

퍼지 ART 알고리즘을 이용한 한방 자가 진단 및 학습 시스템

황병주 · 김광백

신라대학교 컴퓨터정보공학부

Self Health Diagnosis and Learning System of Oriental Medicine

Using Fuzzy ART Algorithm

Byong-ju Hwang · Kwang-baek Kim

Division of Computer and Information Engineering, Silla University

E-mail : vosnos@hanmail.net, gbkim@silla.ac.kr

요 약

본 논문에서는 질병에 대한 전문적인 지식이 부족한 일반인들을 대상으로 스스로 자신의 건강 상태를 쉽게 파악하고, 조금씩 진화하는 질병 바이러스에 따른 증상의 변화를 진단할 수 있는 퍼지 ART 알고리즘을 이용한 한방 자가 진단 및 학습 시스템을 제안한다. 제안된 한방 자가 진단 및 학습 시스템은 72가지 한방 질병과 각 질병에 대한 증상을 분석하여 데이터베이스로 구축하고 구축된 데이터베이스 정보를 기반으로 퍼지 ART 알고리즘을 적용하여 사용자의 질병을 도출한다. 본 논문에서는 사용자가 자신의 대표 증상을 제시하면 해당 증상을 포함하는 질병들을 도출한다. 도출된 질병들의 세부 증상들을 사용자가 입력 벡터로 제시하면 퍼지 ART 알고리즘을 적용하여 세부 증상에 대한 질병들을 클러스터링한 후, 세부 증상에 대한 질병의 소속 정도를 제공한다. 본 논문에서 제시한 시스템을 한의학 전문의가 분석한 결과, 본 논문에서 제시한 시스템이 한방 질병의 보조 진단으로서의 가능성을 확인하였다.

키워드

질병, 증상, 퍼지 ART 알고리즘, 한방 자가 진단 및 학습 시스템

1. 서 론

인간의 신체는 바이러스의 침투나, 갑작스런 환경 등의 변화로 인해 질병이 발병하게 되면 그에 따라 신체에 부정적인 변화가 발견되게 된다.

이런 부정적인 변화를 '증상'이라 하는데 초기에는 미비한 증상에서부터 발견되다가 점차 증상의 심각도가 커지게 된다. 이 때 적절한 조치 없이 질병을 방치하게 되면 질병의 악화로 인해 결국 정상적인 생활을 영위하지 못하게 되므로 미비한 증상이 감지되었을 때 우리는 신속히 치료하여야만 하는데, 현대인의 바쁜 현실 속에서 병원을 직접 방문할 여유가 부족한 사람이나, 도서 산간 지역과 같이 병원을 찾기 어려운 제한적 지역에 거주하는 사람들은 치료가 늦어져 결국 큰 고통과 손실을 경험하게 된다[1].

따라서 본 논문에서는 이러한 고통과 손실을 막기 위하여 우리 몸에 증상이 발견되었을 때, 발견된 증상들로 하여금 집에서 간단히 해당 질병을 조기에 추론하여 진단할 수 있는 한방 자가 진단 및 학습 시스템을 제안한다.

한의학은 예로부터 한국인의 몸을 진단하고 치

유하던 의학으로서 서양 의학과는 다르게 우리의 몸에 맞는 한의학적 데이터가 수백 년에 걸쳐 쌓여 왔기 때문에 한국인 고유의 신체적 특성에 맞게 병을 진단하고 처방할 수 있는 과학적인 의학이다[2]. 한의학의 기준에 의해 질병을 분류함으로써 서양 의학에 의해 질병을 분류하여 진단하는 것보다 한국인의 신체에서 발병하는 질병에 대한 진단의 정확도를 높일 수 있다.

제안한 한방 자가 진단 및 학습 시스템은 병원에 방문하여 진료 받을 때까지 소요될 시간비용과 진료비용을 줄일 수 있으며, 발생 가능한 질병에 대한 지식의 부재로 인해 생기게 되는 불안감을 해소 시킬 수 있고, 질병에 대한 정확한 정보를 조기에 파악하여 치료 받음으로서 더 큰 질병으로의 악화를 막을 수 있다.

제안한 한방 자가 진단 및 학습 시스템은 한의학에 의한 질병 분류에 따라 데이터베이스를 구축하고, 퍼지 ART 알고리즘을 적용하여 질병 및 증상의 갱신을 신속하고 효율적으로 이루어지게 하였으며, 바이러스와 환경의 변화에 따른 질병 및 증상의 변화를 시스템 스스로가 학습하여 판단하도록 하였다.

II. 제안한 한방 자가 진단 및 학습 시스템

2.1 질병과 증상의 조사 및 데이터베이스 구축

제안한 한방 자가 진단 및 학습 시스템은 한의학 관련, 신뢰성 높은 '동의보감'과 '한방의학백과' 서적을 근거로 하여 72가지 한방 질병들을 선정하고, 선정된 한방 질병들의 증상 중 뚜렷이 나타나는 167가지의 증상을 선정하여 데이터베이스로 구축하였다[2,3] 본 논문에서 제시한 질병과 증상의 수집 과정은 그림 1과 같다.

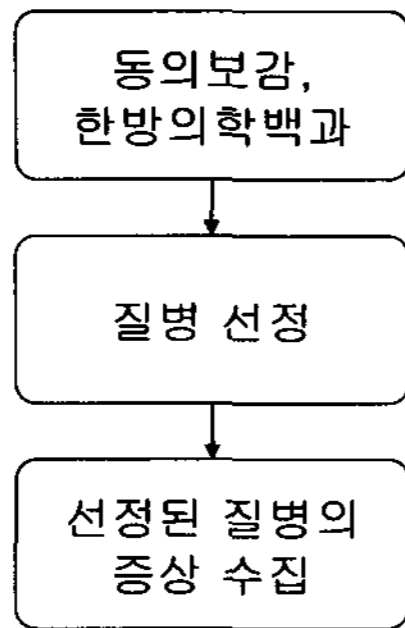


그림 1. 질병과 증상의 수집과정

본 논문에서 설계한 한방 질병과 증상의 데이터베이스 테이블 및 가중치 테이블의 구조는 표 1, 2, 3 과 같다.

표 1. 질병 테이블

질병ID	질병명	증상코드	필수증상 코드
1	감기	32-55-56-58-1-2-31	NULL
2	독감	32-55-56-58-3-2-31-5	NULL
3	늑막염	3-59-58-33-116-12-7	116
...
72	자궁암	27-167-148-37-33	NULL

표 2. 증상 테이블

증상ID	부위	증상
1	전신	발열
2	전신	오한
3	전신	식욕부진
...
167	기타	월경이 불규칙하다

표 3. 가중치 테이블

인덱스	질병ID	증상ID	가중치
1	1	1	1
2	1	2	1
3	1	3	0
...
72	1	167	0
73	2	1	0
...
12024	72	167	1

표 1의 질병 테이블의 증상코드는 표 2의 증상 테이블의 증상코드를 가리키는 것으로 해당 질병에 나타나는 증상을 의미하고 필수 증상 코드는 해당 질병의 증상 중 반드시 나타나야 하는 증상을 의미한다.

표 3의 가중치 테이블은 본 논문에서 개선한 퍼지 ART 알고리즘에서 학습에 의해 변화하는 가중치를 저장하고 있는 테이블로서 모든 질병에 대한 증상들의 카티션 곱만큼의 행을 가진다.

2.2 제안한 한방 자가 진단 및 학습 시스템

본 논문에서는 대표 증상을 선정하여 대표 증상과 관련된 질병을 도출하고 도출된 질병들에 대해 다시 세부 증상을 선정한다. 그리고 필수 증상과 수반 증상을 분리하여 진단의 정확도를 개선한다.

본 논문에서 제안한 한방 자가 진단 및 학습 시스템의 처리 과정은 그림 2와 같다. 본인의 신체에서 발생하는 증상 중 대표 증상을 선택하면, 대표 증상을 포함하는 질병들의 세부 증상이 도출되고, 도출된 세부 증상에서 사용자가 최종 증상을 선택한다. 이와 같이 대표 증상을 선택한 후, 최종 증상을 선택하게 함으로서 질병과 연관성이 없는 증상을 선택하는 것을 사전에 방지하여 질병 진단의 정확도를 높인다. 최종 선택된 증상을 개선된 퍼지 ART 알고리즘의 입력 벡터로 적용하여 질병들을 클러스터링하고 가장 소속도가 높은 한방 질병 3가지를 도출하여 오름차순으로 정렬한다.

본 논문에서 제안한 한방 질병 자가 진단 및 학습 방법은 필수 증상과 수반 증상을 구분하여 처리하므로 질병 진단의 정확도를 높일 수 있다.

질병과 증상의 관계를 분석하면 수반 증상들 로만 나타나는 질병과 수반 증상과 함께 필수 증상이 반드시 나타나는 질병으로 구분된다. 이런 필수 증상은 그림 3과 같이 해당 질병에 대한 필요조건임을 의미한다.

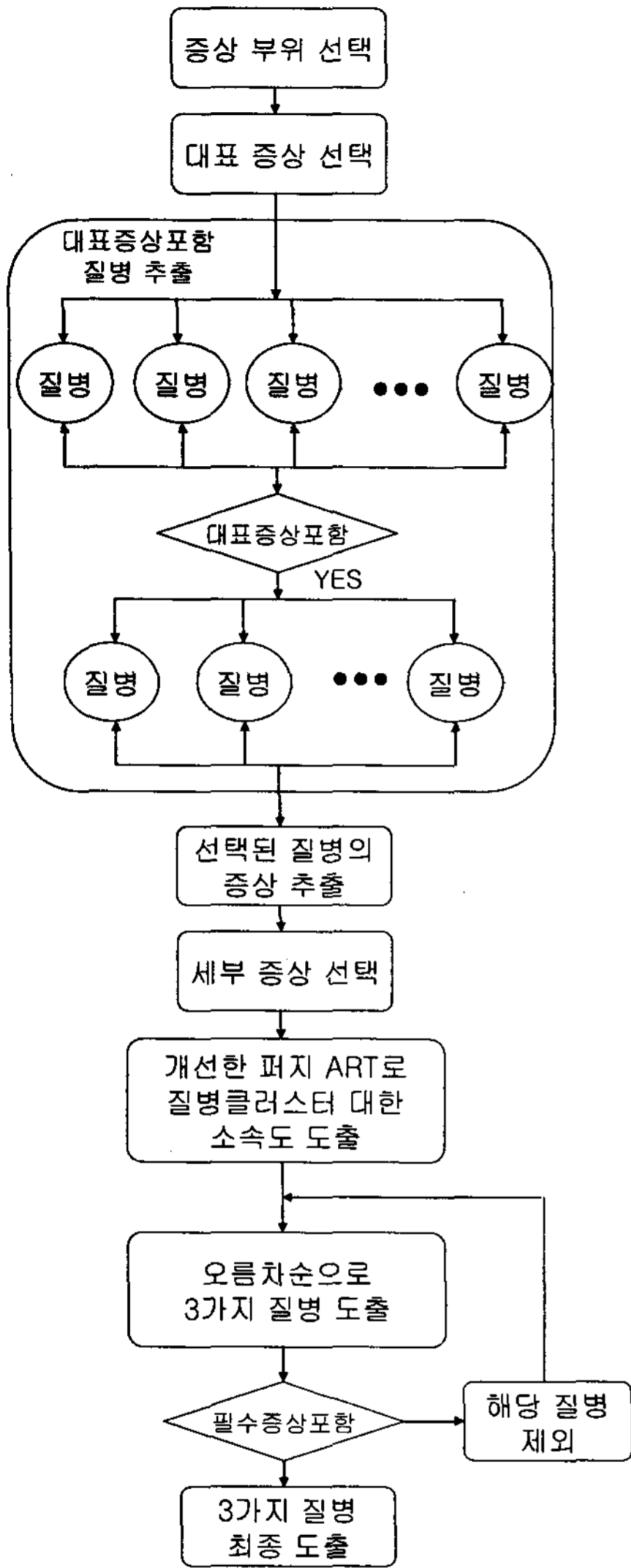


그림 2. 제안한 한방 자가 진단 및 학습 시스템

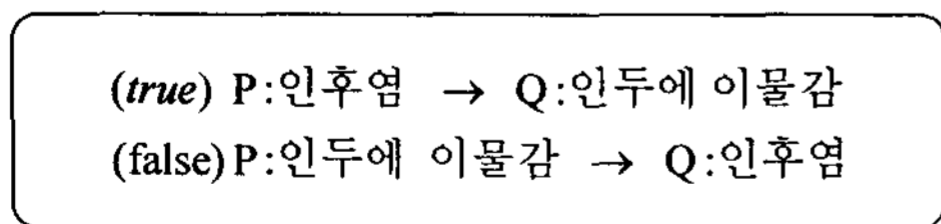


그림 3. 필수증상과 질병관계의 예

예를 들어 인후염 증상의 경우 감기의 증상과 거의 흡사하지만 '인두에 이물감'이라는 증상이 반드시 발생하여야 한다.

인두에 이물감이 있더라도 반드시 인후염이라 할 수는 없지만, 인두에 이물감이 없다면 다른 수반 증상이 모두 발생한다 하더라도 인후염이라 할 수 없다.

따라서 퍼지 ART 알고리즘을 적용하여 증상에 대한 질병을 도출하는 과정에서 승자가 된 클러스터가 필수 증상을 가지고 있음에도 불구하고, 사용자의 입력 벡터에 해당 필수 증상이 포함되어 있지 않을 경우에는 출력값을 0로 설정해서 승자 노드로 선택되는 것을 방지하여 정확한 질병을 도출할 수 있도록 한다.

2.4 개선한 퍼지 ART 알고리즘

본 논문에서 제안한 한방 자가 진단 및 학습 시스템은 제시된 증상들에 대해 질병들을 정확히 분류하기 위하여 기존의 퍼지 ART 알고리즘을 개선하여 적용한다. 사용자가 선택한 증상을 입력 벡터로 제시하면 개선한 퍼지 ART 알고리즘으로 각 한방 질병 클러스터의 소속도를 산출하여 질병들을 도출한다. 퍼지 ART 알고리즘의 일반적 특징은 다음과 같다[4].

1. 자율 학습 모델인 퍼지 ART는 입력 패턴에 대하여 목표값 없이 실시간적으로 자율 학습을 통해 클러스터링하는 알고리즘이다.
2. 퍼지 ART 모델은 주어진 입력 패턴과 생성된 클러스터간의 유사도 측정을 통해 새로운 클러스터를 생성하거나 기존의 클러스터로 통합하는 클러스터 생성 규칙을 적용한다.

지도 학습 알고리즘은 질병과 증상들을 수정 및 삭제, 추가 및 갱신하는 과정에서 모든 증상과 질병의 관계를 재분류하는 비효율성과 함께 학습이 되지 않는 경우가 발생한다. 반면, 퍼지 ART는 자율 학습 방법으로 수정, 삭제 및 추가 시 새로운 질병은 기존의 질병 클러스터에 영향을 주지 않고 새로운 클러스터를 생성하여 질병을 진단할 수 있으므로 지도 학습 방법에서 나타나는 비효율적인 문제점을 해결할 수 있다[5].

본 논문에서 한방 자가 진단 및 학습에 적용하기 위한 개선된 퍼지 ART 알고리즘의 처리과정은 다음과 같다.

- 단계 1 : 각 인수들을 초기화 한다.
- 단계 2 : 입력 벡터가 제시되면 출력 가중치 벡터를 구한다.
- 단계 3 : 출력 벡터를 계산하고 가장 큰 출력값을 가지는 노드를 승자 노드로 선정한다.
- 단계 4 : 승자노드로 선택된 출력 벡터와 경계 변수를 적용하여 유사성을 검증한다.
- 단계 5 : 유사성이 인정되면 학습률과 연결강도를 조정하고 유사성이 인정되지 못하면 새로운 승자 뉴런을 찾아 학습과정을 반복한다.
- 단계 6 : 모든 패턴에 대해 유사성을 만족하는 클러스터가 없으면 새로운 클러스터를 생성한다.

단계 2의 출력 가중치는 이진 벡터로 구성된 입력 벡터와 연속적인 특성을 가지는 연결 가중치 벡터간의 불일치 때문에 출력 벡터의 최대값이 1이 되지 못하는 경우를 해결하기 위하여 최대 출력 벡터를 1로 설정하기 위한 가중치로서, 질병 진단 및 학습 시스템의 질병별 진단 소속 정도에 적용한다[6].

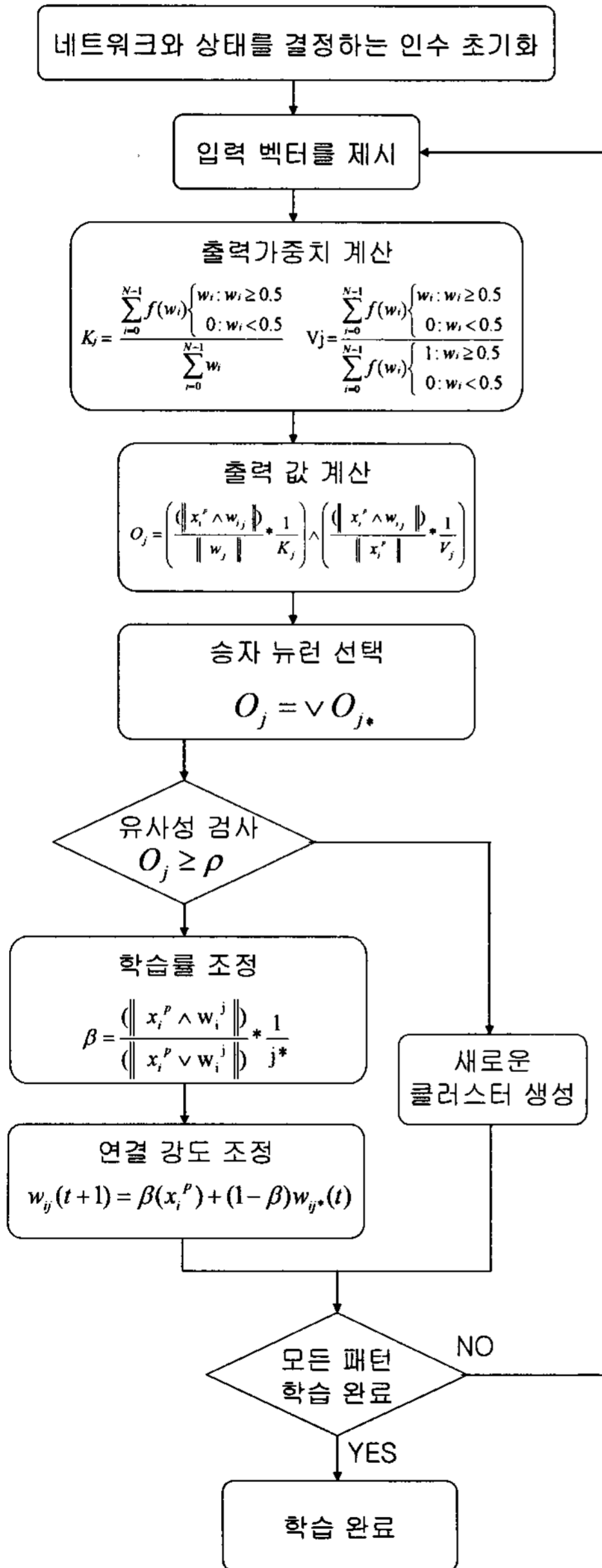


그림 4. 개선한 퍼지 ART 알고리즘

단계 3의 출력 벡터는 퍼지 귀속도에 의해 계산된 값에 해당 출력 가중치로 나눈 값과 경계 변수를 구하는 값에 해당 출력 가중치로 나눈 값간의 최소값으로 계산한다.

구해진 출력값 중 최대값을 승자뉴런으로 선택하고 경계변수와 유사성을 검사한 후, 새로운 클러스터의 생성 유무를 결정하는데, 유사도 검사를 통해 유사성이 인정되면 학습률과 연결 강도를 조정한다. 학습된 최종 출력 벡터는 각 증상에 대한 질병의 소속 정도를 의미한다. 학습률은 입력 벡터와 가중치 벡터의 최소값을 최대값으로 나눈 값에 승자 빈도수로 나눈 값이다. 본 논문에서 제안한 퍼지 ART 알고리즘에서는 승자 노드로 선택되는 빈도수가 증가할수록 학습률은 점차적으로 작아진다. 본 논문에서 한방 질병 분류에 적용되는 개선된 퍼지 ART 알고리즘은 그림 4와 같다.

2.5 학습에 따른 가중치 변화와 활용

질병에 따라 발생하는 증상은 오랜 기간의 의학적 연구에 의해 정형화 되어있는 데이터로 존재한다. 그러나 환경의 변화가 미비했던 과거에는 질병의 변화 역시 미비하여 증상의 변화가 거의 없었으나, 급속히 나빠지는 현재 지구환경의 부정적 변화로 인하여 과거 어느 때 보다 빠른 속도로 바이러스 균이 변하여 질병과 그에 따른 증상도 변화고 있다.

본 논문의 자가 진단 및 학습 시스템은 정형화 되어 있는 증상 데이터를 기반으로 시스템의 데이터베이스가 초기화 되지만 증상이 조금씩 변하는 것과 마찬가지로 학습에 의해 동적으로 수정되게 된다. 그러므로 본 논문의 자가 진단 및 학습 시스템은 관리자의 별도의 증상 추가 행위 없이도 사용자들의 증상 입력에 의해 스스로 학습하여 변경되게 된다. 또 반대로 이러한 학습된 데이터들을 바탕으로 질병에 대한 증상이 변하고 있는 근거로 제시될 수도 있다.

학습된 가중치의 갱신되는 과정은 그림 5에서 확인할 수 있다. 그림 5에서 붉게 표시된 Label은 해당질병의 증상 중 눈에 띄게 나타나는 증상을 의미하고 노란색 Label은 간혹 나타나는 증상을 의미한다.

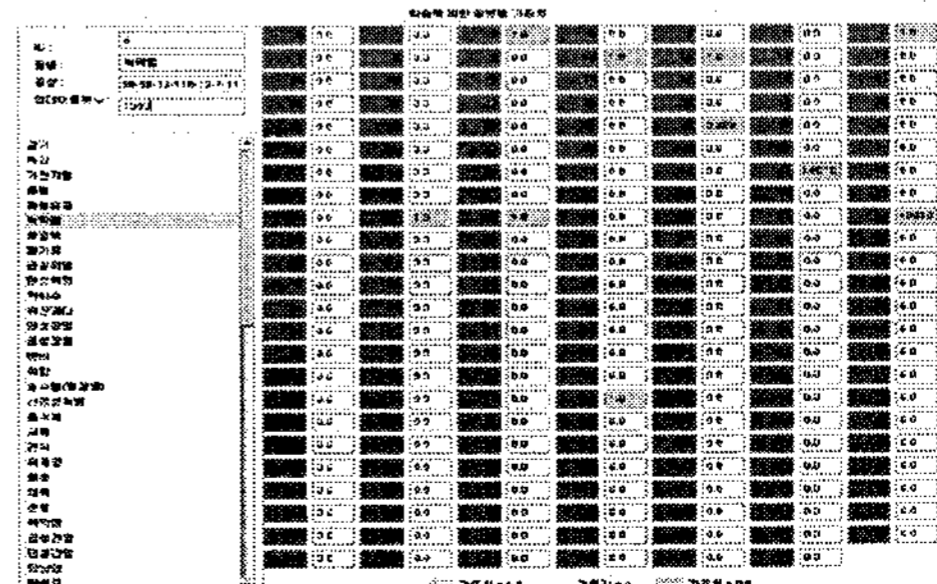


그림 5. 질병별 가중치 업데이트 현황

III. 실험 및 결과 분석

실험 환경은 Intel Pentium IV 3.0GHz CPU와 448MB RAM이 장착된 IBM 호환 PC상의 Java 6.0과 Oracle 9i 환경에서 구현하였다.

실험에 사용된 질병과 증상은 한의학의 질병 분류에 의해 72가지 질병과 그에 따른 167가지의 증상으로 분류하였다.

본 논문에서의 한방 자가 진단 및 학습 시스템은 부위별 증상 중 대표 증상과 그에 관련된 세부 증상을 사용자에게 질의하여 선택된 증상들을 토대로 가장 소속도 높은 상위 3가지 질병을 결과로 도출하였다.

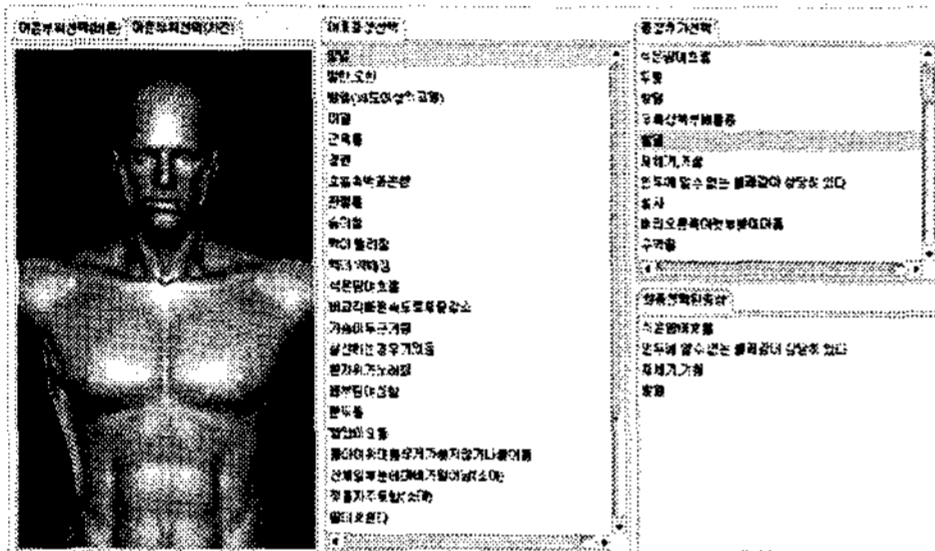


그림 6. 증상 질의 화면

그림 6은 제안된 자가 진단 시스템의 증상 질의 화면이다. 신체 그림 중 한 부분을 선택하면 해당 부위의 증상이 모두 출력되고 그 중 사용자가 대표증상을 선택하면 해당 대표증상이 나타나는 질병들의 모든 세부 증상이 추출되어, 사용자가 최종적으로 세부 증상을 선택하게 된다.

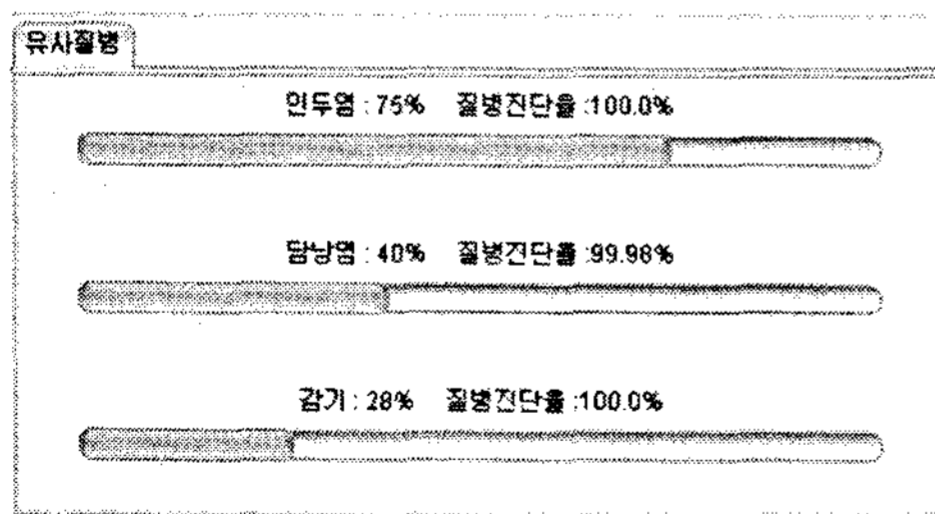


그림 7. 질병 도출 결과 화면

여기서 선택된 최종 증상들을 입력 벡터로 하여 개선한 퍼지 ART 알고리즘에 의해 소속도가 높은 세 가지 질병을 그림 7과 같이 오름차순으로 출력한다. 본 논문의 자가 진단 및 학습 시스템은 결과 값과 함께 질병 진단율도 함께 출력한다.

증상만으로 질병을 판단할 수 있는 진단율은 질병마다 다른데, 해당 질병에 모든 증상이 뚜렷이 나타나서 가중치가 1에 가깝다면 진단율 역시 100%에 가깝게 나타날 수 있지만, 현실의 질병

들은 반드시 발생하지 않는 애매한 증상들을 가지고 있으므로 그 애매성으로 인하여 증상만으로 질병을 판단할 경우에는 질병 진단율이 낮아진다. 따라서 질병의 진단율을 따로 출력함으로써 진단율이 낮은 질병의 경우 정밀 검사를 다시 받을 것을 권장하도록 한다. 그러나 두 가지 이상의 질병이 발병할 경우, 질병의 진단율이 100%이고 해당 질병의 모든 증상을 선택하여도 현저히 낮은 결과가 도출되는 문제점이 발생하였다. 그림 8은 신경성위병과 급성장염에서 발생하는 모든 증상을 입력하였을 때의 결과 화면이다.

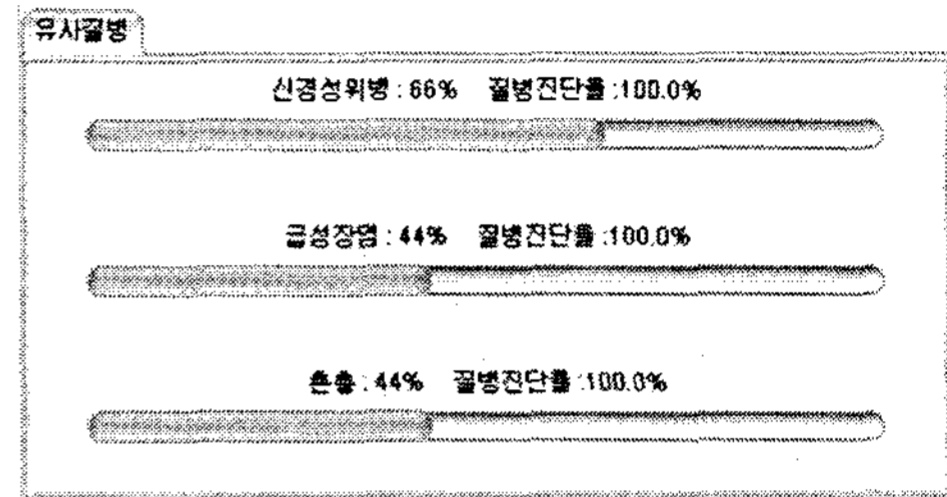


그림 8. 두가지 질병이 발병하였을 때의 질병 도출 결과 화면

IV. 결 론

본 논문에서는 한의학에 대한 전문적인 지식이 없어도 증상을 가지고 한의학에 따른 본인의 건강 상태를 검사할 수 있는 자가 진단 및 학습 시스템을 제안하였다.

본 시스템은 대표 증상을 증상 질의로 받은 후, 개선된 퍼지 ART 알고리즘을 적용하여 입력 벡터와 질병 클러스터간의 소속도를 계산하여 질병을 도출하였다.

구현 및 실험 분석을 통하여 본 논문에서 제안한 한방 자가 진단 및 학습 시스템의 효율성을 보였고, 학습을 통한 가중치 변화량 정보가 의학적으로 활용될 수 있음을 확인하였다. 그러나 본 논문에서는 대표 증상으로 세부 증상을 도출하므로 다수의 질병을 가진 사용자의 경우 각 질병별 소속도가 전반적으로 떨어져서 정확도가 낮아짐을 확인하였다.

따라서 향후 연구 과제는 도출된 질병과 상관 없는 증상이 입력되는 경우에도 해당 증상의 발생이 단순 예외 상황인지 또 다른 질병을 내포하고 있는지를 증상과 증상, 증상과 질병, 질병과 질병의 관계를 퍼지 기법을 적용하여 추론하고 진단할 수 있도록 개선할 것이다.

참고문헌

- [1] 김광백, 김주성, "FCM 알고리즘과 퍼지 소속도를 이용한 지능형 자가 진단 시스템," 한국

- 지능정보시스템학회논문지, 13권 1호, pp.81-90, 2007.
- [2] 김영섭, (허준) 동의보감, 솔빛 출판사, 2003.
- [3] 이철호, CHINESE MEDICINE 한방의학백과, 민중서관, 1999.
- [4] C. M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
- [5] K. B. Kim, H. W. Yun, "A Study on Recognition of Bronchogenic Cancer Cell Image Using a New Physiological Fuzzy Neural Networks," Japanese Journal of Medical Electronics and Biological Engineering, Vol.13, No.5, pp. 39-43, 1999.
- [6] K. B. Kim, S. Kim, K. B. Sim, "Nucleus Classification and Recognition of Uterine Cervical Pap-Smears Using Fuzzy ART Algorithm," Lecture Notes in Computer Science, LNCS 4247, Springer, pp.560-567, 2006.