MLM Fine Tuning

A Multi-Stage Transfer Learning Approach for Personality Classification
Using Masked Language Modeling and Task-Specific Fine-Tuning

FIVE GUYS

김민규 성재영 이영주 이준성 이현준

Method

MLM Fine Tuning

Task

MBTI Classification

Models

Transformer

Attention is all you need

Google (2017)

BERT

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Google (2018)

ELECTRA

ELECTRA: Pre-Training Text Encoders as Discriminators rather than Generators

Google Research (2020)

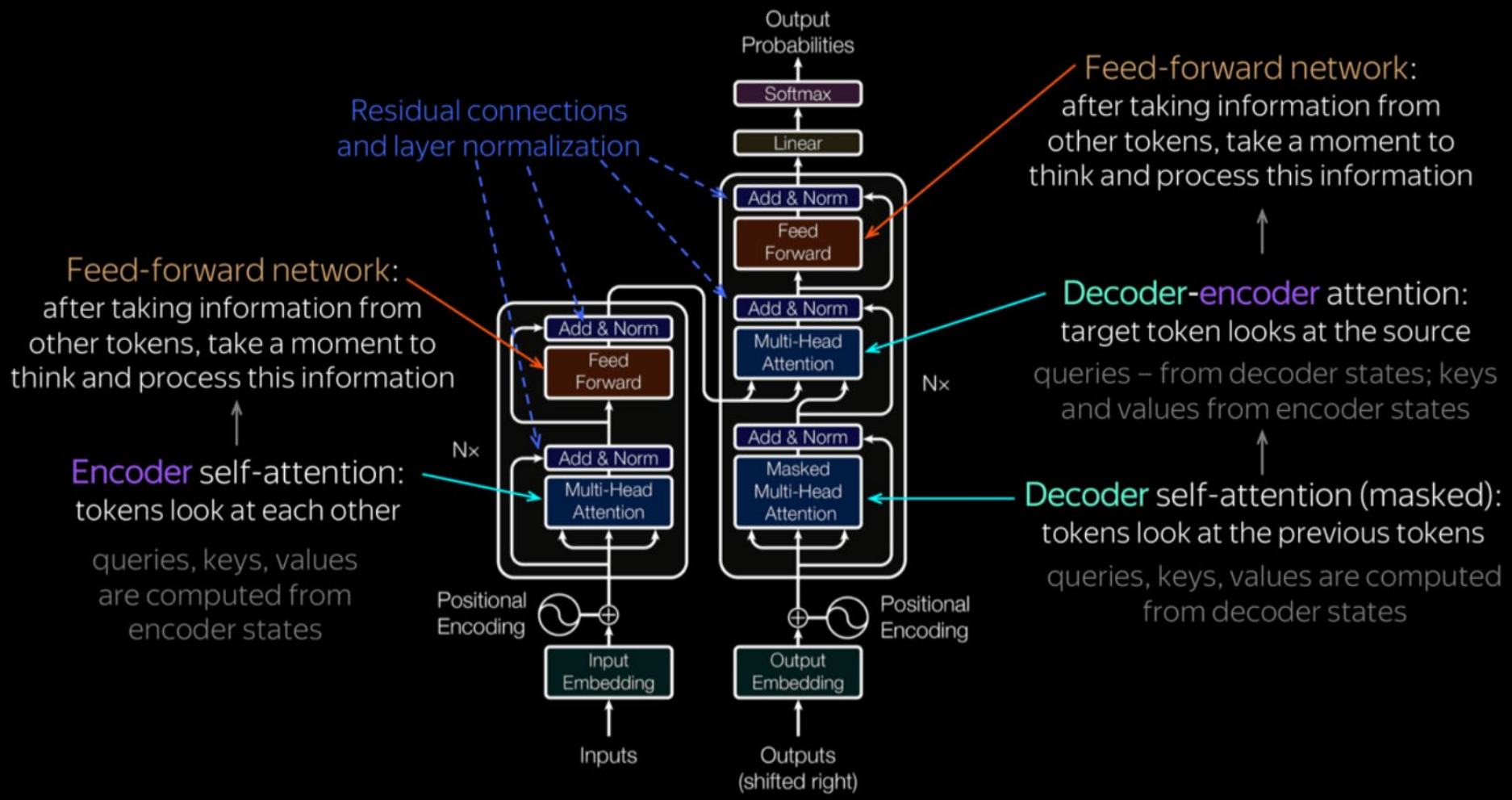
KcELECTRA

KcELECTRA: Korean Comments ELECTRA

Lee Jun Bum (2022)

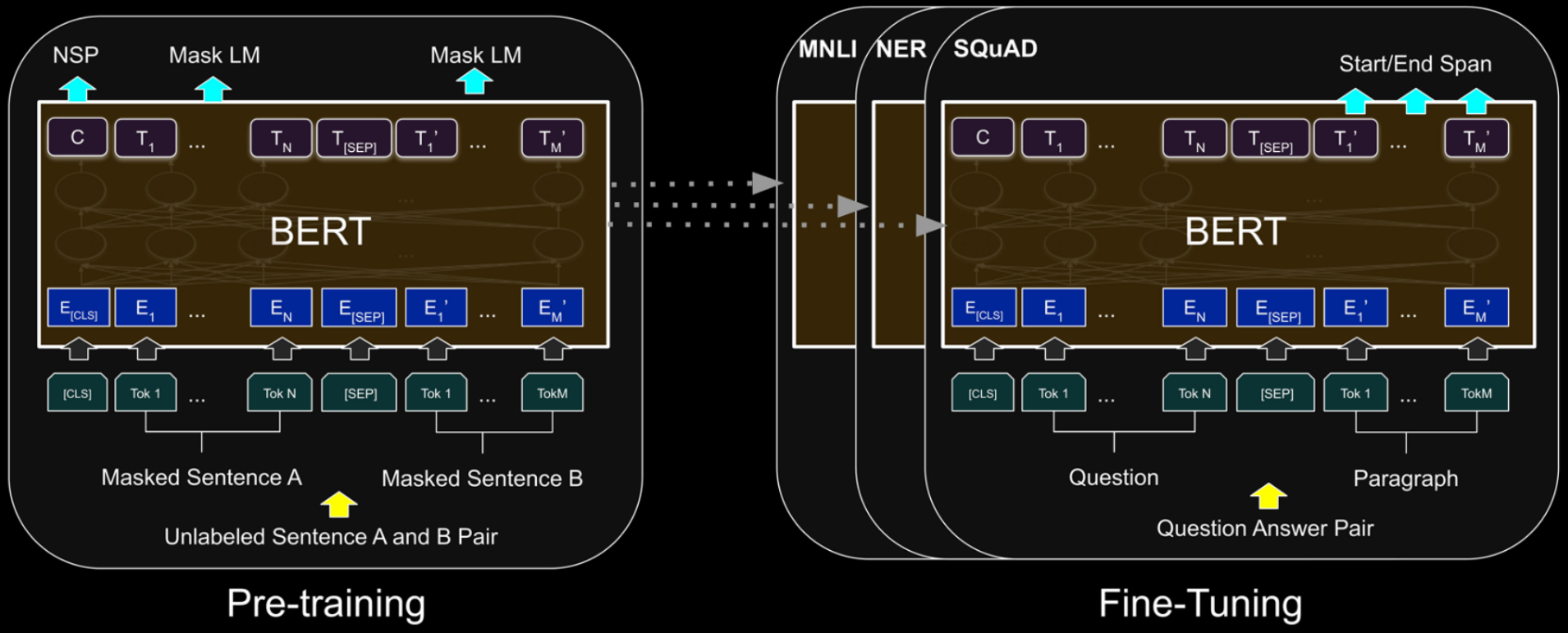
Transformer

Google (2017) Attention is all you need



BERT

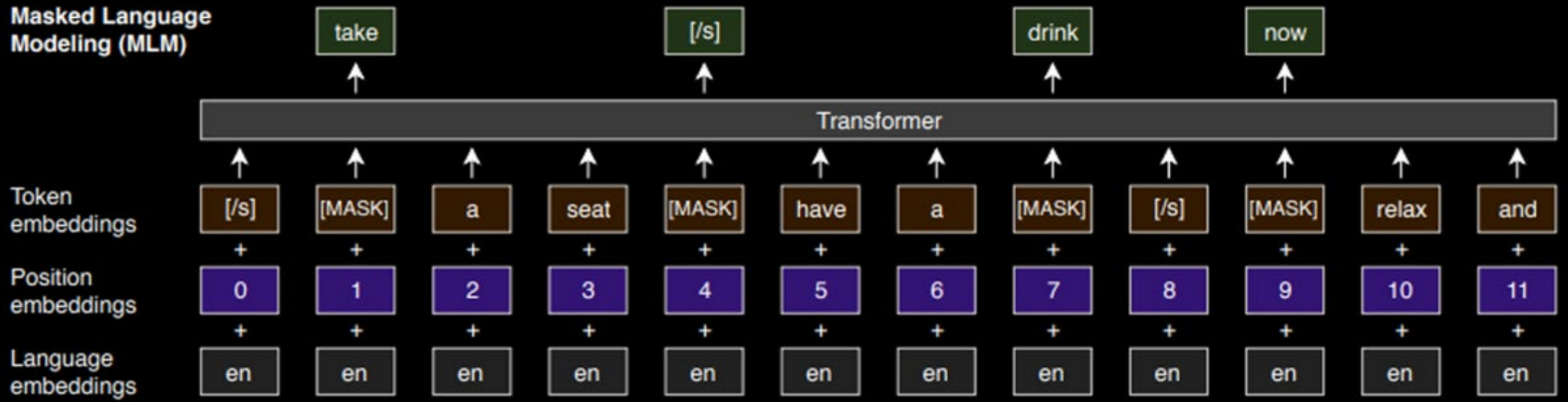
Google (2018) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding



MLM (Masked Language Model)

NSP (Next Sentence Prediction)

MLM (Masked Language Modeling)



Masking ratio

15 %

MLM in BERT

Masking Rates			Dev Set Results		
			MNLI	NER	
MASK	SAME	RND	Fine-tune	Fine-tune	Feature-based
80%	10%	10%	84.2	95.4	94.9
100%	0%	0%	84.3	94.9	94.0
80%	0%	20%	84.1	95.2	94.6
80%	20%	0%	84.4	95.2	94.7
0%	20%	80%	83.7	94.8	94.6
0%	0%	100%	83.6	94.9	94.6

Masking ratio

15 %

MASK (80%) my dog is hairy → my dog is [MASK]

RND (10%) my dog is hairy → my dog is apple

SAME (10%) my dog is hairy → my dog is hairy

Domain Specific MLM

I went to [MASK] after lunch

Model

ISTJ	ISFJ	INFJ	INTJ
ISTP	ISFP	INFP	INTP
ESTP	ESFP	ENFP	ENTP
ESTJ	ESFJ	ENFJ	ENTJ

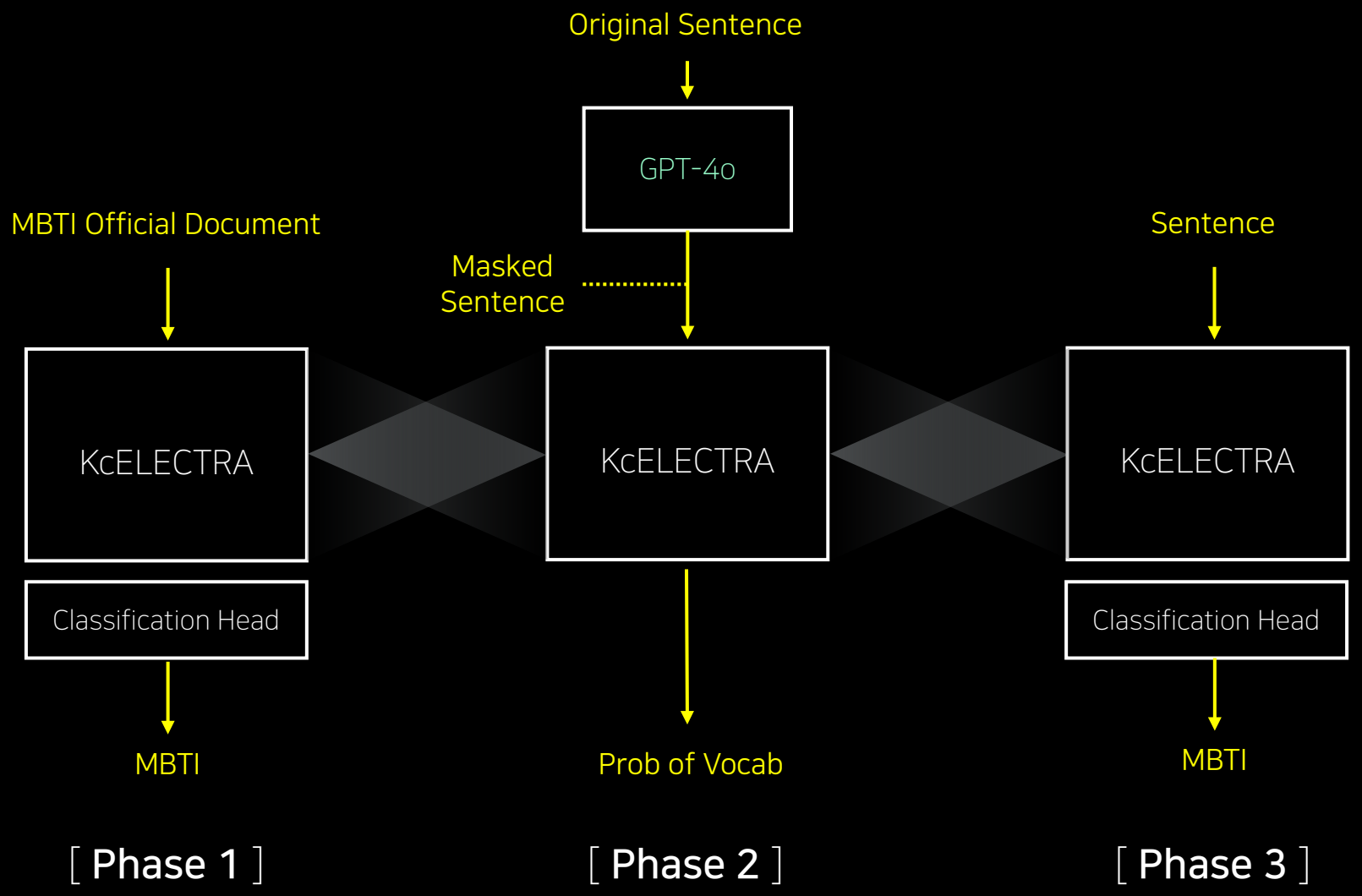
MBTI

Model'

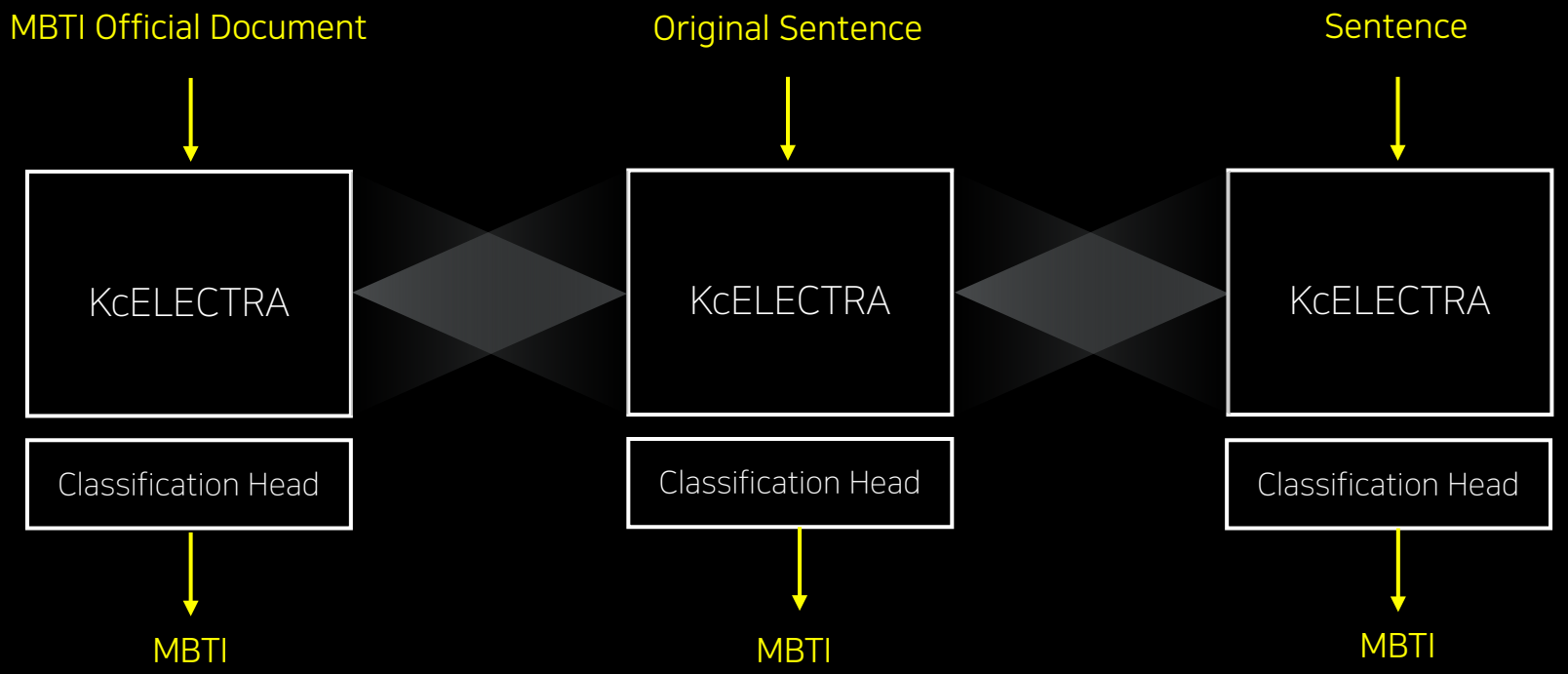
MLM Fine Tuning

A Multi-Stage Transfer Learning Approach for Personality Classification Using Masked Language Modeling and Task-Specific Fine-Tuning

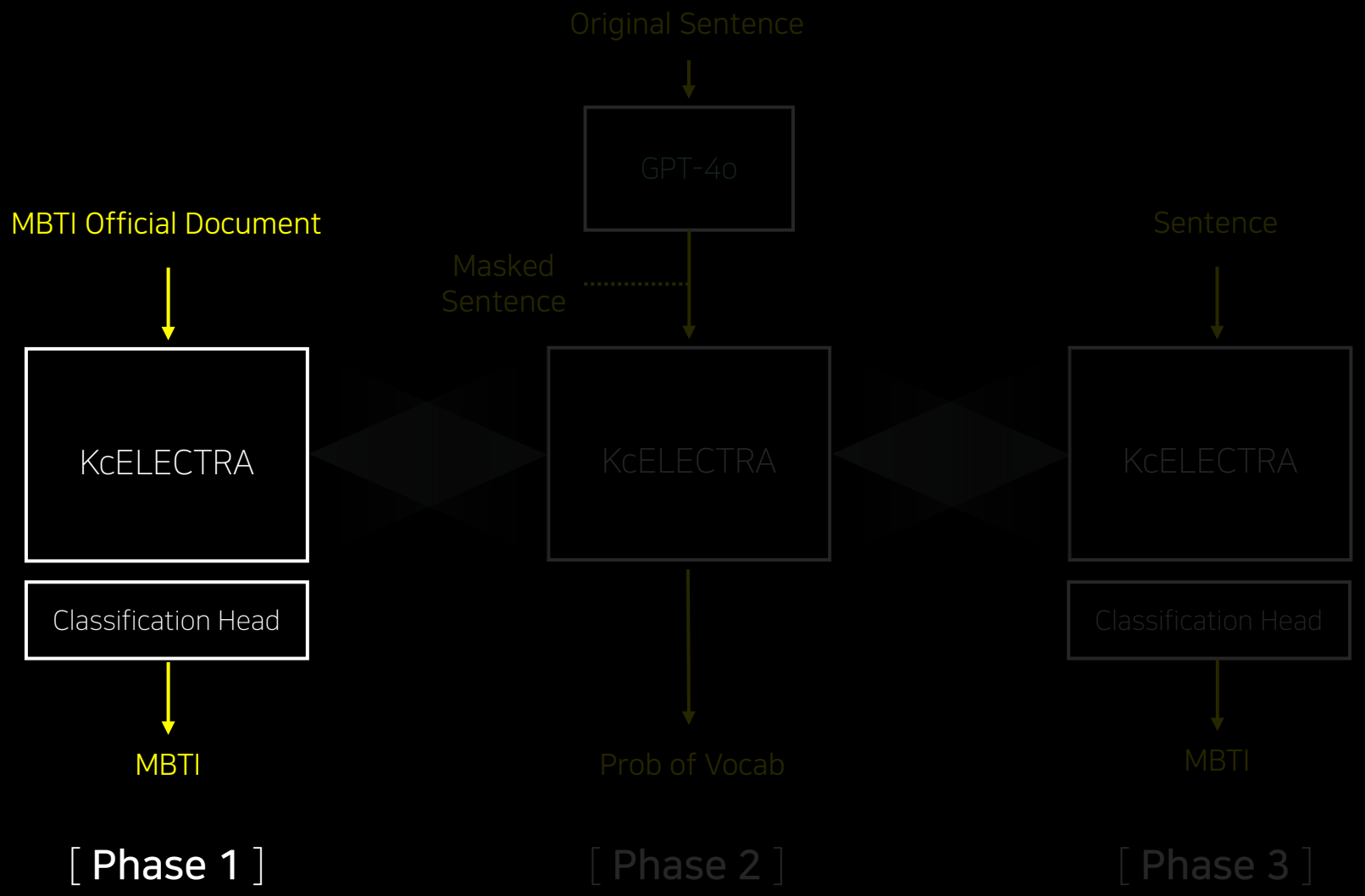
MLM Fine Tuning



Original


MBTI Official Document
Original Sentence
Sentence
KcELECTRA
KcELECTRA
KcELECTRA
Classification Head
Classification Head
Classification Head
MBTI
MBTI
MBTI
[Phase 1]
[Phase 2]
[Phase 3]

Phase 1


Original Sentence
GPT-4o
MBTI Official Document
Masked Sentence
Sentence
KcELECTRA
KcELECTRA
KcELECTRA
Classification Head
Classification Head
MBTI
Prob of Vocab
MBTI
[Phase 1]
[Phase 2]
[Phase 3]

MBTI Official Document Dataset

그들은 급격한 합리적 자극 반응 과정에서 사실적이고 즉각적인 현실을 바탕으로 비판적 결정을 내리는 경향이 있습니다.

이것은 학교와 다른 고도로 조직화 된 환경을 ESTP에 대한 도전으로 만듭니다. 확실히 똑똑하지 않기 때문이 아니며, 그곳에서 잘 할 수는 있지만, 공식 교육의 연대적인 강의 접근은 이러한 성격이 일반적으로 즐기는 실습 학습과는 거리가 멉니다.

ESTP

Dataset

1440 =

Article

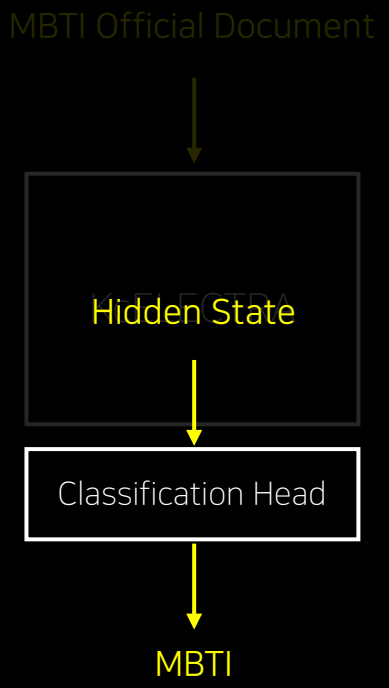
90 X

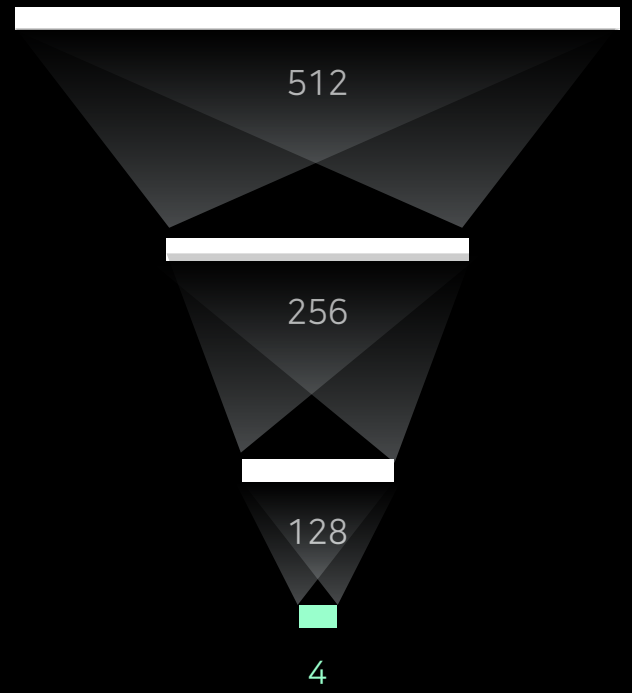
MBTI

16

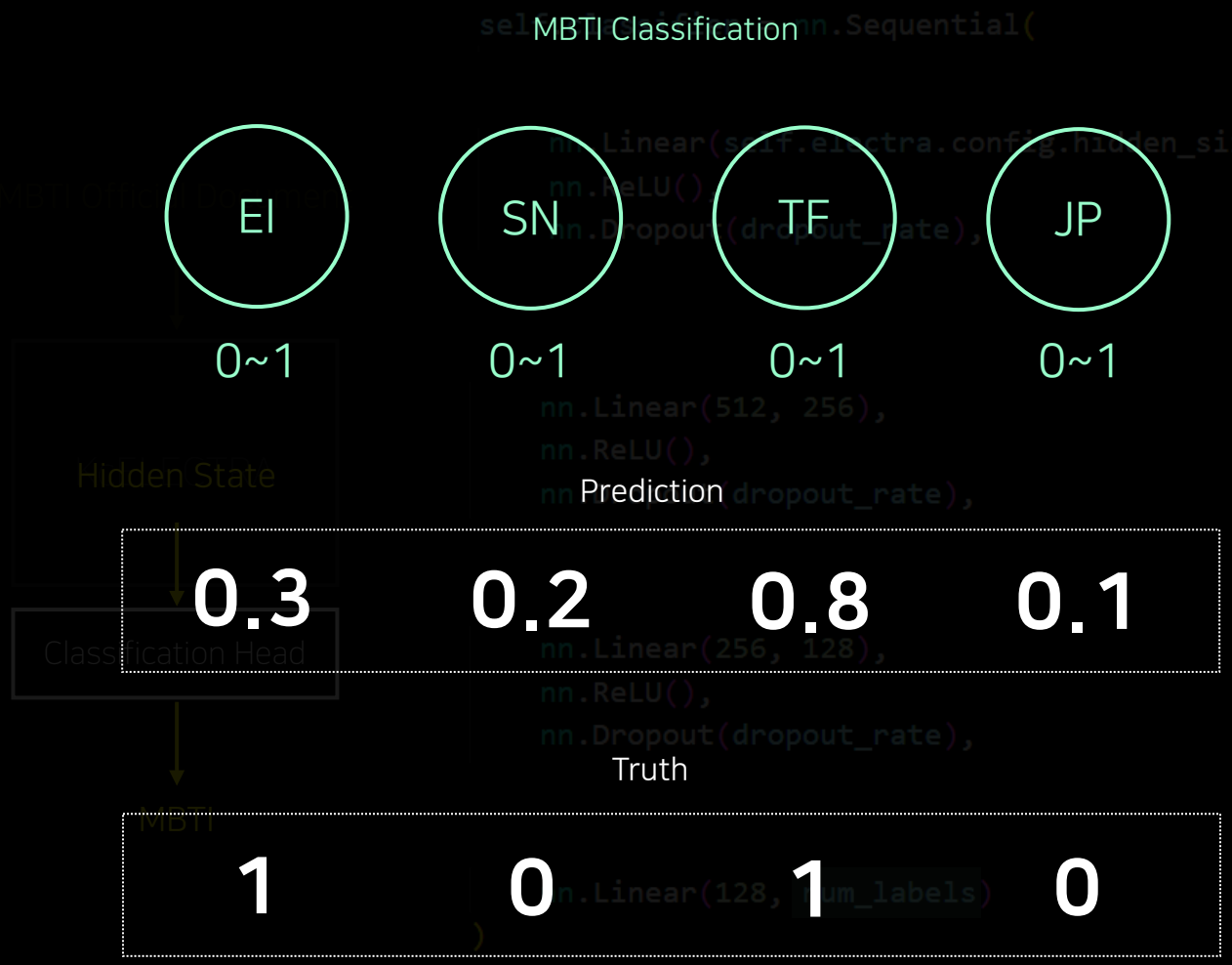
Classification Head

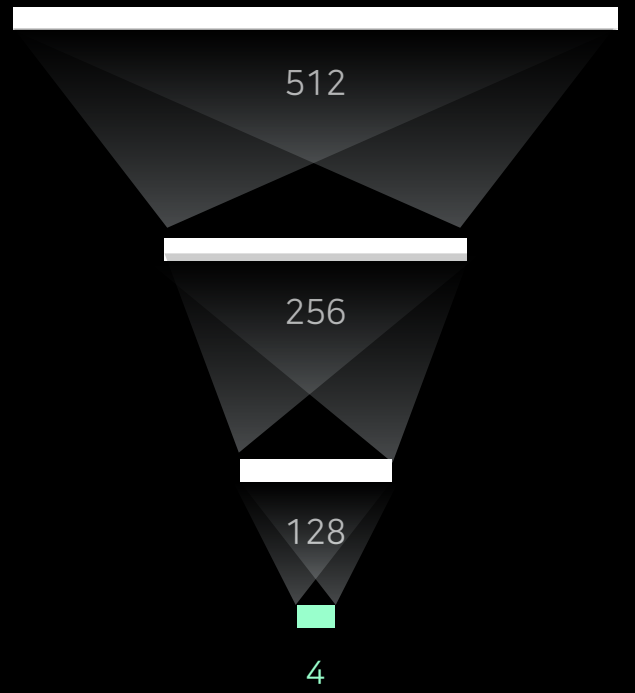

```
self.classifier = nn.Sequential(
    nn.Linear(self.electra.config.hidden_size, 512),
    nn.ReLU(),
    nn.Dropout(dropout_rate),
    nn.Linear(512, 256),
    nn.ReLU(),
    nn.Dropout(dropout_rate),
    nn.Linear(256, 128),
    nn.ReLU(),
    nn.Dropout(dropout_rate),
    nn.Linear(128, num_labels)
)
```

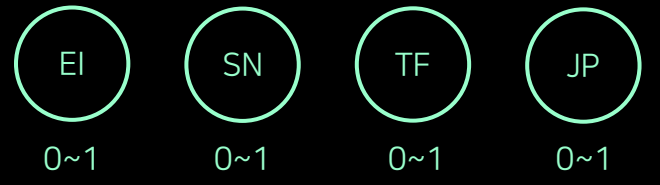




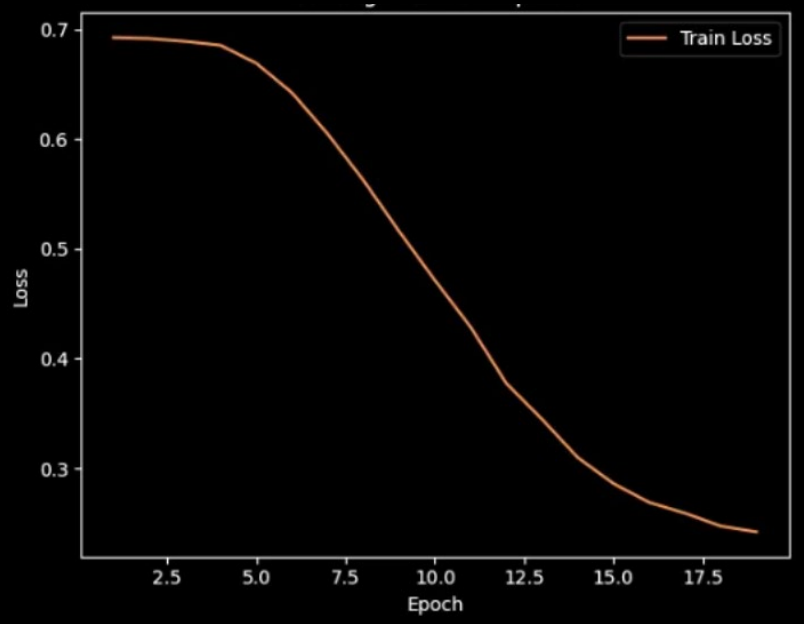
Classification Head



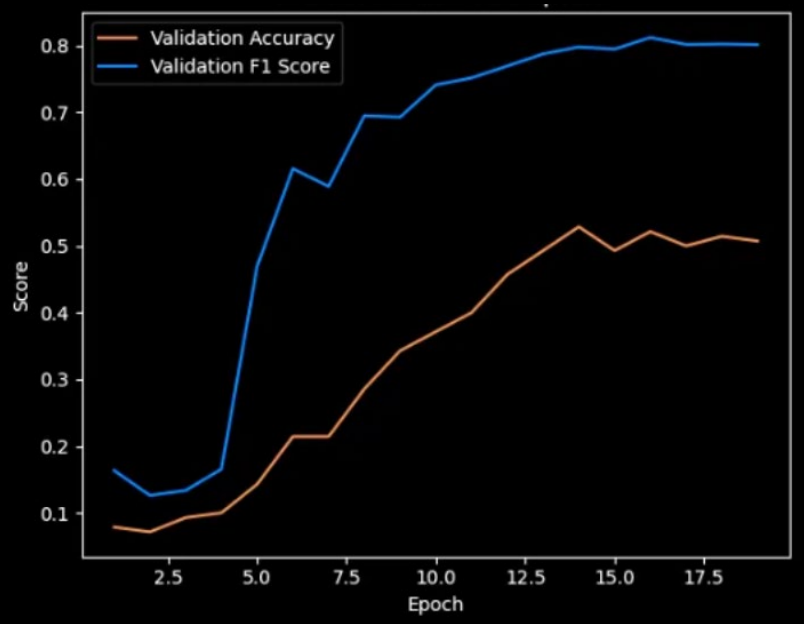




Training



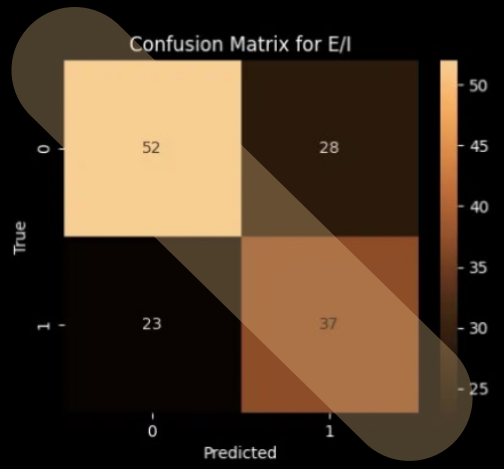
Training Loss Over Epochs



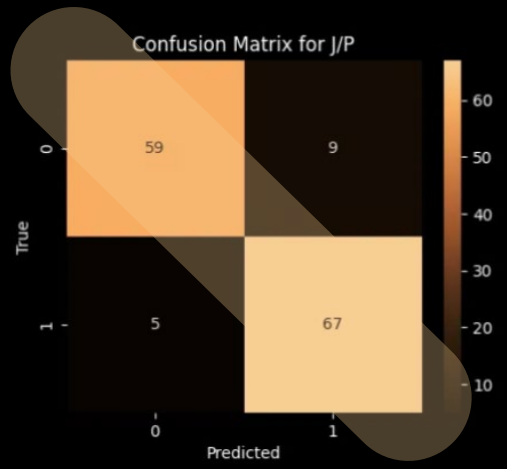
Validation Metrics Over Epochs

Training

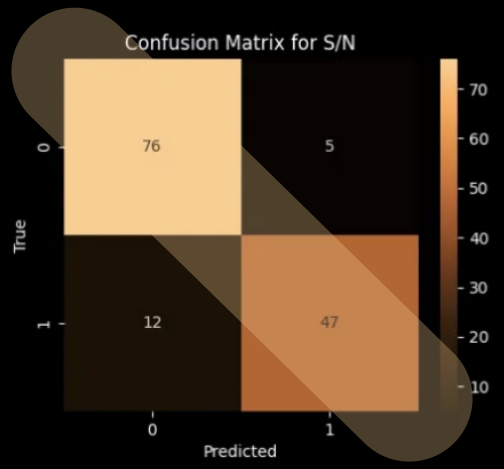
EI



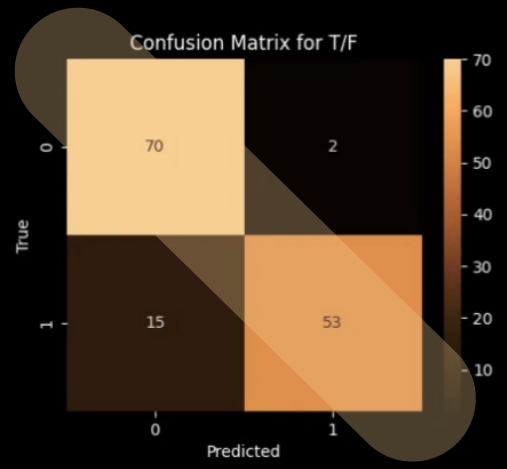
JP



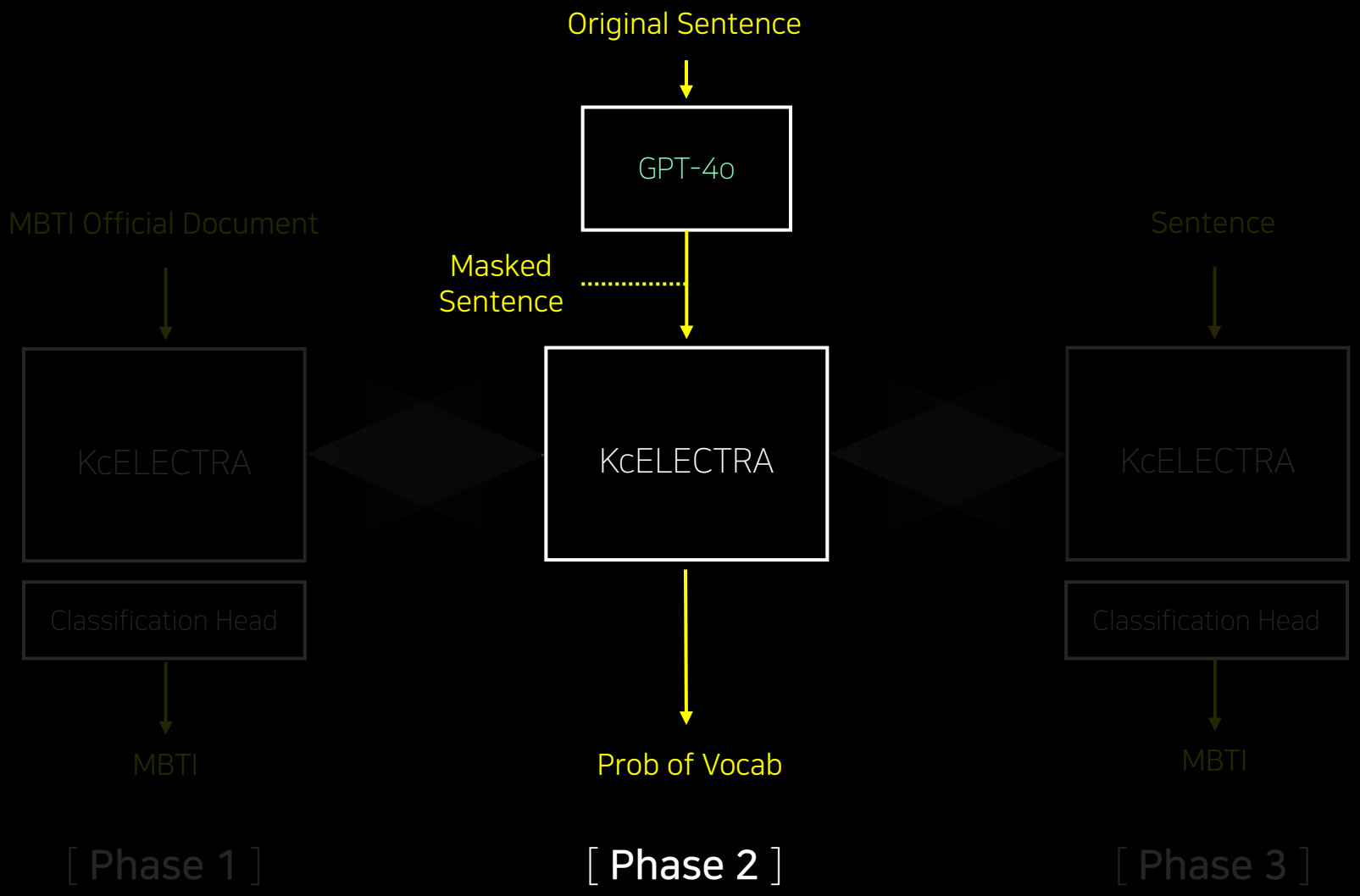
SN



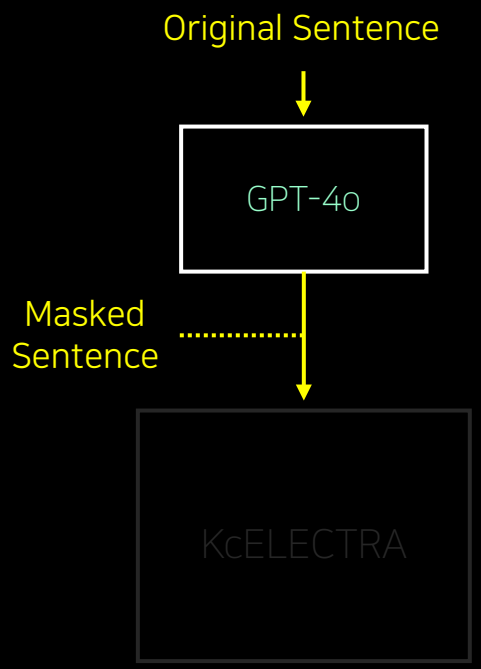
TF



Phase 2


Original Sentence
GPT-4o
MBTI Official Document
Masked Sentence
Sentence
KcELECTRA
KcELECTRA
KcELECTRA
Classification Head
Classification Head
MBTI
Prob of Vocab
MBTI
[Phase 1]
[Phase 2]
[Phase 3]

Masked Dataset



Original Sentence

이 일은 효율적으로 처리해야 해. 우선순위를 확실히 정하자.

MBTI

ESTJ

GPT-4o

"MBTI의 특성에 맞는 가장 중요한 토큰 1개를 [MASK] 해 줘"

Masked Sentence

이 일은 효율적으로 처리해야 해. [MASK] 를 확실히 정하자.

Dataset

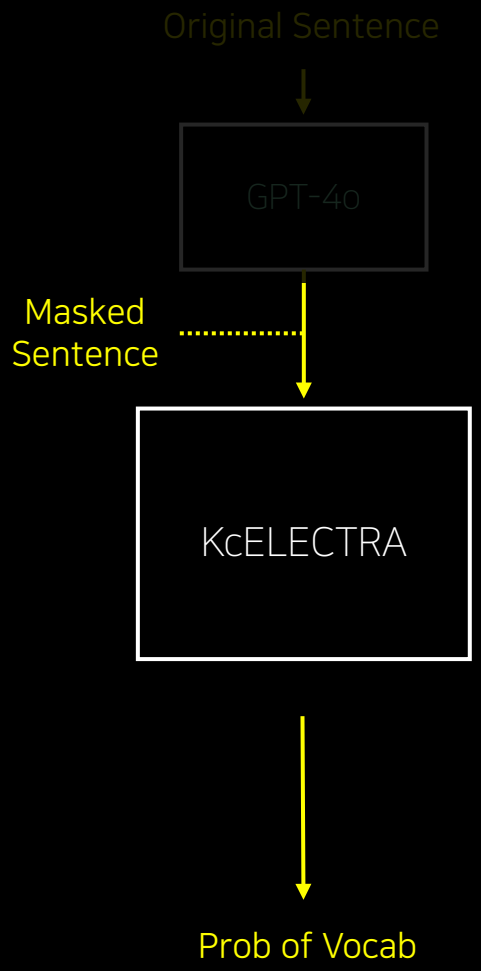
14,400 = **900** X **16**

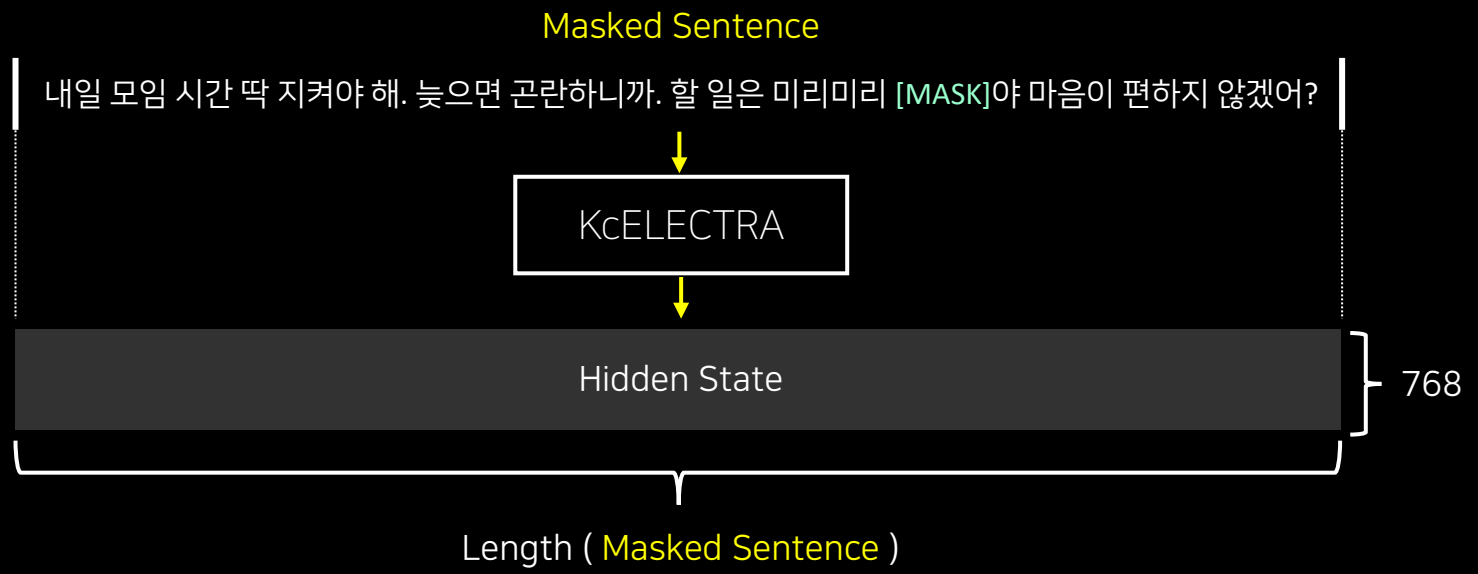
Original Sentence

MBTI

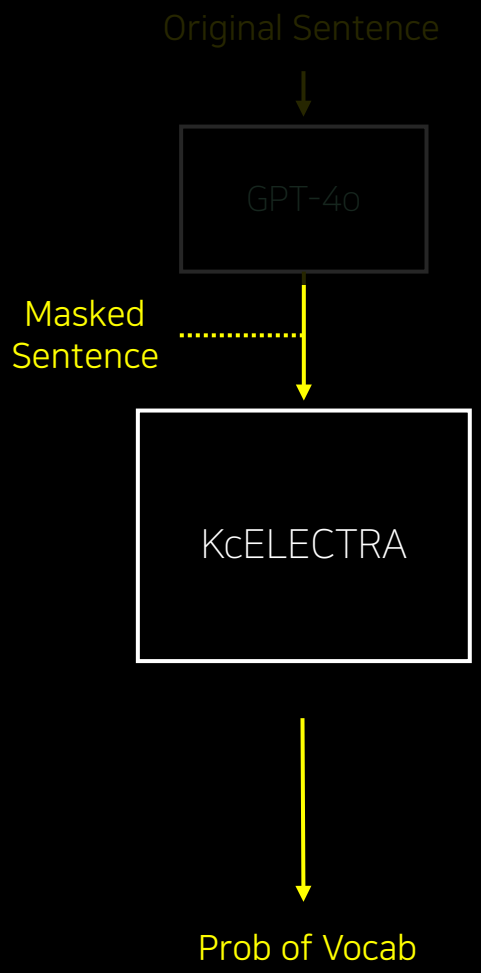
Prob of Vocab

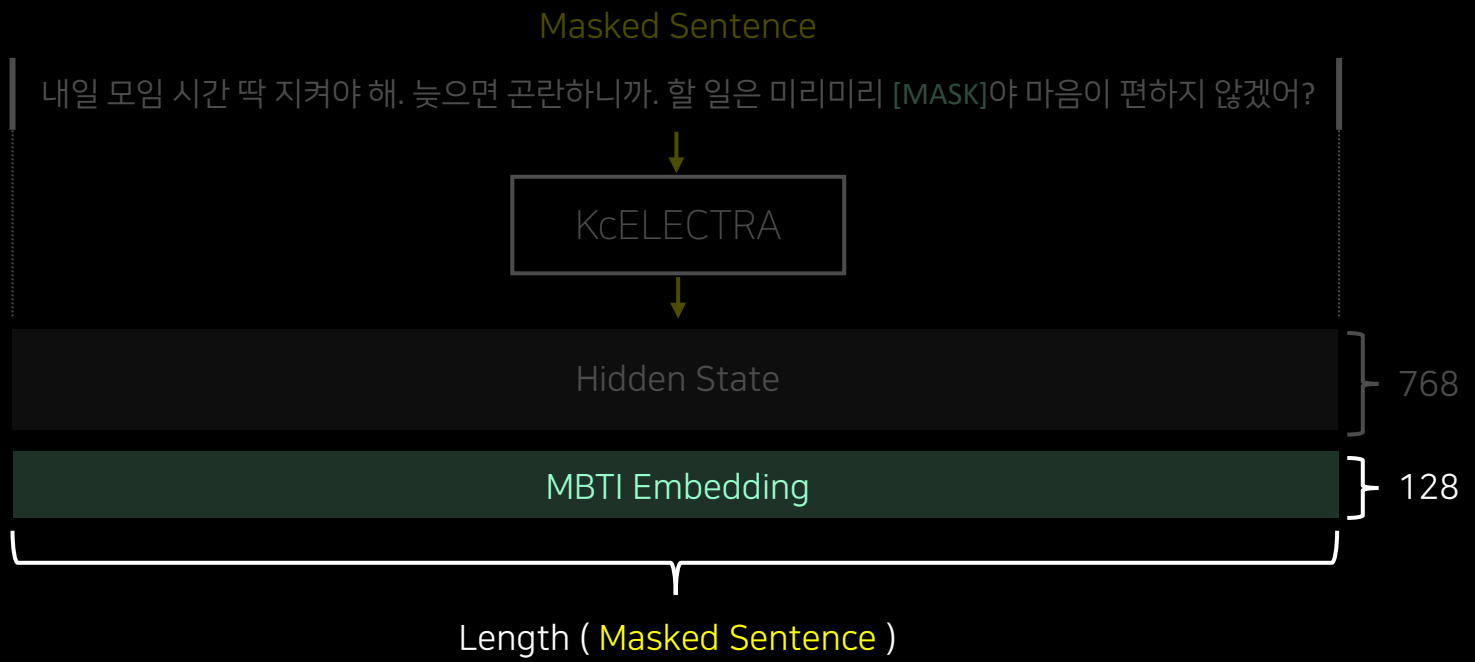
Hidden State





MBTI Embedding

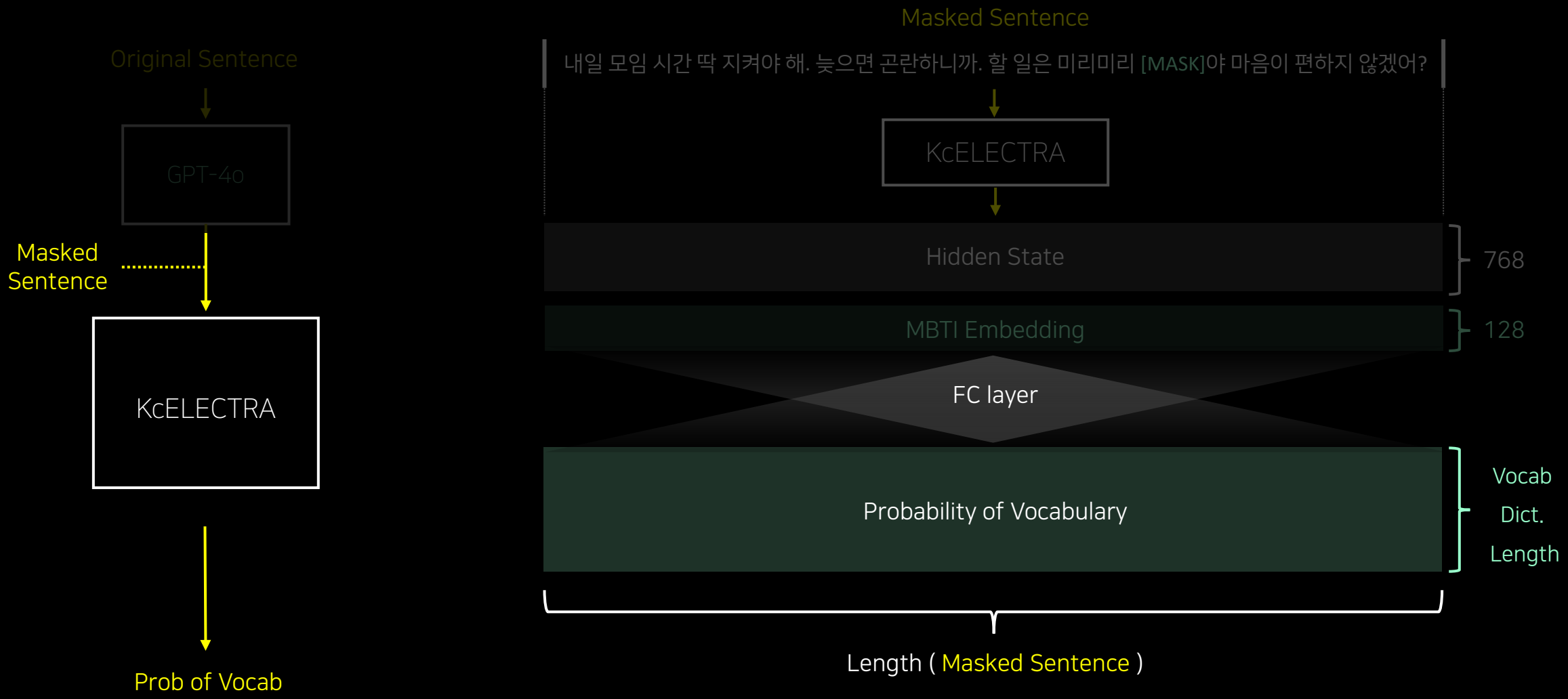




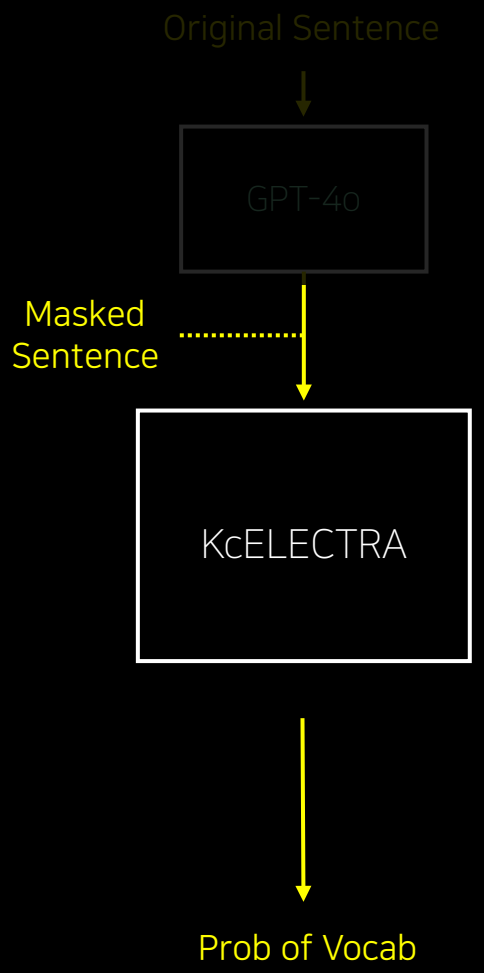

```
MBTI_TYPES = [
    'INTJ', 'INTP', 'ENTJ', 'ENTP', 'INFJ', 'INFP', 'ENFJ', 'ENFP',
    'ISTJ', 'ISFJ', 'ESTJ', 'ESFJ', 'ISTP', 'ISFP', 'ESTP', 'ESFP'
]
```

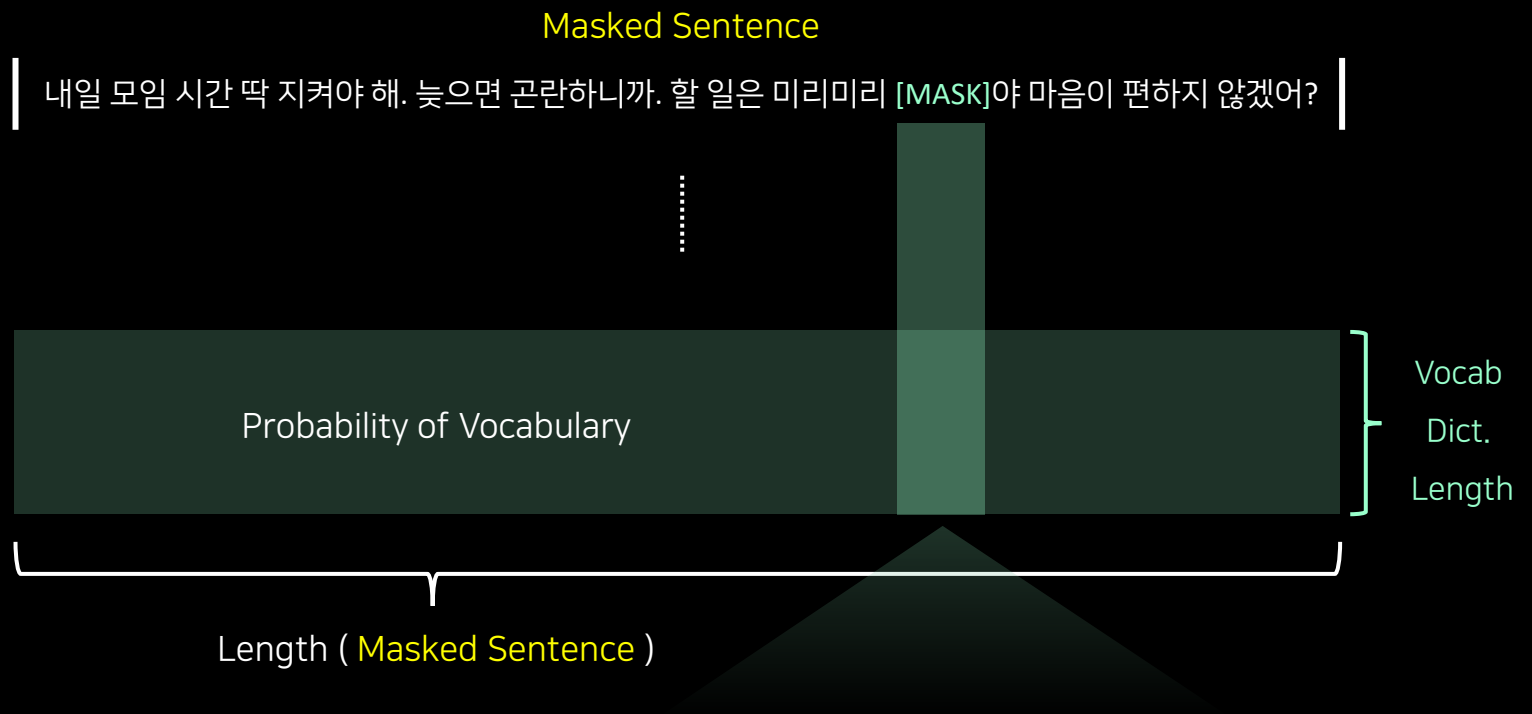
0~15

Probability of Vocabulary



Probability of Vocabulary





ISTJ

Vocab 1	잘해	(Prob: 0.0143)
Vocab 2	정리해	(Prob: 0.0119)
Vocab 3	기록해	(Prob: 0.0080)
Vocab 4	체계	(Prob: 0.0075)
Vocab 5	세워	(Prob: 0.0071)

Probability of Vocabulary

Masked Sentence

내일 모임 시간 딱 지켜야 해. 늦으면 곤란하니까. 할 일은 미리미리 [MASK]야 마음이 편하지 않겠어?

ISTJ

Vocab 1	잘해	(Prob: 0.0143)
Vocab 2	정리해	(Prob: 0.0119)
Vocab 3	기록해	(Prob: 0.0080)
Vocab 4	체계	(Prob: 0.0075)
Vocab 5	세워	(Prob: 0.0071)

ENFP

Vocab 1	자유롭게	(Prob: 0.0118)
Vocab 2	풀어	(Prob: 0.0111)
Vocab 3	감동	(Prob: 0.0111)
Vocab 4	잘	(Prob: 0.0111)
Vocab 5	신기	(Prob: 0.0085)

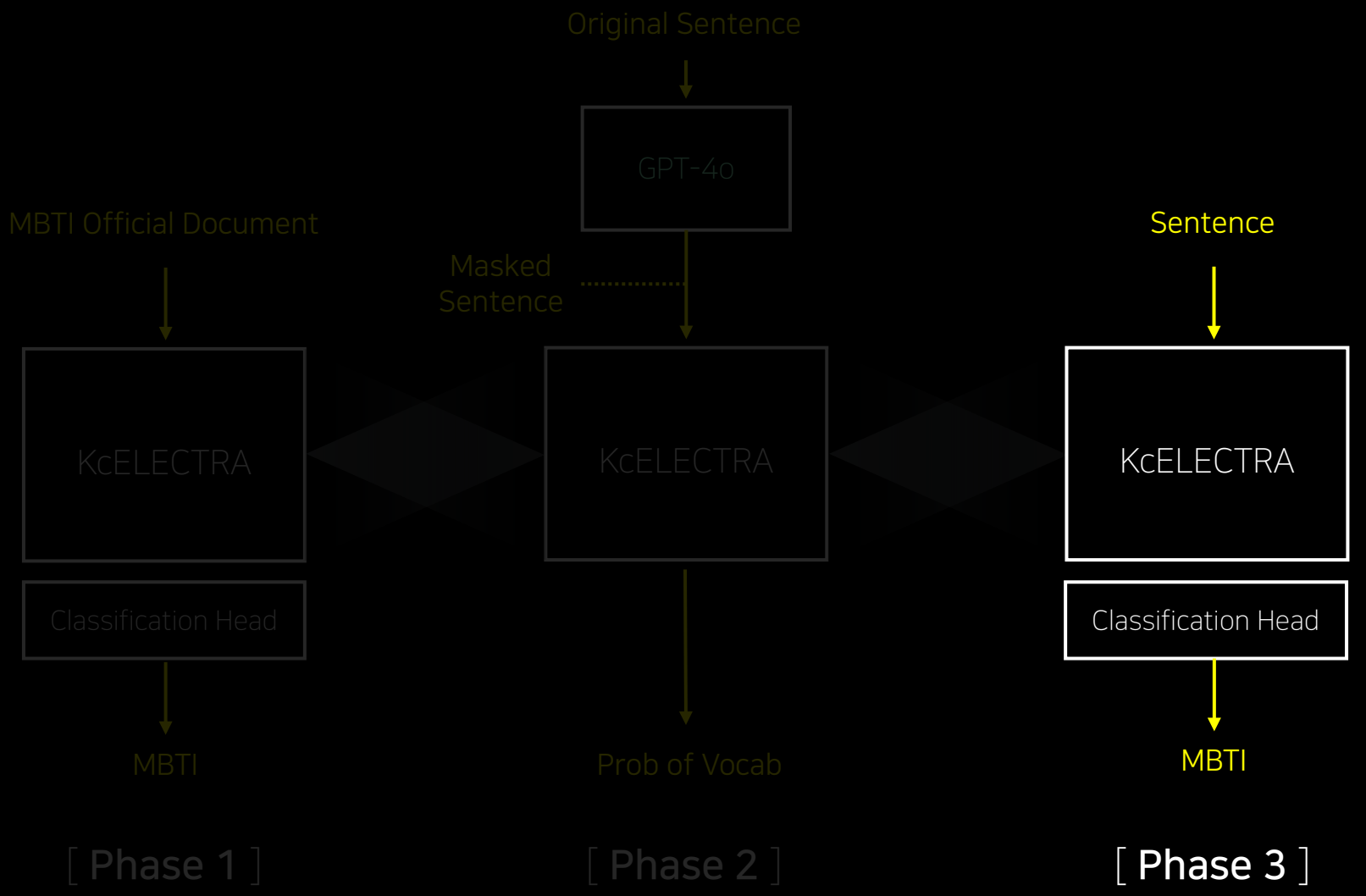
INTJ

Vocab 1	적으로	(Prob: 0.0201)
Vocab 2	돼야	(Prob: 0.0201)
Vocab 3	잘	(Prob: 0.0168)
Vocab 4	체계	(Prob: 0.0164)
Vocab 5	정리	(Prob: 0.0154)

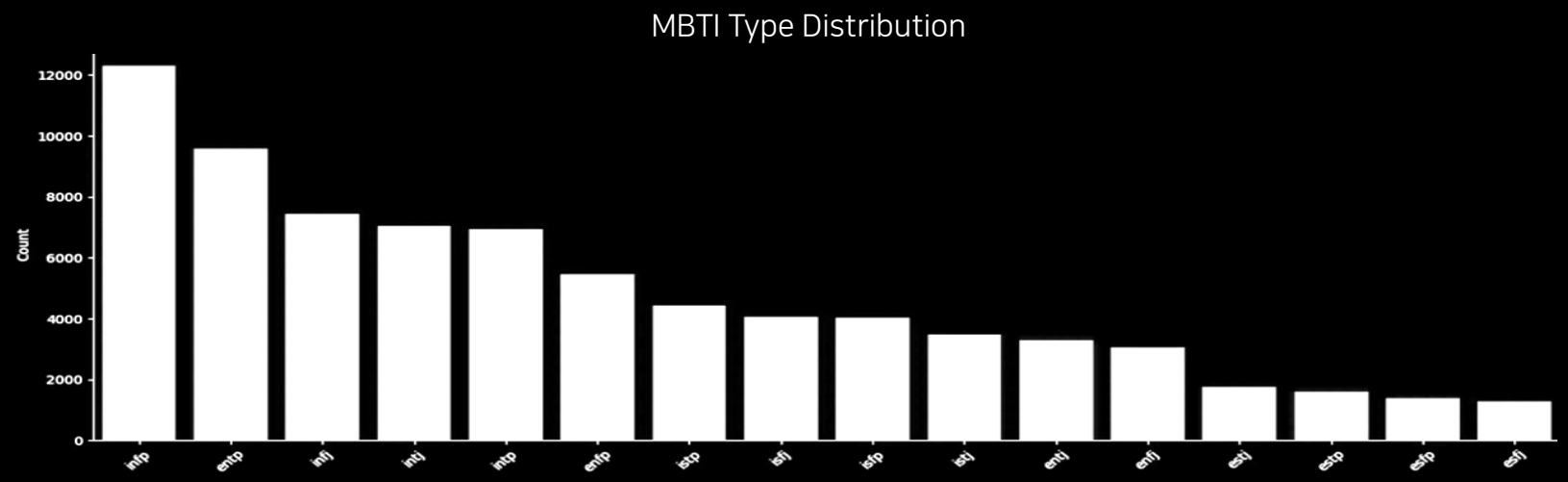
ESFP

Vocab 1	자유롭게	(Prob: 0.0188)
Vocab 2	잘	(Prob: 0.0133)
Vocab 3	풀어	(Prob: 0.0094)
Vocab 4	정리	(Prob: 0.0079)
Vocab 5	효율	(Prob: 0.0068)

Phase 3


Original Sentence
GPT-4o
MBTI Official Document
Masked Sentence
Sentence
KcELECTRA
KcELECTRA
KcELECTRA
Classification Head
Classification Head
MBTI
Prob of Vocab
MBTI
[Phase 1]
[Phase 2]
[Phase 3]

Article Dataset



님들이 과자나 치킨 같은 거 사 와서 먹으라고 했는데 동생이 안 먹겠다고 해서 방치해 놨던 먹을거리가 다음날 보니 말도 없이 없어졌을 때 화나 세영???

Article

ISTJ

MBTI

80,000

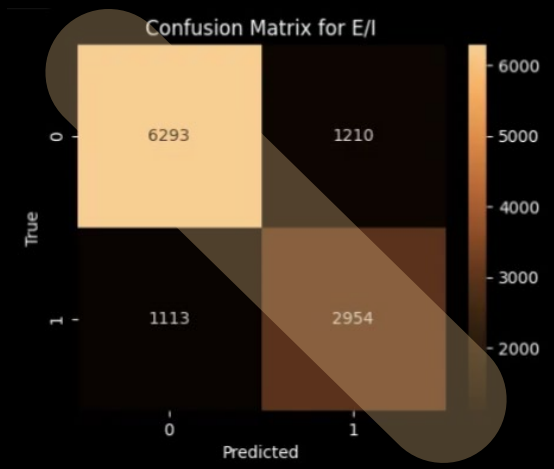
Dataset

Result

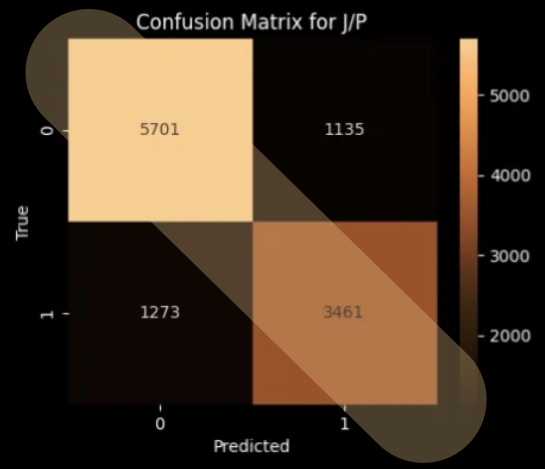
	MLM Fine Tuning
Test Loss	1.1752
Test Accuracy	58.26

Result

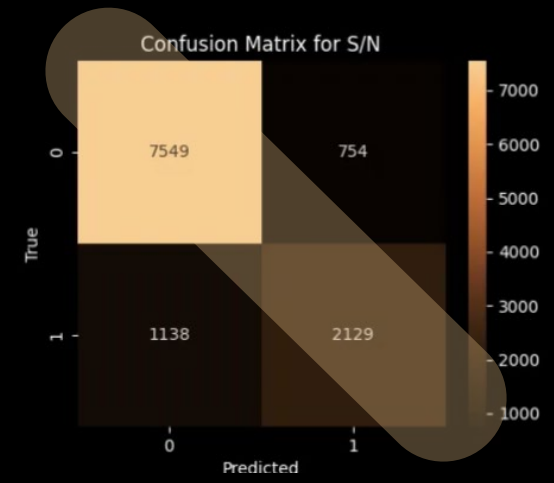
EI



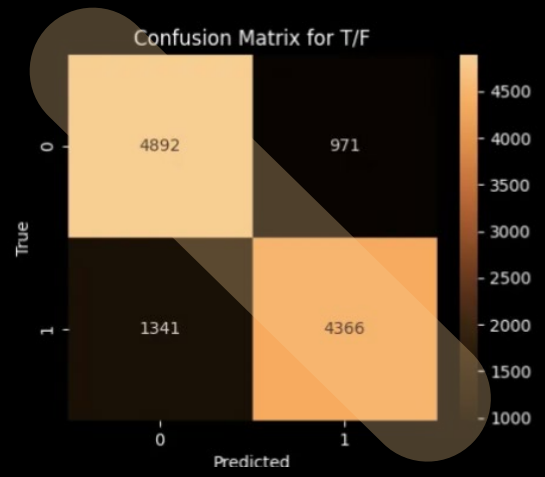
JP



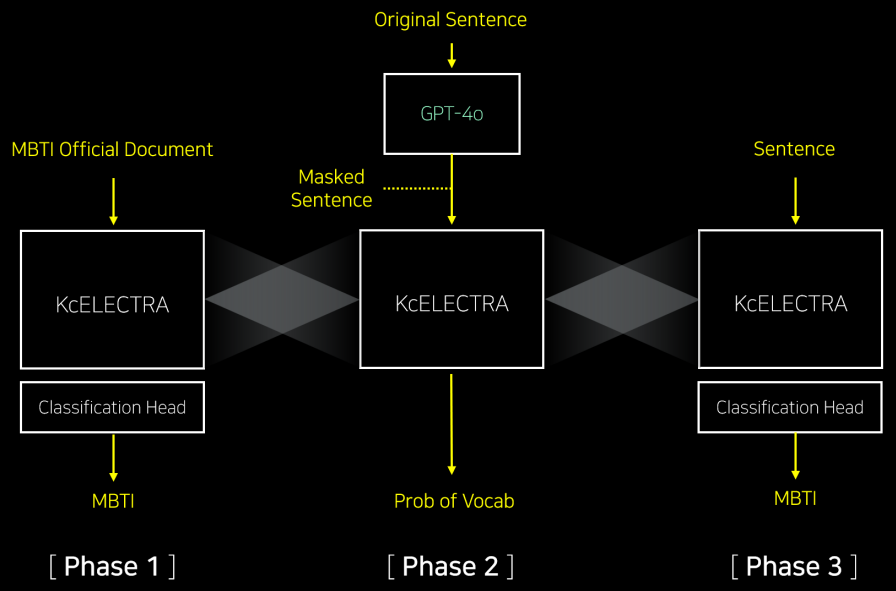
SN



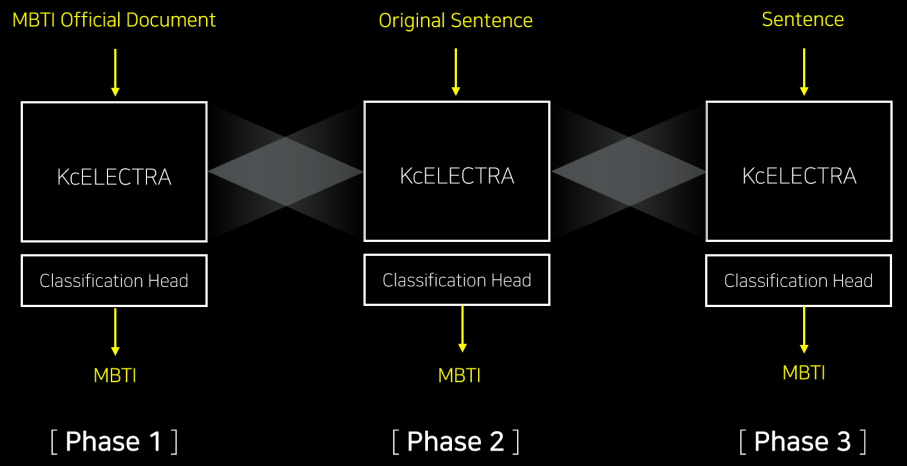
TF



Comparison



MLM Fine Tuning

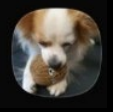


Fine Tuning

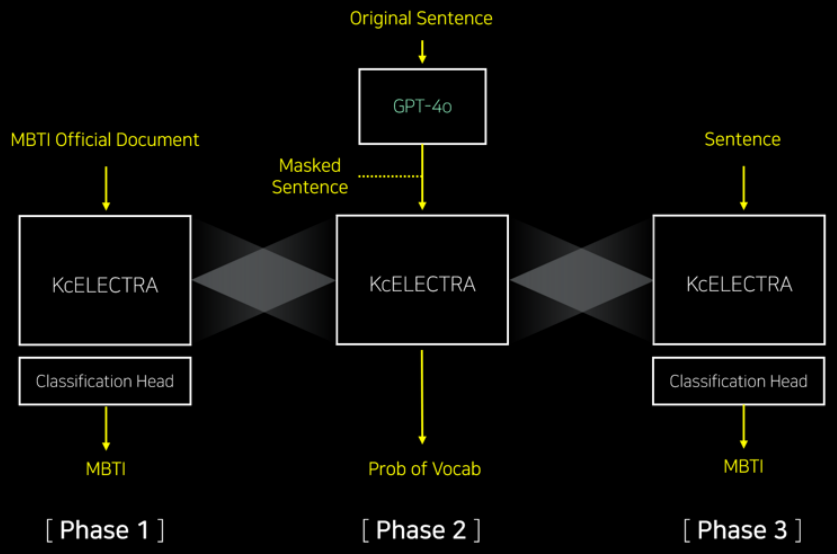
Comparison

	MLM Fine Tuning	Fine Tuning
EI Accuracy	80.86	**81.31**
SN Accuracy	82.31	**82.82**
TF Accuracy	**81.88**	81.78
JP Accuracy	**81.50**	79.29
Test Loss	**1.0580**	1.2715
Test Accuracy	**81.64**	81.30

Apply

		Truth	MFT	FT
 김민규	너가 기존에 masking한 방식은 gpt한테 부탁한거라 그랬지??! 오케이오케이 그런 식으로 갈 것 같어 고마워!	ISFJ	ISFP	ESFP
 성재영	나 1시넘어서 들어갈것 같은데 너 피곤하면 내일 아침에 할까?	INTP	INTP	INTP
 이영주	다 했는데 MBTI 16개별로 나눠서 업로드 할까 아니면 다 합쳐서 업로드 할까? 컴비 강의실만가서 올릴겡	ESTP	ESTJ	ENFP
 이준성	더깔끔하게 할거면 솔직히 우리 그 마스킹한 데이터셋 랜덤 마스킹 해서 한번 더 돌려보는것도 나브지않은디 어떻게생각해 데이터팀한테 부탁해놓을까	INFP	INTP	ENFP
 이현준	굳굳 두 문장이 딱 적당한 길이인 것 같아 이거 두 문장 번역과정에서 한국어로는 세 문장 돼도 괜찮아!	ISFJ	INFJ	INFP

MLM Fine Tuning



A Multi-Stage Transfer Learning Approach for Personality Classification
Using Masked Language Modeling and Task-Specific Fine-Tuning