# 상품 평가 텍스트에 암시된 사용자 관점 추출

장경록<sup>0</sup>, 이강욱, 맹성현 한국과학기술원

kyoungrok.jang@kaist.ac.kr, chaximeer@kaist.ac.kr, myaeng@kaist.ac.kr

# Extracting Implicit Customer Viewpoints from Product Review Text

Kyoungrok Jang<sup>o</sup>, Kangwook Lee, Sung-Hyon Myaeng Korea Advance Institute of Science and Technology

#### 요 약

온라인 소비자들은 amazon.com과 같은 온라인 상점 플랫폼에 상품 평가(리뷰: review) 글을 남김으로써 대상 상품에 대한 의견을 표현한다. 이러한 상품 리뷰는 다른 소비자들의 구매 결정에도 큰 영향을 끼친 다는 관점에서 볼 때, 매우 중요한 정보원이라고 할 수 있다. 사람들이 남긴 의견 정보(opinion)를 자동으로 추출하거나 분석하고자 하는 연구인 감성 분석(sentiment analysis)분야에서 과거에 진행된 대다수의 연구들은 크게는 문서 단위에서 작게는 상품의 요소(aspect) 단위로 사용자들이 남긴 의견이 긍정적 혹은 부정적 감정을 포함하고 있는지 분석하고자 하였다. 이렇게 소비자들이 남긴 의견이 대상 상품 혹은 상품의 요소를 긍정적 혹은 부정적으로 판단했는지 여부를 판단하는 것이 유용한 경우도 있겠으나, 본 연구에 서는 소비자들이 '어떤 관점'에서 대상 상품 혹은 상품의 요소를 평가했는지를 자동으로 추출하는 방법에 초점을 두었다. 본 연구에서는 형용사의 대표적인 성질 중 하나가 자신이 수식하는 명사의 속성에 값을 부여하는 것임에 주목하여, 수식된 명사의 속성을 추출하고자 하였고 이를 위해 WordNet을 사용하였다. 제안하는 방법의 효과를 검증하기 위해 3명의 평가자를 활용하여 실험을 하였으며 그 결과는 본 연구 방향이 감성분석에 있어 새로운 가능성을 열기에 충분하다는 것을 보여주었다.

주제어: 전자상거래, 감성 분석, 상품 속성 추출, 상품 리뷰 E-commerce, Sentiment Analysis, Aspect Extraction, Product Reviews

### 1. 서론

웹은 사람들이 모여 다양한 주제에 대한 의견을 교환 하는 장소가 되었다. 사용자들이 웹상에서 상품이나 서 비스에 대해 표현한 의견을 담고 있는 평가 텍스트는 전 자상거래 분야에서 특히 중요한데, 그 이유는 그러한 의 견들이 다른 온라인 소비자들의 구매 결정에 직접적인 영향을 미치기 때문이다. 웹의 개방적인 특성 덕분에 소 비자들이 온라인상에 남기는 '상품 리뷰'를 담고 있는 텍스트의 양이 급격하게 증가하였고, 자연스럽게 상품 리뷰에 담긴 방대한 양의 의견 정보를 자동으로 분석하 방법이 필요해졌다. 감성 분석 (sentiment analysis)[3, 4, 7, 8, 13, 14, 16]은 이러한 필요에 맞 춰 대두된 연구 분야로서, 텍스트 형태로 표현된 의견 정보를 자동으로 추출하고 분석하는 것을 목표로 한다. 주로 텍스트에 포함된 의견이 상품을 긍정적이거나 부정 적으로 판단했는지 여부를 판단하는 것을 판단하게 된 다.

감성 분석 연구의 초기에는 개별 '문서' 단위별 분석에서 시작하였으나, 최근에는 상품의 특성 (aspect) 단위로 분석하고자 하는 특성 기반 감성 분석(aspect-based sentiment analysis) 연구가 많은 주목을받기 시작했다. 특성 기반 감성 분석은 소비자가 대상상품의 어떤 특성 (예: 화면, 외관)에 대해 긍정적이거나 부정적 의견을 표했는지를 분석하는 것을 목표로 한다. 예를 들어 "이 스마트폰의 화면은 밝지만, 배터리수명이 너무 짧다"라는 문장에서, 두 가지의 다른 의견

이 대상 스마트폰의 두 가지 특성에 대해 표현되었다. 화면(특성)은 밝다는 긍정적 의견과 배터리 수명(특성) 은 짧다는 부정적 의견이 같이 존재하므로 특성 별 판단 이 중요하다. 같은 상품이라도 각 특성 별로 표현된 평 가는 다를 수 있기 때문에, 대상 상품의 각 특성 별로 어떤 의견이 표현되었는지를 인식하는 것은 매우 중요하다.

비록 현재로선 특성 기반 감성 분석 [4, 6, 7, 10, 15]이 가장 상세한 수준의 분석 방식이지만, 이 분석 방식에도 여전히 한계는 남아 있다. 대부분의 관련 연구들이 대상 측면에 대해 표현된 의견의 극성 (polarity)을 판별하는데 그쳤기 때문이다. 물론 이러한 형태의 분석이 유용한 경우도 있겠지만 좀 더 상세한 형태의 분석도가능하다. 예를 들어 "이 엔진은 시끄럽다"라는 문장에서, 우리는 의견을 표현한 소비자가 엔진을 부정적으로 평가하고 있다는 것뿐만 아니라, 엔진을 소음이라는 '관점 (viewpoint)'에서 평가하고 있다는 것 역시 알수 있다. 소비자들은 같은 상품이라도 다양한 관점에서 평가하기 마련이므로, 이러한 관점들을 파악하는 것은 소비자들이 왜 상품에 대해 긍정적이거나 부정적인 의견을 표현하는지를 이해하는데 도움을 줄 수 있다.

본 논문에서는 사람들이 상품을 평가하기 위해 표현한 의견에 암시된 관점을 추출하는 문제를 정의하고 그 방법을 제안한다. 본 논문에 등장하는 '대상 (target)'이라는 용어는, 상품 그 자체(예: '자동차')를 나타내거나 상품의 특정 aspect(예: '엔진')를 가리키기 위해

사용된다. 본 연구에서 제안하는 방식은 의견이 나타난 문장에 포함된 형용사로부터 관점 (예: 소음)을 추출하는 것인데, 이는 형용사가 의견을 표현하기 위해 가장흔히 사용되는 품사이기 때문이다. 의견에 암묵적으로 표현되는 관점을 추출하게 되면 사람들이 상품의 어떤 측면에 관심을 기울이고 평가하였는지를 이해할 수 있게되어 다양한 의견 텍스트에 존재하는 평가 정보를 요약하는데 도움이 된다 (그림 1).

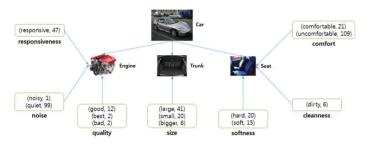


그림 1 관점(viewpoint) 기반 감성 분석

언어학분야에서 수행된 연구에 따르면, 형용사가 가진 유일한 역할은 그것이 수식하는 명사의 특정 속성 (attribute)에 어떠한 값 (value)을 부여하는 것이다 [11]. 예를 들어 "이 카메라는 무겁다"라는 문장에서, 형용사인 '무겁다'는 그것이 수식하는 명사인 카메라의 속성 중 하나인 '무게'에 값을 부여하고 있다. 본연구에서는 이러한 속성이 사람들이 상품을 평가할 때 가졌던 관점과 일치한다는 가설을 세웠다. 즉 사람들이 '무겁다'라는 형용사를 사용하여 의견을 표현한 이유는, 그들이 카메라를 '무게'의 관점에서 평가하고자하였기 때문인 것이다. 따라서 우리가 제안하는 태스크는 형용사가 가지고 있는 값을 사용해서 수식하고 있는 대상이 가지고 있는 속성을 추출하는 것이라고도 말할 수 있다.

형용사가 값을 부여하는 속성을 추출하기 위해, 본 연 구에서는 거대한 영문 어휘 리소스인 워드넷 (WordNet) [12]을 활용하였다. 워드넷에서 모든 형용사 synset은 2,500개 이상의 클러스터로 구성되어 있으며, 각각의 클 러스터들의 중심에는 head synset이라는 특수한 유형의 synset이 한 쌍 존재한다. 그 한 쌍의 head synset은 서 로 반의어 관계를 가지고 있어 반대말 쌍(antonymous pair)이라고 불리며, 해당 클러스터에 속한 형용사들이 값을 부여하는 속성의 양 극값 (bipolar values)을 나타 낸다. 예를 들어, 'small-large'로 이루어진 반대말 쌍에서는 'size'라는 속성의 양 극값을 나타내며, 이 반대말 쌍을 중심으로 그와 유사한 의미를 가진 형용사 synset (satellite synset)들이 뭉쳐 하나의 클러스터를 이루고 있다. 따라서 같은 클러스터에 포함된 모든 형용 사 synset들은 모두 양 극값 중 하나를 나타내고 있다. 따라서 형용사가 값을 부여하는 속성을 알아내기 위해서 는, 먼저 해당 형용사가 어떤 클러스터에 속하는지를 알 아내고, 그 클러스터를 대표하는 head synset을 인식한 다음, 마지막으로 그 head synset이 나타내는 속성을 추 출하면 된다. head synset이 나타내는 속성은 워드넷 상 에 코딩되어 있어 쉽게 알아낼 수 있다.

하지만 워드넷 상에서 모든 head synset이 자신과 연 관된 속성을 가지고 있지는 않다. 워드넷 상에 있는 2.500개 이상의 형용사 클러스터 중 오직 620개의 head synset들만이 연관된 속성을 가지고 있다 (워드넷 3.0 기준), 이 커버리지 문제를 해결하기 위한 방법으로, 사 용하려고 하는 head synset과 연관된 속성이 워드넷에 없는 경우 그와 유사한 다른 head synset에서 대신 연관 된 속성을 추출하도록 하였다. 워드넷에서 유사한 head synset끼리는 'also see'라는 관계를 통해 서로 연결 되어 있으므로 이 관계를 활용하면 된다. 또 다른 문제 는 같은 형용사라 할지라도 그것이 나타내는 의미 (sense 혹은 synset)가 여러 개 있을 수 있기 때문에 여 러 개의 head synset과 연관되어 있을 수 있다는 것이 다. 해당 형용사의 어떤 head synset을 선택하느냐에 따 라 최종적으로 추출되는 속성(attribute)이 달라지므로 사용된 의미에 맞는 head synset을 선택하는 것을 매우 중요하다. 본 연구에서는 사람들이 상품에 대한 의견을 표현하기 위해 형용사들을 사용한 만큼 극성 점수 (polarity score)가 가장 높은 의미로 형용사를 사용했 을 것이라고 간주 하였고, 그러한 의미를 나타내는 head synset을 선택하는 것이 속성 추출의 정확도에 기여할 것이라고 가정하였다. 형용사의 각 의미가 가진 극성 점 수를 계산하는 데는 SentiWordNet [1]을 활용하였다. SentiWordNet은 워드넷에 포함된 모든 synset들에 극성 점수를 부여한 영문 리소스이다.

본 연구에서는 위에서 설명한 방식으로 형용사로부터 추출한 '속성'이, 실제로 사람들이 상품에 대해 표현한 의견에 암시된 '관점'과 일치하는지를 검증하고자하였다. 평가를 위해 110개의 테스트 케이스들이 생성되었고 각 테스트 케이스는 상품을 평가한 문장과, 그 문장에 포함된 형용사로부터 추출한 속성으로 구성되어 있다. 우리는 평가자들에게 생성한 테스트 케이스들을 보고 "추출된 속성이 의견 문장에 암시된 관점과 일치하는가?"를 질문하였다. 실험 및 평가 결과는 본 논문에서 제안하는 관점 추출 태스크가 쉽지는 않지만 충분히발전 가능성이 있다는 점을 보여주고 있다.

### 2. 관련 연구

본 연구는 특성 기반 감성 분석 [7, 10]과 관련되어 있다. 해당 태스크는 '몇 퍼센트의 사람들이 타겟을 긍정적이거나 부정적으로 평가했는지'와 같은 식으로 의견을 '정량적'으로 요약하는 것을 목표로 한다. 이런수준의 분석도 물론 유용하지만, 만약 우리가 "사람들이어떤 관점에서 타겟을 평가하였는지"라는 질문에 대답할 수 있다면, 사람들의 의견에 대해 더욱 자세하게 이해할 수 있을 것이다. 우리가 아는 한 우리의 연구는 이러한 질문에 답하고자 한 최초의 연구이다.

암시된 특성 검출 (implicit aspect extraction) 연구 [4, 6, 15] 역시 우리의 연구와 관련되어 있다. 해당 태스크는 텍스트 속에서 명시적으로 나타나지 않았지만 의견 속에 암시된 특성을 추출하는 것을 목표로 한다. 예를 들어 우리는 "이 가방은 무겁다"라는 의견 문장 속

에 '무게'라는 타켓이 암시되어 있다는 것을 알 수 있다. 본 논문에서 추출하고자 하는 관점 역시 문장에 명시적으로 나타나지 않은 암시적인 것이지만, 우리가 말하는 '관점'은 상품의 특성과는 다르다. 관점은 "사람들이 상품을 어떻게 평가하는가?"와 관련이 있는 개념이며, 그러므로 이는 일반적으로 말하는 '상품의 스펙'과는 다른 의미를 지닌다. 이런 측면에서 볼 때, 우리의 태스크는 사람들이 타켓을 어떤 관점에서 바라보고 있는지를 이해하는데 더 적합하다고 할 수 있다.

형용사들로부터 속성을 추출하기 위해, 우리는 [4]에서 사용된 것과 유사한 접근방식을 사용하였다. 이 연구는 우리의 연구와 마찬가지로 형용사로부터 그것이 값을 부여하는 속성을 추출하는 것을 목표로 하였는데, 워드넷이라는 단일 리소스만 이용해서는 커버리지가 충분치않을 수 있다는 점에 주목하였다. 속성 추출 커버리지를 증가시키기 위해, 해당 연구의 저자들은 복수의 온라인사전 리소스를 활용하는 방식을 제시한다. 그들이 제시한 방식은 크게 3개의 단계로 나눠진다:

- 1. 복수의 온라인 사전으로부터 형용사의 '주해 (glosses)'를 가져온다.
- 2. 주해에 포함된 모든 명사를 추출하고, 추출된 명사들을 형용사의 속성 후보들로 간주한다.
- 3. 추출된 속성 후보들의 사전적 관계 (예: 동의어/ 반의어 관계)를 이용하여, 후보들이 실제로 속성을 나타내는지 아닌지를 자동으로 분류한다.

우리의 연구는 다음과 같은 점에서 그들의 것과 다르다: (1) 그들은 형용사로부터 암시된 상품의 특성을 추출하는 것을 목표로 하였다. 반면, 우리 연구의 목표는 사람들이 표현한 의견 속에 암시된 관점을 추출하고자하였다. 의견 문장에 의견의 타겟이 명시적으로 나타났는지 아닌지 여부와는 상관이 없다. (2) 그들의 연구가최대한 많은 형용사로부터 속성을 추출하는데 초점을 둔반면, 우리 연구는 형용사가 값을 부여하는 속성이 의견에 암시된 관점과 일치하는지를 검증하고자 하였다.

## 3. 제안된 방법

이 단원에서는 먼저 워드넷에서 형용사들이 어떻게 구성되어 있는지에 대해 설명한다. 그 다음, 관점 추출에 필요한 단계를 설명한다. 마지막으로, 우리가 관점 추출의 커버리지와 정확도를 향상시키기 위해 취했던 조치들에 대해 설명한다.

## 3.1 워드넷의 형용사 구성

워드넷에는 descriptive, relational 이렇게 두 가지유형의 형용사가 있다 [11]. 본 연구에서 우리는 descriptive 형용사에 초점을 맞추었는데, 이 유형의 형용사들이 우리가 원하는 형용사의 특성, 즉 그것이 수식하는 명사의 속성에 값을 부여하는 특성을 가지고 있기때문이다. 예를 들어 "이 가방은 작다"라는 문장에서, '작다'라는 형용사는 가방의 '크기' 속성에 값을 부여하고 있다. 다음은 가방의 '크기' 속성에 '작다'라는 값을 부여하는 것을 함수 형태로 표현한 것이다.

크기(가방) = 작다

워드넷에서 형용사들은 '반의어 관계 (antonymy)'를 중심으로 조직되어 있는데, 이는 대부분의 경우 형용사 가 값을 부여하는 속성은 양 극의 값 (bipolar values) 을 갖기 때문이다 [11]. 예를 들어, '작다'의 반의어 는 '크다'이고, '작다-크다'는 함께 '무게'(속성) 의 양 극값을 나타낸다. 이러한 반의어 쌍을 워드넷에서 는 antonymous pair라고 부르며, antonymous pair를 구 성하고 있는 synset들을 head synset이라고 부른다. (그 림 2)는 형용사 클러스터가 어떻게 구성되어 있는지를 보여준다. 워드넷에는 이와 같은 형용사 클러스터가 2,500여개 있고, 각 클러스터의 중심에는 앞서 설명한 대로 antonymous pair가 있다. 그 주위로 head synset과 유사한 의미를 가진 satellite synset들이 모여 하나의 클러스터를 구성하게 된다. 워드넷 상에서는 오직 head synset들만이 연관된 '속성'을 가지고 있으며, 따라서 특정 형용사로부터 속성을 추출하기 위해선 먼저 해당 형용사가 속한 클러스터의 head synset을 찾아야만 한 다.

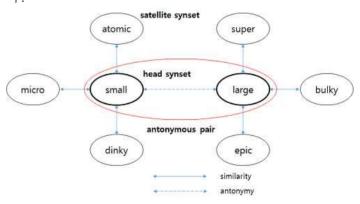


그림 2 형용사 클러스터

### 3.2 관점 추출 과정

형용사로부터 속성을 추출하기 위해서는 먼저 대상 형용사의 head synset을 알아내야 하는데, 그 때 발생하는 두 가지 이슈는 다음과 같다: (1) 찾아낸 head synset과 연관된 속성이 워드넷에 없을 때의 대처법, (2) 형용사가 여러 개의 head synset과 연관되어 있을 경우, 그들중 하나를 선택하는 법. 본 단원에서는 형용사로부터 추출하는 과정과 각 이슈에 대처하는 법에 대해 자세히 설명한다.

# 3.2.1 전처리

먼저 형용사와 그것이 수식하는 명사를 판별하기 위한 품사 (POS) 태깅을 수행한다. 그리고 우리는 '의견 문장'을 '형용사와 그것이 수식하고 있는 명사를 포함하고 있는 문장'이라고 정의하고, 품사 태그로 구성된 {\NN\times VB\times JJ\times} 패턴을 사용하여 그러한 문장들만을 골라 추출한다. 예를 들어, 앞에서 설명한 패턴을 이용하면 "engine is quiet"와 같은 문장이 추출된다. 문장에 포함된 형용사는 의견을 나타내며, 명사는 타겟을 나타낸다. 또한 우리는 텍스트에 포함된 노이즈를 줄이기 위해, 추출된 의견(형용사) 및 타겟(명사) 단어에 대해 스펠링 체크 및 lemmatization을 수행하였다.

# 3.2.2 형용사의 head synset 판별

형용사로부터 속성을 추출하기 위해선, 먼저 형용사의 head synset이 무엇인지를 알아내야 한다. 그러기 위해 우리는 대상 형용사를 lemma로 가지고 있는 head synset을 먼저 찾는다. Synset의 lemma는 그 synset이 나타내는 의미를 가지고 있는 구체적인 단어를 지칭하며, 따라서 어떤 형용사가 한 head synset의 lemma라는 것은, head synset이 나타내는 의미를 그 형용사가 가지고 있다는 것을 의미한다. 이 과정에서 워드넷을 사용하기 위해, 우리는 Natural Language Toolkit (NLTK) [2]에서 제공하는 워드넷 인터페이스를 활용하였다.

## 3.2.3 형용사의 head synset 선택

하나의 형용사는 여러 개의 sense를 가질 수 있으며, 그런 경우 여러 개의 head synset과 연관되게 된다. 본 논문에서 우리는 속성 추출에 사용할 head synset을 선 택하는 두 가지 정책을 제시한다.

Top synset: 가장 자주 사용되는 head synset을 선택하는 정책이다. 워드넷에서 head synset은 사용 빈도순으로 정렬되어 있기 때문에, 가장 자주 사용되는 head synset을 쉽게 판별할 수 있다.

Polar synset: 가장 높은 극성 (polarity)를 가지고 있는 head synset을 선택하는 정책이다. 이 정책은 사람들이 상품을 평가하기 위해 형용사를 사용한 만큼, 높은 극성 (polarity)을 가진 의미로 형용사를 사용하였을 것이라는 가정에 기반하고 있다. head synset의 극성 점수를 계산하기 위해, 우리는 SentiWordNet [1]을 활용하였다. SentiWordNet는 워드넷의 모든 synset에 대한 객관성 점수 (objective score)를 계산해 놓은 리소스이다. 객관성 점수는 synset이 얼마나 객관적인 의미인지를 점수화해서 표현한 것으로, synset의 객관성 점수가 낮을수록 synset의 극성 점수가 높다.

### 3.2.4 속성 추출

바로 전 단계에서 선택된 하나의 head synset에서 연 관된 속성을 추출한다. 워드넷에서 head synset은 자신 이 연관된 속성과 'attribute'라는 관계로 연결되어 있다. 이 attribute 관계를 이용해서 head synset과 연 관된 속성을 가져오면 된다.

#### 3.2.5 유사한 head synset의 속성 추출

워드넷에 있는 모든 head synset이 그와 연관된 속성을 가지고 있지는 않다. 저자들이 파악한 바에 따르면 오직 620개의 head synset들만이 연관된 속성을 가지고 있다. 워드넷에 2,500개 이상의 형용사 클러스터가 있고, 각 클러스터가 반의어 관계에 있는 한 쌍의 head synset을 중심으로 가지고 있다는 점을 감안하면, 워드넷의 형용사 속성 커버리지는 낮은 편이라고 할 수 있다. 형용사의 속성 추출률을 높이기 위해, 우리는 워드넷에서 head synset 간에 존재하는 'also see'관계를 활용하였다. also see 관계는 서로 다른 클러스터에 있지만, 관련된 의미를 가진 head synset를 사이를 연결하고 있다. 우리는 형용사의 head synset과 연관된 속성이 없는 경우, also see 관계를 이용해서 다른 유사한 head synset을 찾은 후, 그 head synset에서 다시 연관된 속성을 추출하도록 하였다. (그림 3)이 그 예시이다.

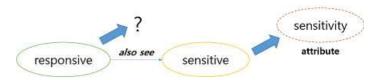


그림 3 유사한 head synset의 속성 추출

(그림 4)는 전체적인 속성 추출 과정을 그림으로 보여 준다.

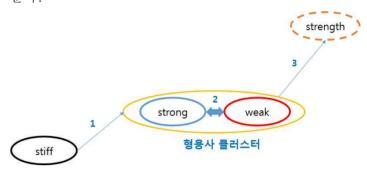


그림 4 속성 추출 과정

여기서 번호 1은 형용사('stiff')가 속한 클러스터를 알아내는 과정, 번호 2는 클러스터의 중심을 이루고있는 한 쌍의 head synset인 antonymous pair ('strong-weak')를 찾는 과정, 마지막으로 번호 3은 head synset과 연관된 속성 ('strength')를 찾는 과정을 나타내고 있다.

### 4. 실험 결과

실험에는 Ganesan과 ChengXiang의 연구 [5]에서 사용된 데이터셋이 활용되었다. 정확하게는 전체 데이터셋 중 자동차 리뷰 데이터셋만이 활용되었는데, 데이터셋에는 edmund.com에서 수집한 2007년, 2008년, 2009년 식자동차들에 대해 남겨진 리뷰들이 포함되어 있다. 각 년도 별로 140개에서 250개까지의 자동차가 있으며, 전체자동차 리뷰의 개수는 42,230개이다 [5].

우리는 분석하고자 하는 타겟이 이미 주어져 있다고 가정하고 실험을 진행하였다. 우리는 11개의 타겟을 데 이터셋에서의 출현 빈도수에 기반하여 선정하였다. 타겟 에는 자동차 자체를 나타내는 car가 포함되어 있고, 자 동차의 10가지 특성이 포함되어 있다. 선정된 11개 타겟 의 리스트는 다음과 같다 (그림 5): car, door, cabin, engine, navigation, seat, suspension, trunk, transmission, stereo, system.

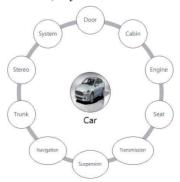


그림 5 타겟 리스트

분석할 타겟을 선정한 다음, 우리는 3.2.1 전처리 단원에서 설명한 패턴을 이용해서 데이터셋으로부터 타겟이 포함된 '의견 문장'및 의견 문장에 포함된 형용사를 추출하였다. 예를 들어 대상 타겟이 engine이라면, 패턴을 이용하여 데이터셋으로부터 "engine is quiet"라는 문장을 추출하고, 이 문장에 포함된 형용사인 'quiet'역시 함께 추출한다. 그 다음, 우리는 각 타겟을 수식하는 형용사들을 출현 빈도순으로 정렬하였다. 이 단계를 거치는 이유는 각 타겟을 평가할 때 주로 사용되는 형용사가 무엇인지를 파악하기 위해서이다. 최종적으로 각타겟 별 테스트 케이스를 생성할 때에는 빈도수 기준으로 top 10 형용사만을 선택하여 사용하였다.

형용사로부터 속성을 추출하기 위해, 우리는 3.2 단원에서 소개한 추출 방식들을 조합하여 총 네 가지의 알고리즘을 고안하였다:

- 1. top\_head: 가장 자주 사용되는 head synset으로부 터 속성 추출
- 2. top\_head\_similar: 가장 자주 사용되는 head synset으로부터 속성 추출 & 실패할 경우 그와 유사 한 head synset으로부터 속성 추출 시도
- 3. polar\_head: 가장 높은 극성을 가진 head synset 으로부터 속성 추출
- 4. polar\_head\_similar: 가장 높은 극성을 가진 head synset으로부터 속성 추출 & 실패할 경우 그와 유사한 head synset으로부터 속성 추출 시도

위에서 제시한 네 가지 알고리즘의 성능을 비교하기 위해, 우리는 각 알고리즘별로 총 110개의 테스트 케이 스로 이루어진 테스트 셋을 생성하였다. 각 테스트 케이 스는 다음과 같은 세 개의 요소를 포함하고 있다:

- 의견 문장: "engine is responsive"
- 알고리즘이 추출한 속성: sensitivity
- 속성의 정의: the ability to respond to physical stimuli

(표 1)은 타겟 중 하나인 engine에 대해 생성된 테스트 케이스 일부를 보여주고 있다.

의견 문장	관점	관점 정의	
engine is smooth	evenness. n.02	a quality of uniformity and lack of variation	
engine is powerful	strength. n.01	the property of being physically or mentally strong	
engine is great	quality.n. 01	an essential and distinguishing attribute of something or someone	

표 1 테스트 케이스 예시 ( 'engine' )

평가를 위해, 우리는 세 명의 평가자에게 형용사로부터 추출된 속성이 의견 문장에 암시된 관점과 일치하는 지를 평가해달라고 요청하였다. 측정 기준으로는 precision, recall, F1 점수와 더불어, 평가 agreement를 알아보기 위한 fleiss' kappa [9] 점수를 사용하였

다. 실험 결과는 (표 2)에 나타나 있다.

알고리즘	precision	recall	F1	kappa (n=3)
top_head	0.6000	0.8182	0.6923	0.7033
top_head_similar	0.6182	0.8727	0.7237	0.6347
polarity_head	0.6455	0.8909	0.7486	0.5989
polarity_head_similar	0.6636	0.9182	0.7704	0.5927

표 2 실험 결과

실험 결과는 비록 우리가 제안한 태스크가 쉽지는 않 지만, 발전 가능성이 있음을 보여주고 있다. 평가자들의 agreement 정도가 그리 높지 않다는 점에서 볼 때 이 태 스크가 쉽지 않다는 것을 알 수 있다. 계산된 kappa 점 수는 0.5에서 0.7사이로, 평가자들 사이의 agreement 수 치가 보통에서 꽤 높음 사이라는 것을 의미한다. 사용한 방식 중 polarity\_head 방식이 top\_head 방식보다 높은 precision과 recall을 보였다는 것은 우리의 가설 - 사 람들은 형용사를 극성이 높은 의미로 사용했을 것이다 -이 합리적이었다는 것을 증명해준다. 더불어, 속성 추출 에 실패할 경우 유사한 head synset으로부터 또다시 추 출을 시도하는 정책은 precision 및 recall 향상에 기여 했다는 것을 알 수 있다. recall 뿐만 아니라 precision 이 함께 향상된 이유는, 만약 한 알고리즘이 속성 추출 에 실패했을 경우, 다른 알고리즘과의 비교를 위해 그 알고리즘은 무조건 틀렸다고 간주되었기 때문이다.

### 5. 결론

본 논문에서는 의견에 암시되어 있는 상품 평가의 관점을 추출하는 태스크를 정의하고 하나의 방법론을 제안하였다. 의견 문장에 포함된 형용사로부터 추출된 속성이 사용자가 평가할 때 가졌던 관점과 일치한다는 가정하에, 워드넷을 활용하여 형용사로부터 속성을 추출하는 방법을 제안하였고, 평가자들을 통해 그 둘이 실제로 일치하는지를 평가하였다. 세 명의 평가자들로부터 받은평가 결과는 이러한 접근 방식이 발전 가능성이 있지만,여전히 개선할 여지가 많다는 것을 보여주었다. 보다 정확한 속성 추출을 위해선 단어의미 중의서 해소방법이도입되어야 한다고 예상하고 있으며 이를 향후 연구로수행 해나갈 예정이다.

### 참고문헌

- [1] Baccianella, S. et al. 2010. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. Proceedings of the 7th conference on International Language Resources and Evaluation (LREC' 10) (2010), 2200-2204.
- [2] Bird, Steven, E.L. and E.K. 2009. Natural Language Processing with Python. O' Reilly Media Inc.
- [3] Dave, K. et al. 2003. Mining the peanut gallery: Opinion Extraction and Semantic Classification of Product Reviews. Proceedings of the twelfth international conference on World Wide Web (WWW' 03) (New York, New York, USA, May. 2003),

519.

- [4] Fei, G. et al. 2012. A Dictionary-Based Approach to Identifying Aspects Implied by Adjectives for Opinion Mining. COLING (Posters)' 12. 2, December 2012 (2012), 309-318.
- [5] Ganesan, K. and Zhai, C. 2011. Opinion-based entity ranking. Information Retrieval. 15, 2 (Aug. 2011), 116–150.
- [6] Hai, Z. et al. 2011. Implicit Feature Identification via Cooccurrence Association Rule Mining. Computational Linguistics and Intelligent Text Processing SE 31. 6608, (2011), 393–404.
- [7] Hu, M. and Liu, B. 2004. Mining and summarizing customer reviews. Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining (KDD'04) (New York, New York, USA, Aug. 2004), 168.
- [8] Hu, M. and Liu, B. 2004. Mining opinion features in customer reviews. Proceedings of the 19th national conference on Artifical intelligence (Jul. 2004), 755–760.
- [9] Joseph, F.L. 1971. Measuring nominal scale agreement among many raters. Psychological bulletin. 76, 5 (1971), 378-382.
- [10] Liu, B. et al. 2005. Opinion Observer: Analyzing and Comparing Opinions on the Web. Proceedings of the 14th international conference on World Wide Web (WWW' 05) (New York, New York, USA, May. 2005), 342.
- [11] Miller, G.A. et al. 1993. Introduction to WordNet: An Online Lexical Database. International Journal of Lexicography . 3 , August (Dec. 1993), 235–244.
- [12] Miller, G.A. 1995. WordNet: A Lexical Database for English. Communications of the ACM. 38, 11 (1995), 39–41
- [13] Moghaddam, S. and Ester, M. 2010. Opinion digger. Proceedings of the 19th ACM international conference on Information and knowledge management (CIKM' 10) (New York, New York, USA, Oct. 2010), 1825.
- [14] Pang, B. et al. 1988. Thumbs up? Sentiment Classification sing Machine Learning Techniques. Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10. (1988), 79-86.
- [15] Su, Q. et al. 2008. Hidden Sentiment Association in Chinese Web Opinion Mining. Proceedings of the 17th international conference on World Wide Web (WWW' 08) (New York, NY, USA, 2008), 959-968.
- [16] Zhuang, L. et al. 2006. Movie Review Mining and Summarization. Proceedings of the 15th ACM

international conference on Information and knowledge management (CIKM' 06)(New York, New York, USA, 2006), 43.