



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0086992
(43) 공개일자 2025년06월16일

(51) 국제특허분류(Int. Cl.)

G16C 20/70 (2019.01) G06N 3/0442 (2023.01)
G06N 3/092 (2023.01) G16C 20/50 (2019.01)
G16C 20/90 (2019.01) H04L 9/32 (2006.01)

(52) CPC특허분류

G16C 20/70 (2019.02)
G06N 3/0442 (2023.01)

(21) 출원번호 10-2023-0176323

(22) 출원일자 2023년12월07일

심사청구일자 2023년12월07일

(71) 출원인

인천대학교 산학협력단

인천광역시 연수구 갯벌로 27, 인천대학교 이노베이션센터(송도동)

(72) 발명자

김지범

경기도 광명시 광명역로 26 광명역파크자이아파트 102동 1701호

박진영

인천광역시 부평구 이규보로 79 주공뜨란채아파트 105동 301호

(74) 대리인

특허법인시공

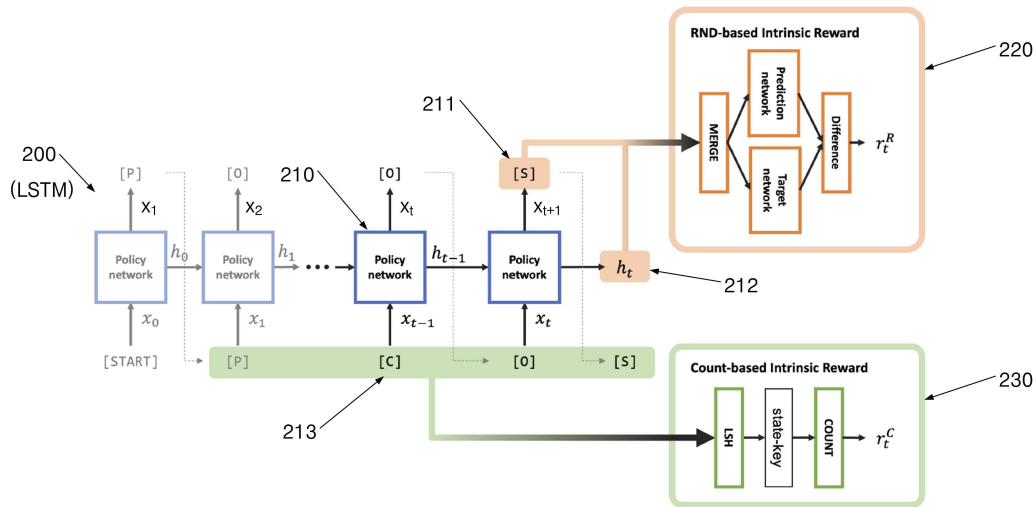
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 보상 기반의 강화 학습을 통해 마련된 신규 분자식 생성 모델에 기초하여 신규 분자식을 생성할 수 있는 전자 장치 및 그 동작 방법

(57) 요약

본 발명은 보상 기반의 강화 학습을 통해 마련된 신규 분자식 생성 모델에 기초하여 신규 분자식을 생성할 수 있는 전자 장치 및 그 동작 방법을 제시함으로써, 사용자가 보다 간편하게 다양한 신규 분자식을 획득할 수 있도록 지원할 수 있다.

대표도



(52) CPC특허분류

G06N 3/092 (2023.01)

G16C 20/50 (2019.02)

G16C 20/90 (2019.02)

H04L 9/3226 (2013.01)

H04L 9/3236 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711180669
과제번호	2022R1A4A5034121
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	집단연구지원
연구과제명	첨단 바이오소재 기반 뇌졸중 치료 및 기전 연구
기 여 율	1/2
과제수행기관명	인천대학교
연구기간	2023.03.01 ~ 2024.02.29

이 발명을 지원한 국가연구개발사업

과제고유번호	1711198471
과제번호	00260175
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	AI+X융합인재 연구교육과정 개발 및 운영
기 여 율	1/2
과제수행기관명	인천대학교산학협력단
연구기간	2023.07.01 ~ 2023.12.31

명세서

청구범위

청구항 1

사전 설정된 분자식 표기법에서 정의하고 있는 복수의 문자들 각각에 대한 원-핫 벡터가 저장되어 있는 문자 저장부;

사전 설정된 시작 토큰을, LSTM(Long Short-Term Memory) - 상기 LSTM을 구성하는 각 타임 스텝은, 이전 타임 스텝의 출력 토큰을 입력 토큰으로 인가받아, 상기 복수의 문자들 각각에 대응되는 분류(classification) 확률을 생성한 후, 상기 복수의 문자들 중, 각 분류 확률에 따라 랜덤하게 선택된 어느 하나의 문자에 대한 원-핫 벡터를, 당해 타임 스텝의 출력 토큰으로 산출하는, 신경망 기반의 정책 네트워크(Policy Network)로 구성되어 있음 - 의 첫 번째 타임 스텝에 입력으로 인가한 후, 상기 LSTM의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 당해 타임 스텝에서 산출된 출력 토큰과 히든 스테이트(hidden state)를, RND(Random Network Distillation) 기반의 내적 보상(intrinsic reward)을 산출하기 위한 제1 보상층(reward layer)에 입력으로 인가하여, 제1 내적 보상을 산출하고, 상기 첫 번째 타임 스텝부터 당해 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 문자열을, 카운트 기반(Count-based)의 내적 보상을 산출하기 위한 제2 보상층에 입력으로 인가하여, 제2 내적 보상을 산출한 후, 상기 제1 내적 보상과 상기 제2 내적 보상을 합산함으로써, 상기 LSTM의 각 타임 스텝별 보상을 산출하는 보상 산출부;

상기 LSTM의 각 타임 스텝별 보상의 산출이 완료되면, 상기 LSTM의 각 타임 스텝별 보상을 모두 합산하여 합산 보상을 산출한 후, 상기 합산 보상이 최대가 되도록 상기 LSTM에 대한 강화 학습(Reinforcement Learning)을 수행함으로써, 분자식 생성 모델을 생성하는 모델 생성부; 및

상기 분자식 생성 모델의 생성이 완료된 이후에, 사용자에게 의해, 신규 분자식의 생성 명령이 인가되면, 상기 분자식 생성 모델을 구성하는 상기 LSTM의 첫 번째 타임 스텝에 상기 시작 토큰을 입력으로 인가함으로써, 상기 LSTM을 구성하는 각 타임 스텝별로 산출되는 출력 토큰에 대응되는 문자들로 구성되는 문자열을, 신규 분자식으로 생성하는 분자식 생성부

를 포함하는 전자 장치.

청구항 2

제1항에 있어서,

상기 분자식 표기법은

SMILES(Simplified Molecular Input Line Entry System) 또는 SELFIES(Self-Referencing Embedded Strings)인 것을 특징으로 하는 전자 장치.

청구항 3

제1항에 있어서,

상기 보상 산출부는

상기 LSTM의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 상기 제1 내적 보상과 상기 제2 내적 보상을 산출한 후, 당해 타임 스텝에서 산출된 출력 토큰에 대응하는 문자를, 사전 설정된 단백질과 문자 간의 연관도 스코어를 외적 보상(Extrinsic reward)으로 산출하는 외적 보상층에 입력으로 인가하여 외적 보상을 산출한 다음, 상기 제2 내적 보상에 상기 외적 보상을 합산함으로써 상기 제2 내적 보상을 보정하고, 상기 제2 내적 보상의 보정이 완료되면, 상기 제1 내적 보상과 상기 보정된 제2 내적 보상을 합산함으로써, 상기 LSTM의 각 타임 스텝별 보상을 산출하는 것을 특징으로 하는 전자 장치.

청구항 4

제1항에 있어서,

상기 보상 산출부는

상기 LSTM을 구성하는 타임 스텝들 중 어느 하나인 제1 타임 스텝에서의 보상을 산출해야 하는 순서가 되는 경우, 상기 제1 타임 스텝에서 산출된 출력 토큰과 히든 스테이트를, 상기 제1 보상층을 구성하는 신경망 기반의 예측 네트워크(Prediction Network)와 타겟 네트워크(Target Network)에 각각 입력으로 인가하여, 상기 예측 네트워크로부터 산출되는 출력과 상기 타겟 네트워크로부터 산출되는 출력 간의 오차를, 상기 제1 타임 스텝에서의 상기 제1 내적 보상으로 산출하고,

상기 첫 번째 타임 스텝부터 상기 제1 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 제1 문자열을 확인하고, 상기 제1 문자열을, 상기 제2 보상층을 구성하는 LSH(Locality Sensitive Hashing)층에 통과시켜, 상기 제1 문자열에 대해 LSH를 수행함으로써 제1 해시 데이터를 생성한 다음, 상기 첫 번째 타임 스텝부터 상기 제1 타임 스텝까지의 각 타임 스텝별로, 보상을 산출하는 과정에서 상기 LSH가 수행됨에 따라 생성된 해시 데이터들 중, 상기 제1 해시 데이터가 등장하는 빈도수를 카운트하여, 카운트된 빈도수의 제곱근의 역수를, 상기 제1 타임 스텝에서의 상기 제2 내적 보상으로 산출하는 것을 특징으로 하는 전자 장치.

청구항 5

제4항에 있어서,

상기 모델 생성부는

상기 합산 보상이 최대가 되도록 상기 LSTM에 대한 강화 학습을 수행함과 동시에, 상기 제1 보상층을 구성하는 상기 예측 네트워크에 대한 강화 학습도 함께 수행하는 것을 특징으로 하는 전자 장치.

청구항 6

제1항에 있어서,

상기 모델 생성부를 통해, 상기 분자식 생성 모델의 생성이 완료된 이후에, 상기 사용자로부터 상기 분자식 생성 모델의 이용을 위한 비밀번호 설정 명령이 인가되면, p (p 는 2이상의 자연수임)자리수의 제1 번호와 q (q 는 2이상의 자연수임)자리수의 제2 번호를 랜덤하게 생성하고, 상기 제1 번호를 구성하는 각 자리 숫자를 모듈로 2 연산한 값을 성분으로 갖는 벡터와 상기 제2 번호를 구성하는 각 자리 숫자를 모듈로 2 연산한 값을 성분으로 갖는 벡터 간의 합성곱을 수행하여 결과 벡터를 산출한 후, 상기 결과 벡터의 해밍 무게를 사전 설정된 해시 함수에 입력으로 인가하였을 때 산출되는 해시 값을, 사용자 인증 값으로 지정하여, 사전 지정된 클라우드 저장 서버에 저장하는 저장 처리부;

상기 제1 번호와 상기 제2 번호를 서로 연결한 번호를, 상기 분자식 생성 모델의 이용을 위한 제1 비밀번호로서 생성한 후, 상기 제1 비밀번호를 화면 상에 표시하면서, 상기 제1 비밀번호를 숙지할 것을 지시하는 메시지를 화면 상에 표시하는 표시부;

상기 제1 비밀번호와 상기 메시지가 화면 상에 표시된 이후, 상기 사용자로부터 상기 분자식 생성 모델의 이용을 위한 인증 요청 명령이 인가되면, 화면 상에, 인증을 위한 비밀번호를 입력할 것을 지시하는 메시지를 표시한 후, 상기 사용자에 의해, 상기 제1 비밀번호가 입력되면, 상기 제1 비밀번호로부터 상기 제1 번호와 상기 제2 번호를 분리한 후, 상기 제1 번호를 구성하는 각 자리 숫자를 모듈로 2 연산한 값을 성분으로 갖는 벡터와 상기 제2 번호를 구성하는 각 자리 숫자를 모듈로 2 연산한 값을 성분으로 갖는 벡터 간의 합성곱을 수행하여 상기 결과 벡터를 산출한 다음, 상기 결과 벡터의 해밍 무게를 상기 해시 함수에 인가하여 상기 해시 값을 연산한 후, 상기 해시 값이 상기 클라우드 저장 서버에 저장되어 있는 상기 사용자 인증 값과 일치하는지 확인하여, 서로 일치하는 것으로 확인되면, 상기 사용자에 대한 인증을 완료 처리하는 인증 완료부; 및

상기 사용자에 대한 인증이 완료 처리되면, 화면 상에, 인증이 완료되었음을 지시하는 메시지를 표시함과 동시에, 상기 사용자로 하여금 상기 신규 분자식의 생성 명령을 인가할 것을 안내하는 안내 메시지를 표시하는 안내 메시지 표시부

를 더 포함하는 전자 장치.

청구항 7

사전 설정된 분자식 표기법에서 정의하고 있는 복수의 문자들 각각에 대한 원-핫 벡터가 저장되어 있는 문자 저장부를 유지하는 단계;

사전 설정된 시작 토큰을, LSTM(Long Short-Term Memory) - 상기 LSTM을 구성하는 각 타임 스텝은, 이전 타임 스텝의 출력 토큰을 입력 토큰으로 인가받아, 상기 복수의 문자들 각각에 대응되는 분류(classification) 확률을 생성한 후, 상기 복수의 문자들 중, 각 분류 확률에 따라 랜덤하게 선택된 어느 하나의 문자에 대한 원-핫 벡터를, 당해 타임 스텝의 출력 토큰으로 산출하는, 신경망 기반의 정책 네트워크(Policy Network)로 구성되어 있음 - 의 첫 번째 타임 스텝에 입력으로 인가한 후, 상기 LSTM의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 당해 타임 스텝에서 산출된 출력 토큰과 히든 스테이트(hidden state)를, RND(Random Network Distillation) 기반의 내적 보상(intrinsic reward)을 산출하기 위한 제1 보상층(reward layer)에 입력으로 인가하여, 제1 내적 보상을 산출하고, 상기 첫 번째 타임 스텝부터 당해 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 문자열을, 카운트 기반(Count-based)의 내적 보상을 산출하기 위한 제2 보상층에 입력으로 인가하여, 제2 내적 보상을 산출한 후, 상기 제1 내적 보상과 상기 제2 내적 보상을 합산함으로써, 상기 LSTM의 각 타임 스텝별 보상을 산출하는 단계;

상기 LSTM의 각 타임 스텝별 보상의 산출이 완료되면, 상기 LSTM의 각 타임 스텝별 보상을 모두 합산하여 합산 보상을 산출한 후, 상기 합산 보상이 최대가 되도록 상기 LSTM에 대한 강화 학습(Reinforcement Learning)을 수행함으로써, 분자식 생성 모델을 생성하는 단계; 및

상기 분자식 생성 모델의 생성이 완료된 이후에, 사용자에게 의해, 신규 분자식의 생성 명령이 인가되면, 상기 분자식 생성 모델을 구성하는 상기 LSTM의 첫 번째 타임 스텝에 상기 시작 토큰을 입력으로 인가함으로써, 상기 LSTM을 구성하는 각 타임 스텝별로 산출되는 출력 토큰에 대응되는 문자들로 구성되는 문자열을, 신규 분자식으로 생성하는 단계

를 포함하는 전자 장치의 동작 방법.

청구항 8

제7항에 있어서,

상기 분자식 표기법은

SMILES(Simplified Molecular Input Line Entry System) 또는 SELFIES(Self-Referencing Embedded Strings)인 것을 특징으로 하는 전자 장치의 동작 방법.

청구항 9

제7항에 있어서,

상기 보상을 산출하는 단계는

상기 LSTM의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 상기 제1 내적 보상과 상기 제2 내적 보상을 산출한 후, 당해 타임 스텝에서 산출된 출력 토큰에 대응하는 문자를, 사전 설정된 단백질과 문자 간의 연관도 스코어를 외적 보상(Extrinsic reward)으로 산출하는 외적 보상층에 입력으로 인가하여 외적 보상을 산출한 다음, 상기 제2 내적 보상에 상기 외적 보상을 합산함으로써 상기 제2 내적 보상을 보정하고, 상기 제2 내적 보상의 보정이 완료되면, 상기 제1 내적 보상과 상기 보정된 제2 내적 보상을 합산함으로써, 상기 LSTM의 각 타임 스텝별 보상을 산출하는 것을 특징으로 하는 전자 장치의 동작 방법.

청구항 10

제7항에 있어서,

상기 보상을 산출하는 단계는

상기 LSTM을 구성하는 타임 스텝들 중 어느 하나인 제1 타임 스텝에서의 보상을 산출해야 하는 순서가 되는 경우, 상기 제1 타임 스텝에서 산출된 출력 토큰과 히든 스테이트를, 상기 제1 보상층을 구성하는 신경망 기반의 예측 네트워크(Prediction Network)와 타겟 네트워크(Target Network)에 각각 입력으로 인가하여, 상기 예측 네트워크로부터 산출되는 출력과 상기 타겟 네트워크로부터 산출되는 출력 간의 오차를, 상기 제1 타임 스텝에서의 상기 제1 내적 보상으로 산출하고,

상기 첫 번째 타임 스텝부터 상기 제1 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 제1 문자열을 확인하고, 상기 제1 문자열을, 상기 제2 보상층을 구성하는 LSH(Locality Sensitive Hashing)층에 통과시켜, 상기 제1 문자열에 대해 LSH를 수행함으로써 제1 해시 데이터를 생성한 다음, 상기 첫 번째 타임 스텝부터 상기 제1 타임 스텝까지의 각 타임 스텝별로, 보상을 산출하는 과정에서 상기 LSH가 수행됨에 따라 생성된 해시 데이터들 중, 상기 제1 해시 데이터가 등장하는 빈도수를 카운트하여, 카운트된 빈도수의 제곱근의 역수를, 상기 제1 타임 스텝에서의 상기 제2 내적 보상으로 산출하는 것을 특징으로 하는 전자 장치의 동작 방법.

청구항 11

제10항에 있어서,

상기 분자식 생성 모델을 생성하는 단계는

상기 합산 보상이 최대가 되도록 상기 LSTM에 대한 강화 학습을 수행함과 동시에, 상기 제1 보상층을 구성하는 상기 예측 네트워크에 대한 강화 학습도 함께 수행하는 것을 특징으로 하는 전자 장치의 동작 방법.

청구항 12

제7항에 있어서,

상기 분자식 생성 모델을 생성하는 단계를 통해, 상기 분자식 생성 모델의 생성이 완료된 이후에, 상기 사용자로부터 상기 분자식 생성 모델의 이용을 위한 비밀번호 설정 명령이 인가되면, p (p 는 2이상의 자연수임)자리수의 제1 번호와 q (q 는 2이상의 자연수임)자리수의 제2 번호를 랜덤하게 생성하고, 상기 제1 번호를 구성하는 각 자리 숫자를 모듈로(modulo)-2 연산한 값을 성분으로 갖는 벡터와 상기 제2 번호를 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터 간의 합성곱을 수행하여 결과 벡터를 산출한 후, 상기 결과 벡터의 해밍 무게를 사전 설정된 해시 함수에 입력으로 인가하였을 때 산출되는 해시 값을, 사용자 인증 값으로 지정하여, 사전 지정된 클라우드 저장 서버에 저장하는 단계;

상기 제1 번호와 상기 제2 번호를 서로 연결한 번호를, 상기 분자식 생성 모델의 이용을 위한 제1 비밀번호로서 생성한 후, 상기 제1 비밀번호를 화면 상에 표시하면서, 상기 제1 비밀번호를 숙지할 것을 지시하는 메시지를 화면 상에 표시하는 단계;

상기 제1 비밀번호와 상기 메시지가 화면 상에 표시된 이후, 상기 사용자로부터 상기 분자식 생성 모델의 이용을 위한 인증 요청 명령이 인가되면, 화면 상에, 인증을 위한 비밀번호를 입력할 것을 지시하는 메시지를 표시한 후, 상기 사용자에 의해, 상기 제1 비밀번호가 입력되면, 상기 제1 비밀번호로부터 상기 제1 번호와 상기 제2 번호를 분리한 후, 상기 제1 번호를 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터와 상기 제2 번호를 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터 간의 합성곱을 수행하여 상기 결과 벡터를 산출한 다음, 상기 결과 벡터의 해밍 무게를 상기 해시 함수에 인가하여 상기 해시 값을 연산한 후, 상기 해시 값이 상기 클라우드 저장 서버에 저장되어 있는 상기 사용자 인증 값과 일치하는지 확인하여, 서로 일치하는 것으로 확인되면, 상기 사용자에 대한 인증을 완료 처리하는 단계; 및

상기 사용자에 대한 인증이 완료 처리되면, 화면 상에, 인증이 완료되었음을 지시하는 메시지를 표시함과 동시에, 상기 사용자로 하여금 상기 신규 분자식의 생성 명령을 인가할 것을 안내하는 안내 메시지를 표시하는 단계를 더 포함하는 전자 장치의 동작 방법.

청구항 13

제7항 내지 제12항 중 어느 한 항의 방법을 컴퓨터와의 결합을 통해 실행시키기 위한 컴퓨터 프로그램을 기록한 컴퓨터 판독 가능 기록 매체.

청구항 14

제7항 내지 제12항 중 어느 한 항의 방법을 컴퓨터와의 결합을 통해 실행시키기 위한 저장매체에 저장된 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 보상 기반의 강화 학습을 통해 마련된 신규 분자식 생성 모델에 기초하여 신규 분자식을 생성할 수 있는 전자 장치 및 그 동작 방법에 대한 것이다.

배경 기술

- [0002] 제약 기술의 발달로 치료 효과가 큰 약물들이 많이 개발되고 있다.
- [0003] 하지만, 최근에는 다양한 신종 질병이 등장함에 따라, 이러한 신종 질병에 대응하기 위한 빠른 치료약 개발의 필요성이 증가하고 있다.
- [0004] 빠른 치료약을 개발하기 위해서는 약물을 구성하는 새로운 분자식을 생성할 필요가 있다.
- [0005] 최근에는 인공지능 기반의 기계학습 기술이 등장하고 있다는 점에서, 신약 개발을 위한 새로운 분자식을 만드는 데 있어서도 이러한 인공지능 기술의 활용을 고려할 수 있다.
- [0006] 분자식은 SMILES(Simplified Molecular Input Line Entry System) 또는 SELFIES(Self-Referencing Embedded Strings)와 같은 소정의 분자식 표기법에 따라 소정의 문자열의 형태로 표기할 수 있다. 이로 인해, 새로운 분자식을 생성하는 분자식 생성 모델은, 소정의 문자열을 생성하는 모델이라고 볼 수 있고, 이에 따라, 분자식 생성 모델은 자연어 처리 분야에서 자주 사용되는 LSTM(Long Short-Term Memory)을 기반으로 구현할 수 있다. 즉, LSTM의 각 타임 스텝별로 산출되는 문자들로 구성되는 문자열을 새로운 분자식으로 산출하는 형태로, 분자식 생성 모델을 구현할 수 있다.
- [0007] 이때, 분자식 생성 모델은, 최대한 다양한 형태의 분자식이 만들어지도록 구성되어야 한다는 점에서, 학습 과정에서, 다양한 형태의 분자식이 생성되도록 하는 방향으로 학습이 수행되어야 한다. 이와 관련해서, 분자식 생성 모델을 학습시키는데 있어서, 강화 학습(Reinforcement Learning)을 이용하는 방안을 고려할 수 있다. 구체적으로, 다양한 형태의 분자식이 생성되도록 하는 방향으로 보상(reward)이 산출되도록 프레임워크를 구성하여, 이 보상이 최대가 되도록 강화 학습을 수행하는 기법을 고려할 수 있다.

선행기술문헌

비특허문헌

- [0008] (비특허문헌 0001) Luca A. Thiede et al., "Curiosity in exploring chemical spaces: intrinsic rewards for molecular reinforcement learning", arXiv:2012.11293, 2020.
- (비특허문헌 0002) Yuri Burda et al., "Exploration by Random Network Distillation", arXiv:1810.12894, 2018.

발명의 내용

해결하려는 과제

- [0009] 본 발명은 보상 기반의 강화 학습을 통해 마련된 신규 분자식 생성 모델에 기초하여 신규 분자식을 생성할 수 있는 전자 장치 및 그 동작 방법을 제시함으로써, 사용자가 보다 간편하게 다양한 신규 분자식을 획득할 수 있도록 지원하고자 한다.

과제의 해결 수단

- [0010] 본 발명의 일실시예에 따른 전자 장치는, 사전 설정된 분자식 표기법에서 정의하고 있는 복수의 문자들 각각에 대한 원-핫 벡터가 저장되어 있는 문자 저장부, 사전 설정된 시작 토큰을, LSTM(Long Short-Term Memory) - 상기 LSTM을 구성하는 각 타임 스텝은, 이전 타임 스텝의 출력 토큰을 입력 토큰으로 인가받아, 상기 복수의 문자들 각각에 대응되는 분류(classification) 확률을 생성한 후, 상기 복수의 문자들 중, 각 분류 확률에 따라 랜덤하게 선택된 어느 하나의 문자에 대한 원-핫 벡터를, 당해 타임 스텝의 출력 토큰으로 산출하는, 신경망 기반의 정책 네트워크(Policy Network)로 구성되어 있음 - 의 첫 번째 타임 스텝에 입력으로 인가한 후, 상기 LSTM의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 당해 타임 스텝에서 산출된 출력 토큰과 히든 스테이트

(hidden state)를, RND(Random Network Distillation) 기반의 내적 보상(intrinsic reward)을 산출하기 위한 제1 보상층(reward layer)에 입력으로 인가하여, 제1 내적 보상을 산출하고, 상기 첫 번째 타임 스텝부터 당해 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 문자열을, 카운트 기반(Count-based)의 내적 보상을 산출하기 위한 제2 보상층에 입력으로 인가하여, 제2 내적 보상을 산출한 후, 상기 제1 내적 보상과 상기 제2 내적 보상을 합산함으로써, 상기 LSTM의 각 타임 스텝별 보상을 산출하는 보상 산출부, 상기 LSTM의 각 타임 스텝별 보상의 산출이 완료되면, 상기 LSTM의 각 타임 스텝별 보상을 모두 합산하여 합산 보상을 산출한 후, 상기 합산 보상이 최대가 되도록 상기 LSTM에 대한 강화 학습(Reinforcement Learning)을 수행함으로써, 분자식 생성 모델을 생성하는 모델 생성부 및 상기 분자식 생성 모델의 생성이 완료된 이후에, 사용자에 의해, 신규 분자식의 생성 명령이 인가되면, 상기 분자식 생성 모델을 구성하는 상기 LSTM의 첫 번째 타임 스텝에 상기 시작 토큰을 입력으로 인가함으로써, 상기 LSTM을 구성하는 각 타임 스텝별로 산출되는 출력 토큰에 대응되는 문자들로 구성되는 문자열을, 신규 분자식으로 생성하는 분자식 생성부를 포함한다.

[0011] 또한, 본 발명의 일실시예에 따른 전자 장치의 동작 방법은, 사전 설정된 분자식 표기법에서 정의하고 있는 복수의 문자들 각각에 대한 원-핫 벡터가 저장되어 있는 문자 저장부를 유지하는 단계, 사전 설정된 시작 토큰을, LSTM - 상기 LSTM을 구성하는 각 타임 스텝은, 이전 타임 스텝의 출력 토큰을 입력 토큰으로 인가받아, 상기 복수의 문자들 각각에 대응되는 분류 확률을 생성한 후, 상기 복수의 문자들 중, 각 분류 확률에 따라 랜덤하게 선택된 어느 하나의 문자에 대한 원-핫 벡터를, 당해 타임 스텝의 출력 토큰으로 산출하는, 신경망 기반의 정책 네트워크로 구성되어 있음 - 의 첫 번째 타임 스텝에 입력으로 인가한 후, 상기 LSTM의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 당해 타임 스텝에서 산출된 출력 토큰과 히든 스테이트를, RND 기반의 내적 보상을 산출하기 위한 제1 보상층에 입력으로 인가하여, 제1 내적 보상을 산출하고, 상기 첫 번째 타임 스텝부터 당해 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 문자열을, 카운트 기반의 내적 보상을 산출하기 위한 제2 보상층에 입력으로 인가하여, 제2 내적 보상을 산출한 후, 상기 제1 내적 보상과 상기 제2 내적 보상을 합산함으로써, 상기 LSTM의 각 타임 스텝별 보상을 산출하는 단계, 상기 LSTM의 각 타임 스텝별 보상의 산출이 완료되면, 상기 LSTM의 각 타임 스텝별 보상을 모두 합산하여 합산 보상을 산출한 후, 상기 합산 보상이 최대가 되도록 상기 LSTM에 대한 강화 학습을 수행함으로써, 분자식 생성 모델을 생성하는 단계 및 상기 분자식 생성 모델의 생성이 완료된 이후에, 사용자에 의해, 신규 분자식의 생성 명령이 인가되면, 상기 분자식 생성 모델을 구성하는 상기 LSTM의 첫 번째 타임 스텝에 상기 시작 토큰을 입력으로 인가함으로써, 상기 LSTM을 구성하는 각 타임 스텝별로 산출되는 출력 토큰에 대응되는 문자들로 구성되는 문자열을, 신규 분자식으로 생성하는 단계를 포함한다.

발명의 효과

[0012] 본 발명은 보상 기반의 강화 학습을 통해 마련된 신규 분자식 생성 모델에 기초하여 신규 분자식을 생성할 수 있는 전자 장치 및 그 동작 방법을 제시함으로써, 사용자가 보다 간편하게 다양한 신규 분자식을 획득할 수 있도록 지원할 수 있다.

도면의 간단한 설명

[0013] 도 1은 본 발명의 일실시예에 따른 전자 장치의 구조를 도시한 도면이다.
 도 2는 본 발명의 일실시예에 따른 전자 장치의 동작을 설명하기 위한 도면이다.
 도 3은 본 발명의 일실시예에 따른 전자 장치의 동작 방법을 도시한 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0014] 이하에서는 본 발명에 따른 실시예들을 첨부된 도면을 참조하여 상세하게 설명하기로 한다. 이러한 설명은 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였으며, 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 본 명세서 상에서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 사람에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다.

[0015] 본 문서에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있다는 것을 의미한다. 또한, 본 발명의 다양한 실시예들에 있어서, 각 구성요소들, 기능 블록들 또는 수단들은 하나 또는 그 이상의 하부 구성요소로 구성될

수 있고, 각 구성요소들이 수행하는 전기, 전자, 기계적 기능들은 전자회로, 집적회로, ASIC(Application Specific Integrated Circuit) 등 공지된 다양한 소자들 또는 기계적 요소들로 구현될 수 있으며, 각각 별개로 구현되거나 2 이상이 하나로 통합되어 구현될 수도 있다.

[0016] 한편, 첨부된 블록도의 블록들이나 흐름도의 단계들은 범용 컴퓨터, 특수용 컴퓨터, 휴대용 노트북 컴퓨터, 네트워크 컴퓨터 등 데이터 프로세싱이 가능한 장비의 프로세서나 메모리에 탑재되어 지정된 기능들을 수행하는 컴퓨터 프로그램 명령들(instructions)을 의미하는 것으로 해석될 수 있다. 이들 컴퓨터 프로그램 명령들은 컴퓨터 장치에 구비된 메모리 또는 컴퓨터에서 판독 가능한 메모리에 저장될 수 있기 때문에, 블록도의 블록들 또는 흐름도의 단계들에서 설명된 기능들은 이를 수행하는 명령 수단을 내포하는 제조물로 생산될 수도 있다. 아울러, 각 블록 또는 각 단계는 특정된 논리적 기능(들)을 실행하기 위한 하나 이상의 실행 가능한 명령들을 포함하는 모듈, 세그먼트 또는 코드의 일부를 나타낼 수 있다. 또, 몇 가지 대체 가능한 실시예들에서는 블록들 또는 단계들에서 언급된 기능들이 정해진 순서와 달리 실행되는 것도 가능함을 주목해야 한다. 예컨대, 잇달아 도시되어 있는 두 개의 블록들 또는 단계들은 실질적으로 동시에 수행되거나, 역순으로 수행될 수 있으며, 경우에 따라 일부 블록들 또는 단계들이 생략된 채로 수행될 수도 있다.

[0017] 도 1은 본 발명의 일실시예에 따른 전자 장치의 구조를 도시한 도면이고, 도 2는 본 발명에 따른 분자식 생성 모델을 설명하기 위한 프레임워크를 도시한 도면이다.

[0018] 도 1을 참조하면, 본 발명에 따른 전자 장치(110)는 문자 저장부(111), 보상 산출부(112), 모델 생성부(113) 및 분자식 생성부(114)를 포함한다.

[0019] 문자 저장부(111)에는 사전 설정된 분자식 표기법에서 정의하고 있는 복수의 문자들 각각에 대한 원-핫 벡터가 저장되어 있다.

[0020] 이때, 본 발명의 일실시예에 따르면, 상기 분자식 표기법은 SMILES(Simplified Molecular Input Line Entry System) 또는 SELFIES(Self-Referencing Embedded Strings)일 수 있다.

[0021] 관련해서, 문자 저장부(111)에는 하기의 표 1과 같은 형태로 정보가 저장되어 있을 수 있다.

표 1

문자	원-핫 벡터
A	[1 0 0 0 0 ...]
B	[0 1 0 0 0 ...]
C	[0 0 1 0 0 ...]
...	...

[0025] 보상 산출부(112)는 사전 설정된 시작 토큰을, LSTM(Long Short-Term Memory)(200)의 첫 번째 타임 스텝에 입력으로 인가한 후, 상기 LSTM(200)의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 당해 타임 스텝에서 산출된 출력 토큰(211)과 히든 스테이트(hidden state)(212)를, RND(Random Network Distillation) 기반의 내적 보상(intrinsic reward)을 산출하기 위한 제1 보상층(reward layer)(220)에 입력으로 인가하여, 제1 내적 보상을 산출하고, 상기 첫 번째 타임 스텝부터 당해 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 문자열(213)을, 카운트 기반(Count-based)의 내적 보상을 산출하기 위한 제2 보상층(230)에 입력으로 인가하여, 제2 내적 보상을 산출한 후, 상기 제1 내적 보상과 상기 제2 내적 보상을 합산함으로써, 상기 LSTM의 각 타임 스텝별 보상을 산출한다.

[0026] 여기서, 상기 LSTM(200)을 구성하는 각 타임 스텝은, 도면부호 210에 도시된 그림과 같이, 이전 타임 스텝의 출력 토큰을 입력 토큰으로 인가받아, 상기 복수의 문자들 각각에 대응되는 분류(classification) 확률을 생성한 후, 상기 복수의 문자들 중, 각 분류 확률에 따라 랜덤하게 선택된 어느 하나의 문자에 대한 원-핫 벡터를, 당해 타임 스텝의 출력 토큰으로 산출하는, 신경망 기반의 정책 네트워크(Policy Network)로 구성되어 있다. 이때, 첫 번째 타임 스텝에는 시작 토큰이 입력으로 인가되는데, 이 시작 토큰은 '[0 0 0 0 ...]'의 형태를 갖는 제로 벡터로 구성되어 있을 수 있다.

- [0027] 이때, 각 타임 스텝을 구성하는 정책 네트워크를 간단하게 설명하면 다음과 같다. 우선, 설명의 편의를 위해, t 번째 타임 스텝에 대한 정책 네트워크에 대해 설명하면, 상기 정책 네트워크는 이전 타임 스텝인 $t-1$ 번째 타임 스텝의 정책 네트워크를 통해 산출된 출력 토큰인 x_t 를 입력 토큰으로 입력받을 수 있게 되고, 아울러, $t-1$ 번째 타임 스텝의 정책 네트워크를 통해 산출된 히든 스테이트인 h_{t-1} 를 전파받게 된다. 그 이후, 상기 정책 네트워크는, x_t 와 h_{t-1} 를, 정책 네트워크를 구성하는 신경망에 통과시켜, 문자 저장부(111)에 저장되어 있는 상기 복수의 문자들 각각에 대응되는 분류 확률을 생성하게 된다. 관련해서, 상기 복수의 문자들이 총 10개라고 하는 경우, 정책 네트워크는, x_t 와 h_{t-1} 를 정책 네트워크를 구성하는 신경망에 통과시켜 10개의 출력을 생성한 후, 이 10개의 출력을 softmax와 같은 분류 확률을 산출하는 활성 함수에 통과시켜, 상기 복수의 문자들 각각에 대응되는 총 10개의 분류 확률을 생성하게 된다. 그리고 나서, 상기 정책 네트워크는, 상기 복수의 문자들 중, 각 문자에 대응되는 분류 확률에 따라 어느 하나의 문자를 선택하여, 선택된 문자에 대한 원-핫 벡터를, t 번째 타임 스텝의 출력 토큰인 x_{t+1} 로 산출하게 된다. 즉, 상기 복수의 문자들에 대한 분류 확률이 각각 '0.1, 0.2, 0.3, ...'과 같이 산출되었다고 하는 경우, 상기 정책 네트워크는 각 문자에 대한 분류 확률인 '0.1, 0.2, 0.3, ...'의 확률로 어느 하나의 문자를 랜덤하게 선택할 수 있다. 관련해서, 'A'라는 문자의 분류 확률이 '0.1'이었다고 하는 경우, 상기 정책 네트워크는 'A'라는 문자에 대해서, 상기 복수의 문자들 중에서 선택될 확률을 '0.1'로 적용해서, 랜덤하게 선택할 수 있다. 즉, 'A'라는 문자는 상기 복수의 문자들 중에서 선택될 확률이 '0.1'이 되는 것이다. 이렇게, t 번째 타임 스텝의 정책 네트워크를 통해, 어느 하나의 문자에 대한 원-핫 벡터가 출력 토큰 x_{t+1} 로 산출되면, 이 출력 토큰 x_{t+1} 는 다음 타임 스텝인 $t+1$ 번째 타임 스텝의 입력 토큰으로 전달되고, 동시에, t 번째 타임 스텝의 정책 네트워크를 통해 산출된 히든 스테이트인 h_t 도 $t+1$ 번째 타임 스텝으로 전달된다. 이러한 형태의 반복을 통해, 상기 LSTM(200)은 각 타임 스텝별로, 상기 복수의 문자들 중에서, 각 문자에 대해서 산출된 분류 확률에 따라 랜덤하게 선택된 어느 하나의 문자에 대한 원-핫 벡터로 구성된 출력 토큰을 산출하게 된다.
- [0028] 이때, 본 발명의 일실시예에 따르면, 보상 산출부(112)는 상기 LSTM(200)을 구성하는 타임 스텝들 중 어느 하나인 제1 타임 스텝에서의 보상을 산출해야 하는 순서가 되는 경우, 상기 제1 타임 스텝에서 산출된 출력 토큰(211)과 히든 스테이트(212)를, 상기 제1 보상층(220)을 구성하는 신경망 기반의 예측 네트워크(Prediction Network)와 타겟 네트워크(Target Network)에 각각 입력으로 인가하여, 상기 예측 네트워크로부터 산출되는 출력과 상기 타겟 네트워크로부터 산출되는 출력 간의 오차를, 상기 제1 타임 스텝에서의 상기 제1 내적 보상으로 산출할 수 있다.
- [0029] 그리고, 보상 산출부(112)는 상기 첫 번째 타임 스텝부터 상기 제1 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 제1 문자열(213)을 확인하고, 상기 제1 문자열(213)을, 상기 제2 보상층(230)을 구성하는 LSH(Locality Sensitive Hashing)층에 통과시켜, 상기 제1 문자열에 대해 LSH를 수행함으로써 제1 해시 데이터를 생성한 다음, 상기 첫 번째 타임 스텝부터 상기 제1 타임 스텝까지의 각 타임 스텝별로, 보상을 산출하는 과정에서 상기 LSH가 수행됨에 따라 생성된 해시 데이터들 중, 상기 제1 해시 데이터가 등장하는 빈도수를 카운트하여, 카운트된 빈도수의 제곱근의 역수를, 상기 제1 타임 스텝에서의 상기 제2 내적 보상으로 산출한다.
- [0030] 관련해서, 상기 제1 타임 스텝을, 도 2의 t 번째 타임 스텝이라고 하는 경우, 보상 산출부(112)는, 도 2에 도시된 그림과 같이, t 번째 타임 스텝에서 산출된 출력 토큰(211)인 x_{t+1} 와 히든 스테이트(212)인 h_t 를, 결합하여 하나의 입력으로 구성한 후, 이 입력을, 상기 제1 보상층(220)을 구성하는 예측 네트워크와 타겟 네트워크에 각각 입력으로 인가할 수 있다. 여기서, 상기 예측 네트워크와 상기 타겟 네트워크는, 전 연결 계층(fully connected layer)로 구성된 신경망으로, 두 네트워크는 모두 랜덤하게 초기화된 후, 상기 예측 네트워크는 추후 강화 학습을 통해 변화되고, 상기 타겟 네트워크는 학습 없이, 초기화된 상태를 그대로 유지하게 된다. 이렇게, 상기 예측 네트워크와 상기 타겟 네트워크에 각각 입력이 인가됨에 따라, 각 네트워크로부터 출력이 산출되면, 보상 산출부(112)는 상기 예측 네트워크로부터 산출된 출력과 상기 타겟 네트워크로부터 산출된 출력 간의 오차를, 하기의 수학적 1의 연산에 따라, t 번째 타임 스텝에서의 제1 내적 보상으로 산출할 수 있다.

수학식 1

$$r_t^R = \frac{1}{2} \|\hat{f}(x_{t+1}, h_t) - f(x_{t+1}, h_t)\|^2$$

여기서, $\hat{f}$ 은 상기 예측 네트워크, f 는 상기 타겟 네트워크를 의미한다.

또한, 보상 산출부(112)는, 첫 번째 타임 스텝부터 상기 t번째 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 제1 문자열(213)을 확인할 수 있다. 관련해서, 도 2에서는 상기 제1 문자열이 'P...COS'로 구성된 예가 도시되어 있다. 이렇게, 상기 제1 문자열(213)의 확인이 완료되면, 보상 산출부(112)는 상기 제1 문자열(213)을, 상기 제2 보상층(230)을 구성하는 LSH(Locality Sensitive Hashing)층에 통과시켜, 상기 제1 문자열(213)에 대해 LSH를 수행함으로써 제1 해시 데이터를 생성할 수 있다. 여기서, LSH는 비슷한 항목을 같은 버킷에 넣는 해쉬 알고리즘을 의미한다. 이렇게, 상기 제1 해시 데이터의 생성이 완료되면, 보상 산출부(112)는, 상기 첫 번째 타임 스텝부터 상기 t번째 타임 스텝까지의 각 타임 스텝별로, 보상을 산출하는 과정에서 상기 LSH가 수행됨에 따라 생성된 해시 데이터들 중, 상기 제1 해시 데이터가 등장하는 빈도수를 카운트한 후, 카운트된 빈도수의 제곱근의 역수를 상기 t번째 타임 스텝에서의 제2 내적 보상으로 산출할 수 있다. 이와 관련해서, 보상 산출부(112)가 상기 제2 내적 보상을 산출하는 과정을 수식으로 표현하면, 하기의 수학식 2와 같이 나타낼 수 있다.

수학식 2

$$r_t^C = \frac{1}{\sqrt{\text{Count}(\text{LSH}(\text{MF}(\text{mol}_t)))}}$$

mol_t 는 상기 제1 문자열(213)을 의미하고, MF는 'Morgan fingerprints'를 의미하는 것으로서, 보상 산출부(112)는 상기 제1 문자열(213)에 대해 Morgan fingerprints를 사용해 특징 벡터로 인코딩한 후, 인코딩된 특징 벡터에 대해 LSH를 수행함으로써, 이산적이고 낮은 차원의 데이터(discrete and lower dimensional data)로 해시하여 상기 제1 해시 데이터를 생성할 수 있다. 그리고 나서, 보상 산출부(112)는 상기 첫 번째 타임 스텝부터 상기 t번째 타임 스텝까지에 대한 해시 데이터들 중, 상기 제1 해시 데이터가 등장하는 빈도수를 카운트 한 후, 카운트된 빈도수의 제곱근의 역수를 상기 t번째 타임 스텝에서의 제2 내적 보상으로 산출할 수 있다.

이러한 방식으로, 보상 산출부(112)는 상기 LSTM(200)을 구성하는 각 타임 스텝별로, 출력 토큰이 산출될 때마다, 각 타임 스텝에서의 상기 제1 내적 보상과 상기 제2 내적 보상을 산출하는 과정을 반복할 수 있다.

이렇게, 보상 산출부(112)는, 상기 LSTM(200)을 구성하는 각 타임 스텝별로, 상기 제1 내적 보상과 상기 제2 내적 보상이 산출되면, 상기 제1 내적 보상과 상기 제2 내적 보상을 합산함으로써, 상기 LSTM(200)의 각 타임 스텝별 보상을 산출할 수 있다.

이때, 본 발명의 일실시예에 따르면, 보상 산출부(112)는 상기 LSTM(200)의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 상기 제1 내적 보상과 상기 제2 내적 보상을 산출한 후, 당해 타임 스텝에서 산출된 출력 토큰에 대응하는 문자를, 사전 설정된 단백질과 문자 간의 연관도 스코어를 외적 보상(Extrinsic reward)으로 산출하는 외적 보상층에 입력으로 인가하여 외적 보상을 산출한 다음, 상기 제2 내적 보상에 상기 외적 보상을 합산함으로써 상기 제2 내적 보상을 보정하고, 상기 제2 내적 보상의 보정이 완료되면, 상기 제1 내적 보상과 상기 보정

된 제2 내적 보상을 합산함으로써, 상기 LSTM(200)의 각 타임 스텝별 보상을 산출할 수 있다.

- [0043] 관련해서, 소정의 질병에 대해서 특정 분자식을 갖는 약물이 효과를 거두기 위해서는, 해당 질병을 구성하는 단백질과 분자식 간의 결합 친화도(Binding Affinity)가 높아야 한다. 이로 인해, 본 발명에서는 상기 LSTM(200)의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 상기 제1 내적 보상과 상기 제2 내적 보상 외에, 해당 출력 토큰에 따른 문자가, 타겟하는 단백질에 얼마나 결합 친화도를 갖는지에 대한 스코어를 외적 보상으로 산출하는 외적 보상층을 포함할 수 있다. 예컨대, 특정 질병을 치료할 목적의 약물에 따른 분자식을 생성하는 것이 목표라고 하고, 해당 질병은 'X'라는 단백질로 구성되어 있다고 한다면, 본 분자식 생성 모델은, 'X'라는 단백질에 높은 결합 친화도를 갖는 분자식이 생성되도록 학습될 필요가 있다. 이로 인해, 본 발명에서는 'X'라는 단백질과, 분자식 생성에 사용되는 문자들 간의 결합 친화도를 표상하는 연관도 스코어를 외적 보상으로 산출하는 외적 보상층이 존재할 수 있다. 이 외적 보상층은, 소정의 함수식에 따라, 'X'라는 단백질과, 분자식 생성에 사용되는 문자들 간의 결합 친화도를 표상하는 연관도 스코어를 외적 보상으로 산출하도록 구성될 수 있다. 이로 인해, 보상 산출부(112)는 상기 LSTM(200)의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 상기 제1 내적 보상과 상기 제2 내적 보상을 산출한 것 외에, 당해 타임 스텝에서 산출된 출력 토큰에 대응하는 문자를, 상기 외적 보상층에 입력으로 인가하여 외적 보상을 추가로 산출할 수 있다.
- [0044] 그리고, 보상 산출부(112)는 각 타임 스텝별로, 상기 외적 보상의 산출이 완료되면, 상기 제2 내적 보상에 상기 외적 보상을 합산함으로써 상기 제2 내적 보상을 보정하고, 상기 제2 내적 보상의 보정이 완료되면, 상기 제1 내적 보상과 상기 보정된 제2 내적 보상을 합산함으로써, 상기 LSTM(200)의 각 타임 스텝별 보상을 산출할 수 있다.
- [0045] 즉, 보상 산출부(112)는 t번째 타임 스텝에서 상기 제1 내적 보상과 상기 제2 내적 보상이 산출되면, t번째 타임 스텝에서 산출된 출력 토큰에 대응하는 문자를, 상기 외적 보상층에 입력으로 인가하여, t번째 타임 스텝에서의 외적 보상을 추가로 산출한 후, 상기 제2 내적 보상에 상기 외적 보상을 합산함으로써 상기 제2 내적 보상을 보정하고, 상기 제2 내적 보상의 보정이 완료되면, 상기 제1 내적 보상과 상기 보정된 제2 내적 보상을 합산함으로써, t번째 타임 스텝에서의 보상을 산출하게 된다.
- [0046] 모델 생성부(113)는, 보상 산출부(112)를 통해, 상기 LSTM(200)의 각 타임 스텝별 보상의 산출이 완료되면, 상기 LSTM(200)의 각 타임 스텝별 보상을 모두 합산하여 합산 보상을 산출한 후, 상기 합산 보상이 최대가 되도록 상기 LSTM(200)에 대한 강화 학습(Reinforcement Learning)을 수행함으로써, 분자식 생성 모델을 생성한다.
- [0047] 이때, 본 발명의 일실시예에 따르면, 모델 생성부(113)는 상기 합산 보상이 최대가 되도록 상기 LSTM(200)에 대한 강화 학습을 수행함과 동시에, 상기 제1 보상층(220)을 구성하는 상기 예측 네트워크에 대한 강화 학습도 함께 수행할 수 있다.
- [0048] 여기서, 상기 제1 내적 보상은 상기 예측 네트워크와 상기 타겟 네트워크 간의 오차가 얼마나 큰지를 표상하는 보상이고, 상기 제2 내적 보상은 문자열의 중복 빈도가 얼마나 낮은지를 표상하는 보상이며, 상기 외적 보상은 특정 타임 스텝의 문자가 타겟하는 단백질에 얼마나 연관성을 갖는지를 표상하는 보상이라는 점에서, 상기 분자식 생성 모델을 상기 합산 보상이 최대가 되는 방향으로 강화 학습하게 된다면, 상기 분자식 생성 모델은, 이전에 중복되지 않는 다양한 분자식을 생성하도록 학습될 동시에, 타겟하는 단백질에 연관성이 높은 분자식을 생성하도록 학습되게 된다고 볼 수 있다.
- [0049] 이렇게, 모델 생성부(113)를 통해, 상기 분자식 생성 모델의 생성이 완료된 이후에, 사용자의 의해, 본 발명의 전자 장치(110) 상에, 신규 분자식의 생성 명령이 인가되면, 분자식 생성부(114)는 상기 분자식 생성 모델을 구성하는 상기 LSTM(200)의 첫 번째 타임 스텝에 상기 시작 토큰을 입력으로 인가함으로써, 상기 LSTM(200)을 구성하는 각 타임 스텝별로 산출되는 출력 토큰에 대응되는 문자들로 구성되는 문자열을, 신규 분자식으로 생성한다.
- [0050] 앞서 언급한 바와 같이, 상기 분자식 생성 모델은 상기 합산 보상이 최대가 되는 방향으로 강화 학습이 수행된 모델이므로, 상기 분자식 생성 모델은 이전에 중복되지 않는 다양한 분자식을 생성하도록 학습되어 있음과 동시에, 타겟하는 단백질에 연관성이 높은 분자식을 생성하도록 학습되어 있는 모델이라고 볼 수 있다. 이로 인해, 상기 사용자는, 전자 장치(110)에 신규 분자식의 생성 명령을 인가함으로써, 상기 분자식 생성 모델을 통해, 다양한 형태의 분자식을 획득할 수 있으면서도, 타겟하는 단백질에 친화도(연관도)가 높은 분자식을 획득할 수 있게 된다.
- [0051] 본 발명의 일실시예에 따르면, 전자 장치(110)는, 소정의 인증된 사용자만이, 학습이 완료된 상기 분자식 생성

모델을 이용하여, 신규 분자식을 획득할 수 있도록 하는, 사용자 인증과 관련된 구성을 더 포함할 수 있다.

- [0052] 이와 관련해서, 본 발명의 일실시예에 따르면, 전자 장치(110)는 저장 처리부(115), 표시부(116), 인증 완료부(117) 및 안내 메시지 표시부(118)를 더 포함할 수 있다.
- [0053] 저장 처리부(115)는, 모델 생성부(113)를 통해, 상기 분자식 생성 모델의 생성이 완료된 이후에, 상기 사용자로부터 상기 분자식 생성 모델의 이용을 위한 비밀번호 설정 명령이 인가되면, p (p 는 2이상의 자연수임)자리수의 제1 번호와 q (q 는 2이상의 자연수임)자리수의 제2 번호를 랜덤하게 생성하고, 상기 제1 번호를 구성하는 각 자리 숫자를 모듈로(modulo)-2 연산한 값을 성분으로 갖는 벡터와 상기 제2 번호를 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터 간의 합성곱을 수행하여 결과 벡터를 산출한 후, 상기 결과 벡터의 해밍 무게를 사전 설정된 해시 함수에 입력으로 인가하였을 때 산출되는 해시 값을, 사용자 인증 값으로 지정하여, 사전 지정된 클라우드 저장 서버(20)에 저장한다.
- [0054] 표시부(116)는 상기 제1 번호와 상기 제2 번호를 서로 연결한 번호를, 상기 분자식 생성 모델의 이용을 위한 제1 비밀번호로서 생성한 후, 상기 제1 비밀번호를 화면 상에 표시하면서, 상기 제1 비밀번호를 숙지할 것을 지시하는 메시지를 화면 상에 표시한다.
- [0055] 예컨대, p 를 3, q 를 5라고 하고, 상기 제1 번호가 '213', 상기 제2 번호가 '45678'과 같이 생성되었다고 하는 경우, 저장 처리부(115)는 '213'을 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터인 '[0 1 0]'과, '45678'을 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터인 '[0 1 0 1 0]' 간의 합성곱을 수행하여 '[0 0 1 0 1 0 0]'라는 결과 벡터를 생성한 후, 상기 결과 벡터의 해밍 무게인 '2'를 상기 해시 함수에 입력으로 인가하였을 때 산출되는 해시 값을, 사용자 인증 값으로 지정하여, 사전 지정된 클라우드 저장 서버(20)에 저장할 수 있다.
- [0056] 그리고 나서, 표시부(116)는 상기 제1 번호인 '213'과 상기 제2 번호인 '45678'을 서로 연결한 번호인 '21345678'을, 상기 분자식 생성 모델의 이용을 위한 제1 비밀번호로서 생성한 후, 상기 제1 비밀번호를 화면 상에 표시하면서, 상기 제1 비밀번호를 숙지할 것을 지시하는 메시지를 화면 상에 표시할 수 있다. 이를 통해, 상기 사용자는 상기 제1 비밀번호인 '21345678'을 숙지해 둘 수 있다.
- [0057] 인증 완료부(117)는 상기 제1 비밀번호와 상기 메시지가 화면 상에 표시된 이후, 상기 사용자로부터 상기 분자식 생성 모델의 이용을 위한 인증 요청 명령이 인가되면, 화면 상에, 인증을 위한 비밀번호를 입력할 것을 지시하는 메시지를 표시한 후, 상기 사용자에 의해, 상기 제1 비밀번호가 입력되면, 상기 제1 비밀번호로부터 상기 제1 번호와 상기 제2 번호를 분리한 후, 상기 제1 번호를 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터와 상기 제2 번호를 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터 간의 합성곱을 수행하여 상기 결과 벡터를 산출한 다음, 상기 결과 벡터의 해밍 무게를 상기 해시 함수에 인가하여 상기 해시 값을 연산한 후, 상기 해시 값이 클라우드 저장 서버(20)에 저장되어 있는 상기 사용자 인증 값과 일치하는지 확인하여, 서로 일치하는 것으로 확인되면, 상기 사용자에 대한 인증을 완료 처리한다.
- [0058] 관련해서, 전술한 예와 같이, '21345678'라는 상기 제1 비밀번호와 상기 메시지가 화면 상에 표시된 이후, 상기 사용자로부터 상기 분자식 생성 모델의 이용을 위한 인증 요청 명령이 전자 장치(110)에 입력으로 인가되면, 인증 완료부(117)는 상기 제1 비밀번호로부터 '213'이라는 제1 번호와 '45678'이라고 하는 제2 번호를 분리한 후, '213'을 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터인 '[0 1 0]'과, '45678'을 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터인 '[0 1 0 1 0]' 간의 합성곱을 수행하여 '[0 0 1 0 1 0 0]'라는 결과 벡터를 산출할 수 있다. 그리고 나서, 인증 완료부(117)는 '[0 0 1 0 1 0 0]'라는 결과 벡터의 해밍 무게인 '2'를 상기 해시 함수에 입력으로 인가함으로써, 해시 값을 연산한 후, 상기 해시 값이 클라우드 저장 서버(20)에 저장되어 있는 상기 사용자 인증 값과 일치하는지 확인하여, 서로 일치하는 것으로 확인되면, 상기 사용자에 대한 인증을 완료 처리할 수 있다.
- [0059] 안내 메시지 표시부(118)는 상기 사용자에 대한 인증이 완료 처리되면, 화면 상에, 인증이 완료되었음을 지시하는 메시지를 표시함과 동시에, 상기 사용자로 하여금 상기 신규 분자식의 생성 명령을 인가할 것을 안내하는 안내 메시지를 표시한다.
- [0060] 이를 통해, 상기 사용자는 상기 안내 메시지를 확인한 후, 신규 분자식의 생성 명령을 전자 장치(110) 상에 인가할 수 있고, 이때, 분자식 생성부(114)는, 상기 사용자에 의해, 상기 신규 분자식의 생성 명령이 입력으로 인가되면, 상기 분자식 생성 모델을 기초로, 신규 분자식을 생성할 수 있다.

- [0061] 도 3은 본 발명의 일실시예에 따른 전자 장치의 동작 방법을 도시한 순서도이다.
- [0062] 단계(S310)에서는, 사전 설정된 분자식 표기법에서 정의하고 있는 복수의 문자들 각각에 대한 원-핫 벡터가 저장되어 있는 문자 저장부를 유지한다.
- [0063] 단계(S320)에서는, 사전 설정된 시작 토큰을, LSTM(상기 LSTM을 구성하는 각 타임 스텝은, 이전 타임 스텝의 출력 토큰을 입력 토큰으로 인가받아, 상기 복수의 문자들 각각에 대응되는 분류 확률을 생성한 후, 상기 복수의 문자들 중, 각 분류 확률에 따라 랜덤하게 선택된 어느 하나의 문자에 대한 원-핫 벡터를, 당해 타임 스텝의 출력 토큰으로 산출하는, 신경망 기반의 정책 네트워크로 구성되어 있음)의 첫 번째 타임 스텝에 입력으로 인가한 후, 상기 LSTM의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 당해 타임 스텝에서 산출된 출력 토큰과 히든 스테이트를, RND 기반의 내적 보상을 산출하기 위한 제1 보상층에 입력으로 인가하여, 제1 내적 보상을 산출하고, 상기 첫 번째 타임 스텝부터 당해 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 문자열을, 카운트 기반의 내적 보상을 산출하기 위한 제2 보상층에 입력으로 인가하여, 제2 내적 보상을 산출한 후, 상기 제1 내적 보상과 상기 제2 내적 보상을 합산함으로써, 상기 LSTM의 각 타임 스텝별 보상을 산출한다.
- [0064] 단계(S330)에서는, 상기 LSTM의 각 타임 스텝별 보상의 산출이 완료되면, 상기 LSTM의 각 타임 스텝별 보상을 모두 합산하여 합산 보상을 산출한 후, 상기 합산 보상이 최대가 되도록 상기 LSTM에 대한 강화 학습을 수행함으로써, 분자식 생성 모델을 생성한다.
- [0065] 단계(S340)에서는, 상기 분자식 생성 모델의 생성이 완료된 이후에, 사용자에게 의해, 신규 분자식의 생성 명령이 인가되면, 상기 분자식 생성 모델을 구성하는 상기 LSTM의 첫 번째 타임 스텝에 상기 시작 토큰을 입력으로 인가함으로써, 상기 LSTM을 구성하는 각 타임 스텝별로 산출되는 출력 토큰에 대응되는 문자들로 구성되는 문자열을, 신규 분자식으로 생성한다.
- [0066] 이때, 본 발명의 일실시예에 따르면, 상기 분자식 표기법은 SMILES(Simplified Molecular Input Line Entry System) 또는 SELFIES(Self-Referencing Embedded Strings)일 수 있다.
- [0067] 또한, 본 발명의 일실시예에 따르면, 단계(S320)에서는, 상기 LSTM의 각 타임 스텝에서 출력 토큰이 산출될 때마다, 상기 제1 내적 보상과 상기 제2 내적 보상을 산출한 후, 당해 타임 스텝에서 산출된 출력 토큰에 대응하는 문자를, 사전 설정된 단백질과 문자 간의 연관도 스코어를 외적 보상으로 산출하는 외적 보상층에 입력으로 인가하여 외적 보상을 산출한 다음, 상기 제2 내적 보상에 상기 외적 보상을 합산함으로써 상기 제2 내적 보상을 보정하고, 상기 제2 내적 보상의 보정이 완료되면, 상기 제1 내적 보상과 상기 보정된 제2 내적 보상을 합산함으로써, 상기 LSTM의 각 타임 스텝별 보상을 산출할 수 있다.
- [0068] 또한, 본 발명의 일실시예에 따르면, 단계(S320)에서는, 상기 LSTM을 구성하는 타임 스텝들 중 어느 하나인 제1 타임 스텝에서의 보상을 산출해야 하는 순서가 되는 경우, 상기 제1 타임 스텝에서 산출된 출력 토큰과 히든 스테이트를, 상기 제1 보상층을 구성하는 신경망 기반의 예측 네트워크와 타겟 네트워크에 각각 입력으로 인가하여, 상기 예측 네트워크로부터 산출되는 출력과 상기 타겟 네트워크로부터 산출되는 출력 간의 오차를, 상기 제1 타임 스텝에서의 상기 제1 내적 보상으로 산출하고, 상기 첫 번째 타임 스텝부터 상기 제1 타임 스텝까지, 각 타임 스텝을 통해 산출된 출력 토큰에 대응하는 문자들로 구성되는 제1 문자열을 확인하고, 상기 제1 문자열을, 상기 제2 보상층을 구성하는 LSH층에 통과시켜, 상기 제1 문자열에 대해 LSH를 수행함으로써 제1 해시 데이터를 생성한 다음, 상기 첫 번째 타임 스텝부터 상기 제1 타임 스텝까지의 각 타임 스텝별로, 보상을 산출하는 과정에서 상기 LSH가 수행됨에 따라 생성된 해시 데이터들 중, 상기 제1 해시 데이터가 등장하는 빈도수를 카운트하여, 카운트된 빈도수의 제곱근의 역수를, 상기 제1 타임 스텝에서의 상기 제2 내적 보상으로 산출할 수 있다.
- [0069] 이때, 본 발명의 일실시예에 따르면, 단계(S330)에서는, 상기 합산 보상이 최대가 되도록 상기 LSTM에 대한 강화 학습을 수행함과 동시에, 상기 제1 보상층을 구성하는 상기 예측 네트워크에 대한 강화 학습도 함께 수행할 수 있다.
- [0070] 또한, 본 발명의 일실시예에 따르면, 상기 전자 장치의 동작 방법은, 단계(S330)을 통해, 상기 분자식 생성 모델의 생성이 완료된 이후에, 상기 사용자로부터 상기 분자식 생성 모델의 이용을 위한 비밀번호 설정 명령이 인가되면, p (p 는 2이상의 자연수임)자리수의 제1 번호와 q (q 는 2이상의 자연수임)자리수의 제2 번호를 랜덤하게 생성하고, 상기 제1 번호를 구성하는 각 자리 숫자를 모듈로(modulo)-2 연산한 값을 성분으로 갖는 벡터와 상기 제2 번호를 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터 간의 합성곱을 수행하여 결과 벡터를 산출한 후, 상기 결과 벡터의 해밍 무게를 사전 설정된 해시 함수에 입력으로 인가하였을 때 산출되는

해시 값을, 사용자 인증 값으로 지정하여, 사전 지정된 클라우드 저장 서버에 저장하는 단계, 상기 제1 번호와 상기 제2 번호를 서로 연결한 번호를, 상기 분자식 생성 모델의 이용을 위한 제1 비밀번호로서 생성한 후, 상기 제1 비밀번호를 화면 상에 표시하면서, 상기 제1 비밀번호를 숙지할 것을 지시하는 메시지를 화면 상에 표시하는 단계, 상기 제1 비밀번호와 상기 메시지가 화면 상에 표시된 이후, 상기 사용자로부터 상기 분자식 생성 모델의 이용을 위한 인증 요청 명령이 인가되면, 화면 상에, 인증을 위한 비밀번호를 입력할 것을 지시하는 메시지를 표시한 후, 상기 사용자의 의해, 상기 제1 비밀번호가 입력되면, 상기 제1 비밀번호로부터 상기 제1 번호와 상기 제2 번호를 분리한 후, 상기 제1 번호를 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터와 상기 제2 번호를 구성하는 각 자리 숫자를 모듈로-2 연산한 값을 성분으로 갖는 벡터 간의 합성곱을 수행하여 상기 결과 벡터를 산출한 다음, 상기 결과 벡터의 해밍 무게를 상기 해시 함수에 인가하여 상기 해시 값을 연산한 후, 상기 해시 값이 상기 클라우드 저장 서버에 저장되어 있는 상기 사용자 인증 값과 일치하는지 확인하여, 서로 일치하는 것으로 확인되면, 상기 사용자에게 대한 인증을 완료 처리하는 단계 및 상기 사용자에게 대한 인증이 완료 처리되면, 화면 상에, 인증이 완료되었음을 지시하는 메시지를 표시함과 동시에, 상기 사용자에게 하여금 상기 신규 분자식의 생성 명령을 인가할 것을 안내하는 안내 메시지를 표시하는 단계를 더 포함할 수 있다.

[0071] 이상, 도 3을 참조하여 본 발명의 일실시예에 따른 전자 장치의 동작 방법에 대해 설명하였다. 여기서, 본 발명의 일실시예에 따른 전자 장치의 동작 방법은 도 1 내지 도 2를 이용하여 설명한 전자 장치(110)의 동작에 대한 구성과 대응될 수 있으므로, 이에 대한 보다 상세한 설명은 생략하기로 한다.

[0072] 본 발명의 일실시예에 따른 전자 장치의 동작 방법은 컴퓨터와의 결합을 통해 실행시키기 위한 저장매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다.

[0073] 또한, 본 발명의 일실시예에 따른 전자 장치의 동작 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 본 발명을 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.

[0074] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다.

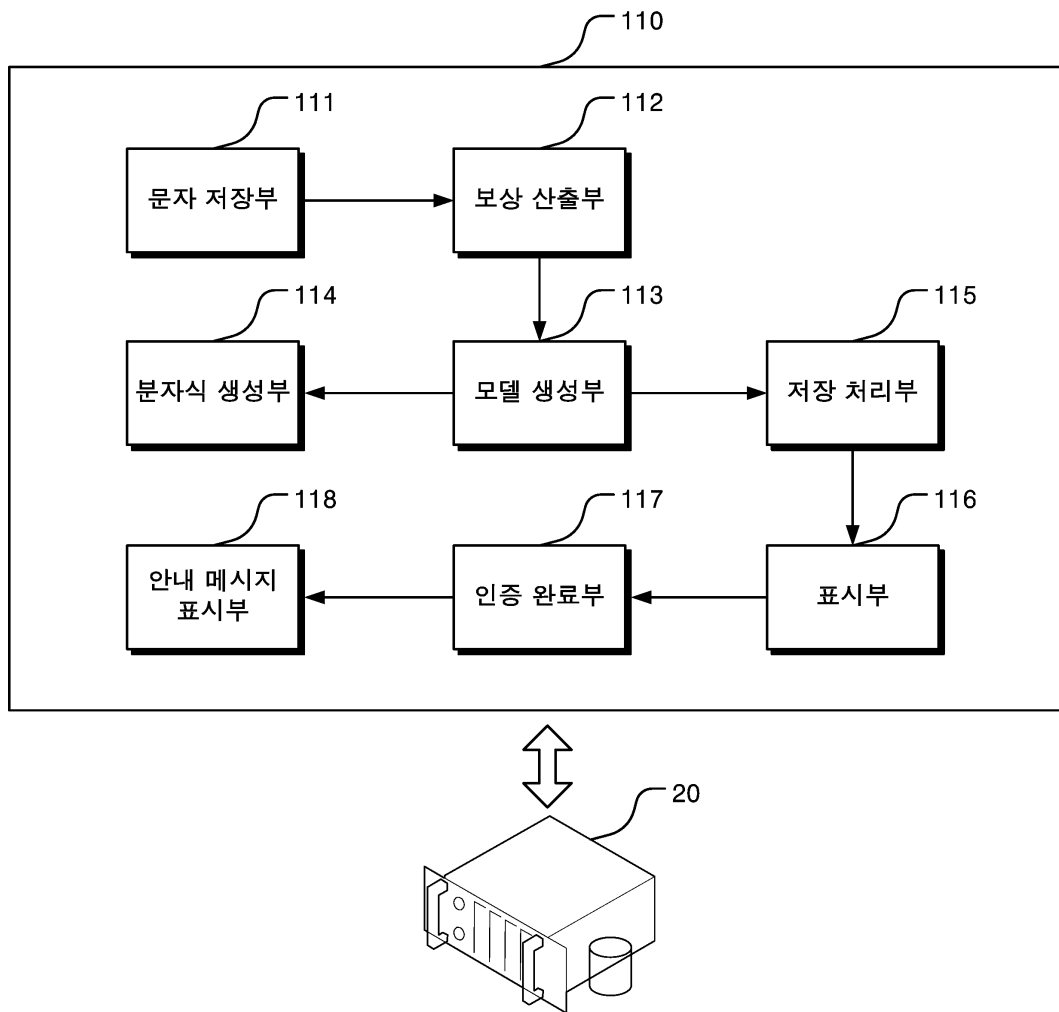
[0075] 따라서, 본 발명의 사상은 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

부호의 설명

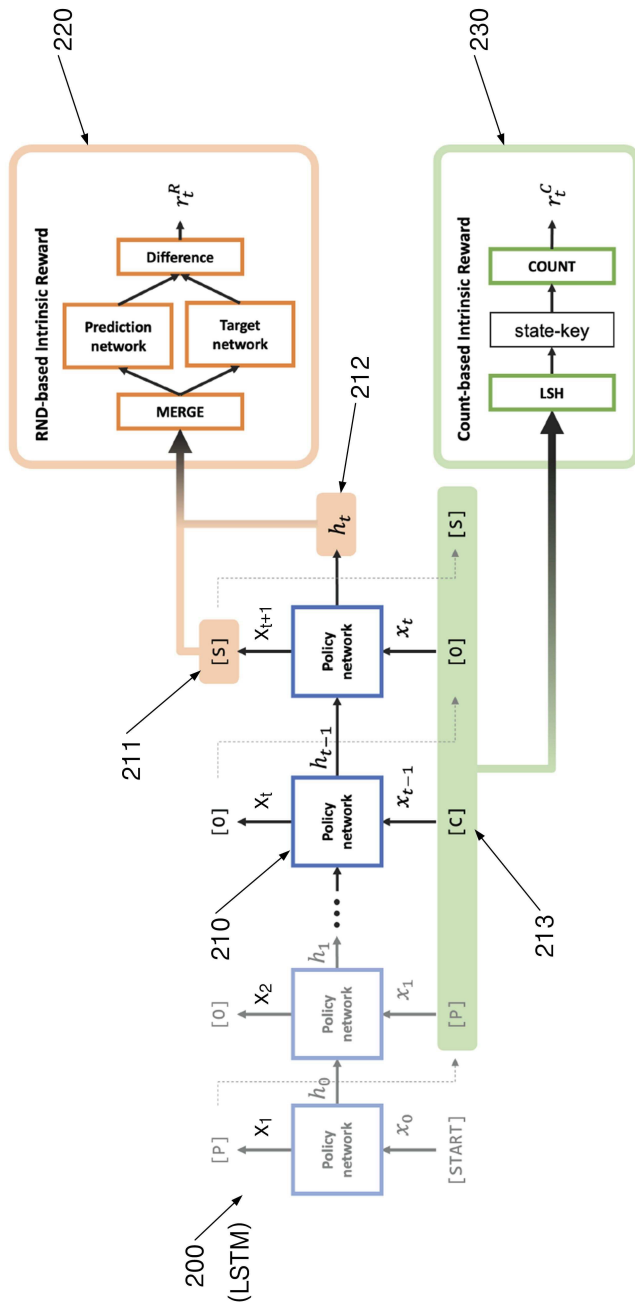
[0076] 110: 전자 장치
111: 문자 저장부 112: 보상 산출부
113: 모델 생성부 114: 분자식 생성부
115: 저장 처리부 116: 표시부
117: 인증 완료부 118: 안내 메시지 표시부
20: 클라우드 저장 서버

도면

도면1



도면2



도면3

