

통합 로열티 프로그램 앱 로그 데이터를 이용한 이탈 고객 예측

권오현*, 박동현**, 조 신***

Customer Churn Prediction using App Log Data of a Coalition Loyalty Program

Oh-Hyun Kwon*, Dong-Hyun Park**, and Shin Cho***

요 약

고객 확보 및 유지는 기업의 이윤 극대화에 필수적이기 때문에 이탈 가능성이 큰 고객을 사전에 예측하는 것은 이론적, 실무적으로 중요한 과제이다. 본 연구는 국내 최대 통합 로열티 프로그램인 OK캐쉬백의 비식별 로그 데이터를 활용하여 앱 이용 고객의 이탈을 예측하는 모델을 개발하였다. 이용 고객의 인구통계학적 정보와 앱 내 행동 데이터를 변수로 사용하고 Naïve Bayes Classification, Support Vector Machine, Deep Neural Network 기법을 활용하였으며, 특히 RFM(Recency, Frequency, Monetary) 스코어를 예측 변수로 추가하여 예측력을 향상시켰다. 분석 결과에 따르면, Deep Neural Network가 여러 층의 가중치 계산으로 고차원적 특성이 추출되어 높은 F1-스코어를 나타냈으며, 행동 로그와 인구통계학적 정보를 활용함으로써 고객 이탈 예측이 가능함을 확인하였다.

Abstract

Churn prediction is an important task both theoretically and practically since customer acquisition and retention is essential to maximizing a company's profits. This study developed a model to predict the churn of app users by using the mobile app log data of OK Cashbag, the largest coalition loyalty program in Korea. The model used customer demographic information and in-app behavior data as variables, and utilized Naïve Bayes Classification, Support Vector Machine, and Deep Neural Network techniques. In particular, the prediction accuracy was significantly improved by adding RFM(Recency, Frequency, Monetary) scores as predictors. The analysis confirmed that the Deep Neural Network showed the highest F1-score by extracting high-dimensional characteristics, and it was also confirmed that customer churn could be effectively predicted by using the behavior log and demographic information.

Keywords

churn prediction, customer behavior analysis, loyalty program, machine learning, deep learning

* 연세대학교 정보대학원 비즈니스 빅데이터 분석
- ORCID: <http://orcid.org/0000-0001-7005-0925>
** 연세대학교 정보대학원 디지털경영
- ORCID: <http://orcid.org/0000-0001-8537-0905>
*** 연세대학교 정보대학원 교수(교신저자)
- ORCID: <http://orcid.org/0000-0001-7944-4349>

· Received: Jun. 08, 2021, Revised: Jun. 30, 2021, Accepted: Jul. 03, 2021
· Corresponding Author: Shin Cho
Graduate School of Information, Yonsei University, New Millennium Hall,
50 Yonsei-ro, Seodaemun-gu, Seoul, 03722, Korea,
Tel.: +82-2-2123-4546, Email: shincho@yonsei.ac.kr

I. 서 론

고객 행태를 분석하고 예측하는 것은 고객 확보 및 유지, 고객 가치 증진, 잠재고객 활성화 등을 통해 기업이 이윤 극대화를 하는데 필수적이다. 특히 한 번 유입된 고객을 지속적으로 유지하는 것이 매출액 및 수익성에 큰 영향을 미치기 때문에 이탈 가능성이 큰 고객을 사전에 예측함으로써 고객을 유지하고 관리할 필요성이 크다[1].

이런 필요성 때문에 고객 이탈에 대한 연구는 오래 전부터 광범위하게 이루어져왔다. 최근까지는 고객의 인구통계학적 특성과 행동 패턴을 바탕으로 주로 고객 세분화나 회귀분석 등의 분석 방법을 활용하여 고객 이탈에 영향을 미치는 요인들에 관한 연구가 주류를 이루었다[1][2].

최근에는 머신 러닝 및 딥러닝 모델을 이용한 예측이 활발하게 이루어지고 있는데, 고객 이탈을 예측하는데도 이 방법을 이용한 연구가 많이 이루어지고 있다. 고객의 거래이력, 행동 데이터 및 인구통계학적 정보를 바탕으로 소프트웨어 기업의 고객 이탈을 예측한 연구[3]에서는 머신 러닝 기법인 의사 결정 나무를 비롯하여 나이브 베이저안 분류(NBC, Naïve Bayes Classification)를 활용하였다. 그리고 신용카드 이용 고객 중 이탈고객과 유지고객들의 데이터를 활용한 연구[4]에서는 서포트 벡터 머신(SVM, Support Vector Machine) 기법을 사용하여 고객 이탈을 예측하였다. 또한 전자상거래 데이터를 이용한 고객 이탈 연구[5]는 딥러닝 기법인 합성곱 신경망(CNN, Convolutional Neural Network)을 이용하여 고객 정보를 비정형 정보와 결합해 고객의 행동 예측력을 높였다.

한편 로열티 프로그램은 신규 고객을 획득하고 기존 고객을 유지하기 위한 수단으로 가장 널리 쓰이는 방법 중 하나다. 전통적으로 로열티 프로그램은 단일 기업의 로열티 프로그램을 의미했지만 최근에는 다수의 가맹점과 소비자들을 연결하는 통합 로열티 프로그램이 늘어나고 있다. 우리나라에서는 OK캐쉬백이 대표적인 통합 로열티 프로그램으로서 4만 4천여 가맹점과 약 3,500만 명 회원을 보유하고 있다[6].

통합 로열티 프로그램에서의 고객 행태에 대한

실증 연구는 주로 OK캐쉬백을 대상으로 이루어졌는데, 앱 접속, 이벤트 참여, 쿠폰 다운로드 등 이용 패턴에 대한 연구[6], 포인트 적립 및 사용 행위에 대한 연구[7][8] 등이 대표적인 예다. 그러나 아직까지 로열티 프로그램에서 고객 이탈 예측에 대한 연구는 없었다. 그리고 방법론 관점에서는 로열티 프로그램에서의 고객 행태 분석에 머신 러닝 및 딥러닝 모델을 사용한 사례도 발견되지 않았다.

본 연구는 개인을 식별할 수 있는 정보를 포함하지 않은 OK캐쉬백 모바일 앱 로그 데이터를 SK플래닛으로부터 제공받아 로열티 프로그램에서 이탈 고객을 예측하는 모델을 개발하고자 하고자 한다. 이탈 고객 예측은 고객을 이탈 고객과 유지 고객으로 분류하는 이진 분류 모델로 일반화할 수 있으며 인구통계학적 변수와 앱 로그 데이터를 이용하여 NBC, SVM, 인공 신경망(DNN, Deep Neural Network)와 같은 모델링 기법으로 이탈 고객을 예측한다. 특히 모델 예측력을 높이기 위해 마케팅에서 널리 쓰이는 RFM(Recency, Frequency, Monetary) 스코어를 예측 변수로 추가한 결과를 제시한다.

로열티 프로그램의 주요 목표가 고객 유지에 있음을 감안할 때, 머신 러닝 및 딥러닝 기법을 활용하여 고객 이탈을 예측하는 연구는 실무적 필요성이 클 뿐 아니라 딥러닝 기법의 활용 범위를 확대했다는 점에서 학문적 기여도 있을 것으로 기대된다.

II. 관련 연구

2.1 고객 행태 예측과 고객 세분화

고객 행태 예측에 관한 선행 연구들은 확률적인 방법론을 사용한 연구들이 많은 편이다. 먼저 베이저안 네트워크를 구축해 스마트폰 사용자의 라이프로그를 분석해 현재의 활동들을 판별한 후 그 사용자의 미래 행동을 예측한 연구가 있다[9]. 이 연구에서는 사용자의 활동을 노드로 하여 이전의 활동 후 다음 활동을 할 때 사전 확률 값을 바탕으로 사후 확률 값을 계산한다. 이후 연관 분석을 통해 다음의 활동을 예측한다. 단순히 확률만을 가지고 고객의 활동을 예측했기 때문에 예측 성능이 낮을 수 있다는 한계점이 존재한다. 한편, 실시간으로 올라

인 쇼핑물의 고객들의 특정 상품에 대한 집중도를 분석해 구매의지가 있는 관심고객을 찾아내 전체 고객이 아닌 관심고객 중심으로 프로모션을 진행하는 시스템에 대한 연구[10]에서는 특정 상품 페이지에 자주 접속하는 고객을 관심고객으로 정의하여 실제 구매 고객 여부인지를 검증하는 방법을 이용하였다. 즉, 고객이 특정 상품 페이지 조회를 많이 할수록 구매할 확률이 높다는 가설을 정밀도, 재현율, F1-스코어로 검증했다. 다음으로 고객 행동을 웹 로그 분석을 통해 고객의 특정 콘텐츠 이용 선호를 파악할 수 있는 U-스코어, P-스코어, M-스코어를 계산하여 고객에게 개인화 추천 서비스에 활용한 연구가 있다[11]. 이 연구는 고객을 파악하는데 있어 지수화를 통해 추천 시스템에 활용할 수 있는 방법을 마련하였지만 단순히 지수화에 그쳤다는 한계가 있다.

고객 세분화에 관한 연구와 관련해서는 우선 고객의 구매행동을 기간이 고려된 기법으로 세분화해 마케팅 방안을 수립한 연구가 있다[12]. 해당 연구는 인터넷 쇼핑물의 구매 이력을 바탕으로 RFM 스코어를 계산하여 이를 바탕으로 고객 세분화 방법을 제안하고 있다. 한편 인구통계정보 뿐만 아니라 고객들의 거래 패턴, 채널 접촉 패턴을 활용해 고객을 세분화하는 연구[13]에서는 머신 러닝 방법론을 사용하여 고객 세분화를 수행했다는 점에서 의의가 있지만 클러스터링에 그쳐 맞춤형 추천 모형 개발이나 고객 이탈 예측까지는 이어지지 못했다.

2.2 고객 이탈 예측

기업에게 고객은 수익과 직결되는 요소로 고객 중심적 비즈니스 전략이 중요하다. 기업은 고객과의 긍정적 관계를 유지하기 위해 고객관계 관리에 많은 비용을 투자한다. 이는 고객이 이탈하게 되면 기업은 매출 감소로 이어지고 기존 고객 유지보다 신규 고객 유지에 들어가는 비용이 5배에서 많게는 6배까지 필요하기 때문이다[14]. 따라서 다양한 산업에서 고객의 이탈을 관리하기 위해 예측 모델을 개발하여 고객 이탈을 방지하는 방법에 대한 연구가 이루어지고 있다.

통신 서비스 산업에서는 이탈 고객들과 유지 고

객들의 실제 통화량 사용내역을 바탕으로 다양한 머신 러닝 방법을 비교해 고객 이탈 예측에 적합한 모델을 찾는 연구가 진행되었다[15]. 이탈 고객 예측에 있어 여러 가지 통화 관련 데이터를 사용하였지만 인구통계학적 정보나 다른 비정형 데이터를 함께 사용하여 분석하지는 못했다. 한편 게임 산업에서는 고객이 이탈하기 전까지의 행동 패턴을 분석하여 딥러닝 방식을 활용한 고객 이탈 예측 모델 연구가 있다[16]. 이 연구는 고객 행동 패턴과 통계적 특징을 모두 사용한 앙상블 모델을 통해 고객 이탈 예측의 정확도를 높일 수 있음을 보였다. 그리고 전자상거래 산업에서 정형 정보와 비정형 정보를 결합하여 CNN 모델을 활용해 고객행태를 예측한 연구[5]에서는 텍스트로 이루어진 VOC 데이터를 사용하여 다양한 고객의 유형과 이탈 고객을 딥러닝 기법을 이용해 예측하였다.

이처럼 최근에 머신 러닝 및 딥러닝 기법을 활용한 고객 이탈 예측 연구들이 이루어지고 있으나, 아직까지 로열티 프로그램에서 고객 이탈 예측에 대한 연구는 없었다.

2.3 로열티 프로그램 분석

로열티 프로그램 기업은 고객들의 정보나 서비스 활동 로그를 데이터로 활용할 수 있기 때문에 이를 고객 행태 예측에 활용하고 있다. 기업 입장에서는 신규고객을 유치하는 것보다 기존고객을 유지하는 것이 더 효과적이기 때문에 주로 수익 극대화를 위해 고객들의 이탈을 방지하려고 한다.

기업은 고객 유지를 위해 로열티 프로그램을 운영하고 있기 때문에 로열티 프로그램을 이용하는 고객이 로열티 프로그램을 어떻게 생각하고 이용하는지에 대한 분석이 필요하다. 로열티 프로그램의 혜택과 관련한 연구[17]에서는 고객의 태도와 만족도 및 지속적 이용 의향에 대해 연구하였다. 로열티 프로그램 업체별로 고객의 이용 태도와 만족도를 설문조사를 통해 분석했으며, 만족도와 지속적 이용 의향에 대해 집단별로 비교하였다. 그러나 설문조사 결과가 실제 이용 행태와 상당한 격차를 보이는 경향이 있음을 감안할 때 신뢰도에는 한계가 있다. 한편, 로열티 프로그램에서 포인트를 적립하는 노력

수준에 따라 포인트 소진의 금액 및 속도가 다르다는 점을 분석한 연구[18]에서는 거래 행동 데이터를 이용하여 행동을 분석했다는 점에서 의의가 있으나, 실제 행동이 어떻게 이루어질지에 대한 예측까지 이어지지는 못했다.

2.4 RFM 스코어

고객관계 관리와 고객 세분화를 위한 방법으로 널리 쓰이는 RFM 스코어는 고객별로 점수를 부여하는 방식으로 분석하는 방법이다. 일반적으로 고객이 얼마나 최근에 상품을 구입했는지를 나타내는 최근 구매일(Recency), 고객이 얼마나 자주 상품을 구입했는지를 나타내는 구매 빈도(Frequency), 고객이 상품을 구입하는데 사용하는 금액인 구매 금액(Monetary)을 기준으로 하여 가중 평균한 값을 지표로 계산하여 사용한다. 기업은 고객을 분석하면서 이익에 가장 크게 기여하는 변수를 판단하는데 있어 어려움이 있는데, 통계적인 방법을 이용하는 RFM 스코어는 다양한 변수들을 몇 개의 단순한 요인으로 집약할 수 있다는 점에서 효과적이다. RFM 스코어는 이러한 점수 계산을 통해 고객의 20%가 기업의 매출 80%를 가져온다는 이론인 파레토 법칙을 적용하여, 높은 점수를 기록한 세분화된 상위 고객층을 집중적으로 마케팅하려는 목표를 가진다[19].

하지만 로열티 프로그램 고객의 이탈을 예측하는데 있어 RFM 스코어를 활용하는 연구는 아직 없는 것으로 보인다. 기존의 인구통계학적 변수와 고객의 앱 내 행동 데이터를 이용한 모델링에 RFM 스코어를 이용하여 고객의 주기성을 파악하고 점수를 계산하여 변수화한 것을 추가하여 고객이 이탈할지에 대한 여부를 예측한다면, 고객 세분화와 마케팅 측면에서 높은 성능의 고객 이탈 예측 모델링이 가능할 것으로 기대된다.

III. 분석 기법

이탈 고객을 예측하는 분석은 앱 로그 데이터로부터 추출 가능한 고객의 행동 변수들을 이용하여 이탈 고객 및 유지 고객을 판별하는 분류 문제로 단순화할 수 있다. 분류 문제는 머신 러닝 분석 방

법 중 모델 학습 시 종속변수가 주어지는 지도 학습의 일종이며, 본 연구에서는 분류 문제의 여러 가지 분석 기법 중 조건부 확률을 기반으로 하는 NBC, 데이터 간 경계 평면을 설정하여 분류하는 SVM, 신경망 이론을 이용하여 입력 데이터를 최종 출력층에서 분류하는 DNN를 이용하여 이탈 고객 예측 모델의 성능을 확인해보고자 한다.

3.1 NBC

NBC는 조건부 확률에 베이지 정리를 적용하고 모든 특성에 대한 독립성을 가정하여 분류를 하는 확률적 분류 방법이다. NBC는 분류하려는 대상 x 의 속성 $x_1, x_2, x_3, \dots, x_n$ 이 서로 독립적이라 가정하고, x 를 카테고리 C_k 로 분류하는 작업은 아래 식 (1)과 같이 계산한다[20].

$$\begin{aligned} x &= [x_1 x_2 x_3 \dots x_n] \\ \arg_k \max P(C_k | x) \\ &= \arg_k \max P(x_1 x_2 x_3 \dots x_n | C_k) P(C_k) \\ &\approx \arg_k \max \prod_i P(x_i | C_k) P(C_k) \end{aligned} \tag{1}$$

$P(x_1 x_2 x_3 \dots x_n | C_k)$ 는 각 속성 변수들이 카테고리 C_k 에 속할 확률의 곱을 의미하고, 이 확률이 가장 큰 C_k 를 찾는 것이 NBC의 알고리즘이다. NBC의 장점은 연산 속도가 빠르고 분류 정확도가 높다는 점이다. 또한 노이즈와 누락 데이터가 있어도 비교적 수행 성능이 좋으며 다중 분류 문제에도 잘 동작한다[21].

3.2 SVM

SVM은 두 카테고리 중 어느 하나에 속한 데이터의 집합이 주어졌을 때, 주어진 데이터 집합을 바탕으로 하여 새로운 데이터가 어느 카테고리에 속할지 판단하는 비확률적 이진 선형 분류 모델을 만드는 알고리즘이다[22]. 만들어진 분류 모델은 데이터가 사상된 공간에서 경계로 표현되는데 SVM 알고리즘은 그 중 가장 큰 폭을 가진 경계를 찾는다. 입력 데이터를 커널 함수를 사용하여 고차원 형상 공간에 먼저 투영하고 해당 형상 공간에 선형 분류

문제를 형성함으로써 비선형 의사결정 기능을 학습하도록 SVM을 확장할 수 있다. 학습 데이터를 통해서 분류 모델을 만들면, 학습에 사용되지 않은 새로운 데이터를 어떤 카테고리에 속하는지 예측할 수 있게 된다. SVM을 통한 분류의 결과의 정확도가 높으려면 학습 데이터의 특성이 분류에 필요한 정보를 효율적으로 표현하고 있어야 한다.

SVM 모델은 고차원의 특성을 가진 데이터가 분류되는데 있어서 경계를 설정하는 함수인 커널 함수를 선형 함수 이외에 다양한 비선형 함수로 설정할 수 있기 때문에[23], 로열티 프로그램 고객 데이터처럼 많은 변수로 이루어져 있어 이탈 고객의 특성 분포가 일정하지 않은 경우에도 이전 분류의 성능을 높일 수 있다. 이러한 SVM의 특징을 활용하여 분류 예측 모델을 만들고 예측의 정확도를 확인하는 연구는 다양한 분야에서 진행되고 있다. 신용카드 이용고객들의 고객관계 관리를 위해 이탈고객과 유지고객들의 데이터를 활용하여 고객 이탈을 예측한 연구[4]에서도 SVM을 이용해 분류한 뒤 NBC 모델과 결합하여 예측 모델링을 진행하였다. 하지만 로열티 프로그램에 있어서 고객 이탈 예측을 SVM 모델을 이용해서 예측한 연구가 존재하지 않았기 때문에, 본 연구에서는 다른 모델과의 비교를 위해 예측 모델링에 사용하고자 한다.

3.3 DNN

인공 신경망을 통한 분류는 오류 역전과 알고리즘을 통해 학습되는 기본적인 방법을 통해 분류한다. 특히, 인공 신경망은 복잡한 입력에 대해서도 충분한 학습 과정을 통하여 입력된 데이터를 분류할 수 있다. DNN은 인공 신경망에서 입력층과 출력층 사이에 여러 개의 은닉층을 쌓아놓은 구조이다. 인공 신경망의 각 은닉층은 이전 계층의 출력값을 다음 계층의 입력 값으로 전달하는 구조를 가지고 있다. 이런 구조로 인해 데이터의 고유한 특징을 추상화할 수 있으며, 학습 과정에서 자동으로 특성 추출이 이루어지고 데이터 간의 관계를 예측함으로써 단일 신경망으로는 할 수 없던 식별문제를 해결한다.

DNN의 학습 과정 속 데이터 특성 추출 방식은

기존의 분류 문제 해결 모델에서는 확인할 수 없었던 고차원적인 특성을 추출할 수 있기 때문에, 많은 변수로 이루어진 로열티 프로그램 고객 데이터에 있어서도 측정된 인구통계학적 변수나 행동 변수를 넘어서는 새로운 특성 추출이 가능하다. 이러한 DNN 모델의 특징을 이용하여 분류 예측 모델을 만들고 예측의 정확도를 확인하는 연구는 다양한 분야에서 진행되고 있다. 그동안 자연어 처리나 이미지 인식에서 주로 활용되던 CNN을 사용하여 고객 정보를 비정형 정보와 결합해 고객의 행동 예측력을 높인 연구[5]에서도 다층의 신경망을 이용하여 전자상거래 데이터에서 심층적인 특성 추출로 고객 이탈을 더 정확하게 예측하려는 시도를 하기도 했다. 따라서 본 연구에서도 DNN이 가장 우수한 예측 모델링 성능을 보일 것으로 기대한다.

IV. 연구 방법

본 연구는 OK캐쉬백의 비식별 앱 로그 데이터를 이용하여 통합 로열티 프로그램 이용 고객의 앱 상에서의 행동을 분석하고, 이탈 고객(Churn)에 대해 정의한 후 이를 예측하는 모델을 만들어 통합 로열티 프로그램 기업이 고객 세분화를 함으로써 효율적인 마케팅 활동을 할 수 있도록 하는데 기여하고자 한다. 분석에는 2019년 1월부터 12월까지 OK캐쉬백 앱을 이용한 고객 20,494명의 비식별 앱 로그 데이터를 이용하였다.

개괄적인 연구 순서는 그림 1과 같다. 우선 종속 변수인 이탈 고객에 대한 정의를 내리고 기준에 따라서 전체 고객을 이탈 고객과 유지 고객(Retention)으로 구분한다. 고객의 인구통계학적 변수는 독립변수로 사용되며, 앱 로그로부터 추출되는 고객의 행동 데이터도 행동별, 포인트별, 시간대별로 수치화하여 추가적인 독립변수로 사용된다. 기본적인 고객의 인구통계학적 변수와 앱 내 행동 데이터를 이용한 모델링을 수행하고, 이후 RFM 스코어를 변수로 추가하여 새로운 모델의 성능을 비교한다.

전체 고객을 학습 데이터와 테스트 데이터로 나누어 각각 NBC, SVM, DNN으로 고객의 이탈 여부를 종속변수로 하는 모델을 학습한 후, 각 분석 기법을 사용했을 때의 예측 결과에 대해 평가한다.

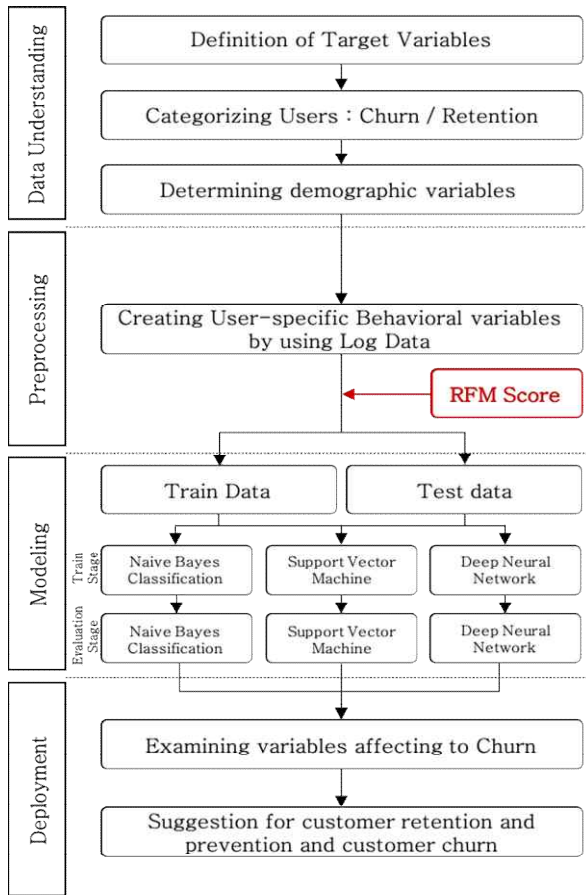


그림 1. 연구 프로세스
Fig. 1. Research process

이후 RFM 스코어가 반영되지 않았을 때와 반영됐을 때의 모델 성능을 비교하고 최근 구매일, 구매빈도, 구매 금액을 분석하는 것이 고객 이탈 예측에 미치는 영향이 어느 정도인지 확인한다.

로그 데이터는 각 유저별로 특정 페이지 혹은 쿠폰을 클릭할 때 발생하는 로그로, 각 고객별로 인구통계학적 변수로는 성별, 연령대 등과 앱을 접속할 때 이용한 휴대폰의 통신사 및 제조사가 있으며, 로그로부터 생성한 행동 변수는 크게 Action, Point, Holiday, Time Division, Event Type이 있다. 분석에 사용한 종속 변수 및 독립 변수들은 표 1과 같다.

먼저 종속변수인 이탈 고객 여부는 고객이 생성한 고객별 로그 데이터가 특정 날짜를 기준으로 90일 이내에 없을 경우 이탈로 보았다.

90일을 기준으로 잡은 것은 일반적으로 앱 내 잔존하는 고객 비율이 90일 뒤에는 10% 미만이라는 조사 결과[24]와 웹 정보 접근을 위한 인증 정보의 유지기간이 90일이기 때문에 웹사이트 비밀번호 변

경 기간도 보통 90일을 기준으로 한다는 논리를 참고했다. 본 연구에 쓰인 데이터의 경우 2019년 한 해의 로그 데이터이기 때문에, 2019년 9월 30일을 기준 날짜로 하여, 해당일 이전 가장 마지막으로 로그가 생성된 날짜로부터 90일 이내에 앱 로그가 생성되지 않은 고객은 이탈 고객으로 보았다.

그 이외의 인구통계학적 변수 및 행동 변수들은 모두 독립변수로 활용되었다. 먼저 Action은 접속, 검색, 매거진 확인, 이벤트 확인, 두툼 쿠폰 확인, 두툼 쿠폰 발급, 쿠폰 확인, 쿠폰 발급으로 나뉘며, 각각의 행동에 대해 고객별로 1년간의 행동 횟수를 구하여 전체 Action 횟수에 대한 비율로 변수를 생성하였다. 횟수를 직접 변수로 사용하지 않고 비율로 생성한 이유는 각 고객별로 생성하는 로그 개수의 차이가 심하기 때문에 스케일을 맞추기 위함이다. 비율 변수 대신 로그값을 취하는 방법도 고객별 스케일 차이를 줄여줄 수 있는 방법 중 하나이나, 실험 결과 로그값을 사용한 방법이 비율 변수 이용 방법에 비해 성능 차이가 거의 없었기 때문에 본 논문에서는 비율 변수를 사용한 결과를 제시하였다.

Point는 각 고객의 포인트 등급의 변화가 상승할 경우 ‘적립’으로 하락할 경우 ‘사용’으로 가정하여 각각의 변화 횟수를 파악한 후 전체 변화 횟수 중 적립에 대한 비율과 하락에 대한 비율로 각각 변수를 생성했다.

Holiday는 고객이 로그를 생성한 날이 공휴일 혹은 주말일 때의 로그 횟수와 평일일 때의 로그 횟수를 고객별 전체 생성 로그 횟수에 대한 비율로 변수를 생성하였다.

Time Division은 고객이 로그를 생성한 시간대가 2~6시일 경우 Daybreak, 7~11시일 경우 Morning, 12~15시일 경우 Midday, 16~20시일 경우 Twilight, 21~1시일 경우 Midnight으로 하여, 고객별 전체 생성 로그 횟수별 각 시간대의 비율로 변수를 생성하였다.

Event Type은 각 고객이 접속한 이벤트 페이지의 종류에 따라서 각각 전체 이벤트 중 어느 정도의 비율로 접속했는지에 대한 비율 변수로 생성하였다.

표 1에 정의된 이탈 고객 기준에 따라 분류된 전체 고객의 이탈 여부는 표 2와 같이 9,339명의 이탈 고객과 11,155명의 유지 고객으로 나타난다.

표 1. 종속 변수 및 독립 변수의 정의

Table 1. Dependent & Independent variables

	Variables		Definition
Dependent	Churn		Users whose activity logs have not been generated within 90 days(3 months) since the last activity log date
Demographic	Gender	Male, Female	User's gender
	Age	10s, 20s, 30s, 40s, 50s, 60s, 70s	User's age
	Carrier	SKT, KT, LGU+, Other Carr	User's mobile phone carrier
	Manufacturer	Samsung, LG, Apple, Other manu	User's mobile phone manufacturer
Behavioral	Action	Welcome ratio	Welcome log divide by action log
		Search ratio	Search log divide by action log
		Magazine view ratio	Magazine view log divide by action log
		Event view ratio	Event view log divide by action log
		Dutum view ratio	Dutum view log divide by action log
		Dutum issue ratio	Dutum issue log divide by action log
		Coupon view ratio	Coupon view log divide by action log
	Point	Coupon down ratio	Coupon down log divide by action log
		Point accumulation ratio	Point accumulation log divide by total log
	Holiday/Working day	Point use ratio	Point use log divide by total log
		Holiday ratio	Holiday log divide by Total log
	Time division	Working day ratio	Working day log divide by total log
		Daybreak ratio	2 a.m. ~ 6 a.m. log divide by total log
		Morning ratio	7 a.m. ~ 11 a.m. log divide by total log
		Midday ratio	12 a.m. ~ 3 p.m. log divide by total log
		Twilight ratio	4 p.m. ~ 8 p.m. log divide by total log
	Event type	Midnight ratio	9 p.m. ~ 1 a.m. log divide by total log
		Attend ratio	Attend log divide by Event type log
		Default event ratio	Default event log divide by event type log
		Direct ratio	Direct log divide by event type log
		Game ratio	Game log divide by event type log
Coupon ratio		Coupon log divide by event type log	
Dutum ratio		Dutum log divide by event type log	
Magazine ratio	Magazine log divide by event type log		
Purchase ratio	Purchase log divide by event type log		

표 2 이탈/유지 고객 분포

Table 2. User's distribution of churn / retention

Class	Counts
Churn	9,339
Retention	11,155

기존 사용변수 이외에 추가적으로 예측의 성능을 비교하기 위해 사용한 RFM 스코어는 기준일을 2019년 9월 30일로 하여 고객 별로 각각 그 이전 가장 최근에 로그가 생성된 일자를 최근 구매일, 로그를 생성한 모든 날짜의 수의 합을 구매 빈도, 로그 중 포인트를 적립하는 활동을 하여 적립된 모든

포인트 금액의 합을 구매 금액으로 정의하여 계산하였다.

R 스코어는 가장 최근에 로그가 생성된 순서부터 나열하여 약 20%에 가까운 균등한 5개 구간으로 나누었다. 가장 최근에 로그가 생긴 고객부터 5점을 부여하고, 가장 오래 전에 로그가 생긴 고객의 경우 1점을 부여하였다. 표 3에서 확인할 수 있듯, 2019년 9월 30일을 기준으로 기준일 당일이 가장 최근 로그인 경우 5점, 8일 전부터 1일 전 사이의 경우 4점, 39일 전부터 9일 전 사이의 경우 3점, 117일 전부터 40일 전 사이의 경우 2점, 118일 전부터 그 이전까지 모든 경우 1점을 부여하였다.

F 스코어는 2019년 9월 30일보다 이전의 모든 로그가 생성된 날짜의 수의 합이 큰 순서로 나열하여 약 20%에 가까운 균등한 5개 구간으로 나누어 가장 큰 횟수를 가진 고객부터 5점을 부여하고, 가장 적은 횟수를 가진 고객에게 1점을 부여하였다. 표 4에서 확인할 수 있듯, 69회에서 273회 사이의 고객은 5점, 27회에서 68회 사이의 고객은 4점, 12회에서 26회 사이의 고객은 3점, 5회에서 11회 사이의 고객은 2점, 1회에서 4회 사이의 고객은 1점을 부여하였다.

M 스코어가 기존의 온라인 쇼핑물 기준으로는 구매 금액을 의미했는데, 로열티 프로그램에서는 적립 활동이 고객이 앱을 사용하는 주된 목적이기 때문에 본 연구에서는 구매 금액을 대신하여 고객이 적립한 포인트의 합을 M 스코어로 설정했다.

포인트 적립 활동의 기준은 보유 포인트 구간의 상승하는 변동이 있을 경우로 정했다. 적립 포인트의 합이 큰 순서로 나열하여 5개의 구간으로 나누어 가장 큰 고객부터 5점을 부여하고, 가장 적은 고객에게 1점을 부여하였다. 분석 대상 기간 내 보유 포인트 구간의 상승 변동이 있지 않은 고객이 많았기 때문에 1점인 고객의 비율이 66.6%가 되었고, 나머지 2점에서 5점 사이의 고객들을 균등한 분포로 나누었다. 표 5에서 확인할 수 있듯, 1년 간 적립한 포인트의 합이 70포인트에서 18296포인트 사이의 고객은 5점, 12포인트에서 69포인트 사이의 고객은 4점, 3포인트에서 11포인트 사이의 고객은 3점, 1포인트에서 2포인트 사이의 고객은 2점, 적립한 포인트가 없는 고객은 0점을 부여하였다.

V. 분석 결과

표 3. R 스코어

Table 3. R score

R score	Days before	Counts	Ratio(%)	Acc. ratio(%)
5	0	4,804	23.4%	23.4%
4	1 ~ 8	3,868	18.9%	42.3%
3	9 ~ 39	3,920	19.1%	61.4%
2	40 ~ 117	3,955	19.3%	80.7%
1	118 ~ 272	3,947	19.3%	100%

표 4. F 스코어

Table 4. F score

F score	Days	Counts	Ratio(%)	Acc. ratio(%)
5	69 ~ 273	4,036	19.7%	19.7%
4	27 ~ 68	4,136	20.2%	39.9%
3	12 ~ 26	4,178	20.4%	60.3%
2	5 ~ 11	4,270	20.8%	81.1%
1	1 ~ 4	3,874	18.9%	100%

표 5. M 스코어

Table 5. M score

M score	Points	Counts	Ratio(%)	Acc. ratio(%)
5	70 ~ 18296	1,676	8.2%	8.2%
4	12 ~ 69	1,666	8.1%	16.3%
3	3 ~ 11	1,727	8.4%	24.7%
2	1 ~ 2	1,785	8.7%	33.4%
1	0	13,640	66.6%	100%

학습 데이터와 테스트 데이터에 포함되는 고객의 비율은 각각 70%, 30%로 나누었다. 각각의 데이터 셋에는 이탈 고객 7,270명, 유지 고객 7,075명이 학습 데이터에, 이탈 고객 3,034명, 유지 고객 3,115명이 테스트 데이터에 포함되었다. 학습 데이터를 학습하는 과정에서는 각 변수들의 분산을 맞춰주는 표준 스케일링을 진행하고, 인구조계학적 변수의 경우 범주형 변수이므로 원-핫 인코딩을 진행했다. 이탈 고객 예측은 정확도와 F1-스코어로 평가하였다. F1-스코어는 모델이 이탈 고객으로 예측한 것 중 실제 이탈 고객의 비율인 정밀도와 실제 이탈 고객 중 모델이 이탈 고객으로 예측한 재현율의 조화 평균으로, 분류 예측 모델에서 평가 지표로 주로 사용한다. 본 연구에서는 실제 이탈한 고객을 모델이 이탈한 고객이라고 얼마나 정확하게 예측하는지 확인하는 것이 중요하기 때문에 F1-스코어를 주요 평가 지표로 설정했다.

5.1 RFM 스코어가 반영되지 않은 경우

NBC를 이용한 예측 모델의 정확도는 65.46%, F1-스코어는 50.14%로 나타났고, SVM을 이용한 예측 모델의 정확도는 69.39%, F1-스코어는 61.10%로 나타났다. DNN은 은닉층 2층으로 이루어진 구조로

모델링을 했으며, 각 층의 노드 수는 64개로 하였다. 옵티마이저는 adam을 사용하고 손실 함수로는 이진 분류이므로 교차 엔트로피를 이용하였다. 이를 이용한 예측 정확도는 69.39%, F1-스코어는 65.92%로 나타났다. 각 모델별 정확도와 F1-스코어는 다음 표 6과 같다.

각 예측 모델별 혼동 행렬을 살펴보면 다음과 같다. NBC 모델의 결과는 표 7, SVM 모델의 결과는 표 8, DNN 모델의 결과는 표 9와 같다.

NBC 모델을 이용한 예측에서는 유지 고객을 정확하게 예측하는 경우가 많았다. SVM 모델을 이용한 예측에서는 이탈 고객을 정확하게 예측하는 경우가 더 늘어났으며, DNN 모델을 이용한 예측에서는 이탈 고객을 정확하게 예측하는 경우가 가장 많았다.

표 6. RFM 스코어가 반영되지 않은 경우 모델별 정확도
Table 6. Accuracy for each model when RFM score is not reflected

Model	Accuracy	F1-score
Naive bayes classification	0.6546	0.5014
Support vector machine	0.6939	0.6110
Deep neural network	0.6939	0.6592

표 7. NBC 혼동행렬
Table 7. Confusion matrix of NBC

		Predicted	
		Churn	Retention
Actual	Churn	1,068	1,671
	Retention	453	2,957

표 8. SVM 혼동행렬
Table 8. Confusion matrix of SVM

		Predicted	
		Churn	Retention
Actual	Churn	1,478	1,261
	Retention	621	2,789

표 9. DNN 혼동행렬
Table 9. Confusion matrix of DNN

		Predicted	
		Churn	Retention
Actual	Churn	1,820	919
	Retention	963	2,447

본 연구에서 중점적으로 예측하려는 고객은 이탈 고객이기 때문에, 이탈 고객을 가장 잘 예측하는 DNN 모델의 성능이 가장 적합하다는 것을 알 수 있다.

5.2 RFM 스코어가 반영된 경우

기존의 독립변수 Gender, Age, Carrier, Manufacturer, Action Ratio, Point Ratio, Holiday Ratio, Time Division Ratio, Event Type Ratio에 각 고객별로 계산된 R 스코어, F 스코어, M 스코어를 변수로 추가하여 고객의 이탈 여부를 종속변수로 하는 예측 모델을 새롭게 진행했다. 모델링 과정에서 변수를 새롭게 추가하는 것 이외의 다른 조건은 RFM 스코어가 반영되지 않은 경우와 동일하게 진행하였다.

NBC를 이용한 예측 모델의 정확도는 77.38%, F1-스코어는 71.61%로 나타났고, SVM을 이용한 예측 모델의 정확도는 88.13%, F1-스코어는 84.61%로 나타났다. 인공신경망은 RFM 스코어가 반영되지 않은 경우와 마찬가지로 은닉층 2층으로 이루어진 구조로 모델링을 했으며, 각 층의 노드 수는 64개로 하였다. 옵티마이저는 adam을 사용하고 손실 함수로는 이진 분류이므로 교차 엔트로피를 이용하였다. 이를 이용한 예측 정확도는 88.40%, F1-스코어는 86.76%로 나타났다. 각 모델별 정확도와 F1-스코어는 다음 표 10과 같다.

각 예측 모델별 혼동 행렬을 살펴보면 다음과 같다. NBC 모델의 결과는 표 11, SVM 모델의 결과는 표 12, DNN 모델의 결과는 표 13와 같다.

RFM 스코어를 반영하지 않았을 때와 마찬가지로 NBC 모델, SVM 모델, DNN 모델 순서로 예측의 정확도와 F1-스코어가 높았다. 또한, 세 가지 모델 모두 RFM 스코어를 반영하지 않았을 때에 비해 정확도와 F1-스코어가 상승하였다. NBC 모델을 이용한 예측에서는 기존의 RFM 스코어를 반영하지 않았을 때보다 정확도는 11.92% 상승하고, F1-스코어는 21.47% 상승하였다. 기존에는 유지 고객을 정확하게 예측하는 경우가 많았다면, RFM 스코어를 반영한 후에는 이탈 고객을 정확하게 예측하는 경우가 늘어났다. SVM 모델을 이용한 예측에서는 기

존의 RFM 스코어를 반영하지 않았을 때보다 정확도는 18.74% 상승하고, F1-스코어는 23.51% 상승하였다. 유지 고객을 정확하게 예측하는 경우와 이탈 고객을 정확하게 예측하는 경우 모두 상승하여 향상된 성능을 보였다. DNN 모델을 이용한 예측에서는 기존의 RFM 스코어를 반영하지 않았을 때보다 정확도는 19.01% 상승하고, F1-스코어는 20.84% 상승하였다. 다른 모델에 비해 이탈 고객을 정확하게 예측하는 경우가 가장 많았고, RFM 스코어를 반영하지 않았을 때의 결과와 같이 중점적으로 예측하려는 고객이 이탈 고객이기 때문에, 이탈 고객을 가장 잘 예측하는 DNN 모델의 성능이 가장 적합하다는 것을 다시 한 번 확인할 수 있었다.

표 10. RFM 스코어가 반영된 경우 모델별 정확도
Table 10. Accuracy for each model when RFM score is reflected

Model	Accuracy	F1-score
Naive bayes classification	0.7738	0.7161
Support vector machine	0.8813	0.8461
Deep neural network	0.8840	0.8676

표 11. NBC 혼동행렬
Table 11. Confusion matrix of NBC

		Predicted	
		Churn	Retention
Actual	Churn	1,754	985
	Retention	406	3,004

표 12. SVM 혼동행렬
Table 12. Confusion matrix of SVM

		Predicted	
		Churn	Retention
Actual	Churn	2,171	568
	Retention	162	3,248

표 13. DNN 혼동행렬
Table 13. Confusion matrix of DNN

		Predicted	
		Churn	Retention
Actual	Churn	2,336	403
	Retention	310	3,100

VI. 결론 및 논의

본 연구에서는 통합 로열티 프로그램인 OK캐쉬백 앱 이용 고객 행태를 고객의 이탈 여부 측면에서 이탈 고객 예측 모델링을 통해 분석하였고, RFM 스코어를 결합하여 고객의 세분화에 따른 고객 이탈 예측 모델의 성능 향상을 확인했다. 이탈 고객을 예측하는 이진 분류 모델로 NBC, SVM, DNN 세 가지 방법을 사용하였으며, 인구통계학적 변수와 앱 내 행동 데이터를 이용하여 이탈 고객을 예측해보고 부가적으로 RFM 스코어를 계산하여 변수로 추가한 예측 모델링을 통해 이를 평가하였다.

정확도와 F1-스코어 측면에서 모델의 성능을 각각 평가해본 결과로는 RFM 스코어가 반영된 경우와 아닌 경우 모두 DNN이 가장 성능이 우수한 것으로 확인되었다. 이는 DNN의 학습 과정 속 데이터 특성 추출 방식이 기존의 분류 문제 해결 모델에서는 확인할 수 없었던 고차원적인 특성을 추출할 수 있기 때문에, 많은 변수로 이루어진 멤버십 고객 데이터에서 인구통계학적 변수나 행동 변수, RFM 스코어 이외에도 사람이 식별하지는 못하지만 신경망 층 간 연결된 가중치 계산을 통한 새로운 특성이 추출되면서 다른 모델에 비해 우수한 성능을 나타낸 것으로 보인다. DNN의 정확도와 F1-스코어로 보았을 때, 로열티 프로그램 앱을 이용하는 고객의 행동 로그와 인구통계학적 정보로 고객 이탈을 예측하는 것이 상당 부분 가능한 것임을 확인할 수 있었으며, RFM 스코어를 추가하여 성능의 높은 향상을 이룬 것으로 보아 고객의 활동을 점수로 나타내는 세분화가 고객 이탈 예측에 중요한 영향을 미친다고 할 수 있다.

본 연구 결과를 활용하여 로열티 프로그램 기업은 고객의 최근 이용, 이용 주기, 포인트 적립 금액을 주요 고객 세분화 지표로 설정하여 이용 고객 관리를 통한 고객 이탈 방지 및 수익 창출이 가능해질 것으로 기대한다. 비즈니스 특성 상 고객의 진입과 진출이 활발한 로열티 프로그램 운영에서 고객 이탈 관점에서의 비용 최소화는 수익 극대화 전략에 기여하는 바가 클 것으로 예상된다.

향후 연구에서는 본 연구에서 사용한 수치화된

고객 행동 데이터 이외에도 고객의 검색 기록이나 VOC 같은 텍스트 형태의 비정형 데이터를 함께 사용해서 고객 행동 특성 파악을 다차원으로 접근한다면, 고객 이탈에 영향을 미치는 요소에 대한 보다 자세한 분석이 가능할 것으로 보인다.

본 연구는 로열티 프로그램에서 딥러닝 기법을 활용하여 처음으로 고객 이탈을 예측함으로써 관련 연구에 기여하였는데, 다른 비즈니스에서도 고객 행동 로그 데이터를 통한 이탈 고객 예측에 본 연구에서 사용한 방법이 활용될 수 있기를 기대한다.

References

- [1] S. Y. Kim, J. Y. Song, and G. I. Lee, "A Study of Customer Churn by Analysing CRM Customer Data", Korean Marketing Association, Vol. 7, No. 1, pp. 21-42, Apr. 2005.
- [2] M. J. Kim, J. I. Kim, and S. Y. Park, "A Study on the Influencing Factors that Determine Customer's Churning Behavior on IPTV Services", Korea Society of Consumer Studies, Vol. 2013, No. 5, pp. 295-302, May 2013.
- [3] C. Cimpoeru and A. Andreescu, "Predicting Customers Churn in a Relational Database", Informatica Economică, Vol. 18, No. 3, pp. 5-16, Mar. 2014.
- [4] M. A. H. Farquod, V. Ravi, and S. Baqi Raju, "Churn prediction using comprehensible support vector machine: An analytical CRM application", Applied Soft Computing, Vol. 19, pp. 31-40, Feb. 2014, <https://doi.org/10.1016/j.asoc.2014.01.031>.
- [5] S. S. Kim and J. W. Kim, "Customer Behavior Prediction of Binary Classification Model Using Unstructured Information and Convolution Neural Network: The Case of Online Storefront", Journal of Intelligence and Informations Systems, Vol. 24, No. 2, pp. 221-241, Aug. 2018, <https://doi.org/10.13088/jiis.2018.24.2.221>.
- [6] J. H. Jo, H. J. Jang, and S. Cho, "An Empirical Analysis of Mobile App-Based Coalition Loyalty Program Usage Pattern", The e-Business Studies, Vol. 22, No. 1, pp. 83-105, Feb. 2021.
- [7] J. Y. Kim, J. Lee, and S. Y. Kim, "Impacts of Reward Accrual Effort on Redemption Behavior in a Multi-Vendor Loyalty Program", Asia Marketing Journal, Vol. 18, No. 4, pp. 77-98, Jan. 2017, <https://doi.org/10.15830/amj.2017.18.4.77>.
- [8] K. R. Koo and W. S. Woo, "Managing Customer's Usage Behavior in a Multi-vendor Loyalty Program", Journal of Distribution Science, Vol. 13, No. 5, pp. 5-14, May. 2015, <http://dx.doi.org/10.15722/jds.13.5.201505.5>.
- [9] J. G. Bang and B. M. Kim, "A user behavior prediction technique using mobile-based Lifelog", Journal of the Korea Industrial Information Systems Research, Vol. 19, No. 6, pp. 63-76, Dec. 2014, <http://dx.doi.org/10.9723/jkiis.2014.19.6.063>.
- [10] H. J. Lee, H. K. Park, and W. S. Lee, "Implementation of Customer Behavior Evaluation System Using Real-time Web Log Stream Data", The Journal of Korean Institute of Information Technology, Vol. 16, No. 12, pp. 1-11, Dec. 2012, <http://dx.doi.org/10.14801/jkiit.2018.16.12.1>.
- [11] D. Y. Seo, D. J. Kim, J. K. Yun, J. H. Kim, K. S. Moon, and J. H. Oh, "Method for Preference Score Based on User Behavior", CRM Research, Vol. 4, No. 1, pp. 55-68, Mar. 2011.
- [12] Y. P. Jung and C. S. Yum, "Customer Relationship Management of the Internet Shopping Mall Using Customer Segmentation", The Journal of Korean Institute of Information Technology, Vol. 11, No. 12, pp. 159-167, Dec. 2013.
- [13] M. S. Chang and H. J. Kim, "A Customer Segmentation Scheme Base on Big Data in a Bank", Journal of Digital Contents Society, Vol. 19, No. 1, pp. 85-91, Jan. 2018, <http://dx.doi.org/http://dx.doi.org/10.9728/dcs.2018.19.1.85>.
- [14] A. D. Athanassopoulos, "Customer Satisfaction Cues To Support Market Segmentation and

Explain Switching Behavior", Journal of Business Research, Vol. 47, No. 3, pp. 191-207, Mar. 2000.

[15] B. Huang, M. T. Kechadi, and B. Buckley, "Customer churn prediction in telecommunications", Expert Systems with Applications, Vol. 39, No. 1, pp. 1414-1425, Jan. 2012.

[16] K. T. Kim and J. H. Lee, "Predictive Models for Customer Churn using Deep Learning and Boosted Decision Trees", Journal of Korean Institute of Intelligent Systems, Vol. 28, No.1, pp. 7-12, Feb. 2018, <http://dx.doi.org/10.5391/JKIIS.2018.28.1.7>.

[17] Y. M. Park and H. S. Lee, "Business Management and Consumer Use of Membership Services", Journal of Consumption Culture, Vol. 17, No. 1, pp. 113-138, Mar. 2014, <http://dx.doi.org/http://dx.doi.org/10.17053/jcc.2014.17.1.006>.

[18] J. Y. Kim, J. H. Lee, S. Y. Kim, and B. H. Lee, "The Effects of Point Accumulation Effort Level on Redemption Behavior in Loyalty Program", Korean Journal of Marketing, Vol. 27, No. 1, pp. 85-106, Mar. 2012.

[19] Y. G. Lee and H. K. Choi, "A Comparative Study of RFM Model for using Statistical Techniques", Journal of Mathematics and Statistics, Vol. 9, pp. 95-118, Dec. 2002.

[20] S. Yu, "A Study on Recommendation Method for Real Estate Using Naive Bayes Classification", The Journal of Korean Institute of Information Technology, Vol. 17, No. 10, pp. 115-120, Oct. 2019, <http://dx.doi.org/10.14801/jkiit.2019.17.10.115>.

[21] P. Domingos and M. Pazzani, "Beyond Independence: Conditions for the Optimality of the Simple Bayesian Classifier", Proceedings of 13th International Conference on Machine Learning, San Francisco, CA, United States, Vol. 1996, pp. 105-122, Jul. 1996.

[22] H. I. Lim, "Design of Similar Software Classification Model through Support Vector Machine", Journal of Digital Contents Society,

Vol. 21, No. 3, pp. 569-577, Mar. 2020, <http://dx.doi.org/10.9728/dcs.2020.21.3.569>.

[23] Sagar S. Nikam, "A Comparative Study of Classification Techniques in Data Mining Algorithms", Oriental Journal of Computer Science & Technology, Vol. 8, No. 1, pp. 13-19, Apr. 2015.

[24] Braze Customer Engagement Platform, <https://www.braze.com/resources/articles/app-customer-retention-spring-2016-report> [accessed: Jun. 22, 2021]

저자소개

권 오 현 (Oh-Hyun Kwon)



2018년 8월 : 성균관대학교
프랑스어문학과(문학사)
컴퓨터공학과(공학사)
2019년 9월 ~ 현재 : 연세대학교
정보대학원 석사 4학기
관심분야 : 머신러닝, 데이터 분석,
스마트 모빌리티

박 동 현 (Dong-Hyun Park)



2019년 2월 : 세종대학교
경영학과(경영학사)
2020년 3월 ~ 현재 : 연세대학교
정보대학원 석사 3학기
관심분야 : 디지털 전략,
데이터기반 전략, 디지털 플랫폼

조 신 (Shin Cho)



1982년 : 서울대학교 경제학사
1989년 : Washington University
경제학 박사
2013년 4월 ~ 현재 : 연세대학교
정보대학원 교수
관심분야 : 디지털 전환 전략,
기술혁신전략, 정보통신산업분석