

종합 설계

(주제 선정 및 소개)

Type Talk: MBTI

Analysis Using Chat

MBTI

ENTJ



ESFP



ENFP



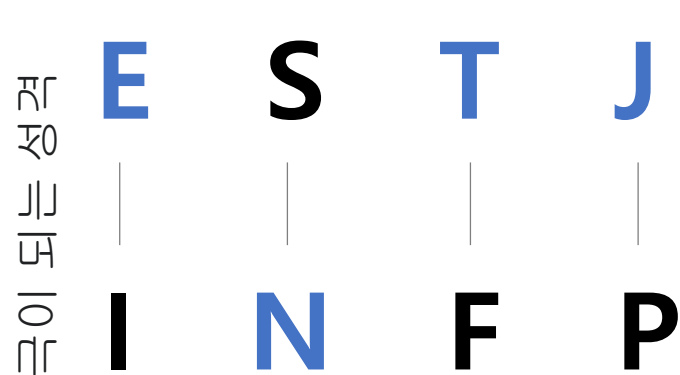
INTJ



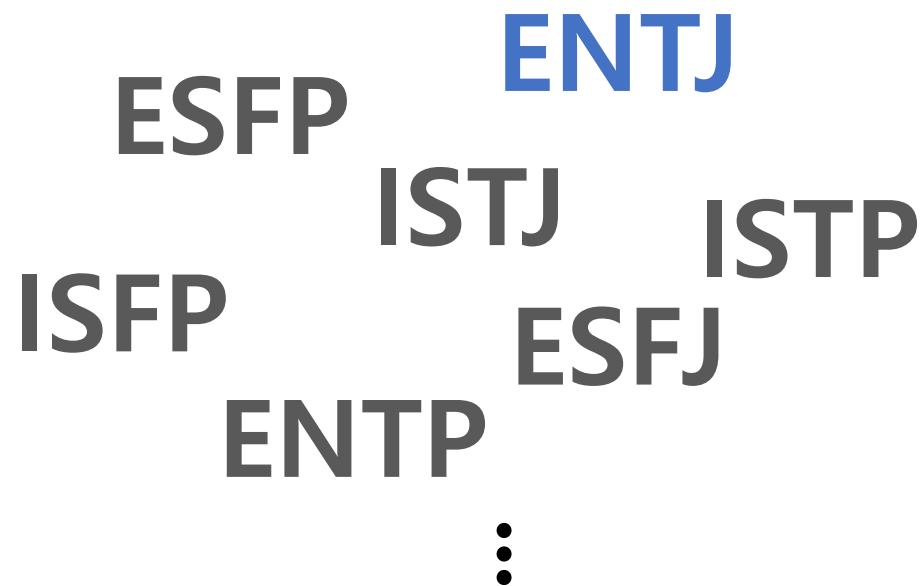
MBTI (Myers-Briggs Type Indicator)

: 분석심리학을 바탕으로 만들어진 **성격 유형 검사**

(자기보고형 성격 유형 검사로, 사람의 성격을 16가지의 유형으로 나누어 설명)



네 가지 척도로 성격을 표시



4개의 문자로 구성된 개인의 성격유형

I ntroversion (내향)		E xtroversion (외향)		iN tuition (직관)		S ensing (감각)		F eeling (감정)		T hinking (사고)		P erceiving (인식)		J udging (판단)	
주의 초점 (에너지의 방향)		인식기능 (사람이나 사물 인식 방식)		판단기능 (판단의 근거)		생활양식 (선호하는 삶의 패턴)									
내면 세계		세상과 타인		이상주의적 (영감)		현실적인 타입(경험)		관계와 사람 위주		사실과 진실 위주 (논리, 객관)		즉흥적인 생활		계획적인 생활	
새로운 사람을 보면?		사과하면 떠오르는 것은?		나 이번에 시험 떨어졌어		여행 가기 전 준비할 것은?									
일단 가만히 있는다.		먼저 다가가서 말을 건다.		백설공주, 뉴턴		빨간색, 맛있다		어떡해,, 괜찮아		뭔 시험?		그냥 가서 정해		1시간단위로 세세하게 계획짜기	



16가지 유형
으로 분류 가능

유형		T		F	
		J	P	J	P
I	S	ISTJ	ISTP	ISFJ	ISFP
	N	INTJ	INTP	INFJ	INFP
E	S	ESTJ	ESTP	ESFJ	ESEP
	N	ENTJ	ENTP	ENFJ	ENFP

57% 외향형 E

57%

외향형

43%

내향형

외향형의 사람들은 단체 활동을 즐기고 사회적 교류를 중시합니다. 이들은 외부로 열정을 표출하고 자신의 즐거움을 표현합니다.

57% 사고형 T

57%

사고형

43%

감정형

사고형의 사람들은 객관성과 합리성을 중시하며 논리에 집중하느라 감정을 간과할 때가 많습니다. 이들은 사회적 조화보다는 효율성이 더 중요하다고 생각하는 경향이 있습니다.

성격 유형:

경영자 ESTJ-T



71% 현실주의형 S

29%

직관형

71%

현실주의형

관찰형의 사람들은 실용적이고 현실적인 성격을 지니고 있습니다. 이들은 현재 발생하고 있거나 발생할 가능성이 매우 높은 일에 집중하는 경향이 있습니다.

54% 계획형 J

54%

계획형

46%

탐색형

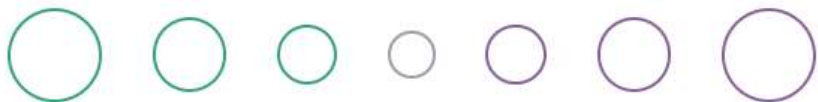
계획형의 사람들은 결단력이 높고 철저하며 계획적인 성격을 지니고 있습니다. 이들은 명확성과 예측 가능성을 중시하고 어떤 일을 제대로 끝내고자 하며 즉흥적인 일보다는 체계적으로 계획을 세우는 일을 선호합니다.

무료 성격유형검사

NERIS Type Explorer®

주기적으로 새로운 친구를 만든다.

동의



비동의

자유 시간 중 상당 부분을 다양한 관심사를 탐구하는 데 할애한다.

동의



비동의



1/28

다른 사람들에게 자신을 소개하는 것이 어렵습니다.

그렇다

아니다

Q. “혼자보다는 다른 사람과 시간을 보내고 싶어 하는가?”

기존 검사

혼자보다는 다른 사람과 시간을 보내고 싶어한다.

동의 ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ 비동의

1/28

다른 사람들에게 자신을 소개하는 것이 어렵습니다.

☐ 그렇다

☐ 아니다

객관식

단 1개의 답만 선택 가능

개선된 검사

Show me your **MBTI**

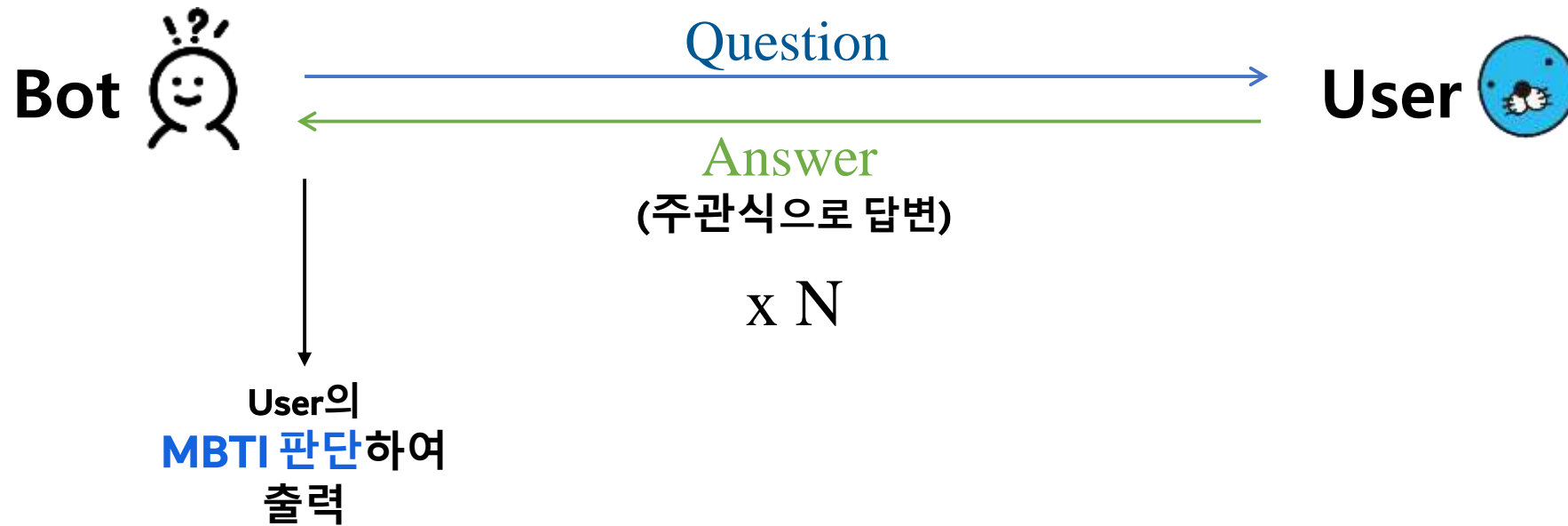
그렇지 않다. 혼자 있는 걸 더 선호하는 편이다.
혼자 있는 게 아무래도 더 편하다고 생각한다.
독립적인 게 가장 편하다.
다른 사람과 있으면 불편함을 평소 느끼는 편이다.
친구들과 약속 잡기가 귀찮아서 집에 주로 있는 편이다.
그렇다. 친구들과 노는 것이 가장 즐겁다.
아무래도 타인과 있을 때 에너지를 얻는다.
친구들이 내 삶의 활력소이다.

...

주관식

무한가지의 답 작성 가능

chatting for MBTI - TEST



Show me your MBTI

Bot 

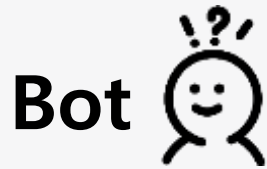
User 

풍선을 보면 떠오르는 생각은?

풍선 안에 들은 헬륨가스를 마셔보고 싶다..

시험이 다가오는데 공부가 손에 안 잡힌다면, 무슨 생각을 하는가?

망했다. 이번에도 벼락치기를 해야겠네!!
인터넷으로 벼락치기 방법을 찾아본다.



Bot

Show me your
MBTI



User

사용자에게 질문

질문에 대한 응답


사용자의 응답으로 <MBTI 분석>

사용자에게 질문

질문에 대한 응답

·
·
·

User의 MBTI가
판별될 때까지 반복해서 질
문

Show me your 
MBTI

⋮

당신의 MBTI는 **ESTJ** 입니다.

Show me your MBTI

당신의 MBTI는 **ESTJ** 입니다.



엄격한 관리자

ESTJ-A / ESTJ-T

사물이나 사람을 관리하는 데 타의 추종을 불허하는 뛰어난 실력을 갖춘 관리자형.

순서	기능
주기능	Te (외향 사고)
부기능	Si (내향 감각)
3차기능	Ne (외향 직관)
4차기능	Fi (내향 감정)



엄격한 관리자
ESTJ 연예인



한채영



이지혜



대프콘



김민규



JYJ
김준수



GOT7
뽐뽐



AB6IX
박우진



모모랜드
제연



ESTJ 사업가형

체계적으로 일하고 성과지향적인 이들은
엄격한 관리자입니다.
옳다고 생각하는 일은 거침없이 밀고 나가는
추진력도 가지고 있어요.
감정이나 느낌이 아닌
합리적 이성으로 행동하며
사업체나 단체를 이끌어가는 재능이 있어요.

추천 직업:
프로젝트 매니저, 보험설계사, 변호사, 약사 등

<ESTJ 공합 차트>



딥러닝 기반의 MBTI 성격유형 분류 연구

김정민, 박지민, 이로운, 조서원, 심재형
이화여자대학교 컴퓨터공학과

[표1] 기존 MBTI 분류 연구와의 접근 방식 및
정확도 비교

	접근방법	분류 모델	분류언어	데이터양	전체정확도
본 논문	binary	LSTM	한국어	33,144	20.29%
관련 논문	[2] binary	LSTM	영어	430,000	21%
	[3] binary	LSTM	영어	422,845	38%
	[4] binary	Linear/Logistic Regression	영어	17,640,062	45%
	[5] multiclass	LR	영어	8676 (frequency of data set)	66.59%

★
pandora
(Reddit)

kaggle

LSTM 모델을 이용한 감정분류

E/I 유형 65.33%
N/S 유형 66.88%
F/T 유형 66.07%
P/J 유형 70.32%

전체 정확도 20.29%

Survey Analysis of Machine Learning Methods for Natural Language Processing for MBTI Personality Type Prediction

Our best performing models were as follows (Table 6):

Model Type	Train Accuracy	Test Accuracy
Softmax (baseline)	19%	17%
Naive Bayes	32%	26%
Regularized SVM	34%	33%
Deep Learning	40%	38%

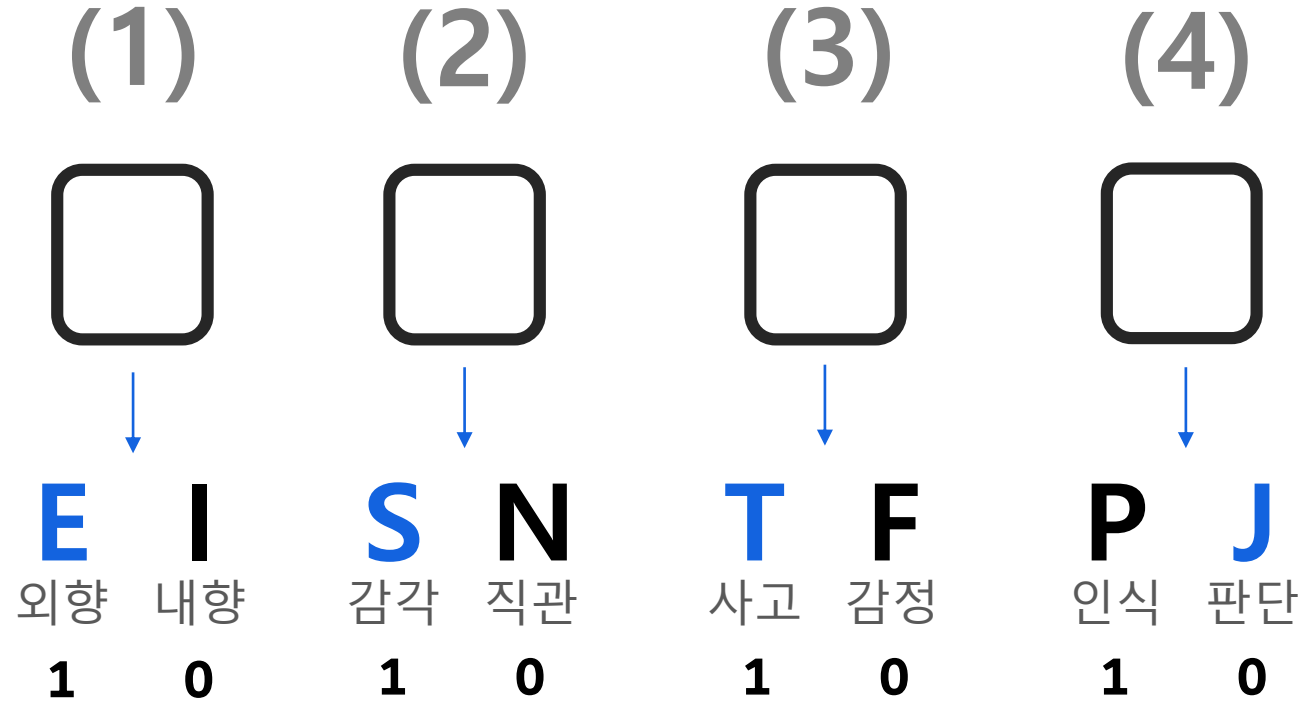
Table 6: Comparison of different methods for MBTI-classification

16-CLASS CLASSIFIER	Training Accuracy	55%
	Test Accuracy	23%
4 BINARY CLASSIFIERS	Training Accuracy	40%
	Test Accuracy	38%

유형별 각 모델들보다, 전체 모델의 정확도가 낮아지는 문제점

-> 미해결

(클래스 불균형, 데이터 품질 등이 원인으로 예상..)
/ MBTI 검사의 한계 (낮은 신뢰성)



**4 Binary
Classifiers**

E S T J

Dataset

Question Data

E data	S data	T data	J data
I data	N data	F data	P data

write & create

Answer Data

E data	S data	T data	J data
I data	N data	F data	P data

Label

E/I
S/N
T/F
J/P

KoGPT-2

EDA

Data Augmentation

Text preprocessing

data = [Question data _ Answer data _ Label]

Transformer

Encoder

Answer Analysis Model



KoNLPy
(Morpheme Analysis)

Word Embedding Vector
values for each Key-word

Sentiment
Analysis

LSTM

Frequency

Modifier
(adverbs)

4-BINARY CLASSIFIERS Model

E/I
model

S/N
model

T/F
model

J/P
model

UI

Bot

User

E/I Question

Answer Analysis
Model

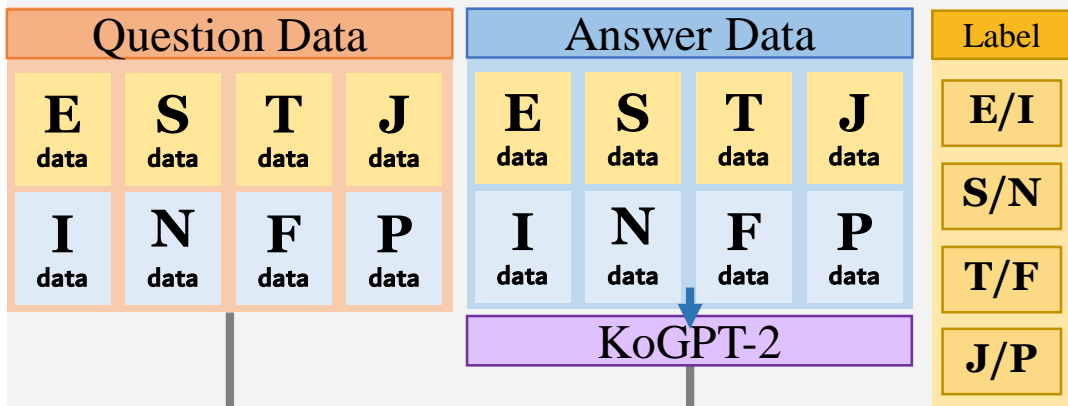
Identifying
a single MBTI type

Output

: User's MBTI

User's Answer

Dataset



Text preprocessing

data = [Question data_Answer data_Label]

KoGPT-2
Data Augmentation

Transformer

Encoder

4-BINARY CLASSIFIERS
Model

E/I
model

S/N
model

T/F
model

J/P
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

Dataset

Question Data

E data	S data	T data	J data
I data	N data	F data	P data

write & create

Answer Data

E data	S data	T data	J data
I data	N data	F data	P data

Label

E/I
S/N
T/F
J/P

KoGPT-2

EDA

Data Augmentation

Text preprocessing

data = [Question data _ Answer data _ Label]

Transformer

Encoder

Answer Analysis Model



KoNLPy
(Morpheme Analysis)

Word Embedding Vector
values for each Key-word

Sentiment
Analysis

LSTM

Frequency

Modifier
(adverbs)

4-BINARY CLASSIFIERS Model

E/I
model

S/N
model

T/F
model

J/P
model

UI

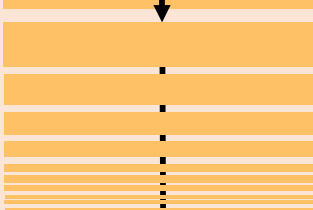
Bot

User

E/I Question

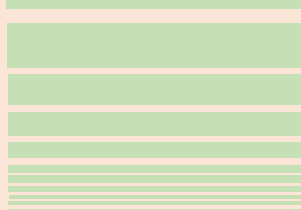
Answer Analysis
Model

Identifying
a single MBTI type

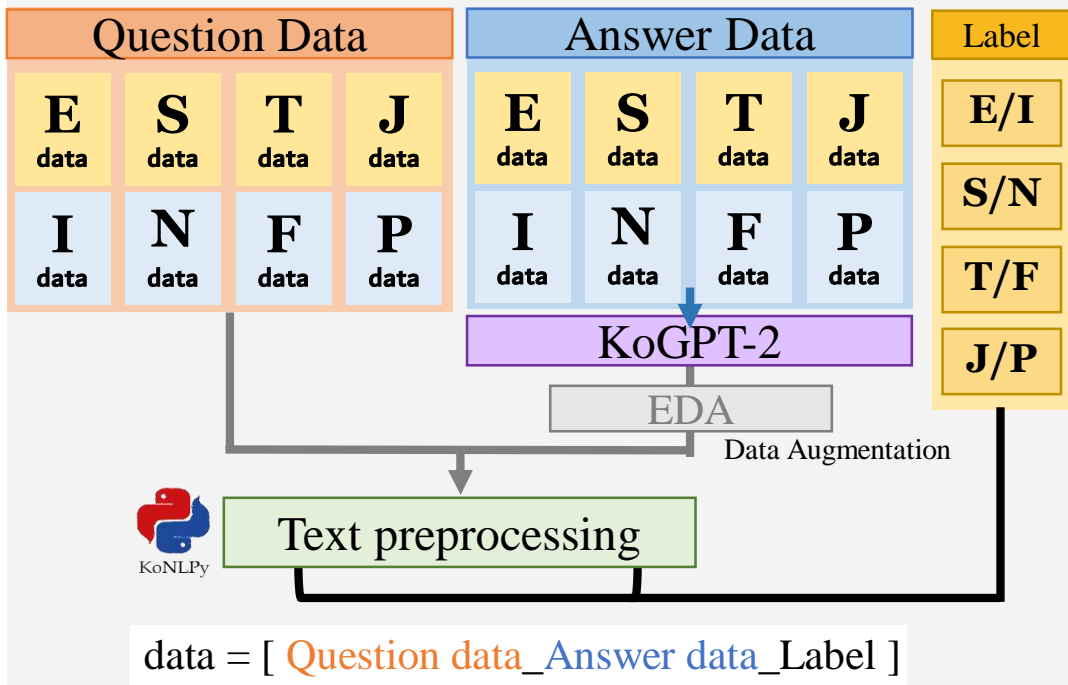


Output
: User's MBTI

User's Answer



Dataset



Transformer

Encoder

4-BINARY CLASSIFIERS Model

E/I
model

S/N
model

T/F
model

J/P
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

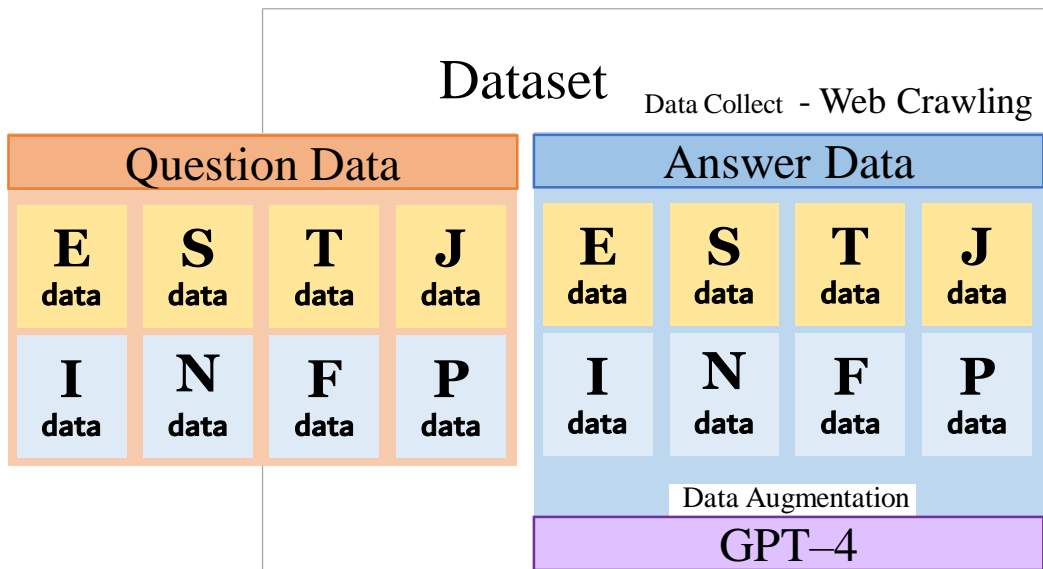
Answer Analysis
Model

Identifying
a single MBTI type

Output

: User's MBTI

User's Answer



Text preprocessing

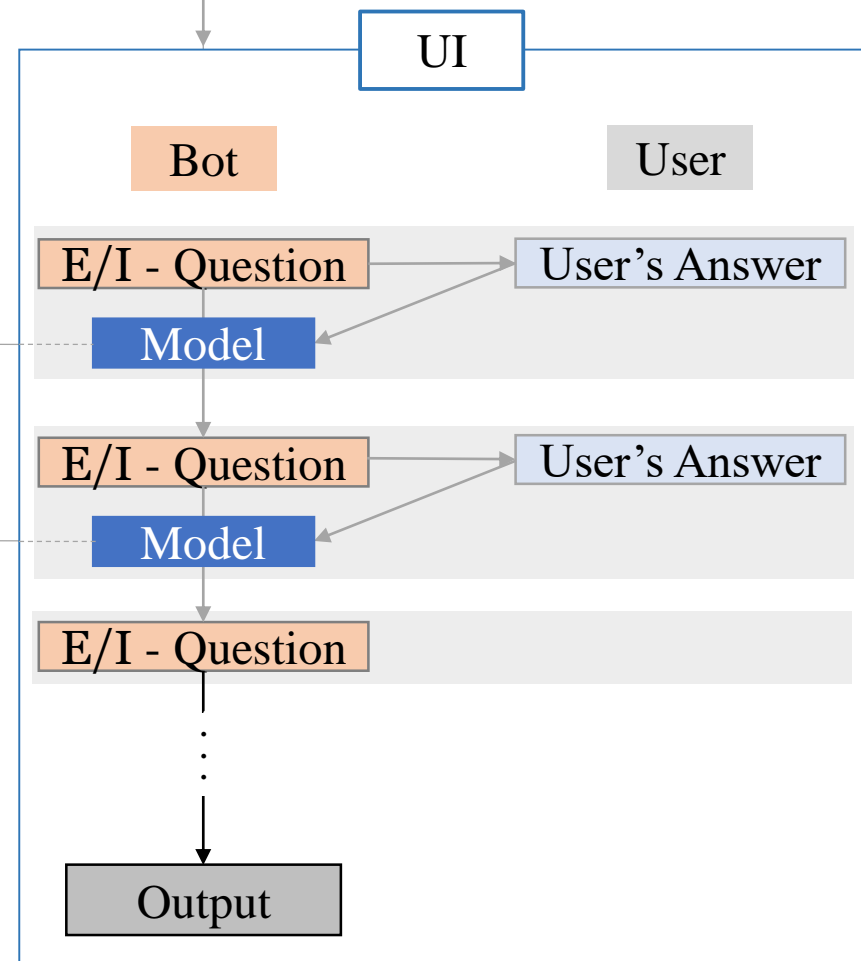
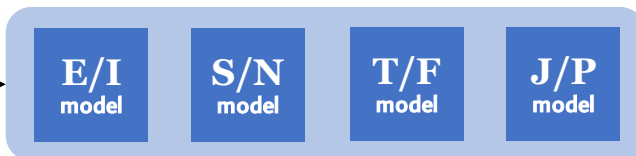
< Question data_Answer data_label >

Transformer

Word2Vec-ko,
FastText-KR, KoBERT

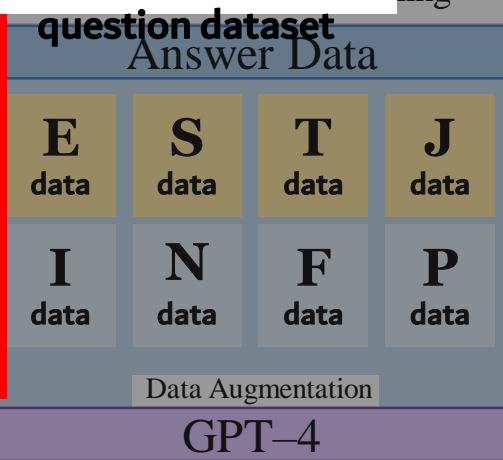
Encoder

**4-BINARY CLASSIFIERS
Model**



Mechanism

: 다양한 질문을 하



Transformer 모델

Word2Vec-ko,
FastText-KR, RoBERTa

“Answer” 분석

E/I
model

S/N model

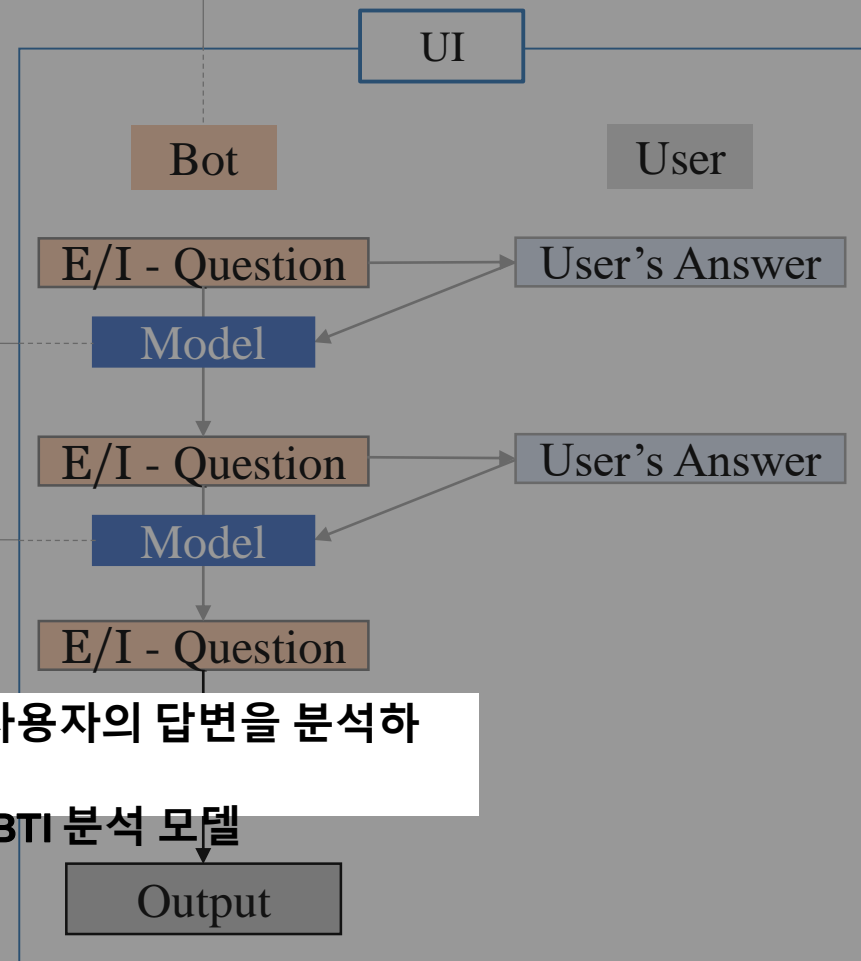
T/F
model

J/P
model

: 사용자의 답변을 분석하

MBTI 분석 모델

Output





“Question” data

1. Bot의 질문은 사전에 준비

- 사전에 질문들은 모두 labeling
- 랜덤하게 선정하여 질문

2. 특별한 상황을 가정한 질문 일부 추가

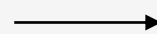
- 직접 작성
- GPT 사용 **미정**

3. 기본 질문 자동 생성 **미정**

- koGPT를 이용한 자동질문 생성

1. 사전에 준비된 질문

혼자보다는 다른 사람과 시간을 보내고 싶어 하는가?



E
문

vs

I

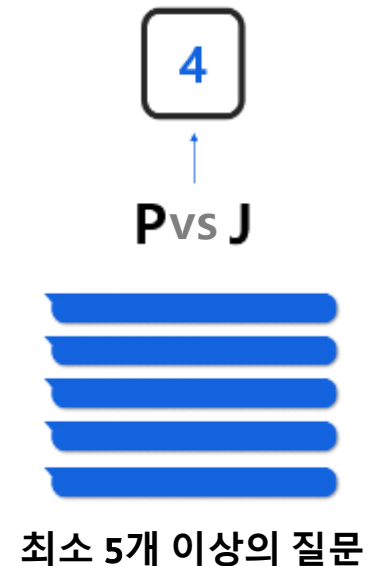
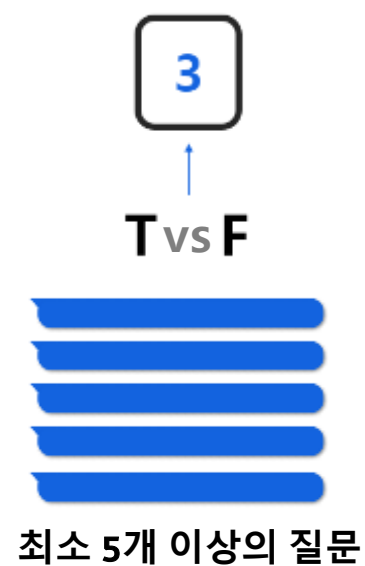
를 판단하기 위한 질

-> 사전에 질문 '유형' 라벨링

Question Data

Index	Question	Label_1
1	혼자보다는 다른 사람과 시간을 보내고 싶어 하는가?	E/I
2	사람들과 만나는 약속이 끝난 후 집에 들어와 혼자 있을 때, 나의 기분은?	E/I
3	친한 친구가 근처에 아는 지인이 있는데 불러도 되는지 묻는다면?	E/I
4	오랜만에 참석한 동창회. 나는 뭐하고 있을까?	E/I
...		
45	친구와 여행중인 당신. 전날 밤을 새서 너무 피곤하다. 하지만 아직 일정은 남아있다. 당신의 선택은?	J/P

사용자에게



•
•
•

총 20개 이상의 질문

1
E I

30개씩 준비

2
S N

30개씩 준비

3
T F

30개씩 준비

4
P J

30개씩 준비

총 약 120개의 질문

Bot은 랜덤하게 질문
-> User의 답변을 통한
MBTI 판별이 정확해질 때 까지

2. 특별한 상황을 가정한 질문

상 황

질 문

- 1) 사전에 준비한 질문
- 2) KoGPT를 이용한 질문 및 상황 자동생성

이번에는 상황을 한번 가정 해볼
게~

당신은 <포켓몬스터>의 주인공 '지우' 가 되었다.
이 포켓몬스터 세계는 그리 호락호락하지 않다.
당신은 포켓몬스터를 몬스터 볼로 잡고 키워가며, 다른 포켓몬 트레이너들과 전투를 벌여
야 한다.

이것이 이 <포켓몬스터> 세계에서야 생존 방식이다.

포켓몬 트레이너가 된 당신.
전원이 모인 입단식에서 신입생 가운데 3명을 뽑는
가벼운 실력 테스트를 진행한다고 한다. 당신의 반
응은?

(단절되는 질문)

아무리 연습해도 포켓몬을 다루는 능력
이 부족한 것 같다.. 자꾸만 전투 도중에
포켓몬이 도망가버린다. 당신의 선택은?

E 내 실력을 보여줄 좋은 기회! 가장 먼저 손들어 지원
하다.

I 아이 부끄러워.. 괜히 나섰다가 못하면 어떡해. 일단
다른 신입생들의 반응을 살핀다.

E 처음보는 얼굴인 옆자리 웅이한테 코칭을 부탁
하다.

I 난 독학이 좋아. 혼자서 이전 시합 영상을 돌려
본다.

이번에는 상황을 한번 가정 해볼게~

당신은 <포켓몬스터>의 주인공 '지우'가 되었다.

이 포켓몬스터 세계는 그리 호락호락하지 않다.

당신은 포켓몬스터를 몬스터 볼로 잡고 키워가며, 다른 포켓몬 트레이너들과 전투를 벌여야 한다.

이것이 이 <포켓몬스터> 세계에서야 생존 방식이다.

포켓몬 트레이너가 된 당신.
전원이 모인 입단식에서 신입생 가운데 3명을 뽑는 가벼운 실력 테스트를 진행한다고 한다. 당신의 반응은?

E 내 실력을 보여줄 좋은 기회! 가장 먼저 손들어 지원하다.

I 아이 부끄러워.. 괜히 나섰다가 못하면 어떡해. 일단 다른 신입생들의 반응을 살핀다.

앗차차! 나는 조무래기에 불과했다.. 생각보다 다들 잘하는 걸? 다시 자리로 돌아갈까?

답변에 따른
질문 자동생성

웬걸? 다 나와 같은 마음인가보다. 아무도 나서지 않네? 지금 나가면 우승할 수 있는 찬스이지 않을까?

1

E I

2

S N

3

T F

4

P J

상

1개의 성격 유형을
판별할 수 있는 질문 여러 개
(최소 10개 이상)

→ 약 3~5개의 상황

→ 약 3~5개의 상황

→ 약 3~5개의 상황

→ 약 3~5개의 상황

‘일반 질문’ 과 ‘상황 질문’
혼합해서 질문

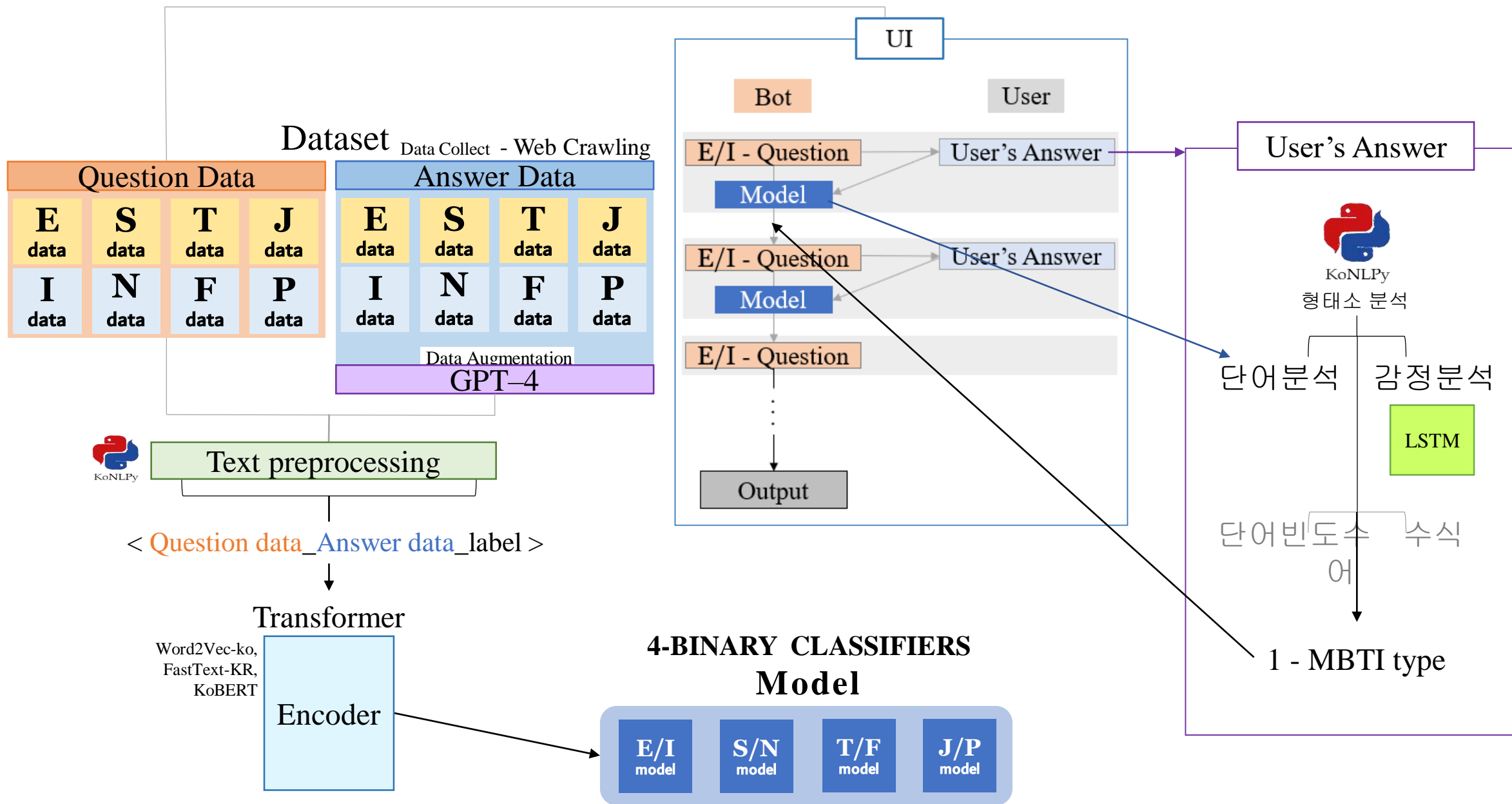


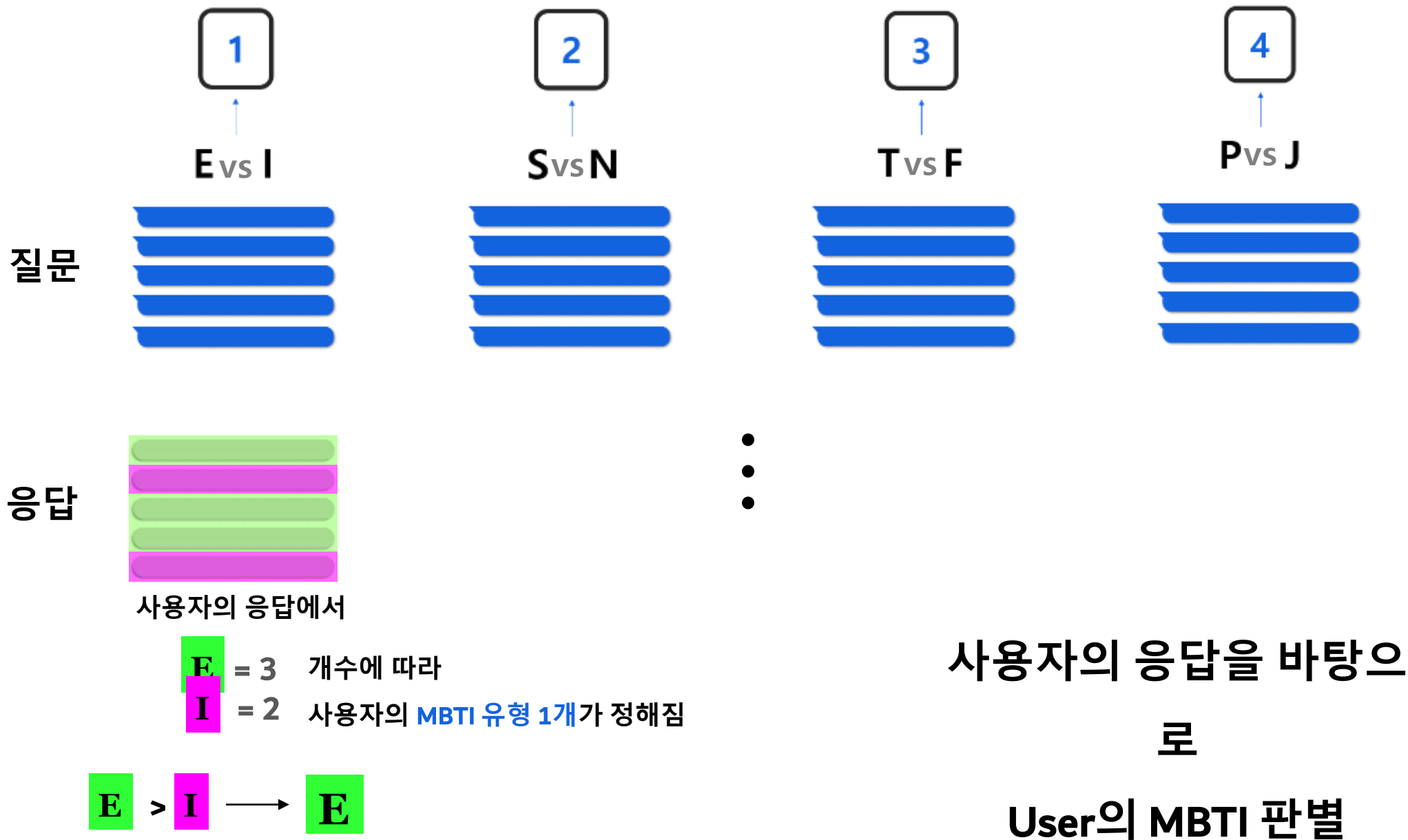
“Answer” 분석 model

유형 1) **일반적인** 응답

유형 2) **중립적**으로 대답

유형 3) **완전히 다른** 대답 (else)





E

I

혼자보다는 다른 사람과 시간을 보내고 싶어하는가?

그렇다. 친구 랑 노는 게 더 재밌다 E

나는 그냥 혼자 있는 게 더 좋아 I

난 보통 독립적으로 있는 편이 I

사색을 하는 것을 선호하는 편이 I

몰라 나도? ?

개수에 따라 E = 1
사용자의 MBTI 유형 1개가 정해 I = 3

집
E < I → I

혼자보다는 다른 사람과 시간을 보내고 싶어
하는가?

유형 1) 일반적인 응답

판단 기준

1) 단어의 임베딩 벡터 값

E응답, I응답 구분 → 워드 임베딩 도구,
Transformer의 워드 임베딩 활용

2) 감정분석

긍정, 부정 구분 → LSTM 사용

3) 단어 출현 빈도수

가중치 부여

4) 수식어 - 부사

가중치 부여

학교에 특별한 전학생이 왔다. 당신의 반응은?

나는 평소 조용함을 즐기는 편이지만,
어쩌면 이번 기회를 통해 새로운 인연을 맺을 수 있을 수도 있다는 생각에 엄청나게 기
분이 좋았다.

체언

NNG, NNP

용언

VV, VA

부사

어근

XR

[('나', 'NP'), ('는', 'JX'), ('평소', 'NNG'), ('조용', 'XR'), ('하', 'XSA'),
('ㄱ', 'ETN'), ('을', 'JKO'), ('즐기', 'VV'), ('는', 'ETD'), ('편', 'NNB'),
('이', 'VCP'), ('지만', 'ECE'), (~~','~~, '~~SP~~'),

('어쩌면', 'MAG'), ('이번', 'NNG'), ('기회', 'NNG'), ('를', 'JKO'), ('통하', 'VV'),
('어', 'ECS'), ('새롭', 'VA'), ('ㄴ', 'ETD'), ('인연', 'NNG'), ('을', 'JKO'),
('맺', 'VV'), ('을', 'ETD'), ('수', 'NNB'), ('있', 'VV'), ('을', 'ETD'), ('수', 'NNB'),
('도', 'JX'), ('있', 'VV'), ('다는', 'ETD'), ('생각', 'NNG'), ('에', 'JKM'),
('엄청나', 'VA'), ('게', 'ECD'), ('기분', 'NNG'), ('이', 'JKS'), ('좋', 'VA'),
('았', 'EPT'), ('다', 'EFN'), (~~','~~, '~~SF~~')] }

N~ : 체언

V~ : 용언

E~ : 어미

J~ : 조사

M~ : 관형사+부사

X~ : 접두사, 접미사, 어근

- 나는 평소 조용함을 즐기는 편이지만,
- 어쩌면 이번 기회를 통해 새로운 인연을 맺을 수 있을 수도 있다는 생각에 엄청나게 기분
이 좋았다.

체언	NNG, NNP
용언	VV, VA
부사	
어근	XR

1. 평소
조용
즐기

2. 어쩌면
이번
기회
통하 (여)
새롭
인연
맺 (다)
있 (다)
있 (다)
생각
엄청나
기분
좋 (다)

- 워드 임베딩 벡터 값

체언	어근
----	----
- 감정분석 (긍/부정)

용언

- 단어 출현 빈도수(가중치)
- 수식어 – 부사 (가중치)

감독명	대표작품	추출한 한줄평 개수
알폰소 쿠아론(A)	그래비티, 칠드런 오브 맨, 해리포터 3 등	9,971

Word2Vec로 워드 임베딩 후
“감독” 과 유사한 단어 상위 19개 출력

쿠아론	0.097
알폰소	0.118
연출	0.201
능력	0.453
알권	0.683
장르	0.747
대역	0.824
기법	0.826
정적	0.831
경이	0.841
연기	0.841
과학	0.844
탄생	0.851
전달	0.865
물타이크	0.897
개원	0.910
작품	0.927
상징	0.975
음향	1.000

- 나는 평소 조용함을 즐기는 편이지만,
- 어쩌면 이번 기회를 통해 새로운 인연을 맺을 수 있을 수도 있다는 생각에 엄청나게 기분이 좋았다.

체언	NNG, NNP
용언	VV, VA
부사	
어근	XR

1. 평소
조용
즐기

I 0.9
긍정

$$I\ 0.9 + (\text{긍정}) = I$$

2. 어쩌면
이번
기회
통하 (여)
새롭
인연
맺 (다)
있 (다)
있 (다)
생각
엄청나
기분
좋 (다)

E 0.5

$$E\ 0.5 + (\text{긍정}) = E$$

긍정

혼자보다는 다른 사람과 시간을 보내고 싶어
하는가?

유형 2) 중립적으로 대답하는 경우

난 친구랑 노는 것도 좋지만, 혼자 있는 것도 좋아!

E

I

—————> 더 많은 질문을 통해 MBTI 파악

57% 외향형



혼자보다는 다른 사람과 시간을 보내고 싶어
하는가?

유형 3) 완전히 다른 대답을 하는 경우

아 배고프다

집에 언제 가지..

치킨 먹고 싶다



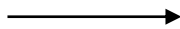
질문에 대한 답이 적절하지 않습니다.

-> 재 질문

예외) 새로운 장소를 탐험하는 것을 좋아해요.

I 혼자 새로운 장소 탐험

E 여럿이 새로운 장소 탐험

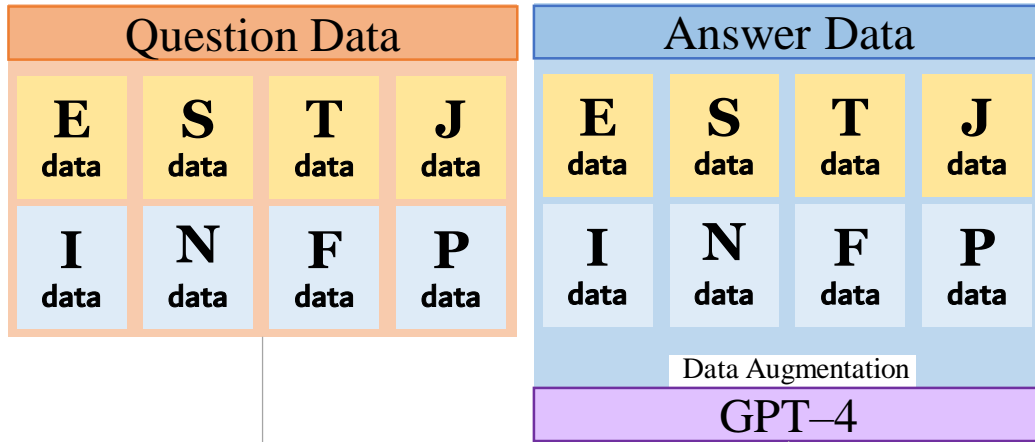


질문에 대한 적절하지 않은 대답으로 판
단

-> 재 질문

Transformer

Data Collect - Web Crawling



< Question data_Answer data_label >



Text preprocessing

Transformer

Word2Vec-ko,
FastText-KR, KoBERT

Encoder

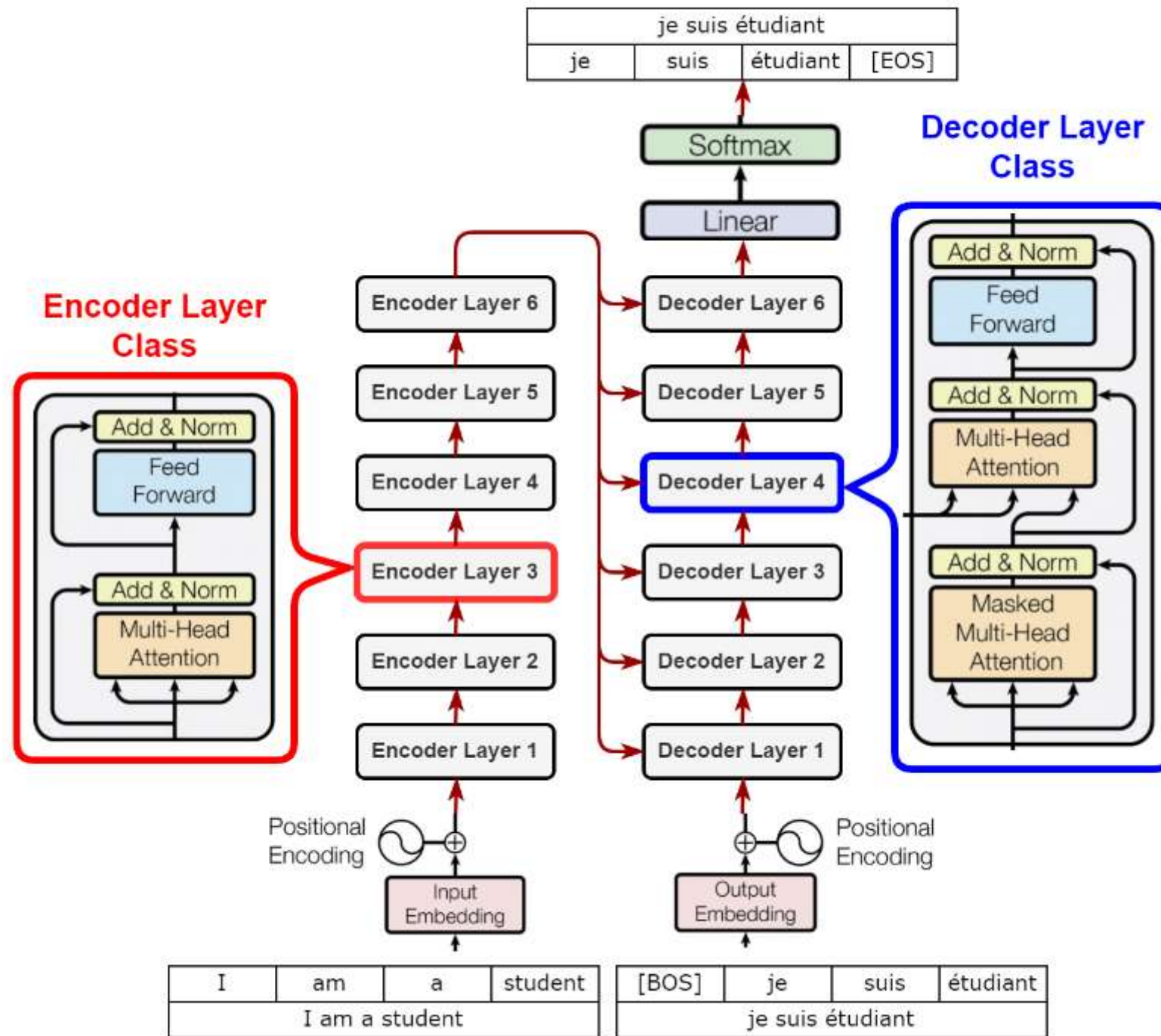
4-BINARY CLASSIFIERS
Model

E/I
model

S/N
model

T/F
model

J/P
model



Data_1 = [‘Question data’, ‘ Answer data’, ‘label’]

Training Dataset

Index	Answer_Index	Question	Answer	Label
1	1	혼자보다는 다른 사람과 시간을 보내고 싶어 하는가?	그렇다. 다른 사람들과 함께 시간을 보내는 것을 즐기는 편이다.	E/I
	2		난 혼자 있는 시간을 더 선호하는 것 같다. 조용한 게 최고야!	
	
	250		그래도 사람들과 어울리는 게 더 행복하지?	
2	1	사람들과 만나는 약속이 끝난 후 집에 들어와 혼자 있을 때, 나의 기분은?	약속이 끝나고 집에 돌아와 혼자 있을 때, 대개 나는 안정감과 평온함을 느낀다.	E/I
	2		집에 돌아와 혼자 있는 순간에는 조금 외로움을 느낄 때도 있다.	
	
	250		집에 돌아와서 집안 정리를 한 뒤 나만의 시간을 갖는 편이다.	
...				

Answer

Web Crawling

E data	S data	T data	J data
I data	N data	F data	P data

8가지 유형에 대한

특징 data 수집



Answer 일부

직접 수집 및 정제

Data Augmentation

GPT-4

Answer
그렇다. 다른 사람들과 함께 시간을 보내는 것을 즐기는 편이다.
난 혼자 있는 시간을 더 선호하는 것 같다. 조용한 게 최고야!
...
그래도 사람들과 어울리는 게 더 행복하지?
약속이 끝나고 집에 돌아와 혼자 있을 때, 대개 나는 안정감과 평온함을 느낀다.
집에 돌아와 혼자 있는 순간에는 조금 외로움을 느낄 때도 있다.
...
집에 돌아와서 집안 정리를 한 뒤 나만의 시간을 갖는 편이다.

감사합니다