

# **POLITECNICO DI MILANO**

Scuola di Ingegneria Industriale e dell'Informazione  
Corso di Laurea Magistrale in Ingegneria Informatica



**CitBot:**  
design e sviluppo di un chatbot in azienda

Relatore: Prof. Pier Luca LANZI

Tesi di Laurea di:

Simone BRUZZECHESSA  
Matr. 875363

Anno Accademico 2018 - 2019



# Ringraziamenti

Ringrazio la mia azienda, Revevol, per avermi dato la possibilità di lavorare a questo progetto e di portarlo come argomento di tesi di laurea. Ringrazio inoltre tutte le persone con cui ho collaborato e che mi hanno supportato durante la realizzazione del progetto.

Ci tengo inoltre a ringraziare l'Associazione Luca Coscioni, con cui ho collaborato nell'ultimo anno, per avermi dato la possibilità di realizzare questo progetto, è stato un piacere lavorare con voi.

Ringrazio il professore Pier Luca Lanzi per la fiducia e la disponibilità che mi ha concesso.

Ringrazio la mia famiglia perché anche grazie al loro sostegno e al loro aiuto sono diventato la persona che sono oggi.

Infine grazie te, Giulia, per tutto.

*Milano, Ottobre 2019*

S. B.



# Indice

<b>1</b>	<b>Introduzione</b>	<b>1</b>
<b>2</b>	<b>Cosa sono i Chatbot</b>	<b>3</b>
2.1	Chatbot . . . . .	3
2.2	Vantaggi dell'utilizzo dei chatbot . . . . .	6
2.3	Tipologie di chatbot . . . . .	6
2.3.1	Generative-based o Retrieval-based . . . . .	7
2.3.2	Open o Closed domain . . . . .	9
2.4	Chatbot di altre Associazioni . . . . .	10
2.4.1	Mencap . . . . .	10
2.4.2	Shelter Scotland . . . . .	11
<b>3</b>	<b>Obiettivi</b>	<b>13</b>
3.1	L'Associazione Luca Coscioni . . . . .	13
3.1.1	Temi trattati dall'Associazione . . . . .	14
3.2	Requisiti dell'associazione e obiettivi dell'elaborato . . . . .	15
<b>4</b>	<b>Soluzione</b>	<b>17</b>
4.1	Architettura . . . . .	17
4.2	Servizi . . . . .	18
4.2.1	Google Spreadsheet GSuite . . . . .	18
4.2.2	Dialogflow . . . . .	20
4.2.3	Kommunicate . . . . .	21
4.2.4	AppEngine . . . . .	21
4.2.5	Cloud Datastore . . . . .	22
4.3	L'intelligenza artificiale in CitBot . . . . .	23
4.3.1	Neural Architecture Search . . . . .	24

---

4.3.2	Transfer Learning . . . . .	25
4.3.3	Cloud AutoML . . . . .	26
4.3.4	Cloud Translate API . . . . .	28
4.3.5	Cloud Text-to-speech e Speech-to-text . . . . .	29
4.4	Addestramento dei modelli di ML . . . . .	29
4.4.1	Modello per la classificazione della tematica . . . . .	30
4.4.2	Modello per la classificazione delle domande di un argomento . . . . .	30
4.5	Flusso della conversazione . . . . .	31
4.6	Piattaforme integrate . . . . .	35
4.7	Integrazioni su siti Web . . . . .	36
<b>5</b>	<b>Valutazione</b> . . . . .	<b>39</b>
5.1	Valutazione dei modelli . . . . .	39
5.1.1	Valutazione modello per la classificazione delle aree tematiche . . . . .	42
5.1.2	Valutazione modello per la classificazione delle risposte sul Biotestamento . . . . .	43
5.2	Valutazione da parte dell'Associazione e degli utenti . . . . .	44
<b>6</b>	<b>Conclusioni</b> . . . . .	<b>49</b>

# Elenco delle figure

2.1	Ricerca della parola chatbot dal 2011 ad oggi, fonte Google Trends . . . . .	4
2.2	Panoramizza Chatbot 2017, fonte KeyReply . . . . .	5
2.3	Grafico degli utenti attivi mensilmente per le principali 4 app di messaggistica e social network, fonte Business Insider	5
2.4	Panoramica delle tipologie di chatbot in commercio, fonte Chatbotslife . . . . .	7
2.5	Esempio di architettura seq2seq, fonte Chatbotslife . . . . .	8
2.6	Chatbot di Mencap: the voice of learning disability, sito Mencap . . . . .	11
2.7	Ailsa: l'assistente di Shelter Scotland, sito Shelter Scotland	12
4.1	Architettura di CitBot . . . . .	18
4.2	Foglio di lavoro Google Spreadsheet . . . . .	19
4.3	Console di gestione di Dialogflow . . . . .	21
4.4	Interfaccia del Cloud Datastore su GCP . . . . .	23
4.5	Neural Architecture Search . . . . .	25
4.6	Differenze fra ML tradizionale e Transfer learning . . . . .	27
4.7	Cloud Auto ML . . . . .	28
4.8	Set di dati . . . . .	30
4.9	Set di dati per il primo modello . . . . .	31
4.10	Set di dati per il secondo modello . . . . .	32
4.11	Flusso logico delle informazioni in CitBot . . . . .	32
4.12	Chatbot integrato all'interno della pagina facebook dell'associazione . . . . .	35
4.13	CitBot integrato in Telegram . . . . .	36
4.14	Sito web dedicato al progetto CitBot . . . . .	37

4.15	Integrazione del chatbot sul sito web dell'Associazione Luca Coscioni . . . . .	37
5.1	Matrice di confusione . . . . .	41
5.2	Precision e Recall del modello di classificazione delle aree tematiche . . . . .	42
5.3	Matrice di confusione . . . . .	43
5.4	Precision e Recall del modello sul biotestamento . . . . .	44
5.5	Matrice di confusione . . . . .	45
5.6	Percentuale di domande relative alle diverse aree tematiche	46
5.7	Percentuali di successo nel riconoscimento delle aree tematiche e delle risposte . . . . .	47



# Sommario

Questo elaborato tratta lo sviluppo di *CitBot*, nome utilizzato per identificare il progetto commissionato all'azienda *Revevol Italia s.r.l.* dall'*Associazione Luca Coscioni*, una no profit che si occupa di promuovere le libertà civili e i diritti umani. Il significato del nome, che si compone del prefisso "Cit", trova ragione nello scopo del progetto, quello di diventare il chatbot dei "cittadini", italiani e non, uno strumento che sia in grado di dare risposte riguardo ai loro diritti e a tutte le tematiche legate all'Associazione.

La soluzione realizzata utilizza prodotti quasi interamente offerti dalla *Google Cloud Platform (GCP)*, ovvero la suite di servizi di *cloud computing* messa a disposizione dall'azienda di Mountain View. In particolare, le funzionalità di *machine learning* e di *natural language processing* sono state realizzate utilizzando "Cloud AutoML", un prodotto della GCP rilasciato di recente in grado di creare modelli di classificazione sfruttando tecniche avanzate di *transfer learning* e *neural architecture search*.

La tesi descrive l'intero processo dello sviluppo della soluzione, dalla prima fase di raccolta dei requisiti, avvenuta ad Ottobre 2018, fino alla messa in produzione, avvenuta a Maggio 2019. Si conclude con degli spunti sulle evolutive future.

# Abstract

This paper deals with the development of CitBot, a name used to identify the project commissioned to the company Revevol Italia s.r.l. by the Luca Coscioni Association, a non-profit organization that promotes civil liberties and human rights. The meaning of the name, composed by the prefix "Cit", is linked to the purpose of the project, namely to become the chatbot of the "citizens", Italian and not only, a tool that is able to give answers about their rights and all the topics covered by the Association.

The solution implemented uses products almost entirely offered by the Google Cloud Platform (GCP), the suite of cloud computing services offered by the Mountain View company. In particular, the machine learning and natural language processing features were implemented using Cloud AutoML, a recently released GCP product designed to create textual classification models using advanced transfer learning and neural architecture search techniques.

The thesis describes the entire development process of the solution, from the first phase dedicated to the definition of the requirements, which took place in October 2018, until the final release in production, which took place in May 2019. It concludes with some ideas on future evolution and enhancements.

# Capitolo 1

## Introduzione

L'attuale panorama complessivo della messaggistica istantanea è caratterizzato da una forte spinta verso grandi innovazioni tecnologiche e queste continue trasformazioni, a tendere, potranno avere un'enorme influenza su ogni settore della nostra vita. L'innovazione della messaggistica istantanea poggia sul concetto di automazione e "machine learning" e il chatbot, software progettato per simulare una conversazione con un essere umano, è un'espressione concreta di questi concetti.

L'obiettivo dell'elaborato è proprio quello di descrivere la realizzazione di CitBot, un chatbot che sia in grado di informare i cittadini, italiani e non, dei propri diritti; da qui il nome. Il progetto è stato realizzato durante la mia esperienza lavorativa ed è stato commissionato dall'Associazione Luca Coscioni, una no profit che si occupa delle libertà civili e i diritti umani (diritto alla scienza, all'assistenza personale autogestita, all'abbattimento della barriere architettoniche, ecc.). Revevol, l'azienda in cui ho sviluppato il progetto, è un *system integrator* partner Google perciò i servizi utilizzati per la realizzazione della soluzione sono offerti dalla *Google Cloud Platform* (GCP), ovvero la suite di servizi di *cloud computing* dell'azienda di Mountain View. In particolare, le funzionalità di *machine learning* e di *natural language processing* sono state realizzate sfruttando "Cloud AutoML", un servizio della GCP rilasciato recentemente in grado di creare modelli di classificazione sfruttando tecniche avanzate di *transfer learning* e *neural architecture search*.

L'elaborato descrive l'intero arco di sviluppo della soluzione realizzata, dall'analisi dei requisiti, avvenuta ad Ottobre 2018, fino alla messa in

produzione il 31 Maggio 2019. In dettaglio:

**Nel secondo capitolo** è delineato lo stato dell'arte commerciale dei chatbot. Dopo una prima parte introduttiva sui chatbot e sulla loro diffusione, se ne descrivono le diverse categorie esistenti ed alcuni esempi realizzati da altre associazioni benefiche.

**Il terzo capitolo** , dopo un'introduzione dell'Associazione Luca Coscioni, riassume i requisiti e gli obiettivi del progetto.

**Nel quarto capitolo** si descrive in dettaglio la soluzione realizzata, sia in termini di architettura, servizi utilizzati e *machine learning* ma anche flussi delle conversazioni e integrazioni con le piattaforme di messaggistica e i siti web.

**Nel quinto capitolo** viene valutato l'effettivo funzionamento di alcuni modelli di *machine learning* realizzati e presenta un'analisi del prodotto basata sul feedback dell'Associazione e degli utenti finali.

**Nel sesto ed ultimo capitolo** si riassumono i risultati ottenuti con uno sguardo verso ulteriori sviluppi futuri.

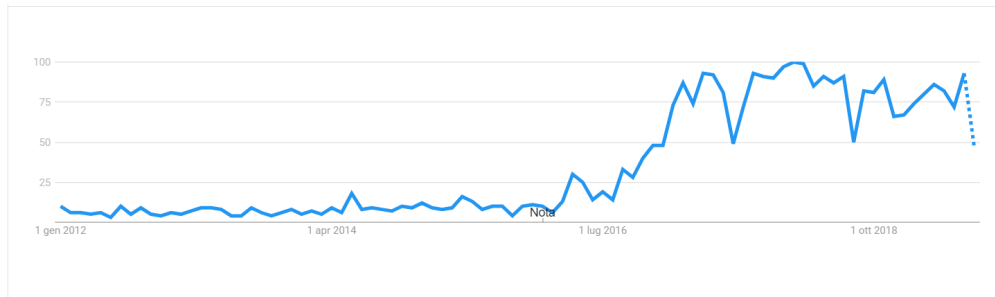
# Capitolo 2

## Cosa sono i Chatbot

### 2.1 Chatbot

La parola chatbot è unione di due parole distinte: "bot" (abbreviazione di robot) che indica un programma o uno script utilizzato per l'automazione di compiti ripetitivi che spesso vengono svolti dall'uomo e "chat" (dall'inglese "dialogo") in riferimento alle piattaforme di messaggistica. Un chatbot è quindi un software progettato per interagire con gli esseri umani all'interno di una piattaforma di messaggistica utilizzando linguaggio naturale. I chatbot, i quali possono essere realizzati per diversi scopi (il supporto per i clienti, rispondere alle FAQ degli utenti, eseguire un ordine per conto di una persona), fanno parte della famiglia delle interfacce conversazionali, le quali consentono ad un utente di comunicare con un computer come se fosse un essere umano. L'interesse del mercato per questa tecnologia è dimostrato in modo quantitativo da Google Trend (grafico in figura [2.1 nella pagina seguente](#)) attraverso uno studio sulla frequenza di ricerca della parola "chatbot" e in cui si nota come dal 2017 ci sia stata una vera e propria esplosione senza soluzione di continuità fino ad oggi.

Come ulteriore testimonianza dell'interesse da parte del mercato per questa tecnologia ci sono le centinaia di aziende, di prodotti e di *framework* che sono nati e che sono stati sviluppati per il design, la realizzazione e la gestione dei chatbot. Per comprendere la diffusione del fenomeno è riportata in figura [2.2 a pagina 5](#) una panoramica, dal cerchio più interno a quello più esterno, rispettivamente di: piattaforme di messaggistica



**Figura 2.1:** Ricerca della parola chatbot dal 2011 ad oggi, fonte Google Trends

che permettono l'integrazione, grandi aziende che hanno adottato questa tecnologia, fornitori in grado di realizzare queste soluzioni ed infine gli strumenti che ne permettono lo sviluppo.

I motivi che sono dietro a questo fenomeno sono molteplici e dipendono sia da fattori tecnologici che sociali.

Rispetto al sociale, la diffusione dei chatbot è legata al fatto che, negli ultimi anni, il numero di persone che utilizzano le applicazioni di messaggistica è maggiore rispetto a quelle che fanno uso dei social network. In figura 2.3 a fronte il sorpasso, avvenuto nel 2015, degli utenti attivi su applicazioni di messaggistica ai danni dei social network.

Un altro fattore è legato ad un progresso tecnologico negli ambiti di *machine learning*, *sentiment analysis*, *natural language processing* e *natural language understanding*, che ha reso i chatbot in grado di interpretare con maggiore efficacia il linguaggio umano e di individuare più facilmente le necessità e gli stati d'animo degli utenti. Queste innovazioni, ampliando le funzionalità dei chatbot, ne hanno permesso l'utilizzo in molti più settori innescando un'inevitabile crescita della domanda da parte del mercato.

Un'ultima motivazione riguarda i costi ed i tempi di sviluppo. Infatti, questi sono molto spesso inferiori ai costi di realizzazione di altre tipologie di soluzioni, in quanto utilizzano tool e framework il cui servizio in alcuni casi è offerto gratuitamente.

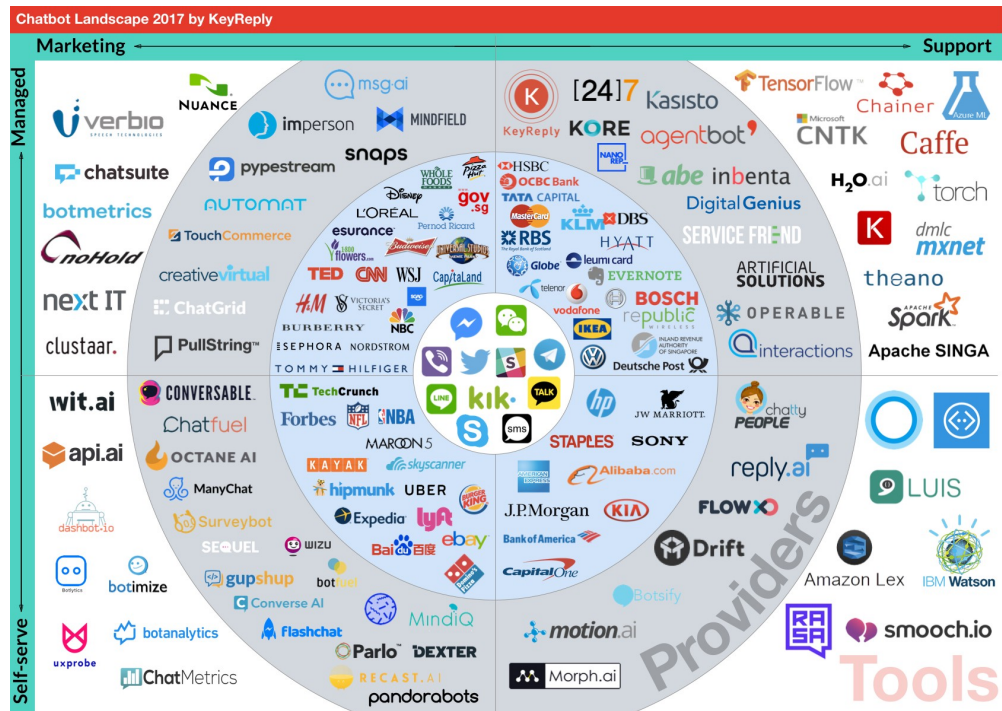
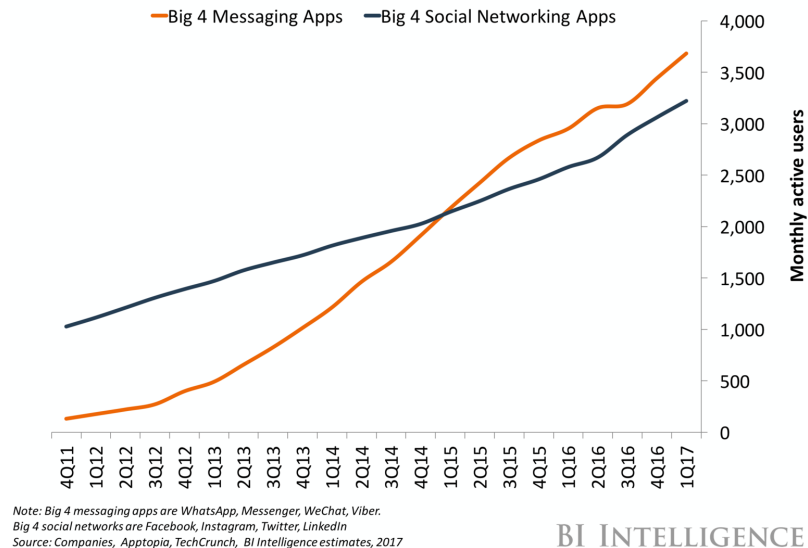


Figura 2.2: Panoramica Chatbot 2017, fonte KeyReply

**Messaging Apps Have Surpassed Social Networks**

Global monthly active users for the top 4 messaging apps and social networks, In millions



BI INTELLIGENCE

Figura 2.3: Grafico degli utenti attivi mensilmente per le principali 4 app di messaggistica e social network, fonte Business Insider

## 2.2 Vantaggi dell'utilizzo dei chatbot

I vantaggi principali derivanti dall'utilizzo di un chatbot dipendono chiaramente dalle specifiche implementazioni e dalla tipologia di servizio offerto. Si possono comunque individuare alcuni vantaggi generali ed indipendenti dalla specifica implementazione perché intrinseci alla soluzione chatbot, che sono stati riportati di seguito:

- disponibilità del servizio 24 ore su 24, 7 giorni su 7 anche nei giorni di festività;
- un servizio più rapido rispetto a quello umano e in grado di processare un numero elevato di richieste contemporaneamente in modo da ridurre il tempo di attesa dei clienti e, di conseguenza, migliorare la *user experience*;
- facilità nell'utilizzo dovuta all'assenza di curva di apprendimento da parte degli utenti.
- facilità nel trovare le informazioni necessarie relative ad un determinato argomento (molto più facile che navigare le FAQ di un sito ad esempio), oppure facilità nell'ottenere un determinato risultato rispetto ad esempio all'utilizzo di un'applicazione dedicata;
- riduzione del costo per il committente (ad esempio i *customer service*). Infatti realizzare e mantenere un chatbot è, in genere, molto più economico che assumere un dipendente. L'avanzamento della tecnologia e dei framework per la realizzazione dei chatbot renderanno questo divario ancora più evidente;

## 2.3 Tipologie di chatbot

Nella figura [2.4 nella pagina successiva](#) è rappresentata una tabella contenente le 4 principali tipologie di chatbot presenti in commercio, le cui differenze dipendono sia dalla complessità delle conversazioni in grado di gestire, che dalle risposte fornite agli utenti.



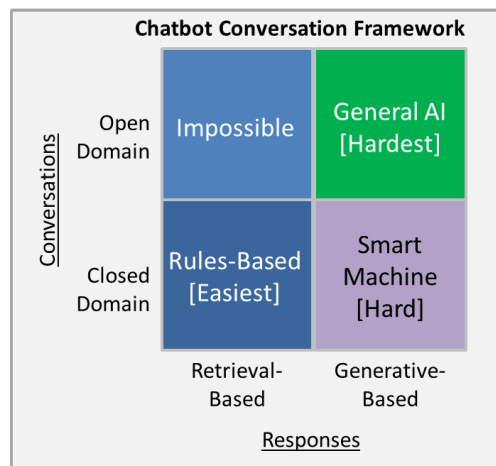


Image from Medium

**Figura 2.4:** Panoramica delle tipologie di chatbot in commercio, fonte Chatbotlife

### 2.3.1 Generative-based o Retrieval-based

Una prima distinzione fra le due categorie *generative-based* e *retrieval-based* di chatbot è in riferimento al modello con cui vengono prodotte le risposte fornite all'interlocutore umano.

#### Generative-based models

I chatbot di tipo *generative-based* sono caratterizzati dalla generazione dinamica della risposta che avviene ad ogni nuova interazione con l'utente. Questi modelli si basano su tecnologie molto simili a quelle del *Machine Translation* ma, invece di fare traduzioni tra due lingue diverse, traducono da una frase in input detta "contesto" ad una di output che sarà la risposta fornita all'utente.

Spesso questi sistemi utilizzano un'architettura di reti neurali di tipo *sequence-to-sequence* come quella in figura 2.5 nella pagina seguente, dove il contesto (ovvero l'input dell'utente) è codificato dalla prima rete in un vettore, il quale è a sua volta decodificato nella risposta da fornire all'utente nella seconda.

Il vantaggio principale di questa soluzione è l'assenza della definizione di

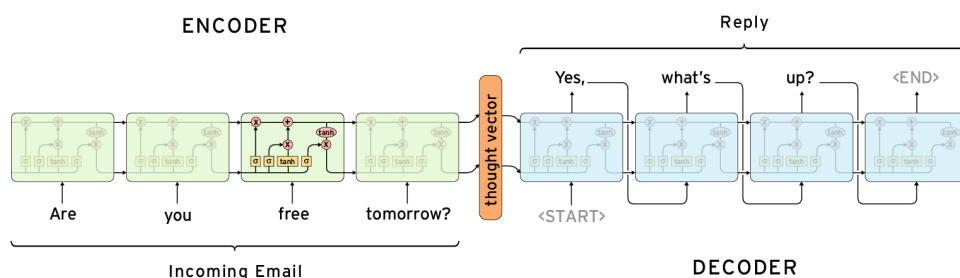


Image from Medium

**Figura 2.5:** Esempio di architettura seq2seq, fonte Chatbotlife

un set di risposte predefinite e delle loro regole di selezione sulla base delle domande dell'utente, necessarie invece nella categoria definita successivamente. Un chatbot di questo tipo è autonomo nella generazione delle risposte, le uniche informazioni necessarie al suo funzionamento sono lo storico delle conversazioni utilizzate durante la fase di addestramento.

Il problema principale di questo modello è la difficoltà di realizzazione, sia in termini di costi che di performance. Una soluzione di questo tipo può infatti generare risposte grammaticalmente scorrette (o errate) oppure rispondere in maniera diversa alla stessa domanda. Una ulteriore complessità è la necessità, per questi modelli, di utilizzare un set di dati di dimensioni elevate al fine di ottenere un risultato in grado di fornire risposte accurate.

### Retrieval-based models

A differenza dei modelli di tipo *generative-based*, quelli di tipo *retrieval-based* non generano nuove risposte, bensì si limitano a selezionarne una da una lista predefinita. A partire da una frase inviata dall'utente e secondo un determinato criterio, questa soluzione deve individuare la migliore risposta possibile rispetto a quelle predefinite (chiamate *candidate responses*). La loro selezione si basa su un meccanismo che consiste nell'associare loro un punteggio, determinato in base al criterio definito in precedenza, utilizzato per l'ordinamento dalla più adeguata alla meno. Infine viene selezionata la risposta con il punteggio più elevato che sarà quella inviata all'utente finale.

Il criterio per l'assegnazione dello score alle *candidate response* può essere

più o meno sofisticato: dai più semplici sistemi *rule-based* basati sulla corrispondenza testuale a classificatori complessi di *machine learning* e *natural language understanding*.

Il beneficio, rispetto al modello precedente, è la correttezza delle risposte inviate dal chatbot, trattandosi in questo caso di una lista predefinita non si verificheranno problemi grammaticali o di risposte errate dal punto di vista grammaticale. Inoltre i sistemi di questo tipo sono più facili da realizzare e richiedono meno dati per l'addestramento.

Lo svantaggio è la necessità di realizzare il set di risposte del chatbot e la scelta di un buon meccanismo di assegnazione degli score alle risposte, in modo da replicare alle domande effettuate dall'utente con quelle più corrette.

### 2.3.2 Open o Closed domain

A seguire l'ulteriore divisione dei chatbot nelle due categorie *open domain* e *closed domain* in base al dominio delle conversazioni possibili.

#### Open Domain

I chatbot di tipo *open domain* sono quelli per cui non esiste un limite negli argomenti e nelle tematiche alle quali il bot è in grado di rispondere. Questi sistemi devono poter rispondere a qualsiasi domanda rivolta dall'utente, anche in assenza di un contesto specifico oppure di un obiettivo definito.

Per questo motivo la realizzazione di questi sistemi è molto complessa e legata ad un approccio di tipo *generative based* (sarebbe impossibile realizzare un set di risposte per un chatbot *retrieval based*).

#### Closed Domain

La tipologia *closed domain* prevede lo sviluppo di un chatbot per uno specifico dominio o problema. La soluzione realizzata sarà responsabile di rispondere alle domande rivolte al solo dominio di competenza, un chatbot realizzato per la gestione degli ordini di un ristorante sarà in grado di

rispondere a: "vorrei ordinare delle pizze da asporto" ma non a "come è il tempo oggi?".

La categoria adottata per la realizzazione di CitBot è di tipo *retrieval-based* e *closed domain*, come approfondito nei capitoli successivi dell'elaborato, il chatbot sarà responsabile di rispondere alle domande rivolte dagli utenti selezionando una delle risposte predefinite realizzate dal committente, tutte relative ad un numero limitato di tematiche.

## 2.4 Chatbot di altre Associazioni

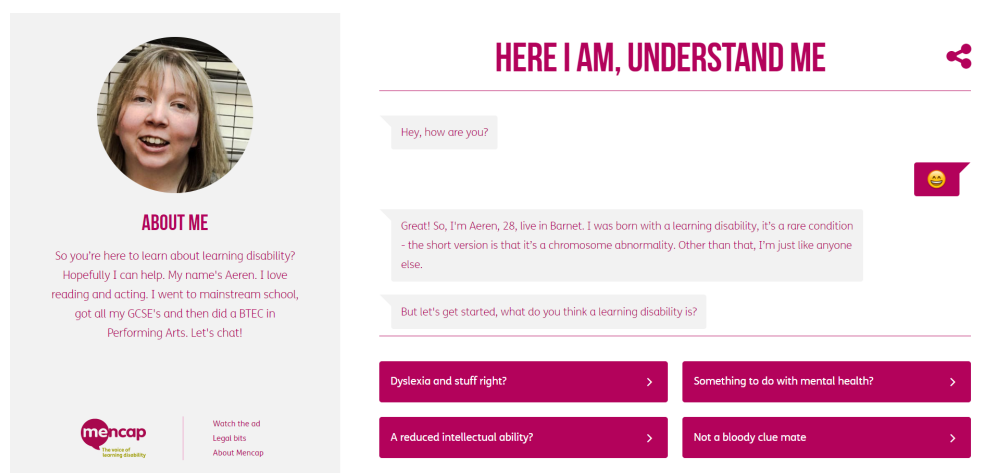
All'interno di questa sezione saranno presentati alcuni esempi di chatbot disponibili online realizzati, come CitBot, da associazioni benefiche e che ne condividono, seppur utilizzando tecnologie diverse, la stessa tipologia descritta in precedenza.

### 2.4.1 Mencap

La Royal Mencap Society è un'associazione benefica con sede nel Regno Unito che lavora con persone affette da difficoltà di apprendimento. Il chatbot realizzato fa parte della campagna 'Here I Am', la sua funzione è quella di aiutare le persone ad avere informazioni sui problemi legati all'apprendimento e a superare molti stereotipi che affliggono questa tematica.

Il chatbot in questione è disponibile solo sul sito dell'associazione (non è stata predisposta nessuna integrazione con piattaforme come Facebook Messenger), inoltre non viene fatto uso di NLP (*natural learning processing*). L'utente che si interfaccia alla soluzione viene semplicemente guidato attraverso una lista di domande e risposte predefinite (una sorta di albero decisionale). Nonostante non sia possibile personalizzare l'esperienza di conversazione, i dialoghi realizzati sono sorprendentemente curati e l'umorismo con il quale risponde il chatbot dà all'utente l'idea di avere un dialogo con una vera persona.

Secondo Mencap il numero di persone che utilizza il chatbot è ancora basso, nonostante ciò il livello di coinvolgimento è comunque elevato. Inoltre



**Figura 2.6:** Chatbot di Mencap: the voice of learning disability, sito Mencap

sostengono che questa soluzione abbia contribuito ad un incremento del 3% nella consapevolezza dell'associazione e che sia riuscito nel suo obiettivo principale, quello di educare gli utenti.

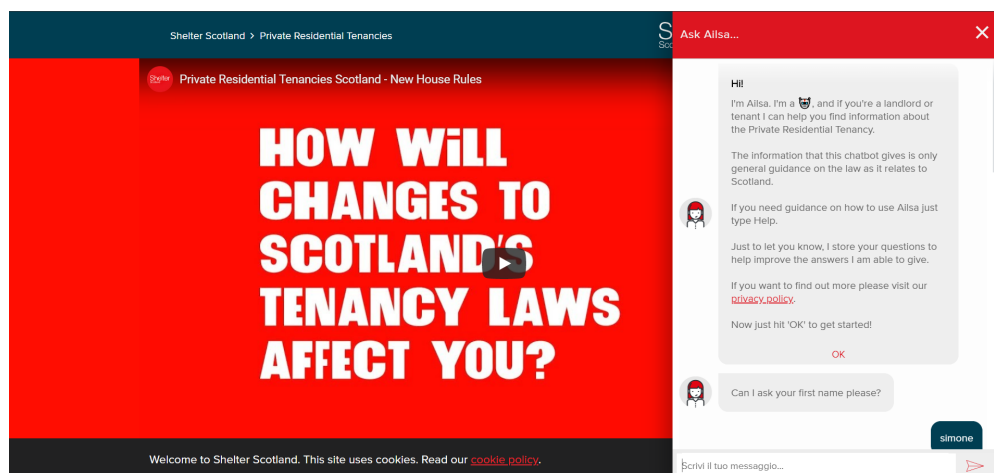
In figura 2.6 un'immagine della pagina del chatbot con le prime interazioni guidate.

### 2.4.2 Shelter Scotland

Shelter Scotland è un'associazione scozzese che aiuta oltre mezzo milione di persone ogni anno che abitano in edifici in cattive condizioni o sono senzatetto attraverso servizi di consulenza, supporto e legali. Lo scopo dell'associazione è quello di fare in modo che in futuro nessuno debba rivolgersi a loro per chiedere aiuto.

Contrariamente alle soluzioni di altre associazioni benefiche realizzate per aumentare la consapevolezza o aumentare il coinvolgimento per alcune tematiche, Shelter Scotland ha realizzato un chatbot per un obiettivo ben preciso. L'associazione ha infatti implementato il chatbot "Ask Ailsa" sul suo sito Web, sviluppato per informare tutti i proprietari ed inquilini degli immobili privati in Scozia sulle modifiche attuate alla legislazione sull'affitto privato e per rispondere a domande specifiche su questo argomento.

Il lancio del chatbot ha permesso all'associazione di semplificare il servizio di informazione e ridurre il numero di richieste in arrivo al servizio di assistenza.



**Figura 2.7:** Ailsa: l'assistente di Shelter Scotland, sito Shelter Scotland

Ne sono state realizzate due versioni entrambe basate su *Microsoft's Azure bot framework* <sup>1</sup> e *LUIS* <sup>2</sup> (come *natural language processing model*). In figura 2.7 il chatbot in funzione sul sito dell'associazione.

<sup>1</sup>Microsoft Azure è la piattaforma cloud pubblica di Microsoft, che offre servizi di cloud computing.

<sup>2</sup>Language Understanding (LUIS) è un servizio di API basato sul cloud che, attraverso l'apprendimento automatico utilizzato sul testo in linguaggio naturale, è capace di estrarre il significato generale ed le informazioni più rilevanti e dettagliate.

# Capitolo 3

## Obiettivi

In questo capitolo viene presentata l'Associazione Luca Coscioni, committente del progetto, successivamente saranno riassunti gli obiettivi dell'elaborato.

### 3.1 L'Associazione Luca Coscioni

L'Associazione Luca Coscioni è un'associazione no profit di promozione sociale iscritta al Registro persone giuridiche a carattere nazionale e a quello per la trasparenza UE.

Tra le priorità dell'Associazione vi sono l'affermazione delle libertà civili e i diritti umani, in particolare quello alla scienza, l'assistenza personale autogestita, l'abbattimento della barriere architettoniche, le scelte di fine vita, la legalizzazione dell'eutanasia, l'accesso ai cannabinoidi medici e il monitoraggio mondiale di leggi e politiche in materia di scienza e autodeterminazione.

Fondata nel 2002 da Luca Coscioni, leader Radicale e docente universitario, malato di sclerosi laterale amiotrofica, l'Associazione negli anni ha promosso azioni a livello nazionale, come il referendum del 2005 per abrogare la Legge 40 sulla "procreazione medicalmente assistita" o le proposte di legge d'iniziativa popolare per la legalizzazione dell'eutanasia o la produzione, consumo e commercio della cannabis per tutti i fini, solo per citare alcune iniziative.

L'Associazione Luca Coscioni non è una associazione che si occupa di assistenza, bensì promuove l'iniziativa di disabili e malati, rendendoli protagonisti della vita politica. Luca Coscioni e, l'ex co-presidente Piergiorgio Welby hanno guidato mobilitazioni non violente che hanno gettato le basi per le disobbedienze civili di Marco Cappato e Mina Welby che hanno assicurato all'Italia le disposizioni anticipate di trattamento e portato la Corte Costituzionale a pronunciarsi sull'articolo 580 del codice penale relativo al suicidio assistito.

### 3.1.1 Temi trattati dall'Associazione

A seguire il dettaglio delle tematiche sulle quali è impegnata l'Associazione Luca Coscioni, la lista non è riportata al solo scopo informativo in quanto buona parte di esse sono già state integrate all'interno della soluzione realizzata e le altre lo saranno in futuro.

**Libertà di ricerca scientifica** l'Associazione nasce con l'obiettivo di ottenere libertà per la ricerca scientifica perché pensa che in Italia, ancora oggi, troppe ideologie e assolutismi di carattere soprattutto religioso, ostacolano l'avanzamento della ricerca che potrebbe portare un giorno benefici per persone affette da malattie che oggi non lasciano speranza di vita.

**Fecondazione assistita** l'Associazione è da anni in prima linea per abolire la legge 40/2004, che impedisce di fatto le buone pratiche mediche della fecondazione assistita in Italia.

**Fine vita e eutanasia** l'Associazione crede che le decisioni di fine vita siano decisioni personalissime e, in quanto tali, debbano essere prese con la massima libertà dalla persona per sé stessa. In Italia, la Costituzione riconosce che nessuno può essere obbligato ad alcun trattamento sanitario contro la propria volontà e prevede altresì che la libertà personale è inviolabile.

**Diritti dei disabili** l'Associazione Coscioni si batte da anni per il riconoscimento dei diritti dei disabili, per la lotta contro le discriminazioni, per garantire i dispositivi idonei alle persone con disabilità. È stata anche realizzata un'applicazione "No Barriere" per aiutare i cittadini a segnalare violazioni, barriere architettoniche o disservizi.



**Aborto e contraccezione** per l'Associazione decisioni personali e private della persona in merito ad un tema così delicato come la contraccezione e l'aborto devono essere sempre rispettate, in particolare per le questioni che attengono alle scelte sul proprio corpo e alle relative scelte riproduttive. Si batte perché tutti possano prendere queste decisioni in modo responsabile, sia con una adeguata informazione sessuale e riproduttiva, sia accedendo con pienezza di diritti all'interruzione di gravidanza.

**Antiproibizionismo** l'antiproibizionismo è un concetto ampio, che abbraccia anche il concetto di libertà di ricerca scientifica e si fonda su regole. In Italia non è vietato l'uso della Cannabis terapeutica per i malati o per la cosiddetta terapia del dolore, nonostante questo sia dal punto di vista burocratico che culturale, il ricorso a queste pratiche vede posizionata l'Italia fra gli ultimi paesi nel mondo.

## 3.2 Requisiti dell'associazione e obiettivi dell'elaborato

Durante i primi contatti avvenuti con l'associazione sono stati definiti una lista di requisiti da soddisfare durante la realizzazione della soluzione presentata in questa tesi.

L'obiettivo principale dell'associazione è quello di realizzare un chatbot che sia in grado di informare il cittadino sui propri diritti fondamentali e sulle proprie libertà civili al fine di consentirgli la loro difesa, sopperendo a una lacuna dello Stato su alcune di queste tematiche.

La soluzione da realizzare è posta al centro di una vera e propria campagna di informazione e di comunicazione su tutte le tematiche per cui è impegnata sia l'Associazione Luca Coscioni che, potenzialmente, anche altre associazioni. Il sistema deve essere in grado di rispondere in maniera autonoma alle domande che vengono rivolte dalle persone all'associazione, centinaia di mail e messaggi vengono ricevuti ogni mese sul sito e sulle varie piattaforme in cui è attiva, molti dei quali di carattere informativo generale. Utilizzando il chatbot, le persone dovrebbero essere in grado di avere le informazioni necessarie in maniera autonoma, in questo modo andrebbe a diminuire il numero di contatti ricevuti dall'associazione, alleggerendola dalla maggior parte del lavoro più ripetitivo e permettendo ai suoi membri

di concentrarsi solo sulle questioni più complesse, per le quali è necessario l'intervento di una persona.

A seguire la lista dettagliata dei requisiti per la soluzione da realizzare emersi durante gli incontri preliminari con il committente, alcuni dei quali sono in realtà le caratteristiche principali di un sistema chatbot.

**Massima disponibilità della soluzione** : deve essere in grado di rispondere alle domande rivolte dalle persone in maniera autonoma e rapida senza interruzioni di servizio;

**Facilitare l'interazione per l'utente** : l'utilizzo del servizio deve risultare semplice ed intuitivo, la soluzione deve rimpiazzare ed espandere una soluzione FAQ esistente sul sito la cui fruizione risulta al momento complessa per gli utenti, soprattutto quelli più inesperti.

**Realizzare un sistema flessibile** : la soluzione deve garantire all'associazione l'incremento del numero di aree tematiche gestite (rispetto a quelle attuali) e deve essere possibile la sua integrazione con altri sistemi senza la necessità di dover modificare l'architettura iniziale;

**Massima accessibilità del chatbot** : il chatbot deve essere integrato nel maggior numero di piattaforme di messaggistica e deve permettere all'utente di interfacciarsi attraverso diverse tipologie di interazione. Oltre a quella testuale il chatbot deve permettere all'utente di comunicare attraverso la voce sia al momento dell'invio delle domande che durante la ricezione delle risposte. Questo requisito è fondamentale considerando che il target della soluzione sono i cittadini italiani in qualsiasi stato di salute, quindi anche persone non vedenti, con problemi di udito o mutismo.

I requisiti descritti in precedenza sono, di fatto, gli obiettivi dell'elaborato, a partire da essi è stata progettata e realizzata la soluzione descritta in dettaglio nel capitolo successivo.

# Capitolo 4

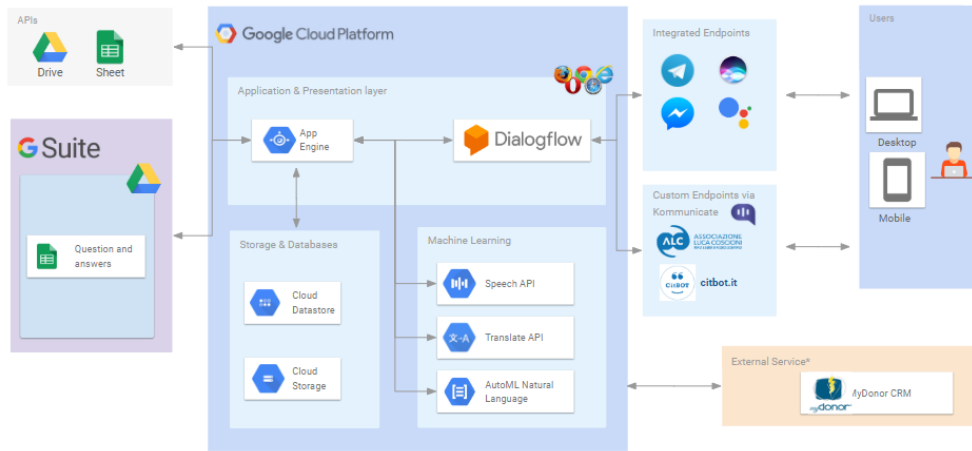
## Soluzione

Questo capitolo tratta la soluzione realizzata sulla base dei requisiti definiti nel capitolo precedente. La prima parte sarà dedicata all'architettura della piattaforma e una descrizione sommaria del funzionamento, a seguire il dettaglio dei servizi e delle componenti mostrate in precedenza con un focus sulla parte di *machine learning* all'interno del chatbot. Successivamente verrà descritto il funzionamento della piattaforma dal punto di vista tecnico ed infine verranno mostrate le integrazioni attuali del chatbot ed alcune immagini della soluzione attuale.

### 4.1 Architettura

Nella figura [4.1 nella pagina successiva](#) viene illustrata la soluzione architetture adottata per l'implementazione del chatbot. I singoli servizi e prodotti utilizzati, mostrati in figura, saranno oggetto di approfondimento nella prossima sezione, a seguire una descrizione di alto livello del sistema.

A destra gli utenti possono interfacciarsi al chatbot attraverso molteplici piattaforme di messaggistica e siti web, i messaggi inviati vengono poi processati dall'applicazione realizzata in questo elaborato. L'applicativo responsabile della logica del chatbot è anche in grado di comunicare con altri servizi dell'ecosistema Google, nello specifico: Google Spreadsheet e Google Drive che fanno entrambi parte della *G Suite*. Oggetto di integrazioni future è anche il componente "myDonor", soluzione CRM attualmente



**Figura 4.1:** Architettura ad alto livello

adottata dall'Associazione.

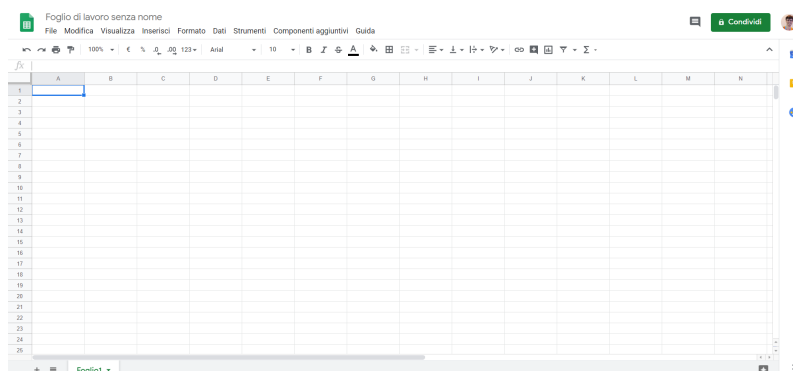
## 4.2 Servizi

Questa sezione è dedicata alla presentazione di tutti i servizi mostrati all'interno dell'architettura della soluzione. Oltre a presentarli, il secondo obiettivo di questo paragrafo è l'analisi dei vantaggi e dei benefici ricavati dal loro utilizzo e le problematiche o requisiti che sono stati in grado di risolvere.

### 4.2.1 Google Spreadsheet GSuite

*Google Sheets* è un programma offerto da Google all'interno dell'offerta G Suite per la gestione dei fogli elettronici (o di calcolo) equivalente a Microsoft Excel. Il software è offerto come applicazione web (quindi Software as a service o SaaS) e permette il salvataggio in maniera automatica dei file su Google Drive.

Google Spreadsheet consente agli utenti di creare e modificare i file direttamente online, aprendo a scenari di *collaboration* fra gli utenti in real-time. In ogni momento è possibile visualizzare le modifiche fatte dalle altre persone che stanno lavorando al foglio, tracciando le azioni svolte da ogni utente in una *history* condivisa del documento.



**Figura 4.2:** Foglio di lavoro Google Spreadsheet

Il proprietario del documento è in grado di gestire i permessi del file (utilizzando la stessa funzionalità di Google Drive) concedendo alle altre persone i permessi di modifica, visualizzazione o commento del file.

I vantaggi che derivano dall'utilizzo di questa soluzione sono stati molteplici, nello specifico:

**Collaborazione e revisione** : permettere a più persone dell'Associazione di lavorare contemporaneamente allo stesso foglio di lavoro su più tematiche senza doverne duplicare il contenuto, avendo la possibilità di consultare le modifiche fatte dagli altri;

**Interfaccia semplice ed intuitiva** : l'interfaccia offerta dal servizio è di facile utilizzo da parte degli utenti, inoltre nessuna formazione è stata necessaria per l'utilizzo di questo strumento già adottato dall'Associazione;

**Interazione via API** : possibilità di leggere e scrivere dal foglio di lavoro in maniera programmatica via API. L'applicazione sarebbe in grado di interagire con il foglio di lavoro dell'associazione che può essere utilizzato come interfaccia grafica per il caricamento delle domande e delle risposte utilizzate per l'addestramento del chatbot.

In figura 4.2 un esempio di foglio di calcolo di Google Spreadsheet.

### 4.2.2 Dialogflow

Dialogflow è una suite di sviluppo end-to-end basata sul principio "build-once deploy-everywhere" (crea una volta ed esegui il deployment ovunque) per la realizzazione di interfacce conversazionali (chatbot) su siti web, applicazioni per dispositivi mobili, piattaforme di messaggistica e dispositivi IoT.

Dialogflow supporta nativamente oltre 20 lingue e l'integrazione con almeno 14 piattaforme diverse fra le quali Messenger, Telegram, Skype, Slack, Google Assistant e Alexa.

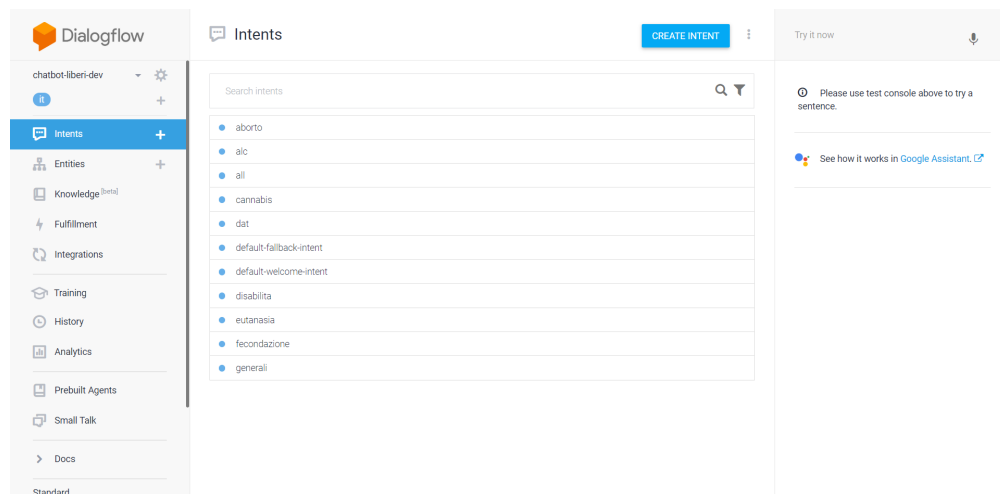
Una funzionalità importante offerta dal prodotto è la correzione automatica dell'ortografia e degli errori di battitura dei messaggi inviati dalle persone. In questi casi il servizio riesce a correggere automaticamente gli errori ortografici utilizzando una tecnologia simile a quella impiegata da Google per raffinare le ricerche effettuate dagli utenti sul suo motore di ricerca.

Il servizio supporta anche l'analisi real-time del sentiment delle frasi inviate dagli utenti grazie alla tecnologia *Google Cloud Natural Language*. I risultati delle analisi sul *sentiment* possono essere utilizzati per individuare utenti insoddisfatti o essere messi in relazione alle tematiche discusse dalle persone per determinare quali temi portano ad un sentimento "più positivo" da parte dei clienti.

Dialogflow permette di realizzare applicazioni serverless attraverso l'utilizzo delle *Cloud Functions for Firebase*, in alternativa viene fornita agli sviluppatori la possibilità di registrare la propria soluzione ad un "custom WebHook" per la ricezione, attraverso chiamate autenticate, delle informazioni collezionate durante le interazioni con gli utenti. Quest'ultima è la soluzione adottata all'interno di questo elaborato per garantire la massima flessibilità dell'applicazione realizzata, consentendole logiche di funzionamento complesse ed integrazioni con altri servizi.

La scelta di questo servizio risiede nella possibilità di integrare il chatbot realizzato attraverso Dialogflow su, almeno, 14 piattaforme diverse attraverso una semplice configurazione che Google definisce "*one-click*". L'utilizzo di questo prodotto ha semplificato il raggiungimento di uno dei requisiti del cliente, ovvero la possibilità di integrare il chatbot nel maggior numero di piattaforme e canali possibili.

In figura [4.3 nella pagina successiva](#) un'immagine della console di configu-



**Figura 4.3:** Console di gestione di Dialogflow

razione e monitoraggio dei chatbot realizzati utilizzando questo strumento.

### 4.2.3 Kommunicate

Kommunicate è una piattaforma sviluppata dall'omonima azienda per la gestione e l'automazione del servizio clienti mediante l'utilizzo di chatbot. La piattaforma offre diversi servizi fra i quali: integrazione nativa con chatbot di Dialogflow, integrazioni con molteplici software per il supporto clienti, report e analytics delle conversazioni avvenute fra gli utenti ed i chatbot, "handoff" fra i chatbot e le persone del supporto.

La soluzione realizzata utilizza il widget sviluppato dall'azienda per l'integrazione del chatbot all'interno dei siti web, in sostituzione a quello più limitato offerto nativamente dal servizio Dialogflow.

### 4.2.4 AppEngine

Google App Engine (GAE) è una piattaforma di tipo *Platform as a service* (PaaS) che offre servizi di *cloud computing* per lo sviluppo e l'hosting di applicazioni web gestite ed orchestrate da Google nei suoi Data Center. Le applicazioni eseguono su più server all'interno di *sandbox* per motivi di sicurezza. Il vantaggio di una soluzione PaaS come App Engine è la

scalabilità automatica delle applicazioni che avviene in base al numero di richieste ricevute dagli utenti.

Il numero di richieste a cui è soggetto un chatbot può variare di molto durante l'arco della giornata, è inoltre prevedibile la presenza di picchi di traffico che possono avere luogo in concomitanza, ad esempio, con il lancio di una dedicata campagna promozionale. Per questi motivi una soluzione basata su AppEngine riesce a garantire una gestione ottimale del numero di risorse responsabili del funzionamento del sistema. Nel momento di picco l'allocazione di più istanze permetterà di non incorrere in rallentamenti o interruzioni del servizio mentre, nel caso opposto di traffico ridotto, sarà automatico lo spegnimento delle risorse inutilizzate al fine di limitare i costi.

La scalabilità automatica, legata ad un modello di pagamento di tipo "pay per use" basato sull'utilizzo orario delle risorse, ha determinato la scelta di questa piattaforma per ospitare l'applicazione realizzata per CitBot.

#### 4.2.5 Cloud Datastore

*Cloud Datastore* è un database NoSQL ad alta scalabilità in grado di gestire automaticamente il partizionamento orizzontale e la replica. Si tratta di un database ad alta disponibilità e durabilità, le cui peculiarità sono le seguenti:

**Transazioni atomiche** : possibilità di eseguire una serie di operazioni in maniera atomica, il cui risultato finale è di successo per tutte oppure nessuna delle modifiche viene effettuata;

**Scalabilità e performance** : utilizzo di un'architettura distribuita per la gestione automatica dello *scaling*. L'utilizzo di indici e di specifici vincoli permettono alle *query* di scalare a seconda della dimensione dei risultati e non dell'intero set di dati;

**Flessibilità e linguaggio di query semplificato** : il Datastore è un database senza schema che riduce le attività manuali di aggiornamento della struttura dei dati sottostante durante la naturale evoluzione dell'applicazione. L'utilizzo di un motore di query avanzato consente il caricamento dei dati attraverso l'utilizzo di filtri ed ordinamenti;



Name/ID	Parent	response	responseHash
name-0eddb24860ff6e607a7bd2931344f2	Key(Intent, 'aborto')	La legge 194/1978 garantisce il tuo diritto all...	5e4c4bcff44ba0ac4caea872bc070ab
name-117bf134d0a2539ff4729aa3689f238	Key(Intent, 'aborto')	Puoi trovare la pillola dei 5 giorni dopo press...	8ae053b0bdea3287f173bb693546568
name-319b6bd430022abe381030c439e11...	Key(Intent, 'aborto')	Puoi trovare la pillola del giorno dopo press...	16200bf8ead513bda50f0a2870c4652
name-3a17046f2d450168b19ff0c372fac7	Key(Intent, 'aborto')	L'obiezione di coscienza è prevista dalla Leg...	bf32d0b50147be5c9a53d4316e9f008
name-3d5a2f072004cda14e99f4c94706c7d3	Key(Intent, 'aborto')	L'unica cosa che puoi fare è rivolgerti ad un a...	a9e9c4041456379ce0a304c4730ddca
name-5d98503a6dd62733e4d981d7ee687b...	Key(Intent, 'aborto')	Non esiste un registro dei medici obiettori di ...	7e0c529826c8773cea26d95599de5252
name-6b7bc0880c3842b4c53cae994c51e...	Key(Intent, 'aborto')	L'Associazione Luca Coscioni aiuta i cittadin...	dcaf1adccc52211451470a6631260647
name-7af9f1bcc5e9a26c10f8e240a4ba14	Key(Intent, 'aborto')	Norlevio, cioè la pillola del giorno dopo, è un ...	8f8b94ab28f8db503a349aee6bc24f5
name-8bfba5f0009ba3a18d1c0b6eea8bf6	Key(Intent, 'aborto')	Non siamo soliti segnalare cliniche specific...	b57daa4e781f6259e5e3b7e770a2c0ba
name-b90a41a461fee524432caed2478d6f9b	Key(Intent, 'aborto')	EllaOne, cioè la pillola dei 5 giorni dopo, è un ...	cd61f969c8164915cd07943ea3845a8b5
name-cd753585837a70afa1d39ea52a6d10...	Key(Intent, 'aborto')	Nel caso in cui il farmacista si rifiuti di dispe...	20a6a8c54087f3ede4d7280ec8b53c66
name-ea8744e69d689ca01a224cc06c5cc33	Key(Intent, 'aborto')	Per ottenere l'IVG con RU486 puoi rivolgerti a...	4f243e62774a59f70eb53b76389e0673
name-f9c3e3d1b21fbb8a9b5052f484e4be6d	Key(Intent, 'aborto')	Si può ricorrere all'aborto farmacologico entr...	f4245affa651050935a3a3b10044c35
name-041f4ab76d71f6898a23fda26ff56c50	Key(Intent, 'alc')	E' possibile inviare un'email a donazioni@ass...	9507877ec2fb74e092f67b1b7abf6040
name-04ff901ee729c94af9068cc0495eda11	Key(Intent, 'alc')	Tutte le persone fisiche con residenza in Itali...	b07d19516422efcae2987871108fca7

Figura 4.4: Interfaccia del Cloud Datastore su GCP

**Server-Side Encryption** : tutti i dati vengono cifrati prima di essere scritti sul disco e decodificati durante la lettura prima di renderli disponibili all'applicazione;

**Completamente gestito e senza nessun *downtime* pianificato** : l'amministrazione del servizio è gestita direttamente da Google, gli sviluppatori sono responsabili solo della logica applicativa.

La soluzione realizzata adotta il Datastore per la persistenza dei dati, il motivo è l'assenza di vincoli relazionali o necessità di ricerche complesse per le quali sarebbe stato vantaggioso l'utilizzo di un DB SQL-like.

Il vantaggio principale dato dall'utilizzo di questo prodotto, oltre quelli descritti in precedenza, è la possibilità di avere, anche in questo caso, un modello di pagamento di tipo pay-per-use senza nessun costo fisso o di gestione come nel caso dell'utilizzo di un database SQL tradizionale.

In figura 4.4 la pagina dedicata alla visualizzazione dei dati contenuti nel Datastore all'interno della console GCP.

## 4.3 L'intelligenza artificiale in CitBot

Questa sezione descrive la parte più rilevante della soluzione realizzata, quella di intelligenza artificiale e dei servizi utilizzati per l'interpretazione delle domande fatte dagli utenti e la conseguente selezione della risposta

più appropriata fra quelle a disposizione del sistema.

I servizi mostrati successivamente sono quelli utilizzati nella versione finale del chatbot, che hanno mostrato le performance migliori e garantito al contempo la scalabilità futura della soluzione. Prima della presentazione dei servizi saranno descritti due tecnologie sulle quali è basato uno dei due prodotti utilizzati.

### 4.3.1 Neural Architecture Search

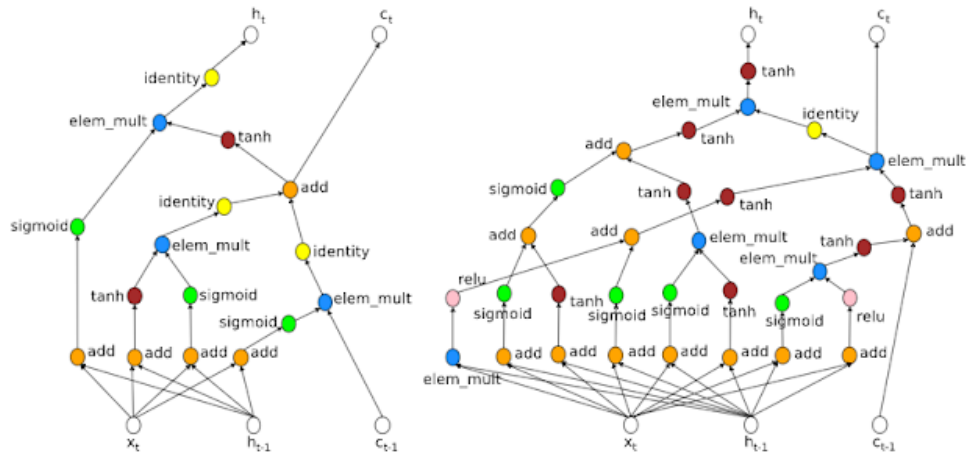
Le reti neurali sono dei modelli potenti e flessibili che hanno ottimi risultati in molti problemi di Machine Learning complessi, come la classificazione delle immagini oppure il processamento del linguaggio naturale (*Natural Language Processing*).

Nonostante le reti neurali siano riconosciute come uno strumento indispensabile per i problemi affrontati oggi giorno dall'intelligenza artificiale, rimangono comunque complesse nel design e nella realizzazione che richiede uno studio ed una preparazione tali da restringere queste attività solamente a *data scientist* di esperienza. A confermare questa teoria è il CEO di Google, Sundar Pichai, il quale ha scritto che "progettare reti neurali richiede molto tempo e richiede una competenza che ne limita l'utilizzo a piccola comunità di scienziati e ingegneri. Ecco perché abbiamo creato un approccio chiamato AutoML, che dimostra come sia possibile utilizzare reti neurali per progettare altre reti neurali."

L'ultima frase di Pichai, ovvero "utilizzare reti neurali per progettare altre reti neurali" si riferisce alla *Neural Architecture Search* o NAS. Con NAS ci si riferisce ad una tecnica per automatizzare il design delle reti neurali, in grado di avere performance simili, se non superiori in alcuni casi, a quelle delle reti neurali progettate dai *data scientist*. Le modalità con cui viene implementata questa tecnica possono essere categorizzate a seconda dello spazio di ricerca, della strategia di ricerca e del metodo utilizzato per la valutazione delle performance. Di seguito un dettaglio:

**spazio di ricerca** : definisce il tipo o i tipi di reti neurali (ANN) che possono essere realizzati e ottimizzati;

**strategia di ricerca** : definisce l'approccio utilizzato per esplorare lo spazio di ricerca;



**Figura 4.5:** A sinistra la cella originale di LSTM, a destra quella trovata tramite NAS, fonte Le and Zoph's blog post

**metodo di valutazione delle performance** : definito sulle performance di una possibile rete neurale a partire dalla sua definizione (ovvero senza averla realizzata in precedenza).

Una delle implementazione di questa teoria è stata fornita da uno degli scienziati che ha lavorato ad AutoML, Quoc Le, il quale ha sviluppato una *Recurrent Neural Network* o *RNN* per il design di nuove reti neurali utilizzando il *Reinforcement Learning* per addestrare la RNN a progettare ANN (*Artificial Neural Network*) sempre più performanti, sulla base della loro accuratezza definita su un set di dati di "valutazione".

Questa tecnica, anche se dispendiosa da un punto di vista computazione, è utile perché permette di scoprire architetture complesse, potenzialmente più complesse di quelle realizzate fino ad ora dagli scienziati, che possono essere ottimizzate per specifici problemi o determinati set di dati. L'idea che è alla base di questa tecnica è che ogni problema o set di dati ha bisogno, per raggiungere il massimo delle performance, di una specifica architettura di reti neurali.

### 4.3.2 Transfer Learning

La seconda tecnica si basa su un paradigma completamente opposto rispetto a quello visto in precedenza. Se con la *Neural Architecture Search*

l'idea è che ogni problema abbia bisogno di una specifica ANN per raggiungere le performance migliori, il *Transfer Learning* si basa sul principio inverso, ovvero che una specifica architettura di reti neurali è in grado di generalizzare su tipi di problemi simili fra loro.

Il *Transfer learning* consiste nel creare un modello di *Machine Learning* a partire da un set di dati, anche limitato, "riutilizzando" un modello già esistente e già allenato a risolvere problemi simili. In questo caso anche in presenza di pochi dati è comunque possibile raggiungere buoni risultati, il motivo è che alcune delle informazioni del problema che il nuovo modello deve risolvere possono essere "ereditate" dal modello di partenza.

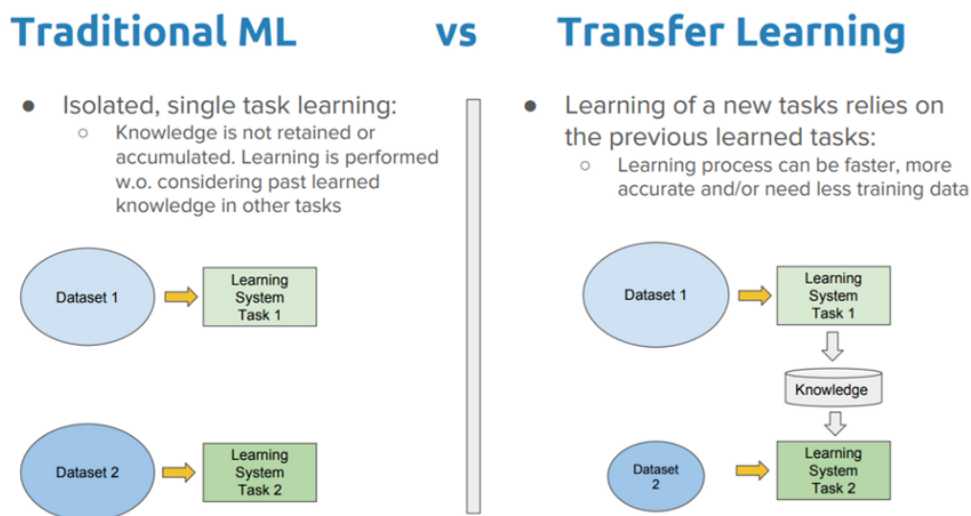
Ad esempio durante lo sviluppo di un modello per la classificazione dei Camion, invece di realizzare un nuovo modello da zero, potrebbe essere utile utilizzare le informazioni già collezionate da un precedente modello in grado di classificare con elevata precisione delle auto. Trattandosi di problematiche simili, sia in termini di dati da processare che di tipologia di problema, applicare il *transfer learning* potrebbe voler dire raggiungere delle performance sul nuovo modello simili a quelle del vecchio, anche a fronte di un numero esiguo di dati in input. Utilizzare un modello già addestrato in precedenza comporta la riduzione della capacità computazionale richiesta per l'addestramento, il nuovo modello non richiede di essere allenato con tutti i dati, ma con solo con quelli del nuovo problema.

Nella figura [4.6 nella pagina successiva](#) si evidenzia la differenza fra l'approccio "tradizionale" al *machine learning* e quello che fa uso del *transfer learning*.

### 4.3.3 Cloud AutoML

Il servizio che utilizza le tecnologie analizzate in precedenza è *Cloud AutoML*, un prodotto offerto da Google all'interno della GCP che consente di realizzare modelli di *machine learning* con performance elevate senza necessità per gli utilizzatori di avere esperienza in questo settore.

Il prodotto permette agli utenti, mediante una semplice ed intuitiva *Graphic User Interface (GUI)*, di lanciare l'addestramento di modelli complessi di classificazione in pochi minuti ed averli a disposizione per l'utilizzo in poche ore, senza necessità di ulteriori *deployments* (in figura [4.7 a pagina 28](#) viene riportato il semplice schema di utilizzo del servizio). La gestione del modello realizzato è demandata a Google, in questo modo gli



**Figura 4.6:** A sinistra l'approccio tradizionale al ML, a partire da due set di dati vengono generati due modelli. A destra un esempio di Transfer Learning in cui le informazioni acquisite dal primo modello, addestrato a partire dal set di dati numeroso, viene "trasferita" ed utilizzata per addestrare un secondo modello con un numero di dati molto inferiore, fonte Towardsdatascience

sviluppatori non devono preoccuparsi della scalabilità e dei costi relativi all'infrastruttura, il sistema di pagamento è infatti pay-per-use.

Le ottime prestazioni di *Cloud AutoML* sono motivate dal fatto che questo servizio sfrutta oltre 10 anni di tecnologia proprietaria di Google fra cui le due descritte in precedenza: *Transfer learning* e *Neural Architecture Search (NAS)*.

Il prodotto ha due concetti fondamentali, il primo è set di dati (*dataset*) mentre il secondo è il modello di *machine learning*. Per set di dati si intende un insieme di contenuti testuali ai quali sono stati associati una o più *label*, questo sarà utilizzato per durante la fase di addestramento del modello. Il dataset può essere realizzato utilizzando come sorgenti di dati da caricare sul servizio sia file csv che zip o file di testo.

A partire da un set di dati può essere addestrato un modello di *machine learning* in grado di classificare contenuto testuale simile a quello utilizzato durante l'allenamento. Il servizio permette di gestire il versionamento dei modelli di *machine learning* realizzati dallo stesso set di dati che incrementa



**Figura 4.7:** Schema di funzionamento di *Cloud AutoML*, fonte Google

nel numero di dati con il passare del tempo.

*Cloud AutoML* offre delle API e delle librerie per interagire in modo programmatico ed automatizzare le operazioni descritte in precedenza. Nello specifico è possibile demandare la classificazione di un contenuto testuale, aggiornare o creare nuovi dataset, lanciare un nuovo addestramento per un modello oppure verificare le performance di un modello.

La scelta di AutoML dipende dalle possibilità di automatizzare le procedure descritte in precedenza e di avere un prodotto in grado di offrire performance elevate, senza compromessi sulla dimensione dei dati utilizzati per l'addestramento o sul numero di tematiche gestite dal chatbot.

#### 4.3.4 Cloud Translate API

Il secondo servizio di *Machine Learning* utilizzato è quello delle Cloud Translate API. Come suggerito dal nome, questo utilizza le reti neurali già addestrate di Google per permettere la traduzione immediata di contenuto testuale fra più di 100 lingue diverse. La necessità del suo utilizzo nasce da una limitazione attuale del servizio AutoML, che al momento della scrittura è in grado di gestire (con performance elevate) solo contenuto testuale in lingua l'inglese.

Questa limitazione di AutoML rappresenta un problema in quanto, sia le domande e risposte di esempio dell'Associazione, che la lingua delle interazioni degli utenti del chatbot sono in italiano. Le Translate API sono state utili per la traduzione degli esempi forniti dall'Associazione utilizzati come dataset per l'allenamento del modello e per la conversione del testo inviato dall'utente durante la conversazione del chatbot da sottomettere al modello per la classificazione.

Una caratteristica importante delle *Cloud Translation API* è la funzionalità di determinare la lingua del testo in input, nel caso in cui questa non sia nota a priori. Sarebbe inoltre possibile utilizzare questa funzionalità per generare dinamicamente la risposta nella lingua individuata in precedenza, a partire da un template di risposta in inglese.

### 4.3.5 Cloud Text-to-speech e Speech-to-text

*Cloud Text-to-speech* e *Cloud Speech-to-text* sono due servizi complementari in grado di convertire del contenuto in formato testuale in audio e viceversa.

Seppure non utilizzati direttamente all'interno della soluzione realizzata, queste tecnologie sono comunque presenti nell'integrazione con Google Assistant. Quest'ultimo permette all'utente di interagire con il chatbot tramite comandi vocali, a seguito della domanda posta vocalmente, l'utente riceve la risposta dal chatbot tramite una delle voci configurate in Google Assistant. Questa funzionalità fa uso dei servizi descritti in precedenza, i quali supportano oltre 100 lingue, il riconoscimento automatico della lingua utilizzata dall'utente e l'identificazione e la trascrizione di nomi propri, date e numeri di telefono.

## 4.4 Addestramento dei modelli di ML

Questa sezione descrive le fasi di creazione e di addestramento dei modelli di *machine learning* utilizzati all'interno dell'applicazione.

La prima fase è quella relativa alla generazione del set di dati utilizzato per addestrare il modello di AutoML, a tal fine sono stati collezionati e processati tutti gli esempi di domande e risposte inviate dall'Associazione tramite un Google Spreadsheet. Nella figura [4.8 nella pagina successiva](#) un esempio di file compilato dall'Associazione contenente per alcune tematiche (una per foglio di lavoro).

I dati collezionati sono stati elaborati e tradotti in inglese per essere utilizzati durante l'addestramento di due tipologie di modelli con funzionalità e tipologia di classificazione diverse.

Domanda	Risposta
Devo pagare per fare le Dichiarazioni Anticipate di Trattamento (DAT)?	Le Disposizioni Anticipate di Trattamento (DAT) depositate in Comune non costano nulla. Sono esenti dall'obbligo di registrazione tributaria, dall'imposta di bollo e da qualsiasi altro tributo, imposta, diritto o tassa. Se ti rivolgi a un notaio per l'autentica, quest'ultimo può chiederti un compenso.
Devo pagare per autenticare le Dichiarazioni Anticipate di Trattamento (DAT)?	Le Disposizioni Anticipate di Trattamento (DAT) depositate in Comune non costano nulla. Sono esenti dall'obbligo di registrazione tributaria, dall'imposta di bollo e da qualsiasi altro tributo, imposta, diritto o tassa. Se ti rivolgi a un notaio per l'autentica, quest'ultimo può chiederti un compenso.
posso conservare la dat altrove oltre che all'ufficio preposto	Le Disposizioni Anticipate di Trattamento (DAT) depositate in Comune non costano nulla. Sono esenti dall'obbligo di registrazione tributaria, dall'imposta di bollo e da qualsiasi altro tributo, imposta, diritto o tassa. Se ti rivolgi a un notaio per l'autentica, quest'ultimo può chiederti un compenso.
Quanto costa le Dichiarazioni Anticipate di Trattamento (DAT)?	Le Disposizioni Anticipate di Trattamento (DAT) depositate in Comune non costano nulla. Sono esenti dall'obbligo di registrazione tributaria, dall'imposta di bollo e da qualsiasi altro tributo, imposta, diritto o tassa. Se ti rivolgi a un notaio per l'autentica, quest'ultimo può chiederti un compenso.
Devo pagare per fare le Dichiarazioni Anticipate di Trattamento (DAT)?	Le Disposizioni Anticipate di Trattamento (DAT) depositate in Comune non costano nulla. Sono esenti dall'obbligo di registrazione tributaria, dall'imposta di bollo e da qualsiasi altro tributo, imposta, diritto o tassa. Se ti rivolgi a un notaio per l'autentica, quest'ultimo può chiederti un compenso.
devo pagare?	Le Disposizioni Anticipate di Trattamento (DAT) depositate in Comune non costano nulla. Sono esenti dall'obbligo di registrazione tributaria, dall'imposta di bollo e da qualsiasi altro tributo, imposta, diritto o tassa. Se ti rivolgi a un notaio per l'autentica, quest'ultimo può chiederti un compenso.
Il biotestamento si autentica a pagamento?	Le Disposizioni Anticipate di Trattamento (DAT) depositate in Comune non costano nulla. Sono esenti dall'obbligo di registrazione tributaria, dall'imposta di bollo e da qualsiasi altro tributo, imposta, diritto o tassa. Se ti rivolgi a un notaio per l'autentica, quest'ultimo può chiederti un compenso.

Figura 4.8: Domande e risposte inviate dall'Associazione Luca Coscioni

#### 4.4.1 Modello per la classificazione della tematica

Il primo modello realizzato è in grado di associare un quesito fatto dall'utente ad una tematica. I dati utilizzati per l'addestramento sono stati generati a partire dai dati forniti dall'Associazione utilizzando la traduzione della domanda di esempio e un identificativo dell'area tematica come *label* associata. Nella figura 4.9 a fronte è stato riportato un esempio di set di dati utilizzati per l'addestramento del modello.

#### 4.4.2 Modello per la classificazione delle domande di un argomento

La seconda tipologia di modelli realizzati sono responsabili di individuare la risposta per una determinata domanda, associandole la *label* più appropriata sulla base dei dati con i quali sono stati addestrati.

Nella soluzione sono stati realizzati ed utilizzati un modello specifico per ogni area tematica fornita dall'Associazione, utilizzata come *label* nel precedente modello. Ogni modello è stato allenato utilizzando uno specifico set di dati contenente le domande per quell'area, associate ad una *label* che in questo caso è un identificativo univoco della risposta per quella domanda.



TEXT ITEMS	TRAIN	EVALUATE	PREDICT
All texts	1643		
Labeled	1643		
Unlabeled	0		

Text	Label
ok, but can the municipality refuse to register the advance treatment declarations (dat)?	dat
is it normal not to accept the early treatment declarations (dat) in the municipality?	dat
i already told you, i don't want to make a bequest! i want to know if: having established the fact that i am a resident, the registrar does not agree to register my dats in common. do you know it or not?	dat
my question is: having verified that i am resident in the municipality, the registrar may not accept the advance declarations of treatment (dat) and not register them?	dat
if my municipality does not comply with the early declarations of treatment (dat), what can i do?	dat
i do not want to fill in a new living will, i want to know if it is legal not to accept the advance declarations of treatment (dat) to the municipality of residence	dat
i went in common, but they didn't accept my living will	dat
in common they refuse me the living will, what should i do?	dat
can the municipality not be equipped for early statements of treatment (dat)?	dat
i want to distrust the municipality, because they don't accept my dat!	dat

**Figura 4.9:** Domande e rispettive label utilizzate per l'addestramento del primo modello utile alla classificazione della tematica

L'identificativo utilizzato durante la classificazione delle domande è poi utilizzato dall'applicazione realizzata su AppEngine per recuperare la risposta da fornire all'utente. L'utilizzo di quest'ultimo permette di svincolare il processo di classificazione della domanda, effettuato dal modello, dalla risposta da inviare all'utente. In questo modo, anche a fronte di un aggiornamento del testo della risposta, il modello di classificazione può essere lasciato invariato (cambierà solamente il testo all'interno dell'applicativo). L'addestramento del modello rimane così necessario solo nel caso di inserimento di nuove domande, siano esse relative a risposte nuove oppure già presenti all'interno del modello. Nella figura 4.10 nella pagina successiva è stato riportato un esempio di set di dati utilizzati per l'addestramento del modello sul testamento biologico e fine vita.

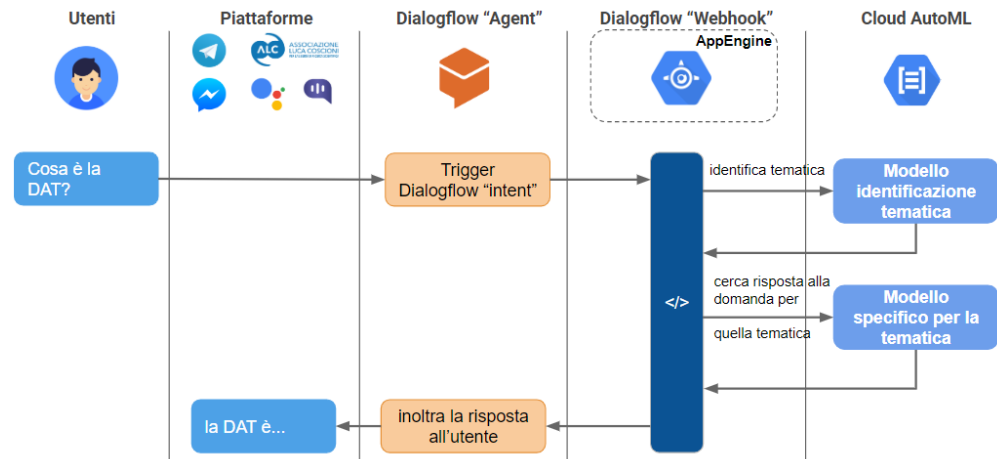
## 4.5 Flusso della conversazione

Questa sezione tratta la logica realizzata per la gestione dell'interazione fra gli utenti ed il chatbot, in figura 4.5 a pagina 25 il flusso dei dati e delle logiche realizzate per il funzionamento di CitBot.

Lo schema riportato è generale e si applica a tutte le piattaforme con le quali è stata predisposta un'integrazione (Telegram, Facebook Messenger, ecc.), ed è relativo ad una interazione completa fra l'utente ed il chatbot. A seguire il dettaglio del flusso riportato in figura:

Text	Label
Can the trustee accept the appointment at a later time?	83359d11ca1f20c676ade001000d7fb0
I want to die! Can I write it in the living will?	fcce2b1996c80885fe65415ad5b4bdc2
What happens if I write that I want euthanasia in my dat?	fcce2b1996c80885fe65415ad5b4bdc2
Does this will also include euthanasia? if yes, I am not interested.	fcce2b1996c80885fe65415ad5b4bdc2
Can I write in the Advance Declarations of Treatment (DAT) that I want assisted suicide?	fcce2b1996c80885fe65415ad5b4bdc2
Can I write that I want to die in the Early Statements of Treatment (DAT)?	fcce2b1996c80885fe65415ad5b4bdc2
What happens if I write that I want to die in my living will?	fcce2b1996c80885fe65415ad5b4bdc2
Can I write in the Advance Declarations of Treatment (DAT) that I want euthanasia?	fcce2b1996c80885fe65415ad5b4bdc2
If I write that I want euthanasia (assisted suicide) will the Advance Declarations of Treatment (DAT) be canceled?	fcce2b1996c80885fe65415ad5b4bdc2
If I write that I want to die in the bioassay, am I invalidating it?	fcce2b1996c80885fe65415ad5b4bdc2

**Figura 4.10:** Domande e rispettive label utilizzate per l'addestramento del secondo modello per la classificazione delle risposte



**Figura 4.11:** Dettaglio dei flussi logici con i quali CitBot è in grado di rispondere alle domande degli utenti

1. L'utente, accedendo ad una delle piattaforme integrate, è in grado di porre una domanda al sistema, in alcuni casi utilizzando i comandi vocali;
2. La domanda posta dall'utente (eventualmente trascritta mediante Cloud Speech-to-text API) viene inviata attraverso la specifica piattaforma integrata a Dialogflow;
3. Dialogflow riceve in input la domanda posta dall'utente e invia il contenuto, insieme ad alcuni metadati come la sorgente della conversazione e la sessione dell'utente, ad un determinato *Endpoint* sul quale sarà in ascolto l'applicazione web sviluppata su piattaforma *AppEngine*;
4. L'applicazione è responsabile di tutta la logica applicativa della soluzione realizzata, a valle dell'elaborazione della richiesta dell'utente verrà restituita la risposta all'utente finale seguendo le logiche ed il flusso descritto in seguito;
5. La domanda posta al chatbot è tradotta in lingua inglese facendo uso delle *Cloud Translate API*, in questo modo il testo può essere utilizzato come input per i modelli di AutoML analizzati in precedenza;
6. A seguito della traduzione della domanda viene effettuata una prima classificazione attraverso l'utilizzo del primo modello di AutoML per riconoscere l'area tematica di riferimento. A seconda della classificazione effettuata da AutoML si possono verificare due casi diversi:
  - (a) La risposta ricevuta da AutoML viene processata dall'applicazione per determinare a quale delle aree tematiche fa riferimento la domanda ricevuta, in questo caso è stata predisposta una soglia di confidenza configurabile al di sotto della quale la classificazione effettuata da AutoML non viene considerata attendibile. In caso di mancato riconoscimento della tematica viene fornita una risposta di *default* all'utente informandolo del fatto che il sistema non è riuscito ad individuare una risposta adeguata per quella domanda, invitandolo a modificare o riformulare la domanda fatta in precedenza;
  - (b) Nel caso in cui sia stato possibile determinare con precisione elevata l'area tematica della domanda viene fatta una seconda

chiamata ad un secondo modello di AutoML che in questo caso è specifico per la tematica. Anche in questo caso sono possibili due scenari diversi:

- i. Anche in questo caso la risposta ricevuta da AutoML viene processata dall'applicazione al fine di individuare la risposta con una confidenza maggiore. In questo caso una seconda soglia è stata predisposta in modo che una risposta di *default* venga fornita all'utente in caso non sia stata raggiunta una precisione elevata nella classificazione. La risposta fornita a questo punto è diversa da quella precedente in quanto la tematica è nota, si possono quindi fornire all'utente esempi di domande da sottoporre al chatbot relative a quell'argomento;
- ii. Nel caso in cui anche la seconda classificazione ha avuto esito positivo (quindi con una confidenza nella classificazione superiore alla soglia definita) è possibile restituire la risposta finale all'utente come *response* opportunamente formattata per Dialogflow;

7. Dialogflow è responsabile di inviare la risposta all'utente;
8. La risposta sarà poi visualizzata in maniera differente a seconda della piattaforma utilizzata per la comunicazione con il chatbot.

Un caso particolare rispetto a quello descritto in precedenza è quello di una prima interazione con il chatbot determinata non da una domanda posta dall'utente, ma da un *trigger* automatico come l'apertura della chat di Messenger o l'apertura automatica del widget del chatbot sul sito ufficiale dell'Associazione. Questi eventi sono gestiti tramite specifiche logiche che specifici messaggi che vengono inviati da Dialogflow all'applicazione a fronte di questi eventi e che permettono di riconoscere questo tipo di interazione per fornire una risposta quanto più adeguata al contesto. Nel caso specifico un messaggio di presentazione e di breve introduzione è stato configurato per aiutare l'utente a comprendere come utilizzare il servizio e con quali finalità. Un esempio di questo messaggio è visibile nelle figure [4.14 a pagina 37](#) e [4.15 a pagina 37](#) nella sezione successiva.



**Figura 4.12:** Chatbot integrato all'interno della pagina facebook dell'associazione

## 4.6 Piattaforme integrate

Il successo di un chatbot dipende, oltre che dalle funzionalità offerte, anche dalla facilità di accesso e di utilizzo per gli utenti e dalla numerosità dei canali dai quali è raggiungibile. La scelta di utilizzare Dialogflow come orchestratore dei canali di ingresso delle comunicazioni verso CitBot ha facilitato l'integrazione dell'applicativo con molte delle più famose piattaforme di messaggistica attualmente disponibili in commercio.

In pochi minuti è possibile integrare il proprio chatbot con Google Assistant, Facebook Messenger, Slack, Skype, Telegram, Twitter, Twilio ed altre applicazioni. La soluzione attuale è stata resa disponibile sulle seguenti piattaforme di messaggistica:

- Facebook Messenger [4.12](#)
- Telegram [4.13](#) nella pagina seguente

Sono già previste integrazioni future con Google Assistant e WhatsApp Business, la prima è già in approvazione da parte di Google e sarà disponibile prossimamente.

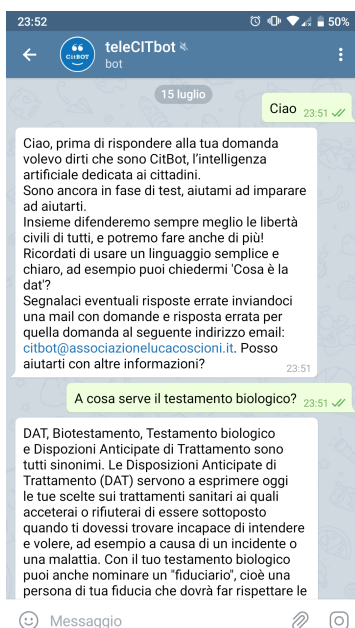


Figura 4.13: Immagine del chatbot integrato all'interno di Telegram

## 4.7 Integrazioni su siti Web

La scelta dell'utilizzo di Kommunicate per la gestione del widget relativo alla comunicazione con l'utente è stata necessaria vista la necessità dell'associazione di integrare il chatbot all'interno del proprio sito web e le limitazioni imposte dall'utilizzo del widget di *default* di Dialogflow (ad esempio non sono gestiti i pulsanti o i collegamenti ipertestuali). Un secondo sito web dedicato è stato realizzato e reso disponibile agli utenti sia in ambiente di produzione che di test, dove alcune persone legate all'associazione hanno avuto la possibilità di testare le nuove funzionalità o tematiche prima del rilascio sulla versione utilizzata dagli utenti finali.

Di seguito viene riportata l'immagine 4.14 nella pagina successiva del sito ufficiale realizzato per CitBot.

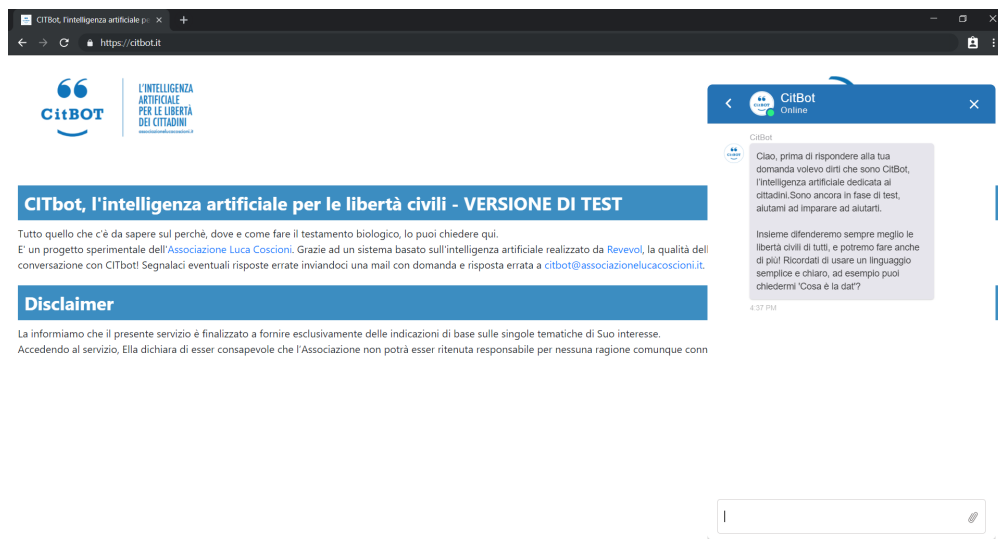


Figura 4.14: Sito web dedicato al progetto CitBot

A seguire una seconda immagine 4.15 relativa all'integrazione del chatbot all'interno di alcune sezioni del sito dell'Associazione.



Figura 4.15: Integrazione del chatbot sul sito web dell'Associazione Luca Coscioni





# Capitolo 5

## Valutazione

In questo capitolo sarà presentata una valutazione della soluzione realizzata in due sezioni distinte, la prima tratta l'analisi oggettiva di alcuni dei modelli di *Machine Learning* realizzati, la seconda tratta il giudizio finale dato dall'Associazione e dagli utenti finali utilizzatori del chatbot.

### 5.1 Valutazione dei modelli

La prima tipologia di valutazione effettuata si riferisce all'accuratezza della classificazione effettuata dai modelli addestrati utilizzando AutoML. Le analisi delle *performance* del modello realizzato si basano sulle informazioni disponibili nell'apposita sezione di *Evaluation* presente all'interno della *suite* di AutoML. In seguito della fase dell'addestramento è disponibile una sezione apposita dove è possibile analizzare i risultati ottenuti. A seguire vengono riportati i risultati di due modelli utilizzati rispettivamente per identificare la tematica della domanda dell'utente e per la classificazione delle domande relative al testamento biologico.

In entrambi i casi le performance del modello sono state valutate sulla base delle metriche descritte in seguito, l'obiettivo non è fornire una definizione ed una spiegazione esaustiva ma semplicemente fornire una panoramica sugli strumenti di valutazione del modello di *machine learning* utilizzati in questo capitolo. Sono requisiti per affrontare la definizioni delle metriche le definizioni dello *score threshold* e la categorizzazione delle predizioni

effettuate dal modello durante l'addestramento, riportate sotto.

*Score Threshold* si riferisce al livello di "confidenza" richiesta dal modello per l'assegnazione di una categoria (o *label*) ad una frase in ingresso. Questa è un parametro per la valutazione delle due metriche descritte successivamente, in generale aumentando la *threshold* aumenta conseguentemente la *precisione* del modello a scapito della *recall* e viceversa. Il valore di *threshold* utilizzato a regime dal modello dipende dal tipo di problema che deve essere affrontato, nel caso in cui il requisito è di evitare record classificati erroneamente la *threshold* sarà elevata, mentre sarà inferiore nel caso in cui è importante individuare il maggior numero di record classificati con una *label* specifica, anche a scapito del numero più elevato di errori.

A seguito della definizione della *score threshold* è possibile suddividere le classificazioni effettuate dal modello su un set di dati di test (di cui si conoscono le *label* associate), per una specifica *label* in esame, nelle seguenti categorie:

- *True positive* (tp) frasi per cui il modello ha correttamente associato la *label* corrispondente
- *True negative* (tn) frasi per cui il modello non ha giustamente associato la *label* in esame
- *False positive* (fp) frasi per cui il modello ha erroneamente associato la *label* in esame
- *False negative* (fn) frasi per cui il modello avrebbe dovuto associare la *label* che, erroneamente, non è stata assegnata

Nella figura 5.1 [nella pagina successiva](#) relativa alla matrice di confusione in cui vengono le categorie descritte in precedenza.

Di seguito l'elenco delle metriche utilizzate nella misurazione delle performance del modello:

- *Precision* la definizione di precisione è la seguente:

$$Precision = \frac{tp}{tp + fp}$$

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

**Figura 5.1:** Matrice di confusione

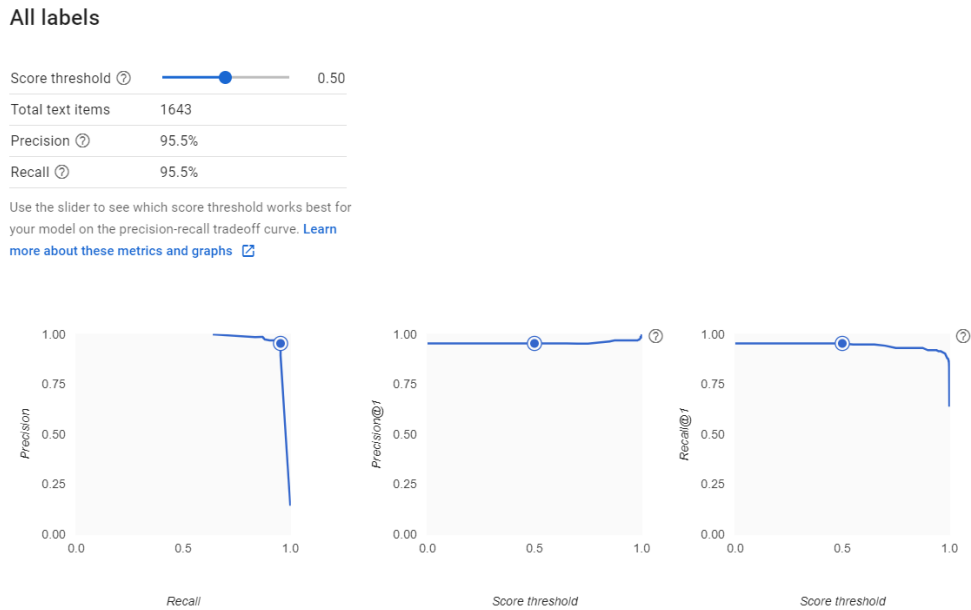
La precisione indica, rispetto a tutti gli input che sono stati classificati con una label, quelli che sarebbero dovuti essere effettivamente associati a quest'ultima.

- *Recall* la definizione è la seguente:

$$Recall = \frac{tp}{tp + fn}$$

La *recall* indica, rispetto a tutti gli input che sarebbero dovuti essere associati ad una *label*, quanti effettivamente lo sono stati.

- *Average Precision* indica l'area definita sotto la curva relativa al grafico della *precision-recall*. Questo grafico viene costruito a partire da tutti i valori di *precision* e *recall* per ogni possibile *threshold* fra 0 e 1. Questo indicatore permette di capire le performance del modello su tutte le possibili soglie. Più questo valore è vicino ad 1 e più il modello è efficace sul set di dati di test, un modello random ha una *average precision* intorno allo 0.5.



**Figura 5.2:** Precision e Recall del modello di classificazione delle aree tematiche

### 5.1.1 Valutazione modello per la classificazione delle aree tematiche

Nelle figure 5.2 e 5.3 nella [pagina successiva](#) vengono riportati i risultati del modello in esame. Nella prima immagine sono mostrati i risultati prendendo come riferimento 0.5 come *score threshold*, questo valore permette di massimizzare sia la *precision* che la *recall* che raggiungono la stessa percentuale del 95.5%. In questo caso il risultato è particolarmente positivo perché le tipologie di domande fornite per le aree tematiche sono diverse fra loro e permettono all'algoritmo di identificare con precisione elevata le parole chiave all'interno del testo che influiscono maggiormente sull'assegnazione della specifica area tematica.

Nella seconda immagine viene riportata la matrice di confusione ottenuta durante l'addestramento del modello, dalle informazioni contenute nella matrice si intuisce come la precisione del modello sia elevata (la maggior parte delle percentuali di assegnazioni avvenute con successo presenti sulla diagonale secondaria sono vicine al 100%). L'unica problematica riscontrata riguarda le domande dell'area tematica di carattere generale che nel 33.3% dei casi vengono erroneamente assegnate alla *label* del Biotestamento.

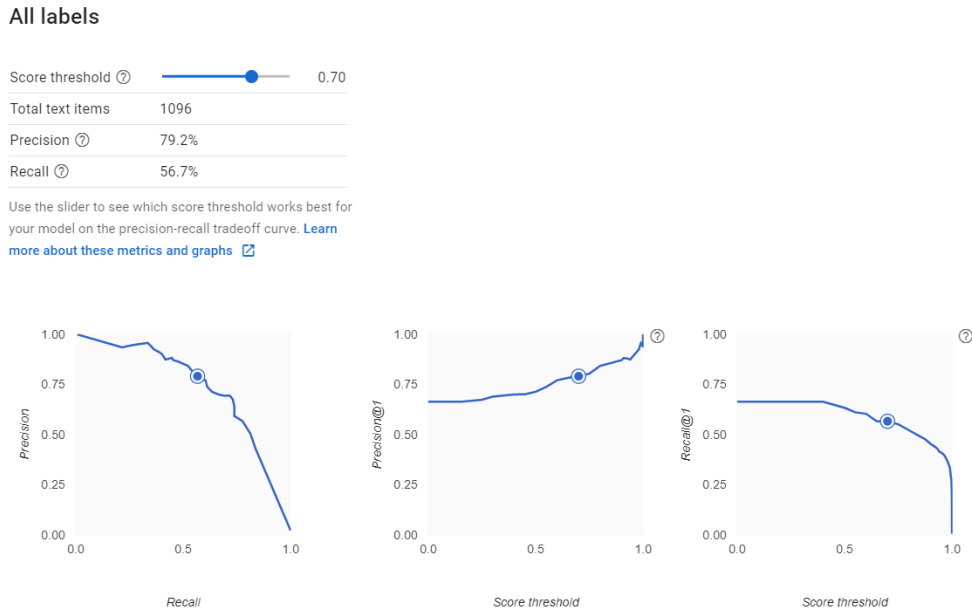
True label	Predicted label						
	aborto	cannabis	dat	disabilita	eutanasia	fecondazione	generali
aborto	86.4%	-	13.6%	-	-	-	-
cannabis	-	100.0%	-	-	-	-	-
dat	-	-	99.1%	-	0.9%	-	-
disabilita	-	-	14.3%	71.4%	-	-	14.3%
eutanasia	-	-	-	-	100.0%	-	-
fecondazione	-	-	-	-	-	100.0%	-
generali	-	-	33.3%	-	-	-	66.7%

**Figura 5.3:** Matrice di confusione per il modello della classificazione della tematica

### 5.1.2 Valutazione modello per la classificazione delle risposte sul Biotestamento

Nelle figure 5.4 nella pagina seguente e 5.5 a pagina 45 vengono mostrati i risultati del secondo modello realizzato per il testamento biologico. Come in precedenza, nella prima immagine vengono mostrati i risultati del modello utilizzando 0.7 come valore di riferimento della *score threshold*, in questo modo è possibile ottenere una *precision* del 79.2% ed una *recall* del 56.7%. Rispetto al risultato precedente si nota immediatamente una riduzione sia della precisione che della *recall*, l'80% nella precisione è stato raggiunto incrementando il valore dello *score threshold* a 0.7, a scapito della *recall* che è di poco superiore al 50%.

Durante lo sviluppo della soluzione sono state realizzate diverse versioni di questo modello, allenate su un set di dati via via sempre di dimensioni maggiori, l'incremento nel numero di risposte differenti ha contribuito negativamente alle performance del modello. Consultando le informazioni presenti nella matrice di confusione nella seconda figura è stato possibile individuare problemi all'interno del set di dati utilizzato. In alcuni casi conoscere sia l'insieme di domande classificato in maniera errata che la *label* alle quali venivano assegnate, ha contribuito a far emergere problematiche nel dataset come insiemi di domande che potevano essere "uniti" a formare una nuova risposta unica e, in generale, ad una sua pulizia rispetto ai dati



**Figura 5.4:** Precision e Recall del modello sul biotestamento

iniziali.

La diminuzione delle performance, rispetto al modello precedente, è legato al fatto che le domande utilizzate per l'addestramento sono spesso composte dalle stesse parole che, in relazione al contesto, assumono significati diversi rappresentando un ulteriore fattore di complessità per il modello.

Rispetto al caso precedente la matrice di confusione, in questo caso, risulta più complessa da analizzare perché la *label* utilizzata non è significativa, come accennato in precedenza si tratta di un identificativo univoco di una risposta. Per questo modello sono state effettuate analisi dettagliate solo nei casi in cui le percentuali di errore sono state significative.

## 5.2 Valutazione da parte dell'Associazione e degli utenti

Questa sezione descrive la valutazione fatta dall'Associazione sulla soluzione realizzata e i giudizi collezionati dagli utenti utilizzatori del chatbot.

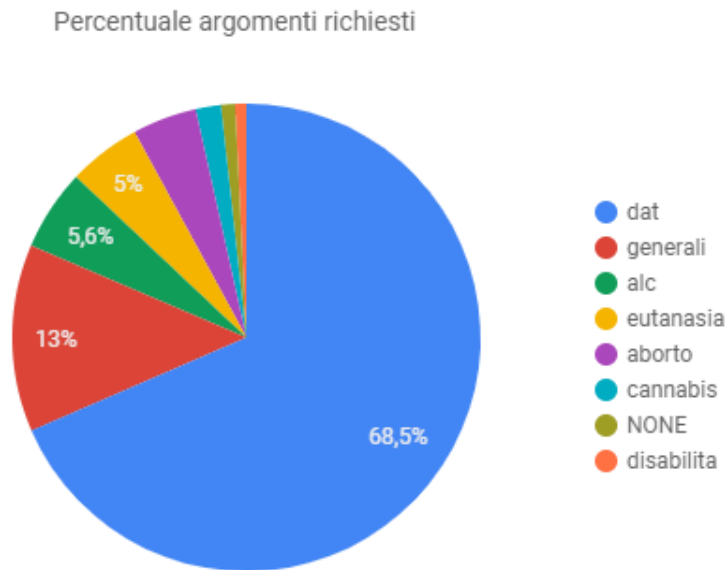
True label	Predicted label	08678ec45a1fede32...	090715c644e1bad432...	29a8eca97a3c9fa49f8547ce8f766fe4	499e1178b1e4543ef74213fa6b0219dd	4e5754bc6fe42165472891372c2e8182	51ebf9e86a9c895892bce3219f0a4d53	5c1e97b3f82cb9f00d18dfc57164fb9	6ac5a756a315fb4a56988adb5640278	f3ad4750ba89635e055c51311ab364dd	fcce2b1996c80885fe65415ad5b4bdc2
08678ec45a1fede32b47109b6597b0a	50.0%	-	-	-	-	50.0%	-	-	-	-	-
090715c644e1bad43bc885b70f5aff1	-	50.0%	-	-	-	-	-	-	50.0%	-	-
29a8eca97a3c9fa49f8547ce8f766fe4	-	-	100.0%	-	-	-	-	-	-	-	-
499e1178b1e4543ef74213fa6b0219dd	-	-	-	100.0%	-	-	-	-	-	-	-
4e5754bc6fe42165472891372c2e8182	-	-	-	33.3%	50.0%	-	-	16.7%	-	-	-
51ebf9e86a9c895892bce3219f0a4d53	-	-	-	-	11.1%	88.9%	-	-	-	-	-
5c1e97b3f82cb9f00d18dfc57164fb9	-	-	5.9%	5.9%	-	5.9%	82.4%	-	-	-	-
6ac5a756a315fb4a56988adb5640278	-	-	-	-	-	66.7%	-	33.3%	-	-	-
f3ad4750ba89635e055c51311ab364dd	-	-	-	-	-	-	-	-	100.0%	-	-
fcce2b1996c80885fe65415ad5b4bdc2	-	-	-	100.0%	-	-	-	-	-	-	-

**Figura 5.5:** Matrice di confusione per il modello di classificazione sul biotestamento

La valutazione effettuata dall'associazione è stata effettuata su un set di domande e risposte collezionate dal chatbot durante una delle sessioni di test realizzata con un insieme di utenti selezionati, sostenitori dell'associazione, che si sono messi a disposizione effettuando delle interazioni e dei quesiti al sistema. Le domande e le conseguenti risposte sono state collezionate in un Google spreadsheet e sottoposte alla verifica delle persone dell'associazione che hanno indicato ogni risposta come:

- Corretta: qualora la risposta fornita dal chatbot fosse stata adeguata alla domanda;
- Errata: qualora la risposta fornita dal chatbot non fosse adeguata e ne esistesse, all'interno del set di riposte, una più corretta per la domanda posta dall'utente;
- Mancante: qualora la risposta fornita dal chatbot non fosse adeguata, ma non fosse presente nessuna risposta ritenuta valida per la domanda effettuata (sarebbe quindi stato necessario aggiungere una nuova risposta per quella domanda).

Quelle identificate come mancanti sono state successivamente analizzate per determinare la necessità o meno dell'inserimento di una nuova tipologia



**Figura 5.6:** Percentuale di domande relative alle diverse aree tematiche

di risposta per quella domanda. Le risposte corrette ed errate sono state utilizzate per definire la percentuale di successo delle interazioni avvenute con il chatbot, nello specifico di una delle tante fasi di test avvenute che ha collezionato 517 quesiti posti dagli utenti sono stati raggiunti i seguenti risultati:

- 83.9% di aree tematiche identificate correttamente
- 82% di risposte fornite con successo agli utenti

Complessivamente il chatbot è stato in grado di rispondere in modo corretto a circa 357 risposte sulle 517 interazioni totali, delle risposte errate buona parte erano quelle fornite dal chatbot a "default" e non risposte completamente errate o forvianti. Sono stati riportati alcuni grafici nelle figure 5.6 e 5.7 riguardanti la fase di test appena discussa. Come evidenziato nel grafico in figura 5.6 la maggior parte delle domande rivolte dagli utenti al chatbot sono relative all'area tematica del testamento biologico, questa proporzione non è casuale in quanto dipende dalla volontà dell'Associazione di porre l'attenzione inizialmente su questo tema.

Durante la fase di test sono emersi molti *feedback* da parte degli utenti che ne hanno sia elogiato alcune funzionalità, che messo in evidenza





**Figura 5.7:** Percentuali di successo nel riconoscimento delle aree tematiche e delle risposte

delle criticità. I feedback positivi ricevuti dagli utenti riguardano, in generale, un apprezzamento per questa "entità" in grado di rispondere a domande per le quali regolamentazione e risposte sono difficili da recuperare online. Alcuni utenti hanno apprezzato la possibilità di poter interagire direttamente attraverso una chat, che la maggior parte degli utenti utilizza quotidianamente. Alcuni ringraziamenti sono stati ricevuti dal chatbot durante la conversazione con gli utenti a seguito di una o più risposte esaustive.

La maggiore criticità che è emersa durante le ultime fasi di test con le persone dell'associazione è la mancanza di una introduzione da parte del chatbot sulle aree tematiche e sulle tipologie di domande alle quali è in grado di rispondere. Questa problematica è stata affrontata prima del rilascio finale in produzione modificando alcune risposte fornite da CitBot ed aggiungendo una breve frase di apertura che informa l'utente sulle tematiche gestite, enunciando anche alcune frasi di esempio come: "Cosa è la DAT?" o "A cosa serve il testamento biologico?".

Come ulteriore evidenza del successo del progetto e della sua effettiva utilità nella difesa dei diritti dei cittadini e delle loro libertà viene riportato un articolo [11] pubblicato da Rosita Rijtano sul sito Repubblica.it in cui si descrive il progetto realizzato, gli scopi del progetto ed in particolare una serie di quesiti sottoposti al chatbot ai quali quest'ultimo è stato in grado di rispondere. Nell'articolo viene evidenziato come a fronte di alcune domande come "L'eutanasia è possibile in Italia?" o "cosa sono i DAT?", il chatbot sia stato in grado di rispondere correttamente ed esaustivamente.



# Capitolo 6

## Conclusioni

La realizzazione di CitBot è stata un'esperienza importante sia da un punto di vista personale che da quello aziendale. La soluzione realizzata è un esempio di come oggi sia possibile realizzare sistemi tecnologici all'avanguardia, anche per quanto riguarda tematiche come il *machine learning* e il *natural language processing*, senza essere necessariamente esperti del settore.

La seconda tematica importante riguarda i costi di realizzazione. Sebbene lo sviluppo di un chatbot "*custom*" (senza l'utilizzo di un framework esistente) possa sembrare un progetto complesso ed oneroso, questo elaborato ha dimostrato come, sfruttando servizi e prodotti messi a disposizione dai grandi *cloud provider* come Google, sia possibile realizzare una soluzione completa in settimane, e non in mesi o anni.

L'arco temporale del progetto ha infatti coperto un periodo di circa 6 / 7 mesi fra i primi contatti avuti con l'associazione e la messa in produzione di CitBot. In questi mesi solo una minima parte ha interessato lo sviluppo vero e proprio, mentre la maggior parte del tempo è stato investito inizialmente nella creazione di un set di dati (partendo dalle FAQ presenti sul sito dell'associazione) e, successivamente, nelle numerose fasi di test. Queste ultime sono avvenute con gli utenti dell'associazione e hanno consentito la realizzazione di un dataset importante, sul quale è basato il sistema attuale.

A valle della messa in produzione, sono state pianificate modifiche e ulteriori funzionalità che saranno implementate in futuro. A seguire una descrizione

delle attività concordate:

- *Gestione del multilingua da parte del chatbot*: introdurre la possibilità di identificare, attraverso l'utilizzo di API offerte da Google, la lingua del testo inserito dall'utente utilizzatore del servizio per rispondere nella lingua dell'utente, ove possibile, o utilizzando l'inglese come "default";
- *Sentiment Analysis*: introdurre l'analisi del *sentiment* dell'utente che sta interagendo con il chatbot in modo da riuscire a far intervenire un membro dell'associazione qualora dovesse essere riscontrato un *sentiment* particolarmente negativo. Questo poiché alcune delle tematiche, come l'eutanasia e l'aborto, sono particolarmente delicate e può essere necessario l'intervento di una persona;
- *Integrazione con MyDonor*: myDonor è il CRM (*Customer relationship management*) utilizzato dall'associazione, in questo caso lo scopo è di personalizzare l'esperienza utente grazie alle informazioni che si dispongono già sulla persona che sta interagendo con il chatbot;
- *Estensione del chatbot a nuove aree tematiche*: aumentare il numero delle aree gestite da CitBot e la quantità di risposte fornite per le stesse;
- *Sviluppo di un'applicazione web per la diretta gestione delle aree tematiche*: rispetto all'espansione delle aree discussa nel punto precedente, la fase di aggiornamento delle domande e delle risposte relative alle nuove tematiche richiede una serie di operazioni in carico a Revevol. La nuova funzionalità consiste nella realizzazione di una applicazione web dedicata che permetta all'associazione la gestione autonoma di CitBot in termini di: modifica delle risposte alle domande esistenti, aggiunta di nuove domande e risposte, inserimento di nuove tematiche e nuove lingue, addestramento di nuovi modelli di *machine learning*, test delle performance dei nuovi modelli e rilascio delle nuove versioni di CitBot in produzione.

# Bibliografia

- [1] Valerio Boffino. *Tra utopia e futuro: Chatbot e Natural Language Processing*. 2019. URL: [https://botmanism.altervista.org/tra-utopia-e-futuro-chatbot-e-natural-language-processing/?doing\\_wp\\_cron=1565512094.2370479106903076171875](https://botmanism.altervista.org/tra-utopia-e-futuro-chatbot-e-natural-language-processing/?doing_wp_cron=1565512094.2370479106903076171875).
- [2] Associazione Luca Coscioni. *Associazione Luca Coscioni*. URL: <https://www.associazionelucacoscioni.it/cosa-facciamo/>.
- [3] Nikki Gilliland. *Five examples of charity chatbots*. 2018. URL: <https://econsultancy.com/five-examples-of-charity-chatbots/>.
- [4] Google. *AutoML Natural Language Beginner's guide*. 2019. URL: <https://cloud.google.com/natural-language/automl/docs/beginners-guide#evaluate>.
- [5] Google. *Dialogflow Enterprise Edition*. URL: <https://cloud.google.com/dialogflow-enterprise/>.
- [6] Business Insider Intelligence. *THE MESSAGING APPS REPORT: Messaging apps are now bigger than social networks*. 2016. URL: <https://www.businessinsider.com/the-messaging-app-report-2015-11?IR=T>.
- [7] Stefan Kojouharov. *Ultimate Guide to Leveraging NLP & Machine Learning for your Chatbot*. 2016. URL: <https://chatbotslife.com/ultimate-guide-to-leveraging-nlp-machine-learning-for-you-chatbot-531ff2dd870c>.
- [8] Mencap. *Mencap: the voice of learning disability*. URL: <https://www.mencap.org.uk/get-involved/campaign-mencap/here-i-am>.
- [9] Sundar Pichai. *Making AI work for everyone*. 2017. URL: <https://www.blog.google/technology/ai/making-ai-work-for-everyone/>.

- [10] Allison Ragan. *Taking the Confusion Out of Confusion Matrices*. 2018. URL: <https://towardsdatascience.com/taking-the-confusion-out-of-confusion-matrices-c1ce054b3d3e?gi=d27279682f47>.
- [11] Rosita Rijitano. «CitBOT, il chatbot italiano che risponde alle domande sul testamento biologico». In: *Repubblica* (2019). URL: [https://www.repubblica.it/tecnologia/2019/06/05/news/citbot\\_il\\_chatbot\\_italiano\\_paladino\\_dei\\_diritti\\_risponde\\_alle\\_domande\\_sul\\_testamento\\_biologico-228038439/](https://www.repubblica.it/tecnologia/2019/06/05/news/citbot_il_chatbot_italiano_paladino_dei_diritti_risponde_alle_domande_sul_testamento_biologico-228038439/).
- [12] Madhu Sanjeevi. *Chapter 11: ChatBots to Question & Answer systems*. 2018. URL: <https://medium.com/deep-math-machine-learning-ai/chapter-11-chatbots-to-question-answer-systems-e06c648ac22a>.
- [13] Dipanjan Sarkar. *A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning*. 2018. URL: <https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a>.
- [14] Shelter Scotland. *Private Residential Tenancies Scotland - New House Rules*. 2017. URL: <https://scotland.shelter.org.uk/newhouserules>.
- [15] Kumar Shridhar. *Generative Model Chatbots*. 2017. URL: <https://medium.com/botsupply/generative-model-chatbots-e422ab08461e>.
- [16] Rachel Thomas. *An Opinionated Introduction to AutoML and Neural Architecture Search*. 2018. URL: <https://www.fast.ai/2018/07/16/auto-ml2>.
- [17] Rachel Thomas. *Google's AutoML: Cutting Through the Hype*. 2018. URL: <https://www.fast.ai/2018/07/23/auto-ml-3/>.
- [18] Wikipedia contributors. *Google App Engine*. [Online]. 2019. URL: [https://it.wikipedia.org/wiki/Google\\_App\\_Engine](https://it.wikipedia.org/wiki/Google_App_Engine).
- [19] Wikipedia contributors. *Neural architecture search*. [Online]. 2019. URL: [https://en.wikipedia.org/wiki/Neural\\_architecture\\_search](https://en.wikipedia.org/wiki/Neural_architecture_search).
- [20] Barret Zoph e Quoc V. Le. «Neural Architecture Search with Reinforcement Learning». In: 2017. URL: <https://arxiv.org/abs/1611.01578>.