

# 자동차산업의 고객분류 및 타겟 마케팅 모델

## Model of Customer Classification Target Marketing in Automotive Corporation

이병엽\*, 박용훈\*\*, 유재수\*\*\*

배재대학교 전자상거래학과\*, 충북대학교 정보통신학과\*\*,  
충북대학교 전기전자컴퓨터공학부\*\*\*

Byoung-Yup Lee(bylee@pcu.ac.kr)\*, Yong-Hoon Park(yhpark@netdb.chungbuk.ac.kr)\*\*,  
Jae-Soo Yoo(yjs@chungbuk.ac.kr)\*\*\*

### 요약

최근 정보기술이 발달하면서 수많은 자료들이 체계화된 데이터베이스에 저장되고, 기업의 데이터베이스의 규모는 폭발적으로 증가되고 있다. 데이터 마이닝(Data Mining)은 이런 방대한 자료의 분석을 통해, 그 속에 숨어있는 의미를 찾는 과정이라고 볼 수 있다. 본 논문에서는 자동차 마케팅에서 이용 가능한 데이터를 데이터 마이닝 분석 기법중의 하나인 연관규칙(association rule)에 따라 분석하였다. 본 논문에서 제시하고자 하는 바는 기존 고객에 대한 분류 및 고객 속성과약, 고객 분류 및 분석에 따른 고객의 연관규칙을 수행하여 해당 기업의 전략적 마케팅 수립을 통해 경영 과학적으로 접근할 수 있는 데이터 마이닝 분석에 관한 사례 연구이다. 본 논문의 분석 사례를 통하여 자동차 분야의 특성에 따라 효과적인 분석 및 의사결정과 더 나아가 CRM마케팅, 동향분석 및 예측 등에 유용한 정보를 분석할 수 있는 사례로 활용될 수 있다.

■ 중심어 : | 데이터 마이닝 | 분류 규칙|연관규칙 |

### Abstract

Recently, According to computer technology has been improving, Massive customer data has stored in database. Using this massive data, decision maker can extract the useful information to make a valuable plan with data mining. Data mining offers service providers great opportunities to get closer to customer. Data mining doesn't always require the latest technology, but it does require a magic eye that looks beyond the obvious to find and use the hidden knowledge to drive marketing strategies.

Automotive market face an explosion of data arising from customer but a rate of increasing customer is getting lower. therefore, we need to determine which customer are profitable clients whom you wish to hold. This paper builds model of customer loyalty detection and analyzes customer patterns in automotive market with data mining using association rule and basic statistics methods. With the help of information technology.

■ keyword : | Data Mining | Classification | Association Rule |

## I. 서론

2000년대 초반부터 시작된 기업의 전산화에 따른 기

업의 고객 데이터는 가히 기하급수적으로 축적되었다. 이에 따라 기업의 최근 데이터 수집 및 저장기술의 발달, 데이터베이스 관리시스템과 데이터웨어하우징 기

술의 광범위한 사용은 기업내부의 대량의 데이터를 축적할 수 있도록 하고 있으며, 축적된 데이터는 의사결정에 필요한 새롭고 가치 있는 정보와 지식을 획득할 수 있는 잠재적인 원천으로 인정되고 있다[26]. 기업들은 급변하는 경영환경에서 기업의 경쟁력을 강화하기 위해서 축적된 데이터를 분석하고 정보와 지식을 획득할 수 있는 능력과 정보기술을 보유해야 한다[8].

현재 우리나라는 2007년 기준 한국 자동차산업 생산량 409만대로서 일본, 미국, 중국 등에 이어 세계5위의 자동차 생산국이다. 최근 글로벌 경쟁 심화와 업계판도의 변화, 시장구조의 변화가 예상되지만 향후 내수는 경기둔화, 고용불안정, 고유가 지속등의 요인들이 크게 개선되기 어려울 것으로 보이나, 차량 노후화에 따른 대체 수요 발생, 신차 효과등이 긍정적인 요인으로 작용할 것으로 예상된다. 하지만 자동차 산업은 한미 FTA 및 여러 가지 어려운 글로벌한 시장과 더불어 자동차 생산 및 내수의 보급이 급격하게 증가 되고 있는 추세이고, 현재 우리나라 자동차 보급률은 IMF외환위기 이전 150만대 수준이던 내수시장은 외환위기 이후 '02년 160만대 수준까지 증가하였으나 최근 고유가 등으로 인한 내수 시장 침체로 인하여 부진한 상황을 면치 못하고 있다. 이에 본 논문은 데이터 마이닝을 통한 고객 분석을 통하여 자동차 산업의 고객 마케팅의 좋은 분석 방법으로 활용하기를 기대한다.

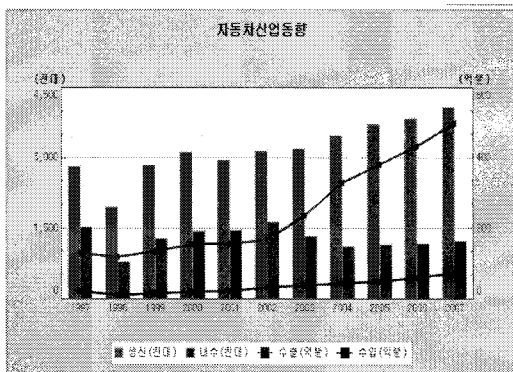


그림 1. 자동차 산업 동향

국내 자동차 보급률은 07년 기준 자동차 1대당 인구 수 3.0명, 1세대당 자동차 보급대수 0.88대로서 2014년

경 자동차등록 대수는 2,000만대에 도달할 것으로 예측되며 이에 자동차 업계는 기업간 보유고객의 수가 늘어남에 따라 고객 데이터와 정보의 양도 기하급수적인 증가를 보여 왔으며 이러한 데이터와 정보의 증가는 데이터 활용에 대한 방법과 그 영역의 증대를 가져왔다[13]. 따라서 자동차 산업의 경쟁방식은 기존고객의 유지, 서비스의 차별화와 고도화를 통해 이루어져야 하며 신규 고객 1명 유치에 드는 비용은 기존고객 1명을 유지하는데 드는 비용보다 4-6배 더 들게 되는 것을 감안한다면 양적 성장인 신규고객 유치보다는 보유고객에 대한 다각적인 분석에 따른 효율적인 마케팅 방법 적용으로 심화되는 자동차 완성 업체간의 경쟁에서 우위를 차지할 수 있을 것이다[18].

본 논문에서는 자동차 완성업체에서 이용 가능한 전산 데이터를 이용하여 데이터마이닝 방법론 중의 하나인 연관규칙을 이용하여 기존고객을 세분화한 다음 고객 개인의 특성에 맞는 마케팅 프로모션을 하게하고 신규고객을 획득할 때는 신규고객의 특성을 미리 예측하여 세분화함으로써 고객의 평생가치를 촉진하여 기업과 고객과의 관계성을 높여서 기업은 안정된 고객층으로부터 수익을 창출하고 고객들은 해당 기업으로부터 더 많은 혜택을 받게 하는데 목적이 있다. 그러므로 본 논문에서 제시하고자 하는 것은 기존 고객에 대한 연관규칙, 및 고객속성 파악, 기존 고객 분류에 따른 신규고객의 분류 예측 두 가지를 수행하여 해당 기업의 마케팅 전략수립에서부터 경영 과학적으로 접근할 수 있는 데이터마이닝 분석 방법을 제시하고자 한다[26].

본 논문은 국내 자동차 완성 업체인 A사의 자료를 입수하여 실제 데이터를 중심으로 연구하였다. 또한 SAS를 이용하여 고객 data에서 변수 추출 및 정제작업을 실시하였다. 데이터 마이닝은 SAS사의 데이터 마이닝 툴인 Enterprise Miner를 이용하여 연관규칙에 의한 고객그룹속성 및 고객 그룹에 대한 분류 및 유의한 연관 규칙을 제시하였다.

본 논문의 구성은 I장 서론 에서는 연구 목적, 연구 범위와 방법을 제시, 데이터 마이닝의 개념과 CRM, 데이터 마이닝 응용분야 등을 살펴보고, 본문에서는 자동차 완성 업체의 사례 연구 및 본 연구에서 사용한 데이

터 마이닝 기법들에 대한 세부적인 내용을 기술하였으며 연관규칙의 결과를 살펴보고 의미를 분석한다. 결론에서는 연구결과를 종합하고 미래의 연구방향을 제시하였다.

## II. 데이터마이닝(Data Mining)

CRM(Customer Relationship Management)은 고객 전략을 정의하고, 경영자원의 분배를 최적화하며, 고객과 관련된 모든 부분에서 최상의 서비스를 제공하기 위해, 고객의 욕구와 행동, 고객의 수익성을 규명해 가는, 지속적이고 자동화된 프로세스로서 CRM은 고객과 관련된 기업의 내외부 자료를 분석, 통합하여 고객특성에 기초한 마케팅활동을 계획하고, 지원하며, 평가하는 과정이다. 또한 CRM은 데이터 마이닝을 통해 고객데이터의 세분화를 실시하여 신규고객획득, 우수고객 유지, 고객가치증진, 잠재고객 활성화, 평생 고객화와 같은 사이클을 통하여 고객을 적극적으로 관리하고 유도하며 고객의 가치를 극대화 시킬 수 있는 전략을 통하여 마케팅을 실시한다. 이러한 CRM의 정의로 볼 때 CRM에서의 데이터 마이닝은 필수 불가결한 도구라고 볼 수 있다[15][24].

2000년대에 접어들어 비즈니스 세계에서, 고객중심의 시각에서 모든 것을 바라보며, 그것에 맞추어 경영 전략을 수립하는 일들이 이제 거의 상식의 관점에서 이루어지고 있다. 이러한 관점에서 비즈니스의 당위 목적들이 제품 또는 서비스의 단순한 판매에만 있는 것이 아니라 고객과 관계를 맺어가는 과정임을 인식하는 것이라고 볼 수 있다. 이 CRM은 크게 2가지 영역의 형태로 나누어 볼 수 있다. 첫 번째는 기능적 CRM, 다음은 분석적 CRM으로 나누어 볼 수 있다[15][24]. 기능적 CRM을 요약하면 비즈니스에서 고객과의 관계를 기능적인 단위 측면에서 초점을 맞춘 경우를 의미한다. 예를 들면 모든 고객들의 접점에서 어떤 유용한 특정 고객의 상호 작용에 대한 정보만을 생성하여 이를 활용하는 경우를 말한다. 분석적 CRM은 개인적인 측면과 전체적인 측면에서 고객의 이해를 더욱더 높여줄 수 있는

정보를 제공한다[21]. 어떤 전략적인 측면에서든지 고객 개개인에게 발생하는 모든 수준의 상황에 대해서 비즈니스적인 요구만 있다면 언제든지 정보를 제공할 수 있는 경우를 의미한다.

따라서 자동차산업에서 데이터 마이닝과 CRM의 활용을 요약 하면 다음과 같다. 첫째, 고객 세분화와 고객가치의 이해증대로 더 많은 고객의 획득이 예상된다. 특정한 고객과 집단의 세분화와 가치의 극대화는 타겟 마케팅의 캠페인의 목표가 될 수 있다. 둘째, 고객 세분화와 신규 자동차 구매간의 관계에 대한 더 나은 이해로 교차 판매의 효율이 증대 될 수 있다. 셋째, 언제 그리고 왜 고객들이 이탈을 하는지, 그리고 어떤 경우에 긴급한 대처 방안은 무엇인지에 대한 개념 파악으로 우수고객의 유지율의 향상을 도모 할 수 있다. 이러한 요약된 개념들은 데이터 마이닝 또는 분석적 CRM의 기법에서 얻어진 결과로서 운영 CRM에 의해 활용 될 수 있다[15].

데이터 마이닝은 각각의 산업의 특성에 따라 표준화된 분석방법과 사례들이 있다. 본 논문에서 제시하는 데이터 마이닝 분석자들이 새로운 모형을 구축하는데 소요되는 노력을 최소화 할 수 있고, 더 나아가 기존 고객 및 신규 고객관리에 측면에서 마케팅 자료로 활용 될 수 있다.

## III. 분석의 절차 및 변수 선정

### 3.1 연구 절차와 방법

본 연구에서는 국내 완성자동차 업체인 A사의 자료를 입수하여 연구하였고 수행하는 연구의 진행절차는 데이터 마이닝 프로세스의 각 단계별 과정에 따라 진행을 하고자 하며 과정은 첫째 고객자료 수집 및 통합, 둘째 수집 자료에 대한 추출 및 정제 작업, 셋째 변수의 선정, 넷째, 분류에 의한 고객 그룹속성 분석 및 고객그룹에 대한 연관규칙의 분석의 순서로 진행 되었다.

### 3.2 자료의 구성

본 연구에서 사용한 데이터는 국내 완성형자동차 업

체의 샘플자료로써 소비자구매 행동에 대한 일반적인 데이터 즉 구매시 자동차 비교 여부, 구매시 비교차종, 비교사유, 현소유차 구매 사유 등이며, 자동차에 대한 만족도를 위한 항목은 전반적인 만족도, 소유차 제조회사 이미지 개선도, 이미지 적합도, 개인 속성별 만족도, 연비/유지비 평가 등이고 재구매 측정을 위한 항목은 소유차 재구매 의도, 소유차 제조회사 재구매 의도로 구성되었으며 과거 200X년 9월부터 10월까지 발생한 신차 구입자를 대상으로 이루어진 설문을 대상으로 하였다. 데이터를 이루고 있는 고객의 수는 3,159명으로 총 고객에서 샘플링을 하여 선정하였고, 조사 결과중 null 값은 평균값으로 일괄 처리, 또는 데이터 마이닝 알고리즘에 적합한 값으로 변환 하였다. 따라서 대체 평균값이나, 기타 값들에 의해서 추출된 데이터 마이닝 룰은 분석/해석 대상에서 제외를 하였다. 반면 무응답의 경우도 일정한 패턴의 규칙성을 가질 수 있고, 의미 있는 결과를 가질 수도 있지만 분석에서는 일괄적으로 무응답인 경우는 룰 생성에서 제외하였다. 최초 원시 데이터로부터 고객의 기본정보의 데이터 셋(background), 고객 성향들의 유의한 값을 가진 변수들의 집합, 그리고 이 두 데이터들을 합한 데이터를 최종 연관규칙 데이터 마이닝 모델에 적용하였다.

### 3.2 기초 통계 분석

데이터 마이닝 알고리즘을 적용하기 위해서는 데이터에 대한 사전처리가 필요하다. 표본의 각 사례의 특정 변수에 누락된 데이터가 있는지 확인하여 누락된 데이터에 대해서는 변수의 평균값, 중위수 혹은 중간값 등으로 결측치를 대체시킨다. 만약 결측치를 지닌 변수가 이산형이라면 해당 변수의 최빈값으로 대체하는 것이 일반적이다. 또한 중복된 레코드를 제거하고 변수의 값이 일관성이 없으면 이를 바로 잡아야 한다. 잠정적으로 선택된 많은 변수 중에서 원하는 모형에 가장 적합한 변수를 선정해야 하는데 선정방법에는 대표적으로 입력변수와 목표변수간의 결정계수를 이용하는 방법과 카이제곱 통계량을 이용하는 방법이 있다[1]. 또한 기초통계 분석을 함으로써 원시데이터 변수들 중 유의한 변수들을 정의하고 효과적이고 이상적인 마이닝 분

석 결과를 얻기 위한 과정으로서 주로 분석한 통계분석은 mean, freq, correlation, ttest, cronbach's alpha 등을 분석하였다.

#### 3.2.1 평균값

설문 데이터중 특별히 정리를 요하는 고객의 기본데이터, 유의한 확률 통계치를 가지고 있는 변수들의 평균값을 살펴보면 다음 [그림 2]와 같다.

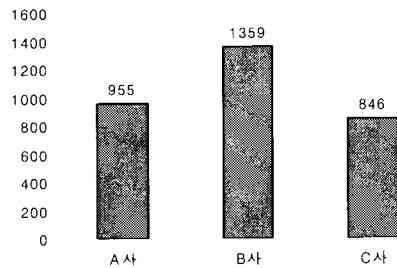
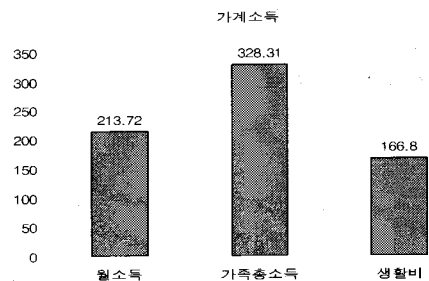


그림 2. 자동차 회사별 샘플수

가계소득에 관한 평균값을 살펴 보면 평균 328만원 가량의 총 가계소득을 살펴볼수 있고, 평균연령은 30대,20대의 순으로 조사되었고, 가족수는 4명의 가족구성원이 가능 높은 평균값의 분포를 보였으며 직업으로는 전문직, 공무원, 자영업/회사원의 순으로 조사되었다. 평균값중 특이한 사항은 차량구입자중 전체 3159명 중 신문을 구독하는 구성이 2,447(78%) 임을 알수있다. 이는 중요한 데이터 마이닝의 변수로 도출될 수 있고 신문의 구독여부에 따른 새로운 데이터 마이닝의 연관 규칙을 예측할 수 있다.



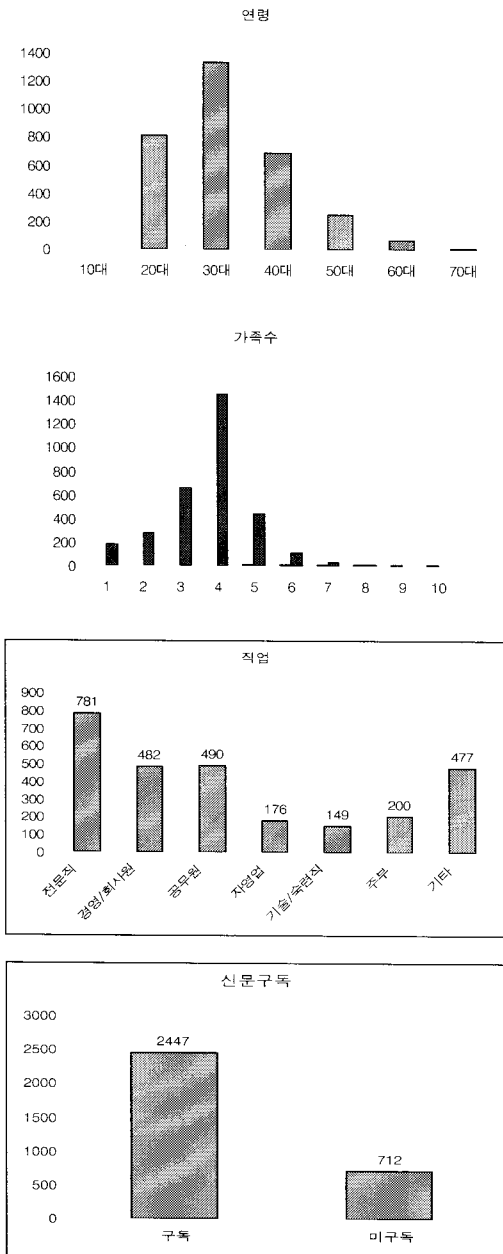


그림 3. 샘플의 평균값

### 3.2.2 유의변수 도출

유의한 변수를 도출하기 위하여 선택된 변수들의 대한 상관관계(correlation)를 알아보기 위하여 분석한 결과 [표 2]의 결과를 볼 수 있다.

표 2. 변수의 Correlation

변수	Correlation Coeff	변수인수
차명 VS 마지막고려차종	0.7252	2385
고려차종 VS 마지막고려차종	0.92082	2385
소유차종 VS 고려차종	0.75543	2385
소유차종 VS 마지막고려차종	0.72874	2385
마지막고려차종 VS 소유차종	0.72874	2385
조작용이성 VS 조작편의성	0.79233	3159
가족총소득 VS 생활비	0.70127	3078
연령 VS 가족형태	0.77786	3159

\* Prob > 0.0001, Correlation Coeff < 0.7

[그림 3][표 2]의 평균값과 상관관계(Correlation) 분석을 통하여 간단히 정리되기 어려운 변수들이 있었지만 중요 변수들과 통계분석 결과 얻어낸 유의한 변수들을 정리하면 [표 3]와 같다.

표 3. 산출된 유의 생성 변수

구분	변수
고객배경정보	성별, 연령, 가족내에서의 위치, 가족형태, 직업, 회사 규모, 집소유, 가족수, 결혼, 맞벌이, 신문구독, 월소득, 가족총소득, 생활비
구매 경험	차명, 배기량, 차량타입, 변속기, 새 차 구입이유, 차량구입시 가장 중요하게 생각하는점, 차량 선택시 의사결정, 구매비교차종, 현재소유차종, 고려차종, 마지막고려차종, 마지막고려 이유, 만족도, 회사이미지, 안전성, 성능, 소음, 외부 디자인, 내부디자인, 조작편의성, 용도(출퇴근, 등하교, 쇼핑용, 데이트, 업무/집대, 여행용, 종교활동, 친구/친척방문, 기타)

## IV. 문제정의 및 모델링 분석

문제정의의 과정에서 데이터의 특성을 나열하고 데이터 마이닝의 문제정의 제약 조건을 정의 하면 다음과 같다. 첫째, 데이터 마이닝의 문제를 해결하는 가장 일반적인 5가지(Regression, Association, C4.5, Episode, Sequence) 알고리즘이 있다. 이러한 각각의 알고리즘 별로 데이터의 일정한 형식의 타입과 변수들이 필요한데, 현재 분석을 하기위해 보유하고 있는 로우 데이터의 특성을 살펴보면 고객의 의견이나 생각을 묻는 설문조사 형식의 자료들로서, 고객의 거래 데이터 변수들이 결여되어 있어 적용할 수 있는 알고리즘이 극히 제한적

이다. 앞서 언급한 이유 이외에도 CRM 또는 마이닝에 대한 해석은 고객의 거래 데이터에 대한 분석이 주를 이루고 있고, 또한 고객의 성향을 알 수 있는 데이터들이 부족해서 알고리즘 적용시 특별한 룰들을 얻어내기가 어렵다. 따라서 제한 조건을 벗어난 데이터 마이닝의 문제점들을 도출한 결과 Association Rule을 통한 고객의 성향 분석, 전체적인 데이터의 성격들의 결과를 도출할 수 있었다.

#### 4.1 Association Rule 문제 정의

연관성 규칙은 한 항목 그룹과 다른 항목 그룹 사이에 존재하는 강한 연관성을 찾아내는 마이닝 기법으로 적용된 연관규칙은 Agrawal의 Apriori 알고리즘을 사용하였고 implication의 형태로 표현되고 표현식은 다음과 같다.

- Form:  $A \rightarrow B$  'A implies B' ('A이면 B이다')

연관규칙의 파라미터에서 지지율과 유의도를 설정할 수 있는데 각각히 설명하자면 지지율이란 전체 로우 데이터 중에서 어떤 일정한 규칙 중 이에 해당하는 ID의 수를 가리키고, 유의도는 이러한 룰의 ID중 분석에 설정된 유의도에 맞게 룰들을 생성한다. 분석 고객 데이터들의 변수들간의 상호연관성을 통한 데이터 전체의 고객성향의 인지 및 새로운 변수들간의 룰들을 생성하고, 이에 맞는 CRM 측면에서의 접근이 가능수이다. 따라서 적용 알고리즘 문제 정의는 분석된 유의 변수를 토대로 다음과 같은 연관규칙의 문제를 정의 하였다.

- 첫째, 소유차종에 따른 고객의 성향은 무엇인가? 타겟 마케팅을 할 수 있는 유의한 변수는 무엇인가?
- 둘째, 신규고객, 재 구입 잠재고객 및 새로운 DM 발송 그룹들의 선별
- 셋째, 잠재 자동차 구매 고객의 만족도를 통한 재 구매 고객 선별로 문제를 정의 하였다.

#### 4.2 Association 분석 및 결과

정제된 데이터를 가지고 Association Rule 을 적용해

본 결과를 정리하면 다음과 같다. 데이터 마이닝 분석 과정은 먼저 앞서 정의한 데이터들을 각각 분석하고 유의한 변수들과 타겟 변수들과의 마이닝을 통해 얻어진 결과이다. Association Rule의 분석절차는 적절한 지지율과 유의도를 알아내기 위해 각각의 지지율과 유의도 별로 데이터 마이닝 결과를 비교하여 가장 적합한 룰이 나오는 조건(Support: 20%, Confidence: 60%)로 정의하여 아래와 같은 룰들을 생성하였다. Association Rule은 적절한 조건이 주어지지 않으면 적게는 몇 십개에서 많게는 몇 만개 이상의 룰들이 생성되어 Association Rule이 적절한 파라미터의 조건과 이에 맞는 변수들의 선정이 가장 중요한 문제이다. 따라서 본 알고리즘을 적용 하기전 선행된 분석에 맞는 변수들의 통계적인 수준을 분석하고 이에 맞는 변수들을 선정하여 Association Rule을 적용하였다. 생성된 각각의 룰들의 의미와 비교가 가능한 자동차 3사의 룰들을 비교하였고 생성된 룰들은 다음과 같다.

##### (1) 타겟마케팅 측면에서의 고객 성향 분석

타겟마케팅 측면에서 적용변수를 살펴보면 성별, 연령, 본인의 가족위치, 가족형태, 직업, 직급, 회사규모, 집소유, 가족수, 결혼, 맞벌이, 신문구독, 정기신문구독, 자주 읽는 신문, 일요일 신문 보는 시간, 즐겨 보는 순서, 소유차량가격, 월 소득, 가족총소득, 생활비를 적용하였고, Confidence 60% <, Support Mean 20% <의 유의한 룰들을 정리하면 다음과 같다.

- Rule(Support:1065(0.337), confidence:0.784) : 현재 보유차종이 A자동차사 인 사람은 신문구독을 하고 있다.
- Rule(Support:743(0.234), confidence:0.779) : 현재 보유차종이 B자동차사인 사람은 신문구독을 하고 있다.
- Rule(Support:639(0.337), confidence:0.755) : 현재 보유 차종이 C자동차사인 사람은 신문을 구독하고 있다.

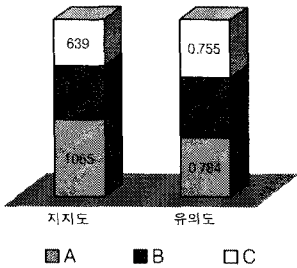


그림 4. 소유차종에 따른 신문 구독율

위의 유의한 결과 이외에 물을 요약하면 대부분의 자동차를 소유하고 있는 사람들은 기존남성으로서 집에서 위치는 가장의 형태를 띠고 있음을 알 수 있다. 또한 위의 자동차 3사별 신문 구독의 측면에서 분석해보면 지지율이나 유의도에서 A자동차를 소유한 고객이 신문 구독율이 가장 많음을 알 수 있다. 이에 맞는 CRM 전략을 수립하면 CRM은 고객만족과 고객의 요구조건을 관리하고 충족시킴으로써 회사에서는 이익을 극대화하고 회사의 이미지 상품의 이미지를 극대함에 기본 조건을 가지고 있다고 가정하고, 일정한 광고를 통해 신차에 대한 광고를 전달한다면, 위의 분석의 전체적인 면을 고려해 대부분의 차량을 구입한 사람들은 기존의 남성이고 일간신문을 반드시 구독하는 것으로 조사되어 졌고, 좀더 세부적으로 일간신문을 보는 순서가 정지면, 경제면, 즐겨보는 순서등을 고려할 때 신차 출시나 마케팅은 신문의 광고는 정치면이나 경제면에 자동차 광고를 내는 것이 분석의 결과 좀 더 나은 마케팅 효과를 가지고 올 수 있다고 할 수 있다.

(2) 잠재 신규고객, 재 구입 잠재고객 그룹의선별

잠재 및 신규고객, 재 구입 잠재고객 그룹의 선별의 측면에서 적용된 변수들을 살펴보면 제조회사, 차명, 배기량, 차종, 구매비교차종, 마지막고려차종, 구입이유, 만족도, 좋은점, 재구입, 재구입이유, 성별, 연령, 가족형태, 직업, 직급, 회사규모, 가족수, 맞벌이, 월소득, 가족총소득, 생활비를 적용하였고 Confidence 60%<, Support Mean 20%< 의 유의한 물들을 정리하면 다음과 같다.

- Rule(Support:756(0.239), Confidence:0.639) : 현재 소유 차종이 1500cc인 자동차를 소유하고 있는 사람은 재 구매시점에서도 1500cc 자동차를 구매할 것이다.
- Rule(Support:1019(0.323), Confidence:0.75) : 제조 회사가A인 경우 소유하고 있는 자동차 구입시 비교차종이 있었다.
- Rule(Support:735(0.233), Confidence:0.77) : 제조회사가 B인 경우 소유하고 있는 자동차 구입시 비교차종이 있었다.
- Rule( Support:631(0.200), Confidence:0.746) : 제조 회사가 C인 경우 소유하고 있는 자동차 구입시 비교차종이 있었다.

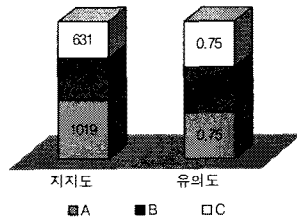


그림 5. 완성업체별 지지도 및 유의도

위와 같은 결과를 보고 유추하자면 국내 3개 자동차 회사차를 구입한 사람들 중 B사나 C사자동차를 구입하는 사용자들보다 A사차를 구매하는 사람들이 소유하고 있는 자동차를 구매하기 전 비교차종이 더 많았다는 것을 알 수 있다. 이러한 고객들을 잠재고객으로 설정하고 지속적인 관리를 함으로써 다음 자동차 구입시 비교 대상 차종의 범주에서 좀더 유리한 결과를 가져 올 것으로 사료 된다.

(3) 잠재 자동차 구매 고객의 만족도를 통한 재 구매고객 선별

잠재 자동차 구매 고객의 만족도를 통한 재 구매 고객 선별의 적용문제에서 적용된 변수를 살펴 보면 제조 회사, 차명, 배기량, 차종, 구매비교차종, 마지막고려차종, 구입이유, 만족도, 좋은점, 재구입, 재구입이유, 성

별, 연령, 가족형태, 직업, 직급, 회사규모, 가족수, 맞벌이, 월소득, 가족총소득, 생활비를 적용하였고 유의한 물들을 정리하면 다음과 같다.

- Rule(Support:669(0.212), Confidence:0.623) : 차량 구입시 비교차종이 있고, 현재 소유하고 있는 자동차에 만족도가 만족스러운 사람은 현재 구입한 차를 다시 구매한다.
- Rule(Support:631(0.200), Confidence:0.746) : 현재 소유차에 대한 만족도가 "만족"인 사람 중 재구매시 현재 구입차를 다시 구입할 의사가 있는 남자는 반드시 재구매 의사가 있다.
- Rule(Support:901(0.285), Confidence:0.623) : 소유 자동차의 만족도가 "만족"인 사람은 차량구입당시로 돌아가도 현재 승용차를 재구매 할 것이다.

위의 결과를 해석하면 현재 소유차에 대한 만족도가 만족이라고 대답한 사람과 재 구매의사라고 한 남성 중 A사차를 소유한 고객그룹 생성하면 그 그룹의 사람들은 반드시 재구매 의사를 가지고 있다고 해석 가능하다. 따라서 A차의 구매고객 중 현재 소유차의 만족도가 만족인 고객을 대상으로 타겟마케팅을 한다면 보다 좋은 매출 효과를 가져 올 수 있다.

## V. 결론

완성 자동차 시장에서 기업간 경쟁이 더욱 심화되고 최근 글로벌 경쟁 환경에서 급격한 내수 감소 추세를 나타냄에 따라 기업이 마케팅 활동을 수행함에 있어 기존 고객의 유지가 중요한 과제가 되고 있다[19]. 본 논문에서는 국내의 완성자동차 구매 고객을 대상으로 데이터 마이닝을 위한 기법 중 연관규칙 이용하여 기존 고객을 분석하였고 그 결과는 다음과 같다.

첫째, 데이터 마이닝에 의한 고객분석 결과 타겟마케팅 측면에서의 고객 성향 분석을 통하여 몇가지 유의한 결과를 살펴볼수 있었다. 이는 완성자동차 업체는 자동차 구입고객층의 대다수가 신문을 구독함을 알수 있었고 정치면이나 경제면을 활용하면 보다 나은 마케팅 효과를 보임을 입증하였다. 둘째, 잠재 신규고객, 재 구입

잠재고객 그룹의선별의 문제에서는 A사의차를 소유하고 있는 고객들이 구매전 비교차종이 더 많다는 유의한 연관규칙을 발견하였고, A사는 잠재고객을 대상으로 좀더 면밀한 경쟁사와의 비교에서 우위를 차지 할수 있는 부분을 마케팅 한다면 좋은 데이터 마이닝 모델의 활용을 할수 있다고 판단된다. 셋째, 잠재 자동차 구매 고객의 만족도를 통한 재구매 고객의 선별의 문제에서는 특정 자동차 회사의 소유차 만족도가 높은 고객을 대상으로 마케팅 한다면 신규 고객의 유치 보다 유리한 마케팅을 할수 있는 유의한 규칙이라고 판단된다. 추후 연구과제로는 다수의 변수 중에서 고객의 특정 속성을 인공신경망에 적용함으로써 우수, 비우수 고객의 유형의 추출 및 입력 변수가 단순화 됨으로써 모형을 쉽게 파악하고 우수고객과 비우수고객을 구분함으로써 보다 효율적인 마케팅 전략의 수립이 가능할수 있도록 세부적인 분석이 필요하다.

## 참고 문헌

- [1] 강현철, 한상태, 최종후, 데이터 마이닝 방법론 및 활용, 자유아카데미, 1999.
- [2] 김영만, "통신서비스 시장에서 데이터 마이닝을 이용한 이탈고객 분석", 한국과학기술원, 석사학위논문, 1999.
- [3] 김정수, 통합 데이터베이스 마케팅 시스템, 범우사, 1997.
- [4] 노상규, "데이터 마이닝 기법을 이용한 인공 신경망 입력변수 선정에 관한 연구", 서울대학교대학원, 석사학위논문, 1999.
- [5] 박종덕, 지태창, "데이터 마이닝 수행절차에 관한 연구", LG-EDS 시스템 기술연구부문 연구개발팀, 1999.
- [6] 박찬욱, "데이터베이스 마케팅", 연암사, 1996.
- [7] [http://magic.iin.co.kr/info\\_dong/index.htm](http://magic.iin.co.kr/info_dong/index.htm)
- [8] 장남식, 홍성완, 장재호, 데이터 마이닝, 대청, 1999.
- [9] 조인호, SAS 연습과 활용, 성안당, 1996.



[10] 조재희, 박성진, *데이터 웨어하우징과 OLAP*, 대청, 1996.

[11] 지원철, 김민용, “데이터 마이닝과 의사결정 지원 시스템”, *한국정보과학회지*, Vol.16, pp.24-36, 1998.

[12] 최중후, 한상태, *데이터 마이닝 기능과 사용법*, 자유아카데미, 2000.

[13] <http://www.hri.co.kr/m3/vip/9928/ex2.htm>

[14] <http://www.dpc.or.kr/dbworld/document/9709/spec.html>

[15] 허준, *통신시장에서의 데이터 마이닝 Telco-CAT*, SPSS아카데미, 2000.

[16] C. Apte and S. Weiss, “Data mining with decision trees and decision rules,” *Future Generation Computer Systems*, Vol.13, pp.197-210, 1997.

[17] M. Berry and G. Linoff, “Data Mining Techniques For Marketing, Sales and Customer Support,” John Willy & Sons, 1997.

[18] J. P. Bigus, *Data Mining with Neural Networks*, McGraw-Hill, 1996.

[19] E. Carl and E. Katz, “Customer spending behavior and telecommunications services,” *Telecommunications Policy*, Vol.22, pp.23-46, 1998.

[20] A. Famili and W. M. Shen, “Data Preprocessing and Intelligent Data Analysis,” *Intelligent Data Analysis*, Vol.1, pp.3-23, 1997.

[21] U. M. Fayyad and G. P. Shapiro, *Advanced in Knowledge Discovery and Data Mining*, AAAAI Press/MIT Press, 1996.

[22] A. Feelders, H. Daniels, and M. Holsheimer, “Methodological and practical aspects of data mining,” *Information & Management*, Vol.37, pp.271-281, 2000.

[23] B. A. Jain and B. N. Nag, “Performance Evaluation of Neural Network Decision Models,” *Journal of Management Information*

Systems, Vol.14, pp.201-216, 1997.

[24] H. Lan, *Data Mining*, Morgan Kaufmann Publishers, 1999.

[25] D. Pyle, *Data Preparation for Data Mining*, Morgan Kaufmann Publishers, 1999.

[26] M. Rob, *Data Warehousing and Data Mining for Telecommunications*, Artech House, 1987.

[27] S. Zhang, “Aggregation and maintenance for database mining,” *Intelligent Data Analysis*, Vol.3, pp.475-490, 1999.

저 자 소 개

이 병 엽(Byoung-Yup Lee)

정회원



- 1991년 2월 : 한국과학기술원 전 산학과(공학사)
- 1993년 2월 : 한국과학기술원 전 산학과(공학석사)
- 1997년 2월 : 한국과학기술원 경영정보공학(공학박사)

- 1993년 1월 ~ 2003년 2월 : 대우정보시스템 차장
- 2003년 3월 ~ 현재 : 배재대학교 전자상거래학과 부교수

<관심분야> : XML, 지능정보시스템, 데이터베이스 시스템, 전자상거래학

박 용 훈(Yong-Hoon Park)

정회원



- 2005년 : 호원대학교 정보통신공학과 및 건축공학과(공학사)
- 2007년 : 충북대학교 정보통신공학과(공학석사)
- 2007년 ~ 현재 : 충북대학교 정보통신공학과 박사과정

<관심분야> : 데이터베이스 시스템, 정보검색, 시공간 데이터베이스, 센서 네트워크 및 RFID

유 재 수(Jae-Soo Yoo)

종신회원



- 1989년 : 전북대학교 컴퓨터공학과(공학사)
  - 1991년 : 한국과학기술원 전산학과(공학석사)
  - 1995년 : 한국과학기술원 전산학과(공학박사)
  - 1995년 ~ 1996년 8월 : 목포대학교 전산통계학과 전임강사
  - 1996년 8월 ~ 현재 : 충북대학교 전기전자컴퓨터공학부 부교수
- <관심분야> : 데이터베이스 시스템, XML, 멀티미디어 데이터베이스, 분산 객체 컴퓨팅