

소셜 빅데이터 분석을 통한 정신건강정책 제안

Proposal of Mental Health Policy through Analysis of Social Big Data

류현희¹ · 황규범² · 현요한³ · 오하영⁴

Hyunhee Ryu · Kyubum Hwang · Yohan Hyun · Hayoung Oh

요약: 본 연구는 코로나19 이후 우울증세, 스트레스 등의 증가와 빅데이터 분석을 이용한 의사결정 방안이 다양한 분야에서 활용 및 확대됨에 따라 소셜 빅데이터 분석을 이용해 정신건강정책을 제안하고자 한다. 2021년 11월 01일에서 2022년 10월 31일 동안의 소셜 빅데이터를 수집하여 자연어전처리 과정을 거쳐 BERT 모델을 활용한 키워드 추출, 토픽모델링을 진행하였다. 이를 통해 노인과 정신 장애를 겪는 사람들도 위한 복지 시설 마련, 정신건강 관련 프로그램 구축 및 강화의 필요성 등의 정책을 도출하였다. 또한, 현 정신건강정책과의 비교 분석을 통해 노인·환자·정신질환자를 위한 복지 시설의 확충 및 복지의 강화 등 정책의 방향을 제시하였다.

Abstract: This research aims to propose mental health policies with the application of social big-data analysis. Especially since COVID-19, mental health issues such as depression rates and stress have increased, and also big-data analysis has been used and expanded in various fields. In this research, we used social big-data collected from 2021.11.01 to 2022.10.31. After preprocessing and natural language processing, we used the BERT model for keyword extraction and topic modeling. With the findings, this paper shares results to improve the welfare for elderlies and mentally disabled people as well, including the provisions of facilities and mental health programs. In addition, new policies and policy directions for the elderly, sick and mentally-challenged are further recommended through comparative analysis with the current mental health policies.

¹ 성균관대학교 경영학과

² 성균관대학교 경영학과

³ 성균관대학교 경영학과

⁴ 교신저자, 성균관대학교 인공지능융합학과 부교수

I. 서론

코로나19 이후 사회에 다양한 변화가 발생하였으며, 빈곤·질병·실업 등에 의해 우울 증세가 증가하였다(박상미, 2020). 2·30대 청년층의 스트레스가 취업과 대인관계 등으로 인해 증가하였고(통계청, 2021), 높은 우울 수준은 자살 생각을 높이는 매개 역할을 하고 있다(김효연·박한중, 2022). 심지어 코로나 발생 후 이전보다 정신건강 상태가 많이 악화되었다. 인제대학교 보건대학원 조사에 따르면 코로나 장기화로 인해 국민 55% 이상이 우울감을 느꼈고 작년인 2021년도 관측된 결과인 46%보다 증가한 수치이다 (남동현, 2022).

하지만 사람들의 정신건강에 대한 인식 수준은 높지 않다. “2017년도 한국에서는 거의 4명 중 1명이 정신 장애를 앓고 있지만 10명 중 1명만이 치료받는 국가(Nagar, 2022)”라고 Harvard International Review에서 설명했다. 이를 해소하기 위해 보건복지부에서는 관련 예산을 증대시키고 있으며, 정신건강복지기본계획(‘21~‘25)에 따르면 2025년까지 5년 간 2조 원(연평균 4000억)이 정신건강 증진을 위해 매년 투입될 예정이다. 정신건강 문제는 육체적 건강과 달리 외적으로 드러나지 않아 조기 발견이 어렵고 사회적으로 낙인효과가 발생할 수 있기 때문에 정신건강 관련 문제를 정책을 통해 효과적으로 해결해야 한다(이상영, 2012).

또한, 현재 다양한 분야에서 빅데이터와 인공지능을 기반으로 하는 의사결정 방안을 활용하고 확대하고 있다. 이는 정책 결정 분야에서도 마찬가지이다(김병조·은종환, 2020). 빅데이터는 3V(volume, velocity, variety)를 특징으로 하는 데이터이다. 지난 몇 년간 빅데이터에 대한 논의에 따라 위의 3V에 2V(value, veracity)를 추가한 개념으로 정의되고 있다. 이 중 소셜 빅데이터는 소셜미디어와 블로그, 커뮤니티 등 온라인 상에서 실시간으로 생산하는 대규모 데이터를 지칭한다. 소셜 빅데이터는 개개인이 직접 생산하는 데이터이기 때문에 온라인 상에서 시민들의 생각을 수집할 수 있다는데에 의미가 있다. 과학기술통신부의 정책·기술 동향에 따르면 정신건강관리 분야에서 코로나19로 사회적 거리두기와 봉쇄 조치가 이뤄지고 불안과 우울증을 겪는 사람들이 늘어남에 따라 원격 솔루션이 필수적인 실태이다. 현 정부의 중점 추진 과제 중 하나로 인공지능·데이터 기반의 정책 의사결정 지원체계 구축과 국가 현안·난제 해결을 위한 민관 협업 활성화가 있으며 국민들에게 맞춤형으로 간편하게 서비스를 제공하는 서비스가 필요하다. 이에 더해 데이터 기반 국정운영 확대를 위해 소셜 빅데이터와 빅데이터 분석을 통한 정책 수립은 중요한 요소 중 하나가 될 것이다. 공공서비스 또한 민간 서비스처럼 상황과 필요에 따라 빠르게 변화하고 구현될 수 있도록 하기 위해서도 소셜 빅데이터 분석을 통한 정책 수립이 필요하다.

소셜 빅데이터 분석은 기본적으로 텍스트 분석을 기반으로 하는 경우가 많다. 이에 분석의 효율성 및 정확도를 높이기 언어모델을 사용할 수 있다. 언어모델이란 언어를 수학적으로 표현하기 위해 제작한 모델을 의미한다. 언어모델 사용 사례로는 재택근무로 인한 내부자 위협을 감소하기 위해 BERT의 사전학습 방식인 마스크 언어 모델을 이용한 내부자 위협행위 탐지 모델과(손민혁, 2022), 딥러닝을 이용한 기학습 언어모델을 활용한 기계 독해 기술이 있다(김성현, 2019).

증거기반의사결정과 증거의 다원화의 필요성이 요구되는 현시점에서 본 연구는 온라인 채널을 통해 수집한 소셜 빅데이터 분석과 정신보건 정책 관련 키워드에 대한 분석을 통해 국민들의 정신건강 문제에 대한 예방대책 및 완화 방안 중 하나인 정책 제안을 위한 연구를 진행하고자 한다.

II. 선행연구 고찰

제목	발행연도	분석방법 및 사용모델	사용데이터	한계점
소셜 빅데이터 기반 보건복지 정책 미래신호 예측	2016	EDA: 단어빈도, 문서빈도, TF-IDF 연관분석: Apriori 알고리즘 분류예측: Random Forest	- 소셜 빅데이터	- 3개월에 한정된 소셜 빅데이터 - 집단을 기준으로 분석하여 개인에게 적용할 경우 생태학적 오류가능성 - 문서내 의미와 이론적 모형 내의 의미차이 발생 가능성
기후변화 정책 수립 지원을 위한 소셜 빅데이터 분석	2019	NLP: KoNLPy, Doc2Vec(word embedding) 토픽 모델링 토픽 모델링 평가: 응집성 분석 시각화: pyLDAvis	- 소셜 빅데이터	- 데이터를 수집한 네이버 블로그의 이용자수가 적음 - 데이터를 해석 및 분석할 때 연구자의 주관 개입가능성 - 인터넷과 SNS 이용자의 의견만 수집
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	- 인터뷰 데이터 - 온라인 커뮤니티 데이터	- 여성 미성년자의 삶의 질을 통합적으로 고려하지 않음 - 가설과 검증이라는 기존의 연구방법론에 적합하지 않음 - 필요한 모든 자료와 변수 수집 불가능 - 변수의 관계를 찾는 것에 한계가 있음

한국미혼모에 대한 관점 변화와 정부정책의 방향: 1995년~2020년 소셜미디어 빅데이터 분석	2021	EDA: 단어빈도, TF-IDF, N-gram 네트워크 분석 의미연결망 분석(CONCOR분석)	- 소셜 빅데이터	- 빅데이터 활용이 높지 않았던 시기도 연구 기간에 포함하였기 때문에 당시 사회를 대변하지 못함 - 키워드인 '미혼모', '싱글맘', '비혼모'가 시대별로 사용되었기 때문에 정확하게 비교하는 데에 한계가 있음
---	------	--	-----------	--

기존의 연구에서는 주로 뉴스, 블로그, 트위터, 커뮤니티와 같은 온라인 채널을 통해 웹 문서를 크롤링하거나 설문조사와 같은 데이터 수집이 이루어졌다. 주제분석 및 감정분석을 실시하거나 Apriori 알고리즘으로 연관분석을, Random Forest를 활용해 분류예측을, word2vec 알고리즘을 활용해 word embedding, paragraph embedding을 시행하였다. 또한, 백본 모델을 통해 특징을 대표하는 핵심 의미망을 만들어 상대적으로 파급력이 강한 키워드를 찾기 위해 보나시치 영향력 지수를 사용하는 등 다양한 방법을 활용하였다.

보건정책과 다른 분야 정책 수립에 관한 선행연구를 살펴보았을 때 나타난 한계점들은 다음과 같다. 첫 번째로 데이터를 수집한 기간이 짧거나 이용자 수가 적은 소셜 네트워크 서비스의 데이터를 사용한 점이다.(송태민·송주영, 2016). 두 번째로는 데이터를 해석 및 분석할 때, 연구자의 주관이 개입할 가능성이 있다는 것이다(정예림·강정은, 2019). 세 번째는 변수 간의 관계를 찾는 것에 한계가 있다는 점이다(김학준·나세훈, 2017).

word2vec 알고리즘은 단어를 다차원 공간에 임베딩하는 알고리즘이다. 하지만, 이 알고리즘은 문맥을 고려하지 못한다는 한계점이 존재한다. 이를 보완하기 위해 사전학습 언어모델인 BERT를 활용하여 정책수립 지원을 위한 모델의 성능을 향상시키고자 한다. BERT 모델은 2018년 구글이 공개한 사전 훈련 언어모델로 문맥을 반영한 임베딩을 사용한다. 또한, BERT는 단어보다 더 작은 단위로 쪼개는 subword tokenizer를 사용하기 때문에 오타자와 OOV(Out of Vocabulary)를 임베딩하지 못하는 문제를 해결할 수 있다. 하지만, BERT로 정확한 모델을 구현하기 위해서는 데이터 크기가 커야 하며 큰 데이터를 학습하려면 시간이 오래 걸린다는 단점도 존재한다.

BERTopic과 같은 알고리즘 변형들은 단어와 문장을 벡터화 하는 데에 큰 기여를 했다. BERTopic은 사전학습 언어모델을 사용하여 데이터 임베딩 후 클래스 기반 TF-IDF를 활용하여 클러스터링 할 수 있다. 클러스터에서 중요 의미를 유지하며 쉽게 해석할 수 있는 텍스트를 도출할 수 있으며 전반적으로 안전하게 처리하고 광범위한 상황에서 사용할 수 있다는 장점을 가지고 있다. 또한 사용 및 미세 조정에서 상당한 유연성을 보여준다(Grootendorst, 2022). 백 개 이상의 언어로 데이터 훈련을 하고 모델을 구현할 수 있다는 큰 장점이 있지만 한 문서에 여러 주제가 존재할 수 있다는 점을 고려하지 못하는 단점이 있다.

III. 연구방법

첫 번째 단계는 데이터 수집이다. 소셜 빅데이터 수집은 네이버 뉴스, 블로그 게시글의 출처에서 크롤링을 통해 2021년 11월 1일부터 2022년 10월 31일까지의 데이터를 대상으로 하였다. 파이썬의 **Selenium** 패키지, **requests** 패키지를 사용해 **HTML**을 수집하고 **BeautifulSoup4** 패키지를 이용해 **HTML**에서 텍스트 데이터를 추출하였다. 이후, 수집한 텍스트 데이터를 **re** 모듈을 이용한 정규표현식으로 데이터 정제, **KoNLPy** 등의 모듈을 이용한 형태소 분석을 통한 토큰화 등의 전처리를 진행하였다. 형태소 분석을 통해 명사와 동사만을 추출한 문서들을 **n-gram**을 추출할 수 있는 카운트 기반 벡터화 알고리즘을 사용하기 위해 파이썬의 **sentence_transformers** 패키지를 사용하여 단어를 추출하였다. 문서와 문서로부터 추출한 키워드를 **SBERT**를 통해 문서와 가장 유사한 키워드들을 추출하였다. 이후 텍스트에서 잠재적 의미구조의 발견을 위해 토픽 모델링을 진행하였다. 토픽 모델링은 **BERT embeddings**와 클래스 기반 **TF-IDF**를 활용하여 중요한 단어를 유지하면서 밀집된 클러스터를 만들 수 있는 **BERTopic** 알고리즘을 활용하였다. 이는 파이썬의 **bertopic** 패키지를 사용하여 진행하였다. 토픽 모델링 결과를 해석하여 도출된 결과와 현 정부가 시행하고 있는 정신보건 정책과의 비교를 통해 현재 정책과 비교·분석하였으며, 이러한 절차를 거쳐 최종적인 정책 수립 방향과 정책을 제안하였다.

IV. 소셜빅데이터 분석결과

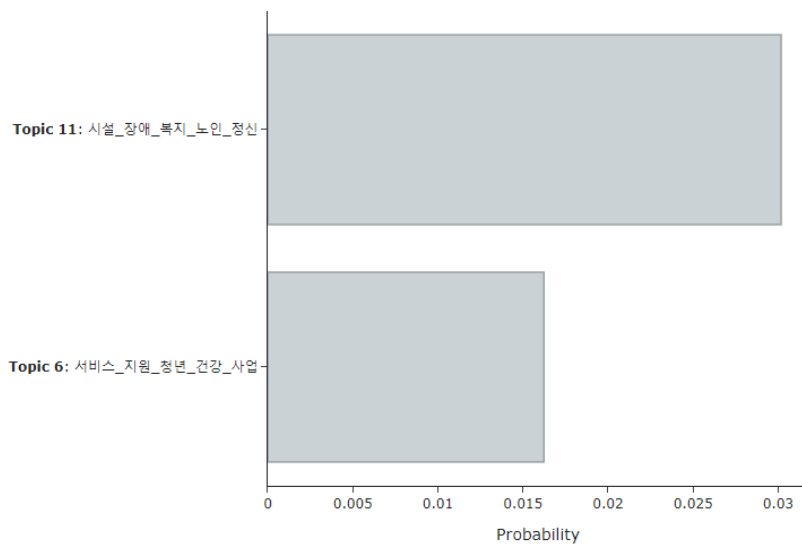
keybert를 통한 분석에서는 데이터의 전처리 부족 및 메모리 관련 문제로 인해 유의미한 결과를 얻지 못하였지만 **BERTopic**을 통한 토픽 모델링에서는 유의미하다고 보여지는 결과를 확인할 수 있었다.

아래는 토픽 모델링 결과의 **Distance map**과 **topic probability distribution**에 대한 시각화이다.

Fig 1. Blog_Intertopic Distance Map



Fig 2. Blog_Topic Probability Distribution



네이버 뉴스 데이터 토픽모델링 결과 [‘지원’, ‘건강’, ‘정책’, ‘코로나’, ‘정신’]의 키워드가 도출되었다. 이러한 키워드는 현재 한국에서 3년동안 지속되는 코로나19 상황이 반영된 것으로 보인다. 그만큼 정신건강 복지를 지원하기 위한 구체적인 정책의 필요성을 알 수 있다.

네이버 블로그 데이터를 통해서는 네이버 뉴스 키워드와 유사하게 [‘코로나’, ‘지원’, ‘정신’, ‘복지’, ‘정책’, ‘의료’]의 키워드가 도출되었으며 더 나아가 [‘장애’, ‘불안’, ‘자살’, ‘우울’]을 포함한 50여 개의 정신건강 토픽 키워드를 확인할 수 있다. 또한, [‘확대’, ‘강화’] 등의 키워드와 같이 정책에 대한 수요 키워드가 도출되었다. 이를 통해 누구나 겪을 수 있는 다양한 정신건강문제에 대한 정책 강화의 수요를 알 수 있다.

V. 현재 정책과 비교 및 정책제안

정신건강종합대책(2016)은 입원, 퇴원과 관련있는 자기결정권 보호를 주요 문제로 삼았고, 정신건강복지법(2016)은 취약한 복지지원을 문제로 규정하였다. 전 세계적으로 많은 정부와 기업들도 심리적 안전에 많이 투자하고 있다. 그만큼 개인이 혼자서 정신건강을 치료하는 것에는 한계가 있고 정부가 나서서 손을 내밀어야 한다는 것이다. 안타깝게도 오늘날 많은 한국인들이 다양한 정신건강 문제로 고통 받고 있지만 현대 정신정책은 많이 미진하다. 2016년 2월에 한국 정부는 청소년들의 정신건강을 위한 종합대책을 발표하였다. 정신건강 뿐만 아니라 육체적 건강도 밀접한 연관성이 있다는 것을 파악하였으나 이를 제대로 고려한 정책 혹은 사업은 없었다(최지희, 2017). 전체적으로 정신건강에 대한 인식과 정책적 노력은 아직까지도 부족하다. 정책전환기인 2016년~2020년에는 정신건강복지법 제정을 통해 복지지원이 주요 정책문제로 규정되어 긍정적 변화가 있었으나 복지 지원 영역의 협소함, 복지 지원 방법의 구체성 부족 등의 한계가 있다(김문근, 2021).

코로나 팬데믹 이후 정신건강에 대한 인식이 높아져 국가에서도 지원 정책을 발표하였고, 2020년 이후 여러 지원 방향과 전망이 보인다. 코로나 이후 우울, 불안 증상 관련 정신질환이 확장되고 있으며 정부는 ‘코로나19 대응 심리지원정책(2021)’을 발표하였다. 하지만 이 대응도 격리자와 감염자 대상으로만

초점이 맞추어져 있고 포괄적인 정신적 프로그램을 실행하지 않고 아직 방역에 집중하는 모습이다(이민우, 2021).

현재 정부에서 시행하는 정신건강 관련 정책과 BERTopic으로 생성된 키워드들의 조합으로 도출된 정책, 그리고 현 정책과 도출된 정책의 비교는 아래 표와 같다.

현재 정책	도출된 정책	방향 및 제언
청소년들의 정신건강을 위한 종합대책 발표	- [청소년, 정신, 아동, 건강, 증진] 아동과 청소년의 정신건강을 위한 정책과 지원 프로그램의 증진이 필요하다.	청소년 정신질환과 자살률은 꾸준히 증가하는 추세이다. 현재 정책의 정신질환실태조사 주기(5년)를 줄이는 등의 지원으로 조기발견 및 치료가 필요하다.
2022년 청년마음건강지원사업	- [청년, 자립, 지원, 사업, 예산] 청년들의 건강한 자립을 위해 여러 복지 사업을 확장하고 있다. 국가는 건강한 미래를 위해 끊임없는 지원과 예산을 투자해야 한다.	2021년부터 2022년까지의 데이터에서 다음과 같은 정책이 도출되었으므로 2022년 시행된 청년마음건강지원사업의 효과성에 대해 좀 더 지켜보고 미흡하다면 강화하기 위한 지원이 필요할 것으로 보여진다.
2022년 청년마음건강지원사업	- [청년, 상담, 심리, 부담금, 건강] 국가의 부담금을 통해 청년들의 정신건강을 위한 심리 상담 센터 등 많은 지원이 필요한 상황이다.	우울감이나 취업으로 인해 심리적 어려움을 겪고 있는 청년을 대상으로 하는 전문심리상담서비스를 제공하고 있지만 우선 지원 대상이 있기 때문에 지원 대상의 확장이 필요하다.
노인, 환자, 정신질환자 위한 정신건강종합대책	- [시설, 장애, 복지, 노인, 정신] 노인이나 정신 장애를 겪는 사람을 위한 다양한 복지 및 시설이 필요하다.	과거부터 노인, 환자, 정신질환자를 위한 정신건강대책이 시행되었으나, 여전히 관련된 시설과 복지의 필요성이 나타났다. 따라서 관련 정책의 강화 및 보완이 필요한 것으로 보여진다.
코로나19 대응 심리정책지원	- [지원, 건강, 정책, 코로나, 정신] - [확대, 지원, 강화, 추진, 구축] 코로나 이후 정신건강을 위한 지원이 필요하다. 또한, 관련 서비스의 체계적인 구축과 지원확대를 추진 역시 필수적이다.	현재 정부는 코로나 감염 확진자 등에 대한 심리 지원을 실행하고 있으나 지역사회 주민 전체를 대상으로 정신건강복지서비스를 체계적으로 제공하지 못하고 있다. 심리상담소와 같은 시설 개설 및 관련 사업에 초점을 맞춰야 한다.

2022 자살예방사업	- [자살, 우울, 정신, 증가, 건강] 우울감 경험을, 정신질환인 스트레스의 인지율, 자살률은 꾸준히 증가하고 있다. 이를 완화하기 위한 구체적인 정신건강 정책이 필요하다.	제5차 자살예방정책위원회에서는 자살예방에 관한 보다 효과적인 정책의 탐색과 수립이 필요한 것으로 보인다.
-------------	--	--

VI. 결론 및 한계점

본 연구는 소셜 빅데이터를 이용하여 키워드 추출, 토픽 모델링을 통해 정신건강 정책을 제안하고 방향을 수립하고자 하였다. BERTopic을 통한 토픽 모델링을 통해 ['건강', '정책', '정신', '우울', '확대'] 등 5가지 단어로 구성된 토픽 키워드를 도출하였고 이를 해석하여 정책을 도출하고 방향을 제안하였다.

정부에서는 2016년 2월 '행복한 삶, 건강한 사회를 위한 정신건강 종합대책'에서 청소년들의 정신건강 증진을 위한 방안을 발표하였음에도 불구하고, 청소년 정신질환은 꾸준히 증가하였으며, 토픽모델링 결과에서도 [청소년, 정신, 아동, 건강, 증진]의 토픽으로 이에 관련된 내용을 확인할 수 있었다. 이에 따라 더욱 효과적인 지원이 필요하며, 이를 위해 현재 5년인 정신질환실태조사의 주기를 짧게 하여 조기발견을 목표로 해야 할 것이다. 더불어, 2022년 청년의 정신건강을 위해 청년마음건강지원사업을 진행하였음에도 불구하고, [청년, 자립, 지원, 사업, 예산], [청년, 상담, 심리, 부담금, 건강] 등의 토픽을 통해 여전히 상담과 지원에 대한 수요가 큰 것으로 확인되었으며, 해당 정책의 효과성을 확인하고 개선할 필요가 있다.

정부는 노인, 환자, 정신질환자를 위한 정신건강종합대책을 시행하였으나 [시설, 장애, 복지, 노인, 정신]의 토픽을 통해 여전히 노인과 정신 장애를 겪는 사람을 위한 복지 지원과 시설이 미비함을 짐작할 수 있다. 이는 구체적인 지원을 위한 근거마련을 통해 현재 시행중인 정책을 보완하여야 함을 의미한다.

토픽 [지원, 건강, 정책, 코로나, 정신], [확대, 지원, 강화, 추진, 구축]은 코로나로 인한 우울 관련 정책에 대한 수요가 더 크다는 것을 암시한다. 현재 정부는 '코로나19 심리지원 정책'을 발표하며 코로나 감염 확진자 등에 대한 심리 지원을 실행하고 있으나 지역사회 주민 전체를 대상으로 정신건강복지서비스를 체계적으로 제공하지 못하고 있다. 심리상담소와 같은 시설 개설 및 관련 사업에 초점을 맞춰야 한다.

정부에서 매년 자살예방사업을 시행하고 있고, 2022년도 역시 자살예방사업을 시행하고 있지만, 세계적으로 대한민국의 자살률과 우울감을 느끼는 정도는 크게 줄어들고 있지 않다. 또한 [자살, 우울, 정신, 증가, 건강]의 토픽을 통해 시민들이 느끼는 자살, 우울감의 증가, 정신건강에 대한 우려 역시 찾아볼 수 있었다. 이에 따라 제5차 자살예방정책위원회에서는 보다 효과적이고 실용적인 관련 정책의 탐색과 수립이 필수적인 것으로 보여진다.

사전학습 언어모델을 활용한 정책제언 연구는 향후 소셜 빅데이터 분석을 통한 정책제언 방법론의 정교화와 일반화를 통해 다양한 분야에서의 정책수립의 기반이 될 수 있다.

하지만 위 연구의 한계점도 분명히 존재한다. 첫째로, 데이터를 크롤링하는 과정에서 ‘정신건강정책’을 검색어로 모든 블로그 게시글과 네이버 뉴스데이터를 수집하였으나 분석에 불필요한 데이터가 다수 포함되어 분석결과의 정확도를 낮아졌을 가능성이 크다. 두번째로, 토픽모델링의 결과 해석에 있어 단어를 기반으로 정책을 도출해야 하므로 주관적 견해가 포함될 수 밖에 없다. 마지막으로 소셜 빅데이터를 수집할 수 있는 채널이 네이버 내의 자료로 한정적이었기 때문에 다른 채널을 이용하는 사용자들의 견해를 반영할 수 없었다.

참고문헌

김문근. (2021). 국가 정신건강복지정책의 문제정의 특성과 변화에 관한 연구 -의료관점과 인권관점에 근거한 문제구조화 특성 분석-. 사회복지정책, 48(3), 5-38.

김성현, 김성만, 황석현. (2019). 대규모 지식그래프와 딥러닝 언어모델을 활용한 기계 독해 기술. 한국정보처리학회 학술대회논문집, 26(2), 922-925.

남동현. "코로나-19 팬데믹 전후 개인특성과 생활습관, 정신건강 간의 관련성 : 성향점수매칭법을 활용하여." 국내석사학위논문 인제대학교 보건대학원, 2022. 경상남도

문유경(You Kyung Moon);전기택(Ki Taek Jeon);배호중(Ho Joong Bae);김근태(Keun Tae Kim). (2017). 여성가족 빅데이터의 활용을 통한 정책 개선과 발굴(Ⅰ). 한국여성정책연구원 연구보고서, 2017(53), 75-119.

서동희, 전복선.(2021).한국미혼모에 대한 관점 변화와 정부정책의 방향 : 1995년~2020년 소셜미디어 빅데이터 분석.한국융합학회논문지,12(12),305-313.

손민혁, 이상준. (2022). BERT 마스크 언어 모델을 활용한 내부자 위협행위 탐지. 한국지식정보기술학회 논문지, 17(4), 703-714.

송태민, 송주영.(2016).소셜 빅데이터 기반 보건복지 정책 미래신호 예측. 보건정보통계학회지,41(4),417-427.

이민우. (2021).코로나19 대유행이 가져온 정신건강 위기와 대응 정책과제.국회입법조사처 제1880호, 3.

이상영.(2012).사회적 여건변화에 따른 정신건강증진정책의 추진방향과 과제.보건복지포럼 2012년 5월 통권 제187호,68-77.

이형하.(2021).빅데이터를 활용한 빈곤, 실업, 질병, 우울증과의 구조모형 연구 : Google 트렌드를 중심으로.한국컴퓨터정보학회 학술발표논문집 ,29(1),119-120.

정예림, 강정은.(2019).기후변화 정책 수립 지원을 위한 소셜 빅데이터 분석.환경정책,27(4),211-239.

최지희, 전진아.(2017). 청소년의 정신건강 현황과 건강행태와의 관련성. *Health and Welfare Policy Forum*, 2017(3), 72–83. <https://doi.org/10.23062/2017.03.7>

Grootendorst, Maarten. (2022, March 11). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. Cornell University - ArXiv. <https://doi.org/10.48550/arxiv.2203.05794>

Kim, T.-H., Cho, D.-B., Lee, H.-Y., Won, H.-J., & Kang, S.-S. (2020). BERT 모델을 이용한 감정 분석 시스템. 2020 온라인 추계학술발표대회 논문집 제27권 제2호, 1.

Nagar, S. (2022, March 11). *The struggle of mental health care delivery in South Korea and Singapore*. Harvard International Review. Retrieved October 6, 2022, from <https://hir.harvard.edu/the-struggle-of-mental-health-care-delivery-in-south-korea-and-singapore/>

Scarpino, I., Zucco, C., Vallelunga, R., Luzzza, F., & Cannataro, M. (2022). *Investigating Topic Modeling Techniques to Extract Meaningful Insights in Italian Long COVID Narration*. BioTech, 11(3), 41. <https://doi.org/10.3390/biotech11030041>



류현희(Hyunhee Ryu)

성균관대학교 경영학과

관심분야: 경영전략

선행연구 논문 수집, 분석 결과, 현재 정책 비교 및 제안 작성



현요한(Yo Han Hyun)

성균관대학교 경영학과

관심분야: AI 마케팅, 경영전략

보건복지부정책 및 정신건강정책 관련 논문자료 수집

초록, 선행 연구 요약, BERT 모델, 현 정책 비교 내용 작성



황규범(Kyubum Hwang)

성균관대학교 경영학과

관심분야: Natural Language Processing, Data Analysis, Deep Learning

연구주제 선정, 선행연구논문 수집, 크롤링 코드 구현, 데이터

전처리, 분석 코드 구현, 최종보고서 작성