**유튜브 댓글을 이용한 국회의원 선거 예측**

- 키워드 : '나경원 류삼영'

- 기간 : 24-03-01 ~ 24-06-03

- 총 255개 유튜브 영상 크롤링,

총121050개 댓글(대댓글 포함)

텍스트, 영수증, 번호, 폰트이(가) 표시된 사진

자동 생성된 설명텍스트, 번호, 스크린샷, 폰트이(가) 표시된 사진

자동 생성된 설명텍스트, 번호, 스크린샷, 폰트이(가) 표시된 사진

자동 생성된 설명

Top 50 words

한단어 '류', '영', '나' 등, 뜻이 존재하는 한단어는 일부러 제외하지 않음

텍스트, 스크린샷, 번호, 폰트이(가) 표시된 사진

자동 생성된 설명

참고 : 위키피디아

https://ko.wikipedia.org/wiki/%EB%8C%80%ED%95%9C%EB%AF%BC%EA%B5%AD\_%EC%A0%9C22%EB%8C%80\_%EA%B5%AD%ED%9A%8C%EC%9D%98%EC%9B%90\_%EC%84%A0%EA%B1%B0\_%EC%97%AC%EB%A1%A0\_%EC%A1%B0%EC%82%AC

(1)

real\_event\_times =

['2024-03-11', '2024-03-12', '2024-03-18', '2024-03-25', '2024-03-26', '2024-03-29', '2024-04-01']

텍스트, 번호, 폰트이(가) 표시된 사진

자동 생성된 설명텍스트, 번호, 폰트이(가) 표시된 사진

자동 생성된 설명텍스트, 번호, 폰트, 스크린샷이(가) 표시된 사진

자동 생성된 설명

Top 50 words in odds ratio

라인, 텍스트, 그래프, 도표이(가) 표시된 사진

자동 생성된 설명

순서대로

(Number of comments per minute,

event scores per minute estimated with log odds ratio,

event scores per minute estimated with log odds ratio based on relative frequency)

빨간점 : 실제 이벤트 시점

기존 연구와 다른 점 : 축구 경기는 이벤트 시점(골 들어갈 때)에 댓글이 많이 달리지만,

선거 데이터는 선거 바로 전에 댓글이 많이 달린다.

->>따라서 outlier로 이벤트 시점을 예측하는 것은

축구 경기 골 들어가는 것은 예측 가능,

선거 여론조사 시점 불가능..???

이유 : 여론조사 때 댓글이 많이 달리지 않고

특정 단어에 odds ratio를 반영해서 예측하더라도, 그래프를 보면 여론조사 때(빨간점)이 높은 수치가 아니다

텍스트, 스크린샷, 도표, 라인이(가) 표시된 사진

자동 생성된 설명텍스트, 라인, 그래프, 친필이(가) 표시된 사진

자동 생성된 설명

#k=1.5, residuals\_list\_lin

좌 : 국소선형, 국소 상수회귀 적합 그래프

우 : 이상점 탐지 그래프

텍스트, 평행, 번호, 도표이(가) 표시된 사진

자동 생성된 설명