ARTICOLO DI ELEARNINGNEWS

Year 4 - number 134 Wednesday 11 november 2020

DynDevice LMS e l'Intelligenza Artificiale: DyBot

Pubblicato su rivista scientifica internazionale il paper sul nuovo chatbot di DynDevice LMS

L'impegno profuso da Mega Italia Media nello studio e l'applicazione dell'**Intelligenza Artificiale (A.I.)** nella formazione online tramite la **piattaforma eLearning DynDevice LMS** ha portato alla pubblicazione del paper "**Evaluating different Natural Language Understanding services in a real business case for the Italian language**" (Valutazione di diversi servizi di comprensione del linguaggio naturale in un vero e proprio business case per la lingua italiana) sulla rivista internazionale di settore Procedia Computer Science.

Il paper è stato presentato il 16-18 Settembre 2020 alla conferenza internazionale "24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems".

L'esperienza, sviluppata grazie alla stretta collaborazione tra il **dipartimento di Ingegneria Informatica dell'Università degli Studi di Brescia** (UNIBS) e Mega Italia Media Spa, eLearning company di Castel Mella (BS), è legata alla creazione di un chatbot per DynDevice LMS: **DyBot**. DyBot è stato concepito per dare una risposta automatica ai questiti degli utenti durante l'erogazione di corsi eLearning.

Cos'è un Chatbot?

Abbiamo già parlato di chatbot in diverse occasioni, ad esempio qui, qui e qui.

Il termine chatbot deriva dall'unione di due parole: "chat" (sistema di messaggistica istantanea) e "bot" (diminutivo del termine "robot"); quest'ultimo termine è generalmente utilizzato per definire una macchina, o nel caso dell'informatica un software, in grado di automatizzare uno o più compiti.

Quindi un chatbot è un software che permette di conversare in modo autonomo con una persona attraverso un sistema testuale. Generalmente questo tipo di software si appoggia a complessi algoritmi di Machine learning e Intelligenza artificiale al fine di interpretare le frasi che l'utente gli sottopone in linguaggio naturale (linguaggio umano).

Il core di un chatbot sono i complessi algoritmi di Natural Language Understanding (NLU) che permettono al chatbot di interpretare il linguaggio umano e carpirne le informazioni essenziali. Le informazioni che un motore NLU può estrarre da una frase sono principalmente due: l'intento e le entità. L'intento`è il significato generale della frase, quello che l'utente vuole dire al chatbot mentre le entità non considerano il significato generale, bensì cercano informazioni specifiche da estrarre (ad esempio, una data o il nome di una città). La potenza incredibile di queste tecnologie è che permettono, con un certo grado di errore, di interpretare correttamente una frase scritta in qualunque modo anche con errori di ortografia.

Ecco un esempio per chiarire meglio questo concetto: l'utente vuole sapere se può interrompere la fruizione del corso e riprenderlo successivamente oppure se deve guardare tutte le lezioni in un'unica sessione; in questo caso il chatbot dovrà essere in grado di classificare correttamente tutte le seguenti affermazioni:

? "Buongiorno, posso riprendere il corso in un secondo momento?"
? "Se interrompo un video, quando mi riconnetto alla piattaforma dovrò iniziare nuovamente dal principio?"
? "Ho poco tempo in uffcio posso guardarli a casa i vedeo?"

Per le tre frasi precedenti il motore NLU deve riconoscere lo stesso intento, ovvero "Riprendere corso successivamente" (da notare che le frasi sono scritte in modo completamente diverso e la terza presenta errori di ortografia evidenti). Quindi, come possiamo capire dall'esempio, il punto di forza dei chatbot è la capacità potenziale di interpretare correttamente infinite combinazioni di frasi diverse aventi lo stesso scopo e di essere robusti agli errori di ortografia e sintattici.

Addestrare un Chatbot

Tantissime tecnologie studiate nella computer science moderna prevedono di utilizzare algoritmi che sono in grado di auto-apprendere attraverso l'esperienza. Per capire in modo intuitivo questo complesso concetto, possiamo pensare agli algoritmi con auto-apprendimento come a dei bambini molto piccoli che apprendono tramite l'esperienza diretta e la gratificazione che ricevono dai genitori o educatori. Ad esempio, un bimbo piccolo impara a dire mamma o altre parole perchè i genitori lo spronano a dire quella stessa parola ripetendola decine di volte e, a seguito dei numerosi tentativi, quando il bimbo riesce a dire la parola riceve dai genitori delle gratificazioni d'affetto che permettono al bambino di comprendere di aver fatto un'azione corretta. Questo metodo di apprendimento funziona bene, tanto che, durante i primi anni di vita, pur non conoscendo la grammatica, i bambini riescono ad imparano a parlare correttamente grazie agli esempi fornitogli dal mondo circostante e dalle figure di riferimento che lo aiutano in questo processo.

Le tecnologie utilizzate dai motori NLU sono molto simili al bambino del nostro esempio, perchè per imparare ad interpretare correttamente l'intento sottostante ad una frase hanno bisogno di ricevere diverse decine o centinaia di esempi di frasi (adeguatamente classificate) aventi lo stesso intento, ma diversa formulazione. In gergo tecnico questa fase viene chiamata "training". Lo sviluppatore, alla stregua del genitore dell'esempio qui sopra, deve supervisionare questo processo e capire se gli algoritmi apprendono correttamente. Se così non fosse, deve provvedere a fornire ulteriori esempi che aiutino l'algoritmo ad apprendere ulteriormente. Questo processo è ripetuto finché non si raggiunge un livello di performance adeguato (quest'ultima fase è chiamata testing).

Lo sviluppo di Dybot

Il chatbot sviluppato da Mega Italia Media (MIM) si chiama DyBot e nella sua prima versione permette di interpretare e rispondere autonomamente ad una classe specifica di richieste provenienti dai discenti. In futuro, permetterà di rispondere a un numero sempre crescente di richieste da parte degli utenti. Infatti, le richieste provenienti dai discenti possono evolversi nel tempo (ad esempio, l'introduzione di una nuova funzionalità in DynDevice LMS potrebbe portare alla nascita di una nuova classe di richieste da parte dei discenti). Dybot è una tecnologia progettata per rispondere a questi bisogni, a breve e lungo termine. DyBot non sostituisce i tutor umani di Mega Italia Media, ma si affianca ad essi, rispondendo alle domande frequenti che prevedono risposte standardizzate. Tuttavia, quando non è in grado di fornire la miglior risposta, inoltra automaticamente il quesito al tutor umano per far sì che l'utente possa essere assistito nel migliore dei modi.

Lo sviluppo di Dybot è stato sviluppato seguendo un solido e rigoroso processo scientifico e sperimentale (Analisi, Classificazione, Implementazione, Training e Testing). In particolare:

- in prima istanza si è proceduto ad analizzare lo storico delle richieste pervenute ai tutor tramite la chat implementata nella piattaforma eLearning DynDevice LMS. La fase di analisi ha permesso di stabilire quali classi di domande erano le più frequenti e quindi concentrare l'attenzione del chatbot alla risoluzione di quelle richieste.
- nella fase di classificazione le richieste pervenute sono state "etichettate", cioè ad ogni richiesta è stato associato uno specifico intento.
- di conseguenza, è stato sviluppato il codice utile ad integrare il chatbot nel sistema di messaggistica di DynDevice LMS.
- infine, il motore NLU è stato addestrato a riconoscere gli intenti più comuni sottostanti le richieste che gli utenti di DynDevice LMS sottopongono normalmente ai tutor.

Il contenuto del paper internazionale su Dybot

Il contributo parte dall'assunto che, nell'ultimo decennio, le principali aziende private operanti nel settore IT hanno sviluppato piattaforme cloud-based con funzioni di **comprensione del linguaggio naturale (NLU)**, ampiamente utilizzate per scopi di ricerca e commerciali. Una delle ragioni principali del successo delle piattaforme NLU è che permettono di semplificare lo sviluppo di **Chatbots** o Spoken Dialogue Systems (SDS), in molti casi, senza la conoscenza dei linguaggi di programmazione.

L'articolo presenta una descrizione generale e una tassonomia che riunisce le caratteristiche e i vincoli dei diversi servizi NLU basati su cloud disponibili sul mercato. Inoltre, fornisce una valutazione e un confronto sulla capacità di riconoscere gli intenti sottostanti delle diverse frasi. Le frasi utilizzate sono una raccolta di richieste fatte dagli utenti in italiano sulla piattaforma eLearning DynDevice LMS.

L'analisi risulta molto utile ad aiutare l'azienda, proprietaria della piattaforma eLearning, a capire in quale piattaforma NLU è meglio costruire una **chatbot per la lingua italiana** con l'obiettivo di rispondere alle domande più frequenti. Precedentemente, studi solidi sul tema sono stati svolti solamente su piattaforme in lingua inglese

L'articolo è scaricabile qui.