

머신러닝 예측모형을 활용한 간편결제 서비스 이용 요인 분석

박희대

한국해양수산개발원 전문연구원

Factors of Simple Payment Service Use: Application of Machine Learning Prediction Model

Hee-Dae Park

Associate Research Fellow, Korea Maritime Institute, 26, Haeyangro 301 beon-gil, Yeongdo-gu, Busan, 49111, Korea

[요 약]

본 연구에서는 최근 성장하고 있는 간편결제 서비스에 영향을 미치는 요인을 머신러닝 예측 모형으로 분석하고, 간편 결제 서비스 시장 발전을 위한 정책적 시사점을 도출하고자 하였다. 이를 위해 2017년~2019년 한국방송진흥공사의 소비자 행태조사 자료를 분석에 사용하였다. 간편결제 서비스 이용에 대한 의사결정에 영향을 미칠 수 있는 요인을 인구 특성, 온라인 이용환경, 온라인 콘텐츠 수용도로 구분하고 이 중 어떤 요인이 주로 영향을 주는지를 비교 및 분석하였다. 분석결과 개별적인 요인만 고려된 모형보다 모든 특성 요인을 동시에 고려한 모형이 가장 높은 예측 성과를 보였다. 또한 간편결제 서비스 이용에 유의한 영향을 미치는 요인 중 온라인 이용환경 및 이용 정도와 관련된 변수가 가장 중요함을 확인하였다. 이를 통해 온라인 환경에 대한 경험이 간편결제 서비스 이용을 증가시키는데 매우 중요한 요인임을 파악하였다. 이는 소비자의 온라인 환경 및 서비스에 대한 노출을 증가시키는 정책 및 전략이 간편결제 서비스 시장 성장에 필요함을 시사한다.

[Abstract]

This study aims to analyze the factors affecting the use of simple payment service and suggest policy implications. To do this, the Consumer Behavior Survey from 2017 to 2019, provided by Korea Broadcasting Advertising Corporation, was used to analyze. This paper divides factors into three categories, namely demographic factors, accessibility and exposure to online service and acceptance of online contents, to figure out which factors mostly affect the use of simple payment service. The estimation results reveal that the model including all factors outperforms the models with subsets of factors. Furthermore, it is found that the most important factors of using simple payment service are variables related to accessibility and exposure to online service. Based on the results, it is suggested that the policies and strategies, which encourage consumers to experience up-to-date online service, are necessary for the online simple payment service to develop.

색인어 : 간편결제, 핀테크, LASSO, 예측모형, 머신러닝

Key word : Simple payment service, FinTech, LASSO, Prediction, Machine Learning

<http://dx.doi.org/10.9728/dcs.2020.21.5.921>



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 02 April 2020; Revised 15 May 2020

Accepted 25 May 2020

*Corresponding Author; Hee-Dae Park

Tel: +82-51-797-4752

E-mail: heedae85@kmi.re.kr

I. 서론

다양한 기술의 발전과 플랫폼 서비스의 등장으로 결제 시스템은 많은 변화를 겪고 있다. 1990년대에는 현금이 주요한 결제 수단이었으나 2000년대에는 신용카드가 주요한 결제 수단이었다. 2008년 금융위기 이후 모바일 기술의 발달로 핀테크(FinTech) 산업이 성장하였으며, 핀테크 산업의 주요 발전 분야 중 하나가 결제 시스템이다. 간편결제 서비스는 온·오프라인에서 신용카드나 계좌정보를 스마트폰 앱 등에 미리 등록해 지문 인식, 안면인식 등을 통해 개인정보를 확인해 결제하는 서비스를 통칭한다. 최근 카카오페이, 네이버페이 등의 정보통신기반 전자금융업자와 삼성페이, SSG페이 등과 같은 유통·제조 기반 전자금융업자 등 다양한 시장 참여자가 경쟁적으로 간편결제 서비스를 출시하였다.

소비자 입장에서 간편결제 서비스는 다른 결제 시스템에 비해 매우 편리한 결제 환경을 제공한다. 스마트폰만 있으면 쇼핑, 배달음식 주문, 보험가입, 은행 서비스 등 실생활에 필요한 대부분의 서비스를 이용할 수 있다. 또한 기술의 발달로 지문인식, 홍채인식, 안면인식을 통해 개인을 식별할 수 있어 결제 단계에서의 보안이 크게 강화되었다. 이에 다른 결제 시스템보다 안전하게 이용할 수 있다는 점도 간편결제 서비스의 성장에 큰 요인으로 작용하였다. 기업들은 경쟁적으로 간편결제 시스템을 도입하고 적극적인 프로모션을 통해 소비자들을 유치하였다. 간편결제 시스템 공급자들은 간편결제 시스템을 통해 획득한 소비자의 기호, 구매패턴 등의 빅데이터를 수집 하여 이를 기반으로 새로운 비즈니스 모델을 구축하는 등 기존 비즈니스 모델을 고도화 하는데 다시 활용하는 등 핀테크 산업 발전의 선순환을 이끌었다. 기업들은 자사가 구축한 하나의 플랫폼 안에서 소비자가 원하는 모든 것을 가능하게끔 환경을 구축하고, 한번 이용한 고객들을 충성심이 높은 고객으로 전환하려 하였다.

정부도 간편결제 서비스 확장에 많은 역할을 하였다. 특히 2015년 4월 공인인증서 의무 사용을 폐지하여 Active X 기반의 공인인증서 모듈을 사용하지 않은 간편결제 시스템이 금융감독원의 보안성 심의기준을 통과 하는 등 제도적인 지원도 간편결제 시장의 성장에 기여하였다. 이후 많은 기업들이 간편결제 서비스 시스템을 구축하였으며 정부도 제로페이(Zero Pay) 출시하는 등 적극적으로 간편결제 시스템을 도입하고 정책 장벽을 낮추었다.

간편결제 서비스 산업은 최근 급성장하였다. 표 1은 한국은행에서 발표한 2019년 상반기 중 전자지급 서비스 이용현황 보도자료 중 간편결제 서비스 이용 추이 통계이다. 한국은행에 따르면 간편결제 서비스의 일평균 이용건수는 2019년 상반기 기준 약 535만건, 1,628억원으로 작년 하반기 대비 약 18.2%, 15.8% 상승하였다. 간편결제 서비스가 도입된 이래로 꾸준한 상승을 유지하고 있음을 확인할 수 있다.

본 연구에서는 최근 금융산업에 나타난 간편결제 서비스의 성장 요인을 분석하고자 한다. 지속적인 성장을 보이고 있는 간

편결제 서비스의 이용에 미치는 요인을 분석함으로써 향후 간편결제 서비스 산업의 발전을 위해 필요한 요인이 무엇인지를 파악하고자한다. 구체적으로 연구에서는 소비자들이 간편결제 서비스를 이용할 때 어떤 요인이 가장 중요했는지를 파악하고자한다. 간편결제 서비스 이용에 미치는 개인 특성 및 제반 특성을 파악함으로써 어떤 요인이 간편결제 서비스의 성장에 기여하고, 어떤 요인이 장애가 되는지를 파악하여, 정책적 시사점을 제시하고자 한다.

표 1. 간편결제서비스 이용 추이(일평균)

Table 1. Trend of simple payment

	2017		2018		2019
	First half	Latter half	First half	Latter half	First half
No of usages	1,601	2,566	3,317	4,522	5,345
Information and Communication	407	489	601	810	1,016
Distribution and Manufacture	1,194	2,078	2,715	3,712	4,331
amount of money	50,748	843,33	111,745	140,679	162,847
Information and Communication	12,824	177,84	21,335	27,125	33,235
Distribution and Manufacture	37,923	66,549	90,410	113,554	129,613

II. 선행연구의 검토

최근 급 성장한 분야인 만큼 많은 연구에서 간편결제 서비스에 대해 분석하였다. 다만, 대부분의 선행연구가 간편결제 시스템에 대한 영향요인 및 수용도에 대한 분석에 집중되어 있었다. 황보충 외(2016)[1]에서는 모바일 간편결제 서비스의 지속적 사용의도에 영향을 미치는 요인을 분석하였다. 설문조사 데이터를 이용하여 자기 효능감, 적합성 등 사용자 특성과 가시성, 외부영향 등 사회적 특성, 편재성 등 모바일 특성이 지각된 유용성에 긍정적인 영향을 미친 것으로 분석하였다. 진로와 이동명(2017)[2]은 중국에 거주하는 모바일 간편결제 서비스 이용자를 대상으로 한 설문조사 자료를 이용하여 모바일 간편결제 서비스의 수용의도에 미치는 요인을 분석하였다. 분석결과 온라인 자기 효능감, 접근성, 적합성 등이 서비스 만족도에 양의 영향을 미친다고 분석하였다. 이성준(2019)[3]에서는 플랫폼 기반 간편결제 서비스로의 전환의도에 미치는 요인에 대하여 분석하였다. 플랫폼 기반 서비스 중에서 카카오페이 자료를 사용하였다. 분석결과 밀집 요인으로서 기존 시스템에 대한 불만족, 당김 요인들로서 인지된 간편성 및 서비스 신뢰, 그리고 계류 요인들로서 감성적·인지적 관성 및 비효율적 주관적 규범이 영향을 미치는 것으로 분석되었다. 박상우 외(2020)[4]에서는 소비자의 불확실성 회피성향을 모형에 고려하여 간편결제 수용의지에 미치는 영향을 분석하였다. 49개 국가의 여섯 가지 문화차원에 대한 국가별 점수를 분석에 이용하였다. 분석결과 집단 불확실성 회피성향과 개인 특성으로의 불확실성 회피성향 모두 모바일 간편결제 서비스 수용에 영향을 미치는 것으로 도

출하였다.

본 연구에서는 기존 선행연구에서 간편결제 서비스 이용 및 수용에 미치는 영향을 기존 연구에서 이용한 변수 외 다양한 설명변수를 추가하여 보다 고도화된 방법으로 접근하였다.

이를 위해 본 연구에서는 주요 분석 방법으로 LASSO(Least Absolute Shrinkage and Selection Operator)를 이용하였다. LASSO 모형은 예측이 필요한 다양한 분야에 활용되고 있다. 강수연 외(2015)[5]에서는 LASSO 모형을 포함한 다양한 머신러닝 모형을 적용하여 국내 드라마 시청률 예측모형을 제시하였다. 드라마 방영 전에 수집되는 기본 정보를 이용해 드라마 초반 시청률을 예측하는데 유의한 변수를 파악하였다. 김미림 외(2017)[6]에서는 유사한 방식으로 국내 예능 시청률 예측 모형을 제시하고, 예측에 주요한 변수를 머신러닝 기법을 이용해 찾았다. 실시간으로 데이터가 수집되어 많은 양의 데이터가 활용가능하고, 예측이 중요한 금융산업에서는 LASSO 모형을 이용하여 대출 부도, 신용카드 부도, 주가 등을 예측하고 있다. 윤종문(2019)[7]에서는 신용카드 부도위험을 예측하기 위해 LASSO를 포함한 머신러닝 방법론과 딥러닝 신경망 모형을 이용했다. 박재연 외(2016)[8]에서는 LASSO를 포함한 머신러닝 및 딥러닝 방법론으로 KOSPI주가지수를 예측하였다. 차성재와 강정석(2018)[9]에서는 딥러닝 알고리즘을 적용하여 기업부도예측 모형을 제시하였다. 고차원 자료의 차원 축소를 위해 기업의 재무정보 중 기업부도를 예측하는데 유의한 변수를 찾는 것이 선행되어야 했고, 이를 위해 LASSO 모형을 이용하였다.

본 연구에서는 기존 선행연구에서 예측과 주요 지표 선별에 가장 많이 이용되는 LASSO 모형을 적용하여 간편결제 사용자의 이용 의도에 영향을 미치는 유의한 요인을 파악하고자 한다. 기존에 수집된고객 자료를 바탕으로 간편결제 이용에 영향을 미치는 유의한 요인을 파악하여, 새로운 고객들의 특성을 바탕으로 간편결제 서비스를 얼마나 이용할지를 예측할 수 있다. 또한, 조훈 외(2020)[10]과 마찬가지로 요인 변수를 특성 별로 그룹화하여 분석하였다. 이를 통해 어떤 특성을 가진 요인 변수 집단이 간편결제 서비스 이용 의도에 주된 요인인지를 파악하고자 한다. 이때, 어떤 요인들에 의해 가장 크게 영향을 받는지를 파악하고 있다면 간편결제 서비스 산업을 부흥시키기 위해 정책 당국과 기업어떤 요인을 고려해야 하는지에 대해 의미 있는 정보를 제시할 수 있다. 이러한 점에서 본 연구는 정책적, 실무적 공헌도가 있을 것이다.

III. 분석자료 및 연구가설

3-1 연구가설

본 연구에서는 최근 성장하고 있는 간편결제 서비스 이용 여부에 미치는 영향 요인을 파악하고자 한다. 간편결제 서비스 이용에 미치는 요인은 다양한 채널에서 발생할 수 있다. 본 연구에서는 간편결제 서비스 이용에 영향을 주는 요인들을 크게 인

구학적 특성, 온라인 서비스의 이용 및 노출 정도, 온라인 콘텐츠의 수용도로 구분하여 어떤 매커니즘을 통해 소비자들이 간편결제 서비스를 이용하고 있는지를 파악하고자 한다. 그림 1에 본 연구의 연구모형을 도식화하여 나타내었다.

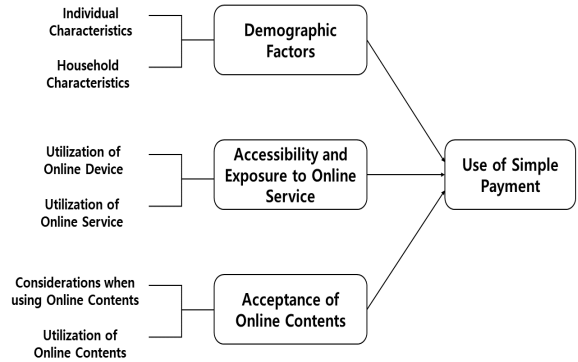


그림 1. 연구 모형
Fig. 1. Research Design

본 연구에서는 분석 대상인 간편결제 서비스 이용 의도를 설명하기 위해 ‘간편결제 서비스 이용 여부’를 종속변수로 삼았다. 간편결제 서비스를 한번이라도 이용한 적이 있는지를 변수화하여 연구모형에 활용하였다. 또한, 소비자들이 간편결제 서비스 이용에 대한 의사결정과정에 크게 3가지 특성이 영향을 미칠 것이라고 가정하였다. 3가지 특성 중 첫 번째인 인구학적 인 특성(Demographic Factors)에는 학력, 소득, 성별, 나이 건강 등의 개인특성과 거주지역, 주거형태, 결혼, 자녀유무 등의 가구특성과 관련된 요인들을 포함한다. 두 번째 온라인 서비스의 이용 및 노출 정도(Accessibility and Exposure to Online Service)는 현재 소비자들이 온라인 서비스를 이용하는데 필요한 제반 환경이 갖추어져 있는지, 실제 인터넷 서비스를 얼마나 이용하고 있는지 등의 변수를 포함한다. 구체적으로 이동통신가입여부, 스마트폰유무, PC 인터넷 사용시간, 모바일 인터넷 사용시간 등의 변수가 이용되었다. 마지막으로 온라인 콘텐츠 수용도(Acceptance of Online Contents)는 모바일 어플리케이션과 같은 콘텐츠를 사용함에 있어서 주로 고려하는 사항이 무엇인지, O2O 어플리케이션이나 SNS를 얼마나 활용하고 있는지 등의 변수를 포함한다.

이렇게 구분된 세 유형의 특성 중에서 간편결제 서비스 이용에 어떤 요인이 영향을 주는지를 비교하여 파악하고자 한다. 이를 위해 본 연구에서는 다음의 두 가지 연구 문제에 대해 살펴보고자 한다.

1. 간편결제 서비스 이용에 어떤 개별적인 채널(특성 집단)이 가장 많은 영향을 미치는가?
2. 두 가지 이상의 채널(특성 집단)이 동시에 고려되었을 때 간편결제 서비스 이용 의도를 더 잘 설명할 수 있는가?

즉, 구분된 특성 집단에 포함되어 있는 다양한 요인들의 개

별적인 영향 정도를 파악하여 어느 특성 집단에 속하는 요인들이 간편결제 사용자의 이용 의도를 가장 잘 설명하는지를 비교·분석한다. 나아가 여러 특성집단이 동시에 영향을 미칠 수 있다는 점을 반영하여, 여러 특성집단에 속하는 변수들이 상호작용하여 복합적으로 간편결제 이용 의도에 영향을 미치고 있는지를 분석하고자 한다.

3-2 분석자료

본 연구에서는 한국방송광고진흥공사의 소비자행태조사 자료를 분석에 사용하였다. 소비자행태조사 자료는 전국 소비자들의 미디어 이용행태, 광고에 대한 반응, 제품이용 및 구매형태, 라이프 스타일 등에 대한 설문조사를 통해 구축된 자료이다. 소비자행태조사는 연 1회, 개별 면접방식의 설문조사로 진행되며 전국 만 13세에서 만 64세 남·여 5,000명을 대상으로 한다. 2017년까지 5,000명을 대상으로 시행되었으며, 2018년부터 조사표본이 4,000명으로 축소되었다. 인구센서스에 기반한 할당 추출을 통해 표본을 구성하며 추적조사가 아닌 횡단면(Cross-Section) 자료의 형태이다. 소비자행태조사는 1999년부터 시행되었으나 간편결제 관련 문항은 2017년부터 반영되어 조사되었다. 이에 본 연구에서는 2017~2019년 3개년 자료를 분석에 이용하였으며 3개 연도 자료에서 동일하게 수집된 변수만을 분석에 이용하였다. 사용된 변수는 총 65개이며 분석에 이용된 표본은 3개 연도의 조사표본 13,000개 중 결측치를 포함한 일부 표본을 제외한 12,141개(2017:4,671개; 2018:3,745개; 2019:3,745개)의 자료를 포함한다. 이용된 변수의 형태는 설문 문항에 따라 다양하게 나타난다. 나이, 소득, 어플리케이션 이용 횟수, 인터넷 사용 시간 등은 연속형 변수로 분석에 이용되었고, 만족도 등의 변수는 1~6점 척도 변수로 분석에 이용되었다. 또한, 범주형 변수들은 개별 더미변수(Dummy variable)로 변환하여 분석에 이용하였다. 주요 변수에 대한 기초 통계량은 표 2에서 제시한다.

표 2. 기초통계량
Table 2. Summary Statistics

Variable	Mean	SD	Min	Max
Gender	0.51	0.50	0	1
Age	40.01	14.26	13	69
Married	0.63	0.48	0	1
Household Income(Won)	522.11	197.02	50	1050
No of kids	1.12	0.97	0	5
Using internet or Not	0.67	0.47	0	1
Monthly SNS use	11.39	11.80	0	30
Using simple payment	0.31	0.46	0	1

IV. 분석방법

본 연구에서는 머신러닝(Machine Learning) 기반의 예측 알고리즘을 분석 모형으로 이용한다. 머신러닝 예측 알고리즘은 주어진 데이터를 학습데이터(Training data)와 검증데이터(Test data)로 나누어 학습데이터를 이용해 모형을 추정하고, 추정된 결과를 검증데이터에 예측하여 예측 성과를 살펴보는 분석 방법론이다. 본 연구에서는 소비자들의 간편결제 이용 여부를 예측 대상으로 삼아 머신러닝 예측 모형을 구성하고자 한다.

본 연구에서 사용한 예측 모형 선형모형(Linear model) 기반 머신러닝 방법론 중에 하나인 LASSO이다. LASSO는 Tibshiran(1996)이 개발한 차원 축소 방법으로 수많은 독립변수 중에서 종속변수에 영향을 미치는 최적의 변수를 도출하는 방법이다.[11] 광범위한 설문조사 자료와 같이 많은 정보들이 수집된 고차원 자료에서 어떤 정보가 유의미한 정보인지를 파악할 때 유용성이 높은 방법론이다. 이에 차원 축소가 필요한 다양한 연구 분야에서 사용되고 있다. 특히 종속변수를 예측하는 것이 중요한 연구 분야에서는 어떤 설명변수가 종속변수에 영향을 주는 지 파악하고, 이를 바탕으로 종속변수를 예측하는데 주로 활용된다. LASSO 모형도 전통적인 계량경제학에서 주로 쓰이는 선형회귀모형과 유사하게 평균제곱오차(Mean squared error)를 최소화하는 계수추정치를 찾는 방식을 이용하고 있으나, 비용함수(Cost Function)에 계수추정치에 대한 제약 조건이 추가된다는 점에서 차이점이 있다. 식(1)은 일반적인 LASSO 모형의 비용함수이다.

$$Min \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{i,j})^2 + \lambda \sum_{i=1}^p |\beta_j| \quad (1)$$

일반적인 선형회귀모형의 비용함수에 새로운 항이 더해져 있는 형태로, 이 항은 식(1)과 같은 비용함수를 최소화함에 있어서 제약으로 작용한다. 이에 식(1)을 최소화하는 계수추정치를 추정한다는 것은 모든 계수추정치의 절대 값의 합이 일정 수준 이하가 되도록 제약을 가한 상태에서 계수 추정치를 추정하는 것이다. 식 (1)의 λ 는 제약의 수준을 결정하는 정규화 파라미터(Regularization parameter)로서 λ 의 크기에 따라 추정 계수에 가해지는 제약의 크기가 달라지며, 이에 추정 계수 값이 변화하게 된다. 식(1)에 가해진 제약은 모든 계수 추정치의 절대 값의 합이 일정 수준 이하가 되도록 요구하기 때문에 계수추정치들이 실제 가져야하는 값보다 축소되어 추정되고, 극단적인 경우에는 0으로 수축되는 경우도 발생하게 된다. 따라서 LASSO 방법론을 적용하게 되면 종속변수에 미치는 영향 정도가 큰 변수들은 제약을 가했음에도 계수추정치가 크게 추정될 것이고, 영향정도가 약한 변수들은 제약을 가했을 때 계수추정치가 작아지거나 0으로 수축될 것이다. 이러한 과정을 통해 LASSO 모형은 종속변수에 영향을 미치는 주요 변수들만을 포함하는 모형을 구성 할 수 있다. 이렇게 계수 추정단계에서 제

약을 가하는 이유는 본 모형이 ‘예측’을 위해 이용되기 때문이다. 예측에 중요하지 않은 변수가 종속변수에 미치는 영향을 축소시키도록 모형을 추정함으로써 새로운 데이터 셋(out-of-sample)에서의 예측력을 높일 수 있다. 이에 정규화 파라미터 λ 의 값을 결정할 때도 분석 모형의 목적이 ‘예측’을 위한 것임을 인지하고 외표본에서의 예측력을 가장 높일 수 있는 수준의 λ 로 제약 정도를 정할 필요가 있다. 이에 본 논문에서는 데이터를 기반으로 예측 성과를 가장 높일 수 있는 파라미터를 찾는 알고리즘을 구성하고 이를 적용하여 간편결제 서비스 이용을 예측하는 데 주요한 지표들을 찾고자 한다.

먼저 본 분석에서는 데이터를 8:2의 비율로 나누어서 사용한다. 전체 데이터의 80%를 학습데이터(Training data)로 정의하고 식(1)과 같은 비용함수를 최소화하는 계수를 추정한다. 우선 학습데이터 중 75%의 표본만을 이용하여 모형을 추정한다. λ 의 값을 임의로 부여한 후 도출된 계수 값을 통해 남겨진 학습데이터 25%에서의 맞춤값(Fitted Value)을 도출한다. 종속변수가 이항변수(Binary variable)인 경우에는 맞춤값이 확률로 나타나게 된다. 본 연구의 종속변수는 간편결제 서비스 이용 여부이므로 추정된 계수를 이용한 맞춤값이 간편결제 서비스 이용 확률을 의미한다. 이때, 추정된 맞춤값이 항상 0과 1사이의 값을 가지도록 로지스틱함수(Logistic function)으로 맞춤값을 변환하였다.

이렇게 추정된 간편결제 서비스 이용 확률을 어느 수준을 기준으로 1로 할당할 것인지에 따라 본 모형의 예측력이 달라지게 된다. 이에 0에서 1사이에서 임계값(Threshold)을 다양하게 조정하여 수신기 동작 특성(Receiver Operating Characteristics; ROC) 곡선을 그리고 그 아래의 면적인 AUC(Area Under the ROC curve)값을 도출한다. 그리고 해당 ROC 곡선 상에서 예측 성과(특이도와 민감도)를 가장 크게 만들어주는 최적 임계값(Optimal Threshold)를 찾을 수 있다. 이 과정을 통해 모든 λ 값에 대하여 특정 λ 값에서의 최적 임계값을 찾을 수 있다. λ 값을 바꿔가며 전체 과정을 반복 시행하게 되면 λ 값과 임계값의 조합마다의 예측의 정확성을 계산하여 비교할 수 있고, 가장 높은 예측성과를 보이는 λ 와 임계값의 조합을 선택할 수 있다. 이러한 과정을 통해 위 모형은 첫째, 외표본에서의 예측성과를 높여주는 계수추정치들을 제시함으로써 어떤 변수가 예측에 유의한지를 파악하게 한다. 둘째, 외표본에서의 예측성과를 가장 높여주는 제약 수준을 알 수 있다. 셋째, 외표본에서의 예측성과를 가장 높일 수 있는 최적 임계값을 제시한다.

최종적으로 학습데이터에서 선택된 계수추정치와 정규화 파라미터, 최적 임계값을 이용하여 검증데이터(Test data)에서의 모형의 예측력을 평가한다.

본 연구에서는 앞서 그림 1에서 제시한 특성 집단을 이용하여 7개의 모형을 제시하여 위 과정을 동일하게 수행하고, 검증 데이터를 이용하여 7개 모형의 예측 성과를 비교함으로써 본 연구의 연구문제에 대한 답을 찾고자 한다. 분석에 이용된 7개의 모형은 다음과 같다.

모형 1: 인구특성학적 요인(A)만 포함

모형 2: 온라인 이용환경 및 이용 정도 변수(B)만 포함

모형 3: 온라인 콘텐츠 수용도 변수(C)만 포함

모형 4: A+B 함께 포함

모형 5: A+C 함께 포함

모형 6: B+C 함께 포함

모형 7: A+B+C 모두 포함

V. 분석결과

연구가설에 입각하여 앞서 제시된 7개 모형의 분석결과를 제시하고자 한다. 먼저 간편결제 서비스 이용에 어떤 특성이 개별적으로 영향을 받는지에 대한 분석결과는 표 3에 제시하였다.

표 3. 각각의 특성별 분석결과

Table 3. Individual characteristics Results

	Model 1	Model 2	Model 3
Optimal lambda	0.035	0.020	0.096
Optimal threshold	0.358	0.302	0.302
Test set accuracy at optimal threshold & lambda	0.722	0.691	0.727
AUC(test set)	0.748	0.736	0.690
No. of Not selected vars	22	5	26
No. of Selected vars	5	10	1
Total No. of vars in the model	27	15	27

학습과정에서 선택된 최적 λ 와 최적 임계값은 개별 모형에 따라 다르게 추정되었다. 선택된 최적 제약수준(λ)하에서 추정된 계수를 바탕으로 어떤 요인이 선택되었는지가 결정이 되었고, 이를 바탕으로 검증데이터 표본에 예측한 결과를 예측 정확도(Accuracy)와 AUC 값으로 제시하였다. 정확도는 학습데이터에서 선택된 특정 임계값을 기준으로 판단한 결과이며 학습 데이터를 통해 추정된 계수와 파라미터의 예측력을 비교하는 기준이 된다. 그러나 본 연구에 이용된 종속변수와 같이 0과 1의 분포가 균형적이지 않은 경우에 정확도만을 이용하여 모형을 평가하는 것은 바람직 하지 않다. 이에 AUC 값을 이용하여 해당 모형이 종속변수를 0과 1로 얼마나 잘 분류하는지를 평가할 필요가 있다. AUC 값은 검증데이터에서 가능한 모든 임계값에서의 예측력을 전반적으로 보여준다.

그림 2에서 본 연구에서 추정된 7개 모형 중 AUC값이 가장 높을 때, 즉 예측 모형의 성능이 가장 좋을 때와 AUC 값이 가장 낮을 때, 즉 모형의 성능이 가장 낮을 때의 ROC 곡선과 해당되는 AUC 값을 제시한다. 그림 2에서 Y축의 민감도(Sensitivity)는 실제 간편결제를 이용한 사람 중 해당 모형에서 간편결제를 이용할 것으로 예측되는 사람의 비율을 의미하며, 이 수치가 높

을수록 예측 성과가 좋은 모형이다. X축은 (1-특이도)를 나타내고 있다. 특이도(Specificity)는 실제로 간편결제를 이용하지 않은 사람 중 해당모형에서 간편결제를 이용하지 않을 것으로 예측되는 사람의 비율을 의미한다. 즉, 특이도가 높을수록 예측성과가 좋은 모형을 의미하며, 그림 2에서는 X축의 값이 작을수록 예측성과가 높음을 의미한다. 민감도와 특이도는 임계값에 따라 결정되며, 모든 임계값에서 더 높은 민감도와 특이도를 보이는 모형이 예측 성능이 좋은 모형으로 평가된다. 즉, ROC 곡선이 좌측 상단으로 치우쳐져 있어, 아래 면적인 AUC 값이 더 클수록 좋은 모형이라는 의미로 해석된다.

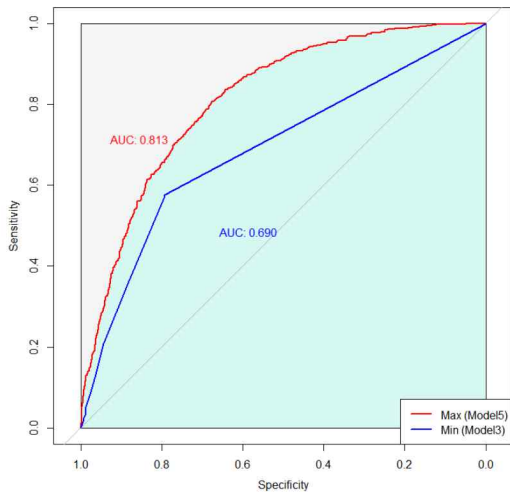


그림 2. 모형별 ROC 곡선
Fig. 2. ROC curves by model

표 3에서는 모형 별로 몇 개의 요인변수가 최종적으로 선택되었는지를 함께 제시하였다. 개별적인 특성집단만 이용하여 예측모형을 구성한 1,2,3번 모형의 예측성과를 비교하였을 때 온라인 콘텐츠 수용도만을 예측 변수로 이용한 모형3의 예측정확성이 72.7%로 가장 높게 나왔으나 검증데이터에서의 AUC 값이 가장 낮았고, 선택된 변수가 1개 밖에 되지 않았다. 이때 선택된 변수는 월평균 배달앱 이용 횟수였다. 다음으로 예측성과가 가장 높은 모형 1의 경우 5개의 변수가 선택되었으며 이는 대졸 여부, 거주지역, 자가 주택유무, 나이 등을 포함하였다.

다음으로 특성집단을 2개씩 조합하여 예측 모형에 포함한 분석결과를 표 4에 제시한다. 앞선 표 3에서 개별 특성집단을 이용했을 때보다 전반적으로 예측 성과가 높아짐을 확인할 수 있었고, 선택된 예측 변수의 개수도 증가함을 알 수 있었다. 4,5,6 번 모형의 예측정확성 및 AUC값을 비교해보면 인구특성학적 요인 변수가 포함되는 경우에 예측성과가 더 높아진다는 것을 알 수 있다. 특히 인구특성요인과 온라인 콘텐츠 수용도 관련 변수를 포함하는 모형 5의 경우 예측성과가 74.1%로 가장 높게 나타났고, AUC 값도 0.813으로 타 모형에 비해 높은 수준으로 나타났다. 또한, 모형에 포함된 총 52개 변수 중 39개가 선

택되었다. 선택된 인구특성학적 요인은 학력, 카드지출액, 거주지역, 성별, 나이, 건강상태, 소득 등이며 온라인 콘텐츠 수용도 관련 변수는 모바일 콘텐츠 이용, 모바일 티비 시청여부, SNS 이용 정도, 배달앱 이용정도, 인터넷 쇼핑 선호도 등 대부분의 변수가 선택되었다. 구체적인 선택 변수 목록은 그림 3에서 제시하도록 한다.

표 4. 특성 조합의 분석결과
Table 4. Combinations of characteristics Results

	Model 4	Model 5	Model 6
Optimal lambda	0.003	0.005	0.043
Optimal threshold	0.323	0.330	0.320
Test set accuracy at optimal threshold & lambda	0.737	0.741	0.735
AUC(test set)	0.801	0.813	0.786
No. of Not selected vars	9	13	28
No. of Selected vars	31	39	12
Total No. of vars in the model	40	52	40

마지막으로 모든 특성집단의 요인을 포함하여 분석한 모형 7의 분석 결과를 표 5에 제시한다. 모든 특성집단을 이용하였을 때의 검증데이터 예측 정확도가 75.3%로 다른 모형에 비해 가장 높다는 것을 확인할 수 있었다. 또한 총 65개중 51개의 변수가 선택되었다는 것을 감안할 때, 인구특성학적 요인과 온라인 콘텐츠 수용도 관련 변수를 이용한 모형 5에 온라인 사용 환경과 노출도 변수를 추가하는 경우 예측성과를 더 높이는 방향으로 모형이 개선되었음을 알 수 있었다. 이를 통해 간편결제 서비스 이용 의도에 미치는 영향을 파악함에 있어서 모든 특성집단 요인 변수를 모두 고려하는 것이 의미있는 작업을 확인하였고, 소비자의 간편결제 서비스 이용에는 세 유형의 특성집단이 복합적으로 영향을 미치고 있음을 파악할 수 있었다.

표 5. 모든 특성의 분석결과
Table 5. All characteristics Results

	Model 7
Optimal lambda	0.004
Optimal threshold	0.341
Test set accuracy at optimal threshold & lambda	0.753
AUC(test set)	0.812
No. of Not selected vars	14
No. of Selected vars	51
Total No. of vars in the model	65*

*17 demographic characteristics, 10 online accessibility variables, 22 online contents acceptance variables and 2 year dummies are selected.

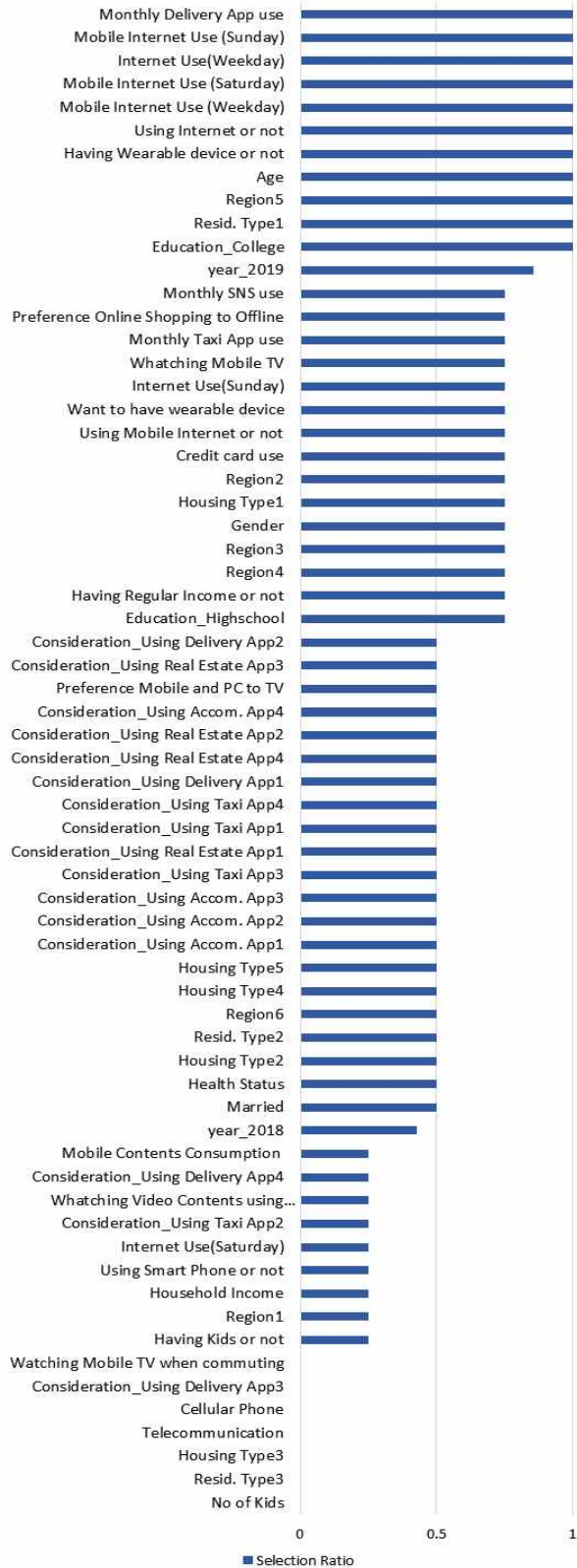


그림 3. 변수별 선택비율
Fig. 3. Selection Ratio by Variable

마지막으로 그림 3에서 분석에 활용된 모든 요인변수 목록과 해당 변수가 얼마나 자주 주요 변수로 선택이 되었는지를 그래프로 나타내었다. 세로축은 분석에 이용된 모든 변수의 목록이고 가로축은 변수 선택 비율(Selection ratio)이다. 각 변수는 최대 4개의 모형에 이용되었다. 그 중 몇 개의 모형에서 선택되었는지를 비율로 계산하여 막대그래프로 나타내었다. 모형에 포함될 때마다 선택되는 경우 이 비율이 1의 값을 가지며, 한번도 선택되지 않은 경우 0의 값을 가진다. 즉, 이 비율이 높을수록 간편결제 서비스 이용의도에 영향을 미치는 주요 요인변수라고 할 수 있다. 4번 모두 선택된 변수는 11개이며 3가지의 특성 중에서 온라인 이용환경 및 이용정도가 6개로 가장 많은 변수를 포함하고 있었다. 특히 모바일 기기와 모바일을 이용한 인터넷 사용 관련 변수들이 반복적으로 선택되었다. 이는 간편결제 서비스 이용 촉진에 여건 조성이 중요하다는 것을 의미한다. 한번도 선택되지 않은 변수는 7개이다. 3개 특성 중에서 인구특성 요소가 가장 많이 선택 되지 않았다. 자녀수, 주거 형태 등의 변수 등은 간편결제 서비스 이용에는 영향을 주지 않는 것으로 보여진다.

VI. 결론

본 연구에서는 크게 두 가지 가설을 검증함으로써 최근 급 성장하고 있는 간편결제 서비스의 이용 요인에 대한 분석을 하였다. 우선, 간편결제 서비스 이용에 어떤 개별적인 채널이 가장 많은 영향을 미치는지를 분석한 결과, 온라인 이용 환경 및 이용 정도 채널이 가장 많은 영향을 미친다는 것을 확인하였다. 다음으로, 두 가지 이상의 채널이 동시에 고려되었을 때 간편결제 서비스 이용 의도에 대한 설명력이 증가하는지 살펴본 결과 독립적인 채널만 고려한 모형보다 세 가지 채널 모두를 고려한 모형이 간편결제 서비스 이용 의도를 더욱 잘 설명하는 것으로 나타났다. 이는 간편결제 서비스 사용의도에 미치는 영향 요인이 모든 특성집단에 고르게 있다는 것을 알 수 있다. 이는 소비자들이 간편결제 서비스를 이용함에 있어서 다수의 요인이 복합적으로 작용하고 있음을 시사한다.

본 연구의 결과를 바탕으로 다음과 같은 정책적 시사점을 제안한다. 간편결제 서비스 시장의 활성화를 위해서는 무엇보다도 모바일 환경에 대한 경험을 향상 시키는 정책 및 전략이 필요하다. 간편결제 서비스를 이용한 경험이 없거나 이용 빈도가 낮은 소비자에게 다양한 혜택을 통해 초기 경험 및 반복적 사용을 할 수 있는 정책 및 마케팅 전략이 수립된다면, 반복적인 사용으로 간편결제 서비스의 가장 큰 장점인 간편성 및 안정성을 소비자들이 인식하여 간편결제 시장이 지속적으로 발전 할 수 있을 것으로 생각된다.

기존 선행연구에서는 제한된 변수만을 이용하였으나 본 연구에서는 머신러닝의 변수 선택 및 예측 모형인 LASSO를 이용하여 수많은 변수 중 가장 유의한 변수들을 선택할 수 있었다.

이러한 방법론적 유용성을 이용하여 보다 다양한 변수를 동시에 고려하여 간편결제 서비스 사용자의 이용 의도를 분석했다는 점에서 기존연구와 차별점을 갖는다. 다만, 본 연구는 간편결제 서비스가 다양한 산업에 다양한 방법으로 도입되고 있음에도 불구하고, 자료의 제약으로 인해 간편결제 서비스의 이용 유·무만으로 연구를 진행했다는 점에서 한계가 있다. 추후에 간편결제 서비스를 세분화하여 분석할 수 있다면 간편결제 서비스 시장 성장을 위한 구체적인 시사점을 제안할 수 있을 것이다.

참고문헌

- [1] B. C. Hwang, K. S. Chung, and M. J. Noh, "An Empirical Study on Relationships between Characteristics of Mobile Payment Services and Continuance Intention," *Korean Business Education Review*, Vol. 31, No. 4, pp.591-615, August 2016.
- [2] C. Lu, and D. M. Lee, "An Empirical Study on the Factors Influencing the Acceptance of Mobile Easy Payment Service: A Case of Chinese User," *Journal of the Korea Contents Association*, Vol. 17, No. 8, pp.1-13, August 2017.
- [3] S. J. Lee, "An Analysis of Factors Influencing Switching Intention toward Online Platform-based Easy Payment Service with Moderating Effects of Policy Expectations: Focusing on Kakao Pay," *Journal of the Korea Contents Association*, Vol. 19, No. 5, pp.426-442, May 2019.
- [4] S. W. Park, S. H. Kim, and W. J. Choi, "The Effect of Consumers' Uncertainty Avoidance on Their Acceptance to Use Mobile Payments," *Korean Marketing Association*, Vol. 35, No. 1, pp.53-68, February 2020
- [5] S. Y. Kang, H. J. Jeon, J. H. Kim, and J. W. Song, "A Study on Domestic Drama Rating Prediction," *The Korean Journal of Applied Statistics*, Vol. 28, No. 5, pp.933-949, October 2015.
- [6] M. L. Kim, S. Y. Lim, C. H. Jang, and J. W. Song, "A study on entertainment TV show ratings and the number of episodes prediction," *The Korean Journal of Applied Statistics*, Vol. 30, No. 6, pp.809-825, December 2017.
- [7] J. M. Yoon, "Effectiveness Analysis of Credit Card Default Risk with Deep Learning Neural Network," *Journal of Money & Finance*, Vol. 33, No. 1, pp.151-183, March 2019.
- [8] J. Y. Park, J. P. Ryu, and H. J. Shin, "Predicting KOSPI Stock Index using Machine Learning Algorithms with Technical Indicators," *Journal of Information Technology and Architecture*, Vol. 13, No. 2, pp.331-340, April 2016.
- [9] S. J. Cha, and J. S. Kang, "Corporate Default Prediction Model Using Deep Learning Time Series Algorithm, RNN and LSTM," *Journal of Intelligence and Information Systems*, Vol. 24, No. 4, pp.1-32, December 2018.
- [10] H. Cho, K. S. Han, Y. J. An, and S. H. Hong, "A Study on the Factors Affecting the Satisfaction and Usage Intention of Ultra High Definition VR Music Performance Video," *Journal of Digital Contents Society*, Vol. 21, No. 3, pp.539-549, March 2020.
- [11] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B*, Vol. 58, No. 1, pp.267-288, January 1996.



박희대(Hee-Dae Park)

2014년 : 서울대학교 대학원 (경제학석사)

2019년 : 서울대학교 대학원 (경제학박사)

2019년~현 재: 한국해양수산개발원

※관심분야 : 머신러닝(Machine Learning), 구조적 추정(Structural Estimation), 수요 추정(Demand Estimation) 등