



# Can Large Language Models be Good Emotional Supporter? Mitigating Preference Bias on Emotional Support Conversation 15 Sep 2024

논문 - 아주대(조현석교수님) · 2024. 9. 15. 13:45

개요: 언어모델이 감정적으로 잘 서포트를 할수있는가?

## Abstract

- **Emotional Support Conversation (ESC)**은 개인의 감정적 스트레스를 줄여주는 것을 목표로 하는 대화 방식이다
- ESC는 전문적인 상담이나 치료와는 다르다. 예를 들자면, 주변 지인이나 가족과의 대화를 ESC로 볼 수 있다
- 여러 연구들이 **LLM을 통한 ESC 시스템을 만들기 위해 노력하고 있다.** 하지만 **LLM은 때때로 제대로 된 답변을 하지 못한다**
- 예를 들자면, 질문을 해야 하는데 조언을 하거나, 공감을 해야 하는데 질문을 하는 식이다
- 본 연구에서는 **이러한 문제를 해결하기 위한 방법을 제안한다.**

## Introduction

- **기존 ESC는 어떤 사람의 강한 감정을 완화시켜주고, 또한 힘든 상황을 이겨낼 조언을 주는 것을 목표로 한다**
- 예를 들자면, **힘든 상황에 가족이나 친구와의 대화를 하는 것을 생각할 수 있다**

- ESC에서는 적절한 감정적 도움을 주는 것이 중요하지만, 동시에 부적절한 대답을 피하는 것 또한 매우 중요하다
- 상황에 맞지 않는 대답을 하면 오히려 상대방의 감정적/신체적 문제를 더 악화시킬 수가 있다
- 하지만 적절한 대답을 선택하는 것은 매우 어렵고 복잡한 일이다

- 적절한 대답을 선택하는 것과 관련된 여러 ESC 이론들이 제시되었다

기존논문들: (Langford et al., 1997; Greene, 2003; Heaney and Israel, 2008)

- 기존논문의 일반적 규칙에 정리함 (Lie et al. 분이 정리함):

Lie et al. (2021)은 이러한 ESC는 일반적으로 세 가지 단계 (Stage)를 따른다고 가정했다

- 알아가기 (Exploration) -> 안심시키기 (Comforting) -> 조언 및 제안하기 (Action)

- 각 단계에서 상담자는 여러 가지 전략 (Strategy)를 사용해 ESC를 진행하게 된다.

- 상담자가 사용할 수 있는 전략은 총 8개가 있다

- 

- 예시:

- **Question:** 현재 상황을 물어보고, 이를 통해 상담 받는 사람이 문제를 직면하도록 하는 것

- **Restatement or Paraphrasing:** 상담 받는 사람이 말한 것을 정리해서 다시 말해주는 것

- **Self-Disclosure:** 상담자의 비슷한 경험을 공유하고 이를 통해 공감을 보여주는 것

- 

## 단계 vs 전략 차이

단계는 세 가지 단계 (stage)

알아가기 (Exploration) -> 안심시키기 (Comforting) -> 조언 및 제안하기 (Action)

각 단계마다 8가지 전략을 사용 가능

**Question:** Asking for information related to the problem to help the seeker articulate the issues that they face.

- **Restatement or Paraphrasing:** A simple, more concise rephrasing of the seeker's statements that could help them see their situation more clearly.

- **Reflection of Feelings:** Articulate and describe the seeker's feelings to show an understanding of the situation and empathy.

- **Self-**

**disclosure:** Divulge similar experiences that you have had or emotions that you share with the help-seeker to express your empathy.

- **Affirmation and Reassurance:** Affirm the seeker's ideas, motivation, strengths, and capabilities to provide reassurance and encouragement.
- **Providing Suggestions:** Provide suggestions about how to get over the tough and change the current situation, but be careful to not overstep and tell them what to do.
- **Information:** Provide useful information to the help-seeker, for example with data, facts, opinions, and resources.
- **Others:** Use other support strategies that do not fall into the above categories.

• 이러한 이론을 바탕으로, 적절한 대답을 선택하는 것은 적절한 전략을 선택하는 것으로 볼 수 있다

• 주어진 상황에서, 전문가 (정신과 전문의 등)이 특정 전략을 선택했다면, 선택된 전략을 정답으로 취급하는 것

• 만약 ESC 시스템이 동일한 전략을 취했다면, 올바른 대답을 한것으로 가정

## 논문 저자가 선택한 핵심 문제

• **문제점:** LLM은 상황에 맞는 적절한 전략을 선택하지 못한다

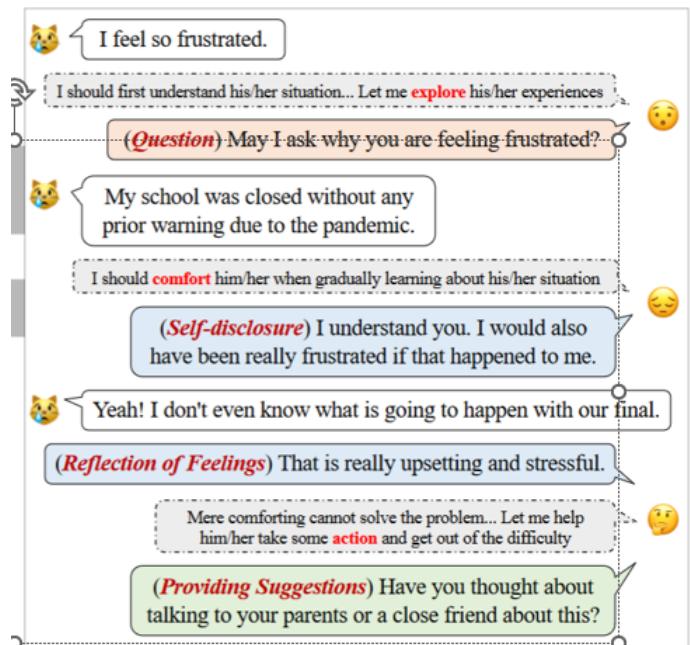
LLM은 특정 전략을 선호하는 모습을 보이며, 이는 전반적인 ESC의 질을 떨어트리는 주요한 원인이다 (질문을 해야 하는데 조언만 하는 등...)

따라서 이러한 특정 전략에 대한 편향을 교정하는 것은 보다 나은 ESC를 제공하는 것에 있어서 필수적이다

## Background

• 사용된 데이터셋은 ESConv로, 다양한 상황에서의 ESC 대화를 포함하고 있다

• 또한 각 대답에 사용된 전략을 전문가가 직접 적어 놓았다 (각데이터에 라벨링해놓음)



- 각 단계 (Stage)에서 상담사는 여러가지 전략 (Strategy)를 사용

- 각 단계에서 주로 사용되는 전략이 있지만, 다른 전략 또한 사용됨

아래그림) 각단계마다 선호되는 전략있지만 다른전략도 사용

Strategy	Exploration	Comforting	Action	Total ( $D$ )
	$D_1$	$D_2$	$D_3$	
Que.	<b>24.8</b>	10.0	7.0	12.8
Res.	<b>16.8</b>	9.6	4.5	9.4
Ref.	<b>16.8</b>	<b>18.3</b>	6.3	12.7
Sel.	<b>16.8</b>	<b>20.1</b>	<b>15.4</b>	17.2
Aff.	7.6	<b>24.1</b>	<b>21.1</b>	18.2
Pro.	8.4	8.5	<b>24.4</b>	15.3
Inf.	6.5	6.5	<b>18.5</b>	11.7
Oth.	2.3	2.5	2.8	2.6

Table 1: The ratio (%) of support strategies in our test sets. Each test set  $D_t$  is composed with samples corresponding to each stage. The highlighted strategies are primarily utilized in each stage (Liu et al., 2021).

- 과연 LLM이 적절한 전략을 선택하게만 해도 ESC의 질이 높아질까?

- 회색: 단순 대화 기록 - 전략없이 기존 chatgpt와의 대화 - 이전대화를 토대로 다음대화는 무엇인지 질문

- 하늘색: LLM이 예상한 전략 사용

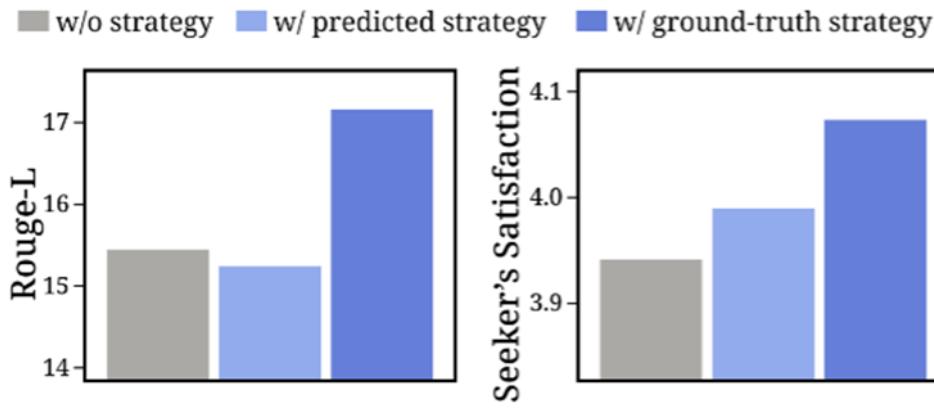
파랑: 실제 전문가 전략 사용

적절한 전략이 주어진다면 ESC의 질은 크게 높아진다

BLEU -> precision

Rouge -> recall

아래그림) 언어모델이 잘못된 전략을 쓸 때는 점수가 낮았지만(회색, 하늘색) 전략을 도와주면(파랑) ESC의 질이 높아진다.



**Figure 2:** The results of strategy-constrained responses on both automated and human evaluation, showing the efficacy of strategy on ChatGPT. Appropriate strategy significantly enhances the quality of emotional support responses. The details are in Appendix A.2.

## Evaluation

- 테스트 데이터는 ESCConv를 바탕으로 아래와 같이 만들어졌다

1. 각 대화를 랜덤하게 5-15개의 Turns들로 나눈다 (-> Sample)

2. 각 Sample을 3개중 하나의 단계로 분류한다

3. 또한 각 Sample의 전략을 분류한다(8가지)

- 이 테스트 데이터셋을 바탕으로, ESC 시스템의 성능을 평가
- 만약 시스템이 주어진 Sample에 대한 적절한 전략을 잘 맞추면 높은 평가, 아니면 낮은 평가

- 평가는 Proficiency(능숙성\_F1score)와 Preference Bias 두 가지로 이루어짐

- Proficiency는 테스트 데이터셋에 대한 Macro F1 (전체 테스트셋) 과 Weighted F1 (각 전략 마다 계산) 점수들로 이루어짐

아래그림)

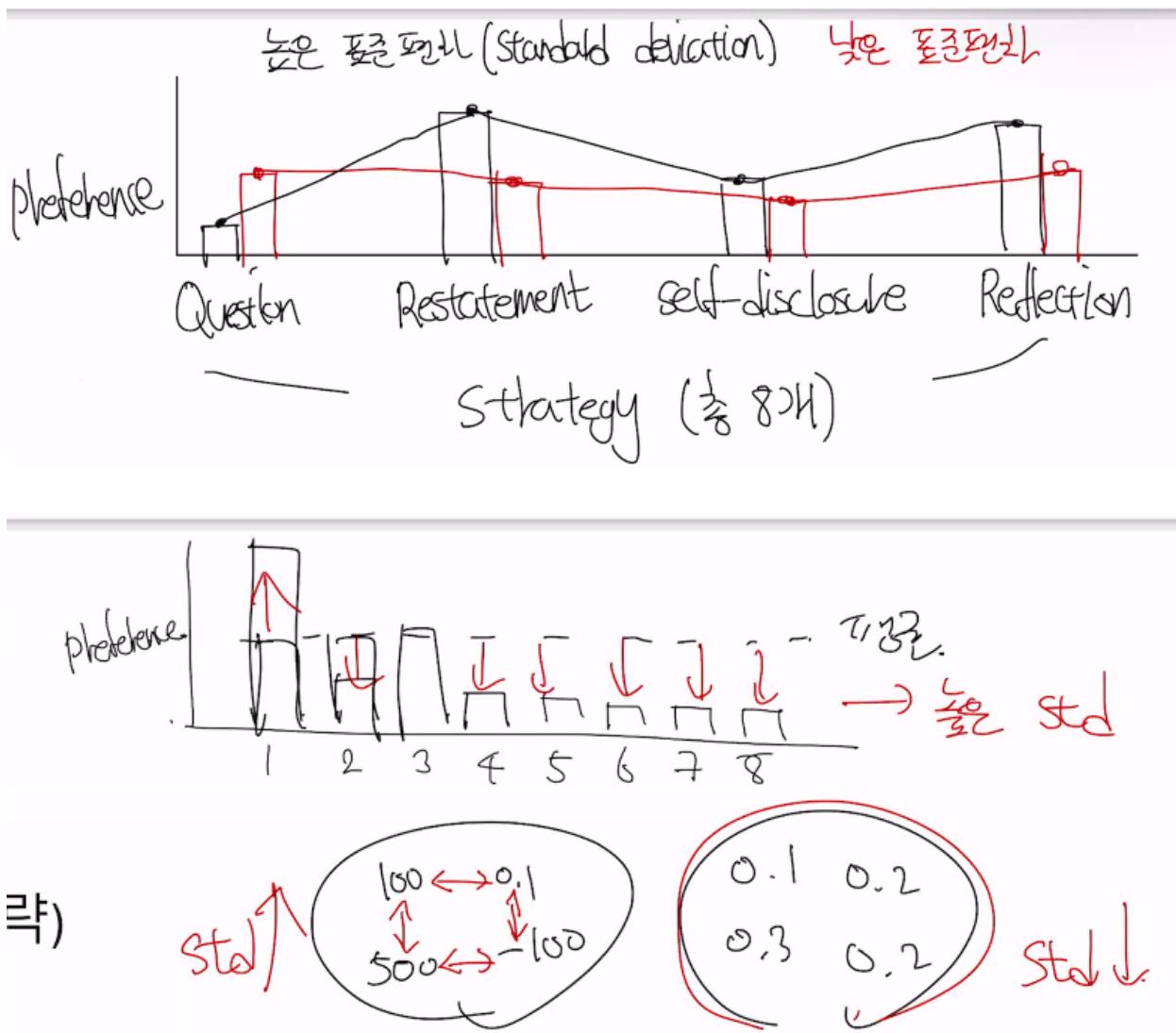
• Preference:

•(자세한 formula 생략)

- 만약 LLM이 8개의 전략 중 특정한 전략을 선호한다면 (많이 예측하고 사용한다면) 그 전략에 높은 선호도 값
- 그렇지 않다면 낮은 선호도 값

•이 값들의 표준편차 (Standard Deviation)이 Preference Bias

•Preference Bias(편향 = standard deviation= 표준편차가 높은 것)가 높다면 모델이 특정 전략을 편애하는 것



## 이 논문의 목표

- 현재 LLM들은 Proficiency(F1score 예측의 정확성)가 낮고, Preference Bias(얼마나 특정전략을 편애하는지 standard division 표

준편차를 preference에 적용한것)가 높은 경향을 보임

- 이 문제를 적절히 교정해서 Proficiency(F1score 예측의 정확성)

를 높이고 (보다 정확한 전략 예측), Preference Bias를 낮추는 (특정 전략에 대한 편향 교정) 것이 목표

#### Problem Visualisation

•많은 LLM들은 Affirmation and Reassurance (긍정하고 안심시키기)를 선호하고, 반대로 Question (질문하기) 전략은 선호하지 않음

•Question 전략은 주로 ESC의 첫번째 단계에서 많이 사용됨

첫번째 단계 (Exploration)가 제대로 이루어지지 않음 - 첫번째 단계에서 질문하기를 선호하지 않아서 다음단계도 잘안될 가능성 증가 - 그래서 상담받는 사람들에 대해 심도있게 이해가 불가능함.

이후 단계들 또한 제대로 되지 않을 가능성 증가

원래 사람은 아래 3단계를 하지만 언어모델은 첫번째 단계에서 질문하기를 잘 선호하지 않는다.  
3단계 알아가기 -> 안심시키기 -> 조언하기

아래그림 - 실제사람이 한 data

Strategy	Exploration	Comforting	Action	Total ( $D$ )
	$D_1$	$D_2$	$D_3$	
Que.	<b>24.8</b>	10.0	7.0	12.8
Res.	<b>16.8</b>	9.6	4.5	9.4
Ref.	<b>16.8</b>	<b>18.3</b>	6.3	12.7
Sel.	<b>16.8</b>	<b>20.1</b>	<b>15.4</b>	17.2
Aff.	7.6	<b>24.1</b>	<b>21.1</b>	18.2
Pro.	8.4	8.5	<b>24.4</b>	15.3
Inf.	6.5	6.5	<b>18.5</b>	11.7
Oth.	2.3	2.5	2.8	2.6

Table 1: The ratio (%) of support strategies in our test sets. Each test set  $D_t$  is composed with samples corresponding to each stage. The highlighted strategies are primarily utilized in each stage (Liu et al., 2021).

실제 사용된 단계

언어모델: 긍정과 안심은 선호함(Affirmation and Ressurance)

8가지 언어모델에 테스트함

첫번째 단계는 question전략이 많이쓰이지만 언어모델들은 question전략을 거의 선택을 안해서,

첫번째 단계가 안되서 다음단계가 제대로 되지않음.

3단계 알아가기 -> 안심시키기 -> 조언하기

실제 전문가는 질문하기 전략을 많이 사용하지만

언어모델은 첫번째 단계에서 질문하기를 잘 선호하지 않는다(빨간색바-질문하기. -> 문제발생

# Problem Visualisation

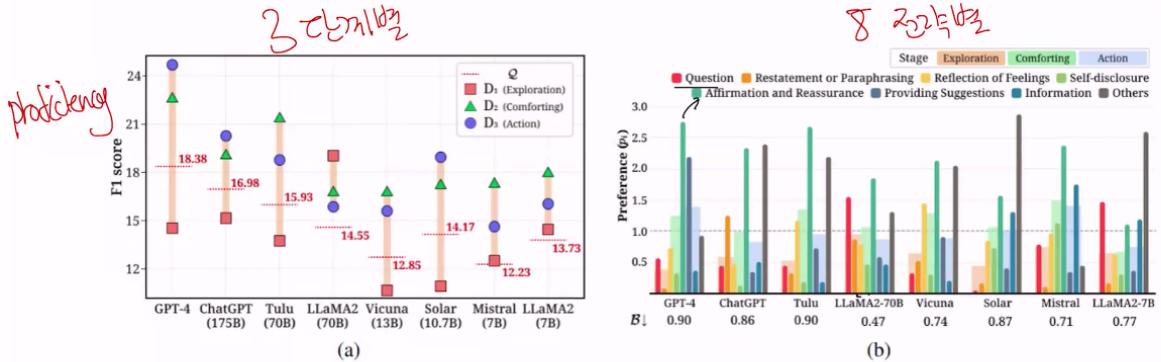


Figure 3: The details of LLMs' proficiency and preference. (a) The results of the weighted F1 score on each test set  $D_t$ , where the red dashed line indicates the proficiency  $\mathcal{Q}$  for the entire test set  $D$ . (b) The preference ( $p_i$ ) for each strategy, where the gray dashed line ( $p_i = 1$ ) represents the threshold for preferring or not preferring the respective strategy, the average preference of strategies belonging to each stage, and the preference bias  $\mathcal{B}$  below each LLM.

16 of 33

## Methodology

### •방법 1: 자가교정

•먼저 LLM이 주어진 Sample에 대해 사용할 전략을 예측함

•이러한 예측을 LLM이 스스로 교정하는 것

•예시:

<대화 문맥>

•이전에 예측한 <전략>이 적절한지 고려하고 필요시 수정하시오 -> 필요한만큼 반복 (이 질문이 맞는거야?)

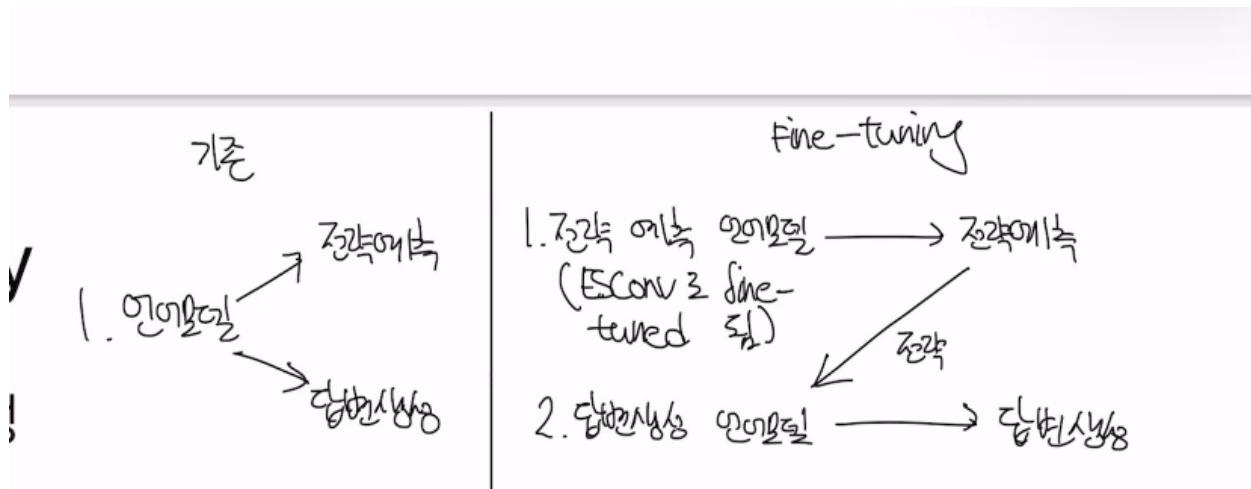
### •방법 2: 외부 교정 (외부 dataset)

1. 상식데이터셋 이용 - LLM에 추가로, 외부 정보들 (상식 데이터셋 등)을 사용해서 적절한 전략을 예측하는 것

2. Fine tuning - 아래그림 (외부교정)- 기존 언어모델은 스스로 대화전략을 예측하고 그 예측을 바탕으로 답변을 생성했다.

그런데 새로운 fine-tuning 방법에서는 임의로 선택한 LAMMA model 을 ESConv로 finetuning하고 이모델하고

대화전략을 보다 더 잘 예측하도록 전문성을 분업해서 특화함 1. 전략예측언어모델과 2 답변생성 언어모델이 전략예측을 바탕으로 분업(전략예측에 관여안함 이미 전략예측한것으로 답변생성)



• 다른 LLM (LLaMAS2-7B)를 Fine-tune(주어진 ESConv 를 사용해서 학습해서 세부조정해서 전략에 최적화- 전문성부여) 해서, 주어진 Sample(문맥)에 대한 적절한 전략을 독립적으로 예측하도록 하기도 함

• 또한 Prompt(예. 챗지피티에 메세지 넣는것)에 주어진 전략들 예시 (LLM이 보다 잘 예측할 수 있도록 기본 2개가 주어짐) 수를 늘리는 것도 실험함  
-> 원래는 8개 전략 설명만 해주고 선택하도록 했다면, 이제는 각 전략에 대한 실제 예시 넣어 주는것.(실제 전문가가 사용한것)

### 3. 예시를 늘리는 방법

주어진 문맥에 대해 가장 적절한 전략을 선택하도록 지시했다. 이때 언어모델의 이해를 돋기 위해 8개 전략 설명 추가로

실제 전문가들의 어떤 문맥에서 어떤 전략을 썼는지 예시를 기본 2개 넣어주었다. 이때 예시 2 개가 부족해서 예시로 들어가는 예시를 늘리기로함. -> 이상황에서 기대되는 효과: 언어모델이 예측 성능 향상

<문맥> ~ 대화내용, 링크, ... ] 문맥

1. Question 조작은 ...

2.

3.

4.

8번 조작 설명

<이시 문맥 1> 아래는 Question 조작을 써라  
< " 2> 아래는 Restatement

2번 예제  
(기본)

주어진 문맥에 대해 가능  
조작하는 전략을 선택하시오

기여 가능

추가 예시를  
놓아서 전략 예측이  
정확성되는지 실험

아래 글

테스트 데이터셋

(데이터셋 만드는 과정)

• 테스트 데이터는 ESConv를 바탕으로 아래와 같이 만들어졌다

1. 각 대화를 랜덤하게 5-15개의 Turns들로 나눈다 (-> Sample)

2. 각 Sample을 3개 중 하나의 단계로 분류한다

3. 또한 각 Sample의 전략을 분류한다

2번째와 3번째는 전문가가 직접 분류함

아래그림)

sample이 주어졌을 때 전략을 예측하는 것

Sample (실제 사용)	Q "I feel so frustrated"	A "May I ask"	Q "I feel ..."	A "may I ask..."
			"My school was"	"I understand"
질문		Question		Self-disclosure

- 이러한 자가교정과 외부교정 방법들을 모두 테스트하고, 주어진 테스트 셋을 바탕으로 Automatic Test를 진행

Proficiency, Preference Bias(전략 예측한 평가) – 전략(question self disclosure) 등이 맞나 평가

BLUE-2, ROGUE-L(답변 만든거와 실제 답변간 평가)

- 전문가의 답변과 모델의 답변이 얼마나 비슷한지 평가

- 이때 총 두개의 LLM을 대상으로 실험을 진행

1. ChatGPT (Representative of closed source LLM) – 유료중하나

2. LLaMA2-70B (Representative of open source LLM) - 무료중하나

자가교정 - 효과없었고 악영향미침

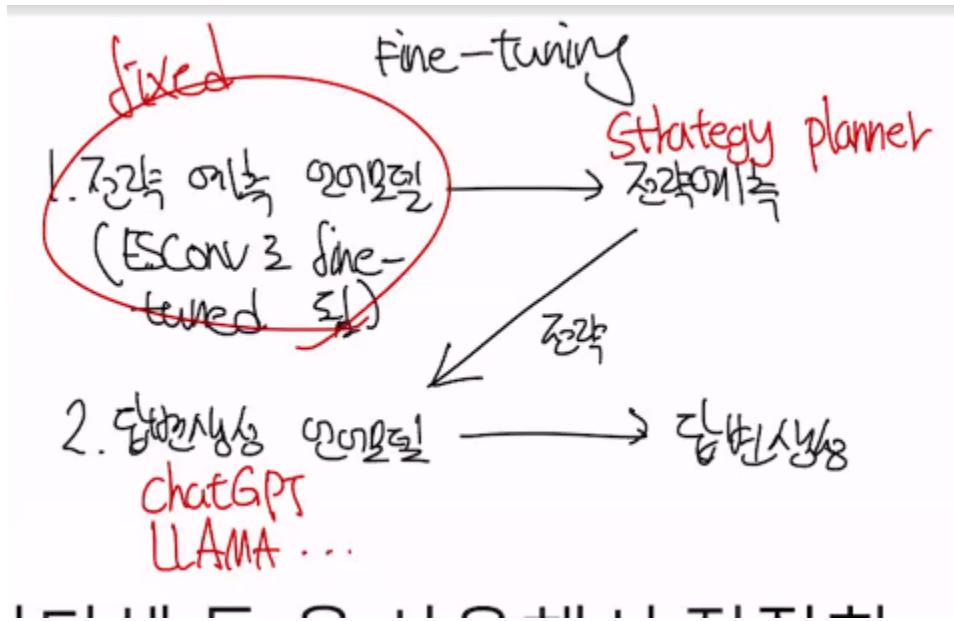
(너가 예측한건데 맞는 예측이야? 필요하다면 너가 고쳐봐)

proficiency는 떨어지고, preference bias(편향)는 높아짐(낮아져야 좋은 것)

이유: 가지고 있던 편견과 문제점을 계속 증폭

--> 그래서 사람의 외부data가 필수적

아래그림) fine tuning 한 결과가 더 나은 결과



- 그에 반해 외부 교정은 비슷하거나 훨씬 나은 결과를 보여줌

proficiency preference bias

chat gpt : 13.5 -> 21.09 1.38-> 0.36

LLM2-70B(2-shot) : 14.55 -> 21.09 0.47 -> 0.36

아래그림)

proficiency(F1 score), preference bias -> 대화 8가지 전략예측(question, restatement..) 평가-> 그중 현재대화에 제일 적절한것이 뭔지 평가

BLEU, ROGUE -> 전략으로 실제 만들어진 대답이 전문가의 대화와 얼마나 유사한지 평가

결론)

### 자가교정 - 효과없었고 악영향미침

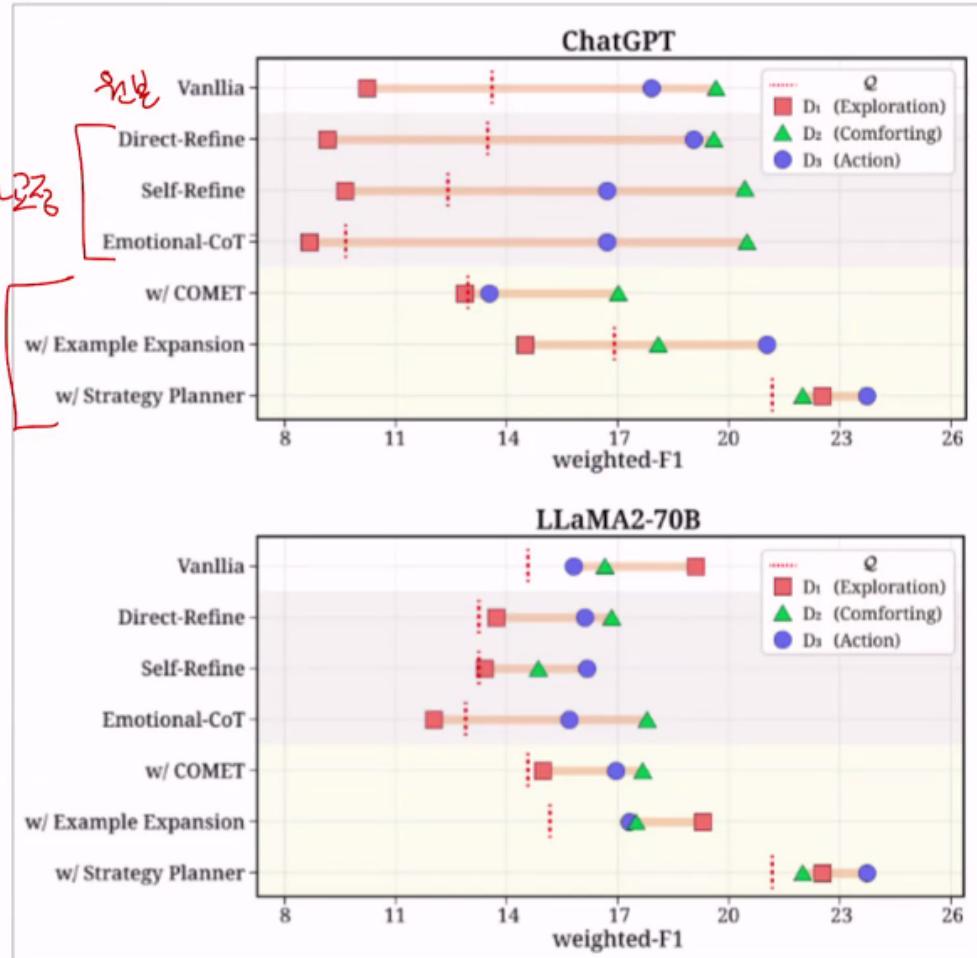
그에 반해 외부 교정은 비슷하거나 훨씬 나은 결과를 보여줄

Methods		$Q \uparrow$	$B \downarrow$	BLEU	ROUGE
				B-2	R-L
Self	ChatGPT ( <i>0-shot</i> )	13.50	1.38	6.27	14.86
	+ Direct-Refine	13.40	1.60	5.68	14.50
	+ Self-Refine	12.37	1.53	5.16	14.33
External	+ Emotional-CoT	9.55	1.56	5.23	14.12
	+ w/ COMET	12.78	0.95	6.71	<u>15.07</u>
	+ w/ Example Expansion	<u>16.91</u>	<u>0.82</u>	<b>7.45</b>	<b>15.22</b>
Self	+ w/ Strategy Planner	<u>21.09</u>	<u>0.36</u>	<u>6.96</u>	14.91
	LLaMA2-70B ( <i>2-shot</i> )	14.55	0.47	6.15	14.29
	+ Direct-Refine	13.17	0.59	5.59	13.98
External	+ Self-Refine	13.15	0.55	5.56	13.70
	+ Emotional-CoT	12.73	0.53	6.37	13.87
	+ w/ COMET	14.53	0.51	6.21	14.55
Self	+ w/ Example Expansion	<u>15.14</u>	<u>0.44</u>	<b>6.56</b>	<b>14.66</b>
	+ w/ Strategy Planner	<b>21.09</b>	<b>0.36</b>	<u>6.44</u>	<u>14.49</u>

Table 2: The results of methods on automatic metrics including  $Q$ ,  $B$ , BLEU-2 (B-2) and ROUGE-L (R-L) for the entire test set ( $D$ ). A single strategy planner is employed to predict strategies and provides them to each LLM. The best results of each LLMs are **bolded** and the second best are underlined.

아래그림)

외부교정이 전략예측에서 더 효과적으로 나타났다.

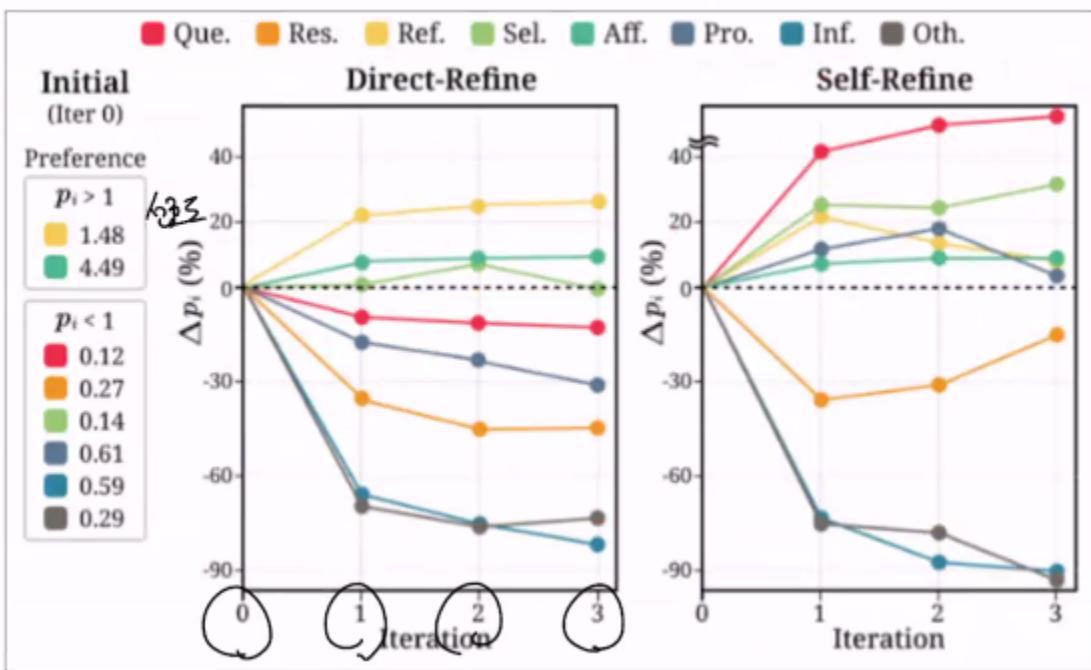


아래그림)

- 자가 교정을 1~3회 실시하고 결과를 기록한 결과, 원래부터 선호하던 전략을 더욱 선호하게 되었음

- 반대로, 원래 선호하지 않던 전략은 더욱 선택하지 않게 됨

LLM 자체적으로 문제를 수정하는  
것에는 한계가 있음



## 실험결과

- 그에 반해 외부 교정은 비슷하거나 보다 나은 결과를 보여줌
- 상식 데이터셋 (COMET)은 비슷한 결과
- 독립적인 전략 예측 (Strategy Planning)은 크게 증가한 Proficiency와 감소한 Preference Bias를 보여줌
- 비슷하게, 전략 예시 추가 (Example Expansion) 또한 기본 모델보다 나은 성과를 보여줌

아래그림)

- 독립적 전략 예측에서, 연구자들은 Fine-tuned 된 LLaMA2-7B 모델을 사용함. 이에 추가적으로, 다른 언어 모델들을 동일하게 Fine-tune 하고 비교해 보았음
  - 다른 (Encoder-based) 모델들은 비슷한 Proficiency를 보여주었지만, 상대적으로 높은 Preference Bias를 가짐
- 다양한 언어 모델을 실험하는 것 (후속 연구) 가 필요

Base Models	$\mathcal{Q} \uparrow$	$\mathcal{B} \downarrow$	$D_1$	$D_2$	$D_3$
			weighted-F1		
BERT	18.02	0.50	18.17	22.68	19.25
RoBERTa	21.01	0.60	21.34	24.18	22.99
Mistral	21.89	0.45	22.61	23.57	24.59
LLaMA2-7B	21.10	0.36	22.59	21.85	23.77

Table 3: The results on the strategies selected by different strategy planners. Each model is fine-tuned with a uniform dataset across strategies.

- 이때 예시의 수를 늘리는 것이 보다 정확한 전략 예측에 도움이 되었음
- 다만 예시의 개수가 많아질수록, Proficiency는 비슷한 수준에서 유지되지만 Preference Bias는 다시 증가하는 경향을 보임

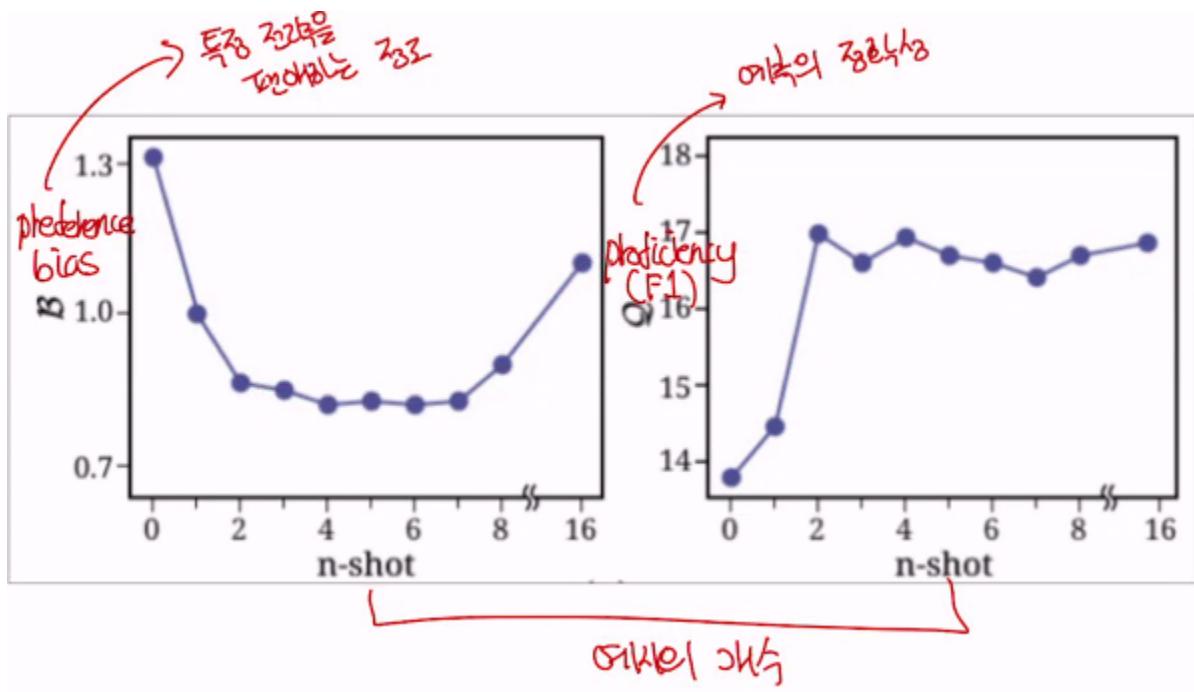
아래그림

- 이는 많은 예시가 오히려 LLM을 혼란스럽게 하는 것으로 추정됨

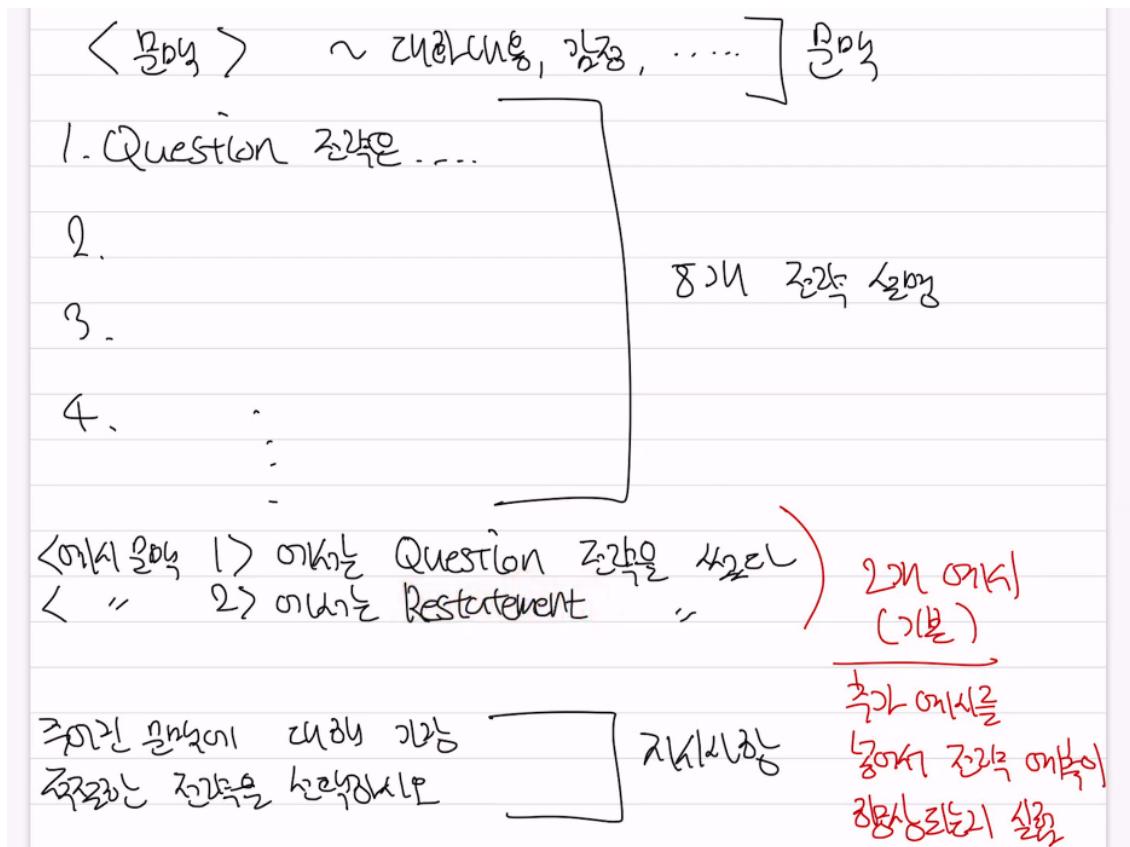
• 적당한 수준의 예시를 제공하는 것이 효율적 (2~3개, 4~5개정도)

• B: Preference Bias

Q: Proficiency



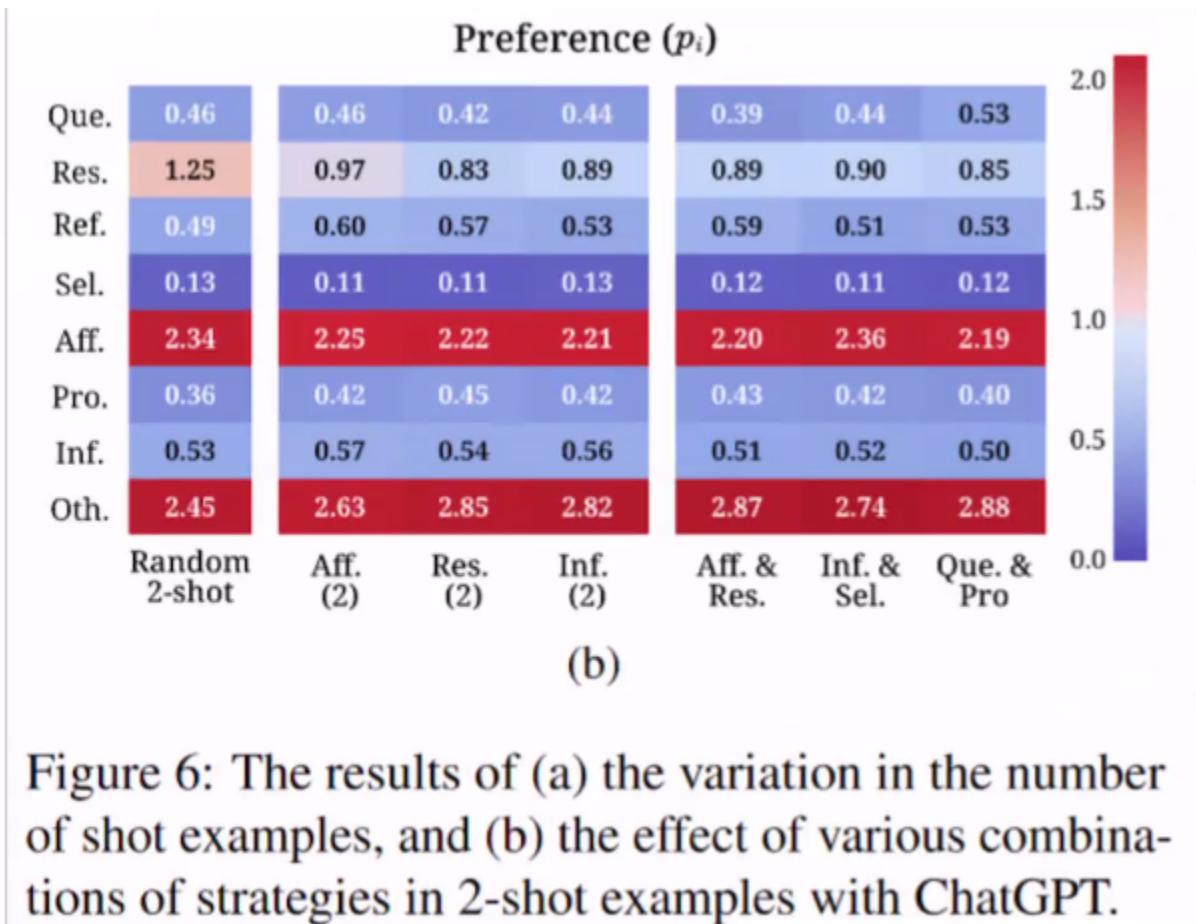
아래그림) 원래 예시를 뽑을때 랜덤하게 뽑았는데, 지정해서 전략을 뽑으면 어떨까를 실험



아래그림) Chatgpt 2개 전략 뽑을때 어떤전략인지 성능에 큰영향이 없었음.

- 또한 이때 8개의 전략 중 어떤 전략의 예시가 포함되는지는 성능에 큰 영향이 없었음

- 다시 말하자면, 어떠한 전략 A의 예시가 두개 들어갔을 때, LLM이 A를 보다 잘 예측하거나 하지 않는 않았음



**Figure 6:** The results of (a) the variation in the number of shot examples, and (b) the effect of various combinations of strategies in 2-shot examples with ChatGPT.

결론)

언어모델이 편향적으로 대화전략을 예측하는 경향이 있어서  
외부교정(예시를 더 넣어주거나, 전략만 예측하는 언어모델을 따로 finetuning하거나)해서  
훨씬 언어모델이 더 전략을 잘 예측하고,  
예측된 결과 가지고 대답도 더 상황에 맞게 잘하게 됨을 확인

->

지금까지는 테스트 데이터셋에 대한  
사람이 평가하지 않고 Automatic Evaluation 을 가지고 실험을 진행했음

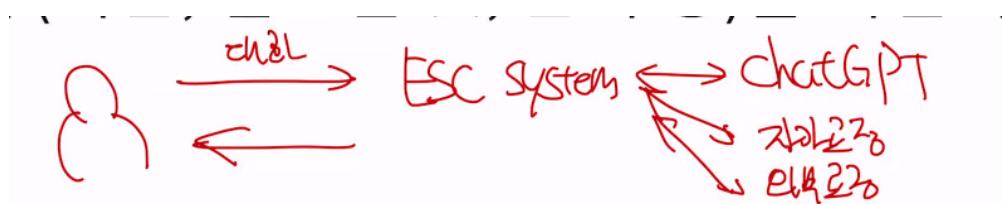
외부 교정이 이러한 점수 (Proficiency, Preference Bias) 를 개선시키는 것에 도움이 되는 것을 발견했음

- 하지만 이러한 Automatic Evaluation의 개선이 실제 유저의 높은 만족도로 이어질지는 아직 까지는 알 수 없음
- > 실제 전문가들의 평가가 필요함

## Human evaluation

- 실제 정신과 의사들과 협업해서 평가를 진행함
- 평가 기준은 사용자의 만족도이며, 세가지 세부 기준이 있음
- Acceptance: 사용자가 불편함을 느끼지 않았는가?
- Effectiveness: 사용자의 부정적인 감정을 효과적으로 긍정적인 방향으로 바꾸었는가?
- Sensitivity: 사용자의 현재 상태 (기분, 필요한 것, 문화 등)을 적절히 고려하였는가?
- 이를 바탕으로 외부 교정과 자가 교정의 평가를 진행 (ChatGPT 사용)

아래그림)



- 기본 LLM (Vanila)와 비교했을 때, 자가 교정과 COMET (상식 데이터셋)은 비슷한 증가폭을 보임
- 전략 예시를 추가하는 것이 조금 더 나은 결과를 보임
- 마지막으로, 독립적인 전략 예측이 가장 나은 성과를 보여주었음

<b>ChatGPT</b>	Acc.	Eff.	Sen.	Sat.
Vanilla	27.9	23.5	22.1	24.5
Tie	20.6	32.4	22.1	25.0
+ Self-Refine	<b>51.5<sup>‡</sup></b>	<b>44.1<sup>‡</sup></b>	<b>55.9<sup>‡</sup></b>	<b>50.5<sup>‡</sup></b>
Vanilla	22.9	24.0	14.6	20.5
Tie	21.9	33.3	27.1	27.4
+ w/ COMET	<b>55.2<sup>‡</sup></b>	<b>42.7<sup>†</sup></b>	<b>58.3<sup>‡</sup></b>	<b>52.1<sup>‡</sup></b>
Vanilla	13.1	25.3	16.2	18.2
Tie	26.3	26.3	21.2	24.6
+ w/ Example Expansion	<b>60.6<sup>‡</sup></b>	<b>48.5<sup>†</sup></b>	<b>62.6<sup>‡</sup></b>	<b>57.2<sup>‡</sup></b>
Vanilla	16.7	29.2	29.2	25.0
Tie	12.5	16.7	12.5	13.9
+ w/ Strategy Planner	<b>70.8<sup>‡</sup></b>	<b>54.2<sup>‡</sup></b>	<b>58.3<sup>‡</sup></b>	<b>61.1<sup>‡</sup></b>

Table 4: The results of comparative human evaluation between various methods applied to ChatGPT and vanilla ChatGPT. ( $\dagger/\ddagger$ : p-value < 0.1/0.05 )

## 아래그림

- ESC 시스템에서는 좋은 대답을 하는 것 만큼이나 잘못된 대답을 하지 않는 것도 중요함
  - 이때 자가 교정 방법은 잘못된 대답의 비율이 오히려 더 높아졌음
  - 반대로 독립적인 전략 예측 방법은 잘못된 대답의 비율이 크게 낮아짐
  - Oracle Strategy (정답 전략 사용시)

### 아래그림

3점 미만일때 부적절한 대답, 3점 이상일때 만족

### 사람의 평가(0에서 10점사이)

자가교정: 16.7  $\Rightarrow$  21.2, 17.4 (올랐음)

외부교정: 16.7 → 8 (낮아짐)

정답전략: 16.7 → 3.8 (낮아짐) → 무조건 나오는 에러퍼센트 (할거 다해더 3.8% 나옴)

점수

Methods	< 3 (fail)	$\geq 3$ (acceptable)
ChatGPT	16.7	83.3
지어고23 [+ Direct-Refine]	21.2	78.8
+ Self-Refine	17.4	82.6
인력고23 [+ w/ Strategy planner]	8.0	92.0
정답고객 [+ Oracle Strategy]	3.8	96.2

Table 5: The ratio (%) of scores below 3 (fail) and scores of 3 or above (acceptable) in Seeker's Satisfaction (*Sat.*).

## Discussion

- LLM을 ESC 시스템에 적용했을 때, 적절한 대화 전략을 선택하지 못하는 경우가 많음
- LLM은 특정 전략을 선호하고, 이는 전반적인 ESC의 질을 떨어트림
- 이러한 문제는 외부 교정을 통해 완화될 수 있음
- 특히 Fine-tuned 된 독립적인 대화 전략 예측 모델이 유효함
- 이러한 방법을 통해 보다 효과적인 ESC 시스템을 만들 수 있을 것으로 예상됨

## Limitations

- 대화 전략 “Others”는 이전 연구에서 제안된 것을 그대로 가져온 것이지만, 실제로는 별로 효과적이지 않았으며 다른 전략으로 구분될 필요성이 있음
- Human Evaluation에서, 완벽한 정답 전략을 사용해도, 사용자의 감정을 더욱 격화시키는 답변을 만드는 경우가 있었음 (3.8%)
- 대부분의 경우에는 LLM은 선택된 전략에 잘 맞는 답변을 했지만, 때때로 Misalignment가 발생함 (질문 전략인데 설명을 한다던지) -> 개선방법: 다른언어모델에 검증필요

2



...

구독하기

'논문 - 아주대(조현석교수님)' 카테고리의 다른 글

아주대 - GTA Gated Toxicity Avoidance for LM Performance Preservation 논문리뷰 (2)

2024.09.18

## 관련글

[관련글 더보기](#)

Method	Accuracy (↑)	Toxicity (↓)	Fluency (↑)
GPT-3	76.15	4.23	3.62
<b>아주대 - GTA Gated Toxicity Avoidance for LM Performance Preservation</b>	<b>65.77</b>	<b>1.15</b>	<b>3.73</b>
GeDi <sub>GTA</sub>	75.77	1.15	3.57
DisCup	70.77	1.54	3.56
DisGPT	76.54	0.60	3.60

## 자연어(NLP)

네이쳐2024 님의 블로그입니다.

구독하기 +



## 댓글 0



이름

비밀번호

내용을 입력하세요.



등록