(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0086107
(43) 공개일자 2012년08월02일

(51) 국제특허분류(Int. Cl.)
G06F 9/06 (2006.01) *G06F 17/30* (2006.01)
(21) 출원번호 10-2011-0007348
(22) 출원일자 2011년01월25일
심사청구일자 2011년01월25일

(71) 출원인
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
백성욱
서울특별시 강남구 도곡로93길 12, 우성 2차 203동 602호 (대치동)
이정중
서울특별시 종로구 종로 347, 롯데캐슬 天동 504호 (숭인동)
(뒷면에 계속)
(74) 대리인
특허법인태백

전체 청구항 수 : 총 18 항

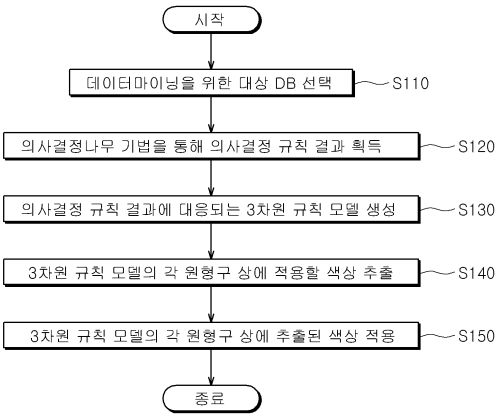
(54) 발명의 명칭 **OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법 및 시스템**

(57) 요 약

본 발명은, 데이터마이닝을 위한 대상 DB를 선택하는 단계와, 상기 대상 DB에 의사결정나무 기법을 적용하여 복수의 규칙 조건으로 상기 DB를 분류한 의사결정 규칙 결과를 얻는 단계, 및 상기 의사결정 규칙 결과에 대응되는 3차원 규칙 모델을 생성하는 단계를 포함하며, 상기 의사결정 규칙 결과는, 상기 분류에 따라 이진 트리 형의 나무 가지 모양으로 표현되고, 상기 나무 가지의 각 종단 노드는 특정 규칙 조건을 만족하는 대상 DB의 집합을 나타내는 개별 클래스로 구분되며, 상기 3차원 규칙 모델은, 상기 규칙 조건이 유사한 클래스끼리 유사 색상으로 표시되는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법 및 시스템을 제공한다.

본 발명에 따른 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법에 따르면, 의사결정나무 기법을 이용한 데이터마이닝 분석 결과를 3차원 규칙 모델로 효과적으로 시각화하여 표현함에 따라 일반 사용자로 하여금 분석 결과의 직관적인 파악이 가능하게 하고 데이터 접근의 편의성을 향상시키는 이점이 있다.

대 표 도 - 도1



(72) 발명자

김승규

서울특별시 성동구 성수2가 1동 376번지 2층

김성국

서울특별시 강동구 상암로47길 13 (명일동)

이범석

경기도 고양시 일산서구 일산로636번길 37-48, 102호 (대화동)

강경일

인천광역시 계양구 장기동 14블럭 대경이즈뷰 401호

이 발명을 지원한 국가연구개발사업

과제고유번호	JP090972M0214831
부처명	서울시
연구사업명	중소기업 제품개선 지원사업
연구과제명	분산형 OLAP 시스템에서의 eCube 시각화 기술 개발
주관기관	세종대학교
연구기간	2009.12.01 ~ 2010.11.30

특허청구의 범위

청구항 1

데이터마이닝을 위한 대상 DB를 선택하는 단계;

상기 대상 DB에 의사결정나무 기법을 적용하여 복수의 규칙 조건으로 상기 DB를 분류한 의사결정 규칙 결과를 얻는 단계; 및

상기 의사결정 규칙 결과에 대응되는 3차원 규칙 모델을 생성하는 단계를 포함하며,

상기 의사결정 규칙 결과는,

상기 분류에 따라 이진 트리 형의 나무 가지 모양으로 표현되고, 상기 나무 가지의 각 종단 노드는 특정 규칙 조건을 만족하는 대상 DB의 집합을 나타내는 개별 클래스로 구분되며,

상기 3차원 규칙 모델은,

상기 규칙 조건이 유사한 클래스끼리 유사 색상으로 표시되는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법.

청구항 2

청구항 1에 있어서,

상기 3차원 규칙 모델은 동일한 클래스들끼리 3차원 원형구 형태로 연결되어 표시되고,

상기 3차원 원형구들의 개별 크기는 해당 클래스들의 규칙 조건에 부합되는 레코드의 개수에 비례하여 결정되는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법.

청구항 3

청구항 2에 있어서,

YUV 기반의 색상 좌표계로부터 상기 3차원 규칙 모델의 각 원형구 상에 적용할 색상을 추출하는 단계; 및

상기 3차원 규칙 모델의 각 원형구 상에 상기 추출된 색상을 적용하는 단계를 더 포함하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법.

청구항 4

청구항 3에 있어서,

상기 색상을 추출하는 단계는,

상기 YUV 기반의 색상 좌표계의 원형 궤도의 범위를 설정하는 단계;

색상 구간의 설정을 위하여, 최소각도 및 최대각도를 초기화하는 단계;

상기 최소각도 및 최대각도를 이용하여, 상기 의사결정 규칙 결과를 구성하는 각 계층별 왼쪽 자식 노드와 오른쪽 자식 노드의 색상 구간을 각각 설정하는 단계;

해당 색상 구간 상의 각 자식 노드들의 위치에 대응되는 대표각도를 각각 산출하는 단계;

상기 원형 궤도의 범위를 이용하여, 상기 대표각도에 대응되는 상기 색상 좌표계 상의 좌표 정보를 상기 각 자식 노드들에 대해 각각 산출하는 단계; 및

상기 산출된 좌표에 해당되는 상기 각 자식 노드들의 색상을 각각 추출하는 단계를 포함하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법.

청구항 5

청구항 4에 있어서,

상기 왼쪽 자식 노드와 오른쪽 자식 노드에 대한 부모 노드의 색상은,

상기 왼쪽 자식 노드의 색상과 오른쪽 자식 노드의 색상의 중간값에 해당되는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법.

청구항 6

청구항 4에 있어서,

상기 왼쪽 자식 노드와 오른쪽 자식 노드의 색상 구간을 각각 설정하는 단계는 아래의 수학식을 이용하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법:

$$\begin{cases} L(left) = L \\ H(left) = \dfrac{L+H}{2} \end{cases}$$

$$\begin{cases} L(right) = \dfrac{L+H}{2} \\ H(right) = H \end{cases}$$

여기서, L 은 상기 최소각도, H 는 상기 최대각도이고, $L(left)$ 와 $H(left)$ 는 상기 왼쪽 자식 노드의 색상 구간의 설정을 위한 최소각도 및 최대각도이고, $L(right)$ 와 $H(right)$ 는 상기 오른쪽 자식 노드의 색상 구간의 설정을 위한 최소각도 및 최대각도이다.

청구항 7

청구항 6에 있어서,

상기 대표각도는 아래의 수학식을 이용하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법:

$$M(left) = \frac{L(left) + H(left)}{2}$$

$$M(right) = \frac{L(right) + H(right)}{2}$$

여기서, $M(left)$ 는 상기 왼쪽 자식 노드의 위치에 대응되는 대표각도이고, $M(right)$ 는 상기 오른쪽 자식 노드의 위치에 대응되는 대표각도이다.

청구항 8

청구항 7에 있어서,

상기 좌표 정보는 아래의 수학식을 이용하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법:

$$\begin{cases} U = r \times \cos\left(\frac{low + high}{2}\right) \\ V = r \times \sin\left(\frac{low + high}{2}\right) \end{cases}$$

여기서, r은 상기 원형 궤도의 범위이고, low와 $high$는 상기 왼쪽 자식 노드의 경우 $L(left)$와 $H(left)$ 값, 상기 오른쪽 자식 노드의 경우 $L(right)$와 $H(right)$ 값을 각각 나타낸다.

청구항 9

청구항 8에 있어서,

상기 원형의 색상 좌표계에서,

상기 노드의 길이에 따라 밝기정보가 변경되도록 상기 노드의 길이가 증가할수록 상기 색상의 종류가 더욱 세분화되고, 상기 r 값은 노드의 길이에 비례하여 증가하도록 설정된 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법.

청구항 10

청구항 1 내지 청구항 9 중 어느 한 항에 기재된 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

청구항 11

데이터마이닝을 위한 대상 DB를 선택하는 데이터선택부;

상기 대상 DB에 의사결정나무 기법을 적용하여 복수의 규칙 조건으로 상기 DB를 분류한 의사결정 규칙 결과를 얻는 결과획득부; 및

상기 의사결정 규칙 결과에 대응되는 3차원 규칙 모델을 생성하는 모델생성부를 포함하며,

상기 의사결정 규칙 결과는,

상기 분류에 따라 이진 트리 형의 나무 가지 모양으로 표현되고, 상기 나무 가지의 각 종단 노드는 특정 규칙 조건을 만족하는 대상 DB의 집합을 나타내는 개별 클래스로 구분되며,

상기 3차원 규칙 모델은,

상기 규칙 조건이 유사한 클래스끼리 유사 색상으로 표시되는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템.

청구항 12

청구항 11에 있어서,

상기 3차원 규칙 모델은 동일한 클래스들끼리 3차원 원형구 형태로 연결되어 표시되고,

상기 3차원 원형구들의 개별 크기는 해당 클래스들의 규칙 조건에 부합되는 레코드의 개수에 비례하여 결정되는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템.

청구항 13

청구항 12에 있어서,

YUV 기반의 색상 좌표계로부터 상기 3차원 규칙 모델의 각 원형구 상에 적용할 색상을 추출하는 색상추출부; 및

상기 3차원 규칙 모델의 각 원형구 상에 상기 추출된 색상을 적용하는 색상적용부를 더 포함하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템.

청구항 14

청구항 13에 있어서,

상기 색상추출부는,

상기 YUV 기반의 색상 좌표계의 원형 궤도의 범위를 설정한 다음, 색상 구간의 설정을 위하여, 최소각도 및 최대각도를 초기화하고,

상기 최소각도 및 최대각도를 이용하여, 상기 의사결정 규칙 결과를 구성하는 각 계층별 왼쪽 자식 노드와 오른쪽 자식 노드의 색상 구간을 각각 설정한 다음, 해당 색상 구간 상의 각 자식 노드들의 위치에 대응되는 대표각도를 각각 산출하고,

상기 원형 궤도의 범위를 이용하여, 상기 대표각도에 대응되는 상기 색상 좌표계 상의 좌표 정보를 상기 각 자식 노드들에 대해 각각 산출한 다음, 상기 산출된 좌표에 해당되는 상기 각 자식 노드들의 색상을 각각 추출하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템.

청구항 15

청구항 14에 있어서,

상기 색상추출부는,

상기 왼쪽 자식 노드와 오른쪽 자식 노드의 색상 구간의 설정 시, 아래의 수학식을 이용하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템:

$$\begin{cases} L(left) = L \\ H(left) = \dfrac{L+H}{2} \end{cases}$$

$$\begin{cases} L(right) = \dfrac{L+H}{2} \\ H(right) = H \end{cases}$$

여기서, L 은 상기 최소각도, H 는 상기 최대각도이고, $L(left)$ 와 $H(left)$ 는 상기 왼쪽 자식 노드의 색상 구간의 설정을 위한 최소각도 및 최대각도이고, $L(right)$ 와 $H(right)$ 는 상기 오른쪽 자식 노드의 색상 구간의 설정을 위한 최소각도 및 최대각도이다.

청구항 16

청구항 15에 있어서,

상기 색상추출부는,

상기 대표각도의 산출 시, 아래의 수학식을 이용하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템:

$$M(left) = \frac{L(left)+H(left)}{2}$$

$$M(right) = \frac{L(right)+H(right)}{2}$$

여기서, $M(left)$는 상기 왼쪽 자식 노드의 위치에 대응되는 대표각도이고, $M(right)$는 상기 오른쪽 자식 노드의 위치에 대응되는 대표각도이다.

청구항 17

청구항 16에 있어서,

상기 색상추출부는,

상기 좌표 정보의 산출 시 아래의 수학식을 이용하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템:

$$\begin{cases} U = r \times \cos\left(\frac{low + high}{2}\right) \\ V = r \times \sin\left(\frac{low + high}{2}\right) \end{cases}$$

여기서, r은 상기 원형 궤도의 범위이고, low와 $high$는 상기 왼쪽 자식 노드의 경우 $L(left)$와 $H(left)$ 값, 상기 오른쪽 자식 노드의 경우 $L(right)$와 $H(right)$ 값을 각각 나타낸다.

청구항 18

청구항 17에 있어서,

상기 원형의 색상 좌표계에서,

상기 노드의 길이에 따라 밝기정보가 변경되도록 상기 노드의 길이가 증가할수록 상기 색상의 종류가 더욱 세분화되고, 상기 r 값은 노드의 길이에 비례하여 증가하도록 설정된 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템.

명 세 서

기 술 분 야

[0001] 본 발명은 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법 및 시스템에 관한 것으로서, 보다 상세하게는 OLAP(Online Analytical Processing) 시스템에서의 의사결정나무 기법을 이용한 데이터마이닝 분석 결과를 3차원 규칙 모델로 효과적으로 시각화하는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법 및 시스템에 관한 것이다.

배 경 기 술

[0002] 일반적으로 OLAP(Online Analytical Processing)란 대용량의 다차원 데이터를 효율적으로 분석 처리하는 컴퓨터 소프트웨어를 의미하며, 이러한 OLAP는 데이터를 축적 가공하는 모든 분야(ex, 일반 비즈니스 분야, 제조 공정, 생명 공학)에 적용된다.

[0003] 한편 의사결정나무 기법은 데이터마이닝의 대표적인 분류기법으로서, 데이터베이스에 저장된 수많은 데이터들의 관계를 규칙 형태의 패턴으로 구분하여 이진 나무 구조로 표현한다. 이러한 의사결정나무 기법은 규칙에 의해 분류 구조가 명확하게 구분되고, 구조 표현에 소요되는 시간이 짧은 장점을 지니고 있어, 데이터마이닝 분석 작업에 주로 사용된다. 그러나, 데이터의 표본과 규칙의 복잡성에 따라 나무 구조 길이가 결정되므로, 연관 규칙이 많아질수록 데이터 간의 규칙 상관관계를 한눈에 파악하기 어렵다.

[0004] 종래의 데이터마이닝 분석 결과의 시각화 표현방법의 경우, 복잡한 텍스트, 도표 등으로 이루어져 있으며, 일반적으로 숙련된 분석가에 의해 해석되어져 왔다. 그 표현 방법으로는 분석 결과를 2차원 자율 곡선 그래프로 표현하는 방법과, 데이터마이닝에 필요한 모든 규칙조건을 콘 형태의 3차원 모델링으로 시각화하는 방법 등이 있다. 그런데, 이러한 종래의 경우는 데이터마이닝에 관해 전반적인 설명 및 지식 없이는 해석에 많은 어려움이

따르는 문제점이 있다.

발명의 내용

해결하려는 과제

[0005] 본 발명은, 의사결정나무 기법을 이용한 데이터마이닝 분석 결과를 3차원 규칙 모델로 효과적으로 시각화하여 표현함에 따라 분석 결과의 직관적인 파악이 가능하게 하고 데이터 접근의 편의성을 향상시키는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법 및 시스템을 제공한다.

과제의 해결 수단

[0006] 본 발명은, 데이터마이닝을 위한 대상 DB를 선택하는 단계와, 상기 대상 DB에 의사결정나무 기법을 적용하여 복수의 규칙 조건으로 상기 DB를 분류한 의사결정 규칙 결과를 얻는 단계, 및 상기 의사결정 규칙 결과에 대응되는 3차원 규칙 모델을 생성하는 단계를 포함하며, 상기 의사결정 규칙 결과는, 상기 분류에 따라 이진 트리 형의 나무 가지 모양으로 표현되고, 상기 나무 가지의 각 종단 노드는 특정 규칙 조건을 만족하는 대상 DB의 집합을 나타내는 개별 클래스로 구분되며, 상기 3차원 규칙 모델은, 상기 규칙 조건이 유사한 클래스끼리 유사 색상으로 표시되는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법을 제공한다.

[0007] 여기서, 상기 3차원 규칙 모델은 동일한 클래스들끼리 3차원 원형구 형태로 연결되어 표시되고, 상기 3차원 원형구들의 개별 크기는 해당 클래스들의 규칙 조건에 부합되는 레코드의 개수에 비례하여 결정될 수 있다.

[0008] 그리고, 상기 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법은, YUV 기반의 색상 좌표계로부터 상기 3차원 규칙 모델의 각 원형구 상에 적용할 색상을 추출하는 단계, 및 상기 3차원 규칙 모델의 각 원형구 상에 상기 추출된 색상을 적용하는 단계를 더 포함할 수 있다.

[0009] 여기서, 상기 색상을 추출하는 단계는, 상기 YUV 기반의 색상 좌표계의 원형 궤도의 범위를 설정하는 단계와, 색상 구간의 설정을 위하여, 최소각도 및 최대각도를 초기화하는 단계와, 상기 최소각도 및 최대각도를 이용하여, 상기 의사결정 규칙 결과를 구성하는 각 계층별 왼쪽 자식 노드와 오른쪽 자식 노드의 색상 구간을 각각 설정하는 단계와, 해당 색상 구간 상의 각 자식 노드들의 위치에 대응되는 대표각도를 각각 산출하는 단계와, 상기 원형 궤도의 범위를 이용하여, 상기 대표각도에 대응되는 상기 색상 좌표계 상의 좌표 정보를 상기 각 자식 노드들에 대해 각각 산출하는 단계, 및 상기 산출된 좌표에 해당되는 상기 각 자식 노드들의 색상을 각각 추출하는 단계를 포함할 수 있다.

[0010] 이때, 상기 왼쪽 자식 노드와 오른쪽 자식 노드에 대한 부모 노드의 색상은, 상기 왼쪽 자식 노드의 색상과 오른쪽 자식 노드의 색상의 중간값에 해당될 수 있다.

[0011] 또한, 상기 왼쪽 자식 노드와 오른쪽 자식 노드의 색상 구간을 각각 설정하는 단계는 아래의 수학식을 이용할 수 있다.

$$\begin{cases} L(left) = L \\ H(left) = \dfrac{L+H}{2} \end{cases}$$

$$\begin{cases} L(right) = \dfrac{L+H}{2} \\ H(right) = H \end{cases}$$

[0012]

[0013] 여기서, L 은 상기 최소각도, H 는 상기 최대각도이고, $L(left)$ 와 $H(left)$ 는 상기 왼쪽 자식 노드의 색상 구간의 설정을 위한 최소각도 및 최대각도이고, $L(right)$ 와 $H(right)$ 는 상기 오른쪽 자식 노드의 색상 구간의 설정을 위한 최소각도 및 최대각도이다.

[0014] 또한, 상기 대표각도는 아래의 수학식을 이용할 수 있다.

$$M(left) = \frac{L(left) + H(left)}{2}$$

$$M(right) = \frac{L(right) + H(right)}{2}$$

[0015]

[0016] 여기서, $M(left)$는 상기 왼쪽 자식 노드의 위치에 대응되는 대표각도이고, $M(right)$는 상기 오른쪽 자식 노드의 위치에 대응되는 대표각도이다.

[0017] 또한, 상기 좌표 정보는 아래의 수학식을 이용할 수 있다.

$$\begin{cases} U = r \times \cos\left(\frac{low + high}{2}\right) \\ V = r \times \sin\left(\frac{low + high}{2}\right) \end{cases}$$

[0018]

[0019] 여기서, r은 상기 원형 궤도의 범위이고, low와 $high$는 상기 왼쪽 자식 노드의 경우 $L(left)$와 $H(left)$ 값, 상기 오른쪽 자식 노드의 경우 $L(right)$와 $H(right)$ 값을 각각 나타낸다.

[0020] 또한, 상기 원형의 색상 좌표계에서, 상기 노드의 길이에 따라 밝기정보가 변경되도록 상기 노드의 길이가 증가할수록 상기 색상의 종류가 더욱 세분화되고, 상기 r 값은 노드의 길이에 비례하여 증가하도록 설정될 수 있다.

[0021] 그리고, 본 발명은 상기 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.

[0022] 또한, 본 발명은, 데이터마이닝을 위한 대상 DB를 선택하는 데이터선택부, 상기 대상 DB에 의사결정나무 기법을 적용하여 복수의 규칙 조건으로 상기 DB를 분류한 의사결정 규칙 결과를 얻는 결과획득부, 및 상기 의사결정 규칙 결과에 대응되는 3차원 규칙 모델을 생성하는 모델생성부를 포함하며, 상기 의사결정 규칙 결과는, 상기 분류에 따라 이진 트리 형의 나무 가지 모양으로 표현되고, 상기 나무 가지의 각 종단 노드는 특정 규칙 조건을 만족하는 대상 DB의 집합을 나타내는 개별 클래스로 구분되며, 상기 3차원 규칙 모델은, 상기 규칙 조건이 유사한 클래스끼리 유사 색상으로 표시되는 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템을 제공한다.

[0023] 또한, 상기 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템은, YUV 기반의 색상 좌표계로부터 상기 3차원 규칙 모델의 각 원형구 상에 적용할 색상을 추출하는 색상추출부, 및 상기 3차원 규칙 모델의 각 원형구 상에 상기 추출된 색상을 적용하는 색상적용부를 더 포함할 수 있다.

발명의 효과

[0024] 본 발명에 따른 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법에 따르면, 의사결정나무 기법을 이용한 데이터마이닝 분석 결과를 3차원 규칙 모델로 효과적으로 시각화하여 표현함에 따라 일반 사용자로 하여금 분석 결과의 직관적인 파악이 가능하게 하고 데이터 접근의 편의성을 향상시키는 이점이 있다.

도면의 간단한 설명

[0025] 도 1은 본 발명의 실시예에 따른 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법의 흐름도이다.

도 2는 도 1을 위한 시스템 구성도이다.

도 3은 도 1의 S110단계를 위한 대상 DB의 예시도이다.

도 4는 도 1의 S120단계로서, 도 3에 의사결정나무 기법을 적용한 결과의 예시도이다.

도 5는 도 4의 결과를 XML기반의 스크립트로 저장하는 예시도이다.

도 6은 도 1의 S130단계로서, 도 4의 결과에 대응되는 3차원 규칙 모델의 예시도이다.

도 7은 본 발명의 실시예에 따른 3차원 규칙 모델의 공생 및 경쟁 관계 표현의 예시도이다.

도 8은 도 1의 S140단계를 위한 YUV 기반의 색상표 예시도이다.

도 9는 도 1의 S140 단계를 위한 상세 흐름도이다.

도 10은 도 1의 S140단계에 따른 색상 추출 결과를 나타내는 예시도이다.

도 11은 도 10의 ①, ②, ③ 노드의 색상 추출 과정을 설명하는 개념도이다.

도 12는 본 발명의 실시예에 따라 색상 표현된 원형구들에 대한 규칙 조건의 유사도 관계를 나타낸 예시도이다.

발명을 실시하기 위한 구체적인 내용

[0026] 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다.

[0027] 본 발명은 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법에 관한 것으로서, 특히 대량의 데이터로부터 통계 및 수학적 분석 방법을 적용하여, 데이터들 사이의 관계, 성향, 패턴 등 다양하고 가치있는 정보를 사용자에게 효과적으로 제공한다. 이러한 본 발명은 의사결정을 지원하는 의사결정나무 기반 데이터마이닝 결과를 효과적으로 시각화하여 표현할 수 있다.

[0028] 도 1은 본 발명의 실시예에 따른 OLAP 시스템에서의 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법의 흐름도이다. 도 2는 도 1을 위한 시스템 구성도이다. 상기 시스템(100)은 데이터선택부(110), 결과획득부(120), 모델생성부(130), 색상추출부(140), 색상적용부(150)를 포함한다.

[0029] 이하에서는 도 1 및 도 2를 참조로 하여 상기 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법에 관하여 상세히 알아본다.

[0030] 먼저, 데이터마이닝을 위한 대상 DB를 상기 데이터선택부(110)를 통해 선택한다(S110). 상기 대상 DB는 다양한 분야의 분석 대상 DB에 해당될 수 있다.

[0031] 도 3은 도 1의 S110단계를 위한 대상 DB의 예시도이다.

[0032] 이러한 도 3은 개인ID 정보, 성별, 직업, 거주지, 결혼 여부, 연봉, 취미, 종교, 총계좌수, 여신잔액, 수신잔액 등을 주요 데이터로 하는 대상 DB를 그 예로 한 것이다. 물론, 본 발명이 반드시 이러한 예로 한정되는 것은 아니다.

[0033] 여기서, 복수의 분석 항목(개인ID 정보, 성별, 직업, 거주지, 결혼 여부, 연봉, 취미, 종교, 총계좌수, 여신잔액, 수신잔액)들 중에서 분석 대상이 되는 항목을 클래스로 결정한다. 예를 들어, 거주지를 기본으로 하여 데이터마이닝을 하고자 하는 경우, 구로구를 클래스 A, 마포구를 클래스 B, 강서구를 클래스 C, 강남구를 클래스 D 등으로 설정하도록 한다.

[0034] 상기 S110 단계 이후, 상기 대상 DB에 의사결정나무 기법을 적용하여 복수의 규칙 조건으로 상기 DB를 분류한 의사결정 규칙 결과를 얻는다(S120). 이러한 S120단계는 결과획득부(120)를 통해 수행한다.

[0035] 도 4는 도 1의 S120단계로서, 도 3에 의사결정나무 기법을 적용한 결과의 예시도이다. 이는 일련의 규칙 패턴들로 이루어진 의사결정나무 분석 결과이다. 이때, 사용된 규칙 조건으로는 연봉, 수신잔액, 여신잔액, 직업, 거주지, 계좌수, 종교유무, 취미 등이 있다.

[0036] 이로부터, 대용량의 대상 DB를 의사결정나무 기반의 데이터마이닝 기법으로 분석하여 얻어지는 상기 의사결정 규칙 결과는, 상기 규칙 조건에 의한 분류에 따라 이진 트리(Binary Tree) 형의 나무 가지 모양으로 표현됨을 알 수 있다.

[0037] 또한, 상기 나무 가지의 각 종단 노드는 특정 규칙 조건을 만족하는 대상 DB의 집합을 나타내는 개별 클래스(class)로 구분됨을 알 수 있다. 다시 말해서, 상기 의사결정나무의 각 종단 노드는 일련의 규칙 조건 집합이라

볼 수 있고, 이는 클래스로 구분된다.

[0038] 도 5는 도 4의 결과를 XML기반의 스크립트로 저장하는 예시도이다. 이렇게 클래스로 구분된 규칙 조건 집합들은 효율적인 데이터 관리와 구조적 범용성이 높은 특성을 지닌 XML기반의 스크립트로 저장되며, 고유 번호와 애트리뷰트 이름, 인스턴스 값, 비중수치로 구성된다.

[0039] 상기 S120단계 이후에는, 상기 의사결정 규칙 결과에 대응되는 3차원 규칙 모델을 상기 모델생성부(130)에서 생성한다(S130).

[0040] 도 6은 도 1의 S130단계로서, 도 4의 결과에 대응되는 3차원 규칙 모델의 예시도이다. 여기서, 상기 3차원 규칙 모델은 동일한 클래스들끼리 군집되어 있는 3차원 원형구 형태로 연결되어 표시된다. 이때, 각각의 클래스는, 동일한 클래스들끼리 하나의 줄로 연결된 원형구들의 집합 형태로 표현된다. 또한, 상기 3차원 규칙 모델은 상기 규칙 조건이 유사한 클래스끼리 유사 색상으로 표시된다.

[0041] 상기 3차원 원형구들의 개별 크기는 해당 클래스들의 규칙 조건에 부합되는 레코드의 개수에 비례하여 결정된다. 도 6에서 원형구의 직경 크기가 클수록 해당 레코드의 개수가 많음을 의미한다. 이때, 상기 레코드 개수에 비례하는 비중수치를 부가하여 상기 원형구들의 개별 크기(ex, 직경)를 결정할 수 있다.

[0042] 도 7은 본 발명의 실시예에 따른 3차원 규칙 모델의 공생 및 경쟁 관계 표현의 예시도이다. 즉, 의사결정나무기반 규칙 데이터로 생성된 상기 3차원 규칙 모델은 특정한 방향성을 띄도록 각각의 클래스에 대하여 생태계 기반의 생명체 형태인 3차원 시각화 공생관계 형태 또는 3차원 시각화 경쟁관계 형태로 표현될 수 있다.

[0043] 이러한 도 7은 분석된 데이터에 생명체와 같은 역동적인 표현을 부가하여 데이터 분석자 즉, 사용자들에게 다차원/대용량 데이터들 사이의 생동감 있는 인터랙션을 느낄 수 있게 해준다. 이러한 생명체의 모습은 특정 방향성을 가지고 그룹을 짓는 철새들의 모습과 유사하며, 큰 머리와 가는 꼬리로 물속을 헤엄치는 올챙이 모습과 같은 복합적인 모양새로 표현된다.

[0044] 상기 S130단계 이후에는, YUV 기반의 색상 좌표계로부터 상기 3차원 규칙 모델의 각 원형구 상에 적용할 색상을 상기 색상추출부(140)를 통해 추출한다(S140). 이러한 S140 단계는 3차원 규칙 모델로 시각화된 원형구의 규칙 데이터 유사 정도를 측정하고 추후 이를 색상으로 표현하기 위함이다.

[0045] 도 8은 도 1의 S140단계를 위한 YUV 기반의 색상표 예시도이다. 상단 왼쪽은 YUV 방식의 색상표의 예이며, 상단 오른쪽은 왼쪽의 색상표에서 밝기정보(Y)와, 원형궤도(r)를 표현한 좌표계이다. 본 발명에서는 색차정보를 표현하는 방식 중 하나인 YUV 기반의 색상 좌표계를 사용하여 색상을 추출한다. 이러한 YUV 기반의 색상 좌표계는 원형의 궤도를 따라 색상 구간들로 나누어져 있다.

[0046] 상기 YUV 방식은 원색 신호에 대한 정보를 가지고 있는 RGB에 비해 색상 품질 저하가 낮은 이점이 있다. 인간의 눈은 컬러정보에 비해 밝기정보에 민감하기 때문에, 본 실시예에서는 밝기 정보(Y)와, 컬러 정보(U, V)를 분리하여 표현이 가능한 YUV방식을 선택한다. 본 발명에서는 의사결정나무의 노드 길이(도 8의 아래 그림 참조)에 따라 밝기정보(Y) 값이 변경(더욱 세분화)되며, YUV기반 색상좌표계의 각 원형 궤도의 크기(r)를 결정한 뒤 이를 색 구간으로 설정하였다. 이는 상기 색상의 조합을 증가시킬 수 있다.

[0047] 상기 S140단계에 따르면, 각 종단 노드까지 계층화되는 의사결정나무의 특성에 따라 각 계층별로 자식 노드들의 색상 구간을 분리하되, 상기 원형의 궤도를 따라 상기 규칙 조건이 유사한 클래스끼리 유사 색상으로 추출하도록 한다. 이러한 원리는 도 8의 아래 그림 즉, 색상표현이 적용된 의사결정나무 그림에서도 확인이 되며, 의사결정나무의 각 클래스는 상기 원형 궤도 내에서 색상이 추출되었음을 알 수 있다.

[0048] 또한, 도 8을 통해, 의사결정나무에서 왼쪽 자식 노드와 오른쪽 자식 노드에 대한 '부모 노드'의 색상은 상기 왼쪽 자식 노드의 색상과 오른쪽 자식 노드의 색상의 중간값에 해당됨을 알 수 있다.

[0049] 상기 S140단계 이후에는, 상기 3차원 규칙 모델의 각 원형구 상에 상기 추출된 색상을 적용한다(S150). 이는 상기 색상적용부(150)를 통해 수행한다. 이러한 S150 단계에 따라, 각 원형구의 색상은 다양한 색상을 표출하되, 유사 클래스 끼리는 유사 색상으로 표현된다.

[0050] 이하에서는, 상기 색상을 추출하는 단계에 관하여 보다 상세히 알아보도록 한다. 도 9는 도 1의 S140 단계를 위한 상세 흐름도이다. 먼저, 상기 YUV 기반의 색상 좌표계의 원형 궤도의 범위를 설정한다.(S910). 본 실시예에서는 YUV기반 색상 좌표계의 밝기 정보(Y)의 경우 의사결정노드의 왼쪽 자식노드는 증가시키고, 오른쪽 자식노드는 감소시켰다. 그리고, 원형 궤도의 범위(r)는 원점(r=0)에서 시작하여 최대 0.5의 범위로 증가하되, 노드의

길이가 하나씩 늘어날 때마다 "0.5/총 노드의 길이" 만큼 r 값이 점차 증가하도록 설정하였다.

[0051] 상기 S910단계 이후, 색상 구간의 설정을 위하여, 최소각도 및 최대각도를 초기화한다(S920). 본 실시예에서 최소각도 L 은 90° , 최대각도 H 는 450° 로 초기화하였다.

[0052] 다음, 상기 최소각도 및 최대각도를 이용하여, 상기 의사결정 규칙 결과를 구성하는 각 계층별 왼쪽 자식 노드와 오른쪽 자식 노드의 색상 구간을 각각 설정한다(S930). 이러한 S930 단계는 수학식 1을 이용한다.

수학식 1

$$\begin{cases} L(left) = L \\ H(left) = \dfrac{L+H}{2} \end{cases}$$

$$\begin{cases} L(right) = \dfrac{L+H}{2} \\ H(right) = H \end{cases}$$

[0053]

[0054] 여기서, L 은 상기 최소각도, H 는 상기 최대각도이다.

[0055] 그리고, $L(left)$ 와 $H(left)$ 는 상기 왼쪽 자식 노드의 색상 구간의 설정을 위한 최소각도 및 최대각도이다. 또한, $L(right)$ 와 $H(right)$ 는 상기 오른쪽 자식 노드의 색상 구간의 설정을 위한 최소각도 및 최대각도이다.

[0056] 이후, 해당 색상 구간 상의 각 자식 노드들의 위치에 대응되는 대표각도를 각각 산출한다(S940). 상기 대표 각도의 산출은 수학식 2를 이용한다.

수학식 2

$$M(left) = \frac{L(left)+H(left)}{2}$$

$$M(right) = \frac{L(right)+H(right)}{2}$$

[0057]

[0058] 여기서, $M(left)$ 는 상기 왼쪽 자식 노드의 위치에 대응되는 대표각도이고, $M(right)$ 는 상기 오른쪽 자식 노드의 위치에 대응되는 대표각도이다.

[0059] 이후, 상기 원형 궤도의 범위를 이용하여, 상기 대표각도에 대응되는 상기 색상 좌표계 상의 좌표 정보를 상기 각 자식 노드들에 대해 각각 산출한다(S950). 상기 좌표 정보의 산출은 수학식 3을 사용한다.

수학식 3

[0060]
$$\begin{cases} U = r \times \cos\left(\frac{low + high}{2}\right) \\ V = r \times \sin\left(\frac{low + high}{2}\right) \end{cases}$$

[0061] 여기서, r은 상기 원형 궤도의 범위이고, low와 $high$는 상기 왼쪽 자식 노드의 경우 $L(left)$와 $H(left)$ 값, 상기 오른쪽 자식 노드의 경우 $L(right)$와 $H(right)$ 값을 각각 나타낸다. 이때, 상기 원형 궤도의 범위는 상기 원형 좌표계의 반지름 값에 대응된다. r은 최대 0.5의 값을 갖는다. 또한, 상기 노드의 길이 하나씩 증가할 때마다 r 값이 0에서 최대 0.5까지 증가한다. 만약, 데이터 마이닝된 결과, 총 노드의 길이가 7개인 경우에는, 노드의 길이가 하나씩 증가할 때마다 r 값은 0.07(=0.5/7) 만큼 증가한 값을 수학식 3에 적용한다. 즉, 노드의 길이가 0인 최초 부모 노드의 경우 r=0(따라서 (0,0)의 좌표에 위치), 다음의 자식 노드에 대해서는 노드의 길이가 1이므로 r=0.07, 그 다음의 자식 노드에 대해서는 노드의 길이가 2이므로 r=0.14 값이 적용된다.

[0062] 다음, 상기 산출된 좌표에 해당되는 상기 각 자식 노드들의 색상을 각각 추출한다(S960). 이에 따라, 상기 의사결정나무 내의 전체 노드들의 개별 색상은 YUV기반 색상 좌표계의 원형 궤도 안에 위치하는 각각의 좌표값을 통해 결정된다.

[0063] 도 10은 도 1의 S140단계에 따른 색상 추출 결과를 나타내는 예시도이다. 시작 노드부터 단말 노드까지 내려오는 일련의 규칙조건들은 각각 하나의 클래스로 정의된다. 각 노드들은 앞서 도 9의 과정을 통해 도 10에 도시된 YUV 기반 색상 좌표계의 해당 구간에서 각각 색상이 추출되어 개별 색상으로 할당된다.

[0064] 각각의 규칙에 따라 처음 두 개의 자식노드 ①과 ②번 노드가 존재하며, 이 ①과 ②번 자식 노드는 앞서 도 9의 과정을 통해 도 10에 도시된 YUV 기반 색상 좌표계의 해당 구간에서 각각 색상이 추출되어 개별 색상으로 할당된다.

[0065] 그리고, ①번 노드는 각 규칙에 따라 ③과 ④번 자식 노드로 분류되어 있다. 이때, ①번 노드는 ③과 ④번 노드에 대한 부모 노드에 해당된다. ③과 ④번 자식 노드 또한 도 9의 과정을 통해 도 10에 도시된 YUV 기반 색상 좌표계의 해당 구간에서 각각 색상이 추출되어 개별 색상으로 할당된다.

[0066] 즉, 도 10의 색상을 적용한 의사결정나무의 예제 그림에서 볼 수 있듯이, 일련의 규칙에 의해 종단 노드에서 클래스A, 클래스B, 클래스C 등으로 나뉘고, 이는 오른쪽에 표시한 대로 각각 일련의 규칙정보를 포함하고 있다.

[0067] 도 10의 좌측 아래 그림을 참조하면, 이 원형의 색상 좌표계는, 그 밝기정보(Y)가 상기 노드의 길이에 따라 변경되도록, 상기 노드의 길이가 증가할수록 상기 색상의 종류가 더욱 세분화된 것을 알 수 있다. 이때, 앞서 설명한 바와 같이 상기 r 값은 노드의 길이에 비례하여 증가한다.

[0068] 이러한 도 10의 원형 좌표계는 노드의 총 길이(depth)가 7일 때 YUV좌표계의 원형궤도를 표현된 그림이다. 즉, 좌표계의 중심에서 외곽을 향해 궤도가 7단계로 구분되어 있다(이 원형 좌표계의 2와 4번 표시는 도 10의 상단에 있는 노드(②,④)를 의미하는 것이 아니라, 노드의 길이(Depth) 별 대응되는 좌표계 적용 범위 부분을 표시한 것이다). 이는 도 8의 오른쪽 그림보다 노드의 총 길이가 작은 경우로서 설명의 편의를 위하여 더욱 간소화한 예를 표현하기 위함이다. 참고로 도 10의 상단에 있는 의사결정나무는 노드의 길이가 4인 부분 까지를 간단히 도시한 것이다.

[0069] 도 10의 하단 오른쪽 그림은 왼쪽의 원형 궤도 내에서 실제 색상을 추출하기위해 의사결정나무의 왼쪽 자식노드와 오른쪽 자식노드의 밝기정보(Y) 값을 서로 일정 비율로 가감하여 표현한 것이다. 즉, 앞서 설명한 바와 같이, 부모 노드를 기준으로 왼쪽 자식노드는 밝기값이 점점 증가되고, 오른쪽 자식노드는 밝기값이 점점 감소

된다. 이는 도 10의 상단 의사 결정나무의 각 노드(①,②,③ 등)와 하단 오른쪽 그림의 해당 노드(①,②,③ 등)의 할당된 색상 정보를 비교함으로써 알 수 있다. 그리고, 노드의 길이(depth)가 늘어날수록 왼쪽 자식노드는 밝기정보(Y)가 미세하게 증가하며, 오른쪽 자식노드는 밝기정보(Y)가 미세하게 감소한다.

[0070] 도 11은 도 10의 ①, ②, ③ 노드의 색상 추출 과정을 설명하는 개념도이다. 이를 앞서 수학식 및 도 10과 연계하여 설명하면 다음과 같다.

[0071] 먼저, 도 10에서 최초의 왼쪽 자식 노드 ①의 색상 추출 과정에 대해 알아본다. 이때, 도 11의 시작 부분은 도 10의 Yearly Income 부분에 대응되는 것으로서, 노드의 길이는 0인 부분이며, r=0의 값을 갖는다. 우선, 초기화된 최소각도 L = 90° , 최대각도 H = 450° 를 이용하여, 상기 왼쪽 자식 노드 ①에 대한 $L(left)$값과 $H(left)$ 값을 수학식 1을 통해 우선 산출한다(S930). $L(left)$는 90° , $H(left)$는 270° 로 산출되고, 이에 따라 상기 왼쪽 자식 노드 ①에 대한 색상 구간은 상기 색상좌표계 상에서 90° ~ 270° 범위로 설정된다.

[0072] 다음, 상기 구하여진 $L(left)$값과 $H(left)$ 값을 수학식 2에 대입한 결과, $M(left)$값은 180° 로 산출된다(S940). 이는 곧 상기 왼쪽 자식 노드 ①의 위치에 대응되는 대표각도에 해당된다.

[0073] 그런 다음, 상기 구하여진 $L(left)$와 $H(left)$ 값을 수학식 3에 대입하되, 상기 왼쪽 자식 노드 ①은 노드의 길이(Depth)가 1인 경우므로 r=0.07을 적용하면, 상기 왼쪽 자식 노드 ①의 좌표 정보(U,V)가 (-0.07,0)으로 산출된다(S950).

[0074] 이에 따라, 도 10의 색상 좌표계 상에서 (U,V)=(-0.07,0)의 위치 해당되는 색상(①번 위치의 색상)을 추출하고(S960), 이를 상기 노드 ①에 대한 원형구 색상으로 적용한다.

[0075] 다음, 도 10에서 최초의 오른쪽 자식 노드 ②의 색상 추출 과정에 대해 알아본다. 초기화된 최소각도 L = 90° , 최대각도 H = 450° 를 이용하여, 상기 오른쪽 자식 노드 ②에 대한 $L(right)$값과 $H(right)$값을 수학식 1을 통해 산출한다(S930). $L(right)$는 270° , $H(right)$는 450° 로 산출되고, 이에 따라 상기 오른쪽 자식 노드 ②에 대한 색상 구간은 상기 색상좌표계 상에서 270° ~ 450° 범위로 설정된다.

[0076] 다음, 상기 구하여진 $L(right)$값과 $H(right)$ 값을 수학식 2에 대입한 결과, $M(right)$값은 360° 로 산출된다(S940). 이는 곧 상기 오른쪽 자식 노드 ②의 위치에 대응되는 대표각도에 해당된다.

[0077] 그런 다음, 상기 구하여진 $L(right)$값과 $H(right)$ 값을 수학식 3에 대입하되, 상기 오른쪽 자식 노드 ② 또한 노드의 길이(Depth)가 1인 경우므로 r=0.07을 적용하면, 상기 오른쪽 자식 노드 ②의 좌표 정보(U,V)가 (0.07,0)으로 산출된다(S950).

[0078] 이에 따라, 도 10의 색상 좌표계 상에서 (U,V)=(0.07,0)의 위치 해당되는 색상(②번 위치의 색상)을 추출하고(S960), 이를 상기 노드 ②에 대한 원형구 색상으로 적용한다.

[0079] 물론, ①번에 대한 왼쪽 자식 노드 ③번에 대해서도 상술한 방법과 동일한 방식이 이용된다. 이때, 상기 수학식 1에 적용되는 최소각도 L과 최대각도 H는 ①번 노드에 대하여 산출되었던 최소각도와 최대각도 값 즉, 90° 와 180° 를 각각 사용한다.

[0080] 이러한 도 11의 결과는 도 10의 하단 오른쪽 결과와 동일한 결과를 도출한다.

[0081] 이러한 방식을 통해, 의사결정나무의 각 종단 노드에 대응되는 모든 3차원 원형구들의 색상은 상기 색상 좌표계의 색상 구간 내에서 추출되며, 의사결정나무 구조의 복잡도에 따라 원형 궤도 범위가 결정된다. 이에 따라, 최소각도와 최대각도의 사이 각도 즉, 해당 노드에 대해 결정되는 색상 구간은 의사결정나무의 아래 방향으로 갈수록 좁아지게 된다. 물론, 상기한 일련의 과정은 최후의 종단 노드를 발견할 때까지, 색상구간의 분리 과정 내지 원형구에 색상을 적용하는 과정을 재귀적으로 연산한다.

[0082] 상기과 같은 원리의 과정을 반복함에 따라, 도 10의 결과물 즉, 각각 노드에 대한 좌표를 각각 산출할 수 있고, 이로부터 색상을 각각 추출하여 3차원 규칙 모델 즉, 각각의 원형구에 색상을 적용할 수 있게 된다.

[0083] 참고로, 도 10에 도시된 의사결정나무를 보면, 적용 규칙의 분류가 상이하더라도 클래스의 종류는 동일할 수 있으며, 또한 클래스의 종류가 다르더라도 적용된 색상은 유사할 수 있다.

[0084] 도 12는 본 발명의 실시예에 따라 색상 표현된 원형구들에 대한 규칙 조건의 유사도 관계를 나타낸 예시도이다. 도 12는 클래스 A,B,C에 대한 3차원 규칙 모델을 나타내는 것으로서, 3차원 규칙 모델의 개별 색상은 의사결정나무의 클래스에 해당되는 종단 노드의 색상을 나타낸다.

[0085] 이를 참조하면, 특정 클래스 내의 원형구 색상은 각 규칙 조건 사이의 유사정도를 나타내며, 이는 다른 클래스에 해당하는 규칙조건들 중 일부분과 유사하게 표현될 수 있다.

[0086] 도 12에서 (1)과 (2)의 원형구는 모두 클래스 A로서 클래스의 종류는 동일하지만 색상은 달리 할당되어 있다. 이는 규칙 간의 유사도가 떨어짐을 의미한다. 이와는 달리, (3)과 (4)의 원형구는 클래스는 C와 A로 상이하지만, 규칙 조건 간의 유사도가 높아서 거의 유사한 색상으로 표현된다.

[0087] 이상과 같은 본 발명에 따르면, 데이터마이닝에 의해 생성된 생태계 기반 3차원 모델을 규칙 조건의 유사도에 따라 다양한 색상으로 표현함으로써, 클래스 간의 상관관계를 보다 명확하게 드러낼 수 있다. 이는 단순한 색상의 분리 표현이 아니라 노드 간의 규칙 조건의 유사도에 비례하여 색상이 중복 없이 변화하는 특징을 갖는다는 점에서 주목할 만하다.

[0088] 동일 클래스의 종단 노드임에도 규칙조건의 유사도가 떨어질 경우에는 색상이 상이하게 달라질 수 있고, 서로 다른 클래스에 속한 종단 노드임에도 규칙 조건의 유사도가 클 경우는 비슷한 색상으로 표현되어 사용자에게 색상 구분을 통한 직관적인 이해를 돕는다.

[0089] 즉, 본 발명은 데이터마이닝 결과의 표현에 있어서 대량의 데이터 및 정보를 2차원에 제한하지 않고 가시적 효과가 우수한 3차원 형태로 시각화하여 효율적이고 적절한 의사결정을 지원한다. 뿐만 아니라, 분석 결과에 대한 접근성을 향상시켜서, 숙련된 해석 기술을 보유하지 않은 일반 사용자에게도 분석 및 활용의 폭을 넓힐 수 있게 한다.

[0090] 이상과 같은 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 방법은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 매체로서 데이터가 저장되는 모든 종류의 기록장치를 포함한다.

[0091] 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CO-ROM, 자기 테이프, 플로피디스크, 광데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

[0092] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 다른 실시예가 가능한 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호범위는 첨부된 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

부호의 설명

[0093] 100: 의사결정나무 기법을 이용한 3차원 규칙 모델 생성 시스템

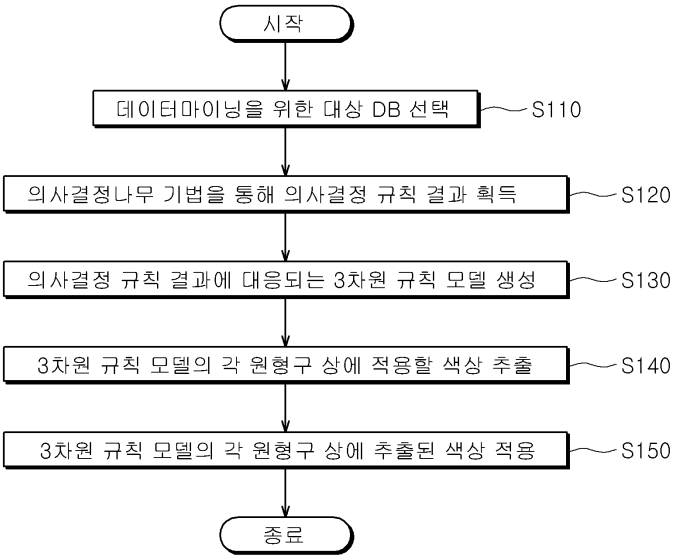
110: 데이터선택부 120: 결과획득부

130: 모델생성부 140: 색상추출부

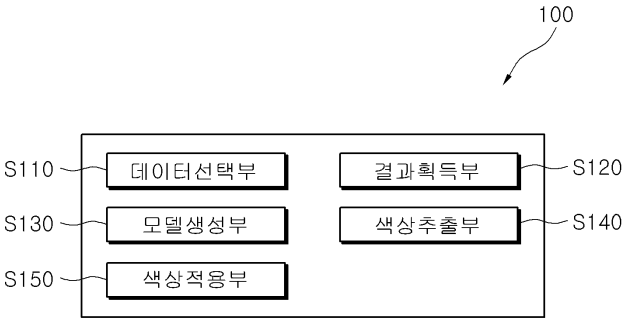
150: 색상적용부

도면

도면1



도면2



도면3

Table A

ID	성별	직업	거주지	결혼여부	연봉
321084	남	교육	구로구	Y	4900
322037	남	컴퓨터	마포구	N	2600
322427	남	자영업	강서구	N	2500
325416	여	서비스	강남구	N	3700
320708	남	교육	성동구	Y	3000
323372	여	서비스	송파구	N	4500
321056	남	영업	송파구	Y	3600
320764	남	자영업	광진구	Y	2700
320068	여	교육	동작구	Y	2700
323105	남	교육	은평구	N	3100
...					

Table B

CUST_ID	취미	종교여부	총계좌수	여신잔액	수신잔액
322427	여행	무	1	6400	9700
321056	독서	무	2	2900	9200
321084	여행	무	3	2700	12100
320068	사진	유	3	1600	9100
320708	자동차	무	2	7200	8000
323105	영화	유	5	4600	10200
325416	골프	무	3	3400	8000
322037	독서	무	4	4500	8400
323372	영화	유	4	2600	6100
320764	골프	유	3	2400	8100
...					

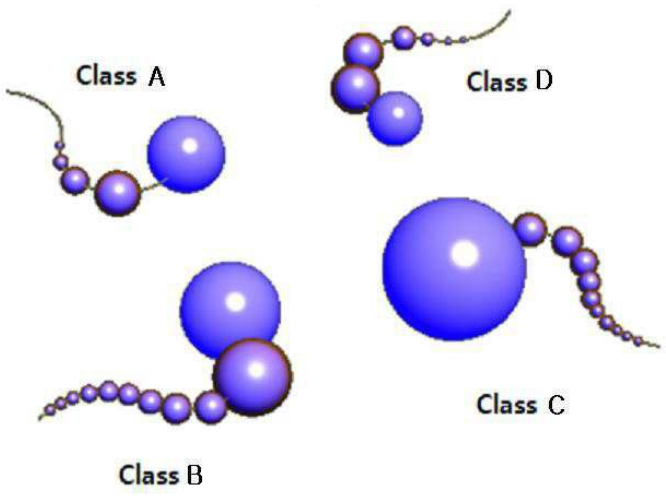
도면4



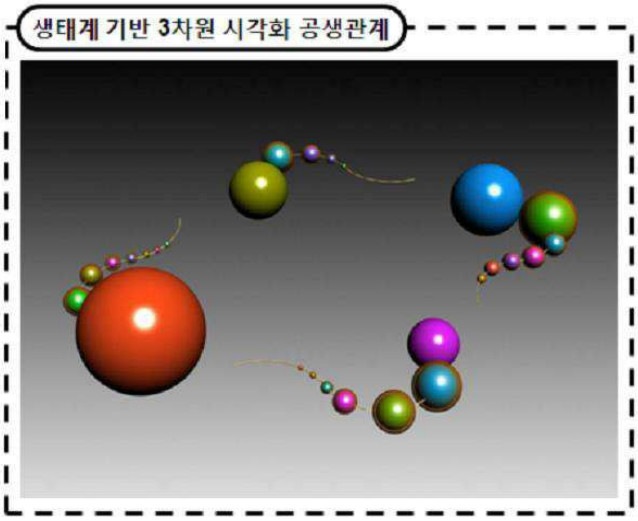
도면5

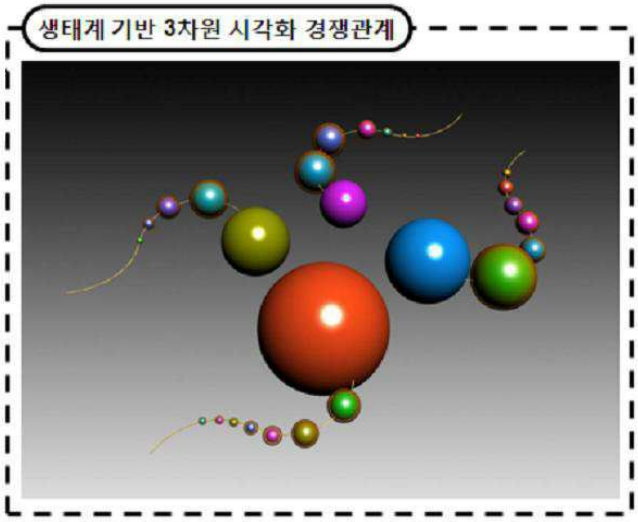
```
<?xml version="1.0" encoding="euc-kr" ?>
<rules total_rules = "38" num_of_classes = "4">
  <class ID = "0" attr_name = "CRD_GRADE" value = "A" num_of_rules = "5">
    <rule ID = "21" content = "2300.00<=YEARLY_INCOME<4400.00 & SELF_DP_BAL<8000.00 & SELF_LN_BAL<3400.00 & JOB IS Product & ACCNT<6.00"
    weight = "0.03">
    <rule ID = "27" content = " YEARLY_INCOME<4400.00 & SELF_DP_BAL>=8000.00 & JOB IS NOT Education & SELF_LN_BAL>=3000.00 & RLGN_CD IS N
    & HOBY_CD IS Golf" weight = "0.03">
    <rule ID = "29" content = "YEARLY_INCOME<4400.00 & SELF_DP_BAL>=8000.00 & JOB IS Education & SELF_LN_BAL<3000.00" weight = "0.21">
    <rule ID = "32" content = "YEARLY_INCOME>=4400.00 & SELF_LN_BAL<2000.00" weight = "0.77">
  </class>
  <class ID = 1 attr_name = "CRD_GRADE" value = "B" num_of_rules = "6">
    <rule ID = "4" content = "YEARLY_INCOME<2300.00 & SELF_DP_BAL<8000.00 & REGION IS NOT Jongro & JOB IS Service & SELF_LN_BAL<1500.00 &
    REGION IS NOT Gangseo" weight = "0.06">
    <rule ID = "8" content = "YEARLY_INCOME<2300.00 & SELF_DP_BAL<8000.00 & REGION IS Jongro" weight = "0.32">
    <rule ID = "25" content = "YEARLY_INCOME<4400.00 & SELF_DP_BAL>=8000.00 & JOB IS NOT Education & SELF_LN_BAL<3000.00 & ACCNT>=4.00"
    weight = "0.56">
    <rule ID = "28" content = "YEARLY_INCOME<4400.00 & SELF_DP_BAL>=8000.00 & JOB IS NOT Education & SELF_LN_BAL>=3000.00 & RLGN_CD IS Y"
    weight = "0.04">
    <rule ID = "31" content = "YEARLY_INCOME<4400.00 & SELF_DP_BAL>=8000.00 & JOB IS Education & SELF_LN_BAL>=3000.00 & RLGN_CD IS Y"
    weight = "0.04">
    <rule ID = "33" content = "YEARLY_INCOME>=4400.00 & SELF_LN_BAL>=2000.00 & REGION IS NOT Songpa" weight = "0.26">
  </class>
                                                    ...
```

도면6

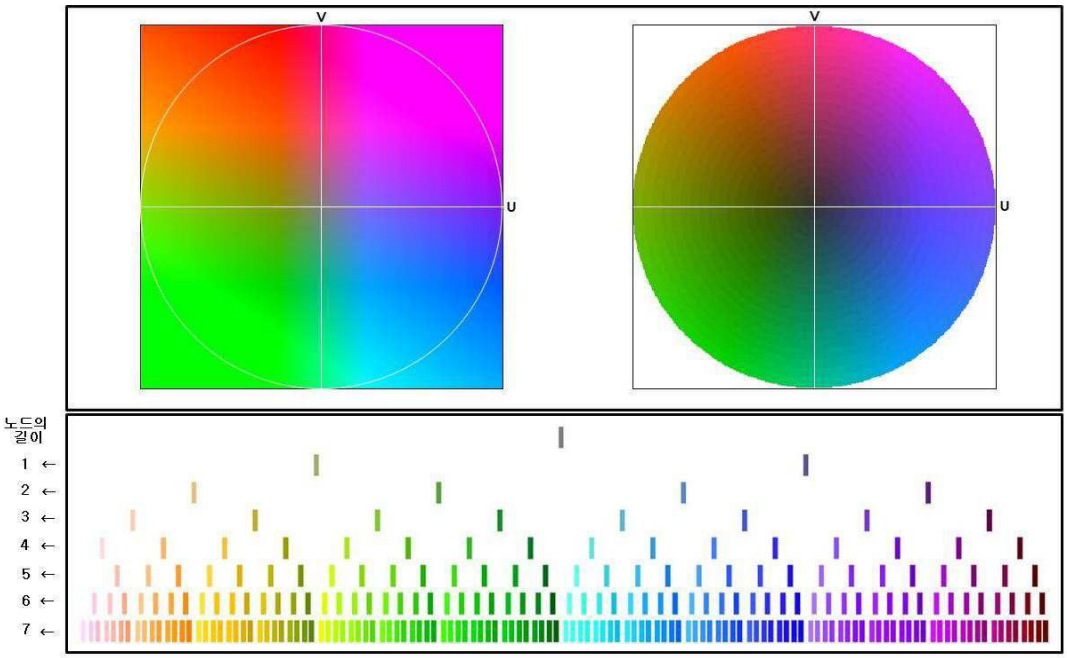


도면7

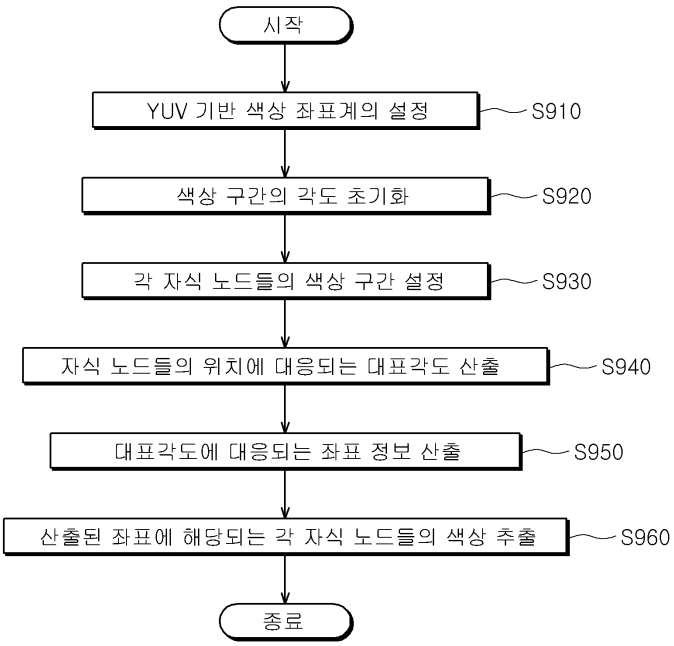




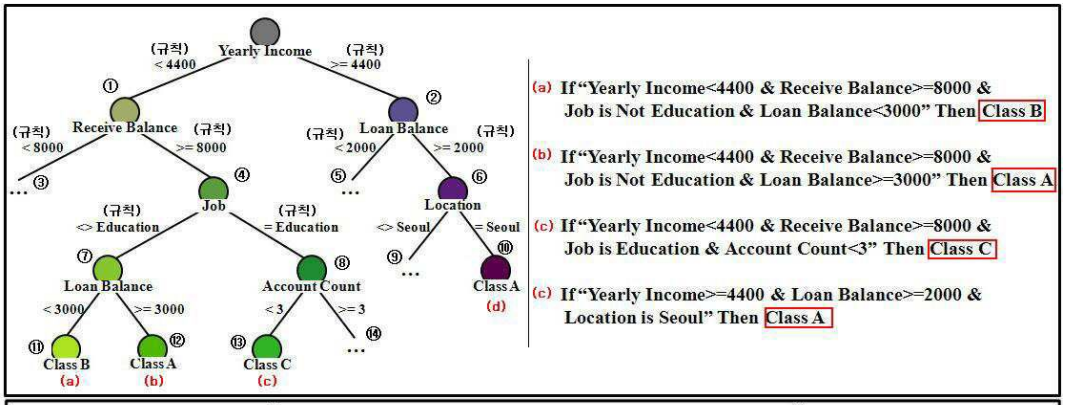
도면8

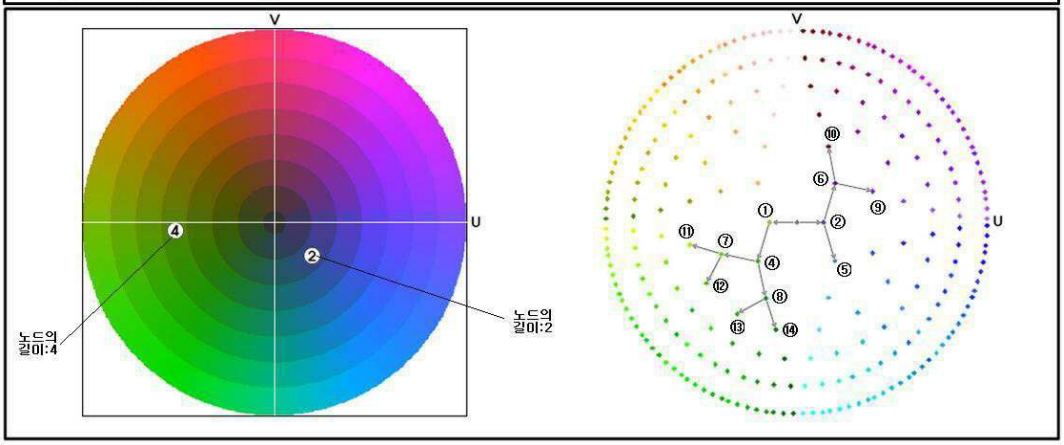

V
U
V
U
노드의 깊이
1
2
3
4
5
6
7

도면9

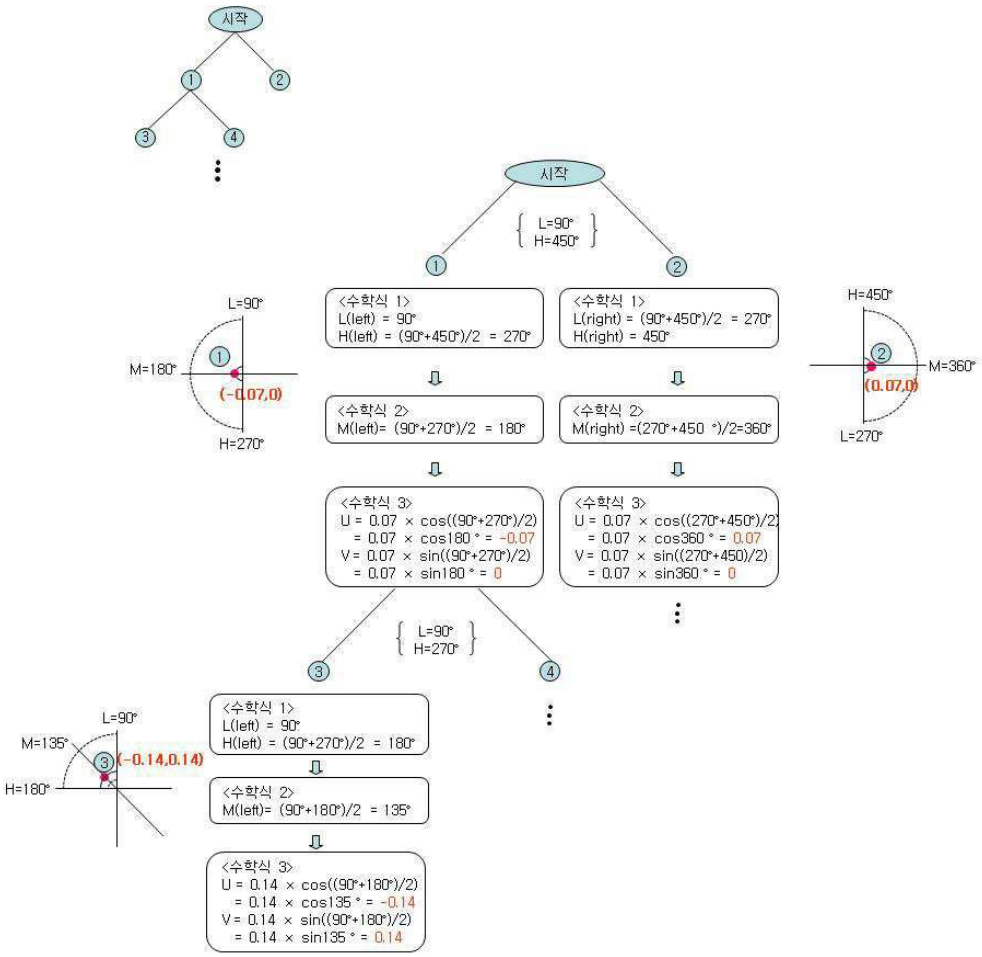

시작
YUV 기반 색상 좌표계의 설정
S910
색상 구간의 각도 초기화
S920
각 자식 노드들의 색상 구간 설정
S930
자식 노드들의 위치에 대응되는 대표각도 산출
S940
대표각도에 대응되는 좌표 정보 산출
S950
산출된 좌표에 해당되는 각 자식 노드들의 색상 추출
S960
종료

도면10

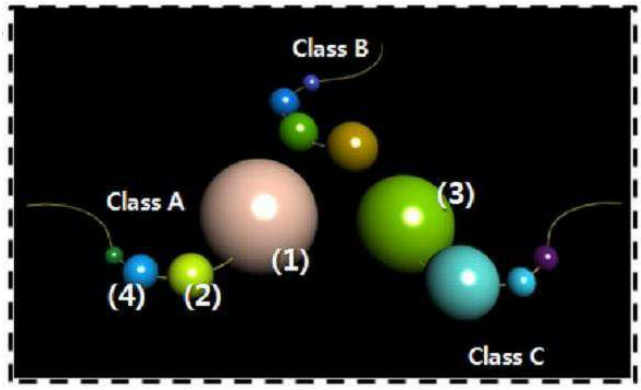

(규칙) < 4400
Yearly Income
(규칙) >= 4400
①
②
(규칙) < 8000
Receive Balance
(규칙) >= 8000
(규칙) < 2000
Loan Balance
(규칙) >= 2000
③
④
⑤
⑥
(규칙) <> Education
Job
(규칙) = Education
<> Seoul
Location
= Seoul
⑦
⑧
⑨
⑩
Loan Balance
Account Count
Class A
(d)
< 3000
>= 3000
< 3
>= 3
⑪
⑫
⑬
⑭
Class B
(a)
Class A
(b)
Class C
(c)
(a) If "Yearly Income<4400 & Receive Balance>=8000 & Job is Not Education & Loan Balance<3000" Then Class B
(b) If "Yearly Income<4400 & Receive Balance>=8000 & Job is Not Education & Loan Balance>=3000" Then Class A
(c) If "Yearly Income<4400 & Receive Balance>=8000 & Job is Education & Account Count<3" Then Class C
(c) If "Yearly Income>=4400 & Loan Balance>=2000 & Location is Seoul" Then Class A

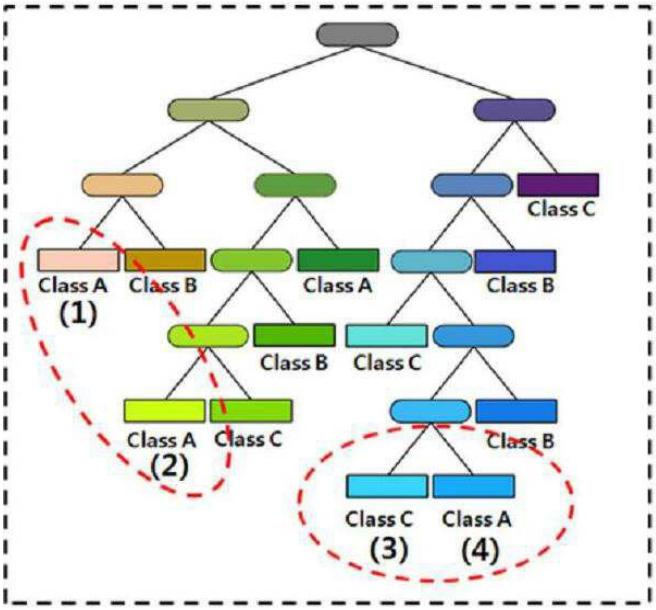

V
U
4
2
노드의 길이:4
노드의 길이:2
V
U

도면11


시작
1
2
3
4
시작
L=90°
H=450°
1
2
<수학식 1>
L(left) = 90°
H(left) = (90°+450°)/2 = 270°
<수학식 1>
L(right) = (90°+450°)/2 = 270°
H(right) = 450°
L=90°
M=180°
H=270°
(-0.07,0)
H=450°
M=360°
L=270°
(0.07,0)
<수학식 2>
M(left)= (90°+270°)/2 = 180°
<수학식 2>
M(right) =(270°+450 °)/2=360°
<수학식 3>
U = 0.07 × cos((90°+270°)/2)
= 0.07 × cos180 ° = -0.07
V = 0.07 × sin((90°+270°)/2)
= 0.07 × sin180 ° = 0
<수학식 3>
U = 0.07 × cos((270°+450°)/2)
= 0.07 × cos360 ° = 0.07
V = 0.07 × sin((270°+450)/2)
= 0.07 × sin360 ° = 0
L=90°
H=270°
3
4
L=90°
M=135°
H=180°
3
(-0.14,0.14)
<수학식 1>
L(left) = 90°
H(left) = (90°+270°)/2 = 180°
<수학식 2>
M(left)= (90°+180°)/2 = 135°
<수학식 3>
U = 0.14 × cos((90°+180°)/2)
= 0.14 × cos135 ° = -0.14
V = 0.14 × sin((90°+180°)/2)
= 0.14 × sin135 ° = 0.14

도면12


Class B
Class A
(1)
(2)
(3)
(4)
Class C


Class C
Class A
Class B
(1)
Class A
Class B
Class B
Class C
Class A
Class C
(2)
Class B
Class C
Class A
(3)
(4)