

앱 이용실적과 앱 리뷰 감성분석의 통합적 모델 구축에 관한 연구

A Study on Building an Integrated Model of App Performance Analysis and App Review Sentiment Analysis

김동욱*, 김성범**

한국과학기술원 기술경영전문대학원*, 금오공대 IT융합학과**

Dongwook Kim(yooksonbaekon@gmail.com)*, Sungbum Kim(it89@kumoh.ac.kr)**

요약

이 연구의 목적은 모바일 앱 실적 변수 간의 관계를 파악하여 예측 가능한 앱 실적 변수의 추정 모델을 구성하고 앱 리뷰가 앱 실적 지표에 미치는 영향을 검증하는 것이다. 연구1과 2에서는 상관분석과 기계학습의 랜덤 포레스트 회귀 추정을 사용하여 앱 실적 간의 관계를 도출하고 앱 실적 추정 모델링을 수행하였다. 연구3에서는 앱 리뷰를 텍스트 마이닝의 감성분석을 사용하여 일별 감성 점수를 도출한 후 다변량 시계열분석을 사용하여 앱 리뷰의 감성점수가 앱의 일일 설치 횟수에 선행하여 영향을 주는 것을 발견하였다. 앱을 개발하고 서비스하는 기업은 앱 실적 지표와 앱 리뷰에서 제기되는 불만족과 고객 니즈를 검토하여 적기에 앱을 개선하고 마케팅 판매촉진활동의 시점과 방향성을 도출할 수 있다.

■ 중심어 : | 앱 실적 지표 | 앱 리뷰 | 기계학습 | 감성분석 | 다변량 시계열 |

Abstract

The purpose of this study is to construct a predictable estimation model that reflects the relationship between the variables of mobile app performance and to verify how app reviews affect app performance. In study 1 and 2, the relationship between app performance indicators was derived using correlation analysis and random forest regression estimation of machine learning, and app performance estimation modeling was performed. In study 3, sentiment scores for app reviews were by using sentiment analysis of text mining, and it was found that app review sentiment scores have an effect one lag ahead of the number of daily installations of apps when using multivariate time series analysis. By analyzing the dissatisfaction and needs raised by app performance indicators and reviews of apps, companies can improve their apps in a timely manner and derive the timing and direction of marketing promotions.

■ keyword : | App Performance Indicator | App Review | Machine Learning | Sentiment Analysis | Multivariate Time Series |

I. 서론

1. 연구목적

스마트폰의 대중화와 네트워크의 고도화로 모바일 앱은 생활 속의 필수적인 서비스가 되었고 소비자는 모바일 앱 없이는 편리한 생활을 하지 못하는 시대가 되

* 이 논문은 2019년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임(NRF-2019S1A5A8035991)

접수일자 : 2021년 10월 05일

수정일자 : 2021년 11월 04일

심사완료일 : 2021년 11월 04일

교신저자 : 김성범, e-mail : it89@kumoh.ac.kr

었다. 모바일 앱 없이는 택시를 잡기도 힘들고 음식을 빠른 속도로 주문할 수도 없다. 반면 스마트폰과 모바일 앱은 이것을 사용하기 어려워하는 계층에게는 새로운 정보격차의 벽으로 작용하게 되었다.

모바일 앱이 생활의 필수재가 되면서 이를 제공하는 기업의 경쟁도 치열해지고 있으며 앱 개발 서비스 기업의 경쟁과 생존을 위한 방식도 다양해지고 있다. 기업은 앱의 유료화를 통한 수익화를 모색하거나 가입자를 늘려서 많은 사용자가 앱을 사용하게 하고 가입자를 기반으로 새로운 수익원의 창출을 도모한다. 기업은 앱을 통하여 어떻게 수익을 창출할지, 앱과 관련한 어떤 지표를 관리하는 것이 효과적일지 고민을 시작하게 되었다.

그러나 기존의 모바일 앱에 관한 선행연구들은 앱 사용 의도와 영향 요인[1][2], 이용 동기의 품질요인[3], 앱 리뷰에 대한 토픽 모델링, 감성분석[4][5], 앱 리뷰가 개발사 성과에 미치는 영향 분석[6] 등의 주제에 국한하여 시도되었다. 그리고 앱 이용에 대한 시계열 데이터에 대한 구축과 분석에 대한 시도는 찾아보기 힘들며, 앱 감성분석 점수의 시계열 모형 접근은 존재하지 않았다.

이 연구는 앱 이용실적에 대한 다양한 지표의 통합과 앱 리뷰 감성 점수의 결합을 통해서 앱 개발자, 마케터와 기획자에 대한 실무적 시사점을 도출하는 것을 목표로 한다. 구체적으로 모바일 앱을 기획 및 개발하고 모바일 서비스를 진행하고 있는 기업에 필요한 앱 실적과 성장에 관한 관리지표를 제시하는 것을 목적으로 한다.

2. 연구문제와 접근방법

모바일 앱 실적을 추적하는 것은 매우 어려운 일이다 [7]. 또한 자사의 앱 실적을 다른 경쟁 서비스의 실적과 비교해야 할 필요성도 있다. 이런 이유로 다양한 앱의 실적을 추적하고 그 자료를 수집 정리하여 앱 개발자 및 앱 서비스 회사에 경쟁사 앱 실적 데이터를 제공하는 비즈니스가 하나의 사업 아이템으로 등장하고 있다.

사용자들은 앱을 사용하면서 다양한 사용의 디지털 흔적을 남긴다. 새로운 앱을 설치하는 행위, 일간, 주간, 월간 기간의 앱 사용의 시간, 하루에 몇 번을 구동시키는지, 사용 후 얼마 만에 삭제했는지에 대한 로그 정보

를 남긴다. 앱 사용의 패턴은 인구학적 정보와 시간 데이터와 결합하여 앱 사용의 패턴에 관한 보다 유용한 정보를 생산한다.

이 연구에서는 이와 같은 지표들이 서로 어떤 관계가 있는지, 앱 스토어의 앱 평가 평점과 어떤 관계가 있는지 다양한 앱 실적 횡단면 자료를 사용하여 변수 간에 유의미한 관계를 탐색한다. 또한 사용자가 리뷰에서 서술한 내용이 앱 실적 지표와 어떤 관계를 맺는지를 탐구한다. 이와 같은 연구목적에 달성하기 위하여 다음과 같은 연구질문을 제기한다.

연구문제 1, 앱 스토어의 평균 평점과 앱 사용 실적 지표 간의 관계는?

연구문제 2, 앱 이용실적으로 사용되는 다양한 지표 간의 관계는 어떠한가? 앱 실적 지표 간에 관련성이 높은 지표는 무엇이며 다른 지표를 입력변수로 하여 가장 안정적으로 목표변수로 추정할 수 있는 앱 실적 지표는 무엇인가?

연구문제 3, 앱 실적은 앱 리뷰로부터 영향을 받는가?

이와 같은 연구 질문에 대한 답을 찾기 위하여 자료 수집과 분석을 포함하여 아래와 같이 3단계의 연구로 구성하여 단계마다 연구목적에 가장 적합한 프로세스와 방법론을 사용하여 접근한다.

연구문제 1은 앱 스토어의 각 앱에 대한 평균 평점은 사용자가 특정 앱을 내려받아 사용할지를 고민할 때 사용자의 판단에 영향을 미칠 수 있는 가시적인 위치에 있음을 중시한다. 이와 같은 사실로부터 평균 평점은 사용자의 앱 사용 실적에 영향을 미칠 것이라는 가설을 설정할 수 있다. 그리고 이를 검증하기 위해서 다른 앱 실적과 상관관계를 분석하고, 평균평점을 목표변수로 설정하고 다른 앱 실적 지표를 설명변수로 설정하여 유효한 추정 모델을 구성할 수 있는지 검증한다. 또한 역으로 다른 앱 실적 변수 추정 시 평균 평점의 설명변수로의 중요도를 검증한다.

연구문제 2의 해결을 위해서 앱 실적 지표 중 다른 앱 실적 지표와 가장 상관관계가 높거나 다른 앱 실적 지표로부터 안정적으로 추정이 가능한 지표는 무엇인지 탐색한다. 실무적으로 관리변수로 사용하는 실적 변수 중에서 목표변수(종속변수)로 사용할 수 있는 변수

표 1. 연구질문, 방법론, 연구단계

	세부 연구질문	검증 방법	방법론	연구단계
RQ1	앱 스토어의 평균평점과 앱 사용 실적 지표 간의 관계는?	M1: 평균평점과 다른 앱 실적 변수 간의 상관분석	상관분석	연구1
		M2: 평균평점을 다른 앱 실적 변수로 추정	머신러닝	연구2
		M3: 다른 앱 실적 변수 추정 시 평균평점의 중요도 분석	머신러닝	연구2
RQ2	앱 실적지표 간에 관련성이 높은 지표는?	M4: 앱 실적지표 간의 상관분석	상관분석	연구1
		M5: 특정 앱 실적지표를 다른 앱 실적지표로 추정	머신러닝	연구2
RQ3	앱 실적은 앱 리뷰로부터 영향을 받는가?	M6: 시계열 상에서 특정 감성점수가 앱 실적에 선행하는지 여부	다변량 시계열	연구3

를 찾는다. 앱 실적 지표로서 일일활성사용자수(DAU), 월간활성사용자수(MAU)와 앱 이용자의 실질적 앱 사용을 의미하는 앱 전환지표로서 고착도(DAU/MAU)를 검토하고, 추가로 일일평균실행횟수(activities)와 기타 변수들의 목표변수 사용 가능성을 머신러닝을 이용하여 분석한다.

연구문제 3에서 앱 리뷰가 앱 실적에 영향을 미치는 지를 분석하기 위해서는 앱 리뷰의 양과 기간을 설정해야 한다. 너무 많은 리뷰는 오히려 의사결정에 역작용을 초래할 수 있다. 앱 리뷰가 사용을 고민하는 미래의 사용자에게 영향을 미쳐 앱 실적에 영향을 미치기 위해서는 모바일 앱에 접근하는 소비자들이 앱 리뷰를 읽을 수 있는 앱과 모바일 사이트의 가용 화면 안에 리뷰가 등장해야 한다. 대부분의 앱이 일일 단위로 많은 수의 리뷰가 누적되므로 월 단위로 측정해서는 영향 관계를 분석 하기가 어렵다. 따라서 앱 사용에 영향을 주는 범위에 있는 리뷰를 일별로 정리하여 분석하는 것이 적합하다.

텍스트 마이닝의 감성분석과 다변량 시계열 분석을 사용하여 위의 연구문제 3을 검증한다. 사용자가 앱 스토어에서 앱을 내려받아 설치하기 전에 앱 리뷰를 읽어본다는 가정하에 앱 리뷰가 앱 실적에 영향을 줄 수 있을 것으로 가정한다. 앱 리뷰를 감성점수로 표현한 후 앱 리뷰의 다양한 감성점수가 특정 앱 실적 지표에 시계열적으로 선행하는지 탐구한다. 연구3은 앱 이용실적과 앱 리뷰 감성점수 간의 통합적 분석이다.

기존의 연구 중 특별히 앱 이용에 대한 시계열 데이터에 대한 분석은 좀처럼 찾아보기 힘들며 앱 실적과 앱 리뷰를 연계하여 통합적 모델을 구축한 연구 역시 이 분야에서는 새로운 시도이다. 이 연구는 앱 이용실적에 대한 다양한 지표와 앱 리뷰의 감성을 분석하여

도출한 감성 변수를 통합하여 시계열 상에서 의미 있는 지표를 도출한다. 이를 사용하여 앱 마케팅과 기획자에 대한 실무적 시사점을 도출하는 것을 목표로 한다.

II. 선행연구

1. 모바일 앱 연구

연구자들은 모바일 앱을 새로운 기술과 제품, 서비스로 인식하면서 이론적으로는 기술수용모형(technology acceptance model: TAM)[8]과 통합기술수용모형(Unified theory of acceptance and use of technology: UTAUT)[9][10]에 기초하여 모바일 앱에 대한 사용자 수용성 중심의 연구를 진행하였다. 이들 연구는 특정 모바일 앱의 유용성, 신뢰, 실용주의적 동기, 쾌락주의적 동기, 사회적 관계 [11]등의 변수가 사용 의도[12] 또는 지속적 사용 의도 [11]에 어떤 영향을 미치는가에 집중하였다. 이 연구들은 인지된 품질(perceived quality)[13][14]과 만족도(satisfaction) [15][16]를 추가적인 변수로 사용하여 확장되었고 새로운 카테고리의 새로운 모바일 서비스가 등장할 때마다 소비자들이 어떻게 반응하는지를 다양한 변수를 추가하면서 진행되었다.

실무적인 활용성에 좀 더 접근한 연구들은 대중들이 어떤 모바일 앱을 좀 더 빈번하게 사용하는지, 어떤 서비스가 사용자의 지속적 사용을 획득하는지에 관심을 가지기 시작했다. 지속적 사용과 잦은 사용이 기업의 매출과 수익으로 연결되면서 모바일 앱 사용자가 만약 사용을 멈춘다면 무엇 때문에 서비스 사용을 중지하는지 그 원인의 파악도 중요해졌다[17].

모바일 앱과 사용자와의 관계는 기존에 모바일 앱을

사용하지 않았던 영역에서 새롭게 사용하게 된 영역은 어디인지, 이를 통해 새로이 공급자와 수요자가 연계 되는 혜택은 무엇인지에 주목하면서 확장되어 나갔다. 새로운 영역 중 의료와 헬스 서비스를 대표적인 사례로 선정할 수 있다[18]. 특별히 의료는 사용자의 생명과 건강과 관련 있는 분야이기 때문에 모바일 앱의 품질에 대한 높은 관심을 가지게 되었고[19-21], 비교적 신경과 분야[22]나 안과 분야[23]의 전문적인 분야까지도 모바일 앱을 사용자 측정 도구로 사용하려고 하는 노력이 지속되었다. 그러나 의료와 헬스 분야는 앱이 제공하는 효용뿐만 아니라 모바일 앱 사용의 적정성도 확보되어야 하므로 이에 대한 규제의 정도와 방향성 수립도 중요한 이슈로 인식되고 있다[24-26].

마지막으로 모바일 앱에 관하여 연구자들이 주목한 주제는 앱을 개발하고 서비스하는 기업의 관점이 강하게 반영된 연구 들로서 앱을 운영하면서 비용의 최소화 또는 수익의 극대화를 위하여 어떤 활동과 지표 측정이 필요한가를 탐구하는 연구들이다. 앱 운영체제(OS) 플랫폼에 따라 추가로 발생하는 개발상의 비용을 절감하기 위한 고민은 대표적인 개발 비용의 관점에서 접근한 연구이다[27]. 그러나 기업의 생존이라는 관점에서는 비용의 최소화도 중요하지만, 기업의 성장을 위한 가입자의 증대와 수익의 확보가 더 본질적인 주제이다. 앱을 개발 후 앱 스토어에 업로드 후 서비스를 시작하는 기업의 경우 아이폰의 생태계와 구글 안드로이드 마켓 중 어느 쪽이 매출 증대에 도움이 되는지에 관한 고민을 하게 된다. 이것은 앱을 어떤 시장에 출시해야 앱 수익 극대화를 달성할 수 있을까 하는 기업의 관심사를 잘 반영하고 있다[28].

모바일 앱 서비스 기업은 궁극적으로 어떤 활동을 통해서 매출을 증대할지를 고민한다. 기업의 전략에 부합하는 앱 개발과 서비스 전략을 수립하고 어떤 지표를 통해서 프로모션과 수익증대를 점검할지 성장과 수익화를 위한 지표설정과 관리체계를 확립해야 한다.

학술적인 연구에서는 두 가지 지표에 주목하고 있는데 첫째는 앱 실적의 지표로서 앱 스토어의 앱 평점에 주목하여 개별 앱의 품질과 유용성을 측정하여 점수화하는 것이고, 둘째는 앱 리뷰에 관련된 연구로서 텍스트 마이닝 분석의 하나인 감성분석을 사용하여 사용자

의 긍정 및 부정의 감성을 점수화하여 앱 평점을 보완하는 것이다.

2. 앱 평가와 기업의 수익

앱 스토어가 앱의 평가 정보를 제공하는 것은 앱 스토어의 필수적인 활동이며 앱을 개발 및 서비스 하는 회사는 이 점수에 관심을 가질 수밖에 없다. 앱 스토어가 앱 평가를 객관적이고 공정하게 진행하는 것은 앱 스토어가 신뢰를 얻는 데 매우 중요한 요인이다.

앱 스토어는 제품 생명 주기가 짧고 버전 업이 빈번한 앱을 공정하고 객관적으로 평가하기 위한 앱 평가(app rating)시스템의 구축을 위한 노력을 계속하고 있다[29][30].

앱의 평가에 관한 연구는 실험환경을 설정하고 실제 사용자들에게 앱을 사용하게 한 후 고객의 사용패턴과 좌절 포인트를 발견하여 서비스 기획의 인사이트와 개선점으로 연결하려는 노력으로 확대되었다[31]. 앱의 사용자 평가는 사용성(usability) 평가로 대표되며 앱이 얼마나 사용하기 편리한지, 배우기 편리한지, 사용자의 기대 만족도에 부응하는 지로 평가한다[32].

모바일 앱 평가(rating)에 관한 연구는 모바일 앱의 다양한 카테고리의 분야별로 확장되었으며 재무(finance) 분야의 앱 평가지표를 기능, 미적, 신뢰, 행태의 관점에서 개발하여 앱 스토어의 별점 평가(star rating)[33]를 보완하기도 하였고, 디지털 헬스 분야의 모바일 앱 평가를 기능성과 미학(functionality and aesthetics), 참여와 정보(Engagement and information)의 관점에서 연구하기도 하였다[34].

모바일의 앱 스토어에서 앱에 대한 개별 사용자 평가는 앱의 기능과 성능에 더욱 중점을 둔 평가이기 때문에 이러한 평가 결과가 기업의 수익에 어떻게 영향을 미치는지는 쉽게 판단하기 어렵다. 이러한 이유로 기업에서는 앱 소지율을 관리지표로 사용하여 사용자가 앱을 사용을 중지하고 왜 이탈하는지 또는 지속해서 사용하게 하는 요인은 무엇인지를 추가 조사하기도 한다. 이를 통해 기업이 지속 가능한 수익 증가를 위해서 무엇을 개선해야 하는지에 관한 추진과제를 도출하고 있다[35].

모바일 앱의 보유율(retention rate)은 궁극적으로

앱 서비스 성장과 수익증대에 연결되므로 모바일 앱 서비스 기업은 어떻게 하면 이 비율을 높일지 고민하게 된다. 간단하게 말하면, 앱 소지율의 촉진 요인은 증가시키고 저해 요인은 제거하는 것이 필요하다.

앱 소지율에 영향을 미치는 요인 중에서는 앱을 서비스하는 기업이 통제할 수 있는 요인이 있고 통제 불가능한 요인이 있다. 앱 소비 환경인 단말기의 성능과 네트워크 환경은 모바일 앱을 구동시키는 단말기의 배터리의 성능과 네트워크 지연(latency)을 통해서 모바일 앱의 소지율에 영향을 주지만 이러한 요인들은 앱 서비스 기업이 완벽하게 통제가 어려운 부분이다[36], 이러한 통제가 어려운 앱 소비의 외부 환경적 요인은 앱을 출시하는 시점의 단말기와 네트워크 상황이 앱 기능과 특성을 결정한다는 의미도 가진다. 이러한 소비 환경을 서비스 기획에 반영하지 못하면 너무 앞선 서비스로서 제대로 된 앱의 기능이 구현되지 못하여 소비자의 외면을 받을 수도 있다.

모바일 앱의 소지율은 앱의 유료와 무료의 정책에 영향을 받기도 하는데 사용 중에 발생하는 추가 아이템 유료화와 프리미엄 버전의 출시가 소지율에 영향을 미칠 수 있다[37-39]. 기업이 앱을 출시하는 경우 궁극적으로 앱의 유료화의 시기와 범위를 고민하여야 하며 단기간 수익의 증대와 가입자 증가라는 두 지표는 기업이 결정해야 할 어려운 선택 사항이기도 하다[40].

요약하면 앱 스토어는 앱 스토어 기존의 앱 평가(app rating)에 대한 개선을 고민하고 있고 개별적인 앱의 품질 평가는 전문가에 의한 4가지 축의 다차원 평가 체계인(engagement, functionality, aesthetics, information) Mars의 평가 체계와 인간 컴퓨터 상호작용(human computer interaction: HCI)관점의 유용성 평가를 중심으로 진행되고 있다[21][33][41][42]. 반면에 앱 출시 기업은 앱 서비스의 보유율(retention rate)과 수익 규모에 초점을 맞추며 앱 수익의 극대화를 위한 전략을 추진하고 있다[36][40].

3. 앱 리뷰와 감성분석

앱 리뷰는 모바일 앱 사용자가 그들의 의견을 표출하는 유일한 소통 창구이다. 그것은 칭찬일 수도 있고 불만일 수도 있으며 특별히, 기능상 예러에 대한 불만이

나 새로운 기능의 요구일 수도 있다. 앱이 작동하지 않거나 네트워크상의 문제를 보이는 경우 기업은 사용자의 리뷰를 모니터링 하여 즉각적인 보완조치를 취할 수 있다[43]. 이러한 대응이 늦어지는 경우 사용자는 앱을 다시 사용하지 않거나 삭제해 버릴 수 있으므로 앱 리뷰는 모바일 앱을 서비스하는 기업의 입장에서는 서비스의 운명을 좌우하는 고객의 중요한 피드백이다.

앱 리뷰는 앱 스토어의 앱 평가(app rating)와 상호 보완적이다. 앱 스토어 리뷰의 긍정 부정의 감성(sentiment)분석을 통하여 앱 평가를 보완할 수도 있고[44], 모바일 앱 리뷰를 토픽 모델링을 사용하여[45], 사용자가 어떤 이야기를 하고 있고, 불만족과 만족 요인은 무엇이고, 사용자가 원하는 추가 기능은 무엇인지를 추출할 수 있다. 감성분석과 토픽 모델링의 결과는 앱 스토어의 평점과 연계를 통해서 앱 평점 점수를 보완하는 역할을 한다.

앱의 리뷰와 평점이 중요해 짐에 따라 사용자가 작성한 앱의 평가(rating)와 리뷰(review)의 신뢰성이 중요한 문제로 등장한다. 리뷰 내용 중의 거짓 리뷰(fraud)의 발생을 빅데이터 분석으로 찾아내려는 노력이 앱 리뷰를 관리하는 기업의 필수적인 임무가 되고 있다[46]. 거짓 리뷰와 진실한 리뷰를 분류하는 기술을 사용하여 유니콘 기업이 된 엘프(Yelp)는 이런 진실한 리뷰의 중요성을 보여주는 사례이다. 최근 엘프 데이터가 정말 도움을 주는가" 라는 주장을 제기한 연구는 새로운 시각을 제기하고 있지만, 역설적으로 앱 리뷰의 중요성을 다시 한번 강조하는 사례이다[47].

앱 실적을 평점 또는 소지율을 가지고 측정하거나, 앱 리뷰를 사용하여 앱의 퍼포먼스를 측정하고 모니터링 하는 것은 모바일 앱 서비스 전략과 성장을 위해 중요한 작업이다. 우리의 연구는 앱의 실적 모니터링이 고객관계관리(CRM) 수준이 될 수 있도록[48] 각 측정치 간에 관계를 횡단면적으로 그리고 시계열적으로 탐색한다.

III. 방법론

1. 데이터와 변수

이 연구는 앱 실적 자료를 수집하고 정리하는 기업인 풀러(Fuller)사의 자료를 사용하였다. 2018년 6월의 기준으로 한국 시장 구글플레이 스토어의 5389개 앱을 대상으로 분석을 진행하였다.

분석된 변수는 사용자가 평가한 앱의 평가점수(average score), 설치사용자수(install user), 월간활성사용자수(MAU), 일간활성사용자수(DAU), 신규설치자수(new installs), 소지율(own rate), 월간활성사용자율(MAU율=추정MAU/추정설치사용자수 *100), 일간활성사용자율(DAU율=추정 DAU/추정설치사용자수 *100), 설치율(rate install: 추정 신규 인스톨 수/추정설치사용자수 *100), 일일평균실행횟수(average activities), 고착도(DAU/MAU)이다.

이 연구에서는 연구문제와 관련하여 앱 스토어의 평균평점(average score)이 다른 실적변수와의 관계에서 목표변수로서 사용할 수 있을지와 다른 변수를 추정할 때 중요 설명변수로 작동할 수 있는지의 역할에 주목하며, 추가적으로 일간활성사용자수(DAU), 일간활성사용자율(rate DAU), 월간활성사용자수(MAU), 월간활성사용자율(rate MAU), 평균실행횟수(average activities), 고착도(DAU/MAU)를 관리지표로 사용할 수 있을 지에 관하여 탐색한다.

일간활성사용자(DAU)는 한 달 평균 수치로 계산하는 데 앱 출시 초기에 중요한 지표이며 메신저나 게임과 같은 반복적 사용이 중요한 앱의 경우에 중요 지표가 된다. 일간활성사용자(DAU)는 서비스 유지 보수(service maintenance)의 관점에서도 서버의 트래픽을 예측하고 대비할 수 있는 지표이다. 월간활성사용자(MAU)는 사용자의 규모와 앱의 성장세를 측정할 수 있는 지표로서 중요하게 사용된다. 고착도는 일간활성사용자수(DAU)를 월간활성화사용자(MAU)로 나누는 것으로 계산되며(DAU/MAU), 보통 해당 앱을 계속 사용하게 될 사용자의 크기를 의미하는 것으로 인식된다. 지속 사용자로 전환된 전환을 의미하는 지표로 고착도(stickiness)로 표현한다. 이 지표는 서비스의 활성화를 가능하는 지표이고, 고착도가 낮으면 서비스 품질이 낮다는 것을 암시한다. 고착도는 그 수치 자체도 중요하지만, 시간이 지남에 따라 수치의 증감이 중요할 수 있다.

앱 리뷰 데이터는 구글플레이 스토어를 대상으로 파이썬을 사용하여 크롤링하며 LIWC 감성사전을 사용하여 감성 점수를 도출한다.

2. 상관관계분석

상관분석은 SPSS, R, 파이썬에서 모두 분석이 가능하다. 상관분석은 두 변수 간에 선형적 상관관계가 있는지를 조사하기 위한 통계적 기법이며 가설검정보다는 투입되는 변수 간의 관계를 탐색적으로 조사하기 위하여 사용된다. 0~±1 사이의 값을 가지며 1에 가까울수록 높은 상관관계를 나타낸다.

3. 기계학습

이 연구에서는 회귀 추정을 위해 사용할 방법론으로 머신러닝 기법을 사용하였고 R과 파이썬을 사용하였다. 기본적으로 회귀를 이용한 추정을 위해서 랜덤 포레스트(random forest) 회귀 추정을 사용하였고 평균 제곱근편차(Root mean Square Deviation:RMSE)와 분산점수(variance score)를 평가지표로 사용하였다.

RMSE는 오류의 제곱 값에 루트를 씌워서 계산한다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2} \quad (1)$$

파이썬 사이킷런에서 제공되는 variance score는 주어진 데이터에서 오류분산을 의미하며 1에 가까울수록 좋은 값이다.

$$variance\ score(y, \hat{y}) = 1 - (Var(y - \hat{y})) / (Var(y)) \quad (2)$$

기계학습은 파이썬을 사용하여 수행하였다.

표 2. 변수 설명

	변수 명	변수 설명
1	평균평점(average score)	구글 플레이스토어 리뷰의 평균 평점
2	설치자수(install user)	월말에 해당 앱을 모바일 단말기에 설치중인 사용자 수
3	추정 MAU (monthly active user)	한달에 한 번 이상 앱을 실행한 이용자수
4	추정 DAU (daily active user)	하루에 한 번 이상 앱을 실행한 이용자수
5	신규 설치자수 (new installs)	데이터 기준 월의 신규 인스톨 수

6	소지율 (own rate)	전체 패널 중 해당 앱을 소지하고 있는 사용자의 비율
7	월간활성사용자율 (rate MAU)	추정 MAU/추정 설치 사용자 수 *100
8	일간활성사용자율 (rate DAU)	추정 DAU/추정 설치 사용자 수 *100
9	설치율 (rate install)	추정 신규 installs /추정 설치 사용자 수*100
10	평균 실행횟수 (average activities)	데이터 기준 월의 한 달간 해당 앱이 실행된 평균 횟수
11	DAU/MAU	추정 DAU를 추정 MAU로 나눈 수(1에 가까울수록 좋은 실적으로 간주)

4. 텍스트마이닝과 감성분석

텍스트 마이닝은 크게는 주제 분석, 클러스터링, 감성 분석으로 나누어질 수 있으며 정성적 텍스트 데이터를 사용하여 정량적 분석을 진행하며 정성적 해석으로 보완하는 방법론이다. 그중에서 감성분석은 긍정, 부정 등 특정 감성 극에 해당하는 점수를 도출하기 때문에 정성적 해석의 한계를 보완하는 정량적 해석이 가능한 방법론이다.

분석을 진행하기 전 단계로서 텍스트 수집이 이루어져야 하는데 텍스트 추출 방법은 앱 리뷰 크롤러를 제작하여 사용한다. 파이썬을 사용하여 앱 크롤러를 제작하며 requests, selenium, BeautifulSoup 모듈을 사용한다. 한국 구글플레이 마켓의 각 모바일 앱에 대한 리뷰를 크롤러를 사용하여 수집한다. 감성분석은 자체적으로 해당 텍스트의 성격을 반영하는 자체 사전을 만들어 진행할 수 있으나 그 경우 감성의 극성을 긍정, 부정, 중립의 단순한 지표만을 사용할 수 밖에 없다. 이런 한계를 극복하기 위해서 객관성이 확보된 영문 감성분석 사전으로서 LIWC 2015를 사용한다. LIWC 감성사전은 기본적으로 영문을 분석하는 도구로서 이 연구에서는 한국 시장 구글 플레이 스토어 앱의 한글 리뷰를 번역기에 의해 영어로 번역한 후 분석을 시도한다[49].

LIWC 2015는 긍정, 부정 감성뿐만 아니라 다양한 감성 점수를 도출할 수 있도록 구성되어 있다. LIWC 감성사전은 아래 [표 3]이 보여주는 것처럼 긍정 부정, 사회적 관계, 권력(power), 위험(risk)에 대한 태도, 시간적 관점의 점수(time-orientation), 개인적 관심사로서의 돈(money), 종교(religion), 등 총 40개 이상의 변수를 텍스트로부터 생성한다.

표 3. LIWC 2015 감성변수

Positive emotion, Negative emotion(anxiety, anger, sadness), Social processes(family, friends, female references, male references), Cognitive processes(insight, causation, discrepancy, tentative, certainty, differentiation), Perceptual processes(see, hear, feel), Biological processes(body, health, sexual, ingestion), Drives (affiliation, achievement, power, reward, risk), Time orientation(past focus, present focus, future focus), Relativity(motion, space, time), Personal Concerns(work, leisure, home, money, religion, death, informal language, swear words, netspeak, non-fluency, fillers)

5. 다변량 시계열 분석

다양한 앱 실적 시계열 자료와 기간별 앱 리뷰를 감성점수로 도출한 감성점수의 시계열 자료 간에 선행 후행 관계와 장단기 관련성(relationship)을 분석한다. 하나의 시계열 자료가 아닌 복수의 시계열 변수이므로 다변량 시계열 분석을 실행한다. 벡터자기회귀모형(Vector Autoregressive Model: VAR)이나 선형결합이 존재하는 벡터오차수정모형(Vector Error Correction model: VEC)을 검토하여 공적분이 존재하면 벡터오차수정모형(VEC)을 선택하고 그렇지 않으면 벡터자기회귀모형(VAR)모형을 사용한다. 공적분이 존재한다는 것은 장기적 선형결합이 존재함을 의미하며 두 변수의 추세가 장기적으로 유사함을 의미한다.

이 연구에서는 두 변수가 장기적 추세를 가지지 않고 있고, 일일 실적변수 변수인 설치(install)가 없거나 리뷰가 없는 구간이 존재하기 때문에 벡터오차수정모형(VEC) 모형의 사용은 불가하므로 벡터자기회귀모형(VAR) 모형을 사용하여 장단기선행 후행 관계를 분석한다. 벡터자기회귀모형(VAR)은 경제이론이 뒷받침되지 않는 경우의 탐색적 연구에서 의미를 도출할 때 유용하다. VAR 모형은 회귀분석의 특징과 시계열 분석의 특징을 결합한 모형으로써 변수 간의 상관관계와 인과 관계를 추정할 수 있는 모형이다. 벡터자기회귀모형(VAR)을 이용하여 몇 시차 전에 변수가 영향을 주는지를 검증하며 원인 결과검증을 위해 그랜저 테스트(Granger test)를 사용하여 원인변수와 결과변수를 확인한다. 마지막으로 충격반응 분석을 통해서 시간의 흐름에 따른 동태적 반응을 분석한다. 벡터자기회귀모형(VAR)의 수행은 STATA 통계 프로그램을 사용하였다.

IV. 연구 결과

1. 연구 1: 앱 실적변수 간 상관분석

한국의 구글플레이 스토어의 2018년 6월 기준 5389개의 구글플레이 스토어 앱을 대상으로 분석하였다. 평균 평점(average score)은 5점 만점의 리커트(likert) 척도로 측정되었으며 다른 10개의 변수와 달리 사용자가 앱 스토어에서 정성적으로 평가한 수치라는 면에서 앱 실적 10개의 변수와 다른 특성을 가진다. 앱 소지율(own rate), 월간활성사용자율(rate MAU), 일간활성사용자율(rate DAU), 설치율(rate install), 고착도(DAU/MAU)는 백분율로 표시된다. 이 연구에서 첫 번째 목표변수로 주목한 변수는 연구문제 1에서 제기한 소비자의 앱 평균평점(average score)이고, 두 번째로 일간활성사용자(DAU), 일간활성사용자율(rate DAU), 월간활성사용자(MAU), 월간활성사용자율(rate MAU), 일일평균실행횟수(average activities), 고착도(DAU/MAU) 변수 간의 관계와 추정 가능한 목표변수로서 기능 여부이다.

피어슨 상관분석의 결과 연구문제1에서 제기한 소비자의 평가로 나타난 평균평점과 다른 실적변수와의 관계는 상관관계가 높지 못하였다. 다른 실적변수 간 관계는 일간활성사용자(DAU)와 월간활성사용자(MAU) 간의 상관관계가 높으며 일간활성사용자(DAU)가 거의 모든 변수와 높은 상관관계에 있음을 보인다. 월간활성사용자(MAU)는 일간활성사용자(DAU)와 소지율(own

rate)과 상관관계가 높았다. 신규 설치자 수(new installs)는 기존의 설치자 수와 정(+)의 상관관계를 보였으며 월간활성사용자(MAU), 일간활성사용자(DAU)와도 정(+)의 상관관계를 보인다. 월간활성사용자(MAU)와 일간활성사용자(DAU), 신규설치자수(new installs), 설치율(rate install)이 타 변수와 긍정의 상관관계가 비교적 높았으며 사용자의 전환을 의미하는 고착도(DAU/MAU)는 일간활성사용자율(rate DAU)과 높은 정(+)의 상관관계를 나타냈다.

2. 연구 2: 앱 실적변수의 예측(기계학습 기법에 따른 목표변수 예측)

기계학습을 사용하여 첫째는 앱 스토어의 평균 평점을 목표변수로 설정하였을 때 다른 앱 실적 지표가 평균평점을 잘 예측하는지, 둘째는 다른 앱 실적 지표를 목표변수로 사용할 때 평균평점이 중요 설명변수로 사용되는지를 탐색하였다. 셋째는 앞에서 제시한 앱 실적 지표 중 하나의 변수를 목표변수로 설정하고 나머지 변수를 설명변수로 사용하여 설명력이 높은 모델은 어떤 앱 실적 지표가 목표변수가 되는 경우인지를 탐구하였다. 실무적으로 현업에서 관리의 중요성이 제기된 앱 실적 지표를 하나씩 목표변수로 간주하여 모델링을 시도하였다.

회귀에 의한 추정은 랜덤포레스트(Random forest) 회귀 추정을 이용하여 RMSE와 Variance score를 도출하였다. RMSE는 변수의 측정 단위에 영향을 받아 변

표 4. 앱 실적 변수 간 상관관계 Pearson Correlations

	average score	install user	MAU	DAU	new installs	own rate	rate MAU	rate DAU	rate install	average activities	DAU/MAU
average score	1	-.044**	-.025	.000	-.026	-.044**	.082**	.132**	-.025	.036**	.114**
Install user	-.044**	1	.838**	.597**	.425**	1.000**	-.051**	-.079**	.422**	.002	-.122**
MAU	-.025	.838**	1	.881**	.495**	.838**	.071**	.010	.489**	.022	-.032*
DAU	.000	.597**	.881**	1	.377**	.597**	.091**	.079**	.371**	.042**	.056**
new installs	-.026	.425**	.495**	.377**	1	.425**	.092**	-.023	.993**	.038**	-.105**
own rate	-.044**	1.000**	.838**	.597**	.425**	1	-.051**	-.079**	.422**	.002	-.122**
rate MAU	.082**	-.051**	.071**	.091**	.092**	-.051**	1	.695**	.094**	.148**	.290**
rate DAU	.132**	-.079**	.010	.079**	-.023	-.079**	.695**	1	-.024	.193**	.630**
rate install	-.025	.422**	.489**	.371**	.993**	.422**	.094**	-.024	1	.038**	-.109**
average activities	.036**	.002	.022	.042**	.038**	.002	.148**	.193**	.038**	1	.163**
DAU/MAU	.114**	-.122**	-.032*	.056**	-.105**	-.122**	.290**	.630**	-.109**	.163**	1

* 0.05수준(양측)에서 유의, ** 0.01수준(양측)에서 유의

수별로 크기를 비교하기는 어려우며 variance score를 사용하여 설명력을 검토하였다. 전체 앱을 대상으로 평균평점(average score), 평균실행횟수(average activities), 일간활성사용자(DAU), 일간활성사용자율(rate DAU), 월간활성사용자(MAU), 월간활성사용자율(rate MAU), 설치율(rate install), 고착도(DAU/MAU)를 목표변수로 사용하여 안정적인 예측모델링이 가능한지 검증하였다.

평균평점(average score)과 평균실행횟수(average activities)는 variance score가 0.005와 0.017로 매우 낮게 도출되어 목표변수로 적절치 않은 것으로 나타났다. 특히 평균평점(average score)은 다른 변수를 추정하는 데도 중요한 변수로 사용되지 않은 것으로 나타났다. 일간활성사용자율(rate DAU)의 variance score가 74%이며 나머지 일간활성사용자(DAU), 월간활성사용자(MAU), 월간활성사용자율(rate MAU), 고착도(DAU/MAU), 설치율(rate install)의 variance score는 90% 이상으로 분석되었다. 변수 중요도는 목표변수를 월간활성사용자율(rate MAU), 일간활성사용자율(rate DAU), 설치율(rate install), 고착도(DAU/MAU)인 경우를 분석하였다. 월간활성사용자율(rate MAU)과 일간활성사용자율(rate DAU)는 각각을 추정 시 상호 중요한 변수로 작용하고 있으며, 일간활성사용자율(rate DAU)를 추정하는 경우 평균실행횟수(activities)가 일정부분 영향을 미친다는 점이 주목할 만하다. 전반적으로 설치 사용자수(install users)가 추정 모델링 시 중요변수로 등장하고 있는 점을 고려한다면 기업에서는 일단 앱을 다운로드 받아 설치하는 것을 1차 목표변수로 사용하고, 2차적으로 설치된 앱을 자주 활성화시키는 것을 주요 관리 지표화 할 수 있음을 보여준다.

표 5. 목표변수 별 모델 성능 평가

목표 변수	RMSE	Variance Score
average score	0.643	0.005
average activities	16.045	0.017
DAU	271200	0.942
rate DAU	0.094	0.735
MAU	448552.710	0.930
rate MAU	0.152	0.916
DAU/MAU	0.095	0.956
rate install	0.000	0.971

3. 연구 3: 앱 실적변수와 감성점수를 통합한 다변량 시계열 분석

연구 3을 위해서는 한 가지 앱을 선정하여 사례 분석을 시행하였다. 여러 앱 중에서 앱 리뷰의 길이를 고려하였는데, 모바일 단말기 화면에 1~3페이지에 걸쳐서 일일 리뷰가 누적되는 앱을 선정하였고 이 연구에서는 네이버 TV 앱을 선정하였다. 감성점수의 도출은 LIWC2015 감성사전을 사용하였으며, 총 감성변수 중 앱 실적에 영향을 줄 수 있다고 판단한 변수를 선정하였다. 사용된 변수는 긍정을 의미하는 Posemo, affiliation, achieve, power, reward와 부정을 의미하는 negemo, anx, anger, sad, swear, netspeak, assent 변수이며, 긍정변수의 합으로 이 연구에서 새로 생성한 sum_positive와 부정적 의미의 변수의 합으로 만들어진 sum_negative를 추가로 생성하여 사용하였다. 이 연구는 다변량 시계열 분석을 사용하여 변수 간 선후행 관계를 분석하였다. 변수가 복수이므로 다변량 시계열 분석을 사용하고 두 시계열 간에 장기적 추세가 존재하지 않는 경우에 해당되어 VAR를 사용하여 분석하였다.

앱 리뷰가 앱 실적에 영향을 줄 것이라는 가정은 기본적으로 앱을 내려받으려고 시도할 때, 즉 리뷰를 읽고 사용행태를 변화시키거나 하고자 하는 행동을 강화할 것이라는 사고에 기반한다. 리뷰가 읽을 수 있는 범위 안에 놓여야 하므로 기간의 측면에서 앱 실적과 앱 리뷰를 기반으로 도출된 감성점수와와의 관계는 일별 분석이 적합하다. 기본적으로 앱의 인기도나 성장 상황에 따라 차이점이 존재할 수 있지만, 앱 리뷰가 사용자의 단말기에 보이는 영역은 첫 화면이나 2번째 3번째 화면일 것이며 5페이지를 넘어가지 않는 것이 기본적인다. 인기가 많은 앱은 5번째 페이지가 넘어가기까지 전부 1일 차 리뷰일 것이며 너무 인기가 많아서 일일 앱의 리뷰 숫자가 수 천 개가 넘어갈 때에는 사용자가 그 모두를 읽는 것은 한계가 있고 일별 앱 리뷰를 정량화하는데도 문제가 발생한다. 보통의 앱이라면 대략 1일에서 5일 이내의 리뷰를 읽는 경우가 많을 것이며 분석을 위해서 지나치게 앱 리뷰가 많지 않은 앱이 분석에 적합하다. 앱 실적 중 분석의 대상으로서 가장 적합한 실적은 앱 스토어에서 리뷰를 읽고 앱을 내려받는 상황

표 6. Feature importance by Random Forest

Random Forest labeling	R ²	Important feature 1	Important feature 2	Important feature 3	Important feature 4	Important feature 5
rate MAU	94%	rate_DAU	DM	install User	MAU	DAU
rate DAU	74%	rate MAU	DM	activities	install User	DAU
rate Install	97%	new installs	-	-	-	-
DAU/MAU	96%	rate_DAU	MAU	DAU	rate MAU	install user

을 가정한다면, 감성점수와와의 관계가 가장 의미 있는 앱 실적 지표는 다운로드 즉 앱 설치와 관련된 변수이다. 따라서 이 연구에서는 앱 감성 리뷰의 일별 점수와 앱 일별 설치 실적을 비교하여 분석하였다. 첫 번째 일일 설치회수(install)와 감성 변수 중 achieve와의 관계를 아래의 4단계의 순서로 분석을 진행하고 같은 방법으로 일일 설치회수와 다른 감성 변수와 관계를 탐색하였다. 1단계는 변수 간의 최적 시차를 검증하며 AIC, HPIC, SBIC에 의해서 결정한다. installs와 achieve는 1 시차가 최적 시차임을 보여준다[표 7]. 2단계 분석을 통하여 1시차 전 achieve 변수가 installs에 통계적으로 유의한 영향을 주고 1시차전 installs는 achieve에 통계적으로 유의한 영향을 주지 못함을 분석하였다[표 8].

3단계는 인과관계 검증, 정규성, 자기상관성을 검증한다. achieve 감성점수가 원인이고 installs가 결과로 검증이 되었으며[표 9], 자기상관이 없는 것으로 검증이

되었으나[표 10], Jarque-Bera test에 의한 정규성은 만족하지 못하는 것으로 나타났다[표 11]. 벡터오차수정모형(VEC)의 경우 정규성이 매우 중요한 가정이 되지만 벡터자기회귀모형(VAR)의 경우 정규성을 충족하지 못하더라도 일정부분 의미 있는 결과 도출이 가능하다고 보기 때문에 변수간의 관계가 의미가 있는 것으로 판단하였다.

표 9. Granger causality Wald tests

Equation	Excluded	chi 2	df	prob > chi ²
installs	achieve	11.151	1	0.001
installs	ALL	11.151	1	0.001
achieve	installs	1.1672	1	0.280
achieve	ALL	1.1672	1	0.280

표 10. Autocorrelation test(Lagrange-multiplier test

lag	chi 2	df	prob > chi 2
1	2.1673	4	0.7050

H0: no autocorrelation at lag order

표 7. Selection-order criteria

Lag	LL	LR	df	P	FPE	AIC	HQIC	SBIC
0	-8099.41				2.8 e + 11	32.0214	32.0279	32.0381
1	-8079.21	40.412*	4	0.000	2.6 e + 11*	31.9573*	31.9770*	32.0075*
2	-8077.14	4.1304	4	0.389	2.6 e + 11	31.9650	31.9977	32.0485
3	-8076.14	2.0124	4	0.733	2.6 e + 11	31.9768	32.0227	32.0938
4	-8074.76	2.7520	4	0.600	2.7 e + 11	31.9872	32.0462	32.1375
5	-8072.39	2.7435	4	0.602	2.7 e + 11	31.9976	32.0697	32.1813
6	-8072.34	2.1024	4	0.717	2.7 e + 11	32.0092	32.0092	32.2264
7	-8071.80	1.0812	4	0.897	2.8 e + 11	32.0229	32.1212	32.2735
8	-8070.56	2.4715	4	0.650	2.8 e + 11	32.0338	32.1452	32.3178

표 8. Short-term relationship between installs and achieve

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
installs					
installs L1.	0.6240	0.0410	1.52	0.128	-0.0179 0.1427
achieve L1.	112.3647	33.6484	3.34	0.001	46.4148 178.3145
_cons	11764.21	1072.008	10.97	0.000	9663.1170 13865.31
achieve					
installs L1.	-0.0005	0.0004	-1.08	0.280	-0.0014 0.0000
achieve L1.	0.2471	0.0401	6.16	0.000	0.1684 0.3257
_cons	10.0147	1.2789	7.83	0.000	7.5080 12.5214

표 11. Jarque-Bera test

Equation	chi 2	df	prob > chi 2
installs	7.3 e + 05	2	0.00000
achieve	3933.014	2	0.00000
ALL	7.4 e +0.5	4	0.00000

4단계는 충격효과분석으로서 그림에서 보는 바와 같이 감성점수 achieve는 installs에 향 후 1 시차까지 강한 영향을 주고 다음 시차부터는 영향이 급속히 감소하는 것으로 나타났다(그림 1). 이상의 분석 프로세스에 따라서 앱 실적지표 중 일일 install 숫자의 시계열 자료와 다른 감성점수와 관계를 추가적으로 분석하였다. 긍정을 의미하는 감성점수인 posemo, affiliation, achieve, power, reward와 부정을 의미하는 negemo, anx, anger, sad, swear, netspeak, assent, 긍정변수의 합으로 새로 도출한 sum_positive와 부정변수의 합으로 만들어진 sum_negative가 분석의 대상이 되었다. 의미 있는 결과를 도출한 내용을 변수별로 정리하면 아래와 같다. 아쉽게도 정규성을 만족하는 관계는 없었으나 벡터오차수정모형(VEC)에 비하여 벡터자기회귀모형(VAR)이 정규성에 엄격하지 않은 점을 고려하여 그랜저 인과검정(Granger test)을 통과하고 자기상관되어 있지 않은 경우를 의미 있는 결과로 정리하였다. Install과 1시차 전 achieve감성점수, install과 1시차전 reward 감성점수가 install에 영향을 주는 것으로 나타났다(표 12).

achieve는 win, success, better와 유사한 의미의 단어 조합에 의해서, reward는 take, prize, benefit 등의 단어 조합에 의해서 점수로 계산되었다. 이러한 단어의 조합이 많이 나올수록 1시차 후에 앱 설치에 긍정적 영향을 주는 것으로 해석할 수 있다.

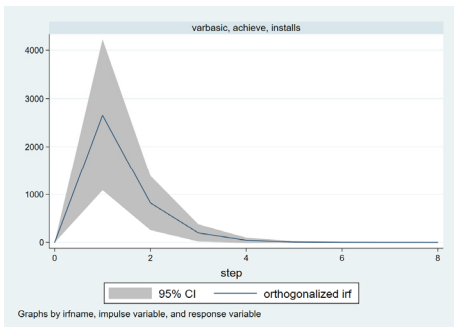


그림 1. 충격분산효과

표 12. 일별 앱 실적변수와 감성 변수의 다변량 시계열 분석 결과

구분	사례1	사례2
앱 실적 변수	install	install
감성변수	achieve	reward
시차	1시차 전	1시차 전
Granger test	○	○
자기상관 Test	○	○
정규성	X	X

V. 토론과 결론

1. 실무적 시사점

이 연구는 앱 이용실적 변수와 앱 리뷰 감성분석을 통해 도출된 변수를 사용하여 실적 변수 간 관계 및 감성점수와의 관계를 분석하였다.

사용자 평균평점, 설치 사용자 수, 월간, 일간 활성 사용자, 신규 설치 사용자 수, 평균실행횟수 등은 앱을 개발하고 앱 스토어에서 서비스를 시작한 후 기업이 서비스의 성장을 위해 반드시 관리해야 할 변수이다. 변수 간 관계 및 감성점수와의 관계는 앱과 서비스의 성장과 마케팅의 전략 방향을 수립하기 위해 반드시 참조해야 하는 변수이다. 지표의 하락과 상승, 특별히 선행지표의 하락과 상승이 제시하는 서비스에 대한 지표들은 언제 마케팅 프로모션이 필요하고 언제 위기 대응이 필요한지를 제시함으로써 성장과 위기 관리의 가이드 라인을 제공한다. 이 연구를 통해 제시된 관리 지표 로서의 변수들은 우선적 관리 목표로서 KPI(key performance index)를 의미한다.

앱 실적의 횡단면 자료를 사용하여 만들어진 실적 변수 예측 모델은 어떤 머신러닝 모델과 변수가 관리되어야 할지를 판단하는데 경험적 기준을 제시하며 고객의 전환을 의미하는 지표로 사용할 수 있는 고착도(DAU/MAU) 지표 등은 앞으로 중요한 관리지표로서 기능할 것이다.

다변량 시계열 자료 분석을 통해서 도출된 감성분석의 선행지표 가능성은 월별 분석보다는 일별 분석이 의미 있는 것으로 발견되었으며 특정 분야 앱을 통해 도출된 결과의 한계를 고려하더라도 선행지표로써 감성점수의 중요성과 앱에 대한 고객 평가로서의 리뷰 관리 중요성을 실무자들에게 제시하고 이를 기반으로 한 마케팅 계획을 수립할 수 있게 한다.

2. 이론적 방법론적 시사점

이 연구는 앱 실적과 앱 리뷰를 사용하여 데이터의 수집부터 전처리, 분석과정까지 데이터분석 전반에 걸친 프로세스를 진행하였다. 데이터 생애주기에 걸쳐서 데이터의 형태에 적합한 다양한 방법론을 제시하고 텍스트 형태의 비정형데이터인 리뷰 데이터를 감성점수의 형태로 정형 데이터 화하여 기존의 데이터와 통합하여 분석하는 방법론을 제시하고 활용 가능성을 제시하였다.

데이터분석에 있어서 횡단면 자료와 시계열 자료를 통합적으로 사용하였으며 머신러닝 기법과 다변량 시계열 방법론(VAR)을 사용하여 그들 변수 간의 시차 선행 여부를 밝히는 데 집중하였다. 이 연구는 방법론적으로 횡단면 자료와 시계열 자료를 종합적으로 사용하였고 머신러닝 기법과 시계열 기법을 통합하여 사용함으로써 문제를 해결하기 위해서 종합적인 방법론의 사용이 필요하고 필수적임을 제시하였다.

방법론을 구현하는 사용 도구에 있어서도 SPSS, R, STATA, 파이썬의 사용 예를 실질적인 문제 해결에 사용하는 모습을 제시함으로써 데이터분석 교육의 실질 사례로 기능할 수 있도록 하였다.

프로그래밍을 기본 전공과목으로 하는 학생 뿐만 아니라 프로그래밍 비전공자를 대상으로 각각의 위치와 미래 진로에 적합한 데이터분석 교육의 중요성이 강조되고 있다. 이 연구는 다음의 몇 가지 측면에서 교육 커리큘럼의 구성과 활용 측면에서 의미를 가진다. 첫째는 데이터의 형태와 수집의 관점에서 정형 데이터와 비정형데이터의 수집, 비정형데이터의 정형화, 정형 데이터와 비정형데이터 통합의 과정을 제시한다. 둘째, 횡단면 데이터의 분석과 기계학습의 활용 예를 제시한다. 이 연구에서는 회귀를 이용한 추정을 위해서, 랜덤 포레스트 회귀추정 방법론을 사용하였으며, 이런 방법론을 사용한 분석은 데이터분석 방법론으로 교육안으로 기능할 수 있다. 셋째 시계열 방법론으로 사용된 다변량 시계열 방법론(VAR)을 시계열 자료 분석에 관한 기본적인 사례 접근의 교육안으로 활용할 수 있다. 마지막으로 SPSS, R, STATA에 대한 툴 사용 방법과 파이썬을 활용한 데이터 크롤링 방법론은 데이터분석에 필요한 프로그래밍과 툴 교육에 필수적인 커리큘럼으로 활용

할 수 있다.

3. 연구의 한계와 미래연구의 방향성

위와 같은 기여점에도 불구하고 이 연구는 몇 가지 한계를 가지고 있다. 첫째는 자료수집의 한계로서 연구의 소재로 사용된 데이터가 일정 기간에 국한되어 수집되었다는 시간적인 제약과 자료수집 기관이 제공하는 데이터의 정확성과 완전성의 문제를 가지고 있다는 점이다. 이러한 이유로 앱 전반에 대한 일반화된 결론을 도출하기에는 무리가 있다. 이 연구는 특정 시점의 데이터와 특정 앱을 대상으로 진행되었기 때문에 실무에 그대로 적용하기 위해서는 자료수집의 상시화가 필요하고 실시간 분석을 진행하고 의미를 도출할 필요가 있다.

둘째는 감성사전을 통한 감성점수 도출 방법론의 한계로서 미리 만들어진 감성사전에 의존하여 감성분석을 진행하였기 때문에 감성사전이 대응하지 못하는 신조어와 이중 부정에 의한 긍정의 감성에 대해서 정확한 감성점수를 확보하기 어렵다. 이런 점을 극복하기 위해서는 자연어처리에 대한 추가적인 고민과 협업이 필요하다. 또한 이 연구에서 사용한 감성사전이 영어에 기반하고 있어서 한글 앱 리뷰의 영어 번역의 한계를 가지고 있다. 서로 다른 언어를 하나의 기준언어를 사용하여 텍스트마이닝을 할 때는 번역기를 사용하여 하나의 기준언어로 분석하는 것이 일반적인 방법론이기는 하지만[49], 번역에 의한 감성사전 사용은 한국어의 독특한 의미를 영어로 전환하면서 정확히 전달하지 못하는 단점이 있다. 한국어 감성사전이 부재하여 발생하는 문제점은 앞으로 다양한 극성을 가진 한글 감성점수를 도출하는 후행 연구의 필요성을 제기한다.

이와 같은 한계에도 불구하고 기존의 연구와 비교하여 이 연구는 실무적 시사점을 도출하기 위하여 수집 가능한 앱 실적 데이터를 횡단면과 시계열 기준으로 최대한 수집하였으며 사용 가능한 다양한 방법론을 동원하여 통합적으로 접근하였다면 점에서 실질적인 현업의 문제를 해결하는 연구로서 의미가 있다.

앱을 개발하고 서비스하는 기업에 앱 리뷰는 고객의 살아 있는 목소리이며 요구(needs)라는 측면에서 매우 소중한 마케팅 자원이다. 기업은 앱 리뷰를 이용하여

앱의 불만족에 대한 개선점을 찾아내고 이것을 반영하여 소비자의 니즈를 만족하는 앱을 개발 및 서비스할 수 있다. 기업은 앱 실적 간의 관계를 분석하여 주요 지표를 예측하고, 앱 리뷰의 내용을 분석하여 앱 리뷰와 관련된 앱 자체의 실적 변수를 예측하고 관리함으로써 앱의 업데이트와 필수적인 마케팅 활동을 수행할 수 있게 된다. 관리되고 예측되는 주요 지표는 기업의 적시 개발과 마케팅의 필수적인 활동을 가능하게 하는 기업 생존의 중요한 나침반이다.

참고 문헌

- [1] 이희서, 광나연, 이증정, “스마트폰 앱 사용 및 추천의 도 영향 요인에 관한 연구,” 한국콘텐츠학회논문지, Vol.15, No.8, pp.481-494, 2015.
- [2] K. Gi Jin, B. Gwang In, and Y. Jung Mee, “A Study on the Effect of Easiness to Use the Food Service Related Application of Smart Phone on Intentions of Use: Focused on the Mediation Effect of Familiarity and Usefulness,” (in 한국어), Korean Journal of Hospitality & Tourism, Vol.20, No.6, pp.61-81, 2011(12).
- [3] 유로, 이종호, “모바일 메신저 앱의 이용 동기와 품질요인이 몰입과 사용자의 만족도와 충성도에 미치는 영향,” e-비즈니스연구, Vol.15, No.3, pp.255-276, 2014.
- [4] 김광국, 김용환, 김자희, “사용자 리뷰 토픽분석을 활용한 모바일 쇼핑 앱 고객만족도에 관한 연구,” 한국전자거래학회지, Vol.23, No.4, pp.41-62, 2018.
- [5] 조혁준, 강주영, “한국과 미국 간 모바일 앱 리뷰의 감성 차이에 관한 탐색적 비교 분석,” 한국 IT 서비스학회 학술대회 논문집, pp.105-109, 2016.
- [6] 이동일, 최승훈, “앱스토어에서 소비자 리뷰와 전문가 리뷰가 앱 개발사의 성과에 미치는 영향,” JOURNAL OF KOREAN MARKETING ASSOCIATION, Vol.27, No.2, pp.113-136, 2012.
- [7] L. Ravindranath, J. Padhye, S. Agarwal, R. Mahajan, I. Obermiller, and S. Shayandeh, “Appinsight: Mobile app performance monitoring in the wild,” in Proceedings of the 10th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2012, pp.107-120, 2012.
- [8] Heo, J. Seo, Cho, H. Chang, Kim, and Y. Ju, “Advertising Implications of Beacon Technology - Focusing on TAM to Predict User Acceptance of Beacon Application,” The Korean Journal of Advertising and Public Relations, Vol.17, No.3, pp.98-137, 2015(7).
- [9] P. Tak and S. Panwar, “Using UTAUT 2 model to predict mobile app based shopping: evidences from India,” Journal of Indian Business Research, Article, Vol.9, No.3, pp.248-264, 2017.
- [10] S. Yuan, W. Ma, S. Kanthawala, and W. Peng, “Keep Using My Health Apps: Discover Users' Perception of Health and Fitness Apps with the UTAUT2 Model,” Telemedicine and e-Health, Article, Vol.21, No.9, pp.735-741, 2015.
- [11] C. H. Hsiao, J. J. Chang, and K. Y. Tang, “Exploring the influential factors in continuance usage of mobile social Apps: Satisfaction, habit, and customer value perspectives,” Telematics and Informatics, Article, Vol.33, No.2, pp.342-355, 2016, Art. No.730.
- [12] N. Valaei, S. R. Nikhashemi, G. Bressolles, and H. H. Jin, “A(n) (a)symmetric perspective towards task-technology-performance fit in mobile app industry,” Journal of Enterprise Information Management, Article Vol.32, No.5, pp.887-912, 2019.
- [13] J. K. Carmody, L. A. Denson, and K. A. Hommel, “Content and usability evaluation of medication adherence mobile applications for use in pediatrics,” Journal of Pediatric Psychology, Article Vol.44, No.3, pp.333-342, 2019.
- [14] D. Lin, C. P. Bezemer, Y. Zou, and A. E. Hassan, “An empirical study of game reviews on the Steam platform,” Empirical Software Engineering, Article Vol.24, No.1, pp.170-207, 2019.
- [15] T. Alessa, S. Abdi, M. S. Hawley, and L. D.

- Witte, "Mobile apps to support the self-management of hypertension: Systematic review of effectiveness, usability, and user satisfaction," *Journal of Medical Internet Research*, Review Vol.20, No.7, 2018, Art. no. e10723.
- [16] C. Y. Lee, C. H. Tsao, and W. C. Chang, "The relationship between attitude toward using and customer satisfaction with mobile application services: An empirical study from the life insurance industry," *Journal of Enterprise Information Management*, Article Vol.28, No.5, pp.680-697, 2015.
- [17] P. Krebs and D. T. Duncan, "Health app use among US mobile phone owners: A national survey," *JMIR mHealth and uHealth*, Review Vol.3, No.4, 2015, Art. no. e101.
- [18] J. A. Cafazzo, M. Casselman, N. Hamming, D. K. Katzman, and M. R. Palmert, "Design of an mHealth app for the self-management of adolescent type 1 diabetes: A pilot study," *Journal of Medical Internet Research*, Conference Paper Vol.14, No.3, 2012.
- [19] A. E. Roberts, T. A. Davenport, T. Wong, H. W. Moon, I. B. Hickie, and H. M. LaMonica, "Evaluating the quality and safety of health-related apps and e-tools: Adapting the Mobile App Rating Scale and developing a quality assurance protocol," *Internet Interventions*, Article Vol.24, 2021, Art. no. 100379.
- [20] A. Salazar, H. de Sola, I. Failde, and J. A. Moral-Munoz, "Measuring the quality of mobile apps for the management of pain: Systematic search and evaluation using the mobile app rating scale," *JMIR mHealth and uHealth*, Article Vol.6, No.10, 2018, Art. no. e10718.
- [21] F. Mandracchia, E. Llauradó, L. Tarro, R. M. Valls, and R. Solà, "Mobile phone apps for food allergies or intolerances in app stores: Systematic search and quality assessment using the mobile app rating scale (MARS)," *JMIR mHealth and uHealth*, Article Vol.8, No.9, 2020, Art. no. e18339.
- [22] A. Vianello, L. Chittaro, S. Burigat, and R. Budai, "MotorBrain: A mobile app for the assessment of users' motor performance in neurology," *Computer Methods and Programs in Biomedicine*, Article Vol.143, pp.35-47, 2017.
- [23] M. K. Schmid, M. A. Thiel, K. Lienhard, R. O. Schlingemann, L. Faes, and L. M. Bachmann, "Reliability and diagnostic performance of a novel mobile app for hyperacuity self-monitoring in patients with age-related macular degeneration," *Eye (Basingstoke)*, Article Vol.33, No.10, pp.1584-1589, 2019.
- [24] R. Onodera and S. Sengoku, "Innovation process of mHealth: An overview of FDA-approved mobile medical applications," *International journal of medical informatics*, Vol.118, pp.65-71, 2018.
- [25] A. K. Yetisen, J. Martinez-Hurtado, F. da Cruz Vasconcellos, M. E. Simsekler, M. S. Akram, and C. R. Lowe, "The regulation of mobile medical applications," *Lab on a Chip*, Vol.14, No.5, pp.833-840, 2014.
- [26] A. J. Barton, "The regulation of mobile health applications," *BMC medicine*, Vol.10, No.1, pp.1-4, 2012.
- [27] M. Willocx, J. Vossaert, and V. Naessens, "Comparing performance parameters of mobile app development strategies," in *Proceedings - International Conference on Mobile Software Engineering and Systems*, MOBILESoft 2016, pp.38-47. 2016.
- [28] P. Roma and M. Vasi, "Diversification and performance in the mobile app market: The role of the platform ecosystem," *Technological Forecasting and Social Change*, Article Vol.147, pp.123-139, 2019.
- [29] I. J. Mojica Ruiz, M. Nagappan, B. Adams, T. Berger, S. Dienst, and A. E. Hassan, "Examining the Rating System Used in Mobile-App Stores," *IEEE Software*, Article Vol.33, No.6, pp.86-92, 2016.
- [30] M. Bardus et al., "The Arabic version of the

- mobile app rating scale: Development and validation study,” *JMIR mHealth and uHealth*, Article Vol.8, No.3, 2020, Art. no. e16956.
- [31] N. Henze, M. Pielot, B. Poppinga, T. Schinke, and S. Boll, “My app is an experiment: Experience from user studies in mobile app stores,” *International Journal of Mobile Human Computer Interaction (IJMHCI)*, Vol.3, No.4, pp.71-91, 2011.
- [32] F. Nayebi, J. M. Desharnais, and A. Abran, “The state of the art of mobile application usability evaluation,” in *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp.1-4, 2012.
- [33] J. Huebner, C. Schmid, M. Bouguerra, and A. Ilic, “Finmars: A mobile app rating scale for finance apps,” in *ACM International Conference Proceeding Series*, pp.6-11, 2019.
- [34] A. Escriche-Escuder et al., “Assessment of the quality of mobile applications (Apps) for management of low back pain using the mobile app rating scale (mars),” *International Journal of Environmental Research and Public Health*, Article Vol.17, No.24, pp.1-16, 2020, Art. no. 9209.
- [35] S. Sigg, E. Lagerspetz, E. Peltonen, P. Nurmi, and S. Tarkoma, “Exploiting usage to predict instantaneous app popularity: Trend filters and retention rates,” *ACM Transactions on the Web (TWEB)*, Vol.13, No.2, pp.1-25, 2019.
- [36] A. Zuniga, H. Flores, P. Hui, J. Manner, and P. Nurmi, “Tortoise or hare? Quantifying the effects of performance on mobile app retention,” in *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, pp.2517-2528, 2019.
- [37] L. Hsu, and J. C. C. Lin, “Effect of perceived value and social influences on mobile app stickiness and in-app purchase intention,” *Technological Forecasting and Social Change*, Vol.108, pp.42-53, 2016.
- [38] K. L. Hsiao and C. C. Chen, “What drives in-app purchase intention for mobile games? An examination of perceived values and loyalty,” *Electronic Commerce Research and Applications*, Vol.16, pp.18-29, 2016.
- [39] L. Dogruel, S. Joeckel, and J. Vitak, “The valuation of privacy premium features for smartphone apps: The influence of defaults and expert recommendations,” *Computers in Human Behavior*, Vol.77, pp.230-239, 2017.
- [40] Y. J. Lee, H. Ghasemkhani, K. Xie, and Y. Tan, “Switching decision, timing, and app performance: An empirical analysis of mobile app developers’ switching behavior between monetization strategies,” *Journal of Business Research*, Article Vol.127, pp.332-345, 2021.
- [41] E. M. Messner et al., “The german version of the mobile app rating scale (MARS-G): Development and validation study,” *JMIR mHealth and uHealth*, Article Vol.8, No.3, 2020, Art. no. e14479.
- [42] B. Y. B. Kim, A. Sharafoddini, N. Tran, E. Y. Wen, and J. Lee, “Consumer mobile apps for potential drug-drug interaction check: Systematic review and content analysis using the mobile app rating scale (MARS),” *JMIR mHealth and uHealth*, Review Vol.6, No.3, 2018, Art. no. e74.
- [43] H. Khalid, E. Shihab, M. Nagappan, and A. E. Hassan, “What do mobile app users complain about?,” *IEEE Software*, Article Vol.32, No.3, pp.70-77, 2015, Art. no. 6762802.
- [44] W. Luiz et al., “A feature-oriented sentiment rating for mobile app reviews,” in *The Web Conference 2018 - Proceedings of the World Wide Web Conference, WWW 2018*, pp.1909-1918, 2018.
- [45] B. Fu, J. Lin, L. Liy, C. Faloutsos, J. Hong, and N. Sadeh, “Why people hate your App - Making sense of user feedback in a mobile app store,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Vol.Part F128815, pp.1276-1284, 2013.
- [46] L. Chandra Sekhar Reddy, D. Murali, and J.

Rajeshwar, "A Review on Mobile App Ranking Review and Rating Fraud Detection in Big Data," in Lecture Notes in Networks and Systems Vol.74, ed, pp.551-556, 2019.

- [47] K. M. Altenburger and D. E. Ho, "Is Yelp actually cleaning up the restaurant industry? A re-analysis on the relative usefulness of consumer reviews," in The World Wide Web Conference, pp.2543-2550, 2019.
- [48] P. A. Salz, "Monitoring mobile app performance," Journal of Direct, Data and Digital Marketing Practice, Article Vol.15, No.3, pp.219-221, 2014.
- [49] D. Spry and T. Dwyer, "Representations of Australia in South Korean online news: a qualitative and quantitative approach utilizing Leximancer and Korean keywords in context," Quality and Quantity, Article in Press pp.1-20, 2016.

김 성 범(Sungbum Kim)

정회원



- 2013년 2월 : KAIST 기술경영(공학박사)
- 2013년 9월 ~ 현재 : 금오공대 IT 융합학과 부교수

〈관심분야〉 : ICT 전략, 기업가정신, 신제품/신사업 개발

저 자 소 개

김 동 욱(Dongwook Kim)

정회원



- 2012년 2월 ~ 2015년 9월 : KT 기술전략담당 전임연구원
- 2015년 10월 ~ 2020년 3월 : GSMA 네트워크 기술전문가
- 2020년 2월 : KAIST 기술경영(공학박사)

〈관심분야〉 : ICT 전략/정책, 기술표준, 에너지 기술 전략/정책