

박사학위논문

데이터마이닝기법을 활용한 주택담보대출  
연체가능성 분석에 관한 연구

2015년

한성대학교 대학원  
경제부동산학과  
부동산학전공  
유재술

박사학위논문  
지도교수 안정근

# 데이터마이닝기법을 활용한 주택담보대출 연체가능성 분석에 관한 연구

The Default Possibility Analysis of Housing Mortgage Loans  
Using Data Mining Techniques

2014년 12월 일

한성대학교 대학원  
경제부동산학과  
부동산학전공  
유재술

박사학위논문  
지도교수 안정근

# 데이터마이닝기법을 활용한 주택담보대출 연체가능성 분석에 관한 연구

The Default Possibility Analysis of Housing Mortgage Loans  
Using Data Mining Techniques

위 논문을 부동산학 박사학위 논문으로 제출함

2014년 12월 일

한성대학교 대학원  
경제부동산학과  
부동산학전공  
유재술

유재술의 부동산학 박사학위논문을 인준함

2014년 12월 일

심사위원장 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

# 국 문 초 록

## 데이터마이닝기법을 활용한 주택담보대출 연체가능성 분석에 관한 연구

한성대학교 대학원  
경제·부동산 학과  
부 동 산 학 전 공  
유 재 술

우리나라 주택담보대출의 확대는 2008년 미국의 서브프라임 모기지사태가 발생하면서 부동산시장이 재조명되었으며 단기간에 크게 증가한 주택담보대출의 리스크관리에 대한 관심이 높아지고 있다. 또한 그동안 주택보급을 증가로 주택시장은 안정세를 유지하고 있지만 주택수요는 둔화된 현상을 보이고 있다. 최근 경기침체로 인해 저성장이 지속되고 있으며 인구증가율은 감소되고 생산인구 또한 빠르게 감소할 것으로 전망되고 있어, 고령사회 진입을 앞두고 있는 등 여러 환경이 변화하고 있다. 사회적, 경제적 환경변화와 더불어 가계대출의 대부분을 차지하고 있는 주택담보대출의 커다란 규모와 빠른 증가의 건전성에 대한 관리의 필요성이 높다고 판단되고 있다.

정부에서도 2014년 관계부처 합동으로 가계부채 구조개선 촉진방안을 제시하였고 부채부분 가계건전성 제고를 위한 대출 증가속도의 관리, 대출구조의 개선, 취약계층 채무조정 등 부분별 대책을 순차적으로 마련하고 있는 시점으로 금융기관에서는 가계대출 비중이 높은 주택담보대출 관리가 필요한 상황이다. 기존의 선행연구가 주로 거시경제변수와 연체율 관계를 분석한 연구로서 금융기관에서 대출심사 및 연체예측성 등 자료로 적용하기에는 부적합하였다. 이에 본 연구에서는 주택담보대출 자료를 가지고 연체가능성 여부를 예

측하기 위하여 회귀모형인 이항로짓모형을 적용하였고 데이터마이닝기법에는 인공신경망모형 및 의사결정나무모형을 적용하여 활용여부를 각 분석모형을 통하여 예측하였다.

연구결과 비연체차입자에 대한 예측은 로짓모형과 의사결정나무모형이, 연체차입자에 대한 예측은 인공신경망모형과 의사결정나무모형이 높게 나타났다. 상대적으로 편차를 보이긴 하였으나 전반적인 예측력은 의사결정나무모형이 가장 우수한 것으로 나타났다. 로짓모형과 인공신경망모형에서는 대출기간, 금리조정기간, 주택규모 등이 공통적으로 연체에 영향을 많이 미치는 것으로 나타났으며, 의사결정나무모형에서는 대출시기, 대출금액, 대출목적 등이 상위 분류기준으로 나타나 연체여부에 영향을 주는 것으로 나타났다. 각 모형의 특성에 따라 장점을 발견하였으며 각 분석결과를 모두 고려하여 판단할 필요가 있을 것으로 보인다. 로짓모형은 통계적인 활용도에서, 인공신경망모형은 통계적 유의성에 대한 제약 없이 다양한 변수의 영향력을 활용할 수 있다는 점에서 의의를 가질 수 있으며, 의사결정나무모형은 정상 차입자와 연체자의 단계별 분류기준을 확인 할 수 있다는 점에서 의미가 있다고 본다.

**【주요어】** 주택담보대출, 연체율, 데이터마이닝, 이항로짓모형, 인공신경망모형, 의사결정나무모형.

# 목 차

제 1 장 서 론 .....	1
제 1 절 연구의 배경 및 목적 .....	1
제 2 절 연구범위 .....	4
제 3 절 연구방법 .....	5
제 2 장 이론적 검토 .....	7
제 1 절 주택담보대출의 개념 및 유형 .....	7
1) 주택담보대출의 개념 .....	7
2) 주택담보대출의 유형 .....	8
제 2 절 주택담보대출의 위험관리 .....	14
1) 주택담보대출에 대한 위험관리 필요성 .....	14
2) 사전적 위험관리 .....	22
3) 사후적 위험관리 .....	29
제 3 절 국·내외의 가계부채 대응정책 .....	37
1) 우리나라의 가계부채 대응정책 .....	37
2) 해외 주요국가의 가계부채관리 .....	40
3) 해외 사례의 시사점 .....	52
제 4 절 데이터마이닝기법의 개념과 특징 .....	56
1) 데이터마이닝 모형의 개념 .....	56
2) 데이터마이닝 모형의 특징 .....	58
제 5 절 선행연구 .....	59
1) 주택담보대출에 관한 선행연구 .....	59
2) 주택담보대출 리스크요인에 관한 선행연구 .....	62

3) 주택담보대출 연체가능성에 관한 선행연구 .....	65
4) 선행연구와의 차별성 .....	67
<b>제 3 장 연구모형 및 분석자료 .....</b>	<b>70</b>
제 1 절 연체가능성 분석모형 .....	70
1) 로지스틱 회귀분석 모형 .....	71
2) 인공신경망 모형 .....	75
3) 의사결정나무 모형 .....	79
제 2 절 연체율 산정기준과 연체율 현황 .....	84
1) 산정기준 .....	85
2) 연체율 현황 .....	87
제 3 절 분석자료와 변수 .....	89
1) 분석자료 .....	89
2) 분석변수 .....	93
3) 분석자료의 특징 .....	95
제 4 절 연구대상의 속성 .....	97
1) 성별간 차이 .....	97
2) 용도간 차이 .....	98
3) 시기별 차이 .....	100
4) 연체 및 비연체가구간 특성 .....	102
<b>제 4 장 주택담보대출 연체가능성 분석 .....</b>	<b>105</b>
제 1 절 로지스틱 회귀분석 결과 .....	105
1) 연체가능성 변수 .....	105
2) 분석모형의 예측정확도 .....	108
제 2 절 인공신경망분석 결과 .....	109
1) 10일 이상 연체기준 .....	110

2) 20일 이상 연체기준 .....	111
3) 30일 이상 연체기준 .....	112
4) 분석모형의 예측정확도 .....	113
제 3 절 의사결정나무분석 결과 .....	116
1) 10일 이상 연체기준 .....	116
2) 20일 이상 연체기준 .....	122
3) 30일 이상 연체기준 .....	126
4) 분석모형의 예측정확도 .....	129
제 4 절 분석결과 정리 .....	131
제 5 장 결 론 .....	135
제 1 절 요약 및 시사점 .....	135
제 2 절 연구의 한계 및 향후 연구과제 .....	139
참고문헌 .....	140
ABSTRACT .....	148

## 표 목 차

〈표 1〉 주택담보 금융의 저당방식 .....	7
〈표 2〉 주택담보대출 변동금리에 적용되는 기준금리의 종류 .....	10
〈표 3〉 주택담보대출 금리특징 및 장·단점 .....	11
〈표 4〉 주택담보대출 상환유형 .....	12
〈표 5〉 원금균등분할상환과 원리금균등분할상환의 차이점 .....	13
〈표 6〉 주택담보대출 추이 .....	15
〈표 7〉 2013년 주택담보대출 상환구조 및 상환비율 .....	17
〈표 8〉 주택담보대출의 만기구조 현황 .....	18
〈표 9〉 주택담보대출 금리구조 잔액 현황 .....	18
〈표 10〉 국가별 고령화 진행 현황 .....	22
〈표 11〉 금융기관의 주택담보대출 담보인정비율 .....	24
〈표 12〉 금융기관의 주택담보대출 DTI 규제내용 .....	25
〈표 13〉 LTV와 DTI 규제 변천과정 .....	26
〈표 14〉 DTI 규제와 금융기관 채무상환능력평가 적용 .....	27
〈표 15〉 DTI 계산기 .....	28
〈표 16〉 주택담보대출 기한의 이익상실 시기 .....	31
〈표 17〉 주택담보대출의 연체기간에 따른 건전성 분류 .....	33
〈표 18〉 시중은행의 주택담보대출 단계별 연체관리 .....	34
〈표 19〉 주택담보대출의 전이억제율과 회수율 .....	35
〈표 20〉 주택담보대출 고정이하 연체비율 추정 산식 .....	36
〈표 21〉 국가별 가처분소득 대비 가계부채 비율 .....	37
〈표 22〉 정부의 가계부채 구조개선 촉진방안 .....	38
〈표 23〉 정부의 은행권 주택담보대출 구성비 목표 .....	38
〈표 24〉 정부의 보험권 및 상호금융권 비거치식 분할상환 목표 .....	39
〈표 25〉 주택담보대출 구조개선 확대방안 .....	39
〈표 26〉 주택담보대출 취약계층 채무조정방안 .....	40
〈표 27〉 미국의 공공 모기지 지원 프로그램 .....	43

<표 28> 영국의 가계부채 구성비 .....	44
<표 29> 아이슬란드 채무재조정 프로그램 .....	49
<표 30> 국가별 가계부채 원인과 영향 .....	53
<표 31> 가계부채에 대한 대응법 .....	56
<표 32> 데이터마이닝기법의 종류 .....	57
<표 33> 주택담보대출에 관한 선행연구 .....	61
<표 34> 주택담보대출의 리스크 요인에 관한 선행연구 .....	64
<표 35> 주택담보대출의 연체가능성에 관한 선행연구 .....	67
<표 36> 데이터마이닝 모형의 연구내용 .....	69
<표 37> 인공신경망 모형의 금융관련 적용분야 .....	76
<표 38> 의사결정나무 모형의 분석단계 .....	80
<표 39> 의사결정나무 모형의 적용 .....	81
<표 40> 의사결정나무 분석의 알고리즘 비교 .....	82
<표 41> 의사결정나무 분석에 사용된 변수 .....	83
<표 42> 금융감독원 연체율 산정기준 .....	86
<표 43> 분석대상 주택담보대출 자료의 특성 .....	92
<표 44> 분석사례의 기술적 통계분석 .....	93
<표 45> 분석모형의 변수 .....	95
<표 46> 연도별 LTV 평균사용률 현황 .....	96
<표 47> 연령 분포현황 .....	96
<표 48> 차입금액 분포현황 .....	96
<표 49> 대출기간 분포현황 .....	96
<표 50> 담보제공평형 분포현황 .....	96
<표 51> 성별간 주요변수의 평균 .....	97
<표 52> 평균비교 결과 : t-검정 .....	98
<표 53> 구입목적별 주요변수의 평균 .....	99
<표 54> 평균비교 결과 : t-검정 .....	99
<표 55> 대출시기별 주요변수의 평균 .....	101
<표 56> 평균비교 결과 : 분산분석 .....	102

〈표 57〉 연체여부에 따른 주요변수의 평균 .....	103
〈표 58〉 평균비교 결과 : t-검정 .....	104
〈표 59〉 로지스틱 회귀분석 결과 .....	106
〈표 60〉 로짓모형을 통한 연체가능성 예측정확도 .....	109
〈표 61〉 인공신경망분석 결과표(10일 이상 연체) .....	111
〈표 62〉 인공신경망분석 결과표(20일 이상 연체) .....	112
〈표 63〉 인공신경망분석 결과표(30일 이상 연체) .....	113
〈표 64〉 인공신경망 모형을 통한 연체가능성 예측정확도 .....	114
〈표 65〉 10일 이상 연체기준 Rule set .....	120
〈표 66〉 20일 이상 연체기준 Rule set .....	125
〈표 67〉 30일 이상 연체기준 Rule set .....	129
〈표 68〉 의사결정나무 모형을 통한 연체가능성 예측정확도 .....	130
〈표 69〉 오분류율 예측정확도 비교 .....	132
〈표 70〉 오분류율 표에 의한 예측 .....	132
〈표 71〉 오분류율 표에 의한 모형별 통계량 .....	133

## 그 립 목 차

〈그림 1〉 연구의 구성 .....	6
〈그림 2〉 연도별 가처분소득대비 가계신용비율 증가 추이 .....	16
〈그림 3〉 주택가격상승률과 주택보급률 .....	19
〈그림 4〉 아파트 경매건수와 매각가율 .....	20
〈그림 5〉 생산가능 인구비율 및 고령인구 추계 .....	21
〈그림 6〉 금융위기 이전의 가계부채 증가와 향후 소비깎과의 관계 .....	41
〈그림 7〉 영국의 가계부채 추이 .....	44
〈그림 8〉 영국의 주택가격 및 개인대출과 가계소비지출 추이 .....	45
〈그림 9〉 영국의 주택가격 상승률과 평균 주택가격 .....	46
〈그림 10〉 영국의 가처분소득대비 가계부채 비중 추이 .....	47
〈그림 11〉 아이슬란드의 경제성장률과 주택가격 상승률 및 주택가격 지수 .....	48
〈그림 12〉 스페인의 가처분소득 대비 가계부채 추이 .....	50
〈그림 13〉 스페인의 주택가격 추이 .....	51
〈그림 14〉 국가별 주택담보대출 금리 유형 .....	54
〈그림 15〉 연체가능성 분석모형의 흐름도 .....	71
〈그림 16〉 로지스틱 회귀분석 모형 .....	72
〈그림 17〉 다층 퍼셉트론 인공신경망의 구조 .....	78
〈그림 18〉 국내 원화대출 부문별 연체율 현황 .....	88
〈그림 19〉 주택담보대출 연체율 현황 .....	89
〈그림 20〉 인공신경망모형 구축 .....	110
〈그림 21〉 의사결정나무 분석의 모형구축 결과(10일 이상 연체) .....	117
〈그림 22〉 의사결정나무 분석의 트리 결과1(10일 이상 연체) .....	118
〈그림 23〉 의사결정나무 분석의 트리 결과2(10일 이상 연체) .....	119
〈그림 24〉 의사결정나무 분석의 모형구축 결과(20일 이상 연체) .....	123
〈그림 25〉 의사결정나무 분석의 트리 결과(20일 이상 연체) .....	124
〈그림 26〉 의사결정나무 분석의 모형구축 결과(30일 이상 연체) .....	127
〈그림 27〉 의사결정나무 분석의 트리 결과(30일 이상 연체) .....	128
〈그림 28〉 AUC 모형과 ROC Curve 결과 .....	134

# 제 1 장 서 론

## 제 1 절 연구의 배경 및 목적

우리나라는 1970년대 이후로 급격한 경제성장을 이루면서 도시화 현상이 급격하게 진행되어 왔다. 이로 인해 인구의 도시집중, 가구 분화 등으로 도시 내 주택이 부족한 현상이 나타나기 시작했고 심각한 사회문제로 대두되었다. 따라서 이 시기에는 정부가 주택공급을 주도하였으며 주로 분양시장을 중심으로 공급확대 정책이 실행되었다(권병조, 2006). 이러한 분양위주의 주택공급은 당시 개별가계의 자산축적이 많지 않았던 시기로 필수적으로 부동산금융을 통한 자금 공급이 필요하게 되었다.

부동산금융이란 부동산을 취득하거나 개발에 필요한 자금을 융통하는 것을 말한다. 부동산의 취득과 개발에는 많은 자금이 소요되므로, 사람들은 자기자본(equity capital)뿐만 아니라 대부분 부채자본(debt capital)을 활용한다.<sup>1)</sup> 하지만 우리나라 경우 금융기관을 통한 부동산금융 지원은 경제성장 동력마련을 위해서라는 명목으로 기업에 집중되었다. 그 결과 외환위기 이전까지 가계에 대한 자금지원은 미미했었다. 정부에서는 이러한 문제를 선분양제도와 청약통장제도, 국민주택기금 설치 등을 통해서 보완했으나 많은 주택구입자들이 자금조달에 어려움을 겪었다.

우리나라 고유의 전세제도가 시장에서 빠르게 자리를 잡게 된 것도 제도화된 주택금융의 미비가 한 원인이 되었다. 그 결과 외환위기 직전인 1996년 공식부문의 주택금융이 전체 시장의 38.4%에 불과하며 전세보증금이 전체 금융의 절반을 초과하였다(김정인 외, 2002, 손재영 외, 2008). 이렇게 주택금융시장의 규모도 작았던 데다가 경제성장과 함께 가계소득과 주택가격이 빠르게 상승하면서 주택담보대출에 대한 리스크관리 필요성도 높지 않았었다. 절대적인 대출규모가 크지 않은 상황에서 상승하고 있는 주택가격을 고려하면 LTV 비율(loan to value ratio)<sup>2)</sup>도 낮았으며, 금융기관에서는 대출채권의

1) 안정근(2014), 『현대부동산학』 제5판, 양현사, p.359.

2) 담보인정비율 또는 주택담보대출비율이라고도 한다.

부실화 가능성이 크게 낮았기 때문이다. 대출을 받는 가계입장에서도 주택공급이 부족한 상황에서는 지렛대 효과(leverage effect)가 가능하였고, 주택매각을 통한 자본차익을 누릴 수 있으므로 채무불이행 가능성이 높지 않았다. 그러나 지렛대 효과는 타인자본을 이용함으로써 투자자의 위험-수익의 관계를 증폭시키는 역할을 함으로써 투자수익률을 높이기도 하지만, 다른 한편으로는 그 만큼 위험도 증대시킨다.<sup>3)</sup>

이러한 상황은 외환위기를 거치면서 크게 변화되었다. 대기업들이 무리한 사업확장과 자금조달을 감당하지 못하여 연쇄적으로 부도가 나기 시작하고 경제도 급격하게 침체되었다. 이로 인해 관치금융과 기업에 대한 시설자금지원을 중심으로 한 우리나라 금융시스템도 변화될 수밖에 없었다. 정부에서도 경제활성화를 위해서 분양가규제해제, 전매제한 폐지, 외국인의 취득 및 이용자유화 등 각종 규제를 완화하면서 금융규제완화의 일환으로 부동산대출을 자유화하였다(손재영 외, 2008).

금융기관의 가계에 대한 부동산대출은 기업대출 리스크에 비해서 담보의 건전성과 회수의 안정성이 높다고 판단하여 가계대출과 주택구입에 필요한 주택담보대출로 크게 확대되었으며 2013년말 예금은행이 취급한 주택담보대출잔액 비중은 가계대출의 68.4%를 차지하고 있다. 이러한 주택담보대출의 확대는 2000년대 후반까지 주택시장 활성화에 긍정적인 영향을 미치고 주택공급을 확장시켜 2010년에 주택보급률 100%를 달성하는데 기여하였다.

하지만 2008년 미국의 서브프라임 모기지(subprime mortgage)사태가 발생하면서 단기간에 크게 증가한 주택담보대출 리스크 관리에 대한 관심이 높아지고 있다. 또한 2010년대 들어 주택시장은 장기안정 추세 양상을 나타내고 있고 인구증가율이 점차 감소하면서 고령사회 진입을 앞두고 있는 등 시장환경도 변화하고 있다.

정부에서도 2009년부터 가계부채 관리를 위한 정책방향을 발표하여 이자부상환과 변동금리가 대부분이었던 주택담보대출 구조를 원리금균등상환과 고정금리대출로 개선하는 노력을 하고 있다. 하지만 주택담보대출의 건전성확보를 위해서 가장 중요한 역할이 필요한 기관은 대출주체인 금융기관이 되어야

---

3) 안정근, 전계서, p.275.

한다. 우리나라는 이미 대부분 금융기관이 민영화되었으므로 자체적인 리스크 관리방법을 개발하고 대응방안을 마련할 필요가 있다. 그런데 금융기관은 아직까지도 정부의 대출규제방안인 LTV비율과 DTI비율(debt to income ratio)<sup>4)</sup>정도를 대출 전에 고려하고 채무불이행 시에는 경매시장을 통해 채권을 회수하는 전통적 방식이 리스크 관리의 주된 방법으로 사용되고 있다.

앞으로 주택시장이 안정화되면서 정부의 금융규제도 점차 완화될 것으로 예상되고 있으며 경매시장에서도 경락가율에서 지역별, 유형별 양극화현상이 나타나고 있어, 금융기관은 능동적으로 리스크 관리방안을 마련할 필요가 있다. 최근 관심이 높아지고 있는 빅데이터(big data)<sup>5)</sup> 등의 활용가능성을 고려할 때, 고객들의 대출상환실적과 인구통계학적 특성정보를 보유하고 있는 금융기관이 리스크관리를 위한 다양한 시도를 할 필요가 있다.

이에 본 연구에서는 주택을 구입하면서 발생하는 주택담보대출의 리스크 관리를 위해 연체가능성을 분석할 수 있는 방안을 제시하고자 한다. 이를 위해서 먼저 현재 우리나라의 가계부채현황 및 구조를 살펴볼 필요가 있다. 주택담보대출 특성은 해외와는 다르기 때문에 분석에 필요한 변수와 방법도 차별화 되어야 하기 때문이다. 또한 분석방법도 선행연구에서 사용되고 있는 회귀분석방법과 함께 새롭게 대두되고 있는 분석방법인 데이터마이닝(data mining) 분석을 통해서 효과적인 방안을 도출하고자 한다.

데이터마이닝은 다양한 목적으로 사용될 수 있으며, 목적에 따라 분류할 때, 예측(prediction), 분류(classification), 군집화(clustering), 연관규칙(association rule) 추출로 구분할 수 있다.<sup>6)</sup> 결과적으로는 본 연구를 통해서 금융기관 주택담보대출 리스크관리의 중요성과 건전성을 알아보고 다양한 분석방법을 통해서는 최적의 연체관리 예측모형을 제시하는 데 목적이 있다.

---

4) 총부채상환비율이며 소득에 비해 대출원금과 이자를 얼마나 상환할 수 있는가의 비율이며 정부의 규제수단이다.

5) 빅데이터란 대량의 데이터를 의미함, 실시간으로 끊임없이 다양한 형태의 데이터가 수집되는 것을 말함, 한국경제연구원(2013), KIPA 연구보고서 2013-04, <표 2-7> 빅데이터 개념 정의 정리, p.42.

6) 전치혁(2013), 『데이터마이닝기법과 응용』, 한나래, p.15.

## 제 2 절 연구범위

주택담보대출에 대한 연구는 다양한 측면에서 접근할 수 있다. 먼저 금융정책적 측면에서는 국가에서 주택담보대출에 대한 건전성 규제를 어느 수준까지 개입해야 하며 적절한 규제수단이 무엇인지에 대한 연구가 있다. 주로 LTV와 DTI 및 대출총액규제 등의 적절성 및 효과성을 분석하는 연구 등이 해당한다. 이러한 연구는 연체가능성이 높은 가구별 특징을 파악하거나 개별 금융기관의 특성을 고려하지 못하는 한계가 있다.

다음으로 주택가격상승률과 금리 등 거시적 변수와 금융권 전체연체율과의 관계를 살펴보는 연구가 있다. 거시적 변수를 고려하는 경우 연구범위를 거시적 측면으로 설정하므로 가구별 특성이나 금융기관의 특수성 등을 고려하지 못하게 된다.

마지막으로 금융기관에서 직접적으로 실행된 실적을 통해서 연체가능성이 있는 차주의 특성과 대출유형, 대출규모 등을 결합하여 분석하는 방법이 있다. 이 경우 자료취득에는 한계가 있으나 허용되는 자료를 이용하므로 가장 적용가능성이 높은 방법으로 볼 수 있다.

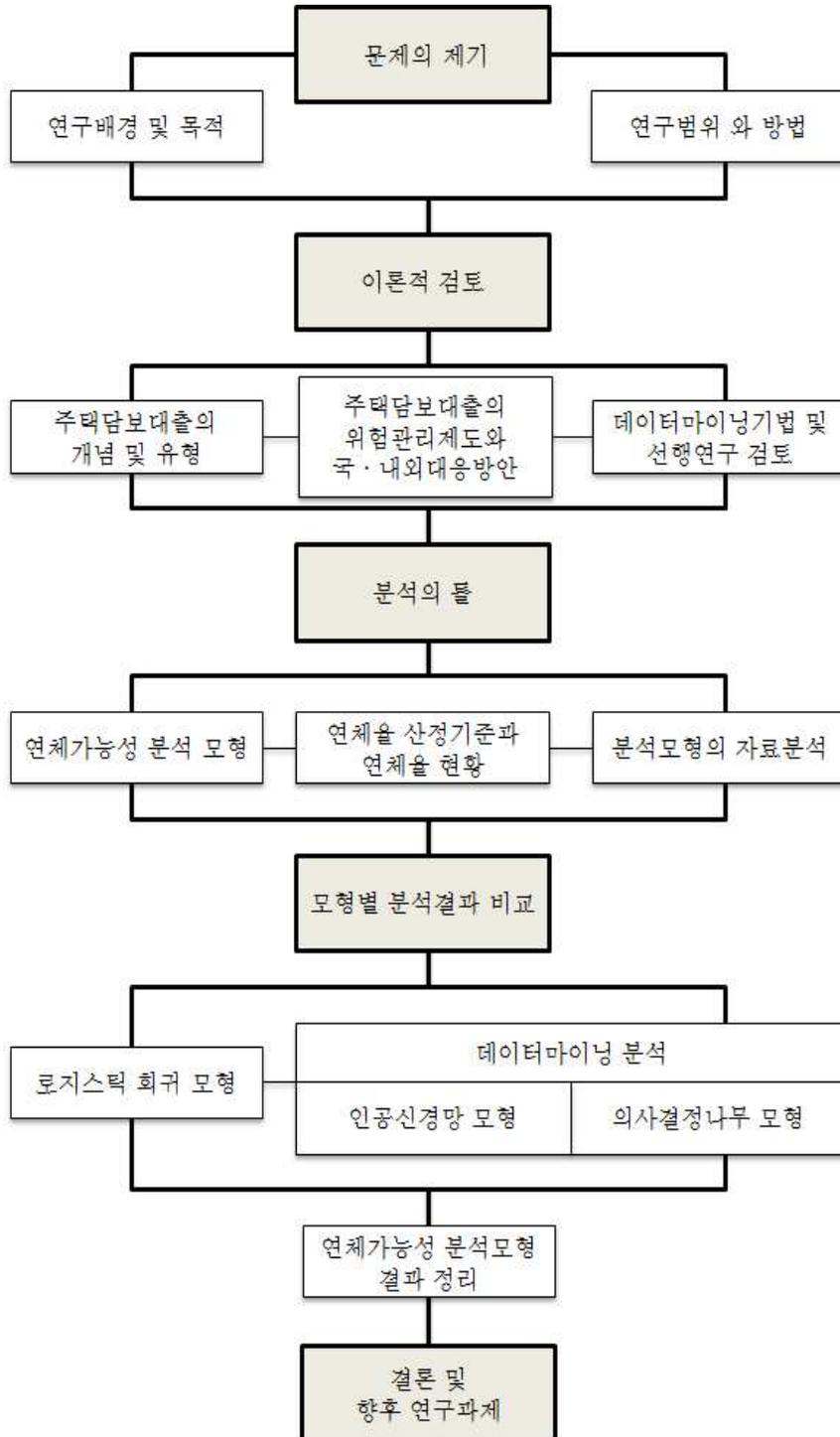
각각의 연구범위가 분야별로 제공하는 시사점이 다양하지만 본 연구의 목적이 금융기관에서 연체가능성을 분석할 수 있는 다양한 방법론의 적용가능성을 살펴보는 것이므로 적합한 연구범위를 설정 하였다. 먼저 분석대상은 허용된 자료로서 접근가능성과 정확성을 고려하여 시중금융기관 서울소재 대출 자료를 활용하였다. 일반적으로 대출이 실행된 후 일정기간이 지난 후에야 연체가능 여부를 판단할 수 있으므로 2007년부터 2010년 사이에 이루어진 주택담보대출 데이터를 통해서 연체상태를 기준으로 분석을 실행하였다. 주택담보대출이 이루어진 주택의 유형별로도 연체가능성의 차이가 있을 수 있으나 아파트가 아닌 경우 대출사례가 충분하지 않아 아파트를 중심으로 분석을 실행하였다.

### 제 3 절 연구방법

주택담보대출에 대한 리스크관리를 위한 분석방법으로도 다양한 방법이 있다. 거시경제지표와의 관련성을 분석하는 경우 시계열분석방법이 많이 사용되었으며 대출데이터를 활용한 연구에서는 회귀분석모형이 주로 사용되었다. 특히 금융기관의 관점에서 접근하는 경우 자체적인 대출데이터를 활용할 수 있어 다양한 통계적 분석방법을 적용할 수 있다. 본 연구에서도 일반적으로 주택담보대출의 연체가능성을 분석하기 위해서 주로 사용되는 회귀모형을 이용하여 분석하였으며 여기에 최근 대용량데이터분석 등을 통해서 중요도가 높아지고 있는 데이터마이닝 모형을 적용하였다.

구체적인 연구방법을 살펴보면, 먼저 이론적 검토를 위해서 주택담보대출의 개념과 유형을 살펴보고 리스크발생의 가능성 및 대응방안을 파악하고자 한다. 이를 통해서 향후 금융기관에서 리스크관리가 중요한 역할을 하게 될 것이라는 것을 제시한다. 다음으로 분석대상 자료들과의 관계를 분산분석 등을 통해서 살펴보고, 연구모형인 로지스틱 회귀모형과 인공신경망모형, 의사결정모형을 통해서 연체가능집단의 특성을 파악하고자 한다. 또한 데이터마이닝 모형인 인공신경망모형과 의사결정나무모형을 통하여 로지스틱 회귀모형과의 모형별 차이를 제시하고 향후 연체가능성 집단을 선별하는 우수한 모형을 제시 하고자한다.

〈그림 1〉 연구의 구성



## 제 2 장 이론적 검토

### 제 1 절 주택담보대출의 개념 및 유형

#### 1) 주택담보대출의 개념

주택담보대출은 부동산의 전형적인 부채금융방식으로 저당금융과 신탁금융이 있다.<sup>7)</sup> 저당금융을 금융기관은 주택담보금융이라고 하며 차입자가 소유권을 보유한 채 대상 부동산에 근저당을 설정하여 금융기관으로부터 직접대출을 받는 것을 말한다.

〈표 1〉 주택담보 금융의 저당방식

구 분	저당금융	신탁금융
소유권자	차입자	신탁회사
등기방식	근저당권 설정	신탁등기
부동산관리	금융기관	신탁회사
채무불이행시	경매처분	공매처분
처분방법	절차복잡, 비용많음	절차간편, 비용저렴

자료 : 안정근(2014), 『현대부동산학』 제5판, 양현사, p.367.

부동산담보금융은 부동산과 관련된 권리를 담보로 제공하고 자금을 조달하는 방법을 의미한다. 부동산담보금융은 제공되는 담보물건에 따라서 주택담보대출, 상가담보대출, 토지담보대출 등으로 구분되며 주택담보대출은 주택을 담보로 하여 자금을 조달하는 방법을 의미한다. 우리나라의 금융기관에서 주택담보대출은 계정분류상 ‘주택자금대출’과 ‘가계일반자금대출’로 구분된다. ‘주택자금대출’은 금융회사의 자체자금을 이용하여 주택을 담보로 자금을 차입해주는 상품으로 자금의 사용목적은 주택취득 및 전세자금 등으로 한정된 경우에 해당한다. ‘가계일반자금대출’은 거주용 주택구입목적 이외에 투자용

7) 안정근, 전계서, p.367.

주택구입, 생활자금마련 등 다양한 대출의 용도가 혼재되어 있다(손재영 외, 2008).

금융기관 대출상품들은 주택담보대출 외에도 공적목적으로 제한된 대상에게 공급되는 국민주택기금대출과 주택금융공사 모기지론이 있다. 이러한 국민주택기금대출은 정부의 융자성 자금계정인 국민주택기금에서 조성된 자금을 이용하여 국민주택규모 이하(85m<sup>2</sup>이하) 주택의 취득 시 자금을 지원하고 있다. 대부분 대출이 고정금리·장기대출방식으로 지원되고 있으나 최근 디딤돌대출과 공유형모기지 등 다양한 지원방안이 개발되고 있다.

한국주택금융공사가 공급하는 모기지론(보금자리론)은 자본시장에 대출채권을 유동화 하여 자금을 조달하고 있으며 시장은행을 통해서 모기지론을 판매하고 있다. 그리고 유동화를 전제로 대출이 이루어지므로 증권화에 유리한 장기고정금리·원리금분할상환형으로 대출이 이루어지고 있다. 지원조건을 다양화하여 적격대출, e보금자리론, 기금과 연계한 디딤돌대출 등 다양한 대출상품을 제공하고 있다(고성수, 2008).

이렇게 주택담보대출의 큰 범주아래 여러 기관에서 다양한 상품이 출시된 이유는 주택공급확대와 금융자율화로 우리나라 주택담보대출시장이 커졌기 때문이다. 하지만 주택담보대출시장의 빠른 성장에 비해서 만기구조 및 상환방식의 안정화 등 대출구조 선진화 등은 미진한 상황이다.

## 2) 주택담보대출의 유형

### 가) 금리형태에 따른 유형

주택담보대출의 유형을 구분하는 데는 여러 가지 기준이 있으나 금리의 변동가능성을 통해서 구분하는 방법이 많이 사용되고 있다. 주택담보대출 금리는 크게 고정금리방식, 변동금리방식 및 고정과 변동금리를 혼합한 혼합형방식 등의 종류가 있다.

고정금리방식은 대출기간 중 금리가 변하지 않고 만기까지 유지되는 대출로서, 대출 당시에는 변동금리방식보다 금리가 높다는 단점을 가지고 있으나,

시장금리 변동에 영향을 받지 않고, 매일 납입하는 이자액이 균일하여 시장금리 상승 시에는 변동금리대출 보다 유리할 수 있다는 장점이 있다.<sup>8)</sup>

변동금리방식은 일정 주기인 3개월, 6개월, 12개월 등의 주기변동마다 금리가 변동되는 특성을 가지고 있으며, 대출초기 금리는 고정금리방식과 비교하여 낮다는 장점이 있지만, 매일 납입해야 하는 이자금액이 일정 주기마다 변동될 수 있고 시장금리 상승 시에는 이자부담이 증가할 수 있다는 단점이 있다.<sup>9)</sup> 변동금리 구조<sup>10)</sup>에 사용되는 시장금리는 양도성예금증서(CD)<sup>11)</sup>와 코픽스(COFIX)<sup>12)</sup>, 코리보(KORIBO)<sup>13)</sup>, 국·공채, 금융채 등이 있으며, 금융기관에 따라 내부금리기준을 정하여 사용되는 내부기준금리와 최우량고객에 사용되는 프라임 레이트(prime rate)등이 있다.

이러한 시장금리들은 <표 2>와 같이 기준금리에 영향을 주는 특징을 가지고 있어서, 차입자 입장에서 변동금리를 선택할 때는 기준금리에 적용되는 변동성의 크기와 장점을 파악하여 안정성이 고려되는 정교한 금리선택이 필요할 수 있다.

그리고 변동금리의 경우 시장금리에 연동되고 있으므로 적용된 기준금리 변동성이 클 경우 차입자 이자율위험이 커진다는 문제점이 있다. 따라서 최근 정부에서는 가계부채의 구조개선을 위해서 변동금리 위주의 주택담보대출을 고정금리로 전환하는 방안을 추진 중이다.

---

8) 전국은행연합회(2010), 『주택담보대출 핸드북』, p.11.

9) 상계서, p.11.

10) 변동금리 구조는 기준금리 + 금융기관이 정하는 가산금리인 스프레드(spread)로 이루어지고 있다.

11) 양도성예금증서(CD : certificate of deposit)는 은행이 단기의 자금 조달을 위해 발행하는 무기명 정기에금증서로 한국에는 1984년 도입됐다. 2014년 현재 CD금리는 신용 AAA등급 7개 시중은행이 발행한 CD에 대해 10개 증권사가 금리를 평가해, 하루에 두 번 수익률을 금융투자협회에 보고하는 형식으로 결정된다.(자료 : 금융투자협회, <http://www.kofia.or.kr>, 경제학사전).

12) COFIX(Cost of Funds Index)는 국내 9개 정보제공은행들이 제공한 자금조달 관련 정보를 기초로 하여 산출되는 자금조달비용지수로서, "잔액기준 COFIX", "신규취급액기준 COFIX" 및 "단기COFIX"로 구분 공시되며, 주택담보대출 금리 사용은"잔액기준 COFIX", "신규취급액기준 COFIX"가 사용되고 있다.(자료 : 전국은행연합회, <http://www.kfb.or.kr>).

13) 코리보(KORIBOR : KOREA Inter-Bank Offered Rates)는 영국의 'LIBOR' 처럼 한국의 은행간 기준 금리이다. 2014년 3월 현재 시중 금융기관 13곳의 금융기관 금리의 평균으로 결정되고 있다. 13곳은 시중은행 7개, 특수은행 3개, 지방은행 3개이며 금리는 이중 상위 3개, 하위 3개를 제외하고 나머지 7개 금융기관 금리의 평균으로 결정된다.(자료 : 전국은행연합회 금융자료실, <http://www.kfb.or.kr>).

〈표 2〉 주택담보대출 변동금리에 적용되는 기준금리의 종류

기준금리 종류		특 징
시장금리	양도성예금증서(CD)	양도성예금증서(CD) + 금융기관이 정하는 가산금리 스프레드(spread) = 양도성예금증서에 연동된 변동금리
	코픽스(COPIX)	잔액기준 코픽스(COPIX) 또는 신규취급액기준 코픽스(COPIX) + 금융기관이 정하는 가산금리 스프레드(spread) = 코픽스(COPIX)에 연동된 변동금리
	코리보(KORIBO)	코리보(KORIBO) + 금융기관이 정하는 가산금리 스프레드(spread) = 코리보(KORIBO)에 연동된 변동금리
	국공채, 금융채	국공채, 금융채 + 금융기관이 정하는 가산금리 스프레드(spread) = 국공채, 금융채에 연동된 변동금리
내부기준금리		금융기관 내부기준 금리 + 금융기관이 정하는 가산금리 스프레드(spread) = 내부기준금리에 연동된 변동금리
프라임 레이트(prime rate)		금융기관 최우량 금리 + 금융기관이 정하는 가산금리 스프레드(spread) = 프라임 레이트(prime rate)에 연동된 변동금리

자료 : 이재연 외(2010), 『주택담보대출 기준금리개선 방안』, 한국금융연구원, 내용참고.

고정금리와 변동금리 혼합형은 고정금리방식과 변동금리방식이 결합된 것으로, 일정기간 동안은 고정금리방식이 적용되다가 이후에는 변동금리방식이 적용되거나 반대로 변동금리방식이 적용되다가 고정금리방식이 적용되는 것이다. 현재 금융기관에 따라서는 대출기간 중 횟수에 관계없이 금리방식을 전환할 수 있는 상품을 갖추고 있다.<sup>14)</sup>

14) 전국은행연합회, 전계서, p.17.

〈표 3〉 주택담보대출 금리특징 및 장·단점

금리종류	금리형태	특 징	장/단점
고정금리		- 대출 기간 중 동일한 이자율로 적용한 금리가 변하지 않고 만기까지 유지되는 대출	장점 : 대출만기까지 동일한 금리유지 단점 : 변동금리 보다 금리가 높다
변동금리		- 대출주기인 3개월, 6개월, 12개월 등의 주기마다 금리가 변동되며, 실세금리와 연동되는 특성을 가지고 있음	장점 : 실세금리 하락에 따라 금리가 하락되며, 고정금리보다 금리가 낮음 단점 : 실세금리 상승에 따른 금리상승
혼합형금리		- 고정금리방식과 변동금리방식이 일정기간 동안 서로 결합된 혼합금리임	장점 : 금리예측과 자금계획에 맞춘 금리선택 가능 단점 : 금리예측 및 자금계획이 어려움

자료 : 전국은행연합회(2010), 『주택담보대출 핸드북』, pp.16-17.

### 나) 상환방식에 따른 유형

금리의 변동가능성과 함께 주택담보대출의 유형을 구분하는 방법으로 대출금의 상환방식에 따른 유형이 있다. 대출금의 상환방식에 따른 유형은 원금균등분할상환, 원리금균등분할상환, 만기일시상환, 원금일부분할상환 등이 있으며, 일반적으로 원금균등분할상환방식과 원리금균등분할상환방식이 비중 높게 활용되고 있다.

각 유형별 특징은 아래 〈표 4〉에 나타난 바와 같다. 우리나라에서는 2000년대 중반까지 만기일시상환이 주로 사용되었으나 가구부채의 안정성을 위해서 대출기간 동안 일정한 원금을 상환하는 것이 필요하므로 정부에서는 정책적으로 원리금균등상환으로 전환하는 것을 추진하고 있다.

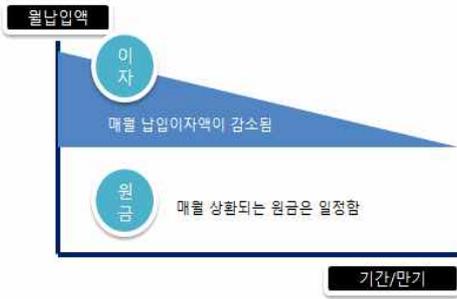
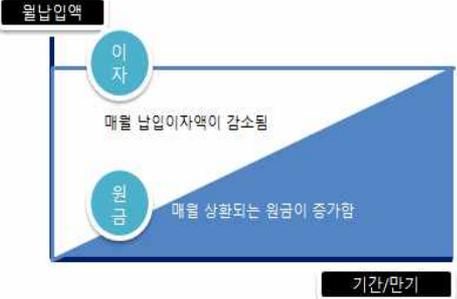
〈표 4〉 주택담보대출 상환유형

상환유형	상환특징
원금균등분할상환	대출기간 동안 매번 원금이 균등하게 상환되어 이자가 매월 줄어드는 방법
원리금균등분할상환	대출기간 동안 일정한 균등금액을 상환하는 방법
만기일시상환	대출만기까지 이자만 납입하다가 만기에 원금일시 상환하는 방법
원금일부분할상환	일정기간 거치 후 원금 중에 일정 부분만 상환하는 방법

자료 : 전국은행연합회(2010), 『주택담보대출 핸드북』, p.19.

주택담보대출의 상환방법은 매월 납입하는 상환액과 밀접한 관련이 있으므로 차입자 입장에서는 적합한 상환유형을 활용하는 것도 중요하다. 점차비중이 증가하고 있는 상환방법은 원금균등분할상환과 원리금균등분할상환으로 아래 〈표 5〉와 같으며 나타난 바와 같이 초기 납입액은 원금균등분할상환이 크므로 초기 채무불이행가능성이 높은 것으로 알려져 있으며 대출기간 동안 전체원리금상환액이 적다는 특징을 가지고 있다. 반면에 원리금균등분할상환은 기간전체의 상환액이 동일하다는 장점이 있지만 원금의 감소는 원금균등분할상환에 비해서 느리게 진행된다는 단점이 있다.

〈표 5〉 원금균등분할상환과 원리금균등분할상환의 차이점

원금균등분할상환	원리금균등분할상환
	
<ul style="list-style-type: none"> <li>- 초기 이자부담비용이 원리금균등분할상환 보다 크고, 원금은 매월 균등하게 상환된다.</li> <li>- 대출약정기간 동안 전체상환액은 원리금균등분할상환액 보다 적게 상환되는 특징이 있다.</li> <li>- 이는 원금상환액을 균등하게 만기까지 상환하기 때문이다.</li> </ul>	<ul style="list-style-type: none"> <li>- 초기 이자부담비용이 원금균등분할상환보다 적으며, 원금상환도 적게 한다.</li> <li>- 대출약정기간 동안 전체상환액은 원금균등분할상환액 보다 많게 상환되는 특징이 있다.</li> <li>- 이는 원금상환액을 초기에 적게 하고 만기기간에 다가갈수록 원금상환액이 많기 때문이다.</li> </ul>

자료 : 전국은행연합회(2010), 『주택담보대출 핸드북』, pp.19-20.

## 제 2 절 주택담보대출의 위험관리

### 1) 주택담보대출에 대한 위험관리 필요성

우리나라의 경우 주택담보대출 등 급격하게 가계대출이 증가하여 경제전반에 위협요인이 되고 있다. IMF에서 2012년에 발표한 분석에서도 가계부채율 증가가 소비의 추정치와 실제치의 차이를 발생시키며 ‘부’의 상관관계를 나타낸다는 결과를 발표하여 과도한 가계부채, 특히 주택담보대출은 경기침체의 악화와 회복에 주요 요인이 된다는 점을 제시하였다.<sup>15)</sup>

또한 우리나라의 경우 주택담보대출의 만기 및 상환구조와 급격한 고령화로 인한 신규대출 수요 진입의 한계 등으로 주택담보대출에 대한 많은 위험요인을 안고 있어 위험관리의 필요성이 증가하고 있다.

### 가) 주택담보대출의 규모 및 증가속도

우리나라는 외환위기 이후 금융권의 가계대출이 증가하고 주택가격이 상승하면서 주택담보대출도 크게 증가하였다. 이로 인해 2013년 말을 기준으로 우리나라의 가계신용은 1,000조원을 초과했으며, 가계대출에서 가장 큰 비중을 차지하고 있는 주택담보대출은 전체 가계대출의 70%에 달하고 있다. 주택가격의 상승과 자가소유비율의 증가 등으로 인해서 주택담보대출 증가가 지속되는 현상과 단기간 내 과도하게 증가한 측면을 고려한다면 부실가능성이 내재되어 있을 가능성이 높다. 실례로 미국의 경우에도 주택담보대출이 단기간 내에 증가되었고, 서브프라임모기지(sub prime mortgage)<sup>16)</sup> 등이 판매되면서 부실로 이어졌다. 따라서 급격한 주택담보대출 증가는 위험관리 필요성

15) IMF보고서에서는 2002년과 2006년의 가계부채증가율과 1996년부터 2004년 추세에 따른 2010년 가계소비 추정치와 2010년 실제 가계소비와의 차이에 대한 상관관계를 분석한 결과 10%의 DTI증가는 -2.1%p이상의 소비갭을 나타내 가계부채 증가가 향후 가계소비 감소로 이어져 경기침체의 원인이 될 수 있음을 제시하였으며 특히 영국과 스페인 등 유럽의 주요 국가도 대출증가로 인한 소비갭이 크게 나타나 최근 이들 국가의 경제위기가 가계부채와 무관하지 않은 점을 제시하였다.

16) 신용도가 낮은 사람들을 대상으로 비교적 높은 이자를 받기로 약정하고 주택을 담보로 빌려주는 대출을 말한다.

을 가져온다고 판단할 수 있다.

우리나라 가계신용은 2005년부터 급속한 성장을 보이며 증가하고 있으며, 2013년말 규모는 1021.3조원으로 전년말 대비 6.0%가 증가되었다. 가계대출 규모를 보면 은행과 비은행을 합한 주택담보대출 규모가 418조원으로 가계신용 규모의 40.9%를 차지하고 있으며 예금은행의 가계대출은 481.1조원이나, 주택담보대출은 328.9조원으로 그 비율이 68.4%에 달해 예금은행 주택담보대출 비중은 매우 높다고 할 수 있다.

〈표 6〉 주택담보대출 추이

(단위 : 조원, %)

구 분	2008	2009	2010	2011	2012	2013
가계신용 <sup>17)</sup> (a) (증가율)	723.5 (8.7)	776 (7.3)	843.2 (8.7)	916.2 (8.7)	963.8 (5.2)	1,021.3 (6.0)
주택담보대출 <sup>18)</sup> (b) (증가율) (비중, b/a)	311.2 (6.3) (43.0)	338.5 (8.8) (43.6)	362.8 (7.2) (43.0)	392 (8.0) (42.8)	404.2 (3.1) (41.9)	418.1 (3.4) (40.9)
가계대출 (증가율)	683.6 (8.5)	734.3 (7.4)	793.8 (8.1)	861.4 (8.5)	905.9 (5.2)	963 (6.3)
예금은행(c) (주택담보대출)(d) (비중, d/c)	388.6 254.7 (65.5)	409.5 273.7 (66.8)	431.5 289.6 (67.1)	455.9 308.9 (67.8)	467.3 318.2 (68.1)	481.1 328.9 (68.4)
비은행(e) (주택담보대출)(f) (비중, f/e)	126.7 56.4 (44.5)	140.3 64.8 (46.2)	162.1 73.2 (45.2)	183.7 83.1 (45.2)	192.6 86.0 (44.7)	206.1 89.2 (43.3)
판매신용	39.9	41.7	49.4	54.8	57.9	58.3

자료 : 한국은행 경제통계시스템, <http://ecos.bok.or.kr>, 2014. 10월 현재.

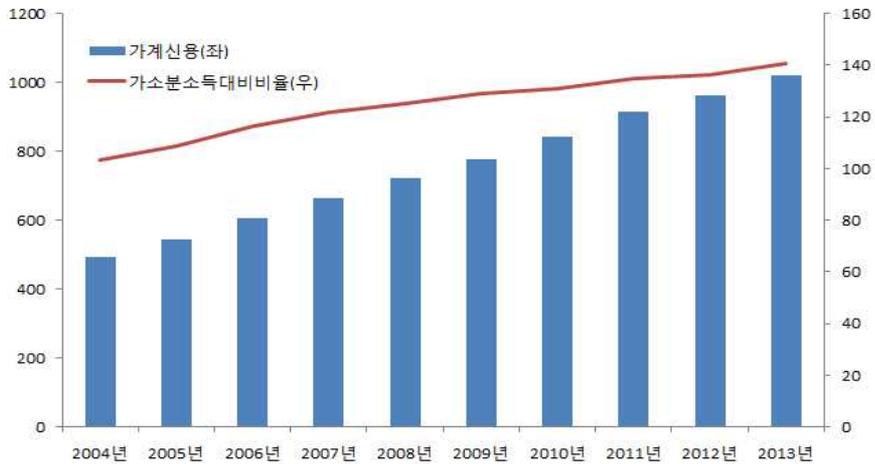
이렇게 가계부채가 증가하면서 가계부채에 대한 위험지표인 가처분대비 가

17) 가계대출과 판매신용을 합산한 대출을 가계신용이라 함.

18) 은행과 비은행의 합산한 대출.

계신용 비율도 크게 증가하였다. 우리나라의 가계부채는 10년간 527.1조원이 증가하였으며 연평균 8.4%가 증가하였다. 이에 비하여 가처분소득은 같은 기간 동안 249.2조원 증가하였으며 연평균 4.7% 증가하였다. 이에 따라서 2013년말 현재 가처분소득대비 가계부채 비율은 140.5%로 10년전 보다 37.1%가 증가하여 이는 가처분소득 보다 가계부채의 증가 속도가 빠름을 보여주고 있다.

〈그림 2〉 연도별 가처분소득대비 가계신용비율 증가 추이



자료 : 한국은행 경제통계시스템, <http://ecos.bok.or.kr>, 2014. 10월 현재.

정부는 가계부채의 규모와 주택담보대출의 규모가 지속적으로 증가하고 건전성 지표가 심각할 정도로 악화되었다고 판단하여 2011년 6월 ‘가계부채 연착륙 종합대책’을 통해 가계부채의 안정적 관리를 위한 대응방향과 정책수단을 발표하였다.

금융기관에 대하여 가계대출의 적정증가를 유도하는 고위험 주택담보대출의 BIS 자기자본비율<sup>19)</sup> 위험가중치를 상향 적용하였고, 채무자의 상환능력 확인 정착을 위한 지도와 금융기관 영업점 성과평가 지표인 가계대출 실적지표를 삭제하는 조치를 취했다. 이후에도 지속적으로 시장을 점검하고 있다.

19) BIS(Bank for International Settlements) 자기자본비율은 은행의 건전성 지표의 하나로 사용되고 있으며 산식은 다음과 같다. BIS 자기자본비율 = 자기자본/위험가중자산 × 100.

## 나) 주택담보대출 구조의 편중

우리나라 주택담보대출의 위험관리가 시급한 원인은 만기구조, 금리구조, 상환구조가 취약하다는 점이다. 주택담보대출 구조는 만기 길이를 의미하는 만기구조와 고정금리와 변동금리 비중의 편중현상을 살펴보는 금리구조, 상환 방식 비중을 살펴보는 상환구조 등으로 구분할 수 있다. 일반적으로 만기구조가 장기이고, 고정금리에 원금을 함께 상환하는 방식이 안정적인 구조로 평가받고 있다.

2013년 현재 주택담보대출 만기구조는 만기일시상환 대출이 31.4%, 거치기간중인 대출이 38.3%로서 이자만 상환중인 대출이 약 70%에 달하고 있다. 따라서 원금상환 중인 대출의 2배 이상 규모가 이자만 상환하고 있다. 이자부상환의 경우 소액의 이자만 낼 때는 부담이 없지만, 만기 시에는 거액의 원금을 상환해야 하며 상환부담이 크게 늘어나게 되므로 개선이 필요한 상황이다.

〈표 7〉 2013년 주택담보대출 상환구조 및 상환비율

납입방법	상환비율		비 고
만기일시상환 중	69.7%	31.4%	이자상환
거치기간 중		38.3%	
원금상환 중	30.3%		원금상환

자료 : 관계부처합동, 금융위원회(2014), 『가계부채 구조개선 촉진방안』, p.3.

대출의 만기구조는 2012년 3년 이하의 단기대출이 27.7%로서, 2007년에 37.8%에 비교한다면 감소한 비율이나, 5년 이하의 단기대출까지 포함한다면 30%가 넘는 비율로서 단기대출이 상대적으로 높은 편이다. 하지만 10년 초과 주택담보대출의 비중도 약 59%로 높은 편으로 만기구조가 개선되어가는 가운데 양극화현상이 나타나고 있다.

즉 주택담보대출 대부분이 단기대출이거나 10년을 초과하는 장기대출로 구성되어 있다. 이러한 이유는 정부에서 장기대출을 장려하기 위해서 15년 이상 대출의 경우에만 소득공제를 제공하면서 나타난 현상으로 정부차원 인센티브 제공이 만기구조 개선에 효과가 있다는 점을 확인할 수 있다.

〈표 8〉 주택담보대출의 만기구조 현황

(단위 : 조원, %)

구 분	3년 이하	3년 초과 ~5년 이하	5년 초과 ~ 10년 이하	10년 초과	잔 액
2007년	82.8 (37.8)	15.2 (6.9)	31.6 (14.4)	89.7 (40.9)	219.3 (100)
2008년	75.5 (33.3)	10.7 (4.7)	17.5 (7.7)	122.6 (54.2)	226.3 (100)
2009년	95.7 (36.5)	9.4 (3.6)	21.8 (8.3)	135.7 (51.7)	262.6 (100)
2010년	95.3 (33.5)	11.5 (4.1)	24.4 (8.6)	153.3 (53.9)	284.4 (100)
2011년	94 (31.0)	14.5 (4.8)	27.7 (9.1)	167.2 (55.1)	303.4 (100)
2012년	87.2 (27.7)	13.5 (4.3)	28.5 (9.1)	185.3 (58.9)	314.4 (100)

자료 : 금융감독원.

대출의 금리구조는 2013년 4월 기준 고정금리대출이 23.8%, 변동금리대출이 76.2%로서, 2012년보다는 변동금리대출에 대한 편중이 완화되고 있으나 70%가 넘는 비중을 보이고 있다. 변동금리는 경기변동이 그대로 반영되는 민감한 구조를 가지고 있어, 이러한 구조의 편중은 금리상승 시 이자부담이 가중될 수 있다. 특히 2008년에 발생한 금융위기에서 어려움을 겪었던 아일랜드, 스페인, 영국의 경우 유럽내 타국가에 비해서 변동금리 비중이 높아 우리나라와 유사성을 나타낸 점을 고려할 때 금리구조의 개선도 필요한 상황이다.

〈표 9〉 주택담보대출 금리구조 잔액 현황

(단위:%)

구 분	2012.12	2013.6	2013.12	2014.1	2014.2	2014.3	2014.4
고정금리	19.8	23.2	21.3	20.8	21.0	21.4	23.8
변동금리	80.2	76.8	78.7	79.2	79.0	78.6	76.2

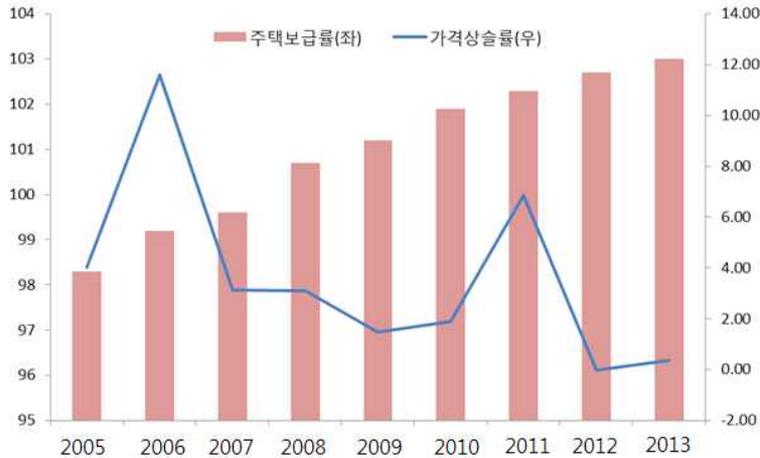
자료 : 한국은행 경제통계시스템, <http://ecos.bok.or.kr/>, 2014. 4월 현재.

#### 다) 주택시장의 환경변화

주택담보대출의 리스크관리가 필요한 이유는 우리나라 주택시장의 환경이

빠르게 변화하고 있기 때문이다. 주택시장의 환경변화는 다양한 방면에서 나타나고 있으며, 특히 주택가격의 안정화가 지속되고 있는 점과 빠르게 진행되고 있는 고령화 현상 등이 장기적으로 주택시장 및 주택담보대출의 안정성에 영향을 미칠 것으로 예상된다. 먼저 우리나라는 주택보급률이 점차 증가하면서 주택가격의 변동 폭이 크게 감소하고 있다. 이에 따라서 향후 주택시장은 과거와 같이 높은 가격상승률을 기대하기 힘든 상황이 되고 있으며, 따라서 가격상승을 예상하고 과도하게 대출을 받은 가구의 경우 채무불이행으로 진행될 가능성이 높아지고 있다.

〈그림 3〉 주택가격상승률과 주택보급률

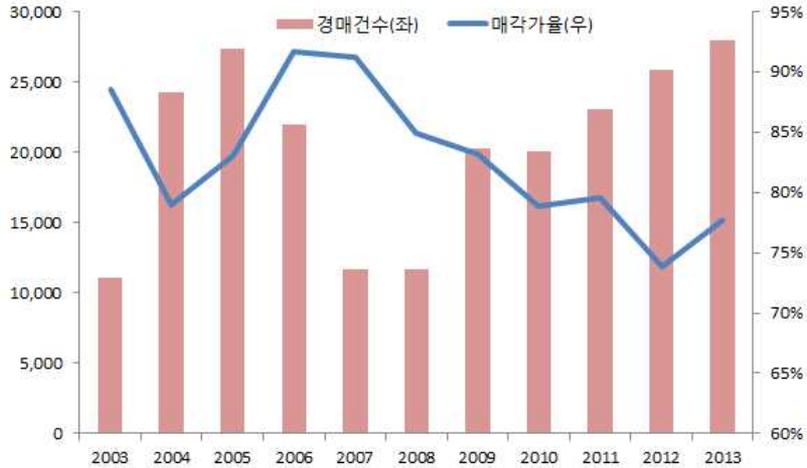


자료 : 한국은행 경제통계시스템, <http://ecos.bok.or.kr/>,  
KB 국민은행, <http://nland.kbstar.com/>, 2014. 10월 현재.

주택시장의 안정화로 인한 가계부채부실화의 징후는 경매실적 및 매각가율에서도 확인할 수 있다. 2003년 이후 경매건수와 매각가율을 살펴보면 금융위기 이후인 2009년부터 아파트 경매건수가 크게 증가하면서 매각가율이 하락하고 있다. 특히 매각가율의 경우 아파트의 매각가율은 80%를 초과한다는 심리적 마지노선이 2010년부터 무너지고 70%대를 유지하고 있으며, 2012년에는 73%까지 하락하여 금융기관에서도 자체적인 주택담보대출의 리스크 관리가 필요하다는 점을 시사하고 있다. 즉 정부의 규제로 인하여 대출시점에서

주택담보대출의 LTV는 50~60%수준을 유지하고 있으나 주택가격이 하락할 경우 매각가율 이상으로 상승할 가능성이 있어 사전에 주택담보대출의 건전성관리를 위한 추가적인 절차가 필요한 시기가 된 것으로 판단된다.

〈그림 4〉 아파트 경매건수와 매각가율



자료 : 대법원 법원경매정보(2003~2013), <http://www.courtauction.go.kr/>.

현재 빠르게 진행되고 있는 고령화현상도 주택담보대출 리스크관리가 필요한 요인이 되고 있다. 한국은 전세계에서 가장 빠른 속도로 고령화가 진행되고 있으며 고령화는 경제 전반에 영향을 미치는 것으로 알려져 있다.

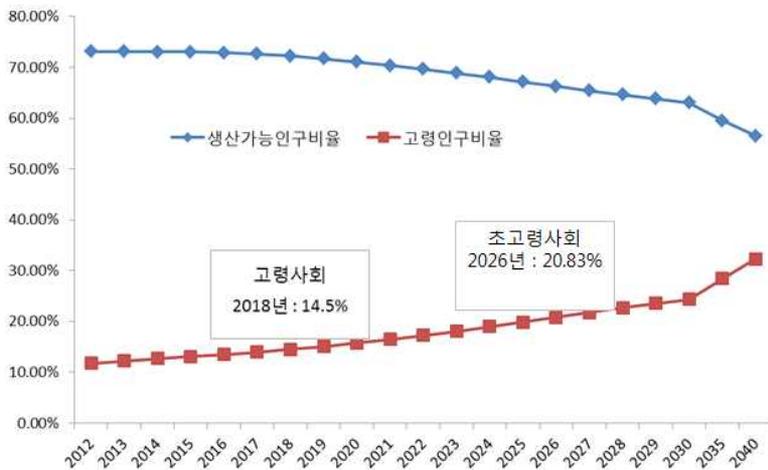
이혜훈(2002)과 성명기(2009)는 우리나라는 인구증가율이 급속히 감소되고 생산가능 인구에 비해 고령자 비중이 빠르게 증가함에 따라 청·장년층의 노령 인구 부양부담이 증가된다고 하였다. 또한 저축성향이 낮은 노년층 및 비근로 계층 비중의 증가는 저축률 감소를 가져오며, 생산가능 인구 감소는 재정수입이 감소되고 정부지출은 확대된다고 제시하였다. 이러한 세수감소와 재정지출 증가는 재정수지 악화로 이어져 재정 건전성이 저하될 것으로 연구결과를 제시하였다.

홍석철·전한경(2013)은 고령인구일수록 빈곤율이 높아지고 소득수준이 악화되는 현상은 향후 고령인구의 비중이 증가할수록 사회전체의 소득 불평등이 심화될 수 있다는 연구결과를 제시하였다. 이러한 고령인구의 증가가 생산

기능의 감소, 소득의 감소, 저축의 감소로 이어지는 현상을 발생시키고 국가적 재정 건전성 저하를 가져온다는 연구결과를 제시하고 있어 개인소득이 감소함에 따라 고령대출자들의 가계대출 및 주택담보대출자들의 상환능력 문제가 고령사회의 속도와 더불어 빠르게 진행될 수 있다.

사회의 고령화정도를 구분하는 방법으로는 65세 이상인구가 7%이상이면 고령화사회이며 14%이상이면 고령사회, 20%이상이면 초고령사회로 보고 있다. 우리나라는 추계인구기준으로 2011년 기준 국내 총 인구대비 65세 이상의 인구비중이 11.3%로 고령화사회에 이미 진입하였고, 2026년에는 20.83%, 2037에는 30.13%로 초고령사회로 진입할 것으로 예상되고 있다.

〈그림 5〉 생산가능 인구비율 및 고령인구 추계



자료 : 통계청, e-나라지표(2014), <http://kostat.go.kr/>, <http://kosis.kr/>.

박형수·홍승현(2011)은 선진국의 경우 고령화사회인 65세 인구비중이 7% 이상에서 초고령사회인 20%이상으로 진입하는데 캐나다 79년, 일본 36년, 독일 76년, 이탈리아 80년, 미국 89년, 프랑스가 154년이 소요되는 것에 비해, 우리나라의 경우 약 27년 내로 초고령사회에 진입하게 될 것으로 전망되므로 매우 빠른 고령화추세를 나타낸다고 제시하였다.

따라서 급격한 고령화로 인한 생활비 등의 조달을 목적으로 주택을 매각하

고자 하는 고령층에 비해서 신규수요층의 진입이 많지 않을 경우 주택시장침체 및 주택담보대출 부실화로 전이될 수 있으므로 주택담보대출에 대한 리스크관리가 필요하다고 판단된다.

〈표 10〉 국가별 고령화 진행현황

구 분	도달연도			소요연수		
	*7%	*14%	*20%	7%→14%	14→20%	7%→20%
우리나라 (통계청)	1999년 (2000년)	2018년 (2018년)	2026년 (2026년)	19년 (18년)	8년 (8년)	27년 (26년)
캐나다	1945년	2010년	2024년	65년	14년	79년
프랑스	1864년	1978년	2018년	114년	40년	154년
독일	1932년	1972년	2008년	40년	36년	76년
이탈리아	1927년	1988년	2007년	61년	19년	80년
일본	1970년	1995년	2006년	25년	11년	36년
영국	n.a	1975년	2028년	n.a	53년	n.a
미국	1942년	2014년	2031년	72년	17년	89년
전세계	2002년	2039년	n.a	37년	n.a	n.a

\*주) UN은 65세이상 인구비중에 따라 고령화사회(Ageing society : 고령인구비율 7% 이상), 고령사회(Aged society : 고령인구비율 14%이상), 초고령화사회(Post-aged society : 고령인구비율 20%이상)로 분류함.

자료 : 박형수·홍승현(2011), 『고령화 및 인구감소가 재정에 미치는 영향』, 한국조세연구원, p.48.

## 2) 사전적 위험관리

주택담보대출의 사전적 위험관리방안은 주택담보대출의 취득시점에서 위험을 관리하는 방안이다. 일반적으로 대출심사 과정에서 위험관리가 이루어지게 되며 LTV와 DTI를 통해 대출금의 적정규모를 산정한 후 대출이 이루어져야 한다.

### 가) 담보인정비율

담보인정비율(LTV : loan to value ratio)은 주택담보대출 취급시 담보가치에 대한 대출취급가능금액 비율을 말한다. 주택담보가치 대비 대출금액 비율을 의미하는 것으로 담보대출을 취급하는 하나의 기준이다. 금융기관은 대출채권에서 부도가 발생하는 경우 담보자산을 처분하여 대출채권 상환에 충당하며, 이때 대출채권 상환에 부족분이 발생하지 않도록 일정의 담보인정비율 이내에서 담보대출을 취급하고 있다.<sup>20)</sup>

일반적으로 금융기관 경매 처분시에는 담보자산의 시가보다 낮은 수준에서 낙찰가격이 결정된다는 점을 고려하고 있으며, 부동산가격 하락시에도 대출채권 상환에 부족분이 발생하지 않도록 하기 위해 금융기관은 담보인정비율을 자체적인 기준보다 낮은 수준에서 운용되고 있다.<sup>21)</sup>

주택의 종류 및 주택의 소재지역에 따라 담보자산의 시가대비 처분가액 비율이 달라지므로, 주택담보대출에 대해서는 주택의 종류 및 담보가 소재한 지역에 따라 담보인정비율을 차등하여 적용하고 있으며, 담보인정비율을 다음과 같이 산정하도록 하고 있다.

$$LTV = (\text{주택담보대출금액} + \text{선순위채권} + \text{임차보증금 및 최우선변제 소액 임차보증금}) \div \text{주택가격(담보가치)}^{22)}$$

이렇게 산정된 담보인정비율에 대해서 정부에서는 지역과 만기 등에 따라서 차등화 된 기준을 적용하고 있다. 현재는 투기지역과 투기과열지구에 지정된 지역이 없으므로 60~70% 수준의 LTV가 적용되고 있다. 하지만 앞서 살펴본 바와 같이 최근 주택시장 침체 및 경매 매각가율 하락으로 금융기관 자체적인 리스크 분석에 대한 필요성이 높아지고 있다.

20) 전국은행연합회, 전계서, p.8.

21) 전국은행연합회, 전계서, p.8.

22) 상계서, p.8.

〈표 11〉 금융기관의 주택담보대출 담보인정비율

만 기		투기지역		투기과열지구		기타지역	
		주택 <sup>1)</sup>	아파트	주택 <sup>1)</sup>	아파트	주택 <sup>1)</sup>	아파트
3년 이하 <sup>2)</sup>		50%이내	40%이내 <sup>2)</sup>	50%이내	50%이내 <sup>2)</sup>	60%이내	60%이내 <sup>2)</sup>
3년 초과~10년 이하		60%"	40%"	60%"	60%"	60%"	60%"
10년 초과	담보가액 6억원 초과	60%"	40%"	60%"	60%"	60%"	60%"
	담보가액 6억원 이내	60%"	60%"	60%"	60%"	60%"	60%"
10년 이상 분할상환 <sup>3)</sup>		70%"	70%"	70%"	70%"	70%"	70%"
2014.8.1변경		70% 4)					

주 : 1) 아파트(주상복합아파트 포함)를 제외한 주택을 말함.

2) 아파트 중도금대출 및 이주비대출의 경우 만기 구분 없이 해당비율 적용.

3) 1년 이내에 한국주택금융공사에 매각하거나 자체 유통화계획이 있는 고정금리부 대출.  
(거치기간 1년 이내, 총부채상환비율 40% 이하)을 말함.

4) 만기, 금액, 지역에 관계없이 일괄적용.

자료 : 금융감독원, 은행업감독규정 제29조의2(주택담보대출에 대한 리스크관리), 은행업감독  
업무시행세칙 제18조의2(별표 18 : 주택담보대출에 대한 리스크관리 세부기준, 2013.12.30).

#### 나) 총부채상환비율

총부채상환비율(DTI : debt to income ratio)은 연간소득대비 연간 대출원리금 상환액의 비율이다.<sup>23)</sup> 금융감독원은 연간 대출원리금의 합계액이 연간 가계소득의 비율 이내가 되도록 채무수준을 정하였으며, 금융기관들은 자체적인 개인별 신용평가 기준에 따라, 연간 대출원리금이 연간소득의 DTI 비율 이내가 되도록 담보대출한도를 정하고 있다. DTI는 2005년 주택가격이 급등하면서 정부는 부동산 가격안정을 위해 LTV 규제 강화와 함께 DTI 규제를 도입하였으며, 총부채상환비율은 다음과 같이 산정하도록 하고 있다.

$$DTI = (\text{해당 주택담보대출 연간 원리금 상환액} + \text{기타부채의 연간 이자 상환액}) \div \text{연소득}^{24)}$$

23) 전국은행연합회, 전계서, p.5.

24) 상계서, p.5.

DTI는 LTV와 달리 세분화된 규제내용을 가지고 있다. 주요내용을 살펴보면, 규제대상지역을 투기지역, 투기지역을 제외한 서울지역, 인천·경기지역으로 크게 구분하였고, 차입대상자가 희망하는 대출금액 기준은 1억원 초과와 담보대출을 기준으로 주택의 종류는 신규분양 아파트와 재고주택으로 한정하였다.

〈표 12〉 금융기관의 주택담보대출 DTI 규제내용

구 분		내 용		
대상주택	주택유형	아파트		
	주택종류	- 신규분양주택, 재고주택 - 규제제외 : 집단대출(이주비, 중도금, 잔금), 미분양 담보대출, 기금대출		
대상지역	수도권	투기지역	서울 (투기지역제외)	인천·경기
대출금액	1억원 초과 담보대출	40%	50%	60%
		60% 주) - 고정금리와 1년이내의 분할상환상품을 선택하면 +10%p 가산비율이 추가로 가능함		

주) 2014년 8월 1일 현재.

#### 다) LTV와 DTI 규제의 변천과정

우리나라에서는 LTV와 DTI 등 주택담보대출 리스크관리를 위한 제도 도입이 상대적으로 늦게 나타났다. 본격적으로 LTV가 도입된 것은 2002년이었으며, DTI는 2005년에 도입되어 주택시장 상황에 따라서 강화와 완화가 반복되었다. 2009년 이후에는 전반적으로 주택시장에서 규제완화가 이루어졌으나 가계부채 문제가 심각해지면서 규제가 강화되기도 하였다.

LTV와 DTI 규제는 주택시장 과열방지로 시작되었지만 그 결과로 2008년 금융위기에 도 주택담보대출은 큰 위험요인이 발생하지 않는 부수적 효과가 나타났다. 즉, LTV의 도입으로 담보자산의 가치를 유지하는 데 큰 어려움이 없었으며, DTI의 도입으로 주택담보대출의 상환불이행 가능성을 조정하는 데

도 중요한 역할을 하였다.

〈표 13〉 LTV와 DTI 규제 변천과정

규제일자	규제강화 및 해제 내용	
2002년 9월	LTV 도입	투기과열지구 LTV 60%
2002년 10월	LTV 강화	전 지역 확대적용 LTV 60%
2003년 6월	LTV 강화	투기지역 만기 3년 이내 LTV 60 → 50%
2003년 10월	LTV 강화	투기지역 만기 10년 이내 LTV 50 → 40%
2005년 6월	LTV 강화	투기지역 만기 10년 이내 및 6억원 초과 LTV 50 → 40%
2005년 8월	DTI 도입	투기지역 DTI 40% 적용
2006년 3월	DTI 강화	투기지역 6억원 초과 주택 소유권 3개월 이내 DTI 40%
2006년 11월	LTV 강화	비은행 투기지역 6억원 초과 LTV 60 → 50%
2007년 2월	DTI 강화	감독규정 규제대상 대출 6억원 이하 아파트에 대해서도 DTI 40~60%를 적용
2007년 8월	DTI 강화	감독규정 규제대상 외 대출 6억원 이하 아파트에 대해서도 DTI 40~70%를 적용
2008년 11월	투기지역해제	강남3구(강남구, 서초구, 송파구)를 제외한 전지역
2009년 7월	LTV 강화	투기지역 6억원 초과 아파트 LTV 60 → 50%
2009년 9월	DTI 강화	DTI 수도권도입, 서울 DTI 50%, 인천·경기 DTI 60%
2009년 10월	LTV 강화 DTI 강화	LTV, DTI 비은행권으로 적용확대, LTV 60~70% → 50~60%, DTI 확대적용(투기지역 → 수도권 비투기지역)
2010년 4월	DTI 완화	기존 주택구매시 LTV한도, DTI 한도를 초과해 대출가능 하도록 주택 금융공사 보증지원
2010년 8월	DTI 완화	투기지역제 외 DTI 생략가능 금융권 규제자율화 시행(무주택, 1가구1주택) : 2010.8~2011.3
2011년 4월	DTI 강화	DTI 규제 부활
2012년 5월	투기지역해제	강남3구(강남구, 서초구, 송파구), LTV, DTI는 40 → 50%, 강남3구 해제로 투기지역은 한 곳도 없음
2014년 8월	LTV 완화 DTI 완화	LTV비율 : 지역별, 기간별, 가격별, 담보별 구분비율 50 ~ 60% → 70%로 통일 DTI비율 : 서울 50%, 경기/인천 60% → 60%로 통일 단, DTI는 대출방법 선택으로 가산 비율 활용 시 최고 70%까지 가능 (고정금리 : +5%p, 1년 이내 분할상환 : +5%p = +10%p)

자료 : 금융감독원 보도자료, 금융기관 주택담보대출 내용정리.

라) 금융기관 DTI 규제 운용현황

금융감독원은 금융기관의 주택담보대출 취급시 차입자의 채무상환능력 확인관행 정착을 위하여 획일적인 DTI 규제보다는 금융기관이 자율적으로 규제의 기본틀을 유지하면서 채무상환능력평가 기준을 마련하여 변경하도록 유도하는 정책을 펼쳤으며 연구조사 결과 규모가 큰 시중금융기관들은 <표 14>와 같이 당초의 DTI 규제보다 적용범위가 확대되었다.

이러한 채무상환능력평가 적용의 취지는 차입자의 상환능력을 감안하여 부채를 조달하도록 함으로써 차입자의 가계는 물론 금융기관과 금융시장 등 우리경제의 건전한 성장과 발전을 도모하기 위한 금융의 기본원칙<sup>25)</sup>이라고 볼 수 있어 빠른 기간내 여신전문기관까지 정착되어져야 할 것으로 본다.

<표 14> DTI 규제와 금융기관 채무상환능력평가 적용

구 분	DTI 규제	채무상환능력평가 적용
대상지역	투기지역, 서울, 인천·경기	전국 모든지역
평가대상	아파트, 신규분양 및 기존주택	주택담보대출 신규대상
담보물건	아파트	모든주택 (아파트, 연립, 다세대, 단독, 다가구, 주상복합)에 적용
평가기준	대출금액 1억원 초과시	기준금액에 관계없이 채무상환능력평가

자료 : 시중 금융기관 조사자료, 2014. 11월 현재.

주택담보대출을 받는 일반인에게 있어 DTI라는 뜻도 익숙하지 않지만 그 내용도 복잡하여 쉽게 이해 할 수 없으며, 대출실행 전에 대출한도와 금리도 계산하기가 어려워서 대출접수 후에 금융기관의 시스템에 의존하는 실정에 있다. 이러한 점을 고려하여 금융감독원에서는 <표 15> DTI 계산기 화면을

25) 금융감독위원회, 금융감독원(2012), 『총부채상환비율(DTI) 규제 보완방안』, p.1.

만들어 운용하고 있으나 대출실행 전에 정확한 계산을 하기에는 부족한 화면으로 보여진다. 또한 금융기관 채무능력상환평가 시스템은 규모가 있는 시중은행에서는 운용되고 있지만 저축은행 등에는 이러한 시스템이 미미한 것으로 알려지고 있다.

주택담보대출을 취급하는 해당 금융기관에서 <표 15>와 같은 계산화면을 개인신용 평가하는 CSS평가<sup>26)</sup> 제도와 연계하여 적용하는 노력을 기울인다면 차입자는 상시적으로 주택담보대출에 대한 상환능력과 금리와 이율 등을 알 수 있다고 본다.

<표 15> DTI 계산기

부동산시가	( )만원	대출예상금액	( )만원
연간소득	( )만원	기존대출원금	( )만원
신용대출이자율	( )%	기존대출이자율	( )%
평형	( )85m <sup>3</sup> 미만 ( )85m <sup>3</sup> 이상		
신용도	( )상 ( )중 ( )하		
<div style="border: 1px solid black; padding: 5px; display: inline-block;">계산하기</div>			
귀하의 DTI 는 ( )%이며 원리금균등상환조건에서 대출가능액은	( )만원입니다(10년 상환기준)		
	( )만원입니다(20년 상환기준)		
	( )만원입니다(30년 상환기준)		

자료 : 금융감독원, <http://www.fss.or.kr/>, 가계 재무관리 화면, DTI계산기, 2014년 11월 현재.

26) CSS란 Credit Scoring System의 줄임말로 개인의 직업과 소득수준, 자산정도, 신용상태, 금융기관 거래정보 등을 종합 평가하고 금융기관별 신용등급에 따라 금리 및 대출 여부를 결정해주는 전산 시스템을 말한다.

### 3) 사후적 위험관리

주택담보대출의 사후적 위험관리방안은 이미 채무불이행이 발생한 경우 금융기관이 피해를 최소화하기 위한 방안으로 볼 수 있다. 현재 금융기관에서 사용하고 있는 방안으로는 주택담보대출의 기한의 이익상실제도와 연체관리가 있다.

#### 가) 주택담보대출 기한의 이익상실

주택담보대출 기한의 이익상실제도는 금융기관이 거래처에 대한 기한이익을 상실시키는 경우로, 법률의 규정 또는 당사자 간의 특정에 의해 일정한 사유가 발생함으로써 당연히 상실하는 경우와, 채권자가 이행을 청구함으로써 상실하는 경우가 있다.<sup>27)</sup>

주택담보대출의 경우는 약정에 의한 기한의 이익상실로서, 금융기관거래에서 여러 이유로 대출시 정한 기한보다도 빨리 변제기를 도래시켜 상계 또는 담보권실행을 하여야 할 경우가 있게 된다.<sup>28)</sup>

금융기관은 주택담보대출 취급시 여신거래기본약관이나 개별약정서 등에 기한이익상실에 관한 내용을 정해두고 있으며 이자지급 연체사항은 기한이익당연 상실사유 중 하나의 규정이다. 이자지급을 13일을 초과하여 지체하거나, 여신기간(기한연장의 경우 연장기간을 포함)중 이자지급의 지체회수가 4회에 달한 때 및 분할상환금 또는 분할상환원리금 지체회수가 2회에 달한 때의 규정이 적용되고 있다.<sup>29)</sup>

표준약관 변경전 차입자는 주택담보대출이 실행이 되고 이날로부터 1달이 경과되는 시점 이전에 이자후취방식에 의한 방법으로 이자지급 약정일에 1회 차를 상환해야하며, 이때 상환되지 않으면 매월 상환해야하는 이자에 대한 지연배상금 이자가 발생된다. 즉 매월 상환되는 이자에 대한 지연이자만 발생되

27) 행정규칙, 대부거래 표준약관 제12조, 금융기관 여신거래기본약관, 기한도래 전 채무변제 의무 내용 참조.

28) 상계서, 행정규칙, 대부거래 표준약관.

29) 공정거래위원회, 은행여신거래기본약관(가계용), 표준약관, 제10006호(2013.12.30, 개정) 참조.

는 것이다. 그리고 이자지급 약정일에 이어서 다시 1개월이 연체되면 분할상환금 또는 분할상환원리금 지체회수가 2회가 되어 기한의 이익상실사유가 발생이 되며 주택담보대출 전체잔액 대출액에 대한 이자를 내야 했다.

그러나 공정거래위원회는 2013년 12월 주택담보대출 차입자에게 부담이 되었던 기한의 이익상실 기준을 연장하고, 담보가치 하락시 금융기관에서 요구하는 담보물 보충 청구권요건을 강화하여 차입자의 권익이 증진되도록 ‘은행여신거래기본약관(가계용)’을 개정하였다. 이에 금융위원회는 금융기관에게 2014년 4월 1일부터 개정된 약관을 적용하도록 하였다.

개정된 약관에서는 기한의 이익상실 기간을 연체미납후 1개월에서 2개월로 늘렸으며, 금융기관이 고객에게 기한의 이익상실을 통보하는 기간도 종전 3영업일 전 통지함에서 7영업일 전까지 대출잔액 전부에 연체료가 부과된다는 사실을 반드시 서면으로 통지토록 변경하였다. 또한 담보가치가 하락했더라도 채무자 또는 보증인의 책임이 있는 사유로 신용이 악화되거나 담보물 가액감소가 현저한 경우에만 추가 담보제공을 요구할 수 있도록 개정하였다.

〈표 16〉 주택담보대출 기한의 이익상실 시기

구 분	2014. 3. 31 이전	2014. 4. 1. 이후
기한의 이익상실 시기	이자지급 약정일부터 1개월경과 (분할상환금 2회 연체)	이자지급 약정일부터 2개월경과 (분할상환금 3회 연체)
기한의 이익상실 前 사전통지기간	기한이익 상실일 前 3영업일까지 도달	기한이익 상실일 前 7영업일까지 도달 (채무이행 지체사실 및 대출잔액 전부에 연체료가 부과 될 수 있 다는 사실을 반드시 서면으로 통지토록 함)
은행의담보물 보충청구권 행사요건	-	담보가치가 하락한 경우 고객의 책임유무를 불문하고, 대출기간 중 추가 담보제공 요구 가능한 것을 고객의 책임있는 사유로 신용악화 와 담보가치 감소가 현저한 경우 에만 추가 담보제공 요구가 가능 하도록 함

The diagram illustrates the change in the interest loss period for housing loans. It shows a timeline with two main periods. The first period, labeled '종전' (End), shows that interest loss begins 1 month after the due date (2 missed payments) and continues until 3 business days before the loss period ends. The second period, labeled '변경 후' (After Change), shows that interest loss begins 2 months after the due date (3 missed payments) and continues until 7 business days before the loss period ends. Below the timeline, three boxes specify the interest payment terms: '이자지급 약정일 이자미납부시 (1회연체)', '이자지급 약정일 + 1개월경과 (2회연체)', and '이자지급 약정일 + 2개월경과 (3회연체)'. Arrows indicate the duration of '기한의 이익상실' (Interest Loss Period) for each scenario.

자료 : 공정거래위원회, 은행여신거래기본약관(가계용), 표준약관, 제10006호(2013,12,30, 개정).

## 나) 주택담보대출 연체관리

주택담보대출 차주의 연체가 현실화되면 금융기관은 정해진 절차에 따라서 연체관리에 들어간다. 연체관리는 금융기관뿐만 아니라 금융기관의 건전성관리를 책임지고 있는 정부에서도 중요한 관리대상이 된다. 정부는 자산의 건전

성을 확보하기 위하여 IMF와의 합의하에 1999년 12월말부터 차주의 과거 원리금 상환실적 뿐만 아니라 미래 채무상환능력을 반영하여 자산건전성을 분류하는 FLC(forward looking criteria)에 의한 자산건전성 분류기준<sup>30)</sup>을 마련하여 시행하고 있다.

주택담보대출 여신관리도 자산건전성 분류를 하게 되며 주택담보의 건전성 정도에 따라 5단계로 나누어 관리를 하게 되며, 금융기관에서는 이러한 기준에 따라 영업점과 연체관리를 집중하는 후선센터에서 단계별로 관리를 진행한다.<sup>31)</sup>

첫번째 단계인 정상이란 신용상태가 양호한 대출금으로 연체기간이 1개월 미만인 초기연체의 경우다. 두번째 요주의단계는 연체기간이 3개월 미만으로 현재는 원리금 회수에 문제가 없으나 앞으로는 신용상태가 악화될 가능성이 있어 세심한 주의나 사후관리가 필요한 대출금을 가리키며, 세번째 고정단계는 연체기간이 3개월 이상으로 차입자의 신용상태가 이미 악화돼 채권회수에 상당한 위험이 발생한 것으로 판단되는 대출금과 회수의문 또는 추정손실 대출금 중 회수할 수 있는 예상금액을 말하고 있다. 네번째 회수의문단계는 연체기간이 3개월 이상 1년 미만이면서 차입자의 채무상환 능력이 현저하게 악화돼 채권회수에 심각한 위험이 발생한 대출금 중 회수예상금액을 초과하는 대출금을 가리킨다. 마지막으로 추정손실 단계는 연체기간이 1년 이상으로 상환능력이 심각하게 나빠져 손실처리가 불가피한 대출금 중 회수예상금액을 초과하는 부분을 말한다.<sup>32)</sup>

고정여신(sub-standard) 이하의 단계인 고정, 회수의문, 추정손실(연체 1년 이상)을 부실채권이라 하고 이를 고정이하 여신이라 부른다. 이러한 고정이하 여신의 비율은 금융기관 자산건전성을 평가하는 지표 중 하나이다.

---

30) 은행업감독규정 27조 1항, 은행은 정기적으로 차주의 채무상환능력과 금융거래내용 등을 감안하여 보유자산의 건전성을 “정상”, “요주의”, “고정”, “회수의문” 및 “추정손실”의 5단계로 분류하고, 적정한 수준의 대손충당금(지급보증 충당금 및 미사용 약정충당금 포함)을 적립·유지해야 한다.

31) 은행업감독규정, <별표 3> 자산건전성 분류기준, 자산건전성 분류단계별 정의 참조.

32) 상계서, 은행업감독규정 <별표 3>.

〈표 17〉 주택담보대출의 연체기간에 따른 건전성 분류

구 분		주택담보대출 연체기간	건전성 내용
정상채권	정 상	1월 미만	신용상태 양호 대출금
	요 주 의	1월 이상 3월 미만	신용상태 악화 가능성
부실채권	고 정	3월 이상	채권회수 위험발생 (회수예상가액 해당부분)
	회수의문	3월 이상 12월 미만	심각한 회수위험 (회수예상가액 초과부분)
	추정손실	12월 이상	손실이 불가피함 (회수예상가액 초과부분)

자료 : 금융기관 조사자료.

주택담보대출의 연체관리는 금융기관별로 다를 수 있겠으나, 콜센터를 기반으로 한 후선연체관리센터가 마련된 금융기관은 체계적인 연체관리를 하고 있다. 지동현(2007)도 ABC 금융기관의 사례를 통하여 금융기관의 주택담보대출 초기 사후관리는 회수노력이 주종을 이룬다고 하였으며, 3개월 미만의 단기연체고객에 대해서는 영업점 또는 후선관리센터에서 관리를 하며, 주로 콜센터를 기반으로 한 ACS(auto call system)센터를 통하여 전화안내를 하고, 영업점은 방문과 DM발송 등의 방법으로 관리를 하고 있다고 하였다.

영업점과 후선센터의 회수노력에도 불구하고 연체가 지속된 3개월 이상의 장기연체대출은 고정여신으로 분류되고 내부이전을 통해 내부이전가격과 채권서류까지도 여신관리 전담부서로 이관이 된다. 구체적인 연체관리 절차는 아래 〈표 18〉과 같으며 총 8단계의 관리절차로 진행된다.

〈표 18〉 시중은행의 주택담보대출 단계별 연체관리

연체관리 단계		내 용
1단계	유선독촉 DM(direct marketing) SMS(short message service)	대출이 실행되고 납입기일 이후 연체자로 분류가 되면 유선안내와 DM, SMS를 병행하여 연체관리를 함
2단계	서면독촉	유선독촉과 병행한 연체관리로 잔액연체 또는 연체 장기화될 때
3단계	심방독촉	유선 및 서면독촉에도 연체가 지속될 때 이거나 채권관리 시 필요에 따라서 함
4단계	재산조사	채무자의 직장 및 자택 또는 과세지 물건 재산조사
5단계	가압류	채권자인 금융기관이 금전채권을 확보하기 위하여 부동산 또는 급여 등을 임시로 가압류를 함
6단계	집행권원 신청	강제집행의 사전단계에서 집행력의 근거인 채무이행판결문 등을 신청하고, 법원을 통해 주택담보대출 채권 채무 사실을 확인받고 해당 신청서를 채무자에게 보내 상환을 요구하는 지급명령 신청을 함
7단계	담보권 실행	담보권을 실행하기 위한 임의경매를 신청하고, 채권자인 금융기관은 담보권을 행사하여 경매로 매각한 후 그 매각대금에서 다른 채권자보다 우선하여 주택담보대출 채권을 회수하는 강제집행 절차를 함
8단계	강제경매 신청	채권압류 및 추심단계로서, 집행권원을 가진 채권자인 금융기관은 증명 서류를 갖추어 주택담보대출이 취급된 물건 소재지의 관할 지방법원에 경매를 신청함

금융기관은 내부 부실율을 주택담보대출 사후관리 측면에서 기업정보로 다루고 있어 상세하게 알 수 없었으나 초기에 공통적으로 관리하는 방법은 〈표 19〉와 같이 전이억제율과 회수율 등으로 추정된 자료이다. 주택담보대출 연체관리 방법에는 여러 가지 형태가 있을 수 있으며 관리지표로 전이억제율 또는 전이율과 회수율의 지표가 보편적으로 사용되고 있다. 전이 억제율이란

주택담보대출 신규 연체발생 후 10일 이상 또는 1개월 이상 등의 기간을 정하여 연체로 전이되지 않는 비율이며, 회수율이란 연체발생 된 관리채권의 회수금액 비율을 말한다.

〈표 19〉는 전이 억제율과 회수율을 통하여 전이율을 추정한 내용으로 주택담보대출 신규취급 후 처음 연체발생 10일 시점에서는 5.67%가 전이되어 연체를 하였으며, 이 연체 대상자를 포함한 1개월 시점에서 연체자는 10.34%가 연체하고 0.60%가 전이된 것으로 추정하였다. 또한 1개월 이상 연체대상채권은 40.37%의 회수가 이루어져 1개월 시점의 장기연체 전이율은 0.34%라고 추정해 볼 수 있다.

〈표 19〉 주택담보대출의 전이억제율과 회수율

전이 억제율		회수율
10일 이상	1개월 이상	1개월 이상
94.33%	89.67%	40.37% ③
평균 연체전이율 5.67% ①	평균 연체전이율 10.34% ②	
①5.67% × ②10.34% × ③(1-40.4) = 0.34%		
* 1개월 이상 장기연체 전이율은 0.34%로 추산		

자료 : 금융기관 조사자료, 자료의 기간은 2013년 평균.

주택담보대출의 장기적인 부실정도를 파악하려면 금융기관의 연체율뿐만 아니라 부실채권으로 분류되는 고정이하의 비율이 중요할 수 있다. 금융기관의 주택담보대출 연체율은 매일 금융감독원에 보고가 이루어지고 있으나, 고정이하 비율의 자료는 거의 없는 실정이다.

지동현(2007)은 ABC은행의 사례를 통하여 2004년도 주택담보대출의 최초 평균연체발생률 1.96%의 수치를 가지고, 3개월 미만까지 회수되지 않고 고정이하로 남아있는 연체자의 비율을 3.64%로 추정한 바 있다. 최초 연체발생 1개월 이내에 60.5%가 회수되고, 1~2개월 기간에 55.7%, 2~3개월 기간에 28.4%가 회수된다고 보았을 때, 연체 후 3개월 이전에 84.5%가 회수되므로서, 3개월 이전에 회수되지 않고 3개월 이후에도 연체로 남아 있는 3.64%의

연체비율을 추정하였다.<sup>33)</sup> 그러나 위의 사례를 <표 20>을 통하여 회수기간을 구분하고 회수율을 분석해 본 결과 2.94% 고정이하 연체비율의 값을 구하였다. 약간 상이한 값을 구하였으나, 연체율이 정해지고 구간별 회수율만 구하면 고정이하 연체비율을 추정할 수 있었다.

부실채권이 고정이하로 편입되면 연체채권은 내부이전되고 여신관리전담부서는 단계별 연체관리를 하게 된다. 최종 경매진행 전에도 지속적인 회수노력을 하고 있어 연체의 비율은 감소할 수 있다.

<표 20> 주택담보대출 고정이하 연체비율 추정 산식

회수기간	회수율	고정이하 연체비율 추정내용 (평균 연체발생을 1.96% 가정)
1개월 이내	60.5%	$100\% - 60.5\% = 39.5\%$
1~2개월 이내	55.7%	$39.5\% \times 55.7\% = 22.0\%$
2~3개월 이내	28.4%	$100\% - (60.5\% + 22\%) = 17.5\% \times 28.4\% = 4.97\%$
고정이하 연체비율		$= (60.5\% + 22.0\% + 4.97\%) = 87.47\%$ $= 1.96\% \times (1 - 87.4\%) \times 12$ $= 2.94\%$

자료 : 지동현(2007), 『주택담보대출의 리스크관리』, KB국민은행연구소, p.57.

주택담보대출의 손실률을 추정한 방두완·박연우(2012)는“국내 주택담보대출 LGD의 기본특성과 결정요인 분석”에서 주택담보대출 LGD(loss given default ; 부도대출 손실률)을 실증분석한 결과 아파트담보물건의 평균 LGD는 30%라고 추정하였다. 이것은 부실이 발생된 채무자 아파트의 담보가치(LTV)가 70%라는 의미로 해석할 수 있어서, 아파트의 가격이 10%이상 하락하고 LTV 60%로 취급된 대출이 부실이 된다면, 금융기관은 원금을 100% 회수할 수 없는 사례가 발생할 수 있다. 그리고 3개월 이내의 고정이하 연체가 장기화되어 비율이 점차 증가할수록 주택담보대출 부실 가능성은 높아진다고 볼 수 있다.

33) 추정연체비율 공식 =  $1.96(1 - 84.5) \times 12 = 3.64\%$ .

### 제 3 절 국·내외의 가계부채 대응정책

#### 1) 우리나라의 가계부채 대응정책

##### 가) 소득 대비 부채 비율 관리

우리나라의 가처분소득 대비 가계부채 비율은 2012년말 현재 163.8%로 서브프라임 모기지가 발생했던 미국보다 50%p OECD평균보다 30%p 가량 높은 수준이다. 정부는 2014년 2월 가계부채는 오랜기간 누적된 구조적 위험으로 기업부채와 달리 단기간 해결이 어려운 문제이며, 시스템 리스크 완화를 위한 단기대책과 함께 중장기적 구조개선 노력을 병행 추진하겠다고 발표했다. “경제혁신 3개년 계획”에서 제시한 바와 같이 가계 “소득 대비 부채비율”을 가계부채 핵심 관리지표로 설정하였고, 2017년까지 동 비율을 현재보다 5%p 인하된 수준으로 관리하는 것을 목표로 가계소득 개선, 부채부담 경감 등의 내용을 발표하였다.<sup>34)</sup>

〈표 21〉 국가별 가처분소득 대비 가계부채 비율

(단위 : %)

한국	스웨덴	캐나다	미국	OECD평균
163.8	172.1	154.3	114.9	134.8

참고 : 2012년 말 기준.

자료 : 관계부처합동, 금융위원회(2014), 『가계부채 구조개선 촉진방안』, p.4.

또한 정부는 “경제혁신 3개년 계획” 후속조치로 “가계부채 구조개선 촉진방안”을 발표하였으며 가계건전성 종합대응의 일환으로 자산, 부채, 소득, 지출의 4가지 방향 구조개선안을 발표하였고 이중 우선적으로 부채부분의 구조개선을 추진하는 내용을 발표하였다.

34) 관계부처합동, 금융위원회(2014), 『가계부채 구조개선 촉진방안』.

〈표 22〉 정부의 가계부채 구조개선 촉진방안

자 산	주택시장 활성화, 실물자산 유동화, 금융자산 축적지원
부 채	증가속도 관리, 대출구조의 개선, 취약계층 채무 재조정
소 득	가계소득 개선, 자영업 경쟁력 강화, 일자리 창출
지 출	주거 및 교육비의 가계지출 합리화, 비 소비지출 부담 경감

자료 : 관계부처합동, 금융위원회(2014), 『가계부채 구조개선 촉진방안』, p.4.

나) 고정금리, 비거치식 분할상환 비중 확대

대출구조의 개선으로 상환구조를 개선하겠다는 내용이며 은행권의 장기 및 분할상환대출 비중을 확대하고 고정금리대출 비중을 확대하겠다는 내용이다.

기존 30~50%의 주택담보대출 BIS 위험가중치를 전액 만기연장한 일시상환대출에 대해 70%까지 상향하고 분할상환 및 고정금리 상품판매 비중의 목표를 2016년까지 30%확대에서 2017년까지 40%까지 확대한다는 내용이다.

〈표 23〉 정부의 은행권 주택담보대출 구성비 목표

(단위 : %)

구 분	분 할		비분할		계	
	2013년	2017년	2013년	2017년	2013년	2017년
고정금리	6.6	15	9.3	25	15.9	40
변동금리	12.1	25	72	35	84.1	60
계	18.7	40	81.3	60	100	100

자료 : 관계부처합동, 금융위원회(2014), 『가계부채 구조개선 촉진방안』, p.7.

또한 보험권 및 상호금융권도 비거치식 분할상환 비중을 확대하는 목표를 세웠으며, 보험권은 40%까지 상호금융권은 15%까지의 구성비 목표를 정하였다.

〈표 24〉 정부의 보험권 및 상호금융권 비거치식 분할상환 목표  
(단위 : %)

구 분	비거치식 분할상환	
	2013년	2017년
보험권	26.1	40
상호금융권	2.0	15
계	18.7	40

자료 : 관계부처합동, 금융위원회(2014), 『가계부채 구조개선 촉진방안』, p.7.

정부는 주택담보대출 만기기간을 소득공제 확대를 통하여 장기로 유도하며, 준고정금리 대출상품 및 금리상한 대출상품 등 다양하게 출시하여 변동금리 상품의 선택을 증가시켰다. 또한 MBS 시장 활성화를 통하여 현재 고정금리 대출상품의 추가적인 금리인하 여력을 확보한다는 내용이다.

〈표 25〉 주택담보대출 구조개선 확대방안

주택담보대출 구조개선 확대방안	내 용
주택담보대출 소득공제 확대	<ul style="list-style-type: none"> <li>- 만기기간 15년 이상 : 소득공제 한도 확대 (최대 1,500만원 → 1,800만원으로 상향)</li> <li>- 만기기간 10~15년 : 소득공제 혜택 신규 부여 (한도 추후 결정)</li> </ul>
다양한 준고정금리 대출상품 출시	<ul style="list-style-type: none"> <li>- 금리변동주기 5년 이상인 대출 출시</li> <li>- 금리상한대출 출시 (대출 후 금리상승폭을 일정 수준으로 제한 → 소비자들의 금리변동 위험 완화됨)</li> </ul>
MBS 시장 활성화	<ul style="list-style-type: none"> <li>- 한국주택금융공사의 주택저당증권(MBS: Mortgage Backed Securities)을 한국은행 공개시장조작 대상증권 (환매조건부매매)으로 편입 → MBS의 신인도가 높아지고 국·내외 투자수요확대 됨 → MBS 발행금리가 낮아짐 → 고정금리대출 금리의 추가인하 여력이 확보됨 → 주택담보대출의 이자 부담이 완화됨</li> </ul>

자료 : 관계부처합동, 금융위원회(2014), 『가계부채 구조개선 촉진방안』, pp.11-12.

다) 취약계층 채무조정

고위험·고금리 대출을 이용 중인 취약차주의 대출구조를 개선하는 내용으로 금리부담 경감을 지원하기 위하여 영세 자영업자의 바뀔드림론 금리 인하와 대출한도를 상향 조정하였고 고금리대출 지원대상을 확대하였다. 또한 제2금융권의 가계대출 건전성 제고, 구조개선 등을 추진하고 취약계층의 제2금융권 주택담보대출 구조를 개선하기 위하여 단기적이며 일시상환 대출을 장기적이며 분할상환 대출로 전환하는 시범사업을 실시하였다.

〈표 26〉 주택담보대출 취약계층 채무조정방안

취약계층 채무조정방안	내 용
영세 자영업자 바뀔드림론 활성화	- 영세 자영업자 바뀔드림론 금리 인하, 대출한도 상향 조정 - 고금리대출 지원대상 확대(연 20% 이상→15% 이상)
취약계층 제2금융권 주택담보대출 구조개선	- 단기, 일시상환 대출 → 장기, 분할상환 대출로 전환 하는 시범사업 실시 (2014년 중 1,000억원 규모)

자료 : 관계부처합동, 금융위원회(2014), 『가계부채 구조개선 촉진방안』, pp.8-9.

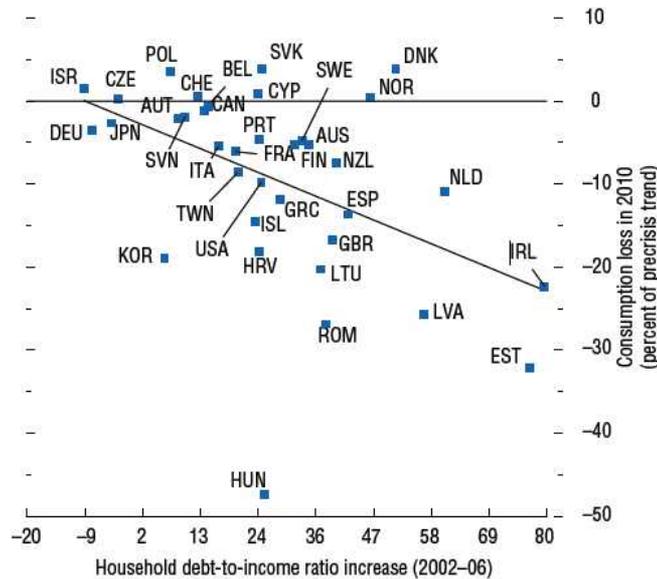
2) 해외 주요국가의 가계부채 관리

미국과 유럽의 일부 국가들은 금융위기 이전에 급격한 주택가격 상승을 하였지만, 서브프라임 사태 이후 주택가격 폭락으로 이미 한국보다 먼저 주택담보대출과 관련한 문제를 경험하였다. 이에 선행되었던 국가별 사례를 발견해 보고, 이를 거울삼아 주택담보대출과 관련된 문제를 해결하는 방안으로 모색할 수 있다고 본다.

주택담보대출 문제가 발생한 대부분 국가의 주택가격은 2000년대 들어 높은 상승세를 보였다. 그러나 2006년 이후 상승세가 큰 폭으로 둔화되었으며 상당수 국가에서는 하락세로 반전되면서 연체율이 증가하고 채무불이행이 발

생하였다. 특히, 영국, 아이슬란드, 스페인 등 유럽의 주요국가와 미국에서는 주택시장의 거품붕괴가 금융시스템까지 영향을 미쳐 국가재정위기로 확산되는 현상이 나타났다. IMF의 분석에서도 과도한 가계부채, 특히 주택담보대출은 경기침체의 악화와 회복의 주요 원인으로 제시되었다.

〈그림 6〉 금융위기 이전의 가계부채 증가와 향후 소비갭과의 관계



자료 : IMF, Ch. 3 “Dealing with Household Debt”, World Economic Outlook, April, 2012.

〈그림 6〉에서 가로축은 금융위기 이전인 2002년과 2006년의 가계부채증가율을 의미하고 세로축은 1996년부터 2004년 추세에 따른 2010년 가계소비 추정치와 2010년 실제 가계소비와의 차이를 나타내고 있으며 양축의 상관관계 결과 10%의 DTI증가는 -2.1%p 이상의 소비갭을 나타내고 있어, 가계부채 증가가 향후 가계소비 감소로 이어져 경기침체 원인이 될 수 있음을 보여주고 있다.

특히 영국(GBR)과 스페인(ESP) 등 유럽의 주요국가도 대출증가로 인한 소비갭이 크게 나타나 최근 이들 국가의 경제위기가 가계부채와 무관하지 않은 점을 확인할 수 있다. 따라서 본 연구에서는 위기당시 미국과 유럽의 국가별 가계부채 상황과 대응방안 등을 확인하고 가계부채관리의 필요성을 제시하고

자 한다.

### 가) 미국

미국은 금융위기 이후 초기에는 금융시장 안정과 경기부양 차원의 정책을 추진하였다. 2009년까지 주택시장 안정정책은 생애최초 주택구입자에 대한 주택구입시 부담을 경감시키기 위해서 세액공제를 실시하여 경기회복을 유도하였다. 부시정부 말기와 오바마정부 초기까지 총 1억 5천만 달러의 세제지원이 이루어졌다. 실행결과 민간 소비증가는 미미하고 재정적자만 키웠다는 비판이 제기되면서 이후에는 주택금융 지원정책 중심으로 가계부채 관리가 진행되고 있다. 미국의 공공 모기지 지원프로그램은 조정방식과 옵션에 따라서 크게 리파이낸싱과 채무조정, 압류 및 퇴거방지로 구분되며 세부적으로는 8개형태로 운영되고 있다.

리파이낸싱 프로그램 중 HARP(Home Affordable Refinance Program)는 모기지 연체가 발생하지는 않았지만 주택가격 하락으로 기존 주택자금대출의 어려움이 커진 경우 대출조건을 변경하여 대환하는 프로그램이며, FHA Short Refinance는 주택가격보다 모기지 대출이 많아진 대출자를 위한 연방 주택청의 원금삭감형 프로그램이다. 이 방법을 이용하면 대출에 대한 연체는 없지만 주택가격보다 대출금이 많은 경우 원금의 10% 이상을 삭감하고 대출조건을 개선하여 리파이낸싱을 할 수 있다. 그리고 FHA2LP는 선순위 모기지가 FHA Short Refinance을 받은 경우 후순위 추가대출을 지원하는 방식이다.

채무조정방식 중 HAMP(Home Affordable Modification Program)는 주택담보대출 연체자들의 원 상환액을 월 소득의 31%를 상회하지 않게 조정하면서 대출조건을 완화하는 다양한 옵션을 제고하는 방식이다. PRA는 HAMP의 지원방식 중 하나로 원금삭감을 지원하는 프로그램이며, 2MP는 후순위 대출 지원 프로그램으로 FHA2LP와 유사한 구조이다.

〈표 27〉 미국의 공공 모기지 지원 프로그램

구분	재 융자(refinancing)	채무조정	압류 및 퇴거방지
대표 프로그램	HARP (Home Affordable Refinance Program)	HAMP (Home Affordable Modification Program)	HAFA (Home Affordable Foreclosure Program)
원금삭감	FHA Short Refinance	PRA (Principal Reduction Alternative)	
후순위 대출지원	FHA2LP (FHA Second Lien Program))	2MP (Second Lien Modification Program)	
실업자 연체유예		HAUP (Home Affordable Unemployment Program)	
HARP	- 1년 無연체, LTV 80% 이상, 적격대출 - 금리인하, 대출기간 변경 등		
FHA <sup>35)</sup> Short Refinance	- 월상환액/월소득<50%, LTV<97.75% - 금리인하, 기간 변경 및 원금 10% 삭감		
FHA2LP	- 리파이낸싱 후 LTV<115% - 후순위 추가 대출		
HAMP	- 월상환액/월소득>31% - 월상환액 월소득의 31%로 조정		
PRA	- 월상환액/월소득>31%, LTV>100% - 월상환액 월소득의 31%로 조정, 원금 삭감		
2MP	- HAMP 3회 연속 無연체 - 후순위 추가 대출		
HAUP	- 실업급여 수령 증명 가능자 - 월 상환액 12개월 이상 유예		
HAFA	- 통상 HAMP 거절 소유주 - 숏세일 <sup>36)</sup> 이나 금융권 DIL <sup>37)</sup> 선택 가능		

자료 : 허윤경·이승우(2012), 『금융위기 이후 미국의 모기지지원 프로그램의 시사점』, 한국건설산업연구원, p.26. 천대중(2014), 『미국은행의 주거용 모기지 비즈니스 현황과 특징』, 우리금융경제연구소, p.15.

35) FHA(Federal Housing Administration)미국 연방주택청.

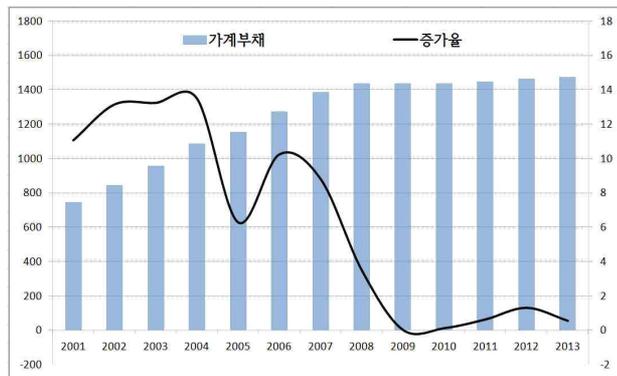
36) 숏세일은 모기지 사후관리 기관과 대출자 간 합의하에 압류 대신 대출자의 자율적 판매를 허용하는 제도임. 모기지 잔액 이하로 주택을 매도함으로써 채권채무 관계를 종결하는 제도로 모기지 잔액보다 주택가격이 낮을 때 시행됨.

37) DIL(Ded in Lieu of Foreclosure)은 친절한 차압(Friendly Foreclosure)이라고도 불리며 권리 양도의 형태임. 통상, 숏세일로 팔기도 어렵고 재정적으로 더 이상 주택을 보유하기 어려울 때 선택하는 옵션임, 숏세일을 시도한 서류가 필요한 경우가 다수임.

## 나) 영국

영국의 가계부채 증가원인은 저금리 지속에 따른 개인대출 증가가 주택가격 상승을 초래했다고 볼 수 있다.<sup>38)</sup> <표 28>의 가계부채의 구성비에서 알 수 있듯이 2005년~2010년까지 평균 80%가 넘는 높은 구성비를 보이며 부동산담보대출 취급을 하고 있었다. 이러한 주택담보대출 증가는 새로운 수익모델이 필요했던 금융기관들이 주택담보대출 확대에 적극적으로 나서면서 가계부채는 2000년 이후 7년간 연평균 9.9%의 높은 증가율을 기록하였다(남영우, 2012).

<그림 7> 영국의 가계부채 추이



자료 : Oxford Economics, 지역경제포커스(2014), 『영국 경기회복의 배경과 전망』 재인용, P.10.

<표 28> 영국의 가계부채 구성비

구분	Mortgage Loan <sup>39)</sup>	Revolving Credit, Automobile Loan	Others
2005	80.7%	15.4%	3.9%
2008	81.9%	13.2%	4.9%
2010	82.7%	12.0%	5.3%

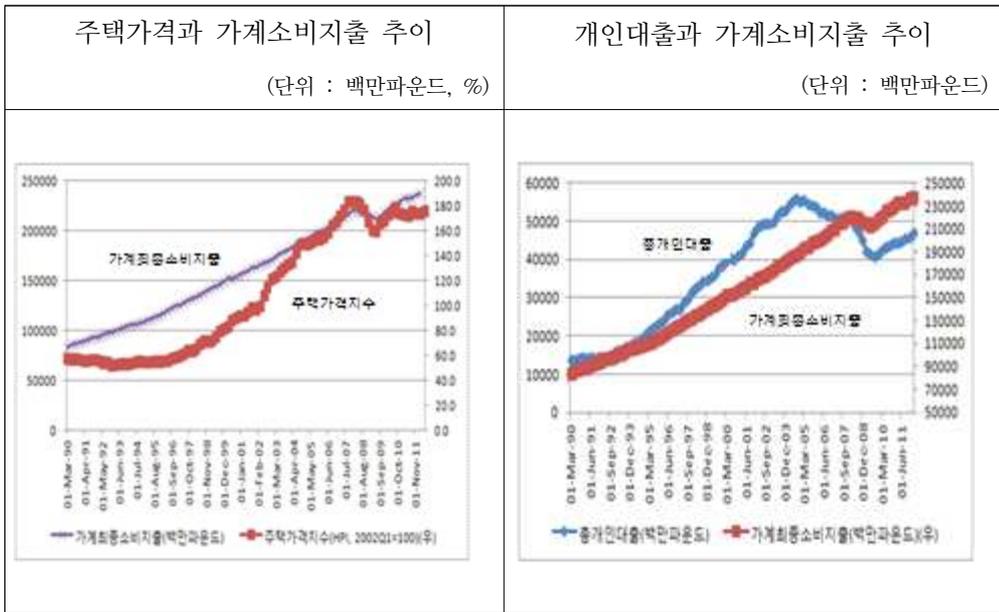
자료 : 국회예산정책처(2012), 『주요국의 가계부채 조정과정 및 대응정책분석』, p.17.

38) 국회예산정책처(2012), 『주요국의 가계부채 조정과정 및 대응정책분석』, p.18.

39) 부동산을 담보로 한 대출금융이며, 주택담보대출을 의미함.

영국의 주택가격은 2007년 하반기 이후 하락폭이 커졌으며, 글로벌 금융위기 여파로 주택가격이 급락하면서 금융기관의 부실이 심화되고 금융권도 높은 해외자산 의존도로 인해 재무건전성이 악화되었다.<sup>40)</sup> 이와 같은 주택가격 하락은 가계대출과 가계소비를 감소시켜 경기에 영향을 주었다.

〈그림 8〉 영국의 주택가격 및 개인대출과 가계소비지출 추이

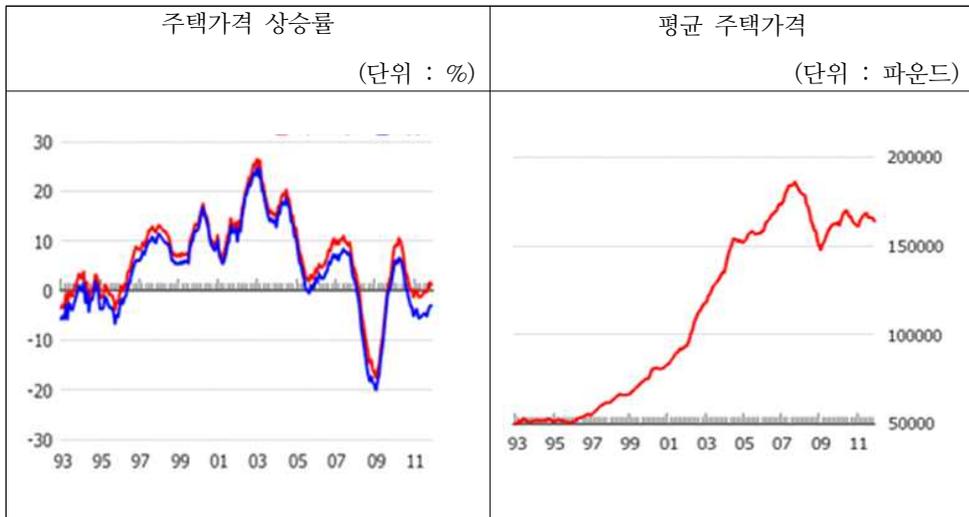


자료 : 국회예산정책처(2012), 『주요국의 가계부채 조정과정 및 대응정책분석』, pp.20-21.

〈그림 8〉을 보면 주택가격의 하락은 가계대출에 영향을 주고 이후 가계소비를 감소시키는 모습을 보여주고 있다. 가계부채가 크게 늘어난 상황에서 주택가격 하락은 주택담보대출의 원리금 부담을 높였으며, 가계 소비여력까지 위축시켜 개인파산 증가로 연결되었다. 2011년까지도 영국의 소비는 침체되고 경제의 불확실성이 지속되는 상황이었다.

40) 국회예산정책처(2012), 전계서, pp.16-22.

〈그림 9〉 영국의 주택가격 상승률과 평균 주택가격



자료 : <http://www.globalpropertyguide.com/>, 참고.

그러나 영국정부의 가계부채 규모는 꾸준히 증가하고 있으나, 2008년을 기점으로 점차 줄어드는 모습을 보이고 있으며, 가처분소득 대비 가계부채 비중 또한 지속적으로 하락하는 모습이다.

〈그림 7〉에서 볼 수 있는 것처럼 영국의 주택시장은 2011년부터 지속적으로 상승하고 있으며, 이것이 가계소비에 긍정적 영향을 끼치고 있는 것으로 평가되고 있다.<sup>41)</sup> 또한 연간 주택가격지수는 기타 유로존 국가들의 지수추이를 상회하는 전망을 보이고 있으며, 영국에 대한 주요기관들의 경제성장률 전망도 긍정적으로 보고 있어 경기개선은 이어질 것으로 보인다. 다만 이러한 긍정적인 전망에도 불구하고 지속성을 강화하기 위해서는 민간소비 뿐 아니라 투자가 활성화되고 경제성장을 이끌어갈 수 있는 추가적인 요소가 필요해 보이고 있다.

41) 지역경제포커스(2014), 『영국 경기회복의 배경과 전망』, 재인용, P.11.

〈그림 10〉 영국의 가치분소득대비 가계부채 비중 추이



주: 가계의 총 가치분소득 중에서 가계부채가 차지하는 비중임.

자료 : 영국중앙은행(BoE), 지역경제포커스(2014), 『영국 경기회복의 배경과 전망』, 그림 재인용, P.10.

#### 다) 아이슬란드

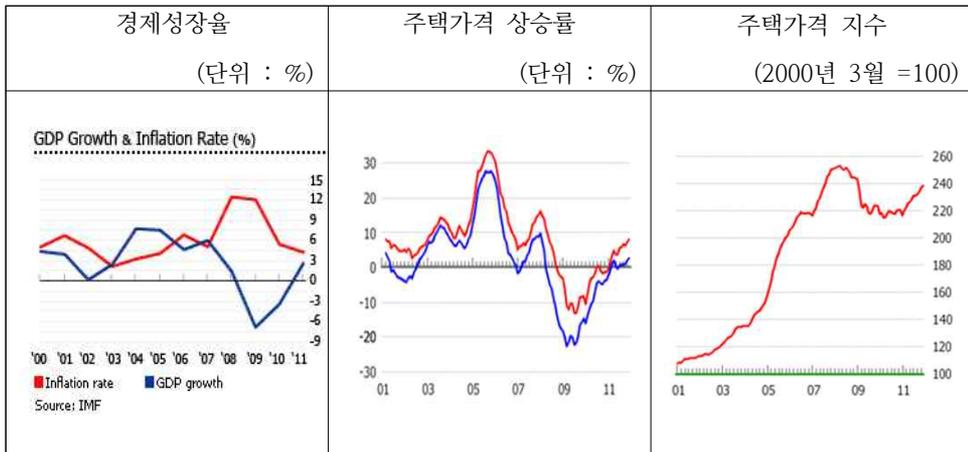
아이슬란드는 금융위기 여파로 2008년 유럽국가 중 가장 먼저 IMF로부터 구제금융을 받는 등 금융시스템붕괴와 경제가 침체하면서 주택가격이 급격하게 하락하였고, 주택담보대출의 LTV가 급격하게 상승하는 등 가계부채가 증가되고 하우스푸어 문제가 나타났다.<sup>42)</sup>

이러한 아이슬란드의 금융시스템 붕괴와 경제침체 원인을 2006년 당시의 가계부채 현황에서 돌아볼 필요가 있다. 당시 아이슬란드 은행은 경쟁적인 가계대출 확대로 2005년 말 가계부채는 전년대비 23.4%증가한 1조 800억 크로나를 기록하는 등 가계부채 규모가 크게 확대되었다. 2004년 예금은행들의 총여신 규모는 총수신의 약 4배에 이르며 그 중 가계대출이 차지하는 비중은 약 70% 수준 이었다(남영우, 2012). 아이슬란드 상황은 현재 우리나라 금융환경과 닮은 모습으로 금융기관의 경쟁적인 주택담보대출 확대와 주택담보대출 비중이 높은 점은 시사하는 바가 크다고 할 수 있다.

42) 남영우(2012), 『가계부채 문제해결을 위한 주택담보대출 관리방안』, 두리안, pp.54-55.

그러나 아이슬란드는 2008년 구제금융지원을 받고난 후인 2009년까지도 주택가격이 하락을 하였으나, 2010년부터 상승하기 시작했으며 최근까지도 상승현상이 이어지고 있다.

〈그림 11〉 아이슬란드의 경제성장률과 주택가격 상승률 및 주택가격 지수



자료 : Statistics Iceland, <http://www.globalpropertyguide.com/>, 그림 재인용.

주택가격 상승은 아이슬란드의 전반적인 경기회복도 있지만 가계채무에 대한 과감한 구조정책에 기인한 것이라 할 수 있다. 압류유예와 물가연동대출의 일시적 중지, 채무상환 재조정의 구조정책과 채무자는 옴브즈만 사무소를 이용하여 직접 채권자와 채무를 재조정하도록 하였다. 그리고 은행은 자발적으로 채무재조정을 실시하여 상환능력이 있는 채무자는 담보가치의 100%까지 부채를 감면하였다.

또한 해외투자 자본의 유출통제로 인해 외국자본이 부동산시장으로 집중되어 경제가 회복되고, 신규투자까지 유입되는 등 주택시장의 회복이 빠르게 나타났다. 이러한 정부주도의 채무조정 정책은 해외자본 유출통제와 함께 주택시장이 정상화되는데 중요한 역할을 하게 되었다.

채무조정 프로그램은 4가지가 있으며 상환부담조정 프로그램으로 상환과 만기연장에 대한 부담조정을 하였으며, 대출감면 협약으로 다중채무와 채무상환에 어려움이 있는 가구는 채무를 감면하였다. 또한 저소득층에 대해서도 대출감면 지원을 하였으며, LTV 110% 초과가구는 대출금 감액으로 만기연장을 유도하였다.

〈표 29〉 아이슬란드의 채무재조정 프로그램

프로그램	수익자	채무재조정	동기유인 및 부담경감
Payment Smoothing (상환부담조정)	소비자물가지수 및 외환연계대출과 자동차 대출	상환계획조정 및 만기연장	- 소비자물가지수 연계대출 - 외환연계대출
Sector Agreement (대출감면협약)	다중채무가구 및 채무상환에 어려움이 있는 가구	담보가액의 100%까지 대출금액 감면	- 대형대출기관 사이에서 정부가 주도함 - 자발적 참여 - 대출기관 대출금액 감면 부담
DO (대출감면지원)	Sector Agreement와 유사하지만 보다 저소득층	부분협정과 유사하지만 더 많은 대출금액 감면	- 재조정 실패할 경우 법원주관하에 채무재조정 - 대출금액 감면부담
Mortgage Write-down (대출금 감액)	2010년 12월 현재 LTV 110% 초과 가구	대출원금을 담보 자산의 110%까지 감소	- 모기지대출자와 정부간의 협정 - 자발적 참여 - 대출기관 대출감면 부담

자료 : IMF, Ch. 3 “Dealing with Household Debt”, World Economic Outlook, April, 2012.

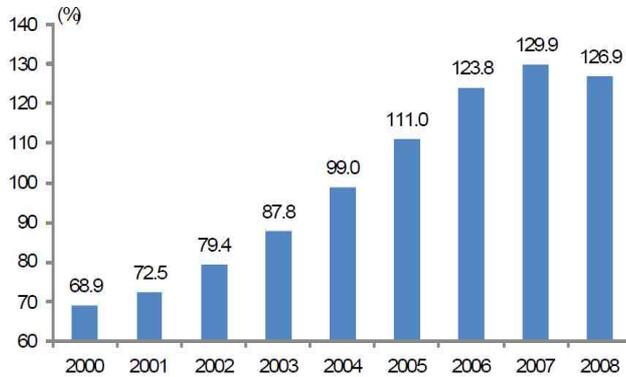
#### 라) 스페인

스페인 재정악화 주요원인은 크게 세가지 정도로 들 수 있다. 부동산 버블 붕괴, 지방재정의 악화, 재정의 펀더멘탈 약화로 볼 수 있다. 먼저 부동산 버

블은 2000년 이후 부동산가격이 급등하면서 가계부채도 증가하였다.<sup>43)</sup> 이러한 원인은 1999년 유로화 도입 이후 저금리기조에 해외자금이 부동산시장으로 유입되었으며 저축은행들의 부동산담보대출 확대가 주택관련대출 급등의 원인이 되었다.

이후 2008년 부동산 버블붕괴로 주택가격은 급락하였고 저축은행의 부실채권이 급증하였다. 지방재정 악화의 경우는 중앙정부의 통제력 약화와 방만한 재정운용이 원인으로 지방정부 적자가 지속되었다. 그리고 재정의 펀더멘탈 약화는 노동시장에서 과도한 고용보장과 높은 실업수당 이었으며 취약한 산업경쟁력과 환율 조정기능의 부재로 볼 수 있다.

〈그림 12〉 스페인의 가처분소득 대비 가계부채 추이



자료 : Eurostat, 대한상공회의소(2012), 『가계부채현황과 정책과제』, 재인용.

스페인의 주택가격 하락원인은 재정적자와 함께 가계부채조정 등으로 스페인 경제의 불확실성이 증가하면서 하락되었다. 주택가격지수는 2008년 2,102.1포인트를 정점으로 하락하기 시작하여, 2010년에는 1,863.4포인트로

43) 스페인의 가처분소득 대비 가계부채 비율은 2007년 129.9%로, 2000년 68.9%에 비해 1.9배나 상승하였다, 대한상공회의소(2012), 『가계부채현황과 정책과제』.

11.4% 하락하였다. 주택가격과 함께 주택시장에 과잉투자가 이루어져 미분양 주택은 100만채를 상회한 것으로 나타났다.<sup>44)</sup>

〈그림 13〉 스페인의 주택가격 추이



주: 증감률은 전년동기대비임.

자료 : Reuters, 대한상공회의소(2012), 가계부채현황과 정책과제 재인용.

스페인은 이러한 문제를 해결하기 위해서 가계부채를 감당하지 못하는 가구들에 대해서 2008년부터 2010년까지 GDP의 7%에 해당되는 여신을 디레버리지 하는 구조조정을 실시하였으며, 저축은행 등 금융기관들이 자체적으로 채무를 조정하고 대출대상 주택을 매입하여 재임대하는 방법을 사용하였다.

그러나 한계 채무자에 대한 직접적인 지원보다는 소극적인 지원으로 채무불이행 가구에 대해서 채무의 일시적 감면이나 채무조정으로 원리금상환을 유도하였다. 일부 은행들은 채무불이행 방지를 위해 해당주택을 매입하고 재판매 시점까지 임대해주는 방법도 활용하였다.

이러한 방법은 채무자의 부실화를 차단하기 위한 관리 및 정책지원이 미흡하여 은행 부실화의 주요인으로 작용하였으며, 국가차원에서 통계대상이나 부실채권 비중에 포함되지 않는 문제가 나타났다. 은행은 매입주택의 재판매를

44) 남영우(2012), 전계서, p.53.

위해서 낮은 금리의 대출을 실행하면서 결과적으로 금융기관 부실 악화의 원인이 되었다.<sup>45)</sup>

OECD 보고서에 따르면 스페인의 2012년 가계부채비율은 141.1%로서 2007년 149.3%의 최고치를 기록한 이후로 점차감소를 보이고 있다. 그러나 지방재정의 건전화와 금융기관의 정상화를 위해서는 이들에 대한 과감한 구조개혁이 요구되고 있다.

### 3) 해외 사례의 시사점

미국과 유럽의 사례로 본 가계부채의 원인은 저금리 정책으로 인하여 주택 수요가 증가되고 이로 인해 주택가격이 상승 되었다고 볼 수 있다.

주택가격 상승으로 발생하는 가계부채의 증가는 주택가격이 하락하면서 가계의 순자산 감소와 가계소비를 감소시키고 하우스푸어가 발생하는 문제까지도 영향을 주고 있다. 스페인의 경우는 대출금리가 인상되어 주택가격이 하락하고 경제성장률도 낮아지는 현상을 나타냈다.

---

45) 남영우(2012), 전계서, p.53.

〈표 30〉 국가별 가계부채 원인과 영향

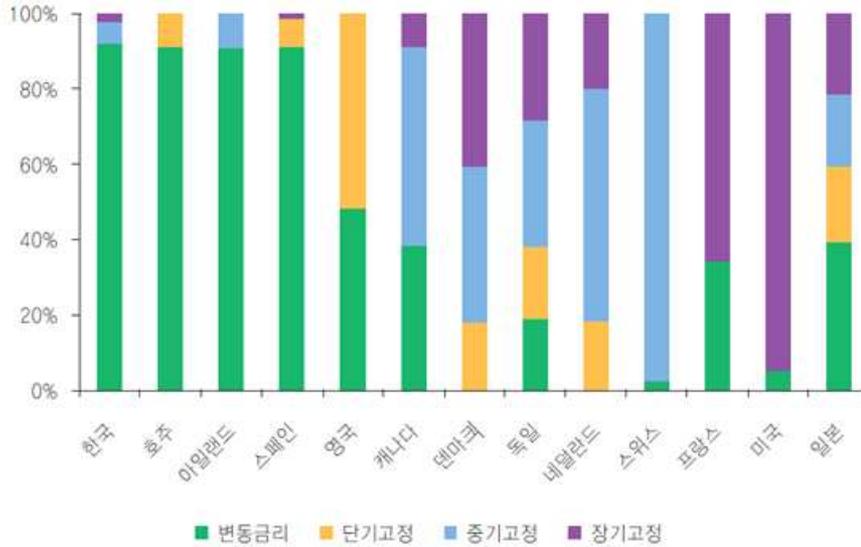
구분	가계부채 원인	가계부채 영향
미국	<ul style="list-style-type: none"> <li>- 외국의 투자자금 유입으로 과잉유동성이 발생하고 소비증가로 부채 증가를 유발함</li> <li>- 저금리로 인한 주택수요의 증가와 주택가격 상승</li> </ul>	<ul style="list-style-type: none"> <li>- 주택가격의 하락으로 가계의 순자산 감소와 개인소비지출을 감소시킴</li> </ul>
영국	<ul style="list-style-type: none"> <li>- 저금리 지속으로 인한 주택수요의 증가와 주택가격 상승</li> </ul>	<ul style="list-style-type: none"> <li>- 주택가격의 하락은 가계대출을 감소시켰으며 가계소비를 감소시켜 경기에 영향을 줌</li> </ul>
아이슬란드	<ul style="list-style-type: none"> <li>- 은행의 경쟁적인 가계대출확대</li> </ul>	<ul style="list-style-type: none"> <li>- 가계부채가 증가되고 하우스푸어 문제가 발생됨</li> </ul>
스페인	<ul style="list-style-type: none"> <li>- 1990년이후 2005년까지 주택시장의 호황으로 주택수요증가</li> <li>- 금리인하와 주택구입시 세제혜택이 주택수요 증가원인</li> </ul>	<ul style="list-style-type: none"> <li>- 2008년부터 대출금리가 인상되었으며 이 영향으로 주택가격이 하락하고 경제성장률도 낮아짐</li> </ul>

자료 : 국회예산정책처(2012), 『주요국의 가계부채 조정과정 및 대응정책분석』 내용을 참조.

주목할 점으로는 가계부채가 증가한 후 주택가격이 하락하면서 나타나는 문제는 변동금리 비중이 높은 국가인 영국, 스페인 등에서 나타난 것을 볼 수 있다. 〈그림 14〉에서 조사된 자료는 미국의 경우 변동금리 비중이 낮지만 한국, 호주, 아일랜드, 스페인, 영국 순으로 변동금리 비중이 높았다.

특히 주택을 재산증식의 수단으로 이용하는 국가에서는 조기상환수수료가 부담이 없는 변동금리를 선호하는 경향이 있어 변동금리 비중이 높은 우리나라와 유사성을 나타내고 있다.

〈그림 14〉 국가별 주택담보대출 금리 유형



자료 : RBA, CHMC, EMF, GPG, MBA AND S&P, 2009년 신규대출기준.

출처 : Lea, International Comparison of Mortgage Product Offerings, Research Institute for Housing America, 2010. 이재연(2011), 『주택금융제도의 국제간 비교 및 정책 제안』, p.18 <표1> 재인용.

주요 국가들의 가계부채 및 하우스푸어에 대한 대책을 한국은행이 발표한 가계부채 대응법을 이용하여 분류하면 다음과 같은 방법 및 평가를 나타내고 있다.

먼저 긴축정책은 가계부채의 감축(디레버리징; deleveraging)으로 대출상환 기간을 연장하고 대출금리 인하와 채무를 감면하는 등 채무를 조정하는 정책이다. 이러한 정책은 금융기관의 수익성이 악화되고 또한 건전성도 악화될 수 있다. 그리고 신규대출의 감소는 가계부채 조절에 긍정적인 측면이 있으나 주택시장이 침체할 수 있는 가능성이 높기 때문에 신중한 접근이 필요하다. 미국의 경우 2008년 금융위기 이후 정부 주도하에 부실자산과 금융기관 구조조정을 추진하여 안정화를 이루었지만 직접적인 정부의 개입이 없었던 것이다.

그러나 영국의 경우 부실채권 정리절차의 번거로움과 가계부채에 대한 금융기관의 소극적인 대처로 가계부채의 감축은 효과가 크지 않고 있다. 현재 미국의 경우 효율적인 부실채무정리 절차에 힘입어 빠른 경기회복 전망을 보이고 있다.

반면 고성장의 경우에는 주택시장을 활성화하고 부동산 거래를 촉진시켜 문제를 해결하는 방법으로 미국과 영국에서 도입하였고 최근 우리나라에도 규제완화를 포함한 많은 부동산대책을 발표하고 있다. 하지만 대부분 정책이 단기적으로 시장회복에 한계가 있는 점을 고려할 때 효과적인 성과를 확신하기 어려운 상황이다. 특히 금융위기 이후 영국에서 성장주도 방법을 적극적으로 사용하고 있으나 뚜렷한 성과가 없으므로 성장기조를 유지하면서 다른 방법과 병행해야 할 것으로 판단된다.

가계부채를 정부부채로 이전하는 방안은 주택담보대출에 대한 위험을 정부에서 감당하는 방법이다. 유럽국가들이 금융기관의 부실을 정부에서 지원하는 방법을 사용하였으나 오히려 국가재정위기로 확산되는 경우도 나타났다. 우리나라의 경우, 가계부채 규모가 문제로 지적되고 있지만 정부부채 규모 또한 관리가 필요한 시점이라고 밝히고 있다. 2010년 정부부채는 392조원으로 GDP대비 34.5%로 낮은 수준이나 공공기관 부채와 통안채를 포함한다면 국가부채의 규모는 941조원으로 GDP대비 82.6%로 크게 상승되어 정부부채에 대한 체계적관리가 필요한 시점이다.<sup>46)</sup> 국민의 세금으로 지원한다는 비판의 여지도 있고 국가부채를 줄여야 할 시점에 가계부채를 정부부채로 이전하는 방법은 바람직하지 않다고 본다.

채무재조정의 경우 IMF에 따르면 거시적인 경기부양책 보다는 효과적인 방안으로 제시되고 있으며 미국과 스페인, 아이슬란드 등에서 적극적으로 적용되었다. 다만 스페인의 경우 금융기관이 독자적으로 채무조정을 하였고 아

---

46) 남주하(2012), 『가계부채 해소방안』, pp.61-62.

이슬란드와 미국의 경우 정부 주도하에 이루어졌으며 정부의 주도로 이루어진 경우가 가계부채 관리에 더욱 유용한 것으로 나타났다.

〈표 31〉 가계부채에 대한 대응법

대응방안	방 법	평 가
긴 축	경제주체의 지출억제 등으로 부채의 증가율 감소 및 절대규모 축소	디레버리징(deleveraging) <sup>47)</sup> 과정 과정에서 소비투자의 감소와 경기침체로 경제주체의 부채상환 능력이 더욱 악화 될 위험
고성장	부채를 줄이지 않고 높은 성장을 통해 부채상환 능력 향상	가장 이상적인 방법이나 수출급증 등 상당한 행운이 따라야 함
정부부채로의 이전	민간 경제주체들의 부채를 구제금융, 경기부양책 등을 통해 정부가 부담	정부의 부채수준이 높은 경우 채택이 어렵고 국가 부도위기를 초래할 수 있음
고 인플레이션	인플레이션을 통해 부채의 실질가치를 하락시키는 것으로 주로 정부 채무부담을 줄이는 데 활용	국제경쟁력 저하, 기축통화국이 아닌 경우 외환위기 등에 처할 우려
채무재조정	채권자들의 합의를 통해 부채의 만기, 규모, 금리 등을 상환이 가능한 범위 내로 조정	채무재조정에 따른 신용훼손으로 장기간 상당한 비용지출 불가피

자료 : 한국은행, 『부채경제학과 한국의 가계 및 정부부채』, 2012년 4월호.

## 제 4 절 데이터마이닝기법의 개념과 특징

### 1) 데이터마이닝 모형의 개념

데이터마이닝은 적용분야에 따라서 개념정의에 다소 차이가 있으나, 일반적

47) 부채 정리를 말함, 보유한 자산을 상회하는 부채를 끌어들여 이를 지렛대로 삼아 투자수익률을 높이는 레버리지(leverage)의 반대말이다.

으로 데이터베이스, 데이터웨어하우스, 데이터마트 등 자료저장소에 저장되어 있는 방대한 양의 데이터로부터 의사결정에 도움이 되는 유용한 정보를 발견하는 일련의 작업집합을 의미한다(이준, 2012). 또한 데이터마이닝은 데이터에 잠재되어 있으나 발견되지 않고 있는 유용한 정보를 추출함으로서 의사결정을 지원하는데 적용될 수 있으며, 주로 과거의 데이터를 중심으로 미래의 행위를 예측하는데 사용되는 경우가 많다. 데이터마이닝의 활용을 위해서는 대용량데이터의 구축이 필수적이며, 특정한 분석기법을 활용하는 것이 아니라 다양한 기법을 활용한 일련의 분석과정을 의미한다. 따라서 데이터마이닝을 위한 다양한 방법들이 개발되고 있어, 광범위한 데이터를 통해 유용한 정보를 추출할 수 있는 경우 데이터마이닝 기법으로 인정되므로 포괄적인 개념이라고 할 수 있다.

〈표 32〉 데이터마이닝기법의 종류

데이터 마이닝 기법	내 용
인공신경망 (artificial neural network)	인공신경망이란 인간의 뇌를 흉내내어 구성된 네트워크로 그 학습 능력과 추론능력이 매우 뛰어난 것으로 알려져 있다. 신경망은 단순한 처리인자들인 뉴런(neuron)의 전이함수(transfer function)와 신경망의 구조를 나타내는 계층(layer)의 수와 처리인자들 간의 연결 상태(connection) 및 연결강도(weights)로 구성되어있다. 이들을 적절하게 조정하는 학습규칙에 따라 여러 가지 형태를 갖게 된다(SAS Inc., 1998).
의사결정나무 (decision tree)	의사결정나무는 의사결정 규칙을 도표화하여 관심대상이 되는 집단을 몇 개의 소집단으로 분류하거나 예측을 수행하는 분석방법이다. 분석과정이 나무구조에 의해서 표현되기 때문에 판별분석(discriminant analysis), 회귀분석(regression analysis), 신경망(neural networks) 등과 같은 방법들에 비해 연구자가 분석과정을 쉽게 이해하고 설명할 수 있다는 장점을 가지고 있다.

자료 : 최중후·소선하(2005), 『사례로 배우는 데이터 마이닝』, 자유아카데미, p.25, p.69.

과거에는 주로 개념적으로 연구되었던 데이터마이닝이 실제 데이터를 활용하여 실용적인 분야로 확장되어 가고 있는 이유는 자료의 수집 및 처리비용이 급격하게 낮아지고 있기 때문이다. 또한 자동화된 자료수집과 전자도구의 일상화, 자료저장 구조에 대한 기술적 발전은 데이터를 보다 분석에 적합하도록 축적하는데 도움을 주고 있다. 이에 대한 관심은 데이터마이닝과 그 관련 분야인 데이터 웨어하우징(data warehousing)의 발전을 가속화 하고 있다.

## 2) 데이터마이닝 모형의 특징

데이터마이닝은 다음과 같은 특징이 있다(구자용 등 2011). 먼저 데이터마이닝은 대용량의 관측가능한 자료를 활용한다. 다만 기본자료 등은 데이터분석을 염두에 두지 않고 시간의 흐름에 따라 축적되는 것이 일반적이다. 따라서 분석을 위한 자료의 클린징 및 코드관리 등의 후속작업이 필요하다.

두번째 특징은 데이터마이닝이 컴퓨터중심적기법이라는 점이다. 20세기 이후 컴퓨터의 성능발전과 가격하락으로 다양한 분야에서 데이터마이닝의 구현이 가능해졌다.

또한 데이터마이닝은 그 분석기법의 특성상 경험적 방법에 근거하고 있다는 특징이 있다. 기본적으로 원천자료로 과거 데이터를 사용해야 한다는 측면에서 신경망모형이나 의사결정나무와 같은 데이터마이닝을 위한 기법은 이론적 원리에 기초하여 개발되었다기 보다는 경험에 기초하여 개발되었다. 이렇게 경험적 방법에 근거한 분석방법은 수리적인 수단으로 증명하는데 한계가 있다는 특징도 있다.

활용 측면에서 데이터마이닝의 특징은 일반화와 의사결정지원에 있다. 일반화는 예측모형이 새로운 자료에 얼마나 잘 적용되도록 하는 것인가를 의미하며, 의사결정지원은 데이터마이닝의 결과가 기업 등의 다양한 경영상황하에서 경쟁력확보를 위한 의사결정에 활용될 수 있어야 한다는 것이다.

## 제 5 절 선행연구

### 1) 주택담보대출에 관한 선행연구

주택담보대출 및 가계부채에 관련된 선행연구는 가계대출과 관계가 있는 주택관련 대출이며 가계대출 리스크관리와 차입자의 금리선택에 대한 내용으로 구분된다. 가계대출 증가요인을 제시한 연구는 Dynan and Kohn(2007)으로 가계부채는 저축과 소비와 밀접한 관계가 있다고 보았으며, 가계부채가 가처분소득과 GDP 증가율보다 빠르게 증가하거나, 소득대비 부채비율이 증가되는 현상을 규명하였다. 그리고 미국의 부채가 급증한 원인으로 저금리 보다는 자산가격 상승에 따른 주택관련 대출이 증가한 것으로 보았으며, 고학력과 청장년층 인구비중의 증가 및 금융혁신도 증가되는 일부원인으로 제시하였다.

Brown and Taylor(2008)는 미국, 영국, 독일의 가구자료를 이용한 모형을 통해 자산과 부채의 규모가 동시에 결정된다고 하였으며, 이러한 자산과 부채 규모의 결정에는 가구의 특성과 총소득, 가구원 등이 유의한 영향을 미친다고 제시하였다. 분석결과 자산규모에 영향을 주는 것으로, 높은 교육수준과 기혼, 그리고 가구주 규모가 큰 가구일수록 순자산이 커지는 것으로 나타났다. 또한 소득 4분위 분석에서는 소득분위가 낮을수록 소득이 감소 될 때, 순자산의 감소효과가 커진다고 제시하고 있다.

김현정(2010)은 우리나라 가계부채의 문제점을 가계부채 현황과 추이 그리고 가계부채의 특성을 살펴본 다음 거시금융환경 변화에 따른 가계재무건전성의 민감도를 분석하여 가계부채 문제를 제시하였다. 우리나라의 가계는 실물자산 구입을 위해 상당부분의 부채를 보유하고 있고, 가계대출의 만기구조 및 상환방식과 관련되는 취약성과 소득하위계층의 경우 금리 및 소득충격에 더 취약한 것으로 나타났다.

주택담보대출에 관한 두 번째 연구유형은 차입자 입장에서 대출상품 및 구조에 대한 연구이며 주택담보대출 금리선택 선호연구이다. Alm and Follain(1987)과 Brueckner(1986)는 차입자 입장에서 대출특성을 연구하였다. 차입자가 대출위험에 대하여 감수하는 수준, 시간선호도 비율, 주택에 대한

선호 등의 차입자 개별변수가 주택담보대출 고정금리 선택에 영향을 미친다고 제시하였으며 개인적인 부(wealth)의 정도, 대출만기기간, 대출금액 등도 금리선택에 영향을 미치는 것으로 나타났다.

Dhilon, Shilling and Sirmans(1987)의 연구에서는 주택담보대출 전문은행의 자료를 이용하여 분석하였으며, 차입자 특성변수와 이자율 차이의 시장변수 중에서 어느 변수가 금리선택에 영향을 미치는지를 분석하였다. 시장의 변수에는 대출시점에서 고정금리 수준이 영향을 미쳤으며, 변동금리 선택에 미치는 영향으로는 변동금리 대출에 대한 금융기관의 마진이며, 차입자 특성변수 중에서 단기보유 목적과 결혼여부, 공동명의 차입여부 등이 영향을 미치는 변수로 제시되었다.

박성욱외 2인(2009)은 금융기관 주택담보대출 자료를 이용하여 차입자의 위험관련 특성변수와 금리차이가 차입자 금리선택에 영향을 미치는 것으로 추정하였다. 고정금리와 변동금리 차이가 차입자의 금리선택에 영향을 미치며, 차입자 변수들 중에는 나이, 성별, 직업유형, 주택구입가격 등이 금리선택에 영향을 미치는 것으로 나타났다. 즉, 나이가 많을수록, 남성일수록, 자영업 및 개인기업 일수록, 그리고 부(wealth)의 수준이 높을수록 변동금리를 선택하는 것으로 나타났다. 그러나 대출금액과 대출만기기간이 증가할수록 고정금리를 선택하는 것으로 나타났다.

손 옥(2009)은 주택담보대출 변동금리형 기준금리의 적절성을 평가하였으며, 편중되어있는 변동금리형 현상을 완화할 수 있는 방안을 제시하였다. 주택담보대출의 기준금리가 되는 시장금리를 대상으로 적절성을 평가하였으며, 금리조건 선택에 영향을 미치는 요인으로 금리 등 가격변수와 차입자 개인의 특성변수인 대출위험에 대한 태도를 제시하였다. 그리고 차입자에 대한 금융정보제공과 교육방식의 정교화, 고정·변동금리 혼합형 대출의 활성화, 고정금리와 변동금리 간의 금리격차 축소 유도를 변동금리대출 편중현상을 완화할 수 있는 대안으로 제시하였다.

유승동(2013)은 주택담보대출 잠재적 차입자 경우 주택보유 여부에 따라 선호하는 금리유형이 있으며, 주택을 보유한 잠재적 차입자는 고정금리보다 변동금리를 선호한다고 제시하였다. 이는 투자수요가 높은 차입자의 경우 금

리변동에 따른 이자상환 위험을 감당할 수 있는 유인이 높기 때문으로 보았다.

〈표 33〉 주택담보대출에 관한 선행연구

저 자	내 용
Brueckner (1986) Alm and Follain (1987)	차입자 입장에서 대출상품 및 구조에 대한 선호도를 분석함. 차입자의 대출위험 감수 수준, 시간선호도 비율, 주택에 대한 선호 등의 변수가 주택담보대출 고정금리 선택에 영향을 미침.
Dhillon, Shilling and Sirmans(1987)	차입자 특성변수와 이자율 차이의 시장변수 중 금리선택에 영향을 미치는지를 분석함. 시장변수는 고정금리 수준에 영향을 미침, 변동금리는 단기보유의 목적, 결혼여부, 공동명의 등이 영향을 미침.
Dynan and Kohn(2007)	가계부채는 저축과 소비와 밀접한 관계가 있음. 가계부채가 가처분소득과 GDP 증가율보다 빠르게 증가하거나, 소득대비 부채비율이 증가되는 현상을 규명함.
Brown and Taylor(2008)	자산과 부채의 규모는 동시에 결정됨. 자산과 부채규모의 결정에는 가구주의 특성과 총 소득, 가구원 등이 유의한 영향을 미친다고 제시함.
박성욱, 박갑제, 김태혁(2009)	주택담보대출 차입자의 금리선택에 대해서 연구함. 차입자의 특성변수와 금리차이가 차입자의 금리선택에 영향을 미침, 나이, 성별, 직업유형, 주택구입가격 등이 영향을 미침.
손 옥(2009)	주택담보대출의 기준금리 및 편중완화에 대한 연구함. 금리선택의 요인은 가격변수와 차입자 개인의 특성변수인 대출위험에 대한 태도라고 제시함.
김현정(2010)	거시금융환경 변화에 따른 가계재무 건전성의 민감도를 분석함. 가계대출의 만기구조 및 상환방식과 관련되는 취약성과 소득하위 계층의 경우 금리 및 소득충격에 더 취약한 것으로 나타남.
유승동(2013)	주택담보대출 잠재적 차입자의 경우 주택의 보유여부에 따라 선호하는 금리유형이 있음을 제시함. 주택보유 잠재적 차입자는 변동금리를 선호하며, 투자수요가 높은 경우 변동금리 이자상환 위험을 감당한다고 제시함.

## 2) 주택담보대출의 리스크 요인에 관한 선행연구

주택담보대출 위험에 대한 연구로는 크게 연체율요인을 분석한 연구와 주택담보대출이 가계에 미치는 영향을 분석한 연구가 있다. 먼저 주택담보대출의 연체율에 대한 연구이다.

위중범·백흥기(2008)는 부동산담보대출 연체율과 금리 등 거시 변수들의 관계를 실증적으로 분석하여, 금리정책을 통하여 경기조절과 부동산 가격을 고려하는 방안을 분석한 결과, 부동산담보대출 연체율이 부동산 관련 지표로서 금리정책에 이용하는 가능성을 제시하였다. 연체율은 금융과 부동산 지표로서 유용한 정보를 담고 있으며, 부동산담보대출 연체율은 소득이나 실업률 등보다 금리에 가장 민감하고 직접적으로 영향을 받는 것으로 나타났다. 연체율은 거시변수들뿐 아니라 부(wealth)의 효과, 금융상황, 소득의 분배구조 등도 반영하며, 부동산 가격 지표인 주택매매가격지수보다 거시경제변수들과 밀접하게 관련되어 있다고 보았다.

심종원의 2인(2009)은 2005년부터 2009년까지 주택담보대출연체율, 실업률, 소비자물가지수, 대출금리, 주택매매가격지수 등의 변수를 통해 주택담보대출 연체율에 영향을 주는 요인을 분석하였다. 분석결과 실업률과 소비자물가지수는 연체율에 양(+)의 영향을 미쳤으며 주택매매가격지수는 음(-)의 영향을 미치는 것으로 나타났다. 분해분석 결과에서는 소비자물가지수가 주택담보대출 연체율에 가장 많은 영향을 미치는 것으로 나타났다.

김희철·박형근(2010)은 패널데이터 모형을 통한 국내 일반은행의 연체율 분석에서 은행대출을 기업대출과 가계대출 및 신용카드대출로 분류하고 거시경제변수 중 연체율에 영향을 미치는 요인에 대해서 분석하였다. 분석결과 국내 일반은행 연체율에 대해서 소비자물가지수는 정(+)의 영향을 미치는 유의한 변수로 나타나고 동행(경기)종합지수와 종합주가지수는 음(-)의 영향을 나타내는 유의적인 변수이지만 환율, 국민주택채권 그리고 고용률은 각각 음(-)의 영향을 나타내는 비유의적 변수로서 연체율에 큰 영향을 미치지 않은 것으로 나타났다.

Brent William et al.,(2011)은 2005년부터 6년간 미국의 주요지역에 대해

서 주택담보대출의 연체율결정요인을 분석하였다. 분석결과로 차입자에 대해서는 인종, 성별, 신용평가점수, 소득 등이 영향을 미치는 것으로 나타났다. 대출구조에서는 자가거주 여부, 변동금리 여부 등이 영향을 미치고 경제변수로서 대출불승인율, 실업률, 건축허가건수, 전문·사무직 종사자수의 증가율 등이 영향을 미치는 것으로 나타났다.

허석균(2012)은 DTI, LTV 및 대출상환 조건이 주택담보대출의 연체율에 미치는 영향을 분석하고 LTV, DTI 및 주택담보대출의 변수들이 조기연체에 미치는 영향을 추정하였으며, 그 결과 LTV와 DTI가 가계대출 연체율을 높이며, 장기·원리금분할상환 방식은 연체율을 낮추는 것으로 나타났다.

주택담보대출이 가계에 미친 영향에 대한 연구로 장민·이규복(2010)은 기준금리인상이 가계건전성에 주는 영향으로서, 향후 기준금리 1%p 상승 시 전체 가계의 연간 이자부담액은 4조원 증가하고 국민처분가능소득(2009년 기준)대비 비중은 약 0.4%p 확대될 것으로 예상하였으며, 저소득계층의 경우 상대적으로 높은 부채수준을 지니고 있어 금리상승 시 부작용이 가장 빠르게 파급될 것으로 예상하였다. 또한 주택을 소유한 가구의 총부채 규모가 커서 가계수지를 상대적으로 빠르게 악화시킬 것으로 추정하였고, 저소득층의 경우 DTI가 40%를 상회하는 가구의 비율이 부채보유 가구의 31%로 고소득층에 비해 크게 높은 것으로 분석되어 DTI 규제 완화는 저소득층의 부채를 증대시켜 장기적인 부작용을 초래할 가능성이 높음을 제시하였다.

김일수·남영우(2011)는 이자율변화가 가계에 미치는 영향을 분석하였으며, 이자율상승이 가계의 흑자율에 미치는 영향은 소득분위가 하락할수록 흑자율 감소폭은 크게 나타나며, 이자율상승이 가계소득에 더 많은 영향을 미친다는 결과를 제시하였다.

남영우·성상준(2012)은 가계대출 구조의 변화가 가계에 미치는 영향에 대하여 분석하였으며, 기존의 변동금리와 거치식 상환방식을 고정금리 원리금균등 상환방식으로 개선할 경우 단기적으로는 가계위험과 가계의 건전성을 강화시킬 수 있으나 대출이자 상환액이 증가되어 가계부담이 증가되고 주택의 수요가 감소된다고 제시하였다. 저소득층 일수록 소득감소가 가계의 상환부담에 미치는 영향이 더 큰 것으로 나타나 대출구조 개선과 차별화된 지원방안의

필요성을 제시하였다.

〈표 34〉 주택담보대출의 리스크 요인에 관한 선행연구

저 자	내 용
위중범·백홍기 (2008)	부동산담보대출 연체율이 부동산 관련 지표로서 금리정책에 이용 가능성을 제시함.
심종원·정의철·정현정(2009)	주택담보대출 연체율, 실업률, 소비자물가지수, 대출금리, 주택매매가격지수 등의 변수를 통해 주택담보대출 연체율에 영향을 주는 요인을 분석함.
장민·이규복(2010)	기준금리인상이 가계 건진성에 주는 영향을 분석함. 기준금리 1% 상승시 가계의 연간 이자액은 4조 증가, 저소득층은 많은 부채보유로 부작용이 크고, 부채가구는 가계수지 악화됨.
김희철·박형근 (2010)	국내일반은행의 연체율을 분석함. 소비자물가지수는 정(+)의 영향으로 유의한 변수, 환율, 고용율은 음(-)의 영향으로 영향을 미치지 않은 변수로 제시함.
김일수·남영우 (2011)	이자율 상승이 가계에 미치는 영향을 분석함. 주택담보대출가구 중 소득분위가 낮을수록 이자율상승이 가계소득에 더 많은 영향을 미친다는 결과를 제시함.
Brent William et al.,(2011)	미국 주요지역의 주택담보대출 연체율 결정요인을 분석함. 차입자에 대해서는 인종, 성별, 신용평가점수, 소득 등이 영향을 미치는 것으로 나타났다.
남영우·성상준 (2012)	차입자가 선택하는 대출구조에 따른 가계의 DTI를 분석함. 저소득층의 경우에 소득감소가 가계의 상환부담에 미치는 영향이 더 큰 것으로 나타나 차별화된 지원방안의 필요성을 제시함.
허석균(2012)	대출상환 조건이 주택담보대출의 연체율에 미치는 영향을 분석함. LTV와 DTI가 가계대출 연체율을 유의하게 높이는 한편 장기·원리금 분할상환 방식은 연체율을 유의하게 낮추는 것으로 나타남.

### 3) 주택담보대출의 연체가능성에 관한 선행연구

지규현·김정인·최창규(2006)는 은행 및 외부에서 이용가능한 차입자 관련 자료를 기초로 주택담보대출 고객의 위험특성과 위험요인을 분석하였으며 분산분석(ANOVA)을 통하여 연체그룹과 연체하지 않은 그룹의 금융이용 및 차입자 특성을 분석하였고, 연체를 결정하는 공통적인 요인을 찾기 위한 요인분석(factor analysis)과 차입자의 유사성 및 상이성을 규명하는 군집분석(cluster analysis), 그룹별 연체요인분석을 위해 이항로짓분석(binary logit analysis)을 통하여 LTV와 PTI<sup>48)</sup>가 특정 구간을 초과하는 경우에는 연체율이 급격히 상승한다는 결과를 제시하였고, 주택자금대출 선진화를 위해서는 차입자의 상환능력이 중요하며 대출과정에서 차입자의 정보수집 강화가 필요하다고 제시하였다.

방두완·박세운·박연우(2010)는 생존분석 모형을 통해서 한국 모기지시장의 채무불이행 가능성을 분석하였으며 차입자의 신용등급이 낮을수록, 현시점의 LTV가 높을수록, DTI가 높을수록 채무불이행율이 증가하고 서울과 부산의 경우 주택가격상승률이 높을수록 채무불이행율은 감소한다는 분석결과를 제시하였다. 또한 차입자의 직업에서는 자영업자일 경우 채무불이행율이 더 높았다.

Ronel Elul et al.,(2010)은 모기지 부도에 영향을 미치는 요인을 분석하였으며 분석결과 모기지 채무불이행에 영향을 미치는 요인은 자산불건전성 보다는 실업과 같은 개인의 상황변화에 따른 비유동성이라고 제시하였다. 이러한 결과는 서브프라임 모기지와 같은 외부충격으로 인해서 주택담보대출이 부실화될 가능성이 높아진다는 점을 제시하였다.

강성진·이주석(2012)은 한국주택금융공사에서 시행한 2010년“주택금융 및 보금자리론 수요실태조사”를 활용한 주택금융의 수요자 특성에 관한 연구에서 주택을 구입한 시기는 2003년 전후 수도권과 아파트를 중심으로 주택담보대출을 활용하여 주택을 구입하였으며, 연령이 높을수록 주택담보대출 이용확

---

48) PTI(payment to income)=(월평균 원리금 상환액)/월평균소득\*100, 월평균 소득에 대한 주택자금대출의 월평균 상환지불능력임, 국내에서는 PTI를 연간소득대비 연간대출원리금 상환액을 의미하는 DTI개념으로 사용하고 있음.

률과 이용금액은 증가하며, 젊은세대, 미혼세대 일수록 주택담보대출의 이용 확률이 적었다. 또한 대출의 만기기간이 길수록, 원금상환은 유예하고, 거치기간이 길수록 대출의 규모가 큰 것으로 나타났다. 소득수준을 고려하지 않고 대출을 이용하고 있으며, 시세차액보다 담보대출의 규모가 크기 때문에 원금상환기간이 도래되거나, 만기가 도래할 경우 대출상환에 어려운 주택소유자가 증가할 가능성이 있다고 제시하였다.

방두완·박연우(2012)는 주택담보대출의 손실률을 추정한 주택담보대출 LGD<sup>49)</sup>를 실증분석한 결과 아파트담보물건의 평균 LGD는 30%라고 추정하였으며, 자산가치 대비 부채비율이 LGD를 설명하는 가장 중요한 변수로 제시하여 대출특성변수인 주택담보대출 부채비율이 높을수록 LGD에 가장 큰 영향을 미친다는 결과를 제시하였다.

이동걸 외(2013)는 우리나라의 부채보유 가구의 연체결정 요인에 관한 연구에서 보유가구의 연체를 주택담보와 신용대출로 구분하여 연체결정 요인을 추정하였으며, 주택담보의 경우 최초의 약정금액, LTV, 약정만기를 주택담보대출 속성 3가지로 뽑았다. 분석결과 약정금액과 LTV가 유의적인 관계로 분석되었으며, 소득수준별 연체결정 모형으로 추정한 결과는 LTV와 약정금액은 모든 소득군에서 유의한 결과를 가져온 반면 약정만기는 중·고소득군에서만 유의한 결과를 나타냈다.

---

49) LGD(loss given default) : 부도대출 손실률을 말함.

〈표 35〉 주택담보대출의 연체가능성에 관한 선행연구

저 자	내 용
지규현·김정인·최창규(2006)	주택담보대출 고객의 위험특성과 위험요인을 분석함. LTV와 PTI가 특정구간을 초과하는 경우 연체율이 급격히 상승하며 주택자금대출은 차입자 상환능력이 중요함을 제시함.
방두완·박세운·박연우(2010)	생존분석모형을 통하여 한국모기지 시장의 채무불이행 가능성을 분석함 저소득층의 소득감소가 상환부담에 영향이 더 큰 것으로 제시함.
Ronel Elul et al.,(2010)	모기지 부도에 영향을 미치는 요인은 자산불건정성 보다는 실업과 같은 개인의 상황변화에 따른 비유동성이며, 서브프라임 모기지와 같은 외부 충격이 주택담보대출의 부실화 가능성이 높다는 점을 제시함.
강성진 · 이주석 (2012)	주택금융의 수요자 특성에서 대출의 만기기간이 길수록, 원금상환은 유예하고, 거치기간이 길수록 대출규모가 큰 것으로 나타남. 시세차액보다 담보대출의 규모가 커서 원금상환기간과 만기가 도래할 경우 대출상환이 어려울 수 있음을 제시함.
방두완 · 박연우 (2012)	주택담보대출의 손실률을 추정한 주택담보대출 LGD를 실증 분석함. 대출특성 변수인 주택담보대출 부채비율이 높을수록 LGD에 가장 큰 영향을 미친다는 결과를 제시함.
이동걸 외(2013)	우리나라의 부채보유가구의 연체결정 요인을 추정함. 주택담보대출 모든 소득군에서 약정금액과 LTV가 유의적인 관계이며, 약정만기는 중·고소득군에서만 유의한 결과를 제시함.

#### 4) 선행연구와의 차별성

본 연구는 이미 다양한 분야에서 활용하고 있는 데이터마이닝 기법인 인공 신경망모형과 의사결정나무모형을 주택담보대출 연체가능성 분석에 활용하였다. 최적의 모형을 선별하기 위하여 전통적 회귀모형인 로짓모형과 비교함으로써 다양한 모형개발이 시도되었으며, 이를 통하여 주택담보대출 연체가능성 분석모형을 선별하였다.

주택담보대출의 연체가능성을 예측하기 위해서는 차입자의 데이터가 필요하나, 데이터의 접근 및 활용성은 제한적이라고 볼 수 있다. 본 연구의 자료는 제한적으로 허용된 범위에서 금융위기 전후 시점에서, 금융기관에서 발생한 주택담보대출 대상자를 분석대상 데이터로 활용하였으며, 이를 차입자의

고객특성과 주택담보대출 특성변수로 분류하여 연체가능성 원인을 구분하여 연구하였다. 또한 세가지 모형을 통하여 연체가능성에 미치는 영향과 요인을 분석하였으며, 연구의 적합성 및 모형별 차별화를 시도하였다.

주택담보대출 연체가능성 적용모형은 여러 가지 다양한 변수와 속성이 포함되고 대형화된 데이터를 통하여 모형개발이 발전적으로 이루어질 것으로 전망이 되고 있다. 본 연구는 연체가능성이 있는 데이터를 찾을 수 있고 연체가능성에 관련한 유사한 관계나 패턴을 탐색하여 이상적인 모형을 찾아내고 있다. 이러한 점은 주택담보대출의 연체자와 비연체자간의 두 집단 차이를 비교함에 있어서 유용한 분석방법으로 볼 수 있다. 주택담보대출 연체가능성분석에 데이터마이닝 기법인 인공신경망모형과 의사결정나무모형 적용은 의미가 있다고 본다.

또한 본 연구에 적용된 모형별 장점을 발견하였다. 로지스틱 회귀모형은 통계적인 측면에서 유의성 판단을 할 수 있다는 점이며, 인공신경망모형은 통계적 활용도보다는 선택변수를 다양하게 활용할 수 있다는 점이다. 의사결정나무모형은 주택담보대출 연체자와 비연체자를 단계별로 구분할 수 있으며, 이러한 연체가능성 분류기준을 도표로서 쉽게 확인 할 수 있는 점이다.

그리고 모형별 장점을 통하여 활용성을 발견하였다. 기존의 데이터마이닝 모형 연구는 <표 36>에서 볼 수 있는 것처럼 금융, 부동산, 자연과학, 경영학 등 여러 분야에서 연구가 이루어져왔으나 주택담보대출 연체가능성에 관련한 연구는 연체율과 거시경제변수와의 관련성을 시계열적으로 분석한 연구가 대부분이었으며 금융기관 등에서 적용하기에도 한계가 있었다. 그러나 본 연구에서 사용한 주택담보대출 연체가능성 모형은 금융기관 실무에 적용할 수 있다는 점이 기존 선행연구와 차별성이라 할 수 있다.

〈표 36〉 데이터마이닝 모형의 연구내용

연구모형	저자	연구내용	연구분야
인공신경망 모형	김경태 외(2001)	주가수익률 예측	금융
	정재호 외(2002)	신용카드 부정사용에 대한 예측	
	정유석(2003)	기업도산 예측	
	노태협 외 2인(2005)	KOSPI 200 주가지수 변동성에 대한 예측	
	최병남(1994)	토지평가시스템 구축	부동산
	김태훈 외(2004)	아파트의 특성변수 예측	
	정화미 외(2004)	개별공시지가 산정 예측	
	남영우외(2006)	거시경제변수와 주택가격과의 관계를 예측 분석함	
	정원구외2인(2007)	공동주택 가격지수 예측	
	우종덕 외(2011)	분양보증사고 사업장에 대한 예측	
의사결정나무 모형	변성호 외 3인(2008)	홍수에 영향을 미치는 요인을 분석	자연과학
	송영석 외 3인(2009)	급경사지에 대한 재해요인을 분석	
	이문주 외(2010)	고객경험을 근거로 테마파크평가 분석	
	조성빈(2010)	채용결정요인 분석	경영
	황진수 외 2인(2011)	고객의 패밀리레스토랑 선택요인 분석	
	박준호 외 2인(2011)	부동산 입지요인 분석	부동산
	최상일 외 2인(2011)	노년기 주거유형 선택 예측	

## 제 3 장 연구모형 및 분석자료

### 제 1 절 연체가능성 분석모형

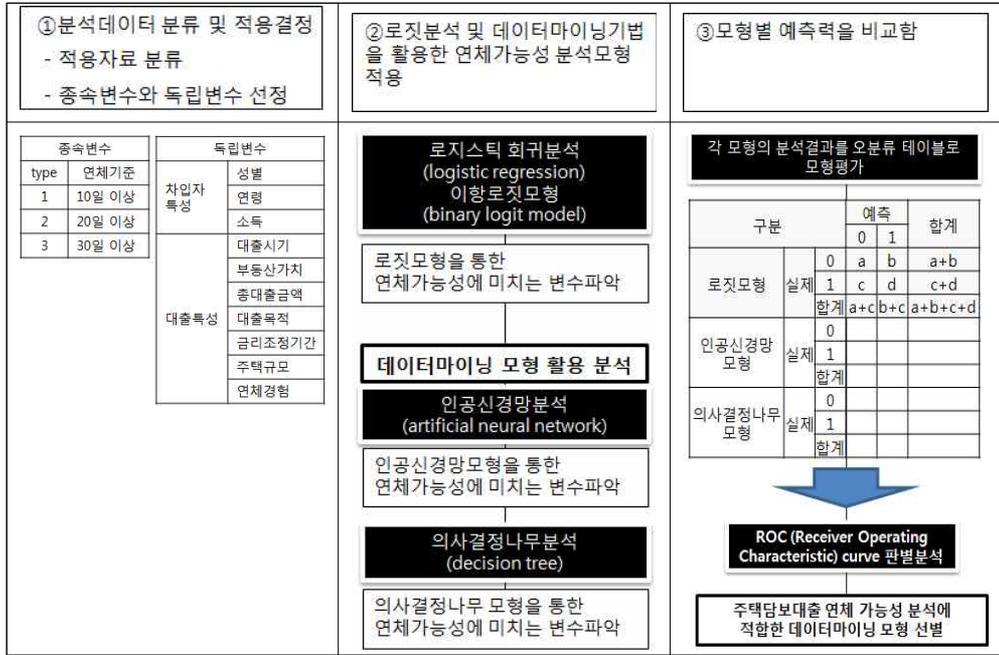
주택담보대출에 대한 거시경제 변수와 연체율을 분석하는 모형개발은 다양한 분석방법이 적용되어 왔으며 선행연구 등을 통해서 나타난 가장 일반적인 적용모형은 회귀분석모형으로 볼 수 있다.

본 연구가 주택담보대출 차입자의 연체가능성에 대한 분석이므로 종속변수를 연체와 비연체로 구분하였고, 이를 적용가능한 모형인 이항로짓모형(binary logit model)을 사용하였다. 독립변수로 차입자특성에는 성별과 연령, 소득을 반영하였으며 대출특성에는 대출기간, 대출시기, 부동산가치, 대출목적, 금리기간조정, 주택규모, 연체경험이 반영되었다.

이러한 모형은 시계열방식을 적용한 모형으로도 개발되고 있으며, 거시경제 변수와 연체율간의 관계를 분석하는 데 유용한 모형으로 평가받고 있다. 최근에는 데이터마이닝기법 등 여러 모형이 제시되고 있으나 주택담보대출 리스크관리에는 아직까지 구체적인 연구결과나 모형개발이 나타나지는 않고 있다.

본 연구에서는 회귀분석인 이항로짓모형과 함께 데이터마이닝기법인 인공신경망모형과 의사결정나무모형을 적용하여 주택담보대출 연체가능성에 미치는 변수를 파악해 보고 세가지 모형별 예측력을 비교하여 가장 우수한 분석모형을 선별하고자 한다.

〈그림 15〉 연체가능성 분석모형의 흐름도

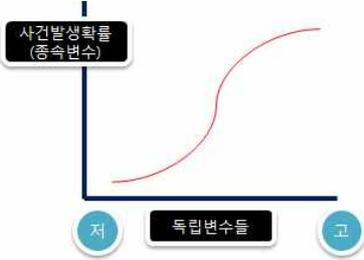
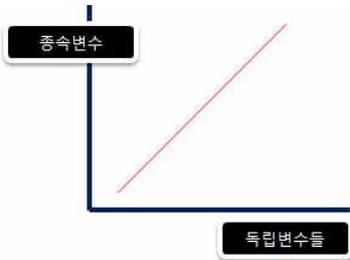


1) 로지스틱 회귀분석 모형

가) 모형의 특징

로지스틱 회귀분석(logistic regression)은 어떤 사건이 발생하는지 안하는지를 직접 예측하는 것이 아니라, 그 사건이 발생할 확률을 예측하며(박희정 외, 2009), 일반적으로 종속변수의 범주가 2개인 경우에 적용된다. 예를 들어 제품의 불량여부(양호/불량), 거래의 사기여부(정상/사기), 선거에서 특정후보에 대한 지지 여부(지지함/지지하지 않음), 시험에서 합격여부(합격/불합격)등을 들 수 있다. 종속변수가 2개의 범주를 가질 때 사용되는 회귀모형을 이항로짓모형(binary logit model) 또는 로짓모형(logit model)이라하며, 종속변수와 독립변수의 관계를 단순회귀분석과 다중회귀분석은 선형으로 가정하는 데 비해, 로지스틱 회귀분석은 S자형으로 가정한다.

〈그림 16〉 로지스틱 회귀분석 모형

로지스틱 회귀분석 모형	단순회귀분석과 다중회귀분석 모형
	
어떤 집단에 속할 것 인지 예측	다중 선형 회귀분석은 연속형 Y값을 예측

로지스틱 회귀분석은 분석척도로 종속변수에는 명목척도를 사용하며, 독립 변수에는 명목, 서열, 간격, 비율척도의 측정대상을 단순히 범주로 구분하기 위하여 사용한다.

#### 나) 모형의 설정

본 연구는 주택담보대출 연체가능성 여부에 대한 예측 또는 분류분석이다. 범주형인 이항로짓모형을 사용하여 종속변수에 '1'을 연체가 발생한 경우로 '0'을 비연체의 경우로 설정한 후 분석하고자 하는 변수들을 독립변수로 처리하여 적용함으로써 연체에 영향을 미치는 변수를 통계적으로 측정할 수 있다. 이러한 분석모형의 설정은 대출이 발생된 후 연체자와 그렇지 않은 비연체자인 2분형으로 구분하여 볼 수 있기 때문이다.

연체가능성 여부에 대하여 예측을 할 수 있는 독립변수는 차입자의 특성인 성별, 연령, 소득과 대출의 특성인 대출기간, 대출시기, 부동산가치, 대출금액, 대출목적, 금리조정기간, 주택규모, 연체여부를 변수로 선정하였다.

분석모형인 로짓모형은 종속변수와 독립변수 간의 관계를 다음과 같이 가정한다. 먼저 연체발생사건을  $y$ 라고 가정할 경우 연체발생에 영향을 미칠 수 있는 차입자특성과 대출유형에 대한 특성은 아래와 같이 독립변수인  $x$ 로 나타낼 수 있다.

$$y = \beta_1 X_1 + \beta_2 X_2 + \dots + \epsilon, \quad y = \begin{cases} 1 \\ 0 \end{cases} \quad (1)$$

1 : 연체가 발생한 경우

0 : 비연체의 경우

이러한 관계식으로부터 기대되는 사건  $y$ 가 일어날 확률  $\text{prob}(y=1)$ 은 식(2)와 같이 유도된다.

$$\begin{aligned} \text{prob}(y=1) &= \text{prob}\left(\sum_{k=1}^K \beta_k X_k + \epsilon > 0\right) \\ &= 1 - F\left(-\sum_{k=1}^K \beta_k X_k\right) \\ &= F\left(\sum_{k=1}^K \beta_k X_k\right) \end{aligned} \quad (2)$$

로짓모형은 연속확률분포함수의 형태로 정의되는 로짓분포를 연계함수로 가정하므로 독립변수들의 선형결합으로부터 사건  $y$ 가 일어날 확률을 식(3)으로 예측할 수 있다.

위의 식은 다음과 같은 Odds와 Log Odds를 유도할 수 있으며 Odds는 이항선택에 있어 한 사건에 일어나지 않을 경우 대비 사건이 일어날 경우의 확률비로 해석할 수 있다.

$$F\left(\sum_{k=1}^K \beta_k X_k\right) = \frac{1}{1 + e^{-\sum_{k=1}^K \beta_k X_k}} = \frac{e^{\sum_{k=1}^K \beta_k X_k}}{1 + e^{\sum_{k=1}^K \beta_k X_k}} \quad (3)$$

$$prob(y = 1) = F(\theta) = \frac{e^\theta}{1 + e^\theta} \quad (4)$$

$$e^\theta = \frac{prob(y = 1)}{1 - prob(y = 1)}$$

$$\log\left(\frac{prob(y = 1)}{1 - prob(y = 1)}\right) = \sum_{k=1}^K \beta_k \chi_k$$

$$(\text{단, } \theta = \sum_{k=1}^K \beta_k \chi_k)$$

따라서 식(2)와 식(4)를 이용하여 임의의 사건  $y$ 가 일어날 경우( $y=1$ )와 일어나지 않을 경우( $y=0$ )의 사건 확률형태는 식(5)와 식(6)으로 계산된다.

$$\begin{aligned} prob(y = 1) &= 1 - L\left(-\sum_{k=1}^K \beta_k \chi_k\right) \quad (5) \\ &= L\left(\sum_{k=1}^K \beta_k \chi_k\right) = \frac{e^{\sum_{k=1}^K \beta_k \chi_k}}{1 + e^{\sum_{k=1}^K \beta_k \chi_k}} \end{aligned}$$

$$\begin{aligned} prob(y = 0) &= 1 - prob(y = 1) \quad (6) \\ &= L\left(-\sum_{k=1}^K \beta_k \chi_k\right) = \frac{1}{1 + e^{\sum_{k=1}^K \beta_k \chi_k}} \end{aligned}$$

식(4)의  $\sum_{k=1}^K \beta_k \chi_k$ 는 종속변수의 형태가 응답(response)이 아닌 사건/시행(event/trial)의 형태로 입력될 수 있으며 일어날 사건이 이항반응범주 형태인 경우, 비율의 추론이란 목적에 한해 이용될 수 있다(송문섭·조신섭, 2004).

## 2) 인공신경망 모형

### 가) 모형의 특징

인공신경망모형(artificial neural network)<sup>50)</sup>은 인간의 두뇌를 인공적으로 모형화하여 만든 신경세포모형이다. 인간의 뛰어난 인식력과 판단력은 단순한 기능을 가진 신경세포조합에서 행하여진다는 사실이 밝혀지고, 이후 많은 학자들은 이 같은 인간의 신경세포조직을 인위적으로 만들 수 있는 연구방법을 고민하였고 그 결과 컴퓨터를 활용한 인공신경망모형이 탄생되었다.

이러한 인공신경망 모형은 인간의 두뇌구조를 모방한 지도학습 방법으로서 여러 개의 뉴런(neuron)<sup>51)</sup>들이 상호 연결하여 입력에 상응하는 최적의 출력값을 예측한다. 출력값은 수많은 반복적인 훈련과정을 통하여 데이터에 숨어 있는 패턴을 찾아내어 예측하는 기법이다. 또한 예측, 분류, 군집화 문제에 적용하는 범용 도구중 하나로서 언어 및 문자인식, 의학진단, 프로세스제어, 레이더 목표식별, 환율예측 및 사기거래 탐지와 같은 분야에 사용하고 있다.

최근 인공신경망모형의 활용도는 금융관련 분야에서도 다양하게 적용되고 있어 활용도가 높다고 할 수 있다. 금융업계는 데이터가 다양하고 많은 반면 고객의 금융거래 이외의 일반적인 행동에 관한 데이터는 규칙적으로 정리돼 있지 않기 때문에 인공신경망모형을 활용하여 금융거래 이외의 행동을 예측할 수 있다. 대출거래자에 대한 연체가능성분석과 금융사기의 탐색 및 예방, 신용카드 부정사용 파악, 고객 신용평가 등을 분석하는데 활용이 확대될 것으로 보인다.

---

50) 인공신경망이란 뇌기능의 특성 몇 가지를 컴퓨터 시뮬레이션으로 표현하는 것을 목표로 하는 수학 모델이다. 신경망이라고도 하며, 생물학이나 뇌 과학과 엄밀하게 구별하려는 경우에는 앞에 '인공'을 붙인다(구승환 외, 2009).

51) 뉴런(neuron)이란 신경세포로서 정보처리의 단위이다.

〈표 37〉 인공지능망 모형의 금융관련 적용분야

적용분야	적용방법	내용
금융 및 보험사기 적발	분류, 모형화	사기가능성 지수화 및 사기형태를 모형화
종합주가지수 예측	분류, 시계열모형	데이터 및 시계열모형을 분석하여 주가지수 예측
기업의 부도 예측	모형화	정상기업과 부도기업의 재무재표 내용 등을 변수로 예측모형 개발
개별기업 주식의 평가	모형화	개별기업의 개별적 특성을 맞춘 평가모형 개발
주식시장의 상황 파악	모형화	주식시장의 불안정지수 개발
대출기업의 신용평가	모형화	기업의 신용등급을 모형으로 산출 신용등급별 부실률과 점유율 예측
부동산가격지수 예측	모형화	검정 데이터(testing data)를 이용 예측력 평가
신용카드 부정사용 적발	모형화	정상사용자와 부정사용자의 회원정보와 가맹점정보 등을 변수로 모형화

나) 모형의 분석방법

제3장 제1절에서 로지스틱 회귀분석을 통하여 연체가능성 요인을 찾아내기 위해 모형방법을 설정하였고 이를 비교검증하기 위하여 인공지능망모형을 사용하였다. 인공지능망모형을 사용한 이유는 연체가능성 분석모형의 다양한 시도와 더불어 인공지능망모형과 로지스틱회귀모형 간 차이를 통하여 예측력이 우수한 모형을 선별하고자 한다.

앞장에서 로지스틱회귀모형은 종속변수인 연체사실에 대하여 연체여부를 판별하는데 효과적이라고 생각되는 여러 개의 독립변수들의 값을 가지고 연체자들의 특성을 판별함수로 만들었으며, 이러한 값이 연체여부에 속하는지를 판별하는 모형으로 사용되었다.

인공신경망구조는 역전파(back propagation) 알고리즘 적용이 가능한 구조로서 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)이 서로 연결된 다층 퍼셉트론(multi-layer perceptron)모형을 가지고 있으며, 각 계층에는 여러 뉴런들이 서로 연결되어 주어진 연체가능성 문제를 학습하고 추론하는 기능을 수행하는 구조를 가지도록 하였다.

인공신경망의 핵심은 뉴런 간 연결과 연결강도(connection weights)와 이에 따른 다계층의 구조에 달려 있으며, 각 뉴런은 주어진 학습자료를 학습하는 기능과 상호 연결된 또 다른 뉴런에 그 처리결과를 보내는 기능이 존재한다. 이러한 전달기능을 위하여 학습시 사용되는 함수는 활성화함수<sup>52)</sup>로서, 가장 많이 사용되고 있는“S”자형 시그모이드 함수<sup>53)</sup>를 사용하였으며, 은닉층과 출력층의 개수를 달리하여 실험하였다.

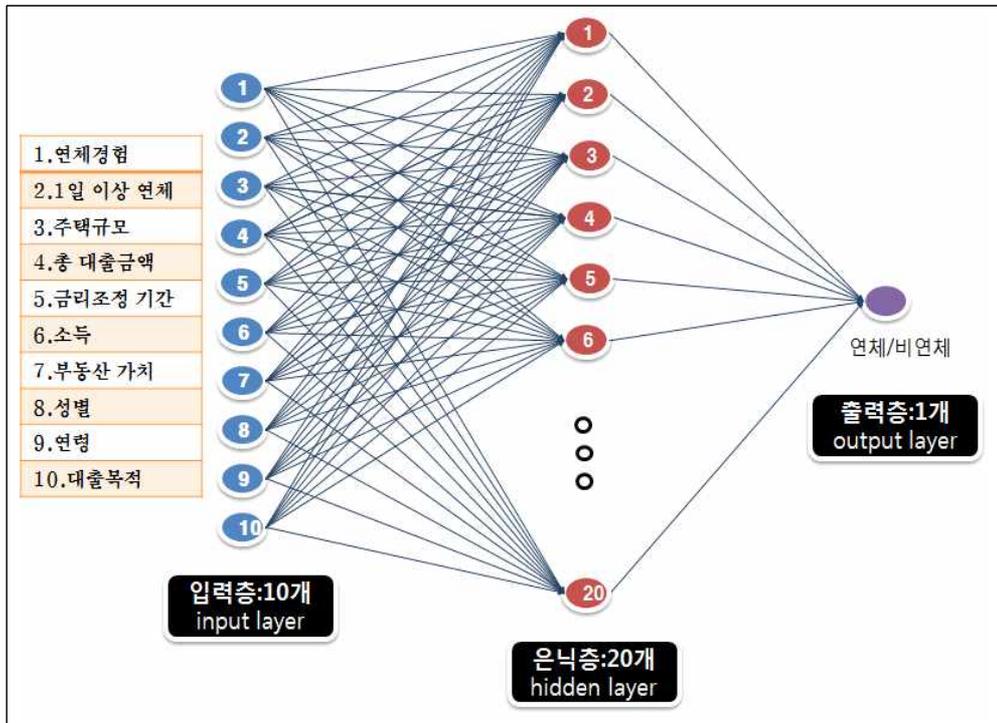
본 연구에 사용된 인공신경망 연체가능성 예측 모형은 입력층이 10개이며, 각층의 은닉노드가 20개인 다층 퍼셉트론 알고리즘을 적용하였으며 아래의 <그림 17>과 같다.

---

52) 결과 값을 내보낼 때 함수로 전이함수라고도 한다.

53) 시그모이드 함수는 단극성, 양극성 두 종류가 있다. 단극성 시그모이드 함수는 증가함수이며 출력값이 0과 1사이의 값을 갖고 로지스틱함수와 유사하다.

〈그림 17〉 다층 퍼셉트론 인공지능망의 구조



#### 다) 모형의 적용사례

인공신경망모형은 초기 인공지능의 개발을 위하여 공학분야에서 발전하였으나 경영학분야에서 많이 적용되어 활용되고 있다.

특히 금융분야에서 적극적으로 연구가 진행되어, 김경태 외(2001)는 추가수익률을, 정재호외(2002)는 신용카드 부정사용에 대한 예측을 위해서 신경망분석을 적용하였다. 정유석(2003)은 인공신경망모형을 이용하여 기업도산을 예측하는 분석을 실행하였으며, 노태협 외 2인(2005년)은 신경망모형과 금융시계열을 이용한 KOSPI 200 추가변동성에 대한 예측모형 분석결과를 제시하였다. 분석결과 변동성의 방향성 예측면에서 금융시계열의 GARCH 모형이 인공신경망 모형보다 우수한 성과를 나타내었으며, 반면 인공신경망모형은 변동의 예측정확성면에서 더 우수한 결과를 제시하였다.

부동산분야에서도 인공신경망모형이 적용된 연구가 진행되었다. 최병남

(1994)은 토지에 대한 각종 정보를 통해 신경망모형을 활용하여 토지평가시스템을 구축하였으며, 정확미 외(2004)는 당시 건설교통부의 개별공시지가 산정에 신경망분석을 응용하는 방법을 제시하였다. 주택에 대한 연구로는 김태훈 외(2004)가 아파트의 특성변수를 이용하여 회귀분석과 신경망모형을 구축하여 이를 비교하는 연구를 실행하였으며 남영우 외(2006)는 거시경제변수와 주택가격과의 관계를, 정원구 외(2007)는 공동주택 가격지수 예측에 관한 연구로 거시경제지표변수와, 서울, 강북, 강남의 공동주택가격지수를 사용하였으며 인공신경망모형을 구축하여 공동주택가격지수를 예측하여 분석하였다. 우종덕 외(2011)는 분양보증시 사고사업장에 대한 예측모형 개발을 신경망모형을 적용하여 분석하였다.

### 3) 의사결정나무 모형

#### 가) 모형의 특징

의사결정나무는 의사결정 규칙을 나무구조로 도표화하여 분류와 예측을 수행하는 분석방법이다. 이 방법은 분류 또는 예측과정이 나무구조에 의한 추론 규칙에 의해서 표현되기 때문에 분석자가 그 과정을 쉽게 이해하고 설명할 수 있다는 장점이 있다(홍성두 외, 2008). 따라서 판별분석 또는 회귀분석, 신경망분석 등과 같은 분석에 있어서 사전에 이상치를 확인하거나 분석에 필요한 변수를 찾아내는데 이용될 수 있고 자체적으로 예측모형을 개발하는데도 활용될 수 있다 또한 이러한 분석들에 비해 연구자가 분석과정을 쉽게 이해하고 설명할 수 있다는 장점을 가지고 있다.<sup>54)</sup>

의사결정나무는 하나의 나무구조를 이루고 있으며 마디(node)라고 불리는 구성요소들로 이루어져 있다. 즉, 나무구조가 시작되는 마디로서 분석대상이 되는 전체자료로 이루어진 뿌리마디(root node), 하나의 마디로부터 분리되어 나간 2개 이상의 마디들을 의미하는 자녀마디(child node), 각 자녀마디들의 상위마디를 의미하는 부모마디(parent node), 각 나무줄기의 마지막 마디를

54) 최중후·소선하(2005), 『사례로 배우는 데이터마이닝』, 제1판, 자유아카데미, p.25.

의미하는 최종마디(terminal node), 나무구조의 중간에 위치하고 있는 중간 마디(internal node), 하나의 마디로부터 마지막 마디까지 연결된 일련의 마디들을 의미하는 가지(branch)로 구성되어 있다. 그리고 전체 나무가지를 이루고 있는 마디의 개수를 깊이(depth)라고 한다.

이러한 구조는 분리의 기준에 따라 중요도가 가장 높은 변수부터 값에 따라 나누게 되며 어떤 예측변수를 이용하여 어떻게 분리하는 것이 목표변수의 분포를 가장 잘 구별해주는지를 파악하여 자식마디가 형성된다. 그리고 더 이상 분리가 일어나지 않고 현재의 마디가 끝 마디가 되도록 지정하는 규칙을 정지기준 이라한다. 분리가 일어나면서 가지치기가 발생하며 지나치게 많은 마디는 예측오차가 커질 우려가 있어 이러한 경우에는 적절하지 않은 마디를 제거하여 적절한 크기의 나무구조를 가지는 모형을 선택한다.

〈표 38〉 의사결정나무 모형의 분석단계

세분화 (segmentation)	관측개체를 비슷한 특성을 갖는 몇 개의 그룹으로 분할하여 각 그룹별 특성을 발견하고자 하는 경우
분류 (classification)	여러 예측변수에 근거하여 목표변수의 범주를 몇 개의 등급으로 분류하고자 하는 경우
예측 (prediction)	자료로부터 규칙을 찾아내고 이를 이용하여 미래의 사건을 예측하고자 하는 경우
차원축소 및 변수선택 (data reduction and variable screening)	매우 많은 수의 예측변수 중에서 목표변수에 큰 영향을 미치는 변수들을 골라내고자 하는 경우
교호작용 효과의 파악 (interaction effect identification)	여러 개의 예측변수들이 결합하여 목표변수에 작용하는 교호작용을 파악하고자 하는 경우
범주의 병합 또는 연속형 변수의 이산화 (category merging and discretizing continuous variable)	범주형 목표변수의 범주를 소수의 몇 개로 병합하거나, 연속형 목표변수를 몇 개의 등급으로 범주화하고자 하는 경우

자료 : 최종후·서두성(1999), 『데이터마이닝 의사결정나무의 응용』, 통계청, 통계분석연구, pp.62-63.

의사결정나무분석은 탐색과 모형화라는 두 가지 특성을 모두 가지고 있다고 할 수 있고 자료의 세분화와 분류 및 예측은 탐색의 결과를 통한 모형화 단계에서 나타난다. 일반적으로 의사결정나무분석은 <표 38>과 같은 단계를 거친다(최종후·서두성, 1999).

#### 나) 모형의 분석방법

로지스틱회귀분석과 인공신경망분석을 통하여 연체가능성 모형을 설정하였고, 이번에는 두가지 모형을 비교검증하기 위하여 데이터마이닝모형인 의사결정나무모형을 추가하여 연체가능성 예측을 비교해보았다. 의사결정나무모형을 분석한 이유는 연체가능성모형의 다양한 시도 측면과 전통적 회귀분석모형인 로짓모형과 데이터마이닝모형인 인공신경망모형과도 차이를 발견하여 최종적으로는 3가지 모형의 차이점에서 예측력이 우수한 모형인지를 발견하고자 한다. 의사결정나무모형이 가지고 있는 사용 분야의 특징을 통하여 연구자료를 세분화, 분류, 예측, 변수선택에 적용하여 분석하였다.

<표 39> 의사결정나무 모형의 적용

세분화 (segmentation)	주택담보대출자 데이터를 세분화하여 특성집단을 파악하고 고객을 세분화 한다
분류 (classification)	주택담보대출 데이터에 근거하여 종속변수의 범주를 2개의 종류로 연체 및 비연체로 분류 한다
예측 (prediction)	주택담보대출 자료로부터 규칙을 찾아내고 이를 이용하여 연체가능성을 예측하고 대출고객의 속성에 따라서 연체자와 비연체자를 예측 한다
변수선택 (variable selection)	주택담보대출실행에 발생된 여러 가지의 독립변수들 중 이들 중에서 유의한 영향을 주는 변수를 발견 한다

본 연구에서 사용된 의사결정나무 분석방법 알고리즘은 C5.0을 사용하였으며 엔트로피(entropy)통계량을 분리기준으로 삼는다.<sup>55)</sup> 엔트로피지수(entropy

index)라고 하며 오분류율을 불순도의 측도로 사용하여 의사결정나무모형을 형성하게 되며 회귀모형과 동일한 방식으로서 의사결정나무 성장단계에서 최적의 분리를 할 때 사용된다.

모형의 분리기준은 연체가능성이 가장 높은 변수부터 값에 따라 나뉘게 되고 이에 따라 의사결정나무 분석은 구축모형과 트리 그리고 Rule set이 작성되어지는 알고리즘이다. 이러한 방법은 범주형 속성 전체에 대해서 하위 노드 (node)를 생성하기 때문에 선택된 속성의 범주값이 많을 경우 하위 노드 가지(branch) 수가 많아지는 단점을 가지고 있다.

〈표 40〉 의사결정나무 분석의 알고리즘 비교

구 분	C5.0	CHAID	CART	QUEST
입력변수에 의한 분리형식	다중분할	다중분할	이진분할	이진분할
연속형 목표변수	처리가능	처리불가	처리가능	처리가능
나무구조 생성시 오분류 비용 사용여부	사용	사용안함	사용	사용안함
결측치 처리	하나의 범주로 처리	하나의 범주로 처리	결측치 대체	결측치 대체
사전확률 사용여부	사용안함	사용안함	사용	사용

자료 : 최종후·소선하(2005), 『사례로 배우는 데이터 마이닝』, 자유아카데미, p.57.

의사결정나무모형 구축은 처음 단계에서 먼저 차입자 548개 데이터가 연체자와 비연체자로 구분되고, 연체여부에 대하여 다시 연체자층과 그렇지 않은 비연체자층으로 나뉘게 된다. 이렇게 차례대로 구분된 층이 의사결정나무 구조모형으로 구축된다.

이러한 모형은 연체가능성에 중요도가 가장 높은 변수부터 값에 따라 차례대로 나누어지며 의사결정나무 분석에 사용된 변수는 〈표 41〉과 같다.

55) 최종후·소선하(2005), 전계서, p.56.

〈표 41〉 의사결정나무 분석에 사용된 변수

변수	측정형태	변수값
성별	명목형	차입자 성별 더미 (남성=0, 여성=1)
연령	연속형	차입자 연령
소득	연속형	차입자 소득(백만원)
대출시기	명목형	2007년, 2008년, 2009년, 2010년
부동산 가치	연속형	담보대상 부동산 가치(천원)
총 대출금액	연속형	대출실행 시 총대출(천원)
대출목적	명목형	대출목적 더미 (구입자금=0, 가계자금=1)
금리조정 기간	명목형	금리조정기간 더미 (6개월이상=0, 6개월미만=1)
주택규모	연속형	대출시점 구입 주택규모(m <sup>2</sup> )
연체경험	명목형	1회 이상 연체 여부 더미 (있음=1, 없음=0)

#### 다) 모형의 적용사례

의사결정나무모형은 데이터마이닝 모델 중 예측력이 우수하고 분류기준에 대한 정보를 제공해주기 때문에 다양한 분야에서 연구가 진행되고 있다.

먼저 자연과학분야에서 변성호외 3인(2008)은 홍수에 영향을 미치는 요인들을 분석하여 구조적 홍수방어대책에 대한 모형을 구축하여 제시하였으며, 송영석 외 3인(2009)은 급경사지에 대한 재해요인을 분석하여 사면경사와 투수계수, 간극비 등의 요인이 재해가능성을 높이는데 중요한 요인이 될 수 있다는 결과를 제시하였다.

경영학분야에서는 주로 대안의 평가 및 고객의 선택요인에 대한 분석을 위해서 의사결정나무모형이 활용되었다. 이문주·한준수(2010)는 고객경험을 근거로 한 테마파크 평가를 위해서 의사결정나무모형을 사용하여 평가의 가장

중요한 요인으로 안전성이 선택된 결과를 제시하였다. 황진수 외 2인(2011)은 고객의 패밀리레스토랑 선택요인에 대해서 분석하였다. 분석결과에서 레스토랑이 가지고 있는 특성이 중요한 선택요인으로 제시되었고 레스토랑 선택속성이 고객의 인구통계속성과 이용행태와 밀접한 관련이 있는 것으로 나타났다. 조성빈(2010)은 채용결정요인 분석에서 성별, 나이, 대학, 전공, 평점, 토익, 실무면접 등이 채용에 어떠한 영향을 주고 있는 지를 확인하였다. 분석결과 남성지원자와 여성지원자의 채용에 차이가 존재하고 있음을 확인할 수 있었고, 출신대학과 토익 등이 채용에 주요한 요인으로 분석되었다.

부동산관련 분야에서는 입지요인 분석과 주거선택분야에서 의사결정나무모형을 적용한 연구가 실행되었다. 박준호 외 2인(2011)은 벤처기업의 입지특성을 분석한 결과 벤처기업의 입지는 종업원 1인당 매출과 업종특성 등에 따라서 다른 특성을 나타낸다고 제시했다. 최상일 외 2인(2011)은 노년기 주거유형 선택을 예측하기 위한 모형구축을 위해서 의사결정나무모형을 적용하였으며, 노후준비금액의 규모와 현재의 직업, 노년기 자녀와의 주거인식 등에 따라서 선호하는 주거입지가 달라지는 것으로 나타났다.

## 제 2 절 연체율 산정기준과 연체율 현황

연체율은 전체대출에 대한 상환의무 위반대출 비율로 가계자산 대비 부채비율 및 처분가능소득 대비 부채비율과 함께 가계부채의 건전성을 판단하는 주요지표 중의 하나이다.

특히, 가계부채 연체율 증가는 직접적으로 금융회사의 부실을 초래하고, 또한 금융회사의 부실은 실물 경제에 부정적 영향을 미치게 되므로 가계부채가 과다한 우리나라의 경우 연체율 추이가 중요한 의미를 지닌다. 따라서 가계부채 연체율에 선행하는 주요 변수를 찾아 면밀히 모니터링하는 등 선제적으로 대응할 필요가 있다(김석기, 2013).

이러한 가계대출 연체율의 증가는 가계대출 비중이 높은 주택담보대출 연체율도 동반하여 상승하는 현상을 보일 수 있어 주택담보대출 연체율에도 변

수를 찾아 모니터링하고 선제적 대응이 필요하다고 본다.

### 1) 산정 기준

현재 우리나라에서는 금융감독원에서 공식적으로 연체율을 산정하여 발표하고 있다. 연체율은 산정하는 기준에 따라서 결과가 다르게 나타날 수 있으므로 산정기준은 연체율 산정의 중요한 요소가 된다.

현재 금융감독원에서 발표하고 있는 연체율은 2014년 3월 이후부터 새롭게 변경된 기준을 적용하고 있다. 변경 전 연체율 산정기준은 2007년 이후부터 변경된 금융감독원의 감독기준에서 정한 연체율의 산정기준과 다르게 적용되면서 혼란이 가중되어 이를 개선하였다.

2007년 이후 은행의 감독기준은 종전 1일 이상 원금 연체시 연체채권으로 분류하던 것을 1개월 이상 원금 또는 이자 연체시 연체채권으로 분류토록 변경하여 시행하고 있으나 연체율에 대한 보도자료 발표시에는 과거 시계열과의 일관성 유지 등을 위해 종전기준인 1일 이상 원금 연체시 연체채권으로 분류하는 연체율 산정 방식이 유지되어 왔다.<sup>56)</sup> 하지만 금융위기 이후 상승했던 국내연체율이 안정되고 기존 방식이 국제기준에도 부합하지 않아 2014년 3월 이후부터는 감독기준 상 연체율기준을 적용하여 발표하고 있다.

기존방식은 이자가 연체될 경우에는 원금 전체를 연체로 보지 않고 원금이 1일 이상 연체되는 시점부터 동 원금을 연체로 분류하는 '1일 이상 원금기준' 방식이 적용되었다. 다만, 이자를 지급해야 할 때부터 일정기간 지체(가계여신 1개월, 기업여신 14일)한 경우에 한해서는 원금 전체를 연체로 분류하였다.

이에 반해 새로운 기준에서는 원금 또는 이자 중 어느 하나라도 연체되는 시점부터 원금전체를 연체로 분류한다. 다만, 연체기간이 1개월 미만인 경우 연체율 산정시 이를 연체대출채권 범주에서 제외하여 연체율산정에서는 '1개월 이상 원리금기준' 방식을 적용하고 있다.

---

56) 금융감독원 보도자료(2014), 『국내은행 대출채권 연체율 보도자료 발표시 산정기준 변경』 pp.1-9.

해외의 경우를 살펴보면 미국, 영국 및 홍콩 등의 경우에도 이자 또는 원금 중 하나라도 연체되는 경우 원금 전체를 연체로 간주하고 있으며, 다만 연체율 산정시에는 30일 미만 연체대출채권은 제외<sup>57)</sup>하고 있으므로 새로운 기준은 국제기준에도 부합하는 것으로 판단된다.

〈표 42〉 금융감독원 연체율 산정기준

구분	1일 원금 기준 (2014년 3월 이전)	1개월 원리금 기준 (2014년 3월 이후)
원금	약정기일에 상환되지 않으면 익일부터 연체채권으로 계상	약정기일에 상환되지 않으면 익일부터 연체채권으로 계상하되, 연체율 산정시에는 1개월 미만 연체채권은 제외
이자 (기업대출)	연체기일 14일 초과시 해당 원금전체를 연체채권으로 계상	약정기일에 상환되지 않으면 익일부터 해당 원금전체를 연체채권으로 계상하되, 연체율 산정시에는 1개월 미만 연체채권은 제외
이자 (가계대출)	연체기일 1개월 초과시 해당 원금전체를 연체채권으로 계상	약정기일에 상환되지 않으면 익일부터 해당 원금전체를 연체채권으로 계상하되, 연체율 산정시에는 1개월 미만 연체채권은 제외
분할상환채권	약정기일에 상환되지 않으면 2회차 미납일부터 해당 원금전체를 연체채권으로 계상	약정기일에 상환되지 않으면 익일부터 해당 원금전체를 연체채권으로 계상하되, 연체율 산정시에는 1개월 미만 연체채권은 제외

자료 : 금융감독원, 『국내은행 대출채권 연체율 보도자료 발표시 산정기준 변경』, 2014.5.7. p.9.

57) 상계서, p.2.

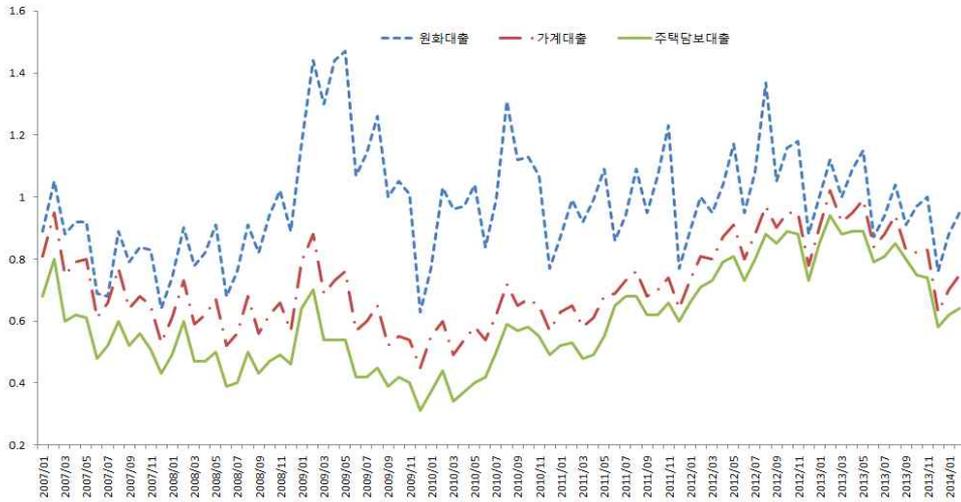
## 2) 연체율 현황

국내은행의 연체율은 거시경제변수와 정부정책 등 다양한 변수에 영향을 받으므로 시기별로 변동성을 가지고 있다. 2007년 우리나라의 원화대출을 부문별로 살펴보게 되면 가계대출의 증가에 따른 우려에도 불구하고 전체 원화대출의 연체율에 비해서 가계대출의 연체율이 낮게 나타나고 있으며, 주택담보대출은 가계대출 중에서도 낮은 연체율을 나타내고 있다.

이러한 현상은 가계대출에 비해서 상대적으로 기업대출이 경기변수에 대한 민감도가 높은 것으로 보여지며, 주택담보대출은 거주하고 있는 주택을 담보로 제공하는 대출특성 상 가계대출 중에서도 연체율이 낮게 나타나고 있는 것으로 판단된다.

시기별로 살펴보면 전반적으로 2007년까지는 국제적인 유동성 확대에 금리가 점차 낮아지는 추세를 보였으나 2008년 금융위기 이후 큰 폭으로 상승하는 추세가 나타나고 있다. 특히 가계대출에 비해서 전체원화대출의 상승폭이 매우 크게 나타나고 있으며 기업대출의 연체율이 매우 높게 나타난 것을 확인할 수 있다.

〈그림 18〉 국내 원화대출 부문별 연체율 현황



자료 : 금융감독원 보도자료, <http://www.fss.or.kr/>.

금융위기 이후 국제적으로 경기회복을 위해서 통화확장 정책을 실시하면서 금리가 낮아지고 다시 연체율도 낮아지기 시작했으나 2010년부터 2012년까지 가계부채를 중심으로 연체율이 다시 상승하였다. 이는 급격하게 증가한 가계부채규모와 경기침체가 원인이 된 것으로 보인다.

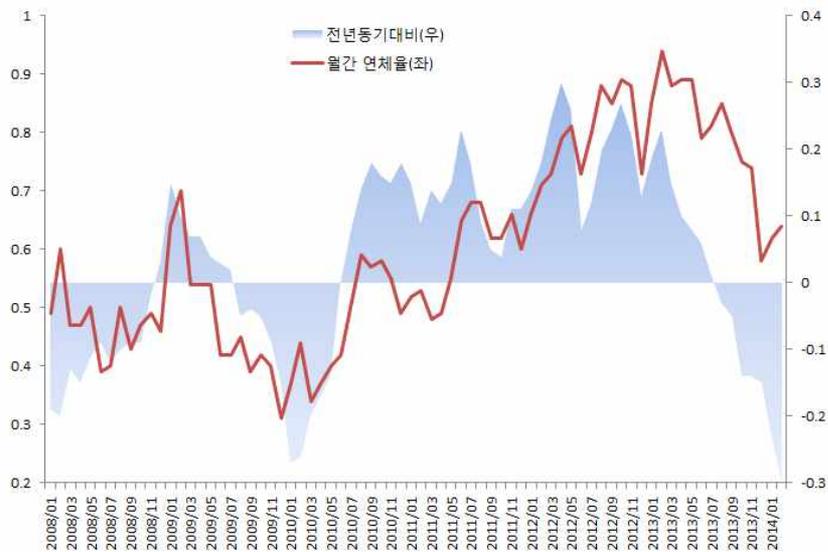
하지만 정부에서 지속적으로 가계부채 관리정책을 실행하면서 가계부채의 연체율도 다시 하락세로 전환되고 있다. 그러나 가계부채 및 주택담보대출의 연체율과 전체 원화대출연체율 간의 간극은 상당부분 축소되어 있어 앞으로 도 지속적인 관리가 필요한 상황으로 판단된다.

주택담보대출의 연체율에 대한 전년 동기대비 연체율과의 비교를 살펴보면 시기별 변화방향을 더욱 구체적으로 살펴볼 수 있다. 매월 연체율은 시기별 추세에 따라서 상승과 하락세를 지속하기 때문에 기간별 비교 시에는 전년 동기대비 연체율변화가 중요한 정보를 제공할 수 있다. 2008년 이후 전년 동월대비 주택담보대출 연체율을 살펴보면 금융위기 발생시기에는 약 10개월 동안 전년 동월대비 연체율이 높게 나타났다. 이에 비해 2010년 이후에는 약 3년 동안 전년 동월에 비해서 연체율이 높게 나타났다.

따라서 국내 주택담보대출의 연체율 상승은 미국발 금융위기의 영향보다는

가계부채의 증가 및 경기침체 등에 더 큰 영향을 나타내고 있는 것으로 판단된다. 즉 시장의 추세를 제거하게 되면 금융위기의 영향을 벗어난 이후에 상당기간 연체율이 상승한 것으로 판단할 수 있다. 2013년 하반기 이후 전년 동월대비 연체율도 하락세에 있으나 정부의 부동산 경기활성화 정책 등이 지속되고 있으므로 연체율의 추이에 대한 점검의 중요성이 대두되고 있다.

〈그림 19〉 주택담보대출 연체율 현황



자료 : 금융감독원 보도자료, <http://www.fss.or.kr/>.

### 제 3 절 분석자료와 변수

#### 1) 분석자료

주택담보대출이 현황에서 나타난 바와 같이 금융기관에서도 체계적인 연체 관리에 대한 필요성이 높아지고 있다. 주택담보대출에 대한 연체율예측을 선

행연구에서는 거시경제변수를 이용하여 전체 주택시장을 대상으로 분석하는 방법과 차입자 및 대출특성을 분석하는 모형으로 구분하였다. 거시경제변수를 통한 분석의 경우 금융정책을 수립하거나 전체 금융기관에 대한 건전성을 관리하는 측면에서는 유리하나 개별적인 금융기관에서 적용하는 데는 한계가 있다.

따라서 본 연구에서는 금융기관의 주택담보대출 자료를 활용하여 연체가능성을 분석하고자 한다. 금융기관의 연체율에 대한 효과적 관리는 대출에 대한 실제 자료를 통해서 모형을 구축하고 이를 분석하는 작업이 필요하며, 연체율에 대한 예측모형개발을 위해서는 주택담보대출 사례를 통해 차입자특성과 대출특성에 대한 자료가 구축되어야 한다.

이를 위해서 본 연구에서는 금융기관의 2007년부터 2010년까지 4년간 대출데이터를 통해서 분석을 실행하였다. 입력자료의 이상치와 결측치 등을 고려하여 분석에 실질적으로 사용된 자료의 수는 548개이며 분석대상의 정형화를 위하여 아파트에 대한 대출사례를 활용하였다. 시기적으로는 대출 후 일정 시간 이후의 연체율을 측정하기 위해서 대출자료를 분석모형에 적용하였다.

분석대상의 주택담보대출자료 특성은 8가지로서 지역, 성별, 대출실행연도, 대출목적, 대출기간, 대출금리, 직업으로 분류하였다. 지역적 특성에서는 서울 지역에서 대출이 이루어져 대부분 수도권지역에 소재하였으며 차입자의 성별은 남자가 79.4%의 비중으로 다수가 차지하고 있으나 여자의 비중도 20.6%로서 적지 않은 비중을 보이고 있다.

대출실행 된 건수는 2008년에 가장 낮게 나타났으며 2007년과 2009년에 실행된 담보대출의 비중이 상대적으로 높게 나타났다. 2008년의 대출건수 감소는 금융위기 이후 주택시장이 침체상태에 빠지면서 영향을 준 것으로 파악된다.

대출목적이 구입자금 용도보다는 가계자금으로 활용된 용도가 더 많아 아파트의 구입보다는 아파트를 담보로 가계자금을 활용하는 비중이 높다고 판단되며, 대출기간은 대출만기가 16년~20년 비중이 전체의 41.79%를 차지하였고 15년을 초과하는 비율은 81.43%로서 높은 비중이긴 하나, 5년이내의 단기대출도 8.9%로서 적지 않은 비율을 가지고 있다.

대출금리 선택은 전체의 73.2%가 3개월 CD를 선택하여 높은 비중을 보이고 있으며 6개월 CD, 코픽스 순으로 취급되었다. 고정금리의 비중이 불과 0.4%에 불과하여 CD금리가 기반이 되어 연동된 변동금리의 편중현상이 두드러져 있었다. 차입자의 직업은 회사원이 81.7%로서 회사가 밀집된 도심소재의 금융기관에서 발생된 자료로서 구체적인 대출관련 자료의 특성 및 항목별 비중은 다음 <표 43>와 같다.

〈표 43〉 분석대상 주택담보대출 자료의 특성

구분		빈도	비중(%)
지역	서울	357	65.1
	경기	167	30.5
	인천	22	4.0
	기타	2	0.4
	합계	548	100.0
성별	남	435	79.4
	여	113	20.6
	합계	548	100.0
대출실행 연도	2007년	148	27.0
	2008년	91	16.6
	2009년	186	33.9
	2010년	123	22.4
	합계	548	100.0
대출목적	구입자금대출	213	38.9
	가계자금대출	335	61.6
	합계	548	100.0
대출기간	5년 이내	49	8.9
	6~10년	21	3.83
	11~15년	86	15.7
	16~20년	229	41.79
	20년 초과	163	29.74
	총계	548	100
대출금리	3개월 CD	401	73.2
	6개월 CD	85	15.5
	cofix 6개월	60	10.9
	고정금리	2	0.4
	합계	548	100.0
직업	회사원	447	81.6
	자영업	50	9.1
	주부	25	4.6
	공무원	11	2.0
	기타	15	2.7
	합계	548	100.0

모형 구축을 위해서 대출 및 차입자특성에 대한 기술적 통계분석을 실행하였다. 분석결과 평균 주택규모는 80.72m<sup>2</sup>였으며 차입자 연령은 약 37세, 평균대출기간은 20년으로 나타났다. 또한 분석된 대상 물건지가 수도권으로 대출당시 주택의 가격은 약 425백만원으로 일반지역보다 상대적으로 높았으며 평균 대출액은 약 97백만원, 차입자의 평균 연소득은 54백만원으로 나타났다. 연체관련 정보에서는 1일 이상 연체횟수는 평균 2.7회이고 연체일수는 평균 11.8일로 나타났다.

〈표 45〉 분석모형의 변수

종속변수(산출변수)		10일, 20일, 30일 이상 연체발생 여부 (연체발생=1,미발생=0)	
독립변수 (투입변수)	차입자 특성	성별	차입자 성별 더미 (남성=0,여성=1)
		연령	차입자 연령
		소득	차입자 소득(백만원)
	대출 특성	대출시기	2007년~2010년
		부동산 가치	담보대상 부동산 가치
		총 대출금액	대출실행 시 총대출
		대출목적	대출목적 더미 (구입자금=0, 가계자금=1)
		금리조정 기간	금리조정기간 더미 (6개월이상=0, 6개월미만=1)
		주택규모	대출시점 구입 주택규모(m <sup>2</sup> )
		연체경험	1회 이상 연체 여부 더미 (있음=1, 없음=0)

## 2) 분석변수

연체율 예측모형 개발에 필요한 로짓분석과 인공신경망분석을 실행하기 위해서 분석에 적용되는 변수를 선정하였다. 분석변수는 연체발생 여부를 종속변수와 산출변수로 하고 변수 간의 상호관계와 변수의 특성 등을 고려하여 차입자 특성과 대출 특성 관련 변수를 독립변수와 투입변수로 선정하였다.

종속변수로는 연체발생 여부를 결정하는 기준을 사용해야 하므로 전체 발생 연체일이 20일 이상인 경우<sup>58)</sup>를 기준으로 하되 비교분석을 위해서 단기인 10일 이상 연체와 장기인 30일 이상 연체의 경우를 함께 분석하였다.

연체 후 20일 이상을 기준으로 할 경우 연체가능성이 높은 가구는 분석데이터기간(2007년~2010년)의 전체 대출가구 중 12.8%가 해당하게 된다. 연체 발생기준을 현 시점의 상환불능가구가 아닌 전체 누적연체일을 기준으로 선정한 이유는, 사후적인 결과뿐만 아니라 사전적인 연체 가능성까지 고려했기 때문이다.

다음으로 로짓모형의 독립변수와 인공신경망모형의 투입변수는 차입자특성인 성별, 연령, 소득을 선정하였다. 투입변수로 선정한 이유는 성별의 경우 주택담보대출 여성대출자의 증가로 분석자료의 20.6%가 해당되어 차입자의 성별에 따른 연체가능성에 차별화가 나타나는지에 대해서 분석할 필요성이 높기 때문이다.

또한 지속적인 생산가능인구의 감소와 고령인구의 증가로 인해서 차입자의 연령을 분석대상 변수에 포함하였으며 소득정도에 따른 연체가능성을 예측하기 위해서 소득도 분석변수에 포함하였다.

대출관련 특성으로 대출기간, 대출시점의 부동산 가치, 총 대출금액, 대출목적, 금리조정기간, 주택규모 등을 선정하였다. 대출기간 중에 나타난 특성을 포함시키기 위해서 한번이라도 연체한 경험이 있는지에 대해서 더미변수화하여 독립변수와 투입변수에 포함시켰다.

일반적으로 일시적인 연체는 연체관리에서 크게 중요하게 고려되지 않았으나, 일시적인 연체경험이 최종적인 연체발생에 영향을 미치는지 확인하기 위해서 변수를 선정하였다.

---

58) 주택담보대출은 대출실행 후 이자지급 약정일에 이자를 미납부하는 경우 납부일의 익일부터 연체에 편입되며 연체관리의 대상이 된다. 금융기관에서는 일반적으로 10일 또는 20일 이내 연체기간을 단기연체로 구분하며 30일 이상 연체기간을 장기연체로 구분하고 있다. 연체관리는 연체편입 후 20일 이내 시점에서 중점 연체관리가 이루어지므로, 20일 이내 시점의 연체자를 종속변수 기준으로 사용함.

〈표 45〉 분석모형의 변수

종속변수(산출변수)		10일, 20일, 30일 이상 연체발생 여부 (연체발생=1,미발생=0)	
독립변수 (투입변수)	차입자 특성	성별	차입자 성별 더미 (남성=0,여성=1)
		연령	차입자 연령
		소득	차입자 소득(백만원)
	대출 특성	대출시기	2007년~2010년
		부동산 가치	담보대상 부동산 가치
		총 대출금액	대출실행 시 총대출
		대출목적	대출목적 더미 (구입자금=0, 가계자금=1)
		금리조정 기간	금리조정기간 더미 (6개월이상=0, 6개월미만=1)
		주택규모	대출시점 구입 주택규모(m <sup>2</sup> )
연체경험	1회 이상 연체 여부 더미 (있음=1, 없음=0)		

### 3) 분석자료의 특징

분석대상의 주택담보대출자료 특성과 별도로 연도별 LTV 평균사용률 현황, 연령분포, 차입금액분포, 대출기간분포, 담보제공 평형분포를 통하여 분석자료의 특징을 분석하였다.

4개년 동안 이루어진 대출의 차입형태는 가계자금으로 사용된 목적이 구입자금으로 사용된 목적보다 많았으며 담보로 제공된 LTV 평균사용률을 보면 가계자금이 32%를 사용하였고 구입자금은 57%를 사용하고 있어 구입자금으로 활용시 LTV 비율이 상대적으로 높이 사용되었다. 차입자연령은 은퇴를 앞둔 시기라고 할 수 있는 55세 초과인 차주가 19.9%를 차지하고 있었다.

대출금액은 1억원 이하 대출비율이 66.4%를 차지하고 있었으며, 대출기간은 단기대출기간이라고 할 수 있는 3년이하 대출비율이 8.6%에 달하고 있었다. 담보로 제공된 아파트평형은 국민주택규모라고 할 수 있는 85(m<sup>2</sup>)이하

제공이 79.6%를 차지하고 있었다.

〈표 46〉 연도별 LTV 평균사용률 현황

년도	대출건수			전체사용 LTV(%)	가계자금 LTV(%)	구입자금 LTV(%)
	전체	가계자금	구입자금			
2007	148	80	68	49	37	63
2008	91	51	40	41	30	55
2009	186	113	73	40	33	52
2010	123	91	32	39	32	59
합계(평균)	548	335	213	(42)	(32)	(57)

〈표 47〉 연령 분포현황

나이	차입자수	비율
45세 이하	259	47.3%
55세 이하	439	80.1%
55세 초과	109	19.9%
합계	548	-

〈표 48〉 차입금액 분포현황

금액	차입자수	비율
5천만원 이하	220	40.1%
1억원 이하	364	66.4%
2억원 이하	495	90.3%
3억원 이하	538	98.2%
3억원 초과	10	1.8%
합계	548	-

〈표 49〉 대출기간 분포현황

대출기간	차입자수	비율
1년 이하	3	0.5%
3년 이하	47	8.6%
5년 이하	50	9.1%
10년 이하	70	12.8%
15년 이하	154	28.1%
20년 이하	381	69.5%
20년 초과	167	30.5%
합계	548	-

〈표 50〉 담보제공 평형 분포현황

아파트평형	차입자수	비율
60(m <sup>2</sup> ) 이하	172	31.4%
85(m <sup>2</sup> ) 이하	436	79.6%
102(m <sup>2</sup> ) 이하	471	85.9%
135(m <sup>2</sup> ) 이하	537	98.0%
135(m <sup>2</sup> ) 초과	11	2.0%
합계	548	-

## 제 4 절 연구대상의 속성

### 1) 성별간 차이

성별에 따른 주요변수간 평균의 차이를 확인하기 위해서 t-검정을 실행하였다. 분석결과 차입자 연령과 대출만기, 대출시 주택가격, 대출금액, 대출취급시소득, 대출취급시 부채, 연체횟수 등에서 남성과 여성의 차이가 나타났다. 남성 차입자의 경우 여성에 비해서 주택규모와 가격, 대출규모 등은 더욱 크게 나타났으나 차입자연령은 낮고 만기는 짧은 것으로 나타났다. 남성의 경우 세대주로서 차입하는 비율이 여성보다 높고 또한 여성보다는 이른 나이에 내집마련 목적으로 대출을 활용하고, 대출기간은 가계자금용도로 대출을 하면서 대출기간을 짧게 활용한 것으로 보인다.

〈표 51〉 성별간 주요변수의 평균

	성별	N	평균	표준편차
주택규모(m <sup>2</sup> )	남성	435	81.97	22.83
	여성	113	75.90	23.48
차입자 연령	남성	435	46.45	8.41
	여성	113	49.83	12.93
대출기간	남성	435	19.57	7.99
	여성	113	21.64	8.48
대출당시가격 (단위:천원)	남성	435	436,182	226,317
	여성	113	381,861	214,966
대출사용LTV	남성	435	42.44	29.29
	여성	113	42.37	27.08
총대출금액 (단위:원)	남성	435	99,851,264	82,318,906
	여성	113	85,440,707	65,113,602
취급시 소득	남성	435	61.41	51.95
	여성	113	26.53	26.24
취급시 부채	남성	435	88.14	99.34
	여성	113	70.58	93.39
현 부채	남성	435	187.26	160.46
	여성	113	131.75	104.96
1일이상 연체횟수	남성	435	2.49	5.70
	여성	113	3.55	6.14
연체일수	남성	435	10.89	39.03
	여성	113	15.14	50.75

〈표 52〉 평균비교 결과 : t-검정

구분	t	자유도	유의확률 (양쪽)	평균차 (남성 - 여성)
주택규모(m <sup>2</sup> )	2.502	546	.013**	6.068
차입자 연령	-2.640	137.547	.009***	-3.384
대출기간	-2.334	167.252	.021**	-2.067
대출당시가격 (단위:천원)	2.296	546	.022**	54,321
대출사용 LTV	.023	546	.981	.071
총대출금액 (단위:원)	1.978	214.755	.049**	14,410,556
취급시 소득	9.945	359.806	.000***	34,883
취급시 부채	1.695	545	.091*	17,565
현 부채	4.434	264.159	.000***	55,508
1일이상 연체회수	-1.735	546	.083**	-1.061
연체일수	-.964	545	.335	-4.248

유의수준 : \*(p<0.1) \*\* (p<0.05) \*\*\* (p<0.01)

## 2) 용도간 차이

대출목적에 따른 평균의 차이에서는 주택규모, 차입자연령, 대출기간, 대출 사용 LTV, 대출금액, 총대출금액 및 대출당시 부채에서 차이가 있는 것으로 나타났다.

구입자금은 가계자금에 비해서 주택규모와 차입자연령, 취급시부채 등은 작은 것으로 나타났으나 대출당시 LTV와 총 대출금액은 더 큰 것으로 나타났다.

구입자금의 경우 주택구입시 허용된 LTV를 모두 활용하여 차입을 하지만 가계자금의 경우는 구입자금보다는 소액으로 활용되어지기 때문으로 보인다.

〈표 53〉 구입목적별 주요변수의 평균

구분	대출목적	N	평균	표준편차
주택규모(m <sup>2</sup> )	구입자금	213	76.83	21.666
	가계자금	335	83.19	23.629
차입자 연령	구입자금	213	44.44	8.153
	가계자금	335	48.87	10.066
대출기간	구입자금	213	22.25	7.371
	가계자금	335	18.56	8.280
대출당시가격 (단위:천원)	구입자금	213	408,799	204,323
	가계자금	335	435,270	236,794
대출사용LTV	구입자금	213	57.25	30.132
	가계자금	335	33.01	23.573
총대출금액 (단위:원)	구입자금	213	125,528,169	82,631,459
	가계자금	335	78,664,478	71,362,838
취급시 소득	구입자금	213	54.73	53.983
	가계자금	335	53.90	47.067
취급시 부채	구입자금	213	62.79	87.021
	가계자금	335	98.26	102.609
현부채	구입자금	213	171.74	149.859
	가계자금	335	178.40	153.979
1일이상 연체횟수	구입자금	213	2.84	5.917
	가계자금	335	2.62	5.741
연체일수	구입자금	213	13.09	46.965
	가계자금	335	10.93	38.040

〈표 54〉 평균비교 결과 : t-검정

구분	t	자유도	유의확률 (양쪽)	평균차 (구입자금-가 계자금)
주택규모 (m <sup>2</sup> )	-3.172	546	.002***	-6.361

차입자 연령	-5.644	515.011	.000***	-4.424
대출기간	5.296	546	.000***	3.685
대출당시가격 (단위:천원)	-1.344	546	.180	-26,471
대출사용 LTV	9.962	373.222	.000***	24.240
총대출금액 (단위:원)	6.817	403.196	.000***	4,686,691
취급시 소득	.190	546	.850	.829
취급시 부채	-4.328	500.748	.000***	-35.464
현 부채	-.499	546	.618	-6.661
1일이상 연체횟수	.416	546	.678	.212
연체일수	.592	545	.554	2.166

유의수준 : \*(p<0.1) \*\* (p<0.05) \*\*\* (p<0.01)

### 3) 시기별 차이

대출시기에 따른 주요변수간 차이를 확인하기 위해서는 분산분석을 실행하였다. 분산의 동질성이 확보되는 변수에 대한 분산분석 결과에서 주택규모와 대출시 사용한 LTV에서 통계적으로 유의미한 차이가 있는 것으로 나타났다. 대출이 실행된 주택규모는 점차 커지는 반면 대출에 적용된 LTV는 점차 낮아지므로 금융위기 이후 중대형 주택소유자들의 대출이 증가하는 반면 금융기관에서는 금융위기이후 보수적으로 대출을 실행한 것으로 사료된다.

신규로 공급되는 중대형아파트(85㎡)의 공급비율은 2003년 21.0%였으나 이후부터 급속히 증가하여 2007년에는 그 비율이 50.2%에 달했으며 2011년에도 31.5%였다.<sup>59)</sup>

59) 자료 : 닥터아파트, 『수도권 중대형 아파트 공급비율』, 2013년 2.20, 조선일보 보도자료 재인용.

〈표 55〉 대출시기별 주요변수의 평균

구 분	연 도	N	평균	표준편차	표준오차
주택규모 (㎡)	2007	148	78.7	21.9	1.8
	2008	91	74.8	21.0	2.2
	2009	186	83.2	24.6	1.8
	2010	123	83.8	22.8	2.1
	합계	548	80.7	23.1	1.0
대출당시 가격 (단위:천원)	2007	148	398,813	258,042	21,211
	2008	91	403,330	198,006	20,757
	2009	186	455,593	212,178	15,558
	2010	123	426,195	216,263	19,500
	합계	548	424,981	224,910	9,608
대출사용 LTV	2007	148	48.8	29.4	2.4
	2008	91	40.8	28.7	3.0
	2009	186	40.3	28.6	2.1
	2010	123	39.1	27.7	2.5
	합계	548	42.4	28.8	1.2
취급시 부채	2007	148	76.6	92.9	7.6
	2008	91	68.5	83.2	8.8
	2009	186	91.8	98.2	7.2
	2010	123	94.7	112.8	10.2
	합계	547	84.5	98.3	4.2
현 부채	2007	148	166.7	163.5	13.4
	2008	91	147.3	114.9	12.0
	2009	186	193.0	164.7	12.1
	2010	123	181.9	140.3	12.7
	합계	548	175.8	152.3	6.5

〈표 56〉 평균비교 결과 : 분산분석

		제곱합	df	평균 제곱	F	유의확률
주택규모 (㎡)	집단-간	6,052	3	2,017	3.85	0.01**
	집단-내	285,206	544	524	.	.
	합계	291,258	547	.	.	.
대출당시 가격 (단위:천원)	집단-간	3.18E+11	3	1.06E+11	2.11	0.10
	집단-내	2.74E+13	544	5.03E+10	.	.
	합계	2.77E+13	547	.	.	.
대출사용 LTV	집단-간	8,422	3	2,807	3.42	0.02**
	집단-내	446,168	544	820	.	.
	합계	454,590	547	.	.	.
취급시 부채	집단-간	55,169	3	18,390	1.91	0.13
	집단-내	5,222,504	543	9,618	.	.
	합계	5,277,673	546	.	.	.
현 부채	집단-간	145,669	3	48,556	2.11	0.10
	집단-내	12,540,090	544	23,052	.	.
	합계	12,685,759	547	.	.	.

유의수준 : \*(p<0.1) \*\*\*(p<0.05) \*\*\*\*(p<0.01)

#### 4) 연체 및 비연체가구간 특성

연체여부에 따른 주요변수간 평균의 차이를 확인하기 위해서 20일 이상 연체를 기준으로 t-검정을 실행하였다. 분석결과 주택규모, 대출당시주택가격, 대출금액, 연체횟수에서 차이가 있는 것으로 나타났다. 연체가 없는 가구의 경우 주택규모가 크고 주택가격이 높았으며 대출금액이 컸지만 대출취득시 소득도 더욱 높은 것으로 나타났다. 따라서 대출규모 보다는 대출을 감당할 수 있는 소득여부가 연체여부에 영향을 미치는 것으로 판단된다.

〈표 57〉 연체여부에 따른 주요변수의 평균

구분	20일 이상연체	빈도	평균	표준편차
주택규모(m <sup>2</sup> )	비연체	478	82	23
	연체	70	71	19
차입자 연령	비연체	478	47	9
	연체	70	48	12
대출당시가격 (단위:천원)	비연체	478	437,093	227,383
	연체	70	342,269	188,750
대출사용 LTV	비연체	478	42	29
	연체	70	44	31
총대출금액 (단위:원)	비연체	478	99,474,059	80,683,119
	연체	70	79,164,286	66,386,807
취급시 소득	비연체	478	57	52
	연체	70	34	25
취급시 부채	비연체	478	86	99
	연체	70	78	93
현 부채	비연체	478	176	150
	연체	70	172	165
1일이상 연체횟수	비연체	478	1	2
	연체	70	15	8
연체일수	비연체	478	2	4
	연체	70	78	93

〈표 58〉 평균비교 결과 : t-검정

구분	t	자유도	유의확률 (양쪽)	평균차 (비연체-연체)
주택규모(m <sup>2</sup> )	3.642	546	.000***	10.637
차입자 연령	-.500	80.502	.618	-.766
대출당시가격 (단위:천원)	3.325	546	.001***	94,825
대출사용 LTV	-.513	546	.608	-1,895
총대출금액 (단위:원)	2.321	101.393	.022**	20,309,773
취급시 소득	5.976	176.256	.000***	22.642
취급시 부채	.637	545	.524	8.024
현 부채	.212	546	.833	4.127
1일이상 연체횟수	-13.790	69.937	.000***	-13.979

유의수준 : \*(p<0.1) \*\*\*(p<0.05) \*\*\*(p<0.01)

## 제 4 장 주택담보대출 연체가능성 분석

### 제 1 절 로지스틱 회귀분석 결과

#### 1) 연체가능성 변수

앞서 선정된 분석변수를 통해서 로지스틱 회귀분석을 실시한 결과는 아래의 <표 59>와 같다. 전체적인 모형의 적합도를  $-2$  로그우도값( $-2LL : -2 \log \text{likelihood}$ )<sup>60)</sup>의 변화량에 대한 검증을 실시하였으며, 모형에 독립변수들이 전혀 포함되지 않았을 경우의  $-2LL$ 과 독립변수들이 모두 포함된  $-2LL$ 의 변화량은 종속변수 10일 이상에서 266.30, 20일 이상에서 233.95, 30일 이상에서 202.96으로서 유의수준에서 유의미한 것으로 나타났다.

독립변수들이 종속변수를 설명하는 정도를 나타내는 결정계수( $R^2$ )의 경우도 Cox & Snell  $R^2$  과 Nagelkerke  $R^2$  2개 모형 모두의 설명력을 비교한 결과 모형은 통계적으로 적합한 것으로 나타났다.

종속변수 10일 이상 연체기준 분석결과 2009년과 2010년에 발생한 대출의 경우가 연체가능성이 낮은 것으로 유의하게 나타났다.

종속변수 20일 이상 연체기준 분석에서는 연체가능성에 영향을 미치는 변수로 주택규모가 작고, 차입자성별이 남성인 경우 연체가능성이 높은 것으로 나타났으며, 대출시기가 2009년인 경우 연체가능성이 낮은 것으로 나타났다.

종속변수 30일 이상 연체기준 분석에서는 연체가능성에 영향을 미치는 변수로 20일 이상 연체기준과 같이 주택규모가 작고, 차입자성별이 남성인 경우 연체가능성이 높은 것으로 나타났다. 또한 대출시기가 2009년인 경우 연체가능성이 낮은 것으로 나타났다.

---

60) 모형의 적합도 판정(goodness of fit of the model)에서 어느 특정한 로짓모형이 얼마나 데이터구조에 적합 한모형인가하는 판별은 그 모형의 우도(likelihood)값에 의존한다. 로짓 모형의 적합도 측정을 위한 통계로는  $-2LL : -2 \log \text{likelihood}$ 을 사용한다. 『사회과학의 고급계량분석 : 원리와 실제』, 이제열 외(2005), 서울대학교 출판부, p.98.

〈표 59〉 로지스틱 회귀분석 결과

구분	10일 이상 연체		20일 이상 연체		30일 이상 연체	
	B	유의확률	B	유의확률	B	유의확률
차입자 성별	-.406	.262	-.756*	.067	-.980**	.035
차입자 연령	-.008	.609	.009	.581	.019	.267
차입자 소득	-.002	.674	-.003	.589	-.002	.740
2008년 더미	-.847	.046	-.418	.347	-.435	.357
2009년 더미	-1.164***	.003	-1.309***	.004	-1.552***	.003
2010년 더미	-1.674**	.049	-1.037	.310	-.979	.464
부동산 가치	.000	.174	.000	.911	.000	.416
총 대출금액	.000**	.013	.000	.172	.000	.187
대출목적	.441	.207	.560	.146	.460	.267
금리조정기간	-.706	.302	-1.003	.236	-1.464	.184
주택규모	-.009	.325	-.021**	.050	-.027**	.028
연체경험 더미	21.118	.992	20.215	.992	19.764	.992
상수항	-19.913	.992	-19.408	.993	-19.456	.992
-2LL	266.30		233.951		202.96	
Cox & Snell R <sup>2</sup>	0.391		0.286		0.227	
Nagelkerke R <sup>2</sup>	0.625		0.536		0.486	

유의수준 : \*(p<0.1) \*\*\*(p<0.05) \*\*\*(p<0.01)

로지스틱 회귀분석결과를 살펴보면 통계적으로 유의한 변수와 연체가능성 여부는 다음과 같다. 차입자특성 변수로 사용된 성별, 연령, 소득 중에서 성별만이 20일 이상 연체와 30일 이상 연체기준에서 연체에 미치는 영향이 통계적으로 유의한 것으로 나타났다. 차입자가 여성일 경우 연체발생 가능성이 낮은 것으로 나타났으며 이러한 이유는 여성 차입자가 남성에 비해서 대출을

이용하여 주택을 구입하는데 보수적으로 접근하는 측면이 있기 때문에 추정된다. 또한 기혼자의 경우 여성이 차주가 되는 경우에는 1가구이상 주택을 보유하는 경우로 경제적인 사정이 양호하기 때문으로도 판단할 수 있다(유재술·안정근, 2013).

반면 차입자소득의 경우 예상과 달리 연체에 미치는 영향이 통계적으로 유의하지 않았으며 이는 분석대상 대출이 실행된 시기가 정부의 DTI 정책이 강력하게 추진되었던 시기와 일치하여 차입자소득이 대출에 미치는 영향이 차별화되어 나타나지 않았기 때문으로 판단된다. 차입자연령이 미치는 영향이 통계적으로 유의하지 않은 것은 분석대상으로 활용된 금융기관이 밀집해 있는 지역으로 차입자가 대부분 직장인이기 때문에 나타난 결과로 판단된다.

대출과 관련된 독립변수에서는 대출대상 주택규모와 대출실행시기, 대출규모가 통계적으로 유의한 영향을 미치는 것으로 나타났다. 먼저 주택규모는 종속변수 20일 이상과 30일 이상 연체기준 분석에서 나타났으며 주택규모가 클수록 연체발생 가능성이 낮은 것으로 나타났다. 이러한 결과는 중대형 아파트의 경우에 대출부실 가능성이 높다는 일반적인 예상과는 상반되는 결과로 큰 주택을 구입한 차입자의 경우 상대적으로 경제적 수준이 높아 연체가능성도 낮아지기 때문으로 판단된다(유재술·안정근, 2013).

대출시기의 경우 종속변수 10일 이상 연체기준 분석에서 2008년과 2009년에 대출이 실행된 경우 연체가능성이 낮게 나타났으며 20일과 30일 이상 연체기준은 2009년에 대출실행 된 경우가 연체가능성이 낮게 나타난 것을 알 수 있었다. 이는 금융위기 이후 대출실행에 대한 기준이 강화되고 주택시장도 침체되면서 대출을 활용하여 주택을 구입하는 경우가 감소했기 때문으로 판단된다.

총 대출금액의 경우 종속변수 10일 이상 연체기준 경우에만 통계적으로는 유의한 것으로 나타났으나, 독립변수의 계수가 제로(0)에 가까워 실질적인 연체가능성을 결정하는 데는 영향을 미치지 못하는 것으로 나타났다. 이는 대출금액 자체보다 대출조건에 대한 특성이 연체가능성에 더욱 영향을 미치게 됨을 보여주는 결과로 보여진다. 그 이유는 LTV 규제로 담보대상 부동산의 일정 수준까지만 대출이 이루어지게 하는 우리나라 부동산정책 특성에 기인한

것으로 판단된다(유재술·안정근, 2013).

## 2) 분석모형의 예측정확도

각 분석방법별 예측의 정확도를 살펴볼 경우 전체적인 연체율에 대한 전체적인 예측의 정확도는 종속변수 30일 이상 연체를 기준으로 한 모형이 91.6%로 가장 높게 나타났으며, 연체자그룹에서는 10일 이상에서 비연체자그룹에서는 30일 이상에서 높게 나타났다. 그러나 30일 이상의 경우는 연체자에 대한 예측은 28.8%로 가장 낮게 나타났다. 10일 이상 연체기준도 전체적인 예측정확도는 87.6%로 낮게 나타났으나 연체자에 대한 예측정확도는 67.9%로 높게 나타났다. 또한 20일 이상 연체기준의 경우 전체 예측정확도는 89.1%, 연체자에 대한 예측정확도는 41.4%로 나타났다.

이러한 모형으로 볼 때 종속변수의 연체일자를 구분하여 정확도 높은 연체자그룹을 선별할 수는 없어 보인다. 전체적인 예측정확도는 30일 이상 연체자그룹이 가장 높았으나 연체자에 대한 예측에서는 10일 이상 연체자그룹 예측정확도가 67.9%로서 가장 높게 나타났기 때문이다.

〈표 60〉 로짓모형을 통한 연체가능성 예측정확도

구분			예측			정확도 (%)
			0	1	합계	
실제	10일 연체	0	408	34	442	92.3
		1	34	72	106	67.9
		합계	442	106	548	87.6
	20일 연체	0	459	19	478	96.0
		1	41	29	70	41.4
		합계	500	48	548	89.1
	30일 연체	0	487	9	496	98.2
		1	37	15	52	28.8
		합계	524	24	548	91.6

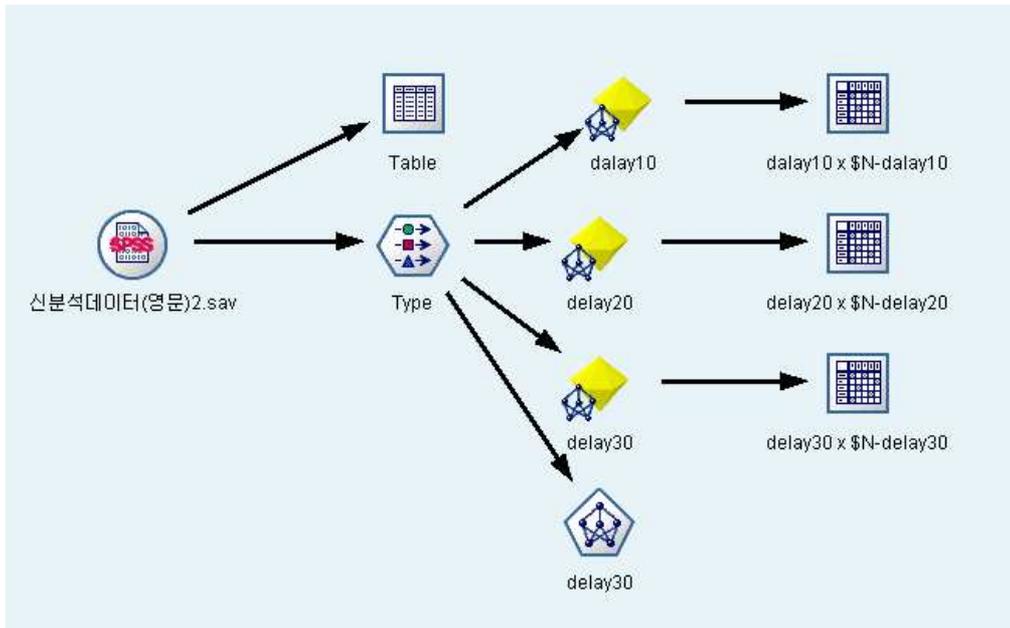
참고 : 미발생=0, 연체발생=1

## 제 2 절   인공신경망분석 결과

인공신경망모형은 로짓모형의 독립변수들을 투입변수로 하고 연체발생여부를 출력변수로 하여 예측모형을 개발하였다. 출력변수의 경우 로짓모형과 같이 10일 이상, 20일 이상, 30일 이상 연체를 별도로 구분하여 변수를 산정하였다. 인공신경망모형의 경우 로짓모형과 달리 변수의 통계적 특성을 고려할 필요가 없으므로 모든 투입변수의 중요도를 확인할 수 있으며 구축된 모형을 통한 결과의 정확도를 측정하였다.

분석에는 데이터마이닝 프로그램인 클레멘타인(Clementine)11.1이 이용되었으며 모형개발을 위한 스트림구축은 다음과 같다. 먼저 SPSS프로그램을 통해 분석에 필요한 변수를 입력하고 분석자료를 불러온 후 입력변수와 출력변수로 상용할 변수를 정의한다. 그리고 유형(Type)을 설정하고 모형을 실행하여 최적화된 모형을 선정한다. 최종적으로 모형에 대한 결과를 확인하는 방법으로 분석을 하였다.

〈그림 20〉 인공지능망모형 구축



1) 10일 이상 연체기준

먼저 10일 이상 연체자에 대한 분석결과 최적화 모형의 입력층은 투입변수의 수와 같이 10개로 나타났으며 Hidden Layer는 20개로 구성되었고 예측의 정확도는 92.4%로 나타났다. 예측에서 가장 중요한 변수로는 1회 이상 연체 경험이 가장 높게 수치로 나타났고 주택규모와 차입자연령, 주택가격, 대출규모 등이 그 뒤를 이었다. 인공지능망모형에서 상대적으로 중요한 것으로 나타난 변수 중에서 로짓모형에서는 통계적으로 유의한 변수와 일치하지 않는 것으로 나타났다.

이러한 결과는 로짓모형은 독립변수의 통계적 유의성을 중시하지만 인공지능망모형은 예측정확도를 중심으로 모형을 구축하기 때문으로 판단된다. 전반적으로 중요도가 높은 항목은 차입자특성 보다는 대출관련 특성으로 나타났다. 차입자특성 중 차입자연령은 모형구성에서 중요한 변수로 제시되었으나, 그 외의 변수들은 상대적으로 낮은 중요도를 나타냈다.

〈표 61〉 인공신경망분석 결과표(10일 이상 연체)

Neural Network architecture	
Estimated accuracy(정확도) : 92.453	
Input Layer(입력층) : 10 neurons	
Hidden Layer 1(은닉층) : 20 neurons	
Output Layer(출력층): 1 neurons	
Relative Importance of Inputs	
연체경험	0.421051
주택규모	0.237521
차입자연령	0.225452
주택가격	0.21361
대출규모	0.210772
차입자소득	0.158102
대출시기	0.143044
금리조정기간	0.123051
차입자성별	0.107439
대출목적	0.105774

## 2) 20일 이상 연체기준

20일 이상 연체자에 대한 분석결과 최적화 모형의 Hidden Layer는 앞서 10일 이상 연체와 같이 20개로 구성되었고 예측의 정확도는 96.3%으로 상승한 것으로 나타났다. 예측에서 가장 중요한 변수로는 차입자연령이 가장 높게 수치로 나타났고 주택가격과 연체경험, 주택규모, 대출시기 등이 그 뒤를 이었다. 10일 이상 연체시 경우와 주요 영향변수가 다르게 나타났으나 주택규모는 로짓모형에서도 통계적 유의성이 있는 변수로 분석되어 두 모형에서 모두 연체여부에 영향력이 높은 변수로 나타났다.

이번 모형에서도 전반적으로 중요도가 높은 항목은 차입자특성 보다는 대출관련 특성으로 나타났으며 차입자특성 중 가장 높은 영향력을 보인 항목은 10일 이상 연체모형과 같이 차입자연령으로 나타나 통계적으로 유의한 변수로 선정되지 않은 로짓모형과는 차별화된 결과를 나타냈다.

〈표 62〉 인공지능망분석 결과표(20일 이상 연체)

Neural Network architecture	
Estimated accuracy(정확도): 96.364	
Input Layer(입력층): 10 neurons	
Hidden Layer 1 (은닉층): 20 neurons	
Output Layer(출력층): 1 neurons	
Relative Importance of Inputs	
차입자연령	0.509229
주택가격	0.393452
연체경험	0.268404
주택규모	0.201775
대출시기	0.175753
대출규모	0.158385
차입자소득	0.154412
대출목적	0.13148
금리조정기간	0.114073
차입자성별	0.077161

### 3) 30일 이상 연체기준

30일이상 연체자에 대한 분석결과 최적화 모형의 Hidden Layer는 앞서 나타난 두 모형과 같이 20개로 구성되었고 예측의 정확도는 94.5%로 나타났다. 예측에서 가장 중요한 변수로는 10일 이상 연체모형과 같이 1회 이상 연체경험이 가장 높은 수치로 나타났고 첫 번째 모형과 같이 주택규모와 차입자소득이 그 뒤를 이었다. 이번 모형에서는 차입자소득이 세 번째로 영향력이 높은 변수로 나타난 것이 다른 모형과 차별화된 특성이며, 연체여부와 주택규모는 10일 이상, 20일 이상, 30일 이상 연체모형에서 모두 높은 변수로 나타냈다.

〈표 63〉 인공지능망분석 결과표(30일 이상 연체)

Neural Network architecture	
Estimated accuracy(정확도): 94,545	
Input Layer(입력층): 10 neurons	
Hidden Layer 1 (은닉층): 20 neurons	
Output Layer(출력층): 1 neurons	
Relative Importance of Inputs	
연체경험	0.31309
주택규모	0.20787
차입자소득	0.094855
금리조정기간	0.089471
대출시기	0.078796
차입자연령	0.077560
대출규모	0.067175
대출목적	0.04674
차입자성별	0.046130
주택가격	0.014101

#### 4) 분석모형의 예측정확도

인공신경망모형은 시행착오법을 통해서 변수간 가중치를 조정하여 예측모형을 구축하므로 각 모형별로 중요한 투입변수는 차별화되어 나타나는 것으로 보인다. 각 모형의 연체가능성에 대한 예측정확도를 살펴본 결과 〈표 64〉에 나타난 바와 같이 비연체에 대한 예측정확도는 97.4%를 나타낸 30일 이상 연체모형이 가장 높게 나타났으며, 연체에 대한 예측정확도는 69.8%를 나타낸 10일상 연체모형이 가장 높은 정확도를 나타냈다. 20일 이상 연체모형의 정확도도 68.6%를 나타내 41.4%를 나타낸 로짓모형에 비해서 현저하게 높은 정확도를 나타냈다.

연체기준을 짧게 설정하여 모형을 구축한 경우일수록 연체자에 대한 비중이 높아지므로 연체자에 대한 예측정확도가 높게 나타나는 반면 연체기준이 길수록 전체 변수에서 차지하는 연체자에 대한 비중이 작아지므로 모형에 투

입된 변수가 작아져 예측의 정확도가 낮아지는 것으로 판단된다.

〈표 64〉 인공지능망 모형을 통한 연체가능성 예측정확도

구분			예측			정확도 (%)
			0	1	합계	
실제	10일 연체	0	411	31	442	93.0
		1	32	74	106	69.8
		합계	443	105	548	88.5
	20일 연체	0	455	23	478	95.2
		1	22	48	70	68.6
		합계	477	71	548	91.8
	30일 연체	0	483	13	496	97.4
		1	33	19	52	36.5
		합계	516	32	548	91.6

참고 : 미발생=0, 연체발생=1.

이상의 결과를 볼 때 인공지능망모형과 로짓모형은 각각의 장점을 가지고 있는 것으로 보인다(유재술·안정근, 2013). 데이터마이닝기법인 인공지능망모형은 개별변수에 대한 통계적 유의성을 제한받지 않아 모든 투입변수를 통해 모형을 구축한다는 장점이 있다.

본 연구에서도 차입자소득 변수의 경우 로짓모형에는 통계적 유의성이 없는 변수로 나타났으나 인공지능망모형에서는 상대적으로 설명력이 높은 변수로 분석되었다. 그 내용을 살펴보면 10일 이상 연체자그룹에서는 연체여부, 주택규모, 차입자연령, 주택가격, 대출규모 순서 다음으로, 20일이상 연체자그룹에서는 차입자연령, 주택가격, 연체경험, 주택규모, 대출시기, 대출규모 순서 다음으로, 30일이상 연체자그룹에서는 연체여부, 주택규모 다음으로 설명력이 보다 높은 변수로 나타났다.

이러한 인공지능망 모형은 변수의 상대적 중요성을 양(+)수로만 가중치를 제시하고 있으나 이러한 모형은 인간의 두뇌구조를 모방한 시행착오법을 사용하는 모형으로 여러 개의 연체가능성의 뉴런들이 상호 연결하여 입력에 상응하는 최적의 출력 값을 예측하는 것이다. 이러한 출력 값은 수많은 반복훈련과정을 통하여 변수의 데이터에 숨어있는 패턴을 찾아내어 연체가능성을

예측하는 장점을 가지고 있다(유재술·안정근, 2013).

로짓모형의 경우는 연체와 비연체 여부에 대한 이항변수를 종속변수로 다른 독립변수들의 연체가능성을 추정하는 분석모형으로 연속형 변수와 범주형 변수들을 동시에 포괄할 수 있는 점과, 연체가능성에 미치는 변수를 이론적으로 설명할 수 있다는 장점도 있다.

따라서 로짓모형은 통계적 유의성을 살펴볼 수 있으며 인공신경망을 통해서 예측의 정확도를 높일 수 있다고 본다. 향후 지속적인 연구를 통해서 두 가지 방법의 결과를 적절하게 활용할 수 있는 방안을 마련할 필요가 있을 것으로 보인다(유재술·안정근, 2013).

### 제 3 절 의사결정나무분석 결과

의사결정나무분석을 위하여 앞서 실행한 로짓분석과 인공신경망분석에서 사용한 변수를 적용하였다. 분석은 앞서 두 분석과 동일하게 10일 이상, 20일 이상, 30일 이상 연체를 기준으로 분석을 실행하고 3가지 분석의 차이점을 확인하였다.

#### 1) 10일 이상 연체기준

10일 이상 연체자를 기준한 분석은 연체경험에서 첫 번째 분리가 발생하고 연체가 한번이라도 있는 경우의 차입자들은 금리조정기간을 기준으로 두 번째 가지로 분류되었고 다음으로 대출시기를 기준으로 세 번째 가지로 분리되었다.

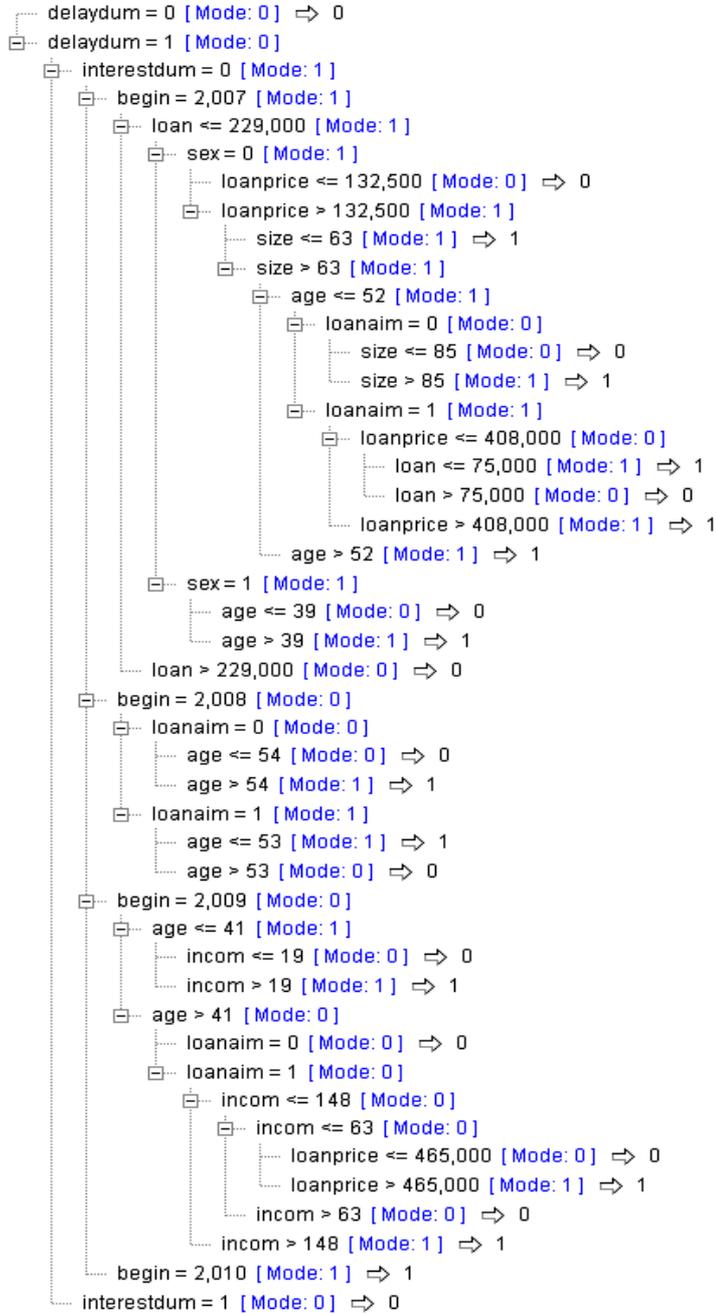
다음단계로 2007년 연체자들의 경우는 차입자성별로, 2008년 연체자들의 경우는 대출목적으로, 2009년 연체자들의 경우는 대출규모로 분류되었다.

대출시기가 중요한 의사결정요인으로 제시된 점은 2009년의 대출이 상대적으로 연체가능성이 낮은 것으로 나타난 로짓모형과 유사한 결과가 나타난 것으로 판단할 수 있으며 2007년 대출에서 상위단계 분류기준인 차입자성별은 로짓모형에서도 유의한 변수로 제시되었다. 상세한 의사결정분석의 트리는 아래 <그림 22><그림 23>과 같다.

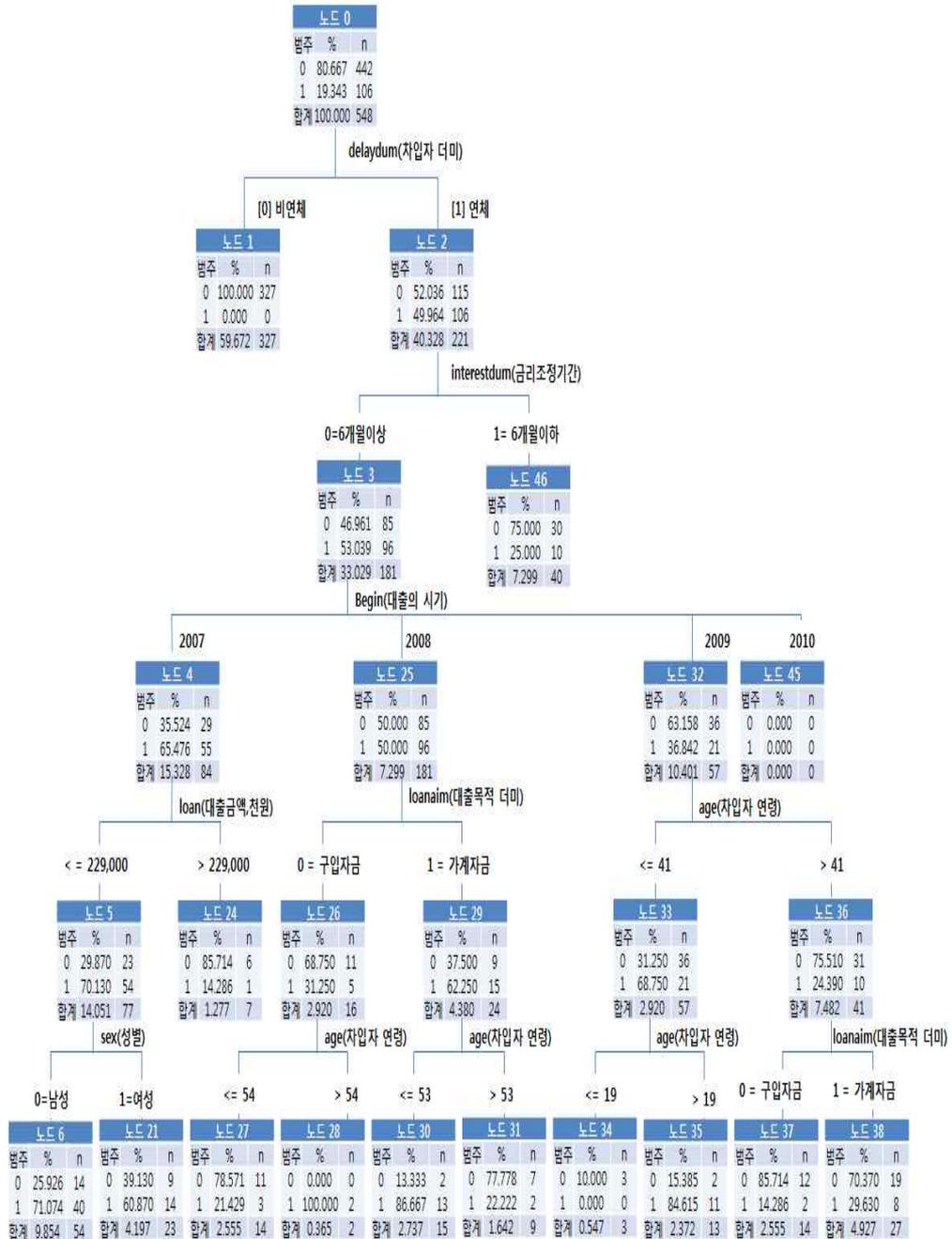
의사결정나무 분석트리에서 확인가능 한 바와 같이 의사결정나무분석은 연체가능성을 확인하기 위한 변수들의 위계 순서를 확인할 수 있다는 특징을 가지고 있다.

즉 로짓모형은 연체가능성 분석을 위해서 통계적으로 유의한 변수를 제시하고 인공신경망모형은 예측에서 상대적으로 중요성이 있는 변수를 제시하는 반면 의사결정나무분석은 변수들 간의 위계순서를 통해서 연구목적인 연체가능성이 높은 유형의 대출을 찾아낸다는 차이점이 있다.

〈그림 21〉 의사결정나무 분석의 모형구축 결과(10일 이상 연체)



〈그림 22〉 의사결정나무 분석의 트리 결과1(10일 이상 연체)





의사결정나무모형 결과를 정리하면 연체가능성이 높은 대출의 규칙을 도출할 수 있으며 <표 65>에서 확인할 수 있다. 먼저 Rule 1 for 1에서 연체가 발생할 수 있는 첫 번째 규칙을 살펴보면 연체경험이 있는 가구로서 6개월 이상 금리조정기간을 가지고 있고 2007년에 대출 취급된 가구이다.

다음 단계로는 대출액이 229백만원 이하이고 차입자성별이 남성이며 주택가격은 132.5백만원을 초과한다. 그리고 주택규모는 63m<sup>2</sup>이하인 경우 연체가능성이 높은 것으로 나타났다. 10일 이상의 연체를 기준으로 한 경우 이러한 규칙의 집단은 <표 65>연체기준 Rule set에서 11개가 나타났으며 대출시기가 큰 줄기로 나타난 것이 특징이다. 2007년에 6개가 2008년에는 2개 2009년 3개가 나타나 연체가능성이 높은 집단으로 구분되었다.

<표 65> 10일 이상 연체기준 Rule set

Rule 1 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,007 and loan <= 229,000 and sex = 0 and loanprice > 132,500 and size <= 63 then 1	Rule 2 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,007 and loan <= 229,000 and sex = 0 and loanprice > 132,500 and size > 63 and age <= 52 and loanaim = 0 and size > 85 then 1	Rule 3 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,007 and loan <= 229,000 and sex = 0 and loanprice > 132,500 and size > 63 and age <= 52 and loanaim = 1 and loanprice<=408,000 and loan <= 75,000 then 1
Rule 4 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,007 and loan <= 229,000	Rule 5 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,007 and loan <= 229,000	Rule 6 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,007 and loan <= 229,000

and sex = 0 and loanprice > 132,500 and size > 63 and age <= 52 and loanaim = 1 and loanprice > 408,000 then 1	and sex = 0 and loanprice > 132,500 and size > 63 and age > 52 then 1	and sex = 1 and age > 39 then 1
Rule 7 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,008 and loanaim = 0 and age > 54 then 1	Rule 8 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,008 and loanaim = 1 and age <= 53 then 1	Rule 9 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,009 and age <= 41 and incom > 19 then 1
Rule 10 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,009 and age > 41 and loanaim = 1 and incom <= 148 and incom <= 63 and loanprice > 465,000 then 1	Rule 11 for 1 if delaydum = 1 and interestdum = 0 and begin = 2,009 and age > 41 and loanaim = 1 and incom > 148 then 1	

- 주) 1. 용어의 설명 : delaydum(연체자 더미, 연체발생=1, 비연체=0), interestdum(금리 조정 기간 더미, 6개월 이상=0, 6개월 이하=1), begin(대출의 시기), loan(대출금액, 천원), sex(차입자 성별 더미, 남성=0, 여성=1), loanprice(담보대상 부동산의 가치, 천원), size(주택규모, m<sup>2</sup>), age(차입자 연령), loanaim(대출목적 더미, 구입자금=0, 가계자금은 =1), incom(차입자 소득, 백만원).
2. 단위 : <표 44> 분석사례의 기술적 통계분석 단위 적용.
3. 이하 <표 66>과 <표 67> 연체기준 Rule set 에서도 동일한 용어 및 단위적용.

## 2) 20일 이상 연체기준

20일 이상 연체자를 기준으로 한 의사결정나무분석에서도 먼저 Rule 1 for 1에서 연체경험에서 첫 번째 분리가 발생하고 연체가 한번이라도 있는 경우의 차입자들은 대출시기를 기준으로 두 번째 가지로 분류되었다. 다음 단계로는 대출액이 229백만원 이하이고 차입자의 성별이 남성이며, 대출목적중 구입자금인 경우 연체가능성이 높은 것으로 나타났다.

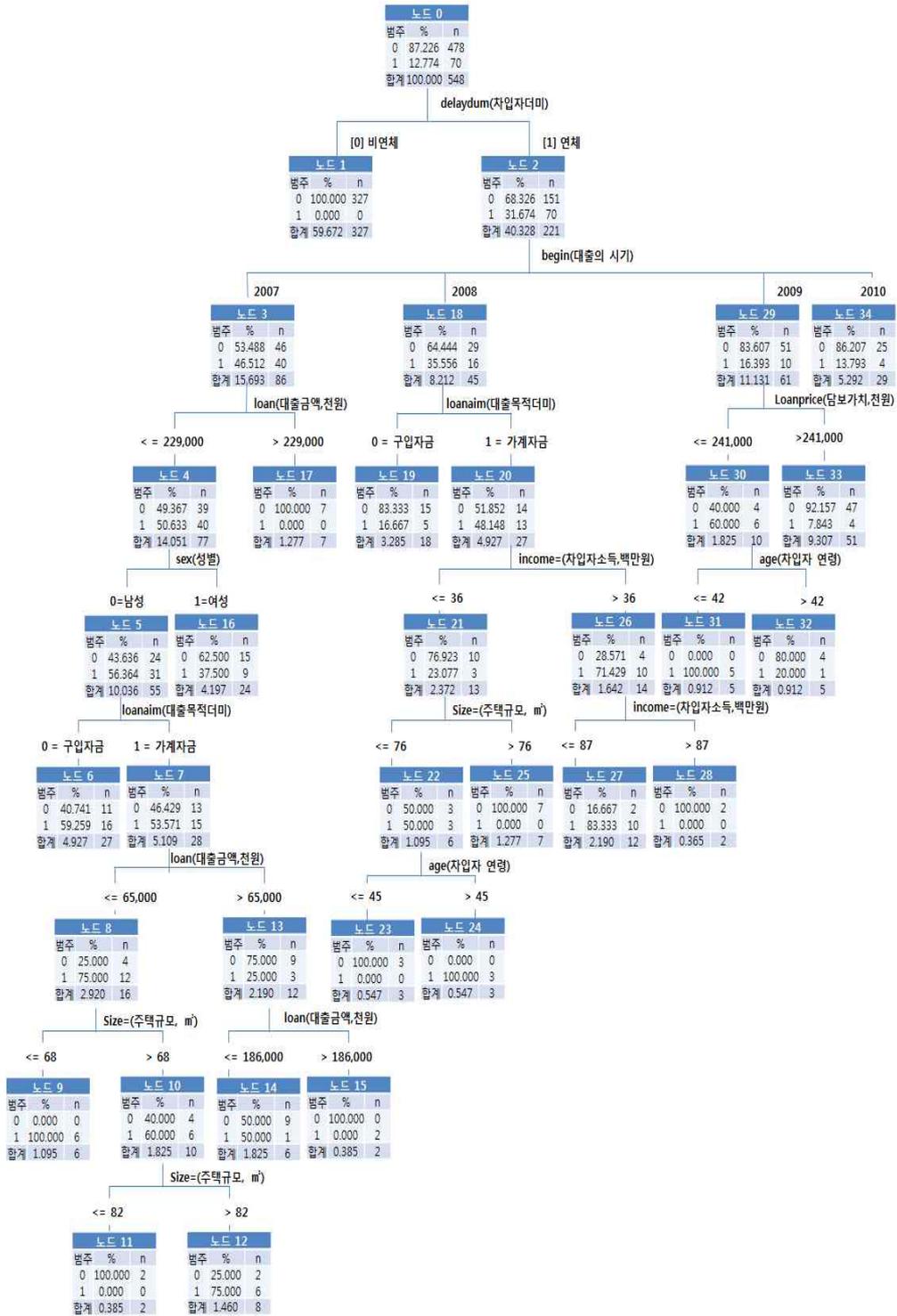
또한 2007년 연체자의 경우는 대출금액으로, 2008년 연체자의 경우는 Rule 5 for 1에서 대출목적으로 분류되었지만, Rule 7 for 1에서는 2009년 연체자의 경우 대출당시 담보가치로 분류되었다.

대출시기의 다음단계 기준에서 2008년과 2009년에는 10일 이상 연체분석과 동일하게 나타났으며 연체경험이 있고 2010년에 대출을 받은 가구의 경우 전체 29가구 가운데 25가구가 10일 이상 연체가 발생하지 않아 연체가 없는 집단으로 분류되었다. 상세한 의사결정분석 트리는 <그림 25>과 같다.

20일 이상 연체가 발생한 대출을 도출하기 위한 규칙은 연체기준 Rule set <표 66>에서 총 7개가 도출되었으며 10일이상의 연체기준과 마찬가지로 대출시기가 큰 줄기로 나타난 것이 특징이다. 2007년에 실행된 대출의 규칙이 4개, 2008년의 경우 2개, 2009년의 경우 1개가 도출되었다.



<그림 25> 의사결정나무 분석의 트리 결과(20일 이상 연체)



〈표 66〉 20일 이상 연체기준 Rule set

<p>Rule 1 for 1          if delaydum = 1          and begin = 2,007          and loan &lt;= 229,000          and sex = 0          and loanaim = 0          then 1</p>	<p>Rule 2 for 1          if delaydum = 1          and begin = 2,007          and loan &lt;= 229,000          and sex = 0          and loanaim = 1          and loan &lt;= 65,000          and size &lt;= 68          then 1</p>
<p>Rule 3 for 1          if delaydum = 1          and begin = 2,007          and loan &lt;= 229,000          and sex = 0          and loanaim = 1          and loan &lt;= 65,000          and size &gt; 68          and size &gt; 82          then 1</p>	<p>Rule 4 for 1          if delaydum = 1          and begin = 2,007          and loan &lt;= 229,000          and sex = 0          and loanaim = 1          and loan &gt; 65,000          and loan &gt; 186,000          then 1</p>
<p>Rule 5 for 1          if delaydum = 1          and begin = 2,008          and loanaim = 1          and incom &lt;= 36          and size &lt;= 76          and age &gt; 45          then 1</p>	<p>Rule 6 for 1          if delaydum = 1          and begin = 2,008          and loanaim = 1          and incom &gt; 36          and incom &lt;= 87          then 1</p>
<p>Rule 7 for 1          if delaydum = 1          and begin = 2,009          and loanprice &lt;= 241,000          and age &lt;= 42          then 1</p>	

### 3) 30일 이상 연체기준

30일 이상 연체자를 기준으로 한 의사결정나무분석에서 먼저 Rule 1 for 1에서는 연체경험에서 첫번째 분리가 발생한다. 연체가 한번이라도 있는 경우의 차입자는 금리조정기간 중 6개월 이상을 기준으로 두번째 가지로 분류되고 대출시기를 기준으로 세번째 가지로 분리되었다.

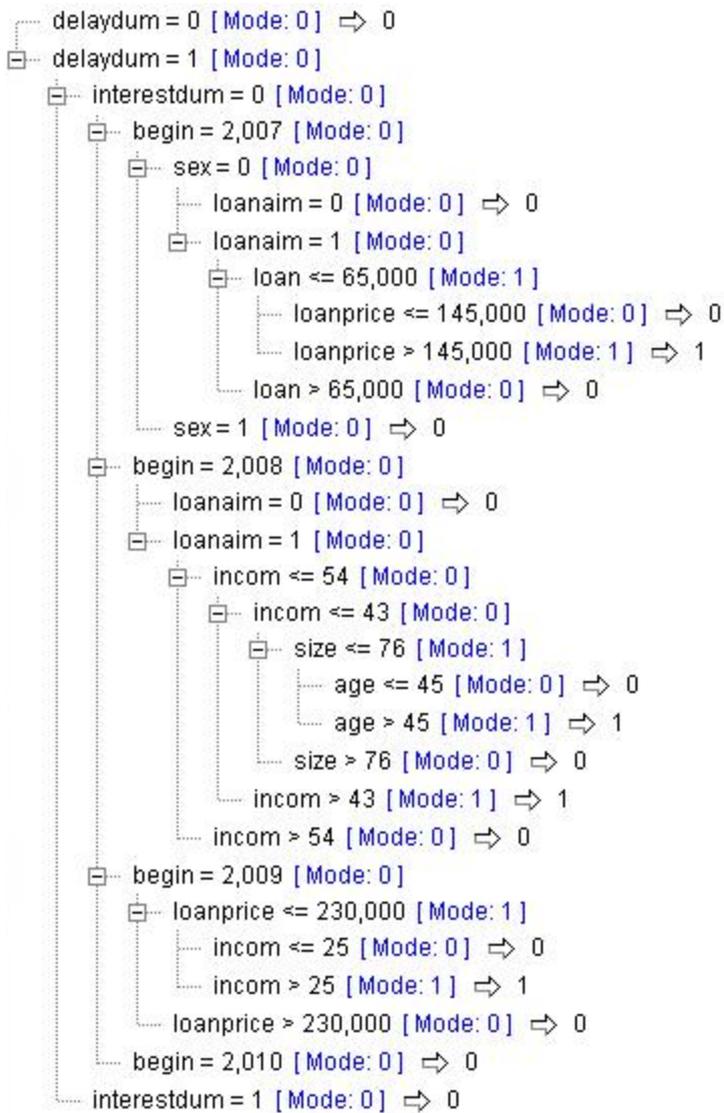
다음단계로 차입자성별이 남성이며, 대출목적중 가계자금인 경우와 대출액이 65백만원 이하이고 대출당시 담보의 가치가 145백만원 보다 큰 경우 연체가능성이 높은 것으로 나타났다.

2007년 연체자의 경우 연체가능성 다음단계로 차입자성별이 분류되었고, 2008년 경우 대출목적이 가계자금인 경우로, 2009년 경우 차입자의 대출당시 담보가치로 분류되었다.

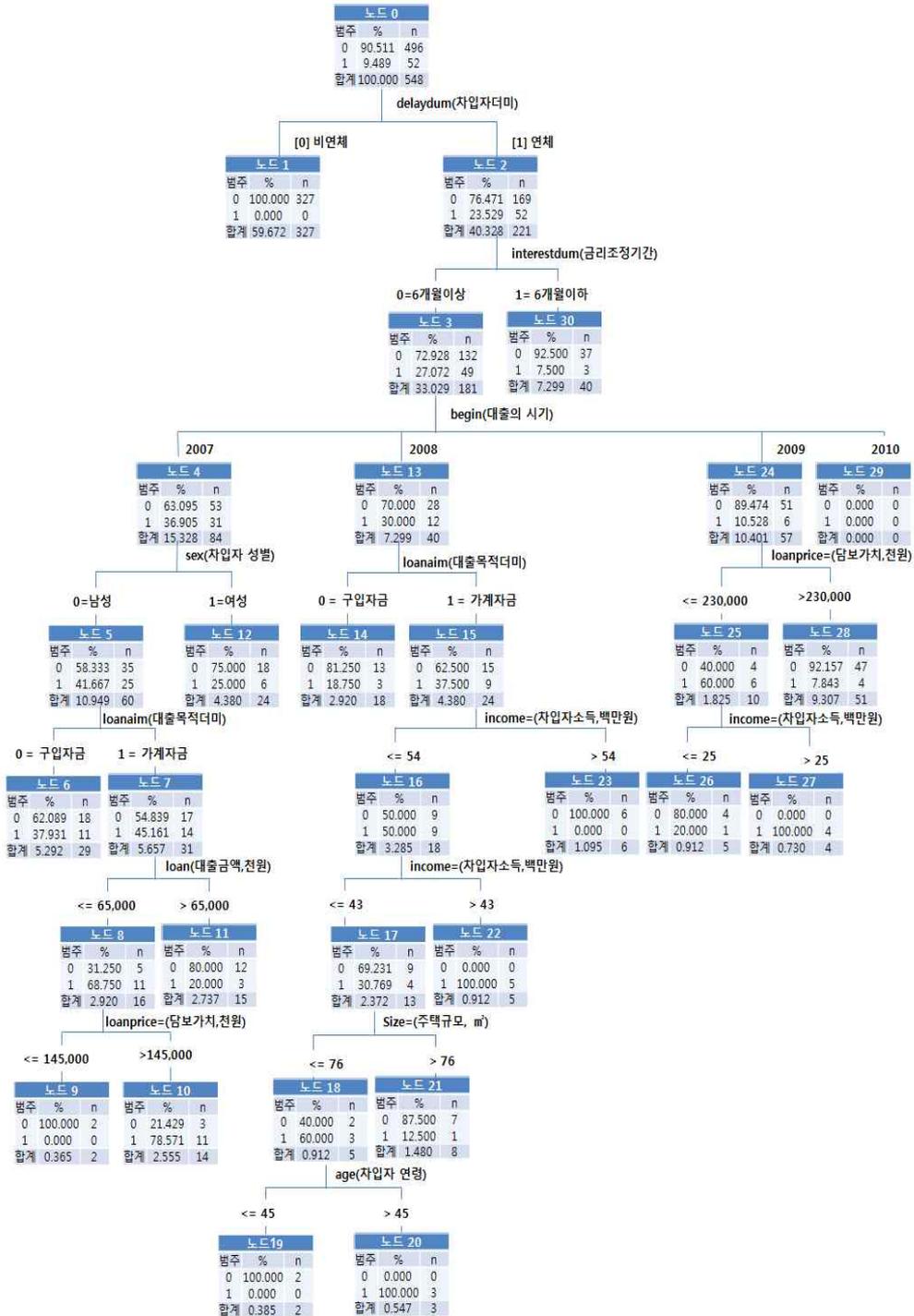
특히 2008년의 경우에는 10일, 20일, 30일 이상 연체기준 모두에서 대출목적단계로 구분될 때 구입자금보다는 가계자금대출이 상대적으로 연체가능성이 높은 것으로 분류되었다. 이러한 결과는 2008년 금융위기로 인해서 가계자금대출이 증가하고 경기악화로 연체가 증가했기 때문으로 판단된다.

연체가 발생한 규칙은 30일 이상 연체기준 Rule set<표 67>에서 총 4개가 도출되었으며 10일과 20일 이상의 연체기준과 마찬가지로 대출시기가 큰 줄기로 나타난 것이 특징이다. 그리고 연체경험이 있는 가구와 금리조정기간이 6개월 이상인 가구 중에서 2007년에 실행된 대출규칙이 1개, 2008년의 경우 2개, 2009년의 경우 1개가 도출되었다.

〈그림 26〉 의사결정나무 분석의 모형구축 결과(30일 이상 연체)



<그림 27> 의사결정나무 분석의 트리 결과(30일 이상 연체)



〈표 67〉 30일 이상 연체기준 Rule set

<p>Rule 1 for 1                  If delaydum = 1                  and interestdum = 0                  and begin = 2,007                  and sex = 0                  and loanaim = 1                  and loan &lt;= 65,000                  and loanprice &gt; 145,000                  then 1</p>	<p>Rule 2 for 1                  If delaydum = 1                  and interestdum = 0                  and begin = 2,008                  and loanaim = 1                  and incom &lt;= 54                  and incom &lt;= 43                  and size &lt;= 76                  and age &gt; 45                  then 1</p>
<p>Rule 3 for 1                  if delaydum = 1                  and interestdum = 0                  and begin = 2,008                  and loanaim = 1                  and incom &lt;= 54                  and incom &gt; 43                  then 1</p>	<p>Rule 4 for 1                  if delaydum = 1                  and interestdum = 0                  and begin = 2,009                  and loanprice &lt;= 230,000                  and incom &gt; 25                  then 1</p>

#### 4) 분석모형의 예측정확도

앞서 분석한 두가지 모형에서 각각 연체가능성을 분석하기 위한 규칙을 도출하였으며 의사결정나무분석에서도 동일한 방법으로 예측정확도를 도출한 결과 비교적 높은 정확성을 나타냈다.

비연체자 예측의 경우 10일 이상 연체기준은 97.3%, 20일 이상 연체기준은 96.9%, 30일 이상 연체기준의 경우 99.4%의 예측정확도를 나타냈으며 연체자 예측의 경우는 10일 이상 연체기준이 81.1%, 20일 이상 연체기준은 68.6%, 30일 이상 연체기준이 44.3%의 예측정확도를 각각 나타내 연체기준을 짧게 설정할 경우 예측정확도가 높았다.

〈표 68〉 의사결정나무 모형을 통한 연체가능성 예측정확도

구 분			예 측			정확도 %
			0	1	합계	
실 제	10일 연체	0	430	12	442	97.30
		1	20	86	106	81.10
		합계	450	98	548	94.10
	20일 연체	0	463	15	478	96.90
		1	22	48	70	68.60
		합계	485	63	548	93.20
	30일 연체	0	493	3	496	99.40
		1	29	23	52	44.20
		합계	522	26	548	94.20

참고 : 미발생=0, 연체발생=1.

## 제 4 절 분석결과 정리

본 절에서는 최적의 모형을 얻기 위해 로짓모형과 인공신경망모형, 의사결정나무 모형을 비교·평가하고 이를 통해 하나의 모형이 선택되면 선택된 모형이 다른 모형에 비해 우수하다는 것을 입증할 수 있다. 따라서 예측을 위해 만든 모형이 서로 다른 모형들 중 어느 것이 가장 우수한 예측력을 보유하고 있는지를 비교분석하고 앞서 설정된 세가지 기법을 이용한 모형 예측 결과를 오분류율 예측정확도를 통하여 비교해 보았다.

각 모형별 예측정확도를 비교하기 위해서 금융기관에서 적용되는 주택담보대출 연체율과 적용되는 변수 등을 고려하여 연체의 기준일자를 10일, 20일, 30일로 분류한 후 연체모형의 결과를 분석해 보았으며, 분석모형별 예측정확도 비교에서는 금융기관에서 중점적으로 연체관리를 하는 20일 시점을 연체 기준일자로 비교하였다.

분석한 결과 세가지 기법에 대한 예측을 크기순서는 <표 70> <표 71>에서 보는 바와 같이 의사결정나무모형 > 인공신경망모형 > 로짓모형 순으로 나타났다. 결과적으로 전체적인 분석의 정확도는 의사결정모형이 가장 높게 나타났으며, 연체자예측의 경우에는 인공신경망모형과 의사결정나무모형이 유사한 것으로 나타났다. 이에 비해서 로짓모형의 경우 전체적인 정확도와 연체자 선별에 대한 정확도에서 상대적으로 낮은 것으로 나타났다.

이러한 결과는 앞서 제시한 바와 같이 로짓모형의 경우 통계적인 유의성을 중심으로 하는 모형인데 비해서 데이터마이닝 모형인 인공신경망모형과 의사결정나무모형은 예측의 정확도를 높이는 경우 변수사용에 대한 유연성이 높기 때문으로 사료된다.

아래 <표 69>을 통하여 모형별 비교표를 작성하였으며 <그림 28>은 AUC 모형을 통하여 ROC Curve 결과를 예측하였다.

〈표 69〉 오분류율 예측정확도 비교

구분		예측		합계	정확도 (%)	
		0	1			
로짓모형	실제	0	459	19	478	96.0
		1	41	29	70	41.4
		합계	500	48	548	89.1
인공신경망모형	실제	0	455	23	478	95.2
		1	22	48	70	68.6
		합계	477	71	548	91.8
의사결정나무모형	실제	0	463	15	478	96.9
		1	22	48	70	68.6
		합계	485	63	548	93.2

참고 : 미발생=0, 연체발생=1.

〈표 69〉에 대한 오분류율의 다양한 예측은 다음과 같다.

〈표 70〉 오분류율 표에 의한 예측

정분류율 (accuracy)	(실제 0, 예측 0)빈도 + (실제 1, 예측 1)빈도 / 전체빈도	
	로짓모형	$(459+29)/548 = 488/548 = 89.1\%$
	인공신경망모형	$(455+48)/548 = 503/548 = 91.8\%$
	의사결정나무모형	$(463+48)/548 = 511/548 = 93.2\%$
오분류율 (error rate)	(실제 0, 예측 1)빈도 + (실제 1, 예측 0) 빈도 / 전체빈도	
	로짓모형	$1-488/548 = 60/548 = 10.9\%$
	인공신경망모형	$1-503/548 = 45/548 = 8.2\%$
	의사결정나무모형	$1-511/548 = 37/548 = 6.8\%$
민감도 (sensitivity)	(실제 1, 예측 1)의 빈도/실제 1의 빈도	
	로짓모형	$29/70 = 41.4\%$
	인공신경망모형	$48/70 = 68.6\%$
	의사결정나무모형	$48/70 = 68.6\%$
특이도 (specificity)	(실제 0, 예측 0)의 빈도/실제 0의 빈도	
	로짓모형	$459/478 = 96.0\%$
	인공신경망모형	$455/478 = 95.2\%$
	의사결정나무모형	$463/478 = 96.9\%$

참고 : 민감도는 1에서 정분류율로, 특이도는 0에서의 정분류율로 예측함.

자료 : 박창이·김진석(2008), R을 이용한 데이터마이닝 내용참조, p.14.

위 <표 70>를 통하여 모형별 통계량을 아래 <표 71>과 같이 작성할 수 있다. 정분류율, 오분류율, 민감도, 특이도를 살펴보면 의사결정나무모형이 로짓 모형이나 인공신경망모형에 비해 우수함을 알 수가 있다.

<표 71> 오분류율 표에 의한 모형별 통계량

모 형	정분류율 (accuracy)	오분류율 (error rate)	민감도 (sensitivity)	특이도 (specificity)
로짓 모형	0.8905	0.1094	0.4143	0.9603
인공신경망 모형	0.9178	0.0821	0.6857	0.9519
의사결정나무 모형	0.9324	0.0675	0.6857	0.9686

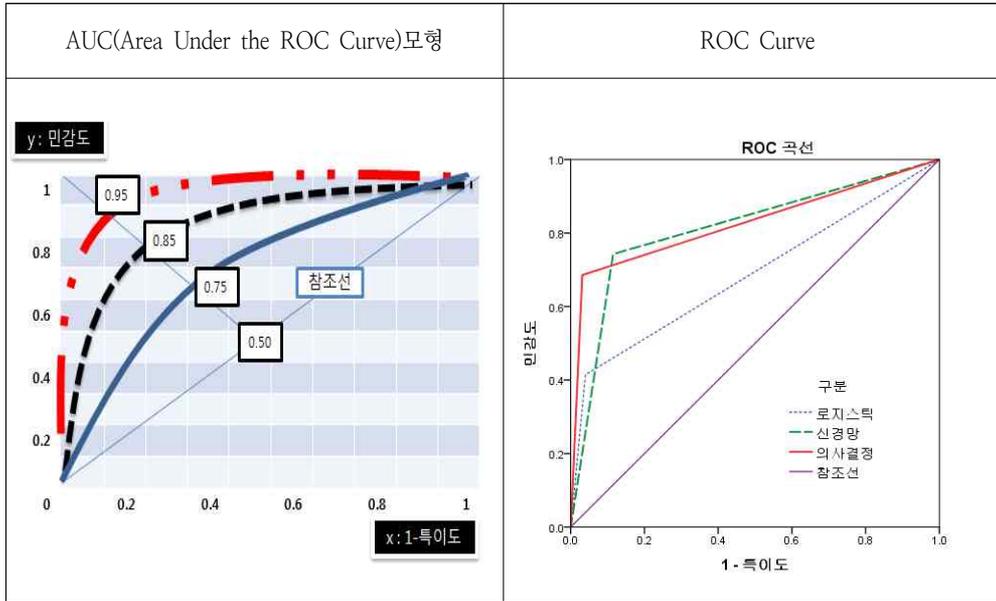
또한 최적의 모형판별을 위해서 세가지 모형을 비교·평가할 수 있는 ROC(Receiver Operating Characteristic)곡선을 사용하였다. 본 연구와 같이 주택담보대출의 연체자 집단과 비연체자 집단으로 구분되는 이진형 종속변수를 가지는 모형일 때 판별분석을 할 수 있으며 예측모형별 성능을 도표로 평가하는데 유용하고 민감도와 특이도에 의해 판단하는 곡선이다.

또한 각 분류 기준값에서 오분류표를 통해 얻어진 민감도와 특이도를 이용하여 ROC곡선을 작성할 수 있으며, 이 곡선의 X축에는 '1-특이도'가, Y축에는 민감도를 각 분류 기준값에 나타낸 것으로 그래프가 도표의 왼쪽 상단으로 가까울수록 성능면에서 우수한 모형으로 판단된다(최종후 외, 1999).

본 연구에서도 마찬가지로 로짓모형과 인공신경망모형, 의사결정나무모형의 ROC곡선을 비교하여 모형의 성능을 비교하고자 하였다. <그림 28>을 보면 ROC곡선 아래의 영역이 참조선의 기준인 0.5보다 크므로, 본 연구의 모델은 적합하다고 볼 수 있다. 이 면적을 나타내는 것이 <그림 28> AUC(Area Under the ROC Curve)이다.

ROC Curve 분석 결과 의사결정나무모형이 로짓모형 및 인공신경망모형보다 분류 및 예측 성능이 우수함을 알 수 있다.

〈그림 28〉 AUC 모형과 ROC Curve 결과



# 제 5 장 결 론

## 제 1 절 요약 및 시사점

주택담보대출의 확대는 2008년 미국의 서브프라임 모기지사태가 발생하면서 부동산시장이 재조명되었으며 단기간에 크게 증가한 주택담보대출의 리스크 관리에 대한 관심이 높아지고 있다. 또한 그동안 주택보급률의 증가로 주택시장은 안정세를 유지하고 있지만 주택수요는 둔화된 현상을 보이고 있다.

최근 경기침체로 인한 저성장이 지속되고 있으며 인구증가율은 감소되고 생산인구 또한 빠르게 감소할 것으로 전망되고 있어, 고령사회 진입을 앞두고 있는 등 여러 환경이 변화하고 있다. 사회적, 경제적 환경변화와 더불어 가계대출의 대부분을 차지하고 있는 주택담보대출의 커다란 규모와 빠른 증가는 건전성에 대한 관리의 필요성이 제기되고 있다.

정부에서도 2014년 관계부처합동으로 가계부채 구조개선 촉진방안을 제시하였고 부채부분의 가계건전성 제고를 위한 대출증가속도의 관리, 대출구조의 개선, 취약계층 채무조정 등 부분별 대책을 순차적으로 마련하고 있는 시점으로 금융기관에서는 가계대출의 비중이 높은 주택담보대출관리가 필요한 상황이다.

기존의 선행연구가 주로 거시경제변수와 연체율관계를 분석한 연구여서 금융기관에서 대출심사 및 연체예측성 등의 자료로 적용하기에는 부적합하였다. 본 연구에서는 주택담보대출 차입자 자료를 가지고 연체가능성 여부를 예측하기 위하여 데이터마이닝기법을 적용해보았으며 활용여부를 각 분석모형을 통하여 예측하였다.

본 연구는 분석에 앞서 먼저 주택담보대출에 대한 이론적인 검토를 통하여 대출에 대한 종류와 용도, 금리형태에 따른 유형, 상환방식을 알아보았으며, 변동금리와 고정금리의 특징과 장·단점을 비교하였고, 변동금리의 경우 고정금리에 비해서 대출초기 금리가 낮다는 장점이 있으나 기준금리의 변동성이 큰 경우에는 차입자의 이자율위험이 커지는 문제점이 있다는 것을 발견하였다. 또한 상환방식의 유형 및 특징에서는 원금균등분할상환과 원리금균등분할

상환의 상환방식에서 차이점이 있었으며 방식에 따라 원금상환금액이 차이를 보이고 있었다.

주택담보대출에 대한 위험관리 필요성으로 주택담보대출의 규모가 크고 빠르게 진행되고 있다는 점이다. 가처분소득보다 가계부채의 증가가 크고 빠르게 증가하고 있으며 주택담보대출 구조의 문제가 발견되고 있다. 일반적으로 만기구조가 장기이고, 고정금리에 원금을 함께 상환하는 방식이 안정적인 구조로 평가받고 있으나 현재는 만기일시상환비율이 높아 개선이 필요한 상황이다.

또한 주택시장의 환경변화는 주택가격의 안정화가 지속되고 있는 점과 생산인구의 감소와 고령화 현상 등이 장기적으로 주택시장 및 주택담보대출의 안정성에 영향을 미칠 것으로 예상되며, 가격상승을 예상하고 과도하게 대출을 받은 가구의 경우 채무불이행으로 진행될 가능성도 높아지고 있다.

주택담보대출 위험관리방안에는 대출시점에서 점검하는 사전관리방안이 있으나 이는 대출의 실행시점에서 위험을 관리하는 방안이며 금융기관의 대출 심사과정에서 위험관리가 이루어지게 된다. 대표적인 규제 방법으로 LTV와 DTI를 통해 대출금의 적정규모를 산정한 후 대출이 이루어진다. 사후적 위험관리방안은 채무불이행이 발생한 경우 피해를 최소화하기 위한 방안으로 볼 수 있으며 주택담보대출의 기한의 이익상실 및 연체관리가 있다.

우리나라 정부는 가계부채의 위험성을 인식하였고 가계부채를 오랜 기간 누적된 구조적 위험으로 보았다. 단기간 해결보다는 시스템리스크 완화를 위한 단기 대책과 중장기적 구조개선 노력을 병행 추진하는 내용으로 소득 대비 부채비율 관리와 고정금리와 비거치식 분할상환의 비중을 확대하고 취약계층 채무조정 등 가계소득과 대출구조 개선 및 부채부담 경감 등의 대응 내용을 발표하고 추진하고 있다.

해외 주요국가의 가계부채관리는 금융위기 이후의 경우로서 영국정부는 침체된 주택시장의 가계부채 및 하우스푸어 문제를 주택경기 활성화를 통해 해결하는 방식을 선택하였다.

스페인 은 가계부채를 감당하지 못하는 가구들에 대해서 저축은행 등 금융기관들이 부실채권의 증가를 억제하기 위해 자체적으로 채무를 조정하고 대

출대상 주택을 매입하여 재임대하는 방법을 사용하였으며, 아이슬란드는 가계 부채 및 하우스푸어에 대한 대책으로 공적자금 투입보다는 차입자에게 적극적인 채무조정을 실시하면서 채권자들도 책임을 분담하게 하였으며, 금융기관에 대한 공적자금 투입을 지양하였다.

미국은 금융위기 이후 초기에는 금융시장 안정과 경기부양차원의 정책을 추진하였으며 효과를 거두지 못하자 오바마정부에서 주택금융지원정책 중심으로 가계부채관리가 진행되고 있다. 미국의 공공 모기지 지원 프로그램은 조정방식과 옵션에 따라서 크게 리파이낸싱과 채무조정, 압류 및 퇴거방지로 구분되며 세부적으로는 8개의 형태로 운영되고 있다.

분석의 틀을 만들기 전에 선행연구를 살펴보았으며 주택담보대출이 가지고 있는 특성과 리스크와 연체가능성 특성으로 분류하여 고찰하였다. 선행연구를 통하여 나타난 주택담보대출에 대한 모형은 거시경제 변수와 연체율을 분석한 모형이 주류를 이루었으며 가장 일반적인 모형은 회귀분석모형이 많이 사용되었다.

본 연구에서는 주택담보대출 차입자의 연체가능성에 대한 분석이므로 연체자를 예측할 수 있는 모형이 필요했으며, 분석집단의 대상자를 연체자와 비연체자로 구분할 수 있고 이를 두개의 범주형 종속변수로 적용 가능한 로지스틱 회귀분석인 이항로짓모형을 우선적으로 적용하였다.

두번째 연구모형은 이러한 로짓모형과 차별성을 살펴보기 위하여 데이터마이닝모형인 인공신경망모형을 적용하였다. 이러한 모형은 대출거래자에 대한 신용분석과 금융사기의 탐색 및 예방, 신용카드의 부정사용 파악, 고객의 신용평가 등 금융관련분야에서도 다양하게 적용되고 있어서 연체가능성 모형에 적용하기에 적합하다고 판단하였다.

세번째 연구모형인 의사결정나무모형은 의사결정 규칙을 나무구조로 도표화하여 분류와 예측을 수행하는 분석방법이다. 이 방법은 분류 또는 예측과정이 나무구조에 의한 추론 규칙에 의해서 표현되기 때문에 분석자가 그 과정을 쉽게 이해하고 설명할 수 있다는 장점이 있는 것이 특징이다.

따라서 이러한 모형들은 연체가능성 판별분석 또는 회귀분석 등과 같은 분

석에 있어서 사전에 이상치를 확인하거나 분석에 필요한 변수를 찾아내는데 이용될 수 있고 다양한 예측모형을 개발하는데도 활용될 수 있다. 또한 큰 틀에서는 로짓모형과 데이터마이닝모형간의 차이를 살펴볼 수 있었으며 개별적으로는 세가지 모형의 결과를 통하여 모형의 차이점과 예측력을 발견하고 연체가능성 여부에 영향을 미치는 변수와 가장 우수한 모형을 선별하였다.

분석결과 먼저 연체가능성 여부에 영향을 미치는 변수를 보면 로짓모형과 인공신경망모형에서는 대출기간, 금리조정기간, 주택규모 등이 공통적으로 연체에 영향을 많이 미치는 것으로 나타났으며, 의사결정나무모형에서는 대출시기, 대출금액, 대출목적 등이 상위분류기준으로 나타나 연체여부에 영향을 주는 것으로 나타났다.

연체가능성 분석모형에 대한 예측율의 순서를 보면 의사결정나무모형이 가장 우수하였고 인공신경망모형, 로짓모형 순으로 나타났다. 연체자 예측의 경우에는 인공신경망모형과 의사결정나무모형이 유사한 것으로 나타났으며 로짓모형의 경우 전체적인 정확도와 연체자 선별에 대한 정확도에서 상대적으로 낮은 것으로 나타났다.

분석내용으로 비연체차입자에 대한 예측은 세가지 연구모형에서 모두 95% 수준을 넘었다. 로짓모형이 96.0%, 인공신경망모형이 95.2%, 의사결정나무모형이 96.9%로 나타났으며, 연체차입자에 대한 예측은 로짓모형이 41.4%, 인공신경망모형이 68.6%, 의사결정나무모형이 68.6%로 나타났다. 비연체차입자에 대한 예측은 로짓모형과 의사결정나무모형이, 연체차입자에 대한 예측은 인공신경망모형과 의사결정나무모형이 높게 나타나 상대적으로 편차를 보이긴 하였으나 전반적인 예측력은 의사결정나무모형이 가장 우수한 모형으로 나타났다.

결론적으로 로짓모형과 데이터마이닝기법을 활용한 연체가능성분석 모형 중 가장 우수한 모형은 의사결정나무모형으로 분석되었다. 이러한 점은 향후 주택담보대출 연체가능성을 판단하는 모형을 적용할 경우 우선적으로 적용할 수 있다는 점이 본 연구로서 의의가 있다고 본다. 그러나 모형의 특성에 따라 장점이 있으므로 각 분석결과를 모두 고려하여 판단할 필요가 있을 것으로 보인다.

로짓모형은 통계적인 활용도에서, 인공신경망모형은 통계적 유의성에 대한 제약 없이 다양한 변수의 영향력을 활용할 수 있다는 점에서 의의를 가질 수 있으며, 의사결정나무모형은 비연체차입자와 연체차입자를 단계별 분류기준으로 확인 할 수 있다는 점에서 의미가 있다고 본다. 또한 이러한 로짓모형과 데이터마이닝모형은 향후 금융기관의 주택담보대출 리스크관리에 적용가능성이 높다고 보여지며 향후 다양한 연구가 진행될 필요가 있다고 본다.

## 제 2 절 연구의 한계 및 향후 연구과제

본 연구가 가지고 있는 한계로는 먼저 투입된 자료가 많지 않다는 점이다. 데이터의 수가 총 548개로 예측모형의 정확도를 높이는데 한계가 있었으며 차입자에 대한 예측이 상대적으로 낮게 나타났다. 이러한 문제는 정부의 개인 정보보호에 강화에 따라 대출자료를 획득하기가 어렵기 때문이다. 따라서 분석에 투입된 자료는 허용된 범위 내에서 분석대상수를 확보할 수 있는 아파트에 한정되었다. 또한 로짓모형의 독립변수와 인공신경망모형 및 의사결정나무모형의 투입변수가 제한적일 수밖에 없어 좀 더 정교한 모형을 개발하는데는 한계가 있었다.

향후 아파트와 다세대, 다가구 및 단독주택, 상가 등 부동산담보대출에 활용되어지는 담보자료가 포함되고 차입자 자료는 개인의 신용등급에 따라 세분화된 거래현황과 연체발생 이후 자료가 분석에 포함될 필요가 있다. 그리고 데이터마이닝 분석에 적합한 수준의 데이터 용량과 차입자에 대한 다양한 정보를 수집하여 금융기관의 리스크관리에 적합가능 한 모형으로 발전시킬 필요가 있다고 본다.

# 참 고 문 헌

## 1. 국내문헌

- 강성진·이주석. (2012). 주택금융의 수요자특성에 관한 연구. 『부동산학연구』, 18(3): 37-50.
- 고성수. (2008). 『한국의 부동산 금융의 현황과 과제 ; 한국의 부동산 금융의 현황과 과제』. 서울: 건국대학교 출판부.
- 공정거래위원회. (2014). <http://www.ftc.go.kr/>
- 관계부처합동. 금융위원회. (2014). 『가계부채 구조개선 촉진 방안』.
- 구승환·김현수·장성용. (2009). 국내 남자 프로농구 승패 예측 모형 비교 연구. 『체육과학연구』, 20(4): 704-711.
- 국가통계포털. (2014). <http://kosis.kr/>.
- 국회예산정책처. (2012). 『주요국의 가계부채 조정과정 및 대응정책 분석』.
- 권병조. (2006). 『국민임대주택백서』. 서울: 건설교통부.
- 금융감독원. (2014). 『은행업감독규정』. <http://www.fss.or.kr/>.
- 금융위원회. (2011). 『가계대출 연착륙 종합대책』.
- 금융위원회. (2013). 『은행권 주택담보대출 구조개선 추진실적』.
- 금융투자협회. (2014). <http://www.kofia.or.kr/>.
- 김경태·이장형·최현오. (2001). 신경망에 의한 주가수익률 예측모형. 『한국기업경영학회』, 14(-): 127-142.
- 김석기. (2013). 채무상환비율(DSR)의 가계부채 연체율 선행성과 시사점. 『주간경영브리프』, 22(47): 10-11.
- 김일수·남영우. (2011). 이자율 변화가 가계의 소득 및 흑자율에 미치는 영향에 대한 연구. 『부동산·도시연구원』, 4(1): 57-74.
- 김정인 외. (2002). 『미국의 주택금융제도』. 서울 : 국민은행 연구소.

- 김태훈·홍한국. (2004). 『회귀모형과 신경망모형을 이용한 아파트가격모형에 관한 연구』. 『국토연구원』, 43(-): 183-200.
- 김현정. (2010). 『우리나라의 가계부채의 특징과 민감도분석』, 『한국경제포럼』, 3(3): 77-94.
- 김희철·박형근. (2010). 패널 데이터모형을 적용한 국내 일반은행 연체율 결정 요인 추정에 관한 연구. 『한국 산학 기술학회 논문지』, 11(1): 272-277.
- 남영우. (2012). 『가계부채 문제해결을 위한 주택담보대출 관리방안』. 서울: 두리안.
- 남영우·성상준. (2012). 대출상환방법의 변화가 DTI에 미치는 영향에 대한 연구. 『부동산·도시연구원』, 4(2): 101-115.
- 남영우·이정민. (2006). 아파트시장예측을 위한 신경망분석 적용 가능성에 대한 연구. 『한국건설관리학회』, 7(2): 162-170.
- 남주하. (2012). 가계부채 해소방안. 『한국경제포럼』, 5(2): 61-86.
- 노태협·이택호·한인구. (2005). 금융시계열 모형을 이용한 KOSPI 200 주가지 수의 변동성 예측. 『한국경영학회』, 34(3): 683-713.
- 노형식. (2012). 『스페인 은행부분 위기전개의 시사점』. 서울 : 한국금융연구원. 21(27).
- 닥터아파트. (2013). 『수도권 중대형 아파트 공급비율』. 조선일보 보도자료 (2013.2.20.) 재인용.
- 대법원 법원경매정보. (2003~2013), <http://www.courtauction.go.kr/>.
- 대한상공회의소. (2012). 『가계부채현황과 정책과제』.
- 박성욱·박갑제·김태혁. (2009). 주택담보대출 차입자의 금리선택에 관한 연구. 『한국금융학회』, 23(2): 133-147.
- 박양수 외. (2012). 『부채경제학과 한국의 가계 및 정부부채』. 서울 : 한국은행.

- 박준호·김태호·주용진. (2011). 의사결정나무분석을 이용한 벤처기업의 공간적 입지특성에 관한 연구. 『국토지리학회』, 45(1): 175-188.
- 박창이·김진석. (2008). 『R을 이용한 데이터마이닝』.
- 박형수·홍승현. (2011). 『고령화 및 인구 감소가 재정에 미치는 영향』. 서울 : 한국 조세연구원.
- 박희정·강호정. (2009). 로지스틱회귀분석을 이용한 코스닥기업의 부실예측모형 연구. 『한국콘텐츠학회』, 9(3): 306-311.
- 방두완·박세운·박연우. (2010). 한국 모기지시장의 채무불이행 및 조기상환 분석. 『한국금융학회』, 24(4): 87-118.
- 방두완·박연우. (2012). 국내 주택담보대출 LGD의 기본특성과 결정요인 분석: 아파트시장을 중심으로. 『한국은행, 경제 분석』, 18(4): 51-83.
- 변성호·강현직·한정우·김태웅. (2008). 의사결정나무모형을 이용한 유역내 구조적 홍수방어 대안 도출. 『대한토목학회 논문집』, 28(1B): 33-40.
- 성명기. (2009). 『저출산 및 고령화가 경제에 미치는 영향과 시사점』. 저출산 고령사회 대응연구, 서울: 국회예산정책처.
- 손욱. (2009). 변동금리형 주택담보대출 기준금리 및 편중완화에 관한 연구. 『한국경제연구학회』, 25(-): 113-157.
- 손재영 외. (2008). 『한국의 부동산 금융의 현황과 과제』. 서울: 건국대학교 출판부.
- 송문섭·조신섭. (2004). 『SAS를 이용한 통계자료분석』. 자유아카데미.
- 송영석·조영찬·서용석·안상로. (2009). 의사결정나무모형을 이용한 급경사지재해 예측프로그램 개발. 『대한토목학회 논문집』, 29(2C): 59-69.
- 심종원·정의철·정현정. (2009). 주택담보대출 연체율결정 요인에 관한 연구. 『부동산분석학회』, 15(2): 81-96.
- 심형섭. (2012). 『채무조정 프로그램의 국제 비교 및 시사점』. 서울 : KDI.
- 안정근. (2014). 『현대부동산학』. 제5판, 양현사.

- 오태연·지균태. (2008). 『아이슬란드 금융위기 원인 및 전망』. 서울 : 지역경제포커스. 2(37).
- 우종덕·이정민. (2011). 주택분양보증 예측모형에 대한연구. 『한국부동산학회』, 47(-): 285-295.
- 위정범·백흥기. (2008). 금리정책과 부동산담보대출 연체율. 『한국기업경영학회』, 15(2): 17-40.
- 유승동. (2013). 『주택대출 금리유형에 대한 선호분석』.
- 유재술·안정근.(2013). 주택담보대출의 연체가능성 분석방법에 대한 연구. 『지적』, 43(2): 237-251.
- 이동걸·전성인·정재욱·변동준. (2013). 『국내 부채보유가구의 연체결정요인에 관한 연구』. 서울: 한국금융학회.
- 이문주·한진수. (2010). 고객 경험을 기반으로 한 테마파크 평가의 주요 속성 연구. 『대한관광경영학회』, 25(4): 317-333.
- 이재연 외. (2010). 『주택담보대출 기준금리개선 방안』. 서울 : 한국금융연구원.
- 이재연. (2011). 『주택금융제도의 국제간 비교 및 정책 제언』. 서울 : 한국금융연구원.
- 이재열 외. (2005). 『사회과학의 고급계량분석 : 원리와 실제』. 서울 : 서울대학교 출판부.
- 이혜훈. (2002). 고령화의 경제적 영향에 대한 소고. 『노동정책연구』, 2권 2호.
- 장민·이규복. (2010). 『기준금리 인상이 가계 건전성에 미치는 영향』. 서울: 한국금융연구원.
- 전국은행연합회. (2010). 『주택담보대출 핸드북』. <http://www.kfb.or.kr/>.
- 전치혁. (2013). 『데이터마이닝기법과 응용』. 서울 : 한나래.
- 정원구·이상엽. (2007). 인공지능망을 이용한 공동주택 가격지수 예측에 관한

- 연구 . 『한국주택학회』, 15(3): 39-64.
- 정유석. (2003). 인공지능망을 이용한 기업도산예측: IMF후 국내 상장회사를 중심으로. 『경희대 박사학위논문』.
- 정재호·신택수. (2002). 인공지능망을 이용한 신용카드부정사용 색출기법에 관한 연구. 『POSRI경영경제연구』, 2(2): 201-224.
- 정화미·허윤경·이성호. (2001). 신경망을 이용한 개별공시지가 산정에 관한 연구. 『국토연구원』, 36(7): 55-66.
- 조성빈. (2010). 의사결정나무추론을 이용한 채용결정요인 분석. 『서강경영논총』, 21(2): 151-164.
- 지규현·김정인·최창규. (2006). 주택담보대출 위험관리를 위한 차입자 특성 분석. 『대한국토·도시계획학회』, 41(4): 117-128.
- 지동현. (2007). 『주택담보대출의 리스크관리』. KB 국민은행연구소.
- 천대중. (2014). 『미국 은행의 주거용 모기지 비즈니스 현황과 특징』. 서울 : 우리금융경제연구소.
- 최병남. (1994). 인공지능망에 의한 토지평가시스템. 『국토연구원』, 22(-): 37-50.
- 최상일·박태진·강정규. (2011). 의사결정나무모형을 이용한 노년기 주거유형선택 예측에 관한 연구. 『한국주거환경학회지』, 9(2): 231-251.
- 최종후·서두성. (1999). 데이터마이닝 의사결정나무의 응용. 『통계청, 통계분석연구』, 4(1): 61-83.
- 최종후·소선하. (2005). 『사례로 배우는 데이터 마이닝』. 서울 : 자유아카데미.
- 최종후·한상태·강현철·김은석·김미경. (1999). E-Miner version 3.0의 변화와 향상된 기능. 『제 10회 한국SAS 사용자 컨퍼런스 발표 논문집』.
- 한국경제연구원. (2013). 『정부3.0 실현을 위한 빅데이터 활용방안』. 서울: KIPA 연구보고서.

- 한국금융연구원 동향실. (2012). 『아이슬란드경제의 회복과 시사점』. 서울 : 한국금융연구원. 21:(23).
- 한국은행. (2012). 『부채경제학과 한국의 가계 및 정부부채』.
- 한국은행. (2014). 경제통계시스템, <http://www.bok.or.kr/>.
- 하나금융연구소. (2006). 『아이슬란드 금융위기 가능성 검토』. 서울: 하나금융 정보.
- 허윤경·이승우. (2012). 『금융위기 이후 미국의 모기지 지원 프로그램의 시사점』. 서울 : 한국건설산업연구원.
- 허석균. (2012). DTI, LTV 및 대출상환 조건이 주택담보대출의 연체율에 미치는 영향. 『규제연구』, 21(2): 39-77.
- 홍석철·전한경. (2013). 인구고령화와 소득불평등의 심화. 『한국경제의 분석』, 19(1): 71-122.
- 홍성두·김동일. (2008). 학습장애 진단의 대안적 통합모형에 대한 의사결정나무 기법의 활용 가능성 탐색. 『특수교육학연구』, 43(1): 211-231.
- 황진수·최영진·황성훈. (2011). 의사결정나무 분석을 이용한 패밀리 레스토랑 선택속성에 따른 시장 세분화에 관한 연구. 『대한관광경영학회』, 23(7): 225-241.
- 한국경제연구원. (2014). 『영국 경기회복의 배경과 전망』. 서울: KIEP 지역경제포커스.
- KB 국민은행. (2014). <https://www.kbstar.com/>.

## 2. 국외문헌

- Alm, J. R. and J. R. Follain. (1987). Consumer Demand for Adjustable-Rate Mortgages. *Housing Finance Review* 6, 1-16.
- Andra C Ghent and Marianna Kudlyak. (2011). Recourse and residential mortgage default : *Review of Financial Studies, Federal Reserve Bank of Richmond Working Paper* No. 09-10R.
- Brent, William, Lynne Kelly. (2011). Debby Lindsey-Taliefero Russel price, "Determinants Of Mortgage Delinquency", *Journal of Business & Economics Research - February*, 9(2), 27-48.
- Brown, S. and K. Taylor. (2008). Household Debt and Financial Asset: Evidence from Great Britain, Germany, and United States, *Journal of Royal Statistical Society*, 171, 615-643.
- Bruckner, J. K. (1986). "The Pricing of Interest Rate Caps and Consumer Choice in the Market for Adjustable-Rate Mortgage". *Housing Finance Review* 16, 119-136.
- Canner, G. Dynan, K. and Passmore, W. (2002). Federal Reserve Bulletin, Mortgage Refinancing in 2001 and Early 2002, 469-481.
- Dhillon U.S and J. D. Shilling and C. F. Sirmans. (1987), Choosing between Fixed and Adjustable Rate Mortgages : *Note. Journal of Money, Credit and Banking*, 19(2), 260-267.
- DTZ New Zealand. (2004). "Housing Costs and Affordability in New Zealand". Centre for Housing Research Aotearoa New Zealand.
- Dynan, K. E. and D. L. Kohn. (2007). The Rise in U.S. Household Indebtedness: *Cause and Consequences. Finance and Economics Discussion Series*.

- Epstein Richard A. (1984). "Toward a Revitalization of the Contract Clause". *University of Chicago Law Review*(*The University of Chicago Law Review*, Vol. 51, No. 3) 51(3), 703–751.
- FHFA. (2012). Refinance Report.
- Green, Richard K. and Wachter, Susan M. (2005). The American Mortgage in Historical and International Context : *Journal of Economic Perspectives*, 19(4), 93–114.
- <http://www.globalpropertyguide.com/Europe/United-Kingdom>.
- <http://www.globalpropertyguide.com/real-estate-house-prices/>.
- IMF, "Ch. 3 Dealing with Household Debt", *World Economic Outlook*, April, 2012.
- Lea, International Comparison of Mortgage Product Offerings, *Research Institute for Housing America*, 2010.
- Ronel Elul, Nicholas S. Souleles , Souphala Chomsisengphet, Dennis Glennon and Robert Hunt (2010), WHAT “TRIGGERS” MORTGAGE DEFAULT?, Working Paper No. 10-13, *Research Department, Federal Reserve Bank Of Philadelphia*.
- Statistics Iceland, <http://www.globalpropertyguide.com/>
- The National Mortgage Servicer's Reference Directory(2004). 21th edition.

# ABSTRACT

## The Default Possibility Analysis of Housing Mortgage Loans Using Data Mining Techniques

Yoo, Jae-Sool

Major in Real Estate

Dept. of Economics and Real Estate

The Graduate School

Hansung University

After the 2008 subprime mortgage insolvency crisis, enlargement of housing mortgage loans shed new light on the Korean real estate market, and interest in risk management has been rising along with the rapid increase of housing mortgage loans. As diffusion ratio of house is consistently increasing over periods, the Korean housing market remains stable, however, housing demand has been slowing down. Especially, socio-economic environment are rapidly changing. Due to the recent worldwide economic recession, the Korean economy is also stuck on a plateau of slow growth. Decrease in population growth rate and decrease in productive population have been accelerating the change into an "Aged Society". The rapid increase in housing mortgage loans which make up most of domestic retail loans raise the importance of soundness management. The Korean government presented interagency structural improvement plans of domestic household debts in 2014, and is

successively preparing countermeasures for management of increasing rate of households debts, structural improvement of debts, and debt settlement for vulnerable social groups, and so on. Therefore, the financial agencies also have to pay attention to the soundness management. Because most advanced researches were focusing on analysis of relationship between macro-economic variables and default rate, they were not appropriate for data which explains relationship between the loan appraisals of financial agencies and the prediction of default rate. In order to predict the possibilities of default rate by using data of actual housing mortgage loan borrowers, binary logit model, neural network model and decision tree model, the latter two of them are data mining techniques, are applied for this research. There is a deviation in the result of research : binary logit model and decision tree model was better for borrowers with no arrears, otherwise neural network model and decision tree model was better for borrowers in arrears. However, decision tree model was superior in the overall predictive performance. In both binary logit model and neural network model, the term of loans, the term of interest rate adjustment, and scale of house are major factors of generating delinquency. In decision tree model, however, time, amount, and purpose of loans are highly affected factors of generating delinquency. Each model has it's own strength, so final conclusion should be made by considering all the results of analysis. Binary logit model is of use statistically, and neural network model has significance in that diverse variables can be utilized with no statistical restriction. Decision tree model has significance in that we can find the step-by-step standard of classifying borrowers with no arrears and borrowers in arrears.

**【Keyword】** : housing mortgage loan, default rate, binary logit model, neural network model, decision tree model.