텍스트, 스크린샷, 번호, 소프트웨어이(가) 표시된 사진

자동 생성된 설명

텍스트, 폰트, 라인, 스크린샷이(가) 표시된 사진

자동 생성된 설명

* preprocess\_new.py의 input 파일 경로

텍스트, 라인, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

* preprocess\_new.py의 output 파일이자, opt\_num\_topics.py, run\_topic\_modeling.py 의 input 파일 경로

#Topic modeling 돌리기

텍스트, 스크린샷, 소프트웨어, 번호이(가) 표시된 사진

자동 생성된 설명

1. opt\_num\_topics.py 에서 경로 및 변수 설정 지정
2. opt\_num\_topics.sh 돌리기 (sh opt\_num\_topics.sh)

🡪 결과 .opt\_num\_topics 에 저장됨

1. 2의 결과로 나온 opt\_num\_topics 폴더 내의 결과 보고, 토픽개수 선정 후

🡪 perplexity는 낮을 수록 coherence는 높을수록 좋음 (그 어디 교차점 선택하면 됨)

선정된 토픽개수에 해당하는 .bin 파일을 ‘models’ 폴더에 옮김

1. run\_topic\_modeling.py 돌림

🡪 .results 에 결과 저장됨

#Topic clustering 돌리기

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

1. topic\_clustering/topic\_clustering.py 코드 실행시키기: