

정신과 환자 처방분석을 통한 의병전역 결정지원 시스템

(DSS of Discharging from Military Service
 using the Analysis of Mental Patient's Prescription)

전 영 희(Young-Hee Jeon)*, 박 건 우(Gun-Woo Park)**, †이 상 훈(Sang-Hoon Lee)***

초 록

최근 정신질환에 대한 사회적 편견이 감소하면서 이차성 이득(secondary gain), 즉 군복무 기피를 위한 목적으로 정신과적 증상을 호소하는 환자가 늘어나고 있다. 특히 객관적 진단도구가 없는 정신의학적 문제의 경우 환자의 주관적 증상호소와 의사의 전문적 판단이 실제 유일한 진단 기준이다. 본 논문에서는 데이터마이닝을 이용하여 과거 정신과 입원 환자들의 처방데이터 분석을 통해 의병전역 결정을 위한 객관적 기준을 제공하여 의무조사의 신속한 의사결정에 도움을 주고자 한다. 따라서 정신과 환자 처방분석을 통한 의병전역 결정지원 시스템을 제안한다.

ABSTRACT

Recently, mental patients are increasing by secondary gain, that is, purpose for avoiding a military service as the social prejudice about mental disease decreases. In particular, it is unique diagnostic tool to a patient's subjective symptoms complaint and a doctor's special judgment, in occasion of psychiatry problem that an objective diagnostic tool does not exist. In this paper, we provide an objective basis to help in a quick decision-making of discharging from military service using the datamining, that analyzes mental patient's prescription to find a special rule. Therefore, we propose the decision support system of discharging from service using the analysis of mental patient's prescription.

Keywords : 데이터마이닝, 정신과 처방데이터, 의병전역 결정지원 시스템

논문접수일 : 2008년 11월 24일 논문게재확정일 : 2008년 12월 10일

* 국군군의학교

** 국방대학교 전산정보학과

*** 국방대학교 국방과학부 교수

† 교신저자

1. 서론

정보화 시대의 도래와 함께 현대의 병원들에 있어서 정보시스템은 빼놓을 수 없는 병원 내 주요 기반시설로 등장하게 되었다. 정보시스템의 효과적인 도입은 병원 내 업무 운영 효율의 혁신적인 향상을 가져와 운영경비의 절감과 함께 환자에 대한 양질의 서비스를 제공하는 것이 가능해지기 때문이다. 정보 시스템의 도입 효과를 극대화하기 위해서는 부분적인 전산화 수준에서 벗어나 병원 내 모든 정보를 통합 관리할 수 있는 병원의 전반적인 분야에 대해 처리 가능한 정보시스템 구현이 요구되고 있다[4].

이와 같은 정보시스템과 관련하여 최첨단 디지털 병원들이 개원되는 현재는 디지털 병원 구축의 근간인 전자의무기록의 시대라 할 수 있다. 의료기관 내에서 발생하는 모든 진료기록을 전산화하여 통합의료정보시스템 구축을 목표로 하고 있다 [11]. 이를 위해 의료정보 통계와 각종 연구목적 및 의료서비스 경쟁력 향상 등을 위한 신속 정확한 의사결정 지원 시스템인 데이터 웨어하우스(Data Warehouse: DW)와 이러한 방대한 데이터로부터 유용한 정보를 추출하는 데이터마이닝은 필수적인 구성요소라 할 수 있다. 현재 우리 군에서도 국방의료정보체계(DEMIS)가 군 병원뿐만 아니라 사단급 의무부대까지 확산됨에 따라, 이를 연계한 군 의료 데이터웨어하우스 구축 사업을 올해 말부터 추진해 나갈 계획이다[8, 17].

최근 군내에서는 정신질환에 대한 사회적 편견이 감소하면서 이차성 이득(secondary gain), 즉 군복무 기피를 위한 목적으로 정신과적 증상을 호소하는 환자가 늘어나고 있다. 특히 객관적 진단도구가 없는 정신의학적 문제의 경우 환자의 주관적 증상호소와 의사의 전문적 판단이 실제 유일한 진단 기준이 되고 있어, 정신과적 증상의 이차성 이득 여부에

대한 판단은 매우 어려운 현실이다[9].

따라서 정신질환자의 진단 특성을 고려시, 의병전역을 위해 객관성을 부여할 수 있는 기준 마련의 필요성이 대두되고 있다.

본 논문에서는 군 병원 전자의무기록의 원천시스템인 국방의료정보체계로부터 정신과 입원환자의 처방데이터를 추출하여, 데이터마이닝을 이용한 약품 처방분석을 통해 의병전역 결정을 위한 객관적 기준을 제공하여 의무조사의 신속하면서도 적극적인 의사결정 지원에 도움을 주고자 한다. 따라서 정신과 환자 처방분석을 통한 의병전역 결정지원 시스템을 제안한다.

2. 관련연구

이 절에서는 정신과 환자 처방분석을 통한 의병전역 결정지원 시스템의 선행연구로서 국방의료정보체계의 특성과 의무조사, 그리고 본 연구의 핵심인 데이터마이닝 기법에 대해 설명한다.

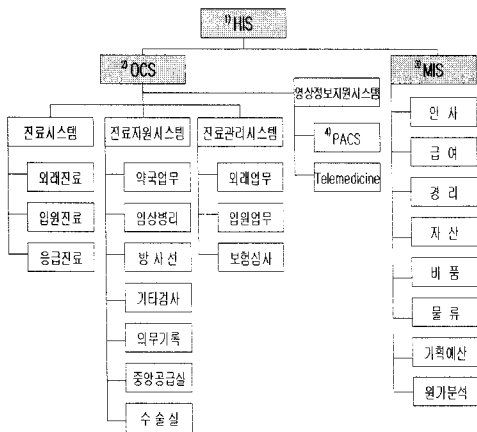
2.1 국방의료정보체계(DEMIS)

국방의료정보체계(Defense Medical Information System: DEMIS)는 군 병원 전자의무기록의 원천시스템(Data Source)으로써 군병원 정보화를 통한 효율적인 의무지원체계의 구축과 군병원 관리능력 향상을 목적으로 기존의 수기문서를 DB화하여 필요한 의료자료를 온라인(On-Line)으로 실시간(Real time) 입·출력하는 시스템이라 할 수 있다.

그 기능을 살펴보면 <그림 1>과 같이 크게 진료시스템, 진료지원 시스템, 진료관리 시스템, 병원전반적 행정업무를 담당하는 병원행정관리 시스템으로 구분되며, 진료 지원 시스템에 대한 보완 정보시스템으로 영상정보지원시스템이 있다 [3].

1) HIS : Hospital Information System(병원정보시스템)

2) OCS : Order Communication System(처방전달시스템)



〈그림 1〉 국방의료정보체계(DEMIS)의 구성

DEMIS 서버와 각 부처의 PC들은 <그림 2>와 같이 UNIX 운영체제를 기반으로 하는 통합 DB 서버를 중심으로 Giga Bit 백본망을 이용하여 하나의 네트워크를 구성한다.

국방의료정보체계는 의무기록과 조회의 편리성을 향상시켜 업무의 효율을 높였다고 할 수 있다. 그러나 각각의 독립된 17개 군 병원으로부터 받

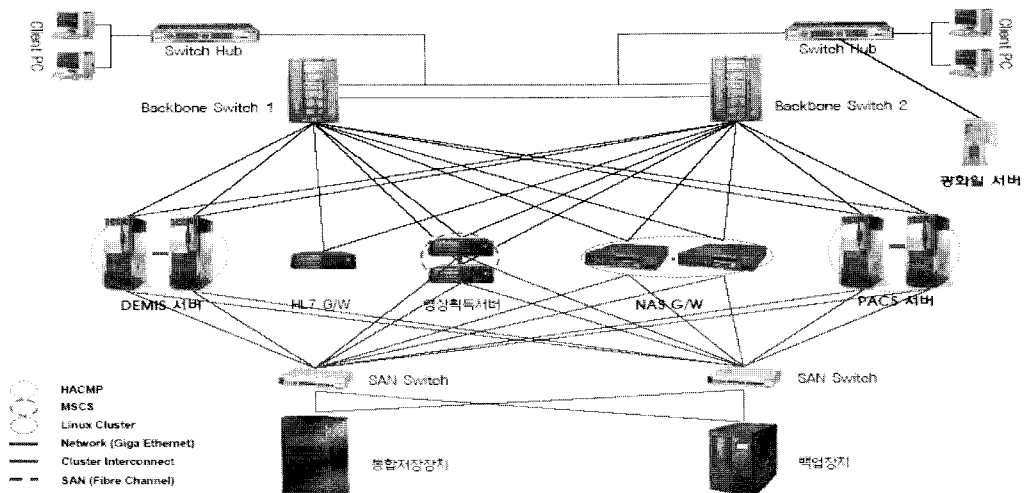
생하는 보고서 및 통계 자료 등이 각각 존재하고, 단순 현황 조회로 이력 데이터에 의한 추세 분석이 어려우며, 보고서 생성을 위해서 다수의 인력과 시간이 소요되고, 동일한 분석 항목에 대해 각각 다른 결과가 발생하는 등의 문제점을 가지고 있다.

따라서 각 병원과의 데이터 통합 및 공유를 위한 군 의료 데이터웨어하우스는 데이터마이닝의 정확한 정보 도출과 안정적인 환경 제공을 위한 필수 조건이라 할 수 있다[10].

2.2 의무조사

심신장애인 전역은 장병들이 질병 및 장애로 인해 군 복무가 불가능한 경우에 군 병원에 입원하여 심신장애전역 절차에 따라 전역을 하는 경우를 말하며, 이후 본 연구에서는 의병전역으로 표현하였다.

그 절차는 <그림 3>과 같이 환자가 군병원에 입원하여 신체등급 판정을 받은 후에 환자진료 중



〈그림 2〉 네트워크 구성도

3) MIS : Management Information System(경영정보시스템)

4) PACS : Picture Archiving Communication System(의료영상저장전송시스템)

결권이 부여된 군병원에 설치된 의무조사위원회 의 의무조사를 거쳐 각 군 본부의 병역처분 변경 심사를 통해 병역처분을 하게 된다.



<그림 3> 의병전역 절차

즉, 의무조사란 진료종결 병원인 군단지원병원, 군지원병원, 국군후방병원에서 의무조사 위원회를 설치하여 군병원 또는 민간의료기관에 입원한 심신장애인에 대하여 심신장애의 정도와 진료의 계속 및 퇴원여부 등을 결정하는 것이다.

병사의 경우에는 국방부령 제645호 징병 신체 검사등 검사규칙에 의해 5급~6급에 해당되면 의병전역 대상자가 되고, 간부의 경우에는 국방부령 제 631호 군인사법 시행규칙 별표 1 심신장애 등급표, 별표 2 심신장애 종합평가 등급표에 의해 1급~9급에 해당되면 의병전역 대상자가 된다. 단, 8, 9급은 본인의 의사에 따라 복무가 가능하다[1, 2, 6].

2.3 데이터마이닝 기법

데이터마이닝에 대한 정의는 여러 학자들에 의해 많이 정의되었는데, 간략하게 정리하면 다음과 같다.

- 대용량 데이터베이스로부터 이제까지 몰랐던 정보를 추출하는 과정
- 패턴 인식 기술이나 통계 기법, 수학적 알고리즘을 이용하여 의미 있는 새로운 상관관계, 패턴, 추세 등을 발견하고 예측하는 과정
- 자동화되고 지능을 갖춘 데이터베이스 분석 기법으로 대량의 데이터로부터 새롭고 의미 있는 정보를 추출하여 의사결정에 이용하는 작업

즉, 데이터마이닝은 하나의 기법으로 이루어진 것이 아니다. 여러 가지 기법들은 서로 다른 목적으로 이용될 수 있으나 현재 주목되는 기법들은 질의도구(query tools), 통계적 기법(statistical technique), 가시화(visualization), 온라인 분석 처리(OLAP: online analytical processing), 사례-기반 학습(case-based learning), 동시발생 매트릭스 (Co-Occurrence Matrix), 의사결정나무(decision tree), 연관규칙(association rule), 신경망(neural network), 유전자 알고리즘(genetic algorithm) 등이 있다. 본 연구에서는 의병전역에 있어 결정적인 영향요소 파악하고 구체적인 의병전역의 분류 기준을 찾기 위해 회귀분석과 의사결정나무를 이용하였다[7, 12, 14, 15].

데이터 마이닝을 통하여 생성된 지식을 활용하여 일반 기업에서는 소비자의 미래 소비 성향이나 제품에 대한 판매 예측을 하며, 금융 업계에서는 고객의 신용 등급, 금리, 환율 변동 등을 예측한다. 그리고 공장에서는 제품 생산비용에 대한 예측과 설비에 대한 결함 예측을 하며 의학분야에서는 환자의 병원 방문 예측과 효과적인 의학치료방법 예측에 이용한다[5, 13, 16].

3. 시스템 설계

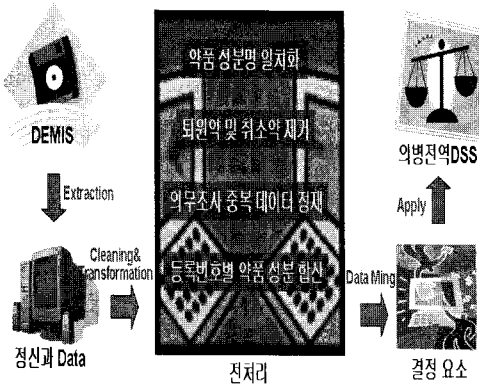
3.1 실험 데이터

본 연구의 데이터는 3개 군병원의 정신과 입원 환자를 대상으로 2006년도부터 2008년도 7월까지 2년 7개월간의 처방정보를 포함한 트랜잭션 데이터를 추출하여 사용하였다. 추출된 데이터는 13개 변수를 포함한 28,842건이며 전처리 과정을 거쳐 최종 실험데이터는 859건을 사용하였다.

3.2 실험 방법

본 연구는 <그림 4>과 같이 크게 3단계로 나누

어 분석하였고, 데이터마이닝 툴은 SAS 9.1 Enterprise Miner를 이용하였다.



〈그림 4〉 실험 PROCESS

첫째, 군병원 전자의무기록의 원천시스템인 국방의료정보체계에서 정신과 처방정보를 포함하는 트랜잭션 데이터를 추출한다. 추출된 데이터는 병원, 등록번호, 약품성분명, 복용일수, 처방일, 재원기간, 의무조사구분, 입원횟수, 상병코드, 상병명, 약품함량, 처방코드, 세부내용의 필드를 포함한다.

둘째, 데이터마이닝 결과 산출되는 정보의 정확도를 높이기 위해 추출데이터에 존재하는 오류값이나 특이값을 보정 및 중복데이터를 제거하고, 데이터마이닝에 필요한 형태로 전처리 작업 수행하였다.

셋째, 정신과 입원환자 처방데이터 중 객관성을 부여할 수 있는 약품 복용량과 재원기간 데이터를 데이터마이닝 툴 SAS Enterprise Miner에 적용하여 의병전역 판단을 위한 기준을 찾아내고 유효성을 검증한다.

넷째, 검증된 의병전역 판단 기준 데이터를 의병전역 결정지원 시스템에 적용한다.

3.3 의병전역 결정지원 시스템 설계

의병전역 결정지원 시스템이란 의무조사 대상자 처방데이터를 입력 받으면, 등록번호별 의병전

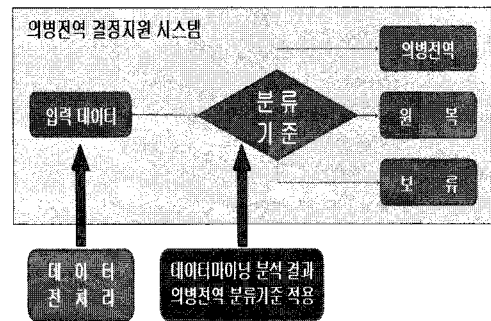
역 결정 요소를 비교하여 자동으로 의병전역 여부를 판단 및 분류해주는 시스템이다. 즉, 사람이 수작업으로 검토하여 고민하는 단계를 자동화함으로써 신속한 의사결과와 의무조사 업무 수행에 도움을 주는 시스템이라 할 수 있다.

본 연구의 의병전역 결정지원 시스템은 <그림 5>와 같이 구성된다.

먼저, 데이터마이닝 분석결과 의병전역 결정요소 및 기준이 도출되면 그것을 의병전역 결정지원 시스템의 분류 기준으로 적용한다.

그리고 전처리 된 처방 데이터를 엑셀 파일로 입력 받으면, 스트링 배열로 변환 후 벡터 처리하고 시스템의 의병전역 결정 기준 조건과 비교한 후, 자동으로 “의병전역”, “원복”, “보류(의무조사 심의 대상)”로 분류 및 출력하는 시스템을 설계하였다.

따라서 본 연구의 의병전역 결정지원 시스템을 구현하기 위해서는 입력데이터의 전처리와 데이터마이닝을 통한 구체적인 의병전역 분류기준 결정은 사전에 이뤄져야 한다.



〈그림 5〉 의병전역 결정지원 시스템 설계

4. 실험결과

4.1 추출데이터 전처리

추출데이터의 전처리 결과는 <표 1>과 같다.

〈표 1〉 추출데이터 전처리 결과

구 분	데이터 ROW (건)			약품수 (개)
	소 계	원복 (A)	의병전역 (J)	
전 처 리 전	28,842	14,230	14,612	417
정신과 약품	22,915	11,500	11,415	38
성 분 별	3,285	1,564	1,721	38
효 능 별	2,181	1,110	1,071	8
복용량 합산	859	472	387	1
증 감	-27,983	-13,758	-14,225	- 416

산출물 결과의 정확도를 높이기 위해 <그림 6>과 같이 처방데이터에서 복지부분류코드별로 분류하여 정신과 의약품만을 선별하고(417개 약품 28,842건→ 38개 약품 22,915건), 의무조사 이후 처방된 퇴원약과 취소약 등 불필요한 데이터의 제거와 중복된 의무조사 식별 데이터(A: 원복/J: 의병전역)를 정제하였다(22,915건→ 20,191건).

그리고 약품 성분과 효능에 대한 영향과 복용량에 대한 영향을 분석하기 위해 다음과 같이 데이터를 합산하였다.

약품 성분별 합산 과정에서는 같은 성분이라도 등록번호별로 사용한 약품의 단위 함량이 상이한 경우가 있어, 표준화를 위해 같은 성분의 가장 작은 함량을 기준으로하여 1로 하고, 함량에 비례하

〈표 2〉 약품 성분별 복용량 표준화

등록번호	성분명	함량 (mg)	가중치 (표준화 점수)	총 복용량 (개)
200783010602	Lorazepam	0.5	1	28
200071000365	Lorazepam	1	2	1
200071000365	Lorazepam	4	8	1

등록번호	총 복용량 (표준화 적용) (개)
200783010602	28 = 28 × 1
200071000365	2 = 1 × 2
200071000365	8 = 1 × 8

※ 총 복용량(표준화 적용) = 총 복용량 × 가중치

여 가중치(표준화점수)를 부여하여 복용량과 가중치를 곱해 표준화된 복용량을 구하였다. <표 2>는 약품 성분별 복용량 표준화의 예이다.

그리고 다음 단계로 <그림 7>과 같이 등록번호별로 같은 성분은 합산하였다. 그 결과 실험데이터 수는 38개 성분 3,285개 데이터로 대폭 감소하였다.

약품 효능별 합산과정에서는 먼저, 약품정보에서 성분별 1일 권장 복용량(mg)과 대표 효능을 찾고, 성분별 1일 권장 복용량(mg)을 기준 함량(mg)으로 나누어 1일 복용 약 수(개)를 구한다. 그리고, 각기 다른 성분을 같은 효능별로 합산하

성분명(원형)	복지부분류코드	의병부분류코드
Chloramphenicol 50mg/10ml/BT	615	주요 감염성성 음질균, 리케치아, 비
Chloramphenicol 5mg/ml, Dexamethasone 1mg/ml	131	안과용제
Chlorhexidine gluconate	281	외피음질균소독제
Chlorphenesin carbonate 125mg	122	물국괴이원제
Chlorpheniramine maleate 4mg/2ml	141	항히스타민제
Chlorpromazine 100mg	117	정신신경용제
Clopirox	285	기생성 피부질환용제
Cineidine 200mg	232	소화성제염용제
Ciprofloxacin	629	기타의 화학요법제
Clarithromycin 250mg	619	기타의 항생물질제
Clindamycin phosphate 1.138g/100ml(1%), 30ml	619	기타의 항생물질제
Clobetasol Propionate	284	진통, 소염, 소염제
Clozapine 25mg	117	정신신경용제
Clozapine 0.5mg	118	항전진제
Clobazam	285	비노성식기근용제(성병예방제 포함)
Clozapine 100mg	117	정신신경용제
Colohe chlorate 20mg	811	아편알카로이드제 제제



성분명(원형)	복지부분류코드	의병부분류코드
Alprazolam	117	정신신경용제
Amitriptyline	117	정신신경용제
Arpiprazole	117	정신신경용제
Benzocaine	118	기타의 중추신경용약
Biperiden HCl	119	기타의 중추신경용약
Bupropion HCl	117	정신신경용제
Buspirone	117	정신신경용제
Carbamazepine	113	항전진제
Chlorpromazine	117	정신신경용제
Clomipramine	117	정신신경용제
Clozapem	113	항전진제
Clozapine	117	정신신경용제
Diazepam	117	정신신경용제
Escitalopram	117	정신신경용제
Fluoxetine	117	정신신경용제
Fluvoxamine maleate	117	정신신경용제
Gabapentin	113	항전진제

〈그림 6〉 처방 성분코드별 분류

1	등록번호	약품명	정신과	약품명	복용일수	복용량	복용일수	복용량
2	200071000305	Haloperidol	정신과	정신과	300	1	3	5.5
3	200071000305	Haloperidol	정신과	정신과	300	1	3	5
4	200071000305	Haloperidol	정신과	정신과	300	1.5	3	5
5	200071000305	Haloperidol	정신과	정신과	300	2	3	5
6	200071000305	Haloperidol	정신과	정신과	300	2	4	16
7	200071000305	Haloperidol	정신과	정신과	300	2	5	20
8	200071000305	Haloperidol	정신과	정신과	300	2	6	5
9	200071000305	Haloperidol	정신과	정신과	300	2	7	26
10	200071000305	Haloperidol	정신과	정신과	300	2	7	29
11	200071000305	Haloperidol	정신과	정신과	300	2	7	29
12	200071000305	Haloperidol	정신과	정신과	300	2	7	29
13	200071000305	Haloperidol	정신과	정신과	300	2	7	29
14	200071000305	Haloperidol	정신과	정신과	300	2	7	29
15	200071000305	Haloperidol	정신과	정신과	300	2	7	29
16	200071000305	Haloperidol	정신과	정신과	300	2	7	29



1	등록번호	약품명	정신과	정신과	복용일수	복용량	복용일수	복용량
2	200071000305	Haloperidol	정신과	정신과	300	1	3	5.5
3	200071000305	Lithium carbonate	정신과	정신과	300	1	3	5
4	200071000305	Clonazepam	정신과	정신과	300	1.5	3	5
5	200071000305	Olanzapine	정신과	정신과	300	2	3	5
6	200071000305	Bontrilone	정신과	정신과	300	2	3	5
7	200071000305	Clonazepam	정신과	정신과	300	2	3	5
8	200071000305	Risperidone	정신과	정신과	300	2	3	5
9	200071000305	Sodium Valproate	정신과	정신과	300	2	3	5
10	200071000305	Lithium carbonate	정신과	정신과	300	2	3	5
11	200071000305	Olanzapine	정신과	정신과	300	2	3	5
12	200071000305	Somilene	정신과	정신과	300	2	3	5
13	200071000305	Haloperidol	정신과	정신과	300	2	3	5
14	200071000305	Lithium carbonate	정신과	정신과	300	2	3	5
15	200071000305	Clonazepam	정신과	정신과	300	2	3	5
16	200071000305	Olanzapine	정신과	정신과	300	2	3	5

<그림 7> 등록번호별 약품 성분별 데이터 합산

기 위한 표준화 방법으로 복용량(개)을 복용일수(일)로 변환하였다. 즉, 성분별 총 복용량(개)을 1일 복용 약 수(개)로 나누어 총 복용량(개)을 총 복용일수(일)로 표준화 및 수치화 하였다. 그리고 다음 단계로 <그림 8>과 같이 등록번호별로 같은 효능의 복용일수를 합산하였다. 그 결과 실험데이터 수는 8개 효능 2,181개 데이터로 감소하였다.

한 개 효능의 상대적인 복용량이 구해진다. 이처럼 전체 8개의 효능에 대해 등록번호별로 상대적 복용량을 구한 다음 <그림 9>와 같이 등록번호별로 상대적 복용량을 합산하면 표준화된 복용량이 구해진다. 그 결과 실험 데이터는 859개로 감소하였다.

1	등록번호	약품명	정신과	정신과	복용일수	복용량	복용일수	복용량
2	200071000305	Haloperidol	정신과	정신과	300	1	3	5.5
3	200071000305	Lithium carbonate	정신과	정신과	300	1	3	5
4	200071000305	Clonazepam	정신과	정신과	300	1.5	3	5
5	200071000305	Olanzapine	정신과	정신과	300	2	3	5
6	200071000305	Bontrilone	정신과	정신과	300	2	3	5
7	200071000305	Clonazepam	정신과	정신과	300	2	3	5
8	200071000305	Risperidone	정신과	정신과	300	2	3	5
9	200071000305	Sodium Valproate	정신과	정신과	300	2	3	5
10	200071000305	Lithium carbonate	정신과	정신과	300	2	3	5
11	200071000305	Olanzapine	정신과	정신과	300	2	3	5
12	200071000305	Somilene	정신과	정신과	300	2	3	5
13	200071000305	Haloperidol	정신과	정신과	300	2	3	5
14	200071000305	Lithium carbonate	정신과	정신과	300	2	3	5
15	200071000305	Clonazepam	정신과	정신과	300	2	3	5
16	200071000305	Olanzapine	정신과	정신과	300	2	3	5

1	등록번호	약품명	정신과	정신과	복용일수	복용량	복용일수	복용량
2	7100	200071000305	A	정신과	1.3	0.0024	58	58
3	7100	200071000305	A	정신과	0.0024	0.0024	58	58
4	7100	200071000305	A	정신과	0.0024	0.0024	58	58
5	7100	200071000305	A	정신과	17.6	0.01481	58	58
6	7100	200071000305	A	정신과	0.0024	0.0024	58	58
7	7100	200071000305	A	정신과	0.0024	0.0024	58	58
8	7100	200071000305	A	정신과	13.3	0.00897	196	196
9	7100	200071000305	A	정신과	15	0.00217	196	196
10	7100	200071000305	A	정신과	0.0024	0.0024	196	196
11	7100	200071000305	A	정신과	0.0024	0.0024	34	34
12	7100	200071000305	A	정신과	24.4	0.00519	34	34
13	7100	200071000305	A	정신과	31.6	0.00101	34	34
14	7100	200071000305	A	정신과	3	0.0084	51	51
15	7100	200071000305	A	정신과	57	0.00163	51	51
16	7100	200071000305	A	정신과	0.0024	0.0024	51	51
17	7100	200071000305	A	정신과	0.0024	0.0024	51	51
18	7100	200071000305	A	정신과	0.0024	0.0024	51	51
19	7100	200071000305	A	정신과	0.0024	0.0024	51	51
20	7100	200071000305	A	정신과	0.0024	0.0024	51	51
21	7100	200071000305	A	정신과	0.0024	0.0024	51	51
22	7100	200071000305	A	정신과	0.0024	0.0024	51	51
23	7100	200071000305	A	정신과	0.0024	0.0024	51	51
24	7100	200071000305	A	정신과	0.0024	0.0024	51	51
25	7100	200071000305	A	정신과	0.0024	0.0024	51	51
26	7100	200071000305	A	정신과	0.0024	0.0024	51	51
27	7100	200071000305	A	정신과	0.0024	0.0024	51	51
28	7100	200071000305	A	정신과	0.0024	0.0024	51	51
29	7100	200071000305	A	정신과	0.0024	0.0024	51	51

<그림 8> 등록번호별 약품 효능별 데이터 합산

<그림 9> 등록번호별 정신과 약품 복용량 합산

정신과 약품 복용량 합산과정에서는 각기 다른 효능을 하나의 약품으로 합산하기 위해 다음과 같이 각 효능의 전체 복용량을 1로 봤을때 등록번호에 해당되는 효능의 상대적 복용량을 구하였다. 즉, 한 개의 효능 복용일수를 모두 합산한 다음, 등록번호 해당 복용일수를 총 합산일수로 나누면,

4.2 데이터마이닝 사용 데이터 변수 및 분할

전처리 결과 데이터마이닝에 사용한 데이터 변수(Value)와 분할(Partition)은 <그림 10>, <그림 11>과 같다. 입력데이터는 등록번호를 ID로, 의 무조사를 Target으로, 복용량과 재원기간을 Input으로 변수를 지정하였고, 데이터 셀은 데이터마이

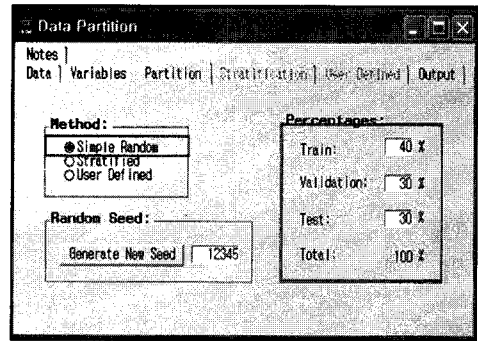
Name	Model Role	Measurement	Type	Format	Informat	Variable Label
VAR1	rejected	ordinal	num	BEST12	BEST12	병원
VAR2						
VAR3						
VAR4						
VAR5						
VAR6	rejected	ordinal	date	YYMMDD	YYMMDD	처방일자

〈그림 10〉 데이터 변수(Value)

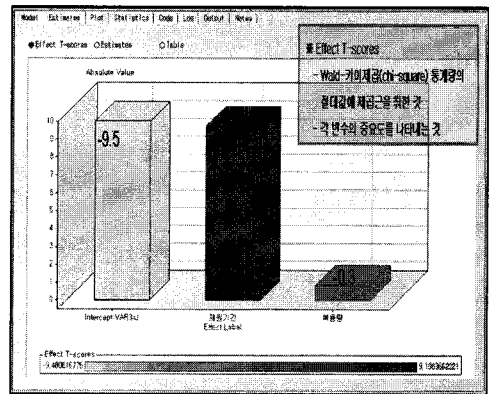
닝 분석용(Train) 40%, 모델 평가용(Validation) 30%, 검증용(Test) 30%로 분할하여 사용하였다.

4.3 정신과 처방 데이터 분석 결과

정신과 입원환자 처방 데이터 분석을 위한 데이터마이닝 분석 흐름도는 <그림 12>와 같다. 결정적인 영향요소 및 Target 데이터의 구체적인 분류 기준을 찾기 위해 회귀분석과 의사결정나무를 이용하였다.

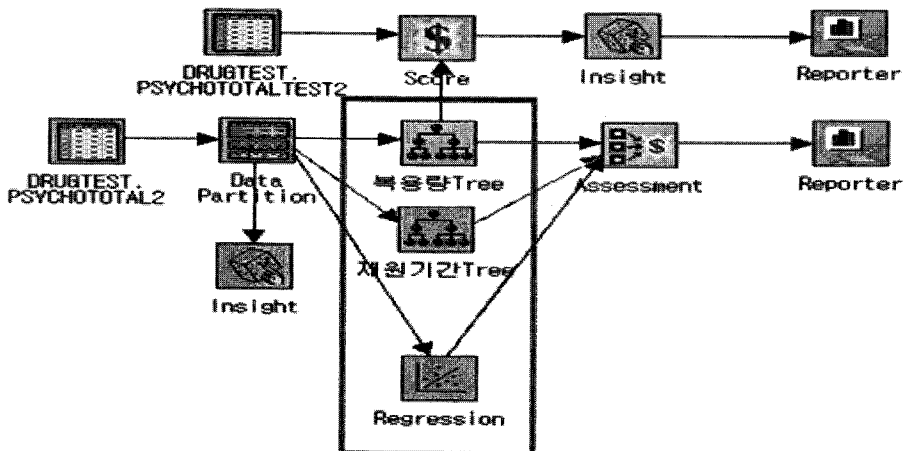


〈그림 11〉 데이터 분할(Partition)

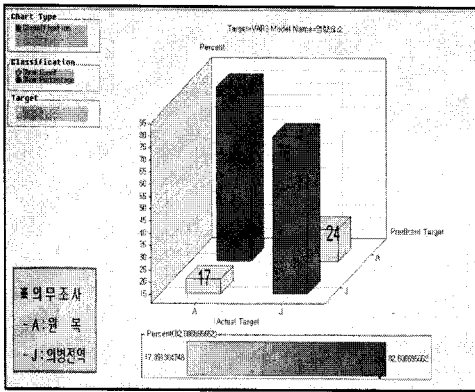


〈그림 13〉 회귀분석 결과 Effect T-scores

회귀분석을 통해 <그림 13>과 같이 입력변수 재원기간이 목표변수 의병전역 여부에 가장 큰 영



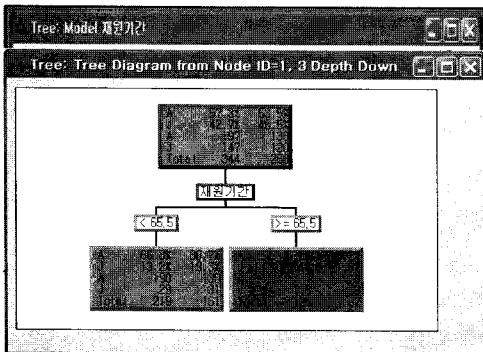
〈그림 12〉 데이터마이닝 분석 흐름도



〈그림 14〉 회귀분석 진단 도표

향을 미치는 요소임을 알 수가 있었고, 진단도표는 <그림 14>와 같이 나타났다. 즉, 원복자가 원복인 경우가 83%, 의병전역자가 의병전역인 경우는 76%로 비교적 예측률이 높다는 것을 알 수 있다.

다음은 재원기간의 구체적인 의병전역 결정 분류 기준을 찾아내기 위해 재원기간을 Input으로, 의무조사를 Target으로 하여 의사결정나무 분석을 해 본 결과 <그림 15>와 같이 65.5일을 기준으로 의병전역 여부가 확연히 구분되어지는 것을 알 수가 있다. 재원기간 ≥ 65.5 일인 경우, 분석용 (Train) 모델에서는 전체 125명 중에 의병전역자가 118명, 즉 재원기간이 65.5일 이상인 환자의 94%가 의병전역을 했음을 알 수가 있고, 평가용 (Validation) 모델에서도 재원기간이 65.5일 이상인 환자의 92%가 의병전역했음을 알 수가 있다.



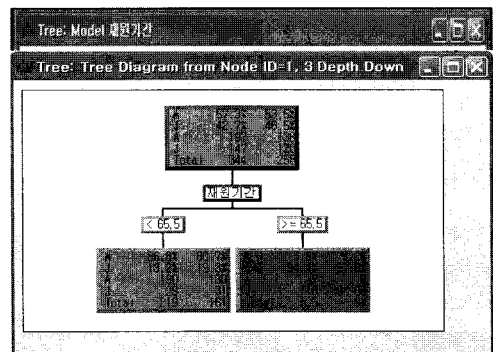
〈그림 15〉 재원기간 의사결정나무 분석 결과

역으로 재원기간 < 65.5 일인 경우, 분석용 모델에서는 전체 219명 중에 원복자가 190명, 즉 재원기간이 65.5일 미만인 환자의 87%가 원복했고, 평가용 모델에서는 81%가 원복했음을 알 수가 있다.

다음은 약품 복용량의 구체적인 의병전역 결정 분류 기준을 찾아내기 위해 복용량을 Input으로, 의무조사를 Target으로 하여 의사결정나무 분석을 해 본 결과 <그림 16>과 같이 0.00292를 기준으로 의병전역 여부가 구분되어 진다.

분석용 모델에서 보여주는 바와 같이 복용량이 0.00292이상인 경우 의병전역자와 원복자가 55% 대 45%로 크게 차이가 나지 않아 재원기간처럼 확연한 분류기준 값으로 보기에는 어렵지만, 복용량이 0.00292 미만인 경우에는 분석용 모델에서 전체 153명 중에 원복자가 111명, 즉 복용량이 0.00292 미만인 환자의 73%가 원복을 했고, 평가용 모델에서는 74%가 원복 했음을 알 수가 있다. 즉, 복용량도 의병전역 결정요소에 영향을 미치는 것을 알 수 있다.

따라서 데이터마이닝 실험결과 회귀분석을 통해 재원기간이 의병전역 결정에 있어 가장 많은 영향을 미친다는 것을 알 수 있었고, 의사결정나무 분석을 통해 의병전역의 결정요소인 약품 복용량과 재원기간의 구체적인 분류기준을 다음과 같이 도출할 수 있었다. 재원기간 65.5일을 기준으로 그 이상일 경우는 의병전역으로, 그 미만일 경



〈그림 16〉 약품 복용량 의사결정나무 분석 결과

우는 원복으로 분류기준을 정할 수 있고, 복용량 0.00292를 기준으로 그 이상일 경우는 의병전역으로, 그 미만일 경우는 원복으로 분류기준을 정할 수 있다.

기타 약품 성분과 효능과의 영향관계를 실험해 보았으나 의병전역과 관련된 결정적인 요인을 찾지 못했다.

4.4 정신과 약품 복용량과 재원기간 영향 평가

데이터마이닝 실험결과 의병전역의 결정적인 영향요소로 나타난 약품 복용량 0.00292와 재원기간 65.5일의 유효성을 비교 평가 데이터 <표 3>을 이용하여 아래 조건을 만족하는 원복자와 의병전역자의 빈도 수와 정확도를 비교하여 평가하였다.

<표 3> 비교 평가 데이터

구 분	비 율 (%)	데이터 ROW (건)		
		소 계	원 복(A)	의병전역(J)
소 계	100	859	472	387
분 석 용 (Train)	40	344	197	147
평 가 용 (Validation)	30	258	138	120
검 증 용 (Test)	30	257	137	120

- 조건 A : 약품 복용량 ≥ 0.00292 → 전역
- 조건 B : 재원기간 ≥ 65.5 일 → 전역
- 조건 C : $A \cap B$ → 전역
- 조건 X : 약품 복용량 < 0.00292 → 원복
- 조건 Y : 재원기간 < 65.5 일 → 원복
- 조건 Z : $X \cap Y$ → 원복

$$\text{※정확도}(\%) = \frac{\text{전역(원복) 빈도수}}{\text{전역(원복) 대상자}} \times 100$$

비교 평가 결과는 <표 4>, <표 5>와 같다.

<표 4> 조건 A, B, C 비교 평가결과

구 분	A (복용량 \geq 0.00292)			B (재원기간 ≥ 65.5 일)			C (A \cap B)			비 고
	계	원 복	전 역	계	원 복	전 역	계	원 복	전 역	
분 석 용 (Train)	191	86	105	125	7	118	98	5	93	건
	100	45	55 (71)	100	6	94 (80)	100	5	95 (63)	% (전체 전역 대비)
평 가 용 (Validation)	146	55	91	97	8	89	78	5	73	건
	100	38	62 (76)	100	8	92 (74)	100	6	94 (61)	% (전체 전역 대비)
검 증 용 (Test)	151	60	91	103	8	95	80	2	78	건
	100	40	60 (76)	100	8	92 (79)	100	2	98 (65)	% (전체 전역 대비)

<표 4>에서 조건 A를 만족하는 경우 의병전역자의 비율이 55%~62%로 낮은 편이고, 조건 B를 만족하는 경우는 92%~94%로 높은 편이다. 즉, 재원기간은 의병전역 결정의 중요 요소임이 확실하고, 복용량은 정확도가 다소 떨어지는 듯하나, 조건 C에서 보는 바와 같이 조건 B를 만족하는 대부분이 조건 A를 만족한다는 것을 알 수가 있으므로, 조건 A, 즉 약품 복용량 또한 의병전역 결정에 있어 중요 요소로 볼 수 있음을 알 수 있다. 그리고 조건 A와 B를 동시에 만족하는 조건 C의 경우, 의병전역자의 빈도수는 약간 줄어드는 반면 데이터 분류결과의 정확도는 94%~98%로 향상됨을 알 수 있다.

<표 5>에서 조건 X를 만족하는 경우 원복자의 비율은 73%~74%이고, 조건 Y를 만족하는 경우는 81%~87%로 높은 편이다. 그리고, 조건 X와 Y를 동시에 만족하는 조건 Z의 경우, 원복자의 빈도수는 줄어드는 반면, 데이터 분류결과의 정확도는 85%~86%로 향상됨을 알 수 있다.

〈표 5〉 조건 X, Y, Z 비교 평가결과

구 분	X (복용량 < 0.00292)			Y (재원기간 < 65.5일)			Z (X ∩ Y)			비고
	계	원복	전역	계	원복	전역	계	원복	전역	
분석용 (Train)	153	111	42	219	190	29	114	97	17	건
	100	73 (56)	27	100	87 (96)	13	100	85 (49)	15	% (전체 원복 대비)
평가용 (Validation)	112	83	29	161	130	31	93	80	13	건
	100	74 (60)	26	100	81 (94)	19	100	86 (58)	14	% (전체 원복 대비)
검증용 (Test)	106	77	29	154	129	25	83	71	12	건
	100	73 (56)	27	100	84 (94)	16	100	86 (52)	14	% (전체 원복 대비)

4.5 의병전역 결정지원 시스템 적용 및 구현

위 평가 결과를 토대로 정확도가 가장 높은 조건 C와 Z를 의병전역 결정지원 시스템의 분류기준으로 적용했을 때, <표 6>은 자동분류 결과를 나타낸다.

〈표 6〉 의병전역 결정지원 시스템 분류결과

구 분	소 계		원 복		의 병 전 역		비 고	
	건수	%	건수	%	건수	%		
원 본	859	100	472	100	387	100	-	
조건 만족 (CUZ)	계	541	63	290	61	251	65	분류
	일치	492	91	248	86	244	97	정확도
	불일치	49	9	42	14	7	3	-
조건 외 (보류)	318	37	182	39	136	35	의무조사	

약품 복용량과 재원기간 기준을 동시에 만족할 경우 즉, 조건 C를 만족하는 경우는 의병전역으로, 조건 Z를 만족하는 경우는 원복으로, 그리고

기타의 경우는 보류, 즉 실제 의무조사 대상자로 자동분류 할 수 있다. 그 결과 의무조사 대상 인원의 63%를 자동으로 분류할 수 있다. 이때 분류 결과의 정확도는 91%이다. 따라서 실제 심의는 나머지 보류자 37%에 대해서만 시행하면 되므로, 기존 의무조사 업무량의 63%를 감소시킨 결과를 가져온다.

5. 결론 및 향후 과제

정신의학적 문제의 특성상 유일한 진단 기준은 환자의 주관적 증상호소와 의사의 전문적인 판단 뿐이다. 이러한 정신과 환자의 객관적 진단 도구가 없는 상황에서 이차성 이득, 즉 군복무 기피를 위한 목적으로 정신과적 증상을 호소하는 환자보다 신속하고 객관적으로 분별하기 위한 시스템이 필요하다.

본 연구에서는 군병원 정신과 입원환자에 대한 처방데이터를 이용하여 데이터마이닝 기법을 통해 의병전역 판단시 결정적인 영향을 미치는 요소로 재원기간과 약품 복용량을 도출할 수 있었고, 특히 재원기간 65.5일과 약품 복용량 0.00292를 기준으로 의병전역 결정지원 시스템에 적용시, 91%의 정확도로 전체 의무조사 대상자의 63%는 자동으로 분류해 낼 수 있었다. 즉, 의무조사는 나머지 37%의 환자에 대해서 집중적으로 분석하면 되므로 의무조사의 신속성에 기여할 수 있다.

그러나 임상에서는 예외적인 사례가 빈번히 발생 가능하므로 정신과 환자 진단시 당연 전문의의 판단이 최우선시 되어야 한다. 따라서 본 연구에서 제시한 정신과 환자 처방분석을 통한 의병전역 결정지원 시스템은 전문의의 의병전역 대상 여부 판단시 신뢰성 있는 객관적 자료를 제공함으로써 보다 신속하게 의사결정을 할 수 있도록 도와준다는 것에 큰 의의가 있다.

향후 연구과제로는 보다 정확하고 안정적인 정보 도출을 위해 군 의료 데이터웨어하우스 구축

후 본 의병전역 결정지원 시스템을 연계하여 유효성을 검증할 필요가 있다.

그리고 정신과 환자의 진단에 따른 주요(Major) 약품만을 추출하여, 추가 증상에 따른 부가(Miner)약품을 배제함으로써 여러가지 변수에 의한 영향을 최소화한 상태에서의 실험이 필요하다.

또한, 다양한 각도로 데이터마이닝 기법을 적용하여 보다 신뢰성 높은 모델을 지속적으로 개발 및 일반화 할 수 있도록 해야 한다.

참고문헌

[1] 국방부령 제 631호 “군인사법 시행규칙”, 2007.

[2] 국방부령 제 645호 “징병신체검사등 검사규칙”, 2008.

[3] 김준우, 김승기, 전동진 “군 병원정보시스템 구현의 중요 성공요인 분석 연구”, 경영정보학연구, 2004. 11.

[4] 김희중, 이창래, “PACS와 의료영상디스플레이 시스템”, 한국정밀공학회지, 2008.1.

[5] 노기용 외, “심장 질환 진단을 위한 데이터마이닝 기법”, 한국감성과학회, 2007.

[6] 안희만, “지휘관이 꼭 알아야 할 의료정보”, 국방부, 2006.

[7] 이상호, “C4.5 알고리즘을 이용한 의사결정나무 구성 및 탐색 컴포넌트 개발”, 울산대학교 컴퓨터 정보통신 공학부 석사논문, 2007.

[8] 이상훈 외, “국방의료종합정보체계(DW) 사전연구 결과”, 2008.

[9] 임태훈, “군대 내 인권상황 실태조사 및 개

선방안 연구”, 군의무발전 관련 토론회, 2006.

[10] 장남식, 홍성완, 장재호, “성공적인 지식경영을 위한 핵심 정보기술 데이터마이닝”, 대청, 2000.

[11] 전진옥, “디지털 병원 구현을 위한 전자 의무기록의 현황 및 전망”, 전자공학회지, 2006.

[12] Agrawal, R., arning, A., Bollinger, T., Mehta, M., Schafer, J., and Srikant, R., “The Quest Data Mining System,” Proceeding of the 2nd International Conference on Knowledge Discovery in Databases and Data Mining, Portland, Oregon, August, 1996.

[13] Han J., Kanmer M., “Data Mining : Concepts and Techniques”, Morgan Kamfmann, 2000.

[14] Hasan, H. and P. Hyland, “Using OLAP and multidimensional data for decision making”, IT Professional, Vol.3

[15] Mitra, S., S.K. Pal and P. Mitra, “Data mining in soft computation framework: aurbay”, IEEE Transactions on Neural Networks, Vol. 13, pp.3-14, 2002.

[16] Riccardo, B., and Blaz, Z., “Predictive data mining in clinical medicine: Current issue and guideline”, International Journal of Medical Informatics, 2006.

[17] T. B. Pedersen. Warehousing The World-A Few Remaining Challenges. In proc. of DOLAP, 2007.

■ 저자 소개 ■

전영희 (E-mail: fommy_hee@hanmail.net)

- 1998 인제대학교 보건대학 의용공학과 졸업(학사)
- 2009 국방대학교 국방관리대학원 전산정보학과 졸업(석사)
- 현재 국군군의학교 학생연대 운영과장
- 관심분야 데이터마이닝, 데이터웨어하우스, Business Intelligence

박건우 (E-mail: pgw4050@hotmail.com)

- 1997 충남대학교 컴퓨터과학과 졸업(학사)
- 2007 연세대학교 컴퓨터과학과 정보통신 연구실 졸업(석사)
- 현재 국방대학교 국방관리대학원 전산정보학과 박사과정
- 관심분야 정보검색, 소셜 네트워크, 네트워크, 네트워크 보안

이상훈 (E-mail: hoony@kndu.ac.kr)

- 1978 성균관대학교 전자공학과 졸업(학사)
- 1989 연세대학교 산업대학원 전산학과 졸업(석사)
- 1997 일본 교토대학교 정보공학 졸업(박사)
- 서일대학 겸임교수
- 충남산업대학교 교수
- 현재 국방대학교 국방관리대학원 전산정보학과 교수
- 현재 국방과학부 학부장
- 관심분야 정보검색, 데이터베이스, HCI, 멀티미디어