

NLP를 활용한 콘텐츠 플랫폼 내에서의 댓글 시스템 개선

김예진¹, 손소현², 오수현³, 이현진⁴, 장예지⁵
¹이화여자대학교 휴먼기계바이오공학부 학부생, ²통계학과 학부생,
³수학과 학부생, ⁴수학교육과 학부생, ⁵경영학과 학부생
 alice1004u@ewha.ac.kr, thgus1598@ewhain.net, o.ss.u2056@gmail.com,
 summit723@ewhain.net, yeaje0831@gmail.com

Improving the comment system within the content platform using NLP

Ye-Jin Kim¹, So-Hyun Son², Su-Hyun Oh³, Hyeon-Jin Lee⁴, Ye-Ji Jang⁵
¹Dept. of Mechanical and Biomedical Engineering, ²Dept. of Statistics,
³Dept. of Mathematics, ⁴Dept. of Mathematics Education, ⁵Dept. of , Ewha
 School of Business, Ewha Womans University

요약

본 연구는 콘텐츠 플랫폼에서 제작자와 소비자 간의 건설적인 소통 촉진을 위해 악플을 순화하여 유의미한 피드백으로 전환하고, 댓글을 카테고리화하여 다양한 의견 교환을 가능하게 한다. 이를 통해 악플로 인한 부정적 영향을 최소화하고, 생산적인 콘텐츠 제작 및 소통 문화를 형성하고자 한다.

1. 서론

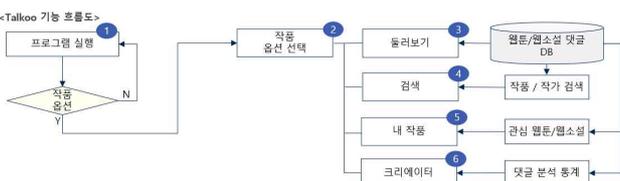
콘텐츠 플랫폼에서 악플은 제작자에게는 심리적 스트레스와 피드백 왜곡을, 소비자에게는 의견 교환의 왜곡과 정서적 부담을 유발하는 심각한 문제이다. 기존의 악플 처리 방식은 주로 삭제나 차단에 의존하여, 유의미한 의견마저 소실되는 부작용을 초래한다. 또한, 댓글을 공감순으로 정렬하는 기존 시스템은 필터 버블과 의견 편향 문제를 심화한다..

본 연구는 이러한 문제를 해결하기 위해, 악플 탐지 및 순화 기능과 댓글의 카테고리화 및 주요 토픽 정리를 통해 보다 체계적이고 다각적인 의견 교환을 가능하게 하는 시스템을 개발하였다. 이를 통해 플랫폼 내에서 악플의 부정적 영향을 최소화하고, 다각적인 의견 교환이 가능한 시스템을 제안한다.

2. 서비스

본 연구에서 개발된 Talkoo 시스템은 다음과 같은 주요 기능을 제공한다.

[그림 1] Talkoo 기능 흐름도



[그림 2] Talkoo 웹사이트 주요 기능 사진



2.1 전체 댓글 조회 및 댓글 키워드 검색

등록된 작품의 회차별 댓글을 조회하고 특정 키워드가 포함된 댓글만을 선별하여 확인할 수 있다.

2.2 악플 순화 및 댓글 필터링

User-generated 데이터를 기반으로 사전 훈련되어 데이터 노이즈에 강한 KcELECTRA 모델을 기반으로 콘텐츠 플랫폼에 특화된 악플 탐지 기능을 도입하여 기존 악플 탐지 알고리즘의 한계를 극복하고자 했다. 특히 기존 악플 감지 시스템을 우회한 악플 데이터를 추가로 라벨링하여 더욱 정교한 탐지가 가능하도록 개선했으며, ChatGPT 기반의 악플 순화 기능을 통해 악플을 유의미한 피드백으로 전환

한다. 이 과정에서, ChatGPT 프롬프트를 세밀하게 조정하여 사용자 생성 콘텐츠의 유머나 독창적인 표현이 유지되도록 설계되었다.

2.3 회차별 주요 토픽 제공 및 카테고리 필터링

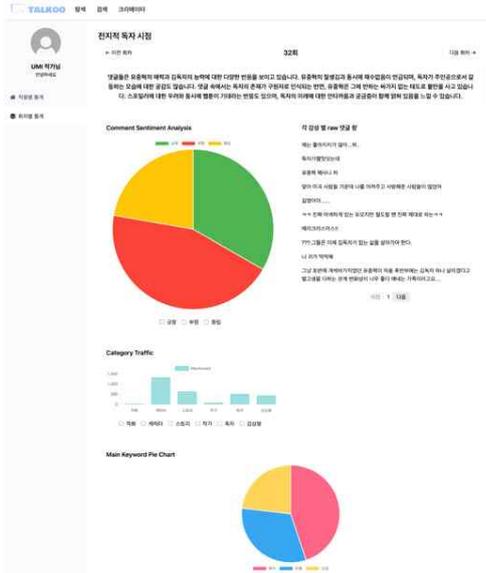
회차별로 댓글에서 추출된 주요 토픽을 중심으로 이용자들이 자유롭게 소통할 수 있는 채팅 공간을 제공한다.

또한, 댓글 필터링 시스템에서는 RoBERTa 모델을 사용하여 작화, 작가, 스토리, 캐릭터, 독자, 감상의 6가지 카테고리로 댓글을 자동으로 분류한다. RoBERTa는 대규모 사전 훈련을 통해 복잡한 언어 패턴과 비정형화된 데이터를 정확하게 이해하고 분류하는 데 강점을 지니고 있어, 각 카테고리의 특성을 잘 반영하여 정확한 필터링을 가능하게 한다.

이 방식은 사용자로 하여금 관심 있는 주제에 쉽게 접근하고 다양한 주제의 댓글을 체계적으로 제공할 수 있도록 한다.

2.4 작품 통계를 포함한 대시보드 기능

[그림 3] 작품 통계를 포함한 대시보드 기능



각 회차에 대한 댓글 데이터를 시각화하여, 주요 키워드와 감정 비율 등을 분석할 수 있는 대시보드를 제공한다. 또한, ChatGPT를 활용해 각 회차의 주요 내용을 요약하여 제공함으로써, 제작자가 회차별로 핵심 피드백과 트렌드를 신속하게 파악할 수 있도록 지원한다.

3. 사용자 기반 댓글 시스템 및 대시보드 설계

메인탭, 댓글 상세페이지, 크리에이터 대시보드의 세 가지 주요 구성 요소로 나뉘며, 웹 기반 인터페이스를 통해 접근성과 편의성을 높였다.

[그림 4] 메뉴 구성도



- 사용자 대상 댓글 시스템: 사용자는 댓글을 카테고리별 필터링하거나, 키워드 뮤트 기능을 활용해 관심 있는 댓글만 선별하여 볼 수 있다. 주요 토픽과 악플 순화 기능을 통해 보다 건설적인 의견 교환을 촉진한다. 순화된 댓글을 통해 더 긍정적이고 유의미한 정보를 확인할 수 있다.
- 크리에이터 대시보드: 작품별 및 회차별로 댓글 통계 및 트렌드를 분석하고, 댓글 요약 및 감성 분석 결과를 한눈에 파악할 수 있는 기능을 제공한다. 이를 통해 크리에이터가 독자의 반응을 효율적으로 수집하여 대응할 수 있다.

4. 결론 및 기대효과

본 연구에서는 웹툰 콘텐츠에 특화된 악플 탐지 및 댓글 분석 모델을 개발하여, 단순히 악플을 숨기거나 삭제하는 방식에서 벗어나 순화된 피드백을 제공하고 다양한 의견이 활발하게 공유될 수 있는 플랫폼을 구현하였다. 또한, 웹툰 댓글 데이터를 6가지의 카테고리로 분류하고 감성 분석 모델을 적용하여 기존의 단편적인 댓글 시스템을 개선하였다. 이를 통해 콘텐츠 제작자는 효율적으로 작품에 대한 피드백을 확인할 수 있으며, 콘텐츠 소비자는 건설적인 댓글 문화를 경험할 수 있다. 본 연구는 콘텐츠 제작자나 소비자가 기존의 악플로 인해 받는 정서적 피해를 줄이면서도 피드백의 가치를 보존하여, 콘텐츠 제작 환경에 실질적 변화를 가져올 것으로 기대된다.

※ 본 논문은 과학기술정보통신부 대학디지털교육역량강화 사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.