

상품평의 언어적 분석을 통한 상품 평가 요약 시스템

이 우 철* · 이 현 아** · 이 공 주***

요 약

본 논문에서는 폭발적으로 증가하고 있는 상품평을 효과적으로 활용하기 위해 언어적 분석을 통하여 상품 평가를 요약하는 시스템을 제안한다. 시스템에서는 스커트 상품 분류의 경우 '디자인'과 '재질'과 같이, 상품을 평가하는 기준이 되는 항목에 대한 상품평의 부정과 긍정의 극성 여부를 판별하여 그래프 형태로 요약하여 제시한다. 본 논문에서는 작은 시드 어휘와 문맥에 기반한 자동 확장 방법을 사용하여 평가 항목 별 평가 어휘 극성 사전을 구축하여 평가 항목에 대한 상품평의 극성을 판정한다. 제안한 방식은 여러 온라인 쇼핑몰의 실제 상품평에 대한 실험에서 극성 사전 추출에서 평균 69.8%의 정확율과 문장별 극성 식별에서 평균 81.8%의 정확율을 보였다.

키워드 : 평가 항목, 상품평 요약, 극성 판별, 감성 분류, 전자상거래

Product Evaluation Summarization Through Linguistic Analysis of Product Reviews

Woo Chul Lee* · Hyun Ah Lee** · Kong Joo Lee***

ABSTRACT

In this paper, we introduce a system that summarizes product evaluation through linguistic analysis to effectively utilize explosively increasing product reviews. Our system analyzes polarities of product reviews by product features, based on which customers evaluate each product like 'design' and 'material' for a skirt product category. The system shows to customers a graph as a review summary that represents percentages of positive and negative reviews. We build an opinion word dictionary for each product feature through context based automatic expansion with small seed words, and judge polarity of reviews by product features with the extracted dictionary. In experiment using product reviews from online shopping malls, our system shows average accuracy of 69.8% in extracting judgemental word dictionary and 81.8% in polarity resolution for each sentence.

Keywords : Product Features, Product Review Summarization, Polarity Resolution, Sentiment Classification, Electronic Commerce

1. 서 론

상품을 직접 살펴볼 수 없는 인터넷 쇼핑에서 기존 구매자가 작성한 상품평은 구매 예정자의 구매 결정에 영향을 미치는 중요 정보원이다. 상품평의 높은 구매 영향력에 의해 인터넷 쇼핑몰에서는 기존 구매자로부터 가능한 많은 수의 상품평을 획득하기 위해 다양한 유인책을 사용하고 있다. 실제 구매자에 의해 작성되는 상품평의 양적인 증가는 상품평의 객관성을 보장하여 근래에는 상품평만을 참고하여 구매 여부를 결정하는 경우가 늘고 있다. 국내의 경우 이용자의 94.3%는 다른 이용자의 구매 경험과 평가를 기반으로

최종 구매 여부를 결정하고 있으며, 59%는 쇼핑 후 상품평 및 이용 후기 등을 작성함으로써 온라인 구전의 형성과 확산에 참여하고 있다[1].

인터넷 쇼핑 시장의 급속한 성장과 상품평의 양적 증가에 의하여 근래에는 하나의 상품에 천여 개 이상의 상품평이 기술되는 경우들이 늘어나 이를 일일이 읽어 상품 구매 여부를 판단하기 어려워졌다. 대부분의 쇼핑몰에서 별점 등의 형태로 요약된 선호도를 제시하고 있으나, 상품평에 맞지 않는 별점이 매겨진 경우가 많고, 상품의 다양한 특성에 대한 평가를 하나의 점수 체계로 표현하여 구매에 필요한 정확한 정보를 파악하기 힘들다. 영어권에서는 상품평에 대한 다양한 연구들이 이루어졌으나 대량의 언어 지식을 필요로 하는 등 국내 실정에 적합하지 않으며, 국내에서의 상품평에 대한 최근 연구들 역시 수동이나 반자동 지식 구축을 전제로 하고 있다.

본 논문에서는 한국어 상품평을 자동으로 요약하기 위한

* 본 연구는 금오공과대학교 학술연구비에 의하여 연구된 논문
† 정 회 원 : (주)유승토타일솔루션 대리
** 중 심 회 원 : 금오공과대학교 컴퓨터공학부 조교수(교신저자)
*** 정 회 원 : 충남대학교 전기정보통신공학부 부교수
논문접수: 2009년 11월 23일
수 정 일 : 1차 2009년 12월 28일
심사완료: 2009년 12월 28일

시스템을 제안한다. 시스템에서는 평가 항목, 즉 상품의 다양한 특성별로 상품평을 긍정과 부정 중 어느 극성(polarity) 인지 판별하여 사용자가 원하는 정보를 알아보기 쉽게 제시한다. 시스템은 1) 상품평을 수집하여 정제한 뒤, 2) 상품평 내 어휘 정보와 웹 검색을 활용하여 평가 항목을 추출하고, 3) 언어처리기법을 이용하여 평가 항목별 평가 어휘와 평가 어휘의 극성을 자동으로 추출하여, 4) 평가 항목별 극성을 그래프 형태로 요약하여 출력한다. 각 단계는 상품평과 웹 검색 결과를 활용하여 모두 자동으로 처리된다.

본 논문은 다음과 같이 구성된다. 2장에서 상품평 요약에 대한 기존 연구를 살펴보고, 3장에서는 본 연구의 상품평 요약 방법을 설명한다. 4장에서는 실험 결과를 보이며 5장에서 결론 및 향후 연구에 대해 논의한다.

2. 기존 연구

텍스트를 분석하여 의견을 추출하고, 추출된 의견에 긍정과 부정 극성을 부여하여 구조화하는 연구는 오피니언 마이닝(Opinion Mining) 또는 문서 감성 분류(Document Sentiment Classification)로 대표된다. 감성 분류의 초기 연구는 문서 수준[2-4]이나 문장 수준[5, 6]에서의 접근이 대부분이었으나, 근래에는 의견 추출 대상의 다양한 특징별로 극성을 판별하는 특징 기반 요약(Feature Based Summarization)에 대한 연구가 활발하게 진행되고 있다. 상품평에서의 의견 추출에서는 세부적인 상품 특징 - 예를 들어, 스커트의 경우 '사이즈', 디지털카메라의 경우 '액정' - 에 대한 기존 구매자의 평가가 구매 예정자의 구매 결정에 중요한 판단 요소가 될 수 있어 특징 기반 요약이 필요하다. 특징 추출에서는 주로 통계적 기법을 사용하며 TF-IDF 기반 방식, 연관마이닝 기반 방식, PMI(Point-wise Mutual Information) 기반 방식 등이 있다[14].

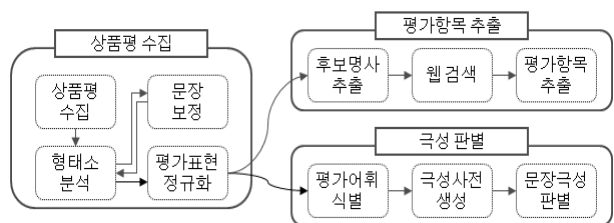
특징 추출을 통해 얻어진 상품의 각 특징을 기반으로 상품에 대한 평가를 요약하기 위해서는 각 특징에 대한 평가가 긍정적인지 부정적인지 분석해야 한다. 이를 위해서는 "디자인이 좋다"의 경우 '좋다'와 같이 평가에 사용되는 어휘, 즉 평가 어휘의 극성 판별이 필요하다. 문장이나 어휘의 극성을 판별하는 기존 연구에는 기계학습을 이용한 방법, WordNet을 이용한 방법, PMI를 이용한 방법, 자연어처리기법을 결합한 방법 등이 있다. 기계학습을 이용한 방법[3, 4]에서는 별점이나 추천과 비추천 값을 가지는 상품평을 학습 데이터로 사용하여 극성을 예측하고 있으나, 별점 데이터의 신뢰도가 떨어지는 국내 쇼핑몰의 상품평에는 적용하기 어렵다. WordNet을 이용한 방법[7, 8]에서는 WordNet의 유의어, 반의어 관계를 이용하여 형용사의 극성을 판별하는데, 이 방식은 평가 항목에 따라 극성이 달라지는 형용사에 대한 처리가 불가능하다. 예를 들어 디지털카메라의 경우 '작다'가 크기에 대한 평가 어휘라면 긍정적인 표현이지만 액정에 대한 평가 어휘라면 부정적인 표현이다. 또한, 한국어에서는 WordNet과 같은 언어 자원을 실용적으로 사용하기 어렵기 때문에 이 방식은 한국어 상품평 처리에 그대로 적

용하기 힘들다. PMI를 이용한 방식[2]에서는 평가 어휘가 될 수 있는 형용사 또는 부사를 추출하고, PMI를 이용하여 각 단어의 의미극성은 판단하며, 자연어처리기술을 결합한 방식에는 구 수준에서 문맥상 극성 식별을 시도한 연구[9]와 사전의 용어 주해를 이용한 연구[10], 형용사와 접속사 간 연결 정보를 이용한 연구[11] 등이 있다. 이러한 방식들 역시 평가 항목에 따라 달라지는 평가 어휘의 극성을 처리할 수 없으며, 일부 연구에서는 지도 학습에 기반하여 대량의 학습 데이터를 요구하는 문제점이 있다. 또한 영어권의 연구의 경우 정확한 문법과 구조화된 형태로 작성된 긴 길이의 상품평에 대한 접근이 많아 비문법적 단문 형태의 상품평 위주인 국내 인터넷 쇼핑몰에 적용하기 어렵다.

상품평에 대한 영어권의 활발한 연구에 비해 상품평에 대한 국내의 연구는 시작 단계에 있다. Appraisal theory에 근거하여 작성된 Semantic Clause 형태의 극성 사전을 정의하여 한국어 상품평을 분석하고자 하는 연구[12]는 극성 사전의 수동 구축 비용과 의미 분류 객관성의 문제 등이 존재한다. 감정 자질을 이용한 연구[13]에서는 영어 극성 정보를 한국어로 변환하여 초기 감정 자질을 획득하고 사전을 이용한 자질 확장을 시도하였으나, 해결되지 않는 자료 부족 문제와 오류 수정을 위해 인간 주석자에 의존하는 문제점이 있다.

3. 한국어 상품평 요약 시스템

본 장에서는 평가 항목에 기준한 한국어 상품평 요약 방법을 제안한다. 요약 과정은 크게 상품평 수집, 평가 항목 추출, 극성 판별의 세 부분으로 나뉜다. (그림 1)은 시스템의 구조와 흐름을 나타낸다. 상품평 수집부에서는 온라인상의 상품평을 수집하고, 구어체, 띄어쓰기 등의 오류를 보정한 후, 형태소를 분석하고 재구성한다. 평가 항목 추출부에서는 평가 항목 후보 명사와 상품과의 연관도 점수를 이용하여 평가 항목을 추출한다. 극성 판별부에서는 평가 항목별 평가 어휘 극성 사전을 생성한 후, 생성된 사전을 이용하여 문장 극성을 판별하고 평가 항목별로 요약한다. 평가 항목 추출에서는 상품평의 어휘 빈도와 웹 검색 결과만을 사용하고, 극성 사전 생성에서는 작은 크기의 초기 극성 사전에 기반한 상품평 내의 문맥 정보를 활용한 사전 자동 확장을 수행하여, 한국어 상품평 처리에서 발생하는 언어 자원 획득의 문제를 최소화시킨다.



(그림 1) 한국어 상품평 요약 과정

3.1 상품평 수집 및 보정

상품평 수집부에서는 온라인상의 상품평을 수집하고, 상

품평의 문장을 보정하여 저장한다. 인터넷에서 수집한 상품평은 깨진 문자, 띄어쓰기와 구어체 등의 다수의 오류를 포함하므로, 올바른 상품평 요약을 위해서는 문장 보정이 필요하다. 시스템의 보정 과정은 아래와 같다.

1) 수집 중 보정

상품평 수집 과정에서 HTML태그를 제거하고 캐릭터 개체(character entity)로 표현된 특수문자, 인코딩 변환 과정에서 깨진 문자들을 원래 문자로 치환한다. 치환 불가능한 문자는 공백 처리한다. 수집된 상품평은 형태소분석기를 이용하여 일차적으로 형태소를 분석한 뒤 다음 단계로 전달한다.

2) 문장 중단 1차 보정

한국어 상품평은 문장 중단 기호가 생략되거나 잘못 사용된 문장들이 많아 문장 구분이 힘들다. 마침표 없이 이모티콘 형태로 종료되는 문장이 특히 많고, 문장 부호를 두 번 이상 반복 사용하는 경우도 흔하다. 시스템에서는 문장 종결로 사용가능한 어구 뒤에 이모티콘 리스트에 포함된 기호 열이나 중복된 문장 부호가 등장하면 기호열을 마침표로 대체하거나 문장 부호 중복을 제거하여 문장 종단을 보정한다.

3) 구어체 표현 정규화

일반인들이 작성한 상품평에는 ‘하네요’에서 어미 ‘네요’를 ‘네욘’, ‘네용’, ‘네염’, ‘네여’ 등으로 표기한 구어체 표현들이 자주 발생하며 이는 형태소 분석에서 잘못된 결과를 내기 쉽다. 문장 기호 바로 앞에 나타나면서 미등록 명사로 태깅된 단어가 ‘여’, ‘염’, ‘영’, ‘용’, ‘욘’으로 종료되는 경우 이를 ‘요’로 치환한다.

4) 띄어쓰기 보정 및 형태소 분석

보정 과정이 끝난 상품평은 자동띄어쓰기 모듈을 이용하여 띄어쓰기를 보정한다. 띄어쓰기 보정이 완료되면 문장 별로 분리하여 형태소 분석과 품사 태깅을 수행한다.

5) 평가 표현 정규화

다양한 형태의 평가 표현을 효율적으로 추출하고 처리하기 위하여 복합적인 형태소로 구성된 표현들을 한 단어로 묶는 정규화를 수행한다.

5-1) 관용적 표현 처리

상품평에서 자주 등장하는 평가 표현들을 정규화한다. 평가 표현은 일반적으로 동사나 형용사로 구성되는데, ‘마음에 든다’의 경우 형태소 분석 결과 ‘마음+에 들+는다’로 분석되어, 형용사나 동사만을 추출하면 동사 ‘들’만 추출되어 정확한 평가 어휘를 얻을 수 없다. 시스템에서는 ‘마음에 들다’, ‘맘에 듭니다’ 등의 표현을 허위(pseudo) 형용사 ‘마음들’을 포함하는 표현으로, ‘별로다’, ‘별로이다’ 등의 표현을 허위 형용사 ‘별로’를 포함하는 표현으로 치환하여 평가 어휘로 추출한다. 부사 ‘별로’도 품사를 허위 형용사로 바꾸어 평가 어휘로 추출한다.

5-2) 보조동사 정규화

‘~하(다)’, ‘~되(다)’, ‘~스럽(다)’ 등의 동사파생접미사나 형용사파생접미사는 앞쪽 명사와 결합하여야 의견 표현이 가능하다(예: 추천+하+다, 세련+되+다, 촌+스럽+다). 이러한 접미사 앞에 단일 명사가 나타나는 경우 이를 결합하여 평가 어휘로 추출한다.

5-3) 부정 표현 반전 처리

상품평에서 부정형으로 사용된 용어는 용어 자체의 극성과 반대의 극성으로 판별해야 한다. 시스템에서는 부정을 의미하는 부사 ‘안’의 수식을 받거나 ‘않다’, ‘아니다’를 보조 용언구에 포함한 용어는 반전 표시한 후 해당 부정어를 제거한다. 예를 들어 ‘크지 않다’, ‘안 좋다’, ‘퍼지는 건 아니다’는 각각 ‘!크다’, ‘!좋다’, ‘!퍼지다’로 치환된다.

3.2 상품 평가 항목 추출

수집과 보정 과정을 거친 상품평에서 평가 항목을 추출한다. 본 연구에서는 상품평 내의 명사의 빈도에 기반한 내부 연관도와 웹 검색 결과를 활용하는 PMI의 외부 연관도를 결합한 PMI-RTF 방식[14]으로 평가 항목을 추출한다. 상품을 평가할 수 있는 평가 항목은 대부분 명사이므로, 상품 분류 카테고리 c 의 전체 상품평(review∈ c)에 나타나는 명사 t_i 를 평가 항목 후보로 보고, t_i 의 평가 항목으로의 점수를 수식 (1)로 계산한다. t_i 의 상품평 내 출현 빈도 $f_{review \in c}(t_i)$ 를 상품평에서 등장한 후보 명사 빈도의 최고값 $MAX(f_{review \in c})$ 으로 나누어 내부 연관도를 산출한다. 검색 엔진에 후보 명사 t_i 를 단일 검색하여 얻어진 페이지수 $f_{web}(t_i)$ 를 분류명 c_{name} 과 후보 명사 t_i 를 결합하여 검색한 결과 $f_{web}(c_{name}, t_i)$ 로 나누어 PMI 값, 즉 외부 연관도를 산출한다. 내부 연관도와 외부 연관도 두 값을 곱하여 얻어지는 PMI-RTF 값의 상위 단어를 평가 항목으로 추출한다.

$$PMI-RTF(t_i, c) = \frac{f_{review \in c}(t_i)}{MAX(f_{review \in c})} \times \frac{f_{web}(c_{name}, t_i)}{f_{web}(t_i)} \quad (1)$$

3.3 극성 사전 생성

극성 사전 생성에서는 평가 항목에 대한 평가 어휘를 추출하고 극성을 결정하여 극성 사전을 구축한다. 상품평에서 구매자의 평가 의견은 주로 형용사와 동사로 표현되므로, 형용사와 동사, 3.1에서 생성한 허위 형용사의 단일 어휘를 평가 어휘의 후보로 사용한다.

평가 어휘는 상품 평가 항목에 따라 그 극성이 달라질 수 있다. 예를 들어 디지털카메라의 경우 ‘작다’라는 평가 어휘는 평가 항목 ‘화면’이나 ‘액정’에 대해서는 부정적인 의견을 표현하는 반면에 ‘크기’에 대해서는 긍정적인 의견을 표현한다. 이와 같이 평가 항목에 따라 극성이 달라지는 평가 어휘를 반영하기 위해서는 각 상품 분류의 평가 항목별로 독립적인 극성 사전이 필요하다. 본 논문에서는 평가 항목별 극성 사전의 필요성을 지적하고 (그림 2)와 같은 구조로 평가 항목 사전을 자동으로 구축한다.

극성 사전은 수동으로 극성을 분류한 소량의 시드(seed) 집합을 이용한 자동 확장 방식으로 생성한다. 시드 집합은 상품 분류와 평가 항목에 상관없이 고유한 극성을 가지는 단어만을 포함한다. 극성 별 시드 집합은 25개 정도의 단어로 구성되며, 각 극성에 대한 시드 목록은 <표 1>과 같다.

시드 집합을 이용한 평가 항목별 극성 사전 생성에서는 문맥을 고려한 자동 확장 방식을 사용한다. 평가 항목이 포함된 문장을 해당 상품 분류 내 전체 상품평에서 검색하고 검색된 문장에서 추출한 평가 어휘 집합에서 긍정과 부정

시드 집합에 포함된 단어의 개수를 기준으로 문장에 포함된 후보 어휘의 극성을 결정한다. 예를 들어 “크기가 작아서 좋아요”의 문장에서 평가 어휘가 될 수 있는 형용사나 동사는 ‘작다’와 ‘좋다’이며, ‘좋다’는 긍정 어휘 시드에 포함되므로 문장 내 시드 어휘 일치 여부에서 긍정이 우세 극성으로 판단하여 극성이 결정되지 않은 ‘작다’를 평가 항목 ‘크기’에 대한 긍정 어휘로 분류한다.

추출 과정에서 역접어(‘~지만’, ‘~으나’, ‘~은데’ 등)가 발견될 경우, 평가 항목과 평가 어휘의 위치를 고려하여 극성을 결정하는 휴리스틱을 적용한다. 역접어에 대한 처리는 역접어로 연결된 두 개의 절이 동일한 평가 항목에 대한 의견을 기술하는지의 여부에 따라 두 가지 경우로 나뉜다. 두 개의 절에 다른 평가 항목이 포함된 경우에는 역접어를 기준으로 문장을 나누어 처리한다. 예를 들어 “크기가 좀 커서 불편하지만 화질은 나름대로 뛰어나서 괜찮아요”라는 문장의 경우, 역접어 ‘~지만’을 기준으로 평가 항목 ‘크기’에 대해서는 ‘크다’와 ‘불편하다’를, ‘화질’에 대해서는 ‘뛰어나다’와 ‘괜찮다’를 평가 어휘로 추출한다. 이와 같은 문장 분리를 통해 ‘크기’에 대한 ‘크다’는, 긍정 시드 ‘괜찮다’에 영향을 받지 않고, 부정 시드 ‘불편하다’에 의해 부정 극성으로 판별할 수 있다. 두 개의 절 사이에 역접어가 포함되어 있지만 평가하는 항목이 하나인 경우는 극성을 반전시킨다. 예를 들어 “기장이 좀 짧아 보이지만 입어보니 잘 맞네요”에서는 평가 항목 ‘기장’ 뒤에 역접어 ‘~지만’이 있지만 역접어 뒤로 다른 평가 항목(명사)이 따르지 않으므로 뒷 절도 ‘기장’에 대해 평가하는 문장으로 판단하며, 이러한 경우 역접어는 극성의 반전을 의미하므로 역접어를 기준으로 앞뒤 평가 어휘가 반대 극성을 가지게 한다. 예제에서는 ‘맞다’가 긍정 시드 사전에 포함되어 있으므로 ‘짧다’는 부정 극성으로 판단한다.

상품평의 모든 문장에 대한 평가 어휘 추출과 극성 판별 결과를 통합하면 동일 어휘가 부정과 긍정의 양쪽 극성 사전에서 모두 포함되는 경우가 발생할 수 있다. 예를 들어 대부분의 문장에서 ‘만족스럽다’는 긍정 어휘와 같이 나타나지만, “가격이 비싸서 만족스러운 정도가 낮기는 하지만”에서는 부정 시드 ‘비싸다’에 의해 ‘만족스럽다’의 극성이 부정으로 판별된다. 이러한 단어의 최종 극성을 결정하기 위해 수식 (2)~(3)을 이용하여 평가 어휘의 극성 점수를 계산한다. 문장별 극성 판별 과정에서 더 많이 판단된 극성일수록 적합할 가능성이 높으며, 평가 항목과의 거리가 멀수록 해당 어휘는 평가 어휘로서의 기능이 떨어진다는 점에 착안하여, 극성 점수는 각 극성으로 판별된 빈도에 비례하고 평가 항목과의 거리값에 반비례하는 값을 가지게 한다.

$$score(p,+) = \frac{f(p^+)}{f(p^+) + f(p^-)} \times \frac{1}{\sum dist(e, p^+)} \quad (2)$$

$$score(p,-) = \frac{f(p^-)}{f(p^+) + f(p^-)} \times \frac{1}{\sum dist(e, p^-)} \quad (3)$$

수식에서 $score(p,+)$ 는 평가 어휘 p 의 긍정 극성의 점수, $score(p,-)$ 는 평가 어휘 p 의 부정 극성의 점수를 표시한다. $f(p^+)$ 는 문장별 판별 과정에서 긍정 어휘로 분류된 빈도를

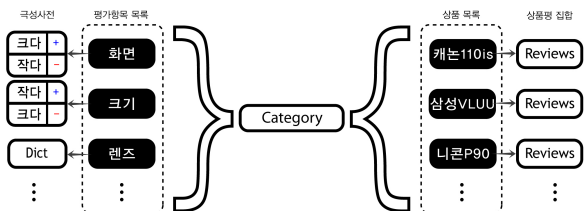
표시하며, $\sum dist(e, p^+)$ 는 긍정 어휘로 분류된 경우 평가 항목 e 와의 거리 값의 합을 나타낸다. 최종적으로 얻어진 긍정 극성 점수가 부정 극성 점수보다 크면 평가 어휘 p 의 극성을 긍정으로, 작으면 부정으로 결정한다. 만일 두 값이 동일하면 거리 정보를 제외하고 각 극성에서 등장한 빈도만을 비교하여 극성을 결정한다. 빈도까지 같은 경우 해당 평가 어휘는 극성 판별에 부적합한 것으로 보고 양쪽 극성 사전 모두에서 제거한다.

3.4 극성 판별/요약

상품평 요약의 마지막 단계에서는 자동 생성된 극성 사전을 적용하여 평가 항목에 대한 평가 극성을 판별한다. 극성 판별은 평가 항목이 포함된 문장 단위로 시행한다. 평가 항목 e 에 대한 문장의 극성은 수식 (4)를 이용하여 문장에 포함된 평가 어휘의 우세한 극성을 계산하여 결정한다. 평가 항목 e 를 포함하는 문장 s 를 상품평에서 추출하고, 문장에 나타나는 평가 어휘 p 의 극성을 극성 사전에서 얻는다. 평가 어휘 p 가 긍정 사전 dic^+ 에 포함된다면 평가 항목 e 로부터 평가 어휘의 거리 $dist(e, p)$ 값의 역수를 더해주고 부정 사전 dic^- 에 포함된다면 거리 값의 역수를 빼는 방식으로 평가 항목에 가까이 있는 평가 어휘의 극성 점수를 높게 부여한다.

$$SO(s, e) = \sum_{p \in dic^+} \frac{1}{dist(e, p)} - \sum_{p \in dic^-} \frac{1}{dist(e, p)} \quad (4)$$

평가 항목 e 에 대한 최종 극성을 판별하기 위해, 수식 (4)의 단일 문장에 대한 극성값 $SO(s, e)$ 의 값이 양수이면 해당 문장의 극성을 긍정으로 보고 평가 항목의 긍정 극성 점수 $score(s, e, +)$ 를 1만큼 증가시키고, 음수이면 평가 항목의 부정 극성 점수 $score(s, e, -)$ 를 1만큼 증가시킨다. 평가 항목 e 의 최종 긍정 극성 점수 $score(e, +)$ 와 부정 극성 점수 $score(e, -)$ 는 수식 (5)와 (6)을 이용하여 상품평에 포함된 모든 문장($s \in review$)에서의 각 극성의 점수의 비율로 결정한다. 시스템에서는 이 값을 (그림 3)의 그래프 형태로 사용자에게 제시한다.



(그림 2) 상품 분류별 극성 사전 구조

<표 1> 평가 어휘 추출을 위한 시드 목록

긍정 시드	좋다, 괜찮다, 마음들다, 만족하다, 적당하다, 무난하다, 깔끔하다, 편리하다, 예쁘다, 편하다, 쉽다, 시원하다, 튼튼하다, 멋지다, 어울리다, 귀엽다, 맛다, 친절하다, 훌륭하다, 우수하다, 꼼꼼하다, 저렴하다, 감쪽하다, 편안하다
부정 시드	아쉽다, 불편하다, 짜증나다, 싫다, 나쁘다, 어렵다, 이상하다, 별로다, 귀찮다, 후회하다, 실망하다, 어정쩡하다, 떨어지다, 지저분하다, 촌스럽다, 모호하다, 답답하다, 어설피다, 불친절하다, 비싸다, 반쯤하다, 애매하다, 다르다, 부담스럽다, 아깝다

$$score(e,+) = \frac{\sum_{s \in review} score(s,e,+)}{\sum_{s \in review} score(s,e,+) + \sum_{s \in review} score(s,e,-)} \quad (5)$$

$$score(e,-) = \frac{\sum_{s \in review} score(s,e,-)}{\sum_{s \in review} score(s,e,+) + \sum_{s \in review} score(s,e,-)} \quad (6)$$



(그림 3) 상품평 요약 출력의 예 (긍/부정 그래프 출력)

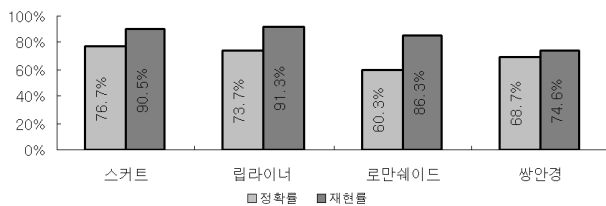
4. 실험 및 평가

(그림 3)은 제안된 방식으로 구축된 한국어 상품평 요약 시스템의 화면을 보인다. 시스템 구축에서는 자동띄어쓰기 모듈[15]과 형태소분석기[16], 네이버 웹문서 검색기를 활용하였다. 제안한 시스템의 성능을 평가하기 위해 극성 사전 생성, 문장 극성 판별, 요약 결과에 대한 사용자 만족도의 세 가지 측면에서 실험을 진행하였다. 실험을 위한 데이터는 여러 온라인 쇼핑몰들의 상품과 상품평을 수집하여 제공하는 가격비교사이트[17]에서 수집되었으며, 실험 데이터에 대한 상세한 정보는 <표 2>와 같다. 실험데이터는 평가 항목 추출 방식[14]에서 사용한 것과 동일하다.

극성 사건의 정확성에 대한 실험은 네 개의 상품 분류에서 무작위로 선택한 6개 평가 항목의 평가 어휘를 대상으로 수행하였다. 해당 평가 항목이 포함된 모든 상품평을 대상으로 평가 어휘 추출과 각 어휘의 극성 판별을 수동으로 수행하여

<표 2> 실험 데이터 상세 정보 (단위 : 개)

상품 분류	상품	상품평	문장	상품평 당 평균 문장 수
스커트	567	2,990	8,748	2.92
쌍안경	82	486	1,406	2.89
립라이너	71	1,477	4,317	2.92
로맨셰이드	157	1,438	4,525	3.05



(그림 4) 극성 사전 구축의 정확율과 재현율

정답 극성 사전을 구축하였다. 수동 구축된 정답 사전과 시스템에서 추출한 사전을 비교하여 정확율과 재현율을 얻었다. 24개의 평가 항목(분류 4개 x 평가 항목 6개)에 대한 극성 사건의 극성 분류 실험 결과 평균 정확율 69.8%, 평균 재현율 85.7%를 보였다. 각 분류별 실험 결과는 (그림 4)와 같다.

극성 사전 추출에서는 부정 사건의 단어수와 재현율이 긍정 사전에 비해 평균적으로 약간 낮은 수치를 보였으며, 이는 상품평에 긍정적 표현이 부정적 표현보다 상대적으로 많기 때문으로 분석되었다. <표 3>은 로맨셰이드 상품의 평가 항목 '사이즈'에 대한 평가 어휘 사건의 예를 보인다. 문장 보정에도 불구하고 문장 종결이 명확하지 않아 평가 항목 '사이즈'는 적합하지 않은 긍정 단어 '화사하다'가 부적합한 부정 극성으로 추출된 경우나, 평가 어휘로 보기 힘든 '생각하다', '받다' 등의 단어가 평가 어휘로 추출된 경우들을 확인할 수 있다.

문장 극성 판별에 대한 실험에서는 스커트 상품 분류의 2개 상품의 5개의 평가 항목에 대해 수동으로 정답 데이터를 구축하여 정확성을 평가하였다. 정답 데이터는 평가 항목을 포함하는 문장을 제시하고, 해당 문장과 문장의 각 평가 어휘가 긍정인지 부정인지 표시하게 하는 수동평가시스템을 통해 구축하였다. 문장 극성 판단 정확도 비교는 시스템의 극성 판단 결과가 사용자의 판단과 일치할 경우 1점, 사용자는 극성이 있다고 판단했으나 시스템은 중립으로 표시한 것은 0.5점을 주었다. 실험 결과에서는 평균 정확율 81.8%를 얻었으며 결과는 <표 4>와 같다. 결과 분석에서는 평가 어휘의 극성 판별 오류에 의한 오류가 대부분이었으며 평가자에 따라 극성 판정이 달라질 수 있는 경우(예를 들어, "길이는 약간 짧은 스타일인데, 디자인이 괜찮으니깐 맘에 들면 구매하세요"에서 '길이'에 대한 극성)로 인한 오류도 다수 포함하는 것으로 나타났다.

마지막으로 시스템에 의해 요약되어 그래프로 출력되는 결과가 사용자에게 얼마나 도움이 되는지를 알아보기 인터넷 쇼핑을 자주하는 피실험자 7명을 대상으로 만족도 평가를 수행하였다. 평가에서는 스커트 상품 분류에 속한 상품 중 다섯 개를 무작위 추출하여 해당 상품의 모든 상품평을 피실험자에게 보이고 8개 평가 항목에 대한 시스템의 요약 결과에 대해 0점부터 10점까지의 점수를 매기게 하였다. 만족도 평가에서 평균 7.6점을 받아, 제안된 시스템을 통해 상품평을 자동으로 요약한 결과가 실제 사용자의 상품평 분석

<표 3> 로맨셰이드의 '사이즈'에 대한 평가 어휘 사전 추출의 예

긍정	좋다, 맞다, 저렴하다, 구입하다, 예쁘다, 원하다, 달다, 넉넉하다, 남겨주다, 마음들다, 고민하다, 상관없다, 불다, 팬찮다, 설치하다, 생각하다, 비하다, 받다, 말다, 정하다, 주문했다, 추천하다, 친절하다, 나누다, 꼼꼼하다, 깔끔하다, 기다, 크다, 튼튼하다, 팔다, 해주다, !주문하다, !작다
부정	다르다, 주문하다, !맛다, 찾다, 치다, 비싸다, 아쉽다, 애매하다, 어렵다, 작다, 이상하다, 별나다, 끝내다, 걱정하다, 허전하다, 2개다, !적당하다, 화사하다

<표 4> 문장 극성 판별 정확도 평가 결과

상품	스커트 1			스커트 2		평균
	길이	디자인	재질	길이	배송	
정확율	79.6%	73.5%	90.0%	77.7%	88.4%	81.8%

에 도움이 되는 것으로 나타났다.

5. 결 론

본 논문에서는 근래 폭발적으로 증가하고 있는 상품평을 효과적으로 활용하기 위한 상품평 요약 시스템을 제안하였다. 웹 문서 검색을 통한 평가 항목 추출과 작은 크기의 시드 사전에 기반한 극성 사전 자동 확장을 통해 한국어 상품평 요약 연구가 가지는 지식 구축의 어려움을 해소하였다. 시스템의 성능 측정을 위한 실험 결과는 평가 어휘 극성 사전 분류의 정확도 69.8%와 채현율 85.7%, 문장 극성 분류 정확도 81.8%, 사용자 만족도 7.6점을 얻어 본 시스템이 상품평 요약에 유용함을 보였다. 상품평 요약은 모바일기기를 이용한 온라인 쇼핑에서의 상품평 제공에 활용될 수 있으며, 각 기업에 자사 또는 경쟁사의 제품에 대한 평가 파악을 용이하게 할 것으로 기대된다.

향후 연구로는 사전과 웹 문서와 같이 획득하기 쉬운 추가 지식을 활용한 극성 사전 추출의 정확도 향상, 상품평의 다양한 비문법적 표현의 추가적인 보정 등이 필요하다. 근래 국내 쇼핑몰의 경우 많은 상품평을 획득하기 위해 단기간에 상품평 작성을 요구하여 장기간 사용해야 품질을 검증할 수 있는 상품에 대한 유용한 상품평을 얻기 어려운 경우가 있다. 적절하지 않은 상품평을 제외시키고 쇼핑몰과 상품의 특성에 따라 각 상품평을 다른 중요도로 사용하기 위한 추가적인 모델 개발이 필요하다.

참 고 문 헌

[1] 한국인터넷진흥원, “웹 2.0시대의 네티즌 인터넷 이용 현황 - 참여와 공유의 인터넷,” http://www.nida.or.kr/doc/issue_sum.pdf, 2006.

[2] Turney, “Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews,” In Proc. of the ACL, pp.417-424, 2002.

[3] Pang, L. Lee, and S. Vaithyanathan, “Thumbs up? Sentiment classification using machine learning techniques,” In Proc. of EMNLP, pp.79-86, 2002.

[4] Kushal Dave, Steve Lawrence, David M. Pennock, “Mining the peanut gallery : opinion extraction and semantic classification of product reviews,” In Proc. of the WWW, pp.519-528, 2003.

[5] Yu and V. Hatzivassiloglou, “Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences,” In Proc. of EMNLP, pp. 129-136, 2003.

[6] Esuli, A. and Sebastiani, F., “Determining term subjectivity and term orientation for opinion mining,” In Proc. of the EACL, pp.193-200, 2006.

[7] Hu, M. and Liu, B. “Mining and summarizing customer reviews,” In Proc. of the 10th ACM SIGKDD Conf., pp.168-177, 2004.

[8] Bing Liu, Mingqing Hu and Junsheng Cheng. “Opinion Observer: Analyzing and comparing opinions on the web,” In Proc. of WWW, pp.342-351, 2005.

[9] Wilson, T., Wiebe, J., and Hoffmann, P, “Recognizing contextual polarity in phrase-level sentiment analysis,” In Proc. of the Conference on HLT and EMNLP, pp.347-354, 2005.

[10] Esuli, A. and Sebastiani, F. “Determining term subjectivity and term orientation for opinion mining,” In Proc. of the EACL. pp.193-200, 2006.

[11] Hatzivassiloglou, V. and McKeown, K. R, “Predicting the semantic orientation of adjectives,” In Proc. of the EACL, pp.174-181, 1997.

[12] 명재석, 이동주, 이상구, “반자동으로 구축된 의미 사전을 이용한 한국어 상품평 분석 시스템,” 정보과학회논문지 : 소프트웨어 및 응용, 제 35권 제 6호(2008. 6), pp.392-403, 2008.

[13] 황재원, “한국어 문서 감정 분류를 위한 감정 자질 추출 및 가치 강화 기법에 관한 연구,” 동아대학교, 석사학위논문, 2008.

[14] 이우철, 이현아, 이공주, “효율적인 상품평 분석을 위한 어휘 통계 정보 기반 평가 항목 추출 시스템,” 정보처리학회논문지 제16-B권 제 6호, 2009.

[15] Naver Lab, 자동 띄어쓰기, <http://s.lab.naver.com/autospacing/>

[16] 강승식, HAM, “한국어 형태소 분석기와 한국어 분석 모듈,” 국민대학교 자연언어 정보검색연구실, <http://nlp.kookmin.ac.kr>.

[17] 온라인 가격비교 사이트 BB.co.kr, <http://www.bb.co.kr>.

이 우 철



e-mail : lee256@naver.com

2009년 금오공과대학교 소프트웨어공학과 (공학석사)

2009년~현 재 (주)유승토타솔루션 대리
관심분야: 자연언어처리, 지식공학, 웹서비스

이 현 아



e-mail : halee@kumoh.ac.kr

1996년 연세대학교 컴퓨터과학과(학사)
1998년 한국과학기술원 전산학과(공학석사)
2004년 한국과학기술원 전산학과(공학박사)
2000년~2004년 (주)다음소프트 언어연구 팀장

2004년~현 재 금오공과대학교 컴퓨터공학부 조교수
관심분야: 자연언어처리, 정보검색, 지식공학, 기계번역

이 공 주



e-mail : kjoolee@cnu.ac.kr

1992년 서강대학교 전자계산학과(학사)
1994년 한국과학기술원 전산학과(공학석사)
1998년 한국과학기술원 전산학과(공학박사)
1998년~2003년 한국마이크로소프트(유) 연구원
2003년 이화여자대학교 컴퓨터학과 대우전임강사

2004년 경인여자대학 전산정보과 전임강사
2005년~현 재 충남대학교 전기정보통신공학부 부교수
관심분야: 자연언어처리, 자연어인터페이스, 기계번역, 정보검색