

국내 연체경험자의 정상변제 요인에 관한 연구

A Study on the Factors of Normal Repayment of Financial Debt Delinquents

최 성 민 (Sungmin Choi) 한국신용정보원 선임조사역
김 호 영 (Hoyoung Kim) 한국신용정보원 선임조사역, 교신저자

요 약

국내 개인신용평가회사들은 과거와 현재 시점의 다양한 금융거래 정보를 활용하여 개인의 신용을 평가하고 있는데, 이 중 과거에 대출을 실행하여 이를 상환 또는 연체한 이력에 대한 정보를 의미하는 ‘상환이력정보’는 신용평가에 활용되는 다른 항목들에 비해 상대적으로 활용 비중이 높은 항목이다. 그러나 개인이 연체된 채무를 모두 변제하여 현재 연체중인 상태가 아닌 경우에도 과거의 연체 이력이 부정적인 요인으로 최장 5년간 평가에 반영되고 있어 금융소비자에게 과도한 불이익을 준다는 지적이 지속적으로 있어 왔다. 실제로 연체 이력이 있는 개인의 경우, 연체된 채무를 성실하게 변제한 개인(정상변제)과 그렇지 않은 개인(비정상변제)으로 구분할 수 있는데, 이들 간에는 신용도의 차이가 존재하므로 ‘정상변제’하는 개인의 특징을 확인하여 이들에게 ‘상환이력정보’의 활용기간을 단축시켜 주는 등의 혜택을 제공하는 것이 바람직하다고 판단된다. 본 연구는 이러한 문제의식에서 출발하여 한국신용정보원에서 보유하고 있는 2019년 12월 말 기준, 개인의 대출·연체·변제 정보에 기반하여 국내 연체경험자의 정상변제 요인을 분석하였다. 방법론은 개인신용평가모형에서 주로 사용하는 로지스틱 회귀모형을 기본으로 하여 의사결정나무, 신경망 모형 등의 머신러닝 방법론을 추가로 활용하였으며, 각 방법론별 성능을 비교해보았다. 실증분석 결과, 연체건수, 대출·연체유형 등이 정상변제 여부에 영향을 미치는 유의한 변수들로 확인되었으며 방법론 중에는 신경망 모형의 성능이 가장 높은 것으로 나타났다. 이러한 연구결과는 연체된 개인 차주의 정상변제 여부에 영향을 미치는 요인을 확인하여 개인신용평가모형을 고도화하는데 도움이 될 수 있을 것으로 보이며 연체 후 성실하게 변제하는 개인을 정책적으로 지원하기 위한 기초자료로도 활용될 수 있을 것으로 보인다. 향후에는 정상변제 요인을 추가 발굴하여 금융업권별 정상변제 요인의 세부적인 차이를 확인하고 이를 실제 모형에 반영하는 연구가 필요할 것이다.

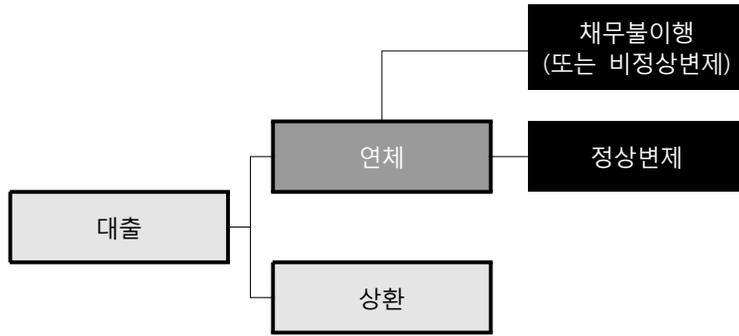
키워드 : 신용정보, 대출, 연체, 정상변제, 채무상환 이력

I. 서 론

1.1 연구배경 및 목적

한국금융연구원(2013)에 따르면, 최근 수년간

국내 가계부채 규모는 장기화된 저금리 기조와 이에 따른 개인들의 대출수요 증가, 금융회사의 이자수익 확대를 위한 가계 여신 비중 확대 등의 이유로 지속적으로 증가하는 추세이다. 그러나 이러한 가계부채의 증가는 연체를 경험하거나 연체된 채



<그림 1> 대출경험자의 대출-연체(상환)-변제 절차

무를 변제하지 못하는 개인도 함께 증가하게 될 위험성이 있다.

현재 개인이 은행 등의 금융회사에서 대출을 받기 위해서는 본인의 신용평가가 결과가 필요한데, 이를 위해 국내 개인신용평가회사(Credit Bureau; CB)에서는 통계적인 방법론을 기반으로 한 자동화된 신용평가모형 시스템을 통해 개인의 신용평점과 등급을 산출하여 금융회사와 본인에게 제공하고 있다. 이때 활용하는 정보에는 상환이력정보, 현재부채수준, 신용거래기간, 신용형태정보 등이 있는데, 이 중 활용비중이 상대적으로 높은 ‘상환이력정보’는 현재의 연체 여부와 과거의 채무 상환이력을 반영하는 정보로서 현재 채무를 모두 변제하여 부채가 없는 개인이라 할지라도 과거에 연체한 이력이 있으면 신용평가지 부정적 요소로서 반영된다.

‘상환이력정보’는 신용정보법령상 개인신용평가회사에서 최대 5년간 활용이 가능¹⁾하므로, 연체 이력이 있는 개인은 연체를 모두 변제하더라도 최대 5년간 연체 이력이 신용평가에 활용되는 불이익을 받아왔으며 이로 인해 상환이력정보의 활용기간이 과도하게 길다는 내용의 민원이 개인신용평가회사와 금융감독원에 지속적으로 제기²⁾되어 왔다.

한편 과거에 3개월 이상 장기 연체한 이력이

존재하여 신용평가지 소위 채무불이행자로 간주되는 개인은 본인의 신용을 관리하기 위해 연체중인 채무를 채무자 개인이 직접 변제하는 경우와 그렇지 않은 경우로 다시 구분할 수 있는데, 특히 채무자 개인이 정상적으로 채무를 변제한 ‘정상변제’의 경우, 개인의 채무상환 의지를 반영하는 중요한 지표로 활용될 수 있다.

대출은 <그림 1>과 같이 실행 이후, 연체 또는 상환으로 진행되며 연체된 이후에는 최종적으로 채무불이행 상태로 진행되거나 정상적인 변제가 이루어진다. 현재 국내외에서 신용정보를 활용한 연구들을 살펴보면, 개인의 금융거래 절차 측면에서 대출이 발생한 이후의 연체 또는 채무불이행 요인에 대한 연구는 지속적으로 이루어져 왔으나 ‘연체 이후 변제’에 대해서는 데이터 확보의 어려움 등의 이유로 연구가 부족한 상황이다.

본 연구에서는 연체 이후, ‘정상변제’로 채무를 상환한 개인의 특징을 살펴보고, 이를 신용평가에 반영할 수 있는 방법에 대해 검토해보았다. 또한 현재 금융당국에서는 연체 이력이 있는 개인이 다시 신용을 회복하여 정상적인 금융활동을 수행할 수 있는 방법에 대한 정책적 고민이 부족한 측면이 있으므로, 실제 개인의 금융거래 데이터로 분석한 본 연구가 기존의 개인신용평가모형을 보완할 수 있는 실무적인 기여 뿐만 아니라, 연체경험자의 신용회복을 위한 정책 수립의 기초자료로도 활용될 수 있을 것이다.

1) 신용정보법 감독규정 제19조(오래된 신용정보의 삭제).
 2) 금융감독원 보도자료(2018.11.16.), 「개인신용평가 민원동향 및 주요 민원사항 Q&A」 발표.

본 연구의 구성은 다음과 같다.

제 I 장 서론에서는 연구의 배경과 목적을 제시하고, 제 II 장 이론적 배경에서는 국내 개인신용정보의 수집·관리·활용체계와 개인신용평가모형의 발전 현황, 개인의 연체·부도·변제 관련 선행연구 동향 등을 서술한다. 한편 제 III 장 연구설계 및 방법론에서는 실증연구를 위한 데이터의 특성, 구성, 분석절차 및 방법론 등을 설명하고, 제 IV 장에서는 국내 연체경험자의 정상변제 요인 확인 및 모형 간 성능비교 등 실증분석 결과를 설명한다. 마지막 제 V 장에서는 본 연구의 결론과 함께 실무적·정책적 시사점을 제시하고 향후 연구과제에 대해 고찰한다.

II. 국내 개인신용평가체계 현황 및 관련 연구동향

2.1 개인신용정보의 개념과 활용

2.1.1 개인신용정보의 정의

개정 신용정보법³⁾에 의하면 신용정보는 ‘금융거래 등 상거래에 있어서 거래 상대방의 신용을 판단할 때 필요한 정보’를 의미한다. 이는 특정 신용정보 주체를 식별할 수 있는 정보, 신용정보 주체의 거래내용을 판단할 수 있는 정보, 신용정보 주체의 신용도를 판단할 수 있는 정보, 신용정보 주체의 신용거래능력을 판단할 수 있는 정보 등을 포함하는 내용이다.

한편 신용정보는 대상에 따라 기업과 개인으로 다시 구분할 수 있는데 이 중 개인신용정보는 ‘기업 및 법인에 관한 정보를 제외한 살아있는 개인에 관한 신용정보’로 신용정보법 상에서 별도로 정의하고 있으며, 신용정보 중 개인의 신용도와 신용거래능력 등을 판단할 때 필요한 정보로서 성명, 주민등록번호 등을 통해 개인을 알아볼 수 있는

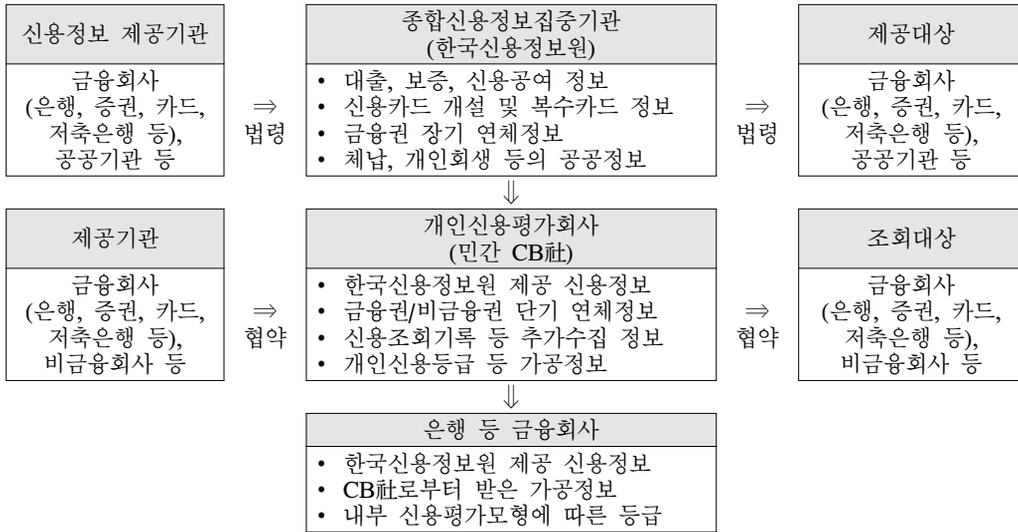
정보를 말한다. 다만, 성명 등 신용도 판단과 무관한 개인에 관한 정보도 다른 개인신용정보와 결합하는 경우, 개인신용정보로 취급한다. 법령상 정의된 개인신용정보의 범위는 <표 1>과 같다.

<표 1> 개인신용정보의 범위

구분	집중관리·활용 대상 정보	
식별정보	• 성명 및 개인 식별번호	
신용거래 정보	대출·당좌거래	• 대출 현황 • 당좌예금 및 가계당좌예금의 개설 및 해지 사실 • 담보 및 채무보증 현황
	신용카드	• 신용카드 발급 및 해지 사실, 결제/미결제 금액 • 복수카드의 신용카드 이용금액, 이용한도, 신용카드 현금 유통한도 등 • 신용카드 분실 및 도난 등의 사고발생, 사고종결에 따른 보상, 사고처리 사실 등
	보험상품	• 보험계약의 체결에 관한 정보 • 보험금의 청구 및 지급에 관한 정보
신용도 판단정보	연체	• 대출금, 신용카드 대금 등의 연체 사실
	대위변제·대지급	• 대위변제 및 대지급 발생 사실
	어음·수표	• 어음 및 수표의 거래정지처분 사실 및 부도 사실
	증권	• 증권의 투자매매업·투자중개업에 관한 정보
	금융질서 문란정보	• 부정 대출 등을 통해 신용질서를 문란하게 한 사실, 거짓이나 부정한 방법으로 신용카드를 발급받거나 사용한 사실, 보험사기 사실 등
	공공정보	• 법원의 회생과 관련된 결정, 파산선고, 면책, 복권과 관련된 결정, 채무불이행자명부 등재, 말소 사실 • 국세, 지방세, 관세 또는 국가채권과 벌금, 과태료, 과징금, 추징금 등의 체납 관련 정보 • 사회보험료, 공공요금, 수수료 등의 관련 정보 • 주민등록 관련 정보 • 다른 법령에 따라 국가, 지방자치단체, 공공기관 등으로부터 받은 행정처분 중 금융거래 정보

자료: 국가법령정보센터(2020).

3) 「신용정보의 이용 및 보호에 관한 법률」 개정안은 국회를 통과(2020.1.9.)하여 8월 5일 시행.



자료: 금융위원회, 금융감독원, 한국신용정보원 재구성.

<그림 2> 개인신용정보의 수집·관리·활용체계

2.1.2 개인신용정보의 수집·관리·활용체계
 현재 국내 개인신용정보의 수집·관리·활용 체계는 신용정보법 상에 명시된 종합신용정보집중기관(Public Credit Registry; PCR)에서 개인신용평가사 활용해야 하는 기초적인 개인신용정보의 수집·관리를 담당하고 있으며 민간 신용정보회사(Credit Bureau; CB)와 금융회사가 종합신용정보집중기관의 정보와 사적 계약에 기초한 추가 수집정보를 활용하여 실제 개인신용평가를 수행하는 형태로 이루어져 있다.

한국금융연구원(2014)에 따르면, 이러한 국내 개인신용정보 수집·관리·활용체계 구조는 PCR과 민간 CB사가 병존하는 방식의 신용정보 유통구조이며, 자발적으로 발달한 민간 CB사가 신용정보의 수집·관리·활용을 모두 담당하고 있는 미국 등 개인금융시장의 역사가 오래된 선진국의 경우나 개인금융시장의 역사가 짧고 금융위기를 경험했던 중남미나 ASEAN 지역의 국가들에서 흔히 보이는 정부 주도의 PCR 인프라 중심 신용정보 유통구조와는 다른 새로운 형태이다.

한편 2016년 1월 설립된 한국신용정보원⁴⁾은 중

합신용정보집중기관으로서 금융회사와 사전에 약속된 계약⁵⁾에 기초하여 대출, 연체 등의 여신 관련 금융거래정보들을 수집하고 있으며 국세청이나 지방자치단체, 대법원 등의 공공기관으로부터 체납 정보나 회생, 파산선고, 면책, 복권 등의 공공정보들도 함께 수집하고 있다. 한국신용정보원은 이러한 정보들을 민간 CB사와 금융회사에 제공하며, 민간 CB사는 한국신용정보원에서 제공받은 자료와 자체 수집한 자료 등을 가공한 데이터, 평가결과 등을 다시 금융회사에 판매하고 있다. 또한 내부적인 개인신용평가 프로세스가 존재하는 금융회사의 경우, 한국신용정보원의 신용정보 기초자료와 민간 CB사로부터 구입한 정보, 개인 고객들로부터 수집한 식별정보와 자사의 거래이력 정보들을 종합하여 자체적인 내부신용등급을 산정하고 이를 여신심사에 활용하고 있다.

<그림 2>는 국내 개인신용정보의 수집·관리·활용체계를 도식화한 것이다.

4) 금융위원회 보도자료(2016.1.5.), “한국신용정보원, 종합신용정보집중기관으로서 첫 발을 떼다.”
 5) 한국신용정보원, 『일반신용정보관리규약』.

2.2 개인신용평가

2.2.1 개인신용평가모형의 발전

Elizabeth Mays(2004)에 따르면, 개인신용평가 모형은 금융소비자의 미래 채무불이행 위험을 평점 형태로 예측하는 통계적·수리적 모형을 말한다. 개인신용평가모형은 1941년 David Durand가 은행, 자동차 할부금융회사 등 여러 금융회사의 개인신용정보를 이용하여 판별분석 방법론에 의해 우·불량⁶⁾ 고객을 구별해낼 수 있는 신용평가모형을 개발⁷⁾한 것을 시작으로 본격적으로 도입되기 시작하였다. Li and Zhong(2012)에 따르면 1960년대 후반부터 신용카드가 소개되면서 신용카드 발급허가와 관련한 신용평가모형이 활용되기 시작하였으며 1980년대부터는 개인대출 뿐만 아니라, 주택구매, 개인사업자 신용평가 등으로 개인신용평가모형의 활용범위가 넓어지기 시작하였다. 그리고 1990년대에 이르러 현재 대다수의 금융회사에서 사용하는 스코어카드(scorecard) 방식의 신용평가모형이 소개되었다.

해외에서는 1989년에 FICO사에서 개인신용평가를 위한 FICO Score가 처음 출시되었으며, 이후 3대 CB사로 불리는 Equifax, Experian, TransUnion이 설립되어 FICO Score에 기반한 신용평가모형을 운영하고 개인별 신용평점을 산출하게 되면서 이를 이용한 금융회사의 개인 여신심사가 이루어지게 되었다.

국내에서는 1990년대 중반부터 개인에 대한 금융회사의 대출 관행이 담보 위주에서 무담보 신용기반의 대출로 변화하면서 개인 차주에 대한 신용평가의 필요성이 높아지게 되었다. 특히 2003년 카드 사태, 신용불량자 급증 등에 대응하여 통계모형에 기반한 개인신용평가가 도입되면서 빠르게 성장해 왔는데, 신용정보 인프라가 강화되고 개인

CB업이 도입⁸⁾되면서 다양한 신용정보에 기초한 개인신용평가체계가 본격적으로 마련되기 시작하였다. 이에 따라 금융회사의 리스크 관리가 개선되었을 뿐만 아니라, 장기·분할상환 여신관행 정착 등 여신구조도 건전화되었으며 신용이력의 중요성이 커짐에 따라 채무자인 개인의 건전한 신용생활 및 신용관리를 유도할 수 있게 되었다.

최성민(2018)에 따르면, 개인신용평가모형 관련 연구는 데이터와 방법론 측면에서 다양한 연구들이 진행되어 왔다. 먼저 데이터 측면에서는 금융거래 정보 외에 온라인 활동 정보, 통신료 및 공공요금 납부정보, 부동산 보유 정보 등으로 평가에 활용되는 데이터의 범위가 확대되는 추세이다. 방법론 측면에서는 도입 초기에 평가자의 주관적 판단에 의존했던 것이 조건표 기준의 평점 평가, 통계적 방법론에 의한 평가모형 등으로 진화해왔으며, 최근에는 머신러닝 방법론을 적용한 평가모형이 등장하기 시작하였다.

또한 Dastile et al.(2020)의 리뷰 논문에 따르면, 최근 신용평가모형 관련 연구 동향은 방법론 측면에서 전통적인 로지스틱 회귀모형 외에 다양한 머신러닝 방법론을 적용한 연구들이 진행되고 있으며, 변수선택 기법, 불균형 데이터 가공, 모형 투명성 등의 이슈에 대해서도 다양한 연구가 진행되어 왔다.

국내의 개인신용평가모형 관련 연구로는 금융거래정보가 부족한 고객(Thin-Filer)에 대한 신용평가시 개인신용평가모형에 전기요금 납부정보를 추가로 활용하여 신용평가모형의 예측력을 향상시킨 권영준 등(2011)의 연구와 함준호 등(2015)의 거시경제지표를 활용한 개인신용평가모형 고도화 연구사례 등이 있다.

이 외에도 PERC(2006)에서는 신용정보 외의 공공요금이나 통신료 납부정보, 온라인 상거래 데이터 등 다양한 비금융정보를 활용하여 전통적인 신

6) 불량은 일반적으로 90일 이상 장기연체를 의미
7) D. Durand(1941), Risk Elements in Consumer Instatement Financing, National Bureau of Economic Research, New York.

8) 2002년 9월 한신정·한신평정(現 NICE평가정보)에서 개인신용평가 업무를 시작하였으며, 2005년 국내 8개 금융사 연합으로 KCB가 설립.

<표 2> 개인신용평가모형의 평가항목 및 평가항목별 활용비중

평가요소	상세내용	활용비중		
		FICO	NICE	KCB
상환이력정보	현재 연체 여부 및 과거 채무상환 이력 등	35.0%	40.3%	28.0%
현재부채수준	현재 보유채무 건수 및 금액 등(대출금액, 신용카드 금액 등)	30.0%	23.0%	26.0%
신용거래기간	최초/최근 개설일로부터의 기간 등	15.0%	10.9%	14.0%
신용형태정보	신용거래 종류 및 형태 등(상품별 건수 및 비중)	10.0%	25.8%	32.0%
신용조회정보*	신규 신용거래를 위한 활동(최근 조회건수 등)	10.0%	-	-
가점 부여	현재 지출수준을 고려한 가처분소득 추정, 신용관리 교육 등 신용개선 노력 여부, 증빙소득 및 비금융정보 성실상환 정보 등	-	+a	+a

* 금융위원회의 「서민금융 기반강화 종합대책(2011.4.15.)」에 따라 신용조회정보는 개인신용평가 평가항목에서 제외. 자료: FICO, NICE평가정보, KCB, 한국신용정보원 재구성.

용평가모형의 정확성과 신뢰도를 높이기 위한 연구가 진행되었으며 우리나라에서도 개인 CB사 중 한 곳인 KCB에서 국세청 소득금액, 세금·건강보험·국민연금 납부내역, 통신요금, 공공요금 등의 정보를 활용하여 개인의 신용도에 긍정적인 영향을 주는 가점 방식의 K-Score를 2013년에 출시⁹⁾한 바가 있다.

2.2.2 평가활용 데이터 및 방법론

현재 국내 민간 CB사들은 <표 2>와 같이 개인 신용평가지 상환이력정보, 현재부채수준, 신용거래기간, 신용형태정보 등을 평가항목으로 활용하고 있다. 평가항목 내 세부 변수들의 활용은 CB사별 신용평가모형에 따라 상이하나, 대부분 공통적으로 사용하는 변수들이 많다.

상환이력정보의 경우, 3개월 이상 연체를 의미하는 장기연체 여부와 3개월 미만 연체를 의미하는 단기연체 여부를 구분하여 평가에 활용하고 연체건수, 연체금액, 연체상환 후 경과일수 등을 주로 반영한다. 현재부채수준은 대출건수, 대출금액 등과 신용카드 이용실적, 단기카드대출 잔액, 대출규모 대비 상환비율 등을 반영하며, 신용거래기간의 경

우, 연체없이 정상적인 신용거래를 지속한 기간을 확인하여 충분한 신용거래가 축적되어 있음을 확인한다. 신용형태정보는 대출의 종류(업권, 금융상품 등)나 대출 개설빈도, 신용카드 이용패턴(이용금액, 이용개월 수 증가 등)을 반영한다. 이 외에도 본인의 증빙소득이나 국민연금, 건강보험, 통신비 납부 등 비금융 분야의 성실상환정보를 통해 개인의 성실도를 판단하고 이를 가점요인으로 활용하고 있다.

신용평가모형 방법론의 경우, 로지스틱 회귀모형이 실무적으로 가장 많이 사용되고 있는데, 이 영섭(2003)에 따르면, 이는 로지스틱 회귀모형의 정확성이 우수하고, 과대 적합(overfitting) 가능성이 적으며, 개발과정이 다른 방법론에 비해 간편하기 때문이다. 또한 개발 이후, 모형에 대한 해석과 설명이 용이하다는 점도 장점이다. 최근에는 비금융정보들을 평가에 활용하고 머신러닝 등 새로운 방법론을 적용한 연구들도 시도되고 있으나, 개정 신용정보법에서 본인 신용평점에 대한 설명요구권이 포함¹⁰⁾되면서 설명가능한 신용평가모형에 대한 이슈가 있다.

9) 코리아크레딧뷰로 보도자료(2013.11.18.), “신용과 사람이 소통하는 새로운 개인신용평가, K-Score”.

10) 개정 신용정보법 제36조의2(자동화평가 결과에 대한 설명 및 이의제기 등).

2.2.3 개인신용평가회사의 상환이력정보 활용기준

개인신용평가지 활용되는 평가항목 중 현재 연체 여부와 과거 채무상환능력을 반영하는 상환이력정보는 개인의 신용도 및 거래능력을 판단할 수 있는 중요한 정보에 해당한다. 아래의 <표 3>에서도 알 수 있듯이 상환이력정보는 다른 정보들에 비해 평가항목의 가중치가 상대적으로 높으며 일반적으로 연체이력이 있는 경우, 그렇지 않은 일반 개인에 비해 상환이력정보와 현재부채수준 관련 평가항목의 가중치가 상대적으로 더 높다.

<표 3> 연체이력 유무에 따른 평가항목별 활용비중

평가항목	NICE		KCB	
	일반	연체 이력 有	일반	연체 이력 有
상환이력정보	37%	55%	24%	55%
현재부채수준	23%	25%	28%	15%
신용거래기간	12%	3%	15%	4%
신용형태정보	28%	17%	33%	26%
신용조회정보	-	-	-	-
가점 부여	증빙소득, 비금융정보 성실상환 정보 등			

자료: NICE평가정보, KCB.

개인이 대출 또는 신용카드를 3개월 이상 장기 연체하는 경우, 개인신용평가회사는 연체를 변제한 이후에도 3~5년까지 개인의 연체이력을 평가에 활용할 수 있으며, 상환이력정보는 최대 5년까지 활용 가능¹¹⁾하다. 때문에 연체이력이 있는 개인은 연체를 변제한 이후에도 연체하기 이전의 신용도로 회복되는 기간이 최장 5년까지 소요된다.

CB사가 신용평가모형을 통해 신용등급과 평점을 산정함에 있어 해당 정보를 일정 기간 활용하는 것은 금융회사의 여신 건전성을 관리와 우리 경제의 건전한 신용질서를 확립하기 위해서도 필

요하다고 판단되나, 여전히 연체를 상환했음에도 불구하고, 해당 연체이력을 3~5년간 평가에 지속적으로 활용하여 신용등급 또는 평점의 회복기간을 장기간으로 가져가고 있다는 사실은 문제점으로 지적할 수 있다.

해당 내용은 2019년 국정감사 이슈 분석 자료¹²⁾에서도 언급되었는데 “스스로 신용을 관리하는 사람과의 신용평가상의 역차별이 발생하지 않도록 장기연체이력은 신용평가에 활용이 필요할 수 있지만, 채무불이행자의 특성에 따라 신용평가를 고도화함으로써 위험도가 낮은 경우에는 좋은 신용평가를 받을 수 있도록 할 필요가 있다.”라고 지적한 바가 있으며, “채무불이행 이력이 ‘채무자 변제’사유로 해제된 경우 신용평가에 긍정적으로 반영되거나 이력 활용기간을 단축하는 등의 개선을 검토할 필요”가 있다는 의견을 제시하였다.

그간 상환이력정보의 활용기준은 민원과 국정감사 의견 등을 반영하여 여러 차례 변화해 왔다. 2011년에는 신용평가지 소액·단기 연체정보의 반영을 축소하는 정책을 시행하여 10만 원 미만의 연체정보는 평가에 미반영하고, 90일 미만의 연체정보는 채무상환시, 3년간만 신용평가에 반영하도록 하였다.¹³⁾

2016년에는 국세, 지방세 등 체납정보는 체납해제 이후 평가에 활용할 수 있는 상환이력정보기간을 5년에서 3년으로 단축하고, 한국장학재단의 학자금 장기연체 발생시, 연체 해제 후 평가에 활용할 수 있는 상환이력정보기간을 5년에서 1년으로 단축하는 등 정책적인 목적에 의해 특정 유형의 연체에 대해 신용등급 또는 평점의 회복기간을 줄여주는 규제¹⁴⁾를 진행해왔다.

11) 신용정보법 감독규정 제19조(오래된 신용정보의 삭제).

12) 국회입법조사처의 「2019 국정감사 이슈 분석 IV - 정무위원회, 기획재정위원회」자료 중 “채무불이행자의 변제 후 신용회복 개선 방안” 내용.

13) 금융위원회 보도자료(2011.4.15.), 「서민금융 기반 강화 종합대책」발표

14) 금융위원회 보도자료(2016.1.6.), “신용정보 활용방식을 개선하여 국민 불편을 적극 해소하겠습니다.”

한편 2018년에는 연체금액과 건수에 대한 등록 기준이 다시 강화¹⁵⁾되었다. 단기연체의 경우, 기존에 10만 원 이상 5영업일 이상 연체시 CB사에 단기연체로 등록되고 상환(또는 등록해제) 후 3년간 상환이력정보가 신용평가에 반영되던 것을 30만원 이상 10영업일 이상 연체시 등록되도록 변경(연체이력이 1건이고, 30만 원 미만 30일 미만 연체시 평가에 미활용)하였으며, 장기연체의 경우에는 기존에 50만 원 이상 3개월 이상 연체시 신용정보원에 장기연체로 등록되고 상환(또는 등록해제) 후에도 5년간 상환이력정보가 신용평가에 반영되던 것을 100만 원 이상으로 금액 기준을 상향 조정하였다.

본 논문에서는 이러한 정책적 흐름 속에서 연체이력이 변제 등의 이유로 연체정보가 해제된 이후 3~5년간 유지되어 신용도 회복에 어려움을 겪고 있는 개인 중 정상변제 가능성이 높아 신용도 개선 의지가 있는 개인을 확인하여 이들의 신용도 회복에 도움을 줄 수 있는 방법을 찾고자 한다.

2.3 개인신용정보를 활용한 연체·부도·변제 연구

2.3.1 개인신용정보를 활용한 연체·부도 관련 연구

현재 국내외에서 신용정보를 활용한 연구들을 살펴보면 개인 차주보다는 기업 차주의 신용정보를 활용한 연구결과들이 다수를 차지하고 있으며, 개인신용정보와 관련한 연구주제들은 개인 차주의 연체 결정요인과 부도위험 분석 등 연체 및 부도에 대한 연구가 대부분이다.

실제 개인의 연체와 부도에는 채무상환능력, 채무상환심리, 신용관리역량 등 다양한 요인이 작용한다. 그러나 현재 대부분의 선행연구들은 연체가 경제적 어려움에서 발생한다는 점에서 Campbell

and Dietrich(1983), Campbell and Cocco(2015), Krainer *et al.*(2009) 등의 연구에서는 경제적 요인, 즉, 채무상환능력에 초점을 두어 이루어져 왔다. 또한 Getter(2003), Gross and Souleles(2002), Bhutta *et al.*(2010) 등 경제적 요인에 초점을 둔 연구들에서는 소득, 부채상환금액 등의 절대적 수치나 소득, 자산과 부채와의 상대적 비중이 연체에 중요한 영향을 끼치게 됨을 밝힌 바 있다. 이러한 연구들은 단순히 소득이나 자산 등 경제적 능력의 부족만이 아니라 부채와의 상대적 비중이 중요한 요인임을 밝혀낸 것이다. 그러나 해외 연구들에서도 채무상환능력만을 초점을 맞춰 실증분석한 연구는 크게 부족한 실정이며, 몇몇 연구에서 부채부담을 채무상환능력의 일환으로 다룬 뿐이다.

개인 연체결정요인에 대한 국내의 실증분석 연구에는 KCB의 데이터를 활용하여 개인 단위가 아닌 부채를 보유한 가구 단위별 연체결정 요인을 임의추출 방식의 샘플링 자료로 분석한 이동걸 등(2014)의 연구사례가 있었으며 이장연, 임영주(2014)에서는 7등급 이하 저신용등급의 가계차주에 대한 연도별 동태적 분석사례도 있었다. 또한 패널데이터 모형을 적용하여 기업대출, 가계대출, 신용카드 대출 등 유형별로 국내 일반은행 연체를 결정요인을 추정한 김희철, 박형근(2010)의 연구와 가계금융복지조사 데이터를 활용하여 연체기간과 연체사유에 따른 연체현황을 조사한 성영애(2019)의 연구, 가구 특성 및 소득 수준, 대출상품에 따른 연체사유를 분석한 조재현(2020)의 연구, 연체채무자의 연체기간에 따른 완제 행태를 분석하여 채무조정제도에 대해 실증분석한 김형준 등(2019)의 연구 등도 있다. 또한 최근에는 한국신용정보원의 신용정보 표본DB와 기계학습 알고리즘을 사용한 개인의 채무불이행 예측모형을 생성하고 각 모델 간 성능을 비교한 박소희, 최대선(2019)의 연구도 있었다.

한편 대출상품별로는 주택담보대출과 관련하여 DTI, LTV 및 대출상환 조건이 주택담보대출의 연체율에 미치는 영향에 대해 연구한 허석균(2012)

15) 금융위원회 보도자료(2018.1.30.), 「개인신용평가체계 종합 개선방안」 발표.

의 연구와 김문년(2015)의 주택담보대출의 연체 위험 분석 연구, 민경록 등(2007)의 국내 신용카드 사용자의 연체요인에 대한 실증연구, 한준태, 정진아(2016)의 학자금 대출 연체의 신용위험 평점 모형 개발 연구 등이 있다.

개인 부도리스크에 대한 연구는 개인도산 확정 정보 등을 활용하여 개인도산 현황과 특성을 분석한 결과를 발표한 최종원(2014)의 연구와 개인회생 신청자의 특성을 반영한 NICE평가정보 CB연구소(2012)의 도산신청 예측 모델 연구 등이 있다.

2.3.2 개인신용정보를 활용한 변제 관련 연구

본 연구의 분석주체인 연체경험자의 변제와 관련된 연구의 경우, 연구에 활용할 수 있는 데이터의 부족과 변제에 대한 금융회사 내부 데이터 축적이 이루어지지 않아 관련 선행연구나 실증분석이 매우 부족한 실정이다. 이는 금융소비자의 개별 특성을 파악할 수 있는 활용가능한 공개 데이터가 부족하기 때문이기도 하지만, 각 금융회사가 보유한 여신 관련 정보들이 분산되어 있어 이를 아우를 수 있는 데이터의 수집이 어렵기 때문이기도 하다.

때문에 개인 차주의 변제 관련 연구는 데이터에 의한 실증연구보다는 가정준(2016)의 제3자의 변제와 관련한 비교법적 고찰, 손병환(2018)의 대위 변제와 근저당권의 이전 연구, 홍봉주(2017)의 제3취득자의 변제자 대위 등과 같이 주로 대위변제, 근저당권 등 법률상 이슈를 다룬 연구들이 주로 진행되어 왔다.

III. 연구설계 및 방법론

3.1 데이터 특성

현재 한국신용정보원으로 집중되는 일반신용정보, 기술신용정보, 보험신용정보 등 3가지 유형의 신용정보 중 본 연구의 분석대상인 일반신용정보

는 신용거래정보, 신용도판단정보, 신용거래능력판단정보, 공공정보 등의 대항목으로 구분할 수 있으며 이 중 대출 관련 정보가 신용거래정보에, 연체 관련 정보가 신용도판단정보에 포함되어 있다.

본 연구에 활용된 데이터는 2019년 12월 말 기준 한국신용정보원에 등록된 대출·연체·변제 정보로 대출 및 연체 개별 건에 포함된 항목에는 대출, 연체 건수, 금액 등의 정보와 개인 차주의 식별정보에서 추출한 성별과 연령 정보가 함께 포함된다.

3.1.1 대출경험자와 연체경험자

본 연구에서 언급하는 대출경험자는 2019년 12월 말 기준으로 금융회사를 통해 대출을 실행하여 한국신용정보원에 대출 관련 정보가 1건 이상 등록된 개인을 의미한다. 연체경험자는 2019년 12월 말 기준으로 금융회사를 통해 받은 대출 건에 대해 연체가 발생하여 한국신용정보원에 연체 관련 정보가 1건 이상 등록된 개인을 말한다.

3.1.2 연체 후 변제한 유형의 구분

한국신용정보원의 일반신용정보관리규약에 의해 3개월 이상 연체하여 해당 정보가 등록된 개인 차주에게 등록해제 사유가 존재하면 등록된 연체 정보는 삭제된다. 이러한 등록해제 사유는 결국 ‘연체 후 변제한 사유’로 볼 수 있으며, 연체정보를 삭제하는데 필요한 등록해제 사유에 대한 코드별 상세내역은 다음의 <표 4>와 같다.

본 연구에서는 이 중 해제코드 ‘01’에 해당하는 ‘채무자 변제’ 사유만을 ‘정상변제’한 신호로 간주하였으며 해제코드 ‘02’부터 ‘99’까지는 ‘비정상 변제’의 신호로 판단하였다.

3.1.3 신용거래정보 중 대출 관련 정보

일반신용정보 항목 중 하나인 신용거래정보에는 개인이 가계당좌예금이나 신용카드 등을 개설, 발급한 정보와 은행 등의 금융회사를 통해 실행한 대출정보와 단기카드대출(현금서비스), 장기카드

〈표 4〉 연체 등록해제 사유 구분

해제코드	등록해제 사유	등록해제 사유 세부 내용
01	채무자 변제	채무자가 채무를 직접 변제한 경우
02	보증인 변제	보증인이 채무를 변제한 경우
03	강제회수	금융기관이 채권을 법적 절차 등에 의해 회수한 경우
04	손실처리	금융기관이 채권을 감면, 면제, 포기한 경우
06	신용회복지원	개별 금융기관의 신용회복지원이 있는 경우
10		신용회복위원회의 신용회복지원이 있는 경우
11		한마음금융의 신용회복지원이 있는 경우
07	채무자 변제 및 손실처리	
08	보증인 변제 및 손실처리	
09	강제회수 및 손실처리	
12	회생/변제계획 인가	회생/변제계획인가 결정이 있는 경우
13	소멸시효 만료	소멸시효가 완성된 경우
14	파산	파산면책 결정이 있는 경우
15	사망/파산	개인의 사망 또는 법인의 파산이 있는 경우
16	양도	채권을 금융기관이 아닌 자에게 양도한 경우
17		채권을 금융기관에 양도한 경우
99	기타	등록사유 발생일로부터 관리기간이 종료되거나 기타의 이유

자료: 한국신용정보원, 「일반신용정보관리규약」.

〈표 5〉 대출유형 구분

구분(등록코드)	대출유형		유형코드	등록시기
대출현황(0031)	신용대출	신용대출	100	등록사유 발생일로부터 5영업일 이내
		학자금대출	150	
	담보대출	예·적금 담보대출	200	
		유가증권 담보대출	210	
		주택 담보대출	220	
		주택외부동산 담보대출	230	
		지급보증(보증서) 담보대출	240	
		보급자리론	245	
		학자금(지급보증담보) 대출	250	
		주택연금대출	260	
		전세자금담보대출	270	
		전세보증금담보대출	271	
		기타 담보대출	290	
	보험계약대출	400		
	할부금융	신차 할부금융	500	
		중고차 할부금융	510	
		기타 할부금융	590	
리스	금융리스	700		
	운용리스	710		
카드론(0037)	카드론 거래사실			
현금서비스(0041)	신용카드 현금서비스 거래사실			

자료: 한국신용정보원, 「일반신용정보관리규약」.

〈표 6〉 연체유형 구분

등록사유	등록코드	등록시기	보존기간
대출원금 및 이자 3개월 이상 연체	0101	등록사유 발생일로부터 7영업일 이내	1. 연체 등록금액 1천만 원 초과 - 등록사유 발생일로부터 90일 이내 해제시 미보존 - 채무를 변제하지 않은 기간(단, 해당 기간이 1년 초과시 1년간 보존) 2. 연체 등록금액 1천만 원 이하 - 보존하지 않음
5만원 이상의 카드론 대금 3개월 이상 연체	0103		
5만원 이상의 신용카드 대금 3개월 이상 연체	0104		
5만원 이상의 할부금융 대금 3개월 이상 연체	0105		
기타 사유(신용보증기금 보증 청년창업대출 6개월 이상 연체, 농림부 지원 학자금 용자사업 대출원금 등 6개월 이상 연체, 외환 연체 등)	-		

자료: 한국신용정보원, 「일반신용정보관리규약」.

대출(카드론) 등 카드사를 통한 대출정보, 채무보증정보, 복수카드정보 등이 포함되어 있다. 본 연구에서는 은행, 카드사, 보험 등 여러 금융업권 중 대다수의 대출정보가 존재하는 은행, 카드사만을 분석에 이용하고 한국신용정보원의 대출유형을 활용하였다.

3.1.4 신용도판단정보 중 연체 관련 정보

신용도판단정보에는 연체정보, 대위변제 및 대지급 정보, 부도정보 등이 포함되는데 이 중 연체정보를 활용하여 분석에 이용하였으며 연체유형은 특수한 형태의 연체유형을 제외하고 대출원금 및 이자 3개월 이상 연체(0101), 신용카드 대금(0104) 및 카드론(0103), 할부금융(0105) 연체만을 대상으로 삼았다.

3.2 데이터 구성

한국신용정보원에서 보유중인 일반신용정보의 대출 정보와 연체 정보는 데이터 수집 목적으로 기존에 마련된 DB 구조상 신용거래정보 DB 테이블과 신용도판단정보 DB 테이블로 구분하여 운영되고 있으며 데이터의 단위가 개인이 보유한 계좌별로 구성되어 있어 본 연구에 필요한 개인 차주 단위별 데이터 구성이 이루어지지 않은 상태였다.

때문에 분석에 용이하도록 2개의 DB 테이블을

개인식별번호(주민등록번호를 해시코드로 비식별화한 key)를 기준으로 연결하여 하나의 개인식별번호를 중심으로 개인 차주 단위별로 대출 정보와 연체 정보를 연결한 2019년 12월 말 기준 연체경험자의 데이터셋을 구성하였으며 다음의 <표 7>에서와 같이 독립변수와 종속변수를 구성하였다.

종속변수의 경우, 연체 후 변제한 여러 유형 중 채무자가 채무를 직접 변제한 경우만을 정상변제로 분류하였으며, 그 외에 강제회수, 손실처리, 신용회복지원 등의 이유로 변제한 경우는 비정상변제로 분류하였다. 또한 연체 건이 다수 존재하는 경우, 연체 건 모두를 정상변제한 경우만 1로 표시하고 그 외에 해당 연체 건 중 일부만 정상변제한 경우는 0으로 표시하였다.

독립변수는 한국신용정보원에서 수집하는 대출, 연체 관련 정보 항목들을 중심으로 연체경험자의 식별정보인 성별, 연령 정보와 대출·연체 유형, 대출·연체건수, 대출·연체 금액(총 금액, 평균 금액), 금리 등의 정보 항목을 변수로 선별하였으며 이러한 변수들 중 다중공선성 여부를 점검하여 다중공선성이 있고 설명력이 약한 변수는 제거하였으며 결측값이 없고 분석에 이용하기 적합한 변수를 최종적으로 선택하여 분석에 활용하였다.

한편 실제 기초통계량을 비교해본 결과, 연체 경험자의 연체금액과 건수를 연체경험자 중 정상

〈표 7〉 변수 설명

구분	변수 영문명	변수 의미	비고
종속변수	RLS_YN	정상변제 여부	정상변제(채무자 직접 변제)=1, 비정상변제(강제회수, 손실처리 등)=0
독립변수	GENDER	성별	남=1, 여=2
	AGE	연령	개인 차주의 연령
	LOAN_XXX(3자리) (대출유형 코드)	대출유형	개인 1명이 보유중인 대출유형의 dummy 변수 (해당 유형 보유=1, 미보유=0)
	RGST_XXXX(4자리) (연체유형 코드)	연체유형	개인 1명이 보유중인 연체유형의 dummy 변수 (해당 유형 보유=1, 미보유=0)
	LN_CNT	대출건수	개인 1명이 해당 년월에 보유중인 대출 총 건수
	ODU_CNT	연체건수	개인 1명이 해당 년월에 보유중인 연체 총 건수
	LN_AMT	총 대출금액	개인 1명이 보유한 총 대출금액
	ODU_AMT	총 연체금액	개인 1명이 보유한 총 연체금액
	LN_AVG_AMT	평균 대출금액	개인 1명의 계좌별 대출 평균금액
	ODU_AVG_AMT	평균 연체금액	개인 1명의 계좌별 연체 평균금액
AVG_DSR_RE_RATE	대출 평균금리	개인 1명이 보유한 대출 계좌별 평균금리	
통제변수	GG_YM	등록년월	한국신용정보원에 등록된 년월
	ID_NO_IPIN_NUM	개인식별번호	주민등록번호를 비식별화한 key
	IND_ITT_CD	금융업권 구분	금융업권별 구분코드

변제자와 비교해보면, 정상변제자의 경우가 연체 금액과 건수 측면에서 현저히 낮게 나타나는 것을 확인할 수 있는데, 이를 통해 정상변제자의 신용도가 전체 연체경험자에 비해 높을 수 있음을 확인할 수 있다.

3.3 분석 방법론

본 연구에서 적용한 방법론은 개인신용평가모형에서 가장 많이 사용하는 로지스틱 회귀모형(logistic regression model)과 의사결정나무 모형

〈표 8〉 분석대상 변수의 기초통계량

구분	대출경험 (A)	연체경험 (B)	연체경험 비율 (B/A)	정상변제 경험* (C)	정상변제** (D)	정상변제 비율 (D/B)
건수 기준	17,750,551	723,496	4.08%	23,118	4,228	0.58%
사람 기준	15,244,750	575,003	3.77%	15,575	3,571	0.62%

* 연체경험자 중 현재 연체된 건에서 최소 1건이라도 정상변제한 경우

** 연체경험자 중 현재 연체된 건 모두 정상변제한 경우

구분	대출경험자		연체경험자				연체미경험자	
	대출금액	대출건수	대출금액	대출건수	연체금액	연체건수	대출금액	대출건수
사람 기준	5,690	2.25	1,500	2.23	9,116	5.09	5,868	2.26

구분	연체경험자				연체경험자 중 정상변제자			
	대출금액	대출건수	연체금액	연체건수	대출금액	대출건수	연체금액	연체건수
사람 기준	1,500	2.23	9,116	5.09	3,393	1.86	6,780	1.12

〈표 9〉 분석에 활용한 표본의 개수

구분	학습용 데이터(70%)			검증용 데이터(30%)			총계
	Y=0	Y=1	소계	Y=0	Y=1	소계	
표본 개수	503,045	510,448	1,013,493	216,223	218,469	434,692	1,448,185

(decision tree model), 신경망 모형(neural network model) 등 3가지이다. 본 연구의 경우, 종속변수에 해당하는 ‘정상변제 여부(정상변제=1, 비정상변제=0)’가 이진형(binary) 변수이므로 선행연구를 참고하였을 때, 로지스틱 회귀모형으로 우선 분석하는 것이 적합하다고 보았다. 또한 머신러닝 기법에서 분류 방법론으로 주로 사용되는 의사결정나무 모형과 신경망 모형을 활용하여 로지스틱 회귀모형 분석결과와 비교하였다.

분석대상인 데이터셋의 경우, 연체경험자 중 정상변제를 한 개인 차주의 비중이 매우 낮아, 데이터 불균형이 존재한다. 이 경우, 학습과정의 왜곡이 발생하게 되어 분석결과의 정확성과 안정성이 저해되는 위험에 노출될 우려가 있다. 때문에 이러한 문제의 해결을 위해 정현승 등(2008)의 연구를 참고하여 ‘RLS_YN’ 변수의 0과 1의 비율이 1:1을 유지하도록 정상변제의 비율을 172.4배 증폭시켜 분석에 활용하는 오버샘플링(over-sampling) 기법¹⁶⁾을 적용하였다. 또한 모형의 성능 확인을 위해 <표 9>와 같이, 전체 표본을 학습용 데이터 70%, 검증용 데이터 30%로 나누어 분석에 활용하였다.

방법론 중 우선 적용한 로지스틱 회귀모형은 종속변수의 확률사건 발생이 이진형(발생=1, 미발생=0)인 경우, 개별 관측치들이 어느 집단으로 분류될 수 있는가를 분석하고 이를 예측하는 모형을 개발하는데 활용되는 모형이다.

한편 의사결정나무 모형은 의사결정 규칙을 나무 가지 형태로 도식화하여 분류와 예측을 수행하는 방법론이다. 의사결정나무 모형을 사용하면 본 연구의 종속변수인 ‘정상변제 여부’에 대한 의사결정 과정을 독립변수를 이용하여 도식화할 수 있어 직관

적인 이해가 가능하다는 장점이 있다. 의사결정나무 모형에는 CHAID¹⁷⁾, CART¹⁸⁾, C5.0 등 여러 가지 알고리즘이 존재하는데 이 중 분류의 정확도가 높은 모형으로 분류되는 CART 방식을 적용하였다.

신경망 모형은 머신러닝 기법 중 주로 분류와 예측을 위해 사용되는데, 신경망 알고리즘 중 가장 널리 사용되는 다층 퍼셉트론(Multi-Layer Perceptron; MLP)를 활용하였다. 신경망 모형의 장점은 자료간의 비선형적인 관계를 찾아낼 수 있다는 점이나, 과대적합(over-fitting)의 단점이 있어 학습된 모형에 새로운 데이터를 적용하면 예측력이 떨어질 수 있다. 독립변수를 dummy 변수로 구성한 본 연구의 경우, 김의중(2016)의 연구에서 신경망 모형에 적합하다고 언급된 ‘범주형 독립변수의 범주 빈도가 유사한 경우’에 해당되어 신경망 모형을 적용하는 것이 유리할 것으로 판단하였다.

3가지 방법론의 성능 비교는 오분류표¹⁹⁾와 이익 도표²⁰⁾, AUC, GINI 등의 성능지표를 기준으로 하였으며, 분석도구는 IBM의 SPSS Modeler 17.1을 이용하여 데이터 가공 및 전처리, 모델링, 분석 등 전 과정을 수행하였다.

IV. 실증분석 결과

4.1 정상변제에 영향을 미치는 요인

분석결과, 정상변제에 영향을 미치는 주요 변

16) 소수 군집의 데이터를 다수 군집 데이터의 크기까지 증폭 발생시키는 방법.

17) CHI-squared Automatic Interaction Detection.

18) Classification And Regression Trees.

19) Confusion Matrix, 종속변수의 실제값과 모형에 의해 산출된 예측값 사이의 관계를 나타내는 표.

20) Lift Chart, 균일하게 n등분한 각 집단에 분류된 관측치의 분포된 상태를 누적으로 표시한 그림.

〈표 10〉 독립변수의 상대적 중요도

선택된 독립변수 (영문명/변수 의미)	적용한 방법론별 중요도		
	로지스틱 회귀모형	의사결정나무 모형 (CART 알고리즘)	신경망 모형 (MLP 알고리즘)
GENDER (성별)	< 0.01	< 0.01	< 0.02
AGE (연령)	< 0.01	< 0.01	0.07
LOAN_100 (신용대출)	0.01	< 0.01	0.05
LOAN_0037 (카드론)	< 0.01	< 0.01	0.02
LOAN_0041 (현금서비스)	0.03	< 0.01	0.02
RGST_0101 (대출원금 및 이자 3개월 이상 연체)	0.01	0.02	0.02
RGST_0103 (카드론 연체)	0.04	< 0.01	0.02
RGST_0104 (신용카드대금 연체)	0.11	0.02	0.03
LN_CNT (대출건수)	< 0.01	< 0.01	0.12
ODU_CNT (연체건수)	0.78	0.92	0.48
LN_AVG_AMT (평균 대출금액)	0.03	0.02	0.17

수는 성별, 연령, 대출·연체유형 일부, 대출·연체건수, 평균 대출금액 등으로 확인되었으며 <표 10>은 종속변수에 영향을 미치는 독립변수 간 상대적 중요도를 수치로 표시²¹⁾한 것이다.

3가지 방법론 모두 ‘연체건수’가 가장 높은 중요도를 보이고 있으며 특히 로지스틱 회귀모형과

의사결정나무 모형의 경우, 각각 0.78, 0.92로 압도적인 중요도를 보이고 있다.

변수에 대한 해석이 용이한 로지스틱 회귀모형을 기준으로 단계적 선택법(Stepwise)으로 선택된 독립변수의 의미에 대해 살펴보면, 성별의 경우, 남자(1)에 비해 여자(2)인 경우가 정상변제 가능성이 높음을 확인할 수 있으며, 독립변수 중 상대적 중요도가 가장 높은 것으로 확인된 연체건수의 경우, 추정치가 음수를 보이고, 오즈비가 1보다 작은 수치를 보이고 있어, 연체건수가 적을수록 정상변제 가능성이 높아지는 것으로 해석될 수 있다. 대출유형 중에는 ‘카드론’이 있는 경우, 연체유형 중에는 ‘대출원금 및 이자 3개월 이상 연체’가 있는 경우, 연체경험자의 정상변제 가능성이 낮아지는 것으로 해석 가능하다. 한편 신용대출이 있는 연

21) 분석도구인 SPSS Modeler에서 Saltelli et al.(2004)의 연구에 근거하여 산출한 예측변수 중요도(Predictor Importance) 수치로서 특정 독립변수가 종속변수에 미치는 상대적인 영향력에 대한 정보를 제공.

※ 민감도 분석(Sensitivity Analysis)과 유사한 개념으로 종속변수의 분산(분모) 대비 특정 독립변수가 종속변수에 미치는 기대값에 대한 분산(분자)을 계산한 수치를 산출하며 이를 다시 모든 독립변수에 대한 수치의 합(분모) 대비 특정 독립변수에 대한 수치(분자)로 정규화하는 방식임.

〈표 11〉 로지스틱 회귀모형에서의 분석결과

선택된 독립변수 (영문명/변수 의미)	추정치	표준오차	유의확률	오즈비
GENDER (성별)	0.062	0.006	< 0.0001	1.064
AGE (연령)	0.017	0.000	< 0.0001	1.017
LOAN_100 (신용대출)	0.504	0.008	< 0.0001	1.655
LOAN_0037 (카드론)	-0.091	0.008	< 0.0001	0.913
LOAN_0041 (현금서비스)	0.582	0.008	< 0.0001	1.790
RGST_0101 (대출원금 및 이자 3개월 이상 연체)	-0.485	0.008	< 0.0001	0.616
RGST_0103 (카드론 연체)	0.554	0.010	< 0.0001	1.740
RGST_0104 (신용카드대금 연체)	0.972	0.009	< 0.0001	2.644
LN_CNT (대출건수)	0.030	0.002	< 0.0001	1.030
ODU_CNT (연체건수)	-1.589	0.005	< 0.0001	0.204
LN_AVG_AMT (평균 대출금액)	0.000	0.000	< 0.0001	1.000

체경험자는 그렇지 않은 경우에 비해 약 1.66배 정상변제 가능성이 높고, 신용카드대금 연체가 있는 연체경험자는 그렇지 않은 경우에 비해 약 2.64배 정상변제 가능성이 높음을 확인할 수 있다.

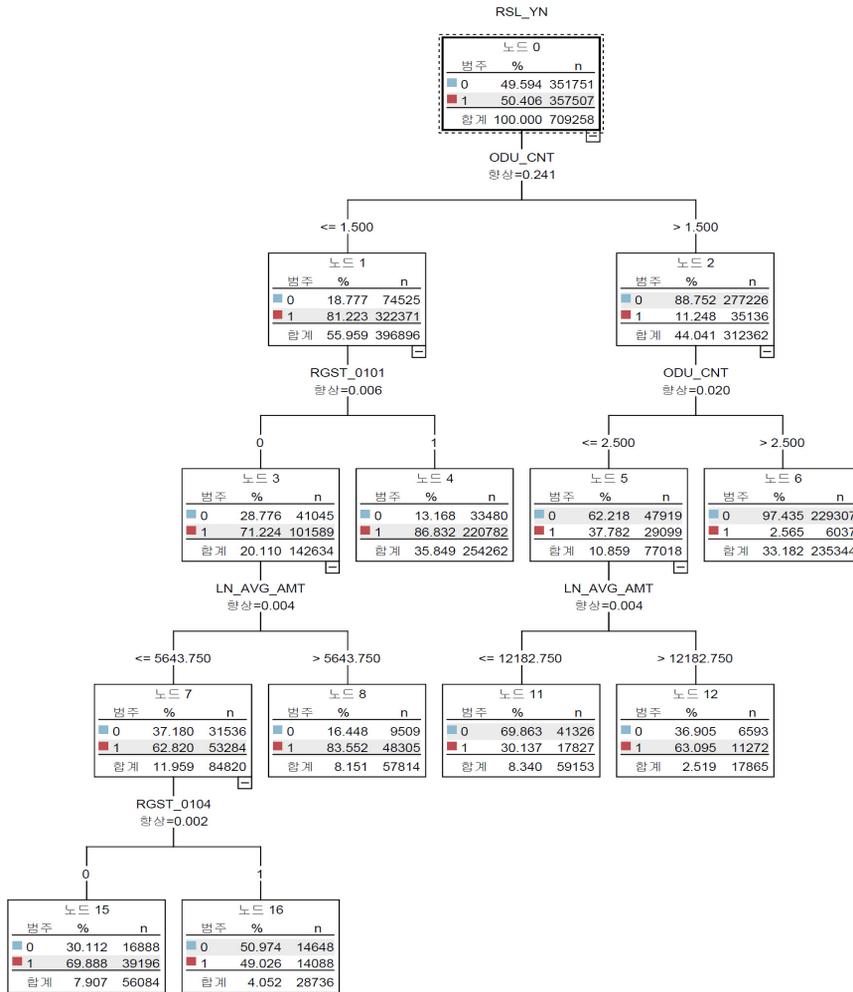
의사결정나무 모형은 최대 깊이를 4로 한정하여 분석하였으며, 분석결과, 로지스틱 회귀분석과 동일하게 가장 높은 영향도를 보인 ‘연체건수’에 대해 1.5건을 기준으로 1차 분기가 발생하여 ‘연체건수’ 1.5건 이하인 경우가 정상변제 가능성이 높으며 ‘연체건수’ 1.5건 이하인 경우 중에서는 ‘대출원금 및 이자 3개월 이상 연체’가 있는 연체경험자가 그렇지 않은 경우에 비해 정상변제 가능성이 높아짐을 알 수 있었다. 또한 ‘연체건수’ 1.5건을 초과하는 경우에는 다시 분기가 이루어져 ‘연체건수’ 2.5건 이하일 경우 정상변제 가능성이 높아

지고 2.5건 이하인 경우 중에서는 평균 대출금액이 약 1,200만 원을 초과하는 경우, 정상변제 가능성이 높아짐을 알 수 있었다.

한편, 신경망 모형은 은닉층 7개의 다중 퍼셉트론으로 모형이 구성되었으며, 신경망 모형의 특성상 변수의 의미는 확인할 수 없었다.

4.2 모형별 성능 평가

오분류표를 기준으로, 학습용 데이터로 구축된 각각의 모형을 검증용 데이터 기준으로 성능을 비교한 결과, 정확도 측면에서는 신경망 모형(86.63%), 민감도 측면에서는 로지스틱 회귀모형(94.33%), 특이도 측면에서는 의사결정나무 모형(80.93%)이 가장 높은 수치를 보인다.



〈그림 3〉 의사결정나무 모형에서의 분석결과

〈표 12〉 오분류표(Confusion Matrix)

성능비교 지표	로지스틱 회귀모형	의사결정나무 모형 (CART 알고리즘)	신경망 모형 (MLP 알고리즘)
<학습용 데이터>			
정확도(Accuracy)	86.12%	85.29%	86.80%
민감도(Sensitivity)	94.39%	89.39%	93.90%
특이도(Specificity)	77.72%	81.14%	79.59%
<검증용 데이터>			
정확도(Accuracy)	86.07%	85.21%	86.63%
민감도(Sensitivity)	94.33%	89.45%	93.88%
특이도(Specificity)	77.74%	80.93%	79.29%

〈표 13〉 이익 도표(Lift Chart) - 검증용 데이터 기준

성능비교 지표	로지스틱 회귀모형	의사결정나무 모형 (CART 알고리즘)	신경망 모형 (MLP 알고리즘)
<학습용 데이터>			
상위 10% 반응률	93.48%	86.78%	92.44%
상위 10% Lift	1.85배	1.72배	1.83배
<검증용 데이터>			
상위 10% 반응률	92.96%	86.74%	92.82%
상위 10% Lift	1.85배	1.73배	1.85배

〈표 14〉 모형 성능지표(AUC, GINI)

성능비교 지표	로지스틱 회귀모형	의사결정나무 모형 (CART 알고리즘)	신경망 모형 (MLP 알고리즘)
<학습용 데이터>			
AUC	0.914	0.900	0.924
GINI	0.828	0.800	0.849
<검증용 데이터>			
AUC	0.914	0.900	0.923
GINI	0.828	0.800	0.846

또한 이익 도표(Lift Chart)를 활용하여 3가지 방법론에 따른 성능을 비교한 결과, 모형 적용 이전의 정상변제 반응률을 50%로 설정²²⁾하면, 검증용 데이터 기준으로 모형 적용 후, 1등급(상위 10%)의 연체경험자에 대한 정상변제 반응률은 로지스틱 회귀모형은 92.96%, 의사결정나무 모형은 86.74%, 신경망 모형은 92.82%로 증가한다.

한편 Lift 값의 경우, 1등급(상위 10%)의 연체경험자에 대한 정상변제자 출현 빈도²³⁾는 검증용 데이터 기준으로 로지스틱 회귀모형이 1.85배, 의사결정나무 모형이 1.73배, 신경망 모형은 1.85배를 보인다. 3가지 방법론 모두 하위 등급으로 이동할 수록 Lift 값이 감소하는 것으로 나타나고 있어 3가지 방법론에 의한 모형 모두 안정된 형태임을 확인하였다.

모형 성능지표로 사용되는 AUC, GINI²⁴⁾를 이용하여 평가한 결과, 다음과 같이 AUC는 0.9 이상, GINI 0.8 이상의 높은 수치가 나와 3가지 방법론 모두 성능이 우수한 것으로 확인되었으며, 특히 검증용 데이터 기준으로 AUC, GINI 수치가 각각 0.923, 0.846으로 가장 높게 나온 신경망 모형이 타 모형에 비해 성능이 우수한 것으로 확인되었다.

따라서 오분류표와 이익 도표, 모형 성능지표 등을 종합해볼 때, 모형의 성능은 신경망 모형, 로지스틱 회귀모형, 의사결정나무 모형의 순으로 우수하다고 판단된다.

V. 결론 및 시사점

5.1 요약 및 결론

한국신용정보원이 보유한 개인신용정보로 국

22) Baseline %Response = 50%.

23) 상위 10% Lift 값은 모형을 적용하기 이전과 비교했을 때, 1등급(상위 10%)에서 정상변제할 연체경험자임을 해당 배수만큼 사전에 예측할 수 있다는 의미.

24) GINI = 2×AUC-1.

내 연체경험자의 정상변제 요인을 분석한 결과, 유의한 변수로 성별, 연령, 대출·연체유형 일부, 대출·연체 건수, 평균 대출금액 등이 확인되었으며, 이 중 ‘연체건수’가 가장 영향력이 높은 변수로 밝혀졌다.

방법론에 따라 산출된 모형의 해석이 다소 상이한데, 로지스틱 회귀모형의 경우에는 연체경험자의 연체건수가 적을수록 정상변제 가능성이 높고, 대출유형 중 ‘카드론’이 있는 경우, 연체유형 중 ‘대출원금 및 이자 3개월 이상 연체’가 있는 경우, 연체경험자의 정상변제 가능성이 낮아지는 것으로 해석할 수 있다. 또한 신용대출이 있는 연체경험자는 그렇지 않은 경우에 비해 약 1.66배 정상변제 가능성이 높고, 신용카드대금 연체가 있는 연체경험자는 그렇지 않은 경우에 비해 약 2.64배 정상변제 가능성이 높음을 확인하였다.

의사결정나무 모형에서는 ‘연체건수’가 1.5건을 기준으로 1차 분기가 발생하여 1.5건 이하이면서 ‘대출원금 및 이자 3개월 이상 연체’가 있는 연체경험자가 우선적으로 정상변제 가능성이 높은 것으로 확인되었다. 또한 연체건수 1.5건 이하이면서 ‘대출원금 및 이자 3개월 이상 연체’가 없는 개인은 대출금액을 기준으로 정상변제 가능성이 구분됨을 확인하였다.

본 연구결과는 다음과 같이 활용할 수 있을 것으로 보인다. 첫째, 간접적인 혜택을 제공하는 방법으로 연체경험자 중 정상변제 가능성이 높은 특징을 가진 개인들을 선별하여 기존 3~5년의 ‘상환이력정보’ 활용기간을 1~2년으로 단축시켜줌으로써 연체이력 활용에 따른 불이익을 경감시키고 신용도 회복을 빠르게 할 수 있도록 지원하는 것이다. 둘째, 신용도에 긍정적인 영향을 미칠 수 있는 비금융정보 등을 개인신용평가회사에 추가로 제공하면 신용평점에 가산점을 부여하는 현재의 방식을 차용하여 연체경험자 중 정상변제 가능성이 높은 개인에게는 기존 신용평점에 가산점을 부여함으로써 보다 직접적인 혜택을 제공하는 방법이 있을 수 있다.

셋째, 은행 등의 금융회사는 개인 여신심사시 정상변제 요인을 확인하여 리스크 관리에 활용한다면, 개인 대출채권의 부실률을 낮추고 수익성을 높이는데 도움이 될 것이다. 넷째, 정책적 측면에서도 연체된 채무를 성실하게 변제하여 정상적인 금융활동을 하고자 하는 개인을 선별하는 자료로 활용하여 연체건수 증가에 따른 신용도 하락을 방지하고 정상변제를 할 수 있는 차주의 신용도 회복을 지원하는 정책을 개발하는데 도움이 될 수 있을 것이다. 뿐만 아니라, 경제적 취약계층에 대한 서민금융 지원제도를 시행하는 경우에도 서민 금융에 사용되는 자원을 정상변제 가능성이 높은 성실상환자에게 집중할 수 있게 되어 효과적인 자원 활용이 가능할 것이다.

5.2 한계점 및 향후 연구방향

본 연구는 한국신용정보원이 보유한 정보만으로 정상변제 요인을 확인하기 위한 변수들을 구성하였기 때문에 대출과 연체, 정상변제 등에 영향을 미치는 추가적인 요인이 존재할 수 있음을 간과한 측면이 있다. 따라서 향후 대출, 연체와 관련된 신규 변수들을 추가로 발굴하여 분석하고 개인 신용평가회사가 산출하는 신용평점 등을 반영하여 금융업권별 정상변제 요인의 차이 등을 세부적으로 확인한다면, 통계모형의 보완과 정교화가 가능할 것이다.

개인이 금융회사의 채무를 1개월 이상 연체하면 대부분의 금융회사는 개인과의 여신거래를 기한이익 상실²⁵⁾로 처리하여 사실상 거래계약을 종료하고 금융회사는 원리금과 연체 이자를 일시에 청구할 수 있는 권리를 갖게 된다. 본 연구에 활용된 자료는 3개월 이상 연체한 장기연체 정보이므로, 이미 채무불이행중인 개인으로 낙인된 연체경험자가 분석대상이 된다. 따라서 3개월 미만의 단

25) 채무자의 신용위험 증가시, 금융기관이 대출만기 이전에도 채무를 일시에 회수할 수 있는 권리.

기연체 정보와 정상변제 정보를 확보할 수 있다면 정상변제 가능성이 높은, 즉 상환의지가 높은 개인을 식별하여 이들에게 혜택을 부여할 수 있는 실질적인 연구가 가능할 것이다. 또한 향후에 월별 패널자료를 활용하여 최초 연체경험자가 2회 이상 재연체를 하고 채무불이행 단계까지 진행되는 일련의 과정에 대한 분석이 이루어진다면 본 연구에서 영향력이 높은 변수로 확인된 연체건수에 대해 추가적인 시사점을 발견할 수 있을 것으로 보인다.

특히 올해 8월 시행된 개정 신용정보법에는 검증위원회를 통해 신용평가모형을 검증하고 CB사에 개선점을 권고할 수 있는 법적 근거²⁶⁾가 마련되었다. 따라서 향후 본 연구와 유사한 연구들이 축적된다면, 금융소비자에게 혜택을 줄 수 있는 다양한 방법들이 금융소비자 정책에 반영될 수 있을 것이다.

참 고 문 헌

- [1] 가정준, “제3자의 변제와 관련한 비교법적 고찰”, *비교사법*, 제23권, 제1호, 2016, pp. 189-218.
- [2] 국가법령정보센터, “신용정보의 이용 및 보호에 관한 법률”, 2020, Available at <http://www.law.go.kr/LSW/lsInfoP.do?lsiSeq=213953#0000>.
- [3] 국회입법조사처, “2019 국정감사 이슈 분석 4권 - 정무위원회, 기획재정위원회”, 2019, pp. 134-135.
- [4] 권영준, 남재현, 조민정, “개인신용평가에서의 비금융정보의 경제적 효과”, *한국경제연구*, 제29권, 제2호, 2011, pp. 81-107.
- [5] 금융감독원, “개인신용평가 민원동향 및 주요 민원사항 Q&A”, 2018, Available at http://www.fss.or.kr/fss/kr/promo/bodobbbs_view.jsp?seqno=21848&no=1&s_title=%B0%B3%CE%B
- [6] 금융위원회, “「개인신용평가체계 종합 개선 방안」 발표”, 2018, Available at http://www.fsc.go.kr/info/ntc_news_view.jsp?bbsid=BBS0030&page=1&sch1=subject&sword=%EA%B0%9C%EC%9D%B8%EC%8B%A0%EC%9A%A9%ED%8F%89%EA%B0%80%EC%B2%B4%EA%B3%84%20%EC%A2%85%ED%95%A9&r_url=&menu=7210100&no=32285.
- [7] 금융위원회, “「서민금융 기반강화 종합대책」 발표”, 2016, Available at <https://www.fsc.go.kr/no010101/69922?srchCtgr=&curPage=&srchKey=sj&srchText=%EC%84%9C%EB%AF%BC%EA%B8%88%EC%9C%B5+%EA%B8%B0%EB%B0%98%EA%B0%95%ED%99%94+%EC%A2%85%ED%95%A9%EB%8C%80%EC%B1%85&srchBeginDt=&srchEndDt=#>.
- [8] 금융위원회, “신용정보 활용방식을 개선하여 국민 불편을 적극 해소하겠습니다”, 2016, Available at <https://www.fsc.go.kr/no010101/71909?srchCtgr=&curPage=&srchKey=sj&srchText=%EC%8B%A0%EC%9A%A9%EC%A0%95%EB%B3%B4+%ED%99%9C%EC%9A%A9%EB%B0%A9%EC%8B%9D%EC%9D%84&srchBeginDt=&srchEndDt=#>.
- [9] 금융위원회, “한국신용정보원 종합신용정보 집중기관으로서 첫 발을 떼다”, 2016, Available at http://www.fsc.go.kr/info/ntc_news_view.jsp?bbsid=BBS0030&page=1&sch1=subject&sword=%ED%95%9C%EA%B5%AD%EC%8B%A0%EC%9A%A9%EC%A0%95%EB%B3%B4%EC%9B%90&r_url=&menu=7210100&no=30848.
- [10] 김문년, “주택담보대출의 연체위험 분석”, 주

26) 개정 신용정보법 제26조의3(개인신용평가체계 검증위원회).

- 택연구, 제23권, 제2호, 2015, pp. 103-142.
- [11] 김의중, *알고리즘으로 배우는 인공지능, 기계 학습, 딥러닝 입문*, 위키북스, 서울, 2016.
- [12] 김형준, 류두진, 조훈, “개인부실채권 상환약 정채결 및 완제 행태 분석”, *재무연구*, 제32권, 제2호, 2019, pp. 187-219.
- [13] 김희철, 박형근, “패널 데이터모형을 적용한 국내 일반은행 연체율 결정요인 추정에 관한 연구”, *한국산학기술학회*, 제11권, 제1호, 2010, pp. 272-277.
- [14] 민경록, 고희진, 이재현, 위경우, “국내 신용카드 사용자의 연체요인에 관한 실증연구”, *대한경영학회지*, 제20권, 제4호, 2007, pp. 1953-1976.
- [15] 박소희, 최대선, “개인신용정보 표본DB 기반의 대출 현황 분석 및 채무불이행 예측성능 비교”, *정보과학회논문지*, 제46권, 제7호, 2019, pp. 627-635.
- [16] 이장연, 임영주, “금융위기 이후 저신용 가계 차주 현황”, *BOK 이슈노트*, 제2014-2호, 2014, pp. 1-11.
- [17] 성영애, “가계부채 연체유형과 관련요인에 관한 연구”, *소비자학연구*, 제30권, 제3호, 2019, pp. 27-49.
- [18] 손병환, *대위변제와 근저당권의 이전에 관한 연구* (석사학위 논문), 부산대학교, 2018.
- [19] 이동걸, 전성인, 정재욱, 변동준, “우리나라 가계부채 연체결정요인 및 취약성 연구”, *금융연구*, 제28권, 제2호, 2014, pp. 137-178.
- [20] 이영섭 역, *데이터마이닝 Cookbook*, 교우사, 서울, 2013.
- [21] 정현승, 강창완, 김규곤, “불균형 데이터에 대한 오버샘플링 효과 연구”, *Journal of Korea Data Analysis Society*, 제10권, 제4호, 2008, pp. 2089-2098.
- [22] 조재현, “가계대출 가구의 연체사유별 특징 및 연체 결정요인 분석”, *부동산분석*, 제6권 제1호, 2020, pp. 1-17.
- [23] 최성민, “개인신용평가모형의 변화와 시사점”, *한국신용정보원 CIS이슈리포트*, 2018-4호, 2018, pp. 1-12.
- [24] 최종원, “개인도산 현황 및 예측”, *NICE Credit Insight*, 2014년 가을호, 2014, pp. 22-43.
- [25] 코리아크레딧뷰로, 개인신용평가체계 공시, Available at https://www.allcredit.co.kr/download/KCB_Credit_Evaluation_System.pdf.
- [26] 한국금융연구원, “가계부채 백서”, *KIF 금융리포트*, 제2013-1호, 2013.
- [27] 한국금융연구원, “신용정보 집중관리체계 등 제도개선을 위한 연구”, 금융위원회 정책연구자료, 2014, Available at http://www.prism.go.kr/homepage/entire/retrieveEntireDetail.do?pageIndex=1&research_id=1160100-201400035&leftMenuLevel=160&cond_research_name=%EC%8B%A0%EC%9A%A9%EC%A0%95%EB%B3%B4+%EC%A7%91%EC%A4%91%EA%B4%80%EB%A6%AC%EC%B2%B4%EA%B3%84+%EB%93%B1+%EC%A0%9C%EB%8F%84%EA%B0%9C%EC%84%A0%EC%9D%84+%EC%9C%84%ED%95%9C+%EC%97%B0%EA%B5%AC&cond_research_start_date=&pageUnit=10&cond_order=3.
- [28] 한국신용정보원, 일반신용정보관리규약(제2020-5차 신용정보집중관리위원회), 2020, Available at <http://www.kcredit.or.kr/info/allRawView.do?contentNo=838>.
- [29] 한준태, 정진아, “학자금 대출 연체의 신용위험 평점 모형 개발”, *한국데이터정보 과학회지*, 제27권, 제5호, 2016, pp. 1293-1305.
- [30] 함준호, 조현철, 권영철, “가계부채의 경기순응성과 거시정보결합 신용평가에 대한 연구”, *한국경제의 분석*, 제21권, 제1호, 2015, pp. 1-63.
- [31] 허석균, “DII, LTV 및 대출상환 조건이 주택담보대출의 연체율에 미치는 영향”, *규제연구*, 제21권, 제2호, 2012, pp. 39-77.

- [32] 홍봉주, “제3취득자의 변제자 대위”, *일감법학*, 제38권, 2017, pp. 455-475.
- [33] Bhutta, N., J. Dokko, and H. Shan, “The depth of negative equity and mortgage default decisions”, *Finance and Economics Discussion Series Working Paper*, Federal Reserve Board of Governors, 2010-35, 2010.
- [34] Campbell, J. Y. and J. F. Cocco, “A model of mortgage default”, *CFS Working Paper*, No.452, 2015.
- [35] Campbell, T. S. and J. K. Dietrich, “The determinants of default on insured conventional residential mortgage loans”, *The Journal of Finance*, Vol.38, No.5, 1983, pp. 1569-1581.
- [36] Dastile, X, T. Celik, and M. Potsane, “Statistical and machine learning models in credit scoring: A systematic literature survey”, *Applied Soft Computing Journal*, 2020, pp. 1-21.
- [37] Durand, D., *Risk Elements in Consumer Instate-ment Financing*, National of Bureau of Economic Research, 1941.
- [38] Elizabeth Mays, 김민정 역, *Credit Risk Modeling 신용위험평가모형: 이론과 실제*, 넥스트웨이브, 서울, 2004.
- [39] Getter D. E., “Contributing to the delinquency of borrowers”, *Journal of Consumer Affairs*, Vol.37, No.1, 2003, pp. 86-100.
- [40] Gross, D. B. and N. S. Souleles, “An empirical analysis of personal bankruptcy and delinquency”, *Review of Financial Studies*, Vol.15, No.2, 2002, pp. 319-347.
- [41] Krainer, J., S. LeRoy, and O. Munpyung, “Mortgage valuation and mortgage default”, *FRBSF Working Paper*, 2009-20, 2009.
- [42] Li, X. L. and Y. Zhong, “An overview of personal credit scoring: Technique and future work”, *International Journal of Intelligence Science*, 2012, pp. 181-189.
- [43] NICE CB연구소, “NICE 개인도산 예측 평점”, *NICE Credit Insight*, 2012년 여름호, 2012, pp. 74-79.
- [44] NICE평가정보, 개인신용평점체계 공시, 2020, Available at http://www.niceinfo.co.kr/creditrating/cb_score_1_1.nice.
- [45] Political and Economic Research Council (PERC), “Give Credit where Credit is Due: Increasing Access to Affordable Mainstream Credit Using Alternative Data”, 2006.
- [46] Saltelli, A., S. Tarantola, F., F. Campolongo, and M. Ratto, *Sensitivity Analysis in Practice - A Guide to Assessing Scientific Models*, John Wiley, 2004.

A Study on the Factors of Normal Repayment of Financial Debt Delinquents

Sungmin Choi* · Hoyoung Kim**

Abstract

Credit Bureaus in Korea commonly use financial transaction information of the past and present time for calculating an individual's credit scores. Compared to other rating factors, the repayment history information accounts for a larger weights on credit scores. Accordingly, despite full redemption of overdue payments, late payment history is reflected negatively for the assessment of credit scores for certain period of the time. An individual with debt delinquency can be classified into two groups; (1) the individuals who have faithfully paid off their overdue debts(Normal Repayment), and (2) those who have not and as differences of creditworthiness between these two groups do exist, it needs to grant relatively higher credit scores to the former individuals with normal repayment.

This study is designed to analyze the factors of normal repayment of Korean financial debt delinquents based on credit information of personal loan, overdue payments, redemption from Korea Credit Information Services. As a result of the analysis, the number of overdue and the type of personal loan and delinquency were identified as significant variables affecting normal repayment and among applied methodologies, neural network models suggested the highest classification accuracy. The findings of this study are expected to improve the performance of individual credit scoring model by identifying the factors affecting normal repayment of a financial debt delinquent.

Keywords: Credit Information, Loan, Delinquencies, Normal Repayment, Repayment History

* Manager, Korea Credit Information Services

** Corresponding Author, Manager, Korea Credit Information Services

● 저 자 소 개 ●



최 성 민 (abcdchoi@gmail.com)

서울대학교 산업공학과를 졸업하고 한국과학기술원에서 경영학 석사를 취득하였으며 연세대학교 기술경영학협동과정에서 박사과정을 수료했다. LG CNS, 하나금융경영연구소 등에서 근무했으며, 현재 한국신용정보원 선임조사역으로 재직 중이다. 주요 관심분야는 신용평가모형, 머신러닝, 금융혁신, 데이터 기반의 의사결정 지원 등이다.



김 호 영 (hoyoung.kim@kcredit.or.kr)

서강대학교에서 경영학 석사(Business Analytics 전공)를 취득하였으며, 현대정보기술, 한국재정정보원을 거쳐 현재 한국신용정보원 선임조사역으로 재직 중이다. 주요 관심분야는 금융 빅데이터 분석, 신용정보 데이터의 가공 및 처리, 분석 등이다.

논문접수일 : 2020년 10월 20일

게재확정일 : 2020년 11월 16일

1차 수정일 : 2020년 11월 09일