

**(19) 대한민국특허청(KR)**
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0142374

(43) 공개일자 2023년10월11일

(51) 국제특허분류(Int. Cl.)

G06Q 30/02 (2023.01) G06N 3/0475 (2023.01)

G06N 3/08 (2023.01) G06N 3/094 (2023.01)

G06Q 10/06 (2012.01)

(52) CPC특허분류

G06Q 30/0202 (2023.01)

G06N 3/0475 (2023.01)

(21) 출원번호 10-2023-0042911

(22) 출원일자 2023년03월31일

심사청구일자 2023년03월31일

(30) 우선권주장

1020220041011 2022년04월01일 대한민국(KR)

(71) 출원인

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

주식회사 메디밸류

서울특별시 영등포구 여의공원로 115, 6층(여의도동, 세우빌딩)

(72) 발명자

석준희

서울특별시 강남구 역삼로 315-1, 502동 2404호

김인수

서울특별시 성북구

(74) 대리인

김등용

전체 청구항 수 : 총 6 항

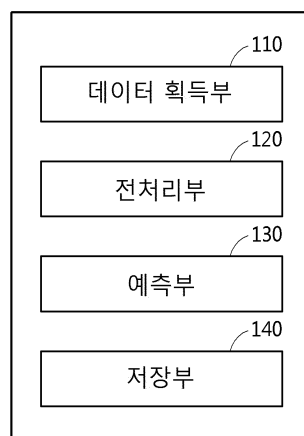
(54) 발명의 명칭 인공지능 생성 모델 기반의 의약품 재구매 시점 및 수량 예측 장치 및 방법

(57) 요약

의약품 재구매 시점 및 수량 예측 장치 및 방법이 개시된다. 상기 의약품 재구매 시점 및 수량 예측 방법은 컴퓨팅 장치에 의해 수행되고, 제1 구매기록 데이터를 획득하는 단계, 상기 제1 구매기록 데이터를 전처리하는 단계, 전처리된 제1 구매기록 데이터로부터 예측 모델의 학습 데이터를 생성하는 단계, 상기 학습 데이터를 이용하여 적대적 생성 신경망(Generative Adversarial Network, GAN)을 학습하여 상기 예측 모델을 생성하는 단계, 및 제2 구매기록 데이터에 대한 예측을 수행하여, 의약품 재구매 시점 및 수량을 도출하는 단계를 포함한다.

대표도 - 도1

100



(52) CPC특허분류

G06N 3/088 (2023.01)

G06N 3/094 (2023.01)

G06Q 10/06315 (2013.01)

명세서

청구범위

청구항 1

컴퓨팅 장치에 의해 수행되는 의약품 재구매 시점 및 수량 예측 방법에 있어서,

제1 구매기록 데이터를 획득하는 단계;

상기 제1 구매기록 데이터를 전처리하는 단계;

전처리된 제1 구매기록 데이터로부터 예측 모델의 학습 데이터를 생성하는 단계;

상기 학습 데이터를 이용하여 적대적 생성 신경망(Generative Adversarial Network, GAN)을 학습하여 상기 예측 모델을 생성하는 단계; 및

제2 구매기록 데이터에 대한 예측을 수행하여, 의약품 재구매 시점 및 수량을 도출하는 단계를 포함하는 의약품 재구매 시점 및 수량 예측 방법.

청구항 2

제1항에 있어서,

상기 의약품 재구매 시점 및 수량을 도출하는 단계는,

상기 제2 구매기록 데이터에 대응하는 입력 데이터를 상기 예측 모델에 입력하여 제1 예측 값 및 상기 제1 예측 값에 대한 신뢰도(k)를 도출하는 단계;

상기 제2 구매기록 데이터에 대한 평균(m)을 산출하여 제2 예측 값을 도출하는 단계; 및

상기 제1 예측 값과 상기 제2 예측 값의 가중 평균하여 최종 예측 값을 도출하는 단계를 포함하는,

의약품 재구매 시점 및 수량 예측 방법.

청구항 3

제2항에 있어서,

상기 최종 예측 값(T_{Final})은 수학적식을 이용하여 산출되고,

상기 수학적식은 $T_{Final} = G(T_P, z) * k + (1 - k) * m$ 이고,

상기 $G(T_P, z)$ 는 상기 입력 데이터(T_P)와 노이즈(z)를 입력으로 하는 적대적 생성 신경망 기반의 상기 예측 모델에 포함된 생성기(Generator)의 출력인,

의약품 재구매 시점 및 수량 예측 방법.

청구항 4

제3항에 있어서,

상기 학습 데이터를 생성하는 단계는,

수학적식 2를 이용하여 상기 전처리된 제1 구매기록 데이터로부터 상기 적대적 생성 신경망의 출력을 생성하는 단계; 및

상기 전처리된 제1 구매기록 데이터를 미리 정해진 크기로 분할하여 상기 적대적 생성 신경망의 입력을 생성하는 단계를 포함하는,

의약품 재구매 시점 및 수량 예측 방법.

청구항 5

컴퓨팅 장치에 의해 수행되는 의약품 재구매 시점 및 수량 예측 방법에 있어서,

구매기록 데이터를 획득하는 단계;

상기 구매기록 데이터를 전처리하는 단계;

전처리된 구매기록 데이터를 예측 모델에 입력하여 제1 예측 값 및 제1 예측 값에 대한 신뢰도(k)를 도출하는 단계;

상기 구매기록 데이터의 평균(m)을 산출하여 제2 예측 값을 도출하는 단계; 및

상기 제1 예측 값과 상기 제2 예측 값으로부터 최종 결과를 산출하는 단계를 포함하고,

상기 최종 결과(T_{Final})는 수학적식을 이용하여 산출되고,

상기 수학적식은 $T_{Final} = G(T_P, z) * k + (1 - k) * m$ 이고,

상기 $G(T_P, z)$ 는 상기 전처리된 구매기록 데이터(T_P)와 노이즈(z)를 입력으로 하는 적대적 생성 신경망 기반의 상기 예측 모델에 포함된 생성기(Generator)의 출력인,

의약품 재구매 시점 및 수량 예측 방법.

청구항 6

제1 구매기록 데이터를 획득하는 데이터 획득부;

상기 제1 구매기록 데이터를 전처리하고, 전처리된 제1 구매기록 데이터로부터 예측 모델의 학습 데이터를 생성하는 전처리부;

상기 학습 데이터를 이용하여 적대적 생성 신경망(Generative Adversarial Network, GAN)을 학습하여 상기 예측 모델을 생성하는 예측부를 포함하고,

상기 예측부는 제2 구매기록 데이터에 대한 예측을 수행하여, 의약품 재구매 시점 및 수량을 도출하는,

의약품 재구매 시점 및 수량 예측 장치.

발명의 설명

기술 분야

[0001] 본 발명은 인공지능 기반의 의약품의 재구매 시점 및 수량을 예측하는 장치 및 방법에 관한 것으로서, 보다 상세하게는 과거의 의약품 구매 수량과 구매 시점의 정보들을 바탕으로 학습된 적대적 생성 신경망과 선형 예측 모델의 예측 정보를 앙상블하여 사용자의 향후 의약품 재구매 시점과 수량을 예측하는 장치 및 방법에 관한 것이다.

배경 기술

[0002] 일반적으로 구매자들이 행하는 소비행위는 구매자들의 특성을 반영하는 것으로 알려져 있다. 따라서 과거의 소비 패턴을 분석하는 것으로 구매자의 향후의 패턴을 예측하는 것이 일정 부분 가능하다.

[0003] 이에 따라, 근래에 들어 사용자들의 소비 패턴 데이터들을 분석하여 향후의 소비 패턴을 예측하는 인공지능 시스템들이 많이 차용되고 있다. 흔히, 인터넷 쇼핑몰이나 영화 추천 시스템이 그러한 소비 패턴 빅데이터를 분석하여 예측을 진행한다. 이와 같은 구매 패턴 데이터를 이용한 향후 재구매 시점을 예측하는 종래의 기술이 대한민국 공개특허 제2019-0012894호에 개시된 바가 있다.

[0004] 하지만, 현존하는 기술들은 학습을 위해 다량의 데이터를 필수적으로 요구하는 다층퍼셉트론 모델이나 순환신경망 모델과 같은 인공지능 기법들이나 협업 필터링이나 확률적 행렬 분해 모델과 같은 데이터 기반의 기법들을 이용하여 개발되었다.

[0005] 이에 따라, 기존의 기술들은 인터넷 쇼핑몰의 구매 기록데이터나 영화 평론 사이트 영화 평점 데이터와 같이 빅

데이터 취합이 용이한 분야들에 한해서 개발되고 있다는 한계점을 가지고 있다.

[0006] 따라서, 비교적 적은 데이터만을 취합할 수 있는 분야들에 한해서도 소비 패턴 데이터를 좀 더 상세하게 고려하는 것으로 향후의 재구매 시점과 수량을 예측할 수 있는 시스템의 개발이 요구된다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 대한민국 공개특허 제2019-0012894호 (2019.02.11. 공개)

발명의 내용

해결하려는 과제

[0008] 본 발명이 이루고자 하는 기술적인 과제는 인공지능 생성 모델 기반의 의약품 재구매 시점 및 수량을 예측하는 장치 및 방법을 제공하는 것이다.

과제의 해결 수단

[0009] 본 발명의 일 실시예에 의한 의약품 재구매 시점 및 수량 예측 방법은 컴퓨팅 장치에 의해 수행되는 의약품 재구매 시점 및 수량 예측 방법으로, 제1 구매기록 데이터를 획득하는 단계, 상기 제1 구매기록 데이터를 전처리하는 단계, 전처리된 제1 구매기록 데이터로부터 예측 모델의 학습 데이터를 생성하는 단계, 상기 학습 데이터를 이용하여 적대적 생성 신경망(Generative Adversarial Network, GAN)을 학습하여 상기 예측 모델을 생성하는 단계, 및 제2 구매기록 데이터에 대한 예측을 수행하여, 의약품 재구매 시점 및 수량을 도출하는 단계를 포함한다.

[0010] 또한, 본 발명의 다른 실시예에 의한 의약품 재구매 시점 및 수량 예측 방법은 컴퓨팅 장치에 의해 수행되는 의약품 재구매 시점 및 수량 예측 방법으로, 구매기록 데이터를 획득하는 단계, 상기 구매기록 데이터를 전처리하는 단계, 전처리된 구매기록 데이터를 예측 모델에 입력하여 제1 예측 값 및 제1 예측 값에 대한 신뢰도(k)를 도출하는 단계, 상기 구매기록 데이터의 평균(m)을 산출하여 제2 예측 값을 도출하는 단계, 및 상기 제1 예측 값과 상기 제2 예측 값으로부터 최종 결과를 산출하는 단계를 포함하고, 상기 최종 결과(T_{Final})는 수학식을 이용하여 산출되고, 상기 수학식은 $T_{Final} = G(T_P, z) * k + (1 - k) * m$ 이고, 상기 $G(T_P, z)$ 는 상기 전처리된 구매기록 데이터(T_P)와 노이즈(z)를 입력으로 하는 적대적 생성 신경망 기반의 상기 예측 모델에 포함된 생성기(Generator)의 출력인 것을 특징으로 한다.

[0011] 또한, 본 발명의 일 실시예에 의한 의약품 재구매 시점 및 수량 예측 장치는 제1 구매기록 데이터를 획득하는 데이터 획득부, 상기 제1 구매기록 데이터를 전처리하고, 전처리된 제1 구매기록 데이터로부터 예측 모델의 학습 데이터를 생성하는 전처리부, 상기 학습 데이터를 이용하여 적대적 생성 신경망(Generative Adversarial Network, GAN)을 학습하여 상기 예측 모델을 생성하는 예측부를 포함하고, 상기 예측부는 제2 구매기록 데이터에 대한 예측을 수행하여, 의약품 재구매 시점 및 수량을 도출하는 것을 특징으로 한다.

발명의 효과

[0012] 본 발명의 실시예에 따른 의약품 재구매 시점 및 수량 예측 장치 및 방법에 의할 경우, 기존에 존재하는 인공지능기반의 추천 시스템의 한계인 부족한 데이터에서의 학습 불능을 새로운 인공지능 기법인 적대적 생성 신경망 기술을 시스템에 도입함으로써 해결할 수 있다.

[0013] 또한, 의약품의 구매 일자와 구매 수량으로 된 간단한 구매 이력 정보만을 활용하여 의약품의 재구매 시기를 예측하기 때문에 반복구매가 빈번한 어떤 도메인에서도 보편적으로 활용 가능하여 범용성이 높다는 장점이 있다.

[0014] 또한, 의약품의 구매 일자와 구매 수량으로 된 간단한 구매 이력 정보만을 활용하여 의약품의 재구매 시기를 예측하기 때문에 현재 우려를 낳고 있는 개인정보 보호의 문제로부터 상당히 자유로운 편이라는 장점이 있다.

도면의 간단한 설명

- [0015] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.
- 도 1은 본 발명의 일 실시예에 따른 예측 장치의 기능 블록도이다.
- 도 2는 도 1에 도시된 전처리부의 전처리 동작을 설명하기 위한 도면이다.
- 도 3은 적대적 생성 신경망 기반의 제1 예측 모델의 구조를 도시한다.
- 도 4는 테스트셋에서 본 발명의 성능을 보여주는 산점도 플롯이다.
- 도 5는 테스트셋에서 본 발명과 다른 기계학습 기반 모델들의 성능 비교 결과를 도시한다.
- 도 6은 도 1에 도시된 예측 장치에 의해 수행되는 의약품 재구매 시점 및 수량 예측 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0016] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태들로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.
- [0017] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.
- [0018] 제1 또는 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1 구성 요소는 제2 구성 요소로 명명될 수 있고 유사하게 제2 구성 요소는 제1 구성 요소로도 명명될 수 있다.
- [0019] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.
- [0020] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0021] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 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.
- [0022] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시예들을 상세히 설명한다. 그러나, 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0023] 도 1은 본 발명의 일 실시예에 따른 예측 장치의 기능 블록도이다.
- [0024] 도 1을 참조하면, 의약품 재구매 시점 및 수량 예측 장치 등으로 명명될 수도 있는 예측 장치(100)는 데이터 획득부(110), 전처리부(120), 예측부(130), 및 저장부(140)를 포함할 수 있다. 또한, 예측 장치(100)는 적어도 프로세서(processor) 및/또는 메모리(memory)를 포함하는 컴퓨팅 장치로 구현될 수 있다. 컴퓨팅 장치는

PC(Personal Computer), 랩탑 컴퓨터, 태블릿 PC, 서버(server), 노트북 등을 포함한다.

- [0025] 데이터 획득부(110)는 소정의 유무선 통신망을 통하여 데이터를 수신할 수 있다. 실시예에 따라, 데이터 획득부(110)는 USB 메모리 장치와 같은 저장 장치로부터 소정의 입출력 인터페이스를 통해 데이터를 수신할 수도 있으며, 데이터는 저장부(140)에 미리 저장되어 있을 수도 있다.
- [0026] 데이터는 인공지능 모델의 학습을 위한 제1 데이터 및/또는 인공지능 모델의 학습이 완료된 후에 예측을 위한 입력을 의미하는 제2 데이터를 포함할 수 있다. 또한, 데이터는 의약품에 대한 구매 정보로써, 각각이 고객사(구매자, 사용자 등으로 명명될 수도 있음), 주문 제품(제품명, 제품을 식별할 수 있는 정보 등을 의미할 수 있음), 주문 날짜, 및 주문 수량 중 적어도 하나를 포함하는 복수의 레코드들을 의미할 수 있다.
- [0027] 데이터 획득부(110)에 의해 수신된 데이터는 저장부(140)에 저장될 수 있다.
- [0028] 전처리부(120)는 데이터 획득부(110)에 의해 수신된 데이터 또는 저장부(140)에 저장되어 있는 데이터에 대한 전처리 동작을 수행할 수 있다. 전처리 동작을 설명하기 위한 도면인 도 2를 참조하면, 구매기록 데이터에서 구매 패턴을 추출하기 위해서는 데이터들을 품목별로 정리하는 것이 요구된다. 구매기록 데이터를 품목별로 정리하게 되면 품목별 구매기록의 길이가 모두 다르게 나타나게 된다. 이렇게 다양한 길이를 갖는 데이터들을 정리하기 위하여 본 발명에서는 제품명(또는 고객사와 제품명)을 키(key)로 갖고 구매기록 데이터들을 밸류(value)로 갖는 사전(dictionary) 구조로 데이터를 전처리하였다.
- [0029] 또한, 전처리부(120)는 전처리된 데이터로부터 예측 모델의 입력 데이터를 생성할 수 있다. 대부분의 기계학습 기반의 예측모델들은 입력 데이터들의 구조가 모두 일치해야 한다. 하지만, 구매기록 데이터들의 경우에는 구매 횟수가 모두 다르므로 데이터의 구조가 모두 다르다. 따라서, 전처리부(120)는 다양한 구조를 갖는 구매기록 데이터들을 가공하여 동일한 길이(1)을 갖는 데이터들로 절단할 수 있다. 즉, 미리 정해진 구매 횟수를 포함하도록 데이터를 분할하여 동일한 길이를 갖는 입력 데이터를 생성할 수 있다.
- [0030] 또한, 본 발명은 한정된 수량의 데이터들만을 수집할 수 있는 분야에서도 높은 성능을 갖는 재구매 시점 및 수량 예측 모델의 개발을 목표로 하였다. 현존하는 재구매 시점 예측 모델은 이전의 구매 기록을 바탕으로 향후의 재구매 시점을 예측하는 구조로 되어 있다. 따라서, 이러한 예측 모델을 학습시키기 위하여 구매기록들을 절단하여 이전 시점을 바탕으로 향후 구매 시점을 예측하는 방식으로 학습 데이터를 생성하게 된다. 그러나, 이러한 재구매 시점 예측에서 한정된 수량의 데이터들만을 수집할 수 있는 경우에는 현존하는 구조의 학습 데이터 생성은 이상치에 매우 취약하여 학습 자체가 제대로 되지 않는다는 단점을 갖게 된다. 따라서, 가상의 구매주기와 수량을 구매기록 데이터로부터 계산해낸 후 구매기록 데이터들을 일정한 길이 1로 절단하여 계산된 구매주기를 예측하는 방식으로 학습 데이터를 구축하였다. 여기서 가상의 구매주기와 수량은 후술할 예측 모델(생성기)의 출력을 의미할 수 있다.
- [0031] 예시적으로, n 개의 구매기록을 갖는 제품명 i 의 구매 날짜(구매주기를 의미할 수도 있음)를 $T_{i_1}, T_{i_2}, \dots, T_{i_n}$ 라고 하고, 구매 수량을 $N_{i_1}, N_{i_2}, \dots, N_{i_n}$ 이라 할 때, 가상의 구매주기 T_i 와 수량 N_i 는 수학적 식 1과 같은 방법으로 계산된다.
- [0032] [수학적 식 1]
- $$T_i = \frac{\min(T_{i_1}, T_{i_2}, \dots, T_{i_n}) + \sum_k \frac{T_{i_k}}{n}}{2}, N_i = \frac{\min(N_{i_1}, N_{i_2}, \dots, N_{i_n}) + \sum_k \frac{N_{i_k}}{n}}{2}$$
- [0033]
- [0034] 이와 같이, 가상의 구매주기와 구매 수량을 데이터들로부터 계산하여 이를 예측하는 방식으로 학습 데이터를 구축하는 것으로 n 개의 구매기록을 갖는 제품명 i 에 대하여 총 $(n-1)$ 개의 학습 데이터를 생성할 수 있다. 이러한 학습 데이터 구축 기법은 한정된 수량 데이터에서 이상치에 대한 내성을 높이는 효과를 갖는다. 또한, 최소값과 평균값의 평균을 적용하는 방식으로 이상치가 존재하지 않는 상황에도 학습 데이터의 계산된 구매주기와 실제 구매주기와 큰 차이를 보이지 않도록 보정하였다.
- [0035] 상술한 바와 같이, 전처리부(120)는 획득된 데이터에 대한 전처리 동작을 수행한 후 예측 모델의 입력 데이터(학습 데이터 및/또는 예측을 위한 타겟 데이터)를 생성할 수 있다.
- [0036] 예측부(130)는 학습 데이터를 이용하여 소정의 신경망 모델을 학습함으로써 제1 예측 모델을 생성할 수 있다. 소정의 신경망 모델은 적대적 생성 신경망(Generative Adversarial Networks)을 의미할 수 있다.

- [0037] 적대적 생성 신경망 기반의 제1 예측 모델의 구조는 과거 1회의 구매기록과 노이즈 벡터를 입력으로 받아 구매 주기(및/또는 구매수량)를 예측하는 생성기(generator, G)와 구매기록과 구매주기(및/또는 구매수량)를 입력으로 받고 이러한 구매기록과 구매주기가 적절한 구매주기인지(및/또는 이러한 구매기록과 구매수량이 적절한 구매수량인지) 판단하여 데이터의 진위 여부를 판단하는 판별기(discriminator, D)로 구성되어 있다. 적대적 생성 신경망 기반의 제1 예측 모델의 구조는 도 3에 도시되어 있다.
- [0038] 과거의 구매 기록을 T_P , 계산된 구매주기를 T , 노이즈 벡터를 z , 생성기를 G , 판별기를 D 라고 할 때, 판별기의 손실 함수(loss function)는 수학적식 2와 같고 생성기의 손실 함수는 수학적식 3과 같다.
- [0039] [수학적식 2]
- [0040]
$$L_D = -E[\min(0, -1 + D(T_P, T))] - E[\min(0, -1 + D(T_P, G(T_P, z)))]$$
- [0041] [수학적식 3]
- [0042]
$$L_G = -E[D(T_P, G(T_P, z))]$$
- [0043] 이러한 적대적 생성 모델의 손실 함수로부터 제1 예측 모델은 생성기를 통해 구매 기록으로부터 품목의 구매 주기(및/또는 구매수량)를 예측할 수 있다.
- [0044] 또한, 본 발명에서는 판별기를 통해 생성기의 성능을 끌어올리는 것 외에도 생성기가 예측한 품목의 구매 주기(및/또는 구매수량)의 신뢰도를 계산한다. 적대적 생성 모델의 판별기는 수학적식 2의 손실 함수를 최소화하는 방향으로 학습하게 되는데 이는 결국 과거의 구매 기록을 보고 계산된 구매주기의 신뢰도를 계산하는 방향으로 학습하는 것과 유사하다.
- [0045] 따라서, 본 발명의 제1 예측 모델의 재구매 시점 및 수량 예측은 적대적 생성 모델의 생성기에서 이루어지게 되며 이 값의 신뢰도 계산은 적대적 생성 모델의 판별기를 통해 계산된다.
- [0046] 또한, 예측부(130)는 학습이 완료된 제1 예측 모델과 제2 예측 모델을 이용하여 입력 데이터(제2 데이터를 의미할 수 있음)에 대한 예측값(재구매 시점 및/또는 재구매 수량)을 출력할 수 있다. 즉, 본 발명의 일반성 유지를 위하여 전체 입력 데이터의 평균을 고려하여 예측을 수행한다. 이러한 시도는 선형 기반 예측 모델(제2 예측 모델)과 신뢰도 기반 앙상블 기법으로 구현된다. 제2 예측 모델은 입력 데이터에 대한 평균(구매 주기에 대한 평균 및/또는 구매 수량에 대한 평균)을 도출하는 모델을 의미할 수 있다.
- [0047] 구체적으로, 예측부(130)는 입력 데이터를 제1 예측 모델에 입력하여 제1 예측값과 제1 예측값에 대한 신뢰도를 도출하고, 입력 데이터를 제2 예측 모델에 입력하여 제2 예측값을 도출한다. 또한, 예측부(130)는 신뢰도를 기반으로 제1 예측값과 제2 예측값을 앙상블하여 최종 예측 결과를 도출할 수 있다. 즉, 예측부(130)는 제1 예측 모델의 생성기에서 예측된 구매주기(및/또는 수량), 제1 예측 모델의 판별기에서 계산된 생성기의 결과에 대한 신뢰도, 그리고 선형 기반 예측 모델(제2 예측 모델)에서 계산한 입력 데이터의 구매주기(및/또는 구매수량)의 평균을 가중평균하여 최종 예측 결과를 도출한다. 선형 기반 예측 모델(제2 예측 모델)에서 계산한 입력 데이터의 평균을 m , 판별기에서 계산된 신뢰도를 k 라고 할 때, 최종 예측 결과는 수학적식 4를 통해 계산된다.
- [0048] [수학적식 4]
- [0049]
$$T_{Final} = G(T_P, z) * k + (1 - k) * m$$
- [0050] 수학적식 4에서 k 는 0 보다 크거나 같고 1 보다 작거나 같은 실수이다.
- [0051] 저장부(140)에는 예측 장치의 동작을 위한 OS(Operating System), 프로그램, 소스코드 등이 저장되어 있을 수 있다. 또한, 저장부(130)에는 데이터, 전처리된 데이터, 예측 모델, 예측의 결과 등이 저장될 수 있다.
- [0052] 도 1에 도시된 예측 장치(100)의 구성들 각각은 기능 및 논리적으로 분리될 수 있음으로 나타내는 것이며, 반드시 각각의 구성이 별도의 물리적 장치로 구분되거나 별도의 코드로 작성됨을 의미하는 것이 아님을 본 발명의 기술분야의 평균적 전문가가 용이하게 추론할 수 있을 것이다.
- [0053] 또한, 본 명세서에서 "~부"라 함은, 본 발명의 기술적 사상을 수행하기 위한 하드웨어 및 상기 하드웨어를 구동하기 위한 소프트웨어의 기능적, 구조적 결합을 의미할 수 있다. 예컨대, 상기 "~부"는 소정의 코드와 상기 소정의 코드가 수행되기 위한 하드웨어 리소스의 논리적인 단위를 의미할 수 있으며, 반드시 물리적으로 연결된

코드를 의미하거나, 한 종류의 하드웨어를 의미하는 것이 아니다.

- [0054] 본 발명이 제안하는 적대적 생성 신경망 기반의 의약품 재구매 시점 및 수량 예측 모델을 한정된 의약품 구매기록 데이터로 학습하여 학습하지 않은 최근의 의약품 구매 기록 데이터를 예측하였을 때, 우수한 성능을 보임을 확인할 수 있다(도 4 참조). 대부분의 경우 계산된 구매주기와 모델을 통해 예측한 구매주기에 큰 차이가 없음을 확인할 수 있다.
- [0055] 또한, 본 발명의 차별성을 도면 5에서 확인할 수 있다. Linear Regression, Lasso, Ridge, Random Forest 등의 기계학습 기반의 예측 모델들과 동일한 데이터로 학습하여 동일한 테스트셋에서 예측 결과와 계산된 구매주기의 차이를 손실값으로 갖는 박스 플롯을 그렸을 때 본 발명(GAN)의 손실값이 다른 기계학습 기반의 모델에 비해 확연히 작음을 확인할 수 있다.
- [0056] 도 6은 도 1에 도시된 예측 장치에 의해 수행되는 의약품 재구매 시점 및 수량 예측 방법을 설명하기 위한 흐름도이다. 예측 방법은 컴퓨팅 장치에 의해 수행되기 때문에, 예측 방법에 포함되는 단계들 중 적어도 일부는 컴퓨팅 장치의 프로세서의 동작으로 이해될 수도 있다. 이하, 예측 방법을 설명함에 있어, 앞선 기재와 중복되는 내용에 관하여는 그 기재를 생략하기로 한다.
- [0057] 도 6을 참조하면, 우선 데이터가 획득된다(S110). 데이터는 유무선 통신망을 통해 수신되거나, 소정의 저장 장치로부터 수신될 수 있다. 또한, 데이터는 미리 수신되어 컴퓨팅 장치 내에 저장되어 있을 수도 있고, 제1 데이터 및/또는 제2 데이터를 포함한다.
- [0058] 수신된 데이터에 대한 전처리 동작이 수행된다(S120). 전처리 동작은 로테이터를 제품별 구매 기록으로 가공하는 것을 의미할 수 있다.
- [0059] 다음으로, 전처리된 데이터로부터 입력 데이터가 생성될 수 있다(S130). 입력 데이터는 제1 예측 모델의 학습을 위한 학습 데이터로, 미리 정해진 크기로 분할된 구매기록 데이터(입력) 및/또는 계산된 구매주기(및/또는 구매수량)(출력)를 포함할 수 있다.
- [0060] 학습 데이터를 이용한 적대적 생성 신경망의 학습을 통해 제1 예측 모델이 생성된다(S140). 실제 데이터를 대상으로 하는 예측을 수행하는 경우라면, S140 단계는 생략될 수도 있다.
- [0061] 마지막으로, 구매 주기 및/또는 구매 수량에 대한 예측이 수행된다(S150). 예측을 위한 입력 데이터는 S110 단계에서 수신된 데이터 또는 S150 단계 이전에 수신되는 데이터로부터 생성될 수 있다.
- [0062] 이상에서 설명된 장치는 하드웨어 구성 요소, 소프트웨어 구성 요소, 및/또는 하드웨어 구성 요소 및 소프트웨어 구성 요소의 집합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성 요소는, 예를 들어, 프로세서, 컨트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크로컴퓨터, FPA(Field Programmable 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(Operation System, 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 기술 분야에서 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.
- [0063] 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, 구성 요소(Component), 물리적 장치, 가상 장치(Virtual Equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(Signal Wave)에 영구적으로, 또는 일시적으로 구체화(Embody)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.
- [0064] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판

독 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), 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. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0065]

본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성 요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성 요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

[0066]

100 : 예측 장치

110 : 데이터 획득부

120 : 전처리부

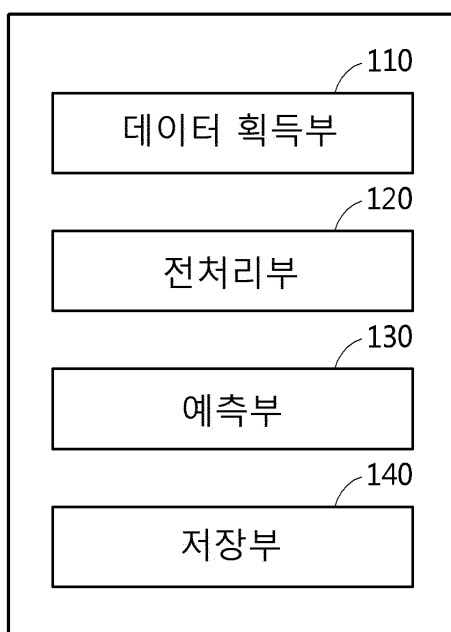
130 : 예측부

140 : 저장부

도면

도면1

100



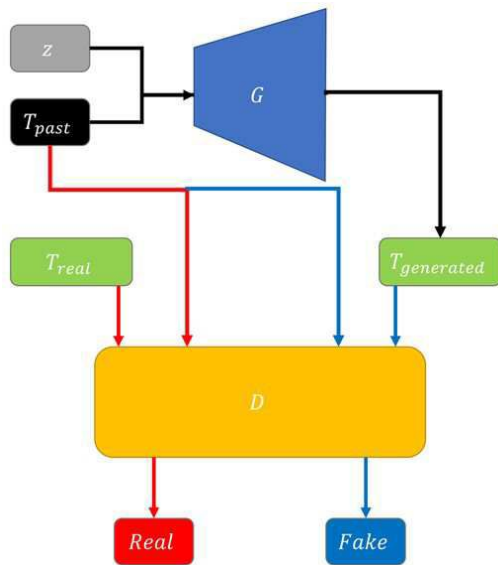
도면2

고객사	주문 제품	주문 날짜	주문 수량
OO기업	A제품	20xx년 x월 x일	n개
XX기업	B제품	20xx년 x월 x일	m개
⋮	⋮	⋮	⋮
△△기업	C제품	20xx년 x월 x일	k개

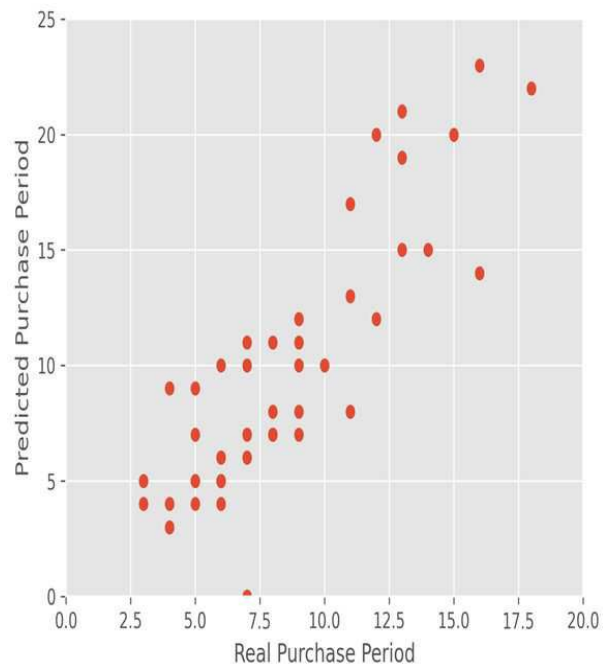
Dictionary

Key	Value
(OO기업, A제품)	[(날짜, 수량), (날짜, 수량), ..., (날짜, 수량)]
(OO기업, B제품)	[(날짜, 수량), (날짜, 수량), ..., (날짜, 수량)]
(XX기업, C제품)	[(날짜, 수량), (날짜, 수량), ..., (날짜, 수량)]
⋮	⋮
(XX기업, Z제품)	[(날짜, 수량), (날짜, 수량), ..., (날짜, 수량)]

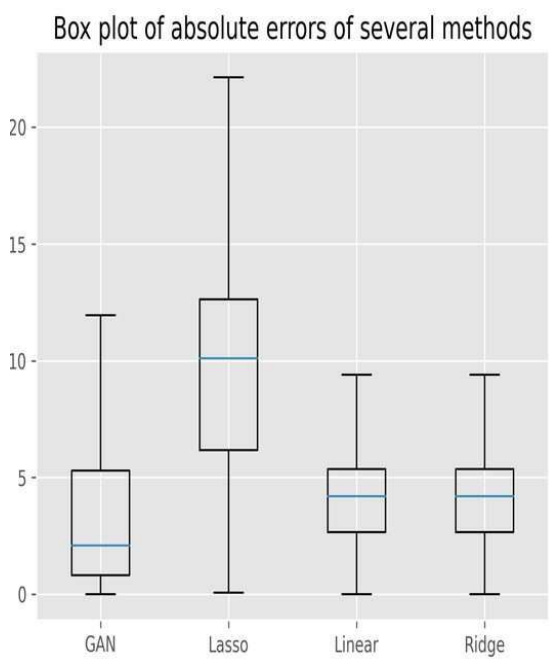
도면3



도면4



도면5



도면6

