한국경영과학회지
제51권 제1호
2026년 2월
https://doi.org/10.7737/JKORMS.2026.51.1.017

효율적 문맥 반영을 위한 워드 임베딩 기반 감성 분석 연구

나 인 · 배장원†
한국기술교육대학교 산업경영학부

A Study on Word Embedding-Based Sentiment Analysis for Efficient Contextual Representation

In Na · Jang Won Bae
Korea University of Technology and Education
School of Industrial Management

■ Abstract ■

With the rapid growth of text generated online, sentiment analysis technology—which automatically classifies users' opinions—has gained increasing importance. This study proposes a sentiment analysis method that effectively reflects the contextual information of sentences without requiring large-scale training data. The proposed model extracts context-aware vector representations through word embeddings and determines sentiment by calculating semantic similarity between review sentences and words in a sentiment lexicon. In addition, to identify the embedding model that best captures Korean contextual semantics, two BERT-based models and two SBERT-based models were compared. Using product review data provided by AI Hub, both sentiment-expressive sentence data and a Korean sentiment lexicon were employed. Cosine similarity between each sentence and the sentiment lexicon words was computed, and four selection criteria were established to identify the most semantically relevant sentiment words. As a result, the KoSBERT model achieved the best performance, with the highest accuracy (0.8919) and F1-score (0.9014) when averaging the sentiment word vectors whose cosine similarity with the sentence was 0.5 or higher.

Keywords : Sentiment Analysis, Text analysis, Word Embedding, BERT

논문접수일 : 2025년 10월 23일 논문게재확정일 : 2025년 12월 04일
논문수정일 : 2025년 11월 18일
† 교신저자, jangwon_bae@koreatech.ac.kr

1. 서 론

감성 분석(Sentiment Analysis)은 텍스트에 드러나는 감성을 긍정, 부정으로 파악하여 자동으로 분류하는 기술을 의미한다. 최근 온라인에서 생성되는 텍스트가 증가함에 따라 텍스트를 자동으로 분류하고 요약하는 기술의 중요성이 주목받고 있다. 기존에는 문서의 주제를 분류하는 것이 중요했지만, 현재는 리뷰와 같이 사용자 생성 콘텐츠에서 작성자의 감성이 핵심 정보로 작용한다. 제품이나 서비스의 리뷰의 감정을 분석해서 긍정인지 부정인지 자동으로 분류하는 기술은 추천 시스템, 여론 분석 등 다양한 텍스트 분야 분석에서 활용할 수 있다[13].

그러나 기존 감성 분석 연구의 상당수는 단순한 키워드 기반 통계 분석이나[3-6, 11], 대규모 학습 데이터에 의존한 모델 학습[2, 7, 8, 9]에 국한되어 있다. 키워드 기반 통계적 접근은 구조가 단순하고 구현이 쉬우며 해석 가능성이 높다는 장점이 있으나, 문장에서 단어의 존재 여부나 빈도를 중심으로 감성을 판단하기 때문에 문맥 정보를 충분히 반영하지 못한다. 이에 따라 동일한 단어라도 상황에 따라 의미가 달라지는 경우 정확도가 크게 저하될 수 있으며, 감성 사전에 등재되지 않은 단어가 등장하면 분석 자체가 불가능하다는 한계를 가진다.

반면 지도 학습 기반 모델은 사전 학습된 언어모델을 활용하여 문맥적 의미를 정교하게 반영할 수 있고, 충분한 규모의 라벨링 된 데이터가 확보될 경우 높은 성능을 보장한다는 장점을 지닌다. 그러나 이러한 방식은 대량의 라벨링 데이터 구축이 필수적이며, 데이터 수집 및 모델 학습 과정에 상당한 자원과 비용이 소요된다는 한계를 가진다.

본 연구는 두 접근법의 장점을 결합하면서 단점을 보완하기 위해 워드 임베딩을 활용한 유사도 기반 감성 분석 방법을 제안한다. 제안한 방법은 감성 사전 기반 접근의 단순성과 가벼운 구조를 유지하면서도, 사전학습된 언어모델을 통해 문장과 감성 단어의 임베딩을 생성하여 문맥 정보를 효과적으로 반영할 수 있다.

이를 위해 AI Hub에서 제공하는 제품 리뷰 데이터인 '속성 기반 감정 분석 데이터'와 국립군산대학교에서 구축한 감성 사전 데이터를 활용하였다. AI 허브(AI Hub)는 과학기술정보통신부와 한국지능정보사회진흥원(NIA)가 구축한 국가 인공지능 데이터 공유 플랫폼으로, 이미지, 음성, 텍스트, 영상 등 다양한 분야의 대규모 공개 데이터셋을 제공한다[12].

각 리뷰 문장과 감성 사전의 단어들은 BERT 및 SBERT 계열의 사전학습 언어모델을 통해 768차원의 임베딩 벡터로 변환되며, 이를 바탕으로 코사인 유사도를 계산하여 감성 분석을 수행한다.

제안한 방법의 주요 특징은 다음과 같다. 첫째, 워드 임베딩을 활용하여 문맥적 의미를 반영함으로써 문장 내 감성 표현을 효과적으로 파악할 수 있다. 또한 벡터 값은 수치 형태이기 때문에 코사인 유사도 계산이 가능하며, 이를 통해 문장과 감성 단어 간의 의미적 관계를 -1부터 1 사이의 정량적 지표로 산출할 수 있다.

둘째, 대규모 라벨링 데이터 구축이 필요 하지 않으며, 모델 학습 과정이 필요하지 않은 간단한 방법이다. 이를 통해 복잡한 학습 절차 없이 효율적인 감성 분류가 가능하다. 이는 레이블 데이터 구축 및 모델 학습에 많은 자원이 소요되는 기존 접근법의 한계를 보완할 수 있다.

이와 같이 본 연구에서 제안하는 방법론은 단순한 구조를 기반으로 문맥을 효과적으로 반영한 감성 분석을 수행할 수 있어, 다양한 한국어 텍스트 분석 분야에 활용될 수 있을 것으로 기대된다.

또한 제안한 감성 분석 방법의 효과성을 검증하기 위해, 제안 방법의 분류 성능을 정량적으로 평가(4장)하고, 연구 문제의 타당성을 확인하였다. 이어서 5장에서는 대표적인 기존의 감성 분석 모델들을 벤치마크로 성능을 비교(5장)하여 제안한 방법의 실효성을 검증하였다.

2. 선행 연구 고찰

한국어 텍스트 감성 분석 방법은 크게 감성 사전

을 활용한 통계 기반 분석 방법, Transformer 기반 언어 모델 활용 방법, 그리고 이 두 가지를 결합한 복합적 방법으로 구분할 수 있다.

2.1 통계 기반 분석 방법

통계 기반 분석 방법은 토픽 모델링, TF-IDF와 같이 문장 키워드의 빈도를 분석하고, 드러난 토픽으로 감성을 분석하는 방법이다. 감성 사전을 함께 활용하는 방법도 있는데, 키워드를 추출한 후 감성 사전으로 감성을 분석하는 방법이다. 예를 들어 '감사합니다.'라는 문장의 키워드가 '감사'이고 감성 사전에 '감사'라는 단어가 긍정으로 제시되어 있다면 해당 문장은 긍정으로 분석하는 방법이다.

지경규[11]는 강의평가에서 수집된 서술형 응답 데이터에서 명사 위주의 키워드를 추출하고 TF-IDF 빈도 분석과 LDA 분석을 통해 해당 강의평의 토픽을 추출하여 간접적으로 감정을 분류하였다.

반건우 등[3]은 제인 에어 작품 영문 텍스트를 수집하여 LDA를 통해 각 장 별로 등장하는 주요 주제 분포 파악하였다. 이후 각 주제에 TF-IDF 분석을 통해 주요 단어를 추출하였고, 감성 사전을 이용하여 각 문장에 포함된 감성 단어와 그에 따른 감성 점수를 기반으로 긍정, 부정을 분류하였다.

서혜선[4]은 '대전'과 '환경'을 중심 키워드로 설정하여 SNS 텍스트 데이터(블로그, 카페, 웹문서 등)를 수집하고, 이를 바탕으로 토픽 모델링을 수행하였다. 이후 KNU 한국어 감성 사전을 활용하여 각 토픽에 포함된 단어를 사전과 대조하여 긍정, 부정, 중립 단어의 빈도를 정량적으로 산출하여 감성 분석을 수행하였다.

오정심[5]은 신안군, 진도군, 완도군의 문화 경관과 관련된 총 8,847건의 경험담을 수집하여 감성 분석을 수행하였다. 감성 사전을 활용해 경험담에 나타난 긍정, 부정 단어의 비율을 계산함으로써 텍스트의 전반적인 감성 경향을 파악하였다.

이대국 등[6]은 문학 작품의 대본 데이터를 대상으로 주요 키워드를 추출한 뒤, 해당 단어가 EmoLex 감성 어휘 사전에 존재할 경우 감성 점수를 부여하는 방식으로 감정을 분류하였다.

2.2 Transformer 기반 언어 모델 활용 방법

이 방법은 Transformer 기반 거대 언어 모델의 지도학습을 통해 감성 분석 모델을 구축하는 방법을 의미한다. BERT나 GPT 계열의 모델에 감정 라벨이 있는 데이터를 사용하여 미세 조정을 통한 감성 분석 모델을 구축한다.

박현정, 신경식[2]은 전자제품 후기 데이터를 기반으로 '배터리' 등 21개의 속성 카테고리를 정의하고, 해당 속성에 대한 감정을 수작업으로 라벨링하였다. 이후 BERT 모델을 활용하여 각 문장의 CLS 토큰을 입력으로 받아 문장의 감정(긍정/부정)을 분류하는 실험을 진행하였다.

이민아 등[7]은 네이버 영화 리뷰 데이터를 수집한 후, 긍정/부정 이진 감정 라벨을 부여하고 KoBERT, KoGPT-2, KoBART 모델을 각각 학습시켜 성능을 비교하였다. 하이퍼파라미터 튜닝을 통해 모델별 성능을 최적화하였다.

전수현, 노미진[9]은 AI hub에서 제공하는 여성의류, 남성의류, 패션슈즈, 잡화 카테고리로 분류된 온라인 패션 리뷰 데이터를 활용하여 감성 분석을 수행하였다. 각 데이터는 속성별 리뷰와 그에 따른 감정 점수가 부여되어 있고, KcELECTRA 모델을 통한 감성 분석 예측을 수행하였다.

이현수 등[8]은 AIhub에서 2021년에 구축한 희곡 데이터를 활용하여 GPT-3 API 기반 감성 분석을 수행하였다. 분석에서는 기쁨과 흥미의 긍정 감성, 슬픔/분노/혐오/공포/놀람/통증/지루함의 부정 감성의 총 9개 감성 범주를 분류하였다.

2.3 복합적 방법

복합적 방법은 통계 기반 키워드 분석과 Transformer 기반 언어 모델을 복합적으로 활용하여 감성을 분석하는 방법이다.

박준성 등[1]은 러닝(running)앱 후기 데이터를 대상으로 LDA를 통해 4가지의 주요 키워드를 추출한 후 AI Hub 감성 데이터를 활용하여 미세조정된 KoBERT 분류 모델을 사용해 리뷰의 감정 분석을 수행하였다. 이후 토픽별 긍정, 중립, 부정 비율에 가중치를 부여하는 방식으로 감성 점수를 산출하였다.

정환웅, 서지훈[10]은 구글 플레이스토어의 게임 리뷰를 대상으로, 명사 중심의 TF-IDF 분석을 통해 상위 5개의 키워드를 추출한 뒤, 키워드에 해당하는 리뷰를 필터링했다. 이후 해당 키워드가 포함된 리뷰를 CLOVA-X 모델로 분석하여 감성 비율을 산출하였다. 각 키워드별로 부정 및 긍정 리뷰 비율을 비교하여, 더 높은 비율을 차지한 감성을 해당 키워드의 대표 감정으로 분류하는 방식을 적용하였다.

2.4 기존 연구 방법과의 비교

위와 같은 방법들은 각 사례에서 우수한 성능을 보여주고 있으나 일반적으로는 다음과 같은 한계점을 지닌다.

통계 기반 키워드 분석 방법은 단순 키워드 추출이므로 텍스트의 문맥을 반영하지 못한다는 한계가 있다. 해당 방법의 감성 분석은 문장의 문맥이 아닌 키워드의 빈도에 의존하기 때문에, 문맥 정보 반영에 한계가 있다. 특히, 명사는 동일하지만 조사 또는 문장 구조에 따라 의미가 달라지는 경우를 충분히 구분하지 못한다는 약점을 지닌다. 또한 문장에 감성 사전을 활용한 방법에서는, 감성 사전에 해당하는 단어가 없을 때는 분석이 불가하다는 한계점이 있다.

Transformer 언어 모델에 기반한 방법은 대규모 언어 모델을 활용함으로써 문맥 정보를 효과적으로 반영할 수 있다는 장점이 있다. 그러나 이들 모델은 지도학습 기반이기 때문에, 대량의 라벨링된 학습 데이터가 필요하다는 것과 미세 조정 시 자원과 시간이 소모된다는 한계를 가진다.

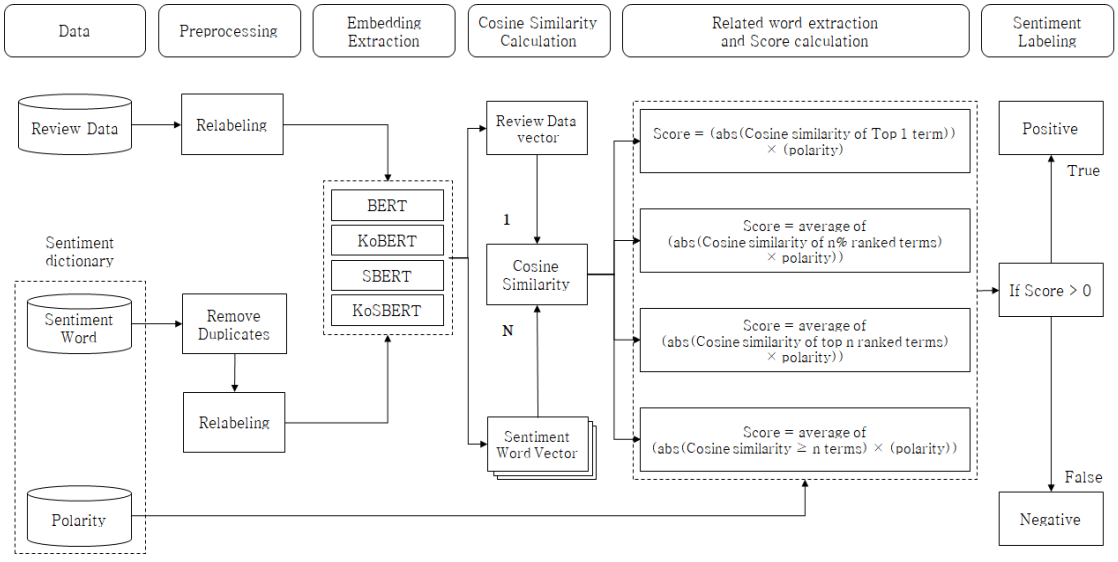
본 연구에서는 두 접근법의 장점을 결합하면서 단점을 보완하기 위해 워드 임베딩을 활용한 유사도 기반 감성 분석 방법을 제안한다. 해당 방법은 워드 임베딩과 코사인 유사도 분석을 통해 간단하고 리뷰의 문맥을 효율적으로 반영할 수 있는 방법론이다. 워드 임베딩은 문장의 의미를 벡터 공간에 매핑시키는 방식으로, 코사인 유사도를 계산을 통해 감성 사전의 단어와 제품 리뷰 간의 의미적 유사도를 측정할 수 있다.

제안된 방법론은 다음과 같은 특징을 가진다. 첫째, 워드 임베딩 방식으로 문장 전체의 의미를 대표하는 임베딩 값을 추출하여 분석에 사용하기 때문에 문맥을 반영한 감성 분석을 수행한다. 이를 통해 기존 연구가 감성 단어의 단순 빈도 통계에 의존함으로써 문맥을 반영하지 못했던 한계를 보완할 수 있다. 또한 감성 사전에 포함되지 않은 단어가 문장에 사용된 경우에도 감성 분석이 가능하다. 본 방법론에서는 임베딩 추출 시 한국어 문맥을 가장 잘 반영하는 모델을 판별하기 위해 다양한 워드 임베딩 모델을 비교 분석하여 감성 분석에 가장 효과적인 모델을 제시한다.

둘째, 라벨링된 데이터가 없는 경우에 활용할 수 있는 감성 분석 방법이다. 본 연구에서 제안한 방법은 문장의 임베딩 값과 감성 사전 단어의 임베딩 값의 유사도를 기준으로 감성 분석을 활용하기 때문에 라벨링 데이터 없이 감성 분석이 가능하다. 또한 데이터를 활용한 별도의 학습 과정이 필요하지 않기 때문에 Transformer 기반 언어 모델 활용 방법에 비하여 비교적 간단하다.

3. 연구 방법

본 연구에서는 제품 리뷰에 대해 문맥을 효과적으로 반영하는 간단한 감성 분석 기법을 고안하였다 (<그림 1> 참고). 전체 연구 과정은 총 7단계로 구성된다. 먼저 제품 리뷰 데이터와 감성 사전 데이터를 수집하고, 결측치 삭제 등 전처리 과정을 거친다. 이후, 전처리된 데이터에 대해 사전학습된 대형 언어 모델인 BERT 계열 모델을 활용하여 문장 임베딩 벡터를 생성한다. 생성된 임베딩을 바탕으로, 각 리뷰와 감성 사전 단어 간의 코사인 유사도를 계산하여 의미적 유사성을 측정한다. 마지막으로, 해당 코사인



〈그림 1〉 연구 구조도

유사도를 가중치로 활용하여 감성 사전의 감성 점수와 곱한 값을 기반으로 최종 감성 라벨을 부여한다. 각 단계의 자세한 과정은 다음과 같다.

3.1 활용 데이터

본 연구에서는 AI hub의 속성 기반 감정 분석 데이터와 감성 사전 데이터를 활용하였다. AI Hub에서 제공하는 '속성 기반 감정 분석 데이터'는 제품 리뷰와 해당 리뷰에 나타난 사용자 감정을 태깅한 데이터셋이다. 사용한 칼럼은 IT 기기 제품 리뷰 칼럼이며, 감정은 점수(score) 형태로 제공된다. 감정 점수 체계는 0점부터 5점까지의 척도와 0점부터 10점 기준의 100점까지의 척도가 혼재되어 있으며 100점 척도의 경우 10점 단위로 구성되어 있다. 이에 따라 본 연구에서는 5점 척도 기준 긍정 평가점수인 4점과 5점, 100점 척도 기준 긍정 평가점수인 70점, 80점, 90점, 100점 해당하는 리뷰를 '긍정'으로 재라벨링하였다. 또한 5점 척도 기준 부정 평가점수인 0점, 1점, 2점과 100점 척도 기준 10점, 20점, 30점, 40점에 해당하는 리뷰를 추출하여 '부정'으로 재라벨링하였다. 이에 따라 가전 제품에 대한 긍정 리뷰 35,169건과 부정 리뷰 2,651건을 연구에 활용하였다(<표 1> 참고). 라벨링 칼럼은 연구 과정에서 사용되지 않지만, 이후 방법론의 성능 검증을 위한 용도로 사용되었다.

〈표 1〉 IT 기기 리뷰 데이터

데이터 크기	전체 리뷰 수	37,820건
레이블 분포	긍정	35,169건
	부정	2,651건
리뷰 길이 통계	최소 값	19자
	최대 값	995자
	평균	106자

감성 사전 데이터는 국립군산대학교에서 제공하는 한국어 감성 사전을 활용하였다(<표 2> 참고) [14]. 이 사전은 표준국어대사전의 뜻풀이를 기반으로 구축되었으며, 일반 단어 외에도 어구, 축약어, 이모티콘 등을 포함한 다양한 형태의 긍정, 부정, 중립 표현을 포함한다. 각 항목에는 감성의 강도를 나타내는 5단계 감성 점수(polarity)가 부여되어 있으며, 매우 부정(-2), 부정(-1), 중립(0), 긍정(1), 매우 긍정(2)으로 구분된다. 전체 14,843개의 단어가 수록되어 있으며, 감성 분포는 매우 부정 4,796개, 부정 5,028

개, 중립 154개, 긍정 2,266개, 매우 긍정 2,597개로 구성되어 있다. 데이터 셋은 단어 원형을 나타내는 word_root, 실제 표현 단어인 word, 감성 점수를 나타내는 polarity 세 가지 칼럼으로 구성되어 있다.

〈표 2〉 감성 사전 데이터

polarity	원본 데이터 크기	전처리 데이터 크기
2 (매우 긍정)	2,597	-
1 (긍정)	2,266	4,863
0 (중립)	154	-
-1 (부정)	5,028	9,826
-2 (매우 부정)	4,796	-
전체	14,843	14,689

감성 사전은 word_root보다 중복되는 단어가 적은 word 칼럼 데이터를 활용하였고, 중복단어를 제거한 후 총 14,689개의 고유 감성 단어를 사용하였다. 해당 사전은 감성 강도에 따라 매우 부정(-2), 부정(-1), 중립(0), 긍정(1), 매우 긍정(2)으로 분포되어 있었으나, 본 연구에서는 감성의 강도보다 감정의 긍정, 부정 이진 분류가 중요하다고 판단하여, 목적에 맞게 재구성하였다. 이에 따라 중립 154개 단어를 삭제하였고, 매우 부정(-2)과 부정(-1)을 부정(-1)으로 통합하고, 매우 긍정(2)과 긍정(1)을 긍정(1)으로 통합하였다. 이를 통해 최종적으로 -1, 1의 두 가지 감성 값을 갖는 감성 단어 사전을 구축하였다. 각각 부정 9,826개, 긍정 4,863개의 단어로 구성되었다.

3.2 워드 임베딩 추출

감성 대화 데이터와 감성 사전 단어의 문맥 정보를 함축하여 유사성을 분석하기 위해 임베딩 벡터로 표현하는 과정이 필요하다. 이 임베딩 벡터 생성 과정에서는 한국어 텍스트의 의미를 가장 정확하게 반영하는 모델을 평가하기 위해 BERT 계열 언어 모델 2가지와 SBERT 계열 모델 2가지를 사용하여 비교 분석 하였다(<표 3> 참고).

BERT(Bidirectional Encoder Representations from Transformers)는 구글에서 제안한 사전학습 언어 모델로, Transformer 인코더 구조를 기반으로 한다. 기존 언어 모델들이 문장을 단방향적으로만 학습한 것과 달리, BERT는 양방향 학습을 통해 단어의 좌우 문맥을 동시에 반영할 수 있다는 특징을 가진다[19]. 이러한 구조로 인해 텍스트 내 단어 간의 의미적 관계를 깊이 있게 파악할 수 있다. 기본 BERT 모델은 한국어를 포함 다국어 데이터셋으로 사전학습되어 범용성이 높은 모델이다.

〈표 3〉 BERT 계열 및 SBERT 계열 언어 모델

embedding model	개발 연도	학습 언어	embedding dimension
BERT	2018	다국어	768
KoBERT	2019	한국어	768
SBERT	2019	다국어	768
KoSBERT	2021	한국어	768

KoBERT 모델은 SK텔레콤에서 한국어 자연어 처리 성능 향상을 위해 개발한 BERT 기반 모델로, 한국어 위키백과, 뉴스 데이터 등 대규모 한국어 텍스트 데이터 기반으로 사전학습되었다[16]. 이 모델은 영어 중심의 wordpiece 토크나이저 대신, 한국어 문맥구조에 더 적합한 sentencepiece 토크나이저를 사용하여 한국어 문장 구조의 특성을 세밀하게 반영할 수 있다.

SBERT(Sentence-BERT)는 문장 임베딩에 특화된 모델이다. 기존 BERT 모델의 구조를 확장하여 문장 단위의 비교, 의미 파악 등에서 높은 성능을 보인다[17].

KoSBERT는 한국어로 학습된 SBERT 모델로, 한국어 문장 단위 분석에 특화된 모델이다. 한국어 자연어 추론 및 유사도 데이터를 통해 사전 학습되어 한국어 문장 표현의 정확성을 높이기 위한 모델이다[18].

본 연구에서는 BERT, koBERT, SBERT, KoSBERT 네 가지 모델을 비교 분석하여 제품 리뷰와 감성 사전 간의 의미적 유사성 측정에 가장 적합한 임베딩 모델을 도출하고자 하였다.

3.3 제품 리뷰와 감성 단어의 유사도 계산

도출된 제품 리뷰와 감성 사전의 임베딩 값으로 코사인 유사도(Cosine Similarity)를 계산한다. 코사인 유사도는 두 벡터 간의 방향적 유사도를 측정하는 지표로, 각 벡터가 이루는 각도의 코사인 값으로 계산된다. -1부터 1 사이의 값을 가지며 두 벡터가 유사할수록 1에 가깝고 반대일수록 -1에 가깝다.

각 리뷰 문장 임베딩 벡터에 대해 감성 사전에 포함된 14,843개의 감성 단어 임베딩과 코사인 유사도를 계산한다. 즉, 하나의 리뷰 문장에 대해 감성 사전의 모든 단어와의 코사인 유사도가 산출되며, 이는 각 감성 단어가 해당 문장과 의미적으로 얼마나 가까운지를 나타낸다.

3.4 제품 리뷰의 관련 감성 단어 선별

이 과정에서는 리뷰와 감성 단어 사전 간의 코사인 유사도 계산 결과를 바탕으로, 유의미한 수준의 연관성을 보이는 '관련 감성 단어'를 선별한다. 여기서 '유의미하다'는 것은 코사인 유사도 값이 일정 수준 이상으로 높아, 해당 단어가 리뷰와 충분히 밀접한 의미적 관련성을 지닌다고 판단되는 경우를 의미한다.즉, 리뷰의 내용을 가장 잘 설명하거나 대표할 수 있는 감성 단어로 간주하는 것이다. 예를 들어 '제품 성능이 매우 좋습니다.' 라는 문장에 대해 '좋다'라는 감성 단어의 코사인 유사도 절대값이 0.8이고 '기쁘다'가 0.7이라면 '좋다'를 더 유의미하다고 판단한다.

본 연구에서는 문장의 관련 감성 단어를 선별하기 위해 4가지 기준을 설정하였다. 이 기준들은 코사인 유사도 절대값을 척도로 하여, 특정 단어가 기준값을 충족할 경우에만 해당 단어를 관련 감성 단어로 판정하기 위한 판단 기준을 의미한다. 이 4가지 방식은 Top-1 방식, Top n% 방식, Top n개 방식, threshold n 방식이다. 각 기준은 구체적으로 다음과 같다(<표 4> 참고).

Top-1 방식은 코사인 유사도 계산 결과 중 절대값이 가장 큰 단어 1개만을 관련 감성 단어로 선별하는 방법이다. 예를 들어, 감성 사전에 포함된 모든 단어 중 코사인 유사도의 절대값이 가장 큰 단어가 '감사'라면, 해당 단어 하나 만을 관련 감성 단어로 선정한다.

Top n% 방식은 코사인 유사도 계산 결과 중 절대값이 상위 n% 이상인 단어만 관련 감성 단어로 선별하는 방법이다. 본 연구에서 n값은 1, 5, 10으로 설정하여 유사도 절대값 상위 1%, 5%, 10% 이내인 단어들을 관련 감성 단어로 선정한다.

Top n개 방법은 코사인 유사도 계산 결과 중 절대값 높은 순으로 상위 n개의 단어만을 관련 감성 단어로 선별하는 방법이다. n값은 10, 30, 50으로 설정하여 상위 10개, 30개, 50개의 단어만을 관련 감성 단어로 선별하는 방법이다.

Threshold n 방식은 코사인 유사도 계산 결과 중 절대값 n이상인 단어만을 관련 감성 단어로 선별하는 방법이다. 여기서 Threshold는 최소 임계값을 의미하며, 0.1, 0.3, 0.5로 설정하였다. 즉, 유사도 절대값이 0.1, 0.3, 0.5 이상인 단어들만 관련 감성 단어로 선별하는 방법이다.

3.5 제품 리뷰의 감성 점수 계산

선정된 관련 감성 단어들은 감성 단어 사전에 표기된 해당 단어의 감성 점수를 활용한다. 이후 각 관련 감성 단어의 코사인 유사도 값을 가중치로 적용하여, 감성 점수와 곱한 값을 산출하고 이를 모두 합산

〈표 4〉 관련 감성 단어 선별 기준

선별 방식	관련 감성 단어 선별 방법	적용 기준
Top-1	코사인 유사도 절대값 최대인 단어 1개	최대값
Top n%	코사인 유사도 절대값 상위 n%인 단어	n = 1, 5, 10
Top n개	코사인 유사도 절대값 상위 n개인 단어	n = 10, 30, 50
Threshold n	코사인 유사도 절대값 n이상인 단어	n = 0.1, 0.3, 0.5

하고 평균내어 리뷰의 최종 감성 점수를 도출한다.

예를 들어, 리뷰 문장이 “제품이 정말 좋습니다.”라고 할 때, 감성 사전에 포함된 감성 단어가 ‘감사’, ‘좋다’, ‘싫다’이며 각 감성 점수는 각각 1, 1, -1라고 하자. 리뷰 문장과 세 감성 단어 간의 코사인 유사도를 계산한 결과가 좋다 0.9, 감사 0.8, 싫다 0.4로 나타났다고 가정한다. 본 연구에서 Threshold를 0.5로 설정하면 유사도가 0.5 이상인 단어만 관련 감성 단어로 채택되므로, ‘감사’와 ‘좋다’가 선정된다.

선정된 단어의 코사인 유사도 값은 감성 점수의 가중치로 반영되어 다음과 같은 식으로 계산된다. 리뷰 문장 S의 최종 감성 점수는 다음과 같이 계산된다.

$$S = \frac{1 \times 0.9 + 1 \times 0.8}{2} = 0.85$$

위와 같은 계산식에 따라 최종 감성 점수는 0.85로 산출되며, 이는 0 이상이기 때문에 해당 리뷰는 긍정으로 분류된다. 본 연구에서 도출되는 최종 감성 점수는 -1에서 1 사이의 값을 가지며, 감성 점수가 0 이상이면 긍정, 0 미만이면 부정으로 라벨링한다.

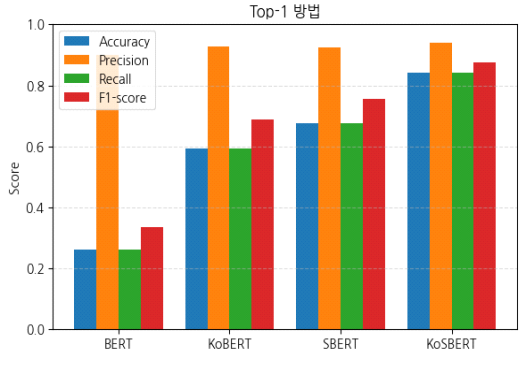
이와 같은 방법으로 도출된 최종 감성 점수는 -1부터 1까지의 값을 가진다. 그리고 최종 감정 점수가 0 이상이면 긍정, 0 미만이면 부정으로 라벨링한다.

4. 제안 방법에 대한 성능 분석

Top-1 방식, Top n% 방식, Top n개 방식, Threshold n 방식, 총 4가지 방법을 적용하여 성능을 비교 분석하였다. 제안한 방법의 성능을 평가하기 위해 Accuracy(정확도), Precision(정밀도), Recall(재현율), F1-score의 네 가지 분류 성능 지표를 활용하였다. Accuracy는 전체 예측 중 정답을 맞힌 비율을 의미하며, Precision은 모델이 정답이라고 예측한 사례 중 실제로 정답인 비율을 나타낸다. Recall은 실제 정답 사례 가운데 모델이 올바르게 정답으로 분류한 비율을 의미하고, F1-score는 Precision과 Recall의 조화평균으로 두 지표 간의 균형을 평가하는 지표이다. 본 연구에서는 이와 같은 지표 정의를 바탕으로 각 방법의 성능을 비교분석하였다. 이를 통해 한국어 제품 리뷰의 감성 분석에 가장 적합한 모델을 식별하고자 하였다.

4.1 Top-1 방법

Top-1 방법은 KoSBERT 모델을 사용한 방법이 가장 높은 성능을 달성하였다. 모든 모델이 0.9 이상의 높은 정밀도를 보였으나, 재현율은 모델 간 편차가 크게 나타났다(<그림 2> 참고). 특히 KoSBERT 모델은 정밀도 0.94, 재현율 0.84로 가장 균형 잡힌 성능을 보이며, F1-score 0.87로 BERT 모델 대비 0.54 증가한 수치를 보였다. 이는 KoSBERT 모델이 다른 모델에 비해 효율적인 분류 성능을 달성했음을 의미한다.



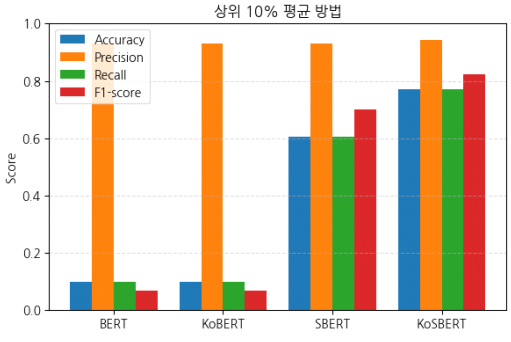
〈그림 2〉 Top-1 방법 결과 비교

4.2 상위 n% 방법

4.2.1 상위 10% 평균 방법

상위 10% 방법은 BERT와 KoBERT 모델에서 모두 낮은 정확도와 f1 score를 보인다(<그림 3> 참고). 그러나 SBERT 모델에서 점차 개선되며 KoSBERT에서 네 가지 성능이 모두 균형 잡힌 것을 확인할 수 있다. 특히 네 모델 모두 정밀도는 0.9 이상으로 높은 수치지만 KoSBERT에서 재현율이 0.0996에서 0.7694로 크게 개선됨을 확인할 수 있다.

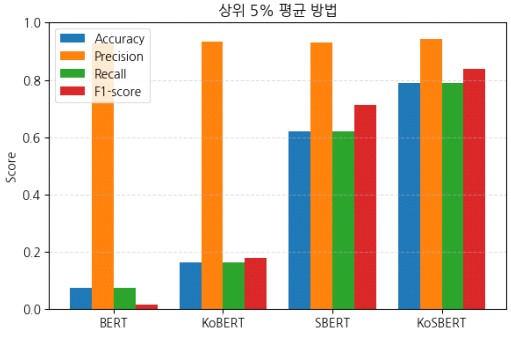
또한 정확도와 f1-score는 SBERT, KoSBERT 모델에서 크게 개선되었다.



〈그림 3〉 상위 10% 방법 결과 비교

4.2.2 상위 5% 평균 방법

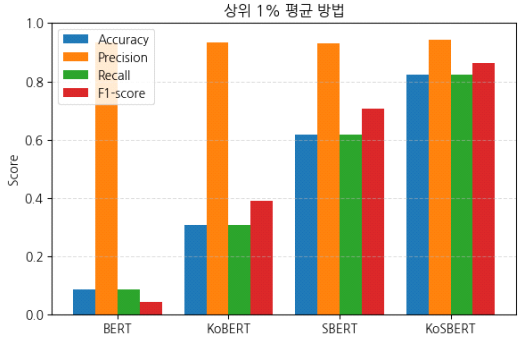
상위 5% 평균 방법도 SBERT 계열 모델인 SBERT와 KoSBERT 모델에서 정확도와 재현율, f1-score가 크게 증가했음을 확인할 수 있다(<그림 4> 참고). BERT 모델과 KoBERT 모델에서 정밀도는 0.90 이상으로 모든 모델에서 높았지만, 재현율과 큰 편차를 보인다. 반면, KoSBERT 모델에서 재현율 0.7899, f1-score 0.8392로 균형 있는 성능을 달성했다.



〈그림 4〉 상위 5% 평균 방법 결과 비교

4.2.3 상위 1% 평균 방법

상위 1% 평균 방법은 BERT 모델에서 0.088로 낮은 정확도를 보이지만 점차 다른 모델에서 개선됨을 확인할 수 있다(<그림 5> 참고). KoSBERT 모델에서 정확도가 0.8235의 수치로 크게 향상되었다. 정밀도는 네 가지 모델에서 모두 0.9 이상의 성능을 보였으나 BERT 계열 모델에서 재현율과 큰 편차를 보인다. SBERT 계열 모델에서 점차 향상되었으며 KoSBERT에서 정밀도 0.9437, 재현율 0.8235, f1-score 0.8628로 균형 잡힌 높은 성능을 보였다.

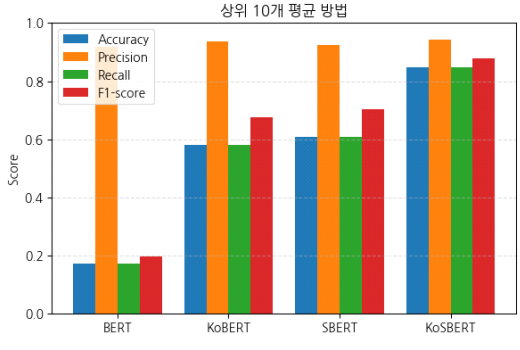


〈그림 5〉 상위 1% 평균 방법 결과 비교

4.3 상위 n개 방법

4.3.1 상위 10개 평균 방법

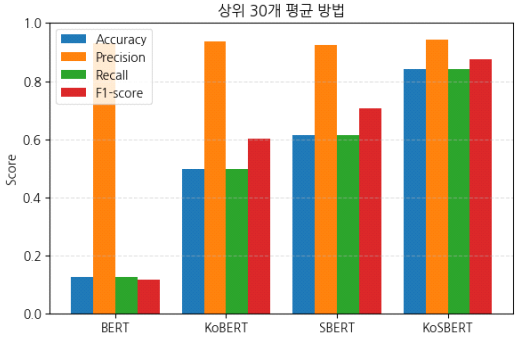
상위 10개 평균 방법은 BERT 모델에 비해 KoSBERT 모델에서 정확도가 크게 개선되었음을 확인할 수 있다(<그림 6> 참고). 모든 모델에서 정밀도는 높지만, BERT 계열의 모델은 재현율과 큰 편차를 보인다. 반면 KoSBERT 모델에서 정밀도 0.9433, 재현율 0.8466, f1-score 0.8787로 가장 균형 잡힌 높은 성능을 보였다.



〈그림 6〉 상위 10개 평균 방법 결과 비교

4.3.2 상위 30개 평균 방법

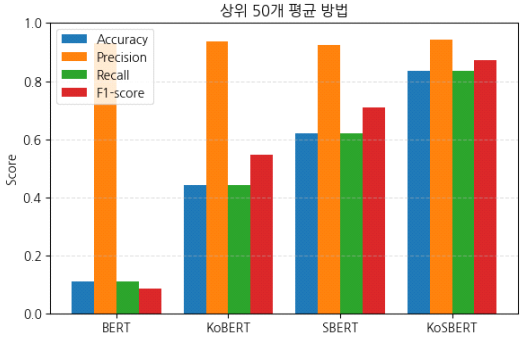
상위 30개 평균 방법도 BERT 계열 모델에 비해 KoSBERT 모델에서 성능이 크게 증가하였다(<그림 7> 참고). 정밀도는 네 가지 모델에서 모두 높았으나 재현율은 KoSBERT에서 크게 개선되었다. KoSBERT는 정확도 0.8405, 정밀도 0.9438, 재현율 0.8405, f1-score 0.8745로 가장 높은 성능을 보였다.



〈그림 7〉 상위 30개 평균 방법 결과 비교

4.3.3 상위 50개 평균 방법

상위 50개 평균 방법도 정밀도는 모두 높았지만 BERT와 KoBERT 모델에서 낮은 정확도와 정밀도, f1-score 수치를 보였다(<그림 8> 참고). SBERT 모델에서 점차 개선되었으며 KoSBERT 모델에서 정확도 0.8367, 정밀도 0.9437, 재현율 0.8367, f1-score 0.8720으로 안정적인 성능을 보인다.

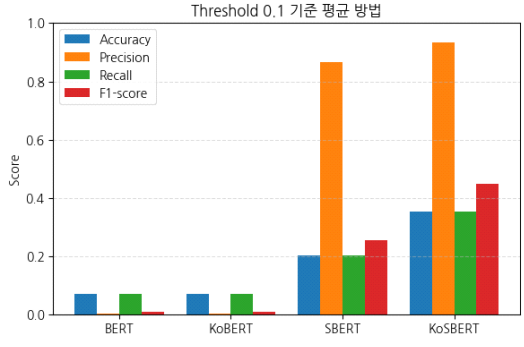


〈그림 8〉 상위 50개 평균 방법 결과 비교

4.4 Threshold 기준 평균 방법

4.4.1 Threshold 0.1 기준 평균 방법

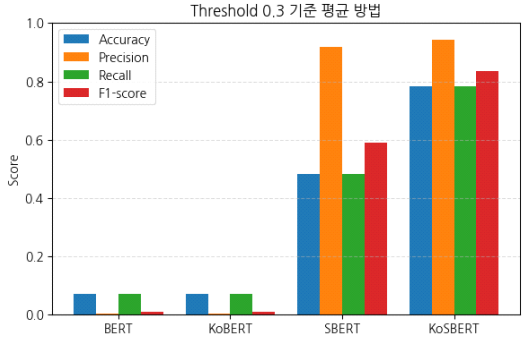
Threshold 0.1 기준 평균 방법은 네 가지 모델 모두 낮은 성능을 보였다(<그림 9> 참고). 다른 방법에서 가장 높은 성능을 보인 KoSBERT 모델에서도 정확도 0.3551, 재현율 0.9348, 재현율 0.3551, f1-score 0.4492를 보이며 낮은 성능을 드러냈다. 이는 언어 모델이 한국어 문맥을 잘 반영하는지의 여부와 관련없이, 해당 기준은 문장의 감성을 평가하는데 부적절한 방법임을 확인할 수 있다.



〈그림 9〉 Threshold 0.1 평균 방법 결과 비교

4.4.2 Threshold 0.3 기준 평균 방법

Threshold 0.3 기준 평균 방법에서는 BERT 계열 모델에 비해 KoSBERT 모델에서 성능이 크게 개선되었음을 확인할 수 있다(<그림 10> 참고). 정밀도

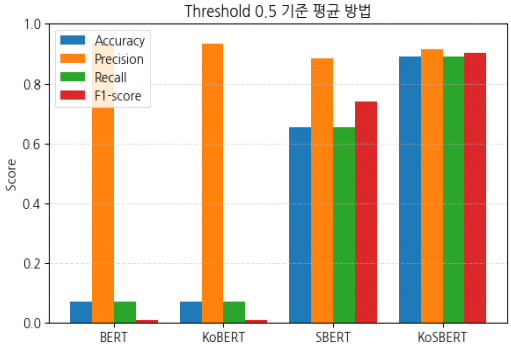


〈그림 10〉 Threshold 0.3 평균 방법 결과 비교

는 SBERT와 KoSBERT에서 0.9 이상으로 크게 증가하였고 특히 KoSBERT 모델은 정확도 0.7843, 정밀도 0.9422, 재현율 0.7843, f1-score 0.8352로 크게 향상된 성능을 보였다.

4.4.3 Threshold 0.5 기준 평균 방법

Threshold 0.5 기준 평균 방법에서는 KoSBERT의 정확도가 BERT 모델에 비해 약 0.89 높게 나타났으나, 정밀도는 BERT(0.9348)보다 다소 낮은 0.9136을 기록하였다(<그림 11> 참고). BERT는 개별 단어의 감성에 민감하게 반응해 코사인 유사도 자체는 높았으나, 문장 전체 문맥을 반영하지 못해 실제 감정 방향과 불일치하여 낮은 정확도를 보인 것으로 해석할 수 있다. 반면 KoSBERT는 문장 단위 임베딩을 통해 감정의 흐름과 극성을 일관되게 반영함으로써, 정밀도는 다소 낮더라도 전체 감정 분류 정확도와 재현율이 크게 향상되었다.



〈그림 11〉 Threshold 0.5 평균 방법 결과 비교

4.5 종합 결과

네 가지 감성 분석 방법 모두에서 KoSBERT 모델이 가장 높은 성능을 기록하였다(<표 5> 참고). 이는 문장 수준 임베딩에 최적화된 SBERT 계열 모델이 기존의 BERT 기반 모델보다 감성 분류 성능에서 우수함을 시사한다. 특히, 한국어 데이터로 학습된 KoSBERT는 한국어 제품 리뷰 데이터에 대한 감성 분석에 가장 적합한 모델임이 실험을 통해 입증되었다. 이는 KoSBERT가 BERT에 비해 문장 내 감성적 문맥을 보다 효과적으로 반영하는 임베딩을 생성하였음을 의미한다.

〈표 6〉 종합 결과 비교

기준	acc	pre	rec	f1-score
Top-1	0.8409	0.9398	0.8409	0.8743
상위 1%	0.8235	**0.9437**	0.8235	0.8628
상위 10개	0.8466	0.9433	0.8466	0.8787
Threshold 0.5	**0.8919**	0.9136	**0.8919**	**0.9014**

한편, 각 방법에서 가장 뛰어난 성능을 보인 기준을 제시하면 다음과 같다. Top n% 방식에서는 상위 1% 단어 평균 방식이 가장 높은 성능을 나타냈으며, Top n개 방식에서는 상위 10개 단어 평균 방식, Threshold 기반 평균 방식에서는 임계값 0.5 기준의 평균 방식이 각각 최적의 결과를 보였다. 또한 Top n%, Top n개, Threshold n 방법에서 단어의 수를 제한할수록 성능이 개선되었는데, Top-1 방법보다 우수한 결과를 보였다. 이는 단일 단어에 과도하게 의존할 경우 문장 전체의 감정 정보를 충분히 반영하지 못하고, 특정 단어의 편향에 의해 감성 점수가 왜곡될 가능성이 높기 때문으로 판단된다. 반면 일정 수 이상의 감성 단어를 활용하는 방법들은 필요한 정보는 유지하면서도 노이즈가 되는 단어는 걸러내는 균형적 구조를 가지므로, 문맥에 더 적합한 감성 점수를 산출하여 상대적으로 안정적이고 우수한 성능을 보이는 것으로 해석된다.

반면 Top n%, Top n, Threshold n 방식은 감성 단어의 수를 적절하게 제한하면서도 단일 단어에 치우치지 않는 균형적 감성 반영이 가능하여, 보다 안정적인 감성 점수를 산출할 수 있었다.

즉, 감성 단어의 수가 너무 적으면 정보 손실이 발생하고, 지나치게 많으면 노이즈가 증가하므로, 일정 수준의 필터링을 적용한 세 방식이 Top-1 방식보다 우수한 성능을 보인 것으로 판단된다.

또한 대부분의 결과에서 BERT 계열 모델은 낮은 성능을 보이지만 정밀도만 매우 높은 것을 확인할 수

있다. 이와 같은 현상은 데이터 불균형 상황에서 모델이 Positive(정답) 클래스라고 판단한 사례가 극히 일부에 불과할 때 발생한다. 모델이 소수의 사례만을 Positive로 예측하고 대부분을 Negative(오답)로 분류하는 경우, Positive로 예측한 소수의 사례가 우연히 맞아 정밀도는 높게 유지될 수 있다. 그러나 실제 Positive 사례의 대부분을 놓치기 때문에 재현율이 매우 낮아지고, 전체 분포를 반영하는 정확도 역시 낮아지는 결과가 나타난다. 이는 불균형한 클래스 분포에서 흔히 발생하는 전형적인 현상이다. 반면 KoSBERT 기반 방식은 정밀도뿐 아니라 재현율과 F1-score까지 고르게 향상된 것으로 나타났다. 이는 KoSBERT가 문맥 정보를 보다 정확하게 반영함으로써, 클래스 불균형 상황에서도 상대적으로 안정적인 성능을 유지할 수 있음을 시사한다.

이에 따라 본 연구에서는 가장 우수한 성능을 기록한 KoSBERT 모델을 사용하고, 각 방식 별로 성능을 비교 분석하였다

실험 결과, Threshold 0.5 기준 평균 방법이 정확도 0.8919, 정밀도 0.9136, 재현율 0.8919, f1-score 0.9014로 가장 균형 잡힌 높은 성능을 달성했음을 확인할 수 있다. 이는 KoSBERT를 통해 텍스트의 문맥을 반영하는 벡터를 생성하고, 유사도 절대값 기준으로 상위 0.5%의 단어를 평균 내는 방법이 한국어 문장 감성 라벨링에 있어 가장 효과적인 방법임을 의미한다.

모델별 성능 비교 결과는 다음과 같다. 이는 각 방법에서 도출된 accuracy, precision, recall, f1-score 값의 평균을 계산한 것이다. 전반적으로 BERT 계열 모델은 매우 낮은 수준의 성능을 보였으며 특히 KoSBERT는 정확도 측면에서 가장 우수한 성능을 나타냈으며, 다른 모델에 비해 현저히 높은 결과를 달성하였다(<표 6> 참고). 이를 통해 한국어 문맥을 가장 효과적으로 반영하는 모델은 문장 임베딩에 최적이고, 한국어로 사전학습된 KoSBERT 모델임을 확인할 수 있다.

〈표 6〉 모델별 종합 결과 비교

Model	acc	pre	rec	f1-score
BERT	0.1142	0.7429	0.1142	0.0892
KoBERT	0.2895	0.7478	0.2895	0.3179
SBERT	0.5703	0.9159	0.5703	0.6573
KoSBERT	**0.7579**	**0.9389**	**0.7579**	**0.8212**

5. 벤치마크 모델과의 성능 비교분석

본 연구에서는 기존 감성 분석 연구에서 널리 활용되는 두 가지 주요 접근법을 벤치마크로 설정하여 제안한 방법의 성능을 검증하였다. 첫 번째는 감성 사전을 이용한 통계 기반 방법이며, 두 번째는 사전학습 언어모델을 활용한 지도학습 기반 감성 분석 방법이다. 두 벤치마크 모두 본 연구에서 사용한 것과 동일한 제품 리뷰 데이터를 기반으로 평가를 수행하였으며, 이를 통해 제안한 비지도 임베딩 기반 방법의 상대적 성능을 정량적으로 비교분석하였다.

첫 번째 벤치마크로 감성 사전을 활용한 전통적인 통계 기반 감성 분석 방법을 적용하였다. 이 방법은 리뷰 문장에 포함된 단어가 감성 사전에 존재하는지를 확인하고, 해당 단어의 감성 점수를 단순 누적하여 문장의 전체 감성 점수를 계산하는 방식이다. 구체적으로는 감성 사전에 포함된 각 단어에 대해 리뷰 문장 내 등장 여부를 탐색하고, 등장할 경우 그 단어에 부여된 감성 점수를 합산하여 리뷰의 감성 점수를 산출하였다. 산출된 최종 감성 점수는 다음과 같은 규칙에 따라 라벨을 부여하였다. 감성 점수가 0보다 큰 경우 긍정, 0보다 작은 경우 부정으로 분류하였다.

두 번째 벤치마크로 사전학습 언어모델인 KR-ELECTRA를 활용하여, 리뷰 텍스트를 입력으로 하는 이진 분류 모델 학습을 수행하였다. KR-ELECTRA는 Google Research에서 제안한 ELECTRA 구조를 한국어에 맞게 학습시킨 한국어 전용 언어모델이다[15]. 전체 데이터는 훈련(70%), 검증(20%), 테스트(10%) 비율로 분할하였다. 학습 과정에서 조기 종료를 도입하여, 검증 데이터 기준으로 5회 연속 성능 개선이 없을 경우 학습을 중단하도록 설정하였다. 학습은 총 6회(epoch) 동안 진행되었으며, 결과는 정확

도 0.958, 정밀도 0.479, 재현율 0.500, f1-score 0.489로 나타났다.

벤치마크와의 비교 결과는 <표 7>과 같다. 먼저 벤치마크 1(감성 사전을 활용한 통계 기반 방법)은 Accuracy 0.6368, F1-score 0.3268로 가장 낮은 성능을 보였다. 이는 감성 단어의 존재 여부만을 기준으로 감성을 판단하는 단순 통계적 방식이 실제 감성을 효과적으로 분석하는 데 한계가 있음을 보여준다.

〈표 7〉 벤치마크 비교 결과

방법	acc	pre	rec	F1-score
Benchmark 1	0.6368	0.3694	0.4280	0.3268
Benchmark 2	0.958	0.479	0.500	0.489
Proposed-best	0.8919	0.9136	0.8919	0.9014

벤치마크 2는(KR-ELECTRA를 활용한 지도학습 기반 분류 모델)는 Accuracy 0.958로 가장 높은 정확도를 기록하였으나, Precision(0.479), Recall (0.500), F1-score(0.489)는 상대적으로 낮았다. 이는 제한된 양의 학습 데이터로는 모델이 충분한 패턴을 학습하기 어렵다는 점을 시사하며, 특히 데이터 불균형 상황에서는 소수 클래스에 대한 학습이 제대로 이루어지지 않음을 의미한다. 또한 지도학습 모델이 높은 일반화 성능을 확보하기 위해서는 더 많은 양질의 라벨링 데이터가 필요함을 보여준다.

반면, 제안한 방법 중 성능이 가장 우수했던 KoSBERT 모델을 사용한 Threshold 0.5 평균 방식은 Accuracy 0.8919, F1-score 0.9014로 전반적인 지표에서 가장 균형 잡힌 성능을 나타냈다. 특히 Precision과 Recall 모두 0.89 이상으로 높은 값을 기록하였다. 이는 제안된 임베딩 기반 접근법이 문맥 정보를 효과적으로 반영하면서도 기존 감성 사전 방식의 한계를 보완하고, 지도학습 기반 방법보다 적은 자원으로도 높은 성능을 낼 수 있음을 보여준다.

6. 결 론

본 연구에서는 감성 사전과 워드 임베딩을 활용한 감성 분석 방법을 제안하였고 다양한 임베딩 모델과 감성 점수 산출 방식을 조합하여 비교 분석하였다. 실험 결과, SBERT 구조 기반의 KoSERT모델이 모든 평가 지표에서 가장 우수한 성능을 기록하였다. 특히 상위 Threshold 0.5 기준 평균 방법이 가장 높은 정확도(0.8919)와 f1 score(0.9014)를 보여, 약 89% 수준의 성능을 입증하였다.

제안한 방법은 별도의 대규모 딥러닝 파인튜닝 과정 없이도 충분히 높은 감성 분류 성능을 달성할 수 있음을 보여준다. 임베딩 벡터를 통해 문장의 문맥 정보를 반영하고, 감성 사전 단어와의 유사성 분석을 통해 대규모 학습 없이도 효과적인 감성 분석을 수행하였다. 이 방법은 기존 감성 분석 연구에서 문맥을 반영하지 못한다는 한계를 보완할 수 있고, 추가적인 학습 과정이 필요하지 않아 라벨링 데이터 구축 및 모델 미세 조정에 필요한 비용과 시간을 절감할 수 있다. 유사도 계산 과정으로 인한 계산량이 증가하지만, 별도의 학습 단계가 필요 없는 가벼운 구조를 통해 실용성과 효율성을 확보할 수 있다.

본 연구에서 제안한 방법은 한국어 텍스트 분석을 다루는 다양한 도메인에도 적용할 수 있다. 향후 연구에서는 특정 제품군이나 서비스의 리뷰에 대해 감성 분석을 수행한 뒤, 긍정 그룹과 부정 그룹으로 구분하여 각각의 강점과 보완점을 도출할 수 있을 것이다. 이를 통해 사용자에게는 개인화된 추천 서비스를 제공하고, 서비스 제공자에게는 데이터 기반의 개선 방향을 제시할 수 있을 것으로 기대된다.

참 고 문 헌

[1] 박준성, 남수연, 이찬호, 구성모(2025), “토픽모델링을 이용한 러닝 앱 리뷰 감성분석,” 한국품질경영학회지, 53(1), 19-32.

[2] 박현정, 신경식(2020), “BERT를 활용한 속성 기반 감성분석: 속성카테고리 감성분류 모델 개발,” 지능정보연구, 26(4), 1-25.

[3] 반건우, 안소미, 박소영(2022), “텍스트마이닝을 통한 제인 에어 의 토픽분석 및 감성분석,”

인문사회 21, 13(2), 2415-2428.

[4] 서혜선(2023), “SNS 텍스트 데이터 기반 대전시 환경 관련 토픽 모델링 및 감성 분석,” 한국데이터분석학회지, 25(3), 2023, 949-958.

[5] 오정심(2024), “텍스트 마이닝을 활용한 ‘문화경관’ 인식과 감성 분석: 전라남도 신안·진도·완도 사례를 중심으로,” 한국콘텐츠학회논문지, 24(3), 255-264

[6] 이대국, 김지영, 전수영(2024), “감정 어휘 사전을 이용한 문학 감정분류,” Journal of the Korean Data Analysis Society, 26(2), 457-469.

[7] 이민아, 박연지, 나준영, 손채봉(2023), “KoBERT, KoGPT-2, KoBART 활용 및 하이퍼파라미터 최적화를 진행한 리뷰 감성분석 애플리케이션 구현,” 디지털콘텐츠학회논문지, 24(11), 2831-2840.

[8] 이현수, 김민하, 서지원, 김정이(2023), “자연어 처리 딥러닝 모델 감정분석을 통한 감성 콘텐츠 개발 연구,” 문화기술의 융합 저널, 9(4), 687-692.

[9] 전수현, 노미진(2025), “온라인 패션 리뷰 데이터를 활용한 속성 기반 감성 분석,” 경영교육연구, 40(1), 183-196.

[10] 정환웅, 서지훈(2025), “LLM 기반의 게임 이용자 경험 개선을 위한 리뷰 감성분석,” Journal of KIIT, 23(3), 51-59.

[11] 지경규(2022), “텍스트 애널리틱스를 활용한 서술형 강의평가 분석,” 융합지식학회논문지, 10(1), 71-80.

[12] AI Hub(2022), “속성기반 감정분석 데이터,” 한국지능정보사회진흥원(NIA), https://www.aihub.or.kr/aihubdata/data/view.do?dataSetSn=71603.

[13] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K.(2019), “BERT: Pre-training of deep bidirectional transformers for language understanding,” In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT) (pp. 4171-4186). Association for Computational Linguistics.

[14] 박상민, 온병언, 나철원. (2025), “KnuSentiLex,” GitHub repository: https://github.com/park1200656/KnuSentiLex/tree/master/KnuSentiLex/data.

[15] Park, J.(2020), “KoELECTRA: Pretrained ELECTRA model for Korean,” GitHub repository, Available at: https://github.com/monologg/KoELECTRA.

[16] Park, S., Park, S., Moon, J., Kim, S., Cho, W. I., Han, J., Park, J., Song, C., Kim, J., Song, Y., Oh, T., Lee, J., Oh, J., Lyu, S., Jeong, Y., Lee, I., Seo, S., Lee, D., Kim, H., Lee, M., Jang, S., Do, S., Kim, S., Lim, K., Lee, J., Park, K., Shin, J., Kim, S., Park, L., Oh, A., Ha, J.-W., and Cho, K.(2021), “KLUE: Korean language understanding evaluation,” arXiv preprint arXiv:2105.09680.

[17] Reimers, N. and Gurevych, I.(2019), “Sentence-BERT: Sentence embeddings using Siamese BERT-networks,” In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP), 3982-3992.

[18] sentence-transformers(n.d.), “jhgan/ko-sroberta-multitask,” [Pretrained sentence embedding model], Hugging Face, From https://huggingface.co/jhgan/ko-sroberta-multitask.

[19] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., and Polosukhin, I.(2017), “Attention is all you need,” In Advances in Neural Information Processing Systems, 5998-6008.