

# 효과적인 애스팩트 마이닝을 위한 다중 레이블 분류접근법

## Multi-Label Classification Approach to Effective Aspect-Mining

원종윤 (Jong Yoon Won) 성균관대학교 경영대학 박사과정

이건창 (Kun Chang Lee) 성균관대학교 경영대학 글로벌경영학과/삼성융합의과학원 융합의과학과 교수, 교신저자

### 요 약

최근의 감성분류 연구는 출력변수가 하나인 단일레이블 분류방법을 사용한 연구가 많다. 특히, 이러한 연구는 하나의 극성 값(긍정, 부정)만을 찾는 연구가 많다. 그러나 한 문장 안에는 다중적인 의미가 내포되어 있다. 그 중에서도 감정과 오피니언이 이러한 특징을 갖는다. 본 논문은 두 가지 연구목적에 제시한다. 첫째, 한 문장 안에 다양한 토픽(주제 또는 애스팩트)이 있다는 사실을 기반으로, 해당 문장을 각 애스팩트 별로 감성을 분류하는 애스팩트 마이닝을 수행한다. 둘째, 두개 이상의 종속변수(출력 값)를 한 번에 분석하는 다중레이블 분류방법을 적용한다. 이에 본 연구는 감성분류의 연구가 단일분류기에 의해서만 이루어진 연구를 개선하고자 다중레이블 분류방법에 의한 애스팩트 마이닝을 수행하고자 한다. 이와 같은 연구목적에 달성하기 위해 국내 뮤지컬 데이터를 수집하였다. 분석결과 문장 안에 있는 다양한 애스팩트별 감성을 추출하였고, 유의한 결과를 얻었다.

**키워드 :** 데이터마이닝, 애스팩트 마이닝, 감성분석, 다중레이블 분류방법, 머신러닝

## I. 서 론

정보기술의 급격한 발전은 소비자의 행동 양식을 변화시킨다(Yadollahi *et al.*, 2017). 인터넷과 스마트폰 등 정보 접근에 대한 용이성이 과거보다 확연하게 향상됨에 따라, 소비자가 습득한 정보는 소비 과정에서 적극적으로 활용된다. 특히, 과거에는 구전으로 표현되던 입소문(Word of Mouth)이 리뷰 등의 형태로 온라인에서 문자로 표현되면서, 소비자는 타 소비자에게 제품과 서비스를 광고하는 자발적 전도자 역할을 하고 있다(Lee *et al.*, 2017).

이처럼, 온라인 리뷰는 소비 과정에서 상당한 영향력을 행사하고 있으며, 기업 디지털마케팅에서 제품과 서비스에 대한 온라인 리뷰 분석의 중요성은 더욱 높아지고 있다(Alqaryouti *et al.*, 2019; Jena, 2020; Liu, 2019; Wang *et al.*, 2014).

온라인 리뷰를 분석하는 연구 방법으로는 토픽 분석(Topic Analysis), 네트워크분석(Network Analysis), 감성분석(Sentiment Analysis) 등 다양하다(Almars *et al.*, 2019; Celardo and Everett, 2019). 이 중 감성분석은 리뷰 문장 속에 내포되어 있는 긍정 또는 부정 감성을 확인할 수 있도록 하기 때문에 최근 많은

관심을 받고 있는 연구분야이다. 기존의 감성분석 연구는 온라인 리뷰의 극성 값(긍정, 부정)을 분류하는데 초점이 맞춰져 있다(Chen and Huang, 2019). 그러나 한 리뷰 안에는 주제(에스펙트)별로 다양한 감성을 내포하고 있다. 이는 에스펙트 별로 다른 감성을 갖는다는 것을 의미하다(Alqaryouti et al., 2019; Tan et al., 2019). 즉, 하나의 온라인 리뷰를 전체적으로 긍정, 부정의 글로 구분할 수 있다. 그렇지만 온라인 리뷰는 에스펙트 별로 다른 감성을 갖고 있기 때문에 리뷰를 전체적으로 평가하는 기존 분류 방법은 정확한 감성분석이라 말하기는 어렵다.

에스펙트에 따라 감성을 분류하는 것을 에스펙트 마이닝(Asspect-Mining)이라고 한다. 에스펙트 마이닝은 에스펙트에 따른 각기 다른 감성을 분류하기 때문에 기존의 감성분석보다 정밀하게 분석한다. 그렇기 때문에 에스펙트 마이닝은 감성분석의 진보된 분석방법으로 여겨진다(Yang et al., 2019). 에스펙트 마이닝의 사례는 에어비앤비의 경영전략에서 찾아볼 수 있다. 에어비앤비는 “집주인은 끔찍하고, 집은 조금 더러웠지만, 전망이 아름답고 지하철역이 집 근처에 있었다.”라는 문장에서 ‘집주인’, ‘청결상태’, ‘전망’, ‘대중교통’의 에스펙트를 추출하고, 이에 따른 감성을 분류하였다. 분류 결과를 활용하여 소비자의 선호를 파악하고, 집주인에게 최적의 가격을 추천하였으며, 소비자에게는 가격 대신 온라인 리뷰를 홍보하였다. 그 결과, 회사의 매출은 크게 상승하였고, 경쟁사보다 시장 우위를 선점하였다(Lawani et al., 2019).

일반적으로 기존의 많은 감성분석 연구는 단일 레이블 분류 방법(Single-Label Classification)을 적용하였다. 단일 레이블 분류 방법은 출력값을 하나로 지정하여 분류하는데, 에스펙트 마이닝은 에스펙트별로 출력값을 갖기 때문에 출력값이 두 개 이상이다. 단일 레이블 분류 방법을 적용한 에스펙트 마이닝은 감성을 에스펙트 별로 따로 분석하여 결과를 도출해야 함을 의미한다. 따라서 단일

레이블 분류 방법을 적용한 에스펙트 마이닝은 문제를 해결하는 것처럼 보이지만, 에스펙트에 따라 다른 출력값들의 상관성을 고려하지 못하며 출력값에 따라 각각 분석하기 때문에 한계점으로 지적된다. 이와 달리 다중 레이블 분류(Multilabel Classification)방법은  $n$ 개 이상의 출력값을 한 번에 분류할 수 있다. 그렇기 때문에 데이터마이닝 연구 분야에서 다중 레이블 분류 방법은 단일 레이블 분류 방법의 진보된 형태로 평가받는다(Zhang and Zhou, 2013). 결과적으로, 에스펙트 마이닝을 수행하는데 있어 단일 레이블 분류 방법보다 다중 레이블 분류 방법이 효과적인 것이다.

더욱이 분류 연구의 최종 목적은 데이터의 자동 분류이다. 에스펙트 마이닝의 연구목적 또한 이와 같다. 자동 분류 시스템은 새로 생성되는 온라인 리뷰는 에스펙트에 따라 자동적으로 감성이 분류되어야 한다. 이는 기업이 온라인 리뷰를 기반으로 제품과 시장에 대해 실시간으로 관찰하고 변화를 파악할 수 있음을 의미한다. 그리고 자동 분류 시스템은 단일 레이블 분류방법을 적용하는 것보다 다중 레이블 분류방법을 적용하는 것이 더욱 효과적인 것이다. 이에 본 연구는 첫째, 한글 감성분류 연구 중 에스펙트 마이닝 연구가 현저히 부족하다는 점, 둘째, 다중 레이블 분류방법을 적용한 한글 감성분류 연구가 이뤄지지 않았다는 연구현황을 고려하여 다음과 같은 연구 질문을 하였다.

- RQ1. 다중레이블 분류방법을 적용한 한글 에스펙트 마이닝을 수행하였을 때, 어느 다중레이블 분류기가 가장 우수한 결과를 보이는가?
- RQ2. 다중레이블 분류를 적용한 한글 에스펙트 마이닝을 수행하였을 때, 어느 단어 벡터 방식(Bag of Words와 Word2vec)이 더 적합한가?
- RQ3. 다중레이블 분류를 적용한 한글 에스펙트 마이닝을 수행할 때 속성선택(Feature Selection)을 적용할 경우 향상된 결과를 얻을 수 있는가?

## II. 관련문헌

### 2.1 감성분석

감성분석 연구는 소셜미디어의 발달과 함께 사람들의 관심을 받는 연구 분야이다. 특히, 페이스북, 트위터, 트립어드바저와 같은 플랫폼의 발달은 사람들이 자신의 의견을 표현하고 공유하는 것을 하나의 문화로 만들었다(Shibuya and Tanaka, 2018). 감성분석 연구는 일반적으로 머신러닝 연구 방법(Machine-Learning Method)과 감성사전 기반 연구 방법(Lexicon-based Method)으로 구분한다(Ali *et al.*, 2019). 머신러닝 방법은 다양한 머신러닝 분류기를 적용하여 데이터를 학습하고, 모델을 구축한다. 그리고 도출된 모델을 이용하여 새로운 데이터로 모델의 우수성을 검증하는 방법이다(Krouska *et al.*, 2017).

감성사전 기반의 연구 방법은 긍정 단어와 부정 단어로 구분되어 있는 어휘 사전을 기반으로 한다. 여러 감성사전이 연구에 적용되는데, 영문 감성사전 연구에는 SentiWordNet이 많이 사용된다(Santosh *et al.*, 2016). 한편 한글 감성사전의 경우 KOSAC, KNU 등이 연구에 사용되지만 한글은 동음이의어가 많고, 복잡한 문장 구조 때문에 어느 감성사전이 연구자와 실무진들에게 인정받는지는 논하기는 어렵다. 감성사전은 적용되는 영역(Domain)에 따라 분석의 오차가 생길 수 있다(Lee *et al.*, 2016). 예를 들면 ‘슬픔’이라는 단어는 일반적으로 부정적인 의미이라고 여겨진다. 그러나 책, 영화, 공연과 같은 예술 영역에서는 오히려 긍정적 의미로 사용되기도 한다. 즉, 감성사전의 단어들은 영역에 따라 의미가 다르기 때문에 감성사전 기반의 감성연구의 문제점으로 지적되기도 한다. 이러한 문제를 해결하고자 감성사전 기반 연구는 도메인에 따라 다양한 감성사전을 구축하는 연구로 발전하고 있다(Jha *et al.*, 2018).

최근 감성분석 연구는 텍스트에서 의미 있는 정보를 선택적으로 추출하는 연구로 발전하고 있다.

이 같은 감성분석 방법이 적극적으로 적용되는 분야 중 하나가 지능형 교통체계(Intelligent transportation system: ITS) 분야이다. 소셜 미디어의 사용자들은 교통과 관련된 다양한 문제(교통 체증, 사고, 도로 공사 등)에 대한 일상을 다른 사람들과 공유한다. 이를 본 다른 사용자들은 해당 게시물에 텍스트 또는 이모티콘으로 댓글을 남겨 정보를 교환하고, 유용성을 평가한다(Das *et al.*, 2016). 이와 같은 사람들의 행동양식에 따라 ITS 분야는 대중들이 소셜 미디어에 공유하는 교통에 대한 감성적인 정보를 취득하고 이를 분석하여, 교통에 관련된 최신 정보를 추출하였다. ITS 분야의 연구는 이러한 최신 정보를 활용하여 실시간 교통정보를 파악하고, 사용자들에게 최신 교통 정보를 실시간으로 알려주는 시스템을 개발하고 있다(Abberley *et al.*, 2017; Das *et al.*, 2016). 이렇듯 감성분석 연구는 문서 전체의 긍, 부정을 분류하는 연구뿐만 아니라 연구자가 원하는 구체적이고, 의미 있는 정보를 선택적으로 추출하고, 이러한 정보를 활용하여 사람들에게 구체적이고 최신성을 갖춘 정보를 제공하는 연구로 발전하고 있다.

과거의 국내 감성분류 연구는 수집하기 용이한 영문 데이터를 활용하여 연구하였거나, 한글 데이터인 경우 특정 영역(영화, 트위터 등)으로 편중되었다. 이는 한글과 영어는 어순이 다르기 때문에 다른 방법으로 연구되어야 한다는 점과 데이터의 영역에 따라 감성적 어휘가 다른 의미로 사용된다는 점에서 국내 감성분석 연구의 한계로 지적되었다(남길임, 조은경, 2017). 그렇지만 최근 국내 감성분석 연구는 다양한 영역의 확장과 다각적인 분석방법으로 발전하였다. 이는 한글 감성분석의 정밀함과 다양함을 시사한다. 최근의 한글 감성분류 연구는 다음과 같다.

이영준, 윤보현(2019)은 금융통화위원회의 의사록 데이터를 분석하여 미래 중앙은행의 금리 의사결정을 예측하는 모형을 제시하였다. 금융통화위원회의 의사록 데이터를 토픽모델링 기법으로 감성을 분류하였고, 테일러 준칙(Taylor rule)을 적

용하여 중앙은행의 금리 의사결정을 예측하는 모델의 정확도를 높였다. 이는 기존의 텍스트 데이터에만 의존하는 감성분류 연구에 경제학적 이론을 접목하여, 시장을 보다 정교하게 예측하였다는 점에서 의의가 있다.

정민지 등(2019)은 모바일 간편결제서비스(삼성페이)에 대한 소비자 반응을 분석하였다. 트위터와 온라인 커뮤니티로부터 수집한 텍스트 데이터를 분석 대상으로 선정하였으며, 빈도분석, 군집분석, 네트워크 분석의 다각적인 방법을 적용하여 제품에 대한 소비자의 반응을 분석하였다. 상품에 대한 소비자 오피니언을 분류하고, 다양한 분석 기법 간의 비교로 소비자의 의견을 심층적으로 파악하였다.

박현정 등(2018)은 딥러닝을 적용한 감성분석 연구를 하였다. 딥러닝(Deep-Learning) 모델을 적용하기 위한 기초 연구로서, 교착어이면서 동음이의어의 비중이 높은 한글도 딥러닝 기법을 적용하였을 때 만족할만한 정확도를 보인다는 것을 발견하였다. 특히, 한글에 대한 특징을 파악하고 딥러닝을 적용한 한글 감성분석 방안을 제시한다는 측면에서 의의가 있다.

요약하면, 한글 감성분석 연구는 최신 연구기법인 딥러닝을 적용하여 감성분류의 정확도를 높

이고자 했다. 그리고 기존연구에서 검증된 빈도분석, 군집분석, 네트워크 분석을 이용하여 다양한 시각에서의 비교는 한글 감성분석의 정확성을 높였다. 마지막으로 감성분석과 경제학적 이론을 결합한 예측 모델을 개발하여 감성분석의 학문적 활용도를 높였다. 지금까지 논의된 바와 같이 다양한 한글 감성분석 연구는 많지만 본 연구와 같이 온라인 리뷰를 애스팩트에 따라 구분하여 분석한 애스팩트 마이닝 연구는 매우 부족하다. 더욱이 본 연구와 같이 다중레이블 분류방법을 적용한 한글 감성분석은 거의 존재하지 않는다.

## 2.2 애스팩트 마이닝

애스팩트 마이닝은 크게 두 가지로 구분할 수 있는데, 애스팩트 탐지(Aspect Detecting)와 제품의 애스팩트에 대한 오피니언 탐지(Detecting the Opinion about the aspect of the product)이다(Garcia-Pablos *et al.*, 2018). 애스팩트 탐지는 애스팩트 마이닝의 초기 연구이다. 일반적으로 토픽모델링(Topic modeling)을 활용하여 애스팩트를 추출하고 이를 분류하는 것이다(Qiu *et al.*, 2011). 토픽모델링에서 많이 적용되는 추출 모델은 pLSA와 LDA가 가장 보편적으로 사용된다. 두 모델은 모두 문서와 단어

〈표 1〉 최근 국내 감성분석 연구

저자	내용	기법	데이터 세트
이영준, 윤보현(2019)	토픽모델링 기법을 적용한 감성분석으로 중앙은행의 금리에 대한 예측모형을 제시	토픽모델링	금융통화위원회 의사록
정민지 등(2019)	모바일 간편결제를 이용하는 소비자의 반응을 분석하고, 반응간 연관성을 파악함.	빈도 분석, 군집 분석, 네트워크 분석	트위터
박현정 등(2019)	형태소 벡터를 도출하고, 맞춤법이나 띄어쓰기 등 전처리를 하고 난후에 딥러닝을 적용하고 나서 감성분석의 분류 정확도가 향상됨	딥러닝 (Convolutional Neural Network)	온라인 쇼핑몰 리뷰
박재수, 이재수(2019)	일간지 및 경제신문의 부동산기사에 대해 감성 분석을 하였으며, 그 결과 서울시 집값을 예측하는 모형을 제시	토픽모델링	신문기사
김다예, 이영인(2018)	Word2Vec 알고리즘을 통해 단어를 수치적으로 표현하고 이를 기반으로 한국종합주가지수의 종가를 예측하는 모델을 제시	군집 분석	네이버 뉴스

사이의 잠재변수를 통해 문서에서 의미있는 토픽을 분석한다. 애스팩트 탐지의 초기연구는 긍정, 부정의 오피니언을 구분하는 방법이 아닌 해답에 대한 ‘예’ 또는 ‘아니오’를 말한다. 즉, 애스팩트 탐지는 텍스트에 대한 감성을 분류하는 것이 아닌, 의견이나 감정의 존재여부만을 판단한다.

이와 달리 제품의 애스팩트를 구분하고 그에 따른 오피니언을 분류하는 연구는 실무적 활용도가 높다. 이러한 애스팩트에 대한 오피니언 탐지 연구는 애스팩트에 따라 각각의 오피니언을 구분한다. 특히 온라인 리뷰를 작성한 사람의 의사를 애스팩트에 따라 오피니언을 분류하기 때문에 기존의 감성분석보다 소비자의 의견을 상세히 알 수 있다. 기존 애스팩트에 대한 오피니언 탐지 연구는 다음과 같다.

소비자가 호텔 리뷰를 작성할 때에 호텔 전반에 대한 평가보다는 세부항목별로 별점을 준다. 예를 들어 식사의 질, 위생상태, 부대시설의 우수함, 교통의 편리함 등에 대해서 각각 별점을 준다. 한편, Marcheggiani *et al.*(2014)은 이러한 호텔 리뷰의 특성을 활용하여 리뷰의 애스팩트에 따른 긍정, 부정의 오피니언을 분류하였다. 분류를 위해 직접 개발한 다중레이블 분류 알고리즘인 CRF(Conditional Random Field)을 적용하여 애스팩트 마이닝을 하였다.

호텔 온라인 리뷰를 작성하는 소비자들이 모든 항목을 체크하며 항목별로 리뷰를 작성하지 않는 경우가 많다. 그렇기 때문에 작성된 일부의 리뷰를 바탕으로 작성되지 않은 부분의 오피니언을 예측하는 연구가 있다. Li *et al.*(2016)은 소비자가 체크하지 않은 별점에 대해 작성된 다른 내용을 분석하여 누락된 별점을 예측하는 모델을 제시하였다. 이를 위해 EM-알고리즘을 적용하여 누락된 별점을 예측하고 다른 연구와 차별성을 나타내었다.

Khan *et al.*(2016)은 스마트폰의 운영체제를 애스팩트로 구분하였다. 구글의 안드로이드와 애플의 IOS로 나누어 소비자 의견에 대해 애스팩트 마이닝을 수행하였다. 분석을 위해 유튜브(Youtube)에 게시된 스마트폰 리뷰 영상을 본 시청자가 남긴 댓글을 수집하였다. 그리고 리뷰에 나타난 긍정적, 부정적, 중립적 의견을 분류하여 스마트폰 운영체제에 따라 애스팩트 마이닝을 하고 이를 비교하여 분석하였다.

이와 달리 Rathan *et al.*(2018)은 애스팩트를 스마트폰의 제조사와 스마트폰의 기능에 따라 구분하였다. 스마트폰의 기능에 따른 애스팩트는 배터리, 카메라, 디스플레이 등으로 나누었다. 또한, 감성을 예측하는데 이모티콘(Emoji)을 함께 분석하였다. 이모티콘은 텍스트와 달리 함축적 의미를

〈표 2〉 최근 애스팩트 마이닝 연구

저자	내용	데이터 세트	다중레이블 분류
Marcheggian <i>et al.</i> (2014)	CRF 알고리즘을 적용하여 호텔 리뷰에 대한 감성분석함.	온라인 리뷰 (Trip Advisor)	Yes
Li <i>et al.</i> (2016)	EM-알고리즘을 통해 호텔의 평점을 예측함.	온라인 리뷰 (Trip Advisor, Dianping)	No
Khan <i>et al.</i> (2016)	스마트폰 운영체제에 대한 비교 분석	온라인 리뷰 (Youtube)	Yes
Rathan <i>et al.</i> (2018)	스마트폰의 제조사와 기능에 따른 분석	온라인 리뷰 (Amazon)	No
Le <i>et al.</i> (2015)	온라인 쇼핑에서 10개의 물품을 선택하여 분석.	온라인 리뷰 (CJ-mall)	No
정폴잎 등(2019)	스마트폰의 기능에 따라 네트워크 분석을 통한 감성분석.	온라인 리뷰 (Amazon)	No

가지기 때문에 텍스트와 함께 쓰일 때 감성을 더 명확하게 분류할 수 있다.

애스팩트 마이닝 연구는 컴퓨터 공학분야에서 주로 다루고 있으며, 단일 레이블 알고리즘을 다중레이블 알고리즘으로 변환하여 적용하는 연구 또는 새로운 알고리즘을 개발하는 연구가 많다. 이와 달리 본 연구는 한글 애스팩트 마이닝 연구가 현저히 부족하다는 점 그리고 한글 감성분석 연구에는 다중 레이블 알고리즘을 적용한 연구가 없는 현황을 고려하여 한글 애스팩트 마이닝에 다중 레이블을 적용하였을 때, 가장 적합한 모델에 대해 논의하고자 한다.

### III. 연구방법

#### 3.1 다중 레이블 분류

$x = R^d$ 은 d차원의 인스턴스 공간을 나타내고,  $y = y_1, y_2, \dots, y_q$  로 q개의 레이블 공간을 나타낸다. 다중 레이블 분류 방법은  $D = \{(x_i, Y_i | 1 \leq i \leq m)\}$ 의 학습 데이터로부터  $h: x \rightarrow 2^y$ 의 함수를 학습하는 것이다. 각각의 객체  $(x_i, Y_i)$ 에 대해  $x_i \in x$ 는 d차원 특징인 인스턴스  $(x_{i1}, x_{i2}, \dots, x_{id})$ 를 나타내고,  $Y_i \subseteq y$ 는  $x_i$ 와 관련된 레이블의 집

합이다. 다중 레이블 분류방법은 이러한 데이터 구조를 통해  $x$ 에 대한 적절한 레이블을  $h(x) \subseteq y$ 을 통해 찾아내는 것을 의미한다(Zhang and Zhou, 2013).

앞에서 이미 언급한 바와 같이, 다중 레이블 분류 방법은 객체에 단일값이 아닌 복수개의 n개의 레이블이 있는 자료를 분류한다. 그리고 해당 n개의 레이블은 벡터로 처리가 가능하다. 대개 해당 레이블은 이진 값으로 표기된다. 단일 레이블 분류 방법과 달리 다중 레이블 분류는 복수개의 여러 레이블을 처리할 수가 있어서 애스팩트 마이닝 연구에 필요하다.

<그림 1>은 이진 분류 데이터셋, 단일 레이블 멀티클래스 데이터 세트, 그리고 다중 레이블 데이터 세트를 알기 쉽게 표현하였다. 이진 분류 데이터셋은 다중 레이블 분류의 입장에서 가장 기본적인 데이터 구조이다. 단일 레이블 멀티클래스 (Multi class) 역시 다중 레이블 분류 데이터 세트와는 차이가 있음을 알 수 있다. 단일 레이블 멀티클래스 구조의 데이터 세트의 경우 이를 적절하게 사용할 경우 다중 레이블 문제를 일부 해결할 수 있는 것처럼 보인다. 그렇지만 다중 레이블에서의 레이블 수가 증가할 때마다 해당 레이블 변수는 2n개의 출력 값으로 표현된다. 따라서, 멀티클래

Binary Data-set						Multiclass Data-set					
$X_1$	$X_2$	...	$X_{f-1}$	$X_f$	class	$X_1$	$X_2$	...	$X_{f-1}$	$X_f$	class
$X_{11}$	$X_{12}$	...	$X_{1f-1}$	$X_{1f}$	1	$X_{11}$	$X_{12}$	...	$X_{1f-1}$	$X_{1f}$	C7
$X_{21}$	$X_{22}$	...	$X_{2f-1}$	$X_{2f}$	0	$X_{21}$	$X_{22}$	...	$X_{2f-1}$	$X_{2f}$	C1
...					1	...					C6
$X_{n1}$	$X_{n2}$	...	$X_{nf-1}$	$X_{nf}$	0	$X_{n1}$	$X_{n2}$	...	$X_{nf-1}$	$X_{nf}$	C3

Multilabel Data-set							
$X_1$	$X_2$	...	$X_{f-1}$	$X_f$	$Y_1$	...	$Y_k$
$X_{11}$	$X_{12}$	...	$X_{1f-1}$	$X_{1f}$	0	1	0
$X_{21}$	$X_{22}$	...	$X_{2f-1}$	$X_{2f}$	1	0	0
...					0	0	1
$X_{n1}$	$X_{n2}$	...	$X_{nf-1}$	$X_{nf}$	1	0	1

<그림 1> 데이터 속성에 따른 특징

스 분류는 단일 레이블 분류방법의 레이블 수를 전부 반영할 수 없다(Zhou *et al.*, 2012).

### 3.1.1 다중 레이블 분류 알고리즘

다중 레이블 분류 알고리즘은 단일 레이블 분류 방법과 유사한 점이 많다. 그렇지만 단일 레이블 분류방법의 알고리즘을 다중 레이블 분류방법에 적용하기 위해서는 알고리즘의 변용이 필요하다. 그렇기 때문에 다중 레이블 분류방법 알고리즘은 크게 문제변환방식과 알고리즘 적용방식으로 나눈다. 본 연구는 기존문헌에서 우수하다고 평가된 7가지 알고리즘을 적용하여 모델을 평가하였다. BR(Binary Relevance)은 레이블 값이 이진값으로 되어 있는 간단한 형태의 다중 레이블 문제를 각각의 레이블마다 독립적으로 평가하는 방식이다. CC(Classifier Chain)는 이진 분류기의 가장 잘 알려진 다중 레이블 앙상블로, 데이터에서 많은 레이블들이 무작위 순서로 선택되어 k개의 모델을 훈련시켜 평가하는 방식이다(Read *et al.*, 2009). LP(Label Powerset)는 레이블 조합을 기반으로 하는 가장 기본적인 변환 방법이다(Read *et al.*, 2011). LP는 일반적으로 레이블 종속성 정보를 고려하므로 BR과 비교하여 평가한다. RAKEL(Random-k-Lablesets)은 레이블의 무작위 부분집합에 대해 다계층 분류기를 학습하는 방법이다. CLR(Calibrated Label Ranking)은 이진분류기 앙상블이다(Furnkranz *et al.*, 2008). 레이블 쌍의 관련성을 구별하는 방식인 OVO(One Versus One) 접근 방식으로 가상 레이블을 만들어, 최종 분류기를 보정하고 관련성이 없는 레이블과 관련이 있는 레이블을 분리하여 평가한다(Hüllermeier *et al.*, 2008). BP-MLL(Back-Propagation)은 인공신경망 모델을 기반으로 만들어진 알고리즘이며, ML-kNN은 k-최근접 이웃(kNN)방법을 전환하여 다중레이블에 적용시킨 것이다(Huang *et al.*, 2017).

### 3.1.2 다중 레이블 분류 평가지표

다중 레이블 평가지표는 크게 예제기반 평가지

표와 레이블 기반 평가지표로 분류된다. 성과를 측정하기 위한 다양한 지표가 있지만, 일반적으로 다중 레이블 분류방법 연구 분야에서 주로 사용되는 성과측정치는 Hamming Loss(HL)와 Ranking Loss(RL)가 보편적이다. 이에 본 연구는 HL와 RL을 주평가지표로 성과를 측정하며, 보조지표로 One Error(OE)를 사용한다.

HL은 잘못 분류된 인스턴스와 레이블쌍의 비율을 평가한다. 즉, 관련 레이블이 누락되었거나 무의미한 것을 예측하는 것이다. i번째 인스턴스의 실제 레이블 집합인  $Y_i$ 와 예측된  $Z_i$ 사이의 대칭차집합(Symmetric difference set)의 차이를 토대로 계산된다(Zhang and Zhou, 2007).

$$HammingLoss = \frac{1}{n} \frac{1}{k} \sum_{i=1}^n |Y_i \Delta Z_i|$$

RL은 인스턴스에 대해 관련성이 있는 레이블과 관련성이 없는 레이블의 모든 조합을 취하고, 분류기 예측에서 관련성이 없는 레이블이 관련성이 있는 레이블보다 이전에 얼마나 나타났는지 계산한다(Pereira *et al.*, 2018). 계산은 인스턴스의 관련이 있는 레이블과 관련성이 없는 레이블의 곱으로 나눈 값을 정규화 한 다음에 평가된 인스턴스의 수를 평균한 것이다. 이에 RL은 측정값이 낮을 수록 분류가 더 잘 수행되었다고 판단한다.

$$RankingLoss = \frac{1}{n} \sum_{i=1}^n \frac{1}{|Y_i| \times |\bar{Y}_i|}$$

$$|y_a, y_b : rank(x_i, y_a) > rank(x_i, y_b), (y_a, y_b) \in Y_i \times \bar{Y}_i|$$

OE는 다중 레이블 분류기에 대한 신뢰할 수 있는 레이블의 긍정오류(False positive)인 경우의 비율을 나타낸다(Zhang and Zhou, 2013). 분류기에 의해 주어진 예측값의 최상위 레이블이 집합에 속하지 않는 경우, 그 다음에 나오는 식을 1로 반환한다. 이러한 예상치 못한 누적 횟수가 누적되고,

이 값을 평균화한다.

$$OneError = \frac{1}{n} \sum_{i=1}^n \|\text{argmax} \langle \text{rank}(x_i, y) \notin Y_i \rangle\|$$

### 3.1.3 다중 레이블 분류 속성선택

속성은 분류시에 고려하여야 할 독립변수 또는 설명변수를 의미한다. 따라서, 주어진 분류문제의 속성의 개수가 많아지면 그만큼 분류시 요구되는 컴퓨팅의 부담이 커지고 시간도 오래 걸린다. 속성의 개수를 차원 (dimensionality) 이라고도 하는데 차원은 이진(Binary)분류 문제, 멀티클래스(Multi-class)분류뿐만 아니라 다중 레이블(Multi-label)분류를 수행하는데도 많은 영향을 준다(Liu *et al.*, 2010; Liuc and Motoda, 2012). 따라서 속성중 분류의 정확성을 올리고 분류시간을 줄이는데에 기여하는 속성만을 골라내는 속성선택 문제는 중요하다. 차원의 수를 크게 줄일 수 있기 때문이다. 속성선택이 성공적으로 이뤄지면 분류 정확성에 큰 관련성이 없는 속성을 제거할 수 있어서 분류시간을 줄이고 분류 정확성을 크게 올릴 수 있다.

기존 속성선택 방법은 LDA(Linear Discriminant Analysis)나 CCA(Canonical Correlation Analysis) 등과 같이 통계적 기법을 기반으로 한다. 속성선택 방법은 크게 두 가지 방법으로 구분한다. 각 기능이나 기능의 하위집합을 Information gain 그리고 Relief와 같은 필터(Filter) 방법에서 포괄적인 평가모형을 사용하는 방식과 기본 학습 알고리즘에 의해 제공되는 결과에 따라서 모형을 평가하는 래퍼(Wrapper)방식이 있다.

속성선택은 다중 레이블 분류에도 적용되는데, 다중 레이블 분류는 두 개 이상의 레이블이 있기 때문에 단일 레이블 분류 속성선택 보다 더 복잡하다. 다중 레이블 분류 속성선택은 알고리즘 분류방식과 마찬가지로 문제변환방식과 알고리즘 적용방식으로 구분하여 설명할 수 있다. 본 연구는 문제변환방식과 알고리즘 적용방식에서 효과적이라 판단되는 속성방법을 각각 하나씩 선택하

여 적용하였다. 선택한 속성선택 알고리즘은 문제변환방식의 ReliefF(Spolaor *et al.*, 2013)와 알고리즘 적용방식의 SCLS(Scalable criterion for large set)(Lee and Kim, 2017)이다.

ReliefF 방법은 단일 레이블 분류방법의 속성선택방법으로 많이 사용되고 있다. ReliefF는 타겟 레이블과 인스턴스 간의 최근접 이웃 점수를 산출하여 점수가 높은 속성을 선택하는 방식이다. 본 연구는 이를 적용하기 위해 문제변환방식인 BR 기법과 결합하여 이를 적용하였다(Spolaor *et al.*, 2013). 반면, 알고리즘 적용방식인 SCLS 속성선택 방법은 각 레이블에 대한 속성들의 의존성에 따른 계산 값을 포함하면서, 계산의 반복성을 배제하는 방법이다. 이는 기존의 다른 속성선택 방법보다 효율성을 극대화하였다고 평가받는다. 결과 SCLS는 특정 조건하의 관련성을 보다 정확하게 평가하는 관련성 평가 프로세스를 구축하여 학습의 정확도를 높이는 방식이다(Lee and Kim, 2017).

## 3.2 연구모형

### 3.2.1 워드임베딩

텍스트 마이닝은 비정량적 데이터인 텍스트를 정량적 데이터로 변환하는 것이 필수적이다. 이러한 과정을 워드 임베딩이라 한다. 최근 자연어 처리방법이 다양해지면서, 워드임베딩 방법도 다양하다. 이러한 연구는 머신러닝 기법을 적용하여 워드임베딩 간의 정확도 비교로 우수성을 입증한다. 본 연구는 기존의 많은 논문에서 검증된 두 가지 모델을 통해 애스팩트 마이닝을 위한 워드임베딩 모델의 효과성을 입증하고자 한다.

첫째로, 기존의 연구에서 우수하고 안정적이라고 여겨지는 Bag-of-words(BOW). 둘째로, 인공지능 스피커에도 적용되는 방식으로 딥 러닝 방식의 Word2Vec이다. 두 가지 모델은 기존 문헌에서도 많은 비교연구가 이뤄졌으며, 어느 모델이 더 우수한지 검증되지 않았다. 뿐만 아니라 다중 레이블을 적용하여 두 모델 간 비교를 한 연구는 부족하다.



BOW는 단어를 행렬공간에 표현하는 방법으로 특정 단어가 문서에 나타나는 출현빈도를 산출하는데 유용하다. 예를 들어보자. 단어를 W로 표시하고 특정 문서에 나타난 단어 빈도가 W1, W1, W2, W2, W2, W3로 나타났다면 이는 첫 번째 단어가 2번, 두 번째 단어가 3번, 세 번째 단어가 1번 나타났음을 의미한다. 따라서,  $X = \{2, 3, 1\}$ 로 표현한다. 이를 활용하여 행렬공간에 표현하는 방법도 다양하다. 그 중 해당 문장에서 단어의 빈도수를 나타내는 Term Frequency(TF), 단어가 출현했는지 여부를 나타내는 Term Occurrence(TO), 문장 내에서 단어의 중요성을 나타낼 수 있는 Term Frequency-Inverse Document Frequency(TF-IDF) 등이 있다(Ngo-Ye and Sinha, 2014). 각 벡터의 값은 문서의 특수한 정보를 추측할 수 있으며, 연구목적에 따라 다양하게 적용된다. 본 연구는 문서 내에서 단어의 중요성을 나타낼 수 있는 TF-IDF 방법을 적용하였다.

Word2Vec은 Mikolov *et al.*(2013)이 개발하여 널리 사용되는 워드임베딩 알고리즘이다. 비슷한 위치에 등장하는 단어와 해당 위치에 등장하는 다른 단어들간 유사성을 감안한 분포가설에 근거한 방법이다. 비슷한 위치에 등장하는 단어들은 그 의미도 비슷할 것이라는 전제를 갖고 그에 대한 값을 산출한다. Word2vec은 연속 말뭉치인 CBOW(Continuous Bag of Words)와 스킵그램인 Skip-Gram 두 가지 방식으로 구분된다. CBOW는 타겟단어 주변에 있는 단어를 확인하면서 타겟단어를 분석하는 방식이다. 스킵그램 방식은 타겟단어를 먼저 확인을 하고 해당 단어 주변에 있는 여타 단어를 추정하는 방법이다. 본 연구는 두 가지 방식 중 많은 기존연구에서 그 우수성이 입증된 스킵그램 방식을 적용한다.

### 3.2.2 데이터 세트

한글 애스팩트 마이닝을 수행하기 위해 본 연구는 국내 뮤지컬 공연에 대한 구매자 온라인 리뷰 수집하였다. 뮤지컬의 경우 대다수의 사람이 예매

를 통해 공연을 관람한다. 뮤지컬은 연극, 영화와 같은 공연물보다 관람에 따른 가격이 상대적으로 높다. 따라서, 관객들은 뮤지컬을 관람하기 전에 관람 후기를 면밀히 찾아본다. 생활필수품이나 전자제품과 달리 뮤지컬과 같은 여가생활에 따른 소비는 구매자가 더 많은 정보를 찾아보고 결정하며 다른 상품들에 비해 더 상세하고 다양하게 리뷰를 작성하는 것을 확인하였다.

본 연구는 국내 뮤지컬 예매 플랫폼 중 시장점유율 80%인 1위 업체인 인터파크 플랫폼을 선정하였으며, 2013~2017년도에 공연된 국내 뮤지컬 중 8편을 선정하여, 총 3만 건의 리뷰를 수집하였다. 수집한 리뷰는 실제 공연의 관람자여부도 알 수 있어 객관성이 높은 리뷰라고 판단하였다. 이 중 본 연구의 분석목적에 가장 적합한 데이터를 선택하기 위해 전처리한 결과, 리뷰 1,500건을 선정하였다. 이는 애스팩트 마이닝을 원활하게 수행하기 위함이다.

많은 연구에서 적용하는 기본감정 6개(행복감, 공포, 놀람, 슬픔, 화남, 혐오)를 본 연구의 애스팩트로 선정하여 적용하였다(Giatsoglou *et al.*, 2017). 각 리뷰 어떠한 감정이 내포되어 있는지 탐지하여 적용하였으며, 선정된 1,500건의 리뷰는 최소 두 가지 이상의 감정이 내포되어 있는 데이터로 다중 레이블 분석에 적합하였다.

## IV. 연구결과

본 연구의 다중 레이블 분류방법을 적용한 한글 애스팩트 마이닝 결과는 <표 3>을 통해 볼 수 있다. RQ1은 다중 레이블을 적용한 한글 애스팩트 마이닝을 수행하였을 때 어떤 워드 임베딩 방식이 보다 효과적인지 검증하였다. 텍스트 분석에서 사용되는 워드임베딩 방법에 따른 두 가지 모델은 BOW와 Word2vec로 구분하였는데, 첫째, HL을 기준으로 한 평가는 다음과 같다. 분류기마다 조금씩 상이하지만, 본 연구에서 적용된 7가지의 분류 알고리즘 중에 ML-kNN을 제외한 BR, CC, RAKEL,

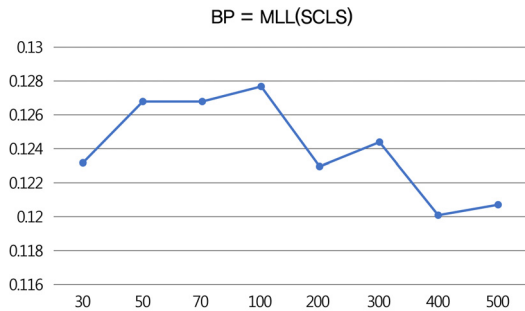
CLR, LP, BP-MLL 등 알고리즘의 경우, BOW 모델에서 가장 낮은 오차값을 보인다. 둘째, RL지표의 경우, HL와 마찬가지로 ML-kNN을 제외한 모든 알고리즘이 BOW 모델의 우수성을 지지한다. 마지막으로, OE의 경우, CLR와 ML-kNN을 제외한 나머지 알고리즘이 BOW 모델이 한글 애스펙트 마이닝을 효과적으로 수행한다는 결론을 나타낸다.

다음으로, RQ2는 한글 애스펙트 마이닝에 보다 적합한 모델을 구분하기 위해 어느 다중 레이블 분류기가 우수한지 검증하였다. 그 결과 첫째, HL을 기준으로, BP-MLL이 BOW 모델에서 0.1171로 가장 우수한 분류기로 평가되며, 그 다음으로 ML-kNN이 Word2Vec에서 두 번째로 우수한 0.1309의 값으로 적합한 알고리즘이다. 둘째로, RL는 BP-MLL이 BOW 모델에서 0.0503으로 다른 분류기보다 우수한 결과를 제시한다. 그 다음으로 우수한 분류기는 ML-kNN으로 Word2Vec모델에서 0.0695의 결과를 나타낸다. 마지막으로 OE를 기준으로 평가하였을 때, 위의 결과들과 마찬가지로 BP-MLL이 BOW 모델에서 0.0554로 다른 분류기보다 우수한 결과를 제시한다. 이러한 결과는 각각의 알고리즘에 대한 성능의 차이를 볼 수 있으며, 크게는 문제변환방식보다 알고리즘 적용방식이 더 높은 우수성을 나타냄을 보여주었다.

마지막으로 RQ3은 효과적인 애스펙트 마이닝을 위한 속성선택방법의 우수성을 검증하고자 하였다. 적용한 결과, 속성선택 방법을 적용한 분석은 앞의 결과를 기반으로 한다. 속성선택 방법은

워드임베딩 기법에서 우수하다고 평가하는 BOW 모델의 데이터를 기반으로 속성을 추출하였으며, 다중레이블 분류기중 앞서 성능이 우수하다고 평가한 BP-MLL와 ML-kNN을 기반으로 속성선택 방법에 따른 결과를 비교하였고, 평가지표는 HL를 기준으로 하였다.

최근 연구물인 알고리즘 적용방식 속성선택 방법인 SCLS는 자동적으로 최적의 속성을 추출하는 방법이 아닌, 추출하고자 하는 속성 수를 입력하여 속성을 추출한다. 이에 본 연구는 <표 4>와 같이 BOW의 1,000개의 속성을 골고루 8가지의 수로 추출하였다. 이는 전체적인 그래프를 통해 현저히 오류가 낮은 지점을 찾기 위함이다. <그림 2>에서와 같이 SCLS는 400개의 속성을 선택한 지점에서 오류가 낮아지는 것을 확인하였다. 반면, 문제변환방식인 ReliefF 방법의 속성선택의 결과는 최적의 속성 값을 추출하는 방식으로, 하나의 결과만을 제시한다.



<그림 2> SCLS 속성선택

<표 3> BOW 모델과 Word2vec 모델의 결과

	Bag-of-Word			Word2Vec		
	Hamming Loss	Ranking Loss	One Error	Hamming Loss	Ranking Loss	One Error
BR	0.163	0.177	0.235	0.218	0.222	0.316
CC	0.166	0.133	0.178	0.213	0.174	0.239
RAkEL	0.214	0.112	0.147	0.346	0.123	0.138
CLR	0.279	0.099	0.11	0.311	0.113	0.125
LP	0.186	0.161	0.233	0.221	0.18	0.26
BP-MLL	0.1171	0.0503	0.0554	0.1597	0.1211	0.0926
ML-kNN	0.1437	0.0919	0.0918	0.1309	0.0695	0.0724

〈표 4〉 속성선택 결과

	SCLS								RF_BR
	30	50	70	100	200	300	400	500	
BP-MLL	0.1232	0.1268	0.1268	0.1277	0.123	0.1244	<b>0.1201</b>	0.1207	<b>0.1185</b>
ML-kNN	0.1323	0.1364	0.1341	0.1357	0.1537	0.1536	<b>0.1519</b>	0.1507	<b>0.1174</b>

〈표 4〉의 속성선택의 결과를 요약하면 다음과 같다. 알고리즘 적용방식인 SCLS에서 400개의 속성을 추출한 값과 Relief 방식의 값을 추출한 속성을 비교하였을 때, BP-MLL의 값은 각각 0.1201, 0.1185로 Relief 방식이 보다 최적의 속성선택 성능을 보인다. 또한, ML-kNN을 기준으로 각 0.1519, 0.1174인 것을 확인할 수 있다. 이는 Relief 방식이 ML-kNN에서도 더 효과적인 것을 확인할 수 있다.

## V. 결 론

본 연구결과를 요약하면 다음과 같다. 첫째, 다중 레이블 분류기를 이용하여 한글 애스팩트 마이닝을 할 때, 워드임베딩 방식은 Word2Vec보다는 BOW가 더 효과적이다. 둘째, 다중 레이블 분류기를 이용하여 한글 애스팩트 마이닝을 할 때, 다중 레이블 분류기는 문제변환방식보다 알고리즘 적용 방식을 적용할 때 더 높은 성능을 보인다. 셋째, 다중 레이블 속성선택 방법을 적용하였을 때, 문제변환방식의 알고리즘 적용 방식이 효과적이다.

기존의 애스팩트 마이닝은 단일 레이블 분류 방법을 적용한 연구가 많다. 이는 레이블 간의 관계를 고려하지 않고, 각각의 애스팩트를 개별적으로 분석함을 말한다. 반면, 본 연구의 다중 레이블을 적용한 한글 애스팩트 마이닝은 레이블 간의 상관관계까지 고려하여 한 번에 분석이 가능하다. 그렇기 때문에 단일 레이블 분류방법을 적용한 애스팩트 마이닝과는 다르다. 본 연구의 기여는 학문적 의의와 실무적 의의로 구분한다.

학문적 의의는 다음과 같다. 첫째, 본 연구는 한글 애스팩트 마이닝을 효과적으로 수행하기 위해 다중레이블 분류 방법을 적용하였다. 한글 감

성분류 연구는 단일 레이블 분류 방법을 적용한 연구로 문장 전체의 감성을 평가하는 연구에 편중되어 있다. 그리고 애스팩트에 따라 각각의 애스팩트의 오피니언을 분석한 연구도 부족하다. 본 연구에서 제안한 방법은 중복된 의견이나 글 안에 포함된 다양한 의견을 탐지할 수가 있어서 기존연구와 차별화된다.

둘째, 다중 레이블을 적용한 기존 애스팩트 마이닝 연구에서 속성선택방법을 적용한 연구는 매우부족하다. 앞에서 이미 기술한 바와 같이 데이터 마이닝 분야에서 속성선택은 결과의 질을 올리고 분석속도를 줄이는 효과적이다. 본 연구는 다중 레이블 속성선택 방법간 비교를 통하여 어떤 방법이 한글 감성분석시 더 효과적인지를 실증적으로 입증하였다.

본 연구가 가지고 있는 실무적 가치는 다음과 같다. 최근 많은 기업은 온라인 리뷰, 신문 기사의 댓글, 블로그 등의 자료를 기반으로 시장을 분석하고 기업의 브랜드를 높이고자 치열하게 경쟁하고 있다. 기존의 설문조사와 달리 감성분석 방법은 기업의 입장에서 고객의 감성변화와 시장의 판도변화를 파악하는데 시간 및 비용측면에서 더 효과적이기 때문이다. 또한 스마트폰 기반의 모바일 네트워크 및 인터넷 환경이 크게 개선되면서 기업이 관심을 갖는 온라인 데이터의 양이 크게 증가하고 다양성 또한 매우 커졌다. 이러한 환경에 처한 기업의 입장에서 본 연구에서 제안하는 다중 레이블 기반의 애스팩트 마이닝은 온라인 데이터로부터 보다 풍부한 결과 도출이 가능하여 전략적 대응을 더욱 용이하게 한다. 즉, 타겟 제품과 서비스에 대한 애스팩트 별로 소비자의 반응을 구체적으로 알 수 있다. 더욱이, 본 연구에서 제안한 감성

분석 모델은 감성분류 자동화 시스템 구축에 적용이 가능하여서, 온라인상에서 실시간으로 업로드되는 새로운 데이터를 자동으로 분류할 수가 있어서 소비자의 반응을 실시간으로 애스펙트 별로 세밀하게 관찰할 수 있다.

본 연구는 지도학습 방법을 적용하였기 때문에 사전에 알 수 있는 레이블이 이미 존재하여야만 적용이 가능하다는 한계점이 있다. 따라서, 보다 고도화된 소비자 감성분석 자동 분류 시스템 구축을 위해서는 비지도학습 또는 강화학습을 적용하여 분석기법의 효과성을 올릴 필요가 있다. 이는 향후 연구주제로 더 연구될 가치가 있다.

## 참고문헌

- [1] 김다예, 이영인, “Word2Vec을 활용한 뉴스 기반 주가지수 방향성 예측용 감성 사전 구축”, *한국빅데이터학회지*, 제3권, 제1호, 2018, pp. 13-20.
- [2] 남길임, 조은경, *한국어 텍스트 감성분석*, 커뮤니케이션북스, 2017.
- [3] 박재수, 이재수, “아파트 매매가격과 부동산 온라인 뉴스의 교차상관관계와 인과관계 분석: 온라인 뉴스 기사의 비정형 빅데이터를 활용한 감성분석 기법의 적용”, *Journal of Korea Planning Association*, 제54권, 제1호, 2019, pp. 131-147, Available at DOI: 10.17208/jkpa.2019.02.54.1.131.
- [4] 박현정, 송민채, 신경식, “CNN을 적용한 한국어 상품평 감성분석: 형태소 임베딩을 중심으로”, *지능정보연구*, 제24권, 제2호, 2018, pp. 59-83, Available at DOI: 10.13088/jiis.2018.24.2.059.
- [5] 이영준, 윤보현, “토픽모델링과 감성분석에 기반한 금통위 의사록 분석”, *Journal of The Korean Data Analysis Society*, 제21권, 제2호, 2019, pp. 889-900.
- [6] 정민지, 이윤희, 유채민, 김지원, 정재은, “텍스트 마이닝 기법을 이용한 모바일 간편결제 서비스에 대한 소비자 반응 분석: 삼성페이를 중심으로”, *디지털융복합연구*, 제17권, 제1호, 2019, pp. 9-27, Available at DOI: 10.14400/JDC.2019.17.1.009.
- [7] 정폴잎, 안현철, 광기영, “텍스트 마이닝과 소셜 네트워크 분석을 이용한 스마트폰 디자인의 핵심속성 및 가치 식별”, *대한경영학회지*, 제32권, 제1호, 2019, pp. 27-47, Available at DOI: 10.18032/kaaba.2019.32.1.27.
- [8] Abberley, L., N. Gould, K. Crockett, and J. Cheng, “Modelling road congestion using ontologies for big data analytics in smart cities”, *International Smart Cities Conference*, 2017, pp. 14-17, Available at DOI: 10.1109/ISC2.2017.8090795.
- [9] Ali, F., D. Kwak, P. Khan, S. El-Sappagh, A. Ali, S. Ullah, K. H. Kim, and K. S. Kwak, “Transportation sentiment analysis using word embedding and ontology-based topic modeling”, *Knowledge-Based Systems*, Vol.174, 2019, pp. 27-42, Available at DOI: 10.1016/j.knosys.2019.02.033.
- [10] Almars, A., X. Li, and X. Zhao, “Modelling user attitudes using hierarchical sentiment-topic model”, *Data & Knowledge Engineering*, Vol.119, 2019, pp. 139-149, Available at DOI: 10.1016/j.datak.2019.01.005.
- [11] Alqaryouti, O., N. Siyam, A. A. Monem, and K. Shaalan, “Aspect-based sentiment analysis using smart government review data”, *Applied Computing and Informatics*, 2019 Online, Available at DOI: 10.1016/j.aci.2019.11.003.
- [12] Celardo, L. and M. G. Everett, “Network text analysis: A two-way classification approach”, *International Journal of Information Management*, 2019 pp. 1-8, Available at DOI: 10.1016/j.ijinfomgt.2019.09.005.
- [13] Chen, F. and Y. Huang, “Knowledge-enhanced

- neural networks for sentiment analysis of Chinese reviews”, *Neurocomputing*, Vol.368, 2019, pp. 51-58, Available at DOI: 10.1016/j.neucom.2019.08.054.
- [14] Das, S., X. Sun, and A. Dutta, “Text mining and topic modeling of compendiums of papers from transportation research board annual meetings”, *Transportation Research Record*, Vol.2552, No.1, 2016, pp. 48-56, Available at DOI: 10.3141/2552-07.
- [15] Furnkranz, J., E. Hullermeier, E. L. Mencia, and K. Brinker, “Multilabel classification via calibrated label ranking”, *Machine Learning*, Vol.73, No.2, 2008, pp. 133-153.
- [16] Garcia-Pablos, A., M. Cuadros, and G. Rigau, “W2VLDA: almost unsupervised system for aspect based sentiment analysis”, *Expert Systems with Applications*, Vol.91, 2018, pp. 127-137, Available at DOI: 10.1016/j.eswa.2017.08.049.
- [17] Giatsoylou, M., M. G. Vozalis, K. Diamantaras, A. Vakali, G. Sarigiannidis, and K. C. Chatzisavvas, “Sentiment analysis leveraging emotions and word embeddings”, *Expert Systems with Applications*, Vol.69, 2017, pp. 214-224, Available at DOI: 10.1016/j.eswa.2016.10.043.
- [18] Huang, J., G. Li, S. Wang, Z. Xue, and Q. Huang, “Multi-label classification by exploiting local positive and negative pairwise label correlation”, *Neurocomputing*, Vol.257, 2017, pp. 164-174, Available at DOI: 10.1016/j.neucom.2016.12.073.
- [19] Hüllermeier, E., J. Fürnkranz, W. Cheng, and K. Brinker, “Label ranking by learning pairwise preferences”, *Artificial Intelligence*, Vol.172, No.16, 2008, pp. 1897-1916, Available at DOI: 10.1016/j.artint.2008.08.002.
- [20] Jena, R., “An empirical case study on Indian consumers’ sentiment towards electric vehicles: A big data analytics approach”, *Industrial Marketing Management*, 2020, pp. 1-12, Available at DOI: 10.1016/j.indmarman.2019.12.012.
- [21] Jha, V., R. Savitha, P. D. Shenoy, K. R. Venugopal, and A. K. Sangaiah, “A novel sentiment aware dictionary for multi-domain sentiment classification”, *Computers & Electrical Engineering*, Vol.69, 2018, pp. 585-597, Available at DOI: 10.1016/j.compeleceng.2017.10.015.
- [22] Khan, A. U. R., M. Khan, and M. B. Khan, “Naive multi-label classification of YouTube comments using comparative opinion mining”, *Procedia Computer Science*, Vol.82, 2016, pp. 57-64, Available at DOI: 10.1016/j.procs.2016.04.009.
- [23] Krouska, A., C. Troussas, and M. Virvou, “Comparative evaluation of algorithms for sentiment analysis over social networking services”, *Journal of Universal Computer Science*, Vol.23, No.8, 2017, pp. 755-768.
- [24] Lawani, A., M. R. Reed, T. Mark, and Y. Zheng, “Reviews and price on online platforms: Evidence from sentiment analysis of Airbnb reviews in Boston”, *Regional Science and Urban Economics*, Vol.75, 2019, pp. 22-34, Available at DOI: 10.1016/j.regsciurbeco.2018.11.003.
- [25] Le, H., J. Lee, and H. K. Lee, “Purchase process aspect-based opinion mining: An application for online shopping mall”, *The Journal of Internet Electronic Commerce Research*, Vol.15, No.2, 2015, pp. 15-28.
- [26] Lee, J. and D. W. Kim, “SCLS: Multi-label feature selection based on scalable criterion for large label set”, *Pattern Recognition*, Vol.66, 2017, pp. 342-352, Available at DOI: 10.1016/j.patcog.2017.01.014.
- [27] Lee, S., J. H. Lee, S. H. Jung, and J. Park, “The role of entropy of review text sentiments on online WOM and movie box office sales”, *Electronic Commerce Research and Applications*, Vol.22,

- 2017, pp. 42-52, Available at DOI: 10.1016/j.elerap.2017.03.001.
- [28] Lee, S. H., J. Cui, and J. W. Kim, "Sentiment analysis on movie review through building modified sentiment dictionary by movie genre", *Journal of Intelligence and Information Systems*, Vol.22, No.2, 2016, pp. 97-113, Available at DOI: 10.13088/jiis.2016.22.2.097.
- [29] Li, Y., C. Shi, H. Zhao, F. Zhuang, and B. Wu, "Aspect mining with rating bias", *European Conference on Machine Learning and Knowledge Discovery in Databases*, 2016, pp. 458-474.
- [30] Liu, H., H. Motoda, R. Setiono, and Z. Zhao, "Feature selection: An ever evolving frontier in data mining", *In Feature Selection in Data Mining*, 2010, pp. 4-13.
- [31] Liu, X., "Analyzing the impact of user-generated content on B2B Firms' stock performance: Big data analysis with machine learning methods", *Industrial Marketing Management*, 2019, pp. 1-10, Available at DOI: 10.1016/j.indmarman.2019.02.021.
- [32] Liuc, H. and H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*, Springer Science & Business Media, 2012.
- [33] Marcheggiani, D., O. Tackstrom, A. Esuli, and F. Sebastiani, "Hierarchical multi-label conditional random fields for aspect-oriented opinion mining", *European Conference on Information Retrieval*, 2014, pp. 273-285.
- [34] Mikolov, T., K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space", arXiv preprint arXiv:1301.3781, 2013.
- [35] Ngo-Ye, T. L. and A. P. Sinha, "The influence of reviewer engagement characteristics on online review helpfulness: A text regression model", *Decision Support Systems*, Vol.61, 2014, pp.47-58, Available at DOI: 10.1016/j.dss.2014.01.011.
- [36] Pereira, R. B., A. Plastino, B. Zadrozny, and L. H. Merschmann, "Correlation analysis of performance measures for multi-label classification", *Information Processing & Management*, Vol.54, No.3, 2018, pp. 359-369, Available at DOI: 10.1016/j.ipm.2018.01.002.
- [37] Qiu, G., B. Liu, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation", *Computational Linguistics*, Vol.37, No.1, 2011, pp. 9-27, Available at DOI: 10.1162/coli\_a\_00034.
- [38] Rathan, M., V. R. Hulipalled, K. R. Venugopal, and L. M. Patnaik, "Consumer insight mining: Aspect based twitter opinion mining of mobile phone reviews", *Applied Soft Computing*, Vol.68, 2018, pp. 765-773, Available at DOI: 10.1016/j.asoc.2017.07.056.
- [39] Read, J., B. Pfahringer, G. Holmes, and E. Frank, "Classifier chains for multi-label classification", *Machine Learning and Knowledge Discovery in Databases*, 2009, pp. 254-269.
- [40] Read, J., B. Pfahringer, G. Holmes, and E. Frank, "Classifier chains for multi-label classification", *Machine Learning*, Vol.85, No.3, 2011, pp. 333-359.
- [41] Santosh, D. T., K. S. Babu, S. D. V. Prasad, and A. Vivekananda, "Opinion mining of online product reviews from traditional LDA topic clusters using feature ontology tree and sentiwordnet", *International Journal Education and Management Engineering*, Vol.6, 2016, pp. 34-44, Available at DOI: 10.5815/ijeme.2016.06.04.
- [42] Shibuya, Y. and H. Tanaka, "Public sentiment and demand for used cars after a large-scale disaster: Social media sentiment analysis with Facebook pages", *Social Web in Emergency and*

- Disaster Management*, 2018.
- [43] Spolaor, N., E. A. Cherman, M. C. Monard, and H. D. Lee, "ReliefF for multi-label feature selection", *IEEE Brazilian Conference on Intelligent Systems*, 2013, pp. 6-11.
- [44] Tan, X., Y. Cai, J. Xu, H. F. Leung, W. Chen, and Q. Li, "Improving aspect-based sentiment analysis via aligning aspect embedding", *Neurocomputing*, 2019, Available at DOI: 10.1016/j.neucom.2019.12.035.
- [45] Wang, G., J. Sun, J. Ma, K. Xu, and J. Gu, "Sentiment classification: The contribution of ensemble learning", *Decision Support Systems*, Vol.57, 2014, pp. 77-93, Available at DOI: 10.1016/j.dss.2013.08.002.
- [46] Yadollahi, A., A. G. Shahraki, and O. R. Zaiane, "Current state of text sentiment analysis from opinion to emotion mining", *Association for Computing Machinery Computing Surveys*, Vol.50, No.2, 2017, pp. 25-58, Available at DOI: 10.1145/3057270.
- [47] Yang, C., H. Zhang, B. Jiang, and K. Li, "Aspect-based sentiment analysis with alternating coattention networks", *Information Processing & Management*, Vol.56, No.3, 2019, pp. 463-478, Available at DOI: 10.1016/j.ipm.2018.12.004.
- [48] Zhang, M. L. and Z. H. Zhou, "ML-KNN: A lazy learning approach to multi-label learning", *Pattern Recognition*, Vol.40, No.7, 2007, pp. 2038-2048, Available at DOI: 10.1016/j.patcog.2006.12.019.
- [49] Zhang, M. L. and Z. H. Zhou, "A review on multi-label learning algorithms", *IEEE Transactions on Knowledge and Data Engineering*, Vol.26, No.8, 2013, pp. 1819-1837, Available at DOI: 10.1109/TKDE.2013.39.
- [50] Zhou, Z. H., M. L. Zhang, S. J. Huang, and Y. F. Li, "Multi-instance multi-label learning", *Artificial Intelligence*, Vol.176, No.1, 2012, pp. 2291-2320.

## Multi-Label Classification Approach to Effective Aspect-Mining

Jong Yoon Won\* · Kun Chang Lee\*\*

### Abstract

Recent trends in sentiment analysis have been focused on applying single label classification approaches. However, when considering the fact that a review comment by one person is usually composed of several topics or aspects, it would be better to classify sentiments for those aspects respectively. This paper has two purposes. First, based on the fact that there are various aspects in one sentence, aspect mining is performed to classify the emotions by each aspect. Second, we apply the multiple label classification method to analyze two or more dependent variables (output values) at once. To prove our proposed approach's validity, online review comments about musical performances were garnered from domestic online platform, and the multi-label classification approach was applied to the dataset. Results were promising, and potentials of our proposed approach were discussed.

**Keywords:** *Data-Mining, Aspect-Mining, Sentiment-Analysis, Multi-Label Classification, Machine Learning*

---

\* Ph D. Student, SKK Business School, Sungkyunkwan University

\*\* Corresponding Author, Professor, Global Business Administration/Department of Health Sciences & Technology, SAIHST(Samsung Advanced Institute for Health Sciences & Technology), Sungkyunkwan University



## ○ 저 자 소 개 ○



**Jong Yoon Won (yoonjbest1@gmail.com)**

He is now pursuing PhD degree at the SKK Business School in Sungkyunkwan University (Seoul, South Korea). He is actively working on neuroscience based decision-making mechanism, business problem solving creativity, Big-Data analytics, and AI-driven analysis of business problem-solving processes, etc.



**Kun Chang Lee (kunchanglee@gmail.com)**

He is a full professor in Sungkyunkwan University, Seoul, South Korea. He is also affiliated with SAIHST (Samsung Advanced Institute for Health Sciences & Technology). He has published many papers at journals like Decision Support Systems, Journal of MIS, IEEE Transactions on Engineering Management, to name a few.

논문접수일 : 2020년 01월 24일

게재확정일 : 2020년 04월 10일

1차 수정일 : 2020년 02월 17일