

문장을 분석하여 악의적인 댓글 분류

B489078 최봉근
B387028 오민규
B489032 설세형
B389088 홍정우

Contents

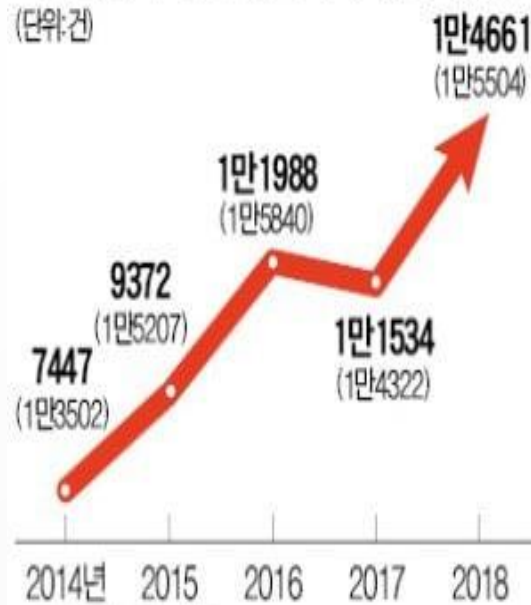
1. 주제 선정
2. 프로젝트 구조 및 사용한 기술 설명
3. 결론

주제 선정

Need

5년 만에 두배 늘어난 사이버 명예훼손

(단위:건)



※()안은 일반 명예훼손

자료:대검찰청

명예훼손 악플에 대한 법강화

[정원시작: 2019-10-15] [정원마감: 2019-11-14] [정원인원: 784]

10월14일 설리가 자택에서 자살을 했습니다. 많은 악플들과 모욕을 당했습니다. 죽고나서도 악플은 수없이 많이 올라왔습니다. 정신적으로도 많이 피해를 입었을 것이고 매우 힘들었을 것입니다.더이상 이런일이 발생안했으면 좋겠습니다.악플로 인해 ***의 **도 자살을 택하였습니다.이런 극단적인 자살을 막아주시고 악

HOME > 국민소통광장 > 국민청원 및 제언

가수설리 연예인 사망사건, sns악플 싸이버명예훼손,처벌강화해주세요

[정원시작: 2019-10-15] [정원마감: 2019-11-14] [정원인원: 821]

가수 설리가 악플과 루머를 견디다 못해 10월14일 향년 만24세 꽃다운 나이로 사망하였습니다. 비단, 과거 악플, 싸이버명예훼손 악성루머, 싸이버악플러들에 의해, 몇명의 연예인이 자살하였습니까, 너무 많습니다. 이전에 이슈가 되었던, 음담패설 네이버 다음 카페등, 불법성적인싸이트에 연예인이 아닌 일반인 여성

HOME > 국민소통광장 > 국민청원 및 제언

인터넷 실명제를 도입하여 주십시오

[정원시작: 2019-10-15] [정원마감: 2019-11-14] [정원인원: 369]

10월 14일 오후 배우이자 가수인 설리양이 세상을 떠났습니다. 이유는 자살이라고 하지만 과연 다들 자살이라고 생각할지 의문입니다. 악플, 꾸준히 달리는 악플들은 흔히 키보드 워리어라고 불리는 많은 사람들로부터 시작되어 연예인, 주변사람들을 괴롭히고 괴롭히고 괴롭히고 꾸준히 괴롭히고 있었습니다. Sns

-왼쪽 그래프는 대 검찰청에서 발표한 자료로 악플로 인한 사이버 명예훼손이 5년 동안 얼마나 증가하였는지를 알 수 있다.

-얼마 전 악플로 인해 가수 설리가 자살한 사건이 있었다 이 사건을 계기로 악플에 대한 관심이 더 많아졌고 예방하기 위한 움직임이 늘어나고 있다.

-따라서 악플을 막거나 예방하는 기술은 꼭 필요하고 중요한 일이라고 할 수 있다.

Competitor

nlty**** >

2019.11.18. 00:29 신고

자존심의 가치가 그 정도인가요

답글 193

👍 9868 🗨 994

iek1**** >

2019.11.18. 01:02 · 신고

내복증정품에 세계적으로 나라의 존심과 바꿔치기하는거지.. 유니클로 방문자들 인생 값어치가 내복한장 수준이네

답글 72

👍 7261 🗨 435

lusi**** >

2019.11.18. 00:26 · 신고

그리고싶을까..불쌍하다는생각이 드는건 나만그런건가..생각없는사람들..

답글 52

👍 3389 🗨 329

- 욕설 자동 필터링

너무 심한 욕설을 작성이 불가능하고 그렇지 않는 것은 ooo로 변형되거나 그냥 작성 가능하다

- 이용자들의 신고 기능

댓글 옆에 신고 버튼이 있어서 이용자들이 직접 신고 할 수 있다

프로젝트 구조 및 사용 기술

Architecture





학습 데이터 크롤링

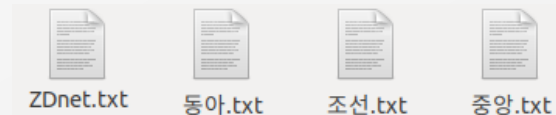
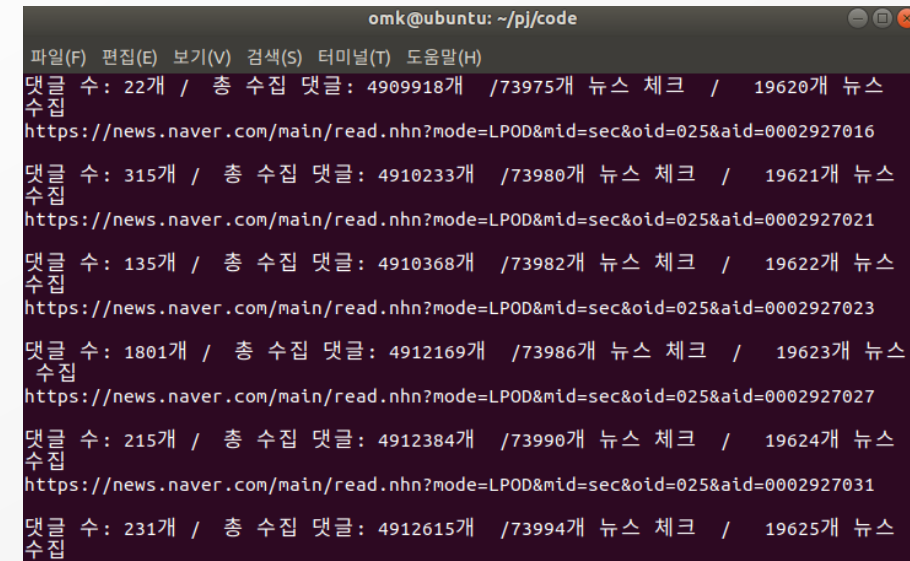


2017년 1월 ~ 2019년 5월 네이버 뉴스 댓글

카테고리: 정치, 경제, 사회, 생활/문화, IT, 세계

수집한 댓글의 개수: 1,812,483

크롤러 제작 및 학습 댓글 데이터 수집





데이터 라벨링(댓글 분류)

악성 댓글 분류 기준

공격성이 있고 비속어 포함 시	악플 댓글
공격성은 있으나 비속어 없을 시	의심 댓글
공격성과 비속어가 없을 시	일반 댓글

지금뉴스에 청계광장에 대규모경찰을 투입한다고 합니다 나가계신분들
여기는 신대방역 현재 이동 중
틀딱이 새끼들아
생긴것도 ㅇ같이 생겨가지고ㅇ같은 소리만 하고 자빠졌네사람이 그렇게
아직 그자리를 지키라는 메일을 받으셨는지 이번만은 상황을보고 생각C
작작좀해라 명 죽였다자나 ㅅㅂ
꼭참석해서 개 되지들도 화낼줄안다는것을 보여줘야겠다
변질되지말고 평화시위로 끌어내리자
최순실을 재물로 개돼지의 먹잇감으로 바치자 이년은 대중을 개돼지로
집회참가한 대학생들도 시민들도 다 빨갱이 중복이라고 떠들때 되지 않
살인마를 들이면 개지랄 헐놈덜이개소리허구 지랄이네 쳐자 덜



댓글 분류
라벨링

__label__일반 지금뉴스에 청계광장에 대규모경찰을 투입한다고 합니다 나가계신분들 안전이 최
__label__일반 여기는 신대방역 현재 이동 중
__label__악플 틀딱이 새끼들아
__label__악플 생긴것도 ㅇ같이 생겨가지고ㅇ같은 소리만 하고 자빠졌네사람이 그렇게도 없더니
__label__일반 아직 그자리를 지키라는 메일을 받으셨는지 이번만은 상황을보고 생각이라는것을
__label__악플 작작좀해라 명 죽였다자나 ㅅㅂ
__label__일반 꼭참석해서 개 되지들도 화낼줄안다는것을 보여줘야겠다
__label__일반 변질되지말고 평화시위로 끌어내리자
__label__의심 최순실을 재물로 개돼지의 먹잇감으로 바치자 이년은 대중을 개돼지로 우습게 보
__label__일반 집회참가한 대학생들도 시민들도 다 빨갱이 중복이라고 떠들때 되지 않았나
__label__악플 살인마를 들이면 개지랄 헐놈덜이개소리허구 지랄이네 쳐자 덜



데이터 전처리(초중종성 분리/ 의미 없는 특수문자 등 제거)

- 1 자살할놈이 뭐 방탄조끼 저거누가 좀 안죽이나
- 2 경위인데도 방탄 조끼하나 지급안되나 아무리 총기 소지가 불가능해!
- 3 공권력강화 절대 찬성
- 4 제발 처벌강화해서 조지자 제발부탁이다
- 5 사형하자 저런 쓰레기때문에 선량한 사람들이 고통받고 출동한 경찰레기는 사형하자
- 6 신발가게 직원분 대단
- 7 자살하려했다면서 방탄조끼 입었네
- 8 경찰 조끼는 뚫리고 범죄자 방탄은 막히고 엄청난 모순이다
- 9 죽을려고했다면서 방탄조끼는 왜입었대 ㅋㅋ



1 자-사-라-하-리-니-로-이- . -로-기- . 바-오-타-니-조-기- . 기- . 자- . 기- . 니- . 기- . 조-
2 기-기-오-기- . 이-니-데- . 니- . 바-오-타-니- 조-기- . 하- . 니- . 지- . 기-바-아-니-데-
기- . 니- . 오-하- . 니- . 마-니-아-기-오-니- 데- . 바- . 하- . 오- . 자- .
3 기-오-기-니-라-기-기-오-하- . 자-기-데- . 차-니-사-기-
4 제- . 바-리- 차- . 바-기-기-오-하- . 하- . 사- . 조- . 지- . 자- . 제- . 바-리-바- . 타-기-
5 사- . 하-기-오-하- . 자- . 자- . 리-니- . 쓰- . 리- . 기- . 따- . 로-니- . 하- . 사-기-리-기-오-하-니-
오-하-니- 기-기-오-차-리-기-라-니- . 로-기-사-기-로-니- . 이- . 기- . 오- . 이- . 리- . 니- . 이- . 리- . 니-
니- . 기- . 기- . 니- . 하- . 바- . 리- . 니- . 오- . 사- . 오- . 기- . 리- . 리- . 자- . 니- . 니- . 쓰- . 리- .
6 사- . 니-바-리-기- . 기- . 지- . 기- . 기- . 바- . 리- . 데- . 니- .
7 자- . 사-라-하- . 리- . 하- . 하- . 하- . 니- . 바-오-타-니-조-기- . 이-바-기- . 하- .
8 기-기-오-차-리- 조-기- . 니- . 따- . 기- . 바- . 리- . 자- . 바-오-타-니-로-니-
9 자-기-로-리- . 기- . 하- . 하- . 하- . 니- . 바-오-타-니-조-기- . 니- . 오- . 하- . 오- .
10 자- . 리- . 자- . 바-리- . 니- . 카-리- . 자-리- . 리- . 기- . 기- . 기- . 기- . 오- . 차-리- . 바-
차-리-기-라-니- . 자-기- . 바- . 리- . 니- . 사- . 제- . 바-오-타-니-바-기- . 바-
니- .
11 사-기-바-리- . 자- . 하- . 기- . 리- . 기- . 자- . 바-리- . 하- . 하- . 하- . 기- . 따- .
하- . 니- . 지- . 기- . 기- . 리- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- .
하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- . 하- .

Word embedding

워드 임베딩(Word Embedding)은 단어를 벡터로 표현하는 대표적인 방법

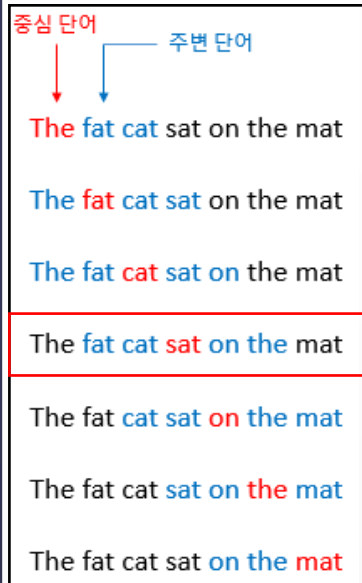
기존	분산 표현
표현하고자 하는 단어의 인덱스의 값만 1	단어의 의미를 다차원에 벡터화
↓	↓
하나에 단어를 표현하는데 큰 벡터가 필요 (사전 단어 x개면 x 차원이 필요하다)	벡터의 차원수가 상대적으로 줄어든다.
각 단어간 유사성을 표현할 수 없다는 단점	단어 간 유사도를 계산 가능하다.
  [1, 0, 0, 0, 0]	  [0.5, 0.0, 1.0, 1.0, 0.3]
  [0, 1, 0, 0, 0]	  [0.5, 0.0, 1.0, 1.0, 0.0]
  [0, 0, 1, 0, 0]	  [0.2, 0.9, 0.5, 0.0, 1.0]

비슷하다
다르다

-Word2Vec : 분산표현을 사용하는 대표적인 방식 중 하나
(NNML -인공신경망 사용)

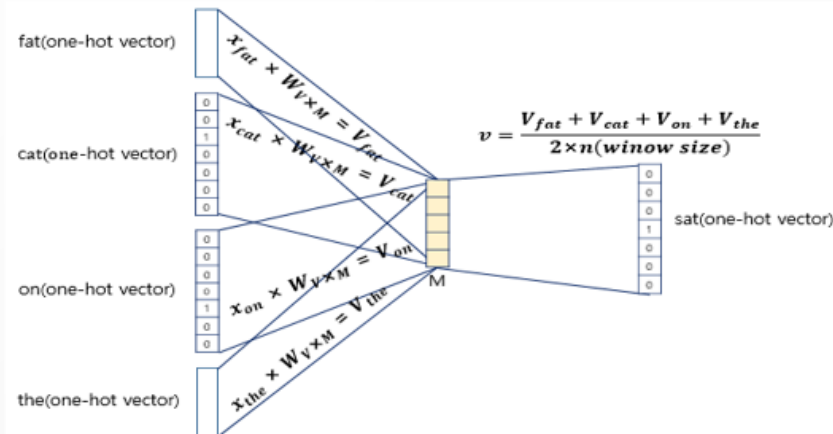
Word2Vec -CBOW(Continuous Bag of Words) , Skip-Gram

- ✓ Word2Vec에는 두 가지 예측 방식이 존재



CBOW

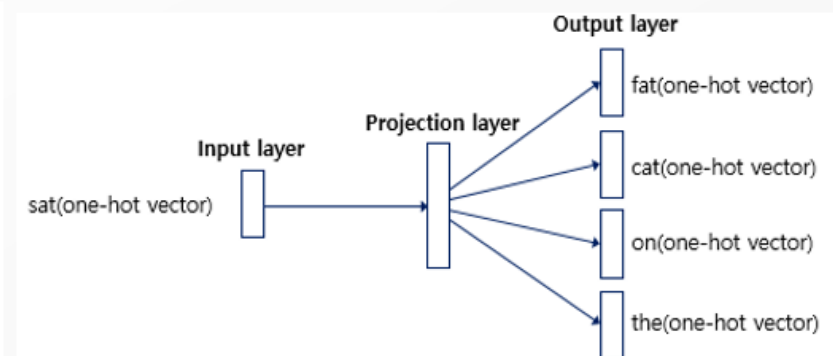
주변의 단어를 통하여 타깃 단어를 예측



fat, cat, on, the 를 통해 sat을 예측

Skip-Gram

중심 단어를 통해 주변 단어를 예측



sat을 통해 fat, cat, on, the 예측

Word2Vec의 단점

모르는 단어 (out of vocabulary) 에 대해서는 벡터 값을 얻을 수 없다.

드물게 나오는 단어에 대하여 학습이 불안정

구조적 특성상 자주 나오는 단어에만 집중 - **오타 자나 드문 단어에 취약하다**

한 단어의 벡터 값을 직접 학습



fastText

단어를 구성하는 **부분단어(subwords)의 벡터 합**으로 표현 -> 오타 자와 새로운 단어에 대응 가능

[where] 학습 시 `where -> <wh, whe, her, ere, re>` `<whe, wheh, here, ere>`

단어 [어디야] 학습 시 [어디야]라는 한 단어가 아닌 [어디][디야]와 같은 subword를 통해 추정 학습
[어딘야] 라고 오타가 날 경우 [어딘][딘야] 는 [어디][디야] 와 비슷한 단어로 인식되지 않을 수 있다.

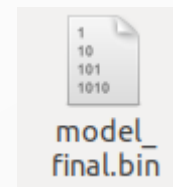


한글의 경우 오타 자는 초성/중성/종성 중 한곳이 틀리기 때문에
subword 통한 학습을 활용하기 위하여 **자모를 분리하여 학습하는** 하는 방안 채택



전 처리(라벨링/자소 분리) 된 데이터를 사용하여 fastText 모델 생성

```
omk@ubuntu:~/pj/code/txt/통과/fastText$ ./fasttext sup
Progress: 100.0% Trials: 3 Best score: 0.964274 ET
Training again with best arguments
Read 22M words
Number of words: 3653080
Number of labels: 3
Progress: 100.0% words/sec/thread: 48243 lr: 0.0000
```



테스트 파일을 통한 정확도 측정 결과
[자소 분리 전 처리 적용 시]

VS

테스트 파일을 통한 정확도 측정 결과
[자소 분리 전 처리 미적용 시]

```
omk@ubuntu:~/pj/
N      363120
P@1    0.964
R@1    0.964
```

```
omk@ubuntu:~/pj/
N      363120
P@1    0.929
R@1    0.929
```

꼬끼오꼬끼오 오늘얏닭이저녁부터울겠네—꼬끼오 여염병하산하래민중이
label 일반

VS

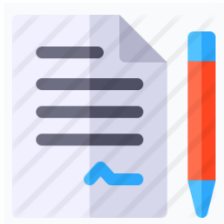
꼬끼오 | 꼬끼오 | 오늘 | 얏닭 | 이저 | 녁부 | 터울 | 겿네 | — | 꼬끼오 | 여 | 엽 | 병 | 하 | 산 | 하 | 래 | 민 | 중 | 이 |
오티르 | 게 | 쓰 | 니 | 게 | — | — | 꼬끼오 | 꼬끼오 | 여 | 엽 | 병 | 하 | 산 | 하 | 래 | 민 | 중 | 이 |
저 | 토 | 오 | | — |
label 욕설

자소 분리한 데이터를
학습시킨 모델에서
정확도가 4% 향상되었다.

프로그램 동작과정



네이버 뉴스 주소를 입력 시
해당 뉴스 댓글 수집



수집 댓글 전처리
(초중중성 분리/
의미 없는 특수문자 등 제거)



fastText

학습시켜 둔
fastText 모델을 통하여
악플/ 의심 / 일반 댓글 분류

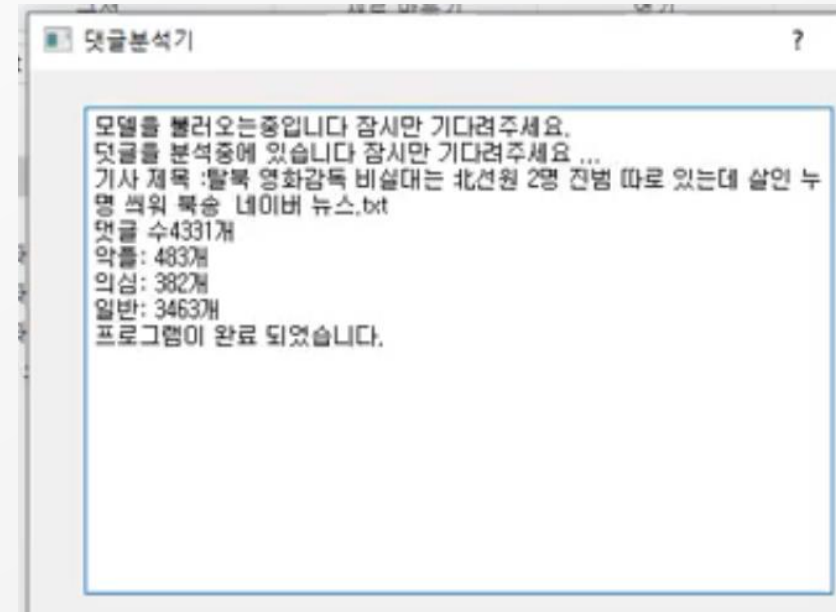


악플 의심 일반

fastText 모델에서 분류된 댓글을
Text파일로 출력

```

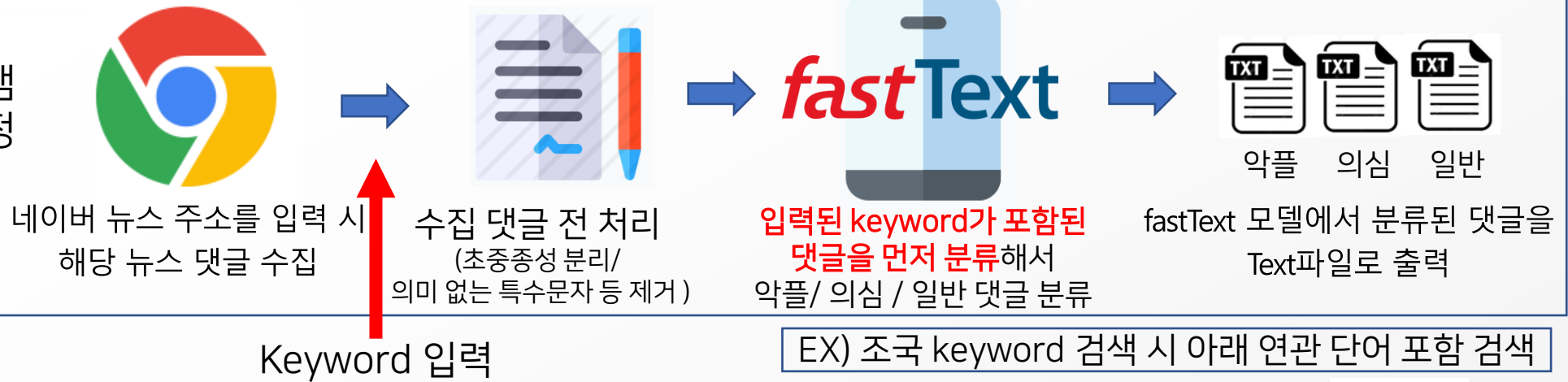
모델 로드 중 ...
기사링크 입력
https://news.naver.com/main/ranking/read.nhn?rankingType=
d=0011110316&date=20190930&type=1&rankingSectionId=1058
[ 1 / 4 ] 댓글 크롤링 중 ...
기사 제목 : 日제품 불매운동 소강 국면 진입했다...소셜미디어
스
댓글 수 1309개
[ 2 / 4 ] 댓글 자소분리 중 ...
[ 3 / 4 ] 자소분리 댓글 fasttext 모델에서 분류 중 ...
악플 : 20개
의심 : 14개
일반 : 317개
[ 4 / 4 ] 분류된 파일 자소분리 복구 중 ...
    
```



악플_日제	의심_日제	일반_日제
품 불매운동	품 불매운동	품 불매운동
소강 국면 진	소강 국면 진	소강 국면 진
입했다...소	입했다...소	입했다...소
셜미디어 연	셜미디어 연	셜미디어 연
관어 급감 네	관어 급감 네	관어 급감 네
이버 뉴스	이버 뉴스	이버 뉴스

추가 기능

프로그램
동작과정



EX) 조국 keyword 검색 시 아래 연관 단어 포함 검색

Keyword로 입력한 단어만 검색하는 것이 아닌
연관 단어를 포함하여 검색
(fastText Word representations)

```
omk@ubuntu:~/pj/code/txt/통과/fastText
Query word? 조국-가ㅏㅓ
조국-가ㅏㅓ-가ㅏㅓ - 0.884797
조국-가ㅏㅓ-가ㅏㅓ - 0.884398
조국-가ㅏㅓ-가ㅏㅓ - 0.884152
조국-가ㅏㅓ-가ㅏㅓ - 0.8834
조국-가ㅏㅓ-가ㅏㅓ - 0.882094
조국-가ㅏㅓ-가ㅏㅓ - 0.877626
조국-가ㅏㅓ-가ㅏㅓ - 0.877516
조국-가ㅏㅓ-가ㅏㅓ - 0.875292
조국-가ㅏㅓ-가ㅏㅓ - 0.873418
조국-가ㅏㅓ-가ㅏㅓ - 0.873242
조국-가ㅏㅓ-가ㅏㅓ - 0.87109
조국-가ㅏㅓ-가ㅏㅓ - 0.870013
조국-가ㅏㅓ-가ㅏㅓ - 0.869626
조국-가ㅏㅓ-가ㅏㅓ - 0.866814
조국-가ㅏㅓ-가ㅏㅓ - 0.866339
```

조국
조국애는
조국캐슬
조국이념
조국
조국너는
조국
조국정경심
조국년
조국짱
조국가족
조국비호
조국외
조국은
왜조국

추가기능 의도

- 키워드로 자신의 제품이나 경쟁사 제품 뉴스에 있는 자신의 제품평을 알 수 있다
Ex)아이폰 뉴스에서 삼성이라는 키워드로 삼성 핸드폰 제품과 비교되는 평을 알 수 있다.
- 댓글 중 자신의 이름이 들어간 댓글을 따로 찾아서 악플과 의심 일반 댓글에 대해 알 수 있다.
모욕죄 또는 명예훼손죄 같은 경우 다수가 객체가 피해자임을 알 수 있어야한다.
따라서 자신의 이름이 있는 악플일 경우 모욕죄 또는 명예훼손죄에 해당한다.

예시)

```
기사링크 입력
https://news.naver.com/main/ranking/read.nhn?rank
d=0004518153&date=20191117&type=1&rankingSectionI
키워드 입력
설리
[ 1 / 4 ] 댓글 크롤링 중 ...
기사 제목 :설리 남친 주장 논란 유투버 그알 등장연
댓글 수 380개
[설리] 키워드 댓글 [ 184 ] 개
[ 2 / 4 ] 댓글 자소분리 중 ...
[ 3 / 4 ] 자소분리 댓글 fasttext 모델에서 분류 중
악플: 19개
의심: 12개
일반: 153개
[ 4 / 4 ] 분류된 파일 자소분리 복구 중 ...
```

댓글분석기

모형을 불러오는중입니다 잠시만 기다려주세요.
댓글을 분석중에 있습니다 잠시만 기다려주세요 ...
기사 제목 :설리 남친 주장 논란 유투버 그알 등장연예인 악플 감내해야
네이버 뉴스.txt
댓글 수 380개
[설리] 키워드 댓글 [184] 개
악플: 19개
의심: 12개
일반: 153개
프로그램이 완료 되었습니다.

-
- 악플_자모결합_설리 남친 주장 논란 유투버 그알
 - 의심_자모결합_설리 남친 주장 논란 유투버 그알
 - 일반_자모결합_설리 남친 주장 논란 유투버 그알

악플_자모결합_설리 남친 주장 논란 유투버 그알 분류 파일 일부 내용

서얼리전남친 ㅋㅋㅋ아무도 안믿어 불량감자닭아가지고ㅋㅋㅋ넌 다시태어나도 설리 발톱만큼도 못따라갈꺼다 불량감자야 거울이나보고 티비나
진짜 정신 나갔나 설리 남자친구라고 어그로 끌고 논란 되니까 남친은 아니고 그 정도로 좋아한 팬이었다 ㅇㅈㄹ 어떤 생각을 가지고 살면 저렇게
생기다가 만 세키가 어디서 설리 이름팔고있어 이세키야 진심 처맞고싶냐
최자가머라한줄알았네ㅋㅋ 설리가아무리외롭고 힘들어도 널만나겠냐 좆도 이쑤시게 만하개 생겨가지고
설리는 정말 안타깝지만 너새기는악플받고 좀 ㄷ져야해

결론

시연

키워드 없이 댓글 분류 :

<https://www.youtube.com/watch?v=PRDsxupsVSI&feature=youtu.be>
<https://www.youtube.com/watch?v=k44nyEB8kj0&t=122s>

키워드를 이용하여 댓글 분류:

<https://youtu.be/VLx0hKOC9ag>
https://www.youtube.com/watch?v=0hSvWh-_8Mo&t=1s

자세한 사항이나 프로그램을 다운받고 싶다면:

<https://github.com/passergy14/Malicious-comments-Block-System>

개선해야할 점

- ❖ 데이터 라벨링에 주관성이 들어가서 객관성 떨어진다.
- ❖ 객관적이고 정확한 기준이 필요로 하다.
- ❖ Gui 개선이 필요하다. 좀더 보기 좋고 편안한 디자인이 되어야 한다.

감사합니다