

Cara Menjalankan

Buat folder tmp, dan virtual environment, dan install requirements untuk mac dan gnu/linux:

```
1 mkdir tmp
2 python3 -m venv venv
3 source venv/bin/activate
4 pip3 install -r requirements.txt
```

Untuk windows:

```
1 mkdir tmp
2 python -m venv venv
3 venv\Scripts\activate
4 pip install -r requirements.txt
```

atau:

```
1 mkdir tmp
2 virtualenv venv
3 venv\Scripts\activate
4 pip install -r requirements.txt
```

Kalau yang windows venvnya salah commandnya, bisa di browsing kak hehe saya ga pake windows soalnya maaf kak 🙏

Selanjutnya jalankan search.py.

Perhatikan bahwa signature fungsi retrieve:

```
def retrieve_[method scoring](self, query, k=10, optimize=True, debug=False):
```

Flag optimize digunakan untuk menggunakan Top-K Wand Algorithm.

Hasil menggunakan `retrieve_tfidf`

```
1 Query : alkylated with radioactive iodoacetate
2 Results:
3 Document Name          Score
4 ./collection/6/507.txt  24.036
5 ./collection/6/554.txt  11.771
6 ./collection/11/1003.txt 9.404
7 ./collection/4/388.txt  8.386
8 ./collection/4/387.txt  8.386
9 ./collection/11/1018.txt 6.765
10 ./collection/2/144.txt  6.765
11 ./collection/2/119.txt  6.765
12 ./collection/3/247.txt  6.765
```

```

13 ./collection/6/512.txt          6.765
14
15 Query : psychodrama for disturbed children
16 Results:
17 Document Name                  Score
18 ./collection/9/820.txt         20.276
19 ./collection/10/918.txt        12.965
20 ./collection/9/821.txt         12.948
21 ./collection/1/36.txt          11.402
22 ./collection/9/817.txt         11.243
23 ./collection/10/926.txt        10.335
24 ./collection/8/799.txt         10.335
25 ./collection/10/927.txt        10.072
26 ./collection/1/99.txt          9.199
27 ./collection/2/100.txt         9.199
28
29 Query : lipid metabolism in toxemia and normal pregnancy
30 Results:
31 Document Name                  Score
32 ./collection/1/7.txt           28.687
33 ./collection/2/159.txt         18.274
34 ./collection/4/328.txt         17.877
35 ./collection/4/306.txt         17.659
36 ./collection/1/81.txt          16.615
37 ./collection/4/329.txt         12.836
38 ./collection/3/288.txt         12.630
39 ./collection/4/371.txt         11.406
40 ./collection/9/881.txt         10.723
41 ./collection/3/201.txt         10.566

```

Hasil menggunakan `retrieve_bm25`

```

1 [nltk_data] Downloading package stopwords to /home/hocky/nltk_data...
2 [nltk_data] Package stopwords is already up-to-date!
3 [nltk_data] Downloading package punkt to /home/hocky/nltk_data...
4 [nltk_data] Package punkt is already up-to-date!
5 Query : alkylated with radioactive iodoacetate
6 Results:
7 Document Name                  Score
8 ./collection/6/507.txt         21.465
9 ./collection/11/1003.txt       9.548
10 ./collection/6/554.txt         8.549
11 ./collection/6/506.txt         7.108
12 ./collection/8/793.txt         7.082
13 ./collection/2/119.txt         6.724
14 ./collection/4/387.txt         6.574
15 ./collection/3/247.txt         6.178

```

16 ./collection/8/745.txt 6.015
17 ./collection/11/1018.txt 5.882

18

19 Query : psychodrama for disturbed children

20 Results:

21	Document Name	Score
22	./collection/9/820.txt	19.885
23	./collection/9/821.txt	10.632
24	./collection/1/36.txt	9.860
25	./collection/10/918.txt	9.477
26	./collection/9/817.txt	8.179
27	./collection/10/926.txt	8.087
28	./collection/9/814.txt	7.642
29	./collection/1/99.txt	7.552
30	./collection/9/894.txt	7.358
31	./collection/5/493.txt	7.355

32

33 Query : lipid metabolism in toxemia and normal pregnancy

34 Results:

35	Document Name	Score
36	./collection/1/7.txt	26.316
37	./collection/4/328.txt	17.556
38	./collection/4/306.txt	14.148
39	./collection/2/159.txt	13.300
40	./collection/1/81.txt	10.982
41	./collection/7/602.txt	10.343
42	./collection/4/371.txt	9.948
43	./collection/4/329.txt	9.542
44	./collection/3/288.txt	8.923
45	./collection/1/12.txt	8.415