www.calsec.or.kr
http://dx.doi.org/10.7838/jsebs.2013.18.4.315

편향된 의견 문서 검출을 위한 이상치 탐지 기법

Outlier Detection Techniques for Biased Opinion Discovery

연종흠(Jongheum Yeon)*, 심준호(Junho Shim)**, 이상구(Sang goo Lee)***

초 록

소셜 미디어에서는 상품평, 영화평 등의 다양한 종류의 의견이 표현되고 있으며, 사용자들이 물품 구매 등에 있어 이러한 의견을 참고로 하여 결정을 내리는 것은 일반적이 되었다. 하지만 의견 정보의 활용도가 높아질수록 이를 부적절하게 왜곡하는 사례 또한 증가하고 있다. 예를 들어, 홍보를 목적으로 과도하게 긍정적인 의견이 포함된 리뷰를 작성하거나, 반대로 일반적인 평가에서 벗어나 과도하게 부정적인 의견을 게시하는 경우 등이다. 편향된 의견은 소셜 미디어의 신뢰성과 연결 되기 때문에 이를 검출하는 것은 점차 중요한 문제로 대두되고 있다. 기존의 오피니언 마이닝 혹은 감성 분석은 문서를 분석하여 그 문서가 가지고 있는 의견의 성향을 판단하는 기법이다. 하지만 기존의 연구는 의견을 단순히 긍정/부정으로만 분류하는 방향으로 연구가 이루어져 왔으며, 특히 사전에 의견 성향에 따라 분류된 충분한 양의 학습 데이터가 필요하다는 단점이 있다. 본 논문에서는 학습데이터가 없는 경우에, 전체 문서의 의견 성향 분포에서 벗어난 의견 문서를 검출하는 기법을 제안한다. 여기에는 각도기반 이상치 탐지와, 개인화된 페이지랭크 방법을 활용한다. 또한 영화 리뷰 문서를 대상으로 실험을 수행하여 제안한 방법들의 성능을 분석하였다.

ABSTRACT

Users in social media post various types of opinions such as product reviews and movie reviews. It is a common trend that customers get assistance from the opinions in making their decisions. However, as opinion usage grows, distorted feedbacks also have increased. For example, exaggerated positive opinions are posted for promoting target products. So are negative opinions which are far from common evaluations. Finding these biased opinions becomes important to keep social media reliable. Techniques of opinion mining (or sentiment analysis) have been developed to determine sentiment polarity of opinionated documents. These techniques can be utilized for finding the biased opinions. However, the previous techniques have some drawback. They categorize the text into only positive and negative, and they also need a large amount of training data to build the classifier. In this paper, we propose methods for discovering the biased opinions which are skewed from the overall common opinions. The methods are based on angle based outlier detection and personalized PageRank, which can be applied without training data. We analyze the performance of the proposed techniques by presenting experimental results on a movie review dataset.

키워드 : 이상치 탐지, 오피니언 마이닝, 감성 분석, 개인화된 페이지랭크
Outlier Detection, Opinion Mining, Sentiment Analysis, Personalized Page Rank

---

본 연구는 숙명여자대학교 교내연구비 지원에 의해 수행되었음(과제번호 1-1303-0136).
본 논문은 한국전자거래학회 2013 춘계학술대회에서 발표된 “편향된 의견 문서 검출을 위한 열쇠 : 다차원 접근” 논문이 우수 논문으로 추천되어 확장 수정된 것임.

* First Author, School of Computer Science and Engineering, Seoul National University (jonghm@europa.snu.ac.kr)
** Corresponding Author, Department of Computer Science, Sookmyung Women’s University (jshim@sookmyung.ac.kr)
*** School of Computer Science and Engineering, Seoul National University(sglee@europa.snu.ac.kr)
2013년 10월 29일 접수, 2013년 11월 18일 심사완료 후 2013년 11월 22일 게재확정.

1. 서 론

소셜 미디어에서는 상품평, 영화평 등과 같이 다양한 종류의 의견이 표현되고 있으며 그 양 또한 기하급수적으로 증가하고 있다. 사용자들이 물품 구매 등에 있어 이러한 의견을 참고로 하여 결정을 내리는 것은 일반적이 현상이 되었다. 이러한 의견은 크게 두 부류로, 긍정적인 의견과 부정적인 의견으로 나눌 수 있다. 긍정적인 의견의 경우 제품이나 서비스의 이익에 실제적으로 도움이 되지만, 반대로 부정적인 의견은 제품의 구매 동기를 낮추는 요인 중 하나로서 작용하는 것이 밝혀져 있다[1]. 그렇기 때문에 기업들은 자사의 제품이나 기업 이미지에 대한 소셜 미디어 상의 평판을 관리하는데 많은 노력을 기울이고 있다. 구체적으로 자사의 제품에 대한 긍정적인 의견을 마케팅에 활용하거나, 부정적인 의견에 대해서 제품 개선 등의 형태로 대응하는 것이 그 예이다.

하지만 사용자들의 의견 정보에 대한 활용도가 높아질수록 이를 부적절하게 왜곡하는 사례 또한 증가하고 있다. 예를 들어, 특정 제품, 서비스, 단체, 인물에 대해서 홍보를 목적으로 과도하게 긍정적인 의견을 표현 하거나, 반대로 폄하하기 위한 목적으로써 일반적인 평가에서 벗어나 과도하게 근거 없는 부정적인 의견을 게시하는 경우가 있다[2]. 이러한 의견들은 일반적인 의견과는 다르게 본래 글을 작성한 사람의 의도를 숨긴 채 표현되기 때문에 편향된 의견(biased opinion)이라고 할 수 있다. 왜곡된 의견을 찾아내는 것은 소셜 미디어의 신뢰성을 확보하는데 있어 필요한 작업으로, 소셜 미디어의 분석 등에 있어 점차 중요한 프로세스가 되어가고 있다.

웹문서, 이메일을 대상으로 스팸을 검출하는 기법은 많은 연구가 이루어졌으며, 실용적인 수준에서 활용되고 있다. 하지만 일반적인 스팸 문서와 편향된 의견 문서는 지니고 있는 특성에 차이가 있어 기존의 기법들을 직접 적용하기에 적절하지 않다[3, 4]. 기존의 스팸 검출과 편향된 의견 검출의 가장 주요한 차이점으로는 내용을 사람이 직접 읽는 것으로 검출하는 것이 거의 불가능에 가깝다는 것이다. 예를 들어, 한 사용자가 특정 서비스에 대해서 어떤 커뮤니티에서 좋은 평가를 내리는 리뷰를 작성 하였을 때, 그 커뮤니티 대다수의 의견이 나쁜 평가를 이루고 있다면 이는 편향된 의견으로 판단할 수 있다. 하지만 만약 다른 커뮤니티에서는 좋은 평가와 나쁜 평가가 비슷한 비율을 이루고 있다면 이 의견은 편향되어 있다고 판단하기 어렵다.

오피니언 마이닝(Opinion Mining) 혹은 감성 분석(Sentiment Analysis)과 관련된 연구[5, 6]들은 문서를 분석하여 그 문서가 가지고 있는 의견의 성향을 판단하는 기법이다. 의견 문서의 검출과 분류에 이 방법은 널리 사용되고 있다. 하지만 이 기법들은 문서를 단순히 긍정/부정으로 분류하는데 집중하여 연구가 이루어져 왔기 때문에, 그 보다 다양한 종류의 의견에는 대응하기 어렵다. 또한, 분류를 위해 사전에 의견 성향에 따라 분류된 충분한 양의 학습 데이터가 필요하다는 단점이 있다.

본 논문은 문서 집합의 의견 성향 분포를 고려하여 편향된 소수의 의견 문서를 검출하는 기법을 제안한다. 편향된 의견 문서는 대다수의 일반적인 의견의 문서에 비해 적은 숫자를 이루고 있다. 이때, 문서에서 추출한 키

워드 빈도수와 같은 통계치는 다수의 의견 문서에 비해 소수 의견의 문서가 다른 분포를 보인다. 따라서 편향된 의견 문서를 일종의 이상치(outlier)라 보고, 문서의 통계치를 활용한 이상치 탐지 기법을 적용한다.

특히, 일반적으로 대다수의 의견 문서들은 의견 성향에 따라 분류가 되어 있지 않기 때문에, 학습 데이터가 필요 없는 랭킹 기반의 비감독(Unsupervised) 이상치 탐지 기법을 적용하는 것이 필요하다. 본 논문에서 두 가지 랭킹 기법을 사용한다. 1) 이상치 탐지 기법 중 하나인 각도 기반 이상치 탐지(Angle Based Outlier Detection) 기법과 2) 문서 키워드 그래프를 활용한 개인화된 페이지 랭크(Personalized Page Rank) 기법을 활용한다. 또한 영화 리뷰 문서를 대상으로 실험을 수행하여 제안 방법을 검증한다.

본 논문은 다음과 같이 구성되어 있다. 제 2장은 감성 분석 분야에서 의견 스팸을 검출하는 기법을 관련 연구로서 살펴본다. 제 3장은 문서 모델링 방법과 편향 문서의 검출 기법을 설명하며, 제 4장에서 이를 검증하기 위한 실험 결과를 기술하였다. 제 5장은 논의 및 향후 과제에 대해 기술한다.

2. 관련 연구

오피니언 마이닝 혹은 감성 분석은 문서 내에서 나타나는 의견 텍스트를 추출하고, 텍스트가 긍정인지 부정인지 판별하는 극성 판단을 기본 프로세스로 한다. 이렇게 추출된 의견은 요약이나 시각화 절차를 통해 사용자에게 전달된다[15]. 의견 텍스트 추출에는 크게 문장 구조 정보를 활용하는 자연언어처리 기반의 방법[7, 8]들과 어휘 빈도수, TF IDF 등을 활용하는 통계적 기반 방법[9, 10]이 있다. 극성 판단에는 사전에 정의된 사전을 활용하는 방법[7, 8]이나 기계학습에 기반한 분류기(Classifer)[11]를 활용하는 방법이 이용되고 있다. 이외에 감성 분석을 원할하게 하기 위한 센티워드넷(SentiWordNet)[12] 등의 언어 자원도 구축되어 활용되고 있다.

이러한 의견 문서를 다루는 연구들에서, 사용자들의 일반적인 평가가 아닌, 어느 방향으로 사용자들의 의견을 유인하기 위해 쓰여지는 리뷰를 의견 스팸(Opinion Spam) 또는 가짜 리뷰(Fake Review) 등으로 부르며 이를 탐지하고자 하는 방법이 제안되어 왔다. [2]에서는 이와 같은 개념이 처음으로 제시되었으며, 상품 리뷰를 중심으로 의견 스팸이 웹이나 이메일 스팸과 다른 점을 밝히고, 검출하기 위해 기계 학습에 기반한 방법을 제안하였다.

[13]에서는 상품 리뷰를 중심으로 스팸 리뷰를 작성하는 사용자의 행동 특성을 활용하여 스팸 작성자를 검출하는 방법을 제시하였다. 스팸 작성자들의 행동은 크게 두 가지 유형으로, 상품 점수에 영향을 주기 위해 일반적인 사용자들과는 다른 양상으로 상품 점수를 매기는 행위와, 리뷰 텍스트에 영향을 미치기 위해 여러 번 리뷰를 작성하는 행위로 나뉘어진다. 유사도에 기반한 점수화로 스팸 지수를 정의하였으며, 이를 머신 러닝 기법에 입력 특징으로 사용하였다.

[14]에서는 개별적인 의견 스팸이나 그 작성자를 찾아내는 것에서 더 나아가, 집단적으로 가짜 리뷰를 작성하는 집단을 찾아내는 방법을 제안하였다. 빈발 항목 집합 마이닝(Frequent

Itemset Mining)을 통해 후보 집단을 찾아낸 후, 각 집단의 행위, 집단 간의 관계, 개별 리뷰 작성자, 리뷰 대상 상품 간의 관계를 활용한다.

하지만 이 기법들을 활용하기 위해서는 의견 성향에 따라 분류된 충분한 양의 학습 데이터가 필요하다는 단점이 있다. 특히, 이러한 데이터 셋을 만드는 것은 매우 어려운 작업에 해당한다. 예를 들어, 한 사용자가 특정 서비스에 대해서 어떤 커뮤니티에서 좋은 평가를 내리는 리뷰를 작성 하였을 때, 그 커뮤니티 대다수의 의견이 나쁜 평가를 이루고 있다면 이는 편향된 의견으로 판단할 수 있다. 하지만 만약 다른 커뮤니티에서는 좋은 평가와 나쁜 평가가 비슷한 비율을 이루고 있다면 이 의견은 편향되어 있다고 판단하기 어렵다. 즉, 같은 텍스트가 경우에 따라서는 편향되거나 편향되지 않을 수 있기 때문에, 전체 문서 집합에 대한 배경 지식 없이 텍스트 자체만으로 일관성 있는 문서의 분류가 매우 어렵다.

따라서 본 논문에서는 학습 데이터가 없는 경우에, 전체 문서의 의견 성향 분포에서 벗어난 의견 문서를 검출하는 기법을 제안하고자 한다. 여기에는 각도 기반 이상치 탐지(Angle-Based Outlier Detection)와, 개인화된 페이지랭크(Personalized PageRank) 방법을 활용한다.

3. 편향된 의견 검출

의견 문서를 이상치 탐지 알고리즘에 적용하기 위해서는 문서를 수치로 표현된 벡터로 표현해야 한다. 일반적으로, Bag of Words 방식의 모델링을 사용한다. $\{w_1, w_2, \cdots, w_r\}$를 모든 문서의 어휘 집합이라고 하였을 때, 각각의 문서는 다음 벡터로 표현된다.

$$D = (d_{i1}, d_{i2}, \cdots, d_{ir})$$

이때 d_{i1} 은 문서 d_i 에서 어휘 w_1에 대한값이다. 대표적으로 단어에 대한 빈도수나 TF-IDF 등을 사용한다.

의견 문서의 성향을 벡터에 보다 적절히 반영하기 위해서, 모든 어휘로 벡터를 구성하지 않고 긍정 또는 부정과 같이 의견 극성을 지니고 있는 감성 어휘(Sentiment Lexicon)만을 사용하는 방법이 있다. 이때의 어휘는 사용자가 별도로 정의할 수도 있으며, 센티워드 넷(SentiWordNet)[7]과 같은 사전 데이터를 활용하기도 한다. 수치화에도 어휘 빈도수 외에 어휘가 가지는 의견 극성 값의 합 또는 평균 등의 함수를 적용한다.

3.1 각도 기반 이상치 탐지

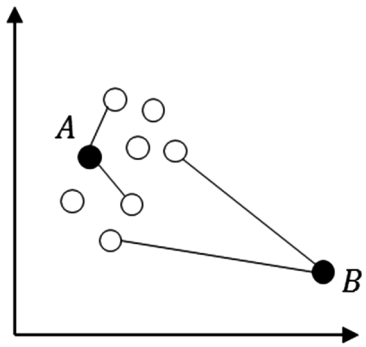
이상치 탐지는 데이터의 집합 내에서 특이하거나 변칙적인 개체를 찾아내는 작업으로, 신용카드 부정 적발, 네트워크 공격 탐지, 건강 감시 등의 응용에 활용되고 있다.

이상치 탐지 방법은 이진방식과 랭킹 방식을 포함해 다양한 방식이 존재한다. 이진 탐지 기법은 초기에 연구된 방식들로, 통계 분포에서 일정 범위를 벗어나는 것과 같이 특정한 조건을 만족하는 개체들을 이상치로 판단하는 방법이다[10]. 랭킹 방식은 최근 활발히 연구가 되고 있는 방법으로, 각 개체의 이상치 정도(Outlierness)를 수치화하여 정의하고, 이 수치에 따라 상위에 위치한 개체들을 이상치로 판별한다. 이 기법은 임계값(Threshold)의 정

의를 통해 쉽게 이진 탐지 기법으로 확장 가능하다.

각도 기반 이상치 탐지는 이러한 방식 중 하나로, 특히 고차원 데이터에 효과적이다[6]. 이 방법은 <Figure 1>과 같이 한 개체를 중심으로 다른 개체들간의 각도의 편차를 통해 이상치를 판별한다. 개체 B와 같이 다른 개체들과의 떨어진 정도가 크다면 서로간에 이루는 각도의 변화가 작지만, 개체 A와 같이 군집 내부에 개체가 있다면 그 각도의 변화는 커진다.

이러한 특성을 활용하여 각도 기반 이상치



〈Figure 1〉 Overview of the Angle Based Outlier Detection

정도 ABOF(Angle Based Outlier Factor)를 정의하는데, 한 개체와 이를 제외한 모든 두 개체 쌍에 대해서 각도를 측정하고, 각도의 변화량을 나타내는 분산에 거리 가중치를 계산한 값으로 정의한다.

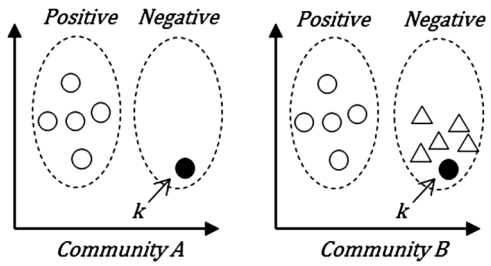
$$\mathrm{ABOF}(\vec{A}) = VAR_{\vec{B},\vec{C}\in D}\left(\frac{<\overline{AB},\overline{AC}>}{|\overline{AB}|^2\cdot|\overline{AC}|^2}\right)$$

$$= \frac{\sum_{\vec{B}\in D}\sum_{\vec{C}\in D}\left(\frac{1}{|\overline{AB}|\cdot|\overline{AC}|}\cdot\frac{<\overline{AB},\overline{AC}>}{|\overline{AB}|^2\cdot|\overline{AC}|^2}\right)^2}{\sum_{\vec{B}\in D}\sum_{\vec{C}\in D}\frac{1}{|\overline{AB}|\cdot|\overline{AC}|}}$$

$$-\left(\frac{\sum_{\vec{B}\in D}\sum_{\vec{C}\in D}\frac{1}{|\overline{AB}|\cdot|\overline{AC}|}\cdot\frac{<\overline{AB},\overline{AC}>}{|\overline{AB}|^2\cdot|\overline{AC}|^2}}{\sum_{\vec{B}\in D}\sum_{\vec{C}\in D}\frac{1}{|\overline{AB}|\cdot|\overline{AC}|}}\right)^2$$

모든 개체들에 대해서 각각의 ABOF의 값을 구해 정렬할 수 있으며, 값이 작을수록 이상치 정도가 높은 것을 의미한다.

의견 편향은 앞서 살펴본 것처럼 문서가 포함되는 집합의 의견 분포에 따라 편향이 결정된다. 이상치 역시 거리 등과 같은 다른 개체들과의 유사도에 따라 정상 또는 이상이 결정된다.



〈Figure 2〉 Biased Opinions in Different Document Sets

<Figure 2>는 문서 벡터를 좌표평면에 사영했을 때의 모습이다. 개체 k는 부정적인 의견을 담고 있는 문서이다. 이때 k는 커뮤니티 A에서는 다른 문서들과 확연히 다른 형태의 의견 성향을 보인다. 하지만 커뮤니티 B에서는 부정적인 의견을 지닌 다른 문서들이 상당수 있기 때문에 이 개체는 편향된 의견이라 보기 어렵다.

다차원 데이터를 고려한 랭킹 기반의 이상치 탐지 기법들은 이러한 개체를 찾아내는데 적용할 수 있다. <Figure 1>과 유사하게, k는 커뮤니티 A에서는 다른 문서들과 거리가 멀리 떨어져있게 되므로 ABOF 값에 따라 높은 랭

크에 위치한다. 하지만 커뮤니티 B에서는 k 와 가까운 거리에 다른 문서들이 다수 분포하기 때문에, ABOF 값은 높은 값을 지니지 않는다. 따라서 각도 기반 이상치 탐지를 편향된 의견의 검출에 적용하는 경우에는 알고리즘에 따라 상위에 랭크 되는 몇 개의 개체를 살펴보고, 그것들의 의견이 다른 것들과 차이가 많이 나는 경우 편향되어 있다고 판단할 수 있다.

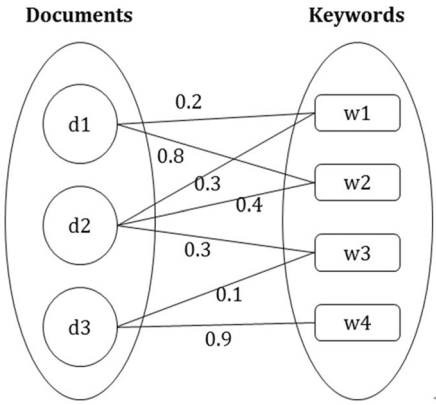
3.2 개인화된 페이지 랭크

페이지랭크(PageRank)는 웹의 페이지의 하이퍼링크로 이루어진 그래프를 통해 각각의 웹페이지의 중요도를 구하는 알고리즘이다. 이들은 검색 시스템에 사용될 뿐만 아니라, 추천 시스템, 소셜 네트워크 분석 등 다양한 응용에서 활용되고 있다. 개인화된 페이지랭크는 이를 확장한 알고리즘으로서, 특정 노드를 시작 지점으로 랜덤 워크를 수행했을 때 다른 노드에 도달할 확률을 계산하고, 이는 곧 시작 노드와 다른 노드들 간의 유사도 또는 중요도를 나타낸다.

구체적으로, 페이지랭크는 그래프에서 ϵ의 확률로 임의의 노드로 점프하고, $1-\epsilon$ 확률로 현재 노드에서 아웃 엣지(Out Edge)로 이동하는 랜덤 워크(Random Walk)를 반복하였을 때 얻게 되는 안정 분포(Stationary Distribution) 확률이다. 개인화된 페이지랭크는 페이지랭크와 같으나, 랜덤하게 점프하는 노드가 항상 동일한 시드 노드(Seed Node)로 가게 될 때의 확률이다. 시드 노드를 u라 할 때, 노드 v의 페이지랭크는 다음과 같다.

$$\pi_u(v) = \epsilon\delta_u(v) + (1-\epsilon)\sum_{\{w|(w,v)\in E\}} \pi_u(w)\alpha_{w,v}$$

앞서 문서를 표현하기 위해 사용한 문서 벡터의 경우, 그래프 형태가 아니기 때문에 개인화된 페이지랭크를 직접 적용하기 어렵다. 이를 위해, 문서 키워드로 이루어진 이분 그래프(Bipartite Graph)로 표현하여 적용하여야 한다. 즉, 문서들을 하나의 노드 집합, 키워드들을 다른 하나의 노드 집합으로 하고, 그들 사이에 빈도수 등의 수치를 엣지(Edge)의 가중치(Weight)로 하는 그래프를 생성하게 된다.



〈Figure 3〉 Example of document-Keyword Graph

<Figure 3>은 이렇게 생성된 그래프의 한 예시이다. 개인화된 페이지랭크를 통한 문서들의 값은 각각의 문서와 가장 유사한 문서를 의미하게 된다. 서로간에 비슷한 키워드를 많이 공유하고 있다면, 랜덤 워크를 통해 도달할 확률이 높게 된다. d1의 경우 d3보다는 d2의 확률이 높게 된다. 이 방법은 직접적인 벡터간의 비교 하는 코사인 유사도(cosine similarity) 방법에 비해, 직접적으로 공유하는 키워드가 없더라도 유사 문서를 찾을 수 있다는 장점이 있다. 키워드 벡터와 같이 벡터가 희소(sparse)한 경우에 이와 같은 방법은 더욱 적절하다.

최종적인 랭크는, 각각의 시작 노드에 대해서 모든 개인화된 페이지랭크를 구한 후 그 랭크 순서를 누적하는 방식으로 구한다. 다수의 문서들에서 페이지랭크가 낮게 나타나는 문서는, 그 문서들과 가장 동떨어져 있다고 볼 수 있다. 예를 들어, d1과 유사한 문서 순서가 d2 d3, d2와 유사한 문서 순서가 d1 d3와 같이 나타난다면, d3가 순위에 대한 누적 값이 가장 크기 때문에 이상치라고 판단한다.

4. 실험 및 분석

실험은 감성 분석 분야에서 많이 사용된 영화리뷰 데이터로 수행하였다[5]. 이 문서 집합은 사전에 나누어진 긍정 문서 1000개와 부정 문서 1000개로 나누어져 있다. 실험은 크게 2가지로, 다수의 긍정적인 문서에서 소수의 부정적인 문서를 검출 하는 것과 다수의 부정적인 문서에서 소수의 긍정적인 문서를 찾는다.

실험 결과는 <Table 1>과 같다. 기본적인 감성 분석 방법(이하 SA(Full))는 전체 문서에서 긍정 800개, 부정 800개의 문서를 학습 데이터로 한 나이브 베이즈 분류기(naïve Bayes classifier)를 적용한 것이다. SA(Small 1)과 SA(Small 2)는 학습셋의 크기와 변화에 따른 성능을 알아보기 위한 실험으로, 긍정문서 25개와 부정문서 25개만을 학습 데이터로 사용한 결과이다.

각도기반 이상치 탐지 기법을 적용한 실험(이하 ABOD)과 개인화된 페이지랭크 방법을 적용한 실험(이하 PPR)은 모두 품사 태깅(POS tagging) 후 형용사와 동사만을 추출 한 후, 이들에 대한 빈도수로 생성한 문서 벡터를 사용하였다. 다른 품사 조합이나, SentiWordNet 등을 사용하였으나, 형용사와 동사만을 사용한 경우가 가장 좋은 성능을 보였다.

<Table 1>의 각 열은 개별적인 실험 셋에 대한 결과들이다. 예를 들어, 긍정문서 20, 부정문서 2, 이상치 Neg.인 경우는 긍정문서 20개와 부정문서 2개를 포함하는 문서셋에서 이상치를 부정문서로 보고 2개인 문서를 찾아내는 것에 대한 정확도를 측정하는 것이다. 실험셋은 나이브 베이즈 분류기에 사용된 학습 데이터를 제외한 나머지 400개의 문서로서 실험 대상을 구성하였다. 이상치가 아닌 문서가 20개인 경우 10개의 실험 셋, 30개인 경우 6개, 50개인 경우 4개, 100개인 경우 2개로 구성하였으며, 결과는 이들에 대한 평균 값이다. 이상치는 각각의 실험셋의 10%, 20%를 상정하고 그 숫자를 적용하였다.

평가는 SA의 경우 실험 대상에 대한 분류 정확도(precision)를 측정하였다. <Table 1>에서처럼 가장 높은 성능을 보이며, 긍정 문서에서 부정문서를 찾는 경우가 그 반대의 경우보다 높은 성능을 보였다. 이는 학습 데이터가 필요한 분류기를 사용한 다른 감성 분석 연구에서와 유사한 결과이다.

ABOD와 PPR는 대상들간의 순서를 매기는 것이기 때문에 분류기인 SA와 동일한 평가방법을 적용할 수는 없다. 이상적인 랭킹은 이상치가 가장 높게 위치하는 것이기 때문에, 찾아야 하는 이상 문서의 개수가 상위에 얼마나 있는지를 정확도로서 측정하였다. 예를 들어, 긍정문서 20개, 부정문서 4개의 문서들에서 부정문서를 찾아야 하는 경우, 1위부터 4위 안에 부정 문서가 몇 개나 있는지를 측정한다.

<Table 1>의 각 열은 개별적인 실험 셋에 대한 결과들이다. 예를 들어, 긍정문서 20, 부정

〈Table 1〉 Experimental Analysis

Data set			Precision				
Number of positive docs	Number of negative docs	Outlierness	SA (Full)	SA (Small 1)	SA (Small 2)	ABOD	PPR
20	2	Neg.	0.941	0.132	0.659	0.050	0.100
20	4	Neg.	0.900	0.204	0.671	0.100	0.125
30	3	Neg.	0.929	0.136	0.667	0.111	0.056
30	6	Neg.	0.903	0.208	0.662	0.083	0.028
50	5	Neg.	0.936	0.132	0.668	0.100	0.050
50	10	Neg.	0.908	0.204	0.654	0.125	0.100
100	10	Neg.	0.927	0.132	0.659	0.050	0.050
100	20	Neg.	0.883	0.204	0.663	0.225	0.100
2	20	Pos.	0.509	0.923	0.750	0.200	0.150
4	20	Pos.	0.550	0.838	0.742	0.250	0.225
3	30	Pos.	0.490	0.909	0.732	0.056	0.111
6	30	Pos.	0.551	0.838	0.731	0.167	0.222
5	50	Pos.	0.514	0.918	0.732	0.100	0.200
10	50	Pos.	0.550	0.846	0.713	0.125	0.300
10	100	Pos.	0.514	0.914	0.741	0.000	0.100
20	100	Pos.	0.546	0.833	0.738	0.150	0.300

문서 2, 이상치 Neg.인 경우는 긍정문서 20개와 부정문서 2개를 포함하는 문서셋에서 이상치를 부정문서로 보고 2개인 문서를 찾아내는 것에 대한 정확도를 측정하는 것이다. 실험셋은 나이브 베이즈 분류기에 사용된 학습 데이터를 제외한 나머지 400개의 문서로서 실험 대상을 구성하였다. 이상치가 아닌 문서가 20개인 경우 10개의 실험 셋, 30개인 경우 6개, 50개인 경우 4개, 100개인 경우 2개로 구성하였으며, 결과는 이들에 대한 평균값이다. 이상치는 각각의 실험셋의 10%, 20%를 상정하고 그 숫자를 적용하였다.

평가는 SA의 경우 실험 대상에 대한 분류 정확도(precision)를 측정하였다. <Table 1>에서처럼 가장 높은 성능을 보이며, 긍정 문서에서 부정문서를 찾는 경우가 그 반대의 경우보다 높은 성능을 보였다. 이는 학습 데이터가 필요한 분류기를 사용한 다른 감성 분석 연구에서와 유사한 결과이다.

ABOD와 PPR는 대상들간의 순서를 매기는 것이기 때문에 분류기인 SA와 동일한 평가방법을 적용할 수는 없다. 이상적인 랭킹은 이상치가 가장 높게 위치하는 것이기 때문에, 찾아야 하는 이상 문서의 개수가 상위에 얼마나 있는지를 정확도로서 측정하였다. 예를 들어, 긍정 문서 20개, 부정 문서 4개의 문서들에서 부정 문서를 찾아야 하는 경우, 1위부터 4위 안에 부정 문서가 몇 개나 있는지를 측정한다.

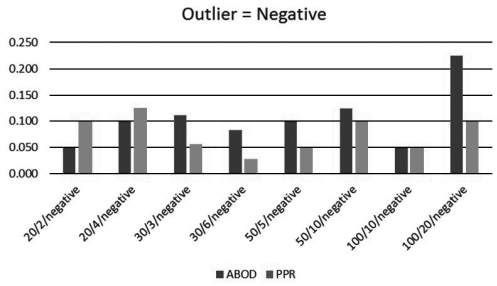
ABOD와 PPR은 SA와 비교하여 낮은 결과를 보였다. 하지만 SA를 적용하기 위해서는 대상의 약 10배가 넘는 1600개 문서에 해당하는 학습 데이터가 필요하다는 단점이 있다. 또한 긍정/부정 두 가지로 분류하는 경우이기 때문에 기본 성능(baseline)이 50%라고 볼 때, 긍정문서가 이상치인 경우의 결과는 기본 성능보다 향상이 거의 없음을 알 수 있다.

또한, SA(Small 1)과 SA(Small 2)는 각각 50개의 학습 데이터를 사용하지만, 학습 데이터의 구성에 따라 성능의 편차가 심해지는 것을 알 수 있다. SA(Small 2)의 경우 성능이 65%~75%로 안정적이다. 그렇지만, SA(Small 1)은 부정 문서에서 긍정 문서를 찾는 것은 높은 정확도를 보이는데 반해, 반대의 경우에는 정확도가 매우 낮다.

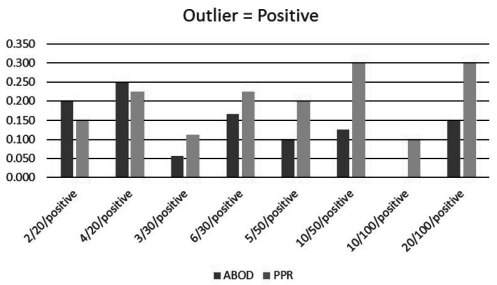
상세히 살펴보면, 실험에서 긍정 문서 30개와 부정 문서 3개로 구성되어 부정 문서를 찾는 경우, 긍정 문서는 대부분 정확히 분류하였지만, 부정 문서는 맞춘 것이 0개였다. 부정문서를 찾는 실험셋의 경우 이와 유사한 경향을 보였다. 따라서 정확도는 90% 정도가 나왔지만, 실제 소수의 의견인 부정 문서를 찾아내기 위해 적용하는 데는 어려움이 있다고 할 수 있다. 이는 학습 데이터의 특성에 따라 분류기가 과적합(Over-fitting)되기 때문에 발생하는 문제로 보인다.

그러므로, 데이터가 충분히 갖추어지지 않은 경우에, ABOD나 PPR과 같은 랭킹 기반의 이상치 탐지 방법을 적용하는 것이 적합하다.

<Figure 4>와 <Figure 5>는 실험 결과 중 ABOD와 PPR의 결과를 차트로 나타낸 것이다. 차트에 따르면, 부정 문서 검출시에는 ABOD의 성능이 더 좋으며, 긍정 문서 검출시에는 PPR의 성능이 더 좋은 것으로 나타난다. 이는 실험 데이터의 특성에 기인한 것으로 보여진다.



〈Figure 4〉 Performance on Finding Negative Documents



〈Figure 5〉 Performance on Finding Positive Documents

5. 결 론

의견 문서는 매우 양질의 정보이며, 이를 유용하게 활용하기 위해서 감성 분석 또는 오피니언 마이닝과 관련된 연구가 활발히 진행되었다. 하지만 이러한 연구는 문서의 긍정, 부정을 판별하는 기법들에 집중하였으며, 연구 스팸 의견이나 편향된 의견을 탐지하는 것은 그 중요성에 비해 상대적으로 적은 관심을 받아왔다. 본 논문에서는 전체 의견 문서에서 편향된 문서를 찾아내는 방법으로, 각도 기반의 이상치 탐지와 개인화된 페이지랭크 기법을 적용하는 방법을 제안하였다. 이렇게 찾아진 의견 문서는 향후 소셜 미디어 분석을 더욱 정교하게 하는데 활용될 수 있다.

References

[1] Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H. and Jin, C., "Red Opal : Product Feature Scoring from Reviews," In Proceedings of the 8th ACM conference on Electronic Commerce, 2007.

[2] Jindal, N. and Liu, B., "Opinion Spam and Analysis," In Proceedings of the international conference on Web search and web data mining, 2008.

[3] Castillo, C. and Davison, B. D., "Adversarial Web Search," Foundations and Trends in Information Retrieval, Vol. 4, No. 5, 2010.

[4] Liu, B., "Web Data Mining : Exploring Hyperlinks, Contents, and Usage Data," Springer, 2011.

[5] Pang, B., Lee, L. and Vaithyanathan, S., "Thumbs up? Sentiment Classification using Machine Learning Techniques," In Proceedings of the ACL 02 conference on Empirical methods in natural language processing, Vol. 10, 2002.

[6] Ding, X., Liu, B., and Yu, P. S., "A holistic lexicon based approach to opinion mining," In Proceedings of the international conference on Web search and web data mining, 2008.

[7] Hu, M. and Liu, B., "Mining and summarizing customer reviews," In Proceedings of the 10th ACM SIGKDD international conference on Knowledge Discovery and Data mining, 2004.

[8] Liu, B., Hu, M. and Cheng, J., "Opinion observer : analyzing and comparing opinions on the Web," In Proceedings of the 14th international on World Wide Web, 2005.

[9] Scaffidi, C., Bierhoff, K., Chang, E., M. Felker, Ng, H. and Jin, C., "Red Opal : Product Feature Scoring from Reviews," In Proceedings of the 8th ACM conference on Electronic Commerce, 2007.

[10] Jin, W., Ho, H. and Srihari, R., "OpinionMiner : a novel machine learning system for web opinion mining and extraction," In Proceedings of the 15th ACM SIGKDD international conference on Knowledge Discovery and Data mining, 2009.

[11] Esuli, A. and Sebastiani, F., "Determining Term Subjectivity and Term Orientation for Opinion Mining," In Proceedings of 11th conference of the European chapter of the Association for Computational Linguistics, 2006.

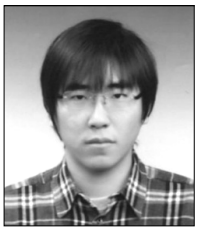
[12] Denecke, K., "Using SentiWordNet for Multilingual Sentiment Analysis," In Proceedings of the International Conference on Data Engineering : ICDE, Workshop on Data Engineering for Blogs, Social Media, and Web 2.0, 2008.

[13] Lim, E., Nguyen, V., Jindal, N., Liu, B., and Lauw, H., "Detecting product review spammers using rating behaviors," In Proceedings of the 19th ACM international conference on Information and knowledge management, 2010.

[14] Mukherjee, A., Liu, B. and Glance, N., "Spotting fake reviewer groups in consumer reviews," In Proceedings of the 21st international conference on World Wide Web, 2012.

[15] Yeom, J., Lee, D. Shim, J., Lee, S. g., "Product Review Data and Sentiment Analytical Processing Modeling," The Journal of Society for e-Business Studies, Vol. 16, No. 4, 2011.

저 자 소 개



연종흠 (E-mail : jonghm@europa.snu.ac.kr)
2008년 서울대학교 컴퓨터공학부 졸업 (학사)
2008년~현재 서울대학교 컴퓨터공학부 대학원 박사과정
관심분야 데이터베이스, 자연언어처리, 데이터 마이닝

심준호 (E-mail: jshim@sookmyung.ac.kr)
1990년 서울대학교 계산통계학과 졸업 (학사)
1994년 서울대학교 계산통계학과 전산과학전공 (석사)
1998년 Northwestern University, Electrical and Computer Engineering (박사)
2001년~ 현재 숙명여자대학교 컴퓨터과학부 교수
관심분야 데이타베이스, 전자상거래, 상품정보, 온톨로지

이상구 (E-mail : sglee@europa.snu.ac.kr)
1985년 서울대학교 계산통계학과 졸업 (학사)
1987년 Northwestern University, Computer Science (석사)
1990년 Northwestern University, Computer Science (박사)
1992년~현재 서울대학교 컴퓨터공학부 교수
관심분야 데이터베이스, 상황 인지, 추천 시스템