KE-T5 기반 한국어 대화 문장 감정 분류

임영범, 김산, 장진예, 신사임, 정민영O

한국전자기술연구원, 인공지능연구센터

{warf34, kimsan0622, jinyea.jang, sishin, minyoung.jung}@keti.re.kr

KE-T5-Based Text Emotion Classification in Korean Conversations

Yeongbeom Lim, San Kim, Jin Yea Jang, Saim Shin, Minyoung JungO

Korea Electronics Technology Institute, Artificial Intelligence Research Center

요 약

감정 분류는 사람의 사고방식이나 행동양식을 구분하기 위한 중요한 열쇠로, 지난 수십 년간 감정 분석과 관련된 다양한 연구가 진행되었다. 감정 분류의 품질과 정확도를 높이기 위한 방법 중 하나로 단일 레이블링 대신 다중 레이블링된 데이터 세트를 감정 분석에 활용하는 연구가 제안되었고, 본 논문에서는 T5 모델을 한국어와 영어 코퍼스로 학습한 KE-T5 모델을 기반으로 한국어 발화 데이터를 단일 레이블링한 경우와 다중 레이블링한 경우의 감정 분류 성능을 비교한 결과 다중 레이블 데이터 세트가 단일 레이블 데이터 세트보다 23.3% 더 높은 정확도를 보임을 확인했다.

주제어: 텍스트 감정 분류, 감정 다중 레이블링, 사전 학습 모델

1. 서론

감정은 어떤 사건이나 일에 대한 느낌이나 마음으로, 지난 수십 년간 감정 분석과 관련된 다양한 연구가 수행되었다. 감정 분류는 사람의 사고방식과 행동양식을 구분하기 위한 중요한 요소로써 비즈니스, 사회 복지, 정치 등 다양한 분야에 활용되고 있어, 감정 분류의 품질과 정확도를 높이기 위해 감정 종류를 구분하는 방법에 대한 많은 연구가 수행되어왔다. 복합적인 특징을 가지는 감정은 한 발화 내에서도 다양한 감정을 포함할 수 있기 때문에, 다중 레이블 데이터 세트가 단일 레이블 데이터 세트보다 더 효과적이라는 주장이 제기되었다[1]. 따라서 본 논문에서는 T5 모델[2]을 한국어와 영어 코퍼스로 학습한 KE-T5 모델[3]을 기반으로, AI Hub 멀티모달 영상 데이터 세트[4]에서 추출한 한국어 발화 데이터 세트(원본 데이터 세트)를 단일 레이블링(단일 레이블 데이터 세트), 다중 레이블링(다중 레이블 데이터 세트)한 후, 두 데이터 세트의 감정 분류 성능을 비교한다.

2. 관련 연구

감정을 어떻게 분류할 것인지에 대해 다양한 연구가 행해졌고, 대표적인 감정 분류 방법이 몇 가지 제안되었다. Ekman는 인간의 얼굴 표정과 그 감정에 대한 연구를 통해 분석된 9가지 특징으로 인간의 보편적인 감정을 6가지(분노, 공포, 슬픔, 기쁨, 혐오, 놀람)로 분류했으며[5], 이후 확장 연구를 통해 인간의 감정을 7가지 감정(분노, 공포, 슬픔, 행복, 혐오, 놀람, 경멸)으로 분류할 수 있다고 주장했다[6]. 그 외 Harmon-Jones 등은 이산 설문 조사를 통해 8가지 감정(분노, 혐오, 공포, 불안, 슬픔, 행복, 편안함, 욕망)을 포착했으며[7], Cowen 과 Keltner는 감정을 27가지로 분류하였다[8].

이처럼 감정을 분류하는 명확한 기준이 없기 때문에, 하나의 인스턴스를 하나의 감정으로 단일 레이블링하는 대신 여러 감정으로 다중 레이블링하는 방법이 제안되었다. Ye 등은 최근 연구에서 단일 레이블 기반 감정 분류 정확도를 향상시키기 더욱 어려워지고 있고, 하나의 텍스트가 단일 레이블링되면 여러 감정 카테고리 중 하나의 카테고리에만 속하는 것으로 간주되지만 실제로는 다양한 감정을 포함하고 있다고 주장하면서 뉴스 기사에서 독자의 감정을 예측하기 위해 다중 레이블 감정 분류 기법을 제안하였다[1]. Jabreel과 Moreno도 대부분의 감정 분석 연구가 단일 레이블 기반 감정 분류에 초점이 맞춰져 있어 다른 여러 감정의 공존을 무시한다고 지적하면서, 딥 러닝 기반 접근 방식을 통해 트윗의 다중 레이블 기반 감정 분류를 제안하였다[9].

3. 데이터 세트

AI Hub에서 제공하는 멀티모달 영상 데이터 세트[4]를 기반으로 한 발화 텍스트 데이터를 사용한다. 멀티모달 영상 데이터 세트는 110시간 분량의 총 6천개 영상 클립으로 구성되어 있으며, 집, 강의실, 카페, 사무실 등에서 발생하는 상황을 주제로 시나리오를 구성하여 총 300명의 연기자가 2인 1조를 이루어 촬영 제작한 85,077개의 발화 정보로 이루어진 데이터 세트이다. 본 논문에서는 이 중에서 발화 텍스트와 이에 상응하는 감정으로 이루어진 발화 텍스트 데이터(원본 데이터 세트)를 추출하였고, 해당 데이터 세트를 전문가가 텍스트 기반의 감정으로 단일 레이블링(단일 레이블 데이터 세트), 다중 레이블링(다중 레이블 데이터 세트)하여 실험에 사용한다. 사용된 모든 발화 데이터는 총 8가지 감정(중립, 놀람, 혐오, 분노, 공포, 행복, 슬픔, 경멸)으로 분류한 Ekman의 기본 감정 분류[6]를 기반으로 레이블링되었다.

본 실험에서는 우선 원본 데이터 세트를 전문가가 텍스트 기반 단일 감정으로 다시 레이블링한다(단일 레이블 데이터 세트). 표 1은 원본 발화 데이터 세트와 재태깅된 단일 레이블 데이터 세트의 각 감정 별 비율을 비교해서 보여준다.

표 1. 데이터 세트의 감정 별 비율 비교(%)

	중립	놀람	혐오	분노	공포	기쁨	슬픔	경멸
원본	38.0	7.3	8.9	5.2	14.3	12.4	11.5	2.4
단일	60.8	6.8	6.1	8.6	2.8	7.8	6.7	0.3

이후 재태깅된 단일 레이블 데이터 세트를 9명의 전문가가 다중 레이블이 가능하도록 최대 2개의 감정 카테고리로 다시 레이블링해 다중 레이블 데이터 세트를 만들어, 단일 레이블 데이터 세트와 비교 실험을 진행한다. 다중 레이블 데이터 세트를 만드는 과정에서 재태깅한 발화 텍스트 수는 10,652개이며, 이 중 다중 레이블링된 발화 텍스트 수는 3,293개이다. 표 2는 다중 레이블 데이터 세트의 한 샘플을 보여준다.

표 2. 다중 레이블링된 발화 텍스트 예시

발화	감정
그만해. 숨막힌다고.	화남, 혐오

4. 실험 및 결과

본 논문에서는 KE-T5 모델을 기반으로 각 발화 데이터를 학습해 감정을 분류할 수 있는 모델을 만들어 실험 및 평가를 진행하였다.

4.1 모델 세팅

한국전자기술연구원에서 Google의 T5 모델을 한국어와 영어 코퍼스로 학습하여 배포한 사전 학습 모델 KE-T5를 사용한다[3]. 인코더와 분류기로 구성된 감정 분류 모델은 학습 배치 사이즈(training batch size) 16, 평가 배치 사이즈(evaluation batch size) 8, 드랍 아웃 비율(dropout rate) 0.1, 학습률(learning rate) 0.001로 설정하였다. 문맥을 고려한 감정 분류를 위해 이전 대화 문맥을 20문장까지 보도록 하였다.

4.2 실험 결과

본 실험에서는 80,169개의 데이터 세트를 약 10:1로 나누어 72,911개를 학습 데이터로, 나머지 7,258개를 테스트 데이터로 사용한다. 표 3은 단일 레이블 데이터 세트와 다중 레이블 데이터 세트를 활용한 감정 분류 비교 실험 결과로, 단일 레이블 데이터 세트가 69.4%, 다중 레이블 데이터 세트가 92.7%의 정확도를 보였다. 이를 통해 다중 레이블 데이터 세트가 단일 레이블 데이터 세트보다 23.3% 더 높은 정확도를 보이는 것을 확인하였다.

표 3. 정확도 비교 실험 결과

	단일 레이블	다중 레이블
정확도	69.4%	92.7%

5. 결론

본 논문에서는 T5 기반으로 한국어와 영어 코퍼스를 학습한 사전 학습 모델인 KE-T5 모델을 통해 단일 레이블 한국어 발화 데이터 세트와 다중 레이블 한국어 발화 데이터 세트의 감정 분류 정확도를 비교하였다. 그 결과 다중 레이블링된 데이터 세트가 더 높은 정확도를 보임으로써 한국어 텍스트 감정 분류를 위해 고품질 다중 레이블 데이터 세트가 필요함을 증명하였다.

감사의 글

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2021-0-00354, 비정형 텍스트를 학습하여 쟁점별 사실과 논리적 근거 추론이 가능한 인공지능 원천기술)

참고문헌

[1] Lu Ye, Rui-Feng Xu, and Jun Xu. Emotion prediction of news articles from reader's perspective based on multi-label classification. In: 2012 international conference on machine learning and cybernetics. IEEE, p. 2019-2024. (2012).

[2] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu. Exploring the limits of transfer learning with a unified text-to-text transformer. arXiv preprint arXiv:1910.10683. (2019).

[3] KE-T5: Korean-English T5, https://github.com/AIRC-KETI/ke-t5

[4] 멀티모달 영상 데이터 세트, https://aihub.or.kr/aidata/137

[5] Ekman, P. An argument for basic emotions. Cognition & emotion, 6(3-4), 169-200. (1992).

[6] Ekman, P., & Cordaro, D. What is meant by calling emotions basic. Emotion review, 3(4), 364-370. (2011).

[7] Harmon-Jones, C., Bastian, B., & Harmon-Jones, E. The discrete emotions questionnaire: A new tool for measuring state self-reported emotions. PloS one, 11(8), e0159915. (2016).

[8] Cowen, A. S., & Keltner, D. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. Proceedings of the National Academy of Sciences, 114(38), E7900-E7909. (2017).

[9] Jabreel, M., & Moreno, A. A deep learning-based approach for multi-label emotion classification in tweets. Applied Sciences, 9(6), 1123. (2019).