

This is a 'smart' chat bot program

pip install nltk

```
Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (3.8.1)
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk) (1.4.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk) (2023.12.25)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk) (4.66.4)
```

pip install newspaper3k

```
Requirement already satisfied: beautifulsoup4>=4.4.1 in /usr/local/lib/python3.10/dist-packages (from newspaper3k) (4.12.3)
Requirement already satisfied: Pillow>=3.3.0 in /usr/local/lib/python3.10/dist-packages (from newspaper3k) (9.4.0)
Requirement already satisfied: PyYAML>=3.11 in /usr/local/lib/python3.10/dist-packages (from newspaper3k) (6.0.1)
Collecting cssselect>=0.9.2 (from newspaper3k)
  Downloading cssselect-1.2.0-py2.py3-none-any.whl (18 kB)
Requirement already satisfied: lxml>=3.6.0 in /usr/local/lib/python3.10/dist-packages (from newspaper3k) (4.9.4)
Requirement already satisfied: nltk>=3.2.1 in /usr/local/lib/python3.10/dist-packages (from newspaper3k) (3.8.1)
Requirement already satisfied: requests>=2.10.0 in /usr/local/lib/python3.10/dist-packages (from newspaper3k) (2.31.0)
Collecting feedparser>=5.2.1 (from newspaper3k)
  Downloading feedparser-6.0.11-py3-none-any.whl (81 kB)
81.3/81.3 kB 5.5 MB/s eta 0:00:00
Collecting tldextract>=2.0.1 (from newspaper3k)
  Downloading tldextract-5.1.2-py3-none-any.whl (97 kB)
97.6/97.6 kB 9.2 MB/s eta 0:00:00
Collecting feedfinder2>=0.0.4 (from newspaper3k)
  Downloading feedfinder2-0.0.4.tar.gz (3.3 kB)
  Preparing metadata (setup.py) ... done
Collecting jieba3k>=0.35.1 (from newspaper3k)
  Downloading jieba3k-0.35.1.zip (7.4 MB)
7.4/7.4 MB 28.3 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Requirement already satisfied: python-dateutil>=2.5.3 in /usr/local/lib/python3.10/dist-packages (from newspaper3k) (2.8.2)
Collecting tinysegmenter==0.3 (from newspaper3k)
  Downloading tinysegmenter-0.3.tar.gz (16 kB)
  Preparing metadata (setup.py) ... done
Requirement already satisfied: soupsieve>1.2 in /usr/local/lib/python3.10/dist-packages (from beautifulsoup4>=4.4.1->newspaper3k) (2.5)
Requirement already satisfied: six in /usr/local/lib/python3.10/dist-packages (from feedfinder2>=0.0.4->newspaper3k) (1.16.0)
Collecting sgmlib3k (from feedparser>=5.2.1->newspaper3k)
  Downloading sgmlib3k-1.0.0.tar.gz (5.8 kB)
  Preparing metadata (setup.py) ... done
Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk>=3.2.1->newspaper3k) (8.1.7)
Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk>=3.2.1->newspaper3k) (1.4.2)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk>=3.2.1->newspaper3k) (2023.12.25)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk>=3.2.1->newspaper3k) (4.66.4)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests>=2.10.0->newspaper3k) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests>=2.10.0->newspaper3k) (3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests>=2.10.0->newspaper3k) (2.0.7)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests>=2.10.0->newspaper3k) (2024.2.2)
Collecting requests-file>=1.4 (from tldextract>=2.0.1->newspaper3k)
  Downloading requests_file-2.0.0-py2.py3-none-any.whl (4.2 kB)
Requirement already satisfied: filelock>=3.0.8 in /usr/local/lib/python3.10/dist-packages (from tldextract>=2.0.1->newspaper3k) (3.14.0)
Building wheels for collected packages: tinysegmenter, feedfinder2, jieba3k, sgmlib3k
  Building wheel for tinysegmenter (setup.py) ... done
  Created wheel for tinysegmenter: filename=tinysegmenter-0.3-py3-none-any.whl size=13538 sha256=49bfc980319f34800ec87d1163610a2f9a7e8ae6e114f2l
  Stored in directory: /root/.cache/pip/wheels/c8/d6/6c/384f58df48c00b9a31d638005143b5b3ac62c3d25fb1447f23
  Building wheel for feedfinder2 (setup.py) ... done
  Created wheel for feedfinder2: filename=feedfinder2-0.0.4-py3-none-any.whl size=3340 sha256=e18972575d8a88b487c34e9a019027c31cd6e240b6c71c517f
  Stored in directory: /root/.cache/pip/wheels/97/02/e7/a1ff1760e12bdbaab0ac824fae5c1bc933e41c4ccd6a8f8edb
  Building wheel for jieba3k (setup.py) ... done
  Created wheel for jieba3k: filename=jieba3k-0.35.1-py3-none-any.whl size=7398382 sha256=ab48d8920aa3a84083b9cb77491c0f9ac83aecd740a73197b333f
  Stored in directory: /root/.cache/pip/wheels/7a/c4/0c/12a9a314ecac499456c4c3b2fcc2f635a3b45a39dfbd240299
  Building wheel for sgmlib3k (setup.py) ... done
  Created wheel for sgmlib3k: filename=sgmlib3k-1.0.0-py3-none-any.whl size=6049 sha256=1a2d7b4294e04bfa812b817a4ef062fe5e7fd43a5b572ddc01432c
  Stored in directory: /root/.cache/pip/wheels/f0/69/93/a47e9d621be168e9e33c7ce60524393c0b92ae83cf6c6e89c5
Successfully built tinysegmenter feedfinder2 jieba3k sgmlib3k
Installing collected packages: tinysegmenter, sgmlib3k, jieba3k, feedparser, cssselect, requests-file, feedfinder2, tldextract, newspaper3k
Successfully installed cssselect-1.2.0 feedfinder2-0.0.4 feedparser-6.0.11 jieba3k-0.35.1 newspaper3k-0.2.8 requests-file-2.0.0 sgmlib3k-1.0.0
```

```
from newspaper import Article
import random
import string
import nltk
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics.pairwise import cosine_similarity
import numpy as np
import warnings
warnings.filterwarnings('ignore')
```

nltk.download('punkt',quiet=True)

True

GET THE ARTICLE Url

```
article=Article('https://www.mayoclinic.org/diseases-conditions/chronic-kidney-disease/symptoms-causes/syc-20354521') #url about black holes
article.download() #Downloads the HTML content of the article from the web
article.parse() #Parses the downloaded HTML content to extract the article's text
article.nlp() #Applies natural language processing to extract keywords, summary, and other insights from the article's text
corpus=article.text #property contains the main body text of the article after it has been parsed
```

```
print(corpus)
```



```
I'm Dr. Andrew Bentall, a kidney doctor at Mayo Clinic. I look after patients with kidney disease, either in the early stages, or with more advanced chronic kidney disease. Chronic kidney disease is a disease characterized by progressive damage and loss of function in the kidneys. It's estimated that chronic kidney disease affects about 15 percent of the U.S. population. There are quite a few things that can cause or put you at higher risk for chronic kidney disease. Some of them are not things that can be avoided. Often there are no outward signs in the earlier stages of chronic kidney disease, which is grouped into stages 1 through 5. Generally, earlier stages of the disease are easier to manage. Even before any symptoms appear, routine blood work can indicate that you might be in the early stages of chronic kidney disease. And the earlier you find out, the better off you'll be. Treatment is determined by what is causing your kidneys to not function normally. Treating the cause is key, leading to reduced complications and better outcomes. If you have kidney disease, there are lifestyle choices. Namely quit smoking. Consuming alcohol in moderation. If you're overweight or obese, then losing weight can help.
```



```
#tokenization
text=corpus
sentence_list=nltk.sent_tokenize(text)
```

```
print(sentence_list)
```



```
["I'm Dr. Andrew Bentall, a kidney doctor at Mayo Clinic.", 'I look after patients with kidney disease, either in the early stages, or with more advanced chronic kidney disease. Chronic kidney disease is a disease characterized by progressive damage and loss of function in the kidneys. It's estimated that chronic kidney disease affects about 15 percent of the U.S. population. There are quite a few things that can cause or put you at higher risk for chronic kidney disease. Some of them are not things that can be avoided. Often there are no outward signs in the earlier stages of chronic kidney disease, which is grouped into stages 1 through 5. Generally, earlier stages of the disease are easier to manage. Even before any symptoms appear, routine blood work can indicate that you might be in the early stages of chronic kidney disease. And the earlier you find out, the better off you'll be. Treatment is determined by what is causing your kidneys to not function normally. Treating the cause is key, leading to reduced complications and better outcomes. If you have kidney disease, there are lifestyle choices. Namely quit smoking. Consuming alcohol in moderation. If you're overweight or obese, then losing weight can help.']
```



```
#creating function to return a greeting response
def greeting_response(text):
    text=text.lower()

    bot_greetings=['howdy','hi','hey','hello','hola']
    user_greetings=['hi','hey','hello','hola','greetings','wassup']

    for word in text.split():
        if word in user_greetings:
            return random.choice(bot_greetings)
```

```
def index_sort(list_var):
    length=len(list_var)
    list_index=list(range(0,length))
    x=list_var
    for i in range(length):
        for j in range(length):
            if x[list_index[i]] > x[list_index[j]]:
                temp=list_index[i]
                list_index[i]=list_index[j]
                list_index[j]=temp
    return list_index
```

```
#bot response function is designed to generate a response from a chatbot by calculating the similarity between the user input
def bot_response(user_input):
```

```
# similarity_scores_list

    bot_response=''

# index

    similarity_scores_list=similarity_scores_flatten()

print('Dr Bot: I am Scie Bot. I will answer your queries about Black holes. If you want to exit, type bye.')
```

```
exit_list=['exit','see you later','bye','quit','break']
while(True):
    user_input=input()
    if user_input.lower() in exit_list:
        print('Dr Bot: Chat with you later...')
        break
```

```
➡ Dr Bot: I am Scie Bot. I will answer your queries about Black holes. If you want to exit, type bye.
bye
Dr Bot: Chat with you later...

if response flag==0:

while(True):
    user_input=input()
    if user_input.lower() in exit_list:
        print('Dr Bot: Chat with you later...')
        break
    else:
        if greeting_response(user_input) != None:
            print('Dr Bot: '+greeting_response(user_input))
```

```
➡ hi
Dr Bot: howdy
hello
Dr Bot: howdy
hey
Dr Bot: hello
bye
Dr Bot: Chat with you later...
```

```
while(True):
    user_input=input()
    if user_input.lower() in exit_list:
        print('Dr Bot: Chat with you later...')
        break
    else:
        if greeting_response(user_input) != None:
            print('Dr Bot: '+greeting_response(user_input))
        else:
            print('Dr Bot: '+bot_response(user_input))
```

```
➡ hi
Dr Bot: hey
what is chronic kidney disease
Dr Bot: Chronic kidney disease is a disease characterized by progressive damage and loss of function in the kidneys.
what are the risk factor
Dr Bot: There are things that are more under our control: Heavy or long-term use of certain medications, even those that are common over-the-coun
bye
Dr Bot: Chat with you later...
```

