

한국 근현대시 감정 라벨링 데이터셋 구축: 문학 텍스트의 컴퓨터 기반 감정 분류와 생성형 AI 활용을 위한 기초 연구

(한국학중앙연구원 한국학대학원 | 임이로 지해인 김병준)

1. 문제의식

- 최근 거대 언어 모델(LLMs)의 등장과 함께, 생성형 AI를 활용한 문학 창작과 그 사회문화적 영향에 대한 논의가 활발 (Porter & Machery, 2024). 인간만의 고유 영역으로 여겨졌던 ‘창작’에 대한 사회적 인식의 변화가 일어나고 있음.
- 인공지능 기술의 국가 편향 현상에서, **한국어 문학**의 고유한 정서와 미학을 체계적으로 **분석하고 데이터 형태로 보존하는 것은 중요한 학술적 과제**.
- 기존의 KOTE(Korean Online That-gul Emotions) 데이터셋은 50,000개의 현대 한국인의 온라인 댓글을 44개의 감정으로 분류한 유용한 데이터, 그러나 정제되지 않은 일상 언어를 기반으로 하여 문학 텍스트 분류에 적용하기에는 한계가 있음.

2. 연구 목적과 의의

본 연구는 디지털 인문학적 방법론을 활용하여 한국 근현대시의 감정 분류를 위한 전문화된 데이터셋을 구축하여, 컴퓨터 기반 문학 연구와 인공지능 시 창작에 기여하고자 함.

구축한 데이터셋과 모델은 전산 문학 연구의 ‘멀리서 읽기’ 방법론을 통해 한국문학 텍스트의 감정 표현을 거시적으로 분석하고, 한국인의 정서를 반영한 생성형 AI 문학 창작 도구 개발에 기여하기를 기대함.

3. 연구 방법

요약: KOTE의 감정 분류 체계를 적용한 근현대시 문학 감정 데이터셋을 구축하여, KcELECTRA 언어 모델에 KOTE 데이터셋을 학습한 모델에 본 연구에서 구축한 데이터셋을 파인튜닝 (fine-tuning)하여 시 텍스트 감정 분류에 더 적합한 모델을 개발.

작품목록 QR



(1) 데이터 구축

- 시인 선정 :** 김소월, 윤동주, 이상, 임화, 한용운
- 휴먼 라벨링 :** 2명의 전문 연구자가 KOTE의 감정 분류 체계를 적용하여 라벨링을 수행

KOTE 감정 분류 ['불평/불만', '환영/호의', '감동/감탄', '지긋지긋', '고마움', '슬픔', '화남/분노', '존경', '기대감', '우쭐댐/무시함', '안타까움/실망', '비장함', '의심/불신', '뿌듯함', '편안/쾌적', '신기함/관심', '아껴주는', '부끄러움', '공포/무서움', '절망', '한심함', '역겨움/징그러움', '짜증', '어이없음', '없음', '패배/자기혐오', '귀찮음', '힘듦/지침', '즐거움/신남', '깨달음', '죄책감', '증오/혐오', '흐뭇함 (귀여움/예쁨)', '당황/난처', '경악', '부담/안_내김', '서러움', '재미 없음', '불쌍함/연민', '놀림', '행복', '불안/걱정', '기쁨', '안심/신뢰']

본문	제목	저자	rater1	rater2
만족(滿足)은 언제든지 인생(人生)과 수직(豎的) 평행(平行)이다	만족	한용운	깨달음, 감동/감탄	깨달음, 없음
남들은 자유(自由)를 사랑한다지마는 나는 복종(服從)을 좋아하여요	복종	한용운	깨달음, 비장함	깨달음, 기대감, 불안/걱정, 뿌듯함

총4,166개

개수	윤동주	김소월	임화	이상	한용운	통계
행	895	958	972	155	816	3796
작품전체	89	86	55	31	109	370
작품수	88	84	19	12	89	292

LLM 감정분류
전체보기 QR



(2) 분석 모델

a. LLM 모델로 감정분류

예시 본문	저자	제목	v1(GPT-o1)	v2(GPT-o1+KOTE train)	v3(GPT-o1+KOTE train)+시라벨링 데이터
내 고장 칠월은 청포도가 익어가는 시절			기대감: 0.78	기대감: 0.82	기대감: 0.82
이 마을 전설이 주절이주절이 열리고 먼데 하늘이 꿈꾸며 알알이 들어와 박혀	이육사	청포도	편안/쾌적: 0.65	행복: 0.75	편안/쾌적: 0.76
하늘 밑 푸른 바다가 가슴을 열고 흰 뜬단배가 곱게 밀려서 오면			즐거움/신남: 0.55	흐뭇함 (귀여움/예쁨): 0.55	흐뭇함 (귀여움/예쁨): 0.72
내가 바라는 손님은 고달픈 몸으로…			행복: 0.52	즐거움/신남: 0.52	행복: 0.65

b. 파인튜닝

- | | Train | Validation | Test |
|--|----------------|------------|---------|
| • 연구 데이터 세팅 : | 데이터개수 2817건(행) | 352건(행) | 353건(행) |
| 시 라벨링 데이터를 랜덤하게 섞어서 80% 학습, 10% 검증, 10% 테스트용으로 분배 + 2명의 라벨러가 일치하는 감정만 평가지표에 사용 | | | |

- 모델 성능 테스트 조건 : 감정 분류 점수 0.3점을 초과하기

	모델 1 (KcELECTRA2023 + KOTE + 시라벨링 데이터)	모델 2 (KcELECTRA2023 + KOTE 데이터셋)	모델 3 (KcELECTRA2021 + KOTE 데이터셋)
성능 평가	- Accuracy(micro) : 0.9136 - F1(macro) : 0.1232 - MCC(flattened) : 0.2972	- Accuracy(micro) : 0.4911 - F1(macro) : 0.0772 - MCC(flattened) : 0.0094	- Accuracy(micro) : 0.7960 - F1(macro) : 0.0501 - MCC(flattened) : 0.0061

예시 본문	저자	제목	모델 1 (KcELECTRA2023 + KOTE + 시라벨링 데이터셋)	모델 2 (KcELECTRA2023 + KOTE 데이터셋)	모델 3 (KcELECTRA2021 + KOTE 데이터셋)
아주 오랜 세월이 흐른 뒤에 힘없는 책갈피는 이 종이를 떨어뜨리리 그때 내 마음은 너무나 많은 공장을 세웠으니 어리석게도 그토록 기록할 것이 많았구나			당황/난처: 0.319 기질투는 형 나의 도 힘	불안/걱정: 0.430 안타까움/실망: 0.393	감동/감탄: 0.466 경악: 0.358 고마움: 0.508 공포/무서움: 0.343 귀찮음: 0.349 기쁨: 0.589 깨달음: 0.434 부끄러움: 0.479 …(외 15개 감정은 생략)

4. 연구 의의

- 본 연구는 기존 사회과학 분야에서 주로 사용되는 연구 방법인 텍스트 감정 분류 모델을 만들어 문학 분야 ‘시’에 적용한 국내 첫 시도 사례.
- 본 연구에서 휴먼 라벨링으로 구축한 문학 감정 데이터셋이 향후 문학 텍스트에 대한 정량적 연구의 정확도와 객관성을 높이는 지표가 되기를 기대함.
- 연구 데이터셋에 활용된 한국 근현대시의 문학적 표현과 감수성을 기계 가독형 데이터로 변환하여 그 가치를 지속하고 보존하는데에 기여할 것.