

시스템 설계 및 구현 보고서

[AI 기반 독거노인 대화형 정서 지원 시스템]

[안녕 보리야!]

지도 교수

장주용

조원

[박수빈], 임채은, 유예성, 지민주

목차

1. 문제 정의 및 필요성
2. 시스템 설계
3. 데이터 및 모델 성능
4. 문제점 및 해결방안
5. 성과 요약 및 향후 과제
6. 참고 문헌 및 데이터 출처

1. 문제 정의 및 필요성

1.1. 문제 정의

우리 사회는 빠른 속도로 고령화가 진행되고 있으며 독거노인의 수 또한 급격히 증가하고 있다. 독거노인은 가족, 이웃, 지역사회와의 상호작용이 제한되어 심리적 고립과 외로움, 우울감 등의 정서적 어려움을 겪기 쉬운 환경에 처해 있다. 특히 일상 속에서 타인과의 정서적 교류가 부족한 상태가 지속되면, 심리적 안정감이 약화되고 정신건강 문제로 이어질 수 있다.

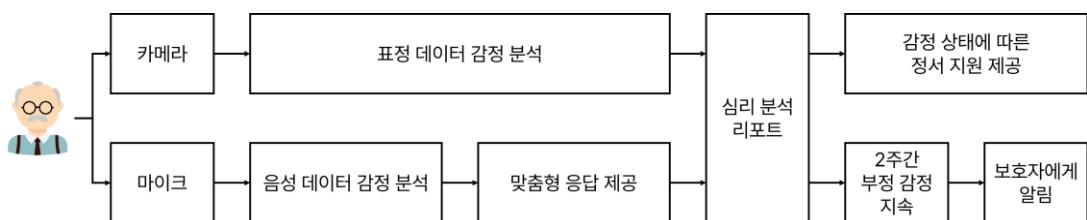
1.2. 필요성

정서적 고립은 단순히 외로운 감정에 그치지 않고, 고령자의 전반적인 삶의 질과 건강에 직결되는 중요한 문제이다. 따라서 고령자, 특히 독거노인의 삶의 질을 향상시키기 위해서는 정서적 상태를 능동적으로 파악하고, 그에 맞춘 상호작용이 가능한 시스템이 필수적이다.

이러한 배경에서, 단순한 알림 기능을 넘어 음성과 표정을 이용해 다양한 감정 신호를 실시간으로 분석하고, 상황에 맞는 정서적 피드백을 제공하는 AI 기반 정서 지원 시스템의 필요성이 대두된다. 감정의 인식과 대응이 가능한 시스템은 독거노인의 심리적 안정과 정서적 연결감을 강화하고, 나아가 심리 문제의 조기 발견 및 대응에도 기여할 수 있다.

2. 시스템 설계

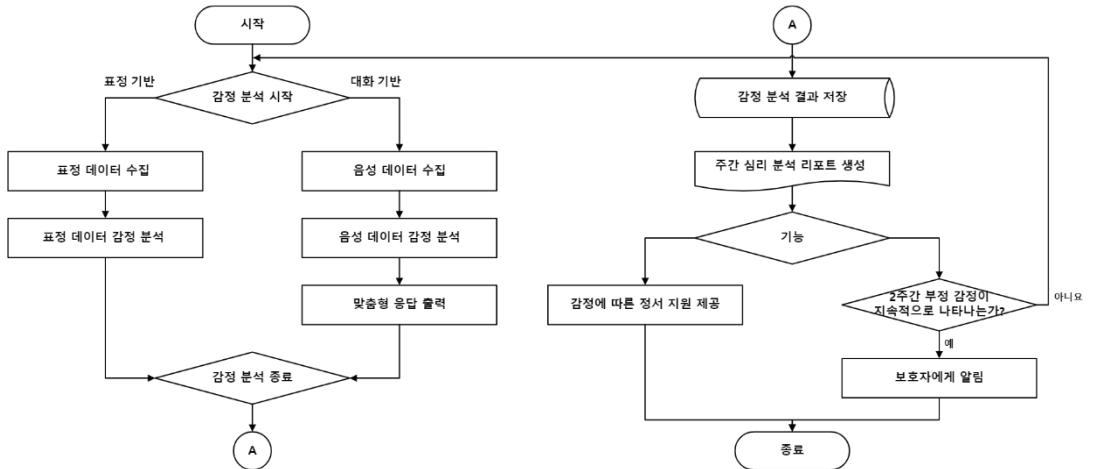
2.1. 구성도



구성도의 흐름은 다음과 같다. 먼저 사용자의 음성 및 영상 데이터는 시스템에 설치된 마이크와 카메라를 통해 실시간으로 수집된다. 수집된 데이터로 표정 감정 분석과 음성 감정 분석이 각각 수행된다. 음성 감정 분석 결과를 바탕으로는 사용자

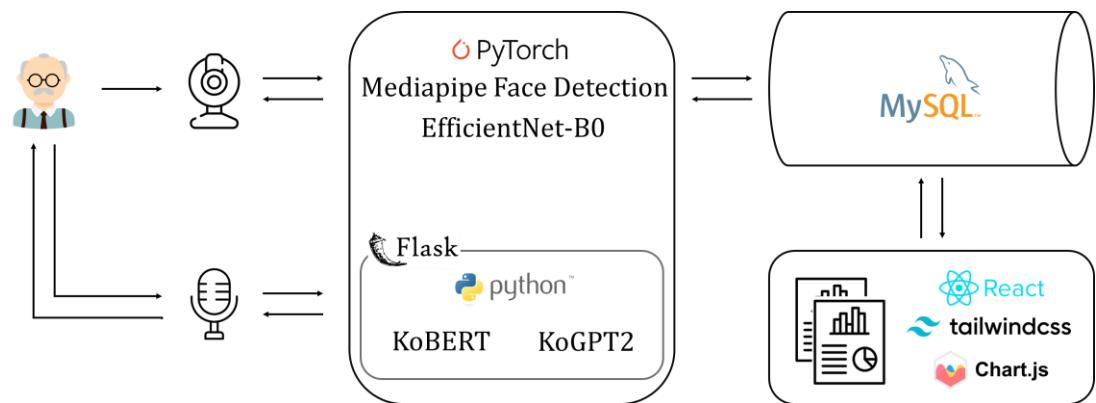
의 상태에 맞춘 맞춤형 응답이 제공된다. 이후, 음성과 표정에서 추출된 감정 정보를 통합 분석하여 사용자의 정서 상태를 종합적으로 해석하며, 이를 기반으로 심리 상태 리포트가 생성된다. 생성된 심리 분석 리포트를 토대로 감정 상태에 따른 정서적 지원이 이루어지며, 만약 2주간 부정 감정이 지속되는 경우 보호자에게 알림을 전송하는 구조로 시스템이 작동한다.

2.2. 순서도



순서도를 보면 먼저 카메라와 마이크를 통해 사용자의 표정 데이터와 음성 데이터를 실시간으로 수집한다. 이후 각각의 데이터를 기반으로 표정 감정 분석과 음성 감정 분석을 병렬적으로 수행한다. 분석이 완료되면 사용자와의 상호작용을 위해 감정 상태에 적합한 맞춤형 응답을 출력한다. 감정 분석 결과는 데이터베이스에 저장되며 주간 단위로 심리 분석 리포트를 생성하는 데 활용된다. 이를 통해 사용자의 감정 상태를 체계적으로 파악할 수 있으며 분석된 결과를 바탕으로 맞춤형 정서 지원 서비스를 제공한다. 또한, 부정 감정이 2주 이상 지속적으로 나타날 경우에는 보호자에게 자동으로 알림이 전달되어 적절한 대응이 가능하도록 한다.

2.3. 시스템 아키텍처 및 기술 스택

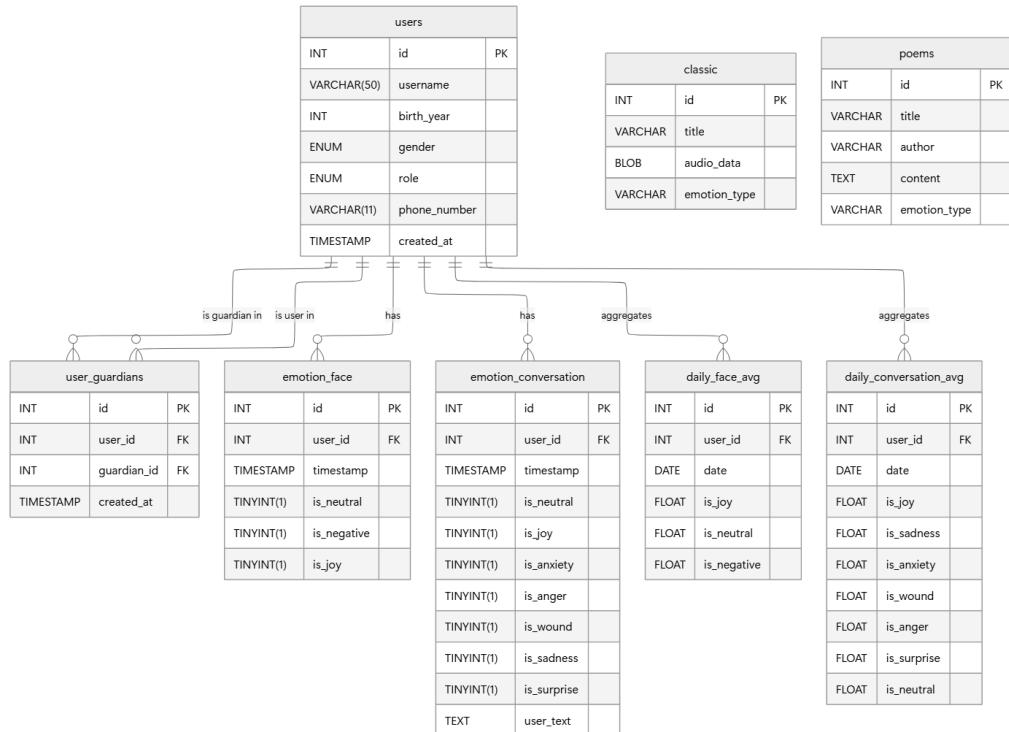


사용자의 음성과 표정을 입력으로 받아 정서 상태를 분석하고 이를 분석하여 시각적으로 제공하는 구조로 설계되었다. 마이크와 카메라를 통해 음성 및 영상 데이터를 실시간으로 수집하며 영상 데이터는 로컬 장치에서 Mediapipe Face Detection을 이용해 얼굴을 인식한 후, EfficientNet-B0 기반의 표정 분석 모델을 통해 감정 분류를 수행한다.

음성 데이터는 Python 기반의 Flask 서버에서 Whisper 모델로 텍스트로 변환된 후 KoBERT 모델을 통해 감정 분류, 이어서 KoGPT2 모델을 통해 정서적 응답을 생성한다. 각각의 분석 결과는 MySQL 데이터베이스에 저장되며 이는 심리 분석 리포트 생성과 감정 변화 추적에 활용된다.

저장된 데이터는 React 기반의 웹 인터페이스를 통해 시각화되며, 화면 구성은 Tailwind CSS를 사용하여 설계되었다. 분석된 감정 데이터는 Chart.js를 활용해 다양한 형태의 차트로 시각화되어 사용자와 보호자가 쉽게 이해할 수 있도록 구성되었다.

2.4. 데이터베이스 설계



2.5. 기능별 상세 설계 및 구현

2.5.1. 음성 인식 (Speech to Text)

- 목적

독거노인의 발화를 음성으로 입력받아 실시간으로 텍스트로 변환함으로써, 이후 감정 분석 및 대화 응답 생성의 입력값으로 활용하고자 하였다. 이를 통해 자연스러운 음성 기반 상호작용을 구현하고, 정서적 고립감 완화를 위한 대화형 시스템의 기반 기능을 마련하고자 하였다.

- 사용 기술

Wake Word 감지를 위한 PvPorcupine, 음성-텍스트 변환을 위한 OpenAI Whisper(medium) 모델, ngrok 기반 Flask server를 이용한 Google Colab GPU 연동

- 구현 내용

시스템은 항상 Wakeword(보리야)를 감지하는 대기 상태를 유지하다가 호출이 인식되면 “네 말씀하세요”라는 음성 응답과 함께 약 6초간 사용자 음성을 녹음하도록 설계되었다.

녹음된 음성은 ngrok 기반 Flask server를 통해 Google Colab으로 전송되며 Colab의 GPU 환경에서 Whisper(medium) 모델을 활용해 텍스트로 변환된다. 변환된 텍스트는 음성 기반 감정 분석의 입력값으로 사용된다.

Wakeword 감지에는 PvPorcupine을 사용하였으며 CPU 환경에서도 실시간 처리가 가능하다는 장점을 갖는다. 또한, Whisper 모델은 한국어 성능이 우수하고 오픈소스로 제공되어 무료로 사용할 수 있다는 점에서 선택되었다.

2.5.2. 감정 분석(KoBERT 3/5중 분류)

- 목적

음성으로부터 변환된 텍스트 데이터를 기반으로 사용자의 감정 상태를 분석하고 이를 통해 정서적 반응 및 응답을 생성하는 핵심 기능을 구현하자 하였다. 단순 감정 분류를 넘어 정확한 부정 감정을 식별하기 위해 KoBERT 기반 2중 감정 분류 구조를 도입하였다.

- 사용 기술

KoBERT 사전학습 모델, Huggingface Transformers, PyTorch, Flask API 서버, Google Colab GPU 환경

- 구현 내용

감정 분석은 2단계 분류 구조로 설계되었다.

1차 분류에서는 텍스트를 기쁨, 증립, 부정의 3가지 감정으로 분류하며 2차 분류는 1차 결과가 부정일 경우에만 수행되어 불안, 상처, 분노, 슬픔, 놀람의 5가지 세부 감정으로 정밀 분석한다.

감정 분석 요청은 Flask 기반 REST API 서버에서 처리되며, STT 모듈에서 전달받은 텍스트는 /predict/conversation API 엔드포인트로 전송

되어 KoBERT 모델로 실시간 감정 예측을 수행한 뒤 JSON 형식으로 응답된다.

2.5.3. 응답 생성(KoGPT2)

- 목적

감정 분석 결과와 사용자의 발화 내용을 바탕으로 정서에 공감하고 대화를 이어갈 수 있는 자연스러운 문장을 생성하여 사용자와의 상호작용을 구현하고자 하였다. 단순 반응형 응답을 넘어 맥락 유지와 감정 맞춤형 대화를 가능하게 하기 위해 KoGPT2 기반의 언어 생성 모델을 도입하였다.

- 사용 기술

KoGPT2 사전학습 모델, Huggingface Transformers, PyTorch, Flask API 서버, Google Colab 환경

- 구현 내용

응답 생성 모델은 사용자의 발화를 STT 및 감정 분석 모듈을 통해 처리한 후, "<usr> 사용자 발화 <sys> 응답" 형식의 입력 시퀀스를 생성하여 KoGPT2에 전달한다.

생성된 응답은 TTS(Text-to-Speech) 모듈을 통해 음성으로 출력되며, 사용자에게 말벗 기능을 제공한다.

응답 생성은 Flask 서버에서 처리되며, 클라이언트에서 감정 분석 결과와 함께 텍스트 입력을 전송하면 /generate/response API를 통해 KoGPT2로 전달되어 실시간 응답 문장을 생성한 후 JSON 형식으로 반환된다.

2.5.4. 표정 분석(HSEmotion)

- 목적

사용자의 얼굴 표정을 통해 비언어적 감정 신호를 실시간으로 인식하고 대화 감정 분석 결과와 통합하여 정서 상태를 보다 정밀하게 파악하는 것을 목표로 하였다. 실시간 처리와 적정 정확도, 경량화를 동시에 고려해 CPU 환경에서도 실행 가능한 표정 분석 시스템을 구축하고자 하였다.

- 사용 기술

Mediapipe Face Detection, HSEmotion 사전학습 모델 (enet_b0_8_va_mtl), PyTorch, OpenCV, 밝기 기반 조건 제어, Flask 연동

- 구현 내용

시스템은 웹캠 영상에서 Mediapipe를 활용해 얼굴 랜드마크와 경계 박스를 실시간으로 추출하며, 얼굴 영역만을 분리하여 전처리한 뒤, EfficientNet-B0 기반의 HSEmotion 모델에 입력한다.

표정은 기쁨, 중립, 부정의 3종 감정으로 분류되며 조명 조건에 따른 오류 방지를 위해 감정 분석 활성화 조건으로 평균 밝기 15 이상일 경우에만 모델이 작동하도록 설정하였다.

2.5.5. 심리 분석 리포트 및 경고 시스템

- 목적

사용자의 표정 및 대화 데이터를 종합적으로 분석하여 정서 상태를 직관적으로 확인할 수 있는 리포트를 제공하고, 특정 감정 패턴이 지속될 경우 보호자가 조기에 인지할 수 있도록 경고 알림 기능을 추가하고자 하였다. 이를 통해 정서적 고립과 심리적 위기를 사전에 감지하고 대응하는 체계를 마련하였다.

- 사용 기술
MySQL 기반 감정 로그 저장, Flask REST API, React.js 기반 리포트 UI, Recharts, WordCloud, 경고 트리거 로직(Python), 주차별 통계 집계 쿼리

- 구현 내용

시스템은 매일 감정 분석 결과를 MySQL에 저장하고, 주간 단위로 데이터를 집계하여 리포트를 생성한다. 리포트에는 다음과 같은 구성 요소가 포함된다:

- 어제의 감정 비율 (원 그래프)
- 최근 7일 간 감정 추이 (누적 막대 그래프)
- 주차별 감정 평균 및 단어 빈도 (누적 막대 그래프)
- 사용자의 사용 단어 빈도 (WordCloud)
- 감정 경고 여부 및 피드백 메시지

감정 경고 로직은 다음 두 조건을 모두 만족할 경우 작동한다:

- 표정 감정 조건
두 번째 주차(최근 7일)의 부정 감정 비율이 0.4 이상
- 대화 감정 조건
두 주차 동안 대화 감정의 최빈값이 동일하며, 해당 감정이 불안·상처·슬픔·분노·놀람 중 하나일 경우

이 기준은 표정은 최근 상태, 대화는 2주 간의 경향성을 반영하기 때문에, 두 요소가 모두 부정적으로 일치할 때를 고위험 신호로 판단한다. 경고 발생 시 보호자에게 SMS 또는 리포트를 통해 알림을 전송할 수 있도록 연동이 구성되어 있다.x

2.5.6. 가상 동반자 기능 (알림 및 정서 지원 콘텐츠 제공)

- 목적

독거노인의 일상 리듬 유지를 돋고, 감정 상태에 따른 정서적 피드백을 제공함으로써 심리적 고립을 완화하고 정서적 안정감을 지원하는 것을 목적으로 하였다. 단순한 알림을 넘어 사용자 감정에 따라 맞춤형 콘텐츠를 자동 제공하는 가상 동반자 기능을 구현하고자 하였다.

- 사용 기술

Python 스케줄러, gTTS(TTS 변환), 클래식 음악·시 구절 데이터베이스, 시간 기반 트리거 조건 설정

- 구현 내용

상 동반자 기능은 다음과 같은 일상 생활 지원 및 정서 개입 기능으로 구성된다:

- 식사 및 복약 알림

사전에 등록된 시간에 자동 음성 알림

- 일정 리마인더

일정 등록 후 해당 시간에 TTS 음성으로 알림 제공

- 하루 마무리 대화

저녁 시간에 하루 회고 질문과 감정 응답 기능 제공

- 정서 지원 콘텐츠 제공

어제의 감정 분석 결과 우세한 감정이 부정적 감정일 경우 아침/저녁 시간에 감정별 시 또는 클래식 음악을 자동 제공

콘텐츠 제공 대상 감정은 불안, 상처, 분노, 슬픔, 놀람이며 각각에 대해 적절한 시 문구 및 클래식 음악을 DB에 사전 저장하고 해당 감정이 일정 비율 이상이면 자동으로 출력되도록 구성하였다.

3. 데이터 및 모델 성능

3.1. KoBERT 기반 감정 3중 분류 모델

3.1.1. 목적

KoBERT 기반 감정 분류 모델은 사용자의 음성 입력을 텍스트로 변환한 후, 해당 발화를 기쁨, 부정, 중립 중 하나로 분류하기 위해 사용하였다. 이 모델은 전체 정서 지원 시스템에서 1차 감정 판단을 수행하는 모듈로 작동한다.

3.1.2. 데이터셋 구성

AI-HUB에서 감정이 분류된 한국어 대화 데이터셋을 수집하였다 전체 대화 데이터셋을 다음 기준과 같이 재분류하였다.

중립	중립
부정	슬픔, 공포, 혐오, 분노, 놀람, 불안, 당황
기쁨	기쁨, 행복

- 한국어 감정 정보가 포함된 연속적 대화 데이터셋

웹 크롤링 등의 방법으로 수집한 대화에 대한 감정이 레이블링 되어있다. '행복, 중립, 슬픔, 공포, 혐오, 분노, 놀람'의 7가지 감정으로 분류되어 있다.

중립	43,786
부정	10,784
기쁨	1,030
총	55,600

- 한국어 감정 정보가 포함된 단발성 대화 데이터셋

SNS 글 및 온라인 댓글에 대한 웹 크롤링을 실시하여 선정된 문장들이 포함되어 있다. '기쁨, 슬픔, 놀람, 분노, 공포, 혐오, 중립'의 7가지 감정으로 분류되어 있다.

중립	4,830
부정	27,727
기쁨	6,037
총	38,594

- 노년층 대상 감정 분류 모델 데이터셋

노년층이 포함된 대화 내용이 포함되어 있다. '불안, 기쁨, 분노, 슬픔, 상처, 당황'의 6가지 감정으로 분류되어 있다.

중립	0
부정	8,434
기쁨	1,716
총	10,150

- 공감형 대화

챗봇이 사용자의 감정에 적절하게 공감하는 대화를 하도록 설계된 데이터셋으로, 다양한 화자 관계를 포함하고 있다. '기쁨, 당황, 분노, 불안, 상처, 슬픔'의 6가지 감정으로 분류되어 있다. 전반적으로 '기쁨' 데이터가 부족하다고 판단하여, '기쁨'으로 분류된 문장만 포함하여 학습을 진행하였다.

중립	0
부정	0
기쁨	6,260
총	38,249

- 감성 대화 말뭉치

크라우드 소싱 수행으로 일반인 1,500명을 대상으로 세대 별 감성 대화 텍스트가 구축되어 있다. 아래 이미지와 같이 60가지 세부 감정이 포함되어 있다.

60가지 감정 분류						
기분	분노	슬픔	불안	상처	당황	기쁨
1	툴툴대는	실망한	두려운	질투하는	고립된	감사하는
2	좌절한	비통한	스트레스 받는	배신당한	남의 시선 의식하는	사랑하는
3	짜증나는	후회되는	취약한	고립된	외로운	편안한
4	방어적인	우울한	혼란스러운	충격 받은	열등감	만족스러운
5	악의적인	마비된	당혹스러운	불우한	죄책감	흥분되는
6	안달하는	염세적인	회의적인	희생된	부끄러운	느긋한
7	구역질 나는	눈물이 나는	걱정스러운	억울한	혐오스러운	안도하는
8	노여워하는	낙담한	조심스러운	괴로워하는	한심한	신이 난
9	성가신	환멸을 느끼는	초조한	버려진	혼란스러운	자신하는

중립	869
부정	44,633
기쁨	6,126
총	51,628

- 감정 분류를 위한 대화 음성 데이터셋

감성대화 어플리케이션을 통해 일정 기간 동안 사용자들이 어플리케이션과 자연스럽게 대화한 데이터가 포함되어 있다. 'happiness, angry, disgust, fear, neutral, sadness, surprise'의 7가지 감정으로 분류되어 있다.

중립	2,529
부정	35,797
기쁨	5,665
총	43,991

- 최종 학습 데이터셋 구성

균형 조정 전 데이터셋은 다음과 같이 부정과 기쁨의 문장이 중립에 비해 많아 다운샘플링을 진행하였다.

(1) 균형 조정 전

중립	52,014
부정	127,375
기쁨	58,823
총	186,198

(2) 균형 조정 후

중립	52,014
부정	52,014
기쁨	52,014
총	156,042

3.1.3. 학습 환경

모델	SKT KoBERT (transformers 기반 BERT 계열 한 국어 사전학습 모델)
Optimizer	AdamW
Learning Rate	2e-5
Batch Size	64
Epoch	10 (Early Stopping 적용)
Loss Function	CrossEntropyLoss
환경	Colab T4 GPU

Early Stopping 조건

- Validation Loss가 3 epoch 연속으로 개선되지 않는 경우
- Train Accuracy가 Validation Accuracy보다 10% 높은 경우

3.1.4. 결과 및 분석

```
100% | 157/157 [00:32<00:00,  4.64it/s]
**Final Test Accuracy: 0.8417**

**잘못 예측한 데이터 샘플 (실위 10개) **
원본 문장 실제 라벨 예측 라벨
0 그만 뒀다. 부정 중립
1 나 딱 한번만 더하고 가면 안돼? 한번만 더하면 저거 짐을 수 있을 거 같애! 중립 부정
2 조사하는 동안 패션쇼는 끝난다구요! 부정 중립
3 아... 돈.. 너무 미안한데 내가 급한 사정이 좀 생겨서 그런데 이따 올래요? 이따... 부정 중립
4 그거야 제가 묻고 싶은 말이죠... 중립 부정
5 얼마한테 다 이른다. 부정 중립
6 저한테 무슨 일이 생기면 제 자식들 보살펴줄 수 있는 분은 아버지뿐이라고 생각했어요. 부정 중립
7 너 군대에서 말뚝 한 번 박아 블래? 엄마 바꿔봐. 부정 중립
8 돈 많고, 잘 생기고, 거기다 성격까지 끝내주는 남자란다! 긍정 부정
9 앞으로도 이런식이면 아무도 니 말을 들어주지 않을 거야. 그때마다 한 명씩, 죽어나... 중립 부정

**총 잘못 예측한 데이터 개수: 1590**

**클래스별 잘못된 예측 개수**
실제 라벨
중립 830
부정 592
긍정 168
Name: count, dtype: int64
```

- 최종 테스트 정확도: 84.17%
- 전체 테스트 문장 수: 총 31,209개

- 클래스별 오분류:

중립	830
부정	592
기쁨	168
총	1,590

- 중립 발화가 가장 많은 오분류를 기록하였다. 이는 중립과 부정/긍정 간의 문맥적 경계가 모호한 경우가 많기 때문으로 판단된다. 예를 들어, 이미지의 “그만 뒀다”와 같은 발화는 부정적 의도로도 해석될 수 있어 중립 ↔ 부정 간 혼동이 잦았다. 기쁨 클래스의 오분류 수는 상대적으로 적으며, 이는 긍정 표현이 비교적 명확하게 구별되었기 때문으로 보인다.

3.2. KoBERT 기반 감정 5중 분류 모델

3.2.1. 목적

KoBERT 기반 부정 감정 세분화 모델은 1차 감정 분류에서 '부정'으로 분류된 발화를 보다 정밀하게 분석하기 위해 설계되었다. 단순한 기쁨, 부정 분류를 넘어 실제 정서 지원에 필요한 감정 상태를 파악하고 적절한 대응을 유도하는 것이 목표이다. 이 모델은 2단계 분류기로 사용하여 1차 분류 결과가 '부정'인 경우에만 활성화된다.

3.2.2. 데이터셋 구성

- 한국어 감정 정보가 포함된 연속적 대화 데이터셋(excel 파일)
 '행복, 중립, 슬픔, 공포, 혐오, 분노, 놀람'의 7가지 감정으로 분류되어 있다.
 '슬픔, 혐오, 분노'의 부정 감정을 학습에 사용하였다.

행복	중립	슬픔	공포	혐오	분노	놀람
1,030	43,786	1,972	98	220	3,628	4,866

- [노년층 대상 감성 분류 모델\(csv 파일\)](#)

'분노, 슬픔, 불안, 상처, 당황, 기쁨'의 6가지 감정으로 분류되어 있다. '분노, 슬픔, 불안, 상처, 당황'의 부정 감정을 학습에 사용하였다.

분노	슬픔	불안	상처	당황	기쁨
1,623	1,749	1,818	1,631	1,613	1,716

- [감성 대화 말뭉치\(json 파일\)](#)

'분노', '슬픔', '불안', '상처', '당황'의 6가지 감정으로 크게 분류되어 있다. '분노, 슬픔, 불안, 상처, 당황'의 감정을 학습에 사용하였다.

분노	슬픔	불안	상처	당황	기쁨
9,160	9,125	9,319	9,142	8,756	6,126

- [공감형 대화\(json 파일\)](#)

'기쁨, 당황, 분노, 불안, 상처, 슬픔'의 6가지 감정으로 분류되어 있다. '당황, 분노, 불안, 상처'의 부정 감정을 학습에 사용하였다.

기쁨	당황	분노	불안	상처	슬픔
18,783	17,445	21,944	16,494	199,34	12,926

- 최종 학습 데이터셋 구성

균형 조정 전 데이터셋은 '분노, 슬픔, 불안, 상처, 당황, 공포'의 6가지 부정 감정으로 구성되어 있었다. 그러나 '공포' 클래스의 문장 수가 현저히 적어 학습에 적합하지 않다고 판단되어 제외한 후 학습하였다.

(1) 균형 조정 전

분노	슬픔	불안	상처	당황	공포
22,576	20,253	24,020	26,981	24,078	98

(2) 균형 조정 후

분노	슬픔	불안	상처	당황
22,576	20,253	24,020	26,981	24,078

3.2.3. 학습 환경

모델	SKT KoBERT (transformers 기반 BERT 계열 한 국어 사전학습 모델)
Optimizer	AdamW
Learning Rate	5e-5
Batch Size	32
Epoch	10 (Early Stopping 적용)
Loss Function	CrossEntropyLoss
환경	Colab T4 GPU

Early Stopping 조건

- Validation Loss가 3 epoch 연속으로 개선되지 않는 경우
 - Train Accuracy가 Validation Accuracy보다 10% 높은 경우

3.2.4. 결과 및 분석

100%	716/716 [02:25<00:00, 4.94it/s]
Final Test Accuracy: 0.7826	
**잘못 예측한 데이터 샘플 (상위 10개) **	
0	병원 문장 실제 라벨 예측 라벨
1	병이 더 심해져 거동이 불편해지기 전에 돈을 더 벌어야 하는데 일자리를 구하기 어려... 분노 상처
2	정말 좋은 사람인 줄 알았는데. 분노 상처
3	회사에서 옆 자리 생활을 했던 동기가 퇴사를 해서 마음이 허전해. 당황 상처
4	여자친구가 보는 앞에서 날 치는데 그 자식이 진짜 역겹다는 생각이 들었어. 당황 분노
5	나는 이른바 회피형 인간이라고 부르는 타입인데 누군가와 연애하는 게 너무 어려워. 분노 불안
6	친구에게 괜한 말을 해서 난처한 표정을 짓는데 내가 먼저 사과를 해야 할 것 같아.... 불안 당황
7	아침에 눈을 뜨면 출근하기 싫다는 생각만 나. 불안 슬픔
8	나는 정신에 문제가 있나 봐. 친구들 사이에서 이러지도 저러지도 못하고 정말 힘들어. 불안 상처
9	재정 문제로 남편이 불만이 많다. 그동안 번 돈이 다 어디 갔냐고. 분노 불안
	나 회사 면접에서 탈락했어. 상처 슬픔

- 최종 테스트 정확도: 78.26%
- 전체 테스트 문장 수: 총 22,891개
- 클래스별 오분류:

당황	1,073
상처	1,032
슬픔	1,029
분노	971
불안	873
총	4,978

- 상처, 당황, 슬픔 클래스 혼동이 많았다. 부정 감정은 감정적으로 유사한 표현을 공유하는 경우가 많기 때문에 분류 경계가 모호하였을 것으로 예상된다. 예를 들어, 이미지의 "병이 더 심해져서 거동이 불편해지기 전에 돈을 더 벌어야 하는데..."는 상황적으로는 불안이지만, 심리 상태에 따라 상처나 분노, 슬픔으로도 해석 가능하다. 오분류 비율이 높았던 것은 감정 간 미묘한 차이를 텍스트 정보만으로 구별하기 어려운 구조적 한계를 반영한다.

3.3. HSEmotion 기반 표정 분석 모델

3.3.1. 목적

본 프로젝트에서는 텍스트 기반 감정 분석 이외에 사용자의 표정을 기반으로 실시간 감정 상태를 파악하기 위해 HSEmotion 모델을 사용하였다. 표정은 대화의 비언어적 단서로 작용하여 텍스트 기반 감정 분석을 보조하는 역할을 한다. 해당 모델은 실시간 카메라 영상에서 얼굴을 검출한 후, 검출된 얼굴 이미지를 기반으로 감정을 예측하는 데 활용한다.

3.3.2. 입력 및 처리 방식

입력은 실시간 카메라 영상에서 MediaPipe 를 이용해 얼굴 영역을 검출한 후, 해당 영역을 224×224 크기로 리사이즈하여 모델에 입력하는 방식으로 처리된다. 모델은 EfficientNet-B0 기반 사전학습 모델(enet_b0_8_va_mtl)을 사용하여 GPU 없이도 로컬 CPU 환경에서 실시간 추론이 가능하다.

HSEmotion 의 기본 출력 감정 라벨은 총 8 개(Happiness, Surprise, Anger, Fear, Sadness, Neutral, Disgust, Contempt)였으나, 모델을 로컬 CPU 에서 사용하기 때문에 실시간 시스템에서 보다 명확한 기준과 경량화된 처리 구조가 중요하다고 판단하였다. 이에 따라 'Anger', 'Fear', 'Sadness', 'Disgust', 'Contempt'를 부정 감정 범주로 통합하고 'Happiness'는 기쁨으로, 'Neutral'과 'Surprise'는 중립 감정 범주로 통합하여 3 개의 감정 범주로 재정의하였다. 이러한 분류 방식은 감정 간 모호한 경계를 줄이고, 분류 정확도를 상대적으 높일 수 있다. 또한 실시간 응답 구조에 적합하도록 출력 해석을 단순화할 수 있다는 장점이 있다.

3.3.3. 테스트 결과 및 분석

- 테스트 데이터셋

AI-HUB 의 '한국인 감정인식을 위한 복합 영상 데이터셋' 중에서 총 1,200 장의 이미지를 추출하여 테스트에 활용하였다.

- 모델

실시간 처리를 위해 GPU 없이 CPU 기반 환경에서 동작 가능한 모델(enet_b0_8_va_mtl)을 사용하였다. 이 모델은 EfficientNet-B0 기반으로, 로컬 환경에서도 실시간으로 얼굴 인식 및 감정 예측이 가능한 경량화 모델이다.

- 성능

1200 장으로 구성된 test 이미지를 사용하여 테스트하였다.

	f1-score		
	Natural (0)	Negative (1)	Positive (2)
enet_b0_8_va_mtl	83%	99%	86%

- F1-score는 Precision과 Recall의 조화 평균으로, 두 지표 간의 균형을 종합적으로 평가할 수 있는 지표이다. 테스트 결과 CPU 기반 모델임에도 전체적으로 고른 성능을 보였으며, 특히 부정 감정에 대한 민감도가 높게 나타났다. 이는 본 시스템이 로컬 CPU 환경에서도 실시간 처리가 가능하고, 독거노인 정서 지원 시스템에 적합함을 시사한다.

3.4. KoGPT2 기반 대화 응답 생성 모델

3.4.1. 목적

본 모델의 목적은 사용자(특히 노년층)의 발화를 음성으로 입력 받아, 그 문맥에 어울리는 맞춤형 자연어 응답을 자동으로 생성하는 모델을 구현하는 데 있다. KoGPT2는 언어 생성에 특화된 모델로서, 본 프로젝트에서는 사용자의 음성 인식 이후의 반응을 생성하는 역할을 담당하며, 정서 지원 시스템의 응답 생성 모듈로 활용되었다. 해당 모델은 노인 대상 대화, 감정 공감형 일상 대화 등을 포함한 데이터셋을 기반으로 학습되어, 실제 상황에 적절한 응답을 생성 할 수 있도록 설계되었다.

3.4.2. 데이터셋 구성 및 전처리

- 한국어 대화 데이터셋

응급상황 및 오피스 도메인 대화 데이터

응급 상황 데이터: 748개 대화세트, 단발성 4,975개 문장 포함오피스 데이터: 10,202개 대화 세트, 총 46,414개의 대화 쌍

상황별 연속 대화 세트	10,950세트
System – user 대화 쌍	58,008쌍
문장	120,991개

- 감성 대화 말뭉치

일반인 1,500명을 대상으로 한 음성 15,700문장 및 코퍼스 27만 문장,
 세대별 감성 대화 텍스트 구축을 통한 세대별 감성 대화 서비스 제공
 우울증 관련 언어 의미 구조화 및 대화 응답 시나리오 동반
 수집된 데이터
 대화 말뭉치 중 노년연령층에 해당하는 대화셋만 선별하여 사용

상황별 연속 대화 세트	16,640세트
System – user 대화 쌍	49,920쌍
문장	99,8401개

- Wellness 대화 데이터셋

한국어 웰니스 챗봇 학습용 데이터셋

KoGPT2 모델을 활용한 감정 기반 대화 모델 학습을 위해 구성된 데이터
 데이터 구성: 100개 대화 샘플 (감정 라벨(12개)에 대해 각 100개 샘플
 존재개) 총 1,200개

데이터 구성	Wellness 대화 데이터셋(chatbot_dataset_1.csv) 감정/상태 기반 질문-응답 쌍으로 구성된 한국어 상담 용 대화 데이터셋
데이터 출처	hoit1302/kogpt2-wellness-chatbot (GitHub)
라이선스	공개 데이터셋으로 비상업적 용도에 사용 가능 (MIT License)

- 공감형 대화 데이터셋

인간과 같이 공감할 수 있는 능력을 갖고 자연스러운 공감 대화 인터페이스를 위한 인공지능 학습용 대화 데이터

상황별 연속 대화 세트	31,821개
문장 수 별 발화 개수	472,821개
문장	1,213,872개

- 데이터셋 전처리

사용자 발화(usr)와 시스템 응답(sys)을 구분할 수 있도록 태그를 부착하여 입력/출력 구조를 통일하였다. 모든 문장을 시퀀스 기반 학습에 적합하도록 토크나이즈(tokenize) 처리하고, 문장 길이는 64 토큰 이하로 제한하였으며 특수문자, 공백, 중복 문장 등 불필요한 요소를 제거하여 텍스트 품질을 향상시켰다.

3.4.3. 모델 학습

- 모델 파라미터 및 학습 설정

Model	KoGPT2 model (kogpt2-base-v2 기반 사전학습 모델, Huggingface의 Transformers 라이브러리를 통해 로딩)
Optimizer	AdamW
Learning Rate	3e-5
Weight decay	1e-4
Batch size	16
Epoch	10(Early Stopping)
Loss Function	CrossEntropyLoss

- 학습 환경

Training Platform	Google Colab GPU T4
Processor	Intel 13th Gen i7-1360P @ 2.20GHz
Memory	16 GB RAM (15.7 GB usable)
Operating System	64-bit, x64-based processor
Frameworks	Python 3.10, TensorFlow 2.x, Huggingface Transformers

- 성능 평가 지표

accuracy	정답과 일치하는 예측 비율
perplexity	다음 문장(단어) 예측의 자연스러움을 평가하는 지표

3.4.4. 결과 및 분석

- 한국어 대화 데이터셋과 감성말뭉치 데이터셋(노인 발화)를 중심으로 학습을 수행한 경우, 총 223,878 개의 대화 쌍을 기반으로 초기 fine-tuning 정확도는 0.1808 을 기록하였다. 이후 학습을 거듭하면서 최대 0.7845 까지 정확도가 향상되었으며, 모델은 감정 표현이 포함된 발화에 대해 점차 자연스러운 응답을 생성하는 경향을 보였다.

- 반면, 웰니스 대화 데이터셋(Wellness chatbot dataset)을 병합하여 학습한 결과, 모델의 최종 정확도는 0.8191에 도달하였으며, 테스트 당시 생성된 응답은 사용자 발화의 정서적 맥락을 효과적으로 반영하고 있는 것으로 확인되었다.

```

100% ██████████ 1730/1731 [39:12<00:01, 1.31s/it]
[Epoch: 1] cost = 0.368994296
✨ 새로운 최고 검증 정확도 갱신! Best Valid Acc: 0.7650

100% ██████████ 1730/1731 [39:21<00:01, 1.28s/it]
[Epoch: 2] cost = 0.325419039
✨ 새로운 최고 검증 정확도 갱신! Best Valid Acc: 0.7850

100% ██████████ 1730/1731 [38:37<00:01, 1.28s/it]
[Epoch: 3] cost = 0.315241814
⌛ 1/3 epochs 동안 성능 개선 없음.

100% ██████████ 1730/1731 [39:21<00:01, 1.37s/it]
[Epoch: 4] cost = 0.306455493
✨ 새로운 최고 검증 정확도 갱신! Best Valid Acc: 0.7878

100% ██████████ 1730/1731 [38:08<00:01, 1.33s/it]
[Epoch: 5] cost = 0.299158573
✨ 새로운 최고 검증 정확도 갱신! Best Valid Acc: 0.8184

100% ██████████ 1730/1731 [38:24<00:01, 1.27s/it]
[Epoch: 6] cost = 0.294150323
⌛ 1/3 epochs 동안 성능 개선 없음.

```

- 정량적인 측면에서는 학습 반복에 따라 정확도 수치는 꾸준히 향상되었으나, 정성적인 관점에서 확인된 대화의 매끄러움이나 표현의 자연스러움은 충분히 확보되지 못하였다. 특히 일부 응답은 문법적으로는 타당하지만 맥락 상 어색하거나 감정적 공감이 결여된 문장으로 나타나는 경우가 있었다.
- 이러한 한계를 극복하기 위해 공감형 대화 데이터셋을 추가로 활용하였으며, 감정 주제에 기반한 질문-응답 쌍을 포함시켜 정서적 응답 표현의 품질을 개선하고자 하였다. 또한 기존의 정확도 중심 평가 지표 외에 perplexity 값을 함께 도입하여, 생성된 응답이 문법적, 어휘적 측면에서 자연스러운지를 종합적으로 판단할 수 있도록 평가 체계를 강화하였다.

[Epoch 10] 손실: 0.272030
 검증 정확도 (토큰 단위): 0.7730
 Perplexity (평균 손실 1.0113): 2.7491
 새로운 최고 검증 정확도 갱신! Best Valid Acc: 0.7730

- 최종적으로 파인튜닝된 KoGPT2 모델은 검증용 데이터셋을 기준으로 약 77%의 정확도를 기록하였다. 또한 언어 생성 모델의 예측 성능을 나타내는 척도인 perplexity는 2.75로 측정되었으며, 이는 모델이 다음 단어를 자연스럽게 예측할 수 있을 정도로 낮은 수준으로, 응답의 문법적 자연스러움과 일관성을 일정 수준 확보하였음을 의미한다.

- 대화 응답 결과 예시

```
[8] return_answer_by_chatbot('집에 아무도 없이 혼자 있어서 너무 무서워 .')  
☞ '혼자 있어 무서우시군요 .'  
  
[8] return_answer_by_chatbot('집에 아무도 없이 혼자 있어서 너무 무서워 . 나 놀아주면 안돼?')  
☞ '혼자 지내기 힘들어 하시는군요 . 지금 기분에서 벗어나기 위해 할 수 있는 일이 뭐가 있을까요?'
```

응답 생성 결과를 예시로 살펴보면, “오늘도 혼자 밥 먹었어”라는 발화에 대해 “혼자 드시면 마음이 허전하시죠.”라는 응답이 생성되었고, “어떤 사람 때문에 화나서 그냥 집 왔어.”라는 입력에 대해서는 “속이라도 버릴까 걱정되네요. 스트레스 해소를 위해 산책하시는 건 어떨까요?”와 같은 공감적 표현이 도출되었다. 이는 단순한 정보 제공을 넘어, 사용자 감정에 반응하는 정서적 연결감을 기반으로 한 대화 생성이 가능함을 시사한다.

- 노인 대상 데이터셋을 반영한 결과, 응답 표현이 보다 현실적이고 따뜻한 방향으로 변화하였다. 발화가 짧고 단순한 경우에도 내포된 감정을 적절히 해석하여 정서적으로 어울리는 응답을 생성하였다. 또한 다양한 감정이 섞인 발화에 대해서도, 학습된 문장 구조를 바탕으로 적절한 공감이나 위로 표현을 포함시킬 수 있었다.
- 기존의 사전학습 KoGPT2 모델만을 사용했을 경우에는 주제와 감정에 관계없는 일반적 응답이 출력되는 경향이 있었지만, 파인튜닝 이후에는 특정 주제(예: 외로움, 건강 불안, 가족 문제 등)에 대해 구체적이고 실제적인 응답이 생성되었다는 점에서 큰 개선이 있었다.

4. 문제점 및 해결방안

4.1. Whisper 모델 실행 환경 개선

STT(Speech-to-Text)를 수행하는 Whisper 기반 음성 인식 모델을 로컬(Windows PC) 환경에서 실행할 경우, CPU 기반의 연산 한계로 인해 처리 속도가 느리고, 시스템 성능에 따라 인식 정확도가 저하되는 문제가 발생하였다. 특히, 실시간 상호작용이 요구되는 시스템의 특성상, 응답 지연으로 인해 사용자 경험이 저해되는 결과를 초래하였다. 이를 해결하기 위해 Whisper 모델의 실행 환경을 로컬에서 Google Colab 기반의 원격 서버로 이전하였다. Colab의 GPU 자원을 활용함으로써 음성 인식 처리 속도가 크게 향상되었고, 인식 정확도 또한 안정적으로 유지하였다. 이와 같은 구조 개선을 통해 전체 시스템의 실시간성을 확보하고, 사용자 응답 지연 문제를 효과적으로 해결할 수 있었다.

4.2. KoBERT 병렬 추론 구조 개선

KoBERT 이중 분류 모델을 병렬로 로드해 추론하던 구조에서 GPU 자원 부족과 세션 충돌 문제가 발생했다. 이를 해결하기 위해 1차 분류가 '부정'일 때만 2차 모델을 실행하는 조건 분기 구조로 수정하고, `torch.no_grad()`를 적용해 불필요한 연산을 최소화하였다.

4.3. KoGPT2 기반 감정 응답 생성 모델의 문장 다양성 향상

KoGPT2를 기반으로 감정 응답 생성 모델을 개발하는 과정에서 가장 큰 시행착오는, 입력이 달라져도 유사한 문장이나 반복적인 표현이 출력되는 등 표현의 다양성과 맥락 반영이 부족했던 점이다. 특히 짧은 프롬프트 입력에 대해 무의미한 일반 응답이 반복되거나, 사용자 상황과 동떨어진 내용이 생성되는 문제가 있었다. 이러한 문제를 해결하기 위해, 다양한 상황을 반영한 학습 데이터를 충분히 확보하고 모델 학습을 진행하였다. 학습 후에는 다양한 입력을 기반으로 샘플 출력을 수집하고, 그 결과를 직접 확인하면서 응답 문장의 품질을 정성적, 정량적으로 평가하였다. 또한, 출력 결과의 문체나 어투가 지나치게 단조로워지지 않도록 지속적으로 확인·보완하였다. 이러한 반복적인 테스트와 조정 과정을 통해 모델이 보다 자연스럽고 실제 사용자 응답에 적합한 문장을 생성할 수 있도록 개선할 수 있었다.

4.4. 심리 분석 리포트 API 연동 오류 해결

심리 분석 리포트 구현 과정에서 API 연동에 문제가 발생하였다. 감정 평균값이 저장되기 전에 리포트 API가 호출되면서 감정 결과가 NULL 또는 0%로 출력되는 현상이 발생하였고 사용자 요청 날짜와 데이터베이스에서 조회된 날짜가 일치하지 않아 잘못된 정보가 리포트에 노출되었다. 이를 해결하기 위해 API 응답 순서를 조정하여 감정 분석이 완료된 이후에만 리포트가 생성되도록 하였으며 사용자 ID 와 날짜가 정확히 일치하는 데이터만 조회되도록 쿼리 구조를 수정하였다. 또한, 감정 데이터가 존재하지 않을 경우에도 오류 없이 안내 문구를 출력하도록 처리하여 사용자 혼란을 방지하였고, 감정 필드 형식을 프론트와 통일함으로써 데이터 해석 오류를 방지하였다.

5. 결과 및 향후 과제

5.1. 시연 동영상

시연자의 초상권을 위해 영상 링크를 기재하지 않고 캡처본으로 대체한다.

[시스템 사용 전, 사용자와 보호자 정보 등록]

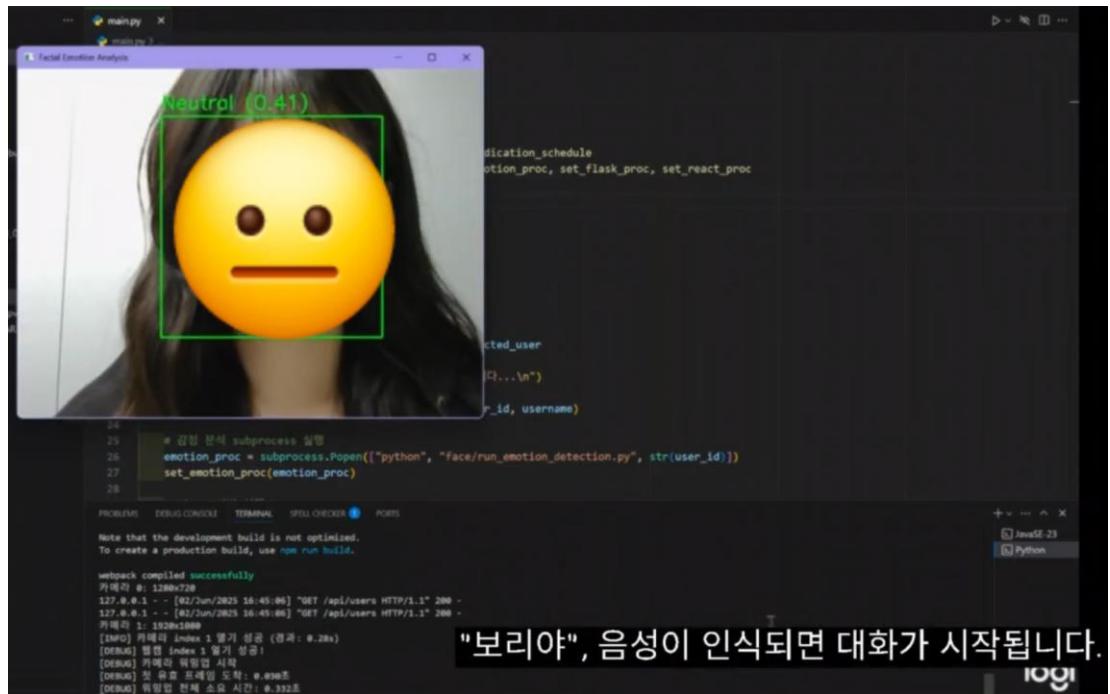


[복용 중인 약이 있다면 약 종류와 알림 시간을 설정]



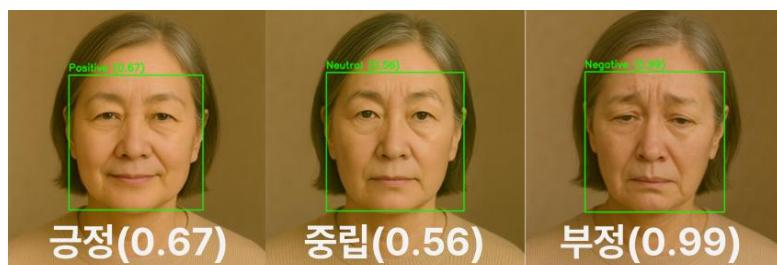
시스템 구동 시 심리 분석 리포트가 자동 생성

[“보리야” 음성이 인식되면 대화 시작]



실시간으로 카메라 속 표정을 인식해 감정을 분석

[모델 표정 분석 예시]



[보리와의 대화]

보리와의 대화 대표 사진

```
인식된 텍스트: 그들은 사람이 많은데 다 떠나갔어  
감정 분석 결과: 슬픔  
챗봇 응답: 후회가 크셨겠어요.  
[DEBUG] stop_intent: False  
[DB] 감정 분석 결과 저장 완료! (User ID: 1, 감정: 슬픔)
```

```
인식된 텍스트: 창문 열어보니까 밖에 꽃이 많이 피었더라  
감정 분석 결과: 기쁨  
챗봇 응답: 창밖 물결이 환해지니 기분이 조금 나아질 것 같아요.  
[DEBUG] stop_intent: False  
[DB] 감정 분석 결과 저장 완료! (User ID: 14, 감정: 기쁨)
```

세부 대화 내용은 아래에 기재

사용자 발화	챗봇 응답	감정 분석 결과
그리운 사람이 많은데 다 떠나갔어	후회가 크셨겠어요.	슬픔
옛날 사진 보니까 가슴이 먹먹해지고 눈물이 나더라고	우는 건 나쁜 일이 아니 예요. 나쁜 감정이 있다면 눈물로 흘려 보내는 것도 좋아요.	슬픔
아까 근처 공원 잠깐 돌고 왔어	산책하면서 무슨 일은 없 으셨나요?	중립
사람도 별로 없고 조용하더라고	조용한 장소는 생각 정리 해 도움이 되죠.	중립
요즘 혼자 있으면 자꾸 우울한 생 각이 들어	사람들과 함께 하는 것도 도움이 될 거예요.	슬픔
사람들 많은 곳에 가기 쉽지 않아	그래도 한 번 도전해보세 요. 제가 옆에 있어 드릴 게요.	불안
창문 열어보니까 밖에 꽃이 많이 피었더라	창밖 풍경이 환해지니 기 분이 조금 나아졌을 것 같아요.	기쁨

5.2. 심리 분석 리포트 결과

홍길동님의 심리 분석 리포트

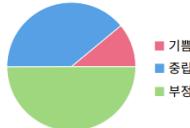
주차별 감정 변화와 부정 감정의 세부 분석

보호자: 임꺽정 (01000000000)

사용자 선택: 홍길동

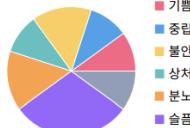
어제의 감정 비율 (2025-06-01)

표정



기쁨 (11.0%)
중립 (39.0%)
부정 (50.0%)

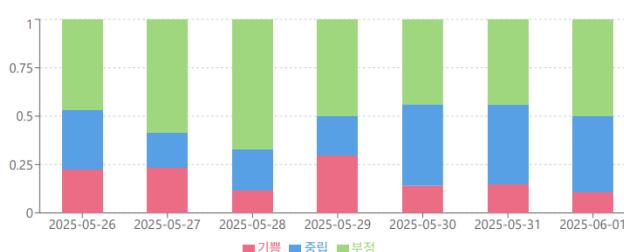
대화



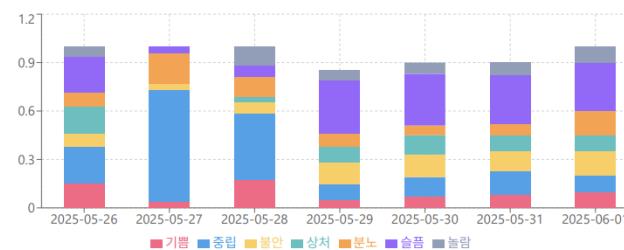
기쁨 (10.0%)
중립 (10.0%)
불안 (15.0%)
상처 (10.0%)
분노 (15.0%)
슬픔 (30.0%)
놀람 (10.0%)

어제 감정 분석 결과, 표정에서는 부정 감정이 높게 나타났고, 대화에서는 '슬픔' 감정이 우세하여 추천 콘텐츠로 시 제공하였습니다.

최근 7일 표정 감정 추이



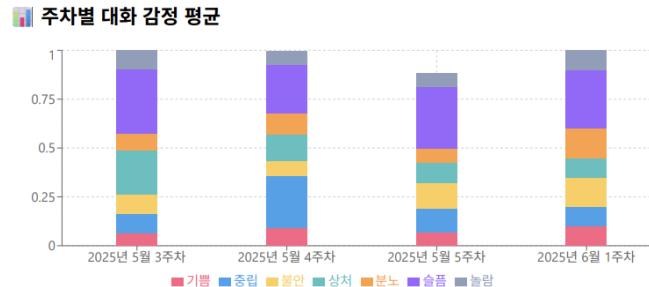
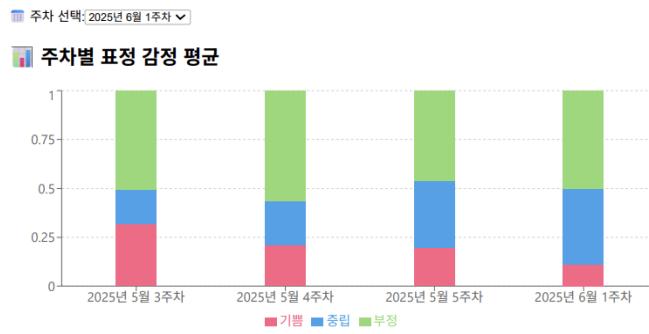
최근 7일 대화 감정 추이



표정과 대화 모두에서 부정 감정이 우세하게 나타났습니다. 이는 스트레스, 피로, 정서적 불균형의 신호일 수 있으며, 전반적인 감정 상태의 악화를

의미합니다. 초기 개입 및 심리적 지지가 필요한 상황으로 판단됩니다. 또한 부정 감정 중 '슬픔' 감정이 평균 22.6%로 가장 두드러지게

나타났습니다. 해당 감정에 대한 정서적 개입이 권장됩니다.



⚠ 최근 2주간 표정에서 부정 감정이 지속적으로 높게 나타났고, 대화에서는 '슬픔' 감정이 평균 30.8%로 우세하게 관찰되었습니다. 무기력감으로 이어지지 않도록 관심이 필요합니다.

▶ 주차별 자주 사용하는 단어



5.3. 결론 및 시사점

대화와 표정을 기반으로 사용자의 감정 상태를 분석하고 정서적 피드백과 일상 지원 기능을 제공하는 AI 기반 정서 지원 시스템을 구현하였다. 기술적으로는 Wakeword 기반 음성 입력 구조를 구성하고 Whisper와 KoBERT, KoGPT2, HSEmotion 등의 딥러닝 모델을 모듈화하여 통합 시스템을 설계하였다. 또한, 분석 결과를 시각적으로 표현하는 심리 분석 리포트와, 감정 상태에 따른 콘텐츠 재

생 기능까지 포함함으로써, 단순 기능 구현을 넘어 사용자 경험을 고려한 시스템을 구현할 수 있었다.

이번 프로젝트를 통해 감정이라는 추상적인 개념을 분석 가능한 형태로 구현하는 과정에서 기술적 구현뿐만 아니라 표현 방식, 윤리성, 사용자 피드백 구조 등 다양한 측면을 고려해야 함을 배울 수 있었다.

또한, 시스템이 실제 환경에서 작동하기 위해서는 응답 속도, UI 직관성, 연산 효율 등도 중요하다는 것을 인식하게 되었으며 이를 개선하기 위한 서버 구조 변경, 밝기 기반 조건 제어, 인터페이스 단순화 등 다양한 시도를 해왔다. 이러한 경험은 단순히 작동하는 시스템을 넘어서 어떻게 하면 오래 쓸 수 있고 사람에게 자연스럽게 다가갈 수 있는가까지 고민하게 되었다.

무엇보다 이 프로젝트는 단순한 기능 구현을 넘어서, 실제 사회 문제에 대해 기술로 접근해본 경험이라는 점에서 큰 의미가 있었다. 향후 사용자 테스트 확대, 감정 분석 체계 고도화, 외부 복지 연계 서비스 통합 등을 통해 완성도를 높인다면 실질적인 정서 지원 시스템으로 확장 가능할 것으로 기대된다.

5.4. 향후 과제

5.4.1. 감정 분석의 정밀도 향상

- 표정 분석은 3종 분류, 대화 분석은 7종 분류 체계를 사용하고 있어 감정 간 해석 기준이 불일치하는 문제가 존재한다. 향후에는 표정 분석의 감정 분류를 7종 분류 체계화 통합하는 것이 필요하다.
- 감정 경계까 모호한 상황을 정량적으로 판단할 수 있도록 신뢰도 출력값 기반 감정 보정 알고리즘 도입이 필요하다.

5.4.2. 대화 기반 응답의 자연스러움 및 문맥 유지 강화

KoGPT2 응답 생성 결과는 일정 부분 반복적이거나 맥락에서 벗어나는 경향이 있다. 이를 개선하기 위해

- Transformer 기반 장기 문맥 모델 도입 (예: KoAlpaca, LLaMA 기반 모델)

- 대화 히스토리 기억 기능과 프롬프트 튜닝 기법 적용 필요

사용자 성향에 따른 맞춤형 응답을 위해 사용자별 감정 히스토리를 반영한 응답 전략 수립도 고려해야 한다.

5.4.3. UI/UX 개선 및 일정 등록 기능 고도화

현재 일정 등록 방식은 단순한 시간 입력 기반으로, 사용자의 이해도가 낮을 경우 입력 오류가 발생할 수 있다. 이를 해결하기 위해:

- 캘린더 UI 도입, 음성 명령 기반 일정 입력 기능
- 일정 확인·편집 기능 통합

5.4.4. 지속 가능한 운영 환경 확보

현재 Whisper 모델은 Colab GPU 서버에서 운영되며, 장기 운영에는 한계가 있으며 경량화된 STT 모델의 로컬 전환이 필요하다.

또한, 감정 분석 결과를 보관·조회하는 기능은 확장성 있는 백엔드 구조로 개편이 필요하다.

6. 참고 문헌 및 데이터 출처

- 질병관리청. “우울감의 정의 및 원인.” 국가건강정보포털.
https://health.kdca.go.kr/healthinfo/biz/health/gnrlzHealthInfo/gnrlzHealthInfo/gnrlzHealthInfoView.do?cntnts_sn=5294
 - 국립정신건강센터. “주요 우울장애의 DSM-5 진단 기준.” 국가트라우마센터.
https://nct.go.kr/distMental/rating/rating01_2_2.do
 - 통계청. “독거노인 자료.” e-지역지표 KOSIS.
https://kosis.kr/visual/eRegionJipyo/themaJipyo/eRegionJipyoThemaJipyoView.do?AreaChoiceCombo=A_01_02&chartGbn=DTypeChart&graphTypeGbn=THEMA&jipyo1PrdDe=04bcd4bcd&jipyold=5440_6763&menuThemald=A_01_02_02®ionChkVal=00%40&themaGbn=subjectJipyo&themaGbnMenu=subjectJipyo&themald=A_01_02
 - 통계지표포털. “지표 정보.” 지표나라.
<https://www.index.go.kr/unity/potal/indicator/IndexInfo.do?idxCd=4233>
 - Ekman, P. “An argument for basic emotions.” *Cognition and Emotion*, vol.6, no.3–4, pp.169–200, 1992.
 - 노현서. “필사를 활용한 시치료 프로그램 효과 연구.” 석사학위논문, 한국외국어대학교 교육대학원, 2020. 서울. <https://www.riss.kr/link?id=T15645660>
 - 최혜경. “정서적 치유를 위한 시의 실용적 활용.” *문학치료연구*, vol.22, pp.217–268, 2012. <https://kiss.kstudy.com/Detail/Ar?key=3512474>
 - 문형주. “메타 정보 기반 이미지 얼굴 표정 인식 개선.” 석사학위논문, 조선대학교 산업기술창업대학원 소프트웨어융합공학과, 2024. 조선대학교 학술정보원.
 - 이원민, 온병원. “감성 및 감정 단어 마스킹 기반 BERT와 GPT 파이프라인 방식을 통한 감정 문장 생성.” *한국정보기술학회논문지*, vol.19, no.9, pp.29–40, 2021.
 - GitHub – SKT-AI. “KoGPT2 모델.” <https://github.com/SKT-AI/KoGPT2>
 - GitHub – SKTBrain. “KoBERT 모델.” <https://github.com/SKTBrain/KoBERT>
 - GitHub – google-ai-edge. “Mediapipe.” <https://github.com/google-ai-edge/mediapipe>
 - Bazarevsky, V. et al. “BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs,” arXiv preprint, arXiv:1907.05047, 2019.

<https://arxiv.org/abs/1907.05047>

- GitHub – serengil. "DeepFace." <https://github.com/serengil/deepface>
- GitHub – oarriaga. "Real-time Convolutional Neural Networks for Emotion and Gender Classification."
https://github.com/oarriaga/face_classification/tree/master
- GitHub – sb-ai-lab. "HSEmotion: Efficient High-Speed Emotion Recognition."
<https://github.com/sb-ai-lab/EmotiEffLib>
- Sabanovic, S. et al. "HSEmotion: Efficient High-Speed Emotion Recognition," arXiv preprint, arXiv:2203.13436, 2022. <https://arxiv.org/abs/2203.13436>
- AI Hub. "웰니스 대화 스크립트 데이터."
<https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=267>
- GitHub – e9t. "네이버 영화 리뷰 기반 긍/부정 감정 분석 데이터셋."
<https://github.com/e9t/nsmc>
- AI Hub. "한국어 감정 정보가 포함된 단발성 대화 데이터셋."
<https://www.aihub.or.kr/aihubdata/data/view.do?dataSetSn=270>
- AI Hub. "노년층 대상 감성 분류 모델."
<https://www.aihub.or.kr/aihubdata/data/view.do?dataSetSn=453>
- AI Hub. "한국어 감정 정보가 포함된 연속적 대화 데이터셋 (.xlsx)." <https://www.aihub.or.kr/aihubdata/data/view.do?dataSetSn=271>
- AI Hub. "한국어 대화 데이터셋."
<https://www.aihub.or.kr/aihubdata/data/view.do?dataSetSn=272>
- AI Hub. "감성 대화 말뭉치."
<https://www.aihub.or.kr/aihubdata/data/view.do?dataSetSn=86>
- AI Hub. "한국인 감정인식을 위한 복합 영상."
<https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=82>
- AI Hub. "한국인 얼굴 합성을 위한 발화 모습 이미지."
<https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=71427>
- AI Hub. "공감형 대화."
<https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=71305>
- 연아롬. "AlphaPose 모델 개요 및 적용 사례 분석."
<https://yeonarom.tistory.com/19>

- Google AI Edge. "Mediapipe."
<https://ai.google.dev/edge/mediapipe/solutions/guide?hl=ko>
- Velog – lighthouse97. "ResNet의 이해."
<https://velog.io/@lighthouse97/ResNet%EC%9D%98-%EC%9D%B4%ED%95%B4>
- 노현서. "필사를 활용한 시치료 프로그램 효과 연구." 国外學術論文 韓國外國語
 大學校 教育大學院, 2020. 서울
- <https://www.riss.kr/link?id=T15645660>
- 최혜경. (2012). 정서적 치유를 위한 시의 실용적 활용. 문학치료연구, 22, 217-268.
- <https://kiss.kstudy.com/Detail/Ar?key=3512474>
- 나태주. (2020). 꽃을 보듯 너를 본다. 지혜
- 나태주. (2017). 가장 예쁜 생각을 너에게 주고 싶다. 알에이치코리아
- 나태주. (2022). 너의 초록으로, 다시. 더블북
- 나태주. (2023). 끝까지 남겨두는 그 마음. 북로그컴퍼니
- 나태주. (2022). 자기돌봄의 시. 북로그컴퍼니
- 나태주. (2019). 마음이 살짝 기운다. 알에이치코리아
- 나태주. (2022). 너무 잘하려고 애쓰지 마라. 열림원
- 박광수. (2014). 문득 사람이 그리운 날엔 시를 읽는다. 걷는나무
- 진은영. (2022). 나는 오래된 거리처럼 너를 사랑하고. 문학과지성사
- kogpt2-wellness-chatbot
- <https://github.com/hoit1302/kogpt2-wellness-chatbot/tree/master>
- 서경현, 박정양 and 이재구. (2007). 음악이 스트레스로부터의 심혈관계 반응과 정서 반응 회복에 미치는 영향. 한국심리학회지: 건강, 12(2), 395-409.
- <https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artId=ART001058999>
- 문형주, (2024), 메타 정보 기반 이미지 얼굴 표정 인식 개선 [석사학위 논문, 조선대학교 산업기술창업대학원 소프트웨어융합공학과], 조선대학교 학술정보원
- Ekman, P. (1992). An argument for basic emotions. Cognition and Emotion, 6(3–4), 169–200
- 이원민, 온병원. (2021). 감성 및 감정 단어 마스킹 기반 BERT와 GPT 파이프라인 방식을 통한 감정 문장 생성. 한국정보기술학회논문지, 19(9), 29–40
- 쇼팽. 녹턴 No.2 in E-flat Major, Op. 9 No. 2

- 프렐류드.D♭장조 Op. 28 No. 15 ("빗방울")
- 드뷔시.달빛 (Clair de Lune)
- 에릭 사티.짐노페디 No.1 (Gymnopédie No.1)
- 라흐마니노프.보칼리제 (Vocalise)
- 바흐.아리오소 (Arioso from Cantata BWV 156)
- 헨델.라르고 (Largo from Xerxes)
- 차이콥스키.현을 위한 세레나데, 2악장 (Waltz)
- 모차르트.피아노 소나타 K.545 2악장
- 멘덴스존.'6개의 노래', 작품번호 34 중 두 번째 곡 노래의 날개 위에
- spikeekips. (2016). 한국어 불용어 리스트. GitHub Gist.
- <https://gist.github.com/spikeekips/40eea22ef4a89f629abd87eed535ac6a>