

텍스트 마이닝 기법을 활용한 웹툰 댓글 분석 : 네이버 베스트 도전 웹툰을 중심으로

이윤주[○], 소현정^{*}, 곽기영^{*}, 안현철^{*}

[○]국민대학교 비즈니스IT전문대학원 비즈니스IT,

^{*}국민대학교 비즈니스IT전문대학원 비즈니스IT

e-mail: {nmmlas0ui, hyeon-jeongso, kykwahk, hcahn}@kookmin.ac.kr^{○*}

A Study of Webtoon comments using Text mining : Focusing on Naver's Best Challenge Webtoon

Yunju Lee^{*}, Hyeonjeong So^{*}, Kee-Young Kwahk^{*}, Hyunchul Ahn^{*}

[○]Graduate School of Business IT, Kookmin University,

^{*}Graduate School of Business IT, Kookmin University

● 요 약 ●

웹툰 시장의 성장에 따라, 웹툰을 주제로 다양한 연구가 진행되고 있다. 그러나 웹툰의 댓글을 분석한 연구는 특정 웹툰에 한정된 연구가 많아 일반적인 웹툰 독자의 댓글 특징을 보기에는 한계가 있다. 또한 웹툰의 흥행과 관련하여, 흥행에 성공한 웹툰과 그렇지 못한 웹툰의 독자 반응을 파악하는 연구는 부족한 실정이다. 따라서 본 연구는 웹툰의 흥행에 주목하여, 흥행의 지표를 웹툰 플랫폼 정식연재로 판단하고, 정식연재가 된 웹툰과 되지 못한 웹툰의 댓글을 비교 분석하였다. 분석 결과, 정식연재가 된 웹툰은 긍정적인 감상평 외에 2차적 저작물을 언급하고 등장인물의 이름 언급이 높았으나 정식연재가 되지 못한 웹툰은 부정적인 감상평 외에, 웹툰 요소에 대한 부정적 언급과 웹툰 장르와 다른 장르의 언급이 나타나 웹툰에 대한 독자의 불만족 요인을 파악할 수 있었다.

키워드: 텍스트 마이닝(Text mining), 웹툰(Webtoon), Word-of-Mouth(WoM)

1. 서론

현대 사회에서 제품 및 서비스를 이용하는 사용자들은 다방면으로 자신의 후기, 즉 리뷰, 댓글을 자유롭게 남길 수 있다. 이에 따라 기업들은 사용자의 리뷰를 이용하여 이익을 더 극대화하기 위하여 리뷰 데이터를 다양한 방법으로 분석하고 있다. 선행연구에서는 뉴스, 온라인소셜물, 영화 댓글 분석을 통하여 댓글의 중요성에 대한 연구가 이루어지고 있다. 댓글의 영향력을 볼 수 있는 연구는 대표적으로 3가지로 볼 수 있다. 첫째, 뉴스에서의 댓글은 사용자들의 의견표명을 통하여 여론의 방향을 알려주는 데 멈추는 것이 아니라, 뉴스 매체의 사회적 영향력에 관한 판단에도 영향이 미치는 것으로 나타났다[5]. 또한, 사용자들은 뉴스 기사의 내용보다는 다른 사람의 댓글에 시간을 많이 할애하는 성향을 보이기도 한다. 이에 따라 댓글에 대한 영향력은 크다고 할 수 있다[9]. 둘째, 온라인 시장에서 중요한 정보원으로 사용되고 있는 댓글은 특히, 화장품 제품 사용 후기에 대한 의존도가 높은 것으로 나타났다[3, 4]. 온라인 리뷰의 평가적인 정보 특성은 리뷰 수용자들에게 영향을 미치는 것으로 나타났다[6, 18, 19]. 셋째, 영화 리뷰를 통해 흥행성과의 관계를 나타냈다[14]. 문성민·하효지

(2015)는 영화를 보고 관객들이 느끼는 감정과 영화 흥행과의 관계를 규명하기 위해 영화 장르별로 감정 어휘의 분포를 탐색하고 영화의 흥행도와 감정 어휘의 관계를 파악하였다. 이러한 선행연구를 통해 리뷰 및 댓글의 중요성이 점점 커져가고 있음에 주목하여 본 연구는 네이버 웹툰 댓글을 통한 리뷰를 분석하고자 한다. 웹툰(Webtoon)은 PC 및 모바일에서 자유롭게 만화를 즐길 수 있는 콘텐츠 서비스를 의미하며, 현재 국내 만화 시장의 대표적인 장르이다[12]. 특히 한국의 웹툰 시장은 출판 만화의 온라인 배급을 위한 해외 시장과는 다르게, 온라인 환경에 맞추어 제작된 콘텐츠를 제공하고 있고, 서비스를 제공하는 플랫폼과 콘텐츠 창작자인 작가, 소비자 구조가 갖추어져 하나의 산업으로 성장하였다는 점에서 국내 고유의 콘텐츠로 평가되고 있다[16]. 국내 웹툰 시장규모는 매년 상승세를 기록하고 있으며, KT경제경영연구소(2015)에 따르면 2018년 8,805억에서 2020년 매출 1조 원을 기대하고 있다. 최근에는 웹툰 플랫폼이 해외진출에 성공하게 되면서 해외 디지털 만화 시장 선점에도 성과를 보이고 있다.

II. 이론적배경

웹툰 시장 성장에 따라, 웹툰을 주제로 다양한 연구가 진행되고 있다. 윤기현 외(2015)의 연구는 국내 웹툰의 전반적인 특성을 분석하였으며, 장현석·이건창(2018)은 웹툰이 콘텐츠 산업에서 베이스 캠프(Base Camp) 역할을 할 수 있다는 점에 주목하여, 웹툰의 모바일게임화 가능성을 예측하였다. 또한, 웹툰의 흥행을 예측하는 연구도 진행되었는데, 양지훈 외(2016)의 연구는 웹툰의 흥행을 결정하는 요인을 분석하고자 평점, 좋아요의 수 등 여러 정형 데이터를 활용하여 분석하였다. 최근에는 웹툰의 댓글에 주목하는 연구들이 증가하고 있다[2, 10, 11]. 웹툰의 댓글은 작가와 독자 사이의 의사소통 매개체가 되기 때문인데, 댓글을 통해 독자는 자신의 의견을 작가에게 전달할 수 있고, 작가는 독자들의 의견을 확인할 수 있다. 이러한 댓글의 특징에 주목하여, 이동균 외(2018)는 금연 홍보 웹툰의 댓글을 분석하여 독자들의 반응을 분석하였고, 이지현·황장선(2017)은 브랜드 웹툰의 댓글을 분석하여 브랜드 웹툰의 광고 효과를 파악하고자 하였다. 최기숙(2019)의 연구 역시 웹툰의 댓글을 분석하여 독자의 반응 및 소통 방식을 밝혔다. 선행 연구 중 웹툰과 관련된 연구는 증가하는 추세이지만, 웹툰의 댓글을 주제로 진행된 연구는 부족한 상황이다. 따라서 본 연구는 웹툰의 흥행에 주목하여 정식연재에 성공한 웹툰과 그렇지 못한 웹툰의 댓글을 비교 분석하고자 한다. 국내 웹툰 시장에서 가장 큰 점유율을 차지하는 네이버 웹툰(NAVER WEBTOON)은 정식 연재중인 웹툰 콘텐츠 외에, 누구나 제작이 가능한 창작만화(UCC)를 게시할 수 있는 ‘도전만화’ 서비스를 제공한다[13]. 그중 인기 작품은 ‘베스트 도전만화’로 게시되며, 독자에게 완성도와 인기를 인정받고, 네이버 자체 기준에 적합한 웹툰은 네이버 정식 웹툰으로 승격된다. 정식 웹툰으로 승격된 웹툰은 네이버 웹툰 플랫폼에서 연재되며, 공식적으로 5개 언어로 서비스되어 전 세계 독자에게 연재가 가능하다. 이에 따라, 본 연구는 웹툰의 흥행을 측정할 수 있는 지표를 정식연재 승격으로 판단하였다. 네이버 웹툰 플랫폼 특성상 조회수를 공개하지 않고, 각 웹툰 평점의 경우 소수점 단위로 비슷한 양상을 띄기 때문에 그 차이가 크지 않아서 흥행지표로 보기 어렵다는 점이 있다. 또한 평점이 높은 웹툰이라고 해서 독자의 수가 많다고 할 수 없으며, 그 웹툰이 인기가 많다고 보기 어려울 수 있다. 이러한 이유로 본 연구는 웹툰의 정식연재 여부를 흥행요인으로 판단하였으며 이를 승격으로 지칭하여 분석을 진행하였다.

III. 연구방법

본 연구에 사용된 데이터는 네이버 ‘베스트 도전만화’에서 마지막 연재 기준 2017년부터 2020년까지 연재된 웹툰 정식연재 승격된 웹툰 7편과 승격에 실패한 웹툰 7편을 두 그룹으로 나누어, 총 14편의 댓글을 파이썬으로 크롤링하여 수집하였다. 전체 댓글 수는 36,194개로, 웹툰 댓글을 텍스트 마이닝(text mining) 기법으로 분석하여 각 웹툰에서 보이는 독자의 반응 특징과 승격을 기준으로 나눈 두 그룹의 댓글 특징을 분석하여 독자의 반응 차이를 알아보려 한다.

Table 1. 분석에 사용된 웹툰과 댓글 수

승격여부	웹툰	댓글 수
승격된 웹툰	삶이 우리를 속일지라도	4,910
	아홉수 우리들	2,935
	혼밥 혼술 혼자 사는 이야기	2,090
	결혼생활 그림일기	5,255
	꼬맹이를 부탁해!	3,152
	취사병 전설이 되다	8,605
	어글리후드	1,556
승격이 되지 않은 웹툰	내 여동생은 귀여워	5,089
	스펙타클 내인생	770
	학살던전마	269
	그대 곁으로	214
	당할연	1,086
	우리들만의 싸이커	200
	키치 꼬치	63
총 웹툰		36,194

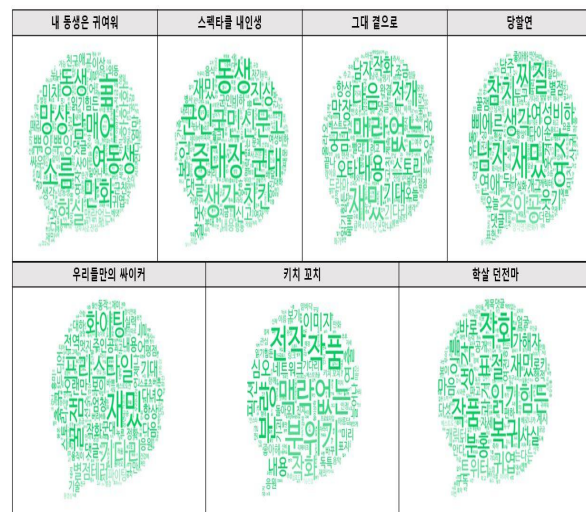


Fig. 1. 승격이 된 웹툰 Wordcloud

Fig. 1.을 보면, 승격에 성공한 웹툰의 경우 공통적인 특징으로 정식연재를 암시하거나 승격에 관련된 단어들 이 등장하였다. 또한 긍정적인 감상평도 많이 등장하였는데, 특히 ‘꿀잼’과 같은 신조어를 포함한 ‘공감’, ‘재밌’과 같은 단어가 등장하였다. 웹툰 ‘결혼생활 그림일기’의 경우 웹툰의 2차적 저작물이라 할 수 있는 이모티콘을 언급하여 웹툰이 콘텐츠의 베이스 캠프가 될 수 있음을 확인할 수 있었고, 웹툰 ‘삶이 우리를 속일지라도’, ‘꼬맹이를 부탁해’에서는 등장인물의 이름이 많이 언급되는 점을 보아 독자가 웹툰에 몰입감을 갖고 댓글을 작성하였음을 알 수 있다.

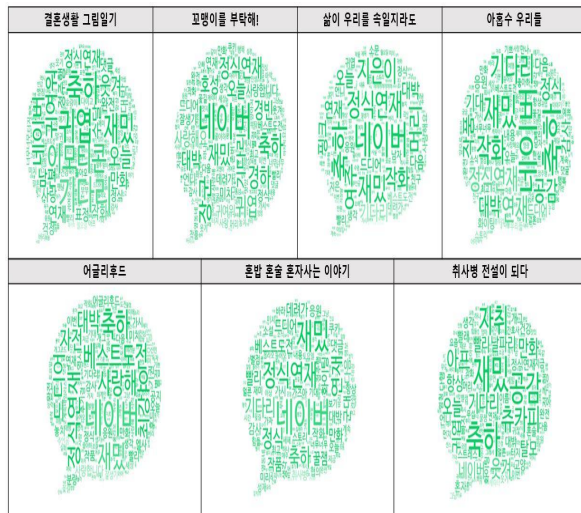


Fig. 2. 승격이 되지 않은 웹툰 Wordcloud

Fig. 1.을 보면, 승격에 성공한 웹툰의 경우 공통적인 특징으로 정식연재를 암시하거나 승격에 관련된 단어들이 등장하였다. 또한 긍정적인 감상평도 많이 등장하였는데, 특히 ‘꿀잼’과 같은 신조어를 포함한 ‘공감’, ‘재밌’과 같은 단어가 등장하였다. 웹툰 ‘결혼생활 그림일기’의 경우 웹툰의 2차적 저작물이라 할 수 있는 이모티콘을 언급하여 웹툰이 콘텐츠의 베이스 캠프가 될 수 있음을 확인할 수 있었고, 웹툰 ‘살이 우리를 속일지라도’, ‘꼬맹이를 부탁해!’에서는 등장인물의 이름이 많이 언급되는 점을 보아 독자가 웹툰에 몰입감을 갖고 댓글을 작성하였음을 알 수 있다. 전체 웹툰의 댓글을 TF-IDF 분석한 결과, 각 웹툰의 특징적인 키워드를 확인할 수 있었다. 승격에 성공한 웹툰의 댓글 분석 결과, 워드클라우드와 비슷한 결과를 보였다. 대부분의 웹툰이 등장인물 이름 언급 수가 높고, 웹툰을 소재로 한 2차 창작물에 대한 키워드가 등장하였다. 웹툰 ‘결혼생활 그림일기’와 ‘어글리후드’에서는 각각 ‘굿즈’, ‘공식굿즈’라는 단어가, 웹툰 ‘아홉수 우리들’은 ‘드라마화’, 라는 단어가 등장하였다. 반면 승격이 되지 않은 웹툰의 TF-IDF 분석 결과는, 첫 번째로 웹툰의 장르와 다른 장르를 나타내는 단어가 등장하는 것을 볼 수 있다. 웹툰 ‘내 여동생은 귀여워’의 경우, 일상을 그린 에피소드 장르의 웹툰이지만 ‘호러물’, ‘판타지’, ‘공포’ 라는 단어가 등장한다. 또한 독자의 비평이 나타나는 단어도 볼 수 있었는데, 웹툰 ‘그대결로’의 경우 ‘급한전개’라는 단어가, 웹툰 ‘당할연’의 경우 ‘시대착오’, ‘회풍’이라는 단어가 등장하였고, 웹툰 ‘키치 꼬치’의 경우 ‘징그럽’, ‘혼란’이 등장하였다. 또한 웹툰 ‘학살던전마’의 경우 TF-IDF 결과에서 ‘대추’, ‘양념장’과 같은 요리재료가 등장하였는데, 독자가 웹툰의 전개에 불만을 표시할 때 댓글에 웹툰에 대한 감상평이 아닌 레시피를 작성한다는 점에서 흥미로운 결과로 보인다.

결혼생활 그림일기	꼬맹이를 부탁해!	살이 우리를 속일지라도	아홉수 우리들	어글리후드	취사병 전설이 되다	혼밥 혼술 혼자 사는 이야기
다이고	장군	지은	봉우리	어글리	디어	츠크
호랑	경빈	지은이	차우리	기권	취사	츠크피
굿즈	호성	인공	공시	어글리후드	대장	탈모
달이	꼬맹이	추잡	승무원	추기경	성재	날파리
호랭이	홍삼	은사	연봉	미애	켈리	혼술
애용	말랑	지은아	하핑	네임리스	재밌습니	간호사
이비인후과	꼬부	남준일	김우리	뚜벅	담당	무도
라이거	이한	용도	수박	로그	장이	개그우먼
알아먹	꼬부해	스토킹	청혼	주교	계급	고양시
장모	꽃잎	권지은	퇴사	뚜벅뚜벅	해안	맥주

Fig. 3. 승격이 된 웹툰 TF-IDF

그대결로	내 여동생은 귀여워	당할연	스펙타클 내인생	우리들만의 싸이커	키치 꼬치	학살던전마
죽사	근친	빠에르	신문고	프리스타일	심오	로키
형사	호러물	다이슨	민원	축구	네트워크	행적
참이슬	낭랑	두나	페미	공중	키치 꼬치	베로니카
번지 점프	공상	탈탈	국민신문고	재정	기괴	루이
타자	스릴	찰리	비속어	티아	토미자와 히 도시	흑막
궁금합니다	판타지 만화	한솔	소방	재민	아포칼립스	학살
대천	뿌잉을	브리칸	소방관	누가하는지모 르겟지만	최유진	성하
고속	동인	성민	반입	갈아엮	형이상학	무능력자
스피드	구가	당할연	피소	빨갇네	꾸짖	홍초
철강철렁으로	헌터	탈탈탈	모욕	하필이면	취적당해버렸네요	학살던전마

Fig. 4. 승격이 되지 않은 웹툰 TF-IDF

IV. 결론 및 시사점

본 연구는 정식연재로 승격된 웹툰과 그렇지 못한 웹툰의 댓글을 텍스트 마이닝 기법을 적용하여 승격과 관련하여 어떠한 차이가 있는지 분석하고자 하였다. 이를 위해 총 14편의 웹툰의 댓글을 수집하였고 36,194개의 댓글을 형태소 분석하여 빈도분석 및 TF-IDF 분석을 진행하였다. 그 결과 웹툰 독자들이 작성한 댓글의 특성과 승격된 웹툰 그룹과 그렇지 못한 웹툰 그룹의 댓글 특징을 확인할 수 있었다. 독자가 댓글을 통해 작가에게 자신의 의견을 표명한다는 점에서, 각 웹툰 댓글에 대한 텍스트 분석은 독자의 의견을 한눈에 파악할 수 있었으며, 승격이 되지 못한 웹툰의 경우에는 독자가 웹툰의 어느 요소에 불만족한 의견을 표출하는지 파악할 수 있었다. 선행 연구에서는 특정한 한 편의 웹툰의 댓글로 연구를 했다면 본 연구는 14편의 웹툰의 댓글을 분석을 진행하여 새로운 시도를 했다는 점에서 학문적 의미가 있다. 또한 승격이 된 웹툰의 경우에는 독자가 웹툰의 어느 요소에 만족하고 있고, 댓글 중 등장인물 이름의 빈도에 따라 얼마나 몰입하여 감상하고 있는지 확인할 수 있었다. 이러한 댓글을

통해 독자들이 2차적 저작물에 대한 언급한다면 웹툰 플랫폼은 웹툰의 2차적 저작물 기획 단계에서 활용될 수 있다는 점에서 실무적 의미가 있다. 본 연구에서는 웹툰의 댓글을 워드클라우드와 TF-IDF만 적용하여 분석했다는 점에서 한계점을 갖는다. 추후 연구에서는 승격요인에 영향을 미치는 다른 정량적인 수치와 댓글을 통하여 분석한다면 더 정확하게 승격요인을 확인할 수 있을 것으로 기대한다.

REFERENCES

- [1] C. N. Shin, Y. E. Kim, and Y. K. Park, "A Study on the Effects of After-purchase Feedback About Customer Service Quality on Purchase Process-Focusing on Internet Shopping Mall," *Journal of Distribution Research*, Vol. 14, No. 1, pp. 27-44, 2009.
- [2] D. K. Lee, S. J. Lee, I. Choi. "Response Analysis of Stop Smoking Campaign Webtoon Using Text Mining Technology: Focused on the "Tale of Cigarette" Comments," *Journal of Health Informatics and Statistics*, Vol. 43, No. 1, pp. 70-79, 2018.
- [3] E. A. Park, "Effects of the direction of online reviews on purchase decision: The role of product type and brand reputation," *The Korean Journal of Advertising*, Vol. 18, No. 5, pp. 7-22, 2007.
- [4] E. J. Lee, and W. S. Shim, "A study on the behavior characteristics of point of purchase, postpurchase and trust evaluation of internet shopping afternotes," *The e-business studies*, Vol. 8, No. 3, pp. 155-170, 2007.
- [5] E. Lee, and Y. J. 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, pp. 50-71, 2009
- [6] E. Y. Lee, and T. M. Lee, "The effects of information characteristics on word of mouth performance in online environment: Focused on the moderating effect of consumer knowledge," *The Korean Journal of Advertising*, Vol. 16, No. 2, pp. 145-171, 2005.
- [7] H. S. Jang and K. C. Lee, "Data Mining Approach to Predicting Serial Publication Periods and Mobile Gamification Likelihood for Webtoon Contents," *Journal of the Korea Society of Computer and Information*, Vol. 23, No. 4, pp. 17-24, 2018.
- [8] J. H. Yang, J. Y. Lee, and S. W. Lee. "Factors Affecting Webtoon's Success : An Empirical Study," *JOURNAL OF THE KOREA CONTENTS ASSOCIATION*, Vol. 16, No. 5, pp. 194-204, 2016.
- [9] J. Kang, and S. Kim, "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, pp. 143-166, 2012.
- [10] J. Lee, J. Hwang, "Characteristics of and Consumer's Responses to Brand Webtoons. *Advertising Research*," No. 115, pp. 246-296, 2017.
- [11] K. S. Choe, "About the Multi-layered Communication of Princess Pari on the Webtoon Platform of Daum?Focusing on Analysis of Narrative Structure and Comments," *Journal of Popular Narrative*, Vol. 25, No. 3, pp. 303-345, 2019.
- [12] K. Yoon, K. Jung, I. Choi, and H. Choi, "Features of Korean Webtoons through the Statistical Analysis," *Cartoon and Animation Studies*, pp. 177-194, 2015.
- [13] National IT Industry Promotion Agency, "Study on globalization strategy of webtoon platform," National IT Industry Promotion Agency, 2017.
- [14] P. Lee, and S. Park. "An Exploratory Study on the Critics's Reviews Reported in the Press : Focusing on the Relationship Between Opinion Quality of Film Reviews and Box Office Performance," *Journal of the Korea Entertainment Industry Association*, Vol. 13, No. 7, pp. 1-13, 2019.
- [15] S. C. Seong, and J. Kim, "Webtoon, Dream of 1 Trillion Won Market," Seoul: KT Economic Management Institute Issue & Trend, pp. 1-10, 2015.
- [16] S. Lee. and J. B. Gwan. "The Current Service Status and the Developmental Direction of Webtoon 2.0: Focusing on the Changing of User Interface and the Applying of Multimedia Effects," *The Journal of the Korea Contents Association*, Vol. 15, No. 8, pp. 96-108, 2018.
- [17] S. M. Mun, H. J. Ha, K. Lee, "Relationship Analysis between the Box Office Performance and Sentimental Words in Movie Review," *Society of Design Convergence*, Vol. 14, No. 4, pp. 1-16, 2015.
- [18] T. Hennig-Thurau, K. P. Gwinner, G. Walsh, and D. D. Gremler, "Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet?," *Journal of interactive marketing*, Vol. 18, No. 1, pp. 38-52, 2004.
- [19] W. Y. Chun, and H. J. Jung, "Effects of online reviews on evaluation and purchase intention of a product in internet shopping: The role of gender differences," *The Korean Journal of Consumer and Advertising Psychology*, Vol. 7, No. 1, pp. 113-129, 2006.