CS410 Final_Project - Professional Health Search

1. Overview

In this project we have tried to implement a search system that aims to find medical health professionals specialized in the given research areas. The underlying data we are collecting from web crawler from different hospitals like https://nyulangone.org/ . We have also ranked the data. We utilized the common code from Expert Search system (https://github.com/CS410Fall2020/ExpertSearch).

2. Implementation Details

The application is web based. We have deployed both the front end and the back end in AWS cloud. But the application can be set up even on the local machine, the details have been provided later in this document. The front end of the project has been developed using JavaScript, CSS3 and bootstrap. The backend has been developed using Gunicorn, Flask and Python.

The code comprises of the following parts:
Front End – It's a simple UI where the user can enter the search criteria related to any healthcare provider specialty for e.g. Cardiology, neurology, etc. and click the Search button.

Web Crawler – It crawls the pages in https://nyulangone.org/ to find the pages related to the health care providers.

Web scraper – It cleans, parses the data obtained above and stores in the form of text files.

Ranker – It performs the ranking of the documents using the BM25 ranking function. It then returns the list of the doctors that satisfy the search criteria.

Improvements:

- The UI could have been made a bit more elaborate.
- We could have displayed a bit more information about each healthcare provider that satisfy the search criteria.
- The system can be extended for searching not just the healthcare providers but also other medical services like scan, x-ray etc that are typically done in each hospital or medical facility.

3. Usage

The project is quite simple to use.

We have deployed both the front end and back end in the cloud. The project can be accessed using the following link:

http://34.238.135.24:5000/

The project can also be installed using the following steps:

Make sure you have python 2.7.x installed on your machine.

Step 1 - Clone code

git clone https://github.com/toskaushik/ExpertSearch.git

Step 2 - Navigate to ExpertSearch

cd ExpertSearch

Step 3 - Install dependencies

pip install -r requirements.txt

Step 4 - Run the application using Gunicorn

gunicorn server:app -b 127.0.0.1:5000

Note - We tested Application in Mac and Linux OS.

4. Team

- Sharad Kaushik – Team Leader. Responsibilities include implementation of the Web crawler, Web scraper that provides the data to the system.

- Faizan Ali Danish Khan – Team member. Responsibilities include implementation of the UI, the backend functionalities for the ranker and deployment in AWS.

- Ambalika Roy – Team member. Responsible for Installation of the system, testing the application, project presentation and project documentation.