

소셜 미디어 상에서의 인공지능 관련 사회적 여론에 대한 다 범주 감성 분석*

이상원

한양대학교 일반대학원
비즈니스인포매틱스학과
(lsw417@hanyang.ac.kr)

최창욱

한양대학교 일반대학원
비즈니스인포매틱스학과
(winehouse@hanyang.ac.kr)

김동성

한양대학교 일반대학원
경영학과
(paulus82@hanyang.ac.kr)

여운영

한양대학교 일반대학원
비즈니스인포매틱스학과
(ywy31730@hanyang.ac.kr)

김종우

한양대학교 경영대학
경영학부
(kijw@hanyang.ac.kr)

인공지능 기술의 비약적인 발전으로 인하여, 사용자의 편의성 증대를 목적으로 다양한 분야에서 관련된 제품과 서비스들의 개발이 이루어지고 있다. 이러한 기술의 발전에는 긍정적인 과급 효과에 대한 기대감이 존재하나, 향후 발생 가능한 부정적인 측면에 대한 논의도 활발히 이루어지고 있다. 예를 들어, 인공지능 기술 기반의 자율주행 자동차의 경우 안정성의 향상이라는 측면에서 많은 관심을 받고 있으나, 트롤리 딜레마, 시스템 보안 문제 등의 사회적 이슈 또한 활발히 논의되고 있다. 이에 따라, 인공지능 관련 기술의 발전과 사회적 수용을 위해서는 사회적으로 논의되는 주요 관련 이슈들에 대한 확인과 효과적인 분석이 요구된다.

이를 위해, 본 연구에서는 ‘이세돌 vs 알파고’ 시점인 2016년 3월을 포함하여 2016년 1월부터 2017년 12월까지 2년 동안의 인공지능과 관련된 사회적 이슈들을 파악하고 온라인상에서 발생하는 사회적 여론에 대하여 다 범주 감성을 분석하고자 한다. 이를 위하여 국내 대표적인 포털 사이트에서 인공지능 관련 뉴스의 수와 관련된 뉴스 제목, 뉴스의 댓글을 웹 크롤링(Web Crawling) 하였다. 사회적 여론에 대한 다 범주 감성 분석은 논의되는 이슈들의 중요성을 고려하여 단순 긍정 또는 부정이 아닌, 분노, 혐오, 두려움, 행복, 중립, 슬픔, 놀라움의 7가지 다 범주 감성으로 분석하였다.

분석 결과, 대부분의 이벤트 기간에 대하여 1위 감성은 ‘행복’으로 나타났지만 각 키워드에 대하여 나오는 감성이 상이함을 볼 수 있었다. 또한 2016년 상반기, 하반기, 2017년 상반기, 하반기로 나누어 보았을 때 시간이 지남에 따라 ‘분노’의 감성이 낮아짐을 확인하였다. 이러한 분석 결과를 바탕으로 인공지능과 관련하여 현재 논의되고 있는 다양한 이슈와 동향 파악이 가능하며, 이에 대한 대응 방안 마련에 활용이 가능할 것이다. 향후 감성 분석기의 성능 향상과 댓글에 대한 공감 및 비공감도의 가중치를 추가하여 분석한다면 사회적 여론을 보다 세밀하게 파악 할 수 있을 것이다.

주제어 : 인공지능, 감성 분석, 사회적 여론, 온라인 뉴스, 온라인 댓글

논문접수일 : 2018년 5월 29일 논문수정일 : 2018년 9월 9일 게재확정일 : 2018년 10월 24일
원고유형 : 학술대회(급행) 교신저자 : 김종우

* 이 논문 또는 저서는 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017S1A3A2066740).

1. 서론

컴퓨터 처리속도의 개선과 알고리즘의 개발로 오늘날 인공지능은 오목과 체스를 넘어 바둑에 이르기까지 인간과 대결 가능하게 되었고 2016년 3월에는 이세돌 9단과의 대국에서 승리하면서 인공지능(AI: Artificial Intelligence) 기술은 대중들에게 화제가 되었다(Kim and Kim, 2016). 이러한 인공지능 기술의 발전과 함께 각 분야에서는 인공지능 제품, 서비스의 개발이 이루어지고 실생활에서 체감할 수 있도록 가시화됨에 따라 다양한 사회적 논란이 대두되었다. 하지만 이 논란과는 별개로 인공지능과 관련된 이슈들에 대해 대중이 어떤 생각을 가지는가를 분석하기 위한 연구는 아직 미비한 편이다.

2016년 3월 이세돌과 알파고의 대국 이후에 인공지능에 대한 관심이 증가하여 다양한 사회적 이슈가 제기되고 있다. 예를 들어 자율 주행의 윤리적 딜레마와 같은 문제들과 더불어 음성 인식 및 번역을 하는 다양한 서비스 및 상품에 대한 이슈, 세종대에서 열린 ‘인간 vs 인공지능’의 스타크래프트 대결 등 최근 인공지능과 관련하여 다양한 사회적 이슈들이 발생하고 있다. 이러한 인공지능과 관련된 여러 이슈에 대하여 각 이슈의 발생원인을 파악하고 이에 대한 사회적 여론의 감성을 파악 할 수 있다면 기업이나 국가에서 관련 이슈에 대한 대응책을 마련하는데 기초가 될 수 있을 것이다.

본 연구에서는 인공지능 관련 여론을 파악하기 위한 방법으로 뉴스 댓글을 선정하였고 국내에서 가장 대표적인 포털 사이트에서 뉴스 기사에 달린 댓글을 수집하였다. 인터넷 댓글은 여론을 파악하는 방법으로 활용된다(Yang, 2008). 그리고 뉴스 기사 속의 댓글은 댓글 자체가 여론으

로의 역할을 할 뿐만 아니라, 뉴스 이용자들은 뉴스 댓글을 읽고 개인의 의견 변화에 영향을 받는다(Kim and Sun, 2006; Kang and Kim, 2012). 이러한 선행 연구를 참고하여 본 연구에서는 국내에서 가장 대표적으로 사용되는 웹 포털의 뉴스와 뉴스의 댓글을 활용하여 인공지능에 대한 사회적 여론을 분석하고자 한다.

본 연구에서는 인공지능이 우리 사회에 자리매김한 ‘이세돌 vs 알파고’ 시점인 2016년 3월을 포함하여 2016년 1월부터 2017년 12월까지 2년 동안의 인공지능과 관련된 사회적인 이슈들을 파악하고 각 이슈에 대한 댓글에 대하여 7가지 다 범주 감성 분석(분노, 혐오, 두려움, 행복, 중립, 슬픔, 놀라움)을 실시하였다. 인공지능과 관련된 사회적 이슈의 파악은 국내 대표 포털 서비스를 활용하였고, 온라인 뉴스의 기사 수를 기반으로 기사의 수가 급증하면 해당 이벤트에 대하여 이슈가 되었다고 판단하고 이벤트 기간으로 설정하였다. 이벤트 기간을 파악 후 해당 이벤트 기간에 대한 뉴스 제목 크롤링 및 이에 대한 형태소 분석, 빈도 분석을 활용하여 이벤트 기간과 관련된 키워드를 추출하였다. 마지막으로 각 이벤트 기간에서 관련된 키워드로 뉴스 검색 후 댓글을 수집하여 감성 분석을 실시 하였다. 뉴스 댓글에 대한 감성 분석은 합성곱 신경망(CNN: Convolutional Neural Networks)을 활용해 댓글별로 7가지 감성의 확률을 예측하였고 가장 가능성이 높은 감성에 대한 빈도분석을 통해 각 이슈에 대한 감성을 파악하였다. 이러한 여론 파악은 향후 정부의 정책적 방향을 잡거나 기업의 의사결정에 앞서 여론에 대한 동향 파악을 보다 세밀하게 하는 것을 가능하게 할 것이다. 추가적으로 본 연구에서는 ‘인공지능’의 키워드로 인공지능 관련 여론을 파악하였지만, 다른 키워드에 대해

서 본 연구 방법을 사용하여 분석하고자 하는 키워드에 대한 특정 기간, 특정 이슈에 대해 여론 파악도 가능하다.

본 논문의 구성은 다음과 같다. 2장에서는 온라인 뉴스 기사의 특성, 댓글의 효과에 대한 기존 관련 연구들을 검토하고 인공지능 동향 및 감성 분석에 관련된 선행 연구를 살펴본다. 3장에서는 인공지능과 관련하여 이벤트 기간과 키워드를 파악 및 감성 분석의 연구 절차에 대하여 설명하고 4장에서는 인공지능 각 이슈별 대중의 감성에 대한 연구 결과를 제시한다. 마지막 5장에서는 결론과 연구의 한계점에 대하여 설명하고 보완점과 향후 연구 방향에 대하여 제시한다.

2. 관련 연구

2.1 온라인 뉴스 댓글

소셜 네트워크 서비스(SNS, Social Networking Service)에서 생산되는 데이터의 양이 빠르게 증가하고 영향력이 커지면서 이에 대한 데이터 활용에 많은 관심과 연구가 수행되고 있다. 그리고 소셜 네트워크 서비스의 공유된 데이터에서 여론을 파악하기 위한 방법으로 온라인 댓글의 활용 방안과 영향에 관한 연구들은 지속적으로 수행되어 왔다(Jeong and Kim, 2006; Koh, 2018; Kim and Lee, 2011; Park, 2013).

국가의 정책적 결정이나 기업의 전략적 의사 결정 이전의 여론 파악은 국가의 정책 방향에 영향을 줄 뿐 아니라 기업 활동의 방향을 잡고 기업이 속한 경제 상황의 변화에도 영향을 주게 되므로 필수 불가결한 요소라고 할 수 있다. 이러한 의사 결정에 중요한 영향을 끼치는 여론 파악

은 소셜 네트워크 서비스에서 생산되는 댓글을 통해 가능하다(Yang, 2008).

소셜 네트워크 서비스 중 특히 뉴스 기사에 달린 댓글을 통해서도 여론을 파악 할 수 있는데, 이러한 인터넷 뉴스 댓글을 통한 여론 파악은 댓글 자체 만으로도 여론 파악을 하는 것뿐만 아니라 온라인 뉴스를 이용하는 사람들에게 영향을 미치기도 한다(Kang and Kim, 2012). 또한 인터넷 뉴스 댓글은 단순히 여론 파악을 위한 도구일 뿐만 아니라 인터넷 뉴스를 이용하는 사람들에게 여론의 동향을 추정하는 단서가 되기도 한다(Lee and Jang, 2009). 이처럼 인터넷 댓글은 여론 파악의 역할 뿐만 아니라 인터넷 뉴스를 이용하는 사람들에게 해당 이슈에 대한 여론을 추정하고 댓글을 통해 영향을 받고 있다.

뉴스 댓글과 뉴스 이용자 간의 영향은 다수의 선행 연구들을 볼 때 댓글 자체가 뉴스 이용자에 의견 형성에 영향을 미친다고 볼 수 있다. 이슈에 대한 관여도와 관련하여 해당 이슈와 이용자 간의 관여도가 높다면 개인의 의견 변화에 대한 가능성이 낮을 수 있고, 관여도가 낮은 이슈에 대해서는 이용자의 의견 변화의 가능성이 높을 수 있다. 그러나 국내 연구들의 공통된 의견은 뉴스의 댓글이 영향력이 있고 뉴스를 이용하는 사람들에게 영향을 미쳐 기존의 태도에 변화를 일으킨다는 것이다(Lee and Jang, 2009; Kim and Sun, 2006; Kang and Kim, 2012). 그러므로 이러한 뉴스의 댓글을 통한 감성 분석은 국내 여론 동향을 추정하는 단서가 될 것이다.

기존 다수의 국내 연구들의 경우 뉴스의 댓글의 자체에서 여론을 파악하기 위하여 다 범주 감성 분석을 하는 연구는 미비한 편이다. 또한 특정 이슈에 대하여 보다 다양한 각도에서 여론을 파악하기 위해서는 댓글에 대한 긍정, 부정, 대

한 여론 파악을 진행하기보다는 다 범주의 감성 분석이 요구된다.

2.2 인공지능 기술 동향

인공지능이라는 분야는 1956년 개최된 Dartmouth Conference에서 AI (Artificial Intelligence)라는 용어가 채택되면서 확립되었다. 우리나라의 경우 2016년 3월 이세돌 기사와 ‘구글 딥마인드 (Google Deepmind)’가 개발한 인공지능 바둑 프로그램 알파고의 대결을 진행하면서 인공지능이라는 단어는 우리 사회에 자리매김 하게 되었다.

인공지능의 폭발적인 성장과 동향에 반응하여 미래부와 정보통신기술진흥센터에서도 인공지능 관련 기술들을 5개의 핵심 기술과 15개의 세부 기술로 분류하게 되었고, 국가 전략 프로젝트의 9개의 과제 중 하나로 인공지능은 선정되었다. 급격한 성장과 동시에 미래 육성 핵심 기술로 자리 잡은 인공지능은 산업 전반에 대한 영향력이 광범위하고 이에 대한 파급효과가 다양하여 산업, 정치, 사회, 문화적 논의가 활발히 이루어지고 있고 인공지능과 관련된 기술 동향에 대한 연구 역시 지속적으로 이루어지고 있다(Won and Lee, 2016; Cho 2013).

초기 규칙 기반(Rule Based) 프로그램에서 기계학습(Machine Learning), 딥러닝(Deep Learning)에 이르기까지 진화한 인공지능은 현재 실시간 번역, 영상인식, 음성인식, 감성 분석 등 다양한 분야에서 뛰어난 성능을 보이고 있다. 그러나 이러한 인공지능의 발전 초기에는 긍정적인 전망이 주를 이루었지만 점차 이에 대한 윤리, 도덕적 문제나 일자리 문제 등 현실적인 문제들에 직면하면서 이에 대한 다양한 사회적 여론이 존재하고 있다(You, 2017).

본 논문에서는 이러한 인공지능의 발전과 함께 다양한 사회적 여론이 존재함을 인지하고 이에 대한 세부적인 여론 파악을 위하여 단순 긍정, 부정이 아닌 각 이슈별로 구분된 다 범주 감성 분석(분노, 혐오, 두려움, 행복, 중립, 슬픔, 놀라움)을 실시하였다. 이러한 인공지능에 대한 여론 파악은 시간 흐름에 따른 인공지능의 인식 변화와 더불어 각 이슈에 대한 사회적 여론을 상세히 파악하고 이를 바탕으로 기업이나 국가의 정책적 방향을 잡거나 의사 결정 함에 있어서 사회적 여론에 대한 깊이 있는 이해가 가능할 것이다.

2.3 감성 분석

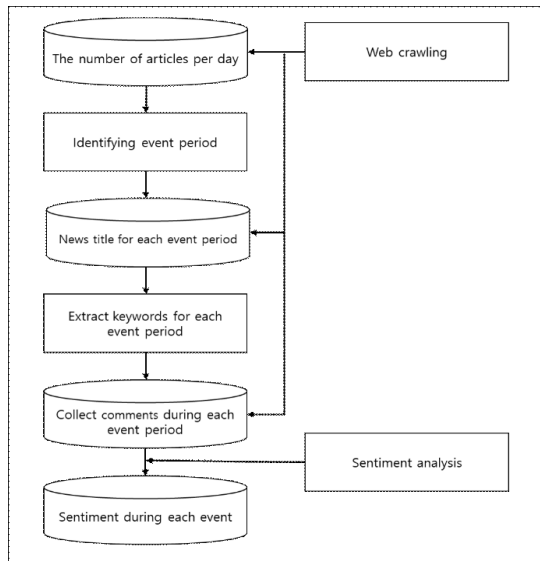
감성 분석이란 많은 사람으로부터 작성되는 대용량의 텍스트 데이터에서 주요 의견이나 감성, 태도, 성향을 추출하는 텍스트 마이닝(Text Mining)의 한 영역으로 오피니언 마이닝(Opinion Mining)이라고도 한다. 소셜미디어와 같은 온라인 환경에서 생성되는 텍스트 데이터의 양이 증가함에 따라, 자연어 처리 기술의 중요도가 높아지고, 이에 따라 감성 분석에 관한 연구는 국내외적으로 활발히 이루어지고 있다(Kim and Kim 2014; Kim et al. 2016; Han and Kim 2018).

감성 분석의 대표적인 방법으로 감성 사전을 구축하여 텍스트 데이터의 감성을 추출하는 방안과 구문 분석을 통해 분석하는 방안 등이 있으나, 최근에는 감성 사전 구축 및 구문 분석을 하지 않고, 단어 혹은 글자를 기반으로 한 합성곱 신경망(CNN: Convolutional Neural Network)이 활용되고 있다(Kim et al. 2017). 특히 글자를 기반으로 한 합성곱 신경망은 오타자나 초성체만 사용된 글 내에서도 감성 분석이 가능하다.

3. 연구 방법

3.1 연구 절차

본 연구는 다음과 같은 연구 절차를 통해 수행되었다(<Figure 1> 참조). 먼저 인공지능 관련 이벤트 기간을 파악하기 위하여 국내 대표적인 온라인 포털인 네이버에서 인공지능과 관련된 일별 기사 수를 파악하였다. 수집된 데이터는 2016년 1월 1일부터 2017년 12월 31일까지의 뉴스의 수를 웹 크롤링(Web Crawling)하였다.



<Figure 1> Research Process

이렇게 얻어진 일별 인공지능 관련 기사의 수는 이슈가 되는 기간을 파악하기 위하여 일별 기사 수의 평균에서 $+2\sigma$ (표준편차) 이상인 날짜를 추출하여 이벤트 기간으로 설정하였다. 이벤트 날짜가 연속적으로 출현하는 경우에는 하나의 이벤트로 묶어 활용하였다. 인공지능과 관련하여 이슈가 되는 기간을 파악 후 해당 기간에 어

떤 뉴스가 많았는지 확인하기 위한 뉴스 제목 웹 크롤링을 실시하였다.

결과적으로 얻어진 각 이벤트 기간별 뉴스 제목에 대한 데이터를 형태소 분석, 빈도 분석을 활용하여 빈도가 많은 단어를 추출하였다. 각 이벤트 기간의 뉴스 제목 데이터에서 형태소 분석과 빈도분석을 통해 얻어진 단어들을 인공지능과 관련된 해당 이벤트 기간의 키워드로 설정하였다. 그리고 해당 키워드를 이전 단계에서 파악한 이벤트 기간으로 크롤링하여 키워드 관련 뉴스와 이에 대한 댓글을 수집하였다.

이렇게 얻어진 각 이벤트 기간에 대한 댓글을 활용하여 댓글에 대한 7가지 다 범주 감성 분석을 실시하였고 가장 높은 확률의 감성을 파악하였다. 하나의 댓글에 대하여 가장 높은 확률의 감성을 해당 댓글의 감성이라고 설정하고 이벤트 기간에 대한 댓글들의 7가지 감성 비율을 확인하였다. 각 이벤트별 감성의 비율은 ‘이벤트 기간의 각 감성빈도 / 이벤트 기간의 전체 감성빈도’로 측정하였다(Kim et al. 2016, <Formula 1> 참조).

$$\text{Sentiment ratio} = \frac{\text{Frequency of each sentiment}}{\text{Total sentiment frequency}} \quad (1)$$

3.2 연구 데이터

본 연구는 국내 대표 포털 사이트인 네이버의 뉴스에서 2016년 1월 1일부터 2017년 12월 31일까지 수집한 인공지능 관련 네이버 뉴스의 수와 기사 제목, 댓글 데이터를 활용하였다. 그러나 AI와 인공지능 관련 뉴스 기사의 수를 파악할 때 AI 키워드는 동음이의어로 기존의 조류인플루엔자 바이러스(AI, Avian Influenza)와 키워드가 섞여 검색어에서 AI는 제외하고 ‘인공지능’을 활용

<Table 1> Extract keywords based on event period

Number	Event Period	Event keywords			
		1	2	3	4
1	2016-01-28	이세돌	컴퓨터		
2	2016-03-09 ~ 2013-03-17	이세돌	알파고	인간	
3	2016-08-31	누구	음성인식	SK텔레콤	NUGU
4	2016-12-15	파파고	음성인식	번역	
5	2016-12-29	빅데이터	의료	셀바스	
6	2017-01-06	자율주행	엔디비아	자동차	
7	2017-01-09	스마트폰			
8	2017-01-16 ~ 2017-01-18	지니	비서		
9	2017-03-29 ~ 2017-03-30	갤럭시	빅스비		
10	2017-10-11	대통령	뽀통령		
11	2017-10-31	송병구	스타크래프트		

하여 해당 이벤트 기간을 추출하였다. 그 결과 23개의 이벤트 기간을 추출하였고 키워드가 없거나 댓글이 없는 이벤트 기간은 제외하여 최종적으로 11개의 이벤트 기간을 추출하였다. 최종적으로 추출된 이벤트 기간과 키워드는 다음과 같다(<Table 1> 참조).

이벤트 기간에 인공지능과 관련된 키워드를 추출 후 해당 키워드를 활용한 뉴스 재검색을 통

해 키워드와 관련된 뉴스를 수집하였다. 키워드 중에는 이벤트 기간이라고 하더라도 컴퓨터, 네이버, 자동차 등의 키워드의 경우 인공지능과 관련되지 않는 기사가 추출되어 검색어에 인공지능을 더하여 웹 크롤링을 진행하였다. 키워드 한 개 이상인 경우에는 뉴스 링크가 중복되어 링크에 대한 중복 제거 후 뉴스 댓글을 수집하였다. 수집된 댓글 데이터는 총 171,454개로 이벤

<Table 2> News comment and sentiment (31 October 2017)

Number	Comments	Sentiment
1	어디 쓰레기 같은 프로그램 가지고 와서 ㅋㅋ.....	disgust
2	진짜 해도해도 너무하네.. 블리자드 ai는 개발비 얼마나 쓴거냐?	disgust
3	말이 AI지 그냥 컴터랑 한 거 아닌가요?	neutral
4	뭘 어떻게 만들었길래 하나같이 4드론 찍고 달리냐ㅋㅋㅋㅋㅋㅋ	happiness
5	송병구 몰래카메라ㅋㅋㅋㅋㅋㅋㅋ별풍선 5만개 ㅋㅋㅋ	happiness
6	이걸 왜 인공지능이라 하지?? 누구는 인류를 압승하고 누구는 완패당하면 AI 라고 해도, 기술력으로 자기 얼굴에 칩빨기인데?	anger
7	4드론 ㅋㅋㅋㅋㅋㅋㅋㅋ	happiness
8	al 연구비가 얼마라구요?????ㅋㅋㅋ	happiness
9	이거 횡령 아니고선 설명이 안되는데?	surprise
10	알파고도 아니고 그냥 한국산명칭이가지고 이겼다고 호들갑이네	anger

〈Table 3〉 The frequency and ratio of sentiment in news comments (31 October 2017)

Sentiment	Happiness	Disgust	Neutral	Anger	Surprise	Sadness	Fear	Total
Frequency	163	112	86	66	65	37	18	547
Ratio(%)	29.8	20.5	15.7	12.1	11.9	6.8	3.3	100

트 기간별로 감성 분석기를 통해 7가지 다 범주 감성으로 분류 하였다(<Table 2> 참조). 마지막으로 해당 이벤트 기간의 감성을 비율 분석하여 이벤트 기간에 어떤 감성이 주를 이루었는지 감성 순위를 비교하였다(<Table 3> 참조).

3.3 활용한 감성 분석기

감성 분석의 대표적인 방법은 감성 사전을 구축하여 텍스트 데이터의 감성을 추출하는 방안과 구문 분석을 통해 분석하는 방안 등이 있으나, 본 연구에서는 딥러닝 기법 중 하나인 글자 기반의 합성곱 신경망(CNN: Convolutional Neural Network)을 활용하였다(Kim et al. 2017).

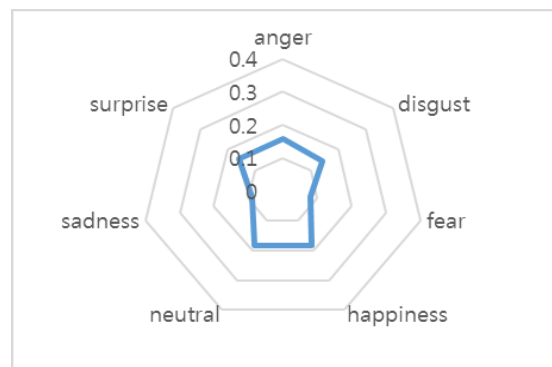
기존에 구축한 글자 기반 합성곱 신경망 감성 분석 모델을 활용하여 7가지 감성(분노, 혐오, 두려움, 행복, 중립, 슬픔, 놀라움)으로 분류된 SNS 글, 영화 및 대본 등의 데이터를 수집하여 학습과 테스트를 실행하였다. 7가지 감성은 Ekman et al. (1987)의 연구에서 제시된 6가지 감성(분노, 혐오, 두려움, 행복, 슬픔, 놀라움)과 ‘중립’의 감성을 더하여 7가지 감성으로 분류되었다. 학습 데이터는 11,765개로 학습시켰으며 2,945개의 테스트 데이터에 대하여 74.25%의 정확도를 나타냈다. 이러한 글자 기반의 합성곱 신경망은 기존의 감성 사전 기반의 감성 분류기에 비해 정확도가 높은 장점이 있다(Kim et al. 2017). 또한 맞춤법이 틀리거나 초성체만 사용한 감정 표현, 글자

를 이용한 이모티콘의 표현에도 감성 분석이 가능하여 뉴스 댓글의 감성 분석에 적합하다고 판단된다.

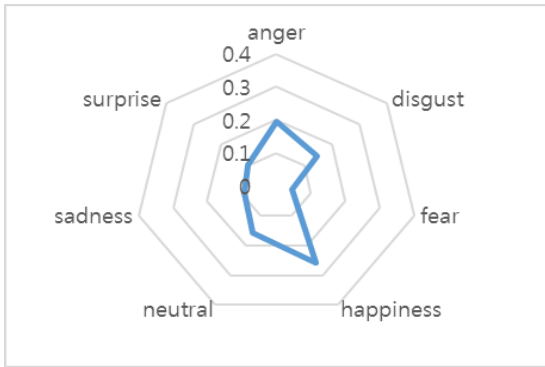
4. 분석 결과

본 논문에서 분석한 인공지능과 관련된 사회적 여론의 다 범주 감성 분석의 결과는 다음과 같다(<Figure 2~12> 참고).

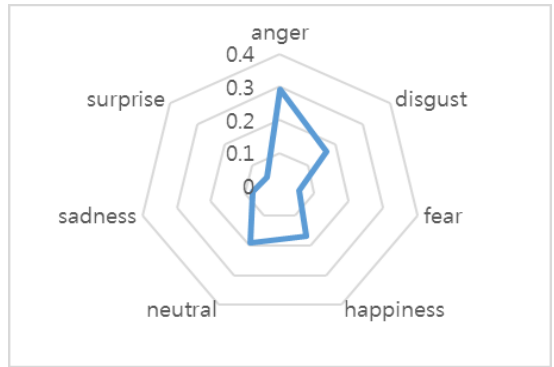
이벤트 기간별 키워드를 추출한 후 해당 키워드로 이벤트 기간의 뉴스를 검색하였다. 이를 통해 뉴스 링크를 추출하였고 중복 링크를 제거한 후 총 171,454개의 댓글을 수집하였다. 분석 결과 11개의 인공지능 관련 이벤트에 대하여 10개의 이벤트는 ‘행복’, 1개의 이벤트는 ‘분노’가 가장 많은 빈도의 감성으로 확인되었다.



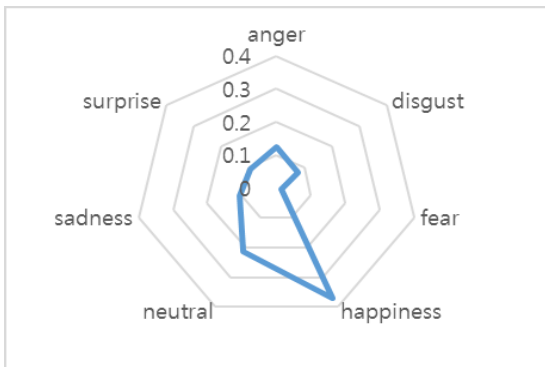
〈Figure 2〉 Sentiment distribution of 28 January 2016



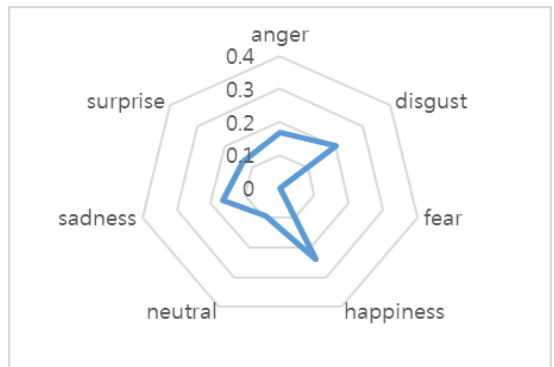
〈Figure 3〉 Sentiment distribution from March 7 to 17, 2016



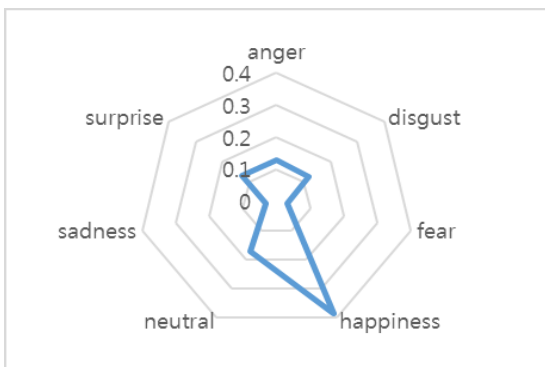
〈Figure 6〉 Sentiment distribution of 29 December 2016



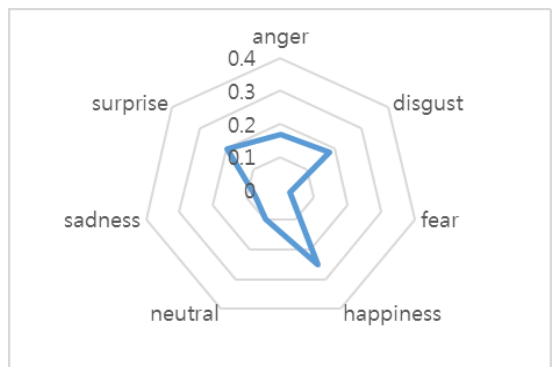
〈Figure 4〉 Sentiment distribution of 31 August 2016



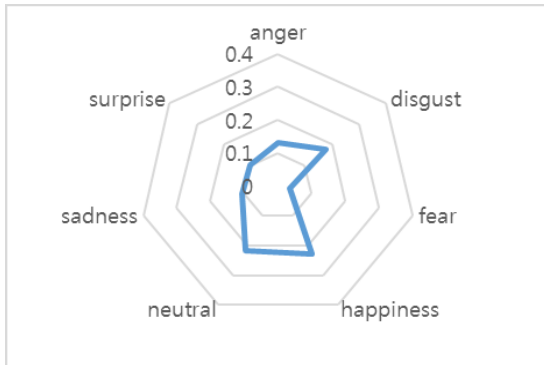
〈Figure 7〉 Sentiment distribution of 6 January 2017



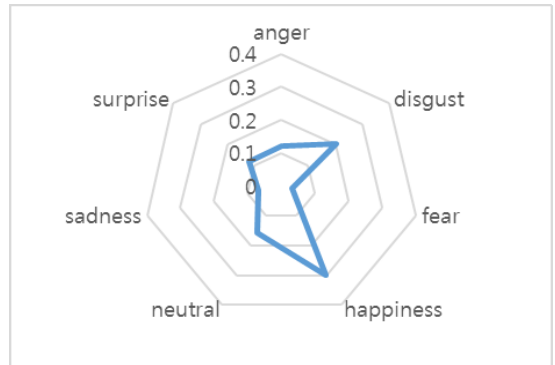
〈Figure 5〉 Sentiment distribution of 15 December 2016



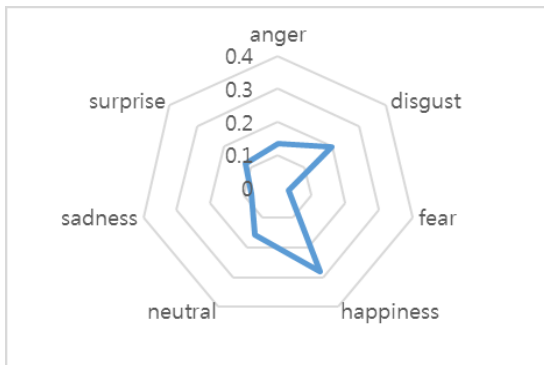
〈Figure 8〉 Sentiment distribution of 9 January 2017



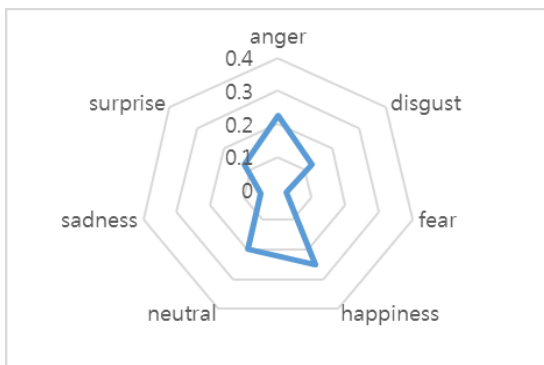
〈Figure 9〉 Sentiment distribution from January 16 to 18, 2017



〈Figure 12〉 Sentiment distribution of 31 October 2017



〈Figure 10〉 Sentiment distribution from March 29 to 30, 2017



〈Figure 11〉 Sentiment distribution of 11 October 2017

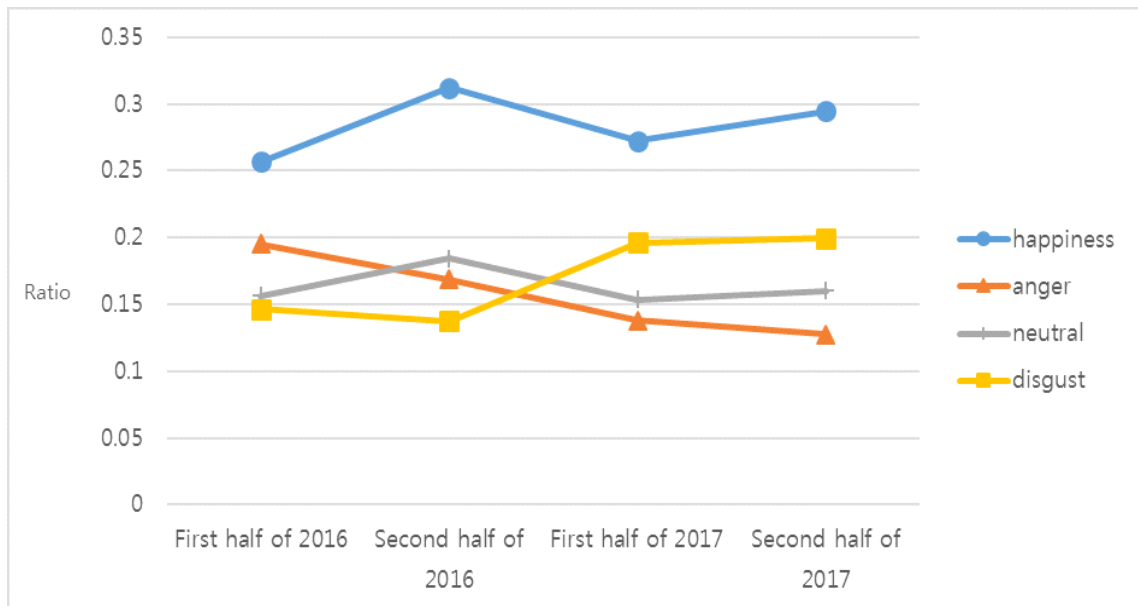
이벤트 기간별로 추출된 11개의 이슈를 살펴 보면 2016년 1월 26일의 이벤트의 경우 이세돌 기사와 ‘구글 딥마인드(Google Deepmind)’가 개발한 인공지능 바둑 프로그램 알파고의 대결을 공식 발표하는 기사가 주를 이루었다. 대국이 시작되기 전 인공지능에 대한 여론은 대국이 진행되면서 파악한 댓글의 여론보다 감성 비율이 크게 분포한 것을 확인 할 수 있었다. 2016년 3월 7일에서 17일 사이의 이벤트는 이세돌 기사와 알파고의 대국이 진행되면서 파악한 여론이다. 대국을 발표하기 전보다 확연히 ‘행복’, ‘분노’가 증가한 것을 볼 수 있었고 이세돌 기사에 대한 응원의 메시지가 다수 있는 것을 확인하였다. 2016년 8월 31일의 인공지능 관련 이벤트는 SK텔레콤에서 최초 인공지능 음성인식 디바이스 ‘NUGU’를 출시한다는 것이었다. 이에 대한 여론은 기존에 없던 인공지능 제품에 대한 긍정적인 반응 및 기능에 대한 기대감이 주를 이루었다. 2016년 12월 15일의 경우 파파고 및 인공지능 음성인식, 번역에 관련된 기사가 이벤트로 선정되었고, 2016년 12월 29일에는 빅데이터를 활용하여 구제역 확산 방지 및 의료 기술 관련 기

사가 많았다. 해당 이벤트 기간은 빅데이터를 활용한 규제역 확산 방지에 대한 기술의 허점 및 지난 정부에 대한 비난 여론이 많은 것을 확인할 수 있었다.

2017년 1월 6일의 인공지능 관련 이벤트는 자율주행 관련된 내용이 이벤트로 선정되었다. 2017년 1월 9일의 인공지능 관련 이벤트는 스마트폰과 관련된 이벤트로 스마트폰에 인공지능 기능이 탑재되는 것을 이벤트로 확인할 수 있었다. 2017년 1월 16일에서 2017년 1월 18일의 이벤트는 SK텔레콤의 음성인식 디바이스 ‘NUGU’에 이어 KT의 ‘기가 지니’ 출시 및 이에 대한 기사임을 확인하였다. 2017년 3월 29일부터 2017년 3월 30일의 이벤트는 갤럭시S8과 이에 탑재된 음성인식 플랫폼 ‘빅스비’로 인해 인공지능 관련 이슈가 발생함을 확인하였다. 2017년 10월

11일의 이벤트는 4차산업혁명위원회 출범식에서 문재인 대통령과 인공지능 캐릭터 로봇인 ‘뿌로봇’이 대화한 기사가 인공지능 관련 이슈로 선정되었다. 마지막으로 2017년 10월 31일에는 세종대에서 열린 ‘인간 vs 인공지능’의 스타크래프트의 대결로 송병구 선수와 국내외 총 4개의 인공지능 프로그램이 맞붙은 경기의 뉴스가 주를 이루며 이벤트 기간으로 선정되었다.

반기 별 상위 4개의 감성 변화는 다음 그림과 같다(<Figure 13> 참고). 반기 별 상위 4개의 감성 변화를 보면 키워드별 감성뿐 아니라 전체 감성이 ‘행복’이 1위인 것을 볼 수 있다. 그러나 2016년 하반기에서 2017년 상반기를 볼 때 인공지능에 대한 여론 중 ‘행복’에 대한 수치가 낮아짐을 볼 수 있다.



<Figure 13> Ratio of the top four sentiment by half

5. 결론 및 추후 연구 방안

본 연구는 인공지능 관련된 사회적 여론을 파악하기 위하여 국내 유명 포털 사이트에서 인공지능과 관련된 뉴스의 수와 뉴스 제목, 댓글을 웹 크롤링하여 감성 분석을 실시하였다. 분석 결과 대부분의 이벤트 기간에 대하여 1위 감성은 ‘행복’으로 확인되었다. 또한 일반적으로 연속적인 양방향성 대화를 하는 채팅과 다른 특성을 가진 댓글은 ‘슬픔’과 ‘두려움’ 같은 감성의 빈도가 낮은 것을 볼 수 있었다. 이는 댓글 자체가 풍자나 냉소의 역할을 함으로써 나타나는 현상이라고 판단 된다(Jeon, 2007).

이세돌 기사와 알파고의 대국 이전 시점과 대국을 진행하는 시점의 두 결과의 특징을 살펴보면 대국 이전에는 ‘슬픔’, ‘두려움’을 제외한 대부분의 감성이 고르게 분포한 것을 볼 수 있었으나 대국을 진행하면서 ‘행복’이 증가하고 인간을 뛰어넘는 대국이 보여지면서 이에 대한 ‘분노’의 수치가 높아짐을 볼 수 있었다. 이벤트별 감성 중 빅데이터 관련 이벤트 기간인 2016년 12월 29일은 지난 2014년 인공지능 시스템을 도입하여 구제역 확산을 막으려고 했으나 문제점이 있었다는 기사들을 바탕으로 이에 대한 ‘분노’의 감성과 지난 정부의 비난 여론이 강하게 나타난 것을 확인 할 수 있었다.

국내 최초로 인공지능 스피커 출시를 발표한 SK텔레콤의 인공지능 음성인식 디바이스 ‘NUGU’와 같이 신규로 시장에 진입한 경우 이에 대한 기대를 나타내는 댓글이 주를 이루었으나 이후 후발 주자로 2017년 상반기에 출시를 발표한 KT의 ‘기가 지니’의 경우에는 새로운 제품에 대한 호기심이나 신기함보다는 제품의 기능상의 문제점을 짚거나 비하, 풍자하는 글들을 찾아볼 수

있었다. 같은 카테고리의 제품이더라도 이미 기능상의 문제에 대하여 여론이 형성되어 있어서 앞서 출시한 SK텔레콤의 ‘NUGU’와 비교해 보았을 때 ‘행복’의 비율은 낮아지고 ‘혐오’의 비율이 높아짐을 확인 할 수 있었다.

본 연구는 4차 산업 혁명으로 크게 대두되는 인공지능과 관련된 사회적 여론 파악을 위해 이벤트 기간 및 키워드를 추출하고 해당 기간의 뉴스 댓글을 수집하였다. 사회적 여론을 파악하기 위한 방법으로 댓글을 활용한 분석은 이에 대한 보다 세밀한 구분으로 나누었을 때 사회적 여론을 확실하게 파악 할 수 있다고 판단된다. 위와 같은 이유로 사회적 여론을 세밀하게 관찰하기 위하여 긍정, 부정이 아닌 7가지 다 범주 감성 분석을 실시하였다. 또한, 본 연구는 ‘인공지능’의 키워드로 사회적 여론을 파악하였지만 새로운 이슈에 대하여 여론 파악을 하고자 한다면 ‘인공지능’이 아닌 다른 주제어를 선정하여 사회적 여론을 파악 할 수 있을 것이라고 생각 된다.

앞에서 제시한 방법론은 기존의 한 주제에 대한 여론 분석에서 더 나아가 특정 주제에 대하여 시간에 따른 사회적 여론의 변화를 파악하고 주제에 대한 사건 식별 및 사건별 감성 분석을 하였다는 것에 학술적 의의가 있다. 또한 기존 긍정-부정의 일차원적 여론 분석이 아닌 7가지 다 범주의 감성을 통해 사회적 여론을 보다 세밀하게 파악하였다는 것에 의미가 있다. 이러한 여론 파악 방법은 특정 대상에 대하여 사회적 여론을 파악하고 기업이나 정부가 정책적 방향을 잡거나 새로운 의사결정을 함에 있어서 여론에 대한 동향 파악을 보다 세밀하게 할 수 있을 것이라고 판단된다.

본 연구 역시 몇 가지 한계점을 지니고 있다. 뉴스 댓글의 감성 분석을 실시하여 국내 여론을

파악하였으나 뉴스의 각 댓글에는 공감과 비공감 표시가 존재한다. 어떠한 댓글은 많은 공감을 얻어 대중들에게 동의를 얻는 댓글도 존재하는 반면 비공감이 압도적으로 많은 댓글도 존재하고 공감이나 비공감을 받지 못한 댓글도 존재한다. 이러한 공감 및 비공감이 존재하는 댓글들을 각 뉴스 시청의 수와 댓글의 공감, 비공감도에 관련된 지수를 만들어 댓글의 감성 분석을 한다면 보다 정확하게 사회적 여론을 파악할 수 있을 것이라고 판단된다. 또한 본 연구에는 인공지능 관련 댓글을 분석하여 ‘드루킹’ 사건과 같이 매크로 프로그램으로 인한 여론 조작이 일어나지 않았다. 추후, 정치적 주제에 대한 여론 파악을 한다면 이에 대한 대응 방안 마련을 통해 정확한 사회적 여론을 파악할 수 있을 것이다. 마지막으로 본 연구에 사용된 감성 분석기는 인터넷 댓글 전용의 감성 분석기를 사용한 것이 아니라 범용 감성 분석기를 사용하여 풍자의 댓글 역시 ‘행복’으로 판단하는 경향을 보였다. 또한 인터넷 댓글은 연속적으로 양방향성 대화를 주고받지 않는 경우가 많기 때문에 향후에는 댓글 전용의 훈련 세트를 활용한다면 더 정확한 여론 파악이 가능할 것이라고 생각된다.

참고문헌(References)

Byeong-Cheol Jeon, "A Study on the Form of Abusive Words in the Reply," *Studies in Linguistics*, No. 11(2007), 189~205.

Chang Ho Park, "An Analysis on the Pros and Cons Presented in the Responses of Replies as Intermediated Interaction of the Internet," *Discourse, The Korean Association of*

Socio-Historical Studies, Vol.16, No.2(2013), 135-164.

Dong Sung Kim, Kitae Kim, and Jong Woo Kim, "Character-based multi-category sentiment analysis on social media using deep learning algorithms," *Korean Institute of Industrial Engineers Spring Joint Conference Journal Issue*, (2017), 5082~5084.

Ekman, Paul, Friesen, Wallace V. , M. O'sullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider, R. Krause, W. A. LeCompte, T. Pitcaim, P. E. Ricci-Bitti, K. Scherer, M. Tomita, and A. Tzavaras, "Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion," *Journal of personality and social psychology*, Vol.53, No.4(1987), 712~717.

Eun-Ju Lee and Yoon Jae Jang, "Effects of Others' Comments on Internet News Sites on Perceptions of Reality : Perceived Public Opinion, Presumed Media Influence, and Self-Opinion," *Korean Journal of Journalism & Communication Studies*, Vol. 53, No. 4 (2009), 50-71.

Eun-Mee Kim and Yu-Hwa Sun, "The Effect of Replies in Internet News on the Audience," *Korean Journal of Journalism & Communication Studies*, Vol. 50, No. 4 (2006), 33-64.

Hye Seung Yang, "The Effects of the Opinion and Quality of User Postings on Internet News Readers' Attitude Toward the News Issue," *Korean Journal of Journalism & Communication Studies*, Vol. 52, No. 2 (2008), 25-281.

Hyemi Kim and June Woong Rhee, "The Impacts of Frame Fusion of News Text and Comments on News Interpretation and

- Attitude”, *Korean Journal of Journalism & Communication Studies*, Vol.55, No.2(2011), 32-55.
- Ir-Kwon Jeong and Young-Seok Kim, "Impact of "Datgeul" of Online Media on Public Opinion : An Examination of Perception of Public Opinion and Third Person Effect," *Korean Journal of Journalism & Communication Studies*, Vol. 50, No. 4(2006), 302-327.
- Jaeho Han, Woojeong Kim and Kyungsik Han, "Development of a Multi-Emotional Model for Online News Comments and Its Application," *Human Computer Interaction Korea Conference*, (2018), 893~897.
- Kang Jae-Won and Kim Sun-Ja, "A Study on the Effect of Comments Posted under Internet News Articles : In Consideration of the Degree of Involvement in Issues and Whether or not Opinions are in Accord," *Korean Journal of Journalism & Communication Studies*, Vol. 56, No. 2(2012), 143-166.
- Kim, D. S., and J. W. Kim, "Public Opinion Sensing and Trend Analysis on Social Media: A Study on Nuclear Power on Twitter”, *International Journal of Multimedia and Ubiquitous Engineering*, (2014), 373-384.
- Kim, Daewon and Kim, Beom-jin, "How AlphaGo does Change People’s Perception of Introduction of Artificial Intelligence into Intellectual Work?," *Journal of Cybercommunication Academic Society*, Vol. 33, No. 4(2016), 107-158.
- Moonjung Koh, "Who comments on or shares online news stories? : Application of statistical learning methodology," *Journal of Cybercommunication Academic Society*, Vol. 35, No. 1(2018), 5-51.
- S.J.You, “4th Industrial Revolution and Artificial Intelligence,” *Journal of Korea Multimedia Society*, Vol.21, No.4(2017), 1-8.
- Seo In Kim, Dong Sung Kim and Jong Woo Kim, "Public Sentiment Analysis of Korean Top-10 Companies : Big Data Approach Using Multi-categorical Sentiment Lexicon," *Journal of Intelligence and Information Systems*, Vol. 22, No. 3(2016), 45~69.
- Top Websites Ranking, Available at <https://www.similarweb.com/top-websites/korea--republic-of> (Accessed May 15, 2018).
- Won dong-kyu and Sang Pil Lee, "Implication of Artificial Intelligence and 4th Industrial Revolution,” *Industrial Engineering Magazine*, Vol. 23, No. 2(2016), 13-22.
- Young-Im Cho, “Understanding Big Data and main issues,” *Journal of Korean Association for Regional Information Society*, Vol.16, No.3(2013), 43-65.

Abstract

Multi-Category Sentiment Analysis for Social Opinion Related to Artificial Intelligence on Social Media*

Sang Won Lee** · Chang Wook Choi** · Dong Sung Kim*** ·
Woon Young Yeo** · Jong Woo Kim****

As AI (Artificial Intelligence) technologies have been swiftly evolved, a lot of products and services are under development in various fields for better users' experience. On this technology advance, negative effects of AI technologies also have been discussed actively while there exists positive expectation on them at the same time. For instance, many social issues such as trolley dilemma and system security issues are being debated, whereas autonomous vehicles based on artificial intelligence have had attention in terms of stability increase. Therefore, it needs to check and analyse major social issues on artificial intelligence for their development and societal acceptance. In this paper, multi-categorical sentiment analysis is conducted over online public opinion on artificial intelligence after identifying the trending topics related to artificial intelligence for two years from January 2016 to December 2017, which include the event, match between Lee Sedol and AlphaGo. Using the largest web portal in South Korea, online news, news headlines and news comments were crawled. Considering the importance of trending topics, online public opinion was analysed into seven multiple sentimental categories comprised of anger, dislike, fear, happiness, neutrality, sadness, and surprise by topics, not only two simple positive or negative sentiment. As a result, it was found that the top sentiment is "happiness" in most events and yet sentiments on each keyword are different. In addition, when the research period was divided into four periods, the first half of 2016, the second half of the year, the first half of 2017, and the second half of the year, it is confirmed that the sentiment of 'anger' decreases as goes by time. Based on the results of this analysis, it is possible to grasp various topics and trends currently discussed on artificial intelligence, and it can be used to prepare

* This work was supported by the Ministry of Education of Korea and the National Research Foundation of Korea in 2017(NRF-2017S1A3A2066740).

** Dept. of Business Informatics, Graduate School, Hanyang University

*** Dept. of Business Administration, Graduate School, Hanyang University

**** Corresponding Author: Jong Woo Kim

School of Business, Hanyang University

222 Wangsimni-ro, Seongdong-gu, Seoul 133-791, Korea

Tel: +82-2-2220-1067, Fax: +82-2-2220-1169, E-mail: kjw@hanyang.ac.kr

countermeasures. We hope that we can improve to measure public opinion more precisely in the future by integrating empathy level of news comments.

Key Words : Artificial Intelligence, Sentimental Analysis, Social Opinion, Online News, Online comments

Received : May 29, 2018 Revised : September 9, 2018 Accepted : October 24, 2018

Publication Type : Conference(Fast-track) Corresponding Author : Jong Woo Kim

저 자 소개



이상원

현재 한양대학교 일반대학원 비즈니스인포매틱스학과 석사과정으로 재학 중이다. 서경대학교에서 산업공학, 전자공학(복수) 전공으로 공학사 학위를 취득하였다. 주요 연구 관심분야는 데이터 마이닝 기법과 응용, 오피니언 마이닝, 자연어 처리, 딥 러닝 기법의 활용 등이다.



최창욱

현재 한양대학교 일반대학원 비즈니스인포매틱스학과 석사과정으로 재학 중이다. 한양대학교에서 경영학 학사 학위를 취득하였다. 주요 관심분야는 데이터마이닝, 기계학습, 딥러닝, 등이다.



김동성

현재 한양대학교 일반대학원 경영학과에서 경영정보시스템 전공으로 박사과정에 재학 중이다. 동 대학원에서 경영정보시스템 전공으로 석사학위를 취득하였다. 주요 연구 관심분야는 데이터 마이닝 기법과 응용, 기계 학습, 오피니언 마이닝, 사회 네트워크 분석, 딥 러닝 기법의 활용 등이다.



여운영

현재 한양대학교 일반대학원 비즈니스인포매틱스학과 석사과정 중이다. 한국해양대학교에서 데이터정보학과에서 이학사를 취득하였고, 주요 관심분야는 데이터마이닝, 기계 학습, 딥러닝, 자연어처리 등이다.



김종우

현재 한양대학교 경영대학 경영학부 교수로 재직 중이다. 서울대학교 수학과에서 학사를 마쳤으며, 한국과학기술원에서 경영과학으로 석사학위를, 산업경영학으로 박사학위를 취득하였다. 주요 연구 관심분야는 데이터마이닝 기법과 응용, 기계학습과 딥러닝, 오피니언 마이닝, 상품추천기술, 지능형 정보시스템, 집단지성, 사회 네트워크 분석, 클라우드 컴퓨팅 서비스 등이다.