Abstract

Question answering (QA) systems [1] are developed to enable the communication between a human and a computer. Since the dawn of computers, the communication between man and machine has evolved with leaps and bounds. This evolution has opened up the windows to an entirely new field known as Natural Language Processing(NLP). NLP makes it possible for the computer to understand a human language in its natural form. Today NLP gets used in some way or the other in almost all walks of human life. A simple QA system is a good example of the same.

In today's fast paced and technologically dominated world, the internet, especially google search has become the 'go-to' solution for almost all human problems. However, the vast amount of information on the internet does more harm than good in many cases. An example of the same is people trying to check their disease symptoms online prior to consulting a doctor. This sometimes leads to people panicking over a simple headache or worst case ignoring life threatening conditions.

The aim of this project is to develop a QA system that makes use of the information provided by a person's primary physician. Development of such a system will help the patient to get basic answers to their health questions tailored to their specific case.

The project aims to develop a QA system that would be beneficial in the field of medicine. In age of information explosion, such a system would play the role of a moderation force and help the medical professional prevent spread of incorrect information. Such a system has the potential to help reduce the ever increasing number of people that try self-treatment based on the unsupervised information from the internet.
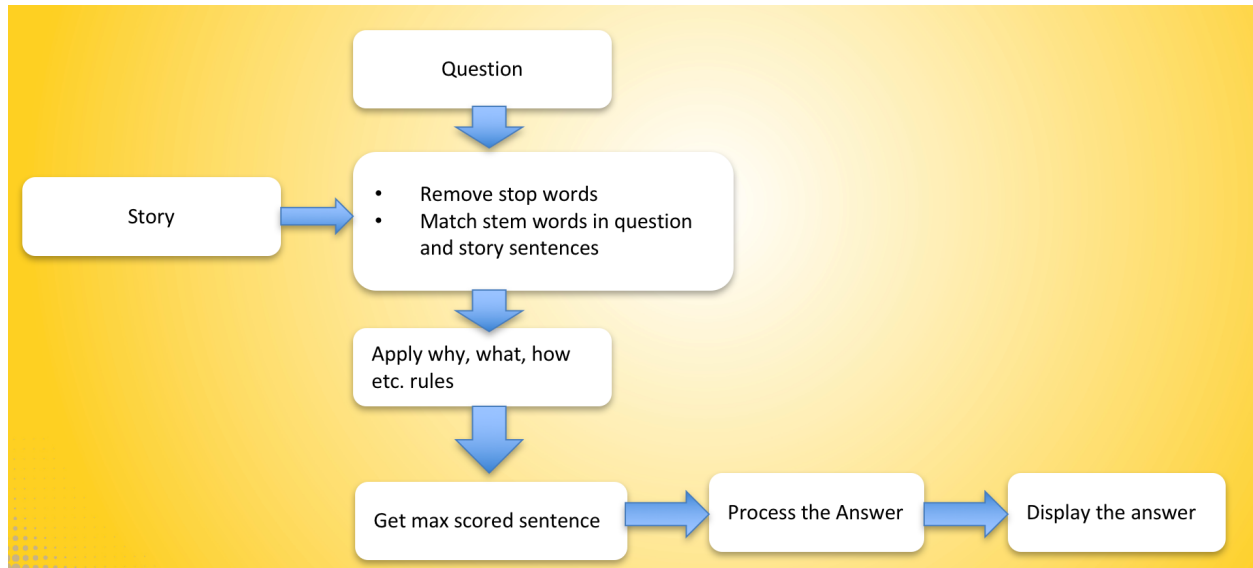
Introduction

My motivation to develop such system comes from my own family. It's just that when you care about someone in the family, you always want to help them by looking for solutions to their problem. As internet is nowadays at a fingertip distance we always tend to lookout for solutions on the internet. Though it is the first consultant in many of the problematic situations, it is also spread across vast amount of information sources, some write and some misleading.

Internet sometimes provides helpful information and sometimes misleading. But in case of health no one wants to end up in a situation where the problem is critical but you ignore it trusting some irrelevant sources or where it is not critical and you create a panicking situation.

This is where Question Answering system comes into picture. Your personal Physician can keep record of your health history, your family history and guide you accordingly. The document provided by the doctor can include caution in case of search for some critical question, Like see you doctor immediately etc. In this case, things won't be ignored and your day today health related questions will be answered under you doctors supervision.

Block diagram

Output

```
Q:Does hypertension increase risk of other diseases?
Ans: High blood pressure or hypertension increases the risk of heart disease and stroke


Q:Is obesity a risk factor for hypertension?
Ans:  Hypertension risk factors include obesity, drinking too much alcohol, smoking, and family history


Q:What is the common treatment for hypertension?
Ans:  Beta-blockers are a common treatment for hypertension.  Many doctors say prehypertension should be treated


Q:What are the symptoms to look out for hypertension?
Ans: If your blood pressure is extremely high, there may be certain symptoms to look out for, including Severe headache, Fatigue or confusion
If you have any of these symptoms, see a doctor immediately
```

Evaluation

I have measured accuracy of the system using Precision and recall

We tested this system on two different data, one was based on Hypertension [3] and other was on Type 2 Diabetes [4].

We found that out of the 10 questions we asked we got results for all 10 out of which two were incorrect and it was because of the incorrect tags.

$$Precision = 10/12 = 0.833 \qquad\qquad Recall = 8/10 = 0.8$$

Discussion

1. It is required that all the sentences should be grammatically correct so that we can match heuristic scores.

2. We also found that the tags assigned by pos_tags are not correct in cases where "test" is a VB but is treated as NN by pos_tag similar case with prevent as well. This leads to incorrect score and ultimately to the incorrect or incomplete information.

3. Deep learning can be further implemented to make system smarter.

4. Going further this project can be extended as an interactive Chatbot or an android application.

5. This could also be developed as a Text based system where a patient can send a mail or text to the System with bunch of queries and then system replies back by addressing each question individually. This Question Answering system can be maintained by Physician's office so it can intelligently address the question.

References

[1] Daniel Jurafsky and James H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall, 2017.

[2] Riloff, E. and Thelen, M., (2000) "A Rule-based Question Answering System for Reading Comprehension Tests", ANLP/NAACL-2000 Workshop on Reading Comprehension Tests as Evaluation for Computer-Based Language Understanding Systems.

[3] webmd.com "Hypertension overview facts" https://www.webmd.com/hypertension-high-blood-pressure/guide/hypertension-overview-facts"

[4] webmd.com "Type 2 Diabetes" https://www.webmd.com/diabetes/type-2-diabetes#1"