



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

2024년 2월
석사학위 논문

텍스트 데이터 감성 분류 개선을 위한
Fine-tuning 방법 제안

조선대학교 산업기술창업대학원

소프트웨어융합공학과

박 정 일

텍스트 데이터 감성 분류 개선을 위한 Fine-tuning 방법 제안

Fine-tuning method proposed to improve text data
sentiment classification

2024년 2월 23일

조선대학교 산업기술창업대학원

소프트웨어융합공학과

박 정 일

텍스트 데이터 감성 분류 개선을 위한
Fine-tuning 방법 제안

지도교수 김 판 구

이 논문을 공학 석사학위신청 논문으로 제출함.

2023년 10월

조선대학교 산업기술창업대학원

소프트웨어융합공학과

박 정 일

박정일의 석사학위논문을 인준함

위원장 조선대학교 교수 최준호 (인)

위 원 조선대학교 교수 김판구 (인)

위 원 조선대학교 교수 신주현 (인)

2023년 11월

조선대학교 산업기술창업대학원

목 차

ABSTRACT

I. 서론	1
A. 연구 배경 및 목적	1
B. 연구 내용 및 구성	3
II. 관련 연구	4
A. 감성 분류	4
1. 음성 기반 감성 분류	4
2. 얼굴 표정 기반 감성 분류	6
3. 텍스트 기반 감성 분류	7
B. 감성 분류 모델	9
1. 비지도 학습 모델	9
2. 지도 학습 모델	10
3. 앙상블 모델	13
4. Fine-tuning 모델	15
III. 텍스트 감성 분류 Fine-tuning 모델	18
A. 연구 구성도	18
B. 데이터 전처리	20
C. 감성 분류 Fine-tuning 모델	26

IV. 실험 및 결과	30
A. 실험 데이터셋	30
B. 실험 평가 방법 및 결과 분석	39
1. 실험 평가 방법	39
2. 실험 평가 결과 분석	40
V. 결론 및 향후 연구	48
참고문헌	49

표 목 차

[표 2-1] Performance Comparison Results Between Existing and New Models	14
[표 3-1] Quantity of Training and Testing Dataset Preprocessing	25
[표 3-2] Hyper Parameter of KoELECTRA	27
[표 3-3] Configuration of KoELECTRA-base	28
[표 4-1] Development Environment	30
[표 4-2] Naver Sentiment Movie Corpus Summary	33
[표 4-3] Naver Shopping Corpus Summary	35
[표 4-4] Steam Corpus Summary	37
[표 4-5] Confusion Matrix Terms	39
[표 4-6] Dataset Testing and Accuracy Confirmation	41
[표 4-7] Dataset Cross-testing and Accuracy Confirmation	42
[표 4-8] Summary of 3 Dataset Characteristics	44
[표 4-9] Dataset 2 Cross-testing and Accuracy Confirmation	45
[표 4-10] Dataset 3 Cross-testing and Accuracy Confirmation	46

그림 목 차

[그림 2-1] An Overview of Replaced Token Detection in ELECTRA	11
[그림 2-2] ELECTRA’s Replaced Token Detection Pre-training Results	12
[그림 2-3] Four Fine-tuning Methods Presented by Keras	16
[그림 3-1] Research Structure for Improving Emotional Classification	18
[그림 3-2] Naver Shopping Corpus Original Label Format	20
[그림 3-3] Naver Shopping Corpus Preprocessing Label Format	20
[그림 3-4] Check Distribution of Positive and Negative Labels After Preprocessing Naver Shopping Corpus	21
[그림 3-5] Train, Test Original Data from Naver Sentiment Movie Corpus ..	22
[그림 3-6] Train and Test Preprocessing Data from Naver Sentiment Movie Corpus ..	22
[그림 3-7] Naver Shopping Corpus Original Data	23
[그림 3-8] Naver Shopping Corpus Preprocessed Data	23
[그림 3-9] Steam Corpus Original Data	24
[그림 3-10] Steam Corpus Preprocessed Data	24
[그림 3-11] Message Using HuggingFace’s Transformers Library	26
[그림 4-1] Naver Sentiment Movie Corpus Data Example	32
[그림 4-2] Naver Shopping Corpus Data Example	32
[그림 4-3] Steam Corpus Data Example	32
[그림 4-4] Naver Movie Ratings and Reviews Screen	34
[그림 4-5] Naver Shopping Ratings and Reviews Screen	36
[그림 4-6] Steam Game Ratings and Reviews Screen	38
[그림 4-7] Accuracy Calculation Formula	39
[그림 4-8] Model Accuracy Fine-tuning Graph by Dataset	43
[그림 4-9] Model Accuracy Fine-tuning Graph per Dataset	43
[그림 4-10] Comparison of Accuracy Between Random F-T and HLSN F-T in Two Types	45
[그림 4-11] Comparison of Accuracy Between Random F-T and HLSN F-T in Three Types	47

ABSTRACT

Research on efficient fine-tuning methods for Korean sentiment analysis

Park, Jung Il

Advisor : Prof. Kim, Pan Koo Ph.D.

Department of Software Convergence
Engineering

Graduate School of Industrial Technology
and Entrepreneurship, Chosun University

Modern enterprises are increasingly adopting sentiment analysis as a crucial task, highlighting the accurate understanding of consumer opinions across various domains such as social media, product reviews, and customer feedback as a key challenge for success in competition. Sentiment analysis helps enhance products and services by grasping diverse opinions and sentiments of consumers.

The utilization of large-scale text data collection and fine-tuning with pre-trained language models plays a vital role in improving performance. With recent technological advancements, sentiment analysis models demonstrate high performance, and the ELECTRA model, in particular, provides outstanding results through efficient learning methods and fewer computational resources.

The purpose of this study is to optimize the performance of the model through efficient fine-tuning on various datasets using the KoELECTRA model that learned Korean in ELECTRA. An AI model to classify human sentiments as positive and negative and predict them. Through this, we expect to be able to achieve higher and faster performance in NLP tasks by efficiently fine-tuning various datasets suitable for the purpose.

I. 서론

A. 연구 배경 및 목적

감성분류는 현대 기업들에게 더 이상 선택이 아닌 필수적인 작업으로 자리매김하고 있다. 특히 소셜 미디어 플랫폼, 제품 리뷰, 그리고 고객 피드백과 같은 다양한 도메인에서 소비자들의 의견을 정확히 이해하는 것은 기업이 경쟁에서 성공을 거두기 위해 핵심적인 과제로 부각되고 있다.

소비자들의 다양한 의견과 감성을 파악하고 분석하는데 감성분류는 중요한 도구로 작용한다. 이는 기업이 제품이나 서비스를 지속적으로 향상시키기 위해 소비자들의 니즈와 선호도를 정확하게 파악하는 데 도움이 된다. 최근에는 대량의 텍스트 데이터가 손쉽게 수집되고 있어, 감성분류 모델을 더 정교하게 조정하는 데 활용될 수 있는 풍부한 자료가 확보되고 있다.

또한, 사전 훈련된 언어 모델들이 다양한 자연어 처리 작업에서 우수한 성능을 보여주고 있다. 이러한 모델들을 기반으로 한 Fine-tuning은 감성 분류 능력을 향상시키는 데 매우 효과적인 전략으로 부각되고 있다. 이는 감성분류 모델이 다양한 도메인에서 빠르게 대응하면서도 특정 업무에 특화된 성능을 발휘할 수 있도록 도와준다.

사전 훈련된 언어 모델과 강화학습 알고리즘 등의 발전으로 감성분류 모델이 높은 성능을 발휘하고 있고, 클라우드 기술의 진보로 대규모의 텍스트 데이터를 신속하게 처리하고 저장하는 것이 가능해졌다. 이러한 기술적 진보는 기업이 소셜 미디어, 고객 피드백, 제품 리뷰 등에서 발생하는 다양한 텍스트 데이터를 실시간으로 분석하고 이해하는 데 도움을 준다.

ELECTRA는 Efficiently Learning an Encoder that Classifies Token Replacements Accurately의 약자로, 자연어 처리(NLP) 분야에서 사용되는 언어 모

델이다. 이 모델은 텍스트 데이터를 이해하고 생성하는 데 사용되며, 다양한 NLP 작업에 적용될 수 있다. ELECTRA는 기존의 언어 모델보다 효율적인 학습 방법을 제공하며, 더 적은 계산 리소스를 사용하여 더 높은 성능을 달성할 수 있다.

Fine-tuning은 사전 학습된 모델을 특정 작업에 맞게 조정하는 과정으로 이 과정에서는 모델의 가중치를 작업 데이터에 맞게 조정하며, 이를 통해 모델이 특정 작업에 대해 더 정확한 예측을 할 수 있게 된다. ELECTRA 모델을 Fine-tuning 하는 것은 특정 NLP 작업에 대한 모델의 성능을 크게 향상시킬 수 있다.

따라서, 본 연구의 목적은 ELECTRA에 한국어를 학습한 KoELECTRA 모델을 사용하여 다양한 데이터셋에 대해서 효율적인 Fine-tuning을 통해 모델의 성능을 최적화하는 것으로 사람의 감성을 긍정, 부정으로 분류하고 이를 예측하기 위한 AI 모델을 훈련한다. 이를 통해 우리는 목적에 맞는 다양한 데이터셋을 효율적으로 Fine-tuning 하여 NLP 작업에서 더 높고 빠르게 좋은 성능을 달성할 수 있을 것으로 기대한다.

B. 연구 내용 및 구성

본 연구의 주요 내용은 사용자의 감성을 좀 더 정확하게 분석하기 위한 AI모델 학습 방법을 제시하기 위해 본 논문은 다음과 같은 구성으로 작성되었다.

다양한 주제에서 느끼는 텍스트 리뷰 감성에 대한 수집 데이터를 확인하고 이를 긍정과 부정으로 구분하여 기준을 잡고 텍스트 데이터 전처리를 진행한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구로 감성 분류에 대한 음성 기반, 얼굴 표정 기반, 텍스트 기반과 관련된 기존 연구를 기술하고 감성 분류 모델에서는 지도 학습과 비지도 학습에 관련된 기존 연구와 더불어 앙상블 모델과 연속적인 학습을 위한 Fine-tuning 방법에 대한 기준을 기술한다.

3장에서는 다양한 종류의 다른 데이터셋을 통해 텍스트 감성 분류 개선을 위한 Fine-tuning 모델을 제안한다. 감성분석의 성능 향상을 위해 데이터셋에 라벨링 된 감성에 대한 구조를 분석하고 그러한 텍스트 데이터를 효율적으로 Fine-tuning 하여 실제 감성 분류가 가능한 텍스트로부터 감성 분류의 정확도 향상을 위한 방법에 관해 기술한다.

4장에서는 제안한 Fine-tuning 학습 처리 방법과 감성분류 모델을 실험 및 평가 하고, 마지막으로 5장에서는 결론과 향후 연구에 관해 기술하고 마무리한다.

II. 관련 연구

A. 감성 분류

감성 분석(Sentiment Analysis)은 일련의 텍스트, 음성 또는 이미지와 같은 매체에서 사람의 감성을 자동으로 식별하고 분석하는 기술이다. 이는 기계 학습과 자연어 처리(NLP) 기술을 활용하여 이루어지며, 텍스트, 음성의 억양, 이미지의 얼굴 표정 등을 분석하여 감성 상태를 인식하려고 한다.

소셜 미디어, 고객 응대, 마케팅, 정치 분석, 의료 및 심리학 등 다양한 분야를 통해 기업은 제품에 대한 고객의 감성을 이해하여 제품 개선에 활용하거나 고객 서비스를 개선하기 위해 고객의 감성을 파악할 수 있고 소셜 미디어에서의 감성 반응을 파악하여 특정 이벤트나 제품의 평판을 이해하고 대응하는 데에도 활용된다.

기술이 계속 발전함에 따라 인간의 감성을 보다 정확하게 이해하고 해석할 수 있도록 발전하고 있으나 여전히 인간의 감성을 완벽하게 이해하는 것은 어려운 과제이며 이를 위한 연구와 기술 개발은 계속적으로 진행되고 있다.

1. 음성 기반 감성 분류

음성 신호를 분석하고 해석하여 말하는 사람의 감성 상태를 감지하고 인식하는 기술 또는 시스템을 의미한다. 이 기술은 주로 음성 처리, 기계 학습 및 심층 학습 기술을 활용하여 구현된다. 음성 기반 감성 인식은 다음과 같은 몇 가지 주요 목표와 응용 분야를 가지고 있다.

첫 번째로, 감성 분류 측면에서, 이 기술은 음성 신호에서 감성 특성을 추출하고 해당 감성을 식별하는 데 사용된다. 감성을 인식할 수 있는 일반적인 요소로 단어, 톤, 음성신호의 피치, 포먼트, 그리고 발음 속도 및 음질 등이 있고 이를 통해 음성 데이터를 분석하여 사용자의 감성을 이해하고, 이를 통해 감성적인 반응을 분석하고 해석할 수 있다. [1]

두 번째로, 상호 작용 및 인터페이스 측면에서, 음성 기반 감성 인식은 음성 인식 시스템, 가상 비서 및 챗봇과 같은 응용 프로그램에서 사용자와의 상호 작용을

향상시키는 데 활용된다. 사람의 감성은 주관적인 욕구와 의도를 무의식적으로 표현하며, 이는 다양한 상황 정보 중에서 사용자의 의도를 가장 많이 반영하는 정보라고 볼 수 있다. [2] 예를 들어, 사용자의 감성에 따라 시스템이 더 자연스럽게 응답할 수 있으며, 이는 사용자 경험을 향상시키는 데 도움이 된다.

세 번째로, 감성 분류 기반 마케팅 측면에서, 이 기술은 광고, 제품 설계, 및 마케팅 전략 개발에 활용된다. 소비자의 감성 반응을 이해함으로써 제품 또는 서비스에 대한 효과적인 마케팅 전략을 수립할 수 있으며, 이는 제품의 성공과 소비자 만족도를 높일 수 있다. 추천 시스템은 사용자의 선호도를 이해하여 제품 구매 결정을 지원하는 역할을 하는데 더불어, 비즈니스 전략 측면에서도 중요한 역할을 하므로 많은 기업과 기관이 이에 큰 관심을 가지고 있습니다. 최근에는 다양한 추천 시스템 연구 중에서 특히 NLP와 딥러닝을 결합한 하이브리드 추천 시스템 연구가 증가하고 있다. [3]

네 번째로, 감성 지원 및 건강 관리 측면에서, 감성 분류를 통해 개인의 감성적인 상태를 모니터링하고 분석함으로써, 스트레스, 우울증 및 기타 정신 건강 문제를 조기에 감지하고 지원을 제공하는 데 도움을 줄 수 있다. 감성 표현은 보편적이며, 감성 상태는 우리 생활의 모든 영역에 중요한 영향을 미치는 매우 핵심적인 요소이다. 현재까지 감성이 유발된 상황에서 획득한 뇌파를 분석하고 해당 결과를 토대로 감성 상태를 정의하는 노력은 주로 심리학자들에 의해 수행되었으나 최근 연구결과에 따르면, 감성과 관련된 정보는 정신활동을 지배하는 뇌가 활성화될 때 발생하는 뇌파를 통해서도 파악할 수 있다는 것이 밝혀졌다. [4]

마지막으로, 보안 및 인증 측면에서, 음성 기반 감성 인식은 음성을 사용한 생체 인증 시스템의 한 부분으로 사용될 수 있다. 음성인식 기술의 활발한 적용은 관련된 사이버 공격과 위협에 대한 우려도 높이고 있고 기존 연구는 자동화 화자 인증 (ASV: Automatic Speaker Verification) 기술의 개발과 정확성 향상에 주로 집중되어 왔으나, 현재 실제 음성인식 서비스에서 사용되는 자동화 화자 식별 기술에 대한 사이버 공격 및 위협에 관한 분석 및 연구는 다양하고 체계적으로 수행되지 않는 경우가 있다. [5] 사용자의 감성 상태를 고려하여 보안 시스템을 보다 견고하게 만들어 더 안전한 인증 절차를 제공할 수 있어야 한다.

2. 얼굴 표정 기반 감성 분류

기술과 인공지능 알고리즘을 사용하여 사람의 얼굴 표정을 분석하여 해당 얼굴의 감성을 감지하고 인식하는 기술을 가리킨다. 이 기술은 컴퓨터 비전과 기계 학습 기술을 결합하여 사용된다.

얼굴 표정 기반 감성 인식의 주요 목표 중 하나는 사람의 감성을 이해하고 상호 작용하는 컴퓨터 시스템 및 응용 프로그램을 개발하는 것으로 이를 통해 컴퓨터나 로봇은 사용자의 감성 상태를 파악하고 이에 적절하게 대응할 수 있다.

얼굴 표정 인식 연구의 표정인식을 위한 특징 추출 방법은 크게 통계적 성분 분석 기반의 방법과 특징점 기반의 방법으로 구분할 수 있다. 먼저, 성분 분석을 기반으로 하는 특징 추출 방법은 추출된 얼굴 영역의 전체 또는 일부 픽셀 정보를 활용하며, 주로 고유 벡터(eigen vector)를 활용하여 특징을 추출한다. 반면에 특징점 기반의 특징 추출 방법은 얼굴의 특정 부분인 눈, 입과 같은 특징점의 위치 정보를 활용하며, 특징점 구성 방법에 따라 여러 가지 알고리즘이 존재한다. [6]

이러한 알고리즘을 사용하여 얼굴 표정 기반 감성 인식은 다음과 같은 주요 단계로 구성된다. 먼저, 이미지나 비디오에서 얼굴을 식별하고 추출하는 단계를 통해 컴퓨터 비전 기술을 활용하여 얼굴 영역을 인식하고 분리하여 얼굴 특징 추출이 이어진다. 이 단계에서 얼굴의 구체적인 특징을 추출하고, 눈, 코, 입, 눈썹 등의 위치를 파악하며, 표정 특징을 식별한다.

이후, 추출된 얼굴 특징을 사용하여 해당 얼굴의 감성을 분류하고 인식하는 감성 분류 단계가 진행된다. 예를 들어, 웃고 있는 얼굴은 '기쁨'으로 분류될 수 있으며, 얼굴의 표정 특징과 표현을 분석하여 다양한 감성 중 하나로 분류한다.

마지막으로, 감성 분류 결과를 활용하여 컴퓨터 시스템이 적절하게 대응하거나 상호 작용하도록 설계되는데 이를 통해 사용자 경험을 개선하거나, 마케팅 및 광고 분석을 위한 데이터 수집에 활용하며, 감성 인식 로봇 및 인공지능 비서의 인터랙션 개선, 그리고 안전 및 보안 분야에서 위험 감지 및 대응에 활용된다. 이러한 기술은 다양한 응용 분야에서 중요한 역할을 하며, 사람과 컴퓨터 간의 상호 작용을 더욱 자연스럽게 만들어주는 역할을 한다.

얼굴 표정 기반 감성 인식은 다양한 분야에서 사용될 수 있으며, 이를 통해 사용자 경험 개선, 마케팅 및 광고 분석, 감성 인식 로봇 및 인공지능 비서, 안전 및 보안 분야 등 다양한 응용 분야에서 활용된다.

3. 텍스트 기반 감성 분류

인간의 감성을 분류 및 분석하기에 앞서 감성 분석은 텍스트에 나타난 주관적 요소인 긍정/부정 감성을 판별해 수치, 도식, 등급 등으로 정량화하는 작업이다. 이는 단순히 긍정/부정을 추출하는 것이 아니라, 대상이 되는 개체나 속성을 추출하거나, 감성을 표현하는 이들의 감성, 감정, 태도, 입장 등을 분석한다. 대중이 만들어낸 텍스트에 사용된 감성 표현 언어들을 '추출-범주화 및 분류-예측'의 단계를 거쳐 정량화한다. 일반적으로 극성에 따라 '이분법적(긍정/부정) + 중립(선택적)' 또는 '정도를 부여해 0 ~ 10의 수치'로 나타낸다.

감성 분석의 주요 요소로는 감성 표현(sentimental Expression), 대상(target), 개체(Entity), 개체 속성(Aспект), 의견보유자(opinion holder)의 속성, 시간에 따른 추이 등이 있다.

감성 분석은 다음과 같은 단계로 이루어진다:

1. 데이터 수집 : SNS와 같은 웹사이트 매체에서 정보를 수집한다.
2. 주관성 탐지 : 수집된 정보에서 작성자의 주관에 드러난 부분만 걸러낸다.
3. 주관성 극성 또는 정도 분류 : 주어진 문자 데이터의 패턴을 학습해 감성 분류 예측 모델을 구축한다.

구축한 모델에 학습 데이터와 유사한 특성을 갖는 분석 대상 데이터를 넣어 분석 수행한다. 모델의 학습을 위해 문서 또는 문장에 '긍정/부정'의 라벨로 분류된 데이터가 충분히 있어야 한다. 없다면 모델 구축이 불가능하므로 감성사전 기반 분석으로 적용해야 한다.

수집된 문자 데이터를 전처리해 정리한 후 미리 구축된 감성 사전과 매칭하여 극성감성단어 빈도를 정량화하여 감성 분석한다. 감성 사전은 감성단어에 긍정/부정과 그 극성 범주 값과 특성을 나타내는 감성 점수로 구성된다. 감성 사전의 품질이 분석의 성과에 직결된다.

전반적인 긍/부정이 아닌 속성 단위의 감성분류로 분석 대상을 더욱 정교하게 분석한다. 다수의 문서에서 나타나는 속성별 감성/평가를 해당 속성별로 취합해 제공하여 정교한 분석결과 제시가 가능하다.

기계학습기반 중 하나인 지도학습(Supervised Learning) 기법에는 결정트리 분류

기(Neural Network), 나이브 베이즈 분류기(Naive Bayes Classifier), 서포트 벡터 분류기(Support Vector Machine), 최대 엔트로피 모델(Maximum Entropy) 등이 있다.

최근에는 그런 과정 없이 단어의 특성을 저차원상의 실수 벡터 값으로 나타내는 워드 임베딩(Word Embedding) 기법이 등장하며 NLP 분야의 주목을 받고 있다. 단어들의 의미가 유사하면 벡터 공간상에 가깝게 배치하여 단어들 간의 어휘적 관계를 벡터로 표현하는 표현 방법이다.

감성 분석은 다양한 분야에서 활용될 수 있는데 예를 들어, 화장품 관련 상품의 구매후기 데이터를 수집하여 감성 분석 기법을 적용하면 마케팅 대시보드를 제공하는 것이 목표가 될 수 있다. 이를 위해 데이터 수집, 데이터 전처리, 형태소 사전 구축, 속성어 사전 구축, 감성어 사전 구축, 감성 분석 모형 구축 및 결과 도출 등의 과정이 필요하다.

감성 분석을 성공적으로 수행하기 위해서는 적용 분야별 특성을 살린 사전을 잘 구축해야 한다. '형태소 사전, 감성어 사전, 속성어 사전'같은 인프라 사전을 잘 구축하는 것이 매우 중요하며, 이는 지속적으로 업데이트되어야 한다. 또한, 데이터 수집 전략을 세울 필요가 있으며, 다른 데이터와 연계해 다양한 분석을 수행해야 한다.

감성 분류는 컴퓨터가 텍스트 데이터를 분석하여 해당 텍스트에서 어떤 감성이나 감성 상태가 표현되고 있는지 인식하고 이해하는 기술을 가리킨다. 이 기술은 자연어 처리(Natural Language Processing, NLP) 및 기계 학습 기술을 기반으로 하며, 주로 다음과 같은 작업을 수행한다.

텍스트로부터 감성을 인식하는 능력은, 기계가 인간의 감성을 인식하는 과정에서 음성이나 비전과 같이 단순히 현재 입력 데이터에 대한 감성 분류를 해주는 수준을 넘는 중요한 역할을 한다. 이 능력은 과거 기억, 감성 주체의 관점, 그리고 감성 주체의 성향과 성격과 같은 다양한 요소를 고려하여 감성을 지능적으로 이해하고 인식하게 하는데 필요하다. [7]

감성 분류는 주어진 텍스트가 긍정적인지, 부정적인지 또는 중립적인지를 분류한다. 예를 들어, 소셜 미디어 글이나 제품 리뷰를 분석하여 긍정적인 리뷰와 부정적인 리뷰를 식별할 수 있다.

B. 감성 분류 모델

1. 비지도 학습 모델

비지도 학습 기반 방법은 감성 분류를 수행할 때 레이블이 지정되지 않은 데이터를 사용하여 감성적인 패턴이나 특징을 찾는 방법으로 이 방법은 데이터에 대한 사전 정보 없이 감성을 탐색하고 모델링할 수 있다.

감성 클러스터링은 비지도 학습에서 감성 분류를 수행하는 일반적인 방법 중 하나로 이 방법은 데이터 포인트를 서로 비슷한 감성을 가진 그룹 또는 클러스터로 그룹화한다. 클러스터링 알고리즘(예: K-평균 클러스터링, DBSCAN)을 사용하여 데이터를 클러스터로 분할하고, 각 클러스터에 어떤 감성이 관련되어 있는지를 파악한다.

감성 어휘 사전은 감성 어휘 사전은 감성과 관련된 단어와 문구를 수집하고 분석하는 방법 중 하나로 각 단어나 문구는 긍정, 부정 또는 다른 감성 카테고리와 연관시킬 수 있다. 이러한 사전을 사용하여 주어진 텍스트에서 어떤 감성 단어가 나타나는지를 분석하고, 이를 통해 전체 텍스트의 감성을 파악한다.

차원 축소 기술(예: 주성분 분석, t-SNE)을 사용하여 데이터의 특성을 시각화하고 감성적인 패턴을 발견하는 데 도움을 줄 수 있다. 이러한 기술을 사용하면 데이터를 저차원 공간으로 투영하여 패턴을 눈으로 확인할 수 있다.

비지도 학습 기반 감성 분류에서 기계 학습 기술을 사용할 수도 있다. 예를 들어, 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA)과 같은 토픽 모델링 알고리즘을 사용하여 텍스트 데이터에서 주요 주제와 감성적인 측면을 파악할 수 있다.

한국어의 의미역이 태깅된 대량의 말뭉치를 구하기 어려운 경우 이를 직접 구축하기 위해서는 상당한 시간과 노력이 필요하기 때문에 [8]의 연구를 보면 말뭉치를 이용하여 통계적으로 의미역 결정(semantic role labeling)을 수행하기 위해서 의미역을 태깅하는 작업을 필수적으로 진행하고 비지도 학습의 하나인 self-training 알고리즘을 적용하여, 의미역이 태깅되지 않은 말뭉치로부터 의미역을 결정하는 방법을 제안한다. 이를 위해, 세종 용언 전자사전의 격률정보를 활용하여 자동으로 학습 말뭉치를 구축하였으며, 확률 모델을 적용하여 점진적으로 학습하였다. 이러한

과정에서 4개의 부사격 조사에 대해 평균적으로 83.00%의 정확도를 달성한다.

하지만 비지도 학습은 데이터에 대한 사전 정보가 제한적인 상황에서 유용하며, 레이블링 작업을 필요로 하지 않으나 정확도와 해석 가능성 측면에서 한계가 있을 수 있으며, 모델의 성능을 평가하기 어려울 수 있다.

2. 지도 학습 모델

지도 학습 기반 방법은 감성 분류를 수행하기 위해 레이블이 지정된 훈련 데이터를 활용하는 방법으로, 이 데이터를 기반으로 모델을 학습시키는 과정을 포함한다.

데이터 수집 및 레이블링 방법으로 레이블이 지정된 훈련 데이터가 필요하고 감성 분류의 경우, 텍스트나 음성 데이터에는 각각 어떤 감성(긍정, 부정, 중립 또는 더 세분화된 감성 카테고리)이 포함되어 있어야 한다. 이 레이블은 수작업으로 또는 자동으로 붙일 수 있다.

특징 추출을 위해 훈련 데이터에서 감성을 예측하는 데 도움이 되는 특징(feature)을 추출하고 텍스트 데이터의 경우, 단어, 문장, n-그램(n-grams), 단어 임베딩(word embeddings) 등의 특징을 사용할 수 있다. 각 특징은 데이터를 수치화하고 모델이 학습할 수 있는 형식으로 변환된다.

감성 분류에는 다양한 모델이 사용될 수 있으며, 선택한 모델에 따라 예측 성능이 달라진다. 예를 들어, 텍스트 데이터에 대한 감성 분류에는 순환 신경망(RNN), 컨볼루션 신경망(CNN), BERT, ELECTRA와 같은 트랜스포머 모델 등이 사용된다. [9, 10, 11]

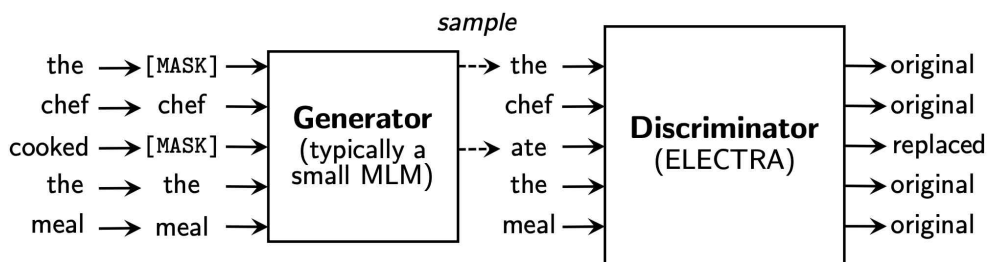
모델을 훈련 데이터로 학습시키고, 검증 데이터를 사용하여 모델의 성능을 평가한다. 이를 통해 모델의 하이퍼파라미터를 조정하거나 과적합을 방지하기 위한 조치를 취할 수 있다.

모델이 훈련된 후, 새로운 텍스트나 음성 데이터에 대한 감성을 예측할 수 있고 모델은 입력 데이터를 분석하고 해당 데이터의 감성을 출력으로 반환한다.

지도 학습 기반 감성 분류는 레이블 된 데이터가 필요하고, 모델의 성능은 데이터의 품질과 양에 크게 의존한다. 따라서 데이터 수집과 레이블링 과정이 중요하며, 높은 정확도를 얻기 위해서는 충분한 양의 훈련 데이터가 필요할 수 있다.

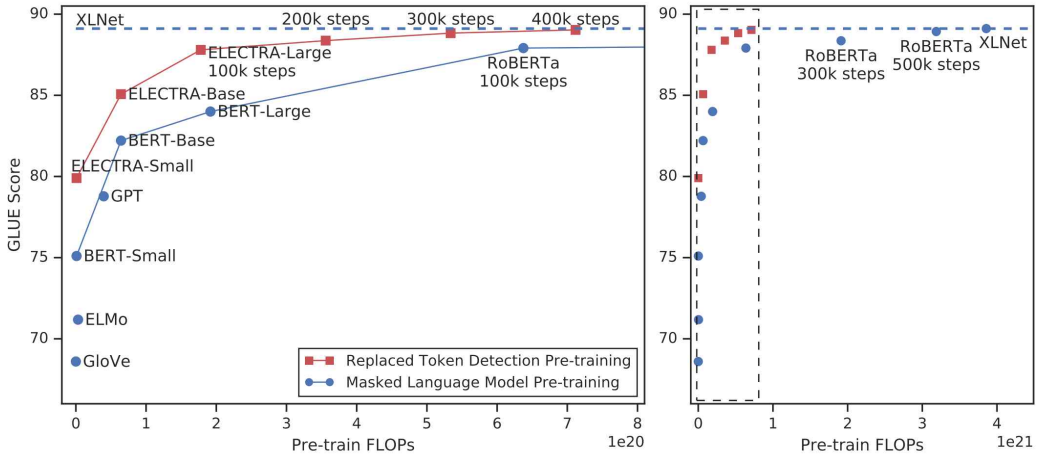
모델의 정확도와 함께 학습의 효율성에도 주목할만한 새로운 Pre-training 기법을 적용한 language model로 ELECTRA (Efficiently Learning an Encoder that Classifies Token Replacements Accurately) [11]의 연구에 따르면, BERT를 포함한 대부분의 언어 모델들은 입력을 마스크 토큰으로 대체하고, 이를 원래 토큰으로 복원하는 Masked Language Modeling(MLM)을 통해 Pre-training을 진행한다.

Masked language modeling(MLM)은 양방향 정보를 고려하여 Autoregressive Language Modeling 학습에 비해 효과적인 학습 방법일 수 있지만, 이 방법에는 몇 가지 문제점이 존재한다. 문제점으로는 전체 토큰의 15%만 학습하여 나머지 부분에 대한 손실의 발생과 손실로 인해 생기는 compute budget 부담 그리고 학습 시 MASK 토큰을 참고하여 모델을 통해 예측하지만 실제 추론(inference)을 할 때에는 MASK 토큰이 존재하지 않는다. 이는 모델이 실제 문장을 처리할 때 MASK 토큰을 만나지 않기 때문이다.



[그림 2-1] An Overview of Replaced Token Detection in ELECTRA

이러한 문제를 해결하기 위해 논문에서는 학습 효율을 높이기 위해 [그림 2-1]의 방법으로 Replaced Token Detection(RTD)라는 새로운 Pre-training task을 제안하고 있다. RTD는 Generator를 활용하여 실제 입력 토큰의 일부를 진짜가 아닌 토큰으로 치환하며, Discriminator를 통해 이 토큰들이 원래 입력에 속한 진짜 토큰인지 또는 Generator가 생성한 진짜가 아닌 토큰인지를 이진 분류 문제로 판단한다. ELECTRA는 입력의 전체 토큰에 대해 RTD 작업을 수행하므로, 입력의 15%만을 대상으로 하는 방법보다 효율적이고 효과적이다.



[그림 2-2] ELECTRA's Replaced Token Detection Pre-training Results

이 연구에서는 General Language Understanding Evaluation(GLUE) 벤치마크와 Stanford Question Answering(SQuAD) 데이터셋을 활용하여 모델의 성능을 평가하였다. 모델의 Pre-training에는 주로 Wikipedia와 BooksCorpus를 사용하였으며, 특히 Large 모델에서는 XLNet에서 사용된 ClueWeb, CommonCrawl, Gigaword 등의 데이터셋을 추가로 활용하였다. 이는 BERT와 동일한 모델 구조와 하이퍼파라미터를 사용하기 위함이고, 일부 평가 데이터셋의 크기가 작아 랜덤 시드의 영향을 크게 받을 수 있어, 이를 보완하기 위해 10번의 Fine-tuning 결과 중 중간값(median)을 최종 성능으로 선택한다.

[그림 2-1]을 통해 BERT와 ELECTRA의 [10,11] 성능과 효율성을 확인해 보면 ELECTRA는 모든 학습 과정을 데이터, 모델, 계산량을 동일한 조건으로 진행했을 때 높은 GLUE 성능을 보여준다. ELECTRA-Small은 BERT-Large에 비해 파라미터 수가 1/20이고, 계산량은 1/135인데도 불구하고 단일 GPU를 사용하여 4일 만에 학습이 가능했다. 또한, ELECTRA-Small은 BERT-Small보다 GLUE 성능이 5 point 높고, 더 큰 모델인 GPT보다도 성능이 우수하다는 걸 확인할 수 있다. 마지막으로 이 모델 또한 훌륭한 성능을 보여주는데 ELECTRA-Large는 RoBERTa나 XLNet보다 파라미터가 적고, 계산량은 1/4로 학습했음에도 불구하고 이들과 유사한 성능을 보이는 등 GLUE에서는 ELECTRA를 능가하였고, SQuAD 2.0에서는 최고의 성능을 보여주었다. 따라서, ELECTRA는 새롭게 제안된 RTD 태스크를 통해

기존의 representation learning 기법들보다 더 효율적으로 학습할 수 있음을 보이고 이는 매우 빠르게 성능이 향상되는 것을 확인할 수 있다.

이를 통해 ELECTRA는 더욱 빠르고 효과적으로 학습할 수 있고 동일한 모델 크기, 데이터, 컴퓨팅 리소스를 사용하면서도 기존 BERT의 성능을 초과함을 확인할 수 있고 이러한 성과는 Small 모델 설정에서 특히 두드러짐이 인상적이다.

3. 앙상블 모델

감성 분류를 위한 모델 앙상블 학습은 여러 다른 머신 러닝 모델을 결합하여 더 나은 예측 성능을 달성하기 위한 기술로 감성 분류는 주로 텍스트, 오디오, 비디오 또는 이미지와 같은 데이터에서 감성 정보를 추출하는 작업을 포함한다.

데이터를 더 조금 사용하여 유사하거나 향상된 성능을 보이는 새로운 모델을 위해 [12]의 연구에서는 코로나19 이슈로 지난 2년 동안 꾸준히 뉴스 도메인에서 다루어져 왔으며, 감염 확산 상황에 따라 다양한 감성을 반영한 기사가 생산되었다. 이러한 이유로, 코로나19를 감성 분류의 대상으로 선정하여 데이터 수집은 네이버 포털 검색 시스템을 통해 코로나19 관련 뉴스 기사를 월별로 약 200~300건 크롤링하여 진행하였다. 이후, 분석의 정확도를 높이기 위해 추출된 뉴스 기사 제목에서 관용구와 특수문자 등을 제거하는 전처리 과정을 거쳐 언어 모델의 Fine-Tuning을 위한 학습 데이터를 준비하였다. 총 8,000건의 데이터를 라벨링 하고, 이를 학습 데이터와 검증 데이터로 분리하여 데이터셋을 구성하였다.

감성분류를 위해 감성사전 기반의 방법을 사용하였고 성능 향상을 위해 도메인 특화 단어를 추가적으로 적용하는 방식을 구현하였다. 범용 감성사전으로는 KNU 감성 사전[13]을 사용하여 이를 통해 매칭률 기반의 감성분류를 수행하였다. 그러나 범용 감성사전은 일반적인 감성 표현에 따른 긍정, 부정 분류에 중점을 두므로, 특정 도메인에 대한 감성분류 성능을 향상시키기 위해서는 도메인 특화 단어 사전의 개발이 필요하다. 따라서 본 연구에서는 166개의 도메인 특화 단어를 개발하여 범용 감성 사전을 보완한다.

자연어 처리 분야에서는 기계학습이 중요한 역할을 하고 있고 특히, 언어 모델을 활용한 자연어 처리는 뛰어난 성능을 보여주고 있다. 그중에서 BERT[10]에 비해 학습 자원이 적게 들면서도 우수한 성능을 발휘하는 ELECTRA[11] 모델을 활용하여 한국어를 학습한 KoELECTRA[14]를 사용한다. 라벨링 된 코로나19 관련 기사

제목 데이터를 활용하여 KoELECTRA 모델을 Fine-Tuning 하였고, 이를 통해 감성 분류를 진행하였다.

KoELECTRA 언어 모델을 사용하여 NSMC[15] 데이터를 기반으로 감성 분류 모델을 만들었다. 이 모델은 코로나19 데이터를 사용하여 추가 학습(Post Training)을 거쳤고 초기에는 NSMC 데이터로 사전 학습된 모델을 사용하여 코로나19 뉴스 기사의 감성 분류를 시도했지만, 뉴스 기사 제목에 대한 감성 분류의 정확도가 만족스럽지 않았다. 이에 따라, 코로나19 뉴스 기사 제목을 사용하여 모델에 추가 학습(Post Training)을 진행하였다. 이 과정을 통해, 코로나19 뉴스 기사 데이터만을 사용하여 학습한 감성 분류 모델보다 더 우수한 성능의 모델을 구현하고자 했다.

[표 2-1] Performance Comparison Results Between Existing and New Models

모델	F-score	Accuracy	Precision	Recall
단어사전	0.4384	0.2807	0.2809	0.9981
KoELECTRA	0.8144	0.8961	0.8183	0.8106
KoELECTRA (Post Learn)	0.8092	0.8903	0.7916	0.8276
양상블 모델	0.8144	0.8961	0.8183	0.8206

범용감성 사전과 도메인사전, KoELECTRA 언어모델, KoELECTRA 언어모델 Post Training의 평가를 바탕으로 모델의 분석된 결과를 라벨링 데이터로 구현하여 학습데이터의 라벨링에 소요되는 비용을 절감하는데 활용되도록 하는 방안을 연구하여 양상블 모델의 결과를 확인한다.

4. Fine-tuning 모델

Fine-tuning은 머신 러닝 모델을 특정 작업이나 데이터에 맞게 조정하는 프로세스를 말한다. 기본적으로 미리 훈련된(pre-trained) 모델을 가져와서 새로운 작업이나 데이터에 대해 추가로 훈련시킨다. 이는 미리 학습된 모델이 갖고 있는 일반적인 특성을 살려서 특정 작업에 빠르게 적응할 수 있게 해 준다.

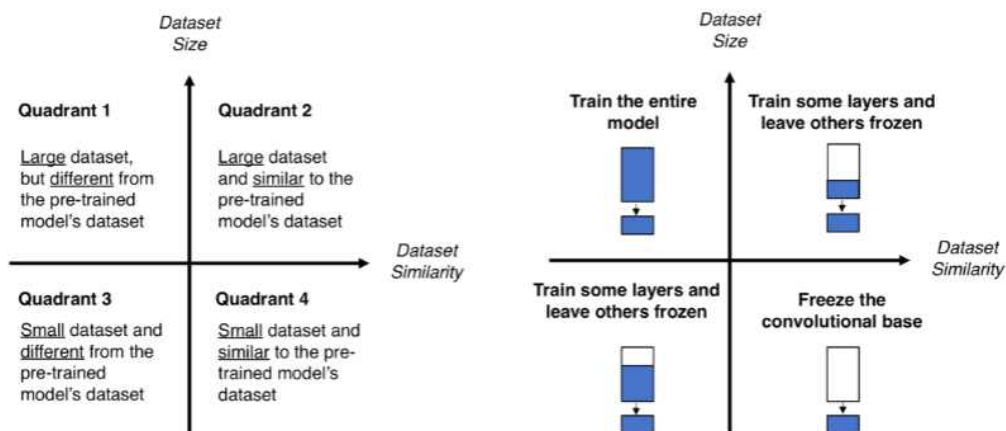
Fine-tuning은 여러 분야에서 사용되고 있으며, 자연어 처리(NLP), 이미지 분류, 음성 인식 등 다양한 작업에 적용됩니다. 아래는 NLP 분야에서의 Fine-tuning에 대한 구체적인 설명이다. 설명의 순서는 미리 훈련된 모델 선택, 새로운 작업에 맞게 모델 구성, 새로운 데이터로 모델 훈련, 하이퍼파라미터 조정으로 서술한다.

미리 훈련된 모델 선택은 대규모 데이터셋에서 사전 훈련된 모델을 선택하고 이 모델은 일반적인 언어 구조나 이미지 특징 등을 학습한 상태로 대표적인 예시로는 OpenAI의 GPT 모델이나 ELECTRA, BERT 모델이 있다.

새로운 작업에 맞게 모델 구성은 선택한 모델을 기반으로 새로운 작업에 적합한 형태로 모델을 수정한다. 이때, 모델의 일부 레이어를 동결시키거나 새로운 레이어를 추가하는 등의 조정을 할 수 있다.

새로운 데이터로 모델 훈련은 선택한 모델을 새로운 데이터셋으로 추가로 훈련한다. 이때, 미리 학습된 가중치는 유지하면서 새로운 작업에 특화된 가중치를 조절하게 된다. 하이퍼파라미터 조정은 훈련 과정에서 모델의 성능을 최적화하기 위해 하이퍼파라미터를 조정한다. 이는 학습률, 배치 크기, 에폭 수 등을 조절하는 것을 포함한다. 그리고 최종적으로 성능 평가를 진행하여 훈련된 모델을 평가하고 새로운 작업에 대한 성능을 측정한다. 이 단계에서는 검증 데이터셋이나 테스트 데이터셋을 사용한다.

Fine-tuning의 장점은 새로운 작업에 대해 매우 효율적으로 모델을 적용할 수 있다는 점이고 적절한 미리 학습된 모델과 충분한 양의 훈련 데이터가 필요하며, 모델을 지속적으로 업데이트해야 하는 상황에서 유용하게 활용된다.



[그림 2-3] Four Fine-tuning Methods Presented by Keras

Keras는 딥러닝 모델을 쉽게 구축하고 훈련시키기 위한 고수준 신경망 API로, Python에서 사용 가능한 오픈 소스 라이브러리이다. 초기에는 독립적인 라이브러리였지만 TensorFlow와 통합된 후 TensorFlow의 주요 고수준 API로 채택되었다. 현재는 TensorFlow의 일부로서 공식적으로 tf.keras를 통해 사용할 수 있다.

간단하고 직관적인 API를 제공하여 입문자나 중급자가 쉽게 딥러닝 모델을 개발할 수 있도록 도와준다. 모델 정의, 계층 구성, 최적화, 손실 함수 등 딥러닝에 필요한 다양한 기능을 제공하며, 모듈성과 확장성을 갖추어 다양한 신경망 아키텍처를 실험하고 구축하는 데 적합하다. 이러한 Fine-tuning은 언어 모델에서 많은 문서를 학습시킨 것을 내 것에 맞게 변형시키고자 할 때 사용하여 더 많은 데이터로 학습된 것을 기초로 조금씩 수정을 진행한다.

전이학습(Transfer Learning) 및 미세조정(Fine-tuning)에 대한 가이드에서는, 제시하는 Task의 성격에 따른 Fine-tuning 4가지 전략에 대해서 4분위로 나눠 설명한다.

1사분면에서 제시하는 기준을 보면 데이터셋이 크고, 유사성이 작은 경우는 모델 전체를 학습시키는 것이 낫다. 데이터셋의 크기가 크기 때문에 충분히 학습이 가능하고 데이터셋끼리의 유사성이 작다고 해도 모델의 구조와 파라미터들은 여전히 재사용 가능하므로, 재학습을 시켜주는 편이 낫다.

2사분면에서는 데이터셋이 크고, 유사성이 큰 경우 Convolution base의 뒷부분과

Classifier를 학습시킨다. 사실 최적의 경우이기 때문에 모든 옵션을 선택할 수 있지만, 데이터셋이 유사하기 때문에 전체를 학습시키기보다는 강한 feature가 나타나는 Convolution base의 뒷부분과 Classifier만 새로 학습시키는 것이 최적이다.

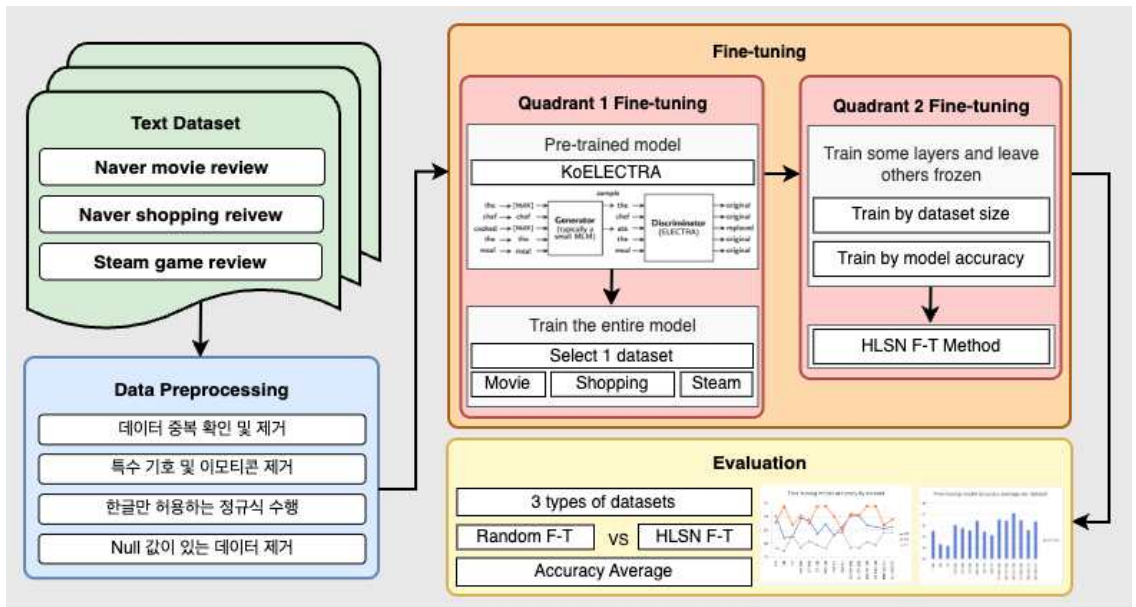
3사분면에서는 데이터셋이 작고, 유사성이 작은 경우로 Convolution base의 일부분과 Classifier를 학습시킨다. 가장 나쁜 상황으로, 제일 어렵다. 데이터가 적기 때문에, 적은 레이어를 Fine tuning 하면 별 효과가 없고, 많은 레이어를 Fine tuning 했다가는 오버피팅이 발생할 것이다. 따라서 Convolution base의 어느 정도를 새로 학습시켜야 할지를 적당히 잡아주어야 한다.

마지막으로 4사분면에서는 데이터셋이 작고, 유사성이 큰 경우로 Classifier만 학습시킨다. 데이터가 적기 때문에 많은 레이어를 Fine tuning 할 경우 오버피팅이 발생한다. 따라서 앞부분의 Feature Extraction은 그대로 쓰고, 최종 Classifier의 FC Layer들에 대해서만 Fine tuning을 진행한다.

Ⅲ. 텍스트 감성 분류 Fine-tuning 모델

A. 연구 구성도

본 절에서는 제안하는 텍스트 데이터 감성 분류 개선을 위한 Fine-tuning 모델에 대해 기술한다. [그림 3-1]은 본 논문에서 제안하는 효율적 Fine-tuning을 위한 데이터 학습 방법을 설명한다.



[그림 3-1] Research Structure for Improving Emotional Classification

비대면 상황에서 텍스트만으로 감성을 정확히 인식하는 것은 효율적인 의사소통을 위한 중요한 과제이다. 감성 대화 데이터를 통한 연구는 자연어 처리 및 기계학습 모델의 감성 이해 능력을 향상시키는데 중요한 역할을 한다. 이러한 연구는 다양한 감성 표현과 문맥을 이해하고 해석함으로써 감성 인식 알고리즘의 정확성과 신뢰성을 높일 수 있다. 또한, 동일한 감성이라도 긍정적인 면과 부정적인 면이 있을 수 있음을 고려하는 것이 필요하다. 따라서 네이버 감성 영화 리뷰 데이터셋 (Naver sentiment movie corpus), 네이버 쇼핑 리뷰 데이터셋 (Naver Shopping

Corpus) 그리고 게임 유통 서비스인 Steam의 각종 게임에 달린 리뷰 데이터셋 (Steam Corpus)의 긍정 및 부정 데이터를 사용하여 KoELECTRA 모델로 이러한 데이터셋을 교차로 Fine-tuning 한다.

이 과정에서 단일 데이터셋으로 학습한 모델과 Fine-tuning을 거친 모델의 성능을 비교하여, 실제 Fine-tuning을 선행하는 것이 학습하려는 데이터셋의 크기와 종류에 따라 얼마나 효율적인 성능을 나타내는지 확인한다. 이 기술은 비대면 상호작용에서 감성을 더 정확하게 파악하고 이해하는 데 기본적인 문제인 감성의 긍정과 부정을 더 정확하게 분류함으로써 더 나은 해결책을 제시할 것으로 기대된다.

B. 데이터 전처리

본 절에서는 다양한 리뷰 문장을 통해 극성 데이터에 해당하는 긍정과 부정 라벨링 및 한글 텍스트 위주의 데이터 전처리를 진행한다. 전처리 과정에 대한 예시는 Naver Shopping Corpus를 통해서 진행한다.

전체 데이터의 개수와 label 열의 개수를 먼저 확인하고 label 열이 0 또는 1로 긍정과 부정을 나타내는 두 가지 값으로만 구성되지 않은 경우 해당 label 값에 대해서 파악한다. Naver Shopping Corpus의 경우 별점이 1에서 5까지의 숫자로 표현되는 경우이고 이에 따라 별점 1과 2를 부정으로 간주하며, 별점 3은 중립으로 간주하여 제외한다. 그리고 별점 4와 5를 긍정으로 간주하여 데이터를 처리한다.

전체 리뷰 개수 : 200000

label		document
0	5	배송빠르고 굿
1	2	택배가 엉망이네용 저희집 밑에층에 말도없이 놔두고가고
2	5	아주좋아요 바지 정말 좋아서2개 더 구매했어요 이가격에 대박입니다. 바느질이 조금 ...
3	2	선물용으로 빨리 받아서 전달했어야 하는 상품이었는데 머그컵만 와서 당황했습니다. 전...
4	5	민트색상 예뻐요. 옆 손잡이는 거는 용도로도 사용되네요 ㅎㅎ

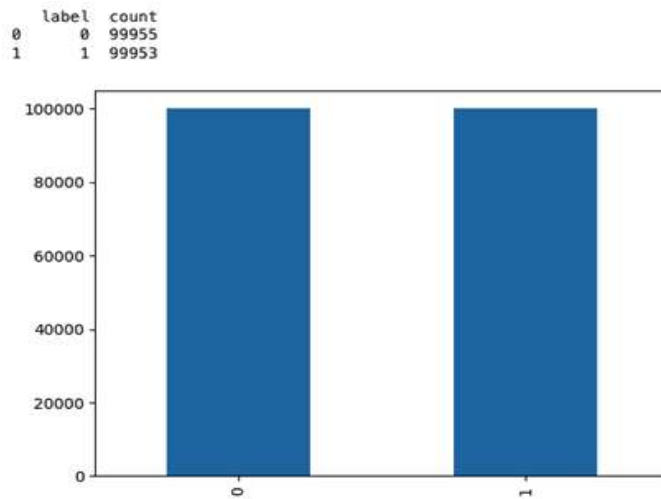
[그림 3-2] Naver Shopping Corpus Original Label Format

label		document
0	1	배송빠르고 굿
1	0	택배가 엉망이네용 저희집 밑에층에 말도없이 놔두고가고
2	1	아주좋아요 바지 정말 좋아서2개 더 구매했어요 이가격에 대박입니다. 바느질이 조금 ...
3	0	선물용으로 빨리 받아서 전달했어야 하는 상품이었는데 머그컵만 와서 당황했습니다. 전...
4	1	민트색상 예뻐요. 옆 손잡이는 거는 용도로도 사용되네요 ㅎㅎ
...
199995	0	장마라그런가!!! 달지않아요
199996	1	다이슨 케이스 구매했어요 다이슨 슈퍼소닉 드라이기 케이스 구매했어요가격 괜찮고 배송...
199997	1	로드샵에서 사는것보다 세배 저렴하네요 ㅏㅏ 자주이용할게요
199998	1	넘이쁘고 세련되보이네요~
199999	1	아직 사용해보지도않았고 다른 제품을 써본적이없어서 잘 모르겠지만 ㅎㅎ 배송은 빨랐습니다

200000 rows × 2 columns

[그림 3-3] Naver Shopping Corpus Preprocessing Label Format

다음으로, 데이터 샘플에서 document열에 중복된 값과 Null 값을 제거한다. 또한, 한글 범위를 고려한 정규 표현식을 사용하여 불필요한 글자를 제거하고, 구두점 및 특수 문자도 제거한다. 리뷰 업로드 특성상 한글뿐만 아니라 영어, 숫자, 특수문자 등으로도 리뷰를 업로드할 수 있으므로 정규 표현식 실행 후 빈 값이 남는 경우가 있을 수 있다. 따라서 공백이나 빈 값이 있는 행은 Null 값으로 대체하고, Null 값이 있는 행의 데이터를 제거한다.



[그림 3-4] Check Distribution of Positive and Negative Labels After Preprocessing Naver Shopping Corpus

네이버 감성 영화 리뷰 데이터셋(Naver sentiment movie corpus)[15]은 감성 분석에 사용되는 대규모 한국어 데이터셋 중 하나로, 한국의 대표적인 포털 사이트인 네이버에서 수집한 영화 리뷰 데이터를 포함하고 있다. 이 데이터셋은 영화 리뷰에 대한 감성(긍정적 또는 부정적)을 분석하기 위해 사용되고 각 리뷰에는 긍정적인 감성 또는 부정적인 감성을 나타내는 레이블이 할당되어 있으며, 이를 통해 감성 분석 알고리즘의 학습과 평가에 사용된다.

20만 건의 전체 리뷰 데이터를 15만 건은 학습용으로, 나머지 5만 건은 테스트용으로 분리했다. 이 데이터는 영문 기준 140자 미만으로 구성되어 있으며 각각의 감성 클래스는 무작위로 분리되었다. 이 중 10만 건은 원래 1에서 4까지의 평점을 가진 부정적인 리뷰이며 나머지 10만 건은 원래 9에서 10의 긍정적인 평점을 받은

리뷰이다. 5에서 8의 평점을 가진 중립적인 리뷰는 제외되었다.

모든 전처리 과정을 거쳤을 때 15만 건의 학습용 데이터는 145,248건으로, 5만 건의 테스트 데이터는 48,779건으로 확인되었다.

```

id    document    label
9976970 아 더빙.. 진짜 짜증나네요 목소리    0
3819312 홈...포스터보고 초딩영화줄...오버연기조차 가법지 않구나    1
10265843 너무재밌었다그래서보는것을추천한다    0
9045019 교도소 이야기구면 ..솔직히 재미는 없다..평점 조정    0
6483659 사이몬페그의 익살스런 연기가 돋보였던 영화!스파이더맨에서 늘어보이긴 했지만 커스틴 던스트가 너무나도 이뻐보였다    1
5403919 막 걸음마 댄 3세부터 초등학교 1학년생인 8살용영화.ㅋㅋㅋ...별반개도 아까움.    0
7797314 원작의 긴장감을 제대로 살려내지 못했다.    0
9443947 별 반개도 아깝다 욕나온다 이응경 김용우 연기생활이몇년인지..정말 발로해도 그것보단 낫겠다 남치.감금만반복만복..이드라마는 가족도없다 연기 못하는사람만모였네    0
7156791 액션이 없는데도 재미 있는 뜻안되는 영화    1

```

```

id    document    label
6270596 굳 ㅋㅋ    1
9274899 GONTOPLCLASSINTHECLUB    0
8544678 뭐야 이 평정들은... 나쁘진 않지만 10점 짜리는 더더욱 아니잖아    0
6825595 지루하지는 않은데 완전 막장임... 돈주고 보기에...    0
6723715 3D만 아니었어도 별 다섯 개 줬을텐데.. 왜 3D로 나와서 제 심기를 불편하게 하죠??    0
7898805 음악이 추가 된, 최고의 음악영화    1
6315043 진정한 쓰레기    0
6097171 마치 미국에서에서 튀어나온듯한 창의력없는 로봇디자인부터가, 고개를 찡جه한다    0
8932678 길수록 개판되가는 중국영화 유치하고 내용없음 품집다 끝남 말도안되는 무기에 유치한cg남무 아 그림다 동사서독같은 영화가 이건 3류아류작이다    0

```

[그림 3-5] Train, Test Original Data from Naver Sentiment Movie Corpus

```

id    document    label
9976970 아 더빙 진짜 짜증나네요 목소리    0
3819312 홈포스터보고 초딩영화줄오버연기조차 가법지 않구나    1
10265843 너무재밌었다그래서보는것을추천한다    0
9045019 교도소 이야기구면 솔직히 재미는 없다평점 조정    0
6483659 사이몬페그의 익살스런 연기가 돋보였던 영화스파이더맨에서 늘어보이긴 했지만 커스틴 던스트가 너무나도 이뻐보였다    1
5403919 막 걸음마 댄 세부터 초등학교 1학년생인 8살용영화별반개도 아까움    0
7797314 원작의 긴장감을 제대로 살려내지 못했다    0
9443947 별 반개도 아깝다 욕나온다 이응경 김용우 연기생활이몇년인지정말 발로해도 그것보단 낫겠다 남치감금만반복만복이드라마는 가족도없다 연기 못하는사람만모였네    0
7156791 액션이 없는데도 재미 있는 뜻안되는 영화    1

```

```

id    document    label
6270596 굳 ㅋㅋ    1
8544678 뭐야 이 평정들은 나쁘진 않지만 점 짜리는 더더욱 아니잖아    0
6825595 지루하지는 않은데 완전 막장임 돈주고 보기에...    0
6723715 만 아니었어도 별 다섯 개 줬을텐데 왜 로 나와서 제 심기를 불편하게 하죠    0
7898805 음악이 추가 된 최고의 음악영화    1
6315043 진정한 쓰레기    0
6097171 마치 미국에서에서 튀어나온듯한 창의력없는 로봇디자인부터가고개를 찡جه한다    0
8932678 길수록 개판되가는 중국영화 유치하고 내용없음 품집다 끝남 말도안되는 무기에 유치한남무 아 그림다 동사서독같은 영화가 이건 류아류작이다    0
6242223 이별의 아픔뒤에 찾아오는 새로운 인연의 기쁨 모든 사람이 그렇지는 않네    1

```

[그림 3-6] Train and Test Preprocessing Data from Naver Sentiment Movie Corpus

네이버 쇼핑 리뷰 데이터셋(Naver Shopping Corpus)[16]은 네이버의 쇼핑 서비스에서 수집된 대규모의 한국어 데이터셋으로 이 데이터는 주로 상품 설명, 리뷰, 평가 등 다양한 형태의 쇼핑 관련 텍스트로 구성되어 있고 상품 카테고리와 브랜드, 소비자들의 의견과 평가를 포함하고 있어 한국어 텍스트에 대한 광범위한 이해와 분석에 도움이 될 데이터를 포함하고 있다.

20만 건의 전체 리뷰 15만 건은 학습용으로, 나머지 5만 건은 테스트용으로 각각의 감성 클래스를 무작위로 분리했다. 이 중 10만 건은 1에서 2까지의 별점을 가진 부정적인 리뷰이며 나머지 10만 건은 4에서 5의 긍정적인 별점을 가진 리뷰이다. 3의 별점을 가진 중립적인 리뷰는 제외되었고 긍정과 부정의 비율이 1:1에 가깝도록 샘플링 되었다.

모든 전처리 과정을 거쳤을 때 15만 건의 학습용 데이터는 149,930건으로, 5만 건의 테스트 데이터는 49,976건으로 확인되었다.

```

5   배공빠르고 굿
2   택배가 엉망이네용 저희집 밑에층에 말도없이 놔두고가고
5   아주좋아요 비지 정말 좋아서2개 더 구매했어요 이가격에 대박입니다, 바느질이 조금 앙성하긴 하지만 편하고 가성비 최고예요.
2   선물용으로 빨리 받아서 전달했어야 하는 상품이었는데 머그컵만 와서 당황했습니다. 전화했더니 바로주신다했지만 배송도 누락되어있네요.. 확인안하고 바로 선물했으면 큰일날뻔했네요..
5   민트색상 예뻐요. 옆 손잡이는 거는 용도로도 사용되네요 ㅎㅎ
2   비추합니다 계란 뒤집을 때 완전 불편해요 ㅋㅋ 코팅도 묻어가고 보기엔 예쁘고 실용적으로 보였는데 생각보다 진짜 별로입니다.
1   주문을 11월6에 시켰는데 11월16일에 배송이 왔네요 ㅎㅎㅎ 여기 회사측과는 전화도 안되고 아무런 연락을 받을수가 없으니 답답하신 분들은 다른곳에서 사시는거 추천드립니다
2   넉넉한 길이로 주문했는데도 안 맞네요 별로예요
2   보풀이 계속 때쳐림 나오다가 지금은 안나네요~
2   110인데 전문속옷브랜드 위생팬티105보다 작은듯해요. 불편해요. 밴딩부분이 다 신축성없는 일반실로 되어있어 빅사이즈임에도 빅사이즈같지않아요. 입고벗을때 편하게 밴딩부분이 늘어나고

```

[그림 3-7] Naver Shopping Corpus Original Data

```

1   배공빠르고 굿
0   택배가 엉망이네용 저희집 밑에층에 말도없이 놔두고가고
1   아주좋아요 비지 정말 좋아서개 더 구매했어요 이가격에 대박입니다 바느질이 조금 앙성하긴 하지만 편하고 가성비 최고예요
0   선물용으로 빨리 받아서 전달했어야 하는 상품이었는데 머그컵만 와서 당황했습니다 전화했더니 바로주신다했지만 배송도 누락되어있네요 확인안하고 바로 선물했으면 큰일날뻔했네요
1   민트색상 예뻐요 옆 손잡이는 거는 용도로도 사용되네요
0   비추합니다 계란 뒤집을 때 완전 불편해요 코팅도 묻어가고 보기엔 예쁘고 실용적으로 보였는데 생각보다 진짜 별로입니다
0   주문을 월에 시켰는데 월일에 배송이 왔네요 여기 회사측과는 전화도 안되고 아무런 연락을 받을수가 없으니 답답하신 분들은 다른곳에서 사시는거 추천드립니다
0   넉넉한 길이로 주문했는데도 안 맞네요 별로예요
0   보풀이 계속 때쳐림 나오다가 지금은 안나네요
0   인데 전문속옷브랜드 위생팬티보다 작은듯해요 불편해요 밴딩부분이 다 신축성없는 일반실로 되어있어 빅사이즈임에도 빅사이즈같지않아요 입고벗을때 편하게 밴딩부분이 늘어나고 입었

```

[그림 3-8] Naver Shopping Corpus Preprocessed Data

스팀의 각종 게임에 달린 리뷰 데이터셋(Steam Corpus)[16]은 컴퓨터 게임 디지털 유통 플랫폼인 Steam에서 발생하는 대규모의 텍스트 데이터 모음이다. 이 데이터는 게임 리뷰, 게임 요약, 사용자 댓글, 포럼 게시물 등 다양한 형태의 텍스트로 구성되어 있다.

이 모든 데이터셋은 일반적으로 감성 분석 모델의 학습 및 테스트를 위해 사용되며, 딥러닝 및 기계 학습 모델을 포함한 다양한 모델에 적용될 수 있다. 이를 통해 모델은 문장 또는 문서의 감성을 올바르게 분류하고, 특정 리뷰나 문장이 긍정적인지 또는 부정적인지를 정확하게 예측할 수 있도록 학습이 가능하다.

10만 건의 전체 리뷰 7만 5천 건은 학습용으로, 나머지 2만 5천 건은 테스트용으로 각각의 감성 클래스를 무작위로 분리했다. 긍정과 부정의 비율이 1:1에 가깝도록 샘플링되었다.

모든 전처리 과정을 거쳤을 때 7만 5천 건의 학습용 데이터는 74,918건으로, 2만 5천 건의 테스트 데이터는 24,972건으로 확인되었다.

```
0 노래가 너무 적음
0 돌겠네 진짜. 황숙아, 어크 공장 그만 돌려라. 죽는다.
1 막노동 체험판 막노동 하는사람인데 장비를 내가 사야돼 뭐지
1 차악!차악!!차악!!! 정말 이래서 왕국을 되찾을 수 있는거야??
1 시간 때우기에 좋음.. 도전과제는 50시간이면 다 깰 수 있어요
1 역시 재미있네요 전작에서 할수 없었던 자유로운 덱 빌딩도 좋네요^^
1 재미있었습니다.
1 은근 쉽지만 은근 어려운 게임
```

[그림 3-9] Steam Corpus Original Data

```
0 노래가 너무 적음
0 돌겠네 진짜 황숙아 어크 공장 그만 돌려라 죽는다
1 막노동 체험판 막노동 하는사람인데 장비를 내가 사야돼 뭐지
1 차악차악차악 정말 이래서 왕국을 되찾을 수 있는거야
1 시간 때우기에 좋음 도전과제는 시간이면 다 깰 수 있어요
1 역시 재미있네요 전작에서 할수 없었던 자유로운 덱 빌딩도 좋네요
1 재미있었습니다
1 은근 쉽지만 은근 어려운 게임
```

[그림 3-10] Steam Corpus Preprocessed Data

이렇게 모든 전처리를 마친 3종의 Naver sentiment movie corpus, Naver shopping corpus, Steam corpus 데이터셋은 학습 데이터셋으로 75%, 테스트 데이터셋으로 25% 구성한다.

[표 3-1] Quantity of Training and Testing Dataset Preprocessing

Dataset	전체	전처리	Train(75%)	Test(25%)
Naver sentiment movie corpus	200,000	194,027 (-5,973)	145,248	48,779
Naver shopping corpus	200,000	199,906 (-94)	149,930	49,976
Steam corpus	100,000	99,890 (-110)	74,918	24,972

C. 감성 분류 Fine-tuning 모델

본 절에서는 효율적인 Fine-tuning을 위한 방법과 사전 학습(Pre-trained) 모델로 사용된 KoELECTRA의 모델 학습 설정에 대해 설명한다.

한국어 뉴스, 위키백과, 나무위키, 모두의 말뭉치 등 약 34GB의 말뭉치로 사전 학습한 KoELECTRA-base-v3 모델을 KoELECTRA 모델 중에 선정하여 사용한다.

Keras에서 제시하는 Fine-tuning의 4가지 방법 중 첫 번째에 해당하는 기준을 보면 데이터셋이 크고, 유사성이 작은 경우는 모델 전체를 학습시키는 것이 낫다고 제시한다. 이유는 데이터셋의 크기가 크기 때문에 충분히 학습이 가능하고 데이터 셋끼리의 유사성이 작다고 해도 모델의 구조와 파라미터들은 여전히 재사용 가능하므로, 재학습을 시켜주는 편이 좋다. 그래서 사전 학습(Pre-trained) 모델로 KoELECTRA 모델을 Fine-tuning 하기 위해 Hugging Face의 Transformers 라이브러리를 사용하여 모델을 설정하였다.

```
Downloading pytorch_model.bin: 100% 452M/452M [00:36<00:00, 20.1MB/s]
Some weights of ElectraForSequenceClassification were not initialized from the model checkpoint at monologg/koelectra-base-v3-discriminator and are newly initialized:
['classifier.out_proj.weight', 'classifier.out_proj.bias', 'classifier.dense.bias', 'classifier.dense.weight']
You should probably TRAIN this model on a down-stream task to be able to use it for predictions and inference.
```

[그림 3-11] Message Using HuggingFace's Transformers Library

이 메시지는 KoELECTRA 모델을 불러오기 위해 Hugging Face의 라이브러리 Transformers를 사용할 때 발생하는 경고 메시지이다. 이 메시지는 모델의 일부 가중치(weights)가 미리 학습된(pre-trained) 체크포인트에서 초기화되지 않고 새로 초기화되었다는 것을 나타낸다.

구체적으로, 'classifier.out_proj.weight', 'classifier.out_proj.bias', 'classifier.dense.bias', 'classifier.dense.weight'의 가중치들이 초기화되지 않았고, 대신에 새로운 초기화가 이루어졌다는 내용이다. 모델의 가중치는 일반적으로 훈련 과정에서 최적화를 통해 업데이트된다. 사전 학습(Pre-trained) 모델의 체크포인트에서 가중치를 가져올 때는 이미 어떤 작업에서 학습된 가중치를 사용하게 된다. 하지만 여기서는 해당 가중치들이 초기화되지 않았고, 대신에 모델을 처음부터 새롭게 초기화했다는 것을 의미한다.

이 경우, 모델이 처음부터 학습을 시작하는 것과 같아진다. 그러므로 이 모델을 사용하려면 명시적으로 다운스트림 작업에서 모델을 학습시켜야 한다. 사전 학습 (Pre-trained) 모델의 가중치를 가져와서 Fine-tuning을 통해 특정 작업에 맞게 조정하는 것이 일반적인 방법이다. 이러한 다운 스트림 작업을 통해 모델을 특정 작업에 맞게 조정하고 성능을 향상시키기 위해 첫 번째로 선택된 데이터셋의 학습을 진행한다. 이렇게 함으로써 저장된 가중치가 새로운 모델에 적용되어 모델의 초기 상태가 설정된다.

두 번째 기준은 데이터셋이 크고, 유사성이 큰 최적의 경우로, 모든 옵션을 선택할 수 있지만, 데이터셋이 유사하기 때문에 이진 분류 작업을 위한 가중치를 새로 학습시키는 것이 최적이다. 첫 번째 기준에 맞춰 Fine-tuning을 진행한 모델을 3가지의 model_weights.pt 파일에 각각 순서를 매겨, 저장된 가중치를 현재 모델에 불러온다. 이를 통해 추론이나 추가 Fine-tuning을 수행할 수 있다.

사전 학습의 트랜스포머 레이어 수는 12이며, 임베딩 사이즈는 768이다. 판별 모델의 히든 사이즈는 768이고, 피드-포워드 네트워크의 히든 사이즈는 3072, 어텐션 헤드는 12로 생성 모델은 각 수치의 1/3인 256, 1024 4를 사용한다.[11] 또한 사용된 KoELECTRA의 Fine-tuning 시 Hyper parameter는 학습 배치 크기로 16, 평가 배치 크기로 16, 에포크 5, 학습률 5e-6, 최대 시퀀스 길이는 제한을 두지 않게 설정하였다.

[표 3-2] Hyper Parameter of KoELECTRA

Train batch size	12
Validation batch size	768
Number of epochs	5
Learning rate	5e-6

Train Batch Size (훈련 배치 크기): 훈련 데이터를 작은 배치로 나누어 신경망을 반복적으로 업데이트하는 데 사용되는 데이터 포인트의 수이다. 배치 크기는 메모리 사용, 훈련 시간 및 모델의 안정성에 영향을 미친다. 일반적으로 큰 배치 크기는 더 높은 계산 효율성을 제공 가능하고 더 작은 배치 크기는 모델의 더 나은 일반화를 도울 수 있다.

Validation Batch Size (검증 배치 크기): 검증 데이터에 대한 성능을 평가할 때

사용되는 배치 크기로 훈련과는 달리, 검증은 모델이 아직 보지 않은 데이터에 대한 성능을 평가하는 데 사용된다. 일반적으로 검증에는 작은 배치 크기가 사용되며, 이는 검증 데이터에 대한 평가를 효율적으로 수행할 수 있게 도와준다.

Number of Epochs (에폭 수): 전체 훈련 데이터셋을 한 번 통과하는 것을 의미하는 수이다. 에폭 수는 전체 데이터셋을 몇 번 반복하여 학습할 것인지를 결정한다. 에폭이 너무 작으면 모델이 충분히 학습되지 않을 수 있고, 에폭이 너무 크면 과적합(overfitting)의 위험이 있다. 일반적으로 훈련 손실이 더 이상 감소하지 않을 때까지 에폭을 늘리거나, 검증 데이터에 대한 성능이 떨어질 때까지 에폭을 늘린다.

Learning Rate (학습률): 가중치를 업데이트할 때 사용되는 스케일 인자로, 모델이 데이터에서 학습하는 속도를 결정한다. 학습률이 너무 작으면 학습이 더디게 진행되거나 수렴하지 않을 수 있고, 너무 크면 발산하여 모델이 수렴하지 않을 수 있다. 일반적으로 학습률은 여러 실험을 통해 조정되며, 옵티마이저와 함께 튜닝되어야 한다.

[표 3-3] Configuration of KoELECTRA-base

Discriminator	
Number of layers	12
Embedding size	768
Hidden size	768
Number of heads	12
Generator	
Number of layers	12
Embedding size	768
Hidden size	256
Number of heads	4

Discriminator:

Number of Layers (레이어 수): 판별기의 신경망 구조에서 사용되는 레이어의 수이다. 일반적으로 많은 레이어를 사용하여 더 복잡한 특징을 학습할 수 있다.

Embedding Size (임베딩 크기): 입력 데이터의 임베딩 차원 크기이다. 임베딩은 입력 데이터를 고차원 벡터로 매핑하는 기술로, 주로 텍스트나 범주형 데이터에서 사용된다.

Hidden Size (은닉층 크기): 각 레이어의 은닉층의 크기를 나타낸다. 이는 모델이 학습하는 매개 변수의 양과 모델의 표현 능력에 영향을 미친다.

Number of Heads (어텐션 헤드 수): 어텐션 메커니즘을 사용하는 경우, 어텐션 헤드의 수를 나타낸다. 어텐션은 모델이 입력의 다양한 부분에 주의를 기울이게 하는데 도움이 된다.

Generator:

Number of Layers (레이어 수): 생성기의 신경망 구조에서 사용되는 레이어의 수이다. 생성기가 얼마나 복잡한 모델인지를 결정한다.

Embedding Size (임베딩 크기): 생성기에서 사용되는 임베딩의 차원 크기이다. 이 임베딩은 생성기의 입력으로 사용된다.

Hidden Size (은닉층 크기): 각 레이어의 은닉층의 크기를 나타낸다. 이는 생성기의 표현 능력에 영향을 미친다.

Number of Heads (어텐션 헤드 수): 어텐션 메커니즘을 사용하는 경우, 어텐션 헤드의 수를 나타낸다. 어텐션은 모델이 입력의 다양한 부분에 주의를 기울이게 하는데 도움이 된다.

이러한 매개 변수는 모델의 구조와 학습 능력에 큰 영향을 미치며, 실제로는 특정 작업 및 데이터에 대한 실험을 통해 최적의 값을 찾는 것이 일반적이다.

여러 번의 Fine-tuning 단계를 거치면서 새로운 레이어가 추가되지는 않는다. 대부분의 경우, Fine-tuning은 모델의 기존 레이어의 가중치를 조절하여 새로운 작업에 적응하도록 하는 과정이다. 이를 통해 모델은 초기에 학습한 지식을 유지하면서 새로운 데이터셋이나 작업에 민감하게 반응하도록 조정된다.

IV. 실험 및 결과

본 장에서는 제안한 텍스트 데이터의 긍·부정 감성을 고려한 데이터 처리 방법과 여러 종류의 데이터셋을 통한 Fine-tuning으로 텍스트 데이터 감성 분류 개선에 도움이 되는 Fine-tuning 방법을 실험한다. 실험에서 사용한 데이터셋에 대하여 설명하고 제안한 방법의 효율성을 입증하기 위해 분류 모델 성능 평가 지표를 활용하여 정확도(Accuracy)를 측정하여 성능을 평가한다.

본 연구에서 실험에 사용한 환경은 [표 4-1]과 같다.

[표 4-1] Development Environment

Division		Contents
HW	CPU	Intel(R) Xeon(R) CPU @ 2.30GHz
	GPU	NVIDIA Tesla V100
	RAM	12 GB
	HDD	200 GB
SW	OS	Ubuntu 22.04.3 LTS
	Program Language	Python

A. 실험 데이터셋

본 절에서는 실험에 사용된 세 개의 리뷰 데이터셋에 대해서 설명한다. 네이버 감성 영화 리뷰 데이터셋(Naver sentiment movie corpus), 네이버 쇼핑 리뷰 데이터셋(Naver Shopping Corpus) 그리고 게임 유통 서비스인 Steam의 각종 게임에 달린 리뷰 데이터셋(Steam Corpus)이다.

네이버 감성 영화 말뭉치(Naver Sentiment Movie Corpus)[15], 네이버 쇼핑 리뷰 데이터셋(Naver Shopping Corpus) 그리고 Steam의 게임에 달린 리뷰 데이터셋(Steam Corpus)[16]은 한국어 감성 분석 연구를 위해 수집된 대규모의 텍스트 데이터 세트이다. 이 데이터는 온라인 영화 리뷰나 온라인 쇼핑몰 리뷰, 게임 사이트 리뷰에서 추출되었으며, 해당 웹사이트에서 제공하는 사용자들이 작성한 리뷰들을 포함하고 있다. 이러한 리뷰들은 영화에 대한 다양한 의견과 감성 표현을 포함하고 있어, 감성 분석 및 자연어 처리 연구에 중요한 자원으로 활용되고 있다.

이러한 말뭉치는 긍정적인 감성, 부정적인 감성, 중립적인 감성 등 다양한 감성적인 표현을 포함하고 있어, 모델의 성능을 평가하고 개선하는 데에 사용된다.

말뭉치의 크기와 다양성은 연구자들이 다양한 실험과 분석을 수행할 수 있도록 도와주며, 특히 한국어 자연어 처리 분야에서 모델의 성능을 향상시키기 위한 기반 자료로 활용된다. 이러한 데이터 세트를 통해 모델은 특정 도메인(영화 리뷰)에서의 감성 분석 작업에서 뛰어난 성과를 달성할 수 있도록 학습되고 이러한 리뷰 말뭉치는 텍스트 데이터의 전처리 및 특성 추출, 감성 어휘 구축 등 다양한 자연어 처리 작업에서 활용될 수 있다. 연구자들은 이 데이터를 활용하여 감성 분석의 성능 향상을 위한 다양한 기술적 접근 방식을 탐구하고 개발할 수 있다.

종합적으로, 리뷰 말뭉치는 한국어 감성 분석 연구를 위한 중요한 자원으로 인정되며, 다양한 연구 및 응용 분야에서 활용되고 있다.

```

$ head ratings_train.txt
id      document      label
9976970 아 더빙.. 진짜 짜증나네요 목소리      0
3819312 흙...포스터보고 초당영화줄....오버연기조차 가볍지 않구나      1
10265843      너무재밌었다그래서보는것을추천한다      0
9045019 교도소 이야기구만 ..솔직히 재미는 없다..평점 조정      0
6483659 사이몬베그의 익살스런 연기가 돋보였던 영화!스파이더맨에서 늙어보이기만 했던 커스틴 던스트가 너무나도 이뻐보였다      1
5403919 막 걸음마 땀 3세부터 초등학교 1학년생인 8살용영화.ㅋㅋㅋ...별반개도 아까움.      0
7797314 원작의 긴장감을 제대로 살려내지못했다.      0
9443947 별 반개도 아깝다 욕나온다 이웅경 길용우 연기생활이몇년인지..정말 발로해도 그것보단 낫겠다 납치,감금만반복반복..이드라마는 가족도없다
7156791 액션이 없는데도 재미 있는 몇안되는 영화      1

```

[그림 4-1] Naver Sentiment Movie Corpus Data Example

```

5      배공빠르고 굿
2      택배가 엉망이네용 저희집 밑에층에 말도없이 놔두고가고
5      아주좋아요 버지 정말 좋아서2개 더 구매했어요 이가격에 대박입니다. 바느질이 조금 엉성하긴 하지만 편하고 가성비 최고예요.
2      선물용으로 빨리 받아서 전달했어야 하는 상품이었는데 머그컵만 와서 당황했습니다. 전화했더니 바로주신디했지만 배송도 누락되어있었네요.. 확인안
5      민트색상 예뻐요. 옆 손잡이는 거는 용도모도 사용되네요 ㅎㅎ
2      비추합니다 계란 뒤집을 때 완전 불편해요 ㅠㅠ 코팅도 묻어나고 보기엔 예쁘고 실용적으로 보였는데 생각보다 진짜 별로입니다.
1      주문을 11월6에 시켰는데 11월16일에 배송이 왔네요 ㅎㅎㅎ 여기 회사측과는 전화도 안되고 아무런 연락을 받을수가 없으니 답답하신 분들은 다른
2      녀석한 길이로 주문했는데도 안 맞네요 별로예요
2      보풀이 계속 때처럼 나오다가 지금은 안나네요~
2      110인데 전문속옷브랜드 위생팬티105보다 작은듯해요. 불편해요. 밴딩부분이 다 신축성없는 일반실로 되어있어 빅사이즈임에도 빅사이즈같지않아요.

```

[그림 4-2] Naver Shopping Corpus Data Example

```

0      노래가 너무 적음
0      돌겠네 진짜. 황숙아, 어크 공장 그만 돌려라. 죽는다.
1      막노동 체험관 막노동 하는사람인데 장비를 내가 사야돼 뭐지
1      차악!차악!!차악!!! 정말 이래서 왕국을 되찾을 수 있는거야??
1      시간 때우기에 좋음.. 도전과제는 50시간이면 다 깰 수 있어요
1      역시 재미있네요 전작에서 할수 없었던 자유로운 덱 빌딩도 좋네요^^
1      재미있었습니다.
1      은근 쉽지만 은근 어려운 게임
0      베즈스다 이 개새끼들아, 시작할 때 체스판 돌아가는거 5분동안 3번 봤더니 ㅈㄴ 백치네 진짜 무한로딩 버그 안쳐고치나 겸하지말라는거나
1      1시간 40분동안 스트레스 다 풀림 ㄹㅇ

```

[그림 4-3] Steam Corpus Data Example

[표 4-2] Naver Sentiment Movie Corpus Summary

데이터 설명	
id	Naver에서 제공한 리뷰 ID
document	실제 리뷰 내용
label	리뷰의 감성 클래스 (0: 부정적, 1: 긍정적)
열은 탭으로 구분(.tsv 형식)	
파일 확장자는 .txt	
전체 200,000개의 리뷰로 구성	
특징	
모든 리뷰는 140자 미만	
각 감성 클래스는 동일하게 샘플링(랜덤 예측 50% 정확도)	
100,000개의 부정적인 리뷰(원래 1-4 등급의 리뷰)	
100,000개의 긍정적인 리뷰(원래 9-10 등급의 리뷰)	
중립적인 리뷰(원래 5-8 등급의 리뷰)는 제외	
파일	
ratings.txt: 모든 200,000개의 리뷰	
ratings_test.txt: 테스트를 위해 50,000개의 리뷰	
ratings_train.txt: 훈련을 위해 150,000개의 리뷰	

아이언맨
영화 · Iron Man · 2008

전체 기본정보 감독/출연 **평점** 무비클립 포토 리뷰 명대사 시리즈 보러가기 >

평점 [관람추이](#)

네티즌 평점

8.93

★★★★★
10,841명 참여

남자 ★ 8.82 여자 ★ 9.13

감상포인트

- OST
- 영상미
- 스토리
- 연기
- 연출

나이별 만족도

10대	★ 9.22
20대	★ 8.74
30대	★ 8.66
40대	★ 8.91
50대	★ 8.62

[관람객/네티즌](#) [평론가](#) [MY](#)

리뷰

별점을 선택해주세요.

★★★★★

감상평을 등록해주세요. 등록

전체 ▾ 최신순 ▾ 스포일러 포함

★★★★★ 10
망한 MCU 에 조의를 표합니다.다시 정주행...갑니다
medi**** · 2023.11.05. 14:24 · 신고

👍 2 🗨 0

★★★★★ 10
마블 영화 정주행 시작했습니다.아이언맨 최고입니다
alwa**** · 2023.11.01. 00:59 · 신고

👍 1 🗨 0

★★★★★ 10
마블 시네마틱 유니버스의 전설작
eunh**** · 2023.10.29. 17:26 · 신고

👍 3 🗨 0

[그림 4-4] Naver Movie Ratings and Reviews Screen

[표 4-3] Naver Sentiment Movie Corpus Summary

데이터 설명			
언어	한국어		
출처	네이버 쇼핑(https://shopping.naver.com/)		
수집 기간	2020.06 ~ 2020.07		
데이터 분포		별점	건수
	긍정(99,963)	5	81,177
		4	18,786
	부정(100,037)	2	63,989
		1	36,048
계		200,000	
특징			
첫번째 필드에는 별점(1~5)			
두번째 필드에는 텍스트가 위치			
긍/부정 분류를 위해 애매한 3점의 텍스트를 제외			
긍정(4~5점)과 부정(1~2점)의 비율이 1:1에 가깝도록 샘플링			



삼성전자 삼성 75인치 TV 스마트 사이니지 UHD 4K LED TV 1등급 BE75C-H
1,199,000원 1,679,000원

N 구매하기

상세정보	리뷰 1,053	Q&A 527	반품/교환정보
------	-----------------	---------	---------

리뷰 1,053건 ✓랭킹순 | 최신순 | 평점 높은순 | 평점 낮은순

[전체보기](#) | [포토/동영상](#) | [스토어PICK](#) | [한달사용리뷰](#) | [재구매](#)
[상품옵션](#) ▾ | [#만족도](#) | [#가격](#) | [#배송](#) | [#화질](#) | [#디자인](#) | [#조작성](#) | [#품질](#) | [#성능](#) | [#화면](#) | [#사용성](#) | [#기능](#) | [#무게](#) | [#제어](#)

리뷰 클렌징 시스템 작동 중입니다. >

★★★★★ 5
 na**** · 23.10.08. | 신고
 삼성 TV 설치 옵션: 삼성 75인치 UHD 4K 스탠드형
 주거형태 30평대위화자네 가정

BEST! 11년만에 티비 바꾸네요. 큰 화면으로 영화보고 싶다는 애들의 성화에 못이겨 검색하다가 장바구니에 한창동안 담아 뒀었는데요. 사이니지 티비라는 것도 처음 들어서 검색해보고 비슷한 가격대의 중소기업 티비와 비교하며 한참을 고민하다가 주문했습니다. 지정일 배송 잘 되었구요. ...

더보기 ▾

판매자 23.10.10. | 신고
 고객님 상품을 구매해주셔서 감사합니다. 좋은 제품과 좋은 서비스로 보답드리겠습니다. 항상 건강하시고 행복하세요 ~ ^^

2

★★★★★ 5
 ma**** · 23.10.30. | 신고
 삼성 TV 설치 옵션: 삼성 75인치 UHD 4K 스탠드형
 주거형태 40평대대가족

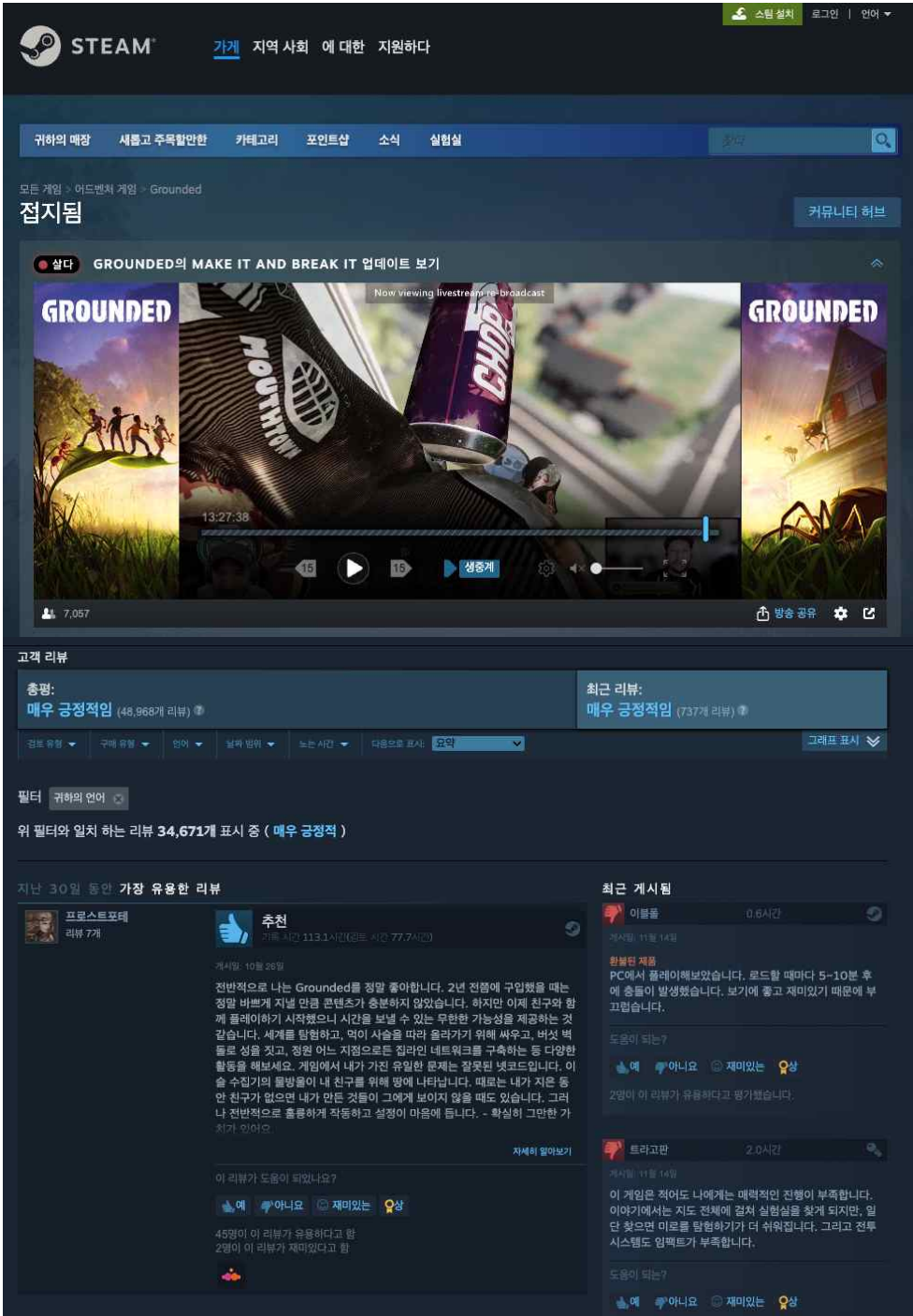
BEST! **재구매** 50명정도가 식당서 사용하기 위해서 75인치를 구매하였는데 아무래도 크기에 신경을 많이 쓰게 되더군요. 75인치를 막상 받고나니 이정도면 딱 되었다 할정도로 크기에 만족했고 특히 베젤도 적다보니 엄청 더 크게 보여서 만족스럽고 거기다가 와이파이 연결이되 되니 넷플릭스나 여러 OTT까지 볼수있어서 더 만족스럽더군요. 두께도 얇고 무게도 무겁지 않아서 장착하는데는 어렵지도 않았고 쉽게 장착하게 되었네요. 무엇보다 티비를 장착하고나 식당이 더 고급스러워지네요^^. 이만한 크기에 가격도 이 정도면 집에 티비도 바꾸고 싶네요. 망설이시는 분들 75인치 이 제...

0

[그림 4-5] Naver Shopping Ratings and Reviews Screen

[표 4-4] Steam Corpus Summary

데이터 설명		
언어	한국어	
출처	Steam (https://store.steampowered.com/)	
수집 기간	2020.05 ~ 2020.06	
데이터 분포	건수	
	긍정	49,996
	부정	50,004
	계	100,000
특징		
게임 커뮤니티 특성 상 비속어 및 은어가 많이 사용됨		
첫번째 필드에는 긍/부정(1=긍정, 0=부정)		
두번째 필드에는 리뷰 텍스트가 위치		
긍정과 부정의 비율이 1:1에 가깝도록 샘플링		



[그림 4-6] Steam Game Ratings and Reviews Screen

B. 실험 평가 방법 및 결과 분석

1. 실험 평가 방법

본 연구의 성능은 데이터가 균등한 경우로 정확도(Accuracy)라는 분류성능평가 지표를 통해 비교하고 제시하는 Fine-tuning 방법의 사용 유무에 따라 모델의 성능 차이를 통해 본 연구의 성능을 평가한다. 또한, 여러 가지 데이터셋을 통해 KoELECTRA 모델을 사용하여 Fine-tuning 해서 만들어진 모델의 비교 분석을 통해 감성분석의 성능을 검증한다. [그림 4-8]은 정확도(Accuracy) 계산에 사용되는 공식이다.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

[그림 4-7] Accuracy Calculation Formula

혼동행렬에서 TP, TN, FP, FN의 의미는 [표 4-5]에서 설명하고 있다.

[표 4-5] Confusion Matrix Terms

TP	True Positive	실제 Positive를 Positive로 예측(정답)
TN	True Negative	실제 Negative를 Negative로 예측(정답)
FP	False Positive	실제 Negative를 Positive로 예측(오답)
FN	False Negative	실제 Positive를 Negative로 예측(오답)

2. 실험 평가 결과 분석

본 절에서는 본 논문에서 제안한 텍스트 데이터의 긍·부정 감성을 고려한 데이터 처리 방법과 여러 종류의 데이터셋을 통한 Fine-tuning으로 텍스트 데이터 감성 분류 개선에 도움이 되는 Fine-tuning 방법에 대해서 성능 평가한 결과를 기술한다.

결과를 기술하는 순서는 총 3개의 데이터셋을 Fine-tuning 한 후 테스트를 진행한 결과를 먼저 확인하고, 이후에 데이터셋을 교차하여 2가지 형태 또는 3가지 형태의 Fine-tuning을 진행한다. 마지막에 학습된 데이터셋의 테스트 데이터셋을 통해 정확도를 확인한 모델을 구분하기 위해 표에서 두꺼운 실선으로 표시를 하였다.

Fine-tuning을 1번 진행한 모델 테스트 결과는 각각 정확도 Movie 90.32, Shopping 93.77, Steam 83.05를 기록했다. 해당 기록을 통해 모델별 성적 구분을 하기 위해 Shopping은 가장 높은 기록으로 High라 표기하고, Movie는 중간으로 Middle로 표기한다. 그리고 Steam은 가장 낮은 기록으로 Low라 표기한다. 이 기록을 기준으로 해당 테스트 데이터셋을 사용한 열에서 다른 Fine-tuning을 진행한 모델의 정확도를 비교해 낮은 경우 파란색으로 표기하였고, 더 높은 정확도인 경우 빨간색으로 표기하였다. 그중에서 밑줄은 해당 테스트 데이터셋으로 평가한 가장 높은 점수를 의미한다.

각 Fine-tuning 모델별 정확도는 기준이 되는 Fine-tuning 1번 진행한 모델의 순서에 맞게 Movie, Shopping, Steam 데이터셋 순서로 실험을 진행한다.

이후에 진행될 실험 결과에는 가독성을 위해 Movie는 ‘MV’로 Shopping은 ‘SP’로 Steam은 ‘ST’로 약어를 지정하여 사용한다. Fine-tuning 1번 진행한 모델을 ‘기준이 되는 모델’이라 하고, 실험에 사용된 3개의 데이터셋의 규모에 따라 규모를 표현하는 약어를 사용한다. 데이터셋 규모가 20만인 Movie와 Shopping 데이터셋은 일반 규모 데이터셋으로 Normal Dataset의 약어로 ‘ND’라 표현하고 상대적으로 규모가 10만으로 적은 Steam 데이터셋은 적은 수량 데이터셋으로 Small Dataset의 약어로 ‘SD’라고 표현 한다.

[표 4-6] Dataset Testing and Accuracy Confirmation

	Fine-tuning Model	Test Dataset 실험 결과			
		Movie	Shopping	Steam	평균
ND	Movie(MV)	90.32	87.85	78.40	85.52
	Shopping(SP)	82.03	93.77	77.19	84.33
SD	Steam(ST)	82.47	87.06	83.05	84.19

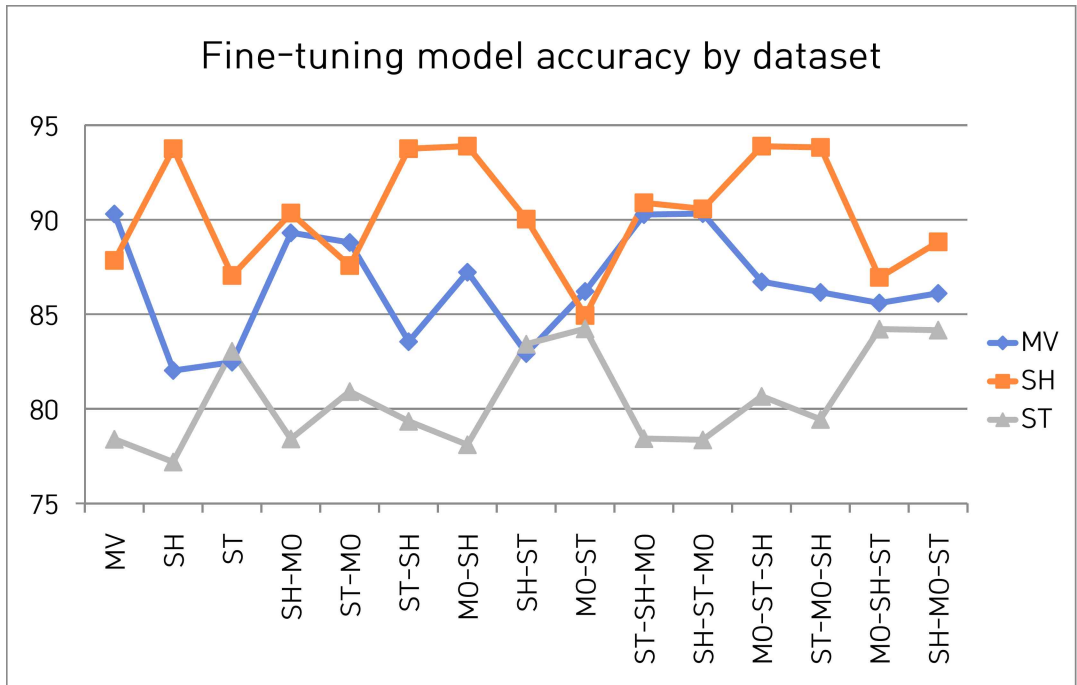
각 테스트 데이터셋의 성적을 기준이 되는 모델 정확도와 비교 결과, Movie 데이터셋 테스트에서는 기준이 되는 Movie 모델의 정확도 90.32 보다 +0.1 포인트 차이를 보인 모델은 SP(ND,High)→ST(SD,Low)→MV(ND,Middle) 순서로 Fine-tuning한 모델로 90.33이라는 정확도를 보였다.

Shopping 데이터셋 테스트는 기준이 되는 Shopping 모델의 정확도 93.77 보다 MV(ND,Middle)→SP(ND,High)와 MV(ND,Middle)→ST(SD,Low)→SP(ND,High) 모델에서 각각 +0.14 포인트와 +0.13 포인트의 높은 정확도 차이를 확인하였고 MV(ND,Middle)→SP(ND,High) 모델이 MV(ND,Middle)→ST(SD,Low)→SP(ND,High) 보다 0.01 포인트의 근소한 차이인 93.91로 가장 높은 정확도를 보였다.

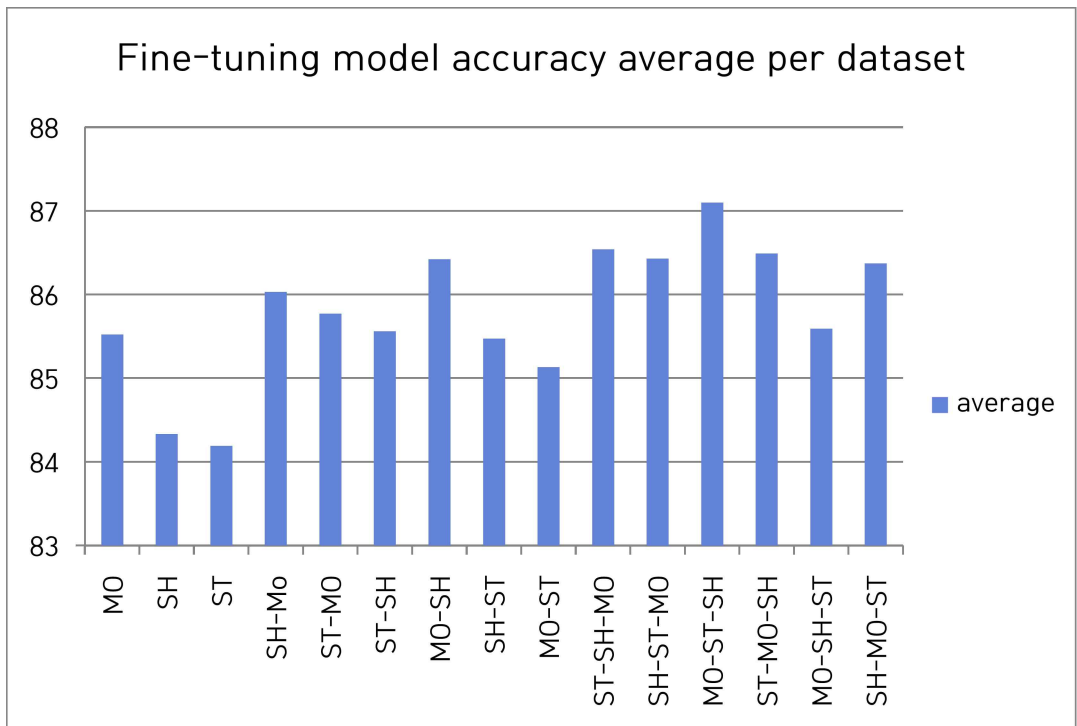
마지막으로 Steam 데이터셋 테스트는 기준이 되는 Steam 모델의 정확도 83.05 보다 MV(ND,Middle)→ST(SD,Low)와 MV(ND,Middle)→SP(ND,High)→ST(SD,Low) 모델에서 각각 +1.19 포인트와 +1.17 포인트의 비교적 높은 차이를 보였다. 이 두 모델은 0.02 포인트의 근소한 차이를 보이고 MV(ND,Middle)→ST(SD,Low) 모델이 84.24로 가장 높은 결과를 보였다.

[표 4-7] Dataset Cross-testing and Accuracy Confirmation

	Fine-tuning Model	Test Dataset 실험 결과			
		Movie	Shopping	Steam	평균
2가지 형태 (ND, SD)	SP → MV	89.32	90.36	78.41	86.03
	ST → MV	88.80	87.58	80.92	85.77
	ST → SP	83.56	93.77	79.35	85.56
	MV → SP	87.24	93.91	78.10	86.42
	SP → ST	82.94	90.05	83.42	85.47
	MV → ST	86.22	84.94	84.24	85.13
3가지 형태 (ND, SD, ND)	ST → SP → MV	90.28	90.91	78.43	86.54
	SP → ST → MV	90.33	90.59	78.36	86.43
	MV → ST → SP	86.73	93.90	80.67	87.10
	ST → MV → SP	86.17	93.84	79.45	86.49
	MV → SP → ST	85.60	86.96	84.22	85.59
	SP → MV → ST	86.12	88.84	84.16	86.37



[그림 4-8] Model Accuracy Fine-tuning Graph by Dataset



[그림 4-9] Model Accuracy Fine-tuning Graph per Dataset

위 실험은 데이터셋이 총 3개이고 데이터셋의 규모가 ND, ND, SD인 경우를 바탕으로 확인한 결과이다. 이러한 결과를 토대로 데이터셋이 2가지 형태인 경우는 데이터 규모가 비슷한 ND + ND 인 경우와 규모가 크고 작은 ND + SD 인 경우가 있을 수 있고 데이터셋이 3가지 형태인 경우는 데이터 규모가 비슷한 ND 2개와 SD 1개를 조합한 6가지 경우가 있다.

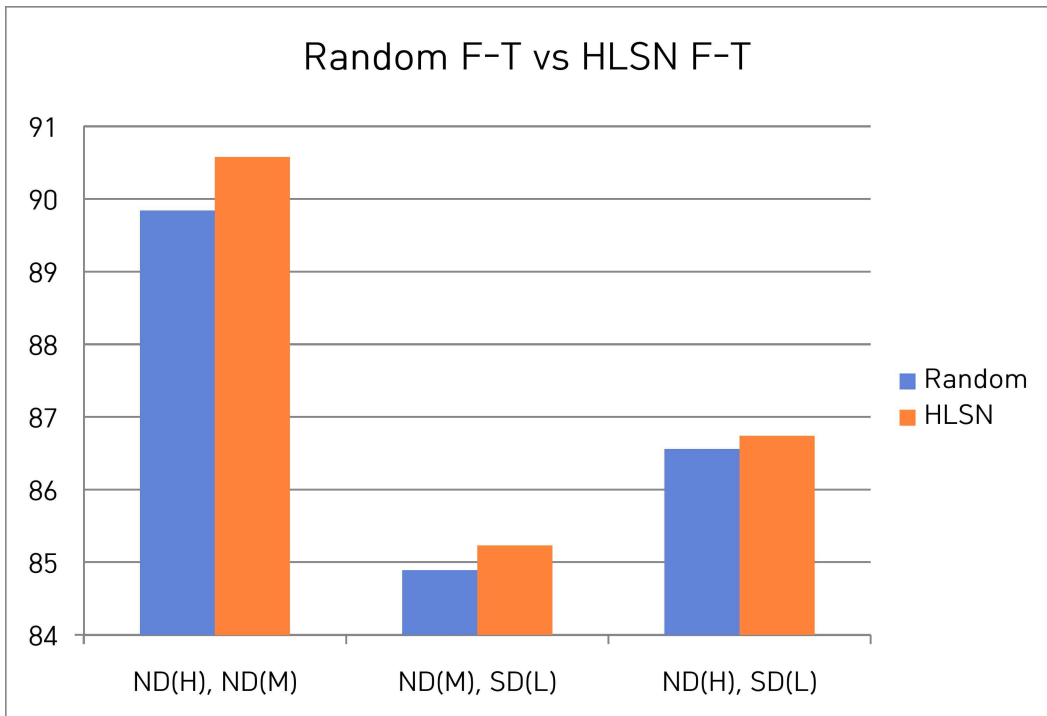
데이터셋이 2가지 형태인 경우에 효율적인 Fine-tuning을 위해 먼저, 각 데이터셋이 가지고 있는 학습 정확도를 확인한다. ND 인 MV는 90.32, ND 인 SP는 93.77 그리고 SD 인 ST은 83.05를 기록했다. 첫 번째 ND + ND의 경우 마지막에 학습되는 데이터셋은, 1가지 형태인 경우에 정확도가 가장 좋았던 모델의 데이터셋으로 Fine-tuning을 진행하는 것이 그러지 않은 방식에 비해서 학습 데이터 테스트 정확도가 평균 +0.74 포인트 높게 나와 같은 크기의 데이터셋이라도 정확도에 도움을 주는 것으로 확인하였다. 두 번째 ND + SD의 경우 ND를 우선적으로 Fine-tuning 해주고 SD를 마지막에 학습해주는 방법이 SD를 먼저 Fine-tuning 하는 방법보다, ND(Middle) + SD(Low)에서는 +0.34 포인트, ND(High) + SD(Low)는 +0.18 포인트 더 높은 학습 정확도 평균을 보여줬다. 확인결과 2가지 형태 중에 같은 데이터셋 규모를 가지고 있는 ND + ND의 경우 HLSN(High Last Small Next) Fine-tuning 방식 통해 SD가 있다면 SD는 첫 번째 학습이 아닌 다음에 학습, SD가 없다면 High Score를 마지막에 학습하는 경우 Random Fine-tuning 모델에 비해 높은 정확도 평균을 확인하였다.

[표 4-8] Summary of 3 Dataset Characteristics

Fine-tuning Model	데이터셋 규모	모델 정확도
Shopping(SP)	20만건(ND)	93.77(High)
Movie(MV)	20만건(ND)	90.32(Middle)
Steam(ST)	10만건(SD)	83.05(Low)

[표 4-9] Dataset 2 Cross-testing and Accuracy Confirmation

	Fine-tuning Model	Test Dataset 실험 결과			
		MV	SP	ST	평균
2가지 형태 (ND + ND)	SP(ND/High) → MV(ND/Middle)	89.32	90.36	-	89.84
	MV(ND/Middle) → SP(ND/High)	87.24	93.91	-	90.58
2가지 형태 (ND + SD)	ST(SD/Low) → MV(ND/Middle)	88.80	-	80.92	84.89
	MV(ND/Middle) → ST(SD/Low)	86.22	-	84.24	85.23
	ST(SD/Low) → SP(ND/High)	-	93.77	79.35	86.56
	SP(ND/High) → ST(SD/Low)	-	90.05	83.42	86.74



[그림 4-10] Comparison of Accuracy Between Random F-T and HLSN F-T in Two Types

마지막으로 데이터셋이 3가지 형태인 경우에 효율적인 Fine-tuning을 위해서는 기본이 되는 데이터셋의 각각의 학습 결과를 먼저 확인하고 가장 좋은 정확도를 가지는 데이터셋을 맨 마지막에 두어 학습 순서를 잡는다. 남은 2개의 데이터셋은 위에서 확인된 실험 결과와 같이 ND + SD로 Fine-tuning을 진행한다.

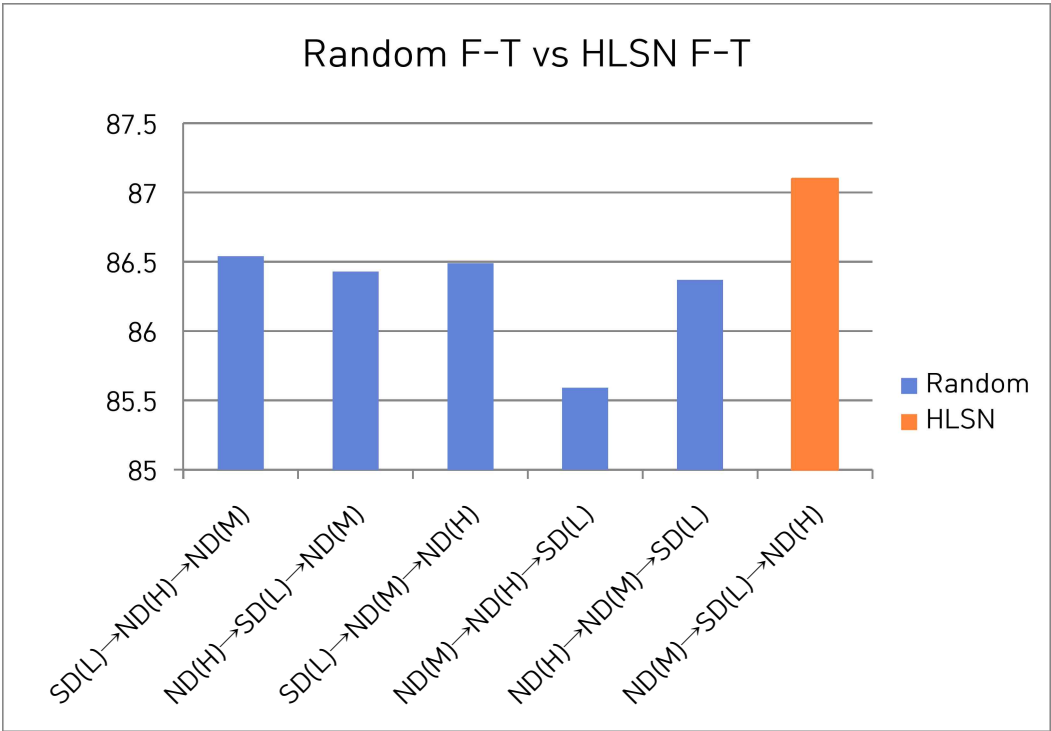
확인 결과 효율적인 HLSN(High Last Small Next) Fine-tuning 방식을 사용하여 3가지 형태의 경우에도 SD가 있다면 첫 번째 학습이 아닌 다음에 학습 SD가 자리를 잡았다면 High Score는 마지막에 학습하는 모델과 Random Fine-tuning을 진행한 모델의 가장 큰 정확도 평균 차이는 1.51 포인트이고 나머지 다른 모델의 경우 0.56~0.73 포인트 차이를 보임을 확인하였다.

[표 4-8] Summary of 3 dataset characteristics

Fine-tuning Model	데이터셋 규모	모델 정확도
Shopping(SP)	20만건(ND)	93.77 (High)
Movie(MV)	20만건(ND)	90.32 (Middle)
Steam(ST)	10만건(SD)	83.05 (Low)

[표 4-10] Dataset 3 Cross-testing and Accuracy Confirmation

	Fine-tuning Model	Test Dataset 실험 결과			
		MV	SP	ST	평균
3가지 형태 (ND + SD + ND)	ST(SD/Low) → SP(ND/High) → MV(ND/Middle)	90.28	90.91	78.43	86.54
	SP(ND/High) → ST(SD/Low) → MV(ND/Middle)	90.33	90.59	78.36	86.43
	MV(ND/Middle) → ST(SD/Low) → SP(ND/High)	86.73	93.90	80.67	87.10
	ST(SD/Low) → MV(ND/Middle) → SP(ND/High)	86.17	93.84	79.45	86.49
	MV(ND/Middle) → SP(ND/High) → ST(SD/Low)	85.60	86.96	84.22	85.59
	SP(ND/High) → MV(ND/Middle) → ST(SD/Low)	86.12	88.84	84.16	86.37



[그림 4-11] Comparison of Accuracy Between Random F-T and HLSN F-T in Three Types

V. 결론 및 향후 연구

본 논문에서는 여러 개의 각기 주제가 다른 감성 데이터셋이 주어졌을 때 더 정확한 감성분석을 위해 라벨 데이터의 특징을 고려한 효율적인 Fine-tuning 학습 방법을 제안하였다. 데이터 셋으로는 Naver에서 확인 가능한 ‘naver movie review’, ‘naver shopping reivew’ 데이터와 Steam에서 확인 가능한 ‘steam game review’ 데이터를 통해 한국어 감성 정보가 포함된 데이터 셋을 사용하였다. 해당 감성 클래스는 전처리 이후에도 긍·부정 데이터의 수가 고르게 균형을 이룬다.

주제가 다른 3가지 리뷰 데이터셋을 전처리하고 해당 데이터셋에 대해 KoELECTRA 모델을 이용한 감성 분류 Fine-tuning을 진행한다. 실험은 분류 성능을 높이기 위해 데이터셋을 다양하게 Fine-tuning 하였으며, Fine-tuning Task에 맞게 첫 Fine-tuning의 경우 사전 학습(Pre-trained) 모델인 KoELECTRA를 전체 학습하고 이후 학습부터는 학습된 모델을 불러와 배치 크기와 학습률을 고정하여 같은 학습 조건일 때 최적의 모델인 HLSN Fine-tuning 방법을 통한 모델을 얻는다.

이후, 사용자의 텍스트 리뷰 데이터에 대한 감성 분류 시 해당 모델을 사용하고자 한다. 실험을 통해 확인된 바와 같이, 제안하는 HLSN Fine-tuning 방법은 3가지 데이터셋의 경우만 있다는 한계가 있다.

후속 연구에서는, 추가로 확보될 데이터셋의 클래스별 불균형 문제를 고려하여 가중치를 다르게 설정하거나, Fine-tuning을 통한 동결된 학습 조건이 아닌 다양한 학습 조건을 두어 이미 학습된 모델이 추가되는 데이터만을 어떤 방식으로 학습을 하는 것이 효율적인지에 대한 방법을 연구하고자 한다. 확장 가능성이 있는 논문에서 학습한 KoELECTRA HLSN Fine-tuning 감성 분류 모델을 바탕으로 사용자의 리뷰에 보다 정확한 긍·부정 감성 분류뿐만 아니라 더 나아가, 리뷰 데이터 크롤링과 전처리를 자동화 하여 일정 데이터 규모 간격으로 효율적인 Fine-tuning 까지도 스스로 학습하는 자동화 모델을 개발 가능할 것으로 기대된다.

참고 문헌

- [1] Cho Dong-wook, Kim Bong-hyun, Lee Se-hwan, “Comparison and Analysis of Speech Signals for Emotion Recognition”, The 25th Korean Information Processing Society Spring Academic Presentation Conference Proceedings Vol. 13, No. 1, pp. 533-536, 2006. 5.
- [2] Ryu Yoon-ji, Lim so-yeon, Kim Sang-wook, “Context-aware Smartphone Inferring Activity according to User Emotion”, Proceedings of the 34th Autumn Conference of the Korean Information Processing Society Vol. 17, No. 2, 2010. 11.
- [3] Park Ho-yeon, Kim Kyung-jae, “Recommender system using BERT sentiment analysis”, Journal of intelligence and information systems v.27 no.2, pp. 1-15, 2021. 3.
- [4] Ji Hoon, Lee Chung-heon, Park Moon-kyu, Ahn Young-joon, Lee Dong-hoon, “Recognition of the emotional state through the EEG”, Korea Information and Communication Society 2015 Spring Conference, pp.958-961, 2015.
- [5] Hong Cheon-ho, Cho Young-ho, “Cyber Threats Analysis of AI Voice Recognition-based Services with Automatic Speaker Verification”, Journal of Internet Computing and Services Vol. 22, No. 6, pp. 33-40, 2021.
- [6] Heo Kyung-moo, Kang so-min, “Facial expression recognition technology”, Journal of Control Robot Systems v.20, no.2, pp. 39-45, 2014.
- [7] Seol Yong-so, Kim Dong-joo, Kim Han-woo, Park Jung-jung, “Emotion Recognition from Natural Language Text Using Predicate Logic Form”, Proceedings of the Summer Conference of the Korean Computer Information Society Vol. 18, No. 2, pp. 411-412, 2010. 7.
- [8] Kim Byung-so, Lee Yong-hoon, Lee Jong-hyuk, “Unsupervised Semantic Role Labeling for Korean Adverbial Case”, Journal of the Society for Information Science. Journal of KIISE. Software and Applications v.34, No.2,

pp. 112–122, 2007. 2.

- [9] Hong Sung-jun, “Korean Spoken Sentences Emotion Classification using Recurrent - Convolutional Neural Network”, master, Kyungpook National University, 2017.
- [10] Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, NAACL-HLT, 2019.
- [11] Kevin Clark, Minh-Thang Luong, Quoc V. Le, Christopher D. Manning, “ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators”, ICLR, May 2020.
- [12] Ha Jae-ryong, “A study on sentiment labeling techniques for news article titles”, master, Korea University, 2022. 2.
- [13] Park Sang-min, Na Cheol-won, Choi Min-seong, Da-hee Lee, On Hospital, “Building a sentiment dictionary based on Bi-LSTM”, Kunsan National University, Intelligence Information Research Vol. 24, No. 4, pp. 219–240, 2018. 12.
- [14] “KoELECTRA”, Github, <https://github.com/monologg/KoELECTRA> (accessed. Oct, 13, 2023)
- [15] “nsmc”, Github, <https://github.com/e9t/nsmc/> (accessed, Oct, 13, 2023)
- [16] “말뭉치 모음”, <https://github.com/bab2min/corpus/tree/master/sentiment> (accessed, Oct, 13, 2023)