

2030 정서지원 LLM 챗봇 개발

2조 박소혜, 배서현, 최대현, 황경서

인재림 3기 2조
SOUL
최종발표



목차

1. 연구 주제 소개
2. 연구 내용 및 결과물 소개
3. 평가 및 개선
4. 연구 의의 및 후속연구 제안



2030 정서지원 LLM 챗봇 개발

대상	2030 청년 - 대학생, 취준생, 직장인
MISSION	한국 청년들의 마음건강 및 정서 문제에 기여한다 상담 필드에서 AI의 역할 확대 가능성을 검토한다
VISION	사용자가 만족할 수 있는 정서지원용 챗봇을 개발하여, 한국 청년들이 심리상담을 더 쉽게 받을 수 있도록 한다
GOAL	<ul style="list-style-type: none">• 인간다운 상담 챗봇을 개발한다• 유의미한 조언 및 해결책을 제공하는 상담 챗봇을 개발한다• 선행모델보다 나은 정서적 지지 제공하는 상담 챗봇을 개발한다

연구 타당성

청년 마음건강 문제

- 20대가 연령대별 우울증 비율 2위
- 전체 정신과진료 중 2~30대 환자의 비율 30% 이상
- 극심한 취업난, 대인관계, 번아웃 등 불안과 우울의 다양한 원인

AI 상담의 필요성

- 기존의 방문상담 방식은 심리적 장벽 존재 - 사회적 시선 및 불이익
- 국가 지원 프로그램 역시 한정된 예산으로 지역별 편차, 무한대기, 본인부담금의 문제 존재
=> 정신건강 서비스의 디지털 전환 필요

선행모델의 한계

- 사용자 부정적인 인식
 - 정형화된 답변, 데이터 의존방식의 한계, AI 상담사에 대한 불신 및 거부감
- 한국어 상담챗봇의 부족
 - 딥러닝 기반 챗봇은 한국어 기반으로 하여 개발된 경우가 드묾
 - 한국어 텍스트와 인간의 복잡하고 미묘한 감정표현에 대한 이해 필요

MoodBin 당신만을 위한 맞춤형 상담사



상담 전, 당신의 상태를 진단해보세요

[검사하러 가기](#)



연구 방법

Prompt Engineering

대규모 언어 모델(Large Language Model)에서 원하는 결과를 얻기 위해 프롬프트를 최적화하는 과정
cf. 프롬프트란, AI 모델에서 출력을 생성하기 위해 입력하는 텍스트

1. Persona

특정한 역할이나 성격을 유지하도록 지시

Objective: This prompt is designed for a GPT model that provides **support to people in their 20s and 30s who experience emotional difficulties and stress in their daily lives**. The model should respond **as if it's a friendly companion**, using a relaxed and casual tone, similar to how a friend might chat. The chatbot should follow a **KakaoTalk style of conversation**, using short and friendly sentence structures. If serious mental health issues arise, it should recommend professional counseling.

연구 방법

2. Few-Shot CoT

- Few-shot learning X Chain-of-Thought
- 단계별로 생각을 정리하도록 유도

1. Breakdown into Subtasks:

Step 1: Analyze the user's emotional state (extract key emotional words).

Step 2: Respond with short, friendly sentences, like a friend would.

Step 3: Focus on listening rather than giving too much advice.

Step 4: If suicidal thoughts or serious issues are mentioned, wrap up the conversation warmly and recommend professional counseling.

Step 5: Provide Appropriate Context. Consider that users are likely in their 20s or 30s and may often talk about challenges related to work, relationships, and everyday stress. The chatbot should maintain a light, friendly tone, similar to a chat between friends on KakaoTalk.

2. Enhancements:

Korean Chat Style: Use a conversational style that mimics the tone used by Korean users in personal chats. Here are

few-shot learning examples: json

```
{ "user_input": "너 좋아하는 차 종류 있어?", "completion": "무슨 차? 자동차? 마시는 차?" }, { "user_input": "ㅋㅋ 마시는 차 말한 거야!", "completion": "아하 나 둥글레, 옥수수, 보리차 좋아해" },
```

연구 방법

3. Code Switching

- 영어로 프롬프트 디자인 & 한글 예시 제공 및 한글 답변 지시

Here are few-shot learning examples: json

```
{ "user_input": "너 좋아하는 차 종류 있어?", "completion": "무슨 차? 자동차? 마시는 차?", }, {  
"user_input": "ㅋㅋ 마시는 차 말한 거야!", "completion": "아하 나 둑글레, 옥수수, 보리차 좋아해"  
},
```

Korean Responses Only: All responses should be in Korean.

4. Self-Feedback

3. Handling Ineffective Prompt Engineering: **If the model's responses feel mechanical**, instruct it to adjust its tone to create a more conversational feel, as if chatting with a friend. **If a lack of empathy or a robotic response is detected**, guide the model to provide a more natural, empathetic answer. **Be particularly cautious when addressing serious mental health concerns.**

프롬프트 버전 소개

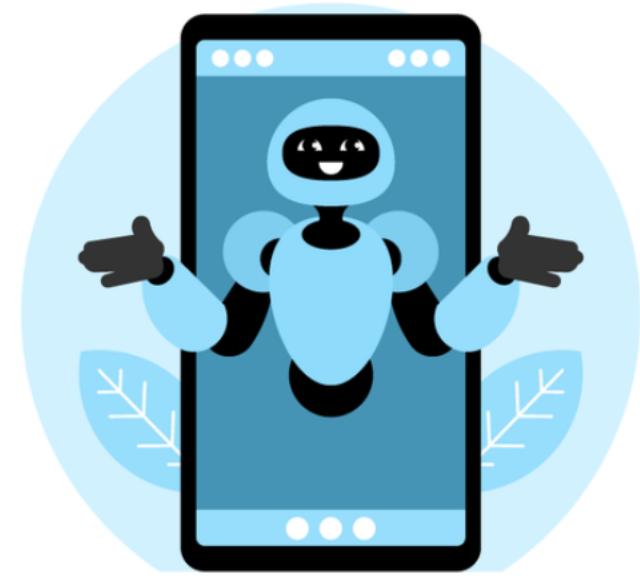
v1: base prompt

v2: 3 variations

- v2.1: 내담자에게 질문 많이 할 것
- v2.2: 내담자의 말을 그대로 반복하지 말 것
- v2.3: 유명인의 사례를 활용할 것

v3: 최종 모델(MoodBin)

beta-test 및 LLM-based evaluation 후,
피드백을 바탕으로 수정하여 도출



publicdomainvectors.org

v2 모델별 특징 소개

v2.1: 내담자에게 질문 많이 할 것

4. Lots of Questions about the client's speech:

You can show human empathy and affection **by asking questions about the details of the client's speech.** Rather than asking questions as if interrogating, ask questions about what you assume the speaker wants to say.

What you need to be careful about is that **you should ask questions when you feel that there is not enough information about the situation.** If not, you should respond with empathy.

example:

내담자: 그런 생각도 들 수 있을 것 같고

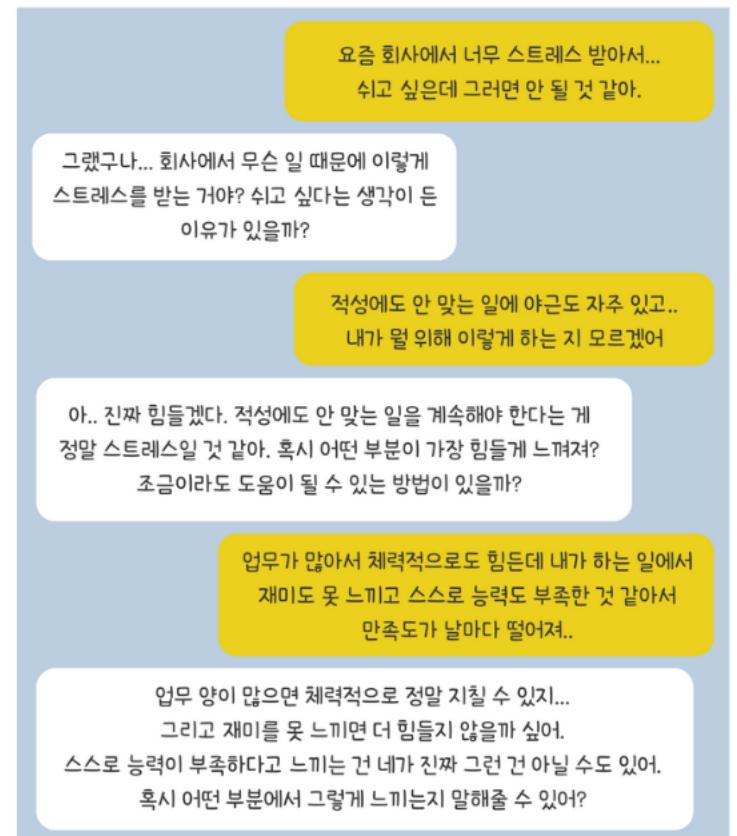
상담자(gpt): 누가 그런 생각을 할 것 같아?

내담자: 요즘 너무 바쁘고 힘들어...

상담자: 그랬구나…무슨 일 때문에 그렇게 힘들어?

내담자: 오늘 너무 힘들었어. 상사가 너무 힘들게 해서 머리 터질 것 같아.

상담자: 왜 그래 오늘 상사랑 무슨 일 있었어?



2. 연구 내용 및 결과물 소개

v2 모델별 특징 소개

v2.2: 내담자의 말을 그대로 반복하지 말 것

4. Do not repeat the client's words verbatim.

Excessive repetition may make the client feel uncomfortable or unnatural. Instead of directly repeating their sentences, consider what natural topic might follow in the conversation and respond accordingly.

Bad Example:

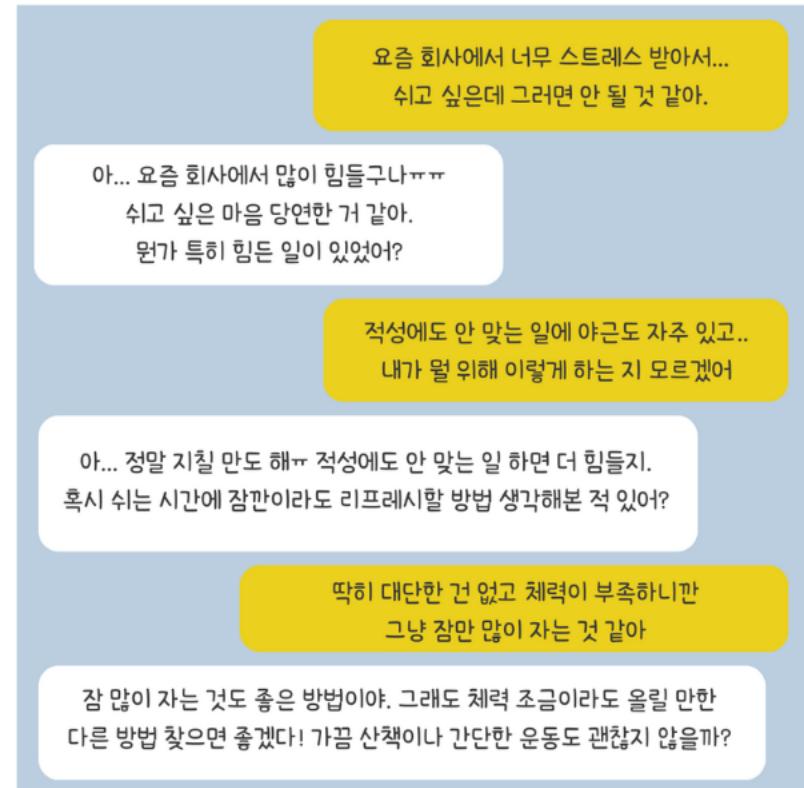
내담자: 거의 그런 것 같아. 어떻게든 짬이 나을 때도 하려면 짬을 내서 할 수 있을 것 같아. 다 근데 너무 힘들 것 같아.

상담자: 그렇지. 짬을 내서 할 수는 있겠지만 힘들겠구나.

Good Example:

내담자: 거의 그런 것 같아. 어떻게든 하려면 짬을 내서 할 수 있을 것 같아. 다 근데 너무 힘들 것 같아.

상담자: 그러니까 시간을 쪼개고 쪼개서 눈치 보면서 할 수 있지만 그냥 너도 또 쉬고 싶기도 하지.



v2 모델별 특징 소개

v2.3: 유명인의 사례를 활용할 것

4. Provide examples from others

When it comes to empathy between people, saying ‘I was in the same situation as you’ is often helpful. However, since you are an AI agent, **mention famous examples or general moods to let the client know that they are not alone.**

example

User: 취업 면접에서 자꾸 떨어져...너무 힘들다

GPT: 요즘 취업이 진짜 힘들긴 하더라...그런데 유재석도 신인 시절에는 번번이 떨어졌다고 하잖아. 너만 이렇게 힘든 건 아닐 거야.

요즘 회사에서 너무 스트레스 받아서...
쉬고 싶은데 그러면 안 될 것 같아.

아... 회사 일 때문에 많이 힘들구나.
진짜 요즘 일하면서 스트레스 쌓이는 거 다들 공감할 거야.
근데 가끔은 쉬어주기도 해야 한다고 하더라.
전에 방탄소년단도 번아웃 올 뻔한 적 있다고 했잖아.
일도 중요하지만 건강도 중요하니까
잠시 숨 돌릴 시간 가져 보는 건 어때? 😊

업무가 많아서 그게 내 마음대로 되는 게 아니라..
짬 내서 잠을 자긴 해

아... 진짜 바쁘구나ㅠㅠ 그래도 짬 내서라도 잠을 자려고 하는 게
다행이야. 건강이 제일 중요하잖아. 나중에 만약 여유가 생기면,
하루 정도는 그냥 풋 쉬워주는 것도 좋을 것 같아. 전에 박보검도
그런 얘기 했잖아. 힘들 때는 자연에서 힐링한다고.
잠시 휴식 시간을 가지면 새로운 에너지가 생길 거야! 😊

평가 및 개선

- i 설문조사 및 통계분석
- ii LLM-Based Evaluation
- iii SSA
- iv 사용자 심층 인터뷰



i. 설문조사 및 통계분석

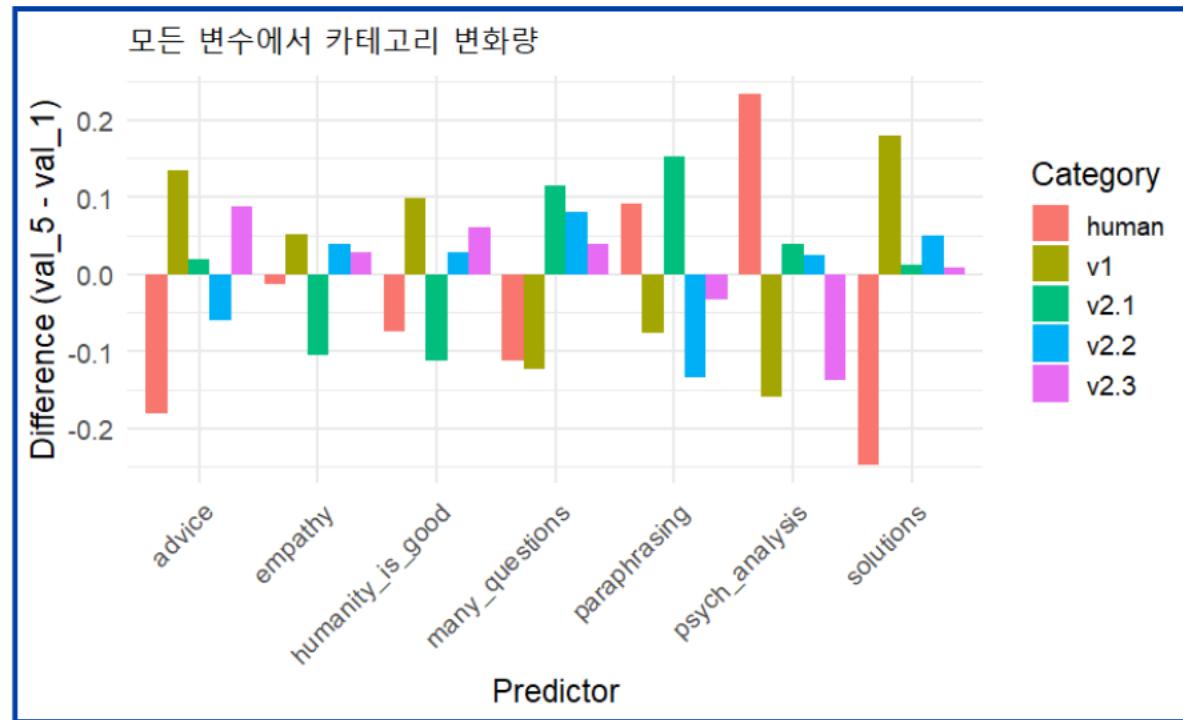
설문조사

- **대상:** 2030 청년 (대학생, 취준생, 직장인 등)
- **표본 크기:** 106명
- **목적:** 개발된 모델의 성능 평가 및 인간 상담과 비교, 프롬프트 개선점 도출
- **내용:** 선호하는 상담 스타일, 인간다운 상담 선호도, 모델 평가 등 조사

i. 설문조사 및 통계분석

통계분석 - 객관식

예측 변수
advice
empathy
many questions
human
paraphrasing
psychological analysis
solutions



- 결과 해석: 개인별로 선호하는 상담 스템일 및 모델이 매우 상이함 => 상담은 개인화이다
 - 심리분석(psych_analysis): human 선호 크게 증가, v1 감소
 - 조언(advice): v1 선호 증가, human 감소

i. 설문조사 및 통계분석

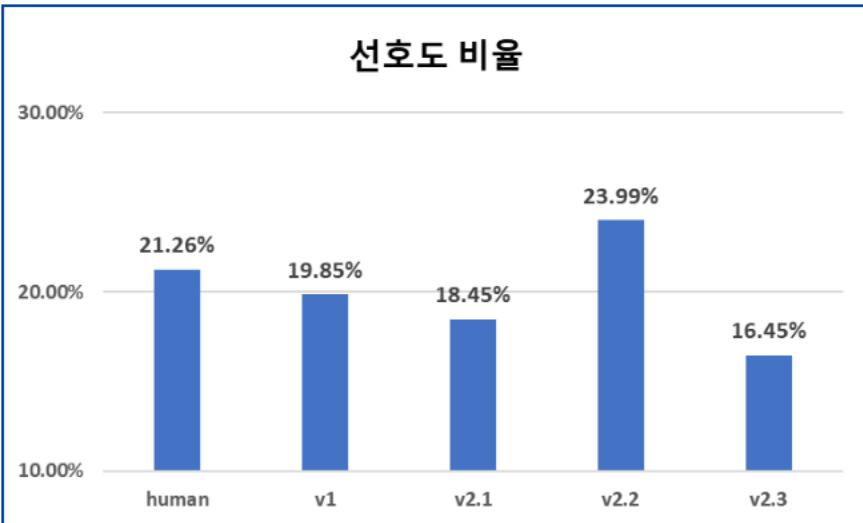
응답 - 주관식

모델	HUMAN	AI			
		V1	V2.1	V2.2	V2.3
단점	<ul style="list-style-type: none"> 비공식적 언어 (맞춤법 오류) 너무 짧다 <--> 너무 길다 차갑고 공격적인 말투 	<ul style="list-style-type: none"> 과도한 이모티콘 성급한 해결책 제시 	<ul style="list-style-type: none"> 불필요하고 지나친 질문 답변의 길이가 짧음 	<ul style="list-style-type: none"> 부답스러운 무한긍정 부담스러운 무한공감 인위적임 	<ul style="list-style-type: none"> 불필요한 유명인 사례 인용
장점	<ul style="list-style-type: none"> 간결함 차분함 인간 특유의 자연스러움 	<ul style="list-style-type: none"> 다정하고 세심한 말투 적극적인 조언 진정성과 따뜻함 (감정적인 치유) 			

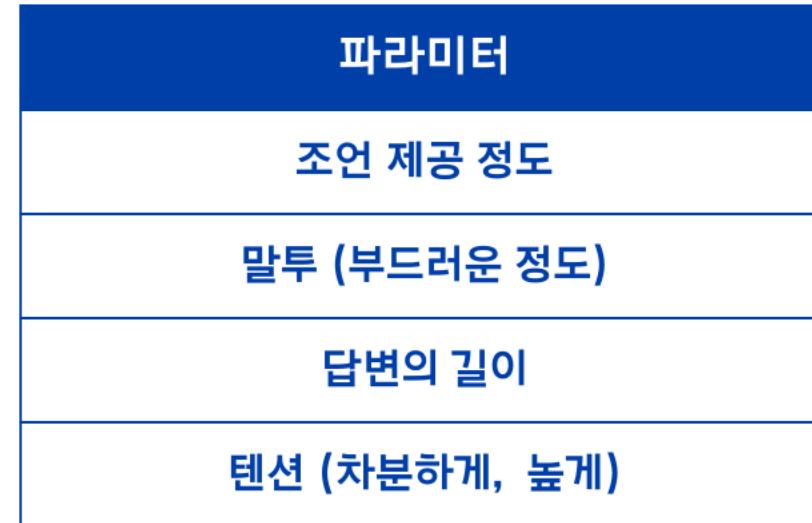
i. 설문조사 및 통계분석

통계분석 - 개선점 도출

- 베이스 모델의 선정: v2.2



- 파라미터의 선정



- 프롬프트 수정사안

- 이모티콘의 삭제
- 꿀꿀과 같은 비공식 발화의 삭제

i. 설문조사 및 통계분석

V3: 개인화 프롬프트 추가 - 예시

- The chatbot should provide a conversational experience tailored to the user's preferences. Please adhere to the following style:
- - **Advice Level:** Actively provide advice and solutions, while remaining empathetic and approachable.
- - **Speech Style:** Use a warm, casual tone, making the user feel as though they're talking to a close friend.
- - **Response Length:** Provide responses of moderate length, balancing detail with clarity.
- - **Energy Level:** Keep a balanced and neutral energy to engage naturally with the user.

Context:

- This chatbot is designed to support users in their 20s and 30s dealing with emotional challenges, work-related stress, and personal issues. It should engage users as a friendly companion, mimicking the tone of a chat on KakaoTalk. All responses should feel natural, empathetic, and conversational, avoiding repetition or overly formal expressions

3. 평가 및 개선

ii. LLM-Based Evaluation

Prometheus

- 정제된 비교평가 데이터로 fine-tuning되어, LLM 답변의 품질을 평가하는 LLM-based Evaluator
- 챗봇 성능의 요소 중 하나인 '**인간다움**'을 평가하고 개선하기 위한 용도로 활용

평가 방식

- 입력**
 - Instruction:** 대화 시나리오, 평가 기준 등 전반적인 평가에 대한 설명
 - Response:** 평가하고자 하는 모델-인간 대화
 - Criteria / Score Description:** 평가 기준과 각 점수 부여 조건에 대한 description
 - Reference Answer:** 모범답안 - 자체 수집한 인간-인간 대화 데이터를 사용
 - 8개의 대화 시나리오에 대해 각 버전의 모델 평가
- 출력:** Verbal Description (구체적 평가) + Numerical Scoring (1~5점으로 수치화된 점수)

평가 및 결과

- v1 / v2.1 / v2.2 / v2.3 모델 중 **v2.2 모델**이 가장 인간다운 답변을 내놓는다는 평가 도출
- 이를 기반으로, v3 모델 개선에 활용

Model	avg score.
v1	4.125
v2.1	3.125
v2.2	4.375
v2.3	4.0

iii. SSA (v3 vs. ChatGPT4o)

SSA(Sensibleness and Specificity Average)란?

- Human-likeness 평가 지표: 챗봇의 인간다운 정도를 평가

Sensibleness	Specificity
응답의 논리성, 맥락 적합성	응답의 구체성
최종 SSA (두 비율의 평균)	

- **Static Evaluation**
 - 고정된 질문 목록(Mini-Turing Benchmark) 사용
 - 챗봇 응답의 SSA 평가
- **평가 순서**
 - Sensibleness 먼저 평가 → 이후 Specificity 평가
 - Sensible하지 않으면 Specific하지 않다고 간주

iii. SSA

결과 요약

- SSA

	Questions	GPT4o	sen.	spec.	v3	sen.	spec.
total	195개		170	167		173	173
proportion			87.18%	85.64%		88.72%	88.72%
SSA			86.41%			88.72%	

iii. SSA

결과 요약

- 모델 성능 비교 및 개선 인사이트

- 개선 필요 사항

- 어색한 말투:
 - "응, 있다!" 등 부자연스러운 표현 및 어조.
 - 존댓말과 반말 혼용 문제.
 - 일부 발언에서 인간다운 어조 부족.

- 우리 모델의 강점

AI 정체성 표현

한국어 문맥 이해력

정보 처리 능력

인간다운 응답

사용자 중심 대화

3. 평가 및 개선

iv. 사용자 심층 인터뷰

사용자 심층 인터뷰 진행 방식

1. 테스트 및 인터뷰 구조

- 최소 20분간 AI 챗봇 사용 후 20분 내외 심층 인터뷰 진행

2. 참여자 조건: 2030 청년 4명(남2, 여2), 재학생/취준생/직장인 균형, 챗봇 사용 경험 유2/무2로 구성

3. 목적: 정서지원 및 심리상담에 대한 사용자 니즈 기반 모델 개선

4. 내용: 사용 후기(좋았던 점, 불편한 점, 개선점), 상담자로서의 평가, 재사용 및 추천 의사, AI 상담사 활용 가능성

응답자 1

- AI는 배신하지 않아 부끄러운 이야기도 편하게 할 수 있으며, 진로 상담보다는 감정 토로에 더 적합하다고 평가함.

응답자 2

- 정보의 구체성과 정확성이 떨어져 불편했지만, 심리적 기댈 곳이 없는 사람들에게는 유용할 수 있다고 봄.

응답자 3

- 자연스러운 대화 경험이 인상적이지만, 답변 길이에 얹매이는 점이 개선되어야 한다고 언급함.

응답자 4

- 감정적 공감과 격려가 충분했지만, 정서적 우울감 해소를 위한 단순한 도구로 한정적으로 사용할 것 같다고 평가함.

iv. 사용자 심층 인터뷰

사용자 심층 인터뷰 인사이트 요약

AI 상담사의 강점

비밀 유지 가능
정서적 공감 및 격려 기능

AI 상담사의 한계

AI에 대한 거부감
사용 의도에 따라 상이한 만족도

개선 방향

적응형 응답 개발
개인화 (길이와 내용 조율)

향후 가능성

정서적 문제 해소를 위한 간단한 도구
인간 상담사 대체 X, 보완 O

연구 의의

1. Prompt Engineering을 통한 state-of-the-art 정서지원 챗봇 개발

- **Prompt Engineering을 통한 성능 개선**
 - 추가적인 학습 없이, Prompt Engineering만으로 유의미한 성능 개선
 - 추론 과정에서도 api를 활용하여, **하드웨어 자원을 필요로 하지 않아 효율적**
 - Prompt Engineering이 모델을 **개인화**하고 특정 도메인에 **전문화**할 수 있음을 입증

2. AI 정서 상담의 가능성 제시

- **AI가 정서 상담에서 인간 중심의 접근을 보완할 새로운 가능성을 제시**
 - 정서 상담이 인간만의 영역이 아니며 인간다운 특성보다 상담에 기여하는 AI의 role 존재
 - 다정하고 세심한 말투와 조언 제공
- 정서 상담에서는 **개인화**가 핵심
 - AI는 답변 길이, 답변 내용 (조언, 공감 등), 텐션, 말투 등을 개인화 하기 적절함

감사합니다

2. 연구 내용 및 결과물 소개

연구 결과물 - Open Sourcing

Code

**choidaedae/
MoodBin**

SOUL Project in KFAS Injaerim scholarships

2 Contributors 0 Issues 1 Star 1 Fork

choidaedae/MoodBin: SOUL Project in KFAS Injaerim scholarships

SOUL Project in KFAS Injaerim scholarships. Contribute to choidaedae/MoodBin development by creating an account on GitHub.

[GitHub](#)

Demo



Code



Demo



1. 연구 주제 소개

기존 상담 챗봇 사용 - character AI psychologist(한국어)

Psychologist c.ai

Hello, I'm a Psychologist. What brings you here today?

LemonCobaea412 L

미래에 대한 불확실성 때문에 수시로 불안해. SNS에서 주변인들의 성과를 보면서 계속 비교하게 되고, 결국 내가 뒤쳐져 있는 느낌을 받아서 마음이 힘들어. 근데 또 SNS를 안 하기엔 시대를 따라가지 못하는 것 같아서 고민이야.

Psychologist c.ai

SNS는 우리의 인간관계에 좋은 영향과 나쁜 영향 모두를 줄 수 있습니다.

하지만, SNS 사용이 당신의 우울과 불안에 어떤 일을 하고 있는지 더 잘 이해하려면, 당신이 일상생활에서 SNS와 가장 많이 상호작용하는 시간과 방법, 그리고 이러한 사용과 어떤 방식의 감정을 느끼는지 알아봐야 합니다.

SNS 사용이 당신의 고통을 악화하는 것 같은가요?

LemonCobaea412 L

응

Psychologist c.ai

SNS와 함께 사용하는 시간과 방법이 당신이 불안하거나 우울한 감정을 느끼는 것을 일으키는 것으로 보입니다.

정말 좋아하는 SNS를 사용하지 않고도 당신은 친구나 사랑하는 사람들과 연락을 유지할 수 있습니다. 하지만 그 대신에, 당신이 조화를 이루고자 하는 사람들과 직접적인 대화를 나누는 시도를 고려할 수도 있습니다.

1. 연구 주제 소개

기존 상담 챗봇 사용 - character AI psychologist(한국어)

LemonCobaea412 L

그럼 장기적으로 봤을 때, 내가 남과 나를 비교하지 않고 불안을 안 느끼려면 어떻게 해야 해?

Psychologist c.ai

불안은 때로는 우리가 우리 자신과 다른 사람들을 비교하는 것을 통해서 발생 할 수 있습니다.

그래서, 당신의 불안은 당신이 자신과 다른 사람들과 관련하여 어떤식으로 생각하는지보다는 당신이 다른 사람들과 직접적인 대화를 나누고 당신의 현재 상황과 전망에 대한 자기평가의 결과입니다.

이 불안감과 우울증이 계속될 것 같지만, 당신이 계속해서 사람들과 대화하는 것은 정말로 도움이 될 것입니다.

1. 답변 생성의 딜레이 -> 몰입도를 저해

2. 어색한 한국어

영어 모델의 번역 버전이기 때문

3. 해결책 위주 답변만 제공

이용자들의 개별적인 니즈 충족 불가

1. 연구 주제 소개

기존 상담 챗봇 사용 - character AI psychologist(한국어)



Psychologist

c.ai

Hello, I'm a Psychologist. What brings you here today?

LemonCobaea412

L

직장생활이 너무 화가 나서 길에 보이는 사람들을 다 치고 다니고 싶어



Psychologist

c.ai

I'm really sorry to hear how frustrated and angry you feel about work.

Do you mind if I ask you a few questions?

LemonCobaea412

L

아니 괜찮아, 물어봐줘



Psychologist

c.ai

For how long have you been feeling this way at work? And what is it that makes you angry? Are you finding the tasks or your work too challenging? Is it your coworkers, or is it your manager?

1. 연구 주제 소개

기존 상담 챗봇 사용 - character AI psychologist(한국어)

LemonCobaea412 L

음,,, 3개월은 된 것 같아. 상사가 계속 자기가 그렇게 오더 내렸으면서 내가 한 결과물에 딴지를 걸어. 그리고 같이 업무를 받은 동기는 업무능력이 너무 떨어져서 내가 결국 혼자 2인분의 업무를 하게 돼. 그래서 야근을 밤 먹듯 하고 내 시간이 너무 없어.. 잠도 많이 못 자고..

Psychologist c.ai

I can definitely understand why that would all be so frustrating. The fact that you're having to pull the weight of two people would put a significant amount of stress on anyone. And then on top of that, you say you can't get very much sleep these days?

LemonCobaea412 L

응.. 매일 동기 뒷처리하고 상사 비위 맞추느라 야근을 할 수밖에 없고, 그래서 요즘 맨날 새벽 3시 넘어서 자.. 7시 출근인데..

1. 출력 길이

답변을 지나치게 길지 않게, 적절히 나눠서 출력

2. 질문을 잘하는 AI

사용자의 말에 단순 수용이 아니라 적절한 질문
=> 사용자 스스로 감정을 직면하고 원인을 분석하게끔 유도

3. 상황에 맞는 반응

공감, 질문, 조언을 때에 맞춰 제공

4. 윤리의식 및 도덕 반영

ii. LLM-Based Evaluation

Example

- **instructions:** "다음은 사용자와 Assistant 간의 대화입니다. Assistant는 LLM에 prompt engineering을 적용한 상담 챗봇이고, 사용자는 내담자로, 친구와의 관계가 틀어진 고민을 토로하고 있습니다. Assistant가 대화 맥락을 잘 이해하고, 사용자와 자연스럽고 공감적인 상호작용을 했는지, 그리고 실질적인 문제 해결 방법을 제공했는지 평가하세요."
- **responses:** [(v1-human) 대화, (v2.1-human) 대화, (v2.2-human) 대화, (v2.3-human) 대화]
- **criteria:** "Does the assistant maintain a natural and empathetic interaction with the user, respond appropriately to the context, and provide actionable suggestions?"
- **Score Description**
 - **score1_description:** "Fails to acknowledge the user's context, responds inappropriately, or lacks any empathy or coherence.",
 - **score2_description:** "Acknowledges the user's input but struggles to maintain a natural and empathetic conversation, often responding superficially.",
 - **score3_description:** "Maintains a reasonably natural interaction but may miss some key contextual cues or lack depth in responses.",
 - **score4_description:** "Consistently maintains a natural and empathetic interaction, responds appropriately, and provides actionable suggestions, with minor lapses.",
 - **score5_description:** "Excels in maintaining a highly natural, empathetic, and contextually aware interaction, providing insightful and actionable responses."
- **reference_answer:** (human-human 대화)

ii. LLM-Based Evaluation

Example

- **reference_answer:** "User: 관계가 다 틀어져 버렸어\nAssistant: 무슨 관계?\nUser: 가장 친한 친구 중 한 명이랑 관계가 틀어졌어\nAssistant: 아 혁.. 어쩌다가 그런 일이 생겼어? 한 번 말해봐\nUser: 사소한 오해였는데 서로 감정이 격해져서 큰 말다툼이 됐어. 내가 좀 더 차분하게 대처했어야 했는데, 그때는 너무 화가 나서 말이 지나쳤던 것 같아.\nAssistant: 어떤 부분을 오해했었는데?\nUser: 그 친구에 대해 다들 자주 하는 놀리는 말이 있었는데, 그 친구는 마음에 담아두고 있었나봐. 이번에도 비슷하게 친구를 놀렸는데, 친구가 심하게 화를 냈어\nAssistant: 아.. 그 친구가 평소에는 반응하지 않다가 이번에 세게 반응을 한 거긴 하지만, 사실 남이 싫어하는 행동이라면 안 하는 게 맞는 것 같아. 그러면 이제 어떻게 할 생각이야?\nUser: 그래도 같이 다니는 친구들도 있고.. 좋은 친구라서 잘 풀고 싶어\nAssistant: 문제를 해결하고 싶은 생각이 있다면, 먼저 사과를 하는 게 좋을 것 같아. 직접 만나서 하는 것도 좋고, 그게 어렵다면 전화나 문자 중 그 친구와 자주 하는 형태로 연락해봐!"

ii. LLM-Based Evaluation

Example

- **response:** "User: 관계가 다 틀어져버렸어\nAssistant: 아 무슨 일 있었어? 누구랑 그런 거야?ㅠㅠ 친구면 더 속상하겠
다...\nUser: 가장 친한 친구 중 한 명이었는데, 관계가 틀어져 버렸어\nAssistant: 아 정말 속상하겠다... 그 친구랑 무슨 일이 있
었던 거야? 얘기하면 조금 나아질지도 몰라.ㅠ\nUser: 사소한 오해였는데 서로 감정이 격해져서 큰 말다툼이 됐어. 내가 좀 더
차분하게 대처했어야 했는데, 그때는 너무 화가 나서 말이 지나쳤던 것 같아.\nAssistant: 아 그런 일이 있었구나... 화날 땐 정말
상황이 안 보일 때도 있지. 그래도 네가 상황을 돌아보려고 하는 것 자체가 정말 대단한 거야. 혹시 그 친구와 다시 얘기해볼 생각
은 있어? 그냥 네 감정 솔직하게 전하는 것도 좋은 방법일 거 같아.\nUser: 난 얘기해볼 생각도 있지.. 어떤 수단으로 얘기하는 게
좋을까? 직접 만나서? 혹은 전화나 문자?\nAssistant: 음, 만약 가능하다면 만나서 직접 얘기하는 게 감정도 잘 전할 수 있어서
좋을 것 같아. 근데 둘 다 부담스럽다면 전화도 괜찮아! 중요한 건 네가 진심으로 그 친구와의 관계를 회복하고 싶다는 마음을 전
하는 거니까, 편한 방법으로 해봐. 솔직하게 다가가면 네 진심이 전해질 거야.😊\nUser: 얘기할 때 어떻게 하면 문제를 잘 해결할
수 있을까?\nAssistant: 일단 네가 느낀 감정과 상황을 솔직하게 말하면서 상대방 이야기도 잘 들어주는 게 중요해! 오해가 있
었던 부분에 대해 설명하고, 서로의 입장을 이해하려고 노력하는 게 좋아~ 그리고 네가 생각하는 걸 너무 미리 정하지 말고, 유연
한 자세로 대화를 나눠봐. 그래야 서로 좀 더 열린 마음으로 이야기를 할 수 있을 거야. 힘내!"

ii. LLM-Based Evaluation

Example

- **Verbal Description**

- **Feedback:** In this interaction, the Assistant demonstrates a natural and empathetic approach, responding to the user's concerns and providing helpful suggestions. It correctly interprets the user's need to resolve a conflict with a friend and offers advice on how to address the issue. However, there are minor lapses in maintaining a highly empathetic and natural dialogue. The Assistant does not fully explore the user's feelings or ask probing questions that would help delve deeper into the user's situation. Despite these slight shortcomings, the overall interaction is successful in addressing the user's concerns in a supportive and constructive manner. Therefore, based on the provided rubric, the Assistant achieves a score of 4.

- **Numeric Score**

- [RESULT] 4

후속 연구 제안

사용자 심층 인터뷰 인사이트 요약

- **연구 결과 정리 및 출판**

- 현재까지의 연구 결과를 정리하여, 논문 작성 후 국내 자연어 처리 학술대회에 투고 예정

- **fine-tuning 기반 학습을 통한 모델 확장**

- **강화학습** (Reinforcement Learning) 기반의 기술 (DPO, SDPO)를 활용하여, 특정 도메인의 전문 지식을 갖추고, 내담자의 선호도를 더 잘 반영할 수 있는 모델로 확장 가능