

온라인 뉴스 베스트 댓글의 특성 분석

김진우¹ · 조혜인² · 이봉규^{3*}¹연세대학교 정보대학원 ICT·콘텐츠트랙 석사과정²연세대학교 정보대학원 ICT·콘텐츠트랙 석사과정³연세대학교 정보대학원 ICT·콘텐츠트랙 교수

Analyzing the Characteristics of Online News Best Comments

Jin Woo Kim¹ · Hye In Jo² · Bong Gyou Lee^{3*}^{1, 2, 3}Graduate School of Information, Yonsei University, Seoul, South Korea

[요 약]

온라인 뉴스에 개인의 참여가 활성화되면서 댓글의 중요성은 더욱 커지고 있다. 특히 이용자들에게 많은 공감을 받는 '베스트 댓글'은 주된 여론으로 인식되고 큰 영향력을 가진다. 따라서 본 연구는 온라인 뉴스 댓글 데이터를 이용하여 베스트 댓글의 특성을 알아보고자 하였다. 이를 위해 일반 댓글과 차이점을 보일 가능성이 있는 요소를 설정 후, 데이터를 수치화하여 일반 댓글과 베스트 댓글의 차이를 분석하였다. 본 연구는 최근 댓글 조작 등의 문제 해결에 실마리를 제공하고 개인 및 학술단체, 정부기관 등을 주제로 하여금 기초 자료로 활용될 것으로 기대된다.

[Abstract]

The importance of comments is constantly growing as a participation of individual in Online News that being invigorated. The 'Best Comments', which strongly related by major participants are recognized as a primary public opinion, and obtains the power. Thus this study is aimed to analyze the characteristics of the 'Best Comments' by utilizing the data of comments on Online News. For this study, a possible element that may reveal the difference between 'general' comments and 'best' comments were set up, digitalized the data, and examined the difference between 'general' and 'best' comments. This study is expected to provide a clue for the problematic issues, such as 'online comment rigged scandal' in recent; also as a basic data that subjected by the individual, academic society, government, and etc.

색인어 : 온라인뉴스, 댓글, 베스트 댓글, 내용분석

Key word : Online News, Comments, Best Comments, Content Analysis

<http://dx.doi.org/10.9728/dcs.2018.19.8.1489>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 25 July 2018; Revised 20 August 2018

Accepted 28 August 2018

*Corresponding Author; Bong Gyou Lee

Tel: +82-2-2123-6524

E-mail: jin.woo.kim@yonsei.ac.kr

I. 서론

과거의 대표적인 정보 전달 매체였던 TV 또는 신문은 단방향적인 정보 전달 기능을 가지고 있었다. 하지만 인터넷은 이용 주체들이 서로 상호작용하며 양방향적 정보를 전달하는 기능을 가지고 있으며, 이는 인터넷이 기존의 정보를 전달하는 매체들과 구별되는 특징 중 하나이다[1]. 그 중 댓글은 이용자들 간에 즉각적인 양방향 소통을 가능하게 하는 중요한 요소이다. 댓글이란, 인터넷 이용자들이 물리적 장소와 시간에 구애 받지 않고 가상의 토론 공간에서 자유롭게 인터넷에 접속해서 정보와 의견을 교환하고 의사소통을 할 수 있는 수단이다[2]. 최근 댓글은 네티즌들의 토론과 지식 공유 수단으로 자리 잡음으로써 '댓글 저널리즘¹⁾'이라는 새로운 문화 트렌드를 형성했다[3]. 특히 온라인 뉴스 사이트의 회원가입을 하지 않아도 개인의 SNS 계정으로 댓글을 작성할 수 있는 소셜 댓글 서비스가 활성화되면서 뉴스 댓글 작성이 용이해졌다. 이로 인해 개인의 온라인 뉴스 댓글 참여가 더욱 확대되었고, 개인의 실시간 참여가 증가하는 등 다양한 여론 형성에 크게 기여하고 있다. 하지만 이러한 개인의 참여 확대와 온라인 뉴스 댓글의 증가는 정보 교류의 긍정적인 영향 뿐 아니라, 익명성을 악용한 무분별한 비난, 욕설, 인신공격 등의 악성 댓글 증가로 댓글문화 형성에 악영향을 주고 있다. 많은 언론사와 포털은 악성 댓글, 트롤, 여론조작의 성격을 가진 기사 댓글들로 인해 댓글 란을 유지하고 관리하는데 어려움을 겪고 있다. 미국 공영 라디오(National Public Radio)의 경우 초기에는 청취자들이 댓글 란을 통해 건전한 토론을 할 것이라 예상하였으나 약 0.06%의 일부 사용자가 대량의 댓글을 올리는 등 부정적인 결과로 나타나 관리자 기능을 강화하여 댓글 란을 유지시키려 노력하였지만 결국 댓글 란을 폐지시켰다[4]. 국내는 최근 포털 뉴스의 댓글 조작 및 여론 조작에 대해 문제가 더욱 심각해지고 있다. 온라인 뉴스 댓글 참여가 가장 활성화 되어 있는 네이버의 경우, 기사 댓글에 누를 수 있는 '공감 및 비공감' 횟수를 계정(아이디)당 24시간 기준 50개로 제한하고, 연속해서 댓글을 작성할 때 시간 간격을 60초로 늘리는 등 긍정적인 댓글활동을 장려하기 위해 노력했다. 하지만 2018년 4월 13일 파워블로거 '드루킹'을 비롯한 더불어민주당 당원들이 "네이버 뉴스에서 댓글로 여론 조작을 시도"한 혐의로 검찰에 구속된 사건을 화두로 이러한 조치는 임시방편이라는 지적이 나오고 있다. 댓글 통계 분석 사이트인 위드미터에 따르면 2017년 10월부터 2018년 4월까지 네이버의 댓글 작성자, 작성 글, 공감 수 등을 분석한 결과 댓글 상위 작성자 100명이 작성한 댓글의 수가 23만 487건에 달한다고 하였다.

1) 저널리즘이란, 매스 미디어를 통해 공공적인 사실이나 사건에 관한 정보를 보도하고 논평하는 활동으로, 본 연구에서는 댓글이 온라인 기사가 제공하는 정보 이외의 새로운 정보를 제공한다는 점에서 '댓글 저널리즘'이라는 표현을 인용하였다.

시간 순으로 등록되는 일반적인 댓글과 달리 베스트 댓글은 상위에 게시됨으로써 비교적 영향력이 크다[5]. 소수의 이용자들이 작성한 글이지만 이용자들에게 많은 공감을 받는 '베스트 댓글'은 주된 여론으로 인식이 되며 글을 읽는 이용자들에게 큰 영향을 준다. 특히, 이용자는 관심과 배경지식이 적은 분야에서 댓글에 강하게 동조하는 경향이 있으며 깊이 고민하지 않고 타당성이 낮은 댓글조차도 인식에 영향을 받는다[6][7]. 이와 같은 베스트댓글의 중요성에도 불구하고 인터넷 뉴스, 댓글, SNS 연구에 비해 베스트 댓글에 관한 연구는 아직 초기단계이다[5][8]. 베스트 댓글의 영향력을 검증하기 위한 연구가 진행되고 있지만, 어떠한 댓글이 베스트 댓글로 채택되는지, 베스트 댓글만의 특징은 무엇인지와 같이 베스트 댓글의 특성을 규명하기 위한 논의는 충분히 이루어지지 못하였다. 따라서 본 연구는 인터넷 뉴스 댓글을 대상으로 내용분석방법을 이용하여 베스트 댓글이 가지는 특성을 알아보고자 한다.

본 연구의 구성은 다음과 같다. 1장은 연구배경 및 연구목적을 설명하고, 2장에서는 이론적 배경이 되는 온라인 뉴스 댓글 관련 선행연구를 분석한다. 이를 바탕으로 3장에서 연구가설을 제시한 후, 4장에서 연구방법을 설명하고, 5장에서 분석 결과를 제시한다. 마지막으로 6장에서는 결론 및 시사점, 연구의 한계점에 대하여 서술한다.

II. 이론적 배경

온라인뉴스에서 댓글은 온라인상의 이용자들에게 공감과 반론 등의 의견을 자유로이 표현할 수 있는 대표적인 도구이다. 이용자들이 남기는 댓글들은 1~2개 문장 정도의 길지 않은 글이지만 다수의 댓글이 모이게 되면 큰 영향력을 발휘할 수 있다[10]. 댓글에 대한 연구는 개별적으로 이루어지기보다 온라인 뉴스의 연구의 일부분으로서 온라인상에서 여론이 어떻게 형성되었는지, 어떻게 이용자들에게 확산되었는지에 대하여 주로 연구되어졌다[10]. 하지만 소셜 댓글 서비스 활성화를 통해 댓글 작성이 확대됨에 따라 온라인 뉴스 댓글에 대한 연구도 증가하는 추세이다. 온라인 뉴스 댓글에 대한 연구는 크게 댓글의 영향을 실증 분석한 양적 연구와, 내용분석 방법을 통해 이를 특성별로 분류하는 질적 연구로 나누어진다. 내용분석을 통한 질적 연구로는 포털 뉴스기사의 장르와 기사에 사용된 프레임에 따라 전반적으로 정치기사가 다른 기사들에 비해 이성적 댓글이 적고 욕설 댓글이 다수이며, 포털뉴스, 소셜미디어 등 댓글이 작성되는 게시글에 따라 댓글의 어조, 댓글의 길이, 댓글의 품질 등 다양한 요소에서 차이가 존재하였다[11][12]. 댓글의 영향을 실증 분석한 연구를 통해 인터넷 뉴스 댓글은 전혀 근거가 없고 타당하지 않은 경우에도 이용자들은 댓글에 의해 크게 영향을 받는 것으로 나타났다, 특정 기사에 대한 온라인 댓글은 해당 기사뿐 아니라

다른 기사를 읽을 때에도 영향을 미쳤다. 또한 기사에 달린 댓글이 이용자의 견해와 다를 때 댓글의 영향력이 강하게 작용하고, 댓글의 논리성보다 다수의 의견을 통해 댓글이 조성하는 분위기가 이용자들의 의견 형성에 더 큰 영향을 준다 [6][7][13]. 이러한 특징은 인터넷 공간에서 일반적인 이용자들이 ‘댓글 쓰기’보다는 ‘댓글 읽기’ 활동을 더 많이 하기 때문이다. 또한 이용자들은 누가 의견을 제시했는지, 얼마나 많은 사람들이 읽었는지도 댓글을 받아들이는데 있어서 중요한 요소로 고려하며, 포털 및 검색 엔진에서 제시하는 정보의 순위에 의존적이다[14]. 실험을 통해 이용자들은 한 페이지에 제시되는 정보의 순위를 뒤집어 보여주어도 여전히 상위에 제시된 결과에 주목하였다[15][16]. 이는 일반적으로 상위 뉴스기사에 조회 수가 더 많은 현상과 베스트 댓글의 동조효과가 발생하는 현상 등 인터넷 상의 상단에 노출되는 정보들이 더 큰 영향력을 가진다는 것을 뒷받침한다. 특히 많은 사람들이 읽은 베스트 댓글의 경우 동조효과가 존재하고 베스트 댓글을 통한 온라인 여론 형성이 가능하며, 부정적인 베스트 댓글에 비하여 긍정적인 베스트댓글이 이용자들에게 더 강한 동조효과가 나타난다[8]. 또한 ‘다른 사람들이 옳다고 생각하면 아마도 옳을 것이다.’ 라는 발견적 사고도 발생한다 [17]. 베스트 댓글은 일반 댓글과 다르게 유희적으로 말하기, 과장하여 말하기, 은어로 말하기 등의 특징이 존재함을 알 수 있었다[18].

언론분야 뿐만 아니라 마케팅 분야에서도 베스트 댓글의 영향력에 대하여 연구가 이루어지고 있으며 베스트 댓글의 내용은 소비자들의 제품 품질 지각 과정에 영향을 주었다. 제품 간 품질 차이가 큰 경우에는 베스트 댓글의 영향력이 감소하고, 중립적인 의견과 생생한 구매후기를 전달하는 내용일 경우 더 많은 이용자들이 댓글을 공감하였다[5][19]. 베스트 댓글의 영향력을 검증하는 연구는 진행되고 있지만, 어떠한 댓글이 베스트 댓글로 채택되는지, 베스트 댓글만의 특징은 무엇인지와 같은 연구는 부족한 현실이다.

III. 연구가설 설정 및 측정방법

베스트 댓글이 가지고 있는 특성을 알아 보기위해 네이버 뉴스 댓글 데이터를 수치화하여 일반 댓글과 베스트 댓글의 차이를 분석하고자 한다. 이를 위해 5가지 연구가설을 설정하였다. 먼저, 댓글의 길이는 댓글을 구성하는 글자의 수로 정의하였다[20]. 일반적으로 글의 길이는 정보의 깊이와 내용을 판단하는 기본적인 지표로 작용 할 수 있는 만큼 댓글의 특징을 구분하는 중요한 요소 중 하나라고 할 수 있다. 일반 댓글의 경우 글자 수의 범위가 제각각이겠지만 많은 이에게 공감을 받은 베스트 댓글의 경우 일반 댓글에 비해 상대적으로 많은 글자 수로 작성되어 있을 것이라 예상했고, 이를 바탕으로 다음과 같은 가설을 설정하였다.

가설 1. 베스트 댓글은 일반 댓글에 비해 길이가 더 길다.

댓글의 태도는 기사나 기존의 댓글에 대해 긍정성 혹은 부정성을 표출하는 정도를 의미한다[9]. 선행연구에 따르면 온라인 뉴스에 달린 댓글 중 상당한 의견이 부정성을 표출하는 댓글인 것으로 나타났다[21]. 해당 척도는 긍정적인 언어와 부정적인 언어의 기준을 정해 형용사와 명사를 위주로 긍정적인 형태소의 수와 부정적인 형태소의 수를 세어 긍정적인 형태소 수에서 부정적인 형태소 수의 차이를 보았다. 사회 문제 뉴스 기사의 특성상 객관적이고 중립적인 내용 또는 긍정적인 내용보다 부정적인 댓글의 비율이 높을 것으로 예상하였다. 특히 대부분의 댓글이 부정성을 표출하는 만큼 많은 이들에게 공감 받는 베스트 댓글도 일반 댓글에 비해 부정적인 댓글의 비율이 높을 것이라 판단하여 다음과 같은 연구가설을 설정하였다.

가설 2. 베스트 댓글은 일반 댓글에 비해 부정적이다.

표 1. 변수의 조작적 정의 및 측정방법

Table 1. Operational Definition and Measurement of Variables

Variables	Operational Definition	Measurement	Reference
Number of Characters	The number of characters in the comment	The count of characters in the comment	H. Y. Lee et al. [22]
Attitude	Number of positive and negative words in comments	Subtraction of positive words and negative words in comment	H. Y. Lee et al. [22]; E. H. Yeo et al. [13]
Quality	The usefulness and readability of comments	Content Analysis	E. M. Kim et al. [9]; Lang et al. [23];
Article Relevance	Relation of comments and articles	Content Analysis	H. M. Kim et al. [21]
Written Time	Time the comment was written	Standardize to measure when comments were written	-

선행연구를 살펴보면 댓글의 품질은 다양한 방법으로 측정할 수 있다[9][23]. [표1]을 보면 TV 뉴스, 기사 그리고 온라인 커뮤니티 등 많은 매체들에서 각기 다른 요소를 측정할 것을 볼 수 있다. 본 연구에서는 선행연구를 바탕으로 온라인 뉴스 기사 댓글의 품질을 분석하기 위해 내용적 측면에서는 유용성을, 표현적 측면에서는 가독성을 바탕으로 측정하였다. 내용분석의 방법을 활용하여 유용성과 가독성을 동시에 고려하여 측정하였으며, 내용분석의 특성상 평가자의 생각, 의견이 반영될 수 있다는 왜곡 문제를 최대한 해결하기 위해서 5명의 석사 과정 이상의 연구원들을 통해 댓글에 유용한 정보가 있는지, 읽기 좋은지 등의 코딩 스킴을 구체적으로 정한 뒤, 각자 숫자로 평가한 후 평균값을 분석하여 객관성을 확보하려 노력하였다. 분석에 앞서 베스트 댓글이 일반 댓글에 비해 유용성과 가독성을 기준으로 측정할 댓글의 품질이 좋을 것으로 예상하여 다음과 같은 연구가설을 설정하였다.

가설 3. 베스트 댓글은 일반 댓글에 비해 품질이 높다.

기사 관련성은 댓글 내용이 기사와 연관되어 있는지를 의미한다. 온라인 뉴스 기사의 이용자가 댓글을 작성할 때, 기사의 내용이 반영되는가를 고려한 것이다[24]. 이를 통해 댓글 작성자가 댓글 작성 시, 기사의 내용을 얼마나 이해하고 반영하였는지, 작성자의 댓글에 기사의 내용을 어느 정도 고려하였는지를 판단 할 수 있다. 해당지표 또한 평가자 개인의 생각 및 의견이 반영될 수 있다는 주관적인 왜곡 문제가 존재해 5명의 석사 과정 이상의 연구원들을 통해 평가하였다. 해당 댓글이 기사와 연관이 있다고 판단되면 '1', 그렇지 않으면 '0'으로 코딩 후 연구원들의 평균값을 분석하여 객관성을 확보하려 노력했다. 따라서 본 연구에서는 일반적인 뉴스 댓글의 경우 기사 본문 내용과 상관없는 댓글을 작성 할 수 있다는 점을 근거로 고려하여 일반 댓글과 베스트 댓글간의 기사 관련성에 차이가 있을 것이라 예상하였고, 다음과 같은 연구가설을 도출하였다.

가설 4. 베스트 댓글은 일반 댓글에 비해 기사 관련성이 높다.

마지막으로, 본 연구에서는 기존연구와의 차별성을 위해 댓글이 작성된 시간이라는 변수를 고려하였다. 인터넷 뉴스의 가장 큰 특징 중 하나는 실시간성이다. 댓글의 작성된 시간이라는 요인은 작성된 댓글유형과 밀접한 관계가 있을 것으로 예상되어 진다. 베스트 댓글의 특성상 많은 사용자에게 공감

을 받아야하기 때문에 작성된 시간이 빠른 댓글의 경우 더 많은 사용자에게 노출되며 이에 본 연구에서는 많은 이들에게 공감을 받는 베스트 댓글의 경우 일반 댓글과 작성된 시간에 있어서 차이가 있을 것으로 예상하여 다음과 같은 연구가설을 도출하였다.

가설 5. 베스트 댓글은 일반 댓글에 비해 빠른 시간에 작성되었다.

IV. 연구방법

4-1 기사주제 선정 방법

최근 문제인 정부의 근로시간 52시간으로 단축 정책, 최저임금 인상 정책과 함께 청년 실업률, 청년 실업, 취업난에 대한 기사가 많이 게시되고 있다. 통계청에 의하면 15세~29세의 청년실업률은 약 10.5%로 5월 기준 2,000년 이후 관련 통계 최고치를 기록했다. 우리나라 청년실업자 비중은 23.3%로 OECD 국가 중 가장 높은 것으로 나타났다. OECD국가 중 청년실업 수치가 20%를 넘는 나라는 우리나라가 유일하다. 사회문화적 측면에서 살펴보면 청년들에게는 실업에서 벗어나고자 스펙을 쌓는 것이 가장 큰 과제가 되었고, 과거에는 당연했던 취직-결혼-출산이라는 필수 과정이 선택의 영역으로 전환되었다[25]. 이와 같이 현재 '청년실업' 문제는 지속적으로 사람들의 관심을 받고 있다. 정부는 청년들의 불안정성 문제를 해결하기 위해 주로 단기고용을 제공하는 인턴제 등의 고용정책에만 집중해왔다[26]. 정부뿐만 아니라 다양한 학계에서도 청년실업 해결을 위한 연구가 진행되고 있다. 과거 정부에서 추진하는 청년실업 정책은 오히려 부작용을 낳고 있다는 점을 지적하며 기업 비용절감이 아닌 사회적 책임 분담 구조를 통해서 해결방안 제안하기도 하며, 과거와 현재의 한국 정부의 청년실업 정책을 비교해 청년실업 문제 해결을 위한 시사점을 도출, 1인 창조기업 창업 등 청년실업 문제 해결 방법, 청년들의 심리적 불안 상태 해결 방안 연구 등 주로 사회 현상에 관해 문제점을 분석하고 해결방안을 탐색하는 연구가 이루어졌다[25][27]. 이처럼 '청년실업' 주제는 사회문화, 정책적 측면 등 국내의 다양한 학문 영역에서 연구되고 있으며 본 연구는 자영업자, 구직자, 취직자 등의 관점에 따라 다양한 의견이 나타날 수 있다는 점에 주목하여 '청년실업' 기사를 대상으로 연구를 진행하였다.

표 2. 기사목록

Table 2. Article Lists

Article Title	Comment (Number, %)			
	General comment	Ratio (%)	Best Comment	Ratio (%)
'Work Hell' Liberal Arts Students, Engineering College Major Increase 9 Times, MONEYTODAY, 2018.04.16	576	98.29	10	0.90
One of the 7 Graduates of College Graduation "Experience of Suicidal Impulse by Employment Stress", Yonhapnews, 2018.05.20	650	98.48	10	1.71
High School Students Go to 'Employment Institute', Sedaily, 2018.05.22	1,103	99.10	10	1.51
Employment Rate Dramatic Drama. Korea's Elderly Population Ranked First in the OECD, Yonhapnews, 2018.06.02	1,182	99.16	10	0.33
Hungry Young People, Living in a Worrying Part-time Job life, The Kyunghyang Shinmun, 2018.06.03.	2,994	99.67	10	0.84
Total	6,505	99.23	50	0.76

4-2 데이터 수집 및 처리

댓글의 수집은 ‘크롤링’을 이용하였다. 최종 선정된 5개의 기사에 대한 댓글의 구체적인 데이터 수집과정은 다음과 같다. ‘Visual Studio 2017 version’을 이용하여 보다 쉽게 댓글을 크롤링하기 위해 C#의 WindowsFormApp 컨트롤을 활용하여 UI를 구성하였다. 네이버 뉴스 기사의 URL을 입력하면 .csv 파일로 댓글 내용 및 댓글 작성 시간, 댓글 작성자 등이 저장되게 코딩하였다.

본 연구의 수집된 대상은 청년실업을 주제로 선정된 네이버 뉴스 기사 5개에 달린 6,555개의 댓글이다. 네이버 뉴스의 경우, 댓글 정책에 따라 댓글을 보는 방법이 매우 다양하다. 공감 수에서 비공감 수를 뺀 수치가 많은 순으로 정렬되는 ‘순공감순’, 최신 작성 댓글이 앞 페이지에 위치하는 ‘최신순’, 공감 수와 비공감 수의 비율을 따졌을 때 공감 비율이 높은 순으로 정렬되는 ‘공감비율순’, 댓글에 답글이 많이 달린 순으로 정렬되는 ‘답글순’, 마지막으로 작성한 순서대로 정렬되는 ‘과거순’으로 네이버 뉴스 이용자들은 자신이 사용하고 싶은 방식을 통해 다른 사람들의 의견을 볼 수 있다. 이러한 기능들 중 공감 수에서 비공감수를 뺀 수치가 가장 많은 순으로 정렬되는 ‘순공감순’에서 상위 10개까지를 베스트 댓글 집단으로 보았다. 이외의 나머지 댓글은 일반 댓글 집단으로 분류하였다.

베스트 댓글은 각각 기사의 순공감순의 10개 댓글로 선정하였으며 총 50개를 사용하였고, 연구에 사용된 5개의 네이버 뉴스 기사는 2018년 4월 16일부터 2018년 6월 3일까지 게시된 것으로 하였다. 네이버의 댓글 정책상 정치기사의 경우 선거기간에는 댓글의 순공감순을 알 수 없어 베스트 댓글을 알 수 없고, 정치적인 논란이 될 만한 기사는 제외하여 선별하였다. 네이버의 경우, 포탈에서 키워드를 검색하여 뉴스

를 보면 기사가 아웃링크 방식으로 제공되고, 뉴스 페이지 내에서의 검색을 통해서 접근하면, 인링크 방식으로 뉴스를 볼 수 있게 되어있다. 아웃링크 방식의 뉴스 기사의 경우, 댓글이 많이 달리지 않는다는 한계점이 있어 뉴스 페이지에서의 검색이 가능한 주제로 선정해 인링크 방식의 온라인 기사 5개를 선별하였다. 선별한 5개의 기사는 [표2]와 같다. 본 연구는 일반 댓글과 베스트 댓글의 차이를 보기 위한 연구이기 때문에 분석을 위한 댓글이 최소 500개에서 최대 3,000개 이상까지 달린 기사를 위주로 하였다. 언론사 별로 댓글의 품질이 다를 가능성을 고려하여 특정 언론사만을 대상으로 하지 않고, 최대한 중복되지 않도록 기사를 수집하였다. 5개의 기사에서 댓글 작성자는 평균적으로 여자(18.029%)보다 남자(81.971%)가 많았고, 청년 실업 기사의 댓글의 특성 상 10대의 비율(1.878%)을 제외하고 20대(26.169%), 30대(35.925%), 40대(22.941%), 50대(12.885%)로 전 연령층에서 고르게 댓글을 작성한 것으로 나타났다.

V. 연구결과

5-1 상관관계 분석결과

본 연구에서는 가설검증을 위해 선정된 네이버 뉴스 기사 5개의 댓글을 수집하여 베스트 댓글과 일반 댓글로 나누었고, 5개의 가설 검증에 앞서 수치화된 데이터 간에 상관관계 분석을 실시하였다. 상관관계를 분석한 결과는 [표3]과 같다. 댓글의 길이와 댓글의 품질간의 상관관계수 값이 0.628로 측정되어 양의 상관관계가 있는 것으로 나타났지만 다른 측정 변수 간에는 상관관계가 없는 것으로 나타났다.

표 3. 댓글의 변수간의 상관관계 분석

Table 3. Correlation between Variables in Comments

		Number of Characters	Quality	Article Relevance	Attitude	Written Time
Correlation	Number of Characters	1.000				
	Quality	0.628	1.000			
	Article Relevance	0.280	0.373	1.000		
	Attitude	-0.276	-0.007	0.008	1.000	
	Written Time	0.047	0.108	0.078	0.039	1.000

표 4. t-test 분석결과

Table 4. Results of t-test Analysis

Variables	Group	Number	Average	Standard Deviation	F-value	p-value
Number of Characters	General Comment	6505	79.680	70.3290	9.069	0.006**
	Best Comment	50	115.200	87.5140		
Quality	General Comment	6505	1.2003	0.4327	22.431	0.000***
	Best Comment	50	1.6320	0.6251		
Article Relevance	General Comment	6505	0.3665	0.4584	10.172	0.000***
	Best Comment	50	0.6720	0.4233		
Attitude	General Comment	6505	-1.0806	1.5000	1.709	0.000***
	Best Comment	50	-1.8640	1.5843		
Written Time	General Comment	6505	0.0034	0.9457	8.152	0.000***
	Best Comment	50	-0.4447	0.1947		

*p<0.05, **p<0.01, ***p<0.001

5-2 가설검증 결과

본 연구는 가설 검증을 위하여 수집된 데이터를 이용하여 일반 댓글 그룹과 베스트 댓글 간의 통계적 차이가 있는지 SPSS Statistics의 independent samples t-test 분석기법을 활용하여 검증하였다. 각 변수에 대한 t-test 결과는 [표 4]와 같다. 일반댓글의 길이 평균은 79.680, 베스트 댓글의 길이평균 115.200로, 이 차이는 통계적으로 유의미하며 베스트 댓글의 길이가 일반 댓글이 비해 길 것이라는 가설 1이 채택되었다. 댓글의 품질의 경우, 일반댓글의 평균은 1.2003, 베스트 댓글의 평균은 1.6320로 나타나 베스트 댓글의 품질이 일반댓글에 비해 높을 것이라는 가설 2도 채택되었다. 댓글의 기사 관련성(일반댓글 평균=0.3665, 베스트댓글 평균=0.6720), 댓글의 태도(일반댓글 평균=-1.0806, 베스트댓글 평균=-1.8640), 댓글의 작성된 시간(일반댓글 평균=0.0034, 베스트댓글 평균=-0.4447) 모두 통계적으로 유의미한 것으로 나타났으며 모든 가설이 채택됨을 알 수 있었다.

VI. 결 론

6-1 연구결과 논의

본 연구는 네이버 뉴스 기사의 일반 댓글과 베스트 댓글 간의 비교분석을 통해 베스트 댓글이 가지고 있는 특징을 실증적으로 검증하였다. 상관관계 분석결과 변수들 간 댓글의 품질과 댓글의 길이의 항목에서만 양의 상관관계가 나타나는 것을 확인하였다. 댓글의 길이는 정보의 깊이와 내용을 판단하는 기본적인 지표로 작용 할 수 있는 만큼 유용성과 가독성을 중심으로 평가한 댓글의 품질과 양의 상관관계가 나타나는 것으로 해석되어진다.

연구 결과, 댓글의 길이(가설 1)는 정보의 깊이와 내용을 판단할 수 있는 기본 지표로 작용하는 만큼 베스트 댓글이 일반댓글과 비교하여 긴 댓글 길이를 가지는 것으로 나타났다. 댓글의 길이가 더 긴 데에는 다양한 이유가 있었지만 댓글의 품질과 댓글의 길이에 양의 상관관계를 갖는 것으로 살펴봐 아 더 많은 정보를 제공하는 글에 이용자들은 더 많은 공감을 하는 것으로 해석할 수 있다.

댓글의 품질 차이 검증 결과(가설 2), 내용적 측면에서의

유용성, 표현적 측면에서의 가독성이라는 요소가 베스트 댓글의 채택에 중요한 요소임이 확인되었다. 베스트 댓글은 많은 이들에게 공감을 받는 만큼 글의 품질이 보장되지 않는 경우 채택되기 힘든 것을 알 수 있었으며, 대부분의 베스트 댓글은 품질이 평균 이상 요구되는 것을 확인 할 수 있었다.

기사 관련성(가설 3)의 경우 베스트 댓글이 일반 댓글에 비해 더 높은 기사 관련성을 나타나는 것을 확인 할 수 있었다. 대부분의 베스트 댓글은 일반 댓글에 비해 높은 기사 관련성과 더불어 기사에서 제공하지 않는 추가적인 정보를 제공하기도 하며 이용자들에게 다른 의견을 제공한다[28]. 하지만 ‘T-T 진짜 죽고싶어요ㅠㅠㅠㅠㅠㅠㅠㅠ 우울해서 잠도 안와요...’, ‘취업문제가 잘 해결되길.. 취업생들 힘내세요’ 등과 같이 기사 관련성이 떨어지더라도 이용자들에게 공감을 받을 수 있는 댓글은 베스트 댓글로 채택됨을 알 수 있었다.

댓글의 태도(가설 4)의 경우 일반 댓글에 비해 베스트 댓글이 더욱 많은 부정적인 형태소를 이용하는 것으로 나타났다. ‘쓰잘대 없는 대학보다 훨 낫다 기술배워 일찌기 취직하고 빚 없이 사는게 맞다 대학간다가 학자금 빌려도 졸업해서 취직 안되 금융노예로 사는것 보다 낫다’와 같이 부정적인 형태소를 긍정적인 형태소에 비해 많이 이용함으로써 베스트 댓글은 일반 댓글에 비해 더욱 자극적인 요소를 포함하며 많은 이들의 이목을 받는 것을 알 수 있었다. 가설 검증을 통해 부정적인 댓글이 긍정적인 댓글보다 영향력이 큰 것을 검증하였다.

마지막으로 댓글이 작성된 시간(가설 5)의 경우 베스트 댓글의 작성된 시간이 일반 댓글의 작성된 시간에 비해 빠른 것으로 나타났다. 특히 연구에 사용된 50개의 베스트 댓글 전부가 평균값보다 모두 빠르게 작성된 것으로 나타났으며 작성된 시간 항목의 경우 베스트 댓글로써의 채택에 중요한 요소인 것으로 확인되었다. 5개의 연구가설 검증결과, 일반 댓글과 베스트 댓글의 다섯 가지 변수 중 댓글의 길이, 댓글의 태도, 댓글의 품질, 기사 관련성, 댓글의 작성시간 모든 항목에 있어서 가설이 모두 채택되었다.

6-2 시사점

본 연구의 시사점은 다음과 같다. 첫째, 베스트 댓글과 일반 댓글의 차이를 내용분석을 했다는 점이다. 선행연구에서 보았듯, 베스트 댓글에 대한 연구는 부족한 실정이며, 내용분석 연구 역시 부족한 상황에서 베스트 댓글 및 내용 분석 연구를 수행하였다는 점에서 의미가 있다. 선행연구를 통해 베스트 댓글의 특성을 분석 할 변수를 구성하였으며 특히 기존 연구에서 고려하지 않았던 작성된 시간이라는 유의미한 변수를 도출하여 베스트 댓글과 일반 댓글 간에 유의미한 차이가 있음을 검증하였고, 향후 온라인 뉴스 기사 댓글과 관련된 연구의 발판이 될 수 있을 것으로 기대한다.

둘째, 실제 데이터를 활용하여 연구를 수행했다는 점에서 의미가 있다. 댓글 관련 선행 연구는 주로 설문조사를 이용해

댓글이 이용자의 인지에 미치는 영향을 실증 분석한 연구였다. 이러한 연구는 선별된 소수의 댓글을 이용하여 연구를 수행하였기 때문에 연구결과 일반화에 있어서 한계점이 존재하였다. 하지만 본 연구에서는 선정된 5개 기사의 모든 댓글 데이터를 수집하여 내용분석 함으로써 더욱 타당한 분석결과를 제시하여 기존 연구의 한계점을 보완하고자 하였다.

셋째, 감성사전을 이용한 데이터 분석이 아닌 6,555개의 댓글을 모두 읽어 적절한 코딩 스킴을 근거로 댓글의 품질과 기사 관련성을 분석하였다는 점에서 연구의 의미가 있다. 감성사전을 이용해 데이터를 분석하는 경우, 긍정적 또는 부정적 감성사전 각각이 포함하고 있지 않은 단어나 형태소, 은어, 이모티콘 등은 의미 있는 데이터에 포함되지 않는다는 한계가 있다. 예를 들어, 모음과 자음을 모두 분리하여 댓글을 작성한 경우, ‘헬조선’, ‘ㅈㅈ’, ‘개꿀’과 같은 시대 상황이 반영된 은어의 경우, 특정 커뮤니티 언어의 경우, 긍정적 언어로 부정적인 감정을 드러내는 비꼬는 어투의 경우 등을 모두 고려하여 감성사전의 한계를 극복하고자 하였다.

6-3 연구의 한계점 및 향후 연구방향

본 연구는 베스트 댓글의 특징을 도출하는 과정에 있어서 최대한 객관적으로 분석하고자 하였으나, 다음과 같은 연구의 한계점을 내포하고 있다. 첫째, 댓글의 내용 분석 중 긍정적, 부정적 형태소의 기준을 전부 구축할 수 없다는 점이다. 개개인마다 명사 또는 형태소에 대해 긍정적 또는 부정적이라고 생각하는 기준이 다를 뿐 아니라 가독성이나 유용성의 경우에도 사람마다 이해하는 것과 판단하는 것이 각각 다를 수밖에 없다. 또한, 댓글을 읽다보면 평가자 개인의 주관적인 의견이 반영될 가능성이 존재한다. 내용 분석의 코딩 스킴이 있다 하더라도 객관성이 부족하다고 지적 될 수 있으며, 이는 연구의 한계점으로 제시될 수 있다.

둘째, 사회적 이슈가 되는 뉴스기사만을 토대로 연구를 수행하였기에 타 분야의 베스트 댓글과 일반 댓글의 차이점에 대해서 대표성이 확보되었다고 일반화할 수 없다. 향후 연구에서는 사회 뉴스뿐만 아니라 환경, IT, 연예, 과학 분야 등 추가적인 분야의 데이터를 수집하여 주제별 베스트 댓글과 일반 댓글간의 차이를 연구할 필요성이 있다. 또한, 본 연구의 결과를 토대로 지속적인 관련 연구주제들이 과생될 것으로 기대된다. 기존에 진행되었던 온라인 뉴스 댓글 관련 연구뿐만 아니라 온라인 커뮤니티, SNS, 아웃링크 뉴스기사 연구 등에서도 작성된 시간이라는 새로운 변수의 활용과 실제 데이터를 이용했다는 점에서 후속연구의 가이드라인 역할을 수행하며, 향후 다양한 분야의 댓글 관련 연구의 시발점이 될 수 있기를 기대한다.

참 고 문 헌

- [1] J. S. Lee, "Research on Factors Affecting the Effects of Reading Replies", *Korean Journal of Communication & Information*, Vol. 42, pp. 249-279, 2008.
- [2] I. K. Jeong, "Impact of "Datgeul" of Online Media on Public Opinion : An Examination of Perception of Public Opinion and Third Person Effect", *Korean Society For Journalism And Communication Studies*, Vol. 49, No. 1, pp. 29-56, 2006.
- [3] I News, "Commentary Journalism' Comes"[Internet], Available:
<https://news.naver.com/main/read.nhn?mode=LSD&mid=sec&sid1=105&oid=031&aid=0000022158>
- [4] Huffington, "Why did they remove the news comment box"[Internet], Available:
https://www.huffingtonpost.kr/entry/news_kr_5ae01a9be4b061c0bfa2dc46
- [5] I. S. Lee, "Effectiveness of Online Best Reply on Consumers' Perception of Product Quality Message Sidedness, Strength and Consumer Experience", *Advertising Research*, Vol. 110, pp. 60-83, 2016.
- [6] J. W. Kang, "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 Society For Journalism And Communication Studies*, Vol. 56, No. 2, pp. 143-166, 2012.
- [7] W. Y. Chun, "Effects of Online Comments on Judgments of a Politician: The Role of Correction Messages", *Korean Journal of Social and Personality Psychology*, Vol. 24, No. 2, pp. 133-150, 2010.
- [8] K. S. Moon, "An Effect of the Valence of Best Reply on the Conformity of General Reply", *Journal of The Korea Contents Association*, Vol. 13, No. 12, pp. 201-211, 2013.
- [9] E. M. Kim, "The Effect of Replies in Internet News on the Audience", *Korean Society For Journalism & Communication Studies*, Vol. 50, No. 4, pp. 33-64, 2006.
- [10] K. M. Kim "Online news diffusion and public opinion formation : A case study of 'controversy over Park Kyung Sin blog posting'", *Journal of Communication Science*, Vol. 12, No. 4, pp. 35-72, 2012.
- [11] J. S. Lee, "The Influence of the Frame and Genre of Portal News Articles on the Type of Online Replies", *Korean Journal of Journalism & Communication Studies*, Vol. 54, No. 2, pp. 116-137, 2010.
- [12] S. D. Kim, "A Comparative Analysis between General Comments and Social Comments on an Online News Site", *Journal of The Korea Contents Association*, Vol. 15, No. 4, pp. 391-406, 2015.
- [13] E. H. Yeo, "Influence of Online News Replies : Accessibility and Applicability", *Journalism & Communication*, Vol. 15, No. 2, pp. 237-259, 2011.
- [14] H. K. Park, "Public Perceptions of Public Social Workers in Comments of the Internet Media Discussion Rooms after Welfare Embezzlement Cases in 2008", *Korean Journal of Social Welfare*, Vol. 62, No. 1, pp. 391-415, 2010.
- [15] Pan, B, "In google we trust: Users' decisions on rank, position, and relevance", *Journal of Computer-Mediated Communication*, Vol. 12, pp. 801-823, 2007.
- [16] J. Y. Oh. "The Effects of Search Engine Credibility and Information Ranking on Search Behavior", *Korean Journal of Journalism & Communication Studies*, Vol. 53, No. 6, pp. 26-49, 2009.
- [17] M. N. Lee, "The False Consensus Effect of Internet Replies When Varying the Issue Position, Message Quantity, and Prior Attitude", *Journal of Cyber communication Academic Society*, Vol. 27, No. 3, pp. 47-87, 2010.
- [18] H. N. Choi, "The Dialog Strategies of Best Replies from the Perspective of Maxim Flouting", *Korean Literature*, Vol. 60, pp. 79-117, 2015.
- [19] D. H. Cheon, "A Study of effect on trust, acceptance and spread of word-of-mouth and purchase intention by characteristics of electronic word-of-mouth on travel industry", *The Academy of Customer Satisfaction Management*, Vol. 14, No. 1, pp. 83-100, 2012.
- [20] W. B. Ji. "A Study on Advertising Criticism of Netizen: A Focus on Comments on Commercials", *Journal of Practical Research in Advertising and Public Relations*, Vol. 5, No. 1, pp. 26-51, 2012.
- [21] K. N. Kim, "Analyzing Readers' Comments in Internet Newspaper Ohmynews", *Korean Journal of Broadcasting and Telecommunication Studies*, Vol. 19, No. 3, pp. 7-41, 2005.
- [22] H. J. Lee, "Factors Influencing Subscribers' Voluntary Payment Behavior on an Online News Site: Focusing on the Role of Appreciation", *Knowledge Management Society*, Vol. 14, No. 4, p1-17, 2013.
- [23] A. Lang, D. Potter, and M. E. Grabe, "Making News Memorable: Applying Theory to the Production of Local Television News," *Journal of Broadcasting & Electronic Media*, Vol.47, No.1, pp.113-123, 2003.
- [24] H. M. Kim, "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, pp. 32-55, 2011.

- [25] Y. J. Kim, “A Comparative Study on the Unemployment Policy of Korean Youth”, *The Korean Association For Public Administration*, Seoul, pp. 753-769, 2013.
- [26] S. Y. Lee, “Analysis of Precariousness in Korean Youth Labour Market”, *Journal of Critical Social Policy*, Vol. 54, pp. 487-521, 2017.
- [27] S. H. Kim, “Labor Market Restructuring of Park’s Government and the Revision of Employment Policies for Young”, *The Journal of Labor Studies*, Vol. 31, pp. 5-37, 2015.
- [28] J. S. Lee, “The Relationships Among Emotional Feelings Presented in Photo News, the Valence of Online Replies, Empathy, Attitudes, and Behavioral Intention ; Focusing on the Mediative Role of Empathy”, *Journal of Media Economics & Culture*, Vol. 9, No. 1, pp. 135-170, 2011.

김진우(Jin Woo Kim)



2017년 : 고려대학교 정보통계학과 (학사)

2017년~현 재: 연세대학교 정보대학원 ICT·콘텐츠트랙 (석사과정)
※관심분야 : 빅데이터 분석, 소셜 데이터, 미디어 정책 등



조혜인(Hye In Jo)

2018년 : 순천향대학교 컴퓨터공학과 (학사)

2018년~현 재: 연세대학교 정보대학원 ICT·콘텐츠트랙 (석사과정)
※관심분야 : 텍스트 마이닝, 데이터 아키텍처, 데이터 분석 등



이봉규(Bong Gyou Lee)

1988년 : 연세대학교 상경대학 경제학과 (학사)
1992년 : Cornell University, Dept. of CRP (MS)
1994년 : Cornell University, Dept. of CRP (Ph.D)

1997년~2005년: 한성대학교 공과대학 정보전산학부 교수
2008년~2011년: 연세대학교 정보대학원 부원장
2016년~2017년: 연세대학교 정보대학원 원장
2005년~현 재: 연세대학교 정보대학원 ICT·콘텐츠트랙 교수
2018년~현 재: 연세대학교 학술정보원 원장
※관심분야 : IT 정책·산업, 방송통신융합정책