經營科學
第21卷 第3號
2004年 11月

백화점 고객의 구매 분석 및 고객관계관리 전략 적용

하성호* · 백경훈**

Analyzing Customer Purchase Behavior of a Department Store and Applying Customer Relationship Management Strategies

Sung Ho Ha* · Kyung Hoon Baek**

■ Abstract ■

This study analyzes customer buying-behavior patterns in a department store as time goes on, and predicts moving patterns of its customers. Through them, it suggests in this paper short-term and long-term marketing promotion strategies. RFM techniques are utilized for customer segmentation. Customers are clustered by using the Kohonen's Self Organizing Map as a method of data mining techniques. Then C5.0, a decision tree analysis technique, is used to predict moving patterns of customers. Using real world data, this study evaluates the prediction accuracy of predictive models.

Keyword : Customer Relationship Management, Customer Analysis, Data Mining, Customer Segmentation, Marketing Strategy

1. 서 론

국내 백화점 업계는 1990년대 초반 이후 소비자 중심의 적극적인 고객만족(Customer satisfaction : CS) 경영을 목표로 정하고, 불특정 다수의 고객을 대상으로 전통적인 TV, 신문, 잡지, 전단 채널 등을 통하여 광고 및 판촉 활동을 펼치는 매스마케팅(Mass marketing)을 실시하였다. 하지만, 전통적인 매스마케팅은 시간이 지남에 따라 마케팅 전략 방향의 부재, 일원화된 고객 관리의 어려움, 마

논문접수일 : 2004년 7월 31일 논문게재확정일 : 2004년 10월 15일
* 경북대학교 경상대학 경영학부
** 화성산업(주) 동아백화점 전산실

케팅 효율성 감소, 성과 측정의 어려움, 외부 환경 및 소비 동향에 대한 구체적인 대응 수단이 부족하다는 등 여러 가지 문제점을 노출시켰다. 또한, 매스마케팅 전략은 고객의 욕구를 만족시키기 보다는 고객에게 상품이나 행사를 안내하는 수준의 마케팅 접근 방법으로 진정한 의미의 고객 만족을 이끌어내지 못하는 전략적 접근 방법을 구사하였다.

게다가 최근 백화점 업계의 동향을 살펴보면 시장의 수요는 늘어나지 않는 상황에서 고객이 백화점을 선택할 수 있는 폭은 넓어졌고, 백화점간의 경쟁은 더욱 심화되고 소비자들의 욕구는 더욱 다양화 되었으며, 기존 소비자를 위한 서비스와 각종 행사는 쉽게 모방될 수 있게 되었다. 따라서 유통시장에서 공급자와 수요자와의 관계는 과거 공급자인 백화점이 시장을 주도하던 시대에서 수요자인 고객이 중심이 되는 고객 중심의 시장 환경으로 변화하고 있다. 고객 중심의 시장 환경에서 기존의 일회성이고 무차별적인 매스마케팅 전략으로는 더 이상 고객의 욕구를 정확히 충족할 수 없게 됨으로써 정확한 고객 정보를 바탕으로 전략적 고객세분화를 통해 마케팅의 목표 고객군을 찾고, 더 나아가서 개별화된 목표 고객을 선별하고 이에 맞는 적절한 마케팅프로모션 전략을 통해 고객 자체에 초점을 맞추어 고객의 만족을 이끌어내는 진정한 고객만족 경영의 전략적 접근방법이 필요하게 되었다. 이러한 새로운 경영패러다임의 등장으로 나타난 것이 고객관계관리(Customer Relationship Management : CRM)기법이다. 고객관계관리는 고객의 욕구를 충족시킴으로서 고객과의 신뢰 관계를 구축하고 궁극적으로 수익성 높은 고객을 장기적으로 유지시키는 광범위한 비즈니스 활동을 말한다[7, 25].

고객관계관리의 마케팅적 접근 방법은 단기적인 관점에서 어떤 고객 집단을 표적화(targeting)하고 마케팅프로모션을 실시하여 즉각적인 구매를 유도할 것인지, 중기적인 관점에서 신규 고객의 획득전략, 기존고객을 유지하기 위한 전략으로 어떤 고객 집단을 표적화하여 마케팅프로모션 비용을 최소화하고 이익을 최대화할 것인지에 대한 방안을 결정하고, 장기적인 관점에서 고객생애가치(Life Time Value : LTV)를 극대화 할 수 있는 여러 가지 전략적 방안을 결정하는데 이용할 수 있다. 그러므로 고객관계관리에서 장단기적인 관점을 염두에 두고 고객들에게 적절한 마케팅프로모션을 실시하기 위해서는 시간의 변화에 따른 고객들의 구매 행위 패턴을 예측할 수 있는 동적인 고객 분석 방법이 필요하다. 즉, 고객들은 시간이 지남에 따라 변화하고 그들의 구매행위 또한 시간이 지남에 따라 변화하기 때문에 특정 시점에서 추출한 고객의 구매행위 분석 내용은 특정 시점이 지나면 가치를 잃어버린다. 따라서 시간이 흐름에 따라 변화하는 고객의 구매행위 패턴을 미리 예측할 수 있는 고객의 구매행위 분석 방법인 동적인 고객 분석방법이 요구된다.

본 연구는 국내 A 백화점에서 구매 고객들의 동적인 분석을 위해서 구매행위의 패턴을 시간의 흐름에 따라 분석하여 동일 구매행위 패턴의 고객들을 군집화하고, 미래 고객군집의 이동패턴을 예측하여 고객관계관리기법을 활용한 장단기적인 관점에서의 마케팅 프로모션 방안을 제시하고자 한다.

본 논문은 모두 5장으로 구성되었으며, 서론에서는 연구의 배경과 목적 및 연구 범위와 방법을 제시하였다. 제2장에서는 일반 유통업에서의 고객관계관리 도입과 국내 백화점의 고객관계관리 도입 사례를 살펴보았다. 제3장에서는 연구 대상 백화점의 실제 고객 자료를 근거로 RFM 분석 기법과 다변량 분석 기법 중 하나인 데이터마이닝을 활용한 동적인 고객 분석 방법론을 제시하였다. 제4장에서는 분석 결과를 기술하였고 장단기적인 고객관계관리 전략을 제시하였다. 제5장에서는 본 연구의 의의와 향후 연구 방향을 제시하였다.

2. 고객관계관리

2.1 고객관계관리의 전략적 활용

고객관계관리는 신규 고객을 획득하고 고객의

요구를 함께 공유하며 고객과의 상호작용을 통하여 장기적으로 기업의 이익을 증대하는 것이다[15]. 또한, 진정한 고객이 누구인지, 그들이 어떤 행동을 하는지 그리고 그들이 무엇을 원하는지를 미리 인식하기위하여 비즈니스 프로세스와 정보 기술을 개선하는 것이다[24]. 일반적인 고객관계관리에 대한 정의는 신규고객획득, 기존고객유지 및 고객의 수익성을 증대시키기 위하여 지속적인 커뮤니케이션을 통해 고객 행동을 이해하고 영향을 주기위한 광범위한 접근을 말한다[22].

고객관계관리 전략은 고객과 관계를 형성하는 부분과 이를 업무 영역에 활용하는 부분으로 나눌 수 있다. 고객과 관계를 형성하는 전략에는 첫째, 신규 고객을 발굴하는 전략으로 기존의 고객 데이터를 중심으로 기업의 수익 증대에 긍정적인 영향을 줄 수 있는 고객 유형을 찾아내어 그들과 관계를 맺기 위해 다양한 채널을 통하여 접근한다. 둘째, 기존 고객의 충성도를 높이는 전략으로 지속적인 고객 관계 강화 활동과 로열티 프로그램을 통하여 고객의 충성도를 높이고 장기적인 관계를 강화한다. 셋째, 교차판매(cross-selling)나 업셀링(up-selling)을 통하여 수익성을 증대하는 전략으로 고객의 거래 데이터를 분석하여 고객 구매 패턴을 발견하고 이탈 고객을 재획득하거나 휴면 고객을 재활성화 하는 것으로 다양한 고객 정보를 기반으로 고객의 이탈 시점을 예측하고 이탈이 예상되는 고객에게 이탈 방지 마케팅 프로모션을 실시한다.

영업에 활용하는 전략에는 첫째, 고객의 가치를 분석, 평가하여 고객 세분화를 실시하고 타켓 마케팅이나 일대일 마케팅을 실천한다. 둘째 고객 접점에서 제품 및 서비스에 관한 다양한 정보를 제공할 수 있도록 과거 고객의 구매 패턴 및 정보를 분석하여 매출을 유도한다. 셋째, 고객 서비스에 있어서 고객의 다양한 접촉 정보를 통해 고객과의 일대일 커뮤니케이션을 실천하여 관계를 우호적으로 유지한다. 이러한 전략적 활용들은 모든 고객의 데이터를 기반으로 수립되어진다[6].

2.2 일반 유통업에서의 고객관계관리

김영아[2]는 백화점의 시장 세분화를 위한 데이터마이닝 응용에 관한 연구에서 백화점 고객의 매출 거래 자료를 이용하여 개별 점포 차원에서 바겐세일이 고객들의 구매 행동에 미치는 영향을 연구하였다. 방문율(frequency), 구매액(amount), 기간(term)에 따른 고객들의 방문율과 구매액의 패턴을 그래프로 분석하는 추세분석(trend analysis)을 실시하고 고객 차원에서 바겐세일에 대한 반응의 특성을 기준으로 계층적 군집분석(hierarchical clustering)기법 중 집합적(agglomerative)군집 분석을 이용하여 고객들을 세분화하여 각 집단에 적절한 개별 마케팅 전략을 수립하였다.

우근식[5]은 백화점에서 데이터마이닝을 이용한 고객충성도 분석에 관한 연구에서 전통적인 고객 분석 방법인 RFM 분석을 통하여 백화점의 고객 매출 자료를 근거로 계산한 고객의 스코어로 고객 세분화를 실시하였고, 의사결정나무 분석방법을 이용하여 고객의 특성을 분류하였다.

가상 상점에서 고객 행위 연관성 분석을 위한 데이터마이닝 기법에 관한 김종우[3]의 연구는 가상 상점에서 고객의 구매 이력 및 행동을 모니터링 하여 고객의 다양한 취향을 파악하고 이를 마케팅에 활용하였다. 그는 데이터마이닝 기법의 하나인 장바구니 분석을 활용하여 고객 유지 및 이탈 방지를 위해서 가상 상점의 기존 고객 및 잠재 고객을 대상으로 고객 행위를 분석하여 지식을 추출하는 기법을 제시하였다.

곽재민[1]은 통신판매업체의 데이터베이스 마케팅에 관한 연구를 통해 국내 통신판매업체의 데이터베이스 마케팅 수행을 위해 구체적인 방법으로 데이터마이닝 기법 중에서 의사결정나무 분석방법을 이용하였다. 이를 통해 고객 데이터에 대한 다양한 통계 분석을 수행하고, 우수 고객을 평가, 관리하기 위해 현실적인 RFM 분석 모형을 수립하고 분석하여, 고객의 생애 가치를 측정하는데 필요한 고객별 잔존가능성을 예측하였다.

박병훈[4]은 RFM 점수를 이용한 고객 반응 분석에 관한 연구를 하였는데 RFM 점수를 기초 자료로 이용하고 데이터마이닝 기법 중에서 의사결정나무 분석방법과 신경망 분석, 로지스틱 회귀분석방법을 이용하여 고객들의 마케팅 캠페인에 대한 반응을 예측하여 RFM 점수에 의한 고객 세분화의 방법과 비교하여 결과를 유도하였다. RFM 점수만을 고려한 반응 예측율은 74.54%로 나타났고 가용한 모든 변수를 고려한 예측에서 C5.0 방법의 예측율이 78.98%로 가장 높게 나타났다.

고객 군집화의 이론들 중에서 최근 신경망(Neural network)을 이용한 연구가 과거 스코어로 고객을 군집화하는 방법이나 회귀분석을 통한 분석 보다 더욱 정확한 고객군집화의 방법으로 이용되며[1, 9, 10, 26] 신경망 분석 방법 중 자기조직화맵(Self-Organizing Map : 이하 SOM)을 이용하여 고객 군집화를 실시하는 것이 고객 군집화의 기준을 임의로 설정한 분석 방법 보다는 더 좋은 고객 군집화 방법으로 제시되고 있다[8, 16, 19].

2.3 국내 백화점 업계의 고객관계관리 도입

2.3.1 현대 백화점

현대백화점은 1998년 CRM 시스템 구축을 시작하여 매출액 기준으로 신규고객, 이탈고객, 휴먼고객, 재활성화고객 등 전략적 고객군으로 고객을 세분화하고, 고객활성화전략, 고객애호도제고전략, 고객유지전략, 신규고객확보전략 등 마케팅전략을 수립하였다. 2001년부터는 기존 백화점 사업 부문은 물론 인터넷 쇼핑몰, 홈쇼핑, 여행 사업 및 호텔 부문까지 확대 통합한 전사적 CRM시스템 구축을 목표로 CRM을 진행하였다.

2.3.2 신세계 백화점

신세계백화점은 1995년 일본에서 데이터 분석 패키지를 도입하여 활용하였으며, 1999년 12월 신세계 I&C가 인터넷 환경에 적합한 웹 기반 시스템으로 CRM을 구축하였다. 신세계백화점은 POS (point-of-sales) 시스템을 실시간으로 연결하여 구축한 "신고객마케팅시스템"을 이용하여 고객 한사람 한사람을 대상으로 일대일마케팅을 펼칠 수 있게 되었다. 신세계백화점의 CRM의 특징은 저비용, 고효율 구조의 마케팅이 가능한 고객관리프로그램을 제공하고 최신 정보 기술의 통합과 최적화된 시스템을 개발, 운영, 유지하는 프로세스들을 기반으로 다년간의 활용과 적용으로 검증된 풍부한 노하우와 사례를 가지고 있다는 점이다.

2.3.3 롯데 백화점

롯데백화점은 롯데정보통신을 주축으로 한국후지쯔와 공동 개발한 고객관리 데이터베이스 시스템을 1998년 도입하였다. 이 시스템을 이용하여 각종 고객 분석 자료를 수 분 내지 수십 분에 생성할 수 있게 되었다. 과거 매출액만 기준으로 DM(direct mail) 행사의 대상 고객을 추출하던 방식에서 벗어나 다양한 고객 자료를 분석하여 그 대상을 추출함으로써 고객의 성향과 욕구에 맞는 마케팅이 가능하게 되었다. 통합된 고객관리 데이터베이스 시스템을 활용하여 회원고객의 장바구니분석, 구매회수분석, 휴면고객분석, 고액구매고객의 매출을 분석하여 상품 구성과 매장 배치를 결정하는데 활용하여 고객에 대한 서비스를 향상시키고 휴면고객의 매출 활성화를 유도하였다.

2.3.4 갤러리아 백화점

갤러리아백화점은 1999년부터 CRM의 초보 단계로 불리는 데이터베이스마케팅을 전개하였고, 2001년 후반 명품관, 패션관 등에 대해 차별화된 고객 서비스를 발굴하였다. 고객 분석 정보시스템이 구축된 2002년 후반부터는 백화점 제휴업체는 물론 서울프라자호텔, 한화국토개발, 투어몰(tour mall) 등 그룹 내 레저 사업 부문과 연계하여 종합적인 서비스를 제공하고 있다. 갤러리아백화점은 파레토 법칙을 적용한 CRM 전략을 실시하여 이익공헌도가 높은 고객을 대상으로 각 개인의 상세한 정보

를 분석하고 우수 고객을 집중적으로 관리해 고객 로열티를 계속 높여 나가는 전략을 펼치고 있다. 또한, 각종 공연 초청, 생일과 결혼기념일 축하 이벤트, 멤버십 카드를 통한 우대 혜택 등 부가적인 서비스를 도입하는데도 CRM 시스템을 활용하고 있다.

2.3.5 애경 백화점

애경백화점은 1997년 1월 경쟁점의 신규 출점에 따른 고객 유지의 필요성과 판촉 경쟁의 가속화에 따른 고비용 구조의 마케팅 전략에서 벗어나 고효율 구조의 마케팅 전략을 펼치기 위해서 CRM을 도입하게 되었다. 1998년 ACOM이라는 CRM 시스템을 오픈하여 상권정보관리, 고객정보관리, 구매정보관리, 판촉관리, 고객세분화, DM 발행, 텔레마케팅, 설문조사분석, 고객과 브랜드의 ABC 분석, 고객유지율분석, 판촉반응분석, 가격민감도 분석에 활용하였으며, 매출 분석, 이익 분석, 경쟁점 분석, 행사장 분석, 상품 교차 분석, 브랜드 교차 분석을 실시하였다.

3. 고객 분석과 고객관계관리

본 연구에서 사용한 고객 자료는 대구 시내에 소재한 A 백화점의 2002년 7월부터 2003년 12월까지의 자사 카드 매출거래와 고객의 식별이 가능한 현금매출거래 자료이며, A백화점의 자사카드 중 법인카드를 제외한 개인회원카드 소지자를 대상으로 정규 직원 및 파견 직원은 제외한 일반회원의 매출을 기초 자료로 하였으며, 본인과 가족을 분리하여 개인별로 분석하는 기준을 정하였다.

본 연구의 목적인 고객관계관리기법을 활용한 동적인 고객 분석 및 전략적 활용을 위한 고객 군집화와 예측을 위한 분석 방법은 <그림 1>과 같다. 고객 군집화를 위한 기초 자료 분석을 위해 RFM 분석 방법을 채택하였고, 코호넨의 SOM을 이용하여 고객을 군집화하였다. 우수고객군집과 이탈가능군집 등 시간의 흐름에 따라 각 군집의 변화 모습을 분석하고, 고객의 군집 간 이동을 예측하기 위해서 의사결정나무 분석방법인 C5.0을 이용하였다.

3.1 고객 군집화

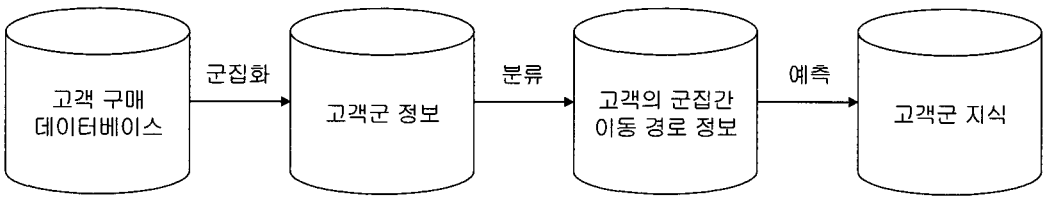
<그림 2>는 고객 군집화 방법을 상세하게 설명하고 있다.

3.1.1 매출 분석 기간의 설정

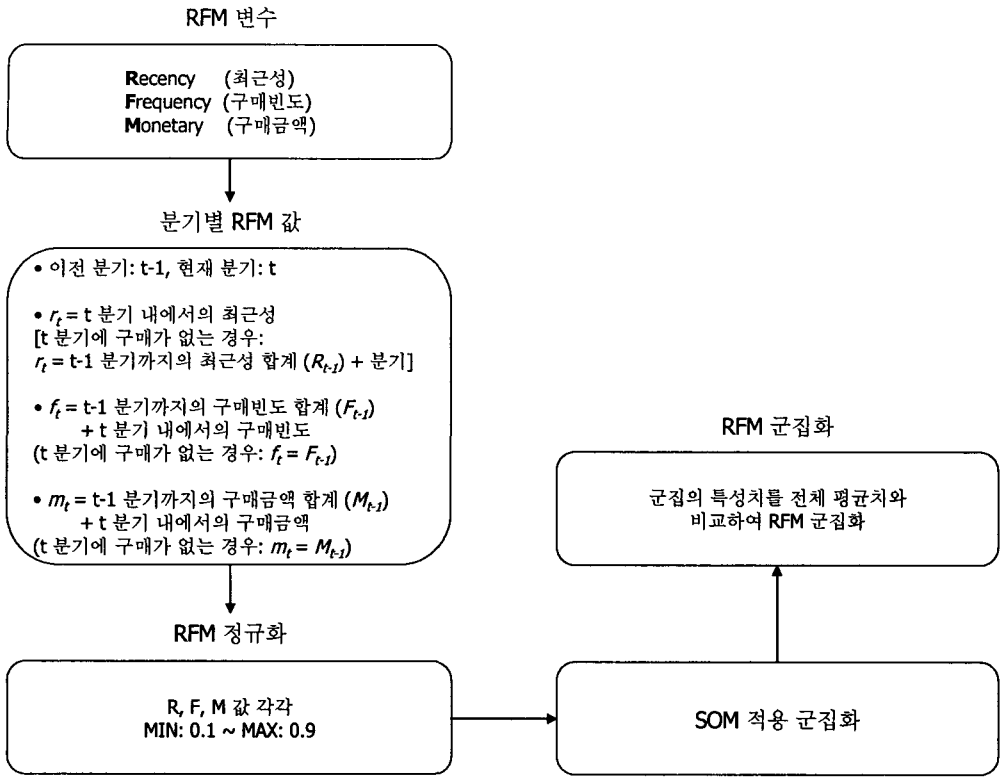
분석 기간은 분기(3개월) 단위로 2002년 7월부터 2002년 9월까지 기간을 처음 매출분석기간으로 정하고, 2003년 9월까지 총 5개의 분기에 대해 고객 매출을 분석하였다.

3.1.2 RFM 분석기법 적용

고객의 구매 분석을 위해서 데이터베이스 마케팅에서 가장 널리 사용하는 방법인 RFM 분석 방법을 채택하여 고객 군집분석의 기초값으로 이용하였다[14]. RFM 분석에서의 R(ecency)은 최근에 구매를 하였는가, F(requency)는 얼마나 자주 구



〈그림 1〉 고객 군집화와 예측을 위한 분석 방법



〈그림 2〉 상세한 고객 군집화 방법

매를 하였는가, M(onetary)은 구매 금액은 얼마나 되는가를 나타내는 것으로 고객 구매 행위를 매출의 최근성, 구매빈도, 구매금액을 기준으로 분석하는 방법이다. 예를 들어 분석 기간이 2002년 7월 1일부터 2002년 9월 30일까지의 3개월인 경우, 분석 기간의 마지막 일자인 9월 30일에 매출이 발생했으면 R값은 0의 값을 가지고 이 기준일에서 분석 고객별로 매출일자까지의 기간을 R값으로 구하였다. F값은 분석 기간 중 고객별로 매출 일수를 합하여 계산하는데 1일 2회 이상 매출이 발생하여도 일자를 기준으로 1회 방문한 것으로 간주하였다. M값은 분석 기간 중 고객별 구매 금액을 합산하였다.

3.1.3 분기별 RFM 값 계산

분기별로 RFM 값을 계산할 때 고객이 해당 분기에 매출이 없으면 분석 대상에서 제외되는 경우가 발생한다. 본 연구의 분석 기간 단위는 분기(3개월)이며 고객이 3개월간 구매가 없다고 분석 대상에서 제외하기에는 짧은 기간이므로 분석 대상에서 고객을 제외하거나 이전 분기의 우수고객이 우수고객군에서 제외되는 것을 방지하기위해서 RFM 값을 계산할 때 누적값을 적용하는 기준을 정하였다(Ha 등, 2002).

이전 분기를 *t-1*, 현재 분기를 *t*라고 하였을 때 아래의 수식과 같이 R, F, M의 누적값을 계산 적용하였다.

- *t*분기의 Recency 값 (r_t)
  [단, *t*분기에 매출이 없는 경우 : r_t = *t-1*분기까지의 최근성 합계 (R_{t-1}) + 분기]
- *t*분기의 Frequency 값 (f_t) = *t-1*분기까지의

구매빈도 합계 (F_{t-1}) + t분기 내에서의 구매빈도 (단, t분기에 매출이 없는 경우 : $f_t=F_{t-1}$)

- t분기의 Monetary 값 (m_t) = $t-1$분기까지의 구매금액 합계 (M_{t-1}) + t분기 내에서의 구매금액(단, t분기에 매출이 없는 경우 : $m_t = M_{t-1}$)

3.1.4 RFM값의 정규화

R, F, M값을 고객 군집화를 위한 입력값으로 이용하기 위해서는 R값, F값에 비하여 M값이 상대적으로 스케일이 크므로 R, F, M값을 동일 스케일로 정규화 하여야 한다. 분석 기간 중 R값에서 가장 큰 값을 0.9로, 가장 작은 값을 0.1로 정하고, 여타 R값을 0.1과 0.9사이의 값으로 정규화 하였다 (식 1 참조). 동일한 방식으로 F값, M값의 정규화 작업을 수행하였다.

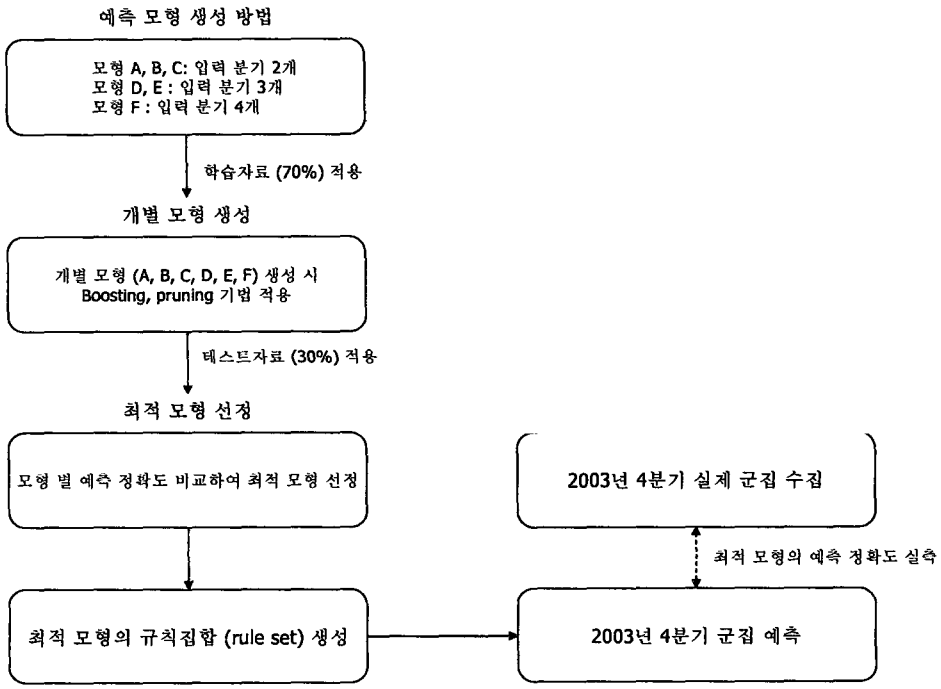
$$\text{정규화된 r} = \frac{(0.9-0.1)[\text{정규화전 r} - \text{M}_n(\text{r})]}{\text{Max(r)} - \text{Min(r)}} + 0.1 \quad \text{(식 1)}$$

3.1.5 SOM을 이용한 고객 군집화

고객군집화에 사용된 군집 알고리즘은 코호넨의 SOM이다. SOM은 비모수 군집화 알고리즘으로 고객이나 상품을 속성이 비슷한 집단으로 군집화 할 때 적합한 방법인데[13, 17] 여기서는 정규화한 RFM 값을 입력 값으로 고객군집화를 실시하였다. 그리고 나서 각 분기별로 형성된 군집의 평균 R값을 해당 분기의 전체 평균 R값과 비교하여 크면 '↑'로, 작으면 '↓'로 기호화하였다. 마찬가지로 F값과 M값도 동일한 방법으로 기호화하여 고객 군집을 RFM 군집으로 특성화 하였다.

3.2 고객군집의 예측모형 생성

본 연구에서는 고객의 군집간 이동 패턴을 동적으로 분석하기위해 최적의 예측모형을 생성하여야 하는데 <그림 3>과 같은 예측모형 생성 및 검증 방법을 이용한다.



〈그림 3〉 예측 모형 생성 및 검증 방법

3.2.1 시간 창을(Time window) 달리하는 복수 모형 설계

분석에 이용가능한 분기별 자료를 시간 창을 달리하는 방법으로 모형 A~모형 F까지 6개의 모형으로 생성하였다. 예를 들어 모형 A를 생성할 때 2002년 3분기와 2002년 4분기의 고객 군집을 분류 알고리즘에(여기서는 C5.0) '입력 속성'으로, 2003년 1분기 소속 군집을 '결과 클래스'로 사용하였다. <표 1>은 각 모형별 입력 속성과 결과 클래스를 보여주고 있다.

3.2.2 모형 생성을 위한 데이터 세트(data set) 구성

고객의 군집 간 이동 패턴을 동적으로 분석하기 위해 예측 모형에 적용하는 자료는 고객 자료 97,856건 중 68,542명(70%)의 자료를 예측 모형을 만들기 위한 학습용 자료 세트(training data set), 29,314명(30%)을 테스트 자료 세트(test data set)로 구성하였다. 예측의 정확도를 높이기 위해서 무작위로 추출하기 보다는 층화표본추출방법(stratified sampling)으로 샘플의 특성이 모집단의 특성을 반영하도록 하였다.

3.2.3 예측모형 생성을 위한 C5.0 적용

시간의 흐름에 따라 변화하는 고객 군집의 예측 모형을 만들기 위해 의사결정나무 분석 방법 중 C5.0을 이용하였다. C5.0은 의사결정나무 분석의 원리에 따라 자료를 분류하여 적절한 분류 규칙을 생성하는 방법이다. 규칙은 일반적으로 "A이면 B이다"라는 형태로 구성되며 생성된 예측모형에 데이터를 입력하면 분류 결과를 쉽게 예측할 수 있다. 학습 알고리즘의 정확성을 개선하는 방법으로 학습 시 부스팅(Boosting) 기법을 사용하였다. 부스팅은 처음 학습규칙을 생성하고 학습규칙에서 벗어나는 자료들을 다음 학습규칙에서 처리하는 기법으로 반복적인 학습규칙을 통해 가장 정확성이 높은 최적의 학습규칙을 제안하여 학습규칙을 통한 예측의 정확성과 효율성을 높일 수 있다. 이외에도 가지치기(pruning)를 하면 예측력을 향상시킬 수 있는데 이를 통해 모형을 일반화하였다. 가지치기는 데이터에 포함된 모든 레코드들을 완벽하게 분류할 때까지 나무를 확장하여 분류모형을 만들게 되면 불필요한 속성들이 나무의 마디에 포함되기 때문에 모형이 제공하는 규칙 자체가 엉뚱한 의미를 갖게 될 수 있으며, 이로 인해 새로운 레코드에 대한 예측력이 감소하는 결과를 초래하는 현상을 방지하기 위해 과잉맞춤(Overfitting)으로 인해 불필요하게 복잡해진 의미 없는 마디를 제거하는 것을 의미한다[18, 20, 23].

3.2.4 예측모형의 검증을 위한 정오분류행렬

최적의 모형을 찾아내기 위해서 정오분류행렬(confusion matrix)을 이용하여 예측 모형의 정확성을 측정하였다. 정오분류행렬이란 목표변수의 실제범주와 모형에 의해 예측된 분류 범주사이의 관계를 나타내는 표이다. 목표변수의 범주가 3개인 경우 3×3개의 셀로 이루어진 테이블 형식의 결과로 나타낼 수 있다. 정오분류행렬표에서 대각선상에 나타나는 수는 실제범주와 예측범주가 일치하

〈표 1〉 예측 모형 생성을 위한 모형별 '입력'과 '결과 클래스'

분기 / 모형	2002년 3분기	2002년 4분기	2003년 1분기	2003년 2분기	2003년 3분기
모형 A	입력 속성	입력 속성	결과 클래스		
모형 B		입력 속성	입력 속성	결과 클래스	
모형 C			입력 속성	입력 속성	결과 클래스
모형 D	입력 속성	입력 속성	입력 속성	결과 클래스	
모형 E		입력 속성	입력 속성	입력 속성	결과 클래스
모형 F	입력 속성	입력 속성	입력 속성	입력 속성	결과 클래스

는 개체의 수이고, 비대각에 존재하는 수는 예측이 어긋난 개체의 수를 나타내는 방법이다.

4. 적용 결과

4.1 분기별 군집 분석

본 연구에서 고객 군집화를 하기 위해 사용된 기초 자료는 고객번호, 고객별 R값, F값, M값, 정규화 R값, 정규화 F값, 정규화 M값, 분기구분코드로 구성되어있다. 그 중에서 SOM의 학습을 위해 정규화 R값, 정규화 F값, 정규화 M값을 투입하였다. SOM의 학습 파라메타는 입력 노드 3개, 출력 노드를 9개로 하였다. <표 2>는 SOM에 의해 군집화된 2002년 3분기의 고객군집별 인원수와 비율을 보여주고 있다.

동일한 방법으로 2002년 4분기부터 2003년 3분기까지 4개 분기의 군집화를 실시하였다. <표 3>은 각 분기별 군집화의 결과로서 군집의 개수, 인원수, 형성된 군집의 평균 R값, 평균 F값, 평균 M값을 나타내고 있다. 각 분기별로 군집의 개수가 차이 나는 것은 SOM을 통하여 형성된 군집이 경쟁학습을 통하여 입력 값의 특성이 유사한 군집으로 형성되므로 군집의 개수는 차이가 나타날 수 있으며 따라서 분기별로 항상 동일한 개수의 군집이 나오는 것은 아니다.

또, 분기별로 형성된 군집의 위치가 X-Y 평면상에서 동일한 지역에 위치한 군집이라고 해서 동일한 의미를 가진 군집이라 할 수 없다. 이것을 해결하기 위해, 각 분기별로 형성된 군집의 R값을 해당 분기의 전체 평균 R값과 비교하여 크면 '↑', 작으면 '↓'로 기호화하였다. 마찬가지로 F값과 M값도 동일한 방법으로 기호화하여 고객군집을 식별하도록 하였다. <표 4>는 2002년 3분기의 고객군집을

〈표 2〉 2002년 3분기의 군집별 인원수와 비율

군집	1	2	3	4	5	6	7	8	9	합계
인원수	8,694	891	18,169	20,221	1,881	11,784	15,115	12,627	8,474	97,856
비율(%)	8.88	0.91	18.57	20.66	1.92	12.04	15.45	12.90	8.66	100

〈표 3〉 분기별 군집 현황과 기초 통계치

분기	군집 수	자료 수	전체 평균 R값	전체 평균 F값	전체 평균 M값
2002년 3분기	9개	97,856	23.84	6.18	540,764
2002년 4분기	8개	124,787	35.88	9.49	966,445
2003년 1분기	7개	140,214	41.16	12.58	1,316,949
2003년 2분기	8개	147,947	49.74	16.05	1,649,751
2003년 3분기	8개	153,664	50.35	19.14	1,922,581

〈표 4〉 2001년 3분기 군집화 결과

X-Y 위치	인 원 수	군집 평균 R값	군집 평균 F값	군집 평균 M값	RFM 군집
00	8,694	2.18	26.39	1,530,595	R↓F↑M↑
01	891	32.61	8.99	693,321	R↑F↑M↑
02	18,169	72.37	1.54	196,135	R↑F↓M↓
10	20,221	4.67	9.77	885,755	R↓F↑M↑
11	1,881	21.95	8.18	762,462	R↓F↑M↑
12	11,784	39.05	2.27	268,372	R↑F↓M↓
20	15,115	2.79	2.99	344,091	R↓F↓M↓
21	12,627	11.17	2.66	385,047	R↓F↓M↓
22	8,474	22.55	2.48	337,304	R↓F↓M↓
전체	97,856	23.84	6.18	540,764	

〈표 5〉 전체 고객을 대상으로 분석한 분기별 군집현황

군집 \ 분기	2002년 3분기	2002년 4분기	2003년 1분기	2003년 2분기	2003년 3분기
R↑F↓M↓	29,953	46,827	63,159	65,274	77,337
R↓F↑M↑	30,796	36,435	39,408	51,534	64,196
R↓F↓M↓	36,216	41,525	25,858	17,306	12,131
R↓F↓M↑	0	0	11,789	13,833	0
R↑F↑M↑	891	0	0	0	0
합 계	97,856	124,787	140,214	147,947	153,664

〈표 6〉 2002년 3분기 고객 97,856명의 군집간 이동 현황

군집 \ 분기	2002년 3분기	2002년 4분기	2003년 1분기	2003년 2분기	2003년 3분기
R↓F↑M↑	30,796(31.5%)	33,288(34.0%)	33,127(33.9%)	41,040(41.9%)	48,836(49.9%)
R↑F↓M↓	29,953(30.6%)	36,585(37.4%)	39,069(39.9%)	36,428(37.2%)	41,372(42.3%)
R↓F↓M↓	36,216(37.0%)	27,983(28.6%)	17,423(17.8%)	10,755(11.1%)	7,648(7.8%)
R↓F↓M↑	0	0	8,237(8.4%)	9,633(9.8%)	0
R↑F↑M↑	891(0.9%)	0	0	0	0
합계	97,856	97,856	97,856	97,856	97,856

RFM 값을 이용하여 표시한 결과이다.

군집의 결과표에서 확연하게 구별되는 군집은 'R↓F↑M↑' 군집과 'R↑F↓M↓' 군집, 'R↓F↓M↓'군집이다. 'R↓F↑M↑' 군집은 군집 평균 R값이 전체 평균보다 작고, F값이 전체 평균보다 크며, M값이 전체 평균보다 큰 군집이다. 이 군집은 최근에 자주 구매를 하며, 구매 금액도 비교적 큰 고객군집이므로 우수고객군집으로 분류할 수 있다. 비슷한 방식으로 R↑F↓M↓ 군집은 이탈가능고객군집으로, R↓F↓M↓ 군집은 신규고객군집으로 각각 분류할 수 있다. <표 5>는 분기별 전체 인원의 군집 현황을 요약한 것이다. 전반적으로 우수고객(R↓F↑M↑)의 수와 이탈가능고객(R↑F↓M↓)의 수가 증가하고 있음을 알 수 있다.

4.2 분기별 고객의 군집간 이동

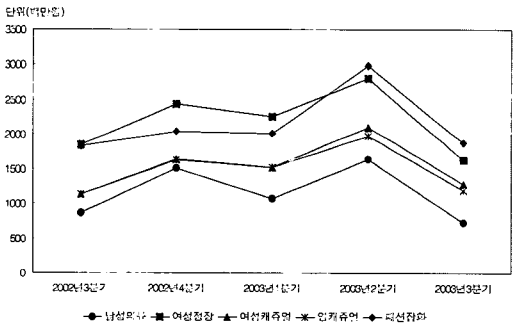
2002년 3분기의 97,856명의 고객이 다음 분기에 어떤 고객군에 속하게 될 것인지를 알아보기 위해 고객 97,856명을 분기별로 추적하여 군집간 이동 인원수의 변동 현황을 살펴보았다.

<표 6>을 보면 시간이 경과함에 따라 우수고객군집(R↓F↑M↑)과 이탈가능군집(R↑F↓M↓)에 속하는 고객의 수가 꾸준히 증가함을 알 수 있다. 이것은 2002년 3분기 고객 중 신규고객군집(R↓F↓M↓)에 속하던 고객이 꾸준히 우수고객군집과 이탈가능군집으로 이동하기 때문이다. 게다가 우수고객군집과 우수고객가능군집 (R↓F↓M↑)에 속하는 고객의 증가폭이 이탈가능고객의 증가폭보다 상회하여 신규고객의 우수고객으로의 전환비율이 이탈가능고객으로의 전환비율보다 월등함을 알 수 있다.

또, 우수고객군집에 속하게 된 고객은 다음 분기에도 우수고객군집에 속할 확률이 가장 높은 것으로 나타났는데 이를 통해 백화점에서 우수고객군집에 더 많은 우수 고객이 유지될 수 있도록 고객유지전략을 수립하고 마케팅프로모션을 집행하는 것이 필요하다. 한편으로, 이탈가능군집에 속하는 고객에 대해서도 동일한 분석을 하였는데 이탈가능군집에 속하는 고객은 이탈가능군집을 벗어나지 않고 머무르는 비율이 높았고 상품구매력이 지속적으로 감소하는 것으로 분석되었다. 즉, 한번 이탈가능고객은 이탈가능군집으로 이동할 확률이 가장 높으므로 이들에게는 이탈가능군집의 특성을 처음 보이는 시점에서 이탈가능군집에 유입되는 것을 방지하는 고객이탈방지전략과 과거 구매 상

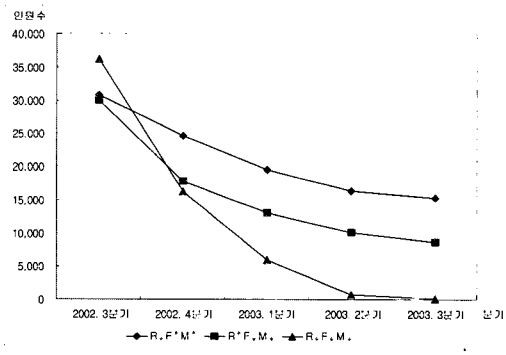
품군의 매출을 유도하는 고객재활성화전략이 필요하다.

<그림 4>는 2002년 3분기에서 2003년 3분기까지 우수고객으로 분류된 고객의 특성을 상위 5대 구매 품목 위주로 조사한 결과를 보여주고 있다. 이들 고객은 주로 30대~50대의 (93.3%) 남성(66.4%)으로 패션잡화, 여성캐주얼, 여성정장 등의 상품군에서 높은 매출을 보이는 경향을 가지고 있었다. 이들에 대해서는 패션잡화, 여성캐주얼, 여성정장의 상품군에 지속적인 관심을 가지도록 마케팅프로모션을 실시하여 지속적인 상품 구매를 유도하는 고객유지전략을 펼치는 것이 바람직하다.



〈그림 4〉 우수고객군집의 상품군별 매출 현황

2002년 3분기의 97,856명을 대상으로 각 군집에 속한 고객이 분기별로 동일 고객군집에 머무르는 현황을 <그림 5>와 같이 분석하였다. 2002년 3분기에 우수고객군집(R↓F↑M↑)에 속했던 고객 중 다음 분기에 동일 군집에 머무르는 고객은 25.1%로 감소하고 이후에 20.0%, 16.7%, 15.6%로 지속적으로 감소함을 알 수 있다. 이것은 시간의 경과에 따라 신규 유입에 따른 백화점 전체 우수고객의 수는 증가하지만 한번 우수고객이었던 고객은 더 이상 우수고객으로 남아있질 않고 점차 줄어든다는 것을 보여주는 것이다. 이탈가능군집(R↑F↓M↓)도 2002년 3분기 고객 중 30.6%에서 18.2%, 13.4%, 10.2%, 8.6%만의 고객이 이탈가능군집을 유지하였다. 이것은 고객재활성화전략의 효과로 우수고객군집 혹은 기타 군집으로 전환된 이탈가능 고객이 존재함을 나타낸다. 신규고객(R↓F↓M↓)은 2002년 3분기에 37.0%에서 2003년 3분기에 0.1%로 꾸준히 감소하여 여타 군집으로의 이동이 상당히 진행되었음을 보여주고 있다.



〈그림 5〉 동일 군집을 유지한 군집 그래프

4.3 고객군집 예측

고객의 군집간 이동을 동적으로 분석하기 위한 예측 모형을 만들기 위해서 97,856건의 고객자료를 학습용 자료 68,542건(70%), 테스트용 자료 29,314건(30%)으로 나누었다. 학습용 자료세트 로 예측모형을 생성하기 위해 C5.0을 사용하였는데 예측모형의 정확도를 높이기 위해 부스팅 기법을 사용하였고 가지치기의 정도를 조정하였다. 동일한 학습 방법을 모형 B, 모형 C, 모형 D, 모형 E, 모형 F에 적용하여 각각의 모형을 생성하였고 예측율이 가장 뛰어난 최적의 모형을 선정하였다. <표 7>의 정오분류행렬표는 모형 A의 군집 예측 정확도를 나타내고 있다.

<표 7>에서 알 수 있듯이 테스트용 자료 중에서 올바르게 분류한 신규고객군집(R↓F↓M↓)의 고객 1,757명, 우수고객군집(R↓F↑M↑)의 고객 7,085명, 이탈가능고객군집(R↑F↓M↓)의 고객 8,613명을 합한 17,455건을 전체 건수인 29,314건으로 나눈 비율을 테스트용 자료 세트에 대한 예측 정확도로 표시하였다. 즉, [(1,757 + 7,085 + 8,613)

/ 29,314 × 100 = 59.74%]로 모형 A의 예측 정확도를 계산하였다. 동일한 방법으로 모형 B~모형 F의 예측율을 계산하였고 결과는 모형 B(61.68%), 모형 C(71.27%), 모형 D(62.28%), 모형 E(71.38%), 모형 F(71.13%)로 나타났다. 6개의 모형 중 최적 모형을 찾기 위해 각 모형의 정오분류행렬표에서 정확도가 가장 높은 71.38%의 예측율을 보인 모형 E를 최적의 예측모형으로 결정하였다. <표 8>에 최적 모형 E의 규칙집합(Rule Set)을 나타내었다.

모형 E에서는 우수고객군집에 대해서 모두 16개의 규칙이 생성되었으며, 이탈가능군집에 대해서 7개의 규칙이 생성되었다. 우수고객군집에 대한 규칙 중 "*if* C == R↓F↑M↑ *and* B == R↓F↑M↑ *then* → R↓F↑M↑(14066, 0.881)"은 이전 분기에 R↓F↑M↑ 군집이며 현재 분기에도 R↓F↑M↑ 군집인 경우 다음 분기에 R↓F↑M↑ 군집으로 이동한 건수가 14,066건이고 이 규칙에 대한 신뢰도는 0.881의 값을 가진다는 의미이다. 이탈가능군집에 대한 규칙 중 "*if* C == R↑F↓M↓ *then* → R↑F↓M↓(21857, 0.743)"은 현재 R↑F↓M↓ 군집에 속하고 다음 분기에도 R↑F↓M↓ 군집에 속하는 경우는 21,857건이고 신뢰도는 0.743이라는 것을 의미한다.

<표 9>은 모형 형성 과정에서 정확도가 가장 높게 나타난 모형 E를 적용하여 2003년 4분기에 예측된 군집과 실제 군집을 비교하여 군집 예측 정확도를 실측한 결과를 보여주고 있다. 실제 군집은 4개이지만 모형 E는 2개의 군집을 예측하였고,

〈표 7〉 모형 A의 테스트용 자료세트에 대한 정오분류행렬표

테스트 자료 세트				
실제값 \ 예측값	R↓F↓M↓	R↓F↑M↑	R↑F↓M↓	합계
R↓F↓M↓	1,757	1,121	2,359	5,237
R↓F↓M↑	476	917	1,046	2,439
R↓F↑M↑	966	7,085	1,906	9,957
R↑F↓M↓	1,587	1,481	8,613	11,681
합계	4,786	10,604	13,824	29,314
정확도(%)	59.74%			

〈표 8〉 최적 모형의 규칙집합(Rule set)

```
Rule #1 for R↓F↑M↑ :
if C == R↓F↑M↑and B == R↓F↑M↑then → R↓F↑M↑ (14066, 0.881)
Rule #2 for R↓F↑M↑ :
if C == R↓F↑M↑and B == R↓F↓M↓then → R↓F↑M↑ (4707, 0.722)
Rule #3 for R↓F↑M↑ :
if C == R↓F↓M↑and B == R↓F↑M↑then → R↓F↑M↑ (2114, 0.705)
Rule #4 for R↓F↑M↑ :
if C == R↓F↑M↑and B == R↓F↓M↑then → R↓F↑M↑ (2113, 0.698)
Rule #1 for R↑F↓M↓ :
if C == R↑F↓M↓ then → R↑F↓M↓ (21857, 0.743)
Rule #2 for R↑F↓M↓ :
if C == R↓F↑M↑ and B == R↑F↓M↓ and A == R↑F↓M↓
                        then → R↑F↓M↓ (1636, 0.491)
Rule #3 for R↑F↓M↓ :
if C == R↓F↓M↓ and B == R↑F↓M↓ and A == R↑F↓M↓
                        then → R↑F↓M↓ (1354, 0.614)
Rule #4 for R↑F↓M↓ :
if C == R↓F↓M↓ and B == R↑F↓M↓ and A == R↓F↓M↓
                        then → R↑F↓M↓ (975, 0.515)
 중략
```

예측의 정확도는 우수고객군집(R↓F↑M↑)을 정확히 예측한 36,286명, 이탈가능군집(R↑F↓M↓)을 정확히 예측한 33,133명을 전체 인원인 97,856명으로 나눈 70.9%이다. 군집별로는 우수고객군집의 예측 정확도는 79.8%, 이탈가능군집의 정확도는 80.1%이다. 모형 E가 2개의 군집만 예측한 원인을 살펴보면 군집별 인원 비율이 높은 차이를 나타내고 있기 때문인데 구매력이 낮은 신규고객군집(R↓F↓M↓)은 4.2%, 구매력이 높은 신규고객군집(R↓F↓M↑)은 7.1%, 우수고객군집은 46.5%, 이탈가능군집은 42.2%로 상대적으로 신규고객군집의 비율이 낮다.

따라서 최적의 예측모형으로 선정한 모델 E의 경우 비율이 낮은 군집은 제거되면서 일반화된 모형이라 할 수 있다. 군집 인원이 많은 군집은 예측율이 높게 나타났고 군집 인원이 적은 군집은 예측 모형에서 규칙을 생성하지 못한 결과가 발생하였는데 결국 고객은 시간이 지남에 따라 우수고객군집과 이탈가능군집으로만 유입되는 것으로 예측되었다. 군집 예측율을 우수고객군집과 이탈가능군집만의 인원인 86,767명으로 측정하면 예측의 정확도는 80.0%를 나타낸다.

고객군집의 예측을 통하여 고객이 다음 분기에 어느 군집으로 이동하는지를 예측할 수 있다면 고객 패턴에 맞는 타켓마케팅이 가능하다. <표 10>은 모형 E를 적용하여 샘플 개별고객에 대한 2003년 4분기 군집 예측한 결과를 보여주고 있다.

우수고객집단에 속할 것으로 예측되는 고객들에게는 그들만을 위한 차별화된 서비스를 제공하여 고객 만족도를 향상시키고 그들과의 관계를 장기적으로 유지하는 전략을 펼쳐야 한다. 우수고객을 위한 마케팅 프로모션에는 백화점내에 VIP룸을 운영하고, 발렛파킹 서비스나 고객의 취미를 파악하여 각종 문화 행사 초청권을 제공하는 것뿐만 아니라, 생일 축하, 기념일에 차별화된 편지 발송이나 이메일을 통해서 우수 고객으로 유지 될 수 있도록 하여야한다.

이탈이 시작된 고객에게 이탈 방지를 위한 프로모션 활동으로 설문 조사 등을 실시하여 고객의 소리를 수집하고 다양한 고객 이탈 원인을 분석하여

〈표 9〉 2003년 4분기 군집 예측 결과

실제군집 \ 예측군집	R↓F↑M↑	R↑F↓M↓	합계	실제 군집의 고객 비율
R↓F↓M↓	2,102	2,009	4,111	4.2%
R↓F↓M↑	3,907	3,071	6,978	7.1%
R↓F↑M↑	36,286	9,165	45,451	46.5%
R↑F↓M↓	8,183	33,133	41,316	42.2%
합계	50,478	47,378	97,856	100%

〈표 10〉 고객별 예측 군집과 실제 군집의 비교

고객번호	2002년 4분기	2003년 1분기	2003년 2분기	2003년 4분기 (예측군집)	2003년 4분기 (실제군집)
A00000339	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑
A00000611	R↓F↓M↓	R↓F↓M↑	R↓F↓M↓	R↓F↑M↑	R↓F↑M↑
A00000894	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑
A00001318	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑
A00001681	R↑F↓M↓	R↓F↑M↑	R↑F↓M↓	R↑F↓M↓	R↓F↑M↑
A00002437	R↓F↓M↑	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↑F↓M↓
A00002718	R↓F↓M↓	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↑F↓M↓
A00003065	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑	R↓F↑M↑
A00003341	R↑F↓M↓	R↑F↓M↓	R↓F↑M↑	R↑F↓M↓	R↑F↓M↓
A00006240	R↓F↓M↓	R↓F↑M↑	R↑F↓M↓	R↑F↓M↓	R↑F↓M↓

이탈방지 전략 수립의 기초 자료로 이용하는 등 그에 상응하는 프로모션을 제공함으로써 이탈고객들에게 지속적으로 관심을 표현하여 이탈을 방지할 수 있도록 한다.

신규고객을 확보하기 위해 첫 구매 시 할인 혜택과 감사의 인사장을 발송하고, 재구매를 유도하기 위한 할인 구폰을 발송하거나, 구매 상품의 연관성을 분석하여 상품을 추천해주는 프로모션을 실시한다.

5. 결 론

본 연구는 시간의 변화에 따른 고객의 구매행위 패턴을 분석하여 동일 구매행위 패턴의 고객들을 군집화하고 미래의 고객군집의 이동 패턴을 예측하여 고객관계관리 관점의 마케팅프로모션 방안을 제시하였다. 기본적으로 고객 매출행위를 최근성, 구매빈도, 구매금액으로 분석하는 RFM분석방법을 채택하였고, 데이터마이닝 기법 중 신경망 분석 방법인 코호넨의 SOM을 이용하여 고객군집화를 하였다. 또한, 의사결정나무 분석 방법인 C5.0을 이용하여 고객의 군집간 이동을 예측하여 군집별로 우수고객군집의 예측율은 79.8%, 이탈가능군집의 예측율은 80.1%를 보였다. 실제 고객 자료를 바탕으로 고객의 향후 군집 이동을 예측하여 고객 군집간 이동에 대해 동적인 분석이 가능하다는 것을 실증적으로 보여주었다.

본 연구는 다음과 같은 의의를 갖는다. 첫째, 분기별로 고객 군집 현황을 파악하여 시간이 지남에 따라 고객 군집의 증감을 확인할 수 있었다. 우수고객군집의 증감 추세, 이탈가능군집의 증감 추세, 신규고객군집의 증감 추세, 신규고객의 유입 추세를 파악하여 우수고객군집을 유지하고 이탈가능군집으로의 유입을 줄이는 기업의 장단기적인 비즈니스 전략을 제시하였다.

둘째, 우수고객은 우수고객군집을 유지하는 확률이 높으므로 고객유지전략의 마케팅프로모션을 실시하는 것이 타당하고, 이탈가능고객이 이탈가능군집으로 접어들기 전에 이탈을 미연에 방지할 수 있는 적절한 시점에 개별 고객의 자료를 제공할 수 있도록 하였다.

넷째, 미래의 고객 군집의 패턴을 예측함으로써 우수고객군집에 대한 고객유지전략, 신규고객군집에 대한 고객성숙전략, 이탈가능고객군집에 대한 이탈방지전략 등 고객별로 선별적인 타켓마케팅을 구사할 수 있도록 지원하여 고객 접근 비용을 최소화하고 기업의 이익을 최대화할 수 있는 고객관계관리 전략으로 활용할 수 있도록 하였다.

본 연구의 향후 연구 방향은 다음과 같다.

첫째, 현재까지 RFM 분석은 고객 구매 행위를 분석하는 방법 중에서 가장 효율적인 방법이었지만 앞으로의 연구에서는 RFM 값을 포함한 여러 변수와 새로운 군집 분류 기준을 적용하여 군집 변화를 예측할 수 있도록 확장하는 연구가 필요하다.

둘째, 상품군별로 세분화된 RFM 분석을 통하여 고객군을 더욱 세분화하여 해당 상품군에 적합한 고객을 찾아내고 고객들에게 개별적인 프로모션을 할 수 있도록 한다.

셋째, 예측모형을 통한 군집의 프로모션 반응과 예측모형을 통하지 않은 고객 선별 방법의 프로모션 결과를 비교하여 예측 모형의 마케팅 성과를 증명하고자 한다.

넷째, 고객관계관리는 온라인 기업이나 오프라인 기업 모두에게 중요한 경영 전략으로 인식되고 있으므로 향후 e비즈니스로 확대 적용하여 온라인과 오프라인에서의 적용 결과를 비교하는 연구가 필요하다.

참 고 문 헌

[1] 곽재민, "통신판매업체의 데이터베이스 마케팅에 관한 연구", 동국대학교 석사학위논문, 2001.

[2] 김영아, "시장세분화를 위한 데이터마이닝 응용에 관한 연구", 서강대학교 석사학위논문, 2001.

[3] 김종우, "가상상점에서 고객 행위 연관성 분석을 위한 데이터마이닝 기법", 「한국전자거래학회지」, 제4권, 제1호(1999).

[4] 박병훈, "RFM 점수를 이용한 고객 반응 분석," 중앙대학교 석사학위논문, 2002.

[5] 우근식, "백화점에서 데이터마이닝을 이용한 고객 충성도 분석에 관한연구", 동국대학교 석사학위 논문, 2000.

[6] 최정환, 이유재, 「죽은 CRM 살아있는 CRM」, 한언, 2001.

[7] Anderson, J., *Customer Relationship Management*, Prentice Hall, 1999.

[8] Adriaans, P. and D. Zantinge, *Data Mining*, Addison-Wesley, 1996.

[9] Berry, M.J.A. and G.S. Linoff, *Mastering Data Mining : The art and Science of Customer Relationship Management*, John Wiley & Sons, 2000.

[10] Berson, A., S. Smith and K. Thearling, *Building Data Mining Applications for CRM*, McGraw-Hill, 1999.

[11] Ciesielski, V. and G. Palstra, "Using a Hybrid Neural/Expert System for Data Base Mining in Marketing Survey Data," *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, 1996, pp.38~43.

[12] Ha, S.H., S.M. Bae and S.C. Park, "Customer's Time-Variant Purchase Behavior and Corresponding Marketing Strategies : An Online Retailer's Case" *Computers & Industrial Engineering*, Vol.43, No.4(2002), pp.801~820.

[13] Han, J. and M. Kamber, *Data mining concepts and techniques*, Morgan Kaufmann Publishers, 2000.

[14] Hughes, A.M., *Strategic Database Marketing*, 2nd ed., McGraw-Hill, New York, 2000.

[15] Jutal, D., J. Craig and P. Bodorik, "Enabling and measuring electronic customer relationship management readiness," *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, 2001.

[16] Kiang, M.Y., R.T. Chi and K.Y. Tam, "A Decision Support System for Market Segmentation - A Neural Networks Approach," *Proceeding of Americas Conference on Information System*, 1995.

[17] Kohonen, T., T.S. Huang and M. R. Schroeder, *Self-Organizing Maps*, Springer Verlag, 2000.

[18] Loh, W.Y. and Y.S. Shin, "Split selection methods for classification trees," *Statistica Sinica*, Vol.7(1997), pp.815~840.

[19] Mangiameli, P., S.K. Chen and D. West, "A Comparison of SOM Neural Network and Hierarchical Clustering Methods," *European Journal of Operational Research*, Vol.93, Iss.2(1996), pp.402~417.

[20] Mitchell, T.M., *Machine Learning*, McGraw-Hill, 1997.

[21] Mulier, F. and V. Cherkassky, "Self-Organization as an Iterative Kernel Smoothing Process," *Neural Computation*, Vol.7(1995), pp.1165~1177.

[22] Peppard, J., "Customer Relationship Management in Financial Services," *European Management Journal*, Vol.18(2000), pp.312~327.

[23] Quinlan J.R., *C4.5 : Programs for Machine Learning*, Morgan Kaufmann Publishers, 1993.

[24] Ryals, L. and S. Knox, "Cross-functional Issues in the Implementation of Relationship Marketing through Customer Relationship Management," *European Management Journal*, Vol.19(2001), pp.534-542.

[25] Sharp, D, E., *Customer Relationship Management System Handbook,* Auerbach Publications, 2003.

[26] Tam, K.Y. and M.Y. Kiang, "Managerial Applications of Neural Networks : The Case of Bank Failure Predictions," *Management Science*, Vol.38(1992), pp.926~947.