



온라인 댓글의 주제 분석을 위한 토픽 모델링

이슈 포착과 분류에 활용 가능한 LDA와 BTM의 비교와 검증

이신행 중앙대학교 미디어커뮤니케이션학부 부교수

Topic Modeling for Analyzing Online Comments*

Comparing and Validating LDA and BTM for Capturing and Classifying Issues

Shin Haeng Lee**

(Associate Professor, School of Media & Communication, Chung-Ang University)

Using computers to rapidly and efficiently build a model to organize massive volumes of textual data, topic modeling is an unsupervised machine learning technique that can be used to classify texts into related themes or to analyze the nature and distribution of topics. However, topic modeling's usage in media research has recently come under fire for failing to take into account reliable and valid measures of theoretically defined concepts. This means that topic modeling needs methodological validation and reliability in order to be employed in media research as a tool for investigating and summarizing massive volumes of textual material. Additionally, it is helpful to be able to group online comments into "issues" and list their important points in order to quickly identify social issues and monitor discourse patterns in real-time on digital platforms. For this reason, attempts to diagnose the methodological validity of topic modeling for analyzing the topics of comments are of great significance. Therefore, this study validates topic modeling for analyzing online comments by verifying its performance as follows. First, we discussed why topic analysis of comments is necessary and what the implications are through the conceptualization of "issues" in the context of online comments. Then, with an emphasis on the Latent Dirichlet Allocation (LDA) model, we reviewed the principle of topic modeling to estimate the topic of text and the assumptions of statistical models that affect topic estimation. Additionally, we contrasted the merits and drawbacks of LDA and the Biterm Topic Model (BTM) to suggest topic modeling as a means of identifying the subject of comments and categorizing them as "issues." Based on the above theoretical discussion, we applied topic modeling to analyze 9,000 online news comments on articles covering nine social issues and validate whether the topics

* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea(NRF-2017S1A6A3A01078538)(이 논문은 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2017S1A6A3A01078538)).

** shinlee@cau.ac.kr

are useful to classify comments according to the "issues" of the news. The results are as follows. First, compared to BTM, LDA is highly dependent on the hyperparameter, α , with lower values leading to better model performance. Second, both BTM and LDA were able to estimate the optimal number of topics (K), but BTM showed less variation in performance with value selection than LDA, and performance degradation was worse when the value was lower than the optimal K than when it was higher. Third, both BTM and LDA performed better when adding bigrams along with unigrams to the vocabulary, but the difference was more pronounced for LDA. Based on these validation results, we assessed the validity of topic modeling for analysis of comments and discussed its implications.

Keywords: Online Comment, Topic Modeling, LDA, BTM, Model Validation

국문초록

토픽 모델링은 컴퓨터를 이용해 빠르고 쉽게 방대한 양의 정보를 구성하거나 파악하는 모델을 추정하는 비지도 기계학습 방법으로 문서를 비슷한 토픽별로 묶어내거나 토픽의 내용과 분포를 파악하고자 할 때 활용된다. 그런데 토픽 모델링을 미디어 연구에 활용함에 있어 이론적으로 정의된 개념을 타당하고 신뢰할 수 있는 측정하는 방법으로써의 고민이 부족했다는 지적이 최근 제기되고 있다. 대단위 텍스트 자료를 탐색하고 요약하는 도구인 토픽 모델링을 미디어 연구에 활용하기 위해서 방법적 타당성 검증과 신뢰성 확보가 필요하다는 것이다. 더욱이, 온라인 플랫폼을 통해 거대한 규모로 빠르게 발생하고 있는 댓글을 '이슈'에 따라 분류하고 그 주요 주제를 요약하는 방법은 사회적 쟁점을 즉각적으로 파악하고 담론 양상을 통시적으로 추적하는데 유용하다. 이러한 이유에서 댓글의 주제를 분석하기 위한 토픽 모델링의 방법적 타당성을 진단하는 시도는 그 의미가 크다고 할 수 있다. 이에 본 연구는 온라인 댓글의 주제 분석을 위한 토픽 모델링 방법을 제시하고 그 성능과 타당성을 다음과 같이 검증했다. 우선, 온라인 댓글의 맥락에서 다뤄지는 '이슈'에 대한 개념화를 통해 댓글의 주제 분석이 왜 필요하고 어떠한 함의가 있는 것인지에 대해 논의했다. 그리고 토픽 모델링이 텍스트의 주제를 추정하는 원리와 주제 추정에 영향을 미치는 통계 모델로써의 가정들에 대해 Latent Dirichlet Allocation(LDA) 모델을 중심으로 살펴봤다. 또한 댓글의 주제를 포착하고 '이슈'별로 분류하기 위한 토픽 모델링을 제안하고자 LDA와 Biterm Topic Model(BTM)의 성능과 한계를 비교했다. 이상의 이론적 논의를 토대로 9개의 사회적 이슈를 다루고 있는 기사에 달린 9,000건의 온라인 뉴스 댓글을 토픽 모델링으로 분석해 주제를 추정하고 이에 따라 댓글이 뉴스의 '이슈'에 따라 분류되는지를 모델별로 비교 검증했다. 그 결과는 다음과 같다. 첫째, BTM에 비해 LDA는 초모수 α 에 많은 영향을 받았는데 α 값이 낮아질수록 모델의 성능이 좋아졌다. 둘째, BTM과 LDA 모두 최적의 주제의 개수(K)를 추정할 수 있었으나, BTM이 K 값 선정에 따른 성능 변화가 LDA보다 적었고 K 값이 최적의 값보다 클 때보다 낮을 때 성능 저하가 심해졌다. 셋째, BTM과 LDA 모두 분석 단어 목록에 단일 형태소와 함께 바이그램(bigram)을 추가할 때 성능이 좋아졌으나 그 차이는 LDA에서 더욱 뚜렷했다. 이러한 검증 결과를 토대로 댓글의 주제 분석을 위한 토픽 모델링의 활용 가능성을 진단하고 그 함의를 논의했다.

핵심어 : 온라인 댓글, 토픽 모델링, LDA, BTM, 타당성 검증

1. 서론

토픽 모델링(topic modeling)은 컴퓨터를 이용해 빠르고 쉽게 방대한 양의 정보를 구성하거나 파악하는 모델을 추정하는 비지도 기계학습 방법으로 이를 이용해 문서를 비슷한 토픽별로 묶어 내거나 토픽의 내용과 분포를 파악하고자 할 때 활용된다. 여기서 문서란 뉴스 기사나 논문, 또는 블로그 게시물 혹은 댓글 등 길어와는 상관없이 하나의 독립된 텍스트 개체를 의미한다. 또한 토픽은 문서의 주제로서 문서에서 자주 등장하는 단어들로 추론되는 내용적 특성을 뜻한다. 이에 토픽 모델링은 디지털 미디어에서 생산되는 텍스트의 폭발적 증가와 함께 그 활용 범위와 가치가 매우 높아지는 추세다. 전통적 내용분석의 방법으로는 가능하지 않았던 방대한 크기의 텍스트 자료를 효율적으로 다뤄 자료에 기반하여 그 의미를 파악하고 주제적 특성을 도출하는 방법으로 토픽 모델링이 활용될 수 있기 때문이다. 무엇보다 토픽 모델링은 의제(agenda) 분석이나 프레임 분석과 같이 내용분석으로 미디어 텍스트의 특성을 도출해 의미를 추론하는 연구에 있어 전통적 방법을 보완하는 것을 넘어 대안으로서까지 그 가능성이 활발하게 논의되고 있다(이준웅·김성희, 2018; Maier et al., 2018; Walter & Ophir, 2019).

그러나 토픽 모델링을 미디어 연구에 활용함에 있어 이론적으로 정의된 개념을 타당하고 신뢰할 수 있는 측정하는 방법으로써의 고민이 부족했다는 지적이 최근 제기되고 있다(Chen, Peng, Kim, & Choi, 2023; Ying, Montgomery, & Stewart, 2022). 이는 대단위 텍스트 자료를 탐색하고 요약하는 도구인 토픽 모델링을 미디어 연구에 활용하기 위해 방법적 타당성 검증과 신뢰성 확보가 필요함을 의미한다. 따라서 본 연구는 토픽 모델링의 방법 확립과 지평 확장을 위한 실증 분석을 시도하며 댓글의 주제 분석에 적용코자 한다. 댓글은 소셜미디어와 스마트폰의 확산과 함께 광범위한 영역에서 대단위로 빠르게 생산되고 수집되는 단문(short texts) 형식의 대표적 온라인 텍스트이다. 특히, 뉴스 댓글은 특정 ‘이슈’에 대해 혹은 이를 보도하는 지사나 매체에 대해 찬반 의견 또는 긍·부정의 정서를 표현하기 때문에 이를 통해 온라인 여론 동향을 조망하는 지표로 작용한다(이은주·장윤재, 2009). 더욱이 뉴스 이용자는 ‘이슈’에 대한 공중 담론 형성에 있어 댓글을 통해 적극적 참여자로 관여할 수 있다(김혜미·이준웅, 2011; 황태연·유정모·김화연, 2022). 여기에서 ‘이슈’란 사회적 문제에 대해 서로 반대되는 입장을 취하는 집단 사이에서 논쟁이 일어나거나 그 여지가 있는 사안을 의미한다(Miller & Reichert, 2000). 따라서 댓글이 ‘무엇에 대해 말하고 있는가’(What to talk about)를 파악하고 해당 ‘이슈’에 따라 댓글을 분류하는 방법은 미디어 연구에 있어 중요한 함의를 갖는다. 결국, 온라인 플랫폼을 통해 거대한 규모로 빠르게 발생하고 있는 댓글을 ‘이슈’에 따라 분류하고 그 주요 주제를 요약하는 방법은 사

회적 쟁점을 즉각적으로 파악하고 담론 양상을 통시적으로 추적하는데 유용하다. 이러한 이유에서 댓글의 주제를 분석하기 위한 토픽 모델링의 방법적 타당성을 진단하는 시도는 그 의미가 크다고 할 수 있다.

이러한 댓글 분석의 가치와 의미와는 달리 빅데이터 관점에서 댓글을 내용적으로 분석한 국내 연구는 작성자의 정서를 지도기계학습이나 딥러닝 방식으로 측정(이신행·현아연, 2023; 황태연 외, 2022)하거나 댓글에서 추출된 단어들의 공기(co-occurrence)관계에 기반해 의미연결망 분석으로 주요한 언어적 표현을 도출(안순태·임유진·이하나, 2020)하는 연구에 한정된 측면이 있다. 반면 대단위 텍스트의 주요 의미 구조를 파악하기 위해 활발하게 이용돼 온 토픽 모델링은 댓글 분석에 적용된 경우는 미미했다. 대표적인 토픽 모델링 기법으로 이용되고 있는 LDA(Latent Dirichlet Allocation)와 STM(Structural Topic Model)의 특성상 댓글과 같은 단문 형식의 텍스트에서 주제를 정확하게 파악하기 어렵다는 점 때문이다(안순태·이하나·정순달, 2022). 문서에 등장하는 단어들로 잠재적 주제를 추정하는 토픽 모델링 방법은 상대적으로 희소한 단어들로 구성되는 단문에 적합하지 않다는 것이다.

이 문제를 해결하기 위해 제시된 방법이 BTM(Biterm Topic Model)이다. BTM은 각 문서에서 잠재된 주제에 의해 단어 출현 패턴이 결정되는 것이 아닌 문서 전체 집합에 분포된 주제에 의해 단어들의 공기(co-occurrence) 패턴이 나타난다고 전제함으로써 문서 내의 단어 희소성 문제를 극복한다(Yan, Guo, Lan, & Cheng, 2013). 다시 말해, 주제를 한 문서 내에서 출현하는 단어들로 추정하는 것이 아니라 여러 문서들에 걸쳐서 동시에 출현하는 단어쌍(biterm)의 패턴으로 추정한다는 것이다. 이 방법은 트윗이나 질문과 같은 단문은 물론 비교적 장문의 온라인 게시글에 대해서도 도출한 주제의 의미적 일관성과 주제에 따른 문서 분류에 있어서도 LDA보다 좋은 성능을 보이는 것으로 나타났다(Yan et al., 2013). 그러나 분석 대상이 영어 텍스트였고 댓글에 대한 주제 분석에는 적용되지 않았다는 점 등의 이유는 한국어 단문을 대상으로 한 토픽 모델의 검증은 요구한다. 이에 문길성(2021)은 LDA와 BTM의 성능을 한국어 기반 질문과 기사 제목에 대해 비교해 BTM이 주제 일관성과 문서 표현 측면에서 상대적으로 우수한 모델임을 제시했다. BTM을 다양한 텍스트를 대상으로 주제를 분석하는데 활용돼 온 LDA와 비교하여 그 성능을 평가하고 단문에서의 적합성을 검증한 것이다. 그러나 주제 개수(K) 이외에 LDA의 초모수 조절을 통한 성능 비교가 없었고 텍스트 토큰화 방식의 효과를 살피지 않았다는 점, 그리고 무엇보다 성능 평가 척도가 자동화 방식의 양적 지표로만 구성되어 도출된 주제에 대한 인간의 해석 가능성은 고려되지 않았다는 점 등을 고려할 때 LDA와의 성능 비교를 통한 BTM의 검증은 여전히 요구된다고 할 수 있다. 더욱이 LDA와 달리 BTM은 문서 전체에 걸쳐

주제들의 분포를 추정할 뿐 문서 내에서의 주제 분포를 추정하지 않기 때문에 이슈에 따라 댓글을 분류하는 방법으로서의 타당성 검증이 필요하다.

이상의 문제의식에서 본 연구는 온라인 댓글의 주제 분석을 위한 토픽 모델링 방법을 다음과 같이 검토하고 그 성능과 타당성을 검증했다. 우선, 온라인 댓글의 맥락에서 다루지는 ‘이슈’에 대한 개념화를 통해 댓글의 주제 분석이 왜 필요하고 어떠한 함의가 있는 것인지에 대해 논의했다. 그리고 토픽 모델링이 텍스트의 주제를 추정하는 원리와 주제 추정에 영향을 미치는 통계 모델로서의 가정들에 대해 Latent Dirichlet Allocation(LDA) 모델을 중심으로 살펴봤다. 또한 댓글의 주제를 포착하고 ‘이슈’별로 분류하기 위한 토픽 모델링을 제안하고자 LDA와 Bitern Topic Model(BTM)의 성능과 한계를 비교했다. 이상의 이론적 논의를 토대로 9개의 사회적 이슈를 다루고 있는 기사에 달린 9,000건의 온라인 뉴스 댓글에 토픽 모델링을 적용하여 추정된 댓글의 주제가 뉴스의 ‘이슈’에 따라 분류되는지를 모델별로 비교 검증했다. 그리고 검증 결과를 토대로 댓글의 주제 분석을 위한 토픽 모델링의 활용 가능성을 진단하고 그 함의를 논의했다.

2. 이론적 배경

1) 댓글의 주제 포착과 이슈별 분류는 왜 중요한가?

댓글은 인터넷 게시물에 대해 이용자들이 반응 차원에서 남기는 단문(short texts)으로 디지털 미디어 환경에서 개인들이 의견을 자유롭게 표현하는 도구이자 이전의 매스미디어에서는 볼 수 없었던 쌍방향식 커뮤니케이션 산물이다(이재신·김지은·류재미·강재혁, 2010). 특히, 댓글은 인터넷을 매개로 생산 및 유통되는 게시물에 대해 이용자들이 어떠한 의견과 태도를 보이는지 엿볼 수 있는 수단일 뿐만 아니라 댓글이 게시물 하단에 위치해 누적된 형태로 공개되고 게시물과 함께 소비되며 콘텐츠로서의 효과를 가진다(염정운·김류원·정세훈, 2020; 정일권·김영석, 2006). 가령, 온라인 뉴스기사의 댓글은 기사의 내용과는 구별된 효과를 뉴스 이용자에게 미치게 되는데, 그 이유는 이용자들이 기사를 해석하는데 댓글을 참고하여 기사의 내용을 평가하고 해당 이슈를 해석하기 때문이다(이은주·장윤재, 2009; 이재신·성민정, 2007). 더욱이 댓글이 이용자의 의견과 일치할 경우에는 기존 의견이 강화되고 일치하지 않는 경우에도 태도 변화를 유도하는 등의 효과를 보인다(강재원·김선자, 2012). 그리고 기사에 달린 댓글의 전반이 부정적 성격일 경우에는 뉴스 내용에 대한 신뢰성을 저하시킬 수 있다(이재신·성민정, 2007), 이러한 이유로 댓글은 인터넷 사용자들 간의 상호작용 및 쌍방향 커뮤니케이션을 유도하는 도구로써 해

당 이슈에 대한 담론과 여론을 조성하는 작용을 한다(Papacharissi, 2004).

온라인 뉴스 댓글은 그 내용이 이용자의 의견과 태도 형성에 영향을 미치기도 하지만 얼마나 많은 댓글이 달리는가가 해당 이슈에 대한 관심도나 주목도를 보여주기도 한다. 또한 온라인 뉴스의 유통 채널인 네이버와 같은 포털 플랫폼에서는 댓글 수 지표를 이용자들에게 공개하고 댓글 많은 뉴스를 추천해주는 큐레이션 서비스를 제공한다는 점에서 댓글은 인터넷을 매개로 이루어지는 뉴스의 생산과 전파에 영향을 미친다. 더 나아가 누적된 댓글의 대다수가 비슷한 방향의 반응으로 구성될 경우엔 뉴스를 해석하는데 있어 프레임으로 작용할 수 있고(염정윤 외, 2020), 다른 뉴스 이용자들에게는 일종의 여론으로 읽히게 된다(김은미·선유화, 2006; 김혜미·이준웅, 2011). 결국, 댓글 집합은 기술적 규범(descriptive norms)으로 작용하며 이용자 개인이 다수 의견에 동조하는 효과를 발휘할 수 있다는 것이다(정일권·김영석, 2006). 이러한 댓글 효과는 전통적 매체를 통한 뉴스 소비 구조에서는 이뤄질 수 없었던 대인 커뮤니케이션 요소가 댓글로써 결합돼 이용자 간의 사회적 상호작용이 이뤄지는 공간을 제공함과 동시에 댓글이 공중의 의견으로 비춰지며 사회적 이슈의 여론 형성 기제로 작용하고 있음을 의미한다(최동성·최성은·최용준, 2008). 물론, 댓글 작성자는 소수에 불과해 댓글 주류의 의견이 사회 여론이라고 보기엔 무리가 있다는 지적 또한 타당하다(고문정, 2018). 그럼에도 불구하고 댓글에 노출된 이용자들은 특정 이슈에 대한 사회적 의견을 댓글에 기초해 유추하는 경향이 있어 댓글은 여론 조성과 변화에 영향을 미칠 수 있다(염정윤 외, 2020).

댓글이 수용자에게 가지는 효과를 바탕으로 뉴스 댓글의 내용을 분석한 최근 연구에서는 이용자 개인의 자유로운 의견으로서 기사의 내용과는 관계없이 작성자가 적극적으로 자신의 솔직한 의견을 댓글로써 적극적으로 표현할 뿐만 아니라(안순태 외, 2022), 이를 통해 댓글의 대향적 메시지가 사회문화적 체제에 맞서는 대안 담론의 형성을 주도한다고 주장한다(김은진·강주현, 2019). 이러한 이유로 댓글은 게시글에 대한 반응 정도를 보이는 양적인 측면과 함께 어떠한 관점과 해석으로 구성되고 있는가를 나타냄으로써 여론의 동향과 추세를 파악할 수 있는 지표가 된다. 다시 말해, 댓글은 온라인 공간에서 구성원들의 목소리가 어떠한 이슈를 중심으로 표출되고 있는가를 보여줌과 동시에 개인 사용자들의 이슈 지각과 태도에 영향을 미친다는 것이다(양혜승, 2008). 따라서 댓글을 주제적 특성에 따라 분류하는 방법은 온라인 공간에서 형성되고 있는 담론의 형성과 흐름을 추적하는데 유용할 뿐만 아니라 우리 사회의 주요 이슈와 여론 양상을 포착하는데 기여할 수 있다. 이 관점에서 토픽 모델링은 뉴스 댓글을 주제에 따라 분류하여 해당 이슈에 대한 태도를 탐색하거나(안순태 외, 2022), 혐오표현의 내용적 특성을 포착하는데(양혜승, 2022) 활용되어 왔다. 그러나 주제에 따른 댓글 분류가 타당한 방식으로 이뤄지는가에

대한 검증이 부재했고 모델 조정에 따른 성능 평가 역시 부족했기 때문에 이에 대한 실증 연구가 요구되는 실정이다.

물론, 댓글의 주제를 파악하고 이슈에 따라 분류하기 위해서는 댓글이 게시글에 대한 반응이라는 점을 고려해야 한다. 댓글의 주제는 곧 게시글의 주제일 수 있기 때문이다. 이러한 이유로 댓글의 주제를 파악하는 방법이 왜 필요한 것인가에 대한 의문이 제기될 수 있다. 댓글의 주제를 파악하고 이슈에 따라 분류하는 방법이 원 게시글에 대한 주제 분석으로 대체될 수 없는 까닭은 무엇일까? 첫째는 댓글의 경우 광고성 정보나 스팸(spam), 또는 게시글 내용과는 전혀 상관없는 단순 욕설이나 웃음 표시, 이모티콘 등으로만 구성된 메시지가 포함될 수 있다는 점이다. 이렇게 게시글의 주제에서 벗어난 댓글은 분석에서 제거되어야 댓글 작성자들의 의도와 여론을 보다 정확하게 포착할 수 있다. 둘째, 비교적 단일한 의제를 다루고자 게시되는 뉴스 기사나 온라인 커뮤니티 또는 블로그 게시글과는 달리 여러 주제를 포함하는 뉴스 영상이나 1인 미디어의 동영상 콘텐츠, 또는 생방송 스트리밍의 경우에는 같은 게시글이라도 댓글들에서 여러 주제가 복합적으로 나타날 수 있어 이를 주제별로 분류할 필요성이 있다. 셋째, 댓글과 같은 단문의 주제를 파악하는 방법은 트위터와 같이 단문 기반의 온라인 플랫폼을 통해 빠르게 생성 및 축적되고 있는 개인들의 메시지를 효율적으로 분석할 수 있는 도구로써 그 활용 가치가 크다고 할 수 있다. 이상의 이유를 토대로 본 연구는 댓글의 주제를 파악하고 이에 따라 분류할 수 있는 방법으로 토픽 모델링을 제안하고자 한다. 이를 위해 우선 토픽 모델링이 어떠한 원리와 과정으로 주제를 포착하고 문서를 분류하는지 설명하고자 한다.

2) 텍스트의 주제 탐색과 분류를 위한 토픽 모델링의 원리

토픽 모델링은 디지털 텍스트 자료의 생산과 저장, 그리고 사용이 보편화됨에 따라 대단위의 정보를 정리하고 탐색하거나 파악하기 위한 도구의 필요로 등장 및 발전해 온 대표적 컴퓨테이션얼(computational) 도구이다. 보다 구체적으로, 토픽 모델링은 통계적 가정에 기초해 각 문서에서 출현하는 단어들과 그 빈도를 생성하는 모델의 모수들(parameters)을 추정함으로써 전체 문서 집합에 잠재하는 주제의 구조와 특성을 포착한다. 이 중 대표적 토픽 모델은 LDA로서 이 모델은 다음의 세 통계적 가정에 기초해 주제의 구조를 추정한다: 1) 각 문서에는 복수의 주제가 존재하고 2-1) 각 문서의 단어 출현은 전체 문서에서의 주제 분포에 영향을 받는데, 2-2) 이때 주제가 각 문서에 할당된 정도 역시 단어 출현에 영향을 미친다(Blei, 2012). 이러한 가정 하에 LDA는 주제의 잠재적 구조에 따라 문서에서 단어들이 출현하게 되는 생성적 모델(generative model)을 추정한다. 여기서 주제의 잠재적 구조는 1) 각 주제(k)의 단어들에 대한 할당 분포

($\beta_{1:K}$), 2) 각 문서(d)에 대한 주제들의 할당 분포(θ_d), 그리고 3) d 번째 문서에서 n 번째 단어에 대한 주제 할당($z_{d,n}$)의 변수들로 구성된다(Blei, 2012). 결국, LDA는 잠재적 주제 구조에 의해 문서에서의 특정 단어가 출현할 확률이 결정된다고 보고 관찰 변수인 문서의 단어들로부터 잠재 변수들의 사후 분포(posterior distribution)를 산출한다. 즉, 문서의 잠재적 주제 구조를 추론하고자 문서에 대한 잠재 변수들의 조건부 분포인 사후 분포를 계산하는 것이다(Blei, 2012).

LDA가 추정하는, 즉 관찰된 문서의 단어 출현을 고려해 주제 구조의 사후 분포를 산출하는 계산은 이론적으로는 가능하나 기하급수적으로 많은 경우들에 대한 확률 계산을 거쳐야하는 문제가 있다. 따라서 LDA는 관찰 변수를 고려하며 임의 변수의 사후 분포를 반복 표집하면 실제 분포와 가깝게 수렴하게 되는 마르코브 연쇄(Markov chain) 원리에 기초한 깁스 표집(Gibbs sampling)을 주로 사용해 잠재적 주제 구조를 근사화(approximation)한다. 물론 토픽 모델링은 깁스 표집으로 사후 분포의 근사치를 추정하는 방법 이외에도 잠재 구조의 분포들에 대한 모수(매개변수)군을 가정한 후 실제와 가장 가까운 분포의 모수군을 찾는 변분 방법(variational method)이 사용된다. 이 변분 방법은 모수군의 최적화를 도출하기 때문에 문서 차원의 변수들에 대한 모수들까지 포함한 모델 추정이 가능하고, 이를 위한 기댓값 최대화(expectation-maximization) 알고리즘을 사용하는 STM(Structural Topic Model)이 제안되었다(Roberts, Stewart, & Airoldi, 2016). 특히, STM은 LDA와 같이 문서에서의 단어 출현을 예측하는 생성적 모델을 가정하지만 여기에 문서 단위의 모수를 추정함에 따라 저자나 시기 등 문서 단위의 특성에 따라 주제가 만연한 정도(topical prevalence)나 그 내용적 성격(topical content)이 어떻게 변화하는지도 측정한다. 결국, 토픽 모델링은 모두 전체 문서 집합에서 추출된 단어들에 대해 각기 다른 할당 확률을 가지는 잠재적 주제들이 각기 다른 비중으로 할당된 각 문서를 생성하는 모델을 충분히 반복해서 추정한 후,¹⁾ 실제 문서와 가장 가까운(likely) 결과를 도출하는 모델을 도출하는 <Figure 1>의 과정을 거친다(Blei, 2012). 이 모델은 문서와 주제의 관계(θ_d) 그리고 주제와 단어의 관계(β_k)를 나타내는 모수를 추정하기 때문에, 최적 모델에서 추정된 β_k 로부터 주제의 의미를 추론하고 θ_d 를 이용해 문서에서 어떠한 주제가 현저하게 다뤄지는지 파악할 수 있다.

1) 깁스 표집에 따른 모수군의 반복 추정이 최소 1,000회 이상이 되어야 신뢰성 있는 토픽 모델의 성능을 기대할 수 있다(Maier et al., 2018).

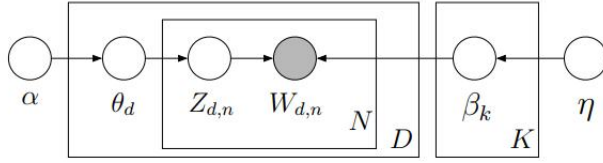


Figure 1. The process of iteratively generating the observed variable

Note. $W_{d,n}$, from the set of words in the document, D , according to LDA. This generative process is driven by the latent variables of θ_d as the distribution of topics in the document D , $Z_{d,n}$ which assigns words according to this distribution, and β_k as the distribution of assignments of words to topics. From "Probabilistic topic models," by D. M. Blei, 2012, *Communications of the ACM*, 55(4), p. 81. All rights of the source are protected.

LDA로 대표되는 생성적 토픽 모델링 기법은 주제 구조를 구성하는 잠재적 변수들의 사후 분포를 기계학습으로 근사화하는데 이 과정은 디리클레(Dirichlet) 분포를 따르는 모수의 초기 상태의 사전 분포(prior distribution)로부터 시작된다. 그리고 이러한 사전 분포에 성격을 규정하는 초모수(hyperparameter)에는 α 와 η , 그리고 K 가 있다. α 는 주제들에 대한 각 문서의 확률 분포(θ_d)를 η 는 단어들에 대한 각 주제의 확률 분포(β_k)의 사전 분포를 결정하는 초모수로서 잠재 변수에 대한 사후 분포의 형태(shape)와 특이성(specificity)에 영향을 미친다 (Maier et al., 2018). 구체적으로, α 값은 각 문서에 대한 주제들의 분포에 영향을 미치는데 그 값이 작을수록 소수의 주제들에 집중되게 되며 η 값은 각 주제에 대한 단어들의 분포에 영향을 미치고 그 값이 작을수록 소수의 단어가 할당된다. K 값은 주제의 개수를 의미하는데 높은 값일수록 문서에서 많은 수의 주제를 추정해 보다 구체적이고 세부적인 내용을 파악할 수 있지만, 너무 큰 K 값은 의미적 관련성이 적거나 중복되는 주제들까지 추정하게 된다(Maier et al., 2018).

이러한 LDA의 선택적 초모수 결정은 주제 구조 추정에 영향을 미치기 때문에 성능 평가를 통해 최적화된 값을 찾기 위한 조정(tuning)이 필요하다. 이 조정 작업은 다양한 초모수 값으로 구성된 모델로 토픽 모델을 일부 데이터(90%)로 학습해 추정한 후 그 결과 생성된 데이터와 학습에 사용되지 않은 나머지 데이터(10%)를 비교하여 모델의 적합도(goodness of fit)를 가능도(likelihood)로 평가하게 된다. LDA의 경우 이러한 모델 평가에 가능도와 반비례 관계인 혼잡도(perplexity)를 이용하는데, 이는 혼잡도가 낮을수록 모델이 더 좋은 성능임을 의미한다. 그러나 혼잡도를 토대로 도출한 토픽 모델이 인간 코더의 주제 해석 결과와 꼭 일치하는 것은 아니라는 지적이 있다. 최적의 토픽 모델로 각 주제에 대한 단어들의 할당 분포를 추정해 각 분포의 상위 단어들로 주제의 의미를 해석하는데 있어, 단어들이 의미적으로 유기적인 관계를 꼭 형성하는 것은 아닐 수 있기 때문이다(Chang, Boyd-Graber, Gerrish, Wang, & Blei, 2009). 따라서

각 주제에 있어 높은 할당 확률의 단어들이 의미적 일관성(semantic coherence)을 띄고 있는지 그리고 의미적으로 중복된 주제들이 도출되지 않는지(exclusivity) 역시 감안하여 초모수를 결정하게 된다(Roberts, Stewart, & Tingley, 2019).

결국, 토픽 모델의 예측 성능 척도를 토대로 초모수를 조정하는 과정은 초모수 조정이 텍스트의 주제를 포착하는데 있어 미치는 영향에 대한 고려를 요구한다. 가령, k 값을 너무 작게 설정할 경우 도출된 주제 간의 차별성이 떨어지는 반면 너무 크게 설정할 경우 의미적으로 중복되거나 해석이 어려운 주제들까지 도출될 수 있다. α 값 또한 토픽 모델의 성능에 중요한 역할을 하는데 문서에서의 주제 분포에 영향을 미치기 때문이다. 이에 초모수 비고를 통해 최적 모델을 도출함에 있어 η 값은 고정하는 반면 α 값에 대해서는 조정 작업을 거치며 최적 모델을 탐색하는 방안이 제안되었다(Maier et al., 2018). 이상의 토픽 모델링의 주제 추정 원리와 초모수 조정은 댓글의 주제를 포착하고 이슈별로 분류하는 방법에 필요한 통계적 가정들을 제시한다. 이를 토대로 다음 절에서는 댓글이라는 단문으로부터 주제를 추정하는데 요구되는 토픽 모델링의 가정에 대해 고찰해보고 타당하면서도 효과적인 댓글의 주제 분석 방법을 논의하고자 한다. 나아가 통계적 가정과 초모수 조정이 댓글에 대한 토픽 모델링 성능에 어떠한 영향을 미칠지에 대해 고찰할 것이다.

3) 이슈에 따라 댓글을 분류하기 위한 토픽 모델링

초모수 결정 과정은 댓글에 대한 토픽 모델링 성능에 어떠한 영향을 미치게 될까? 토픽 모델링은 선형적으로 주어진 k 값에 따라 각 주제가 각각의 문서에서 어느 정도의 비중으로 할당되는지를 추정한다. 즉, 모든 댓글은 각 주제들에 대한 확률적 할당 분포를 갖게 되는 것이다. 이를 통해 가장 높은 할당 확률의 주제가 같은 댓글들을 같은 이슈에 대한 댓글들로 분류할 수 있는 것이다. 이때 댓글의 주제에 대한 할당 확률은 초모수 k 와 α 에 영향을 받게 되는데, 앞서 언급했듯 k 값이 작아질수록 더 소수의 주제에 대한 할당 확률을 추정하게 되며 α 값이 작아질수록 소수의 주제에 집중된 확률 분포가 도출될 수 있기 때문이다. (Figure 2)는 초모수 α 값이 작을수록 점으로 표현된 문서들이 3개의 주제(꼭지점)로 분산되는 경향을 보이는 반면 α 값이 클수록 주제들의 중앙에 위치해 문서의 주제들 거리가 가까워짐을 보여준다. 즉, α 값이 커질수록 복수의 주제들이 문서에 영향을 미쳐 주요 주제 구분이 모호해지는 반면 값이 작아질수록 문서의 주된 주제를 잘 구분할 수 있음을 의미한다. 결국, 단문이라는 댓글의 특징을 고려할 때 소수의 주제가 출현할 가능성이 높고 이는 초모수 α 값을 작게 설정할수록 토픽 모델의 성능이 좋아질 수 있으리란 추정이 가능하다. 이에 다음의 연구질문을 제시하였다.

연구질문 1. 토픽 모델링의 초모수 K 와 α 에 대한 조정은 모델 성능에 어떠한 영향을 미칠까?

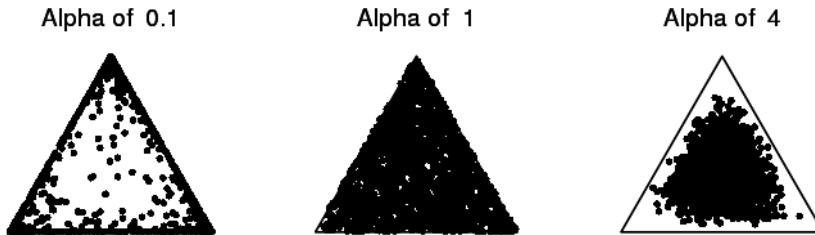


Figure 2 How the alpha parameter affects Dirichlet distribution

From "Experiments with Latent Dirichlet Allocation," by J. Moller-Mara, Aug. 3. 2014, Retrieved from <https://molleremara.com/blog/lda/>. All rights of the source are protected.

또한 문서당 소수의 주제에 대해 높은 할당 확률을 추정하는 토픽 모델은 K 값 선정에도 비교적 자유로울 수 있다. 최적의 K 값보다 큰 값을 설정하더라도 댓글에서 눈에 띄게 두드러지는 특정 주제와 그 외 나머지 주제들 사이에는 현저한 할당 확률 차이가 발생할 가능성이 크기 때문이다. 따라서 많은 주제의 수라도 댓글 전반에서 평균적으로 높은 할당 확률의 주제들을 중심으로 주제 분포와 내용을 파악하는 방법에 대한 타당성은 검증해 볼 만하다. 반면, 너무 낮은 K 값의 토픽 모델은 의미적 차별성이 적은 주제들을 도출하기 때문에 댓글에서 주된 주제별로 분류된 댓글이 이슈별로 구별되기 어려울 수 있다. 이에 최적의 K 값보다 낮은 K 값의 초모수로 구성된 토픽 모델은 좋은 성능을 보이지 않을 것이라는 추정이 가능하다. 따라서 다음의 연구질문을 수립하였다.

연구질문 2. 최적의 K 값보다 작거나 크게 K 값이 설정된 경우 토픽 모델의 성능은 어떠할까?

그런데 이러한 초모수 조정만으로 토픽 모델링을 이용한 댓글의 주제 분석에 좋은 결과를 기대할 수 있을까? 앞서 LDA는 대단위의 문서자료를 비슷한 주제별로 묶어내는데 유용한 방법이라고 설명했다. 이는 LDA가 댓글마다의 주제 분포를 추정하기 때문에 가능한 것이다. 그리고 이 분포 추정은 각 주제에 높은 할당 확률을 갖는 단어들(단어)이 문서에서 등장하는 출현 빈도에 기반한다. 다시 말해, LDA는 문서 단위에서 어떠한 단어들(단어)이 얼마나 등장하는가에 대한 패턴으로 주제의 분포를 추정한다. 문제는 댓글과 같은 단문에서는 상대적으로 소수의 단어들(단어)이 사용되고 그 빈도 역시 차이가 크지 않기 때문에 단어들의 문서 출현 빈도로 주제를 추정하는데 무리가 있다는 점이다. 즉, 댓글에서 주제를 추정하는데 관련성 높은 단어(또는 주제를 의미하는 단어)가 등

장하는 빈도가 드물기 때문에 이를 바탕으로한 주제 추론은 정확성이 떨어질 수밖에 없다는 것이다.

BTM(Biterm Topic Model)은 단문 데이터에 대해 LDA가 갖는 한계점을 지적하며 이를 극복하는 방안으로 제안된 토픽 모델이다(Yan et al., 2013). BTM은 각 단문에서 어떠한 단어들인 어느 정도의 빈도로 출현하는가가 아닌 전체 단문 집합에서 단어들인 공기하는 (co-occurring) 패턴으로 도출된 바이텀(biterm)을 바탕으로 주제 분포를 추정한다. 여기에서 바이텀이란 단문에서 그 순서와 관계없이 함께 사용된 단어쌍(an unordered pair of words)을 의미한다. 즉, BTM은 전체 단문 집합에 대해 단어들의 공기 관계에서 도출된 모든 바이텀의 생성을 추론하는 확률론적 생성 모델(probabilistic generative model)인 것이다. <Figure 3>은 문서의 주제 분포를 고려해 단어 출현을 생성하는 LDA와 달리 문서 전체에 대해 바이텀의 출현을 생성하는 BTM의 주제 추정 과정을 보여준다. 결국 전체 댓글을 대상으로 바이텀을 도출하고 단어와 바이텀의 관계, 그리고 바이텀과 전체 댓글의 관계를 추정하는 것이다. 이를 통해 BTM은 단문에서 함께 등장하는 단어 w_i 과 w_j 로 구성된 바이텀 $w_i + w_j$ 에 대해 주제 K_1 이 할당될 확률을 산출해 각 주제의 전체 문서에 대한 할당 분포를 추정한다. 또한 그리고 주제 K_1 에 대해서도 단어 w_i 과 w_j 이 각각 할당될 확률, 즉 주제에 대한 단어의 관계 정도를 추정해 주제의 의미를 추론하게 된다. 이렇게 단문 특징을 고려해 제안된 BTM은 보다 일관되고 분명한 주제 도출이 가능할 뿐만 아니라 장문 텍스트에서까지 LDA보다 더 나은 성능을 보여 범용적 활용 가능성이 제기되었다(문길성, 2021; Yan et al., 2013).

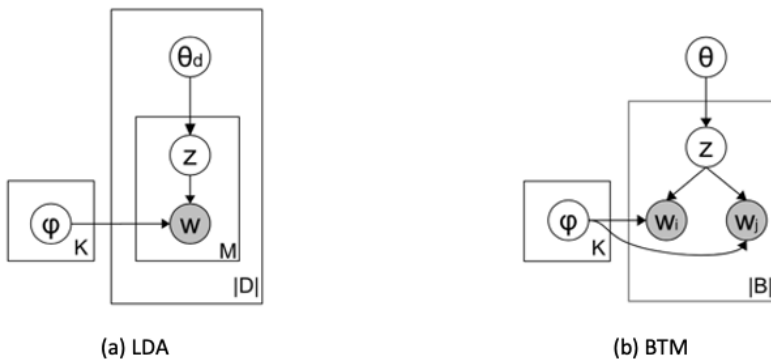


Figure 3. Graphical representations of LDA and BTM

From "A Biterm Topic Model for Short Texts," by X. Yan, J. Guo, Y. Lan, and X. Cheng, 2013, *Proceedings of the 22nd International Conference on World Wide Web*, p. 1447. All rights of the source are protected.

BTM은 주제의 분포나 단어와의 관계 추정에 있어 문서에서 단일 단어가 출현하는 빈도가 아닌 단어와 단어가 함께 사용되는 경향에 기반한다. 다시 말해, 두 단어의 동시 발생(co-occurrence) 패턴으로 주제를 추정한다는 것이다. 함께 사용된 두 단어가 단일 단어보다 주제를 더 잘 드러낼 수 있을 뿐만 아니라 전체 자료를 대상으로 주제 분포의 학습을 향상시킬 수 있기 때문이다(Yan et al., 2013). 물론 BTM은 댓글에서의 주제 분포를 추정하지는 않는다. 그러나 추정된 단어와 바이그램과 관계를 기반으로 댓글에 대한 주제 분포를 예측할 수 있어 이를 이용해 댓글을 주요 주제별로 분류할 수 있다.

이상의 논의를 통해 LDA의 주요한 통계적 가정과 단문 분석의 한계, 그리고 이에 대한 해결책으로 제시된 BTM의 특징을 살펴봤다. 댓글이 어떠한 ‘이슈’에 대한 내용인지를 포착하고 주제에 따라 분류하는 방법은 단문 위주의 디지털 텍스트를 분석하고자 함에 있어 그 방법적 활용도가 크다고 할 수 있다. 그러나 대표적 토픽 모델인 LDA의 통계적 가정이 단문에 적합하지 않다는 이유로 댓글에 대한 주제 분석은 뉴스 기사나 논문 등에 비해 상대적으로 매우 저조했다. LDA의 단점을 해결하는 모델로서 BTM을 사용하여 댓글의 주제를 분석한 연구가 나타나고 있지만(안순태 외, 2022), 두 모델의 성능을 한국어 단문을 이용해 비교 검증한 연구는 미미한 현실이다. 그리고 BTM의 장점을 고려해 바이그램(bigram)으로 측정된 두 단어의 동시 발생 빈도를 LDA에 적용시켰을 때 단문에 대한 주제 분석 성능이 향상될지에 대한 평가 또한 해결되어야 할 과제다. 따라서 본 연구는 다음의 연구질문을 제시하고자 한다.

연구질문 3. 토픽 모델의 단어 목록에 단일 형태소(unigram)와 함께 바이그램(bigram)을 포함해 구성하면 모델 성능이 향상될 것인가?

3. 방법

본 연구는 토픽 모델링을 이용해 댓글의 주요 주제를 포착해 이에 따라 댓글을 분류하는 방법을 제시하고 타당성 검토를 목표로 한다. 따라서 본 연구는 다음의 절차에 따라 분석을 진행했다. 우선, 2020년부터 2021년까지 한국 사회의 혐오 문제에 대한 공론화가 일어났던 9개의 사회적 ‘이슈’를 선정하고 관련 기사에 달린 댓글을 국내 최대 포털 사이트인 네이버에서 수집했다. 선정된 이슈들은 외국인 혐오, 여성 혐오, 성소수자 혐오 문제가 불거진 사건들로 많은 논란과 침예한 대립이 발생했기 때문에 많은 양의 댓글이 달렸을 뿐만 아니라 각 ‘이슈’의 주요 담론을 반영한다

고 판단했다. 또한 혐오표현은 포털의 사건·사고 뉴스에 달리는 댓글의 상당한 비율을 차지하고 있을 뿐만 아니라 해당 이슈에 대한 개인의 지각과 태도에도 영향을 미치기 때문에(양혜승, 2018), 우리 사회에서 나타나는 혐오의 양상과 특징을 파악하는데 중요한 자료로 활용될 수 있다는 점도 고려했다. 나아가, 성소수자와 여성, 그리고 외국인은 우리 사회의 소수자 집단으로 혐오표현의 주요 대상이 되고 있을 뿐만 아니라(홍주현·나은경, 2016), 이들을 향한 악의적 댓글을 정확하게 포착하기 위해서는 사안의 맥락까지 고려되어야 한다는 점(이신행, 2021)을 감안하여 ‘이슈’에 따라 댓글을 분류하는 방법의 검증과 평가를 수행했다. 이를 위해 수집한 데이터에 대해 토픽 모델링으로 추정된 댓글에 대한 주제 분포상 가장 높은 주제에 따라 각 댓글을 분류했다. 그리고 토픽 모델링의 결과에 따라 분류된 댓글이 실제 수집된 뉴스의 이슈별로 분류된 것인지 그 일치도를 측정하는 방식으로 모델의 타당성을 평가했다. 분석 과정에 대한 보다 자세한 설명은 아래와 같다.

1) 데이터 준비와 처리

본 연구가 댓글의 주제 분석을 위한 토픽 모델링에 사용한 데이터는 다음 9개 이슈를 다룬 네이버 뉴스 기사에 달린 댓글이다. 뉴스에 달린 댓글을 토픽 모델링으로 주제를 추정해 이에 따라 분류했을 때 기사에서 다루고 있는 이슈별로 분류되는지에 대한 평가로 방법적 타당성을 검증하고자 하기 때문이다. 이에 2020년 1월부터 2021년 12월까지의 2년 동안 국내에서 혐오 문제가 불거진 주요 이슈를 선정해 이를 다루고 있는 네이버 뉴스 기사의 댓글을 <Table 1>과 같이 수집했다.

Table 1. NAVER News and Comments on Hateful Issues Targeting Social Minorities

Target	Issue	News	Comments
Foreigner	샘 오취리의 '관짝소년단' 비판	12	2,104
	영화 '청년경찰'의 조선족 비하	36	2,216
	제주도에 들어온 예멘인들의 난민신청	270	14,075
Female	소설 '82년생 김지영'의 타임지 추천서 선정	643	27,503
	대림동 여성경찰의 대처 논란	248	31,180
	이수역 술집 남녀 쌍방 폭행 사건	612	43,113
LGBTQ	변희수 하사의 여군 복무 청원	121	10,932
	퀴어 축제 서울시청 앞 개최 추진	1,125	54,716
	숙명여대 트랜스젠더 지원자 논란	46	6,047
Total		3,111	191,887

수집된 댓글은 형태소 분석으로 토픽 모델링을 위한 전처리 과정을 거쳤다. 토픽 모델링은 기본적으로 문서와 단어 혹은 문서 내 단어의 공기 관계를 바탕으로 주제를 추정하기 때문이다. 따라서 본 연구에서의 문서인 댓글을 형태소 분석으로 단어 단위로 토큰화했다. 이러한 형태소 분석에는 Kiwi(Korean Intelligent Word Identifier)가 사용됐다. Kiwi는 세종계획 말뭉치와 모두의 말뭉치에 대한 학습을 거쳐 문장 분리와 오타 교정 등 웹 텍스트에 대한 정확한 분석(웹 텍스트: 약 87%, 문어 텍스트: 약 94%)이 기대되는 기능을 제공할 뿐만 아니라 속도 또한 다른 형태소 분석기에 비해 빠르다는 장점이 있다(Lee, 2022). 세종 품사 태그 체계에 따라 형태소 분석을 진행해 댓글의 주제를 파악하는데 유용하지 않은 품사(조사, 어미 등)의 형태소와 부호 또는 특수문자 등은 제거하고 체언(명사/대명사)과 용언(동사/형용사)에 해당하는 형태소만 남겼다.²⁾ 그 결과 댓글의 평균 형태소 개수는 35.62($SD = 33.28$)이고 가장 긴 댓글의 형태소 개수는 276개, 최소는 1개였다. 이 중 90% 이상의 댓글이 7개 이상의 형태소로 구성되어 있었다.

이상의 과정으로 처리된 데이터로 토픽 모델링을 이용해 이슈별로 댓글을 분류하고 모델의 정확성을 살피기 위해 지도기계학습 모델의 성능 평가 방식을 활용했다. 우선, 수집된 댓글의 수가 이슈별로 차이가 날 뿐만 아니라 댓글의 길이 역시 제각각이기 때문에 안정적이고 균형있는 분류 성능 평가를 위해 다음의 데이터 처리 작업을 시행했다. 우선, 본 연구가 댓글이라는 단문에서 주제를 포착하고 이에 따라 댓글을 분류하는 방법의 구축과 검증은 목표하고 있으나 주제를 파악하기에는 너무 짧은 댓글은 분석에서 제외했다. 가령, 댓글이 “극혐”과 같은 인터넷 용어나 “대박”과 같은 감탄사로만 이뤄져 있을 때 이 댓글에서 감정을 측정할 수는 있어도 주제를 파악할 수는 없기 때문이다. 따라서 본 연구에서는 댓글에 사용된 형태소가 8개 미만인 경우에는 분석에서 제외했다.³⁾ 반대로 댓글의 길이가 너무 긴 경우 역시 본 연구의 목표를 벗어난다고 판단해 형태소가 112개(상위 5%) 이상인 댓글 또한 분석에서 제외했다.

2) 익명의 심사위원께서 지적하였듯, 토픽 모델링을 위한 형태소 분석은 동의어 처리도 포함하는 경우가 일반적이다(Maier et al., 2018). 같은 의미이지만 변형된 형태의 단어를 동일하게 다룰 때 보다 정확한 결과 도출이 가능하기 때문이다. 특히, 댓글의 경우 언어 번이가 매우 빈번하게 일어나기 때문에 동의어 처리는 결과에 상당한 영향을 미칠 것으로 예상된다. 그러나 본 연구는 이러한 댓글의 조건에서도 주제 포착을 위한 토픽 모델링이 신뢰할 만한 결과를 도출한 것인지를 검증하고자 동의어 처리를 최소화한 후 분석 결과를 평가했다. ‘이슈’에 따라 댓글을 분류함에 있어 토픽 모델링 성능 평가에 보다 엄격한 기준을 적용하고자 한 것이다. 그러나 댓글 분석에 토픽 모델링을 적용하는 후속 연구는 결과의 신뢰성 확보를 위해 동의어 처리를 포함할 것을 제안한다. 적절한 심사의견에 감사드린다.

3) 분석에 포함된 댓글의 길이 기준은 수집된 댓글의 길이 분포상 90% 이상의 댓글이 7개보다 많은 형태소로 구성되어 있었다는 점을 고려해 결정했다. 그러나 이 기준에 대한 평가가 필요함은 물론 단문 데이터 성격에 따라 길이 분포 역시 변할 수 있으므로 후속 연구에서 단문의 범위가 모델 성능에 미치는 영향에 대해 검증해 볼 필요가 있겠다.

이러한 댓글 정제 후 각 이슈별로 댓글을 1,000건씩 무작위 추출했다. 이슈별 댓글의 수가 차이날 경우 기계학습은 물론 모델 평가에 있어 편향이 일어날 수 있기 때문이다. 그 결과, 이슈별 뉴스 댓글을 1,000건씩 총 9,000건의 댓글로 데이터를 구성했다. 그리고 토픽 모델링을 위한 문서-단어 행렬로의 구조화를 다음과 같이 진행했다. 우선, 단어 빈도에 기반하여 주제를 추정하는 토픽 모델링에 있어 너무 드물게 출현하거나 반대로 거의 모든 문서에서 등장하는 단어들에 의해 확률적 추론의 정확성이 떨어질 수 있다(Maier et al., 2018). 이에 토큰화 결과 전체 댓글의 0.5% 이하 또는 99% 이상의 댓글에서 등장하는 형태소들은 제거했다. 그리고 문서로서의 댓글에 토큰화로 추출된 단일 형태소를 단어로 하는 행렬(문서 9,000건 × 단어 1,653개)을 마련했다. 그리고 바이그램 효과를 살피기 위해 각 댓글에서 토큰화된 형태소들 중 인접한 두 형태소들로 바이그램을 추출하여 이를 단일 형태소와 함께 단어로 포함된 행렬(문서 9,000건 × 단어 3,453개)도 준비했다.

2) 토픽 모델의 초모수 조정 및 성능 평가

본 연구가 댓글 분석에 사용한 토픽 모델은 LDA와 BTM이다. 두 모델은 모두 깃스 샘플링 방식으로 모수를 충분히 반복해서 추정해야 신뢰성 있는 결과를 도출할 수 있어 반복 횟수를 2,000회로 설정했다. 또한 주제-단어 관계의 사전 분포에 관계하는 η 값은 상대적으로 그 중요성이 크지 않아 조정 작업을 거치지 않고 각 토픽 모델의 기본값으로 고정했다. 이후 토픽 모델의 성능을 평가하는 기준인 혼잡도와 가능도를 토대로 각각 LDA와 BTM의 초모수 K 와 α 를 조정하여 최적의 모델을 선정했다. 특히, 초모수 α 값이 0.01에서 10까지로 증가함에 따른 토픽 모델의 성능을 탐색함과 동시에 K 값이 9개 이슈(최적값)를 중심으로 최소 3개에서 30개로 설정할 경우의 성능을 비교했다. 또한 K 값이 최적값보다 낮거나 높을 경우에 있어 성능과 타당성이 어떠한지 살펴봤다. 주제 추정을 위한 단어 목록이 단일 형태소로만 구성된 경우와 바이그램을 포함한 경우와의 비교 또한 진행했다. 이를 위한 모델의 성능 평가는 토픽 모델로 추정(LDA) 혹은 예측(BTM)한 댓글의 주제 분포에 따라 가장 높은 확률의 주제로 댓글을 분류한 결과가 댓글이 만들어진 이슈를 반영하는지를 측정하는 방식으로 이루어졌다. 왜냐하면 댓글은 9개 이슈에 대한 뉴스 기사에서 수집되었기 때문에 토픽 모델링으로 추정한 댓글의 주제가 이슈를 적절하게 반영하고 있다면 주제에 따른 분류가 이슈별 구분과 일치할 것이기 때문이다. 이러한 지도 기계학습 모델의 평가 방식으로 본 연구는 토픽 모델의 성능을 검증했다. 성능 평가에 사용한 지표는 정확도(accuracy), 정밀도(precision), 재현성(recall), 그리고 F1-점수다(F1-score).⁴⁾

4. 결과

1) 초모수 K 와 α 값의 조정이 모델 성능에 미치는 영향

첫 번째 연구질문은 토픽 모델링의 초모수 K 와 α 에 대한 조정이 댓글의 주제 포착과 이슈 분류에 어떠한 영향을 미칠 것인지에 대한 것이었다. 즉, BTM과 LDA의 초모수 조정 효과를 비교하는 것이다. 이에 우선 BTM과 LDA의 초모수 조정에 따른 모델 성능을 각각 가능도와 혼잡도를 기준으로 평가했다. <Figure 4>는 다양한 초모수 K 값과 α 값에 따른 BTM의 가능도를 보여준다. 분석 결과, 가능도가 높아질수록 모델 성능이 상승한다는 점을 감안할 때 K 값이 커질수록 모델 성능이 증가하고 있음이 나타났다. 그런데 K 값이 9를 지나면서 성능 상승의 기울기가 크게 꺾이는 형태를 보임에 따라 주요한 주제의 수가 9개임을 유추할 수 있다. 반면 α 값의 변화에 따른 BTM의 성능 변화는 미미했다.

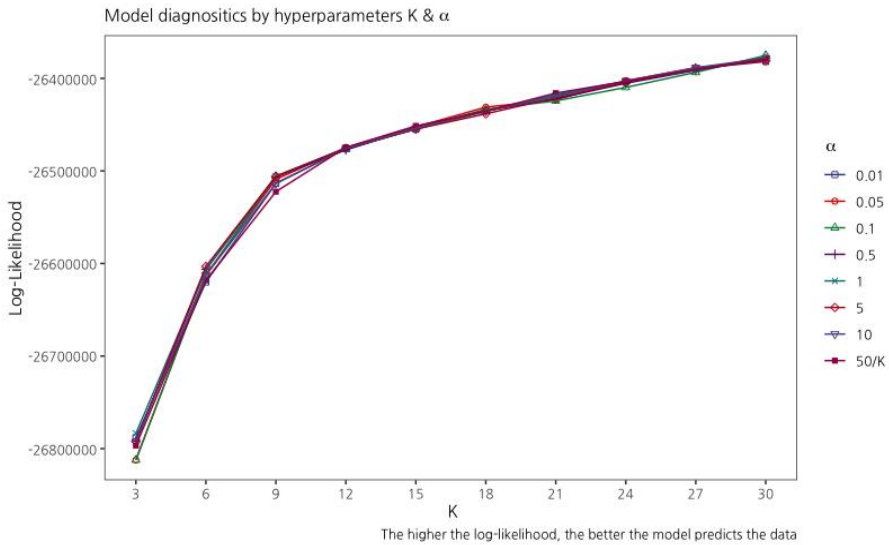


Figure 4. How BTM's performance changes with hyperparameter tuning (including bigrams in the word list of the model)

- 4) 정확도 = (참 긍정 + 참 부정) / (참 긍정 + 거짓 부정 + 거짓 긍정 + 참 부정)
 정밀도 = 참 긍정 / (참 긍정 + 거짓 긍정)
 재현율 = 참 긍정 / (참 긍정 + 거짓 부정)
 F1-점수 = (2 × 정밀도 × 재현율) / (정밀도 + 재현율)

초모수 조정이 LDA의 성능에 미치는 영향은 어떠할까? <Figure 5>에서 볼 수 있듯이 LDA의 초모수 α 값은 BTM에 비해 상당히 큰 영향을 미치고 있었다. α 값을 0.1 이하로 낮게 설정할 경우 그 이상의 값 설정에 비해 눈에 띄는 모델 성능 향상이 나타난 것이다. 특히, LDA 모델이 기본적으로 α 값을 $50/K$ 로 설정하는 경우가 많다는 점을 고려할 때 이 경우의 모델 성능과 0.1 이하의 모델 성능의 차이는 주목할 만하다. 반면 LDA 역시 K 값이 증가함에 따라 모델 성능이 향상되고 있었다. 혼잡도의 경우 값이 낮아질수록 모델 성능이 좋아짐을 의미하는데, 무엇보다 LDA도 K 값이 9를 지나면서 성능 향상의 정도가 크게 줄어들어 BTM과 마찬가지로 최적값을 유추하는데 무리가 없는 것으로 나타났다. 문제는 α 값이 0.1보다 크게 설정된 모델의 경우에 이러한 성능 저하 경향성이 눈에 띄게 줄어들어 BTM과 마찬가지로 최적값을 유추하는데 무리가 없는 것으로 나타났다. 문제는 α 값이 0.1보다 크게 설정된 모델의 경우에 이러한 성능 저하 경향성이 눈에 띄게 줄어들어 BTM과 마찬가지로 최적값을 유추하는데 무리가 없는 것으로 나타났다. 결국, LDA는 α 값이 모델 성능에 직접적인 영향을 미칠 뿐만 아니라 초모수 조정에 따른 혼잡도에 기반하여 K 값을 선정하는 과정에도 중요하게 작용하고 있다는 의미다. 따라서 댓글과 같은 단문에 대해 LDA를 활용하여 주제를 분석하고자 할 때는 0.1 이하의 α 값을 설정해야 좋은 성능을 기대할 수 있겠다.

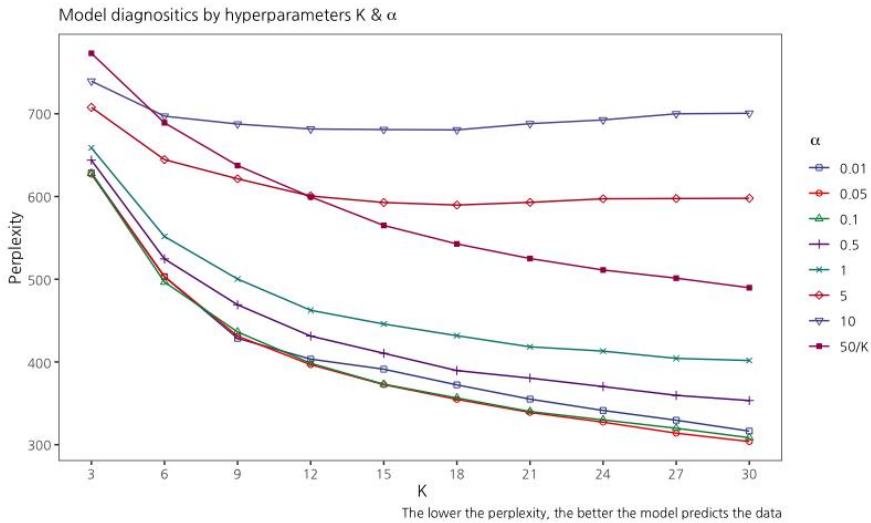


Figure 5. How LDA's performance changes with hyperparameter tuning (including bigrams in the word list of the model)

2) K값 선정이 모델 성능에 미치는 영향

두 번째 연구질문은 최적의 K 값보다 작거나 크게 K 값이 설정된 경우 토픽 모델의 성능이 어떻

게 나타날지에 대한 것이었다. 이에 최적의 K 값인 9를 기준으로 K 값이 3으로 설정된 모델과 30으로 설정된 모델의 성능을 측정 후 비교했다. 이때 BTM의 α 값은 기본값인 $50/K$ 로, LDA의 α 값은 0.01로 설정했다. 우선, 9개 주제를 추정된 토픽 모델의 댓글 분류 정확도는 BTM의 경우 정확도가 90.14%였고, LDA는 91.15%로 나타났고 정밀도, 재현율, F1-점수의 평균 역시 비슷한 수치로 나타났고 표준편차 역시 크지 않아 모든 이슈에 대해서 고른 분류 정확성을 보이고 있었다(〈Table 2〉 참조). 이는 단문 분석을 위해 제안된 BTM뿐만 아니라 α 값 조정을 거친 LDA 역시 댓글을 이슈별로 분류하는데 있어 매우 뛰어난 성능을 보여주고 있음을 밝혀낸 것이다.⁵⁾ 오히려 LDA가 BTM에 비해 약간 더 높은 정확도와 균형을 보여주었는데, 이 결과는 LDA 성능이 α 값 조정이 매우 민감하며 토픽 모델링에 있어 텍스트의 구조와 성격을 고려한 초모수 조정이 중요한 과정임을 제시한다.

Table 2. Performance Evaluation of Topic Models that Categorized News Comments on 9 Issues into 9 Topics

Metric	BTM ($K = 9, \alpha = 50/K$)	LDA ($K = 9, \alpha = 0.01$)
Accuracy	90.14%	91.15%
Precision	$M = 90.44\%, SD = 5.09\%$	$M = 91.27\%, SD = 3.91\%$
Recall	$M = 90.14\%, SD = 5.67\%$	$M = 91.16\%, SD = 3.60\%$
F1-score	$M = 90.12\%, SD = 3.64\%$	$M = 91.16\%, SD = 3.02\%$

다음으로 K 값을 3으로 설정한 토픽 모델의 성능이다. 9개 이슈는 혐오 대상에 따라 3개의 유형으로 유목화가 가능했는데, 이러한 주제 구조의 특성을 토픽 모델이 반영해 3개의 혐오 이슈별로 분류되는지를 평가한 것이다. 〈Table 3〉에서 볼 수 있듯, BTM의 정확도는 72.22%였고 LDA는 63.11%로 토픽 모델 성능의 상당한 저하가 나타났다. 물론 BTM의 정확도가 LDA에 비해 9.11% 포인트 높아 더 나은 성능임을 보였지만 정밀도와 재현율 간의 평균 차이가 7% 포인트 이상 나타났고 표준편차 역시 크게 증가해 댓글 주제에 따라 이슈별로 분류하는 데 있어 불균형이 발생하고 있음이 드러났다.

5) 최적의 BTM과 LDA 모델로 각각 추정된 9개 주제는 의미적 일관성과 차별성 측면에서도 이슈를 잘 반영하고 있었다. 〈Appendix 1〉은 두 모델 모두 각 주제에 대해 높은 확률로 할당된 단어(상위 15개)들이 해당 이슈에 기반한 의미적 일관성을 나타내고 있을 뿐만 아니라 다른 이슈와의 차별성 또한 보여주고 있음을 보여준다. 주제에 따른 댓글 분류 역시 이슈 분포를 잘 반영해 고른 비율을 보여 두 토픽 모델의 정확하고 균형있는 분류 성능을 제시하고 있다.

Table 3. Performance Evaluation of Topic Models that Categorized News Comments on 9 Issues into 3 Topics

Metric	BTM ($K = 3, \alpha = 50/K$)	LDA ($K = 3, \alpha = 0.01$)
Accuracy	72.22%	63.11%
Precision	$M = 79.45\%, SD = 17.61\%$	$M = 60.89\%, SD = 12.50\%$
Recall	$M = 72.21\%, SD = 37.48\%$	$M = 63.10\%, SD = 29.19\%$
F1-score	$M = 68.11\%, SD = 20.05\%$	$M = 61.22\%, SD = 20.45\%$

〈Figure 6〉은 이러한 분류 성능 평가를 시각화한 결과로써, 댓글을 BTM에 따라 3개의 주제로 예측해 분류한 결과가 댓글이 달린 뉴스 이슈와 일치되는 정도를 표현한다. 가령, “Topic 1”이 가장 높은 확률로 할당된 댓글이 “변희수 하사의 여군 복무 청원”이나 “퀴어 축제 서울시청 앞 개최 추진”, “숙명여대 트랜스젠더 지원자 논란” 등 성소수자에 대한 혐오와 관련한 이슈의 뉴스에 달린 댓글인지의 여부를 따져 일치도를 계산해 수치가 높아질수록 채도가 높아지게 된다. 분류 결과는 “Topic 1”과 “Topic 3”이 각각 성소수자 혐오 이슈와 여성 혐오 이슈의 뉴스 댓글들을 비교적 정확하게 반영하고 있었으나 외국인 혐오 이슈로 분류되어야 할 “Topic 2”의 댓글들이 “샘 오취리의 ‘관짝소년단’ 비판” 이슈의 댓글들을 제외하고 다른 혐오 유형의 이슈들로 분류되는 경향이 높았다. 이러한 주제 간의 분류 정확성 불균형은 최적값보다 낮은 K 값의 토픽 모델이 댓글을 비슷한 유형의 이슈별로 묶어주는 데 효과적이지 않음을 의미한다. 결국 토픽 모델링에서 주제 개수의 선정이 이상적인 주제 구조를 충분히 반영할 수 없도록 낮게 설정될 경우 의미적 차별성이 모호한 주제들이 추정되며 이슈에 따른 댓글 분류 타당성이 낮아질 수 있음이 드러났다.

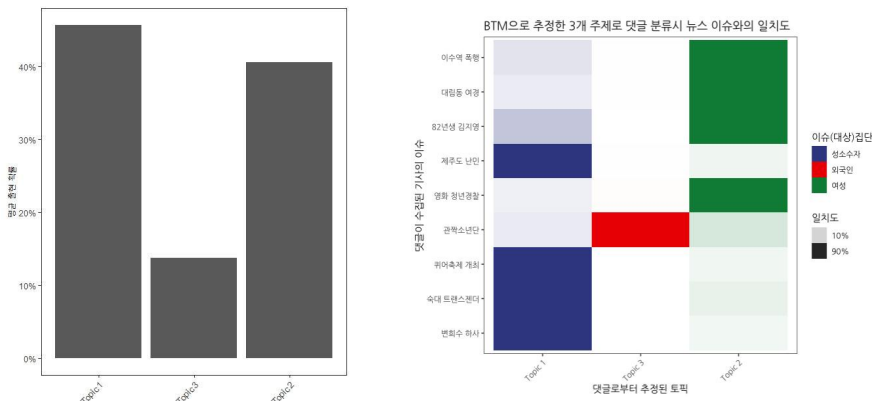


Figure 6. The left graph shows the average assignment probability distribution for comments on three topics. The right graph visualizes the issue prediction performance of comments classified into the three topics by the BTM set to K of 3.

그렇다면 K 값을 최적값보다 높게 설정하면 모델의 성능은 어떻게 될까? K 값을 30으로 설정한 토픽 모델이 추정한 주제들 중 댓글 전체에서 출현할 평균 확률이 높은 상위 9개의 주제를 선정하고 이에 따라 댓글을 분류했을 때 이슈와의 일치도를 평가했다. 다시 말해, 토픽 모델로 추정한 30개의 주제들이 댓글마다 할당되는 확률의 평균값을 계산하여 가장 두드러지게 출현하는 주제 9개를 선정한 후, 각 댓글에서 이들 9개 주제 중 최대 확률 주제를 확인해 이에 따라 분류하고 그 결과를 살핀 것이다. 이 분석은 K 값이 크더라도 토픽 모델이 9개 이슈로 구성되어 있는 댓글에 대해 이들 9개의 주제가 두드러진 확률 분포를 추정할 것이고 나머지 주제들은 이슈 구분에 미미한 영향을 미칠 것이라는 가정에 기반한다.

Table 4. Performance evaluation of topic models that categorized news comments on 9 issues into the top 9 most probable topics out of 30 topics

Metric	BTM ($K = 30, \alpha = 50/K$)	LDA ($K = 30, \alpha = 0.01$)
Accuracy	83.20%	71.13%
Precision	$M = 84.77\%, SD = 8.89\%$	$M = 72.23\%, SD = 7.44\%$
Recall	$M = 83.20\%, SD = 13.06\%$	$M = 71.13\%, SD = 11.78\%$
F1-score	$M = 82.98\%, SD = 7.73\%$	$M = 70.84\%, SD = 5.68\%$

〈Table 4〉에 표시된 것과 같이, 이 모델의 성능은 BTM과 LDA 모두 앞서 언급한 낮은 K 값 모델에 비해 향상된 성능을 보여주고 있었다. 특히, BTM의 성능 향상이 눈에 띄는데 정확도는 83.20%로 약 11% 포인트 상승했으며 정밀도와 재현율, F1-점수의 평균 모두 비슷한 수치로 증가함과 동시에 표준편차 역시 크게 줄어들어 이슈 간 분류 성능의 불균형이 발견되지 않았다. LDA 역시 정확도가 71.13%로 약 8% 포인트 향상됐고 다른 평가 지표의 평균과 표준편차 역시 두드러진 성능 증가를 보여주었다. 그러나 BTM이 LDA보다 전반적으로 더 정확한 성능을 보여주고 있어 K 값 선정이 미치는 영향이 제한적일 뿐만 아니라 최적값보다 낮은 K 값보다 높은 K 값으로 주제 분포를 추정할 때 더 정확한 분류가 가능함이 드러났다.

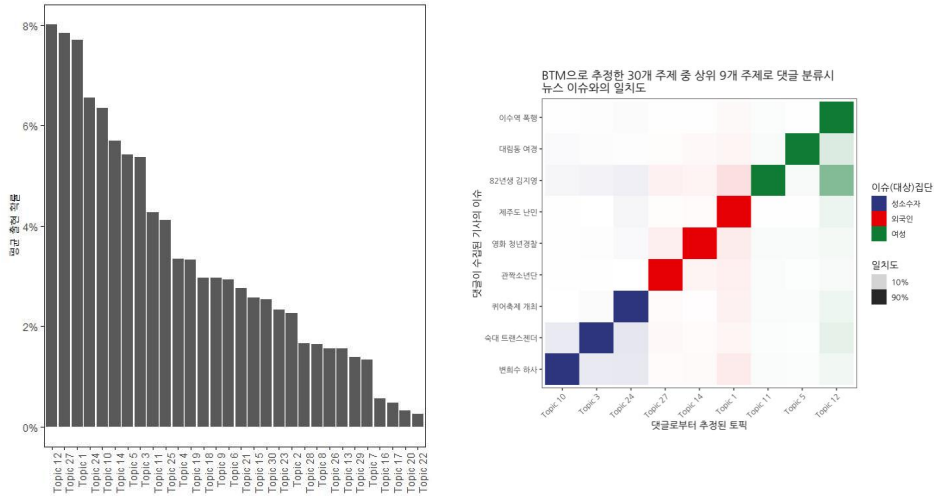


Figure 7. The left graph shows the average assignment probability distribution for comments on 30 topics. The right graph visualizes the issue prediction performance of comments classified into the top 9 topics with average probability by the BTM set to K of 30.

이러한 BTM의 성능을 시각화한 <Figure 7>에서 확인할 수 있듯, 토픽 모델의 주제 분포 상 상위 9개 주제가 대체적으로 이슈를 잘 반영하고 있었을 뿐만 아니라 이슈 혼동이 나타난 경우도 같은 혐오 유형의 이슈 사이에서 발생했다. 예를 들어, “Topic 12”가 가장 두드러진 댓글은 “이수영 폭행” 이슈에 대한 것들이 대부분이었지만 “82년생 김지영” 이슈나 “대림동 여경” 이슈의 댓글이 일부 있는 것으로 나타났다. 그러나 이 이슈들은 모두 여성 혐오 논란과 관련하고 있기 때문에 BTM으로 댓글에서 추정된 주제들은 의미적 유기성을 보인다고 평가할 수 있겠다. 결국, 토픽 모델의 K 값을 선정함에 있어 최적값보다 높게 설정할 경우의 손실은 낮게 설정할 경우 보다 낮으며 특히 BTM은 K 값으로 인한 모델 성능의 변화가 LDA보다 적어 댓글과 같은 단문에서는 효과적인 토픽 모델임을 확인했다.

3) 바이그램이 토픽 모델 성능에 미치는 영향

마지막 연구질문은 단어 목록에 바이그램을 추가한 토픽 모델의 성능 변화에 관한 것이다. 이에 단일 형태소로만 단어 목록을 구성한 LDA와 BTM으로 댓글의 주제를 추정하고 분류한 결과의 이슈 일치도와 단어 목록에 바이그램을 추가한 토픽 모델의 성능을 비교했다. <Table 5>에서 확인할 수 있듯, BTM과 LDA 모두 바이그램을 포함했을 때 성능이 향상됐고 특히 BTM보다 LDA의 성능 개선이 두드러졌다. BTM의 정확도는 단일 단어 기반의 토픽 모델은 88.00%인

반면 바이그램 포함 모델은 90.14로 2.14% 포인트 증가에 그쳤고, 정밀도와 재현율, F1-점수의 평균 역시 비슷한 수준의 상승이 나타났다. 반면 LDA의 정확도는 바이그램을 포함할 경우 83.70%에서 91.15%로 7.45% 포인트 증가했으며 다른 지표의 평균 역시 비슷한 수준의 개선은 물론 표준편차는 더 줄어드는 결과가 나타난 것이다. 이러한 차이는 기본적으로 바이그램을 이용해 주제를 추정하는 BTM에는 바이그램의 포함이 미치는 개선 효과가 상대적으로 미미했지만, LDA의 경우 문서에서 등장하는 단일 단어는 물론 바이그램의 빈도를 통해 주제 분포를 추정하는 방법으로 도출한 댓글의 주제가 이슈를 보다 정확하게 반영했기 때문으로 추론된다. 결국, 댓글에서 출현하는 단어들은 물론 단어들의 조합인 바이그램이 주제 추정에 효과적임이 드러난 것이다.

Table 5. Performance Evaluation of Topic Models with a Word List of Only Unigrams and Topic Models with the Addition of Bigrams

Algorithm	Metric	Unigram only	Bigram included
BTM $K = 9$ $\alpha = 50/K$	Accuracy	88.00%	90.14%
	Precision	$M = 88.28\%, SD = 5.10\%$	$M = 90.44\%, SD = 5.09\%$
	Recall	$M = 87.97\%, SD = 6.34\%$	$M = 90.14\%, SD = 5.67\%$
	F1-score	$M = 88.00\%, SD = 4.05\%$	$M = 90.12\%, SD = 3.64\%$
LDA $K = 9$ $\alpha = 0.01$	Accuracy	83.70%	91.15%
	Precision	$M = 83.75\%, SD = 4.59\%$	$M = 91.27\%, SD = 3.91\%$
	Recall	$M = 83.72\%, SD = 5.90\%$	$M = 91.16\%, SD = 3.60\%$
	F1-score	$M = 83.68\%, SD = 4.76\%$	$M = 91.16\%, SD = 3.02\%$

또한, 바이그램이 댓글의 주제 추정을 개선하는 효과는 이슈에 따라 차이가 나타나지 않았다. 오히려 <Figure 8>에서 볼 수 있듯, 단일 단어만으로는 댓글에서 추정한 주제가 이슈 차별성을 반영하지 못하는 경우가 더 빈번했다. BTM과 LDA 모두 댓글의 바이그램을 고려해 주제를 추정한 모델이 이슈별로 댓글을 정확하게 분류한 것이다. 특히, 이미 언급했듯 BTM보다 LDA의 경우에 댓글 분류 모델의 성능 향상이 두드러졌는데 이 결과는 문서의 특성을 고려해 초모수 α 를 조정한다면 바이그램을 포함한 모델은 BTM보다 더 정확한 주제 추정과 분류가 가능함을 보여준다. 나아가 LDA를 이용한 주제 분석에는 단일 단어와 함께 바이그램이 포함된 단어 목록이 마련되어야 함을 시사한다. 정리하자면, 댓글의 주제 분석에 있어 BTM은 안정적인 성능으로 신뢰성이 확보된 반면 LDA는 초모수 조정과 바이그램 포함 등의 처리 과정이 동반될 때 타당성있는 결과가 기대된다 할 수 있겠다.

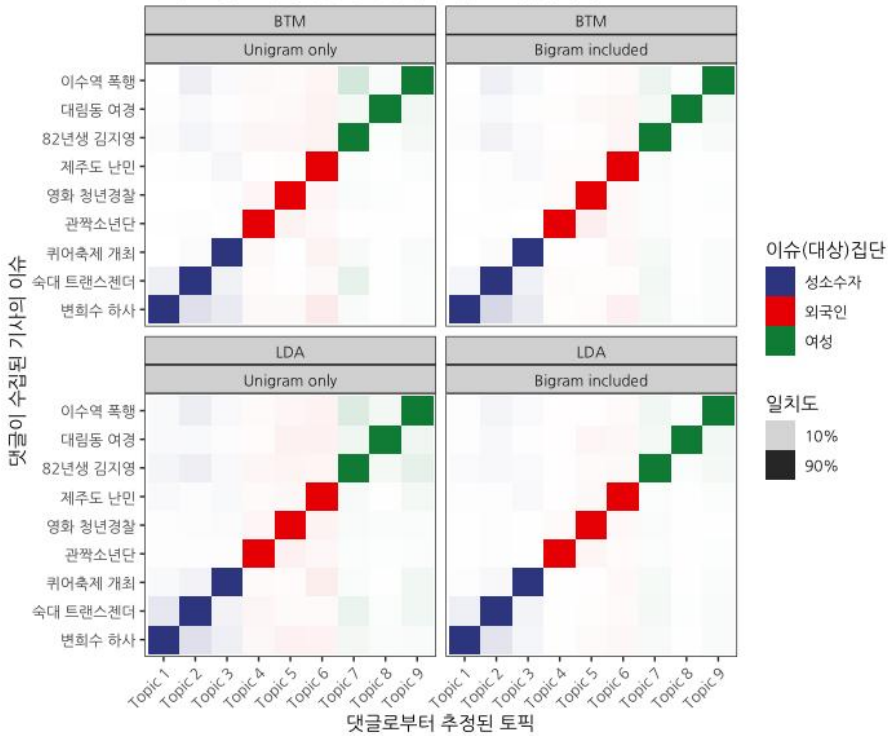


Figure 8. Visualization of comment classification performance for topic models with a word list of only unigrams and topic models with the addition of bigrams

5. 논의 및 결론

본 연구는 디지털 미디어 플랫폼의 증가와 함께 빠르게 확산되고 있는 단문 텍스트에 대해 토픽 모델링으로 주제를 파악하는 방법의 타당성을 검증했다. 대표적 단문 텍스트인 댓글의 주제적 특성을 도출해 온라인 담론의 내용과 방향을 포착해야 할 필요성이 커지고 있지만, 전통적 토픽 모델인 LDA의 한계로 인해 충분한 연구가 진행되지 못했기 때문이다. 확률 추정을 기반으로 한 토픽 모델의 통계적 가정으로 인해 단어들의 낮은 빈도수로 이루어진 단문에 대해서 주제를 정확히 추정하는데 한계가 있다는 점이 주된 이유였다.

이러한 한계를 극복하는 방안으로 제시된 토픽 모델이 BTM이다(Yan et al., 2013). BTM을 이용해 댓글의 주제적 특성을 도출하는 연구가 점차 주목을 받고 있지만(예: 안순태 외, 2022), 정작 한글로 구성된 댓글에 대해 BTM이 도출한 주제 구조의 타당성 검증과 모델의 신

되성 분석은 거의 이뤄지지 않았다. 본 연구는 이에 대표적 토픽 모델인 LDA와 함께 BTM을 이용해 댓글의 주제 도출과 이에 따른 분류가 어떠한 타당성을 보이는지를 검증하고, 모델의 초모수 조정과 단어 처리에 따른 성능 변화는 어떠한지를 탐색했다. 나아가 인터넷 환경에서의 보편적인 텍스트 구조인 대단위 단문 자료를 주제별로 분류할 수 있는 방법적 고민을 토대로 토픽 모델링을 통한 댓글 주제 분석의 신뢰성을 더하고자 하였다.

이를 통해 본 연구가 밝혀낸 결과와 그 함의는 다음과 같다. 첫째, 댓글과 같은 온라인 단문의 주제 분석을 위해 제안된 BTM이 한글 기반의 뉴스 댓글을 관련 이슈별로 분류하는데 정확한 성능을 나타내는지 검증하고 이를 LDA 성능과 비교함으로써 방법적 신뢰성 확보에 기여했다. BTM은 뉴스가 다루는 사건이나 논란에 대해 찬성 혹은 반대의 의견 및 정서로 표현되는 댓글을 '이슈'별로 분류하는데 정확하고 안정적인 성능을 보였는데 적절한 K 값이 주어질 경우 90% 이상의 분류 정확도를 보였을 뿐만 아니라 초모수 조정에 따른 성능 변화의 폭이 적었다. '이슈'에 대한 다양한 의견이 나타나는 댓글의 주제를 포착하고 이에 따라 같은 '이슈'에 대한 댓글끼리 분류하는 방법에 있어 BTM이 효과적임을 확인한 것이다.

물론, 본 연구의 분석 대상이 혐오 이슈가 불거진 사안의 포털 뉴스 댓글로 국한되었기 때문에 다른 내용과 형식의 텍스트 데이터에서도 동일한 결과가 나타날 것인지에 대한 의문이 생길 수 있다. 하지만 토픽 모델링을 이용해 댓글을 성소수자와 외국인, 그리고 여성이라는 대상 차원에서의 속성뿐만 아니라 각 집단 내에서도 '이슈'에 따라 정확하게 분류했기 때문에 내용의 차이가 모델 성능에 미칠 영향은 크지 않을 것이라는 추론이 가능하다. 또한 댓글의 일반적 형식을 고려해 길이 상 이상치(outlier)에 해당하는 사례들을 제거하였고, 온라인 텍스트에서 단어를 추출하는데 정확한 성능을 보여주는 형태소 분석기 Kiwi를 사용했기 때문에 형식적 측면에서도 연구 결과의 외적 타당성은 높다고 평가할 수 있다.

또한 본 연구는 LDA는 댓글과 같은 단문의 주제 분석에 적합하지 않다는 인식과 달리 0.1 이하의 α 값과 조정 작업을 통한 최적의 K 값이 주어질 경우 분류에 우수한 성능을 보임을 확인했다. 토픽 모델링에 있어 K 값에 대한 고민과 초모수 조정 작업은 필수적으로 이루어지는 반면 α 값에 대한 검토는 상대적으로 미미했던 것이 사실이다. 그러나 본 연구의 결과는 분석 대상인 텍스트의 성격에 따라 K 값보다 α 값이 모델 성능에 더 중요한 영향을 미침을 제시한다. 더욱이, K 값이 최적값보다 높게 설정될 경우에는 주요 주제의 분포를 추정하고 이에 따라 댓글을 분류하는데 있어 성능 저하가 크지 않았기 때문에 오히려 α 값에 대한 관심이 더 요구된다 할 수 있다. 이러한 차원에서 토픽 모델의 K 값이 주제 분류 성능에 미치는 영향을 살펴본 결과, 실제 주제의 수보다 낮게 K 값을 설정할 경우에는 도출된 댓글의 주제가 이슈를 정확하게 반영하지 못했지만

높게 설정할 경우 주제에 따른 댓글 분류가 이슈의 차별성을 잘 보여주는 것으로 나타났다. 높은 K 값의 토픽 모델이어도 α 값의 조정을 통하면 문서의 주요 주제를 포착하고 관련성 적은 주제에 대해서는 낮은 분포를 추정하기 때문이다. 이 결과는 토픽 모델링 활용에 있어 최적의 K 값을 찾기 위한 고민을 효용의 측면에서 고려해 볼 필요가 있음을 제시한다.

나아가, 본 연구는 토픽 모델링에 있어 바이그램을 추가하는 처리 과정이 성능 개선에 효과적임을 확인했다. 문서의 주제가 단어 출현에 영향을 미치는 과정에 단일 형태소의 단어 사용뿐만 아니라 인접한 두 단어의 쌍으로 구성된 바이그램에도 영향을 미치기 때문이다. 다시 말해 문서에서 사용된 단일 단어들뿐만 아니라 주제를 추정하기 보다는 단어들끼리 함께 사용되는 패턴을 학습할 때 토픽 모델의 성능이 개선됨을 확인했다.

마지막으로, 토픽 모델링으로 도출한 댓글의 주제는 댓글이 달린 게시글의 이슈를 잘 반영하고 있었다는 점이다. 즉, 댓글이 '무엇에 대해 말하고 있는가(What to talk about)'를 파악하는 방법은 온라인 공간에서 논쟁이 일어나고 있는 '이슈'의 내용은 물론 그 크기에 대해서도 측정할 수 있다는 의미다. 그리고 여기에 토픽 모델링의 방법이 효과적임을 발견했다. 사실 데이터 과학 분야에서 제안된 토픽 모델링이 미디어 연구 분야에서 논의되어 온 의제나 프레임 측정하는 방법으로 타당한지에 대한 의문은 지속되고 있다(DiMaggio, Nag, & Blei, 2013; Maier et al., 2018; Ying et al., 2022). 토픽 모델링으로 도출하는 토픽은 결국 문서에서 자주 등장하는 단어들로 추정되는데 의제나 프레임은 단어 빈도수에 따라서만 측정되는 개념이 아니기 때문이다.

물론, 미디어가 뉴스보도를 통해 공중이 '무엇에 대하여 생각하는가(What to think about)'를 결정한다는 의제설정 효과(McCombs & Shaw, 1972)에 따르면 의제란 언론이 중요하게 다루는 사건이나 이슈로 정의할 수 있고, 이는 결국 뉴스 기사에서 빈번하게 등장해 대중이 자주 노출되는 용어나 키워드들로 의제를 추론할 수 있다는 주장이 가능하다. 실제 비슷한 관점에서 대단위 뉴스 기사를 토픽 모델링으로 분석해 언론의 의제를 파악하는 연구가 활발히 진행되고 있다(이세미·유승의·안순재, 2020; 이종혁·길우영, 2019; 함승경·정사강·김은영, 2022; Grimmer & Stewart, 2013; Koltsova & Koltcov, 2013).

그러나 뉴스 의제를 단순히 기사의 주제가 아닌 언론이 뉴스로 다루는 이슈나 사건들을 통해 다양한 층위에서 드러나는 공중의 주요 안건이라는 관점으로 보았을 때 토픽 모델링으로 추정한 토픽만으로 의제를 파악한다고 하기에는 무리가 있다. 뉴스 기사는 대체로 하나의 의제를 다루는데 같은 기사에 복수의 주제가 존재한다고 가정하는 토픽 모델링은 의제 분석에 개념적으로 타당하지 않기 때문이다. 미디어 연구 분야에서 제시된 언론 의제는 뉴스 기사의 헤드라인과 지면상 위치 또는 보도 순서 등을 결정하는 저널리즘과 뉴스의 가치 등에도 영향받기 때문에 단순

히 많은 빈도로 출현하는 단어들로 의제를 포착하는 접근은 타당하지 않다는 비판도 가능하다.

프레임 분석에 있어서도 토픽 모델링은 대안적 방법으로 주목받아 왔다. 언론보도에 있어 프레임은 특정 주장이나 관점, 그리고(또는) 일부 사실 등을 강조하는 방식에 따른 사건이나 이슈의 선택적 제시로 정의된다(Entman, 1993). 즉, 뉴스 프레임은 기사에서 해당 사안이 어떻게 다뤄지고 있는가를 나타내는 해석적 틀로써 “선택과 강조, 그리고 전달의 방식”으로 볼 수 있다(Gitlin, 1980). 특히, 프레임은 비슷한 커뮤니케이션 상황에서 동일하거나 유사한 단어와 기호를 사용해 같은 대상이나 속성을 반복적으로 상기시키는 뉴스의 패턴으로 드러난다(Entman, Matthes, & Pellicano, 2009). 이러한 관점에 기반하여 토픽 모델링으로 기사에서 도출한 주제들의 군집(cluster)으로 프레임을 파악하고자 한 최근의 연구 동향이 관심을 끌고 있다(이준웅·김성희, 2018; 조민정·이신행, 2021; Walter & Ophir, 2019). 그러나 다른 한편으로는 주제들의 군집화에 있어 연구자의 주관적 선택이 개입할 뿐만 아니라 의제와 마찬가지로 단순히 단어의 빈도수가 아닌 사용 방식과 강조의 맥락 등을 전체적으로 고려할 때 프레임을 정확히 해석할 수 있다는 점에서 방법적 신뢰성에 대한 의문이 제기될 수 있다.

미디어 연구 분야에서 내용분석의 방법으로 측정돼 온 개념들에 대해 토픽 모델링을 적용하는 시도에 대한 우려와 비판은 타당한 측면이 있다. 이러한 지적에 대한 답변의 차원에서 본 연구는 댓글을 통해 형성되는 담론장에서 다뤄지는 ‘이슈’의 특징과 관련 댓글의 분포를 파악하는 방법으로서의 토픽 모델링을 제안하며 그 타당성을 검증했다. 댓글은 ‘이슈’에 대한 찬반 의견의 장을 제공하며 관련한 주제로 적절히 분류될 수 있음이 확인된 것이다. 그 결과, 디지털 플랫폼에서 거대한 크기로 쉽고 빠르게 나타났다가 사라지기를 반복하여 측정이 쉽지 않았던 댓글 담론을 토픽 모델링을 활용해 즉각적으로 요약하고 그 분포를 계량적으로 측정하는 방법을 제시했다. 이를 통해 온라인 여론의 양상을 보다 다각적이고 심도 있게 파악할 수 있는 방법 구축에 기여했다는 점에 본 연구의 의의가 있다 할 수 있겠다.

본 연구의 결과는 포털 뉴스에 달리는 댓글이 ‘이슈’에 따라 그 내용적 특성이 분명하게 구분되고 있음을 보여준다. 특히, 분석 대상이 우리 사회의 혐오 정서가 쉽게 투영되고 있는 소수자 집단을 다룬 뉴스의 댓글임을 고려할 때 혐오 대상과 사안의 맥락이 언어적 표현으로 발현되고 조직되는 장으로의 댓글을 제시한다(이신행, 2021). 더욱이 뉴스 댓글로 표출되고 구체화되는 소수자 집단에 대한 혐오는 개인이 ‘이슈’를 해석하고 태도를 결정하는데 영향을 미치기 때문에 혐오의 관점에서 댓글에 대한 분석이 중요하며 여기에는 대상은 물론 사안의 성격에 따라 댓글을 분류하는 방법이 중요하게 작용함을 의미한다(양혜승, 2022). 다시 말해, 혐오가 향하는 대상은 물론 그 정서가 표현되는 방식을 댓글 분류를 통해 포착할 수 있다는 것이다. 나아가 혐

오는 소수자에 대해 같은 태도와 정서를 공유하는 집단 의식의 발로로서 다양한 사회적 이슈가 그 대상과 맥락에 기초해 혐오로 발전되는데 댓글이 주요한 역할을 수행하고 있음을 시사한다.

이상의 댓글과 같은 온라인 단문의 주제 분류를 위한 토픽 모델링의 의의에도 불구하고 본 연구는 다음의 한계를 가진다. 첫째, 본 연구가 주제 분석의 대상으로 설정한 단문에 대해 분명한 기준을 제시하지 못했다는 점이다. 비록 본 연구는 수집한 댓글 길이의 분포를 바탕으로 그 분석 대상의 범위를 설정하였으나 온라인 단문에 대한 보다 명확한 개념화를 바탕으로 해당 범위를 정할 필요가 있다. 그리고 단문 범위의 변화가 모델 성능에 미치는 영향까지 파악할 때 단문의 주제 분석을 위한 토픽 모델의 범용적 활용도가 높아지리라 기대한다. 나아가 단문이 아닌 뉴스 기사와 같은 장문 데이터에 대해서도 BTM의 성능은 물론 초모수 조정과 바이그림의 효과를 LDA와 비교를 통해 검증하는 후속 연구는 토픽 모델링의 방법적 지평 확장에 유의미한 기여를 할 것으로 본다.

둘째, 본 연구는 LDA와 BTM이라는 깃스 표집 방식으로 모수를 추정하는 토픽 모델의 성능과 타당성을 검증했다. 그러나 변분 방법에 기반한 모수 추정을 사용하는 STM과의 비교는 추후 연구에서 다루볼 필요가 있다. 특히, STM은 텍스트 특성 이외의 저자나 시기 등의 문서 차원의 변수가 주제 분포나 내용에 미치는 효과를 추정하기 때문에 내용분석을 활용하는 사회과학 연구에서 빈번하게 사용되고 있다는 점(Roberts et al., 2019)을 고려한다면 댓글 주제 분석에 대한 성능이 검증될 필요가 있겠다.

셋째, 본 연구가 분석 대상으로 수집한 댓글은 혐오 이슈에 관련한 네이버 뉴스에서만 수집됐다는 점이다. 따라서 보다 다양한 이슈에 대한 댓글을 대상으로 한 주제 분류는 물론 네이버 뉴스의 댓글이 아닌 다음(Daum) 등의 다른 포털 사이트, 또는 유튜브(YouTube)나 인스타그램(Instagram)과 같은 소셜미디어 플랫폼에서 생성되는 댓글들에 대한 모델의 성능 검증이 필요하다. 따라서 후속 연구를 통해 토픽 모델링의 성능과 타당성이 다양한 이슈와 플랫폼에서 수집된 댓글을 이용하여 검증될 필요가 있음을 제안한다.

마지막으로, 토픽 모델링은 이슈별 댓글을 분류하는데 우수한 성능임을 확인했지만, 댓글이 이슈에 대해 찬성인지 반대인지(또는 우호적인지 비우호적인지)를 판별하는 방법으로서의 타당성 평가가 이루어지지 못했다는 점도 언급할 필요가 있겠다. 주제와 태도 및 정서를 구분하여 댓글을 분류해야 비로소 온라인 여론의 정확한 분석이 가능하다는 점을 감안할 때 후속 연구의 필요성이 더욱 두드러진다. 따라서 토픽 모델링을 이용하여 댓글을 이슈별로 분류하는 방법에 더해 찬반의 분포까지 추정하는 방법을 모색할 필요가 있으며 이는 디지털 텍스트를 활용한 컴퓨테이셔널 분석 방법의 확립에도 유의미한 기여를 하리라 본다.

References

- An, S., Lee, H., & Chung, S. (2022). Online public opinion on the conflict between young and old generations through topic modeling: An exploratory study on the media's role in generational solidarity. *Korean Journal of Journalism & Communication Studies*, 66(1), 89-126. doi:10.20879/kjics.2022.66.1.003. [안순태·이하나·정순돌 (2022). 토픽 모델링 분석을 통해 살펴본 세대 갈등에 대한 온라인 댓글 여론의 반응: 세대 연대를 위한 미디어 역할에 관한 탐색적 연구. <한국언론학보>, 66권 1호, 89-126.]
- An, S., Lim, Y., & Lee, H. (2020). A study of viewers' comments on online Mukbang videos: A big-data analysis of perceptions toward eating behavior. *Korean Journal of Journalism & Communication Studies*, 64(2), 269-310. doi:10.20879/kjics.2020.64.2.007. [안순태·임유진·이하나 (2020). 온라인 먹방(먹는 방송, Mukbang)의 댓글 연구: 식행동 관련 인식에 대한 빅데이터 분석. <한국언론학보>, 64권 2호, 269-310.]
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Chang, J., Boyd-Graber, J., Wang, C., Gerrish, S., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In Y. Bengio, D. Schuurman J. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems* (pp. 288-296). Cambridge, MA: The MIT Press.
- Chen, Y., Peng, Z., Kim, S.-H., & Choi, C. W. (2023). What we can do and cannot do with topic modeling: A systematic review. *Communication Methods and Measures*, 17(2), 111-130.
- Cho, M., & Lee, S. H. (2021). Analyzing news frames in the coverage of COVID-19: Data-driven approach to frame analysis. *Journal of Speech, Media and Communication Research*, 20(1), 79-121. doi:10.20879/kjics.2021.65.3.003. [조민정·이신행 (2021). 코로나19 관련 언론 보도 프레임 분석: 자료기반 자동화 프레임 추출 방법을 중심으로. <한국소통학보>, 20권 1호, 79-121.]
- Choi, D. S., Choi, S. E., & Choi, Y. J. (2008). How do comments affect forming public opinion in Internet portal news? *Journal of Political Communication*, 8, 311-358. doi:10.35731/kpca.2008..8.009. [최동성·최성은·최용준 (2008). 인터넷 포털뉴스 댓글의 여론형성 과정과 특성에 관한 연구. <정치커뮤니케이션연구>, 8권, 309-356.]
- DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics*,

41(6), 570-606.

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51-58.

Entman, R., Matthes, J., & Pellicano, L. (2009). Nature, sources and effects of news framing. In K. Wahl-Jorgensen & T. Hanitzsch (Eds.), *The handbook of journalism studies* (pp. 175-190). New York: Routledge.

Gitlin, T. (1980). *The whole world is watching*. Berkeley: University of California.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267-297.

Ham, S., Jung, S., & Kim, E.-Y. (2022). Analysis of media coverage on the issue of "Comfort Women" by daily newspapers in South Korea from 2003 to 2020: A big data study using topic modeling. *Korean Journal of Communication & Information*, 111, 181-215. doi:10.46407/kjci.2022.02.111.181. [함승경·정사강·김은영 (2022). 2003년~2020년 한국 일간지의 '위안부' 보도 분석: 빅데이터를 이용한 의제 분석을 중심으로. <한국언론정보학보>, 111권, 181-215.]

Hong, J., & Na, E. (2016). Online hate speech diffusion network analysis: Issue-specific diffusion patterns, types and intensity of verbal expression on online hatred. *Korean Journal of Journalism & Communication Studies*, 60(5), 145-175. doi:10.20879/kjcs.2016.60.5.006. [홍주현·나은경 (2016). 온라인 혐오표현의 확산 네트워크 분석: 이슈 속성별 확산 패턴 및 혐오표현의 유형과 강도. <한국언론학보>, 60권 5호, 145-175.]

Hwang, T. Y., Yoo, J. M., & Kim, H. Y. (2022). Analysis of changes in domestic public opinion before and after the Yemeni refugee crisis in Jeju: Focusing on sentiment analysis approach using deep neural network. *Korean Public Administration Review*, 56(2), 127-166. doi:10.18333/KPAR.56.2.127. [황태연·유정모·김화연 (2022). 제주 예멘난민 사태 전후 국내 여론변화 분석: 심층 인공신경망을 활용한 감성분석을 중심으로. <한국행정학보>, 56권 2호, 127-166.]

Jeong, I.-K., & Kim, Y.-S. (2006). 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*, 50(4), 302-327. [정일권·김영석 (2006). 온라인 미디어에서의 댓글이 여론에 미치는 영향에 관한 연구: 여론동향 지각과 제3자 효과를 중심으로. <한국언론학보>, 50권 4호, 302-327.]

Kang, J.-W., & Kim, S.-J. (2012). 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*, 56(2), 143-166. [강재원·김선자 (2012). 인터넷 뉴스기사에 달린 댓글의 효과 연구. <한국언론학보>, 56권 2호, 143-166.]
- Kim, E. J., & Kang, J. H. (2019). Grounded theory analysis of portal news and comment discourse on abortion. *Journal of Communication Research*, 56(2), 188-242. [김은진·강주현 (2019). 임신중절에 관한 포털뉴스와 댓글 담론의 근거이론적 분석. <언론정보연구>, 56권 2호, 188-242.]
- Kim, E.-M., & Sun, Y.-H. (2006). The effect of replies in internet news on the audience. *Korean Journal of Journalism & Communication Studies*, 50(4), 33-64. [김은미·선유화 (2006). 댓글에 대한 노출이 뉴스 수용에 미치는 효과. <한국언론학보>, 50권 4호, 33-64.]
- Kim, H., & Rhee, J. W. (2011). The impacts of frame fusion of news text and comments on news interpretation and attitude. *Korean Journal of Journalism & Communication Studies*, 55(2), 32-55. [김혜미·이준웅 (2011). 인터넷 뉴스와 댓글의 뉴스 프레임 융합 효과 연구. <한국언론학보>, 55권 2호, 32-55.]
- Koh, M. (2018). Who comments on or shares online news stories?: Application of statistical learning methodology. *Journal of Cybercommunication Academic Society*, 35(1), 5-51. doi:10.36494/JCAS.2018.03.35.1.5. [고문정 (2018). 누가 온라인 뉴스에 댓글을 작성하거나 뉴스를 공유하는가?: 통계 학습 방법의 적용. <사이버커뮤니케이션학보>, 35권 1호, 5-51.]
- Koltsova, O., & Koltcov, S. (2013). Mapping the public agenda with topic modeling: The case of the Russian livejournal. *Policy & Internet*, 5(2), 207-227.
- Lee, E.-J., & Jang, Y. J. (2009). 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*, 53(4), 50-71. [이은주·장윤재 (2009). 인터넷 뉴스 댓글이 여론 및 기사의 사회적 영향력에 대한 지각과 수용자의 의견에 미치는 효과. <한국언론학보>, 53권 4호, 50-71.]
- Lee, J. H., & Kil, W. (2019). News agenda classification and media diversity analysis using topic modeling: Based on news on the Presidential New Year Press Conference. *Korean Journal of Broadcasting and Telecommunication Studies*, 33(1), 161-196. doi:10.22876/kab.2019.33.1.005. [이종혁·길우영 (2019). 토픽모델링을 이용한 뉴스 의제 분류와 미디어 다양성 분석: 대통령 신년 기자회견 관련 뉴스 분석을 통해. <한국방송학보>, 33권 1호, 161-196.]
- Lee, J. S., Kim, J., Ryu, J., & Kang, J. (2010). The influence of the frame and genre of portal news articles on the type of online replies. *Korean Journal of Journalism & Communication Studies*, 54(2), 116-137.

[이재신·김지은·류재미·강재혁 (2010). 기사 프레임과 장르가 댓글 유형에 미치는 영향. <한국언론학보>, 54권 2호, 116-137.]

Lee, J. S., & Sung, M. J. (2007). The effects of reading replies on the perception of online news articles: Focusing on the PR perspective. *The Korean Journal of Advertising and Public Relations*, 9(4), 7-45.

[이재신·성민정 (2007). 온라인 댓글이 기사 평가에 미치는 영향: PR적 관점을 중심으로. <한국광고홍보학보>, 9권 4호, 7-45.]

Lee, M. (2022). *Kiwi: Korean Intelligent Word Identifier*. Retrieved from 3/1/23 from <https://github.com/bab2min/Kiwi>.

Lee, S. H. (2021). Biased artificial intelligence: Analyzing the types of hate speech classified by ‘Cleanbot’, NAVER AI for detecting malicious comments. *Journal of Cybercommunication Academic Society*, 38(4), 33-75. doi:10.36494/JCAS.2021.12.38.4.33.

[이신행 (2021). 편향적 인공지능: 네이버의 악플 탐지용 인공지능 ‘클린봇’이 판별한 혐오표현의 유형 분석. <사이버커뮤니케이션학보>, 38권 4호, 33-75.]

Lee, S. H., & Hyun, A. (2023). How does YouTube “Cyber-Wrecker” channels generate toxicity?: The effects of YouTuber’s anonymity, normative conformity, and echo-chamber of hate on malicious comments. *Journal of Cybercommunication Academic Society*, 40(2), 53-88. doi:10.36494/JCAS.2023.06.40.2.53.

[이신행·현아연 (2023). 유튜브 ‘사이버렉카’ 채널은 어떻게 악성댓글을 양산하는가? 유튜브 익명성, 규범 동조, 혐오의 반향실 효과를 중심으로. <사이버커뮤니케이션학보>, 40권 2호, 53-88.]

Lee, S.-M., Ryu, S.-E., & Ahn, S. (2020). Mass media and social media agenda analysis using text mining: Focused on ‘5-day rotation mask distribution system’. *Journal of the Korea Contents Association*, 20(6), 460-469. doi:10.5392/JKCA.2020.20.06.460.

[이새미·유승의·안순재 (2020). 텍스트 마이닝을 활용한 매스 미디어와 소셜 미디어 의제 분석: ‘마스크 5부제’를 중심으로. <한국콘텐츠학회논문지>, 20권 6호, 460-469.]

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., ... & Adam, S. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, 12(2-3), 93-118.

McCombs, M. E., & Shaw, D. L. (1972). The agenda-setting function of mass media. *The Public Opinion Quarterly*, 36(2), 176-187.

Miller, M. M., & Riechert, B. P. (2000). Interest group strategies and journalistic norms: News media

- framing of environmental issues. In U. Beck, B. Adam, S. Allan, & C. Carter (Eds.), *Environmental risks and the media* (pp. 45-54). New York: Routledge.
- Mun, G.-S. (2021). Evaluating LDA and BTM topic models for short texts. *Journal of the Korean Data Analysis Society*, 23(3), 1313-1326. doi:10.37727/jkdas.2021.23.3.1313. [문길성 (2021). 단문의 주제 분석을 위한 LDA와 BTM 토픽모형 평가. <Journal of The Korean Data Analysis Society>, 23권 3호, 1313-1326.]
- Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New media & society*, 6(2), 259-283.
- Rhee, J. W., & Kim, S.-H. (2018). News frames in the coverage of fine-dust disaster: Application of structural topic modeling. *Korean Journal of Journalism & Communication Studies*, 62(4), 125-158. doi:10.20879/kjcs.2018.62.4.004. [이준웅-김성희 (2018). 미세먼지 재해 보도의 프레임 분석: 구조적 주제모형(Structural Topic Modeling)의 적용. <한국언론학보>, 62권 4호, 125-158.]
- Roberts, M. E., Stewart, B. M., & Airoidi, E. M. (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, 111(515), 988-1003.
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, 91(2), 1-40.
- Walter, D., & Ophir, Y. (2019). News frame analysis: An inductive mixed-method computational approach. *Communication Methods and Measures*, 13(4), 248-266.
- Yan, X., Guo, J., Lan, Y., & Cheng, X. (2013). A biterm topic model for short texts. in *Proceedings of the 22nd International Conference on World Wide Web*, 1445-1456.
- Yang, H. S. (2008). 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*, 52(2), 254-281. [양혜승 (2008). 인터넷 뉴스 댓글의 견해와 품질이 독자들의 이슈에 대한 태도에 미치는 영향. <한국언론학보>, 52권 2호, 254-281.]
- Yang, H. S. (2018). Hate speech toward specific regions in Korea: Content analysis of comments posted to crime news stories in Naver. *Korean Journal of Journalism & Communication Studies*, 62(6), 7-36. [양혜승 (2018). 포털과 지역혐오: 네이버 범죄뉴스의 지역혐오댓글에 대한 내용분석. <한국언론학보>, 62권 6호, 7-36.]
- Yang, H. S. (2022). A text mining analysis on hate comments targeted at women, immigrants, and the elderly: In the context of NAVER crime news. *Korean Journal of Broadcasting and*

Telecommunication Studies, 36(3), 5-45. doi:10.22876/kab.2022.36.3.001. [양혜승 (2022). 여성, 이주민, 노인 혐오성 댓글에 대한 텍스트 마이닝 분석: 네이버 범죄뉴스를 맥락으로. <한국방송학보>, 36권 3호, 1-41.]

Ying, L., Montgomery, J. M., & Stewart, B. M. (2022). Topics, concepts, and measurement: A crowdsourced procedure for validating topics as measures. *Political Analysis*, 30(4), 570-589.

Yum, J.-Y., Kim, R., & Jeong, S.-H. (2020). A meta-analysis of the effects of user comments. *Journal of Communication Research*, 57(2), 5-49. doi:10.22174/jcr.2020.57.2.5. [엄정운·김류원·정세훈 (2020). 국내 댓글 효과 연구에 대한 메타 분석. <언론정보연구>, 57권 2호, 5-49.]

최초 투고일 2023년 06월 16일

게재 확정일 2023년 08월 04일

논문 수정일 2023년 08월 04일

부록

Appendix 1. Keywords of 9 Topics and Distributions of the Topics

Target	Issue	Model	Keywords	%
FOREIGNER	'관객소년단' 비판	LDA	차별, 인종, 인종 차별, 흑인, 한국, 비하, 아니, 나라, 생각, 취리, 오 취리, 사람, 학생, 샘 오, 분장	11.0
		BTM	차별, 인종, 인종 차별, 흑인, 비하, 아니, 한국, 취리, 생각, 오 취리, 사람, 학생, 샘 오, 의도, 분장	13.0
	영화 '청년경찰' 논란	LDA	조선족, 중국, 한국, 동포, 영화, 나라, 사람, 중국인, 우리, 아니, 사과, 범죄, 생각, 한국인, 중국 동포	11.2
		BTM	조선족, 중국, 한국, 동포, 영화, 나라, 사람, 중국인, 우리, 아니, 사과, 범죄, 생각, 한국인, 인데	10.1
	에멘인 난민 신청	LDA	난민, 나라, 국민, 우리, 사람, 이슬람, 한국, 아니, 인권, 반대, 우리 나라, 무슬림, 자국민, 정부, 국가	11.5
		BTM	난민, 나라, 국민, 우리, 이슬람, 사람, 우리 나라, 한국, 무슬림, 아니, 국가, 자국민, 인권, 반대, 정부	11.9
FEMALE	소설 '82년생 김지영'	LDA	여자, 남자, 여성, 페미, 아니, 차별, 영화, 남성, 한국, 사람, 사회, 생각, 힘들, 세대, 김지영	11.8
		BTM	여자, 남자, 여성, 페미, 한국, 영화, 아니, 힘들, 차별, 사회, 사람, 세대, 생각, 남성, 우리	12.1
	대림동 여성경찰 논란	LDA	경찰, 여경, 남자, 시민, 여자, 아니, 남경, 제압, 문제, 국민, 수 있, 체력, 사람, 필요, 능력	10.3
		BTM	경찰, 여경, 남자, 아니, 시민, 여자, 수 있, 제압, 문제, 남경, 체력, 사람, 국민, 능력, 필요	10.2
	이수영 남녀 쌍방 폭행	LDA	여자, 남자, 사람, 여성, 폭행, 사건, 페미, 사비, 아니, 피해자, 메 갈, 청원, 기사, 영상, 한남	11.0
		BTM	여자, 남자, 사람, 여성, 폭행, 페미, 사건, 사비, 피해자, 기사, 아니, 영상, 메 갈, 청원, 머리	9.2
LGBTQ	변희수 하사 여군 복무 청원	LDA	여군, 인권, 군대, 생각, 군인, 아니, 여자, 남자, 사람, 나라, 성전환, 수술, 입대, 생활, 전역	10.2
		BTM	여군, 인권, 여자, 군대, 생각, 아니, 군인, 남자, 성전환, 사람, 수술, 입대, 나라, 생활, 복무	9.2
	퀴어 축제 비대면 개최	LDA	동성애, 축제, 사람, 반대, 퀴어, 인권, 에이즈, 아니, 동성애자, 입니다, 나라, 우리, 국민, 퀴어 축제, 생각, 소수자	11.3
		BTM	동성애, 축제, 사람, 반대, 인권, 퀴어, 에이즈, 국민, 입니다, 우리, 나라, 아니, 동성애자, 소수자, 생각	12.4
	숙명여대 트랜스젠더 지원자 논란	LDA	여자, 남자, 여대, 여성, 사람, 생각, 아니, 사회, 차별, 성전환, 수 있, 소수자, 페미, 수술, 입학	11.6
		BTM	여자, 여성, 남자, 여대, 아니, 사람, 생각, 사회, 차별, 수 있, 페미, 소수자, 남성, 라고, 평등	12.1

Note. Single-letter words that are less useful for interpreting topic meaning were removed from the list of keywords.