente, possiede solo dei baffi.



looks like a face	99.9 %
appears to be male	99.9 %
age range	42 - 50 years old
not smiling	94.3 %
appears to be confused	62.6 %
not wearing glasses	96.4 %
not wearing sunglasses	99.9 %
eyes are open	93.5 %
mouth is closed	88.4 %
has a mustache	82.2 %
has a beard	99.1 %
face is not occluded	96.2 %

Figura 4.3: Risultato di "Analisi facciale" relativo ad un volto disegnato

4.2.2 Riconoscimento di volti celebri

Un'altra funzione interessante di Amazon Rekognition è quella relativa al riconoscimento di volti di persone famose con buoni livelli di confidenza e precisione, anche in foto dove la celebrità potrebbe risultare meno riconoscibile. Come primo esempio abbiamo preso il volto del famoso attore bambino Macaulay Culkin (Figura 4.4); L'IA è riuscita a riconoscerlo senza alcun problema con una precisione del 95%, nonostante il volto dell'attore oggi risulti meno noto di quello durante le sue apparizioni cinematografiche da giovane.

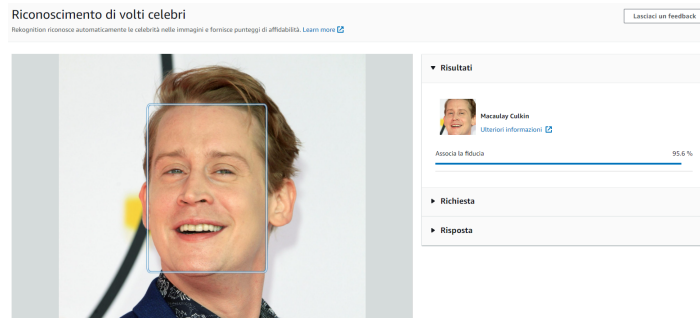


Figura 4.4: Risultato di "Riconoscimento di volti celebri" relativo ad una foto di Macaulay Culkin da adulto

In un altro esempio (Figura 4.5), Rekognition è riuscito senza alcun problema ad identificare il volto del cantante Michael Jackson da giovane, anche in questo caso con un valore di confidenza estremamente alto.

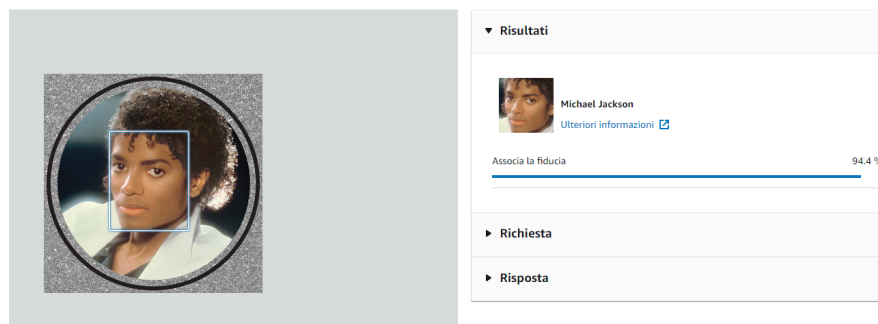


Figura 4.5: Risultato di "Riconoscimento di volti celebri" relativo ad una foto di Michael Jackson da giovane

Una tecnologia del genere potrebbe risultare estremamente utile quando si cerca di comporre dati statistici relativi alle apparizioni di certe figure note in un dato campo, come, ad esempio, quante volte il volto di un calciatore appare all'interno di un determinato sito di notizie sportive, o anche altri casi simili

4.2.3 Confronto facciale

L'ultima funzionalità di Amazon Rekognition che verrà presa in esame è "Confronto facciale". Ad essa devono essere fornite in input due immagini di volti umani, e l'IA proverà a dire se le due foto ritraggono o meno la stessa persona, anche in questo caso con livelli di confidenza variabili. Per il primo caso d'uso abbiamo utilizzato due foto del famoso attore inglese Ian McKellen (Figura 4.6); sono state scelte due foto a più di 60 anni di differenza l'una dall'altra; tuttavia Rekognition è stato in grado di affermare con molta sicurezza che si trattasse della stessa persona.

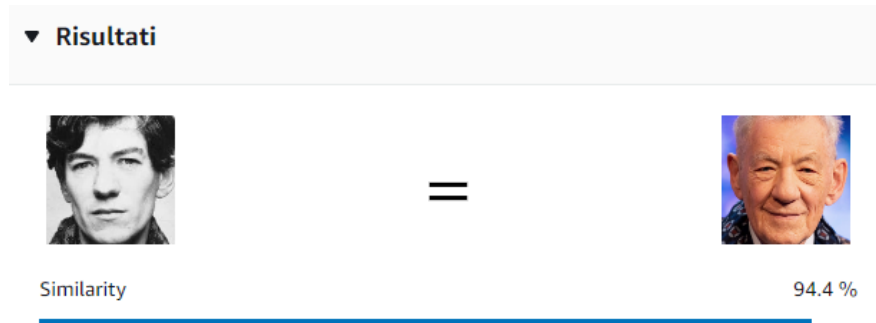


Figura 4.6: Risultato di "Confronto facciale" relativo a due foto di Ian McKellen

Per mettere alla prova lo strumento di Amazon abbiamo anche scelto due foto di Michael Jackson (Figura 4.7); in questo caso, l'IA non è riuscita ad affermare correttamente che si trattava della stessa persona, dicendo, invece, che erano due individui differenti. In questo caso la trasformazione estetica subita da Michael Jackson nel corso degli anni è stata abbastanza notevole da non permettere a Rekognition di fare una predizione accurata.



Figura 4.7: Risultato di "Confronto facciale" relativo a due foto di Michael Jackson

Le potenziali applicazioni di uno strumento del genere sono innumerevoli, soprattutto nel campo della sicurezza, dove una caratteristica fondamentale, come il volto di una persona, può agire da identificativo meglio di altri metodi, come password o anche DNA. Nonostante ciò, l'utilizzo di strumenti simili ha portato spesso a dibattiti molto accesi sulla privacy dell'individuo e su quanto fosse giusto utilizzare l'Intelligenza Artificiale per sorvegliare le persone.

Trascrizione da video a testo

In questo capitolo si esploreranno le possibilità offerte da un altro strumento basato sulla tecnologia dell'Intelligenza Artificiale, cioè la trascrizione di dialoghi e narrazioni nei filmati in formato testuale. Si parlerà del funzionamento generale di questa tecnologia, del modo in cui è in grado di apprendere e del funzionamento effettivo degli strumenti offerti da Amazon Web Services. L'ultima parte tratterà di casi d'uso concreti di questa tecnologia per la creazione di sottotitoli.

5.1 Introduzione alla trascrizione audio

Prima di entrare nel dettaglio del funzionamento degli strumenti che utilizzeremo, è necessario spiegare il funzionamento generale della tecnologia di riconoscimento vocale automatico, dell'elaborazione del linguaggio naturale e dei metodi di apprendimento utilizzati.

5.1.1 Cos'è la trascrizione audio

La tecnologia di trascrizione audio si basa su alcuni concetti dell'Intelligenza Artificiale, in primis su quello dell'elaborazione del linguaggio naturale ("Natural Language Processing" o NLP). Questo settore dell'IA si occupa di permettere alle macchine di comprendere i discorsi umani e di elaborarne il significato per estrapolarne informazioni importanti. L'NLP si basa sul Machine Learning ed è usato come base di tantissime tecnologie, come ad esempio:

- chatbot e altri sistemi di chat automatizzata;
- estrapolazione di informazioni da documenti;
- servizi di traduzione automatica ad alta precisione;
- filtri antispam;
- interpretazione di tracce audio contenenti discorsi umani;

In questo capitolo ci soffermeremo in particolare sull'ultimo esempio, cioè l'elaborazione di tracce audio con voce umana. Gli strumenti che permettono di eseguire compiti del genere fanno uso di una branca dell'NLP specializzata sui discorsi umani la quale prende il nome di "riconoscimento vocale automatico" ("Automatic Speech Recognition" o ASR). Lo scopo dell'ASR è quello di identificare quali parole stanno venendo pronunciate e di ricostruire

il discorso completo a partire da queste informazioni. In caso di parole poco chiare o di dubbi sull'uso di alcuni termini, un IA abbastanza allenata dovrebbe essere in grado di scegliere l'opzione che più si adatta al senso della frase, così da evitare periodi senza senso. Questa tecnologia può operare in una vasta gamma di lingue e dialetti diversi, così da aprire la possibilità a chiunque di poter usufruire di una trascrizione di elementi audiovisivi, soprattutto se si tratta di persone non udenti o con altri problemi simili.

5.1.2 Come funziona la trascrizione audio tramite machine learning

Per addestrare un modello di ASR si adottano algoritmi basati su reti neurali ricorrenti (RNN), tramite i quali avverrà l'elaborazione dell'audio. Per prima cosa si dovrà costruire un dataset abbastanza ricco, il quale dovrebbe contenere discorsi in quante più lingue, accenti e dialetti possibili, così da preparare il modello per funzionare nella maggior parte delle situazioni e dei contesti possibili. Queste tracce audio dovranno essere trascritte a mano, cosicché potranno fungere da base per far capire al modello se sta riuscendo o meno nel suo intento. Una volta preparati i dati, si eseguono alcuni passaggi di elaborazione delle tracce audio, tra cui:

- Si processa l'audio per renderlo più semplice da comprendere per la macchina tramite cancellazione dell'eco, rimozione dei rumori di fondo e normalizzazione del volume.
- Si estraggono le "feature audio", cioè i parametri numerici associati alle caratteristiche acustiche della traccia, come frequenza, intonazione e timbri vocali.
- Si dividono i dati di addestramento in "batch", cioè insieme di esempi, i quali verranno forniti successivamente alla rete neurale durante l'addestramento vero e proprio.

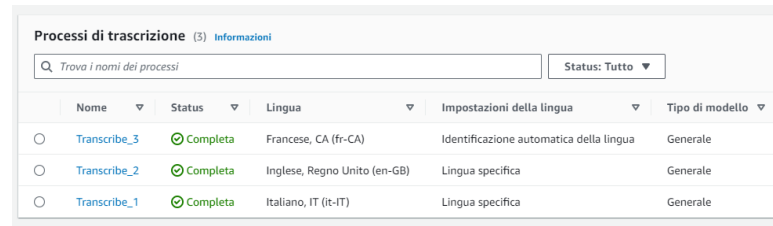
Durante l'addestramento il modello è allenato nell'associare un certo timbro ed intonazione della voce a certe parole, così da poter effettuare predizioni accurate una volta sottoposto a casi d'uso ad esso sconosciuti. Meglio è allenato il modello e più precise ed affidabili saranno le predizioni effettuate da esso in casi reali.

5.2 AWS Amazon Transcribe

Per poter utilizzare la tecnologia dell'Automatic Speech Recognition, adotteremo lo strumento di Amazon Web Services chiamato "Amazon Transcribe", così da poter testare le effettive capacità in un modello di trascrizione audio.

5.2.1 Cos'è Amazon Transcribe

Amazon Transcribe è un servizio di trascrizione da audio in testo offerto dalla piattaforma Amazon Web Services. Questo strumento sfrutta la tecnologia di riconoscimento vocale dell'Intelligenza Artificiale per permettere ai propri utenti di ottenere una copia in formato testuale dei discorsi contenuti all'interno di tracce audio, sia isolate che all'interno di un file video. Amazon Transcribe richiede in input un file precedentemente caricato sulla piattaforma di Amazon e permette agli utenti di scegliere tra molteplici opzioni per l'output della trascrizione. Questo servizio, a differenza di altri, presenta un limite di 60 minuti massimi di audio ogni mese, nel caso si scelga il piano gratuito; altrimenti, subentrano le tariffe classiche di AWS.



Processi di trascrizione (5) Informazioni					
<input type="text" value="Trova i nomi dei processi"/>					Status: Tutto ▼
Nome	Status	Lingua	Impostazioni della lingua	Tipo di modello	
<input type="radio"/> Transcribe_3	Completa	Francese, CA (fr-CA)	Identificazione automatica della lingua	Generale	
<input type="radio"/> Transcribe_2	Completa	Inglese, Regno Unito (en-GB)	Lingua specifica	Generale	
<input type="radio"/> Transcribe_1	Completa	Italiano, IT (it-IT)	Lingua specifica	Generale	

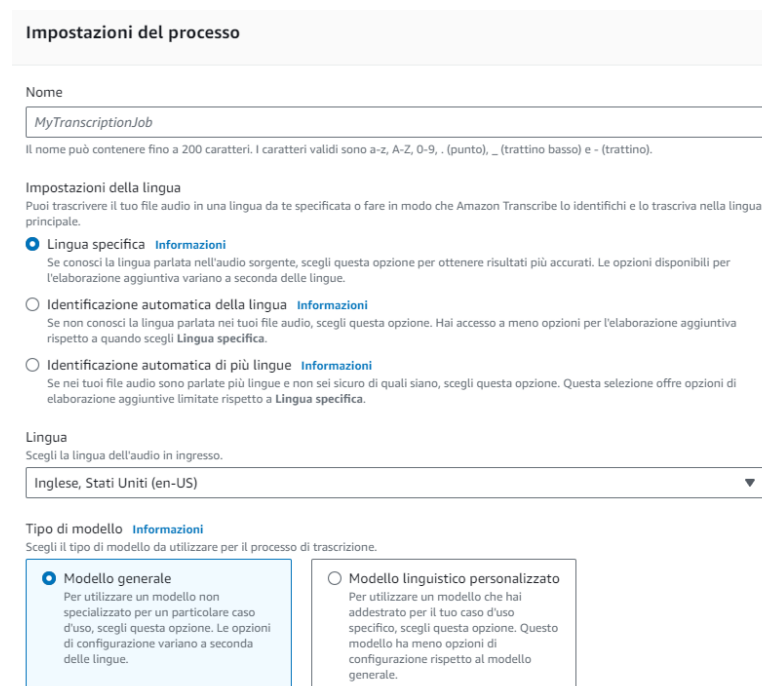
Figura 5.1: Interfaccia di Amazon Transcribe con la lista di processi completati

5.2.2 Funzionalità di Amazon Transcribe

Il servizio di AWS mette a disposizione diverse funzionalità, a partire dal tipo di input che verrà fornito a Transcribe. Sin dall'inizio avremo la possibilità di scegliere tra tre opzioni diverse per la trascrizione audio; queste sono:

- *Analisi in tempo reale*, per flussi di audio real-time.
- *Processo di trascrizione*, per file caricati sulla piattaforma.
- *Analisi post-chiamata*, per registrazioni di chiamate telefoniche.
- *Amazon Transcribe Medical*, una versione di Transcribe specializzata in analisi cliniche.

Oltre a queste opzioni, Transcribe permette all'utente di definire un proprio modello di Intelligenza Artificiale personalizzato, nel caso in cui quelli forniti da Amazon non siano particolarmente adatti al caso d'uso previsto. Un'altra funzionalità interessante è la definizione di un vocabolario personalizzato, così da permettere ai modelli di cercare parole particolari, le quali potrebbero non venir considerate come valide in una trascrizione standard. Per i casi d'uso che analizzeremo in seguito, useremo la funzionalità di "Processo di trascrizione", sfruttando l'upload di file video sulla piattaforma di Amazon (Figura 5.1).



Impostazioni del processo

Nome

Il nome può contenere fino a 200 caratteri. I caratteri validi sono a-z, A-Z, 0-9, . (punto), _ (trattino basso) e - (trattino).

Impostazioni della lingua
 Puoi trascrivere il tuo file audio in una lingua da te specificata o fare in modo che Amazon Transcribe lo identifichi e lo trascriva nella lingua principale.

Lingua specifica [Informazioni](#)
Se conosci la lingua parlata nell'audio sorgente, scegli questa opzione per ottenere risultati più accurati. Le opzioni disponibili per l'elaborazione aggiuntiva variano a seconda delle lingue.

Identificazione automatica della lingua [Informazioni](#)
Se non conosci la lingua parlata nei tuoi file audio, scegli questa opzione. Hai accesso a meno opzioni per l'elaborazione aggiuntiva rispetto a quando scegli **Lingua specifica**.

Identificazione automatica di più lingue [Informazioni](#)
Se nei tuoi file audio sono parlate più lingue e non sei sicuro di quali siano, scegli questa opzione. Questa selezione offre opzioni di elaborazione aggiuntive limitate rispetto a **Lingua specifica**.

Lingua
 Scegli la lingua dell'audio in ingresso.

Tipo di modello [Informazioni](#)
 Scegli il tipo di modello da utilizzare per il processo di trascrizione.

Modello generale

Per utilizzare un modello non specializzato per un particolare caso d'uso, scegli questa opzione. Le opzioni di configurazione variano a seconda delle lingue.

Modello linguistico personalizzato

Per utilizzare un modello che hai addestrato per il tuo caso d'uso specifico, scegli questa opzione. Questo modello ha meno opzioni di configurazione rispetto al modello generale.

Figura 5.2: Interfaccia di Amazon Transcribe per la creazione di una trascrizione audio

Selezionando questa opzione (Figura 5.2) ci vengono presentate fin da subito una miriade di opzioni diverse per il nostro lavoro di trascrizione, come il formato di output desiderato (plain text o formati per sottotitoli, .VTT o .SRT), il modello da utilizzare, la scelta tra rilevazione automatica della lingua o solo una lingua specifica e l'uso di tag.

Nella schermata successiva si hanno ulteriori opzioni per l'elaborazione della traccia audio, in particolare la richiesta di dividere la trascrizione per canali audio o di identificare se persone diverse parlino all'interno del video. Oltre a questo, ci sono ulteriori opzioni per la privacy, come la redazione di informazioni personali o la censura di alcuni termini, entrambe cose che non verranno, quindi, riportate nella trascrizione finale (Figura 5.3).

Configura processo - optional

Informazioni

Impostazioni audio

- Identificazione audio** [Informazioni](#)
Scegli di dividere l'audio multicanale in canali separati per la trascrizione o di suddividere le persone che parlano nell'audio in ingresso.
- Risultati alternativi** [Informazioni](#)
Abilita la visualizzazione di altri risultati di trascrizione
- Identificazione PHI** [Informazioni](#)
Identifica le informazioni sanitarie protette (PHI) in base alla regola sulla privacy HIPAA.

Rimozione dei contenuti

La rimozione dei contenuti nasconde le informazioni nella trascrizione risultante dal file audio sorgente. Amazon Transcribe modifica gli elementi della trascrizione ma non modifica l'audio sorgente.

- Redazione delle informazioni personali** [Informazioni](#)
Etichetta il tipo di informazioni personali e maschera anche il contenuto con il tipo di entità di informazioni personali nell'output della trascrizione. Ad esempio, (123) 456-7890 verrà mascherato come [TELEFONO].
- Filtraggio del vocabolario** [Informazioni](#)
Il filtraggio del vocabolario può rimuovere, mascherare o etichettare le parole specificate nella trascrizione finale.

Figura 5.3: Interfaccia di Amazon Transcribe con opzioni aggiuntive

Una volta scelti tutti i parametri e dopo aver fornito in input un file contenente tracce audio da un Bucket S3 (funzionalità della quale parleremo nella sezione successiva), Amazon Transcribe inizierà ad elaborare il contenuto fornito e, in base alla sua lunghezza e alla complessità della traccia, impiegherà un intervallo di tempo più o meno lungo prima di fornire i risultati nel formato richiesto.

Testo | Identificazione audio | Sottotitoli

VTT

devo usare la macchina. Oggi primo agosto scatta l'obbligo di esposizione nei distributori del prezzo medio del carburante. Vedete qui Diesel uno virgola sette sette La super senza piombo è quasi due euro. Siamo in Puglia, regione con un primato il costo più alto della benzina verde insieme alla regione autonoma di Bolzano. Noi siamo in partenza, faremo il pieno in Croazia trenta centesimi in meno, quindi un bel risparmio. Diciamo che è una mezza estorsione, quindi la vacanza te la fai ugualmente? Riduci un giorno rinunci ad una cena al ristorante e da qualche parte cerchi la quadra. Lei sta andando in vacanza? No, no, purtroppo no. Quanto mi costi? Ma chi ci guadagna in questa grande catena? Beh, sicuramente non i gestori, questo glielo posso, glielo posso dire molto serenamente, nel senso che noi abbiamo un margine fisso e qualunque sia il prezzo esposto alla pompa, per noi non cambia assolutamente nulla. Quanto incide sui bilanci familiari? E diciamo che anche per fare tragitti relativamente brevi. Ti fai i conti? Purtroppo non abbiamo la fabbrica degli euro dentro casa. Quando c'era la lira si arrivava tranquillamente. A fine mese i nostri genitori con l'inflazione a due cifre sono riusciti a farsi le case, le macchine si andava in vacanza e si risparmiava qualcosa.

Ma chi ci guadagna in questa grande catena?

26
00:01:18.709 --> 00:01:21.430
Beh, sicuramente non i gestori, questo glielo posso,

27
00:01:22.260 --> 00:01:25.949
glielo posso dire molto serenamente, nel senso che noi abbiamo un margine fisso

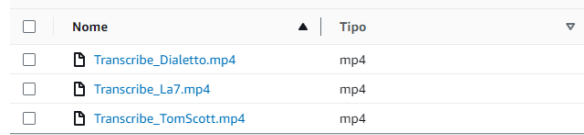
28
00:01:26.230 --> 00:01:28.199
e qualunque sia il prezzo

29
00:01:28.529 --> 00:01:30.360
esposto alla pompa,

Figura 5.4: Esempio di output di Amazon Transcribe in testo semplice ed in formato VTT

5.2.3 Interazione con S3

Come menzionato nella sezione precedente, al fine di fornire a Transcribe il file necessario per l'elaborazione, dobbiamo utilizzare un'altra delle funzionalità messe a disposizione da Amazon Web Services: nello specifico, bisogna usare Amazon S3. Quest'ultimo è un sistema di archiviazione in cloud scalabile, con lo scopo di permettere agli utenti di salvare i loro file in maniera sicura sulla piattaforma di Amazon e di dividerli logicamente in categorie separate tramite l'utilizzo dei "bucket". All'interno di questi bucket i file sono archiviati come "Oggetti" (Figura 5.5), ognuno con i propri metadati associati per distinguerli.



<input type="checkbox"/>	Nome	Tipo
<input type="checkbox"/>	Transcribe_Dialetto.mp4	mp4
<input type="checkbox"/>	Transcribe_La7.mp4	mp4
<input type="checkbox"/>	Transcribe_TomScott.mp4	mp4

Figura 5.5: Esempio di oggetti all'interno di un bucket Amazon S3

Per utilizzare lo strumento di trascrizione, l'utente dovrà creare un nuovo bucket il quale deve possedere un nome completamente univoco all'interno della regione selezionata; successivamente egli dovrà caricare manualmente i file che serviranno nell'elaborazione delle tracce audio. Una volta caricati questi file, basterà passare il collegamento all'oggetto S3 nelle impostazioni del processo di Transcribe, per poi procedere con l'utilizzo vero e proprio dello strumento.

5.3 Test di trascrizione video

Una volta spiegato il funzionamento dello strumento di trascrizione di Amazon Web Services, possiamo passare ai testing su alcuni casi d'uso concreti. In questa sezione vedremo in particolare tre esempi di video per i quali andremo a creare dei sottotitoli, i quali potranno essere implementati con programmi appositi.

5.3.1 Servizio televisivo in italiano

Il primo test è stato effettuato su un servizio televisivo del canale La7 riguardante il caro prezzi del carburante; più precisamente è stata considerata la versione caricata sul canale YouTube dell'emittente televisiva. Dopo aver selezionato la lingua italiana, il processo di elaborazione è iniziato ed ha prodotto con successo i sottotitoli richiesti. Una volta scaricato il file .SRT, lo abbiamo inserito all'interno del video tramite la funzione apposita di VLC Media Player (Figura 5.6).

Per quanto riguarda la qualità dei sottotitoli non sono stati riscontrati problemi eccessivi; l'IA non è riuscita a riconoscere il nome dell'emittente, scrivendola sempre come "la sette"; essa ha, inoltre, avuto un paio di problemi a differenziare tra i diversi interlocutori, specialmente quando essi parlavano uno sopra l'altro, per quanto l'opzione degli interlocutori multipli non fosse stata selezionata nelle opzioni aggiuntive. Transcribe ha anche avuto qualche problema con alcune inflessioni dialettali particolari, ma non in modo particolarmente drastico. Questo esempio risulta molto interessante, specialmente nei casi in cui si dovesse pensare all'accessibilità per i non udenti per servizi televisivi di questo tipo.

5.3.2 Video educativo in inglese

Per questo secondo esempio è stato scelto un video di Tom Scott, un noto creatore di contenuti inglese di tipo educativo e informativo. Il video in questione tratta di un bar che



Figura 5.6: Spezzone dal servizio di La7 con annessi sottotitoli prodotti da Transcribe

serve i propri clienti utilizzando dei tubi pneumatici. Una fatto interessante da notare è che, nonostante sia stata selezionata l'impostazione per la lingua inglese britannica, dato che lo stesso Tom Scott è britannico, buona parte del video viene narrata da uno dei baristi, il quale è neozelandese e quindi presenta un accento completamente diverso da quello dell'altro interlocutore principale; nonostante questo, Transcribe è riuscito a capire perfettamente ogni singola parola, anche quelle tipiche del neozelandese, rendendo questo set di sottotitoli il più vicino alla perfezione possibile.



Figura 5.7: Spezzone dal video di Tom Scott con annessi sottotitoli prodotti da Transcribe

5.3.3 Intervista in dialetto stretto

Per questo ultimo esempio, l'IA è stata sottoposta ad un compito piuttosto estremo, cioè quello di identificare, al meglio delle proprie capacità, le parole pronunciate da un gruppo di anziani pugliesi intervistati riguardo alle preparazioni per il cenone di Natale a Bari. Come prima cosa, è stato chiesto a Transcribe di identificare da sé la lingua parlata nel video; una volta finita l'elaborazione, la lingua che lo strumento di AWS pensava di aver sentito era il francese, quando, chiaramente, non è questo il caso.



Figura 5.8: Spezzione dal video in dialetto barese con annessi sottotitoli prodotti da Transcribe

Durante tutta la prima parte del video, l'IA non è riuscita ad estrapolare neanche una parola dal video, corretta o meno che fosse, mentre, per quanto riguarda il contenuto restante, ogni parola scelta da Transcribe è un qualche termine in lingua francese che assomiglia vagamente all'espressione dialettale utilizzata, ma che spesso non c'entra niente, come nell'esempio della Figura 5.8, dove la signora nel video pronuncia la parola "castagnelle", interpretata male dal modello come "Castagner les".

Discussione sulle esperienze condotte

In questa sezione ci soffermeremo a discutere riguardo ai risultati delle analisi dei casi d'uso effettuate sugli strumenti di Intelligenza Artificiale messi a disposizione da Amazon Web Services, parlando dei loro punti di forza, punti di debolezza e sui prospetti futuri di questa tecnologia.

6.1 Discussione sul riconoscimento di testi

Il riconoscimento di testi all'interno delle immagini è uno dei molti modi per sfruttare la tecnologia di elaborazione delle immagini basata sull'Intelligenza Artificiale. Una tecnologia di questo tipo inizia ad acquisire sempre più importanza all'interno della realtà odierna, ed è applicabile ad una grande quantità di situazioni possibili, come l'automazione di compiti ripetitivi, la sicurezza, il supporto ai clienti ed altro ancora. Questa tecnologia, tuttavia, non è limitata all'uso da parte di grandi aziende; grazie a servizi come Amazon Web Services, chiunque può usufruire di strumenti di Intelligenza Artificiale, anche senza possedere la potenza di calcolo necessaria per addestrare i modelli o mantenere servizi di questo tipo. In questo modo, anche i meno avvezzi alla tecnologia possono provare ad utilizzare gli strumenti basati sull'IA senza troppa difficoltà. Nonostante ciò, l'utilità più significativa della piattaforma di Amazon è sicuramente quella offerta ad aziende con dataset di notevoli dimensioni, considerato quanto sia facile addestrare un proprio modello tramite gli strumenti forniti da Amazon Web Services. Durante i test che abbiamo effettuato utilizzando gli strumenti di AWS per il riconoscimento dei testi nelle immagini, sono trapelate alcune lacune presenti nelle demo messe a disposizione da Amazon Rekognition, probabilmente dovute alla natura stessa di questi test, cioè proprio quella di "Demo". Non per niente, una delle funzionalità più usate di Amazon Rekognition è proprio quella delle Custom Labels, cioè lo strumento che permette agli utenti di creare un modello per il riconoscimento di testi altamente specializzato, così da aumentare la precisione dello strumento di AWS nei casi specifici richiesti dal singolo utilizzatore. Un esempio potrebbe essere quello dei cartelli stradali affrontato nella sezione precedente, il quale ci dimostra che si può facilmente integrare con altri strumenti, anche quelli non basati sull'Intelligenza Artificiale, data la semplicità della struttura delle risposte fornite da Rekognition, cioè dei semplici file JSON. Il futuro sembra sempre più roseo per questo tipo di strumenti, visto e considerato l'enorme interesse che la tecnologia dell'Intelligenza Artificiale ha ottenuto in questi ultimi anni, e si spera che lo sviluppo in questo campo riesca a continuare senza grossi ostacoli per quanto più tempo possibile.

6.2 Discussione sul riconoscimento di volti

Il riconoscimento dei volti umani è una delle branche maggiori della tecnologia di riconoscimento delle immagini basata sull'IA, e non è difficile capirne il perché; La capacità di una macchina di riconoscere, elaborare e compiere operazioni complesse a partire da immagini di volti umani completamente in autonomia può avere risvolti importantissimi per il corso della storia umana, come stanno dimostrando gli utilizzi di questi strumenti che già si stanno applicando oggi. Un'ipotetica macchina allenata su una buona fetta della popolazione mondiale sarebbe in grado di eseguire tantissime operazioni senza il minimo input da parte di operatori umani, sia in positivo che in negativo: infatti per ogni porta ad apertura automatica a riconoscimento facciale, ci può sempre essere un avanzatissimo sistema di sorveglianza messo in atto da individui dalla dubbia morale. A causa di queste preoccupazioni ci sono diversi movimenti sociali che tentano di opporsi allo sviluppo dell'Intelligenza Artificiale, temendo il peggio per l'umanità, nel caso questi strumenti dovessero finire nelle mani sbagliate. Oltre a ciò, tantissime persone hanno anche preoccupazioni su scala minore, ma comunque estremamente rilevanti, che spaziano dalla privacy alla sicurezza personale, fino al semplice malfunzionamento di sistemi che si basano completamente su tecnologie come questa. Detto questo, per quanto abbiamo potuto constatare dalle demo di Rekognition forniteci da Amazon Web Services, il riconoscimento dei volti con l'Intelligenza Artificiale procede in un'ottima direzione, sia dal punto di vista dello sviluppo, che da quello dei suoi utilizzi non negativi. Come per tutti gli strumenti di questo tipo, esso è in grado di gestire i casi più generali in maniera a dir poco eccellente, senza però evitare qualche problema, dovuto ad alcuni utilizzi considerabili più "di nicchia", ma in situazioni di quel tipo si dovrebbe procedere con un modello specializzato. Il riconoscimento dei volti umani è davvero sorprendente, e non ci rimane altro che sperare che tutte le predizioni negative riguardo questo tipo di strumenti non si avverino in futuro.

6.3 Discussione sulla trascrizione da video a testo

La trascrizione delle tracce audio di un video in testo permette all'Intelligenza Artificiale di riconoscere dei pattern vocali nelle voci umane per associarli a parole, le quali verranno poi trascritte in formato testuale. Questo utilizzo della tecnologia dell'IA ha una miriade di applicazioni pratiche molto immediate, tra cui la più notevole è, sicuramente, quella dei sottotitoli autogenerati. Piattaforme come YouTube adottano questo tipo di strumenti ormai da diversi anni, per quanto l'efficacia di questi sottotitoli fosse stata molto limitata in passato, dato che non erano ancora presenti degli strumenti abbastanza avanzati da permettere un perfetto riconoscimento della voce umana da parte di macchine. Strumenti di questo genere potrebbero essere utilizzati come strumenti di accessibilità utilissimi per persone non-udenti, oppure anche come un aiuto per persone che vorrebbero imparare la lingua parlata all'interno del video ed hanno bisogno del supporto dei sottotitoli per farlo. Durante i test effettuati con Amazon Transcribe, è stato subito evidente quanto fosse semplice utilizzare questi strumenti per produrre sottotitoli quantomeno indicativi del contenuto presente all'interno del video, grazie anche all'interfaccia user-friendly fornita da Amazon Web Services e dalla moltitudine di opzioni fra le quali scegliere per ottenere il risultato desiderato. Ovviamente, un modello generale come quello offerto da AWS si presta molto bene a situazioni dove gli interlocutori parlano in maniera chiara, ma mostra le sue debolezze quando si iniziano ad introdurre accenti pesanti e dialetti stretti, come nell'ultimo esempio preso in esame in precedenza. Oltre ai sottotitoli, la trascrizione da audio a testo potrebbe essere anche usata in ambito accademico per l'archiviazione di conferenze o altri interventi in diretta in formato testuale, rendendoli più scorrevoli e facili da analizzare, oppure anche per elaborare enormi moli di

dati in formato audio in maniera estremamente più veloce ed efficiente, specialmente se si è alla ricerca di termini chiave al loro interno. Qualsiasi sia l'utilizzo scelto, l'evoluzione che gli strumenti di trascrizione basati sull'IA possono portare nei relativi campi di applicazione rimangono notevoli, e potrebbero permettere un grande salto in avanti su più fronti.

Lo scopo di questa tesi è stato quello di analizzare la tecnologia dell'Intelligenza Artificiale, cioè la capacità di una macchina di replicare il pensiero umano, oltre a testarne le capacità tramite l'utilizzo di una piattaforma di cloud computing predisposta per questo tipo di attività. Abbiamo definito cosa si intende per "Intelligenza Artificiale", con diversi esempi degli utilizzi che se ne possono fare, oltre ad alcune menzioni di IA famose nella cultura popolare, a riprova di quanto l'IA sia diventata fondamentale nel corso degli anni, anche se nel campo della finzione.

Successivamente abbiamo parlato della storia di questa tecnologia, dei suoi albori, degli obiettivi che si è posta fin dall'inizio, trattando anche degli ostacoli incontrati lungo il cammino, e infine dei numerosi progressi ed evoluzioni, fino ad arrivare a quello che l'IA può fare al giorno d'oggi.

Nella seconda parte abbiamo parlato del mondo del cloud computing e delle sue capacità nello sviluppo e nell'utilizzo dell'Intelligenza Artificiale, soffermandoci, in particolare, sul funzionamento della piattaforma Amazon Web Services, sulla sua divisione in regioni e in zone di disponibilità, oltre che sugli innumerevoli vantaggi che la piattaforma offre ai propri utenti, in particolare alle aziende, le quali sono in grado di tagliare i costi e di semplificare la gestione di alcuni aspetti più ostici dell'IT in maniera considerevole. Oltre a questo, abbiamo trattato dei vari strumenti offerti dalla piattaforma, soffermandoci su quelli più importanti. Nei capitoli finali, abbiamo messo in pratica le tecnologie offerte da AWS per provarne la loro efficacia in casi d'uso reali; ognuno dei tre strumenti che abbiamo utilizzato si è dimostrato più che capace nel gestire casi di tipo più generale, mentre ha avuto più difficoltà quando si è trattato di scendere in casi più specifici, per quanto questo fosse comprensibile, dato che i modelli utilizzati sono proprio stati pensati per essere general-purpose.

Gli strumenti basati sull'Intelligenza Artificiale non finiranno di evolversi; con la loro adozione da parte di un numero sempre maggiore di aziende, gli investimenti nella ricerca e lo sviluppo dell'IA non accennano ad arrestarsi. Secondo *Artificial Intelligence Report 2023* [10] HAI ("Human-Centered Artificial Intelligence") dell'Università di Stanford, questa nuova "era di deployment" delle IA ha portato un investimento privato in questa tecnologia di 91.9 miliardi di dollari solo nel 2022, un numero 18 volte più alto di quello registrato nel 2013. Oltre a questo, lo stesso report dell'università ci parla di come le IA rappresentino sia un grosso beneficio verso certe categorie di lavoratori, che un grosso problema per altre, dato che molti lavoratori temono di perdere il lavoro a causa dell'avanzare di questi strumenti, così come si è potuto vedere con gli scioperi degli scrittori della *Writer's Guild of America*. Il futuro delle IA è fatto di luci ed ombre, e saranno solo gli umani a decidere in quale delle due direzioni esso virerà.

- [1] ASIMOV, I. (1950), *I, Robot*. (Cited at page 5)
- [2] BABBAGE, C. (1982), «On the Mathematical Powers of the Calculating Engine», *The Origins of Digital Computers. Texts and Monographs in Computer Science*. (Cited at page 7)
- [3] BODEN, M. A. (2018), *Artificial Intelligence: A Very Short Introduction*.
- [4] CAMPBELL, M., HOANE, A. J. e HSIUNG HSU, F. (2002), «Deep Blue», *Artificial Intelligence*, vol. 134, p. 57–83.
- [5] CLARKE, A. C. e KUBRICK, S. (1968), *2001: A Space Odyssey*. (Cited at page 5)
- [6] DICK, P. K. (1968), *Do Androids Dream of Electric Sheep?* (Cited at page 5)
- [7] HASHEMIPOUR, S. e ALI, M. (2020), «Amazon Web Services (AWS) – An Overview of the On-Demand Cloud Computing Platform», .
- [8] KRIZHEVSKY, A., SUTSKEVER, I. e HINTON, G. E. (2012), «ImageNet classification with deep convolutional neural networks», *Communications of the ACM*, vol. 60, p. 84 – 90.
- [9] LECUN, Y., BENGIO, Y. e HINTON, G. E. (2015), «Deep Learning», *Nature*, vol. 521, p. 436–444.
- [10] MASLEJ, N., FATTORINI, L., BRYNJOLFSSON, E., ETCHEMENDY, J., LIGETT, K., LYONS, T., MANYIKA, J., NGO, H., NIEBLES, J. C., PARLI, V., SHOHAM, Y., WALD, R., CLARK, J., e PERRAULT, R. (2023), «The AI Index 2023 Annual Report», Rap. tecn., AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. (Cited at page 39)
- [11] MCCARTHY, J. (1960), «Programs with common sense», . (Cited at page 8)
- [12] MCCARTHY, J., MINSKY, M., ROCHESTER, N. e SHANNON, C. E. (2006), «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955», *AI Mag.*, vol. 27, p. 12–14.
- [13] RAO, A. S. e VERWEIJ, G. (2017), «Sizing the prize: what’s the real value of AI for your business and how can you capitalise?», . (Cited at page 1)
- [14] RUSSELL, B. e WHITEHEAD, A. N. (1956), «Principia Mathematica to *56», . (Cited at page 7)

- [15] RUSSELL, S. e NORVIG, P. (1995), «Artificial intelligence - a modern approach: the intelligent agent book», .
- [16] TURING, A. (1937), «On computable numbers, with an application to the Entscheidungsproblem», *Proc. London Math. Soc.* (Cited at page 7)
- [17] VASWANI, A., SHAZEER, N. M., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, L. e POLOSUKHIN, I. (2017), «Attention is All you Need», in «NIPS», .

Siti Web consultati

- Wikipedia – www.wikipedia.org
- Treccani – www.treccani.it
- Amazon Web Services – <https://aws.amazon.com>
- Google Cloud – <https://cloud.google.com>
- Microsoft Azure – <http://azure.microsoft.com>
- IBM – www.ibm.com
- Artificial Intelligence Index – <https://aiindex.stanford.edu>
- PwC – www.pwc.com
- Google DeepMind – www.deepmind.com
- Google Colab – <https://colab.research.google.com>
- OpenAI – <https://openai.com>
- Midjourney – www.midjourney.com
- Stability AI – <https://stability.ai>

Ringraziamenti

Sinceramente non credo che questa sezione sia particolarmente necessaria, e non perché non sia grato a nessuno, ma solo perché non credo di aver fatto qualcosa di così notevole da meritarsi un paragrafo di questo genere; tuttavia, pare che questo sia il protocollo e, tutto sommato, non ho molti altri motivi per rifiutarmi di farlo.

Ringrazio in primis i miei genitori, mia madre Carmen e mio padre Pierluigi, per avermi supportato durante tutto questo percorso, sia economicamente che dal lato dei trasporti.

Ringrazio anche tutti i compagni di studio che ho avuto il piacere di conoscere durante la mia permanenza all'Università di Ancona: Gabriel, Tosca, Luca, Federico, Michele, Nicola, Giorgia, Giulia, Leonardo, Lorenzo, Omar e Matteo (mi sarò sicuramente dimenticato qualche nome).

In particolare, rivolgo un ringraziamento speciale ai miei due compaesani con cui ho confabulato per la maggior parte del mio percorso triennale: parlo ovviamente di Francesco ed Amir; dire che siete stati cruciali al traguardo che ho ottenuto sarebbe un eufemismo.

Un altro ringraziamento esclusivo va rivolto a tre entità esterne a questa facoltà, cioè Nicola O, Matteo G e Matteo S, i quali mi hanno sopportato durante le mie fastidiose tirate riguardo qualsiasi ambito tecnologico verso il quale fossi particolarmente adirato in quel momento (si meritano una menzione particolare la mia scheda video e il mio portatile con Arch Linux).

Vorrei anche ringraziare Larian Studios per aver rilasciato Baldur's Gate 3 in un periodo particolarmente azzeccato della mia vita, così da costringermi a praticare costanza e disciplina, lavorando a questa tesi invece che giocando al loro gioco, e anche per avermi fatto da sollazzo quando il lavoro iniziava a diventare particolarmente pesante. Ah, e anche per aver creato un capolavoro.

Ringrazio il prof. Ursino per avermi guidato attraverso il percorso della mia tesi, risolvendo qualunque dubbio, qualora ne avessi, ed aiutandomi a comporre gli aspetti più cruciali di questo stesso documento.

Detto questo, rivolgo un ultimo ringraziamento a tutti quelli che non ho menzionato in questo passaggio, ma che hanno contribuito in qualche modo al mio percorso universitario, anche se in minima parte.

Grazie a tutti!