

생체 미네랄정보의 다중 퍼지규칙베이스 구축에 의한 지능적 의학진단시스템 구축

An Intelligent Medical Diagnosis System by Multiple Fuzzy Rule Base of Biological Mineral Information Analysis

조영임

수원대학교 IT대학 컴퓨터학과
E-mail: ycho@suwon.ac.kr

요 약

본 논문에서는 모발내에 있는 약 30여가지의 생체 미네랄과 8가지의 중금속 정보 분석을 통해 생체내에 영양상태의 과잉, 결핍 및 불균형 상태를 평가하고, 그 결과가 현재 생체에 미치는 영향을 예측하여, 건강을 유지하는 방향을 제시할 수 있는 의료용 지능적 의학진단 시스템을 구축하였다. 이 논문에서는 생체내 미네랄 정보를 다중 퍼지규칙베이스 시스템으로 구축함으로써 환자에게 보다 효율적으로 치료와 예방방법을 제시할 수 있는 의학진단시스템을 구축하였다.

Key Words : Mineral Information Analysis, Multistep Inference, Bioinformatics

1. 서 론

생물학적으로 우리의 머리카락 속에 있는 30여 가지의 미네랄과 8가지의 중금속이 존재한다. 이때 이들의 양과 중요 미네랄 비율을 분석하여 체내에 과잉, 결핍 및 불균형 상태를 평가하고 대사타입(metabolic type)을 제시하여 건강을 유지하는 방향을 제시하는 방법을 임상 영양학 및 독성학 모발조직 검사방법이라고 한다[1]. 또한 호르몬의 세포수준 활성을 제시(갑상선 및 부신)하며 영양, 신경 및 호르몬의 불균형을 통해 질환발생 가능성을 설명한다.

일반적으로 모발조직 검사방법은 조직 미네랄 검사와 모발 미네랄 검사의 두 가지로 분류된다. 첫째, 조직 미네랄 검사(TMA: Tissue Mineral Analysis)는 조직(머리카락, 일반 조직) 중 미네랄의 결핍, 과잉 및 불균형을 평가하는 방법을 말하며 머리카락만을 갖고 하는 것보다 좀더 포괄적인 의미를 갖는다. 둘째, 머리카락 미네랄 검사(HTMA: Hair Mineral Analysis)는 비용 대비 효율이 가장 높고 정확한 TMA가 HTMA인데, 일반적으로 TMA는 머리카락을 이용한 조직 미네랄 검사[2,3]를 많이 하고 있으므로 HTMA를 TMA라고 하기도 하며 모발분석의 일반적 방법이 된다.

모발분석의 특징 및 장단점은 다음과 같다. 첫째, 모발 표본은 비침해 방법으로 채취되며

로 혈액, 소변, 기타 다른 조직보다 빠르고 쉽게 수집할 수 있다. 둘째, 모발분석방법은 다른 방법의 미네랄 분석보다 비용 대비 효율이 우수하다. 셋째, 모발은 혈액과는 달리 미량원소의 농도에 민감한 영향을 미치는 항상성 기전에 덜 민감하다. 넷째, 미네랄의 저류 또는 손실과 같은 장기간의 편차들은 혈액보다 모발조직에서 쉽게 검출된다. 다섯째, 대부분의 모발조직 미네랄 농도는 혈액이나 다른 조직에서 발견되는 미네랄보다 농도가 매우 높다. 여섯째, 모발조직은 과거뿐만 아니라 현재의 미량원소의 수치기록을 제공한다. 일곱째, 혈청뿐 아니라 외부원인으로부터 모발로 들어가는 물질의 정보를 제공한다. 여덟째, 독성 중금속 수치를 평가하는데 유용하다[4].

그러나 현재 국내 모발분석방법은 분석결과를 해석할 수 있는 의학 전문 프로그램이 개발되어 있지 않고, 일부 국내에서 기구축한 모발분석 관련 데이터베이스가 매우 낮은 수준이어서 의료기관에서 충분한 의료서비스를 제공하기 어렵다. 그러므로 현재 국내에서는 대부분의 모발분석을 미국에 의뢰하게 되는데 이 결과가 서구식 생활패턴에서 비롯된 데이터베이스로부터 구축된 것이므로 한국인에 대해서는 보다 분석적인 한국형 데이터베이스 및 모발분석결과해석이 필요하다.

따라서 본 논문에서는 이러한 문제점을 해결

하기 위해 모발분석을 위한 알고리즘을 개발하고 모발분석결과에 대한 데이터베이스를 구축하여 지능적 의학 진단시스템을 구축하고자 한다. 이것을 위해 국내 모발관련 임상자료를 이용하였으며 통계분석적 오류와 데이터베이스 구축 및 추론상의 오류를 최소화하기 위해 다중 퍼지규칙 베이스를 구축하였고 다단계 통계 분류 방법을 이용하였다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안하는 지능적 추론방법, 다중 퍼지규칙베이스 구성 및 통계적 분류 방법을 설명하며, 3장에서는 제안하는 시스템의 개요와 주요 기능 설명 및 효율성 등을 분석하고, 4장에서 결론을 내리고자 한다.

2. 퍼지 모델

본 논문에서 모발분석을 위한 지능적 의학 전문가 시스템의 연구목표는 그림 1과 같이 모발분석자료의 통계적 분석 및 그룹화와 사용자 중심의 정보검색이 가능한 시스템을 구축한다.

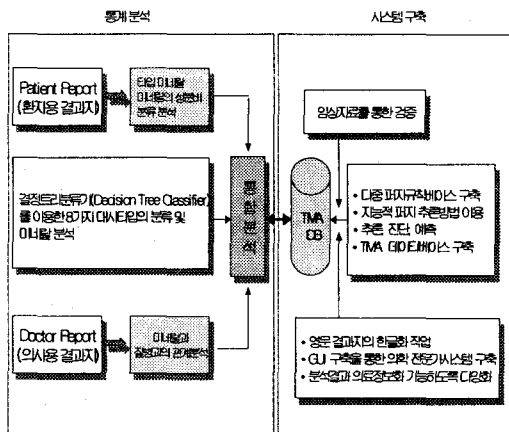


그림 1. 본 논문에서 제안한 시스템

지능적 추론의 배경이 되는 대사타입의 분류 및 미네랄의 분석은 통계적 분류 및 분석 방법 중에서 결정트리분류기(Decision Tree Classifier)를 이용하여 모발분석자료의 대분류에 해당하는 인체의 8가지 대사타입(빠른 대사 4타입, 느린 대사 4타입)을 분류하였다. 이 방법을 사용하는 이유는, 1단계 분류기는 클래스간의 변별력 중에서 평균이 최대가 되는 특징 조합을 이용하나, 이 특징조합이 모든 클래스에 대하여 최적의 변별력을 항상 갖는 것이 아니므로 두 종류이상의 결정을 적절한 수순에 따라 몇 번이고 반복하여 입력패턴의 클래스를 결정해야 하므로 본 논문에서는 다단계분류기인 결정트리 분류방법[5,6]을 응용하여 분류하였으며 결과는 다음 그림 2와 같다.

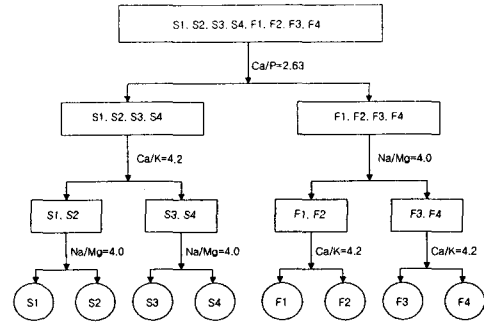


그림 2. 결정트리에 의한 모발분석결과 분류

분석된 통계를 바탕으로 본 논문에서는 불확실한 다량의 정보로부터 지능적 추론이 가능한 데이터베이스 구축을 위해 다중 퍼지규칙베이스 방식을 제안한다. 다중 퍼지규칙베이스는 모발분석 자료와 같이 다양한 미네랄에 의해 의사결정을 해야 하는 시스템인 경우 규칙베이스가 매우 복잡해 질 수 있기 때문에 규칙베이스를 분산시켜서 연관성을 찾아 실시간으로 규칙들을 검색하고 추론하는 방법이 필요하다. 이것을 위해 본 논문에서는 규칙들의 결론부와 다른 규칙의 조건부와의 관계를 트리형태로 구성하여 다단계로 구성함으로써 기존의 퍼지규칙베이스보다 복잡한 규칙[7]을 표현하기에 적합하도록 하는 방법을 제안한다.

본 논문에서는 다중 퍼지규칙베이스 구축을 위해 두 단계가 필요한데, 제 1 단계 구축 후 제 2 단계를 통해 다중 퍼지규칙베이스 구축이 완성된다. 제1단계는 대사타입 분류에 의한 퍼지 데이터베이스 구축하는 단계이며, 제2단계는 각 미네랄 양과 비율에 따른 해당 문장 데이터베이스 검색 및 출력하는 단계이다.

퍼지규칙베이스 구축을 위해 주요 미네랄 비율을 나타내는 언어적 변수인 "높다(high)" "낮다(low)" "적당(acceptable)" 등은 퍼지 집합[8]으로 표현하였는데, 이에 사용되는 수치들은 통계적 분석에 의해 획득하였다.

본 논문에서는 다단계 통계분석방법을 사용하여 분류한 자료로부터 구축한 다중 퍼지규칙베이스에서 퍼지 추론을 효율적으로 하기 위해 분산 목표트리 방식인 TAEMS(Task Analysis Environment Modeling and Simulation)[9]를 적용하였다. TAEMS는 태스크를 분류한 후 분산 목표트리를 구성하여, 서로 다른 목표를 갖는 시스템의 분산목표 수행을 위해 시스템의 목표나 자료들 간의 상호 의존성을 나타냄으로써 효율성을 높이기 위한 분산 인공지능적인 방법이다.

다음 그림 3은 다중 퍼지규칙베이스에서 규칙들간의 상호의존성을 TAEMS 방식을 적용해서 나타낸 그림이며, 본 논문에서 구현한 퍼

지규칙베이스의 구성형태와 FUZZY_INFER()의 추론알고리즘 (FUZZY_INFER_I, FUZZY_INFER_ID)은 다음과 같다. 여기서 Dataset(Input Facts)는 규칙의 조건부에 입력되는 미네랄 또는 미네랄 성분비 값들을 말한다.

<Fuzzy Rule Base 구성 형태>

Rule 1: IF x1 is A1 and y1 is B1 and ... z1 is C1 THEN k1 is R1
 Rule 2: IF x2 is A2 and y2 is B2 and ... z2 is C2 THEN k2 is R2
 Rule n: IF xn is An and yn is Bn and ... zn is Cn THEN kn is Rn

여기서 x1,y1,z1,k1 : 미네랄, 미네랄 비, 내분비선 등
 A1,B1,C1,R1: 퍼지 집합 (high,low,medium...)

<FUZZY_INFER() 알고리즘>

Enter of input facts to Rule i (초기입력);
 TAEMS에 의한 퍼지 규칙 연관관계 트리구조 생성;
 Rule i와 연관된 규칙들의 집합 생성; // 규칙들간 Complexity, Scheduling은 TAEMS 방식의 퍼지 추론 할 수 있는 입력형태로 변환;

If input fact == conditional part of Rule m(연관된 규칙)
 then FUZZY_INFER_I(RuleBase, input facts);
 // 입력이 규칙의 조건부와 모두 일치하고, 결론부가 단일일 경우
 즉, 규칙들간의 다중 상호 의존관계가 없는 경우 퍼지추론하는 과정

If input fact == some conditional part of Rule n(연관된 규칙)
 then FUZZY_INFER_ID(RuleBase, input facts);
 { for(i=0; i<=n; i++) // m, n은 TAEMS에 구축된 연관 규칙들의 갯수
 for(j=0; j<=m; j++)
 { x[i] of Rule[i] = k[j] of Rule[i];
 { y[i] of Rule[i] = k[k] of Rule[k];
 x[i] of Rule[i] = k[m] of Rule[m];
 }

// 입력이 규칙의 조건부의 일부와 유사하고, 다른 규칙의 결론부로부터 값을 유도해야 할 경우
 즉, 규칙들간의 다중 상호 의존관계를 갖는 경우 TAEMS 트리구조에서 생성된 연관 규칙들로 부터 값을 유도하는 과정

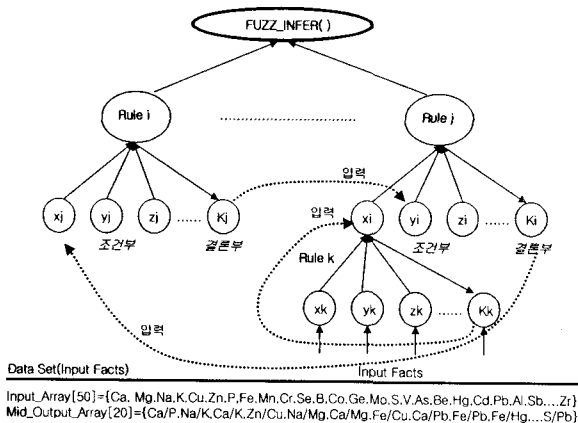


그림 3. 다중 퍼지규칙베이스에서 규칙들 간의 상호의존관계

퍼지규칙베이스에 있는 각 규칙들은 위의 그림과 같이 상호의존관계를 갖는다. 즉, 하나의 규칙을 구성하는 조건부, 결론부 퍼지명제들은 다른 규칙의 조건부에 값을 전달하는 상호의존관계로 구축되어 있다. 이렇게 하는 이유는 모발분석 미네랄 자료가 단순히 몇 개의 미네랄에 의해 분류되지 않고 매우 복잡한 연관관계를 통해 분석되고 측정되기 때문이다.

위 구조에 따라 다중 퍼지규칙베이스에서 10개의 상위레벨로부터 파생된 인체의 대사타입에 따라 분류·구축된 약 1000개의 규칙들의 각 조건부들을 찾아가는 과정을 예를 들어 나

타내면 다음 그림 4와 같다. 만약 ID가 '13002368'인 영어 문장은 출력해야 할 대분류 'Nutrient Mineral Level' 중 두 번째 미네랄인 'Mg'에 관한 368번째 문장을 말한다. 이 ID인 문장이 추론되기 위해서는 위의 그림에서와 같은 3개의 퍼지 규칙이 연관되어 만족되어야 한다. 즉, 'THEN'이하의 결론부인 'S1', 'Ca/K is low', 'Na/Mg is high'가 각각 Rule #2의 조건부에 입력되어 TAEMS방식의 다단계 스케줄링 과정을 거쳐서 해당 ID 문장을 추론하게 된다. 각 규칙의 결론부가 상위레벨의 규칙의 각 조건부 문장들과 일치되는 것들을 찾아가면서 최종적으로 결론부의 문장 ID를 찾게 된다.

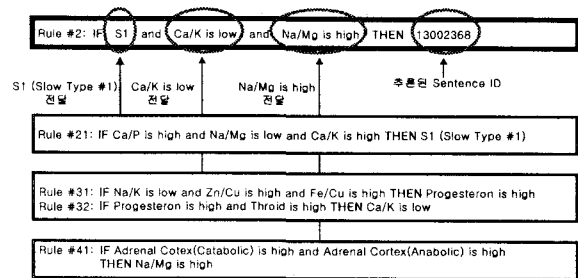


그림 4. 다중 퍼지규칙베이스에서 추론 과정

이로부터 추론된 해당 ID인 문장은 다음과 같다.

[ID: 13002368인 영어문장]

"Magnesium is essential for muscle relaxation, protein synthesis, nerve excitability and energy production on a cellular level....(중략)... When magnesium levels are chronically low in the body, a tendency toward one or more of the following conditions will increase: Noise Sensitivity, Tremors, Body Odor, Muscle Spasms, Irritability, Increased Blood Pressure...."

[ID: 13002368 번역된 한국어 문장]

"마그네슘은 세포레벨에서 근육이완, 단백질 합성, 신경흥분, 에너지생산에 필수요소이다....(중략)... 만약 마그네슘 정도가 인체 내에서 만성적으로 낮으면 다음과 같은 하나이상의 상태가 증가하게 된다. 소음 민감성, 흥분, 압내, 근육경련, 불안정, 혈압상승...."

따라서 환자의 다양한 미네랄 수치와 미네랄들간의 비에 따라 지능적으로 추론된 각 ID에 따른 영어문장들과 해당 문장들이 번역된 한국어 문장들이 데이터베이스로부터 추론되어 레포트형식을 갖추어서 결과지로서 환자와 의사에게 제공된다.

3. 성능 평가

이 시스템의 평균 처리속도는 입력창에서 미네랄 수치를 입력하여 영어 및 번역된 문장을 찾아내는 속도를 측정된 것으로, 기존의 방법은 규칙의 입력부분의 비교횟수가 많아 처리속도가 느린 반면 제안한 방법은 규칙의 조건부 비교횟수를 줄여주어 결론부를 재입력함으로써 평균 처리속도가 약 5배 정도 향상되었다.

표 1. 지능적 의학진단시스템의 효율성 비교

비교항목	다중퍼지규칙 베이스(제안)	단일퍼지규칙 베이스(기존)
평균 처리속도(sec)	0.1sec	0.5sec
결과의 정확성(%)	92%	68%

결과의 정확성은 환자의 임상자료 약 2000건에 대해 측정된 것으로, 시스템에서 전체적으로 한국어로 번역되어 출력되는 최종 결과지가 미국에서 보내오는 영문으로 된 결과지와 일치하는 정도를 측정된 것이다. 환자 한건당 결과지에 포함해서 출력해야 할 총 50개의 문장들 중 단일 퍼지규칙 베이스로 구축한 경우는 34개 정도의 문장만이 일치하였으나 다중 퍼지규칙 베이스로 시스템을 구축한 경우 평균적으로 46개 정도 일치하였고 4개 정도 비슷하거나 잘못된 문장을 출력하였는데, 이것은 제안한 방법이 기존에 비해 약 24% 정도 정확성이 향상된 수치이다.

분석을 위해 이 프로그램을 사용한 경우와 아닌 경우를 실제 실험한 결과, 이 시스템을 사용한 경우 환자 1건당 약 7분 정도가 소요되어 이 시스템을 사용하지 않은 경우보다 86% 일의능률을 가져올 수 있었다.

5. 결론

본 논문은 모발분석을 위한 한국적 데이터베이스 구축 및 지능적 의학 진단 시스템을 연구 개발 하였다.

본 논문에서는 이를 위해 첫째, 통계적 모발 분석을 위해 다단계 분류방법을 활용하여 다중으로 영향을 주는 요인들에 대하여 분류하였다. 둘째, 다중 퍼지규칙베이스를 구축하여 지능적인 퍼지추론 방법에 의해 미네랄들을 찾아내고 관련성을 연구하였다. 셋째, 이러한 방법을 이용하여 지능적 의학진단 시스템을 구축하였다.

현재 개발된 시스템은 국내병원에서 사용 중에 있으며 의학진단시 업무능률이 86%이상 향

상되었고, 사용자 만족도가 92%정도 되었고, 시간당 처리하는 건수도 기존에 비해 약 7.5배 이상 증가하게 되었으며 결과의 정확도도 약 92%정도로 측정되었고, 시스템의 평균 처리속도도 약 5배이상 증가하였다.

앞으로 좀더 많은 임상결과로부터 시스템을 개선해야 할 것이다.

참 고 문 헌

[1] Watts, D.L., Trace Elements and Other Essential Nutrients-Clinical Application of Tissue Mineral Analysis, 2nd Writer's B-L-O-C-K edition, 1999.

[2] Strain, W. H. and Pories, W. J. and Flynn, A. and Hill, O. A., Trace Element Nutriture and Metabolism Through Head Hair Analysis. Trace Substances in Environmental Health. Hemphill, D., Ed. Univ. Mo. Press, Columbia. 1972.

[3] Baumgartner, W. A. and Hill, V.A. and Bland, W. H., "Hair Analysis for Drugs of Abuse". Journal of Forensic Science Vol 34. No 6. 1989.

[4] Vir etal, "Serum and hair concentrations of Copper During Pregnancy", American Journal of Clinical Nutrition Vol. 34, 1991.

[5] Safavian, S.R. and Landgrebe, D., "A survey of decision tree classifier methodology", IEEE Transactions on System, Man, and Cybernetics., Vol.21, No.3, pp.660-674, 1995.

[6] Brown, D.E. and Corruble, V. and Pittard, C. L., "A Comparison of decision tree classifiers with backpropagation neural networks for multinomial classification problems", Pattern Recognition, Vol. 26, No. 6, 953-961, 1993.

[7] Turksen, I.B. and Zhong, Z., "An Approximate Analogical Reasoning Schema based on Similarity Measures and Interval Valued Fussy Sets," Fuzzy Sets and Systems, North-Holland, Vol. 34, pp.323-346, 1990.

[8] Lee, C.C., "Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part I", IEEE Transactions on System, Man, and Cybernetics., Vol.20, No.2, pp.404-418, 1990.

[9] Graham, J. R. and Decker, K.S., "Towards a Distributed Environment-Centered Agent Framework", Umass Technical Report, 1999.