

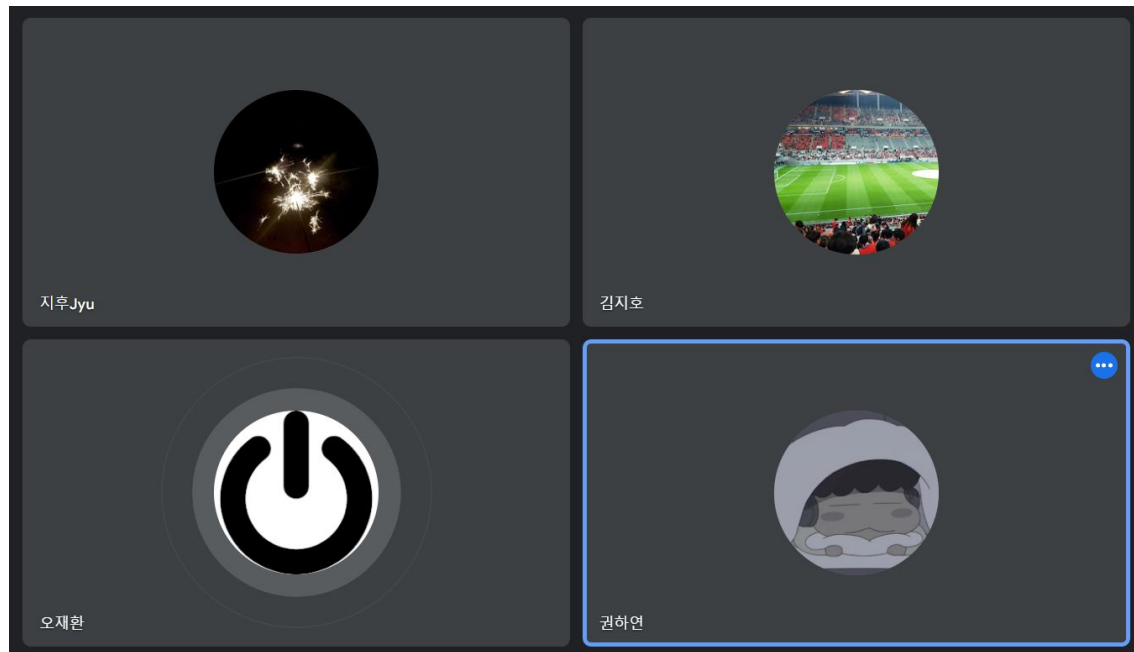


하계컨퍼런스 주제발표 **NLP** 팀

2024.07.30

발표자: 박지후

팀원 소개

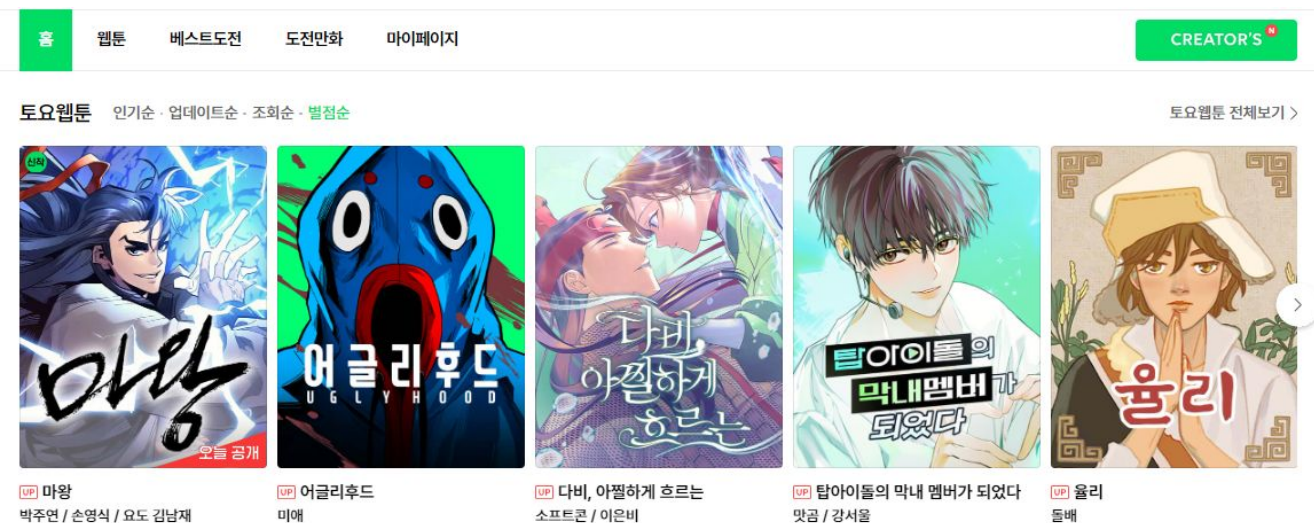


권하연
김지호
박지후
오재환

매주 토요일 12시

컨퍼런스 주제

네이버 웹툰 장르별 댓글 분석을 통한 독자 반응 및 콘텐츠 방향성 제안
: 최신 회차 댓글의 토픽모델링을 중심으로



분석 결과

불용어 처리

- 기존 불용어 사전 활용
- 자체 불용어 사전 제작
- **mecab** 분석기



토픽 모델링

- **LDA**
- **BERTopic**



토픽 라벨링

- 댓글 작성 대상으로

분석 결과

토픽 모델링 결과

> 8개의 장르, 총 24개의 웹툰 댓글로 진행

작품	Topic	Words	Representative Content	댓글 작성 대상
시월드가 내...	0	0.102**"하기"	하기 싫어? 하기 싫어? 하기 싫어? 하기 싫어?	등장인물 행동
시월드가 내...	1	0.018**"뭐" +	그리고보니 긴 시간동안 수많은 여성들	웹툰 스토리
시월드가 내...	2	0.019**"하면"	광공이 좋아요광공이 좋습니다광공이	등장인물 행동
시월드가 내...	3	0.043**"샤샤"	와 눈빛 개싸해서 순간 테오랑 내외함..	등장인물 외양
시월드가 내...	4	0.034**"샤샤"	먼저 선 넘었다는 건 남도 나에게 선을	등장인물 행동
시월드가 내...	5	0.053**"대화"	흐어.... 댜머로 기다리는거 보고 귀여워	등장인물 외양
시월드가 내...	6	0.051**"이제"	흐음~ 감당 가능..? 그니까 죽음이 정해	웹툰 스토리
시월드가 내...	7	0.022**"보고"	테오 ㄹ 양큼팍스아깁고영희 본인이	등장인물 행동
시월드가 내...	8	0.054**"부인"	당신 기다렸어 드르륵 탁... 당신 기다렸	등장인물 행동
시월드가 내...	9	0.143**"셀피"	와...그 샤샤와 닮은 여자가 셀피의 친모	웹툰 스토리
시월드가 내...	10	0.050**"미리"	대공이 눈빛 이쁘게 떠야지 샤샤가 안떠	등장인물 외양

Words (주요 단어)

: 0.043**"샤샤" + 0.016**"대공" +
0.011**"할" + 0.011**"출장" + 0.011**"
테오" + 0.011**"일이" + 0.011**"말" +
0.009**"소피아" + 0.008**"함" +
0.007**"안통"

Representative Content (대표 댓글)

: 와 눈빛 싸해서 순간 테오랑 내외함..
맞다 재 전쟁영웅이지.. 요즘 순둥한
눈빛만 봐서 적응못함 ㅠㅋㅋ

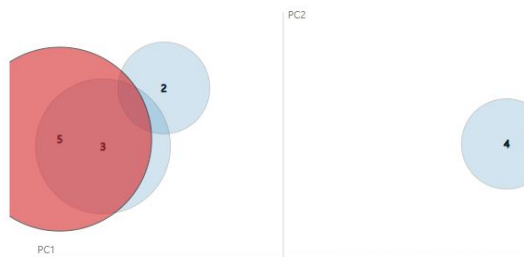
분석 결과

토픽 모델링 결과

> 8개의 장르, 총 24개의 웹툰 댓글로 진행

Selected Topic:

Intertopic Distance Map (via multidimensional scaling)

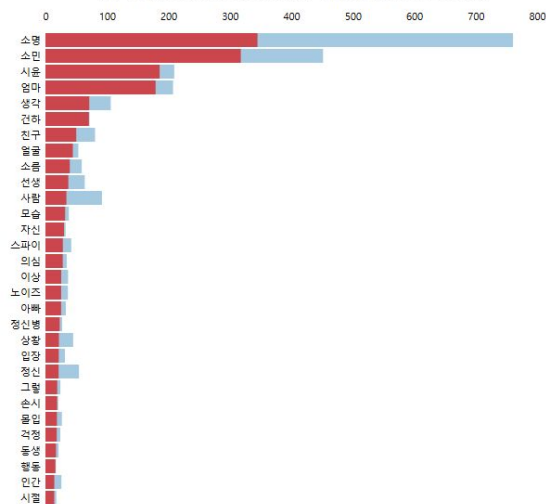


Marginal topic distribution



Slide to adjust relevance metric:⁽²⁾ $\lambda = 1$

Top-30 Most Relevant Terms for Topic 5 (43.9% of tokens)



Overall term frequency
Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t)) for topics t; see Chuang et al (2012)
2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

분석 결과

토픽 모델링 결과

> 8개의 장르, 총 24개의 웹툰 댓글로 진행

댓글 작성 대상

작품	Topic	Words	Representative Content	댓글 작성 대상
시월드가 내...	0	0.102**"하기"	"하기 싫어? 하기 싫어? 하기 싫어? 하기 싫어?"	등장인물 행동
시월드가 내...	1	0.018**"뭐" +	"그리고보니 긴 시간동안 수많은 여성들"	웹툰 스토리
시월드가 내...	2	0.019**"하면"	"광공이 좋아요광공이 좋습니다광공이"	등장인물 행동
시월드가 내...	3	0.043**"샤샤"	"와 눈빛 개싸해서 순간 테오랑 내외함.."	등장인물 외양
시월드가 내...	4	0.034**"샤샤"	"먼저 선 넘었다는 건 남도 나에게 선을"	등장인물 행동
시월드가 내...	5	0.053**"대화"	"흐어.... 덮머로 기다리는거 보고 귀여워"	등장인물 외양
시월드가 내...	6	0.051**"이제"	"흐음~ 감당 가능..? 그니까 죽음이 정해"	웹툰 스토리
시월드가 내...	7	0.022**"보고"	"테오 ㄸ ㄴ 양큼팍스아깁고영히 본인이"	등장인물 행동
시월드가 내...	8	0.054**"부인"	"당신 기다렸어 드르륵 탁... 당신 기다렸"	등장인물 행동
시월드가 내...	9	0.143**"셀피"	"와...그 샤샤와 닮은 여자가 셀피의 친모"	웹툰 스토리
시월드가 내...	10	0.050**"미리"	"대공이 눈빛 이쁘게 떠야지 샤샤가 안떠"	등장인물 외양

구분	
댓글 작성대상	웹툰 스토리
	등장인물 외양
	등장인물 행동
	웹툰 작가
	개인적인 경험
	타 댓글 및 타 이용자
	기타
계	

‘웹툰 스토리, 등장인물 외양, 등장인물 행동, 웹툰 작가, 개인적인 경험, 기타’ 로 **topic** 라벨링

분석 결과

토픽 모델링 결과

8개 장르의 댓글 토픽이 '웹툰 스토리', '등장인물 행동', '등장인물 외양', '웹툰 작가'로 동일하게 나옴. 따라서, 기존 주제였던 '장르별 콘텐츠 방향성 제안'을 다른 주제로 변경

기존

- 8개의 장르

- 토픽 모델링



변경 사항

- 1개의 장르 (로맨스)



- 감정 분석

컨퍼런스 주제

네이버 웹툰 로맨스 장르 댓글 분석을 통한 감정분석
: 기술적, 문화적 **contribution**을 중심으로

#로맨스 총 820

인기순 · 업데이트순 · 조회순 · **별점순**



치즈인더트랩
순끼
★ 9.98



고래별
나유희
★ 9.98



슈퍼 시크릿
이온
★ 9.98



세기말 풋사과 보습학원
순끼
★ 9.98



스피릿 핑거스
한경찰
★ 9.98

8개의 로맨스 작품
: 드라마화된 웹툰 4 + 비드라마화된 웹툰 4

컨퍼런스 주제

네이버 웹툰 로맨스 장르 댓글 분석을 통한 감정분석
: 기술적, 문화적 **contribution**을 중심으로

<기술 part>

1. '스토리', '등장인물', '작가', '기타'
총 4가지의 토픽으로 라벨링

스토리

등장인물

작가

기타

컨퍼런스 주제

네이버 웹툰 로맨스 장르 댓글 분석을 통한 감정분석
: 기술적, 문화적 **contribution**을 바탕으로

<기술 part>

2. 웹툰 댓글의 특성을 고려한
감정 사전 제작

3. 감정 분석 라벨링

4. 여러 종류의 모델로
학습 후 결과 비교

4.1.1. 캐릭터에 대한 원색적인 비난 (부정)

캐릭터에 대한 혐오, 직접적인 욕설, 편견이 포함되어 있는 등 캐릭터를 비난하는 댓글을 포함한다. 댓글 작성자가 작품에 몰입하고 있다는 것의 반증이라 할지라도 캐릭터에 대한 부정감정이 웹툰에 미치는 영향을 살피기 위해 부정으로 분류했다.

<예시 댓글>

- 파닭쿨진짜싫다 —아 답답해 뽀뽀려버려
- 시끄럽고 짜짜있어. 못생긴난아. 온새미로는 헤리꺼야.
- 육근육 극혐 지 때문에 벌어진 싸움인데 지가 뒤에서 쳐 울고있노
- 파티나 복귀 회식이라니ㅋㅋㅋㅋ 희린이 제발 강 꺼져줬으면
- 저 폐기물같은 돼지새끼가 더럽게 뭐하는거야; 더럽게 감기걸린 입으로 핥고 유자차를 타주면 어떡해;; 꼴보기 싫어

컨퍼런스 주제

네이버 웹툰 로맨스 장르 댓글 분석을 통한 감정분석
: 기술적, 문화적 **contribution**을 바탕으로

<문화 part>

등장인물 관련 댓글이 많은 로맨스 장르 웹툰을 바탕으로

등장인물에 대한 독자 반응의 흐름 추적 > 회차별, 등장인물별 감정분석 진행

유의미한 결과 확인 시, 드라마화된 웹툰 **vs** 비드라마화된 웹툰의
분석 비교를 통해 웹툰 산업 방향성 제시

컨퍼런스 일정

X	자매의_사생활_st.xlsx	👤
X	오늘도_사랑스럽개_st.xlsx	👤
X	사장님이_미쳤어요_st.xlsx	👤
X	불편한_관계_st.xlsx	👤
X	반대로_끌리는_사이_st.xlsx	👤
X	독두전_st.xlsx	👤
X	내_ID는_강남미인_st.xlsx	👤
X	내_남편과_결혼해줘_st.xlsx	👤

댓글 크롤링(완)

토픽 라벨링

감성 분석 실행

숏페이퍼 작성

07.30 중간 발표

08.08 컨퍼런스 회의

08.11 공모전 제출

감사합니다

THORAI