

악성 댓글 선별 딥러닝 프로젝트 최종발표

20192905 박채원

프로젝트 예상 진행 순서 -> 실제 프로젝트 진행 순서

1. 데이터 수집 (악성댓글 기준에 따른 악플 여부 판별 레이블링)

-> 데이터 가져오기 (koko) / <https://github.com/kocohub/korean-hate-speech>

2. 데이터 전처리 (토큰화)

-> 데이터 전처리 (기호 삭제, 불용어 제거, 형태소 분석(Okt), 패딩, 토큰화)

3. 여러 모델 구축 & 성능 향상을 위한 여러가지 방법 적용 (모델 선정)

-> 모델링 (임베딩층(Word2Vec), bidirectional lstm, dense층)

4. 실제 댓글에 모델 적용

-> 실제 댓글의 악성 여부 예측

		comments	contain_gender_bias	bias	hate	news_title
0	(현재 호텔주인 심정) 아18 난 마른하늘에 날벼락맞고 호텔망하게생겼는데 누군 계속...		False	others	hate	"밤새 조문 행렬...故 전미선, 동료들이 그리워하는 따뜻한 배우 [종합]"
1	...한국적인 미인의 대표적인 분...너무나 곱고아름다운모습...그모습뒤의 슬픔을...		False	none	none	"'연중' 故 전미선, 생전 마지막 미공개 인터뷰...환하게 웃는 모습 '먹먹'[종합]"
2	...못된 놈들...남의 고통을 즐겼던 놈들...이젠 마땅한 처벌을 받아야지...그래...		False	none	hate	"[단독] 잔나비, 라디오 출연 취소→'한밤' 방송 연기...비판 여론 ing(종합)"
3	1,2화 어설프는데 3,4화 지나서부터는 갈수록 너무 재밌던데		False	none	none	"'아스달 연대기' 장동건-김옥빈, 들끓는 '욕망극플'→논물범벅 '칼끝 대립'"
4	1. 사람 얼굴 손톱으로 긁은것은 인격살해이고2. 동영상이 몰카냐? 메갈리안들 생각...		True	gender	hate	[DA:이슈] '구하라 비보' 최종범 항소심에 영향?...법조계 "'공소권 없음' 아냐"
5	10+8 진짜 이승기랑 비교된다		False	none	none	"[공식입장] '1박2일' 측 "'정준영 출연 중단 결정...심각성 고려'"
6	100년만에 남녀간 성전쟁 한번 크게 치룬 후 일부다처제, 여성의 정치참여 금지, ...		True	gender	hate	"배우 마크 월버그, 성차별 논란에 출연료 16억원 전액 기부"
7	10년뒤 윤서인은 분명히 재평가될것임. 말하나하나가 틀리게없음		False	none	none	"[Oh!쎬 초점] "'정우성 공개 저격'" 윤서인, 이쯤 되면 정우성 안티?"
8	10년만에 재미를 느끼는 프로였는데왜 니들때문에 폐지를해야되냐		False	none	offensive	"'故설리 부재 하에 지속 어려워'...'악플의밤' 측이 폐지를 결정한 이유 [종합]"
9	10년차방탄팬인데 우리방탄처럼 성공은못하겠지만 일단 방탄의 부하가되고싶다는거니 이름...		False	none	offensive	"'""프듀2' 출신 합류'" 노와르 "'롤모델 방탄소년단'"(종합)"
10	10살 차이가 넘을텐데 부부라고? 무슨 내용인지 궁금하네..		False	others	none	"이지훈, '99억의 여자' 출연...오나라와 부부 호흡[공식]"
11	12월이나 1월이더 결혼성수기지 5월은 여름이라 비수기에속한다		False	none	none	"김건모 결혼 내년 5월로 연기..." "보다 많은 분 초청"
12	13살 차이면 요즘 세상에 놀랄일도 아니네요		False	none	none	"[종합]'두산家 며느리' 조수애, 결혼→1000만원 부케→임신→출산...하트이슈ing"
13	1400억 뜯기고 얼마나 억울 했을꼬 뒤질만큼 아플만 하지~~~~ㅋㅋㅋㅋ		False	none	offensive	"[룩@자이나] 판빙빙, 급격히 살빠져...대만 의사에 SOS"
14	15년생인 율아들은 유모차 안타는데		False	none	none	"이병헌♥이민정, 4살 아들과 LA 미국 여행 '이병헌...디즈니랜드서 유모차 끌어'"
15	15살차이면 어릴때사고쳤으면 딸벌인데 ㅋㅋ 능력좋은 전현무 여자들이 전현무스타일이 ...		False	others	offensive	"'연중' 이해성이 밝힌 ♥전현무의 모든 것 #공개열애 #반전매력 [전일야화]"
16	166에 57.4 다이어트 필요없지만건강위해더뵈거야다들 더마르게봄		False	none	none	"'""90kg 인증"'..몸무게 공개→실검 1위 짝은 女연예인[Oh!쎬 탐구]"
17	17년도 아니고 27년이면 진짜 너무했다 이것이 사랑만으로 가능할수 있는지 도저히 ...		False	others	none	"[DA:자이나] 이종성, 2번의 이혼 끝에...27세 연하와 세번째 결혼"
18	180넘어갈정도의 기력지는 아님78에서79180넘는애들은 의심.의구심을 안하고안듣그...		False	none	none	"'키 182cm, 농구선수 제의'..'컬투쇼' 권상우, 바둑배고 다 잘하네(종합)"
19	180이하 호빗 한남들은 결혼 하지마셈 ㅋ 돈없으면 연애도 하지마셈 ㅋ 니들 호빗 ...		True	gender	hate	"'라스' 남주혁 "'중학교때 이미 184cm, 2년만에 키 30cm 폭풍 성장'"
20	18키로감량을해야박수쳐주지 그냥 먹고싶은거다먹고디비져자고하면 18키로는그냥꺼		False	none	hate	"[DAY컷] 'SKY캐슬' 조미녀 "케이 男 아닌 女, 캐릭터 위해 18kg 증량..."
21	1959년 자유당 시절에 하버드 로스쿨에 입학하셨다가 할리우드로 진출하셨어요? 대단...		False	none	none	"'할리우드 1세대' 오순택 별세..." "편히 잠드소서" 주모 물결 [종합]"
22	19에서15로 바뀌고 망트리 탄거...		False	none	none	[단독] 'SNL 코리아' 공식적인 폐지 확정...아름다운 종료
23	1년연예 ㅋ 2년됩니다....		False	none	offensive	[단독]클라라가 밝힌 #남편#신혼집#연예활동(인터뷰)
24	1년전인가 둘이 연남동 술집에서 술마시는거 봤는데 이제야 기사뜨는구나		False	none	none	"[전문] 정진운♥경리 열애 인정 "2년째 열애 맞다, 따뜻한 시선 부탁" (공식입장)"
25	1박2일 재미없고 이제 부도덕한것만 생각나요. 폐지하세요 학부모입장에서폐지비합니다		False	none	none	"'방송하자' 차태현 · 김준호, 팬들이 "'이멤버, 리멤버'" 외치는 이유(종합)[O...
26	1보다 재미없어요 진짜 보다 나울뻔		False	none	offensive	[종합]'신과함께2' 14일만에 1천만명 ↑...최초 쌍천만 시리즈
27	1부는 눈 썩는 거 같아서 2부만 봤다 졸라 웃기지만 ㅋㅋ		False	none	none	'나혼자 산다' 金 獨보적 1위...화사→전현무-이시연-기안84 '웃음 폭탄'

실제 네이버 기사 댓글에 적용 해보기



vete****

2021.08.05. 17:26

김연경선수가 아주큰 외교를 했다고 생각합니다.큰박수를보냅니다

답글 6

```
sentiment_predict('김연경선수가 아주큰 외교를 했다고 생각합니다.큰박수를보냅니다')
```

63.94% 확률로 악플이 아닙니다.



alsr****

2021.08.06. 10:41

진짜 일드럽게 못하고 하는것도 없이 모이지마세요 ㅋㅋㅋㅋㅋㅋㅋㅋ취패고싶내 재앙이

답글 작성

```
sentiment_predict('진짜 일드럽게 못하고 하는것도 없이 모이지마세요 ㅋㅋㅋㅋㅋㅋㅋㅋ취패고싶내 재앙이')
```

96.12% 확률로 악플입니다.



dljt****

2021.08.06. 10:50

절라도다운 절라도 절라디언스럽네

```
sentiment_predict('절라도다운 절라도 절라디언스럽네')
```

80.81% 확률로 악플입니다.



kmc7****

2021.08.06. 11:46

개소리는 달나라 가서 해라 별래만도 못한 것들

답글 작성

```
sentiment_predict('개소리는 달나라 가서 해라 별래만도 못한 것들')
```

86.02% 확률로 악플입니다.

부족한 점

↓ 비꼬는 댓글

```
sentiment_predict('안타깝네요 자연에서 나오지 마시고 편하게 사세요') # 비꼬는 댓글에 대해 판별 불가 / 기사 맥락 파악 필요 / sarcasm
```

62.50% 확률로 악플이 아닙니다.

```
sentiment_predict('나는 켈리플라워 진짜 싫어')
```

58.69% 확률로 악플이 아닙니다.

← 사람이 아닌 특정 물체에
대한 혐오감 표현

```
sentiment_predict('나는 최민식 진짜 싫어') # 사람 이름이 들어가면 선풍..? / 고유명사
```

58.69% 확률로 악플이 아닙니다.

← 사람에 대한 혐오감 표현

```
sentiment_predict('개소리 라는 표현은 잘못된 표현입니다')
```

89.11% 확률로 악플입니다.

← normal한 댓글

이어서 더 해보고 싶은 것

- 악성 댓글의 기준을 세우기
- 악성댓글 기준에 따라 직접 네이버 기사 (or 유튜브) 크롤링 & 댓글 라벨링
- 더 다양한 층 쌓아보기
- **many-to-one** (예를 들어, 기사의 제목을 전처리 해줘서 댓글과 함께 넣어주기)
- 감성 분석에 대한 더 깊은 공부

감사합니다 :)  