

# Cognitive Learning by FabLab

## Scheda progetto

<b>Denominazione azienda/società</b>
Logogramma srl
<b>Titolo del progetto</b>
Automatic Speech Recognition (ASR) per Assistenti Virtuali (AV) nel dominio del commercio online
<b>Descrizione del progetto</b>
<p>Il percorso formativo proposto mira allo sviluppo di nuovi strumenti per il riconoscimento vocale automatico (Automatic Speech Recognition) e affronta, dunque, una sfida scientifica e tecnologica di riferimento internazionale. I moduli di riconoscimento vocale, infatti, sono oggi alla base di moltissime applicazioni digitali, che vanno dai comandi vocali agli assistenti virtuali.</p> <p>Il case-study che si propone si colloca nell'ambito dello sviluppo di Assistenti Virtuali (AV) di tipo task-oriented nel dominio del commercio online e in lingua italiana. In particolare, si intende implementare il modulo di ASR da sviluppare nel presente progetto in un chatbot già realizzato, al fine di testarne le performance in un ambiente di applicazione reale.</p> <p>Il modulo ASR si propone come strumento leggero e personalizzabile, mirato ad accrescere le capacità d'interazione del chatbot. L'architettura del modulo ed i modelli ivi integrati, mirano a fornire uno strumento di base versatile con performance allo stato dell'arte, offrendo la possibilità di estenderlo e personalizzarlo all'occorrenza. La realizzazione del suddetto modulo prevede una fase di sviluppo e test preliminari, una fase di sviluppo con monitoraggio continuo ed una fase finale di valutazione.</p> <p>A partire da un dataset di conversazioni già trascritte e corrette manualmente, e che comprendono scambi comunicativi tra clienti ed operatori telefonici riguardanti informazioni di tipo amministrativo, logistico e commerciale, si procederà allo sviluppo di un modello preliminare di ASR basato su tecnologie open-source che verrà testato in ambiente controllato al fine di rilevarne le performance e gli errori. Sulla base di tali test, verrà quindi migliorato il modello che sarà implementato nell'AV. Questo sarà quindi rilasciato in prova ad un gruppo di clienti, sparsi sull'intero territorio italiano, al fine di validare la soluzione proposta e di procedere alla raccolta di ulteriori dati per la specializzazione del modello finale.</p> <p>Il progetto intende mettere assieme competenze e figure professionali differenti che spaziano dagli informatici ai data scientist ai linguisti (computazionali).</p>
<b>Obiettivi formativi</b>
Il progetto mira allo sviluppo di competenze specifiche nei singoli settori dell'informatica, del data science e della linguistica, mirando parallelamente all'acquisizione, da parte dei tirocinanti, di capacità trasversali che vanno dal team-working al problem-solving, all'organizzazione del lavoro con metodologie "Agile".
<b>Sede svolgimento attività</b>
Logogramma srl - FabLab

<b>Numero di tirocinanti</b>
Due tirocinanti di durata 4 mesi (integrazione del primo bando)

<b>Data inizio indicativa</b>
Marzo 2023

<b>Competenze specialistiche</b>
Informatico: machine learning, python, Deep Neural Network frameworks, architetture distribuite e relativi paradigmi computazionali
Data Scientist: elaborazione dati, text mining, machine learning, Deep Neural Network frameworks
Linguista: linguistica teorica e computazionale, semantica distribuzionale, lessicografia

<b>Attrezzatura per lavoro collaborativo</b>
Server in cloud: macchina virtuale con utenti sulla quale installeremo i software necessari allo svolgimento del progetto.
Server messi a disposizione dal corso di laurea in DataSciences

<b>SSD di riferimento (anche più di uno)</b>
L-Lin/01, Inf/01, Secs/01

<b>Conoscenza lingue</b>
Italiano - Inglese

<b>Referente universitario</b>	<b>Tutor aziendale</b>
Prof.ssa Franco Cutugno	Azzurra Mancini

<b>Realizzabilità del progetto nello spazio FabLab (ad es. e attrezzature ed i materiali necessari sono trasportabili ed utilizzabili nei laboratori del FabLab)</b>
Il progetto è idealmente realizzabile in qualunque spazio dove sia possibile lavorare al computer ed avere una connessione Internet stabile.
Ad ogni modo il progetto sarà parte integrante delle attività della azienda ospite e troverà la sua principale collocazione nello spazio Fablab di Logogramma.

<b>Percentuale delle attrezzature e materiali necessari alla realizzazione del progetto forniti dall'azienda</b>
L'azienda si occuperà di fornire i software necessari allo sviluppo del progetto, implementandoli o sui computer dei tirocinanti oppure in una macchina virtuale in cloud, ovvero sui server accademici messi a disposizione.

<b>Ore uomo di un proprio referente messe a disposizione dall'azienda per la condivisione di know-how verso il tirocinante</b>
1 mese/uomo

## **Modularità del progetto: numero di milestone in cui è possibile articolare il progetto consentendo di ottenere risultati intermedi comunque valutabili in caso di non rispetto dei tempi**

La realizzazione del progetto prevede 3 fasi: una fase di sviluppo preliminare, una di sviluppo con monitoraggio continuo, ed una fase finale di valutazione.

### **Sviluppo preliminare**

Nella prima fase di sviluppo è prevista l'implementazione del modulo base con tecnologie open-source, quali Nvidia Nemo, e relativi modelli pre-addestrati, a partire dalle registrazioni messe a disposizione dall'azienda. In particolare, il cuore dell'architettura prevede 2 componenti fondamentali:

1. Un trascrittore generico, basato su modelli pre-addestrati quali:
  1. Quartznet per la trascrizione dei caratteri.
  2. Conformer-Transducer per la trascrizione di parole.
2. Un modello del linguaggio specializzato e dipendente dal task.

Il modello così sviluppato verrà testato in ambiente controllato grazie al supporto delle trascrizioni di tali conversazioni e si procederà all'analisi degli errori al fine di implementare un modello specifico per i task gestiti dall'AV.

### **Sviluppo**

Durante questa fase il modulo base viene implementato nell'AV e rilasciato in prova ad un gruppo di clienti, con i seguenti obiettivi:

- Valutazione e validazione della soluzione base proposta.
- Raccolta di uno o più dataset di parlato specifico, mirato alla specializzazione del modello.

I dati delle conversazioni di partenza, infatti, sono caratterizzate dall'utilizzo di un'unica varietà dell'italiano. Facendo fruire l'applicazione ad utenti sparsi su tutto il territorio, invece, si mira ad allargare lo spettro di varietà che l'ASR deve essere in grado di riconoscere.

Nella seconda fase si punta, dunque, alla definizione ed all'implementazione della procedura di personalizzazione del modulo, sulla base di parametri quali:

- Task, per la specializzazione del modello del linguaggio.
- Varietà linguistica dei parlanti, per la specializzazione del trascrittore sul parlato specifico.

In questa fase vengono sfruttati i dati raccolti dai clienti per fornire uno strumento ritagliato sulle specifiche esigenze.

### **Valutazione**

La valutazione del modulo avverrà durante entrambe le fasi di sviluppo tramite strumenti di monitoraggio continuo, feedback sia diretti che indiretti da parte dei clienti e meta-dati relativi all'utilizzo dei servizi ed agli utilizzatori degli stessi (ad esempio, tipologia di attività, posizione geografica ed altri).

## **Parallelizzabilità dei moduli in cui è articolato il progetto (al fine di consentire il lavoro simultaneo di tutti i tirocinanti partecipanti al progetto evitando tempi morti)**

I moduli, o fasi, del progetto possono essere parallelizzati di modo che i singoli tirocinanti possano lavorare ai task contemporaneamente.

In particolare, **nella fase preliminare** mentre l'informatico mette in piedi la pipeline di base del modello, il linguista e il data scientist possono procedere all'analisi dei dati di partenza, laddove il linguista si occuperà della selezione del dataset da utilizzare per lo sviluppo del modello e supporterà l'informatico nella scelta dei modelli pretrained da usare per l'implementazione del modello di base, mentre il data scientist analizzerà i fattori rilevanti da tenere in considerazione nell'implementazione del modulo di ASR nell'AV al fine di ottenere, poi, dati e meta-dati utili al miglioramento dello stesso nella seconda fase di sviluppo e in quella di valutazione.

Chiuderà il modulo di sviluppo preliminare una fase di test in ambiente controllato e di analisi degli errori al fine di implementare un nuovo modello.

Nella fase di **sviluppo con monitoraggio continuo**, man mano che i nuovi dati saranno raccolti attraverso l'applicazione finale, l'informatico procederà al cosiddetto CICD (lavorando, per esempio, al controllo del carico di lavoro, valutazione del grado di scalabilità, utilizzo delle risorse, etc.), mentre il data scientist analizzerà i dati raccolti (come posizione geografica, tipologia di attività, feedback clienti, etc.) ed il linguista valuterà quali pattern linguistici sono rilevanti al fine di ottimizzare il modello del linguaggio. Una volta valutati i miglioramenti, il modello sarà riaddestrato per le **valutazioni** finali.

#### **Costo formativo di ingresso del tirocinante (per l'utilizzo di attrezzature o di software non già oggetto di studio durante il percorso curricolare)**

Non si prevedono costi, se non l'eventuale implementazione di una macchina virtuale, qualora questa dovesse essere messa a disposizione dall'azienda.