



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0097515  
(43) 공개일자 2021년08월09일

(51) 국제특허분류(Int. Cl.)

G16H 50/50 (2018.01) G06N 3/08 (2006.01)  
G16H 10/60 (2018.01) G16H 50/30 (2018.01)  
G16H 50/70 (2018.01)

(52) CPC특허분류

G16H 50/50 (2018.01)  
G06N 3/08 (2013.01)

(21) 출원번호 10-2020-0011307

(22) 출원일자 2020년01월30일

심사청구일자 2020년01월30일

(71) 출원인

연세대학교 원주산학협력단

강원도 원주시 흥업면 연세대길 1

(72) 발명자

고상백

강원도 원주시 봉화로 231, 210동 1303호(우산동, 한라비발디2차아파트)

이솔암

강원도 원주시 백간길 17, 104동 1903호(단계동, 원주 봉화산 벨라시티 아파트)

이현주

강원도 원주시 무실로 30

(74) 대리인

특허법인비엘티

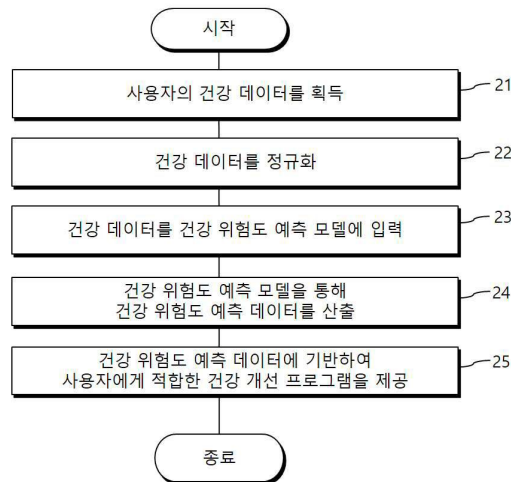
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법

(57) 요약

딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법이 제공된다. 본 발명의 일 면에 따른 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법은, 사용자의 건강 데이터를 획득하는 단계; 상기 획득한 건강 데이터를 미리 기계학습한 건강 위험도 예측 모델에 입력하여, 상기 건강 위험도 예측 모델을 통해 건강 위험도 예측 데이터를 산출하는 단계; 및 상기 건강 위험도 예측 데이터에 기반하여 상기 사용자에게 적합한 건강 개선 프로그램을 제공하는 단계;를 포함한다.

대표도 - 도2



(52) CPC특허분류

*G16H 10/60* (2018.01)

*G16H 50/30* (2018.01)

*G16H 50/70* (2018.01)

---

## 명세서

### 청구범위

#### 청구항 1

딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법에 있어서,

사용자의 건강 데이터를 획득하는 단계;

상기 획득한 건강 데이터를 미리 기계학습한 건강 위험도 예측 모델에 입력하여, 상기 건강 위험도 예측 모델을 통해 건강 위험도 예측 데이터를 산출하는 단계; 및

상기 건강 위험도 예측 데이터에 기반하여 상기 사용자에게 적합한 건강 개선 프로그램을 제공하는 단계;를 포함하는 것을 특징으로 하는 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법.

#### 청구항 2

제1 항에 있어서, 상기 건강 데이터를 정규화하는 단계;를 더 포함하는 것을 특징으로 하는 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법.

#### 청구항 3

제1 항에 있어서, 상기 건강 데이터는 나이, 성별, 키, 몸무게, 음주 여부, 음주량, 음주 횟수, 흡연 여부, 흡연량, 흡연 기간, 혈압, 허리 둘레, 공복 혈당, 당뇨약 복용여부, HDL, 트리글리세리드(triglyceride) 총콜레스테롤, 콜레스테롤 약 복용여부, 고혈압 여부, 진성 당뇨병(Diabetes mellitus) 여부, 이상지질혈증(Dyslipidemia) 여부, 통풍(Gout) 여부, 가족력, 과거 병력, 식습관, 운동습관, 대사당량(metabolic equivalents, METS) 및 수면습관 중 적어도 하나를 포함하는 것을 특징으로 하는 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법.

#### 청구항 4

제1 항에 있어서, 상기 건강 위험도 데이터는 n년 후 대사증후군 위험도 및 복수의 위험인자들 중 적어도 하나의 감소량 별 n년 후 대사증후군 위험도를 포함하고,

상기 복수의 위험인자들은 체중, 흡연, 음주, 식습관, 운동습관 및 수면습관 중 적어도 하나를 포함하는 것을 특징으로 하는 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법.

#### 청구항 5

제4 항에 있어서, 상기 건강 개선 프로그램은 상기 현재 대사증후군 위험도, 상기 복수의 위험인자들 중 적어도 하나의 감소량 별 n년 후 대사증후군 위험도 및 가이드 내용을 포함하는 것을 특징으로 하는 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법.

#### 청구항 6

제1 항에 있어서, 상기 사용자가 사용하는 전자 장치에 건강 안내 플랫폼을 통해 상기 건강 개선 프로그램을 제공하는 것을 특징으로 하는 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법.

#### 청구항 7

제1 항에 있어서,

복수의 사용자들로부터 과거 건강 데이터와 n년 후 건강 데이터를 획득하는 제1 단계;

상기 과거 건강 데이터를 정규화하고 상기 n년 후 건강 데이터에 기반한 n년 후 건강 위험도 예측 데이터를 산출하는 제2 단계; 및

상기 n년 후 건강 위험도 예측 데이터에 기반하여 상기 건강 위험도 예측 모델의 정답 데이터를 생성하고 상기

정규화한 과거 건강 데이터를 상기 건강 위험도 예측 모델에 입력하여, 상기 건강 위험도 예측 모델을 통해 연산한 건강 위험도 예측 데이터를 상기 정답 데이터와 비교하는 제3 단계;를 더 포함하는 것을 특징으로 하는 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법.

#### 청구항 8

제7 항에 있어서, 상기 제1 단계, 상기 제2 단계 및 상기 제3 단계를 반복적으로 수행함으로써 상기 건강 위험도 예측 모델을 지속적으로 기계학습하는 단계;를 더 포함하는 것을 특징으로 하는 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법.

#### 청구항 9

제1 항에 있어서, 상기 건강 위험도 예측 모델은 1차원 CNN(convolutioanal neural network) 또는 RNN(Recurrent Neural Network)인 것을 특징으로 하는 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법.

#### 청구항 10

하드웨어인 컴퓨터와 결합되어, 제1항 내지 제9항 중 어느 한 항의 방법을 실행시키기 위하여 매체에 저장된, 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 프로그램.

### 발명의 설명

#### 기술 분야

[0001] 본 발명은 딥러닝에 기반한 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법에 관한 것으로, 보다 자세하게는 딥러닝 기반 사용자의 건강 데이터를 분석하여 사용자의 위험인자 별 감소량에 따른 건강 위험도를 제공하는 방법에 관한 것이다.

#### 배경 기술

[0002] 최근 현대사회에서는 암, 심혈관 질환, 당뇨병, 비만, 대사증후군 등과 같은 질병이 일반인들에게 쉽게 나타나고 있는 실정이다. 그러나 종래의 경우 일부 질병을 제외하고는 주기적인 건강 검진을 통해서만 해당 질병이 있음을 진단받거나, 검진 결과에 대하여 의사의 확인으로 향후 질병이 발생할 가능성이 있음을 예측하고 있다.

[0003] 이 경우에도 의사의 경험적 판단이나 또는 질병 간의 영향력이나 보편성을 미리 정해진 수치 또는 정상 상태와의 단순 비교를 통해 추후 질병 발생 가능성을 모호하게 예측하고만 있다는 문제가 있었다.

[0004] 이와 같이, 자신의 건강상태를 확인하기 위해서는 비싼 건강검진에만 의존해야 하는 문제가 있는바, 평상시 건강상태를 수시로 확인하고 자신의 건강상태에 적합한 처방을 받아 건강관리를 꾸준히 하고, 이에 따라 최종적으로 건강이 위험해질 가능성(또는 질병 발생 가능성)을 낮추어야 할 개인적, 사회적 요구가 증가하고 있는 실정이다.

### 선행기술문헌

#### 비특허문헌

[0005] (비특허문헌 0001) (특허문헌 0001) US 2015-0154355

### 발명의 내용

#### 해결하려는 과제

[0006] 본 발명이 해결하고자 하는 과제는 사용자의 건강 데이터를 딥 러닝에 기반하여 분석함으로써 사용자에게 적합한 건강 개선 프로그램을 생성하는 방법을 제공하는 것이다.

[0007] 본 발명이 해결하고자 하는 과제들은 이상에서 언급된 과제로 제한되지 않으며, 언급되지 않은 또 다른 과제들

은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

### 과제의 해결 수단

[0008] 상술한 과제를 해결하기 위한 본 발명의 일면에 따른 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법은, 사용자의 건강 데이터를 획득하는 단계; 상기 획득한 건강 데이터를 미리 기계학습한 건강 위험도 예측 모델에 입력하여, 상기 건강 위험도 예측 모델을 통해 건강 위험도 예측 데이터를 산출하는 단계; 및 상기 건강 위험도 예측 데이터에 기반하여 상기 사용자에게 적합한 건강 개선 프로그램을 제공하는 단계;를 포함한다.

[0009] 본 발명의 기타 구체적인 사항들은 상세한 설명 및 도면들에 포함되어 있다.

### 발명의 효과

[0010] 상기와 같은 본 발명에 따르면, 아래와 같은 다양한 효과들을 가진다.

[0011] 본 발명은 현재 시점의 사용자의 건강 데이터를 이용함으로써 사용자의 향후 건강 위험도의 정확도를 높일 수 있다.

[0012] 또한, 본 발명은 사용자에게 적합한 위험인자 별 감소량에 따른 건강 위험도가 포함된 건강 개선 프로그램을 제공할 수 있다.

[0013] 또한, 본 발명은 딥러닝을 통한 사전기계학습을 통해 건강 위험도의 신뢰성을 높일 수 있다.

[0014] 또한, 본 발명은 건강 개선 프로그램을 제공함으로써 사용자의 건강 개선 활동을 유도할 수 있다.

[0015] 본 발명의 효과들은 이상에서 언급된 효과로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

### 도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시 예에 따른 건강 프로그램 제공 장치를 나타낸 블록도이다.

도 2는 본 발명의 일 실시 예에 따른 건강 프로그램 제공 방법을 설명하기 위한 흐름도이다.

도 3은 본 발명의 일 실시 예에 따른 건강 위험도 예측 모델을 기계학습하는 방법을 설명하기 위한 흐름도이다.

도 4는 본 발명의 일 실시 예에 따른 건강 위험도 예측 방법을 나타낸 블록도이다.

도 5는 본 발명의 일 실시 예에 따른 건강 프로그램을 나타낸 예시도이다.

도 6 및 도 7은 본 발명의 효과를 나타낸 예시도이다.

### 발명을 실시하기 위한 구체적인 내용

[0017] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나, 본 발명은 이하에서 개시되는 실시예들에 제한되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술 분야의 통상의 기술자에게 본 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다.

[0018] 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성요소 외에 하나 이상의 다른 구성요소의 존재 또는 추가를 배제하지 않는다. 명세서 전체에 걸쳐 동일한 도면 부호는 동일한 구성 요소를 지칭하며, "및/또는"은 언급된 구성요소들의 각각 및 하나 이상의 모든 조합을 포함한다. 비록 "제1", "제2" 등이 다양한 구성요소들을 서술하기 위해서 사용되나, 이들 구성요소들은 이들 용어에 의해 제한되지 않음은 물론이다. 이들 용어들은 단지 하나의 구성요소를 다른 구성요소와 구별하기 위하여 사용하는 것이다. 따라서, 이하에서 언급되는 제1 구성요소는 본 발명의 기술적 사상 내에서 제2 구성요소일 수도 있음은 물론이다.

[0019] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 발명이 속하는 기술

분야의 통상의 기술자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있을 것이다. 또한, 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.

- [0020] 공간적으로 상대적인 용어인 "아래(below)", "아래(beneath)", "하부(lower)", "위(above)", "상부(upper)" 등은 도면에 도시되어 있는 바와 같이 하나의 구성요소와 다른 구성요소들과의 상관관계를 용이하게 기술하기 위해 사용될 수 있다. 공간적으로 상대적인 용어는 도면에 도시되어 있는 방향에 더하여 사용시 또는 동작시 구성요소들의 서로 다른 방향을 포함하는 용어로 이해되어야 한다. 예를 들어, 도면에 도시되어 있는 구성요소를 뒤집을 경우, 다른 구성요소의 "아래(below)"또는 "아래(beneath)"로 기술된 구성요소는 다른 구성요소의 "위(above)"에 놓여질 수 있다. 따라서, 예시적인 용어인 "아래"는 아래와 위의 방향을 모두 포함할 수 있다. 구성요소는 다른 방향으로도 배향될 수 있으며, 이에 따라 공간적으로 상대적인 용어들은 배향에 따라 해석될 수 있다.
- [0021] 이하, 첨부된 도면을 참조하여 본 발명의 실시예를 상세하게 설명한다.
- [0022] 도 1은 본 발명의 일 실시 예에 따른 건강 프로그램 제공 장치를 나타낸 블록도이다.
- [0023] 도 1을 참조하면, 본 발명의 일 실시 예에 따른 건강 개선 프로그램 제공 장치(100)는 사용자로부터 획득한 건강 데이터를 이용하여 사용자의 건강 위험도를 예측할 수 있다. 예를 들어, 건강 개선 프로그램 제공 장치(100)는 건강 데이터를 입력값으로 하는 건강 위험도 예측 모델을 이용하여 건강 위험도 예측 데이터를 출력할 수 있고, 건강 위험도 예측 데이터에 기반하여 건강 개선 프로그램을 생성하고 사용자에게 제공할 수 있다.
- [0024] 일 실시 예에서, 건강 개선 프로그램 제공 장치(100)는 서버일 수 있고, 딥러닝 학습 모델, 건강 위험도 예측 및 프로그램 생성과 제공 방법을 설정할 수 있는 전용 프로그램이 설치될 수 있다. 예를 들어, 건강 개선 프로그램 제공 장치(100)는 건강 데이터를 획득할 수 있는 데이터 획득부(110), 획득한 데이터들을 처리(예: 데이터 정규화)할 수 있는 데이터 처리부(120), 학습데이터와 건강 위험도 예측 모델을 이용하여 기계학습을 수행하고 건강 위험도 예측 데이터를 산출하는 딥러닝부(130), 건강 개선 프로그램을 생성하고 제공하는 프로그램 제공부(140) 및 학습 결과, 건강 위험도 예측 모델의 구조, 필터 값 및 활성화 함수, 사용자의 건강 데이터, 건강 위험도 예측 데이터 등을 빅데이터화하여 저장할 수 있는 데이터베이스(150)를 포함할 수 있다.
- [0025] 일 실시 예에서, 저혈압 예측 장치(100)는 전자 장치(200)와 네트워크를 통해 연결될 수 있다. 여기서 네트워크는 무선 네트워크 및 유선 네트워크를 포함할 수 있다. 예를 들어, 상기 네트워크는 근거리 통신 네트워크(예: 블루투스, WiFi direct 또는 IrDA(infrared data association)) 또는 원거리 통신 네트워크(예: 셀룰러 네트워크, 인터넷, 또는 컴퓨터 네트워크(예: LAN 또는 WAN))일 수 있다. 전자 장치(200)는 건강 위험도 확인을 원하는 사용자가 사용하는 장치일 수 있다. 전자 장치(200)는 건강 안내 플랫폼을 통해 건강 개선 프로그램을 제공할 수 있다. 여기서 건강 안내 플랫폼은 웹 페이지 또는 전용 앱 어플리케이션일 수 있다. 사용자는 상기 건강 안내 플랫폼을 통해 주기적으로 본인의 건강 데이터를 저혈압 예측 장치(100)에 제공할 수 있다.
- [0026] 예를 들어, 전자 장치(200)는 건강 안내 플랫폼이 설치된 장치일 수 있다. 또한, 전자 장치(200)는, 예를 들면, 스마트폰(smartphone), 태블릿 PC(tablet personal computer), 이동 전화기(mobile phone), 영상 전화기, 전자책 리더기(e-book reader), 데스크탑 PC (desktop PC), 랩탑 PC(laptop PC), 넷북 컴퓨터(netbook computer), 워크스테이션(workstation), 서버(server), PDA(personal digital assistant), PMP(portable multimedia player), MP3 플레이어, 모바일 의료기기, 카메라, 또는 웨어러블 장치(wearable device) 중 적어도 하나를 포함할 수 있다.
- [0027] 도 2는 본 발명의 일 실시 예에 따른 건강 프로그램 제공 방법을 설명하기 위한 흐름도이다. 도 3은 본 발명의 일 실시 예에 따른 건강 위험도 예측 모델을 기계학습하는 방법을 설명하기 위한 흐름도이다. 도 4는 본 발명의 일 실시 예에 따른 건강 위험도 예측 방법을 나타낸 블록도이다. 도 5는 본 발명의 일 실시 예에 따른 건강 프로그램을 나타낸 예시도이다. 도 6 및 도 7은 본 발명의 효과를 나타낸 예시도이다. 도 2 및 도 3의 동작들은 도 1의 건강 프로그램 제공 장치(100)를 통해 수행될 수 있다.
- [0028] 도 2 내지 도 7을 참조하면, 일 실시 예에서, 동작 21에서, 데이터 획득부(110)는 사용자의 건강 데이터를 획득할 수 있다. 예를 들어, 데이터 획득부(110)는 외부 장치로부터 데이터들을 수신하거나 플랫폼(예: 웹 페이지, 어플리케이션)을 통해 사용자로부터 입력 받을 수 있다.
- [0029] 예를 들어, 건강 데이터는 나이, 성별, 키, 몸무게, 음주 여부, 음주량, 음주 횟수, 흡연 여부, 흡연량, 흡연 기간, 혈압, 허리 둘레, 공복 혈당, 당뇨약 복용여부, HDL, 트리글리세리드(triglyceride) 총콜레스테롤, 콜레



스테롤 약 복용여부, 고혈압 여부, 진성 당뇨병(Diabetes mellitus) 여부, 이상지질혈증(Dyslipidemia) 여부, 통풍(Gout) 여부, 가족력, 과거 병력, 식습관, 운동습관, 대사당량(metabolic equivalents, METS) 및 수면습관 중 적어도 하나를 포함할 수 있다.

[0030] 일 실시 예에서, 동작 22에서, 데이터 처리부(120)는 건강 데이터를 정규화할 수 있다. 예를 들어, 건강데이터는 원본 데이터 그대로 건강 위험도 예측 모델에 입력하기 불가능하므로 건강 데이터에 전처리 동작이 수행되고, 표준화 학습 데이터가 생성될 수 있다. 즉, 건강 데이터는 전처리 동작을 통해 인공지능 신경망에 입력될 수 있는 표준화된 형태(즉, 표준화 학습 데이터)로 변형될 수 있다. 예를 들어, 건강 데이터는 미리 정해진 규칙에 따라 수치 값으로 변환되고, 변환된 수치 값은 인공지능 신경망의 입력 값으로 사용되기 위해 백터로 변환될 수 있다. 즉, 정규화 동작은 건강 데이터의 본래 의미에 영향을 주지 않는 범위에서 인공지능 신경망에 적합하게 변경하는 동작을 나타낸다.

[0031] 동작 23 내지 동작 25를 설명하기 전에 도 3을 참고하여 건강 위험도 예측 모델의 기계학습 과정을 구체적으로 설명한다.

[0032] 일 실시 예에서, 동작 31에서, 데이터 획득부(110)는 복수의 사용자들로부터 과거 건강 데이터와 n년 후 건강 데이터를 획득할 수 있다. 예를 들어, 데이터 획득부(110)는 현재 시점 기준으로 누적된 건강 데이터를 획득할 수 있고, 일정 기간 이전의 건강 데이터는 과거 건강 데이터로 분류하고 과거 건강 데이터 이후 데이터는 n년 후 건강 데이터로 분류할 수 있다. 예컨대, 10년 분량의 건강 데이터를 획득할 경우, 최초 기록 시점부터 7년까지를 과거 건강 데이터로 분류하고 8년 시점의 데이터를 1년 후 데이터, 9년 시점의 데이터를 2년 후 데이터, 10년 시점의 데이터를 3년 후 데이터로 분류할 수 있다. 물론 이러한 분류는 일 예에 불과하고 기간 및 시점은 다양하게 변형될 수 있다.

[0033] 일 실시 예에서, 동작 32에서, 데이터 처리부(120)는 과거 건강 데이터를 정규화하고 n년 후 건강 데이터에 기반한 n년 후 건강 위험도 예측 데이터를 산출할 수 있다. 예를 들어, 과거 건강 데이터와 n년 후 건강 데이터는 모두 학습데이터로써 활용될 수 있다. 여기서 과거 건강 데이터는 훈련을 위한 입력 데이터가 될 수 있으므로 정규화가 필요할 수 있다. 또한, 여기서 n년 후 건강 데이터에 기반하여 산출한 n년 후 건강 위험도 예측 데이터는 훈련을 위한 정답 데이터로써 활용될 수 있다. n년 후 건강 데이터는 실제로는 동작 31에서 설명한 바와 같이 현재 시점 기준으로는 과거 데이터이므로 이를 분석할 경우 건강 위험도가 쉽게 도출될 수 있다.

[0034] 일 실시 예에서, 동작 33에서, 딥러닝부(130)는 n년 후 건강 위험도 예측 데이터에 기반하여 건강 위험도 예측 모델의 정답 데이터를 생성하고 정규화한 과거 건강 데이터를 건강 위험도 예측 모델에 입력할 수 있다. 예를 들어, 건강 위험도 예측 모델은 딥 러닝 기법에 기반한 학습 모델일 수 있고, 기계학습을 위해 입력 값과 정답 데이터가 필요하다. 여기서 입력 값은 동작 32에서 정규화한 과거 건강 데이터일 수 있고, 정답 데이터는 n년 후 건강 위험도 예측 데이터에 기반하여 생성되며 현재 대사증후군 위험도 및 복수의 위험인자들 중 적어도 하나의 감소량 별 n년 후 대사증후군 위험도를 포함할 수 있다. 여기서 정답 데이터는 딥러닝 모델의 ground truth를 의미한다.

[0035] 예를 들어, 건강 위험도 예측 모델은 1차원 CNN(convolutional neural network) 또는 RNN(Recurrent Neural Network)인 딥 뉴럴 네트워크, XG boost 또는 Random forest 일 수 있다.

[0036] 예를 들어, RNN(Recurrent Neural Network)은 한 개 이상의 특징값(벡터)이 입력되며, 새롭게 입력된 특징값에 이전에 입력된 특징 값에 대한 은닉층 또는 출력층의 출력값도 함께 입력되는 특징을 가지고 있다. 1차원 CNN(Convolutional Neural Network)은 1차원의 벡터 또는 행렬을 입력으로 받으며, 일반적인 ANN 단계 이전에 Convolution단계 또는 pooling단계 등이 이루어질 수 있다. Convolution 단계에서는 가중치(weight)를 가진 다차원의 행렬과 convolution 연산을 수행한다. pooling 단계에는 max pooling 또는 mean pooling 방식이 있다.

[0037] 예컨대, 1차원 CNN의 경우, 가중치(weight)를 다르게 부여한 1차원 필터를 이용하여 이미지 데이터로부터 1차원적 특징들을 보다 용이하게 추출할 수 있다. 예컨대, 주파수 대역별로 특징이 다른 이미지 데이터가 있는 경우, 각각의 시간 구간에 대하여 주파수 대역별로 다른 가중치를 갖는 1차원 필터를 이미지 데이터에 적용하여 2D CNN에 비해 좀더 세밀하게 주파수에 대한 특징을 추출할 수 있다.

[0038] 예를 들어, Random Forest는 수많은 Decision Tree들이 Forest를 구성하여 각각의 예측결과를 하나의 결과변수로 평균화 하는 알고리즘이고, XGBoost는 Random Forest의 Tree는 독립적이라면 XGBoost의 Tree의 결과를 다음 트리에 적용하는 boost방식의 알고리즘이다.

[0039] 일 실시 예에서, 동작 34에서, 딥러닝부(130)는 건강 위험도 예측 모델을 통해 연산한 건강 위험도 예측 데이터

를 정답 데이터와 비교할 수 있다. 예를 들어, 도 4에 도시된 바와 같이, 건강 위험도 예측 데이터는 n년 후 건강 위험도, 위험인자 감소량 별 건강 위험도 및 위험인자 감소량 조합 별 건강 위험도를 포함할 수 있다. 물론 이외에도 건강 위험도 예측 데이터는 현재 시점의 건강 위험도도 포함할 수 있다. 여기서, 복수의 위험인자들은 체중, 흡연, 음주, 식습관, 운동습관 및 수면습관 중 적어도 하나를 포함할 수 있다.

[0040] 구체적으로, 일 예에서 건강 위험도는 암, 심혈관 질환, 당뇨병, 비만 또는 대사증후군 등을 포함할 수 있다. 설명의 편의를 위해서 대사증후군을 일 예로 들어 하기에서 설명한다.

[0041] 예를 들어, 건강 위험도가 대사증후군 위험도일 경우, 건강 위험도 예측 데이터는 현재 대사증후군 위험도, n년 후 대사증후군 위험도, 복수의 위험인자들 중 적어도 하나의 감소량 별 n년 후 대사증후군 위험도를 포함할 수 있다. 예컨대, 위험인자 감소량 별 대사증후군 위험도는 1개의 위험인자 감소량 별 대사증후군 위험도일 수 있고, 위험인자가 체중일 경우 n년 후 체중조절 감소량에 따른 대사증후군 위험도가 될 수 있다. 또한, 예를 들어, 위험인자 감소량 조합 별 대사증후군 위험도는 2개 이상의 위험인자 감소량 별 대사증후군 위험도일 수 있고, 위험인자가 체중과 흡연일 경우 n년 후 체중 조절 감소량과 금연 기간 조합에 따른 대사증후군 위험도가 될 수 있다.

[0042] 한편, 건강 위험도 예측 데이터에 포함된 위험도는 0~100 사이의 값을 가질 수 있고 이는 해당 건강 위험도가 발생할 확률일 수 있다. 예컨대, 대사증후군 위험도가 90으로 나올 경우 이는 대사증후군이 발생할 확률이 90%라는 의미이다.

[0043] 일 실시 예에서, 동작 35에서, 딥러닝부(130)는 동작 31 내지 동작 34를 반복적으로 수행함으로써 건강 위험도 예측 모델을 지속적으로 기계학습할 수 있다. 이를 통해, 건강 위험도 예측 모델의 필터 값 등을 조정할 수 있고, 건강 위험도 예측 데이터의 신뢰성을 높일 수 있다.

[0044] 상기 동작 31 내지 35를 통해 미리 기계학습된 건강 위험도 예측 모델을 이용하여 하기 동작 23 내지 동작 25가 수행될 수 있다.

[0045] 일 실시 예에서, 동작 23에서, 딥러닝부(130)는 건강 데이터를 미리 기계학습한 건강 위험도 예측 모델에 입력할 수 있다. 예를 들어, 건강 위험도 예측 모델의 입력 값은 사용자의 현재 시점의 건강 데이터일 수 있다.

[0046] 일 실시 예에서, 동작 24에서, 딥러닝부(130)는 건강 위험도 예측 모델을 통해 건강 위험도 예측 데이터를 산출할 수 있다. 예를 들어, 건강 위험도 예측 데이터는 n년 후 건강 위험도, 위험인자 감소량 별 건강 위험도, 위험인자 감소량 조합 별 건강 위험도 및 현재 시점의 건강 위험도를 포함할 수 있다. 건강 위험도가 대사증후군일 경우, 건강 위험도 데이터는 현재 대사증후군 위험도 및 복수의 위험인자들 중 적어도 하나의 감소량 별 n년 후 대사증후군 위험도를 포함할 수 있고, 여기서 복수의 위험인자들은 체중, 흡연, 음주, 식습관, 운동습관 및 수면습관 중 적어도 하나를 포함할 수 있다.

[0047] 예를 들어, 건강 위험도 예측 데이터는 현재 대사증후군 위험도, n년 후 대사증후군 위험도, 위험인자 감소량 별 대사증후군 위험도, 위험인자 감소량 조합 별 대사증후군 위험도를 포함할 수 있다. 예컨대, 위험인자가 체중 1개일 경우, 2년 후 체중 감소량이 5kg일 경우 대사증후군 위험도가 88.02%로 출력될 수 있고, 위험인자가 체중과 흡연일 경우, 2년 후 체중 감소량이 5kg이고 금연 기간이 6개월 이상일 경우 대사증후군 위험도는 70%로 출력될 수 있다. 즉, 위험인자 하나당 감소량이 클수록 대사증후군 위험도는 낮아질 수 있고, 위험인자 조합에서 각각의 위험인자 가중치를 고려하여 대사증후군 위험도가 결정될 수 있다.

[0048] 일 실시 예에서, 동작 25에서, 프로그램 제공부(140)는 건강 위험도 예측 데이터에 기반하여 사용자에게 적합한 건강 개선 프로그램을 제공할 수 있다. 예를 들어, 건강 개선 프로그램이란 사용자에게 건강 위험도를 낮출 수 있는 방법을 위험인자와 관련하여 가이드하는 권고 사항 등을 포함하는 프로그램(또는 일정이나 계획)일 수 있다. 예를 들어, 도 5에 도시된 바와 같이, (A) 건강 프로파일을 갖는 사용자에게 (B) 형태의 메시지를 포함하는 건강 개선 프로그램을 제공할 수 있다.

[0049] 예를 들어 건강 개선 프로그램은 현재 대사증후군 위험도, 복수의 위험인자들 중 적어도 하나의 감소량 별 n년 후 대사증후군 위험도 및 가이드 내용을 포함할 수 있다. 여기서 가이드 내용은 사용자에게 대사증후군의 위험도를 알리는 경고 문구와 내용의 이해를 돕는 문구 등이 포함될 수 있다. 또한, 대사증후군 위험도의 수치 값은 해당 확률 값 이외에도 위험인자 감소가 없을 때 대비하여 감소되는 위험도 확률 값이 될 수 있다.

[0050] 예를 들어, 도 5의 (B)에 도시된 바와 같이, 건강 개선 프로그램은 대사증후군에 해당할 수 있다는 경고문구, 복수의 위험인자 중 어느 한 위험인자도 2년 후에 감소시키거나 중단시키지 못할 경우의 대사증후군 걸릴 확률



(도 5의 90.95%), 위험인자가 체중 1개일 경우 2년 동안 체중 감소량에 따른 대사증후군 걸릴 확률의 감소량(도 5의 2.93% lower risk with 5kg loss 등)을 포함할 수 있다.

[0051] 또한, 위험 인자가 2개 이상일 경우, 예를 들면, 2년 동안 체중 감소량이 5kg이고 금연 기간이 6개월일 경우 대사증후군 걸릴 확률의 감소량은 도 5의 2.93%보다 더 높은 수치(예: 9.6%)가 될 수 있다.

[0052] 즉, 본 발명은 앞서 기재한 바와 같이 1개의 위험인자 감소량에 따른 건강 위험도의 현재 확률 또는 확률의 감소량을 제공할 뿐만 아니라 2개 이상의 위험인자 조합의 감소량에 따른 건강 위험도의 현재 확률 또는 확률의 감소량도 제공할 수 있다. 따라서, 본 발명을 통해 체중, 흡연, 음주, 식습관, 운동습관 및 수면습관 등을 포함하는 복수의 위험인자들의 건강 위험도에 대한 영향을 사용자가 손쉽게 확인할 수 있다.

[0053] 또한, 사용자는 본 발명에서 제공하는 건강 개선 프로그램을 통해 본인의 건강 위험도, 어떤 위험인자들을 어떤 일정(예: 2년까지 몇키로 감량 등)에 따라 수행할 경우 건강 위험도가 얼마나 낮아지는 지를 손쉽게 알 수 있으며 보다 효율적으로 본인의 건강 위험을 낮출 수 있다.

[0054] 본 발명의 효과를 확인하기 위해, 7321명의 참가자들을 27945 번의 신뢰할만한 연속적인 방문을 통해 학습데이터를 수집하였다. 여기서 참가자들은 나이가 65살 이하, 악성 질환이 없고, 2년간의 연속적인 방문이 이루어진 참가자들이다. 이중에서 5124명의 참가자의 건강 데이터를 학습 데이터(train data)로 추출하고 2197명의 건강 데이터를 검증 데이터로 추출하였다. 예를 들어, 학습데이터와 검증 데이터는 하기 표 1과 같다.

표 1

Characteristics	Train Cohort (n=5124)	Validation Cohort (n=2197)	p
Age, year	50.7 ± 7.5	50.6 ± 7.6	0.587
Sex, n (%)			0.401
Male	2468 (48.2%)	1034 (47.1%)	
Female	2656 (51.8%)	1163 (52.9%)	
Height, cm	160.5 ± 8.6	160.5 ± 8.5	0.952
Weight, kg	63.6 ± 10.0	63.5 ± 10.1	0.617
Self-reported lifestyle			
Alcohol intake			
Ever, n (%)	2791 (54.5%)	1167 (53.1%)	0.299
Current, n (%)	2487 (48.5%)	1033 (47.0%)	0.244
Amount, g/week	10.0 ± 23.3	8.9 ± 20.6	0.033
Smoking			
Ever, n (%)	2031 (39.6%)	851 (38.7%)	0.485
Current, n (%)	1222 (23.8%)	514 (23.4%)	0.698
Amount, pack	7.4 ± 11.4	7.3 ± 11.3	0.591
Duration, year	9.4 ± 13.4	9.2 ± 13.1	0.452
Self-reported history, n (%)			
Hypertension	572 (11.2%)	247 (11.2%)	0.953
Diabetes mellitus	105 ( 2.0%)	56 ( 2.5%)	0.212
Dyslipidemia	121 ( 2.4%)	56 ( 2.5%)	0.692
Gout	209 ( 4.1%)	106 ( 4.8%)	0.168
Metabolic syndrome			
Yes, n (%)	1189 (23.2%)	509 (23.2%)	0.997
Component, n (%)			
Waist circumference	1466 (28.6%)	635 (28.9%)	0.822
Triglyceride	1787 (34.9%)	772 (35.1%)	0.849
HDL	2022 (39.5%)	859 (39.1%)	0.791
Glucose	804 (15.7%)	353 (16.1%)	0.712
Blood pressure	1865 (36.4%)	791 (36.0%)	0.768
No. of components, n (%)			0.023

0	1220 (23.8%)	555 (25.3%)	
1	1505 (29.4%)	625 (28.4%)	
2	1210 (23.6%)	508 (23.1%)	
3	797 (15.6%)	313 (14.2%)	
4	332 ( 6.5%)	150 ( 6.8%)	
5	60 ( 1.2%)	46 ( 2.1%)	

- [0056] 상기 자료를 통해 도 6에 도시된 바와 같이, 건강 데이터에 기반한 2년 후와 4년 후 활동 강도(metabolic equivalents : METs)를 예측하였고, 회귀 분석(Logistic regression)에 비하여 본 발명의 건강 위험도 예측 모델에 사용되는 XG boost, 딥 뉴럴 네트워크 또는 랜덤 포레스트(random forest)가 AUC가 더 높음을 확인할 수 있었다.
- [0057] 또한, 상기 자료를 통해 도 7에 도시된 바와 같이 (A) 과체중, (B) 일반체중, (C) 저체중인 참가자를 대상으로 weight에 변화를 주었을 경우에 2년 후 대사증후군의 위험도를 보면, (B) 일반체중에서는 회귀 분석과 본 발명의 XG boost가 동일한 예측을 하고 있지만, Linear assumption을 가정하는 logistic regression에서는 (A)에서 이미 심하게 과체중인 사람이 살이 더 쪼다고 해서 위험도가 계속 가파르게 증가하거나 (C)에서 저체중인 사람이 살을 더 뺀다고 해서 위험도가 계속 감소하는 모습을 보이지만, 실제로는 체중의 양 극단에서는 그 소량의 체중 변화에 의한 영향이 별로 없으므로 본 발명의 XG boost 와 같이 어느 정도 한계점 이후에서는 그 변화량이 거의 나타나지 않는 것이 성능 측면에서 우수함을 나타낸다.
- [0058] 동맥 혈압 데이터만 이용한 학습 모델(1-channel CNN)과 3개의 데이터(동맥 혈압 데이터, 심전도 데이터 및 이산화탄소 분압 데이터)를 이용한 본 발명의 학습 모델(3-channel CNN)을 비교한 결과, 도 12 및 도 13에 도시된 바와 같이 5분 후 저혈압 예측, 10분 후 저혈압 예측, 15분 후 저혈압 예측 모두 본 발명의 학습 모델이 동맥 혈압 데이터만 이용한 학습 모델에 비해 ROC 커브의 면적(AUROC=AUC)이 더 큼을 확인할 수 있었다. ROC 커브의 면적은 AUC(accuracy)로써 학습 모델 성능의 비교 지표로써 더 클수록 학습 모델 성능이 좋다.
- [0059] 본 발명의 일면에 따른 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 방법은, 사용자의 건강 데이터를 획득하는 단계; 상기 획득한 건강 데이터를 미리 기계학습한 건강 위험도 예측 모델에 입력하여, 상기 건강 위험도 예측 모델을 통해 건강 위험도 예측 데이터를 산출하는 단계; 및 상기 건강 위험도 예측 데이터에 기반하여 상기 사용자에게 적합한 건강 개선 프로그램을 제공하는 단계;를 포함한다.
- [0060] 다양한 실시 예에 따르면, 상기 건강 데이터를 정규화하는 단계;를 더 포함할 수 있다.
- [0061] 다양한 실시 예에 따르면, 상기 건강 데이터는 나이, 성별, 키, 몸무게, 음주 여부, 음주량, 음주 횟수, 흡연 여부, 흡연량, 흡연 기간, 혈압, 허리 둘레, 공복 혈당, 당뇨약 복용여부, HDL, 트리글리세리드(triglyceride) 총콜레스테롤, 콜레스테롤 약 복용여부, 고혈압 여부, 진성 당뇨병(Diabetes mellitus) 여부, 이상지질혈증(Dyslipidemia) 여부, 통풍(Gout) 여부, 가족력, 과거 병력, 식습관, 운동습관, 대사당량(metabolic equivalents, METS) 및 수면습관 중 적어도 하나를 포함할 수 있다.
- [0062] 다양한 실시 예에 따르면, 상기 건강 위험도 데이터는 n년 후 대사증후군 위험도 및 복수의 위험인자들 중 적어도 하나의 감소량 별 n년 후 대사증후군 위험도를 포함하고, 상기 복수의 위험인자들은 체중, 흡연, 음주, 식습관, 운동습관 및 수면습관 중 적어도 하나를 포함할 수 있다.
- [0063] 다양한 실시 예에 따르면, 상기 건강 개선 프로그램은 상기 현재 대사증후군 위험도, 상기 복수의 위험인자들 중 적어도 하나의 감소량 별 n년 후 대사증후군 위험도 및 가이드 내용을 포함할 수 있다.
- [0064] 다양한 실시 예에 따르면, 상기 사용자가 사용하는 전자 장치에 건강 안내 플랫폼을 통해 상기 건강 개선 프로그램을 제공할 수 있다.
- [0065] 다양한 실시 예에 따르면, 복수의 사용자들로부터 과거 건강 데이터와 n년 후 건강 데이터를 획득하는 제1 단계; 상기 과거 건강 데이터를 정규화하고 상기 n년 후 건강 데이터에 기반한 n년 후 건강 위험도 데이터를 산출하는 제2 단계; 및 상기 n년 후 건강 위험도 데이터에 기반하여 상기 건강 위험도 예측 모델의 정답 데이터를 생성하고 상기 정규화한 과거 건강 데이터를 상기 건강 위험도 예측 모델에 입력하여, 상기 건강 위험도 예측 모델을 통해 연산한 건강 위험도 데이터를 상기 정답 데이터와 비교하는 제3 단계;를 더 포함할 수 있다.
- [0066] 다양한 실시 예에 따르면, 상기 제1 단계, 상기 제2 단계 및 상기 제3 단계를 반복적으로 수행함으로써 상기 건

강 위험도 예측 모델을 지속적으로 기계학습하는 단계;를 더 포함할 수 있다.

[0067] 다양한 실시 예에 따르면, 상기 건강 위험도 예측 모델은 1차원 CNN(convolutioanal neural network) 또는 RNN(Recurrent Neural Network)일 수 있다.

[0068] 다양한 실시 예에 따르면, 딥러닝 기반 건강 위험도 예측을 통한 건강 개선 프로그램 제공 프로그램은 하드웨어인 컴퓨터와 결합되어, 제1항 내지 제9항 중 어느 한 항의 방법을 실행시키기 위하여 매체에 저장될 수 있다.

[0069] 본 발명의 실시예와 관련하여 설명된 방법 또는 알고리즘의 단계들은 하드웨어로 직접 구현되거나, 하드웨어에 의해 실행되는 소프트웨어 모듈로 구현되거나, 또는 이들의 결합에 의해 구현될 수 있다. 소프트웨어 모듈은 RAM(Random Access Memory), ROM(Read Only Memory), EPROM(Erasable Programmable ROM), EEPROM(Electrically Erasable Programmable ROM), 플래시 메모리(Flash Memory), 하드 디스크, 착탈형 디스크, CD-ROM, 또는 본 발명이 속하는 기술 분야에서 잘 알려진 임의의 형태의 컴퓨터 판독가능 기록매체에 상주할 수도 있다.

[0070] 이상, 첨부된 도면을 참조로 하여 본 발명의 실시예를 설명하였지만, 본 발명이 속하는 기술분야의 통상의 기술자는 본 발명이 그 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 그러므로, 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며, 제한적이지 않은 것으로 이해해야만 한다.

## 부호의 설명

[0071] 100 : 건강 개선 프로그램 제공 장치

110 : 데이터 획득부

120 : 데이터 처리부

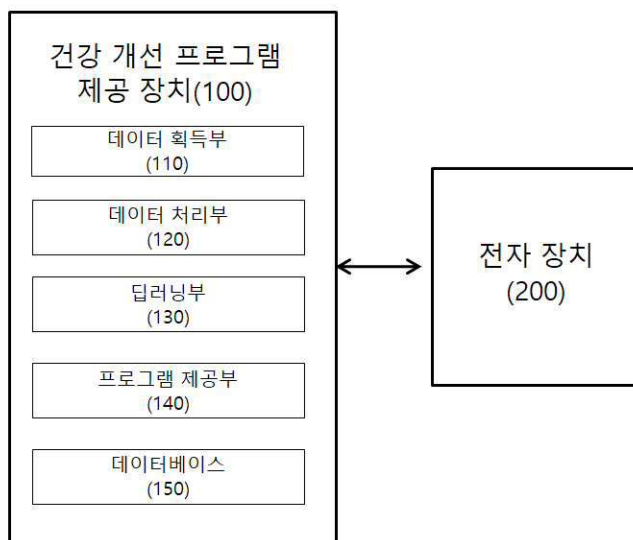
130 : 딥러닝부

140 : 프로그램 제공부

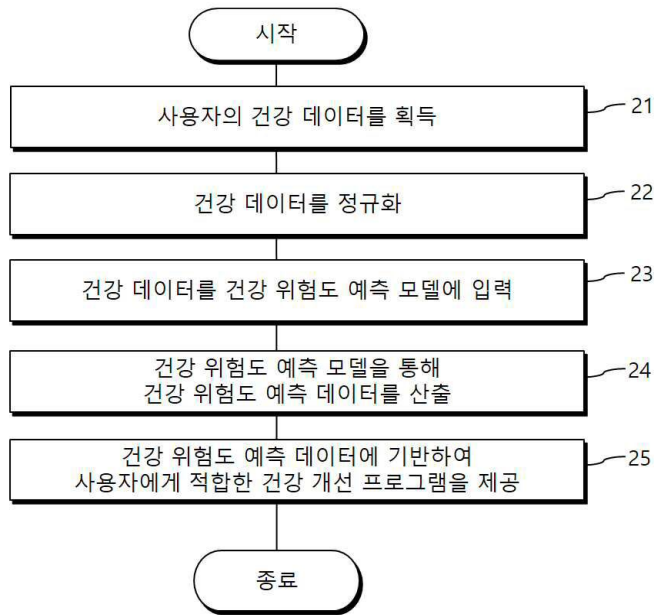
150 : 데이터베이스

## 도면

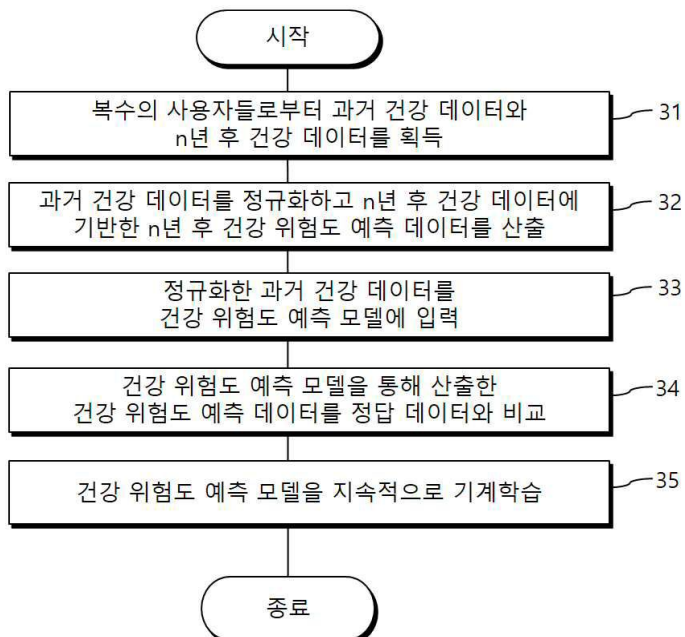
### 도면1



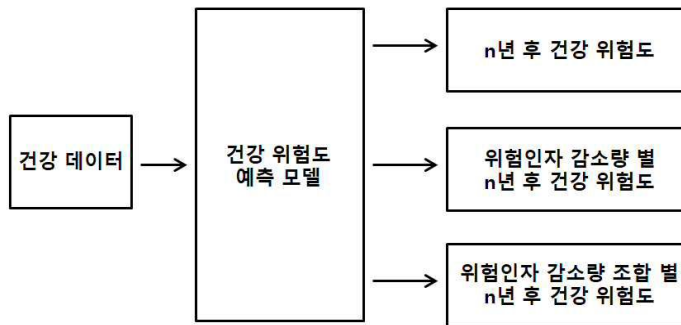
도면2



도면3



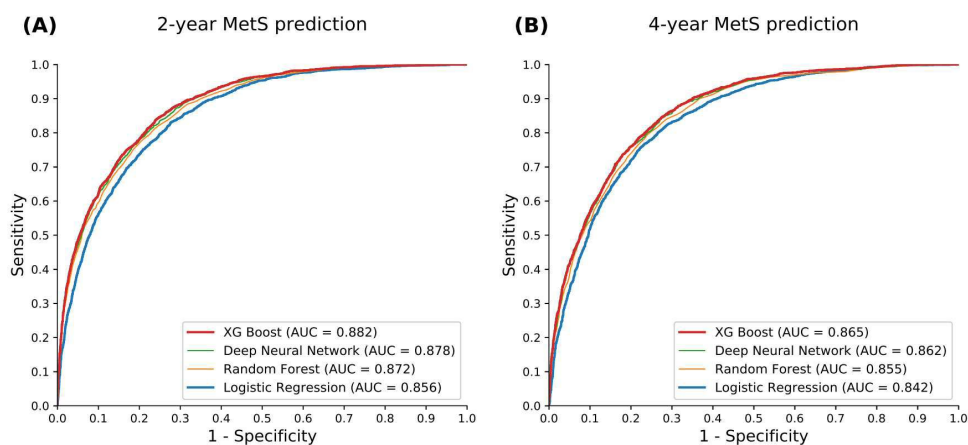
도면4



도면5

(A) Health Profile at Present				(B) Let's Loss Weight!
Age	32	Blood pressure		Based on given information, you could be diagnosed as metabolic syndrome. You have a probability of 90.95% for still having metabolic syndrome after 2 years if without any weight loss.
Sex	Male	- Systolic	140 mmHg	
Height	178.3 cm	- Diastolic	90 mmHg	
Weight	88.5 kg	Waist circumferences	115 cm	However, you can have 2.93% lower risk with 5 kg loss, 11.55% lower risk with 10 kg loss, 25.02% lower risk with 15 kg loss, 42.64% lower risk with 20 kg loss of your body weight!
Drinking	Yes	Fasting glucose	138 mg/dL	
- Amount	3 bottles / week	- Medication	No	
Smoking	Current	Triglyceride	328 mg/dL	
- Amount	1 pack / day	HDL	35 mg/dL	
- Duration	30 years	Total cholesterol	180 mg/dL	
		- Medication	Yes	

도면6



도면7

