

[우울증 감지] 연구 주제 제안

제안 연구

- 1. 텍스트와 표정 인식을 통한 멀티 모달
- 2. 텍스트와 음성 데이터를 이용한 멀티 모달

	연구 1 (표정)	연구 2 (음성)
장점	새로운 형식의 챗봇 제안	선행 연구를 통한 효과 검증
단점	학습 데이터셋 부족	음성 인식 및 텍스트 변환을 위한 API 필요

챗봇뿐만 아니라 우울증 감지에 관련된 다양한 선행 연구를 조사하였습니다. 그 중 위 2가지 주제를 검토해 보았습니다.

1. 텍스트 + 표정 인식

특징

- 이전 프로젝트 경험 활용
- 선행 연구 부족
- 데이터셋 부족

과제

- 실효성에 대한 심리학적 해석이 필요
- 감정 데이터셋의 레이블 재분류
- CPU 환경에서도 동작할 수 있는 구조 고민

이전에 연구했던 “실시간 얼굴 인식을 활용한 대체 입력 기법”의 경험을 살려 적은 비용으로 얼굴 감지 및 감정 분석을 실행할 수 있을 것으로 기대하였습니다. 한 연구[1]에서 최대 87.6%의 정확도를 보이며 유의미한 결과를 보였습니다. 하지만 성능 향상을 위해 사용 가능한 데이터셋이 부족했습니다. 기존의 ‘감정 분류’ 이미지는 기쁨/슬픔/불안... 과 같은 형식인데 이것을 우울/비우울로 재분류하는 작업은 심리학에 대한 명확한 지식 없이는 어려움이 있습니다. 예를 들어, 웃는 이미지는 모두 ‘기쁨’으로 분류되지만 이것이 반드시 비우울 이미지라고 판단하는 것이 심리학적으로 옳은지와 같은 문제들이 남아있습니다. 또한 “AVEC-2016”과 같이 우울/비우울로 라벨링 되어 있는 경우 데이터 불균형 문제가 있습니다. AVEC-2016은 86장의 비우울 데이터와 21장의 우울 데이터로 구성되어 있습니다. 그뿐 아니라 한국인 데이터가 부재하기 때문에 테스트 환경이 아닌 실제 서비스 환경에서도 잘 분류할 수 있을지에 대한 문제가 있습니다. 만약 연구를 진행하게 된다면 실시간으로 이미지를 분류할 수 있도록 모델의 성능 문제도 해결해야할 것으로 보입니다.

추가로 이미지를 활용한 연구는 ‘SNS 이미지를 통한 우울증 감지’[2]가 있었습니다.

2. 텍스트 + 음성 정보

특징

- 선행 연구 존재
- 음성 데이터 처리를 위한 API 필요

과제

- 음성 인식 API
- 음성 데이터 전처리
- 데이터셋 구축

텍스트 기반의 챗봇에서 텍스트 정보뿐만 아니라 음성의 패턴 분석을 통한 정밀한 예측이 가능할 것으로 보입니다. 기존에 선행 연구들[3, 4, 5, 6]이 존재하며, 텍스트나 이미지와 달리 국외 데이터를 적용해도 국내 문제에 동일하게 활용할 수 있을 것이라고 생각됩니다. 다만, 문화적 차이에 따라 음성 패턴이 달라질 수 있는지에 대한 심리학 관점에서의 피드백이 필요해 보입니다. 추가로 마이크를 통해 입력된 음성을 텍스트로 변환하는 문제도 해결해야 할 것입니다.

참고 자료

- [1] Clinical Depression Detection in Adolescent by Face (2018)
- [2] 기계학습 및 소셜미디어 이미지 기반 우울증 분석 모델 개발 (2018)
- [3] 딥러닝을 이용한 음성신호 기반 우울증 진단 (2021)
- [4] 2D-CNN 기반 우울증 감지를 위한 음성데이터 전처리 (2021)
- [5] AI 스피커를 활용한 어텐션 메커니즘 기반 멀티모달 우울증 감지 시스템 (2021)
- [6] DCNN and DNN Based Multi-modal Depression Recognition (2017)
- [7] NHANES 데이터를 이용한 생활 패턴 기반의 우울증 예측 머신러닝 기법 (2021)
- [8] 우울증 치료를 위한 챗봇 개발의 방향성 고찰 (2020)
- [9] 부정적 감정 완화를 위한 BERGPT (2021)