(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0086559
(43) 공개일자 2012년08월03일

(51) 국제특허분류(Int. Cl.)
G06F 19/00 (2011.01) *G06F 17/30* (2006.01)
G06F 17/10 (2006.01) *G06Q 50/22* (2012.01)
(21) 출원번호 10-2011-0007873
(22) 출원일자 2011년01월26일
심사청구일자 2011년01월26일

(71) 출원인
계명대학교 산학협력단
대구광역시 달서구 달구벌대로 1095 (신당동)
(72) 발명자
손창식
대구광역시 중구 동덕로40길 27 (동인동3가)
김윤년
대구광역시 서구 달구벌대로 1707, 광장타운 107동 1203호 (내당동)
(뒷면에 계속)
(74) 대리인
김일환

전체 청구항 수 : 총 11 항

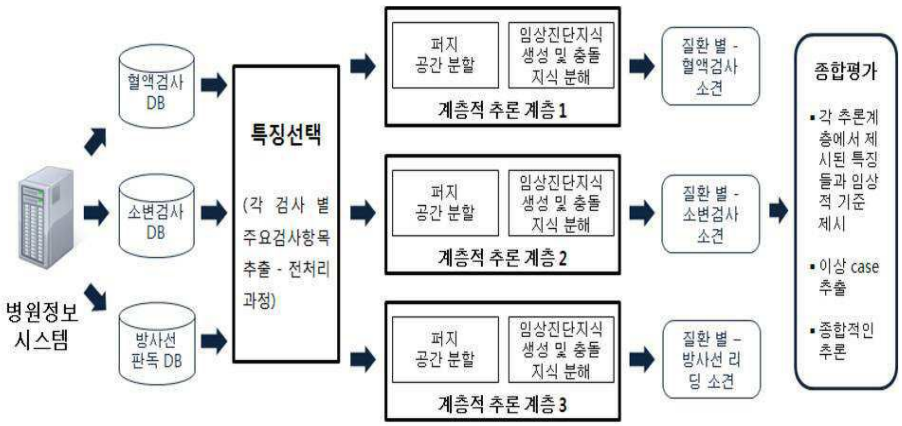
(54) 발명의 명칭 **계층적 퍼지 추론에 기반한 임상진단 지원방법 및 그 시스템**

(57) 요 약

본 발명은 계층적 퍼지추론을 기반한 임상진단 지원방법 및 시스템에 관한 것으로, a) 다수의 환자 각각에 대한 다수의 검사 항목의 검사 데이터를 각 검사항목 데이터 베이스로부터 추출하는 단계; (b) 상기 각 검사 항목의 검사 데이터에서 특징을 선택하는 단계; (c) 상기 특징이 선택된 검사 데이터를 각 항목별로 계층적 퍼지 추론하는 단계; (d) 상기 각 항목별로 상기 퍼지 추론한 출력값을 바탕으로 질환별 상기 검사 소견을 추출하는 단계; 및 (e) 상기 각 소견을 바탕으로 종합적으로 임상 평가하는 단계를 포함하는 것을 특징으로 한다.

이와 같은 본 발명을 제공하면, 그리고, 응급실 등에 내원 한 환자들의 임상정보(혈액, 소변, 방사선 결과 등)를 계층별 추론 결과와 종합적인 추론 결과를 토대로, 질환별 감별진단을 위한 주요 파라미터, 임상적 기준 및 최적화된 지식을 제공함으로써, 응급환자 등의 추론 가능한 질환을 유추할 수 있으며 의대생을 비롯한 의료진의 교육 및 훈련에 도움을 줄 수 임상 진단지원 방법 및 시스템을 제공하게 된다.

대 표 도 - 도1



(72) 발명자

박희준

대구광역시 달서구 새동네로 34, 보람타운 102동 502호 (용산동)

서석태

울산광역시 남구 봉월로 133, 103동 1109호 (신정동, 남산포스코더샵아파트)

김민수

대구광역시 수성구 지범로40길 6, 신화맨션 101동 1310호 (범물동)

이 발명을 지원한 국가연구개발사업

과제고유번호	RTI04-01-01
부처명	지식경제부
연구사업명	지방기술혁신사업
연구과제명	u-Health 기반 진단/예측전문가시스템 상용화 기술개발
주관기관	계명대학교
연구기간	2010.07.01 ~ 2011.06.30

특허청구의 범위

청구항 1

(a) 다수의 환자 각각에 대한 다수의 검사 항목의 검사 데이터를 각 검사항목 데이터 베이스로부터 추출하는 단계;

(b) 상기 각 검사 항목의 검사 데이터에서 특징을 선택하는 단계;

(c) 상기 특징이 선택된 검사 데이터를 각 항목별로 계층적 퍼지 추론하는 단계;

(d) 상기 각 항목별로 상기 퍼지 추론한 출력값을 바탕으로 질환별 상기 검사 소견을 추출하는 단계; 및

(e) 상기 각 소견을 바탕으로 종합적으로 임상 평가하는 단계를 포함하는 것을 특징으로 하는 계층적 퍼지 추론에 기반한 임상지원 진단방법.

청구항 2

제1항에 있어서,

상기 검사 항목은,

혈액검사, 소변검사 및 방사선 판독 검사인 것을 특징으로 계층적 퍼지 추론에 기반한 임상지원 진단방법.

청구항 3

제1항에 있어서,

상기 (b) 단계는,

상기 각 검사 데이터에 대한 임상진단 결과를 표식하는 단계; 및

상기 검사 항목별로 임상진단 결정을 위한 컷 오프(cut-off) 값을 산출하는 단계인 것을 특징으로 하는 계층적 퍼지 추론에 기반한 임상지원 진단방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 (c) 단계는,

복수개의 퍼지 시스템의 계층적 입출력 구조로서,

상기 각 검사 항목의 검사 데이터를 입력으로 하는 하나의 퍼지 시스템의 출력과 또 다른 검사 데이터를 입력을 새로운 입력으로 하여 출력값을 생성하되,

상기 퍼지 시스템의 입력은 적어도 3개 이상인 것을 특징으로 하는 계층적 퍼지 추론에 기반한 임상지원 진단방법.

청구항 5

제4항에 있어서,

상기 (c) 단계는,

상기 각 퍼지 시스템에서,

상기 각 검사 항목의 검사 데이터를 입력값을 퍼지 공간 분할하는 단계; 및

상기 분할된 데이터로부터 임상진단 지식을 생성하는 단계를 포함하는 것을 특징으로 하는 계층적 퍼지 추론에 기반한 임상지원 진단방법.

청구항 6

제5항에 있어서,

상기 퍼지 공간의 분할은 상기 검사 데이터의 통계적 분포를 통해, 정상군(Normal Group)과 비정상군(Abnormal Group)으로 분할하는 것을 특징으로 하는 계층적 퍼지 추론에 기반한 임상지원 진단방법.

청구항 7

제4항에 있어서,

상기 임상진단 지식의 생성은,

상기 분할된 데이터로부터 제1차 임상진단 지식을 생성하는 단계; 및

상기 제1차 임상진단 지식 간 충돌을 회피하기 위한 알고리즘 적용하여 각 계층별 퍼지 시스템에 적용되는 제2차 임상진단 지식을 생성하는 단계를 포함하는 것을 특징으로 하는 계층적 퍼지 추론에 기반한 임상지원 진단방법.

청구항 8

제7항에 있어서,

상기 (e) 단계는,

상기 각 추론 계층에서 제시된 임상적 특징 및 기준을 추출하는 단계;

상기 각 추론 계층의 데이터로부터 이상 케이스(case)를 추출하는 단계; 및

상기 특징, 기준 및 이상 케이스를 바탕으로 종합적인 임상적 추론을 하는 단계를 포함하는 것을 특징으로 하는 계층적 퍼지 추론에 기반한 임상지원 진단방법.

청구항 9

제1항 내지 제3항 중 어느 한 항의 방법을 수행하기 위한 것으로,

복수개의 항목에 따라 적어도 하나의 검사 데이터를 저장한 검사 데이터 베이스;

상기 데이터 베이스로부터 상기 각 항목별 특징 데이터를 선택하기 위한 전처리부;

상기 각 항목별 특징 데이터를 입력으로 하여 계층적 퍼지 추론을 수행하는 퍼지 추론부; 및

상기 퍼지추론을 통해 추출된 각 항목별 검사 소견을 바탕으로 종합적으로 임상 평가하는 종합 평가부를 포함하는 것을 특징으로 하는 계층적 퍼지 추론을 이용하는 임상지원 진단 시스템.

청구항 10

제9항에 있어서,

상기 종합 평가부는,

각 추론 계층의 출력값을 통해 특징 및 임상적 기준을 제공하는 것을 특징으로 하는 계층적 퍼지 추론을 이용하

는 임상지원 진단 시스템.

청구항 11

제9항에 있어서,

각 추론 계층의 출력값을 통해 이상 케이스를 추출하는 것을 특징으로 하는 계층적 퍼지 추론을 이용하는 임상지원 진단 시스템.

명 세 서

기 술 분 야

[0001] 본 발명은 임상진단 지원방법 및 그 시스템에 관한 것으로, 계층적 퍼지 추론의 분류 능력을 개선하기 위한 매개층(intermediate layer)에서의 오류분류 패턴의 수를 최소화하는 퍼지추론 기술에 기반한 임상진단 지원방법 및 그 시스템에 관한 것이다.

배 경 기 술

[0002] 퍼지 시스템의 설계는 일반적으로 지식 생성에 관련된 시간 소비 활동성(a time-consuming activity)과, 전체 규칙수의 줄임, 관련 파라미터 변형에 있다. 이 모든 것들은 퍼지 시스템의 정확성 및 이해력에 영향을 미친다.

[0003] 최근에, 연습(즉 훈련) 데이터로부터 퍼지 규칙을 자동 생성할 수 있는 다양한 접근들이 제안되어 왔으며, 대표적인 방법들로는 휴리스틱 접근, neuro-fuzzy 접근, genetic 접근, '차원의 저주' (curse of dimensionality)의 문제를 극복하기 위한 계층적 접근 등이 있다.

[0004] 대부분의 표준 퍼지 시스템(SFs)의 경우에는 입력변수의 수(즉 차원)가 증가할수록 생성 가능한 후보 퍼지규칙의 수가 지수적으로 증가하는 문제점을 가지고 있기 때문에, neural network와 genetic 알고리즘과 같은 추가적인 학습기술들이 필요할 뿐만 아니라 분류 혹은 추론을 위한 지식베이스를 구축하는데 오랜 시간이 소요되는 문제점을 가진다.

[0005] 계층형 퍼지 시스템에서의 비선형 함수의 근사화 능력들을 포함하는 다양한 관련 연구들이 수행되어져 왔다. 계층적 퍼지 시스템은 표준 퍼지 시스템에서 계산량의 폭발 문제를 극복하기에 유용하다는 것을 확신하는 결과들을 가져오고, 비선형 과정들에서의 파라미터를 최적화하기 위한 진일보한 접근방법들을 역시 포함한다.

[0006] 그러나, 계층형 퍼지 스시템의 주요한 문제는 매개층에서의 결과들이 이전층의 결과와 차후층의 입력값들이 다음 계층들에 대한 변수로서 맵핑될 때 물리적인 의미를 포함하지 않는다는 것이다. 그러므로, 이러한 문제를 줄이기 위해서 계층형 퍼지 시스템에서 중간층과 관련된 규칙의 수를 줄이기 위한 규칙 기반 모습을 묘사해야 하는 문제점이 있다.

발명의 내용

해결하려는 과제

[0007] 상술한 문제를 해결하기 위한 본 발명의 과제는 통계적 특성에 기초한 계층적 퍼지 분류 시스템의 스키마(scheme)를 제공하고자 함이다. 또한, 계층형 퍼지 분류 시스템의 분류 능력을 개선하는 매개층에서의 오분류 패턴의 수를 최소화하는 방법 및 그 시스템을 제공하고자 함에 있다.

과제의 해결 수단

[0008] 상술한 문제를 해결하기 위한 본 발명의 제1 특징은 (a) 다수의 환자 각각에 대한 다수의 검사 항목의 검사 데이터를 각 검사항목 데이터 베이스로부터 추출하는 단계; (b) 상기 각 검사 항목의 검사 데이터에서 특징을 선택하는 단계; (c) 상기 특징이 선택된 검사 데이터를 각 항목별로 계층적 퍼지 추론하는 단계; (d) 상기 각 항목별로 상기 퍼지 추론한 출력값을 바탕으로 질환별 상기 검사 소견을 추출하는 단계; 및 (e) 상기 각 소견을

바탕으로 종합적으로 임상 평가하는 단계를 포함한다.

[0009] 여기서, 상기 검사 항목은, 혈액검사, 소변검사 및 방사선 판독 검사인 것이 바람직하고, 상기 (b) 단계는, 상기 각 검사 데이터에 대한 임상진단 결과를 표식하는 단계; 및 상기 검사 항목별로 임상진단 결정을 위한 컷 오프(cut-off) 값을 산출하는 단계인 것이 바람직하다.

[0010] 또한, 바람직하게는 상기 (c) 단계는, 복수개의 퍼지 시스템의 계층적 입출력 구조로서, 상기 각 검사 항목의 검사 데이터를 입력으로 하는 하나의 퍼지 시스템의 출력과 또 다른 검사 데이터를 입력을 새로운 입력으로 하여 출력값을 생성하되, 상기 퍼지 시스템의 입력은 적어도 3개 이상인 것일 수 있다.

[0011] 또한, 상기 (c) 단계는, 상기 각 퍼지 시스템에서, 상기 각 검사 항목의 검사 데이터를 입력값을 퍼지 공간 분할하는 단계; 및 상기 분할된 데이터로부터 임상진단 지식을 생성하는 단계를 포함하는 것이 바람직하고, 상기 퍼지 공간의 분할은 상기 검사 데이터의 통계적 분포를 통해, 정상군(Normal Group)과 비정상군(Abnormal Group)으로 분할하는 것이 바람직하다.

[0012] 또한, 상기 임상진단 지식의 생성은, 상기 분할된 데이터로부터 제1차 임상진단 지식을 생성하는 단계; 및 상기 제1차 임상진단 지식 간 충돌을 회피하기 위한 알고리즘 적용하여 각 계층별 퍼지 시스템에 적용되는 제2차 임상진단 지식을 생성하는 단계를 포함하는 것이 바람직하다.

[0013] 더하여. 상기 (e) 단계는, 상기 각 추론 계층에서 제시된 임상적 특징 및 기준을 추출하는 단계; 상기 각 추론 계층의 데이터로부터 이상 케이스(case)를 추출하는 단계; 및 상기 특징, 기준 및 이상 케이스를 바탕으로 종합적인 임상적 추론을 하는 단계를 포함하는 것이 바람직하다.

[0014] 그리고, 본 발명의 제2 특징은 상술한 방법을 수행하기 위한 것으로, 복수개의 항목에 따라 적어도 하나의 검사 데이터를 저장한 검사 데이터 베이스; 상기 데이터 베이스로부터 상기 각 항목별 특징 데이터를 선택하기 위한 전처리부; 상기 각 항목별 특징 데이터를 입력으로 하여 계층적 퍼지 추론을 수행하는 퍼지 추론부; 및 상기 퍼지추론을 통해 추출된 각 항목별 검사 소견을 바탕으로 종합적으로 임상 평가하는 종합 평가부를 포함한다.

[0015] 여기서, 상기 종합 평가부는, 각 추론 계층의 출력값을 통해 특징 및 임상적 기준을 제공하는 것이 바람직하고, 각 추론 계층의 출력값을 통해 이상 케이스를 추출하는 것이 바람직하다.

발명의 효과

[0016] 이와 같이 본 발명을 제공하게 되면, 계층적 퍼지 추론 기술을 이용하여 계층별 및 종합적인 임상 추론결과를 제공함으로써, 응급 환자의 상태를 파악하여 가장 적합한 질환을 유추할 수 있는 방법을 제공하게 된다. 또한, 이 모든 지식이 병원 정보시스템 내에 수년 혹은 수십 년간 누적된 정보를 근거로 추론 결과를 제공하기 때문에 정확률을 높일 수 있다는 장점이 있다.

[0017] 그리고, 응급실 등에 내원 한 환자들의 임상정보(혈액, 소변, 방사선 결과 등)를 계층별 추론 결과와 종합적인 추론 결과를 토대로, 질환별 감별진단을 위한 주요 파라미터, 임상적 기준 및 최적화된 지식을 제공함으로써, 응급환자 등의 추론 가능한 질환을 유추할 수 있으며, 의대생을 비롯한 의료진의 교육 및 훈련에 도움을 줄 수 임상 진단지원 방법 및 시스템을 제공하게 된다.

도면의 간단한 설명

[0018] 도 1은 본 발명에 따른 계층적 퍼지 추론에 기반한 임상진단 지원방법의 흐름 및 블록 다이어그램을 나타낸 도면,

도 2는 본 발명에 따른 임상진단 지원방법에 적용되는 계층적 퍼지 시스템의 블록 다이어그램을 나타낸 도면,

도 3은 입력변수 백혈구(WBC)에서 정상군(Normal Group)과 비정상군(Abnormal Group)의 데이터 분포의 예를 나타낸 그래프,

도 4는 본 발명에 따른 방법을 적용하여 백혈구 검사항목의 퍼지 공간 분할의 예를 나타낸 도면,

도 5는 본 발명에 따른 방법을 적용하여 입력변수 백혈구 및 적혈구 2변수 퍼지 공간분할의 예를 나타낸 도면,

도 6은 본 발명에 적용된 방법으로 이전 계층의 퍼지추론 결과들을 활용하여 다음 계층의 입력변수의 퍼지 공간을 분할한 예를 나타낸 도면,

도 7은 본 발명에 적용되는 이전 계층의 퍼지 추론 결과로부터 생성되는 퍼지 함수를 나타낸 모식도이다.

발명을 실시하기 위한 구체적인 내용

[0019] 이하에서 본 발명에 따른 바람직한 실시예를 도면을 참조하여 상세히 설명하기로 한다.

[0020] 도 1은 본 발명에 따른 계층적 퍼지 추론에 기반한 임상진단 지원방법의 흐름 및 블록 다이어그램을 나타낸 도면이다. 도 1에 나타낸 바와 같이, 본 발명에 따른 임상지원 진단방법은 (a) 다수의 환자 각각에 대한 다수의 검사 항목의 검사 데이터를 각 검사항목 데이터 베이스로부터 추출하는 단계; (b) 상기 각 검사 항목의 검사 데이터에서 특징을 선택하는 단계; (c) 상기 특징이 선택된 검사 데이터를 각 항목별로 계층적 퍼지 추론하는 단계; (d) 상기 각 항목별로 상기 퍼지 추론한 출력값을 바탕으로 질환별 상기 검사 소견을 추출하는 단계; 및 (e) 상기 각 소견을 바탕으로 종합적으로 임상 평가하는 단계를 포함하여 구성한다.

[0021] 그리고, 본 발명에 따른 또 다른 발명으로서, 상술한 방법을 수행하기 위한 시스템으로서, 복수개의 항목에 따라 적어도 하나의 검사 데이터를 저장한 검사 데이터 베이스; 상기 데이터 베이스로부터 상기 각 항목별 특징 데이터를 선택하기 위한 전처리부; 상기 각 항목별 특징 데이터를 입력으로 하여 계층적 퍼지 추론을 수행하는 퍼지 추론부; 및 상기 퍼지추론을 통해 추출된 각 항목별 검사 소견을 바탕으로 종합적으로 임상 평가하는 종합평가부를 포함한다.

[0022] 상기 (a) 단계 입력되는 데이터는 다수의 환자의 검사 데이터로 구성되며, 각 환자에 대한 검사 데이터는 다수의 검사 항목으로 구성되며, 각 환자에 대한 임상진단 결정된 상태이다.

[0023] 이와 같은 병원정보 시스템의 검사 데이터 베이스를 통해 다수 환자의 검사 데이터를 검사 항목별로 분류하여 정렬하고, 각 검사 데이터에 임상진단 결과를 표식하고, 특징을 선택한다.

[0024] 특징의 선택방법을 구체적으로 살펴보면, 다수의 환자의 검사 데이터를 검사 항목별로 분류하여 정렬한다. 예를 들어, (I) 소변검사 항목으로는 뇨색(color), 알부빈(albumin) 수치, 백혈구(white blood cell) 수, 적혈구(red blood cell) 수 등과 같은 수많은 검사항목들이 있으며, (II) 혈액검사의 경우, 전혈구 검사항목 (백혈구 수, 적혈구 수, 헤모글로빈 수치 등), 동맥가스혈 검사항목 (나트륨, 칼륨, 아밀라아제 수치 등), 입원환자에 대한 기본검사항목 (칼슘, 글루코스, 단백질, 콜레스테롤 수치 등) 등이 있으며, (III) 방사선 검사소견 항목으로는 흉부 X-선, CT, MRI 등 각 임상 분과별 질환 진단을 위한 중요 리딩 파라미터들을 정렬한다.

[0025] 상기한 바와 같이 다수의 환자 각각에 대한 다수의 검사 항목의 검사 데이터를 입력받아 검사 항목별로 분류하여 값의 크기에 따라 정렬하여 각각 검사 데이터에 임상진단 결과를 표식하면, 상기 임상진단 결과가 충첩되는 영역이 존재한다.

[0026] 상기 중첩되는 영역을 토대로 임상지원 시스템은 상기 임상진단 결정을 위한 적합도를 평가한다. 즉, 중첩영역에 포함되는 다수의 검사 데이터 각각에 대해 임상진단 결과별로 분리될 수 있는 정도를 나타내는 분리가능정도를 산출하고, 상기 분리가능정도의 합을 산출하여 해당 검사 데이터에 따른 검사 항목이 상기 임상진단 결정을 위해 어는 정도 적합한지를 나타내는 적합도를 결정한다.

[0027] 이러한 과정을 통해 모든 검사 항목에 대해 임상진단 결정에 대한 적합도를 산출하고, 그 산출된 적합도를 토대로 해당 임상진단 결정을 위함 검사 항목 순위를 결정할 수 있다. 그리고 상기 검사 항목별 적합도가 결정되면, 상기 임상지원 시스템은 상기 검사 항목별 임상진단 결과를 결정하기 위한 컷 오프 값을 산출한다. 상기 컷 오프 값은 중첩 영역내에서 검사 데이터의 값의 가중치 평균값이다.

[0028] 상기한 바와 같이 검사 항목별 적합도와 검사 항목별 컷 오프 값이 결정되면, 상기 검사 항목들을 통한 임상진단 결정을 위한 특징 또는 규칙패턴을 생성할 수 있다. 이것이 상기 (b) 단계에서 특징을 선택하는 단계이다. 전처리 단계라고도 한다.

[0029] 여기서, 상기 중첩영역의 중복도(degree of redundancy)를 산출하고, 이를 통해 특징을 선택하게 되는데, 중복도(degree of redundancy)는 이하 [수학식 1]과 같이 나타낸다.

수학식 1

$$N_j = \sum_{x_i \in C_k} h\left(x_{ij}\right) \Big/ s,\ x_{ij} \in a_j,\ j = 1, \ldots, n$$

[0030]

[0031] 상기 [수학식 1]에서 $h(x_{ij})$는 i번째 속성에서 j번째 클래스(정상군 혹은 비정상군)에 포함된 유일한 속성값들의 빈도수를 의미하고, s는 이러한 속성값의 전체 수를 나타낸다. 이때 N_j 중첩영역에서 속성값들의 상대적인 분리 가능성 정도를 나타내고, $N_j \in [0,1]$의 값을 가진다.

[0032] 이처럼 (b) 단계에서 (b) 상기 각 검사 항목의 검사 데이터에서 특징을 선택한 후, (c) 단계에서 상기 특징이 선택된 검사 데이터를 각 항목별로 계층적 퍼지 추론하게 되는데, 도 2는 본 발명에 따른 임상진단 지원방법에 적용되는 계층적 퍼지추론 시스템의 블록 다이어그램을 나타낸 도면이다. 도 1 (a)는 n개의 입력 계측적 퍼지 시스템(HFCS)의 구조를 나타낸 것이고, 도 2(b)는 계층적 퍼지 시스템(HFCS)에서 j번째 퍼지 시스템의 하부구조를 나타낸다.

[0033] 도 2에 나타낸 바와 같이, a_1, a_2, a_3, ..., a_n 은 입력변수를 의미하고, 소변검사 항목 또는 혈액검사 항목 또는 방사선 검사 소견 항목들이 하나의 입력변수가 된다.

[0034] 상술한 바와 같이, (I) 소변검사 항목으로는 뇨색(color), 알부빈(albumin) 수치, 백혈구(white blood cell) 수, 적혈구(red blood cell) 수와 같은 수많은 검사항목들이 있으며, (II) 혈액검사의 경우, 전혈구 검사항목(백혈구 수, 적혈구 수, 헤모글로빈 수치 등), 동맥가스혈 검사항목 (나트륨, 칼륨, 아밀라아제 수치 등), 입원환자에 대한 기본검사항목 (칼슘, 글루코스, 단백질, 콜레스테롤 수치 등) 등이 있으며, (III) 방사선 검사소견 항목으로는 흉부 X-선, CT, MRI 등 각 임상 분과별 질환 진단을 위한 중요 리딩 파라미터들이 있다. 이들 주요 3가지(소변, 혈액 및 방사선) 검사항목들의 검사결과 및 소견들이 본 발명의 계층적 퍼지추론 시스템의 입력값이 된다.

[0035] 도 2에서 y_1, y_2,,....., y_{n-1} 은 각 계층의 출력결과 값을 의미한다. 예를 들어, 4가지 소변검사 항목(뇨색상, 백혈구 수, 적혈구 수, 알부민)만으로 의사 결정을 위한 계층형 추론을 수행한다면, a_1 과 a_2는 백혈구 수와 적혈구 수를, a_3는 뇨 색상을, a_4는 알부민 변수라고 정의할 수 있으며, y_1은 백혈구와 적혈구 수치로 추론된 결과값을, y_2는 y_1의 추론결과값과 알부민 수치 값들로부터 추론된 결과값을 의미합니다. 따라서 최종 출력값인 y_3는 이들 4가지 검사항목(백혈구, 적혈구, 뇨 색상, 알부민)을 모두 종합한 추론결과로 해석할 수 있다.

[0036] 이처럼 본 발명에 적용되는 계층적 퍼지추론 시스템은, 상기 각 검사 항목의 검사 데이터를 입력으로 하는 하나의 퍼지 시스템의 출력과 또 다른 검사 데이터를 새로운 입력을호 하는 퍼지시스템을 계속 생성하여 형성된 것으로, 그 출력값이 상기 입력들이 모두 종합한 값을 출력으로 한다.

[0037] 그리고, 도 2에 나타낸 바와 같이, 계층적 퍼지 추론하는 단계는, 복수개의 퍼지 시스템의 계층적 입출력 구조로서, 상기 각 검사 항목의 검사 데이터를 입력으로 하는 하나의 퍼지 시스템의 출력과 또 다른 검사 데이터를 입력으로 출력값을 생성하되, 상기 퍼지 시스템의 입력은 적어도 3개 이상인 것이 바람직하다.

[0038] 즉, 본 발명은 종래의 단순히 입력된 정보와 기존의 누적된 임상 데이터베이스로부터 키워드 매핑된 사례를 통해서 추론결과를 단계적으로 제공하는 방법과 달리, 병원내의 임상 데이터베이스의 질환별 RAW 데이터(기본 임상검사 - 소변, 혈액검사 및 방사선 검사 리딩 결과 등)로부터 임상지식을 자동구성하기 때문에, 종래의 키워드 검색 혹은 매핑 방법으로의 문제점(임상 데이터베이스에 대응된 사례가 없을 경우에 질환, 진단정보 및 치료법에 대한 추론결과를 제공할 수 없음)이 발생되지 않으며, 단계적인 형태의 근사 추론이 가능하다.

[0039] 또한 본 발명은 도 2에 나타낸 바와 같이, 상기 각 퍼지 시스템에서, 상기 각 검사 항목의 검사 데이터를 입력값을 퍼지 공간 분할하는 단계; 및 상기 분할된 데이터로부터 임상진단 지식을 생성하는 단계를 포함한다.

[0040] 도 3은 입력변수 백혈구(WBC)에서 정상군(Normal Group)과 비정상군(Abnormal Group)의 데이터 분포를 나타낸

그래프이다. 도 3 나타낸 바와 같이, MIN은 입력변수 WBC의 최소값, MAX는 최대값을 나타내고, 이 변수의 도메인은 Dj=[min(a_j), max(a_j)]와 같이 나타낼 수 있다.

[0041] 이 도메인의 찾는 단계는 퍼지 공간 분할의 제1 단계로서, 데이터 속성값들의 내부구간을 추출하는 단계이다. j번째 속성 a_j의 도메인 D_j=[min(a_j), max(a_j)]과 임의의 클래스에 대응하는 j번째 속성의 속성값들의 내부구간 $I_{jk} \in c_k$ 들을 추출한다.

[0042] 퍼지 공간 분할의 2번째 단계는 추출된 내부구간들 사이의 중첩영역을 검출하는 단계이다. 상기 추출된 내부구간들 사이에서 중첩영역(overlapping region) O_j=[o_j^{Lj} , o_j^{Uj}] 는 모든 중첩된 영역들 사이에서 하한(lower)과 상한 경계(upper bound)를 나타낸다.

[0043] 그리고 최종적으로 입력변수 WBC 퍼지 공간 분할(fuzzy partitioning)식은 [수학식 2]와 같다.

수학식 2

[0044] $$P_j = round\left(\left\{\max\left(a_j\right) - \min\left(a_j\right)\right\} \middle/ \alpha \cdot \sigma_j\right)$$

[0045] 여기서 j번째 입력변수(즉 백혈구 항목)의 공간 분할 P_j는 백혈구의 도메인의 최대값에서 최소값의 차를 $\alpha \times \sigma_j$($\alpha > 0$)로 나눈 값의 수를 퍼지 공간분할 수로 결정한다. 다시 말해서, 백혈구 항목의 전체 데이터에 대한 표준편차 값 σ_j은 가변적인 α의 값에 따라 백혈구 항목의 퍼지 공간분할 수가 결정되는 것을 의미한다.

수학식 3

[0046] $$W_j = \left(o_j^{U_j} - o_j^{L_j}\right) \middle/ \left(P_j - 1\right)$$

[0047] [수학식 3]은 [수학식 2]에서 결정된 백혈구 항목의 공간분할 수 P_j와 [수학식 3]에서 결정된 중첩영역의 하한값 o_j^{Lj} 과 상한값 o_j^{Uj}]을 이용하여 도 4와 같은 퍼지공간 분할을 할 수 있다. 도 4는 본 발명에 따른 방법을 적용하여 백혈구 검사항목의 퍼지 공간분할의 예를 나타낸 도면이다.

[0048] 퍼지 추론 단계의 두번째 단계는 퍼지 공간 분할 후, 이를 바탕으로 새로운 임상진단 지식을 생성(Rule generation)하는 단계를 수행한다. 도 5는 본 발명에 따른 방법을 적용하여 입력변수 백혈구 및 적혈구의 퍼지 공간분할의 예를 나타낸 도면이다.

[0049] 도 5에 나타낸 바와 같이, Rule 1: WBC is normal, RBC is normal 규칙의 형식은 $R_k^1 : \text{If } a_1 \text{ is } A_{k,1} \text{ and } a_2 \text{ is } A_{k,2} \text{ Then } y_1 \text{ is } c_k$ 의 조건부에 해당한다. 정상군과 비정상군의 수치 데이터가 입력되어, 다음과 같은 후보규칙 패턴이 생성된다고 가정한다.

수학식 4

[0050] $$\mu_k^1(x_i) = \mu_{k,1}(x_{i1}) \times \mu_{k,2}(x_{i2})$$

수학식 5

[0051] $$\mu_k^{1*}(x_i) = \max_k \left\{ \mu_k^1(x_i) : k = 1, \ldots, m \right\}$$

[0052] **<정상군>**

[0053] R1: If WBC is normal and RBC is normal Then output is normal {0.90, sup=20}

[0054] R2: If WBC is normal and RBC is abnormal Then output is normal {0.8, sup=10}

[0055] **<비정상군>**

[0056] R3: If WBC is abnormal and RBC is abnormal Then output is abnormal {1.0, sup=15}

[0057] R4: If WBC is abnormal and RBC is normal Then output is abnormal {0.6, 7}

[0058] R5: If WBC is normal and RBC is abnormal Then output is abnormal {0.8, sup=3}

[0059] 상기와 같은 규칙 패턴에서, {A, B}는 각각 최대 적합도([수학식 4] 및 [수학식 5]로부터 계산됨)와 대응된 규칙의 발생빈도 수(B)를 의미하고, 정상군의 지식 R2와 비정상군의 지식 R5가 서로 충돌된다. 이를 해결하기 위해서 본 발명에서는 [표 1]을 사용한다. [표 1]은 본 발명에서 적용되는 충돌 규칙 해결 전략을 나타낸 표이다.

표 1

[0060]

Maximum degree / Frequency	$\gamma > \delta$	$\gamma = \delta$	$\gamma < \delta$
$\alpha > \beta$	c_p	c_p	c_p
$\alpha = \beta$	c_p	NA	c_q
$\alpha < \beta$	c_q	c_q	c_q

[0061] 상기 [표 1]의 서로 다른 두 그룹들 간의 충돌 규칙의 경우, 정상군의 지식 R2는 최대 적합도에서 , 발생 빈도수에서 α〉β을 만족하기 때문에 후보 규칙패턴 If WBC is normal and RBC is abnormal Then output is normal {0.8, sup=10}으로 결정된다.

[0062] 이와 같은 방법으로 도 2에서 Fuzzy System 1, Fuzzy System 2, ..., Fuzzy System n-1까지의 후보규칙들을 생성하게 된다. 단지 Fuzzy System 2에서 Fuzzy System n-1 모듈에서의 규칙생성은 이전 단계의 출력들 (즉 Fuzzy System 2의 경우에는 Fuzzy System 1의 추론결과 값)을 이용하여 생성하게 되어 모든 입력값들을 종합적으로 고려하여 추론하게 됨으로써, 오진율을 낮출 수 있는 효과가 있다.

[0063] 그리고, 본 발명에서는 다음 계층의 퍼지 공간 분할을 재조정하게 되는데, 그 방법은 2단계로 구성된다.

[0064] 도 6은 본 발명에 적용된 방법으로 이전 계층의 퍼지추론 결과들을 활용하여 다음 계층의 입력변수의 퍼지 공간을 분할한 예를 나타낸 도면이다. 도 6에 나타낸 바와 같이, 퍼지 공간 분할의 첫 번째 단계는 이전 계층의 퍼지 추론 결과로부터 퍼지 공간 분할의 경계점을 찾는 것이다.

[0065] 여기서, τ_j 는 다음의 [수학식 6]에서 얻어지는 결과값의 표준 편차이다.

수학식 6

[0066]
$$y_j^* = \sum_{y_j} y_j \cdot \mu_{\tilde{k}}(y_j) \Big/ \sum_{y_j} \mu_{\tilde{k}}(y_j), \quad j = 1, \ldots, n-1$$
$$\text{where } \mu_{\tilde{k}}(y_j) = \max_k \left[\mu_k^j(x_i) \times \mu_k(y_j) \right]$$

[0067] 도 7은 본 발명에 적용되는 이전 계층의 퍼지 추론 결과로부터 생성되는 퍼지 함수를 나타낸 모식도이다. 도 7에 나타낸 바와 같이, 두번째 단계는 첫 번째 단계에서 발견된 경계점에 기반한 새로운 퍼지 함수를 생성하는 것이다.

[0068] 이처럼 본 발명에서는 이전 계층(layer)의 퍼지 추론 결과로부터 퍼지 함수의 파라미터를 재조정하게 되는데, 이 과정은 본 발명에 따른 계층적 퍼지추론 시스템(HFCS)에서 적용하는 모든 입력 값들이 사용될 때까지 반복된다.

[0069] **실험결과**

[0070] 본 발명에서 제안된 계층적 퍼지 추론 시스템(HFCS)의 효과를 보여주기 위해, 종래의 3가지 형태의 분류 방법(statistical classifiers, SVMs with three kernel functions, and fuzzy rule learning methods with rules weights) 및 제안된 방법에 대한 실험이 잘 알려진 5가지 데이터셋(Haverman' survival, Blood transfusion service center, New thyroid disease, Pima Indians diabetes, and Wisconsin breast cancer original)에서 수행되었다.

[0071] 이와 같은 데이터 셋의 주요 특징은 [표 2]에 나타내었다. 각 데이터 셋은 랜덤하게 10개의 그룹으로 분리하고, 각 항목의 샘플들은 각 그룹에 일정하게 지정된다. 10-fold CV(cross validation) 및 제안된 HFCS의 입력 조합들은 [수학식 1]에 의해 계산되는 중복도(degree of redundancy)에 따라 결정된다.

표 2

[0072]

Dataset	No. of sample	No. of input feature	No. of class
Hab.	306	3	2
Blood	748	4	2
NT	215	5	3
PID	768	8	2
WBCO	683	9	2

[0073] [표 3] 및 [표 4]는 10-fold CV(cross validation)동안 실험된 평균 분류 정확률을 나타낸다.

표 3

Dataset	Statistical classifiers* (%)			SVMs* (%)		
	LDA	QDA	kNN	Poly.	RBF	Sig.
Hab.	74.82	75.13	66.97	26.47	70.61	73.53
Blood	76.74	38.10	64.45	23.79	74.60	76.21
NT	91.26	96.75	97.21	13.96	70.74	69.81
PID	77.10	74.38	70.33	65.11	65.11	65.11
WBCO	96.05	95.02	96.05	93.71	89.46	65.01
Avg.	**83.19**	75.88	79.00	44.61	**74.10**	69.93
±Sd	**±9.74**	±23.62	±16.23	±33.66	**±9.23**	±5.00

[0074]

표 4

Dataset	Fuzzy rule learning methods* (%)				HFCS (%)
	Chi_RW	Ishibuchi99	Fuzzy slave	Shi et al.	
Hab.	(P=2) 73.53, (P=3) 73.19, (P=4) 71.23	(P=2) 73.53, (P=3) 73.20, (P=4) 73.20	(P=2) 73.20, (P=3) 74.49, (P=4) 71.55	(P=2) 71.23, (P=3) 71.56, (P=4) 74.12	α = 1.27, 77.74 (L_1:17.4, L_2:5.9)
Blood	(P=2) 76.07, (P=3) 76.61, (P=4) 76.87	(P=2) 76.21, (P=3) 76.07, (P=4) 76.21	(P=2) 76.47, (P=3) 75.81, (P=4) 76.47	(P=2) 76.74, (P=3) 75.54, (P=4) 75.94	α = 1.62, 76.20 (L_1:22.5, L_2:8.2, L_3:4.0)
NT	(P=2) 84.24, (P=3) 78.16, (P=4) 89.85	(P=2) 78.61, (P=3) 75.84, (P=4) 69.81	(P=2) 94.46, (P=3) 88.94, (P=4) 79.57	(P=2) 84.26, (P=3) 86.97, (P=4) 85.11	α = 0.9, 93.03, (L_1:35.4, L_2:17.2, L_3:19.9, L_4:18.2)
PID	(P=2) 65.76, (P=3) 72.40, (P=4) 73.44	(P=2) 65.11, (P=3) 64.98, (P=4) 68.24	(P=2) 76.71, (P=3) 74.36, (P=4) 72.14	(P=2) 72.78, (P=3) 71.87, (P=4) 72.50	α = 1.3, 75.01, (L_1:17.5, L_2:8.7, L_3:10.5, L_4:12.4, L_5:11.1, L_6:9.4,L_7:8.6)
WBCO	(P=2) 94.29, (P=3) 91.21, (P=4) 73.21	(P=2) 95.17, (P=3) 96.34, (P=4) 95.76	(P=2) 96.50, (P=3) 95.03, (P=4) 96.63	(P=2) 94.58, (P=3) 95.47, (P=4) 93.13	α = 1.2, 95.76, (L_1:8.2, L_2:8.4,L_3:6.0, L_4:4.9, L_5:6.8, L_6:6.0, L_7:4.6, L_8:5.2)
Avg.±Sd	(P=2) 78.78 ± 10.89	(P=2) 77.73 ± 11.00	**(P=2) 83.47 ± 11.08**	(P=3) 80.28 ± 10.54	**83.55 ± 10.00**

[0075]

[0076] [표 3] 및 [표 4]에서 SVMs의 파라미터 값들은(i.e., kernel types: Polynomial, RBF, and Sigmoid functions; c: 1000; eps: 0.001; degree(d): 10; gamma(g): 1.0; coef0(r): 1.0; nu(n): 0.5; p: 1.0; and

shrinking(h): 0.0) KEEL Ver 1.2에 의해 제공되는 디폴트 값들이다.

[0077] 여기서 'Avg±Sd'는 5개의 데이터셋에 대한 각 분류 방법의 평균 예측성능(performance)를 나타내고, [표 4]에서 'P'는 퍼지 공간 분할의 수이고, 'L_i'는 HFCS의 각 계층에서의 평균 규칙 수를 나타낸다.

[0078] [표 4]에서 제안된 본 발명에 따른 HFCS의 분류 예측성능은 파라미터 α기 1.27, 1.62, 0.9, 1.3 및 1.2로 업데이트 됐을 때 최대임을 알 수 있다. 게다가, 비록, 퍼지 규칙 학습 방법들 사이의 최대 분류 예측성능을 갖는 퍼지 슬래이브(P=2) 보다 분류 정확률이 조금 낮지만, 퍼지 공간 분할의 다른 수를 갖는 본 발명에 따른 HFCS는 5개의 데이터 셋에 대한 종래의 3개의 분류방법보다 더 나은 평균 테스팅 예측성능을 산출함을 알 수 있었다. 이러한 결과들은 본 발명에서 적용되는 HFCS의 예측성능이 중첩영역에서 정의된 퍼지 공간 분할의 수와 밀접하게 관련돼 있음을 나타낸다.

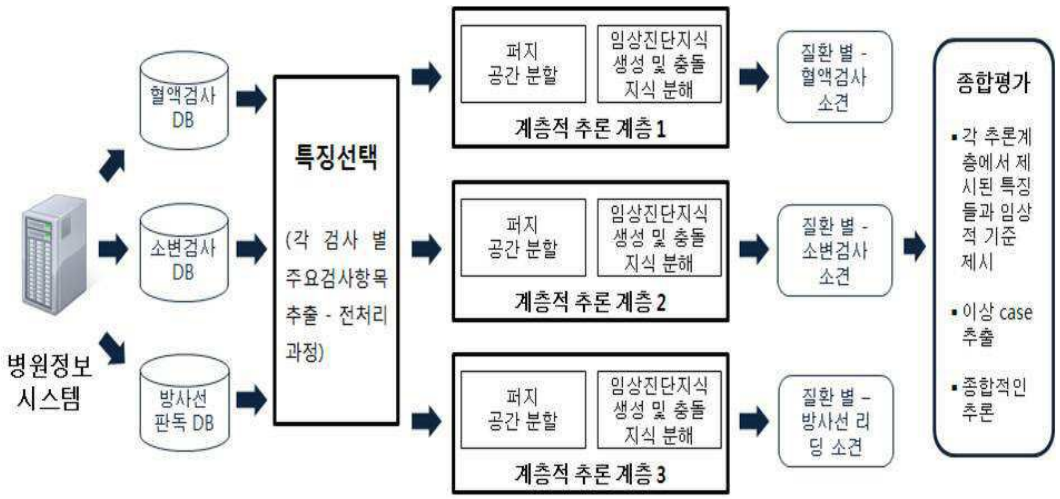
[0079] 이와 같이 본 발명을 제공하게 되면, 계층적 퍼지 추론 기술을 이용하여 계층별 및 종합적인 임상 추론결과를 제공함으로써, 응급 환자의 상태를 파악하여 가장 적합한 질환을 유추할 수 있는 방법을 제공하게 된다. 또한, 이 모든 지식이 병원 정보시스템 내에 수년 혹은 수십 년간 누적된 정보를 근거로 추론 결과를 제공하기 때문에 정확률을 높일 수 있다는 장점이 있다.

[0080] 그리고, 응급실 등에 내원한 환자들의 임상정보(혈액, 소변, 방사선 결과 등)를 계층별 추론 결과와 종합적인 추론 결과를 토대로, 질환별 감별진단을 위한 주요 파라미터, 임상적 기준 및 최적화된 지식을 제공함으로써, 응급환자 등의 추론 가능한 질환을 유추할 수 있으며, 의대생을 비롯한 의료진의 교육 및 훈련에 도움을 줄 수 임상 진단지원 방법 및 시스템을 제공하게 된다.

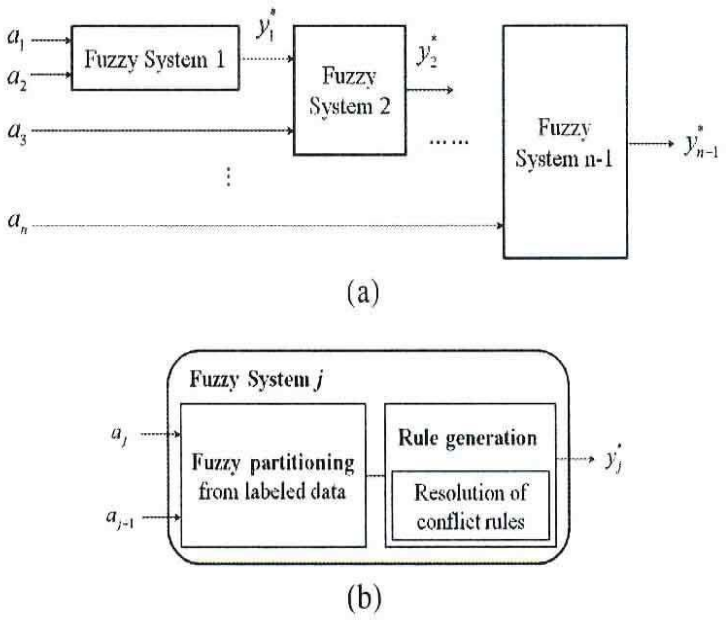
[0081] 이상의 설명에서 본 발명은 특정의 실시 예와 관련하여 도시 및 설명하였지만, 특허청구범위에 의해 나타난 발명의 사상 및 영역으로부터 벗어나지 않는 한도 내에서 다양한 개조 및 변화가 가능 하다는 것을 당 업계에서 통상의 지식을 가진 자라면 누구나 쉽게 알 수 있을 것이다.

도면

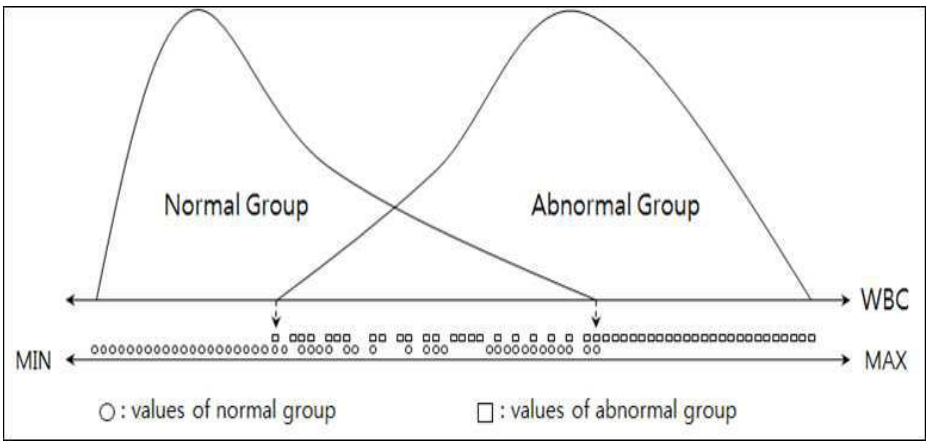
도면1



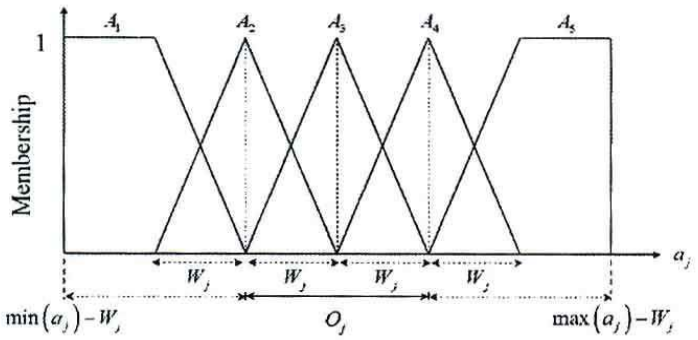
도면2


Fuzzy System 1
Fuzzy System 2
Fuzzy System n-1
(a)
Fuzzy System j
Fuzzy partitioning from labeled data
Rule generation
Resolution of conflict rules
(b)

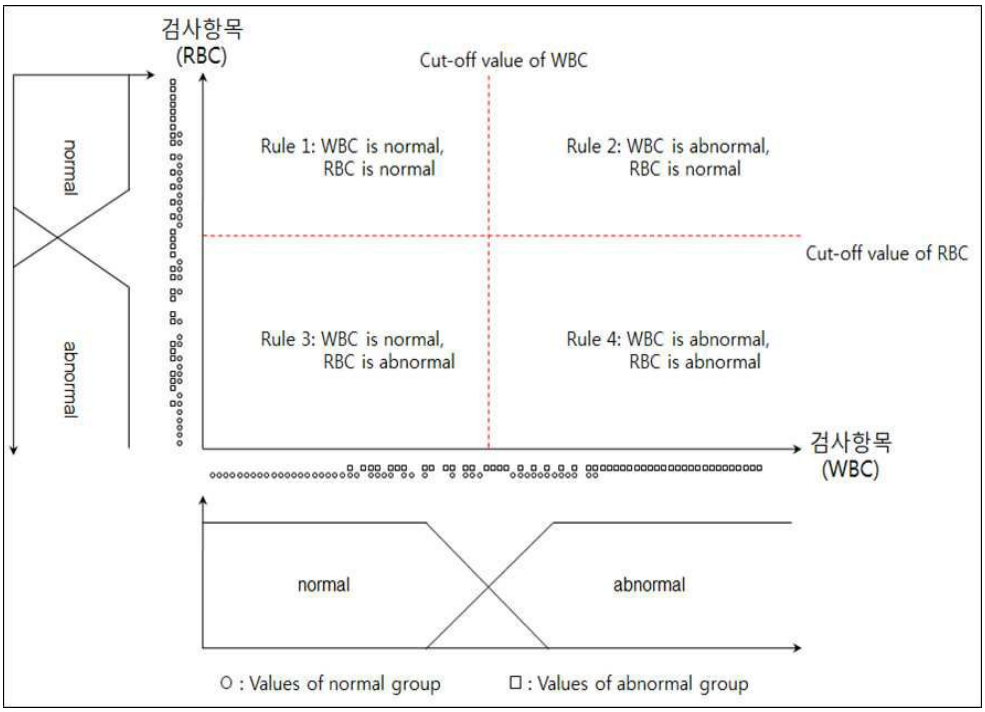
도면3


Normal Group
Abnormal Group
WBC
MIN
MAX
○ : values of normal group
□ : values of abnormal group

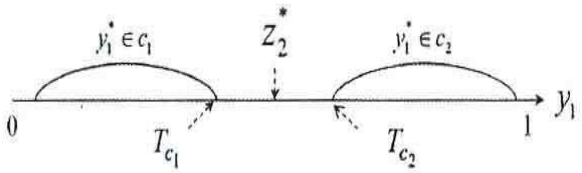
도면4


Membership
1

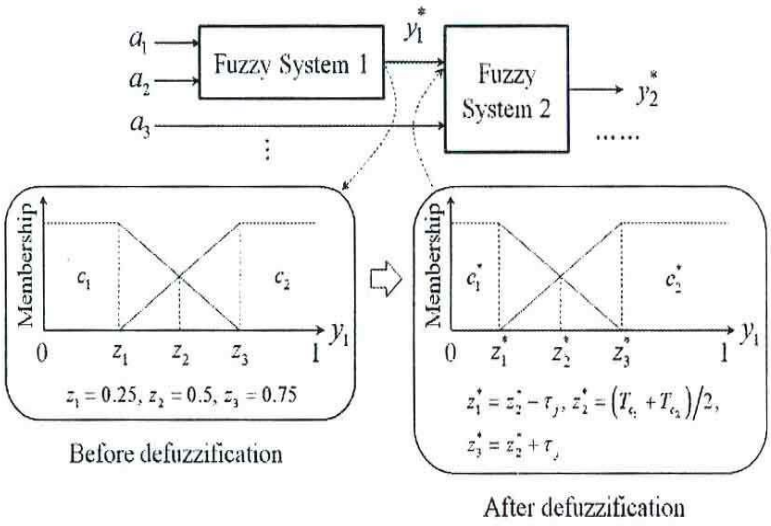
도면5


검사항목
(RBC)
Cut-off value of WBC
normal
abnormal
Rule 1: WBC is normal,
RBC is normal
Rule 2: WBC is abnormal,
RBC is normal
Cut-off value of RBC
Rule 3: WBC is normal,
RBC is abnormal
Rule 4: WBC is abnormal,
RBC is abnormal
검사항목
(WBC)
normal
abnormal
○ : Values of normal group
□ : Values of abnormal group

도면6


$y_1^* \in c_1$
z_2^*
$y_1^* \in c_2$
0
T_{c_1}
T_{c_2}
1
y_1

도면7


a_1
a_2
a_3
Fuzzy System 1
y_1^*
Fuzzy System 2
y_2^*
.....
Membership
c_1
c_2
0
z_1
z_2
z_3
1
y_1
$z_1 = 0.25, z_2 = 0.5, z_3 = 0.75$
Before defuzzification
c_1^*
c_2^*
z_1^*
z_2^*
z_3^*
$z_1^* = z_2^* - \tau_j, z_2^* = (T_{c_1} + T_{c_2})/2,$
$z_3^* = z_2^* + \tau_j$
After defuzzification