관 인 생 략

출 원 번 호 통 지 서

출 원 일 자	2024.10.28
특 기 사 항	심사청구(유) 공개신청(무) 참조번호(Pn2024N0203)
출 원 번 호	10-2024-0148250 (접수번호 1-1-2024-1172018-11) (DAS접근코드DDF5)
출원인 명칭	조선대학교산학협력단(2-2005-017190-3)
대리인 성명	특허법인 지원(9-2017-100021-1)
발명자 성명	신종호 김춘성 이숙영 우태경 소민섭 키벳 던칸
발명의 명칭	예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치 및 그 구동 방법

특 허 청 장

<< 안내 >>

1. 귀하의 출원은 위와 같이 정상적으로 접수되었으며, 이후의 심사 진행상황은 출원번호를 이용하여 특허로 홈페이지(www.patent.go.kr)에서 확인하실 수 있습니다.
2. 출원에 따른 수수료는 접수일로부터 다음날까지 동봉된 납입영수증에 성명, 납부자번호 등을 기재하여 가까운 은행 또는 우체국에 납부하여야 합니다.
※ 납부자번호 : 0131(기관코드) + 접수번호
3. 귀하의 주소, 연락처 등의 변경사항이 있을 경우, 즉시 [특허고객번호 정보변경(경정), 정정신고서]를 제출하여야 출원 이후의 각종 통지서를 정상적으로 받을 수 있습니다.
4. 기타 심사 절차(제도)에 관한 사항은 특허청 홈페이지를 참고하시거나 특허고객상담센터(☎ 1544-8080)에 문의하여 주시기 바랍니다.
※ 심사제도 안내 : https://www.kipo.go.kr-지식재산제도

【서지사항】

【서류명】 특허출원서

【참조번호】 Pn2024N0203

【출원구분】 특허출원

【출원인】

【명칭】 조선대학교산학협력단

【특허고객번호】 2-2005-017190-3

【대리인】

【명칭】 특허법인 지원

【대리인번호】 9-2017-100021-1

【지정된변리사】 김지원, 심성렬

【발명의 국문명칭】 예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치 및 그 구동 방법

【발명의 영문명칭】 ELECTRONIC DEVICE FOR CLASSIFICATION OF ANATOMICAL THERAPEUTIC CHEMICAL(ATC) CODE BASED ON PREDICTIVE MODEL AND METHOD OPERATION THEREOF

【발명자】

【성명】 신종호

【성명의 영문표기】 Shin, Jongho

【주민등록번호】 740613-1XXXXXX

【우편번호】 07648

【주소】 서울특별시 강서구 공항대로 382, 107동 1504호 (화곡동, 우장산롯데캐슬아파트)

【발명자】

【성명】 김춘성

【성명의 영문표기】 KIM, CHUN SUNG

【주민등록번호】 681207-1XXXXXX

【우편번호】 61709

【주소】 광주광역시 남구 서문대로678번길 7, 대주1차아파트 101동 1002호

【발명자】

【성명】 이숙영

【성명의 영문표기】 Lee, Sook Young

【주민등록번호】 600718-2XXXXXX

【우편번호】 61012

【주소】 광주광역시 북구 첨단과기로208번길 50 (오룡동) 조선대 첨단산학캠퍼스 1층 치과 RIC센터

【발명자】

【성명】 우태경

【성명의 영문표기】 Woo, tae kyung

【주민등록번호】 980922-2XXXXXX

【우편번호】 61754

【주소】 광주광역시 남구 효천2로 12, 102동 1702호

【발명자】

【성명】 소민섭

【성명의 영문표기】 So, min seop

【주민등록번호】 921004-1XXXXXX

【우편번호】 61453

【주소】 광주광역시 동구 필문대로 343, 307호

【발명자】

【성명】 키벳 던칸

【성명의 영문표기】 Kibet Duncan

【주민등록번호】 970310-5XXXXXX

【우편번호】 61445

【주소】 광주광역시 동구 지산로57번길 5-10, 영보빌 210호

【출원언어】 국어

【심사청구】 청구

【공지예외적용대상증명서류의 내용】

【공개형태】 논문발표

【공개일자】 2024.08.23

【이 발명을 지원한 국가연구개발사업】

【과제고유번호】 2520000092

【과제번호】 KS211532

【부처명】 해양수산부

【과제관리(전문)기관명】 해양수산과학기술진흥원

【연구사업명】 빅데이터기반해양바이러스제어및마린바이오틱스

【연구과제명】 해조류 기반 바이오 헬스케어 유효성 소재 실증 지원

【과제수행기관명】 조선대학교 산학협력단

【연구기간】 2024.01.01 ~ 2024.12.31

【취지】 위와 같이 특허청장에게 제출합니다.

대리인 특허법인 지원 (서명 또는 인)

【수수료】

【출원료】	0	면	46,000 원
【가산출원료】	47	면	0 원
【우선권주장료】	0	건	0 원
【심사청구료】	10	항	676,000 원
【합계】	722,000원		
【감면사유】	전담조직(50%감면)[1]		
【감면후 수수료】	361,000 원		

【첨부서류】 1.기타첨부서류[위임장]_1통 2.공지예외적용대상(신규성상실의예외, 출원시의특례)규정을 적용받기 위한 증명서류_1통

1 : 기타첨부서류

PDF 파일 첨부

2 : 공지예외적용대상(신규성상실의예외,_출원시의특례)규정을_적용받기_위한_증명서류

PDF 파일 첨부

【발명의 설명】

【발명의 명칭】

예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치 및 그 구동 방법{ELECTRONIC DEVICE FOR CLASSIFICATION OF ANATOMICAL THERAPEUTIC CHEMICAL(ATC) CODE BASED ON PREDICTIVE MODEL AND METHOD OPERATION THEREOF}

【기술분야】

【0001】본 발명의 다양한 실시예는, 예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치 및 그 구동 방법에 관한 것이다.

【발명의 배경이 되는 기술】

【0002】다양한 화학품 및 약품을 개발하기 위해서는 다양한 신약 및 개발 프로세서가 진행되고 있었다. 이러한, 신약 개발 프로세스가 정확하게 검증된 데이터인지를 확인하기 위해서 다양한 사람을 통해 실제 사람 및/또는 동물에 대해 약물을 투여하여 신약을 개발한 것에 대해 검증해왔다.

【0003】그러나 종래의 경우 상기와 같이 반복적인 실험을 통해 목표하는 효능을 도출할 때까지 반복적으로 수행하여 시간, 노력, 자원이 많이 소모되는 문제점이 있었다.

【0004】또한, 이러한 종래의 경우는 전문가마다 다른 판단을 수행하여 정확성 및 일관성에 차이가 있고, 실험에 대한 외부적인 요인으로 인해 실험 결과의 차이가 발생될 우려와 신뢰성이 저하되는 문제점이 있었다.

【0005】따라서, 이러한 문제를 극복하기 위하여, 인공지능 모델을 통해 별도의 사고실험을 직접적으로 수행하지 않으면서도 좋은 결과 값을 얻을 수 있는 발명이 필요한 실정이다.

【선행기술문헌】

【특허문헌】

【0006】**(특허문헌 0001)** 미국공개특허공보 제2024-0013866호(2024.01.11)

(특허문헌 0002) 일본등록특허공보 제7,495,674호(2024.05.28)

【발명의 내용】

【해결하고자 하는 과제】

【0007】따라서, 본 실시예에서는, 화학 데이터를 예측 모델에 입력하여 손쉽게 해당 결과 값을 출력하고, 이를 기반으로 추출된 신규 화학 예측 데이터를 기초로 화학 제품의 사용 여부를 손쉽게 파악할 수 있으며, 인체의 부위에 효과가 있는지를 예측하여 추천할 수 있는 예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치 및 그 구동 방법에 관한 것이다.

【과제의 해결 수단】

【0008】다양한 실시예에 따르면, 예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치에 있어서, 통신 인터페이스; 및 프로세서; 를 포함하고, 상기 프로세서는, 상기 통신 인터페이스로부터, 화학식 구조로 구성된 적어도 하나의 화학 이미지 데이터 및 화학 수식으로 구성된 적어도 하나의 화학 텍스트 데이터를 획득하고, 상기 적어도 하나의 화학 이미지 데이터를 제1 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 구조 결합 데이터를 출력하고, 상기 적어도 하나의 화학 텍스트 데이터를 제2 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 선형 결합 데이터를 출력하고, 상기 적어도 하나의 화학 구조 결합 데이터 및 상기 적어도 하나의 화학 선형 데이터를 상기 예측 모델에 입력하여 적어도 하나의 화학 결과 데이터를 출력하고, 및 상기 적어도 하나의 결과 데이터가 기 설정된 제1 임계 값을 초과하면, 상기 적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합한 기능을 갖춘 화학 제품으로 판단하도록 설정되고, 상기 예측 모델은, 복수의 화학 구조 결합 데이터 및 복수의 화학 선형 데이터와, 복수의 화학 결과 데이터 및 복수의 전문의에 판별된 복수의 화학 제품에 기반하여 학습된다.

【발명의 효과】

【0009】본 실시예는, 적어도 하나의 화학 이미지 데이터와 적어도 하나의 화학 텍스트 데이터를 각각 제1 서브 딥러닝 모델 및 제2 서브 딥러닝 모델에 입력하여 각각 적어도 하나의 화학 구조 결합 데이터와 적어도 하나의 화학 선형 데이터를 추출하고, 적어도 하나의 화학 구조 결합 데이터와 적어도 하나의 화학 선형

데이터를 예측 모델에 입력하여 출력된 결과 데이터로 손쉽게 약물에 적용되는지 유추가 가능함으로써, 개발 기간 및 시간 단축도 가능하면서도 일관성으로 해석될 수 있으며, 별도의 실험 없이 화학 데이터의 결과 값을 확인할 수 있는 이점이 있다.

【도면의 간단한 설명】

【0010】도 1은 본 발명의 다양한 실시예에 따른 전자 장치 및 네트워크의 블록도를 도시한다.

도 2는 다양한 실시예들에 따른, 전자 장치가 동작하는 방법을 설명하기 위한 흐름도이다.

도 3은 다양한 실시예들에 따른, 전자 장치가 화학식을 기초로 예측 모델에 입력하는 방법을 설명하기 위한 예시도이다.

도 4는 다양한 실시예들에 따른, 각 서브 딥러닝 모델을 기반으로 학습되는 예시도이다.

도 5는 다양한 실시예들에 따른, 예측 모델에 입력 데이터, 학습 데이터, 및 출력 데이터를 설명하기 위한 예시도이다.

도 6은 다양한 실시예들에 따른, 화학식을 기초로 학습 데이터를 최소화 및 구별하는 것을 설명하기 위한 예시도이다.

【발명을 실시하기 위한 구체적인 내용】

【0011】 이하, 본 문서의 다양한 실시예들이 첨부된 도면을 참조하여 기재된다. 실시예 및 이에 사용된 용어들은 본 문서에 기재된 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 해당 실시예의 다양한 변경, 균등물, 및/또는 대체물을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다. 본 문서에서, "A 또는 B" 또는 "A 및/또는 B 중 적어도 하나" 등의 표현은 함께 나열된 항목들의 모든 가능한 조합을 포함할 수 있다. "제 1," "제 2," "첫째," 또는 "둘째,"등의 표현들은 해당 구성요소들을, 순서 또는 중요도에 상관없이 수식할 수 있고, 한 구성요소를 다른 구성요소와 구분하기 위해 사용될 뿐 해당 구성요소들을 한정하지 않는다. 어떤(예: 제 1) 구성요소가 다른(예: 제 2) 구성요소에 "(기능적으로 또는 통신적으로) 연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로 연결되거나, 다른 구성요소(예: 제 3 구성요소)를 통하여 연결될 수 있다.

【0012】 본 문서에서, "~하도록 구성된(또는 설정된)(configured to)"은 상황에 따라, 예를 들면, 하드웨어적 또는 소프트웨어적으로 "~에 적합한," "~하는 능력을 가지는," "~하도록 변경된," "~하도록 만들어진," "~를 할 수 있는," 또는 "~하도록 설계된"과 상호 호환적으로(interchangeably) 사용될 수 있다. 어떤 상황에서는, "~하도록 구성된 장치"라는 표현은, 그 장치가 다른 장치 또는 부품들과 함께 "~할 수 있는" 것을 의미할 수 있다. 예를 들면, 문구 "A, B, 및 C를 수행하

도록 구성된(또는 설정된) 프로세서"는 해당 동작을 수행하기 위한 전용 프로세서(예: 임베디드 프로세서), 또는 메모리 장치에 저장된 하나 이상의 소프트웨어 프로그램들을 실행함으로써, 해당 동작들을 수행할 수 있는 범용 프로세서(예: CPU 또는 application processor)를 의미할 수 있다.

【0013】본 문서의 다양한 실시예들에 따른 전자 장치는, 예를 들면, 스마트폰, 태블릿 PC, 데스크탑 PC, 랩탑 PC, 넷북 컴퓨터, 워크스테이션, 서버 중 적어도 하나를 포함할 수 있다.

【0014】도 1을 참조하여, 다양한 실시예에서의, 네트워크 환경(100) 내의 전자 장치(101)가 기재된다. 전자 장치(101)는 버스(110), 프로세서(120), 메모리(130), 입출력 인터페이스(140), 디스플레이(150), 및 통신 인터페이스(160)를 포함할 수 있다. 어떤 실시예에서는, 전자 장치(101)는, 구성요소들 중 적어도 하나를 생략하거나 다른 구성요소를 추가적으로 구비할 수 있다. 버스(110)는 구성요소들(110-160)을 서로 연결하고, 구성요소들 간의 통신(예: 제어 메시지 또는 데이터)을 전달하는 회로를 포함할 수 있다. 프로세서(120)는, 중앙처리장치, 어플리케이션 프로세서, 또는 커뮤니케이션 프로세서(communication processor(CP)) 중 하나 또는 그 이상을 포함할 수 있다. 프로세서(120)는, 예를 들면, 전자 장치(101)의 적어도 하나의 다른 구성요소들의 제어 및/또는 통신에 관한 연산이나 데이터 처리를 실행할 수 있다.

【0015】메모리(130)는, 휘발성 및/또는 비휘발성 메모리를 포함할 수 있다. 메모리(130)는, 예를 들면, 전자 장치(101)의 적어도 하나의 다른 구성요소에 관계

된 명령 또는 데이터를 저장할 수 있다. 한 실시예에 따르면, 메모리(130)는 소프트웨어 및/또는 프로그램(140)을 저장할 수 있다.

【0016】 입출력 인터페이스(140)는, 예를 들면, 환자 또는 다른 외부 기기로부터 입력된 명령 또는 데이터를 전자 장치(101)의 다른 구성요소(들)에 전달하거나, 또는 전자 장치(101)의 다른 구성요소(들)로부터 수신된 명령 또는 데이터를 환자 또는 다른 외부 기기로 출력할 수 있다.

【0017】 디스플레이(150)는, 예를 들면, 액정 디스플레이(LCD), 발광 다이오드(LED) 디스플레이, 유기 발광 다이오드(OLED) 디스플레이, 또는 마이크로 전자기계 시스템 (MEMS) 디스플레이, 또는 전자종이(electronic paper) 디스플레이를 포함할 수 있다. 디스플레이(150)는, 예를 들면, 환자에게 각종 콘텐츠(예: 텍스트, 이미지, 비디오, 아이콘, 및/또는 심볼 등)을 표시할 수 있다. 디스플레이(150)는, 터치 스크린을 포함할 수 있으며, 예를 들면, 전자 펜 또는 환자의 신체의 일부를 이용한 터치, 제스쳐, 근접, 또는 호버링 입력을 수신할 수 있다. 통신 인터페이스(160)는, 예를 들면, 전자 장치(101)와 외부 장치(예: 제 1 외부 전자 장치(102), 제 2 외부 전자 장치(104), 또는 서버(108)) 간의 통신을 설정할 수 있다. 예를 들면, 통신 인터페이스(160)는 무선 통신 또는 유선 통신을 통해서 네트워크(162)에 연결되어 외부 장치(예: 제 2 외부 전자 장치(104) 또는 서버(108))와 통신할 수 있다.

【0018】 무선 통신은, 예를 들면, LTE, LTE-A(LTE Advance), CDMA(code division multiple access), WCDMA(wideband CDMA), UMTS(universal mobile

telecommunications system), WiBro(Wireless Broadband), 또는 GSM(Global System for Mobile Communications) 등 중 적어도 하나를 사용하는 셀룰러 통신을 포함할 수 있다. 한 실시예에 따르면, 무선 통신은, 예를 들면, WiFi(wireless fidelity), 블루투스, 블루투스 저전력(BLE), 지그비(Zigbee), NFC(near field communication), 자력 시큐어 트랜스미션(Magnetic Secure Transmission), 라디오 프리퀀시(RF), 또는 보디 에어리어 네트워크(BAN) 중 적어도 하나를 포함할 수 있다. 한실시예에 따르면, 무선 통신은 GNSS를 포함할 수 있다. GNSS는, 예를 들면, GPS(Global Positioning System), Glonass(Global Navigation Satellite System), Beidou Navigation Satellite System(이하 "Beidou") 또는 Galileo, the European global satellite-based navigation system일 수 있다. 이하, 본 문서에서는, "GPS"는 "GNSS"와 상호 호환적으로 사용될 수 있다. 유선 통신은, 예를 들면, USB(universal serial bus), HDMI(high definition multimedia interface), RS-232(recommended standard232), 전력선 통신, 또는 POTS(plain old telephone service) 등 중 적어도 하나를 포함할 수 있다. 네트워크(162)는 텔레커뮤니케이션 네트워크, 예를 들면, 컴퓨터 네트워크(예: LAN 또는 WAN), 인터넷, 또는 텔레폰 네트워크 중 적어도 하나를 포함할 수 있다.

【0019】제 1 및 제 2 외부 전자 장치(102, 104, 106) 각각은 전자 장치(101)와 동일한 또는 다른 종류의 장치일 수 있다. 다양한 실시예에 따르면, 전자 장치(101)에서 실행되는 동작들의 전부 또는 일부는 다른 하나 또는 복수의 전자 장치(예: 전자 장치(102, 104, 106), 또는 서버(108)에서 실행될 수 있다. 한 실시

예에 따르면, 전자 장치(101)가 어떤 기능이나 서비스를 자동으로 또는 요청에 의하여 수행해야 할 경우에, 전자 장치(101)는 기능 또는 서비스를 자체적으로 실행시키는 대신에 또는 추가적으로, 그와 연관된 적어도 일부 기능을 다른 장치(예: 전자 장치(102, 104, 106), 또는 서버(108))에게 요청할 수 있다. 다른 전자 장치(예: 전자 장치(102, 104, 106), 또는 서버(108))는 요청된 기능 또는 추가 기능을 실행하고, 그 결과를 전자 장치(101)로 전달할 수 있다. 전자 장치(101)는 수신된 결과를 그대로 또는 추가적으로 처리하여 요청된 기능이나 서비스를 제공할 수 있다. 이를 위하여, 예를 들면, 클라우드 컴퓨팅, 분산 컴퓨팅, 또는 클라이언트-서버 컴퓨팅 기술이 이용될 수 있다.

【0021】 도 2는 다양한 실시예들에 따른, 전자 장치가 동작하는 방법을 설명하기 위한 흐름도이다.

【0022】 도 3은 다양한 실시예들에 따른, 전자 장치가 화학식을 기초로 예측 모델에 입력하는 방법을 설명하기 위한 예시도이다.

【0023】 도 4는 다양한 실시예들에 따른, 각 서브 딥러닝 모델을 기반으로 학습되는 예시도이다.

【0024】 도 5는 다양한 실시예들에 따른, 예측 모델에 입력 데이터, 학습 데이터, 및 출력 데이터를 설명하기 위한 예시도이다.

【0025】도 6은 다양한 실시예들에 따른, 화학식을 기초로 학습 데이터를 최소화 및 구별하는 것을 설명하기 위한 예시도이다.

【0027】201 동작에서, 전자 장치(101)(예: 도 1의 프로세서(120))는, 통신 인터페이스(도 1의 통신 인터페이스(160))로부터, 화학식 구조로 구성된 적어도 하나의 화학 이미지 데이터 및 화학 수식으로 구성된 적어도 하나의 화학 텍스트 데이터를 획득할 수 있다. 일 실시예에 따르면, 적어도 하나의 화학 이미지 데이터는, 적어도 하나의 원자 이미지 데이터와, 적어도 하나의 원자 이미지 데이터 간을 각각 연결하는 적어도 하나의 원자 공유 결합 이미지 데이터를 포함할 수 있다. 구체적으로, 적어도 하나의 원자 이미지 데이터는, 일반적으로 사용되는 수소 원자(H), 산소 원자(O), 금 원자(Au)와 같은 일반적인 원자에 대하여 각각 이미지 형태로 표현된 것일 수 있다. 또한, 전자 장치(101)는, 도 3(a)에 표시된 바와 같이, 후술할 제1 노드에 사용되는 V_1, V_2와 같이 변형되는 원래의 이미지 데이터일 수 있다. 또 다른 구체적으로, 적어도 하나의 원자 공유 결합 이미지 데이터는, 수소 원자(H)간의 공유 결합, 수소 원자(H) 및 산소 원자(O)간의 공유 결합으로 통해 생성되는 결합되는 부분을 이미지 형태로 표현된 것일 수 있다. 또한, 전자 장치(101)는, 도 3(a)에 표시된 바와 같이, 후술할 제1 엣지에 사용되는 e_1, e_2와 같이 변형되는 원래의 이미지 데이터일 수 있다.

【0028】 또 다른 일 실시예에 따르면, 적어도 하나의 화학 텍스트 데이터는, 적어도 하나의 원자 이미지 데이터와 매칭되는 적어도 하나의 원자 기호 데이터, 적어도 하나의 화학 이미지 데이터에 복수개의 원자 및 결합 구조에 따른 원자 개수 데이터, 및 적어도 하나의 화학 이미지 데이터의 결합구조와 매칭되는 적어도 하나의 괄호 데이터를 포함할 수 있다. 구체적으로, 원자 기호 데이터는, 수소 원자(H), 산소 원자(O), 금 원자(Au)와 같이 일반적으로 화학기호에 사용되는 문자를 기반으로 상기의 원자 이미지 데이터와 매칭될 수 있다. 또 다른 구체적인 예시로, 적어도 하나의 원자 개수 데이터는 할로겐족(He, Ne 등)과 같은 안정성이 높은 원자가 아닌 일반 원자들의 결합 구조를 나타낼 때 단일객체를 제외한 다수개의 경우 표기되는 수치를 나타낼 수 있다. 즉, 적어도 하나의 원자 개수 데이터는 수소 분자(H_2)의 2를 표시하는 텍스트 데이터로 설정될 수 있다. 또 다른 구체적인 예시로, 적어도 하나의 괄호 데이터는 유기화합물(예: $C_2H_5CC_2H_5$)와 같이 중앙 부분에서 원자 개수 데이터를 제외하고 표기할 때 다른 성질을 표현하는 원자 기호 데이터를 별도로 표기하기 위한 텍스트 데이터로 설정될 수 있다. 이후, 전자 장치(101)는, 적어도 하나의 원자 이미지 데이터 및 적어도 하나의 원자 공유 결합 이미지 데이터를 제1 서브 딥러닝 모델에 입력하고, 적어도 하나의 원자 기호 데이터, 적어도 하나의 원자 개수 데이터, 및 적어도 하나의 괄호 데이터를 제2 서브 딥러닝 모델에 입력할 수 있다.

【0029】203 동작에서, 전자 장치(101)(예: 도 1의 프로세서(120))는, 적어도 하나의 화학 이미지 데이터를 제1 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 구조 결합 데이터를 출력하고, 적어도 하나의 화학 텍스트 데이터를 제2 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 선형 결합 데이터를 출력할 수 있다. 일 실시예에 따르면, 제1 서브 딥러닝 모델은, 복수의 제1 노드, 복수의 제1 엣지, 복수의 제1 행렬, 및 복수의 제2 행렬과, 복수의 화학 구조 결합 데이터에 기반하여 학습될 수 있다. 구체적으로, 제1 딥러닝 모델은, CNN 구조로서, AlexNet, LENET-5, NIN, VGGNet, ResNet, WideResnet, GoogleNet, FractaNet, DenseNet, FitNet, RitResNet, HighwayNet, MobileNet, DeeplySupervisedNet 중 적어도 하나가 이용될 수 있다. 더 구체적으로, LeNet - 5는 LeNet 모델 중 가장 최근 모델이며 Yann LeCun 연구실에서 1990년대에 생성되었고, 우편번호나 숫자를 인식하는데 활용될 수 있다. 또한, LeNet구조가 지금의 CNN과 큰 다른 점이 없다는 것이 핵심이고, Convolution과 subsampling이 쓰이고 feature map을 일자로 펼치는 fully-connection으로 연결을 할 수 있다. 이어서, AlexNet은 ILSVRC 2012에서 우승한 모델이고, 이 때 당시 AlexNet 모델에 인해 딥러닝 모델에 혁명이 일었다고 볼 수 있다. 이는 AlexNet이 CNN 구조로 과거의 top 5 error를 크게 줄일 수 있다. 이어서, AlexNet은 image neural network에서 CNN기법이 적용되기 시작되었고, 이러한 구조로 진행이 되는데 여기서 특이한 점은 다수개의 filter를 한 번에 적용한 것이 아니라 양쪽으로 분할하여 GPU를 2개 사용하여 분석을 진행할 수 있다. ZFNet은 AlexNet의 구조와 거의 유사하고, 실제로 AlexNet에 쓰였던 parameter를 일부 변경

한 정도이면서 성능이 향상되었고, 이는 CNN를 잘 training하려면 intermediate layer에 그 filter들이 어떻게 학습되는지를 확인할 수 있다. 다른 예시로서, GoogLeNet은 LSVRC 2014에서 우승한 모델이고, AlexNet을 비교 대비 depth가 더 깊기도 하고 width도 더 두꺼운데, parameter의 개수가 매우 줄어든 것을 확인할 수 있다. 이는 GoogLeNet에서 inception Module이 개념이 도입되었고, filter 연산이 선형 연산인데 filter 연산을 비선형으로 연산을 수행 시 더 많은 정보를 찾을 수 있지 않을까 라는 컨셉에서 나온 모듈로 이용된 inception module을 적용하였기 때문이다. GoogLeNet의 구조는 Network 구조 안에 Network와 유사한 Layer가 존재하며, 이를 Network In Network(NIN) 구조로 형성되면서, 비선형으로 이뤄질 수 있다. 또한, 딥러닝 모델은 CNN의 구조를 적용하고 있으며, 이에 대한 예시는 한정되지 않는다. 또 다른 일 실시예에 따르면, 전자 장치(101)(예: 도 1의 프로세서(120))는, 도 3(b)에 도시된 바와 같이, 적어도 하나의 원자 이미지 데이터를 각각 적어도 하나의 제1 노드로 변경하면서, 적어도 하나의 제1 노드를 제1 행렬에 결합할 수 있고, 적어도 하나의 원자 공유 결합 이미지 데이터를 각각 적어도 하나의 제1 엣지로 변경하면서, 적어도 하나의 제1 엣지를 제2 행렬에 결합할 수 있으며, 제1 행렬 및 제2 행렬을 제1 딥러닝 모델에 입력하여 적어도 하나의 화학 구조 결합 데이터를 출력할 수 있다. 구체적으로, 전자 장치(101)는, 화합물로 구성된 적어도 하나의 화학 이미지 데이터에서 적어도 하나의 원자 이미지 데이터를 그래프의 제1 노드로, 적어도 하나의 원자들 사이의 결합은 그래프의 제1 엣지로 대체하여 정형화된 데이터로 활용하여 제1 서브 딥러닝 모델에 입력할 수 있다. 또한, 전

자 장치(101)는, 제1 엣지가 방향성을 지니지 않는 그래프를 무방향 그래프로, 제1 엣지가 방향성을 지니는 그래프를 방향 그래프로 설정될 수 있고, 각 방향성의 값에 따라서, 특정 벡터를 출력 값으로 제1 서브 딥러닝 모델에 입력하여 획득될 수 있다.

【0030】 또 다른 일 실시예에 따르면, 제2 서브 딥러닝 모델은, 복수의 제2 노드, 복수의 제3 노드, 및 복수의 제2 엣지와, 복수의 화학 선형 결합 데이터, 복수의 스마일식 코드, 및 ACS|| 인코딩 방식 코드에 기반하여 학습될 수 있다. 구체적으로, 제2 서브 딥러닝 모델은, RNN 구조로서, LSTM, GRU 중 적어도 하나가 이용될 수 있다. 더 구체적으로, RNN(Recurrent Neural Network)은 은닉 계층 안에 하나 이상의 순환 계층을 갖는 신경망을 의미하고, 순환 신경망은 다른 네트워크들과 입력을 받는 방식에 차이가 있으며, 일반적으로 순서가 있는 데이터에 주로 사용되며, 상기 데이터를 입력을 시작으로 하나의 네트워크를 통해서 순서대로 출력을 획득할 수 있다. 또한, RNN에 주로 이용되는 순서 데이터는 소리, 언어, 날씨, 주가 등의 데이터처럼 시간의 변화에 함께 변화하면서 그 영향을 받는 데이터를 의미할 수 있다. LSTM(long short term memory)는, RNN에서 장기 의존성 문제가 발생될 수 있고, 이를 극복하기 위해 변형한 구조일 수 있다. LSTM 레이어는 출력 값 이외에, LSTM셀 사이에서 공유되는 셀의 상태가 존재한다는 점이 특징이고, 셀 상태가 함께 다음 레이어로 전달되면서, 기존의 상태를 보존하여 장기의존성 문제를 해결할 수 있다. GRU(gate recurrent unit)은 LSTM과 비슷한 역할을 하지만, 더 간단한 구조로 이루어져 있어서 계산상으로 효율적으로 사용될 수 있고, GRU는 LSTM에서의 셀

상태(cell state) 역할의 cell state가 별도 존재하지 않고, cell state의 역할을 다음의 출력 h에서 그 역할을 함께 할 수 있으며, GRU에서는 Update Gate, Reset Gate 두 가지만 존재함으로써 구조가 단순하면서도 효율적으로 사용되어 RNN의 기술에서 가장 최근에 생성된 구조일 수 있다. 또한, 제2 딥러닝 모델은 RNN의 구조를 적용하고 있으며, 이에 대한 예시는 한정되지 않는다.

【0031】 또 다른 일 실시예에 따르면, 전자 장치(101)(예: 도 1의 프로세서(120))는, 적어도 하나의 원자 기호 데이터에 결합 요소에 따라 각각 기호를 변경하고, 변경된 적어도 하나의 원자 기호 데이터를 각각 적어도 하나의 제2 노드로 변경하고, 복수개의 원자 기호 데이터가 있는 적어도 하나의 제2 노드에 적어도 하나의 원자 개수 데이터를 적용하여 적어도 하나의 제3 노드를 적용할 수 있으며, 적어도 하나의 괄호 데이터를 각각 적어도 하나의 제2 엣지로 변경하고, 제2 노드, 제3 노드, 및 제2 엣지를 제2 딥러닝 모델에 입력하여 적어도 하나의 화학 선형 결합 데이터를 출력할 수 있다. 구체적으로, 전자 장치(101)는, 도 4에 도시된 바와 같이, 제2 서브 딥러닝 모델이 학습되기 위해서 정규화된 데이터가 필요할 수 있다. 본 실시예에서는, ASC|| 인코딩 방식으로 생성된 SMILES(simplified molecular input line entry system) 표기법을 사용할 수 있다. 이는, 간결하면서도 합리적으로 원자들의 결합을 표현할 수 있고, 데이터 전송을 효율적으로 할 수 있으며, 화합물의 구조를 표준화된 형식으로 표현하여 서로 다른 시스템 간에도 화합물의 정보를 일관되게 교환할 수 있다. 구체적으로, 본 실시예에서의 SMILES 표기법은 원자, 결합, 고리, 방향족, 가지의 5가지로 구성되고, 일정한 규칙이 존재

할 수 있다. 또한, 본 실시예에서는 제2 엣지에 대해서 비결합(.), 단일 결합(-), 이중결합(=), 삼중결합(#), 사중결합($)으로 표기하고, 방향족 결합(:) 입체 결합(/, ₩)로 구별할 수 있으며, 입체결합의 경우 같은 이중 결합인 시스(cis) 이성질체의 경우 /로 표기될 수 있고, 이중 결합이 반대 방향으로 된 트랜스(trans) 이성질체의 경우 ₩로 표기될 수 있다.

【0032】또 다른 일 실시예에 따르면, 전자 장치(101)(예: 도 1의 프로세서(120))는, 적어도 하나의 원자 기호 데이터에서 기 설정된 제2 임계 값을 초과하지 않으면, 적어도 하나의 원자 기호 데이터에서 이성질체가 아닌 것으로 판단하고, ATC(anatomical therapeutic chemical) 레벨 명칭을 기반으로 적어도 하나의 원자 기호 데이터를 제2 노드로 변경하고, 적어도 하나의 원자 기호 데이터에서 제2 임계 값을 초과하면, 적어도 하나의 원자 기호 데이터에서 이성질체인 것으로 판단할 수 있다. 구체적으로, 전자 장치(101)는, ATC의 단계에 따라 각각 세분화될 수 있고, 7개의 단계로써, 총 14개의 그룹이 형성될 수 있다. 이후, 전자 장치(101)는, 도 6에 도시된 바와 같이, 화학식의 물리적 화학적으로 다르게 구성된 이성질체가 있는지 여부를 판단하는 제2 임계 값을 기반으로 적어도 하나의 원자 기호 데이터로부터 판단할 수 있다. 여기서, 제2 임계 값은, 서로 다른 적어도 하나의 원자 기호 데이터로부터 동일한 원자량이 존재하는지 유무 값일 수 있다. 이후, 전자 장치(101)는, 제2 임계 값을 기초로 적어도 하나의 원자 기호 데이터를 비교하여 각각의 적어도 하나의 원자 기호 데이터가 제2 임계 값을 초과하지 않으면, 이성질체가 아닌 것으로 판단하여 제2 노드로 바로 변환하여 학습 데이터로 사용할 수 있다.

【0033】 또 다른 일 실시예에 따르면, 전자 장치(101)(예: 도 1의 프로세서(120))는, 적어도 하나의 원자 기호 데이터가 제2 임계 값을 초과한 경우, 이성질체인 것으로 판단된 적어도 하나의 원자 기호 데이터를 기 설정된 제3 임계 값을 초과하지 않으면, 적어도 하나의 원자 기호 데이터에서 화학 구조식이 별도인 것으로 판단하여, 적어도 하나의 원자 기호 데이터를 괄호 형태에 따라 변경된 상태로 제2 노드로 변경할 수 있다. 구체적으로, 전자 장치(101)는, 도 6에 도시된 바와 같이, 서로 다른 적어도 하나의 원자 기호 데이터가 이성질체인 것으로 판단하면, 각각의 데이터가 시스 구조인지 트랜스 구조인지를 구별하거나, 성질 자체의 차이 유무를 판단하는 제3 임계 값을 기반으로 각각의 적어도 하나의 원자 기호 데이터로부터 판단할 수 있다. 여기서, 제3 임계 값은, 서로 다른 적어도 하나의 원자 기호 데이터로부터 시스 구조인가 또는 트랜스 구조인가에 대한 설정 값 및/또는 구조에 따른 위치의 차이에 따른 값으로 설정될 수 있다. 이후, 전자 장치(101)는, 제3 임계 값을 기초로 적어도 하나의 원자 기호 데이터를 비교하여 각각의 적어도 하나의 원자 기호 데이터가 제3 임계 값을 초과하지 않으면, 이성질체이지만 화학 구조식이 별도이고, 트랜스 기반의 구조인 것으로 판단하여 제2 노드에 변환 전 해당 기호를 입력 후 제2 노드 및 제3 노드, 제2 엣지로 변환하여 학습 데이터로 사용할 수 있다.

【0034】 또 다른 일 실시예에 따르면, 전자 장치(101)(예: 도 1의 프로세서(120))는, 이성질체인 것으로 판단된 적어도 하나의 원자 기호 데이터를 제3 임계 값을 초과하면, 적어도 하나의 원자 기호 데이터에서 화학 구조식이 별도인 것으로

판단하여 적어도 하나의 원자 기호 데이터를 제2 노드로 변경하지 않을 수 있다. 구체적으로, 전자 장치(101)는, 도 6에 도시된 바와 같이, 이성질체이면서 화학 구조식이 동일하면 서로 다른 적어도 하나의 원자 기호 데이터 자체가 동일한 것으로 판단하여 별도의 학습을 수행하지 않음으로써, 제2 서브 딥러닝 모델의 학습 데이터 양을 최소화할 수 있다.

【0035】205 동작에서, 전자 장치(101)(예: 도 1의 프로세서(120))는, 적어도 하나의 화학 구조 결합 데이터 및 적어도 하나의 화학 선형 데이터를 예측 모델에 입력하여 적어도 하나의 화학 결과 데이터를 출력할 수 있다. 일 실시예에 따르면, 예측 모델은, 복수의 화학 구조 결합 데이터 및 복수의 화학 선형 데이터와, 복수의 화학 결과 데이터 및 복수의 전문의에 판별된 복수의 화학 제품에 기반하여 학습될 수 있다. 구체적으로, 적어도 하나의 화학 구조 결합 데이터 및 적어도 하나의 화학 선형 데이터를 기초로 적어도 하나의 화학 결과 데이터를 출력하기 위한 예측 모델은, 기초 데이터에서 지도 학습(unsupervised learning)을 통해 훈련된 AI 신경망 모델로 구동될 수 있다.

【0036】또한, 지도 학습 알고리즘은 예를 통해 학습하도록 설계될 수 있고, 인간인 사용자가 문제에 대한 답을 알고 있고, 인공지능(AI)이 그것을 알아낼 수 있도록 훈련시키고자 할 때 사용할 수 있다. 이는 예시로 교사의 도움을 받아 알고리즘을 '정답'으로 안내하면서 학습시키는 것과 같으므로 아이가 스스로 실험하고 시행착오를 거치며 배우는 비지도 학습 알고리즘과 다를 수 있다. 구체적인 예시로서, 지도 학습 알고리즘은 데이터에서 패턴을 찾아내는 데 능숙하기 때문에 이미지

인식 소프트웨어에 주로 탑재될 수 있다. 예를 들면, 알고리즘에 과일 바구니를 보여주고 구부러진 노란색 과일은 바나나, 둥근 모양의 오렌지색 과일은 오렌지 등으로 인식하도록 훈련하면 각 과일의 특성을 빠르게 학습하고 스스로 식별할 수 있게 될 수 있다. 예측 모델은 훈련 데이터(이 경우 과일 바구니에 해당)를 통해 학습하고 그렇게 학습한 지식을 새로운 테스트 데이터(이 경우 추가로 보여준 새로운 과일)에 적용하는 지도 학습 알고리즘의 좋은 예일 수 있다. 다른 예시로서, 지도 학습 알고리즘은 길찾기 앱들을 이용하여 특정 목적지까지 이동하는데 걸리는 시간을 상당히 정확하게 예측할 수 있다. 예시로, 지도 학습 알고리즘은 사용자에게 정확한 이동 시간을 알려주기 위해 하루 중 해당 시간대, 주중 해당 요일, 목적지까지의 거리, 기상 조건, 교통량 등 모든 종류의 데이터를 분석할 수 있다. 추가로, 지도 학습 알고리즘은 비슷한 조건에서 비슷한 경로, 심지어 동일한 경로로 이동하는 다른 사용자의 여정 데이터까지 분석하여 특정 사용자에게 예상 도착 시간을 알려주게 될 수 있다. 그리고 상단의 모든 과정에 대한 시간 소요는 단 몇 초 만에 행동이 완료될 수 있다. 지도 학습은 일련의 활용 가능한 참조 포인트(즉, 훈련 데이터) 또는 부정할 수 없는 진실(바나나의 모양)이 있어서 알고리즘을 훈련하는 데 사용할 수 있을 때 가장 적합할 수 있다. 다른 예를 들어, 예측 모델은 텍스트 데이터, 이미지 데이터, 및 음원 데이터 중 적어도 하나의 데이터를 기반으로 다중 모달 기반의 모델로 훈련시켜 다양한 이미지화를 구현할 수 있다. 일 실시예에서, 이는 기존의 딥러닝 방식보다 레이블이 지정된 작업별 훈련 데이터의 양을 감소할 수 있고, 한번 구축하면 적은 양의 훈련 데이터로 다양한 훈련이 가능하여 데이터

수집 및 레이블 지정이 용이하며, 정확도가 향상될 수 있다. 또 다른 일 실시예에 따르면, 전자 장치(101)는, 도 4 및 도 5에 도시된 바와 같이, 적어도 하나의 화학 구조 결합 데이터 및 적어도 하나의 화학 선형 데이터를 예측 모델에 입력하여 적어도 하나의 화학 결과 데이터를 출력할 수 있도록 예측하기 위한 예측 모델의 학습 과정으로서, 임의의 가중치들이 부여된 예측 모델을 이용하여 결과값(출력 데이터)을 획득하고, 획득된 결과값을 학습 데이터의 라벨링 데이터와 비교하고, 그 오차에 따라 역전파를 수행하여, 가중치들을 최적화함으로써 수행될 수 있다. 구체적으로, 예측 모델의 학습은 학습 데이터 및 라벨링 데이터 또는 라벨링 되지 않은 데이터를 기초로 예측 모델을 학습시켜, 예측 모델이 입력 데이터에 대하여 출력 데이터를 결정할 수 있도록 하는 과정을 의미한다. 즉, 예측 모델이 상기 데이터들에 대해 규칙을 형성하여 판단하는 것이다. 일 실시예에 따르면, 전자 장치(101)는 예측되는 값을 계산하는 복수의 학습 알고리즘 중 복수의 학습 알고리즘을 사용할 수 있다. 예를 들어, 예측 모델에 앙상블 기법(ensemble method)이 사용될 수 있으며, 학습 알고리즘들을 따로 쓰는 경우에 비해 더 좋은 예측 성능을 얻을 수 있다. 예측 모델을 학습시킨다는 의미는 모델이 가지고 있는 가중치(weight)를 조정하는 것을 의미할 수 있다. 일 실시예에 따라, 학습 방법으로서, 지도 학습(supervised learning), 비지도 학습(unsupervised learning), 강화 학습(reinforcement learning), 모방 학습(imitation learning), 연합 학습(federated-learning) 등 다양한 방법을 이용할 수 있다.

【0037】도시하지 않았지만, 전자 장치(101)는, 예측 모델의 학습 과정에 있어서, 상기 예측 모델의 성능을 평가하는 평가 단계를 포함할 수 있다. 평가 단계에서, 예측 모델은 평가 데이터 세트를 이용하여 평가될 수 있다. 예측 모델의 평가는 상기 학습 단계에 의해 학습된 예측 모델을 평가하고, 예측 모델을 이용하여 새로운 데이터에 대한 예측을 하는 단계일 수 있다. 구체적으로, 평가 단계는 학습된 예측 모델이 새로운 데이터에 대한 일반화(generalization)가 가능한지 측정하는 단계일 수 있다.

【0038】또 다른 일 실시예에 따르면, 예측 모델은, 적어도 하나의 화학 구조 결합 데이터의 제1 노드와 적어도 하나의 화학 선형 데이터의 제2 노드로부터 입력 값으로 사용되는 화합물을 구성할 수 있도록 다양한 특징들을 추출할 수 있다. 구체적으로, 본 실시예에서는, 원자 번호, 원자의 형식 전하, 원자의 반지름 크기, 원자의 방향족성 여부, 해당 원자에 대한 수소 원자의 수, 원자의 혼성화 상태, 동위 원소의 질량, 원자의 부분 전하, 수소 결합의 기부 여부, 수소 결합의 수용체 여부에 따라 결합의 특성을 손쉽게 파악할 수 있도록 제1 노드 및 제2 노드로부터 예측할 수 있다. 이후, 예측 모델은, 적어도 하나의 화학 구조 결합 데이터의 제1 행렬과 적어도 하나의 화학 선형 데이터의 제2 행렬로부터 단일, 이중, 삼중, 및 방향족 결합에 대한 유무를 판단하는 결합 유형과, 단일 결합 방향과, 결합 반지름 크기와, 결합 방향족성 여부와, 결합의 회전 가능 여부와, 결합의 스트레오 화학에 대해서 학습될 수 있도록 입체 화학적 구성을 출력할 수 있다.

【0039】또 다른 일 실시예에 따르면, 예측 모델은, 올바른 범주를 예측하는 분류 모델로서, 혼동 행렬을 평가 지표로 선정하여, 하기의 [수학식 1] 및 [표 1]을 기초로 정확도(accuracy), 정밀도(precision), 재현도(Recall) 및 F1-Score를 도출할 수 있다.

【0040】[표 1]

【0041】

		ATC Code						
		A	B	C	D	G	H	J
평가지표	accuracy	0.77	0.82	0.76	0.77	0.77	0.95	0.88
	Precision	0.796	0.828	0.793	0.77	0.774	0.953	0.88
	Recall	0.767	0.824	0.761	0.766	0.77	0.953	0.877
	F1-score	0.762	0.823	0.754	0.765	0.769	0.953	0.877

		ATC Code						
		L	M	N	P	R	S	V
평가지표	accuracy	0.79	0.78	0.81	0.77	0.8	0.72	0.85
	Precision	0.793	0.781	0.81	0.778	0.803	0.72	0.852
	Recall	0.792	0.778	0.805	0.775	0.798	0.719	0.852
	F1-score	0.792	0.777	0.805	0.774	0.797	0.719	0.852

【0042】[수학식 1]

【0043】(1) $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$

【0044】(2) $Precision = \frac{TP}{TP+TN}$

【0045】(3) $Recall = \frac{TP}{TP+FN}$

【0046】(4) $F1-score=2\times\frac{Precision\times Recall}{Precision+Recall}$

【0047】여기서, TP는 true positive로써, 실제 분류와 예측 분류가 동일하게 정답으로 추출한 값이고, TN은 true negative로써 실제 분류와 예측 분류가 동일하게 정답이 아닌 것으로 추출한 값이고, FP는 false positive로써, 실제 분류에서는 정답이 아니나 예측 분류가 정답인 것으로 추출한 값이고, FN은 false negative 값임.

【0048】즉, 전자 장치(101)는, 도 4 및 도 5에 도시된 바와 같이, 상기의 [수학식 1]과 [표 1]을 통하여, 예측 모델로부터 출력된 적어도 하나의 화학 결과 데이터가 정확한 값을 추출할 수 있고, 정확한 값을 비교하면서도 재현도까지 가능하도록 출력되는 것을 확인할 수 있다.

【0049】207 동작에서, 전자 장치(101)(예: 도 1의 프로세서(120))는, 적어도 하나의 화학 결과 데이터가 제1 임계 값을 초과하는지 여부를 판단할 수 있다. 일 실시예에 따르면, 제1 임계 값은, 소화관 및 대사, 혈액 및 조혈 기관, 심혈관계, 피부과, 비뇨생식기계 및 성호르몬, 성호르몬 이외의 전신성 호르몬 체제, 전신 작용 항감염체, 항종양제 및 조절제, 근 골격계, 신경계, 구충제, 살출제 및 기피제, 호흡기계, 감각기관, 기타 인체 기관에 적합하는지 판단하는 기준점으로 설정될 수 있다. 이때, 전자 장치(101)는, 상단에 기재된 14개의 데이터를 각각 학습하여 설정된 14개의 판단 인공지능 모델을 포함할 수 있다. 여기서, 14개의 판단 인공지능 모델은 제1 서브 딥러닝 모델, 제2 서브 딥러닝 모델, 및/또는 예측 모델

에 사용되는 모델들과 동일한 구성이되 각각 다른 입력 데이터, 출력 데이터, 및 학습 데이터로 구성될 수 있다. 이후, 전자 장치(101)는, 적어도 하나의 화학 결과 데이터를 14개의 판단 각각 입력하여 출력된 판단 데이터 중 적어도 하나 이상 적합한 것으로 판단된 데이터에 따라 14개의 데이터 중 매칭된 데이터를 손쉽게 판단할 수 있다. . 이후, 전자 장치(101)는, 제1 임계 값을 기초로 적어도 하나의 화학 결과 데이터를 비교하여 약품 및 인체에 사용될 수 있는지 유무를 판단할 수 있다.

【0050】209 동작에서, 전자 장치(101)(예: 도 1의 프로세서(120))는, 적어도 하나의 화학 결과 데이터가 제1 임계 값을 초과한 것으로 판단하면, 적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합한 기능을 갖춘 화학 제품으로 판단할 수 있다. 일 실시예에 따르면, 전자 장치(101)는, 소화관 및 대사, 혈액 및 조혈 기관, 심혈관계, 피부과, 비뇨생식기계 및 성호르몬, 성호르몬 이외의 전신성 호르몬 체제, 전신 작용 항감염체, 항종양제 및 조절제, 근 골격계, 신경계, 구충제, 살출제 및 기피제, 호흡기계, 감각기관, 기타 인체기관 데이터를 각각 학습한 14개의 판단 인공지능 모델에 각각 적어도 하나의 화학 결과 데이터를 입력하여 출력된 판단 데이터를 통하여 각각 구별하고, 14개의 판단 인공지능 모델에서 출력된 판단 데이터 중 적어도 하나 이상 적합한 것으로 판단하면 제1 임계 값을 초과한 것으로 판단할 수 있다. 이후, 전자 장치(101)는, 적어도 하나의 화학 결과 데이터가 신체 및/또는 인체에 적합한 화학 제품으로 판단할 수 있다. 이후, 전자 장치(101)는, 예측 모델에 해당 데이터가 명확한 값으로 판단된 적어도 하나의 화학 결과 데이터를 학습 데이터로 사용할 수 있다.

【0051】 한편, 207 동작에서, 전자 장치(101)(예: 도 1의 프로세서(120))는, 적어도 하나의 화학 결과 데이터가 제1 임계 값을 초과한 것으로 판단하면, 211 동작에서, 전자 장치(101)(예: 도 1의 프로세서(120))는, 적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합하지 않은 기능을 갖춘 화학 제품으로 판단할 수 있다. 일 실시예에 따르면, 전자 장치(101)는, 소화관 및 대사, 혈액 및 조혈기관, 심혈관계, 피부과, 비뇨생식기계 및 성호르몬, 성호르몬 이외의 전신성 호르몬 체제, 전신 작용 항감염체, 항종양제 및 조절제, 근 골격계, 신경계, 구충제, 살출제 및 기피제, 호흡기계, 감각기관, 기타 인체기관 데이터를 각각 학습한 14개의 판단 인공지능 모델에 각각 적어도 하나의 화학 결과 데이터를 입력하여 출력된 판단 데이터를 통하여 각각 구별하고, 14개의 판단 인공지능 모델에서 출력된 판단 데이터 전체가 적합하지 않은 것으로 판단하면 제1 임계 값을 초과하지 않은 것으로 판단할 수 있다. 이후, 전자 장치(101)는, 적어도 하나의 화학 결과 데이터가 신체 및/또는 인체에 적합하지 않은 화학 제품으로 판단할 수 있다. 이후, 전자 장치(101)는, 예측 모델에 해당 데이터가 정확하지 않은 값으로 판단된 적어도 하나의 화학 결과 데이터를 학습 데이터로 사용할 수 있다.

【0052】 213 동작에서, 전자 장치(101)(예: 도 1의 프로세서(120))는, 통신 인터페이스(도 1의 통신 인터페이스(160))를 통하여, 적어도 하나의 결과 데이터 및 적합한 기능 여부가 판단된 데이터를 외부 전자 장치(102, 104, 106)에 송신할 수 있다. 일 실시예에 따르면, 전자 장치(101)는, 통신 인터페이스(도 1의 통신 인터페이스(160))를 통하여, 제1 임계 값의 결과 유무에 따른 결과 데이터와 적어도

하나의 화학 결과 데이터를 어플리케이션(미도시)를 통하여 사용자가 손쉽게 입력한 적어도 하나의 화학 데이터에 대한 약물 적합판정을 사용자의 외부 전자 장치(102, 104, 106)를 통해 손쉽게 확인할 수 있다.

【0054】본 실시예는, 적어도 하나의 화학 이미지 데이터와 적어도 하나의 화학 텍스트 데이터를 각각 제1 서브 딥러닝 모델 및 제2 서브 딥러닝 모델에 입력하여 각각 적어도 하나의 화학 구조 결합 데이터와 적어도 하나의 화학 선형 데이터를 추출하고, 적어도 하나의 화학 구조 결합 데이터와 적어도 하나의 화학 선형 데이터를 예측 모델에 입력하여 출력된 결과 데이터로 손쉽게 약물에 적용되는지 유추가 가능함으로써, 개발 기간 및 시간 단축도 가능하면서도 일관성으로 해석될 수 있으며, 별도의 실험 없이 화학 데이터의 결과 값을 확인할 수 있는 이점이 있다.

【0056】다양한 실시예에 따르면, 예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치에 있어서, 통신 인터페이스; 및 프로세서; 를 포함하고, 상기 프로세서는, 상기 통신 인터페이스로부터, 화학식 구조로 구성된 적어도 하나의 화학 이미지 데이터 및 화학 수식으로 구성된 적어도 하나의 화학 텍스트 데이터를 획득하고, 상기 적어도 하나의 화학 이미지 데이터를 제1 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 구조 결합 데이터를 출력하고, 상기 적어도 하나의 화학 텍스트 데이터를 제2 서브 딥러닝 모델에 입력하여

적어도 하나의 화학 선형 결합 데이터를 출력하고, 상기 적어도 하나의 화학 구조 결합 데이터 및 상기 적어도 하나의 화학 선형 데이터를 상기 예측 모델에 입력하여 적어도 하나의 화학 결과 데이터를 출력하고, 및 상기 적어도 하나의 결과 데이터가 기 설정된 제1 임계 값을 초과하면, 상기 적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합한 기능을 갖춘 화학 제품으로 판단하도록 설정되고, 상기 예측 모델은, 복수의 화학 구조 결합 데이터 및 복수의 화학 선형 데이터와, 복수의 화학 결과 데이터 및 복수의 전문의에 판별된 복수의 화학 제품에 기반하여 학습된다.

【0057】다양한 실시예에 따르면, 상기 적어도 하나의 화학 이미지 데이터는, 적어도 하나의 원자 이미지 데이터와, 상기 적어도 하나의 원자 이미지 데이터 간을 각각 연결하는 적어도 하나의 원자 공유 결합 이미지 데이터를 포함하고, 상기 적어도 하나의 화학 텍스트 데이터는, 상기 적어도 하나의 원자 이미지 데이터와 매칭되는 적어도 하나의 원자 기호 데이터, 상기 적어도 하나의 화학 이미지 데이터에 복수개의 원자 및 결합 구조에 따른 원자 개수 데이터, 및 상기 적어도 하나의 화학 이미지 데이터의 결합구조와 매칭되는 적어도 하나의 괄호 데이터를 포함하도록 설정된다.

【0058】다양한 실시예에 따르면, 상기 프로세서는, 상기 적어도 하나의 원자 이미지 데이터를 각각 적어도 하나의 제1 노드로 변경하면서, 상기 적어도 하나의 제1 노드를 제1 행렬에 결합하고, 상기 적어도 하나의 원자 공유 결합 이미지 데이터를 각각 적어도 하나의 제1 엣지로 변경하면서, 상기 적어도 하나의 제1 엣

지를 제2 행렬에 결합하고, 및 상기 제1 행렬 및 상기 제2 행렬을 상기 제1 딥러닝 모델에 입력하여 상기 적어도 하나의 화학 구조 결합 데이터를 출력하도록 설정하고, 상기 제1 딥러닝 모델은, 복수의 제1 노드, 복수의 제1 엣지, 복수의 제1 행렬, 및 복수의 제2 행렬과, 복수의 화학 구조 결합 데이터에 기반하여 학습된다.

【0059】다양한 실시예에 따르면, 상기 프로세서는, 상기 적어도 하나의 원자 기호 데이터에 결합 요소에 따라 각각 기호를 변경하고, 변경된 상기 적어도 하나의 원자 기호 데이터를 각각 적어도 하나의 제2 노드로 변경하고, 복수개의 원자 기호 데이터가 있는 상기 적어도 하나의 제2 노드에 상기 적어도 하나의 원자 개수 데이터를 적용하여 적어도 하나의 제3 노드를 적용하고, 상기 적어도 하나의 괄호 데이터를 각각 적어도 하나의 제2 엣지로 변경하고, 및 상기 제2 노드, 상기 제3 노드, 및 상기 제2 엣지를 상기 제2 딥러닝 모델에 입력하여 상기 적어도 하나의 화학 선형 결합 데이터를 출력하도록 설정되고, 상기 제2 서브 딥러닝 모델은, 복수의 제2 노드, 복수의 제3 노드, 및 복수의 제2 엣지와, 복수의 화학 선형 결합 데이터, 복수의 스마일식 코드, 및 ACS|| 인코딩 방식 코드에 기반하여 학습된다.

【0060】다양한 실시예에 따르면, 상기 프로세서는, 상기 적어도 하나의 원자 기호 데이터에서 기 설정된 제2 임계 값을 초과하지 않으면, 상기 적어도 하나의 원자 기호 데이터에서 이성질체가 아닌 것으로 판단하고, ATC(anatomical therapeutic chemical) 레벨 명칭을 기반으로 상기 적어도 하나의 원자 기호 데이터를 상기 제2 노드로 변경하고, 상기 적어도 하나의 원자 기호 데이터에서 상기 제2 임계 값을 초과하면, 상기 적어도 하나의 원자 기호 데이터에서 이성질체인 것

으로 판단하고, 상기 이성질체인 것으로 판단된 상기 적어도 하나의 원자 기호 데이터를 기 설정된 제3 임계 값을 초과하지 않으면, 상기 적어도 하나의 원자 기호 데이터에서 화학 구조식이 별도인 것으로 판단하여, 상기 적어도 하나의 원자 기호 데이터를 상기 괄호 형태에 따라 변경된 상태로 상기 제2 노드로 변경하고, 및 상기 이성질체인 것으로 판단된 상기 적어도 하나의 원자 기호 데이터를 상기 제3 임계 값을 초과하면, 상기 적어도 하나의 원자 기호 데이터에서 화학 구조식이 별도인 것으로 판단하여 상기 적어도 하나의 원자 기호 데이터를 상기 제2 노드로 변경하지 않도록 설정된다.

【0061】또 다른 다양한 실시예에 따르면, 예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치의 구동 방법에 있어서, 상기 방법은, 통신 인터페이스로부터, 화학식 구조로 구성된 적어도 하나의 화학 이미지 데이터 및 화학 수식으로 구성된 적어도 하나의 화학 텍스트 데이터를 획득하고, 프로세서를 통하여, 상기 적어도 하나의 화학 이미지 데이터를 제1 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 구조 결합 데이터를 출력하고, 상기 적어도 하나의 화학 텍스트 데이터를 제2 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 선형 결합 데이터를 출력하고, 상기 프로세서를 통하여, 상기 적어도 하나의 화학 구조 결합 데이터 및 상기 적어도 하나의 화학 선형 데이터를 상기 예측 모델에 입력하여 적어도 하나의 화학 결과 데이터를 출력하고, 및 상기 프로세서를 통하여, 상기 적어도 하나의 결과 데이터가 기 설정된 제1 임계 값을 초과하면, 상기 적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합한 기능을 갖춘 화학 제품으

로 판단하도록 설정되고, 상기 예측 모델은, 복수의 화학 구조 결합 데이터 및 복수의 화학 선형 데이터와, 복수의 화학 결과 데이터 및 복수의 전문의에 판별된 복수의 화학 제품에 기반하여 학습된다.

【0063】본 문서에서 사용된 용어 "모듈" 또는 "~부"은 하드웨어, 소프트웨어 또는 펌웨어로 구성된 유닛을 포함하며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로 등의 용어와 상호 호환적으로 사용될 수 있다. "모듈" 또는 "~부"는, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는 최소 단위 또는 그 일부가 될 수 있다. "모듈" 또는 "~부"는 기계적으로 또는 전자적으로 구현될 수 있으며, 예를 들면, 어떤 동작들을 수행하는, 알려졌거나 앞으로 개발될, ASIC(application-specific integrated circuit) 칩, FPGAs(field-programmable gate arrays), 또는 프로그램 가능 논리 장치를 포함할 수 있고, 프로세서(120)에 의하여 실행될 수 있다. 다양한 실시예에 따른 장치(예: 모듈들 또는 그 기능들) 또는 방법(예: 동작들)의 적어도 일부는 프로그램 모듈의 형태로 컴퓨터로 판독 가능한 저장 매체(예: 메모리(130))에 저장된 명령어로 구현될 수 있다. 상기 명령어가 프로세서(예: 프로세서(120))에 의해 실행될 경우, 프로세서가 상기 명령어에 해당하는 기능을 수행할 수 있다. 컴퓨터로 판독 가능한 기록 매체는, 하드디스크, 플로피디스크, 마그네틱 매체(예: 자기테이프), 광기록 매체(예: CD-ROM, DVD, 자기-광 매체 (예: 플롭티컬 디스크), 내장 메모리 등을 포함할 수 있다. 명령어는 컴파일러에 의해 만들어지는 코드 또는 인터프리터에 의해 실행될 수 있는 코드를

포함할 수 있다. 다양한 실시예에 따른 모듈 또는 프로그램 모듈은 전술한 구성요소들 중 적어도 하나 이상을 포함하거나, 일부가 생략되거나, 또는 다른 구성요소를 더 포함할 수 있다. 다양한 실시예에 따른, 모듈, 프로그램 모듈 또는 다른 구성요소에 의해 수행되는 동작들은 순차적, 병렬적, 반복적 또는 휴리스틱하게 실행되거나, 적어도 일부 동작이 다른 순서로 실행되거나, 생략되거나, 또는 다른 동작이 추가될 수 있다.

【0064】그리고 본 문서에 개시된 실시예는 개시된, 기술 내용의 설명 및 이해를 위해 제시된 것이며, 본 개시의 범위를 한정하는 것은 아니다. 따라서, 본 개시의 범위는, 본 개시의 기술적 사상에 근거한 모든 변경 또는 다양한 다른 실시예를 포함하는 것으로 해석되어야 한다.

【청구범위】

【청구항 1】

예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치에 있어서,

통신 인터페이스; 및

프로세서; 를 포함하고,

상기 프로세서는,

상기 통신 인터페이스로부터, 화학식 구조로 구성된 적어도 하나의 화학 이미지 데이터 및 화학 수식으로 구성된 적어도 하나의 화학 텍스트 데이터를 획득하고,

상기 적어도 하나의 화학 이미지 데이터를 제1 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 구조 결합 데이터를 출력하고, 상기 적어도 하나의 화학 텍스트 데이터를 제2 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 선형 결합 데이터를 출력하고,

상기 적어도 하나의 화학 구조 결합 데이터 및 상기 적어도 하나의 화학 선형 데이터를 상기 예측 모델에 입력하여 적어도 하나의 화학 결과 데이터를 출력하고, 및

상기 적어도 하나의 결과 데이터가 기 설정된 제1 임계 값을 초과하면, 상기 적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합한 기능을 갖춘 화학 제

품으로 판단하도록 설정되고,

상기 예측 모델은,

복수의 화학 구조 결합 데이터 및 복수의 화학 선형 데이터와, 복수의 화학 결과 데이터 및 복수의 전문의에 판별된 복수의 화학 제품에 기반하여 학습되는,

전자 장치.

【청구항 2】

제1 항에 있어서,

상기 적어도 하나의 화학 이미지 데이터는,

적어도 하나의 원자 이미지 데이터와, 상기 적어도 하나의 원자 이미지 데이터 간을 각각 연결하는 적어도 하나의 원자 공유 결합 이미지 데이터를 포함하고,

상기 적어도 하나의 화학 텍스트 데이터는,

상기 적어도 하나의 원자 이미지 데이터와 매칭되는 적어도 하나의 원자 기호 데이터, 상기 적어도 하나의 화학 이미지 데이터에 복수개의 원자 및 결합 구조에 따른 원자 개수 데이터, 및 상기 적어도 하나의 화학 이미지 데이터의 결합구조와 매칭되는 적어도 하나의 괄호 데이터를 포함하도록 설정되는,

전자 장치.

【청구항 3】

제2 항에 있어서,

상기 프로세서는,

상기 적어도 하나의 원자 이미지 데이터를 각각 적어도 하나의 제1 노드로 변경하면서, 상기 적어도 하나의 제1 노드를 제1 행렬에 결합하고,

상기 적어도 하나의 원자 공유 결합 이미지 데이터를 각각 적어도 하나의 제1 엣지로 변경하면서, 상기 적어도 하나의 제1 엣지를 제2 행렬에 결합하고, 및

상기 제1 행렬 및 상기 제2 행렬을 상기 제1 딥러닝 모델에 입력하여 상기 적어도 하나의 화학 구조 결합 데이터를 출력하도록 설정하고,

상기 제1 딥러닝 모델은,

복수의 제1 노드, 복수의 제1 엣지, 복수의 제1 행렬, 및 복수의 제2 행렬과, 복수의 화학 구조 결합 데이터에 기반하여 학습되는,

전자 장치.

【청구항 4】

제3 항에 있어서,

상기 프로세서는,

상기 적어도 하나의 원자 기호 데이터에 결합 요소에 따라 각각 기호를 변경하고, 변경된 상기 적어도 하나의 원자 기호 데이터를 각각 적어도 하나의 제2 노

드로 변경하고,

복수개의 원자 기호 데이터가 있는 상기 적어도 하나의 제2 노드에 상기 적어도 하나의 원자 개수 데이터를 적용하여 적어도 하나의 제3 노드를 적용하고,

상기 적어도 하나의 괄호 데이터를 각각 적어도 하나의 제2 엣지로 변경하고, 및

상기 제2 노드, 상기 제3 노드, 및 상기 제2 엣지를 상기 제2 딥러닝 모델에 입력하여 상기 적어도 하나의 화학 선형 결합 데이터를 출력하도록 설정되고,

상기 제2 서브 딥러닝 모델은,

복수의 제2 노드, 복수의 제3 노드, 및 복수의 제2 엣지와, 복수의 화학 선형 결합 데이터, 복수의 스마일식 코드, 및 ACS|| 인코딩 방식 코드에 기반하여 학습되는,

전자 장치.

【청구항 5】

제4 항에 있어서,

상기 프로세서는,

상기 적어도 하나의 원자 기호 데이터에서 기 설정된 제2 임계 값을 초과하지 않으면, 상기 적어도 하나의 원자 기호 데이터에서 이성질체가 아닌 것으로 판단하고, ATC(anatomical therapeutic chemical) 레벨 명칭을 기반으로 상기 적어도

하나의 원자 기호 데이터를 상기 제2 노드로 변경하고,

상기 적어도 하나의 원자 기호 데이터에서 상기 제2 임계 값을 초과하면, 상기 적어도 하나의 원자 기호 데이터에서 이성질체인 것으로 판단하고,

상기 이성질체인 것으로 판단된 상기 적어도 하나의 원자 기호 데이터를 기설정된 제3 임계 값을 초과하지 않으면, 상기 적어도 하나의 원자 기호 데이터에서 화학 구조식이 별도인 것으로 판단하여, 상기 적어도 하나의 원자 기호 데이터를 상기 괄호 형태에 따라 변경된 상태로 상기 제2 노드로 변경하고, 및

상기 이성질체인 것으로 판단된 상기 적어도 하나의 원자 기호 데이터를 상기 제3 임계 값을 초과하면, 상기 적어도 하나의 원자 기호 데이터에서 화학 구조식이 별도인 것으로 판단하여 상기 적어도 하나의 원자 기호 데이터를 상기 제2 노드로 변경하지 않도록 설정되는,

전자 장치.

【청구항 6】

예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치의 구동 방법에 있어서,

상기 방법은,

통신 인터페이스로부터, 화학식 구조로 구성된 적어도 하나의 화학 이미지 데이터 및 화학 수식으로 구성된 적어도 하나의 화학 텍스트 데이터를 획득하고,

프로세서를 통하여, 상기 적어도 하나의 화학 이미지 데이터를 제1 서브 딥 러닝 모델에 입력하여 적어도 하나의 화학 구조 결합 데이터를 출력하고, 상기 적어도 하나의 화학 텍스트 데이터를 제2 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 선형 결합 데이터를 출력하고,

상기 프로세서를 통하여, 상기 적어도 하나의 화학 구조 결합 데이터 및 상기 적어도 하나의 화학 선형 데이터를 상기 예측 모델에 입력하여 적어도 하나의 화학 결과 데이터를 출력하고, 및

상기 프로세서를 통하여, 상기 적어도 하나의 결과 데이터가 기 설정된 제1 임계 값을 초과하면, 상기 적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합한 기능을 갖춘 화학 제품으로 판단하도록 설정되고,

상기 예측 모델은,

복수의 화학 구조 결합 데이터 및 복수의 화학 선형 데이터와, 복수의 화학 결과 데이터 및 복수의 전문의에 판별된 복수의 화학 제품에 기반하여 학습되는,

방법.

【청구항 7】

제6 항에 있어서,

상기 적어도 하나의 화학 이미지 데이터는,

적어도 하나의 원자 이미지 데이터와, 상기 적어도 하나의 원자 이미지 데이

터 간을 각각 연결하는 적어도 하나의 원자 공유 결합 이미지 데이터를 포함하고,

상기 적어도 하나의 화학 텍스트 데이터는,

상기 적어도 하나의 원자 이미지 데이터와 매칭되는 적어도 하나의 원자 기호 데이터, 상기 적어도 하나의 화학 이미지 데이터에 복수개의 원자 및 결합 구조에 따른 원자 개수 데이터, 및 상기 적어도 하나의 화학 이미지 데이터의 결합구조와 매칭되는 적어도 하나의 괄호 데이터를 포함하도록 설정되는,

방법.

【청구항 8】

제7 항에 있어서,

상기 방법은,

상기 적어도 하나의 원자 이미지 데이터를 각각 적어도 하나의 제1 노드로 변경하면서, 상기 적어도 하나의 제1 노드를 제1 행렬에 결합하고,

상기 적어도 하나의 원자 공유 결합 이미지 데이터를 각각 적어도 하나의 제1 엣지로 변경하면서, 상기 적어도 하나의 제1 엣지를 제2 행렬에 결합하고, 및

상기 제1 행렬 및 상기 제2 행렬을 상기 제1 딥러닝 모델에 입력하여 상기 적어도 하나의 화학 구조 결합 데이터를 출력하도록 설정하고,

상기 제1 딥러닝 모델은,

복수의 제1 노드, 복수의 제1 엣지, 복수의 제1 행렬, 및 복수의 제2

행렬과, 복수의 화학 구조 결합 데이터에 기반하여 학습되는,

방법.

【청구항 9】

제8 항에 있어서,

상기 방법은,

상기 적어도 하나의 원자 기호 데이터에 결합 요소에 따라 각각 기호를 변경하고, 변경된 상기 적어도 하나의 원자 기호 데이터를 각각 적어도 하나의 제2 노드로 변경하고,

복수개의 원자 기호 데이터가 있는 상기 적어도 하나의 제2 노드에 상기 적어도 하나의 원자 개수 데이터를 적용하여 적어도 하나의 제3 노드를 적용하고,

상기 적어도 하나의 괄호 데이터를 각각 적어도 하나의 제2 엣지로 변경하고, 및

상기 제2 노드, 상기 제3 노드, 및 상기 제2 엣지를 상기 제2 딥러닝 모델에 입력하여 상기 적어도 하나의 화학 선형 결합 데이터를 출력하도록 설정되고,

상기 제2 서브 딥러닝 모델은,

복수의 제2 노드, 복수의 제3 노드, 및 복수의 제2 엣지와, 복수의 화학 선형 결합 데이터, 복수의 스마일식 코드, 및 ACS|| 인코딩 방식 코드에 기반하여 학습되는,

방법.

【청구항 10】

제9 항에 있어서,

상기 방법은,

상기 적어도 하나의 원자 기호 데이터에서 기 설정된 제2 임계 값을 초과하지 않으면, 상기 적어도 하나의 원자 기호 데이터에서 이성질체가 아닌 것으로 판단하고, ATC(anatomical therapeutic chemical) 레벨 명칭을 기반으로 상기 적어도 하나의 원자 기호 데이터를 상기 제2 노드로 변경하고,

상기 적어도 하나의 원자 기호 데이터에서 상기 제2 임계 값을 초과하면, 상기 적어도 하나의 원자 기호 데이터에서 이성질체인 것으로 판단하고,

상기 이성질체인 것으로 판단된 상기 적어도 하나의 원자 기호 데이터를 기 설정된 제3 임계 값을 초과하지 않으면, 상기 적어도 하나의 원자 기호 데이터에서 화학 구조식이 별도인 것으로 판단하여, 상기 적어도 하나의 원자 기호 데이터를 상기 괄호 형태에 따라 변경된 상태로 상기 제2 노드로 변경하고, 및

상기 이성질체인 것으로 판단된 상기 적어도 하나의 원자 기호 데이터를 상기 제3 임계 값을 초과하면, 상기 적어도 하나의 원자 기호 데이터에서 화학 구조식이 별도인 것으로 판단하여 상기 적어도 하나의 원자 기호 데이터를 상기 제2 노드로 변경하지 않도록 설정되는,

방법.

【요약서】

【요약】

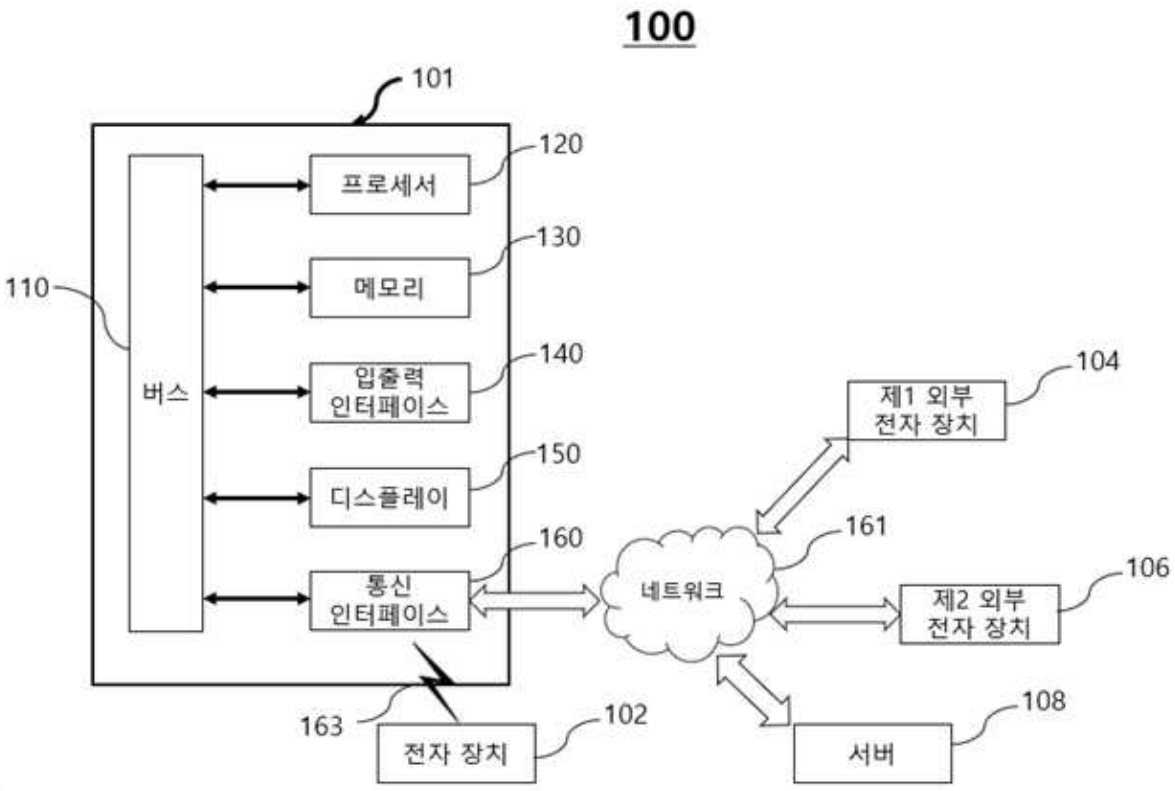
다양한 실시예에 따르면, 예측 모델 기반 ATC(anatomical therapeutic chemical) 코드의 분류를 위한 전자 장치에 있어서, 통신 인터페이스; 및 프로세서; 를 포함하고, 상기 프로세서는, 상기 통신 인터페이스로부터, 화학식 구조로 구성된 적어도 하나의 화학 이미지 데이터 및 화학 수식으로 구성된 적어도 하나의 화학 텍스트 데이터를 획득하고, 상기 적어도 하나의 화학 이미지 데이터를 제1 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 구조 결합 데이터를 출력하고, 상기 적어도 하나의 화학 텍스트 데이터를 제2 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 선형 결합 데이터를 출력하고, 상기 적어도 하나의 화학 구조 결합 데이터 및 상기 적어도 하나의 화학 선형 데이터를 상기 예측 모델에 입력하여 적어도 하나의 화학 결과 데이터를 출력하고, 및 상기 적어도 하나의 결과 데이터가 기 설정된 제1 임계 값을 초과하면, 상기 적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합한 기능을 갖춘 화학 제품으로 판단하도록 설정되고, 상기 예측 모델은, 복수의 화학 구조 결합 데이터 및 복수의 화학 선형 데이터와, 복수의 화학 결과 데이터 및 복수의 전문의에 판별된 복수의 화학 제품에 기반하여 학습된다.

【대표도】

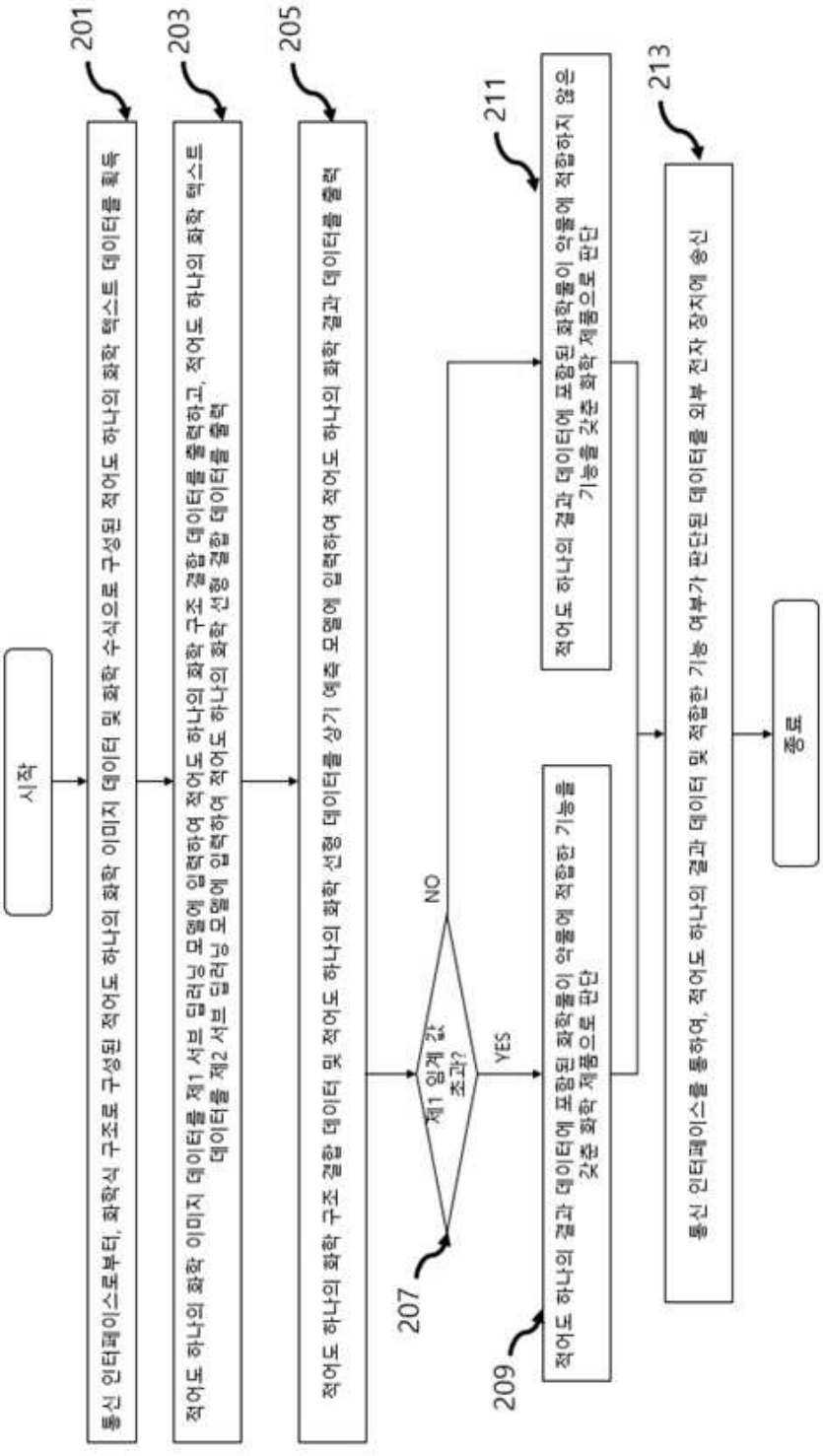
도 5

【도면】

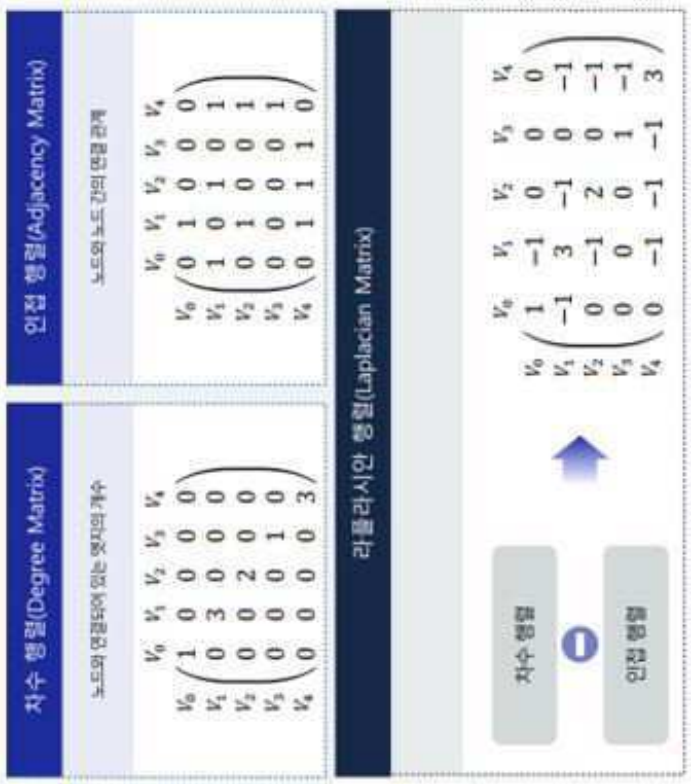
【도 1】


100
101
120
프로세서
130
메모리
110
버스
140
입출력
인터페이스
150
디스플레이
160
통신
인터페이스
163
전자 장치
102
104
제1 외부
전자 장치
161
네트워크
106
제2 외부
전자 장치
108
서버

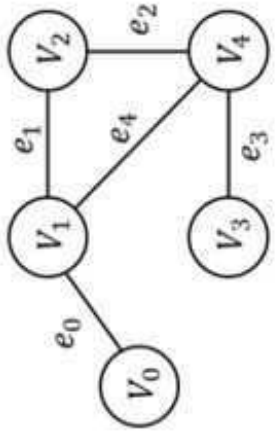
【도 2】


시작
201
통신 인터페이스로부터, 화학식 구조로 구성된 적어도 하나의 화학 이미지 데이터 및 화학 수식으로 구성된 적어도 하나의 화학 텍스트 데이터를 획득
203
적어도 하나의 화학 이미지 데이터를 제1 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 구조 결합 데이터를 출력하고, 적어도 하나의 화학 텍스트 데이터를 제2 서브 딥러닝 모델에 입력하여 적어도 하나의 화학 선형 결합 데이터를 출력
205
적어도 하나의 화학 구조 결합 데이터 및 적어도 하나의 화학 선형 데이터를 상기 예측 모델에 입력하여 적어도 하나의 화학 결과 데이터를 출력
207
제1 임계 값 초과?
NO
YES
209
적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합한 기능을 갖춘 화학 제품으로 판단
211
적어도 하나의 결과 데이터에 포함된 화학물이 약물에 적합하지 않은 기능을 갖춘 화학 제품으로 판단
213
통신 인터페이스를 통하여, 적어도 하나의 결과 데이터 및 적합한 기능 여부가 판단된 데이터를 외부 전자 장치에 송신
종료

【도 3】

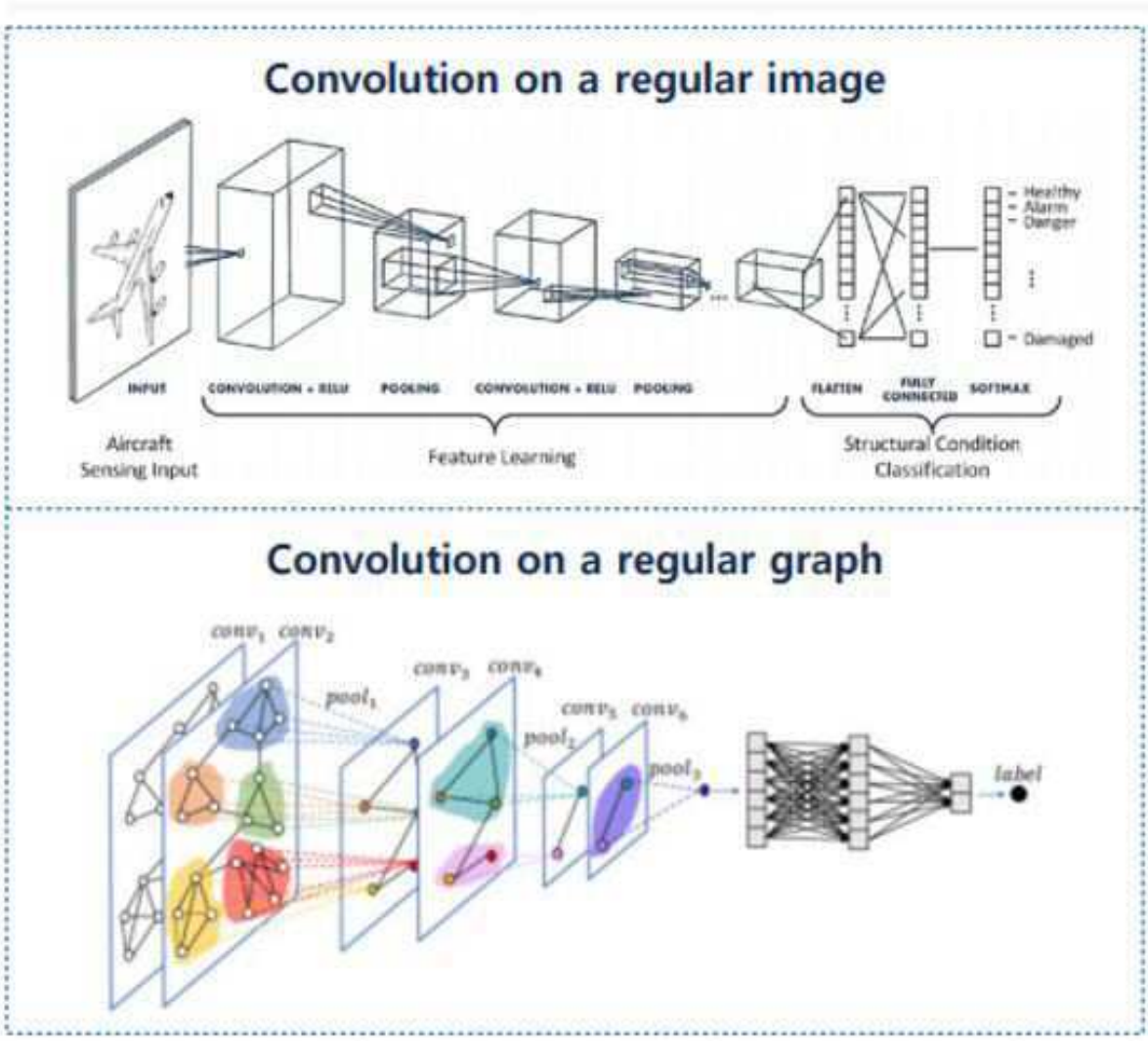

차수 행렬(Degree Matrix)
노드와 연결되어 있는 엣지의 개수
인접 행렬(Adjacency Matrix)
노드와 노드 간의 연결 관계
라플라시안 행렬(Laplacian Matrix)
차수 행렬
인접 행렬

(b)

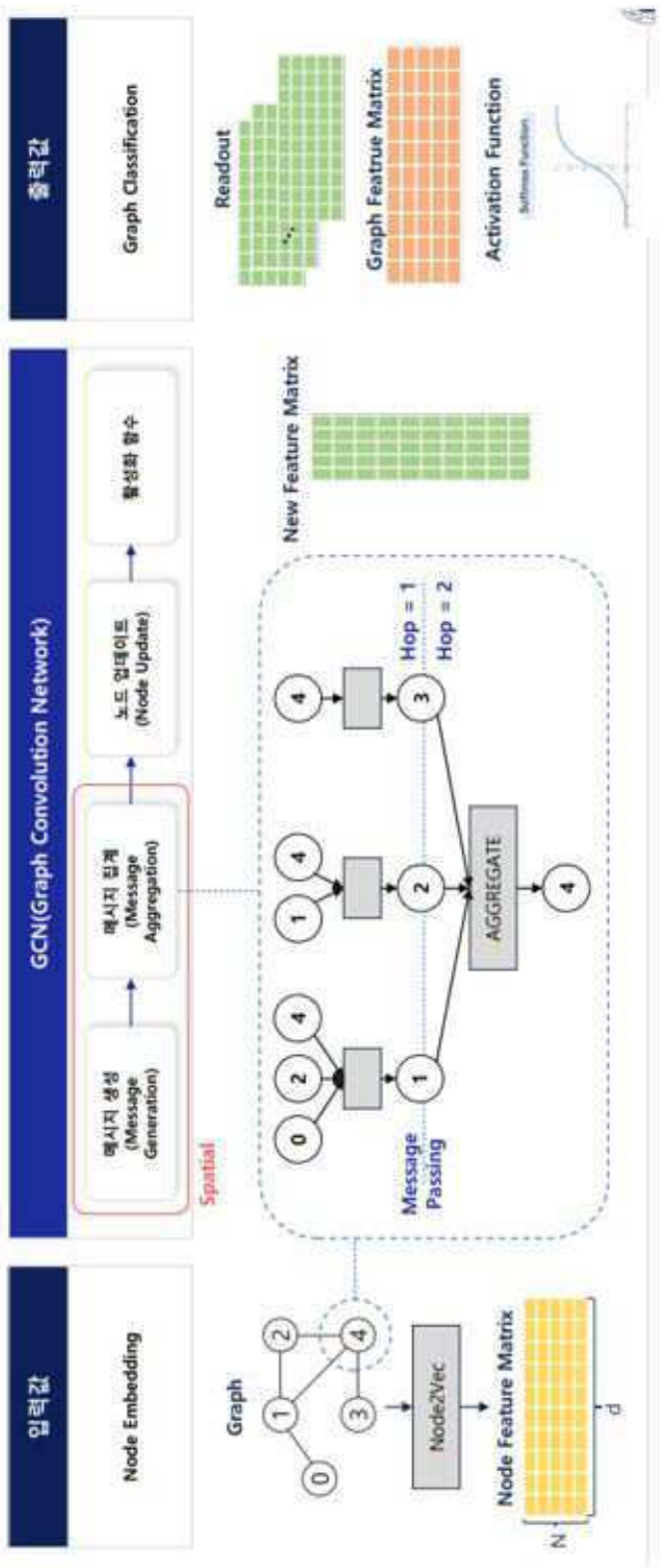

V_0
V_1
V_2
V_3
V_4
e_0
e_1
e_2
e_3
e_4

(a)

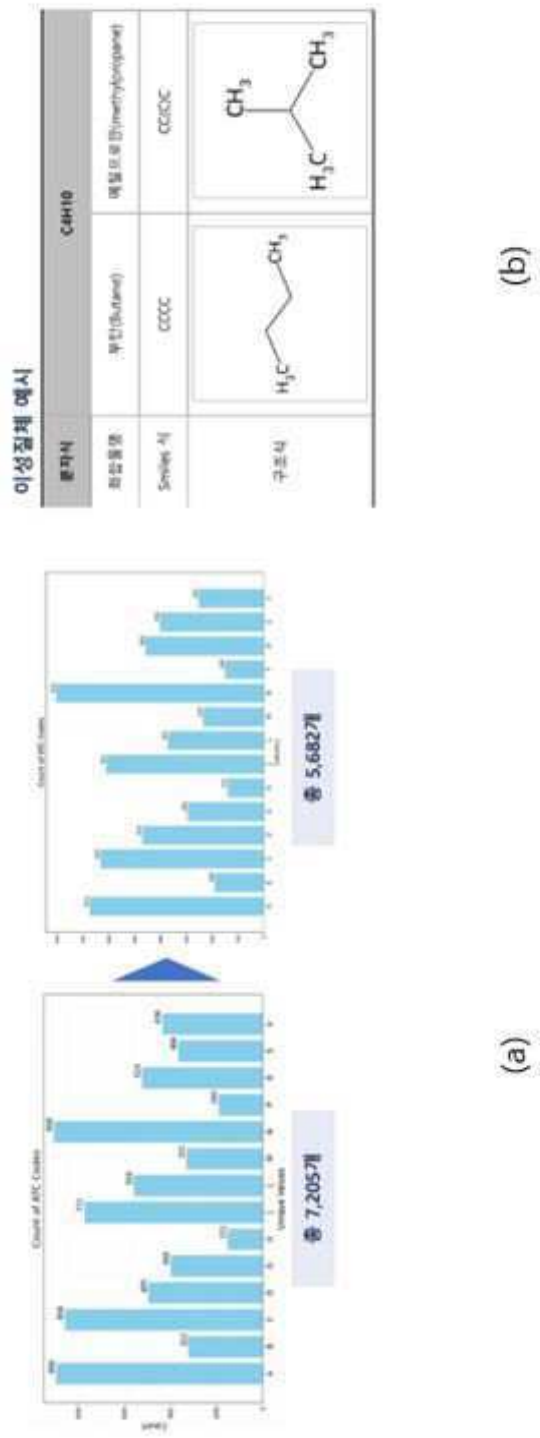
【도 4】


Convolution on a regular image
Healthy
Alarm
Danger
Damaged
INPUT
CONVOLUTION + RELU
POOLING
CONVOLUTION + RELU
POOLING
FLATTEN
FULLY CONNECTED
SOFTMAX
Aircraft Sensing Input
Feature Learning
Structural Condition Classification
Convolution on a regular graph
label

【도 5】


입력값
Node Embedding
Graph
Node2Vec
Node Feature Matrix
N
d
GCN(Graph Convolution Network)
메시지 생성 (Message Generation)
메시지 집계 (Message Aggregation)
노드 업데이트 (Node Update)
활성화 함수
Spatial
Message Passing
AGGREGATE
Hop = 1
Hop = 2
New Feature Matrix
출력값
Graph Classification
Readout
Graph Featrue Matrix
Activation Function

【도 6】


(a)
총 7,205개
총 5,682개
(b)
이성질체 예시
분자식
C4H10
화합물명
부탄(Butane)
메틸프로판(2-methylpropane)
Smiles 식
CCCC
CC(C)C
구조식