

# 심리학적 감정과 소셜 웹 자료를 이용한 감성의 실증적 분류

## Empirical Sentiment Classification Using Psychological Emotions and Social Web Data

장문수

Moon-Soo Chang

서경대학교 컴퓨터학과

### 요 약

소셜 웹이 확산되면서 오피니언 마이닝 혹은 감정 분석 연구가 주목을 받고 있다. 감정 분석을 위해서는 감성을 판별하기 위한 감정 자원이 제공되어야 한다. 기존 감정 분석에서는 감성의 극성에 대한 강도를 표현하는 방법으로 리소스를 구축하고 이를 통하여 의견의 극성을 결정하였다. 본 논문에서는 의견의 극성뿐만 아니라 긍/부정의 근거가 되는 감성의 카테고리를 구성하고자 한다. 본 논문에서는 합리적인 분류를 위하여 심리학적 감정들을 초기 감성으로 정의한다. 그리고 실제로 소셜 웹에서 사용되는 감성의 분포를 얻기 위하여 소셜 웹의 텍스트를 분석하여 감정 정보를 추출한다. 추출한 감정 정보를 이용하여 초기 감정들을 재분류함으로써 소셜 웹을 위한 감정 카테고리를 구성한다. 본 논문에서는 이 방법을 통하여 23개의 감정 카테고리를 제시한다.

**키워드** : 감정 분석, 감정 분류, 오피니언 마이닝, 소셜 웹, 감정

### Abstract

The studies of opinion mining or sentiment analysis have been the focus with social web proliferation. Sentiment analysis requires sentiment resources to decide its polarity. In the existing sentiment analysis, they have been built resources designed with intensity of sentiment polarity and decided polarity of opinion using the ones. In this paper, I will present sentiment categories for not only polarity of opinion but also the basis of positive/negative opinion. I will define psychological emotions to primary sentiments for the reasonable classification. And I will extract the informations of sentiment from social web texts for the actual distribution of sentiments in social web. Re-classifying primary sentiments based on extracted sentiment information, I will organize sentiment categories for the social web. In this paper, I will present 23 categories of sentiment by using proposed method.

**Key Words** : Sentiment Analysis, Sentiment Classification, Opinion Mining, Social Web, Emotion.

## 1. 서 론

오피니언 마이닝(Opinion Mining) 혹은 감정 분석(Sentiment Analysis)은 인터넷이 상업적으로 이용되기 시작하면서부터 연구가 활성화되기 시작하였다. 그리고 모바일 환경과 소셜 네트워크가 급속하게 확산되면서 기존 소비자 조사나 여론 조사를 보완하는 실용적인 기술로 감정 분석 연구[1][2]가 주목받고 있다.

기존 텍스트 마이닝 연구가 정보를 추출하는데 중점을 둔 것과 달리 오피니언 마이닝에서는 텍스트에 내재된 감성으로부터 호/불호의 의견을 추출하는 것을 목적으로 한다. 인터넷 댓글이나 트위터와 같은 SNS의 텍스트는 비교적 짧은 문장에 작성자의 주관적인 의견과 표현들이 많이 포함

된 글이라는 점에서 이 분야의 연구 대상으로 많이 활용되고 있다.

텍스트에서 감정 정보를 추출하는 방법은 정보검색 기술을 이용하는 자동 추출 방법[3][4]와 심리학적 지식을 이용하는 방법, 그리고 컴퓨터 기술과 사람의 인지 능력을 이용한 반자동 추출 방법[5]로 나눌 수 있다. 자동 분류 기법은 구축 비용이 비교적 적게 든다는 것과 대용량의 정보를 이용할 수 있다는 장점이 있지만 자연어 처리 능력과 감정 관련 리소스의 부족으로 일정 이상의 성능을 내기 어렵다. 반면 심리학에서 연구되어온 감정(emotion)<sup>1)</sup> 분류를 그대로 이용하게 되면 현실적인 감정 분류와 일치하지 않는 경우가 많이 발생한다.

위 두 방법의 대안으로 제시되고 있는 반자동 추출 방법

접수일자: 2012년 9월 3일

심사(수정)일자: 2012년 10월 14일

게재확정일자: 2012년 10월 15일

본 논문은 지식경제부 산업융합원천기술개발사업(과제 번호: 10039158)에서 지원하여 연구하였음.

1) Emotion은 사전적으로 감정, 정서 등으로 번역되는데 인문학, 특히 심리학에서는 정서, 감정, 감정으로 사용되고 보편적으로 정서가 많이 사용된다. 인지과학 혹은 전산 분야에서는 일반적으로 감정으로 번역된다. 본 논문에서는 용어의 혼동을 막기 위하여 인문학과 전산학 분야를 구별하지 않고 감정으로 통일한다.

은 일반적으로 대용량 코퍼스 문장들로부터 단어들의 출현 빈도를 측정하여 그 중에서 고빈도 단어에 대해서 사람이 감성어를 찾아내는 방법을 사용한다. 그러나 이 방법에서는 사람이 감성어 후보를 볼 때 그 감성어가 사용된 용례를 확인하기 어렵기 때문에 적절한 감성을 지정하지 못할 수도 있다.

그리고 기존 오피니언 마이닝은 호/불호로 나타나는 극성(polarity)의 강도를 나타내는 연구가 대부분이고 이를 위해 적용되는 감성의 리소스들도 극성 정보를 표현하는데 초점이 맞춰져 있다. 그러나 기술에 대한 사용자의 눈높이가 올라가면 제공되는 의견도 세분화될 필요가 있다. 그 중 한 방안은 최종적인 호/불호의 의견을 제시함과 동시에 근거가 되는 감성의 종류를 제시하는 것이다. 이를 위해서는 감성의 카테고리에 대한 보다 정확한 분류 기준이 필요하다.

본 논문에서는 기존 반자동 감성 분류의 문제점을 보완하기 위하여 심리학적 감정들을 기반으로 하여 소셜 웹에 나타나는 감성들을 분류함으로써 자동 분류의 기초 자료를 구축하고자 한다. 이를 위하여 소셜 웹 데이터로부터 감성을 추출하는 프로세스를 제안하고 이로부터 일련의 감성 카테고리들을 제시한다.

## 2. 관련 연구

### 2.1 오피니언 마이닝에서의 감성 분류

감성 분석 및 오피니언 마이닝에 관한 대표적인 연구로는 Bing Liu의 연구[6]이 있다. 의견 혹은 감성은 주관적인 표현, 즉 개인적인 느낌(feelings), 관점(views), 감정(emotions), 신념(beliefs) 들로 나타난다. 이것에 대한 평가는 일반적으로 긍정과 부정의 강도로 표현되며, 이를 표현하기 위한 요소로 감성값과 함께 상황, 인물, 시간 등의 정보가 사용된다. Liu는 이것을 식 1과 같은 쿼텟플(quintuple)로 표시한다.

$$(e_j, a_{jk}, so_{ijkl}, h_i, t_l) \tag{1}$$

- $e_j$ : 대상의 엔티티
- $a_{jk}$ : 엔티티의 국면/특징
- $so_{ijkl}$ : 나머지 요소들의 상태에서의 감성값
- $h_i$ : 의견의 주체
- $t_l$ : 의견을 표현하는 시점

여기서 가장 중요한 정보는 감성값인데 감성 사전이나 감성 의미망과 같은 리소스로부터 가져오게 된다. 영어권에서 구축된 감성 관련 자료는 SentiWordNet[7], WordNet-Affect[8] 등이 있으며 이들은 기존 단어 시소러스인 WordNet에 감성 정보를 추가한 것이다. 국내에서도 영어 시소러스인 WordNet을 번역하여 감성어를 분류하는 연구[9]가 있으나, 영어와 한국어 단어가 가지는 감성도의 차이에 의한 한계가 존재한다.

SentiWordNet은 감성에 긍정(Positive)과 부정(Negative)의 양 극성을 설정하고, 여기에 감성의 강도를 주관성(subjectivity) 값으로 부여하여 그림 1과 같은 역삼각형 구조를 제안하고 있다. 구축 방법은 반자동 방법을 택하여 일부 핵심 어휘만 사람이 구축하고 나머지는 자동으로 구축하고 있다.

그림 1과 같은 방식은 감성의 극성을 가중치로 표현하기 때문에 해당 감성어가 포함된 문장이 긍정적인지 부정적인지 판단하는데 유용하다. 그러나 로봇의 감성 표현과 같은 감성 응용 분야[10]나 상세한 오피니언 마이닝을 위해 감성을 다수의 카테고리로 분류할 경우에 분류 기준을 제시하기가 어렵다.

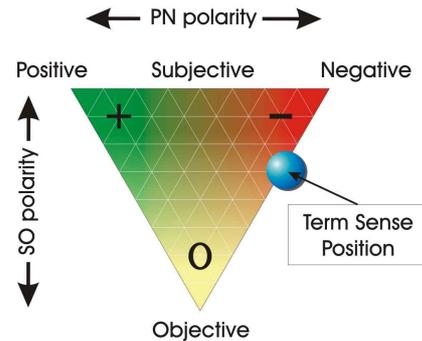


그림 1. SentiwordNet의 감성값 구조  
Fig. 1. The Structure of Sentiment Score for SentiwordNet

### 2.2 심리학적 감정 분류

심리학이나 철학 등의 인문학에서는 감정을 몇 가지 기본 감정으로 분류하고, 나머지 유사한 감정들을 기본 감정들의 조합으로 분류하고 있다. 계층적으로 감정을 분류하고 있는 대표적인 연구로 플러치크(Plutchik)의 연구들[11][12]가 있다. 플러치크는 감정 카테고리를 감정 바퀴(wheel of emotions)로 표현하고 있는데, 그림 2와 같이 8개의 기본 감정을 동심원 형태로 배치하고 있으며 감정의 강도에 따라 원의 바깥으로 갈수록 강도가 약한 감정들이 놓여져 있다. 마주보는 감정은 대체적으로 대립되는 감정을 나타낸다. 그리고 기본 감정들 사이에는 두 감정의 조합 감정을 배치하고 있다. 예를 들어, 기본 감정인 기쁨(joy)과 믿음(trust)의 조합 감정은 사랑(love)이 된다.

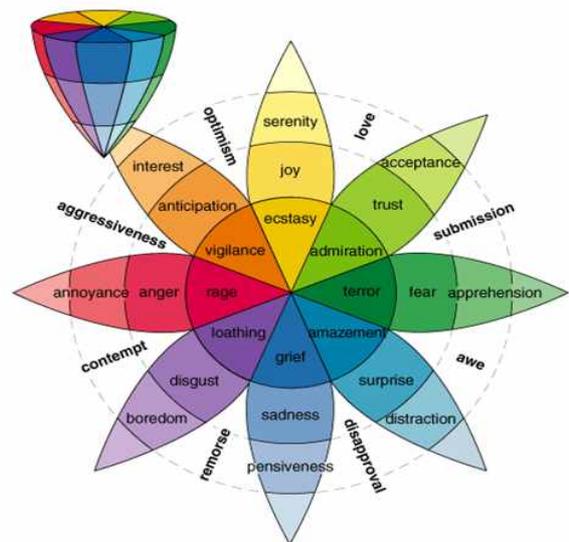


그림 2. 플러치크의 감정 바퀴  
Fig. 2. Plutchik's wheel of Emotions

국내 연구로는 플러칙의 감정 바퀴를 대립 감정 형식으로 24개의 감정으로 분류한 윤애선 연구[13]이 있다. 표 1은 윤애선 연구에서 제안하는 감정 명세로서 대립 형식에 맞추기 위하여 플러칙의 감정 중 일부를 수정하고 있다. 이 연구에서는 대립되는 감정으로 분류를 하고 있으나 극성에 따른 분류는 제시되지 않고 있다. 예를 들어 “화남(anger) “과 “두려움(fear)”는 대립되는 감정으로 표시되고 있지만 감성의 극성으로 보면 둘 모두 “부정”으로 분류된다. 인문학에서 다루는 감정 분류를 오피니언 마이닝에서 사용하기 위해서는 극성에 따라 감정을 재분류할 필요가 있다.

표 1. 대립 감정 명세  
Table 1. Opposition Emotion Specification

	감정명	대립 감정명
기본 감정	JOY_HAPPYNESS	SADNESS
	ANGER	FEAR
	SURPRISE	DESIRE_ANTICIPATION
	TRUST	DISGUST
조합 감정	LOVE	REMORSE
	DISAPPOINTMENT	PRIDE
	CONTEMPT	SHAME
	AGGRESSION	FATALISM
	OPTIMISM	PESSIMISM
	GUILT	GRATEFULNESS
	CURIOSITY	CYNISM
	ENVY	SENTIMENTALITY

### 3. 제안하는 감성 분류

감정이나 감성에 대한 연구는 언제나 연구자들 간에 이견이 있고 합의된 형태는 존재하지 않는다. 감정에 대한 기존 연구[11][14]에서 제시하는 기본 감정들에 대해서 극성으로 나누어 보면 표 2와 같이 대체로 유사하지만 분류 방식에 따라 감정의 군집이 달라진다. 여기서 중립은 긍정과 부정으로 분류되지 않고 다른 문장 요소나 상황에 따라 극성이 달라지는 경우를 나타낸다<sup>2)</sup>. 예를 들어, “놀람(surprise)”은 그 자체로는 긍정과 부정의 판단을 할 수 없지만 주변 요소들에 의해 강한 긍정이 되기도 하고 약한 부정이 되기도 한다. (예시 1)에서 “사실은 ~이었다니”라는 표현에서 “놀람”의 감정이 발견되지만 그 자체로는 극성을 판단할 수 없다. 여기에 “가장 위험한”이라는 표현에 의해 이 “놀람” 감정은 부정적 의견으로 판단된다.

(예시 1)<sup>3)</sup> ... (생략)... 제주가 사실은 가장 위험한 곳이었다니...

2) 본 논문에서의 중립은 SentiWordNet과 같은 기존 감성 분석에서 사용하는 감성값의 중립과는 다르다. 감성값의 중립은 긍/부정의 강도가 없는 글자 그대로의 중립을 뜻한다. 본 논문에서의 중립은 경우에 따라 극성이 달라지는 감성을 중립으로 둔다.

3) 논문에서 제시하는 예시문은 모두 소셜 웹의 하나인 트위터에 올라온 텍스트에서 발췌한 것으로 예시를 위해 특수 기호나 특정 상호 등은 제외하였다.

표 2. 기본 감정들의 극성 분류

Table 2. Polarity Classification of Basic Emotions

	부정적	긍정적	중립
Plutchik	Anger, disgust, fear, sadness,	Joy, acceptance, anticipation	Surprise
Ekman	Anger, disgust, fear, sadness	Joy	Surprise
Frijda	Sorrow	Happiness, interest, desire	Surprise, wonder
Tomkins	Anger, disgust, fear, contempt, shame, distress	Joy, interest	surprise
Watson	Fear, rage	love	
Weiner and Graham	Sadness	Happiness	

앞에서 언급한 것처럼 감성의 종류에 대한 분류는 관점에 따라 다르고, 또한 분류의 목적에 따라 달라진다. 본 논문에서는 이슈가 쉽게 드러나는 제품이나 국가 정책 등을 주제로 하는 소셜 웹의 텍스트를 대상으로 주요 감성을 분류하고자 한다.

상품의 경우에는 상품평에서 호감의 의견이 “감동”에서 출발한 것인지 “기대”에서 출발한 것인지에 따라 관련 회사의 대응 전략이 달라질 수 있다. 따라서 앞으로의 오피니언 마이닝에서는 의견에 대한 단순한 감성의 강도뿐만 아니라 그 의견의 근거가 되는 감성의 종류도 요구될 수 있다. 이를 위하여 본 논문에서는 합리적인 감성의 기준으로 심리학적 기본 감정들을 감성 분류 기준에 맞춰 기본 감성으로 정의하고자 한다. 그리고 실제로 소셜 웹에서 표현되고 있는 감성들의 분포를 반영하기 위하여 소셜 웹의 텍스트를 분석하여 그로부터 기본 감성 카테고리를 재분류하고자 한다. 본 논문에서는 이러한 목적을 위하여 감성 카테고리를 구축하는 방법을 제안하고 그 결과로 새로 분류된 감성 카테고리를 제시한다.

#### 3.1 기본 감정의 극성 분류

오피니언 마이닝에서는 긍정과 부정의 극성 판단이 필요하기 때문에 먼저 플러칙의 기본 감정과 조합 감정을 인지적으로 분석하여 긍정과 부정의 감성으로 분류한다. 표 3은 그 결과를 나타낸 것으로 부정의 감정이 긍정보다 많이 나타나고 있다. 이러한 경향은 다른 연구[9]에서도 나타나는 현상으로 인간의 감정이 부정적인 것에 더 민감하다는 것을 반영한다.

플러칙의 감정 분류는 8개의 기본 감정이 3단계의 강도에 따라 24개의 감정으로 나뉘고 각 기본 감정들의 조합감정들이 추가되어 약 50개의 감정 카테고리로 분류되고 있다. 그러나 과도한 감정의 수는 기계 학습이나 사람의 인지적인 분해 능력을 오히려 저하시키기 때문에 그대로 사용할 수 없다. 본 논문에서는 강도에 따른 분류를 제외한 기본 감정 8개와 조합 감정 중 기본 감정의 강도 변화로 수용될 수 있는 일부 감정을 제외한 17개 조합 감정을 포함하여 구성된 표 3의 감정들을 기본 감성으로 정의한다.

표 3. 플러칙의 감정에 대한 긍/부정 분류  
Table 3. Polarity Classification for Plutchik's Emotions

	부정적	긍정적	중립
기본 감정	공포/두려움 슬픔 혐오 분노	기쁨/즐거움 신뢰 기대	놀람
조합 감정	죄책감 불안/근심 절망 수치/창피 비관 실망 후회 굴복 부러움 냉소 체념 감상	사랑 호기심 자부심/ 긍지 우월감 낙관적임	

3.2 감정 정보 프레임

앞 절에서 정의한 기본 감성은 인지적인 판단으로 구성된 감정들이기 때문에 실제 소셜 웹의 텍스트에서 나타나는 감정과는 차이가 날 수 있다. 본 논문에서는 소셜 웹의 텍스트를 분석하여 감정 정보를 추출함으로써 기본 감정에 없는 새로운 감성을 도출하거나 기존 감성을 삭제하거나 통합하여 기본 감성의 카테고리를 재분류하고자 한다.

먼저 이를 위하여 감정 정보를 표현하는 감정 정보 프레임을 정의한다. 화자의 주관적인 의견이 들어 있는 글에는 감성을 표현하는 부분과 그 감성을 유발하거나 근거가 되는 부분이 있다. 전자를 앵커(anchor)라고 하고 후자를 트리거(trigger)라고 정의한다. 아래 (예시 2)에서 “조금은 설레인다”는 화자의 감성이 기대(anticipation)임을 나타내는 앵커이고 “드디어 ...(생략)... 한글을 깔았다” 부분은 화자의 감성이 형성되는 근거이므로 트리거가 된다. “드디어”, “조금은” 과 같은 수식어는 감성의 강도를 강화시키거나 약화시킨다.

(예시 2) 드디어 OO에서 OO 한글을 깔았다. 정말 하고 싶었던 타자법이어서 조금은 설레인다.

표 4. 감정 정보 프레임의 구성 요소  
Table 4. Elements of Sentiment Information Frame

구성 요소	내용	값	방향
감성도	감성의 강도	1 ~ 5	-
극성	감성의 극성	긍정/부정	-
앵커	화자의 감성을 직접적으로 표현	-	-
트리거	감성 표현의 이유 혹은 원인	-	-
수식어	감성의 강도 변화	1 ~ 3	강화/약화

감성에는 강도가 부여되는데 이것을 감성도라고 정의하고 1~5의 값을 부여한다. 그리고 수식어는 감성도와는 별도로 1~3의 범위의 강도값을 가지며 그 효과의 방향도 감성의

극성을 강조할 경우에는 강화, 반대의 경우에는 약화로 표기한다. 표 4에 감정 정보 프레임의 구성 요소를 나타낸다.

3.3 감정 정보 수집

본 논문에서는 감정 분류를 위해 훈련된 사람에 의해 소셜 웹의 텍스트에 대해서 표 4의 정보를 추출하였다. 분류 작업의 정확성을 위해 다음과 같은 훈련 및 작업 지침에 따라 작업을 진행하였다.

- 지침 1. 작업자는 작업 전에 기본 감성의 의미를 숙지한다.
- 지침 2. 초기 작업자는 숙련자에게 5회 이상 훈련을 받는다.
- 지침 3. 모든 작업은 크로스 체크를 위해 2인 1조로 구성한다.
- 지침 4. 작업 총괄자는 조별 결과를 검토하여 미해결 감정 정보를 전체 회의를 통해 결정한다.
- 지침 5. 새로운 감정 혹은 감정 결합 등 감정 카테고리의 변경은 후보를 설정한 후 2회분 이상의 작업을 진행한 후 결정한다.
- 지침 6. 작업 총괄자는 1회분 작업이 끝날때마다 감정값의 분포를 분석하여 다음 작업에 반영한다.

그리고 작업 과정의 편리성과 작업 상의 오류를 방지하기 위하여 그림 3과 같은 감정 분류 도구를 제작하여 사용하였다. 화면의 상단에 분석 대상 텍스트가 나오고, 팝업창에 긍정/부정/중립의 감정들을 제시하여 작업자가 선택하게 한다. 새로운 감성은 크로스 체크와 전체 회의를 통해 도출된다.



그림 3. 감정 분류 도구  
Fig. 3. Sentiment Classification Tool

감정 정보 추출 과정에서 수집을 위해 사용된 문장은 총 52,159개이고, 그 중에서 6,604개 문장에서 감정 정보를 추출하였다.

3.4 감정 재분류

3.1절에서 정의한 기본 감성은 감정 정보 추출 과정에서 감정 정보의 분포에 따라 재분류된다. 본 논문에서는 재분류 과정에서 표 5에 나타낸 것과 같은 몇 가지 규칙을 제시한다.

첫번째 규칙으로 새로 생성하는 감성들은 보편적인 슬어식 표현을 지향한다. 명사의 경우 그 단어의 의미에 종속되는 경향이 크기 때문에 대표 단어 선택이 어렵다. 그리고 감성이 상태나 행위를 나타내므로 슬어식 표현이 더 적합하다. 플러칙의 8개 기본 감성 중에서 공포/두려움, 슬픔, 분노, 기쁨, 기대는 그대로 남겨나 슬어식 표현으로 변경된다.

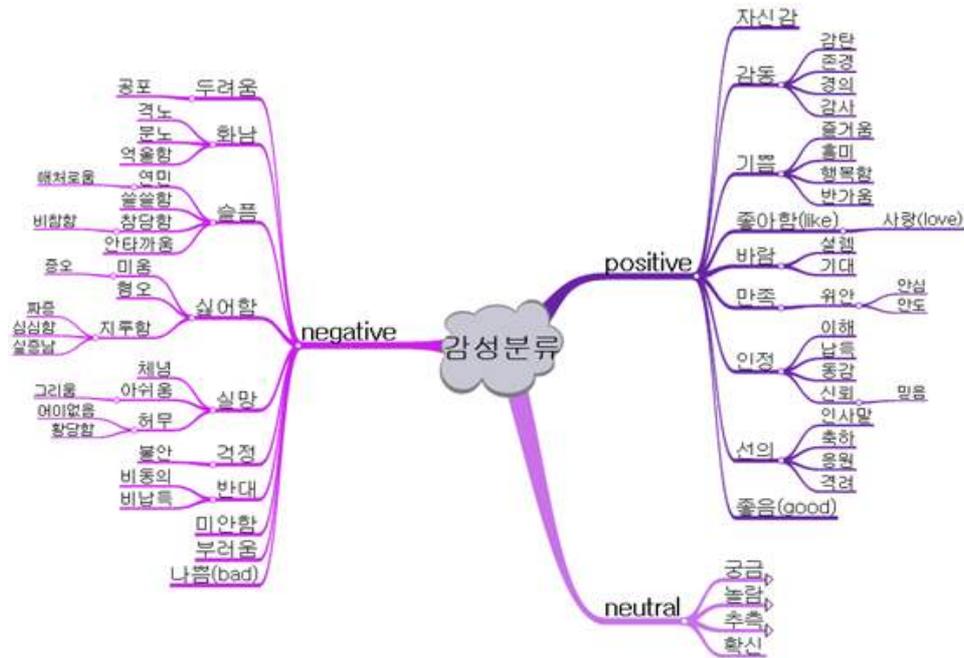


그림 4. 재분류된 감정 카테고리  
Fig. 4. Re-classified Sentiment Categories

혐오는 미움, 지루함 등과 함께 싫어함으로 통합되고, 신뢰는 이해, 납득, 동감 등과 함께 인정으로 통합된다. 통합되는 감정들은 감정 정보 추출에서 구분이 애매하거나 빈번하게 교차 선택되는 감정들을 모아서 회의를 통해 통합한 것들이다. 통합 감성의 또 하나의 기준은 개별적인 감성은 다르지만 공통의 감성이 존재하고 그것이 다른 감정과 구별되는 경우인데, 미움, 혐오, 지루함 등이 여기에 해당한다. 이 감정들은 대상에 대해서 싫어하는 감정이 공통적으로 존재하며 이것은 두려움이나 슬픔 등의 다른 감정들과 구별된다.

그리고 부정적 감정과 긍정적 감성이 대비될 수 있는 감성은 대비되는 단어로 통합한다. 기쁨/슬픔, 좋아함/싫어함, 실망/만족 등이 이에 해당한다. 선의, 미안함, 반대 등은 빈번하게 등장하지만 화자의 감정을 나타내는 표현이 아니기 때문에 기존 감정 분류에는 없으므로 새로운 감정으로 추가한다.

감정 정보 추출에서 화자의 감정을 직접적으로 표현하는 앵커는 생략되고, 언급하고자 하는 대상에 대한 평가만 트리거로 나타나는 경우가 매우 빈번하게 나타났다. 이를 위해 부정 영역의 나쁨(bad)과 긍정 영역의 좋음(good)을 추가한다.

표 5. 감정 분류 수정 규칙  
Table 5. Rules of Sentiment Category Modifying

규칙	내용	사례
1	가급적 술어식 표현을 사용한다.	좋아함(like)
2	대비되는 감성은 대비되는 용어를 사용한다.	기쁨/슬픔
3	자주 혼동되는 감성은 통합 후보로 등록한다.	
4	공통된 감성이 존재하면 통합 후보로 등록한다.	
5	통합 후보들의 대표 감성이 다른 감정과 구별되면 통합한다.	인정(이해, 납득, 동감, 신뢰)
6	앵커가 생략되고 트리거만 존재되 하나의 감정으로 유추가 가능하면 유추되는 감성을 등록한다.	미안함, 반대
7	앵커가 생략된 트리거가 매우 빈번하게 출현하면 감정 후보로 등록한다.	좋음, 나쁨

### 3.5 수정된 감정 카테고리

이상의 과정으로 감정 카테고리를 점진적으로 수정하여 본 논문에서는 그림 4에 나타내는 것과 같이 긍정 9개, 부정 10개, 중립 4개, 합계 23개의 감정으로 분류한다. 그림 4에서 기본 감정들에서 분화된 작은 가지는 각 감정들의 유사 감정들을 나타낸 것이다.

긍정과 부정의 감정들은 대체적으로 대비되는 감정으로 구성되어 있으며, 대표 감정으로 통합된 유사 감정들은 실제 텍스트에 나타난 감정 정보를 반영하고 있기 때문에 기계 학습과 같은 자동 분류에서 활용이 가능하다.

또한, 대표 감정들은 그림 4에서 하단에서 상단으로 갈수록 극성의 강도가 커지도록 배치하고 있다. 단, 본 논문에서는 극성의 강도를 계량화하지 않고 있다. 향후 연구로 계량화 기준을 제시할 필요가 있다.

그림 5는 감정정보 추출 과정에서 각 감정들이 나타난 문장들의 수를 그래프로 나타낸 것이다. 긍정에서는 바람과 좋음이 많이 나타났으며, 부정에서는 싫어함, 실망, 나쁨 등이 비교적 많이 나타났다. 좋음과 나쁨이 많이 나왔다는 것은 본 논문에서 분류의 큰 기준이 되고 있는 화자의 감정 이외에도 극성을 분류할 수 있는 요소가 더 있다는 것을 시

사한다. 이와 같은 요소들은 트리거에 나타나는 것으로 트리거에 대한 분류 기준이 필요한 것으로 사료된다.

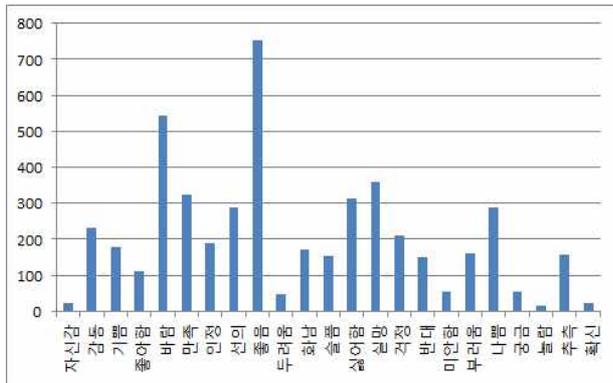


그림 5. 감성 분포 그래프  
Fig. 5. Sentiments Distribution Graph

### 3.6 감성 분류의 활용

본 논문에서 분류한 감성들은 기존 오피니언 마이닝에서 활용하고 있는 극성에 대한 가중치 분류와 달리 심리학에서 제시하는 감정 분류를 바탕으로 감성의 카테고리를 나눈 것이다. 따라서 기존 감성 분석에서 제시하지 못하는 긍/부정의 근거를 제시하거나 결과적으로 나타난 사건에 대한 추론의 근거로 사용될 수 있다. 예를 들어 소셜웹의 이슈들에 대한 대중들의 의견을 분류하고, 나아가 관련 기관이나 기업의 대응 전략을 위한 근거를 제시할 수 있다. 한 기업의 제품에 대한 소비자 평가가 그림 5의 분포보다 실망이 크게 나타난 경우와 화남이 두드러진 경우에 대해서 다른 대응을 할 수 있다.

또한, 본 논문의 자료는 훈련된 사람에게 의해서 구축된 자료로서 자동분류의 학습 자료로 활용될 수 있다. 이를 통하여 대응량의 분류 능력을 갖추면 사회 현상 연구나 마케팅 등의 새로운 분야에 대한 감성 분석도 가능할 것으로 기대된다.

## 4. 결론 및 향후 연구

오피니언 마이닝을 위한 기존 감성 분류는 긍정과 부정의 극성에 대한 감성의 강도를 설정하는 것이 일반적이다. 이 방식은 텍스트의 의견의 방향을 결정하기에는 적합하지만 그 의견에 대한 감성의 종류와 같은 상세한 오피니언 마이닝을 수행하기에는 충분한 정보를 제공하지 못한다. 본 논문에서는 감성의 극성과 함께 감성의 카테고리를 분류하는 방법을 제안함으로써 새로운 감성 분류 방안을 제시하고자 하였다. 이를 위하여 본 논문에서는 심리학에서 제시하는 감정 분류를 기본 감성으로 하고, 감성 정보 프레임 설정하여 이를 기준으로 소셜 웹의 텍스트를 분석하여 기본 감성을 수정하는 방법을 제안하였다. 그리고 이 방법을 통해 구성된 감성 카테고리를 제시하였다.

본 연구에서는 감정 분류를 기준으로 작업을 했기 때문에 트리거에 많이 나타나는 감성, 즉 화자의 감정이 아닌 대상에 대한 평가 의견에 대해서 자세한 검토를 하지 않았다. 향후에는 트리거에 대한 상세한 분석 연구를 계획하고 있다. 그리고 긍/부정으로 대비되는 감성을 추출하게 됨

으로써 긍정과 부정의 감성 수가 비슷하게 나왔다. 이것은 부정의 감성이 자세히 분류되는 일반적인 현상과 배치되는 결과로서 이에 대한 보완 연구가 필요하다.

## References

- [1] Sang-Do Kim, Eeong-Bae Park, Sang-Jo Lee and Kweon-Yang Kim, "A Syllable Kernel based Sentiment Classification for Movie Reviews", *Journal of Korean Institute of Intelligent Systems*, vol. 20, no. 2, pp. 202-207, 2010.
- [2] Kweon-Yang Kim and Chang Suk Kim, "A String Kernel based Sentiment Classification for Blog Text," *Proceedings of KIIS Fall Conference 2009*, vol. 19, no. 2, pp. 199-201, 2009.
- [3] Han-Hoon Kang, Seong-Joon Yoo and Dong-Il Han, "Automatic Extraction of Korean Opinion Words Using PMI-IR and Performance Improvement Method," *Proceedings of KIIS Spring Conference 2010*, vol. 20, no. 1, pp. 318-321, 2010.
- [4] Jong-Seok Song and Soo-Won Lee, "Automatic Construction of Positive/Negative Feature-Predicate Dictionary for Polarity Classification of Product Reviews," *Journal of KIISE: Software and Applications*, vol. 38, no. 3, pp. 157-168, 2011.
- [5] Jae-Seok Myung, Dong-Joo Lee and Sang-Goo Lee, "A Korean Product Review Analysis System Using a Semi-Automatically Constructed Semantic Dictionary," *Journal of KIISE: Software and Applications*, vol. 35, no. 6, pp. 392-403, 2008.
- [6] Bing Liu, *Sentiment Analysis and Opinion Mining*, Morgan & Claypool, 2012.
- [7] Andrea Esuli and Fabrizio Sebastiani, "SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining," *Proc. of the 5th conf on Language Resources and Evaluation(LREC'06)*, pp. 417-422, 2006.
- [8] Carlo Strapparava and Alessandro Valitutti, "Wordnet-Affect: an Affective Extension of WordNet," *Proc. of the 4th Int. Conf. on Language Resources and Evaluation(LREC'04)*, pp. 1083-1086, 2004.
- [9] Jae-Won Hwang and Young-Joong Ko, "A Korean Emotion Features Extraction Method and Their Availability Evaluation for Sentiment Classification," *Korean Journal of Cognitive Science*, vol. 19, no. 4, pp. 499-517, 2008.
- [10] Kwee-Bo Sim, Kwang-Sub Byun and Chang-Hyun Park, "Emotional Expression System based on Dynamic Emotion Space," *Journal of Intelligent and Fuzzy Systems*, vol. 15, no. 1, pp. 18-23, 2005.
- [11] Robert Plutchik and Henry Kellerman, *Emotion: Theory, Research, and Experience Vol.5*, Academic Press, 1990.

- [12] Robert Plutchik, *Emotions and Life*, American psychological association, 2002.
- [13] Ae-Sun Yoon and Hyuk-Chul Kwon, "Component Analysis for Constructing an Emotion Ontology," *Korean Journal of Cognitive Science*, vol. 21, no. 1, pp. 157-175, 2010.
- [14] Andrew Ortony and Terence J. Turner, "What's Basic About Basic Emotions?," *Psychological Review*, vol. 97, no. 3, pp. 315-331, 1990.
- 

저 자 소 개



**장문수(Moon-Soo Chang)**

1992년 : 고려대학교 전자전산공학과 공학사

1994년 : 고려대학교 전자공학과 공학석사

2001년 : 동경공업대학 지능시스템과학전공  
공학박사

2000년 ~ 2003년 : 한국전자통신연구원  
선임연구원

2003년 ~ 현재 : 서경대학교 컴퓨터과학과 조교수

관심분야 : Natural Language Understanding, Knowledge Mining, HCI

Phone : 02-940-7754

E-mail : cosmos@skuniv.ac.kr