

개인신용위험관리에 대한 실증적 연구

연세대학교 대학원

경제학과

김 학 신

개인신용위험관리에 대한 실증적 연구

지도 이 영 선 교수

이 논문을 석사 학위논문으로 제출함

2003년 12월 일

연세대학교 대학원

경제학과

김 학 신

김학신의 석사 학위논문을 인준함

심사위원_____인

심사위원_____인

심사위원_____인

연세대학교 대학원

2003년 12월 일

차 례

그림 차례	iv
표 차례	iv
국문 요약	v
제 1 장 서론	1
제 1 절 연구의 배경과 목적	1
제 2 절 연구의 방법과 범위	2
제 2 장 신용평점의 개념	3
제 1 절 위험관리	3
1. 위험관리(CRM : Credit Risk Management)	3
2. CTI(Computer Telephony Integration)	6
3. 고객관리(CRM : Customer Relationship Management)	6
4. 개인신용평점시스템(CSS : Credit Scoring System)	7
제 2 절 신용평점의 의의와 배경	8
1. 신용평점의 의의	8
2. 신용평점의 배경	11

제 3 장 신용평점시스템의 개요 및 평점표 개발절차	13
제 1 절 신용평점시스템의 구성	13
1. 신청평점시스템	13
2. 행동평점시스템	14
3. 신용평점시스템의 도입배경	15
제 2 절 평점표 개발절차	15
1. 샘플링(Sampling)	16
2. 탐색 및 조정(Exploration & Modification)	16
3. 1차분류(Fine Class)	17
4. 2차분류(Coarse Classing)	17
5. 모형화 및 점수할당(Modeling & Point Assignment)	19
6. 거절점수 설정	20
7. 평점표 검증(Validation)	22
제 4 장 신용평점모형의 실증분석	23
제 1 절 모형개발 개요와 개발절차	23
1. 모형개발 개요	23
2. 모형개발 절차	25
제 2 절 모형의 추정결과	26
1. 선정 변수의 유의성 분석	26

2. 통계적 분석 결과	27
3. 평점표	29
4. 행동평점 분포도	30
5. 행동평점 모형 성과표	32

제 5 장 결론	34
----------	----

참고 문헌	36
영문 요약	40

<그림 차례>

<그림 4-1> 행동평점 분포도	31
-----------------------------	----

<표 차례>

<표 3-1> 결혼기간에 따른 속성	19
<표 4-1> 관찰예측기간표	23
<표 4-2> 기준일과 예측일 현재 우량과 불량 구성비	24
<표 4-3> 변수의 유의성	24
<표 4-4> 변수별 통계값	27
<표 4-5> 모형개발집합(Training Set)과 모형검증집합(Validation Set)의 구성 도	28
<표 4-6> 모형에서의 적중률 비교	28
<표 4-7> 변수별 평점표	29
<표 4-8> 행동평점 모형 성과표	32

국 문 요 약

개인신용위험관리에 대한 실증적 연구

1997년 외환위기 이후 국내 금융환경은 극심한 변화를 겪었다. 수많은 금융기관들이 합병되거나 퇴출을 당하였다. 이런 치열한 경쟁 환경속에서 각 금융기관들은 살아남기 위해 여러 선진금융기법들을 도입하였다. 그 하나가 바로 신용평점제도(Credit Scoring System)이다.

본 연구의 목적은 다음과 같다. 위험관리기법중의 하나인 신용평점제도를 분석하여 이것이 금융기관의 수익성 증대에 얼마나 효과적인지 검증하고 나아가 모형개발의 필요성 및 연구방향을 제시하고자 한다.

본 연구에서는 개인대출정보를 이용하여 그것을 바탕으로 각 속성에 대한 유의성을 검증하고 선택된 속성을 로지스틱 회귀분석(Logistic Regression Analysis)을 통하여 신용평점모형을 만든다. 실증분석에 사용된 로지스틱 회귀분석 모형은 다음과 같다.

소비자 j 의 채무불이행 확률 = $\text{prob.}(Y_j = 0)$

$$= \frac{1}{1 + e^{\beta X_j}}$$

단, $\beta = (\beta_1, \beta_2, \beta_3, \dots, \beta_n)$, $X_j = (X_{1j}, X_{2j}, X_{3j}, \dots, X_{mj})$

소비자 $j = 1, 2, \dots, n$

그리고 그것을 K-folder Cross Validation을 이용하여 검증을 한다. 이렇게 만들어진 신용평점모형과 기존에 정의되어 나누어진 신용우량과 불량률 비교하여 그 모형이 얼마나 효율적으로 신용우량과 불량률 구분하는지 비교할 것이다.

본 연구를 통하여 얻은 결론은 개인신용평점을 도입하기전의 불량률이 13.61%

였으나, 이를 도입하고 적용한 이후의 불량률은 2.24%로 감소할 수 있음을 알 수 있었다.

이와 같은 분석의 결과를 통해 얻을 수 있는 결론은 첫째, 기존의 전통적 신용평가방법에 비해 개인신용평점모형을 도입하여 함께 이용하는 것이 신용우량과 불량률의 더욱 효율적으로 구분하여 금융기간의 수익성을 증대시킨다는 것이다. 둘째 신용평점모형을 이용하면 기존에 비해 비용과 시간을 단축시켜 더욱 효율적이라는 것이다. 마지막으로 이러한 신용평점모형을 이용하여 신용관리 역량을 강화한다면 갈수록 치열해지는 경쟁상황에서 경쟁우위를 점할 수 있다는 것이다.

핵심되는 말 : 개인신용평점모형, 로지스틱 회귀분석, K-Folder Cross Validation

제 1 장 서론

제 1 절 연구의 배경과 목적

1997년 11월 외환위기 이후 국내 금융환경은 격심한 변화를 겪었다. 특히 은행의 퇴출 및 기타 흡수 합병을 통한 금융구조 조정은 앞으로 국내 금융기관이 어떻게 변화를 해야 하는지 경쟁상대가 누구인지 명확하게 보여주고 있다. 즉, 선진금융기관의 국내 진출을 허용하고, 시중은행의 해외매각을 통한 해외자본 유치 정책은 현재 우리가 예상하지 못한 변화를 초래할 것이다. 1998년도 금융기관은 수없이 많은 기업의 부도와 개인의 파산을 통하여 심각한 부실을 떠안았으며, 이러한 부실을 상쇄하기 위하여 기존의 우수 고객(신용도가 높고 위험이 낮은 고객)에게 터무니없는 고이율을 할당하여 이들의 비용구조를 개선하여 막대한 이익을 실현하였다.

90년대 이후 가계금융의 급속한 수요증가는 금융기관의 정책이 도매금융에서 소매금융으로 전환하도록 만들었으며, 이는 위험의 분산을 통한 이익 추구 전략이라고 간주된다. 일례로 씨티은행의 경우 도매금융 집중전략으로 커다란 위기에 봉착하여 시련을 겪은 이후 전반적인 경영전략을 소매금융으로 전향하였으며, 소매금융 전문기관인 퍼스트 시큐리티은행은 고객정보파일에서 고객별 수익성을 분석해 본 결과, 예금잔액 기준 상위 10%의 고객이 은행수익의 85%를 차지하고 있는 것으로 나타났다. 국내 금융기관들도 선진 금융기관과 동일한 소매시장을 대상으로 영업을 하고 있으나, 실현 수익률, 관리비용 및 방법에는 상당한 차이를 보여주고 있다.

선진금융기관과의 치열한 경쟁을 피할 수 없는 세계화 물결 속에서 국내 금융기관들은 기술적인 도구를 이용한 신용과 위험의 분석을 시도하고 있으며, 이미 몇몇 금융기관들은 초기단계로 이를 도입하여 업무에 부분적으로 활용하고 있다. 선진금융기관의 경우, 수십년 전부터 주택자금구입대출, 일반자금대출, 신용카드 등을 취급하는 여신 전문 기관들은 개인고객 대상의 신용을 신중하게 평가하기

위해 노력하여 왔다. 이들 기관이 위험으로 간주하는 것은 무엇이며, 신용이라 여기는 부문은 어떠한 것인지 그리고 신용과 위험의 관계는 어떠한 형태를 나타내는지를 분석하는데 많은 시간, 노력 그리고 인력을 투입하여 왔다. 또한 우수고객의 효율적이고 체계적인 관리방법 및 부실고객을 사전에 방지하기 위해서도 많은 노력을 아끼지 않았다.

이를 총체적으로 위험관리(Credit Risk Management)라 한다. 신용위험관리방법에는 CTI(Computer Telephony Integration), 고객유대관리(CRM : Customer Relationship Management), 신용평점제도(CSS : Credit Scoring System)등이 있다. 본 논문에서는 개인신용평점제도의 이론적 배경과 실증분석을 통하여 금융기관의 수익성 증대(부실고객 방지)에 얼마나 효과적인지 검증하고, 모형개발의 필요성 및 연구방향을 제시하고자 한다.

제 2 절 연구의 방법과 범위

본 논문에서는 고객별 특성에 따라 차별화 된 위험관리 도입에 필요한 신용평점제도(Credit Scoring System)에 대하여 상세히 연구하고자 한다. 2장에서는 위험관리방법의 유기적 관계와 개인신용평점제도의 의의 및 배경에 대해 알아보고, 3장에서는 현재 국내 금융기관들이 도입하고자 하는 신용평점제도에 대한 개요 및 평점표 개발절차에 대해 논하며, 4장에서는 신용평점모형을 실증분석하며, 마지막 5장에서는 본 논문의 결론 및 한계점을 기술할 예정이다.

제 2 장 신용평점의 개념

본 장에서는 각 위험관리의 유기적 관계를 살펴보고, 신용평점시스템의 의의와 배경에 대하여 자세히 설명하였다.

제 1절 위험관리

1. 위험관리(CRM : Credit Risk Management)

금융사업은 위험관리를 통해 수익을 창출하는 사업이다. 여기서 위험은 금융상품 가격을 결정하는 요소로 이는 이자율을 결정하는 중요한 요인이라 할 수 있으며, 고위험에는 고금리의 적용을 원칙(High Risk, High Return)으로 한다. 따라서 금융사업의 체계적이고 효율적인 위험관리는 은행의 경쟁력 및 수익성을 제고시키는 핵심적 역할을 하게 된다. 여기서 효율적이고 체계적인 관리란 고위험을 통한 고수익의 이익창출이 아니라 각종 요인들을 면밀히 분석하여 그들이 수용할 만한 위험의 범주내에서만 위험관리를 수행하고 이러한 범주에 해당하는 고객만을 대상으로 하는 것을 기본 원칙으로 한다. 이는 금융사업의 안정성 원칙에 입각한 것이라 볼 수 있다.

금융위험에는 다섯 가지 유형이 있다. 첫째, 시장 위험(Market Risk)은 내부 통제 불가능한 요소인 금리 또는 환율 등의 변화로 발생하는 손실 및 이익의 위험이다. 둘째, 신용 위험(Credit Risk)은 계약 상대방이 계약을 이행하지 않음으로써 손실을 입게 될 위험이다. 예를 들면, 은행이 대출을 해 주었을 경우 만기일에 원리금(원금 및 이자)의 회수가능성에 대한 위험을 말한다. 셋째, 유동성 위험(Liquidity Risk)은 시장의 자금 유동성 결여 즉, 적절한 현금흐름의 부족으로 인해 입게 되는 위험을 의미한다. 넷째, 경영관리위험(Operational Risk)은 경영관리의 부주의나 비합리적인 내부관리제도로 인하여 발생할 수 있는 위험이다. 다섯째, 법률적 위험(Legal Risk)은 실물경제의 급속한 발전에 비해 제도적인 보완이

시대에 뒤떨어져 있을 경우 발생할 수 있는 국가 총체적인 위험을 말한다.

5가지 위험중 시장위험과 신용위험만이 계량이 가능한 위험이며, 통제가 가능한 위험으로는 신용위험과 경영관리위험이 있다. 나머지 위험들은 계량이 불가능한 위험이다. 특히, 신용위험은 통제가 가능하면서 고객을 대상으로 계량이 가능한 위험이다. 따라서 금융기관의 거래고객 및 거래량의 증가로 인하여 신규고객유치 및 기존고객에 대한 효율적이고 체계적인 위험관리의 방법들이 필요하게 되었다.

위험관리방법에는 4가지 방법이 있는데 구체적으로 살펴보면 아래와 같다.

첫째, CTI(Computer Telephony Integration)로 이는 모든 고객에게 동질의 서비스를 제공하고 위험발생 가능한 고객에 대한 가벼운 위험관리 수단으로 이용할 수 있는 기능을 제공한다.

둘째, 고객관리(CRM : Customer Relationship Management)로서 신상품 및 신시장 정의(Identifying New Products and New Markets)를 위한 고객 데이터베이스를 체계적으로 분석하여 응답률이 높고 수익성이 높은 잠재고객집단을 파악하고 관리하도록 하는데 목적을 두고 있다.

셋째, 개인신용평점제도(CSS : Credit Scoring System)로 신청자의 신용도 평가(Assessing an Applicant's Credit Worthiness)를 통계적 기법을 이용하여 우량고객과 불량고객을 선별하여 손실을 최소화하도록 하는 관리방법이다.

넷째, 위에서 기술한 3가지 방법을 모두 통합하여 정확하고 일관성 있는 시스템의 유지(Maintaining an Accurate and Consistent System) 및 상호연관성을 취적화하여 활용함으로써 신용위험의 최소화로 수익의 극대화를 실현할 수 있는 방법이다. 위의 위험관리방법은 상호 유기적인 관계 속에서 역할을 수행한다고 할 수 있다.

한편, 신진금융 기관들은 80년 초부터 신용위험의 노출을 피하기 위하여 대출취급시 위험관리 시스템을 활용하여 왔으며, 이러한 시스템은 신용관리의 핵심부문이었다. 이와 관련하여 Gradstrom은 신용위험 관리를 효과적으로 수행하기 위한 7가지 준수사항을 제시하였으며, 이는 경영관리 위험을 최소화할 수 있는 방법이라 볼 수 있다.¹⁾

1) Jhon Gradstrom(1996) p.55~60

첫째, 균형 잡힌 신용정책(A Balanced Credit Policy); 은행원의 임무는 한편으론 건전하고 보수적인 신용공여자로서, 다른 한편으로는 사업개발자로서의 이중적인 역할을 동시에 수행하여야 한다. 따라서 부서간의 추진방향이 각 부서의 의도와는 반대로 움직일 수 있다. 의사결정자는 총체적으로 균형있는 신용정책을 구사해야 한다.

둘째, 신용평가과정에 참여하는 전직원(여신담당자, 심사역, 관리자)들이 일반적으로 쉽게 인식할 수 있도록 분명하게 정의된 위험평가체제(A Clearly Defined Risk Rating System); 신용위험과 관련한 용어의 통일성, 이해성 그리고 일반성으로 위험정도를 분명하고 정확한 방법으로 위험측정시스템에 통합하여야 하며, 미국의 위험관리협회에서 제안하는 방법은 부채대가치(Debt-To-Worth)비율로 부채정도를 측정하는 것이다.

셋째, 시의 적절한 위험평가의 적용(Timely Application of Risk Rating); 승인된 대출에 대한 만기까지의 신용변화 측정이 중요하다. 즉, 수용 가능한 위험단계에서 궁극적인 부실로 될 때까지 신용이 적절하게 측정되었는지, 사후관리 활동이 단계별로 적절하게 수행되었는지를 처음부터 파악해야 한다.

넷째, 신용이 정상에서 불량으로 전이할 때 효과적인 조기경보위험 평가의 활용(Use of Early Warning Risk Rating); 위험관리에 있어서 조기경보위험평가는 핵심적이며, 이는 미래 발생 가능한 문제를 예측하는데 있다.

다섯째, 대출전문인력의 효과적인 참여(Effective Loan Officer Participation); 기업의 보상정책이 신규사업이나 단기간의 수익성에만 국한하고 있는 관계로 여신담당자는 소외되는 경향이 강하다. 그러나 기업의 장기적인 안목에서 살펴보면 이들의 참여는 장기적으로 수익성이 높은 고객의 유치를 이끌어 준다.

여섯째, 심사역의 독립적인 신용평가(Independent Credit Review); 대출신청자의 신상정보 및 거래정보 또는 경영관리능력정보와 기업의 시장상황 등은 담당심사역이 가장 많이 알고 있으므로 이들의 신용평가는 독립적으로 수행되어야 한다.

일곱째, 부실대출로부터의 교훈(Learning from Charged-off Loans); 부실에 대한 분석은 여신담당자, 사후관리 담당자 그리고 담당관리자에 의해서 수행되어진다. 부실의 단계별 시의 적절한 정보의 제공 여부, 적절한 조치 여부, 유사한 부실

을 미래에 재발하지 않도록 조치 여부 등이 있다.

이러한 7가지 사항은 효율적인 위험관리가 여신을 전문적으로 취급하는 기관에게 있어서는 기업의 환경을 바꾸게 하는 내용이 될 수 있다. 즉, 전문가의 의견을 중심으로 모형을 구축하고 개발된 모형이 전문가의 관점에서 지속적으로 관리됨으로써 보다 효과적이고 효율적인 여신업무를 운영할 수 있도록 하여준다는 것이다.

2. CTI(Computer Telephony Integration)

CTI란 전화기술과 컴퓨터 기술의 기능적인 결합에 의해 기존에 상담원이 모두 분담했던 고객확인, 요구접수, 정보입수, 거래처리 등과 같은 단위업무를 컴퓨터와 통신기기, 지능형 정보처리장치가 자동 분담 처리함으로써 시간, 효율 대 고객이미지 제고면에서 효과적으로 운영할 수 있는 시스템이다. 국내에서 CTI를 구축한 대표적인 곳은 금융권 및 통신사업자의 고객안내센터, 항공권 예약센터, 자동차회사, 공공기관 등이다. 대표적으로 은행은 일찍이 텔레뱅킹이나 자동연체독촉시스템(ACS : Autocall Collection System)등을 통하여 데이터웨어하우징을 구축하고, 보험업체는 텔레마케팅과 연계된 CTI에 각각 주력하고 있다. 이는 은행은 대출업무와 관련해 체계적인 고객 정보가 필요하고 보험회사는 새 고객을 찾는 것이 중요한 문제로 사람이 해야 할 일들을 크게 줄여 줄뿐 아니라 체계적인 데이터 관리를 통하여 정확하고 다양한 서비스를 제공하는 수단이 되기도 한다. 뿐만 아니라 CTI를 통하여 수집 관리된 자료를 통하여 고객과 지속적인 관계를 유지하는 방법으로 이용되어 질수 있고, 개인신용평점을 산출하기 위한 기초 자료로서 활용되어지기도 한다.

3. 고객관리(CRM : Customer Relationship Management)

고객관리란 은행과 고객간 상호 이익이 될 수 있는 장기적이고 지속적인 관계를 구축하고 유지하기 위한 과정을 의미한다.

고객에 대한 원가 관리나 위험 또는 신용 관리와 같이 협의의 마케팅 범위를 벗어나는 기능과 관련된 부분들을 포함하여 고객과 관련된 사항을 종합적으로 파악함으로써 고객을 능동적이고 효과적으로 관리하고자 하는데 초점을 둔다.

고객관리 과정에서의 전체업무를 크게 두 가지로 나누어 본다면 첫째, 고객 및 기타 환경에 대한 데이터 분석이고 둘째, 결정된 전략을 바탕으로 개별적인 고객과의 관계를 관리하기 위하여 실행하는 구체적인 활동이다.

고객관리를 통해 얻을 수 있는 이익은 매우 다양하다. 우선 우수고객의 유지 비율을 제고할 수 있으며 고객 이탈로 인한 손실의 최소화 및 또한 잠재고객을 활성화시키고 교차판매(Cross-Sell), 재구매(Re-Sell), 판매액 및 판매단가의 증대(Up-Sell)를 유도하고 틈새시장을 개척하는 등 시장점유율(Market Share)을 확대시키고, 우수고객의 유지관리를 통한 수익 증대 등 여러 가지 효과를 기대할 수 있다.

이들 제반 목표들을 달성하기 위한 실행계획에는 단계별 업무 실행 일정과 업무를 지원하기 위한 정보시스템 구축일정이 우선되어야 한다. 즉 데이터를 추출 및 정제하여 데이터 웨어하우스 또는 CRM 통합 데이터베이스를 구축하는 작업과 분석된 정보를 기반으로 실제 고객 관리 업무를 실행하기 위한 CTI에 대한 기반 환경을 구축하는 작업 등이다.

고객관리는 어떤 부분에 상대적으로 중점을 둬으로써 보다 효과적인 그리고 효율적인 고객관리가 가능해질 것인가를 분명히 하는 역할을 하며 장기적으로 고객관리방법을 어떻게 발전시켜 나갈 것인가 제시하여 준다. 고객관리의 성공적인 수행을 위해서는 고객별(행동 예측)점수의 생성, 즉 개인신용평점의 전제조건인 통합 데이터베이스의 구축이 우선되어야 할 것이다.

4. 개인신용평점시스템(CSS : Credit Scoring System)

신용평점이란 고객특성을 분석하여 미래의 신용능력을 예측하는 통계적 방법으로, 고객특성변수는 신청서(고객정보), 거래정보, 외부기관 정보 등에서 파악되고, 각각의 특성은 2-3가지의 속성(Attributes)이라는 것으로 나누어진다. 그리고

각각의 속성은 점수로 되어 있으며, 신청서의 해당 속성의 점수를 합하여 총점을 제공한다.

신용평점은 신청평점시스템(Application Scoring System)과 행동평점시스템(Behavior Scoring System)으로 구분되어 고객의 신용을 계량화한다. 신청평점시스템은 신청서상의 정보 및 외부정보기관의 정보를 이용하여 신용공여 가부를 결정하는데 이용되며, 장점으로는 신청서 처리의 효율 및 효과화, 신용조사비용 절감, 일관된 의사결정, 연체율 및 신용손실 감소화 그리고 고객의 이용실적 예측능력 향상 등이 있다. 행동평점시스템은 금융기관이 제공한 신용한도 내에서 예정한 날짜에 정확하게 신용을 지키는지 등을 파악하여 효과적인 사후관리, 연체상태 접근고객 사전 방지, 신용손실 감소, 고객과의 관계 개선 등 부실을 방지하고 우수고객에 대한관리를 효율적으로 수행하는데 목적을 두고 있다.

다음절에서는 신용평점의 의의 및 배경을 국내외문헌을 통하여 보다 자세하게 고찰하였다.

제 2 절 신용평점의 의의와 배경

1. 신용평점의 의의

신용평점의 변천과정을 살펴보면, 초기 도입단계에서는 신용평점의 가치를 모형개발방법론 자체에 두었으나, 개발 자료의 정확한 축적 및 개발기술의 공유로 인하여 여러 사업 분야의 신용평점은 하나로 수렴하게 되었다. 즉, 신용평점을 만드는 것만으로는 경쟁우위에 있을 수 없다는 것이다. 따라서 신용평점을 적용할 수 있는 자료수, 범위 그리고 창의성을 토대로 신용평점을 적용시킬 수 있을 정도의 이해가 필요하게 되었다. 이렇게 만들어진 신용평점은 신용공여에 대한 의사결정의 신속성, 정확성, 객관성 그리고 일관성을 유지할 수가 있었다.

평점표의 유형에는 일반적으로 Credit Scores(Generic Scores)와 Application Specific Scores(Custom Scores)가 있다. 이의 차이점은 개발 자료도 상이하고, 개

발주체도 상이하다. 첫째, Generic Scores란 신용정보회사의 정보만을 이용하여 고객의 미래 지불능력(future repayment performance)을 예측하기 위하여 만든 신용 평점이다. 이는 신용정보회사의 정보, 과거지불행위, 신용거래 정보, 생활신용정보(전기세, 전화세,...)등이 이용될 수 있고, 인구통계자료(직업, 우편번호등), 소득, 지불능력, 상품정보 등의 활용은 불가능하다. 대표적인 예로는 FICO Credit Score 기법이 있고 비용이 저렴하다는 장점이 있다. 둘째, Custom Scores(Application Specific Scores)란 신청서의 모든 자료를 이용하여 특정한 결과(a specific outcome : foreclosure on the mortgage)를 예측하기 위하여 개발한 평점모형으로 신용정보회사정보, 년소득, 상품, 보증인, 지역경제자료, 담보가치비율(Loan to Value), 담보정보 등이 이용되어질 수 있고, 장점은 Generic Scores보다 2-3배 예측력이 정확하다는 점과 사용자의 충분한 이해가 가능하다는 점이다.²⁾

그리고, 신용평점의 의미는 소비자 신용측면, 경제적 측면, 기술적 측면 세가지 측면에서 살펴볼 수 있다.

(1) 소비자 신용 측면

신용이란 장래 어느 시점에서 그 대금을 지급할 것을 약속하고 현재의 경제적 가치를 획득할 수 있는 능력이라고 정의할 수 있다. 신용의 개념에는 위험(Risk)의 개념이 내재되어 있고, 신용공여기관은 신용을 신청하는 개인들에 대해서 위험을 추정하고, 어느 위험수준에서 이들의 신용사용한도를 결정할 것인지를 판단하여야 한다. 신용공여기관은 어느 수준까지 신용위험을 수용할 것이며, 이에 따른 소비자 신용의 위험을 어느 수준까지 어떤 변수와 연계시켜 가져갈 것인지에 대해서 정책적으로 결정하게 된다.

Golden은 신용공여기관이 가장 많이 이용하고 있는 전통적인 방법의 우량고객 판단요소와 불량고객 판단요소를 연구하였다. 우량고객의 5가지 요소로는 성격(Character : 가족상황, 개인신상, 습관 등), 상환능력(Capacity : 소득, 직업등), 재산(Capital : 채무불이행시 위험에 대비한 재정능력), 담보(Collateral : 유형자산에 대한 가치) 그리고 상태(Condition : 단기간 미래에 대한 예측가능 정보 제공 상

2) William M.Makuch,(1998) p.43~45

태)며, 불량고객을 유발시킬 수 있는 5가지 요소는 방심(Complacency : 우리 고객은 안전하다는 사고), 부주의(Carelessness : 서류미비등), 의사소통왜곡(Communication Breakdown : 의사전달이 명확하지 않을 경우), 우발성(Contingencies : 거시적 변화에 대한 불충분한 주의) 그리고 경쟁(Competition : 양정 팽창 또는 가격경쟁 등으로 인한 동업종 간의 출혈)이라고 하였다.³⁾ 우량고객 및 불량고객의 판단요소는 상호배타적이 아니어서 명확한 구분이 불가능하다는 것이다. 왜냐하면 동일한 자료가 여러 요소에서 반복적으로 사용되고 있기 때문이다.

전통적인 방법들은 개인의 신용도를 측정하고 조사하는데 너무 주관적이고 비체계적이어서 실용성에 있어서 여러 가지 문제가 많았다. 따라서 이러한 문제점을 극복하기 위해서 다양한 접근방법들이 시도되어져 왔다. 이 중 가장 대표적인 접근방법이 개인에 대한 신용평가를 그 개인이 지니고 있는 신용위험과 연계시켜서 개발한 신용평점제이다. 이는 신용위험을 측정해서 규명된 개별적 위험에 신용공여기관들이 적절하게 대처해 나갈 수 있도록 고안된 도구라 할 수 있다. 즉, 신용공여기관은 신청자의 개별적인 위험을 측정하여 신용을 공여해야 한다.

위험도의 분석이 곧 신용평가의 핵심이며, 이러한 어려움은 신용평점의 개발로 해결 할 수 있었다.

(2) 경제적 측면

98년초 미국 소비자가 신용대출 규모(주택저당대출 \$3.7조 포함)는 \$5조에 이르렀으며, 이러한 신용위험에 대한 노출 및 거래량의 증가는 효과적인 신용평가에 대한 관심을 고조시켰다. 신용거래량의 증가는 담당 전문가의 인건비 증가를 유발시켰으며, 관리자는 보다 효과적인 방법을 모색하게 되었다.⁴⁾ 여신취급절차의 가장 큰 변화로는 중앙집중화라고 할 수 있다. 이러한 업무처리방법의 변화는 미국의 90년초에 몰아친 리엔지니어링과 밀접한 관련이 있다. 이는 고객의 신용에 관한 보다 세밀한 평가를 하면서 동시에 여신 담당 인력의 구조조정을 통한 인건비

3) Sam Golden and Harry M.Walker(1996) p.14~20

4) William M.Makuch,(1998) p.66

의 최소화로 비용을 절감하고자 하는 목적에도 이용되고 있다. 소기업 여신규모가 평균적으로 적은 금액일 경우, 여신심사에 따른 심사역의 투입시간과 고임금을 고려할 때 수익성이 결코 높지 않은 것이었다. 그러나 평점모형을 이용할 경우 저임금의 신규 종업원으로도 전문 심사역의 결정과 유사한 위험자산 포트폴리오를 구성할 수 있으며, 보다 신속하고 효율적인 여신업무를 수행할 수 있다는 것을 알게 되었다.

전진금융기관은 거의 20년간 소비자신용대출에서의 신용평점시스템 이용을 성공적으로 이용하고 있다. 1995년 10월이후 여신 3000건을 평점시스템을 이용하여 1800건을 승인하였다. 1년후 이중에서 단 1건만이 30일 연체로 나타났다.⁵⁾

신용평점모형은 객관적인 위험평가도구로서 첫째는 부실예측이 아니라, 지불 문제에 대한 가능성을 예측하고, 둘째는 승인률을 높이는 것이 아니라, 과거의 연체율과 동일하게 유지하면서 고객관리에 있어서 일관성 및 효율성을 제고하는데 있다고 하였다.

(3) 기술적 측면

신용관련 담당 전문가의 노하우를 기계적으로 만들 수 있는 방법이 부단히 연구되어온 결과 90년대이후 대량의 자료를 보관하고 분석할 수 있는 컴퓨터의 제공 및 통계적 기법은 이들 전문가의 의사결정을 전산적으로 동일하게 만들 수 있는 기회를 제공하였다. 즉, 전문가 시스템이라 할 수 있는 이러한 방법은 고임금 전문가의 고용대신 일반사원의 고용만으로도 동일한 결과를 얻을 수 있는 시스템을 지닐 수 있게 되었다. 그러나 개발초기 단계 및 운영상의 정착단계까지는 이들 전문가의 깊은 참여가 필수 불가결한 요소인 것이다.⁶⁾

2. 신용평점의 배경

(1) 신용평점의 역사

5) Beverly Foster(1996) p.61~66

6) William M.Makuch,(1998) p.59~60

신용평가시스템의 역사는 1930년대까지 거슬러 올라간다. 이 시스템의 개척자는 미국의 Spiegel Inc. 의 경영자인 헨리 웰츠였다. 그는 제 2차 세계대전 중에 그 회사의 신용분석가들이 근무 중에 사용할 수 있는 시스템을 개발하였다. 그러나 당시의 모델은 극히 초보적인 수준이었으며, 오늘날과 같은 통계적 기법을 이용한 신용평점제도는 1950년대 이후의 일이었다. 이에 대한 최초의 이론적 발표논문은 1941년에 미국의 Durand의 연구였으며, 그는 은행, 자동차 할부사 등의 신용정보 자료를 활용하여 판별분석(Fisher's Discriminant Analysis)방법으로 신용평점모형을 개발하였다. 이후 1949년에 월버스가 백화점 고객의 신용정보를 기초로 모형을 개발하였으며, 실제 적용 결과 약간의 우량고객을 잃기는 했으나 약 7% 정도의 손실을 줄일 수 있었다.

1950년대 들어서는 이를 전문적으로 실용화하여 판매하는 회사가 등장하기 시작하였으며, 1970년대 이후에는 미국의 거의 모든 신용공여기관에서 이를 도입할 정도로 활성화 되었다.⁷⁾

(2) 신용평점제도의 철학적 배경

신용평점제도는 고객정보를 이용하여 미래 지불능력을 예측하는 것으로 스코어링시스템의 기본적인 사상 및 개발과 관련한 가정에 대하여 이해할 필요성이 있다. 기본적인 가정은 첫째, 과거로부터 정보를 얻기 위하여 분석적인 기술을 이용한다는 것이다. 신규고객에 대한 미래지불능력을 예측하기 위하여 과거자료로 점수를 만든다는 것은 우량과 불량집단간에 서로 다른 이용 가능한 변수들 간의 상관관계가 있다고 가정을 한다. 둘째, 분석적 기술은 정확하게 우량과 불량집단간의 상관관계를 파악할 수 있다는 것이다. 셋째, 우량과 불량집단간 변수가 파악되면 미래는 과거를 닮아가고, 이러한 관계 또한 현재 신청인 집단의 우량과 불량을 분리하는데 이용될 수 있다는 것이다. 따라서 평점은 결코 주관적인 의견을 내포하고 있지 않다. 이러한 이론 및 가정이 성립되면 건전한 스코어카드가 가지는 요인 즉 데이터, 기술 및 개발 전문가, 규정 준수, 신청서들을 쉽게 나열해 볼 수 있다.⁸⁾

7) Noel Capon(1982) p.82~85

제 3 장 신용평점시스템의 개요 및 평점표 개발절차

제 1 절 신용평점시스템의 구성

신용평점시스템은 신청평점시스템과 행동평점시스템으로 나뉜다.

1. 신청평점시스템

신청평점 시스템이란 신규 신용거래 신청자에 대한 현재의 사회적, 경제적 요소들을 분석함으로써 미래의 신용거래 행위(Credit Performance)를 예측하는 통계적 모형으로 과거 신용거래자로부터 얻어진 경험적 정보를 바탕으로 우/불량 예측을 위한 확률을 계산하고 구조화함으로써 현재의 신청자가 우량집단 또는 불량집단으로 분류될 확률을 평점화하여 활용하는 시스템을 말한다.

즉 신규 신용거래 신청자의 신용 위험율을 추정하여 신청평점을 부여하고 이를 통하여 신용거래의 승인여부를 결정하기 위한 심사업무이며, 신용 위험율을 계량화하여 신용평점으로 환산한 신규가입 평점표는 과거 신규가입자들의 자료를 기초로 하여 우/불량 판별에 중요한 항목을 추출하고 통계적 기법을 사용하여 정확한 가중치를 계산하여 도출한다.

적용대상업무는 신용카드의 발급 및 최초한도부여, 대출승인, 보증금면제 또는 차등화, 할부거래 승인여부 결정 및 할부한도부여, 기타 신용공여를 매개로 하는 외상거래 등이다.

운영원리:

- ① 신용거래 개설을 목적으로 하는 신규 신청인의 신용 위험율을 추정하여

8) William M.Makuch,(1998) p.60~61

평점을 부여한다.

② 평점은 신규신청자의 신청서 자료를 기본으로 하여 신용정보회사의 데이터 베이스에서 추출한 신용자료 및 업체의 기본 자료를 이용하여 개발한 통계적 모형을 통하여 산출된다.

③ 평점을 기준으로 업체 측 영업정책을 반영한 업무적용 flow를 개발하여 전산화함으로써 대량의 신용심사업무를 자동화한다.

신청평점시스템은 횡단면적 자료를 이용하여 미래를 예측하는 것이므로 데이터의 질에 크게 종속되는 모형이다. 따라서 객관적이고 정확한 자료의 확보가 시스템의 성능에 매우 중요하다.

2. 행동평점시스템

행동평점 시스템은 기존 신용고객의 현재와 과거 실적(performance)자료를 분석함으로써 신용상태를 주기적으로 측정하여 고객 대응조치를 취하도록 하는 고객관리 시스템이다.

신용상태는 시간이 경과함에 따라 변하기 때문에 신청평점시스템에 의해 도출된 신용평점을 이용하여 계속적으로 신용 위험율을 예측하는 것은 문제가 있으며, 이러한 한계를 극복하기 위해서 회원 거래실적에 관련된 정보를 현재시점 이후의 연체확률로 전환하여 개별고객의 신용 위험율에 행동평점을 부여하며 이를 거래한도 조정, 한도 초과사용 승인, 신용회전(credit revolving), 채권 추심에 이용한다.

적용대상은 기 발급된 신용카드 이용자, 장기대출 수혜자, 장기할부 거래자, 장기간 반복되는 후취거래자(전화료, 통신이용료) 등이다.

운영원리:

① 기존 고객이 거래실적 및 일부 신청서 자료를 분석하여 만든 통계적 평점 모형을 통하여 기존 고객의 연체 위험율을 추정하여 평점의 형태로 산출한다.

② 평점 모형에서 산출되는 평점을 기준으로 업체의 영업 정책을 반영한 업

무적용 FLOW를 설계하여 이를 전산화함으로써 고객관리업무를 자동화한다.

행동평점 시스템은 시계열 자료를 이용하여 미래를 예측하는 시스템이므로 데이터의 질적 객관성이 확보된다. 따라서 시스템의 성능은 신청평점시스템 보다 일반적으로 우수하다. 그러나 시계열 자료라 하더라도 자료관리가 미비하거나 추출이 곤란해지는 경우가 있으므로 데이터 베이스의 관리에 신중을 기할 필요가 있다.

3. 신용평점시스템의 도입배경

신용거래가 대량으로 발생할 수밖에 없는 현실에서 신용 심사 비용과 부실 신용거래의 발생을 동시에 최소화하지 않으면 기업의 수익은 확보될 수 없을 것이다.

따라서 잠재 신용불량자를 사전에 색출하여 부실 신용거래를 사전에 예방하고 또한 신용공여 수혜자의 신용상태를 주기적으로 평가하여 고객관리에 활용할 수 있는 심사 및 관리시스템의 개발의 필요성이 대두되었다.

일관된 심사기준의 적용, 심사 결과의 정확도 제고, 신속한 심사기준의 관리 통제, 심사정책 변경에 따른 시장영향 예측, 부실채권 예방에 따른 수익증대 등 위와 같은 역할을 담당할 수 있는 신용평점 시스템의 개발이란 평점모형을 개발하고 업무적용 flow 및 지침을 확정하여 전산화하는 과정을 말한다.

제 2 절 평점표 개발절차

신용평점모형의 가정은 최근의 과거 행동자료가 미래의 행동을 가장 잘 예측해준다는 것이며, 이를 위한 전제조건으로는 적절한 자료 및 필요한 자료가 무엇인지를 결정해야 한다. 그리고 추출한 표본으로 모형을 개발한 후, 이 모형이 적절한 역할을 수행하고 있는지를 파악하기 위한 일반화작업이 필요하다. 이를 위한 방법으로는 초기 추출한 표본을 분리하여 일부는 개발용 표본으로 사용하고, 나머

지는 모형을 검증하기 위하여 사용하게 되는데 이러한 검증을 대상외 표본 검증(Out-of-Sample Test)이라 하며, 여기에 사용되는 표본을 유보표본(Holdout Sample)이라 한다.

신용평점표 개발과 밀접한 관련이 있는 데이터 마이닝은 대체로 수십메가에서 수십기가에 이르는 대용량의 데이터를 기반으로 한다. 그러나 방대한 양의 데이터를 살피는 것은 시간의 측면에서만 보아도 많은 인내를 요하게 되는 작업이 될 수 있다. 이때 고려해야 하는 과정이 바로 샘플링이다.

1. 샘플링(Sampling)

샘플링이란 방대한 양의 데이터(모집단)에서부터 모집단을 닮은 작은 양의 데이터(샘플표본)를 추출하는 것이다. 이는 데이터마이닝 작업을 하는데 있어서 시간과 비용을 절감하여 효율적인 작업진행을 돕고, 이를 통하여 데이터마이닝 작업의 무게를 조금 가볍게 할 수 있는 절차이다. 그러므로 샘플링이 적절한 가에 대한 질문을 던져보고, 필요하다면 적절한 방법을 이용하여 샘플데이터를 생성한다. 통계학에서는 이미 샘플링을 하는 여러 가지 방법론들이 개발되었는데, 그 중 가장 널리 알려진 방법이 단순임의추출법(Simple Random Sampling)으로, 이는 간단하게 설명하면 모집단을 골고루 섞어놓은 후 원하는 양 만큼을 무작위로 추출해내는 것이다. 이 외에도 층화추출법(Stratified Sampling), 계통추출법(Systematic Sampling)등이 자주 거론되는 방법으로 각 상황에 맞는 샘플링 방법을 선택하여 사용한다.

샘플링 과정에서 가장 유의할 점은 모집단을 닮지 않은, 한쪽으로 치우쳐진 샘플이 추출되는 경우로써, 이는 샘플링 후 기본적인 몇 가지의 자료탐색을 거쳐 반드시 확인하여야 한다.

2. 탐색 및 조정(Exploration & Modification)

데이터의 탐색과정에서는 이미 알고 있는 사실들을 확인하여 수치화하는 작업

을 시작으로 하여 보유하고 있는 수많은 변수들의 관계를 살펴보는 단계이다. 탐색단계는 데이터의 모양을 면밀히 둘러보아 정보화 할 수 있는 기반을 잡아가는 과정이다. 실제 한 모집단으로 알고 있던 것이 두 개 이상으로 나뉠수도 있고, 수십개 이상의 많은 변수를 실제 중요한 정보를 주는 소수의 변수로 축소시킬 수도 있으며, 기존에 있던 변수들 외에 새로운 변수를 생성해야 할 필요성을 느끼게 될 수도 있다. 이렇게 탐색한 정보의 변수들이 조정 단계에서 생성 또는 수정되어 차후 모형화 단계에서 아주 중요한 정보로써 활용될 수 있다. 따라서 신용평점표를 개발하기 위해서는 이미 수집된 특성변수의 가치를 평가해야 한다.

이러한 특성의 평가목적은 첫째, 특성변수의 수를 관리 가능한 크기로 줄이고 (통상적으로 10-20개 이용), 둘째, 서로 상관관계가 높은 특성변수의 수를 줄이는 것이다. 서로 상관관계가 높은 변수를 많이 포함하고 있을 경우 신용평점표는 특정한 자료집합에만 치중하게 되므로 추가적인 자료에 대한 예측력이 빈약할 수 있기 때문이다.

3. 1차분류(Fine Class)

1차분류는 특성변수에 대한 유용한 속성의 범주를 정의하는 것으로, 이는 정의된 특성변수가 가능한 많은 정보를 제공하도록 하기 위한 것이다. 이를 평가하기 위해 이용하는 통계분석 중 정보가치통계(information Value Statistics)모형이 있다. 정보가치란 각각의 특성이 지니고 있는 정보의 양이다. 정보가치통계는 우량과 불량을 판별할 수 있는 특성변수의 능력을 측정하는 지표가 된다. 정보가치통계 계산 후 점수대별로 특성변수를 나열하고, 속성을 많이 지니고 있는 특성변수는 2차분류(Coarse Classing)를 이용하여 적정한 수준으로 변수를 줄인다.

4. 2차분류(Coarse Classing)

2차분류란 최하 2개 최대 20개 사이의 특성변수에 대한 최적의 속성수를 결정하는 일련의 작업과정을 말한다.

속성의 범주를 감소할 필요가 있을 때에는 고객의 신용도와 특성변수와의 상관관계를 평가하여 이들 간의 관계를 잘 지켜줄 수 있도록 속성을 최적화하고, 반면 각 특성변수 당 속성의 수를 감소시킬 수 있다. 속성이 많이 있을 경우 예측력을 떨어뜨리고, 유지관리가 어렵기 때문이다.

회귀분석을 통하여 구한 특성들의 가중치 값을 평점모형개발에 그대로 사용하기에 크게 두 가지 문제점을 내포하고 있다. 첫째, 신용평점모형을 효과적으로 구축한다는 측면에서 초기에 나오는 속성들을 그대로 사용하기에는 그 수가 너무 많다. 보통 5-6개보다 많은 속성들을 이용하게 될 경우 평점오차가 크게 벌어지는 경향이 있다. 둘째 어떠한 특성변수들은 너무나 적은 관측빈도를 보이고 있어 유용한 결론을 추출해 내기가 어렵다.

이러한 점 때문에 증거비중은 속성들을 체계적으로 분류, 재구축하는데 아주 유용한 지표가 되고 있다. 증거비중은 우량고객이 될 수 있는 비율(Odds to be Good)에 자연로그를 취하는 것을 말하는데 여기서 우량고객이 될 수 있는 비율(Odds Rate)은 특성변수내의 각 속성 당 우량고객 백분율(%)을 불량고객 백분율(%)로 나눈 것을 말하고 있다.⁹⁾

$$WE_i = \ln(P_{gi}/P_{bi})$$

단 WE_i = Weight of Evidence for I-attribute

P_{gi} = I 속성에서의 우량고객 백분율

P_{bi} = I 속성에서의 불량고객 백분율

9) I=1에서 m까지, m은 변수의 개수
j=1에서 n까지, n은 변수 I에 대한 속성의 수
 g_{ij} =우량수 b_{ij} =불량수
G=샘플중 우량수
B=샘플중 불량수
이명식 (1999) p.67

<표 3-1> 결혼기간에 따른 속성

속성	우량(%)	불량(%)	증거비중
3년이하	70(9.9)	68(18.2)	-0.609
3년 - 5년	74(10.5)	32(8.6)	0.199
5년 - 20년	424(60.0)	232(62.7)	-0.044
20년이상	139(19.6)	42(11.1)	0.569
합계	770(100.0)	480(100.0)	

자료 : 이명식(1999)

5. 모형화 및 점수할당(Modeling & Point Assignment)

데이터마이닝 과정에서 가장 중요한 단계로서, 앞서 선행되었던 단계에서 설정된 주요한 변수를 사용하고, 각 특성변수의 속성들에 점수를 할당하기 위하여 다양한 모형-Neural Networks¹⁰⁾, CHAID¹¹⁾, CART¹²⁾, 일반선형모형 등의 전통적인 통계적 모형 등을 적합해 보는 단계이다. 이러한 통계적 기법을 이용하여 신용평점표가 개발된다.

첫째, 판별 분석법(Fisher's Discriminant Analysis)이다. 이 기법은 독립변수들이 다중 정규분포로 이루어져야 되고 두 집단간 동일한 분산 및 공분산 구조가 존재해야 한다는 가정하에 선형분석을 할 경우 모델이 직접적인 계수추정이 가능하다. 단 분포에 대한 가정이 현실적으로 맞지 않고 개발된 모델에 의해 계산된 평점이 신용위험도와 직접적인 관계가 없다는 단점을 지니고 있다.

둘째, Zero-One회귀분석법이다. 판별분석법과 마찬가지로 독립변수들이 다중 정규분포로 이루어져야 되고 두 집단간 동일한 분산 및 공분산 구조가 존재해야

10) 신경망모형은 신경생리학분야에서 두뇌의 활동을 이해하고자 하는 목적하에 신경의 작업을 설명하려는 시도에서 출발하여 생물학적인 프로세스를 컴퓨터를 이용하여 모형화하려는 노력에서 비롯된 것으로 80년대 이후 생물학적 활동의 모형발전과 더불어 컴퓨터 성능의 진보 신경망 이론에 대한 통계학적인 접목으로 인해 빠르게 진보되면서 최근에는 데이터마이닝에서 있어서 유용한 기법이 되고 있다.

11) Chi-Square Automatic Interaction Deletion : 여러 변수중 비율을 가장 많이 깨뜨리는 변수가 결국은 응답여부에 가장 많이 미치는 변수

12) Classification and Regression Tree : 분할표를 바탕으로 Q값을 계산한다면 CART에서는 Gini Index라는 지수를 이용한다. Gini Index는 마디(Node)의 순수함의 정도를 재는 척도인데 어느 한 쪽의 카타고리만으로 구성되어 있는 마디일수록 순수하다고 판단한다.

한다는 가정하에서 선형분석일 경우 모델의 직접적인 계수추정이 가능하고 계산 및 기법 적용과정이 편리하다. 뿐만 아니라 개발된 모델에 의한 추정치가 우량 또는 불량일 확률로 나타나서 이해하기가 쉽다. 단점은 개발된 모델에 의한 추정치가 (0,1)을 벗어나는 경우가 많아서 모델의 모순이 발생하기 쉽다.

셋째, 로지스틱회귀분석법(Logistic Regression Analysis)이다. 이 분석법은 개발된 모델에 의해 계산된 추정치가 우량과 불량일 비율로 나타나 이해하기가 쉽고, 판별분석이나 Zero-One회귀분석에 비하여 모집단의 분석에 대한 가정에 구애받지 않기 때문에 판별력이 높다. 그러나 모델개발시 사용하는 식(Maximum Likelihood Analysis)을 사용하기 때문에 비선형 과정을 거치므로 계산상 난해하다는 단점이 있다.

넷째, Entropy분석법이다. 이 기법의 효과적인 적용을 위한 전제조건은 모집단이 정규분포가 되어야 하고 기법자체가 간단하여 연산하기가 쉽다. 그러나 특히 설명변수들의 임의적인 구간 설정에 있어서 개발자의 주관이 개입되어야 한다는 점과 설명변수 구간설정 필요성 때문에 변수의 정보가 상당히 손실될 가능성이 높다.

마지막으로 Non-Parametric기법이다. 이는 모집단의 실제분포와 관계없이 간단한 결과를 산출할 수 있고 모델의 Update가 용이하다. 그러나 개발된 모델이 제공하는 설명변수가 상식과는 맞지 않는 경우가 많고 위에서 서술한 통계적 방법에 비하여 나은 판별력을 갖지 못한다.

신용평점모델 개발에서는 일반적으로 판별분석(Fisher's Discriminant Analysis)과 로지스틱 회귀분석(Logistic Regression)을 많이 사용한다. 본 논문에서는 로지스틱 회귀분석을 이용하여 신용평점모델을 개발할 것이다.

6. 거절점수 설정

개발된 신용평점표를 이용하여 심사를 하게 되면, 심사결과가 거절, 승인 그리고 재심사 대상 등으로 구분된다. 이러한 결정은 거절점수(Cut-off Score)에 의하여 이루어진다. 거절점수는 신용공여기관에서 기꺼이 감수하고자 하는 위험의 범

위(손실금)를 나타내는 지표이다.

다음은 거절점수를 결정하는 방법을 보여준다.

(1) 수익성모형에 의한 거절점수 결정

신용평점모형은 위험을 예측하는 것이며, 평점의 기준설정에 있어 이익이 가장 중요한 요소로 고려되어야 한다는 것이다.

수익성 모델을 이용한다는 것은 예를 들어, 우량고객 건 당 이익이 \$100, 불량고객 당 손실 \$500 이라면, 이 경우 손익분기점을 이루기 위해서 불량 한건 승인하는 위험을 상쇄하기 위하여 우량고객 5명이 필요하다. 즉, \$100의 이익을 내기 위하여 우량고객 6명을 유치해야 한다. 이러한 방법(수익/비용 모델)으로 자산 전체의 목적에 부합되는 거절점수를 결정한다면 최고의 방법이 될 것이다.

(2) 복수거절점수전략(Tiered Cut-off Strategies)

앞장에서는 하나의 거절점수로 승인과 거절을 결정했지만, 이의 대응으로 복수 거절점수를 이용할 수도 있다. 의사결정표를 이용하여 수집된 정보로 전체고객을 고,중,저로 분류할 수 있다. 대부분 승인된 범주의 점수대(A), 대부분 거절된 범주의 점수대(D)로 구분하여, A에 해당하는 신청자의 경우는 세심한 심사 없이 처리하고(자동승인), A와 D사이의 점수대에 있는 신청자의 경우 2차심사 또는 반송 조치, D는 자동거절 하도록 조치한다(자동거절). 이 경우에는 거절점수가 2개 존재하게 된다. 즉, 승인점수와 거절점수. 이 두 점수 사이에 존재하게 되는 점수대를 회색지대(Override, Gray Zone)라 하며, 이는 심사대상자료 심사역이 직접 심사하게 된다. 실질적으로 신용평점에서 중요한 부분이라 할 수 있는 것은 바로 회색지대의 범주 결정 및 운영방법이다. 이러한 범주는 각종 비용과 또는 손실과 직결되는 부분이므로 신중히 결정하여야 한다. 또한, 거절점수는 경기상승, 경기침체 등에 따라 유동적이며, 쉽게 조절가능하다.

7. 평점표 검증(Validation)

적합한 두개 이상의 모형 효과를 비교하여 가장 좋은 모형을 선택할 수 있다. 이렇게 선택된 모형은 실제 모집단에 반영하여 그 효과를 재평가할 수 있다. 이렇게 재평가하는 방법을 검증이라 하며, 이의 목적은 모델링 측정에 있다. 가장 흔히 이용하는 방법으로는 K-fold Cross Validation이 있으며, 이는 모델링 개발에 이용한 자료를 K개로 나누어 다시 검증하는 방법이다. 예를 들어, 1만개의 표본자료가 있고, K값이 3이라면, 1차적으로 1만개의 자료를 3개(가,나,다)나눈 후, 2개는 모델 개발에 1개는 검증에 이용하는 방법이다. 이렇게 분류된 3개중 적중률이 가장 높은 모델을 결정하여 이용하게 된다. 또 다른 방법으로는, 모형 개발에 이용된 표본추출시점 이후의 자료로 신용평점표의 예측력을 검증하는 방법인 기간검증방법(Historical Validation)이 있다.

실증적 분석에서의 검증방법은 K-fold Cross Validation을 이용할 것이다.

제 4장 신용평점모형의 실증분석

제 1절 모형개발 개요와 개발절차

1. 모형개발 개요

현재 신용카드 사업은 국내외환위기 이후 우량고객의 유치보다는 불량고객의 사전방지에 사업의 목적을 두고 있기 때문에 신용도 정의시 부실율 및 회수비용의 감소에 목적을 둔 우량 및 불량을 정의하고자 하였다. 따라서 불량률의 기준은 짧은 연체기간 및 적은 연체금액을 기준으로 설정하였다.

(1) 신용도(Creditworthiness)정의(우량과 불량 정의)

우량정의(Good definition)는 1년간 신용카드 이용실적이 100만원 이상인 회원 중 검증시점(Validation Date)에서 연체자가 아닌 회원이고, 불량정의(Bad definition)는 검증시점에서 연체기간이 2개월 이상이고, 연체금액이 30만원 이상인 회원으로 한다. 불량기준은 현재 카드업계의 일반적인 통계자료를 근거로 하여 설정하였다. 통상적으로 현금서비스 평균이용금액이 30만원이하이므로 30만원한도까지는 불량기준에서 삭제하였고, 2개월 이상 연체자의 경우 대부분이 장기연체로 이어지는 경향이 뚜렷하였으므로 연체기간을 2개월로 결정하였다.

(2)제외대상

이 모형에서는 휴면회원(신규후 1년간 이용실적이 전무한 회원), 우수회원(VIP), 연령제한(18세이하), 외부정보(연합회정보)불량자를 제외대상으로 한다.

(3)관찰/예측 기간

<표 4-1> 관찰예측기간표

관찰기간	관찰시점	예측기간	총개월수
02.8.1-02.10.31 (3개월)	02.11.30	02.12.1-03.2.28 (3개월)	6개월

신용카드의 행동평점 예측기간을 단기간으로 이용한 이유는 가까운 미래를 과거의 이용실적 정보로 예측함으로써 조기에 부실을 방지하고자 하기 때문에 3개월을 예측기간을 사용하였다. 그러나 신용카드는 회원별로 대금결제일이 상이한 관계로 정확하게 3개월을 적용하기에는 불가능하여 결제일 분포도를 참고하여 제일 많은 분포를 보인 매월 27일자 결제회원을 기준으로 관찰시점을 월말로 하였다.

(4) 표본추출

신용카드 회원중 우량회원 2,250명, 불량회원 750명, 총 3천명을 01년 11월 30일 시점에서 단순 임의 추출(Simple Random Sampling)으로 표본을 추출하였는데, 불량률이 우량률에 비해 상당히 작고, 불량자의 특성을 잘 부각시키기 위해 샘플의 불량률을 30%이상 되도록 추출하였다. 02년 11월 현재 우량이 83.20%이고 불량률이 16.80%이며, 03년 2월 현재 우량이 85.93%이고 불량률이 14.07%로 구성되어 있다.

<표 4-2>기준일과 예측일 현재 우량과 불량 구성비

구분		2003년 2월		
		우량	불량	계
2002년 11월	우량	2348	148	2496
	불량	230	274	504
	계	2578	422	3000

(5) 변수선정

<표 4-3> 변수의 유의성

변수(변수명)	유의성(P값)	변수선정
6개월간 연체일수(DATN)	0.0001 ^{*13)}	0
나이(AGE)	0.5697	
연체금액(DLYAMNT)	0.0001 [*]	0
자동차관련금액(KM6)	0.0105	
최근 2개월치 할부구매금액(HKUM)	0.0001 [*]	0
최근 2개월치 일시불구매금액(IKUM)	0.0001 [*]	0
6개월 사용액중 현금서비스사용비율(RSCKUM)	0.0001 [*]	0
6개월 사용액중 할부금액 사용비율(RSHKUM)	0.0001 [*]	0
현재잔액(KUJAN)	0.0001 [*]	0
요구불예금 평균잔액(YMKUPJ)	0.0508	
카드경과기간(DIF)	0.0001 [*]	0
보유카드중 금융권 카드비율(TCARD)	0.0001 [*]	0
평균신용거래간격(SINGAP)	0.0001 [*]	0

다음은 신용카드 관련 용어로 모형개발관련상 이해를 돕기 위하여 참고적으로 설명하였다. 이용실적은 기준유형별로 승인금액, 결제금액, 청구금액, 매출표 처리 일 금액(가맹점에 입금하는 시점)기준이 있다. 본 논문에서의 이용실적은 비용발생 시점주의를 적용하여 표기한 것으로, 신용카드 회원의 이용이 발생하고 은행이 가맹점에 회원의 이용금액을 입금하여주는 시점의 금액을 기준으로 활용하였다. 단, 할부는 청구금액 기준을 이용하였다.

탈회회원에는 카드탈회와 회원탈회 2가지 유형이 있다. 카드탈회는 회원의 이용실적 및 신용도에 따라 등급을 상향하여 발급함으로써 이전에 발급된 카드가 자동탈회된 경우로 이용실적은 있으나 연체금액이 없는 탈회, 이용실적도 있고 연체금액도 있는 탈회 그리고 연체금액만 있는 탈회 등은 모두가 카드탈회라고 볼 수 있다. 반면, 이용실적도 연체금액도 없는 탈회는 회원탈회라고 볼 수 있다. 표본자료 추출시 이러한 유형의 탈회는 모두 제외되었다.

(6) 사전심사 기준

본 논문은 신청평점모형이 아니라 행동평점 모형개발이므로 사전심사 기준을 고려할 필요가 없었다. 단, 행동평점을 측정할 신용카드회원대상을 일정한 기준으로 설정한 수는 있지만 이러한 부문은 이용가능 및 접근이 불가능한 관계로 이를 제외하였다.

2. 모형개발 절차

(1)모형개발을 위한 표본을 추출하였다.

(2)1차변수분석을 이용하여 선정된 변수중 유의한 변수를 선정한다.

(3)Data Set을 3개의 Sub-Data Set으로 분류한다.

(4)2개의 Sub-Data Set으로 모수값을 추정하고, 나머지 Sub-Data Set은 모델링의 적합도를 테스트하는데 사용한다.

(5)모수값을 추정하는데 사용되어지는 Data Set에서의 적중률과 테스트 Set의

13) *는 95% 신뢰구간에 속하는 값

적중률이 동시에 높은 것을 선택한다.

(6)구해진 모형으로 각 변수에 대한 신용평점표를 작성한다.

제 2 절 모형의 추정결과

1. 선정 변수의 유의성 분석

선정된 변수의 유의성을 검증하기 위하여 T-Test 와 Chi-square의 검정을 이용하였으며, P값의 유의성은 $P < 0.005$ 로 하였다. 고려한 변수의 유의성을 검증한 결과 6개월간 연체일수는 우량과 불량을 명확하게 구분하여 주고 있으며, 이용금액의 경우는 우량이 불량보다 평균금액이 높음을 알 수 있다. 또한 신용카드 발급기간이 길수록 우량의 가능성이 높음을 보여주고 있다. 이는 신용카드 이용이 생활의 일부분이 된 우수회원이라 파악된다. 일반적으로 신용카드 회원중 이용금액이 높은 우수고객의 신규후 카드 경과기간이 최소한 2년 이상인 것으로 파악되고 있다. 나이와 요구불예금의 평균잔액의 경우 P값이 0.05보다 크게 나와 유의성이 없는 것으로 판단되나, 사실상 중요한 변수라 고려되어 채택하였다.

<표 4-4> 변수별 통계값

변수명	구분	평균	표준편차	T값	P값
6개월간연체일수	우량	44.34	48.80	-105.66	0.0001
	불량	124.79	43.85		
나이	우량	36.23	10.24	0.56	0.5697
	불량	36.13	10.20		
연체금액	우량	51,523.20	280,718.40	-66.97	0.0001
	불량	1,461,222.56	1,350,088.14		
자동차관련금액	우량	155,530.07	518,118.41	2.55	0.0105
	불량	135,477.29	447,774.93		
최근2개월치 할부 금액	우량	151,921.82	372,868.00	34.86	0.0001
	불량	27,080.97	158,227.91		
최근 2개월치 일시 불구매	우량	213,456.86	573,579.74	36.74	0.0001
	불량	31,631.09	187,028.42		
6개월사용액중현금 서비스비율	우량	0.53	0.36	-7.68	0.0001
	불량	0.57	0.31		
6개월사용액중할부 구매사용비율	우량	0.19	0.25	-4.67	0.0001
	불량	0.20	0.23		
현재잔액	우량	3,645,789.24	21,982,313.22	17.25	0.0001
	불량	751,177.41	4,350,516.47		
요구불잔액	우량	1,625,307.45	115,427,158.52	1.95	0.0508
	불량	49,547.92	1,011,211.52		
카드경과기간	우량	27.61	27.32	22.72	0.0001
	불량	19.55	19.24		
보유카드중 금융권 카드비율	우량	0.82	0.22	-11.79	0.0001
	불량	0.86	0.20		
평균신용거래간격	우량	461.20	554.66	3.81	0.0001
	불량	425.67	539.96		

2. 통계적 분석 결과

모집단 즉 3,000개의 표본을 세 집단으로 나누어 세번 테스트한다. 차례대로 2집단(표본수 2,000개)은 테스트에 이용을 하고, 1집단(표본수 1000개)은 이 테스트의 결과를 검증해 보는 데 이용을 하기로 한다.

즉 <표 4-5>와 같이 구성된다.

<표4-5>모형개발집합(Training Set)과 모형검증집합(Validation Set)의 구성도

구분	Sub-set 1	Sub-set 2	Sub-set 3
I	Training	Training	Validation
II	Training	Validation	Training
III	Validation	Training	Training

<표 4-6>모형에서의 적중률 비교

구분		Sensitivity ¹⁴⁾	Specificity ¹⁵⁾	Hit-Ratio ¹⁶⁾
I	Training	81.0	78.1	80.6
	Validation	78.6	80.6	78.9
II	Training	81.2	77.1	80.7
	Validation	79.3	81.3	79.6
III	Training	80.8	78.6	80.5
	Validation	79.1	78.1	78.9

표본집단을 세개로 나누어 세번에 걸친 테스트에 걸친 테스트 결과, 실제우량을 우량으로 정확하게 예측한 경우와, 실제불량을 불량으로 예측한 경우에 나타나는 모형개발집합의 값에서 결정된 변수의 값으로 모델검증집합에 적용한 경우의 예측값을 비교하여 이 값의 차이가 가장 작은 세번째 집단을 통계적 분석에 이용하고자 한다. 즉 <표 4-7>에서 보는 바와 같다.

14) 실제 우량을 모형결과가 우량으로 판정할 경우

15) 실제 불량을 모형결과가 불량으로 판정할 경우

16) Hit-ratio는 Sensitivity와 Specificity의 평균

3. 평점표

<표 4-7> 변수별 평점표

변수명	모수추정치	평점표 ¹⁷⁾	변수명	모수추정치	평점표
INTERCEPT ¹⁸⁾	5.1242	637.7826	AGE	-0.0081	-0.2532
KM6	8.4430	2.6103	IKUM	-7.6303	-2.3590
HKUM	-2.7873	-9.6176	DATN	-0.0139	-0.4310
DIF	0.0043	0.1356	PLYAMNT	-9.9916	0.0000
KUJAN	2.8095	8.6862	YMKUPJ	-2.8332	-8.7596
RSCKUM	-0.9789	-30.2652	RSHKUM	-1.4508	-44.8562
SINGAP	0.0001	0.0041	TCARD	-0.6432	-19.8859

신용평점표는 다음의 로지스틱회귀함수식을 이용하여 값을 구하였다.

개별 소비자의 채무불이행 여부를 Y라고 하고, 선택된 여러 가지 인구동태학적 변수를 $X_i (i = 1, 2, \dots, m)$ 이라고 하면, 채무불이행 확률은 로지스틱 모델에 의해 다음과 같이 구해진다.

$$Y_j = 1 (\text{소비자 } j \text{ Not default})$$

$$0 (\text{소비자 } j \text{ Default}) \quad \text{단, } j = 1, 2, \dots, n$$

β_i = 선택된 변수 X_i 의 coefficient(score weight)

소비자 j의 채무불이행 확률 = prob.($Y_j = 0$)

$$= \frac{1}{1 + e^{\beta X_j}}$$

단, $\beta = (\beta_1, \beta_2, \beta_3, \dots, \beta_n)$, $X_j = (X_{1j}, X_{2j}, X_{3j}, \dots, X_{mj})$

소비자 j= 1, 2, ..., n

위와 같은 식에 표본에서 추출된 소비자들의 실제 데이터 값을 대입시켜 각각의 계수(coefficient)를 구하고, 구해진 각 변수의 계수(coefficient)값에 의거하여

17) 평점표 = 1,000/변수의 범위(최대값-최소값)×모수추정치

18) 절편값

새로운 고객이 가입시 그 고객의 채무불이행 확률을 추정하는 것이다.

로지스틱회귀함수식:

$$Y = 5.1242 + (AGE \times -0.0081) + (KM6 \times 8.4430) + (IKUM \times -7.6303) + (HKUM \times -2.7873) + (DATN \times -0.0139) + (DIF \times 0.0043) + (DLYAMNT \times -9.9916) + (KUJAN \times 2.8095) + (YMKUPJ \times -2.8332) + (RSCKUM \times -0.9789) + (RSHKUM \times -1.4508) + (SINGAP \times 0.0001) + (TCARD \times -0.6432)$$

평점표의 점수분포는 1,000점을 기준으로 하여 높은 점수는 우량을 낮은 점수는 불량량을 나타내도록 하였다.

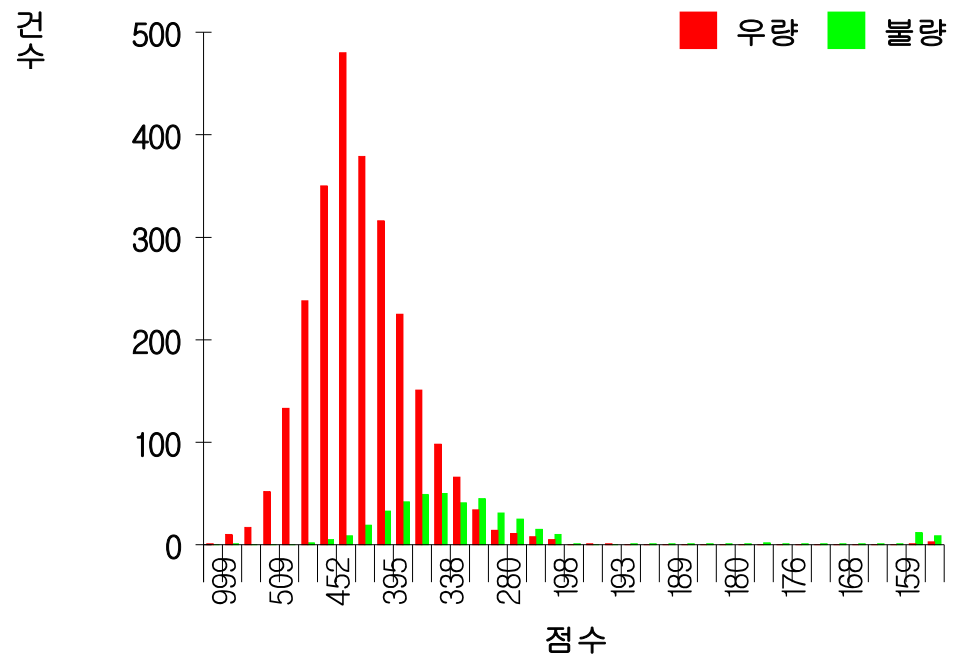
4. 행동평점 분포도¹⁹⁾

행동평점 분포도를 살펴보면, 행동평점이 우량과 불량량을 구분시킬 수 있다는 것을 알 수 있다. 이는 행동평점모형을 이용할 경우, 불량으로 정의한 연체 2개월 이상 될 가능성이 높은 잠재 부실회원을 사전에 알려줄 수 있는 예측력이 상당히 높다는 것을 보여준다.

따라서, 행동평점모형을 이용할 경우 사전부실방지를 효과적으로 막을 수 있으며, 우량고객에게는 보다 나은 서비스를 제공할 수 있다.

19) 우량과 불량량의 행동평점 분포도

<그림 4-1> 행동평점 분포도



5. 행동평점 모형 성과표

<표 4-8> 행동평점 모형 성과표

접수분포			누계				전체	빈도		(우량/ 불량)
접수	우량	불량	우량	백분율	불량	백분율	불량율	누계	백분율	
999	1	-	1	0.02%	-	0.00%	0.00%	1	0.0%	
563	10	1	10	0.40%	1	0.13%	0.02%	11	0.4%	10
528	17	-	27	1.05%	1	0.13%	0.02%	28	0.9%	
509	52	-	79	3.03%	1	0.13%	0.02%	80	2.7%	
490	133	-	212	8.13%	1	0.13%	0.02%	213	7.1%	
471	238	2	450	17.25%	2	0.53%	0.07%	453	15.1%	119
452	350	5	800	30.66%	7	1.74%	0.24%	807	26.9%	70
433	480	9	1,281	49.06%	16	3.88%	0.53%	1,296	43.2%	53
414	379	19	1,659	63.57%	34	8.42%	1.15%	1,694	56.5%	20
395	316	33	1,975	75.66%	67	16.44%	2.24%	2,042	68.1%	10
376	225	42	2,200	84.27%	109	26.74%	3.64%	2,309	77.0%	5
357	151	49	2,350	90.05%	158	38.77%	5.28%	2,509	83.6%	3
338	98	50	2,449	93.81%	208	50.94%	6.93%	2,657	88.6%	2
319	66	41	2,514	96.32%	249	61.10%	8.32%	2,764	92.1%	2
300	34	45	2,549	97.64%	295	72.19%	9.83%	2,843	94.8%	1
280	14	31	2,563	98.18%	326	79.81%	10.86%	2,889	96.3%	0
261	11	25	2,574	98.60%	351	85.96%	11.70%	2,925	97.5%	1
238	8	15	2,582	98.91%	366	89.71%	12.21%	2,948	98.3%	1
198	5	10	2,587	99.10%	376	92.11%	12.54%	2,963	98.8%	1
197	-	1	2,587	99.10%	377	92.25%	12.55%	2,963	98.8%	-
196	1	-	2,587	99.12%	377	92.25%	12.55%	2,964	98.8%	-
193	1	-	2,588	99.14%	377	92.25%	12.55%	2,965	98.8%	-
191	-	1	2,588	99.14%	378	92.51%	12.59%	2,966	98.9%	-
190	-	1	2,588	99.14%	378	92.65%	12.61%	2,966	98.9%	-
189	-	1	2,588	99.14%	379	92.78%	12.63%	2,967	98.9%	-
188	-	1	2,588	99.14%	379	92.91%	12.65%	2,967	98.9%	-
186	-	1	2,588	99.14%	380	93.05%	12.66%	2,968	98.9%	-
180	-	1	2,588	99.14%	380	93.18%	12.68%	2,968	98.9%	-
178	-	1	2,588	99.14%	381	93.32%	12.70%	2,969	99.0%	-
177	-	2	2,588	99.14%	383	93.85%	12.77%	2,971	99.0%	-
176	-	1	2,588	99.14%	384	93.98%	12.79%	2,972	99.1%	-
173	-	1	2,588	99.14%	384	94.12%	12.81%	2,972	99.1%	-
172	-	1	2,588	99.14%	385	94.25%	12.83%	2,973	99.1%	-
168	-	1	2,588	99.14%	386	94.52%	12.86%	2,974	99.1%	-
165	-	1	2,588	99.14%	386	94.65%	12.88%	2,974	99.1%	-
161	-	1	2,588	99.14%	387	94.79%	12.90%	2,975	99.2%	-
159	-	1	2,588	99.14%	388	94.92%	12.92%	2,975	99.2%	-
156	1	12	2,589	99.18%	400	97.86%	13.32%	2,989	99.6%	0
151	3	9	2,592	99.29%	408	100.0%	13.61%	3,000	100.0%	0

행동평점 모형표에서 거절점수를 395점으로 할 경우의 Odds²⁰⁾(10)는 전체표본 추출시 고려한 우량 대 불량 비율의 Odds(6.3)보다 높다. 그러나 행동평점을 이용하여 395점에서의 누적불량율은 2.24%로 전체의 13.61%보다 무려 6배정도의 성과가 있다고 볼 수 있다. 행동평점 성과표 적용방법을 신청평점모형에 적용하여 분석할 경우, 전통적 판단이 가지는 불량률 보다 신청평점표를 함께 이용하여 혼용할 경우 업무성과가 6배 높다고 할 수 있는 것이다. 만일 거절점수 395점으로 인하여 상실할 수 있는 우량고객이 약 25%이며, 이를 재심사하고자 할 경우 일정한 점수의 범주를 재심사대상으로 정할 수 있으며, 이를 회색지대라 한다.

또 다른 방법으로는 위험을 수용할 정도의 거절점수를 정해두고 거절점수보다 어느 정도 거절심사대상과 승인심사대상을 별도의 관점에서 재심사하는 방법이 있다. 거절심사대상에서는 불량인데 우량으로 파악된 건수만 평가하여 거절하며, 승인심사대상은 우량인데 불량으로 평가된 건수를 파악하여 승인하여 주는 방법이다. 이러한 운영적 전략은 적용분야에 따라 다양하게 적용될 수 있으며, 운영방법의 차이에 따라 업무성과가 크게 다를 수 있기도 하다.

20) Odds(우량/불량) : 1개의 불량에 우량이 차지하고 있는 개수. 점수가 높으면 Odds는 높다.

즉, 높은 점수는 위험도가 낮다는 의미이므로 불량률의 수가 적으며, 점수가 떨어질수록 Odds도 하향하는 것이 위험을 올바르게 평가한다고 볼 수 있다.

결론

신용카드사에게 있어 여타 리스크보다 신용리스크가 중요한 이유는 카드사의 입장에서 볼 때, 기본적으로 관리가 가능하다는 점과 관리에 실패했을 경우 여타 통제 불가능한 리스크를 촉발할 수 있기 때문이다. 따라서 카드사의 입장에서 본다면 현재 국내 카드 시장 위기는 근본적으로 개별 카드사들의 신용 리스크 관리 부실이 기폭제가 되었다고 할 수 있다. 즉, 국내 카드사들의 신용리스크 증가가 2002년부터 이들의 대손비용 및 기타 전대환론 증가 등으로 현실화되고 공론화됨으로써 이들의 수익성을 급격히 악화시켰고, 다시 카드사들이 발행한 채권 금리의 상승 및 이의 환매 요구, 부실자산 유동화 실패 등으로 여타 통제 불능의 리스크 즉, 유동성 및 이자율 리스크를 증가시킨 것이다. 결국 이러한 연결고리는 카드사들의 존망을 위협하고 있다.

카드사들이 최근에 경험하고 있는 동일한 신용대란을 되풀이 하지 않기 위해서는 적기에 적절한 방법으로 신용 리스크를 관리할 수 있는 역량을 조속히 강화해야 한다. 신용위험관리는 신용위험의 측정을 토대로 이루어지게 되는데, 사실상 현재까지 표준화된 기법이 확립되어 있는 것은 아니다. 선진금융시장에서 개발된 신용위험 평가방식이라 하더라도 국내사정에 적합하지 않으면 그만큼 우리에게 주는 효용은 감소하게 마련이며, 국내기업 및 개인의 부실화 과정이 선진금융시장의 경우와 다를 것이며 또 이들 기법들을 적용하는데 있어 하부기반으로 필요한 데이터베이스가 충분히 축적되어 있지 않기 때문에 적용이 불가능한 경우도 많다.

이러한 점에서 본 논문에서는 신용위험관리기법 중의 하나인 개인신용평점에 관한 실증적 분석을 통하여 이의 유효성을 검증하였다. 개발모형은 신용카드 고객의 행동평점모형이었으며, 로지스틱스 회귀분석을 이용하였다. 모형의 유효성 검증을 위하여 위험을 표시하는 신용평점순으로 분류한 후 이미 결정한 우량과 불량을 어느 정도 구분할 수 있는지 비교하였다.

개인신용평점을 도입하기전의 불량률이 13.61%였으나, 이를 도입하고 적용한 이후의 불량률은 2.24%로 감소할 수 있음을 알 수 있었다. 따라서 기존의 전통적

방법에 개인신용평점모형을 도입하여 함께 이용하는 것이 더욱 효율적이었음이 판명되었다. 신용위험관리기법의 개발 필요성은 어느 때보다 더 절실하다 하겠다. 또한 신용평점표를 이용한 심사기준은 전통적인 방법보다 비용과 시간을 단축시켜 줄 수 있다. 이러한 신용평점표와 단계별로 계획을 수립하여 성공적으로 신용관리 역량을 강화한다면 나날이 경쟁이 가속화되는 국내의 카드시장에서 경쟁우위를 점할 수 있을 것이다.

끝으로 본 연구의 한계점은

첫째, 모형개발에 이용된 개인 자료의 속성정보 및 거래정보가 부족하여 고객의 속성을 정확히 파악하지 못했다는 것이다. 다양한 개인정보를 이용한 통계적 방법에 의하여 최적의 요소들을 추출하여야 한다. 이를 위하여 개인신용 구성요소들에 대한 데이터웨어하우징 구축이 기업내부 뿐만 아니라 외부정보 제공기관에서도 필요하다.

둘째, 신청 자료의 정확성이다. 현재 데이터는 연령, 직업, 자가용 유무, 결혼유무 등의 구성요소는 비교적 정확한 자료를 가지고 있으나, 소득, 주거상황, 가족수, 근무시간 등은 자료가 불충분하거나, 전혀 없는 경우는 작성자의 임의로 기재하는 수가 많다.

셋째, 2차변수의 생성을 고려하지 않았다. 즉, 1차변수로는 설명력이 떨어지지만 2차변수의 생성을 통한 상관관계를 고려할 경우 보다 더 예측력이 높은 변수를 도출할 수 있기 때문이다. 2차변수란 남성이면서 나이가 20대 혹은 여성이면서 연체일수가 많은 경우 등 개인의 속성정보 조합을 통하여 분석하는 것이다.

참 고 문 헌

-국내 문헌

강돈혁. 1997. 전문가 시스템을 이용한 은행의 개인신용평가모형 개발에 관한 연구. 석사학위논문. 강릉대학교 경영정책과학대학원.

김유진. 2001. 신경망 모형과 그 응용. 석사학위논문. 고려대학교 대학원.

김은숙. 1997. 개인신용평가표의 유용성에 관한 연구 : J은행 가계당좌예금을 중심으로. 석사학위논문. 영남대학교 경영대학원.

김창효. 2000. “고객관리를 위한 새로운 스코어링 기법에 관한 고찰”. 한국통계학회 추계학술발표회 논문집, pp.232-233.

김현규. 2002. 증거비중을 이용한 신용평점모델에 관한 연구. 석사학위논문. 중앙대학교 대학원.

박진우. 1999. “신용리스크 측정 모형과 자본 및 자산의 배분”. 금융, pp.19-25.

신지웅. 1998. “신용카드시스템의 신용평가기법적용에 관한 연구”. 비씨카드 학술 논문집, pp.407-428.

손정현. 1998. 우리나라 신용평점모형의 개선책에 관한 연구. 석사학위논문. 연세대학교 경영대학원.

이명식. 1999. “금융마케팅에서 고객평점제도의 효과성 : 신용 및 수익성을 중심으로

로”. 한국마케팅저널, pp.56-76.

이명식. 1995. “신용평점제(Credit Scoring System)의 의의와 역할(V)”. 신용카드, pp.30-41.

이명식. 1998. “평점제를 이용한 소비자 신용위험 모형개발”. 경영학연구, 6: 1-21.
상명대학교 경영연구소.

이원경. 1998. “신용평점시스템을 위한 통계기법”. 조흥경제연구소.

이창우. 1987. 개인 신용평가 모델을 이용한 의사결정 지원시스템 개발에 대한 연구. 석사학위논문. 한국외국어대학교 경영정보대학원.

정영식. 1998. 신용도 예측 행동 모형 개발을 통한 효과적 여신관리. 석사학위논문.
고려대학교 경영대학원.

정영옥. 1997. “개인에 대한 신용평가 어떻게 해야 하나”. 신용정보, 9: 14-21.

조성환. 2001. 개인신용평가법을 이용한 신용대출 확대방안 : 신협을 중심으로. 석사학위논문. 한남대학교 지역개발대학원.

천인국. 2001. 신용카드 발급에 있어 신용평점모형에 관한 연구. 석사학위논문. 한양대학교 대학원.

최강민. 1998. 개인신용평가모델을 이용한 신용평가 시스템 구축. 석사학위논문. 경북대학교 경영대학원.

최종후, 권기만, 김수택. 2002. “신용평점모형”. 세창출판사.

하성목. 1999. “통합리스크 관리체제로의 발전 및 리스크 현황의 정확한 측정 선행
돼야”. 금융경제, 4: 18-21.

-해 외 문 헌

Allen Jost. 1988. "Data Mining". Credit Risk Modeling. pp. 129-154.

Balvinder S. Sangha. 1998. "A Systematic Approach For Managing Credit
Score Overrides". Credit Risk Modeling. pp. 221-244.

Beverly Hart. 1995. "Successful Collections : Solving The Customer's Credit
Problem". Bankcard Business School Visa Int'L.

Christian Bluhm, Ludger Overbeck, Christoph Wagner. 2002. "An Introduction
to Credit Risk Modeling". CRC press.

Dennis C. Glennon. 1998. "Issues In Model Design And Validation". Credit Risk
Modeling. pp. 207-220.

Edward, M. Lewis. 1994. "An Introduction to Credit Scoring". Fair, Isaac and
Company. 이희만, 박관수 옮김. 1999. “신용스코어링 소개”. 도서출판 피디.

Elizabeth Mays. 2001. "Handbook of Credit Scoring". Glenlake Publishing
Company, Ltd.

Gradstom John. 1996. "Seven Characteristics Of An Effective Credit Risk

Management System And How To Test Them". The Journal Of Lending & Credit Risk Management. pp. 55-60

Hollis Fishelson-Holstine. 1998. "Case Studies In Credit Risk Model Development". Credit Risk Modeling. pp. 169-181.

John M. L. Gruenstein. 1998. "Optimal Use Of Statistical Techniques In Model Building". Credit Risk Modeling. pp. 81-112.

Larry Cordell, Matthew Klena, Timothy J. Malamphy & Wan-Qi Ting. 1998. "Early indication : Design And Validation Of A Comprehensive Behavior Scoring System For Delinquent Mortgages". Credit Risk Modeling. pp. 113-128.

Leonard J. McCahill. 1998. "Organizational Issues In Building And Maintaining Credit Risk Models". Credit Risk Modeling. pp. 13-22.

Lyn C. Thomas, David B. Edelman, Jonathan N. Crook. 2002. "Credit Scoring and Its Applications". Society for Industrial and Applied Mathematics.

Noel Capon. 1982. "Credit Scoring System : A Critical Analysis". Journal Of Marketing. pp. 82-91.

William M. Makuch. 1998. "Scoring Applications". Credit Risk Modeling. pp. 1-12.

William M. Makuch. 1998. "The Basics Of A Better Application Score". Credit Risk Modeling. pp. 59-80.

ABSTRACT

Empirical Research on the Personal Credit Risk Management

Kim, Hak Shin

Dept. of Economics

The Graduate School

Yonsei University

The Korean financial sector has experienced a tremendous change after the financial crisis occurred in 1997. A lot of financial institutions were merged or out of the market. Each institution has adopted various advanced financial techniques to survive such a severe circumstance. One of the advanced financial techniques is so-called the Credit Scoring System(CSS).

The aim of this research is as follows. First, this paper analyzes the CSS that is one of the risk management techniques and tests how effectively the CSS works on enhancement of the financial institution's earnings. In addition, this paper tries to present the necessity of model-development and the direction of research.

This research tests the significance level of each property of data, using private loan data and makes selected Properties of data into the Credit Scoring Model with Logistic Regression Analysis. Logistic Regression Analysis used for a empirical analysis is as follows.

consumer j's default probability = $\text{prob.}(Y_j = 0)$

$$= \frac{1}{1 + e^{\beta X_j}}$$

but, $\beta = (\beta_1, \beta_2, \beta_3, \dots, \beta_n)$, $X_j = (X_{1j}, X_{2j}, X_{3j}, \dots, X_{mj})$

consumer $j = 1, 2, \dots, n$

And this paper tests the model with K-Folder Cross Validation and compares credit good & bad analyzed through the credit scoring model with the one through the existing model, so it will compare how effectively the model divides credit good & bad.

The conclusion through this research is that the bad rate was 13.16% before adopting the Private Credit Scoring, however, we know that the bad rate was dropped to 2.24% after introducing and adopting this system.

The conclusion from the result of this analysis is as follows. First, the model with the private credit scoring is better than the traditional model in the sense that the former increases the earnings of financial institutions, discriminating between 신용우량 and 불량 efficiently. Second, using the CSS is more efficient because it reduces cost and time, comparing with the traditional model. Finally, if strengthening the ability of credit- management by using this CSS, the predominance in financial market can be gained even under more severe competition.

Key words : Credit Scoring System, Logistic Regression Analsys, K-folder Cross Validation