

감성 강도를 고려한 감성 분석 평가집합 구축

Constructing an Evaluation Set for Korean Sentiment Analysis Systems Incorporating the Category and the Strength of Sentiment

김도연, 오영, 박혁로

전남대학교 전자컴퓨터공학과

Do-Yeon Kim(doyeon@ejnu.net), Wu Yong(wuyong@ejnu.net),

Hyuk-Ro Park(hyukro@jnu.ac.kr)

요약

감성 분석은 블로그와 트위터 같은 다양한 소셜 미디어에서 사용자들이 표현하는 감정의 종류를 분석하고 추출하는 연구이다. 현재 감성 분석 연구는 꾸준히 계속되고 있지만, 한국어의 감성 분석 평가 집합은 아직 없다.

본 논문에서는 감성 분석을 평가할 수 있는 평가집합을 구축한다. 평가집합에서는 사용자의 감성에 대한 극성뿐만 아니라 감성의 종류와 강도까지 고려한 평가집합을 구축하였다. 이를 위해 감성의 종류는 긍정에서 7가지의 범주와 부정에서 15가지의 범주를 나누고, 각 범주별로 1~3까지의 강도를 설정하였다. 또한 각 범주에 속하는 어휘에 대해서도 1~3까지의 강도를 설정하였다. 평가집합의 데이터는 다양한 소셜 미디어에서 3,270 문장을 추출하여 구축하였으며, 각 문장에 대해 5 명이 감성의 종류와 강도를 태깅하였다. 구축한 평가집합에서 5명의 일치도는 극성의 경우 93 %, 감성의 종류는 70 %, 강도는 58 % 로 나타났다. 이는 독일어와 스페인어의 평가 집합 보다 일치도가 높게 나타났다. 이 결과는 제안한 평가 집합이 신뢰할 만한 자료로 다른 감성 분석 시스템의 평가데이터로 사용될 수 있음을 보여준다.

■ 중심어 : | 감성 분석 | 감성 강도 | 평가 집합 |

Abstract

Sentiment analysis is concerned with extracting and analyzing different kinds of user sentiment expressed in a variety of social media such as blog and twitter. Although sentiment analysis techniques are actively studied for these days, evaluation sets are not developed yet for Korean sentiment analysis.

In this paper, we constructed an evaluation set for Korean sentiment analysis. To evaluate sentiment analysis systems more thoroughly, each sentence in our evaluation set is tagged with the polarity of the sentiment as well as the category and the strength of the sentiment. We divide kinds of sentiment into 7 positive categories and 15 negative categories. Each category is given the strength of the sentiment from 1 to 3. Our evaluation set consists of 3,270 sentences extracted from various social media. For each sentence, 5 human taggers assigned the category and the strength of the sentiment expressed in the sentence. The ratio of inter-taggers agreement was 93% in the polarity, 70% in the category, 58% in the strength of sentiment. The ratio of inter-taggers agreement our evaluation set is a bit higher than other evaluation sets developed for German and Spanish. This result shows our evaluation set can be used as a reliable resource for the evaluation of sentiment analysis systems.

■ keyword : | Sentiment Analysis | Sentiment Strength | Evaluation Set |

* "본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음"

(NIPA-2012-H0301-12-3005)

접수번호 : #120823-007

접수일자 : 2012년 08월 23일

심사완료일 : 2012년 10월 10일

교신저자 : 박혁로, e-mail : hyukro@jnu.ac.kr

I. 서론

감성이란 '어떤 현상이나 일에 대하여 일어나는 마음이나 느끼는 기분'이다. 감성 어휘는 객관적 가치 평가보다는 주로 내면이나 주관적 감정 또는 심리가 작용하는 의미 영역을 묘사한 것이다. 이러한 감성을 분석한다는 것은 자연언어처리와 전산언어학 그리고 텍스트 분석론을 활용하여 원 자료에서 주관적인 정보를 발견하고 추출하는 과정이다[1].

감성의 분류 체계를 이용하여 감성의 종류를 분석하고 분류하는 연구는 크게 세 가지의 단계로 나눌 수 있다. 첫 번째 단계는 감성 정보가 들어 있는 주관적인 생각이나 느낌을 표현하는 문장을 추출하고, 다음 단계에서 문서 또는 문장의 극성(긍정, 부정)을 나눈다. 마지막 단계는 문서 또는 문장이 어느 정도의 주관성을 갖는지 그 강도를 구하는 강도 분류 연구이다[2][3]. 첫 번째 단계인 감성 정보 분류는 감성 분석 분류과정 중에서 아주 기초적인 과정으로서 문장 중에서 가치 있는 감성 정보들을 추출하는 과정이다. 이런 감성 정보들은 감성 어휘(예 : 기쁘다, 좋다 등), 감성 평가 대상(예 : 핸드폰, 자동차 등), 그리고 관점 소유자(예 : 나, 사용자 등)등이 이에 속한다. 특히 이중에서 감성 어휘의 기준을 설정하는 것은 매우 중요한 일이다. 이 분야는 대부분 심리학에서 분류한 감성 체계를 이용하여 감성 어휘들을 설정한다. 심리학에서 분류하는 인간의 감성에 대한 견해는 학자들에 따라 차이는 있지만 대략 4~8가지로 구별하고 있다. 동양권에서는 희로애락(喜怒哀樂) 또는 칠정(七情)으로 분류하고 있으며, 서양권에서는 즐거움(Joy), 행복(Happiness), 화남(Anger), 슬픔(Sadness) 등으로 분류하고 있다[4][5]. 대부분의 논문에서는 plutchik[6]를 기준으로 감성 분류를 하고 있다. 그러나 서양의 감성으로써 동양권의 감성인 한국어에 적용하기에는 적합하지 않다. 그러므로 본 논문에서는 동양권의 정서에 맞는 공자의 칠정을 기준으로 분류체계를 설정한다.

기존의 대부분의 연구에서는 극성을 분류하여 얼마나 정확한 성능을 보였는지에 주력하고 있다. 성능 평가에 있어서 영어의 경우 신뢰할 만한 평가집합이 공개

되어 있다. Cornell university의 영화리뷰 평가집합은 긍정과 부정 각각 1,000개 문서, 주관성과 객관성으로 태깅된 문장 각 5,000개로 되어 있다[7]. 그리고 MIT대학교에서는 레스토랑에 대한 평가로 총 4,488문장에 대해 음식, 환경, 서비스, 가격, 체험의 5가지 요소로 구분하고, 각 요소별 1~5까지 점수를 매겨 평가한 집합이 있다[8]. 또 다른 GI(General Inquirer) 평가용어사전은 긍정 1,915개, 부정 2,291개의 각 용어에 대해 극성과 강도 등의 속성을 태깅한 평가 집합 등[9] 이외에 다양한 평가 집합이 존재한다. 그러나 아직까지 한국어에 대한 평가 집합은 존재하지 않는다. 이에 본 논문에서는 동양의 정서에 맞는 공자의 분류체계를 바탕으로 한국어에 적합한 감성 분류 체계를 만들고, 이 분류 체계를 바탕으로 극성 분류 뿐만 아니라 세부 감정 분류, 그리고 각 어휘에 대한 강도와 정도 부사어의 강도까지 나타내는 평가 집합을 구축하고자 한다.

본 논문의 구성은 다음과 같다. 2절에서는 심리학에서의 감성 분류 체계와 기존 감성 분석 연구들의 분류 방법, 그리고 영어에서의 평가집합에 대하여 살펴본다. 3절에서는 제안하는 한국어에 맞는 감성 분류 체계와 평가 집합에 대하여 자세히 설명하고, 4절에서는 성능 평가를 보여준다. 마지막 5절에서는 결론과 향후 연구를 제시한다.

II. 관련 연구

심리학에서 정의하는 감성 분류 체계와 이를 이용한 감성 분석의 기존 방법들을 살펴보고, 영어권에서의 평가집합에 대하여 알아본다.

1. 감성 분류체계 및 감성 강도

심리학에서는 인간의 감정을 보통 4~8가지로 구별하며, 동양과 서양으로 크게 나뉘 볼 수 있다. [표 1]은 동양과 서양의 인간의 기본감정을 나타낸 것이다. 비교표를 보면 학자들이 정의한 인간의 기본 감정은 동서양이 거의 유사함을 알 수 있다. 그러나 동양의 칠정과 서양의 기본 감정은 약간 다른 개념이다. 동양은 인간이 느

낄 수 있는 감정들의 대표격을 설정하여 분류한 것이고, 서양의 기본 감정은 선천적인 감정으로서, 이 기본 감정들을 혼합하여 인간의 다른 모든 감정들이 발생한다는 관점이다.

표 1. 동양과 서양의 인간 기본 감정[4][5]

공자	중용	노자	Tompkins (1962) & Izard(1977)	Plutchik (1980)	C (2001)
喜(희) 愛(애)	喜(희) 樂(락)	喜(희) 好(호)	Joy	Joy	Happiness Love Excitement Contentment
怒(노)	怒(노)	怒(노)	Rage Anguish	Anger	Anger
懼(구) 惡(오)		憂(우) 懼(중)	Fear Disgust Shame	Fear Disgust	Fear
哀(애)	哀(애)	悲(비)		Sadness	Sadness Depression
慙(옥)		慙(옥)	Surprise	Surprise Anticipation trust	

기존의 다른 연구를 살펴보면 국어국문학에서 한국어에 대한 감정동사에 대하여 많은 연구가 진행되었다. 김은영[4]의 연구를 살펴보면, 한국의 전통적인 정서인 사단칠정(四端七情)을 기준으로 분류하였다. 사단 중에서는 감정에 의미 영역에 속하는 측은지심(惻隱之心)과 수오지심(羞惡之心)의 2가지와 공자의 칠정을 기준으로 동정, 수치심, 기쁨, 노여움, 슬픔, 두려움, 좋아함, 싫어함, 바람으로 분류하였다. 각각의 영역에 속하는 감정동사의 목록은 그 의미를 영역별로 묶어 분류하고 있다. 또한, 어휘 사이에는 유의관계가 있고 어떤 의미를 공유하는 것은 그것들은 같은 의미 부류로 묶을 수 있다는 것이다. 이를 토대로 김은영은 크게 만족과 불만족으로 나누고 하위에 9가지의 분류체계와 어휘목록들을 설정하였다.

감성 분류에서 비슷한 의미를 띠는 어휘 사이의 강도를 등급으로 표시 할 수 있다. 김기홍[10]의 연구를 보면 Plutchick[6]나 Osgood[11]가 조사한 감성의 강도를 참작하여 앞의 [표 2]과 같이 나타내었다. 표에서 보면, 10.3 ~ 9까지는 굉장히 강한 감정이고 숫자가 낮아질수록 감정의 강도가 약함을 의미한다.

표 2. 감정어휘의 3~4가지 정도의 강도 표시[10]

10.3 - 9.0	8.5 - 7.0	6.5 - 5.0	4.0 -
Rage	Anger	Annoyance	
Ecstasy	Joy	Happiness	
Astonishment	Amazement	Surprise	
Terror	Panic	Fear	Apprehension
Grief	Sorrow	Dejection	
loath	Disgust	dislike	boredom
분통이 터지다	노엽다	귀찮다	
황홀	통쾌	즐거움	
	놀람		
끔찍스럽다	무섭다	두렵다	불안
비애	슬픔	풀죽다(저하)	실중
구역질나다	기분 나쁘다	비위에 거슬리다	

정도 부사어란 상태 동사의 정도를 한층 높이거나 낮추는 기능을 보이는 부사어로서, ‘매우, 훨씬, 대단히, 꽤’ 등이 있다. 정도성의 크기는 ‘가장 높은 등급, 높은 등급, 낮은 등급’으로 정할 수 있다. 높은 등급에 해당하는 부사어는 ‘몹시, 무척, 매우, 아주, 가장’등이 있고, 높은 등급의 부사어는 ‘꽤, 썩, 훨씬, 한결’등이 있다. 낮은 등급의 부사어는 ‘꽤, 제법, 좀’등이 해당한다[12]. 예를 들면, ‘조금 슬프다’나 ‘약간 슬프다’는 ‘울적하다’, ‘우울하다’와 가깝고, ‘아주 슬프다’, ‘매우 슬프다’는 ‘비통하다’, ‘침통하다’와 가깝다고 볼 수 있다. 이러한 정도 부사어의 등급을 이용하여 어휘의 강도를 정하는데 활용하였다.

2. 기존의 평가 집합 연구

2010년 Julia와 2명은 기존의 영어 상품평 말뭉치를 기반으로 다국어 즉, 영어, 독일어, 스페인어의 상품평 말뭉치를 구축하였다. 각 제품의 상품평가에 대한 자질(평가 대상), 자질의 극성, 그리고 강도를 0~3으로 나누었다. 전문가 2명이 태깅하였고, 독일어의 경우, 2명의 극성 일치도가 75.7%(agreement)이다. 극성의 태깅은 비교적 높은 일치율을 보였다. 그러나 강도는 2명의 일치도가 44.3%(agreement)로 비교적 낮게 나타났다[13].

이외의 서론에서 언급한 영화리뷰 평가집합, 레스토랑에 대한 평가집합등 외국의 경우 다양한 평가집합이 존재한다. 이는 주로 상품평들을 대상으로 극성 분류 혹은 극성에 대한 강도를 3~5등급으로 나누는 것이 대부분이다. 그러나 상품평 뿐만 아니라 모든 감정을 표

현한 문장을 분류하기에는 ‘좋다, 나쁘다’의 2가지 감정으로 나누는 것은 너무 포괄적이며 다른 다양한 감정을 표현해 내지 못한다. 이에 본 논문에서는 감정을 표현한 모든 문장을 분류 할 수 있도록, 극성 뿐만 아니라 다양한 세부감정에 대한 분류를 하였으며, 또한, 세부감정이 어느정도의 강도를 가지고 있는지도 등급을 세밀화(9등급)하여 평가집합을 구축한다.

또 다른 연구에서는 모바일 환경에서 설문을 디자인하고 인간의 감성에 가장 가깝게 접근할 수 있는 인공지능의 퍼지 이론을 도입하여 감성을 평가한 방법도 있다[14].

III. 세부 감성 분류의 평가 집합 구축

본 절에서는 극성 분류뿐만 아니라 세부 감성 분류를 하기 위한 분류 기준과 범주별 강도(1~3)에 대하여 설명한다. 또한, 각 범주별 어휘 목록들을 기술하고 어휘들 사이의 강도(1~3)와 정도 부사어의 강도(1~3)에 대하여 자세히 살펴보고자 한다.

1. 감성 분류의 기준

본 연구에서는 문장의 극성만 나타내는 것이 아니라 어떠한 감정을 표현하는지에 대해서도 세밀하게 태깅하는 평가집합이다. 따라서 분류의 범주가 너무 많으면 태깅하기 어렵고, 너무 적으면 정확히 어떠한 감정을 나타내려고 하는지 분류하기 어렵다. 이에 기존의 김은영의 국어 감정동사 연구[4]의 대분류의 범주(만족:2, 불만족:7)의 하위에 있는 소분류의 범주(만족:12, 불만족:34)들을 본 연구에 맞게 재정의 하여, 긍정의 범주 7개, 부정의 범주 15개로 설정하였다.

동양에서의 감정 분류는 감정이 표현된 어휘들을 의미적으로 비슷한 것끼리 모아 하나의 대표격을 설정하여 분류한다[4]. 이는 어휘들끼리 서로 유의어의 관계를 갖는다면 하나의 범주로 묶을 수 있다는 말이다. 그래서 (주)날말 유의어 사전[15]을 기반으로 소분류의 범주에 속하는 어휘들 사이에서 서로 유의어 관계를 가진다면 이들 범주를 통합하였다.

긍정에서의 예를 보면, ‘만족감’의 어휘 목록들은 ‘만족하다’, ‘흡족하다’, ‘뿌듯하다’...등이 있고, ‘자랑스러움’의 어휘 목록들은 ‘대견하다’, ‘자랑하다’... 등이 있다. 여기에서 ‘대견하다’를 사전에서 찾아보면 유의어는 ‘만족하다’, ‘흡족하다’이다. 그리고 ‘반가움’의 어휘목록들은 ‘달갑다’, ‘반갑다’...등이 있고, ‘반갑다’의 유의어는 ‘만족하다’, ‘흡족하다’이다. 그래서 ‘만족감’, ‘자랑스러움’, ‘반가움’을 통합하여 ‘만족감’으로 한다. 또 다른 범주 ‘기쁨’을 보면, 해당 어휘로는 ‘유쾌하다’등이 있고, ‘통쾌함’ 범주의 어휘에 ‘상쾌하다’가 있다. 여기에서 ‘상쾌하다’의 유의어는 ‘유쾌하다’, ‘명쾌하다’, ‘쾌적하다’등이 있다. 이렇게 2개의 범주를 통합하여 ‘기쁨’으로 한다.

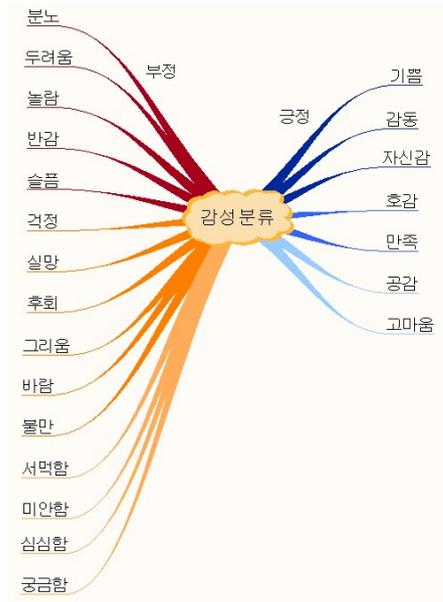


그림 1. 제안하는 감성 분류 체계

부정에서의 예를 보면, ‘동정’의 범주의 어휘들을 살펴보면, ‘애뜻하다’, ‘처량하다’, ‘처절하다’등이 있으며, ‘슬픔’ 범주의 어휘에도 ‘애뜻하다’, ‘처량하다’, ‘처절하다’등이 속해 있다. 이는 동정의 의미이면서 슬픔의 의미도 된다는 뜻이다. 서로 의미적으로 통하므로 하나의 범주로 ‘슬픔’으로 통합하였다. 다른 범주 ‘걱정’의 어휘에는 ‘고민하다’, ‘노심초사하다’, ‘속타다’ 등이 있으며,

‘초조합’의 범주의 어휘는 ‘노심초사하다’, ‘속타다’, ‘싱숭생숭하다’ 등이 있다. 이경우도 앞의 경우와 마찬가지로 하나로 통합할 수 있다. 또 다른 범주 ‘반감’을 살펴보면 어휘에 ‘가소롭다’, ‘지긋지긋하다’, ‘질리다’ 등이 있고, ‘싫증’의 범주의 어휘는 ‘신물나다’, ‘지긋지긋하다’, ‘질리다’, ‘식상하다’ 등이 있다. 비슷한 어휘이므로 통합하여 ‘반감’의 범주로 정하였다. 이와 같이 유의어 사전을 바탕으로 위의 [그림 1]과 같은 범주를 설정하였다.

2. 감성 분류의 강도 정의

기존의 연구에서 보듯이 어느 정도 어휘의 강도를 표시할 수 있다. 이를 기반으로 각각의 범주에 3등급의 기본 강도를 정하고, 그 안의 어휘목록들 사이에서도 3등급의 강도를 정하였다. 긍정과 부정의 분류 목록의 강도는 옆의 [표 3]과 같다. 표에서 괄호안의 숫자는 강도를 나타내며, (3)은 강도가 높고, (2)는 중간 정도의 강도, (1)은 낮은 강도를 나타낸다.

표 3. 감성 범주의 강도와 어휘 목록 사이의 강도

의미	분류	감정어휘
긍정	(3) 기쁨	(3) 들뜨다, 신바람나다, 황홀하다, 통쾌하다 ... (2) 기쁘다, 신나다, 즐겁다, 행복하다, 흥겹다, 설레이다, 실땀나다 ... (1) 반갑다, 우습다 ...
	(3) 감동	(3) 열광하다, 전율이다 ... (2) 감동하다, 감탄하다, 몽클하다 ... (1) 감사하다, 고맙다, 은혜롭다 ...
	(3) 자신감	(3) 기고만장하다, 자신만만하다 ... (2) 떳떳하다, 우쭐하다 ... (1) 자부하다, 자신하다 ...
	(2) 만족	(3) 만족하다, 뿌듯하다, 대견하다 ... (2) 안심하다, 좋다 ... (1) 담담하다, 맑는다...
	(2) 호감	(3) 매료되다, 사랑하다, 흠모하다 ... (2) 감미롭다, 좋아하다, 호감가다 ... (1) 귀엽다, 예쁘다 ...
	(1) 공감	(3) 교감하다 ... (2) 든든하다, 든직하다, 믿다 ... (1) 공감하다, 동감하다 ...
	(1) 고마움	(3) 축하의 어휘들 (2) 격려, 응원의 어휘들 (1) 인사말의 어휘들
부정	(3) 분노	(3) 개탄하다, 격노하다, 노발대발하다, 분통터지다 ... (2) 기막하다, 분하다, 짜증나다, 화나다, 흥분하다 ... (1) 속상하다, 신경질나다, 악오르다 ...
	(3) 두려움	(3) 공포스럽다, 무시무시하다, 섬뜩하다, 소름끼치다 ... (2) 겁나다, 두렵다, 무섭다 ... (1) 스산하다 ...
	(3) 놀람	(3) 경악하다, 기겁하다, 기절초풍하다 ... (2) 놀라다, 당혹하다, 당황하다, 아찔하다, 황당하다 ... (1) 뜨끔하다, 병병하다, 얼떨떨하다 ...
	(3) 반감	(3) 가증스럽다, 경멸하다, 구역질나다, 끔찍하다, 혐오하다 ... (2) 가소롭다, 갈잡다, 거북하다, 꺼림칙하다, 결코럽다, 무시하다, 신물나다, 싫다, 싫증나다 ... (1) 시시하다, 식상하다, 지루하다 ...
	(3) 슬픔	(3) 구슬프다, 비장하다, 비참하다, 비통하다, 뼈아프다, 애절하다, 참담하다, 처절하다 ... (2) 고통스럽다, 괴롭다, 냉정하다, 눈물겹다, 매정하다, 무정하다, 아박하다, 야속하다 ... (1) 아프다 ...

부정	(2) 걱정	(3) 노심초사하다, 안달복달하다, 안절부절하다, 전전긍긍하다 ... (2) 갑갑하다, 걱정하다, 고뇌하다, 긴장하다 ... (1) 망설이다, 아슬아슬하다 ...
	(2) 실망	(3) 낙담하다, 원망하다, 좌절하다 ... (2) 공허하다, 덧없다, 서운하다, 실망하다, 씁쓸하다, 안타깝다, 애석하다 ... (1) 아깝다, 아니꼽다, 아쉽다.
	(2) 후회	(3) 책망하다, 탄식하다 ... (2) 가책하다, 뉘우치다, 자책하다, 회개하다 ... (1) 후회하다...
	(2) 그리움	(3) 고독하다, 고립되다, 삭막하다, 적막하다 ... (2) 그림다, 소외되다, 쓸쓸하다, 외롭다, 적적하다 ... (1) 소소하다 ...
	(2) 바람	(3) 간절하다, 갈망하다, 절절하다 ... (2) 소망하다, 소원하다, 시기하다, 심통나다, 질투하다, 탐나다 ... (1) 바라다, 부럽다, 샘나다, 원하다...
	(2) 불만	(3) 낭패하다, 못마땅하다, 불신하다 ... (2) 곤란하다, 불만족하다, 불편하다, 불평하다, 뽀로통하다, 섣쪽하다, 성가시다, 의아하다 ... (1) 귀찮다, 시들하다, 시큰둥하다, 심드렁하다, 어수선하다...
	(1) 서먹할	(1) (2) 겸연쩍다, 머쓱하다, 멧쩍다, 서먹서먹하다, 소원하다 ... (1) 낯설다, 서먹하다, 어색하다...
	(1) 미안할	(2) 송구하다, 죄송하다, 죄스럽다 ... (1) 미안하다...
	(1) 심심할	(1) (2) 권태롭다, 따분하다, 적적하다 ... (1) 무료하다, 심심하다...
	(1) 궁금할	(1) 궁금하다

3. 정도 부사어와 비속어 처리의 문제점 처리

정도 부사어는 감성 어휘의 강도를 높이거나 낮추는 기능을 하는 부사어를 의미한다. 여기에 해당되는 어휘는 ‘진짜, 훨씬, 굉장히, 정말, 너무, 아주 ...’ 등이 있다. 그러나 아직까지 정도 부사어에 대한 강도의 등급을 정한 연구는 없었다. 본 연구에서는 정도 부사어에 별도로 강도를 주어 감성어휘의 강도와 정도 부사어의 강도를 더하여 사용할 수 있도록 하였다. 이는 시스템의 성격에 따라 가중치를 더 줄 수도 있고, 뺄 수도 있으며 점수로 계산시 강도3= 0.3, 강도2= 0.2, 강도1= 0.1 등으로 조정하여 사용 가능하다. 3명의 연구원이 실제 문장에서 나오는 정도 부사어들을 수집하고 의견을 조합하여 다음과 [표 4]와 같은 강도를 정하였다.

표 4. 정도 부사어의 강도

강도	부사어 어휘 목록
3	진짜, 겁나, 완전, 매우, 제발, 간곡히, 결사, 절대, 굉장히, 엄청, 무지, 짱, 훨씬, 상당하...
2	꼭, 너무, 참, 정말, 더욱, 꽤, 좀더...
1	조금, 약간, 거의, 다소, 겨우, 그다지...

정도 부사어는 아니지만 뒤 혹은 앞의 문장을 꾸며주면서 뜻을 더 강하게 하는 경우도 강도를 주었다. 예를

들어 험, 확, 비속어를 앞에 붙이는 경우(예: 개싫어) 등이다.

비속어의 출현은 부정의 감정을 나타낼 때 특히 그 중에서도 분노, 불만, 실망 등에서 더 자주 나타난다. 이러한 경우는 거의 50~60% 이상의 문장에서 비속어가 나타나며, 비속어가 나타날 때 문장의 강도는 더 강하게 느껴진다. 그러므로 비속어는 기존의 강도에서 1단계 높게 주었다.

4. 평가집합의 난이도 설정

문장에서 감성을 추출할 때 한 문장에서 출현하는 모든 감성어휘들을 분류하고 강도를 나타내어 긍정의 어휘들이 많이 나타날수록 강도는 더 높아 지도록 하였다.

또한 감성 어휘의 의미별 분류와 강도를 정하여 평가집합을 구축하려면 우선 무엇보다 객관성을 유지하는 것이 중요하다. 문장에서 감성 어휘가 직접적으로 드러나 있는 경우도 있고, 직접적인 표현은 없지만 문장 전체에서 분별할 수 있는 감성이 있다. 만약 자동으로 감성 어휘를 추출해내는 경우 본 연구에서 구축한 평가집합을 이용하여 성능을 평가하였을 때, 분류하기 쉬운 문장만 있다면 성능이 아주 좋게 나오게 된다. 반대로 어려운 문장만 있다면 성능은 좋지 않을 것이다. 그래서 난이도 별로 5개의 그룹으로 나누어 평가집합을 구축하였다. 난이도 설정은 다음과 같다.

(1) 문장에 감성 어휘가 직접적으로 나타나고 태깅하기 쉬운 문장

: 예를 들어 “갤럭시 노트 좋다”와 같이 비교적 태깅하기 쉬운 문장들로만 구성하였다.

(2) 문장에 감성이 비유적으로 나타나지만 태깅하기 쉬운 문장

: 예를 들어 “아 진짜 게임 셋다운제 만든 놈들은 제 정신이 아니어”라는 문장은 감성 어휘가 직접적으로 나타나지 않았지만 극성은 분명하게 부정적이고, 분노 혹은 불만 등으로 나타낼 수 있다. 표현 방법이 비유적이지만 비교적 태깅하기 쉬운 문장이다.

(3) 문장에 감성 어휘가 직접적으로 나타나지만 문장이 길고 여러감성이 나타나는 문장

: 예를 들어 “이제 승리가 확실시되고 있으니 오늘밤까지 승리의 기쁨 맘껏 누리고 내일부터 다시 전쟁이다!!!! 한미 FTA 결사 반대!!!” 이 문장에서는 감성 어휘들이 많이 출현되고, 서로 다른 극성도 함께 나타나는 경우이다. ‘기쁨’, ‘결사반대’등의 직접적인 표현이 들어가 있지만 여러 감성이 나오는 경우와 태깅하기 어렵다고 느껴지는 문장들만 모아놓은 경우이다.

(4) 문장에 감성이 비유적으로 나타나면서 문장이 길고 태깅하기 어려운 문장

: 예를 들어 “셋다운제 관련 기사는 어느새 네이버 메인에서 실종되고 검색어 순위는 강등되고 ㅋㅋㅋ 나라 꼴 가관이구나 ~정부기관이 압력 넣으면 알 권리 따위 개나 주는구나 ~그렇구나 ~#이 게_다_여성가족부_때문이다” 이 문장에서 보면 감성을 나타내는 어휘는 출현하지 않았지만 비유적으로 ‘나라 꼴 가관이구나’, ‘알 권리 개나 주는구나’를 보면 극성은 부정이며, 범주는 ‘분노/불만’등에 해당되고, 어휘목록 중에서는 ‘개탄하다/성질나다/못마땅하다’등에 해당될 수 있다.

(5) 문장에 나타난 감성 어휘와 반대의 의미를 지니는 문장

: 예를 들어 ‘난독증 축하드립니다’, ‘여러분들 온라인 게임셋다운제가 시작됩니다! 우와 기쁘다 x같은 여성부 새끼들아!’ 등의 문장을 보면 분명 나타나는 감성 어휘는 긍정이지만 비교는 말로 부정의 의미를 갖는다. 이렇게 난이도 별로 나누어진 평가 집합을 사용할 경우 각 그룹에서 문장들을 정한 비율로 적용하여 객관적으로 얼마나 어려운 문장들을 어느 정도의 정확성으로 분류할 수 있는지 평가 할 수 있도록 하였다.

IV. 평가집합 실험

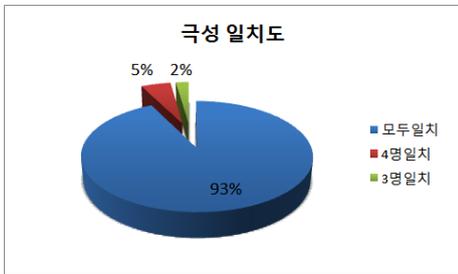
실험 데이터로 트위터, 싸이월드 C로그, 네이버의 영화 리뷰 댓글의 데이터 약 23,000문장을 수집하여 그중

에서 감성이 있는 문장만 선별하여 3,270문장의 평가집합을 구축하였다. 주제별로는 갤럭시탭, 갤럭시 S2, 무상급식, 아이패드, 전자팔찌, 구글폰, 넥서스폰, 아이폰, 아이패드, 모바일 OS, 배터리, 섀터다운제, 한미FTA, 그리고 영화는 박쥐와 해운대의 댓글, 주제가 없는 문장 등으로 구별하였다. 평가집합은 아래 [그림 2]와 같이 구축하였다.

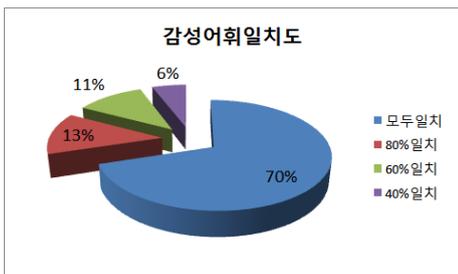
Index	Topic	Category	Text	유무	강도	차이	일치율	난이도	빈도	종류
1	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	높음	1	100%	1	1	제품
2	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품
3	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	중간	2	100%	1	1	제품
4	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품
5	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	중간	2	100%	1	1	제품
6	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품
7	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	중간	2	100%	1	1	제품
8	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품
9	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	중간	2	100%	1	1	제품
10	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품
11	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	중간	2	100%	1	1	제품
12	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품
13	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	중간	2	100%	1	1	제품
14	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품
15	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	중간	2	100%	1	1	제품
16	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품
17	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	중간	2	100%	1	1	제품
18	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품
19	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	유	중간	2	100%	1	1	제품
20	갤럭시	제품	갤럭시 S2가 정말로 훌륭한 제품입니다	무	중간	2	0%	1	1	제품

그림 2. 감성 분류와 강도를 나타낸 평가집합

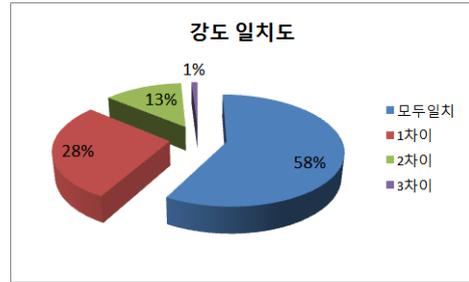
평가집합은 한 문장에 대하여 5명이 모두 태깅하는 방법으로 진행하였다. 난이도 그룹 (1)~(4)는 같이 비율로 적용하고, (5) 감성 어휘와 반대 의미를 지니는 그룹은 데이터가 적어 약1%정도 포함하였다. 이에 대한 결과는 아래 [그림 3]과 같다.



(a) 극성 일치도



(b) 감성 어휘 일치도



(c) 강도 일치도

그림 3. 평가집합의 구축 결과

극성의 경우 5명 모두 일치하는 경우는 93%이고, 4명이 일치하는 경우가 5%로 나타났으며, 3명이 일치하는 경우는 2%이다. 감성 어휘의 일치도를 보면 5명이 모두 같은 감정을 나타내는 경우는 70%이며, 80% 일치(4명이 같은 감정)의 경우는 13%이고, 60% 일치(3명이 같은 감정이면서 2명이 같은 감정이거나 2명이 모두 다른 감정)인 경우는 11%로 나타났으며, 40% 일치(각각 2명씩 일치하고 1명이 다른 감성 일 때 또는 2명이 일치하고 나머지 3명은 모두 불일치)는 6%로 나타났으며, 모두 다르게 태깅한 경우는 없었다. 감성 강도는 모두 같은 강도를 나타내는 경우는 58%이고, 강도의 차이가 1이 차이 나는 경우는 28%이며, 강도가 2차이가 나는 경우는 13%이다. 강도의 차이가 3이 나는 경우는 불과 1%밖에 되지 않으며, 그 이상 차이 나는 경우는 없었다. 2명이 같은 강도를 나타내는 경우는 15.2%이며 강도의 차이는 1의 차이를 보였다. 2이상 차이가 나는 경우는 없었다. 기존의 독일어와 스페인어의 연구[13]에서 2명이 태깅한 결과는 극성이 일치한 경우 75.7%이었고, 본 프로젝트에서는 93%의 일치도를 보였다. 또한 강도 측면에서도 기존 연구의 강도 일치도를 보면 41.3%에 불과하지만 평가셋의 결과는 84.5%의 높은 강도 일치도를 보였다.

앞서 데이터를 난이도에 따라 5그룹으로 나누었다. 이는 평가시 객관성을 증명할 수 있는 중요한 부분이다. 시스템 개발시 시스템의 우수성을 알리기 위해 태깅하기 쉬운 데이터가 많이 섞여있는 집합을 선택한다면 성능이 월등히 좋게 나타난다. 이러한 불공정한 평가를 막기 위해 난이도 설정은 꼭 필요하다. 아래 [그림

4)는 데이터 난이도에 따른 실험 결과이다.

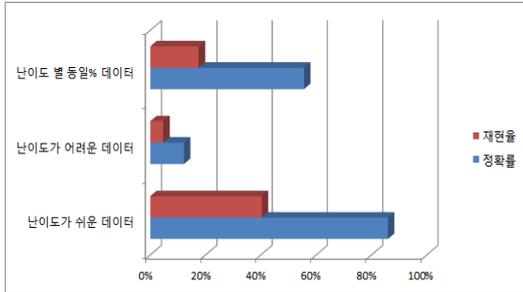


그림 4. 데이터 난이도에 따른 정확률과 재현율

난이도가 쉬운 데이터로만 사용하였을 경우, 정확률은 86%, 재현율은 40.3%이며, 난이도를 똑같은 비율로 섞어 평가하였을 때는 정확률 55.6%, 재현율 17.5%로 나타났다. 정확률은 30%이상 차이가 나며, 재현율의 경우 2배 이상 차이가 나는 것을 볼 수 있다. 이는 평가에서 데이터 난이도가 얼마나 중요한지를 나타내는 결과이다.

V. 결론

본 연구에서는 문장의 극성 뿐만 아니라 세부 감성을 분류하고, 감성의 강도를 정하여 평가집합을 구축하였다. 감성의 분류에서 동양의 정서가 잘 표현될 수 있는 공자의 7가지 감성을 기반으로한 김은영의 논문을 바탕으로 긍정의 세부 분류 7가지 즉, 감동, 기쁨, 자신감, 만족, 호감, 공감, 선의로 분류하였으며, 부정의 세부 분류 15가지 즉, 놀람, 두려움, 반감, 분노, 슬픔, 걱정, 그리움, 바람, 불만, 실망, 후회, 미안함, 서먹함, 심심함, 궁금함으로 분류하였다. 또한 분류에 따라 기본 강도를 정하고, 분류 안의 어휘목록들 간의 강도도 정하여 좀 더 세밀한 강도를 나타낼 수 있도록 하였다. 그리고 정도 부사어에 대해서도 강도를 정하여 구축된 평가집합을 이용할 경우 사용자에게 따라 적응적으로 사용할 수 있도록 하였다. 또한 난이도별로 평가집합을 구축하여 쉬운 문장과 어려운 문장을 적절히 섞어서 사용할 수 있도록 하였다. 다른 장점은 실세계에서 사용하는 데이

터를 이용하여 바로 적용 가능하다는 점이다. 아직까지 한국어에 대한 평가집합이 없는 상황에서 많은 연구에 초석이 될 것으로 기대된다. 향후 연구로는 젊은 층이 주를 이루는 소셜웹에서는 맞춤법과 띄어쓰기가 맞지 않은 경우가 많으며 그로 인해 전처리 과정이 매우 중요하며, 초성어만 쓰는 경우(예:ㅇㅇ, ㄱㅈ)와 인터넷 용어들이 많이 나타난다. 그리고 계속해서 인터넷 용어들이 새로 만들어지므로 이에 대한 적절한 처리가 필요하다.

참고 문헌

- [1] B. Pang, and L. Lee, "Opinion Mining and Sentiment Analysis," Foundations and Trends in Information Retrieval: Vol.2, No.1-2, pp.1-135, 2008.
- [2] 김정호, 김명규, 차명훈, 인주호, 채수환, "한국어 특성을 고려한 감성 분류", 한국감성과학회지 제 13권, 제3호, pp.449-458, 2010.
- [3] H. Tang, S. Tan, and X. Cheng, " A survey on sentiment detection of reviews," Expert Systems with Applications, Vol.36, pp.10760-10773, 2009.
- [4] 김은영, 국어 감정 동사 연구, 전남대학교 대학원, 박사학위논문, 2004.
- [5] P. Harland, "HOW THE BRAIN FEELS," Emotion and Cognition in Neuro-Linguistic Psychotherapy, Rapport, Journal of the Association for NLP (UK), Issue 57, 2002.
- [6] R. Plutchik and H. Kellerman, Emotion: Theory, research, and experience: Vol.1, Theories of emotion.1, New York: Academic, 1980.
- [7] <http://www.cs.cornell.edu/people/pabo/movie-review-data>
- [8] <http://condensr.com>
- [9] <http://www.wjh.harvard.edu/~inquirer/homecat.htm>
- [10] 김기홍, "감정언어와 그의 문법성 고찰", 동서문

화 11, pp.161-181, 1979.

- [11] C. E. Osgood, "Cross-Cultural comparability in Attitude Measurement via Multilingual Semantic Differentials," in *Social Psychology*, pp.95-106, 1965.
- [12] 손춘섭, "정보부사의 의미와 기능에 대한 고찰," *한국어어미학회, 한국어어미학*, 제9권, pp.97-130, 2001.
- [13] J. M. Schulz, C. Womser-Hacker, and T. Mandl, "Multilingual corpus development for opinion mining," In *Proc. of LREC'10*, pp.3409-3412, 2010.
- [14] 김재원, 박훈성, 장재우, "감성어의 비중처리와 퍼지추론에 의한 평가 방법," *한국콘텐츠학회논문지*, 제9권, 제1호, pp.30-35, 2011.
- [15] <http://www.wordnet.co.kr/>

박혁로(Hyuk-Ro Park)

정회원

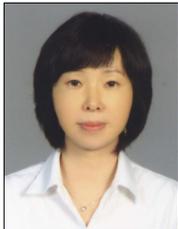


- 1987년 2월 : 서울대학교 컴퓨터 공학과(공학사)
- 1989년 2월 : 한국과학기술원 전산학과(전산학 석사)
- 1997년 2월 : 한국과학기술원 전산학과(전산학 박사)
- 1994년 ~ 1998년 : 연구개발정보센터(한국과학기술정보연구원) 선임연구원
- 1999년 ~ 현재 : 전남대학교 전자컴퓨터공학부 교수
<관심분야> : 정보검색, 자연언어처리, 데이터베이스, 인공지능

저자 소개

김도연(Do-Yeon Kim)

정회원



- 2004년 2월 : 숭실대학교 전자계산원(이학사)
- 2006년 2월 : 동국대학교 컴퓨터공학과 (공학석사)
- 2010년 3월 ~ 현재 : 전남대학교 전자컴퓨터공학과 박사수료

<관심분야> : 정보검색, 이미지 검색, 시맨틱 웹

오영(Wu Yong)

준회원



- 2010년 2월 : 대불대학교 정보보안공학과(공학사)
- 2010년 6월 : 중국 양주대학교 컴퓨터공학과(공학사)
- 2010년 3월 ~ 현재 : 전남대학교 전자컴퓨터공학과 석사과정

<관심분야> : 정보검색, 자연언어처리, 정보추출