SynerGE HACK'ELTH 2020

GEmma for GE

by Team Rethink
Under the mentorship of Mr. Spandan Das

Problem Statement: GESH2020_2

Conversational AI Platform for Medical Equipment Management

To create a conversational AI platform that could help doctors, hospital staff & other medical practitioners get clarifications & answers to their queries around product functionality, operating guidelines, contract terms, service status, etc., for the product / equipment that they have procured from a Medical Equipment manufacturer.

The platform should be 24x7 available and could replace the need to connect with a call center executive, during specific office hrs, to get even basic questions answered.

NEED FOR THE PROJECT: Let’s Watch Why

Moment of Truth…

Demo
Walk-through

Let us show you what we've been upto these past 12 days

We implemented our project as a website for best accessibility. It can also register users so that if at all in the future you wish to integrate this with other products or analyse data retrieved from the website, the database would make it simpler.

Let's go talk to GEmma now!

http//127.0.0.0:5000

The Models behind it:

- BERT Large Uncased (The QnA Model)
- BART Large CNN (The Summarizer)
- DialoGPT-Large (Conversational API)
- BART Large MNLI (Zero Shot Classifier)
- NeuralCoref4 (Conference Resolution)
- ElasticSearchDocumentStore (for faster context retrieval)
- Dense Passage Retriever (Retrieves the relevant context)
- PyPDF2 (for new DocumentStore additions)
- Seq2Seq Bidirectional RNN with Attention Mechanism (Our self built punctuation model)

TECH - STACK

What is GEmma made of?
The How of GEmma's implementation...

We got to learn a lot while we put GEmma together and are really excited to tell you how it all works.

Because,

Technology makes what was once impossible possible

The design makes it real…….

The website and UI:

- Flask (A Python micro-web framework)
- jQuery and Ajax (for Client side retrieval of data)
- Bootstrap (An HTML, CSS and Javascript library)
- MySQL Workbench
- Flask MySQL DB
- HTML5
- CSS3

How do these models gel together? Let’s talk about them

When we open the chat window.
We are technically at a point in the code where the model is waiting for our first communication.

When it receives that text it goes to a Zero Shot Classifier
With labels question, converse or conversational question

ZERO SHOT CLASSIFIER

When it’s a “question”.

Now that it has identifies the message as a question:

1. Decides if its a summarize question or not
2. Uses Summerizer in the first case to return output
3. Uses the BERT QnA Model otherwise
4. In both cases context is retrieved by the Dense Passage Retriver from the ElasticSearchDocumentStore
5. In case it is not able to answer, it will provide the option to look at a summarized version of the retrieved document

BERT Large Uncased (The QnA Model)

BART Large CNN(The Summarizer)

Dense Passage Retriever and ElasticSearchDocumentStore (Getting the context right)

NeuralCoref (Maintaining Flow)

When it’s either “converse” or “conversational question”

DialoGPT-Large (Conversational API)

A professional ChatBot must have punctuation.
For that, We had to improvise

Seq2Seq Bidirectional RNN with Attention Mechanism

(Our self built punctuation model)

Additional functionality that is readily addable:

Speech Recognition
(We already did this with PocketSphynx)

User's Product List
(Can be added at time of registration, whenever a user buys a product they get a GEmma ID or can be added to their existing user)

Automatic Mailing
(Whenever any new products similar to their search or existing products show up we can automatically generate a mail)

An Additional Feature…

To make it truly useful

We have a CLI tool that let's you add more documents to your existing document store

PyPDF2 (for new DocumentStore additions)

Thank You

Gayathri Arvind
Uma Nalini Damodaran
V Abhijit Narayan Rao