



# Upstage AI Lab

## Dialogue Summarization

3조

김지완 최용빈 김도후 박주혁 정혜윤

# Content

---

01. 대회 소개

02. 수행 과정

03. 수행 결과

04. 자체 평가의견



01

# 대회 소개

# 대회 소개

: Dialogue Summarization

개요

- 일상 대화에 대한 요약 효과를 효과적으로 요약하는 생성 모델을 구축

대회 목표 및 기간

목표  
주어진 일상 대화문과 요약문 데이터셋으로부터  
정확하고 일반화된 요약문 생성 모델 구축

기간  
5월 13일 (월) 10:00 ~ 5월 27일 (월) 13:00

모델 및 평가지표

모델  
BART, T5, Gemma, Solar

평가지표  
ROUGE-1-F1, ROUGE-2-F1, ROUGE-L-F1

데이터셋  
train : 12457 / dev : 499 / test : 499

# 대회 소개

: Dialogue Summarization

---

## Metric

- ROUGE-1-F1, ROUGE-2-F1, ROUGE-L-F1

$$\begin{aligned} \text{Score} = & \frac{\sum_i^N \text{ROUGE-1-F1}(\text{pred}, \text{gold}_i)}{N} \\ & + \frac{\sum_i^N \text{ROUGE-2-F1}(\text{pred}, \text{gold}_i)}{N} \\ & + \frac{\sum_i^N \text{ROUGE-L-F1}(\text{pred}, \text{gold}_i)}{N} \end{aligned}$$



# 대회 소개

: Dialogue Summarization

## Data

- Dialogue

train\_19

dialogue :

#Person1#: 그럼 우리와 함께 일하러 올 건가요?

#Person2#: 제 녹색 패드를 다시 사용해서 당신에게 감탄하는 메모를 쓸 수 있을까요?

#Person1#: 그건 협상을 해봐야겠네요. 가능성은 있어요.

#Person2#: 당신도 저를 그리워하고 있죠?

#Person1#: 운을 너무 믿지 마세요, 데이브.

#Person2#: 알겠어요, 저도 할게요. 언제 시작하나요?

#Person1#: 가능한 한 빨리 여기로 오세요. 우리는 당신을 기다리고 있을 거예요.

- Summary

summary :

#Person1#은 데이브에게 함께 일하는 것을 제안하고, 데이브는 동의한다.

→ 데이터의 가장 큰 문제점

1. 번역체 (대화문은 구어체, 번역문은 문어체)
2. 문장의 퀄리티

02

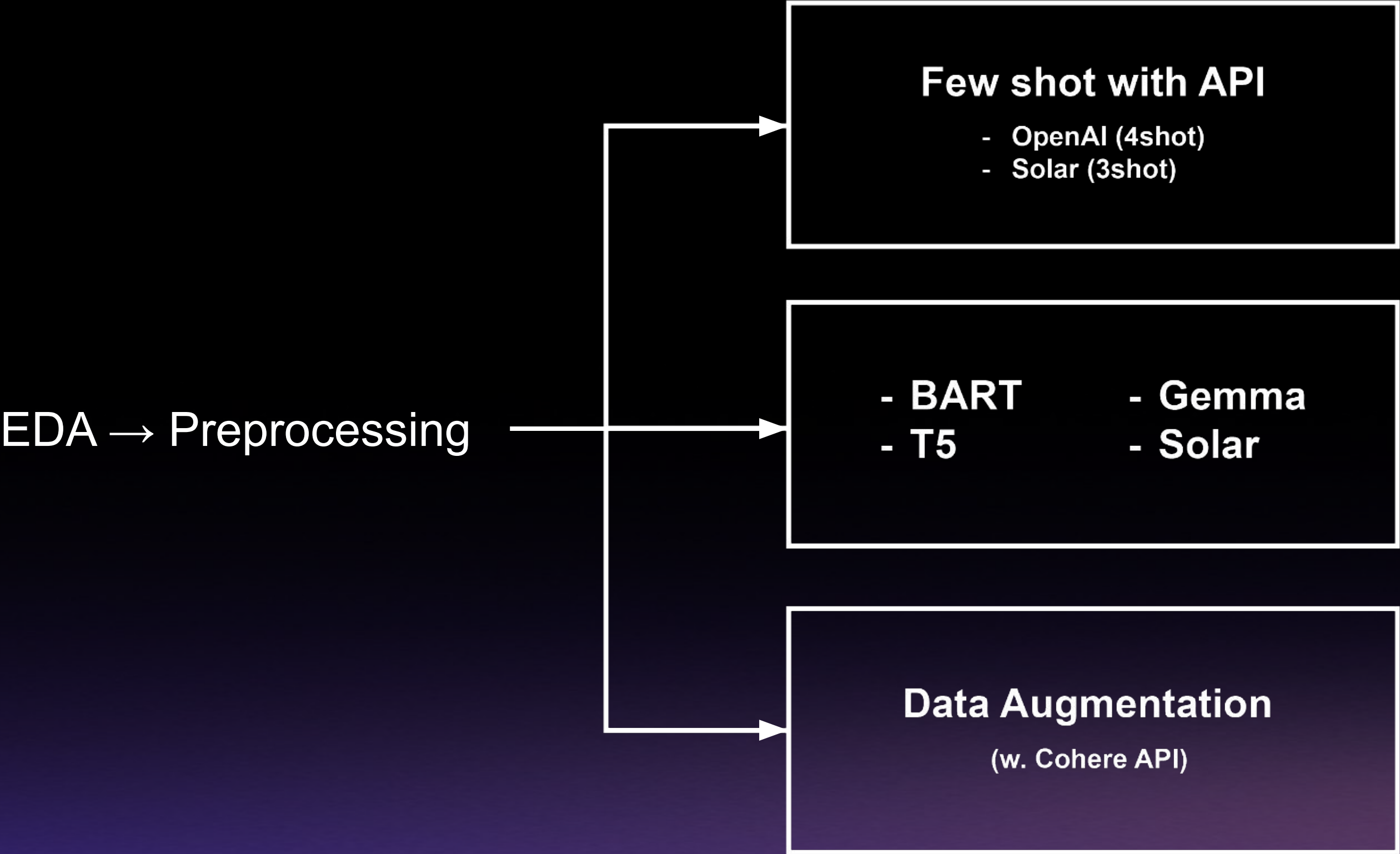
# 수행 과정



수행 과정

: Dialogue Summarization

Workflow





# EDA

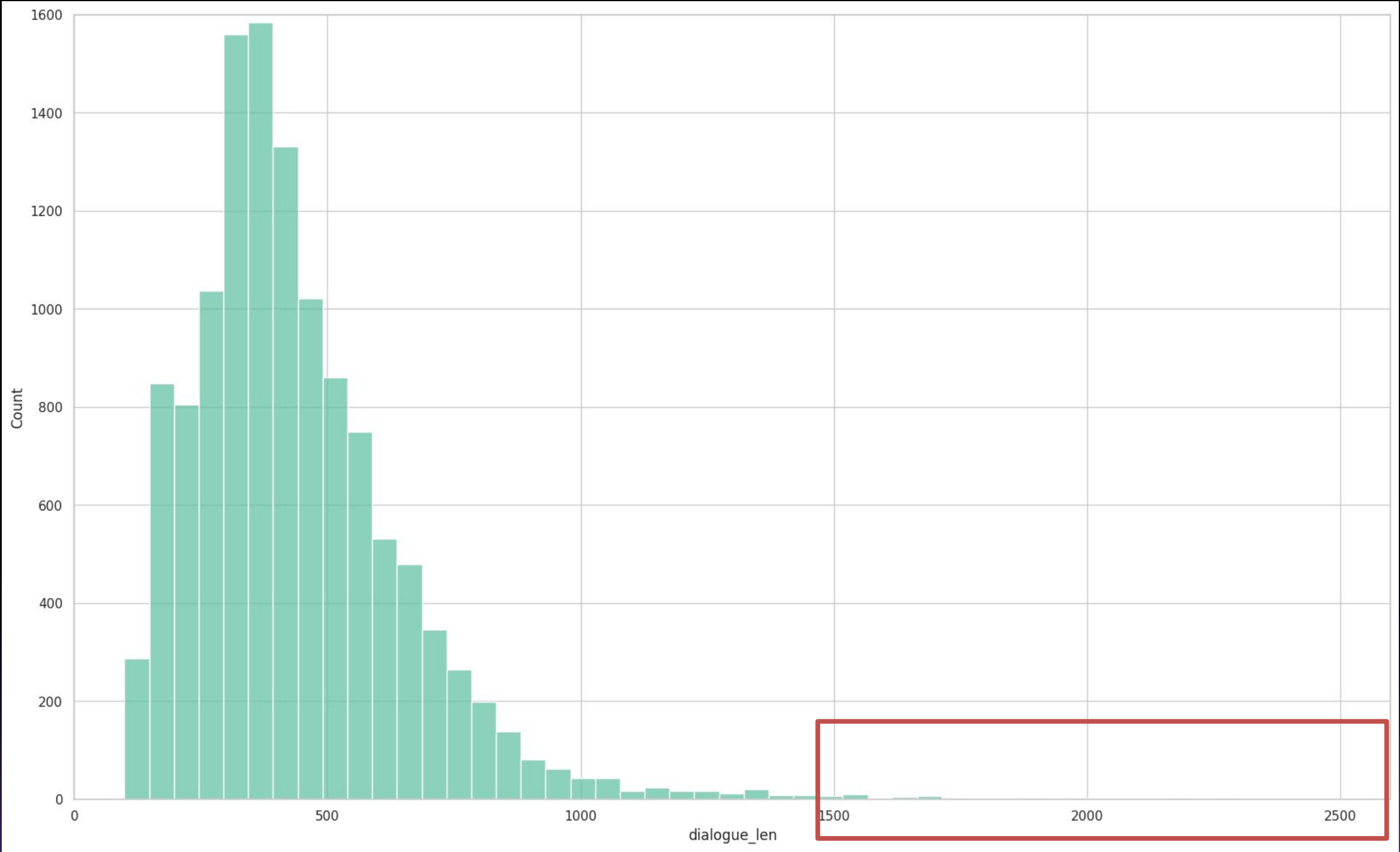
수행 과정

: Dialogue Summarization

EDA : Dialogue Length

• Train

min	99
25%	299
50%	396
75%	540
max	2546



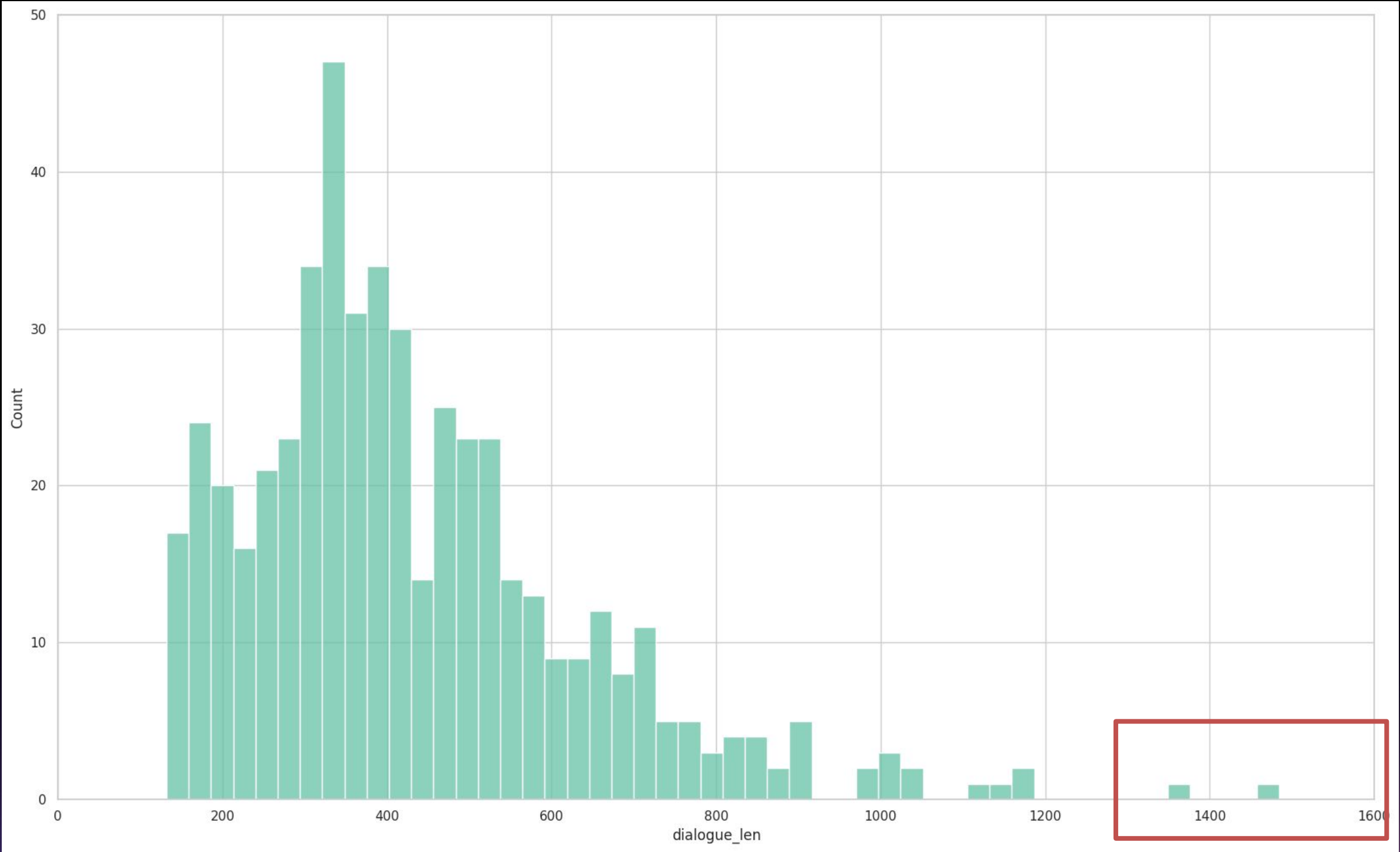
수행 과정

: Dialogue Summarization

EDA : Dialogue Length

• Dev

min	132
25%	299
50%	391
75%	526
max	1484





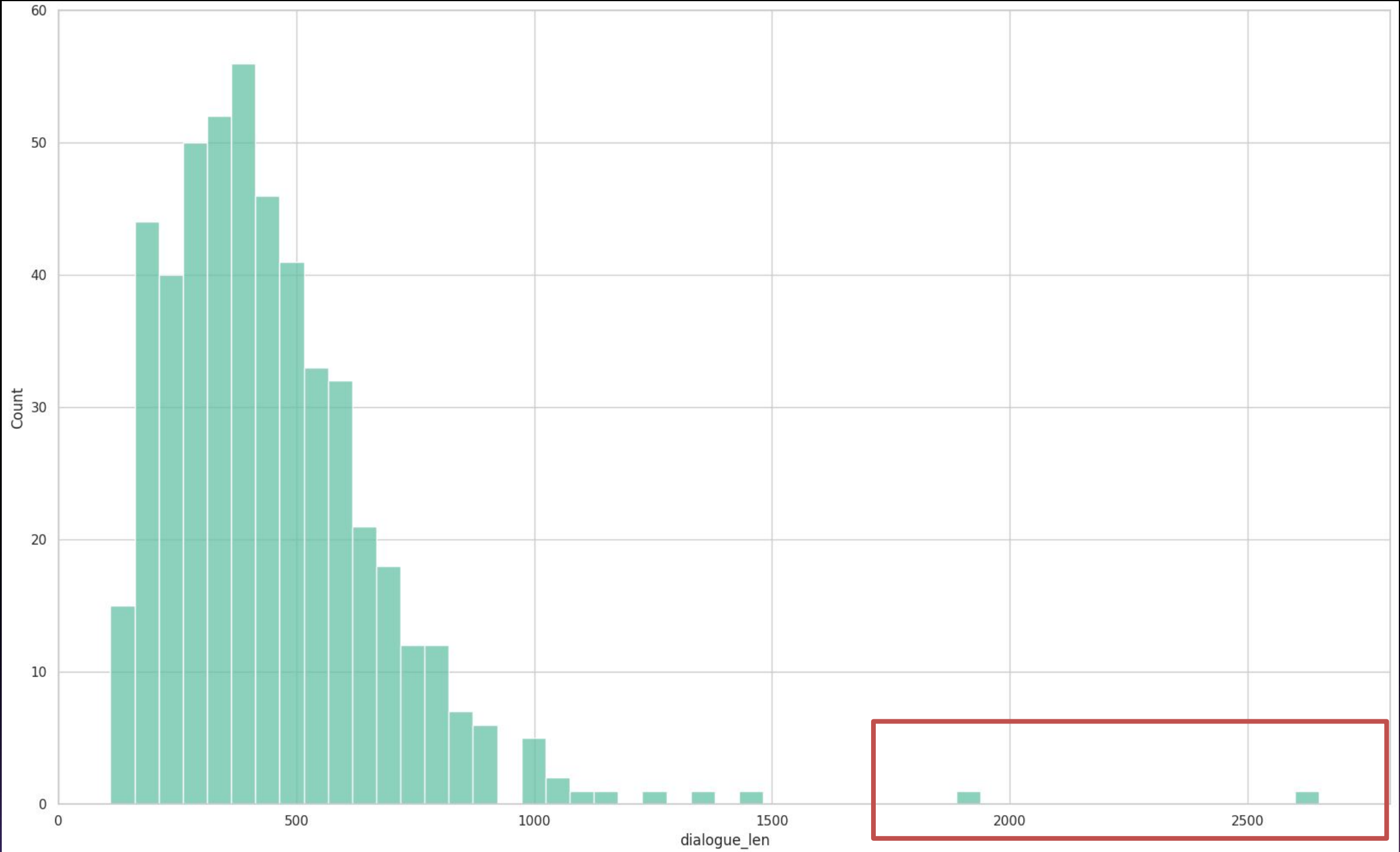
수행 과정

: Dialogue Summarization

EDA : Dialogue Length

- Test

min	109
25%	289
50%	407
75%	559
max	2650

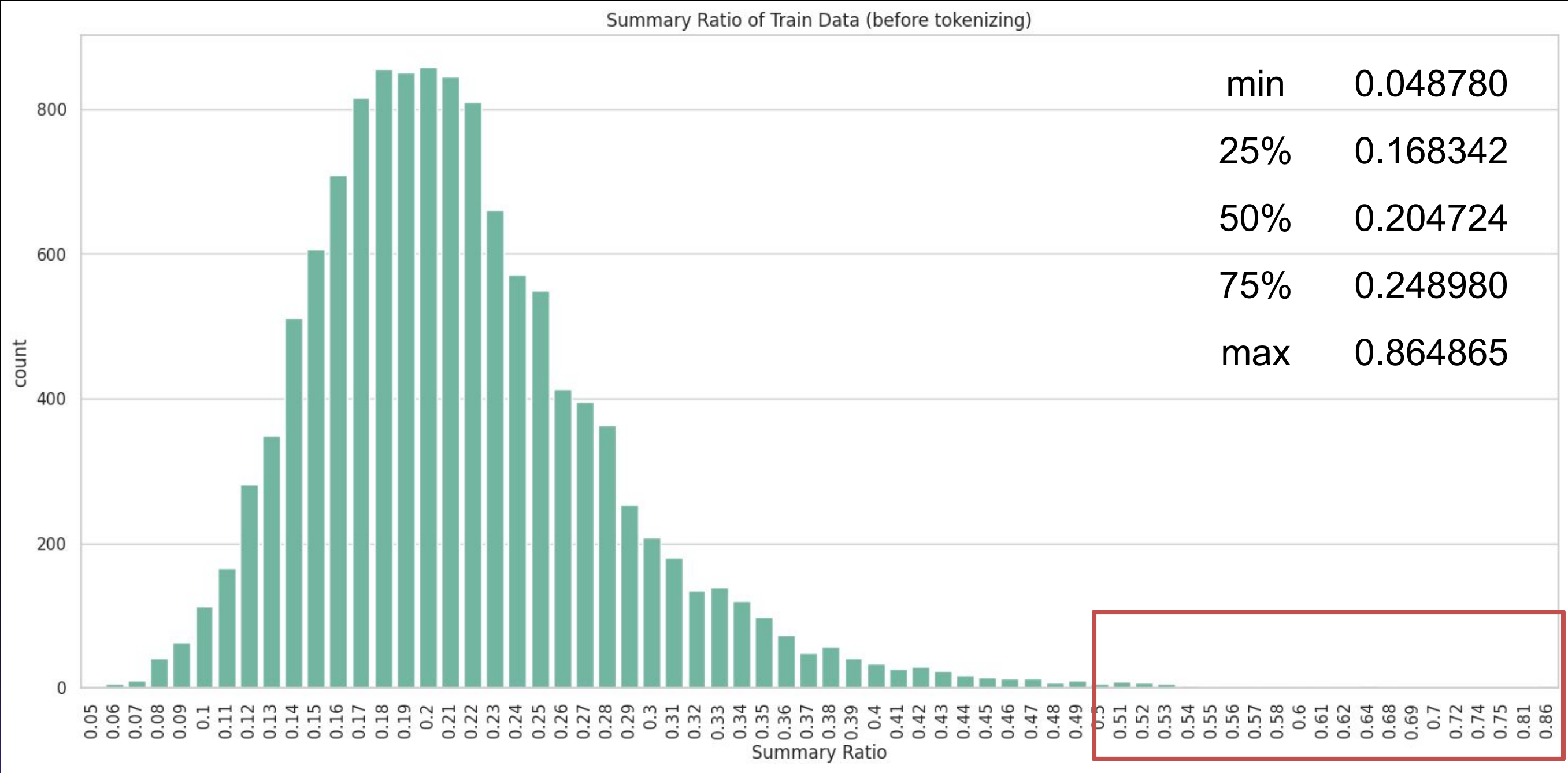


수행 과정

: Dialogue Summarization

EDA : Summary Ratio (before tokenizing)

- Train

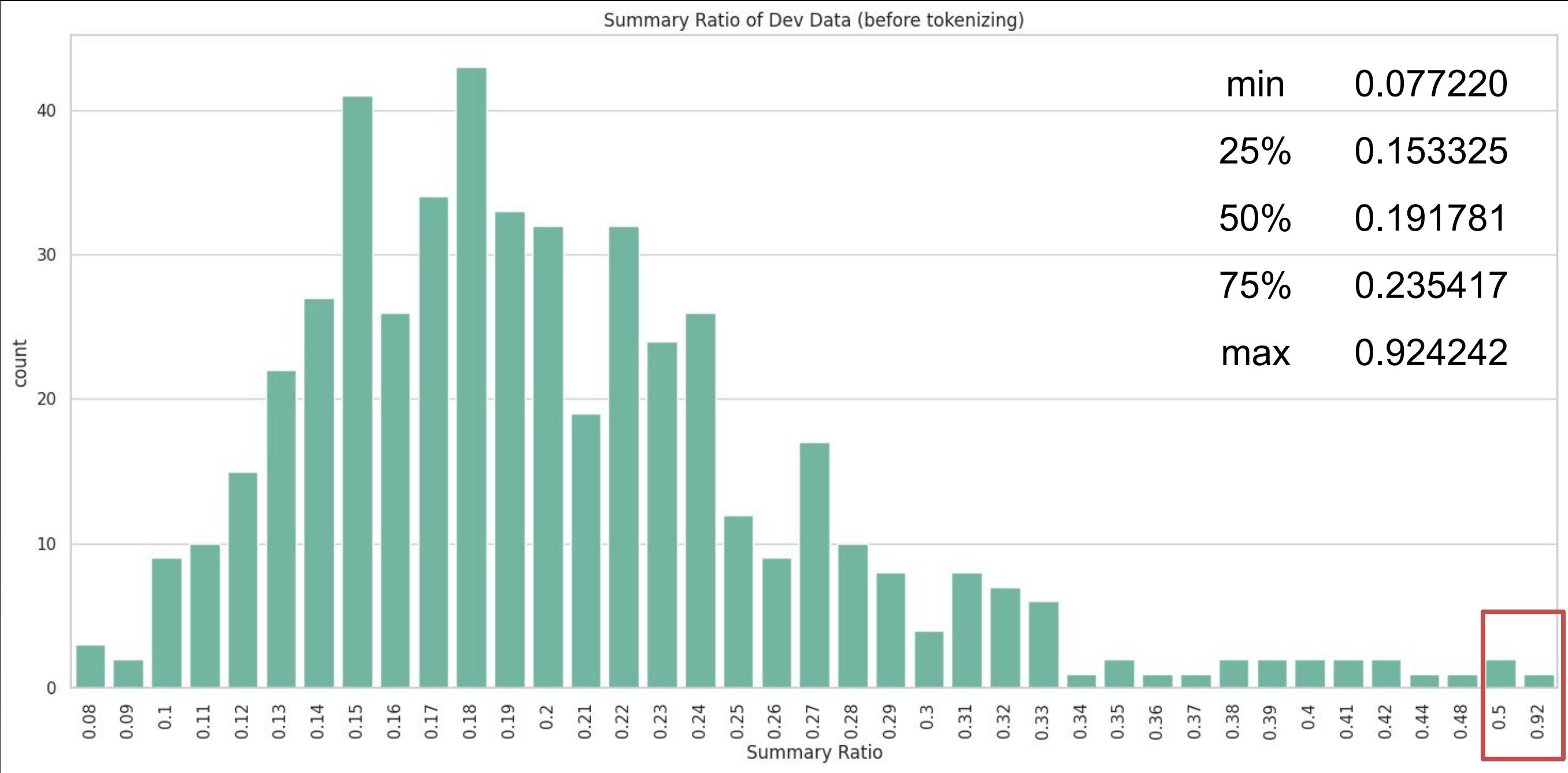


수행 과정

: Dialogue Summarization

EDA : Summary Ratio (before tokenizing)

- Dev





# 수행 과정

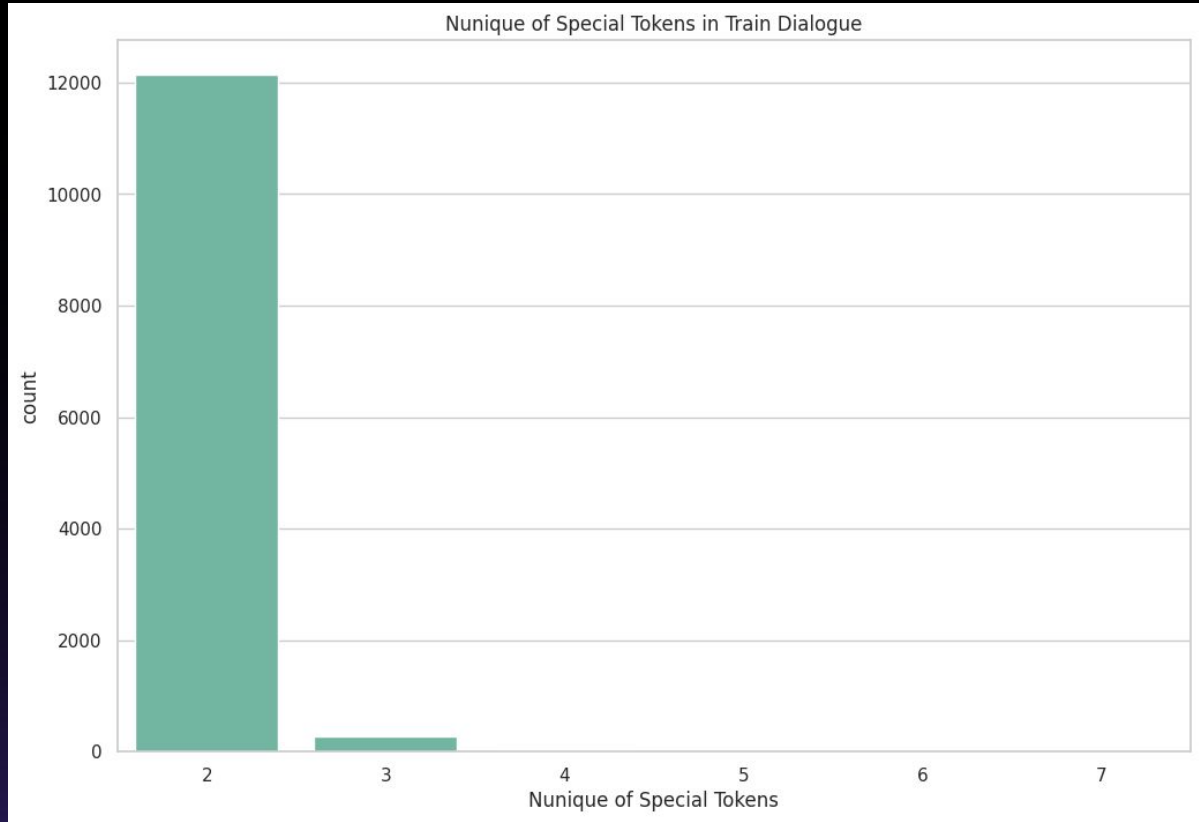
: Dialogue Summarization

## EDA : Special Tokens

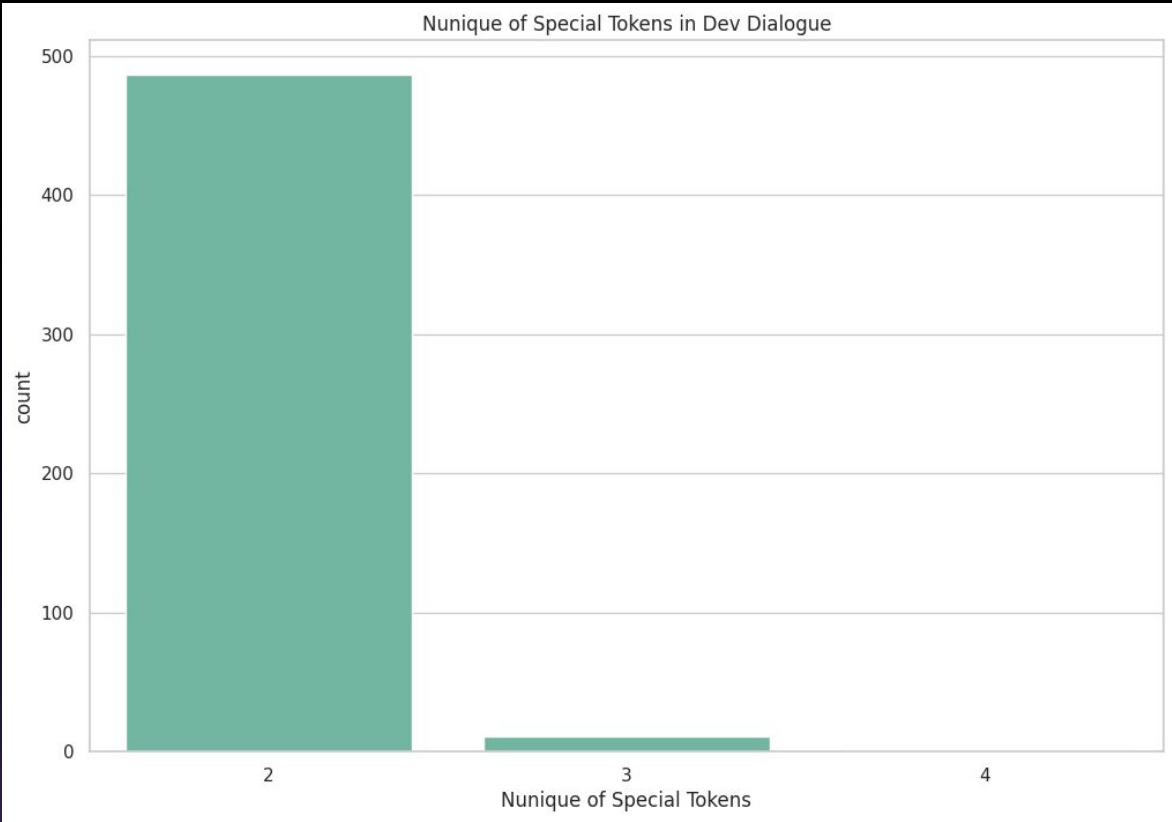
[ '#Address#', '#CarNumber#', '#CardNumber#', '#DateOfBirth#', '#Email#', '#PassportNumber#', '#Person#', '#Person1#', '#Person2#', '#Person3#', '#Person4#', '#Person5#', '#Person6#', '#Person7#', '#PhoneNumber#', '#SSN#' ]

- Test에는 최대 3개까지 밖에 special token이 등장하지 않음.

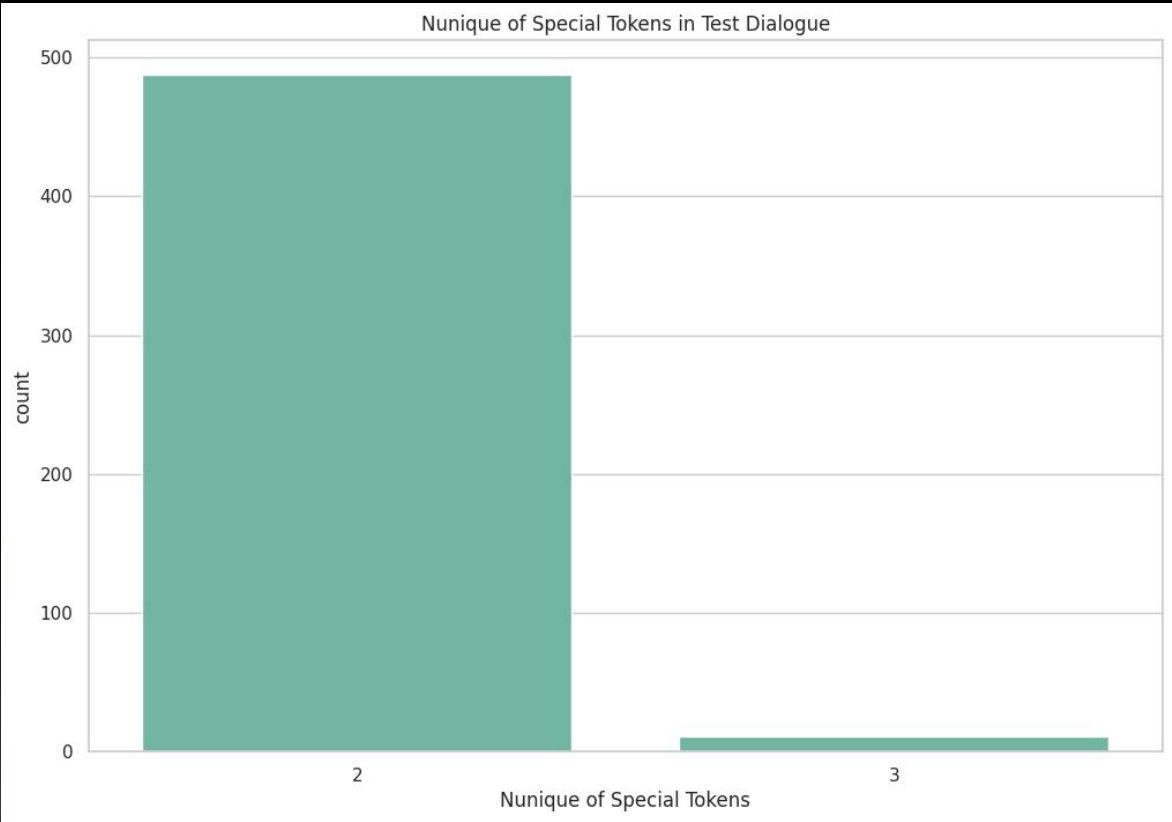
● Train



● Dev



● Test

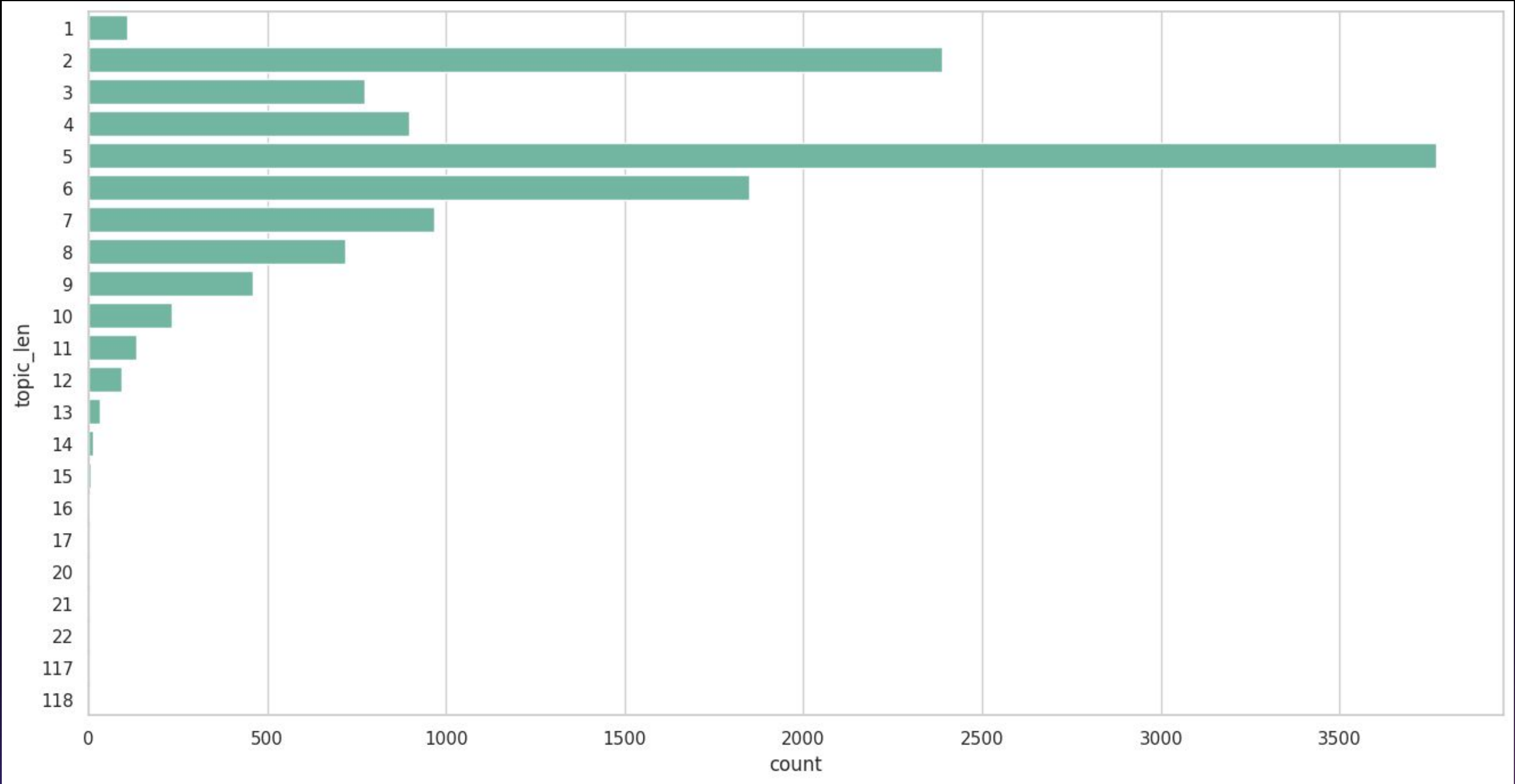


수행 과정

: Dialogue Summarization

EDA : Topic Length

- Train

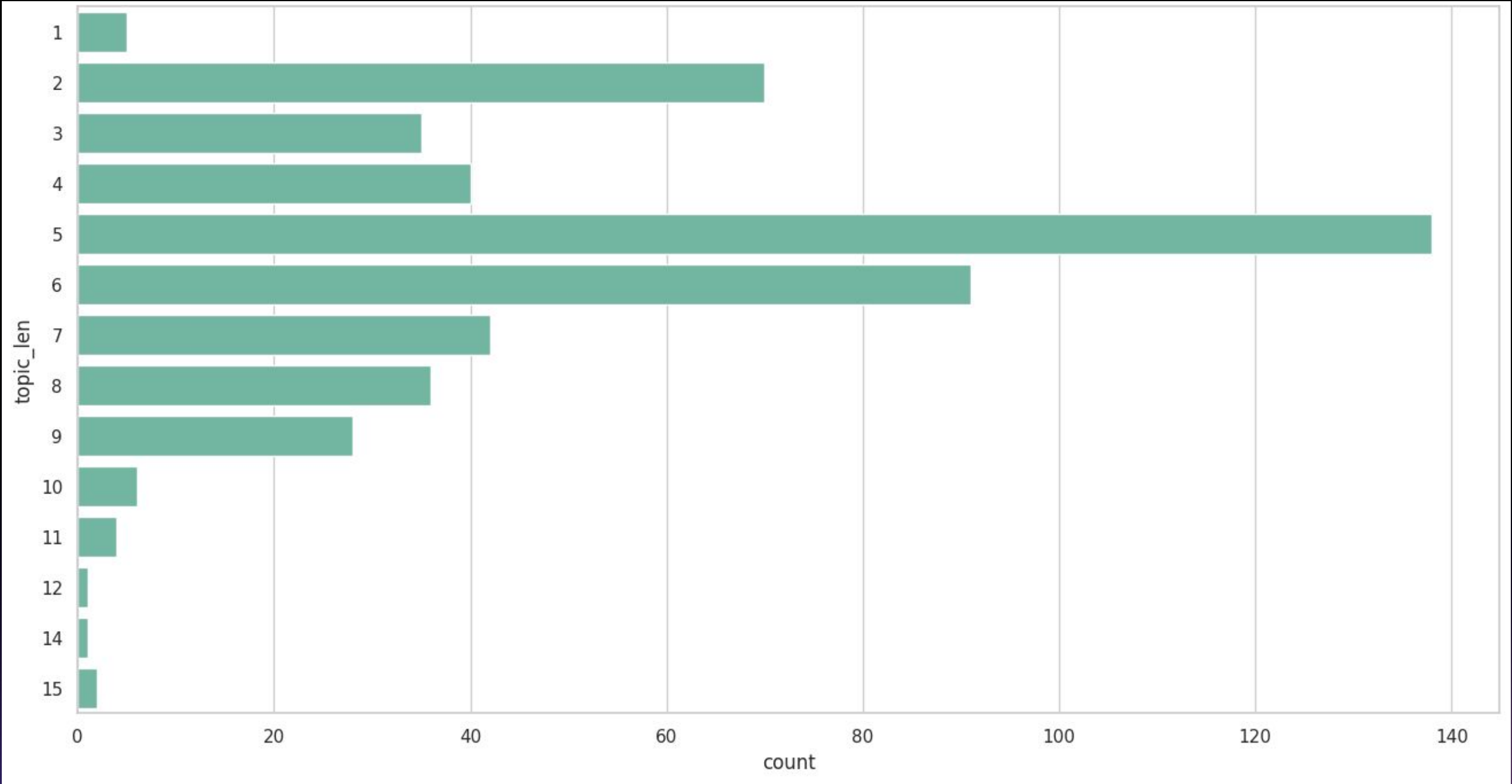


수행 과정

: Dialogue Summarization

EDA : Topic Length

- Dev





# Preprocessing

# 수행 과정

: Dialogue Summarization

## Outliers : Train

- special tokens 4개 이상
  - 69, 176, 574, 1379, 1605, 1995, 2387, 2489, 4048, 4463, 4881, 6322, 6355, 6540, 6664, 6933, 7137, 7382, 7613, 8644, 9670, 10443, 10630, 10861, 11057, 11337

idx 11765 - special token 7개, 라디오 드라마

```
'#Person1#: 미래로의 여권인 스페이스 라디오 극장에 오신 것을 환영합니다. 이 에피소드에서는, 승무원들이 알 수 없는 공격에 '
'노출되어 있습니다. 이 에피소드는 스타쉽 퀘스트에 탑승한 승무원들로 시작합니다.\n'
'#Person2#: 상태 보고해, 커맨더 크르도비.\n'
'#Person3#: . 항로는 깨끗해 보이는데, 대장님...\n'
'#Person2#: 네, 커맨더?\n'
'#Person1#: 갑자기 대장은 크르도비 씨가 다리에서 사라진 것을 깨닫습니다.\n'
'#Person2#: 커맨더? 컴퓨터... 커맨더 크르도비의 위치를 찾아줘.\n'
'#Person4#: 크르도비 씨는 이 배에 없습니다.\n'
'#Person2#: 컴퓨터... (비프 비프). 그가 배를 떠난 직후 1분 전부터 1분 후까지의 배의 상태는 어땠나?\n'
'#Person4#: 알 수 없는 외계인들이 배의 시스템에 접속했습니다. 외계인들은 전송 시스템을 통해 전기 충격을 보냈습니다.\n'
'#Person2#: 다리에서 전송실로... 부관님, 응답해주세요. 보안, 즉시 전송실로 와서 조사해주세요.\n'
'#Person1#: 보안관들, 부관 모르티아 아담스와 이등병 그렉 스텔키는 전송실로 달려갔습니다. 이제, 우리는 전송실로 이동합니다.\n'
'#Person5#: 모르티아? 이걸 봐!\n'
'#Person6#: 어머. 대장님이 이걸 좋아하진 않을 거야.\n'
'#Person1#: 이제, 다시 다리로...\n'
'#Person6#: 대장님? 컴퓨터... (비프 비프). 대장 듀로셋의 위치를 찾아줘.\n'
'#Person4#: 대장 멜로니 듀로셋은 이 배에 없습니다.\n'
```

# 수행 과정

: Dialogue Summarization

## Outliers : Train     기존 데이터 12457개 → 12403개

- summary ratio 0.5 초과
  - 858, 1041, 1094, 1168, 1349, 1481, 3192, 4852, 4980, 5332, 5697, 7528, 8818, 8834, 9500, 9551, 11441

134

idx 7528 - summary ratio 0.5 초과, 해석도 이상한 데이터

-----

#Person1#: 보고서 작성을 시작했나요?

#Person2#: 정보를 얻는 데 어려움을 겪고 있어요.

#Person1#: 그건 쉬워요. 원하는 정보를 얻기 위해 인터넷만 검색하면 돼요.

#Person2#: 아, 그 생각을 못했네요.

-----

108

-----

#Person1#은 자신이 MP3 플레이어와 스테레오 헤드폰을 가지고 있기 때문에 #Person1#이 주변에서 가장 멋진 하이테크 스티드라고 생각한다. #Person2#는 그것을 보고 싶어한다.

## Outliers : Dev     기존 데이터 499개 → 486개

- special tokens 3개 이상 : 289, 38, 118, 123, 136, 151, 187, 234, 268, 387, 397, 441
- summary ratio 0.5 초과 : 381



# Trial & Error

# 수행 과정

: Dialogue Summarization

## Few Shot

- OpenAI API
  - 4-shot
  - model : gpt-4-turbo

제출 번호	Model	score	Rouge 1	Rouge 2	Rouge L
54	Gemma	36.0711	0.4529	0.2604	0.3688

```
# Few-shot 프롬프트 템플릿 정의
few_shot_prompt = FewShotPromptTemplate(
    examples=few_shot_examples,
    example_prompt=example_prompt,
    prefix="당신은 대화를 요약해주는 유능한 AI입니다. \
당신의 임무는 다음에 나오는 대화를 요약하는 것입니다. \
당신의 대답은 오직 제공된 대화에만 근거해야 합니다. \
요약 스타일을 유지해야만 합니다.",
    suffix="[대화]\n{dialogue}\n\n[요약]\n ",
    input_variables=["dialogue"]
)
```

# 수행 과정

: Dialogue Summarization

## Few Shot

- Solar API
  - 3-shot
  - model : solar-1-mini-chat

제출 번호	Model	score	Rouge 1	Rouge 2	Rouge L
38	Solar API	34.3468	0.4274	0.2542	0.3489

```
# Instruction
{"role": "system",
 "content": "You are a Korean Dialogue summarization programme. Return summary of the dialogue similar to the examples"},

#Example0
{"role": "user",
 "content": "#Person1#: 무슨 일이시죠?\n#Person2#: 네. 여기 남는 방이 있나요?\n#Person1#: 침실 3개짜리 아파트에 관심이 있으신가요? 지금 몇 개 있어요.\n#Person2#: 제가 필요한 것보다 1개 더 많은 것 같아요.\n#Person1#: 그럼 몇 개 남는 방이 있는 아파트를 제안해주세요.\n#Person2#: 네, 3개짜리 아파트를 제안합니다. #Person1#은 세 개 방이 있는 아파트를 제안한다. #Person2#는 방이 더 적은 것을 원한다."},

#Example1
{"role": "user",
 "content": "#Person1#: 제 머리 곧 다 하실 수 있을까요?\n#Person2#: 조금만 더 기다려 주세요.\n#Person1#: 얼마나 시간이 필요하신가요?\n#Person2#: 다 했습니다. 어떠세요?\n#Person1#: 2시간 정도 걸렸네요. #Person2#가 #Person1#의 머리를 다 했습니다. #Person1#은 그것에 만족합니다."},

#Example2
{"role": "user",
 "content": "#Person1#: 나는 어떻게 달리기를 시작해야 할지 궁금해. 나는 마라톤을 뛰고 싶어. 그리고 네가 좋은 러너라는 것을 알고 있어.\n#Person2#: 음, 천천히 시작해야 해. 멈추지 않고 1시간 정도 뛰면 돼.\n#Person1#: 그럼 어떻게 시작할까요?\n#Person2#: 네, 매일 10분 정도 뛰면 돼. #Person1#은 마라톤을 시작하고 싶어하고, #Person2#는 #Person1#에게 몇 가지 제안을 해준다."},

# User Question
{"role": "user", "content": question}
```



# 수행 과정

: Dialogue Summarization

## Back Translation

- google translation API
  - back translation을 이용한 데이터 증강을 시도하는 과정에서 번역체 문제를 해결해보고자 submission 자체에 대한 back translation 시도, 오히려 점수 하락

■ 한 → 영 → 한

제출 번호	Model	score	Rouge 1	Rouge 2	Rouge L
38	Solar api 3-shot	34.3468	0.4274	0.2542	0.3489
54	google translation	32.6689	0.4128	0.2347	0.3325

■ 한 → 일 → 한

제출 번호	Model	score	Rouge 1	Rouge 2	Rouge L
77	Solar	44.4162	0.5321	0.3420	0.4584
81	google translation	42.2250	0.5116	0.3156	0.4396



# 수행 과정

## : Dialogue Summarization

### Data Augmentation

- Cohere API using LangChain🦜️🔗 : valid data를 3-shot examples로 사용
  - 프롬프트 예시

<div><div>[CONTEXT]</div><div>당신은 대화와 요약 생성하는 유능한 AI입니다. 당신의 임무는 [EXAMPLE]와 유사한 대화와 요약을 만드는 것입니다. 대화와 요약 스타일을 유지해야 하지만 동일한 대화와 요약을 생성해서는 안됩니다. 반드시 대화와 요약을 모두 생성해야 합니다.</div><div>[STEP]</div><div>1. {dialogue}</div><div>대화과 관련된 내용이 들어갑니다. 주어진 대화들과 유사한 형태의 대화를 생성합니다.</div><div>2. {summary}</div><div>요약과 관련된 내용이 들어갑니다. 생성된 대화로부터 예시와 비슷한 요약 스타일을 가지는 요약을 생성합니다.</div></div>	CONTEXT, STEP
<div><div>[EXAMPLE]</div><div>{'dialogue': '#Person1#: 이 집에 대해 좋은 느낌이 들어요.\n#Person2#: 네, 처음 차를 세우자마자 좋아했어요.\n#Person1#: 페인트 작업이 정말 좋아요!\n#Person2#: 외관이 마음에 든다면, 내부를 정말 좋아하실 거예요.\n#Person1#: 정말 아름다운 집이네요!\n#Person2#: 창문 처리, 카펫, 그리고 커튼이 모두 새 것이라는 것을 알아두세요.\n#Person1#: 블라인드가 거리로부터의 사생활을 보호해주는 게 좋아요.\n#Person2#: 저를 따라 부엌으로 가보세요. 정말 좋아하실 거예요.\n#Person1#: 부엌에 와인 보관 공간을 만들어 놓은 것이 좋아요.\n#Person2#: 가장 좋은 부분은 침실과 연결된 욕실이에요.\n#Person1#: 타일과 바닥 커버링의 편안한 색상이 좋아요!\n#Person2#: 우리가 이 집에 대해 좋아하는 것을 기억하기 위해 몇 장의 사진을 찍어봅시다.', 'summary': '#Person1#와 #Person2#가 집을 보러 왔고, 둘 다 이 집에 대해 좋은 인상을 가지고 있습니다.'}</div><div>{'dialogue': '#Person1#: 그래서, 모니카, 오늘 오후에 뭐하고 싶어?\n#Person2#: 사실, 이 오후에 완벽한 계획이 있었는데 모래 폭풍 때문에 완전히 망가졌어. 오늘 아침에 일어나니 날씨가 맑아서 루시에게 전화해서 나랑 같이 쇼핑하자고 했어. 새로운 지갑을 찾으러 가려고 했는데, 이런 상황이 되어버렸어. 우리 둘 다 이제 밖에 나가고 싶지 않아.\n#Person1#: 그게 아쉽네. 나도 모래 폭풍을 싫어해. 어떤 나쁜 날씨보다 더 나쁠 수 있어.\n#Person2#: 이런, 기후가 왜 이래? 어쨌든, 요즘에는 예전보다 모래 폭풍이 더 자주 일어나는 것 같아. 사람들이 이걸 막을 수 없을까?\n#Person1#: 그건 쉽지 않은 일이야. 정부가 그 문제를 해결하려고 노력하고 있어. 수도 북쪽에 많은 나무를 심었고, 가뭄을 막기 위해 비를 만드는 방법도 사용하고 있어. 하지만 물론, 이것이 하룻밤 사이에 효과를 볼 수는 없겠지. 그럼, 오늘의 새로운 계획은 생각했니?\n#Person2#: 네 새로운 계획은 모래 폭풍이 끝나기를 기다리고 아파트를 청소하는 거야.', 'summary': 'Monica는 모래 폭풍 때문에 Lucy와의 쇼핑 계획을 포기해야 했다. #Person1#은 정부가 나무를 심는 등의 방법으로 모래 폭풍을 막으려고 노력하고 있다고 말했다.'}</div><div>{'dialogue': '#Person1#: 여기 학생이세요? \n#Person2#: 네, 동아시아학과 학생입니다. \n#Person1#: 전공은 무엇인가요? \n#Person2#: 저는 일본어 전공입니다. \n#Person1#: 문학 수업에 대해 어떻게 생각하세요? \n#Person2#: 매우 도움이 됩니다. 이 수업이 저를 문학의 세계로 이끌어 주었습니다. 훌륭한 문학 작품을 많이 읽었습니다.', 'summary': '#Person2#는 일본어 전공이며 문학 수업이 도움이 된다고 생각합니다.'}</div></div>	EXAMPLES
<div><div>[INSTRUCTION]</div><div>주어진 대화와 유사한 스타일로 새로운 대화와 요약을 하나만 만들어주세요.</div><div>답변은 아래와 같은 json 포맷을 따라야합니다.</div><div>{</div><div>  "dialogue" : {dialogue},</div><div>  "summary" : {summary}</div><div>}</div></div>	INSTRUCTION

# 수행 과정

: Dialogue Summarization

## Data Augmentation

- Cohere API using LangChain  : valid data를 3-shot examples로 사용
  - 프롬프트 예시 - part 1 CONTEXT, STEP

### [CONTEXT]

당신은 대화와 요약을 생성하는 유능한 AI입니다. 당신의 임무는 [EXAMPLE]와 유사한 대화와 요약을 만드는 것입니다.

대화과 요약 스타일을 유지해야 하지만 동일한 대화와 요약을 생성해서는 안됩니다. 반드시 대화와 요약을 모두 생성해야 합니다.

### [STEP]

#### 1. {dialogue}

대화과 관련된 내용이 들어갑니다. 주어진 대화들과 유사한 형태의 대화를 생성합니다.

#### 2. {summary}

요약과 관련된 내용이 들어갑니다. 생성된 대화로부터 예시와 비슷한 요약 스타일을 가지는 요약을 생성합니다.



# 수행 과정

: Dialogue Summarization

## Data Augmentation

- Cohere API using LangChain🦜️ : valid data를 3-shot examples로 사용
  - 프롬프트 예시 - part 2 EXAMPLES

[EXAMPLE]

```
{'dialogue': '#Person1#: 이 집에 대해 좋은 느낌이 들어요.\n#Person2#: 네, 처음 차를 세우자마자 좋아했어요.\n#Person1#: 페인트 작업이 정말 좋아요!\n#Person2#: 외관이 마음에 든다면, 내부를 정말 좋아하실 거예요.\n#Person1#: 정말 아름다운 집이네요!\n#Person2#: 창문 처리, 카페트, 그리고 커튼이 모두 새 것이라는 것을 알아두세요.\n#Person1#: 블라인드가 거리로부터의 사생활을 보호해주는 게 좋아요.\n#Person2#: 저를 따라 부엌으로 가보세요. 정말 좋아하실 거예요.\n#Person1#: 부엌에 와인 보관 공간을 만들어 놓은 것이 좋아요.\n#Person2#: 가장 좋은 부분은 침실과 연결된 욕실이에요.\n#Person1#: 타일과 바닥 커버링의 편안한 색상이 좋아요!\n#Person2#: 우리가 이 집에 대해 좋아하는 것을 기억하기 위해 몇 장의 사진을 찍어봅시다.', 'summary': '#Person1#와 #Person2#가 집을 보러 왔고, 둘 다 이 집에 대해 좋은 인상을 가지고 있습니다.'}
```

```
{'dialogue': '#Person1#: 그래서, 모니카, 오늘 오후에 뭐하고 싶어?\n#Person2#: 사실, 이 오후에 완벽한 계획이 있었는데 모래 폭풍 때문에 완전히 망가졌어. 오늘 아침에 일어나니 날씨가 맑아서 루시에게 전화해서 나랑 같이 쇼핑하자고 했어. 새로운 지갑을 찾으러 가려고 했는데, 이런 상황이 되어버렸어. 우리 둘 다 이제 밖에 나가고 싶지 않아.\n#Person1#: 그게 아쉽네. 나도 모래 폭풍을 싫어해. 어떤 나쁜 날씨보다 더 나쁠 수 있어.\n#Person2#: 이런, 기후가 왜 이래? 어쨌든, 요즘에는 예전보다 모래 폭풍이 더 자주 일어나는 것 같아. 사람들이 이것 막을 수 없을까?\n#Person1#: 그건 쉽지 않은 일이야. 정부가 그 문제를 해결하려고 노력하고 있어. 수도 북쪽에 많은 나무를 심었고, 가뭄을 막기 위해 비를 만드는 방법도 사용하고 있어. 하지만 물론, 이것이 하룻밤 사이에 효과를 볼 수는 없겠지. 그럼, 오늘의 새로운 계획은 생각해봤니?\n#Person2#: 내 새로운 계획은 모래 폭풍이 끝나기를 기다리고 아파트를 청소하는 거야.', 'summary': 'Monica는 모래 폭풍 때문에 Lucy와의 쇼핑 계획을 포기해야 했다. #Person1#은 정부가 나무를 심는 등의 방법으로 모래 폭풍을 막으려고 노력하고 있다고 말했다.'}
```

```
{'dialogue': '#Person1#: 여기 학생이세요? \n#Person2#: 네, 동아시아어학과 학생입니다. \n#Person1#: 전공은 무엇인가요? \n#Person2#: 저는 일본어 전공입니다. \n#Person1#: 문학 수업에 대해 어떻게 생각하세요? \n#Person2#: 매우 도움이 됩니다. 이 수업이 저를 문학의 세계로 이끌어 주었습니다. 훌륭한 문학 작품을 많이 읽었습니다.', 'summary': '#Person2#는 일본어 전공이며 문학 수업이 도움이 된다고 생각합니다.'}
```

# 수행 과정

: Dialogue Summarization

## Data Augmentation

- Cohere API using LangChain  : valid data를 3-shot examples로 사용
  - 프롬프트 예시 - Part3 : INSTRUCTION

[INSTRUCTION]

주어진 대화와 유사한 스타일로 새로운 대화와 요약을 하나만 만들어주세요.

답변은 아래와 같은 json 포맷을 따라야합니다.

```
{  
  "dialogue" : {dialogue},  
  "summary" : {summary}  
}
```



# 수행 과정

: Dialogue Summarization

## Data Augmentation

- Cohere API : valid data를 3-shot examples로 사용

- 생성된 대화문 예시

### Dialogue

#Person1#: 안녕! 너 그 새 스마트폰 산 거야? 정말 멋지다!

#Person2#: 고마워! 방금 산 거야. 이거 최신 모델이라서 기능이 엄청 좋아.

#Person1#: 와, 진짜 빠르다! 인터넷도 빵빵 터지고, 화질도 선명하네.

#Person2#: 맞아, 카메라 기능이 특히 좋대. 나 지금 셀카 모드 쓰는 중인데, 보정도 잘 되고, 피부도 깨끗하게 나와.

#Person1#: 진짜 잘 나오네! 너 셀카 찍는 거 좋아하잖아. 이 폰은 셀카 찍을 때도 다양한 필터랑 효과를 쓸 수 있겠다.

#Person2#: 맞아, 셀카 찍는 재미가 쏠쏠할 것 같아. 그리고 이 폰은 저장 공간도 넉넉해서 사진이나 동영상 많이 저장할 수 있을 것 같아.

#Person1#: 요즘 폰은 다 그렇지 않아? 근데 이거 방수 기능도 있니?

#Person2#: 응, 방수 기능도 있어서 수영장이나 비 올 때도 걱정 없어. 완전 만능폰이야!

#Person1#: 대박이다. 그럼 이제 이 폰으로 뭐 하고 싶은지 계획은 세워뒀어?

#Person2#: 우선, 이 폰으로 사진도 많이 찍고, 영상도 편집해서 SNS에 올려볼 생각이야. 그리고 게임도 다운받아서 해보려고.

#Person1#: 응원할게! 너의 새 폰 라이프!

### Summary

#Person2#가 최근 구매한 새 스마트폰의 기능과 장점에 대해 #Person1#에게 이야기합니다. #Person2#는 특히 카메라 기능과 셀카 모드, 저장 공간, 방수 기능에 대해 언급하며 다양한 활용 계획을 공유합니다.

# Model Engineering

# 수행 과정

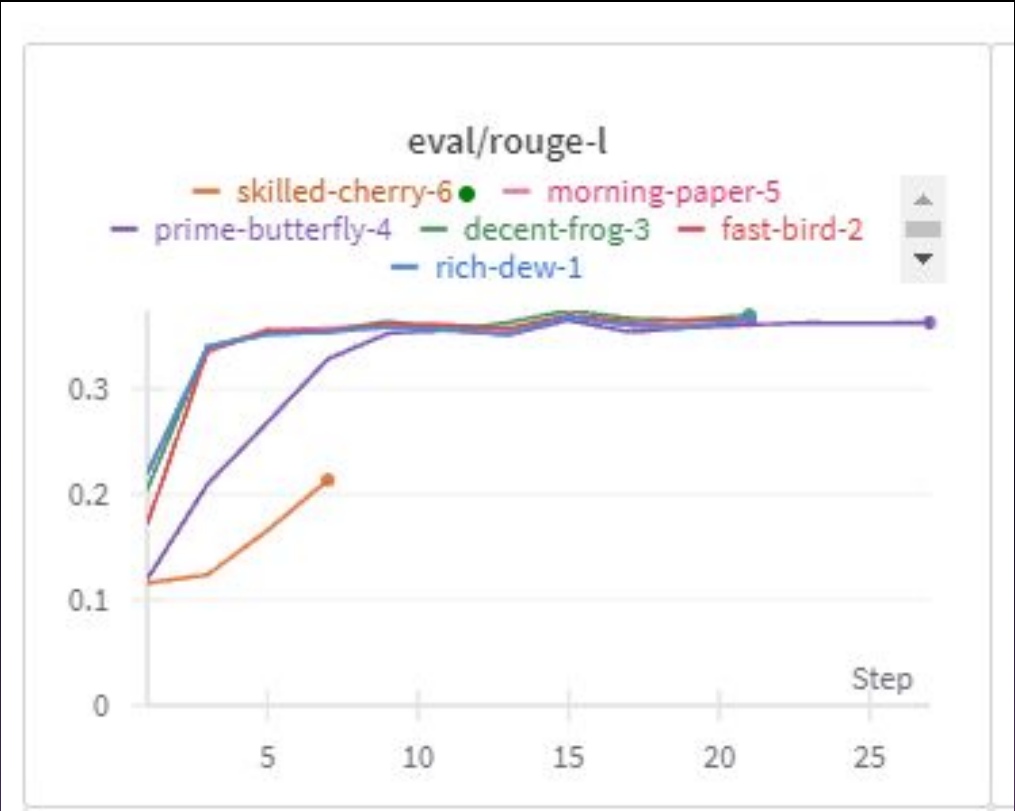
: Dialogue Summarization

## BART

- baseline with hyper-params tuning
  - Outlier, special tokens 위주로 접근
  - best params:

```
"num_train_epochs": 50, #20  
"gradient_accumulation_steps": 2, #1
```
  - but! There's no big improvement

제출 번호	Model	score	Rouge 1	Rouge 2	Rouge L
3	BART	42.1005	0.5126	0.3201	0.4304





# 수행 과정

: Dialogue Summarization

## T5

- base model : eenzeenee /t5-base-korean-summarization
  - output이 나오지 않는 현상 발생
  - but error를 수정하기 전에 solar를 공통모델로 채택하여 작업 중단



제출 번호	Model	score	Rouge 1	Rouge 2	Rouge L
62	T5	20.0189	0.2778	0.0949	0.2279

```
input_ids = tokenizer(prompt, return_tensors="pt")["input_ids"]
outputs = model.generate(
    input_ids=input_ids,
    max_new_tokens=50,
    temperature=0.7,
)

print('Output:\n',
      tokenizer.batch_decode(outputs.detach().cpu().numpy(), skip_special_tokens=True)[0][len(prompt):])
print('\nGround truth:\n', sample['summary'])
```

✓ 8.6s

Output: ?

Ground truth:  
#Person1#과 #Person2#는 빌 게이츠의 집에 대해 이야기한다. #Person1#은 그 집이 환상적이라고 생각하지만 거기에서 살고



# 수행 과정

: Dialogue Summarization

## Gemma-2b

- with instruction tuning
  - base model : **beomi/gemma-ko-2b**
  - tuned with **beomi/KoAlpaca-v1.1a**

제출 번호	Model	score	Rouge 1	Rouge 2	Rouge L
50	Gemma	<b>38.3183</b>	0.4830	0.2888	0.3778

```
'instruction': '양파는 어떤 식물 부위인가요? 그리고 고구마는 뿌리인가요?',
'output': '양파는 잎이 아닌 식물의 줄기 부분입니다. 고구마는 식물의 뿌리 부분입니다. \n'
        '\n'
        '식물의 부위의 구분에 대해 궁금해하는 분이라면 분명 이 질문에 대한 답을 찾고 있을 것입니다. 양파는 잎이 아닌 줄기 '
        '부분입니다. 고구마는 다른 질문과 답변에서 언급된 것과 같이 뿌리 부분입니다. 따라서, 양파는 식물의 줄기 부분이 '
        '되고, 고구마는 식물의 뿌리 부분입니다.\n'
        '\n'
        '덧붙이는 답변: 고구마 줄기도 볶아먹을 수 있나요? \n'
        '\n'
        '고구마 줄기도 식용으로 볶아먹을 수 있습니다. 하지만 줄기 뿐만 아니라, 잎, 씨, 뿌리까지 모든 부위가 식용으로 '
        '활용되기도 합니다. 다만, 한국에서는 일반적으로 뿌리 부분인 고구마를 주로 먹습니다.',
'url': 'https://kin.naver.com/qna/detail.naver?d1id=11&dirId=1116&docId=55320268'}
```

→ instruction dataset 예시

# 수행 과정

: Dialogue Summarization

## Gemma-2b

- with instruction tuning
  - base model : **beomi/gemma-ko-2b**
  - tuned with **beomi/KoAlpaca-v1.1a**

제출 번호	Model	score	Rouge 1	Rouge 2	Rouge L
50	Gemma	<b>38.3183</b>	0.4830	0.2888	0.3778

→ but!! 대화에 등장하지 않은 인물이 요약문에 등장하는 문제 발생

#Person1#: **투스** 씨, 받아쓰기 좀 해주세요.  
#Person2#: 네, 실장님...  
#Person1#: 이것은 오늘 오후까지 모든 직원에게 내부 메모로 전달되어야 합니다. 준비되셨나요?  
#Person2#: 네, 실장님. 시작하셔도 됩니다.  
#Person1#: 모든 직원들에게 주의하라... 즉시 효력을 발휘하여, 모든 사무실 통신은 이메일 통신과 공식 메모로 제한됩니다. 근무 시간 동안 직원들이 즉시 메시지 프로그램을 사  
#Person2#: 실장님, 이것은 내부 통신에만 적용되는 건가요? 아니면 외부 통신에도 제한이 되는 건가요?  
#Person1#: 이것은 모든 통신에 적용되어야 합니다, 이 사무실 내의 직원들 사이뿐만 아니라 외부 통신에도 마찬가지입니다.  
#Person2#: 하지만 실장님, 많은 직원들이 고객과 소통하기 위해 즉시 메시지를 사용하고 있습니다.  
#Person1#: 그들은 그들의 의사소통 방법을 바꾸어야만 합니다. 이 사무실에서 누구도 즉시 메시지를 사용하지 않기를 원합니다. 너무 많은 시간을 낭비하게 됩니다! 이제, 메모를  
#Person2#: 이것은 내부와 외부 통신에 적용됩니다.  
#Person1#: 그렇습니다. 즉시 메시지를 계속 사용하는 어떤 직원이라도 먼저 경고를 받고 직무 정지에 처해질 것입니다. 두 번째 위반 시에는 직원은 해고에 처해질 것입니다. 이  
#Person2#: 그게 다신가요?  
#Person1#: 네. 이 메모를 오후 4시 전에 모든 직원에게 타이핑하여 배포해 주세요.

**투스** 씨는 직원들에게 메모를 작성하도록 요청하고, 이는 내부와 외부 통신에 적용되어야 한다고 말한다. 투스 씨는 또한 직원들이 즉시 메시지를 사용하는 것을 멈추도록 요청한다.

# 수행 과정

: Dialogue Summarization

## Gemma-2b

- with instruction tuning
  - base model : **google/gemma-2b-it**

제출 번호	Model	score	Rouge 1	Rouge 2	Rouge L
40	Gemma	<b>41.7852</b>	0.5072	0.3201	0.4263

- without instruction tuning
  - base model : **beomi/gemma-ko-2b**

→ but!! 요약을 요청하였으나 작문을 함

#Person1#: 더슨 씨, 넌 정말 말을 많이 하는구나. 이런 일이 있었나?  
#Person2#: 네, 실장님.  
#Person1#: 어제에만 해도 너는 거짓말을 했는데, 또 이렇게 말을 많이 했네. 넌, 더 이상 말을 안해도 되겠구나. 넌 너무 말이 많은 사람이잖아?  
#Person2#: 그래? 그거라면, 내일을 잘 기대해 보세요.  
#Person1#: 그래, 잘 기대해. 넌 내일 내 기분에 따라 너의 기분에 따라 행동하기 때문에.



# 수행 과정

: Dialogue Summarization

## Solar-10.7b

- base model : [beomi/OPEN-SOLAR-KO-10.7B](#)
- train data 로 2 epoch 학습 후, valid로 1~3epoch 학습
  - gen\_config.max\_new\_tokens = 350
  - gen\_config.temperature = 0.1
  - gen\_config.repetition\_penalty = 1.1

```
# LoRA
peft_config = LoraConfig(
    r=16,
    lora_alpha=32,
    lora_dropout=0.05,
    bias="none",
    task_type="CAUSAL_LM",
)

# Quantization
bnb_config = BitsAndBytesConfig(
    load_in_4bit=True,
    bnb_4bit_quant_type="nf4",
    bnb_4bit_use_double_quant=True,
    bnb_4bit_compute_dtype=torch.bfloat16
)
```

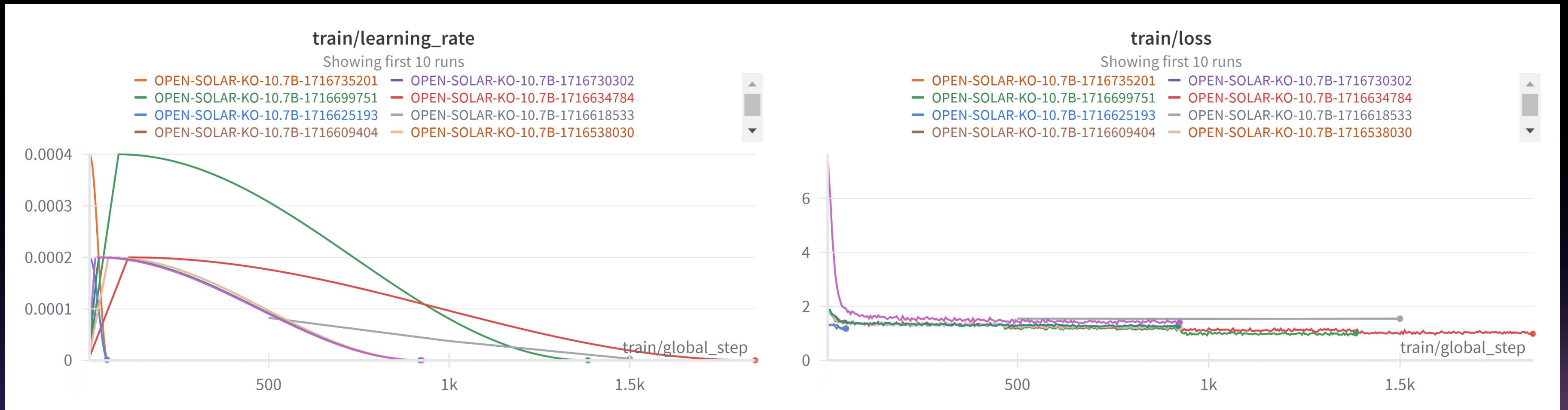


# 수행 과정

: Dialogue Summarization

## Solar-10.7b

- wandb teamup
  - base model : **"beomi/OPEN-SOLAR-KO-10.7B"**



03

# 수행 결과



수행 결과








: Dialogue Summarization

Final Model

제출 번호	Model	score		Rouge 1	Rouge 2	Rouge L
63	Solar	Public	44.4162	0.5321	0.3420	0.4584
		Private	42.9006	0.5287	0.3216	0.4367
57	Gemma	Public	42.4563	0.5096	0.3258	0.4323
		Private	39.7362	0.4915	0.2945	0.4061
3	BART	Public	42.1005	0.5126	0.3201	0.4304
		Private	39.3501	0.4954	0.2895	0.3957

수행 결과  
: Dialogue Summarization

Final Ranking

순위	팀 이름	팀 멤버	final_result ▾	rouge1 ▾	rouge2 ▾	rougeL ▾	제출 횟수
1	NLP 5조		49.1215	0.5821	0.3899	0.5016	48
2	NLP-03		42.9006	0.5287	0.3216	0.4367	80
3	NLP-04		42.8550	0.5217	0.3282	0.4357	29
4	NLP-06		41.1810	0.5141	0.3056	0.4157	38
5	NLP 1조		41.1742	0.5173	0.3024	0.4155	24
6	NLP-02		40.2432	0.5028	0.2996	0.4049	49
7	NLP-07		39.1255	0.4897	0.2878	0.3963	79
8	NLP-08		38.9980	0.4943	0.2855	0.3902	15



04

# 자체 평가의견

# 자체 평가의견

: Dialogue Summarization

## Review

- 아쉬운 점 :
  - 시간 제약으로 **prediction review**까지 진행하지 못 한 점과 **T5를 완성하지 못 한** 것이 아쉽다.
  - 초반 베이스라인 모델의 42.1005 보다 높은 점수를 달성하기 어려웠다. (**아웃라이어**를 빼도 점수 오히려 **하락**)
  - **OOM** 때문에 많은 실험을 시도하기 어려웠다. `OutOfMemoryError: CUDA out of memory.`
  - Back translation을 이용한 **데이터 증강을 시도**하였으나 제약으로 인하여 최종 모델에 합쳐 학습시키지 못 하였다.
- 잘 해결한 점 :
  - **Quantization 성공!**
  - GPU 및 서버 등 **한정된 하드웨어 자원**에 대해 고민하는 계기를 가지게 되었다.

Life-Changing Education

감사합니다.

---