2005 한국경영과학회/대한산업공학회 춘계공동학술대회
2005년 5월 13일~14일, 충북대학교

고객평생가치 모델의 통신산업 응용에 관한 연구

A Study on practice of customer lifetime value models in the telecommunication industry

황성섭, 조성준
서울대학교 산업공학과

Abstract

고객관리란 고객데이터와 정보를 분석, 통합하여 개별 고객의 특성에 기초한 마케팅 활동을 계획, 지원, 평가하는 과정이다. 고객 관리는 기업에 있어서 가장 중요한 전략 중의 하나이며, 1990년대 이후로 그 중요성은 강조되어 왔다. 이러한 고객관리의 출발은 고객의 가치를 정확히 측정하는 것이며, 고객평생가치를 판단하기 위한 많은 연구가 이루어졌다. 우리는 고객평생가치 모델을 한국의 이동통신 산업에 적용하기 위한 이론적 배경과 그 방법론을 분석하였다. 또한 이 논문은 이러한 과정에서 나타나는 문제점의 원인과 그 해결방안을 제시하고 있다. 이러한 연구와 방법론들은 실제 마케팅 전략에 직접적으로 활용될 수 있을 것이다.

1. 서론

고객관리란 우리 회사의 고객이 누구인지, 고객이 무엇을 원하는지를 파악하여 고객이 원하는 제품과 서비스를 지속적으로 제공함으로써 고객을 오래 유지시키고 이를 통하여 고객의 평생가치를 극대화하여 수익성을 높이는 통합된 관리 프로세스를 말한다. 즉 고객데이터와 정보를 분석, 통합하여 개별 고객의 특성에 기초한 마케팅 활동을 계획, 지원, 평가하는 과정이다. 1990년대 이후로 고객관리 개념의 중요성은 강조되어 왔으며, 이것은 기업에 있어서 가장 중요한 전략 중의 하나이다.

이러한 고객관리의 출발은 고객의 가치를 정확하게 측정하는 것이며, 고객의 가치와 가장 이익을 줄 수 있는 고객을 이해하는 것은 고객 유지의 필수적인 요소라고 할 수 있다. 특히, 이동통신 산업과 같이 경쟁이 치열하고 고객 이동이 심한 산업에서는 더더욱 그 중요성이 크다고 할 수 있을 것이다[2].

고객평생가치를 측정하기 위한 많은 연구가 이루어졌다. 우리는 고객평생가치 모델을 한국의 이동통신산업에 적용하기 위한 이론적 배경과 그 방법론을 분석하였다. 이 논문은 이러한 과정에서 나타나는 문제점의 원인과 그 해결방안을 제시하고 있으며, 논문을 통하여 도출된 전략들은 실제 마케팅에 직접적으로 활용될 수 있을 것이다.

기업은 고객 평생가치를 이용하여 고객군의 순위를 매기고, 각각의 세그먼트에 알맞은 마케팅 전략을 구사할 수 있다. 고객평생가치가 높은 세그먼트에 대해서는 각종 할인과 인센티브를 제공하여 고객유지를 위하여 힘써야 할 것이다. 반대로 고객평생가치가 낮거나 이탈확률이 높은 고객군에 대해서는 그 원인을 파악하여 그들이 높은 가치를 형성할 수 있도록 노력을 경주해야 할 것이다. 또한, 각 고객군의 특성을 파악하여 해당 고객군에 대하여 Cross-selling과 Up-selling을 실시할 수 있을 것이다. 또한, 특정 캠페인 실시전과 실시 후의 고객평생가치 측정을 통하여, 그 캠페인이 성공 여부를 판단하는 척도를 제공할 수 있다. 캠페인 실시 후의 고객평생가치는 캠페인 실시 후의 고객평생가치에서 캠페인 비용을 더한 것보다 커야할 것이다. 이러한 개념을 확장하여 누적된 고객의 평생가치에 기초하여 미래 특정 시점의 사업 및 고객의 규모를 예측하거나 계획할 수 있다. 이것은 기업의 설비 및 다른 투자를 위하여 좋은 지침으로서 활용될 수 있을 뿐만 아니라, 경영자의 전략수립에도 큰 도움을 줄 것으로 예상된다.

이 논문은 다음과 같이 구성되어 있다. 우리는 2장에서 고객평생가치의 정의와 그 개념모델에 대하여 살펴볼 것이다. 이러한 접근의 핵심요소인 고객생존분석의 방법론을 3장에서, 신경망을 이용한 생존분석 연구를 4장에서 소개할 것이다. 5장에서는 고객평생가치의 통신산업 응용으로서 세그먼트 기반의 기대서비스기간, 그 수행모델 및 실험방법론에 대하여 기술할 것이다. 마지막으로 6장에서 이러한 방법론의 한계와 문제점을 논하게 될 것이다.

2. 고객평생가치 (LTV)

고객평생가치는 LTV, Customer Lifetime Value, Customer Equity, Customer Profitability 등의 이름으로 논의되고 있다. 이러한 고객평생가치는 여러 학자들에 의하여 다음과 같이 정의되고 있다.

- 고객으로부터 생성되는 미래수익의 현재가치. (Gupta and Lehmann[3], 2003)
- 고객의 트랜잭션 전체기간동안 고객으로부터 발생한 순이익 또는 손실. (Berger and Nasr[4], 1998)
- 고객관리에 관련된 비용을 제외한 고객으로부터의 기대수익. (Blattberg and Deighton[5], 1996)
- 고객이 생성하는 할인된 전체 순이익. (Bitran and Mondscein[6], 1996)
- 고객이 수익에 기여하는 스트림의 순현재가치. (Pearson[7], 1996)
- 고객으로부터 기대되는 비용과 수익에 공헌하는 미래 스트림의 순현재 가치. (Jackson[8], 1994)

▪비용과 수익에 대한 모든 미래 기여의 순현재 가치. (Roberts and Berger[9], 1994)
▪고객으로부터 기대되는 수익과 비용에 대한 모든 미래 기여의 순현재가치. (Courtheoux[10], 1995)
▪트랜잭션 기간동안 고객으로부터 거둬들인 수입에서 고객유치, 판매, 서비스 등의 전체 비용을 차감한 이익의 합. (Hwang, Jung, Suh[2], 2004)

이상의 정의들을 종합해보면 고객평생가치는 "고객의 유지 기간동안 고객으로부터 얻은 수입에서 고객관리에 소요되는 비용을 차감한 순이익의 현재가치"로서 정의할 수 있다.

고객평생가치는 고객의 가치, 기대 서비스기간, 할인인자로서 구성된다. Rosset[1]은 고객평생가치에 대한 모델을 구성함에 있어서 다음과 같은 개념적 접근을 제시하였다.

$$LTV = \int_0^\infty S(t)v(t)D(t)dt \qquad (2.1)$$

각 구성요소에 대하여 살펴보면 다음과 같다. 먼저 t시점에서의 고객의 가치를 나타내는 $v(t)$는 $t \geq 0$의 경우를 나타내며, 현재 시점의 경우는 $v(0)$로 나타낼 수 있다. 고객의 가치는 현재의 데이터로부터 추정한다.

기대서비스 기간(length of service, 이하 LOS)은 고객의 이탈확률로서 나타낼 수 있다. 식 (2.1)에 나타난 $S(t)$는 생존함수이며, t시점에 고객이 회사의 서비스를 유지하고 있는지를 나타낸다. 특정 시점의 순간적인 생존함수 $f(t)$는 $-dS/dt$로 나타낼 수 있으며, hazard 함수 $h(t)$는 $f(t)/S(t)$이다. Helsen과 Schmittlein[11]는 $h(t)$가 왜 $f(t)$에 비하여 적절한지에 대하여 설명했다. LOS는 현재와 과거 데이터를 이용하여 추정할 수 있다.

할인 인자 $D(t)$는 미래의 t시점에서 가치를 현재가치로 할인하기 위한 요소이다. 대표적으로 두 가지 접근 방법이 사용되는데, 그것은 Exponential decay와 Threshold function이다.

3. 생존분석 (Survival Analysis)

생존분석이란 한 개체의 생명과 관련된 자료를 다루는 의학통계의 분야이다. 이것은 생명보험과 같은 금융분야, 신뢰성 분석을 위한 제조분야 등 많은 영역에서 응용되고 있다. 생존분석에서는 생존자료의 특성상 완전한 관측값은 불가능하다. 한 개체가 사망할 때까지 실험기간을 무한정 연장할 수도 없으며, 실험기간 도중 여러 가지 이유로 실험에서 빠져나가는 개체가 발생할 수 있다. 이러한 불완전 관측자료의 분석은 통상적인 완전자료의 분석과 다른 통계적 방법을 필요로 한다. 관찰개체의 생존여부를 통한 관측의 분류는 다음과 같다.

◎ Uncensored case
・Complete Observation = Death
◎ Censored case=Incomplete Observation
・type 1 censoring : survival (cut-off)
・type 2 censoring : follow-up loss

누적 생존율을 산출하는 방법은 크게 직접법과 누적법으로 구분할 수 있다. 직접법은 추적관찰기간이 대부분 5년 또는 10년을 넘는 집단으로부터 그들의 생존율을 직접 산출하는 방법이다. 이동통신 산업과 같이 역사가 짧고 변화가 심한 특성을 가지는 경우 장기간의 데이터를 이용하여 생존율을 구하는 것이 불가능하다. 이러한 이유로 누적 생존율을 산출하기 위하여 주로 누적법을 사용하게 된다. 누적법은 분포를 가정하는지, 모형에 근거하는지에 따라서 모수적 방법, 비모수적 방법, 준모수적 방법으로 분류할 수 있다. 그 구체적인 내용은 다음과 같다.

3.1 모수적 방법 (Parametric method)

대표적인 모수적 방법은 지수분포 또는 와이블 분포를 이용하는 것이다. 지수분포는 가장 단순한 형태의 생존분포이다. 이는 지수분포의 위험함수가 시간에 따라 변하지 않는 상수 값을 가지기 때문이다. 예를 들어 가입 후 t_1시간을 생존해 있었고, 또한 t_2시간까지 생존할 확률은 지수분포를 하는 것이며, t_1과 t_2가 어떤 시간이든 간에 두 시점 사이의 기간에 의해서만 결정됨을 뜻하는 것이다. 그러나 실제로 고객의 생존은 이처럼 단순하지는 않기 때문에 지수분포의 유용성은 한정되어 있다. 와이블분포의 위험함수는 여러 형태를 포함하기 때문에, 지수분포와는 달리 널리 사용된다. 와이블분포는 비교적 단순한 형태를 지니면서도 여러 가지 증가 및 감소의 경우를 표현할 수 있어서 생존분포로 유용하게 사용될 수 있다.

3.2 비모수적 방법 (Non-parametric method)

대표적인 비모수적 생존분석의 방법은 생명표법, Kaplan-Meier법, 신경망을 이용한 방법이 있다. 먼저, 생명표(actuarial)법은 원래 생명보험회사에서 이용하던 방법이다. 관찰기간을 일정단위로 나누어 각 구간마다의 구간생존율을 구하고 이들의 누적으로 일정기간까지의 누적생존율을 구하는 것이다. 연구대상의 크기가 최소한 50표본은 넘어야 되고 관찰단위당 10표본 이상 되는 것이 좋다고 할 수 있다.

Kaplan-Meier법 (product-limit법)은 표본의 크기가 50 이하인 경우에 적용하는 방법으로, product-limit법이라는 어의에서 알 수 있듯이 일정한 간격의 구간생존율을 구하는 것이 아니라 각 사망이 일어난 시점에서의 생존율을 구하고 이들의 누적으로 누적 생존율을 산출하는 방법이다. 이해하기가 쉽고 중도탈락이나 관찰중단 예에 대한 취급이 간단하여 널리 적용되고 있다.

생존분석을 위하여 가장 활발한 연구가 이루어진 비모수적 접근법은 신경망(Neural Network)을 이용한 방법이라고 할 수 있다. 신경망은 인지 구조나 뇌의 학습 과정을 모형화한 분석 기법이다. 이

기법은 과거의 자료로부터 학습이라는 과정을 거쳐서 새로운 관찰값들의 예측을 가능하게 한다. 즉, 과거 고객의 데이터를 통하여 이 고객의 생존 여부를 분석하게 되는 것이다. 신경망을 이용한 생존분석[14]은 4장에서 자세히 살펴보게 될 것이다.

비모수적 생존곡선의 비교는 가설을 검정하는 단계라고 할 수 있다. 생명표법으로 작성된 누적 생존율을 비교하는데는 Mantel- Haenszel법이 주로 사용된다. Kaplan-Meier법에 의한 누적 생존율의 비교에는 두 가지 방법이 가능한데 생존기간이 긴 자료에서는 log-rank법, 생존기간이 짧은 자료에서는 Gehan's generalized Wilcoxon법이 주로 이용된다.

3.3 준모수적 방법 (Semi-parametric method)

생존분석은 여러 가지 혼란변수를 통제한 상태에서 집단들간의 생존확률을 비교하여야 하는 경우가 있으며, 또한 여러 변수들이 동시에 생존시간에 미치는 영향을 알아보고자 하는 경우가 많다. 이와 같이 생존과 관련된 여러 예후변수들의 영향을 동시에 알아보는 다변량분석법으로 Cox가 제안한 비례위험 회귀모형(proportional hazard function)을 사용한다. 이 모형은 생존시간에 대해 어떠한 분포형태도 가정하지 않으므로 비모수적이지만, 모형에 근거하여 회귀계수를 추정한다는 점이 모수적 방법과 유사하여 비모수와 모수의 중간형태인 준모수적 방법이라 일컬어진다.

Cox의 비례위험모형[12]은 모든 고객들이 공통적이고 전체적인 baseline hazard function을 가진다는 개념에서 출발하는 것이다. 먼저, $h_i(t)$를 t시점에서의 위험함수로 정의하고, $\lambda_0(t)$를 모든 독립변수의 값이 0인 경우의 기저 위험함수라 정의하면 회귀모형은 다음과 같다.

$$h_i(t) = \lambda_0(t)\exp\left\{\sum_{c=1}^{C} \beta_c x_{ic}\right\} \qquad (3.1)$$

i: customer i $(i = 1, 2, \cdots, n)$
c: C covariates $(c = 1, 2, \cdots, C)$
$\lambda_0(t)$: underlying hazard function
β_c: parameters estimated during the PH regression

PH coefficient는 logistic regression의 형태를 통하여 추정될 수 있다. 이것은 Complementary-log-log (CLL) model[13]으로 일컬어진다.

$$\log\left[-\log\left\{1 - h_i(t)\right\}\right] = \alpha_t + \sum_{c=1}^{C} \beta_c x_{ic} \qquad (3.2)$$

β_c: coefficients
α_t: baseline age effect.

위의 α_t는 $\lambda_0(t) = 1 - \exp(-\exp(\alpha_t))$를 통하여 baseline hazard component로 변환될 수 있을 것이다.

4. 신경망을 이용한 생존분석

신경망은 데이터마이닝의 한 기법이다. 신경망 기법의 첫 번째 단계는 레이어(layer)와 뉴런(neuron)으로 구성된 특정한 네트워크 구조를 설계하는 것이다. 이렇게 형성된 네트워크는 학습(training) 과정에 이용된다. 이 단계에서 뉴런은 학습에 이용되는 자료를 최적으로 예측하기 위해 네트워크의 가중치를 조정할 수 있도록 입력 자료에 대해 반복적인 과정을 적용한다. 입력 자료로부터 학습하는 단계가 끝나면, 네트워크의 구성이 완료되고, 이를 통해 예측치들을 만들 수 있다. 그러나 전통적인 모형과는 달리, 네트워크에서 이러한 관계들은 변수들간의 관계를 기술하는 통계학이나 방법론에서 사용되는 일반적인 용어로 표현될 수 없다. 이러한 접근 방법은 응용분야에만 관심을 두고 있으며, 숨겨진 메커니즘의 특성이나 현상의 이론적 측면은 고려하지 않는 방법이다. 이러한 신경망을 이용하여 생존을 예측한 기존 연구들을 살펴보면 다음과 같다.

가장 간단한 방법은 특정기간에 대한 생존여부를 검토하는 것[23]이다. 결과적으로 이분류 문제가 되는 것이다. neural network output은 고객이 특정기간에 생존할 확률의 추정치를 제공한다. 50%의 threshold 이상일 때, 고객은 그 기간에 생존할 것이라고 가정되었다. 이러한 접근이 기본적인 것은 명확하지만, 개별적인 생존 또는 위험함수를 산출할 수 없으며, Censoring case를 다룰 수 없다.

Ohno-Machado[15]는 AIDS 환자의 생존분석 문제를 분석하기 위하여 multiple neural network를 이용하였다. 이것은 특정 시점에서 생존을 예측하기 위한 single output을 가진 neural network이다. Censored Observation 데이터는 censoring 시점까지 포함했기 때문에 학습데이터수가 점차적으로 감소하여 예측의 신뢰성을 저하시킨다. 저자들은 neural network를 결합하는 것이 비단조적 곡선의 출현을 감소시킬 수 있는 방법이라고 주장하였다. 즉, 한 neural network의 생존예측은 다른 network의 input으로 사용하게 되는 것이다. 이 연구의 한계는 먼저 단조성으로부터의 이탈이 감소된다고 하더라도 여전히 비단조적인 생존함수를 얻을 수 있다는 점이다. 또한, Neural network를 어떻게 결합하느냐의 문제와 대형의 데이터 셋을 다루기 위하여 부적절한 확장성의 문제이다.

Ravdin과 Clark[16]은 재발(relapse) 또는 죽음에 이르는 시간을 예측하기 위한 생존분석의 수행을 위하여 neural network 기법을 이용하였다. 이 논문에서, 저자들은 neural network이 복잡한 상호작용이 존재하는 데이터 셋에서 그것을 탐지하고, 예측모델을 생성하는데 기여할 수 있을 것이라고 주장하고 있다. 저자들은 생존상태를 나타내는 single output unit을 가진 multi-layer feed-forward neural network를 이용하였다. Time indicator와 survival status indicator가 레코드에 더해졌다. Time indicator는 예측을 위하여 시간 $[1, T_{\max}]$을 기록한다. 완전 관측된 input은 $T_{\max}$까지 replication된 반면, 중도절단 관측된 input은 중도 절단된 시점까지 replication되었다. Survival status는 network의 target이며, 고객이 살아있는 한 0으로, 그렇지 않으

면 1로 설정되었다. 그러나 그들의 연구는 일반화된 생존함수가 단조적으로 감소한다는 어떤 보장도 하지 못했다. 게다가 레코드의 replication는 두 가지 문제를 만들게 되었다. 첫째, late time에서 사망의 수가 과잉대표되었기 때문에, 많은 편의(bias)를 이끌 수 있을 것이다. 둘째, 이 연구는 대형 데이터 셋에 의한 심각한 확장성의 문제를 야기한다.

Ravdin과 Clark의 연구의 변형은 Biganzoli[17]에 의해 제안되었다. Neural network은 유방암의 재발과 사망의 위험에 대한 환자들의 상태예측을 위하여 사용되었다. 그들은 하나의 output과 추가로 time indicator input을 가진 neural network를 학습하였다. 이 논문은 Ravdin과 Clark[16]의 연구를 확장하고, 예후 변수의 하나로서 시간을 코딩함으로써 neural network가 환자의 상태를 예측하기 위하여 censored survival data를 사용할 수 있다는 것을 보여주었다. Neural network는 생존확률을 단조적으로 변환시킬 수 있는 이산적 위험함수를 예측하고, Cox 회귀모형만큼 정확한 결과를 예측하였다. 그러나 많은 데이터 replication요구 때문에 확장성이 부족한 것이 한계이다.

Lapuerta et al.[18]은 censoring case에 대하여 생존 기간을 귀속시키기 위하여 multi-network 전략을 제안하였다. neural network는 고려기간에 대하여 생존 상태가 알려진 관찰만을 이용하여 학습되었다. 따라서 학습된 네트워크는 censoring case에 대한 결과를 예측하기 위하여 사용되었다. 비록 제안된 방법이 Cox 비례위험모형에 견주어 떨어지지 않는다고 할지라도, 유도된 생존확률이 단조적으로 감소한다는 어떠한 보장도 없다. 게다가, 고려되는 시간만큼 많은 neural network을 학습시켜야 하기 때문에 이러한 접근이 대규모의 어플리케이션에는 적절하지 않다.

Faraggi[19]는 전립선 암을 가진 남자들의 생존에 대한 데이터를 이용하여 Cox 비례위험모델의 neural network 확장을 제안하였다. 이 모델은 전통적인 비례위험모형의 모든 이점을 유지하는 것을 가능하게 한지만, 여전히 hazard가 비례한다는 것을 가정한다.

Street[20]는 생존분석문제에 착수하기 위하여 maximum time horizon $T_{\max}$ output unit을 가진 multilayer perceptron을 이용하였다. 모든 output neurons이 -1과 +1 사이의 값을 취하는 output layer에서 hyperbolic tangent activation function이 사용되었다. 0보다 작은 값을 가지는 첫 번째 output neuron은 event time을 예측하는 output neuron이 되는 것이며, 모든 output neuron이 0보다 큰 값을 가진다면 환자는 전체 연구기간에 생존하는 것으로 간주되었다. 따라서 output unit은 기간에 대응하는 생존확률을 나타낸다. 요약하면, training set observation의 output은 다음과 같이 encode되었다.

$$S(t)=\begin{cases} 1 \;\; 1 \leqq t \leqq L \\ -1 \;\; D=1 \text{ and } L<t \leqq T_{\max} \\ S(t-1)\times(1-h(t)) \;\; D=0 \; and \\ \qquad L<t \leqq T_{\max} \end{cases} \qquad (4.2)$$

여기서, $T_{\max}$는 연구와 관련된 최대고려기간을 나타내고, L은 censoring time이며, D는 subject가 중도단절 관측되거나($D=0$) 또는 그렇지 않은 경우($D=1$)를 가리킨다. Neural network는 단조적으로 감소하는 output units를 생성하지 않기 때문에, 여전히 비단조적인 생존곡선이 가능하며 이것은 해석을 어렵게 한다.

Street[20]의 방법에 대한 변형은 Mani[21]에 의해 발전되었다. 이 연구에서 outputs는 다음과 같이 계산되었다.

$$h(t)=\begin{cases} 0 \;\; 1 \leqq t \leqq L \\ 1 \;\; D=1 \text{ and } L<t \leqq T_{\max} \\ \dfrac{d_t}{n_t} \;\; D=0 \text{ and } L<t \leqq T_{\max} \end{cases} \qquad (4.3)$$

여기서, $T_{\max}$는 최대고려기간을 나타내고, L은 censoring time이며, D는 subject가 중도단절 관측되거나($D=0$) 또는 그렇지 않은 경우($D=1$)를 가리킨다. 완전관측에 대하여, hazard는 사망시점까지는 0, 그 이후로는 1로 설정되었다. 중도단절 관측에 대하여 censoring 시점까지는 0, 그 이후로는 Kaplan-Meier 추정치가 설정되었다. 생존확률은 $S(t)=S(t-1)\times(1-h(t))$을 이용함으로써 추정될 수 있다. 생성된 생존곡선은 단조적으로 감소하며, 이것은 해석을 간단하게 하고 robustness를 증가시킨다.

Mani[21]와 비슷하게, Brown[22]은 hazard rate를 예측하기 위하여 multiple outputs를 가진 single neural network를 제안하였다. 완전관측에 대하여, network output는 subject가 살아있는 경우 0으로, event를 겪었을 때 1로 설정되었다. Event 다음의 기간에 대하여, hazard는 제한되지 않았다. Censored observation에 대한 output value는 censoring time까지는 0으로 설정되고, 모든 순차 시간격에 대하여 제한되지 않았다. 저자들은 hazard가 대응하는 에러가 0으로 제한되지 않았을 때, 에러 제곱합을 최소화시키고 가중치 업데이트가 수행되지 않기 위하여 neural network를 학습시키는 것을 제안하였다. 저자들이 제시한 이러한 접근은 확장성이 있으며, 단조적인 생존함수를 이끈다.

5. 통신산업 응용

5.1 고객데이터

이 연구에서 분석한 데이터는 한국의 한 이동통신 회사의 고객 데이터이다. [표 5.1]은 고객 데이터의 구성정보와 그 대표적인 데이터의 속성을 나타내고 있다.

구분	속성
식별정보	· 서비스관리번호
인적정보	· 성별코드 · 연령코드
가입정보	· 서비스 이용종류 코드 · 단말기 모델코드 · 가입연월일 · 서비스 가입해지 여부 · 서비스 해지일자 · 해지연월 · 개인법인구분
사용실적	· 서비스 상태코드 · 선택요금제 코드 · 단말기 획득일자 · 단말기 획득연월 · 청구금액 · 과거 3개월 평균청구금액 · 과거 6개월 평균청구금액 · 수납금액

[표 5.1] 데이터 구성

이러한 데이터를 이용하여 이동통신서비스 고객의 기대수명을 예측하고, 이탈 성향이 비슷할 것으로 예측되는 고객들을 같은 세그먼트로 분류할 수 있을 것이다. 이러한 과정을 통하여 각 세그먼트의 가치를 분석하고 바람직한 마케팅 전략을 수립할 수 있을 것이다.

5.2 세그먼트 기반 LOS

고객평생가치를 측정하기 위한 이론적인 접근방법은 실제적으로 많은 제약이 따른다. 각 고객에 대하여 각각 모든 구성요소들을 측정하는 것은 많은 시간과 비용이 소요되며, 그러한 접근법이 반드시 더 정확한 결과를 도출한다고 보기 어렵다. 이러한 이유로 성향이 비슷한 고객을 군집화하고, 그들의 대표 특성을 반영할 수 있는 모델을 생성하는 것이 합리적이라고 할 수 있다.

2장에서 제시한 고객평생가치 모델을 세그먼트 기반으로 수행하기 위하여 hazard 함수 $h(t)$와 생존함수 $S(t)$가 구체적으로 정의되어야 할 것이다. 이에 Rosset[1]는 아래와 같은 접근방법을 제시하였다. 먼저 harzard 함수 $h(t)$는 다음과 같이 정의할 수 있다.

$$h(t) = \frac{\sum_i I(t_i = t) I(c_i = 1)}{\sum_i I(t_i = t)} \qquad (5.1)$$

이것은 t달 동안 서비스를 유지하고 현재 달에 이탈한 고객의 수를 t달 동안 서비스를 유지한 고객의 수로 나눈 값이다. 식(5.1)의 $h(t)$를 통하여 생존함수 $S(t)$는 다음과 같이 정의할 수 있다.

$$\begin{aligned} S(t) &= \prod_{u<t} \frac{S(u+1)}{S(u)} \\ &= \prod_{u<t} \frac{S(u) - f(u)}{S(u)} \qquad (5.2) \\ &= \prod_{u<t} (1 - h(u)) \end{aligned}$$

여기서 $S(0)$는 1과 같다.

위의 정의를 통하여 우리는 개별적인 고객인 아닌 비슷한 성향의 고객들로 구성된 세그먼트 기반의 LOS를 산출할 수 있다.

5.3 수행모델

Non-parametric 접근의 통신산업 응용을 위한 수행모델을 살펴보기로 하자. 먼저 hazard 함수는 다음과 같이 나타낼 수 있다.

$$h(t) = \frac{d_k}{r_k},\ t_k < t < t_{k+1} \qquad (5.3)$$

r_k: k개월 서비스 유지 고객
d_k: k개월 서비스를 유지하고 이탈한 고객

위의 식 (3.1)의 hazard 함수를 이용하여 t시점에서 고객의 생존율은 다음과 같이 나타낼 수 있다.

$$S(t) = \prod_{k:\, t_k < t} (1 - h(t)) \qquad (5.4)$$

세그먼트 기반의 LOS를 통하여 고객평생가치를 구하기 위하여 우리는 Rosset[1]이 제시한 다음과 같은 모델링 과정을 수행한다. 먼저 t달 동안의 서비스 유지에 대한 이탈자의 비율 p_t를 다음과 같이 정의할 수 있다.

$$p_t = \frac{\sum_i I(t_i = t) I(c_i = 1)}{\left(\begin{array}{c} factor \cdot \sum_i I(t_i = t) I(c_i = 0) \\ + \sum_i I(t_i = t) I(c_i = 1) \end{array} \right)} \qquad (5.5)$$

여기서 factor는 샘플에서 이탈 고객의 비율을 전체 고객의 이탈 비율로 나눈 값이다. 이것은 샘플을 이용하여 구한 값을 보정하기 위한 것이다.

식(5.5)에서 정의한 이탈자의 비율 p_t를 통하여 t달 동안 서비스를 유지할 고객의 생존확률 $S(t)$를 다음과 같이 정의할 수 있다.

$$S(t) = (1 - p_{t-1}) \times (1 - p_{t-2}) \times \cdots \times (1 - p_{t_0}) \qquad (5.6)$$

식 (5.6)를 이용하여 우리가 기대 서비스 기간(expected length of survival)은 다음과 같이 나타낼 수 있다.

$$ELOS = \sum_{t=0}^{h} S(t) \qquad (5.7)$$

여기서 horizon h는 관찰하려는 시점까지의 개월수이다.

우리는 ELOS를 산출하고 최종적으로 세그먼트 기반의 접근을 통하여 LTV를 구할 수 있다. 구체적인 모델은 다음과 같다.

$$LTV = ratio \times \sum_{j \in segment} ELOS_j v_j c_j \qquad (5.8)$$

여기서 v_j는 j번째 고객의 가치, c_j는 j번째 고객의 이탈여부이며, ratio는 세그먼트에서 추출한 샘플에 대한 세그먼트의 비율로 정의할 수 있다.

5.4 모델 평가 및 검증

이 논문에서 논의한 모델의 검증은 실험을 통하여 얻은 각 세그먼트의 ELOS와 실제 데이터를 이용하여 구한 ELOS의 비교를 통하여 이루어질 것이다. 또한 제시한 모델의 평가는 실험을 통하여 얻은 2003년의 LTV와 실제 데이터를 이용하여 구한 LTV의 비교를 통하여 이루어질 것이다. 그 정확성은 다음과 같은 척도를 통하여 측정할 수 있다.

$$\text{정확성} = 1 - \left| \frac{\text{추정값} - \text{실제값}}{\text{실제값}} \right| \qquad (5.10)$$

6. 토의

이 논문에서 제시한 방법은 세그먼트 기반의 고객평생가치 접근이다. 이것은 고객의 세그먼트를 어떻게 분할하느냐에 따라서 현저한 성능의 차이를 보일 수 있다. 여러 가지 방법을 통하여 가장 적절한 분할 방법을 찾는 노력이 필요할 것이다. 또한, 같은 세그먼트에 속하는 고객들은 기대되는 서비스 기간이 유사할 것이라는 가정을 내포하고 있다. 그러나 여러 가지 상황에 따라서 각 고객의 서비스 기간은 달라질 수 있을 것이다. 이러한 부분을 해결하기 위해서는 좀 더 세분화된 고객 세분화를 수행하고, 주기적으로 모델을 갱신해야 할 것이다.

이 논문에서 논의한 기대 서비스 기간을 모두 검정하기 위해서는 실험을 통하여 얻은 최대 서비스 기간이 지난 시점에서 가능할 것이다. 즉 3장에서 논의한 censored case에 대한 보완이 필요하다. 결과 분석에서는 기대되는 서비스 기간이 실험에서 고려하는 최대기간 이하로 나온 고객들에 대해서만 실제검증이 이루어질 수 있다. 따라서, 지속적인 검증과 갱신을 통하여 성능을 확인하고 개선해야 할 것이다.

7. 참고문헌

[1] Saharon Rosset, Einat Neumann, Uri Eick, Nurit Vatnik, Customer Lifetime Value Models for Decision Support, Data Mining and Knowledge Discovery, 7, 321-339, 2003.

[2] Hyunseok Hwang, Taesoo Jung, Euiho Suh, An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry, Expert Systems with Applications, 26, 181-188, 2004.

[3] S. Gupta, D. R. Lehmann, Customers as assets. Journal of Interactive Marketing, 17(1), 9-24, 2003.

[4] P. D. Berger, N. I. Nasr, Customer lifetime value: marketing models and applications, Journal of Interactive Marketing, 12(1), 17-30, 1998.

[5] R. C. Blattberg, J. Deighton, Manage marketing by the customer equity test, Harvard Business Review, Jul-Aug, 136-144, 1996.

[6] G. R. Bitran, S. Mondschein, Mailing decisions in the catalog sales industry, Management Science, 42(9), 1364-1381, 1996.

[7] S. Pearson, Building brands directly: creating business value from customer relationships, London: MacMillan Business, 1996.

[8] D. R. Jackson, Strategic application of customer lifetime value in the direct marketing environment. Journal of Targeting Measurement and Analysis for Marketing, 3(1), 9-17, 1994.

[9] M. L. Roberts, P. D. Berger, Direct marketing management, Prentice-Hall: Englewood Cliffs, NJ, 1989.

[10] R. Courtheoux, Customer retention: how much to invest. Research and the Customer Lifecycle, New York, NY: DMA, 1995.

[11] K. Helsen, D. C. Schmittlein, Analyzing duration times in marketing: Evidence for the effectiveness of Hazard rate models, Marketing Science, 11, 395-414, 1993.

[12] D. R. Cox, Regression Models and Life Tables, Journal of the Royal Statistical Society, B34, 187-220, 1972.

[13] R. L. Prentice, L. A. Gloeckler, Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data, Biometrics, 34, 57-67. 1978.

[14] Baesens B.,Van Gestel T., Stepanova M., Vanthienen J., Neural Network Survival Analysis for Personal Loan Data, Proceedings of the Eighth Conference on Credit Scoring and Credit Control (CSCCVII'2003), Edinburgh, Scotland, September, 2003.

[15] L. Ohno-Machado, Sequential use of neural networks for survival prediction in aids, Journal of the American Medical Informatics Association, 3, 170-174, 1996.

[16] M. De Laurentiis, P. Ravdin, Survival analysis of censored data: neural network analysis detection of complex interactions between variables, Breast Cancer Research and Treatment, 32, 113-118, 1994.

[17] E. Biganzoli, P. Boracchi, L. Mariani, E. Marubini, Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach. Statistics in Medicine, 17, 1169-1186, 1998.

[18] P. Lapuerta, S. P. Azen, L. LaBree, Use of neural networks in predicting the risk of

coronary artery disease, Computers and Biomedical Research, 28, 38-52, 1995.
[19] D. Faraggi, R. Simon, A neural network model for survival data, Statistics in Medicine, 14, 73-82, 1995.
[20] W.N. Street, A neural network model for prognostic prediction, In Proceedings of the Fifteenth International Conference on Machine Learning (ICML), Madison, Wisconsin, U.S., 540-546, 1998.
[21] D.R. Mani, J. Drew, A. Betz, P. Datta, Statistics and data mining techniques for lifetime value modeling, In Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), San Diego, CA, U.S., 94-103, 1999.
[22] S.F. Brown, A. Branford, W. Moran. On the use of artificial neural networks for the analysis of survival data. IEEE Transactions on Neural Networks, 8, 1071-1077, 1997.
[23] L. Bottaci, P.J. Drew, J. E. Hartley, M.B. Hadfield, R. Farouk, P.W.R. Lee, I.M.C. Macintyre, G.S. Duthie, J.R.T. Monson, Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions, The Lancet, 350, 469-472, 1997.