

# 감성분석 결과와 사용자 만족도와의 관계 -기상청 사례를 중심으로-

## Relationship between Result of Sentiment Analysis and User Satisfaction -The case of Korean Meteorological Administration-

김인겸\*, 김혜민\*, 임병환\*, 이기광\*\*  
국립기상과학원 연구기획운영과\*, 단국대학교 경영학과\*\*

In-Gyum Kim(kimig@korea.kr)\*, Hye-Min Kim(hmkim84@korea.kr)\*,  
Byunghwan Lim(weatherman@korea.kr)\*, Ki-Kwang Lee(kiklee@dankook.ac.kr)\*\*

### 요약

기상청에서 현재 시행되고 있는 만족도 설문조사의 한계를 보완하기 위해 SNS를 통한 감성분석이 활용될 수 있다. 감성분석은 2011~2014년 동안 '기상청'을 언급한 트위터를 수집하여 나이브 베이즈 방법으로 긍정, 부정, 중립 감성을 분류하였다. 기본적인 나이브 베이즈 방법에 긍정, 부정, 중립의 각 감성에서만 출현한 형태소들로 추가사전을 만들어 감성분석의 정확도를 향상시키는 방법을 제안하였다. 분석결과 기본적인 나이브 베이즈 방법으로 감성을 분류할 경우 약 75%의 정확도로 학습데이터를 재현한데 반해 추가사전을 적용할 경우 약 97%의 정확성을 보였다. 추가사전을 활용하여 검증자료의 감성을 분류한 결과 약 75%의 분류 정확도를 보였다. 낮은 분류 정확도는 향후 기상 관련의 다양한 키워드를 포함시켜 학습데이터 양을 늘려 감성사전의 질을 높임과 동시에 상시적인 사전의 업데이트를 통해 개선될 수 있을 것이다. 한편, 개별 어휘의 사전적 의미에 기반한 감성분석과 달리 문장의 의미에 기반하여 감성을 분류할 경우 부정적 감성 비율의 증가와 만족도 변화 추이를 설명할 수 있을 것으로 보여 향후 설문조사를 보완할 수 있는 좋은 수단으로 SNS를 통한 감성분석이 활용될 수 있을 것으로 사료된다.

■ 중심어 : | 트위터 | 감성분석 | 나이브 베이즈 | 만족도 |

### Abstract

To compensate for limited the satisfaction survey currently conducted by Korea Metrological Administration (KMA), a sentiment analysis *via* a social networking service (SNS) can be utilized. From 2011 to 2014, with the sentiment analysis, Twitter who had commented 'KMA' had collected, then, using Naïve Bayes classification, we were classified into three sentiments: positive, negative, and neutral sentiments. An additional dictionary was made with morphemes appeared only in the positive, negative, and neutral sentiments of basic Naïve Bayes classification, thus the accuracy of sentiment analysis was improved. As a result, when sentiments were classified with a basic Naïve Bayes classification, the training data were reproduced about 75% accuracy rate. Whereas, when classifying with the additional dictionary, it showed 97% accuracy rate. When using the additional dictionary, sentiments of verification data was classified with about 75% accuracy rate. Lower classification accuracy rate would be improved by not only a qualified dictionary that has increased amount of training data, including diverse keywords related to weather, but continuous update of the dictionary. Meanwhile, contrary to the sentiment analysis based on dictionary definition of individual vocabulary, if sentiments are classified into meaning of sentence, increased rate of negative sentiment and change in satisfaction could be explained. Therefore, the sentiment analysis *via* SNS would be considered as useful tool for complementing surveys in the future.

■ keyword : | Twitter | Sentiment Analysis | Naive Bayes | Satisfaction |

\* 본 연구는 국립기상과학원 2016년도 주요사업인 “기상업무지원기술개발연구(NIMS-2016-3100)” 사업에 의하여 수행되었습니다.

접수일자 : 2016년 07월 08일

수정일자 : 2016년 08월 22일

수정일자 : 2016년 08월 22일

교신저자 : 이기광, e-mail : kiklee@dankook.ac.kr

### I. 서론

기상서비스는 위험기상 현상에 대비하여 사용자들이 적절한 행동을 취하는 의사결정에 활용되지 않는다면 효용이 없다. 그런 의미에서 기상청과 같은 기상서비스 제공자들은 사람들이 재산피해를 최소화하거나 자신의 안전을 보장하는데 기상정보를 어떻게 사용하는지 아는 것이 필수적으로 요구된다. 사용자들이 기상정보를 의사결정에 효율적으로 활용하고 있는지 알아보거나, 앞으로 제대로 활용할 수 있도록 돕기 위해선 기상정보 제공자들과 사용자 간의 커뮤니케이션이 중요하다. 연구자들은 예·특보와 같은 특정 기상정보에 대한 사용자들의 인식을 조사함으로써 효과적인 커뮤니케이션을 지원하고자 하였다[1-8]. 기상청도 일반인과 전문가를 대상으로 기상서비스에 대한 만족도를 설문조사함으로써 대중과의 커뮤니케이션을 시도해 오고 있다[9]. 기상청에서 실시 중인 만족도 조사는 7점 리커트 척도를 적용하여 ‘서비스에 대해 얼마나 만족하십니까?’란 질문에 응답자들이 자신의 만족 수준을 선택하는 방식이다. 그러나 설문조사는 비용이 많이 들고, 조사의 적시성이 떨어지는 단점이 제기될 수 있다[10]. 한편, ‘만족’과 같은 추상적 인식에 대한 대답을 얻는 것이 조사의 목적이라면, SNS(Social Network Services)자료를 활용한 감성분석이 설문조사를 보완하는 좋은 대안이 될 수 있다. 기상청은 2014년 트위터와 블로그자료를 수집하여 기상일반, 기상서비스 등 6개 대분류와 기관, 정책, 장비, 날씨 등 21개 중분류에 대해 210개에 달하는 키워드에 나타난 사용자 감성을 광범위하게 분석하였다. 그리고 [그림 1]은 많은 결과들 중에서 ‘기상청(기상청장 포함)’을 포함하는 글에 나타난 감성분석 결과이다[11].

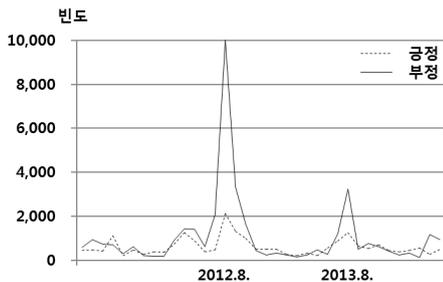


그림 1. ‘기상청’에 대한 감성 추이(기상청, 2014)

[그림 1]에서 2012년 8월과 2013년 8월에 부정적 감성이 증가하였는데, 특히 2012년 8월에 상승폭이 컸던 것을 확인할 수 있다. 보고서에는 2012년 당시 태풍 ‘볼라벤’으로 인한 피해, 2013년은 기상청이 장마 종료를 알린 이후에도 계속 비가 내린 것에 대한 불만이 원인인 것으로 보고되었다. 그런데 비슷한 시기에 수행된 만족도 설문조사 결과는 감성분석 결과와 상반되는 것으로 나타났다[그림 2]. 2012년 11월에 수행된 설문조사에서 20대 일반인의 기상서비스에 대한 만족도는 68.6점이었고, 2013년 하반기 만족도는 62.2점이었다. [그림 1]의 감성분석 결과에 따르면 2012년 8월에 부정적 감성이 크게 늘어났으므로, 시차를 고려하더라도 9~11월에 조사된 만족도가 크게 낮을 것으로 예상할 수 있다. 그리고 2013년 8월의 부정적 감성이 다른 시기에 비해 증가했을지라도 2012년 8월에 비해서는 훨씬 적게 나타났다으므로, 2013년 하반기의 만족도는 2012년보다 높을 것으로 판단할 수 있는데 오히려 크게 하락하였다.

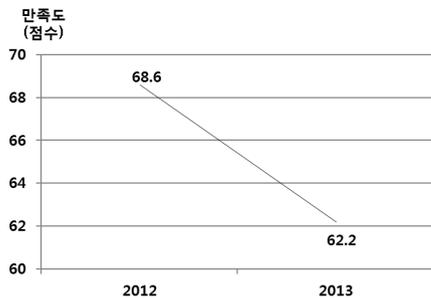


그림 2. 20대 일반인의 만족도 변화

이는 [그림 1]에 나타났던 2012년의 부정적 감성 증가가 기상청에 대한 직접적 불만이 아니라 태풍으로 인한 ‘피해’라는 부정적 의미의 어휘 때문일 수 있다는 것에 주목할 필요가 있다. 실제로 태풍 상륙 전후에는 사람들이 예·경보와 피해 상황에 대한 뉴스에 많이 노출되어 관련 내용을 SNS를 통해 알렸을 수 있다. 한편, 기상청(2014)은 보고서에서 ‘피해가 적어서 다행이니 기상청을 욕할 필요 없다’의 경우처럼 부정적 의미의 개별 어휘를 포함하고 있지만, 실제로는 긍정 혹은 중립으로 분류해야 할 문장이 있는 것으로 보고하기도 하

였다[11]. 결국 감성분석 과정에서 ‘피해’나 비속어와 같은 개별 어휘가 감성분류에 큰 영향을 끼쳐 특정 시기의 부정적 감성이 증가한 결과를 초래한 것이다. 2014년 수행된 분석에서 기상청에 대한 부정적 감성이 사용자 인식을 제대로 반영하지 못한 이유는 210개에 달하는 키워드를 모두 분석하기 위한 감성사전을 활용했기 때문으로 판단된다. 감성사전 제작시 각 어휘에 대한 감성 가중치 설정이 중요한데, 키워드가 다양하여 개별 키워드에 대한 감성분석의 정확성이 낮아졌을 수 있다. 2014년 보고서와 같이 개별 어휘의 사전적 의미를 기준으로 분류한 감성분석 결과는 그 자체로 일반인들이 기상청에 대해 관심이 높았던 시기를 검색하는데 유용할 수 있으나, 만족도 조사 결과와 연계하여 정책에 활용하는 등의 세부분석을 실시할 때 자칫 의사결정자들에게 혼란을 야기할 수도 있음을 알 수 있다.

이에 본 연구에서는 ‘기상청’을 언급한 트위터 만을 검색하여 2014년에 수행했던 어휘 기반의 감성분석이 아니라 실제로 트위터 작성자의 의도가 기상청을 향하고 있는지, 즉 문장이 내포하고 있는 실제 의미를 기반으로 감성분석을 실시하였다. 그리고 감성분석 결과와 설문조사의 만족도와 관련될 수 있는지 살펴봄으로써 향후 실시될 기상청의 대국민 인식조사에 도움이 되고자 한다.

## II. 선행연구

감성분석은 뉴스, 블로그, SNS, 상품 리뷰 등의 텍스트로부터 긍정, 부정, 중립 등의 감성정보를 추출하는 방법론으로서, 분류 결과는 조사 대상에 대한 전반적인 의견, 관점, 인식 등에 관심이 있는 분석가들에게 유용하게 활용될 수 있다[12][13]. 감성분석은 주로 분석 자동화 및 감성사전 구축을 통한 방법론의 정확도 향상에 관한 연구를 시작으로[14-16], 최근에는 범위를 확장하여 주식가격 예측 등 광범위하게 활용되고 있다[17][18]. 감성분석이 자연어 처리를 위한 접근법이라는 점에서 트위터와 같은 마이크로 블로그 자료들이 사용되었는데, 그러한 연구들은 나이브 베이즈, SVM(Support Vector Machine) 등의 모델 성능을 비교하거나, 어휘/

문장/이모티콘 등을 구분하여 감성분석을 시도하였다[19][20].

한편, 마케팅 부문에서 서비스 향상을 위해 감성분석 결과를 만족도 혹은 선호도와 관련시키기도 하였다. O'Connor et al.(2010)은 소비자 신뢰도 및 정치적 의견과 트위터 메시지에 나타난 감성이 빈도 사이에 높은 상관관계가 있음을 보였고, Kang·Park(2014)는 모바일 서비스에 대한 사용자 만족도를 모니터링 하는데 감성분석과 다기준 의사결정모형을 제시하였다[21][22]. 김장석 등(2015)는 화장품 용기 디자인 개발을 위한 사용자 선호 요인을 파악하는데 감성분석을 활용한 바 있다[23].

## III. 연구자료

감성분석을 위해 ‘기상청’을 포함하는 트위터 자료(이하 ‘트윗’)를 수집하였다(출처: search.twitter.com). 수집기간은 2011~2014년 하반기 실시된 만족도 설문조사의 종료시점으로부터 100일 이전까지의 기간으로서 각각 2011.8.23.~11.30., 2012.8.16.~11.23., 2013.7.21.~10.28., 2014.7.18.~10.25. 동안 작성된 트윗이다. 자료 수집기간을 연별로 동일하게 설정하지 않고, 만족도 설문조사가 이루어진 시기에 맞춘 이유는 감성분석 결과를 2012, 2013년의 만족도 변화 추이와 비교하기 위함이다. 수집된 트윗의 수는 2011년 12,571개, 2012년 11,020개, 2013년 9,009개, 2014년 3,672개이다. 트윗 자료와 비교할 만족도는 트위터의 주 사용연령이라 할 수 있는 20대 일반인들의 설문결과를 활용하였다. 만족도의 경우 기상분야 전문가 혹은 고령의 응답자들에게서 비교적 높게 나타나는 경향이 있는데, ‘만족도 관리’ 측면에서 비판적 성향의 응답자들을 세부적으로 분석하는데 SNS를 통한 감성분석이 효과적일 수 있을 것이다.



그림 3. 감성분석 과정

#### IV. 감성분석

감성분석의 첫 번째 단계는 자연어를 문장의 기본단위인 형태소로 구분하는 것이다. 형태소 분석기는 입력된 문장을 어절 단위로 구분한 다음 어휘의 가장 기본단위인 사전의 표제어로 분석하여 품사와 함께 출력해준다. 본 연구에서는 트위터 문장의 형태소 분석을 위해 KAIST(Korea Advanced Institute of Science and Technology)에서 제작한 자바 기반의 “한나눔 형태소 분석기”를 사용하였다[24]. 한나눔 형태소 분석기는 사용자가 단어를 추가하여 분석의 성능을 높일 수 있으나, 본 연구에서는 어떤 사용자 단어도 추가하지 않고 분석기 자체의 성능에만 의존하였다. 소프트웨어 버전은 웹페이지에서 다운로드 받을 수 있는 가장 최신 버전인 0.8.4이다. 형태소 분석기에 의해 구분된 모든 어휘들은 가공 및 수정하지 않고, 분류된 그대로 감성사전에 저장하였다. 만약 형태소분석기가 ‘좋~다’, ‘좋다’, ‘좋-다’ 등의 어휘들을 구분하였다면, ‘좋다’라는 기본형으로 가공하지 않고, 모두 개별적으로 저장하였다.

문장의 형태소 분석과 감성사전 제작에 앞서 각 트윗에 대해 연구자들이 직접 감성을 분류하였다. 분류작업에는 연구원 3인이 참여하였고, 그들은 사전 회의를 통해 감성을 분류할 규칙을 공유한 후 각자가 수집된 모든 트윗에 대해 감성을 분류하였다. 1차로 감성 분류된 트윗들을 병합하여 연구자 모두가 동일한 감성으로 분류한 트윗은 그대로 사전에 저장하고, 그렇지 않은 트윗들은 회의를 통해 2명 이상이 동의하는 감성으로 결정하여 최종 분류하였다. 분류 감성은 긍정, 부정, 중립의 3종류이다.

##### 1. 나이브 베이즈 분석

제작된 감성사전을 기본으로 2012~2014년 트위터에 대해 나이브 베이즈 방법을 사용하여 감성분석을 실시하였다. 나이브 베이즈는 베이즈 정리를 적용하여 문서를 분류하는 확률 분류기의 일종으로 감성분석 분야에서 텍스트 속성을 분류하는데 자주 활용되는 간단한 모델이다[25]. 관련 연구로는 Kang et al. (2012)이 기존의

나이브 베이즈 방법 및 SVM과 비교하여 정확성이 개선될 수 있는 나이브 베이즈 알고리즘을 제시하였고, Dhande·Patnaik(2014)는 나이브 베이즈와 인공신경망을 결합하여 영화 리뷰 데이터셋의 긍정·부정 분류 정확도를 80.65%까지 향상시킬 수 있음을 보였다[25][26]. 최근에는 나이브 베이즈 단일 방법의 분류 정확도 향상뿐만 아니라, 텍스트의 극성 판별 혹은 키워드 추출의 정확도를 향상시키기 위해 나이브 베이즈, SVM 등을 효과적으로 결합하는 앙상블 기법 연구도 진행되고 있다[27][28]. 이처럼 나이브 베이즈는 자연어 처리 분야에서 다양하게 활용되고 있는 방법인데, 그 자체만으로도 잘 전처리된 데이터를 활용할 경우 충분히 효과적인 분석결과를 보여줄 수 있다고 알려져 있다. 본 연구는 기본적인 나이브 베이즈 방법만을 사용하여 트위터 감성을 분류하였다. 나이브 베이즈의 방법은 다음과 같다.

나이브 베이즈는 수집된 원시자료에 대해 학습을 선행한 뒤 입력된 벡터들의 조건부 확률을 계산하여 구분될 클래스 즉 각각의 감성들 중 하나로 분류한다. 입력 벡터  $D$ 는 식(1)처럼 다수의 요소( $e$ )로 이루어지는데, 본 연구의 경우에는 입력 벡터가 문장, 각 요소는 형태소가 된다.

$$D = e_1, e_2, e_3, \dots, e_n \quad \text{식(1)}$$

각각의 트윗을 의미하는  $D$ 가 어느 클래스 즉, 감성에 포함되는지 결정하기 위한 조건부 확률은 식(2)로 나타낼 수 있다. 식(2)에서  $C$ 는 구분될 감성,  $i$ 는  $I$ 번째 감성을 의미한다. 본 연구에서  $C$ 는 긍정, 부정, 중립으로 구분하였다.

$$P(C_i|D) = P(C_i|e_1, e_2, e_3, \dots, e_n) \quad \text{식(2)}$$

식(2)는 베이즈 정리에 의해 식(3)으로 변환된다.

$$P(C_i|D) = \frac{P(D|C_i) \times P(C_i)}{P(D)} \quad \text{식(3)}$$

나이브 베이즈는 트윗의 모든 감성에 대해 식(3)을

기반으로 계산한 뒤 가장 높은 결과값의 감성을 선택한다. 이 과정에서  $P(D)$ 는 감성에 관계없이 동일하므로 생략될 수 있다. 그리고 나이브 베이즈는 트윗을 구성하는 모든 형태소들이 서로 독립적임을 가정한다. 즉, 어느 한 문장 내에서 A 형태소가 출현한 것은 B 형태소의 출현에 어떤 영향도 주지 않는다. 이러한 요소간의 독립성을 가정함으로써 식(3)은 식(4)로 나타낼 수 있다.

$$\frac{P(D|C_i) \times P(C_i)}{P(D)} \rightarrow P(e_1|C_i) \times P(e_2|C_i) \times \dots \times P(e_n|C_i) \quad \text{식(4)}$$

식(4)의  $P(e_n|C_i)$ 는 개별 형태소인  $e_n$ 이 감성 클래스  $C_i$ 에 속할 조건부 확률이고, 식(4)를 통해 개별 문장이 어떤 감성인지 판단할 수 있게 된다. 그런데 식(4)는 두 가지 문제점이 있다. 그것은 ① 입력벡터에서 출현한 특정 요소가 학습 벡터에 존재하지 않으면 해당 조건부 확률은 0으로 계산된다. 즉, 감성사전에 존재하지 않는 형태소가 트윗에 출현하게 되면 해당 형태소가 특정 감성에 속할 확률이 항상 0이 되고, 조건부 확률의 곱으로 나타난 식(4)의 결과가 0으로 도출된다. ② 0보다 작은 값의 확률 값들을 식(4)처럼 모두 곱해줄 때, 요소의 수가 많아지면 결과 값이 0에 가까워지는 언더플로 현상이 발생할 수 있다. 이의 해결을 위해 식(4)에 Laplace smoothing과 Log 변환을 적용해 준다. Laplace smoothing으로  $P(e_n|C_i)$ 를 계산하는 방법은 식(5)와 같다.

$$P(e_n|C_i) = \frac{\text{count}(e_n, C_i) + 1}{|C_i| + |V|} \quad \text{식(5)}$$

$|C_i|$ 는 트윗의 감성 클래스  $C_i$ 에 속한 모든 형태소의 수이고,  $|V|$ 는  $|C_i|$ 에서 중복되는 형태소를 제외한 수이다. 그리고 식(5)에서처럼 Laplace smoothing은 분자에 1, 분모에  $|V|$ 를 공통으로 더해주는 방법이다. 그리고 식(4)의 좌변과 우변에  $\log$ 를 취하면 언더플로 현상을 방지할 수 있게 된다(식 6).

$$\log(P(e_1|C_i)) + \log(P(e_2|C_i)) + \dots + \log(P(e_n|C_i)) + \log(P(C_i)) \quad \text{식(6)}$$

## 2. 추가 감성사전 제작

감성사전에 수록된 형태소들은 품사의 기본형으로 변경되지 않고, 한나눔 형태소 분석기에 의해 분류된 그대로 저장되었다. 전술한 것처럼 ‘좋다’, ‘좋.다.’, ‘좋~다’, ‘좋-다’ 등 트위터 사용자가 임의로 작성한 표현들은 실제로 모두 ‘좋다’라는 기본형으로 구분될 수 있으나, 본 연구에서는 각각의 형태 그대로 감성사전에 저장하였다. 만약 작성자에 의해 변형된 다양한 표현들을 모두 기본형으로 변환한 뒤 저장했다면, ‘좋다’라는 감성어는 긍정, 부정, 중립의 감성사전에 모두 포함될 수 있을 것이다. 하지만, 서로 다른 표현들을 모두 별개로 저장했기 때문에 변형된 어휘들은 특정 감성에서만 사용되었을 수 있다. 즉, ‘좋~다’라는 형태소는 긍정, 부정문에서 사용되었지만, ‘좋.다.’라는 형태소가 부정문에서만 사용된 경우 ‘좋.다.’는 감성사전에서 부정적 어휘로 구분되어 문장을 부정문으로 분류하는데 역할을 할 것이다. 이 때 [그림 4]의 긍정, 부정, 중립 집합에서 음영 부분을 제외한 부분 즉, 각 감성에서만 출현한 형태소들로 추가 사전을 만드는 것을 생각해 볼 수 있다.



그림 4. 감성사전에 포함되는 형태소들의 집합 개념도

이는 매우 간단한 방법으로 기존의 감성사전 외에 추가로 사전을 만드는 것이다. 추가사전을 적용한 나이브 베이즈 방법은 [그림 3]에서 형태소 분석과 감성사전 제작 과정의 사이에 추가사전을 제작·참고하는 단계가 추가된다. 만약 분석할 트윗에 추가사전 내의 형태소가 포함되어 있다면, 감성사전을 사용하기 전에, 추가사전에 포함된 형태소의 감성으로 즉시 분류한다.

V. 결과 및 고찰

감성 분류한 트위터 본문의 예시는 [표 1]에 나타나 있다. 기상청의 블로그 담당자가 작성하거나 기상정보를 객관적으로 전달하는 트윗이 대다수인 중립 감성은 표에 나타나지 않았다. 긍정 트윗에는 짱, 민다, 최고 등 신뢰와 관련된 단어들 포함되었고, 부정 트윗에는 구라청, 욕설, 세금, 슈퍼컴퓨터 등이 문장에 포함되어 조롱 섞인 비난을 하고 있었다. 긍정, 부정 트윗 모두 주로 강우와 관련된 것으로 나타나 2011년 설문조사에서 일기예보 이용시 일반인들이 가장 관심을 갖는 기상요소로 강수유무(강수유무 55.5%, 예상기온 29.2% 등)를 선택했던 것과 관련있는 것으로 판단되었다[29].

표 1. 감성분류된 실제 트윗 내용 예시

감성	트윗
긍정	<ul style="list-style-type: none"> <li>역시....기상청이 짱이얍</li> <li>그래 기상청이 아주 내처럼 바로고 고운말만 하네 역시 믿음만하다</li> <li>오늘은 비 안 온대요. 우리 모두 기상청을 믿어 보아요~</li> <li>날씨 관련에선 뉴스랑 기상청 보는 게 최고지.</li> <li>기상청 닷 할수 없는 노릇이니까.</li> <li>...기상청 진짜 힘드시겠다...</li> <li>요즘 기상청 컨디션 괜찮은가? 소나기 울거라더니 진짜 날이 흐려지네</li> <li>부산에 영향이 얼마나 가느냐가 문제인가... 내 추측인데 우리나라 기상청 쪽이 맞을듯. 일본 재네 항상 희망적으로 예측하더라.</li> </ul>
부정	<ul style="list-style-type: none"> <li>주말에 날씨가 풀린다고? 음, 기상청은 역시 구라청이었어</li> <li>비 풀막 맞고 퇴근했는데 집에 도착하자마자 비그치는 아름다운 이야기 기상청 xxx</li> <li>헤이 기상청군. 비온다며... 내 우산무게 약 714g 어쩔 거야</li> <li>기상청.. 이 쓸데없는...--</li> <li>이쯤되면 기상청에 내는 세금들이 아까워지고 그래요..</li> <li>예허 기상청 니들을 믿은 나만 xx이지 장마가 끝나긴 개뿔..어제 저녁부터 지금까지 계속 비..</li> <li>장마가 끝났다는 일기예보가 나오기 무섭게 오늘 비올 거 같다 ㅋㅋㅋ 슈퍼컴퓨터로 지뢰찾기하는 기상청 ㅋㅋ</li> <li>밤에 온다더니 아침부터 비오는거 보소~기상청...</li> </ul>

[표 2]는 2011~2013년 트위터들에 대한 감성분류 분석결과이다. 감성분류는 긍정, 부정, 중립 3종류이고, 표에서 ‘분류된 트윗’행은 나이브 베이즈 방법으로 분석된 각 감성별 트윗 개수이다. ‘전체트윗’은 연구자들이 모든 트윗에 대해 직접 분류한 결과이고, 이 ‘전체트윗’

이 감성분석의 분류 정확도를 평가하는 기준이 된다. 나이브 베이즈를 활용한 감성분석 결과 부정 감성은 95.98~97.23%의 정확도로 분류되었고, 중립 감성은 70.27~73.76%의 정확도를 보였다. 부정적 트윗에 분류를 명확히 할 수 있는 비속어 등의 단어들 많아 중립 감성보다 정확하게 분류된 것으로 판단된다. 그리고 긍정 감성은 전체 트윗의 수가 애초에 적어 감성분류를 수행하기에 충분한 감성사전에 제작되지 못했기 때문에 정확도가 낮은 것으로 보인다. 세 감성에 대한 전체 정확도인 ‘연별 정확도’는 74.58~78.01%의 정확도로 나타났다. 기본적인 나이브 베이즈 방법만으로는 연별 정확도가 약 75%에 불과하여 원시 데이터를 제대로 재현하지 못하는 것으로 판단된다.

표 2. 2011~2013년 기본 나이브 베이즈 감성분류 결과.

연도	감성	분류된 트윗	전체 트윗	분류 정확도(%)	연별 정확도(%)
2011	긍정	122	339	35.99	74.58
	부정	2,425	2,515	96.42	
	중립	6,828	9,717	70.27	
2012	긍정	87	264	32.95	75.94
	부정	1,504	1,567	95.98	
	중립	6,778	9,189	73.76	
2013	긍정	47	193	24.35	78.01
	부정	2,385	2,453	97.23	
	중립	4,596	6,363	72.23	

표 3. 2011~2013년 추가사전을 활용한 나이브 베이즈 감성분류 결과.

연도	감성	분류된 트윗	전체 트윗	분류 정확도(%)	연별 정확도(%)
2011	긍정	285	339	84.07	96.75
	부정	2,470	2,515	98.21	
	중립	9,407	9,717	96.81	
2012	긍정	220	264	83.33	97.21
	부정	1,527	1,567	97.45	
	중립	8,965	9,189	97.56	
2013	긍정	144	193	74.61	97.79
	부정	2,391	2,453	97.47	
	중립	6,185	6,363	97.20	

[표 3]은 긍정, 부정, 중립의 각 감성에서만 사용된 형태소로 제작한 추가사전을 적용하여 감성분석한 결과이다. [표 1]의 분석결과에서 샘플수가 적어 제대로 분류하지 못했던 긍정 감성의 분류 정확도가 추가사전 적

용을 통해 74.61~84.07%까지 향상되었고, 부정과 중립 감성의 분류 정확도도 최대 약 98%까지 상승하여 학습 데이터의 감성을 충분히 잘 재현한 것으로 나타났다.

추가사전을 사용한 감성분석이 2011~2013년도의 원 데이터를 충분히 재현하는 것으로 나타나 해당 방법으로 2014년도 트윗에 대해 검증을 실시하였다. 그 결과 극히 적은 긍정 트윗은 여전히 거의 분류하지 못했고, 연별정확도는 74.97%로 나타났다. 정확도가 낮아진 원인은 첫째, 2014년 수집된 트윗의 개수가 3,672개로 검증을 수행하기에 충분하지 않고, 둘째, 2011~2013년 트윗으로 제작한 ‘추가사전’의 단어들이 2014년 트윗을 분류하기에는 적절하지 않았을 수 있다. 점점 감소추세에 있는 트위터 사용자 수와 오직 ‘기상청’을 언급한 트윗만을 수집한 원시 데이터의 한계이기 때문에 향후 다양한 매체의 텍스트를 수집하거나 기상현상 혹은 날씨 예보와 같은 키워드를 포함하는 트윗을 수집함으로써 데이터 부족으로 인한 문제를 해결해야 할 것이다. 그리고 월별 혹은 일별로 업데이트되는 어휘들을 실시함으로써 감성사전과 추가사전에 반영한다면 정확도가 높아질 수 있을 것으로 사료된다.

표 4. 2014년 추가사전을 활용한 나이브 베이스 감성분류 결과.

연도	감성	분류된 트윗	전체 트윗	분류 정확도(%)	연별 정확도(%)
2014	긍정	3	57	5.26	74.97
	부정	629	895	70.28	
	중립	2,121	2,720	77.98	

다음으로 추가사전을 활용한 감성분석 결과가 2012, 2013년의 만족도 변화 추이를 설명할 수 있는지 살펴보았다. [그림 1]에서 어휘의 사전적 의미에 기반한 감성 분석결과는 부정적 감성 변화추이와 기상서비스에 대한 만족도 변화 간의 관계를 설명하기 어려웠다. [그림 5]는 2012, 2013년 트윗에 대한 부정적 감성의 비율과 해당 연도 하반기 만족도 변화를 함께 나타낸 것이다. 2013년에 부정적 감성의 비율이 급증하고, 만족도는 급락한 형태를 잘 보여주고 있어 [그림 1]과 달리 문장의 의미를 기반으로 직접 분류한 감성 결과가 만족도 변화를 더 잘 설명할 가능성이 큰 것으로 판단된다. 물론 사

례가 단 2년으로 적기 때문에 일반화는 어렵지만, 앞서 언급한 바와 같이 사전에 포함되는 어휘가 늘어나고, 트위터 뿐만 아니라 뉴스, 댓글, 블로그 등 다양한 매체의 데이터를 확보하여 분석한다면 분석가의 필요에 따라 수시로 기상서비스 사용자들의 인식을 분석할 수 있게 향후 설문조사를 충분히 보완할 수 있을 것이다.

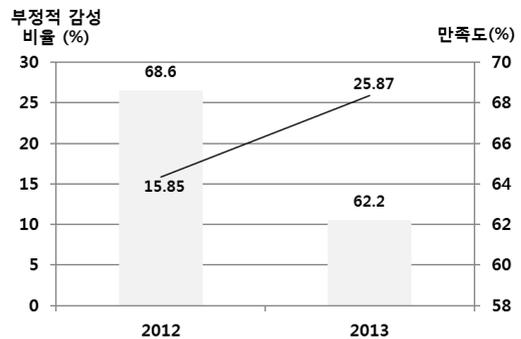


그림 5. 2012, 2013년 추가사전을 사용한 감성분석으로 도출된 부정적 감성 비율(선)과 20대 일반인의 만족도(막대) 변화 추이

## VI. 결론

본 연구는 문장 의미를 고려한 감성분석을 실시하여 만족도 설문조사를 보완할 수 있는지 살펴보고, 트윗에 나타난 감성의 분류 정확도를 향상시킬 수 있는 간단한 추가사전을 제시하였다. 감성분석에 사용된 자료는 2011~2014년 하반기에 기상청이 실시한 만족도 설문 조사의 종료시점으로 부터 과거 100일 동안의 트위터 자료이다.

수집된 트윗을 긍정, 부정, 중립의 감성으로 직접 분류해 본 결과 부정적 어휘를 포함한 긍정문, 혹은 단순히 분류할 수 없는 미세한 의미차이를 포함한 문장 등이 다수 있었다. 해당 트윗들을 연구자가 직접 문장의 의미에 기반하여 감성별로 구분한 뒤 사전을 제작하였고, 나이브 베이스 방법을 적용하여 감성분석을 실시하였다. 학습 데이터인 2011~2013년 트윗의 감성분석 결과 ‘만족도’에 중요한 영향을 미치는 부정적 감성의 분류 정확도는 최고 약 96%, 연별 정확도는 약 76%로

나타났다. 그리고 낮게 나타난 학습데이터의 분류정확도를 높일 수 있는 추가사전을 제시하였다. 추가사전은 긍정, 부정, 중립 문장들에서 다른 감성에는 나타나지 않고 한 가지 감성에서만 출현한 어휘들로 구성되었다. 추가사전을 활용한 감성분석 결과 연별 정확도가 약 97%까지 향상되었다.

학습 데이터를 충분히 재현한 추가사전을 활용하여 2014년 트윗에 대해 검증을 실시하였다. 그 결과 약 75%의 분류 정확도로 나타났는데, 트윗 등 데이터의 추가 수집, 감성/추가사전의 상시 업데이트를 통해 정확도를 향상시키려는 노력이 요구된다.

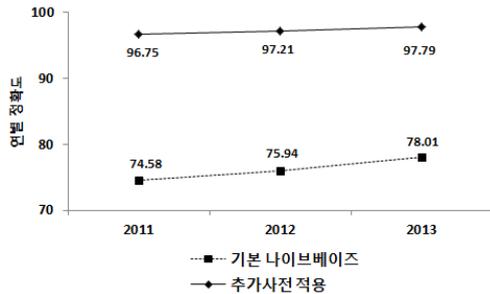


그림 6. 2011~2013년 학습데이터에 대한 기본 나이브 베이즈와 추가사전을 활용한 나이브 베이즈 방법의 연별 감성분류 정확도 비교

추가사전을 통한 방법이 여전히 많은 개선점을 요구하는 것과는 별개로, 문장의 의미에 기반한 본 연구의 감성분석 결과와 설문조사로 얻어진 만족도 변화추이를 살펴보기 위해 2012, 2013년 트윗의 부정적 감성 비율과 만족도를 비교하였다. 그 결과 문장의 의미에 기반한 감성분석 결과가 어휘 기반의 감성분석보다 만족도의 증감을 더 잘 나타낼 수 있는 것으로 나타났다.

본 연구는 빅데이터 분석 및 감성분석 성능향상을 도모한 연구들에 비해서 자료의 수가 적고, 형태소 분석기의 성능을 개선하지 않았으며, 감성사전에 포함될 형태소를 사전의 표제어 즉, 품사 기본형으로 수정하지 않았다. 그리고 감성분석과 만족도 간의 관계 분석에서도 단 2년 동안의 변화만을 비교하여 결과의 일반화에 한계가 있다. 그러나 감성분석의 분류 정확도를 개선할 수 있는 매우 간단한 수정 방법을 제시하였고, 감성분

석 결과와 사용자 만족도를 함께 살펴보면 향후 SNS 데이터를 활용한 감성분석이 설문조사를 보완할 수 있는 가능성을 보인데 의의가 있다. 만족도 설문조사 결과의 효용성에 의문이 제기되는 상황에서 상시적으로 수집될 수 있는 SNS 데이터의 분석 및 활용 가능성은 기상서비스 사용자와의 커뮤니케이션 점점 확대를 위한 시도를 이끌어 내고, 궁극적으로 만족도 제고를 위한 정책 개발에 도움을 줄 수 있을 것이다. 다만, 구체적인 활용을 위해 만족도 변화를 설명할 수 있는 요인 분석을 위해 일별, 월별, 혹은 기상현상 및 기상청 관련 이슈를 세부적으로 분석하여 감성분석 결과와 설문조사 만족도 간의 관계를 통계적으로 증명할 수 있는 추가 연구가 필요할 것이다.

### 참고 문헌

- [1] A. Silver and C. Conrad, "Public perception of and response to severe weather warning in Nova Scotia, Canada," *Meteorological Applications*, Vol.17, pp.173-179, 2010.
- [2] D. Demeritt, S. Nobert, H. Cloke, and F. Pappenberger, "The European Flood Alert System and the communication, perception, and use of ensemble predictions for operational flood risk management," *Hydrological Processes*, Vol.27, pp.147-157, 2013.
- [3] H. Stephanine, B. Rachel, K. Kim, Dr. B. Jerry, and E. Somer, "A Preliminary Look at the Social Perspective of Warn-on-Forecast: Preferred Tornado Warning Lead Time and the general Public's Perceptions of Weather Risks," *weather, climate, and society*, Vol.3, pp.128-140, 2011.
- [4] J. Demuth, J. Lazo, and R. Morss, "Exploring Variations in People's Sources, Uses, and Perceptions of Weather Forecasts," *Weather, Climate, and Society*, Vol.3, pp.177-192, 2011.

- [5] R. Morss, J. Demuth, and J. Lazo, "Communicating Uncertainty in Weather Forecasts: A Survey of the U.S. Public," *Weather and Forecasting*, Vol.23, pp.974-991, 2008.
- [6] S. Joslyn and S. Savelli, "Communicating forecast uncertainty: public perception of weather forecast uncertainty," *Meteorological Applications*, Vol.17, pp.180-195, 2010.
- [7] S. Savelli and S. Joslyn, "Boater Safety: Communicating Weather Forecast Information to High-Stakes End Users," *Weather, Climate, and Society*, Vol.4, pp.7-19, 2012.
- [8] T. Kox, L. Gerhold, and U. Ulbrich, "Perception and use of uncertainty in severe weather warnings by emergency services in Germany," *Atmospheric Research*, Vol.158-159, pp.292-301, 2015.
- [9] 기상청, *2015년도 기상업무 국민 만족도 조사 결과보고서*, 2015.
- [10] 김인겸, 정지훈, 김정윤, 신진호, 김백조, 이기광, "기상예보 정보 사용자 그룹의 만족가치 제고 방안: 강수예보를 중심으로," *한국콘텐츠학회논문지*, 제13권, 제11호, pp.382-395, 2013.
- [11] 기상청, "기상현상과 소셜 데이터 연관성 분석을 위한 기반 연구," 2014.
- [12] S. Kiritchenko, X. Zhu, and S. Mohammad, "Sentiment Analysis of Short Informal Texts," *Journal of Artificial Intelligence Research*, Vol.50, pp.723-762, 2014.
- [13] H. Tang, S. Tan, and X. Cheng, "A survey on sentiment detection of reviews," *Expert Systems with Applications*, Vol.36, pp.10760-10773, 2009.
- [14] 장재영, "온라인 쇼핑몰의 상품평 자동분류를 위한 감성분석 알고리즘," *한국전자거래학회지*, 제14권, 제4호, pp.19-33, 2009.
- [15] 이상훈, 최정, 김종우, "영역별 맞춤형 감성사전 구축을 통한 영화리뷰 감성분석," *지능정보연구*, 제22권, 제2호, pp.97-113, 2016.
- [16] 김도연, 오영, 박혁로, "감성 강도를 고려한 감성 분석 평가집합 구축," *한국콘텐츠학회논문지*, 제12권, 제11호, pp.30-38, 2012.
- [17] 김유신, 김남규, 정승렬, "뉴스와 주가: 빅데이터 감성분석을 통한 지능형 투자 의사결정모형," *지능정보연구*, 제18권, 제2호, pp.143-156, 2012.
- [18] 정지선, 김동성, 김종우, "온라인 언급이 기업 성과에 미치는 영향 분석: 뉴스 감성분석을 통한 기업별 추가 예측," *지능정보연구*, 제21권, 제4호, pp.37-51, 2015.
- [19] A. Go, R. Bhayani, and L. Huang, *Twitter sentiment classification using distant supervision*, Technical report, Stanford, 2009.
- [20] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," *Proceedings of LREC*, 2010.
- [21] B. O'Connor, R. Balasubramanian, B. Routledge, and N. Smith, "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series," in *Proceeding of the Fourth international AAAI Conference on Weblogs and Social Media* Washington, DC, May23-26, pp.122-129, 2010.
- [22] D. Kang and Y. Park, "Review-based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach," *Expert Systems with Applications*, Vol.41, pp.1041-1050, 2014.
- [23] 김장석, 진은미, 이샘, "20대 여성소비자의 화장품 용기디자인 선호도에 관한 연구-기초화장품을 중심으로," *Journal of Integrated Design Research*, 제14권, 제4호, pp.97-106, 2015.
- [24] <http://kddp.net/projects/hannanum/>
- [25] L. L. Dhande and G. K. Patnaik, "Analyzing Sentiment of Movie Review Data using Naive Bayes Neural Classifier," *International Journal*

of Emerging Trends & Technology in Computer Science, Vol.3, pp.313-320, 2014.

[26] H. S. Kang, J. Yoo, and D. Han, "Senti-lexicon and improved Naive-bayes algorithms for sentiment analysis of restaurant reviews," Expert Systems with Applications, Vol.39, pp.6000-6010, 2012.

[27] E. Fersini, E. Messina, and F. A. Pozzi, "Sentiment analysis: Bayesian Ensemble Learning," Decision Support Systems, Vol.68, pp.26-38, 2014.

[28] A. Onan, S. Korukoglu, and H. Bulut, "Ensemble of keyword extraction methods and classifiers in text classification," Expert Systems with Applications, Vol.57, pp.232-247, 2016.

[29] 기상청, "2011년도 기상업무 대국민 만족도 조사," 2011.

저 자 소 개

김 인 겸(In-Gyum Kim) 준회원



- 2006년 2월 : 인제대학교 경영학부(경영학사)
- 2008년 2월 : 인제대학교 경영학과(경영학석사)
- 2010년 1월 ~ 현재 : 국립기상과학원

<관심분야> : 빅데이터, 기상예보가치, SCM

김 혜 민(Hye-Min Kim) 정회원



- 2009년 8월 : 서울과학기술대학교 경영학과(경영학사)
- 2013년 2월 : 서울과학기술대학교 에너지정책학과(경제학석사)
- 2015년 2월 ~ 현재 : 국립기상과학원

<관심분야> : 기상정보 가치평가, 자원·환경경제

임 병 환(Byunghwan Lim) 정회원



- 1998년 2월 : 부경대학교 대기과학과(이학석사)
- 2006년 2월 : 부경대학교 환경대기과학과(이학박사)
- 2000년 6월 ~ 현재 : 국립기상과학원

<관심분야> : 대기대순환, 몬순, 해양기상, 3차원 가시화

이 기 광(Ki-Kwang Lee) 정회원



- 2005년 2월 : KAIST 산업공학과(공학박사)
- 2005년 3월 ~ 2009년 2월 : 인제대학교 경영학부 조교수
- 2009년 3월 ~ 현재 : 단국대학교 경영학부 부교수

<관심분야> : 기상기후경영, SCM, 의사결정론