



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0117005
(43) 공개일자 2022년08월23일

(51) 국제특허분류(Int. Cl.)
G06N 20/20 (2019.01) G06N 5/02 (2006.01)
(52) CPC특허분류
G06N 20/20 (2021.08)
G06N 5/02 (2019.01)
(21) 출원번호 10-2021-0020564
(22) 출원일자 2021년02월16일
심사청구일자 2021년02월16일

(71) 출원인
한국기술교육대학교 산학협력단
충청남도 천안시 동남구 병천면 충절로 1600 (한
국기술교육대학교내)
(72) 발명자
김원태
충청남도 아산시 음봉면 월암로 230
이승기
충청남도 천안시 동남구 병천면 가전7길 1-30,
B-103
김영진
충청북도 청주시 청원구 오창읍 오창공원로 96,
103동 3504호
(74) 대리인
정승교

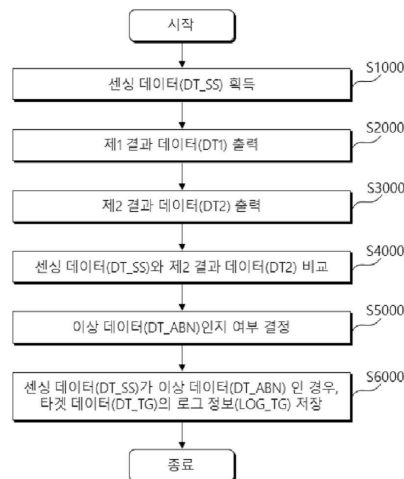
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 기계학습 모델의 이상 데이터 추출 방법 및 장치

(57) 요약

기계학습 모델의 이상 데이터 추출 방법이 제공된다. 본 발명의 몇몇 실시예에 따른 기계학습 모델의 이상 데이터 추출 방법은, 센싱 데이터를 획득하는 단계, 센싱 데이터를 입력으로 구동하는 제1 기계학습 모델에서 제1 결과 데이터를 출력하는 단계, 센싱 데이터를 입력으로 구동하되 제1 기계학습 모델과 상이한 제2 기계학습 모델에서 제2 결과 데이터를 출력하는 단계, 센싱 데이터와 제2 결과 데이터를 비교하는 단계, 비교 결과에 기초하여 센싱 데이터가 이상 데이터인지 여부를 결정하는 단계 및 센싱 데이터가 이상 데이터로 결정된 경우, 이상 데이터로 결정된 센싱 데이터인 타겟 데이터의 로그 정보를 저장하는 단계를 포함한다.

대표도 - 도6



이 발명을 지원한 국가연구개발사업

과제고유번호	1711102838
과제번호	2019-0-01347-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	ICT기반사회문제해결기술개발(R&D)
연구과제명	화재 현장 시뮬레이션 및 지휘 역량 강화 지원형 실감 소방훈련 콘텐츠 기술 개발
기 여 율	1/1
과제수행기관명	한국전자통신연구원
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

기계학습 모델의 이상 데이터 추출 방법으로서,

센싱 데이터를 획득하는 단계;

상기 센싱 데이터를 입력으로 구동하는 제1 기계학습 모델에서 제1 결과 데이터를 출력하는 단계;

상기 센싱 데이터를 입력으로 구동하되, 상기 제1 기계학습 모델과 상이한 제2 기계학습 모델에서 제2 결과 데이터를 출력하는 단계;

상기 센싱 데이터와 상기 제2 결과 데이터를 비교하는 단계;

상기 비교 결과에 기초하여 상기 센싱 데이터가 이상 데이터인지 여부를 결정하는 단계; 및

상기 센싱 데이터가 이상 데이터로 결정된 경우, 상기 이상 데이터로 결정된 센싱 데이터인 타겟 데이터의 로그 정보를 저장하는 단계를 포함하는, 기계학습 모델의 이상 데이터 추출 방법.

청구항 2

청구항 1에 있어서,

상기 제2 기계학습 모델은 과소완전 적층 오토인코더(Undercomplete Stacked Autoencoder)이고,

상기 센싱 데이터와 상기 제2 결과 데이터를 비교하는 단계는, 상기 과소완전 적층 오토인코더의 재현율을 결정하는 단계를 포함하고,

상기 센싱 데이터가 이상 데이터인지 여부를 결정하는 단계는, 상기 재현율이 기정의된 임계값 이상인지 여부에 기초하여 상기 이상 데이터 여부를 결정하는 단계를 포함하는, 기계학습 모델의 이상 데이터 추출 방법.

청구항 3

청구항 2에 있어서,

상기 센싱 데이터가 이상 데이터인지 여부를 결정하는 단계는,

상기 재현율이 상기 임계값 미만인 경우에 상기 센싱 데이터를 상기 이상 데이터로 결정하는 단계를 포함하는, 기계학습 모델의 이상 데이터 추출 방법.

청구항 4

청구항 1에 있어서,

상기 로그 정보는, 상기 타겟 데이터, 상기 타겟 데이터를 입력으로 상기 제1 기계학습 모델에 의해 출력된 제1 타겟 결과 데이터, 상기 타겟 데이터를 입력으로 상기 제2 기계학습 모델에 의해 출력된 제2 타겟 결과 데이터, 상기 타겟 데이터의 입력 시간 및 상기 타겟 데이터의 센싱 시간 중 적어도 하나의 정보를 포함하는, 기계학습 모델의 이상 데이터 추출 방법.

청구항 5

청구항 1에 있어서,

상기 로그 정보를 학습 데이터로 이용하여 상기 제1 기계학습 모델의 전이 학습(Transfer Learning)을 수행하는 단계를 더 포함하는, 기계학습 모델의 이상 데이터 추출 방법.

청구항 6

청구항 1에 있어서,

상기 센싱 데이터는 온도, 습도, 연기량 및 가스량 중 적어도 하나의 정보를 포함하고,

상기 제1 결과 데이터를 출력하는 단계는, 타겟 공간의 화재 발생 여부에 대한 결과를 포함하는 상기 제1 결과 데이터를 출력하는 단계를 포함하는, 기계학습 모델의 이상 데이터 추출 방법.

청구항 7

청구항 6에 있어서,

상기 타겟 데이터의 로그 정보를 저장하는 단계는,

상기 제2 기계학습 모델에 의해 상기 센싱 데이터가 이상 데이터로 결정되고, 상기 제1 기계학습 모델에 의해 생성된 상기 제1 결과 데이터가 상기 타겟 공간에 화재가 발생했다는 정보를 포함하는 경우 상기 타겟 데이터의 로그 정보를 저장하는 단계를 포함하는, 기계학습 모델의 이상 데이터 추출 방법.

청구항 8

청구항 6에 있어서,

상기 타겟 데이터의 센싱 시점의 온풍기 작동 상태, 창문 개폐 상태 및 날씨 정보 중 적어도 하나의 상태 정보를 획득하는 단계; 및

상기 로그 정보 및 상기 상태 정보를 학습 데이터로 이용하여 상기 제1 기계학습 모델의 전이 학습을 수행하는 단계를 더 포함하는, 기계학습 모델의 이상 데이터 추출 방법.

청구항 9

청구항 1에 있어서,

상기 센싱 데이터를 획득하는 단계는, 복수의 센싱 데이터를 획득하는 단계; 및 상기 복수의 센싱 데이터를 기 정의된 구간으로 스케일링한 전처리 데이터를 생성하는 단계를 포함하되,

상기 제1 및 제2 기계학습 모델은, 상기 전처리 데이터를 입력으로 각각 상기 제1 및 제2 결과 데이터를 출력하는, 기계학습 모델의 이상 데이터 추출 방법.

청구항 10

하드웨어와 결합되어 청구항 1 내지 청구항 9 중 어느 하나의 항의 방법을 실행시키기 위하여 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

발명의 설명

기술 분야

본 발명은 기계학습 모델의 이상 데이터 추출 방법에 관한 것으로, 더욱 상세하게는 오토인코더의 재현율에 기

[0001]

초하여 이상 데이터를 추출하고, 추출된 이상 데이터의 로그 정보를 이용하여 기계학습 모델의 전이 학습 성능을 향상시키는 방법에 관한 것이다.

배경 기술

- [0003] 사람의 거주지, 사무실, 공장 등에서 발생하는 화재는 막대한 재산 피해 및 인명 피해를 발생시키고, 이로 인해 화재의 발생을 초기에 판단하여 빠른 대처를 가능하게 하는 화재 감지 알고리즘의 중요성이 부각되어왔다.
- [0004] 기존의 화재 감지기는 단순 룰 베이스를 기반으로 단일 온도센서를 이용해 특정한 온도 값을 감지하거나, 단일 연기센서를 이용해 특정한 연기량(농도)을 감지하는 방식으로 화재를 판단했다. 이러한 화재 감지기는 비용적으로 저렴하다는 장점이 있으나 비화재보(Unwanted Alarm)에 약하다는 단점이 있다. 비화재보, 즉 화재 이외의 요인에 의한 오경보율을 낮추기 위해 여러가지 복합 센서를 이용해 화재를 감지하는 기법을 이용할 수 있으며, 특히 근래에는 복합적인 판단을 위해 기계학습 기법을 이용한 화재 감지 방법이 요구되고 있다.
- [0005] 기계 학습(Machine Learning)은 영상 처리, 음성 인식, 자연어 처리, 화재 감지 등 다양한 목적을 위해 사용된다. 지도 학습(Supervised Learning) 기반의 기계 학습 모델의 장점은 수많은 데이터의 학습을 통해 정확도 높은 분류를 할 수 있다는 것이다. 지도 학습 기반의 기계 학습 모델은 분류하고자 하는 상황에 대한 데이터를 수집하고 적합한 레이블(label)을 부과하여 학습시킬 수 있으며, 이후 학습된 기계 학습 모델은 수신되는 데이터들이 어느 상황에 속하는지 판단할 수 있다. 예를 들어 화재 감지를 위한 기계 학습 모델은 비화재 상황과 화재 상황에 대한 센서 데이터를 수집하고 각각의 레이블을 부과하여 학습(training)시키면 비화재 상황과 화재 상황을 분류할 수 있다. 그러나 학습 단계에서 고려하지 못했던 이상 상황(예를 들어, 센서 오류, 간섭, 고장, 환경 잡음 등)에서는 판단의 오류가 발생할 수 있다. 특히 화재 감지의 경우 이러한 판단 오류는 비화재보에 의한 피해를 야기할 수 있다. 이러한 이유로 판단 오류를 줄이기 위한 기법이 필요하며, 그 기법 중 하나로 지속적으로 수집되는 데이터로부터의 재학습을 수행하는 전이 학습(Transfer Learning) 기법을 통한 판단 오류를 완화 대책이 요구된다.
- [0006] 한편, 전이 학습을 위한 레이블링 작업은 수집되는 방대한 양의 데이터로 인해 많은 시간과 비용(cost)을 소모한다. 또한, 지속적으로 수집되는 방대한 양의 데이터에 포함된 학습 데이터와 유사한 데이터는 기계 학습 모델의 성능 향상에 큰 영향을 미치지 않으므로, 레이블링 과정에서의 낭비를 초래한다. 반면에, 기계학습 모델의 학습 시 고려되지 못한 상황에 직면했을 때 발생하는 이상 데이터를 활용하여 기계학습 모델의 전이 학습을 수행한다면 상대적으로 높은 성능 향상을 기대할 수 있다.

발명의 내용

해결하려는 과제

- [0008] 본 발명이 해결하고자 하는 기술적 과제는 센싱 값의 이상여부를 판단하고, 이에 대한 로그를 저장함으로써 기계학습 모델의 전이 학습 성능을 증대시키는 방법 및 장치를 제공하는 것이다.
- [0009] 또한, 전이 학습에 사용되는 학습 데이터(training data)의 양을 감소시킴으로써 학습의 시간을 단축시키면서 기계학습 모델의 성능을 향상시키는 방법 및 장치를 제공하는 것이다.
- [0010] 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 해당 기술 분야의 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0012] 상기 기술적 과제를 달성하기 위한 본 발명의 몇몇 실시예에 따른 기계학습 모델의 이상 데이터 추출 방법은, 센싱 데이터를 획득하는 단계, 센싱 데이터를 입력으로 구동하는 제1 기계학습 모델에서 제1 결과 데이터를 출력하는 단계, 센싱 데이터를 입력으로 구동하되, 제1 기계학습 모델과 상이한 제2 기계학습 모델에서 제2 결과 데이터를 출력하는 단계, 센싱 데이터와 제2 결과 데이터를 비교하는 단계, 비교 결과에 기초하여 센싱 데이터가 이상 데이터인지 여부를 결정하는 단계 및 센싱 데이터가 이상 데이터로 결정된 경우, 이상 데이터로 결정된

센싱 데이터인 타겟 데이터의 로그 정보를 저장하는 단계를 포함할 수 있다.

- [0013] 몇몇 실시예에 따라, 제2 기계학습 모델은 과소완전 적층 오토인코더(Undercomplete Stacked Autoencoder)이고, 센싱 데이터와 제2 결과 데이터를 비교하는 단계는, 과소완전 적층 오토인코더의 재현율을 결정하는 단계를 포함할 수 있다.
- [0014] 몇몇 실시예에 따라, 센싱 데이터가 이상 데이터인지 여부를 결정하는 단계는, 재현율이 기정의된 임계값 이상 인지 여부에 기초하여 이상 데이터 여부를 결정하는 단계를 포함할 수 있다.
- [0015] 이 때, 센싱 데이터가 이상 데이터인지 여부를 결정하는 단계는, 재현율이 임계값 미만인 경우에 센싱 데이터를 이상 데이터로 결정하는 단계를 포함할 수 있다.
- [0016] 몇몇 실시예에 따라, 로그 정보는 타겟 데이터, 타겟 데이터를 입력으로 제1 기계학습 모델에 의해 출력된 제1 타겟 결과 데이터, 타겟 데이터를 입력으로 제2 기계학습 모델에 의해 출력된 제2 타겟 결과 데이터, 타겟 데이터의 입력 시간 및 타겟 데이터의 센싱 시간 중 적어도 하나의 정보를 포함할 수 있다.
- [0017] 몇몇 실시예에 따라, 로그 정보를 학습 데이터로 이용하여 제1 기계학습 모델의 전이 학습(Transfer Learning)을 수행하는 단계를 더 포함할 수 있다.
- [0018] 몇몇 실시예에 따라, 센싱 데이터는 온도, 습도, 연기량 및 가스량 중 적어도 하나의 정보를 포함할 수 있고, 제1 결과 데이터를 출력하는 단계는, 타겟 공간의 화재 발생 여부에 대한 결과를 포함하는 제1 결과 데이터를 출력하는 단계를 포함할 수 있다.
- [0019] 몇몇 실시예에 따라, 타겟 데이터의 로그 정보를 저장하는 단계는, 제2 기계학습 모델에 의해 센싱 데이터가 이상 데이터로 결정되고, 제1 기계학습 모델에 의해 생성된 제1 결과 데이터가 타겟 공간에 화재가 발생했다는 정보를 포함하는 경우 타겟 데이터의 로그 정보를 저장하는 단계를 포함할 수 있다.
- [0020] 몇몇 실시예에 따라, 타겟 데이터의 센싱 시점의 온풍기 작동 상태, 창문 개폐 상태 및 날씨 정보 중 적어도 하나의 상태 정보를 획득하는 단계와, 로그 정보 및 상태 정보를 학습 데이터로 이용하여 제1 기계학습 모델의 전이 학습을 수행하는 단계를 더 포함할 수 있다.
- [0021] 몇몇 실시예에 따라, 센싱 데이터를 획득하는 단계는, 복수의 센싱 데이터를 획득하는 단계 및 복수의 센싱 데이터를 기정의된 구간으로 스케일링한 전처리 데이터를 생성하는 단계를 포함할 수 있다.
- [0022] 이 때, 제1 및 제2 기계학습 모델은, 전처리 데이터를 입력으로 각각 제1 및 제2 결과 데이터를 출력할 수 있다.
- [0023] 본 발명의 몇몇 실시예에 따라, 하드웨어와 결합되어 상기의 방법을 실행시키기 위하여 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다.
- [0024] 기타 실시예들의 구체적인 사항들은 상세한 설명 및 도면들에 포함되어 있다.

도면의 간단한 설명

- [0026] 도 1은 본 발명의 몇몇 실시예에 따른 이상 데이터 추출 방법이 적용되는 장치를 설명하기 위한 도면이다.
- 도 2는 본 발명의 몇몇 실시예에 따른 판단부를 설명하기 위한 도면이다.
- 도 3은 본 발명의 몇몇 실시예에 따른 제2 기계학습 모델을 설명하기 위한 도면이다.
- 도 4는 본 발명의 몇몇 실시예에 따른 오토인코더를 설명하기 위한 도면이다.
- 도 5는 본 발명의 몇몇 실시예에 따른 입력부를 설명하기 위한 도면이다.
- 도 6은 본 발명의 몇몇 실시예에 따른 기계학습 모델의 이상 데이터 추출 방법을 설명하기 위한 순서도이다.
- 도 7은 센싱 데이터와 제2 결과 데이터의 재현율에 기초하여 이상 데이터를 결정하는 실시예를 설명하기 위한 순서도이다.
- 도 8은 타겟 데이터의 로그 정보를 이용하여 제1 기계학습 모델의 전이 학습을 수행하는 실시예를 설명하기 위한 순서도이다.

도 9는 도 8의 실시예를 설명하기 위한 블록도이다.

도 10은 센싱 데이터 및 제1 결과 데이터에 기초하여 로그 정보의 저장 여부를 결정하는 실시예를 설명하기 위한 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0027] 이하, 본 발명의 다양한 실시예가 첨부된 도면을 참조하여 기재된다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 실시예의 다양한 변경(modification), 균등물(equivalent), 및/또는 대체물(alternative)을 포함하는 것으로 이해되어야 한다.
- [0028] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 해당 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0029] 본 문서의 다양한 실시 예들 및 이에 사용된 용어들은 본 문서에 기재된 기술적 특징들을 특정한 실시 예들로 한정하려는 것이 아니며, 해당 실시 예의 다양한 변경, 균등물, 또는 대체물을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 또는 관련된 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 아이টে에 대응하는 명사의 단수 형은 관련된 문맥상 명백하게 다르게 지시하지 않는 한, 상기 아이টে 한 개 또는 복수 개를 포함할 수 있다. 본 문서에서, "A 또는 B", "A 및 B 중 적어도 하나", "A 또는 B 중 적어도 하나", "A, B 또는 C", "A, B 및 C 중 적어도 하나" 및 "A, B, 또는 C 중 적어도 하나"와 같은 문구들 각각은 그 문구들 중 해당하는 문구에 함께 나열된 항목들 중 어느 하나, 또는 그들의 모든 가능한 조합을 포함할 수 있다. "제1", "제2"와 같은 용어들은 단순히 해당 구성요소를 다른 해당 구성요소와 구분하기 위해 사용될 수 있으며, 해당 구성요소들을 다른 측면(예: 중요성 또는 순서)에서 한정하지 않는다.
- [0030] 실시예들은 퍼스널 컴퓨터, 랩톱 컴퓨터, 태블릿 컴퓨터, 스마트 폰, 텔레비전, 스마트 가전 기기, 지능형 자동차, 키오스크, 웨어러블 장치 등 다양한 형태의 제품으로 구현될 수 있다. 이하, 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0032] 이하에서, 도 1 내지 도 10을 참조하여 본 발명의 몇몇 실시예에 따른 기계학습 모델의 이상 데이터 추출 방법에 대해 설명한다.
- [0033] 도 1은 본 발명의 몇몇 실시예에 따른 이상 데이터 추출 방법이 적용되는 장치를 설명하기 위한 도면이다.
- [0034] 기계학습(Machine Learning)이란 인간의 학습 능력을 컴퓨터 등 기계에 구현하는 학습 방식으로, 기계학습을 구현하는 방법은 신경망(Neural Network), 데이터마이닝(Data Mining), 의사결정 트리(Decision Tree), 유전알고리즘(Genetic Algorithm), 사례기반 추론(Case Based Reasoning), 패턴 인식(Pattern Recognition), 강화 학습(Reinforcement learning), 딥 러닝(Deep Learning) 등이 포함될 수 있다. 본 명세서에서 기계학습으로 표현되는 학습은 특정한 구현 방식이 아닌 상기의 방식들 중 어느 하나로 구현될 수 있다.
- [0035] 인공 신경망이란 사람의 두뇌가 정보를 처리하는 방식을 모방한 연산 시스템이다.
- [0036] 심층 신경망(Deep Neural Network)은 인공 신경망을 구현하는 하나의 방식으로, 복수의 레이어(layer)를 포함할 수 있다. 예를 들어, 심층 신경망은 입력 데이터가 인가되는 입력 레이어(Input Layer), 학습을 바탕으로 입력 데이터에 기반한 예측을 통해 도출된 결과 값을 출력하는 출력 레이어(Output Layer) 및 입력 레이어와 출력 레이어 사이에 다층의 은닉 레이어(Hidden Layer)를 포함한다.
- [0037] 심층 신경망은 정보를 처리하기 위해 이용되는 알고리즘에 따라, 컨볼루션 신경망(Convolutional Neural Network), 리커런트 신경망(Recurrent Neural Network) 등으로 분류된다.
- [0038] 인공 신경망을 학습하는 방식을 딥러닝(Deep Learning)이라 하며, 상술한 바와 같이 딥러닝에는 컨볼루션 신경망, 리커런트 신경망 방식과 같이 다양한 알고리즘이 이용될 수 있다.
- [0039] 이 때, 인공 신경망을 학습한다는 것은 계층간 가중치 및 바이어스 또는 인접한 계층 중 서로 다른 레이어에 속하는 복수의 뉴런들간의 가중치 및 바이어스를 결정하고 갱신하는 것을 나타낼 수 있다.

- [0040] 예를 들어, 복수의 계층적 구조 및 복수의 레이어들, 또는, 뉴런들 간의 가중치 및 바이어스를 총칭하여 인공 신경망의 연결성(connectivity)이라 할 수 있다. 따라서, 인공 신경망을 학습한다는 것은 연결성을 구축하고 학습하는 것을 나타낼 수 있다.
- [0041] 본 발명의 몇몇 실시예에 따른 기계학습 모델의 이상 데이터 추출 장치(10)는 뉴럴 네트워크를 생성하거나, 뉴럴 네트워크를 훈련(train)(또는 학습(learn))하거나, 뉴럴 네트워크를 재훈련(retrain)하는 기능들과 같은 다양한 프로세싱 기능들을 갖는 컴퓨팅 디바이스에 해당될 수 있다. 예를 들어, 기계학습 모델의 이상 데이터 추출 장치(10)는 PC(personal computer), 서버 디바이스, 모바일 디바이스 등의 다양한 종류의 디바이스들로 구현될 수 있다.
- [0042] 도 1을 참조하면, 본 발명의 몇몇 실시예에 따른 기계학습 모델의 이상 데이터 추출 장치(10)는 입력부(100), 판단부(200) 및 저장부(300)를 포함할 수 있고, 입력부(100)를 통해 입력되는 센싱 데이터(DT_SS)가 판단부(200)에 의해 판단(또는 학습)되고, 상기 판단 결과에 기초하여 이상 데이터(DT_ABN)라고 결정된 센싱 데이터인 타겟 데이터(DT_TG)의 로그 정보(LOG_TG)를 저장부(300)에 저장하고, 이를 이용하여 기계학습 모델의 성능을 향상시킬 수 있다.
- [0043] 도시하지는 않았으나, 저장부(300)는 타겟 데이터(DT_TG) 획득 당시의 시간, 타겟 데이터(DT_TG)에 대한 제1 결과 데이터(DT1), 타겟 데이터(DT_TG)의 재현율(또는 비교 데이터(COMP)) 등 로그 정보(LOG_TG)로 생성하는 로깅 모듈과, 생성된 로그 정보(LOG_TG)가 저장되는 데이터 저장소를 포함할 수 있다.
- [0044] 기계학습 모델의 이상 데이터 추출 장치(10)의 각각의 구성요소들(입력부(100), 판단부(200) 및 저장부(300))은 기능 및 논리적으로 분리될 수 있음을 나타내기 위해 별도로 도면에 표시한 것이며, 물리적으로 반드시 별도의 구성요소이거나 별도의 코드로 구현되는 것을 의미하는 것은 아니다.
- [0046] 도 2는 본 발명의 몇몇 실시예에 따른 판단부를 설명하기 위한 도면이고, 도 3은 본 발명의 몇몇 실시예에 따른 제2 기계학습 모델을 설명하기 위한 도면이고, 도 4는 본 발명의 몇몇 실시예에 따른 오토인코더를 설명하기 위한 도면이다.
- [0047] 도 2를 참조하면, 판단부(200)는 제1 기계학습 모델(210) 및 제2 기계학습 모델(220)을 포함할 수 있고, 제1 기계학습 모델(210) 및 제2 기계학습 모델(220)은 센싱 데이터(DT_SS)를 수신하고, 센싱 데이터(DT_SS)에 기초하여 각각 제1 결과 데이터(DT1) 및 비교 데이터(COMP)를 출력할 수 있다. 즉, 제1 및 제2 기계학습 모델(210, 220)은 동일한 학습 데이터인 센싱 데이터(DT_SS)로 학습될 수 있다.
- [0048] 제1 및 제2 기계학습 모델(210, 220)은 주어진 초기 뉴럴 네트워크를 반복적으로 훈련(학습)시킴으로써, 훈련된 뉴럴 네트워크를 생성할 수 있다. 훈련된 뉴럴 네트워크를 생성하는 것은 뉴럴 네트워크 파라미터를 결정하는 것을 의미할 수 있다. 여기서, 파라미터들은 예를 들어 뉴럴 네트워크의 입/출력 액티베이션들, 웨이트들, 바이어스들 등 뉴럴 네트워크에 입/출력되는 다양한 종류의 데이터를 포함할 수 있다. 뉴럴 네트워크의 반복적인 훈련이 진행됨에 따라, 뉴럴 네트워크의 파라미터들은 주어진 입력에 대해 보다 정확한 출력을 연산하기 위해 조정될(tuned) 수 있다.
- [0049] 몇몇 실시예에 따라, 판단부(200)는 모바일 디바이스, 임베디드(embedded) 디바이스 등에 포함될 수 있다. 몇몇 실시예에 따라, 판단부(200)는 뉴럴 네트워크의 구동을 위한 전용 하드웨어일 수 있다.
- [0050] 몇몇 실시예에 따라, 제1 기계학습 모델(210)은 센싱 데이터(DT_SS)에 기초하여 화재 발생여부를 판단할 수 있다. 제1 기계학습 모델(210)은 DNN, RNN, CNN 등과 같이 다양한 기계 학습 모델이 사용될 수 있으며, 본 발명이 적용되는 실시예가 특정 기계학습 모델로 제한되지 않는다. 제1 기계학습 모델(210)의 입력 계층(input layer)의 뉴런 수는 화재 감지 센서에서 입력받는 온도, 습도, 연기, 가스 등의 센서데이터 개수이며, 출력계층의 뉴런 개수는 화재 또는 비화재를 나타내는 뉴런 2개로 구성될 수 있다.
- [0051] 판단부(200)의 각각의 구성요소들(제1 기계학습 모델(210) 및 제2 기계학습 모델(220))은 기능 및 논리적으로 분리될 수 있음을 나타내기 위해 별도로 도면에 표시한 것이며, 물리적으로 반드시 별도의 구성요소이거나 별도의 코드로 구현되는 것을 의미하는 것은 아니다.
- [0052] 도 3 및 도 4를 참조하면, 본 발명의 몇몇 실시예에 따른 제2 기계학습 모델(220)은 오토인코더(221) 및 비교 모듈(222)을 포함할 수 있다. 몇몇 실시예에 따라, 오토인코더(221)는 입력부(100)에 의해 수신되는 센싱 데이터(DT_SS)에 기초하여 제2 결과 데이터(DT2)를 출력하고, 비교 모듈(222)은 제2 결과 데이터(DT2) 및 센싱 데이

터(DT_SS)에 기초하여 비교 데이터(COMP)를 출력할 수 있다.

- [0053] 제2 기계학습 모델(220)의 각각의 구성요소들(오토인코더(221) 및 비교 모듈(222))은 기능 및 논리적으로 분리될 수 있음을 나타내기 위해 별도로 도면에 표시한 것이며, 물리적으로 반드시 별도의 구성요소이거나 별도의 코드로 구현되는 것을 의미하는 것은 아니다.
- [0054] 오토인코더(221)는 입력을 출력으로 그대로 복사하는 비지도학습(Unsupervised Learning)의 일종인 신경망으로, 입력층(Input Layer)과 출력층(Output Layer) 사이의 은닉층(Hidden Layer)을 통해 데이터를 효율적으로 표현(representation)하는 방법을 학습하는 신경망이다. 즉, 오토인코더(221)는 센싱 데이터(DT_SS)를 입력으로 상기 센싱 데이터(DT_SS)를 재현한 제2 결과 데이터(DT2)를 생성하되, 센싱 데이터(DT_SS)와 제2 결과 데이터(DT2)는 재현율이 높은 유사 데이터이거나, 재현율이 낮은 비유사 데이터일 수 있다.
- [0055] 몇몇 실시예에 따라, 오토인코더(221)는 과소완전 적층 오토인코더(Undercomplete Stacked Autoencoder)로 구현될 수 있다. 입력층(Input Layer)은 제1 기계학습 모델(210)과 같이 화재 감지 센서에 의해 생성되는 센싱 데이터(예를 들어, 온도, 습도, 연기, 가스 등) 또는 상기 센싱 데이터의 전처리 데이터를 입력으로 제공받고, 출력층(Output Layer)은 입력층과 동일한 구조(뉴런의 개수 또는 출력되는 데이터의 종류)를 가질 수 있다.
- [0056] 몇몇 실시예에 따라, 제1 기계학습 모델(210) 및 오토인코더(221)는 동일한 학습 데이터를 이용해 훈련된다. 충분한 훈련이 되었다면, 제1 기계학습 모델(210)은 학습 데이터와 비슷한 상황에서는 올바른 판단을 할 수 있을 것이고, 오토인코더(221)는 학습 데이터와 비슷한 상황의 데이터가 입력되면 출력층에서 나오는 제2 결과 데이터(DT2)의 재현율이 높을 것이다. 본 발명에서는 제1 기계학습 모델(210)과 오토인코더(221)의 학습이 완료되었다는 가정을 하며, 제1 기계학습 모델(210) 및 오토인코더(221)의 학습 방법은 제한되지 않는다.
- [0057] 비교 모듈(222)은 센싱 데이터(DT_SS) 및 제2 결과 데이터(DT2)에 기초하여 비교 데이터(COMP)를 출력할 수 있다. 비교 데이터(COMP)는 비교 모듈(222)의 입력 데이터(센싱 데이터(DT_SS) 및 제2 결과 데이터(DT2))의 비교 결과에 대한 정보를 포함한다. 즉, 센싱 데이터(DT_SS) 및 제2 결과 데이터(DT2)의 유사도를 판단하고, 이에 기초하여 결정된 재현율에 대한 정보를 포함할 수 있다. 재현율이란 제2 결과 데이터(DT2)가 센싱 데이터(DT_SS)를 재현한 정도를 나타낸 수치일 수 있고, 이에 대한 판단 조건은 다양하게 설정될 수 있다. 즉, 재현율이 높다면 센싱 데이터(DT_SS)는 기학습된 학습 데이터와 유사한 데이터일 확률이 높고, 이에 따라 센싱 데이터(DT_SS)는 이상치가 낮은 정상 데이터일 확률이 높다. 따라서 제1 기계학습 모델(210)의 판단 결과인 제1 결과 데이터(DT1)의 신뢰도가 높음을 의미한다. 반대로, 재현율이 낮은 경우 센싱 데이터(DT_SS)는 외부 요인 등에 의해 이상치가 높은 이상 데이터일 확률이 높고, 따라서 제1 결과 데이터(DT1)의 신뢰도는 낮음을 의미한다. 본 명세서에서의 "이상 데이터(DT_ABN)"는, 재현율이 기정의된 임계값보다 낮은 데이터를 의미하고, "정상 데이터(DT_NM)"는 이상 데이터(DT_ABN) 이외의 데이터를 의미한다.
- [0058] 예를 들어, 제1 기계학습 모델(210)에 의해 출력된 제1 결과 데이터(DT1)에 "화재 발생"에 대한 정보가 포함되고, 제2 기계학습 모델(220)에 의해 출력된 비교 데이터(COMP)에 재현율이 높다는 정보가 포함되는 경우, "화재 발생"이라는 결과는 비화재보가 아닐 가능성이 높다. 이에 반해, 제1 기계학습 모델(210)에 의해 출력된 제1 결과 데이터(DT1)에 "화재 발생"에 대한 정보가 포함되고, 제2 기계학습 모델(220)에 의해 출력된 비교 데이터(COMP)에 재현율이 낮다는 정보가 포함되는 경우, "화재 발생"이라는 결과는 비화재보일 가능성이 높다.
- [0059] 또한, 재현율이 높은 센싱 데이터, 즉 정상 데이터(DT_NM)의 경우에는 제1 기계학습 모델(210)의 학습 데이터와 차이가 적으므로 추후 재학습(또는, 전이 학습)의 학습 데이터로 사용하더라도 제1 기계학습 모델(210)의 성능 향상에 큰 효과를 기대할 수 없다. 이에 반해, 재현율이 낮은 이상 데이터(DT_ABN)의 경우 제1 기계학습 모델(210)의 학습 데이터로 사용되는 경우 효과적인 학습이 가능하고, 이에 따라 제1 기계학습 모델의 성능을 향상시킬 수 있다.
- [0061] 도 5는 본 발명의 몇몇 실시예에 따른 입력부를 설명하기 위한 도면이다.
- [0062] 도 5를 참조하면, 입력부(100)는 센싱 데이터(DT_SS)를 수신하는 데이터 수집 모듈(110) 및 센싱 데이터(DT_SS)에 대한 전처리 동작을 수행하는 데이터 전처리 모듈(120)을 포함할 수 있다. 본 명세서에서의 "센싱 데이터(DT_SS)"는 전처리 과정을 거친 전처리 데이터(DT_PREP)를 의미할 수도 있다. 즉, 판단부(200)가 수신하는 데이터는 전처리 데이터(DT_PREP)일 수 있다.
- [0063] 입력부(100)의 각각의 구성요소들(데이터 수집 모듈(110) 및 데이터 전처리 모듈(120))은 기능 및 논리적으로

분리될 수 있음을 나타내기 위해 별도로 도면에 표시한 것이며, 물리적으로 반드시 별도의 구성요소가거나 별도의 코드로 구현되는 것을 의미하는 것은 아니다.

- [0065] 도 6은 본 발명의 몇몇 실시예에 따른 기계학습 모델의 이상 데이터 추출 방법을 설명하기 위한 순서도이다.
- [0066] 도 6을 참조하면, 본 발명의 몇몇 실시예에 따른 이상 데이터 추출 방법은, 외부로부터 입력되는 센싱 데이터(DT_SS)에 기초하여 제1 및 제2 결과 데이터(DT1, DT2)를 생성하고, 센싱 데이터(DT_SS)가 이상 데이터(DT_ABN)인지 여부를 판단하여 이상 데이터(DT_ABN)인 타겟 데이터(DT_TG)의 로그 정보(LOG_TG)를 저장할 수 있다.
- [0067] S1000 단계에서, 입력부(100)는 외부로부터 센싱 데이터(DT_SS)를 획득할 수 있다. 상술한 바와 같이, 센싱 데이터(DT_SS)는 화재 감지 센서에 의해 감지되는 데이터일 수 있다. 예를 들어, 센싱 데이터(DT_SS)는 온도, 습도, 연기량 및 가스량 중 적어도 하나의 데이터를 포함할 수 있다.
- [0068] S2000 단계에서, 센싱 데이터(DT_SS)를 입력으로 구동되는 제1 기계학습 모델(210)에서 제1 결과 데이터(DT1)를 출력할 수 있다. 센싱 데이터(DT_SS)는 입력부(100)에 의해 생성되는 전처리 데이터(DT_PREP)일 수 있다. 예를 들어, 전처리 데이터(DT_PREP)는 서로 다른 단위를 갖는 각각의 센싱 데이터(DT_SS)를 동일한 범위로 스케일링한 데이터일 수 있다. 제1 결과 데이터(DT1)는 센싱 데이터(DT_SS)가 화재를 나타내는지 또는 비화재를 나타내는지 판단한 데이터일 수 있다.
- [0069] S3000 단계에서, 센싱 데이터(DT_SS)를 입력으로 구동되는 제2 기계학습 모델(220)에서 제2 결과 데이터(DT2)를 출력할 수 있다. 이 때, 제2 결과 데이터(DT2)는 제2 기계학습 모델(220)에 포함되는 오토인코더(221)에 의해 출력되는 데이터일 수 있다.
- [0070] S4000 단계에서, 센싱 데이터(DT_SS)와 제2 결과 데이터(DT2)를 비교할 수 있다. 구체적으로, 제2 기계학습 모델(220)에 포함되는 비교 모듈(222)에서 상기 센싱 데이터(DT_SS)와 제2 결과 데이터(DT2)를 비교할 수 있다.
- [0071] S5000 단계에서, 센싱 데이터(DT_SS)가 이상 데이터(DT_ABN)인지 여부를 결정할 수 있다. 구체적으로, 비교 모듈(222)에서 결정된 비교 데이터(COMP)에 기초하여, 센싱 데이터(DT_SS)와 제2 결과 데이터(DT2)의 재현율을 결정하고, 상기 재현율에 기초하여 센싱 데이터(DT_SS)가 이상 데이터(DT_ABN)인지 여부를 결정할 수 있다. 즉, 비교 모듈(222)에서 출력되는 비교 데이터(COMP)는 센싱 데이터(DT_SS)가 이상 데이터(DT_ABN)인지 여부에 대한 정보를 포함할 수 있다.
- [0072] S6000 단계에서, 센싱 데이터(DT_SS)가 이상 데이터(DT_ABN)로 판단되는 경우, 타겟 데이터(DT_TG)의 로그 정보(LOG_TG)를 저장할 수 있다. 구체적으로, 이상 데이터(DT_ABN)로 결정된 센싱 데이터(DT_SS)를 타겟 데이터(DT_TG)로 정의하고, 타겟 데이터(DT_TG)의 로그 정보(LOG_TG)를 저장부(300)에 저장할 수 있다.
- [0073] 몇몇 실시예에 따라, 로그 정보(LOG_TG)는 i)타겟 데이터(DT_TG), ii)타겟 데이터(DT_TG)를 입력으로 제1 기계학습 모델(210)에 의해 출력된 제1 타겟 결과 데이터, iii)타겟 데이터(DT_TG)를 입력으로 제2 기계학습 모델(220)에 의해 출력된 제2 타겟 결과 데이터, iv)타겟 데이터(DT_TG)의 입력 시간 및 v)타겟 데이터(DT_TG)의 센싱 시간 중 적어도 하나의 정보를 포함할 수 있다.
- [0075] 도 7은 센싱 데이터와 제2 결과 데이터의 재현율에 기초하여 이상 데이터를 결정하는 실시예를 설명하기 위한 순서도이다.
- [0076] 도 6 및 도 7을 참조하면, 센싱 데이터(DT_SS)와 제2 결과 데이터(DT2)의 재현율에 기초하여 타겟 데이터(DT_TG)를 결정할 수 있다.
- [0077] 도 6의 S4000 단계는, 센싱 데이터(DT_SS)와 제2 결과 데이터(DT2)를 비교하여 재현율을 결정하는 단계(S4100)를 포함할 수 있다.
- [0078] 이후 S5100 단계에서, 재현율이 기정의된 임계값 미만인지 여부를 결정하고, S5200 단계에서, 재현율이 임계값 미만인 경우 센싱 데이터(DT_SS)를 이상 데이터(DT_ABN)로 결정할 수 있다. 다시 말해서, 센싱 데이터(DT_SS)를 타겟 데이터(DT_TG)로 결정할 수 있다. 임계값은 오토인코더(221)의 재현율 0% 내지 100% 사이의 값으로 기정의될 수 있다.

- [0079] 재현율이 임계값 이상인 경우, S5300 단계에서, 센싱 데이터(DT_SS)는 정상 데이터(DT_NM)로 결정될 수 있다. 즉, 재현율이 높다는 것은 외부 요인에 의해 노이즈가 발생하지 않은 데이터일 확률이 높은 바, 제1 기계학습 모델(210)의 결과를 신뢰할 수 있으므로 센싱 데이터(DT_SS)를 정상 데이터(DT_NM)로 결정한다.
- [0080] 이후, 센싱 데이터(DT_SS)가 타겟 데이터(DT_TG)인 경우 타겟 데이터(DT_TG)의 로그 정보(LOG_TG)를 저장부에 저장하고(S6000), 센싱 데이터(DT_SS)가 타겟 데이터(DT_TG)가 아닌 경우 절차를 종료할 수 있다.
- [0082] 도 8은 타겟 데이터의 로그 정보를 이용하여 제1 기계학습 모델의 전이 학습을 수행하는 실시예를 설명하기 위한 순서도이고, 도 9는 도 8의 실시예를 설명하기 위한 블록도이다.
- [0083] 도 8 및 도 9를 참조하면, 로그 정보(LOG_TG)를 이용하여 제1 기계학습 모델(210)의 전이 학습(Transfer Learning)을 수행하는 단계(S7000)를 더 포함할 수 있다. 즉, 로그 정보(LOG_TG)는 제1 기계학습 모델(210)의 재학습 수행 시 학습 데이터(training data)로서 활용될 수 있다. 예를 들어, 타겟 데이터(DT_TG)가 센싱된 시간의 온풍기 사용 여부, 창문 개폐 여부, 가스레인지 등의 사용 여부 등을 조사하여 타겟 데이터(DT_TG)가 특수 환경에 의한 오류로 발생한 데이터인지 여부를 분석하고, 분석 결과를 반영하여 제1 기계학습 모델(210)의 학습 데이터로 활용할 수 있다.
- [0084] 전이 학습이란, 학습 데이터가 부족한 분야의 모델 구축을 위해 데이터가 풍부한 분야에서 훈련된 모델을 재사용하는 학습 기법으로, 본 발명에서는 이상 데이터(DT_ABN)로 결정된 타겟 데이터(DT_TG)의 로그 정보(LOG_TG)를 수집하고, 수집된 정보를 전이 학습의 학습 데이터로 사용할 수 있다. 본 발명의 실시예에 따라 타겟 데이터(DT_TG)의 로그 정보(LOG_TG)를 전이 학습의 학습 데이터로 사용함에 따라, 기존에 학습된 학습 데이터와 유사하지 않은 학습 데이터로 제1 기계학습 모델(210)이 학습되고, 이에 따라 더욱 효율적인 학습이 가능하게 된다. 즉, 기계학습 모델의 성능 향상에 영향을 덜 미치는 데이터(예를 들어, 정상 데이터(DT_NM))는 배제하고, 재학습시 성능 향상에 도움이 되는 데이터(예를 들어, 이상 데이터(DT_ABN) 또는 타겟 데이터(DT_TG))를 학습 데이터로 사용함으로써, 재학습 효율이 증대될 수 있다.
- [0086] 도 10은 센싱 데이터 및 제1 결과 데이터에 기초하여 로그 정보의 저장 여부를 결정하는 실시예를 설명하기 위한 순서도이다.
- [0087] 도 10을 참조하면, 이상 데이터(DT_ABN)로 판단된 센싱 데이터(DT_SS) 중, 제1 결과 데이터(DT1)가 "화재" 정보를 포함하는 센싱 데이터(DT_SS)만을 타겟 데이터(DT_TG)로 결정하고, 상기 타겟 데이터(DT_TG)의 로그 정보(LOG_TG)를 저장할 수 있다.
- [0088] 보다 구체적으로, 제2 기계학습 모델(220)에 의해 생성되는 비교 데이터(COMP)가 제2 결과 데이터(DT2)의 재현율이 임계값 미만이라는 정보를 포함하고, 제1 결과 데이터(DT1)가 "화재"의 정보를 포함하는 경우에만 해당 센싱 데이터(DT_SS)를 타겟 데이터(DT_TG)로 결정할 수 있다. 화재로 판단되는 경우의 센싱 데이터(DT_SS)의 로그 정보(LOG_TG)만을 저장함으로써, 비화재의 발생 가능성을 감소시키고, 제1 기계학습 모델(210)의 재학습 효율을 증대시킬 수 있다.
- [0090] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 컨트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

- [0091] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0092] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.
- [0093] 본 문서에서 사용된 용어 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어로 구현된 유닛을 포함할 수 있으며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로와 같은 용어와 상호 호환적으로 사용될 수 있다. 모듈은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는, 상기 부품의 최소 단위 또는 그 일부가 될 수 있다. 예를 들면, 일 실시 예에 따르면, 모듈은 ASIC(application-specific integrated circuit)의 형태로 구현될 수 있다.
- [0094] 본 문서의 다양한 실시 예들은 기기(machine)에 저장된 하나 이상의 명령어들을 포함하는 소프트웨어로서 구현될 수 있다. 예를 들면, 기기는 저장 매체로부터 저장된 하나 이상의 명령어들 중 적어도 하나의 명령을 호출하고, 그것을 실행할 수 있다. 이것은 기기가 상기 호출된 적어도 하나의 명령어에 따라 적어도 하나의 기능을 수행하도록 운영되는 것을 가능하게 한다. 상기 하나 이상의 명령어들은 컴파일러에 의해 생성된 코드 또는 인터프리터에 의해 실행될 수 있는 코드를 포함할 수 있다. 기기로 읽을 수 있는 저장 매체는, 비일시적(non-transitory) 저장 매체의 형태로 제공될 수 있다. 여기서, '비일시적'은 저장 매체가 실재(tangible)하는 장치이고, 신호(signal)(예: 전자기파)를 포함하지 않는다는 것을 의미할 뿐이며, 이 용어는 데이터가 저장 매체에 반영구적으로 저장되는 경우와 임시적으로 저장되는 경우를 구분하지 않는다.
- [0095] 일 실시 예에 따르면, 본 문서에 개시된 다양한 실시 예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다. 컴퓨터 프로그램 제품은 상품으로서 판매자 및 구매자 간에 거래될 수 있다. 컴퓨터 프로그램 제품은 기기로 읽을 수 있는 저장 매체(예: compact disc read only memory(CD-ROM))의 형태로 배포되거나, 또는 어플리케이션 스토어를 통해 또는 두 개의 사용자 장치들(예: 스마트폰들) 간에 직접, 온라인으로 배포(예: 다운로드 또는 업로드)될 수 있다. 온라인 배포의 경우에, 컴퓨터 프로그램 제품의 적어도 일부는 제조사의 서버, 어플리케이션 스토어의 서버, 또는 중계 서버의 메모리와 같은 기기로 읽을 수 있는 저장 매체에 적어도 일시 저장되거나, 임시적으로 생성될 수 있다.
- [0096] 다양한 실시 예들에 따르면, 상기 기술한 구성요소들의 각각의 구성요소(예: 모듈 또는 프로그램)는 단수 또는 복수의 개체를 포함할 수 있다. 다양한 실시 예들에 따르면, 전술한 해당 구성요소들 중 하나 이상의 구성요소들 또는 동작들이 생략되거나, 또는 하나 이상의 다른 구성요소들 또는 동작들이 추가될 수 있다. 대체적으로 또는 추가적으로, 복수의 구성요소들(예: 모듈 또는 프로그램)은 하나의 구성요소로 통합될 수 있다. 이런 경우, 통합된 구성요소는 상기 복수의 구성요소들 각각의 구성요소의 하나 이상의 기능들을 상기 통합 이전에 상기 복수의 구성요소들 중 해당 구성요소에 의해 수행되는 것과 동일 또는 유사하게 수행할 수 있다. 다양한 실시 예들에 따르면, 모듈, 프로그램 또는 다른 구성요소에 의해 수행되는 동작들은 순차적으로, 병렬적으로, 반복적으로, 또는 휴리스틱하게 실행되거나, 상기 동작들 중 하나 이상이 다른 순서로 실행되거나, 생략되거나, 또는 하나 이상의 다른 동작들이 추가될 수 있다.

부호의 설명

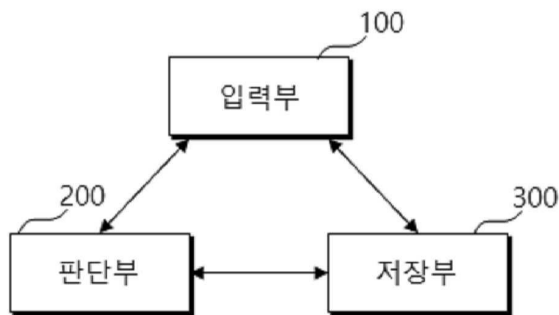
[0098]

100: 입력부 110: 데이터 수집 모듈
120: 데이터 전처리 모듈 200: 판단부
210: 제1 기계학습 모델 220: 제2 기계학습 모델
221: 오토인코더 222: 비교 모듈
300: 저장부

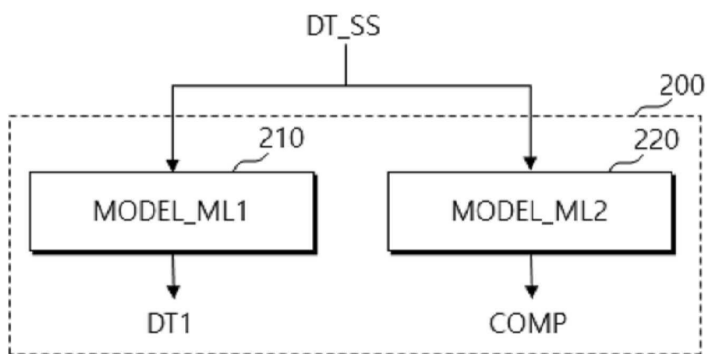
도면

도면1

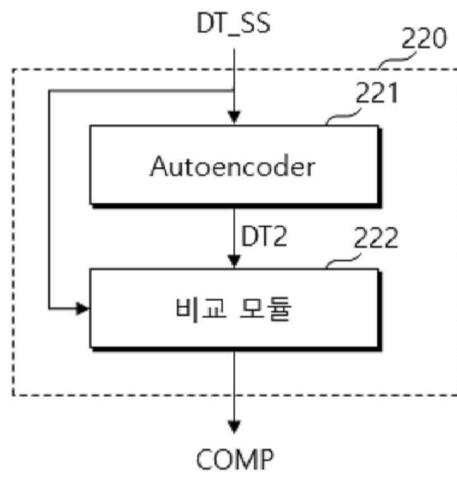
10



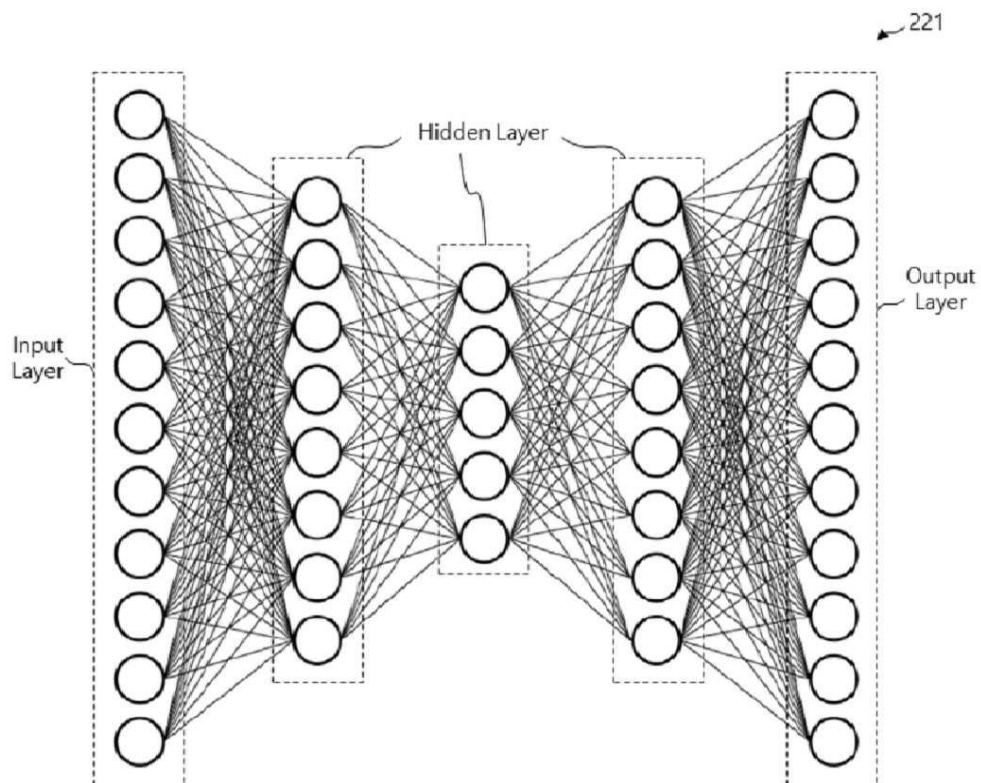
도면2



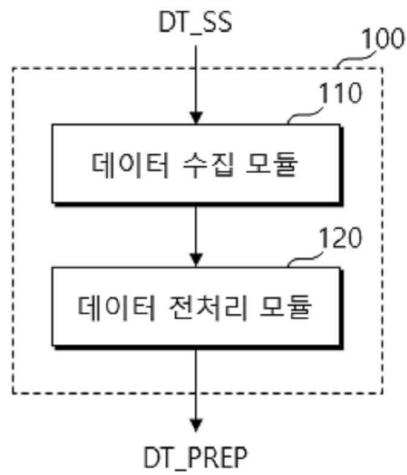
도면3



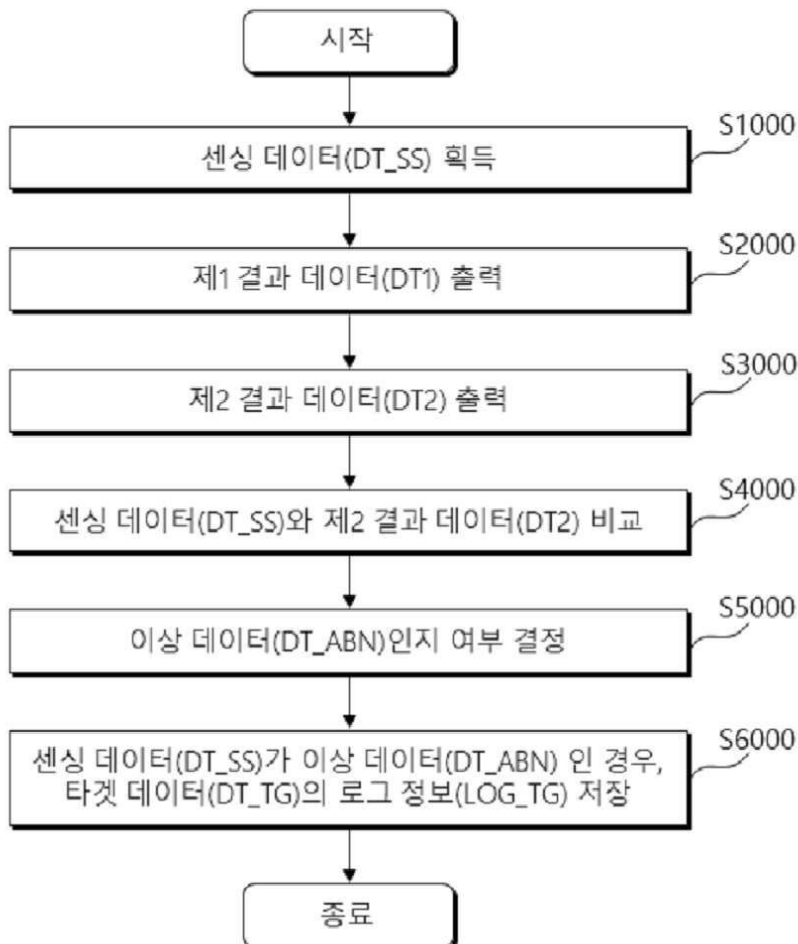
도면4



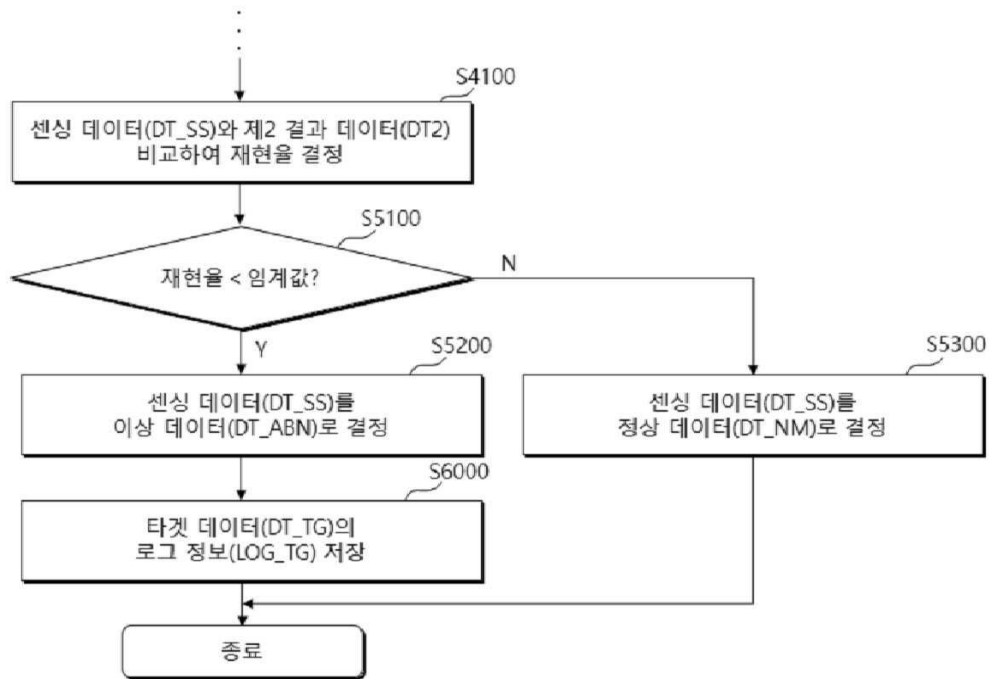
도면5



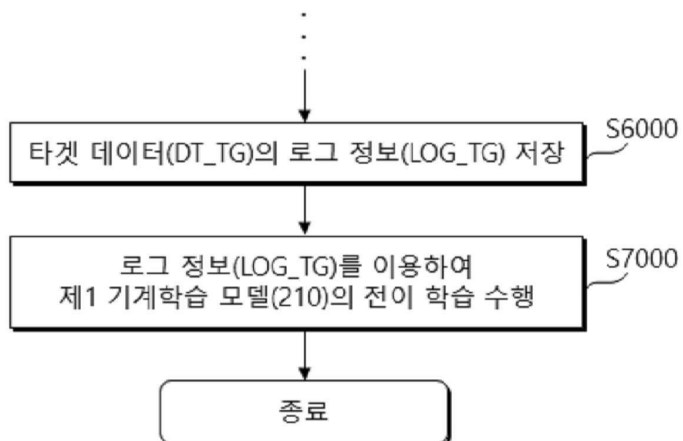
도면6



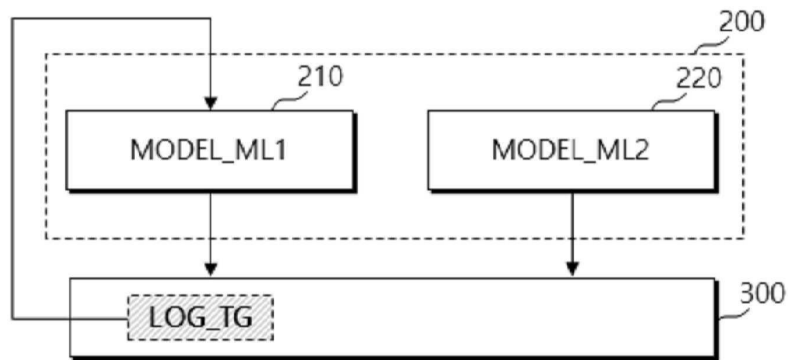
도면7



도면8



도면9



도면10

