

## Transformer 와 TF-IDF 를 활용한 문학 작품 감정 분석

한승민, \*김상철  
국민대학교

handsomemin@kookmin.ac.kr, \*sckim7@kookmin.ac.kr

## Sentiment Analysis of Literary Texts Using Transformer and TF-IDF

Seung Min Han, \*Kim Sang-Chul  
Kookmin University

handsomemin@kookmin.ac.kr, \*sckim7@kookmin.ac.kr

## 요 약

문장 단위 감정 분류 모델과 키워드 감정 귀속 분석을 결합하여 문학 작품의 정서 구조를 정량적으로 해석할 수 있는 분석 프레임워크를 제안한다. 감성 사전 기반 분석 방식은 단어 단위 극성값에만 의존하기 때문에 문맥에 따라 변화하는 감정을 정밀하게 파악하기 어렵고 특정 감정이 어떤 키워드와 연결되어 나타나는지 정량적으로 파악할 수 없다. 본 연구에서는 이러한 한계를 보완하고자 문맥 정보를 반영할 수 있는 Transformer Encoder 기반 감정 분류 모델을 통해 문학 작품의 각 문장을 감정 분류하고, TF-IDF 기반으로 추출한 주요 키워드에 감정 분류 모델로 분류한 각 문장의 감정 레이블을 귀속시켜 키워드 별 감정 분포를 벡터로 수치화하였다. 이를 통해 작품 내 키워드가 어떤 정서적 맥락에서 어느 비중으로 사용되는지 정량적으로 파악할 수 있으며, 문학 작품 내 감정을 키워드 단위로 분석할 수 있는 기반을 마련하였다.

## I. 서론

전지적 작가 시점으로 서술된 「운수 좋은 날」은 인물의 정서 변화가 배경과 사건 전개 속에 간접적으로 드러나는 방식으로 구성되어 있고 감정 주체가 명확하지 않아 기존 감성 사전 기반 분석 방식으로는 작품의 정서를 정확하게 추론하기 어렵다. 본 연구는 Transformer 기반 문장 단위 감정 분류 모델을 학습하여 문맥 정보를 반영한 감정 예측을 수행하고, 이후 감정 분류 결과를 TF-IDF 기반으로 추출한 주요 키워드에 귀속시켜 각 키워드의 감정 분포를 수치화하는 것을 통해 문학 작품의 정서 구조를 정량적으로 분석할 수 있는 기반을 마련하였다.

## II. 관련 연구

## 2.1 감성 사전 기반 감정 분석

감성 사전에 기반한 규칙 기반(rule-based) 접근은 감정 분석에서 널리 활용되어 왔다. 실험 초기에 사용했던 KNU 감성사전은 각 단어에 극성 점수를 부여하여 문장의 정서를 예측하는 데에는 유용했지만, 사전에 저장된 단어 수준의 정보에만 의존하기 때문에 문맥을 고려하지 않고 분석하는 분명한 한계가 있었다.[3] 대표적으로, 작품에 등장하는 강아지 ‘개똥이’가 포함된 문장은 맥락상 중립적이지만 사전 상에 ‘개똥’이라는 단어에 부정 극성값이 부여되어 있어 해당 단어가 포함된 모든 문장의 감정이 왜곡되는 사례가 확인되었다.

## 2.2 감정 귀속 방식과 키워드 기반 정서 분석

감정 분류 결과를 특정 단어나 개체에 연결하는 ‘감정 귀속’ 방식은 상품 리뷰의 문장별 감정 분석 결과를 주요 키워드에 귀속시켜 감정을 추적하거나 SNS 텍스트에서

인물명과 장소명을 중심으로 감정 주체를 추정하는 방식으로, 주로 실용문을 중심으로 일부 연구에서 시도된 바 있다.[5] 그러나 문학처럼 감정 표현이 간접적이고 주체가 모호한 장르에 이를 적용한 사례는 드물다. 또한 생성형 모델의 출력에 대한 기여도를 정량적으로 추정하는 최근 연구는 감정과 키워드 간 관계를 벡터 형태로 수치화할 수 있다는 점에서 본 연구의 설계에 영감이 되었다.[2] 본 연구는 이러한 선행 접근을 문학 작품에 적용함으로써 감정과 키워드 간의 정량적 연계를 시도하였다.

## III. 데이터 및 모델 구성

## 3.1 데이터 구성 및 전처리

웹 기반 문학 아카이브에서 수집한 작품 원문은 인위적인 줄바꿈이 삽입된 형태였다. 이에 마침표('.')를 기준으로 문장 경계를 재구성하여 감정 분류 모델이 문장 단위 감정 분류를 수행할 수 있도록 입력 형식을 맞춰주었다. 이러한 수정을 거쳐 분석 작품은 총 200 문장으로 구성되었다.

## 3.2 감정 분류 모델 설계 배경 및 모델 학습

감정 분류를 수행하는 기존의 사전 학습된 Transformer 기반 모델들이 학습했던 감정 범주는 본 연구에서 사용할 데이터셋의 범주와 일치하지 않았다. 이로 인해 사전학습 모델을 직접 파인 튜닝하는 것은 어렵다고 판단하여 Transformer Encoder 기반의 모델을 직접 구성하여 초기 가중치에서 학습을 시작하였다.

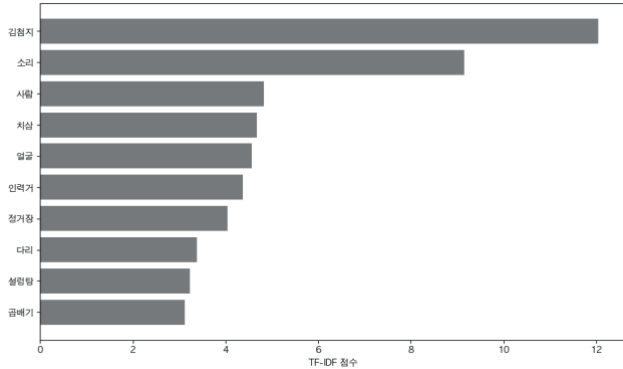
학습에는 약 41 만 문장으로 구성된 emotion\_korTran 데이터셋을 사용하였으며, 손실 함수는 CrossEntropyLoss, 옵티마이저로는 Adam 을 적용하였다. 학습 결과 약 86.3%의 분류 정확도를 달성하였으며 이 모델을 활용해 작품의 문장 별 감정 예측을 수행하였다.

## IV. 실험 결과 및 분석

## 4.1 주요 키워드 및 문장 단위 감정 분류 결과

## 4.1.1 키워드 추출 결과(TF-IDF 기반)

형태소 분석(KLT2023)을 통해 각 문장에서 명사와 고유 명사를 추출한 후 TF-IDF 점수를 기준으로 상위 10 개 키워드를 선정하였다. 추출된 키워드와 TF-IDF 점수는 [그림 1]에 제시되어 있다. 선정된 키워드들은 감정 분류 결과와 연결하여 감정 귀속 분석에 활용되었다.



[그림 1] TF-IDF 기반 상위 키워드 (불용어 제거)

## 4.1.2 문장 단위 감정 분류 결과

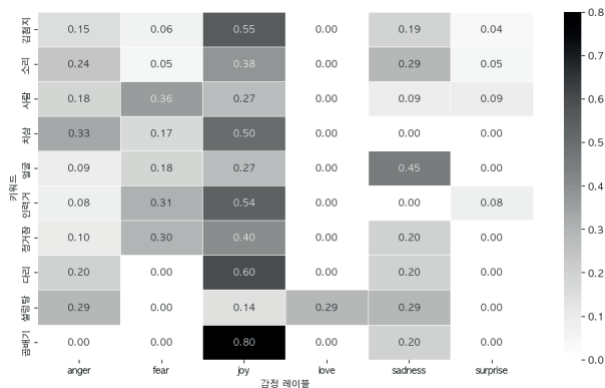
학습된 모델을 통해 「운수 좋은 날」의 총 200 개 문장을 6 가지 감정 범주(joy, sadness, anger, fear, surprise, love) 중 하나로 분류하였다.

[표 1] 감정 분류 결과 : 감정 레이블별 문장 분포

| Emotion Label | Number of Sentences |
|---------------|---------------------|
| Joy           | 96                  |
| Sadness       | 39                  |
| Anger         | 29                  |
| Fear          | 23                  |
| Surprise      | 8                   |
| Love          | 4                   |

## 4.2 감정 귀속 및 정량 분석

문장 단위 감정 분류 결과를 기반으로 각 문장에 포함된 상위 키워드에 해당 문장의 감정 레이블을 귀속하였다. 감정 귀속 대상 키워드는 4.1.1 절에서 추출한 TF-IDF 상위 10 개 단어이며 각 키워드별 감정 분포는 해당 키워드가 포함된 문장에서 나타난 감정 비율로 구성되었다. [그림 2]에서 각 키워드가 6 차원 감정 벡터(anger, fear, joy, love, sadness, surprise)로 정량화된 것을 보여준다.



[그림 2] 문장 별 감정 분류 결과를 TF-IDF 기반으로 추출한 키워드에 귀속한 후 분포를 시각화한 결과

[그림 2]에 나타난 감정 분포를 보면, 예컨대 ‘김첨지’는 joy 감정이 55.3%로 가장 높고, 그 외에 sadness(19.2%)와 anger(14.8%)의 비중도 일정하게 나타난다. 반면 ‘설렁탕’은 joy(14.2%)보다 sadness(28.6%)와 anger(28.6%)의 비율이 더 높게 분포한다. 이와 같이 하나의 키워드에 여러 감정이 나타나는 경향을 통해 작품 내에서 해당 단어가 사용된 문맥의 정서적 복잡성을 반영한다는 것을 알 수 있다. 또한 감정 벡터로 정량화된 이 비율은 각 키워드가 서사 구조 속에서 어떠한 감정과 얼마나 연결되는지를 수치적으로 파악할 수 있다.

## V. 결론 및 향후 연구

본 연구는 문장 단위 감정 분류 모델과 키워드 감정 귀속 분석을 결합하여 문학 작품의 정서 구조를 정량적으로 파악하였다. 감정과 키워드 간 관계를 수치 기반으로 분석하는 이 방식은 감정 분류 결과를 단순히 정성적으로 해석하는 수준을 넘어 문장과 키워드 수준의 감정 정보를 수치화하고 구조적으로 해석할 수 있다.

고유 명사 중심의 키워드 추출과 시계열적으로 감정 분석을 표현하는 것을 결합하면 핵심 등장인물의 감정 변화를 시간에 따라 추적할 수 있으며, 작품의 정서적 흐름을 파악할 수 있을 것이다. 이러한 구조를 활용하여 인기 서사의 감정 흐름 패턴을 유형화하고, 사용자의 정서 선호와의 매칭을 통해 감정 성향에 기반한 맞춤형 문학 작품 추천 시스템으로도 확장 가능할 것이다.

## ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음. (2022-0-00964)

This research was supported by the Ministry of Science and ICT (MSIT) and the Institute for Information & Communications Technology Planning & Evaluation (IITP) under the National Program of Excellence in Software (2022-0-00964).

## 참 고 문 헌

- [1] S. A. Hamad, M. S. Ahmad, "Emotional Text Classification Using TF-IDF and LSTM," Journal of Intelligent & Fuzzy Systems, vol. 39, no. 4, pp. 5403–5413, 2020.
- [2] Li L., Lee K., Park J. S., Devlin J., Toutanova K., "Decentralized Attribution of Generative Models," Transactions of the Association for Computational Linguistics, vol. 11, pp. 837–854, 2023.
- [3] 박상민, 엄창민, 온병원, 정동원, "Bi-LSTM 기반의 한국어 감성사전 구축 방안," 지능정보연구, 제 24 권, 제 4 호, pp. 87–104, 2018.
- [4] 남규현, 이현영, 강승식, "KoNLTK: 한국어 언어 처리 도구," 한글 및 한국어 정보처리 학술대회 논문집, vol. 2018, pp. 611–613, Oct. 2018.
- [5] A. R. Joshi, K. Bhattacharyya, "Transformer-based Deep Learning Models for the Sentiment Analysis of Social Media Data," Procedia Computer Science, vol. 192, pp. 2624–2633, 2021.