

Università degli Studi di Padova

DIPARTIMENTO DI MATEMATICA “TULLIO LEVI-CIVITA”

CORSO DI LAUREA IN INFORMATICA



Utilizzo di Deep Learning per il rilevamento di
anomalie elettrocardiografiche

Tesi di Laurea

Relatore

Prof. Tullio Vardanega

Laureando

Oscar Konieczny

Matricola 2042335

ANNO ACCADEMICO 2023-2024

“I disagree strongly with whatever work this quote is attached to.”

— Randall Munroe

Ringraziamenti

/

Padova, Dicembre 2024

Oscar Konieczny

Sommario

/

Indice

1	Contesto aziendale	1
1.1	L'azienda	1
1.2	Tipologia di clientela	1
1.3	Organizzazione aziendale	2
1.4	Processi aziendali	3
1.5	Tecnologie e strumenti	4
1.5.1	Tecnologie utilizzate	4
1.5.2	Strumenti di supporto ai processi	5
1.6	Propensione all'innovazione	8
2	<i>Stage</i> proposto	9
2.1	Gestione aziendale degli <i>stage</i>	9
2.2	Descrizione progetto	10
2.3	Obiettivi	11
2.4	Vincoli	12
2.4.1	Vincoli tecnologici	12
2.4.2	Vincoli temporali	12
2.4.3	Vincoli organizzativi	12
2.5	Motivazione della scelta	13
3	Svolgimento dello <i>stage</i>	14
3.1	Pianificazione delle attività	14
3.2	Analisi dei requisiti	14
3.3	Progettazione	14
3.4	Codifica	14

3.5	Verifica e validazione	14
3.6	Risultati raggiunti	15
3.6.1	Piano qualitativo	15
3.6.2	Piano quantitativo	15
4	Retrospettiva	i
4.1	Conseguimento degli obiettivi	i
4.1.1	Obiettivi aziendali	i
4.1.2	Obiettivi personali	i
4.2	Competenze acquisite	i
4.3	Divario tra università e lavoro	i
	Glossario	ii
	Bibliografia	iv
	Sitografia	v

Elenco delle figure

1.1	Rappresentazione del <i>framework</i> Scrum utilizzato in azienda . .	4
1.2	Esempio di gestione delle attività con <i>Jira</i>	5
1.3	<i>Screenshot</i> di una <i>repository</i> in <i>Bitbucket</i>	6
1.4	Pagina principale del <i>Confluence</i> di M31 <i>Academy</i>	7
1.5	<i>Screenshot</i> di un canale di comunicazione di M31 <i>Academy</i> . . .	7
2.1	Rappresentazione dei punti principali del progetto	10

Elenco delle tabelle

2.1	Lista dei vari obiettivi aziendali.	11
-----	---	----

Elenco dei codici sorgenti

Capitolo 1

Contesto aziendale

1.1 L'azienda

M31 S.r.l. è un'azienda italiana nata nel 2007 con sede a Padova. Essa si specializza in ingegneria all'avanguardia su una variegata gamma di ambiti, come ad esempio: *mechanical & robotics, software engineering, artificial intelligence, IoT* e molto altro. L'azienda affianca i suoi clienti con cui svolge progetti, partendo dalla fase esplorativa fino alla fase di industrializzazione e certificazione. Nei diciassette anni di operazione, M31 ha realizzato più di cento unici progetti di innovazione che hanno avvantaggiato proprie startup, imprese del territorio e grandi gruppi.

1.2 Tipologia di clientela

La clientela di M31 è abbastanza variegata, ma i settori in cui opera e ha operato sono:

- **Biomedicina:** per la creazione di strumenti diagnostici, piattaforme di *liquid handling* e sistemi di monitoraggio remoto.
- **Monitoraggio e Sicurezza:** sistemi che variano da monitoraggio, antintrusione, videosorveglianza e sensori di allarme.
- **Applicazioni industriali:** per l'evoluzione di processi produttivi, in ambiti come quello delle lavorazioni meccaniche o tessili.

- ***Automation & smart cities:*** per l'automatizzazione di processi in vari ambienti, che siano domestico, di assistenza sanitaria oppure di processi fotografici.
- ***Cloud & digital twins:*** lo sviluppo di una controparte digitale di un prodotto o processo sul *cloud*.
- ***Racing & automotive:*** nell'ambito di vari sistemi e della telemetria per vari sport.

1.3 Organizzazione aziendale

L'organizzazione aziendale di M31 è suddivisa in diversi settori, questi sono: *governance*, *R&D team*, *marketing & sales* e *procurement & supply chain*.

Governance

Questo settore si occupa di amministrare l'azienda, questo avviene in vari modi: presa di decisioni per la direzione dell'azienda, gestione e manutenzione delle risorse finanziarie e la gestione e organizzazione del capitale umano.

I ruoli che possiamo trovare in questo ramo sono principalmente: il *CEO*, anche chiamato amministratore delegato, *accounting specialist* e *HR manager*.

Marketing & sales

Questo è il settore che si occupa di aumentare la visibilità e rapporti con possibili clienti, ma anche di ricercare quali sono le richieste attuali del mercato.

Questo *team* non è relativamente grande, nonostante ciò è possibile notare la presenza del *CBO*, anche detto direttore commerciale in lingua italiana.

Procurement & supply chain

Questa area aziendale si occupa di far sì che le operazioni aziendali possano procedere in modo adeguato, garantendo l'approvvigionamento delle risorse necessarie e la gestione della catena di fornitura.

R&D team

Infine, passiamo al settore aziendale dove sono stato inserito, ovvero il reparto di ricerca e sviluppo. Questo settore è quello più grande dell'azienda, ed è qui che i prodotti effettivi vengono sviluppati.

Ci sono varie aree di specializzazione in questo comparto, esse sono divise in una sorta di gruppi di lavoro, che possono o meno coordinarsi per lavorare insieme ai progetti aziendali attualmente in sviluppo.

Personalmente sono stato inserito sotto al *team di computer vision* e, come ogni altro stagista, anche sotto allo *stage coordinator*.

1.4 Processi aziendali

Durante il mio periodo di *stage* ho potuto assistere a come vengono gestiti i progetti all'interno dell'azienda. Viene utilizzata la metodologia *Agile*, ma in particolare il *framework* Scrum. Attraverso questo *framework*, i progetti vengono suddivisi in varie iterazioni denominati *sprint*, ad ognuno di esse, viene assegnata una certa durata di tempo, che può variare da una o più settimane.

Ogni *sprint* inizia con un *sprint planning*, nel quale vengono definite le attività da svolgere nello stesso *sprint*, queste attività vengono assegnate ai vari membri dello specifico *team* che si occupa di quel progetto.

All'inizio di ogni giornata lavorativa, viene eseguito il *daily stand-up*, nel quale viene discusso il procedimento dei lavori; in particolare viene discusso quello che è stato fatto la giornata precedente e cosa verrà fatto nella giornata attuale.

A fine di ogni *sprint* vengono fatti lo *sprint review* e lo *sprint retrospective*, il primo ha la finalità di valutare se tutte le attività sono state svolte e che siano state fatte correttamente, mentre il secondo considera le problematiche riscontrate durante lo *sprint*, definendo possibili miglioramenti al *way of working*, nel caso ne fosse necessario.

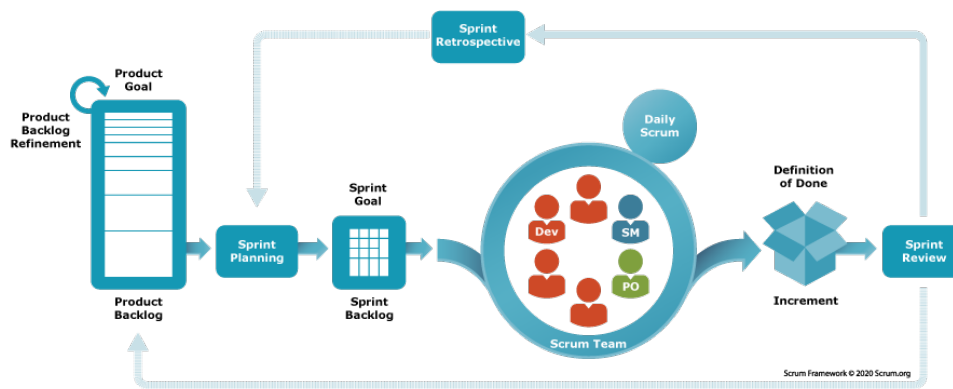


Figura 1.1: Rappresentazione del *framework* Scrum utilizzato in azienda
Fonte: [scrum.org](https://www.scrum.org)

1.5 Tecnologie e strumenti

1.5.1 Tecnologie utilizzate

In azienda vengono usate svariate tecnologie, questo è dovuto principalmente dal fatto che vengono svolti più progetti allo stesso tempo. Non avendo familiarità sugli altri progetti, esporrò le tecnologie che furono a me più vicine.

Python

Python è un linguaggio di programmazione ad alto livello, interpretato e di uso generale. È attualmente il linguaggio più popolare¹ al mondo, questo è dovuto dalla semplicità di utilizzo e dal buon supporto di librerie di terze parti, che semplificano ambiti come ad esempio: *web development*, *data science*, *machine learning*_G e *scripting* di operazioni ripetitive. Perciò non stupisce l'utilizzo di questo linguaggio per l'ambito di *machine learning* e per manipolare i *dataset*.

TensorFlow

TensorFlow è una libreria *open-source* sviluppata da *Google* per il *machine learning* e il *deep learning*_G. È una delle librerie più popolari per la creazione e adde-

¹Fonte: <https://www.tiobe.com/tiobe-index/>

stramento di modelli basati su tecniche di *machine learning*. *TensorFlow* integra al suo interno librerie per sfruttare a pieno processori specifici, come ad esempio le *GPU* e le *TPU*, permettendo così di utilizzare risorse computazionali più adeguate, per questo ambito, in modo semplice.

Keras

Keras è un'altra libreria *open-source* per il *deep learning*, sviluppata per essere semplice e modulare. Di recente è stata riscritta basandosi su *TensorFlow* e, come quest'ultima, permette di creare e addestrare modelli di intelligenza artificiale, ma rende il processo più rapido e intuitivo. Keras supporta una variegata tipologia di reti neurali pronte all'uso, come ad esempio reti neurali convoluzionali e ricorrenti, lasciando comunque la possibilità di modificare qualsiasi cosa si voglia adattare meglio alle proprie esigenze.

1.5.2 Strumenti di supporto ai processi

Jira

Jira è un software proprietario sviluppato da *Atlassian*, è utilizzato come Issue Tracking System (ITS), viene utilizzato per gestire le attività dei vari progetti con la metodologia *Agile*. All'interno di questo software si possono dedicare delle *board* per i vari *sprint* che vengono svolti.

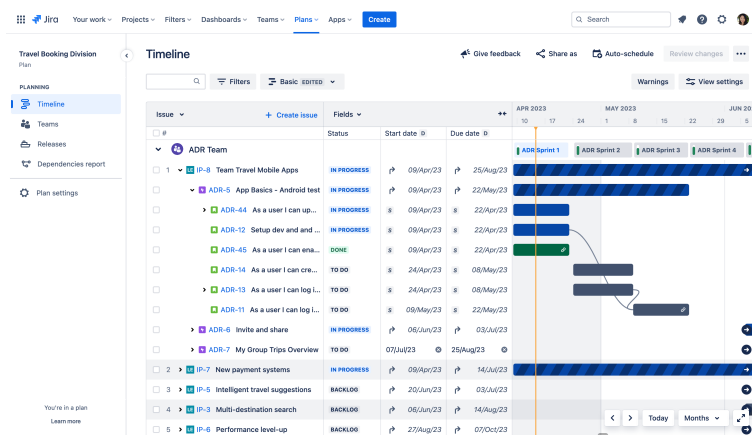



Figura 1.2: Esempio di gestione delle attività con *Jira*

Fonte: atlassian.com

Bitbucket

Un'altro software nella suite di *Atlassian* è *Bitbucket*, che viene utilizzato per il controllo di versione e per la gestione del codice sorgente. Con questo software è possibile creare *repository*, questo spazio permette ai *team* di collaborare in un modo efficiente, avendo anche la disponibilità di utilizzare funzionalità come le *pull request* .

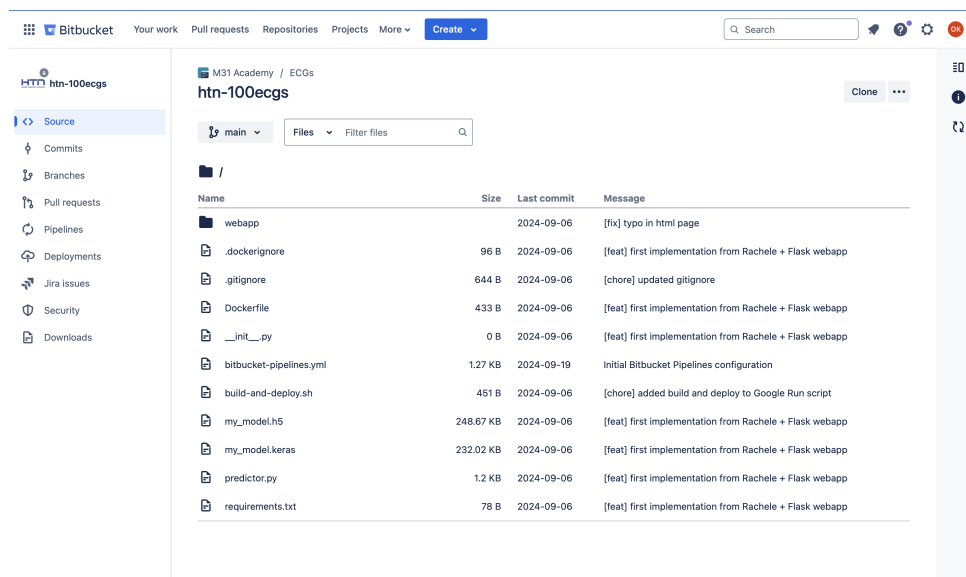


Figura 1.3: Screenshot di una *repository* in *Bitbucket*

Confluence

Confluence è un'altro software sviluppato da *Atlassian* e viene utilizzato per la gestione della documentazione. Facendo parte della *suite* di *Atlassian* è ben integrato con le tecnologie precedentemente espone. All'interno di questo spazio vengono create documentazioni di prodotti e vari documenti contenenti risorse utili anche internamente.

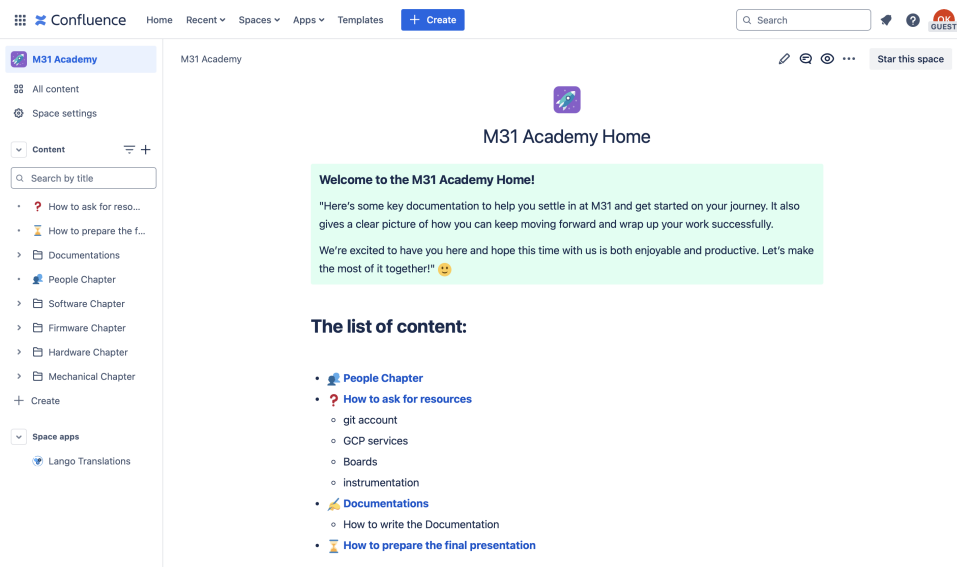


Figura 1.4: Pagina principale del *Confluence* di *M31 Academy*

Microsoft Teams

L'applicativo che viene utilizzato per comunicare è *Microsoft Teams*. Questo software sviluppato da *Microsoft* permette di comunicare, con singoli individui o con gruppi di persone, sia tramite messaggi testuali che tramite videochiamate. Può essere usato anche per trasferimenti di piccoli file, ed è la prima opzione per comunicare con i membri dei team che lavorano da remoto.

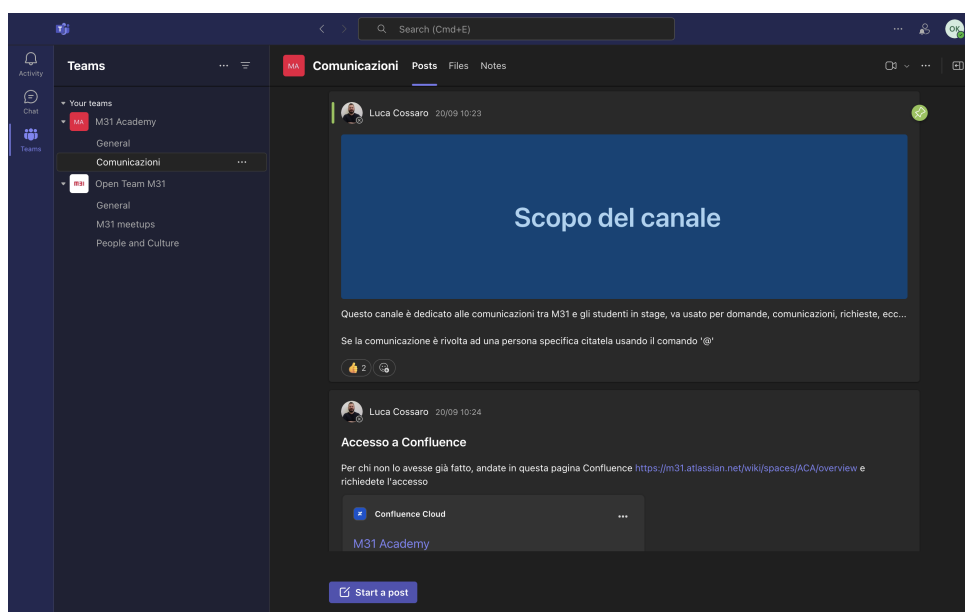


Figura 1.5: *Screenshot* di un canale di comunicazione di *M31 Academy*

1.6 Propensione all'innovazione

M31 si interessa parecchio all'innovazione e questo lo si può vedere dalla filosofia con cui intraprende nuovi progetti di lavoro, come già menzionato nella sezione introduttiva [1.1](#), i progetti che svolge sono tutti con lo scopo di portare progresso nei settori in cui si specializza. Inoltre, durante lo svolgimento del mio tirocinio, ho potuto assistere ad un *meeting* plenario che viene svolto periodicamente, nel quale ho potuto vedere in quali settori l'azienda è intenzionata ad esplorare in futuro, sia dal lato della dirigenza che dal lato dei dipendenti.

Capitolo 2

Stage proposto

2.1 Gestione aziendale degli *stage*

M31 mostra molto interesse verso la creazione di rapporti con le nuove generazioni, questo avviene in diversi modi: tramite la *M31 Academy* e tramite la partecipazione ai progetti di “Ingegneria del *Software*” della laurea di informatica, come azienda proponente.

M31 Academy è una sorta di settore aziendale dedicato alla realizzazione di progetti con studenti e studentesse universitarie o neo-laureati e neo-laureate. Ora come ora, l'azienda sta svolgendo progetti solo con studenti e studentesse universitarie, che prendono il ruolo di stagisti e stagiste, perciò passo a descrivere meglio come essi vengono gestiti.

L'influsso di stagisti e stagiste è principalmente dovuto alla partecipazione dell'azienda allo STAGE-IT, un evento promosso da Confindustria Veneto Est e l'Università di Padova, dove studenti e studentesse vengono introdotti a tre progetti che possono svolgere in azienda. M31 però offre una lista più numerosa di progetti, che vengono esposti quando gli studenti e studentesse contattano direttamente l'azienda. Questi progetti, essendo numerosi, possono essere di vario tipo. Alcuni pongono gli stagisti e stagiste in un progetto su cui l'azienda sta già lavorando. Mentre altri, come quello che ho scelto io, servono all'azienda per esplorare nuove tecnologie o per prepararsi per un progetto aziendale che non ha ancora avuto modo di iniziare.

2.2 Descrizione progetto

Il progetto di *stage* consiste nella realizzazione di un'applicazione per l'analisi di dati tramite tecniche di *machine learning_G* e di *deep learning_G*; in particolare, per l'analisi di segnali elettrocardiografici, tramite elettrocardiogrammi.

Questo progetto di *stage* serve all'azienda per prepararsi ad un progetto che si aspettano di affrontare in futuro, collaborando insieme ad un'azienda con cui sono attualmente in comunicazione.

Questo progetto vede come *focus* principale la gestione di grandi quantità di dati, sotto forma di migliaia di *file* contenenti segnali elettrocardiografici e altre informazioni sui pazienti. Questi dati, facenti parte di un *dataset_G*, vengono utilizzati all'interno di un programma per addestrare *reti neurali artificiali_G*, ovvero di modelli che cercano di replicare il funzionamento del cervello umano. Questi dati non possono subire modifiche inaspettate quando vengono inseriti nel processo di addestramento. Inoltre, questi dati devono essere gestiti in modo efficiente, in modo da non caricare tutto all'interno della *RAM_G*, come viene generalmente fatto nel caso di *dataset* di piccole dimensioni, così da non subire limitazioni di *hardware*.

In secondo luogo, il progetto vede anche l'effettivo addestramento e valutazione di vari tipi di reti neurali artificiali, sia con lo scopo di testare, che di esplorare quali sono le tecniche migliori per l'analisi di elettrocardiogrammi.

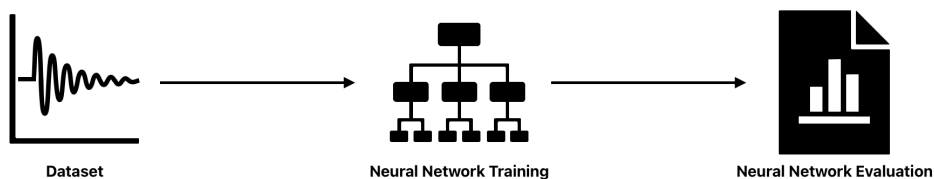


Figura 2.1: Rappresentazione dei punti principali del progetto

2.3 Obiettivi

Il progetto di *stage* era definito da alcuni obiettivi richiesti dall'azienda, ognuno di essi era contrassegnato con una certa importanza; questa può rientrare in una dei tre seguenti tipi:

- **Obbligatorio:** indica un obiettivo primario richiesto dal committente, il suo completamento non è negoziabile
- **Desiderabile:** indica un obiettivo non vincolante o strettamente necessario, ma dal riconoscibile valore aggiunto
- **Facoltativo:** indica un obiettivo facoltativo, che rappresenta valore aggiunto e non strettamente competitivo

Nella tabella 2.1, definita successivamente, vengono esposti i vari obiettivi aziendali che sono stati imposti per il completamento del progetto:

Obiettivo	Importanza
Studio delle tecnologie e del contesto specifico	Obbligatorio
Creazione di un dataset appropriato e di tecniche per la sua gestione	Obbligatorio
Implementazione in <i>Python</i> degli algoritmi basati su tecniche di <i>deep learning</i>	Obbligatorio
Valutazione dell'esecuzione degli algoritmi	Obbligatorio
Realizzazione di <i>unit test</i>	Desiderabile
Idee o suggerimenti su come migliorare in futuro le prestazioni dell'applicativo'	Facoltativo

Tabella 2.1: Lista dei vari obiettivi aziendali.

2.4 Vincoli

2.4.1 Vincoli tecnologici

Per i vincoli tecnologici, alcune delle tecnologie mi sono state imposte per la realizzazione di questo progetto. Queste tecnologie sono relative all'addestramento dei modelli di intelligenza artificiale, nella quale Keras era la tecnologia che dovevo usare. Utilizzando questa tecnologia, però, si è inoltre obbligati a fare uso di *Python* e, per aumentare la semplicità di installazione dei pacchetti di questo linguaggio di programmazione, si è imposto anche l'uso di *Conda*, un sistema per la gestione di pacchetti e *environment*_G, che ha permesso l'installazione di ciò che era necessario senza troppi intoppi.

2.4.2 Vincoli temporali

Per la sua natura, il tirocinio curricolare ha un limite di tempo: ovvero deve essere dalle 300 alle 320 ore di lavoro, e non è possibile sforare 40 ore settimanalmente. Per lavorare al progetto, lo *stage* è stato distribuito in giornate lavorative da 8 ore ciascuno, per un totale di all'incirca 38 giorni.

2.4.3 Vincoli organizzativi

Per il corretto svolgimento del tirocinio curricolare, ho preso, insieme a alle figure che mi supervisionavano, alcuni accorgimenti.

In primis, periodicamente vengono svolti degli incontri di *sprint review* con il mio *tutor* aziendale, nei quali vengono esposte le attività svolte da me durante l'ultimo periodo, in modo da far sì che l'avanzamento del progetto si allineasse con gli obiettivi imposti in modo corretto.

In secundis, ogni cinque giorni lavorativi dovevo mandare una comunicazione al mio relatore di tesi, indicando quali attività erano previste da svolgere durante la settimana lavorativa passata, e come esse sono state svolte rispetto alle aspettative. Questo per monitorare il corretto proseguimento dello *stage*.

2.5 Motivazione della scelta

Sono venuto a conoscenza di M31 attraverso l'evento STAGE-IT, un'iniziativa che mi ha dato la possibilità di incontrare, in un singolo posto, diverse aziende e le loro proposte di *stage*. Tra le aziende presenti, M31 era una di quelle che mi ha interessato di più, questo perchè avevo un maggiore interesse a svolgere un progetto che utilizzava, in qualche maniera, l'intelligenza artificiale.

M31 durante l'evento STAGE-IT aveva proposto un tema di *machine learning* in ambito di *computer vision*_C, che ero intenzionato nel svolgere quando stavo contattando l'azienda, ma vedendo la lista aggiornata dei progetti di *stage* offerti, ho deciso di scegliere un secondo progetto, sempre di intelligenza artificiale, ma in ambito biomedico, questo perchè le tecniche utilizzate per questo specifico ambito mi sarebbero state nuove.

Gli obiettivi personali che mi sono posto prima di svolgere questo tirocinio sono i seguenti:

- **Approfondimento tecnico:** avendo parecchio interesse verso l'intelligenza artificiale, con questo *stage*, desideravo approfondire questo argomento svolgendo un progetto con un'applicazione in questo ambito, per me nuova. Questo obiettivo riflette anche la mia intenzione di proseguire gli studi con la laurea magistrale con indirizzamento all'intelligenza artificiale.
- **Capire l'ambiente lavorativo:** essendo la mia prima esperienza lavorativa in questo settore, ero incuriosito da come, effettivamente, le aziende operano internamente. Non solo apprendere le metodologie aziendali, ma anche la cultura lavorativa.
- ***Problem solving*:** migliorare nell'abilità di risolvere problemi, in modo da trovare soluzioni efficienti a vari problemi a diverse sfide.

Capitolo 3

Svolgimento dello *stage*

3.1 Pianificazione delle attività

Esposizione di come sono state pianificate le attività.

3.2 Analisi dei requisiti

Descrizione dell'attività di analisi dei requisiti.

3.3 Progettazione

Descrizione dell'attività di progettazione.

3.4 Codifica

Descrizione dell'attività di codifica.

3.5 Verifica e validazione

Descrizione delle attività di verifica e validazione.

3.6 Risultati raggiunti

3.6.1 Piano qualitativo

Esposizione dei risultati ottenuti sul piano qualitativo, dove viene visto il prodotto in una visione d'insieme.

3.6.2 Piano quantitativo

Esposizione dei risultati ottenuti sul piano quantitativo, qui vengono illustrate la copertura dei requisiti e dei test.

Capitolo 4

Retrospettiva

4.1 Conseguimento degli obiettivi

4.1.1 Obiettivi aziendali

Esposizione degli obiettivi aziendali raggiunti, facendo riferimento agli obiettivi descritti nelle sezioni precedenti.

4.1.2 Obiettivi personali

Esposizione degli obiettivi personali raggiunti, facendo riferimento agli obiettivi descritti nelle sezioni precedenti.

4.2 Competenze acquisite

Descrizione della maturazione professionale, in ambito di conoscenze e abilità.

4.3 Divario tra università e lavoro

Descrizione della distanza tra le competenze imparate in università e quelle richieste allo *stage*.

Glossario

Computer Vision La *computer vision* è un campo dell'intelligenza artificiale che si occupa di sviluppare sistemi in grado di interpretare e comprendere informazioni visive provenienti da immagini e video. [13](#)

Dataset Un *dataset* è una collezione strutturata di dati, generalmente di grandi dimensioni, organizzata in forma relazionale. [10](#)

Deep Learning Con *deep learning* si intende una sotto-categoria del *machine learning*, che utilizza reti neurali artificiali che possiedono molteplici strati con lo scopo di analizzare e interpretare grandi moli di dati. [4](#), [10](#)

Environment Gli *environment*, anche detti *virtual environment*, sono spazi dedicati per l'installazione di specifici pacchetti e le loro dipendenze, senza influenzare altri progetti o l'intero sistema. [12](#)

GPU L'acronimo GPU (*Graphics Processing Unit*) indica un componente *hardware* creato specificatamente per elaborare immagini e per il rendering grafico. Sono progettate per gestire moltissime operazioni contemporaneamente, tramite l'uso di specifiche librerie è possibile sfruttare questa caratteristica per l'ambito dell'intelligenza artificiale. [5](#)

Machine Learning Con *machine learning* si intende una branca dell'intelligenza artificiale che si occupa di sviluppare algoritmi e modelli statistici per permettere ai computer di apprendere da dati e migliorare le proprie prestazioni, questo senza programmare istruzioni specifiche. [4](#), [10](#)

Pull Request Sono una funzionalità generalmente trovata nei sistemi di controllo di versione, consente agli sviluppatori di proporre modifiche ad una sorgente di codice. Questa funzionalità facilita la collaborazione, permettendo la revisione e discussione di modifiche prima che esse vengano integrate. [6](#)

RAM La *Random Access Memory* è un componente presente nei *computer* che permette di immagazzinare dati che vengono letti e scritti, all'interno di essa, in modo rapido. Questa componente è necessaria per far funzionare programmi e per processare dati. [10](#)

Rete Neurale Artificiale Una rete neurale artificiale è un modello computazionale che si ispira alla struttura e al funzionamento del cervello umano. Questi modelli sono formati da unità di base che si connettono tra di loro per formare una rete. Le unità di base utilizzate si ispirano ad una versione parecchio semplificata dei neuroni biologici. [10](#)

TPU L'acronimo TPU (*Tensor Processing Unit*) indica un componente *hardware* progettato specificatamente per il calcolo di moltissime operazioni matematiche in contemporanea. Esse sono progettate e prodotte specificatamente per essere utilizzate nell'ambito del machine learning. [5](#)

Bibliografia

Sitografia