

웹툰 댓글 분석을 통한 감정 예측 및 토픽 모델링: 드라마화 여부에 따른 독자 반응 비교

CUAI 7기 NLP 팀

권하연(소프트웨어공학), 김지호(미디어커뮤니케이션), 박지후(응용통계학), 오재환(소프트웨어공학)

[요약]

본 연구는 네이버 웹툰 댓글 데이터를 수집하고 감정 분석 및 토픽 모델링을 통해 드라마화된 웹툰과 그렇지 않은 웹툰의 댓글 양상을 비교 분석하였다. 감정 분석에 머신러닝과 LLM 모델을 활용해 성능을 비교했으며, 드라마화 여부에 따른 댓글의 주요 주제 차이를 토픽 모델링으로 분석하였다. 결과적으로, 드라마화된 웹툰에서 독자와 스토리 관련 토픽의 비중이 높았으며, 웹툰 산업의 발전 방향에 대한 시사점을 도출하였다.

1. 서론

최근 웹툰 산업은 빠르게 성장하며 주요한 대중문화 콘텐츠로 자리 잡고 있다. 특히, 웹툰은 원작으로서의 가치를 인정받아 드라마나 영화로 재탄생하는 경우가 많아지고 있으며, 이는 상업적으로도 큰 성공을 거두고 있다. 웹툰의 대중적 인기는 독자들의 즉각적인 반응을 반영한 댓글을 통해 더욱 강화되고 있다.¹ 웹툰 댓글은 독자들의 감정과 의견을 실시간으로 파악할 수 있는 중요한 데이터로, 이를 분석함으로써 웹툰의 인기도와 작품에 대한 반응을 예측할 수 있다. 본 연구에서는 네이버 웹툰의 댓글을 직접 크롤링하여 데이터를 수집한 후, 이를 라벨링하여 감정 분석 모델을 학습하고자 한다. 특히, 드라마화가 많이 이루어지는 '멜로' 장르를 중심으로 연구를 진행하며, 드라마화된 웹툰과 그렇지 않은 웹툰의 댓글 양상을 비교 분석할 것이다.

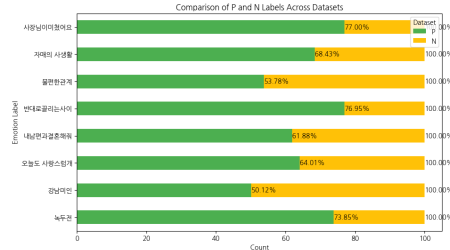
2. 본론

본 연구에서는 크롤링하고 직접 라벨링 댓글 데이터를 이용해 감정 분석을 위해 다양한 모델을 활용하여 모델을 학습하고, 성능을 비교 평가하였다. 이후, 드라마화된 웹툰과 그렇지 않은 웹툰의 댓글을 비교하기 위해 토픽 모델링을 진행하였다.

1) 데이터 수집

본 연구에서 사용된 데이터는 네이버 웹툰에서 댓글을 추출하여 구성하였다. 수집 대상은 "자매의 사생활", "사장님이 미쳤어요", "반대로 끌리는 사이", "불편한 관계", "내 ID는 강남미인", "내 남편과 결혼해줘", "녹두전", "오늘도 사랑스럽개"로, 8개 중 마지막 4개의 웹툰이 드라마화된 작품이다.

각 웹툰에서 추출된 댓글 수가 다양했기 때문에, 특정 회차에 치우치지 않도록 2~3,000개를 선정하였다. 선정된 댓글은 긍정 또는 부정으로 라벨링하였으며, 라벨링 가이드라인은 * 이 문서를 기준으로 일관되게 수행되었다. 최종적으로 수집된 댓글 데이터는 총 20,227개로, 이 중 긍정적인 댓글이 13,161개, 부정적인 댓글이 7,066개였다. 모든 웹툰에서 부정적인 댓글이 긍정적인 댓글보다 훨씬 적은 수를 보였다.



<그림 1.>웹툰별 긍/부정 댓글 비율

데이터의 불균형 문제를 해결하기 위해 부정 데이터에 대해 다운샘플링(downsampling)을 적용하여, 감정 분석 모델이 양측 데이터를 균형 있게 학습할 수 있도록 하였다. 이를 통해 긍정과 부정 감정 모두에 대해 효과적인 분류 성능을 기대할 수 있게 되었다.

2) 감정 라벨링 모델

본 연구에서는 네이버 웹툰 댓글의 감정 분석을 위해 총 4개의 다양한 모델을 사용하여 감정 분류를 수행하였다.

(1) 머신러닝 기반 앙상블 모델

이 모델은 전통적인 머신러닝 기법들을 결합한 앙상블 모델이다. SVM, RandomForest, GradientBoosting, LogisticRegression 모델을 결합하여 앙상블을 구성하였다.

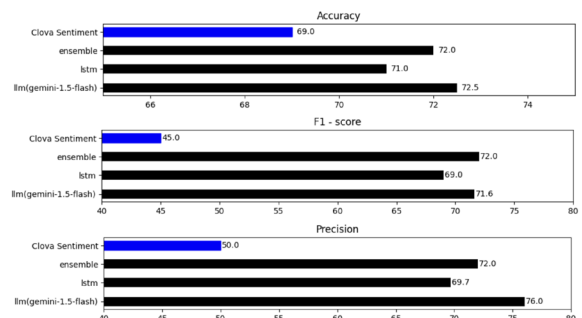
(2) LSTM 기반 분류 모델

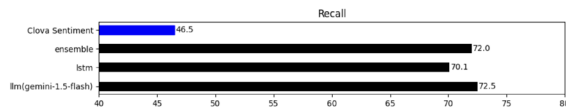
LSTM 기반 분류 모델은 순환 신경망 구조를 활용하여 시퀀스 데이터의 맥락을 이해하는 데 강점을 가진다.

(3) LLM 기반 분류 모델

대형 언어 모델(LLM)인 Gemini-Flash를 활용하여 댓글 감정 분류를 수행하였다. 최근 연구들에서는 LLM을 활용하여 분류를 진행하고자 하는 시도들이 나타나고 있다². 연구에서는 Gemini-Flash 모델에 감정 분류를 위한 가이드라인을 제공한 후, 댓글을 자동으로 분류하도록 하였다.

(1) 모델 성능 비교





<그림 2.> 감정 라벨링 모델 성능 지표

Clova모델은 정확도에서 다른 모델들보다 약 3% 낮은 성능을 보였다. 그러나 이 모델은 F1 Score, Recall 등의 지표에서 현저히 낮은 성능을 보였다. 이는 이 모델이 대부분의 댓글을 부정적으로 예측하는 경향이 강하기 때문으로 분석된다. 이로 인해, 부정적 감정에 대한 예측은 잘 이루어지지만, 긍정적 감정에 대한 예측이 부족하여 전체적인 성능이 저하된 것이다.

네 가지 모델 모두 웹툰 댓글의 감정 분석에서 전반적으로 높은 성능을 보이지는 않았다. 가장 높은 정확도를 기록한 모델은 72.5%로, 모델들 간의 성능 차이는 크지 않았다. 이러한 결과는 웹툰 댓글의 감정 분석이 단순한 문장 분석을 넘어서, 웹툰의 내용, 맥락, 독자들의 감정적인 반응 등 다양한 요소가 복합적으로 작용하기 때문에 예측이 어려운 작업임을 시사한다.

대부분의 모델이 비슷한 성능을 보였지만, 그중에서 LLM 모델이 가장 나은 성능을 기록하였다. 특히, 복잡한 문맥과 감정적인 뉘앙스를 파악하는 데 강점을 보인 이 모델은 다른 모델들에 비해 상대적으로 우수한 지표를 기록하며 웹툰 댓글 감정 분석에 있어 보다 나은 예측 결과를 보여주었다.

3) 토픽모델링

본 연구에서는 웹툰 사용자 댓글에 내재된 주제를 식별하고 분류하기 위해 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA) 기법을 적용하였다. LDA는 주어진 문서 집합에서 단어의 공동 출현 패턴을 분석하여 각 문서가 몇 가지 잠재적인 주제들로 구성되어 있는지를 파악하는 비지도 학습 방법이다. 모델은 문서 내 단어들이 다양한 주제에서 생성될 수 있다고 가정하고, 주제와 단어 간의 분포를 디리클레 분포(Dirichlet distribution)를 이용해 표현한다.

(1) 데이터 전처리

도배성 댓글 처리

도배성 댓글을 처리하기 위해 정규 표현식을 사용하여 반복되는 단어를 감지하고 이를 하나의 단어로 치환하였다. 이 과정은 분석의 정확성을 높이기 위해 수행되었다.

형태소 분석

댓글의 의미 있는 단어를 추출하기 위해 Okt 형태소 분석기를 사용하였다. 이 분석기를 통해 각 댓글에서 주요 형태소인 명사, 형용사, 동사만을 추출하여 사용하였다.

불용어 처리

불용어는 분석에 필요 없는 단어들이므로, 일반적인 불용어 사전과 연구자가 직접 정의한 불용어 사전을 결합하여 제거하였다.

토큰화

형태소 분석 및 불용어 처리를 마친 후, 댓글을 토큰화하여 단어 단위로 분리하였다. 이 과정에서는 의미 있는 단어만을 추출하고, 한글자 단어와 불용어를 제외하여 최종 토큰 리스트를 생성하였다.

위와 같은 전처리 후 댓글에서 가장 많이 사용된 단어

10개를 선정한 결과, 웹툰 ‘불편한 관계’의 부정 댓글을 제외하면 모든 웹툰의 긍정, 부정 댓글에서 작중 등장 인물의 이름이 가장 많이 언급되었다.

(2) 토폭 개수 선정

동일한 토픽 내에서 단어들이 전체 데이터셋에서 얼마나 자주 함께 등장하는지를 측정하는 Coherence Score를 활용하여 최적의 토픽 수를 도출하였다. Coherence Score는 값이 높을수록 해당 토픽이 일관성이 높다는 것을 의미하지만, 토픽의 개수가 지나치게 많으면 오버피팅의 위험이 발생할 수 있다. 따라서, 토픽 수는 10개에서 30개 사이에서 Coherence Score의 변동폭이 가장 큰 지점인 Elbow Point를 선택하였다.

	복두전	내남편과결혼후취위	내 ID는 강남 미인	오늘도 사랑스럽게
이정	16	18	26	15
부정	26	24	26	16
	사장님이 미쳤어요	불면증 관계	자매의 사생활	반대로 끌리는 사이
이정	30	28	28	25
부정	26	12	10	10

<표 1.> 웹툰별 최적의 토픽 개수

(3) 모델 결과

최적의 토픽 개수를 바탕으로 웹툰별 토픽을 추출하고 토픽을 대표하는 단어들을 추출하였다. 단어들은 각 토픽에 포함될 확률과 함께 출력된다. 토픽별 단어 확률 분포는 토픽을 해석하는 데 중요한 정보를 제공하는데, 높은 확률을 가진 단어들은 토픽의 핵심 단어로 간주되며, 이러한 단어들을 통해 토픽의 의미를 파악할 수 있다.

Title	Sentiment	Topic	word1	word2	word3	word4	word5
오늘도 사랑스럽게	1	0	0.087* '애나'	0.073* '선생님'	0.050* '캐시'	0.031* '강매자'	0.014* '귀뚜라미'
오늘도 사랑스럽게	1	0	0.017* '장주랑'	0.013* '미'	0.012* '자기'	0.011* '치름'	0.010* '여매'
강남비엔날레	1	3	0.043* '수아'	0.010* '전설'	0.006* '토틀'	0.007* '둘리'	0.007* '사람'

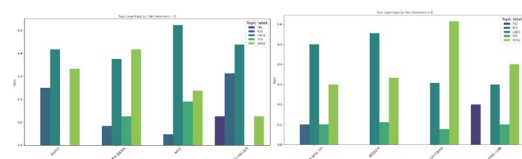
<표 2.> LDA 토픽 모델링 결과

토픽의 의미를 더 정확하게 파악하기 위해 각 문서에 대해 토픽의 확률을 계산하여 대표 댓글을 추출하였다. 최종적으로, 대표 단어와 대표 댓글을 활용하여 ‘스토리, 작가, 캐릭터, 독자, 기타’ 총 5가지의 카테고리별로 각 토픽을 라벨링하였다.

[illegible]

<표 3.> 토픽별 대표 댓글 추출 및 라벨링 결과

(4) 드라마화 여부에 따른 토픽 차이 분석



<그림 3.> 드라마화 유무별 부정 댓글에 대한 토픽 비율

토픽 비율과, 토픽별 댓글 내용을 통해 드라마화 여부에 따른 토픽 차이를 분석하였다. 두 경우 모두 '스토리'와 '캐릭터' 토픽이 주요 비중을 차지했지만, 드라마화된 웹툰에서는 '독자' 토픽의 비중이 비드라마화된 웹툰에 비해 훨씬 더 두드러졌다.

토픽 분류	댓글 내용
스토리	스토리에 대한 예측, 스토리 요약, 시대적 배경 설명, 사회 이슈 언급
작가	작가에 대한 애정, 작품으로 생긴 긍정적 변화, 웹툰에 대한 칭찬
캐릭터	캐릭터의 외형, 캐릭터의 행동, 캐릭터에 대한 평가
독자	캐릭터에 대한 감정이입 및 공감, 자신의 경험
기타	다음화를 기대하는 댓글

<표 4.> 드라마화된 웹툰의 긍정 댓글에 대한 토픽별 댓글 내용

긍정 댓글에서는 '스토리'와 '기타' 토픽에 대한 내용에서 차이를 확인할 수 있었다. 드라마화된 웹툰의 댓글에서는 단순한 스토리에 대한 예측과 요약뿐만 아니라 웹툰에서 다루는 시대적 배경과 사회 이슈에 대한 언급이 등장한다는 것을 확인할 수 있었다. 부정 댓글에서는 드라마화된 웹툰은 시대적 배경 설명과 사회 이슈를, 비드라마화된 웹툰은 플롯의 완성도와 관련된 부정적 피드백이 많았다.

웹툰 산업의 발전을 위해서는 스토리 측면에서는 전개 속도, 분위기, 연출, 개연성 등 플롯의 완성도를 향상시키는 것이 필요하다. 동시에, 시대적 배경을 적극 활용하여 폭넓은 세계관과 주제 의식을 전달하고, 사회 이슈를 반영하여 다양한 의견이 교환될 수 있는 사회적 담론의 장을 마련해야 한다. 독자 측면에서는 감정 이입을 유도하고 소통의 기회를 제공하는 웹툰을 개발하여 작품에 대한 참여도를 높여야 한다. 작가 측면에서는 현실 고증을 통해 작품의 진정성을 확보하고, 일관된 작화와 오타 수정 등 디테일을 철저히 관리하여 웹툰의 품질을 향상시키는 것이 중요하다. 이러한 접근은 웹툰의 전반적인 퀄리티를 높이고, 독자와의 신뢰를 구축하는 데 기여할 것이다.

3. 결론

본 연구는 웹툰 댓글 데이터를 기반으로 한 토픽 모델링과 감정 분석을 통해 웹툰 사용자들의 반응과 선호도를 심층적으로 분석했다. 연구 결과, 개발된 모델들이 긍정적 및 부정적 댓글에서 주요 토픽을 효과적으로 추출하고, 감정 분류에 유의미한 성과를 보였다. 이러한 분석은 웹툰 작가와 플랫폼 운영자들에게 독자들의 반응을 보다 정확하게 파악하고, 이를 바탕으로 웹툰의 품질을 향상시키는 데 중요한 인사이트를 제공할 수 있을 것이다. 또한, 이러한 방법론은 웹툰 산업의 마케팅 전략과 독자 참여 유도, OSMU(One Source Multi-Use) 전략의 최적화에도 기여할 수 있는 잠재력을 지니고 있다.

그러나 본 연구는 특정 플랫폼의 댓글에 한정된 데이

터에 기반하였다는 한계가 있으며, 감정 분석 모델이 댓글의 문맥이나 뉘앙스를 완벽하게 반영하지 못하는 제약이 있다. 향후 연구에서는 다양한 플랫폼과 장르에 대한 데이터를 포함해 분석 범위를 확장하고, 고도화된 자연어 처리 기법을 도입함으로써 분석의 정밀성을 더욱 강화할 필요가 있다.

참고 문헌

1) 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.

2) X. Sun, X. Li, J. Li, F. Wu, S. Guo, T. Zhang, and G. Wang. Text classification via large language models.

3) 웹툰댓글 공부정 감성사전 구축방안,
<https://github.com/hul980/Guide-to-Value-Added-Creation-through-K-Webtoon-Globalization-Strategy?tab=readme-ov-file>