

데이터마이닝을 활용한 한방분야에서의 데이터베이스 마케팅에 대한 연구

이상영*, 이윤석**

A Study on the Database Marketing using Data Mining in the Traditional Medicine

Sang-Young Lee*, Yun-Seok Lee**

요약

본 연구에서는 한방분야에서 검진에 영향을 미치는 요인들에 대하여 의사결정나무기법을 이용하여 도출하였고, 군집 분석을 통해 환자들의 특성에 대하여 알아보았다. 또한 질병간 연관성 분석을 재입원 환자 군들의 질병 구조를 파악하기 위하여 실시하였다. 아울러 도출된 군집들은 병원수익에 어떤 영향을 미치고 있는지에 대한 결과를 도출하였다. 즉 한방분야에 대하여 데이터마이닝 기법을 적용한 데이터베이스 마케팅을 통해 내원 환자들의 특성을 파악하고 병원 수익에 영향을 미치는 요인들을 실증적으로 도출하였다. 본 연구를 통하여 병원의 효율적인 운영과 활성화를 위한 데이터베이스 마케팅을 실시함으로써 병원 경영에 효율성을 도모할 수 있을 것이다.

Abstract

This study is to elicit the factors affected on the medical examination in the traditional medicine using the technical method of the decision tree and characterize the patient subject by clustering analysis technique. And to draw results from the association analysis between the form of diseases in the re-hospitalized patient group. The obtained results were analyzed for their effect on the hospital profits. Thus, through application of the database marketing to the data mining technique in the traditional medicine, the characteristics of patient clients for the objective induction of factors affected on the hospital profits can be identified. Practical application of the database marketing as presented in this study will bring about a fundamental efficiency of hospital management and vitalization.

▶ Keyword : DB 마케팅(DB marketing), 고객관계관리(CRM), 데이터마이닝(Data mining) 등

• 제1저자 : 이상영
• 접수일 : 2005.09.09, 심사완료일 : 2005.10.25
* 남서울대학교 보건행정학과 교수, ** 남서울대학교 보건행정학과 조교수

I. 서론

일반적으로 고객관계관리(CRM: Customer Relationship Management)는 공급자가 시장에서의 경쟁우위를 확보하기 위한 가장 중요한 요소가 고객이라는 것을 의미한다[1]. 즉 고객의 요구가 다양해지고 많은 경쟁자들이 등장하면서 그들과의 경쟁에서 살아남기 위해서는 고객의 욕구를 정확히 파악하고 그것을 충족시킬 필요성이 대두한 것이다[2, 3]. 더욱이 최근에 컴퓨터, 인터넷 등의 IT 기술이 활용되면서 고객관계관리는 상거래를 하는 모든 기업에서 필수적으로 갖춰야 할 전략적 시스템이 되었다[4, 5].

더욱이 의료분야에서도 환자에 대한 적극적인 관계관리의 중요성이 대두되고 있다. 특히 근래에 전 세계적으로 대체의학으로 한방에 대한 관심이 높아짐으로 해서 서구 여러 나라에서도 한방에 대한 연구가 활발히 이루어지고 있다. 또한 우리나라는 전통적으로 한방에 대한 신뢰도가 꾸준히 유지되고 있으며, 지금까지도 이용도가 높은 편이다. 최근 한방 치료의 의미가 예전에 단순한 한약, 침 등의 치료에 국한되지 않고 건강증진이나 질병의 예방으로 변화하면서 소비자의 욕구에 맞는 건강증진사업은 중요한 서비스의 하나로 인식되고 있다. 이러한 변화로 한방분야에서도 건강검진 서비스를 운영하면서, 수동적으로 치료 및 건강검진을 받으러 오는 것을 기다리기보다 능동적 환자에게 만족감을 줄 수 있는 마케팅 전략이 필요하게 되었다.

그러나 대부분의 병원들이 직면하고 있는 경쟁상황의 증가, 병원경영의 악화, 환자들의 보다 향상된 의료서비스의 요구 등은 그동안 병원경영활동에 한계를 드러내고 있다. 이와 같은 환경의 변화로 병원에서의 마케팅활동은 생존을 위한 중요한 요소로 대두되고 있다. 병원은 정보시스템의 발전과 더불어 많은 양의 데이터가 축적되면서 필요한 정보를 찾아내어 의료정보자원의 효율적 활용을 위한 지식경영의 비전 및 전략수립이 요구되고 있다. 과거 병원의 경영전략 수립은 정확한 정보에 근거하기 보다는 경영자의 경험에 근거하여 결정되는 경우가 많았다. 그러나 이제 병원의 규모가 커지며, 나날이 증가하는 경쟁에 우위를 점하기 위해서 정확한 정보에 근거한 경영전략으로 대량의 데이터를 효과적으로 저장, 관리, 활용할 수 있는 지식 탐사 방법인 데

이터마이닝(data mining)에 대한 중요도가 증대되고 있다 [6]. 많은 병원에서 전산화가 구현됨에 따라 차트에 보관되었던 진료정보가 컴퓨터에 보관됨에 따라 데이터를 최신 정보 기술인 데이터마이닝 등의 기법을 이용하여 병원경영에 관한 정보를 얻을 수 있어서 정보기술을 활용하는 병원경영 전략의 수립이 가능하게 되었다. 선진 병원에서는 진료정보의 효율적 활용을 위한 데이터웨어하우스, 데이터마트, 데이터마이닝 등의 정보기술을 기반으로 진료정보를 효율적으로 활용하고 있으나 우리나라에서는 대부분 진료정보 전산화로 구축된 데이터를 단순 업무에만 활용하고 있는 실정이다. 이런 시점에서 병원의 처한 상황이나 마케팅 목적에 따라 전통적인 매스마케팅에 의존하기보다 모든 원인변수를 정확히 계산하고 고객의 평생가치를 높이는 데 그 역할을 하고 있는 데이터마이닝을 통해 데이터베이스 마케팅을 실시함으로써 고객과의 관계를 형성해 나가고, 또한 고객의 욕구변화에 적용할 수 있는 대안이 되리라 생각된다. 의료기관에서의 데이터베이스 마케팅은 환자와 관련 있는 각종 정보들과 내원기록 등을 종합한 정보로 데이터베이스를 구축한 후 이를 바탕으로 목표 환자를 선정하고 전화나 우편물 발송 등을 통하여 특정 고객을 집중 공략하는 마케팅 접근방법을 의미한다. 이러한 추세로 인하여 그 규모가 큰 여러 의료 연구소나 종합병원들은, 이용자들의 여러 데이터를 이용하여 데이터마이닝을 통한 마케팅을 수립하고 있지만, 중소병원 및 한방분야에서는 그러지 못하고 있는 실정이다. 이에 본 논문에서는 한방분야에서의 효과적인 마케팅을 위해 중규모의 한방병원의 데이터를 이용하여, 실증적으로 문제해결 방법을 제시한다. 즉 데이터의 특성 파악 및 전처리 과정 등을 통하여 데이터마이닝 기법으로 제검진을 결정하는 요인을 찾아내고, 군집별 특성을 분석하여 이를 데이터베이스 마케팅에 적용함으로써 데이터베이스를 효과적으로 마케팅에 활용할 수 있는 방안을 제시한다.

II. 관련 연구

고객관계관리를 도입할 시에 가장 선행되어야 할 것은 고객을 정확히 파악하는 것이다[7]. 의료분야에서의 의사, 환자, 보험회사 등과의 관계는 기존의 고객-공급자 모델을 적용하기에는 부적절하며 장기간에 걸친 관계를 설정하는 것이 중요하다[8]. 기존에 고객을 특성별로 정확히 파악하

는 방법에는 여러 가지 기법들이 활용된다. 최근에는 데이터베이스에 저장하여 활용하는 기법들이 사용된다. 이렇게 수집된 자료들은 데이터마이닝 과정을 통하여 고객을 적절한 그룹으로 나누고 이를 바탕으로 캠페인을 실행하여 보다 정확한 고객 군으로 분류하게 된다[9, 10]. 이와 관련된 다양한 분야에서의 기존의 연구들이 있는데 특히 의료분야에서 이러한 캠페인 관리를 통한 고객관계관리 시스템의 실현은 전반적인 의료분야의 발전은 물론 수익 증대에도 큰 도움이 되는 것으로 분석하고 있다[11, 12]. 특히 의료 분야의 경우에 신규 환자에 대한 병원의 신뢰성을 증대시키고, 수익성 있고 충성도가 높은 환자들을 효과적으로 선별하여 관리하기 위하여 이러한 고객관계관리 기법을 적절하게 활용하는 것이 중요하다[13].

국내에서도 타 산업과 마찬가지로 의료분야에 있어서도 상당 수 의료기관들이 수준 높은 의료 서비스를 저렴하고 신속하게 제공하는 것이 경제우위의 전제조건이라는 인식을 공유하면서 의료정보 시스템의 구축을 하였다. 그러나 대부분이 전자차트 시스템이나 보험 처리 시스템 등 비교적 단순한 원무 관리를 지원하는 수준이며, 환자의 증상이나 각종 병력 데이터, 검사 데이터 등을 기반으로 하는 고객관계관리 시스템의 보급은 미흡한 실정이다. 더욱이 본 논문에서 대상으로 하는 한방분야의 경우 일부 병원을 제외하고는 기본적인 시스템조차 미비한 경우가 많다. 이에 반해 대부분의 선진 의료 기관들은 고객관계관리 시스템의 구축을 이미 완료 한 상태이고, 한층 더 나아가 이러한 시스템을 통하여 수집한 환자의 데이터를 분석하여 의료진의 환자 질병 진단을 지원하고자 하는 연구를 꾸준히 진행하여 왔다. 이러한 연구의 동기는 자동화된 의료진단 지원시스템을 제공하고 의료진의 활용을 유도함으로써, 전문 분야라도 익숙한 질병의 진단에만 집착하기 쉬운 의료진의 의뢰지식과 진단의 폭을 넓힐 수 있다는 점, 의료진 개인의 특성에 따라 주관적일 수 있는 질병의 진단이나 예후를 좀 더 객관화함으로써 궁극적인 고객 중심의 서비스가 가능하다는 점이 있다.

III. 한방 데이터베이스 마케팅 설계

본 논문에서의 데이터는 한방병원의 차트를 통해 연구 목적에 맞도록 필요한 항목을 추출하고, 변형하여 구축하였다. 분석을 위한 개략적인 모형은 (그림 1)과 같다.

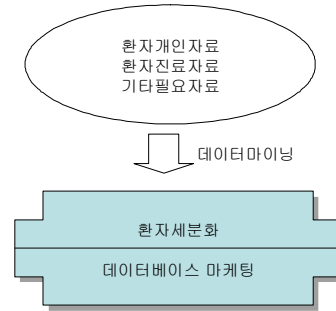


그림 1 분석을 위한 모형화
Fig 1. Modeling for analysis

그림에서 보는 바와 같이 차트 데이터베이스를 이용해 환자 세분화를 수행하였는데 내원한 환자의 특성에 따른 세분화를 수행하고 이를 기반으로 병원 수익 측면에 따라 다시 환자를 세분화하고 이를 바탕으로 군집별 특성에 맞는 마케팅 전략을 모색하였다. 즉 병원에 내원한 환자를 대상으로 병원을 이용한 이력 데이터를 사용하였다. 여기서는 내원환자를 대상으로 하므로 내원 횟수가 높은 환자가 병원에 충성도가 높다는 가정을 하였다. 이때에는 의사결정분석을 통한 재검진 여부를 결정하는 요인분석 및 중요변수 선정 실시하였다. 또한 군집 분석을 통해 요인들에 대한 특성을 파악하였다. 아울러 재검진유무, 총진료횟수, 총내원일수 등을 기준으로 환자를 세분화하고 군집별 환자의 특성을 파악하였다. 또한 군집별 질병의 연관성을 분석하여 진료에 효율성을 높일 수 있도록 하였다. 그리고 이를 바탕으로 병원 수익 측면에서 환자를 다시 세분화하였는데 여기서는 내원한 환자를 대상으로 이들의 요약 데이터와 외래 데이터를 이용하였다.

다음으로 효과적인 데이터마이닝을 지속적으로 수행하기 위해서 데이터베이스 마케팅을 위한 프로토타입 시스템을 구축하였다. 그러므로 우선 여기서는 한방병원에서 저장하고 있는 진료환자의 차트 데이터를 이용하고 기타 필요한 데이터는 분석할 변수를 정의하여 병원 차트에서 정리하여 구축하였다. 이러한 분석에 이용한 변수는 <표 1>과 같다.

표 1. 분석을 위한 변수 속성
Table 1. Variable attribute for analysis

변수명	설명	내용
C.NO	등록번호	
SEX	성별	1.남자 2.여자
AGE	나이	1.10대 2.20대 3.30대 4.40대 5.50대 6.60대 7.70대이상
AREA	거주지	1.전주 2.완주 3.12외 도시 4. 1,2외 농어촌 5.기타
GUA.C	보험형태	10.일반 20.의료보험 30. 산재 40.보호 50.자보 60.해보
JOB	직업	1.전문직, 행정직, 사무직 2.상업, 서비스직, 자영업 3.생산직, 운수업, 단순노무 4.학생, 군인 5.주부, 무, 상세불명
A.PATH	내원경로	1.자의 2.타인소개 3.기타
P.N	질병명	1.소화기 2.비소화기 3.타박상 4.기타
C.O	치료결과	1.완쾌 2.불변 3.기타
S.P	수술여부	0.수술 안함 1.수술함
D.N	타 검진횟수	0.무 1.1회 2.2~5회 6.6회이상
B.P	혈압	1.고 2.저 3.정상
H.N	입원 횟수	0.무 1.1회 2.2~5회 6.6회이상
M.C	주 증상	1.소화장애 2.비소화기 3.발목 4.팔꿈치 5.손목 6.등 7.목 8.어깨 9.허리 10.중복
M.S	투약내역	1. 무 2.침 3.물리치료 4.약물 5.침+물 리치료 6.침+약물 7.물리치료+약물 8.뜸 9. 뜰+침 10. 기타
R.T	재검진일	
T.N	총 내원횟수	조사 기간 동안 총 내원횟수
H.M	한약제조횟수	1.0회 2.1회 3.2~5회 4.6회 이상
F.D	최종내원간격	최근내원일 - 마지막 내원일
CHAR	성격	1. 차분 2.내성적 3.급함 4.기타
L.S	라이프스타일	1.상 2.중 3.하
O.H	사상의학체질	1.태양인 2.태음인 3.소양인 4.소음인

IV. 한방 데이터베이스 마케팅 전략

4.1 내원환자 특성에 따른 세분화

분석대상 한방병원에서 내원환자의 비례에 맞춰 2003년부터 2004년까지 2년간 랜덤하게 선택한 400명의 검진자 수는 <표 2>와 같다. 연도별 변화를 살펴보면 2003년에서 2004년으로 갈수록 매년 내원환자의 수가 늘어나고 있는 추세에 있었다. 성별로는 남자가 31%, 여자가 69%로 여자가 더 많은 것으로 나타났다.

표 2. 대상병원의 연도별 검진자 수
Table 2. Medical examination person number by year of target hospital

년도	검진자 수	남	여
2003	150	34(22.6)	116(77.4)
2004	250	90(36.0)	160(64.0)
계	400	124(31.0)	276(69.0)

(단위: 명(%))

먼저 내원한 환자를 대상으로 이후 병원을 이용한 이력 데이터를 사용하여 재검진 여부를 판단하고 내원횟수가 높은 환자가 병원에 충성도가 높다는 가정을 하였다. 재검진 패턴 모형을 구축하여 상호 비교해 봄으로써 본 분석을 위해 유리한 분석 방법을 선택하였다. 본 논문에서는 기본적으로 E-Miner를 사용하였다. 비교대상 분석법은 회귀분석, 의사결정나무, 신경망 기법으로 구축한 패턴 모형은 (그림 2)와 같다.

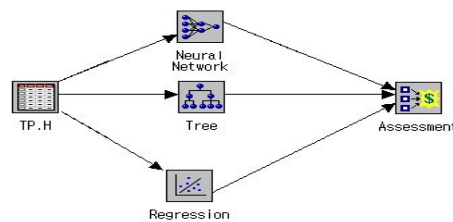


그림 2. 재검진 패턴 모형
Fig 2. Re-medical examination pattern model

이 모형을 분석한 결과 <표 3> 및 (그림 3)에서 보는 바와 같이 의사결정나무가 다른 분석방법보다 대체적으로 좋은 결과를 보였다.

표 3. 모형의 비교 분석
Table 3. Comparative analysis of model

분석법	Root ASE	Response rate(%)		
		상위 10%	상위 20%	상위 30%
회귀분석	0.363	100	97.5	94.16
의사결정나무	0.127	100	98.81	98.81
신경망기법	0.404	95.83	96.25	97.5

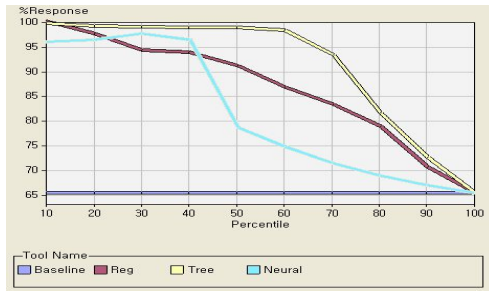


그림 3. 모형의 리프트 차트
Fig 3. Lift chart on model

따라서 Root ASE 값과 Response rate(%)에서 좋은 결과를 보인 의사결정나무를 이용해 재검진 여부를 결정하는 중요변수를 선정하였다. 그 결과가 (그림 4)와 같이 분석 결과 재검진 여부에 가장 큰 영향을 미치고 있는 요인이 주 증상의 여부로 어깨, 허리 그리고 중박된 타박상에 의한 것들이었다. 또한 이를 반영하듯 하위 단계에서 타박상이 영향을 미치고 있으며, 주로 여성이 남성에 비해 재검진을 많이 받는 것으로 나타났다.

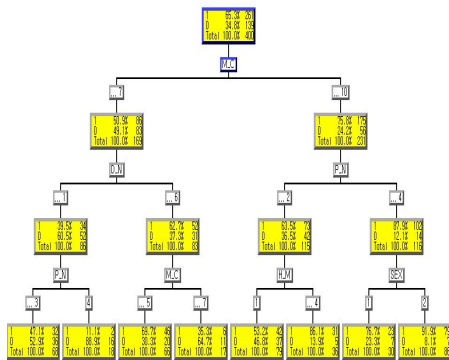


그림 4. 재검진 여부의 의사결정나무
Fig 4. Decision tree of re-medical examination

다음으로 군집분석을 통해 변수들의 특성을 파악한 결과는 다음 <표 4>와 같다.

표 4. 군집 분석을 통한 환자 세분화
Table 4. Patient classification through clustering analysis

구분	군집 I	군집 II	군집 III	군집 IV	군집 V
나이(%)					
10대	1.43	18.09	66.67		
20대	12.6	18.09			2.17
30대	14.29	15.96		50.00	
40대	7.14	13.83			10.87
50대	24.29	23.40	33.33	50.00	16.30
60대	28.57	5.32			31.52
70대 이상	11.43	5.32			39.13
직업(%)					
전문직, 행정직, 사무직	24.29	3.19			3.26
상업, 서비스직, 자영업	32.86	11.70		50.00	28.26
생산직, 운수업, 단순노무	18.57	13.83	33.33	50.00	21.74
학생, 군인	8.57	30.85			
주부, 무상세불명	15.71	40.43	66.67		46.74
치료결과(%)					
완쾌	55.71	92.55	66.67		18.48
불변	30.00	6.38			51.09
기타	14.29	1.06	33.33	100	30.43
타 검진 횟수(%)					
0회	34.29	47.87	66.67	50.00	10.87
1회	21.43	14.89			46.74
2~5회	34.29	27.66			1.09
6회 이상	10	9.57	33.33	50.00	41.30
주 증상(%)					
소화장애	2.86	23.40			2.17
비소화기	4.29	4.26			1.09
발목	2.86	3.19			5.43
팔꿈치	10	8.51			4.35
손목	5.71	14.89			2.17
등	5.71	10.64			6.52
목	11.43	11.70			19.57
어깨	25.71	11.70			1.09
허리	10.00	7.45		50.00	20.65
중박	21.43	4.26	100	50.00	36.96
라이프스타일(%)					
상	35.71	34.04	66.67		3.26
중	57.14	61.70	33.33	50.00	60.87
하	7.14	4.26		50.00	35.87
사상의학(%)					
태양인	1.43	2.13			2.17
태음인	24.29	56.38	100	50.00	66.60
소양인	50.00	30.85		50.00	22.83
소음인	24.29	10.64			8.70

4.2 군집별 특성분석을 통한 환자 특성분석

재검진유무, 총진료횟수 등을 기준으로 환자를 세분화하고 군집별 환자의 질병구조 및 특성의 연관성을 파악하였다. 군집분석을 위한 모형은 (그림 5)와 같다. 여기서는 군집분석을 위한 데이터 정제를 위해서 Filter Outliers 기능을 추가하여 적용하였다.

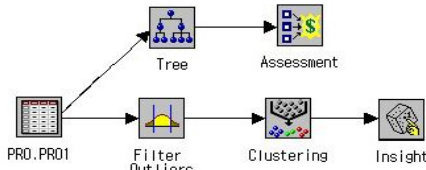


그림 5. 군집분석을 위한 모형
Fig 5. Model for clustering analysis

위 모형을 통해 다섯 개의 군집으로 나누었고 군집별 환자군들에 대한 분석은 <표 5>와 같다.

표 5. 군집별 특성분석
Table 5. Characteristic analysis by clustering

구분	군집 I	군집 II	군집 III	군집 IV	군집 V
재검진일일	3.65	22.66	2.43	1.55	14
총내원일일	4.66	23.66	3.43	2.55	15
최종내원 간격일	5.30	13.33	5.45	4.03	10.62
총진료비용(원)	363,966	671,000	110,297	7,654	39,500
한약제조 횟수회	2.5	60이상	1	0	0
라이프스타일	중상	상	중	중하	중
주 증상					
소화장애	2.86%	23.40%			2.17%
비소화기	4.23%	4.26%			1.09%
발목, 팔꿈치, 손목, 등	24.48%	37.23%			18.47%
목, 어깨, 허리	47.14%	30.85%		50%	41.31%
중복	21.43%	4.26%	100%	50%	36.96%
환자특성	잠재군	충성도군	이탈군	이탈군	충성도군

표에서 보는 바와 같이 재검진일, 총 내원일, 최종 내원 간격, 총 진료비용, 한약제조횟수, 라이프스타일, 주증상 등을 기준으로 군집분석을 실시하였다. 분석을 통해 군집별로 본 병원에 미치는 영향력을 잠재고객군, 충성도군, 이탈환자군으로 분류하였다. 분류를 위한 기준으로 재검진일, 총 내원일, 최종내원 간격, 총 진료비용 등을 사용하였다.

분석결과 군집 I은 재검진일 3.65일, 총 내원일이 4.65일로 비교적 전체 군집의 평균에 속하며 최종내원기간이 5.35일로 짧은 편에 속해 최근까지 지속적으로 병원을 이용하는 환자군이다. 이 환자군은 병원에서 수익을 올릴 수 있는 한약 제조 횟수도 많기 때문에 병원에 이용도를 높일 수 있는 고객군이다. 차별화되고 새로운 치료방법을 개발하여 홍보를 통해 본 병원에 지속적으로 유치할 수 있는 잠재고객의 집단이 되겠다. 질병군을 살펴보면 이 병원에 주 업무인 타박상을 많이 앓고 있다. 주로 행정업무나 서비스 업무를 직업으로 하는 군집으로 비교적 안정된 생활을 이루고 있는 집단이라고 할 수 있다. 그리고 군집 II는 재검진일 및 총 내원일이 5개 군집 중 가장 높았으며, 병원이용률과 함께 진료비도 가장 높은 환자군이다. 또한 여러 질병을 앓고 있는 군집이기도 하다. 최종내원간격이 다른 군집에 비해 높은 값을 보이고 있지만 재검진일이 다른 군집보다 높은 것으로 보아 질병이 발생했을 경우 본 병원을 찾는 주 고객군이 되겠다. 또한 한약 이용도도 다른 군집에 비해 월등히 높기 때문에 병원의 수익에 큰 영향력을 가진 충성도군이라고 할 수 있다. 또한 군집 III은 재검진일이 2.43일, 최종내원 간격이 5.45일로 평균 값을 약간 밑도는 군집이라고 할 수 있다. 여러 가지 원인으로 인해 병원을 이용한 환자로서 간단한 치료와 약 처방으로 완쾌되어 병원이용률이 낮은 환자군이다. 또한 이 환자군은 다른 병원으로 이동한 환자들도 포함되어 있다. 아울러 군집 IV는 병원 이용률이 낮은 환자군으로 그 병세가 미약하여 한 두 번의 치료에 의해 완쾌되어 더 이상 병원을 이용하지 않은 환자들이라고 할 수 있다. 마지막 군집 V는 지병에 의해서 병원을 이용하는 환자들로 재검진일이 14일로 이용도가 높은 환자군들이라 할 수 있다. 주로 상업이나, 서비스 업무를 직업으로 하는 집단으로 계속해서 병원을 이용할 것이라 예상된다. 본 환자군은 지병을 앓고 있는 집단으로 지속적인 치료가 필요하기 때문에 정기적으로 본 병원을 이용한다. 이와 같은 이유로 본 환자군은 지속적인 관리와 새로운 치료방법을 통해 본 병원의 이용률을 유지하도록 할 필요가 있다.

다음으로 군집별 질병의 연관성을 분석을 통해 질병구조를 알아보았다. 이 분석은 환자의 주 질병치료와 함께 연관

된 질병을 병행하여 치료해줄 수 있는 정보가 될 수 있다. 각 군집별 질병구조와 질병의 분포형태가 단순한 군집 III와 군집 IV를 질병구조 연관성의 결과는 다음 <표 6>과 같다.

표 6. 군집별 질병군 분석
Table 6. Disease analysis by clustering
(단위 : %)

구분	군집 I	군집 II	군집 III	군집 V	군집 V
소화장애	2.86	23.40			2.17
비소화기	4.29	4.26			1.09
발목	2.86	3.19			5.43
팔꿈치	10.00	8.51			4.35
손목	5.17	14.89			2.17
등	5.17	10.64			6.52
목	11.43	11.70			19.57
어깨	25.71	11.70			1.09
허리	10.00	7.45		50.00	20.65
중복	21.43	4.26	100	50.00	36.96
한자특성	잠재군	충성도군	이탈군	이탈군	충성도군

위 결과를 살펴보면 군집 I은 일부 질병을 제외하고 그 증상이 고루 분포하고 있음을 알 수 있다. 이 환자군은 전문직, 행정직, 사무직 등 주로 사무적인 일과 상업, 서비스직, 자영업 등 주로 물품들을 취급하는 직업을 가지고 있는 환자로 구성되어 있다. 이는 업무로 인해 부위별 타박상의 증세를 가지고 있음을 반영한다. 군집 II는 소화기와 일부 부위의 타박상을 주로 보이고 있다. 직업은 주로 활동이 적은 직업군을 보이는 환자군이다. 군집 V는 앞에서 언급한 것처럼 지병을 앓고 있는 환자군으로 60대 이후의 인원이 많은 분포를 보이고 있다. 그리고 <표 7>, <표 8>, <표 9>에서는 I, II, V군에 대한 질병구조 연관성 분석결과를 보여준다.

표 7. 질병구조 연관성 분석 - 군집 I
Table 7. Disease structure association analysis - clustering I

순위	질병명	빈도 (%)	부상병 연관성			제점진 병 연관성		
			Rule	Confidence (%)	Support (%)	Rule	Confidence (%)	Support (%)
1	어깨(JJ 6)	25.71	JJ 6=>JJ 5	33.98	2.76	JJ 6=>JJ 8	79.17	5.82
2	중복(JJ 8)	21.43	JJ 8=>JJ 7	37.84	2.14	JJ 8=>JJ 7	73.53	3.83
3	목(JJ 5)	11.43	JJ 5=>JJ 6	33.98	2.76	JJ 5=>JJ 5	80.56	4.44
4	팔꿈치(JJ 2)	10.00	JJ 2=>JJ 3	20.51	1.23	JJ 2=>JJ 8	66.87	3.37
5	허리(JJ 7)	10.00	JJ 7=>JJ 8	37.84	2.14	JJ 7=>JJ 7	92.00	3.52
6	등(JJ 4)	5.17	JJ 4=>JJ 7	60.00	2.76	JJ 4=>JJ 8	80.77	3.86
7	손목(JJ 3)	5.17	JJ 3=>JJ 2	20.51	1.23	JJ 3=>JJ 3	46.34	2.77
8	비소화기(JH 1)	4.29				JH 1=>JH 1	100.00	3.32
9	소화장애(JS 1)	2.86	JS 1=>JJ 8	58.21	5.89	JS 1=>JS 1	92.86	1.90
10	발목(JJ 1)	2.86	JJ 1=>JJ 8	37.04	1.45	JJ 1=>JJ 1	75.41	3.07
		100						

표 8. 질병구조 연관성 분석 - 군집 II
Table 8. Disease structure association analysis - clustering II

순위	질병명	빈도 (%)	부상병 연관성			제점진 병 연관성		
			Rule	Confidence (%)	Support (%)	Rule	Confidence (%)	Support (%)
1	소화장애(JS 1)	23.40	JS 1=>JJ 8	16.42	1.02	JS 1=>JS 1	93.10	2.10
2	손목(JJ 3)	14.89	JJ 3=>JJ 2	10.45	1.02	JJ 3=>JJ 3	46.34	2.77
3	목(JJ 5)	11.70	JJ 5=>JJ 6	90.91	1.75	JJ 5=>JJ 6	92.00	3.52
4	어깨(JJ 6)	11.70	JJ 6=>JJ 5	90.91	1.75	JJ 6=>JJ 5	92.00	3.52
5	등(JJ 4)	10.64	JJ 2=>JJ 1	58.21	5.89	JJ 2=>JJ 2	92.00	3.52
6	팔꿈치(JJ 2)	8.51	JJ 4=>JJ 8	60.00	2.76	JJ 4=>JJ 8	62.54	2.19
7	허리(JJ 7)	7.45	JJ 7=>JJ 8	13.11	1.17	JJ 7=>JJ 7	81.56	4.54
8	비소화기(JH 1)	4.26	JH 1=>JJ 7	22.40	2.10	JH 1=>JH 1	99.10	3.45
9	중복(JJ 8)	4.26	JJ 8=>JJ 7	13.11	1.17	JS 8=>JS 8	94.14	2.89
10	발목(JJ 1)	3.19	JJ 1=>JJ 8	32.24	1.26	JJ 1=>JJ 1	79.17	5.82
		100						

표 9. 질병구조 연관성 분석 - 군집 V
Table 9. Disease structure association analysis - clustering V

순위	질병명	빈도 (%)	부상병 연관성			제점진 병 연관성		
			Rule	Confidence (%)	Support (%)	Rule	Confidence (%)	Support (%)
1	중복(JJ 8)	36.96	JS 8=>JJ 7	73.39	19.20	JS 8=>JS 8	91.67	1.92
2	허리(JJ 7)	20.65	JJ 7=>JJ 8	73.39	19.20	JJ 7=>JJ 7	93.27	1.85
3	목(JJ 5)	19.57	JJ 5=>JJ 8	68.43	3.56	JJ 5=>JJ 5	80.77	3.66
4	등(JJ 4)	6.52	JJ 2=>JJ 8	59.54	3.24	JJ 6=>JJ 7	66.50	3.37
5	발목(JJ 1)	5.43	JJ 1=>JJ 7	46.43	1.90	JJ 1=>JJ 1	92.67	3.89
6	팔꿈치(JJ 2)	4.35	JJ 2=>JJ 3	52.00	5.89	JJ 2=>JJ 3	46.34	2.77
7	소화장애(JS 1)	2.17	JS 1=>JJ 8	10.45	1.02	JJ 1=>JJ 1	92.86	1.90
8	손목(JJ 3)	2.17	JH 3=>JJ 2	52.00	5.89	JH 3=>JH 2	46.34	2.77
9	비소화기(JH 1)	1.09				JH 1=>JH 1	100.00	3.32
10	어깨(JJ 6)	1.09	JJ 6=>JJ 8	37.84	2.14	JJ 6=>JJ 8	73.53	3.83
		100						

먼저 군집 I의 질병구조 연관성 분석의 결과를 살펴보면 주로 부위별 타박상으로 내원한 환자들이다. 대부분의 환자들이 거의 가까운 부위에 연관성을 가지고 질병을 앓고 있는 추세를 보이고 있다. 그리고 한 부위에 질병이 있을 경우 다른 부위에 중복 질병이 발생하고 있기 때문에 병원에서 본 연관성 결과를 토대로 주 질병 치료에만 전념하지 않고 다른 부위의 질병에도 관심을 가져야 할 것이다. 또한 제점진이 주로 같은 질병에 의해 이루어지기 때문에 다양한 치료방법을 개발·도입하여 매 점진마다 차별화된

치료를 시도해야 할 것이다. 그리고 군집 II는 소화장에 환자들이 높은 비율을 보이고 있다. 그만큼 활동이 부족한 집단군으로 주로 손이나 어깨를 많이 사용하거나 경직된 자세로 오랫동안 작업을 하여 질병이 나타나는 경우가 발생하고 있다. 각 질병들이 군집 I에 비해 서로 연관성을 가지고 발생하고 있으며, 제검진 또한 기존의 질병과 다른 질병이 발생하여 이루어지는 경우가 있다. 마지막 군집 V는 충성도 군으로 다른 군집과 질병의 연관성 분석에서 많은 차이를 보이고 있다. 질병의 발생 구조가 다양하게 이루어지고 있으며, 연관성을 가지고 발생하는 질병 또한 다양하다. 이는 다른 군집에 비해 연령대가 높은 군집으로 지병과 함께 여러 부위에서 이상이 발생하는 이유 때문일 것이다. 제검진의 병 연관성을 살펴봐도 기존의 질병과 다른 질병으로 인해 내원한 경우도 많다. 부위별 질병이 많은 환자군으로 관심과 다양한 치료를 통해 지속적으로 내원하게 할 수 있을 것이다.

4.3 병원 수익성 측면에 따른 환자 세분화

여기서는 병원환자를 수익성에 따라서 세분화하였는데 그 기준 변수로는 한약 제조횟수, 총 진료비용, 라이프스타일 등을 사용하였다. 군집분석 결과 <표 10>과 같이 5개의 군집으로 분류되는데 여기서 군집 II가 세 항목 모두 가장 높은 수준을 보이고 있어 고수익군으로 분류되었고 반면에 군집 III은 뚜렷한 특징이 없는 중간수준 또는 일반군으로 분류하였다.

표 10. 군집별 특성분석
Table 10. Characteristic analysis by clustering

구분	군집 I	군집 II	군집 III	군집 IV	군집 V
총진료비용 (원)	363,956	671,000	110,297	7,654	39,500
한약제조횟수(회)	2.5	6이상	1	0	0
라이프스타일	중상	상	중	중	중하
	고수익군	고수익군	일반군	저수익군	중수익군

그리고 고수익군, 일반군, 저수익군으로 분류하는 기준과 패턴을 발견하기 위해 의사결정나무 모형을 적용하였다. 의사결정나무를 이용하여 환자수익성의 영향요인을 분석한 결과 나이, 질병 명, 한약제조횟수, 지역 등이 주요한 특성변

수로 적용되었다. 여기서는 나이가 첫 번째 노드로 분류되어 환자수익성을 설명하는 가장 중요한 변수로 제시되었다. 즉, 나이의 경우에는 모든 연령대에서 일반환자군이 압도적으로 많은 비중을 점하고 있으며, 50대 이전에는 고수익 환자군이 상당한 비중을 차지하고 있다. 이러한 관계는 <그림 6>과 같다.

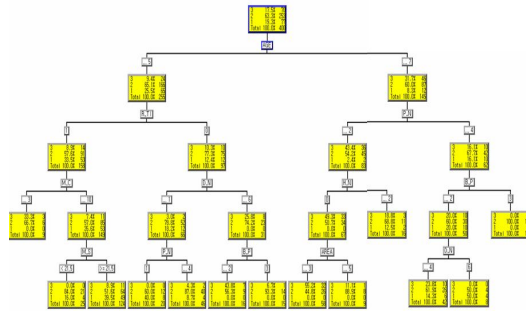


그림 6. 수익성에 대한 의사결정나무
Fig 6. Decision tree about profitability

4.4 데이터베이스 마케팅 전략

분석 결과를 토대로 환자 군집별에 대한 데이터베이스 마케팅 전략 다음과 같이 수립하였다. 먼저 군집 I의 경우 잠재고객으로 연령대가 고루 분포하고 있으며, 직업 또한 집중현상을 보이지 않고 있다. 그리고 모든 요인에 평균 값을 나타내는 환자군으로 우리나라 중산층을 대표하는 집단이 되겠다. 진료내용을 토대로 홍보를 통해 고정 고객으로 유도할 수 있는 집단이라 할 수 있다. 그리고 군집 II는 병원의 수익에 큰 영향을 미치는 집단으로서 충성도가 높은 고객이라고 할 수 있다. 이 집단의 경우에는 별 다른 홍보가 없어도 주기적으로 본 병원을 이용할 것이다. 그러나 이 환자군은 활동이 부족한 집단으로 치료와 함께 간단한 운동을 통한 치료를 병행한다면 병원이용도는 더 높아질 것이라 예상된다. 예를 들면 간단한 기체조나 요가를 할 수 있는 환경을 마련하여 병행치료 한다면 치료 효과 증대와 함께 병원의 환자유치에 많은 영향을 미칠 것이다. 또한 군집 III은 이탈 고객군이지만 이 군집에 속한 환자들은 대부분 본 병원 주위에서 상업을 직업으로 하는 인원들이기 때문에 관심을 가진다면 잠재로알고객이 될 수 있는 가능성이 있다. 예를 들어 유치전략으로 무상진료 일을 선정해서 치료를 한 다든지, 직접 방문하여 상담과 함께 간단한 치료를 병행한

다면 큰 성과를 얻을 수 있을 것이다. 아울러 군집 IV 역시 이탈고객군으로 관심의 대상이 아니다. 마지막 군집 V는 지병을 앓고 있는 환자들로 구성되어 있으며, 정기적으로 병원을 방문하여 간단한 진료를 받는 고객군이 되겠다. 병원에서의 진료항목을 좀 더 보강하여 다양성을 가진다면 이 고객군은 병원 수익에 많은 영향을 미칠 것이다. 이와 같은 병원 수익을 위한 마케팅 전략을 정리하면 <표 11>과 같다.

표 11. 데이터베이스 마케팅 전략 제안
Table 11. Database marketing strategy proposal

군집	전략	개별 전략	공동 전략
군집 I		· e-mail, 전화상담 등을 통한 질병상담 · 지속적 관계유지, 다양한 인센티브제 도입	· 기 치료 및 요가 운영 등으로 환자 내원 유지 및 신환자 유치 전략 수립
군집 II		· 차별화된 약물치료 · 다양한 치료법 개발 및 도입을 통한 치료	· 건강검진 체계 설립을 통한 정기적인 환자 건강체크로 병원의 신뢰도 향상 방안 모색
군집 III		· 무상진료일 배정 · 직접방문으로 질병상담 및 치료 도입	· 무료봉사활동을 통한 병원에 대한 인식 변화
군집 IV		· 치료 후 만족도 조사 · 차별화된 치료 홍보	· 병원 직원의 지속적인 교육을 통해 환자에 대한 친절도 향상
군집 V		· 질병의 연관성 인식을 통한 연계치료 · 다양한 치료를 통한 만족도 향상	

V. 결론 및 향후 연구과제

본 연구는 한방분야를 대상으로 하여 데이터마이닝 기법을 통해 내원 환자들의 특성을 파악하고 병원 수익에 영향을 미치는 요인들을 도출하였다.

그 방법으로는 재검진에 영향을 미치는 요인들에 대하여 의사결정나무기법을 이용하여 도출하였고, 군집 분석을 통해 환자들의 특성에 대하여 알아보았다. 또한 재입원 환자

군들의 질병 구조를 파악하기 위하여 질병간 연관성 분석을 실시하였다.

그리고 도출된 군집별로 병원수익에 어떠한 영향을 미치고 있는지에 대한 결과를 도출하였다. 분석결과 병원을 이용하는 환자들을 5개의 군집으로 나누었고, 그 군집별로 그 특성을 알 수가 있었다. 이러한 일련의 과정을 통해 병원에서 지속적으로 관심을 가지고 유도해야할 고객군을 선정하고 병원 발전에 큰 역할을 할 수 있는 데이터베이스 마케팅 전략을 수립하였다.

그동안 데이터마이닝을 이용한 데이터베이스 마케팅 전략 수립에 대하여 많은 연구가 있었고 현재 의료업계에서도 그 활동이 왕성하게 이루어지고 있다. 그러나 그러한 활동은 규모가 큰 곳에만 국한되어 있고 여러 곳에 산재해 있는 중소 병원 및 본 연구의 대상인 한방분야의 경우 매우 미비한 실정이다. 향후 환자의 특성에 맞도록 개별고객 단위까지 데이터베이스 마케팅 모형을 적합화 한다면 병원의 마케팅이 아직까지 환자의 설문이나 무차별적 블랭킷 마케팅에 머물고 있는 시점에서 향후 정확한 데이터를 바탕으로 한 데이터베이스 마케팅을 수행하는 데 기초 자료로 활용할 수 있을 것이다. 병원의 고객 DB는 이미 방대한 데이터로 구성되어 있어 이의 활용에 따른 잠재적 효과는 매우 크다고 할 수 있다. 병원의 효율적인 운영과 활성화를 위해 데이터베이스를 체계적으로 구축하고 분석한다면 이를 바탕으로 한 데이터베이스 마케팅에 적용함으로써 병원 경영에 효율성을 도모할 수 있을 것이다.

그리고 개개인의 특성에 맞는 보다 차별화 된 의료서비스를 제공하고, 지속적인 건강관리 및 정기적인 검진, 예방을 통해 질병을 조기에 발견, 치료함으로써 병원의 신뢰도를 높이고 환자 또한 건강한 삶을 도모할 수 있도록 할 수 있을 것이다. 추후 연구과제로는 이러한 일련의 고객관리를 위한 데이터베이스 마케팅 수행에 있어 데이터를 정확하게 입력하고 실시간으로 이용할 수 있는 프로그램 개발이 필요하다.

참고문헌

[1] Yim CK, Kannan PK., "Consumer behavioral loyalty: a segmentation model and analysis," *Journal of Business Research*, Vol. 44(2), pp. 75-92, 1999.

[2] Kohli R, Piontek F, Ellington T, VanOsdol T, Shepard M, Brazel G, "Managing customer relationships through e-business decision support applications: a case of hospital-physician collaboration," *Decision Support Systems*, Vol. 32(2), pp. 171-187, 2001.

[3] 김정열, "전자상거래의 주요 성공요인에 관한 연구," *한국컴퓨터정보학회 논문집*, 10권 3호, pp. 267-273, 2005년 7월.

[4] Wells JD, Fuerst WL, Choobineh J., "Managing information technology IT for one-to-one customer interaction," *Information and Management*, Vol. 35(1), pp. 53-62, 1999.

[5] 김종완, 오기욱, "E-CRM을 위한 개인화 마케팅 기법에 관한 연구," *한국컴퓨터정보학회 논문집*, 7권 2호, pp. 109-114, 2002년 7월.

[6] Chen MS, Han J, Yu PS., "Data mining: an overview from a database perspective," *IEEE Trans. on Knowledge and Data Engineering*, Vol. 8(6), pp. 203-213, 1996.

[7] Hendryx MS, Wakefield DS, Murray JF, Uden-Holman T, Helms CM, Ludke R., "Using comparative clinical and economic outcome information to profile physician performance," *Health Services Manage Research*, Vol. 8(4), pp. 213-220, 1995.

[8] MacStravic RS., "Hospital-physician relations: a marketing approach," *Health Care Management Review*, Vol. 11(3), pp. 69-79, 1986.

[9] Bueren A, Schierholz R, Kolbe L, Brenner W., "Customer knowledge management-improving performance of customer relationship management with knowledge management," *Proceedings of*

the 37th Hawaii International Conference on System Sciences, pp. 122-134, 2004.

[10] Gronroos C., "From marketing mix to relationship marketing: towards a paradigm shift in marketing," *Management Decision*, Vol. 32(2), pp. 4-20, 1994.

[11] Ahwah I, Karpel M., "Using profiling for cost and quality management in the emergency department," *Healthcare Financial Management*, Vol. 51(7), pp. 50-53, 1997.

[12] Helfert G, Vith K., "Relationship marketing teams-improving the utilization of customer relationship potentials through a high team design quality," *Industrial Marketing Management*, Vol. 28(5), pp. 553-564, 1999.

[13] Rygielski C, Wang JC, Yen DC., "Data mining techniques for customer relationship management," *Technology in Society*, Vol. 24(1), pp. 483-502, 2002.

저자 소개



이 상 영

1994년 숭실대학교 산업공학과 (공학사)
 1998년 전북대학교 산업공학과 (공학석사)
 2004년 전북대학교 컴퓨터과학과 (이학박사)
 2005년~현재 남서울대학교
 보건행정학과 전임강사
 <관심분야> E-Health, 전자상거래, CRM, 온톨로지 등



이 윤 석

1996년 중앙대학교 의과대학 의학과 (의학사)
 2000년 연세대학교 보건대학원 병원 행정학과(보건학석사)
 2004년 연세대학교 보건대학원 (보건학박사)
 2002년~현재 남서울대학교 보건행정학과 조교수
 <관심분야> 보건행정, 임상, 의무기록, E-Health, 등