

영화평과 평점을 이용한 감성 문장 구축을 통한 영화 평점 추론

Movie Rating Inference by Construction of Movie Sentiment Sentence using Movie comments and ratings

오 연 주¹ 채 수 환^{2*}
Yean-Ju Oh Soo-Hoan Chae

요 약

영화 리뷰 사이트에서 영화 평점은 네티즌들의 주관적 판단으로 결정된다. 이로 인해 그들이 남긴 영화평과 평점 사이의 극성이 서로 불일치하는 경우가 종종 발생한다. 본 논문에서는 이 문제를 해결하기 위해 영화의 평가에 영향을 미치는 감성 문장들의 집합을 만들고, 이들을 영화평에 적용하여 평점을 추론한다. 감성 문장들의 집합을 만들기 위한 과정은 감성 어휘 사전을 구축하는 단계와 감성 문장을 구성하는 단계로 이루어진다. 감성 어휘 사전은 영화평에서 쓰인 형용사와 형용사의 극성을 저장한다. 감성 문장은 영화와 관련된 명사를 주어로 갖고 감성 어휘 사전의 어휘를 서술어로 갖는 문장 구조이다. 감성 문장의 극성과 감성 문장에서 쓰인 서술어의 극성이 다른 문장들은 제거하여 감성 문장들이 감성 어휘 사전 어휘의 극성과 일치되도록 하였다. 영화평에서 쓰인 감성 문장들의 평균 점수를 구하면 영화평이 갖는 감성 점수가 된다. 본 연구 결과를 통해 네티즌들이 매긴 평점에 비해 감성 문장 집합을 적용하여 계산한 영화평의 감성 점수가 영화평에 대한 의견을 더 잘 반영한다는 것을 알 수 있다.

☞ 주제어 : 오피니언 마이닝, 감성 사전, 의미 극성, 문장 점수 측정, 평점 유추

ABSTRACT

On movie review sites, movie ratings are determined by netizens' subjective judgement. This means that inconsistency between ratings and opinions from netizens often occurs. To solve this problem, this paper proposes sentiment sentence sets which affect movie evaluation, and apply sets to comments to infer ratings. Creation of sentiment sentence sets is consisted of two stages, construction of sentiment word dictionary and creation of sentiment sentences for sentiment estimation. Sentiment word dictionary contains sentimental words and its polarities included in reviews. Elements of sentiment sentences are combined with movie related noun and predicate from words sentiment word dictionary. In this study, to make correspondence between polarity of sentiment sentence and sentiment word dictionary, sentiment sentences which have different polarity with sentiment word dictionary are removed. The scores of comments are calculated by applying averages of sentiment sentences elements. The result of experiment shows that sentence scores from sentiment sentence sets are closer to reflect real opinion of comments than ratings by netizens'.

☞ keyword : Opinion mining, Sentiment Dictionary, Sentiment Polarity, measuring sentence score, Rating analogy

1. 서 론

최근 다양한 분야에서 빅데이터 활용이 증가하고 있다. 이에 따라 웹이나 SNS에 올린 글들을 이용해 사람들

의 여론이나 의견 등을 분석하는 오피니언 마이닝 연구의 진행도 활기를 띠고 있다. 오피니언 마이닝은 텍스트에 나타난 감성을 측정하고 특정 기준을 통해 분류, 가공하여 새로운 정보를 얻을 수 있도록 하는 분야이다. 얻어진 정보들로 새로운 어플리케이션을 개발할 수 있고 상업적인 이익을 취하는데 사용하기도 한다[1].

오피니언 마이닝은 텍스트의 주관성 분석을 통해 텍스트에 나타난 글쓴이의 감성이 주관적인지 객관적인지를 판단한다. 주관성을 띠는 텍스트는 극성 분석을 통해 텍스트의 긍정/부정으로 나누어 극성 분석을 진행한다[1].

감성 분석 방법은 단어의 상관관계 분석을 통해 문장의 극성을 알아내는 방법, 자연어처리 방식이나 문장 패턴을 이용하는 방법, 감성 어휘 사전을 구축하여 활용하

¹ Dept. of Computer Engineering, Korea Aerospace University, Kyunggi 412-791, Korea.

² Dept. of Computer Engineering, Korea Aerospace University, Kyunggi 412-791, Korea.

* Corresponding author (ksii@ksii.or.kr)

[Received 26 January 2015, Reviewed 28 January 2015, Accepted 6 April 2015]

☆ “본 연구는 미래창조과학부 및 정보통신산업진흥원의 ICT융합고급인력과정지원사업의 연구결과로 수행되었음” (NIPA-2014-H0401-14-1021)

는 방법 등이 있다. 감성 어휘 사전 구축은 주로 단어의 극성을 측정하여 단어와 극성을 저장하는 방법으로 이루어진다. 외국에서는 SentiWordNet과 같은 영어 감성 어휘 사전이 많이 개발되었다[2]. 하지만 한국어 어휘에 관한 감성 사전 개발의 진행은 아직 그 수준이 미흡한 상태이다.

짧은 텍스트를 통한 오피니언 마이닝 연구에는 주로 뉴스 기사의 댓글이나 트위터, 영화평 등이 주로 쓰인다. 특히, 영화 평점 사이트는 특정 대상(영화)에 대한 여러 의견이 집중되어 있기 때문에 영화평을 활용한 연구가 많이 진행되고 있다.

영화평과 평점은 인터넷 사용자들의 주관성이 개입되기 때문에 영화평과 평점이 함께 연구에 쓰일 경우 연구의 정확성에 영향을 미치게 된다. 예를 들면, 영화평의 의미는 7점을 갖지만 평점으로 1점을 주는 경우와 같이 영화평과 평점 사이의 극성과 극성의 정도가 불일치하는 경우가 발생하는 것을 종종 볼 수 있다.

이에 본 연구에서는 영화의 의견과 평가에 관련된 감성 문장들의 집합을 만들고 영화평에 적용하여 영화평에 대응하는 평점을 추론을 하였다. 감성 문장들의 집합은 감성 어휘 사전을 구축하는 단계와 감성 평가를 위한 감성 문장을 만드는 단계로 이루어진다. 감성 어휘 사전은 임의의 영화의 영화평들에서 공통 극성을 갖는 형용사들을 추출하고 다시 영화평에 적용하여 새로운 형용사들 사전에 갱신하는 구조를 갖는다. 감성 문장은 영화에 대한 의견을 가지면서 ‘주어+주격조사+서술어’의 문장 구조를 갖는 문장이다. 주어는 영화와 관련이 많은 명사로서 영화평들의 주격명사에 해당되며, 서술어는 감성 어휘 사전의 형용사로 구성된다. 감성 문장들을 영화평에 적용하면 영화평의 극성과 대응되는 감성 점수를 측정할 수 있다. 이 연구를 통해 영화평에 대응하는 평점 추론을 가능하게 하여 영화 사이트에 나온 영화들의 정확한 순위를 측정할 수 있게 된다. 더불어 영화 평점 분포도를 통해 영화에 대한 인터넷 사용자들의 감성 정보를 알 수 있게 된다.

2. 관련 연구

2.1 기존 연구

2.1.1 감성 분석

감성 분석은 텍스트의 주관성 측정으로 얻어진 의미 방향을 통해 문서의 성격을 알아내고 이를 이용해 새로

운 정보를 얻을 수 있도록 하는 분야이다. 텍스트의 주관성을 예측하기 위해 패턴을 추출하는 알고리즘을 이용하여 텍스트의 성격이 주관적인지 객관적인지 판별하는 연구와 형용사를 분석하여 주관성을 결정하는 의미기반 특징들을 알아내는 연구가 진행되었다[3][4]. 주관성을 지닌 텍스트는 다시 극성 분석을 통해 긍정인지 부정인지를 판별한다. 극성 판별은 PMI 방법으로 간단하게 측정할 수 있으며, WordNet 등을 이용하여 측정하는 연구도 진행되었다[5][6].

2.1.2 감성 사전 구축

텍스트의 극성을 측정하기 위해 직접 감성 사전을 구축하는 연구가 많이 진행되고 있다. 감성 사전은 형용사처럼 극성을 띠는 서술어와 서술어의 극성을 저장한다. 감성 사전은 문서의 텍스트에 적용하여 의미의 주관성을 분석하는데 쓰인다. 동시에 텍스트에는 나오지만 사전에 수록되지 않은 새로운 어휘의 극성을 판별하는 역할을 한다. 새로 판별된 감성 어휘는 사전에 갱신한다. 우리나라에서는 상품의 도메인과 상품 특징 텍스트를 통해 상품평의 극성을 분류하는 사전을 구축하는 연구가 진행되었다[7]. 자연어처리와 어휘의 언어적 특징을 고려하여 만든 The Highlights도 상품평 분석을 위한 감성 사전이다[8]. 최근에는 상품평이나 영화평에 그치지 않고 주가 분석에서도 감성 사전이 활용되고 있다. 주가 관련 어휘와 주가의 등락 관계를 분석하여 사전을 만들고 경제 뉴스에 따른 주가 지수 예측을 가능하도록 하는 연구가 진행되었다[9].

2.1.3 영화 평점 기반 시스템

짧은 텍스트에 대한 오피니언 마이닝 연구로는 영화평이 주로 쓰인다. 영화평을 이용하여 평점을 예측하고 영화 추천 및 검색을 편리하게 하는 연구가 진행되었다. 대표적인 연구로는 영화평에 나온 단어의 상관관계를 분석하여 평점을 추론하는 연구와 영화평의 감성 어휘를 분석하여 영화 검색 시스템을 개발하는 연구가 진행되었다[10][11].

정확한 영화 평점 분석을 위한 연구도 진행되었다. 리뷰 평점의 신뢰도를 측정하기 위해 리뷰 데이터베이스와 So-PMI 기법을 적용한 연구가 진행되었다[12]. 영화 평점 사이트에서 불일치하는 영화 평점과 영화평을 제거하여 정확한 평점 유추를 할 수 있는 연구가 이루어졌고, 언어

학적 패턴을 통해 네티즌들이 매긴 평점과 실제 평점의 의미 방향을 측정하는 실험이 진행되었다[13][14].

2.2 기존 연구의 한계

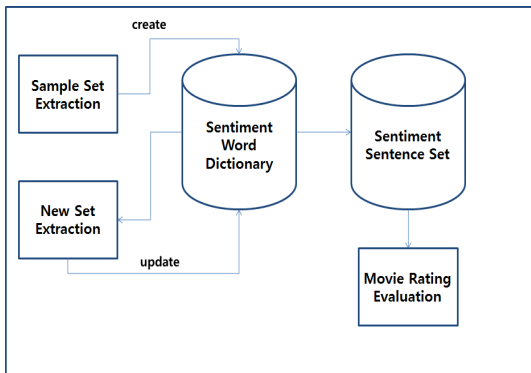
기존의 상품/영화평 극성 분석을 위한 감성 사전 구축 연구들에서는 ‘많다’, ‘작다’와 같이 문맥에 따라 다른 극성을 가질 수 있는 단어에 대해서 실험자가 직접 단어의 극성을 정해주었다. 따라서 단어를 갱신할 때 반자동으로 사전 구축이 이루어졌다.

더불어 기존의 평점 추론 연구들은 영화평의 의미 방향만을 측정하는 실험으로 진행되었다. 이로 인해 단순히 ‘좋다’라고 쓰인 문장과 ‘시나리오가 좋다’라는 영화에 관한 의견이 있는 문장이 서로 같은 점수로 평가된다. 이를 해결하기 위해서는 영화에 대한 의견 반영 정도에 따른 극성의 차별화를 두어야 한다.

본 논문에서는 이 두 가지 문제점을 모두 해결한다. 영화평에서 주로 쓰이는 극성을 설정하여 실험자가 관여하지 않고 어휘를 갱신할 수 있는 자동 업데이트형 감성 사전을 구축하였다. 또한 영화와 관련된 명사와 사전의 단어를 통한 감성 문장을 만들어서 문장들의 영화에 대한 의견 반영 정도를 측정할 수 있도록 하였다.

3. 감성 사전 구축 및 평점 평가 시스템

3.1 시스템 개요



(그림 1) 감성 문장 집합을 통한 평점 추론 시스템
(Figure 1) Sentiment Sentence Set System

그림 1은 본 논문에서 제안하는 감성 문장 집합을 통한 평점 추론 시스템이다. Sample Set Extraction 단계에서

는 임의로 선정된 영화평에서 공통된 의미 방향을 지니는 감성 어휘들을 샘플로 추출한다. 추출한 샘플들은 Sentiment Word Dictionary에 처음으로 수록된다. New Set Extraction 단계에서는 수록된 어휘들을 이용해 새로운 단어들의 극성을 구하고 사전에 수록한다. 이 과정을 반복하여 사전 업데이트를 완료하면 사전 어휘들을 바탕으로 Sentiment Sentence Set을 만든다. 이 Set을 통해 영화평의 문장 점수를 측정하고 네티즌들이 부여한 평점과 비교한다.

3.2 감성 사전 자동 구축

3.2.1 샘플 데이터 집합 만들기

사전 초기에 수록할 샘플 데이터 집합을 만든다. 샘플 데이터를 수집하기 위한 영화 세 편을 골라 각 영화에서 감성 관련 서술어와 서술어의 극성을 구한다. 서술어는 형용사 중에서 극성을 띠는 감정 형용사와 평가 형용사로 지정한다.

각 영화에서 서술어의 극성을 구하기 위해서는 의미 방향의 기준 점수를 정하고, 의미방향을 구하려는 서술어가 기준 점수보다 작으면 부정, 크면 긍정을 부여한다 [7].

$$polarity_x(A,x) = \frac{f(A,x,p)}{review(A,p)} - \frac{f(A,x,n)}{review(A,n)} \quad (1)$$

식 (1)은 영화 A에서 감성 어휘 x의($polarity_x(A,x)$)을 구한 것으로 영화 A에서 전체 긍정 리뷰 개수 ($review(A,p)$) 대비 영화 A에서 감성 어휘 x가 존재하는 부정 리뷰 수($f(A,x,n)$)의 비율의 차이를 구한 것이다. 식(1)의 값이 기준으로 정한 극성 값보다 크면 긍정을 의미하는 1을, 작으면 부정을 의미하는 -1로 지정한다. 각 영화에서 감성 어휘와 극성을 계산하고 세 편의 영화에서 모두 같은 극성을 띠는 어휘들을 고른다. 선별된 어휘 중 어휘가 존재하는 리뷰의 수가 가장 높은 순서대로 30개의 어휘들을 샘플 데이터로 정한다.

표 1은 샘플 데이터 집합의 예를 나타낸 것이다. 샘플 데이터들은 영화에 대한 감정과 평가에 대한 극성을 지닌 감각 형용사와 평가 형용사로 구성된다. 일반적으로 ‘슬프다’라는 서술어는 부정적인 의미로 쓰이지만 영화 평에서는 영화에 대한 긍정 감정 의미로 사용된다는 것을 알 수 있다.

(표 1) 샘플 데이터 집합의 예
(Table 1) Examples of sample data sets

감성 어휘	의미 방향
슬프다	1
재미있다	1
형편없다	-1
똑같다	-1

3.2.2 사전 갱신하기

샘플 데이터 집합을 사전의 초기 어휘들로 설정한다. 초기 상태의 사전을 서술어가 두 개 이상 쓰인 영화평에 적용하여 새로운 감성 어휘의 극성을 구한다. 이 때, 하나의 서술어는 사전에 수록된 단어이어야 한다. 사전을 갱신하는 방법은 다음과 같다.

(표 2) 영화평에서 새로운 감성 어휘의 극성 추론
(Table 2) Inference polarity of new sentimental words in movie opinions

review	재미있었지만 약간 <u>아쉽다</u> .		
word	재미있다	하지만	아쉽다
tag	A	B	B
score	1(set)	-1(new)	

tag는 서술어들의 의미 방향을 나타낸다. 같은 tag 값을 갖는 서술어는 같은 극성을 갖는다. tag는 ‘지만’, ‘나’와 같은 부정 연결 어미나 ‘그러나’, ‘하지만’과 같은 부정 부사에서 의미 방향이 바뀐다. 서술어 중에서 사전에 수록되어 있는 서술어는 그 극성 점수를 부여하며, 그 점수와 tag를 이용해 나머지 서술어의 점수를 추론한다. 사전에 수록되어 있지 않은 형용사들의 tag와 사전에 수록되어 있는 서술어의 tag가 같으면 둘의 점수는 같으며 tag가 다르면 절대값이 서로 반대인 점수를 갖게 된다. 서술어는 점수별로 빈도수를 저장해서 식 (2)와 같이 극성을 구한다.

$$polarity_n(x) = \frac{f(x,p) - f(x,n)}{review(x)} \quad (2)$$

극성($polarity_n(x)$)은 식 (2)에서처럼 사전으로 구한 서술어 x 의 빈도수($review(x)$)에 대한 긍정 점수 빈도($f(x,p)$)와 부정 점수 빈도($f(x,n)$) 차이의 비율로 구해서 극성보다 높으면 긍정, 낮으면 부정을 부여한다. 영

화별로 공통 극성을 갖는 형용사를 사전에 수록한다. 위의 과정을 반복하면서 공통 극성을 갖는 형용사가 모두 나올 때까지 사전을 갱신하여 사전을 구축한다. 갱신을 완료한 최종 사전의 어휘 개수는 실험별로 평균 97개의 형용사를 갖는다.

3.3 문장 점수 평가

3.3.1 감성 문장 만들기

감성 어휘 사전을 이용해 감성 평가에 적용할 감성 문장을 만든다. 영화에 대한 의견과 평가에 관련된 문장이 영화 평가에 영향을 주기 때문에 감성 문장은 ‘(영화)가 (재미있다)’, ‘(작품)이 (좋다)’와 같이 <영화 관련 명사+조사+감성 서술어>로 구성 된다. 영화평의 주격 명사들은 대부분 영화와 관련된 명사들이 많아 주격 명사를 영화 관련 명사로 추출한다[7]. 감성 서술어는 3.2절에서 만든 감성 어휘 사전의 서술어들로 구성된다.

각 감성 문장의 단어가 존재하는 영화평들의 평점 평균이 감성 문장의 점수가 되며, 이로써 주격 명사에 따라 감성 어휘 사전의 단어들의 극성 정도가 세분화 된다. 예를 들어, ‘(연기)가 (좋다)’와 ‘(캐스팅)이 (좋다)’에서 ‘좋다’는 긍정 극성인 1을 갖지만 영화평에서 비중은 7.25와 5.00으로 서로 다른 감성 점수를 갖는다.

또한 문장 점수 의미 방향과 감성 어휘 사전에서의 서술어 극성이 다른 문장 구조를 제거시킨다. ‘(현실감)이 (없다)’라는 문장은 ‘없다’가 서술어 사전에서 부정 극성을 갖지만 문장 점수는 7.5점으로 영화평에서는 주로 긍정적으로 쓰여 사전과 문장 점수가 서로 다른 극성을 갖기 때문에 이러한 문장은 감성 문장에서 제외한다.

3.3.2 문장 점수 측정하기

감성 문장들을 이용하여 영화평의 문장 점수를 측정한다. 아래 식 (3)에서처럼 영화평에 존재하는 감성 문장들의 점수를 통해 영화평에 대응하는 문장 점수를 구한다.

$$rating_{sentiment} = TRUNC\left(\sum_{k} score_{sentiment}\right) \quad (3)$$

식 (3)에서 구한 $rating_{sentiment}$ 은 각 영화평으로 구한 문장 점수가 된다. “영화가 재미있기는 한데 큰 감동은 없는 듯...”이라는 영화평은 네티즌이 평점으로 8점을 부여했다. 이 영화평은 ‘영화가 재미있다’와 ‘감동은 없다’라는 감성 문장이 혼합되어 구성된 문장이다. 각 문장

구조의 점수($score_{basic}$)는 9.13과 -1이며, 영화평의 평점은 이들의 평균(소수점 이하 버림)인 4점이 된다. 이처럼 감성 문장 구조로 추론한 영화 평점과 네티즌들이 임의로 매긴 평점을 비교하여 얼마나 다른지 확인하는 작업을 진행한다.

4. 실험 결과 및 평가

4.1 실험 환경 설정

4.1.1 감성 어휘 사전 구축 실험

본 연구의 실험을 위해 “네이버 영화”에서 임의의 영화 세 편의 영화평과 평점을 수집하였다. 감성 어휘 사전과 감성 문장을 만들기 위한 문장 성분 분석은 서울대학교에서 제공하는 ‘꼬뜨파 형태소 분석기’를 사용하였다.

한 영화에서 1점부터 10점까지 서술어를 갖는 영화평 100개 이상을 고르며, 고른 영화평의 개수가 10,000개 이하가 되도록 영화평과 평점을 추출한다. 추출한 세 편의 영화에서 공통 극성을 갖는 사전을 사전에 수록하며, 이를 영화에 다시 적용해 사전을 갱신한다.

(표 3) 감성 어휘 사전의 극성 판별 정확도
(Table 3) Sentiment polarity identification accuracy of Sentiment Word Dictionary

	S_1	S_2	S_3	pre- cision	recall	F-score
1	1027	6684	9998	69.64	52.70	59.99
2	1031	7330	9877	73.68	60.89	66.67
3	6844	8360	10000	74.07	57.14	64.51
4	6862	9422	10000	81.15	65.88	72.72

표 3은 실험별로 세 영화 영화평의 (S_1, S_2, S_3) 서술어 개수에 따른 완성된 사전 어휘들의 극성 판별에 대한 정확도를 나타낸 것이다. 표3의 precision, recall, F-score은 아래의 식(4)와 같이 계산된다.

$$Precision = \frac{classify(r,r)}{classify(r,r) + classify(r,w)} \quad (4)$$

$$Recall = \frac{classify(r,r)}{classify(r,r) + classify(w,r)}$$

$$F-score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

식 (4)에서 $classify(r,r)$ 은 시스템이 특정 극성에 해당한다고 인식한 단어에 대해 맞게 인식한 서술어의 개수, $classify(r,w)$ 은 시스템이 특정 극성에 해당한다고 인식한 단어에 대해 잘못 인식한 서술어의 개수, $classify(w,r)$ 은 시스템이 특정 극성에 해당하지 않는다고 인식한 단어에 대해 맞게 인식한 서술어의 개수를 나타낸 것이다.

서술어를 갖는 영화평이 많은 영화를 사용하여 감성 사전을 만들수록 F-score이 높아진다. 이는 서술어가 많을수록 서술어에 대한 극성이 분명해지고 공통된 서술어의 개수가 많아져 사전의 성능이 높아지는 것을 의미한다. 또한 precision이 recall보다 높게 나왔는데, 이는 사전을 자동으로 구축할 때 영화마다 다른 극성을 갖는 서술어에 대한 처리의 제약 때문인 것으로 판단된다.

4.1.2 감성 문장 집합 구축 실험

감성 어휘 사전을 만드는데 사용된 영화평들과 감성 어휘 사전으로 감성 문장들을 만든다. 출현 빈도수가 3회 이하인 문장 구조는 제거하며 감성 어휘 사전의 서술어 극성과 다른 극성을 갖는 문장 구조는 제거한다.

(표 4) 감성 문장 집합의 극성 판별 정확도
(Table 4) Sentiment polarity identification accuracy of Basic-Sentence-Form Sets

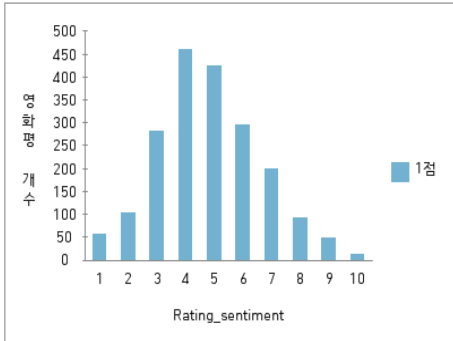
	감성 문장 개수	정확도
1	735	64.35
2	714	66.48
3	751	67.08
4	840	67.14

표 4는 실험별로 완성된 감성 문장들의 극성 판별에 대한 정확도를 나타낸 것이다. 감성 문장의 개수는 감성 어휘 사전에 수록된 어휘가 많을수록 많아진다. 하지만 극성 판별에 대한 감성 문장 개수의 영향은 근소한 차이를 보였다. 이는 주격명사가 자체의 극성을 가져 서술어와 결합되면서 전체문장의 극성이 서술어의 극성과 달라지기 때문인 것으로 판단된다.

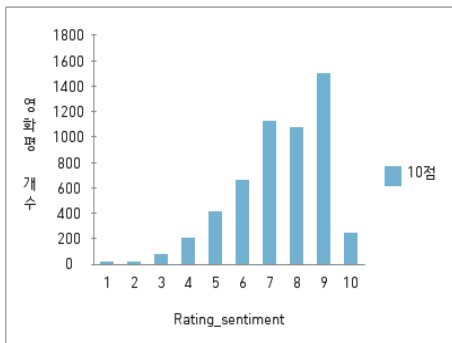
완성된 문장들을 새로운 영화평에 적용시켜 영화의 문장 점수를 구한다. 영화는 31,021개의 영화평을 갖고 있으며 1점부터 10점까지 각각 1%이상 분포한다.

4.2 실험 결과 및 평가

4.2.1 실험 결과



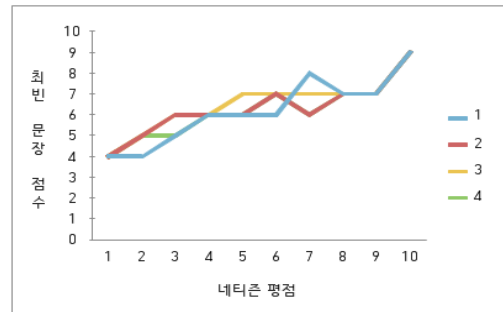
(a) 네티즌들이 매긴 영화평점이 1점일 때 $rating_{sentiment}$ 점수 분포도



(b) 네티즌들이 매긴 영화평점이 10점일 때 $rating_{sentiment}$ 점수 분포도

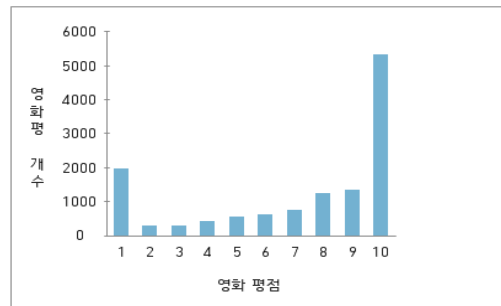
(그림 2) 네티즌 평점별 영화평 대응 문장 점수 분포도의 예
(Figure 2) Example of Distribution of movie comment corresponding sentence score per netizen rating

그림 2의 (a)와 (b)는 첫 번째 실험에서 네티즌들이 매긴 영화평점 1점과 10점에 대한 영화평 대응 문장 점수의 분포도를 나타낸 것이다. 전체 영화평 개수에 대한 영화평 대응 문장 점수와 영화 평점이 일치하는 영화평 개수의 비율은 각각 2.84%(=56/1969)와 4.50%(=240/5326)으로 나타나 네티즌들이 영화평과 평점을 서로 대응하지 않게 부여하는 사실을 알 수 있다. 네티즌들은 1점과 10점을 부여했지만 실제로는 영화평들이 대체로 4점에서 9점의 문장 점수를 갖는다는 것을 알 수 있다.

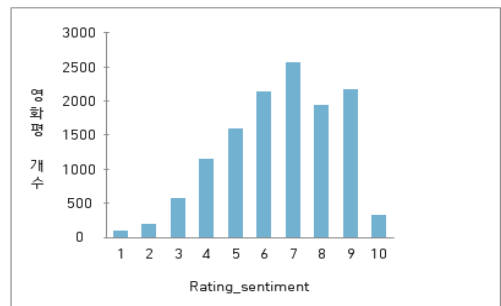


(그림 3) 네티즌 평점별 문장 점수 최빈값
(Figure 3) Mode of movie comment corresponding sentence score rating per netizen rating

그림 3은 실험별로 나타난 네티즌들이 매긴 평점별 영화평 대응 문장 점수의 최빈값을 나타낸 것이다. 온라인 사용자가 매긴 평점 1점은 4점의 문장 점수를 띠며 평점 10점은 9점에 가장 많이 분포한다. 평점 2점에서 9점은 4점에서 8점 사이의 문장 점수를 갖는 것을 확인할 수 있다.



(a) 네티즌들이 매긴 영화 평점



(b) 실험으로 추론한 영화평 대응 문장 점수

(그림 4) 온라인 영화평점과 실험 문장 점수의 차이 비교
(Figure 4) Comparison of online rating and experiment sentence score

그림 4는 12,707개의 영화평을 갖는 임의의 한 영화에 대해 네티즌들이 매긴 영화 평점과 실험으로 추론한 문장 점수를 비교한 것이다. (a)에서 네티즌들이 영화에 대해 1점과 10점을 각각 1,969, 5,326개를 부여해 전체 평점 중 57.4%이상을 차지한다. 하지만 (b)에서 실험으로 추론한 영화평 대응 문장 점수에서는 영화평들이 각 점수대 별로 골고루 분포하는 양상을 띠는 것을 확인할 수 있다.

4.1.2 실험 평가

기존의 영화평을 분석하여 추론한 문장 점수와 네티즌들의 영화 평점을 비교한 연구는 단순히 문장 점수가 기존 평점보다 높으면 긍정, 낮으면 부정을 부여하는 실험이다[14]. 이 실험은 문장의 자체 점수보다는 의미 방향을 구하는데 초점을 맞추었다. 그러나 본 연구는 영화평의 의미방향을 구하는데서 그치지 않고 영화평에 대응하는 점수까지 구하는 실험으로 확장하였다. 또한 기존의 실험은 복잡한 언어학적 지식을 알아야만 문장의 극성을 측정할 수 있었지만 본 연구는 단어 수준에서 문장 점수를 측정할 수 있다.

5. 결론 및 향후 연구

본 연구에서는 영화평에 대응하는 문장 점수를 추론하기 위해 감성 문장 집합을 구축하는 실험을 진행하였다. 감성 어휘 사전은 임의로 선정한 영화들의 공통된 극성을 갖는 서술어를 초기 샘플 데이터로 설정하고, 샘플 데이터를 통해 새로운 서술어들과 서술어의 극성을 갱신하도록 하였다. 감성 어휘 사전은 영화들의 공통 서술어가 많을수록 어휘의 극성 측정의 정확도가 높아진다. 영화의 의견이나 감성 평가를 위한 감성 문장은 영화와 관련된 명사로 이루어진 영화평들의 주격명사를 주어로 갖고 감성 어휘 사전의 형용사들을 서술어로 갖는 구조로써 영화평점에 영향을 주는 문장이다. 이들은 감성 어휘의 개수와 상관없이 어휘의 극성들이 비슷한 정확도를 갖는다. 실험 결과, 영화평의 문장 점수는 6점에서 7점 사이에 분포해 대부분 1점과 10점을 매기는 네티즌들의 평점과는 상당한 차이가 있음을 보여준다. 본 연구를 통해 온라인 영화 사이트의 영화평에 대응하는 평점 측정과 평점을 통한 영화 순위 측정을 할 수 있는 것을 알 수 있다.

하지만 본 연구의 감성 문장이 한정된 품사로만 구성되어있어서 다음과 한계점이 발생한다. 부사구와 같은 수식어의 영향을 고려하지 못하여 ‘영화가 재미있다’와 ‘영화가 매우 재미있다’가 모두 같은 문장 점수를 가지며, 감

성 극성을 갖는 동사나 명사가 포함된 영화평을 측정하지 않아 영화에 대한 평가가 완전히 이루어지지 못하였다. 따라서 보다 정확한 영화의 평점과 순위를 측정하기 위해서는 문장에서의 수식어의 영향력과 형용사 이외의 품사에 대한 극성 연구가 이루어져야 한다. 이로써 온라인 영화 사이트에서의 정확한 영화들의 평점과 순위를 측정하여 영화에 대한 인터넷 사용자들의 정확한 의견 반영을 가능하게 할 것이다.

앞으로의 연구에서는 중립 의미를 갖는 명사와 동사들을 처리하는 연구를 통해 감성 서술어의 범위를 확장할 것이다. 또한 기본 감성 문장 구조에 수식어를 추가하여 수식어구의 문장 점수에서의 영향력을 측정할 것이다. 더 나아가 이모티콘 사용 패턴과 문법에 어긋나는 문장 구조들의 패턴 규칙을 알아내 이들이 영화 평점에 미치는 영향을 알아보는 연구도 진행할 예정이다.

참 고 문 헌 (Reference)

- [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, January 2008. <http://dx.doi.org/10.1561/1500000011>
- [2] A. Esuli and F. Sebastiani, "Sentiwordnet: A publicly available lexical resource for opinion mining" In In Proceedings of the 5th Conference on Language Resources and Evaluation (LREC06), pp. 417-422, 2006. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.61.7217>
- [3] V. Hatzivassiloglou and J. Janyce M, Wiebe "Effects of adjective orientation and gradability on sentence subjectivity" COLING '00 Proceedings of the 18th conference on Computational linguistics, Vol. 1, pp. 299-305, 2000. <http://dx.doi.org/10.3115/990820.990864>
- [4] J. Wiebe and E. Riloff "Creating Subjective and Objective Sentence Classifiers from Unannotated Texts" 6th International Conference, CILing 2005, Vol. 3406, pp. 486-497, February 13-19 2005. http://link.springer.com/chapter/10.1007%2F978-3-540-30586-6_53
- [5] Peter D. Turney "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews" ACL '02 Proceedings of the 40th Annual

- Meeting on Association for Computational Linguistics, pp. 417-424, 2002.
<http://dx.doi.org/10.3115/1073083.1073153>
- [6] Sang-il Song, Dongjoo Lee, Sang-goo Lee, "Identifying Sentiment Polarity of Korean Vocabulary Using PMI" KCC 2010 , Vol.37, No.1 , pp. 260-265, 2010.6.
<http://www.dbpia.co.kr/Article/1308678>
- [7] Jongseok Song, Soowon 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, 2011.3.
<http://www.dbpia.co.kr/Article/1455474>
- [8] Jaeseok Myung, Dongjoo Lee, 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.
<http://www.dbpia.co.kr/Article/838396>
- [9] Eunji Yu, Yoosin Kim, Namgyu Kim, SeungRyul Jeong "Predicting the Direction of the Stock Index by Using a Domain-Specific Sentiment Dictionary" Journal of Intelligence and Information Systems, Vol.19, No.1, pp. 95-110, 2013.
<http://www.dbpia.co.kr/Article/3138319>
- [10] Dumim Yoon, KyungJoong Kim "Prediction of Rating Score from Short Comments on Movies using Word-Rating Correlation Analysis" HCI2011, pp. 484-486, 2011.1.
<http://www.dbpia.co.kr/Article/2548496>
- [11] Sung-Ho Oh, Shin-Jae Kang "Movie Retrieval System by Analyzing Sentimental Keyword from User's Movie Reviews" Journal of the Korea Academia-Industrial cooperation Society, Vol.14, No.3, pp. 1422-1427, 2013.
<http://dx.doi.org/10.5762/KAIS.2013.14.3.1422>
- [12] Jung-Hwa Lee, Ki-Young Lee "The Reliability Measure of Movie Review Ratings" Proceedings of the IEEK Conference, pp. 715-717, 2012.6.
<http://www.dbpia.co.kr/Article/3280417>
- [13] KyongMin Kim, MuHyok Ahn, Younho Lee "Detection of Malicious Rate Evaluation and Prediction of True Rate in Movie Rating" 정, Journal of the Korean Institute of Information Scientists and Engineers: Computing Practices and Letters, Vol.20, No.4, pp. 213-218, 2014.4.
<http://www.dbpia.co.kr/Article/3437599>
- [14] Jee Sun Nam "Study on Linguistic Patterns of Online Reviews on Movie for the Automatic Classification of Human Opinion" The Linguistic Society of Korea, Vol.58, pp. 75-103, 2010.12.
<http://www.dbpia.co.kr/Article/3278672>

● 저 자 소 개 ●



오 연 주 (Yean-Ju Oh)

2014년 한국항공대학교 컴퓨터학과(공학사)
 2014년 ~현재 한국항공대학교 대학원 컴퓨터학과 석사과정
 관심분야 : 오피니언 마이닝, 데이터 마이닝, 자연어 처리
 E-mail : gclass00@kau.ac.kr



채 수 환 (Soo-Hoan Chae)

1973년 한국항공대학교 전자공학과(공학사)
 1985년 미국 Univ. of Alabama 대학원 전산학과(공학석사)
 1988년 미국 Univ. of Alabama 대학원 전기공학과(공학박사)
 1988년~현재 한국항공대학교 전자 및 정보통신공학부 교수
 관심분야 : 분산/병렬처리 시스템, 컴퓨터보안, 데이터마이닝
 E-mail : chae@kau.ac.kr