

챗봇 데이터에 나타난 우울의 범주와 특성에 관한 연구

진효진*, 백금희**, 차지영**, 최정희***, 임현승***, 차미영****
 *기초과학연구원 데이터사이언스그룹, **이화여자대학교 간호대학
 심심이 주식회사, * KAIST 전산학부
 tesschin@ibs.re.kr, dnjsxka486@naver.com, chiyoung@ewha.ac.kr,
 sijay@simsimi.com, hslim@simsimi.com, meeyoung.cha@gmail.com

A study on the categories and characteristics of depressive moods in chatbot data

HyoJin Chin*, Gum-hee Baek**, Chiyoung Cha**, Jeonghoi Choi*** ,
 Hyunseung Im****, Meeyoung Cha*****
 *IBS Data Science Group, **Ewha Womans University College of Nursing,
 SimSimi inc, * KAIST School of Computing

요 약

챗봇의 사용 용도는 일상 대화와 소비자 응대를 넘어서 심리 상담 용도로 확장하고 있다. 이 연구에서는 챗봇-사람 채팅에서 무작위로 추출한 ‘우울’과 관련된 대화 데이터를 텍스트마이닝 기법으로 분석하여 채팅에서의 우울 관련 담론 주제를 파악하였다. 더불어 정성 분석을 통해 사용자들이 챗봇에 털어놓고 있는 ‘우울’의 종류를 범주화하고 분류하여, 트위터의 ‘우울’ 데이터와의 차이점을 비교하였다. 이를 통해 챗봇 데이터의 ‘우울’ 대화만의 특징을 파악하고, 우울 증상 탐지와 그에 따른 적절한 심리지원 정보를 제공하는 서비스 디자인의 착안점을 제시한다.

1. 서론

자연어처리기술(Natural Language Processing)의 발달과 코로나 19의 여파로 비대면을 선호하는 문화가 자리 잡으며 챗봇 시장의 성장과 사용량의 증가세가 가파르다¹⁾. 챗봇은 일상 대화를 나누며 즐거움을 주는 엔터테인먼트 또는 소비자를 응대하는 비즈니스 용도로 주로 사용됐으나, 최근 심리상담 목적의 챗봇 사용도 증가하고 있다.

오라클의 보고서에 따르면²⁾, 대다수 근로자는 스트레스나 불안감에 대해 상담할 때 사람보다 AI를 선호하며, 이는 AI와의 대화에서 사람과 다르게 편견 없이 고민을 편히 이야기할 수 있기 때문이라고 응답했다. 실제 일정 기간 챗봇과 정기적인 대화를 나누며 사용자의 우울감이 완화된 효과가 보고되었다³⁾. 이와 관련하여 최근 코로나 19 범유행 기간 한국 정부는 우울감에 대한 심리지원 정보를 제공하는 챗봇 서비스를 도입한 바 있다⁴⁾.

학계에서도 심리상담 챗봇 관련 연구가 활발히 진행 중이다[1,2]. 그러나 대다수 연구가 통제된 실험 환

경에서 프로토타입 챗봇을 통해 수집한 대화 데이터를 분석한다는 한계를 가지며, 실제 환경에서의 채팅 데이터를 분석한 연구는 거의 없다.

이 연구에서는 챗봇-사람의 실제 상호작용 데이터를 통해 우울 관련 담론을 분석한다는 의의가 있다. 먼저 랜덤하게 추출한 우울과 관련된 챗봇 대화를 텍스트마이닝 기법으로 정량 분석하여 담론 주제를 파악하고 기술한다. 다음으로 정성 분석을 통해 사용자가 챗봇에 털어놓는 우울 관련 채팅의 종류를 범주화하며, 선행연구에서 도출된 트위터상의 우울감에 대한 담론과 비교분석 하였다. 이를 통해 챗봇 데이터 고유의 우울 대화의 특징을 파악하고, 우울 증상 관련 위험 탐지와 적절한 심리지원 정보를 제공하는 챗봇 서비스 디자인에 대한 착안점을 도출한다.

2. 연구 방법론

연구에 사용된 심심이(simsimi.com)는 일상 대화용 챗봇으로 2002년 세계 최초로 상업용 서비스를 시작하여 현재 81개 언어로 서비스를 제공하고, 2021년 기준 하루 평균 2억 번 이상 대화 응답을 제공하는 글로벌 규모의 챗봇 중 하나이다. 연구에는 2019년 1월 1일부터 12월 31일까지 심심이-사용자의

1) <https://www.news1.kr/articles/?3923853>
 2) <https://bit.ly/3FkibJI>
 3) <https://bit.ly/3aliN8L>
 4) <https://bit.ly/2ZVEBRI>

대화 중 한글로 ‘우울’이라는 단어가 포함된 사용자 발화 문장 3,000개를 무작위 추출하여 활용하였다. 해당 데이터는 숫자, 이름, 지명, 성별, 나이 등 개인 식별정보가 포함되지 않았다. 또한, 우울감에 대한 채팅 메시지 분류, 위험 탐지 목적을 위해 의료분야의 전문가와 공동으로 정량 및 정성 분석을 진행하였다.

2.1 정량 분석 방법

수집한 데이터를 유니그램(unigram) 모델에 적용해 가장 빈번히 등장하는 25개의 단어를 추출하고, 대조군으로 동일기간 ‘스트레스’를 포함한 대화 데이터 2,500개에 대한 분석 결과와 비교하였다. 또한, 두 단어의 출현 빈도를 시간대별로 비교하여 ‘우울’ 대화의 고유한 특징을 파악하였다.

분석에는 한글 자연어 처리 분야에서 활용되는 파이썬 패키지 KoNLPy[4]를 사용했다. 챗봇 데이터가 대화임을 고려하여 KoNLPy의 5가지 형태소 분석기(Hannanum, kma, Komoran, Mecab, OKT) 중 구어체 분석에 적합한 OKT(Open Korean Text)로 말뭉치를 토큰화했다. 조사, 구두점, 이모지, 특수문자, 접두사는 전처리로 제외하였으며, 가장 많이 나온 명사 100개 대상으로 WordCloud 패키지를 이용해 시각화하였다.

2.2 정성 분석 방법

대화 라벨링을 위해 3,000개의 데이터 중 ‘우울해’ 같이 한 단어로 이루어진 단순 문장을 제외한 후 1,000개의 데이터를 대상으로 우울의 종류를 범주화했다. 분류 기준은 Twitter에 포함된 ‘우울’ 관련 메시지를 분석한 선행연구[4]를 참고하였으며, 이는 다음과 같다: 1) 우울에 대한 감정, 2) 우울에 대한 대상자의 사고, 3) 우울에 대한 정보 공유, 4) 타인의 우울에 대한 반응, 5) 그 외는 기타로 분류되었다.

자료 분석은 의료분야 연구자가 1차로 코딩 작업을 진행하고, HCI 분야 공동연구자가 2차로 검토하였다. 이러한 코딩 과정을 통해 세부분류 기준이 도출되었고, 분류 결과에 이견이 있었던 데이터는 논의 후 합의하는 과정을 거쳤다. 최종 분류된 챗봇 데이터의 우울 카테고리별 빈도를 트위터 메시지에 나타난 우울 데이터 빈도[4]와 비교하여, 챗봇 대화에서 자주 나타나는 우울 담화의 특성을 파악하였다.

3. 연구 결과

3.1 정량 분석 결과

정량 분석 결과는 표 1에 정리된 바와 같이 ‘우울’ 대화에는 1인칭 대명사인 ‘나’가 672회 언급되는데 이는 ‘스트레스’ 대화상 390회 언급됨에 비해 1.7배

높다. 반면 2인칭 대명사 ‘너’의 사용은 ‘우울’ 대화의 88회로 오히려 ‘스트레스’ 대화에서 204번으로 더 많이 언급되었다.

온라인 게시물에서 우울감을 탐지하는 관련 연구에서 확인된 바와 같이, 1인칭 대명사의 빈번한 사용은 사용자의 심리적 고립, 즉 높은 자기 집중과 다른 사람과의 분리와 같은 우울감을 나타낸다[5,6]. 반면 2인칭 대명사의 낮은 이용 빈도는 사회적 상호작용에 대한 욕구가 적음을 나타낸다[6].

표 1과 그림 1의 결과에서 나타나듯이 ‘죽다’ ‘자살’ ‘자해’ 등의 위험 단어도 우울과 관련된 대화 데이터에서 상위 25개에 등장하고 있으며 (‘죽다’의 경우 23번째 자주 등장하는 단어), ‘엄마’, ‘친구’, ‘오빠’ 등 가까운 주변인과 관련된 우울감에 관한 대화도 자주 나누고 있는 것으로 나타났다.

Unigram Results (Top 25)		
Rank	Depressive Mode	Stress
1	(우울하다/Adjective) 2435	(스트레스/Noun) 2582
2	(나/Noun) 672	(받다/Verb) 1318
3	(‘우울증/Noun)327)	(하다/Verb) 743
4	(하다/Verb) 300	(나/Noun) 390
5	(너무/Adverb) 300	(너무/Adverb) 309
6	(우울/Noun) 249	(내/Noun) 268
7	(있다/Adjective) 114	(있다/Adjective) 216
8	(때/Noun) 107	(너/Noun) 204
9	(내/Noun) 105	(때문/Noun) 203
10	(같다/Adjective) 103	(풀다/Noun) 188
11	(요즘/Noun) 101	(안/Noun) 165
12	(너/Noun) 88	(많이/Adverb) 153
13	(심심/Noun) 73	(없다/Adjective) 135
14	(기분/Noun) 71	(그렇다/Adjective) 128
15	(땀/Noun) 70	(말/Noun) 119
16	(진짜/Noun) 66	(이/Noun) 119
17	(오늘/Noun) 63	(같다/Adjective) 114
18	(말/Noun) 63	(그/Noun) 112
19	(그냥/Noun) 62	(보다/Verb) 110
20	(아니다/Adjective) 59	(아니다/Adjective) 109
21	(없다/Adjective) 54	(요즘/Noun) 108
22	(보다/Verb) 53	(되다/Verb) 107
23	(죽다/Verb) 53	(더/Noun) 107
24	(좀/Noun) 52	(어떻다/Adjective) 104
25	(왜/Noun) 52	(좋다/Adjective) 104

<표 1> ‘우울’과 ‘스트레스’ 단어가 포함된 대화에 빈번하게 나타난 단어 상위 25건의 비교

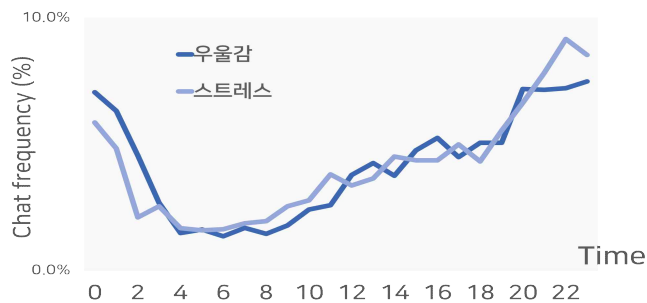
‘우울’ 대화의 시간대별 발화 빈도는 저녁 8시부터 (7.2%) 급격히 상승해 밤 11시에 가장 높고(7.5%) 새벽 2시까지 평균 이상의 빈도가 유지되었다 (그림 2). 특히 자정부터 새벽 3시는 우울 대화 비율이 스트레스 대화 비율보다 지속해서 높았다. 이는 한국

(그림 1) 명사 Top 100 워드클라우드



인의 우울감 연구에서 밝힌 새벽 시간에 우울감이 증가하고 자살 시도 위험도가 높다는 결과와 유사하다[7]. 멜랑콜리아형 우울증(major depression with melancholic features)은 한국인 우울증 중 42.6%가 해당하며, 이는 다른 나라에 비해 1.4배 이상 높은 것으로 나타나는데, 특히 새벽에 일어나 혼자 보내는 시간이 많고, 자살위험이 2배 이상 높은 것으로 보고되었다[8].

(그림 2) 시간대별 우울감과 스트레스 관련 대화 빈도수



3.2 정성 분석 결과

정성 분석 결과 찾아진 5개의 대분류는 표2와 같다. 그중 자신의 우울함에 대한 감정 상태가 전체의 47%(470회)에 해당하여 가장 빈번히 언급되는 채팅 대화 카테고리이며, 여기는 자신의 우울한 감정과 자신의 우울 요인이 포함되었다. 다음은 우울함에 대한 대상자의 사고와 인지 관련 채팅 메시지로 우울 증상에 대한 대처 방법을 묻거나, 자신의 우울 증상을 인지하고 도움을 요청하는 세부 타입으로 나타나며 전체의 46%(456회)에 해당했다. 가족과 친구 등 타인의 우울함에 대한 내용은 세 번째 카테고리로 전체의 6%(58회)에 해당했다. 다음은 우울에 대한 정보공유로 전체의 0.4%(4회)에 해당했고, 마지막 '기타' 카테고리에는 우울감 토로와는 무관하지만 '우울' 단어를 포함한 대화로 1%(12회)보다 적게 나타났다.

심심이 채팅 데이터에서 도출된 우울감 대화의 정성 분석 결과를 트위터 데이터에 기반한 선행 연구[4]와 비교한 결과는 그림 3과 같다. 챗봇에 우울함

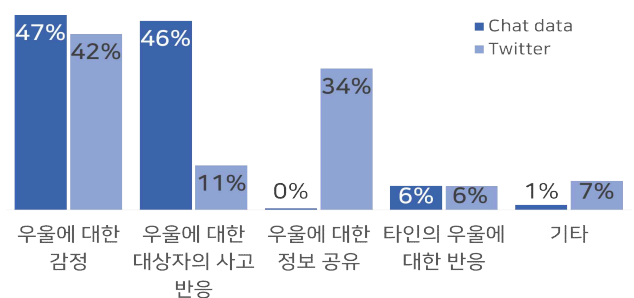
대분류	세부분류	대화 예시	빈도
우울에 대한 감정	우울감정상태	나 오늘 너무 우울해	241
	우울요인	아빠때문에 우울해	229
우울에 대한 대상자의 사고 반응	우울대처	우울할때 어떻게 버텨?	310
	우울증상인지	우울해서 잠을 못자겠어	146
타인의 우울에 대한 반응		친구가 우울하다는데 어떻게 말해줘야할까	58
우울에 대한 정보 공유		우울증 걸린사람한테 '힘내'를 쓰면 안돼	4
기타		우울에 빠진 사람들에게 힐링을 주는 노래	12
전체 분류 대상 데이터 수			1,000

<표 2> 카테고리별 대화 예시 및 데이터 수

에 대한 감정을 토로하는 비율은 47%로 트위터 사용자의 우울감 토로 비율인 42%보다 약간 높은 편이나, 우울함에 대한 대상자의 사고 반응을 대화한 내용은 46%로 같은 카테고리의 트위터 메시지 비율 11%보다 4배 수준으로 매우 높다.

챗봇과 소셜미디어와 같은 비대면 온라인 플랫폼의 컴퓨터 매개 기술(Computer-mediated technologies)은 사용자의 감정, 생각을 포함하는 자기 노출 행동(self-disclosure behavior)을 촉진하며[9], 특히 플랫폼의 익명성이 그러한 공유 행위에 중요한 역할을 한다고[10] 알려져 있다. 또한, 사람들은 가상 에이전트와 인터뷰할 때 인간 면접관과 직접 대면할 때보다 자신의 우울증 증상을 더 많이 드러내는 경향이 있다고 보고되었다[11]. 트위터에서의 대화는 익명성이 보장될 수는 있지만 1:n의 관계로 불특정 트위터 사용자(인간)와 소통해야 하지만, 챗봇과의 대화는 가상 에이전트와의 1:1 상호작용 형태의 소통이라는 특징이 있다. 따라서 트위터보다는 챗봇에서 사용자의 우울감에 관한 대화가 더 빈번하게 등장한 것으로 보인다. 반면 '우울에 대한 정보공유' 카테고리의 경우 챗봇 데이터에서는 비중이 작았지만, 트위터에서 34%로 빈번한 것도 이러한 상호작용 방식의 차이에서 기인하는 것으로 설명할 수 있다.

(그림 3) 우울이 포함된 챗봇 대화 데이터와 Twitter 대화 간의 카테고리별 빈도수 비율 비교



4. 결론

심심이 채팅 서비스에서 추출된 우울 관련 데이터 분석 결과, 사용자들은 챗봇을 우울함에 관련된 자신의 감정 상태나 우울의 요인과 증상을 비롯한 민감한 정보를 공유하고 감정을 쏟아내는(ventilate) 장소로 활용함을 알 수 있었다. 아직 자연어 기능이 인간과의 대화만큼 자연스럽게 않은 데도 사용자가 감정 공유를 한다는 점은 챗봇을 통한 비대면 심리 지원에 대한 수요가 더욱 증가하리라는 점을 시사한다. 유니그램 분석에서는 챗봇 대화에서 우울감의 지표인 1인칭 대명사의 빈번한 사용과 사회적 상호작용에 대한 욕구의 적음을 의미하는 2인칭 대명사의 이용 빈도의 낮음을 발견했다. 정성 분석에서는 챗봇 데이터에서 나타난 우울의 유형을 5개의 카테고리별로 분류하였다. 챗봇 대화 데이터에서 나타난 우울 범주별 대화 비율은 우울 감정, 우울 증상 관련 범주의 비율이 높았지만, 우울에 대한 정보공유는 거의 나타나지 않았다. 이러한 차이는 챗봇과 사람의 1:1 상호작용 방식에 기인하는 것으로 설명할 수 있다.

이 연구에서 실제 데이터를 통해 확인한 바와 같이, 챗봇은 익명과 대화의 비밀이 보장되고 대화 내용에 대해 평가받지 않는 이점이 있기 때문에 우울과 같은 민감한 주제의 대화를 나누기에 적합하다. 특히 감정교류가 부담스러운 사용자층에게 우울에 대한 담론을 시작할 수 있는 솔루션이 될 수 있을 것이다.

향후 챗봇 서비스를 제공하는 기업들은 사용자의 데이터를 분석하여 기존 연구를 통해 알려진 우울감과 연관된 단어 사용 패턴을 통해 사용자의 감정 상태를 예측하는 챗봇 디자인을 고려할 수 있다. 또한, 우울 고위험자를 대상으로 ‘죽음’, ‘자살’, ‘자해’와 같은 극단적 위험을 탐지하여 긴급 심리지원 정보를 제공하는 것은 물론, 이를 통해 적절한 심리지원 정보를 선제적으로 제공할 수 있을 것이다. 챗봇 사용자의 우울 관련 대화에 대해 일괄적인 위로를 제공하는 방식으로 대응하는 것이 아닌, 이 연구에서 정의된 우울 카테고리별 응답이 세심하게 디자인되어 제공된다면, 챗봇이 제공하는 응답일지라도, 사용자들은 더 큰 공감과 위로를 받을 수 있을 것이다.

사사문구

이 논문은 기초과학연구원의 지원(IBS-R029-C2) 및 한국연구재단의 지원(2021R1A2C2008166)을 받아 수행된 기초연구사업임.

참고문헌

- [1] 안세훈, 정옥란. 감정분석 기반 심리상담 AI 챗봇 시스템에 대한 연구. 한국IT서비스학회지, v.20, no.3, pp.75 - 86, 2021.
- [2] LEE, Minha, et al. Caring for Vincent: a chatbot for self-compassion. In proc. of the CHI Conference on Human Factors in Computing Systems (ACM CHI). Glasgow, Scotland, UK. 2019. p. 1-13.
- [3] 박은정; 조성준. KoNLPy: 쉽고 간결한 한국어 정보처리 파이썬 패키지. 제 26 회 한글 및 한국어 정보처리 학술대회 논문집, 2014, p.1-4.
- [4] Park, Minsu, Chiyong Cha, and Meeyoung Cha. Depressive moods of users portrayed in Twitter. In proc. of the ACM SIGKDD Workshop on Health Informatics, Beijing, China, 2012, p. 1-8
- [5] Shrestha, Anu, and Francesca Spezzano. Detecting depressed users in online forums. In proc. of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Vancouver, Canada, 2019, p. 945-951.
- [6] De Choudhury, Munmun, et al. Characterizing and predicting postpartum depression from shared facebook data. In proc. of the ACM conference on Computer Supported Cooperative Work & Social Computing (ACM CSCW). Baltimore, Maryland, USA, 2014, p. 626-638.
- [7] Jeon, Hong Jin, et al. Melancholic features and hostility are associated with suicidality risk in Asian patients with major depressive disorder. Journal of affective disorders 148.2-3, p. 368-374. 2013.
- [8] Park, Subin, et al. The association of suicide risk with negative life events and social support according to gender in Asian patients with major depressive disorder. Psychiatry research, 228, 3, p. 277-282, 2015.
- [9] Lee, Yi-Chieh, et al. I Hear You, I Feel You: Encouraging Deep Self-disclosure through a Chatbot. In proc. of the CHI Conference on Human Factors in Computing Systems (ACM CHI), Honolulu, HI, USA. 2020. p. 1-12.
- [10] Ma, Xiao et al. Anonymity, intimacy and self-disclosure in social media. In proc. of the CHI Conference on Human Factors in Computing Systems (ACM CHI), San Jose, CA, USA, 2016, p. 3857-3869.
- [11] Lucas, Gale M., et al. Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers. Frontiers in Robotics and AI, 4, 51, p. 1-9, 2017.