



*Fine.*

*Fine[피네]: 하루의 끝*

브레멘 음악대 (NLP-05)

변성훈 / 서보성 / 이상민 / 이승우 / 이예원

# INDEX

## 1. Introduction

- Motivation
  - Introduction
  - Timeline
  - Roles
- 

## 2. Models

- Model Pipeline
  - ChatBot
  - Diary Generation
- 

## 3. Serving

- System Architecture
  - Front-End
  - Back-End
- 

## 4. Result

- Demo Video
- Conclusion

---

# 1. Introduction

*Fine.*

- 
- Motivation
  - Introduction
  - Timeline
  - Roles

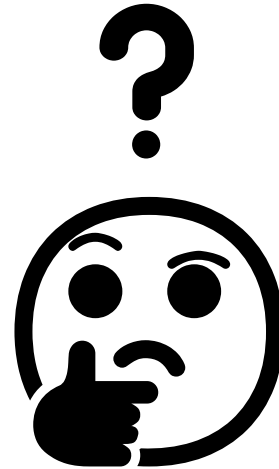
## Motivation

그 때 갔던 맛집 이름이 뭐였지?

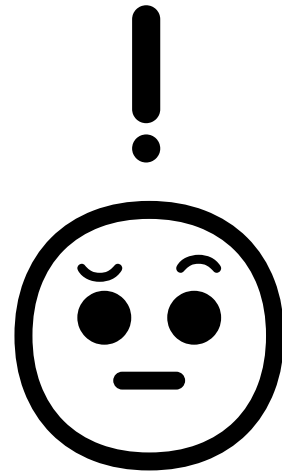
누구랑 갔더라?

그 때 들었던 노래 제목이...

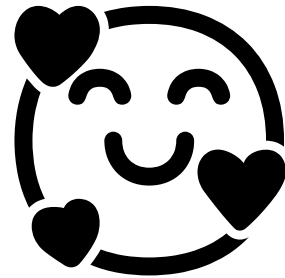
그 때 내가 느꼈던 감정은?



일기를 쓰면 될 텐데!

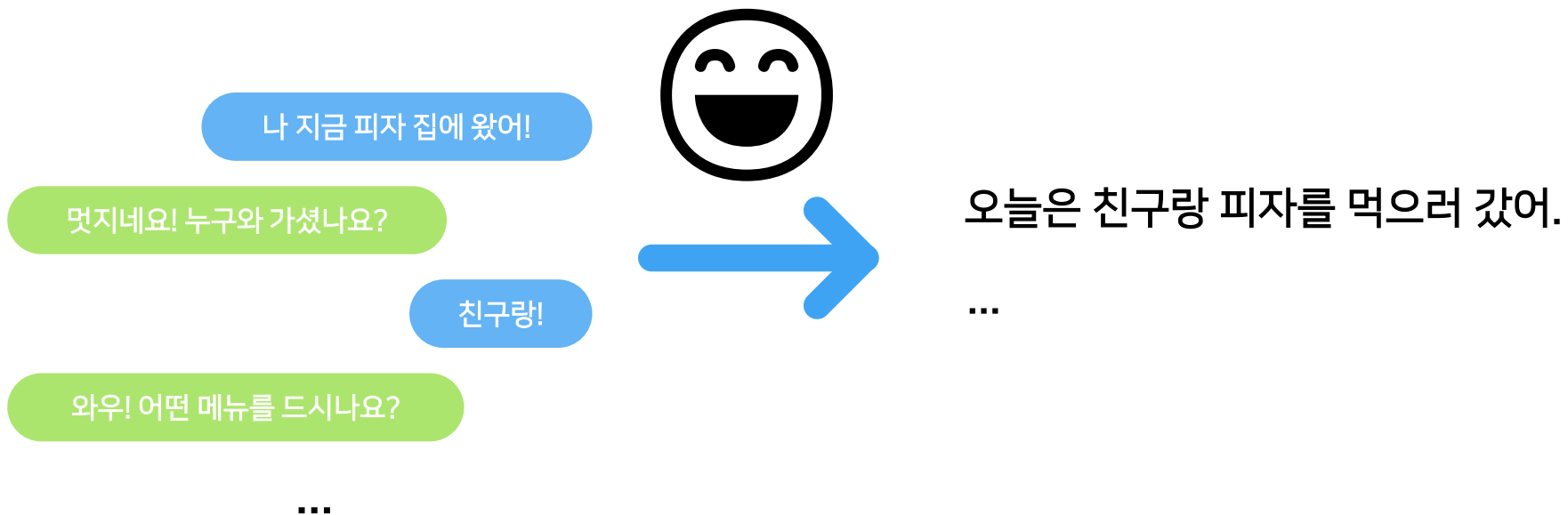


“친구와 하루 종일 수다 떨듯 오늘 일을 말하면,  
일기를 쓰듯 오늘 일을 기록해 줄 수는 없을까?”

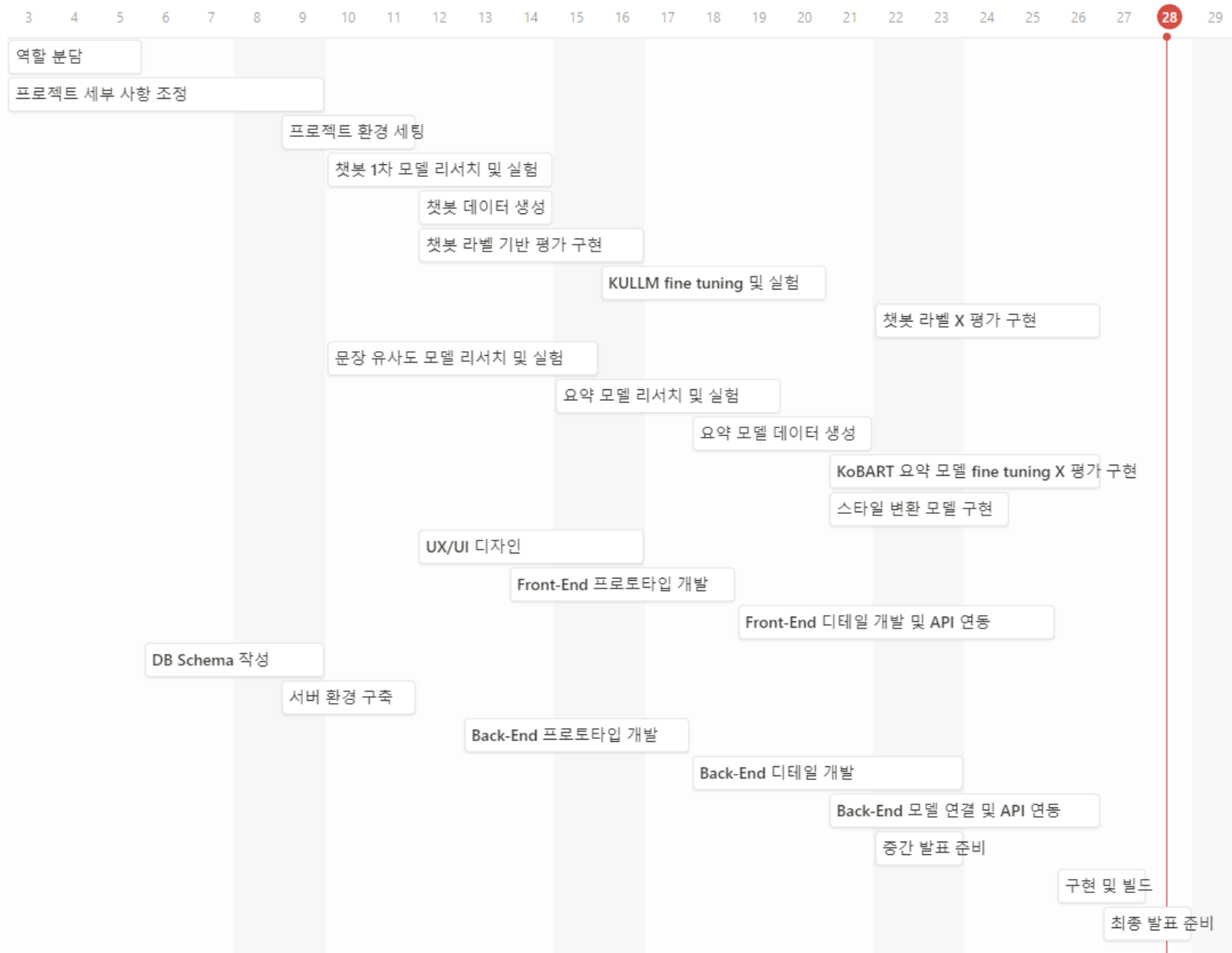


*Fine.*

하루 동안 챗봇과 대화하면, 하루의 끝에 일기를 생성해주는 서비스



# Timeline





## Roles



 변성훈

- Generation Style Transfer Model
- Back-End



 서보성

- Back-End
- Product Serving



 이상민


- Summarization Model



 이승우

- Product Manager
- Front-End



 이예원

- ChatBot Model

---

# 2. Models

*Fine.*

- 
- Model Pipeline
  - ChatBot
  - Diary Generation

## Model Pipeline

  : ChatBot



오늘 친구랑 피자를 먹었어!



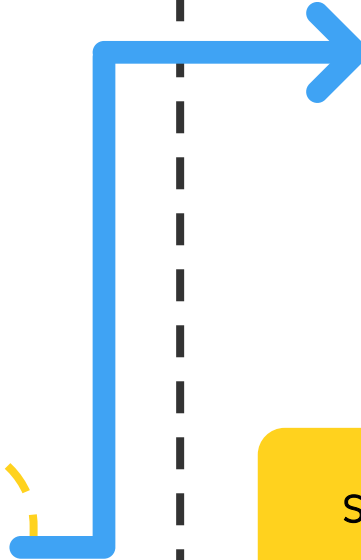
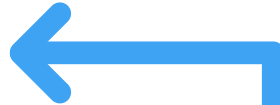
ChatBot  
Model



좋아요! 피자를 먹는 건 항상 즐거운  
일이죠. 어떤 종류의 피자를 먹었나요?



포테이토 피자를 먹었어! 생각보다 맛은 없었지만  
친구랑 같이 먹어서 좋았어.



 : Diary Generation

오늘 친구랑 피자를 먹었어!

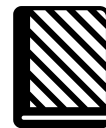
좋아요! 피자를 먹는 건 항상 즐거운  
일이죠. 어떤 종류의 피자를 먹었나요?

...

다음에도 또 도림천을 산책하고 싶어!



Summarization Model + Generation Style Transfer Model



오늘 친구랑 피자를 먹었는데,  
생각보다 맛은 없었지만 친구랑 같이 먹어서 좋았어.  
...

공감 능력을 갖춘 시스템의 적절한 리액션

자연스러운 질문을 통한 오늘 하루의 정보 유도

일기 생성이라는 목적 지향성 대화



Data Generation

## ChatBot

### Dataset Generation Prompt

사용자와 시스템 간의 대화 데이터가 필요해.

**시스템:** 안녕하세요! 오늘 하루는 어떠셨나요?  
무슨 재미있는 일이 있었나요?

**사용자:** 오늘 친구랑 피자를 먹었어!

**시스템:** 와, 그렇군요! 친구랑 피자를 먹는 건 항상 즐겁죠! 어떤 피자를 주문했나요?

**사용자:** 포테이토 피자를 먹었어! 생각보다 맛은 없었지만 친구랑 같이 먹어서 좋았어.

**시스템:** 포테이토 피자도 한 번쯤 시도해볼만 하죠!  
그래도 친구랑 함께하는 시간이 더 중요하니까요.  
먹으면서 어떤 재미있는 주제로 이야기를 나눴어요?

위 예시 같은 형식으로 다른 주제로 대화 데이터를 만들어줘. 대화의 시작은 시스템이 오늘 하루에 대한 질문을 하면서 시작해야 해.

### Prompt

다른 주제로 대화 데이터를 만들어줘. 대화의 시작은 시스템이 오늘 하루에 대한 질문을 하면서 시작해야 해.

gpt-3.5-turbo

Data 생성: 해당 flow로 5개씩, 총 600개

Data 정제:

2개 이상의 data 생성 시 분리  
시스템-사용자 순서의 대화가 아닌 경우 제거  
→ 총 613개

Chat Data

## ChatBot

Dataset



안녕하세요! 오늘 하루는 어떠셨나요? 무슨 재미있는 일이 있었나요?

오늘 공원에서 산책을 했어요.



멋지네요! 공원에서 산책은 마음을 편안하게 해주는 활동이죠. 무슨 경로로 산책하셨나요?

오늘은 공원 일주로를 걸으며 동물들을 관찰했어요.



그렇군요! 동물들을 관찰하면서 어떤 동물을 봤나요?

특히 새들을 많이 봤어요. 각기 다른 종류의 새들이 날으며 알록달록한 깃털이 아름다웠어요.



정말 멋진 경험이었겠네요! 새들의 아름다운 깃털은 참으로 매력적이에요. 산책하며 어떤 생각을 했나요?

새들의 자유로움을 보면서 나도 더 많은 자유를 누리고 싶다는 생각이 들었어요. 오늘부터 조금 더 자유롭게 살아야겠다고 다짐했어요.



멋진 다짐이에요! 자유롭게 살기 위해 노력하는 것은 정말 좋은 선택이에요. 앞으로 더욱 자유롭고 행복한 일들이 가득하길 바랄게요!

	train_sentence_len	train_label_len
count	1696,000000	1696,000000
mean	95,508844	34,052476
std	59,205098	11,548700
min	21,000000	11,000000
25%	31,750000	25,000000
50%	83,000000	33,000000
75%	143,000000	42,000000
max	303,000000	77,000000

	valid_sentence_len	valid_label_len
count	340,000000	340,000000
mean	95,411765	34,414706
std	58,283064	12,442970
min	23,000000	10,000000
25%	32,750000	25,000000
50%	82,000000	33,000000
75%	146,000000	43,000000
max	256,000000	87,000000

gogamza/kobart-base-v2 tokenizer 기준 length

## Input Sentence

최대 3 turn의 이전 대화 히스토리

## Label

Input sentence 다음으로 올  
알맞은 system 응답

→ train 1696개, test 340개

시스템의 응답이 너무 단조로워 최종적으로는 사용하지 않은 "AI HUB 감성 대화 데이터"로 학습한 결과 예시



거의 유사하지만 답변의 길이와 문맥 이해도를 고려해 이후 KoBART로 테스트

## Reference-Based

오늘 공원에서 산책을 했어요.



멋지네요! 공원에서 산책은 마음을 편안하게 해주는 활동이죠. 무슨 경로로 산책하셨나요?

대부분의 응답 구성

간단한 리액션 + 유저 응답 재서술 + 맥락에 맞는 질문



해당 구성의 응답으로 학습하면서 필수적으로  
정답 label과 어느정도의 유사도를 가짐



정답 label 기반 평가

## Reference-Free

다양한 질문과 리액션의 유형 가능



하나의 정답 label 존재 X



label 없이 평가



**ROUGE Score** target과 얼마나 겹치는 지에 관한 지수

**ROUGE-1** tokenizer → Mecab

Recall = target과 prediction의 공통 unigram의 수 / target 의 unigram 수

Precision = prediction에서 target의 공통 unigram의 수 / prediction의 unigram 수

$$F1\ Score = 2 \times \frac{recall \times precision}{recall + precision}$$

**ROUGE-SU** max skip → 2

**Skip Bigram + Unigram**

(ex) target: the brown for jumps → [the, brown, for, jumps, (the, brown),  
(the, for), (brown, for), (brown, jumps), (for, jumps)]

prediction: the fox → [the, fox, (the, fox)]

→ Recall = 1/9, Precision = 1/4, F1 = 0.1538

## Cosine Similarity

korSTS dataset으로 학습한 SBERT

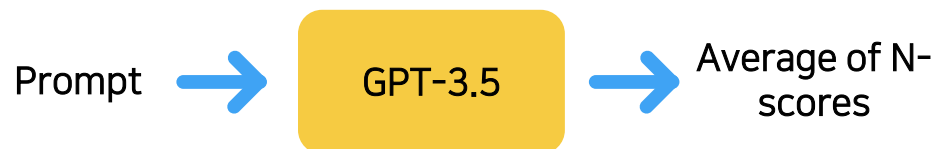
→ prediction과 target의 embedding vector의 cosine similarity

## Perplexity

문장의 길이로 정규화된 문장 확률의 역수: 모델의 **헛갈림**의 수치

$$PPL(W) = \sqrt[N]{\frac{1}{P(w_1, w_2, w_3, \dots, w_N)}} = \sqrt[N]{\frac{1}{\prod_{i=1}^N P(w_i | w_1, w_2, w_3, \dots, w_{i-1})}}$$

**G-Eval** → 생성 데이터와 비교해 얼마나 잘 학습했는지 평가 가능  
gpt-3.5-turbo 모델을 이용하여, 우측 프롬프트로 생성 모델 평가



## Prompt

시스템과 유저의 대화 내역(Conversation History)을 받게 될 것입니다. 그리고 대화 내역에 대한 시스템의 응답(Response)이 제시됩니다. 당신의 작업은 평가 단계에 따라 **응답을 평가**하는 것입니다. 이 평가 기준을 꼼꼼히 읽고 이해하는 것이 중요합니다. 평가하는 동안 이 문서를 계속 열어두고 필요할 때 참조해 주세요.

### 평가 기준:

- 이해 가능성 (0 - 1): Conversation History에 기반하여 Response를 이해 할 수 있나요?
- 맥락 유지 (1 - 3): Conversation History를 고려했을 때 Response가 맥락을 유지하나요?
- 흥미롭기 (1 - 3): Response가 지루한가요, 아니면 흥미로운가요?
- 전반적인 품질 (1 - 5): 위의 답변을 바탕으로 이 발언의 전반적인 품질에 대한 인상은 어떤가요?

### 평가 단계:

1. Conversation History와 Response를 주의깊게 읽습니다.
2. 위의 평가 기준에 따라 Response를 평가합니다.

### Conversation History:

{{Conversation History}}

### Response:

{{response}}

### Result

- 이해 가능성 (0 - 1):
- 맥락 유지 (1 - 3):
- 흥미롭기 (1 - 3):
- 전반적인 품질 (1 - 5):

# KoBART



안녕하세요! 무슨 일이 있으셨나요?

오늘 친구랑 피자를 먹었어!



좋아요! 피자를 먹으면서 친구와 함께 어떤 이야기를 나눴나요?

요즘 서로 뭐하고 지내는지 얘기했어!



좋아요! 피자를 먹으면서 친구와 함께 어떤 이야기를 나눴나요?



- ☑️ 유저의 응답에 대한 적절한 답변이 어려울 경우 이전 시스템 응답을 반복하는 경우가 잦음



문맥 파악을 못하는 근본적 문제



한국어 LLM



KULLM

ChatBot

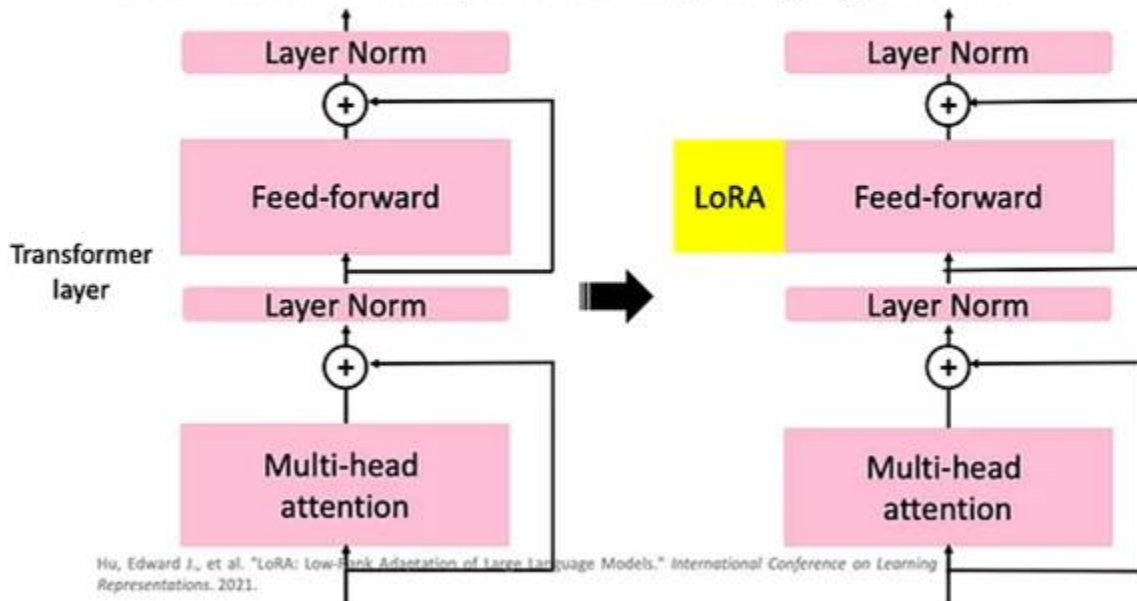
KULLM

Polyglot-ko 모델을 LoRA (Low Rank Adaptation)를 사용하여 학습한 한국어 특화 Large Language Model

Slides credit: Cheng-Han Chiang, Yung-Sung Chuang, Hung-yi Lee, "AAACL\_2022\_tutorial\_PLMs," 2022.

## Parameter-Efficient Fine-tuning: LoRA

- LoRA: Low-Rank Adaptation of Large Language Models



## PEFT (Parameter-Efficient Fine-tuning)

: 대부분의 파라미터를 프리징하고 일부의 파라미터만을 파인 튜닝

- ☑ 적은 데이터로 도메인 밖 데이터에 대해서 일반화 가능
- ☑ 사전 학습된 정보를 망각하는 Catastrophic Forgetting 방지
- ☑ 적은 수의 GPU로도 Fine-tuning 가능

<https://github.com/nlpai-lab/KULLM>

KULLM

## Next Token Prediction

DECODER

- ☑ label에 대해서만 loss 계산
- ☑ embed out과 마지막 layer를 fine-tuning  
(전체 58억 개 중 약 2억 개의 parameters)  
→ 사전 학습된 대규모의 일반화 지식 활용 기대

### Instruction

아래는 작업을 설명하는 명령어와 추가 컨텍스트를 제공하는 입력이 짝을 이루는 예제입니다. 요청을 적절히 완료하는 응답을 작성하세요.

### 명령어: 주어진 문장들은 이전 대화 내용들입니다. 이에 알맞은 시스템 응답을 만들어주세요.

### 입력: **system**: 안녕하세요! 오늘 하루는 어떠셨나요? 무슨 재미있는 일이 있었나요? **user**: 오늘은 새로운 카페에서 커피를 마셨어요!

### 응답: **system**: 좋아요! 새로운 카페에서 커피를 마시면서 어떤 기분이 들었나요? 어떤 종류의 커피를 시도해봤어요?

## KoBART

안녕하세요! 무슨 일이 있으셨나요?

오늘 친구랑 피자를 먹었어!

좋아요! 피자를 먹으면서 친구와 함께 어떤 이야기를 나눴나요?

요즘 서로 뭐하고 지내는지 얘기했어!

좋아요! 피자를 먹으면서 친구와 함께 어떤 이야기를 나눴나요?

model	이해 가능성	맥락 유지	흥미롭기	전반적인 품질
KoBART	0.997	2.937	2.264	3.406
<b>KULLM</b>	<b>0.999</b>	<b>2.957</b>	<b>2.275</b>	<b>3.459</b>
GPT	0.999	2.953	2.315	3.525

model	ROUGE-1	ROUGE-SU	Similarity	Perplexity
KoBART	43.837	29.29	78.034	1.62204
<b>KULLM</b>	<b>45.227</b>	<b>30.695</b>	<b>80.917</b>	<b>1.36967</b>

## KULLM

안녕하세요! 무슨 일이 있으셨나요?

오늘 친구랑 피자를 먹었어!

좋아요! 피자를 먹는 건 항상 즐거운 일이죠. 어떤 종류의 피자를 먹었나요?

포테이토 피자를 먹었어! 생각보다 맛은 없었지만 친구랑 같이 먹어서 좋았어.

...

다음에도 또 도림천을 산책하고 싶어!

좋아요! 도림천은 산책하기에 정말 좋은 곳이에요. 다음에도 도림천을 산책하면서 좋은 시간을 보내세요!

응!

답변의 합리성, 문맥의 이해도 등도 다음 지표와 유의미한 상관관계를 보였기에 모든 평가지표에서 높은 **KULLM** 사용

## KULLM

안녕하세요! 무슨 일이 있으셨나요?

오늘 친구랑 피자를 먹었어!

...

다음에도 또 도림천을 산책하고 싶어!

좋아요! 도림천은 산책하기에 정말 좋은 곳이에요.  
다음에도 도림천을 산책하면서 좋은 시간을 보내세요!

응!

좋아요! 도림천은 산책하기에 정말 좋은 곳이에요.  
다음에도 도림천을 산책하면서 좋은 시간을 보내세요!

## KULLM+

안녕하세요! 무슨 일이 있으셨나요?

오늘 친구랑 피자를 먹었어!

...

다음에도 또 도림천을 산책하고 싶어!

좋아요! 도림천은 산책하기에 정말 좋은 곳이에요.  
다음에도 도림천을 산책하면서 좋은 시간을 보내세요!

응!

오늘 더 특별한 일은 없으셨나요? 더 자세히  
말씀해 주세요!

학습 데이터에 답변의 길이가 짧은 경우가 없어서 이에 대한 적절한 응답을 생성해내지 못함



유저의 응답의 길이가 6보다 작을 경우 **정해진 응답**으로 다른 화제로 넘어가도록 함

## Diary Generation: Summarization Model

안녕하세요! 무슨 일이 있으셨나요?

오늘 친구랑 피자를 먹었어!

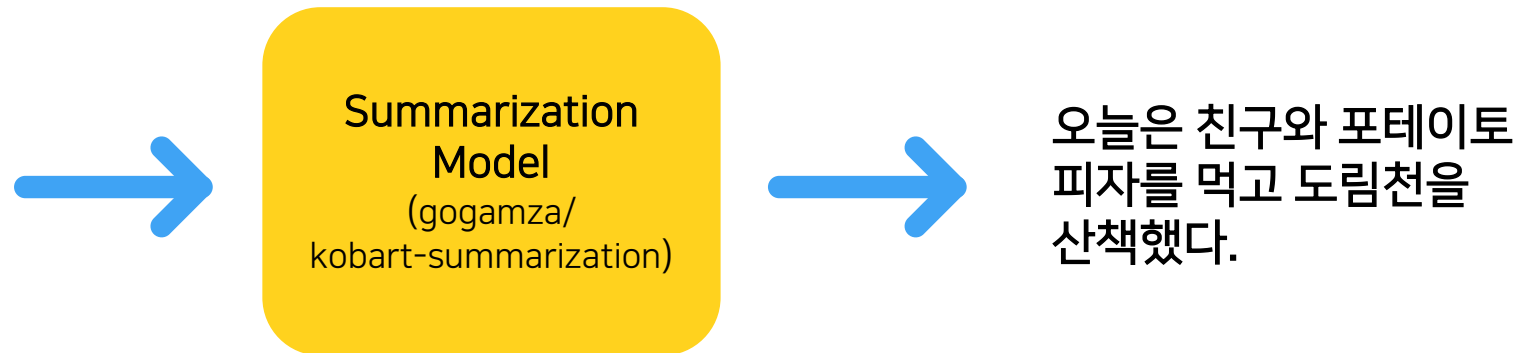
좋아요! 피자를 먹는 건 항상 즐거운 일이죠.  
어떤 종류의 피자를 먹었나요?

포테이토 피자를 먹었어! 생각보다 맛은  
없었지만 친구랑 같이 먹어서 좋았어.

...

피자를 먹고 무엇을 했나요?

피자를 먹고 도림천을 산책했는데  
날씨도 좋고 너무 즐거웠어.



오늘은 친구와 포테이토  
피자를 먹고 도림천을  
산책했다.



## Diary Generation: Summarization Model

Dataset  
Prompt

### 대화를 보고 사용자에게 대한 하루를 요약해줘

**시스템:** 안녕하세요! 오늘 하루는 어떠셨나요? 무슨 재미있는 일이 있었나요?

**사용자:** 저는 오늘 운동을 다녀왔어요!

**시스템:** 멋지게 운동하셨군요! 어떤 종류의 운동을 했나요? 즐거웠나요?

**사용자:** 저는 요가를 했어요. 요가를 하면서 몸과 마음이 편안해지는 느낌이 들어서 정말 좋았어요.

**시스템:** 요가를 하면 몸과 마음이 편안해지니까 정말 좋으시겠어요! 운동하면서 어떤 이야기를 나누거나 다른 사람들과 어떤 경험을 공유했나요?

**사용자:** 요가 수업 중에는 조용한 분위기라 대화는 많이 하지는 않았어요. 하지만 수업 뒤에는 같이 수업을 들은 사람들과 운동 후기를 공유했어요. 다들 요가를 통해 정말 많은 변화를 느끼고 있다고 이야기했어요.

ChatGPT

Summarized Data **생성:** 총 350개

Data **정제:** 비정상 데이터 제거 → 총 341개

사용자는 오늘 요가 수업에 참여하여 운동을 즐겼습니다. 요가를 하면서 몸과 마음이 편안해지는 느낌을 받아서 정말 좋았다고 했습니다. 수업 중에는 조용한 분위기로 대화를 많이 나누지는 않았지만, 수업 후에 같이 수업을 들은 사람들과 운동 후기를 공유하면서 다들 요가를 통해 많은 변화를 느끼고 있다고 이야기했습니다.

## gogamza/kobart-summarization

Chat GPT API는 Controllability 부족 + 금전적 부담  
작은 모델 크기 (약 500MB)

### Hyperparameter

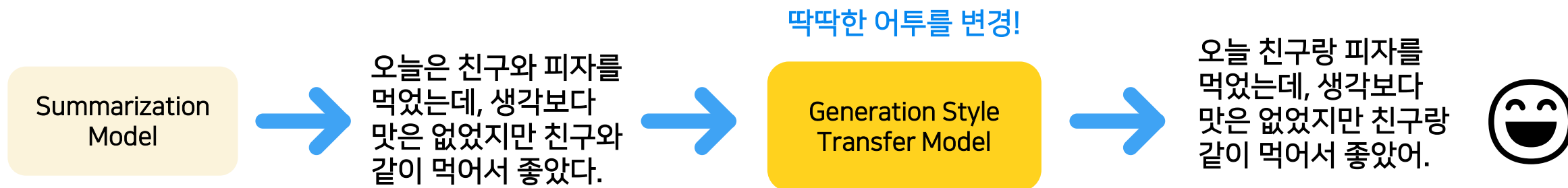
lr	batch_size	epoch	EVAL : ROUGE-1	EVAL : ROUGE-SU
1e-4	32	25	0.7076	0.4969
<b>1e-4</b>	<b>64</b>	<b>25</b>	<b>0.7208</b>	<b>0.5162</b>
3e-4	64	25	0.7215	0.5178
1e-4	64	30	0.7079	0.5074
9e-5	64	25	0.7178	0.5112
7e-6	64	25	0.6840	0.4920

### METRIC

ROUGE-1, ROUGE-SU

target과 얼마나 겹치는 지에 관한 지수

## NHNDQ/bart-speech-style-converter



Pretrained Model에 기본으로 학습되어있는 어투 예시

구어체 오늘 친구랑 피자를 먹었는데, 생각보다 맛은 없었지만 친구랑 같이 먹어서 좋았어.

안드로이드 오늘.. 친구. 피자. 먹음. 생각보다. 맛. 없음. 친구. 동반. 먹음. 좋음.

ENFP 오늘 친구랑 피자 먹었는데,, 생각보다 맛은 없었는데,, 친구랑 같이 먹어서 좋아!!

할아버지 오늘.....친구하고...피자 먹었는데... 생각보다 맛은 없었는데...친구하고...같이 먹어서 좋았던 것 같구먼...

중학생 오늘 친구랑 피자 먹었는데 생각보다 맛없었는데 친구랑 같이 먹어서 좋음

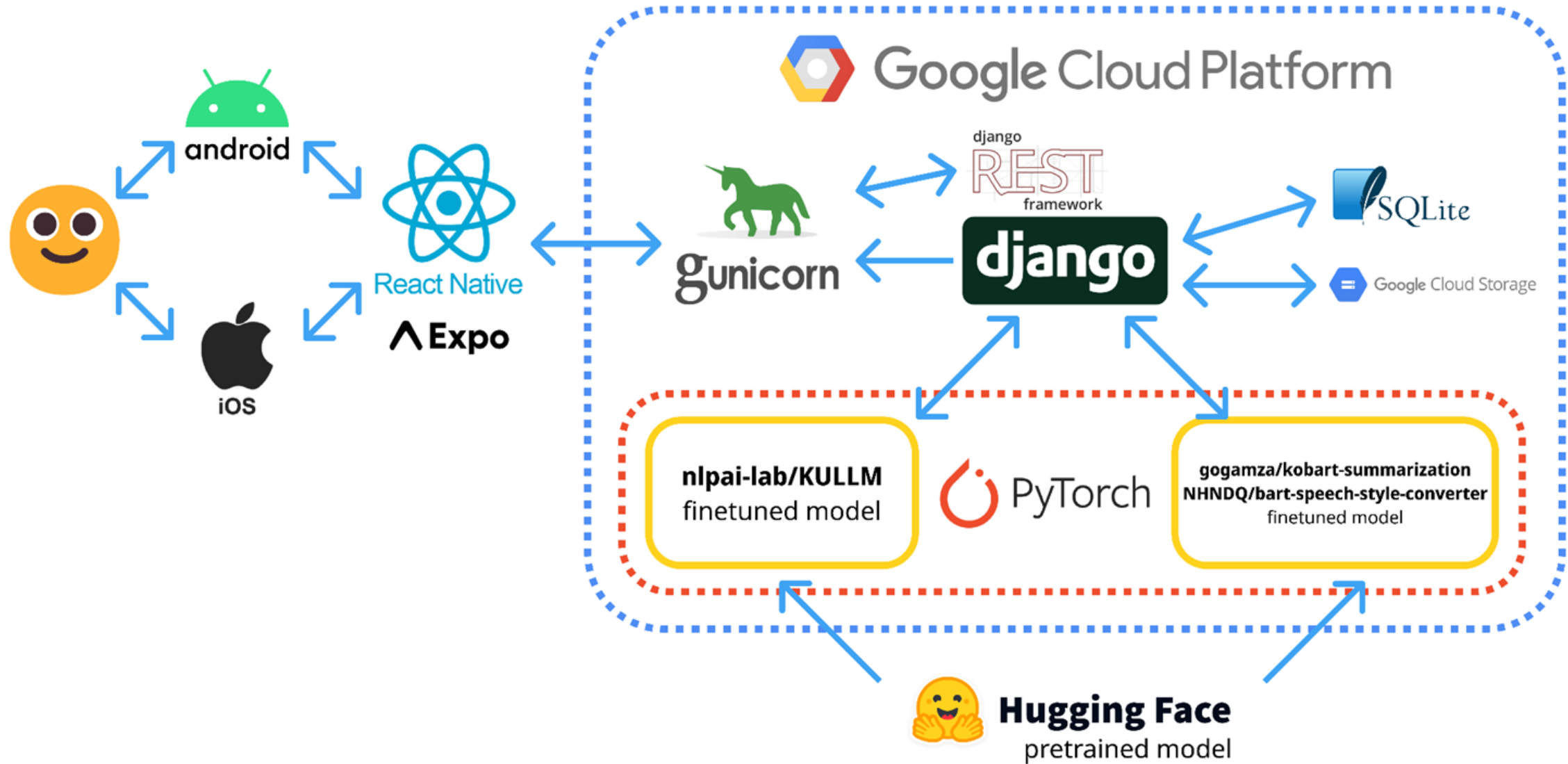
---

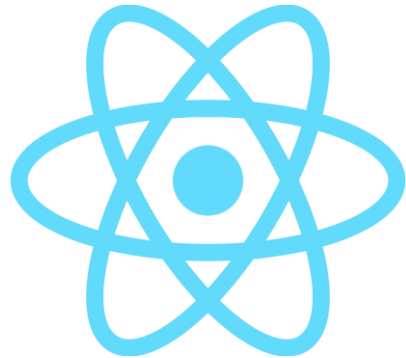
# 3. Serving

*Fine.*

- 
- System Architecture
  - Front-End
  - Back-End

## System Architecture





## React Native

Android, iOS App 개발을 위한 Cross-Platform Framework



# Front-End Design Concept



LOGO

나눔스퀘어 네오

나눔손글씨 다행체

FONTS



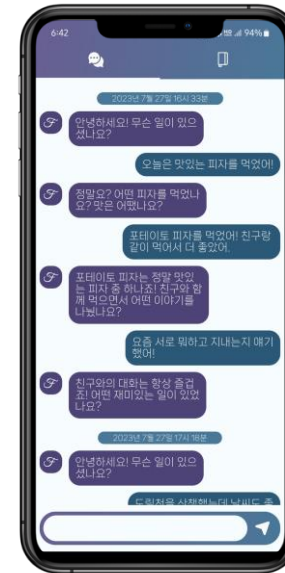
아침



낮



저녁



밤





프로그램들을 빌드하기 위한 기능들을 제공하는  
Google의 Cloud Computing 서비스



파이썬 웹 어플리케이션을 실행하기 위한 WSGI 서버로,  
안정적이고 고성능의 웹 서버를 제공

django



파이썬 웹 프레임워크인 Django를 확장하여  
RESTful API를 구축할 수 있게 해주는 라이브러리



서버가 필요 없는 경량의  
오픈 소스 데이터베이스 관리 시스템



Google Cloud에 객체(이미지)를  
저장하는 서비스



---

# 4. Result

*Fine.*

- 
- Demo Video
  - Conclusion

## Demo Video



<https://youtu.be/5fr-eOwK-7k>

## Conclusion

### 개선 및 발전 방향

#### Modeling

챗봇 데이터 생성 방법이 단순해 데이터 다양성 X  
→ 생성 flow를 재구성해서 응답 유형을 다양화

KULLM 모델 사이즈로 인해 GPU 서버 사용  
→ 유지 비용 비쌈 → **경량화** 필요

Summarization시 동일한 내용을 **중복**으로  
생성하는 경우가 있음  
(같은 내용을 두 번 이상 반복하는 것은 ROUGE  
score에 적극적으로 반영되지 않음)  
→ **평가 지표** 및 **모델 성능** 개선 필요

#### Service

대화 내용 감성 분석을 바탕으로 그 날에 어울리는 노래,  
미디어 등 **컨텐츠 추천** 기능

QA Task 적용을 통한 사용자의 문장형 질문에 알맞는  
일기 내용을 **검색**하는 기능

챗봇 및 일기 **말투 커스터마이징** 기능

음성 인식, 결제 내역, 지도 등 다양한 외부 API 연결

기타 다양한 기능 추가 및 구독 시스템 등 **수익 모델**  
기획을 통해 상업성 도모

---

End of Document  
Thank You.