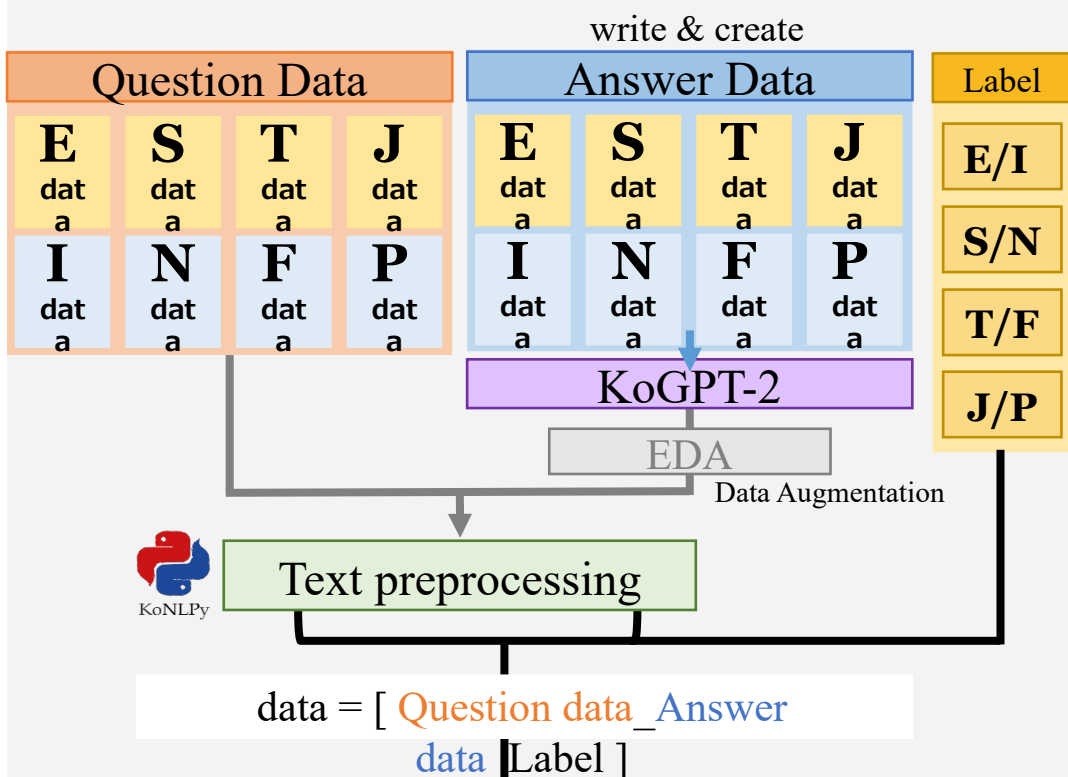


종합 설계

(데이터 수집)

Dataset



Transformer

Encoder

4-BINARY CLASSIFIERS Model

E/I
mode
I

S/N
mode
I

T/F
mode
I

J/P
mode
I

Answer Analysis Model



KoNLPy
(Morpheme Analysis)

Word Embedding Vector
values for each Key-word

Sentiment
Analysis

LSTM

Frequency

Modifier
(adverbs)

UI

Bot

User

E/I Question

Answer Analysis
Model

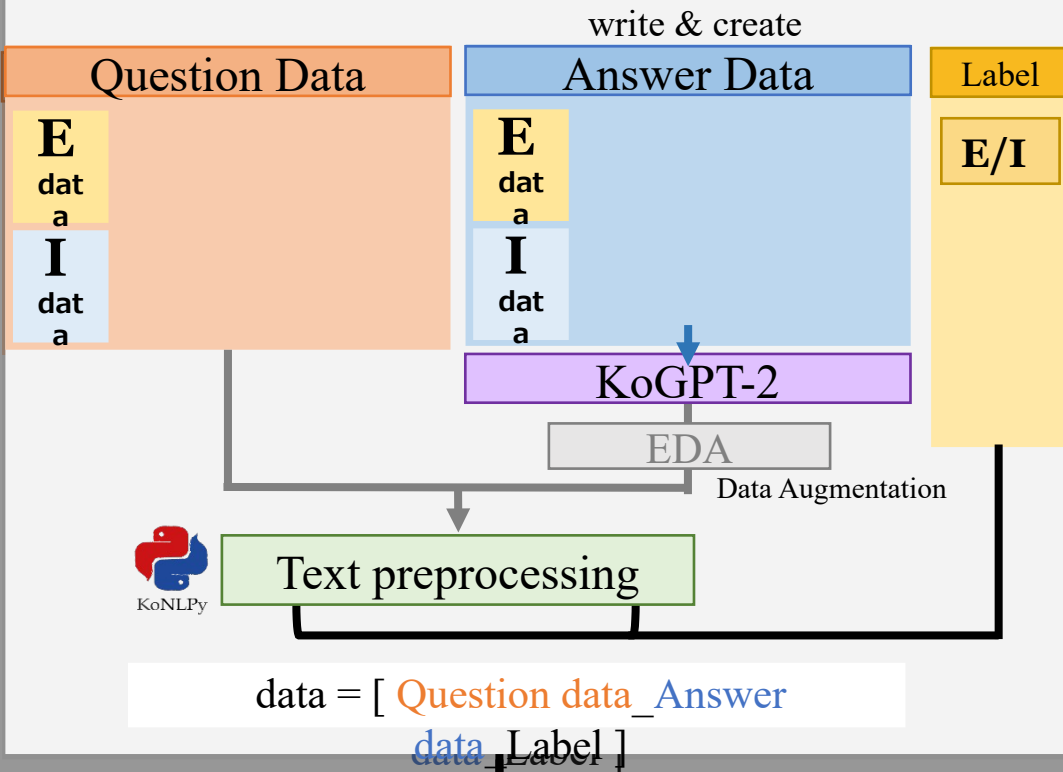
Identifying
a single MBTI type

User's Answer

Output

: User's MBTI

Dataset



Transformer

Encoder

4-BINARY CLASSIFIERS Model



Answer Analysis Model



KoNLPy
(Morpheme Analysis)

Word Embedding Vector
values for each Key-word

Sentiment
Analysis

LSTM

Frequency

Modifier
(adverbs)

UI

Bot

User

E/I Question

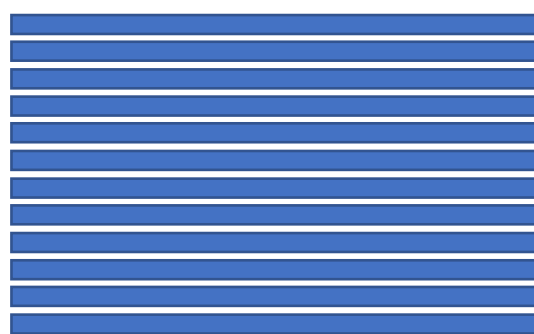
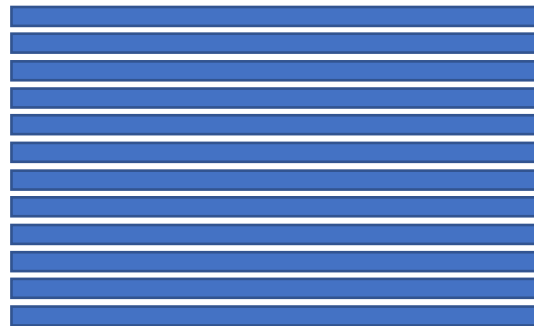
Answer Analysis
Model

Identifying
a single MBTI type

User's Answer

Output
: User's MBTI

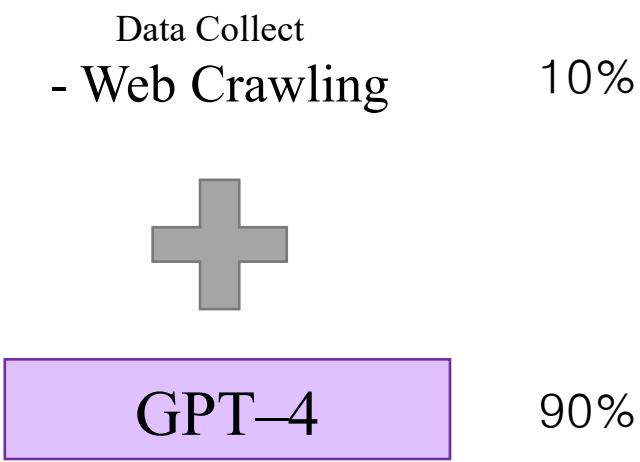
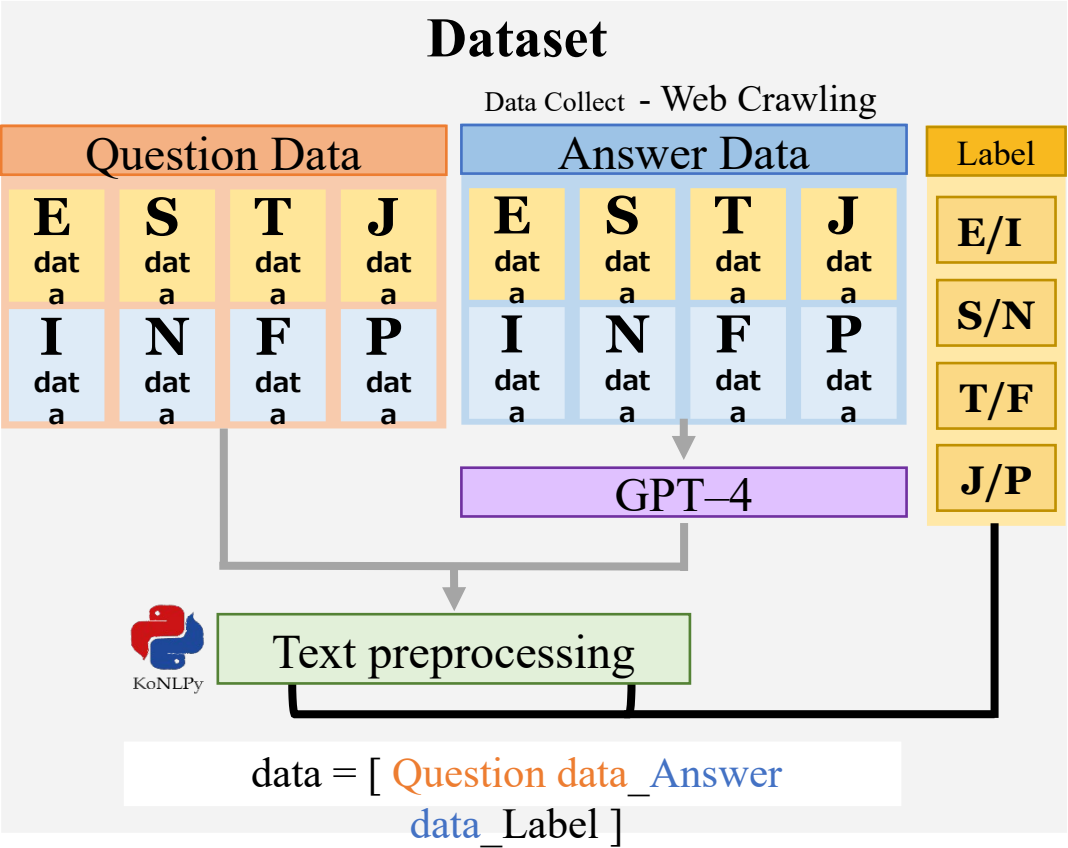
data = [Question data _ Answer data _ Label]



질문	mbti 분류	어떤 유형?
주기적으로 새로운 친구를 만든다.	E,I	E
· 행사에서 새로운 사람에게 먼저 자신을 소개하기보다는 주로 이미 알고 있는 사람과 대화하는 편	E,I	I
다른 사람에게 자신이 어떤 사람으로 보일지 걱정하지 않는 편이다.	E,I	E
단체 활동에 참여하는 일을 즐긴다.	E,I	E
자신보다는 남의 성공을 돕는 일에서 더 큰 만족감을 느낀다.	E,I	E
단체에서 지도자 역할을 맡는 일은 가능한 피하고 싶다.	E,I	I
다른 사람의 주의를 끌지 않으려고 하는 편이다.	E,I	I
혼자보다는 다른 사람과 시간을 보내고 싶어한다.	E,I	E
혼자만의 시간을 보내기보다는 즐거운 파티와 행사로 일주일의 피로를 푸는 편이다.	E,I	E
전화 통화를 거는 일은 가능한 피하고 싶다.	E,I	I
친구에게 먼저 만나자고 연락하는 편이다.	E,I	E
대부분의 시간을 혼자서 일할 수 있는 직업을 원한다.	E,I	I
조용하고 사적인 장소보다는 사람들로 붐비고 떠들썩한 장소를 좋아한다.	E,I	E

Q	A	label
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	친구나 가족에게 전화나 메시지를 보낸다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	SNS를 업데이트하거나 타인의 게시물을 확인한다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	온라인에서 사람들과 채팅하거나 게임을 한다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	다가오는 모임이나 파티의 계획을 세운다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	사진이나 동영상을 편집하여 SNS에 업로드한다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	주변 친구들과의 추억을 돌아보며 사진이나 동영상을 본다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	새로운 사람들을 만나기 위한 모임이나 이벤트를 찾아본다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	음악을 크게 틀어놓고 춤을 춘다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	친구들과의 과거 여행이나 이벤트의 동영상을 만든다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	다가오는 이벤트나 파티의 의상이나 준비물을 준비한다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	친구나 가족에게 보낼 선물이나 편지를 준비한다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	새로운 취미나 활동을 찾아보며 계획을 세운다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	주변 사람들과의 관계를 돌아보며 필요한 경우 연락을 취한다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	다른 사람들과의 경험을 공유하기 위한 블로그나 유튜브 콘텐츠를 제작한다	1
혼자 시간을 보낼 때 주로 어떤 활동을 하나요?	다른 사람들과의 소통을 위해 새로운 언어나 문화를 학습한다	1

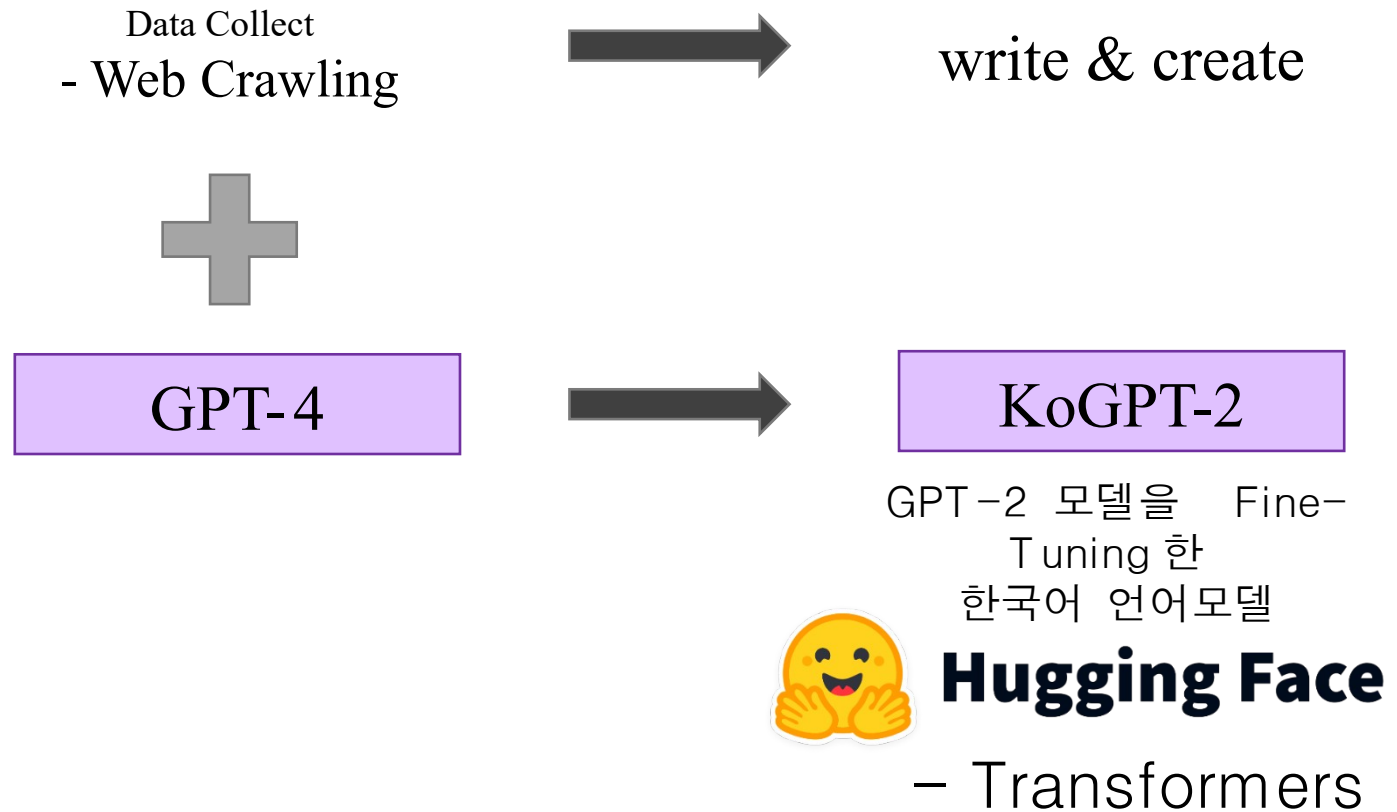
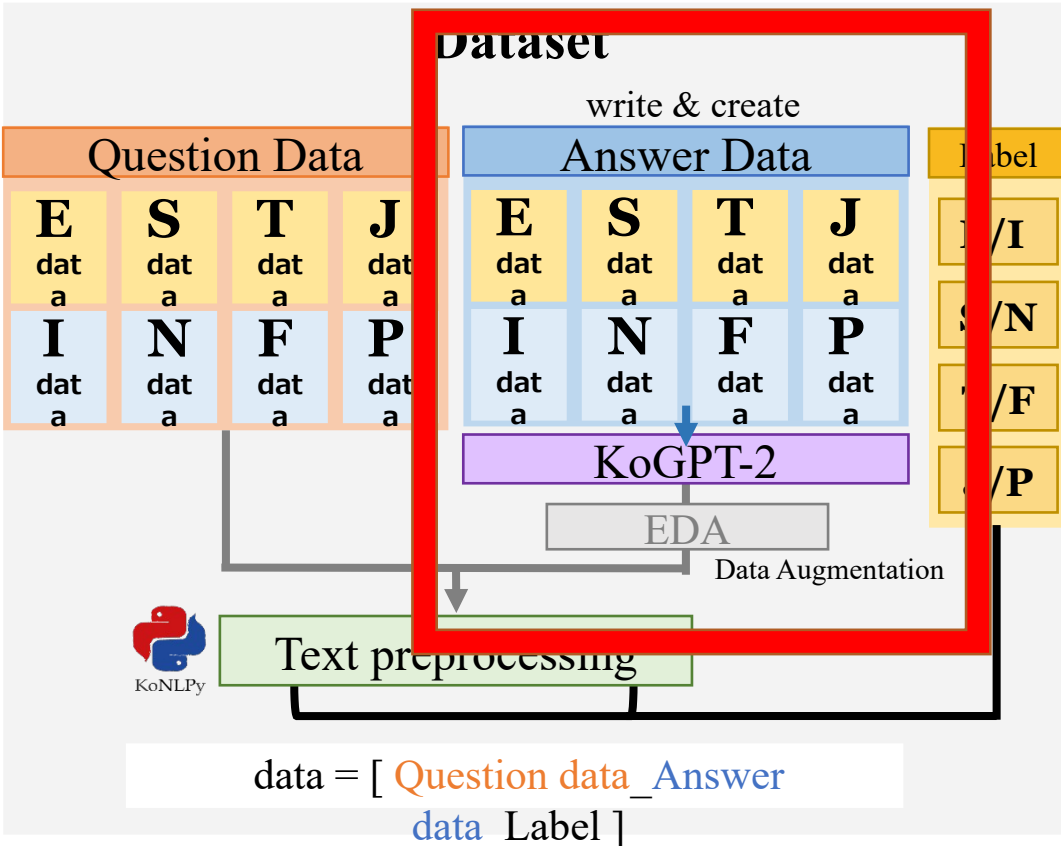
Answer 데이터 생성



GPT 사용이
훨씬 정확하고 편리함

- GPT API 비용문제
- BERT 로도 대체 불가

Answer 데이터 생성



Question Data		write & create		KoGPT-2	합계
E / I	Index	직접 작성	chat GPT		
E	Q-1	10	60	100	240
I		10	60		
E	Q-2	10	19	69	128
I		10	20		
E	Q-3	10	30	97	176
I		10	29		
E	Q-4	10	20	33	93
I		10	20		

Question

새로운 사람들을 만났을 때의 반응은 어떠한가요 ?

Q	A	label
새로운 사람들을 만났을 때의 반응은 어떠한가요?	새로운 사람을 만나면 항상 두근거린다	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	대부분의 경우 먼저 말을 걸려고 해요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	새로운 사람과의 대화는 항상 즐거워요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	처음 만나는 사람에게도 내 이야기를 편하게 공유해요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	새로운 사람을 만나면 에너지를 얻는 느낌이에요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	누군가를 처음 만났을 때 그 사람과 친해지고 싶다는 생각을 해요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	새로운 사람을 만나는 것은 새로운 기회라고 생각해요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	나는 대부분의 사람들과 잘 어울린다고 생각해요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	새로운 친구를 만드는 것은 항상 흥미로워요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	나는 쉽게 사람들과 친밀감을 느낄 수 있어요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	새로운 사람을 만나면 그 사람의 이야기를 듣는 것이 정말 흥미로워요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	나는 쉽게 웃음을 지어서 사람들과 금방 친해져요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	새로운 사람을 만나면 그 사람과의 추억을 만들고 싶어해요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	나는 다양한 사람들과의 경험을 통해 성장한다고 느껴요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	새로운 사람을 만나는 것은 나에게 새로운 모험이에요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	나는 새로운 사람들과의 관계를 맺는 것을 정말 좋아해요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	나는 새로운 사람들과의 대화를 통해 새로운 것을 배우는 것을 좋아해요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	나는 새로운 사람들과의 만남을 통해 내 세상이 확장된다고 느껴요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	나는 새로운 사람들과의 만남을 기다리는 것이 항상 설레요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	나는 새로운 사람들과의 만남을 통해 다양한 문화와 생각을 경험하고 싶어해요	1
새로운 사람들을 만났을 때의 반응은 어떠한가요?	새로운 사람을 만나면 그 사람의 취미나 관심사에 대해 알고 싶어져요	1

Question	Answer	label
새로운 사람들을 만났을 때의 반응은 어떠한가요 ?	새로운 사람을 만나면 항상 두근거린다	1
새로운 사람들을 만났을 때의 반응은 어떠한가요 ?	나는 새로운 사람들을 만나면 항상 적극적으로 대화를 나눈다	1
새로운 사람들을 만났을 때의 반응은 어떠한가요 ?	나는 누군가 나를 찾으면 곧잘 숨어버리고는 한다	0

KoGPT-2

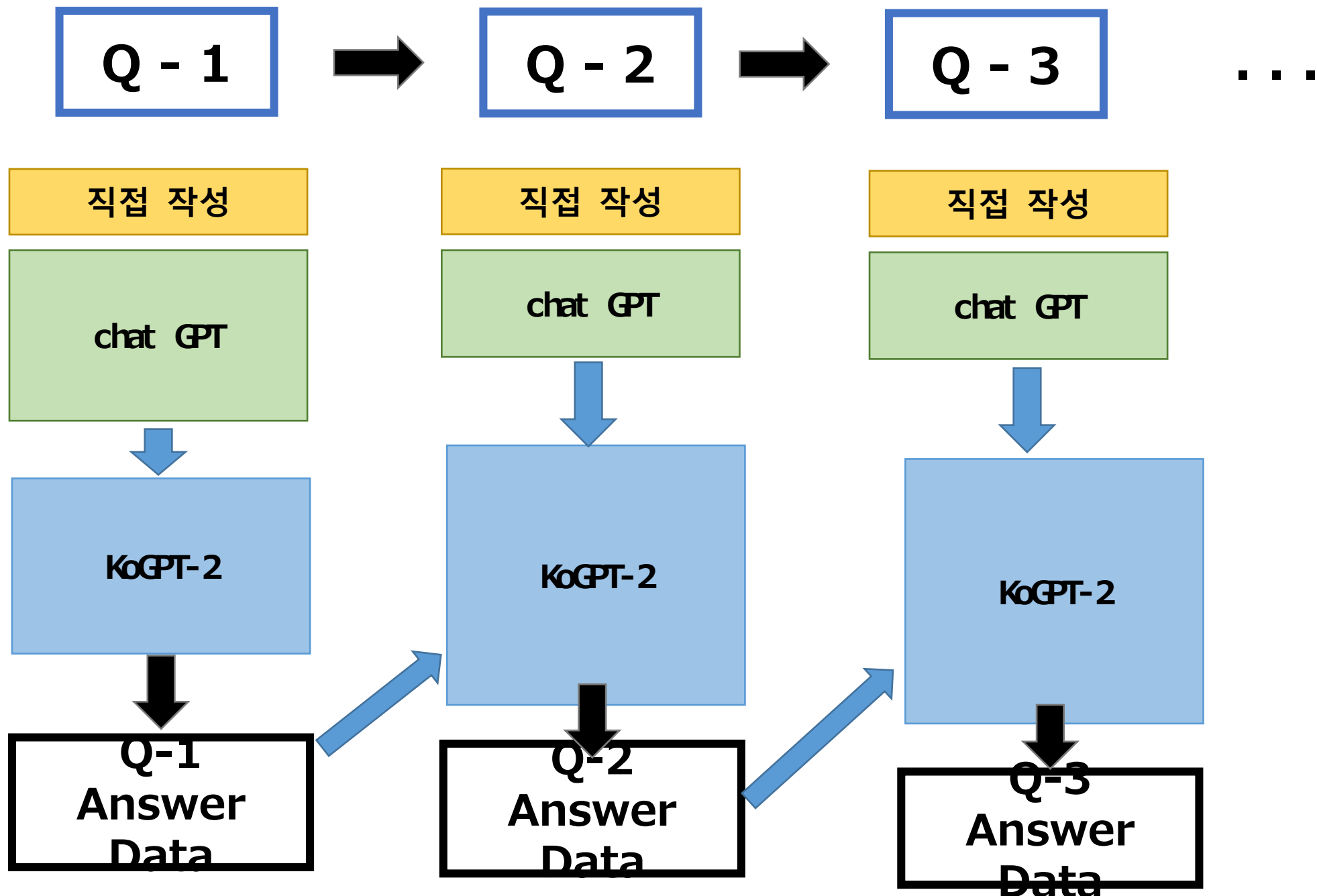
Answer_data .XLSX ☆ 📁 ☁

파일 수정 보기 삽입 서식 데이터 도구 도움말

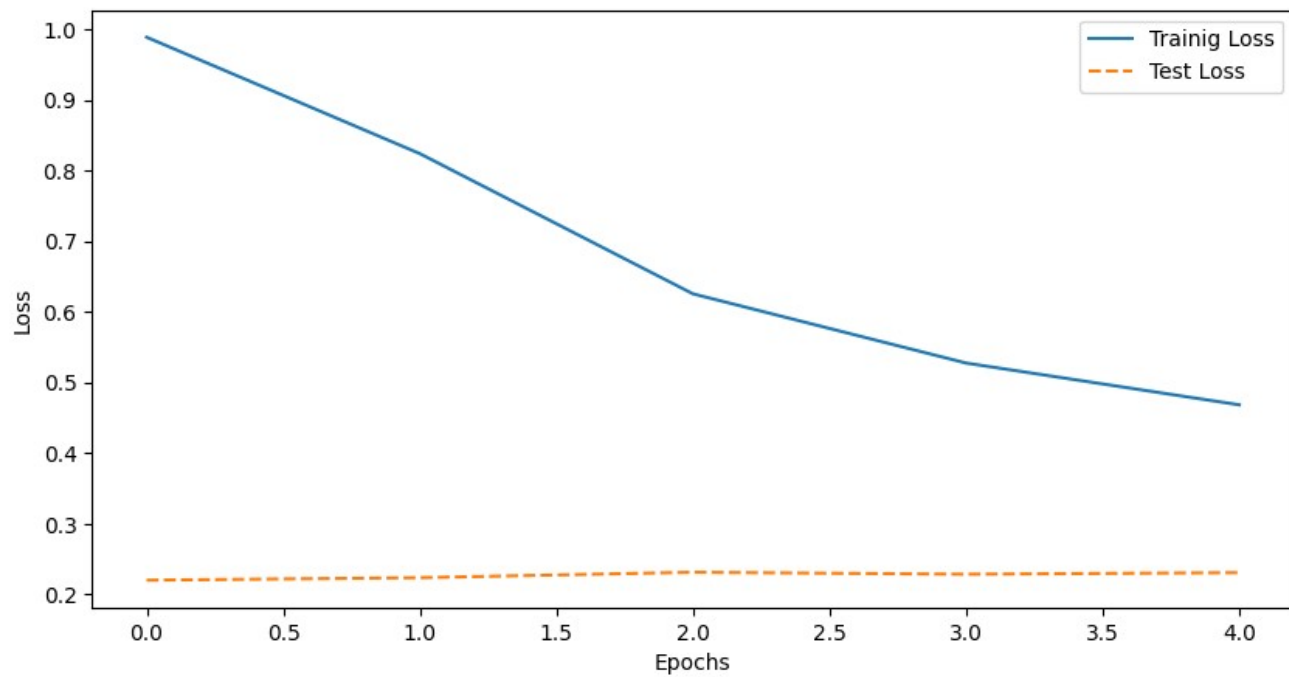
메뉴 ↶ ↷ 🖨 🗑 100% ▾ ₩ % .0_ .00 123 | 기본값... ▾ - 10 + B I 🔍 A

A	B
	Answer
0	새로운 사람을 만나는 것은 새로운 기회가 될 것 같아
1	나는 새로운 사람과의 대화를 통해 다양한 경험을 얻는다고 생각해요
2	나는 새로운 사람을 만나면 그와의 관계를 천천히 그리고 깊게 발전시키려고 해요
3	나는 새로운 사람을 만났을 때 먼저 말을 걸기보다는 상대방의 이야기에 귀를 기울이려고 해요
4	나는 새로운 사람과의 만남에서 내 감정을 쉽게 표출하지 않아요
5	나는 새로운 사람들과 만남에서 서로의 경계를 중요하게 생각해요
6	나는 새로운 사람과의 만남에서 각자의 시간을 내주신다면 그 시간을 정말 좋아해요
7	나는 새로운 사람을 만났을 때의 반응은 어떠한가요?
8	새로운 사람을 만나는 것 자체만으로도 기쁘고 기뻐요
9	새로운 사람을 만나는 것은 항상 특별한 경험을 얻는 것으로 생각해요
10	나는 새로운 사람과의 관계를 맺는 것을 좋아해요
11	새로운 사람을 만나면 그와의 대화보다는 서로의 존재를 느끼는 것이 흥미로워요
12	나는 새로운 사람을 만나면 그 사람을 만나면 그 사람과의 관계를 천천히 발전시키고 싶어해요
13	나는 새로운 사람과의 만남에서 서로의 경계를 느끼곤 해요
14	새로운 사람을 만나면 먼저 말을 걸기보다는 상대방의 이야기를 경청하는 편이다
15	새로운 사람을 만나도 조금은 긴장되지만 조금은 편안하게 대화하고 싶어
16	나는 새로운 사람과의 만남에서 깊은 대화를 통해 새로운 것을 경험하는 것을 좋아해요
17	나는 새로운 사람과의 대화가 많은 사람이에요
18	새로운 사람을 만나면 그 사람에게서 배우는 것이 너무 특별해요
19	나는 새로운 사람과의 만남에서 대화보다는 휴식을 중요하게 생각해요
20	새로운 사람을 만나면 그들과의 관계를 천천히 그리고 깊게 발전시키려고 해요
21	나는 새로운 사람과의 만남에서 서로의 경계를 느끼곤 합니다
22	그들에 대해 조금 더 대화를 나누고 싶어
23	처음 만나면 조금은 긴장되기도 해요
24	나를 찾았을 때의 반응은 어떠한가요?
25	나는 새로운 사람을 만나면 그와의 공통점을 찾는 것을 무척 좋아해요
26	나는 새로운 사람과의 관계에서 친밀감을 느낀다
27	나는 새로운 사람과의 대화가 일상의 일부로 생각해요
28	나는 새로운 사람과의 만남을 통해 새로운 것을 경험한다
29	나는 새로운 사람들과의 첫 만남에서 대화하는 것을 매우 좋아해요

Answer Data
조금씩 생성



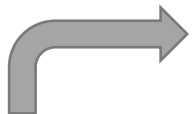
Question Data 4 개 ➡ Answer Data 543 개



1) Loss

2) 실제 성능
평가

~~Accuracy~~



“ 새로운 사람들을 만났을 때의 반응은
어떠한가요 ?”

```
return_answer_by_chatbot('낯선 사람을 만났을 때 당신의 반응은?')
```

```
'새로운 사람을 만나면 그와의 대화보다는 각자의 시간을 갖기를 원한다'
```

```
return_answer_by_chatbot('자신의 생일이나 특별한 날을 어떻게 보내고 싶은가?')
```

```
'사람들과 함께 있을 때 더욱 활기차고 흥미진진해지는 것 같다'
```



X

1) Loss

2) 실제 성능
평가

~~Accuracy~~

EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks

Jason Wei^{1,2} Kai Zou³

¹Protago Labs Research, Tysons Corner, Virginia, USA

²Department of Computer Science, Dartmouth College

³Department of Mathematics and Statistics, Georgetown University

jason.20@dartmouth.edu kz56@georgetown.edu

: 자연어 처리에서의 데이터 증강 기법

- 1) SR : 유의어로 교체 (stopwords 가 아닌 n 개의 단어를 선택해 임의의 단어로 교체)
- 2) RI : 랜덤 삽입
- 3) **RS : 랜덤 교체**
- 4) **RD : 랜덤 삭제**

Label Answer

1 처음 만나는 사람에게도 내 이야기를 편하게 공유해요

원본 문장

1 처음 내 사람에게도 **만나는** 이야기를 편하게 공유해요

유의어로 교체

1 처음 만나는 사람에게도 (**내**) 이야기를 편하게 공유해요

랜덤 삭제

1 처음 만나는 사람에게도 내 **편하게** **이야기를** 공유해요

랜덤 교체

Model	Training Set Size			
	500	2,000	5,000	full set
RNN	75.3	83.7	86.1	87.4
+EDA	79.1	84.4	87.3	88.3
CNN	78.6	85.6	87.7	88.3
+EDA	80.7	86.4	88.3	88.8
<i>Average</i>	76.9	84.6	86.9	87.8
+EDA	79.9	85.4	87.8	88.6

Table 2: Average performances (%) across five text classification tasks for models with and without EDA on different training set sizes.

소규모 데이터 셋 일 때
가장 큰 성능향상

한 문장 당 4 개 이하의 EDA 문장
추가

2000 장 x 3(EDA) = 6000 장

영어기준 -> 한국어에 적용시 의미
변형 多

사전 훈련 모델의 경우 효과 미미

학습모델

data = [Question data _ Answer data _ Label]

