

텍스트마이닝을 활용한 러닝 어플리케이션 사용자 리뷰 분석: Nike Run Club과 Runkeeper를 중심으로

류기문¹, 김일광^{2*}

¹한국체육대학교 일반대학원 석사, ²한국체육대학교 스포츠산업학과 교수

Analysis of User Reviews of Running Applications Using Text Mining: Focusing on Nike Run Club and Runkeeper

Gimun Ryu¹, Ilgwang Kim^{2*}

¹Master, Graduate School of Korea National Sport University

²Professor, Department of Sport Industry, Korea National Sport University

요약 본 연구의 목적은 텍스트마이닝을 활용하여 러닝 어플리케이션 사용자의 리뷰를 분석하였다. 본 연구는 python3의 selenium 패키지를 이용하여 google playstore의 Nike Run Club, Runkeeper의 사용자 리뷰들을 분석자료로 이용하였으며, okt 분석기를 통해 한글 명사만을 남겨 형태소를 분리하였다. 형태소 분리 후 rankNL 사전을 만들어 불용어(stopword)를 제거하였다. 자료 분석을 위해 텍스트마이닝의 TF(빈도분석), TF-IDF(키워드 빈도-문서 역빈도), LDA 토픽모델링을 통해 분석하였다. 본 연구의 결과는 다음과 같다. 첫째, Nike Run Club, Runkeeper 어플리케이션 사용자 리뷰에서 공통적으로 상위 키워드로 '기록', '앱', '운동'의 키워드가 도출되었으며 TF, TF-IDF의 순위에는 차이가 나타났다. 둘째, Nike Run Club의 LDA 토픽모델링으로 '기본 항목', '추가 기능', '오류 사항', '위치기반데이터'의 토픽이 도출되었고 Runkeeper는 '오류 사항', '음성 기능', '러닝 데이터', '사용 혜택', '사용 동기'의 토픽이 도출되었다. 결과를 통해 제안하면 어플리케이션의 경쟁력 향상을 기여하기 위해 오류 및 개선사항을 보완해야 한다.

키워드 : 텍스트마이닝, TF(키워드 빈도), TF-IDF(키워드 빈도-문서 역빈도), LDA 토픽모델링, 사용자 리뷰, Nike run club, Runkeeper

Abstract The purpose of this study was to analyze user reviews of running applications using text mining. This study used user reviews of Nike Run Club and Runkeeper in the Google Play Store using the selenium package of python3 as the analysis data, and separated the morphemes by leaving only Korean nouns through the OKT analyzer. After morpheme separation, we created a rankNL dictionary to remove stopwords. To analyze the data, we used TF, TF-IDF and LDA topic modeling in text mining. The results of this study are as follows. First, the keywords 'record', 'app', and 'workout' were identified as the top keywords in the user reviews of Nike Run Club and Runkeeper applications, and there were differences in the rankings of TF and TF-IDF. Second, the LDA topic modeling of Nike Run Club identified the topics of 'basic items', 'additional features', 'errors', and 'location-based data', and the topics of Runkeeper identified the topics of 'errors', 'voice function', 'running data', 'benefits', and 'motivation'. Based on the results, it is recommended that errors and improvements should be made to contribute to the competitiveness of the application.

Key Words : Text mining, TF(Term frequency), TF-IDF(Term frequency-inverse document frequency), LDA topic modeling, User reviews, Nike run club, Runkeeper

*Corresponding Author : Ilgwang Kim(whyhow2070@gmail.com)

Received January 18, 2024

Accepted April 20, 2024

Revised March 8, 2024

Published April 28, 2024

1. 서론

소프트웨어 기술이 발전함으로 인해, 소셜미디어 플랫폼, 디지털, 스마트폰 기술로 연결된 사물인터넷 등이 증가하고 있다. 그로 인해 사람들은 매일 사용하는 모바일 기기, 기계 센서, 컴퓨터에서 나오는 방대한 데이터로 구성된 빅데이터 시대에 살고 있다.

빅데이터는 높은 다양성, 거대한 규모, 복잡성, 많은 양과 빠른 속도의 특징으로 하는 데이터이다[1]. 빅데이터는 디지털 기술 기반의 인터넷 환경도 다양한 변화를 맞이하고 있다. 이러한 배경으로, 기술 스마트 기기의 업그레이드 속도가 점점 빨라지고, 그 기술 실현 능력도 점점 더 세밀하게 발전하여, 사람들의 스마트폰 이용 및 보유율이 지속적으로 증가하는 현상이 나타나고 있다.

스마트폰 보유율은 2010년 3.8%에서 2021년 93.4%로 지속적으로 증가하여, 보편적 사람들의 매체로서 자리 잡게 되었다[2]. 스마트폰과 같은 스마트 기기를 보유하게 되면서, 인터넷 접속률 및 접근성이 편리해지고 있다. 과학기술정보통신부의 2020 인터넷 이용 실태조사에 의하면, 가구 인터넷 접속률이 99.7%로 나타났고, 개인별 이용 시간이 전년 대비 2.7시간 증가하였다[3]. 스마트폰 이용자가 증가함으로 인해, APP 스토어와 구글 플레이스토어 등과 같은 어플리케이션 시장을 주목하게 되었다[4].

어플리케이션이란 스마트폰에서 실행하는 응용프로그램을 뜻하며, 줄임말로 앱이라고 부르기도 한다. 어플리케이션을 통해 사람들은 정보 탐색의 경험, 다른 사람들과의 정보 공유 기회, 필요한 정보 제공 등을 확대하면서 사회적 네트워크 발판의 장으로 발전시키고 있다[5]. 모바일 분석 및 데이터 플랫폼인 앱애니(App Annie)에 의하면, 전 세계 모바일 어플리케이션 설치 수는 2020년 2,180억건에서 2021년 2,300억건으로 5%의 성장세를 보였다[6]. 이러한 어플리케이션 설치 증가의 영향으로 하루 평균 이용시간이 증가하였다. 한국 사람들의 하루 평균 모바일 앱 사용시간은 2019년 4.1시간에서 2021년 5시간 이상으로 증가했고, 하루 평균 이용시간은 전 세계 3위를 기록할 정도로 많이 사용하고 있다[6]. 스마트폰과 어플리케이션 발전으로 사람들의 삶이 바뀌기 시작했다[7]. 이러한 변화가 어플리케이션을 통해 이용자들이 모이지 않고 비대면으로, 각자가 원하는 장소에서 함께 달리고 기록을 공유하는 등 스포츠 활동을 지원하는 스포츠 어플리케이션에도 영향을 주기 시작했다.

어플리케이션은 모바일 기기에서 작업 효율성을 향상시키는 콘텐츠 응용프로그램이라고 할 수 있고[8], 리뷰는 제품 소비 및 사용 이후 소비자가 경험한 제품과 서비스에 대한 감정을 의견(텍스트)과 점수(별점), 사진 등으로 표현하는 방식이다[9]. 어플리케이션 리뷰는 텍스트 데이터인 댓글과 수치 데이터인 평점으로 어플리케이션을 이용한 후 이용자가 남기는 후기이다[10]. 소비자들은 구매한 제품과 서비스에 대한 느낌을 작성하고, 다른 사람들과 공유하여 정보를 교환하기 시작하였다. 이를 통해 사용자는 다른 사람들의 사용 리뷰를 바탕으로, 소비자가 제품 구매 결정을 내리는데 영향을 미치게 되는 중요한 부분이라고 할 수 있다[11]. 사용자 리뷰에서 개선된 서비스를 제공하기 위해서, 사용자 요구에 적합한 유용한 정보에 대한 온라인 분석 기능이 필요하다[12]. 사용자 리뷰를 분석 및 이해하는 것은 제품, 서비스의 품질 개선 및 발전을 위하여 기업에게 중요한 요소이다[13]. 사용자 리뷰 분석은 어플리케이션 마켓에 대한 구조분석을 파악할 수 있으며, 이는 곧 기업의 마케팅적 시사점으로도 이어질 수 있다[14].

이와 같이 러닝 활동 후 러닝 동선 확인, 속도 및 거리, 기록 공유 등을 할 수 있는 러닝 어플리케이션으로는 나이키의 'Nike Run Club', 아식스의 'Runkeeper', 언더아머의 'MapMyRun'을 사용해서 달리기, 'Strava: 달리기, 라이딩, 하이킹' 등이 있다.

그중 걷기, 러닝에만 특화된 어플리케이션이고, 2023년 4월 기준 1000만 이상의 다운로드가 되었고, 50만개 이상의 리뷰가 있는 어플리케이션인, Nike Run Club과 Runkeeper를 리뷰 분석하고자 한다. Nike Run Club과 Runkeeper를 이용하는 사람들이 어플리케이션을 사용하는데 있어서, 불편함을 감소시키고, 개선사항 도출 및 보완점을 제안하여 러닝 어플리케이션 사용자의 편의성을 높이기 위한 시사점을 제안하고자 한다.

사용자 리뷰는 텍스트 기반으로 한 비정형 데이터로 이루어져 있기에 설문조사와 같은 전통적인 통계 기법을 적용하는 것은 적절치 않다[15]. 또한, 전통적인 통계 기법은 다양한 의견을 반영할 수 없다는 한계점을 지니고 있고, 연구자를 의식하거나 설문조사 환경에 따라 이용자들의 응답이 편향적으로 드러나는 문제를 초래한다[13]. 더불어 사용자 리뷰는 암시적이거나 명시적인 부분은 텍스트 이면의 감정 예측과 주제를 분류하는 특징을 지니기에 텍스트 데이터를 분석하는 기술이 필요하다[15]. 그것

이 텍스트마이닝이다.

텍스트마이닝은 많은 양의 텍스트로 이루어진 비정형 데이터로부터 새로운 정보 등을 찾고 추출하기 위한 기술이다[16]. 이러한 기술을 통해 많은 양의 어플리케이션 리뷰와 같은 텍스트 정보에서 문서의 유사도를 파악하거나, 핵심 키워드를 추출하거나, 의미있는 정보를 파악해내고, 트렌드를 알아내는 등 이용자의 목적에 맞게 활용할 수 있다[17]. 텍스트마이닝을 활용한 스포츠 관련 선행연구로는 운동선수의 일탈에 대한 것으로, 유형별 뉴스 기사의 주요어 분석을 통하여 경기장 내, 외 운동선수 일탈의 행동 분석을 제시한 [18]의 연구가 있다. [19]는 태권도 겨루기 지도자 코칭 전문성 및 개선점 탐색을 통해 코칭 전문성을 높이기 위한 분석을 진행하였고, [20]은 무용학을 지적으로 구조 분석하여 무용학 지식 구조의 공통적 핵심 정보가 무용, 연구, 춤, 한국, 교육 등의 영향을 받는다는 결론을 도출해내었다. 이와 같이 텍스트마이닝 기반의 빅데이터 분석은 스포츠 기사, 메가스포츠이벤트, 무용 등에서 활용되어 진행된 연구들은 있으나 스포츠에 관련된 소비자들의 사용 후기에 대한 스포츠 어플리케이션 리뷰 분석 연구는 미비하다.

본 연구는 스포츠 어플리케이션인 Nike Run Club과 Runkeeper의 전체 100만개, 60만개의 사용자 데이터 리뷰 중 한국에서 사용자 리뷰를 남긴 10,436개, 1,696개 선정하여 연구하였다. 이를 통해 스포츠 어플리케이션의 개선사항 도출 및 보완점을 제시하여, 스포츠 어플리케이션 프로그램 개발 및 운영전략을 위한 실질적인 기초자료를 제공하는데 그 목적이 있다. 한편,

본 연구의 결과는 스포츠의 여가화와 발맞춰 유도 커뮤니티 참여자의 행동을 예측하고 이에 따라 실무적 시사점을 도출하는데 의의가 있을 것으로 파악 된다. 이에 연구문제는 다음과 같다.

첫째, 러닝 어플리케이션 사용자의 리뷰에서 어떤 단어들 이 자주 사용되었는가?

둘째, 러닝 어플리케이션 사용자의 리뷰에서의 단어들은 어떤 토픽들로 이루어져 있는가?

2. 연구방법

2.1 데이터 수집

본 연구는 사용자 리뷰 어플리케이션 개선을 위한 포인트 도출이므로 대상인 어플리케이션을 직접 사용하고

평가로 남긴 안드로이드 구글 플레이스토어에서 2012년 6월부터 2023년 4월(Nike Run Club), 2010년 9월부터 2023년 4월(Runkeeper)까지 한국에서 사용한 사용자 리뷰 데이터 각각 10,436개, 1,696개를 수집하였다. 또한, 웹 크롤링 라이브러리 Selenium을 통해 date(작성일시), rating(별점), contents(내용) 등의 리뷰 기본 항목을 수집하였다.

2.2 데이터 전처리

텍스트마이닝을 활용하여 수집한 텍스트 자료는 데이터 전처리 과정을 통해 분석에 적합한 형태로 만드는 것이 선행되어야 한다[21]. 이에 따라 Okt형태소 분석기를 사용하여 숫자, 이모티콘, 한자, 영어 등 한글 이외에 문자를 제거하고 형태소를 분리하였다. python 3의 rankNL 불용어 사전을 만들어 연구내용에 부적합한 키워드는 추가적인 불용어 처리를 통해 데이터 전처리를 실시하였다.

2.3 TF(키워드 빈도), TF-IDF(키워드 빈도-문서 역빈도)

본 연구는 Nike Run Club, Runkeeper의 각 어플리케이션 리뷰의 키워드 분석을 위해 python 3의 gensim 라이브러리를 사용하여 키워드 빈도(Term Frequency, TF), 키워드빈도-문서 역빈도(Term Frequency-Inverse Document Frequency, TF-IDF) 두 가지 기준을 적용하여 높은 순으로 나열하였다. TF-IDF는 TF값의 치명적 오류를 잡기 위해 IDF를 활용하여 문장의 연관성이 낮음에도 자주 출현하는 경우의 제한을 주기 위한 기법으로 활용한다[22][25].

본 연구에서 각 어플리케이션별로 추출된 상위50위 주요 키워드를 키워드 빈도(TF), 키워드 빈도-문서 역빈도(TF-IDF)의 높은 순서로 나열하였다. 또한 키워드 빈도(TF), 키워드 빈도-문서 역빈도(TF-IDF)의 순위 차이를 확인하기 위해 비교하였다.

2.3.1 LDA 토픽모델링

LDA 토픽모델링은 Nike Run Club, Runkeeper의 어플리케이션 리뷰를 속성별로 구분하여 토픽을 나누는 작업이다. LDA 토픽모델링 분석을 하기 위해서는 단어들을 벡터화하는 작업이 필요하다. python3의 gensim 라이브러리를 이용하여 완성된 데이터셋을 학습시켜 주었다. 데이터를 명사 리스트로 생성하여 문서별 말뭉치를 만들었

다. 만들어진 명사 리스트를 가지고 키워드 분석을 시행하여 의미 없는 단어들이나 비중이 매우 큰 단어들 제거해주는 과정을 진행했다. 이렇게 언어 모델을 만들 때 토픽모델링을 시각화 후 다시 전처리 과정을 반복하여 최적의 LDA 토픽모델을 만들어주었다.

Nike Run Club, Runkeeper 어플리케이션 리뷰의 토픽모델링을 진행하기 위해서는 토픽 수(k)를 정해야 하며, 이때 최적의 파라미터 값인 k값을 정하기 위해서는 Coherence(일관성)과 Perplexity(혼잡도) 값에 근거하여 토픽의 수를 결정한다. k값을 결정하기 위해 파라미터 값의 범위를 1개 ~ 8개까지 설정하여 모델을 실행하였다. Coherence 값의 경우 토픽이 얼마나 일관성이 있는지 파악해준다. 이때 Coherence 값이 클수록 의미론적으로 일관성이 높다고 해석한다. Perplexity의 경우 토픽의 개수마다 학습을 시켜 가장 작은 값을 보이는 구간을 찾아 최적화된 토픽의 개수를 파악해준다. 이는 결과를 얼마나 정확하게 예측하는지를 판단해주며 Coherence 값과 반대로 값이 작을수록 정확도가 높다고 해석된다. 하지만 Perplexity 값이 작다고 해서 무조건 결과가 해석에 용의하다는 것을 의미하지는 않기 때문에 해석에 주의해야 한다.

3. 연구결과

3.1 TF(키워드 빈도), TF-IDF(키워드 빈도-문서 역빈도)

본 연구에서 Nike Run Club, Runkeeper의 사용자 리뷰에서 어떠한 단어들 자주 등장하였는지 파악하기 위해 TF, TF-IDF를 실시하였으며 결과는 Table 1, Table 2와 같다.

Table 1. TF, TF-IDF ranking comparison - Nike run club

Keyword	TF Rank	TF-IDF Rank
Record	1	1
App	2	3
Exercise	3	2
Running	4	6
Applications	5	4
Login	6	7
Error	7	8
Use	8	5
Distance	9	9
Nike	10	10
Save	11	12
Update	12	16
Function	13	15
Loading	14	18
Integration	15	11

Table 2. TF, TF-IDF ranking comparison - Runkeeper

Keyword	TF Rank	TF-IDF Rank
Exercise	1	1
Use	2	2
App	3	3
Applications	4	4
Record	5	5
Distance	6	7
Function	7	8
Running	8	6
Paid	9	12
Free	10	13
Setup	11	15
Walk	12	10
Help	13	9
Speed	14	17
Save	15	18

3.2 LDA 토픽모델링

본 연구에서 Nike Run Club, Runkeeper의 사용자 리뷰에서 어떠한 토픽들로 이루어져있는지 확인하기 위해 LDA 토픽모델링을 실시하였으며 결과는 Table 3, Table 4와 같다. Coherence(일관성) 점수와 Perplexity(혼잡성) 점수를 기반으로 Nike Run Club, Runkeeper Nike Run Club은 '기본 항목', '추가 기능', '오류 사항', '위치기반 데이터'의 토픽이 도출되었으며, Runkeeper는 '오류 사항', '음성 기능', '러닝 데이터', '사용 혜택', '사용 동기'의 토픽이 도출되었다.

4. 논의

본 연구는 텍스트마이닝을 활용하여 러닝 어플리케이션 사용자 리뷰를 분석하였다. 이에 Nike Run Club과 Runkeeper의 사용자 리뷰로 연구를 진행했으며 TF, TF-IDF, LDA 토픽모델링을 통하여 나타난 결과를 다음과 같이 논의하고자 한다.

첫째, Nike Run Club과 Runkeeper의 TF, TF-IDF 순위 차이가 나타났다. Nike Run Club과 Runkeeper 어플리케이션에서 '기록', '앱', '운동'의 단어가 공통적으로 상위 키워드로 도출되었다. 이는 사용자들이 러닝 운동할 때 기록을 확인하기 위해 사용한다고 사료된다.

Nike Run Club의 경우, 상위15 키워드 중에서 '러닝', '로그인', '오류'의 키워드가 TF순위보다 TF-IDF순위가 떨어졌다는 것이다. 이는 많고 다양한 리뷰에 공통적으로 나타나는 경우 점수가 하락하는 TF-IDF의 특성을 고려

Table 3. Topic modeling results (Topic names and reviews) – Nike run club

Topic name		Top 20 keywords, reviews									
Ⓐ	A	Exercise Run	Application Measure	Record Speed	App Friend	Running Help	Distance Voice	Function Recommend	Running Version	Save Fun	Exit Meter
	Ⓐ	① This is the best exercise app! ② I love that it keeps track of my distance. ③ I like that it saves all my different records so I can see how much I've grown over the years. It helps me to keep running and stay healthy. ④ But I keep seeing the distance in miles, how do I change it to kilometers?									
Ⓑ	A	Use Add	Synchronization Star	Nike Use	Integrations Check	Motivation Free	Music Information	iphone Operation	Mile Write	Run Galaxy	Execute Song
	Ⓑ	① I like that it's free ② Can you add a melon to the song? ③ Integration works well ④ It's motivating and good. I like that it automatically runs when I pause and when I use the workout again.									
Ⓒ	A	Error Minutes	Login Share	Update Bug	Android Facebook	Sign up Problem	Edit Connections	Setup Song	Note Kilo	Screen Input	Install Methods
	Ⓒ	① As long as you don't get the occasional error.... Install 3times again ② I keep getting errors when I look at my activity history, please fix it ③ This is the last time I signed up and pressed confirmation, but I didn't even get a message that the membership was entered, and when I went to the login page to see if it was me again, it said that I couldn't sign up again. ④ Since a few days ago, Facebook auto-sharing hasn't worked, so I always have to log in to share. It's pretty inconvenient... Please fix it!									
Ⓓ	A	Korean Service	Display Activity	Map Phenomena	Support Stop	Kilometer Report	Korea Force	People Manage	Battery Interface	Error Mode	Location Tables
	Ⓓ	① I like that it supports Korean. ② It's very easy to use, the interface is simple, and the distance and location judgment is accurate!!! ③ I use it when I walk in the morning. It is useful to know how many kilometers I walked with the location tracking function- I will use it well~^^ ④ I use it well with other apps and connect them. It's easy to use because it displays the location well.									

A: Keywords, B: Reviews

Ⓐ: Default items, Ⓑ: Adding Features, Ⓒ: Errors, Ⓓ: Location-based data

할 때, 많은 Nike Run Club 어플리케이션 이용자가 어플리케이션을 사용하기 위해 로그인시 부정적 경험에서 나타난 결과로 보인다. 본 연구의 결과는 [23]의 아프리카 TV 어플리케이션 사용자 경험의 키워드로 회원가입 및 로그인에 다양한 문제가 많이 제기되어 사용자의 경험 만족도를 낮춘다는 결과와 [24]의 빅데이터 분석을 활용한 보건소 모바일 헬스케어 어플리케이션 사용자들의 경험의 키워드로 로그인 및 실행 등 많은 오류로 인해 서비스를 사용하지 못하는 것에 대한 부정적인 연구결과와 유사하게 나타났다. 따라서 로그인 오류 문제가 개선될 수 있도록 주기적인 어플리케이션 업데이트를 통해 사용자들이 서비스를 이용하는데 불편함이 없도록 개선해야 한다고 사료된다.

Runkeeper 어플리케이션의 경우, 상위10 키워드 중에서 '유료', '무료', '기능' 등의 키워드가 TF순위보다 TF-IDF 순위가 낮다는 것이다. 이는 다양한 리뷰에 공통적으로 출현하는 경우 점수가 떨어지는 TF-IDF의 특성을 고려할 때, 다양한 Runkeeper 어플리케이션 사용자

가 Runkeeper의 무료 및 유료 버전의 각 기능을 이용하고 느낀 경험을 리뷰에 작성한 결과로 보인다. 본 연구의 '유료결제하고 트레이닝 플랜을 짰는데 '트레이닝' 탭 들어가면 아무것도 안떠요 어떻게하면 제 플랜을 볼 수 있는거죠??', '유료의 가치가 없음' 등의 리뷰는 [25]의 스포츠 O2O 서비스 앱인 '에브리핏', '니짐내짐', '아이코젠'의 사용자 리뷰에서 나타난 낮은 품질의 유료 서비스에 대한 불만족 요인으로 나타난 부정적으로 나타난 연구결과와 [15]의 웹툰 어플리케이션 사용자 리뷰 키워드 유료결제 서비스에 대한 부정적인 연구결과와 유사하게 나타났다. 이를 통해 유료 버전을 사용하는 이용자들이 어플리케이션을 편리하게 사용할 수 있도록 유료버전에 맞는 실시간 모니터링 프로그램 및 맞춤형 교육을 개발 및 보완 해야한다고 사료된다.

둘째, Nike Run Club과 Runkeeper의 LDA 토픽모델링의 토픽 수로 각 4개, 5개의 토픽으로 나타났다. 두 개의 어플리케이션 사용자 리뷰에서 공통적으로 '오류 사항' 토픽이 나타났다. [13]의 텍스트마이닝을 활용한 자산

Table 4. Topic modeling results (Topic names and reviews) – Runkeeper

		Top 20 keywords, reviews									
A	Record	Save	Support	Bicycle	Middle	Error	Jogging	Problem	Instruction	Exit	
	Error	Audio	Korean	Modify	Update	Execute	Location	Facebook	Alarm	Delete	
B	① It's good, but it always crashes when I go over 35 kilograms. I can't recover in the middle and it seems to be a chronic disease.										
	② I'm annoyed that it crashes in the middle of use and I can't keep track										
	③ Sometimes the recordings are incorrect and the route is missing.										
	④ There are a lot of errors and it often crashes without locating..										
A	Use	Function	Distance	Free	Speed	Map	Pace	Stop	Runkeeper	Path	
	Android	Manage	Screen	Music	Purpose	Now	Coach	Before	Voice	Performance	
B	① I'd like to see more different voices for coaches.										
	② I feel like I'm running with a coach. Even though I'm only using it for free...										
	③ It's a free app, but it gives me a lot of information^^ I like that the route is also displayed on the map~.										
	④ It's easy to use and the best free app										
A	App	Exercise	Record	Run	Setup	Goal	Run	Recommend	Friend	Momentum	
	Motivate	Activity	Share	Write	Connect	Marathon	People	Check	Solve	Data	
B	① It's great for keeping record when running										
	② Especially with the goal setting and activity setting, even beginners are well motivated to exercise.										
	③ It's great for recording workouts and viewing past data.										
	④ It makes the exercise time not boring and motivation through goal setting is too good.										
A	Start	Walk	Everyday	Running	Help	Paid	Information	Automatic	English	Recommend	
	Running	Version	Screen	Name	Diet	Walk	Program	Money	Development	Training	
B	① It helps me a lot with my diet.										
	② I'm doing it because walking is the best way to lose weight, but I feel more motivated and stronger because I'm using the earplugs.										
	③ Very helpful in training										
	④ It is intuitive and convenient to use. If you create an ID, you can also manage records. I think it's the best among running management programs^^!!!.										
A	Exercise	App	Calorie	Integration	Install	Distance	Measurement	Confirm	Think	Use	
	Fun	Key	Download	Voice	Service	Challenge	Calculate	Strongest	iphone	Health	
B	① I'm using it well because it calculates the distance and tells me the calories.										
	② I always work out with it. It's a basic app that makes working out fun, so I highly recommend it!										
	③ The GPS is very accurate, and the calorie counting is accurate for each type of food, so I use it often.										
	④ It's fun to run while it automatically tells me my pace and other information while I'm running.										

A: Keywords, B: Reviews

ⓐ: Errors, ⓑ: Voice features, ⓒ: Running Data, ⓓ: Using Benefits, ⓔ: Motivation for use

관리 어플리케이션 분석에서 도출된 어플리케이션 연동 오류나 업데이트 시 발생하는 오류로 나타나는 결과와 [26]의 카카오톡 사용자 리뷰의 담론 분석 결과에서 계정, 메시지 전송오류, 로그인, 네트워크 오류로 인해 카카오톡 채팅 전송과 소통상의 불편의 결과는 본 연구결과와 유사하게 나타났다. Nike Run Club의 ‘활동내역 불러하면 계속 오류남니다 수정바랍니다’, ‘몇일 전 부터 페이스 북 자동공유가 안되어서 항상 로그인을 해야 공유가 됩니다. 꽤나 불편하네요.. 고쳐주세요!’와 같은 리뷰와 Runkeeper의 ‘정말 여러 많고 위치 못 잡고 수시로 다운되고ㅠ’, ‘기록이 가끔 오류가 생겨요 경로가 없어졌어요’의 리뷰를 통해 여러 및 기록 오류를 Nike Run Club, Runkeeper 이용자들이 어플리케이션을 사용하고 경험한 것을 확인할 수 있다. 따라서 기록 여러 및 오류를 해결하기 위해 어플리케이션을 점검을 통해 꾸준한 관리가 필요하다고 사료된다.

Nike Run Club의 ‘기본 항목’ 토픽의 경우, [27]의 벅크셀러드 어플리케이션의 사용 결정요인에 관한 연구의 ‘자산 관리’ 토픽과 상이하게 나타났다. 벅크셀러드의 경우 자산 관리 어플리케이션으로 잔고, 금융정보 등 이용자의 자산정보를 확인할 수 있어 자산 관리와 관련된 키워드가 되었다는 것으로 판단된다. 하지만 Nike Run Club의 경우, ‘여러가지 기록이 다 저장되니까 내가 그동안 얼마나 성장했는지 볼 수 있어서 좋아요. 덕분에 꾸준히 러닝으로 건강 유지할 수 있겠어요’, ‘러닝에 필요한 모든 기능이 포함되어있음’ 등의 리뷰를 통해 러닝 후 거리, 기록, 속도 등을 어플리케이션의 기본 항목을 바탕으로 이용자가 어느정도 성장하였는지 파악할 수 있다는 역할을 한다고 사료된다. ‘위치 기반 데이터’의 경우, [28]의 전동키보드 공유 서비스의 키워드와 본 연구 결과의 키워드와 유사하게 나타났다. Nike Run Club의 ‘아침에 걷기 할 때 씩니다. 위치추적기능으로 몇킬로미터를 걸었는지

알 수 있어 유용하네요~ 잘쓰겠습니다~^^', '아주 사용하기 편리하고 인터페이스가 심플하고 거리, 위치 판단이 정확해요!!!' 등의 리뷰를 바탕으로 킬로미터 및 지도 표시의 위치 서비스 기반이 러닝 기록들을 편리하게 확인이 가능하다는 것을 유추할 수 있다.

Runkeeper의 '음성 기능' 토픽의 경우, [29]의 모바일 피트니스 어플리케이션 주요 요인 분석 연구에서 목소리, 음악 같은 청각적 기능은 운동하는데 활력을 제공하는 역할의 결과는 본 연구결과와 유사하게 나타났다. Runkeeper 또한 러닝 시 코치음성 및 음악이 피드백 및 응원 메시지의 역할을 하는 것으로 사료된다. '사용 동기' 토픽 키워드의 경우, [30]의 스마트 관광 활성화를 위한 트립어드바이저 어플리케이션 리뷰 연구의 '사용 용도' 토픽의 키워드와 상이하게 나타났다. 트립어드바이저의 '사용 용도' 토픽으로 'use, easy, travel, love, find'가 도출되었다. 이는 트립어드바이저 어플리케이션을 이용하여 호텔을 찾기 위한 용도를 사용한다고 유추할 수 있지만 본 연구의 Runkeeper는 '운동, 어플, 칼로리, 거리, 측정, 건강, 이용, 재미, 도전' 등의 키워드를 통해 기록에 도전하고 해주고 또한 거리, 칼로리 등의 기록들을 측정 및 계산을 확인하기 위해 사용 동기의 역할을 한다고 유추할 수 있다.

5. 결론 및 제언

5.1 결론

본 연구는 텍스트마이닝을 활용하여 러닝 어플리케이션 Nike Run Club과 Runkeeper의 사용자의 리뷰를 분석하여 어플리케이션 사용에 불편함을 감소시키고, 개선 사항 및 보안점을 도출하여 어플리케이션 경쟁력 향상에 기여하고 사용자의 편의성을 높이는 것의 목적이 있다. 이에 Nike Run Club의 10,436개의 사용자 리뷰와 Runkeeper 1,696개의 사용자 리뷰를 수집하였으며, okt 형태소 분석기를 사용하여 숫자, 영어, 한문 등을 제거하고 명사만 남긴 채 형태소 분리하였으며 rankNL 사전을 사용하여 불용어를 제거하였다. 이후 TF, TF-IDF와 LDA 토픽모델링을 실시한 연구결과는 다음과 같다.

첫째, Nike Run Club의 경우 TF(키워드 빈도) 결과는 기록, 앱, 운동, 러닝, 어플, 로그인, 오류 순으로 나타났으며 TF-IDF(키워드 빈도-문서 역빈도) 결과는 기록, 운동, 앱, 어플, 사용, 러닝, 로그인, 오류 순으로 나타나 TF,

TF-IDF 순서에는 차이가 나타났다. 로그인 오류문제가 개선될 수 있도록 정기적인 어플리케이션 업데이트를 통해 사용자들이 서비스를 사용하는데 불편함이 없도록 만들어야 할 것이다. Runkeeper의 경우 상위 TF(키워드 빈도) 결과는 운동, 사용, 앱, 어플, 기록, 거리, 기능, 달리기, 유료, 무료 순으로 나타났으며 TF-IDF(키워드 빈도-문서 역빈도) 결과는 운동, 사용, 앱, 어플, 기록, 달리기, 거리, 기능의 순서로 나타나 TF, TF-IDF 순서에는 차이가 나타났다. 유료버전에 맞춤형 교육 및 실시간 모니터링 프로그램 기능을 보완 및 개발하여 유료버전을 이용하는 사용자들이 편리하게 어플리케이션을 이용할 수 있도록 유료버전이 보완되어야 한다고 사료된다.

둘째, Nike Run Club의 경우 Coherence Score(일관성 점수), Perplexity Score(혼잡도 점수)에 의해 '기본 항목', '추가 기능', '오류 사항', '위치 기반 데이터'의 4개의 토픽으로 나타났다. 이는 기본 항목이 러닝 운동 이후 속도 및 달린 거리, 기록 등을 통해 사용자가 얼마나 성장하였는지 확인할 수 있는 역할을 한다고 유추할 수 있고 위치 서비스를 기반으로 지도 및 킬로미터 표시가 되어 사용자들이 러닝 활동 후 데이터를 확인하여 편리하고 유용하게 확인이 가능하다고 사료된다. Runkeeper의 경우 Coherence Score(일관성 점수), Perplexity Score(혼잡도 점수)에 의해 '오류 사항', '음성 기능', '러닝 데이터', '사용 혜택', '사용 동기'의 총 5개의 토픽으로 나타났다. 이는 음악 및 코치음성으로 제공되는 응원 메시지나 피드백이 중요한 역할을 하는 것을 유추할 수 있고 어플리케이션에서 제공하는 거리, 칼로리 등의 기록들을 계산 및 측정하여 운동을 재미있게 하고 또한 기록에 도전할 수 있게 사용 동기의 역할을 해준다고 사료된다.

5.2 제언

본 연구의 결과와 논의를 바탕으로 향후 유사한 연구의 방향과 보완점을 제언하면 다음과 같다.

첫째, 본 연구는 Nike Run Club과 Runkeeper 러닝 어플리케이션의 사용자들을 대상으로 한정하여 진행했기 때문에 타 스포츠 종목 어플리케이션 사용자 리뷰들을 일반화시켜 해석하는데 무리가 있을 것으로 판단된다. 이에 각 종목별 스포츠 종목에 맞는 연구가 되어야 할 것이다.

둘째, 본 연구는 Nike Run Club과 Runkeeper 러닝

어플리케이션의 사용자들의 감성 분석까지는 진행되지 않아 사용자들의 긍정, 부정 감성을 파악하지 못하였다. 추후 연구에서는 감성 분석을 진행된다면 보다 세밀한 분석결과를 도출할 수 있을 것이다.

REFERENCES

- [1] Oracle. (2023). What is big data?. <https://www.oracle.com/kr/big-data/what-is-big-data/>
- [2] Y. K. Jang. (2022). South Korea's smartphone ownership rate is 93%, with a surge in older age groups. GAMEVU. <https://www.gamevu.co.kr/news/articleView.html?idxno=23948>
- [3] Y. S. Baek. (2020). Coronavirus has increased internet use in the 50s and 60s. Digital Today. <http://www.digitaltoday.co.kr/news/articleView.html?idxno=265586>
- [4] S. I. Baek, S. H. Bae, Y & Y. Song. (2014). Exploring Moderating Effects of Customer's Previous Knowledge and Involvement on Online Word-of-Mouth Adoption in the Application Markets. *Entrue Journal of Information Technology*, 13(3), 21-34.
- [5] R. S. Lee & M. H. Cho. (2016). The Roles of Information Value, Information Sense, and Prior Knowledge in Relation to the Type of Restaurant Smart Phone Application Contents. *Journal of Tourism Sciences*, 40(7), 31-53. DOI : 10.17086/JTS.2016.40.7.31.53
- [6] B. S. Lim. (2022). Global Consumers Spent \$203 Trillion on Mobile Apps & Spent 3.8 Trillion Hours in 2021...Mobile Market is Stronger Than Ever. SmartPCLove. <https://www.ilovepc.co.kr/news/articleView.html?idxno=42149>.
- [7] D. J. Kim (2013). How One's Lifestyle Affects Satisfaction in Sports Application as well as Re-use and Recommendation of the Application. Master Dissertation, Ma. D. Dissertation, Dongshin University.
- [8] P. K. Han, J. S. Park, B. H. Jun & B. G. Kang. (2010). A Study on the Factors of Mobile Applications Adoption. *Journal of Information Technology Services*, 9(3), 65-82.
- [9] Y. W. Han. (2021). A study on the purchase intention based on customer review and service comment when using delivery application. Master Dissertation, Ma. D. Dissertation, Hongik University.
- [10] H. J. Jang. (2021). A study on the improvement of user experience on mobile payment services : a text mining approach. Master Dissertation, Ma. D. Dissertation, Yonsei University.
- [11] Y. K. Han. (2022). Customer value proposition methodology using text mining of online customer reviews. Doctoral dissertation, Ph. D. Dissertation, Hanyang University
- [12] N. S. Kim, S. A. Lee, S. H. Jo & J. H. Kim. (2014). Multi-Dimensional Keyword Search and Analysis of Hotel Review Data Using Multi-Dimensional Text Cubes. *Journal of Information Technology and Architecture*, 11(1), 63-73.
- [13] D. H. Park. (2021). An analysis of user needs by user's review, based on text-mining : focusing on Bank-Salad and Mint. Master Dissertation, Ma. D. Dissertation, Yonsei University.
- [14] Erdem, T., & Michael P. Keane. (1996). Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Market, *Marketing Science*, 15(1), 1-20. DOI : <https://doi.org/10.1287/mksc.15.1.1>
- [15] H. R. Shin & J. H. Choi. (2022). Analysis of User Reviews for Webtoon Applications Using Text Mining. *The Journal of the Convergence on Culture Technology (JCCT)*, 8(4), 457-468. DOI : 10.17703/JCCT.2022.8.4.457
- [16] Jusoh, S., & Alfawareh, H. M. (2012). Techniques, applications and challenging issue in text mining. *International Journal of Computer Science Issues (IJCSI)*, 9(6), 431.
- [17] J. K. Yoon. (2021). A Study on the Patent Analysis Model for Technology Development Trend Using Text Mining. Master Dissertation, Ma. D. Dissertation, Soongsil University.
- [18] E. H. Cho. (2022). Analysis on Key Words in News Articles by Deviance Types of Athletes Using Text Mining. Doctoral dissertation, Ph. D. Dissertation, Dankook University.

- [19] J. C. Ji & M. G. Jun. (2023). Taekwondo sparring instructor coaching expertise and improvement points using text mining. *Sport Science*, 41(1), 33-40.
DOI : <http://dx.doi.org/10.46394/ISS.41.1.4>
- [20] J. M. Lee.. (2017). An Analysis of Semantic Relations in Knowledge Information in Dance Research Data in Korea from 1958 to 2016, *Korea National Resarch Center for Arts*, (16), 215-237.
- [21] Lucas, C., Nielsen, R, A., Roberts, M, E., Stewart, B, M., Storer, A., & Tingley, D. (2015). Computer-assisted text analysis for comparative politics. *Political Analysis*, 23(2), 254-277.
- [22] S. J. Chae. (2023). Study on the Content Analysis and Development Strategy of Online Dance Video Using Data Mining. Doctoral dissertation, Ph. D. Dissertation, Sangmyung University.
- [23] Jin, W. (2022). Comparative analysis of mobile live streaming application experience based on text mining approach : focusing on TwitchTV and AfreecaTV. Master Dissertation, Ma. D. Dissertation, Yonsei University.
- [24] T. J. Kim. (2022). Analysis of User Perception regarding the Mobile Healthcare Application of a Public Health Center using Big Data. *Korean Journal of Sport Science*, 33(4), 648-658.
- [25] T. J. Kim & M. H. Kim. (2022). An analysis of user reviews on sports O2O app service using big data and text network analysis, *Korean journal of physical education*, 61(4), 117-131.
DOI : 10.23949/kjpe.2022.7.61.4.10
- [26] K. D. Son. (2022). Discourse analysis of user reviews of KakaoTalk : through usability test. Master Dissertation, Ma. D. Dissertation, Hongik University.
- [27] D. E. Bae. (2021). An analysis on determinants of using personal financial management fin-tech application : based on text mining for banksalad users. Master Dissertation, Ma. D. Dissertation, Yonsei University.
- [28] K. A. Seo. (2022). Using user review text mining analysis of success factors of electric kickboard sharing service. Doctoral dissertation, Ph. D. Dissertation, Hoseo University.
- [29] S. H. Lee, J. S. Kim, S. H. Yoon & H. W. Kim. (2020). An Analysis on Key Factors of Mobile Fitness Application by Using Text Mining Techniques : User Experience Perspective. *Journal of Information Technology Services*, 19(3), 117-137.
DOI : 10.9716/KITS.2020.19.3.117
- [30] Y. N. Lee, M. M. C. Han, S. Y. Yu, M. Q. Siow & Y. S. Kim. (2023). Analyzing TripAdvisor application reviews to enable smart tourism : focusing on topic modeling, *Smart media journal*, 12(8), 9-17.
DOI : 10.30693/SMJ.2023.12.8.9

류 기 문(Gimun Ryu)

[정회원]

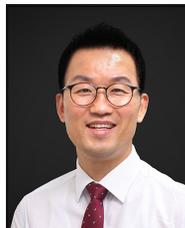


- 2022년 2월 : 한국체육대학교 레저 스포츠학과(학사)
- 2024년 2월 : 한국체육대학교 일반 대학원(석사)

- 관심분야 : 스포츠경영, 스포츠마케팅
- E-Mail : kyleyu10@naver.com

김 일 광(Ilgwang Kim)

[정회원]



- 2023년 12월 : 한국디지털정책학회 웰니스융복합분과 상임이사
- 2023년 12월 : 한국체육대학교 교수

- 관심분야 : 스포츠산업경영, 마케팅
- E-Mail : whyhow2070@gmail.com