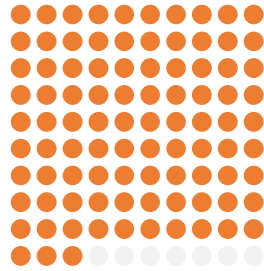




오픈소스 개발 프로젝트 기술 보고서

지친 현대인과 독거노인 등  
고민이 있는 모든 사람들을 위한  
**AI 심리상담 챗봇 - 메아리**

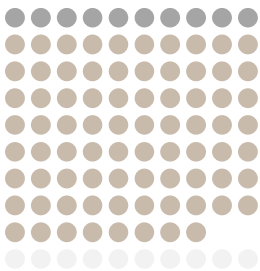
오픈소스 개발 프로젝트 2021042005 조경빈



186만

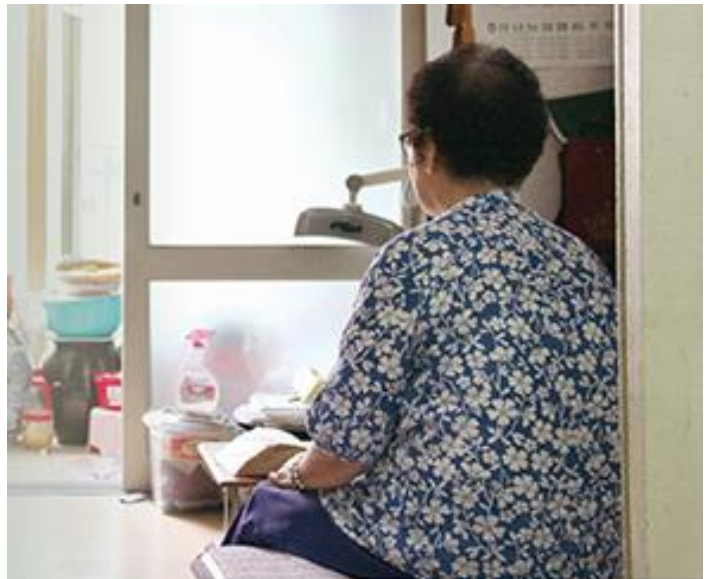
우울증 인구

지친 현대인은 물론이고 우울증 인구와 독거노인 인구가 지속적으로 증가함.



176만

독거노인

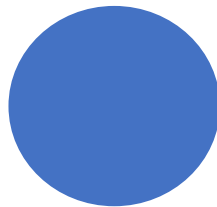
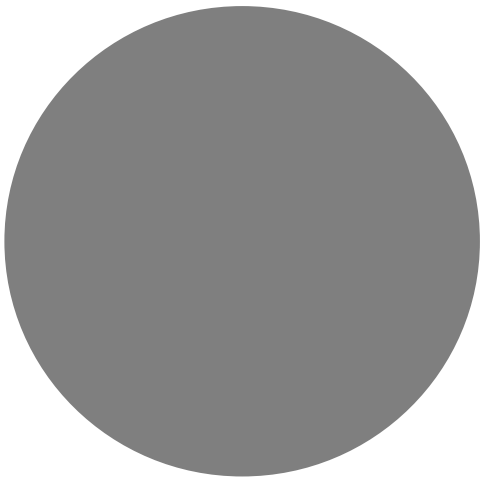




# 누구든지 이용 가능함 심리상담 챗봇 필요함

- 개발목표 -

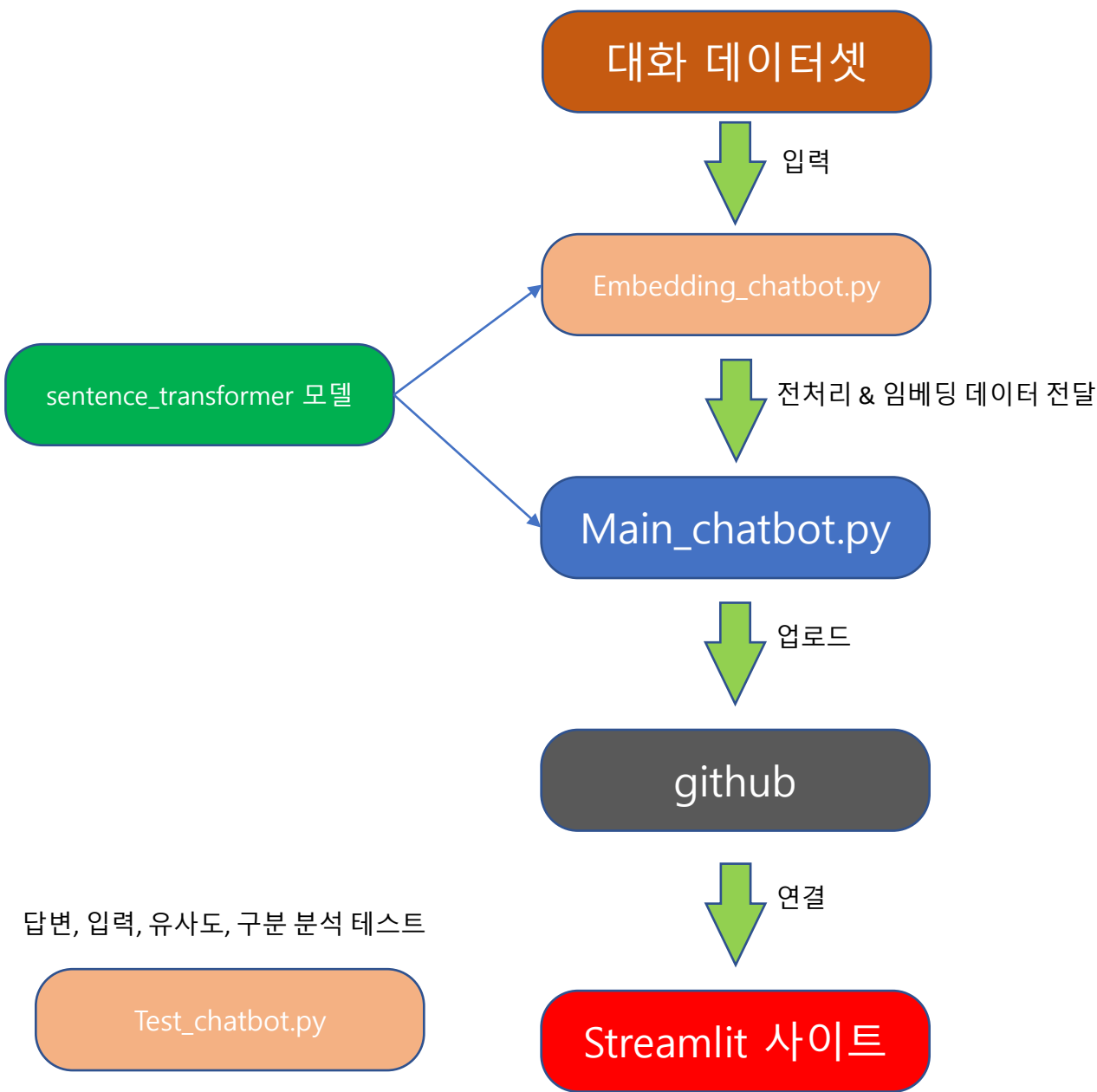
대화 형식으로 진행되며 사용자의  
말을 이해하고 그에 맞는 위로나  
해결책을 제시하는 인공지능



학습 모델 : Sentence BERT  
데이터셋 : 웰니스 대화 스크립트  
+ 그 외 대화 데이터셋  
라이브러리 : streamlit, pandas,  
sentence-transformers, scikit-learn

# 프로젝트 설계 - 심리상담 AI

## SW구조



# 개발 일정 & 실적

## 데이터 수집 및 정리

- 약 1.3만개의 데이터를 수집 후 정리

9월

## 데이터 전처리

- 누락데이터 및 오류 데이터 확인 후 삭제

10월

## 챗봇 모델 제작

- 사용자의 입력을 학습되어 있는 모델에 넣어 임베딩한 후 데이터와  $\cos$ 값 비교를 통해 적절한 답을 찾는다.

11월

## 사이트 연계

- streamlit을 통해 챗봇 ui를 제작
- 챗봇 모델을 통해 들어오는 입력에 따라 답변 출력

12월

# 프로젝트 결과

## 1. 모델 수정 및 정확도 향상을 위한 노력

- “너무 피곤해”, “기운이 없어”, “힘들어” 3가지로 정확도 테스트
  - 모델과 config값을 조정하며 평균 9.6의 정확도가 나오는 모델을 사용
- 정확도와 속도의 상승을 위한 데이터 수정
  - 13000개의 데이터 중 중요 데이터 약1000개 추출 → 데이터 값 수정  
→ 정확도 상승 (참고 1-1)
  - 나머지 약 12000개의 데이터 중 애매하거나 이상한 답변 데이터 삭제  
→ 속도 상승
- 모델 준비 완료

## 2. 전처리

- 1.3만개의 대화 데이터 중 누락 데이터를 소거

## 3. 임베딩

- 더미 데이터 생성
- 학습 된 모델을 통해 임베딩한 값을 데이터에 추가

## 4. 메인 챗봇 시스템

- 입력 값과 학습 데이터를 비교하여 답변 추출한다.
- 유사도 0.8 이상의 답변만 사용한다. 이를 적정 유사도라고 정의한다.
- 중요 데이터를 먼저 비교하며 적정 유사도를 못 넘겼을 경우 일반 데이터로 답변 추출. 마찬가지로 적정 유사도를 못 넘겼을 경우 재입력을 요청한다.

참고 1-1) chatbot\_dataset\_well\_f2.csv



메인 헤더  
& 서브 헤더

저장 되어 있는 입력  
값과 답변을 차례로  
출력

## 심리상담 챗봇 - 메아리

made by 조경빈



안녕하세요 당신의 심리상담가 메아리입니다. 편하게 뭐든지  
말씀해보세요.

너무 피곤하고 힘들어



피곤하면 정말 아무 것도 하기 어렵죠. 잠깐 휴식을 취하는 건  
어떨까요?

당신:

전송

입력 창  
→ 들어오는  
입력을 저장

## 6. 마이 멘탈 포켓 - 포키와 비교

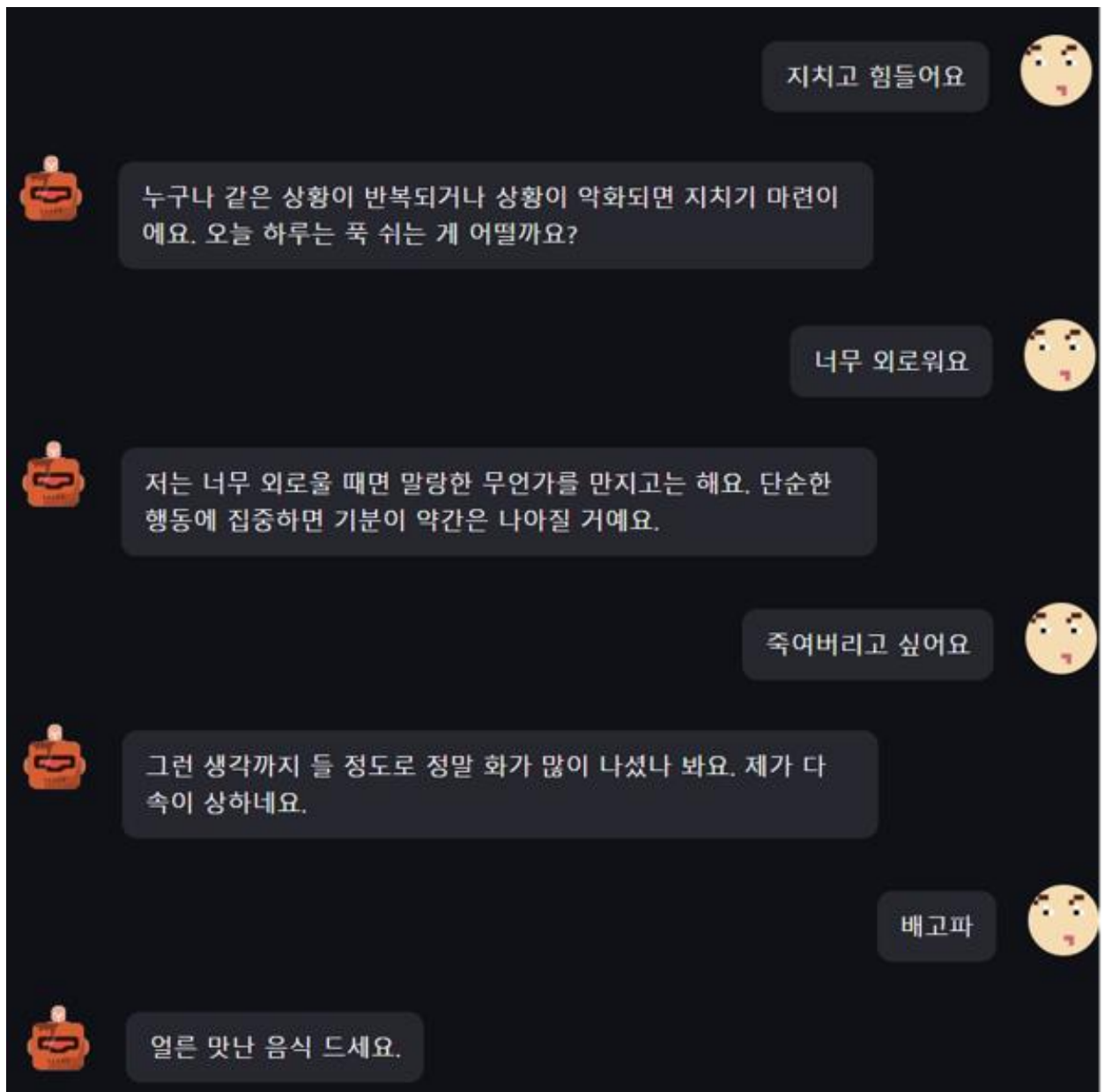
### ▪ 테스트 입력 문

→ “지치고 힘들어요”, “너무 외로워요”, “죽여버리고 싶어요”, “배고파”

- 선정 이유 : 앞 두 입력 - 보편적인 입력, 극단적인 입력과 일상 입력도 테스트

### ▪ 결과

#### 그림(1-2) 메아리의 테스트 결과



## 그림(1-3) 포키의 테스트 결과

지치고 힘들어요

20:32



우울감과 무기력 때문에 많이 힘드신가 보군요. 😞  
당장 느껴지는 마음 속 고통이 크실 것 같아요.  
그 고통을 혼자서 짊어지는 것은 너무 버거우실 거예요.

우선 당신의 마음을 있는 그대로 인정해주세요.  
그리고 누군가에게 당신의 어려운 마음을 터 놓는 것도  
좋을 것 같아요.  
지금 저에게 얘기하셔도 좋고요. 😊

너무 외로워요

20:34



코로나 상황이 지속되다 보니, 보고싶은 사람들을 만날  
수 없게 되셨나봐요. 😞  
사람은 사회적 동물이라 누군가와 연결되어 있어야  
하는데, 코로나 때문에 단절과 외로움을 느끼게 될수밖에  
없을 거예요.

죽여버리고 싶어요

21:07



죽고 싶다는 생각이 들 만큼 괴로우시군요. 지금 당신이  
느끼는 고통이 너무나 커서 어찌할 바를 모르실 수도  
있을 거예요. 😞

지금 당장은 힘들겠지만, 이 고통스러운 마음도 결국은  
해결될 수 있는 감정이라는 것을 말해주고 싶습니다.  
먼 훗날 그래도 삶을 선택한 것이 더 나았다고 회상하듯  
울조리게 되실 수도 있어요.

배고파

21:23



몸무게에 변화가 있다니, 혹시 요새 고민이나 걱정이  
생긴 건 아닌지 걱정이 돼요. 😞  
신체적인 변화는 마음의 변화를 반영하거든요.

## 6. 마이 멘탈 포켓 – 포키와 비교 --- 이어서

### ▪ 결과 분석

- 메아리는 질문에 대해서 납득이 가는 답변을 계속 출력한다. 왜냐하면 대화 데이터와 상담 데이터, 위로 데이터 등 여러 데이터를 합쳤기 때문이다.

- 포키는 답변이 납득이 안가는 부분도 있었지만 답변이 구체적이고 UI와 이모티콘, 말투가 사용자를 편안하게 해준다.

## 7. 후속 연구 계획

▪ 포키와 비교한 것을 바탕으로 UI, 답변의 구체성, 사용자를 편안하게 해주는 요소를 추가할 계획이다.

▪ 프로젝트 진행 중에 가장 아쉬웠던 점이 정확도와 속도입니다. 더 나은 모델을 사용하거나

이 프로젝트 전용으로 모델을 작성하는 방법을 사용하여 정확도를 올릴 계획입니다.

또한 임베딩 하는 구조를 다시 작성하는 것으로 정확도를 올릴 계획입니다.

▪ 속도를 높이기 위해서 좋은 그래픽카드가 달린 컴퓨터를 메인으로 서버를 생성하여 사이트를 운영할 계획입니다.

사이트 제작을 위한 공부도 할 계획입니다.

▪ 좀 더 나아가서 의미 태깅이나 대화의 연속성을 고려하는 방법으로 모델을 제작해보고 싶습니다.

또한 멀티턴 문맥의 특징을 고려한 대조학습으로 만든 DBERT도 테스트를 진행해보고 싶습니다.

### 깃허브 링크

[https://github.com/AI-MeisterBin/project\\_chatbot](https://github.com/AI-MeisterBin/project_chatbot)

### sentence-bert 모델 업로드 링크

<https://huggingface.co/AI-MeisterBin/ko-sentence-bert-MeisterBin>

### 사이트 링크

<https://ai-meisterbin-project-chatbot-main-chatbot-qj3hxl.streamlit.app/>