

eCRM에서 최적화 모형을 이용한 고객 분류 시스템

Customer Classification System using Optimized Form in eCRM

이재훈, 이성주

조선대학교 컴퓨터공학부

Jae-Hoon Lee, Sung-Joo Lee

Dept. of Computer Science Chosun University

E-mail : nuridepo@chosun.ac.kr, sjlee@chosun.ac.kr

요 약

기업들의 고객중심 마케팅 기법중 하나인 고객관계관리(CRM : Customer Relationship Management)가 인터넷의 발전으로 온라인화 되고 있으며 다양하게 발전되어 왔다. 가장 대두되고 있는 문제는 고객 분류를 객관적인 방법으로 어떻게 자동화할 수 있는가 이다. 본 논문은 최적화 모형을 이용하여 고객 분류를 더욱 세밀하게 할 수 있음을 제안하였고 고객 집단 편성 최적화를 반영함으로써 고객을 최적으로 분류할 수 있는 시스템을 설계 및 구축하였다.

1. 서론

이제 인터넷의 광범위한 확산에 따라 수많은 비즈니스 기회들이 열리게 됨으로써 전자상거래 및 인터넷 쇼핑물도 완전히 대중화되어 가고 있는 추세이다.

그러나 이러한 전자상거래에서 고객과의 효과적인 커뮤니케이션의 실현이 이루어지지 않으면 결국 고객들의 호응을 얻지 못할 것이다. 그러므로 사용자의 구미에 맞게 사용자에게 알맞은 정보를 제공한다면 고객에게 좋은 호응을 얻을 수 있을 것이다. 현재 대부분의 인터넷 쇼핑물은 고객의 단계적인 클릭행위를 통해 구매하고자 하는 제품에 접근하는 방식으로 운영되고 있다.

그래서 고객별로 개인화되지 않은 일률적 제품정보를 제공함으로써 고객에게 필요로 하는 제품 정보 탐색에 있어 많은 시간과 노력을 허비하게 한다. 결국 이는 인터넷 쇼핑물을 이용하는 고객의 전반적인 만족도 하락을 초래할 뿐만 아니라, 궁극적으로 한창 가속화되고 있는 전자상거래의 활성화를 저해하는 중요한 원인이 될 것이다.[7][8]

인터넷환경에서 사용자에게 알맞은 정보를 제공하는 마케팅 전략을 세우기 위해서는 사용자 개인의 취향, 접근경향에 관한 정보가 필요하다. 이와 같은 정보를 기반으로 사용자 개인의 특성에 맞은 동적인 웹 페이지 구성이나 링크정보를 제공할 수 있다. 현재 인터넷의 광범위한 확산에 따라 수많은 비즈니스 기회들이 열리게 됨으로써 전자상거래 및 인터넷 쇼핑물이 보편화되어 가고 있으나 이러한 전자상거래에서 고객과의 효과적인 커뮤니케이션이 원활하게 이루어지지 않으면 결국 고객들의 호응을 얻지 못할 것이다.

본 논문은 다음과 같은 구성으로 되어있다. 2장에서 관련 연구로 개인화와 고객 분류 최적화 알고리즘을 살펴본다. 3장에서 고객 분류 시스템을 제안하고 4장에서 결론 및 향후 연구과제를 제시한다.

2. 관련 연구

2.1 개인화

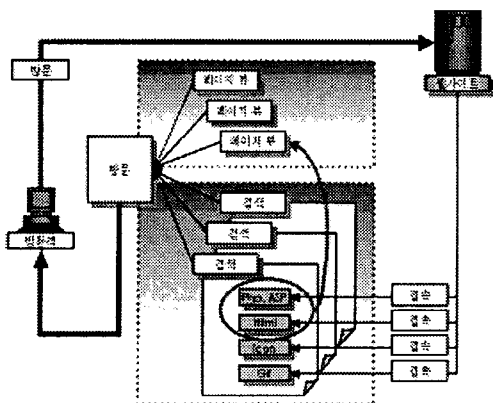
인터넷을 사용하다 보면 여러 가지 사이트들이 다양한 마케팅 방법을 사용해 사용자들이 그 사이트에 머물게 하거나 다시 방문하도록 하고 있

다. 그중에 하나가 사용자의 선호도에 따라 서비스를 선택할 수 있게 하는 방법이 있다. 각종 신문이나 인터넷 서점, 쇼핑몰 등에서 사용자의 구미에 맞게 편집하여 볼 수 있는 기능을 비롯해 사용자의 스타일에 맞는 정보를 선택할 수 있는 것을 보았을 것이다. 이와 함께 사용자의 개인적 취향에 따라 자신의 페이지를 구성하고 사용자의 구매기록, 취향에 맞는 제품을 추천받을 수 있는 기능들을 제공한다. 이렇게 방문자별로 각자의 구미에 맞는 정보를 제공하는 것을 개인화라고 한다.[7]

“개인화” 된 웹 사이트는 이용자에게 무엇을 원하는지 직접 물어 보는 것이 아니라, 이용자의 특성(Log Profiles, Navigation pattern Web Usage, etc.)을 통계학 및 인공지능 기법을 활용하는 데이터마이닝을 통해 분류(Classification) 또는 세그먼트화(Segmentation)하여 개인이 원하는 것을 예측(Prediction)하여 권유하고 다시 이러한 결과에 대해 Feedback하여, 계속적으로 학습(Learning)하는 적응적 웹 사이트(Adaptive Web Site)를 말한다.[9]

2.2 고객 분류 알고리즘

일반적으로 특정 웹 페이지를 보기 위한 사용자의 요구로, 웹 서버는 해당 웹 페이지와 관련된 여러 파일 등에 접근하게 된다. 따라서 사용자가 요청하는 특정 웹 페이지뿐만 아니라 해당 웹 페이지와 관련된 이미지 파일, 이미지 데이터, 모든 연관 파일 등에 대한 정보가 로그 파일에 저장되는 것이다. 이러한 웹 로그 트래킹을 측정하는 단위는 히트(Hits), 페이지뷰(Page View), 체류 시간(Duration Time), 세션(Session), 방문자(Visitor) 등이 있는데 현재 페이지뷰를 측정단위로 가장 많이 이용하고 있다. 로그 분석 방법을 [그림 1]과 같은 방법으로 설계하고 사용자의 페이지뷰를 측정 단위로 사용한다.



[그림 1] 페이지뷰/히트/접속/방문

고객의 로그를 분석하여 고객을 분류를 하면 [표 1]과 같이 12부류의 고객이 편성될 수 있다.

| 고객 분류 | | 의 의 | 등급 |
|------------------|-------|----------------------|----|
| 기 존 고 객 | 귀빈 고객 | 현재 거래가 활발한 고객 | A |
| | 철새 고객 | 잊어버릴 만하면 다시 오는 고객 | A |
| | 꼬리 고객 | 점점 발길이 줄어들는 고객 | B |
| 과 거 고 객 | 휴면 고객 | 잠시, 겨울잠을 자는 고객 | B |
| | 이별 고객 | 뭔가 불만이 있어 가 버린 고객 | C |
| | 번심 고객 | 다른 상표가 좋다고 가 버린 고객 | C |
| 잠 재 고 객 | 희망 고객 | 앞으로 계속 올 것 같은 고객 | A |
| | 유보 고객 | 언제가 꼭 올 것 같은 고객 | B |
| | 혜성 고객 | 한두 번 왔지만, 다시는 안 올 고객 | C |
| | 미지 고객 | 몰라서 안 오는 고객 | B |
| | 오해 고객 | 잘못 알아서 안 오는 고객 | B |
| | 주저 고객 | 알지만, 망설이며 안 오는 고객 | B |

[표 1] 고객의 분류

A등급 고객은 고객에게는 보다 적극적인 관심을 기울여야 한다. 개별적인 접촉과 알아주기 등이 이들을 고정 고객화하는 데 도움이 된다.

B등급 고객들은 잠재력이 충분한 고객들로서, 인내심을 가지고 꾸준하게 접근해야 한다. 고객의 반응을 성급하게 지대하지 말고, 직접 우편이나 광고·홍보를 통해 정보를 전달하는 등 동기 유발을 위한 끊임없는 자극이 필요하다.

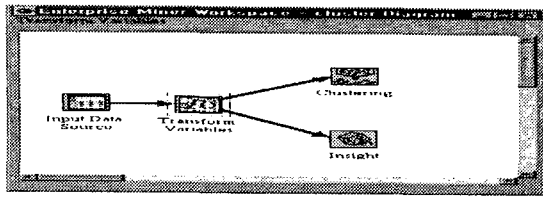
C급 고객들을 붙잡기는 쉽지 않지만 그들이 등을 돌린 이유를 탐색할 필요가 있다.

3. 최적화 모형 적용 고객 분류 시스템

고객의 로그를 분석하여 분류된 고객의 유형을 데이터 마이닝 기법을 적용하여 가공함으로써 집단 편성을 최적화 할 수 있다. [표 1]에서 A, B, C 등급으로 분류된 고객의 특성을 다음 STEP 1, 2, 3을 이용하여 최적화된 집단으로 편성하고 이를 이용하여 고객을 최적 분류하였다.

STEP 1.
Transformation Variable node를 사용해 새로운 분석변수를 생성을 먼저 수행한다.

[그림 2]처럼 Input Data Source Node는 Transform Variable node와 연결되어 있고 Transform Variable node는 Insight node와 Clustering node와 연결되어 있다.



[그림 2] Clustering Diagram

| Name | Keep | Mean | Std Dev | Skew | Kurtosis | CV |
|----------|------|----------|----------|-------|----------|----|
| SALESTOT | Yes | 4302.107 | 366.2132 | 3.77 | 15.62 | 0 |
| ST_RATIO | Yes | -1.63655 | 0.895818 | 2.74 | 8.86 | 0 |
| LE_RATIO | Yes | 0.030569 | 0.299099 | 0.68 | -0.19 | 0 |
| FA_RATIO | Yes | -3.82575 | 0.506011 | 3.10 | 18.80 | 0 |
| FASHION | Yes | 92.2627 | 32.61628 | 0.44 | 0.72 | 0 |
| LEISURE | Yes | 1916.419 | 380.84 | -0.17 | 0.61 | 0 |
| STRETCH | Yes | 444.0466 | 211.6901 | 0.35 | 0.60 | 0 |
| ORIGINAL | Yes | 1849.385 | 291.608 | -0.20 | 0.19 | 0 |

[그림 3] 새로운 변수들이 생성된 결과

STEP 2. 자료의 선택 및 설정한다.
 ① Input Data Source node를 연다.
 ② DUNGAREE 데이터 셋을 선택한다.
 ③ STORID를 ID변수로 설정한다.
 ④ Input Data Source node를 저장하고 닫는다.

이러한 형태의 판매 데이터를 군집화하는 하나의 접근방법은 제품 분류와 제품 판매량을 독립으로 간주하는 방법이다.

Enterprise Miner의 Clustering node는 k-means 군집분석 방법을 사용하고 있다. 많은 군집화 방법(k-means 군집분석 포함)이 입력변수(판매량)와 case(지점)과의 거리를 측정하여 분석의 도구로 사용하고 있다. 거리에 기초하는 방법은 입력변수의 scale에 매우 민감하다. 그러므로 자료에 대한 적절한 변환이 더 좋은 결과를 보이는 경우가 흔히 나타난다.

STEP 3. 군집화를 위한 새로운 변수의 생성
 ① Transform Variable node를 연다.
 ② Action → Create Variable을 선택한다.
 ③ Name Field에 FA_RATIO라고 입력한다.
 ④ Define을 선택한다.
 ⑤ 수식을 정의하는 부분에 수식을 입력한다.

새로 생성된 FA_RATIO 변수는 각 지점의 original jean에 대한 fashion jean의 상대적 scale을 나타내며 이 예에서 log 변환을 한 이유는 분석에 보다 적절한 scale이라고 간주되었기 때문이다.

위의 절차를 반복하여 다음과 같은 변수를 생성한다.

- LE_RATIO=LOG(LEISURE/ORIGINAL)
- ST_RATIO=LOG(STRECH/ORIGINAL)
- SALESTOT=FASHION+LEISURE
+STRETCH+ORIGINAL

[그림 3]은 새로운 변수들로 얻어낸 결과이다.

4. 결론

본 연구에서는 웹 마이닝을 통하여 고객 정보를 재가공하여 최적화하는 알고리즘을 사용하여 고객을 분류하고 다시 한번 가공하여 최적화된 집단 편성을 할 수 있음을 보였다. 이를 통해 온라인과 오프라인을 이용하여 고객분류를 자동화할 수 있고 기업의 의사결정에 많은 영향을 미칠 것으로 사료된다. 제시한 알고리즘을 통해 기업은 고객형태를 쉽게 분류할 수 있을 것이며 고객을 관리하는 지침이 되고 기업간 고객 관계에 많은 영향을 줄 것으로 사료된다. 그러나 집단 편성을 최적화할 수 있다 하더라도 [표 1]의 C등급에 해당하는 고객들은 기업의 노력에도 불구하고 A등급 또는 B등급으로 전향이 어려울 것으로 예상된다. 하지만 C등급 고객들의 형태와 패턴을 파악함으로써 A, B등급 고객들을 C등급 고객으로 전향하지 않을 수 있을 것이다.

기업은 이러한 고객 집단에 대한 적절한 운용과 각 개인의 성향에 따른 방침을 더욱 강화하여 고객 중심의 경제 활동을 도모할 수 있을 것이다.

참고문헌

- [1] 코리아 인터넷마케팅센터,
<http://www.webpro.co.kr>
- [2] Accrue software Inc, "Web Mining Whitepaer", Accrue White Paper, 2000
- [3] 삼성SDS IT Review, "효율적인 로그분석을 활용한 eCRM 마케팅 전략", 2002
- [4] 손소영, 신형원, "범주형 자료에 대한 데이터 마이닝 분류기법 성능 비교", IE Interfaces Vol. 12, No. 4, pp. 551-556, December 1999
- [5] ㈜웹로그, "웹로그 사용자 설명서", 2001
- [6] ㈜웹로그, "Log Cluster 사용자 설명서", 2001
- [7] 차대운, 박유진, "인터넷 쇼핑몰에서 개인화된 정보 제공에 관한 연구", 한국기업경영학회 pp.231-247, 2001.5
- [8] Hyuncheol Kang, Byoung Cheol Jung "A Study of Web Usage Mining for eCRM" *The Korean Communications in Statistics* Vol.8 No.3, 2001 pp 831-840
- [9] Harris Kravatz "Designing Web Personalization Features"