

(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0021189

(43) 공개일자 2022년02월22일

(51) 국제특허분류(Int. Cl.)

G16H 50/50 (2018.01) G06N 20/00 (2019.01)

G06Q 10/04 (2012.01) G16H 10/60 (2018.01)

G16H 40/20 (2018.01)

(52) CPC특허분류

G16H 50/50 (2018.01)

G06N 20/00 (2021.08)

(21) 출원번호 10-2020-0101710

(22) 출원일자 2020년08월13일

심사청구일자 2020년08월13일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

정중문

서울특별시 용산구 이촌로 181, 104동 101호(이촌동, 한강대우아파트)

정현수

서울특별시 강남구 언주로30길 13, A동 3405호(도곡동, 대림아크로빌, 대림아크로텔)

(뒷면에 계속)

(74) 대리인

특허법인우인

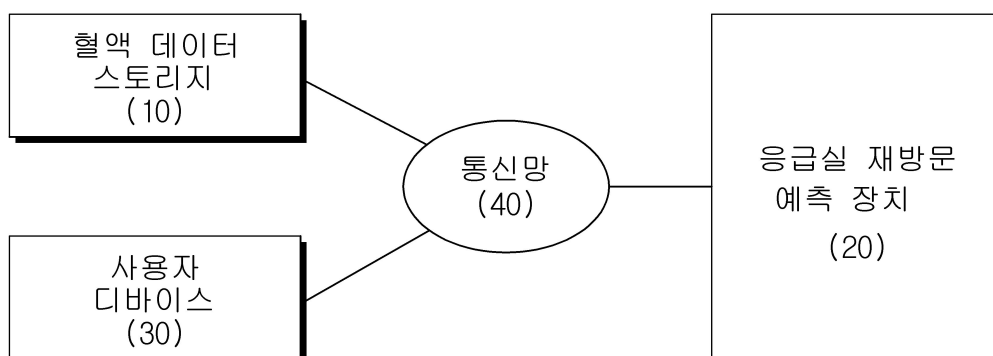
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 응급실 내 재내원 환자 예측을 위한 혈액 기반 모바일 헬스 딥러닝 예측 시스템 및 응급실 재방문 예측 장치

## (57) 요약

본 실시예들은 부스팅 기반의 예측 모델을 혈액 데이터에 접목시켜 혈액과 같이 간단한 데이터로도 응급실 재방문 확률을 예측 가능한 혈액 기반 모바일 헬스 딥러닝 예측 시스템 및 응급실 재방문 예측 장치를 제공한다.

대표도 - 도1



(52) CPC특허분류

*G06N 7/005* (2013.01)

*G06Q 10/04* (2013.01)

*G16H 10/60* (2021.08)

*G16H 40/20* (2021.08)

(72) 발명자

**신영환**

서울특별시 서대문구 연희로10길 24-10, 202호(연희동)

**김상도**

서울특별시 양천구 목동중앙북로7나길 54, 304호(목동, 석경트라움)

**한상길**

경기도 수원시 영통구 광고호수공원로 277, 105동 1906호(원천동, 광고중흥에스클래스)

이 발명을 지원한 국가연구개발사업

과제고유번호 1711116181

과제번호 IITP-2020-2018-0-01799

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원

연구사업명 대학ICT연구센터지원사업

연구과제명 블록체인 비즈니스 서비스 기술 개발 및 인력양성

기 여 율 1/1

과제수행기관명 중앙대학교 산학협력단

연구기간 2018.07.01 ~ 2021.12.31

## 명세서

### 청구범위

#### 청구항 1

응급실 재방문 예측 장치에 있어서,

혈액 데이터를 수집하고 처리하여 부스팅(Boosting) 기반의 재방문 예측 모델을 통해 환자의 재방문 확률을 예측하는 프로세서;

상기 환자의 재방문 확률을 저장하는 저장 매체; 및

상기 환자의 재방문 확률을 사용자 디바이스로 송신하는 통신 인터페이스를 포함하는 응급실 재방문 예측 장치.

#### 청구항 2

제1항에 있어서,

상기 혈액 데이터는 (i) 전혈에서 채취된 복수 유형의 제1 혈액 데이터, (ii) 혈청에서 채취된 복수 유형의 제2 혈액 데이터, (iii) 혈장에서 채취된 복수 유형의 제3 혈액 데이터, (iv) 동맥혈에서 채취된 복수 유형의 제4 혈액 데이터를 포함하는 것을 특징으로 하는 응급실 재방문 예측 장치.

#### 청구항 3

제1항에 있어서,

상기 혈액 데이터는 암호화되고 블록체인 네트워크에 보관되는 것을 특징으로 하는 응급실 재방문 예측 장치.

#### 청구항 4

제1항에 있어서,

상기 부스팅 기반의 재방문 예측 모델은 트리의 균형을 맞추지 않고 최대 손실 값(Max delta loss)을 가지는 리프 노드를 수직적으로 확장시키는 리프 중심(Leaf-wise) 트리 분할 방식이 적용된 그래디언트 부스팅(Gradient Boosting) 예측 모델인 것을 특징으로 하는 응급실 재방문 예측 장치.

#### 청구항 5

제1항에 있어서,

상기 환자의 실제 데이터와 예측 데이터를 이용하여 지도학습을 통해 상기 부스팅 기반의 재방문 예측 모델에 설정되는 하이퍼 파라미터 중에서 (i) 트리마다 다른 특성을 랜덤하게 추출하는 특성 비율, (ii) 최대 깊이, (iii) 리프 개수, (iv) 학습율을 도출하는 것을 특징으로 하는 응급실 재방문 예측 장치.

#### 청구항 6

제5항에 있어서,

상기 환자의 실제 데이터를 이용하여 상기 부스팅 기반의 재방문 예측 모델을 학습하여 상기 재방문 예측 모델의 파라미터를 갱신하는 것을 특징으로 하는 응급실 재방문 예측 장치.

#### 청구항 7

응급실 재방문 예측을 위한 사용자 디바이스에 있어서,

응급실 재방문 예측 장치로 쿼리를 송신하고 환자의 재방문 확률을 수신하는 통신 인터페이스;

상기 환자의 재방문 확률을 수집하고 처리하는 프로세서; 및

상기 환자의 재방문 확률을 출력하는 출력 인터페이스를 포함하는 사용자 디바이스.

## 청구항 8

제7항에 있어서,

상기 환자의 재방문 확률은 상기 응급실 재방문 예측 장치에 의해 예측되며,

상기 응급실 재방문 예측 장치는 혈액 데이터를 수집하고 처리하여 부스팅(Boosting) 기반의 재방문 예측 모델을 통해 환자의 재방문 확률을 예측하는 것을 특징으로 하는 사용자 디바이스.

## 청구항 9

제8항에 있어서,

상기 혈액 데이터는 (i) 전혈에서 채취된 복수 유형의 제1 혈액 데이터, (ii) 혈청에서 채취된 복수 유형의 제2 혈액 데이터, (iii) 혈장에서 채취된 복수 유형의 제3 혈액 데이터, (iv) 동맥혈에서 채취된 복수 유형의 제4 혈액 데이터를 포함하는 것을 특징으로 하는 사용자 디바이스.

## 청구항 10

제8항에 있어서,

상기 혈액 데이터는 암호화되고 블록체인 네트워크에 보관되는 것을 특징으로 하는 사용자 디바이스.

## 청구항 11

제8항에 있어서,

상기 부스팅 기반의 재방문 예측 모델은 트리의 균형을 맞추지 않고 최대 손실 값(Max delta loss)을 가지는 리프 노드를 수직적으로 확장시키는 리프 중심(Leaf-wise) 트리 분할 방식이 적용된 그래디언트 부스팅(Gradient Boosting) 예측 모델인 것을 특징으로 하는 사용자 디바이스.

## 청구항 12

제8항에 있어서,

상기 환자의 실제 데이터와 예측 데이터를 이용하여 지도학습을 통해 상기 부스팅 기반의 재방문 예측 모델에 설정되는 하이퍼 파라미터 중에서 (i) 트리마다 다른 특성을 랜덤하게 추출하는 특성 비율, (ii) 최대 깊이, (iii) 리프 개수, (iv) 학습율을 도출하는 것을 특징으로 하는 사용자 디바이스.

## 청구항 13

제12항에 있어서,

상기 환자의 실제 데이터를 이용하여 상기 부스팅 기반의 재방문 예측 모델을 학습하여 상기 재방문 예측 모델의 파라미터를 갱신하는 것을 특징으로 하는 사용자 디바이스.

## 발명의 설명

### 기술 분야

[0001] 본 발명이 속하는 기술 분야는 응급실 내 재내원 환자 예측을 위한 혈액 기반 모바일 헬스 딥러닝 예측 시스템 및 응급실 재방문 예측 장치에 관한 것이다.

### 배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] 응급 의료 서비스는 인간의 생명과 직결된다. 응급 환자 수의 불규칙성으로 인하여 응급실 과밀화가 문제되고 있다. 응급실에서 퇴원한 환자가 재내원하는 것을 감소시켜서 응급실 과밀화 문제를 해결하려는 연구가 진행되고 있다.

[0004] 포괄적 의료 변수 및 의료 빅데이터에 인공지능 기술을 적용하는 연구가 진행되고 있으나, 의료 빅데이터를 확

보하지 못한 비도심 지역에 적용하는데 한계가 있고, 환자로부터 수백~수천 개의 변수를 수 년에 걸쳐 수집하는 것 역시 쉽지 않은 실정이다.

## 선행기술문헌

### 특허문헌

- [0005] (특허문헌 0001) 한국공개특허공보 제10-2020-0003407호 (2020.01.09)  
(특허문헌 0002) 한국공개특허공보 제10-2010-0103514호 (2010.09.27)

## 발명의 내용

### 해결하려는 과제

- [0006] 본 발명의 실시예들은 부스팅 기반의 예측 모델을 혈액 데이터에 접목시켜 혈액과 같이 간단한 데이터로도 응급실 재방문 확률을 예측하는 데 주된 목적이 있다.
- [0007] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

### 과제의 해결 수단

- [0008] 본 실시예의 일 측면에 의하면, 응급실 재방문 예측 장치에 있어서, 혈액 데이터를 수집하고 처리하여 부스팅(Boosting) 기반의 재방문 예측 모델을 통해 환자의 재방문 확률을 예측하는 프로세서, 상기 환자의 재방문 확률을 저장하는 저장 매체, 및 상기 환자의 재방문 확률을 사용자 디바이스로 송신하는 통신 인터페이스를 포함하는 응급실 재방문 예측 장치를 제공한다.
- [0009] 상기 혈액 데이터는 (i) 전혈에서 채취된 복수 유형의 제1 혈액 데이터, (ii) 혈청에서 채취된 복수 유형의 제2 혈액 데이터, (iii) 혈장에서 채취된 복수 유형의 제3 혈액 데이터, (iv) 동맥혈에서 채취된 복수 유형의 제4 혈액 데이터를 포함할 수 있다.
- [0010] 상기 혈액 데이터는 암호화되고 블록체인 네트워크에 보관될 수 있다.
- [0011] 상기 부스팅 기반의 재방문 예측 모델은 트리의 균형을 맞추지 않고 최대 손실 값(Max delta loss)을 가지는 리프 노드를 수직적으로 확장시키는 리프 중심(Leaf-wise) 트리 분할 방식이 적용된 그래디언트 부스팅(Gradient Boosting) 예측 모델로 구현될 수 있다.
- [0012] 상기 응급실 재방문 예측 장치는 상기 환자의 실제 데이터와 예측 데이터를 이용하여 지도학습을 통해 상기 부스팅 기반의 재방문 예측 모델에 설정되는 하이퍼 파라미터 중에서 (i) 트리마다 다른 특성을 랜덤하게 추출하는 특성 비율, (ii) 최대 깊이, (iii) 리프 개수, (iv) 학습율을 도출할 수 있다.
- [0013] 상기 응급실 재방문 예측 장치는 상기 환자의 실제 데이터를 이용하여 상기 부스팅 기반의 재방문 예측 모델을 학습하여 상기 재방문 예측 모델의 파라미터를 갱신할 수 있다.
- [0014] 본 실시예의 다른 측면에 의하면, 응급실 재방문 예측을 위한 사용자 디바이스에 있어서, 응급실 재방문 예측 장치로 쿼리를 송신하고 환자의 재방문 확률을 수신하는 통신 인터페이스, 상기 환자의 재방문 확률을 수집하고 처리하는 프로세서, 및 상기 환자의 재방문 확률을 출력하는 입출력 인터페이스를 포함하는 사용자 디바이스를 제공한다.
- [0015] 상기 환자의 재방문 확률은 상기 응급실 재방문 예측 장치에 의해 예측될 수 있다.

### 발명의 효과

- [0016] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 간단한 혈액 데이터를 기반으로 응급실 재내원 가능성이 높은 환자를 예측할 수 있는 기술로써 기존 기술에 요구된 필요 데이터양을 획기적으로 줄이고, 나아가 모바일 헬스와 접목시켜 장소에 제한받지 않고 사용할 수 있는 효과가 있다.
- [0017] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에

서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

### 도면의 간단한 설명

- [0018] 도 1은 본 발명의 실시예들에 따른 혈액 기반 모바일 헬스 딥러닝 예측 시스템을 예시한 블록도이다.
- 도 2는 본 발명의 실시예들을 실시하는 컴퓨팅 디바이스를 예시한 블록도이다.
- 도 3은 본 발명의 실시예들에 따른 혈액 기반 모바일 헬스 딥러닝 예측 시스템이 처리하는 데이터를 예시한 도면이다.
- 도 4는 본 발명의 실시예들에 따른 혈액 기반 모바일 헬스 딥러닝 예측 시스템의 동작을 예시한 흐름도이다.
- 도 5 및 도 6은 본 발명의 실시예들에 따라 수행된 시뮬레이션 결과를 도시한 것이다.

### 발명을 실시하기 위한 구체적인 내용

- [0019] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다.
- [0020] 도 1은 본 발명의 실시예들에 따른 혈액 기반 모바일 헬스 딥러닝 예측 시스템을 예시한 블록도이다.
- [0021] 도 1을 참조하면, 혈액 기반 모바일 헬스 딥러닝 예측 시스템은 혈액 데이터 스토리지(10), 응급실 재방문 예측 장치(20), 및 사용자 디바이스(30)를 포함한다.
- [0022] 혈액 데이터 스토리지(10)는 데이터베이스를 포함하며 데이터베이스에 혈액 데이터를 보관한다. 데이터베이스는 데이터의 검색, 추출, 삭제, 편집, 추가 등을 자유롭게 행할 수 있는 데이터의 저장형태를 의미한다. 데이터베이스는 오라클(Oracle), 인포믹스(Infomix), 사이베이스(Sybase), 관계형 데이터베이스 관리시스템(Relational Data Base Management System, RDBMS), 겔스톤(Gemston), 오리온(Orion), 객체 지향형 데이터베이스 관리시스템(Object Oriented Database Management System, OODBMS) 분산 데이터베이스, 클라우드 등을 이용하여 본 실시예의 목적에 맞게 구현될 수 있다.
- [0023] 응급실 재방문 예측 장치(20)는 혈액 데이터 스토리지(10)로부터 혈액 데이터를 수집하고 처리하여 환자의 응급실 재방문 확률을 산출한다.
- [0024] 사용자 디바이스(30)는 응급실 재방문 예측 장치(20)로 쿼리를 송신하고 응급실 재방문 예측 장치(20)는 쿼리에 응답하여 사용자 디바이스(30)로 쿼리에 대응하는 환자의 재방문 확률을 전송한다.
- [0025] 사용자 디바이스(30)는 컴퓨팅 디바이스로 구현될 수 있으며, 스마트폰(Smart Phone), 개인용 컴퓨터(Personal Computer, PC), 태블릿 PC(Tablet PC), 휴대용 개인정보 단말기(Personal Digital Assistant, PDA), 랩톱(Laptop) 등이 있을 수 있으나 이에 한정되는 것은 아니다.
- [0026] 혈액 데이터 스토리지(10), 응급실 재방문 예측 장치(20), 및 사용자 디바이스(30)는 통신망(40)으로 연결되어 있다. 통신망(40)은 혈액 데이터 스토리지(10), 응급실 재방문 예측 장치(20), 및 사용자 디바이스(30) 간의 통신을 가능케 할 목적으로 연결해 놓은 통신설비의 집합을 의미한다. 통신망(40)은 노드(Node), 회선, 간선, 및 위성 등을 포함하고 이들은 서로 연결 및 접속되어 있다.
- [0027] 사용자 디바이스(30)는 유무선통신이 가능하다. 예컨대, 무선통신에는 근거리 무선통신, 원거리 무선통신, 이동통신, 및 무선랜 통신 등 다양한 통신 프로토콜이 사용될 수 있다. 무선통신 프로토콜을 예로 들면, 근거리 무선통신(Near Field Communication, NFC), 지그비(ZigBee), 블루투스(Bluetooth), 와이파이(Wi-Fi), 와이맥스(WiMAX), GSM(Global System For Mobile Communication), 3G(Third Generation) 이동통신, LTE(Long Term Evolution), 4G, 5G 등이 있으나, 이에 한정되는 것은 아니다.
- [0028] 응급실 재방문 예측 장치(20)는 블록체인 네트워크를 통해 데이터를 저장하고 데이터를 송수신할 수 있다.
- [0029] 블록체인 네트워크는 합의 알고리즘을 통해 혈액 데이터를 안전하게 보관한다. 합의 알고리즘은 작업증명(Proof-of-Work), 지분증명(Proof-of-Stake), DPoS(Delegated Proof-of-Work) 등의 블록체인 네트워크에서 사용될 수 있는 다양한 합의 알고리즘을 포괄하는 의미이다. 블록체인 네트워크는 혈액 데이터를 저장하는 역할 뿐 아니라 응급실 재방문 예측 장치(20)의 요청에 상응하여 응급실 재방문 예측 장치(20)에 데이터를 전달하는

역할도 수행한다.

- [0030] 도 2는 본 발명의 실시예들을 실시하는 컴퓨팅 디바이스를 예시한 블록도이다.
- [0031] 혈액 데이터 스토리지(10), 응급실 재방문 예측 장치(20), 및 사용자 디바이스(30)는 컴퓨팅 디바이스(110)로 구현될 수 있다. 컴퓨팅 디바이스(110)는 적어도 하나의 프로세서(120), 컴퓨터 판독 가능한 저장매체(130) 및 통신 버스(170)를 포함한다.
- [0032] 프로세서(120)는 컴퓨팅 디바이스(110)를 동작하도록 제어할 수 있다. 예컨대, 프로세서(120)는 컴퓨터 판독 가능한 저장 매체(130)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 컴퓨터 실행 가능 명령어는 프로세서(120)에 의해 실행되는 경우 컴퓨팅 디바이스(110)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0033] 컴퓨터 판독 가능한 저장 매체(130)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능한 저장 매체(130)에 저장된 프로그램(140)은 프로세서(120)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독한 가능 저장 매체(130)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 컴퓨팅 디바이스(110)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0034] 통신 버스(170)는 프로세서(120), 컴퓨터 판독 가능한 저장 매체(140)를 포함하여 컴퓨팅 디바이스(110)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0035] 컴퓨팅 디바이스(110)는 또한 하나 이상의 입출력 장치를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(150) 및 하나 이상의 통신 인터페이스(160)를 포함할 수 있다. 입출력 인터페이스는 터치 스크린에서 가상 키보드 또는 커서 등을 제공할 수 있다. 입출력 인터페이스는 디스플레이, 키보드, 마우스, 스피커 등에 연결될 수 있다. 입출력 인터페이스(150) 및 통신 인터페이스(160)는 통신 버스(170)에 연결된다. 입출력 장치(미도시)는 입출력 인터페이스(150)를 통해 컴퓨팅 디바이스(110)의 다른 컴포넌트들에 연결될 수 있다.
- [0036] 도 3은 본 발명의 실시예들에 따른 혈액 기반 모바일 헬스 딥러닝 예측 시스템이 처리하는 데이터를 예시한 도면이고, 도 4는 본 발명의 실시예들에 따른 혈액 기반 모바일 헬스 딥러닝 예측 시스템의 동작을 예시한 흐름도이다.
- [0037] 단계 S210에서 응급환자가 발생한 상황인지 판단한다. 단계 S220에서 응급환자가 발생한 상황이 아니면, 이전 환자의 혈액 데이터( $d_{k,k \leq n}$ )에 대한 재내원 결과가 도출되었는지 판단한다. 암호화를 통해 재내원에 따른 혈액 결과 데이터를 구축할 수 있다. 단계 S250에서 이전 환자의 혈액 데이터( $d_{k,k \leq n}$ )에 대한 재내원 결과 데이터를 블록체인 네트워크로 전송한다.
- [0038] 단계 S230에서 응급환자가 발생한 상황이면, 의료진은 환자로부터 채혈을 실시하여 전혈, 혈청, 혈장, 동맥혈 총 네 가지 혈액 데이터를 추출한다. 단계 S240에서 암호화를 통해 새로운 환자의 혈액 데이터( $d_n$ )를 구축한다. 단계 S250에서 새로운 환자의 혈액 데이터( $d_n$ )를 블록체인 네트워크로 전송한다.
- [0039] 단계 S260에서 블록체인 네트워크에 데이터가 성공적으로 추가되었는지 판단한다. 단계 S310에서 블록체인 네트워크에 데이터가 성공적으로 추가되면, 응급실 재방문 예측 장치로 혈액 데이터 또는 암호화된 혈액 데이터를 전송한다. 응급실 재방문 예측 장치는 암호화된 혈액 데이터를 복호화한 후 필요한 혈액 데이터를 추출하여 처리할 수 있다.
- [0040] 단계 S320에서 응급실 재방문 예측 장치는 담당 의료진으로부터 데이터 쿼리가 도착했는지 판단한다. 단계 S330에서 응급실 재방문 예측 장치는 쿼리를 수신하면, 부스팅 기반의 재방문 예측 모델을 통해 응급실 재방문 확률을 도출한다. 단계 S340에서 응급실 재방문 예측 장치는 쿼리를 수신하지 않는 동안에 환자의 혈액 데이터( $d_{k,k \leq n}$ )에 대한 재내원 결과 데이터를 기반으로 훈련을 통해 부스팅 기반의 재방문 예측 모델을 갱신한다.
- [0041] 혈액 데이터 스토리지에는 네 가지 혈액 유형에 대한 총 71개의 변수를 담은 혈액 데이터가 들어있다. 이때 환자의 혈액 데이터는 철저히 암호화된다.
- [0042] 혈액 데이터는 전혈에서 채취된 28개 유형의 제1 혈액 데이터를 포함한다. 제1 혈액 데이터의 특성은 WBC



COUNT, RBC COUNT, Hemoglobin, Hct, MCV, MCH, MCHC, RDW, PDW (fL), PDW (%), MPV, Delta neutrophil 1, Delta neutrophil 2, Thrombotic microangiopathy score, PLT Count, Neutrophil (%), Lymphocyte (%), Monocyte (%), Eosinophil (%), Basophil (%), LUC (%), Neutrophil (#), Lymphocyte (#), Monocyte (#), Eosinophil (#), Basophil (#), LUC (#), ESR (Infrared Ray, Spin)로 구분될 수 있다.

[0043] 혈액 데이터는 혈청에서 채취된 20개 유형의 제2 혈액 데이터를 포함한다. 제2 혈액 데이터의 특성은 CRP (C-Reactive Protein), Calcium, Inorganic P, Glucose, BUN, Creatinine, Uric Acid, Total Protein, Albumin, Alk. Phos, AST (GOT), ALT (GPT), T. Bilirubin, CK, CK-MB, Na, K, Cl, tCO<sub>2</sub>, NT-proBNP로 구분될 수 있다.

[0044] 혈액 데이터는 혈장에서 채취된 6개 유형의 제3 혈액 데이터를 포함한다. 제3 혈액 데이터의 특성은 Troponin-T, PT (Sec), PT (%), PT (INR), Aptt, D-Dimer로 구분될 수 있다.

[0045] 혈액 데이터는 동맥혈에서 채취된 17개 유형의 제4 혈액 데이터를 포함한다. 제4 혈액 데이터의 특성은 pHp, CO<sub>2</sub>, pO<sub>2</sub>, tHb, BE-ECF, BE-B, SBC, HCO<sub>3</sub><sup>-</sup>, TC0<sub>2</sub>, SO<sub>2</sub> (%), O<sub>2</sub> Content, A(alveolar concentration), AaDO<sub>2</sub>, a/A(arterial/alveolar), pO<sub>2</sub>/FIO<sub>2</sub>, Hct, Lactate로 구분될 수 있다.

[0046] 응급실 재방문 예측 장치가 부스팅 기반의 재방문 예측 모델(LightGBM)을 통해 환자의 재방문 확률을 예측하는 동작은 표 1과 같은 의사 코드로 표현될 수 있다.

표 1

### Algorithm 1. BER

1. **SCAN** blockchain network
2. **IF** new patient's blood sample ( $d_n$ ) is detected.
3. **THEN FETCH** blood sample ( $d_n$ )
4. **AND COMPUTE** the probability of emergency room return for new patients  $\hat{o}_n$
5. **IF** Outcome  $o_n$  for new blood sample  $d_n$  is detected
6. **THEN FIND** Proper values of hyperparameters, feature fraction  $f$ , maximum depth  $h$ , number of leaves  $l$ , and learning rate  $\eta$
7. **WHILE** leaving other parameters default value
8. **TRAIN** LightGBM model with new blood sample
9. **END**

[0047]

[0048] BER(Blood sample-based ED Return) 알고리즘이 작동되는 응급실 재방문 예측 장치는 블록체인 네트워크를 상시 감시한다.

[0049]  $d_n$ 은 n번째 환자의 혈액 데이터이고,  $o_n$ 은 n번째 환자의 실제 재내원 여부이고,  $\hat{o}_n$ 은 n번째 환자의 재내원 예측이고,  $f$ 는 LightGBM 내 특성 비율이고,  $h$ 는 LightGBM 내 최대 깊이이고,  $l$ 은 LightGBM의 리프 개수이고,  $\eta$ 는 LightGBM의 학습율을 의미한다.

[0050] 새로운 환자의 데이터가 감지되면, 혈액 데이터를 가져오고 LightGBM을 사용하여 응급실 재내원 확률( $\hat{o}_n$ )을 예측한다. 향후 블록체인 네트워크에서 환자의 재내원 실제 여부가 밝혀지면, 새로운 데이터를 가지고 지도학습을



실시한다. 먼저 하이퍼 파라미터  $f$ ,  $h$ ,  $l$ ,  $\eta$ 을 찾아낸다. 다른 파라미터는 기본으로 설정한다. 그 뒤 새로운 데이터( $d_n$ )이 학습된 더 진보된 LightGBM 모델을 만든다.

[0051] 응급실 재방문 예측 장치가 혈액 데이터 특성을 처리하는 동작은 표 2과 같은 의사 코드로 표현될 수 있다.

**표 2**

**Algorithm 2. Blood sample Feature Engineering**

1. **SET** name of 71 features (string) as a number from 1 to 71(integer) regardless of their order
2. **SET** N/A (not available) values as 0 while neither scale up or down values
3. **CONVERT** file format into CSV (Comma-separated values) file format.
4. **END.**

[0052]

[0053] 응급실 재방문 예측 장치는 혈액 데이터 특성 중에서 손실 데이터를 사용불가항목으로 구분하고 체크하여 학습을 진행한다. 혈액 특성을 넘버링하고, CSV파일로 바꾸고 훈련시킨다.

[0054] LightGBM(Light Gradient Boosting Machine)에서 부스팅(Boosting)은 단순하고 약한 학습기(Weak Learner)를 결합해서 보다 정확하고 강력한 학습기(Strong Learner)를 만드는 방식을 의미한다.

[0055] 학습이란 손실함수를 최소화하는 파라미터를 찾는 것으로 손실함수(Loss Function)는 예측 모델의 오류를 정량화한다. 최적의 파라미터를 찾는 방법에 해당하는 경사 하강(Gradient descent)은 손실함수를 파라미터로 미분해서 기울기를 구하고, 값이 작아지는 방향으로 파라미터를 움직인다 보면 손실함수가 최소화되는 지점에 도달한다.

[0056] LightGBM의 리프 중심 트리 분할 방식은 트리의 균형을 맞추지 않고, 최대 손실 값(max delta loss)을 가지는 리프 노드를 지속적으로 분할하면서 트리의 깊이가 깊어지고 비대칭적인 규칙 트리가 생성된다. 최대 손실값을 가지는 리프 노드를 지속적으로 분할해 생성된 규칙 트리는 학습을 반복할수록 균형 트리 분할 방식보다 예측 오류 손실을 최소화할 수 있다,

[0057] LightGBM는 rf(random forest), gbdt(gradient boosted decision trees), dart(dropout regression trees), goss(gradient-based one-side sampling)을 적용하여 파라미터를 최적화할 수 있다.

[0058] 도 5 및 도 6은 본 발명의 실시예들에 따라 수행된 시뮬레이션 결과를 도시한 것이다.

[0059] 도 5에 도시된 BER 기법을 활용한 AUC-ROC(Area Under the Curve of the Receiver Operating Characteristic)을 참조하면, 기존에 활용된 XGBoost를 활용하는 것보다 LightGBM을 활용하는 것이 더 효율적인 것을 파악할 수 있다. 혈액 샘플만을 이용했음에도 방대한 의료용 행정데이터를 사용한 다른 기법들보다 좋은 결과가 나타난다.

[0060] 16만명의 환자 데이터에 관하여 1500 개의 변수를 53개월동안이나 모아 XGBoost로 응급실 재내원을 예측한 기술은 0.76의 AUC-ROC)를 나타낸다.

[0061] 본 발명에서 제시한 BER 기술은 간편한 71종의 혈액 데이터 변수를 사용하여 환자 데이터 480개 이상이고, ED return 비율이 약 9 퍼센트 이상인 조건을 동시에 충족할 때, 0.7813의 AUC-ROC 성능을 안정적으로 달성할 수 있다. 따라서 본 발명에서 제시하는 BER 기술이 효율적이라는 것을 확인할 수 있다.

[0062] 도 6에 도시된 BER 기법에 LightGBM을 적용했을 때와 XGBoost를 적용했을 때의 소요 시간을 참조하면, 본 발명의 LightGBM 기반 BER 기법의 평균 시간은 0.0437 초를 기록한 반면에 0.1487 초를 기록한 XGBoost 적용 기법

보다 약 30퍼센트의 시간만 소요한 것으로 실시간 모바일 헬스 시스템에서 LightGBM을 채택하는 것이 훨씬 적합하다는 것을 쉽게 알 수 있다.

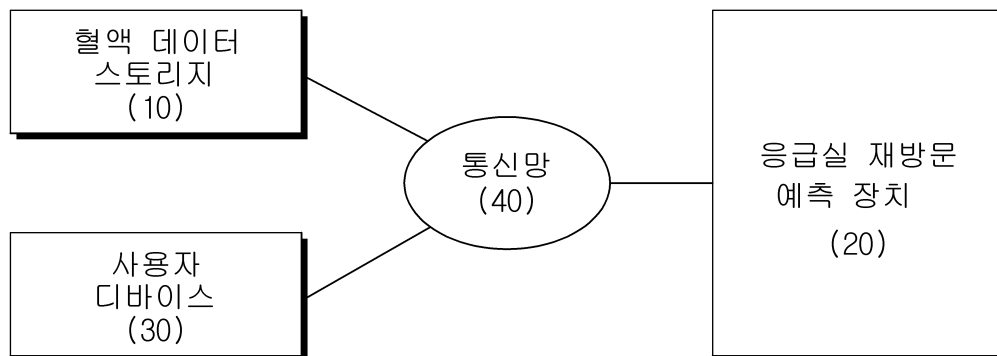
[0063] 도 4에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 4에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.

[0064] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.

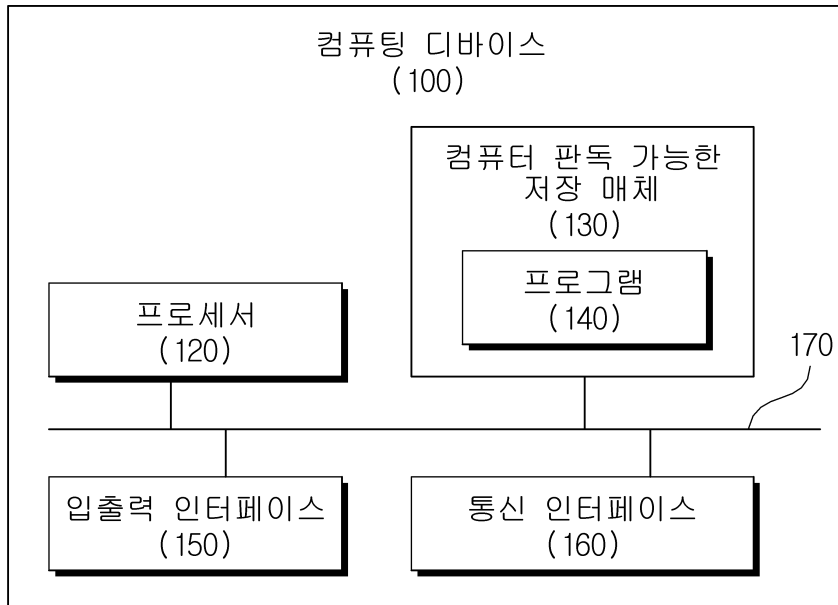
[0065] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

## 도면

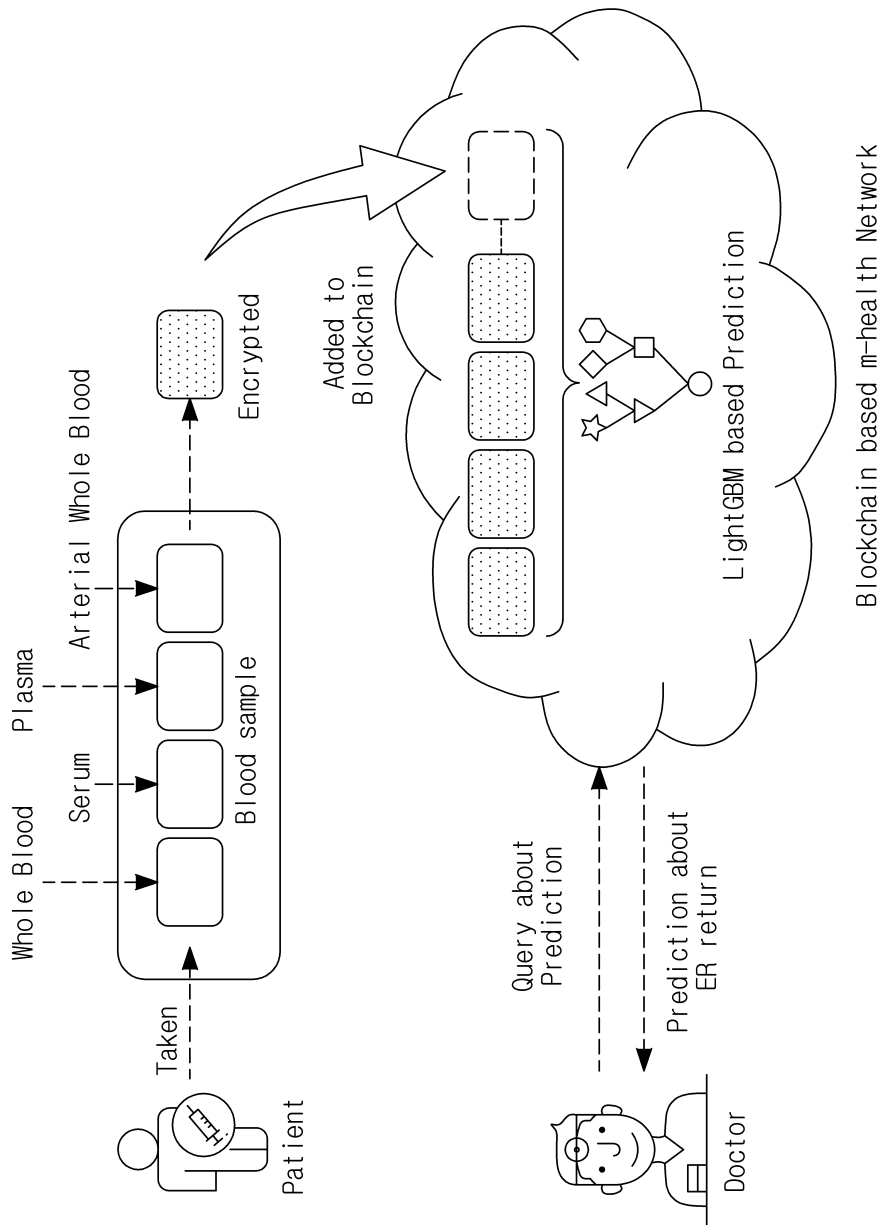
### 도면1



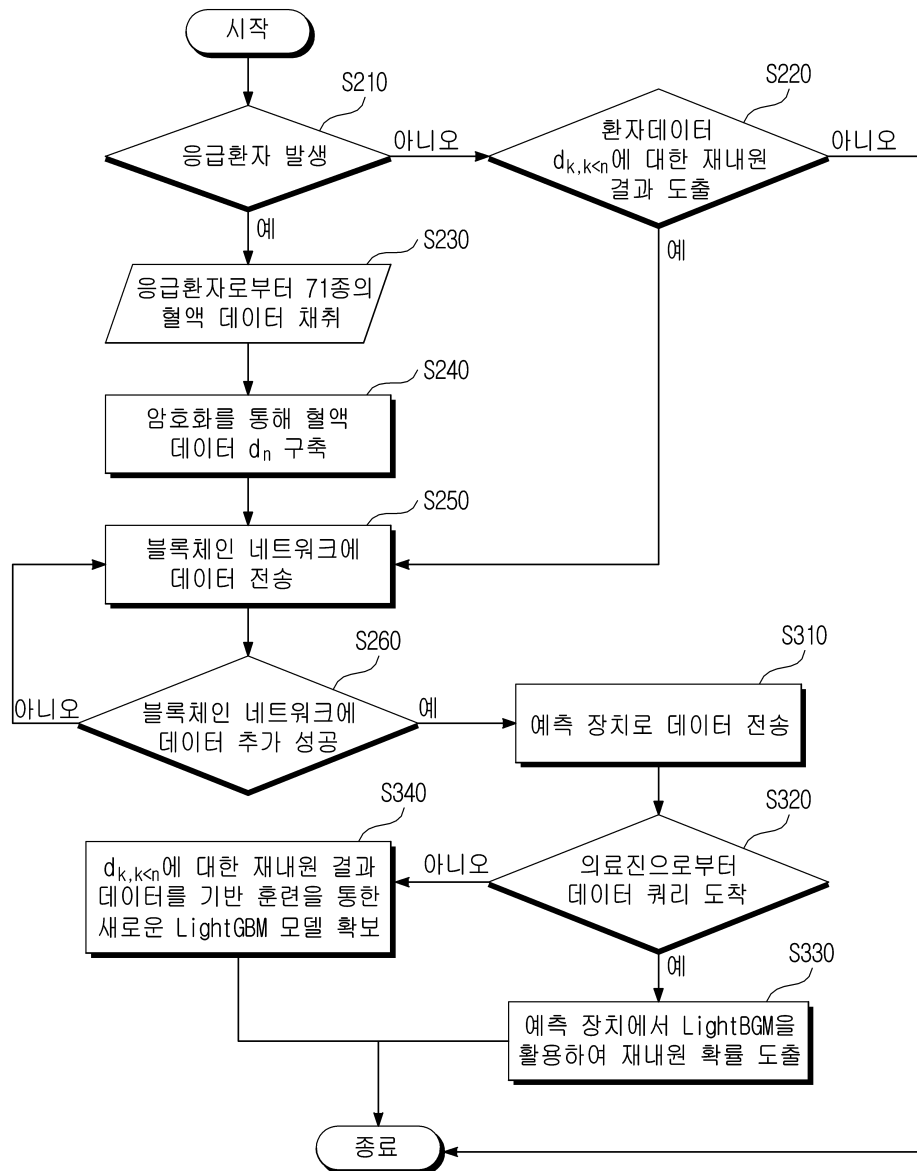
도면2



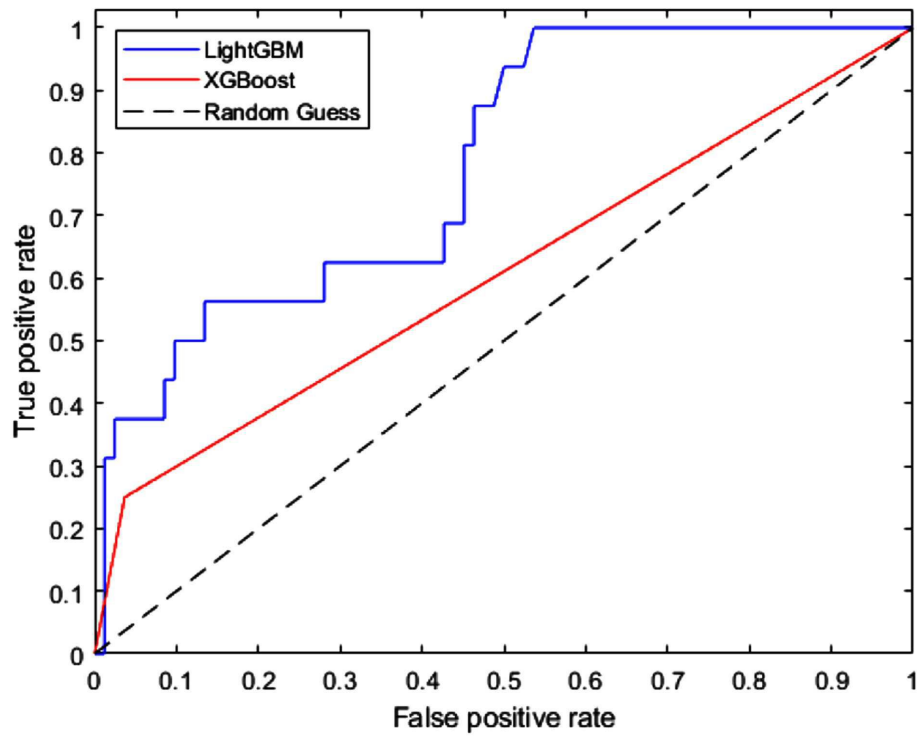
도면3



도면4



도면5



도면6

