

마켓 인사이트를 위한 상품 리뷰의 다차원 분석 방안*

박정현

한양대학교 경영학부
(dutlsdlek@hanyang.ac.kr)

이서호

한양대학교 경영학부
(tjghdia@hanyang.ac.kr)

임규진

한양대학교 경영학부
(kjilincoln@hanyang.ac.kr)

여운영

한양대학교 비즈니스인포매틱스학과
(ywy31730@hanyang.ac.kr)

김종우

한양대학교 경영대학 경영학부
(kjw@hanyang.ac.kr)

인터넷의 발달로, 소비자들은 이커머스에서 손쉽게 상품 정보를 확인한다. 이때 활용되는 상품 리뷰는 사용자 경험을 토대로 작성되어 구매의사결정의 효율성을 높일 뿐만 아니라 상품 개발에 도움을 주기도 한다. 하지만, 방대한 양의 상품 리뷰에서 관심있는 평가차원의 세부내용을 파악하는 데에는 많은 시간과 노력이 소비된다. 예를 들어, 노트북을 구매하려는 소비자들은 성능, 무게, 디자인과 같은 평가차원에 대해 각 차원별로 비교 상품의 평가를 확인하고자 한다.

따라서 본 논문에서는 상품 리뷰에서 다차원 상품평가 점수를 자동적으로 생성하는 방안을 제안하고자 한다. 본 연구에서 제시하는 방안은 크게 2단계로 구성된다. 사전준비 단계와 개별상품평가 단계로, 대분류 상품군 리뷰를 토대로 사전에 생성된 차원분류모델과 감성분석모델이 개별상품의 리뷰를 분석하게 된다. 차원분류모델은 워드임베딩과 연관분석을 결합함으로써 기존 연구에서 차원과 단어들의 관련성을 찾기 위한 워드임베딩 방식이 문장 내 단어의 위치만을 본다는 한계를 보완한다. 감성분석모델은 정확한 극성 판단을 위해 구(phrase) 단위로 긍부정이 태깅된 학습데이터를 구성하여 CNN 모델을 생성한다. 이를 통해, 개별상품평가 단계에서는 구 단위의 리뷰에 준비된 모델들을 적용하고 평가차원별로 종합함으로써 다차원 평가점수를 얻을 수 있다.

본 논문의 실험에서는 대분류 상품군 리뷰 약 260,000건으로 평가모델을 구성하고, S사와 L사의 노트북 리뷰 각 1,011건과 1,062건을 실험데이터로 활용한다. 차원분류모델은 구로 분해한 개별상품 리뷰를 6개 평가차원으로 분류했고, 기존 워드임베딩 방식보다 연관분석을 결합한 모델의 정확도가 13.7% 증가했음을 볼 수 있었다. 감성분석모델은 문장보다 구 단위로 학습한 모델이 평가차원을 면밀히 분석함으로써 29.4% 더 높은 정확도를 보임을 확인했다. 본 연구를 통해 판매자, 소비자 모두가 상품의 다차원적 비교가 가능하다는 점에서 구매 및 상품 개발에 효율적인 의사결정을 기대할 수 있다.

주제어 : 연관분석, 감성분석, 이커머스, 워드임베딩, 텍스트마이닝, 극성탐지

논문접수일 : 2020년 4월 1일 논문수정일 : 2020년 5월 26일 게재확정일 : 2020년 6월 2일
원고유형 : 일반논문 교신저자 : 김종우

* 이 논문은 ‘대학혁신지원사업 비즈니스 랩 기반의 프로젝트 학기제’의 지원을 받아서 수행된 연구임.

1. 서론

2019년 미국 퓨 리서치에 의하면, 한국 국민 100명 중 95명이 스마트폰을 사용하는 것으로 집계된다(Silver, 2019). 스마트폰의 대중화와 인터넷의 발달, IT기술의 발전을 통해 사람들은 쉽고 빠르게 정보를 확인하고 온라인을 통한 편리한 구매가 가능해졌다. 하지만 온라인에서 구매하는 상품은 실물을 확인할 수 없다는 아쉬움과 상품 품질에 대한 불확실성이 존재한다(Lee, 2011). 소비자들은 효율적인 구매의사결정을 위해 기존 사용자의 경험이 담긴 리뷰를 확인하며, 리뷰를 통해 자신이 원하는 상품에 대한 정보를 확인하게 된다(Kovacs, 2008). 즉, 이커머스와 같은 사이트에서 제공하는 리뷰들은 소비자와 판매자 간의 정보 격차를 줄일 뿐만 아니라 자신의 후기를 남김으로써 타인에게 의견을 공유할 수 있는 소통 도구가 되고 있다. 이와 같이 리뷰는 오늘날 의사결정뿐만 아니라 온라인 네트워크를 통해 소통의 기회를 제공하는 도구가 되어 그 중요성과 가치가 더욱 커지고 있다(Henry et al, 2004).

소비자들은 이커머스 상의 리뷰와 평점을 통해 사용자의 경험과 상품 정보를 인지하고 종합적인 평가를 확인한다(Jeong et al, 2019). 하지만 텍스트 리뷰와 같이 정형화되지 않은 리뷰의 경우, 다양한 정보들이 혼재되어 있어 필요한 정보를 식별하는데 수고가 필요하다. 또한 평점 형태로 상품의 평가차원을 일차원적으로 제공하기 때문에 소비자는 구매의사결정에 있어 중요하다고 생각하는 상품의 평가차원을 파악하기 위해, 상당수의 리뷰를 하나하나 읽어보는 과정에서 많은 시간과 노력을 허비하게 된다. 이로 인해, 소비자들은 비교하고자 하는 상품이 있을 경우

여러 상품들의 평가를 종합적으로 확인해야 하며 관심있는 차원의 평가를 파악하는데 어려움을 겪고 있다.

또한 판매자의 입장에서 방대한 양의 상품 리뷰로부터 자신의 상품이 어떤 평가를 받고 있는지 정확히 파악하기 위해서는 많은 노력이 필요하다(Byun et al, 2016). 마켓 센싱을 위해서는 자사뿐만 아니라 경쟁사 상품과의 비교가 필수적이며, 이는 리뷰 데이터를 다차원적으로 활용하여 자사와 타사 상품간의 입체적인 비교 분석이 이뤄져야 한다. 그러나 대부분의 이커머스에서 제공되는 평가는 평가차원별 비교 분석 정보가 아니므로, 제조 기업과 판매자가 경쟁사에 비해 어떤 차원에서 우위에 있는지 강,약점을 파악하고 이를 효율적으로 판단하는 데에 한계가 있다.

이러한 필요성으로 인해서 상품 리뷰를 효과적으로 다차원 분석하기 위한 텍스트 마이닝 연구가 진행되어 왔다. 상품 리뷰의 다차원 분석을 위해서는 차원과 연관 단어들의 매핑이 필수적이다. 이를 위해서 초기에는 차원별 단어사전을 사용하거나(Yeon et al. 2013), WordNet과 같은 온톨로지를 사용하였다. 하지만 사전 구축을 위해서는 많은 시간과 노력을 요하며, 품질에 대한 검증이 어렵고 확장성 측면에서도 한계를 가진다. 이를 극복하기 위해서 최근 연구에서는 워드 임베딩을 통해 문장 내 단어 간의 거리를 계산하여 차원을 분류함으로써, 사전을 사전에 구축해야 하는 한계점을 보완한다(Jeong et al. 2018). 하지만 워드임베딩만 활용하는 경우, 문장 내에서 비슷한 위치에 오지 않는 표현과 품사가 다른 표현의 경우 유사한 차원으로 분류하는데 어려움이 있다.

본 연구는 선행연구들의 문제점을 해결하기 위해서 다음과 같은 해결책을 통해 차별점을 가

진다. 첫째, 리뷰 문장을 구(phrase) 단위로 학습시킨 모델을 구현함으로써 다양한 차원에 대해 선행연구의 문장(sentence) 단위로 학습시킨 모델보다 정확도 높은 분석을 가능하게 하였다. 둘째, 연관분석기법을 워드임베딩과 함께 결합함으로써 동시 출현 빈도를 확인하여 기존 워드 임베딩 기법만 활용한 연구보다 차원분류모델의 정확도를 높인다 셋째, 딥러닝 모델을 활용하여 평가차원별 점수를 파악할 수 있게 함으로써 사전기반 선행 연구와 달리 자동화된 방식으로 모델을 제시하였다. 따라서 본 연구의 모델은 온라인 리뷰의 특질을 잘 반영하였기 때문에 실제 리뷰에 대해 더 정확도가 높으며, 자동화된 과정으로 확장성 있게 활용될 수 있다는 점에서 의의가 있다.

본 논문의 구성은 다음과 같다. 서론을 통해 연구 배경과 필요성을 설명하며, 2장에서 본 연구와 유사한 방식으로 진행된 기존의 연구들을 요약한다. 3장에서는 본 연구에서 제안하고자 하는 상품 리뷰의 다차원 분석 방안에 대해 단계별로 설명한다. 4장에서는 본 모델의 정확도를 비교한 실험결과에 대해 서술하며, 5장에서는 본 연구가 기여하는 바와 한계, 향후 연구 방향에 대한 결론은 정리한다.

2. 관련 연구

2.1 상품 리뷰 감성 분석

상품을 직접 확인하고 구매할 수 있는 오프라인 매장과 달리, 온라인 마켓은 시각적인 정보에만 의존하여 상품을 구매하게 된다. 이러한 한계를 보완하기 위해 소비자들은 상품 리뷰를 확인

하고자 하며, 리뷰는 간접적으로 사용자의 경험 정보를 제공함으로써 정보 격차를 줄이고 효과적인 구매의사결정을 가능하게 한다. 따라서 리뷰는 활용가치가 큰 데이터로써, 이를 통해 사용자의 취향이나 주관 등을 분석하고 유용한 결과를 도출하려는 연구들이 진행되고 있다.

온라인에서 제품 평가정보가 구매의사결정에 미치는 효과를 살펴본 연구에서는 댓글과 리뷰 기사가 구매의도에 유의미한 효과를 보이며, 사이트의 신뢰도에 따라 영향력이 달라지기도 한다(Kim et al. 2007). 이에 이커머스 사이트에서는 이를 효과적으로 활용하고자 상품 리뷰에 대한 시스템을 구축하고 다양한 통계정보를 제공하고 있으며 별도의 체험단을 두어 신뢰성을 높이고자 한다.

또한 실제 리뷰 관리에 있어서는 상품에 대한 별점 통계, 유용한 리뷰 분류 등을 소비자의 참여를 통해 제공하고 있다. 감성 분석(sentiment analysis)은 이러한 일련의 과정을 자동화하는 도구로서, 상품 리뷰의 사용자 의견을 대상으로 그 의견이 긍정적인지, 부정적인지 판단한 후 요약하여 제공한다(Yeon et al. 2011). 이 절에서는 상품 리뷰에 대한 감성 분석 연구들을 단일 차원 감성 분석, 다차원 차원 감성 분석, 딥러닝 기반 감성 분석으로 구분하여 살펴본다.

2.1.1 단일 차원 감성 분석

초기 리뷰 감성분석 연구에서는 상품평을 단일 차원으로 정량화하여 요약한다. 기존 연구에선 상품의 극성(긍정 또는 부정)을 나타내는 단어의 빈도수(Gamon et al. 2005) 또는 이와 평점이 결합된 형태로 분석 결과를 제시하였다(Scaffidi et al. 2007). 어휘의 사용 빈도수를 통한

통계적 기법을 사용하는 방향으로의 연구는 TF-IDF와 같은 방식을 적용한 연구가 진행되었다(Scaffidi et al. 2007). 또한 이를 기반으로 워드넷(wordnet)을 활용한 긍정과 부정, 중립을 판단하는 분류 방안이 제시되었다(Liu et al. 2005). 사전구축을 비롯한 위의 방식은 분석과 예측의 정확도를 올린다는 장점은 있지만 다른 상품군의 리뷰 분석에 적용하기 힘들기에 확장성의 한계가 있다. 또한 여러 가지 차원을 담고 있는 상품 리뷰를 하나의 차원으로 분류하는 것은 소비자의 리뷰를 제대로 대변하지 못하는 포괄성의 한계가 있다.

2.1.2 다차원 감성 분석

최근에는 일차원으로 리뷰를 요약 분석하는 연구의 한계를 보완하기 위해서 상품을 특정하는 여러가지 차원의 평가를 구분하여 리뷰의 감성을 분석하는 연구들이 진행됐다. 기존 다차원 분석 연구에는 호텔 정보 공유 사이트의 리뷰 데이터에서 용어 빈도수를 기반으로 사전을 구축하였고 숙박료, 지역명과 같이 소비자들이 호텔을 선택할 때 중요하게 생각하는 차원별로 분류하고, 평가를 요약하는 시스템이 제안된 바가 있다(Kim et al. 2018).

(1) 사전기반 다차원분석

다차원 상품평을 사전으로 구축하여 분석하는 연구도 진행되었으며 리뷰 내에서 상품의 주요 차원을 표현하는 어휘와 각 어휘들의 극성(polarity) 정보들을 사전으로 저장하여, 전자 상거래 사이트의 리뷰에 담긴 의견 정보가 긍정적인지, 부정적인지 판별하는 모델링 방안을 제시되었다(Myong et al. 2018). 또한 식당의 차원별

긍정/부정 여부를 파악하기 위해 차원별 사전을 구축하고, 이에 따라 5개의 차원(분위기, 장소, 서비스, 음식, 가격)으로 리뷰를 분류하여 분석한 연구도 수행된 바가 있다(Yeon et al. 2013).

하지만 사전구축 방식은 앞선 일차원 분석의 사전기반 연구와 동일하게 사람이 직접 특성(feature)을 선별해야 한다는 점에서 확장성의 한계가 있으며, 리뷰 분석에 자동화가 용이하지 않다. 본 연구는 용어사전으로 리뷰를 선별하는 것이 아닌, 리뷰 전체를 이용하여 다차원에 분류한다는 점에서 리뷰의 정보를 보다 포괄적으로 활용할 수 있다.

(2) Word2Vec을 활용한 다차원 분석

Word2Vec모델은 단어를 계산에 적용할 수 있는 숫자로 변환해서 각 단어에 가중치를 할당할 수 있어 각 단어에 할당된 점수들을 활용하여 텍스트를 분류하는 연구가 활발히 진행됐다(Ju et al. 2019). Word2Vec은 대량의 텍스트를 받아들여 일반적으로 100차원 이상 크기의 벡터 공간을 만들고 각 말뭉치들은 벡터 공간에 알맞게 정렬한다. 비슷한 문맥을 지니는 단어들은 벡터 공간 내에서 가까운 거리에 위치하게 된다. 현재 가장 많이 사용하는 Word2Vec모델은 CBOW와 Skip-gram을 사용해서 학습된다(Ju et al. 2019). CBOW는 주변에 있는 단어들을 중심에 있는 단어에 맞추는 방식이고 Skip-gram은 중심에 있는 단어로 주변 단어를 예측하는 방법이다.

선행연구로는 영화와 호텔 후기 사이트 등의 텍스트 리뷰를 활용하여 Word2Vec모델로 단어 간의 문맥적 유사도를 벡터화하는 연구가(Myong, 2018; Seo, 2016) 있다. 이를 활용한 다차원적 리뷰분석은 각 리뷰의 명사만을 추출하여 워드임

베딩 방식을 사용하여, 각 차원별로 거리가 가까운 단어를 사전으로 구축하였다(Jeong et al. 2018). 이를 통해 리뷰를 사전에 따라 다차원으로 분류하는 시스템을 구현하였다.

워드임베딩(word embedding) 방식을 활용한 연구들은 개별 단어 임베딩만으로 다차원 분석을 진행할 경우, 다수의 단어의 조합이 개별 단어 벡터의 조합으로 표현될 수 없다는 한계를 갖고 있다. 이를 보완하고자 기존 연구에서는 구문에서 동시등장단어(word cooccurrence)를 식별하고 이들을 학습시켰다(Mikolov et al. 2017). 본 연구는 기존 연구의 워드임베딩의 유사도 값과 동시에 연관분석기법을 함께 사용하여 자동화된 모델을 제시하였다.

2.1.3. 딥러닝 기반 감성 분석

온라인 상의 방대한 양의 상품평을 자동으로 분류하고 필요한 정보를 추출하기 위해서는 텍스트의 감성분석이 필요하다. 기존의 감성분석에는 사전 구축방식과 같은 반 자동화된 방법이나 Naive Bayes와 Logistic Regression등의 전통적인 기계학습 방법이 많이 사용되었으나(Kim et al., 2018), 최근에는 딥러닝 기반의 기술들이 기존 반 자동화 방식에서 사람이 직접 선별하였던 feature을 학습할 수 있다는 장점으로 높은 성능을 보이고 있다(Collobert et al. 2016).

CNN(convolutional neural network)모델은 감성 분석을 위한 딥러닝 모델의 하나로서 Kim(2014)의 연구에서 소개된 모델 중 여러 텍스트 데이터 셋에 대해 높은 극성 탐지 성능을 보였다. 기존 연구에선 CNN은 주로 이미지 인식에 응용되었지만 워드임베딩을 통한 단어의 벡터화를 이용해 문장의 감성분석에서도 CNN을 이용한 연구

가 활발히 연구되고 있다 (Goldberg et al. 2014). 리뷰텍스트를 활용하여 CNN 감성 분석을 실시한 예시로는 영화평을 평점 0.0이상 3.0이하는 부정, 평점 7.0이상 10.0이하는 긍정라벨을 붙이는 전처리 과정을 통해 문장을 학습시킨 연구가 있다(Kim et al. 2016). 그러나 해당 연구는 하나의 문장 내에서도 다양한 극성을 가질 수 있음에도 불구하고 리뷰를 평점이라는 하나의 기준으로 학습데이터를 형성시켰다는 한계를 가지고 있다.

본 연구에서는 학습 데이터를 리뷰의 문장을 소비자의 평가를 담고 있는 최소 단위인 ‘구’ 단위로 분해하여 CNN 모델을 구축하여, 문장 단위로 학습데이터를 형성한 CNN모델과 성능을 비교하도록 한다.

2.2 연관규칙분석

연관규칙분석(association rule analysis)은 거나 사건들 간의 일련의 규칙을 생성하는 데이터 마이닝 기법으로서, 대용량 데이터베이스에서 다수의 항목들 간의 동시 발생 규칙을 찾아내는 데 사용된다(Yoo et al. 2017). 연관성의 평가 척도로는 주로 지지도(support), 신뢰도(confidence), 향상도(lift)가 사용된다.

어떤 사건 A, B가 자주 발생하는 사건임을 판별하는 데 쓰는 ‘지지도’는 ‘조건절(A)과 결과절(B)가 동시에 일어날 확률’로서 아래와 같은 수식으로 정리된다.

$$\text{Support}(A \rightarrow B) = P(A, B) \quad (1)$$

어떤 사건 A가 일어난 상황에서 B가 일어날 확률을 판별하는 데 쓰는 신뢰도는 아래와 같은 수식처럼 조건절(A)와 결과절(B)의 조건부확률

로 정의된다.

$$Confidence(A \rightarrow B) = P(A, B) / P(A) \quad (2)$$

어떤 사건 A, B 간에 생성된 규칙이 실제 효용 가치가 있는지를 판별하는데 사용되는 향상도는 아래의 식과 같이 어떠한 사건 A가 주어지지 않았을 때의 사건 B의 확률 대비 사건 A가 주어졌을 때 사건 B의 확률 증가 비율로 정의된다. 사건 A와 B사이에 아무런 관계가 없다면 즉, 독립 관계라면 향상도의 값은 1이며 향상도가 1보다 커질수록 연관성이 높다고 말할 수 있다.

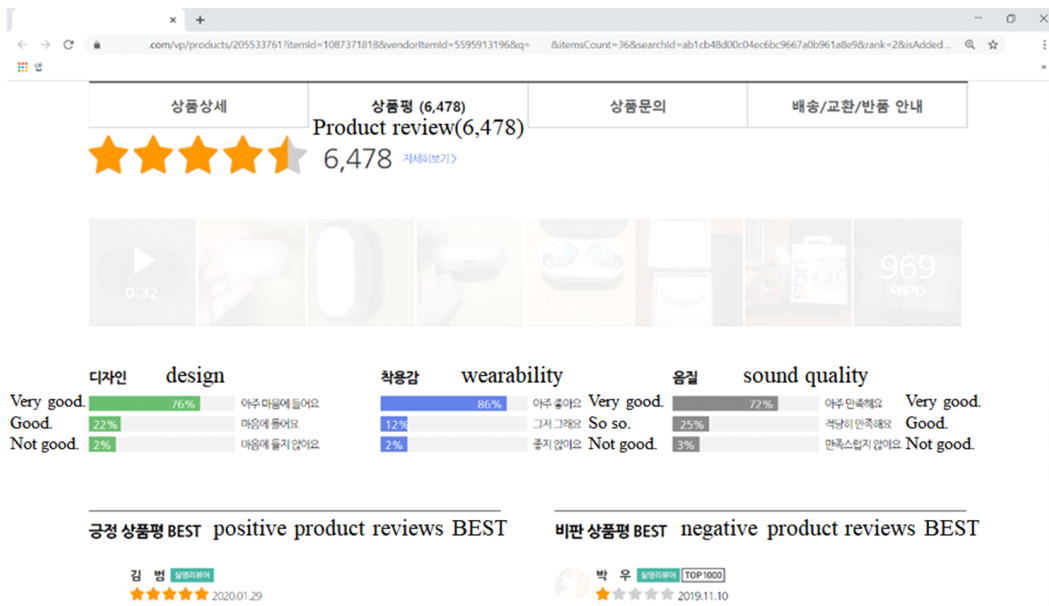
$$Lift(A \rightarrow B) = P(A, B) / P(A) \cdot P(B) \quad (3)$$

본 논문에서는 리뷰를 다차원으로 분석할 시, 앞서 언급한 Word2Vec과 같은 워드임베딩의 한계를 보완하기 위하여 연관분석기법을 함께 사용하도록 한다.

3. 제안 모델

본 연구는 특정 제품의 리뷰 데이터를 이용하여 사용자나 판매자가 그 제품에 대한 고객의 평가를 다차원으로 알아볼 수 있게 하는 것을 목표로 한다. 몇몇 이커머스에서는 <Figure 1>과 같이 일부 차원(디자인, 착용감, 음질)에 대하여 사용자 만족도를 제공하고 있지만, 판매자의 요청에 의해 제공되고 일부 제품과 일부 차원만 확인할 수 있어 소비자들의 입장에서는 객관적인 평가기준을 도출하기 쉽지 않다. 또한 <Figure 2>와 같이 세부 평가에 있어서도 해당 차원(디자인, 착용감)에 대하여 한정된 답변이 이뤄지므로 소비자들의 명확한 평가를 반영하기 어려운 한계점을 갖는다.

본 연구에서 제안하는 모델은 <Figure 3>와 같다. 본 모델은 크게 2단계로 구성되며, 사전에 대



〈Figure 1〉 Coupang Multi-Dimensional Analysis Review

상품평 쓰기 Writing product review ✕

전자, 블루투스 이어폰, SM-
★★★★★ 최고

상품 품질에 대한 경험을 남겨주세요. >

상품 사진을 첨부해보세요. >

Please rate this product in detail.
이 상품을 상세히 평가해주세요.
Did you like the design?
디자인은 마음에 드셨나요?

☐ 아주 마음에 들어요 Very good.
☐ 마음에 들어요 Good.
☐ 마음에 들지 않아요 Not good.

How did you feel when you actually used it?
실제로 사용해보시니 착용감은 어땠나요?

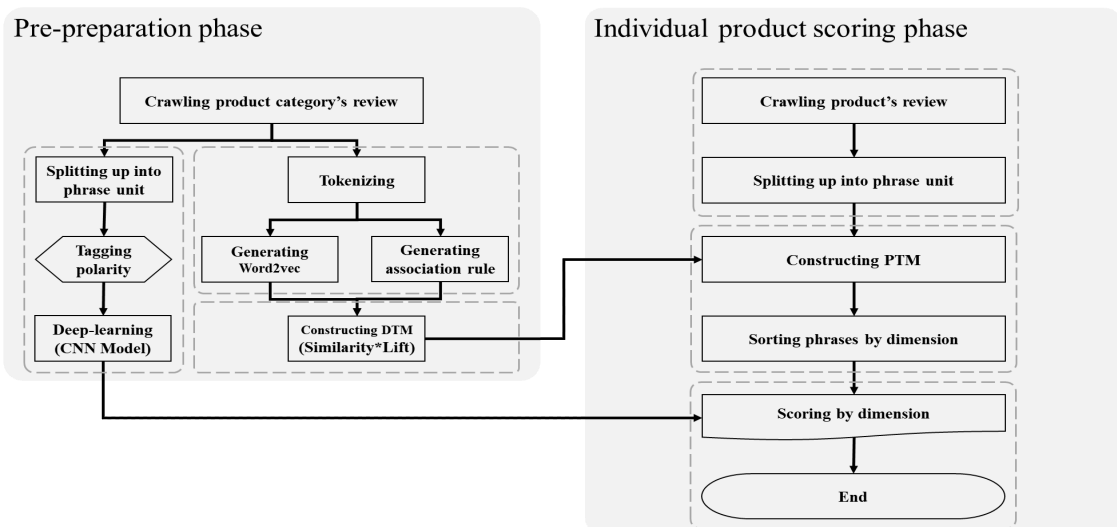
☐ 아주 좋아요 Very good.
☐ 그저 그래요 So so.
☐ 좋지 않아요 Not good.

음질에 만족하셨나요?

상품평 쓰기완료 다음단계 >

〈Figure 2〉 Coupang Multi-Dimensional
Analys Method

분류 상품군 리뷰를 수집하여 모델을 구축하고, 사용자가 비교하고자 하는 개별상품 리뷰에 모델을 적용하여 다차원 상품평가 점수를 얻게 된다. 사전준비 단계에서는 대분류 상품군의 리뷰를 통해 차원분류모델과 감성분석모델을 생성한다. 차원분류모델은 리뷰를 형태소 단위로 나눠 Word2Vec모델을 만들고 차원들에 대한 연관 분석을 실시한다. 차원과 단어 간의 관련성은 앞서 구한 Word2Vec의 유사도(similarity) 값과 연관 분석에서 얻어진 향상도(lift) 값을 곱하여 DTM(dimension term matrix)을 구축한다. 감성분석 모델에서는 구(phrase) 단위의 리뷰에 긍정, 부정, 중립을 태깅하고 이를 학습시킨 딥러닝 모델을 생성한다. 생성된 모델들은 개별상품평가 단계에 적용되어 차원분류 및 스코어링을 진행하고 차원별로 구들의 감성을 스코어링하는데 사용된다. 최종적으로 차원별 스코어를 종합하여 이를 시각화 하는 것이 제시하는 방법의 전체적인 구조이다.



〈Figure 3〉 System Architecture

3.1 사전준비 단계

3.1.1 대분류 상품군 리뷰 수집

기본적인 학습 모델을 만들기 위해 이커머스 사이트에서 리뷰를 수집한다. 본 연구에서는 예시적으로 특정 이커머스 사이트의 ‘노트북’ 상품군을 대상으로 실험을 진행하였으며, 이를 위해 노트북이 속한 ‘디지털/가전/컴퓨터’라는 대분류에 속한 여러 제품들에 대한 리뷰를 수집했다. 각 중분류별로 약 2~5만개씩의 리뷰를 수집했고 중복된 리뷰를 제거한 결과 총 약 26만개의 리뷰를 수집하여 학습을 위한 데이터로 사용한다. 리뷰 수집을 위한 방법으로 파이썬의 BeautifulSoup와 Selenium을 사용하여 크롤링(crawling)하였다. 제품의 리뷰와 별점을 가져오고 별점이 4, 5점이면 긍정, 3점이면 중립, 1, 2점이면 부정이란 감정을 주도록 한다.

3.1.2 감성분석모델 생성

문장으로 이뤄진 리뷰들을 구(phrase)로 나누고 각 구(phrase)들이 어떤 감성을 가지고 있는지 판단하기 위해 본 연구에서는 CNN 딥러닝 기법을 사용했다. CNN은 Convolutional Neural Network의 약자로, 합성곱 신경망 기법이라고도 불린다. 해당 기법은 학습할 때 입력 문장으로부터 핵심적인 N-gram 피처를 추출하는 능력을 갖고 있어 문장의 잠재적인 의미 표현(semantic representation)을 만들게 된다. 이는 리뷰를 분석할 때에도 유용한 방법으로 사용된다.

감성분석에 앞서 수집한 리뷰를 사용하여 CNN모델을 학습시켰다. 별점을 이용한 감성(sentiment)점수를 사용하여 긍·부정에 대한 균형화작업을 실시한 후 6:2:2의 비율로 훈련용 집합,

검증용 집합, 테스트 집합으로 나누어 학습을 진행하였다. 하지만 리뷰를 문장 단위로 CNN모델에서 학습을 시킨 결과, 구(phrase)를 분석했을 때 정확한 결과가 나오지 않았다. 이를 해결하기 위해 약 2만여개의 구(phrase)에 대하여 긍정과 부정, 중립을 직접 태깅하는 작업을 진행하여 성능을 향상시킬 수 있었다.

3.1.3 차원분류모델 생성

(1) 형태소 분석

크롤링으로 수집한 리뷰에서 중복을 제거한 약 26만개의 리뷰 데이터를 차원분류모델을 생성하는데 활용한다. 이를 위해 처리속도가 빠른 분석기 중 하나인 Mecab을 사용하여 리뷰를 형태소 단위로 분석한다.

(2) Word2Vec모델 생성

본 연구에서 Word2Vec모델을 사용한 이유는 각 리뷰들이 어떤 차원에 속하는지를 알아보기 위함이다. 가령 ‘디자인이 예쁘다’라는 리뷰는 ‘디자인’이라는 차원에 속하겠지만 ‘외관이 예쁘다’, ‘모양이 좋다’ 등과 같은 리뷰들도 ‘디자인’이라는 차원에 속하도록 하기 위해서 Word2Vec의 유사도(similarity)라는 개념을 도입한다. Word2Vec기법은 기본적으로 단어 간의 유사도를 측정하는 기법으로 각 단어에 벡터 값을 부여한다. 유사도의 기준이 되는 원리는 기본적으로 해당 단어 대신 그 자리에 다른 단어가 들어가도 얼마만큼 유사한가를 측정한다. 가령 ‘과일이 맛있다’에서 ‘과일’이라는 단어 대신에 ‘음식’이라는 단어가 들어가도 어색함이 없기 때문에 유사도가 높게 측정될 확률이 높다. 이런 식으로 단어 간에 위치, 즉 벡터 값을 부여한다. 부여한 벡

〈Table 1〉 Word Similarity Value by Dimension based on Word2Vec

	Design	Price	Delivery	Weight	Speed	Performance
1	모양 / 0.83	금액 / 0.77	배송 / 0.73	두께 / 0.82	채널 / 0.60	품질 / 0.85
2	외관 / 0.74	가격대/0.67	배달 / 0.69	부피 / 0.71	반응 / 0.59	음질 / 0.78
3	색깔 / 0.74	값 / 0.54	배소 / 0.57	크기 / 0.67	해상도/0.58	화질 / 0.65
4	색상 / 0.74	품질 / 0.49	발송 / 0.51	쿠션 / 0.65	응답 / 0.52	음향 / 0.61
5	색감 / 0.62	성능 / 0.48	피드백/0.50	무게 / 0.63	센서 / 0.51	착용감/0.58
6	외형 / 0.60	화력 / 0.41	택배 / 0.49	그립 / 0.58	배터리/0.50	밥맛 / 0.55
7	디자인/0.59	티나 / 0.40	시일 / 0.48	깊이 / 0.57	감도 / 0.50	사양 / 0.54
...	...	...	...	...	...	...

터 값을 기반으로 단어 간의 유사도를 구할 수 있는데 각 차원에 대한 단어 간 유사도는 다음 <Table 1>과 같다.

(3) 연관규칙 생성 및 편차조정

Word2Vec모델에서는 ‘외관’ 또는 ‘모양’은 ‘디자인’이라는 단어와 유사도(similarity)가 높기에 디자인이라는 차원으로 분류할 수 있게 된다. 하지만 단순히 ‘예쁘다’라는 리뷰가 있다면, 이는 디자인이라는 차원으로 분류하기 힘들어진다. ‘예쁘다’라는 단어가 ‘디자인’이라는 단어와 유사도가 낮게 측정되기 때문인데 그 이유는 ‘예쁘다가 예쁘다’라는 리뷰는 성립되지 않기 때문이다.

이 같은 문제를 해결하기 위해 본 연구에서는 연관분석이라는 기법을 결합적으로 활용하였다. 단어 간의 동시 출현의 평가 척도로는 신뢰도나 향상도를 사용할 수 있다. 신뢰도(confidence)는 조건부확률로, 연관성의 강도를 나타내는 척도이지만 단어의 출현 빈도의 영향을 받는 한계를 가진다. 이에 반해는 향상도(lift)는 임의로 출현할 확률 대비 동시에 출현할 확률로, 단어의 출

현 빈도에 독립적이고, 1이라는 절대적인 기준값을 가진다. 예를 들어, ‘디자인’→‘예쁘다’는 규칙의 경우, 신뢰도는 0.077이고, 향상도는 4.274이다. 반면에 ‘배송’→‘빠르다’는 규칙의 경우, 신뢰도는 0.387이나 향상도는 3.589이다. 두 규칙 모두 유용한 연관규칙으로 판단되나, 개별 신뢰도 값만으로는 유용성 여부를 판단하는데 어려움이 있다. 이에 비해서 향상도는 두 경우 모두 1보다 월등히 큰 값을 가지므로 유용한 규칙임을 쉽게 식별할 수 있다. 따라서 본 연구에서는 연관분석 척도 중 향상도를 활용하였다. 즉, 두 단어 사이에 향상도가 높음을 이용해서 ‘예쁘다’라는 리뷰 역시 ‘디자인’이라는 차원에 속하게끔 했다. <Table 2>는 Word2Vec의 유사도(similarity) 값과 연관분석의 향상도(lift) 값을 결합적으로 사용하여 차원을 분류하는 원리를 설명하는 표이다. ‘빠르다’라는 단어는 ‘배송’과 유사도의 낮지만, 높은 향상도 값이 이를 보완하여 ‘배송’이라는 차원에 분류될 수 있게 된다.

본 연구는 각 차원별로 리뷰에서 언급하는 수의 차이로, 차원마다 유사도와 향상도의 값이 편차가 있음을 확인하였고, 이를 보완하기 위하여

〈Table 2〉 Similarity and Lift

Dimension	Design(디자인)		Delivery(배송)	
	↔ shape(외관)	↔ good(예쁘다)	↔ deliver(배달)	↔ fast(빠르다)
Similarity	0.83	0.20	0.67	0.28
Lift	1.009	4.274	1.065	3.589
Similarity × Lift	0.8374	0.8548	0.713	1.0049

‘편차 조정’을 실시하였다. 전체 리뷰에서 수가 적게 언급되는 차원들은 많이 언급되는 차원들보다 향상도의 값이 항상 크게 나오기 때문에, 다수의 리뷰가 높은 향상도 값에 영향을 받아 적게 언급되는 차원들로 분류되는 오차가 존재하였다. 따라서 6가지 차원에 대해 향상도 값 간의 편차를 조정하여 이를 보완하였다. 조정값을 구하는 방법은 표준화를 통하여 향상도 데이터 간 간격이 감소하게 하였다. 표준화(standardization)의 공식은 다음과 같다.

$$z(lift_{dt}) = \frac{lift_{dt} - \overline{lift_d}}{lift_{s_d}} \quad (4)$$

식 (4)에서 $lift_{dt}$ = t 단어의 d 차원의 향상도 값, $\overline{lift_d}$ = d 차원의 향상도 평균, $lift_{s_d}$ = d 차원의 lift 표준편차 값을 의미한다. 즉, 각 단위의 차원별 향상도를 표준화한 값이다.

(4) DTM(Dimension Term Matrix) 구축

본 연구에서는 각 리뷰들이 어떤 차원에 속하는지를 판단하기 위해 DTM(dimension term matrix)을 도입했다. 앞서 수집된 리뷰에서 단어 목록을 구성하고 Word2Vec 모델과 연관분석기법을 적용했다. 이를 토대로, DTM을 구축했으며 각 차원에 대한 단어의 표준화된 향상도*유사도 값을 나타낸다. 차원분류모델로 구축된 DTM은 개별상품평가 단계에 적용되어 구(phrase) 단위로 구성된 리뷰의 단어들과 행렬곱을 통해 구(phrase)가 속한 차원을 판단하는데 활용된다. 아래의 <Table 3>은 각 차원에 대해 단어가 갖는 표준화된 향상도*유사도의 값을 보여주며, 값이 높을수록 해당 차원에 대한 설명을 담고 있을 가능성이 높다.

DTM의 원소 $d_{ij}(i=1, \dots, d(\text{차원의 수}), j=1, \dots, t(\text{단어의 수}))$ 는 다음과 같이 정의된다.

〈Table 3〉 DTM(dimension term matrix)

Dimension/word	Word1	Word2	Word3	...
Price	0.12	0.33	0.25	
Delivery	2.56	0.22	3.12	
Design	1.20	4.56	1.22	
Speed	1.26	2.70	2.66	
Weight	0.22	0.21	0.44	
Performance	0.11	0.15	1.22	

$$d_{ij}=S_{ij}*z(lift_{ij}) \quad (5)$$

여기서 S_{ij} 는 각 차원에 대해 단어가 갖는 유사도(similarity) 값을 나타낸다.

3.2 개별상품평가 단계

(1) 데이터 전처리

개별상품에서 수집된 리뷰를 사전에 생성된 모델로 평가하기 위해서 해당 리뷰 내에 존재하는 문장들을 형태소 분석기 Mecab을 통하여 구로 분해한다. 분해된 구에서 소비자의 평가가 담긴 동사, 형용사 등이 포함된 구만 선별하여 차원분류모델과 감성분석모델을 적용한다.

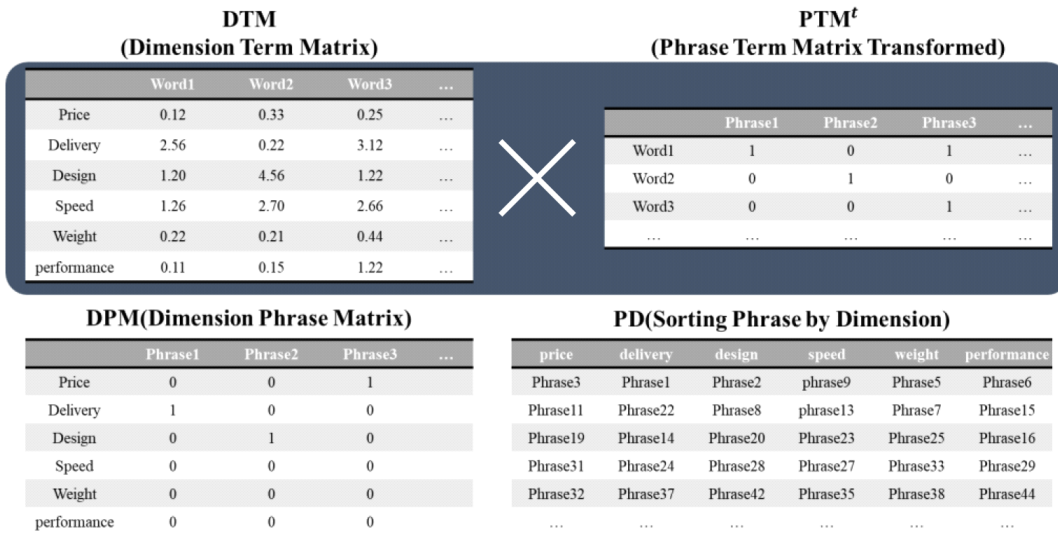
(2) PTM 구축을 통한 차원분류

차원분류를 위해서 선별된 구(phrase)의 단어가 사전준비모델에서 형성된 단어 목록에 있는지 여부를 판단하는 PTM(phrase term matrix)

을 생성한다. PTM은 비교하고자 하는 상품 리뷰들에서 해당 단어가 나오는지 여부를 확인하는 방식이다. 나온다면 1, 나오지 않는다면 0을 주고, 1을 받은 단어들에 대하여 각 차원에 대한 표준화된 향상도*유사도 값을 합산하여 가장 높은 값을 가진 차원으로 지정하는데 활용된다.

차원분류모델에서 단어들과 차원 간의 표준화된 향상도*유사도 값이 매겨진 DTM과 개별상품의 리뷰 상에 단어들로 구성된 PTM의 전치행렬과의 행렬곱을 진행한다. 위 과정을 통해서 각각의 구가 설명하고 있는 차원들을 판별할 수 있으며, 단어의 표준화된 향상도*유사도 값을 종합하여 가장 높은 합을 보인 차원으로 분류된다. 예를 들어, ‘로딩이 빠름’이라는 구는 표준화된 향상도*유사도 값의 합이 6가지 차원 중 ‘속도(speed)’ 차원에서 가장 높게 나왔으므로 ‘속도’ 차원으로 분류된다.

다음 <Figure 4>는 차원분류모델을 통해 개별상품평가모델에 적용됐을 때의 작용을 예시적으로



<Figure 4> Dimension Classification by PTM

로 나타낸 그림이다. DTM과 PTM의 전치행렬의 곱에서 가장 큰 값을 가지는 차원으로 배정되고, 이것은 DPM(dimension phrase matrix) 상의 배정된 차원의 값은 1, 아닌 경우는 0으로 표시된다. 이를 해당 차원으로 정렬함으로써 차원별 리뷰가 구성된다. 구체적으로 DPM은 다음과 같이 계산되며, 아래 식 (8)은 DPM을 통해 d 차원을 설명한다고 판단된 구(phrase)들을 종합하는 방식이다. 이를 토대로 종합된 구의 극성(긍정, 부정, 중립)을 판단하고 차원별 스코어링을 진행하게 된다.

$$p(dpm_{ij}) = dtm_{ik} \times ptm_{kj}^t \quad (6)$$

$$dpm_{ij} = \begin{cases} 1 & \text{when } p(dpm_{ij}) = \max_{k \in \{1, \dots, d\}} p(dpm_{kj}) \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$PD_a = \{p_j | dpm_{aj} = 1, dpm_{aj} \text{ in DPM} \} \quad (8)$$

(3) 스코어링

차원별로 분류된 구를 감성분석모델인 CNN 모델을 이용하여 긍정, 중립, 부정으로 감성분석을 진행한다. 감성분석을 통해 긍정, 중립, 부정 리뷰에 대해 각각 1, 0, -1이 입력되며 시행한 결과로 스코어링을 진행한다. 일례로 전체 리뷰의 개수 중 긍정리뷰가 차지하는 비율을 기준으로 다차원 평가점수를 얻을 수 있다.

4. 실험 설계 및 결과

4.1 실험 내용

제안된 기술에서 리뷰가 어떤 차원에 해당하

는지 분류하는 차원분류모델과 리뷰의 공부정을 판단하는 스코어링 모델의 유용성 평가를 위해 실험을 진행하였다. 평가 데이터는 이커머스 사이트 11번가의 S사 노트북 리뷰 1,011건과 L사 노트북 리뷰 1,062건으로 구성된 구(phrase) 2,001개로 진행하였다.

차원분류모델을 형성하기 위해 대분류 상품군에서 디지털/가전/컴퓨터의 26만 건의 리뷰에 속해 있는 단어들을 바탕으로 Word2Vec모델을 생성하였다. 위 모델을 기반으로 리뷰의 단어들과 6가지 차원인 ‘가격, 디자인, 무게, 배송, 속도, 성능’ 사이의 유사도(similarity)를 구하였다. 이를 연관분석기법을 이용하여 리뷰의 단어들과 6가지 차원간의 형성된 향상도(lift)의 표준화된 값과 곱하여 가중치 행렬을 생성하였다. 각 구에 속해 있는 단어들이 가중치 행렬에서 나타난 값을 합해 그 합이 6가지 차원 중 어떤 차원에서 최댓값을 가지는지를 기준으로 구를 해당 차원으로 분류하였으며, 이를 각 구에 속해 있는 단어와 6가지 차원 간 유사도의 값만 비교하여 차원을 분류하는 모델과 향상도의 값만 비교하여 차원을 분류하는 모델과 정확도를 비교하였다.

감성분석모델은 11번가 대분류 상품군에서 디지털/가전/컴퓨터의 리뷰를 20,080개의 구로 분해한 후, 각 구를 긍정, 중립, 부정으로 태깅하여 학습시킨 구(phrase) 기반 CNN모델이다. 감성분석모델의 정확도를 비교하기 위해서 리뷰를 구로 분해하지 않은 문장전체를 해당 문장이 속해 있는 리뷰가 소비자에게 받은 별점에 따라 긍정, 중립, 부정으로 학습시킨 문장 기반 CNN 모델을 형성하였다. 해당 모델은 별점 5점, 4점을 받은 리뷰에 속해 있는 문장들은 ‘긍정’으로, 별점 3점을 받은 리뷰에 속한 문장들은 ‘중립’으로 별점 2점, 1점을 받은 리뷰에 속해 있는 문장들은 ‘부

정'으로 학습시켰다. 예를 들어, '속도는 그저 그렇지만 디자인이나 색감이 너무 예쁘다!'이라는 리뷰가 별점 5점을 받았다면 해당 리뷰는 문장 기반 모델에서 '긍정'으로 학습된다. 그러나 본 연구의 구(phrase) 기반 모델은 해당 리뷰를 '속도는 그저 그렇지만'과 '디자인이나 색감이 너무 예쁘다'의 두 개의 구로 분해한 후, '속도'에 관해서는 '중립'으로, '디자인'에 관해서는 '긍정'의 의견을 가지고 있는 리뷰로 학습시킨다.

4.2 성과 척도

(1) 정확도(accuracy)

차원분류모델의 성과를 평가하는 방법으로는 정확도를 이용하였다. 정확도는 전체 정답 중 몇 개를 맞혔는가에 집중한다. 평가데이터 2,001개의 구가 6가지 차원 중 어떤 차원으로 분류되는지 사전에 정답을 작성하였다. 그 후 본 차원분류 모델과, 유사도만 고려한 모델, 향상도만 고려한 모델로 차원을 분류하여 차원분류가 정답과 일치하는 비율을 바탕으로 정확도를 평가하였다. 정확도를 구하는 식은 다음과 같다.

$$Accuracy = C(\text{correctly classified phrase})/C(\text{Phrase}) \quad (9)$$

(2) F1 점수

감성분석모델의 성과를 평가하는 방법으로는 F1 점수를 이용하였다. F1 점수는 정밀도(precision)와 재현도(recall)의 조화평균 값이다. 평가데이터 2,001개의 구가 상품에 대해 긍정, 중립, 부정 중 어떠한 평가를 내리고 있는지 사전에 정답을 작성하였다. 그 후 본 모델과, 문장 기반으로 형성된 CNN 모델 각각으로 평가데이터를 긍정, 중

립, 부정 리뷰로 분류하였다. 사전에 작성된 정답과 일치하는 비율을 바탕으로 성과를 평가하였다. F1 점수를 구하는 식은 다음과 같다.

$$F1 \text{ score} = 2 * Precision * Recall / (Precision + Recall) \quad (10)$$

4.3 실험 결과

차원분류 모델 실험 결과 본 연구에서 제시한 방법이 유사도만 고려한 모델과 향상도만 고려한 경우보다 높은 정확도를 보였다. 유사도만 고려한 경우는 정확도가 70%였으며, 향상도만 고려한 경우 정확도가 63%이었다. 한편 유사도와 향상도를 곱하여 함께 고려한 모델은 총 2,001개의 실험데이터 중에서 1,500개의 구의 차원분류를 정확하게 함으로써, 75%의 정확도를 보였다. 최종적으로 편차조정을 적용한 표준화된 향상도 * 유사도의 값을 적용했을 때, 정확도가 85%까지 향상되었음을 알 수 있다.

따라서 리뷰 분석의 과제인 리뷰 데이터를 다차원으로 분류하는 것은 차원과 리뷰 사이의 향상도와 유사도를 동시에 고려하는 단계가 필요함을 보여준다. 나아가 본 모델은 감성분석모델을 형성하는 데에 유의미한 데이터만을 남길 수 있게 한다. 본 모델을 통해서 6가지 차원 중 어떤 차원에도 해당하지 않는 리뷰의 구는 불포함(garbage) 데이터로 분류하여 분석범위의 정확도를 높이기 때문이다. 아래의 <Table 4>는 유사도와 향상도 값을 통한 차원분류 방법의 정확도를 보여주며, 복합적인 활용이 더 유의미했음을 알 수 있다.

감성분석모델 실험결과 기존 연구에서 나타난 문장을 학습데이터로 형성하여 별점으로 학습시켜 구현한 CNN 모델보다 구로 분해하여 각 구

〈Table 4〉 Comparison of Similarity, Lift and Similarity*z(Lift) Accuracy

	Similarity	Lift	Similarity*Lift	Similarity*z(Lift)
The exact number of dimensions	1,416	1,260	1,500	1,680
Accuracy (%)	70.8634	63.0969	75.1567	84.5634

에 대한 긍정, 중립 부정 데이터를 학습시켜 구현한 본 연구의 CNN 모델이 더 높은 정확도를 보였다. 이는 하나의 리뷰에서 다양한 감성을 가지고 있는 소비자의 평가가 별점이라는 하나의 데이터로 분석될 수 없으며, 본 연구와 같이 리뷰에서 유의미한 구문을 추출하여 분석해야 함을 보여주어 본 연구의 유용성을 입증한다. 따라서 리뷰분석의 과제인 리뷰 데이터의 감성 분석은 비정형데이터인 리뷰 데이터를 PTM(phrase term matrix)으로 생성하여 각 차원에 대한 스코어를 산출할 수 있다. 아래 표는 CNN 모델을 통해 리뷰 데이터에 대한 긍정, 중립, 부정을 각각 1, 0, -1로 산출하여 예측값과 실제값을 비교한 결과로, 문장보다 구(phrase) 단위의 학습데이터로 구현한 모델이 55.3%에서 84.7%으로 정확도가 29.4% 향상됨을 확인하였다.

5. 결론

본 연구에서는 이커머스 사이트 내 소비자와 판매자에게 인사이트 제공을 위한 리뷰 파악 다차원 분석 방안을 제시했다. 소비자가 상품을 구매할 때 중요하게 여기는 가격, 성능 등을 차원으로 하는 다차원 분석을 통해 유의미한 정보를 도출하였다. 기존 리뷰 다차원 분석의 대표적 연구 내용인 사전구축 방식이 지닌 한계에서 벗어나기 위해, 별도의 사전 구축 없이 리뷰를 분석하는 방안을 제시하였다. 본 연구는 학습데이터의 리뷰 내에 있는 단어를 차원에 할당하여 사전으로 구축하는 것이 아닌, 워드임베딩 방식으로 차원과 리뷰 내에 있는 단어의 거리 값과 차원과 리뷰 내에 있는 단어의 연관분석을 통한 값을 자동화하여 분류하였다. 한편 기존의 리뷰의 감성

〈Table 5〉 CNN Model Accuracy Improvement

Sentence model					Phrase model				
accuracy	55.3%				84.7%				
True/pred	1	0	-1	total	True/pred	1	0	-1	total
1	867	393	70	1,330	1	1,217	92	21	1,330
0	153	192	99	444	0	110	304	30	444
-1	75	90	62	227	-1	15	51	161	227
total	1,095	675	231	2,001	total	1,342	447	212	2,001

분석을 위해 리뷰 문장 전체를 별점으로 학습시킨 방식과 달리 CNN모델을 기반으로 리뷰 문장을 구로 분해하여 차원으로 분류되는 구와 분류되지 않는 구로 분석 범위를 축소하여 각 구를 긍정, 중립, 부정으로 학습시켰다.

또한 본 연구에서는 제시한 방안의 유용성을 확인하기 위한 실험을 수행하였다. 노트북 리뷰와 차원간의 유사도만 고려한 경우, 연관분석만 고려한 경우를 본 연구의 차원분류 모델과 비교한 결과 유사도와 연관분석의 향상도를 곱하여 함께 고려한 차원분류모델이 분류의 정확도를 높이는데 효과적임을 확인하였다. 한편 본 연구에서 제시하는 방안은 리뷰를 구로 분해하여 이를 긍정, 중립, 부정으로 감성분석을 실시한다. 이는 문장 기반 학습 CNN 모델보다 정확도를 향상시킨다.

연구에서 제시하는 모델은 사용자가 비교하고자 하는 상품들의 리뷰 데이터를 모아 의미 있는 정보 추출을 가능하게 한다. 본 연구는 인공지능이 인간을 대신해 더 많은 온라인 리뷰를 학습하고, 스코어링하여 보여줌으로써 한눈에 상품이 마켓에서 어떠한 평가를 받고 있는지 보여주는 것을 목표로 했다. 따라서 본 연구를 통해 사용자에게 평가차원을 구성하는 세부 내용을 파악하는 데 효과적으로 도움을 줄 수 있는 방안을 제시한다.

제시된 다차원 분석 방안은 이론적 측면에서 다음과 같은 기여점을 가질 것이다. 첫째, 텍스트마이닝 연구에서 비정형 데이터가 갖는 한계를 보완하고 정확도를 높이는 단초를 제공할 것이다. Word2Vec의 워드임베딩과 함께 데이터마이닝에서 사용되는 연관규칙분석을 결합했고, 차원과 단어 간의 유의미한 빈도를 계산하여 정확도를 향상시킬 수 있었다. 다양한 기법들을 연

계하여 더욱 효과적인 접근법을 구상하는데 활용될 가치가 있다. 둘째, 감성분석 연구에 있어 자동화된 시스템을 구상하는데 방향성을 제공할 것이다. CNN 딥러닝 모델을 통해 감성분석을 진행함으로써 사전기반 연구에서 수동작업이 갖는 확장성의 한계를 보완할 수 있었다. 정보의 축적과 함께 대두되는 빅데이터 활용 측면에서도 용이하게 작용할 것이다. 셋째, 텍스트마이닝의 종합적인 연구모델로서 기초연구자료로 활용될 가능성이 높다. 본 연구는 차원분류의 정확도를 높이는 방법뿐만 아니라 한국어의 특질을 고려한 구(phrase) 단위의 감성분석을 시도했으며, 극성탐지 결과를 종합하여 평가차원별 종합점수를 도출하였다. 데이터 전처리에서 종합적인 평가 집계를 통한 스코어까지 제공할 수 있는 연구모델로서 가치가 충분하다고 판단한다.

실무적 측면에서는 이커머스 리뷰 시스템에 다음과 같은 기여점을 가질 것이다. 첫째, 판매자의 입장에서는 비정형데이터인 리뷰를 효과적으로 분석할 수 있게 되며, 중점적으로 다룰 평가차원에 대해 비교 상품과의 경쟁우위를 파악할 수 있다. 이를 통해 시장 인식 조사 및 상품개발 측면에서 방향성에 도움을 받을 수 있다. 둘째, 기존 사전구축 방식보다 용이하며 새로운 리뷰가 다수 형성되는 현 상황에서 별도의 사전을 업데이트 할 필요가 없이 분석을 가능하게 하는 자동화된 방법이다. 셋째, 소비자들로 하여금 자신이 비교하고 싶은 상품의 리뷰를 간편하게 확인할 수 있게 한다. 일일이 읽을 필요없이 분석결과를 바탕으로 빠른 구매의사결정을 내릴 수 있으므로 경제적 비용을 줄일 수 있다.

본 연구와 관련된 한계와 추후 연구과제는 다음과 같다. 본 연구에서 제시한 방법의 유용성을 특정 상품인 ‘노트북’을 한정하여 실험하였다.

따라서 보다 다양한 상품에 대한 검증이 필요하다. 예를 들어 ‘뜨겁다’라는 리뷰에 대해서 본 연구의 ‘노트북’에 대해서는 발열이 심한 제품을 의미하므로 부정적인 의미를 가지고 있지만, ‘전기장판’에 대해서는 발열 효과를 잘 갖고 있는 긍정적인 리뷰로 작용할 수 있다. 그러나 본 연구의 모델로는 ‘노트북’의 리뷰만 학습되어있기 때문에, 전기장판의 ‘뜨겁다’라는 리뷰를 부정적으로 평가할 것이다. 즉, 상품의 속성의 차이를 반영한 확장성 있는 모델의 구현이 필요하다. 또한, 특정 이커머스 사이트뿐만 아니라 리뷰를 다루는 다양한 사이트에서도 이를 적용하여 범용적으로 발전시킨다면 더욱 효과적인 모델로서 자리매김할 것이다.

Reference

- Henry, A., P. Nigel, B. Linda and V. Kevin, “Consumer Behavior - A Strategic Approach. Boston”, M. A.: Houghton Mifflin Company, (2004).
- Byun, S., D. Lee and N. Kim, "Methodology for Identifying Issues of User Reviews from the Perspective of Evaluation Criteria: Focus on a Hotel Information Site", *Journal of Intelligence and Information Systems*, Vol.22, No.3(2016), 23~43.
- Cho, H., S. Kim, J. Lee and J. S. Lee, "Data-Driven Integration of Multiple Sentiment Dictionaries for Lexicon-Based Sentiment Classification of Product Reviews", *Knowledge-Based Systems*, Vol.71, No.8(2014), 61~71.
- Cho, S. G. and S. B. Kim, “Finding Meaningful Pattern of Key Words in IIE Transactions Using Text Mining”, *Journal of Korean Institute of Industrial Engineers*, Vol.38, No.1(2012), 67~73.
- Choi, J., H. Ryu, D. Yu, N. Kim and Y. Kim, “System Design for Analysis and Evaluation of E-Commerce Products Using Review Sentiment Word Analysis”, *KIISE Transactions on Computing Practices*, Vol.22, No.5(2016), 209~217.
- Collobert, R. and J. Weston, *A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning*, In Proceedings of the 25th international conference on Machine learning, 2008. Available at <https://dl.acm.org/doi/abs/10.1145/1390156.1390177> (Accessed 8 March, 2020).
- Gamon, M., A. Aue, S. Corston-Oliver and E. Ringger, “Pulse: Mining Customer Opinions from Free Text”, *In international symposium on intelligent data analysis*, Vol.6, No.12(2005), 121~132.
- Goldberg, Y. and O. Levy, *Word2Vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method*, Cornell University, 2014. Available at <https://arxiv.org/abs/1402.3722v1> (Accessed 15 March, 2020).
- Hu, H. W., Y. L. Chen and P. T. Hsu, “A Novel Approach to Rate and Summarize Online Reviews according to User-Specified Aspects”, *Journal of Electronic Commerce Research*, Vol.17, No.2(2016), 132~152.
- Jeong, H. H., J. Y. Choi and J. Y. Park, "The Effect of Online Consumer Ratings and Quantity of Reviews on Visiting Intention: Focusing on the Types of Restaurant", *Journal of distribution research*, Vol.24, No.4(2019), 1~21.

- Jeong, J., K. H. Mo, S. Seo, C. Y. Kim, H. Kim and P. Kang, "Unsupervised Document Multi-Category Weight Extraction based on Word Embedding and Word Network Analysis: A Case Study on Mobile Phone Reviews", *Journal of the Korean Institute of Industrial Engineers*, Vol.44, No.6(2018), 442~451.
- Ju, M. and S. Youn, "Sentiment Analysis on Movie Reviews Using Word Embedding and CNN", *Journal of the Korea Society of Digital Industry and Information Management*, Vol.15, No.1(2019), 87~97.
- Kim, B. J. and E. R. Hwang, "The Effect of Product Evaluation Information on the Consumer Decision Making in the On-line Context: The Word of Mouth Effect of Product Review and Consumer Replies", *The Korean Academic Society of Business Administration*, (2007), 1~26.
- Kim, G. and C. Lee, "Korean Movie Review Sentiment Analysis Using Convolutional Neural Network", *In Proc. of the KIISE Korea Computer Congress*, (2016), 747~749.
- Kim, J. M. and J. H. Lee, "Text Document Classification based on Recurrent Neural Network Using Word2Vec", *Journal of Korean Institute of Intelligent Systems* Vol.27, No.6 (2017), 560~565.
- Kim, K. H., B. W. Whang and M. C. Kim, "Algorithm Mining Association Rules by Considering Weight Support", *The KIPS Transactions: PartD*, Vol.11, No.3(2004), 545~552.
- Kim, M., J. Byun, C. Lee and Y. Lee, "Multi-Channel CNN for Korean Sentiment Analysis", *Annual Conference on Human and Language Technology*, (2018), 79~83.
- Kovacs, J., "Travel websites usher in new era of guest reviews", *Hotel & Motel Management*, Vol.223, No.19(2008), 29~55.
- Lee, G., J. Jeong, S. Seo, C. Kim and P. Kang, "Sentiment Classification with Word Localization based on Weakly Supervised Learning with a Convolutional Neural Network", *Knowledge-Based Systems*, Vol.152, No.7(2018), 70~82.
- Lee, J., "How eWOM reduces uncertainties in decision-making process: using the concept of entropy in information theory", *The Journal of Society for e-Business Studies*, Vol.16, No.4(2011), 241~256.
- Liu, B., M. Hu and J. Cheng, "Opinion Observer: Analyzing and Comparing Opinions on the Web", *In Proceedings of the 14th international conference on World Wide Web*, (2005), 342~351.
- Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado and J. Dean, "Distributed Representations of Words and Phrases and Their Compositionality", *In Advances in neural information processing systems*, Vol.26, No.347(2013), 3111~3119.
- Myung, J. S., D. J. Lee and S. G. Lee, "A Korean Product Review Analysis System Using a Semi-Automatically Constructed Semantic Dictionary", *Journal of KIISE:Software and Applications*, Vol.35, No.6(2008), 392~403.
- Park, H. J., M. C. Song and K. S. Shin, "Sentiment Analysis of Korean Reviews Using CNN: Focusing on Morpheme Embedding", *Journal of Intelligence and Information Systems*, Vol.24, No.2(2018), 59~83.
- Robinson, R., T. T. Goh and R. Zhang, "Textual Factors in Online Product Reviews: A Foundation for a More Influential Approach

- to Opinion Mining”, *Electronic Commerce Research*, Vol.12, No.3(2012), 301~330.
- Scaffidi, C., K. Bierhoff, E. Chang, M. Felker, H. Ng and C. Jin, “Red Opal: Product-Feature Scoring from Reviews”, *In Proceedings of the 8th ACM conference on Electronic commerce*, (2007), 182~191.
- Seo, D. S., K. H. Mo, J. S. Park, G. C. Lee and P. S. Kang, “Word Sentiment Score Evaluation based on Graph-Based Semi-Supervised Learning and Word Embedding”, *The Journal of the Korean Institute of Industrial Engineers*, Vol.43, No.5(2017), 330~340.
- Silver, L., *Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally*, Pew Research Center, 2019. Available at <https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/> (Accessed 28 December, 2019).
- Son, S. B. and J. H. Chun, “Product Feature Extraction and Rating Distribution Using User Reviews”, *Journal of Society for e-Business Studies*, Vol.22, No.1(2017), 65~87.
- Song, J. S. and S. W. Lee, “Automatic Construction of Positive/Negative Feature-Predicate Dictionary for Polarity Classification of Product Reviews”, *Korean Institute of Information Scientists and Engineers*, Vol.38, No.3(2011), 157~168.
- Song, M., H. Park and K. S. Shin, "Attention-Based Long Short-Term Memory Network Using Sentiment Lexicon Embedding for Aspect-Level Sentiment Analysis in Korean", *Information Processing & Management*, Vol.56, No.3(2019), 637~653.
- Tadano, R., K. Shimada and T. Endo, “Multi-Aspects Review Summarization based on Identification of Important Opinions and Their Similarity”, *In Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation*, Vol.26, No.80(2010), 685~692.
- Xu, H., Y. Zhang and R. DeGroof, “A Feature-Based Sentence Model for Evaluation of Similar Online Products 1”, *Journal of Electronic Commerce Research*, Vol.19, No.4 (2018), 320~335.
- Yeon, J. H., D. J. Lee, J. H. Shim and S. G. Lee, “Product Review Data and Sentiment Analytical Processing Modeling”, *The Journal of Society for e-Business Studies*, Vol.16, No.4(2011), 125~137.
- Yu, S. U. and K. H. Kim, “Analysis of Tourism Reviews by Using Multi-Dimensional Ranking Model”, *Korean Institute of Information Scientists and Engineers*, Vol.33, No.3(2017), 3~14.

Abstract

Multi-Dimensional Analysis Method of Product Reviews for Market Insight

Park Jeong Hyun* · Lee Seo Ho** · Lim Gyu Jin*** ·
Yeo Un Yeong**** · Kim Jong Woo*****

With the development of the Internet, consumers have had an opportunity to check product information easily through E-Commerce. Product reviews used in the process of purchasing goods are based on user experience, allowing consumers to engage as producers of information as well as refer to information. This can be a way to increase the efficiency of purchasing decisions from the perspective of consumers, and from the seller's point of view, it can help develop products and strengthen their competitiveness. However, it takes a lot of time and effort to understand the overall assessment and assessment dimensions of the products that I think are important in reading the vast amount of product reviews offered by E-Commerce for the products consumers want to compare. This is because product reviews are unstructured information and it is difficult to read sentiment of reviews and assessment dimension immediately. For example, consumers who want to purchase a laptop would like to check the assessment of comparative products at each dimension, such as performance, weight, delivery, speed, and design.

Therefore, in this paper, we would like to propose a method to automatically generate multi-dimensional product assessment scores in product reviews that we would like to compare. The methods presented in this study consist largely of two phases. One is the pre-preparation phase and the second is the individual product scoring phase. In the pre-preparation phase, a dimensioned classification model and a sentiment analysis model are created based on a review of the large category product group review. By combining word embedding and association analysis, the dimensioned classification model

* School of Business, Hanyang University

** School of Business, Hanyang University

*** School of Business, Hanyang University

**** School of Business Informatics, Hanyang University

***** Corresponding author: Kim Jong Woo

School of Business, Hanyang University

222 Wangshimni-ro, Seongdong-gu, Seoul 04763, Korea

Tel: +82-2-2220-1067, Fax: +82-2-2220-1169, E-mail: kjw@hanyang.ac.kr

complements the limitation that word embedding methods for finding relevance between dimensions and words in existing studies see only the distance of words in sentences. Sentiment analysis models generate CNN models by organizing learning data tagged with positives and negatives on a phrase unit for accurate polarity detection. Through this, the individual product scoring phase applies the models pre-prepared for the phrase unit review. Multi-dimensional assessment scores can be obtained by aggregating them by assessment dimension according to the proportion of reviews organized like this, which are grouped among those that are judged to describe a specific dimension for each phrase.

In the experiment of this paper, approximately 260,000 reviews of the large category product group are collected to form a dimensioned classification model and a sentiment analysis model. In addition, reviews of the laptops of S and L companies selling at E-Commerce are collected and used as experimental data, respectively. The dimensioned classification model classified individual product reviews broken down into phrases into six assessment dimensions and combined the existing word embedding method with an association analysis indicating frequency between words and dimensions. As a result of combining word embedding and association analysis, the accuracy of the model increased by 13.7%. The sentiment analysis models could be seen to closely analyze the assessment when they were taught in a phrase unit rather than in sentences. As a result, it was confirmed that the accuracy was 29.4% higher than the sentence-based model. Through this study, both sellers and consumers can expect efficient decision making in purchasing and product development, given that they can make multi-dimensional comparisons of products. In addition, text reviews, which are unstructured data, were transformed into objective values such as frequency and morpheme, and they were analysed together using word embedding and association analysis to improve the objectivity aspects of more precise multi-dimensional analysis and research. This will be an attractive analysis model in terms of not only enabling more effective service deployment during the evolving E-Commerce market and fierce competition, but also satisfying both customers.

Key Words : Association Analysis, Sentiment Analysis, E-Commerce, Word embedding, Text Mining, Polarity Detection

Received : April 1, 2020 Revised : May 26, 2020 Accepted : June 2, 2020

Publication Type : Regular Paper Corresponding Author : Kim Jong Woo

저 자 소개



박정현

현재 한양대학교 경영학부, 컴퓨터 소프트웨어학부에 재학중이다. 한양대학교 대학혁신 지원사업 비즈니스 랩 기반 프로젝트 학기제 인턴 과정에서 Business AI Lab에 근무하였다. 주요 연구 관심분야는 금융 산업에서의 기계학습 및 딥러닝 기법 활용, 데이터마이닝 등이다.



이서호

현재 KTDS 경영서비스본부 BI정보서비스팀에서 재직 중이다. 한양대학교 경영학사를 취득하였다. 한양대학교 대학혁신지원사업 비즈니스 랩 기반 프로젝트 학기제 인턴 과정에서 Business AI Lab에 근무하였다. 주요 연구 관심분야는 데이터 분석을 통한 비즈니스에서의 활용, 기계학습 및 딥러닝 기법 활용 등이다.



임규진

현재 한양대학교에서 경영학과 사회혁신융합전공을 배우고 있다. 한양대학교 대학혁신 지원사업 비즈니스 랩 기반 프로젝트 학기제 인턴 과정에서 Business AI Lab에 근무하였다. 주요 연구 관심분야는 데이터마이닝, 딥러닝 기법 활용 및 응용 등이다.



여운영

현재 (주)SK에서 Tech. Pre-Sales그룹에 재직 중이다. 한양대학교 일반대학원 비즈니스인포매틱스학과에서 경영학석사를 취득하였다. 주요 연구 관심분야는 자연어처리, 자연어 이해, 딥러닝 기법 활용 및 응용 등이다.



김종우

현재 한양대학교 경영대학 경영학부 교수로 재직 중이다. 서울대학교 수학과에서 학사를 마쳤으며, 한국과학기술원에서 경영과학으로 석사학위를, 산업경영학으로 박사학위를 취득하였다. 주요 연구 관심분야는 데이터마이닝 기법과 응용, 기계학습과 딥러닝, 오피니언 마이닝, 상품추천기술, 지능형 정보시스템, 집단지성, 사회 네트워크 분석, 클라우드 컴퓨팅 서비스 등이다.