

생명보험사 텔레마케팅 효율성 제고에 관한연구

고봉성¹ · 이석원² · 허정³

¹²³전주대학교 여론정보통계학과

접수 2009년 5월 1일, 수정 2009년 7월 12일, 게재확정 2009년 7월 16일

요약

텔레마케팅의 영업특성상 전화 응대를 한 고객에게 초기 적절한 상품을 추천하여 보험가입을 유치하는 것이 가장 중요하다. 따라서 고객의 선호상품 및 선호시간을 고려한 고객 선별과 타겟팅을 어떻게 진행하느냐에 따라 거수보험료 차이가 발생한다. 본 연구는 데이터마닝 기법 중 로지스틱회귀 모형을 이용하여 생명보험사 아웃바운드 텔레마케팅 영업을 지원하는데 있어서 그 효과를 검증하고자 하였다. 기존의 가입상품에 대한 고객반응정보와 L생명보험사 영업 전략에 따라 연령을 구간화하고 가족사랑보험, 상해보험, 암보험 상품에 대한 모델링을 진행하였다. 설정된 모델을 바탕으로 캠페인 진행시 기존 마케팅 방식과 모델방식으로 고객을 추출하고 실행결과를 통해 모델의 우수성을 입증하였다. 또한 시간이 지남에 따라 설정한 모델이 노후화되고 영업이익이 줄어들어 따라 지속적으로 영업이익을 극대화 할 수 있도록 모델갱신주기를 도출하였다.

주요어: 로지스틱회귀모형, 반응률, 생명보험, 아웃바운드, 텔레마케팅.

1. 서론

최근 우리나라는 사망률과 출생률의 감소로 급격한 고령화 사회로 진입함에 따라 의료비용도 증가추세에 있다. 특히, 암, 뇌혈관질환, 치매 등의 질병은 치료의 장기화가 요구되고 지속적인 간병과 소득상실로 인한 금전적, 정신적 피해가 동반됨에 따라 미래에 발생할지 모르는 불의의 질병과 사고에 대비한 실손 및 보장성 보험에 대한 소비자의 관심이 두드러지게 증가하고 있는 상황이다.

그럼에도 불구하고 영업경쟁력의 한계와 다국적 생명보험사, 다이렉트 생명보험사의 시장진입 등으로 경쟁이 심화되고 있으며, 금융감독원 (2009)에 의하면 글로벌 금융위기로 인한 생명보험사 당기 순이익은 6,172억원으로 전년 (2008년, 2조1,055억원)보다 70% (1조4,883억원)이상 감소하였고, 자본확충 노력에도 불구하고 재무건전성은 크게 훼손되었다. 따라서 생명보험사들은 수익성 개선을 목적으로 소비자 수요에 맞게 고객을 세분화하고 다양한 마케팅전략을 수행할 것으로 생각되며, 인적비용과 유지비용이 크게 소요되는 설계사, 대리점 등 대면영업 중심보다는 고객의 라이프스타일과 접촉시간을 고려한 이메일, 전화 등을 활용한 비대면영업의 비중을 늘려갈 것으로 판단된다.

생명보험협회 (초회 보험료 기준) 자료에 의하면 전화를 이용한 텔레마케팅은 다이렉트 보험사, 홈쇼핑 등에서의 매출 호조로 2004년 점유율 0.4% (175/49,961억)에서 2008년 점유율 4% (1,525/38,180억)로 점유율과 보험료 모두 급성장을 이루고 있고, 텔레마케팅에 적합한 보험상품개발, 텔레마케터 교육,

¹ 교신저자: (560-759) 전북 전주시 완산구 효자동3가 1200, 전주대학교 여론정보통계학과, 교수.
E-mail: bskoh@jj.ac.kr

² (560-759) 전북 전주시 완산구 효자동3가 1200, 전주대학교 여론정보통계학과, 겸임교수.

³ (560-759) 전북 전주시 완산구 효자동3가 1200, 전주대학교 일반대학원 통계학과, 석사과정.

센터 운영, DB마이닝 등의 전략적 접근을 시도하고 있다. 특히 DB마이닝은 기존고객의 과거행동을 바탕으로 고객데이터의 분석을 통해 미래의 보험가입 가망고객을 선별하는 예측모형으로 단기간에 효과를 볼 수 있다는 점에서 선호되고 있다.

DB마이닝에서 로지스틱회귀모형을 이용한 실증연구는 많은 분야에서 폭넓게 응용되고 있으며, 이현정 (2001)은 보험사데이터를 기반으로 신경망, 의사결정나무, 로지스틱회귀모형의 예측력 비교실험을 수행하였으며, 이호영 (2005)은 보험사 이탈고객예측모형을 통해 이탈고객에 대한 스코어를 산출하였고, 윤미례 (2004)는 기업의 부도예측모형을 로지스틱회귀모형과 생존분석을 통해 모형 간 효율성을 비교하였고, Cho와 Lee (2004)는 로지스틱회귀계수를 신경망에 대입하여 신경망의 초기 가중치에 따른 단점을 보완하였고, 김미정과 김진형 (2008)은 로지스틱회귀모형을 이용하여 유족연금 발생을 예측하고 유족연금발생 가능성 정도에 따라 대상자들을 분류하였다.

본 논문에서는 텔레마케팅에서 보험수율 (DB대비 보험가입자수)을 높이기 위하여 로지스틱회귀모형을 이용하여 모델을 생성한 후 현재 진행하고 있는 마케팅 형태와 모델간의 비교를 통해 그 효과를 검증하고자 한다.

2. L사 텔레마케팅 운영과 사전분포

L보험사는 전통적으로 텔레마케팅을 통한 보험영업을 주로 진행하는 보험사이다. 텔레마케팅을 진행하기 위해서는 첫째, 다량의 DB를 보유한 업체와의 제휴관계를 맺어야 한다. 보험사는 보험가입자 고객만을 관리하기 때문에 교차판매는 가능하지만, 신규 영업을 위해서는 양질의 DB를 보유한 대형카드사, 홈쇼핑, 인터넷쇼핑몰, 정유사 등과의 전략적 제휴가 필수적이다. 이들 업체는 공급하는 일정한 DB에 따라 수수료를 지급받으며 마케팅 성과에 따라 추가적인 이익을 남길 수 있어 파트너사로서의 공존이 가능하다.

둘째, 텔레마케팅에 필요한 콜센터 (call center)이다. 콜센터는 고객이 스스로 보험의 필요성에 의해 보험사 대표전화번호로 문의를 할 때 적극적으로 보험상품을 판매하는 인바운드 (inbound) 콜센터가 있다. 인바운드 콜센터는 고객의 문의건수에 대한 예측을 통해 텔레마케터의 수를 배분하는데 최우선적 관심이 있다. 이와 반대로 일정한 DB수를 제휴사로부터 부여받고 텔레마케터를 모집하고 교육시켜 고객의 라이프스타일과 접점을 파악하여 보험상품 가입을 유도하는 아웃바운드 (outbound) 콜센터가 있다. 아웃바운드 콜센터는 다시 원콜 (one-call)과 듀얼콜 (dual-call)로 나뉘는데 원콜은 1차에서 보험상품 가입을 권유할 수 있는 텔레마케터가 바로 상품에 대한 설명과 함께 보험가입을 유도하는 방식이며, 듀얼콜은 1차에서 고객에게 동의여부만 묻고 5초이내에 보험상품 가입을 권유할 수 있는 텔레마케터에게 전환한 후 보험상품 판매를 유도하는 방식이다.

셋째, 양질의 DB선택이다. 가입자수가 많은 제휴사 DB를 일정기간에 다 사용할 수 없으므로 보험이 필요하거나 필요하다고 느낄 수 있는 고객을 선별하여 접촉률을 높일 수 있는 양질의 DB를 선별하는 것이다. 이를 위하여 통계적 모델이 필요한 것이다.

본 논문에서는 L생명보험회사의 제휴사인 S주식회사의 DB를 사용하여 듀얼콜 마케팅 방식에 적용될 가족사랑보험 (24~34세), 상해보험 (35~39세), 암보험 (40~54세)에 대하여 모델을 진행하기로 한다. 연령대를 나누어 상품을 배정한 이유는 주력상품마다 판매단가를 2~3만원대로 설계하여 고객을 쉽게 유인하기 위한 L생명사 정책에 따른 것이다. 2007년 6월에서 2007년 8월까지의 모델설정을 위한 분석용데이터로 사용하고 2007년9월에서 2007년10월까지의 검증데이터로 이용하였다. S주식회사의 정보는 회원정보, 통신정보, 주유/카드정보, 기타정보로 나눌 수 있으며, 주요 변수 정보는 표 2.1과 표 2.2와 같다.

표 2.1 S사 주요변수 내역

구분	주요 변수 정보
회원정보	M월회원등급, 가용/누적포인트, 회원번호, 직업구분, 성별, 생년월일, 카드번호, 자택/이동전화번호, 결혼여부, 카드종류, 온라인회원여부, 우편번호, e-mail주소, 아파트가격, 아파트평수 등
통신정보	통신정상카드보유여부, EBC정상카드보유여부, CB제휴신용카드보유여부, TTL, UTO, CARA, Leaders, Ting정상카드보유여부, 통신매출 등
주유/카드정보	주유매출금액, 주유회수, 신용카드 금액, 신용카드 회수,
기타정보	3/6개월평균먹거리업종 매출/회수, 3/6개월평균 살거리업종 매출/회수, 3/6개월평균 놀거리업종 매출/회수, 3/6개월평균 생활/건강업종 매출/회수 등

표 2.2 S사 기준에 의한 온·오프라인 제휴사 분류

업종 분류	업종
먹거리	주점, 다과점, 음식점
살거리	예술품/선물용품, 식료품, 신변잡화, 의료직물, 서적/문구/인쇄/복사, 의약품, 통신판매, 컴퓨터 사무용기기, 차량용연료/기타연료, 인테리어/건축자재, 자동차/오토바이, 오락/취미, 가전용품/통신기기, 유통점 등
생활/건강	교육기관/교습학원, 광고/이벤트, 개인간병/유사서비스, 의료시설, 여행사, 수리, 관광/숙박시설, 운송업(여객/화물), 법무/회계관리, 이미용/욕탕/물매관리, 예식준비/예식관련 등
놀/쉼거리	공연/전시, 오락 시설, 운동/여가시설, 영화/비디오, 유원지/놀이공원, 놀거리기타 등
온라인	전자상거래, 포인트 판매 등

보험상품별 판매 대상인 연령을 구간화하여 주요지표의 평균값을 살펴본 결과 24~34세 연령대의 경우 6개월평균통신매출금액 (18,465원), 기타정보인 6개월평균먹거리업종매출금액 (1,989원), 6개월평균놀거리/쉼거리업종매출금액 (1,258원), 6개월평균생활/건강업종매출금액 (782원)에서 사용량이 가장 크게 나타나고 있다. 이는 S사 제휴가맹점 이용율이 가장 높은 층으로 대변되기 때문에 사료된다. 35~39세 연령대의 경우 가용포인트 (9,127), 6개월평균주유매출금액 (46,946원), 6개월이용주유소 개수 (1.2회), 6개월평균살거리업종매출금액 (105,612원)으로 외식 및 여행에 대한 라이프스타일을 반영한 것으로 생각되며, 마지막으로 40~54세 연령대의 경우 아파트평수 (14.1), 아파트가격 (12,794만원)등이 높게 나타나고 있어 생활의 안정성이 반영된 것으로 판단된다.

표 2.3 S사 주요 변수 지표

구분	변수항목	24~34세	35~39세	40~54세	전체
회원정보	가용포인트	5431(19218)	9127(15108)	4301(4177)	6187(16921)
	아파트 평수	10(15)	10.5(14.8)	14.1(17.2)	10.8(15.5)
	아파트 가격(만원)	8388(19487)	7684(16600)	12794(24180)	8850(19651)
통신정보	6개월평균통신매출금액	18465(24265)	11805(21755)	11049(17907)	15735(23083)
	6개월평균통신매출회수	0.4(0.5)	0.3(0.4)	0.3(0.5)	0.4(0.4)
주유/카드정보	6개월평균주유매출금액	13407(80240)	46946(18651)	25263(116362)	23467(121318)
	6개월평균주유매출회수	0.3(1.1)	0.8(2.2)	0.4(1.7)	0.4(1.5)
	6개월이용주유소개수	0.6(1.9)	1.2(2.8)	0.8(2.1)	0.8(2.2)
	6개월평균신용카드매출금액	4732(36927)	9385(78260)	3676(41622)	5737(51051)
	6개월평균신용카드매출회수	0.1(0.4)	0.1(0.6)	0.1(0.5)	0.1(0.5)
기타정보	6개월평균먹거리업종매출금액	1989(5759)	1080(4365)	642(2949)	1568(5138)
	6개월평균살거리업종매출금액	42423(170511)	105612(355973)	65855(239142)	61536(241132)
	6개월평균놀거리/쉼거리업종매출금액	1258(4572)	381(2279)	317(12089)	904(5942)
	6개월평균생활/건강업종매출금액	782(11089)	483(8411)	357(8125)	647(10088)

(단위: 원, 회수)

3. 로지스틱회귀모형을 통한 모형구축

종속변수가 범주형 변수일 때 비선형 함수형태의 회귀모형을 설정할 수 있다 독립변수의 수준이 변화할 때 종속변수의 기대값은 0과 1로 수렴할 수 있을 것이다 (성웅현, 2001; 정광모 외, 2003). $P(Y_i = 1 | x_i)$ 을 아래와 같은 승산 (odds)로 변환하면,

$$\Omega(x_i) = \frac{P(Y_i = 1|x_i)}{P(Y_i = 0|x_i)} = \frac{P(Y_i = 1|x_i)}{1 - P(Y_i = 1|x_i)}.$$

$\Omega(x_i)$: 어떤 사건이 발생되지 않을 확률에 대한 발생될 확률의 비율.

이 때 자연대수를 취하면 $\Omega(x_i)$ 의 범위는 $-\infty \leq \Omega(x_i) \leq \infty$ 을 가지며 모형은 로짓에 대하여 선형값을 가진다. 그리고 로짓에 대하여 선형이 되는 이진 로지스틱회귀모형을 구축하면 확률 산출식은 다음과 같이 표현할 수 있다 (이성우외, 2005).

$$P = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}.$$

P : 텔레마케팅시 고객이 보험상품에 가입할 확률, X_k : 보험상품가입 확률에 대한 영향 변수, β_k : 영향 변수의 상관성 및 방향성 계수 (값이 클수록 상관성이 크며, (+)이면 양의 방향으로, (-)이면 음의 방향의 상관성을 가짐).

로지스틱회귀모형에서 목적성 변수는 보험가입여부 (Target)로 보험가입자는 1, 비가입자는 0의 값을 가지며, 설명변수 중 범주형 변수와 연속형 변수 중 구간을 나눈 변수들로 더미 변수를 생성하였으며, 지역, S주식회사 등급, 직업, 연령대, 차종 등이다 (Kim과 Lee, 2008).

이 때, 로지스틱회귀모형에서는 모델의 파라미터를 최대우도법을 사용하여 측정하게 되며, 관찰결과를 가장 그럴듯하게 하는 계수가 선택된다 (이호영, 2005). L생명보험사 텔레마케팅에서 S주식회사의 고객데이터를 이용한 회귀모형계수는 표 3.1과 같다.

표 3.1의 SAS 결과값을 기초로 추정된 회귀계수를 이용하여 구매 고객이 될 확률이 큰 고객을 다음과 같이 예측하였다.

$$\hat{Y}_{\text{가족사랑보험}} = -3.1433 + (0.000007246 * 6\text{개월평균통신매출금액}) + (-0.7906 * 6\text{개월평균통신매출회수}) \\ + (0.1909 * \text{대구}), \dots, + (-0.1063 * (\text{가용포인트} < 935.1)).$$

$$\hat{Y}_{\text{상해보험}} = -4.1569 + (0.0000002265 * 3\text{개월평균신용카드매출금액}) + (1.1871 * \text{생계형}) \\ + (0.5686 * 0(\text{비거주})), \dots, + (-0.00003 * 3\text{개월평균마트매출금액}).$$

$$\hat{Y}_{\text{암보험}} = -3.3228 + (-0.0413 * \text{아파트평수}) + (-0.00764 * \text{아파트평수}) + (-0.098 \\ * 3\text{개월온라인업종유실적여부1}), \dots, + (-0.276 * (30\text{세} \sim 34\text{세}, \text{남성})).$$

가족사랑보험은 암과 상해 등 다보장과 특정질환에 대한 보상금액이 크고 저 연령대일수록 보험가격이 싸기 때문에 저 연령대 구간더미변수인 25~26세 남녀 집단에서 변수 가중치 값이 양의 부호를 띠고 있으며, 연령대가 높아질수록 음의 부호로 값이 전환되어 그 값이 커지는 경향을 보이고 있다. 또한 지역을 살펴보면 주 센터가 대구에 위치한 경향으로 대구 (0.1909), 경북 (0.3182), 경남 (0.2598) 지역에서 양의 부호를 나타내고 있으며, 반면 서울 (-0.1905)은 음의 계수값을 가진다.

상해보험의 경우 차량종류에서 생계형 (1.1871)이 선택되었으며, 아파트 비거주변수와 온라인고객은 음의 계수값을 나타냈다, 이는 소규모 자영업자를 하는 사람들이 필요에 의해 보험가입을 선택하는 것으로 추측할 수 있으며 변수의 선택이 타당한 것으로 보인다.

표 3.1 상품별 로지스틱회귀모형 구축결과

가족사랑보험	상해보험	암보험
zz = -3.1433	zz = -4.1569	zz = -3.3228
0.000007246*6개월평균통신 매출금액	0.0000002265*3개월평균 신용카드매출금액	-0.0413*아파트평수
-0.7906*6개월평균통신 매출회수	0.000001287*6개월평균 살거리업종매출금액	0.654*전남
-0.6924*(31세~34세,남자)	0.000003533*6개월평균 통신매출금액	0.0298*6개월이용주유소개수
0.3479*(25세~26세,여성)	0.000004086*6개월평균 생활/건강업종매출금액	0.5355*(40~41세,남여)
-0.414*(27세~30세,남자)	1.1871*생계형	-0.3076*차량보유여부1
0.1962*EBC제휴1신용정상카드 보유여부1	-0.2394*온라인고객여부1	0.7014*아파트거주여부1
0.3182*경북	-0.8423*아파트보유여부1	0.5271*신규
-0.00429*Offline사용개월수 대소분류_ID	0.3984*ebc카드보유여부1	-0.3372*단기휴면
0.2059*(25세~26세,남성)	-0.3124*3개월 통신유실적여부1	0.00417*주유사용개월 수대소분류_ID
-0.1905*서울	0.3981*6개월온라인업종 유실적여부1	-0.2679*6개월평균통신매출회수
0.2598*경남	0.5686*0(비거주)	0.2347*6개월평균신용카드 매출회수
-3.907083*(18≤아파트평수<30.96)	-0.2905*(0<가용포인트 <17,755)	0.3886*고정
-0.3647*(30.96≤아파트평수<40)	0.5385*(35~37세,남)	0.3855*단골
-0.1337*(935.1≤가용포인트<5194.9)	0.2864*(38~39세,남)	-0.00805*OCBcom사용개월수 대소분류_ID
-0.1274*아파트거주여부1	-0.00003*3개월평균마트매출금액	0.651*제주
0.4192*제주		-0.1032*일반
0.00344*신용사용개월수 대소분류_ID		-1.9804*6개월평균놀거리/ 설거리업종매출회수
0.8157*공무원		-0.1564*우수
0.1909*대구		-0.00507*캐쉬백전체사용개월수 대소분류_ID
0.00377*OCBcom사용개월수대소분류_ID		0.1375*(우수,cb)
-0.0937*6개월먹거리업종유실적여부1		0.2037*(일반,cb)
0.219*전남		-0.1943*온라인고객여부1
-0.218*휴면		
0.1872*강원		
-0.1063*(가용포인트<935.1)		

마지막으로 암보험은 주 판매 연령대상이 40~54세이지만 40대 초반인 40~41세 구간에서 변수가 선택되고 양의 계수값을 가지는 것으로 나타났으며, S주식회사의 고객등급이 고정 (0.3886), 단골 (0.3855)인 고객층이 변수로 선택되어 로알티가 어느 정도 형성된 집단에서 보험가입이 이루어질 것으로 추측된다.

이후 선택된 변수를 가지고 약100가지의 모델 시험을 거쳐 최종적으로 로지스틱회귀모형을 구축하게 되는데 분석용 자료에서 시행한 결과에서 선택된 독립변수들과 추정된 회귀계수를 예측용 자료에 \hat{P} 값을 기준으로 내림차순으로 정렬하고, \hat{P} 값이 큰 순서대로 상위 10%씩 구간을 나누어 구간별 구매고객의 비율을 비교하고 추정된 회귀모형의 성과를 평가한다. 표 3.2는 가족사랑보험, 표 3.3은 상해보험, 표 3.4는 암보험에 대한 분석용 데이터와 예측용 데이터의 구매율을 산출한 것이다. 가족사랑보험의 경

우 분석용 데이터에서 상위 10%의 구간에서의 구매율은 6.99%를 나타내고 있으며, 이 때 전체 구매자 중 21.1%를 차지하는 것으로 해석한다.

표 3.2 가족사랑보험의 분석용데이터와 예측용데이터 적합

Percentile 분석용	1	2	3	4	5	6	7	8	9	10
가입고객수	1627	1138	961	875	757	664	546	508	380	255
누적가입수	1627	2765	3726	4601	5358	6022	6568	7076	7456	7711
전체고객수	23281	23281	23281	23281	23281	23281	23281	23281	23281	23281
누적가입수	23281	46562	69843	93124	116405	139686	162967	186248	209529	232810
부분가입율	6.99%	4.89%	4.13%	3.76%	3.25%	2.85%	2.35%	2.18%	1.63%	1.10%
가입누적%	21.10%	35.86%	48.32%	59.67%	69.49%	78.10%	85.18%	91.77%	96.69%	100.0%

Percentile 예측용	1	2	3	4	5	6	7	8	9	10
가입고객수	158	132	102	89	64	63	63	55	53	34
누적가입수	158	290	392	481	545	608	671	726	779	813
전체고객수	2655	2656	2656	2655	2656	2656	2655	2656	2656	2656
누적가입수	2655	5311	7967	10622	13278	15934	18589	21245	23901	26557
부분가입율	5.95%	4.97%	3.84%	3.35%	2.41%	2.37%	2.37%	2.07%	2.00%	1.28%
가입누적%	19.43%	35.67%	48.22%	59.16%	67.04%	74.78%	82.53%	89.30%	95.82%	100.0%

표 3.3 상해보험의 분석용데이터와 예측용데이터 적합

Percentile 분석용	1	2	3	4	5	6	7	8	9	10
가입고객수	84	42	35	33	28	31	19	16	15	14
누적가입수	84	126	161	194	222	253	272	288	303	317
전체고객수	2,836	2,836	2,836	2,836	2,836	2,836	2,836	2,836	2,836	2,836
누적가입수	2,836	5,672	8,508	11,344	14,180	17,016	19,852	22,688	25,524	28,360
부분가입율	2.96%	1.48%	1.23%	1.16%	0.99%	1.09%	0.67%	0.56%	0.53%	0.49%
가입누적%	26.50%	39.75%	50.79%	61.20%	70.03%	79.81%	85.80%	90.85%	95.58%	100.0%

Percentile 예측용	1	2	3	4	5	6	7	8	9	10
가입고객수	54	39	29	27	31	23	19	15	9	14
누적가입수	54	93	122	149	180	203	222	237	246	260
전체고객수	1,890	1,891	1,890	1,891	1,891	1,890	1,891	1,890	1,891	1,891
누적가입수	1,890	3,781	5,671	7,562	9,453	11,343	13,234	15,124	17,015	18,906
부분가입율	2.86%	2.06%	1.53%	1.43%	1.64%	1.22%	1.00%	0.79%	0.48%	0.74%
가입누적%	20.77%	35.77%	46.92%	57.31%	69.23%	78.08%	85.38%	91.15%	94.62%	100.0%

추정된 로지스틱회귀모형에 따르면 가족사랑보험, 상해보험, 암보험 모두 상위 50%에서 구매고객의 누적비율이 전체 67~70% 사이를 차지하고 있으며, 분석용 자료와 예측용 자료에서의 성과차이는 거의 나타나지 않음을 알 수 있다, 그러므로 모형의 적합도 결과 비교적 모형의 안정성이 있는 것으로 판단된다.

4. 실증분석

L생명보험사 텔레마케팅 제휴사인 S주식회사는 포인트를 통한 가맹점유치를 통해 전국적으로 이만여 개의 제휴상권을 유지하고 있다. 일면식이 없는 고객과의 대화에서 거부감을 최소화하고 친근감을 유지하기 위해 포인트 적립금액과 활용방안에 대하여 알려주며, 향후 본격적으로 표준화된 스크립트를 통해 보험계약유치를 위한 영업을 진행하게 된다.

표 3.4 암보험의 분석용데이터와 예측용데이터 적합

Percentile 분석용	1	2	3	4	5	6	7	8	9	10
가입고객수	1882	1296	1072	1006	884	756	704	563	482	333
누적가입수	1882	3178	4250	5256	6140	6896	7600	8163	8645	8978
전체고객수	28506	28506	28507	28506	28507	28506	28506	28507	28506	28507
누적가입수	28506	57012	85519	114025	142532	171038	199544	228051	256557	285064
부분가입율	6.60%	4.55%	3.76%	3.53%	3.10%	2.65%	2.47%	1.97%	1.69%	1.17%
가입누적%	20.96%	35.40%	47.34%	58.54%	68.39%	76.81%	84.65%	90.92%	96.29%	100.0%

Percentile 예측용	1	2	3	4	5	6	7	8	9	10
가입고객수	247	192	132	129	115	101	89	65	64	59
누적가입수	247	439	571	700	815	916	1005	1070	1134	1193
전체고객수	3734	3734	3735	3734	3735	3734	3734	3735	3734	3735
누적가입수	3734	7468	11203	14937	18672	22406	26140	29875	33609	37344
부분가입율	6.61%	5.14%	3.53%	3.45%	3.08%	2.70%	2.38%	1.74%	1.71%	1.58%
가입누적%	20.70%	36.80%	47.86%	58.68%	68.32%	76.78%	84.24%	89.69%	95.05%	100.0%

본 논문에서는 로지스틱회귀모형을 통하여 구축한 모델과 상담원들이 선호하는 영업방식인 가용포인트가 우수한 고객을 뽑아 두 집단을 상호 비교하였으며, 가족사랑보험, 암보험은 2008년도 12개월을 비교하였다. 상해보험은 2008년 1월부터 6월까지이다.

이 때 반응은 보험가입자수와 가입율을 기준으로 적용하였다. 매월 추출하는 모수 및 추출비율은 달랐으나 각 그룹별 효과를 통하여 모델의 효과성을 입증하고자 하였다.

데이터 추출은 가용포인트를 우선으로 고려하는 집단에서 해당월의 총 모수 중 10%를 추출하고 로지스틱회귀모형에서 90%를 추출하여 평가가 이루어졌으며, 가용포인트를 우선으로 고려하는 집단에서 10% 추출한 이유는 비교대상 점검과 모델의 노후 여부를 판단하기 위해서다.

표 4.1 가족사랑보험 경우 2008년 1월 캠페인을 위한 데이터추출은 2007년 12월12일이며, 총 대상 고객은 160135명이다. 캠페인의 공정성을 기하기 위하여 가용포인트를 우선으로 고려하는 집단 10%와 모델집단 90%를 임의추출방식으로 분배하였다. 모델에서 추출한 인원은 72000명이며, 가용포인트를 우선으로 고려하는 집단에서 추출한 인원은 8000명이다. 두 집단에서 상위 50%씩을 뽑아 캠페인 대상으로 추출하고 적용하였다. 적용결과 모델은 305명, 가용포인트를 우선으로 고려하는 집단은 27명으로 모델이 약26%의 개선효과를 보이는 것으로 나타났다. 또한, 2008년 12개월 동안의 캠페인 결과 가족사랑보험의 경우 66.9%의 효율 개선효과를 보이고 있었으며, 암보험의 경우는 45.3%, 그리고 상해보험의 경우 6개월 동안 평균수율이 49.4% 증가한 것으로 나타났다.

이로써 모델 평가에 대한 효과는 입증되었으나, 사업자가 설정한 수율 변동 허용범위를 벗어난 경우와 지속적으로 예상수율을 유지하기 위하여 시간 경과에 따라 변화된 고객 속성을 새롭게 반영한 갱신모델을 수행할 필요가 존재한다.

갱신모델을 수행하는 경우는 통상적으로 모델 구축 후 6개월 이내에 모델 갱신을 주기적으로 진행하기도 하지만 기준시점 대비 연속 2개월 또는 기준시점대비 급격한 수율 하락이 발생한 경우, 수율 평균선이 하향 추세로 판단될 경우 모델을 재검토 한 후 모델갱신을 수행할 지 여부를 최종 결정하게 된다.

이 때 기준시점은 모델을 설정할 때 예측용 데이터를 적합한 예측치를 기준으로 삼아 실제가입자수를 비교하는 방식으로 진행하며 식은 다음과 같다.

$$\frac{\text{실제가입자수}}{\text{예측가입자수}} > 1.$$

표 4.1 가족사랑보험 월별 캠페인 결과

추출일자	총대상고객	Random sample	추출고객	상위%	가입자	비율	평가	
2007-12-12	160135	모델	144122	72000	50.0%	305	0.42%	26%
		포인트	16013	8000	50.0%	27	0.34%	
2008-01-17	179802	모델	161822	64728	40.0%	217	0.34%	5%
		포인트	17980	7192	40.0%	23	0.32%	
2008-02-12	355672	모델	320105	126000	39.4%	923	0.73%	63%
		포인트	35567	14000	39.4%	63	0.45%	
2008-03-13	163234	모델	146911	58763	40.0%	265	0.45%	9%
		포인트	16323	6529	40.0%	27	0.41%	
2008-04-10	194326	모델	174893	58500	33.4%	285	0.49%	58%
		포인트	19433	6500	33.4%	20	0.31%	
2008-05-16	141887	모델	127698	76619	60.0%	256	0.33%	29%
		포인트	14189	8513	60.0%	22	0.26%	
2008-06-13	166540	모델	149886	104920	70.0%	316	0.30%	21%
		포인트	16654	11658	70.0%	29	0.25%	
2008-07-16	140263	모델	84158	50494	60.0%	82	0.16%	27%
		포인트	56105	33663	60.0%	43	0.13%	
2008-08-14	77855	모델	70070	28028	40.0%	151	0.54%	180%
		포인트	7785	3114	40.0%	6	0.19%	
2008-09-18	83606	모델	50164	25082	50.0%	104	0.41%	36%
		포인트	33442	16721	50.0%	51	0.31%	
2008-10-15	311779	모델	187067	56120	30.0%	311	0.55%	83%
		포인트	124712	37413	30.0%	113	0.30%	
2008-11-16	373433	모델	224060	45412	20.3%	215	0.47%	111%
		포인트	149373	30274	20.3%	68	0.22%	

표 4.2 상해보험 월별 캠페인 결과

추출일자	총대상고객	Random sample	추출고객	상위%	가입자	비율	평가	
2007-12-12	104250	모델	94357	18596	19.7%	52	0.28%	36%
		포인트	9893	1945	19.7%	4	0.21%	
2008-01-17	109338	모델	98404	19680	20.0%	45	0.23%	67%
		포인트	10934	2187	20.0%	3	0.14%	
2008-02-12	110844	모델	99760	22500	22.6%	32	0.14%	78%
		포인트	11084	2500	22.6%	2	0.08%	
2008-03-13	71468	모델	64321	12864	20.0%	36	0.28%	33%
		포인트	7147	1429	20.0%	3	0.21%	
2008-04-10	113630	모델	102267	18000	17.6%	51	0.28%	42%
		포인트	11363	2000	17.6%	4	0.20%	
2008-05-16	73406	모델	66065	13213	20.0%	15	0.11%	67%
		포인트	7341	1468	20.0%	1	0.07%	

가족사랑보험에 대한 모델 구축후 캠페인을 진행한 결과 그림 4.2와 같이 6개월까지는 일정수율의 꺾 상승이 일어난 후 7개월차에 외부변인 (청약기피), 또는 특정변인 (DB주기)으로 인하여 예측가입자수 보다 실제가입자수가 급격히 하락한 결과가 나타났다. 그러므로 모델갱신을 실시하였으며 이후 안정적으로 모델이 운영되는 결과가 도출되었다. 갱신 전 모델은 외부요인을 제거하더라도 선형추세에 의한 그래프에 의해서도 하방경직성을 보이고 있으며, 지속적인 하락을 나타내고 있다. 반면 8개월차부터 적용한 갱신모델은 지속적인 꺾 상승 결과를 보이고 있다.

이 때 모델갱신 전·후 모델변수는 표 4.4와 같다. 모델갱신 전·후 모델에서 연령의 구간화가 조정되었으며, 모델갱신 전에서는 아파트 거주여부와 아파트거래금액의 구간화에 따른 변수 가중치가 많이 반영되었으나 모델갱신 후에는 6개월평균신용카드매출금액, 회원등급 변수구간별 가중치가 많이 반영되었

표 4.3 압보험 월별 캠페인 결과

추출일자	총대상고객	Random sample	추출고객	상위%	가입자	비율	평가
2007-12-12	278662	모델	250796	72000	28.7%	264	0.37%
		포인트	27866	8000	28.7%	17	0.21%
2008-01-17	313722	모델	282350	84704	30.0%	357	0.42%
		포인트	31372	9412	30.0%	23	0.24%
2008-02-12	284124	모델	255712	36000	14.1%	147	0.41%
		포인트	28412	4000	14.1%	10	0.25%
2008-03-13	310959	모델	279863	83958	30.0%	193	0.23%
		포인트	31096	9329	30.0%	17	0.18%
2008-04-10	383306	모델	344975	72000	20.9%	258	0.36%
		포인트	38331	8000	20.9%	8	0.10%
2008-05-16	337371	모델	303634	121453	40.0%	384	0.32%
		포인트	33737	13495	40.0%	35	0.26%
2008-06-13	268841	모델	241957	67500	27.9%	372	0.55%
		포인트	26884	7500	27.9%	31	0.41%
2008-07-16	281277	모델	253149	72000	28.4%	480	0.67%
		포인트	28128	8000	28.4%	50	0.63%
2008-08-14	218138	모델	196324	98162	50.0%	783	0.80%
		포인트	21814	10907	50.0%	75	0.69%
2008-09-18	209711	모델	188740	99000	52.5%	850	0.86%
		포인트	20971	11000	52.5%	59	0.54%
2008-10-15	238458	모델	214612	85845	40.0%	895	1.04%
		포인트	23846	9538	40.0%	40	0.42%
2008-11-16	309743	모델	278769	129883	46.6%	1332	1.03%
		포인트	30974	14431	46.6%	118	0.82%

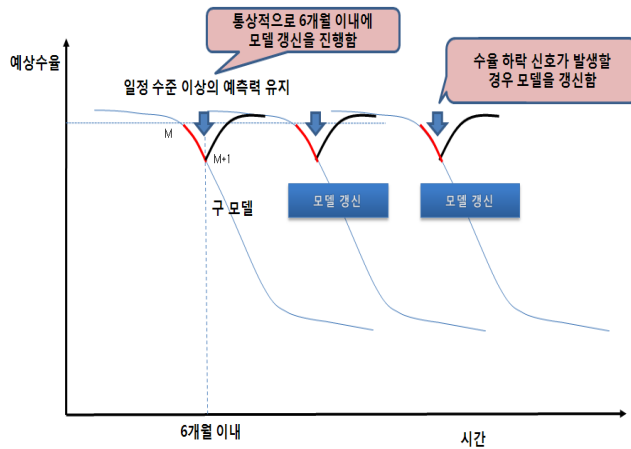


그림 4.1 모델갱신주기

다. 이는 시간이 흐름에 따라서 S주식회사 고객 중 보험가입자의 행동에 변화가 있었다는 것을 의미하며, 100여개의 축약된 변수 중 로지스틱회귀모형에 반영되는 변수와 그 가중치가 변화된 결과를 얻었다. 즉, 과거 모델을 지속적으로 사용하였을 경우 DB주기에 따른 고객들의 피로도와 서브프라임모기지 등 외부변인에 의한 고객성향등을 반영하기 어렵다. 그러므로 최근 고객의 행동성향을 주기적으로 반영하는 것이 지속적인 보험수율에 기여하는 것으로 사료된다.

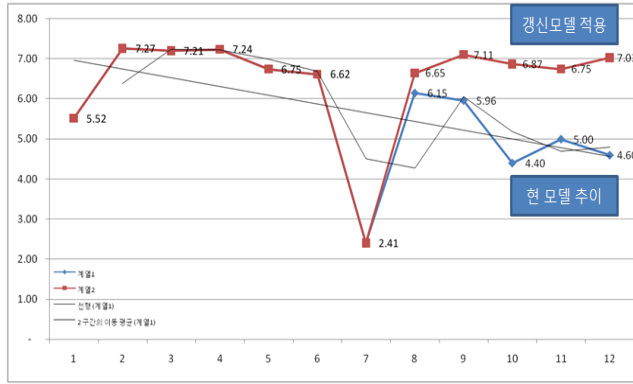


그림 4.2 모델갱신 전·후 예측가입수와 실제가입자수 비교

표 4.4 가족사랑보험 모델갱신 전·후 모델변수

가족사랑보험	가족사랑보험 갱신모델
zz=-3.1433	zz=-4.2754
0.000007246*6개월평균통신매출금액	(0.000007046*가용포인트)
-0.7906*6개월평균통신매출회수	(-0.00764*아파트평수)
-0.6924*(31세~34세, 남자)	(-0.00000261*누적발생포인트)
0.3479*(25세~26세, 여성)	(0.000005557*3개월평균통신매출금액)
-0.414*(27세~30세, 남자)	(-0.000000696*3개월평균신용카드매출금액)
0.1962*EBC제휴신용정상카드보유여부1	(0.5669*3개월평균신용카드매출회수)
0.3182*경북	(0.000001365*6개월평균통신매출금액)
-0.00429*Offline사용개월수대소분류_ID	(0.000001129*6개월평균신용카드매출금액)
0.2059*(25세~26세, 남성)	(-0.00000746*6개월평균거리업종매출금액)
-0.1905*서울	(-0.7174*6개월평균통신매출회수)
0.2598*경남	(-0.8673*6개월평균신용카드매출회수)
-3.907083*(18<=아파트평수<30.96)	(0.00251*캐쉬백전체사용개월수대소분류_ID)
-0.3647*(30.96<=아파트평수<40)	(0.00378*신용사용개월수대소분류_ID)
-0.1337*(935.1<=가용포인트<5194.9)	(-0.00418*Offline사용개월수대소분류_ID)
-0.1274*아파트거주여부1	(0.00325*OCBcom사용개월수대소분류_ID)
0.4192*제주	(-0.2872*서울)
0.00344*신용사용개월수대소분류_ID	(-0.2103*부산)
0.8157*공무원	(-0.2252*대전)
0.1909*대구	(0.1045*경북)
0.00377*OCBcom사용개월수대소분류_ID	(0.2041*제주)
-0.0937*6개월먹거리업종유실적여부1	(0.3424*공무원)
0.219*전남	(0.1372*학생)
-0.218*휴먼	(0.4693*M월회원등급(신규))
0.1872*강원	(0.1025*M월회원등급(단골))
-0.1063*(가용포인트<935.1)	(0.1308*M월회원등급(일반))
	(0.2834*EBC제휴신용정상카드보유여부1)
	(0.1137*리더스정상카드보유여부1)
	(0.1142*3개월살거리업종유실적여부1)
	(-0.098*3개월온라인업종유실적여부1)
	(0.1164*(7118<아파트 실거래가<=13563))
	(0.5513*(25세~27세, 남성))
	(0.5034*(25세~27세, 여성))
	(0.3515*(28세~29세, 여성))
	(-0.276*(30세~34세, 남성))

5. 결론

본 연구는 생명보험사 텔레마케팅 아웃바운드에서 로지스틱회귀분석을 통한 모형을 설정하고 이후 매월 캠페인 수행결과를 토대로 모형에 대한 평가를 진행하였다. 캠페인 진행결과 가족사랑보험, 암보험, 상해보험 3가지 모두 직접적인 보험가입자 증가로 매출의 증대에도 기여하는 것으로 나타났다. 또한 같은 회사의 데이터를 사용하더라도 연령대, 보험종류에 따라 모형에서 취하는 변수는 다르게 나타나고 있으며, 모형을 구축하고 일정한 시간이 흐르면 모델의 노후화가 일어나는 것으로 확인되었다. 그러므로 모델을 주기적으로 갱신할 필요성이 있으며, 가족사랑보험의 경우 약7개월 주기가 적당할 것으로 판단되었다. 하지만 모델이 성과가 노후화 되는 시점을 바로 파악하기에는 최소 2개 이상의 업체가 제휴관계에서 업무를 진행하기 때문에 현실적으로 쉽지 않으며, 이를 위하여 실시간으로 모델효과를 점검할 수 있는 시스템이 갖춰져야 될 것으로 판단된다. 또한 암보험, 상해보험에서도 같은 결과를 나타내는지 확인할 필요성이 존재한다. 텔레마케팅의 경우 반응률 0.1%가 거수보험료 및 십억의 매출규모를 좌우하기 때문에 고객의 속성을 잘 판별하여 반영하는 것이 중요하다고 볼 수 있다.

참고문헌

- 김미정, 김진형 (2008). 로지스틱회귀모형을 이용한 유족연금 수급 분석 연구. <응용통계연구>, 2, 183-200.
- 성용현 (2001). <응용 로지스틱 회귀분석>, 탐진, 서울.
- 윤미례 (2004). <데이터 마이닝을 이용한 부도예측모델>, 석사학위논문, 이화여자대학교, 서울.
- 이성우, 민성희, 박지영, 윤성도 (2005). <로짓·프라빗모형 응용>, 박영사, 서울.
- 이현정 (2001). <데이터마이닝을 이용한 보험회사 고객이탈분석에 관한 연구>, 석사학위논문, 중앙대학교, 서울.
- 이호영 (2005). <생명보험회사 고객관계관리를 위한 고객이탈예측모형에 관한 연구>, 박사학위논문, 한국외국어대학교, 서울.
- 정광모, 최용석 (2003). <로지스틱 회귀와 응용>, 자유아카데미, 서울.
- Cho, Y. J. and Lee, Y. G. (2004). A study on the optimization method of the initial weights in single layer perceptron. *Journal of the Korean Data & Information Science Society*, 2, 331-337.
- Kim, Y. H. and Lee, S. W. (2008). An empirical study on telemarketing business. *Journal of the Korean Data & Information Science Society*, 3, 877-891.

An empirical study on telemarketing efficiency at life insurance

Bong-Sung Koh¹ · Seok-Won Lee² · Jeong-Heo³

¹²³ Department of Public Survey and applied Statistics, Jeonju University

Received 1 May 2009, revised 12 July 2009, accepted 16 July 2009

Abstract

Lower Prices are offered through sales by telemarketing. This is to serve our customers by the fastest and most appropriate referral product that is most important to attract insurance. Therefore, Considering the time the customer's preferred products and preferred customer for screening and targeting, depending on what is the difference between the premiums. This study of the logistic regression model using datamining techniques, the life insurance companies in outbound telemarketing to support sales of the effect you want to validate. To join existing life insurance companies for the customer response and sales strategy based on the L segment and by age group, family-love insurance, accident insurance, and cancer insurance were in progress for the modeling. Set model based on the progress of the campaign to existing customers marketing methods and how to extract and run the model results has proven the superiority of the model. In addition, over time, depending on the aging model is set to a decline in operating profit to maximize the profits th update the model which was derived.

Keywords: Insurance, logistic regression model, outbound, response ratio, telemarketing.

¹ Corresponding author: Professor, Department of Public Survey and applied Statistics, Jeonju University, Jeonju 560-759, Korea. E-mail: bskoh@jj.ac.kr

² Adjunct Professor, Department of Public Survey and applied Statistics, Jeonju University, Jeonju 560-759, Korea.

³ Master of course, Department of Public Survey and applied Statistics, Jeonju University, Jeonju 560-759, Korea.