

Top 뮤직 10 정액제 상품 타겟팅 개선을 위한 결합모델 개발

전희주* · †이재영**

Combining Model Development for Targeting Top Music 10
Additional Service Product of A Mobile Telephone Company

Heuiju Chun* · †Jae-Yeong Lee**

■ Abstract ■

Top music 10 is a additional service product of the A mobile telephone company. Up to now, A company is just selling it by outbound TM to customers which visit any contents of Top Music 10. In this paper, we proposed a targeting method combining two score models by data mining. The proposed combining model is to find customers more likely to respond to outbound TM. The proposed targeting method is expected to improve both from 32.8% to 44.0% in the response rate and from 54.7% to 61.4% in the retention rate.

Keywords : Top Music 10, Data Mining, Target Marketing, Scoring, Combining Model

1. 서 론

A 이동통신회사는 2006년 5월 부가서비스 상품인 Top 뮤직 10 정액제를 출시하였다. 본 상품의

목적은 정보료 정액제 상품 개발로 뮤직 콘텐츠에 대한 고객의 정기적 이용 패턴 형성을 통해 이용량 증대 및 안정적 매출 기반 확보를 위해 개별 서비스 단위로 산재해 있는 뮤직 관련 정액제를 통합하

논문접수일 : 2008년 04월 15일 논문수정일 : 2008년 05월 27일 논문게재확정일 : 2008년 06월 03일

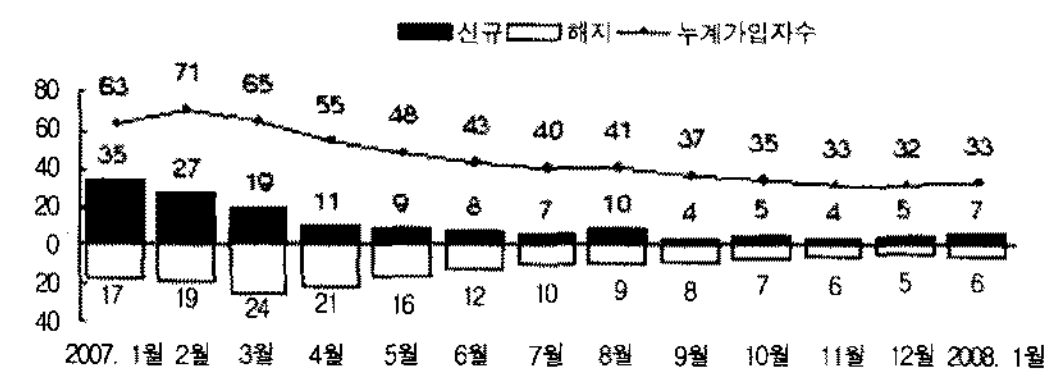
* SK텔레콤 매니저

** 국방대학교 관리대학원 운영분석학과 교수

† 교신저자

여 실질적인 고객가치(Customer Value, CV)를 증대하기 위함이었다. Top 뮤직 정액제는 NATE의 주요 뮤직 서비스를 곡 단위로 통합한 통합 뮤직 유료 콘텐츠인 벨, 링, MOD, 노래방, 뮤직 비디오, 음악 방송에 대하여 월 정액 요금 4000원을 납부하고, 통합 뮤직 곡 페이지 내 유료 콘텐츠 서비스의 종류에 상관없이 총 10건까지 추가 정보 이용료를 납부하지 않고 사용할 수 있는 할인 정액제 서비스이다(단, 해당 월 10건 초과분에 대해서는 종량 과금하며, 통화료는 별도임).

2008년 1월 시점 누계 가입 고객은 약 33만 명으로 2007년 2월 이후 지속 감소 추세에 있다. <그림 1>은 2007년 1월~2008년 1월 동안 Top 뮤직 10 정액제 가입 고객 수, 해지 고객 수와 누계 가입 고객 수를 나타낸다. 신규 가입 고객 수의 감소와 함께 가입고객 대비 해지 고객 수가 많아 누적 가입 고객이 감소 현상을 보이고 있다.



가입 고객 수(단위 : 만명)

<그림 1> Top 뮤직 10 정액제

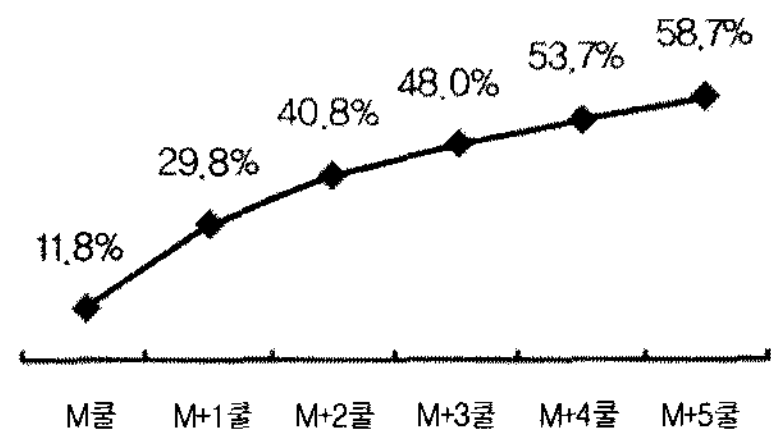
Top 뮤직 10 정액제 가입은 자발적 채널이 14.8%, 유통망이 64.5%, 약 20.4%의 고객이 Outbound TM을 통해 가입하였다. Outbound TM을 통해 가입한 고객의 잔존누계 가입 고객은 2008년 1월 현재 10만 명으로 전체 가입 고객의 31.4%를 차지하고 있다. Top 뮤직 10 부가서비스에 대한 월 평균 TM은 전월 Top 뮤직 10 관련 콘텐츠를 5번 이상 방문하고 전월 정보 이용료를 2000원 이상 납부하며 데이터 통화료 정액제 가입 고객에 한하여 진행하고 있으며 TM 대상 고객은 10만 명 정도로 이루어지고 있다. <표 1>에 나타난 바와 같이 2007년 전체 TM 대상 고객의 Top 뮤직 10 정액제 월 평

균 접촉 대비 성공률은 34.1%로 다른 데이터 오퍼링 상품의 성공률과는 큰 차이는 나지 않았다.

<표 1> 월 평균 Outbound TM 현황

구 분	고객 수	비 중
월 평균 TM 대상	103,597	-
월 평균 진행	98,759	95.3%
월 평균 접촉	49,425	50.0%
월 평균 가입(성공)	16,840	34.1%

TM에 의한 Top 뮤직 10 정액제 가입 고객은 고객의 니즈에 대한 고려 없이 판매되어 가입 후 4개월 차에 48%, 6개월 이후에 58.7%로 정액제 해지율이 증가함으로써 고객의 불만 확대가 우려되었다. <그림 2>는 TM 가입 고객의 해지율 변화를 보여주고 있다.



<그림 2> TM 가입 고객의 해지율

TM 대상 고객에 대한 A사의 6개월 이후 해지율은 정액제 가입 고객이 9.3%, 미가입 고객이 13.5%로 Top 뮤직 10 정액제 가입으로 인한 A사의 Retention 효과는 있는 것으로 판단된다.

본 논문에서는 Top 뮤직 콘텐츠 이용과 전월 정보료 2000원을 납부한 고객을 선정하는 기존의 타겟팅 조건보다 가입을 향상과 가입 후 유지율 증대를 동시에 향상시킬수 있는 데이터마이닝 기법의 2개 모델링을 결합한 스코어링에 의한 타겟팅 방법을 제안하였다. 현재 고객의 특성과 과거 행태 자료를 기초로 데이터마이닝 모델링을 통한 점수화 방법은 신용 관리나 마케팅의 고객 타겟팅 방법으로 많이 이용되고 있다. 진서훈, 김현기[5]는 신

용카드 이용규모 추정을 위해 모델링을 통한 데이터마이닝 기법을 적용하였으며, 한상태 등[6]은 보험회사의 교차 판매 고객의 특성을 파악하기 위해 스코어링 모형을 이용하였고, 김용수[3]는 이동통신 시장에서 데이터 요금제의 해지율을 구하기 위해 예측모형을 사용하였다. 최근에 전희주, 정병철[4], 김정무, 전희주[2]는 데이터마이닝의 모델링에 의한 스코어링 기법을 활용한 효율적인 타겟팅 방법을 제안하였다.

제 2장에서는 타겟 모델의 정의와 모델을 개발하는데 사용되는 변수들에 대한 기초 통계량을 구하고, 제 3장에서는 Top 뮤직 10 가입 및 유지 예측 모델링과 선택된 모델에 대한 검증, 향후 활용 방안을 설명하였으며, 제 4장에서는 결론 및 제언을 기술하였다.

2. 타겟 모델 정의 및 기초 통계 분석

2.1 모델링 개요

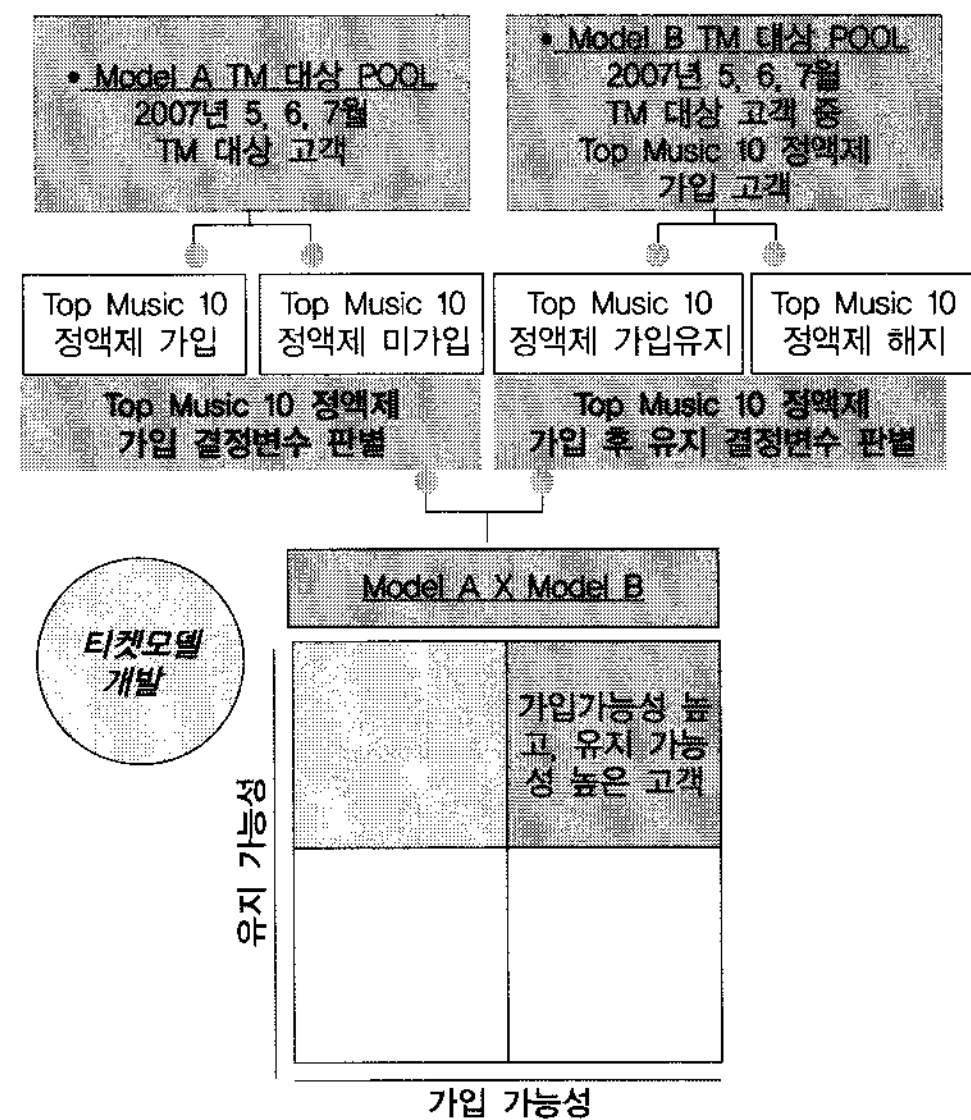
본 연구에서는 Top 뮤직 콘텐츠를 Outbound TM으로 권유할 때 이전에 콘텐츠를 이용해 본 고객을 선택하는 기존의 타겟팅 조건보다 가입율을 높이고 가입 후 유지율 증대를 동시에 향상시키는 타겟팅 방법을 개선하는데 목적이 있다. 즉 Outbound TM 고객에 대해 가입율과 유지율의 교집합 $P(\text{가입} \cap \text{유지})$ 을 최대한 높이하고자 하는 데 있다. $P(\text{가입} \cap \text{유지})$ 은 식 (1)과 같이 표현될 수 있다.

$$P(\text{가입} \cap \text{유지}) = P(\text{유지} / \text{가입}) \cdot P(\text{가입}) \quad (1)$$

$P(\text{가입} \cap \text{유지})$ 을 구하는 모델은 $P(\text{가입} \cap \text{유지})$ 과 $P(\text{가입})$ 을 각각 구하여 모델화하는 방법과 $P(\text{가입} \cap \text{유지})$ 을 바로 모델화하는 방법이 있다. 본 논문에서는 좀 더 직관적이고 현업에서 쉽게 설명할 수 있었으며 적용할 수 있는 모델을 우선 고려하고자 한다. 그래서 예상 가입율과 가입 후 예상 유지율

을 각각 구한 후 일정한 예상 가입율 이상이고 가입 후 예상 유지율이 일정 수준 이상이 되는 모델링 방법을 선택하였다. 본 논문에서 제안된 모델은 예상 가입 확률을 구하기 위한 모델링(모델 A)과 가입 후 예상 유지율을 구하기 위한 모델링(모델 B) 두 개의 파트로 구성된다.

[Top Music 10 정액제 Modeling 개요]

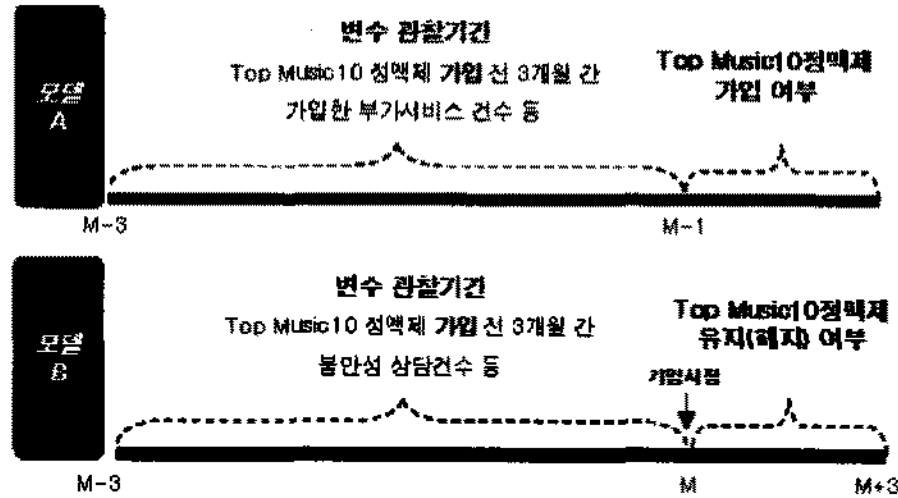


<그림 3> 타겟 모델 정의

<그림 3>과 함께 모델링 개요를 자세히 설명하면 다음과 같다. 우선 2007년 5~7월 자료를 가지고 모델 A를 통해 가입 가능성이 높은 고객을 찾는 다음, 모델 B에 의한 정액제 가입 후 유지율 점수가 높게 예상되는 고객을 선택하는 결합 모델인 타겟 모델을 구하는 것이다. 그리고 타겟 모델은 2007년 8월과 9월 Outbound TM 고객 대상을 가지고 선택된 타겟 모델의 유효성을 검증하게 된다.

모델 A를 구하기 위해 사용한 2007년 5~7월 Outbound TM 대상 고객은 A사 신규 가입 후 3개월이 지난 87,553명이다. 그러나 Top 뮤직 10 정액제 관련 콘텐츠를 TM 전 3개월 동안 사용실적이 없는 미이용 고객(Gray Zone) 7,238명은 이용고객과 미이용 고객의 성향을 분명히 하기 위해 대상에

서 제외하였다. 따라서 남은 80,215명을 가지고 모델링을 구하기 위한 훈련 데이터 세트(Training data set)를 설정하였다.



<그림 4> 타겟 변수 정의

모델 A의 타겟 변수는 TM으로 가입하게 되면 1, 가입을 하지 않으면 0으로 정의하였다. 모델 B의 타겟 변수는 모델 A의 가입 고객 중 25,160명을 대상으로 가입 후 3개월 동안 Top 뮤직 10 정액제 서비스를 유지하면 1, 해지하면 0으로 정의하였다. 모델 B의 타겟 변수를 명확히 정의하기 위해 가입 후 A사의 통신서비스를 해지한 고객은 Gray Zone

<표 2> 모델 설명 변수 정의

인구특성	◦ 성별, 연령
SKT 통화관련	◦ ARPU, DATA ARPU ◦ 음성요금제 유지개월수 ◦ 데이터 선택요금제 유지개월수 ◦ 가입 데이터 선택요금제 ◦ 서비스 가입기간 ◦ 가입 음성 기본 요금제 등
단말기 관련	◦ 단말기 유지기간 ◦ 단말기 사양
NATE 이용수준	◦ NATE 호수대 ◦ Top Music10 콘텐츠 이용량 등
CV 관련	◦ 단순 상담건수 및 불만 상담건수 ◦ TM거부 이력 등
캠페인 관련	◦ Top Music10 정액제 관련 캠페인 진행 횟수 ◦ 총 캠페인 실행 횟수 등
부가서비스	◦ 신규 가입건수, 해지건수 ◦ Top Music10 유사콘텐츠 가입건수 등 ◦ 전 3개월 기간내 가입한 부가서비스 건수 등

으로 분류하여 훈련 데이터 세트에서 제외하였다. 타겟 변수 정의는 가입 모델과 유지 모델을 위한 설명변수로는 성별, 연령의 인구 특성 변수, ARPU (Average Revenue Per User), 데이터 ARPU, 음성요금제 등의 통화 관련 변수, 단말기 관련, NATE 이용수준, CV 관련 변수, 캠페인 관련, 부가서비스 관련 등 총 100개 이상의 변수들을 고려하였다. <표 2>는 가입과 해지 모델을 위한 설명변수 들을 나타내며 설명변수와 타겟변수에 대한 도식화된 설명은 <그림 4>를 참조하기 바란다.

2.2 TM 대상 고객 속성

<표 3>은 2007년 5~7월 TM을 실행한 고객에 대한 Top 뮤직 10 정액제의 가입율과 해지율에 대한 주요 변수의 특성을 보여준다. 남성과 여성의 가입율은 차이가 없지만 유지율은 남성이 여성에 비해 높게 나타난다. 연령은 20대가 가입율이 높지만 반면에 유지율은 30~40대에 비해 낮은 경향을 보인다. 멤버십 등급은 VIP 고객은 가입율도 매우 낮을 뿐더러 유지율 또한 매우 낮은 속성을 나타낸다. 멤버십 미가입 고객은 유지율이 매우 낮아 일단 가입을 하면 미인지등으로 인해 해지율이 낮게 나타나는 현상을 보이고 있다.

과거 캠페인 성공 건수가 많을수록 가입율이 높게 나타나나 유지율은 오히려 낮은 현상을 보인다. 과거 TM 성공률이 높을수록 가입율은 높으나 유지율은 약간 낮게 나타난다. 전 3월 부가서비스 가입 건수가 없다면 가입율은 매우 낮게 나타나나 유지율은 매우 높은 현상을 보인다. 전 3월 부가서비스 해지건수가 많을수록 유지율은 낮게 나타나 과거 부가서비스 가입/해지 이력이 부가서비스 유지율에 큰 영향을 주고 있음을 보여준다.

3. 가입 및 유지 예측 모델링

3.1 모델링 선택 변수

Top 뮤직 10 정액제 가입과 유지를 위한 모델은

<표 3> TM 대상 고객의 가입율 및 해지율 특성

변 수		고객 수	고객 비중	가입율	유지율
성별	남	42,414	48.4%	33.0%	57.1%
	여	45,139	51.6%	33.1%	52.2%
연령	20대	51,500	58.8%	35.4%	50.0%
	30대	30,370	34.7%	27.0%	62.5%
	40대 이상	5,680	6.5%	28.5%	66.6%
가입 기간	1년 미만	22,772	26.0%	37.0%	56.4%
	2년 미만	14,128	16.1%	32.4%	54.1%
	3년 미만	8,410	9.6%	30.7%	55.4%
	4년 미만	9,528	10.9%	31.5%	53.6%
	5년 미만	9,770	11.2%	31.1%	53.5%
	5년 이상	22,945	26.2%	28.0%	53.5%
멤버십 등급	VIP	764	0.9%	16.9%	37.6%
	골드	16,199	18.5%	29.9%	50.2%
	실버	14,141	16.2%	34.7%	46.2%
	일반	22,691	25.9%	31.2%	46.1%
	미가입	33,758	38.6%	32.8%	65.2%
전월 Nate 이용	Heavy	25,395	22.9%	37.4%	49.1%
	Active	20,093	29.0%	37.3%	54.8%
	Trial	3,986	4.6%	18.4%	58.2%
	Potential	38,079	43.5%	27.2%	58.2%
과거 캠페인 성공건수	1건	61,394	70.1%	27.6%	55.7%
	2건	20,830	23.8%	41.1%	54.8%
	3건	4,689	5.4%	47.7%	48.3%
	4건 이상	640	0.7%	47.3%	38.5%
TM성 공여부	성공	7,930	20.3%	44.5%	55.7%
	미성공	20,125	79.7%	28.9%	54.2%
전 3월 부가서비스 가입건수	0	43,795	50.0%	27.1%	60.7%
	1건	22,956	26.2%	37.5%	55.0%
	2건	9,583	10.9%	36.8%	49.6%
	3건	5,158	5.9%	36.7%	42.7%
	4건 이상	6,061	6.9%	35.8%	38.5%
전 3월 부가서비스 해지건수	0	75,066	85.7%	31.3%	57.5%
	1건	7,793	8.9%	36.5%	42.5%
	2건	2,452	2.8%	36.9%	36.6%
	3건	935	1.1%	37.2%	32.7%
	4건 이상	1,307	1.5%	34.9%	32.0%
전체	87,553	28,055	100%	32.0%	54.6%

SAS Enterprise Miner Ver 4.3을 이용하여 먼저 Regression 모델, Decision Tree 모델을 고려하였다. SAS의 Assessment Node를 이용하여 Lift¹⁾값과 ROC(Receiver Operating Characteristic) 등의 모델 평가(Assessment) 도구를 가지고 판단한 결과 Regression 모델이 Decision Tree 모델에 비해 비교우위(Dominance)를 나타내었기 때문에 이 후 부터는 다양한 Regression 모델들 중에서 가장 적합한 모형을 선택하였다.

최종 선택된 로지스틱 회귀모형에 의한 Top 뮤직 10 정액제 가입에 영향을 미치는 변수로는 A사의 통화 이력 및 NATE 이용 관련 변수가 중요하게 고려되었다. 이는 Top 뮤직 10 이용은 우선 데이터 통화료 정액제 가입 고객을 대상에 한해 TM 대상을 설정하였기에 데이터 통화료 정액제 가입 고객은 NATE 이용과 밀접한 관련이 있고 과거 TM을 받은 경험이 이후에 TM에 의한 가입에 영향을 미침을 알 수 있다.

변수의 중요도 우선 순위를 살펴 보면, 전 3월 데이터 ARPU의 정보료 비중, 판매성 캠페인 성공 건수, 3개월 평균 NATE Hit수 5회 이상, 연령, TOP 뮤직 10 콘텐츠 Hit수 5회 이상 여부, 오피 선호도, 접촉률, 전월 ARPU의 데이터 정액료 비중, 전 3월 데이터 선택요금제 종류, 전월 데이터 정액료, TM 성공여부들의 순으로 나타났다. <표 4>는 Top 뮤직 10 가입 모델링을 통해 선택된 변수를 보여 주며 이를 로지스틱 모형에 대한 수식으로 나타내면 다음과 같다.

$$\log\left(\frac{\text{가입율}}{1-\text{가입율}}\right) = -2.045 + 0.081 * \text{Sex}_1 - 0.027 * \text{Age} + 0.202 * A1 + 0.08 * A2 + 0.08 * A3 + 0.12 * A4 + 0.15 * A5 + 0.06 * A6 + 0.03 * B1 + 0.139 * B2 + 0.16 * B3 + 0.19 * B4 - 0.101 * C - 0.104 * D1 - 0.04 * D2 + 0.027 * D3 + 0.04 * D4 +$$

1) Lift : 해당구간의 반응률/전체 평균 반응률로 모델을 선택하지 않은 경우에 비해 모델을 선택한 효과를 나타냄.

$$0.15 * D5 + 0.16 * D6 + 0.07 * D7 + 0.27 * D8 - 0.4 * D9 + 0.480 * E1 - 0.59 * E2 - 0.226 * E3 + 0.005 * E4 + 0.148 * E5 + 0.093 * E6 + 0.121 * E7 + 0.128 * E8 + 0.56 * E9 - 0.09 * F1 + 0.136 * F2 + 0.075 * F3 - 0.025 * F4 + 0.193 * F5 - 0.11 * F6 - 0.089 * G1 - 0.285 * G2 - 0.115 * G3 + 0.019 * G4 + 0.404 * G5 + 0.00003 * H + 0.0000053 * I - 0.006 * J + 0.14 * K1 + 0.159 * K2 - 0.293 * K3 - 0.116 * L1 - 0.152 * M1 + 0.246 * N1 - 0.062 * O - 0.001 * P + 0.459 * Q1 + 0.153 * Q2 + 0.024 * Q3 - 0.233 * Q4 - 0.222 * Q5 + 0.005 * R1$$

Sex_1 : 성별 남성

Age : 연령

A1 : 가입기간 1년 미만

A2 : 가입기간 1년~2년 미만

A3 : 가입기간 2년~3년 미만

A4 : 가입기간 3년~4년 미만

A5 : 가입기간 4년~5년 미만

A6 : 가입기간 5년 이상~10년 미만

A7 : 가입기간 10년~15년 미만

B1 : 멤버쉽 가입 안함

B2 : 멤버쉽 일반

B3 : 멤버쉽 실버

B4 : 멤버쉽 골드

C : (M-1)월 데이터ARPU 비중

D1 : 전 3개월 평균 데이터ARPU 비중 10% 미만

D2 : 전 3개월 평균 데이터ARPU 비중 10~20%

D3 : 전 3개월 평균 데이터ARPU 비중 20~30%

D4 : 전 3개월 평균 데이터ARPU 비중 30~40%

D5 : 전 3개월 평균 데이터ARPU 비중 40~50%

D6 : 전 3개월 평균 데이터ARPU 비중 50~60%

D7 : 전 3개월 평균 데이터ARPU 비중 60~70%

D8 : 전 3개월 평균 데이터ARPU 비중 70~80%

D9 : 전 3개월 평균 데이터ARPU 비중 90% 이상

E1 : 전 3개월 평균 정보료ARPU 비중 10~20%

E2 : 전 3개월 평균 정보료ARPU 비중 10~20%

E3 : 전 3개월 평균 정보료ARPU 비중 20~30%

E4 : 전 3개월 평균 정보료ARPU 비중 30~40%

E5 : 전 3개월 평균 정보료ARPU 비중 40~50%

E6 : 전 3개월 평균 정보료ARPU 비중 50~60%

E7 : 전 3개월 평균 정보료ARPU 비중 60~70%

E8 : 전 3개월 평균 정보료ARPU 비중 70~80%

E9 : 전 3개월 평균 정보료ARPU 비중 90~100%

F1 : 데이터 선택요금제 데이터세이프

F2 : 데이터 선택요금제 안심정액제

F3 : 데이터 선택요금제 네이트프리

F4 : 데이터 선택요금제 이월

F5 : 데이터 선택요금제 텅 데이터프리

F6 : 데이터 선택요금제 미가입

G1 : 음성요금제 일반

G2 : 음성요금제 패밀리

G3 : 음성요금제 TTL커플

G4 : 음성요금제 TTL

G5 : 음성요금제 텅

H : M-1월 데이터 정액료

I : 전 3개월 평균 데이터 통화료

J : 최종 단말 획득 후 보유 기간

K1 : M-1월 NATE 이용 Heavy

K2 : M-1월 NATE 이용 Active

K3 : M-1월 NATE 이용 Potential

L1 : 전 3개월 평균 NATE Hit수 5회 이상 여부_N

M1 : 전 3개월 평균 TOP 뮤직 10 Hit수 5회 이상_N

N1 : 전 3개월 평균 TOP 뮤직 10 Hit수 5회 이상_N

O : 전 6개월 NATE 이용 Heavy 개월 수

P : M-1월 NATE Hit 수

Q1 : 오피션호도 평가없음

Q2 : 오피션호도 매우 불만

Q3 : 오피션호도 불만

Q4 : 오피션호도 보통

Q5 : 오피션호도 만족

R : 접촉율

S1 : 판매성 캠페인 성공건수 실패

S2 : 판매성 캠페인 성공건수 1회

S3 : 판매성 캠페인 성공건수 2회~3회

T1 : TM성공 여부 - 실패

<표 4> Top 뮤직 10 가입모델링 선택 변수

모델 설명변수	자유도	카이제곱 통계량
전 3월 데이터 ARPU의 정보료 비중 그룹	9	1721.3
판매성 캠페인 성공 건수 그룹	2	1021.2
3개월 평균 NATE Hit수 5회 이상		507.5
연령	1	280.8
TOP 뮤직 10 콘텐츠 Hit수 5회 이상 여부		253.7
오피 선호도 그룹	5	251.3
접촉률		239.6
전월 ARPU의 데이터 정액료 비중	1	238.5
전 3월 데이터 선택요금제 종류	6	219.8
전월 데이터 정액료	1	191.3
TM 성공여부	1	173.8
전월 NATE 이용 Hit 수	1	103.9
성별	1	89.1
전월 멤버십 등급	4	68.1
TOP 뮤직 10 콘텐츠 구매 회수 5회 이상 여부	1	63.3
전월 음성요금제 종류	5	60.5
전 3개월 ARPU 중 DATA ARPU 비중	9	56.9
전 6개월 Heavy 개월 수	1	47.8
전월 NATE Hit 수	1	38.2
최종 단말획득 후 보유기간	1	38.2
A사 서비스 가입 기간	7	26.3
전 3개월 데이터 통화료	1	14.5

<표 5> Top 뮤직 10 유지모델링 선택변수

모델 설명변수	자유도	카이제곱 통계량
전월 멤버십등급	4	579.3
전 3개월 부가서비스 해지 건수	4	278.3
연령	1	174.5
전 3개월 단순 상담 건수	1	101.4
전 3개월 부가서비스 가입 건수	4	56.6
TM 제외 채널 캠페인 성공여부	1	53.5
전월 데이터 선택요금제 가입 여부	2	48.8
전3월 데이터 선택요금제 종류	6	47.3
NATE 3개월간 Heavy 또는 Active 유지 여부	1	46.5
전 3개월 평균 ARPU	1	43.5
평균 음성요금제 유지개월 수	1	38.6
3개월간 데이터 선택요금제 가입유지 여부	1	24.5
전월 음성요금제 종류	5	23.0
전 3개월 평균 Top 뮤직 10 관련 콘텐츠 통화료 과금액	1	20.4
최종 단말 획득 후 보유기간	1	20.0
TM성 캠페인 성공율	1	19.3
전 3개월 유사 부가서비스 가입 여부	1	18.2
성별	1	15.9
전 3개월 평균 데이터 통화료	1	13.9
전 3개월 Trial 개월 수	1	13.7

Top 뮤직 10 정액제에 가입 후에 부가서비스를 유지하는 주요 변수로는 고객가치 관련 변수가 중요하게 선택되었다. 이는 가입 후 유지는 가입 고객이 Top 뮤직 10 정액제를 가입한 후 그에 대한 가치를 느꼈을 때는 유지할 확률이 높지만 그렇지 않을 때는 해지할 소지가 높음을 나타낸 것이다.

<표 5>는 Top 뮤직 10 가입 후 유지 모델링을 통해 선택된 변수를 보여 주며, 주요 변수로 선택된 변수들의 우선순위는 전월 멤버십 등급, 전 3개월 부가서비스 해지 건수, 연령, 전 3개월 단순 상담 건수, 전 3개월 부가서비스 가입 건수, 전 3개월

부가서비스 가입 건수, TM 제외 채널 캠페인 성공 여부 등의 순으로 나타났다. Top 뮤직 10 가입 후 유지 모델을 로지스틱 모형의수식으로 표현하면 다음과 같다.

$$\log\left(\frac{\text{가입후유지율}}{1-\text{가입후유지율}}\right) = -2.045 - 0.06 * \text{Sex}_1 - 0.029 * \text{Age} + 0.537 * \text{A1} - 0.093 * \text{A2} - 0.071 * \text{A3} + 0.046 * \text{A4} - 0.206 * \text{B1} - 0.106 * \text{C1} - 0.117 * \text{C2} - 0.018 * \text{C3} - 0.495 * \text{C4} - 0.104 * \text{C5} + 0.199 * \text{C6} - 0.283 * \text{D1} - 0.022 * \text{D2} + 0.092 * \text{E1} - 0.329 * \text{E2} - 0.151 * \text{E3} - 0.051 * \text{E4} +$$

$$0.364 * E5 + 0.0000028 * F + 0.00001 * G - 0.11 * H + 0.106 * I + 0.104 * J1 + 0.046 * K + 0.0001 * L - 0.115 * M + 0.251 * N1 + 0.058 * N2 + 0.052 * N3 - 0.189 * N4 + 0.024 * O1 + 0.035 * O2 - 0.055 * O3 - 0.102 * O4 + 0.124 * P1 + 0.208 * Q + 0.127 * R1$$

Sex_1 : 성별 남성

Age : 연령

A1 : 멤버쉽 가입 안함

A2 : 멤버쉽 일반

A3 : 멤버쉽 실버

A4 : 멤버쉽 골드

B1 : 전 3개월 데이터 선택요금제 가입 유지 여부-N

C1 : M-3월 데이터 선택요금제 데이터세이프

C2 : M-3월 데이터 선택요금제 안심정액제

C3 : M-3월 데이터 선택요금제 네이트프리

C4 : M-3월 데이터 선택요금제 이월

C5 : M-3월 데이터 선택요금제 텡 데이터프리

C6 : M-3월 데이터 선택요금제 미가입

D1 : M-1월 데이터 선택요금제 가입 여부-해지

D2 : M-1월 데이터 선택요금제 가입 여부-미가입

E1 : 음성요금제 일반

E2 : 음성요금제 패밀리

E3 : 음성요금제 TTL커플

E4 : 음성요금제 TTL

E5 : 음성요금제 텡

F : 전 3개월 평균 ARPU

G : 전 3개월 평균 데이터 통화료

H : 평균 음성요금제 유지개월 수

I : 최종 단말 획득 후 보유 기간

J1 : NATE 이용 Heavy 또는 Active 유지 여부-N

K : 전 3개월 NATE 이용 Trial 개월 수

L : 전 3개월 평균 TOP 뮤직 10 통화료 과금액

M : 전 3개월 평균 단순 상담 건수

N1 : 전 3개월 부가서비스 가입 건수-미이용

N2 : 전 3개월 부가서비스 가입 건수-1건

N3 : 전 3개월 부가서비스 가입 건수-2건

N4 : 전 3개월 부가서비스 가입 건수-3건

O1 : 전 3개월 부가서비스 해지 건수-미이용

O2 : 전 3개월 부가서비스 해지 건수-1건

O3 : 전 3개월 부가서비스 해지 건수-2건

O4 : 전 3개월 부가서비스 해지 건수-3건

P1 : 전 3개월 유사 부가서비스 가입 여부-N

Q : TM 캠페인 성공률

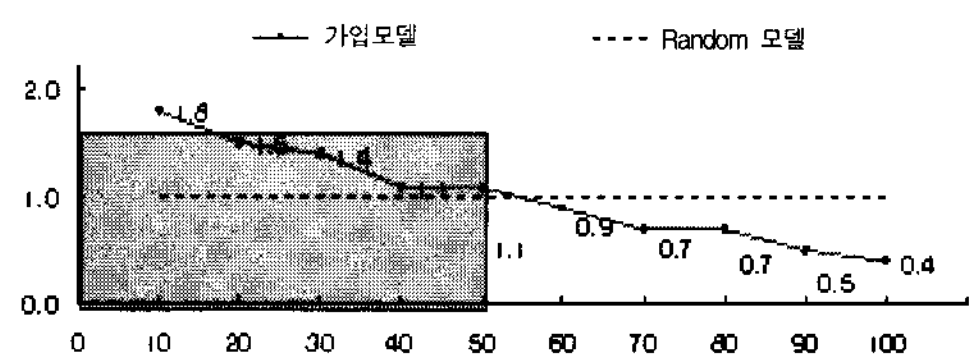
R1 : TM 제외 채널 캠페인 성공 여부-N

3.2 모델 효과 검증

최종 선택된 가입 모델은 22개의 설명변수에 의한 로지스틱 회귀 모형에 의해 결정되었다. 모델 적합도를 나타내는 Akaike's Information Criterion (AIC) 통계량은 65,536, Schwartz Bayesian Criterion(SBC) 통계량은 66,117이며 정분류율은 69.4%로 나타났다. <표 6>은 선택된 로지스틱 회귀모형에 의한 가입 모델을 검증 데이터 세트(validation data set)에 포함된 고객들에게 적용하여 얻은 예상 가입 확률이 높은 순서로 10등분을 하여 표현된 이익 도표(Gain's chart)이다.

<그림 5>는 모델을 사용하지 않은 Random 모델과 비교했을시 가입모델을 이용하는 경우 상위 50%를 선택하는 것이 고객 Pool을 최대한 유지할 수 있기 때문에 매우 효율적임을 보여주고 있다.

가입 모델을 통해 가입을 스코어 상위 50% 고객을 타겟팅할 경우 예상 가입율은 32.8%에서 45.1%로 향상된다. 즉 가입 모델에 의한 스코어 상위 50%를 타겟팅할 경우 모델을 사용하지 않은 경우에 비해 1.4배 타겟팅 효과를 가져오며, 월 평균 TM 대상 고객 수 103,597명에 대해 예상 가입 고객은 월 평균 45,035명 수준이 될 것이다.



<그림 5> 가입모델과 Random모델의 비교

<표 6> 가입 모델링에 의한 이익 도표

Decile	가입율	누적 가입율(%)	Captured Response	누적 Captured Response	Lift	누적 Lift	누적 가입 예상 고객(명)
10	58.7	58.7	17.9%	17.9%	1.8	1.8	60,762
20	50.2	54.4	15.3%	33.2%	1.5	1.7	52,022
30	44.8	51.2	13.6%	46.8%	1.4	1.6	46,374
40	37.4	47.8	11.4%	58.2%	1.1	1.5	38,745
50	34.5	45.1	10.5%	68.7%	1.1	1.4	35,689
60	30.6	42.7	9.3%	78.1%	0.9	1.3	31,745
70	22.2	39.8	6.8%	84.8%	0.7	1.2	22,972
80	21.5	37.5	6.6%	91.4%	0.7	1.1	22,300
90	16.1	35.1	4.9%	96.3%	0.5	1.1	16,704
100	12.1	32.8	3.7%	100.0%	0.4	1.0	12,484
전체	32.8	32.8	100%	100.0%	1.0	1.0	103,597

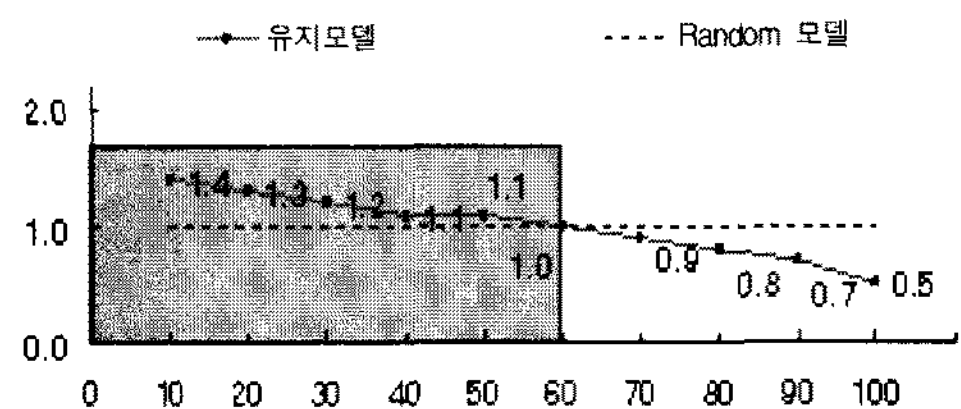
<표 7> 유지 모델링에 의한 이익도표

Decile	가입율(%)	누적 유지율(%)	Captured Response	누적 Captured Response	Lift	누적 Lift	누적 가입 예상 고객
10	73.9	73.9	13.5%	13.5%	1.4	1.4	44,903
20	70.5	72.2	12.9%	26.4%	1.3	1.3	36,653
30	67.7	70.7	12.4%	38.8%	1.2	1.3	31,383
40	60.8	68.2	11.1%	49.9%	1.1	1.2	23,551
50	57.7	66.1	10.6%	60.4%	1.1	1.2	20,606
60	54.3	64.1	9.9%	70.4%	1.0	1.2	17,235
70	50.8	62.2	9.3%	79.7%	0.9	1.1	11,681
80	46.1	60.2	8.4%	88.1%	0.8	1.1	10,276
90	38.8	57.8	7.1%	95.2%	0.7	1.1	6,475
100	26.4	54.7	4.8%	100.0%	0.5	1.0	3,292
전체	26.4	54.7	4.8%	100.0%	1.0	1.0	103,597

최종 선택된 유지 모델은 20개의 설명변수에 의한 로지스틱 회귀 모형에 의해 결정되었다. 모델 적합을 나타낸 AIC 통계량은 22,693, SBC 통계량은 23,004이며 정분류율은 62.9%로 나타났다. <표 7>은 선택된 로지스틱 회귀모형에 의한 유지 모델을 검증 데이터 세트에 있는 고객들에게 적용하여 구한 예상 유지 확률이 높은 순서로 10등분을 하여 표현된 이익 도표이다.

<그림 6>은 유지모델을 사용하지 않은 Random 모델과 비교했을시 유지모델을 이용하는 경우 상

위 60%를 선택하는 것이 고객 Pool을 최대한 활용함과 동시에 가장 효율적임을 보여주고 있다.



<그림 6> 유지모델과 Random모델의 비교

유지 모델을 통해 유지율 스코어 상위 60% 고객을 타겟팅할 경우 Top 뮤직 10 정액제를 가입한 고객의 유지율은 54.7%에서 64.1%로 향상된다. 즉 유지 모델에 의한 스코어 상위 60%를 타겟팅할 경우 모델을 사용하지 않은 경우에 비해 1.2배 타겟팅 효과를 보여 주며, 월 평균 가입고객의 수 103,597명에 대해 예상 유지 고객은 월 평균 20,719명 수준이 될 것이다.

3.3 가입/유지 결합모델

본 연구에서 제안한 가입율과 유지율 두 모델을 이용해 예상 가입율과 예상 유지율이 높은 고객 군을 타겟팅하는 결합모델 방법은 상위 고객을 우선 타겟 순위를 두며 TM 대상 Pool이 더 필요한 경우 순서에 의해 TM 고객 대상을 늘릴 수 있는 장점이 있다.

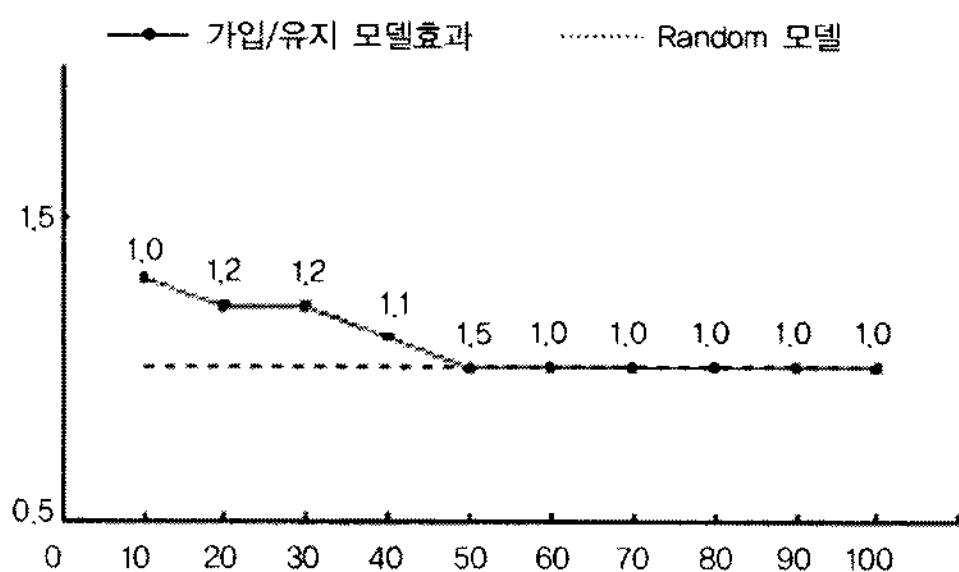
TM 대상 고객 후보 고객 군중에서 예상 가입율 상위 50%와 예상 유지율 상위 60% 이상의 고객을 선택하는 결합모델은 전체 고객 30%에 해당된다.

<표 8>을 보면, 결합모델에 의해 고객을 타겟으로 선택할 경우에 모델을 이용하지 않은 기존 타겟팅 방법보다 가입율은 11.2%, 유지율은 6.7% 향상된 결과를 제시하며, <그림 7>은 가입율과 유지율을 결합한 결합모델에 의해 상위 30% 고객을 타겟팅할 때 모델을 사용하지 않은 Random 모델에 비해 20% 향상된 결과를 보여준다. 본 논문에서 제시한 결합모델에 의한 타겟팅 방법은 고객의 성향 및 이용 패턴을 고려했기 때문에 고객가치를 증대시킬 수 있을 뿐 아니라 A사 입장에서도 고객의 불만을 줄이고 고객 유지(Retention)를 증대시켜 수익성을 더욱 향상시킬 수 있을 것이다.

A사는 2008년 4월부터 기존의 Top 뮤직 10 정액제 부가서비스를 Top 뮤직 10 정액제 플러스라는 상품으로 개편시켜 이용 포인트를 추가 적립해 줌으로써 이용량에 따른 혜택이 확대된다. 이는 과거 Top 뮤직 10 정액제와 전체적인 서비스는 비슷하나 이용에 따른 포인트 제도가 적용되어 고객 가치가 더욱 증대되어 유지율 향상에 도움이 될 것으로 기대된다.

<표 8> 결합모델 효과

	가입율	가입 후 유지율
현안	32.8%	54.7%
개선안	44.0%	61.4%
개선안 효과	11.2%	6.7%
Lift	1.34	1.2



<그림 7> 결합모델과 Random 모델의 Lift 값 비교

4. 결론 및 제언

기존의 Top 뮤직 10 정액제 상품을 Outbound TM을 통해 판매하기 위한 타겟팅 방법은 단순한 Top 뮤직 콘텐츠 별 Hit수에 의해 타겟 대상을 정하였다. 이는 고객의 과거 이용 패턴과 고객의 특성을 반영하지 않았고 고객의 니즈를 반영하지 않았기 때문에 가입 후 유지율이 낮게 나타나 고객 가치를 낮추는 결과를 초래하였다. 본 연구에서는 기존의 타겟팅 방법보다는 과거 Outbound TM에 의한 Top 뮤직 10 가입 고객의 특성과 과거 행태를 반영한 예상 가입율과 예상 유지율이라는 스코어를 통해 타겟팅하는 데이터마닝의 스코어링에 의한 타겟팅 방법을 제안하였다. 본 논문에서 제시된 가입모델과 가입 후 유지모델을 결합한 결합모델을 이용하여 상위 30%를 타겟팅 할 경우에, 기존의 방법보다 가입율은 11.2%, 유지율은 6.7% 향상

된 결과를 보여주었다. 또한 고객의 성향 및 이용 패턴을 고려했기에 고객가치를 증대시킬 수 있을 뿐 아니라 A사 입장에서 고객 불만을 줄이고 고객 유지를 증대시켜 수익성을 더욱 향상시킬 수 있을 것으로 기대된다.

또한 2008년 4월부터 출시되는 Top 뮤직 10 정액제 플러스라는 상품에도 본 논문에서 제시한 모델링에 의한 타겟팅 방법으로 판매한다면, 기존의 Top 뮤직 10 정액제 상품에 포인트가 결합되어 상품력 또한 향상되어 가입 후 유지율이 더욱 더 증대 될 것으로 사료된다.

참 고 문 헌

- [1] 강현철, 한상태, 최종후, 김은석, 김미경, 「데이터마이닝-방법론 및 활용」, 자유 아카데미, 2002.
- [2] 김경무, 전희주, “이동통신사의 M고객센터 활용 증대를 위한 데이터마이닝 활용”, 「Journal of the Korean Data Analysis Society」, 제10권, 제1호(2008), pp.1229-2354.
- [3] 김용수, “로지스틱 회귀분석과 의사결정 나무 분석을 이용한 데이터 요금제의 해지율 예측 모형 수립”, 「Journal of the Korean Data Analysis Society」, 제8권, 제5호(2006), pp.1915-1926.
- [4] 전희주, 정병철, “이동통신사의 로밍 이용 예상 고객 타겟 모델링”, 「Journal of the Korean Data Analysis Society」, 제9권, 제5호(2007), pp.1229-2354.
- [5] 진서훈, 김현기, “신용카드 이용규모 추정 모델 개발을 위한 데이터마이닝의 활용”, 「Journal of the Korean Data Analysis Society」, 제8권, 제6호(2006), pp.2415-2425.
- [6] 한상태, 강현철, 이성건, 정요천, “교차판매 스코어링 모형 개발”, 「응용통계연구」, 제17권(2004), pp.229-238.
- [7] 최종후, 한상태, 강현철, 김은석, 김미경, 이성건, 「데이터마이닝-기능과 사용법」, 자유아카데미, 2001.
- [8] Berry, M.J.A. and G. Linoff, *Data Mining Techniques For Marketing, Sales, and Customer Support*, John Wiley and Sons, New York, 1997.
- [9] SAS Institute, *Data Mining Using SAS Enterprise Miner 4.3*. SAS Institute Inc, 2003.