

<http://dx.doi.org/10.17703/JCCT.2022.8.4.457>

JCCT 2022-7-57

텍스트 마이닝을 활용한 웹툰 애플리케이션 사용자 리뷰 분석

Analysis of User Reviews for Webtoon Applications Using Text Mining

신효림*, 최준호**

Hyorim Shin*, Junho Choi**

요약 웹툰 산업이 급속도로 성장하며, 이러한 성장세와 함께 새로운 웹툰 애플리케이션 모델이 제시되었다. 웹툰 애플리케이션 1.0과 2.0을 지나 3.0의 시대가 시작된 것이다. 이러한 변화에도 불구하고 아직까지 웹툰 애플리케이션을 대상으로 한 사용자 리뷰 분석 연구는 부족한 실정이다. 이에 이 연구는 웹툰 애플리케이션 3.0 모델을 제시한 ‘카카오웹툰(다음웹툰)’을 대상으로 사용자 리뷰를 분석하고자 한다. 분석을 위해 애플리케이션 리뷰 20,382개를 수집한 후 전처리 과정을 버전 별로 TF-IDF, 네트워크 분석, 토픽 모델링, 감성 분석을 실시하였다. 이를 통해 웹툰 애플리케이션 변화에 따른 사용자 경험을 탐구하고 리뷰를 통한 사용성 평가를 진행하였다.

주요어 : 웹툰, 사용자 경험, 앱 리뷰 분석, 텍스트 마이닝, TF-IDF, 토픽 모델링, 감성 분석, 네트워크 분석

Abstract With the rapid growth of the webtoon industry, a new model for webtoon applications has emerged. We have entered the era of webtoon application version 3.0 after ver 1.0 and ver 2.0. Despite these changes, research on user review analysis for webtoon applications is still insufficient. Therefore, this study aims to analyze user reviews for 'Kakao Webtoon (Daum Webtoon)' that presented the webtoon application 3.0 model. For analysis, 20,382 application reviews were collected and pre-processed, and TF-IDF, network analysis, topic modeling, and emotional analysis were conducted for each version. As a result, the user experience of the webtoon application for each version was analyzed and usability testing conducted.

Key words : Webtoon, User Experience, App Review Analysis, Text Mining, TF-IDF, Topic Modeling, Sentiment Analysis, Network Analysis

1. 서론

국내에서 웹툰이 시작된 것은 2002년 다음(Daum)이 온라인 만화 서비스를 시작한 것에서부터 출발한다. 그 후 2005년 네이버가 본격적으로 ‘웹툰(Webtoon)’이라는

이름으로 서비스를 출범하며 ‘웹 페이지로 보는 만화’가 하나의 서비스로 자리잡았다[1]. 한국콘텐츠진흥원에 따르면 2020년 국내 웹툰 시장 규모가 1조원을 넘어섰다고 한다[2]. 웹툰 시장은 2013년 1500억 원에서 6년 만에 7배 가까이 성장한 것으로, 매년 평균 20% 이상

*정회원, 연세대학교 정보대학원 UX전공 박사과정 (제1저자)
**정회원, 연세대학교 정보대학원 정교수 (교신저자)
접수일: 2022년 6월 23일, 수정완료일: 2022년 7월 4일
게재확정일: 2022년 7월 9일

Received: June 23, 2022 / Revised: July 4, 2022

Accepted: July 9, 2022

**Corresponding Author: uxlab.junhochoi@gmail.com

UX track, Graduate School of Information, Yonsei University, Korea

성장해 디지털 콘텐츠 분야 중에서 그 성장세가 두드러지게 나타난다. 현재 글로벌 웹툰 시장 규모는 7조 원 가량이며, 웹툰 IP(Intellectual Property)를 활용한 영화, 드라마 시장까지 100조원에 달할 정도로 커진다.

웹툰 시장의 성장세와 함께 2021년 8월 1일 다음웹툰이 카카오웹툰으로 정식 출범하며 20년 만에 새로운 웹툰 모델을 제시하였다. 웹툰을 모바일로 보기 시작한 때를 기준으로 보면, 10년 만에 완전히 다른 웹툰 애플리케이션이 등장한 것이다. 이들이 내세운 차별화 포인트는 새로운 유저 인터페이스(User Interface, UI)와 사용자 경험(User eXperience, UX)이다. UX 설계의 틀을 바꾸고 완전히 새로운 독창적인 디자인을 구현했다고 한다. 그러나 이러한 변화를 경험한 사용자의 평가는 대체로 부정적이었다[3]. 기존 다음웹툰에 비해 더 많은 동작을 필요로 하기에 사용성이 떨어졌다는 평가하였으며, 무한 스크롤의 조작 방식, 진입화면과 섬네일의 애니메이션 등에 대한 평가도 직관성이 떨어진다는 혹은 매력적이다 등으로 극명하게 나누어졌고, 다크 모드를 디폴트 값으로 제공하며 사용자 설정 권한을 축소시켜 사용자의 불만을 사기도 하였다.

이처럼 애플리케이션(application, app)에 대한 사용자 평가는 사용자 리뷰를 통해 확인할 수 있다. 애플리케이션에 대한 사용자 리뷰에는 애플리케이션에 대한 즉각적인 사용자 경험이 녹아있고, 사용자의 직접적인 피드백을 확인할 수 있다. 사용자 리뷰는 서비스 품질을 개선하고 누락된 기능을 보완하는 데 유용한 정보를 제공하기 때문에 애플리케이션 제작사는 이를 인식하고 사용자 리뷰 데이터로부터 필요한 정보를 획득하여 애플리케이션을 업데이트 하는 것이 매우 중요하다[4-5]. 그러나 제작사와 개발자가 사용자가 작성한 모든 리뷰를 매일 확인하는 것은 현실적으로 어려움이 있다. 다량으로 발생하는 애플리케이션의 모든 리뷰를 이해하기 위해서는 많은 시간이 필요하며 리뷰에 포함된 수많은 노이즈 단어들과 리뷰의 주제를 파악하고 해석하는 과정이 사용자 리뷰에 빠르고 정확하게 대응하기 어렵게 만든다[6-7]. 이와 같은 한계점에도 많은 연구들이 텍스트 마이닝 방법을 활용하여 사용자 리뷰 분석을 시도하고 있다.

따라서 이 연구는 웹툰 애플리케이션의 변화에 따른 사용자 경험 변화를 탐색하고자 한다. 특정 웹툰에 대한 댓글 분석은 연구가 진행된 바 있지만[8-9], 웹툰

플랫폼, 웹툰 애플리케이션 사용자 리뷰를 분석한 연구는 부족한 실정이다. 이에 최근까지 웹툰 애플리케이션에 가장 변화가 컸던 카카오웹툰과 이전 다음웹툰 버전을 대상으로 하여 TF-IDF, 네트워크 분석, 토픽모델링, 감성 분석과 같은 텍스트 마이닝 기법을 사용하여 웹툰 애플리케이션에 대한 사용자 경험을 분석하고자 한다.

II. 이론적 배경

1. 웹툰 애플리케이션 변화

모바일 기기의 확산으로 많은 플랫폼이 웹에서 앱으로 전환되는 시기에 웹툰 또한 모바일 서비스를 시작하였다. 이 시기를 웹툰 애플리케이션 1.0 시기로 웹 환경과 동일하게 세로형 화면으로 웹툰이 제공되었고, 스크롤을 통해 웹툰을 볼 수 있었다. 모바일로 제공될 때, 웹 환경으로부터 이어진 인터페이스를 유지하였기 때문에[10], 사용자가 어려움 없이 애플리케이션을 사용할 수 있었다. 그러나 정지된 시각적 표현만이 가능한 한계와 모바일 기기의 가장 큰 특징인 터치 기반의 인터페이스를 적극적으로 활용하지 못하는 측면이 있어[10], 웹툰 애플리케이션 2.0이 등장하게 되는 배경이 되었다.

웹툰 애플리케이션 2.0은 기술이 발전하고 표현 기법들이 다양해지면서 정지된 시각적 이미지 정보 외에도 배경음악, 모션 인터랙션 등의 효과들이 도입되었다. 정지된 이미지에서 벗어나 다채로운 시각적 정보와 청각적 정보까지 제공되어, 사용자 경험이 확장되었다. 기술과 표현 기법들은 다양해졌지만, 인터페이스의 큰 틀은 변화하지 않았기에 사용자의 웹툰 애플리케이션 1.0에서의 경험이 웹툰 애플리케이션 2.0에도 적용 가능했다. 즉, 사용자 입장에서는 더 편하게 웹툰 애플리케이션을 쓸 수 있었던 시기로, 이 시기를 웹툰 애플리케이션의 발전과 안정화 시기라고 볼 수 있다.

다음웹툰이 카카오웹툰으로 서비스명을 공식적으로 변경하며 웹툰 애플리케이션 3.0이 시작되었다. 지난 10년 간 경험하지 못했던 새로운 변화들이 도입되어 큰 틀에서는 유지되어오던 인터페이스가 완전히 바뀌었으며, 조작 방식도 끝이 없는 무한 확장 스크롤로 변경되었고, 애니메이션 섬네일을 도입하여 웹툰 애플리케이션 2.0에서 보다 더 화려하게 변화하였다. 웹툰 애플리케이션 3.0은 2.0에 비하여 기능과 디자인 측면에서 많은 변화가 있었다.

따라서 이 연구에서는 웹툰 애플리케이션 1.0에서부터 3.0까지 변화를 추적할 수 있는 다음웹툰을 분석 대상으로 선정하여 살펴보고자 한다.

표 1. 다음웹툰과 카카오웹툰 애플리케이션 특징
 Table 1. Daum and Kakao webtoon application features

시기	특징
다음웹툰 애플리케이션 1.0	- 웹에서 앱으로의 전환기 - 웹의 세로 스크롤 형식을 그대로 따름 - 시각적 표현만 가능
다음웹툰 애플리케이션 2.0	- 다양한 멀티미디어 효과 도입 - 터치 기반의 인터페이스 도입 - 모바일 미디어의 특징 반영
카카오웹툰 (구 다음웹툰) 애플리케이션 3.0	- 기존과 완전히 다른 새로운 인터페이스 개발 - 극대화된 모션그래픽 - 터치, 스와이프, 스크롤 등 조작 방식 개선
* 애플리케이션 1.0과 2.0의 특징은 [10] 참고	

2. 텍스트 마이닝을 활용한 사용자 리뷰 분석

애플리케이션 사용자 리뷰는 애플리케이션에 대한 사용자의 생생한 경험을 담고 있는 피드백으로, 사용자가 애플리케이션을 사용한 후 느낀 감정, 생각, 경험, 의견 등을 공유하는 것이다[11]. 대부분의 경우 애플리케이션 사용자 리뷰는 텍스트를 기반으로 한 비정형 데이터로 이루어져 있기 때문에, 기존의 정형 데이터를 가정하여 분석하는 전통적인 통계 기법을 적용하는 것은 적절치 못하다[7]. 또한, 사용자 리뷰는 명시적(explicit)일 수도 있지만, 암시적(implicit)일 수도 있기 때문에 텍스트 데이터에서 주제를 분류하거나 텍스트 이면의 감정을 예측하는 등 텍스트 데이터를 분석하는 기술을 사용해야 한다.

이러한 기술을 텍스트 마이닝(text mining)이라고 한다. 텍스트 마이닝이란 텍스트 데이터로부터 통계적인 의미가 있는 개념이나 특성을 추출하고 데이터 간의 패턴이나 추세 등의 인사이트를 끌어내는 과정이다[12]. 텍스트 마이닝 기법으로는 주제 분류, 정보 추출, 텍스트 요약, 감성 분석 등이 주로 사용된다. 애플리케이션 사용자 리뷰를 대상으로 한 텍스트 마이닝 연구는 주로 주제를 도출하는 토픽 모델링 방법과 텍스트 이면의 사용자의 감정을 파악하는 감성분석을 사용한 연구가 주를 이루고 있다.

텍스트 마이닝을 활용한 애플리케이션 사용자 리뷰 분석은 딜리버리(Delivery), 핀테크(FinTech), OTT

(Over The Top) 등 다양한 산업군을 대상으로 하고 있다. 각 산업군의 최근 연구들을 살펴보면 딜리버리 산업군에서는 배달 애플리케이션 사용자 리뷰를 수집하여 과도한 리뷰가 발생한 시점의 토픽을 중심으로 이슈 요인을 도출하였으며[7], 핀테크 산업군에서는 자산관리 핀테크 기업의 리뷰를 수집하여 토픽 별 감성 분석을 실시해 핀테크 기업에 필요한 핵심 요소를 추출하였다[13]. OTT 산업군에서는 12개의 OTT 서비스 리뷰의 토픽과 감성을 분석하여 서비스 별 강약점을 분석하였다[14].

이처럼 다양한 산업군에서 애플리케이션 사용자 리뷰 분석이 진행되고 있다. 그러나 잠재 시장 가치만 100조원인 웹툰 사업군에서는 웹툰 플랫폼을 대상으로 한 사용자 리뷰 분석 연구가 미비하다. 따라서 이 연구에서 웹툰 애플리케이션 사용자 리뷰를 대상으로 다양한 텍스트 마이닝 기법을 활용하여 분석을 진행하고자 한다.

3. 연구목적

웹툰 애플리케이션 2.0에서 웹툰 애플리케이션 3.0으로 서비스가 진화함에 따라, 웹툰 애플리케이션 3.0에 대한 탐구가 필요한 시점이다. 이에 이 연구는 카카오웹툰과 다음웹툰을 대상으로 웹툰 애플리케이션 1.0부터 3.0까지의 사용자 리뷰를 분석하여 사용자 경험의 변화를 탐색하고, 사용자 리뷰를 통한 사용성 평가를 진행하고자 한다.

이를 위해 네 가지의 텍스트 마이닝 기법을 사용하여 리뷰를 분석하고자 한다. TF-IDF를 통해 웹툰 애플리케이션 버전 별 사용 단어의 차이가 있는지 확인하고, 네트워크 분석을 통해 리뷰 내용에 차이가 있는지 확인하고자 한다. 그 후, LDA 토픽 모델링을 통해 웹툰 애플리케이션 버전 별 토픽을 추출하고, 토픽별 감성 분석을 실시하고자 한다. 이와 같은 분석을 통해 도출된 텍스트 마이닝 결과를 종합하여 사용자 경험의 변화를 정리하고, 현재 웹툰 애플리케이션 3.0의 사용성 이슈를 도출하고자 한다.

III. 연구 방법

이 연구는 텍스트 마이닝을 활용하여 사용자 리뷰를 통한 웹툰 앱 사용성 평가를 진행하기 위해 <그림 1>과

같은 연구 절차를 따라, 데이터 수집, 데이터 전처리, 데이터 분석을 실시하였다. 데이터 수집 단계에서는 구글 플레이 스토어에서 카카오편리뷰를 수집하였고, 데이터 전처리 단계에서는 명사 추출, 불용어 제거, 형태소 분석 등을 통해 데이터를 전처리하였으며, 데이터 분석 단계에서는 TF-IDF, LDA 토픽 모델링, 감성 분석, 네트워크 분석을 진행하였다.

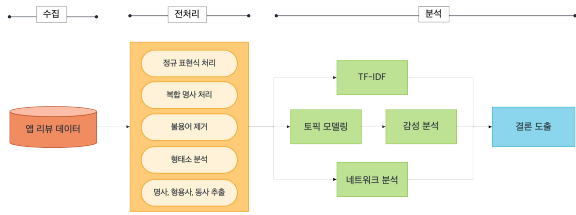


그림 1. 데이터 분석 절차
Figure 1. Process of data analysis

1. 데이터 수집

데이터 수집은 파이썬 기반의 Google-Play-Scrapers API[15]을 이용하여 구글 플레이 스토어의 ‘카카오편리뷰’ 애플리케이션 리뷰를 수집하였다. 데이터 수집은 국가 및 언어를 대한민국으로 설정하여 한국에서 쓰여진 모든 리뷰를 대상으로 하였고, 리뷰 작성자 닉네임, 리뷰 내용, 리뷰 작성 날짜, 리뷰 작성 시 사용된 앱 버전 정보를 수집하였다. 수집된 카카오편리뷰의 작성 기간은 2013년 6월 10일부터 2022년 4월 17일까지이며 데이터 수는 총 20,382개이었다. 웹툰 애플리케이션 변화에 따른 사용자 경험을 분석하고 사용성 평가를 진행하기 위해 수집된 데이터를 앱 버전 정보와 작성 날짜를 기준으로 2013년 6월 10일부터 2017년 9월 4일까지 작성된 리뷰를 애플리케이션 1.0 그룹으로, 2017년 9월 5일 2021년 7월 31일까지 작성된 리뷰를 애플리케이션 2.0 그룹으로, 2021년 8월 1일부터 2022년 4월 17일까지 작성된 리뷰를 애플리케이션 3.0 그룹으로 나누었다.

표 2. 데이터 수집
Table 2. Data collection

그룹명	리뷰 작성 기간	애플리케이션 버전 정보	개수
애플리케이션 1.0	2013년 06월 10일 ~ 2017년 09월 04일	1.x.x	8,265
애플리케이션 2.0	2017년 09월 05일 ~ 2021년 07월 31일	2.x.x	2,723
애플리케이션 3.0	2021년 08월 01일 ~ 2022년 04월 17일	3.x.x	9,394

2. 데이터 전처리

데이터 전처리 단계에서는 정규 표현식 처리를 통하여 특수문자, 기호 등을 제거하였다. 이후 바로 형태소 분석을 실시하면 형태소 사전에 등록되어 있지 않은 단어가 제대로 인식되지 않는 미등록단어 문제(out of vocabulary)가 발생할 수 있다. 즉, 웹툰 애플리케이션과 관련된 용어들이 누락될 확률이 높다는 것이다. 이를 해결하고자 비지도학습 기반의 통계적 패턴을 이용하여 명사를 추출하는 soynlp[16]의 Noun Extractor를 사용하여 단어 사전을 구축하였다. Noun Extractor는 댓글, 리뷰, 기사 등의 문서 집합에서 자주 등장하는 연속된 단어열을 단어로 정의하고 통계를 이용하여 명사를 추출하는 방법이며, 합성명사 또한 인식이 가능하다.

soynlp의 Noun Extractor를 통해 추출된 명사 단어 사전은 python TEANAPS 라이브러리를 사용하여 사전에 구축되어 있던 명사 단어 리스트에 새롭게 추가하였다. TEANAPS는 다양한 패키지를 하나의 API형태로 통합하여 텍스트 마이닝을 지원하는 python 언어 기반의 라이브러리이다[16]. TEANAPS를 통해 KoNLPy 패키지의 Okt 형태소 분석기를 사용하였고, 분석된 형태소 중에서 명사, 동사, 형용사로 태그된 형태소만을 분석에 사용하였다.

마지막으로 분석에 사용할 최종 데이터 셋을 구축하기 위해 데이터 분석에 불필요한 단어를 제거하는 불용어 처리 과정을 거쳤다. ‘웹툰’, ‘어플’, ‘다음’, ‘네이버’ 등의 사용 빈도는 높지만 분석에 유의미하지 않은 단어들과 한 글자 이하의 단어와 형태소, 빈도수가 20 미만인 단어와 형태소를 제거하였다. 이러한 과정을 통해 최종적으로 분석에 사용된 리뷰 수는 19,018개로, 버전 1 7,505개, 버전 2 그룹 2,346개, 버전 3 그룹 9,167개의 리뷰를 분석에 사용하였다.

3. 데이터 분석

1) TF-IDF

TF-IDF(Term Frequency - Inverse Document Frequency)는 정보 검색과 텍스트 마이닝에서 이용하는 가중치로, 문서 내의 각 단어들의 중요한 정도를 통계적 수치로 표현한다. TF(Term Frequency)가 특정한 단어가 문서 내에서 얼마나 자주 등장하는지를 나타내지만, 이것은 단순 단어 출현 빈도이기에 해당 단어가 문서 내에서 중요한 단어인지는 판단하기 어렵다.

이러한 문제를 해결하기 위해 단어가 출현한 문서가 적을수록 가중치를 높게 주는 역문서 빈도(Inverse Document Frequency)를 활용한다. 즉, TF-IDF는 단어가 문서에 출현한 빈도와 단어의 희박성을 동시에 활용하는 방법이다. TF-IDF 값은 TF와 IDF 수치의 곱을 통해 계산하며, 각 단어의 가중치를 나타낸다. 계산된 값을 통해 단어의 중요도를 파악하여 단어를 추출할 수 있다. 이 연구에서는 TD-IDF 가중치를 통해 버전별 웹툰 앱과 관련된 주요 키워드를 파악하고, 데이터에 대한 이해를 위해 이용하고자 한다.

2) 네트워크 분석

네트워크 분석은 사람, 그룹, 데이터 등 객체 간의 관계 및 특성 등을 분석하고 시각화하는 방법으로, 텍스트 마이닝에서는 텍스트 내의 단어들 간의 동시 출현(co-occurrence) 관계를 기반으로 단어 간 연결망을 구성하여 분석하는 방법이다. 즉, 텍스트 네트워크 분석을 통해 전체 문서에서 단어의 영향력을 확인하고, 단어들 간의 관계 속에서 단어의 역할을 파악할 수 있다 [18-22].

네트워크 분석에서는 단어를 노드(node)로 정의하고, 노드 간의 연결을 엣지(edge)로 정의하며, 네트워크 연결 구조를 파악하기 위한 지표로 밀도(Density), 중심성(Centrality) 등을 사용한다. 이 연구에서는 네트워크 연결 구조의 특성을 파악하기 위해 중심성 지표를 사용하였고, 매개 중심성(Betweenness Centrality), 근접 중심성(Closeness Centrality), 연결 중심성(Degree Centrality) 중에서 어떤 단어가 가장 많은 단어들과 같이 쓰였는가를 확인할 수 있는 연결 중심성으로 분석을 실시하였다. 네트워크 분석은 TEANAPS를 통해 NetworkX 패키지를 사용하였다.

3) LDA 토픽 모델링

토픽 모델링은 문서의 집합에서 토픽을 찾아내는 과정으로, LDA(Latent Dirichlet Allocation)는 토픽 모델링의 가장 대표적인 알고리즘으로, 문서가 여러 개의 토픽들의 혼합으로 구성되어 있다고 가정하며, 토픽은 디히클레 확률 분포(Dirichlet distribution)에 기반하여 단어가 토픽에 포함될 확률을 계산한다[23].

토픽 모델링은 토픽의 수를 연구자가 직접 지정해야 하는데, 일반적으로 최적의 토픽 수를 찾기 위해 혼잡도

(Perplexity)와 일관성(Coherence) 점수를 활용한다. 혼잡도는 토픽모델링의 결과가 문서의 내용을 얼마나 잘 반영하는지를 나타내는 수치로, 그 값이 낮을수록 토픽을 잘 반영하고 있는 것이다[24]. 일관성은 토픽의 일관된 정도를 나타내므로 분석 수치가 높을수록 좋은 결과를 나타낸다[25].

이 연구에서는 TEANAPS를 통해 Gensim과 pyLDAvis 패키지를 사용하여 토픽 모델링 분석을 진행하였다. 토픽의 수는 혼잡도와 일관성 값을 참고하여 토픽의 수를 결정하였다. 토픽의 개수를 2개에서 10개까지 비교하여 혼잡도는 감소하고 일관성은 증가하는 지점을 그룹별로 확인하였고, 동료 연구자와 협의하여 최종적으로 버전 1 그룹은 토픽의 수를 6개로, 버전 2 그룹은 토픽의 수를 4개로, 그룹 3은 토픽의 수가 5개로 하였다.

4) 감성 분석

감성 분석은 자연어 처리, 기계 학습 등의 기술을 사용하여 문장 또는 문서에서의 표현이 긍정, 부정 혹은 중립인지를 판별하는 방법이다[26]. 감성 분석은 크게 단어사전 기반 감성 분석과 머신러닝 기반 감성 분석으로 나누어진다. 단어사전 기반 감성 분석은 단어의 감정과 강도가 미리 정의된 감성사전을 기반으로 판단하여 문장의 감성을 판단하는 방식이다. 머신러닝 기반의 감성 분석은 지도 학습과 비지도 학습으로 다시 나누어지며, 지도 학습 기반의 감성분석은 감성이 정의된 텍스트를 학습한 모델을 사용하여 분석하고자 하는 텍스트의 감성을 분류 및 예측하는 방식이다. 비지도 학습 기반의 감성 분석은 문장 패턴이나 단어 간 상관관계 분석, 감성 사전 등을 활용하여 텍스트를 분석하는 여러 가지 방법을 포함한다[27].

이 연구에서 감성 분석은 BERT(Bidirectional Encoder Representations from Transformers)[28] 기반의 사전 학습모델을 활용하는 TEANAPS 라이브러리를 이용하여 감성 분석을 진행하였다. 감성 분석에 사용한 리뷰는 LDA 토픽 모델링 결과를 토대로, 리뷰가 토픽에 할당될 확률을 기준으로 리뷰의 관여도가 0.5 이상인 리뷰만을 감성 분석의 대상으로 하였다. 그 결과 분석에 사용된 리뷰는 총 15,427개로, 애플리케이션 1.0 그룹 리뷰 수는 5,698개, 애플리케이션 2.0 그룹 리뷰 수는 2,117개, 애플리케이션 3.0 그룹 리뷰 수는 7,612개를 감성 분석 대상으로 하였다. 감성 분석의 결과는 각 리뷰에

대한 긍정, 부정, 중립 여부를 판단하고, 그 정도를 0~1 사이 값으로 나타내었다. 정도 값이 0~0.4는 부정(negative) 리뷰로, 0.4~0.7은 중립(neutral) 리뷰로, 0.7~1은 긍정(positive) 리뷰로 산출하였다.

IV. 연구 결과

1. TF-IDF

웹툰 애플리케이션 버전에 따라 사용되는 단어들이 차이가 있는지 확인하기 위해 TF-IDF 분석을 실시하였다. 웹툰 애플리케이션 버전별 TF-IDF 결과는 <그림 2>, <그림 3>, <그림 4>와 같으며, TF-IDF 값을 기준으로 상위 100개의 단어를 추출하여 워드 클라우드(word cloud) 형태로 시각화한 것이다.

애플리케이션 1.0에서 최상위로 도출된 단어는 ‘좋아요’, ‘다음웹툰’, ‘기능’, ‘좋은’ 등과 같은 웹툰 플랫폼이 웹에서 앱으로 전환됨에 따른 만족감과 관련된 단어가 주로 나타났다. 애플리케이션 2.0의 최상위 도출 단어는 ‘캐시’, ‘업데이트’, ‘로그인’, ‘결제’, ‘오류’ 등으로 웹툰 앱에 익숙해진 사용자가 보고한 제 사항과 관련된 단어가 주로 나타났다. 애플리케이션 3.0은 ‘다음웹툰’, ‘업데이트’, ‘유저인터페이스’, ‘불편함’ 등과 같은 단어가 최상위 단어로 도출되었는데, 이는 앱 이름부터 모든 것이 개편된 새로운 카카오페이 웹툰 앱에 대한 평가와 관련된 단어이다.

2. 네트워크 분석

버전별 웹툰 애플리케이션의 리뷰가 어떠한 단어를 중심으로 구성되며, 리뷰 내용에 차이가 있는지 파악하기 위해 동시출현빈도를 계산하고 네트워크 분석을 실시하였다. 분석 결과는 <그림 5>, <그림 6>, <그림 7>과 같다. 원은 노드 사이즈를 표현한 것으로 단어 출현빈도가 높을수록 사이즈가 커지며, 원과 원의 연결선은 단어가 함께 출현함을 보여주고 두께가 두꺼울수록 동시출현빈도가 많다는 것이다.

애플리케이션 1.0의 네트워크 분석 결과를 보면, ‘다음웹툰’, ‘기능’, ‘업데이트’가 중심 노드가 되는 것을 확인할 수 있다. ‘다음웹툰’은 ‘좋아요’, ‘네이버웹툰’, ‘이제’ 등과 함께 사용되었고, ‘기능’은 ‘정주행’, ‘스크롤’, ‘좋네요’ 등과 함께 사용되었으며, ‘업데이트’는 ‘버전’, ‘삭제’ 등과 함께 사용되었다.



그림 2. 애플리케이션 1.0의 워드 클라우드
Figure 2. Word cloud of webtoon app 1.0



그림 3. 애플리케이션 2.0의 워드 클라우드
Figure 3. Word cloud of webtoon app 2.0



그림 4. 애플리케이션 3.0의 워드 클라우드
Figure 4. Word cloud of webtoon app 3.0

애플리케이션 2.0의 네트워크 분석 결과를 보면, ‘업데이트’, ‘캐시’, ‘다음웹툰’이 중심 노드가 되는 것을 확인할 수 있다. ‘업데이트’는 ‘작품’, ‘로딩’ 등과 함께 사용되었고, ‘캐시’는 ‘설치’, ‘결제’, ‘무료캐시’ 등과 함께 사용되었다.

애플리케이션 3.0의 네트워크 분석 결과를 보면, 다른 버전과 다르게 ‘다음웹툰’, ‘유저인터페이스’, ‘캐시’ 등 중심 노드로 볼 수 있는 단어들이 더 많고, 동시출현빈도도 많음을 확인할 수 있다. ‘다음웹툰’은 ‘기존’,

‘카카오웹툰’, ‘유저인터페이스’ 등과 함께 사용되었고, ‘유저인터페이스’는 ‘사용자경험’, ‘불편해요’, ‘불편함’ 등과 함께 사용되었으며, ‘캐시’는 ‘충전’, ‘이벤트’, ‘결제’ 등과 함께 사용되었다.

3. LDA 토픽 모델링

웹툰 애플리케이션의 버전에 따라 다루어지는 주제에 차이가 있는지 파악하기 위해 LDA 토픽 모델링을 실시하였고, 분석 결과는 <표 3>과 같다. LDA 토픽

표 3. 토픽 모델링 결과
 Table 3. Result of topic modeling

그룹	토픽	토픽 키워드	토픽 해석
애플리케이션 1.0*	디자인	좋아요, 깔끔하고, 편해요, 괜찮은, 심플	웹툰 애플리케이션 디자인에 관한 토픽
	연결 오류	와이파이, 네트워크, 계속, 로딩, 연결	연결 오류에 관한 토픽
	기능	기능, 정주행, 최고, 스크롤, 알림,	웹툰 애플리케이션에서만 사용할 수 있는 기능에 관한 토픽
	서비스 런칭	나왔네요, 다음웹툰, 감사합니다, 모바일, 간편하다	웹툰 애플리케이션 런칭에 대한 만족감에 관한 토픽
	사용 오류	오류, 계속, 로그인, 인증, 들어가서,	사용 오류에 관한 토픽
애플리케이션 2.0**	연결 오류	네트워크, 와이파이, 연결, 실행, 안되네요	연결 오류에 관한 토픽
	유료 결제	캐시, 결제, 이벤트, 유료, 무료, 광고	웹툰 유료 결제 서비스에 관한 토픽
	업데이트	업데이트, 리그, 작품, 디자인, 버전	업데이트로 변화된 애플리케이션에 관한 토픽
애플리케이션 3.0***	사용 오류	로그인, 오류, 계속, 설치, 삭제,	사용 오류에 관한 토픽
	연결 오류	계속, 오류, 업데이트, 접속, 연결	사용 오류에 관한 토픽
	기능	화면, 로딩, 스크롤, 처음, 불편해요	웹툰 애플리케이션이 도입한 새로운 기능에 관한 토픽
	유료 결제	캐시, 충전, 결제, 대여권, 유료	웹툰 유료 결제 서비스에 관한 토픽
3.0***	메뉴구조	다음웹툰, 소설원작, 웹툰원작, 작품, 요일별	웹툰 애플리케이션 메뉴 구조에 관한 토픽
	디자인	유저인터페이스, 디자인, 사용자경험, 직관성, 정신없음	웹툰 애플리케이션 디자인에 관한 토픽

* Perplexity : -5.495, Coherence : 0.423
 ** Perplexity : -7.093, Coherence : 0.366
 *** Perplexity : -7.805, Coherence : 0.505

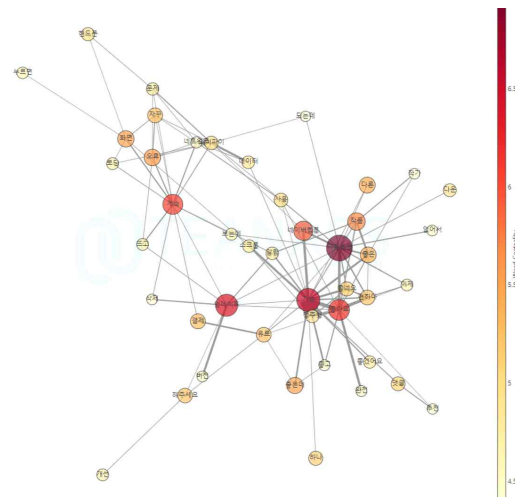


그림 5. 애플리케이션 1.0의 네트워크 분석 결과
 Figure 5. Network analysis result of webtoon app 1.0

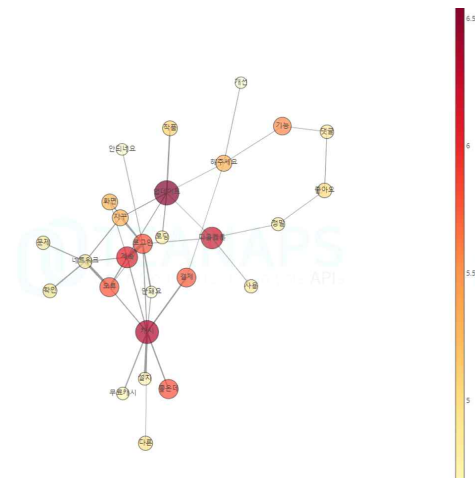


그림 6. 애플리케이션 2.0의 네트워크 분석 결과
 Figure 6. Network analysis result of webtoon app 2.0

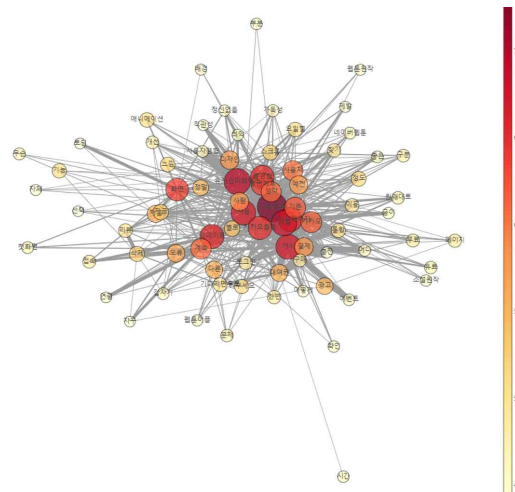


그림 7. 애플리케이션 3.0의 네트워크 분석 결과
 Figure 7. Network analysis result of webtoon app 3.0

모델링을 실시하기 위해 토픽의 개수가 결정되어야 하는데, 혼잡도(Perplexity)와 일관성(Coherence) 점수를 사용하여 버전별 웹툰 애플리케이션의 토픽 수를 선정하였다.

애플리케이션 1.0의 토픽은 ‘디자인’, ‘연결 오류’, ‘기능’, ‘서비스 런칭’, ‘사용 오류’, ‘개선 사항’으로 총 6개의 토픽을 도출하였다. ‘디자인’ 토픽은 ‘좋아요’, ‘깔끔하고’, ‘편해요’, ‘괜찮은’, ‘심플’로 구성되었으며, 사용자가 처음 경험하는 다음웹툰 애플리케이션 디자인에 대해 만족감을 표현하는 토픽이다. ‘연결 오류’ 토픽은 ‘와이파이’, ‘네트워크’, ‘계속’, ‘로딩’, ‘연결’의 키워드로 구성되었으며, 웹툰 애플리케이션에 접속이 어려움을 나타내는 토픽이다. ‘기능’ 토픽은 ‘기능’, ‘정주행’, ‘최고’, ‘스크롤’, ‘알림’의 단어로 구성되었으며, 이는 다음웹툰이 애플리케이션을 제작하면서 추가한 정주행모드, 자동스크롤, 알림 설정 등의 새로운 기능과 관련된 토픽이다. ‘서비스 런칭’ 토픽은 ‘나왔네요’, ‘다음웹툰’, ‘감사합니다’, ‘모바일’, ‘간편’의 단어로 구성되었으며, 사용자의 웹툰 애플리케이션 런칭에 대한 기대감과 만족감을 보여주는 토픽이다. ‘사용 오류’ 토픽은 ‘오류’, ‘계속’, ‘로그인’, ‘인증’, ‘들어가서’의 단어로 구성되었으며, ‘연결 오류’ 토픽과는 다르게 애플리케이션에 접속은 되지 만 그 후에 발생하는 문제에 관한 토픽이다. ‘개선 사항’ 토픽은 ‘업데이트’, ‘기능’, ‘네이버웹툰’, ‘개선’, ‘추가’의 단어로 구성되었으며, 타 웹툰 애플리케이션에 있는 기능이나 새로운 요구 사항들에 관한 토픽이다.

애플리케이션 2.0의 토픽은 ‘연결 오류’, ‘유료 결제’, ‘업데이트’, ‘사용 오류’로 총 4개의 토픽을 도출하였다. ‘연결 오류’ 토픽은 ‘네트워크’, ‘와이파이’, ‘연결’, ‘실행’, ‘안되네요’의 단어로 구성되었으며, 웹툰 애플리케이션 접속의 어려움을 보여주는 토픽이다. ‘유료 결제’ 토픽은 ‘캐시’, ‘결제’, ‘이벤트’, ‘유료’, ‘무료’, ‘광고’의 단어로 구성되었으며, 다음웹툰이 처음으로 돈을 내고 웹툰을 보는 유료 서비스를 시작했기 때문에 나타난 토픽이다. ‘업데이트’ 토픽은 ‘업데이트’, ‘리그’, ‘작품’, ‘디자인’, ‘버전’의 단어로 구성되었으며, 이전 버전에서 디자인적인 변화와 리그 웹툰의 삭제-복구의 내용을 담고 있는 토픽이다. ‘사용 오류’ 토픽은 ‘로그인’, ‘오류’, ‘계속’, ‘설치’, ‘삭제’의 단어로 구성되었으며, 웹툰 애플리케이션 사용의 어려움에 관한 토픽이다.

애플리케이션 3.0의 토픽은 ‘연결 오류’, ‘기능’, ‘유료

결제’, ‘메뉴구조’, ‘디자인’으로 총 5개의 토픽을 도출하였다. ‘사용 오류’ 토픽은 ‘계속’, ‘오류’, ‘업데이트’, ‘접속’, ‘연결’의 단어로 구성되어있으며, 업데이트 후 접속이 잘 되지 않음을 보여주는 토픽이다. ‘기능’ 토픽은 ‘화면’, ‘로딩’, ‘스크롤’, ‘처음’, ‘불편해요’의 단어로 구성되었으며, 이는 카카오펀터이 새로운 기능으로 제공한 무한 스크롤이나 애니메이션 요소가 들어간 첫 화면 등과 관련된 토픽이다. ‘메뉴구조’ 토픽은 ‘다음웹툰’, ‘소셜원작’, ‘웹툰원작’, ‘작품’, ‘요일별’의 단어로 구성되었으며, 이는 다음웹툰 버전일 때는 없었던, 웹툰을 원작으로 구별하는 메뉴와 그 하위에 요일별 메뉴를 추가하였기에 나타난 토픽이다. ‘디자인’ 토픽은 ‘유저인터페이스’, ‘디자인’, ‘사용자경험’, ‘직관성’, ‘정신없음’의 단어로 구성되었으며, 새로운 카카오펀터의 디자인과 관련된 토픽이다.

4. 토픽별 감성 분석

감성 분석은 LDA 토픽 모델링을 통해 도출된 각각의 토픽별로 진행하였다. 토픽별 감성 분석 결과는 <표 4>와 같다. 감성 분석을 통해 각각의 토픽별로 감성 지수를

표 4. 감성 분석 결과

Table 4. Result of sentiment analysis

그룹	토픽*	감성 지수**	긍정 리뷰 수	부정 리뷰 수	중립 리뷰 수	전체 리뷰 수
애플리케이션 1.0	디자인	0.6237	1,087	239	71	1,397
	연결 오류	0.0774	340	313	75	728
	기능	0.9498	886	306	110	1,302
	서비스 런칭	0.9439	576	227	67	870
애플리케이션 2.0	사용 오류	0.4403	318	230	41	589
	개선 사항	0.9532	493	255	64	812
	연결 오류	0.0316	173	272	51	496
	유료 결제	0.0623	139	174	23	336
애플리케이션 3.0	업데이트	0.0316	352	310	53	715
	사용 오류	0.1319	200	309	61	570
	연결 오류	0.4375	307	1,103	149	1,559
	기능	0.0379	201	590	79	870
애플리케이션 3.0	유료 결제	0.1238	420	941	126	1,487
	메뉴구조	0.2843	256	939	11	1,306
	디자인	0.0306	465	1,794	140	2,399

* 리뷰의 토픽 관여도가 0.5 이상인 리뷰만 감성 분석 실시

** 감성 지수 : 0~0.4 부정(negative), 0.4~0.7 중립(neutral), 0.7~1 긍정(positive)

도출하였으며, 이를 통해 토픽의 긍정, 부정 여부를 판단하였다. 그 결과, 애플리케이션 1.0의 '디자인' 토픽의 감성 지수는 0.6237, '연결 오류' 토픽의 감성 지수는 0.0774, '기능' 토픽의 감성 지수는 0.9489, '서비스 런칭' 토픽의 감성 지수는 0.9439, '사용 오류' 토픽의 감성 지수는 0.4403, '개선 사항' 토픽은 0.9532으로 나타났다.

애플리케이션 2.0에서 '연결 오류' 토픽의 감성 지수는 0.0315, '유료 결제' 토픽의 감성 지수는 0.0623, '업데이트' 토픽의 감성 지수는 0.0316, '사용 오류' 토픽의 감성 지수는 0.1319로 나타났다.

애플리케이션 3.0에서 '연결 오류' 토픽의 감성 지수는 0.4375, '기능' 토픽의 감성 지수는 0.0379, '유료 결제' 토픽의 감성 지수는 0.1238, '메뉴구조' 토픽의 감성 지수는 0.2843, '디자인' 토픽의 감성 지수는 0.0306으로 나타났다.

V. 결론 및 논의

이 연구는 다양한 텍스트 마이닝 기법을 사용하여 웹툰 애플리케이션의 버전에 따른 사용자 경험의 차이를 리뷰를 통해 알아보고자 하였다. TF-IDF, 네트워크 분석, LDA 토픽 모델링, 네트워크 분석 결과를 종합해보면, 사용자는 웹툰 애플리케이션 1.0에서는 웹툰 애플리케이션이 런칭한다는 사실에 기뻐하였으며, 애플리케이션이 존재한다는 것에 만족하고 있었음이 나타났다. 또한, 웹에서 앱으로 전환되며 애플리케이션에서만 사용할 수 있는 기능들에 굉장히 만족하는 것으로 나타났다. 정주행 기능과 자동 스크롤 등은 TF-IDF 워드 클라우드와 네트워크 분석 결과에서 구체적으로 언급되었고, 토픽 모델링에서도 '기능' 토픽이 따로 도출되어 긍정적으로 평가를 받았다.

반면, 웹툰 애플리케이션 2.0에 대해서는 사용자 경험이 긍정과 부정으로 극명하게 다르게 나타났다. LDA 토픽 모델링 결과에 따르면 사용자가 모든 토픽에서 부정적인 경험을 한 것으로 나타났지만, 이러한 결과를 다음웹툰 애플리케이션에 큰 문제가 있었다고 해석하는 것보다 긍정 어휘와 부정 어휘가 공존하는 TF-IDF 워드 클라우드와 애플리케이션 2.0 시기가 웹툰 애플리케이션의 안정화 시기라는 점, 그리고 상대적으로 적은 리뷰 수를 고려했을 때, 정말로 불만족한 경우에만 리뷰를 작성했기 때문에 사용자의 부정적인 경험이

두드러지는 결과가 나왔다고 보는 것이 더 적합하다고 보여진다.

마지막으로, 서비스명부터 모든 것이 전면 개편된 카카오페이(구 다음웹툰) 애플리케이션 3.0은 TF-IDF 결과에서도 부정 단어의 출현이 많았으며, 토픽별 감성 분석 결과도 같은 경향성을 보이고 있다. 특히나 카카오페이의 디자인과 기능 측면에서 부정 감성 지수가 높게 나타났다. LDA 토픽 모델링 결과와 카카오페이가 새롭게 추가한 기능들을 종합해보면 대표적으로 상하 좌우 무한 스크롤과 애니메이션 섬네일 화면이 사용자에게 정신없고 불편하다는 평가를 받고 있음을 알 수 있다. 또한, '메뉴구조' 토픽도 부정적인 평가를 받았는데, 사용자 입장에서는 요일별로 구분하던 웹툰이 소셜 원작과 웹툰원작으로 나누어지고 다시 요일별로 구분되다보니, 이전에는 없었던 번거로운 뎀스(depth)가 추가된 것이라 느껴질 수 있다. 그러므로 웹툰을 보는 사용자 입장에서 원작을 구분하는 것이 정말 중요하고 의미가 있는 일인지, IPX(Intellectual Property eXperience)를 새로운 UX라 정의한 카카오페이에게 중요했던 것은 아닌지 의문을 제기한다.

애플리케이션 3.0의 리뷰는 애플리케이션 1.0과 2.0에 비하면 리뷰가 누적된 기간이 짧은 편임에도 불구하고, 가장 많은 리뷰 수를 가지고 있었고 다수의 리뷰가 부정적인 것으로 확인되었다. 이러한 카카오페이에 대한 사용자의 부정적인 경험을 멘탈 모델 측면으로 이해해보고자 한다. 멘탈 모델(Mental Model)은 일종의 인지적 틀(Cognitive Framework)으로써 개인을 둘러싼 여러 사물을 이해하는데 필요한 지식들로 이루어져 있다 [29-31]. 이와 같은 멘탈 모델을 통해 사용자는 서비스를 이해하고 예측하게 되는데, 사용자가 가진 멘탈 모델이 적용되지 않는 경우 사용자는 이해의 어려움을 겪고 당황하게 되며 서비스를 사용하지 않고 이탈할 수도 있다. 즉, 버전 3의 카카오페이(구 다음웹툰) 애플리케이션에 대한 사용자의 부정적인 경험은 10여년 넘게 다음웹툰 애플리케이션과 타 웹툰 애플리케이션을 사용하면서 만들어진 사용자의 멘탈 모델이 카카오페이에는 적용되지 않았고, 그렇기 때문에 카카오페이의 디자인적인 측면과 기능적인 측면에서 불편함과 낯설음을 경험하게 되어 나타난 결과라고 볼 수 있다.

그러나 사용자의 멘탈 모델을 새로운 경험을 통해 수정되거나 사용자가 새로운 멘탈 모델을 만들 수 있다.

애플리케이션을 경험하면 할수록 애플리케이션에 대한 멘탈 모델이 점차 정확해지고 확장된다는 연구결과[32]처럼 현재는 상하좌우 무한 스크롤, 애니메이션 섬네일, 다크모드 등 카카오웹툰이 제시한 새로운 UI와 UX에 불편함을 느끼는 것이 사실이지만, 반복적인 경험으로 새로운 멘탈 모델을 형성하여 다른 새로운 어플리케이션에 응용해서 사용할 수도 있을 것이다. 카카오웹툰 서비스가 시작되고 약 8~9개월이 지난 최근까지도 UI/UX에 불편함을 느끼는 사용자가 여전히 존재하기에, 카카오웹툰의 슬로건인 ‘세상에 없던 새로운 UI/UX’가 사용자의 멘탈 모델의 수정과 확장으로 이어질지, 멘탈 모델의 불일치로 남을지 계속해서 지켜볼 필요가 있다.

이 연구의 한계점과 후속 연구에 대한 제언은 다음과 같다. 첫째, 토픽 모델링을 통한 토픽을 선정하고 해석하는데 연구자의 주관적 판단이 개입되었다는 것이다. 객관적인 지표들을 활용하고 동료 연구자와 합의하여 한계점을 완화하고자 노력하였다. 후속 연구에서는 문맥까지도 파악이 가능한 자연어 처리 기술들을 활용하여 연구자의 주관의 개입을 최소화할 필요가 있다.

둘째, 도출된 토픽은 수집된 리뷰에 한하여 적용이 가능하다는 점이다. 이후 추가로 작성된 리뷰 데이터를 포함하여 분석한다면, 이 연구의 토픽과 다른 토픽이 도출될 수 있다. 그렇지만 후속 연구에서 이 연구의 결과를 참고하여 데이터 전처리를 하거나 해석에 방해가 되지 않도록 불용어 처리를 한다면, 앞으로 작성될 리뷰들을 분석하는데 도움이 될 것이다.

이러한 한계점에도 이 연구는 웹툰 애플리케이션을 대상으로 변화 과정을 추적하여 사용자 경험을 탐구했다는 의의가 있다. 또한, 웹툰 애플리케이션 사용자 리뷰를 다양한 텍스트 마이닝 기법을 활용하여 사용자 경험 분석, 사용성 평가를 시도하였다는 점에서 방법론적 의의를 가진다. 또한, 멘탈 모델을 활용하여 새로운 서비스를 받아들이는 사용자의 행위를 해석한 점에서 이론적 의의가 있다. 이 연구의 결과가 웹툰 산업군이 발전하는데 활용되어 더 나은 웹툰 사용자 경험을 제공할 수 있도록 도움을 줄 수 있기를 기대한다.

References

- [1] J. H. Kim, “Grab the 100 trillion won market! NAVER vs Kakao. ‘Webtoon War’”, *Weekly Donga*, 2021. Available : [https://weekly.donga.com](https://weekly.donga.com/BestClick/3/all/11/2867878/1)
- [2] Korea Creative Content Agency, “2020 A Survey on Webtoon Businesses,” *KOCCA*, 2020.
- [3] J. H. Yoon, “Glamorous but unfamiliar .3 points that raised dissatisfaction with Kakao Webtoon,” *MoneyToday*, 2021. Available : <https://news.mt.co.kr/mtview.php?no=2021081014563988784>
- [4] Gao, C., Zeng, J., Lyu, M. R., and King, I., “Online app review analysis for identifying emerging issues,” In *Proceedings of the 40th International Conference on Software Engineering*, pp. 48–58, 2018.
- [5] Gu, X., and Kim, S., “what parts of your apps are loved by users?,” In *2015 30th IEEE/ACM International Conference on Automated Software Engineering(ASE)*, pp. 760–770, 2015.
- [6] Pagano, D., and Maalej, W., “User feedback in the appstore: An empirical study,” In *2013 21st IEEE international requirements engineering conference*, pp. 125–134, 2013.
- [7] J. H. Jung, H. I. Chung, and Z. K. Lee, “An Analysis of Mobile Food Delivery App ‘Baemin’ by Using Text Mining and ARIMA Model,” *Journal of Digital Contents Society*, Vol. 22, No. 2, pp. 291–299, 2021.
- [8] Lee, Y., So, H., Kwahk, K. Y., and Ahn, H., “A Study of Webtoon comments using Text mining: Focusing on Naver’s Best Challenge Webtoon,” In *Proceedings of the Korean Society of Computer Information Conference*, pp. 219–222, 2020.
- [9] Lee, D. K., Lee, S. J., and Choi, I. Y., “Response Analysis of Stop Smoking Campaign Webtoon Using Text Mining Technology: Focused on the ‘Tale of Cigarette’ Comments,” *Journal of Health Informatics and Statistics*, Vol. 43, No. 1, pp. 70–79, 2018.
- [10] S. J. Lee, and B. G. Jun, “The Current Service Status and the Developmental Direction of Webtoon 2.0 : Focusing on the Changing of User Interface and the Applying of Multimedia Effects,” *JOURNAL OF THE KOREA CONTENTS ASSOCIATION*, Vol. 15, No. 8, pp. 96–108, 2015.
- [11] Morales-Ramirez, I., Munante, D., Kifetew, F., Perini, A., Susi, A., and Siena, A., “Exploiting user feedback in tool-supported multi-criteria requirements prioritization,” In *2017 IEEE 25th International Requirements Engineering Conference*, pp. 424–429, 2017.
- [12] Lee, W. J., “A study on word cloud techniques

- for analysis of unstructured text data,” *The Journal of the Convergence on Culture Technology*, Vol. 6, No. 4, pp. 715-720, 2020.
- [13]Son, A., Shin, W., and Lee, Z, “An Analysis of Key Elements for FinTech Companies Based on Text Mining: From the User’s Review.,” *The Journal of Information Systems*, Vol. 29, No. 4, pp. 137-151, 2020.
- [14]H. S. Cho, S. A. Kang, and M. H. Ryu, “An Analysis of OTT Service Review Using Text Mining: Focusing on the Competitive Advantage of Local Service,” *The Journal of Korean Institute of Communications and Information Sciences*, Vol. 46, No. 4, pp. 722-733, 2021.
- [15]M. G. Jo, “Google-Play-Scraper,” GitHub, 2019. Available : <https://github.com/JoMingyu/google-play-scraper>
- [16]H. Y. Kim, “soynlp”, GitHub, 2021. Available : <https://github.com/lovit/soynlp>
- [17]TEANAPS, “TEANAPS: Text Analysis APIs for Education”, GitHub, 2019, Available : <http://teanaps.com>
- [18]H. G. Shim, Y. C. Kim, H. Y. Shon, and J. Y. Lim, “An Exploratory Usage Pattern Research of Smartphone and Social Media Users through Semantic Network Analysis : Gender and Age Differences in Perception and Evaluation of Usage Pattern,” *Korean Journal of Broadcasting and Telecommunication Studies*, Vol. 25, No. 4, pp. 82-138, 2011.
- [19]C. S. Park, and J. S. Lee, “Steps of Text Data Cleaning: for Network Text Analysis Using Large-scale Data,” *Modern Society and Public Administration*, Vol. 27, No. 4, pp. 35-68, 2017.
- [20]J. S. Han, and J. H. Yoon, “Activation Strategies of the 20th BIFF using Social Big Data Text Mining Analysis,” *Journal of Tourism Sciences*, Vol. 40, No. 1, pp. 133-145, 2016.
- [21]Wasserman, S., Faust, K., Iacobucci, D., & Granovetter, M., “Social Network Analysis: Methods & Applications,” London, England: Cambridge University Press, 1994.
- [22]P. L. Chung, H. C. Ahn, and K. Y. Kwahk, “Identification of Core Features and Values of Smartphone Design using Text Mining and Social Network Analysis,” *Korean Journal of Business Administration*, Vol. 32, No. 1, pp. 27-47, 2019.
- [23]Blei, D. M., Ng, A. Y., and Jordan, M. I., “Latent dirichlet allocation,” *Journal of machine Learning research*, Vol. 3, pp. 993-1022, 2003.
- [24]Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J., and Blei, D., “Reading tea leaves: How humans interpret topic models,” *Advances in neural information processing systems*, Vol. 22, 2009.
- [25]Newman, D., Lau, J. H., Grieser, K., and Baldwin, T., “Automatic evaluation of topic coherence,” In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pp. 100-108, 2010.
- [26]Telecommunications Technology Association, “Information and communication terminology dictionary,” 2022. Available : <http://terms.tta.or.kr>
- [27]M. C. Song, and K. S. Shin, “Construction of Consumer Confidence index based on Sentiment analysis using News articles,” *Journal of Intelligence and Information Systems*, Vol. 23, No. 3, pp. 1-27, 2017.
- [28]Devlin, J., Chang, M. W., Lee, K., & Toutanova, K., “Bert: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint arXiv:1810.04805, 2018.
- [29]Bower, G. H., and Morrow, D. G., “Mental models in narrative comprehension,” *Science*, Vol. 247, pp. 44-48, 1990.
- [30]Carley, K., and Palmquist, M., “Extracting, representing, and analyzing mental models,” *Social forces*, Vol. 70, No. 3, pp. 601-636, 1992.
- [31]Atman, C. J., Bostrom, A., Fischhoff, B., and Morgan, M. G., “Designing risk communications: completing and correcting mental models of hazardous processes, Part I,” *Risk Analysis*, Vol. 14, No. 5, pp. 779-788, 1994.
- [32]W. J. Jung, “The Effects of User Mental Models on the Intention to Use Smartphone Applications,” *The e-Business Studies*, Vol. 15, No. 3, pp. 339-361, 2014.

* APPENDIX (상위 50개 단어 TF-IDF값)

순위	버전 1			버전 2			버전 3		
	word	TF	TF-IDF	word	TF	TF-IDF	word	TF	TF-IDF
1	좋아요	1557	878.2895485	캐시	453	177.2717575	다음웹툰	2014	204.3050487
2	다음웹툰	691	168.4884179	업데이트	406	63.94845948	캐시	1459	192.4754171
3	기능	628	142.5318938	로그인	402	63.60027027	업데이트	1170	167.2169278
4	업데이트	370	113.1308455	좋아요	314	59.33355382	유저인터페이스	1150	161.5195144
5	좋은	292	110.0301869	다음웹툰	196	30.17305377	카카오페이지	1006	136.8638097
6	계속	277	104.9145009	결제	181	29.26520319	작품	927	134.7362748
7	네이버웹툰	274	103.2202903	계속	172	28.97674493	불편함	866	134.4485095
8	최고	265	72.76592677	오류	162	26.4017972	사용	814	121.6484384
9	오류	260	69.2168881	해주세요	130	23.32378417	카카오웹툰	772	104.4938857
10	좋네요	259	65.59316249	좋은데	125	22.45918885	불편해요	767	88.03185539
11	정주행	235	64.85374126	네트워크	120	21.22291533	화면	663	87.76377472
12	편하다	210	62.37390672	댓글	120	19.93262149	카카오	656	87.02879141
13	유료	205	60.81708549	기능	114	18.30267701	기존	645	83.7821926
14	와이파이	201	60.68592832	작품	113	18.27910261	계속	622	83.49175732
15	화면	200	60.33436494	로딩	98	18.17110708	생각	538	81.35025753
16	작품	188	58.64744849	정말	92	18.09735831	결제	530	78.1765601
17	좋은데	187	55.43212254	자꾸	92	17.48054118	디자인	526	76.80352506
18	결제	185	51.05147318	무료캐시	90	17.46140864	사용자	513	75.392438
19	해주세요	185	46.18718353	문제	89	17.34791821	오류	487	73.69410873
20	버전	185	44.29617704	설치	88	17.14612237	대여권	465	70.00601066
21	좋은	170	43.46573657	사용	85	16.71911857	예전	441	66.24355178
22	재미있어요	168	42.36431324	와이파이	79	16.28837544	사람	430	66.19566031
23	재미있는	165	40.65623486	리그	77	15.27980261	통합	387	65.05632805
24	네트워크	154	39.96029238	참여	75	15.15586456	삭제	371	59.92645377
25	자꾸	151	38.16300543	화면	75	14.30259799	처음	367	59.24835515
26	알림	148	37.87048904	다른	74	14.16061562	별로	362	58.95732175
27	완전	147	37.66382498	확인	73	14.01079753	접속	356	58.71375207
28	저장	147	37.21670032	카카오	72	13.92527434	리뷰	355	56.79292412
29	다른	146	37.14302128	사용자	70	13.8881329	기다리면무료	351	54.3802044
30	댓글	144	36.25124737	디자인	70	13.4656068	광고	346	53.81140276
31	삭제	139	36.0482609	연결	67	13.45478096	해주세요	345	53.76665953
32	사용	136	34.94340613	제발	67	13.35636559	로그인	334	51.85355775
33	스크롤	131	33.61773589	무료	64	13.16804044	좋아요	330	51.58567732
34	다운	131	33.56417693	개선	62	13.01501523	정말	327	50.77348999
35	좋겠어요	123	33.49206903	충전	62	12.78165483	다른	321	50.48381047
36	임시	122	33.26721075	이벤트	62	12.17789316	스크롤	319	50.11439582
37	인증	120	33.17869055	네이버웹툰	60	12.15601294	요일별	317	50.02787868
38	좋고	118	32.71986228	좋은	59	11.9903624	구매	312	48.90725995
39	자동	117	32.24852108	유료	58	11.8613306	연결	312	48.87513827
40	작가	117	31.76255676	핸드폰	58	11.84919816	최악	303	47.85646659
41	핸드폰	115	31.62795062	아이디	58	11.75459595	애니메이션	294	47.78835705
42	이제	113	31.56393155	하나	57	11.58815853	느낌	292	47.70875223
43	로딩	112	31.28524517	통합	57	11.49845149	어디	290	44.83493898
44	하나	109	30.20938436	지급	53	11.44942815	무료	288	44.64993483
45	목록	108	28.62899865	리뷰	53	11.28364662	찾기	285	42.66281637
46	문제	107	28.20832314	어떻게	53	10.90158032	충전	284	40.46909856
47	데이터	107	27.47375149	처음	52	10.68965278	지금	274	40.18846391
48	감사합니다	106	27.15484176	갑자기	51	10.66614195	정도	265	39.53199165
49	좋아	106	26.95029558	좋겠어요	51	10.42966634	이용	258	38.94529767
50	좋습니다	104	26.86956294	안되네요	51	10.28457673	사용자경험	255	38.53018408