



seok830621



# 빅데이터 분석을 이용한 정신보건정책 수립 수요 연구(가)

머신러닝의 기초#4

류현희 황규범 현요한



seok830621



## *contents*

1. 연구계획

2. 선행연구조사

  
seok830621

- 다양한 분야에서 빅데이터를 이용한 의사결정 방안 활용 및 확대
- 코로나 19 이후 사회에 다양한 변화 발생
- 빈곤·질병·실업 등에 의해 우울증세가 증가
- 우울, 스트레스, 자살 등은 이전부터 사회의 심각한 문제

  
seok830621

- 빅데이터 분석을 통해 우울, 스트레스, 자살 등 키워드에 대한 감정 분석 실시
- 정신보건정책 수립에 도움을 주기 위함

  
seok830621

- 크롤링을 통한 소셜 빅데이터 수집
- 수집 데이터 자연어 처리
- 키워드 설정 및 감정과 정책 수요 분석
- 분석결과와 현 보건정책 비교분석
- 정책제안



seok830621

# 선행연구조사



5



제목	발행연도	사용모델	사용데이터	한계점
소셜 빅데이터 기반 보건복지 정책 미래신호 예측	2016	EDA: 단어빈도, 문서빈도, TF-IDF 연관분석: Apriori 알고리즘 분류예측: Random Forest	- 소셜 빅데이터	<ul style="list-style-type: none"><li>- 3개월에 한정된 소셜 빅데이터</li><li>- 집단을 기준으로 분석하여 개인에게 적용할 경우 생태학적 오류가능성</li><li>- 문서내 의미와 이론적 모형 내의 의미차이 발생 가능성</li></ul>
기후변화 정책 수립 지원을 위한 소셜 빅데이터 분석	2019	NLP: KoNLPy, Doc2Vec, Paragraph Embedding 토픽 모델링 평가: 응집성 분석 시각화: pyLDAvis	- 소셜 빅데이터	<ul style="list-style-type: none"><li>- 데이터를 수집한 네이버 블로그의 이용자수가 적음</li><li>- 데이터를 해석 및 분석할 때 연구자의 주관 개입가능성</li><li>- 인터넷과 SNS 이용자의 의견만 수집</li></ul>
데이터 사이언스를 활용한 출산 정책 관련 여론 분석	2017	NLP: 유의어-제외어 처리, 한글자연어처리 의미망 분석: 백본모델,거번-뉴먼 알고리즘 키워드 분석: 보나시치 영향력 지수	<ul style="list-style-type: none"><li>- 인터뷰 데이터</li><li>- 온라인 커뮤니티 데이터</li></ul>	<ul style="list-style-type: none"><li>- 여성 미성년자의 삶의 질을 통합적으로 고려하지 않음</li><li>- 가설과 검증이라는 기존의 연구방법론에 적합하지 않음</li><li>- 필요한 모든 자료와 변수 수집 불가능</li><li>- 변수의 관계를 찾는 것에 한계가 있음</li></ul>