

## 배깅 가망고객 모델링을 통한 방카슈랑스 활성화 방안

김태호<sup>1</sup> · 정재화<sup>2</sup> · 김진수<sup>3</sup>

<sup>1</sup>충북대학교 정보통계학과 · <sup>2</sup>남서울대학교 경영학과

접수 2011년 12월 29일, 수정 2012년 1월 13일, 게재확정 2012년 1월 17일

### 요약

그간 보험소비자의 욕구가 다양화, 세분화되어 소비자 특성과 성향에 따른 선호도 파악의 중요성이 증폭되어 왔다. 보험상품의 판매채널 및 마케팅 환경 변화는 고객이 무엇을 중요시하며 상품을 구입하는지에 대한 분석이 요구되지만 이러한 분야에 대한 구체적인 연구가 많이 이루어지지 않아 마케팅 전략 수립에 애로가 있다. 본 연구는 보험 소비자들의 기본성향과 보험가입 패턴에 대한 구체적 정보를 파악하기 위한 설문조사를 실시하였으며, 효율적 통계적 기법을 적용하여 신규 고객 확보를 위한 모형을 구축해 방카슈랑스의 성공적 정착과 활성화 방안을 강구해 보았다.

주요어: 방카슈랑스, 배깅 기법, 의사결정나무.

### 1. 서론

세계 금융산업은 1980년대의 금융자유화 이후 지속적인 규제완화와 개방화로 인해 업종별 장벽이 무너지고 대형화 및 금융 겸업 추세가 가속화되어 왔으며, 우리나라도 금융업 전반에 걸친 인수·합병 추세와 더불어 자유화, 겸업화, 다각화, 디지털혁명 등 금융혁신이 활발하게 이루어지게 되었다. 특히 고객이 필요로 하는 모든 금융거래와 서비스를 동시에 제공하는 원스톱 (one-stop) 종합금융서비스가 일반화되어 금융겸업화가 진척되고 은행, 증권 및 보험으로 분리 운영되던 금융산업의 업무 영역 간 차이가 모호해지면서 종합금융그룹화의 움직임이 나타나게 되었다. 1997년 6월 발표된 금융개혁 세부추진 방안은 금융겸업화를 더욱 촉진시키는 결정적인 요인으로 작용하였으며, 2003년 8월부터 은행과 보험 사와의 제휴를 통해 은행창구에서 고객들이 필요로 하는 금융서비스를 원스톱으로 제공하는 실질적인 방카슈랑스 (bancassurance)가 도입되었다. 이에 따라 은행창구를 통해서 보험상품의 판매가 가능하게 되었으며 본 논문에서는 은행을 통한 방카슈랑스의 판매채널에 초점을 맞추기도 한다. 판매채널 중 방카슈랑스에 대해서는 정세창과 이정환 (2003), 김재현 (2007), 김정동과 손민지 (2008), 정세창과 송영효 (2008) 등이 있으나 모두 보험사에 관련된 연구이다.

방카슈랑스 도입은 은행과 보험사 간의 판매채널 공유에 따른 경영 효율성, 은행과 보험상품의 상호 보완으로 인한 고객 니즈 충족, 고객 신용정보의 공유로 신규 고객 확보시 필요한 탐색비용 절감 등의 이점이 있으므로 효과적 정책이 수립되고 운영된다면 국내 금융산업에 매우 긍정적인 영향을 미칠 것으로 보인다. 은행의 안정성과 효율성에 관해 정재욱과 이지언 (2002), 나동민 (2006) 등의 연구가 있다. 이봉주 등 (2004)의 연구는 방카슈랑스를 통해 은행과 보험회사 모두 이익 효율성을 제고시키는 것으로 나타났다.

<sup>1</sup> 교신저자: (306-736) 충북 청주시 흥덕구 성봉로 410, 충북대학교 정보통계학과, 교수.

E-mail: thkim@chungbuk.ac.kr

<sup>2</sup> (306-736) 충북 청주시 흥덕구 성봉로 410, 충북대학교 정보통계학과, 석사과정생.

<sup>3</sup> (331-707) 충남 천안시 서북구 성환읍 대학로 91번지, 남서울대학교 경영학과, 교수.

금융산업의 자율화, 개방화에 따라 국내 보험시장은 금융산업 간 업무영역의 규제완화, 가격자유화, 새로운 판매채널의 등장과 1997년에 시행된 ENT (Economic Needs Test ; 경제적 수요심사)제도의 폐지 등으로 급격한 환경변화를 겪었으며 국내 보험사의 수적 증가와 더불어 외국 보험사와의 경쟁도 본격화되었다. 또한 보험소비자의 니즈가 다양화·세분화되고 있기 때문에 소비자 특성과 성향에 따른 선호도 파악의 중요성이 더욱 증폭되고 있다. 따라서 금융기관은 장기적인 고객확보책을 정립하고 신규 고객 확보를 위한 고객 지향적 마케팅 전략의 도입이 절실하게 되었다.

방카슈랑스가 도입되고 실질적으로 운영되고 있는 현시점에서 좀 더 현실적인 정보를 바탕으로 실증분석을 할 필요성이 요구되며, 따라서 신규 고객 확보를 위한 모형을 구축하고, 효율적인 통계적 기법을 적용하여 방카슈랑스의 도입효과와 활성화 방안에 대해 구체적으로 분석하고자 한다. 따라서 첫째 방카슈랑스의 개념과 도입배경 및 효과에 대해 간략히 살펴보고, 둘째 모형구축에 이용된 배경 기법에 대한 개념 및 알고리즘을 검토한다. 그리고 설문조사를 통해 얻어진 실제 자료를 사용해 배경 기법을 적용한 모형을 구축하고, 마지막으로 실증분석 결과를 토대로 방카슈랑스의 활성화 방안을 제시한다.

## 2. 방카슈랑스의 도입 배경과 연구방법

세계 금융산업의 겸업화에 부응하여 국내 금융기관의 효율성 증진과 경쟁력 제고를 위한 전략의 일환인 방카슈랑스는 넓은 의미로는 은행과 보험사의 업무제휴 전반을 지칭하는 것으로 그 자체만으로도 종합금융화의 의미를 가지게 된다. 이는 금융기간이 고유업무 영역 뿐 아니라 다른 영역에도 진출하여 금융기관 간 시너지 (synergy)효과를 창출함으로써 효율성을 제고시키며 다양한 욕구를 가진 고객에게 최대한의 선택폭을 부여함으로써 다각화된 금융서비스를 제공하기 위한 것이다.

은행의 경우 금리 자유화, 금융기관 간 경쟁 심화, 관리 및 마케팅 비용 상승 또 업무영역 비중의 축소등으로 인해 금융상품 판매와 증개를 통한 수수료 수입과 같은 새로운 수입원을 창출해야 할 필요성이 증가되었으며, 이는 곧 은행이 방카슈랑스를 적극적으로 도입하는 이유와 직결된다. 보험사도 방카슈랑스를 도입하면 기존 은행의 영업망과 고객을 기반으로 업무 다각화의 효율성을 기할 수 있으며 기존의 판매채널보다 비용절감을 통한 가격 경쟁력 향상으로 은행고객에 접근할 가능성이 커진다. 금융 소비자의 경우에는 동일한 점포에서 금융 및 보험서비스를 제공받을 수 있고 은행상품과 보험상품이 결합된 형태 등 소비자의 욕구에 부응할 수 있는 다양한 금융상품을 선택할 수 있어 편의성이 증가된다.

은행은 우월한 점포망과 낮은 유통비용 등 영업 경쟁력면에서 우월한 위치에 있으며 신뢰성 있는 이미지와 자산운용에서의 노하우가 강점으로 인식되는 반면 보험경영의 노하우, 보험업무를 위한 전문인력과 시스템이 미비하며, 보험상품 개발 및 언더라이팅 능력이 미흡하다는 약점이 있다. 은행이 업무영역 조정을 통한 경영효율성을 제고하며 복합금융상품을 개발하여 고객의 니즈에 부합하는 주문형 서비스체계를 구축한다면 보험시장 진출을 통해 영업기반을 확대하고 이를 더욱 개선·발전시킬 수 있을 것으로 분석된다.

본 연구에서는 409명의 20대 이상 성인 남·녀를 대상으로 금융시장의 환경변화로 인해 다양화·세분화되고 있는 보험소비자들의 기본 성향과 보험 가입 패턴, 또 방카슈랑스에 대한 인지도 등 구체적인 정보를 파악하기 위한 설문조사를 실시하였다. 이 결과를 바탕으로 데이터마이닝의 배경 기법을 이용하여 신규 고객 확보를 위한 모형을 구축하여, 국내에 도입된 방카슈랑스를 성공적으로 정착시키고 활성화시키기 위한 방안을 강구해 보기로 한다.

### 3. 배깅 기법을 이용한 실증분석

#### 3.1. 배깅의 개념 및 알고리즘

배깅 (bagging: bootstrap aggregation)은 Brieman (1996)에 의해 개발되었으며, 하나의 분석용 데이터 (training data)에서  $n$ 개의 붓스트랩 표본을 복원추출하여 생성된  $n$ 개의 모델을 통합 (voting)함으로써 최종모형을 얻는 기법으로 그림 3.1과 같은 알고리즘을 갖는다.

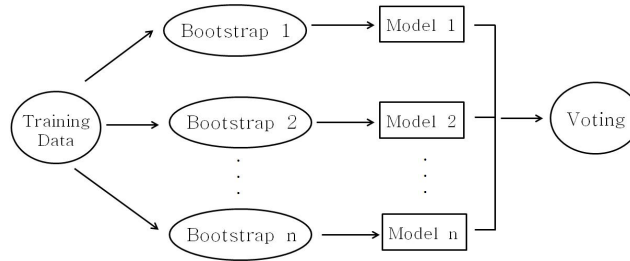


그림 3.1 배깅 알고리즘

분석용 데이터에서 복원추출을 통해 분석용 데이터와 동일한 표본 크기를 갖는 표본을 재추출 (resampling)하여 첫 번째 붓스트랩 표본을 만들고, 같은 방법으로 두 번째 붓스트랩 표본을 재추출하고, 반복하여  $n$ 개의 붓스트랩 표본을 만든다. 첫 번째 붓스트랩 표본으로 첫 모형을, 또 두 번째 붓스트랩 표본으로 둘째 모형을 만드는 식으로 전체  $n$ 개의 모형을 만든다. 여러 독립변수들이 종속변수를 예측하는 각기 다른  $n$ 개의 모형이 만들어지는데, 이 모형들이 가지는 예측값들의 평균으로 최종 예측치를 산출하게 되고 이에 따라 최종모형이 결정된다.

배깅은 복원추출에 의한 붓스트랩 추출 기법이므로 매 추출 시 선택 변수와 모형이 달라지게 되어 최종 모형에는 거의 모든 변수들에 대한 정보를 사용할 수 있게 되며, 따라서 모집단 정보에 근접할 수 있고 정확성을 극대화할 수 있다는 장점이 있다. 배깅은 또한 자료로부터 표본을 추출하여 모형이 생성되므로 표본에 따라 모형이 변하는 단일 의사결정나무의 불안정성을 해결할 수 있으며, 이는 단일 의사결정나무들을 결합하여 예측하는 배깅 기법을 이용할 경우 단일 의사결정나무보다 예측력이 증가하기 때문이다.

#### 3.2. 표본설계 및 기초자료 분석

본 연구에 이용되는 자료는 20세 이상의 전국 성인 남·녀를 대상으로 2006년 9월20일부터 15일 동안 개인별 직접 면접조사를 수행하여 409명의 표본을 수집하였다. 수집된 표본을 가지고 조사를 실시한 결과는 신뢰도 95%수준에서 오차한계가  $\pm 4.8\%$  이내인 것으로 조사되었다. 무응답 등 무성의한 설문지 1매를 제외하고 사용 가능한 설문조사 자료의 인구통계학적 특성 즉, 성별, 연령, 직업, 월 평균소득, 최종학력 등의 분포는 표 3.1과 같다.

보험가입에 대해 조사한 결과를 정리한 표 3.2를 살펴보면, 응답자의 82.6%가 보험에 가입되어있으며, 17.4%는 보험에 가입되어 있지 않는 것으로 나타났다. 보험에 가입된 응답자들 중에서 57.4%가 2개 이하의 보험에 가입하였으며, 한달 동안 평균적으로 납부하는 보험료는 10만원이상에서 30만원미만인 응답자가 가장 많게 조사되었다. 보험설계사의 방문판매로 보험에 가입하는 고객이 57.4%로 가장 많았으며, 은행과 같은 금융기관에서 보험에 가입하는 응답자는 10.9%로 비교적 낮게 나타났다.

표 3.1 인구통계학적 특성

변인	구분	응답자수 (명)	비율 (%)	변인	구분	응답자수 (명)	비율 (%)
성별	남자	209	51.2	결혼 여부	미혼	247	60.7
	여자	199	48.8		기혼	160	39.3
연령	20세 ~ 29세	225	55.1	월평균소득 (단위: 만원)	100 미만	82	20.4
	30세 ~ 39세	105	25.7		100 ~ 200 미만	207	51.5
	40세 ~ 49세	46	11.3		200 ~ 300 미만	91	22.6
	50세 이상	32	7.9		300 이상	22	5.5
직업	공무원	33	8.1	주거 형태	자택	250	61.4
	판매/서비스직	41	10.1		전세	92	22.6
	전문직/기술직	143	35.2		월세	22	5.4
	사무관리직	87	21.4		기타	43	10.6
	생산/운전직	16	3.9	학력	고졸이하	91	22.3
	자영업	22	5.4		대졸이하	290	70.9
직업	농어업종사자	1	0.2	대학원이상	26	6.4	
	가사	21	5.3	자동차 소유 여부	소유	201	49.3
	학생	25	6.2		미소유	207	50.7
	기타	17	4.2				

표 3.2 보험가입실태

변인	구분	응답자수 (명)	비율 (%)	변인	구분	응답자수 (명)	비율 (%)
보험가입여부	가입	338	82.6	보험가입방법	보험설계사	194	57.4
	미가입	71	17.4		보험대리점	43	12.7
가입된 보험상품수	1개	89	26.3	금융기관	37	10.9	
	2개	105	31.1	친분관계	60	17.8	
	3개	70	20.7	인터넷	4	1.2	
	4개 이상	74	21.9	보험회사이미지	18	4.5	
월평균보험료 (단위:만원)	10 미만	103	30.6	보험상품종류	193	48.3	
	10 ~ 30 미만	157	46.6	보험료	83	20.8	
	30 ~ 50 미만	53	15.7	고객관리서비스	96	24.0	
	50 이상	24	7.1	기타	10	2.4	

방카슈랑스에 관한 부문에서는 160 (39.3%)명이 향후 은행에서 보험을 가입할 의사가 있는 것으로 나타났고, 나머지 247 (60.7%)명은 은행에서 보험을 가입하는 것에 대해 부정적인 반응을 보였는데 그 이유로는 은행이 보험전문기관이 아니므로 신뢰성이 의심 (34.7%), 현재 보험가입방법에 만족 (22.4%), 고객센터에 대한 불확신 (17.4%), 직접방문이 귀찮음 (13.8%) 순으로 나타났다.

### 3.3. 배경을 이용한 가입가망 모델링

방카슈랑스를 도입한 은행에서 보험에 가입할 신규 고객 확보를 위한 모형을 구축하기 위해 은행에서의 보험가입 희망 여부를 목표변수로 선정하였다. 그리고 설명변수로 사용된 요인들은 고객정보 사항으로 성별, 연령, 최종학력, 결혼여부, 월 평균소득, 직업, 주거형태, 자동차 소유여부, 주당 은행방문횟수를 또 보험 관련 사항으로는 보험가입여부, 가입된 보험상품의 수, 가입된 보험상품의 종류, 월 평균 보험료를 선택하였다.

은행을 통해 보험에 가입할 고객을 예측하기 위해 SAS Enterprise Miner를 사용하며, 본 연구에서는 의사결정나무의 알고리즘 중 목표변수가 범주형일 때 사용되는 CART, C4.5, CHAID 알고리즘에 배경 기법을 적용하여 모형을 구축해 보았다. 분석기법은 Kass (1980), Brieman (1996), Berry와 Linoff (1997), Denison 등 (1998), Kim (1999), 강현철 등 (2000), 장남식 등 (2000)에 근거한다. 본 설문연

구에서 향후 은행에 보험을 가입할 의사가 있는 사람이 160 (39.3%)명이므로 배경은 C4.5 알고리즘과 CART, CHAID 알고리즘에 각각 25회와 50회로 반복을 실시하여 6개의 모형을 생성하며, 각 알고리즘은 각각 Entropy 지수, Gini 지수, 또  $\chi^2$  통계량을 분리기준으로 사용한다. 그림 3.2는 E-Miner를 통해 모형을 구축한 분석흐름도를 나타낸다.

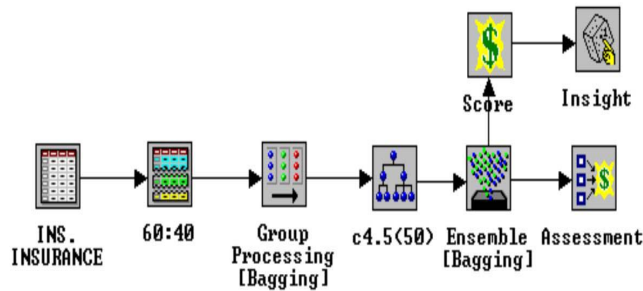


그림 3.2 C4.5 (50)의 분석흐름도

구축한 모형의 결과를 정리한 표 3.3을 보면, 절단값을 0.4로 하여 각 모형의 예측력을 비교했을 때 반복횟수를 50회로 한 C4.5 모형의 예측력이 0.75로 가장 높은 것으로 나타났다. 또한 모형의 누적향상도를 통해 모델링을 하지 않은 경우와 모델링 한 경우의 상대적 효과를 비교해 본 결과 반복횟수를 50회로 한 C4.5 모형이 상위 10%, 20%, 30%에서 가장 우수한 것으로 나타났다.

표 3.3 배깅기법을 이용한 모형의 예측력과 향상도

절단값	CART (25)		CART (50)		C4.5 (25)		C4.5 (50)		CHAID (25)		CHAID (50)	
	예측력	향상도	예측력	향상도	예측력	향상도	예측력	향상도	예측력	향상도	예측력	향상도
0.1	0.63	1.020	0.75	1.218	0.75	1.218	0.82	1.317	0.72	1.122	0.75	1.218
0.2	0.63	1.020	0.53	1.040	0.72	1.188	0.84	1.337	0.63	1.223	0.72	1.268
0.3	0.70	1.056	0.73	1.086	0.73	1.185	0.67	1.254	0.61	1.112	0.66	1.199
0.4	0.55	1.015	0.50	1.020	0.63	1.144	0.75	1.243	0.60	1.115	0.63	1.142
0.5	0.60	1.007	0.66	1.030	0.87	1.198	0.59	1.183	0.67	1.132	0.68	1.135
0.6	0.61	1.003	0.61	1.023	0.57	1.153	0.63	1.155	0.39	1.051	0.60	1.107
0.7	0.45	0.963	0.73	1.047	0.58	1.124	0.67	1.146	0.60	1.039	0.60	1.087
0.8	0.73	0.990	0.63	1.045	0.39	1.064	0.37	1.077	0.58	1.027	0.65	1.082
0.9	0.67	1.001	0.69	1.053	0.49	1.034	0.43	1.034	0.65	1.031	0.43	1.040

- 1) 모형을 통해 계산된 가입확률이 절단값 0.4 이상이면 가입으로 판단, 여기서 절단값은 설문조사 결과로 은행에서 보험을 가입할 의사가 있는 표본의 비율이 39.3%인 것을 근거로 산정.
- 2) 누적향상도 = 해당등급에서 가입의 비율/전체표본에서 가입의 비율

예를 들어 모형의 향상도가 2일 때 모형을 고려하지 않았을 경우 100명의 고객 중 10명이 은행에서 보험에 가입한다면, 모형을 고려하는 경우에는 그 두 배인 20명이 은행에서 보험에 가입할 수 있음을 의미한다. 또한 상위 10%라는 것은 각 모형을 통해 모든 대상자의 은행에서 보험가입 확률을 계산했을 때 그 확률값이 상위 10%인 대상자를 의미한다. C4.5 (50)모형의 상위10% 고객을 대상으로 마케팅 및 홍보활동을 한다면 모형을 고려하지 않은 경우보다 1.317배의 고객이 은행에서 보험을 가입하게 되는 것을 알 수 있었으며, 상위 20%의 고객을 대상으로 마케팅을 한다면 1.337배의 더 많은 고객이 은행에서 보험에 가입하는 것으로 나타났다.

본 연구에서는 배깅모형의 예측력과 향상도를 분석했으나 다른 비교 대상 모형에 비해 어느 정도 안정적인 결과를 보였는지는 논하지 않았다. Cho와 Park (2011)은 교통사고율을 목표변수로 지정한 의사결

정나무모형과 다중매개연관성규칙을 이용하여 불필요한 입력변수를 제외한 모형과의 정확도를 비교한 결과 간결화한 모형이 원모형에 비해 예측정확도에서 약간 떨어지지만 모형의 생성 및 해석 시간과 노력은 단축되는 효율성을 갖는다고 주장하였다. CART, C4.5, CHAID 알고리즘에 배깅 기법을 적용해 모형을 구축한 결과가 주는 정보를 더 잘 이해하기 위해 그래프로 나타낸 결과는 그림 3.3 및 그림 3.4와 같다.

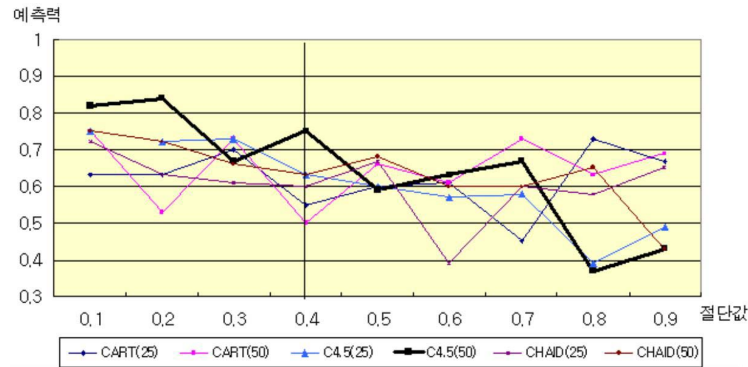


그림 3.3 배깅 기법을 이용한 모형의 예측력그림

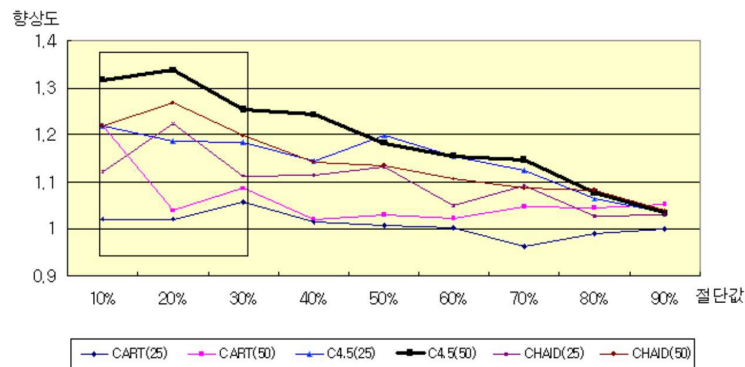


그림 3.4 배깅 기법을 이용한 모형의 향상도

### 3.4. 가입가망 모델링 결과

앞에서 각 모형의 예측력과 향상도를 살펴본 결과, 6가지 모형 중에서 은행에서 보험을 가입할 신규 고객을 확보하기 위해 가장 적합한 모형은 C4.5 (50) 알고리즘에 배깅 기법을 적용한 모형이라는 것을 알 수 있었다. 인공지능망 모형이나 로지스틱 회귀모형에 비해 분석 결과 도출된 배깅모형의 강점은 의사결정나무 그림을 통해 가입규칙을 쉽게 확인할 수 있다는 것이다. 김윤종과 김용철 (2008)은 설문조사를 통해 예측된 수요와 로지스틱 회귀모형을 통해 예측된 수요를 혼합한 수요예측 모형을 제시한 바 있다. Kang과 Jung (2009)은 요인분석과 회귀분석을 사용하여 생명보험 계약자의 만족도, 신뢰도, 충

성도에 영향을 미치는 요인을 분석하였다. Cho 등 (2009)은 고객충성도 예측모형으로 로지스틱 회귀 모형, 의사결정나무 및 신경망 모형을 사용하였다. Park 등 (2011)은 의사결정나무 분석을 적용시킨 예측모형과 업무규칙을 혼합한 하이브리드모형을 국민건강보험 관리모형으로 개발하였다. Cho와 Park (2011)은 의사결정나무 생성 시 목표변수와 입력변수에 대한 관계를 명확히 파악할 수 있는 다중매개연관성규칙을 적용하여 불필요한 입력변수를 제거할 수 있는 방법을 제안하였다.

그림 3.5는 C4.5 알고리즘에 의해 생성된 50개의 의사결정나무 중에서 최적의 모형을 한 개 선택한 것이다. 은행을 통한 보험의 가입여부에 영향을 주는 변수로 직업, 가입한 보험상품 수, 연령, 월 평균 소득, 주당 은행방문횟수가 선택되었다. 도출된 의사결정나무의 가입규칙을 자세히 살펴보면 다음과 같다.

- (i) 직업이 판매/서비스직이거나 자영업, 학생인 고객은 은행에서 보험을 가입할 확률이 65.5%인 것으로 나타났고, 직업이 공무원, 전문직/기술직, 생산/운전직, 사무관리직, 가사가 직업인 고객은 가입확률이 35.3%인 것으로 나타났다.
- (ii) 판매직/서비스직, 자영업, 학생의 직업을 가진 고객 중 가입한 보험상품 수가 1개인 경우 은행을 통해 보험에 가입할 확률이 86.2%인 것으로 나타났고, 2개 이상의 보험에 가입한 고객은 가입확률이 13.8%로 나타났다.
- (iii) 판매/서비스직, 자영업, 학생의 직업을 가진 고객 중 가입한 보험상품 수가 1개 이고, 월 평균 소득이 100만원 이상인 고객이 은행에서 보험을 가입할 확률이 95%인 것으로 나타났다.
- (iv) 공무원, 전문직/기술직, 생산직/운전직, 사무관리직, 가사의 직업을 가진 고객 중 연령이 20-30대 이고, 주당 4회 이상 은행을 방문하는 고객이 은행에서 보험을 가입할 확률이 92.3%인 것으로 나타났다.

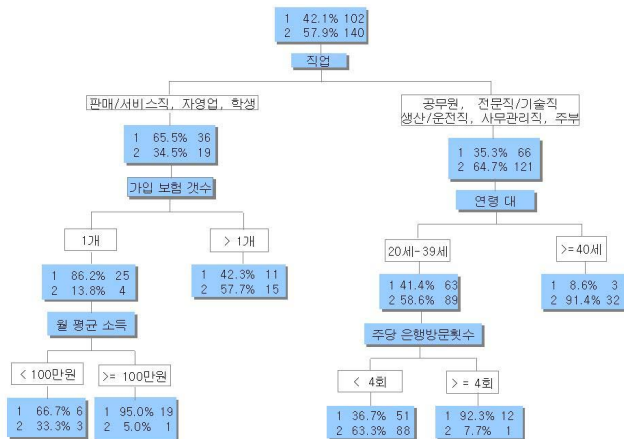


그림 3.5 배깅 기법에 의한 의사결정나무

방카슈랑스에 대한 인식조사를 한 결과를 보면 전체 응답자의 46.6%가 들어본 적이 없다고 응답을 하였고, 들어본 적은 있으나 구체적으로 모른다가 38.2%, 잘 알고 있다가 15.2%로 나타났다. 방카슈랑스의 도입으로 인해 유발되는 긍정적인 면을 부각시켜 방카슈랑스에 대한 인식을 확산시키고, 제도의 성공적인 조기 정착과 시행효과를 극대화하기 위해서는 적극적인 전략이 수립 되어야 한다. 본 연구의 가입가망 모델링 및 고객 니즈 조사결과를 통한 방카슈랑스의 활성화 방안을 살펴보면 아래와 같다.

방카슈랑스 전략의 추진에 가장 중요한 것은 대상 고객을 정하는 것으로 본 연구에서 제안한 배깅을 이용한 신규고객 예측모형을 통해 목표고객을 설정하였다. 은행에서 보험의 가입 여부에 가장 큰 영향을 미치는 변수는 고객의 직업으로 나타났으며, 판매/서비스직, 자영업, 학생의 직업을 가진 고객의 65.5%가 은행에서 보험을 가입하겠다는 의사를 보였다. 이 외에도 현재 가입된 보험상품수가 1개인 경우 가입확률이 높았고, 월 평균 소득이 100만원이상인 경우에, 연령은 20-30대인 경우에 그리고 주당 은행 방문횟수가 4회 이상인 경우에 은행에서 보험을 가입할 확률이 높게 나타났다. 이와 같이 세분화된 고객유형별 가치, 성향, 라이프 스타일에 따른 마케팅 전략과 대상 고객들에 관한 정보와 니즈를 기반으로 특화된 보험상품과 서비스를 개발하여 고객들에게 제안해야 할 것이다.

#### 4. 결론

본 연구에서는 방카슈랑스를 효율적으로 추진하기 위한 신규고객 모형 구축을 통하여 방카슈랑스를 활성화시키는 방안을 제시하였다. 409명의 표본을 대상으로 한 설문조사 자료를 토대로 은행에서 보험을 가입할 고객을 예측해 보기 위해 CART, C4.5, CHAID 알고리즘에 배깅 기법을 적용하여 모델링을 하였다. 그 결과 반복횟수를 50회로 한 C4.5 모형이 예측력과 향상도가 가장 높은 것으로 나타나 이 모형을 최적모형으로 고려하여 분석을 실시하였다.

분석 결과, 은행에서 보험의 가입여부에 가장 큰 영향을 미치는 변수가 고객의 직업으로 선택되었으며, 판매/서비스직, 자영업, 학생의 직업을 가진 고객 중에는 현재 가입한 보험상품의 수가 1개이고 월 평균 소득이 100만원이상인 고객이 은행에서 보험을 가입할 확률이 95%로 가장 높게 나타났다. 또한 공무원, 전문직/기술직, 생산/운전직, 사무관리직, 가사의 직업을 가진 고객 중에서 20-30대이고, 주당 은행을 방문하는 횟수가 4회 이상인 고객이 은행을 통해 보험에 가입할 확률이 92.3%에 달하였다. 이와 같이 배깅 모형에 의해 얻어진 세분화된 고객의 정보에 기반하여 체계적이고 새로운 마케팅 전략을 추진하고, 고객들 각각의 니즈와 특성을 충족시킬만한 다양한 보험상품 및 금융상품을 제시하고 지속적인 종합금융서비스를 제공한다면 방카슈랑스 도입으로 인한 긍정적인 효과를 기대할 수 있을 것이다.

#### 참고문헌

- 강현철, 한상태, 최중후, 김은석, 김미경 (2000). <SAS Enterprise Miner를 이용한 데이터마 이닝>, 자유 아카데미, 파주.
- 김윤중, 김용철 (2008). 서비스 수요조사와 분류모형을 이용한 수요 예측. <응용통계연구>, **21**, 151-157.
- 김정동, 손민지 (2008). 방카슈랑스 도입에 따른 생명보험회사의 효율성 변화. <보험학회지>, **17**, 61-92.
- 김재현 (2007). 방카슈랑스 도입에 따른 생명보험회사의 비용 효율성 및 생산성 변화 연구. <보험개발연구>, **14**, 3-40.
- 나동민 (2006). 방카슈랑스 확대에 따른 은행의 안정성 및 효율성 변화 분석. <한국개발연구>, **28**, 49-93.
- 이봉주, 이순재, 정세창 (2004). 방카슈랑스 도입이 금융겸업화 효율성에 미치는 영향. <경영학연구>, **33**, 449-472.
- 장남식, 홍성완, 장재호 (2000). <데이터마이닝>, 대청, 파주.
- 정세창, 이정환 (2003). 보험회사에 대한 방카슈랑스 도입 효과 및 시사점 분석. <보험개발연구>, **14**, 93-125.
- 정세창, 송영호 (2008). 방카슈랑스 효과와 생명보험회사의 경영전략. <보험개발연구>, **17**, 3-34.
- 정재욱, 이지연 (2002). 방카슈랑스가 은행의 안정성 및 효율성에 미치는 영향. <금융조사보고서>, **8**, 한국금융연구원.
- Berry, M. J. A. and Linoff, G. (1997). *Data mining techniques for marketing, sales, and customer support*, John Wiley and Sons, New York.
- Brieman, L. (1996). Bagging predictors. *Machine Learning*, **24**, 123-140.
- Cho, D., Kim, B., Lee, J., Kim, J. and Kim, S. (2009). A study on the behavior of cosmetic customers. *Journal of the Korean Data & Information Science Society*, **20**, 615-627.



- Cho, K. H. and Park, H. C. (2011). A study on decision tree creation using intervening variable. *Journal of the Korean Data & Information Science Society*, **22**, 671-678.
- Denison, D. G. T., Mallick, B. K. and Smith, A. F. M. (1998). A Bayesian CART algorithm. *Biometrika*, **85**, 363-377.
- Kang, J. C. and Jung, S. C. (2009). An analysis of the factors influencing satisfaction, reliance, and loyalty to the life insurance companies. *Journal of the Korean Data & Information Science Society*, **20**, 713-717.
- Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, **29**, 119-127.
- Kim, S. H. (1999). *Data mining in finance*, Sigma Consulting Group, Louisiana.
- Park, I., Han, J., Sohn, H. and Kang, S. (2011). Developing the administrative model using the data mining technique for injury in national health insurance. *Journal of the Korean Data & Information Science Society*, **22**, 467-476.
- Quinlan, J. R. (1993). *C4.5 programs for machine learning*, Morgan Kaufmann Publishers, Massachusetts.

## Bagging consumer modeling for successive growth and establishment of bancassurance

Tae-Ho Kim<sup>1</sup> · Jae Hwa Jung<sup>2</sup> · Jin Soo Kim<sup>3</sup>

<sup>1,2</sup>Department of Information Statistics, Chungbuk National University

<sup>3</sup>Department of Business Administration, Namseoul University

Received 29 December 2011, revised 13 January 2012, accepted 17 January 2012

### Abstract

As insurance consumers' needs have been diversified and subdivided, it is increasingly important to grasp their preferences by characteristics and properties. Even though changes in sales channels and marketing conditions of insurance require to analyze what consumers take serious views to purchase, it is difficult to devise marketing strategies since not many concrete studies have been done in this field. A questionnaire survey was carried out to learn detailed information about basic disposition and buying patterns of insurance consumers. Applying efficient statistical techniques and then utilizing a model for securing new customers, this study attempts to explore a plan for rapid growth and successive establishment of bancassurance.

*Keywords:* Bagging, bancassurance, decision tree.

---

<sup>1</sup> Corresponding author: Professor, Department of Information Statistics, Chungbuk National University, 410 Sungbong-Ro, Heungduk-Gu, Cheongju, Chungbuk 361-763, Korea.

E-mail: thkim@chungbuk.ac.kr

<sup>2</sup> Graduate student, Department of Information Statistics, Chungbuk National University, 410 Sungbong-Ro, Heungduk-Gu, Cheongju, Chungbuk 361-763, Korea.

<sup>3</sup> Professor, Department of Business Administration, Namseoul University, Cheonan, Chungnam 331-707, Korea.