

# 왜, 어떤 방법으로 고객을 관리할 것인가

정교한 정보분석 기술을 사용하지 않는다고 해서 고객관계관리(CRM)가 이루어질 수 없는 것은 결코 아니다. 그러나 수백만에 이르는 고객들을 각기 특성에 맞도록 차별적으로 관리하기 위해서는 수많은 의 사결정이 필요하며, 또한 이를 위해서 수많은 분석이 이루어져야 하기 때문에 정보기술의 도움을 받을 수 밖에 없다. 이번호에서 생각해 보는 주제는 CRM을 위한 정보의 저장 및 분석을 위한 기술들이다.

- 한재홍/경영학 박사 · 농심데이터시스템 DW팀장 (hanjh@nongshim.co.kr)
- 전용준/프리랜서 컨설턴트 (xyxon@hanmail.net)

## 연재 순서

1. 고객관계관리의 전략적인 가치
2. 고객관계관리를 통한 업무와 조직의 혁신
3. 고객관계관리를 위한 정보분석 기술의 적용 방안
4. 인터넷 비즈니스 기반의 고객관계관리
5. 고객관계관리의 발전 방향

## 1. 데이터 웨어하우징 기술 적용

CRM에서 데이터 웨어하우징이 왜 필요한가? 데이터를 분석하기 위해서는 기본적으로 데이터가 존재해야 하지만 존재하는 모든 데이터가 분석에 유용하지는 않다. 분석을 위해서는 데이터가 분석에 적합한 내용과 형태를 가지고 있어야 한다. 이러한 요구 사항을 충족하기 위해 데이터 웨어하우징을 고려하는 것이다.

데이터 웨어하우스는 수집 가능한 모든 분석용 정보를 한 곳에 통합시킴으로써 적절한 내용과 형태의 분석을 위한 기초 데이터를 제공하는 역할을 한다. 데이터 웨어하우징의 또 한가지 중요한 필요성은 가용성을 제공하는 점이다. 분석용 정보가 업무 운영을 위한 정보와 동일한 시스템에 존재한다면 이를 언제라도 사용하는 것이 가능하도록 하기 어렵기 때문이다.

그렇다면 어떤 정보가 분석을 위해서 필요한가? 쉬운 듯 한데도 불구하고 결코 쉽지않은 질문이다. 필요한 정보는 하고자 하는 분석이 무엇인가에 따라 달라진다. 따라서 분석하고자 하는 내용을 살펴보는 것이 분석용 정보를 정의하기 위해 필요하다. CRM을 위한 분석용 정보를 정의하기 위해서는 CRM과정상에서 무엇을 분석할 것인지를 파악하여야 한다.

분석을 위해 필요한 정보가 무엇인지 파악되면 데이터베이스를 구축해야 한다. 이때 데이터베이스 역시 분석업무를 위한 경

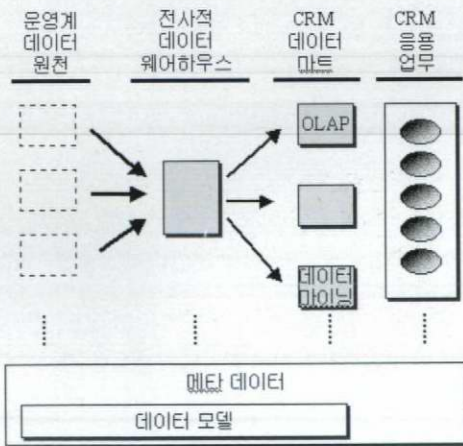
우에는 특별한 요건이 존재한다. 여러 원천으로부터 수집된 데이터 구조가 표준화되고 통합되어야 하며, 시계열적으로 데이터가 축적되어야 하고, 분석에 사용가능 하도록 데이터를 정제하여야 한다.

분석용 정보는 사람 대신 컴퓨터 시스템이 판단하기에 부족하지 않은 수준의 정확성을 가지도록 해야 한다. 다양한 목적으로 활용하기 위해서는 가능한 한 상세한 데이터를 포함시켜서 분석의 신축성을 보장해야 한다.

모델링 과정에서 가능한 한 정규화 시키는 것이 바람직하다. 자주 사용되는 파생 데이터, 즉 요약, 추정, 예측 결과 등은 통제된 중복성을 허용하여야 한다. 그러나 데이터베이스 구조의 결정과정에서 요약된 데이터와 원시 데이터를 별도로 관리할 것인가에 대한 결정에 따라 설계는 달라진다.

### ■ 기존 데이터베이스 통합 모델링의 필요성

CRM을 위한 분석용 데이터베이스를 구축하는 과정에서 첫번째 작업은 기존 데이터베이스가 가지고 있는 항목들 중 분석에 필요한 항목들을 통합하는 모델을 작성하는 것이다. 이 작업의 의미는 분석에 사용될 데이터들에 대한 전사적인 개념 정의를 도출한다는 것이다. 이 과정은 분석용 데이터베이스의 범위와 지원할 수 있는 분석의 범위를 결정하는 데 결정적인 영향을 미친다. CRM에서의 데이터 분석은 전사적으로, 그리고 고객단위로 통합되지 않으면 의미가 없기에 통합 모델링은 매우 중요하다.



(그림 1) 데이터 웨어하우스 아키텍처와 메타 데이터 구성의 일반적인 모습

### ■ 워크샵을 통한 사용자 참여 및 요구 사항의 반영

어떠한 정보를 저장해 두고 사용할 것인가를 결정하는 것은 결

국 사용자의 몫이다. 결코 정보시스템 분석가가 의견을 낼 수 없는 사항이다. 따라서 워크샵 등의 형태로 수행되는 요구정의 과정에서 물론 정보시스템 분석가의 지원을 어느 정도 받기는 하겠지만 사용자 스스로 생각해 내야 한다. 그러나 문제는 분석이라는 업무의 성격상 필요한 모든 데이터를 미리 파악하거나 파악된 정보가 진정으로 유용한지에 대해서도 결정하기가 쉽지 않다는 점이다.

데이터의 필요성 또는 유용성에 대한 불확실성은 분석용 정보의 성격상 불가피한 부분이다. 모든 데이터 분석은 통상 의사결정을 내리기 위한 용도로 이루어 진다. 그러나 CRM에서의 의사결정을 내려야 할 문제라는 것 자체가 비구조적인 성격이며 고객의 행동에 영향을 미칠 수 있는 변수들이 워낙 다양하기 때문에 어떠한 정보가 사용될 것인지를 미리 알 수 없는 경우가 대부분이다. 하지만 정보시스템이라는 정형적인 것을 구현하는 입장에서는 불확실성을 최소화 시켜야만 한다. 가장 현실적인 방법은 결국 틀림없이 필요할 것으로 예상되는 데이터 항목들은 포함시키고, 불확실한 것 중 상대적으로 중요하고 불확실성이 낮은 것을 포함시키며, 상대적으로 반복적으로 사용될 가능성이 높은 것을 포함시키는 것이다.

하지만 이 또한 주관적인 판단에 의존하게 되므로 가능하다면 복수의 업무 전문가 또는 담당자가 상호 의견을 종합해야 한다. 또 내부의 의견만으로 결정하기 어려운 사항이나 경쟁 또는 유사업종의 타사에서도 공통적으로 적용될 부분에 대한 것들은 외부 전문가를 활용하거나 벤치마킹을 수행하는 것이 바람직하다.

이와 같은 불확실성 때문에 의사결정 지원용 데이터베이스는 구조 자체가 비교적 자주 변경될 수 있다. 업무 요구사항의 변화 이유는 두 가지로 나누어 볼 수 있을 것이다. 그 하나는 무엇이 중요한지가 환경 변화에 따라 달라지는 경우이다.

나머지 하나는 CRM 실무 담당자의 자사 CRM에 대한 전문적 이해 부족일 수 있다. 실제로는 이 두가지가 혼재한다. 장기간의 데이터를 저장하고 분석하게 되므로 시간이 가면서 업무 규칙이 변경될 수도 있는 것이며 그에 따른 변화를 담당자가 충분히 파악하지 못할 수도 있는 것이다. 이러한 문제점을 극복하고 서로 다른 시점의 업무 규칙과 그를 바탕으로 발생된 데이터를 모두 반영하기 위해서는 일반화된 형태의 구조로 구축되어야 한다.

### ■ 통합된 고객 데이터베이스의 구축 방안

초기에는 많은 논란들이 있었으나 이제는 더 이상 데이터 웨어

하우스인가, 그저 고객 데이터베이스인가 하는 식의 논란은 큰 의미가 없다고들 인식해가는 상황이며 고객 데이터 웨어하우스 또는 데이터베이스를 구축하기 위해서는 상향식 방법과 하향식 방법을 모두 사용하는 것이 필요하다. 운영상의 현실성과 CRM의 이상향을 모두 반영하는 것이 필요하기 때문이다.

하향식 방법에서는 CRM의 비전으로부터 출발하여 CRM에 필요한 업무 기능 및 데이터 요구 사항을 점진적으로 구체화시켜 나간다. 반대로 상향식 방법에서는 기존 데이터베이스에서 고객 관련 항목을 추출해서 그대로 모델링하고 점진적으로 통합시켜 나간다.

완성된 두 가지 모델의 차이를 조정하여 하나의 모델을 완성함으로써 현실성과 이상적인 모습을 조율한 새로운 모델을 완성하게 되며 이 그림을 바탕으로 새로운 통합된 고객 데이터베이스를 구축하게 된다. 이 과정에서 주의해야 할 사항 중 하나는 현재의 데이터 관련 규칙 및 정의가 합리적이지 않을 수 있다는 점이다. 만일 데이터에 대한 정의가 합리적일 수 없다면 어떠한 분석기법을 적용해서도 좋은 분석이 이루어지는 것은 불가능해진다.

■ CRM에서의 OLAP기술 활용

CRM에서 왜 OLAP(Online Analytical Processing)을 사용하는가? 분명 OLAP기술이 CRM에 제공하는 이익이 존재하기 때문일 것이다. CRM은 고객별로 차별적인 마케팅 및 서비스 등을 제공하기 위해 다양한 측면의 고객, 상품, 구매행동, 서비스이력 등을 분석해야 하며 전사적인 수준에서부터 개별적인 고객에 이르기 까지 다양한 수준의 분석 깊이를 요구한다. 따라서 데이터 베이스에 대한 조작없이 이러한 다양하고 신축적인 분석을 즉시 수행하도록 지원하는 OLAP기술이 필요한 것이다.

OLAP은 생각보다 오랜 역사를 가지고 있다. 1980년대에 IRI 나 로터스 등이 이미 OLAP 기술을 제품화 하였으나 사용은 최근에 와서 급속히 확산되었다. CRM에서의 다양한 분석을 위해 OLAP은 필수적이며 데이터 마이닝과는 보완적인 기능을 수행한다.

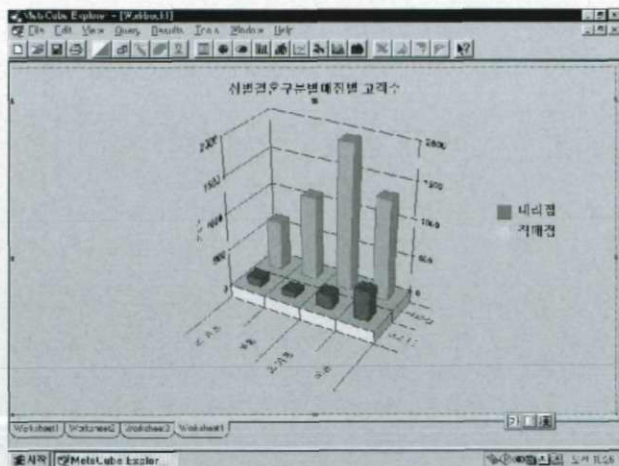
CRM에 참여하는 모든 사람에게 데이터 분석은 필요하지만 방법과 수준, 그리고 분석의 주체에는 차이가 있다. 대상 사용자가 임원인가, 일선관리자인가, CRM 본부 부문의 전문 데이터 분석가인가 등에 따라 서로 다른 도구와 내용을 OLAP을 통해서 분석하게 된다. OLAP적용시의 세가지 중요한 주의사항은 다음과 같다

- 모든 것에 대한 요약 데이터를 만들지 말라.  
여러번 사용할 것에 대해서만 OLAP을 적용하고 아주 복잡한 것은 OLAP으로 분석하려 하지 말아야 한다. 많은 요약 데이터를 가질수록 OLAP의 성능이 좋아지지만 그대신 비용이 증가한다.
- OLAP은 정확하거나 정교한 패턴을 찾기 위한 것이 아니다.  
OLAP은 어디까지나 분석가의 직관에 의존하는 분석 방식이다. 다차원 테이블 또는 그래프로 보여지는 것들에 대해 분석가의 시각적 판단이 이루어져야 하기 때문에 만족할 만한 일정 수준 이상의 답을 파악하는 데 만족할 수 밖에 없다.
- OLAP적용 대상에 대한 근거가 되는 상세 데이터를 준비하라.  
아무리 획기적인 결과가 제시되어도 그 산출 근거를 즉시에 추적할 수 없는 경우에는 사용자의 신뢰를 얻지 못한다. 별도의 상세 데이터와 연계하거나 상세 데이터와의 연계가 OLAP상에서 지원되는 방식을 선택하는 것이 바람직하다.

2. 데이터 마이닝 기술 적용

데이터 마이닝은 이전에 알지 못하던 사실을 발견하기 위한 데이터 분석을 의미한다. 알고 나면 지극히 당연하지만 미리 그 중요성을 인식하지 못했던 사실을 찾아내는 것이 데이터 마이닝의 역할이다. 다시 말하면 데이터 마이닝은 새로운 사실이 밝혀지기 이전까지는 당연하게 여겨지던 사실이 당연하지 않다는 것을 과학적으로 설명해 가는 과정이라고 할 수 있다.

최근 데이터 마이닝에 대한 관심이 부쩍 늘어가고 있지만 아직까지 데이터 마이닝의 가장 큰 적용 영역은 CRM과 관련된 영역이다. 그 이유는 CRM이 다른 업무분야에 비해서 비교적 가시적인 성과를 제공하기 때문이다. 물론 다른 영역에 대한 적용에 문



<그림 2> OLAP을 이용한 CRM의 분석 예

제가 존재하는 것은 아니다. 데이터가 충분히 존재하고 그 데이터를 사용하여 분석하는 것이 가능하며 분석으로부터 얻은 지식이 업무성과를 향상시키는데 충분한 이익이 있다면 데이터 마이닝을 적용하는 것은 가능하다.

비록 아직까지도 데이터 마이닝은 실무에 대한 적용의 초기 단계에 있다고 할 수 있으나 여러 업종에 걸쳐서 성공적인 사례가 상당히 많이 발표되어 있기 때문에 먼 미래의 이야기로 치부할 수 없는 상황이다. 그러나 아직도 데이터 마이닝에 관해 오해가 많이 존재하고 있고 많은 사람들이 정확한 인식을 가지고 있지 못한 것도 사실이며 데이터 마이닝을 위한 도구도 원하는 모든 것들을 해결하기에는 아직 충분치 않으나 빠른 속도로 발전하는 중이다.

데이터 마이닝을 처음 도입하는 조직들이 실패를 경험하는 경우도 적지는 않다. CRM에 적용하는 경우 업무에 대한 매우 높은 수준의 이해를 요구하며, 데이터 준비, 분석기법에 대한 폭넓은 이해, 톨 준비 모두 필요하기 때문에 간단히 끝나지 않는다.

### ■ 누가 참가해서 어떤 일을 해야 하는가?

데이터 마이닝은 통상 프로젝트 단위로 진행된다. 이 과정에 참여하는 사람들을 크게 네가지 일반적인 부류로 나눌 수 있는데 현업전문가는 업무 전반에 관한 상식, 경험 및 전문지식을 보유한 사람이며, 문제 분석 전문가는 해당 문제에 대한 해결 접근방법을 수립하고 모델의 타당성을 평가하는 사람이고, 데이터베이스 전문가는 데이터 마이닝에 사용될 데이터를 준비하는 역할을 하고 마지막으로 분석 기법 전문가는 데이터 마이닝 도구 및 모델링 기법을 적용하는 일을 한다.

실제 CRM에서의 데이터 마이닝 프로젝트에서는 통상 CRM 본부 부문 담당자와 외부 CRM 전문 컨설턴트, 데이터 마이닝 도구 사용 기술자, 내부 데이터베이스 관리자 등이 프로젝트에 참여하게 된다. 이 과정에서 CRM 전문 컨설턴트의 역할은 CRM과 관련된 업무적인 요구사항을 명시적으로 도출하고, 타사 사례 등으로부터 도출된 지침들을 제시하고 고객을 도와서 문제 수립 및 해결 과정 전반을 수행하는 것이다.

데이터 마이닝 도구 사용에 대한 전문 기술만을 가진 인력은 실상 문제에 대한 이해가 부족한 경우가 많아서 현업 담당자의 문제를 그대로 반영하는 모델과 지식을 찾아내지 못하는 경우가 발생하기 때문에 특히 데이터 마이닝을 새로이 도입하는 경우라면 데이터 마이닝 프로젝트 경험을 가진 외부 CRM 전문가를 프로

젝트에 포함시키는 것이 반드시 고려되어야 할 것이다.

하지만 아직까지 국내에서는 문제 분석 전문가나 분석 기법 전문가를 구하는 경우 충분한 자격이나 경험을 갖춘 사람을 만나기가 어려운 상황이다. 학계의 인력들은 대규모 데이터에 대한 분석 경험이 없는 경우 또는 CRM과 같은 현실 문제에 대한 경험이 없는 경우가 많고 업계의 인력들은 기술 및 이론에 대한 기반 지식이 부족한 경우가 많다. 이는 아직까지 데이터 마이닝이 국내 적용의 시작 단계이기 때문일 것이다.

### ■ 고객행동 패턴의 파악 및 예측을 위한 기법

고객의 구매 행동 패턴 파악에 가장 활용도가 높은 방법은 기존의 통계적인 프로파일링이다. 데이터 마이닝이라는 거창한 방법을 적용하는 대신에 과거의 구매 행동을 단순한 방법으로 일반화시키는 것이므로 직관적이며 이해가 쉽고 결과 도출을 위해 필요한 노력의 양도 적다.

그러나 해결하고자 하는 개별적인 문제가 가지고 있는 특성에 적합한 지식 또는 해결책을 찾기 위해서는 보다 정교하고 다양한 기법을 적용하는 것이 필요하다. 이 때문에 <표>에 제시된 것과 같이 인공지능 및 통계를 망라한 여러 기법들이 사용된다. 이러한 다양한 기법들은 상대적인 장단점들을 가지고 있으나 새로운 기법들 중 일부는 과거의 기법에 비해 여러 측면에서 우월한 경우가 많다.

### ■ 교차 판매를 위한 데이터 마이닝 모델링

여기서는 CRM에 데이터 마이닝이 많이 적용되는 두가지 경우

<표> 데이터 마이닝에 사용되는 다양한 기법

|              |   |
|--------------|---|
| ●회귀분석        | 오차를 최소화하는 독립변수에 대한 가중치를 산출하여 최적의 예측치를 도출하는 통계적 기법 |
| ●인공신경망       | 신경망의 학습 방식을 모사하여 자료의 패턴을 학습하고 해를 도출하는 기법          |
| ●의사결정나무      | 나무 형태의 분류 기준을 귀납적으로 생성하는 기법                       |
| ●베이지안망       | 베이지안 이론을 바탕으로 망형의 모형을 구성하여 해결책을 도출하는 기법           |
| ●사례기반 추론     | 과거의 유사한 사례로부터 새로운 문제에 대한 해결책을 결정하는 기법             |
| ●유사성에 의한 군집화 | 속성간의 유사성을 기준으로 유사성이 높은 일련의 군집으로 개체들을 집단체화시키는 기법   |
| ●동시적 연관성 분석  | 개체들간에 어떤 사건에 동시에 관련될 가능성이 높은 경우를 파악하는 기법          |
| ●순차적 연관성 분석  | 사건들간의 일반적인 발생 순서를 파악하는 기법                         |

를 살펴 보기로 한다. 우선 대표적인 CRM 주제 중 하나인 교차 판매(Cross Selling)는 한 제품을 구입한 고객이 다른 제품을 추가로 구입할 수 있도록 유도하는 것을 목적으로 한다. 이로부터 기존 고객이 보다 수익성 있는 제품을 구매하도록 하거나 고객의 거래 폭을 넓혀서 이탈을 어렵도록 하여 충성도 향상의 효과로도 연결된다.

교차 판매는 한 구매가 이루어진 이후의 시점에서 동일 고객의 다음 구매에 대한 제안 내용을 결정하는 문제이다. 따라서 각 제품별 구매 경향 모델을 먼저 개발한 후 고객별로 이 모델을 통합하는 방식을 사용한다. 통합된 모델로부터 가장 높은 점수를 받은 제품이 각 고객에 대한 최적의 다음 제안 대상이 되는 것이다.

그러나 실제로 모델링 하는 과정에서는 고려하여야 할 사항이 많다. 제품의 종류가 많고 서로간의 특성에 차이가 큰 백화점과 같은 경우에는 이런 방식을 그대로 적용하는 데는 어려움이 존재한다. 또 모든 모델이 비교 가능해야만 통합이 가능하다는 점도 고려되어야 한다. 특히 고객입장에서 구매할 필요가 없는 경우, 즉 이미 제품을 가지고 있는 경우이거나, 오른손잡이 제품을 왼손잡이에게 권유하는 경우에 대해서는 결과 점수로 0이 산출될 수 있도록 모델링 되어야 하며 점수가 제품의 대중성을 반영할 수 있어야 한다.

또 시간이 지남에 따라서 고객의 특성이 변화하는 것을 반영하여 모델링 해야 한다. 예를 들어 이탈한 고객은 최근 2년간 구매가 남보다 적다. 그렇다고 최근 2년간 구매가 남보다 적은 사람을 이탈할 것으로 보는 것이 맞겠는가. 이는 이탈 고객이 이탈하기 이전 수준을 보아야 할 것이다.

교차 판매를 위한 적정 제품 모델링에서는 제품의 고객에 대한 구매 가능성과 수익성과 고객의 구매 후 만족 가능성이 모두 고려되어야 할 것이다. 일단 한번 구매한다 해도 다음 구매에서는 같은 제품을 다시 구매하지 않게 된다면 효과는 그다지 크지 않기 때문이다. 또 이 세가지 서로 다른 목표들은 상호 갈등 상황에 놓일 수 있기에 어느 것이 상대적으로 더 중요한가를 결정하는 것도 필요하다.

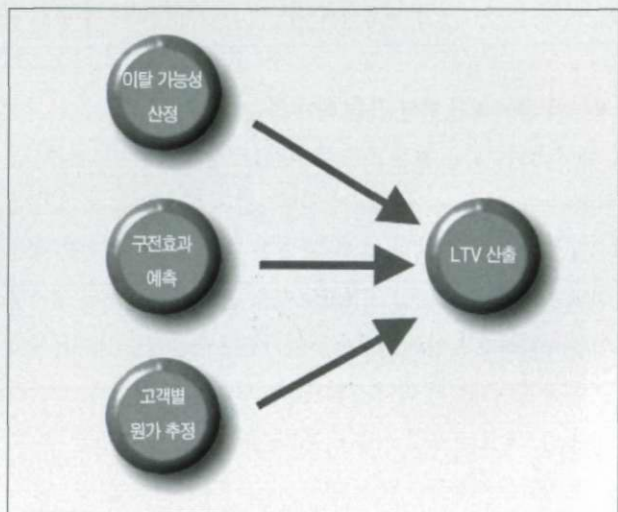
### ■ 고객생애가치의 산출

고객생애가치(Lifetime Value: LTV)는 한 고객이 고객으로서 기업에게 제공할 전 생애에 걸친 이익 금액이다. 이 때문에 CRM의 궁극적인 목표는 각 고객의 LTV를 최대화 시키는 것이다. LTV는 고객을 분류하거나 분석하기 위한 가장 중요한 지표

인 동시에 고객과 기업간의 관계 수준을 표현하는 수치이다.

그러나 이를 산출하는 과정은 말로 정의하는 것처럼 그리 간단하지 않다. 이 때문에 LTV 산출 과정은 데이터 마이닝이나 OLAP 등의 데이터 분석기술을 기반으로 수행된다. LTV를 산정하기 위해서는 각 고객이 얼마나 오랜 동안 기업과의 관계를 유지할 것인지를 파악하여야 하며, 그 기간 동안 얼마나 많은 거래를 할 것이며 그로부터 발생하는 원가가 얼마일 것인가가 산정되어야 한다.

추가적으로 고객의 타 고객에 대한 구전 효과도 LTV에 포함시키는 경우가 있다. 원가의 산정을 위해서는 다양한 수준의 원가 배분 모델이 활용될 수 있다. 이러한 각각의 구성 요소들의 값을 결정하기 위해서는 미래의 고객 행동 및 환경변화에 대한 예측이 필요하며, 집단 단위의 측정치를 개별적인 고객에 배부하기 위한 추정이 필요하다. 이러한 예측 및 추정을 위한 규칙 또는 모델들이 데이터 마이닝이나 OLAP을 활용하여 개발되는 것이다.



(그림 3) 고객생애가치 산출 모델의 예

### ■ 데이터 마이닝에 대한 주의 사항

데이터 마이닝은 기존의 보고서에 대한 분석 중심의 업무 수행 과정을 획기적으로 변화 시킬 수 있을 만큼 큰 잠재력을 가지고 있다. 그러나 매우 정교하게 잘못된 모델링을 빠른 속도로 수행할 수 있게 된다면 복구할 수 없는 곤경에 빠질 수도 있는 위험이 존재한다. 결국 데이터 마이닝 과정에서 내려지는 결론에 대한 정확한 해석을 위해 항상 노력하여야 하며 줄속으로 프로젝트를 수행하는 것은 피하는 것이 바람직하다.

데이터 마이닝은 한번에 결론을 내릴 수 있을 만큼 간단하지 않으며 어차피 지속적인 학습 과정임을 이해하여야 한다. 데이터

마이닝에서 다루는 문제들은 대개 매우 심각한 문제들이며 많은 업무적인 상황에 대한 판단을 필요로 한다. 따라서 어느 기술자나 전문가에게도 전적으로 대신 해달라고 할 수 없음을 인식해야 한다.

CRM과 같이 전략적이며 다양한 변수를 가지고 있고, 또 새로운 업무의 경우에는 더욱더 업무에 대한 이해의 중요성이 크며 문제에 대한 현장 경험을 많이 필요로 한다. 기타 데이터 마이닝 프로젝트 현장에 참여한 전문가들의 의견들을 종합한 결과 중 일부를 제시하면 다음과 같다.

- 데이터 마이닝을 계획하고 있는 대부분의 사람들이 지나치게 낙관적인 기대를 하고 있다. 공급자가 마케팅 과정에서의 과장시킨 메시지를 그대로 믿는 경우가 많다.
- 업무정의가 분명하게 하지도 않은 채로 데이터 마이닝을 시작하려고 한다. 문제가 무엇인지에 대해 명확하고 구체적인 정의가 이루어져야만 데이터 마이닝 성공이 가능하다.
- 도구에 지나치게 의존하는 경향이 있다.
- 다양한 기술에 대한 숙련도를 필요로 하기 때문에 처음부터 데이터 마이닝을 스스로 하려는 것은 바람직하지 않다. 시간이 지나서 기술력이 축적되었다면 당연히 내부에서 독자적으로 수행하는 것이 바람직한 것은 사실이다.
- 한번하고 끝날 수가 없다. 주제도 변동되고 데이터도 변동되고 주변 환경도 변동 된다. 또 한번에 완벽한 답을 얻을 수도 없다. 결국 지속적으로 데이터 마이닝이 이루어져야 한다.

### 3. 가장 중요한 이슈: 데이터 품질의 확보

데이터 분석의 전제는 기본적으로 분석에 필요한 데이터가 충분히 존재하며 그 데이터들이 정확하다는 것이다. 그러나 실제 CRM 프로젝트 상황에서 데이터 웨어하우스를 구축하거나 데이터 마이닝을 위한 데이터베이스를 준비하다 보면 정확하지 못하거나, 수집된 시점이 오래되었기 때문에 현실을 반영하지 못하거나, 심지어는 필요로 하는 항목이 존재하지도 않거나 아니면 단지 항목만 존재할 뿐 거의 입력되지 않은 경우를 흔히 만나게 된다.

데이터가 정확하지 못하면 전체의 성향을 파악하기 어려울 뿐 아니라 개별 고객에 대한 확실한 조치 방안을 수립할 수 없다. OLAP이든 데이터 마이닝이든 분석결과가 전혀 의미 없는 경우가 발생된다. 비교적 많은 경험과 사례를 가지고 있는 미국의 경우에도 최초의 데이터 마이닝 프로젝트에서 한 문제에 대한 답을

내는데 짧아도 수개월이 걸린다고 한다. 이와 같이 오랜 기간이 걸리는 이유는 데이터 분석과정에 많은 시간이 소요되기 때문이라기 보다는 분석에 적합한 데이터 준비를 수행하는 과정이 길었기 때문이다.

결국 데이터 분석 이전에 데이터의 품질을 확보하는 작업이 선행되어야 한다. 실제로 이 과정은 별개의 작업이라기 보다는 데이터 분석이나 통합된 데이터베이스 구조 설계 작업과 밀접하게 연관되어 이루어져야 하는 부분이다. 품질이 낮은 두개의 항목이 존재하더라도 어느 쪽이 분석가치가 높은가에 따라 우선적으로 정제하는 작업을 수행해야 할 것이기 때문이다. 데이터의 정제 역시 우선 순위와 투자수익성에 대한 구체적인 고려를 바탕으로 이루어져야 한다. 하자가 있는 모든 데이터를 완벽하게 정제하려면 CRM을 통한 이익을 모두 투입해도 부족할 것이다.

방법적인 측면에서 보면 모든 데이터를 사람의 눈으로 전부 확인 및 수정할 수 없기 때문에 인공지능 방법을 이용한 데이터 정제도 적용하는 추세이다. 이를 위한 상용의 도구들도 해외에서는 출시되어 있으며 국내에서도 특정 항목, 예를 들어 고객의 주소를 비교하여 중복 고객을 선별하는 등의 기능을 가진 도구들이 출시되고 있다. 이러한 도구가 모든 것을 자동화 하지는 못하지만 데이터 품질 개선 과정의 효율성을 크게 향상시켜 주므로 자동화된 도구를 구입하거나 개발하는 것도 고려할 필요가 있다.

### 4. CRM에서의 데이터 분석: 끝없는 도전

고객은 너무도 많은 설명할 수 없는 특성을 가지고 있다. 그러나 고객이라는 대상을 이해하는 것이 CRM의 가장 큰 숙제이기 때문에 CRM에서는 데이터 분석 기법의 활용이 매우 중요한 성공 요인 중 하나임을 부정할 수 없다. 그러나 항상 분석후의 업무 적용 가능성에 대한 고려가 선행되어야 한다.

데이터 분석을 고려하기 위해서는 적절한 양과 질의 데이터가 존재하는 것이 전제되어야 한다. 이 때문에 데이터 웨어하우스를 구축하여 공통적으로 사용될 데이터의 기반을 미리 확보해 두려는 접근 방법이 타당한 것이다. 데이터 마이닝에는 다양한 기술이 필요하다. 처음부터 무리한 분석을 고려하는 것은 바람직하지 않지만 점진적으로 배워나가기려는 자세는 필요하다.

그러나 아무런 준비를 실제로 해나가지 않는다면 영원히 시작하지 못할 것이다. 끝없는 도전만이 누구나 시작할 수 있는 CRM에서 진정한 경쟁력을 얻을 수 있는 열쇠일 것이다. 