

A Preliminary Study on Clinical Decision Support System based on Classification Learning of Electronic Medical Records

Yang Kyu Shin¹⁾

Abstract

We employed a hierarchical document classification method to classify a massive collection of electronic medical records(EMR) written in both Korean and English. Our experimental system has been learned from 5,000 records of EMR text data and predicted a newly given set of EMR text data over 68% correctly. We expect the accuracy rate can be improved greatly provided a dictionary of medical terms or a suitable medical thesaurus. The classification system might play a key role in some clinical decision support systems and various interpretation systems for clinical data.

Keywords: *decision support, electronic medical records, classification learning, text mining*

1. 서론

최근 3차 의료기관에서 기존 종이 차트를 대신해 모든 의무기록을 전자문서화 한 전자의무기록(EMR, Electronic Medical Records)이 도입되기 시작함에 따라 의학통계 분야도 새로운 연구 대상이 생겨나고 있다. 국내에서는 대구의 D-의료원과 서울의 몇몇 3차 의료기관이 수년 전부터 이러한 환경을 도입해 사용하여 이미 상당한 자료를 보유하고 있다. 또한 문자 위주의 의무기록뿐만 아니라 X-선이나 단층촬영장치에서 생성된 사진과 동영상 자료까지 함께 저장되어 있지만, 축적된 데이터의 처리를 통해 의미 있는 정보와 지식을 추출하는 마이닝에 관한 연구는 아직 미흡한 실정이다.

축적된 임상 데이터나 연구관련 데이터의 가공은 진료정보의 재활용과 정보공유를

1) 경상북도 경산시 유곡동 290번지 대구한의대학교 정보과학부 교수
E-mail : yks@dhu.ac.kr

가능하게 할 뿐만 아니라 새로운 지식의 발견과 의사결정지원시스템 개발 등을 통해 의료서비스의 질적 개선을 이룩할 수 있는 때문이다[6]. 의사결정지원시스템 중 의사에게 가장 중요한 임상 의사결정지원시스템은 의사들이 환자들에 대한 결정을 내릴 때 결정적이고 중요한 정보를 제공하는 프로그램을 가리킨다. 어떤 의미로는, 의료 데이터를 처리하는 프로그램이나 의료 지식을 가진 소프트웨어는 모두 의사결정지원시스템이라 볼 수 있다. 하지만 실제 의미 있는 의사결정지원이 되기 위해서는 크게 의료정보관리, 주의집중, 그리고 특정 환자의 정보제공 등 세 가지 범주에 속하는 것 이어야 한다[8,9].

본 논문에서는 특정 환자에 대한 정보를 제공할 수 있는 의사결정지원 시스템 구축을 목표로 현재까지 연구·개발된 내용을 소개한다. 환자에 따른 임상 지식을 제공하기 위해서는 우선 모집단에 해당하는 일반 환자의 진료기록으로부터 데이터 마이닝 과정을 거쳐야 한다. 하지만 전자의무기록의 특성상 중요한 정보는 전자의무기록 중 의사가 자유로운 스타일로 입력해 놓은 진찰 소견에 숨어 있어서 이 부분의 텍스트 데이터 마이닝 기술이 필수적으로 요구된다. 또한 전자의무기록의 처리를 위해서는 분야별 전문 용어의 통일과 약어 등의 해결을 위해 의학 용어 사전이 필요하다. 하지만 현재까지 국내외에서 전자의무기록을 대상으로 임상정보나 지식의 추출을 연구한 결과는 찾아보기 힘들다.

베이지안 분류 학습을 이용하여 일반적인 텍스트의 자동 분류를 시도한 연구가 있었으며[1], 베르누이 정리와 베이즈 정리를 응용하여 문서들의 계층적인 정보를 추출할 수 있는 TAPPER 알고리즘[4]은 웹 문서를 분류하기 위해 개발되었다. 통계적 학습 기법과 로지스틱 회귀 분석을 이용해 MEDLINE 데이터로부터 의학 용어와 약어 사전의 자동 생성을 연구한 결과가 있으며[5], 정보의 필터링과 검색 기법을 이용하여 특정 환자의 증상과 유사한 진료기록을 찾아주는 연구가 있었다[7].

본 논문에서는 전자의무기록의 분류학습을 통해 질병 분야별로 증상에 관해 의미 있는 단어를 추출하여 통계량을 계산한 다음, 새로운 환자의 증상이 주어질 때 가능성이 높은 질병을 확률을 기준으로 순위별로 나열하게 한다. 또한 질병이 계층적인 구조를 가지기 때문에 질병의 판별에 필요한 증상 단어의 수를 최소화할 수 있는 방법도 함께 제시한다. 즉, 텍스트 마이닝 기술을 이용하여 전자의무기록 데이터에서 새로운 지식의 발견 및 재활용 방법을 제시하고 이에 대한 결과로서 질병별 특징 단어의 추출과 전자의무기록 분류를 통한 의사결정 지원 시스템의 실험적 모델을 제시하는데 목적이 있다.

2. 연구방법

2.1 단어 추출과 의학용어 사전

전자의무기록을 처리하기 위해서는 먼저 의미를 가진 단어를 추출해야 한다. 의무기록에 포함된 단어들 중 의미를 가진 단어들은 대부분 의학 전문 용어와 숫자 및 약어들로 구성된다. 따라서 기본적으로 추출된 명사 중심의 단어들에 대해 의학용어의

관점에서 처리가 이루어져야하며 이를 위해 의학 용어 및 동의어 사전이 필요하다. 의학 분야의 용어 서비스를 위해 미국에서는 UMLS(Unified Medical Language System)라는 통합적 시스템이 개발되어 사용되고 있지만[3], 국내에서는 이에 대한 개발이나 서비스가 거의 이루어지지 않고 있다.

따라서 본 논문에서는 우선 영어 명사만을 추출하여 처리하였으므로 한글 표현, 복합 명사의 처리, 동의어 및 약어의 처리가 이루어지지 않았다는 한계점을 미리 밝혀둔다. 높은 효과성을 가지는 지식의 생성과 더불어 의료정보의 원활한 공유를 위해서는 정확한 의학 용어의 사용이 필수적인 요건이라 볼 수 있으며 이에 대한 연구가 필요하다.

2.2 의사결정 지원 시스템

전자의무기록의 텍스트 문서 처리의 주된 접근 방법은 MedLEE(Medical Language Extraction and Encoding System)[2]와 같이 자연어 처리 기술을 위주로 발전하여 왔다. 하지만 최근 기계학습 알고리즘을 이용하여 문서 데이터 마이닝에 관한 다양한 연구가 활발해지면서, 본 논문에서도 이와 같은 방법을 응용하였다. 특히 여기서는 전자의무기록의 마이닝 결과를 이용하여 궁극적으로는 실제 사용 가능한 의사결정지원 시스템의 구축을 목표로 연구하기 때문에 전자의무기록의 텍스트 데이터를 “진료과”와 “질병명”의 계층관계로 표현하고, 텍스트 문서 학습을 바탕으로 이를 추론하는 시스템에 대해 연구하였다. 즉, 진료기록의 요약 부분을 학습하여 “진료과”와 “질병명”의 계층구조를 형성하는 중요 단어 목록을 생성하므로써 새로운 환자의 임상 정보가 주어질 때 그러한 단어 목록의 처리만으로 질병의 판별에 관한 의사결정 정보를 의사에게 제공할 수 있다.

2.3 전자의무기록 분류

본 연구에서 사용한 데이터는 1998년부터 도입되어 사용하고 있는 D-의료원의 전자의무기록 데이터를 기계학습을 이용하여 처리했다. 학습에 이용된 데이터는 “진단과”와 “질병코드”의 계층적 카테고리 형태로 정리하여 처리하였으며, 결과물로 생성된 특성단어 집합을 질병별 용어사전 대신에 사용하였다.

전자의무기록 데이터를 바탕으로 텍스트 마이닝 기법을 적용하여 생성된 질병별 용어사전을 바탕으로 증상에 대한 전자의무기록 중 요약(summary) 필드의 내용을 사용하였다. 여기서는 추론 알고리즘을 이용하여 환자의 요약이 주어질 때 가장 근접한 대표 질병 명을 추천해주는 의사결정 지원 시스템을 개발하였다. 이 시스템은 그림 1과 같이 왼쪽의 텍스트 박스에 시험 데이터를 넣고 결과를 질의하면 “진료과”와 “질병명”을 알려주는 인터페이스를 가진다. 즉, 그림 1의 왼쪽 텍스트 박스에 임의의 환자에 대한 진료 기록 요약을 자연어 형태로 입력하면 그 환자에 대해 가장 확률이 높은 진료과와 질병 코드를 생성해 준다.

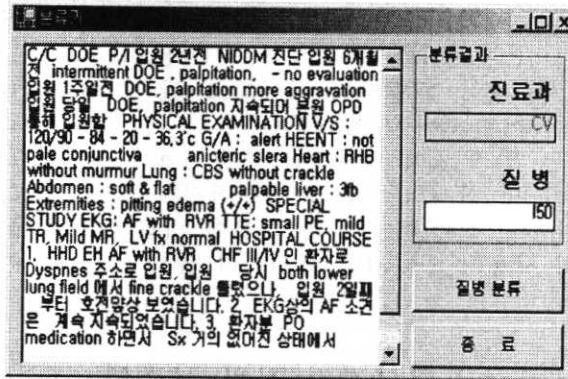


그림 1. 분류 테스트 애플리케이션

2.4 적용 알고리즘

본 연구에서 사용된 알고리즘은 계층적 카테고리 학습을 위해 IBM에서 연구된 TAPER(Taxonomy And Path Enhanced Retrieval system) 알고리즘[4]을 이용하였으며, “진료과” 별로 “질병”에 대해 학습시키는 방법을 적용하였다. TAPER 알고리즘은 베이저안 텍스트 분류를 개선한 알고리즘으로 계층적 구조를 가지는 문서의 분류에 대해 적용할 수 있는 기법이다. TAPER 분류 엔진의 모델은 그림 2와 같으며, 문서 분류를 위한 학습은 위해 베루누이 정리와 베이즈 정리를 응용한 식을 통해 이루어진다.

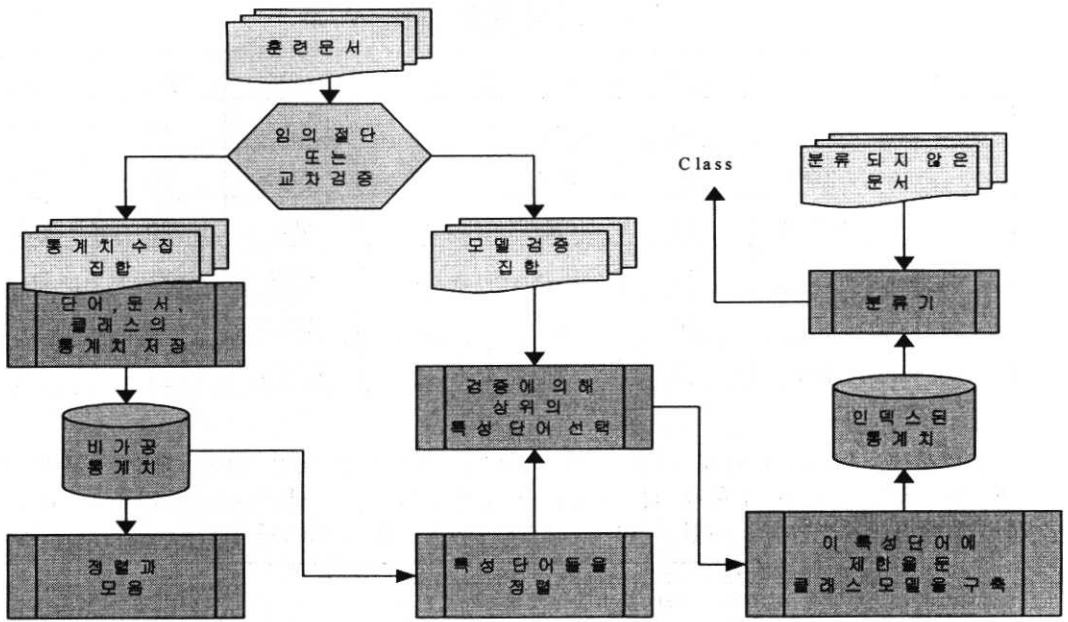


그림 2. 의사결정 지원시스템에 사용된 분류 엔진

전자의무기록 처리 과정에서 “Drug”이나 “Medical” 등과 같이 자주 사용은 되지만 문서의 소속을 결정하는데 큰 역할을 하지 못하는 단어와 “a”나 “and” 등과 같이 사용 빈도는 높지만 문서의 의미 결정에 도움을 주지 못하는 불용어들을 미리 제거하였다. 다음 단계의 학습과정에서 질병별 특성단어들의 리스트를 추출하였다.

3. 실험결과 및 분석

실험은 실제 전자의무기록 데이터를 사용하여 질병별 용어사전 추출과 의사결정 지원 시스템을 통한 질병 질의에 대한 테스트를 중심으로 이루어졌다. 실험은 5개의 진료과에 24개 질병에 대해서 약 5천건의 데이터를 이용하여 학습을 수행하였다. 실험 데이터의 70%는 훈련에 이용되었고, 나머지 30%는 테스트에 이용되었다. 표 1은 실험에 사용된 질병 카테고리를 정리한 표이며, 질병명은 코드로 표현되어 있다.

표 1. 질병별 분류 카테고리 표

번호	코드	번호	코드	번호	코드	번호	코드
0	Z34	12	O42.9	6	C15	18	O60
1	Z92.4	13	K25	7	C56	19	I21
2	N73.6	14	O42	8	I50	20	O80
3	I11	15	I61.6	9	C53	21	D25
4	K74	16	D25.1	10	S02.7	22	I20
5	I10	17	O82	11	S22.4	23	C22

용어사전 구성은 학습에 사용된 특성단어를 추출하는 것을 기본 단계로 한다. 질병별 문서에서 단어를 추출해 각 소속별 사용빈도가 높은 단어순으로 리스트 구성하게 된다. 그림 3은 질병에 대한 단어를 추출하여 나열한 그림이다.

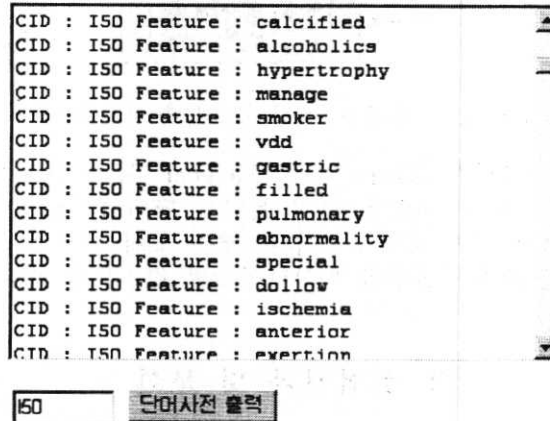


그림 3. 추출된 특성단어

분류의 평균 정확율은 약 68.17% 이었다. 그림 4에서처럼 분류 예측을 실행할 때 유사 질병들로 데이터가 편향되는 경향이 발생하는 것을 볼 수 있는데, 질병명이 유사할 경우 해당 의무기록 데이터들이 비슷한 단어들로 이루어지기 때문이다. 이와 같은 편향성 문제는 용어사전에 대한 수정으로 정확율을 높이거나 카테고리별 모집단 단어와 하부집단 단어의 처리 방법을 수정하여 해결할 수 있을 것으로 기대 한다. 또 다른 해결책은 좀 더 정제된 데이터로 충분히 학습을 실행하여 각 질병별 특성단어의 정확한 목록을 구하는 것이라 볼 수 있다.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
Z34	4												1		3	1		27	6		3				9.09%
Z92.4	5							16	1	5	1	1			1		9	13	2		5	8			0.00%
N73.6			4														14					5			17.39%
I11				11		1										11									91.67%
K74			1	10		8	4		1			1				3								2	37.04%
I10					4											12									100.00%
C15				2	2	804			1	1				2		54								3	98.21%
C56			1					12		5		1					3								54.55%
I50				16					2							18									11.11%
C53	1							10		8							6						1		30.77%
S02.7											2	54				36									3.51%
S22.4											2	54				35									96.43%
O42.9													4						1		27				12.50%
K25				3	5	3	1		2					2		1				1			1	2	10.00%
O42													1		5			16	4		15				12.20%
I61.6													14			9									6.67%
D25.1																20						2			90.91%
O82																1		30							100.00%
O60	1											4		2			7	3		8					12.00%
I21																35					9				100.00%
O80													3		1							44			91.67%
D25			2														5						15		65.22%
I20				1		2				1						41					2			31	83.78%
C22					2	13	5		1							4								10	32.26%
Correct : 889 Out of 1804																									
Accuracy average : 68.1748466257668%																									

그림 4. 질병별 분류 결과

4. 결 론

본 연구에서는 전자의무기록 중 텍스트 데이터의 마이닝과 재활용을 위한 기초연구의 일환으로 질병별 특징단어 추출과 전자의무기록 분류를 통한 의사결정 지원 시스템의 기본 모델을 시험 구축하였다. 현재 D-의료원의 경우 의사마다 한글 혹은 영어, 일반적 서술어 혹은 전문 용어 등 표현하는 용어에 상당한 차이가 있으며, 이로 인해 기계적 처리의 어려움과 자료공유의 문제점이 발생한다. 하지만 이와 같은 어려움은 텍스트 형태의 의료데이터 가공 및 공유 전반에 걸쳐 발생할 수 있는 문제점이라 본다. 본 연구에서는 우선 기계적으로 추출된 특징단어 모음 형태의 용어사전을 임상 의사와의 피드백을 통해 표준 용어사전에 가까운 형태로 발전시킬 수 있다는 가능성을 확인하였다. 또한 텍스트 마이닝을 통해 구축한 의무기록 분류기는 지능적인 의사결정 지원 시스템의 구축 가능성을 보여 주었다.

본 논문에서 현재까지 구축한 시스템은 용어사전 대신 확률적 방법을 이용해 특성 단어 목록을 구성하였고 이를 이용하여 질병에 대한 분류 및 예측을 하고 있다. 또한 영문 용어만을 처리하였고, 복합명사 처리를 포함하지 않았으며, 명사 외의 다른 의미 있는 단어들을 처리하지 않았다. 이와 같은 문제점들은 지속적으로 해결을 시도할 것이다.

나아가 전자의무기록 학습을 통한 질병별 증상 분류 및 관련단어 추출은 의료 전산화의 여러 분야에서 사용될 수 있는 실용적인 표준 용어사전을 만들 수 있고, 우리나라 환경에 맞는 용어사전의 모델을 제안할 수 있을 것이다. 또한 질병별 의학용어 사전이 정의되고 그에 맞게 진료기록에 증상이 기록된다면 이들 증상만으로 질병들에 대한 정확한 분류가 가능하게 되며, 이를 통해서 진단분야에 효율적인 의사결정 지원 시스템을 구축할 수 있을 것이다.

참고문헌

1. 김진상, 신양규(2000), 베이지안 학습을 이용한 문서의 자동분류, 데이터정보과학회지 11(1): 9-30.
2. __ (2003), MedLEE - A Medical Language Extraction and Encoding System, <http://cat.cpmc.columbia.edu/medleexml>.
3. __ (2003), Unified Medical Language System (UMLS), <http://umlsinfo.nlm.nih.gov/>
4. Chakrabarti, S., Dom, B., Agrawal, R., Raghavan, P.(1998), Scalable Feature Selection, Classification and Signature Generation for Organizing Large Text Databases into Hierarchical Topic Taxonomies, VLDB Journal 7(3): 163-178.
5. Chang, J. T., Schutze, H., Altman, R. B. (2002), Creating an Online Dictionary of Abbreviations from MEDLINE, JAMIA 7(3): 326 - 332.
6. Hripcsak G., Friedman C., Alderson P. O.(1995), Unlocking clinical data from narrative reports, Annals of Internal Medicine 122: 681-688.
7. Hsu, W. H. (2003), Activities of the KSU Bioinformatics and Medical Informatics (BMI) Working Group 2001-2002, Kansas State University Technical Report.
8. Musen, M. A., Shahar, Y., Shortliffe, E. H.(1997), Clinical Decision Support Systems, Chapter 16 in [9], 600-639.
9. van Bemmelen, J. H., Musen, M. A.(1997), Handbook of Medical Informatics, Springer Verlag.

[2003년 8월 접수, 2003년 10월 채택]