

일기 감성분석 및 코멘트 생성 서비스 '너하궁(너의 하루가 궁금해)'

NLP-04

4조 참치 | 김해원, 김혜빈, 박준형, 양봉석, 이예령

INDEX

01. INTRODUCTION

02. DATA & MODELING

03. PRODUCT SERVING

04. RESULT & CONCLUSION

05. APPENDIX

01. INTRODUCTION

1-1. 팀 소개

1-2. 프로젝트 개요

안녕하세요, NLP-04조 참치입니다

저희는 '사람과 공감하는 AI'라는 관심사로 모이게 된 팀입니다



김해원

PM

일기 데이터 정제

코멘트 생성
모델 개발



김혜빈

코멘트 어노테이션
Pseudo-labeling

프론트엔드



박준형

감성분석
Pseudo-labeling

감성분석 실험 진행



양봉석

일기 데이터 정제

코멘트 생성
모델 개발



이예령

감성분석
Pseudo-labeling

감성분석 실험 진행

백엔드 구축



우울했던 날에는 따뜻한 위로를, 힘들었던 날에는 격려를
당신의 하루가 궁금해요

일기 감성분석 및 코멘트 생성 서비스 '너하궁(너의 하루가 궁금해)'

01. Introduction

일기 감성분석 및 코멘트 생성 서비스 '너하궁(너의 하루가 궁금해)'

기획 의도

공감과 위로를 받고 싶은 순간,
초등학생 시절 일기장에 따듯한 말
한마디를 적어 주시던 선생님처럼,
누군가 내 하루의 이야기를 들어주고
내가 느낀 감정에 공감해준다면 어떨까?

너 하 궁 🌙

당신의 하루를 이야기해주세요.

<부스트캠프를 마치며>
오늘은 부스트캠프를 수료하는 날이다.
솔직히 말해서 부스트캠프는 내게 너무 어려운 과정이었다.
강의도 어려웠고, 경진대회를 할 때도 나 때문에 팀원들만 더 고생하는 것 같아서 미안한 마음이 들었다.

그렇지만 열정적으로 강의해주시는 마스터님들, 언제나 응원과 조원을 아끼지 않는 멘토님들, 그리고 내가 지칠 때마다 기를 불어넣어준 팀원들 덕분에 무사히 여기까지 올 수 있었던 것 같다.
부캠을 하기 전에는 답자의 '답'자도 몰랐던 내가 어느새 일기를 쓰면 감정을 분석해주고 코멘트를 달아주는 프로젝트에 참여해서 완성까지 오게 되었다.

열심히 한 만큼 꽤 괜찮은 서비스가 만들어진 것 같아서 참 뿌듯한 느낌이 든다.
많은 사람들이 이 서비스를 통해 위로를 받고 힘을 얻었으면 좋겠다.

나의 하루 보내기

일기 입력

당신의 하루에서 느껴지는 감정 해시태그 와 당신에게 들려주고 싶은 제 마음 을 전할게요.

#고마운 하루

부스트 캠프를 마치며 기분이 좋으셨군요. 열심히 한 만큼 좋은 서비스가 만들어진 것 같아 뿌듯한 마음이 드시는군요. 많은 사람들이 이 서비스를 통해 위로를 받고 힘을 얻었으면 좋겠네요.

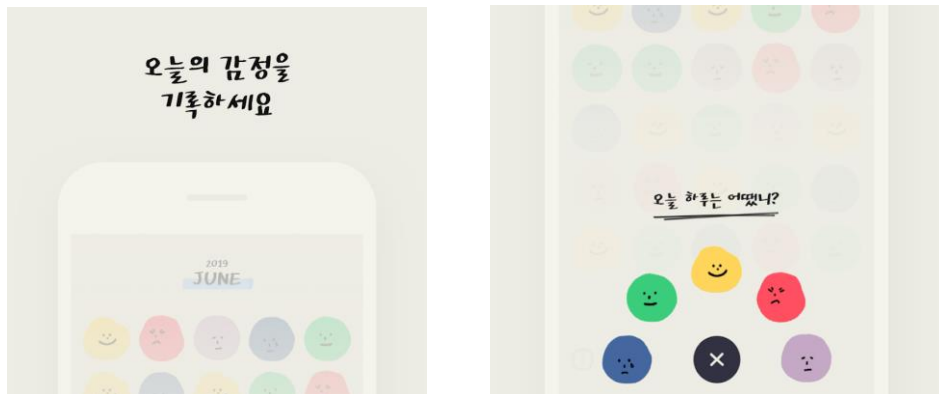
감성분석 결과

코멘트 생성

01. Introduction

일기 감성분석 및 코멘트 생성 서비스 '너하궁(너의 하루가 궁금해)'

기존 서비스와의 차별성



감정기록 다이어리 어플 '무다(MOODA)'

사용 대상 & 기대효과

사용 대상 전 연령층

기대효과 감성 분석을 통한 사용자의 멘탈 헬스케어
개인화된 공감 코멘트를 통한 위로와 격려

Emotion Tracking

사용자가 스스로 오늘의 감정을 선택하는 것이 아닌,
AI가 일기를 기반으로 감성 분석

Sympathetic Comment

Rule-based의 단순한 코멘트가 아닌,
일기의 구체적인 상황에 맞는 공감하는 코멘트를 생성

02. DATA & MODELING

2-1. 데이터

2-2. 감성분석 모델링

2-3. 코멘트 생성 모델링

DATA

데이터셋 수집

일기 데이터

모두의 말뭉치



비출판물 데이터 (ver 1.2)

- 교정을 거치지 않았으며, 어디에도 공개되지 않은 개인적 글쓰기 자료로 구성
(시(동시), 일기, 편지 글, 소설(동화), 감상문, 기타)
- 초등학생부터 80대 성인까지 총 5,937명의 자료
- 대부분이 '일기'인 데이터 수 : 총 4,236개

추가 활용 데이터

AIHub



감성 대화 말뭉치

- 사용자와 챗봇이 감성적인 상황에서 대화하는 데이터셋으로, 최대 3턴의 대화로 이루어짐
- 정신건강을 위한 감성대화 서비스 필요에 의해 구축된 데이터셋이기 때문에 부정적인 감정이 많음 (60가지 감정 상태)
- 총 58,268개의 대화

데이터 정제

세부 카테고리 중 일기 카테고리에 대해서만 사용

제외 대상

길이가 너무 짧은 데이터 (3문장 이하, KSS 활용)
카테고리가 '일기'로 되어있지만 '일기'로 보기 어려운 데이터

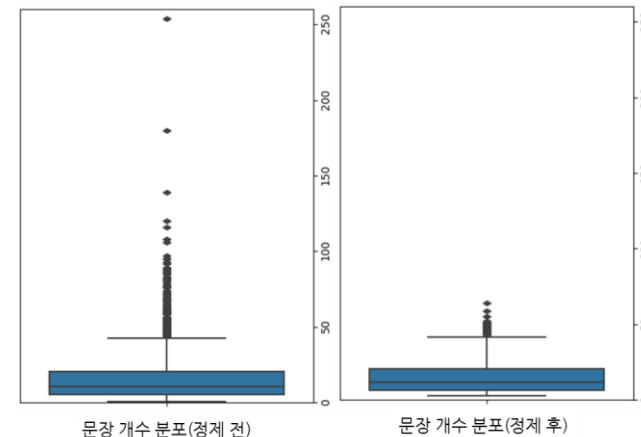
재가공 필요

길이가 너무 긴 데이터 ⇒ 분리

'길이가 길다'의 기준 : KoBART 토큰나이저로 토큰화를 했을 때 토큰 수가 500 이상일 때 분리

여러 개의 일기가 한번에 있는 데이터 ⇒ 분리

비식별화가 필요한 데이터 ⇒ 비식별화 처리



모두의 말뭉치



비출판물 데이터
(ver 1.2)

재가공 예시

2022/09/19
부스트캠프 첫날이다. 오리엔테이션도 하고 뭔가 신기한
감정이 많이 드는 하루이다. 첫 주부터 강의가 많은데 이게
맞나? 수학 너무 어려운 것 같다.
2023/02/14
벌써 부스트캠프 수료식이네. 원래 일기를 꾸준히
쓰려했는데 바쁘다보니 못썼네. 나 자신 진짜 고생했다.
다들 고맙고 강부덕 캠퍼님을 비롯해서 다른 캠퍼님들이
도와주셔서 진짜 많이 배울 수 있었다.

2022/09/19
부스트캠프 첫날이다. 오리엔테이션도 하고 뭔가
신기한 감정이 많이 드는 하루이다. 첫 주부터 강의가
많은데 이게 맞나? 수학 너무 어려운 것 같다.

2023/02/14
벌써 부스트캠프 수료식이네. 원래 일기를 꾸준히
쓰려했는데 바쁘다보니 못썼네. 나 자신 진짜 고생했다.
다들 고맙고 &name1& 캠퍼님을 비롯해서 다른
캠퍼님들이 도와주셔서 진짜 많이 배울 수 있었다.

데이터 정제

재가공 예시

사용자의 발화, 챗봇의 발화를 각각 이어 붙여 일기 및 코멘트로 이용
코멘트와 맞지 않는 챗봇의 발화는 수정 및 제거

- 질문, 추가 정보를 요구하는 경우
- 단순 응답인 경우 ("응", "그래", "아니야" 등)
- 상대방 지칭 언어 변경 ("사용자님" → "당신")
- 단순 상황 요약의 경우 해당 감정에서 해줄 수 있는 상용적인 어구 추가 (ex. 당신이 늘 기분이 좋기를 바랄게요!)

아내가 빨래 개는 일을 시켰는데 생각처럼 잘되지 않아 속상해

아내가 시킨 일이 생각처럼 되지 않아 속상하시군요.

만성질환으로 몸이 아프다 보니 쉬운 일도 집중하기 어렵고 다 끝나치지 못하겠어.

전에 이런 일이 있을 때는 어떻게 하셨나요?

늦더라도 하나씩 차근차근 해결하려고 노력해봐야겠어.

사용자님이 하는 모든 일을 언제나 응원할게요!

일기

아내가 빨래 개는 일을 시켰는데 생각처럼 잘되지 않아 속상해. 만성질환으로 몸이 아프다 보니 쉬운 일도 집중하기 어렵고 다 끝나치지 못하겠어. 늦더라도 하나씩 차근차근 해결하려고 노력해 봐야겠어.

코멘트

아내가 시킨 일이 생각처럼 되지 않아 속상하시군요. 당신이 하는 모든 일을 언제나 응원할게요!

AIHub



감성 대화 말뭉치

02. DATA & MODELING

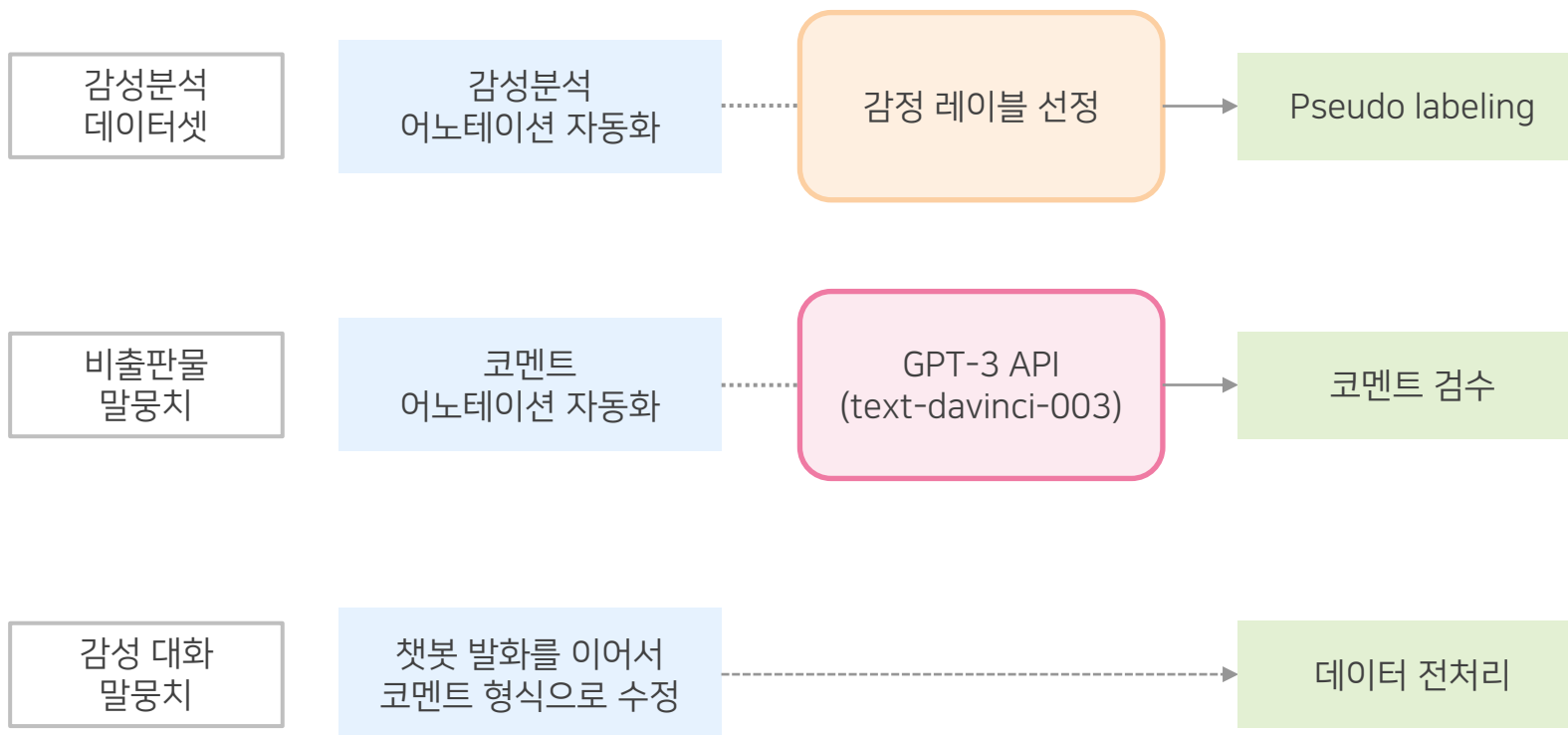
데이터셋 구축 과정

문제 상황

비출판물 데이터의 일기 데이터에 대한 코멘트 데이터 및 감정 레이블 필요
서비스에 맞는 감정 레이블 필요

해결 방안

Pseudo labeling을 통한 어노테이션 자동화



02. DATA & MODELING

감성분석 레이블 선정

1차 레이블

긍정	중립	부정
기쁨(행복한)	일상적인	슬픔(우울한)
감사하는	평온한	힘듦(지침)
설레는(기대하는)	피곤한	짜증남
사랑하는		걱정스러운
뿌듯한		기분 나쁨(불쾌함)



일기 데이터
파일럿 태깅

2차 레이블

긍정	중립	부정
기쁨(행복한)	일상적인	슬픔(우울한)
감사하는	평온한	힘듦(지침)
설레는(기대하는)	피곤한	짜증남
사랑하는	생각이 많은	걱정스러운
뿌듯한	다짐하는	기분 나쁨(불쾌함)
즐거움(신나는)	추억하는	후회하는

- 감성대화 말뭉치의 데이터는 부정적인 감정 위주의 레이블로 구성되어 있음
- 중립/긍정 감정을 위한 레이블 필요

- 감정 레이블 파일럿 태깅 후 필요한 감정 추가 및 삭제

02. DATA & MODELING

감성분석 어노테이션 자동화

감성분석 데이터 수집

AIHub

- 감성 대화 말뭉치
- 감정 분류를 위한 대화 음성 데이터셋
- 한국어 감정 정보가 포함된 단발성 대화 데이터셋
- 한국어 감정 정보가 포함된 연속적 대화 데이터셋

} 에크만의 기본정서 6가지와 '중립'으로 이루어져 있음
(분노, 혐오, 공포, 행복, 슬픔, 놀람) + 중립

우리 서비스와 맞지 않는 레이블로 구성 ⇒ 미리 정의한 감정 레이블 15개 중 4개의 레이블만 분류 가능
기존에 있는 모델로 unlabeled 데이터에 대해 레이블링을 하는 방법론인 **Pseudo Labeling** 진행

KOTE (Korean Online That-gul Emotions) Dataset

- 44개의 감성으로 레이블링
- 미리 정의한 감정 레이블 15개 중 12개 분류 가능

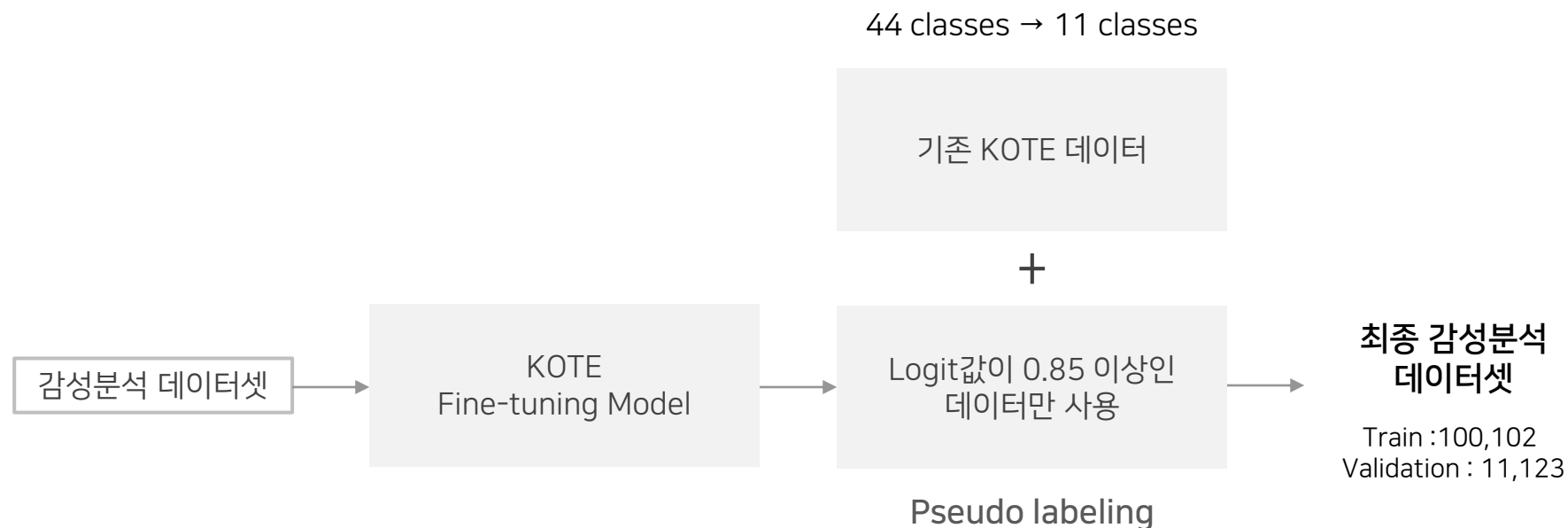
KOTE Fine-tuned model

- KOTE Dataset을 발표한 논문에서 fine-tuning된 모델도 공개
- KcELECTRA를 사용하여 fine-tuning
⇒ 이 모델을 사용하여 Pseudo labeling 진행

Jeon, D., Lee, J., & Kim, C. (2022). User Guide for KOTE: Korean Online Comments Emotions Dataset. arXiv preprint arXiv:2205.05300.

02. DATA & MODELING

감성분석 어노테이션 자동화



Logit값 임계값 설정

데이터의 신뢰도를 확보하기 위해 logit값을 기준으로 데이터를 선별해야 했는데, 각 logit에 대해 0.5부터 0.05단위로 임계치를 높여서 확인해본 결과 임계치 근처의 데이터들이 0.85인 경우 잘못 예측된 데이터가 적었으며 데이터의 개수도 68% 정도로 상당한 수가 유지되었기 때문에 logit threshold 값을 0.85로 지정하였다.

02. DATA & MODELING

감성분석 레이블 선정

2차 레이블			최종 레이블		
긍정	중립	부정	긍정	중립	부정
기쁨(행복한)	일상적인	슬픔(우울한)	기쁨(행복한)	일상적인	슬픔(우울한)
감사하는	생각이 많은	힘듦(지침)	감사하는	생각이 많은	힘듦(지침)
설레는(기대하는)	다짐하는	짜증남	설레는(기대하는)	다짐하는	짜증남
사랑하는	추억하는	걱정스러운	사랑하는	추억하는	걱정스러운
뿌듯한		후회하는	뿌듯한		후회하는
즐거움(신나는)			즐거움(신나는)		

Pseudo-label된
데이터 검수 결과

총 11개의 감정 레이블 선정

KOTE에는 존재하지 않는
레이블 제거

다짐하는, 추억하는, 후회하는

KOTE의 레이블과
유사하지 않은 레이블 제거

뿌듯한

KOTE에서 “뿌듯한”으로 태깅된 데이터 예시
⇒ “별 생각없었는데 검색해보고 알았네..고맙습니다..이젠 거를수있겠네요...”
일기 데이터에서 “뿌듯한”으로 태깅된 데이터 예시
⇒ “다이어리를 보니 이때까지 공부했던 모든 것이 나와있어서 뿌듯했다.”

02. DATA & MODELING

코멘트 어노테이션 자동화

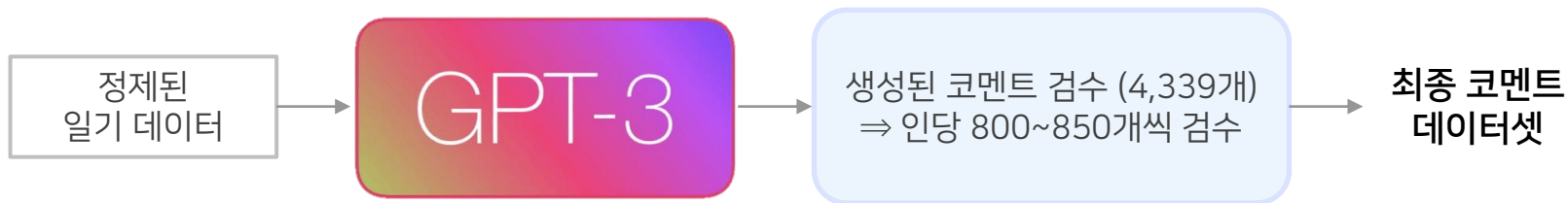


OpenAI의 GPT-3를 기반으로 만든 text-davinci-003 모델 엔진 사용

GPT-3

Our GPT-3 models can understand and generate natural language. We offer four main models with different levels of power suitable for different tasks. Davinci is the most capable model, and Ada is the fastest.

LATEST MODEL	DESCRIPTION	MAX REQUEST	TRAINING DATA
text-davinci-003	Most capable GPT-3 model. Can do any task the other models can do, often with higher quality, longer output and better instruction-following. Also supports inserting completions within text.	4,000 tokens	Up to Jun 2021



- 반말 → 존댓말 치환
- 복수형 주어 → 단수형 주어 (ex. 여러분 → 당신)
- 오타, 맥락/상식에 맞지 않는 문장 제거

MODELING

02. DATA & MODELING

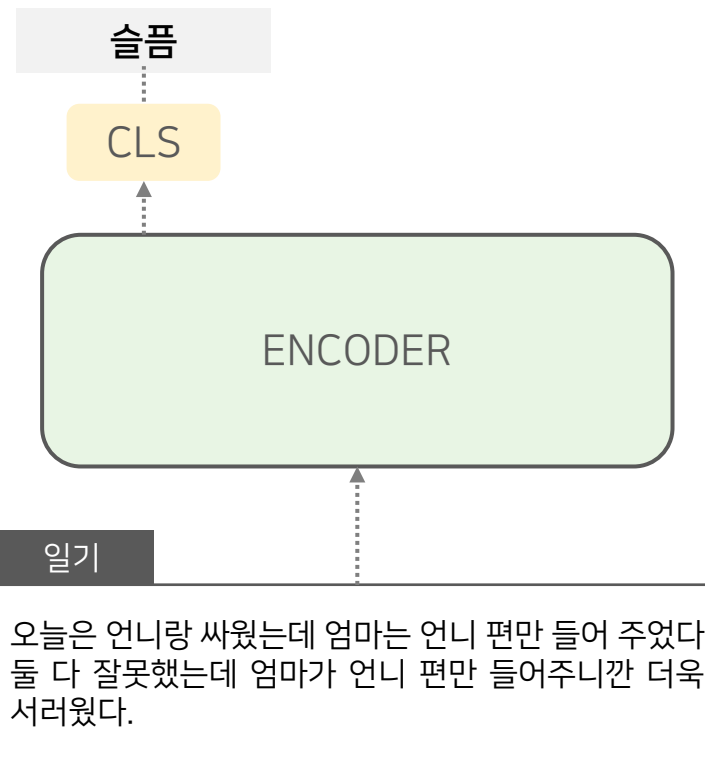
감성분석 모델

KcELECTRA

- 한국어 댓글로 pre-training된 모델 ⇒ 일기와 동일한 구어체
- Pseudo-labeling에 사용했던 KOTE Model에서 활용했던 모델

RoBERTa-large (klue)

- 분류 Task에서 ELECTRA 모델보다 높은 성능을 보임 (KLUE 논문 기준)



[Park, S., Moon, J., Kim, S., Cho, W. I., Han, J., Park, J., ... & Cho, K. \(2021\). Klue: Korean language understanding evaluation. arXiv preprint arXiv:2105.09680.](#)

02. DATA & MODELING

감성분석 모델

Validation Set	F1-Score	Accuracy
RoBERTa-large (klue)	70.601	0.706
KcELECTRA	70.718	0.707

두 모델이 성능지표에서 유의미한 차이가 나지 않았기 때문에
일기 데이터셋에 대해 inference를 하여 결과의 차이를 직접 관찰

⇒ 대분류를 반대되는 감정으로 예측한 경우가 약 3%
(긍정 → 부정, 부정 → 긍정)

⇒ RoBERTa-large가 **전체적인 감정에 초점을 맞춤**

일기

...(중략)...종이냅킨도 오리려고 하는데 오빠가 내말을 따라하면서
놀렸다. 가뜰이나 오리는 것도 자신 없는데 복잡하니까 더 힘들었다.
...(중략)...더 얇아서 풀칠을 잘 하지도 못하고(찢어 질까봐) 그러니
당연히 말라버리고, 결국 푸의 분홍친구 목이 그만..... 등을 완성했다!
성능을 시험해보려고 LED 등을 켜고 어두운 곳에서 켜보았다.
약하지만 예뻐다. 알록달록 하고 색도 가지각색으로 빛났다.
&name1& 이와 &name2& 오빠보다는 못 하지만 못지않게 예뻐다.

RoBERTa-large (klue)

즐거움(신나는)

KcELECTRA

슬픔(우울한)

02. DATA & MODELING

감성분석 모델 평가지표 - Human evaluation

문제 상황

기존의 Accuracy나 F1-score는
11개의 감정 예측을 평가하기에는 한계가 있음

해결 방안

정량평가 지표 설정

감정 레이블

긍정	중립	부정
기쁨(행복한)	일상적인	슬픔(우울한)
감사하는	생각이 많은	힘듦(지침)
설레는(기대하는)		짜증남
사랑하는		걱정스러운
즐거움(신나는)		

평가 가이드라인

2점

일기에서 드러나는 주된 감정이
예측한 감정과 정확히 일치한다고 판단될 때

'감사하는'으로 태깅된 일기에서 평가자가
'감사하는' 감정이 주된 감정이라고 느꼈다면 2점 부여

1점

일기에서 드러나는 주된 감정이 예측한 감정과
다르지만, 같은 대분류(긍정/중립/부정)에
속하는 감정이라고 판단될 때

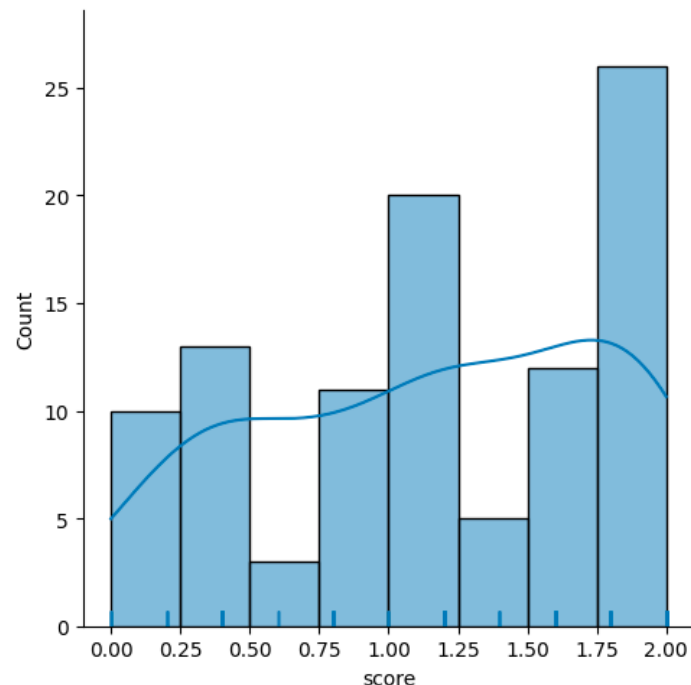
'힘듦(지침)'으로 태깅된 일기에서 평가자가 '슬픔(우울한)' 감정이
주된 감정이라고 느꼈다면 같은 대분류에 속하므로 1점 부여

0점

일기에서 드러나는 주된 감정이 예측한
감정과 전혀 일치하지 않는 경우 0점을 부여

'기쁨(행복한)'으로 태깅된 일기에서 평가자가
'짜증남' 감정이 주된 감정이라고 느꼈다면 0점 부여

감성분석 모델 평가지표 - Human evaluation



각각의 데이터에 대한 평가자가 매긴 점수의 평균 점수 분포

평가 분석

Test data 중 감성 분석을 했을 때
분류하기 어려운 (logit값이 0.451보다 낮은) 데이터 100개 선별하여 평가

Average : 1.16

- 전체적으로 감정의 경향을 맞춘다

Acceptable ratio : 0.61

- 두 명 이상의 평가자가 세부적인 감정을 정확히 맞췄다고 판단하는 샘플의 비율

감정을 맞추지 못한 샘플 중 다양한 감정이 하나의 일기에 들어있는 경우가 많았고
이를 보완하기 위해 Top-1 logit과 Top-2 logit이 비슷한 경우
Top-2에 해당하는 감정도 출력하기로 결정

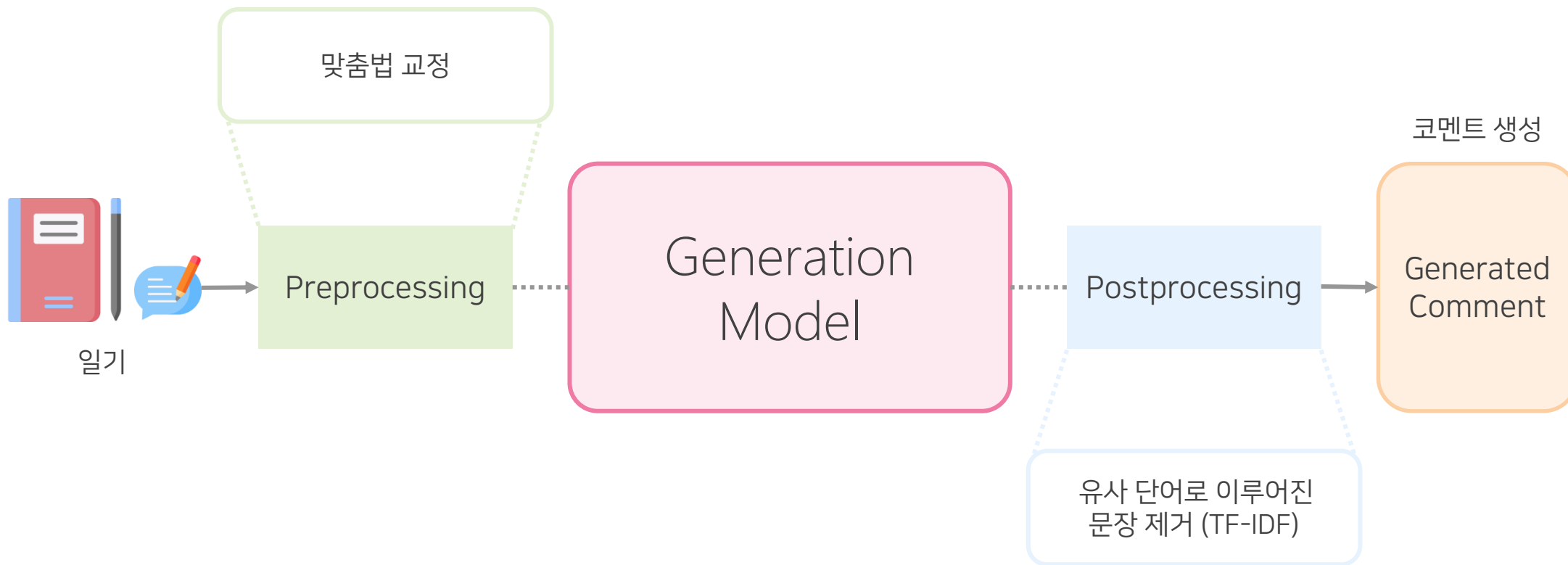
일기

...(중략)... 내가 좋아하는 양념 꼬막이어서 맛있었다.
...(중략)... 그런데 우리집을 지나치고 가버려서
실망했다. 그래서 지하철을 타고 집에 도착했다.
그냥 기차를 탈 걸 그랬다.

다양한 감정이 하나의 일기에 들어있는 예시

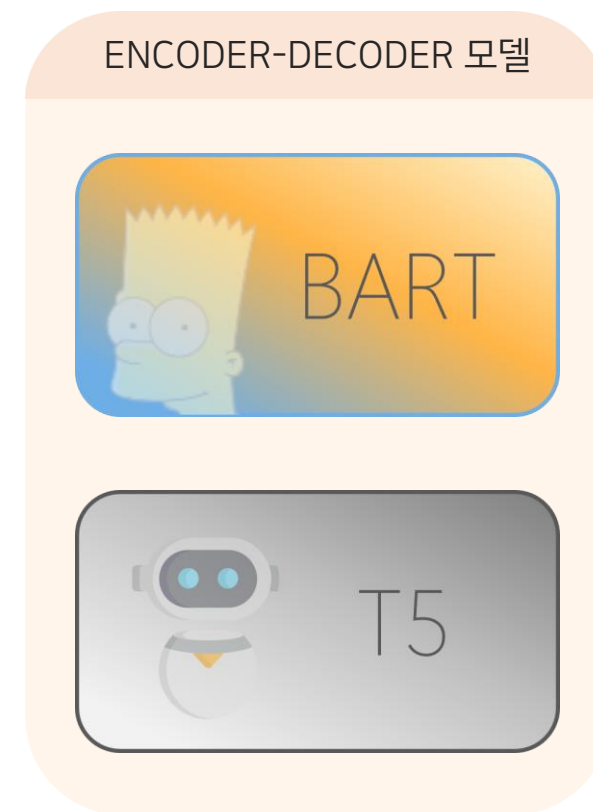
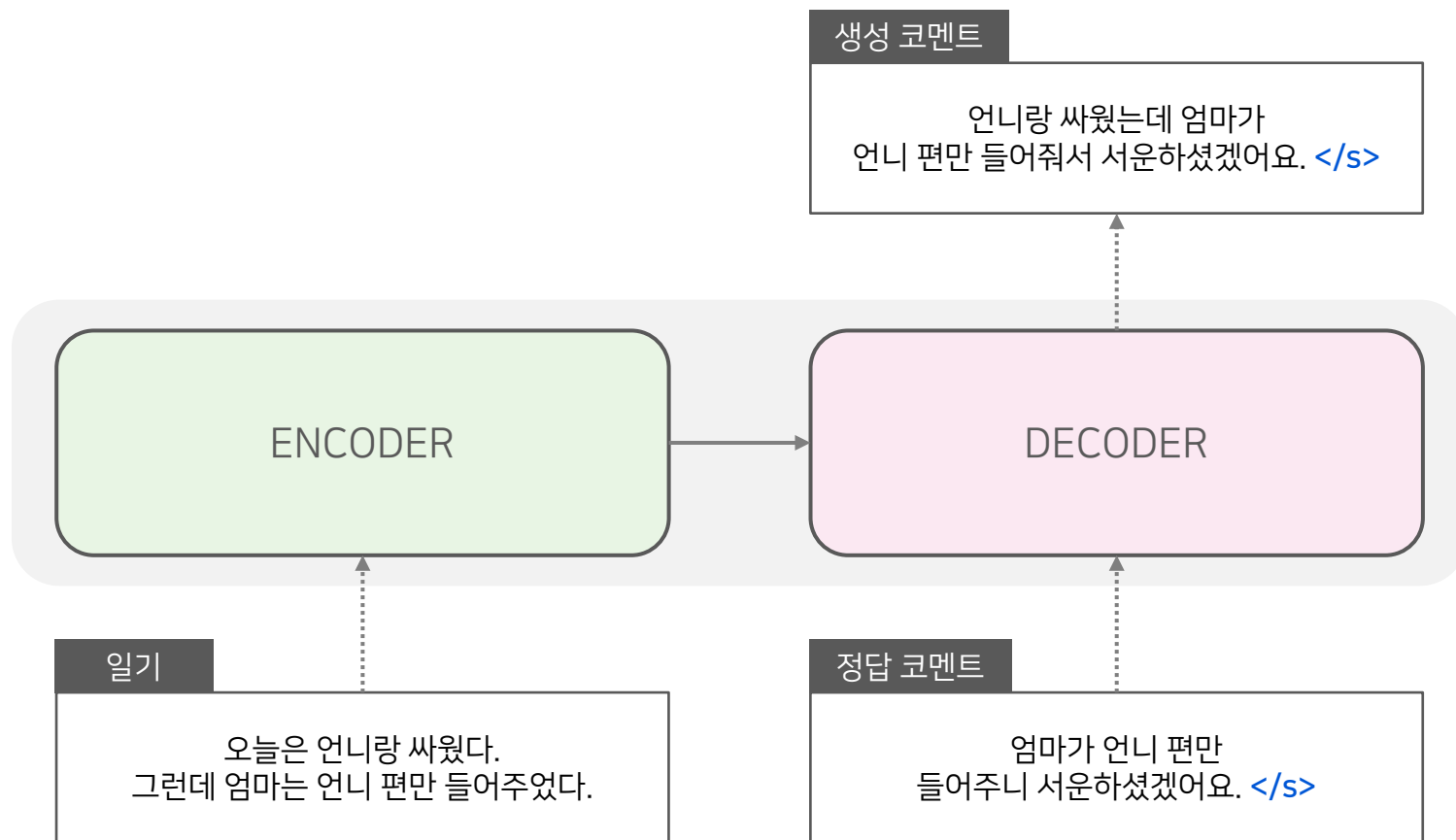
Top-1 감정	Top-1 logit	Top-2 감정	Top-2 logit
즐거움 (신나는)	0.27	슬픔 (우울한)	0.27

코멘트 생성 모델 아키텍처

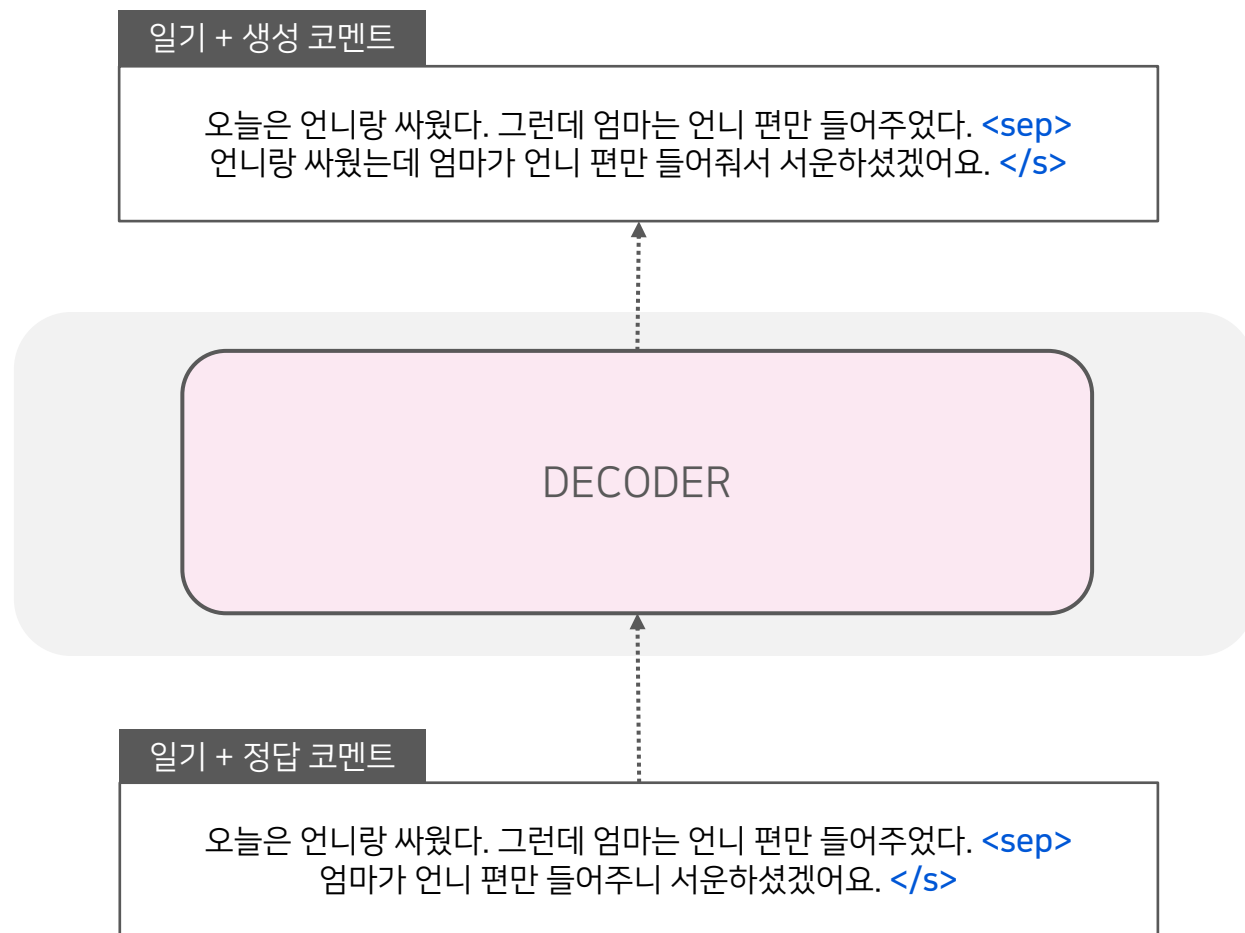


02. DATA & MODELING

코멘트 생성 모델 - Seq2seq

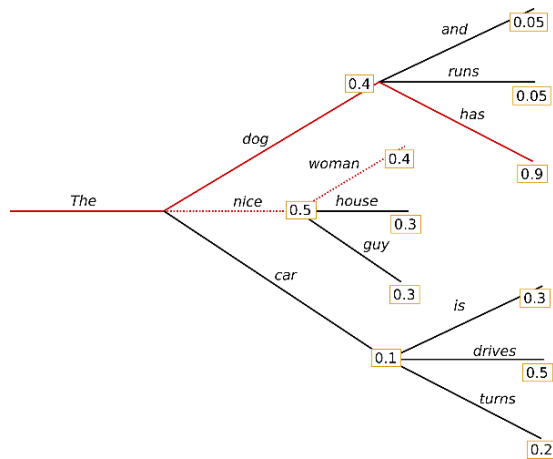


코멘트 생성 모델 - Decoder

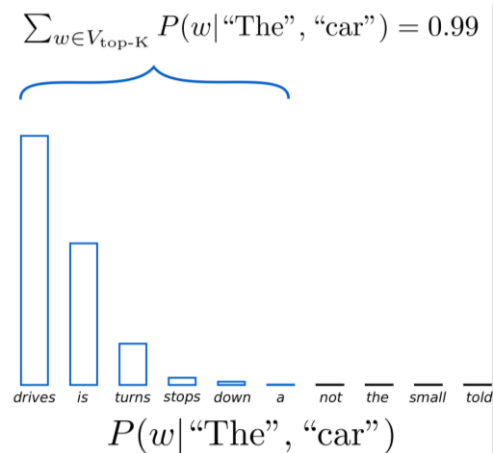


02. DATA & MODELING

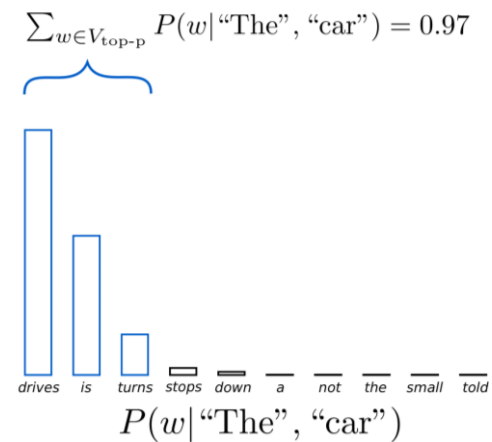
코멘트 생성 옵션



Beam search



Top-K sampling



Top-p sampling

Top-K, Top-p sampling을 통한 틀에 박히지 않은 코멘트 생성

일기에서 묘사되는
다양한 상황 언급

Beam search를 통해 안정적인 코멘트 생성

일기의 감정에 어울리는
공감 코멘트 생성

02. DATA & MODELING

코멘트 생성 모델 평가지표

ROUGE

- 정답 문장과 모델이 생성한 문장의 유사도를 비교, 두 문장 간에 n-gram이 얼마나 많이 겹치는지 계산

$$ROUGE - N = \frac{\sum_{S \in \{Reference\ Summaries\}} \sum_{gram_n \in S} Count_{match}(gram_n)}{\sum_{S \in \{Reference\ Summaries\}} \sum_{gram_n \in S} Count(gram_n)}$$

RDASS

- 문서(v_d), 정답 문장(v_r), 생성 문장(v_p)사이의 관계 고려
- Sentence BERT를 이용해 벡터 v_d, v_r, v_p 추출
- 각 vector 간 cosine similarity의 평균을 지표로 사용

$$s(p, r) = \cos(v_p, v_r) = \frac{v_p^T \cdot v_r}{\|v_p\| \|v_r\|}$$

$$s(p, d) = \cos(v_p, v_d) = \frac{v_p^T \cdot v_d}{\|v_p\| \|v_d\|}$$

$$RDASS = \frac{s(p, r) + s(p, d)}{2}$$

BLEU

- '목표로 하는 문장과 모델이 생성한 문장이 일치하는 정도'를 수치화 한 지표

$$BLEU = \min(1, \frac{output\ length(예\ 측\ 문장)}{reference\ length(정\ 답\ 문장)}) (\prod_{i=1}^4 precision_i)^{\frac{1}{4}}$$

일기

오늘은 언니랑 싸웠다.
그런데 엄마는 언니 편만 들어주었다.

코멘트

언니랑 싸웠는데 엄마가
언니 편만 들어줘서 서운하셨겠어요.

일기 코멘트는 일기에 나타나는 주요 사건들에 대한 이해와 주요 감정에 대한 공감 메시지를 담고 있으므로 요약 TASK와 유사함

⇒ 요약과 같은 Generated Sentence를 평가하는 지표 ROUGE, BLEU, RDASS를 사용

코멘트 생성 모델 평가지표

ROUGE, BLEU를 정량평가 기준으로 선정할 시
의미적 유사도는 높으나 형태학적 유사도가 낮은 문장을 생성하는 경우,
모델의 성능이 낮다고 판단하는 문제 존재

정답 코멘트

오션월드에 가셔서
즐거운 시간을 보내셨군요!
미끄럼틀과 파도에 놀라셨는지 모르겠지만,
여러분이 같이 함께해서 더 재밌게 놀았겠죠!
여러분의 여행기를 기록하셔서 오랜 추억을
간직하세요.

생성 코멘트

정말 재밌는 여행이었군요! 점프하는 것도 힘들고
물에 빠지는 일도 없었으니 정말 즐거웠을 것
같아요. 그리고 미끄럼틀 타는 것도
재미있었겠네요. 다음에 또 가보는 건 어떨까요?
기대해보세요!

ROUGE-1	ROUGE-L	BLEU-1	RDASS
24.53	9.43	23.06	0.58

RDASS 지표를 정량평가의 주요 평가 기준으로 채택

02. DATA & MODELING

코멘트 생성 모델 비교

모델 성능 비교

	ROUGE-1	ROUGE-L	BLEU-1	RDASS
KE-T5-large	28.87	19.41	23.51	0.46
KoGPT2	33.56	22.34	27.25	0.56
KoBART	35.66	23.26	29.79	0.64

ROUGE, BLEU, RDASS 기준 모두 높은 성능을 보였던
KoBART 모델을 최종 모델로 선정

입력 일기

오늘은 언니랑 싸웠다. 그런데 엄마는 언니 편만 들어주었다.
둘 다 잘못했는데 엄마가 언니 편만 들어주니깐 더욱 서러웠다.
맨날 나만 혼나서 억울하다.

KE-T5

어머니께 마음이 안 좋으시군요. 부모님과의 대화로 기분이
풀렸으면 좋겠어요. 잘 해결되었으면 좋겠네요. 좋은 일이 잘
되길 바라요.

KoGPT2

오늘 언니와 싸워서서 속상하셨겠어요. 엄마 편을 들어주지
않아 억울하시군요. 억울함을 풀기 위해 어떻게 하는 것이
좋을까요?

KoBART

엄마가 언니 편만 들어주니 서러우셨겠어요. 맨날 나만 혼나서
억울한 마음이 들었겠네요. 다음번에는 언니랑 다투지 않기를
바래요.

02. DATA & MODELING

코멘트 생성 모델 평가지표 - Human evaluation

좋은 코멘트란?

- ① 문장이 문법적·의미적·논리적으로 자연스럽고
- ② 일기에서 언급하는 사건에 대한 구체적인 언급을 해주며
- ③ 일기에서 드러나는 감정에 대해 공감해주는 코멘트

Sensibleness

Specificity

Sympathy

SSSA
Sensibleness and
Specificity and
Sympathy
Average

평가 분석

* Test dataset 100개에 대해 직접 평가

	Sensibleness	Specificity	Sympathy	SSSA
정답 코멘트	0.80	0.85	0.75	0.80
생성 코멘트	0.65	0.86	0.67	0.73

Sensibleness, Sympathy는 정답 코멘트보다 점수가 낮았지만,
Specificity는 정답 코멘트보다 향상된 점수를 달성함

Adiwardana, D., Luong, M. T., So, D. R., Hall, J., Fiedel, N., Thoppilan, R., ... & Le, Q. V. (2020). Towards a human-like open-domain chatbot. arXiv preprint arXiv:2001.09977.

03. PRODUCT SERVING

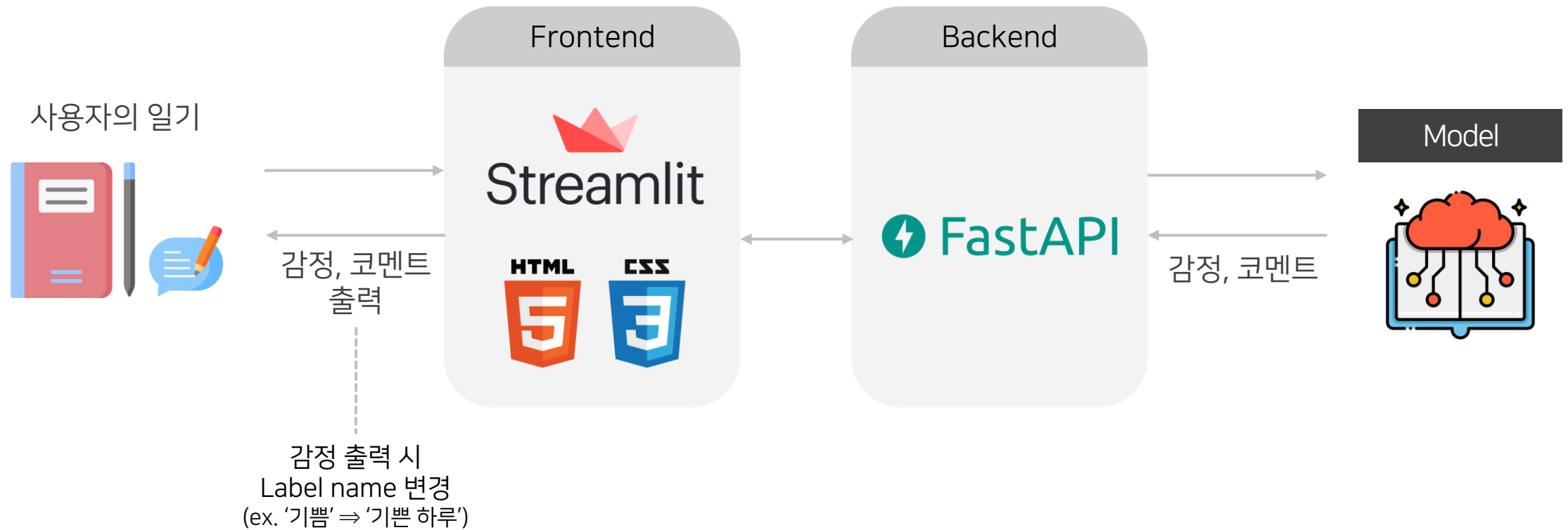
3-1. 서비스 아키텍처

3-2. Frontend

3-3. Backend

03. PRODUCT SERVING

서비스 아키텍처



03. PRODUCT SERVING

Frontend

너하궁

당신의 하루를 이야기해주세요.

일기를 최소 20자 이상 작성해주세요!

나의 하루 보내기

Streamlit, HTML/CSS 사용
일기 text 입력 모듈

<부스트캠프를 마치며>
오늘은 부스트캠프를 수료하는 날이다.
솔직히 말해서 부스트캠프는 내게 너무 어려운 과정이었다.
강의도 어려웠고, 경진대회를 할 때도 나 때문에 팀원들만 더 고생하는 것 같아서 미안한 마음이 들었다.

그렇지만 열정적으로 강의해주는 마스터님들, 언제나 응원과 조원을 아끼지 않는 멘토님들, 그리고 내가 지칠 때마다 기를 불어넣어준 팀원들 덕분에 무사히 여기까지 올 수 있었던 것 같다.
부캠을 하기 전에는 딥러닝의 '딥'자도 몰랐던 내가 어느새 일기를 쓰면 감정을 분석해주고 코멘트를 달아주는 프로젝트에 참여해서 완성까지 오게 되었다.

열심히 한 만큼 꽤 괜찮은 서비스가 만들어진 것 같아서 참 뿌듯한 느낌이 든다.
많은 사람들이 이 서비스를 통해 위로를 받고 힘을 얻었으면 좋겠다.

나의 하루 보내기

당신의 하루에서 느껴지는 감정 해시태그 와 당신에게 들려주고 싶은 제 마음 을 전할게요.

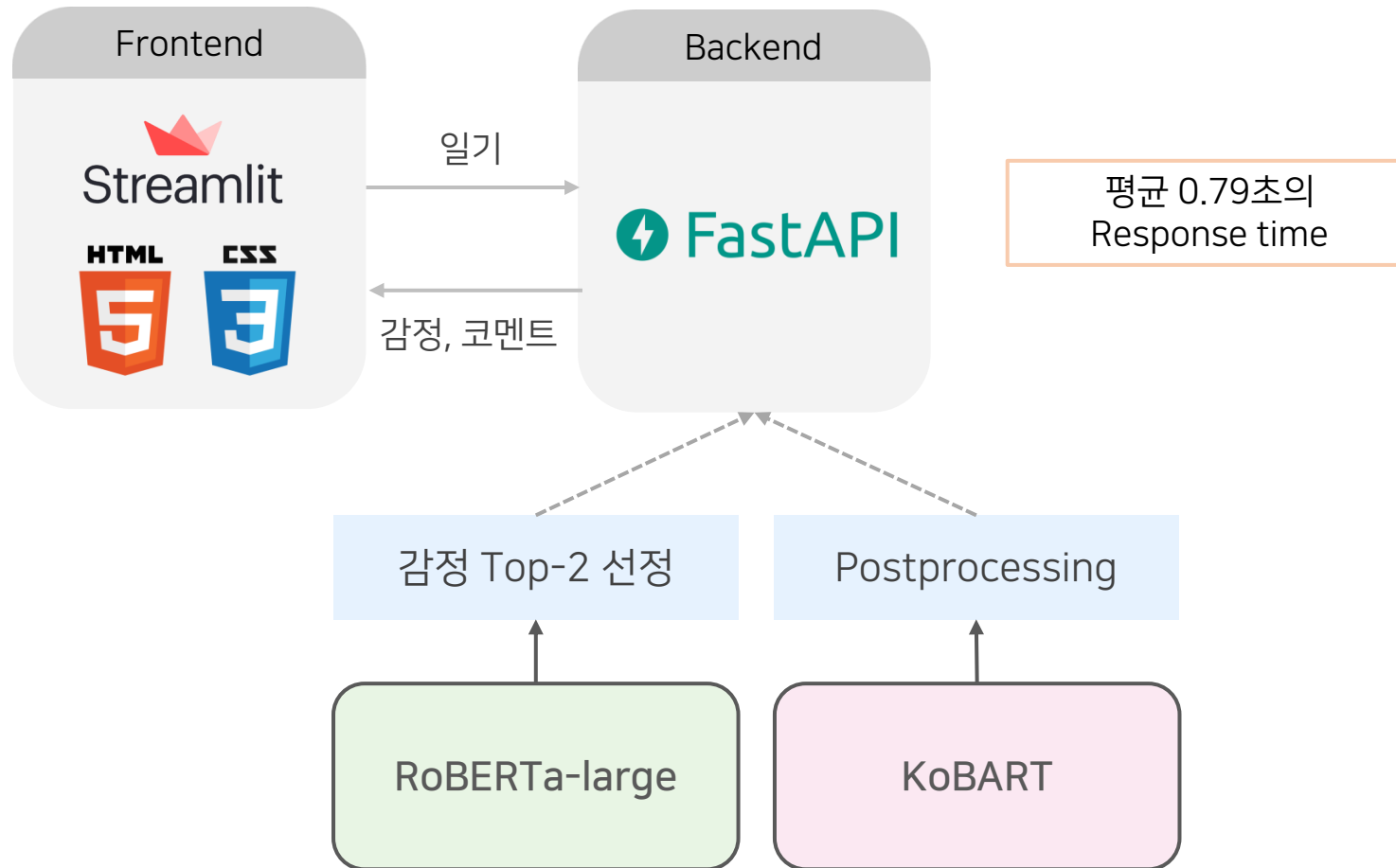
#고마운 하루

부스트 캠프를 마치며 감사한 마음이 드시는군요. 열심히 노력한 만큼 좋은 서비스가 만들어져서 뿌듯하시겠어요. 많은 사람들이 이 서비스를 통해 위로를 받고 힘을 얻었으면 좋겠네요.

사용자의 일기를 토대로 예측된 사용자의 감정 출력
사용자의 일기를 토대로 생성된 위로와 공감의 코멘트 출력

03. PRODUCT SERVING

Backend



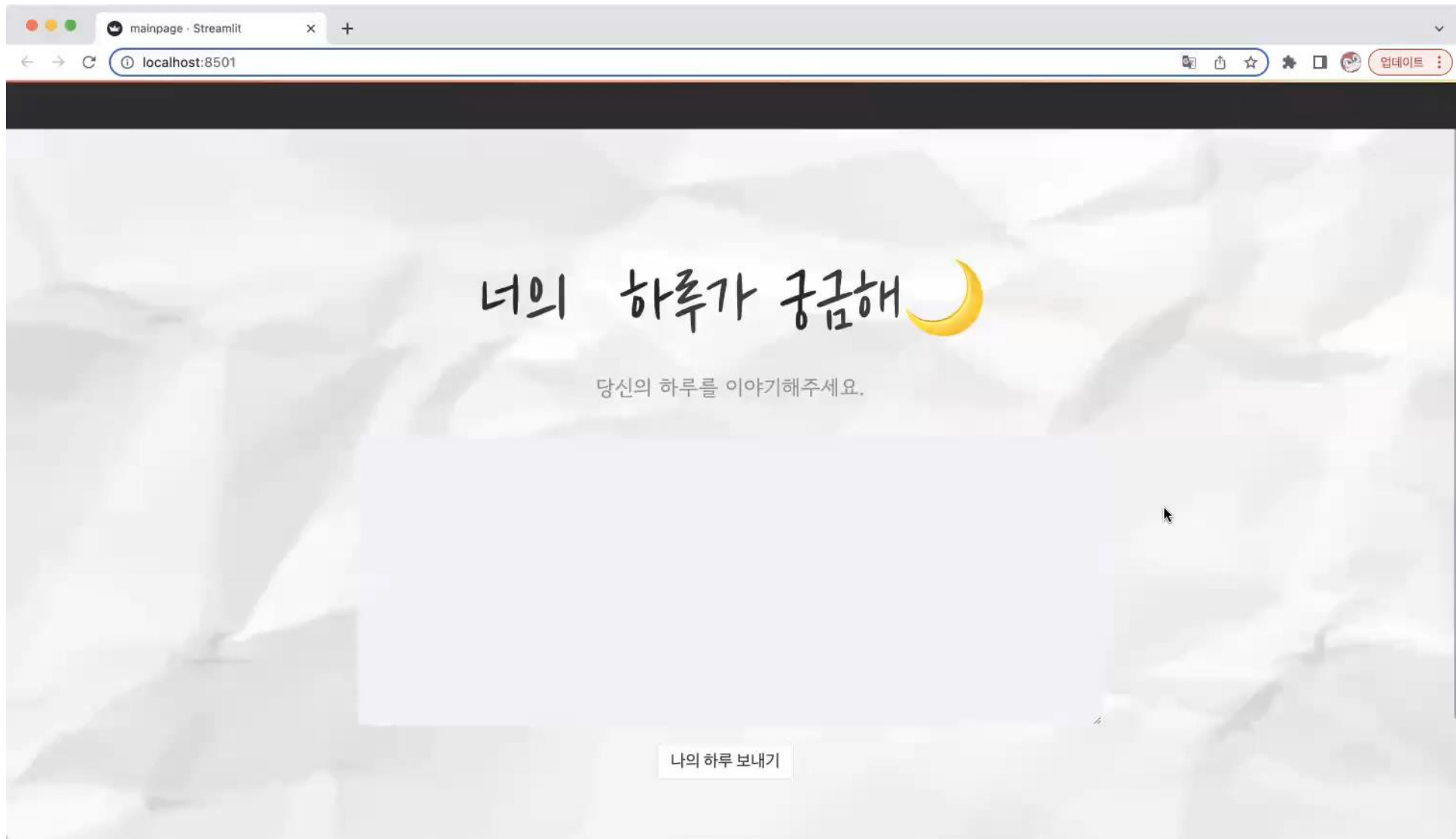
04. RESULT & CONCLUSION

4-1. 시연 영상

4-2. 개선 방향

04. RESULT & CONCLUSION

시연 영상



개선 방향

모델링

- 코멘트의 Sensibleness를 높이는 방안
→ 추가 데이터 학습 (ex. 수필 등)
- 일기의 전체적인 감정보다는 부분적인 감정 표현에 집중하여 분류하는 경향이 존재
→ 긴 글 데이터 추가 학습
- Inference 시간 단축 및 다양한 환경에서의 사용을 위한 모델 경량화 필요

서비스

- DB 연결을 통한 사용자 작성 일기 저장 및 열기 기능 구현
- Response time을 줄일 수 있는 방향으로 백엔드 개선 (현재 평균 0.79초)
- OCR과 STT를 활용하여 손글씨 일기나 음성 일기를 입력으로 받는 기능 추가

04. RESULT & CONCLUSION

참고자료

Models

- <https://huggingface.co/gogamza/kobart-base-v2>
- <https://huggingface.co/KETI-AIR/ke-t5-large-ko>
- <https://huggingface.co/skt/kogpt2-base-v2>
- <https://huggingface.co/klue/roberta-large>
- <https://huggingface.co/beomi/KcELECTRA-base>

Data

- 모두의 말뭉치 - 비출판물 말뭉치 <https://corpus.korean.go.kr/>
- AI hub - 감성대화 말뭉치 <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=86>
- AI hub - 감정분류를 위한 대화음성 데이터셋
<https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=120&topMenu=100&aihubDataSe=extrldata&dataSetSn=263>
- AI hub - 한국어 감정 정보가 포함된 단발성 대화 데이터셋
<https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=120&topMenu=100&aihubDataSe=extrldata&dataSetSn=270>
- AI hub - 한국어 감정 정보가 포함된 연속적 대화 데이터셋
<https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=120&topMenu=100&aihubDataSe=extrldata&dataSetSn=271>
- KOTE <https://github.com/searle-j/KOTE>

ETC

- Adiwardana, D., Luong, M. T., So, D. R., Hall, J., Fiedel, N., Thoppilan, R., ... & Le, Q. V. (2020). Towards a human-like open-domain chatbot. *arXiv preprint arXiv:2001.09977*.
- Lee, D., Shin, M., Whang, T., Cho, S., Ko, B., Lee, D., ... & Jo, J. (2020). Reference and document aware semantic evaluation methods for Korean language summarization. *arXiv preprint arXiv:2005.03510*.
- Park, S., Moon, J., Kim, S., Cho, W. I., Han, J., Park, J., ... & Cho, K. (2021). Klue: Korean language understanding evaluation. *arXiv preprint arXiv:2105.09680*.
- <https://github.com/boostcampaitech3/final-project-level3-nlp-01>

Docs

- <https://huggingface.co/docs>
- https://huggingface.co/docs/transformers/v4.26.0/en/generation_strategies

05. APPENDIX

Training & Deployment 서버 환경

OS

Ubuntu 18.04.5 LTS

CPU

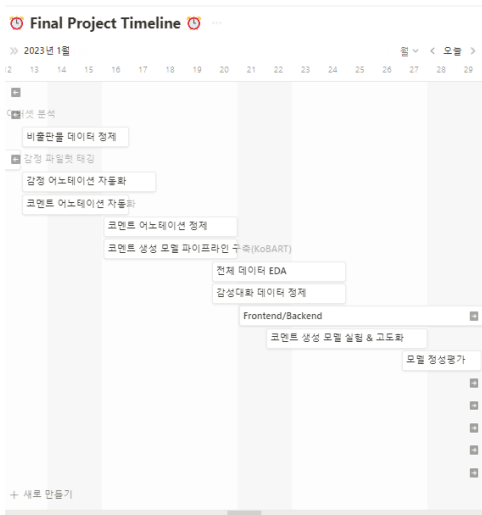
Intel(R) Xeon(R) Gold 5120 CPU @ 2.20GHz

GPU

Tesla V100

05. APPENDIX

협업 과정



- ▶ 스케줄
- ▶ 회의록
- ▶ 개인 노트
- ▶ 멘토링

프로젝트 메인

- Wrap-Up Report
- 최종 프로젝트 소개
- PPT
- 시연 영상
- README
- DATA
- MODELING
- BACK & FRONT

프로젝트 유틸

- [완료] 프로젝트 주제 탐색
- [완료] 최종프로젝트 주제 수합 제
- ▶ 데이터셋
- ▶ 리서치
- ▶ EDA
- ▶ 기획서
- ▶ 프로젝트 서비스
- ▶ QA

Feature/generation #4

Merged 54data merged 42 commits into dev from feature/generation last week

Conversation 0 Commits 42 Checks 0 Files changed 22



54data commented last week

Member

코멘트 생성 (KoBART)

- 생성 예시, 실험 내용, 평가 지표 설명은 팀 노션 KoBART 페이지에 정리해두었으니 참고 부탁드립니다.

Code Structure

```
generation
├── dataset
│   ├── diary_test.csv
│   ├── diary_train.csv
│   └── total_train.csv
├── arguments.py
├── inference.py
├── inference.sh
├── preprocessing.py
├── rouge_utils.py
├── train.py
├── train.sh
└── utils.py
```

train.py 실행

- arguments에 대한 설명을 적입하고 실행하면 arguments.py에서 확인 가능
- train.sh

Github commit message rule

태그 이름	설명
Feat	새로운 기능을 추가할 경우
Fix	버그를 고친 경우
Style	코드 포맷 변경, 세미 콜론 누락, 코드 자체에 대한 수정이 없는 경우
Refactor	프로덕션 코드 리팩토링
Comment	필요한 주석 추가 및 변경
Docs	문서를 수정한 경우
Exp	실험을 진행하는 경우
Rename	파일 혹은 폴더명을 수정하거나 옮기는 작업만인 경우
Remove	파일을 삭제하는 작업만 수행한 경우

Notion을 통한 일정 관리 및 진행 상황 공유

PR을 통한
코드 리뷰 진행

Git Flow 프로세스

commit message rule
설정

05. APPENDIX

협업 과정



nlp04-sajo Non-Profit

+ New Activity Feed Organization settings Unwatch repos

Research interests

None defined yet.

Team members 5



Organization Card

네이버 커넥트재단 부스트캠프 AI Tech NLP-04

Models 57

Sort: Recently Updated

nlp04/final_bart_prepro_fix

Updated 7 days ago · ↓ 2

nlp04/sentiment_roberta_large_with_diary

Updated 7 days ago · ↓ 7

nlp04/gpt_trinity_2_4_3e-5_lp5_nb5

Updated 7 days ago · ↓ 1

nlp04/gpt_trinity_16_4_3e-5_lp5_nb5

Updated 7 days ago

nlp04/kobart_32_4e-5_datav2_min30_lp5_0_temperature1.0

Updated 9 days ago · ↓ 2

nlp04/kobart_32_3e-5_datav2_min30_lp5_0_temperature1.0

Updated 9 days ago

nlp04/gpt_16_5_3e-5_lp5_nb5

Updated 10 days ago · ↓ 12

nlp04/gpt_16_4_3e-5_lp5_nb5

Updated 10 days ago · ↓ 35

nlp04/gpt_16_5_5.6e-5_lp5_nb10

Updated 11 days ago · ↓ 2

nlp04/final_bart

Updated 11 days ago · ↓ 122

Datasets 3

Sort: Recently Updated

nlp04/sentiment_analysis private

Updated 6 days ago · ↓ 36

nlp04/preprocessed_diary_dataset private

Updated 12 days ago · ↓ 118

nlp04/diary_dataset private

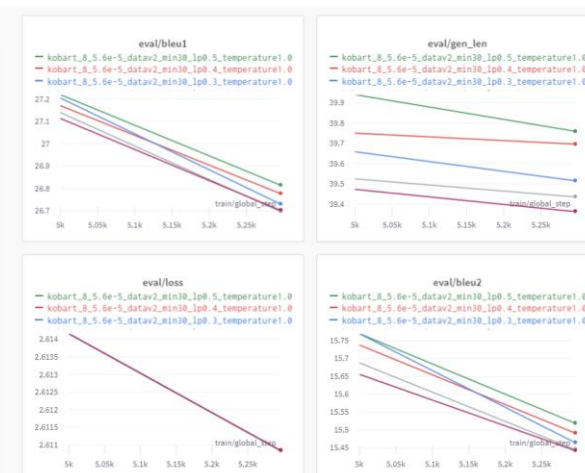
Updated 13 days ago · ↓ 203

HuggingFace Hub을 통한 모델 & 데이터셋 공유

Name (5 visualized)

- kobart_8_5.6e-5_datav2_min30_lp0.5_temperature1.0
- kobart_8_5.6e-5_datav2_min30_lp0.5_temperature1.0
- kobart_8_5.6e-5_datav2_min30_lp0.5_temperature1.0
- kobart_8_5.6e-5_datav2_min30_lp0.4_temperature1.0
- kobart_8_5.6e-5_datav2_min30_lp0.3_temperature1.0
- kobart_8_5.6e-5_datav2_min30_lp0.2_temperature1.0
- kobart_8_5.6e-5_datav2_min30_lp0.1_temperature1.0
- kobart_8_5.6e-5_datav2_min30_lp0.0_temperature1.0

41-48 of 48



WandB를 통한 실험 관리

Model Hyperparameters

Sentiment Analysis Hyperparameters

	Batch size	Learning rate	Epoch
RoBERTa-large (klue)	32 * 4	1e-5	5
KcELECTRA-base	32 * 8	2e-5	10

Generation Hyperparameters

	Batch size	Learning rate	Epoch
KE-T5-large	8	3e-5	5
KoGPT2	16	3e-5	4
KoBART	64	3e-5	5

감성분석 Human evaluation 가이드라인

감정의 대분류와 세부 감정에 대한 정의

대분류	세부 감정	정의
긍정	기쁨	어떠한 원인으로 인해 느껴진 기쁜 감정이 드러나는 글
	고마운	남이 베풀어 준 호의나 도움에 대하여 감사한 마음이 드러나는 글
	설레는(기대하는)	원하는 대로 이루어지기를 바라면서 기다리는 글
	즐거움(신나는)	외부적인 활동 및 경험으로 인해 즐거운 감정이 드러나는 글
	사랑하는	어떤 사람이나 존재를 몹시 아끼고 귀중히 여기며 생각하는 글
중립	일상적인	1. 일상생활에서 느낀 감정 중 위에서 정의한 감정에 해당하지 않는 경우 2. 감정이 드러나지 않은 겪었던 사건을 나열한 글
	생각이 많은	1. 후회하거나 성찰을 통해 자신의 생각 및 다짐을 쓴 글 2. 자신의 철학, 사상 등 생각을 위주로 쓴 글
부정	슬픔(우울한)	원통한 일을 겪거나 불쌍한 일을 보고 마음이 아프고 괴로운 감정이 드러나는 글
	힘듦(지침)	힘든 일을 하거나 어떤 일에 시달려서 기운이 빠지는 감정이 드러나는 글
	짜증남	마음에 탐탁지 않아하고, 화나는 감정이 드러나는 글
	걱정스러운(불안한)	일이 잘못되지 않을까 불안한 감정이 드러나는 글

코멘트 생성 Human evaluation 가이드라인

합리성(Sensibleness)

- 주어진 대화 맥락 내에서 말이 되는 발화인가?
- 일반적인 상식, 사실에 부합하는가?
- 문장이 문법적·의미적으로 자연스러운가?

구체성(Specificity)

- 주어진 맥락(일기)에 대해 구체적으로 언급하는가? 혹은 주어진 맥락과 밀접하게 관련된 언급을 하고 있는가?

공감성(Sympathy)

- 코멘트가 일기에서 드러나는 감정에 대해 잘 이해하고 공감해주거나, 혹은 위로해주고 있는가?

05. APPENDIX

모델 성능 및 오류 검수용 프로토타입

ProtoType

점심으로 중국집에서 탕수육과 짜장면을 시켜먹었다. 탕수육은 겉바속촉이었고 짜장면은 달콤하니 너무 맛있었다. 짜장면과 탕수육의 조합은 정말 좋은 조합인 것 같다. 처음에 탕수육과 짜장면을 동시에 하나의 메뉴로 만들겠다는 생각은 누가 한 걸까? 다음번에도 또 먹으러 와야겠다!

nlp04/kobart_8_5.6e-5_default_option Comment

탕수 면과 짜장면의 조합이 정말 좋은 조합이었군요. 다음번에도 또 맛있게 먹으러 오길 바랄게요!

nlp04/kobart_8_5.6e-5_min30_lp4_sample Comment

맛있는 탕수육과 짜장면을 먹으러 오는 건 정말 좋은 생각이네요! 다음번에도 즐거운 시간 보내시길 바랄게요!

Emotion

즐거움(신나는): 0.4517 기쁨(행복한): 0.4028

감정

- 코멘트 생성 모델 결과 및 감성분석의 logit 값을 확인하기 위해 사용
 - 팀 내부에서 모델 성능 및 오류 검수를 위해 사용
- ⇒ 전처리 및 후처리 필요 케이스 탐색

05. APPENDIX

감성분석 & 코멘트 생성 예시

입력 일기

점심으로 중국집에서 탕수육과 짜장면을 시켜먹었다. 탕수육은 겉바속촉이었고 짜장면은 달콤하니 너무 맛있었다. 짜장면과 탕수육의 조합은 정말 좋은 조합인 것 같다. 처음에 탕수육과 짜장면을 동시에 하나의 메뉴로 만들겠다는 생각은 누가 한 걸까? 다음번에도 또 먹으러 와야겠다!

감성분석

즐거운 하루,
기쁜 하루

코멘트 생성

좋은 조합이에요! 짜장면과
탕수육의 조합이 정말 좋군요.
다음번에도 또 같이 먹으러 가면
좋겠네요

입력 일기

오늘부로 분식집 알바를 그만두기로 했다. 마지막 날이라 대충 일하고 싶었는데 손님이 너무 많이 와서 바빴다. 이제는 점점 심해지는 습진 때문에 더 이상 식당 일은 못하게 될 것 같다. 다른 업종을 찾아봐야지. 그만두게 되어 속이 시원하지만 같이 일하던 사람들과 단골 손님들을 더 이상 못 본다는 생각을 하니 섭섭하기도 하다. 그래도 이제 주말에 여유가 생겨서 좋다.

감성분석

힘든 하루

코멘트 생성

오늘부로 분식집 알바를
그만두기로 하셨군요. 주말에
여유가 생겨서 좋으시겠어요.
다른 업종을 찾아보시는 건
어떨까요?

입력 일기

중간고사를 망쳤다. 내가 공부했던 부분에서는 문제가 안 나오고 '설마 이게 나오겠어?'라고 생각하고 넘어간 부분에서만 문제가 나왔다. 내가 스스로 생각해도 말인지 방구인지 모르겠는 답안을 내고 나왔다. 기말고사 땐 더 잘해야지...

감성분석

슬픈 하루

코멘트 생성

중간고사를 망쳐서 기분이 안
좋으시겠어요. 기말고사 땐 더
잘해서 좋은 결과가 있었으면
좋겠네요. 힘내세요!

05. APPENDIX

Q&A

Human evaluation을 하는 데이터셋을 어떻게 선정했나요?

감성분석 모델 기준으로 분류하기 까다로운 텍스트들을 Human evaluation의 대상으로 선정하는 것이 좋겠다고 판단하여 기존 데이터셋 중 **분류를 했을 때 logit값이 낮은 텍스트 100개를 선정하여** 평가 데이터셋으로 만들었다.

코멘트를 Pseudo labeling할 때 GP3 API에 어떠한 입력값을 주었나요?

"이 일기에 코멘트를 달아줘."라는 질의 뒤에 비출판물 데이터셋의 일기 텍스트를 이어붙여서 입력값으로 주었다.

왜 accuracy, F1-score는 11개의 감정 예측을 평가하기에 한계가 있나요?

accuracy와 F1-score는 감정간의 유사도를 파악하지 못한다. 예를 들어 <기쁨> 이라는 감정을 <즐거움>으로 예측하는 것은 큰 문제가 되지 않지만, <기쁨> 이라는 감정을 <짜증남>으로 예측하는 것은 큰 문제가 된다. 따라서 새로운 지표를 만들어 모델을 평가했다.

End of Document Thank You.