**Programming Assignment**

**1. Goal:** Make a basic information retrieval system with tf-idf weighting.

**2. Data Set**

- The cranfield data set is used (load the data set from the python package **ir-datasets**).

- You can get the answer set using the *get\_answer\_set(dataset, qid)* function.

**3. Detail**

- See the jupyter notebook file.

- Fill in “Your Code” parts.

**4. Due date:** June 9. (Tentative)

**5. Submission Files**

- your jupyter notebook file

- Captured Results

1) Zipf’s law graph

텍스트, 스크린샷, 디스플레이, 직사각형이(가) 표시된 사진

자동 생성된 설명

2) Zipf’s law graph with log scale

라인, 그래프, 스크린샷, 도표이(가) 표시된 사진

자동 생성된 설명

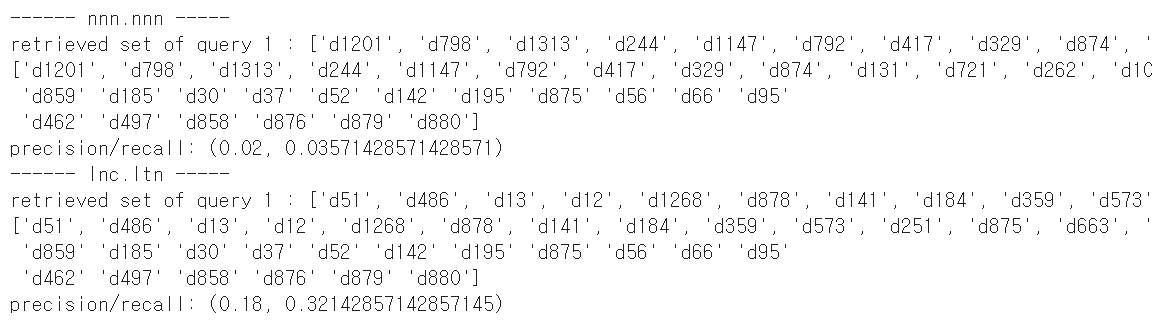
3) count matrix

텍스트, 스크린샷, 폰트이(가) 표시된 사진

자동 생성된 설명

4) Capture the results with mode=’nnn.nnn’ and mode=’lnc.ltn’ for query2.

(Sample Image for query 1)



query 2

텍스트, 스크린샷, 폰트이(가) 표시된 사진

자동 생성된 설명