

모바일 헬스 서비스 사용자 특성 분석 및 이탈 예측 모델 개발

한정현^{1)*}, 이주연²⁾

1) 아주대학교, 2) 아주대학교

Mobile health service user characteristics analysis and churn prediction model development

Jeong Hyeon Han¹⁾, Joo Yeoun Lee²⁾

1) *Ajou University*, 2) *Ajou University*

Abstract : As the average life expectancy is rising, the population is aging and the number of chronic diseases is increasing. This has increased the importance of healthy life and health management, and interest in mobile health services is on the rise thanks to the development of ICT(Information and communication technologies) and the smartphone use expansion. In order to meet these interests, many mobile services related to daily health are being launched in the market. Therefore, in this study, the characteristics of users who actually use mobile health services were analyzed and a predictive model applied with machine learning modeling was developed. As a result of the study, we developed a prediction model to which the decision tree and ensemble methods were applied. And it was found that the mobile health service users' continued use can be induced by providing features that require frequent visit, suggesting achievable activity missions, and guiding the sensor connection for user's activity measurement.

Key Words : Healthcare, Health Management, Mobile Health, Machine Learning, Supervised Learning

Received: October 27, 2021 / **Revised:** December 17, 2021 / **Accepted:** December 21, 2021

* 교신저자 : Joo Yeon Lee/Ajou University/jooyeoun325@ajou.ac.kr

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

1. 서론

헬스케어 산업은 환자들의 병을 치료하기 위한 전통적인 의료 중심 영역에 정보통신기술(ICT : Information and Communication Technologies)을 융합하여 새롭게 용이한 방식으로 다양한 수요자들에게 건강 관련 서비스를 제공하는 영역으로 확장되고 있다.[1]

우리나라의 경우 2012년 80.87세의 기대수명이 2018년 82.7세로 증가하였지만 건강수명은 동일한 기간에 65.7세에서 64.4세로 하락 추이를 보이며 건강수명과 기대수명의 차이가 점점 더 커지고 있는 추세이다.[2] 이러한 기대수명과 건강수명의 차이의 확대, 환경의 변화에 따른 생활 습관 관련 질환의 꾸준한 증가, 그리고 이에 따른 사회 내 노령인구 비율의 증가, 국민 건강 관련 사회적 비용의 상승 효과는 결국 얼마나 오래 사는가에서 어떻게 오래 사는가로 인식을 전환시키는 계기가 되었다.[3] 미 질병통제예방센터는 만성질환으로 인한 의료비 지출이 국가의료비 지출 전체의 약 75%에 이르는 것으로 보고한 바 있으며, 더욱이 만성질환자의 문제는 고령자에 그치지 않고 전 연령층에 걸쳐 지속적으로 확대되는 추세이다.[4]

이렇게 평균 수명의 상승에 따른 인구 고령화, 만성질환자 증가, 사회 문화적인 의료비 지출 문제가 증가하는 한편, 생활 수준이 향상하고 행복 추구를 목적으로 하는 소비자 집단이 증가하면서 건강한 삶을 위한 예방 및 건강 관리의 중요성이 증대되었고 ICT의 발전과 스마트폰 보급 확대라는 기술의 혁신은 결국 모바일 헬스에 대한 폭발적인 관심으로 나타나고 있다.[5],[6]

모바일 헬스는 개인의 건강관리 니즈에 대한 충족 및 사회 전반의 공공 의료 관련 사회경제적 비용의 절감뿐만 아니라 나아가 의료서비스 질의 향상과 질병의 예방 등 산업 혁신 및 사회 정책적 효과를 가져오기 때문에 관련 지원 정책이 연구되고 있으며, 이에 대응하는 제품 및 서비스 비즈니스 영역 또한 크게 확산되고 있는 추세이다.[4],[5]

모바일 헬스 산업이 급성장하며 건강 관리에 관심을 가지는 소비자 그룹을 위한 많은 서비스가 시장에 나오고 있어 사용자들은 넓은 선택의 기회를 가진다. 일반적으로 많은 모바일 서비스 제공 기업은 사용자와의 지속적인 관계를 유지하기 위하여, 그리고 새로운 사용자의 유치를 위하여 높은 마케팅 비용을 지출한다. 모바일 서비스 기업들은 마케팅에 사용되는 비용과 노력을 효율적으로 투입하고자 사용자의 속성을 분석하고 서비스 사용행태를 분석한다. 모바일 헬스 서비스 역시 사용자와의 지속적인 관계를 유지하고 이탈하지 않도록 할 수 있는 방안을 개발하기 위하여, 사용자 및 사용행태 분석 연구가 필요하다.[7]

2. 관련 연구

기술의 발전과 의료 영역이 융합, 확장되면서 헬스케어 패러다임이 변화함과 동시에 헬스케어 산업은 디지털 헬스라는 개념으로 진화해왔다. 1970~2000년에는 의료정보학이 싹트고 병원내 IT시스템이 도입되면서 ‘헬스 IT(Health IT)’라는 용어가 생성되었고, 2000~2020년에는 전자의무기록(EMR : Electronic Medical Record)과 전자상거래 확산 등으로 ‘이헬스(E-Health)’라는 용어가 사용되었다. 2020년 이후는 인공지능(AI : Artificial Intelligence), 로봇, 사물인터넷(IoT : Internet of Thing) 등의 정보통신기술이 건강과 의료서비스에 접목되어 소비자가 어디서든 이용할 수 있도록 서비스를 제공하는 것을 ‘디지털 헬스’로 보고 있다. 디지털 헬스 산업의 범주는 출처의 성격과 국가별 제도 등의 차이로 공통된 범주를 제시하기엔 한계가 있다. 선행 연구들을 살펴보면, 각 분야별 시장 및 산업에 관한 보고서는 정리된 사례가 많으나, 디지털 헬스 산업 전체를 구분하여 정리된 사례는 많지 않다.[8] 그 중 글로벌 컨설팅 기업 딜로이트(Deloitte)가 2015년 영국의 생명과학부에 제시한 보고서가 디지털 헬스 산업 전체를 유형 별로 구분

하고 각 하위 범주까지 제시하는 첫 사례로 볼 수 있다.[9] 딜로이트의 디지털 헬스 산업 분류는 데이터의 흐름(수집 및 측정→저장→분석)을 따라 구성된 것으로 평가되지만, 연구 시점(2015)이 최근의 기술 변화를 반영하지 못한다는 것과 영국의 사례에 중점을 둔 분류라는 것은 한계점으로 볼 수 있다.[8] 딜로이트 보고서는 디지털 헬스를 크게 4가지 산업영역으로 유형을 구분하고 각 유형 별로 하위 범주를 제시하였다.[9] 4가지 디지털 헬스의 유형은 원격 의료서비스의 내용을 포함하는 텔레헬스케어, 모바일 어플리케이션 기반의 모바일 헬스 영역, 빅데이터 기반의 헬스 분석, 그리고 디지털 헬스 정보의 저장과 교류를 디지털화한 헬스시스템 4가지 유형이다. 이 중 모바일 헬스 영역에는 웨어러블 기기와 모바일 어플리케이션 하위 범주가 있다. 웨어러블 기기는 사용자의 활동, 심박수 등 생체신호, 수면 패턴 등의 데이터를 획득하고 모니터링하고 네트워크를 통해 연동하며, 모바일 어플리케이션은 소비자 중심의 건강관리 앱과 전문가 중심의 의료 앱을 포함한다.[10] 본 연구에서 모바일 헬스 서비스는 소비자 중심의 건강관리 앱을 통한 서비스를 대상으로 하였다.

ICT가 헬스케어 산업에 융합, 적용되는 방향에 대한 연구는 각 기술 영역별 연구가 상당히 활발하게 이루어져왔다.[11],[12] 또한 헬스케어 산업의 현황과 발전 방향에 대한 연구도 공공 및 민간 연구기관을 중심으로 진행되었으며, 이러한 연구에는 헬스케어 산업의 서브섹터로서 모바일 헬스의 시장, 기술 현황과 방향에 대한 연구가 포함된다.[4],[6] 모바일 헬스의 사용자에 대하여는 주로 기술수용 관점에서 서비스 수용에 대한 연구가 있었다.[5],[14]

데이터마이닝 기법을 활용한 서비스 사용 수준, 즉 사용자의 지속적 서비스 사용 혹은 이탈에 대한 연구는 주로 통신사, 금융사에서 많이 이루어지고 있는데 모바일 서비스 외의 상품을 제공하는 계약관계의 고객이 존재하고 이들의 이탈이 명확하며 상대적으로 상세한 개인정보를 보유하고 있기 때문에 이탈 예측 모델을 학습하기 쉽다.[7],[15] 이러한 연

구에서 도출된 모델은 비계약관계 회원서비스인 모바일 헬스 서비스에 적용하기에 적절하지 않으며, 산업군 특성에 맞는 고객 유지의 메커니즘을 이해하고 측정 체계를 개발하기 위한 노력이 요구된다.[16] 이러한 근거로 본 연구는 모바일 헬스 서비스 사용자의 지속적 사용에 대한 특성을 분석하고 예측 모델을 개발하고자 한다.

3. 연구 방법론

3.1 연구 대상

본 연구의 대상 자료는 A사 모바일 헬스케어 서비스에 2020년 11월에 가입한 무료 회원 사용자 9,122명의 2020년 11월부터 2021년 3월까지의 서비스 사용 데이터이며 R Studio 프로그램을 사용하여 분석하였다. 또한 사용자들의 접속 로그를 통하여 표 1과 같이 가입 이후 경과월차별 방문율을 구한 결과 3개월차까지 방문율이 급격히 하락하고 3개월차부터 완만하게 하락하는 현상을 확인하고 이에 따라 종속변수 ‘이탈여부’를 생성하여 2개월간 방문이 없는 경우를 ‘이탈’, 2개월 이내 방문이 있는 경우를 ‘유지’로 정의하였다.

<Table 1> Visit rate by month after user subscription

1개월차	2개월차	3개월차	4개월차	5개월차
100%	48%	29%	22%	15%

독립변수는 세 가지 성격의 데이터를 활용하였는데 개인 식별 정보 없이 사용자 기본 정보, 건강 활동 정보, 그리고 서비스 로그 정보가 그것이다.

본 연구에서 사용한 독립변수를 상세히 살펴보면 첫 번째, 사용자 기본 정보는 성별, 연령대, 가입 후 경과 기간 정보가 있다. 두 번째는 건강 활동 정보로서 모바일 헬스 서비스를 이용하며 측정된 사용자의 활동 자료이다. 일 걸음수, 체질량지수 (BMI : Body Mass Index) 등급, 활동 미션 달성률 정보가

포함된다. 세 번째, 서비스 로그 정보는 사용자의 서비스 사용 행태에 따라 시스템에 기록된 정보로서, 방문 횟수, 평균 체류시간, 평균 페이지 체류시간, 평균 방문 간격, 방문당 평균 페이지뷰 수, 센서 연결 여부, 조기(가입 후 30일 이내) 센서 연결 여부 정보가 포함된다. 전체 변수 목록은 아래 표2와 같다.

<Table 2> Variables

Category		Variables
Independent variables	User basic information	Gender
		Age group
		Period since sign-up
	Health activity information	Daily step count
		BMI grade
	Service log information	Activity mission achievement Rate
		Visits
		Average staying time per visit
		Average staying time per page
		Average interval between visits
		Average page views per visit
Dependent variable	Sensor connection status	
	Early sensor connection status	
Churn		

기계학습 알고리즘 적용을 위하여 데이터를 표3과 같이 학습용 자료와 검증용 자료를 70%, 30%로 분할하여 분석을 진행하였다.

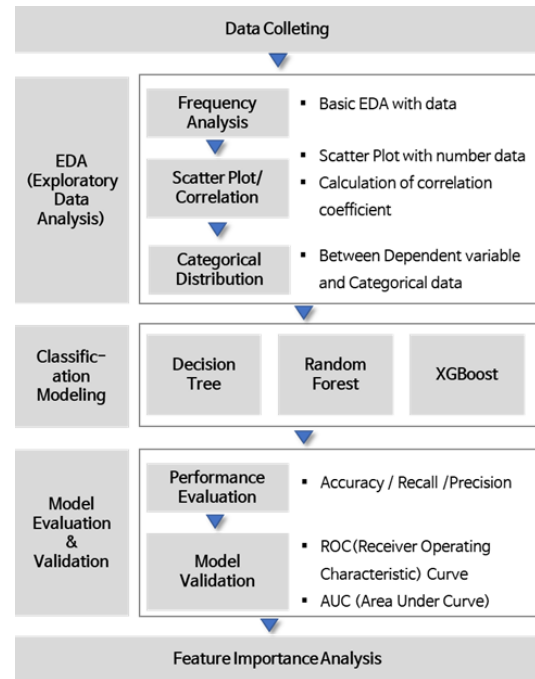
<Table 3> Variables Training and Testing Data

Data Category	# of Rows			%
	Churn	Retention	Sum	
Training	4,269	2,116	6,385	70%
Test	1,863	874	2,737	30%
Total	6,132	2,990	9,122	100%

3.2 연구 절차

본 연구는 그림1과 같이 (1)데이터를 수집하고 (2)탐색적 데이터 분석을 수행한 후 (3)기계 학습 알고리즘인 의사결정나무, Random Forest, XGBoost 기법을 적용하여 이탈 예측 모델을 만들었으며, (4)모델을 평가하고 검증한 후에 (5)변수 중요도를 분석하여 비즈니스 환경에서의 유의미한

해석을 도출하는 단계로 진행하였다.



[Figure 1] Research Process

전통적인 통계적 접근은 모집단에 대한 가설을 검증하는 목적에 적합한 방법론으로서 설명력 관점에서 좋은 방법론이지만 본 연구에서는 학습을 통하여 알 수 없었던 패턴을 찾고 이를 통하여 미래를 예측하고자 하는 목적을 가지고 기계 학습 모델링을 적용하였다.

3.3 데이터 마이닝 프로세스

통계적 규칙이나 의미 있는 패턴을 추출하는 데이터 마이닝 과정은 일반적으로 유사한 절차로 진행된다. Cross Industry Standard Process for Data Mining (CRISP-DM)은 이러한 절차를 여섯 단계의 프로세스로 정립하였다.[17]

비즈니스의 이해 단계에서는 프로젝트의 목적과 비즈니스 요구사항을 이해하고 두 번째 데이터 이해 단계에서는 데이터 수집 및 탐색을 통하여 통찰력을 얻는다. 데이터 준비 단계는 모델링 도구에 공급할 최종 데이터 집합을 준비하며, 네 번째 모델링 단계에서는 여러 모델링 기법을 적용하고 최적화한다.

모델에 대해 정밀한 평가를 수행하는 평가 단계를 거쳐 마지막 배포 단계에서는 모델링을 통하여 확보된 정보를 이해하고 이용할 수 있도록 제공한다. 보고서 작성 혹은 반복 적용 가능한 프로세스 구현 형식이 될 수도 있다.[17],[18]

3.4 데이터 마이닝 기법

데이터마이닝에 적용하는 기계학습 알고리즘은 일반적으로 지도학습과 자율학습으로 구분된다.[13] 본 연구에서는 종속 변수인 ‘이탈’에 대한 정보가 있으므로 지도학습을 수행하였다. 의사결정나무 기법과 앙상블 모델인 Random Forest, XGBoost(Extreme Gradient Boosting) 기법 등 세 가지 모델링 기법을 적용하였고 변수 중요도를 측정하여 ‘이탈’에 중요한 영향을 미치는 변수를 확인하였다.

의사결정나무 모형은 특정 변수에 대한 의사결정 규칙을 나무의 가지가 뻗는 형태로 분류해 나가는 알고리즘이다. 어떤 속성이 목표 변수에 대해 얼마나 많은 정보를 제공하는지가 중요하며, 이러한 속성을 첫번째 분할 기준으로 삼고 연속적으로 이러한 의사결정을 판별하게 된다. 분석과정이 관측되는 대표적인 화이트박스 모델로서 그 과정과 결론이 직관적이고 이해하기 쉽다는 특성을 가진다.[19]

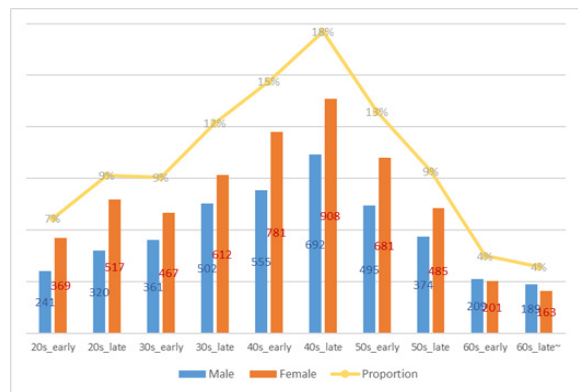
앙상블 기법은 여러 모델의 집합을 이용해서 더 나은 예측 결과에 도달하는 방법으로서 일반적으로 단일모델보다 더 좋은 성능을 보이며, 여러 모델의 결과를 집계하는 Bagging 방법과 이전 모델의 오차를 줄여나가는 방향으로 다음 모델을 생성하는 Boosting 방법 두 가지로 나뉜다.[20] 본 연구에서는 두 가지 앙상블 기법의 대표적인 모델인 Random Forest와 XGBoost(Extreme Gradient Boosting) 기법을 적용하였으며 과적합을 방지하기 위하여 교차타당성검사를 수행하였다. 또한 분석 결과를 실무적으로 유의미하게 해석하기 위하여 세 가지 모델링 기법 별 변수 중요도를 측정하였다.

4. 분석 결과

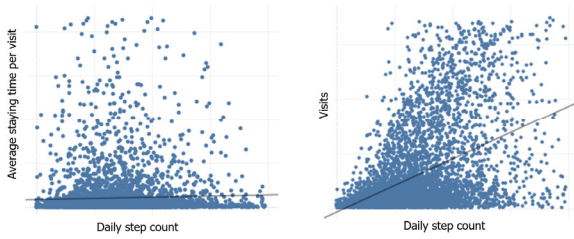
4.1 탐색적 데이터 분석 결과

탐색적 데이터 분석(EDA: Exploratory Data Analysis)을 진행한 결과, 본 연구의 대상이 된 모바일 헬스 서비스의 사용자는 그림2와 같이 연령대 중 40대가 33%로 가장 많고 성별로는 여성이 57% 비율로서 남성보다 더 많다.

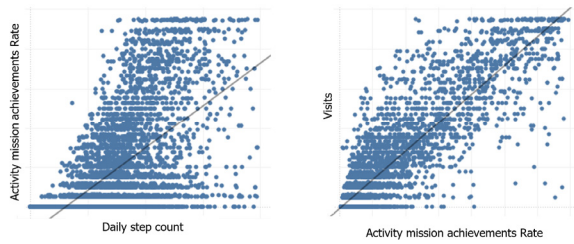
건강 활동 정보인 일 걸음수 변수에 대하여 서비스 로그 정보와의 산점도를 그리고 상관관계를 확인해본 결과, 그림 3과 같이 일 걸음수와 방문 당 체류시간 두 변수의 상관계수는 0.0361로서 상관관계가 거의 없지만, 일 걸음수와 방문 횟수 두 변수의 상관계수는 0.4438로서 약한 상관관계를 가진다. 또한 그림 4와 같이 일 걸음수와 활동 미션 달성률 간의 상관계수는 0.6117로서 상관관계가 있다고 볼 수 있으며, 활동 미션 달성률과 방문 횟수 간의 상관계수는 0.8899로서 강한 상관 관계를 보였다. 즉, 일 걸음수가 많을수록 활동 미션 달성률이 높고, 달성률이 높을수록 모바일 헬스 서비스에 방문을 많이 한다는 사실을 알 수 있다.



[Figure 2] Distribution of users by age and gender

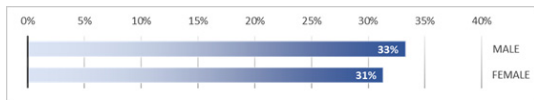


[Figure 3] Scatter plot and trend line of daily step count vs. average staying time, daily step count vs. visits

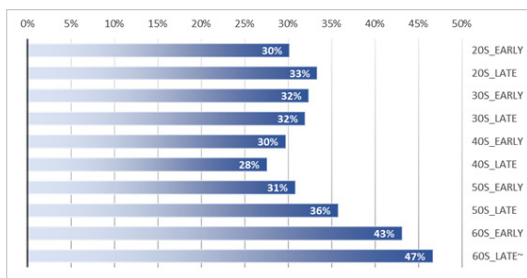


[Figure 4] Scatter plot and trend line of daily step count vs. activity mission achievement rate and visits vs. activity mission achievement rate

범주형 변수들과 종속 변수인 ‘이탈’에 대하여 시각화한 결과, 그림 5와 같이 ‘성별’ 변수의 경우 여성보다 남성의 ‘유지’ 비율이 약간 더 높았으며, 그림 6과 같이 다른 연령대에 비하여 40대 연령대 사용자가 이탈이 많고 60대 이상의 연령대의 미이탈율이 높음을 확인하였다.



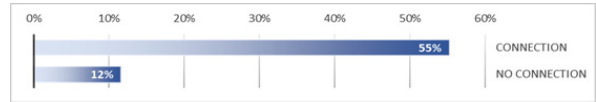
[Figure 5] Retention ratio by gender



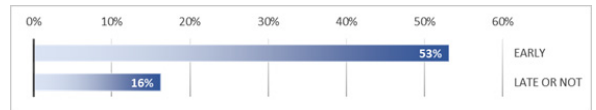
[Figure 6] Retention ratio by age group

또한, 센서 연결 여부와 조기 센서 연결 여부는 모바일 헬스 서비스 사용 수준을 확인할 수 있는 매우 뚜렷한 변수임을 그림 7, 그림 8과 같이 확인하

였다. 즉, 센서를 연결한 사용자 중 55%는 미이탈 사용자가 되지만 연결하지 않은 사용자는 단지 12%만 미이탈 사용자가 된다. 서비스 가입일 30일 이내에 센서를 연결한 경우에는 53%, 30일 이내에 연결하지 않은 사용자는 16% 비율의 사용자만이 미이탈 사용자가 된다.



[Figure 7] Retention ratio by sensor connection



[Figure 8] Retention ratio by early sensor connection

4.2 이탈 예측 모델링 결과

기계 학습 알고리즘인 의사결정나무, Random Forest, XGBoost 기법을 적용하여 이탈 예측 모델을 만들고 튜닝한 후 평가한 결과, 표 4의 모델 성능 지표를 도출하였다.

<Table 4> Model Evaluation

Model	Accuracy	Recall	Precision
Decision Tree	0.8951	0.9028	0.9407
Random Forest	0.9105	0.9331	0.9356
XGBoost	0.9101	0.9356	0.9320

정확도(Accuracy)란 실제로 이탈하는 사용자를 ‘이탈’로 분류하고 실제 유지하는 사용자를 ‘유지’로 분류한 비율이다. 정밀도(Precision)는 ‘이탈’로 분류한 사용자 중에 실제 이탈한 비율이며, 재현율(Recall)은 실제 이탈 사용자 중 모델에서 맞게 분류한 비율이다. 비즈니스 목적에 따라 중요한 성능 지표가 달라질 수 있으며, 이에 따라 최적 모델을 다르게 선정한다. 예를 들어 실제 이탈 사용자를 최대한 ‘이탈’로 예측해야 하는 비즈니스 목적이 있는 경우 재현율이 중요한 지표가 되어야 하며, 그림에

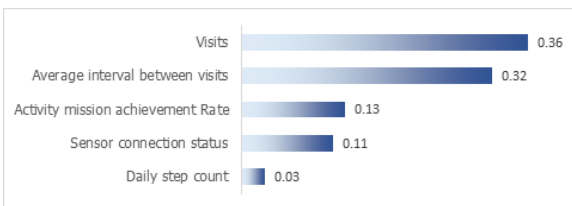
도 다른 성능 지표가 지나치게 낮은 모델은 제외해야 한다. 상호 보완적인 평가지표들은 어느 한쪽을 높이면 다른 지표의 값은 떨어지게 되는 트레이드오프 관계를 가진다. 이러한 트레이드오프를 가시적으로 확인할 수 있는 ROC(Receiver Operating Characteristic) Curve를 그리고 AUC(Area Under Curve) 값을 측정하여 모델의 안정성과 신뢰도를 검증하였다. ROC Curve는 FPR(False Positive Rate)이 변할 때 TPR(True Positive Rate, Recall)이 어떻게 변하는지 나타내는 곡선으로서, 이 곡선 아래의 면적을 구한 값인 AUC가 1에 가까울수록 안정적인 모델로 판단한다. 본 연구에서 도출한 세 가지 모델에 대한 ROC Curve의 AUC는 표5와 같다.

이번 연구 결과에서는 실제 이탈할 사용자를 이탈하지 않을 것이라고 예측하는 것을 방지하기 위한 비즈니스 목적을 고려할 때 재현율 지표가 중요한 성능지표라고 볼 수 있으므로 재현율이 가장 높고 ROC AUC 지표 역시 가장 높게 측정된 XGBoost 모델이 이러한 비즈니스 목적에 최적화된 모델로 선정될 수 있다.

<Table 5> AUC of ROC Curve by Model

Model	Decision Tree	Random Forest	XGBoost
AUC	0.9011	0.9327	0.9431

XGBoost 모델의 변수 중요도를 도출한 결과, 그림 9와 같이 상위 5개의 변수는 방문 횟수, 평균 방문 간격, 활동 미션 달성률, 센서 연결 여부, 일 걸음수 순서로 중요도가 높은 것이 확인되었으며, 연령, 성별, 체질량지수 등급 순으로 중요도가 낮았다.



[Figure 9] Variable importance of model to which XGBoost algorithm is applied

사용자 고유의 기본 정보보다 앱을 사용하면서 발생하는 사용 행태 정보인 건강 활동 정보 및 서비스 로그 정보 데이터가 사용자의 서비스 사용 수준을 예측할 수 있는 핵심 지표가 된다고 판단할 수 있다.

5. 결론

4차 산업혁명의 흐름 속에서 헬스케어 분야는 폭발적인 성장이 전망되는 산업으로 기대되고 있다. 특히 모바일 헬스 분야는 웨어러블 디바이스와 센서 기술의 혁신, 스마트폰의 보급과 함께 사회·경제적인 관점에서 더욱 큰 성장이 기대된다. 본 연구는 이러한 시장에서 모바일 헬스 서비스가 사용자와의 지속적인 관계를 유지할 수 있는 중요한 지표를 확인할 수 있는 분석 결과를 제시하고자 하였다.

탐색적 데이터 분석 결과, 모바일 헬스 서비스 연구 대상 집단에서는 40대 연령대와 여성 사용자의 비율이 상대적으로 높지만, 서비스를 지속 사용하는 ‘미이탈’ 사용자는 60대 이상의 연령대와 남성의 비율이 높았다. 또한 활동 미션 달성률이 높을수록 모바일 헬스 서비스에 방문을 많이 한다는 사실을 알 수 있었으며, 센서의 연결, 특히 서비스 이용 초기에 센서를 연결하여 사용자의 건강 활동을 감지할 수 있도록 하는 것은 사용자가 지속적으로 서비스를 사용, 유지하도록 하는 데에 중요한 요인이 됨을 확인할 수 있었다.

기계 학습 모델링과 검증 평가를 통하여 이탈 예측 모델을 개발하고 변수 중요도를 확인한 결과, 방문 횟수, 평균 방문 간격, 활동 미션 달성률, 센서 연결 여부, 일 걸음수에 따라 모바일 헬스 서비스 사용자의 서비스 사용 수준에 영향이 있음을 확인하였다. 즉, 모바일 헬스 서비스 제공자는 사용자로 하여금 서비스에 방문을 자주 하도록 이끌고, 달성할 만한 적정 활동 미션을 제시하며, 일 걸음수 등 건강 활동을 측정할 수 있는 센서를 조기에 연결할 수 있도록 안내함으로써 사용자의 모바일 헬스 서비스

의 사용 수준을 높일 수 있을 것이다. 변수 중요도는 특히 비즈니스 환경에서 현황을 점검하고 전략을 구상하는 데에 의미있는 지표를 제시할 수 있다. 본 연구에서는 모델 알고리즘에서 도출 가능한 변수 중요도 값을 확인하였으며 향후 연구에서는 변수들 조합에서의 모델에 대한 기여도 측정 등 비즈니스 실무 환경에 더욱 기여할 수 있는 연구를 수행하고자 한다. 이러한 연구를 통하여, 앞으로 중요도가 더욱 높아질 모바일 헬스 서비스의 활성화에 의미 있는 기여가 가능할 것으로 예상된다.

References

1. 스마트 헬스케어 의료기기 기술·표준 전략 보고서, 식품의약품안전평가원, p6-13, 2018
2. 생명표 국가승인통계 제101035호, 통계청, 2020
3. 이은환, 김욱, 질병예방과 건강수명연장을 위한 지역사회 공공 스마트 헬스케어 모델구축, 이슈&진단 제331호, p1-25, 2018
4. 정보통신산업진흥원, 웰니스 산업의 비즈니스모델 분석을 통한 산업 발전 방안 연구, p1-3, 2012
5. Zekun Zheng, 모바일 헬스케어 애플리케이션 수용에 관한 연구, 한국정책과학학회보 제19권 3호, p203-236, 2015
6. 신동진, 헬스케어 산업 지원 정책 및 건강관리 인식 확대에 따른 지속적 성장 전망, 한국기업데이터, p1-23, 2019
7. 이세희, 이지형, RNN을 이용한 고객 이탈 예측 및 분석, 한국컴퓨터정보학회 학술발표논문집 제24권 2호, p45-48, 2016
8. 김지은 외, 디지털 헬스 산업 분석 및 전망 연구, 한국보건산업진흥원, p9-117, 2021
9. 서경화, 디지털 헬스의 최신 글로벌 동향, 의료정책연구소 정책현안분석 2020-2 호, p3-34, 2020
10. Deloitte, Digital Health in the UK : An industry study for the Office of Life Sciences, 2015
11. Accenture, AI and me. Accenture Digital Health Technology Vision 2020, 2020
12. 김문구 외, 인공지능 헬스케어의 산업생태계 분석 및 활성화 방안 연구, 한국정보과학회 학술발표논문집, p720-722, 2016
13. 신수용, 국내 헬스케어 빅데이터 현황과 신규 연구 방향. 정보과학회지 제35권 5호, p16-19, 2017
14. 정계선 외, 모바일 헬스케어 앱의 지속적 수용에 관한 연구: 인지된 용이성의 확장을 중심으로, 사이버커뮤니케이션학보 제36권 2호, p81-117, 2019
15. 박지훈, 딥러닝 기반의 금융회사 고객이탈 예측 모형에 관한 연구, 전북대, 2020
16. 손정민 외, 온라인 게임의 고객 유형 별 이탈 요인 : 신규 고객과 기존 고객을 중심으로, 한국경영과학회지 제39권 제4호, p115-136, 2014
17. Michael J. A. Berry & Gordon S. Linoff, Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, John Wiley & Sons, 2004
18. R. Wirth & J. Hipp, CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining, p29-39, 2000
19. 박희진, 데이터 마이닝 및 암호화된 데이터를 이용한 예측모델 연구, 서울과학기술대학교, 2018
20. Lior Rokach, Ensemble-based Classifiers, Artificial intelligence 33(1), p1-39, 2010