

Word2Vec 및 KoBART를 활용한 게임 추천 시스템

최수지, 김종민, 조민수
광운대학교 정보융합학부

sschoidev@gmail.com, jongmin4422@naver.com, mcho@kw.ac.kr



목차

1. 서론 3
2. 연구 방법 6
3. 연구 결과 14
4. 결론 18
5. 참고 문헌 19

01 서론 (1/3) Introduction

• 1.1 연구 배경 및 목적

- 연구 배경

- 제품 및 콘텐츠 객관적 평가 필수 방법 -> 텍스트 마이닝 기반의 사용자 리뷰 데이터 분석
- 특히, 사용자의 경험에 의존하는 대표적 경험재인 게임산업에서는 비정형 데이터를 바탕으로 다양한 관점에서의 정량적, 정성적 모델링이 필수적 (Kim and Cha, 2020)

- 연구 목적

- 온라인 게임 리뷰 데이터에 대한 새로운 관점 제시
(광고, 접속, 과금, 흥미(재미) 관점 중심으로 평가 기반 추천 시스템 제시)
- Word2Vec으로 핵심 키워드 사전 구축 / 관점 별 감성 사전 구축
- KoBART 기반 학습 모델을 통하여 핵심 관점 별 요약문 제공

01. 서론 (2/3)

- 1.2 기존 연구

- 한국어 리뷰 감성 분석 관련 연구 동향

1. 영화를 통한 리뷰 감성 분석 결과와 평점 비교 (Kim and Song, 2016)
2. 영화 도메인 / 핸드폰 도메인 / 의류 도메인에서 각 도메인의 특징 별 서술어가 가지는 감성 지수 측정 (Song and Lee, 2011)
3. 상품 리뷰의 형태소를 분석하여 대상과 대상의 특징을 세분화하고, 리뷰를 작성한 사용자별 및 요일별로 나누어 분석 (Yeon, Lee, Shim, and Lee, 2011)

- 한국어 요약 모델을 활용한 연구 동향

1. 트랜스포머를 활용하여 한국어 요약 모델 정량평가 및 정성평가 시행 (Kim, Kim, and Bang, 2021)
2. 기존 트랜스포머 기반 모델의 한계점 극복을 위한 KoBERT 모델과의 성능 차이 분석 (Song, Kim, Lee, Ahn, and Kim, 2021)

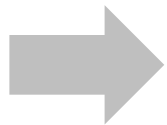
01 서론 (3/3)

- 1.2 기존 연구

- 한국어 텍스트 분석을 통한 추천 시스템 관련 연구 동향

1. 색인어와 TF-IDF 가중치를 활용한 스마트폰 앱 추천 시스템 (Jeon and Ahn, 2015)
2. 모델로 수치화한 리뷰를 반영한
새로운 평점 기반 추천 시스템 (Hyun, Ryu, and Lee, 2019)

감성 지수만으로는 사용자들의 평점 근거를 알기 어렵고,
리뷰 요약만으로는 리뷰의 평가를 수치적으로 알 수 없음



핵심 관점에 대한 감성 지수 추출과 리뷰 요약을 동시 진행하여
감성 지수와 그에 대한 명확한 근거 제시 요약 모델 구축,
최종적으로 이 둘을 활용한 추천 시스템 제안

02. 연구 방법 (1/8)

2.1 연구 절차

Methods

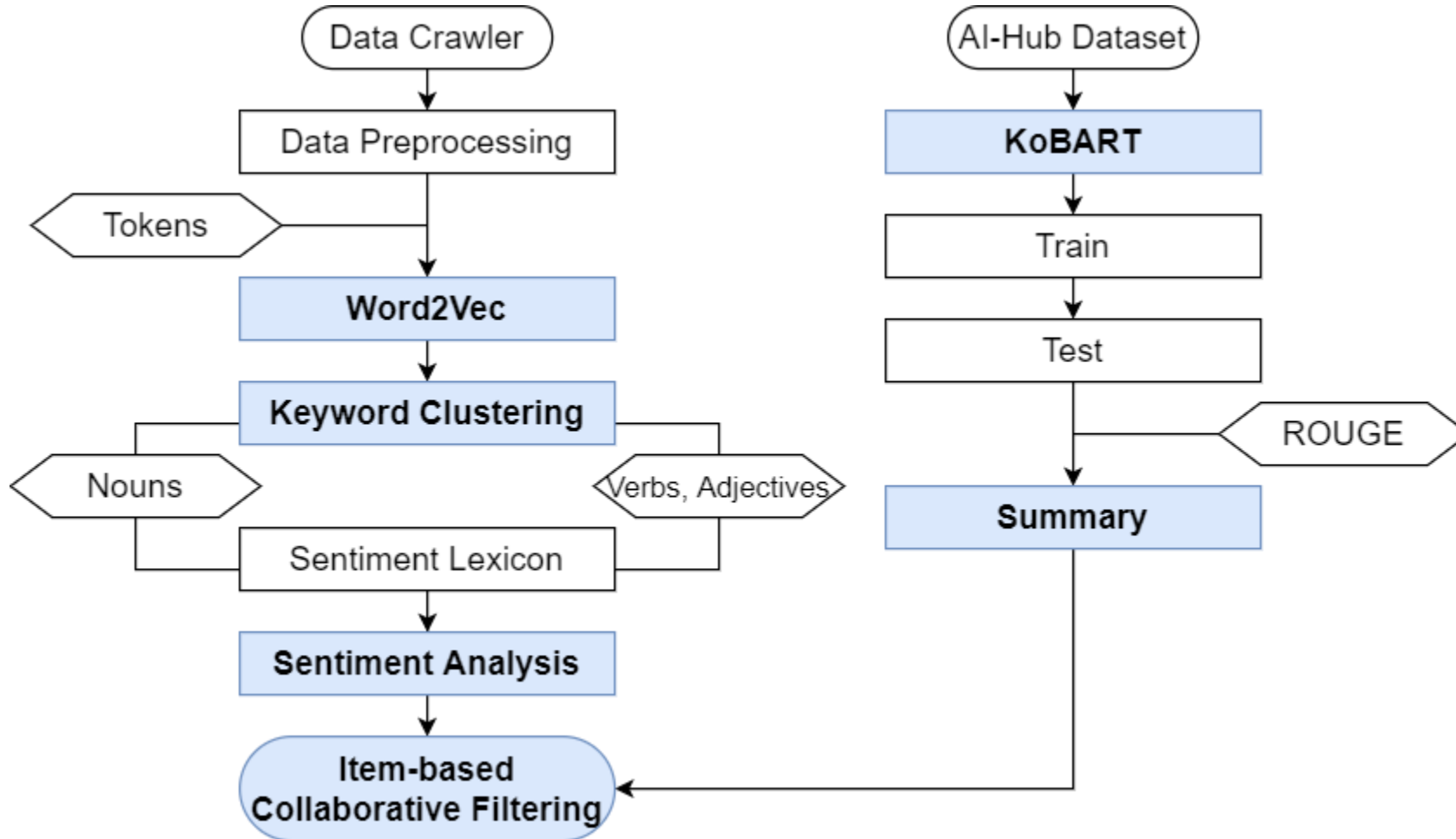


Figure 1. Research Process

02. 연구 방법 (2/8)

2.2 데이터 수집 및 전처리

Step 1

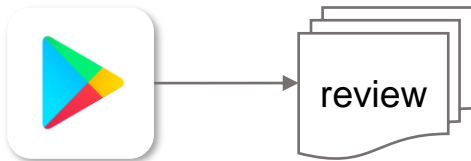


Figure 2. Google Play Review Crawling

- 대상 : 인기 무료 게임(상위 100개)
- 기간 : 2021.06 ~ 2021.07 (3차례)

➡ 총 10,527 개 리뷰 수집

Step 2

- 특수문자 제거
- 이모티콘 제거
- 맞춤법 검사 및 문장 띄어쓰기

Step 3

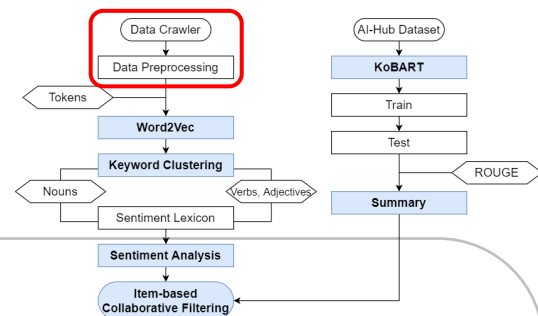
문장 유형별로 서로 다른 한글 형태소 분석기 적용

: 리뷰를 5개의 형태소 분석기로 분석
(Open-Korea_text, Mecab, ...)

예시 문장 : '희귀템 뽑으시면 조금 세집니다.'

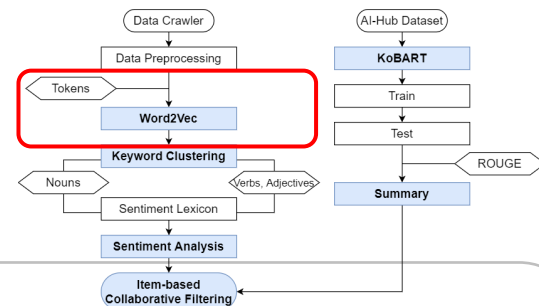
Table 1. Compare Korean-Tokenizer

형태소 분석기	토큰 추출 결과
OKT	'희귀', '템', '뽑으시면', '조금', '세', '집니다', '.'
Mecab	'희귀', '템', '뽑', '으시', '면', '조금', '세', '집니다', '.'



02. 연구 방법 (3/8)

2.3 Word2Vec을 통한 핵심 관점 별 사전 구축



Step 4

Open-Korea-Text 형태소 분석기로 명사 추출

Table 2. Top 100 Nouns

	명사	발생빈도
1	게임	7690
2	광고	4823
⋮	⋮	⋮
100	퍼즐	178

Table 3. Top 100 Nouns

	명사	발생빈도
1	게임	8568
2	광고	4843
⋮	⋮	⋮
100	구글	187

Table 2. 는 맞춤법 검사 진행 O

Table 3. 는 맞춤법 검사 진행 X

▶ 게임 도메인은 '트라하' 같이 특수한 단어 다수
맞춤법 검사 시 '트라하'를 '트라 하' 로 인식하는 문제로
두 가지 방법 동시 진행

도메인 '게임' 명사 제외

명사와 부사로 함께 쓰이는 단어 '진짜' 등 제외

Step 5

$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right) \quad (1)$$

$$\text{cosine similarity}(X, Y) = \frac{X \cdot Y}{\|X\|_2 \cdot \|Y\|_2} \quad (2)$$

TF-IDF로 각 리뷰 별 중요 키워드 추출(1) + 리뷰별로 Top 100 Nouns 존재 여부 추출

Word2Vec의 Skip-Gram방식으로 분석 진행(Kang, & Yang, 2019) 후 코사인 유사도 측정(2)

02. 연구 방법 (4/8)

2.4 핵심 관점 별 감성 사전 구축

네 가지 핵심 관점 선정

광고 = [광고]

접속 = [접속, 오류, 설치, 문제, 서버, 로그인, ... 저장]

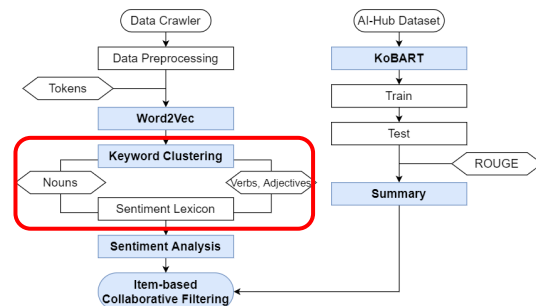
현질 = [현질, 과금, 뽑기, 유도, 확률, 코인, ... 아이템]

흥미 = [흥미, 재미, 난이도, 추천, 최고, 강추, ... 힐링]



핵심 관점과 유사한 단어 클러스터링
접속과 유사한 단어 **32개**
현질과 유사한 단어 **23개**
흥미와 유사한 단어 **12개**

단, 광고는 '광고' 단어로만 진행



Step 6

기존 감성 사전(ex. KNU 한국어 감성사전)은 특정 감성 단어의 경우 게임 도메인과 감성 지수가 상충함
예시 : KNU 한국어 감성사전('없다', 부정), 광고 관점('없다', 긍정)

▶ 따라서 게임 도메인에 맞는 감성 사전 구축
발생 빈도가 높은 순으로 Top 100 Verbs / Top 100 Adjectives 추출
단, '합니다'와 같이 극성을 측정하기 어려운 단어 제외

02. 연구 방법 (5/8)

2.5 감성 평가

Step 7

- 게임 리뷰 데이터셋의 특징으로,
- (1) 리뷰의 어순이 맞지 않거나
 - (2) 특정 게임 내에서만 사용되는 단어 사용
 - (3) 잘못된 문법으로 작성된 리뷰 존재



- (1) 맞춤법 처리 O : 단어(명사)가 변형됨
- (2) 맞춤법 처리 X : 형태소 분석기로 분석한 토큰 적합률 ↓

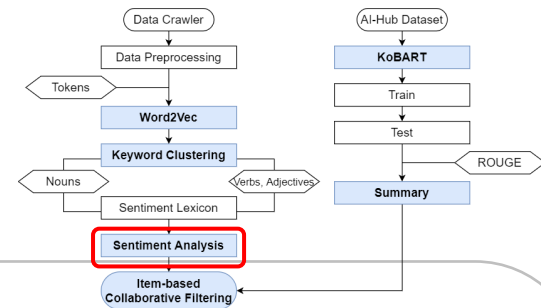


맞춤법 처리+토큰화 : 게임 리뷰는 길이가 길어 명확한 분석 어려움



리뷰 자체를 최대한 문장 단위로 잘게 자르고,
핵심 관점 및 유사 단어가 존재 여부+서술어 존재 여부 조합 방식 채택

각 관점 별 단어의 위치 확인하여 **긍정(1), 부정(-1)**으로 측정



Split Review

첫날에 결제를 하고 꾸준히 대기 열이 많아 접속 자체를 할 수가 없습니다.
 돈은 썼는데 게임을 하다가 질려서 접은 것도 아니고 접속이 돼야 무슨 게임을 하죠

⋮

+

pos(1)	neg(-1)
word1	word1
word1	word1
⋮	⋮

Figure 3. Sentiment Analysis Method

02. 연구 방법 (6/8)

2.6 모델 학습



문서요약 텍스트
뉴스기사 요약
Train : 199,735개
Test : 22,193개

- 실제 게임 리뷰는 요약문 존재 X
- ▶ 요약문이 존재하는 데이터셋 활용

Train

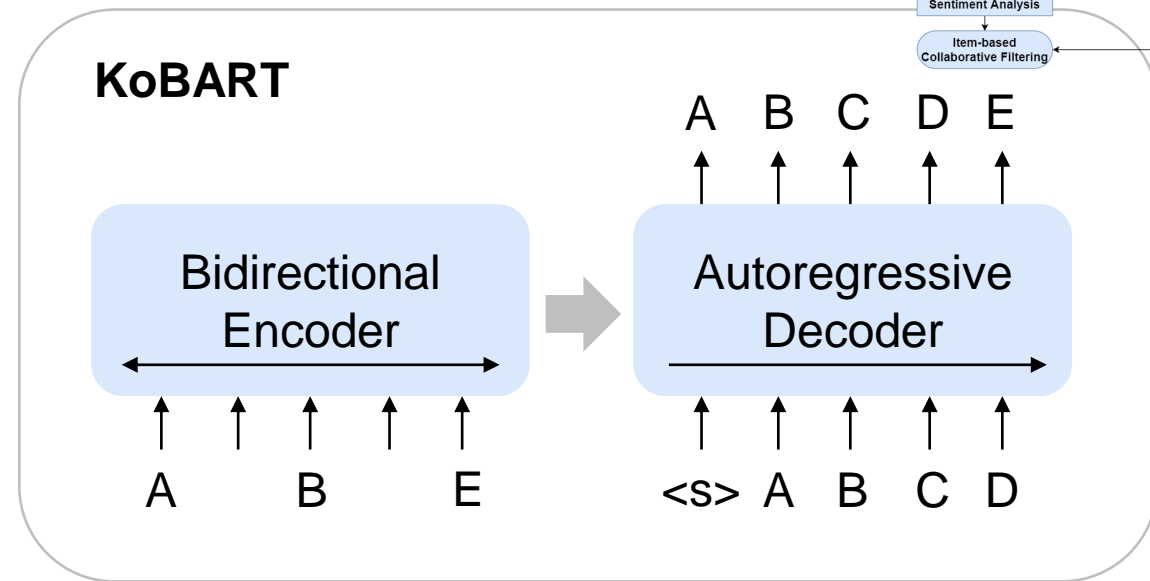
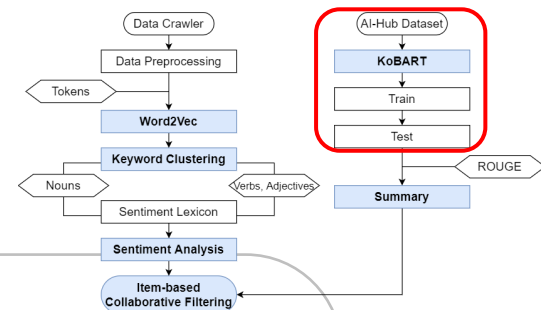


Figure 4. KoBART Structure

Number of parameters : 124Millions
Transformer(Encoder-Decoder)
Number of layers : 6

학습환경 : Google Colab Pro+ (GPU : Tesla V100)
학습시간 : 13:41:28 (시:분:초) 소요



02. 연구 방법 (7/8)

2.7 모델 평가산식

- ROUGE (Recall)

$$\frac{\text{number of overlapping } n - \text{grams}}{n - \text{grams in reference summary}} = \frac{\sum_{s \in \{\text{ReferenceSummaries}\}} \sum_{gram_n \in s} \text{Count}_{\text{match}}(gram_n)}{\sum_{s \in \{\text{ReferenceSummaries}\}} \sum_{gram_n \in s} \text{Count}(gram_n)} \quad (3)$$

- ROUGE (Precision)

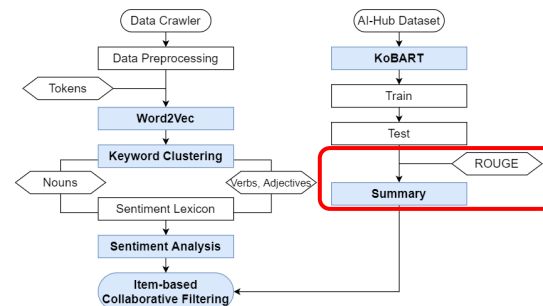
$$\frac{\text{number of overlapping } n - \text{grams}}{n - \text{grams in model summary}} = \frac{\sum_{s \in \{\text{ModelSummaries}\}} \sum_{gram_n \in s} \text{Count}_{\text{match}}(gram_n)}{\sum_{s \in \{\text{ModelSummaries}\}} \sum_{gram_n \in s} \text{Count}(gram_n)} \quad (4)$$

- ROUGE F-1

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (5)$$

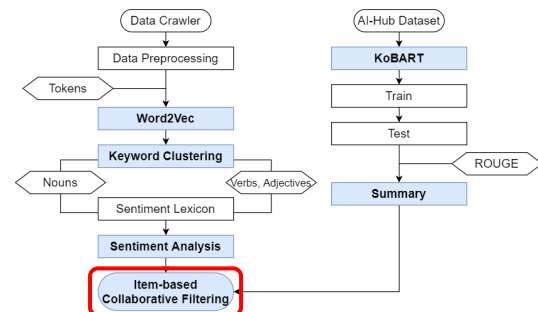
ROUGE (Recall-Oriented Understudy for Gisting Evaluation)

ROUGE 중 세 가지 **ROUGE-1, ROUGE-2, ROUGE-L** 사용
 뒤의 숫자는 각각 N-gram에서 N을 의미, ROUGE-L의 L은 Longest Common Subsequence를 의미
 (Lin, 2004)



02. 연구 방법 (8/8)

2.8 추천 시스템



“게임 추천 시스템을 이용할 사용자들은 ‘광고’가 적고, ‘접속’이 원활하며, ‘현질’을 적게 하고, ‘재미’가 있는 게임을 찾을 것이다.”를 바탕으로 추천 시스템 제안

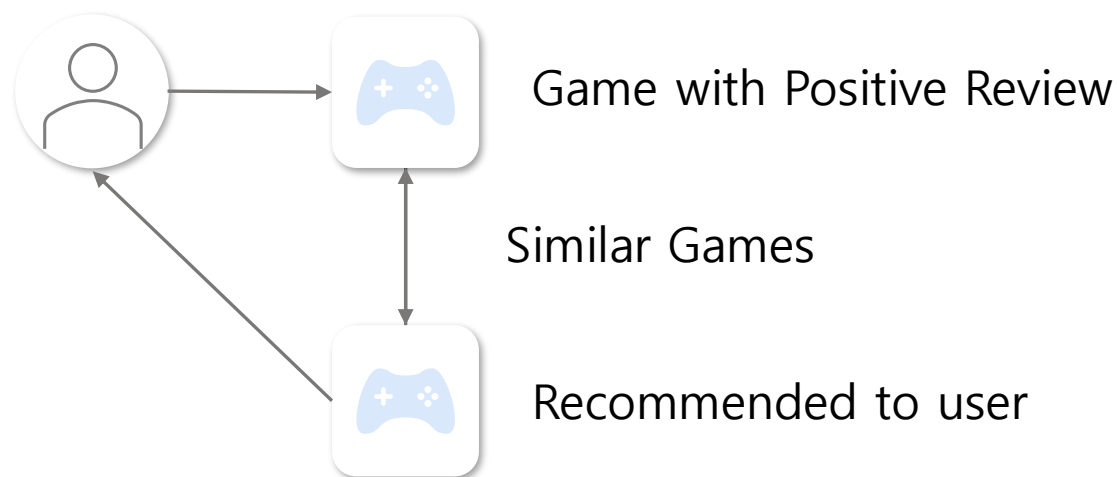


Figure 5. Item-based Collaborative Filtering

랭킹 시스템으로 전체 리뷰 반영 시 관점 별 가장 긍정적으로 평가된 게임 제시 및

협업 필터링으로 사용자가 자신이 긍정적으로 평가한 게임
감성 점수의 *Cosine Similarity*(2)가 높은 게임 추천

03. 연구 결과 (1/4) Results

3.1 모델 정량평가

Table 4. KoBART-Summarization

		ROUGE-1	ROUGE-2	ROUGE-L
문서요약 20,000개 Epoch 5	Recall	0.218341	0.123416	0.20037
	Precision	0.279492	0.161403	0.260129
	F_1	0.232843	0.13242	0.215017
문서요약 220,000개 Epoch 5	Recall	0.331049	0.195779	0.29742
	Precision	0.290762	0.171067	0.262967
	F_1	0.297572	0.175122	0.268205
문서요약 220,000개 Epoch 2	Recall	0.302696	0.179222	0.265614
	Precision	0.312425	0.187828	0.286972
	F_1	0.294466	0.174996	0.269234

Epoch 횟수가 증가할수록 Recall 값이 상승하며 Precision 값이 미세하게 하락함

또한, 데이터의 개수가 증가할수록 Recall 및 Precision 모두 확실한 상승세를 보임

03. 연구 결과 (2/4)

3.2 모델 정성평가

문서 원문

제일 좋아하는 게임인데 그리고 저가 여기에 현질을 많이 했는데... 저희 엄마한테 현질 한 번만 해달라고 해서 엄마가 해주셨어요. 그리고 현질을 잘 안 하는데 그 계정에 10만 원 넘게 현질을 했습니다 최근 들어 지인에게 로퍼를 선물하였지만 로퍼를 받지 않아 환불하라고 해서 환불을 했습니다만.. 네, 어제 10년 정지를 당했습니다. 한국 맵이 출시한다니 이잉!! 마동석 스킨 꼭 얻고 말 겁니다! 한일전도 기대 중이고요! 캐릭터 몸에 선을 탄 파란색 선들이 많이 있어요 제발 보라 등급 스킨 예쁘게 만드세요! 다른 게임과 달리 현질 뺄 거의 안 받는 게임인 거 같아요 저는 이번 스킨 보고 로열 패스 구매했는데 안되니까 속상하네요 이유는 제 주변에 배그 하는 애들이 부모님이 허락을 안 하셔서 현질을 못해서 속상한 친구가 있어요 다른 애들은 현질을 하는데 그 애만 현질 못하고 제 말은 여기까지입니다 현질도 해서 10만 원 정도 썼는데.. 2. 윤과 금과 부과금의 게임 실력차가 나지 않습니다.

KoBART model 요약 결과

제일 좋아하는 좋아하는 게임인데 저가 여기에 현질을 많이 했는데 저번 스킨 보고 로열 패스 구매했는데 안되니까 속상하네요 이유는 제 주변에 배그 하는 애들이 부모님이 허락을 안 하셔서 현질을 못해서 속상한 친구가 있어요 다른 애들은 현질을 하는데 다른 애들은 현질을 하는데 다른 애만 현질을 하는데 제 말은 여기까지입니다 현질도 해서 10만 원 정도 썼는데 제 말은 여기까지입니다 현질도 해서 10만 원 정도 썼는데 제게는 윤과 금과 부과금의 게임 실력차가 나지 않습니다.

Figure 6. ‘과금’ Keyword Abstractive Summary

문서 원문

아~주 재미있습니다! 고의적으로 게임하는 사람들 때문에 스트레스 너무 많이 받아요ㅠ... 다른 게임은 다 지워도 이것만큼은 지운 적이 없는데.. ㅠ.. 일단 정말 재미있습니다! 그래도 재밌으니까 한번 해보는 걸 추천해요! 신규 맵도 많이 나오고 신규 패스랑 신규 크롤러까지 이렇게 재미있는 게임을 만들어줘서 감사하고 앞으로도 재밌게 합시다 파이팅! 아 그리고 점수도 5점 드립니다 고맙습니다 순수 재미를 아주 잘 공략한 게임입니다. 실력 게임이라 차이가 별로 안 나요. 장점 1. 사람이랑 해서 재밌음 3. 재미있음 재미있어요!! 그런데 저희 팀 중 한 명은 5520 그리고 다른 한 명은 12118 네 이분은 저랑 차이가 별로 나진 않았습니다. 정말 재미있는데 레드 와이파이 좀 고치면 좋겠어요. 장점들 재미있다, 프롤 스타즈는 정말 재미있는 게임입니다 3재미 있오, 재미있지만 몇 가지 고쳐야 할 점이 있습니다 일단 게임 자체는 너무 재미있어요 하지만 게임은 재미있어서 별 5개 드립니다^^ 이거 빠면 다 재미있고 심심할 때 해도 좋고 그러니깐 강추예요 저 원래 게임을 별로 안 하는데 이 게임은 진짜 계속하게 돼요. 게임은 재미있습니다 재미있어요! 프롤 하는 거 정말 추천합니다!!! 그래도 재미있으니 싸 별점 4점 드리도록 하겠습니다. 너무 재미있습니다. 앞으로 더 게임 재미있게 바꿔 주십시오!! 너무 재미있는데 레드 와이파이 때문에 움직이지도 못하고 공격도 못해요 게꿀잼 추천!!!!

KoBART model 요약 결과

고의적으로 고의적으로 게임하는 사람들 때문에 스트레스 너무 많이 받아요ㅠ... 다른 게임은 다 지워도 이것만큼은 지운 적이 없는데 진짜 진짜 재미있게 게임하는 게꿀잼 추천!!!!

Figure 7. ‘재미’ Keyword Abstractive Summary

F_1 기준 세 점수 모두 높은 문서요약 텍스트 데이터 220, 000개로 학습한 모델로 정성평가 진행

요약 결과 N-gram에서 중복 제거를 통한 최종 문장을 추출 시
각 관점에 해당하는 단어 관련 문장 생성됨

03. 연구 결과 (3/4)

3.3 감성 분석 결과

$$keyword_n \text{ Score} = \frac{keyword_n \text{ positive Score}}{\text{count of match keyword}_n \text{ review}} \quad (1 \leq n \leq 4, \text{ for } n \text{ is integer})$$



게임별로 4가지 관점에 대한
positive, negative 스코어를
산출함

Table 5. Game-Sentiment Score table

단위: %

Game	Advertisement positive	Advertisement negative	Connect positive	Connect negative	Buying positive	Buying negative	Fun positive	Fun negative
Game1	38.5	84.6	13.8	58.6	30.4	45.7	68.4	36.8
Game2	19.1	44.7	35	60	37.5	50	65.4	34.6
Game3	29.3	51.1	27.8	53.9	48.4	58.1	49.6	27
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

게임별로 감성 점수 벡터를 얻을 수 있음
관점 별로 나누는 방식으로 **평점과 대비해 더욱 세밀함**

03. 연구 결과 (4/4)

3.4 추천 시스템 활용 결과

Table 6. Game-Sentiment Score Cosine Similarity Matrix

	Game1	Game2	Game3	Game4	Game5	...	Game100
Game1	1	0.930197	0.902011	0.575678	0.932348	...	0.917997
Game2	0.930197	1	0.941334	0.642165	0.967358	...	0.941387
Game3	0.902011	0.941334	1	0.683925	0.975525	...	0.919605
Game4	0.575678	0.642165	0.683925	1	0.719087	...	0.537892
Game5	0.932348	0.967358	0.975525	0.719087	1	...	0.933525
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Game100	0.917997	0.941387	0.919605	0.537892	0.933525	...	1

Table 7. Game 'ROBLOX' Cosine Similarity

App Name	Cosine-Similarity
캔디크러쉬사가	0.991906
행복의 저택	0.98469
지오메트리 대시 라이트	0.979084

사용자가 긍정적으로 리뷰를 작성한 게임과
0.95 이상의 코사인 유사도를 보이는 감성 점수 벡터를 가지는 게임 추천 시스템
추천 개수는 최대 3개

Table 7은 게임 'ROBLOX'의 감성지수와
 코사인 유사도가 가장 높은 3개의 게임 추천 결과

04. 결론 (1/1) Conclusion

- 연구 의의

- ▶ 게임 도메인에 적합하고 구체적인 감성 분석 방법을 사용하여 네 가지 핵심 관점(광고, 접속, 과금, 흥미) 별 요약문과 감성 지수를 기반한 추천 시스템을 구축한 점에서 의미가 있음
 - ▶ 이용자들은 관점 별 요약문과 감성 지수를 통하여 원하는 게임을 선택할 수 있음
 - ▶ 게임을 만드는 이들은 요약문 및 감성 지수와 흥행을 비교하여 통찰력을 얻을 수 있음

- 연구 한계 및 추후 연구

- ▶ '접속' 관점의 데이터의 경우 짧은 주기로 데이터 갱신이 필요함
- ▶ KoBART-base의 추가 학습을 게임 도메인 데이터셋을 사용하지 않음
학습할 때 요약문이 존재하는 게임 도메인 데이터셋을 사용할 필요가 있음
- ▶ 쓰인 감성분석 방식은 한 문장에 여러 관점에 대한 다른 극성을 띤 서술어가 존재할 시 점수를 낼 때 오류를 범할 가능성이 높음
게임에 특화된 형태소 분석기를 구축하여 기존 문제점을 해결하거나, 게임에 특화된 감성 사전의 단어 수를 늘리고 음소 단위 또는 음운 단위로 분석하여 개선할 수 있는지에 대한 사례연구가 수행되어야 함

05. 참고문헌 (1/2)

- [1] Hyun, J. Y., Ryu, S. Y., & Lee, S. T. (2019). How to Improve the Accuracy of Recommendation Systems : Combining Ratings and Review Texts Sentiment Scores. *Journal of Intelligence and Information Systems*, 25(1), 219–239.
- [2] Jeon, B. K., & Ahn, H. C. (2015). A Collaborative Filtering System Combined with Users Review Mining : Application to the Recommendation of SmartphoneApps. *Journal of Intelligence and Information Systems*, 21(2), 1–18.
- [3] Kang, H. S., & Yang, J. H. (2019). Optimization of Word2vec Models for Korean Word Embeddings. *Journal of Digital Contents Society*, 20(4), 825–833.
- [4] Kim, I. K., & Cha, S. S. (2020). The Dynamics of Online Word-of-Mouth and Marketing Performance : Exploring Mobile Game Application Reviews. *Journal of The Korea Contents Association*, 20(12), 36–48.
- [5] Kim, S. B., Kim, Y. J., & Bang, J. S. (2021). Analyzing the Impact of Sequential Context Learning on the Transformer Based Korean Text Summarization Model. *The Korean Institute of Information Scientists and Engineers*, 48(10), 1097–1104.
- [6] Kim, Y. Y., & Song, M. (2016). A Study on Analyzing Sentiments on Movie Reviews by Multi-Level Sentiment Classifier. *Journal of Intelligence and Information Systems*, 22(3), 71–89.
- [7] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., ... & Zettlemoyer, L. (2019). Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*.
- [8] Lin, C. Y. (2004, July). Rouge: A package for automatic evaluation of summaries. In *Text summarization branches out* (pp. 74-81).

05. 참고문헌 (2/2)

- [9] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- [10] Song, E. S., Kim, M. S., Lee, Y. R., Ahn, H. C., & Kim, N. G. (2021). Transformer-Based Text Summarization Using Pre-Trained Language Model. *Proceedings of the Korean Society of Computer Information Conference*, 29(2), 395–398.
- [11] Song, S. J., & Lee, S. W. (2011). Automatic Construction of Positive/Negative Feature-Predicate Dictionary for Polarity Classification of Product Reviews. *Journal of KISS : Software and Applications*, 38(3), 157–168.
- [12] Woo, K. J., & Jung, S. H. (2019). Comparison of Korean Morphology Analyzers According to the Types of Sentence. *The Korean Institute of Information Scientists and Engineers*, 1388–1390.
- [13] Yeon, J. H., Lee, D. J., Shim, J. H., & Lee, S. G. (2011). Product Review Data and Sentiment Analytical Processing Modeling. *The Journal of Society for e-Business Studies*, 16(4), 125–137.
- [14] Open-Korean-text : <https://github.com/open-korean-text/open-korean-text>
- [15] Korean Sentence Splitter : <https://github.com/likejazz/korean-sentence-splitter>
- [16] KNU SentiLex : <https://github.com/park1200656/KnuSentiLex>
- [17] KoBART-summarization : <https://github.com/seujung/KoBART-summarization>
- [18] AI-hub : <https://aihub.or.kr/aidata/8054>