

데이터베이스 모델링에 의한 효과와 고객세분화

전희주¹⁾, 김태진²⁾

요 약

상위 20%의 우량고객이 기업수익의 80%를 기여한다는 파레토 법칙을 인용하지 않더라도 고객이 원하는 상품과 서비스를 지속적으로 제공함으로써 고객과의 지속적인 관계를 통한 고객과의 지속적인 관계유지, 특히 우량고객의 확보, 유지는 기업의 수익증대에 깊은 관계가 있으며 결국 기업의 생존을 가늠하는 길이 될 것이다.

최근에는 금융권간 업무영역이 무너지면서 모든 금융기관들은 우수고객 확보를 위해 영업력을 집중시키고 있다. 특히 수익기여도가 높은 우수고객에 대해서는 차별적인 우대서비스를 명시적으로 규정하고 우량고객 확보를 위한 생존을 건 경쟁이 벌어지고 있다. 따라서 우량고객은 자료를 통한 객관적이고 합리적 기준에 의해 확보되어야 한다.

본 연구에서는 이에 대한 방법론으로서 데이터베이스 모델링을 제시하며, 그 효과를 측정하도록 하고, 고객 속성에 따른 고객세분화에 이용될 수 있음을 보여준다.

주요용어 : 데이터베이스 모델링, 데이터베이스 모델, 고객세분화

1. 서론

데이터베이스 모델링에 의한 후보모델중 선택된 모델을 데이터베이스 모델이라 정의할 때, 이 데이터베이스 모델의 효율성을 측정하는 여러 가지 방법이 제시될 수 있을 것이다.

본 논문에서는 모델링 하기 전의 데이터에서 랜덤하게 추출한 대상집단을 편의적으로 랜덤모델이라 정의하고 데이터베이스 모델과 랜덤모델의 비교를 통해서 데이터베이스 모델링의 효과를 제시하고 이를 고객세분화에 어떻게 적용시킬 것인지를 알아보려고 한다.

2. 데이터베이스 모델링에 의한 효과

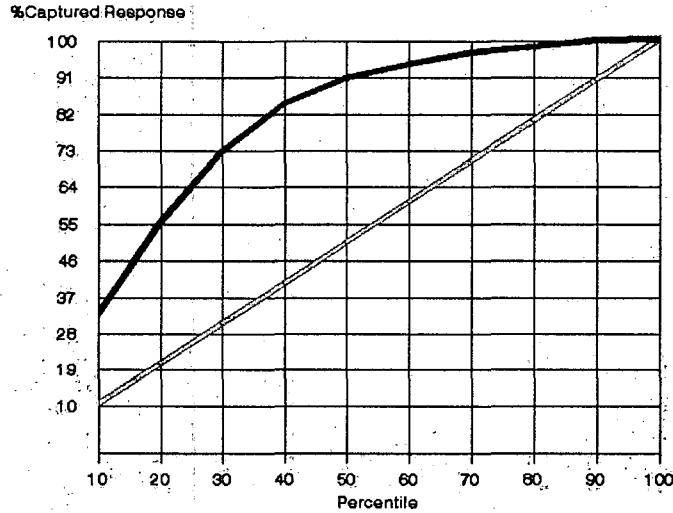
현재 금융업계에서는 우량고객의 확보를 통해 수익은 극대화시키면서도 비용은 줄이는 것이 주 관심사가 되고 있으며 이를 시행하기 위한 여러 가지 방법 중, DM(Direct Mailing)은 데이터베이스 마케팅을 위해 자주 사용하는 방법이다. 불특정 다수의 고객에 대한 DM을 발송하는 것보다는 회사의 수익기여에 도움이 되는 카드이용 가능성이 높은 고객을 마케팅의 우선 대상으로 한다면 이는 마케팅비용을 줄이고 회사의 수익을 높여 시장경쟁력을 높히는데 기여할 것이다.

1) 삼성카드 마케팅팀 CRM Part 과장. 서울 중구 을지로1가 87번지 삼성화재빌딩 12층

2) 삼성카드 마케팅팀 CRM Part. 서울 중구 을지로1가 87번지 삼성화재빌딩 12층

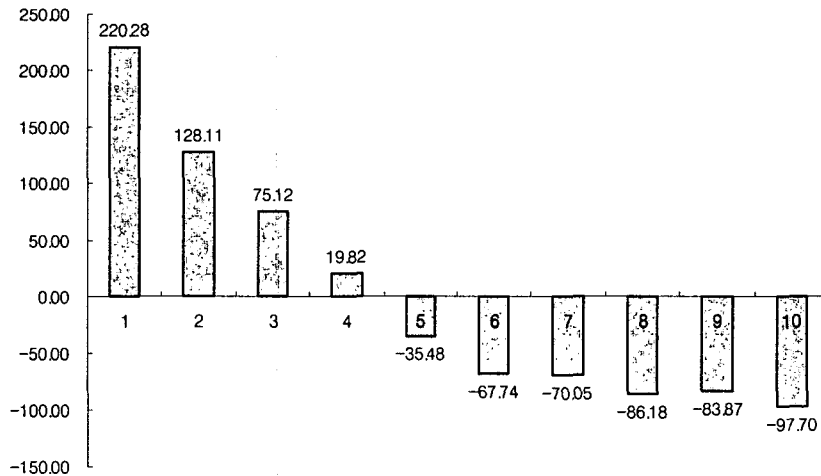
데이터베이스 모델링에 의한 효과와 고객세분화

본 자료는 모카드회사의 2000년 5월 자료를 통해 모델링을 하여, 2000년 8월 대상자에 모델을 적용시켜 DM발송하는 상황설정을 통해 기술하였다.



<그림 1> 데이터베이스모델 vs. 랜덤모델

<그림 1>은 데이터베이스 모델과 랜덤모델에 대한 %Captured Response를 보여주는데, 랜덤모델은 상위 30%의 고객집단에서 실제 30%정도의 이용한 고객을 추출할 수 있으나, 데이터베이스 모델의 경우, 73%의 실제 이용한 고객을 포함해, 랜덤모델에 비해 143%정도 효율이 향상되었음을 쉽게 알수 있다.



<그림 2> Marginal Incremental Gains Over Overall Response

※ Marginal Incremental Gains = (% marginal response / overall response - 1) × 100.

<그림 2>는 순서화된 예측값에 의해 동일한 수를 가진 10개의 집단으로 나누었을 때 랜덤 모델 대비 데이터베이스 모델링에 의한 순수 증가효과를 나타낸 것이다. 상위 4개 집단까지는 (+)효과를 주지만 나머지는 오히려 평균 이하의 효과를 보여준다. 실제 타겟 마케팅을 시행할 때 상위 40%에 주력해야 함은 자명한 사실이다.

Prospect Decile Group	Database Model	Cum % Response	% Gain in Response Due to Model
10	69.5	69.5	220.28
20	49.5	59.5	174.19
30	38	52.3	141.01
40	26	45.7	110.60
50	14	39.4	81.57
60	7	34.0	56.68
70	6.5	30.1	38.71
80	3	26.7	23.04
90	3.5	24.1	11.06
100	0.5	21.7	0.00
total	21.7		

<표 1 > 데이터베이스 모델 Gains Chart

※ %Gains in Response due to model = (Cum % Response/Overall Response - 1) × 100.

<표 1 >은 데이터베이스 모델링에 의한 %Response와 %, 누적 %Response와 랜덤모델 대비 순수 증가 효과를 정리하였다. DM을 발송해야할 집단이 상위 20%까지의 두 집단, 3만명이고 (전체 : 15만명), 각각의 DM당 요금은 320원이라면,

- 데이터베이스 모델:
 - 상위 20% 집단의 누적 %Response은 59.5%이다.
 - 반응자의 수 = 30,000 × 0.595 = 17,850.
 - 비용 / 반응자 = 30,000 × 320 / 17,850 = 538 (원).
- 랜덤모델 :
 - 평균효과를 보여주는 랜덤모델의 누적 %Response은 21.7%이다.
 - 반응자의 수 = 30,000 × 0.217 = 6,510.
 - 비용 / 반응자 = 30,000 × 320 / 6,510 = 1,474 (원).

위에서 기술한 바와 같이, 데이터베이스 모델링을 통한 반응자 수는 랜덤모델에 의한 6,510명보다 11,340명이 많은 17,850명이고, 174% 정도 반응율이 개선된다고 할 수 있다. 하지만 오히려 비용은 1,474원에 비해 936원 줄은 538원으로서 63.5%의 비용절감 효과가 나타났다.

이러한 결과는 데이터베이스 모델링이 고객들의 과거 정보 또는 이력을 사용하지 않는 경우에 비해 실제 타겟마케팅을 실현하는데 수익과 비용 양자에 있어서 기여할 수 있음을 보여주는 것이다. 또한 이렇게 선정된 고객군에 대해서 다시 고객의 속성에 따라 고객세분화를 함으로써 고객의 니즈를 충족시키는 더욱 더 발전된 데이터베이스 마케팅을 실현할 수 있을 것이다.

3. 결론

앞으로 금융기관들은 성장위주의 외형확대 전략에서 탈피해 우수고객에 초점을 맞춘 고객관리전략을 적극 추진해야만 생존할 수 있을 것으로 기대되고, 따라서 금융기관들은 수익기여도를 정확히 평가할 수 없는 신규고객의 획득보다는 기존의 수익성 높은 고객과의 우호적인 관계 구축에 보다 비중을 두는 고객유치전략의 강화에 보다 역점을 기울여야 할 것이다.

자본과 시간의 효율적 투자를 위해 데이터베이스 모델링에 의한 타겟마케팅은 필수적이며, 이를 통한 고객세분화는 향후 보다 세밀하게 적용되어야 할 것이다.

참고문헌

- [1] 최중후, 한상태, 강현철, 김은석, 김미경 (2000). 『SAS Enterprise Miner를 이용한 데이터마케팅』, 서울 : 자유아카데미
- [2] Ian H.Witten and Eibe Frank. (2000). *Data Mining, Practical Machine Learning Tools and Techniques with Java Implementations*, Morgan Kaufmann Publishers.
- [3] Michael J. A. Berry and Gordon Linoff (2000). *Mastering Data Mining, The Art and Science of Customer Relationship Management*, Wiley Computer Publishing.
- [4] Quinlan, J. R. (1993). *C4.5 Programs for machine learning*. San Mateo : Morgan Kaufmann.
- [5] Berry, M. J. A. and Linoff, G. (1997). *Data Mining Techniques for Marketing, Sales, and Customer Support*, New York : John Wiley & Sons.
- [6] Kass, G. (1980). An Exploratory technique for investigating large quantities of categorical data. *Applied Statistics*, 29:2, 119-129.
- [7] Smith, M. (1996). *Neural Networks for Statistical Modeling*, International Thomson Computer Press.