Project Report

Project Guide:

Prof. Radhika Mamidi
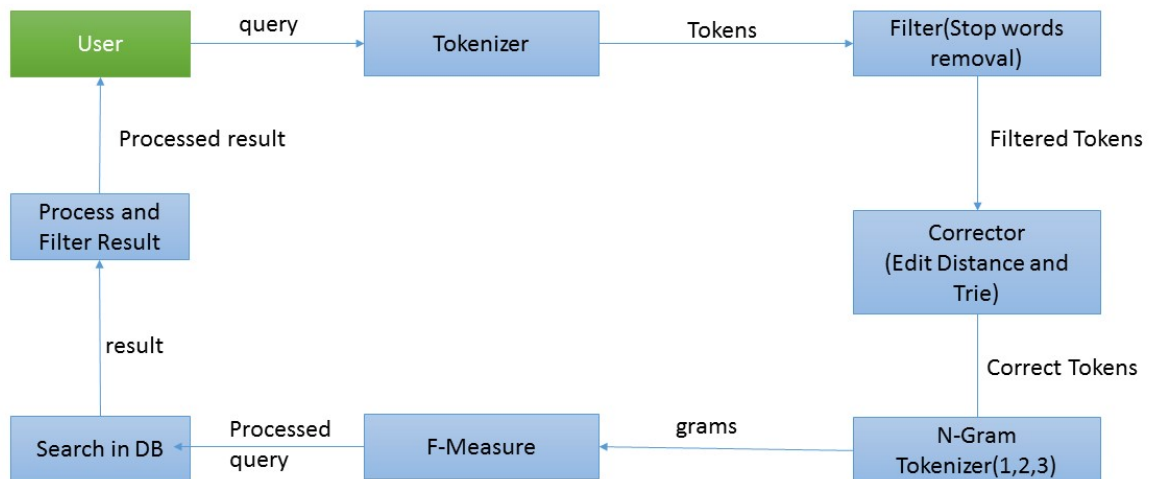
Team Members
Madhudeep Petwal
Vinod Adwani
Satyam Verma

# DoctorBot

**Abstract**

An offline web app which can be used in daily life for querying the common diseases and their associated symptoms. It provides an advice to the patients about the possibility of diseases that could happen.

**Control Flow:**

**Steps Taken:**

1.  Taking input from the user
    User enters the query and it is processed for eg.
    User1 : What are the symptms of cancer?

2.  The query is passed to the tokenizer which generates tokens
    E.g "what", "are", "the", "symptms", "of", "cancer"

3.  Stop words removal
    It removes all the stop words
    E.g "what", "symptms", "cancer"

4.  Correction
    It corrects all the wrong words based on edit distance using trie data structure
    E.g "what", "symptoms", "cancer"

5.  N-gram tokenizer
    It Generates Uni-gram, Bi-gram and Tri-gram tokens
    E.g Unigram - ("what","symptoms", "cancer")
        Bi-gram - ( ("what", "symptoms"),("symptoms","cancer") )
        Tri-gram - ("what", "symptoms", "cancer")

6.  Calculating F-measure

7.  Result is queried in Medical database and output is processed and shown to the user.

1.  **Data Extraction**
    We have extracted data from site http://umm.edu/health/medical/

```
satyam@satyam:~/DoctorBot$ python scrapper.py
satyam@satyam:~/DoctorBot$ python scrapper.py
extracting  A z
extracting  B z
extracting  C z
extracting  D z
extracting  E z
extracting  F z
extracting  G z
extracting  H z
extracting  I z
extracting  J z
extracting  K z
extracting  L z
extracting  M z
extracting  N z
extracting  O z
extracting  P z
extracting  Q z
extracting  R z
extracting  S z
extracting  T z
extracting  U z
extracting  V z
extracting  W z
extracting  X z
extracting  Y z
extracting  Z z
satyam@satyam:~/DoctorBot$ █
```

We have extracted data in python and the data is saved in local disk.
Each disease is a key in dictionary and there is another dictionary where the key is symptoms and various diseases are mappings.

2. **Edit Distance**
   We have applied edit distance algorithm for queries which are not correct in language like misspelled letters etc.

   Thus, query like
   1. symptoms of cancer
   2. Symptoms of canc
   3. symtmp of cancer
   Gives same results.

Below are the screenshots of above queries

```
satyam@satyam:~/DoctorBot$ python queryEngine.py
Enter your query symptoms of cancer
(u'Symptoms of cancer depend on the type and location of the cancer. For example, lung cancer can cause coughing, shortness of breath, or chest pain. C
olon cancer often causes diarrhea, constipation, and blood in the stool.\nSome cancers may not have any symptoms. In certain cancers, such as pancreati
c cancer, symptoms often do not start until the disease has reached an advanced stage.\nThe following symptoms may occur with cancer:\nChills\nFatigue\
nFever\nLoss of appetite\nMalaise\nNight sweats\nWeight loss\n', [u'chest pain', u'constipation', u'chills', u'fatigue', u'fever', u'malaise', u'chills
', u'fatigue', u'fever', u'malaise'])
satyam@satyam:~/DoctorBot$
```

```
satyam@satyam:~/DoctorBot$ python queryEngine.py
Enter your query symptoms of canc
(u'Symptoms of cancer depend on the type and location of the cancer. For example, lung cancer can cause coughing, shortness of breath, or chest pain. C
olon cancer often causes diarrhea, constipation, and blood in the stool.\nSome cancers may not have any symptoms. In certain cancers, such as pancreati
c cancer, symptoms often do not start until the disease has reached an advanced stage.\nThe following symptoms may occur with cancer:\nChills\nFatigue\
nFever\nLoss of appetite\nMalaise\nNight sweats\nWeight loss\n', [u'chest pain', u'constipation', u'chills', u'fatigue', u'fever', u'malaise', u'chills
', u'fatigue', u'fever', u'malaise'])
satyam@satyam:~/DoctorBot$
```

```
satyam@satyam:~/DoctorBot$ python queryEngine.py
Enter your query symtmp of cancer
(u'Symptoms of cancer depend on the type and location of the cancer. For example, lung cancer can cause coughing, shortness of breath, or chest pain. C
olon cancer often causes diarrhea, constipation, and blood in the stool.\nSome cancers may not have any symptoms. In certain cancers, such as pancreati
c cancer, symptoms often do not start until the disease has reached an advanced stage.\nThe following symptoms may occur with cancer:\nChills\nFatigue\
nFever\nLoss of appetite\nMalaise\nNight sweats\nWeight loss\n', [u'chest pain', u'constipation', u'chills', u'fatigue', u'fever', u'malaise', u'chills
', u'fatigue', u'fever', u'malaise'])
satyam@satyam:~/DoctorBot$
```