

텍스트마이닝 기법과 ARIMA 모형을 활용한 배달의 민족 앱 리뷰 분석

정지훈¹ · 정혜인² · 이준기^{3*}

^{1,2}연세대학교 정보대학원 석사과정

^{3*}연세대학교 정보대학원 교수

An Analysis of Mobile Food Delivery App 'Baemin' by Using Text Mining and ARIMA Model

Ji-Hoon Jung¹ · Hye-In Chung² · Zoon-Ky Lee^{3*}

^{1,2}Master's Course, Graduate School of Information, Yonsei University, Seoul 03722, Korea

^{3*}Professor, Graduate School of Information, Yonsei University, Seoul 03722, Korea

[요약]

모바일 애플리케이션 시장의 빠른 성장과 함께 다운로드 수와 리뷰의 수도 급증하고 있다. 개발자는 사용자 리뷰 분석을 통해 서비스 이슈를 감지하고 사용자의 불만족을 개선하고자 하지만, 매일 다량으로 생성되는 리뷰를 처리하기에는 어려움이 있다. 본 연구에서는 수많은 리뷰 중에서 유의미한 서비스 이슈를 탐지하기 위해 텍스트마이닝 기법과 시계열 이상치 탐지 모형 Autoregressive Integrated Moving Average(ARIMA)을 이용하여 최근 빠르게 성장하고 있는 배달 어플리케이션 '배달의 민족' 리뷰를 분석하고자 한다. 이를 위하여 '배달의 민족' 앱 리뷰 71,435건을 크롤링하여 수집한 후, 토픽모델링 기법을 적용하여 5개의 토픽을 도출하였다. 또한, 시계열 기반 이상치 탐지 ARIMA 모델을 활용하여 리뷰수가 급증한 12개의 이상치 일자를 탐지하였고, 해당 일자를 토픽 별 감성분석 결과를 바탕으로 이슈 요인을 분석하였다.

[Abstract]

As the mobile application market is growing rapidly, the number of application downloads and reviews are increasing as well. Developers want to detect and improve service issues and user dissatisfaction through application review analysis, but it is difficult to handle large quantities of reviews that are generated every day. In this paper we analyze review of the delivery application "Baemin", which is growing rapidly in recent years. This study detects significant service issues among numerous reviews by using text mining techniques and ARIMA based time series outlier detection model. 71,435 reviews of the "Baemin" application reviews are collected from Google Play Store, and six topics are derived by applying a topic modeling technique. Furthermore we use ARIMA based outlier detection model and 12 dates are detected as outliers. By applying sentiment analysis and topic modeling, we analyze each outlier to identify app issues.

색인어 : 배달 앱, 모바일 앱 리뷰 분석, 토픽 모델링, ARIMA, 감성 분석

Key word : Delivery App, Mobile App Review Analysis, Topic Modeling, ARIMA, Sentiment Analysis

<http://dx.doi.org/10.9728/dcs.2021.22.2.291>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 28 December 2020; **Revised** 20 January 2021

Accepted 20 January 2021

***Corresponding Author; Zoon-Ky Lee**

Tel: +82-2-2123-4528

E-mail: zlee@yonsei.ac.kr

1. 서론

정보통신 기술(ICT)과 스마트폰 보급의 증가로 모바일 애플리케이션 시장은 빠르게 성장하고 있다. 글로벌 시장조사업체 아이리서치(iResearch)에 따르면, 2020년 전 세계 모바일 앱 시장의 규모는 5,819억 달러로 추산되며, 2014년을 기준으로 2023년까지 년 평균 29% 성장률을 기록할 전망이다[1]. 모바일 애플리케이션 시장의 성장과 함께 앱 스토어 상에서 발생하는 사용자 리뷰의 수도 증가하고 있다. 모바일 애플리케이션의 사용자 리뷰는 애플리케이션에 대한 즉각적인 사용자 경험을 재구성한 사용자들의 직접적인 피드백으로[2], 서비스의 품질을 개선하고 누락된 기능을 보완하는 데 유용한 정보를 제공한다[3]. 사용자 리뷰에서 감지된 불편사항은 개발자들이 애플리케이션을 유지 및 보수하고, 애플리케이션 업데이트 일정을 계획하는데 유용한 근거를 제공한다[4]. 개발자들이 이러한 정보를 적시에 인식하고 행동을 취하기 위하여 사용자 리뷰로부터 필요한 정보를 획득하는 과정은 서비스 운영에 매우 중요하다[3].

앱 리뷰의 중요성을 환기하는 사례로 2014년 페이스북 메신저(Facebook Messenger) 사례가 있다. 페이스북 메신저는 2014년 8월 애플 앱스토어에서 전체 리뷰의 약 94%에서 평점 최저점인 1점을 받으며 대량 고객 이탈이 발생하였다. 당시 페이스북 메신저는 사용자의 전화에 있는 사진과 연락처에 접근이 되는 심각한 사생활 정보 침해 이슈 문제를 갖고 있었다[5]. 그러나 이미 한 달 전 앱 스토어에서 동일한 이슈로 1만 2600여 건의 부정적인 리뷰가 접수된 바가 있었다[5]. 해당 이슈가 한 달 전 사용자 리뷰에서 적시에 탐지되었다면 대량 고객 이탈 전 문제는 해결될 수 있었을 것이다. 이는 사용자 리뷰를 바탕으로 서비스 이슈를 탐지하고 원인을 분석하는 것에 대한 중요성을 인식하는 사건이 되었다.

사용자 리뷰를 통한 이슈 대응이 중요함에도 사용자 리뷰는 빠르고 정확하게 대응하기 어려운 특성을 가지고 있다. 특성은 다음과 같다. 첫째는 양이 매우 많다는 것이다. 모바일 앱 리뷰의 수는, 애플리케이션의 인기에 따라 하루에 몇 천 건까지 증가할 수 있으며[6], 페이스북(Facebook)의 경우 구글 플레이 스토어(Google Play Store)에서 매일 10,000 건 이상의 리뷰를 받는다[7]. 매일 다량으로 발생하는 리뷰를 모두 이해하기 위해서는 많은 시간이 필요하다. 둘째는, 많은 리뷰에는 반복 단어, 한 국어가 아닌 단어 등 수많은 노이즈 단어가 포함되어 있다. 모바일 애플리케이션 리뷰 작성자들은 대부분 스마트폰 단말기를 통하여 작성하기 때문에 길어도 짧아 분석에 어려움이 있다. 셋째는 대부분의 애플리케이션 사용자 리뷰는 주제가 분류되어 있지 않다. 해당 리뷰가 어떠한 주제에 해당하는지 분류하기 위해서는 관리자가 일일이 해석하는 과정이 필요하다.

이와 같은 한계점을 극복하기 위해 다양한 연구가 진행되고 있지만, 토픽 모델링을 활용한 앱의 만족 불만족 요인 도출[8-9], 앱 리뷰 중 주제별 가장 유의한 리뷰를 선별하는 것에 대한 연구[10]로, 특정한 이슈에 대한 리뷰가 다량으로 발생하는

상황을 탐지하기 위한 연구는 미미하다. 따라서, 본 연구에서는 토픽 모델링으로 전체 리뷰의 주제를 분류한 후, 리뷰가 급증하는 일자의 리뷰를 따로 추출하여 분류된 주제와 감성분석 결과를 바탕으로 인사이트를 얻고자 한다.

본 연구의 의의는 다음과 같다. 첫째, 다량의 리뷰 중에서 유의미한 리뷰만을 분석하기 위하여 형태소분석 진행 후, 추출한 명사의 수를 근거로 유의미한 리뷰를 선별하였다. 둘째 토픽 모델링을 통해 모든 리뷰의 주제를 비지도 학습으로 사전에 분류하였다. 셋째 이슈가 발생하는 지점의 리뷰만을 분석하기 위하여, 시계열 이상 탐지 ARIMA(Autoregressive Integrated Moving Average)모델을 이용하여 리뷰 수가 급증하는 일자를 추출하였다. 넷째, 리뷰가 급증한 일자의 리뷰 중에서 특정 주제로 분류된 비율이 높은 일자를 선별하여 특정 주제가 집중적으로 발생한 일자의 서비스 이슈를 파악하고자 하였다. 이로써 시계열 데이터를 기반으로 리뷰 수의 이상치를 탐지함으로써 서비스의 위험 요소를 적시에 감지할 뿐 아니라, 해당 이슈가 속한 주제를 토픽모델링을 근거로 자동 분류함으로써 리뷰 분석을 통한 이슈 탐지에 구체적인 방안을 제시한 것에 의의가 있다.

II. 이론적 배경

2-1 국내 배달 애플리케이션 현황

배달 애플리케이션은 식품·음식 분야 O2O(Online to Offline) 서비스 기업이다. O2O 서비스는 모바일 애플리케이션 등을 통하여 음식을 주문하거나, 택시 또는 렌터카 호출, 호텔 숙박 예약, 청소 도우미, 부동산 계약 요청 등을 받아, 오프라인 서비스를 매칭해주는 플랫폼을 의미한다. 국내 O2O 시장은 국내에서 높은 스마트폰 보급률을 바탕으로 성장하고 있으며, 시장 규모는 지난해 2조 6931억 원 대비 23.2% 늘어나 올해 3조 3,179억 원에 이를 전망이다[11]. O2O 시장에서 2019년을 기준으로 매출액이 가장 많은 분야는 ‘식품·음식’으로 전체 매출 중 28.4%를 차지하였다[12]. 식품·음식 분야 O2O 서비스 기업은 딜리버리 히어로(Delivery Hero)에서 운영하는 ‘배달의민족(Baemin)’, ‘요기요(Yogiyo)’, ‘배달통(Baedaltong)’과 후발 주자로 나타난 ‘쿠팡이츠(CoupangEats)’와 ‘위메프오(Wemakepo)’ 등이 있다. 최근 배달앱 시장은 코로나19의 확산으로 ‘언택트 소비’가 증가하여 고속 성장하고 있으며, 올해 8월 월간 사용자 수는 지난해 2019년 8월 대비 약 25% 증가한 것으로 집계되었다[13]. 국내 배달 애플리케이션 시장은 2019년 12월 독일 업체 딜리버리 히어로가 1위 업체인 ‘배달의 민족’을 약 4조 7,500 억 원에 매각함으로써 시장의 98%를 독점하게 되었다. 딜리버리 히어로는 국내 배달 어플리케이션 시장 2위인 ‘요기요’와 3위인 ‘배달통’도 소유하고 있다[14]. 이와 같은 애플리케이션 시장 독점으로 인해 발생하는 배달 애플리케이션 수수료 상승 문제, 소상공인이 받는 피해와 소비자의 선택권이 줄어드는 등 다양한 각도의 문제에 대한 우려가 제기되고 있다.

2-2 국내 배달 애플리케이션 연구 동향

국내 최초의 배달 애플리케이션은 ‘배달통’으로 2010년에 처음 서비스를 개시하였다. 배달 애플리케이션 시장은 2014년부터 시장 규모가 급격히 성장하여 본격적인 연구가 진행되었다. 초기 연구에서는 배달 애플리케이션의 구현과 설계에 대한 연구가 주를 이루었다. 유진욱(2014)은 소도시의 배달 애플리케이션 설계 및 구현에 관한 연구를 진행하였으며[15], 정진명, 남재원(2015)은 배달 애플리케이션 소비자 보호 방안에 대하여 법학에 근거하여 연구하였다[16]. 이후 배달 애플리케이션이 대중화되고 서비스를 이용하는 이용자가 증가하자 사용자 경험과 서비스 품질에 관한 연구가 이루어졌다. 정혜경(2015)은 국내 배달 애플리케이션 3사(‘배달의 민족’, ‘배달통’, ‘요기요’) 애플리케이션 사용성을 확인하고자 심층 인터뷰를 진행하여 휴리스틱 분석을 하였으며[17], 조미나 등(2017)은 배달 앱 품질에 대한 소비자의 태도와 행동의도를 기술수용모델(TAM)을 중심으로 분석하였다[18].

이와 같은 배달 애플리케이션의 사용자 경험 및 품질 개선과 관련된 선행 연구는 주로 설문조사 또는 인터뷰를 통해 이루어졌다. 설문조사와 인터뷰는 연구 목표에 맞는 답변을 직접적으로 확보할 수 있는 장점이 있어 전통적으로 많은 분야에서 사용되는 방법이지만, 이를 설계하고 진행하기 위하여 많은 비용과 시간, 인력이 요구된다. 따라서 본 연구에서는 텍스트 마이닝 기법을 활용하여, 다량으로 발생하는 애플리케이션 리뷰를 빠르게 분석함으로써 비용과 시간을 절감하고자 한다.

2-3 텍스트 마이닝 기법을 활용한 사용자 리뷰 분석

모바일 애플리케이션의 사용자 리뷰는 애플리케이션에 대한 즉각적인 사용자 경험을 재구성한 사용자들의 직접적인 피드백으로[2], 이용자가 애플리케이션 서비스를 사용한 후 느낀 사용자 경험을 바탕으로 자신의 의견, 생각, 경험을 텍스트 및 이미지 또는 동영상 등을 통해 다른 이용자와 공유한다[19]. 대부분의 애플리케이션 리뷰 데이터는 텍스트를 기반으로 한 비정형 데이터로 이루어져 있기 때문에 기존의 정형 데이터를 가정하여 이루어지는 전통적인 통계 분석 방법은 적절하지 않다. 이러한 데이터를 분석하는 기술로 텍스트 마이닝(Text Mining) 기법이 널리 쓰이고 있으며, 텍스트 마이닝 기법으로는 정보 추출, 텍스트 주제 분류, 감성 분석 등이 있다.

최근 애플리케이션 리뷰에 대한 국내 텍스트마이닝 연구는 주로 토픽 분석을 통한 특정 분야의 애플리케이션의 만족도에 영향을 주는 요인을 조사한 연구가 주를 이루고 있다. 홍정림 외 3인(2019)은 토픽 모델링을 활용하여 모바일 증강현실(AR) 애플리케이션 사용자 리뷰를 분석하였으며[9], 김광국 외 3인(2018)은 애플리케이션 리뷰 토픽분석을 통해 모바일 쇼퍼 앱 고객만족도에 관한 연구를 진행하였다[20]. 이후 애플리케이션

리뷰 분석은 모바일 애플리케이션 시장의 규모의 증가로 인해 급증하는 리뷰 분석을 위해, 프로세스를 자동화하는 방향의 연구가 진행되고 있다. 양유정 외 4인(2019)는 단어 임베딩 및 벡터 유사도를 기반으로 게임 애플리케이션의 리뷰를 자동으로 분류하는 시스템 개발에 관한 연구를 진행하였다[21].

애플리케이션 리뷰 처리의 자동화에 대한 다양한 연구가 진행이 되고 있지만, 주제 분류 단계에만 그치고 있다. 서비스를 운영하는 실무적인 측면에서, 리뷰가 다량으로 발생 시 이를 적시에 탐지하는 것에 대한 연구는 미미한 실정이다. 본 연구에서는 ‘배달의 민족’ 애플리케이션의 리뷰의 개수를 시계열로 분석하여 과도한 리뷰가 발생했을 때 이를 탐지하고 해당 리뷰에 대한 토픽을 자동으로 할당하여 신속한 대응을 하고자 한다. 이를 통해 향후 어플리케이션 리뷰 분석을 통한 서비스 이슈 탐지 프로세스를 정립하는데 유의미한 시사점을 제안하고자 한다.

III. 연구 방법

<그림1>는 본 연구의 연구 절차를 나타내며 크게 데이터 수집, 데이터 전처리, 데이터 분석 단계로 진행하였다. 구체적으로 (1) 데이터 수집에서는 배달의 민족 리뷰 데이터 수집, (2) 데이터 전처리에서는 불용어 제거 및 형태소 분석 등을 통한 전처리를 진행 하였으며, (3) 데이터 분석에서는 LDA 토픽 모델링 후 시계열 이상치 탐지, 감성분석을 진행하였다.

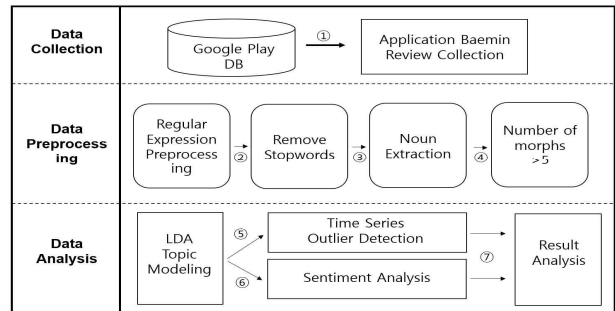


그림 1. 모델링 프로세스

Fig. 1. Modeling Process

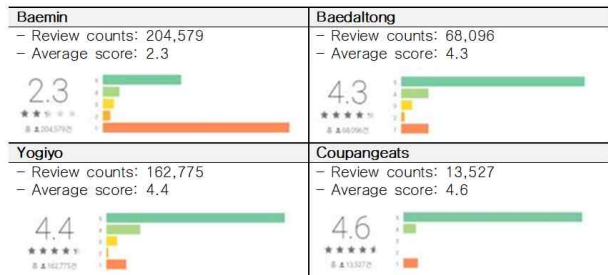


그림 2. 국내 배달 어플리케이션 구글 플레이 사용자 리뷰 현황

Fig 2. Domestic Delivery Application Google Play Store User Reviews Status

3-1 배달 어플리케이션 앱 선정

국내 배달 어플리케이션 중 시장점유율 1~4위 어플리케이션을 대상으로 구글 플레이 스토어에서 확인한 리뷰 평점은 <그림2>와 같다. 시장 점유율 1위인 배달의 민족의 경우, 구글 플레이 스토어에서 발생한 전체 리뷰수가 204,579 건으로 2, 3, 4 위 평균(81,466 건)에 비하여 약 2.5배 많은 리뷰 수를 기록하였다. 하지만, 평점의 경우 2.3점을 기록하며 다른 어플리케이션에 비하여 2점 정도 낮았다. 또한, 다른 어플의 경우 일반적인 온라인 리뷰 분포인 ‘J자 분포(J-shaped Distribution)’를 따르며 5점이 가장 많고 1점이 그 다음으로 많으며, 2,3,4점은 가장 낮은 빈도를 보이고 있다[22]. 하지만, 배달의 민족의 경우 1점이 매우 높고, 5점은 1점의 절반엔 미치지 못하는 분포를 띄고 있다. 배달의 민족이 리뷰 수가 다른 앱에 비하여 매우 많다는 점과 평점이 1점에 매우 집중되어 있다는 점을 고려하여 배달의 민족을 연구 대상으로 설정하였다.

3-2 데이터 수집 및 전처리

데이터 수집 단계에서는 어플리케이션 다운로드 수가 가장 많은 어플리케이션 플랫폼 구글 플레이 스토어에서 배달의 민족 사용자 리뷰 데이터를 크롤링하였다. 이를 위하여 오픈소스 프로그램 Python 3.7과 오픈 패키지 Selenium을 사용하였다. 수집한 데이터는 2016년 6월부터 2020년 10월 기간에 작성된 리뷰로 총 71,435개의 리뷰를 확보하였다. 수집한 항목으로는 작성 날짜, 리뷰 내용, 리뷰 평가가 있다.

데이터 전처리 단계에서는 정규표현식을 사용하여 기타 문장 부호 및 기호를 제외하고 KoNLPy 패키지에서 제공하는 MeCab 형태소 분석기를 이용하여 부사, 관형사, 접속사, 감탄사, 조사, 형용사 등을 제외한 명사만을 추출하였다. 추출한 명사 중 빈도가 높지만 분석에 유의미하지 않은 ‘배민’, ‘민족’, ‘어플’ 등의 단어들은 불용어 처리하였다. 또한 토픽을 분석하기에 너무 짧은 문장을 제외하기 위하여 추출된 명사 수가 5개 미만인 리뷰를 제외하여 15,511 건을 선정하였다. 해당 리뷰에 포함된 명사는 총 145,328건이며, 리뷰는 1건 당 평균 9.37 건의 명사를 포함한다.

3-3 LDA 토픽모델링

LDA(Latent Dirichlet Allocation) 토픽모델링은 의미 있는 주요 토픽들을 찾기 위한 확률 모델적 기법이다. LDA 알고리즘은 문서 내 단어를 바탕으로 변수를 추론하고, 단어들에 대한 잠재 분포를 통해 k개의 토픽을 생성한다[23]. 일반적으로 k개의 토픽의 수를 정하는 기준은 일반적으로 토픽의 혼잡도(Perplexity)와 일관성(Coherence)을 사용한다[24]. 판단 혼잡도는 특정 확률 모델이 결과를 얼마나 잘 예측하는지 알아보기 위한 척도이며, 일관성은 토픽 내 단어의 유사도를 계산하여 해당 주제와 맞는 단어들로 구성되어 있는지 파악하는 척도이다[24]. 본 연구에서는 Python TEANAPS 패키지를 활용하여 일관

성과 응집도 점수를 분석하였다. 토픽의 개수를 2개에서 10개까지 비교 분석해본 결과 토픽의 개수가 5개에서 6개로 변할 때 일관성 점수는 0.297 에서 0.324로 증가하였고, 혼잡도가 -6.6429에서 -6.65로 감소하였다. 따라서 본 연구에서 토픽의 개수가 6개가 타당하다고 판단하였다. 각 토픽의 주제는 추출된 키워드의 의미와 토픽 내 키워드 간의 관계를 고려하여 4명의 연구자가 독립적으로 분석 후에 최종 합의를 통해 선정하였다.

3-4 시계열 이상 탐지

시계열 이상치 탐지는 과거의 시계열 데이터를 바탕으로 모델을 생성하고 모델이 판단한 신뢰수준을 벗어나는 데이터를 이상치라고 판단한다. 개입 ARIMA 모형은 개입이 일어난 시점, 지시개입인지 계단 개입인지에 대한 가정을 해야 한다[25]. 이러한 단점을 보완하기 위해 이상치 형태를 4가지로 구분하여 이상치를 탐색하는 방법론을 사용하였다[26][27]. 4가지 형태의 이상치는 Additive Outlier(AO), Innovational Outlier(IO), Level Shift(LS), Temporary Change(TC)가 있고 다음과 같은 특징을 가진다. AO는 이상치 발생 시점에만 영향을 끼치는 이상치 끼치며, IO는 이상치가 발생한 뒤에도 지속적으로 영향을 끼친다는 점에서 한 시점에만 영향을 끼치는 AO와 차이가 있다. LS는 이상치가 발생한 뒤 영구적으로 데이터의 패턴을 변화시킨다. TC는 이상치가 발생한 뒤 점차적으로 줄어드는 형태를 갖는다.

본 연구에서는 시계열 이상치 탐지를 위해 오픈소스인 R의 tsoutliers 패키지를 활용하였다. 본 연구에서 활용된 표본의 수집 기간은 약 5년에 해당하는 1,612일이므로, Lopez(2016)의 연구에서 표본의 크기가 450보다 높을 때 임계값을 4.0을 사용한다는 근거를 따라[26], 임계값을 4.0으로 설정하여 이상치를 발생시킨 개입 사건을 분석하였다. 이상치 탐지 분석 결과 총 42건의 이상치가 발생하였다. 단순한 서비스 이용량 증가로 리뷰 수가 급증한 경우를 제외하기 위하여 이상치로 분류된 일자 중에서 가장 많이 발생한 토픽이 전체 리뷰 중에서 차지하는 비율이 30%를 초과하는 경우만 유의미한 이상치로 분류하였다. 그 결과, 42 건의 이상치 중에서 12건의 유의미한 이상치를 추출하여 토픽별 나타나는 양상과 구체적으로 관련된 이슈가 무엇인지 분석하였다.

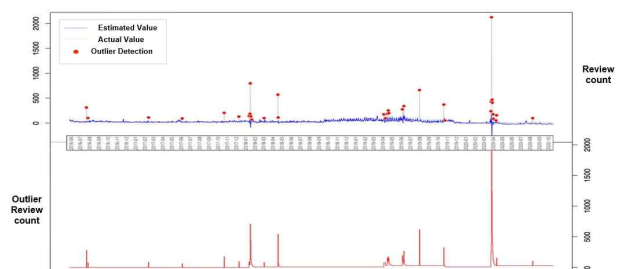


그림 3. 시계열 이상치 탐지 결과
Fig 3. Time Series Outlier Detection Results

3-5 감성 분석

감성분석은 문서, 문단, 문장 등의 텍스트를 긍정 또는 부정으로 분류함으로써 사람의 태도, 성향과 같은 주관적인 데이터를 연구하는 학문이다[28]. 감성분석은 크게 사전 기반 감성분석, 머신러닝 기반 감성분석이 있다. 사전기반 감성분석은 미리 정의된 감성사전을 구축한 뒤, 감성사전을 기반으로 어휘를 판단하여 문장 등의 감성 극성을 판단한다[28]. 머신러닝 기반 감성분석은 문맥(context) 안에서 단어의 감성을 판단하고 문맥 내의 감성의 가중치를 종합하여 문맥의 감성을 판단한다. 본 연구는 오픈소스 패키지 TEANAPS를 활용하여 이상치로 분류된 12개의 일자의 리뷰를 대상으로 감성분석을 진행한 후 일자별 평균 부정 리뷰의 비율을 도출하였다. TEANAPS 패키지는 KoBERT(Korean Bidirectional Encoder Representations from Transformers)를 사용하여 사전 훈련된 언어 모델을 제공한다.

IV. 연구 결과

4-1 토픽 모델링 결과

첫 번째 토픽은 ‘결제’로 ‘현금’, ‘카드’, ‘페이’ 등 결제 방식에 대한 단어들이 함께 추출되었다. 서비스를 매 회 이용할 때마다 거치는 결제 서비스의 절차가 간편하며 사용자가 느끼는 불편을 최소화할 수 있어야 한다. 배달 애플리케이션을 이용하는 구매자의 주 이용 프로세스가 ‘검색-주문-결제’의 3단계로 이루어진다는 점에서 결제 프로세스에 대한 평가는 애플리케이션에 대한 만족도로 직결되며 이를 위한 지속적인 개선 노력이 필요하다.

두 번째 토픽은 ‘기업 윤리’로 ‘수수료’, ‘자영업자’, ‘광고’, ‘소비자’, ‘기업’ 등의 단어가 도출되었으며, 이 중 ‘수수료’ 키워드가 가장 높은 가중치를 기록하였다. 이 토픽은 2020년 4월 배달의 민족을 운영하는 ‘우아한 형제들’이 배달 매출의 5.8%를 수수료로 떼는 정률제를 도입하는 이슈와 관련되어 발생하였다. 또한 독일 기업 딜리버리히어로가 ‘배달의 민족’, ‘요기요’, ‘배달통’을 모두 소유하며 시장의 98%를 점유하게 된 상황에서 자영업자에게 부담을 주는 수수료 정책은 사회적으로 반감을 일으켰고 불매운동이 이어지자, 배달의 민족은 공식 사과와 함께 수습에 나섰다. 이는 서비스를 제공하는 기업의 정책이 기업 윤리를 위반한다고 판단될 때 기업에 치명적인 영향을 줄 수 있음을 보여주고 있다.

세 번째 토픽은 ‘이벤트’로 ‘이벤트’, ‘할인’, ‘쿠폰’, ‘포인트’, ‘사용’ 등의 단어가 추출되었다. 플랫폼 서비스 업체는 사용자 혹은 공급자에게 인센티브를 제공함으로써 플랫폼에 더 많은 고객이 유입 또는 잔류할 수 있도록 마케팅 전략을 낸다[29]. 배달 애플리케이션의 경우 이벤트는 플랫폼 서비스에서 제공하는 이벤트와 공급자인 음식점에서 자체적으로 제공하는 이벤트가 있다.

표 1. 토픽모델링 결과

Table 1. Topic Modeling Results

Topic	Keywords	Explanation
Payment	Payment	Topics related to payment methods or login and authentication
	Error	
	Card	
	Pay	
	Cash	
Business ethics	Commission	Topics related to business ethics such as commission issues
	Advertisement	
	Consumer	
	Business	
	Self-employed	
Events	Discount	Topics related to service events
	Coupon	
	Event	
	Use	
	Point	
Customer reviews	Review	Topics related to reviews written by customers who are users of the service
	Photo	
	Shop	
	Write	
	Customer	
Service error	Customer	Topics related to application service errors
	Center	
	Call	
	Cancel	
	Server	
Restaurant service	Phone,	Service related topics for restaurants that are service providers
	Business	
	Restaurant	
	Menu	
	Shop	

특히 공급자 입장에서는 리뷰 관리를 통한 고객 유입이 중요하기 때문에 거의 모든 음식점에서는 리뷰 이벤트를 실시하고 있다.

네 번째 토픽은 ‘고객 리뷰’로 ‘리뷰’, ‘작성’, ‘사진’, ‘가게’, ‘고객’ 등의 단어가 추출되었다. 구매자는 리뷰를 작성함으로써 공급자에게 의견을 전달하며, 또한 다른 구매자의 리뷰를 읽음으로써 구매 의사결정에 필요한 정보를 획득한다. 온라인 쇼핑몰을 대상으로 수행한 연구에서는 온라인 리뷰를 통해 제품에 대한 태도와 경험을 공유하여 구매 의도와 재방문 의도에 영향력이 미친다는 것을 확인한 바가 있다[30].

다섯 번째 토픽은 ‘서비스 오류’로 ‘고객’, ‘센터’, ‘전화’, ‘취소’, ‘서버’, ‘관리’, ‘연결’ 등의 단어가 추출되었다. 애플리케이션을 사용하면서 발생하는 여러 가지 오류와 그에 대한 서비스 업체의 대응과 관련된 토픽이다. 2019년 한 해 동안 소비자고발센터에 제기된 배달 애플리케이션 배달의 민족, 요기요, 배달통에 대한 소비자 민원을 집계한 결과 가장 많은 민원을 차지한 것은 시스템 오류(36.6%)와 관련된 것이었다[31]. 서비스 오류와 관련하여 소비자 고발센터에 가장 많은 민원이 제기되었다는 점에서 서비스의 오류는 단순히 이용률 감소로 이어지는 것이 아니라, 애플리케이션 자체에 대한 불신과 이탈로 이어질 수 있음을 보여주고 있다.

여섯번째 토픽은 ‘음식점 서비스’로 ‘전화’, ‘업체’, ‘음식점’, ‘메뉴’ ‘가게’ 등의 단어가 추출되었다. 배달의 민족 서비스는 플랫폼 서비스로, 서비스상에서 소비자인 일반 이용자와 공급자인 음식점(업체)의 거래가 이루어지는 곳이다. 외식배달 플랫폼 서비스 공급자에 대한 신뢰는 플랫폼 서비스 자체 대한 만족도에 긍정적인 영향을 미친다. 따라서, 플랫폼 서비스를 제공하는 배달의 민족 사업자뿐만 아니라 공급자인 음식점(업체)가 제공하는 가격, 품질, 서비스가 플랫폼 서비스 고객 만족도에 영향을 미칠 수 있음을 보여준다.

4-2 시계열 이상치 탐지 및 감성 분석 결과

‘서비스 오류’ 토픽을 가진 이상치 일자는 2016-11-27, 2017-06-11, 2018-04-26로, 총 228건의 리뷰가 발생하였다. 해당 날짜 별 감성 분석을 한 결과 부정적인 감성이 차지하는 비율이 각각 69%, 69%, 83%, 가 되었다. 내용을 살펴보면 ‘무한로딩’, ‘서버 접속 오류’, ‘네트워크 오류’ 등의 내용이 대다수를 차지했다. 해당 리뷰로는 “무한로딩 아니 이거 왜 갑자기 인터넷 연결 확인하라면 계속 안되니까 인터넷 잘만되는데 배민만 안되요”, “아니 서버가 왜 이러죠? 데이터를 사용하는 다른 어플과 인터넷은 잘 돌아가는데 왜 배민어플만 데이터가 안돌아갑니까” 등의 리뷰가 있었다. 해당 일자의 리뷰들을 통해서 파악한 서비스 오류는 당일 네트워크 오류가 많이 발생하고, 특정 메뉴 사용 시 로딩이 길게 발생하는 일이 발생하였음을 확인하였다. 또한, 서비스 오류가 발생한 날짜 중 2일은 일요일, 1건은 토요일에 해당하여 모두 주말에 발생하였다. 이에 따라 ‘서비스 오류’ 측면에서 서비스를 개선하기 위해서는 주말에 서비스 모니터링을 강화하는 방법이 필요하며, 특히 ‘네트워크 서비스’와 ‘로딩’ 이 지속되는 현상에 대해 집중적으로 모니터링을 할 필요가 있음을 보여주고 있다.

‘이벤트’, ‘서비스 오류’ 토픽이 동시에 발생한 이상치 일자는 2016-07-27, 2017-02-19, 2017-10-29, 2019-04-14~15로, 총 376건의 리뷰가 발생하였다. 해당 날짜 별 감성 분석을 한 결과 해당 날짜에 발생한 리뷰 중 부정적인 감성이 차지하는 비율이 각각 84.0%, 83.6%, 75.0%, 81.8%이다. 내용을 살펴보면 “이벤트한다고 광고 무지 때리더니 접속도 안된다.”, “이벤트 참 좋다 이벤트가 갑자기 로그아웃 되어있고 서버는 계속 튕기고”, “어플에 이벤트 떠서 카톡 공유 엄청했었고 11시 딱 되서 갔는데 로그아웃되고 뭘까?” 등의 리뷰가 있었다. 배달의 민족 이벤트를 실시한 후, 접속을 했을 때 로그인 안되거나 로그아웃이 되고 이후 접속 장애가 발생하는 서비스 오류가 발생한 것으로 보인다. 이에 따라 이벤트를 기획할 때 시스템의 과부하가 되지 않도록 면밀히 대비하는 것이 필요할 것으로 보인다.

‘기업 윤리’ 토픽을 가진 이상치 일자는 2018-01-24, 2018-03-13, 2019-06-18~20, 2020-04-07로 총 1,113건의 리뷰가 발생하였다. 이는 전체 12건의 이상치 일자로 탐지된 리뷰 중

42.5%에 해당된다. 해당 날짜 별 감성 분석을 한 결과 해당 날짜에 발생한 리뷰 중 부정적인 감성이 차지하는 비율이 각각 83.0%, 90.5%, 71.7%, 73.7% 이다. 부정비율에서는 다른 토픽의 이상치에 비하여 크게 차를 보이지 않지만, 건수가 매우 많다는 점에서 배달의 민족 어플리케이션의 ‘기업 윤리’ 측면에서 발생하는 이슈들을 빠르게 처리할 필요하다는 점을 시사하고 있다.

일자 별 이슈를 살펴보면, 2018년 1월 24일에는 한 음식점에서 고객 신상정보를 확보하여 유출시킨 이슈가 화제가 되었다. 해당 일자에 발생한 리뷰로는 “개인정보 뿌려도 아무도 제재할 수 없다니.. 이 정도면 즉석만남 어플보다 위험한듯 전 요기요로 떠날게요”, “억한 심정으로 개인정보를 고의적으로 유출한 업체에게 강력한 경고 밖에 주지 않는 어플에 어떻게 주소랑 전화번호를 알려줄 수 있겠어요? 다시는 이용 안합니다” 등의 리뷰가 발생하였다.

2018년 3월 13일에 탐지된 이슈는 ‘배민신춘문예’에 출품된 미투 운동을 희화화한 작품들로 인하여 발생한 이슈이다. ‘배민신춘문예’는 음식을 소재로한 짧은 시를 출품하는 공모전이다. 배민신춘문예에서 성희롱과 성폭행에 맞서 사회적으로 고발하는 미투운동을 희화화한 출품작들이 대거 등장하였다. #Meat too -운동 지지-, "저도 당했어요" -미트(meat) 운동/그 맛에 당했어요-, '제 다리를 보더니 침을 삼키면서... -치킨 미투 운동' 등이 그 예이다. 당시 ‘배달신춘문예’에 선정된 카피문구는 배달의 민족 마케팅에 실제로 사용되었기 때문에 그 영향력과 과급효과가 매우 크다는 점에서 해당 부적절한 문구는 사회적으로 큰 이슈가 되었다. 해당 일자의 리뷰로는 “미투운동 희화화에 고문치사사건 희화화라니 배달의 민족 정말 실망입니다. 탈퇴하고 어플 지웁니다”, “미투운동을 희화화하는 콘텐츠를 공개하는 것은 관리자가 공모전에 별 관심이 없거나 윤리적이지 못한 것이겠죠. 저를 비롯한 많은 이들이 이번 사건을 계기로 이 어플을 삭제할 것입니다.” 등의 리뷰가 있었다.

2019년 6월 18일에 발생한 이슈는 할인 이벤트의 공정성 논란이다. 소비를 많이 할수록 할인 쿠폰을 받을 확률이 높아지는 이벤트였으나 VIP회원들이 받은 할인은 천 원인 반면 연예인들에게는 백만 원 상당의 할인을 제공하여 공정성 논란을 불러일으켰다. 해당 일자의 리뷰로는 “앱 사용은 일반인이, 그러나 할인 쿠폰은 연예인에게 쓰는 기업.. 저는 탈퇴하고 삭제”, “연예인한테는 1만원 할인쿠폰 남발하고 일반 소비자에게는 몇 십원씩 적립금 주는 이상한 어플 더 이상 안씁니다.” 등의 리뷰가 있었다.

2020년 4월 7일에 발생한 이슈는 수수료 이슈이다. 배달의 민족은 2019년 12월 13일 독일 딜리버리히어로에 매각되었다. 이로써 딜리버리히어로는 한국 배달 앱 시장 점유율 98% 이상을 확보하며 사실상 독점 상태가 되었다. 2020년 4월 5일 배달의 민족은 수수료 체계를 개편하며 소상공인에게 많은 부담을 주는 정책이라는 의견이 대다수로 청와대 청원에 등장하여 1만 5,475 명의 서명을 받았다. 해당 리뷰로는 “수수료 인상 너무 실망스러워요 삭제합니다 소상공인 다음이 어플입니다 생각 좀 합시다 게다가 이 시국에.. 수수료 인상이라니요”, “독일업체의 독점이라서 그런가 수수료가 너무 비싸요 배달의 민족, 배달통,

요기요가 독일업체꺼라 그런지... 수수료가 너무 부담되네요” 등의 리뷰가 있었다.

‘음식점 서비스’ 토픽과 관련된 이상치 일자에는 2018-01-24, 2019-08-11로, 총 931건의 리뷰가 발생하였다. 해당 날짜 별 감성 분석을 한 결과 부정적인 감성이 차지하는 비율이 각각 83.0%, 91.1%이다. 2018년 1월 24일 발생한 이슈는 ‘기업 윤리’ 토픽과 동시에 발생한 이슈로 한 음식점에서 고객 신상정보를 확보하여 유희시킨 이슈가 화제가 되었다. 이는 플랫폼 서비스에서 거버넌스의 부재로 발생한 이슈이다. 플랫폼 서비스에서는 구매자와 공급자의 거래를 중개함과 동시에 서로 견강한 거래가 성립할 수 있도록 기본적인 규칙을 정하는 거버넌스가 필요함을 시사한다. 고객의 개인정보에 대한 거버넌스를 수립하고, 공급자인 음식점 운영자에게 해당 거버넌스에 대한 교육을 진행하는 것이 필요할 것으로 보인다.

2019년 08월 11일 발생한 이슈는 주문 접수 오류이다. 해당 날짜의 리뷰로는 “아니 주문을 하고 결제까지 했는데 뭐 오류다 매장은 주문 받은 것이 없다. 고객센터는 전화연결도 안되고 뭐 어쩌라는 건가요 다른 곳 시키려고 해도 음식이 올라봐 다른 곳은 시키지도 못하고 취소 빨리 해주세요”, “결제만 되고 주문 내역에 안들어갔네요. 해당 가게에 확인해보니 주문이 전혀 안들어갔다고 합니다 결제 취소해주세요” 등의 리뷰가 있었다. 이용자가 배달의 민족 어플리케이션을 통해 결제를 진행하였지만, 결제 완료가 되어도 음식점에는 접수가 되지 않아서 발생한 사건이다. 이날 사용자들은 음식점에 전화를 걸어 배달 상황을 확인하였지만 아무런 답변을 듣지 못하자 불만이 폭주한 것으로 보인다. 배달의 민족 서비스에서 음식점과 서비스 오류 대응 팀 간의 채널이 있었다면, 해당 오류가 발생했을 때 좀 더 신속하게 대응이 가능할 것이다.

V. 결 론

5-1 연구 요약

모바일 어플리케이션 시장의 규모가 성장함에 따라 앱스토어에서 생성되는 리뷰수도 급증하고 있다. 생성되는 리뷰를 처리하기 위해 소수의 관리자가 리뷰를 일일이 읽고 주제를 분류하여 관련 부서에 전달하여 대응하기에는 시간과 인력이 부족하다. 본 연구는 전체 사용자 리뷰들 중에서 이상치를 탐지하여 리뷰수가 급증한 일자의 리뷰만을 분석하고, 사전에 진행된 토픽모델링을 바탕으로 주제를 분류하였다. 연구 대상으로는 국내 배달 어플리케이션 ‘배달의 민족’을 선정하였다. 구글 플레이 스토어에 존재하는 리뷰 데이터를 수집한 뒤 정규표현식을 활용하여 특이 문자들을 전처리하고 같은 단어가 반복되는 문장을 정제하였으며, 최종적으로 5개 이상의 명사를 갖는 리뷰들을 선택하였다. 데이터 분석은 LDA 토픽 모델링으로 전체

리뷰의 주요 토픽을 도출하였고, 시계열 이상치 탐지를 하여 리뷰수가 급증하는 일자를 추출하였으며, 해당 일자의 감성분석 결과와 함께 일자 별 발생한 이슈를 살펴보았다. 이상치로 탐지된 일자에서 발생한 주요 토픽으로는 ‘서비스 오류’, ‘이벤트’, ‘기업 윤리’, ‘음식점 서비스’이 있으며, ‘이벤트’ 토픽은 ‘서비스 오류’ 토픽과 함께 발생하였다.

각 발생한 이슈를 통해 배달의 민족 서비스 이슈를 짚어본 결과, 첫째 ‘서비스 오류’ 토픽과 관련되어 발생한 이상치는 모두 토요일, 일요일 주말에 발생한 것을 확인하였고, 주요 현상으로는 로딩 시간이 길게 이어지거나 네트워크 오류 및 로딩이 지속되는 오류가 주로 발생한 것으로 보인다. 감성분석 결과 부정적인 비율은 각각 69%, 69%, 83%이다. 둘째, ‘이벤트’와 ‘서비스 오류’ 토픽과 관련되어 발생한 이상치에서는 모두 이벤트 일자에 발생한 서비스 오류에 대한 리뷰가 주를 이루었으며, 이벤트 일자에 서비스 모니터링을 강화할 필요가 있음을 보여주었다. 감성분석 결과는 88%, 69%, 70%, 88%를 기록하였다. 셋째, ‘기업윤리’ 토픽과 관련하여 발생한 이상치 일자의 부정비율은 83.0%, 90.5%, 71.7%, 73.7%를 기록하였으며, 이상치로 분류한 12개의 일자의 리뷰의 42.5%가 ‘기업 윤리’ 토픽에서 발생하였다. 이는 배달의 민족 어플리케이션의 ‘기업 윤리’ 측면에서 이슈를 처리하는 방안이 필요하다는 점을 시사하고 있다. 넷째, ‘음식점 서비스’와 관련되어 발생한 이상치에는 음식점 업체에 대한 인식 및 커뮤니케이션 과정에서 발생하는 불만에 관한 리뷰가 주를 이루었다. 감성분석 결과는 부정적인 답변은 99%, 87%로 비교적 높은 비율을 보여주고 있는데, 이는 플랫폼 서비스인 배달의 민족의 서비스 모니터링에 있어서, 음식점의 서비스를 배제할 수 없음을 시사한다. 이는 배달의 민족 플랫폼 서비스에서 수요자인 고객과 공급자인 음식점 업주 간의 거래를 원활하게 하기 위한 노력을 기울일 필요가 있음을 시사한다.

5-2 한계점 및 향후 연구

본 연구는 다음과 같은 한계점이 있다. 첫째, 단일 변수인 리뷰 수만으로 이상치를 탐지한다. 본 연구에서는 이상치를 탐지하는 모델의 변수로 리뷰 수를 선정하여, 리뷰의 많고 적음에 따라서 새로운 서비스 이슈의 발생 여부를 알고자 하였다. 하지만 리뷰 건수만으로는 이상치를 판단하는 것은 어렵다. 서비스 이슈가 발생했을 때, 감성분석 결과 부정적인 리뷰의 비율의 변화 또는 각 토픽의 평균 비율 등에 변화가 발생할 수 있다. 향후 연구에서는 이와 같은 감성분석 및 각 토픽의 구성 비율의 변화의 이상치를 탐지하여 서비스 이슈를 탐지하는 모델에 대한 연구가 필요한 것으로 보인다.

둘째, 토픽모델링을 통한 토픽을 선정하고 해석하는데 연구자의 주관적 판단이 개입되었다. 텍스트마이닝은 문맥(context)을 고려하지 못하여 모호성이 생기고 이로 인해 연구자의 주관적인 판단이 들어갈 가능성이 있다. 텍스트마이닝의 한계점을 완화하고자 응집도, 혼잡도를 지표로 활용하고 4명의 연구자가 개별 분석 후에 합의하는 방식을 채택하였다. 향후 연구로는 문

맥을 파악할 수 있는 딥러닝과 같은 고도화된 자연어처리 기법을 활용하여 객관성을 확보하고 보다 세부적인 토픽선정을 할 수 있을 것으로 판단된다.

셋째, 기존 데이터에 없던 새로운 이슈가 발생한다면 토픽을 정확하게 분류하지 못한다는 한계점이 있다. 예를 들어 독과점 같은 이슈가 처음으로 발생한 것이라면 기업 윤리 토픽으로 분류하지 못할 가능성이 있다. 향후 연구에서는 처음 등장하는 이슈가 발생하였더라도 이를 탐지하여 리뷰를 분석하는 연구를 진행할 수 있을 것이다.

넷째, 분류된 토픽만으로는 해당 일자 이슈의 구체적인 내용을 알기 어렵다. 향후 연구에서는 생성요약 등을 활용하여 토픽이 할당된 후, 구체적으로 어떤 내용을 포함하고 있는지 파악할 수 있을 것이다.

5-3 연구의 시사점

본 연구의 시사점은 다음과 같다. 첫째, 사용자 리뷰 데이터로 LDA 토픽모델링, 감성 분석, 이상치 탐지를 활용하여 어플리케이션 리뷰 분석을 통한 이슈 대응 프로세스를 정립한 것에 있다. 기존에 존재하는 사용자 리뷰 데이터를 활용한 연구가 대부분 주제 분류를 중심으로 이루어졌던 반면, 본 연구에서는 주제를 분류하는 것에 시계열 이상치를 탐지하는 ARIMA 모델을 추가하여 리뷰를 통해 서비스 이슈를 빠르게 탐지하고 대응하는 연구를 진행하였다는 것에 학문적인 의의가 있다. 모바일 어플리케이션 시장의 빠른 성장과 함께 관련 어플리케이션 서비스 리뷰가 증가하는 가운데 생성되는 리뷰의 주제를 과거 데이터를 기반으로 자동으로 분류하여 내용 파악에 시간을 절감하고, 리뷰 이상 탐지를 통해 어플리케이션 서비스 운영 비용 절감에 이바지하고자 한다는 점에서 실무적인 의의가 있다.

감사의 글

본 연구는 2020년도 연세대 미래융합연구원(ICONS)의 지원에 의하여 이루어진 연구로서, 관계부처에 감사드립니다.

참고문헌

[1] iResearch Industry Overview 2019 June Available: <http://www.iresearchchina.com/>

[2] Thanh-Son Nguyen, Hady Wirawan Lauw, and Panayiotis Tsaparas. "ReviewSynthesis for Micro-Review Summarization." *In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, WSDM 2015, Shanghai, China*, pp. 169-178, 2015 February.

[3] Cuiyun Gao, Jichuan Zeng, Michael R. Lyu, and Irwin King. 2018. "Online app review analysis for identifying emerging

issues." *In Proceedings of the 40th International Conference on Software Engineering (ICSE '18)*. Association for Computing Machinery, New York, NY, USA, 48-58, 2018 May.

[4] Xiaodong Gu and Sunghun Kim. "'What Parts of Your Apps are Loved by Users?' (T)". In: *30th IEEE/ACM International Conference on Automated Software Engineering, ASE 2015, Lincoln, NE, USA*, pp. 760-770, 2015 November

[5] Business Insider. Facebook Messenger is getting slammed by tons of negative reviews. Available: <http://www.businessinsider.com/facebook-messenger-app-store-reviews-are-humiliating-2014-8>.

[6] D. Pagano and W. Maalej, "User feedback in the appstore: An empirical study," in *Proceedings of the 21st IEEE International Requirements Engineering Conference (RE), Rio de Janeiro*, pp. 125-134, 2013 Jul.

[7] App Annie. [Internet] Available: <https://www.appannie.com>.

[8] Kwang-Kook Kim, Yong-Hwan Kim, Ja-Hee Kim, "A Study on Customer Satisfaction of Mobile Shopping Apps Using Topic Analysis of User Reviews", *The Journal of Society for e-Business Studies* 23(4), 41-62(22 pages), 2018.11

[9] Hong Jung-lim, Yoo Mi-rim, Choi Bo-reum, "An Analysis of Mobile Augmented Reality App Reviews Using Topic Modeling", *Journal of the Korean Digital Contents Society, Vol. 20*, 1,417 - 1,427, 2019.07

[10] Chen, Ning and Lin, Jialiu and Hoi, Steven C. H. and Xiao, Xiaokui and Zhang, Boshen, "AR-Miner: Mining Informative Reviews for Developers from Mobile App Marketplace", Association for Computing Machinery, 2014

[11] Korea Mobile Internet Business Association, 2019 Korea Mobile Content Industry Status[Internet], Available: <http://moiba.or.kr/main/bbs/info>

[12] Korea Information Society Development Institute(KISDI), A Study on Economic and Social Value of O2O industry, 2019.02

[13] Mobile Index, Delivery agency service app usage analysis, Available: <https://www.mobileindex.com/report/report>

[14] Jo Mi-na, "Integrating SERVQUAL and Kano's Model into Quality Function Deployment for Delivery Application Service Excellence Development", *Journal of Tourism Sciences*, 44(4), 103-121, 2020.06

[15] Yu jin-uk, Kim dong-hyeok, Kim jung-hyun, Im yang-won, Im hang-yu, "Design and Implementation of Delivery Application for Small City", in *Proceedings of the CICS 2014 Information and Control Conference*, 354 - 356 (3 pages), 2014

[16] Jeong Jin-myung, Nam Jae-won, "An Approach on the Protection for Baedal App Service Users", *Consumer Problems Research* Vol. 46, 207-230 (24 pages), 2015.08

[17] Chung, HaeKyung, "Study on the Usability Test for Mobile Applications - Focused on Delivery Applications" Journal of The Korean Society of Illustration Research, vol. 45, pp.5-14, 2015

[18] Jo Mi-na, Cha Jae-bin, "The Influence of Delivery App Quality Characteristics on Consumer Attitude and Behavior Intention -Focusing on Technology Acceptance Model (TAM)", in *Proceedings of Korean Tourism Association International Conference*, Volume 81, 139-143 (5pages), July 2017

[19] I. Morales-Ramirez, D. Muante, F. Kifetew, A. Perini, A. Susi, and A. Siena, "Exploiting user feedback in tool-supported multi-criteria requirements prioritization," in *Proceedings of the 25th IEEE International Requirements Engineering Conference (RE)*, Lisbon, Sep. pp. 424-429, 2017.

[20] Kim gwang-guk, Kim yong-hwan, Kim jae-hui, "A Study on Customer Satisfaction for Mobile Shopping Apps Using User Review Topic Analysis", Journal of the Korean Society for Electronic Commerce, Vol. 23, 41-62page, 2018.11

[21] Yu-Jeong Yang, Bo-Hyun Lee, Jin-Sil Kim, Ki Yong Lee, "Development of An Automatic Classification System for Game Reviews Based on Word Embedding and Vector Similarity", The Journal of Society for e-Business Studies 24(2), 1-14(14 pages), 2019.05

[22] Nan Hu, Jie Zhang, and Paul A. Pavlou. "Overcoming the J-shaped distribution of product reviews," *Communications of the ACM*, Vol. 52, No. 10, pp. 144-147. October 2009.

[23] Park Ju Seop, Hong Soon-Goo, Kim Jong-Weon. "A Study on Science Technology Trend and Prediction Using Topic Modeling," *Journal of the Korea Industrial Information Systems Research*, 22(4), 19-28. (2017).

[24] Sae-Mi Lee, Soon-Goo Hong. "Analysis of Blockchain Trends Using Topic Modelling Technique," *The Korea Institute of Information and Communication Engineering*, (), 44-47. (2019).

[25] Lee, Ji Seon, Yoon, Yoon Jin. "A Study on the Outliers Detection in the Number of Railway Passengers for the Gyeongbu Line From Seoul to Major Cities Using a Time Series Outlier Detection Technique," *Journal of Korean Society of Transportation*, 35(6), 469-480. (2017).

[26] Lopez J., López M. J., Package 'tsoutliers'. (2016). <https://cran.r-project.org/web/packages/tsoutliers/tsoutliers.pdf>

[27] Chen C., Liu L. M. "Joint Estimation of Model Parameters and Outlier Effects in Time Series.," *Journal of the American Statistical Association*, 88(421), 284-297. (1993).

[28] Sang Hoon Lee, Jing Cui, Jong Woo Kim. "Sentiment analysis on movie review through building modified

sentiment dictionary by movie genre.," *Journal of Intelligence and Information Systems*, 97-113 pages, 2016.

[29] Harvard Business Review. Strategies for Two-Sided Markets [Internet]. Available : <https://hbr.org/2006/10/strategies-for-two-sided-markets>

[30] Angela V. Hausman, Jeffrey Sam Siekpe, "The effect of web interface features on consumer online purchase intentions", *Journal of Business Research* Volume 62, Issue 1, pages 5-13, 2009.

[31] Korea Consumer Agency, Delivery Application Survey, [Internet]. Available : <https://www.kca.go.kr/kca/=6012536>.



정지훈(Ji-Hoon Jung)

2013년 : 숭실대학교 정보통계보험수리학과 (이학사)
2020년 : 연세대학교 정보대학원 (비즈니스 빅데이터 분석 트랙 석사과정)

2020년~현재: 연세대학교 정보대학원 비즈니스 빅데이터 분석 트랙 석사과정
※관심분야 : Big Data Analysis, Data Mining, Machine Learning



정혜인(Hye-In Chung)

2010년 : 성균관대학교 신소재공학부 (공학사)
2020년 : 연세대학교 정보대학원 (비즈니스 빅데이터 분석 트랙 석사과정)

2020년~현재: 연세대학교 정보대학원 비즈니스 빅데이터 분석 트랙 석사과정
※관심분야 : Big Data Analytics, Data Mining, NLP



이준기(Zoon-Ky Lee)

1985년 : 서울대학교 계산통계학(학사)
1991년 : 카네기멜론대학 사회심리학 (석사)
1999년 : 남가주 대학교 경영정보학(박사)

2004년 ~ 현재 : 연세대학교 정보대학원 교수
※관심분야 : Big Data Analytics, Digital Transformation, Open Collaboration