



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년12월06일
(11) 등록번호 10-2053243
(24) 등록일자 2019년12월02일

(51) 국제특허분류(Int. Cl.)
G06F 16/00 (2019.01) G06F 17/27 (2006.01)
G06N 99/00 (2019.01)
(52) CPC특허분류
G06F 16/00 (2019.01)
G06F 17/2785 (2013.01)
(21) 출원번호 10-2017-0019873
(22) 출원일자 2017년02월14일
심사청구일자 2017년11월01일
(65) 공개번호 10-2018-0015561
(43) 공개일자 2018년02월13일
(30) 우선권주장
1020160098926 2016년08월03일 대한민국(KR)
(56) 선행기술조사문헌
KR101551574 B1
KR101542195 B1
KR1020080096859 A
KR1020150095459 A

(73) 특허권자
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
(72) 발명자
김말희
대전광역시 유성구 노은서로210번길 32, 410동
502호 (지족동, 열매마을4단지)
강현중
경상남도 진주시 수곡면 원창길 116
(뒷면에 계속)
(74) 대리인
특허법인지명

전체 청구항 수 : 총 16 항

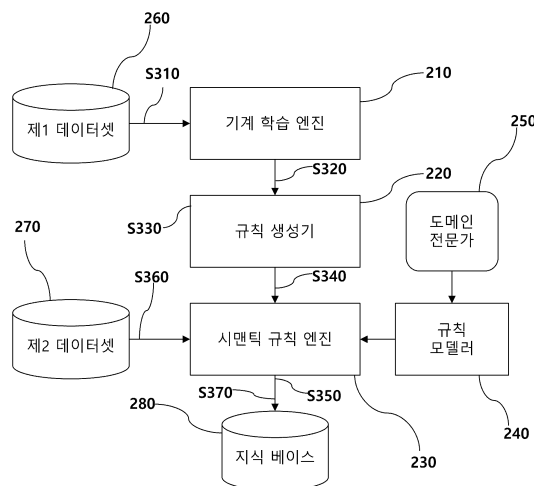
심사관 : 김경완

(54) 발명의 명칭 적응형 지식 베이스 구축 방법 및 그 시스템

(57) 요약

본 발명의 일면에 따른 적응형 지식 베이스 구축 장치는 입력된 제1 데이터 집합을 토대로 학습 모델을 생성하고, 생성된 학습 모델에 따라 제1 데이터 집합에 대한 연관 관계를 분석하는 기계 학습 엔진; 학습된 모델 및 알고리즘을 토대로 규칙을 생성하는 규칙 생성기; 및 생성된 규칙과 제2 데이터 집합을 이용하여 시맨틱 규칙을 생성하는 시맨틱 규칙 생성기;를 포함한다.

대표도 - 도2



(52) CPC특허분류

G06N 20/00 (2019.01)

(72) 발명자

권순현

인천광역시 서구 검암로 53, 105동 204호 (검암동, 풍림아이원2차)

이용준

대전광역시 유성구 가정로 65 (신성동, 대림두레아파트)

김귀훈

대전광역시 서구 만년로 25, 105동 1002호 (만년동, 강변아파트)

표철식

대전광역시 서구 만년로 25, 109동 701호 (만년동, 강변아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호 CRC-15-05-ETRI

부처명 미래창조과학부

연구관리전문기관 국가과학기술연구회

연구사업명 국가과학기술연구회연구운영비지원

연구과제명 자가학습형 지식융합 슈퍼브레인 핵심기술 개발

기 여 율 1/1

주관기관 한국전자통신연구원

연구기간 2015.12.01 ~ 2016.11.30

명세서

청구범위

청구항 1

입력된 제1 데이터 집합을 토대로 학습 모델을 생성하고, 생성된 학습 모델에 따라 제1 데이터 집합에 대한 연관 관계를 분석하는 기계 학습 엔진;

학습된 모델 및 알고리즘을 토대로 규칙을 생성하는 규칙 생성기; 및

생성된 규칙과 제2 데이터 집합을 이용하여 시맨틱 규칙을 생성하는 시맨틱 규칙 생성기;

를 포함하는 적응형 지식 베이스 구축 시스템.

청구항 2

제1항에 있어서,

상기 시맨틱 규칙 생성기는,

트리 기반 규칙 생성 알고리즘, apriori 알고리즘 중 어느 하나의 방법에 의하여 규칙을 생성하는 것인 적응형 지식 베이스 구축 시스템.

청구항 3

제1항에 있어서,

도메인 전문가에 의하여 시맨틱 규칙 생성 방법을 변경하는 규칙 모델러를 더 포함하는 것인 적응형 지식 베이스 구축 시스템.

청구항 4

제1항에 있어서,

상기 시맨틱 규칙 생성기는,

임시로 저장된 기계 학습 규칙, 상기 제2 데이터셋, 규칙 모델을 이용하여 주기적으로 지식 베이스를 확장하는 것

인 적응형 지식 베이스 구축 시스템.

청구항 5

적응형 지식 베이스 구축 시스템에 있어서,

적응형 지식 베이스 구축 모듈을 제공하기 위한 프로그램이 저장된 메모리; 및

상기 프로그램을 실행시키는 프로세서를 포함하되,

상기 프로세서는 상기 프로그램을 실행시킴에 따라,

입력된 제1 데이터셋에 대한 학습 모델을 생성하고, 생성된 학습 모델에 따라 제1 데이터셋의 연관관계 분석 결과를 출력하고, 학습된 모델과, 학습된 알고리즘과 연관관계 분석 결과를 토대로 기계 학습 규칙을 생성하고,

생성된 기계 학습 규칙을 이용하여 시맨틱 규칙을 생성하는 것을 특징으로 하는 적응형 지식 베이스 구축 시스템.

청구항 6

제5항에 있어서,
상기 프로세서는,
RDF, OWL 중 어느 하나의 방법으로 생성된 규칙과 제2 데이터 집합을 이용하여 시맨틱 규칙을 생성하는 것인 적응형 지식 베이스 구축 시스템.

청구항 7

제5항에 있어서,
상기 프로세서는,
제1 데이터셋과 제2 데이터셋에 대한 학습 모델 및 추론 모델을 생성하는 것을 특징으로 하는 적응형 지식 베이스 구축 시스템.

청구항 8

제5항에 있어서,
상기 프로세서는,
도메인 전문가에 의하여 입력된 규칙 모델을 이용하여 시맨틱 규칙을 생성하는 것을 특징으로 하는 적응형 지식 베이스 구축 시스템.

청구항 9

(1) 기계 학습 엔진이 입력된 제1 데이터 집합을 기계 학습하는 단계;
(2) 규칙 생성기가 학습된 결과를 토대로 규칙을 생성하는 단계;
(3) 규칙 생성기가 생성된 규칙과 제2 데이터 집합을 이용하여 시맨틱 규칙을 생성하는 단계; 및
(4) 시맨틱 규칙 엔진이 생성된 시맨틱 규칙을 저장하여 지식 베이스를 구축하는 단계;
를 포함하는 적응형 지식 베이스 구축 방법.

청구항 10

제9항에 있어서,
규칙 모델러가 상기 시맨틱 규칙을 생성하기 위하여, 도메인 전문가에 의하여 규칙을 생성하는 모델을 제공하는 단계;
를 더 포함하는 적응형 지식 베이스 구축 방법.

청구항 11

제9항에 있어서,

상기 시맨틱 규칙을 생성하는 단계는,

상기 규칙 생성기가 RDF, OWL 중 어느 하나의 방법으로 생성된 규칙과 제2 데이터 집합을 이용하여 시맨틱 규칙을 생성하는 것

인 적응형 지식 베이스 구축 방법.

청구항 12

제9항에 있어서,

상기 시맨틱 규칙을 생성하는 단계는,

상기 규칙 생성기가 트리 기반 규칙 생성 알고리즘, apriori 알고리즘 중 어느 하나의 방법에 의하여 규칙을 생성하는 것

인 적응형 지식 베이스 구축 방법.

청구항 13

제12항에 있어서,

상기 apriori 알고리즘은,

상기 규칙 생성기가 지지도에 따라, 각 노드의 분선 단계를 가지치기하는 것을 특징으로 하는

적응형 지식 베이스 구축 방법.

청구항 14

기계 학습 엔진과 시맨틱 규칙 엔진을 이용하여 지식 베이스를 구축하는 적응형 지식 베이스 구축 방법에 있어서,

제2 데이터셋에 대한 시맨틱 추론 모델을 생성하는 단계;

제2 데이터셋에 생성된 시맨틱 추론 모델을 적용한 추론을 하고 예측 결과를 생성하는 단계;

상기 추론된 예측 결과로부터 기계 학습 규칙을 생성하는 단계;

기계 학습 규칙을 시맨틱 규칙으로 변환하는 단계; 및

변환된 시맨틱 규칙을 저장하여 지식 베이스를 확장하는 단계;

를 포함하는 적응형 지식 베이스 구축 방법.

청구항 15

제14항에 있어서,

상기 기계 학습 규칙을 시맨틱 규칙으로 변환하는 단계는,

도메인 전문가에 의하여 제공되는 규칙 모델에 따라 기계 학습 규칙을 시맨틱 규칙으로 변환하는 방법이 결정되는 것

인 적응형 지식 베이스 구축 방법.

청구항 16

제14항에 있어서,
 상기 기계 학습 규칙을 시맨틱 규칙으로 변환하는 단계는,
 트리 구조 모델 또는 apriori 알고리즘 중 어느 하나의 방법에 의하여 시맨틱 규칙을 생성하는 것
 인 적응형 지식 베이스 구축 방법.

발명의 설명

기술 분야

[0001] 본 발명은 적응형 지식 베이스 구축 방법 및 그 시스템에 관한 것으로, 구체적으로 기계 학습을 통하여 생성된 학습 결과를 규칙으로 변환하고, 규칙으로 변환된 지식을 시맨틱 기술을 이용하여 지식 베이스를 구축하는 적응형 지식 베이스 구축 방법 및 그 시스템에 관한 것이다.

배경 기술

[0003] 최근 기계 학습 및 시맨틱 기술에 대한 연구가 활발하게 진행되고 있다. 기계 학습은 주어진 목적에 따라 데이터 기반 학습을 진행하고, 학습된 결과를 토대로 새로운 환경에 필요한 정보를 예측하는 기술이다. 통상적으로 학습 방법은 트리 기반 분석 방법과 연관 기반 분석 방법으로 분류될 수 있다. 트리 기반 분석 방법은 학습 결과 구축된 트리의 노드 정보를 이용하여 규칙(rule)을 생성하고, 연관 기반 분석 방법은 학습 데이터의 패턴을 분석하여 연관 규칙(rule)을 생성한다.

[0004] 시맨틱 기술은 도메인 지식을 가진 설계자에 의하여 규칙(rule)을 생성하고, 생성된 규칙(rule)을 이용하여, 지식을 확장-추론-재활용하여 지식 베이스를 구축하는 기술을 말한다.

[0005] 기계 학습을 이용하여 규칙을 생성하고, 규칙을 이용하여 요청에 따른 결과를 생성할 때, 생성 주기가 있어, 학습 결과를 재활용하지 못하는 문제가 있다. 시맨틱 기술의 경우, 도메인 지식을 갖는 사람이 개입하여 규칙을 변경하기 전에는 사전에 형성된 규칙 기반 지식을 토대로 확장 및 추론의 결과로서 지식 베이스가 구축되는데, 주변 환경에 의하여 실제 환경이 동적으로 변할 수 있어, 주변 환경에 따라 적응적(adaptive)으로 규칙을 변경하고, 변경된 규칙이 반영된 지식 베이스를 구축할 필요가 있다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 전술한 문제를 해결하기 위하여, 적응형 지식 베이스 구축 방법 및 그 시스템을 제공하는 것을 그 목적으로 한다. 구체적으로 본 발명은 기계 학습 및 시맨틱 기술을 이용하여 사람의 개입 없이 동적으로 변화하는 환경을 반영하는 적응형 지식 베이스를 구축하는 방법 및 그 시스템을 제공하는 것이다.

[0008] 본 발명의 목적은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 또 다른 목적들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0010] 전술한 목적을 달성하기 위한 본 발명의 일면에 따른 적응형 지식 베이스 구축 장치는 입력된 제1 데이터 집합을 토대로 학습 모델을 생성하고, 생성된 학습 모델에 따라 제1 데이터 집합에 대한 연관 관계를 분석하는 기계 학습 엔진; 학습된 모델 및 알고리즘을 토대로 규칙을 생성하는 규칙 생성기; 및 생성된 규칙과 제2 데이터 집합을 이용하여 시맨틱 규칙을 생성하는 시맨틱 규칙 생성기;를 포함한다.

[0011] 본 발명의 다른 일면에 따른 적응형 지식 베이스 구축 시스템은, 적응형 지식 베이스 구축 모듈을 제공하기 위한 프로그램이 저장된 메모리; 및 상기 프로그램을 실행시키는 프로세서를 포함하되, 상기 프로세서는 상기 프로그램을 실행시킴에 따라, 입력된 제1 데이터셋에 대한 학습 모델을 생성하고, 생성된 학습 모델에 따라 제1 데이터셋의 연관관계 분석 결과를 출력하고, 학습된 모델과, 학습된 알고리즘과 연관관계 분석 결과를 토대로

기계 학습 규칙을 생성하고, 생성된 기계 학습 규칙을 이용하여 시맨틱 규칙을 생성하는 것을 특징으로 한다.

[0012] 본 발명의 다른 일면에 따른 적응형 지식 베이스 구축 방법은 (1) 입력된 제1 데이터 집합을 기계 학습하는 단계; (2) 학습된 결과를 토대로 규칙을 생성하는 단계; (3) 생성된 규칙과 제2 데이터 집합을 이용하여 시맨틱 규칙을 생성하는 단계; 및 (4) 생성된 시맨틱 규칙을 저장하여 지식 베이스를 구축하는 단계;를 포함한다.

발명의 효과

[0014] 본 발명에 따르면, 기계 학습 기술과 시맨틱 기술을 결합하여 지식 베이스를 구축하여 인간의 개입을 배제하고, 최적의 분석 및 높은 효율성을 달성하는 효과가 있다.

[0015] 본 발명은 IoT 기술, 빅데이터 기술과 연관한 분석, 상황 인식 서비스와 관련된 지능적 서비스 산업 분야에 적용할 수 있고, 기계 학습을 이용한 분석 기술 분야에서 플랫폼으로 활용될 수 있다.

도면의 간단한 설명

[0017] 도 1은 본 발명에 따른 적응형 지식 베이스 구축 방법이 구현되는 컴퓨터 시스템의 구성을 설명하기 위한 예시도.

도 2는 본 발명의 일실시예에 따른 적응형 지식 베이스 구축 시스템의 구성도.

도 3은 본 발명의 다른 실시예에 따른 적응형 지식 베이스 구축 시스템의 구성도.

도 4는 본 발명의 부분 실시예에 따른 트리 학습 모델 기반 적응형 지식 베이스 구축 방법을 설명하기 위한 예시도.

도 5는 본 발명의 부분 실시예에 따른 Apriori 알고리즘을 적용한 적응형 지식 베이스 구축 방법을 설명하기 위한 예시도.

도 6은 본 발명의 다른 실시예에 따른 적응형 지식 베이스 구축 시스템의 구성도.

도 7은 본 발명의 다른 실시예에 따른 적응형 지식 베이스 구축 시스템의 구성도.

도 8은 본 발명의 다른 실시예에 따른 적응형 지식 베이스 구축 방법의 절차 흐름도.

발명을 실시하기 위한 구체적인 내용

[0018] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 한편, 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성요소, 단계, 동작 및/또는 소자는 하나 이상의 다른 구성요소, 단계, 동작 및/또는 소자의 존재 또는 추가를 배제하지 않는다.

[0020] 이하, 본 발명의 바람직한 실시예에 대하여 첨부한 도면을 참조하여 상세히 설명하기로 한다.

[0022] 도 1은 본 발명에 따른 적응형 지식 베이스 구축 방법이 구현되는 컴퓨터 시스템의 구성을 설명하기 위한 예시도이다.

[0023] 한편, 본 발명의 실시예에 따른 적응형 지식 베이스 구축 방법은 컴퓨터 시스템에서 구현되거나, 또는 기록매체에 기록될 수 있다. 도 1에 도시된 바와 같이, 컴퓨터 시스템은 적어도 하나 이상의 프로세서(110)와, 메모리(120)와, 사용자 입력 장치(150)와, 데이터 통신 버스(130)와, 사용자 출력 장치(160)와, 저장소(140)를 포함할 수 있다. 전술한 각각의 구성 요소는 데이터 통신 버스(130)를 통해 데이터 통신을 한다.

[0024] 컴퓨터 시스템은 네트워크(180)에 연결된 네트워크 인터페이스(170)를 더 포함할 수 있다. 상기 프로세서(110)는 중앙처리 장치(central processing unit (CPU))이거나, 혹은 메모리(120) 및/또는 저장소(140)에 저장된 명령어를 처리하는 반도체 장치일 수 있다.

- [0025] 상기 메모리(120) 및 상기 저장소(140)는 다양한 형태의 휘발성 혹은 비휘발성 저장매체를 포함할 수 있다. 예컨대, 상기 메모리(120)는 ROM(123) 및 RAM(126)을 포함할 수 있다.
- [0026] 따라서, 본 발명의 실시예에 따른 적응형 지식 베이스 구축 방법은 컴퓨터에서 실행 가능한 방법으로 구현될 수 있다. 본 발명의 실시예에 따른 적응형 지식 베이스 구축 방법이 컴퓨터 장치에서 수행될 때, 컴퓨터로 판독 가능한 명령어들이 본 발명에 따른 운영 방법을 수행할 수 있다.
- [0027] 한편, 상술한 본 발명에 따른 적응형 지식 베이스 구축 방법은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현되는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록 매체로는 컴퓨터 시스템에 의하여 해독될 수 있는 데이터가 저장된 모든 종류의 기록 매체를 포함한다. 예를 들어, ROM(Read Only Memory), RAM(Random Access Memory), 자기 테이프, 자기 디스크, 플래시 메모리, 광 데이터 저장장치 등이 있을 수 있다. 또한, 컴퓨터로 판독 가능한 기록매체는 컴퓨터 통신망으로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 읽을 수 있는 코드로서 저장되고 실행될 수 있다.
- [0029] 도 2는 본 발명의 일 실시예에 따른 적응형 지식 베이스 구축 시스템의 구성도를 나타낸다.
- [0030] 본 발명에 따른 적응형 지식 베이스 구축 시스템은 기계 학습 엔진; 규칙 생성기; 시맨틱 규칙 엔진; 지식 베이스;를 포함한다. 바람직하게는, 도메인 전문가의 규칙 모델러를 더 포함할 수 있다.
- [0031] 본 발명에 따른 적응형 지식 베이스 구축 시스템은 시맨틱 모델링에 기반을 둔 지식 베이스 구축 환경과 기계 학습에 기반을 두어 학습된 규칙(rule) 기반을 두어 지식 베이스 구축 환경이 병행하여 운영하거나, 각 지식 베이스 구축 환경이 개별적으로 운영될 수 있다. 시맨틱 모델링에 기반을 둔 경우와, 기계 학습에 기반을 둔 경우는 기능적으로 분리될 수 있는 것으로, 전체 시스템을 고려하여, 분산 시스템으로 구축될 수 있다.
- [0032] 기계 학습 엔진(210)은 제1 데이터셋(260)에 대한 최적의 학습 모델을 생성하고, 생성된 학습 모델을 이용하여 제1 데이터셋을 학습하고, 학습된 결과를 토대로 데이터셋에 대한 연관관계 분석 결과를 학습된 모델 결과로서 출력한다.
- [0033] 규칙 생성기(220)는 출력된 학습된 모델 결과를 이용하여 학습된 모델 및 학습된 알고리즘을 규칙화하여 규칙을 생성하고, 분석된 데이터 연관 관계에 대한 규칙도 생성한다. 생성된 규칙을 시맨틱 규칙 엔진(230)에 전달하거나, 생성된 규칙으로부터 시맨틱 규칙을 생성하여 시맨틱 규칙 엔진(230)에 전달하거나, 생성된 규칙으로부터 비정형 데이터를 생성하여 시맨틱 규칙 엔진(230)에 전달할 수 있다. 통상적으로 규칙 생성기(220)는 일반화된 규칙을 생성한다. 다만, 시맨틱 규칙 엔진의 API를 규칙 생성기(220)에 적용하여 직접 시맨틱 규칙을 생성할 수 있고, 이때, 시맨틱 규칙 엔진(230)의 역할을 생성한 시맨틱 규칙을 지식 베이스에 저장할지 여부 및 기존의 시맨틱 규칙에 확장 적용 여부를 결정하게 된다.
- [0034] 본 발명에 따른 적응형 지식 베이스 구축 시스템은 분산 처리 시스템으로 구축될 수 있고, 규칙 생성기(220) 및 시맨틱 규칙 엔진(230)은 분산 처리 시스템 구성 방법에 따라 별도의 서버에 설치될 수 있다. 규칙 생성기(220)에서 시맨틱 규칙을 생성하는지는 통상 각 서버의 로드와 따라 적절하게 선택할 수 있다.
- [0035] 시맨틱 규칙 엔진(230)은 확장 엔진(inference engine)이라고도 하며, 시맨틱 규칙을 확장하는 역할을 한다. 시맨틱 규칙 엔진(230)은 규칙 생성기(220)로부터 전달받은 규칙 및 비정형 데이터를 토대로 시맨틱 규칙을 생성하고, 생성된 시맨틱 규칙을 지식 베이스에 저장한다. 또한, 시맨틱 규칙 엔진(230)은 규칙 생성기(220)로부터 전달받은 시맨틱 규칙을 지식 베이스에 저장한 것인지 저장하지 않을지를 결정한다.
- [0036] 또한, 시맨틱 규칙 엔진(230)은 제2 데이터셋(270)을 입력받고, 구축된 지식 베이스에 저장된 시맨틱 규칙을 확장할 수 있다. 제2 데이터셋(270)은 수작업으로 입력된 새로운 규칙 및 새로운 시맨틱 규칙을 포함할 수 있다.
- [0038] 도 3은 본 발명의 다른 실시예에 따른 적응형 지식 베이스 구축 시스템의 구성도를 나타낸다.
- [0039] 본 발명에 따른 적응형 지식 베이스 구축 방법은 (1) 제1 데이터셋에 대한 학습 모델을 생성하는 단계(S310); (2) 생성된 학습 모델을 이용하여 학습된 모델 결과를 생성 하는 단계; (3) 학습된 모델 결과를 이용하여 규칙을 생성하는 단계(S330); (4) 생성된 규칙을 이용하여 시맨틱 규칙을 생성하는 단계(S340); (5) 생성된 시맨틱 규칙을 저장하여 지식 베이스를 구축하는 단계(S350);를 포함한다.
- [0040] 상기 (1) 단계에서는 기계 학습 엔진(210)이 제1 데이터 집합을 이용하여 학습하여 최적의 학습 모델을 생성한다.
- [0041] 상기 (2) 단계에서는 기계 학습 엔진(210)은 최적의 분석 모델을 학습하고, 추론 단계에서 입력으로 제공되는

제1 데이터 집합에 대한 연관 관계 분석 결과를 출력한다.

- [0042] 상기 (3) 단계에서는 규칙 생성기(220)가 학습된 모델과 학습된 알고리즘을 규칙화하고, 학습된 모델과 학습된 알고리즘을 이용하여 분석된 데이터의 연관 관계도 규칙화한다.
- [0043] 상기 (4) 단계에서는 상기 규칙 생성기(220)에 의해 생성된 일반적인 규칙을 시맨틱 규칙으로 변환한다.
- [0044] 상기 (5) 단계에서는 변환된 시맨틱 규칙을 기존에 저장된 시맨틱 규칙과 병합하여, 지식 베이스를 확장한다.
- [0045] 상기 (1) 단계에서의 기계 학습을 이용하여 규칙을 생성하는 방법은 알고리즘에 따라 상이하다. 본 발명은 기계 학습을 통하여 규칙 생성에 필요한 지식을 획득하여, 획득된 지식도 규칙으로 다시 생성하고, 시맨틱 기술을 적용하여 최종적으로 시맨틱 규칙으로 변환하여 통합하는 방법을 제공한다. 상기 (1) 단계에서 사용되는 기계 학습 방법은 도 4 및 도 5에서 다시 설명하도록 한다. 다만, 이는 예시적인 것이며, 다른 알고리즘 또는 모델에 의하여 규칙을 생성할 수 있고, 발명의 범위를 제한하고자 하는 것이 아니다.
- [0047] 도 4는 본 발명의 부분 실시예에 따른 트리 학습 모델 기반 적응형 지식 베이스 구축 방법을 설명하기 위한 예시도이다.
- [0048] 도 4(a)은 트리 기반의 규칙 생성 방법을 설명하기 위한 예시도이다.
- [0049] 도 4(a)는 입력 데이터로서 A, C가 존재하고, 학습 결과 B, C-1, C-2가 생성되는 것이다. 입력 데이터 A, C에 대하여 도 4(a)와 같이 트리 형태의 데이터 구조를 획득할 수 있다. 각 노드는 데이터가 입력되거나, 출력을 생성할 수 있다. 도 4(a)는 가장 간단한 예시를 제시하는 것이지만, 이진 트리 형태로 더 복잡하고 다양한 트리를 생성할 수 있으며, 각 노드에 대하여 규칙적으로 제어문을 생성하는 것도 가능하다.
- [0050] 예컨대, 도 4(a)에서 생성된 규칙은 다음과 같다. A가 역치(threshold)보다 큰 경우 "1"을 출력한다. 그렇지 않다면, C가 yes라면 "2"를 출력한다. C가 "no"이면 "3"을 출력한다.
- [0051] 도 4(a)의 순서도를 도 4(b)에 나타내었다. 트리 구조의 장점은 재귀적으로 알고리즘을 구성할 수 있다는 점이다. 도 4(b)는 도 4(a)를 설명하기 위하여 제시된 순서도에 불과하지만, 트리 구조를 이용하여, 단순한 규칙을 반복적으로 적용하여 입력되는 제1 데이터셋에 대한 복잡한 기계 학습 모델을 생성할 수 있는 장점이 있다.
- [0053] 도 5는 본 발명의 부분 실시예에 따른 Apriori 알고리즘을 적용한 적응형 지식 베이스 구축 방법을 설명하기 위한 예시도이다. 구체적으로 도 5는 Apriori 알고리즘을 토대로 연관 분석 결과(Association Analysis result)를 규칙(rule)로 생성하는 방법을 설명하기 위한 예시도이다.
- [0054] Apriori 알고리즘은 기계 학습 방법으로 널리 알려진 방법이나, 시맨틱 규칙을 생성하는데 적용하여도 좋은 성능을 낼 수 있다.
- [0055] 도 5(a)에 나타난 숫자는 하기 표 1에 따른 품목(item)에 해당하는 것을 의미하고, A, B, C, D, E는 하나의 거래(transaction)를 의미하고, 하기 표 2는 각 거래에서 구입한 품목을 나타낸다.

표 1

품목 명칭	품목 아이디
soy milk	1
lettuce	2
diapers	3
wine	4
chard	5
orange juice	6

표 2

거래 ID	품목 리스트	품목 아이디 리스트
A	Soy milk, lettuce	1, 2
B	Lettuce, diapers	2, 3
C	Soy milk, wine	1, 4
D	Lettuce, soy milk, diapers, wine	2, 1, 3, 4
E	Lettuce, soy milk, diapers	2, 1, 3

[0060] 상기 표 2를 참고하여 도 5(a)의 시맨틱 웹을 형성할 수 있다. 형성된 시맨틱 웹은 시맨틱 규칙으로 규칙화할 수 있다.

[0061] 예컨대, 가장 아래에 있는 박스를 분석하면, 상기 거래 ID 중 1을 포함하는 것은 총 4개, 2를 포함하는 것은 총 4개, 3을 포함하는 것은 총 4개, 4를 포함하는 것은 총 3개인 것을 알 수 있다. 하기 표는 이를 정리한 것이다.

표 3

품목 집합	지지도
{1}	4
{2}	4
{3}	3
{4}	2

[0065] 이때, 최소 지지도를 4로 설정하면, {1}, {2}은 빈발 품목 집합으로 분류되고, {3}, {4}은 비빈발 품목 집합으로 분류된다. 도 5(a)에서는 빈발 품목 집합을 반전 시켜 표시하였다.

[0067] 마찬가지로, 원소의 개수가 2개인 품목 집합에 대하여 지지도를 연산할 수 있고, 이를 하기 표에 나타내었다.

표 4

품목 집합	지지도
{1,2}	3
{1,3}	2
{1,4}	2
{2,3}	3
{2,4}	1
{3,4}	1

[0070] 이때, 최소 지지도를 3으로 설정하면, {1,2}, {2,3}은 빈발 품목 집합이 되고, {1,3}, {1,4}, {2,4}, {3,4}는 비빈발 품목 집합으로 분류된다. 도 5(a)에서는 빈발 품목 집합을 반전 시켜 표시하였다.

[0072] 마찬가지로, 원소의 개수가 3개인 품목 집합에 대하여 지지도를 연산할 수 있고, 이를 하기 표에 나타내었다.

표 5

품목 집합	지지도
{1,2,3}	2
{1,2,4}	1
{1,3,4}	1
{2,3,4}	1

[0075] 이때, 최소 지지도를 2로 설정하면, {1,2,3}은 빈발 품목 집합이 되고, {1,3}, {1,2,4}, {1,3,4}, {2,3,4}는 비빈발 품목 집합으로 분류된다. 도 5(a)에서는 빈발 품목 집합을 반전 시켜 표시하였다.

[0077] 마찬가지로, 원소의 개수가 4개인 품목 집합에 대하여 지지도를 연산하는 경우, {1,2,3,4}는 지지도가 1이므로, 최소 지지도가 1이라면, 빈발 품목 집합이 되고, 최소 지지도가 2라면, 비빈발 품목 집합이 된다. 도 5(a)에서는 최소 지지도가 1인 경우를 상정하였다.

[0078] 통상 Apriori 알고리즘에서는 비빈발 품목을 가지치기(pruning)하고, 진행하여 연산 속도를 증가시킬 수 있는 기계 학습 방법이다.

[0079] 도 5(b)는 본 발명의 부분 실시예에 따른 Apriori 알고리즘을 적용하는 방법을 설명하기 위한 예시도이다. 구체적으로 도 5(b)는 도 5(a)에서 최소 지지도 미만의 노드에 대하여 가지치기(pruning)를 한 결과를 나타낸다.

[0080] 전체 트리에서 모든 연산을 다 할 경우, 연산 횟수는 품목의 개수에 따라 지수적으로 증가하는데, 이를 컴퓨터

가 연산 속도를 따라갈 수 없는 문제가 있다. 최소 지지도 이하의 노드에 대하여 가지치기(pruning)하고, 최소 지지도 이상의 노드에 대하여만 연산을 계속할 경우, 적은 횟수의 연산에 의하여, 노드 간의 의미관련 규칙 (semantic rule)을 생성할 수 있다. 도 5(b)에서 점선으로 표시한 화살표는 가지치기 될 노드를 알려주고 있다.

- [0082] 도 6은 본 발명의 다른 실시예에 따른 적응형 지식 베이스 구축 시스템의 구성도를 나타낸다.
- [0083] 본 발명에 따른 적응형 지식 베이스 구축 시스템은 분산 처리 시스템으로 구성될 수 있다.
- [0084] 본 발명에 따른 적응형 지식 베이스 구축 시스템은 기계 학습 규칙 생성 서버(610); 시맨틱 규칙 생성 서버(620); 및 규칙 모델러 서버(630);를 포함한다.
- [0086] 기계 학습 규칙 생성 서버(610)는 기계 학습 엔진(210); 및 규칙 생성기(220)를 포함한다. 기계 학습 규칙 생성 서버(610)는 반복적 재학습(repeatitive learning)을 통하여 동적 환경 문맥(context)을 반영한다.
- [0087] 시맨틱 규칙 생성 서버(620)는 기계 학습 규칙으로부터 생성된 시맨틱 규칙을 이용하여 제2 데이터셋에 대한 시맨틱 추론을 통하여 지식 베이스를 확장한다. 또한, 기계 학습 엔진이 제2 데이터셋을 입력 데이터로 하여 예측한 기계 학습 규칙을 시맨틱 규칙 엔진에 전달하고, 시맨틱 추론을 통하여 지식 베이스를 확장할 수도 있다. 이에 관한 설명은 도 7 및 도 8에서 후술한다.
- [0088] 규칙 생성기(220)가 생성된 규칙을 시맨틱 규칙 엔진(240)이 시맨틱 규칙으로 변환하는데, 시맨틱 규칙 엔진은 새롭게 입력된 제2 데이터셋을 이용하여 시맨틱 규칙을 생성할 수 있다. 이때, 새롭게 입력된 제2 데이터셋에 대한 시맨틱 추론을 이용하여 지식 베이스를 확장할 수 있다.
- [0089] 상기 시맨틱 규칙은 RDF 트리플 트리 구조 또는 OWL 구조 중 어느 하나일 수 있다.
- [0090] 시맨틱 규칙 엔진(230)은 규칙 모델러(240)로부터 규칙 모델을 전달받아, 생성된 시맨틱 규칙을 수정할 수 있고, 기계 학습 규칙을 시맨틱 규칙으로 변환하는 방법을 변경할 수 있다.
- [0091] 규칙 모델러 서버(630)은 도메인 전문가에 의하여 입력된 규칙 모델을 시맨틱 규칙 엔진에 전송하는 규칙 모델러를 포함한다.
- [0092] 규칙 모델러는 도메인 전문가가 손쉽게 규칙 모델을 입력할 수 있도록 편리한 사용자 인터페이스(User Interface)를 제공한다. 기계 학습 규칙 및 시맨틱 규칙을 손쉽게 생성할 수 있는 UI를 제공하고, 기계 학습 규칙 및 시맨틱 규칙을 연결하는 UI도 함께 제공한다. 또한 규칙 모델러는 기계 학습 엔진이 생성한 기계 학습 규칙 및 시맨틱 규칙 사이의 연결 관계를 지식 베이스에 저장하기 전에 열람하고, 수정할 수 있는 사용자 인터페이스를 제공한다.
- [0094] 도 7은 본 발명의 다른 실시예에 따른 적응형 지식 베이스 구축 시스템의 구성도를 나타낸다.
- [0095] 도 6과 달리 새롭게 입력된 제2 데이터셋은 시맨틱 규칙 엔진이 아니라 기계 학습 엔진에 직접 입력될 수도 있다. 시맨틱 규칙 엔진에 기계 학습 모듈을 포함시켜 도 6과 같이 처리하는 것도 가능하나, 통상 오랜 시간을 필요로 하는 기계 학습 과정은 단축 시키기 위하여, 제2 데이터 집합을 제1 데이터 집합과 함께 기계 학습 엔진에 모두 입력하여, 기계 학습 과정에서 한번에 처리할 경우, 컴퓨터 자원을 효율적으로 사용하는 것이 가능하다. 본 발명에 따른 적응형 지식 베이스 구축 시스템은 분산 처리 시스템으로 구축할 수 있고, 통상 기계 학습 엔진은 다수의 서버에 병렬형 시스템으로 구축되고, 시맨틱 규칙 엔진은 소수의 서버에서 설치되는 경우, 제2 데이터셋을 기계 학습 엔진에 입력하는 방법이 자원을 효율적으로 분배할 수 있을 것이다.
- [0096] 본 발명의 부분 실시예에 따른 적응형 지식 베이스 구축 시스템은, 제1 데이터셋에 대한 학습 모델을 생성하고, 생성된 학습 모델을 이용하여 제1 데이터셋의 연관 관계를 분석하고, 제2 데이터셋에 대한 시맨틱 추론 모델을 생성하고, 생성된 시맨틱 추론 모델을 이용하여 제2 데이터셋을 이용하여 추론하고 예측 결과를 생성하는 기계 학습 엔진; 상기 기계 학습 엔진이 분석한 학습된 모델 결과로부터 기계 학습 규칙을 생성하고, 상기 추론된 예측 결과로부터 기계 학습 규칙을 생성하는 규칙 생성기; 규칙 생성기로부터 전달받은 기계 학습 규칙을 시맨틱 규칙으로 변환하고, 변환된 시맨틱 규칙을 저장하여 지식 베이스를 구축하는 시맨틱 규칙 엔진;을 포함한다.
- [0097] 바람직하게는 도메인 전문가에 의하여 시맨틱 규칙 생성 방법을 변경하는 규칙 모델러를 더 포함할 수 있다.
- [0098] 필요에 따라 시맨틱 규칙뿐만 아니라, 기계 학습 규칙도 확장될 수 있지만, 기계 학습 규칙 자체는 결정적인 특성이 있어서, 실제 환경이 동적으로 변화하는 환경에서 적응적으로 기계 학습 규칙이 변경될 필요성이 있지만, 기계 학습 엔진을 변경하는 것은 효율적이지 못한 문제가 있다. 규칙 모델러는 기계 학습 규칙을 확장하는

대신, 기계 학습 규칙을 시맨틱 규칙으로 변환하는 방법을 변경함으로써, 동적으로 변화하는 환경에서 적절한 변환 방법을 선택할 수 있도록 도움을 준다. 도메인 전문가는 본 발명에 따른 적응형 지식 베이스 구축 시스템의 구조를 알고 있는 자는 아니나, 시맨틱 규칙에 대한 정보에 대하여 해박한 지식을 가진 자이므로, 생성된 시맨틱 규칙을 지식 베이스에 곧바로 저장하는 대신 도메인 전문가에 의하여 채택 여부를 선택하게 할 수 있고, 이를 토대로 규칙 모델러가 작동할 수 있다. 규칙 모델러는 별도의 학습 과정을 통하여 기계 학습 규칙을 시맨틱 규칙으로 변환하는 방법을 결정하여 시맨틱 규칙 엔진에 전달하고, 시맨틱 규칙 엔진은 규칙 모델러로부터 전달받은 변환 방법을 이용하여 지식 베이스를 구축할 수 있다.

[0099] 시맨틱 규칙 엔진은 규칙 생성기로부터 전달받은 기계 학습 규칙과, 제2 데이터셋과, 규칙 모델러로부터 전달받은 규칙 모델을 토대로 지식 베이스를 구축한다. 이러한 과정은 곧바로 진행되는 것이 아니라, 임의로 각각의 데이터를 저장하고, 적절한 주기로 지식 베이스를 구축하게 된다. 시맨틱 규칙 엔진은 본 발명의 특성에 따라 제1 데이터셋의 입력 순서, 제2 데이터셋의 입력 순서, 규칙 모델이 입력되는 순서 모두에 의하여 구축되는 지식 베이스도 상이해지는 특성이 있다.

[0100] 그러나, 현실적으로 시간에 따라 입력되는 데이터가 달라질 경우, 그 분석 또한 상이해지는 것이 통상적인 것이다. 예컨대, IoT(Internet of Thing) 환경 하에서, 집안의 보일러를 작동하는 경우, 보일러가 작동되는 시간에 따라 기계 학습 내용은 상이해질 수 있다. 즉, 본 발명에 따른 적응형 지식 베이스 구축 시스템은 동적으로 변환하는 환경에서 더욱 강인한 특성이 있다.

[0102] 도 8은 본 발명의 다른 실시예에 따른 적응형 지식 베이스 구축 방법의 절차 흐름도를 나타낸다.

[0103] 본 발명의 부분 실시예에 따른 적응형 지식 베이스 구축 방법은, 제2 데이터셋에 대한 시맨틱 추론 모델을 생성하는 단계; 제2 데이터셋에 생성된 시맨틱 추론 모델을 적용한 추론을 하고 예측 결과를 생성하는 단계; 상기 추론된 예측 결과로부터 기계 학습 규칙을 생성하는 단계; 기계 학습 규칙을 시맨틱 규칙으로 변환하는 단계; 변환된 시맨틱 규칙을 저장하여 지식 베이스를 구축하는 단계;를 포함한다.

[0104] 도 3의 설명과 달리 도 8의 경우, 제1 데이터셋을 입력 데이터로 하여 분석한 지식 베이스가 이미 구축된 경우, 제2 데이터셋을 시맨틱 규칙 엔진이 아닌 기계 학습 엔진에 입력하여 시맨틱 추론이 이루어진다는 점에서 차이가 있다. 제2 데이터셋이 시맨틱 규칙 엔진에 직접 입력될 경우, 구축된 지식 베이스를 이용하여 시맨틱 추론이 이루어질 수 있는 장점이 있으나, 통상 기계 학습 엔진은 고사양의 서버를 요하므로, 기계 학습 엔진의 컴퓨팅 파워가 통상 더 좋은 경우가 많다. 기계 학습 엔진에 제2 데이터셋을 입력하는 경우, 제2 데이터셋을 포함하여 학습 모델을 적용하기 쉽고, 고사양의 서버를 이용하여 빠르게 제2 데이터셋에 대한 추론을 진행할 수 있는 장점이 있다. 미리 추론에 이용된 데이터는 시맨틱 규칙 엔진에 다시 입력되어 구축된 지식 베이스를 이용하여 확인하고, 지식 베이스를 확장할 수 있다.

[0106] 이상, 본 발명의 구성에 대하여 첨부 도면을 참조하여 상세히 설명하였으나, 이는 예시에 불과한 것으로서, 본 발명이 속하는 기술 분야에 통상의 지식을 가진 자라면 본 발명의 기술적 사상의 범위 내에서 다양한 변형과 변경이 가능함은 물론이다. 따라서 본 발명의 보호 범위는 전술한 실시예에 국한되어서는 아니 되며 이하의 특허 청구범위의 기재에 의하여 정해져야 할 것이다.

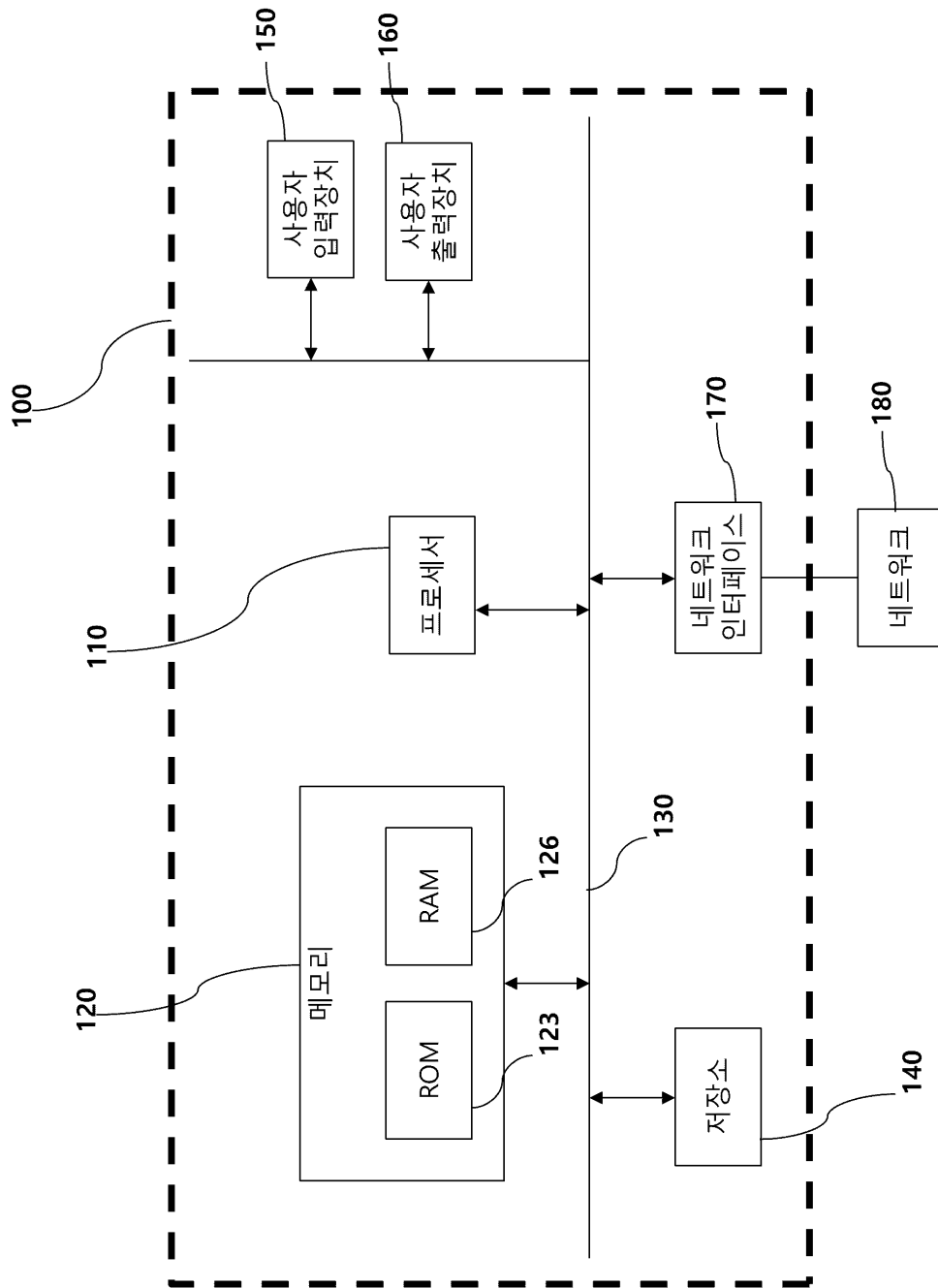
부호의 설명

- [0108] 100: 컴퓨터 시스템
- 110: 프로세서
- 120: 메모리
- 123: ROM
- 126: RAM
- 130: 데이터 통신 버스
- 140: 저장소
- 150: 사용자 입력 장치
- 160: 사용자 출력 장치

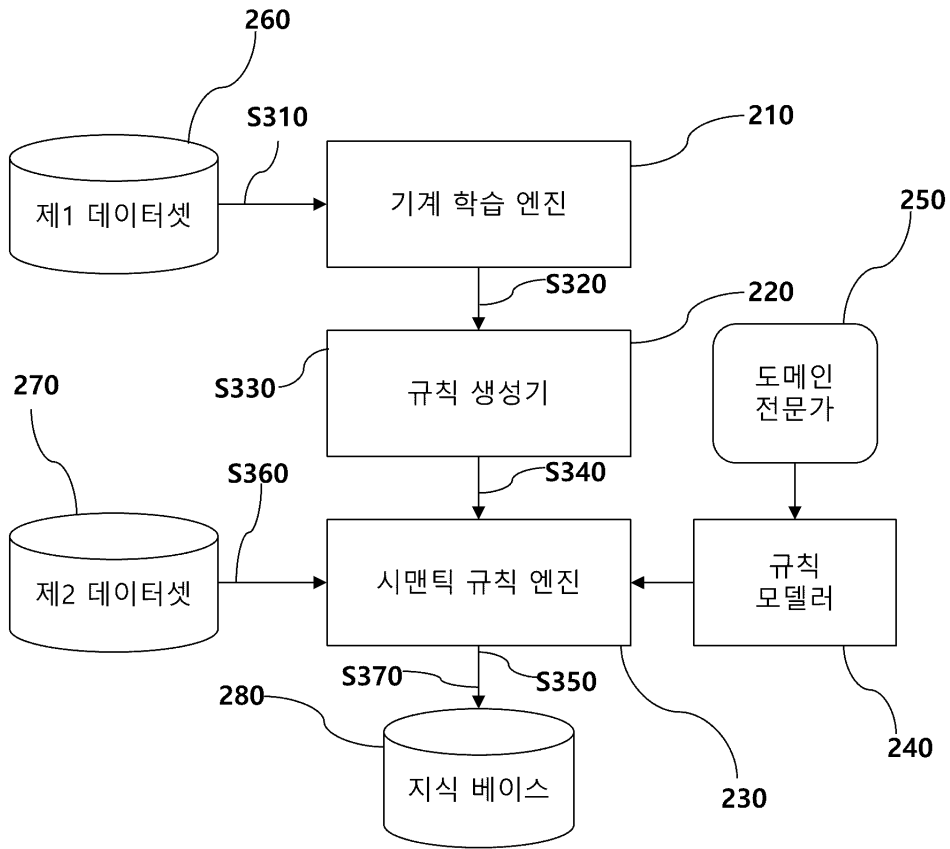
- 170: 네트워크 인터페이스
- 180: 네트워크
- 210: 기계 학습 엔진
- 220: 규칙 생성기
- 230: 시맨틱 규칙 엔진
- 240: 규칙 모델러
- 250: 도메인 전문가
- 260: 제1 데이터셋
- 270: 제2 데이터셋
- 280: 지식 베이스
- S310: 학습 모델 생성 단계
- S320: 학습된 모델 결과 생성 단계
- S330: 기계 학습 규칙 생성 단계;
- S340: 시맨틱 규칙 생성 단계;
- S350: 지식 베이스 구축 단계;
- S360: 제2 데이터셋을 이용한 시맨틱 추론 단계;
- S370: 지식 베이스 확장 단계;
- 610: 기계 학습 규칙 생성 서버
- 620: 시맨틱 규칙 생성 서버
- 630: 규칙 모델러 서버
- S810: 예측 모델 생성 단계;
- S820: 예측 결과 생성 단계;
- S830: 예측 결과의 규칙 생성 단계;
- S840: 시맨틱 규칙 변환 단계;
- S850: 지식 베이스 확장 단계;

도면

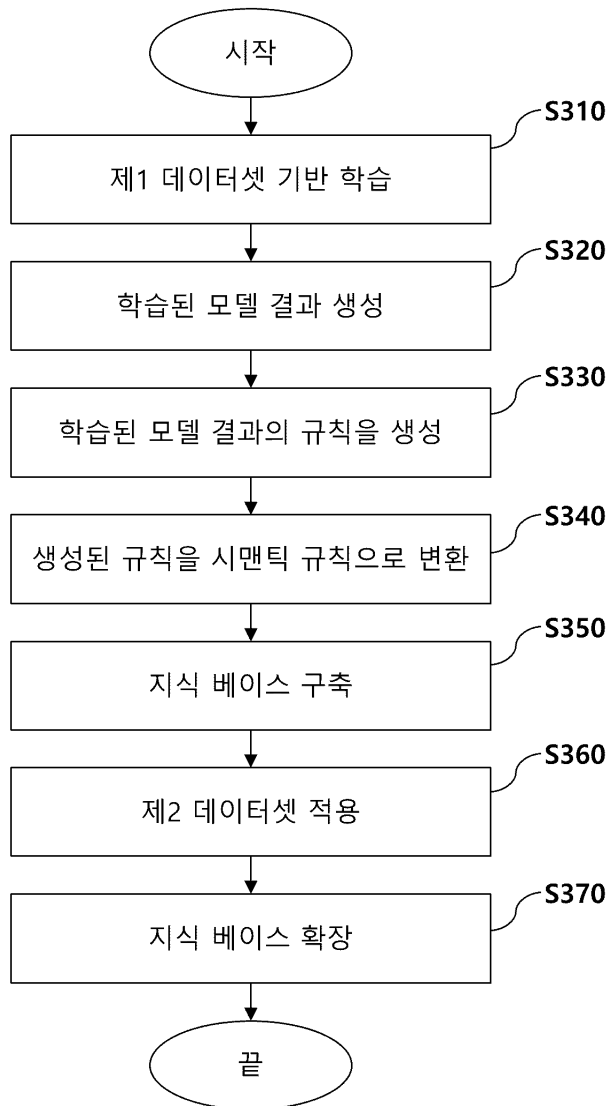
도면1



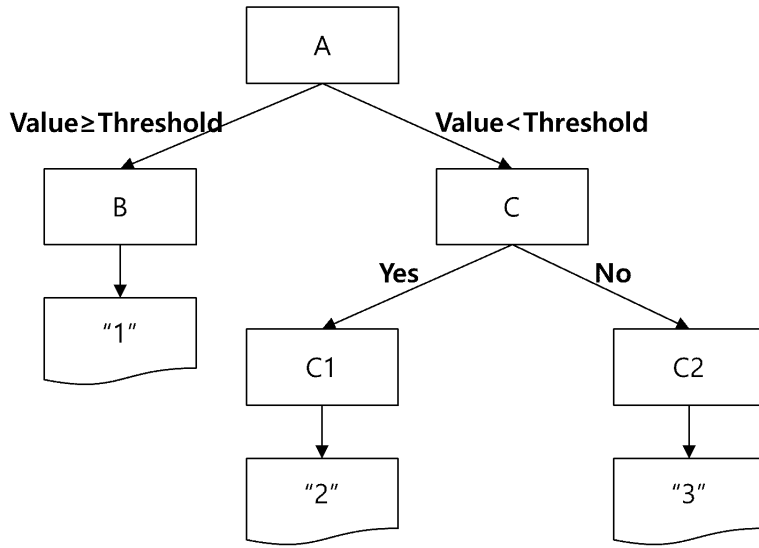
도면2



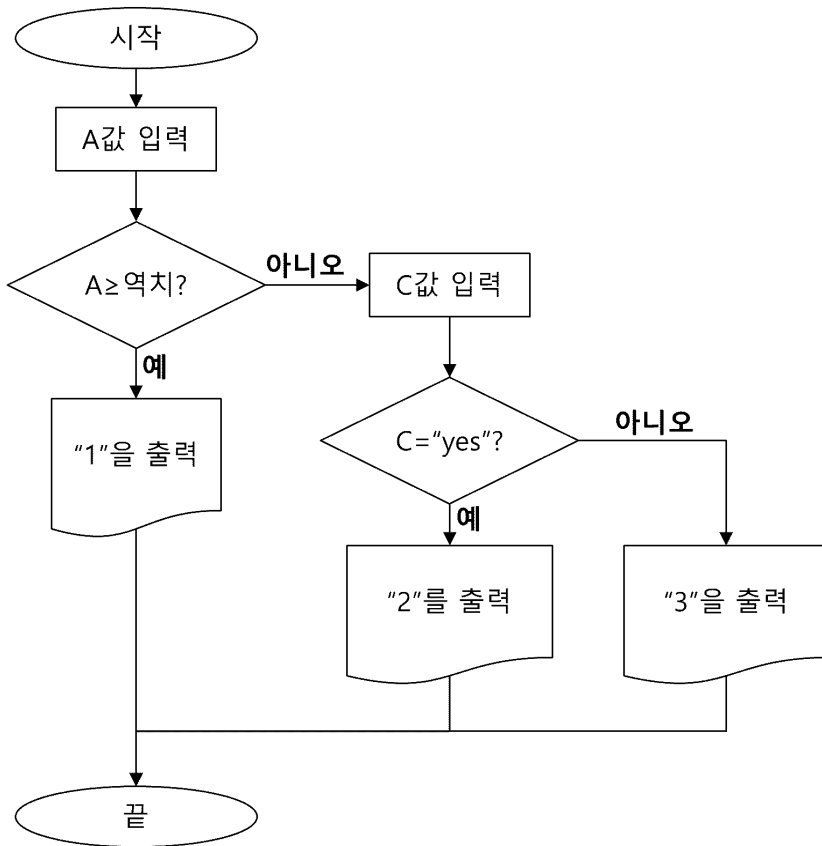
도면3



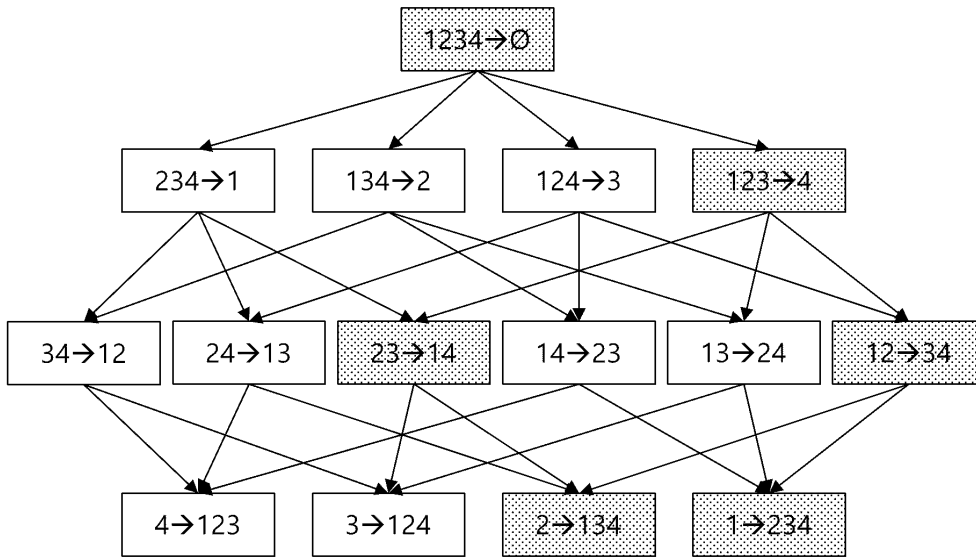
도면4a



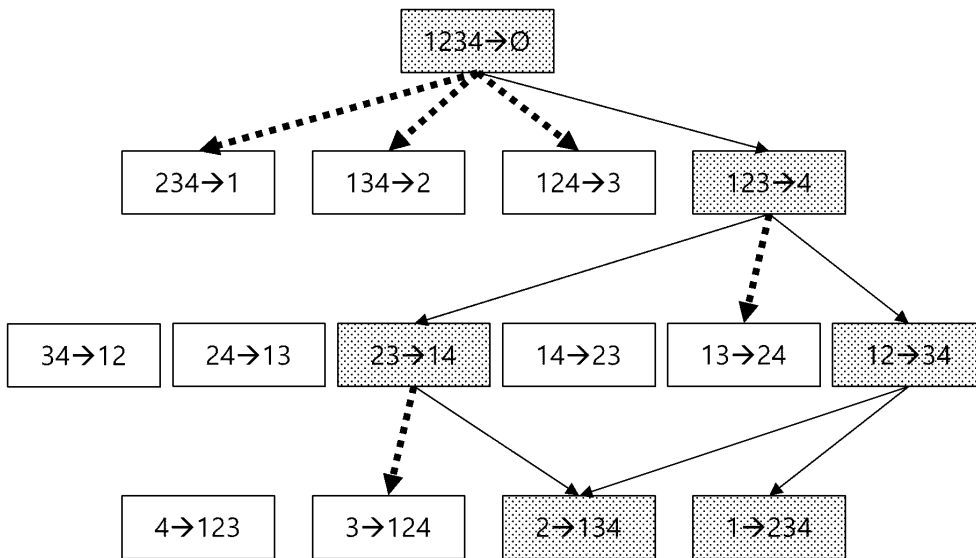
도면4b



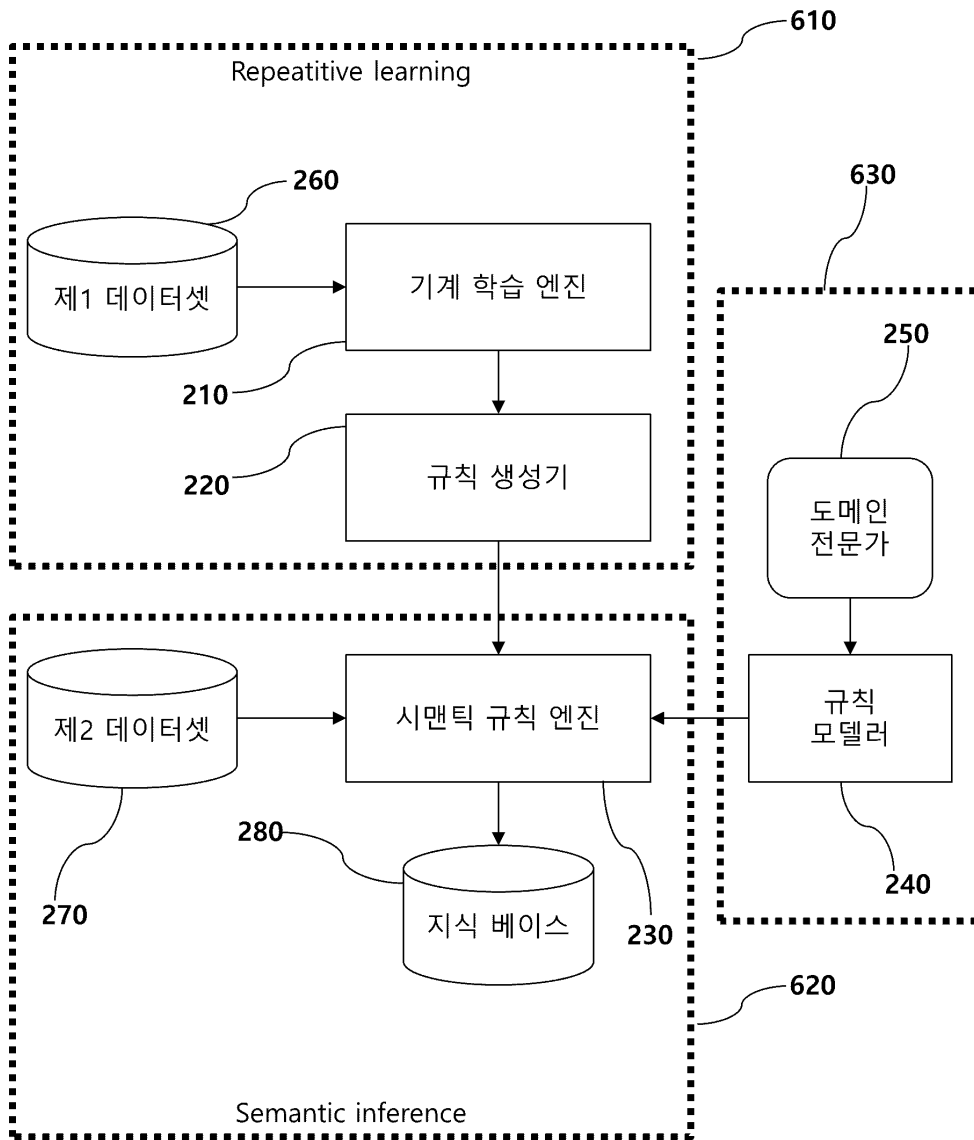
도면5a



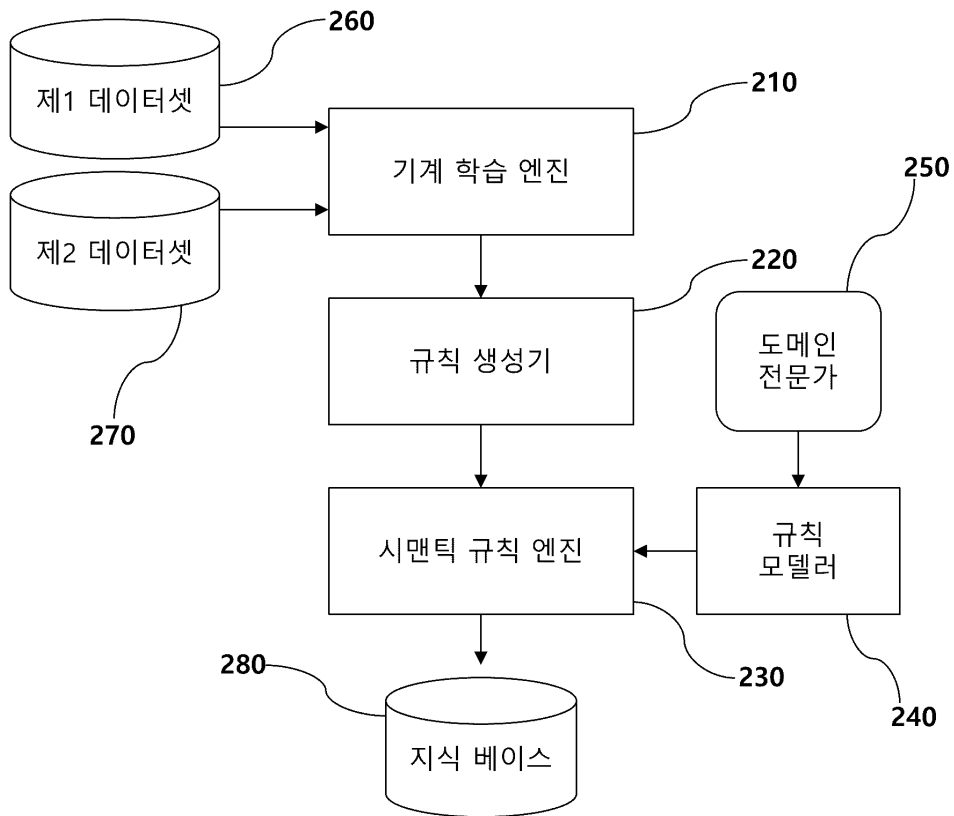
도면5b



도면6



도면7



도면8

