



소셜 미디어를 활용한 COVID-19 백신 부작용 분석

김수정¹, 박정선¹, 송문영¹, 최원경¹, 김현희^{1*}

¹ 동덕여자대학교 정보통계학과

서론

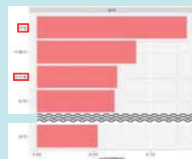
- ❖ 연구 배경 및 목적
 - COVID-19 백신 신속 승인으로 인한 이상반응 연구 부족
 - 이상반응에 대한 데이터가 많지 않은 상황에서 접종 시작
 - 코로나 백신 접종 후 알 수 없는 이상반응 속출
 - 인과성 부족으로 공식적으로 부작용 인정이 어려움
 - 소셜 미디어를 활용하여 일반인들의 즉각적인 반응을 확인하고 도움을 줄 수 있는 방안을 도출하고자 함
- ❖ 연구 내용
 - 소셜 미디어를 통한 COVID-19 백신 이상반응 언급 분석
 - 미디어별 특성에 따른 이상반응 관련 연구 기초 마련
 - 네이버 카페, 네이버 블로그, 트위터, 유튜브

데이터 수집 및 분석

- ❖ 수집 데이터 (수집 기간: 2021.10 ~ 2021.12)
 - 트위터: '코로나 백신 부작용' 키워드로 크롤링
 - 유튜브 댓글: 백신 관련 의료 채널 유튜브 영상 댓글, '백신이름+부작용' 키워드로 크롤링
 - 네이버 카페: '전염병 공유 카페'의 '증상 공유' 게시판의 글
 - 네이버 블로그: '백신 부작용/백신 접종/1차 접종/2차 접종' 크롤링, 코로나 데이터 직접 선별
- ❖ 데이터 전처리 및 분석
 - 텍스트 전처리: KoNLPY의 Okt 이용해 토큰화. 명사 추출
 - 단어사전 구축: 질병청 공식 부작용 기반 단어사전 활용
 - 파이계수 산출: KoNLP의 SimplePos22를 이용해 토큰화. 명사 / 형용사 / 동사 추출
 - 워드 클라우드 시각화: 다빈도 부작용 추출

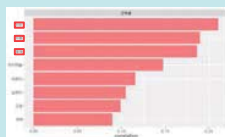
분석 결과

1. 트위터



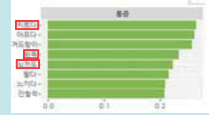
- 워드클라우드 : 생리, 두통, 근육통, 혈전, 흉통 (단어사전 적용시 '생리' 미미함)
- 파이계수 막대그래프: '생리' → '(n)주일', '생리통'과 상관성. 주로 생리 기간 변화(생리불순 등)와 통증에 관한 내용.

2. 유튜브



- 워드클라우드 : 가슴, 근육통, 두통, 심장, 오한, 호흡곤란
- 파이계수 막대그래프 : '근육통' → '오한', '두통', '발열'. '근육통' 발생 시 위 증상이 동반되는 경우 존재.

3. 네이버 카페



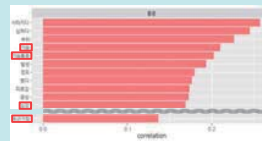
- 워드클라우드 : 심장, 가슴, 두통, 어지러움, 근육통, 흉통
- 파이계수 막대그래프 : '통증' → '찌르다', '왼쪽', '심전도'와 상관성. 주로 심장과 관련된 통증.

4. 네이버 블로그

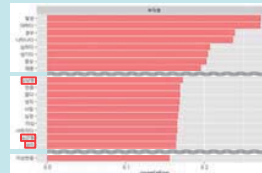


- 워드클라우드 : 두통, 근육통, 오한, 발열, 알레르기

· 파이계수 막대그래프 :



- 1) '통증' → '가슴', '가슴통증', '심장', '두근거림'과 상관성.



- 2) '부작용' → '심낭염', '심근염', '생리' 직접적으로 출현.

결론

- ❖ 주목할 만한 공통적 증상 : '흉통'
 - 흉통의 발현 원인 향후 백신부작용과 관련하여 연구.
- ❖ 소셜 미디어별 특성에 입각한 차이점 :
 - (1) '생리' (트위터)
 - '생리 이상 반응' 타 소셜 미디어에 비해 두드러짐.
 - 현재 부작용으로 인정되는 '이상자궁출혈'을 포함한 '생리불순' 등의 증상도 부작용으로 고려해볼만 함.
 - (2) '심근염, 심낭염' (네이버 블로그)
 - 중증 약물 이상 반응인 만큼, 주로 어떤 경우에 위 증상이 일어났는지 연구할 필요가 있음.
- ❖ <한국어 의료 Bert> 구축 필요성
 - 일반인이 작성한 증상 후기 vs 질병청 공식 부작용
 - 부작용 관련 증상 단어가 부작용 단어로 집계되지 않는 경우 (ex. 머리 - 지끈, 귀 - 먹먹).
 - 한국어 특성에 맞는 <한국어 의료 Bert> 구축하면 향후 연구 진행에 더 좋은 성과를 얻을 수 있을 것.