(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0052457
(43) 공개일자 2025년04월18일

(51) 국제특허분류(Int. Cl.)
G16H 50/30 (2018.01) *A61B 5/00* (2021.01)
G16H 10/40 (2018.01) *G16H 10/60* (2018.01)
G16H 50/20 (2018.01) *G16H 50/70* (2018.01)
(52) CPC특허분류
G16H 50/30 (2018.01)
A61B 5/4088 (2013.01)
(21) 출원번호 10-2025-7010203
(22) 출원일자(국제) 2024년05월31일
심사청구일자 2025년03월27일
(85) 번역문제출일자 2025년03월27일
(86) 국제출원번호 PCT/KR2024/007480
(87) 국제공개번호 WO 2024/248536
국제공개일자 2024년12월05일
(30) 우선권주장
1020230069953 2023년05월31일 대한민국(KR)
1020230158640 2023년11월15일 대한민국(KR)

(71) 출원인
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
박영미
서울특별시 강남구 도곡로78길 22, 112동 1502호
김건하
서울특별시 서초구 동광로19길 51, 1105호
(뒷면에 계속)
(74) 대리인
특허법인더웨이브

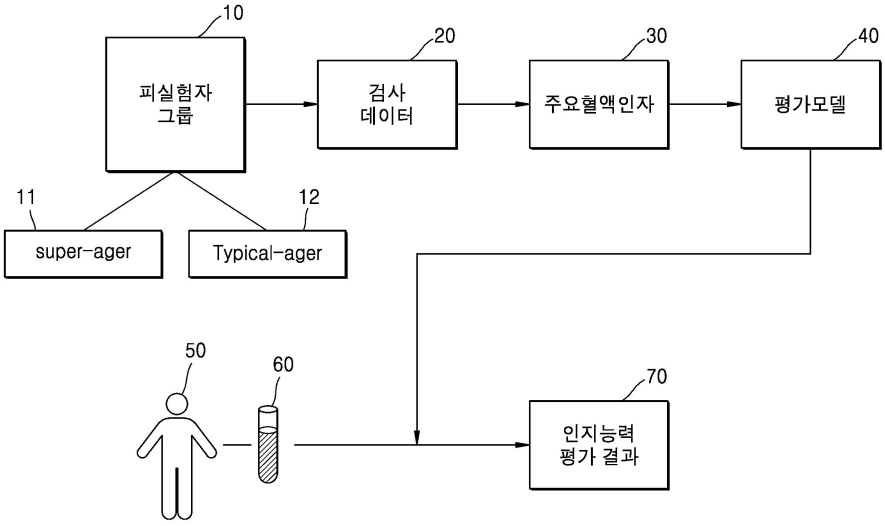
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 **혈액 기반 인지능력 평가 방법 및 장치**

(57) 요 약

본 발명의 일 실시예에 의하면, 피실험자들의 실제 혈액검사 결과 및 인지기능검사 결과를 포함하는 실제 검사 데이터를 획득하는 단계; 상기 실제 검사 데이터에 기초하여, 상기 피실험자들이 제1 그룹에 속하는지 여부와 관련된 인지기능검사의 주요 인지기능 항목을 결정하는 단계; 획득된 상기 실제 검사 데이터에 기초하여 증강 검사 데이터를 생성하는 단계; 상기 증강 검사 데이터의 개별 데이터들에 대하여, 분류(classification) 알고리즘을 이용하여 상기 제1 그룹에 속하는지 여부를 예측하는 단계; 상기 실제 검사 데이터 및 상기 증강 검사 데이터의 개별 데이터들을 제1 그룹 및 제2 그룹 중 어느 하나의 그룹으로 분류하는 단계; 및 분류된 상기 실제 검사 데이터 및 증강 검사 데이터에 기초하여 주요 인지기능 항목별로 상기 제1 그룹에 연관된 주요 혈액인자를 하나 이상 추출하는 단계;를 포함하는, 혈액 기반 인지능력 평가 방법이 제공된다.

대 표 도



(52) CPC특허분류
G16H 10/40 (2021.08)
G16H 10/60 (2021.08)
G16H 50/20 (2018.01)
G16H 50/70 (2018.01)
(72) 발명자
최장환
서울특별시 마포구 마포대로 53, C동 302호
이효빈
서울특별시 서초구 방배선행길 2, 101동 902호

명 세 서

청구범위

청구항 1

피실험자들의 실제 혈액검사 결과 및 인지기능검사 결과를 포함하는 실제 검사 데이터를 획득하는 단계;

상기 실제 검사 데이터에 기초하여, 상기 피실험자들이 제1 그룹에 속하는지 여부와 관련된 인지기능검사의 주요 인지기능 항목을 결정하는 단계;

획득된 상기 실제 검사 데이터에 기초하여 증강 검사 데이터를 생성하는 단계;

상기 증강 검사 데이터의 개별 데이터들에 대하여, 분류(classification) 알고리즘을 이용하여 상기 제1 그룹에 속하는지 여부를 예측하는 단계;

상기 실제 검사 데이터 및 상기 증강 검사 데이터의 개별 데이터들을 상기 제1 그룹 및 제2 그룹 중 어느 하나의 그룹으로 분류하는 단계; 및

분류된 상기 실제 검사 데이터 및 증강 검사 데이터에 기초하여 주요 인지기능 항목별로 상기 제1 그룹에 연관된 주요 혈액인자를 하나 이상 추출하는 단계;

를 포함하는, 혈액 기반 인지능력 평가 방법.

청구항 2

제 1 항에 있어서,

상기 인지기능검사의 주요 항목을 결정하는 단계는,

상기 피실험자들 각각을 제1 그룹 또는 제2 그룹으로 분류하는 단계;

상기 피실험자들을 제1 또는 제2 그룹으로 분류한 것과 상기 실제 실험 데이터에 기초하여, 상기 인지기능검사의 각 항목과 피실험자들이 제 1 그룹에 속하는지 여부 간의 연관관계를 학습하는 단계;

상기 학습에 기초하여 상기 주요 인지기능 항목을 결정하는 단계;

를 포함하는, 혈액 기반 인지능력 평가 방법.

청구항 3

제 1 항에 있어서,

상기 제1 그룹은 피실험자가 본인이 속한 연령그룹의 평균 인지능력 이상의 인지능력을 갖는 슈퍼 에이저(Super Ager)들의 그룹이고, 상기 제2 그룹은 슈퍼 에이저가 아닌 일반 고령자(Typical Ager)들의 그룹인, 혈액 기반 인지능력 평가 방법.

청구항 4

제 1 항에 있어서,

상기 증강 검사 데이터를 생성하는 단계는, 기설정된 데이터 생성 알고리즘을 이용하여 가상의 혈액검사 결과 및 인지기능검사 결과를 포함하는 증강 검사 데이터를 생성하는, 혈액 기반 인지능력 평가 방법.

청구항 5

제 1 항에 있어서,

상기 제1 그룹에 속하는지 여부를 예측하는 단계는, 상기 증강 검사 데이터에 대하여 적어도 2개 이상의 서로 다른 분류 알고리즘을 사용하여 1차 예측 결과를 생성하고, 상기 1차 예측 결과를 적층(stacking)하여 최종 예측 결과를 생성하는, 혈액 기반 인지능력 평가 방법.

청구항 6

제 5 항에 있어서,

상기 적어도 2개 이상의 서로 다른 분류 알고리즘에 입력하기 전, 상기 증강 검사 데이터를 라벨 인코더(LabelEncoder) 및 심플 임퓨터(SimpleImputer)를 이용하여 전처리하는, 혈액 기반 인지능력 평가 방법.

청구항 7

제 1 항에 있어서,

결정된 상기 주요 인지기능 항목은 언어기억(language memory), 시각기억(visual memory) 및 시공간(visuospatial) 항목인, 혈액 기반 인지능력 평가 방법.

청구항 8

제 1 항에 있어서,

상기 주요 인지기능 항목별로 추출된 상기 주요 혈액인자들을 기초로 인지능력을 평가하는 평가 모델을 생성하는 단계;를 추가적으로 포함하는, 혈액 기반 인지능력 평가 방법.

청구항 9

제 1 항의 방법을 컴퓨터에서 실행하기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

청구항 10

혈액 기반 인지능력 평가 장치로서,

상기 장치는 프로세서; 통신부; 및 메모리를 포함하고,

상기 프로세서는,

피실험자들의 실제 혈액검사 결과 및 인지기능검사 결과를 포함하는 실제 검사 데이터를 획득하고,

상기 실제 검사 데이터에 기초하여, 상기 피실험자들이 제1 그룹에 속하는지 여부와 관련된 인지기능검사의 주요 인지기능 항목을 결정하며,

획득된 상기 실제 검사 데이터에 기초하여 증강 검사 데이터를 생성하고,

상기 증강 검사 데이터의 개별 데이터들에 대하여, 분류(classification) 알고리즘을 이용하여 상기 제1 그룹에 속하는지 여부를 예측하며,

상기 실제 검사 데이터 및 상기 증강 검사 데이터의 개별 데이터들을 제1 그룹 및 제2 그룹 중 어느 하나의 그룹으로 분류하고,

분류된 상기 실제 검사 데이터 및 증강 검사 데이터에 긱초하여 주요 인지기능 항목별로 상기 제1 그룹에 연관된 주요 혈액인자를 하나 이상 추출하는, 혈액 기반 인지능력 평가 방법.

발명의 설명

기 술 분 야

[0001] 본 발명의 실시예들은 혈액 기반 인지능력 평가 방법 및 장치에 관한 것이다.

[0002] 본 출원은 2024년도 정부(보건복지부 및 과학기술정보통신부)의 재원으로 한국보건산업진흥원(No. HU20C0271)과 정보통신기획평가원(No. RS-2022-00155966, 인공지능융합혁신인재양성(이화여자대학교), 과제고유번호:1711179344)의 지원을 받아 수행된 연구 결과이다.

배 경 기 술

[0003] 노인 인구가 급격히 증가하면서 노인의 건강관리가 중요한 사회적 관심사로 떠오르고 있으며, 특히 뇌 건강과 관련된 노년기 인지기능 변화에 대한 연구가 활발히 이루어지고 있다. 노년기 인지기능은 연령의 증가에 따라

기능의 손상 혹은 저하가 일반적이라고 하더라도 이러한 변화는 고정된 것이 아니라 여러가지 요인에 의해서 다르게 나타날 수 있는 개인적 속성이다.

[0004] 인지기능을 측정하는 인지기능 테스트는 주로 경도 인지 장애나 치매와 같은 질환을 판별하는 수단으로 활용되며, 서류를 이용하거나 서류를 기반으로 한 컴퓨터 프로그램으로 수행되고 있다. 그러나 이와 같은 서류기반의 인지기능 테스트는 장시간이 소요되어 노년층에게 부담이 되며, 문자를 읽는 것이 불편한 노년층의 경우 활용하기 어려운 문제점이 있다. 또한, 동일한 테스트를 다수 회 수행할 수 없어 인지기능 변화에 대한 획일적인 추적 관찰이 불가능한 문제점이 있었다.

[0005] 한편, 기존의 인지기능 관련 연구는 주로 알츠하이머, 파킨슨, 당뇨, 고혈압 등 질환이 있는 환자를 대상으로 진행되어 왔으나, 최근 해당 연령에 비하여 강인한 인지능력을 갖는 노인을 의미하는 슈퍼에이저(Super-ager)와 관련된 연구들이 주목받고 있다. 슈퍼에이저는 나이에 비해 월등하게 우수한 인지능력을 가짐으로 인해 신체적, 정신적으로 더욱 건강한 노후 생활을 영위한다. 따라서, 이러한 슈퍼에이저는 노인의 인지기능을 평가하는 하나의 척도가 될 수 있다.

[0006] 이에 따라, 질환을 판별하기 위한 것이 아닌 건강의 이상이 없는 대상자가 활용할 수 있는 인지능력 측정 기술 개발에 대한 요구가 높아지고 있으며, 본인의 나이에 비해 강한 인지능력을 갖는지 여부를 평가하는 인지능력 평가 기술의 개발이 필요한 실정이다.

발명의 내용

해결하려는 과제

[0007] 본 발명의 실시예들은 상기한 문제 및/또는 한계를 해결하기 위한 것으로, 혈액검사를 기반으로 인지능력을 평가할 수 있는 인지능력 평가 방법을 제공하는 것을 목적으로 한다.

[0008] 또한, 상기 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공하는 것을 목적으로 한다.

[0009] 그러나 이러한 과제는 예시적인 것으로, 이에 의해 본 발명의 범위가 한정되는 것은 아니다.

과제의 해결 수단

[0010] 본 발명의 일 실시예에 따르면, 피실험자들의 실제 혈액검사 결과 및 인지기능검사 결과를 포함하는 실제 검사 데이터를 획득하는 단계; 상기 실제 검사 데이터에 기초하여, 상기 피실험자들이 제1 그룹에 속하는지 여부와 관련된 인지기능검사의 주요 인지기능 항목을 결정하는 단계; 획득된 상기 실제 검사 데이터에 기초하여 증강 검사 데이터를 생성하는 단계; 상기 증강 검사 데이터의 개별 데이터들에 대하여, 분류(classification) 알고리즘을 이용하여 상기 제1 그룹에 속하는지 여부를 예측하는 단계; 상기 실제 검사 데이터 및 상기 증강 검사 데이터의 개별 데이터들을 상기 제1 그룹 및 제2 그룹 중 어느 하나의 그룹으로 분류하는 단계; 및 분류된 상기 실제 검사 데이터 및 증강 검사 데이터에 기초하여 주요 인지기능 항목별로 상기 제1 그룹에 연관된 주요 혈액 인자를 하나 이상 추출하는 단계; 를 포함할 수 있다.

[0011] 본 발명에 있어서, 상기 인지기능검사의 주요 항목을 결정하는 단계는, 상기 피실험자들 각각을 제1 그룹 또는 제2 그룹으로 분류하는 단계; 상기 피실험자들을 제1 또는 제2 그룹으로 분류한 것과 상기 실제 실험 데이터에 기초하여, 상기 인지기능검사의 각 항목과 피실험자들이 제 1 그룹에 속하는지 여부 간의 연관관계를 학습하는 단계; 상기 학습에 기초하여 상기 주요 인지기능 항목을 결정하는 단계; 를 포함할 수 있다.

[0012] 본 발명에 있어서, 상기 제1 그룹은 피실험자가 본인이 속한 연령그룹의 평균 인지능력 이상의 인지능력을 갖는 슈퍼 에이저(Super Ager)들의 그룹이고, 상기 제2 그룹은 슈퍼 에이저가 아닌 일반 고령자(Typical Ager)들의 그룹일 수 있다.

[0013] 본 발명에 있어서, 상기 증강 검사 데이터를 생성하는 단계는, 기설정된 데이터 생성 알고리즘을 이용하여 가상의 혈액검사 결과 및 인지기능검사 결과를 포함하는 증강 검사 데이터를 생성할 수 있다.

[0014] 본 발명에 있어서, 상기 제1 그룹에 속하는지 여부를 예측하는 단계는, 상기 증강 검사 데이터에 대하여 적어도 2개 이상의 서로 다른 분류 알고리즘을 사용하여 1차 예측 결과를 생성하고, 상기 1차 예측 결과를 적층(stacking)하여 최종 예측 결과를 생성할 수 있다.

[0015] 본 발명에 있어서, 상기 적어도 2개 이상의 서로 다른 분류 알고리즘에 입력하기 전, 상기 증강 검사 데이터를 라벨 인코더(LabelEncoder) 및 심플 임퓨터(SimpleImputer)를 이용하여 전처리할 수 있다.

[0016] 본 발명에 있어서, 결정된 상기 주요 인지기능 항목은 언어기억(language memory), 시각기억(visual memory) 및 시공간(visuospatial) 항목인, 혈액 기반 인지능력 평가 방법.

[0017] 본 발명에 있어서, 상기 주요 인지기능 항목별로 추출된 상기 주요 혈액인자들을 기초로 인지능력을 평가하는 평가 모델을 생성하는 단계;를 추가적으로 포함할 수 있다.

[0018] 본 발명의 방법을 컴퓨터에서 실행하기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

[0019] 본 발명의 다른 일 실시예에 따르면, 혈액 기반 인지능력 평가 장치로서, 상기 장치는 프로세서; 통신부; 및 메모리를 포함하고, 상기 프로세서는, 피실험자들의 실제 혈액검사 결과 및 인지기능검사 결과를 포함하는 실제 검사 데이터를 획득하고, 상기 실제 검사 데이터에 기초하여, 상기 피실험자들이 제1 그룹에 속하는지 여부와 관련된 인지기능검사의 주요 인지기능 항목을 결정하며, 획득된 상기 실제 검사 데이터에 기초하여 증강 검사 데이터를 생성하고, 상기 증강 검사 데이터의 개별 데이터들에 대하여, 분류(classification) 알고리즘을 이용하여 상기 제1 그룹에 속하는지 여부를 예측하며, 상기 실제 검사 데이터 및 상기 증강 검사 데이터의 개별 데이터들을 제1 그룹 및 제2 그룹 중 어느 하나의 그룹으로 분류하고, 분류된 상기 실제 검사 데이터 및 증강 검사 데이터에 각초하여 주요 인지기능 항목별로 상기 제1 그룹에 연관된 주요 혈액인자를 하나 이상 추출할 수 있다.

발명의 효과

[0020] 전술한 본 발명의 과제 해결 수단에 의하면, 본 발명의 실시예들은 혈액검사를 기반으로 대상자의 인지능력을 평가할 수 있으며, 인지능력 평가에 대한 접근성을 높일 수 있다. 또한, 대상자의 비용적 부담을 감소시킬 수 있으며 인지능력 검사에 따른 시간을 절약할 수 있다.

[0021] 또한, 본 발명의 실시예들은 혈액검사를 기반으로 인지능력을 평가함으로써 신뢰도가 높은 평가 결과를 제공할 수 있으며, 시간에 따른 추적관찰이 가능하도록 하여 지속적인 건강관리에 도움을 줄 수 있다.

[0022] 물론 이러한 효과에 의해 본 발명의 범위가 한정되는 것은 아니다.

도면의 간단한 설명

[0023] 도 1은 본 발명의 일 실시예에 따른 인지능력 평가 시스템을 설명하기 위한 도면이다.

도 2는 도 1의 검사 데이터를 예시한 도면이다.

도 3은 본 발명의 일 실시예에 따른 인지능력 평가 장치를 개략적으로 도시한 블록도이다.

도 4는 본 발명의 일 실시예에 따른 인지능력 평가 방법을 설명하기 위한 흐름도이다.

도 5(a)는 주의력 영역의 주요혈액인자를 나타내는 그래프이다.

도 5(b)는 언어 영역의 주요혈액인자를 나타내는 그래프이다.

도 5(c)는 언어기억 영역의 주요혈액인자를 나타내는 그래프이다.

도 5(d)는 시각기억 영역의 주요혈액인자를 나타내는 그래프이다.

도 5(e)는 시공간 영역의 주요혈액인자를 나타내는 그래프이다.

도 5(f)는 전두엽 영역의 주요혈액인자를 나타내는 그래프이다.

도 6 은 본 발명의 다른 일 실시예에 따른 인지능력 평가 방법을 설명하기 위한 흐름도이다.

도 7 은 본 발명의 일 실시예에 따른 인지기능검사의 각 항목별 슈퍼에이저 그룹과의 연관관계를 나타낸 그래프이다.

도 8 은 본 발명의 실시예에 따른 분류 알고리즘에 의한 예측 모델의 정확도를 나타낸 그래프이다.

도 9(a) 내지 도 9(e)는 각 인지기능검사 항목별로 결정된 제1 그룹에 연관된 혈액인자를 설명하기 위한 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0024] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 설명되는 실시 예들을 참조하면 명확해질 것이다. 그러나 본 발명은 아래에서 제시되는 실시 예들로 한정되는 것이 아니라, 서로 다른 다양한 형태로 구현될 수 있고, 본 발명의 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0025] 본 출원에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0026] 또한, 도면에 도시된 구성 요소들 간의 연결 선 또는 연결 부재들은 기능적인 연결 및/또는 물리적 또는 회로적 연결들을 예시적으로 나타낸 것일 뿐이다. 실제 장치에서는 대체 가능하거나 추가된 다양한 기능적인 연결, 물리적인 연결, 또는 회로 연결들에 의해 구성 요소들 간의 연결이 나타내어질 수 있다.

[0027] 또한, 본 개시의 일부 실시예들는 다양한 변환을 가할 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 본 실시예들의 효과 및 특징, 그리고 그것들을 달성하는 방법은 도면과 함께 상세하게 후술되어 있는 내용들을 참조하면 명확해질 것이다. 그러나 본 실시예들은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 다양한 형태로 구현될 수 있다.

[0028] 이하의 실시예에서 제1, 제2 등의 용어는 한정적인 의미가 아니라 하나의 구성 요소를 다른 구성 요소와 구별하는 목적으로 사용되었다. 또한, 이하의 실시예에서 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 또한, 이하의 실시예에서 포함하다 또는 가지다 등의 용어는 명세서상에 기재된 특징, 또는 구성요소가 존재함을 의미하는 것이고, 하나 이상의 다른 특징들 또는 구성요소가 부가될 가능성을 미리 배제하는 것은 아니다. 또한, 이하의 실시예에서 어떤 부분이 다른 부분과 “연결”되어 있다고 할 때, 이는 “직접적으로 연결”되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 “전기적으로 연결”되어 있는 경우도 포함한다.

[0029] 이하, 첨부된 도면을 참조하여 이하의 실시예들을 상세히 설명하기로 하며, 도면을 참조하여 설명할 때 동일하거나 대응하는 구성 요소는 동일한 도면부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다.

[0030] 도 1은 본 발명의 일 실시예에 따른 인지능력 평가 시스템을 설명하기 위한 도면이며, 도 2는 도 1의 검사 데이터(20)를 예시한 도면이다.

[0031] 인지능력 평가 서버(100, 도 3 참조)는 피실험자 그룹(10)으로부터 각 피실험자들의 혈액검사 결과 및 인지기능 검사 결과를 포함하는 검사 데이터(20)를 획득할 수 있다.

[0032] 여기서, 피실험자는 혈액검사 및 인지기능검사를 수행하고 검사 결과를 제공한 사람을 의미한다. 이러한 피실험자들로 이루어진 피실험자 그룹은 복수의 연령그룹으로 분류될 수 있으며, 각 피실험자는 본인이 속한 연령그룹의 평균 인지능력 이상의 인지능력을 갖는 사람과 그렇지 않은 사람으로 분류될 수 있다.

[0033] 즉, 노인으로 이루어진 피실험자 그룹은 슈퍼에이저(super-ager) 그룹과 일반적인 고령자(Typical-ager) 그룹으로 분류될 수 있다. 본 명세서에서 “슈퍼에이저”는 노인이지만 중년의 인지능력을 갖는 사람을 의미한다. 노인(elder)은 60세 이상의 연령을 갖는 사람으로 정의한다. 물론, 노인을 정의하는 연령대는 다른 나이 범위로 설정될 수도 있다. 중년(middle age)은 45세 연령으로 정의한다. 물론, 중년을 정의하는 연령대는 다른 나이 범위로 설정될 수도 있다. 대상자의 인지 능력은 임상적인 평가를 통하여 결정될 수 있다.

[0034] 슈퍼에이저는 인지 예비능(cognitive reserve) 및/또는 뇌 회복력(brain resilience)이 좋은 사람이라고 할 수 있다. 인지 예비능은 뇌 가소성(brain plasticity)에 기반한 개념으로, 노화에 따른 뇌의 변화를 능동적으로 대처한다는 의미를 지닌다. 인지 예비능을 결정하는 뚜렷한 인자는 밝혀지지 않았고, 대상자의 교육 수준과 같은 환경 요인이 인지 예비능에 영향을 미친다고 알려졌다. 뇌 회복력은 신경 세포가 손상된 경우 신경 세포를 복구하는 회복 능력을 의미한다.

[0035] 검사 데이터(20)는 혈액검사 데이터(21), 인지검사 데이터(22) 및 추가 데이터(23)를 포함할 수 있다. 혈액검사 데이터(21)는 피실험자의 혈액검사 결과로, 일반적인 건강검진에서 수행하는 혈액검사로부터 획득되는 것일 수

있다. 예를 들어, 혈액검사 데이터(21)는 혈색소 검사, 공복혈당검사, 간기능검사, 혈청 크레아티닌 검사, 콜레스테롤 검사, 신사구체여과율 검사 항목의 결과값을 포함할 수 있으며, HDL-Cholesterol, Cl, C-peptide, PSK9 등의 인자의 수치를 포함할 수 있다. 다만, 본 발명은 이에 한정되지 않으며, 혈액검사 데이터(21)는 다양한 혈액 검사법에 의한 결과값을 포함할 수 있다.

[0036] 인지검사 데이터(22)는 피실험자의 인지기능검사 결과로, 복수의 검사항목에 대한 결과를 포함할 수 있다. 예를 들어, 인지검사 데이터(22)는 주의력(attention), 언어(language), 언어기억(verbal memory), 시각기억(visual memory), 시공간(visuospatial), 전두엽(frontal)에 대한 검사 결과를 포함할 수 있다.

[0037] 검사 데이터(20)는 혈액검사 데이터(21) 및 인지검사 데이터(22) 외에 추가 데이터(23)로 연령, 성별, 교육 수준 등의 신상 데이터를 포함할 수 있다.

[0038] 인지능력 평가 장치(100)는 피실험자 그룹(10)의 검사 데이터(20)로부터 주요혈액인자(30)를 추출할 수 있다. 구체적으로, 인지능력 평가 장치(100)는 혈액검사 데이터(21)와 인지검사 데이터(22)의 상관관계 분석을 수행하여 혈액검사 데이터(21)에 포함된 복수개의 혈액인자들 중 인지능력에 영향을 미치는 주요혈액인자(30)를 추출할 수 있다.

[0039] 추출된 주요혈액인자(30)는 대상자(50)의 인지능력을 평가하는 평가모델(40)의 생성에 이용될 수 있다. 즉, 인지능력 평가 장치(100)는 혈액검사 결과에 포함된 주요혈액인자(30)들을 이용하여 인지능력을 평가하도록 하는 인지능력 평가모델(40)을 구축할 수 있다.

[0040] 인지능력 평가 장치(100)는 대상자(50)의 혈액 샘플(60)로부터 획득된 혈액 샘플 데이터로부터 주요혈액인자(30)들에 해당하는 값들로 이루어진 혈액정보들을 추출하고, 추출된 혈액정보들 및 평가모델(40)을 이용하여 대상자(50)의 인지능력을 평가할 수 있다. 인지능력 평가 장치(100)는 대상자(50)가 본인이 속한 연령그룹의 평균 인지능력 이상의 인지능력을 갖는지 여부를 평가할 수 있다. 또한, 인지능력 평가 장치(100)는 인지기능검사의 각 항목별로 대상자(50)의 인지능력을 평가할 수 있다.

[0041] 이하에서, 도 3 내지 도 10을 참조하여 인지능력 평가 장치(100)가 동작하는 예를 설명한다.

[0042] 도 3은 본 발명의 일 실시예에 따른 인지능력 평가 장치(100)를 개략적으로 도시한 블록도이다.

[0043] 도 3을 참조하면, 본 발명의 일 실시예에 따른 인지능력 평가 장치(100)는 통신부(110), 프로세서(120), 메모리(130) 및 입출력 인터페이스(140)를 포함할 수 있다. 도 3에 도시된 인지능력 평가 장치(100)는 본 실시예의 특징이 흐려지는 것을 방지하기 위하여 본 실시예와 관련된 구성요소들만을 도시한 것이다. 따라서, 도 3에 도시된 구성요소들 외에 다른 범용적인 구성요소들이 더 포함될 수 있음을 본 실시예와 관련된 기술분야에서 통상의 지식을 가진 자라면 이해할 수 있을 것이다. 또한, 도 3에 도시된 통신부(110), 프로세서(120), 메모리(130) 및 입출력 인터페이스(140)는 독립된 장치로 구현될 수도 있음은 본 실시예와 관련된 기술분야에서 통상의 지식을 가진 자에게 자명하다.

[0044] 본 발명의 일 실시예에 따른 인지능력 평가 장치(100)는 피실험자의 검사 데이터로부터 인지능력 평가에 사용되는 주요혈액인자들을 추출할 수 있으며, 이를 기초로 인지능력을 평가하는 평가모델을 생성할 수 있고, 대상자의 혈액샘플을 이용하여 대상자의 인지능력을 평가할 수 있다.

[0045] 사용자는 인지능력 평가 장치(100)를 통해 대상자의 인지능력 평가 결과를 확인할 수 있다. 여기서 사용자는 대상자의 인지능력 평가 결과의 확인을 원하는 사람이다. 예컨대, 사용자는 의사, 임상학자, 대상자 본인, 대상자의 관리자 등일 수 있다.

[0046] 일 실시예에서, 인지능력 평가 장치(100)는 사용자 단말에 탑재될 수 있다. 이때, 사용자 단말은 컴퓨터 장치로 구현되는 고정형 단말이거나 이동형 단말일 수 있으며, 무선 또는 유선 통신 방식을 이용하여 네트워크를 통해 다른 사용자 단말 및/또는 서버와 통신할 수 있다. 예를 들면, 사용자 단말은 스마트폰(smart phone), 휴대폰, 네비게이션, 컴퓨터, 노트북, 디지털방송용 단말, PDA(Personal Digital Assistants), PMP(Portable Multimedia Player), 태블릿 PC, 웨어러블 기기 등일 수 있으나 이에 한정되지 않는다.

[0047] 다른 실시예에서, 인지능력 평가 장치(100)는 서버에 구비될 수 있다. 이때, 서버는 복수의 사용자 단말들과 네트워크를 통해 통신하여 명령, 코드, 파일, 컨텐츠, 서비스 등을 제공하는 컴퓨터 장치 또는 복수의 컴퓨터 장치들로 구현될 수 있다. 일례로, 인지능력 평가 장치(100)는 두 개의 서버에 구비될 수 있으며, 하나의 서버를 통해 인지능력을 평가하는 평가모델의 학습이 이루어지고, 다른 하나의 서버에서는 학습된 알고리즘을 기초로

대상자의 혈액 샘플 데이터로부터 대상자의 인지능력을 평가할 수 있다.

[0048] 인지능력 평가 장치(100)가 서버에 구비되는 경우, 서버는 네트워크를 통해 접속한 사용자 단말로 어플리케이션의 설치를 위한 파일을 제공할 수 있다. 이 경우, 사용자 단말은 서버로부터 제공받은 파일을 이용하여 어플리케이션을 설치할 수 있다. 또한, 사용자 단말은 사용자 단말이 포함되는 운영체제(Operating system, OS) 및 적어도 하나의 프로그램(일례로 브라우저나 설치된 어플리케이션)의 제어에 따라 서버에 접속하여 서버가 제공하는 서비스나 컨텐츠를 제공받을 수 있다.

[0049] 통신부(110)는 네트워크를 통해 인지능력 평가 장치(100)가 외부 디바이스(미도시)와 서로 통신하기 위한 구성 또는 기능을 제공할 수 있다. 예를 들어, 프로세서(120)의 제어에 따라 제공되는 제어 신호, 명령, 데이터 등이 통신부(110)와 네트워크를 거쳐 외부 디바이스로 전송될 수 있다.

[0050] 통신부(110)는 외부의 디바이스로부터 피실험자들의 혈액검사 결과 및 인지기능검사 결과를 포함하는 검사 데이터를 수신할 수 있다. 또한, 통신부(110)는 외부의 디바이스로부터 대상자의 혈액 샘플 데이터 및 추가 정보를 수신할 수 있다. 또한, 통신부(110)는 인지능력 평가 결과를 사용자 단말과 같은 외부의 디바이스로 송신할 수 있다.

[0051] 프로세서(120)는 기본적인 산술, 로직 및 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 여기서, 명령은 메모리(130) 또는 통신부(110)에 의해 프로세서(120)로 제공될 수 있다. 예를 들어, 프로세서(120)는 메모리(130)와 같은 기록 장치에 저장된 프로그램 코드에 따라 수신되는 명령을 실행하도록 구성될 수 있다. 또한, 프로세서(120)는 인지능력 평가 장치(100)에 포함된 다른 구성요소들의 동작을 전반적으로 제어할 수 있다.

[0052] 프로세서(120)는 다수의 논리 게이트들의 어레이로 구현될 수도 있고, 범용적인 마이크로 프로세서와 이 마이크로 프로세서에서 실행될 수 있는 프로그램이 저장된 메모리의 조합으로 구현될 수도 있다. 예를 들어, 프로세서(120)는 범용 프로세서, 중앙 처리 장치(CPU), 마이크로프로세서, 디지털 신호 프로세서(DSP), 제어기, 마이크로제어기, 상태 머신 등을 포함할 수 한다. 일부 환경에서, 프로세서(120)는 주문형 반도체(ASIC), 프로그램 가능 로직 디바이스(PLD), 필드 프로그램 가능 게이트 어레이(FPGA) 등을 포함할 수도 있다. 예를 들어, 프로세서(120)는 디지털 신호 프로세서(DSP)와 마이크로프로세서의 조합, 복수의 마이크로프로세서들의 조합, 디지털 신호 프로세서(DSP) 코어와 결합된 하나 이상의 마이크로프로세서들의 조합, 또는 임의의 다른 그러한 구성들의 조합과 같은 처리 디바이스들의 조합을 지칭할 수도 있다.

[0053] 메모리(130)는 비-일시적인 임의의 컴퓨터 판독 가능한 기록매체를 포함할 수 있다. 일 예로서, 메모리(130)는 RAM(Random Access Memory), ROM(Read Only Memory), 디스크 드라이브, SSD(Solid State Drive), 플래시 메모리(Flash Memory) 등과 같은 비소멸성 대용량 저장 장치(Permanent Mass Storage Device)를 포함할 수 있다. 다른 예로서, ROM, SSD, 플래시 메모리, 디스크 드라이브 등과 같은 비소멸성 대용량 저장 장치는 메모리와는 구분되는 별도의 영구 저장 장치일 수 있다. 또한, 메모리(130)에는 운영체제(OS)와 적어도 하나의 프로그램 코드(예를 들어, 도 4를 참조하여 후술할 동작을 프로세서(120)가 수행하기 위한 코드)가 저장될 수 있다.

[0054] 입출력 인터페이스(140)는 인지능력 평가 장치(100)와 연결되거나 인지능력 평가 장치(100)에 포함될 수 있는 입력 및/또는 출력을 위한 장치(예를 들어, 키보드, 마우스, 터치스크린 등)와의 인터페이스를 위한 수단일 수 있다. 도 3에서는 입출력 인터페이스(140)가 프로세서(120)와 별도로 구성된 요소로 도시되었으나, 이에 한정되지 않으며, 입출력 인터페이스(140)가 프로세서(120)에 포함되도록 구성될 수도 있다.

[0055] 도 4는 본 발명의 일 실시예에 따른 인지능력 평가 방법을 설명하기 위한 흐름도이다.

[0056] 도 4를 참조하면, 본 발명의 일 실시예에 따른 인지능력 평가 방법은 피실험자들의 검사 데이터를 획득하는 단계(S100), 검사 데이터로부터 인지능력 평가에 사용되는 주요혈액인자들을 추출하는 단계(S200), 추출된 주요혈액인자들을 기초로 인지능력을 평가하는 평가모델을 생성하는 단계(S300) 및 대상자의 혈액 샘플로부터 주요혈액인자에 해당하는 정보들을 추출하고, 추출된 정보들 및 평가모델을 이용하여 대상자의 인지능력을 평가하는 단계(S400)를 포함한다. 이하에서, 연구자의 실제 실험 내용을 구체적으로 예시하여 본 발명의 기술적 사상을 설명하고자 한다.

[0057] 먼저, 본 발명의 일 실시예에 따른 인지능력 평가 방법은 피실험자들의 검사 데이터를 획득한다(S100).

[0058] 인지능력 평가 장치(100)는 통신부(110)를 통하여 외부 디바이스로부터 피실험자들의 검사 데이터를 수신하고, 수신된 검사 데이터를 프로세서(120)에 전달할 수 있다.

[0059] 검사 데이터는 혈액검사 데이터(21), 인지검사 데이터(22) 및 추가 데이터(23)를 포함할 수 있다. 혈액검사 데이터(21)는 피실험자의 혈액검사 결과로, 일반적인 건강검진에서 수행하는 혈액검사로부터 획득되는 것일 수 있다. 다만, 본 발명은 이에 한정되지 않으며, 혈액검사 데이터(21)는 다양한 혈액 검사법에 의한 결과값을 포함할 수 있다.

[0060] 인지검사 데이터(22)는 피실험자의 인지기능검사 결과로, 복수의 검사항목에 대한 결과를 포함할 수 있다. 예를 들어, 인지검사 데이터(22)는 주의력(attention), 언어(language), 언어기억(verbal memory), 시각기억(visual memory), 시공간(visuospatial), 전두엽(frontal)에 대한 검사 결과를 포함할 수 있다.

[0061] 검사 데이터(20)는 혈액검사 데이터(21) 및 인지검사 데이터(22) 외에 추가 데이터(23)로 연령, 성별, 교육 수준 등의 신상 데이터를 포함할 수 있으며, 체질량지수(BMI) 등의 신체검사 데이터를 포함할 수도 있다.

[0062] 구체적 일례로, 연구자가 구성한 피실험자 그룹은 60세 이상의 고령자 81명으로 구성되었으며, 모든 피실험자에 대하여 혈액검사 및 인지기능검사를 수행하였다.

[0063] 다음, 검사 데이터로부터 인지능력 평가에 사용되는 주요혈액인자들을 추출한다(S200).

[0064] 프로세서(120)는 검사 데이터에 포함되어 있는 혈액검사 결과와 인지기능검사 결과의 상관관계 분석을 수행하고, 인지능력 평가에 사용할 주요혈액인자들을 추출할 수 있다. 이때, 프로세서(120)는 상관관계 분석에 검사 데이터에 포함되어 있는 추가 데이터를 더 사용할 수도 있다.

[0065] 구체적으로, 프로세서(120)는 검사 데이터에 포함되어 있는 인지기능검사 결과를 이용하여 본인이 속한 연령그룹의 평균 인지능력 이상의 인지능력을 갖는 피실험자를 분류할 수 있다. 즉, 피실험자 그룹은 복수의 연령그룹으로 분류되며, 각 피실험자는 본인이 속한 연령그룹의 평균 인지능력 이상의 인지능력을 갖는 사람과 그렇지 않은 사람으로 분류될 수 있다.

[0066] 구체적인 일 실험예에서, 연구자가 구성한 피실험자 그룹은 인지기능검사 결과를 이용하여 본인이 속한 연령그룹의 평균 인지능력 이상의 인지능력을 가지는 41명의 슈퍼에이저(Super-ager) 그룹과 40명의 일반 고령자(Typical-ager) 그룹으로 분류되었다. 본 명세서에서는, 설명의 편의를 위하여 슈퍼에이저 그룹은 제1 그룹으로, 일반 고령자 그룹은 제2 그룹으로 정의할 수 있다.

[0067] 슈퍼에이저 그룹과 일반 고령자 그룹의 인지기능검사 결과에 대하여 통계 분석을 수행한 결과는 아래 표 1과 같다. 인지기능검사 결과에서는 언어(language), 언어기억(language memory), 시각기억(visual memory), 및 시공간(visuospatial)이 $p<0.05$로 두 그룹 간에 유의미한 차이를 보였다. 여기서 p는 유의 확률값(p-value)을 의미한다. 또한, 언어기억(language memory), 시각기억(visual memory), 시공간(visuospatial)이 $p<0.01$로 두 그룹 간에 현격한 유의미한 차이를 보였다.

[0068] [표 1]

Variable	TypicalAgers (n = 40)	SuperAgers (n = 41)	p*
Cognitive Domains			
Attention	0.05±0.92	0.45±1.04	0.074
Language	-0.09±0.90	0.33±0.79	0.028
VerbalMemory	0.28±1.03	1.05±0.72	<0.001
VisualMemory	-0.19±0.61	0.89±0.91	<0.001
Visuospatial	-0.19±0.76	0.30±0.63	0.002
Frontal	0.58±0.85	0.81±0.63	0.176

[0069]

[0070] 두 그룹의 혈액검사 결과에 대하여 통계 분석을 수행한 결과는 아래 표 2와 같다. 혈액검사 결과에서는 HDL, Cl, C-peptide, PCSK9이 $p<0.05$로 두 그룹 간에 유의미한 차이를 보였다.

[0071] [표 2]

Variable	Typical Agers (n = 40)	SuperAgers (n = 41)	p*	Variable	Typical Agers (n = 40)	SuperAgers (n = 41)	p*
Blood Analysis				Blood Analysis			
HbA1c	5.90±0.72	5.80±0.63	0.285	C-peptide	2.29±1.15	1.86±0.63	0.039
Total calcium	9.29±0.31	9.35±0.38	0.469	Folic acid	15.06±9.24	14.92±7.40	0.942
Phosphorus	3.70±0.43	3.64±0.37	0.552	Vit B12	868.80±552.97	862.37±300.70	0.948
Glucose	101.90±18.36	102.85±20.69	0.827	WBC	5.50±1.20	5.14±1.16	0.17
BUN	16.48±4.27	16.20±4.46	0.774	RBC	4.36±0.49	4.28±0.44	0.449
Creatinine	0.77±0.19	0.74±0.18	0.481	Hb	13.38±1.39	13.08±1.09	0.281
eGFR-MDRD (IDMS)	83.05±18.75	87.24±19.51	0.327	Hct	39.36±3.91	38.50±3.33	0.29
eGFR-CKD-EPI	80.68±13.93	83.12±11.23	0.386	MCV	90.49±2.88	90.17±3.36	0.643
Triglyceride	110.95±46.13	102.44±44.31	0.4	MCH	30.75±1.12	30.65±1.29	0.71
Total Cholesterol	164.38±36.12	177.63±42.98	0.137	MCHC	33.99±0.90	33.99±0.79	0.985
Total Protein	7.07±0.32	7.08±0.41	0.852	RDW(CV)	12.65±0.49	12.69±0.63	0.725
Albumin	4.48±0.21	4.54±0.22	0.152	Platelet	217.59±50.08	212.60±42.78	0.631
AST	25.95±15.58	24.39±6.88	0.56	ELISA result			
ALT	21.55±11.6	19.29±7.79	0.306	PCSK9 (ng/ml)	99.30±35.53	116.60±42.28	0.05
ALP	70.23±21.82	70.98±20.64	0.874	ApoE (ug/ml)	114.34±75.17	96.76±46.41	0.208
Total Bilirubin	0.61±0.22	0.64±0.26	0.67	CD36 (ng/ml)	105.43±103.10	152.22±169.83	0.138
HDL-Cholesterol	54.33±12.02	60.15±11.17	0.027	EPO (mIU/ml)	10.48±3.74	12.19±4.84	0.079
LDL-Cholesterol	95.33±32.80	102.83±37.52	0.341	AGE (ng/ml)	12.92±5.45	12.92±7.58	0.999
Na	141.45±2.19	141.24±1.68	0.637	RAGE (pg/ml)	82.31±26.39	84.31±28.54	0.744
K	4.48±0.46	4.40±0.35	0.351	Vimentin (ng/ml)	172.10±163.97	163.46±88.99	0.768
Cl	105.90±2.47	104.83±2.14	0.04	FABP-3 (pg/ml)	1218.80±811.65	1215.94±548.00	0.985
CRP	0.11±0.15	0.13±0.13	0.68	FABP-4 (ng/ml)	29.33±13.49	29.84±14.13	0.868
TSH	2.16±1.28	2.45±2.00	0.442	oxLDL (ng/ml)	51.77±5.17	51.66±3.68	0.913
Free T4	1.28±0.39	1.20±0.17	0.224	Adioponectin (ug/ml)	35.98±25.16	43.71±41.15	0.31
Insulin	9.21±6.51	7.63±4.92	0.219	Leptin (ng/ml)	7.80±4.86	8.41±6.60	0.683

[0072]

[0073] 또한, 피실험자 그룹의 인지기능검사 결과와 혈액검사 결과 간 상관관계 분석을 수행하여, 인지기능검사의 각 항목별로 유의미하게 도출된 특징(feature)들은 다음과 같다.

[0074] 주의력(attention)은 HDL Cholesterol(r=0.254, p=0.022), Insulin(r=-0.223, p=0.046)과 유의미한 상관관계를 보였다. 언어(language)는 MCHC(r=0.261, p=0.019)와 유의미한 상관관계를 보였다. 언어기억(verbal memory)는 C-peptide(r=-0.281, p=0.011), EPO(r=0.227, p=0.042), HDL Cholesterol(r=0.219, p=0.049)과 유의미한 상관관계를 보였다. 시각기억(visual memory)는 PCSK9(r=0.281, p=0.011), HDL Cholesterol(r=0.264, p=0.017), Free T4(r=-0.234, p=0.036)와 유의미한 상관관계를 보였다. 시공간(visuospatial)은 HDL Cholesterol(r=0.341, p=0.002), Triglyceride(r=-0.332, p=0.002), Free T4(r=-0.326, p=0.003), ApoE(r=-0.287, p=0.009), Insuline(r=-0.231, p=0.038), AST(r=-0.228, p=0.041)와 유의미한 상관관계를 보였다. 전두엽(frontal)은 oxLDL(r=-0.356, p=0.001), Hb(r=-0.288, p=0.009), Hct(r=-0.262, p=0.018), Vimentin(r=-0.246, p=0.027), K(r=0.237, p=0.034), RBC(r=-0.227, p=0.041), Creatinine(r=-0.226, p=0.042)과 유의미한 상관관계를 보였다. 여기서, r은 피어슨 상관 계수(pearson correlation coefficient)를 의미한다.

[0075] 프로세서(120)는 피실험자 그룹에서 분류된 슈퍼에이저 그룹과 일반 고령자 그룹의 검사 데이터를 이용하여 검사 데이터에 포함된 각 변수들 간에 상관관계 분석을 수행하고, 기 설정된 변수 선택 알고리즘에 기초하여 인지능력 평가에 사용할 주요혈액인자들을 추출(즉, 변수를 선택(feature selection))할 수 있다. 이때, 추출되는 주요혈액인자의 수는 10개일 수 있으나, 이에 한정되는 것은 아니다.

[0076] 머신러닝을 이용하여 평가모델을 생성함에 있어서 변수(features, 또는 variables, attributes)가 사용되는데, 많은 수의 변수 또는 부적절한 변수들이 사용되면 머신러닝 모델이 과적합(Overfitting)되거나 예측 정확도가 감소하는 문제가 발생한다. 이에, 머신러닝 모델이 높은 예측 정확도를 갖기 위해서는 적절한 변수들의 조합을 사용할 필요가 있다. 즉, 예측하고자 하는 예측변수와 가장 연관성이 높은 변수들을 선택하여 가능한 한 적은 수의 변수를 사용하면서 머신러닝 모델의 복잡도(complexity)를 낮출 수 있다.

[0077] 변수 선택 알고리즘은 예를 들어, selectKBest 알고리즘일 수 있으나, 이에 한정되는 것은 아니다.

[0078] 기 설정된 변수 선택 알고리즘을 이용하여 검사 데이터로부터 추출되는 주요혈액인자는 HbA1c, Total calcium, Phosphorus , Glucose, BUN, Creatinine, eGFRMDRD, eGFRCKDEPI, Triglyceride, Total cholesterol, Total Protein, Albumin, AST, ALT, ALP, Total bilirubin, HDLCholesterol, LDLCholesterol, Na, K, Cl, CRP, TSH, Free T4, Insulin, C-peptide, Folic acid, Vit B12, WBC, RBC, Hb, Hct, MCV, MCH, MCHC, RDW(CV), Platelet, PCSK9, ApoE, CD36, EPO, AGE, RAGE, Vimentin, FABP-3, FABP-4, oxLDL, Adioponectin, Leptin 중 적어도 하나를 포함할 수 있다.

[0079] 한편, 프로세서(120)는 기 설정된 변수 선택 알고리즘, 일 실시예에 따르면 selectKBest를 이용하여 인지기능검사의 항목별로 주요혈액인자를 각각 추출할 수 있다. 이하, 도 5(a) 내지 도 5(e)에서 각 인지기능검사의 항목별로 선택되는 주요혈액인자의 일 예를 설명한다.

[0080] 도 5(a)는 주의력 영역의 주요혈액인자를 나타내는 그래프이며, 도 5(b)는 언어 영역의 주요혈액인자를 나타내는 그래프이고, 도 5(c)는 언어기억 영역의 주요혈액인자를 나타내는 그래프이며, 도 5(d)는 시각기억 영역의 주요혈액인자를 나타내는 그래프이고, 도 5(e)는 시공간 영역의 주요혈액인자를 나타내는 그래프이며, 도 5(f)는 전두엽 영역의 주요혈액인자를 나타내는 그래프이다. 각각의 그래프는 각 인지기능검사의 항목에서 각각의 변수가 인지능력을 평가하는 데 기여하는 정도를 기준으로 중요도를 산출하고 중요도가 높은 상위 10개의 변수를 그래프로 나타낸 것이다.

[0081] 도 5(a)를 참조하면, 주의력 영역의 주요혈액인자는 Total calcium, Triglyceride, Total cholesterol, Total bilirubin, HDL-cholesterol, Sodium (Na), Thyroid-Stimulating Hormone (TSH), CD36, Erythropoietin (EPO), Leptin으로 추출되었다.

[0082] 도 5(b)를 참조하면, 언어 영역의 주요혈액인자는 Glucose, Insulin, C-peptide, Red Blood Cell (RBC), Hemoglobin (Hb), Mean Corpuscular Hemoglobin (MCH), Mean Corpuscular Hemoglobin Concentration (MCHC), Receptor for Advanced Glycation End Products (RAGE), Adiponectin, Leptin으로 추출되었다.

[0083] 도 5(c)를 참조하면, 언어기억 영역의 주요혈액인자는 Hemoglobin A1c (HbA1c), Glucose, Alanine Aminotransferase (ALT), HDL-cholesterol, Free-T4, C-peptide, Mean Corpuscular Volume (MCV), MCH, Platelet, EPO로 추출되었다.

[0084] 도 5(d)를 참조하면, 시각기억 영역의 주요혈액인자는 HbA1c, Total cholesterol, HDL-cholesterol, LDL-cholesterol, Free T4, White Blood Cell (WBC), RBC, Hematocrit (Hct), PCSK9, oxidized LDL (oxLDL)로 추출되었다.

[0085] 도 5(e)를 참조하면, 시공간 영역의 주요혈액인자는 HbA1c, Blood Urea Nitrogen (BUN), Triglyceride, Total cholesterol, Aspartate Aminotransferase (AST), HDL-cholesterol, Free T4, Insulin, C-peptide, Apolipoprotein E (ApoE)로 추출된다.

[0086] 도 5(f)를 참조하면, 전두엽 영역의 주요혈액인자는 Triglyceride, Total cholesterol, Potassium (K), Hemoglobin (Hb), Hematocrit (Hct), MCH, MCHC, CD36, Advanced Glycation End Products (AGE), oxLDL로 추출된다.

[0087] 도 5(a) 내지 도 5(e)는 각 인지기능검사 항목별로 추출되는 주요혈액인자의 일 예를 도시한 것이며, 본 발명이 이에 한정되는 것은 아니다.

[0088] 다음, 추출된 주요혈액인자들을 기초로 인지능력을 평가하는 평가모델을 생성한다(S300).

[0089] 프로세서(120)는 주요혈액인자들을 이용하여 평가모델을 학습할 수 있으며, 인지능력을 평가하는 평가모델을 생성할 수 있다. 평가모델은 대상자의 혈액 샘플 데이터를 입력 받아 대상자의 인지능력 평가 결과를 출력하도록 학습(train)될 수 있다.

[0090] 한편, 평가모델을 생성하는 단계(S300)에서 프로세서(120)는 피실험자 그룹으로부터 획득되는 검사 데이터(학습 데이터)의 양적 문제를 해결하기 위하여 실제 피실험자의 검사 데이터를 기초로 증강 데이터를 생성하여 평가모델을 학습할 수 있다. 본 명세서에서는, 데이터의 구분을 위하여 실제 피실험자들로부터 획득한 검사 데이터를 실제 검사 데이터로, 생성된 증강 데이터를 증강 검사 데이터로 지칭할 수 있다.

[0091] 일 실시예에 따르면, 프로세서(120)는 기설정된 증강 데이터 생성 알고리즘을 이용하여 증강 검사 데이터를 생성할 수 있으며, 증강 데이터 생성 알고리즘으로는 Large language model기반 tabular data augmentation인 Generation of REAlistic Tabular data(GReaT) 이 사용될 수 있으나, 이에 한정되는 것은 아니다. 구체적인 일 예로, 연구자는 LLM기반 알고리즘을 이용하여 729개의 증강 데이터를 생성하였고, 전체 810개의 데이터를 이용하여 인지기능 항목 각각의 평가모델을 학습할 수 있다.

[0092] 본 발명의 일 실시예에 따른 평가모델은 기계학습 기술과 인지과학에서, 생물학적 신경망의 구조에 기초하여 구현된 통계학적 학습 알고리즘 또는 그 알고리즘을 실행하는 기계학습모델일 수 있다. 예를 들어, 기계학습모델은 생물학적 신경망에서와 같이 시냅스의 결합으로 네트워크를 형성한 인공 뉴런인 노트(Node)들이 시냅스의 가중치를 반복적으로 조정하여, 특정 입력에 대응한 올바른 출력과 추론된 출력 사이의 오차가 감소되도록 학습함으로써, 문제 해결 능력을 가지는 모델을 나타낼 수 있다. 예를 들어, 기계학습모델은 다층의 노드들과 이들 사이의 연결로 구성된 다층 퍼셉트론(MLP: multilayer perceptron)으로 구현될 수 있다.

[0093] 본 실시예에 따른 평가모델은 MLP를 포함하는 다양한 인공신경망 모델 구조들 중의 하나를 이용하여 구현될 수 있다. 예를 들어, 평가모델은 외부로부터 입력 신호 또는 데이터를 수신하는 입력층, 입력 데이터에 대응한 출력 신호 또는 데이터를 출력하는 출력층, 입력층과 출력층 상이에 위치하며 입력층으로부터 신호를 받아 특성을 추출하여 출력층으로 전달하는 적어도 하나의 은닉층으로 구성될 수 있다. 출력층은 은닉층으로부터 신호 또는 데이터를 수신하여 외부로 출력한다.

[0094] 예를 들어, 평가모델은 선형 회귀(Linear Regression), 라쏘(Lasso), 결정 트리(Decision Tree), RF(Random Forest), SVM(Support Vector Machine), 그래디언 부스트(GradientBoost), 에이다 부스트(AdaBoost), XG부스트(XgBoost) 중 적어도 하나를 이용하여 생성될 수 있으나, 이에 한정되는 것은 아니다.

[0095] 한편, 상술한 바와 같이 평가모델을 생성하는 단계(S300)에서 검사 데이터를 기초로 생성된 증강 데이터를 이용하여 평가모델을 검증할 수 있다. 보다 상세히, 프로세서(120)는 피실험자 그룹으로부터 획득되는 검사 데이터에 기초하여 증강 데이터를 생성할 수 있고 일 예로 증강 데이터 생성 알고리즘으로 LLM 기반 GReaT model이 사용될 수 있다.

[0096] 구체적인 일예로, 연구자는 GreaT 알고리즘을 이용하여 729개의 증강 데이터를 생성하였고, 전체 810개의 데이터를 이용하여 각각의 학습 모델에 대한 10-fold 교차검증을 시행할 수 있다. 교차검증 결과, 앞서 증강한 729개의 증강 데이터로 훈련하여 81개의 기존 데이터로 검증하였다. 자동 머신러닝(Auto ML)으로 잘 알려진 Pycaret을 활용하여 hyperparameter tuning을 시행하였다. 평가함수로는 평균제곱오차(MSE), 루트평균제곱오차(RMSE), 평균절대오차(MAE)를 사용하였다.

[0097] 아래의 [표 3]은 각 인지영역 별 17가지 회귀분석 알고리즘(Linear, Least Angle, Lasso, Lasso Least Angle, Ridge, Bayesian Ridge, Orthogonal Matching Pursuit, Decision Tree, Extra Trees, Random Forest, Huber, Elastic Net, Passive Aggressive, K Neighbors, AdaBoost, Gradient Boost, Light Gradient Boost)에 대한 에러율을 나타내는 표이다. [표 3]을 참조하면, 증강 데이터는 MAE기준 평균 0.15의 에러율을 보이는 것을 알 수 있다.

[0098] [표 3]

[0099]

	Attention			Language			Verbal Memory		
Regression Model	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
Linear	0.0422	0.2053	0.1625	0.0265	0.1629	0.1397	0.0500	0.2237	0.1806
Least Angle	0.0422	0.2053	0.1625	0.0265	0.1629	0.1397	0.0500	0.2237	0.1806
Lasso	0.0437	0.2090	0.1667	0.0266	0.1630	0.1400	0.0479	0.2189	0.1770
Lasso Least Angle	0.0437	0.2090	0.1667	0.0266	0.1630	0.1400	0.0555	0.2357	0.1892
Ridge	0.0424	0.2059	0.1635	0.0266	0.1632	0.1398	0.0499	0.2234	0.1805
Bayesian Ridge	0.0429	0.2071	0.1650	0.0266	0.1631	0.1400	0.0485	0.2203	0.1782
Orthogonal Matching Pursuit	0.0436	0.2088	0.1663	0.0265	0.1626	0.1389	0.0488	0.2209	0.1783
Decision Tree	0.0450	0.2122	0.1678	0.0276	0.1661	0.1424	0.1017	0.3190	0.2560
Extra Trees	0.0450	0.2122	0.1678	0.0291	0.1704	0.1423	0.0502	0.2240	0.1768
Random Forest	0.0437	0.2091	0.1654	0.0287	0.1694	0.1467	0.0512	0.2263	0.1788
Huber	0.0428	0.2068	0.1621	0.0264	0.1624	0.1395	0.0496	0.2227	0.1799
Elastic Net	0.0437	0.2090	0.1667	0.0266	0.1630	0.1400	0.0479	0.2189	0.1770
Passive Aggressive	0.0565	0.2378	0.1902	0.0344	0.1856	0.1589	0.0616	0.2482	0.2049
K Neighbors	0.0433	0.2082	0.1678	0.0274	0.1655	0.1416	0.0541	0.2326	0.1867
AdaBoost	0.0431	0.2076	0.1676	0.0272	0.1650	0.1422	0.0498	0.2232	0.1789
Gradient Boosting	0.0437	0.2090	0.1679	0.0294	0.1716	0.1493	0.0534	0.2311	0.1870
Light Gradient Boosting	0.0437	0.2090	0.1667	0.0266	0.1630	0.1400	0.0585	0.2419	0.1913

	Visual Memory			Visuospatial			Frontal		
Regression Model	MSE	RMSE	MAE	MSE	RMSE	MAE	MSE	RMSE	MAE
Linear	0.0405	0.2013	0.1605	0.0361	0.1901	0.1472	0.0247	0.1570	0.1241
Least Angle	0.0397	0.1991	0.1606	0.0350	0.1871	0.1519	0.0247	0.1570	0.1241
Lasso	0.0390	0.1975	0.1569	0.0361	0.1900	0.1542	0.0246	0.1568	0.1231
Lasso Least Angle	0.0390	0.1975	0.1569	0.0343	0.1852	0.1483	0.0246	0.1568	0.1231
Ridge	0.0390	0.1975	0.1569	0.0343	0.1852	0.1483	0.0245	0.1565	0.1232
Bayesian Ridge	0.0395	0.1987	0.1596	0.0335	0.1831	0.1478	0.0244	0.1562	0.1226
Orthogonal Matching Pursuit	0.0398	0.1996	0.1615	0.0350	0.1871	0.1519	0.0245	0.1565	0.1222
Decision Tree	0.0403	0.2007	0.1635	0.0363	0.1905	0.1547	0.0245	0.1565	0.1230
Extra Trees	0.0403	0.2007	0.1635	0.0363	0.1905	0.1547	0.0245	0.1565	0.1230
Random Forest	0.0398	0.1995	0.1565	0.0348	0.1865	0.1403	0.0245	0.1565	0.1230
Huber	0.0398	0.1995	0.1565	0.0348	0.1866	0.1403	0.0243	0.1558	0.1235
Elastic Net	0.0390	0.1976	0.1594	0.0346	0.1860	0.1474	0.0246	0.1568	0.1231
Passive Aggressive	0.0390	0.1974	0.1569	0.0345	0.1857	0.1487	0.0259	0.1609	0.1283
K Neighbors	0.0408	0.2020	0.1615	0.0351	0.1874	0.1506	0.0254	0.1594	0.1260
AdaBoost	0.0390	0.1975	0.1569	0.0343	0.1851	0.1477	0.0242	0.1556	0.1219
Gradient Boosting	0.0398	0.1995	0.1565	0.0348	0.1865	0.1403	0.0244	0.1563	0.1225
Light Gradient Boosting	0.0398	0.1996	0.1641	0.0497	0.2230	0.1923	0.0246	0.1568	0.1231

[0100] 또한, 프로세서(120)는 평가모델을 구축하기 위하여 인지능력의 각 항목별 점수를 예측하는 것에 있어 혈액인자들이 미치는 영향을 측정할 수 있다. 보다 상세히, 인지능력의 각 항목별로 SHapley Additive explanation(SHAP)을 이용하여 앞서 도출한 항목별 10개의 혈액인자가 미치는 영향력을 평가할 수 있다. 구체적인 일 실시예에서, 프로세서(120)는 인지능력 항목별로 10개의 인자 중에서도 더 많은 영향력을 미치는 혈액인자를 도출할 수 있다. 또한, 프로세서(120)는 각 혈액인자의 수치가 높은 정도 혹은 낮은 정도에 따라 인지능력

점수의 상승에 영향을 미치는지 여부를 도출할 수 있다. 예를 들어, 프로세서(120)는 SHAP을 이용하여 주의력 항목에서 10개의 혈액인자들 중 Leptin이 가장 큰 영향을 미치고, Leptin 수치가 낮을수록 주의력 점수가 높게 나온다고 판단할 수 있다.

[0101] 다음, 대상자로부터 획득된 혈액 샘플 데이터로부터 주요혈액인자들에 해당하는 혈액정보들을 추출하고 추출된 혈액정보들 및 평가모델을 이용하여 대상자의 인지능력을 평가한다(S400).

[0102] 프로세서(120)는 평가모델을 이용하여 대상자의 인지능력을 평가할 수 있다. 또한, 프로세서(120)는 평가모델을 이용하여 대상자의 인지능력을 인지기능검사 항목별로 평가할 수도 있다. 또한, 프로세서(120)는 평가모델을 이용하여 대상자가 본인이 속한 연령그룹의 평균 인지능력 이상의 인지능력을 갖는지 여부를 평가할 수 있다.

[0103] 구체적으로 프로세서(120)는 대상자의 혈액 샘플의 혈액검사 결과로부터 생성되는 혈액 샘플 데이터를 획득할 수 있다. 프로세서(120)는 혈액 샘플 데이터로부터 주요혈액인자들에 해당하는 값을 추출하여 주요혈액인자 값들로 이루어진 혈액정보를 추출할 수 있다.

[0104] 프로세서(120)는 추출된 혈액정보를 평가모델에 입력하여 대상자의 인지능력을 평가 결과를 출력할 수 있다. 프로세서(120)에서 출력되는 인지능력 평가 결과는 각 인지기능검사 항목에 대한 점수일 수 있으나, 이에 한정되는 것은 아니다.

[0105] 다른 실시예로, 프로세서(120)는 대상자의 인지기능 점수를 산출하고, 산출된 인지기능 점수를 대상자의 해당 연령의 평균 인지기능 점수와 비교하여 대상자가 해당 연령의 평균 이상의 인지능력을 갖는지 여부를 판별함으로써 대상자의 인지능력을 평가할 수도 있다. 이때, 프로세서(120)는 대상자의 인지능력이 대상자의 해당 연령의 평균 인지능력에서 벗어나는 정도를 확률 값으로 출력할 수도 있다.

[0106] 도 6 은 본 발명의 다른 일 실시예에 따른 인지능력 평가 방법을 설명하기 위한 흐름도이다.

[0107] 먼저, 본 발명의 인지능력 평가 방법은 피실험자들의 혈액검사 결과 및 인지기능검사 결과를 포함하는 실제 검사 데이터를 획득한다(610). 610 단계는 상술한 도 4 의 S100 단계와 유사한 단계일 수 있으며, 중복되는 설명에 대해서는 생략하기로 한다. 도 4 의 S100단계에서 설명한 일 실험예와 마찬가지로, 피실험자 그룹은 60세 이상의 고령자 81명으로 구성되며 실제 검사 데이터는 실제 피실험자 그룹에 대해 실시한 검사의 혈액검사 및 인지기능검사 데이터일 수 있다.

[0108] 상술한 도 4 의 실시예와 관련하여, 이하에서 설명하는 620 내지 660 단계는 도 4 의 S200 및 S300 단계를 대체할 수 있다. 이에 대해서는 아래에서 상세히 설명하기로 한다.

[0109] 다음으로, 본 발명의 인지능력 평가 방법은 실제 검사 데이터에 기초하여, 피실험들이 제1 그룹에 속하는지 여부와 관련된 인지기능검사의 주요 인지기능 항목을 결정한다(620). 즉, 일 실시예에 따르면 실제 검사 결과에 기초하여 인지기능검사의 각 항목과 슈퍼에이저 여부 간의 연간관계를 학습할 수 있다. 관련하여, 본 발명의 일 실시예에 따르면 전체 피실험자를 슈퍼에이저 그룹(제1 그룹) 및 일반 고령자 그룹(제2 그룹)으로 분류할 수 있다. 따라서, 2개의 그룹으로 분류된 피실험자들의 실제 검사 결과에 기초하여, 개별 피실험자들의 인지기능검사의 각 항목별 점수와 피실험자가 슈퍼에이저 그룹에 속하였는지 여부에 대한 연관관계를 도출할 수 있다. 이를 위해, 본 발명의 일 실시예에 따르면 피실험자들을 제1 또는 제2 그룹으로 분류한 데이터와 실제 검사 데이터에 기초하여, 인지기능검사의 각 항목과 피실험자들이 제1 그룹에 속하는지 여부 간의 연관관계를 학습하고, 학습 결과에 기초하여 주요 인지기능 항목을 결정할 수 있다.

[0110] 도 7 은 본 발명의 일 실시예에 따른 인지기능검사의 각 항목별 슈퍼에이저 그룹과의 연관관계를 나타낸 그래프이다.

[0111] 도 7 을 참조하면, 그래프의 가로축은 각 항목이 슈퍼에이저 여부에 연관된 정도를 나타내는 중요도(Variable Importance)이고, 세로축은 평가 요소(feature)로서 인지기능검사의 각 항목을 나타내고 있다. 도 7 에 나타난 바와 같이, 슈퍼에이저 여부에 높은 연관 정도를 가지는 항목, 즉 주요 인지기능 항목은 언어기억(language memory), 시각기억(visual memory) 및 시공간(visuospatial) 항목일 수 있다.

[0112] 관련하여, 상술한 도 4 의 S200 단계에서 상술한 바와 같이, 프로세서(120)는 검사 데이터에 포함되어 있는 혈액검사 결과와 인지기능검사 결과의 상관관계 분석을 수행하고, 인지능력 평가에 사용할 주요혈액인자들을 추출할 수 있다. [표 1]에 나타난 바와 같이 인지기능검사에서는 언어기억(language memory), 시각기억(visual memory), 시공간(visiopatial)이 $p<0.01$로 두 그룹 간에 상당히 유의미한 차이를 보였다. 이는 도 7 의 실시예

에 따른 인지기능검사 항목별 중요도와 일치하는 결과임을 알 수 있다.

[0113] 다음으로, 본 발명의 인지기능 평가 방법은 실제 검사 데이터를 보충할 수 있는 증강 검사 데이터를 생성한다(630). 보다 상세히, 본 발명의 인지기능 평가 방법은 기설정된 데이터 생성 알고리즘을 이용하여 가상의 혈액검사 결과 및 인지기능검사 결과를 포함하는 증강 검사 데이터를 생성할 수 있다. 관련하여, 앞서 상술한 바와 같이 평가모델을 생성하는 단계(S300)는 데이터 생성 알고리즘을 이용하여 실제 검사 데이터를 기초로 증강 데이터를 생성하여 평가모델을 생성(train) 또는 검증(verify)할 수 있다. 실제 검사 데이터로부터 생성된 증강 데이터를 증강 검사 데이터라 할 수 있으며, 일 실시예에 따르면 81개의 실제 검사 데이터로부터 729개의 증강 검사 데이터를 생성할 수 있다.

[0114] 다음으로, 본 발명의 인지기능 평가 방법은 증강 검사 데이터의 개별 데이터들에 대하여, 분류(classification) 알고리즘을 이용하여 상기 제1 그룹에 속하는지 여부를 예측한다(640). 즉, 생성된 증강 검사 데이터에 대하여 슈퍼에이저 그룹에 속하는지 여부에 대한 결과를 예측할 수 있다. 상술한 610 및 620 단계에서는 실제 개별 피실험자들을 슈퍼에이저 그룹 및 일반 고령자 그룹으로 분류한 결과를 이용할 수 있었지만, 증강 검사 데이터는 실제 피실험자에 대한 것이 아니기 때문에 개별 증강 검사 데이터들이 슈퍼에이저 그룹에 속하는지 여부에 대한 정보가 없다. 따라서, 증강 검사 데이터에 대해서도 슈퍼에이저 그룹 및 일반 고령자 그룹으로 분류하기 위해, 즉 개별 증강 검사 데이터가 슈퍼에이저에 대응하는지 여부를 예측하기 위해 별도의 분류 작업을 수행한 후 슈퍼에이저 여부에 대한 라벨링을 수행하는 단계가 필요하다.

[0115] 이를 위해, 본 발명의 일 실시예에 따르면 각각의 증강 검사 데이터가 슈퍼에이저 그룹에 속하는지 여부를 예측하기 위해 각기 다른 분류 알고리즘을 사용하여 1차 예측 결과를 생성한 후, 1차 예측 결과를 적층(stacking)하여 최종 예측 결과를 생성할 수 있다. 구체적인 일 실시예에 따르면, ExtraTreesClassifier, RandomForestClassifier, GradientBoosting Classifier의 분류 알고리즘을 이용하여 1차 예측 결과를 생성할 수 있지만, 반드시 이에 한정되지 않고 다양한 분류 알고리즘이 1차 예측결과를 생성하는 것에 사용될 수 있다. 보다 상세히, ExtraTreesClassifier는 Extremely Randomized Trees를 사용한 모델로, 무작위로 선정된 일부 요소에 대해 트리를 구성하여 예측하는 것이 가능하다. 또한, RandomForestClassifier는 여러 결정 트리를 앙상블한 모델로, 각 트리가 독립적으로 학습하고 예측하며 그 결과를 종합하여 예측할 수 있다. 또한, GradientBoosting Classifier 이전 트리의 오차를 보완하는 새로운 트리를 만들어 나가면서 예측 성능을 향상시킬 수 있다.

[0116] 또한, 본 발명의 일 실시예에 따르면 2개 이상의 서로 다른 분류 알고리즘을 이용하여 1차 예측 결과를 생성한 후, 예측 결과들을 새로운 요소로 사용하여 메타 모델을 훈련하고, 메타 모델은 예측 결과들을 입력으로 하여 최종 예측 결과를 생성할 수 있다. 이로써, 본 발명의 일 실시예에 따르면 각 분류기가 가진 약점을 상호 보완함으로써 전체적인 예측 성능이 향상된, 증강 검사 데이터가 슈퍼에이저 그룹에 속하는지 여부에 대한 예측 결과를 제공하는 예측 모델을 제공할 수 있다.

[0117] 한편, 본 발명의 일 실시예에 따르면, 증강 검사 데이터에 전처리 작업을 수행한 후 상술한 예측 모델의 분류 알고리즘들에 입력할 수 있다. 보다 상세히, 증강 검사 데이터에 대하여 LabelEncoder를 통해 범주형 데이터를 수치형 데이터로 변환하고, SimpleImputer를 통해 결측값(missing values)을 평균, 중간값, 최빈값 등의 다른 값으로 대체하는 전처리 작업을 수행할 수 있다. 즉, LabelEncoder는 범주형 데이터의 인코딩을 처리하고, SimpleImputer는 결측값을 적절히 처리하여 데이터의 완전성을 유지할 수 있도록 함으로써, 본 발명의 일 실시예에 따른 분류 알고리즘에 의한 예측 모델의 정확도를 높일 수 있다.

[0118] 도 8 은 본 발명의 실시예에 따른 분류 알고리즘에 의한 예측 모델의 정확도를 나타낸 그래프이다.

[0119] 도 8 의 그래프는 본 발명의 예측 모델에 대한 ROC(Receiver Operating Characteristic) 커브로서, ROC 커브는 이진 분류 모델의 성능을 시각화하는 데 사용되되는 그래프이다. ROC 그래프에서, x축(False Positive Rate, FPR)은 실제로는 음성인데 본 발명의 예측 모델이 양성으로 예측한 경우의 비율을 나타낸다. 수학적으로는 FPR = FP / (FP + TN)이다. 여기서 FP는 False Positive(예측 모델이 양성으로 예측했지만 실제로는 음성인 경우의 수), TN은 True Negative(예측 모델이 음성으로 예측하고 실제로도 음성인 경우의 수)이다. 또한, Y축(True Positive Rate, TPR)은 실제로 양성이고 모델이 양성으로 정확하게 예측한 경우의 비율을 나타낸다. 수학적으로는 TPR = TP / (TP + FN)이다. 여기서 TP는 True Positive(모델이 양성으로 정확하게 예측한 경우의 수), FN은 False Negative(모델이 음성으로 예측했지만 실제로는 양성인 경우의 수)이다.

[0120] 도 8을 참조하면, 본 발명의 분류 모델에 따른 ROC 커브는 모델의 성능을 여러 임계값(threshold)에서 평가할

수 있다. 도 8의 그래프에 도시된 각각의 커브는 예측 모델이 특정 임계값에서 얻은 FPR과 TPR의 조합을 나타낼 수 있다. 보다 상세히, 커브가 좌상단에 가까울수록 예측 모델의 성능이 좋다고 판단할 수 있으며, 이는 예측 모델이 높은 TPR을 유지하면서 낮은 FPR을 가졌다는 것을 의미한다. 관련하여, 커브 아래 영역인 AUC(Area Under the Curve)가 클수록 모델의 성능이 우수하다고 평가할 수 있다. 도 8의 일 실시예에서는, ROC 커브의 AUC 값이 모두 0.9로서 본 발명의 예측 모델이 높은 정확도를 가졌다고 평가할 수 있다.

[0121] 또한, 추가적인 도 8의 실시예에서, 예측 모델의 정확도를 평가하기 위해 마이크로 평균 ROC(Micro-average ROC) 커브가 사용될 수 있다. 마이크로 평균 ROC 커브는 다중 클래스 분류에서 각 클래스마다 ROC 커브를 그리는 대신, 모든 클래스에 대한 하나의 ROC 커브를 계산하는 것을 말한다. 즉, 마이크로 평균 ROC 커브는 모든 클래스의 FP(거짓 양성), TP(참 양성) 및 FN(거짓 음성)의 총계를 사용하여 전체적인 평균 ROC 커브를 계산할 수 있다. 또한, 매크로 평균 ROC(Macro-average ROC) 커브는 각 클래스에 대해 개별적으로 ROC 곡선을 계산한 다음, 이를 평균하여 얻은 커브이다. 도 8의 실시예에서는, 마이크로 평균 ROC 커브 및 매크로 평균 ROC 커브의 AUC 값이 모두 0.9 이상으로서 본 발명의 예측 모델이 높은 정확도를 가졌다고 평가할 수 있다. 즉, 본 발명의 일 실시예에 따르면 약 90%의 정확도로 증강 검사 데이터의 개별 데이터들이 슈퍼 에이저 그룹에 속하는지 여부를 예측할 수 있다.

[0122] 다음으로, 본 발명의 인지능력 평가 방법은 실제 검사 데이터 및 증강 검사 데이터를 제1 그룹 및 제2 그룹으로 분류한다(650). 이하의 명세서에서, 설명의 편의를 위하여 실제 검사 데이터 및 증강 검사 데이터를 합쳐 전체 검사 데이터라 지칭할 수 있다. 상술한 바와 같이, 인지기능검사의 주요 항목을 결정하는 단계(620)에서 피실험자들 각각을 제1 그룹 또는 제2 그룹으로 나누었으므로, 실제 검사 데이터를 제1 그룹 또는 제2 그룹으로 분류할 수 있다. 또한, 증강 검사 데이터에 대하여 제1 그룹에 속하는지 여부를 예측하는 단계(640)에서 증강 검사 데이터 각각에 대하여 분류 알고리즘을 이용하여 제1 그룹에 속하는지 여부를 예측하였으므로, 예측 결과에 기초하여 증강 검사 데이터 역시 제1 그룹 또는 제2 그룹으로 분류할 수 있다. 구체적인 일 예시에서, 실제 실험 데이터가 81개이고, 증강 실험 데이터가 791개인 경우, 총 810개의 데이터에 대해 449명은 슈퍼에이저 그룹에 해당하고, 361명은 일반 고령자 그룹에 해당하는 것으로 분류할 수 있다.

[0123] 다음으로, 본 발명의 인지능력 평가 방법은 인지기능검사의 각 주요 인지기능 항목별로 제1 그룹에 연관된 주요 혈액인자를 하나 이상 추출한다(660). 보다 상세히, 본 발명의 일 실시예에 의하면 전체 검사 데이터에 대하여 각 인지기능 항목별로 혈액 인자들의 SHAP (SHapley Additive exPlanations) 값을 도출하고, SHAP 값을 사용하여 슈퍼에이저인지 여부에 영향을 주는 혈액인자를 결정할 수 있다. SHAP은 모델의 예측을 설명하는 데 사용되는 기법으로서, SHAP 값은 Shapley 값이라고도 불리며, 각 혈액인자 특성(feature)이 모델의 예측에 기여하는 정도를 측정할 수 있다. 특히, SHAP은 각 혈액인자 특성의 중요도를 개별적으로 계산하면서 다양한 조합을 고려하여 예측을 설명할 수 있다. 즉, 각각의 인지능력 항목에 대해 SHAP 분석을 수행하여 어떤 혈액 인자가 슈퍼에이저로 분류되는 것에 대한 영향을 정량화할 수 있다.

[0124] 도 9(a) 내지 도 9(e)는 각 인지기능검사 항목별로 결정된 제1 그룹에 연관된 혈액인자를 설명하기 위한 그래프이다.

[0125] 도 9(a)를 참조하면, 슈퍼에이저 그룹을 일반인 그룹과 비교하였을 때 Attention 인지기능 항목에서 혈액 인자들은, Leptin(↓), HDL(↑), Triglyceride(↓), Total bilirubin(↑), CD36(↑), Na(↓)임을 알 수 있다. 이때, (↑)는 해당 혈액인자 수치가 높을수록 해당 인지기능 점수가 높고, (↓)는 해당 혈액인자 수치가 낮을수록 해당 인지기능 점수가 높다는 것을 의미한다.

[0126] 도 9(b)를 참조하면, 슈퍼에이저 그룹을 일반인 그룹과 비교하였을 때 Language 인지기능 항목에서 혈액 인자들은, RAGE(↑), Hb(↓), MCHC(↓)임을 확인할 수 있다.

[0127] 도 9(c)를 참조하면, 슈퍼에이저 그룹을 일반인 그룹과 비교하였을 때 Verbal Memory 인지기능 항목에서 혈액 인자들은 Glucose(↓), C-peptide(↓), MCH(↑), HDL(↑), Platelet(↓), Free T4(↓)임을 확인할 수 있다.

[0128] 도 9(d)를 참조하면, 슈퍼에이저 그룹을 일반인 그룹과 비교하였을 때 Visual Memory 인지기능 항목에서 혈액 인자들은 WBC(↓), RBC(↑),oxLDL(↑)임을 확인할 수 있다.

[0129] 도 9(e)를 참조하면, 슈퍼에이저 그룹을 일반입 그룹과 비교하였을 때 Visuospatial Superager 인지기능 항목에서 혈액 인자들은 Score Visuospatial (↑) : ApoE(↓), HbA1c(↓), AST(↓), Free T4(↑), Triglyceride(↓)임을 확인할 수 있다.

[0130] 도 9(f)를 참조하면, 슈퍼에이저 그룹을 일반입 그룹과 비교하였을 때 Frontal 인지기능 항목에서 혈액 인자들

은 CD36(↑)임을 확인할 수 있다. 즉, [표 1]에서 살펴본 바와 같이, 전두엽(frontal) 인지기능 항목은 일반 고령자와 슈퍼에이저 간의 차이가 크지 않음을 확인할 수 있다.

[0131] 다음으로, 주요 인지기능 항목별로 결정된 주요 혈액인자들을 기초로 인지능력을 평가하는 평가 모델을 생성한다(670). 결정된 주요 혈액인자에 기초하여 인지능력을 평가하는 평가모델을 생성하는 단계는 도 4 의 S300 단계와 유사하므로, 중복되는 설명은 생략하기로 한다. 한편, 도 4 의 실시예와 비교할 때, 도 6 의 실시예에서는 증강 검사 데이터에 대해 분류 알고리즘을 이용하여 슈퍼에이저인지 여부를 예측하고 라벨링하는 단계가 추가적으로 수행되며, 예측된 결과에 기초하여 전체 검사 데이터에 대해 슈퍼에이저 여부를 분류할 수 있다. 따라서, 도 6 의 실시예에 기초하여 추출된 주요혈액인자들은 슈퍼에이저 그룹에서는 레귤러 분포(regular distribution)를 보이지만 일반 고령자 그룹에서는 랜덤 분포(random distribution)을 보이므로, 도 6의 실시예에서 추출된 주요 혈액인자들이 보다 효과적으로 슈퍼에이저의 특징을 보여주는 바이오마커라고 할 수 있다.

[0132] 상술한 바에 따르면, 도 4의 S200 단계에서, 프로세서(120)는 기 설정된 변수 선택 알고리즘, 일 실시예에 따르면 selectKBest를 이용하여 인지기능검사의 항목별로 주요혈액인자를 각각 추출할 수 있다. 한편, 본 발명의 다른 실시예에 따르면, selectKBest가 아닌 하나 이상의 Feature selection을 이용하여 인지영역별 주요 바이오마커를 선정할 수 있다. 이때, 하나 이상의 Feature selection 방법은 Feature Selection with Recursive Feature Elimination (RFE) combined with Simulated Annealing (SA), Feature Selection(인지능) with RFE, Feature Selection(인지능) with BOURUTA일 수 있다.

[0133] 보다 상세히, 본 발명의 일 실시예에 따른 프로세서(120)는 Feature Selection(SA) with RFE를 이용하여 슈퍼에이저 여부에 따라 랜덤포레스트(Randomforest), 아다부스트(Adaboost), lightGBM classifier를 활용해 RFE를 진행하여 얻어진 상위 10개 피처를 추출할 수 있다.

[0134] 구체적인 실시예에 따르면, 프로세서(120)는 랜덤포레스트를 활용해 RFE를 진행하여 EPO, FABP-4, Free T4, Adiponectin, Vit B12, Folic acid, HDLCholesterol, WBC, MCV, Cl 의 10개 상위 인자를 추출할 수 있다. 또한, 프로세서(120)는 아다부스트를 활용해 RFE를 진행하여 HDLCholesterol, FABP-4, EPO, WBC, oxLDL, CD36, RBC, Vit B12, Adiponectin, MCV 의 10개 상위 피처를 추출할 수 있다. 또한, 프로세서(120)는 lightGBM classifier를 활용해 RFE를 진행하여 Vit B12, HDLCholesterol, Platelet, Total cholesterol, EPO, Cl, AST, BUN, CRP, RDW(CV)의 10개 상위 인자를 추출할 수 있다. 또한, 프로세서(120)는 겹쳐지는 인자들(EPO, VitB12, HDLCholesterol)을 추출할 수 있다.

[0135] 한편, 프로세서(120)가 도 4의 S100 단계에서 획득한 피실험자들의 검사 데이터에 따르면, 각 환자에 대해 슈퍼에이저 여부(제1 그룹에 속하는지 여부)가 바이너리(binary)하게 나뉘어져 있지만, 개별 인지기능 점수는 연속적(continuous)이다. 따라서 연속변수를 통해 총 56개의 혈액 인자 중 바이오마커를 검출해내기에는 전체 81명의 데이터셋은 양적으로 부족하다.

[0136] 관련하여 상술한 바와 같이 슈퍼에이저 여부는 Attention, Language, VerbalMemory, VisualMemory, Visuospatial, Frontal의 인지영역에 따른 인지능 점수를 통해 결정된다. 따라서, 해당 인지영역의 인지능 점수를 결정짓는 제일 중요한 시험항목점수 중 상위 25%의 상위 그룹과 나머지 그룹으로 나누어, 각 인지영역에 대한 인지능 점수가 좋은 상위 그룹의 주요 인자를 알아낼 필요가 있다.

[0137] 이를 위하여, 프로세서(120)는 Feature selection(인지능) with RFE를 진행할 수 있다. Feature selection(인지능) with RFE는 각 인지영역에 대한 인지능 점수 상위 그룹 여부에 따라 랜덤포레스트, 아다부스트, lightGBM classifier를 활용해 RFE를 진행하는 것일 수 있다. 프로세서(120)는 이를 통해 얻어진 상위 10개 인자 중, 겹쳐지는 인자들(HDLCholesterol, Leptin, FABP-3, Free T4, Glucose, RDW(CV), CD36, ApoE, EPO, Vit B12, RAGE, Total cholesterol, oxLDL, FABP-4, PCSK9, Triglyceride)을 추출해 낼 수 있다.

[0138] 또한, 프로세서(120)는 Feature selection(인지능) with BORUTA를 진행할 수 있다. Feature selection(인지능) with BORUTA는 각 인지영역에 대한 인지능 점수 상위 그룹 여부에 따라 R기반 BORUTA를 진행하는 것일 수 있다. 프로세서(120)는 이를 통해 10개의 인자(HDLCholesterol, BUN, ApoE, HbA1c, Insulin, EPO, Vimentin, FABP-4, Free T4, Total Cholesterol, RAGE, PCSK9)를 추출할 수 있다.

[0139] 다음으로, 프로세서(120)는 상술한 Feature Selection with Recursive Feature Elimination (RFE) combined with Simulated Annealing (SA), Feature Selection(인지능) with RFE, Feature Selection(인지능) with BOURUTA의 3가지 방법으로 추출된 인자들로 슈퍼에이저 여부를 예측하고, 예측 결과의 정확도를 판단한다. 일 실시예에 따르면 Feature Selection(인지능) with RFE 및 Feature Selection(인지능) with BOURUTA에 의해 추

출된 인자들로 슈퍼에이저를 예측했을 때의 정확도가 가장 높으므로, 프로세서(120)는 EPO, Vit B12, HDLCholesterol, Leptin, FABP-3, Free T4, Glucose, RDW(CV), CD36, ApoE, RAGE, Total cholesterol, oxLDL, FABP-4, PCSK9, Triglyceride, BUN, HbA1c, Insulin, Vimentin 의 총 20개의 인자들을 슈퍼에이저 바이오마커로 선별할 수 있다.

[0140] 한편, 상술한 도 6의 630 단계에 따르면, 평가 모델의 정확도를 높이기 위해 증강 검사 데이터를 생성할 수 있다. 관련하여, 본 발명의 다른 실시예에 따르면, 증강 검사 데이터는 평가 모델의 검증을 위해 사용될 수 있다.

[0141] 보다 상세히, 생성된 증강 검사 데이터를 이용하여 평가 모델을 훈련시키고, 원본 데이터를 이용하여 Large-language model based 'GReaT' 방법을 통해 평가 모델을 검증하였다. 구체적인 실시예에 따르면, 전체 81개의 데이터 중 증강을 위해 사용된 데이터는 49개이고, 70개의 증강 데이터를 얻었으며, 이를 통해 앞서 상술한 총 20개의 바이오마커를 통해 실제 데이터와 증강 데이터의 상관관계분석을 한 결과 차이가 매우 적었다.

[0142] 한편, 본 발명의 일 실시예에 따르면 상술한 20개의 바이오마커 중 일부 바이오마커는 2개 이상의 인지영역에 공통적으로 해당될 수 있다. 이와 같이 복수개의 인지영역에 공통적으로 해당되는 바이오마커는 인지예비능을 측정하는 지표로 사용될 수 있으며, 이를 통해 슈퍼에이저여부를 보다 정확하게 예측할 수 있다.

[0143] 전술한 바와 같이, 본 발명의 실시예들에 따른 인지능력 평가 방법 및 장치는 혈액검사를 기반으로 인지능력을 평가할 수 있으며, 장시간 소요되는 인지기능검사 또는 MRI 스캔 없이 간편하게 인지능력 평가 결과를 획득할 수 있음으로써 인지능력 평가에 소요되는 시간 및 비용을 크게 절감할 수 있다.

[0144] 또한, 본 발명의 실시예들은 인지능력 평가에 대한 접근성이 확대되도록 할 수 있다. 일례로 고령 인구에서 주기적으로 수행되는 혈액검사 결과를 이용하여 인지능력을 평가하도록 함으로써 인지능력에 대한 정기적인 검진을 수행하도록 할 수도 있다.

[0145] 또한, 본 발명의 실시예들은 대상자가 본인이 속한 연령그룹의 평균 인지능력 이상의 인지능력을 갖는지 여부를 평가하여 질환자가 아닌 건강한 사람 역시 간편하게 인지능력을 검사하고 관리하도록 할 수 있으며, 혈액검사를 기반으로 하여 비교적 객관성이 있는 인지능력 평가 방법을 제공함으로써 인지능력의 변화를 추적할 수 있는 효과를 가진다.

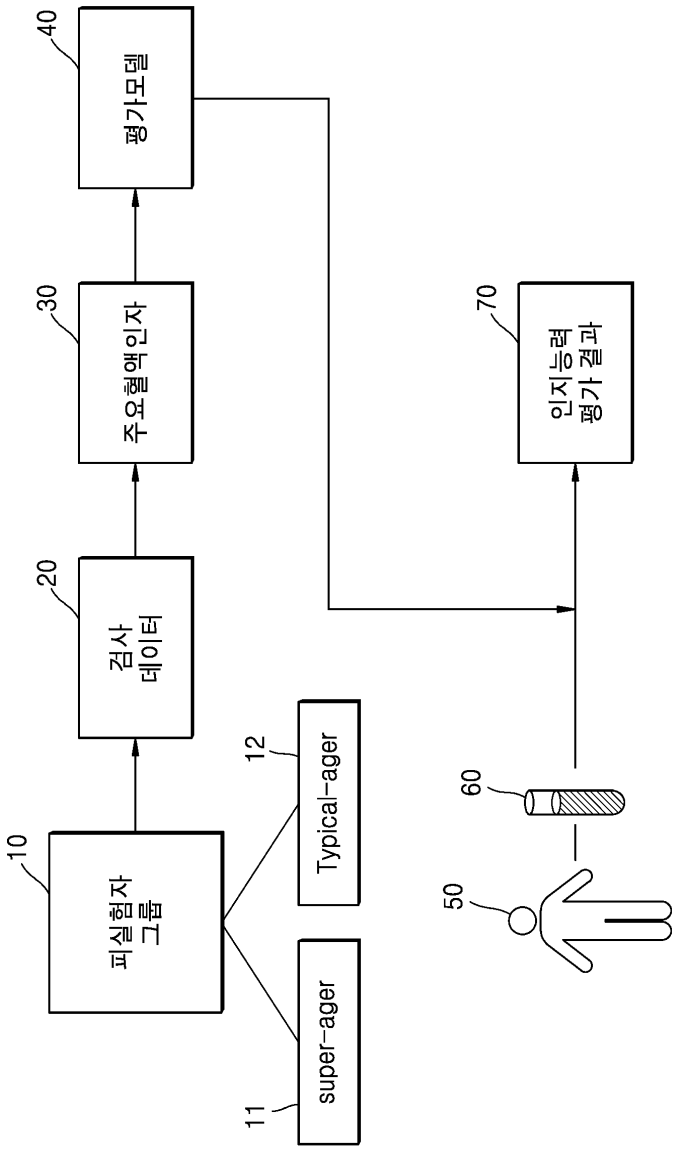
[0146] 본 실시예들은 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터에 의해 실행 가능한 명령어를 포함하는 기록 매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터 판독가능 매체는 컴퓨터 저장 매체 및 통신 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다. 통신 매체는 전형적으로 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈과 같은 변조된 데이터 신호의 기타 데이터, 또는 기타 전송 메커니즘을 포함하며, 임의의 정보 전달 매체를 포함한다.

[0147] 또한, 본 명세서에서, “부”는 프로세서 또는 회로와 같은 하드웨어 구성(hardware component), 및/또는 프로세서와 같은 하드웨어 구성에 의해 실행되는 소프트웨어 구성(software component)일 수 있다.

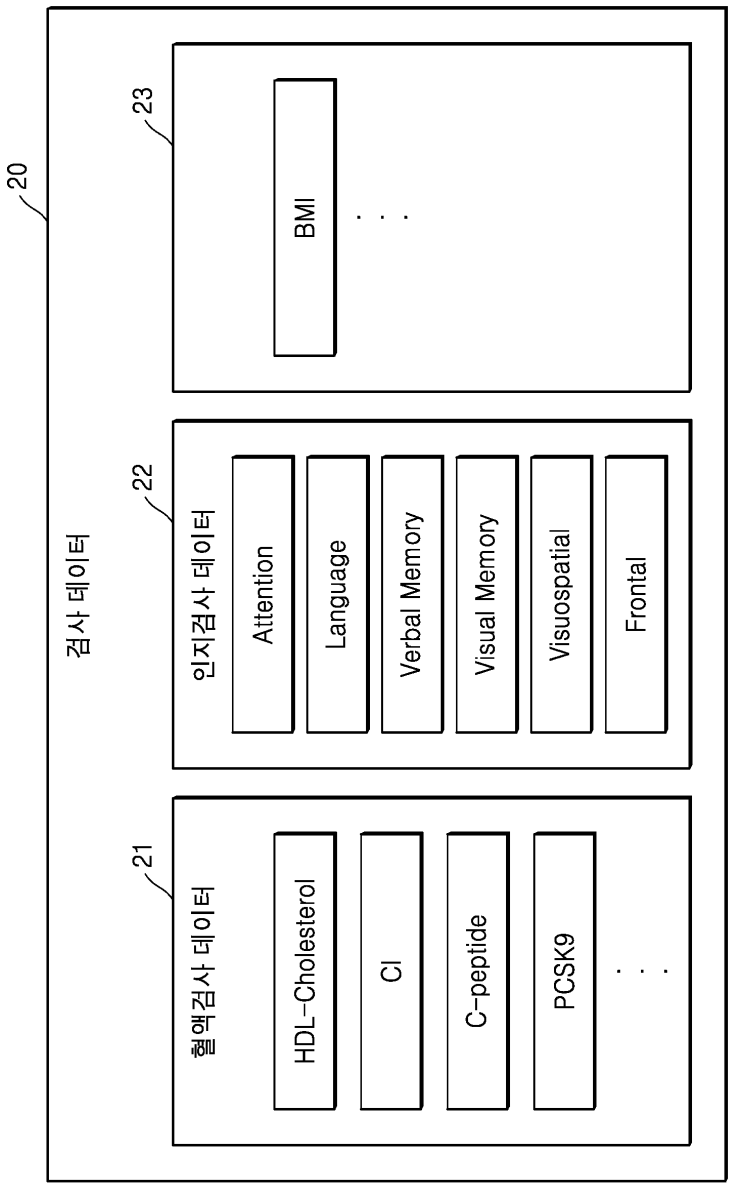
[0148] 이와 같이 본 발명은 도면에 도시된 일 실시예를 참고로 하여 설명하였으나 이는 예시적인 것에 불과하며 당해 분야에서 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 실시예의 변형이 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

도면

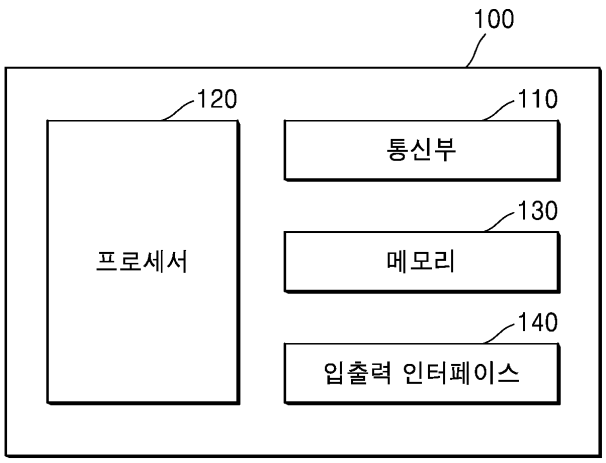
도면1


40
평가모델
30
주요혈액인자
20
검사
데이터
10
피실험자
그룹
12
Typical-ager
11
super-ager
70
인지능력
평가 결과
60
50

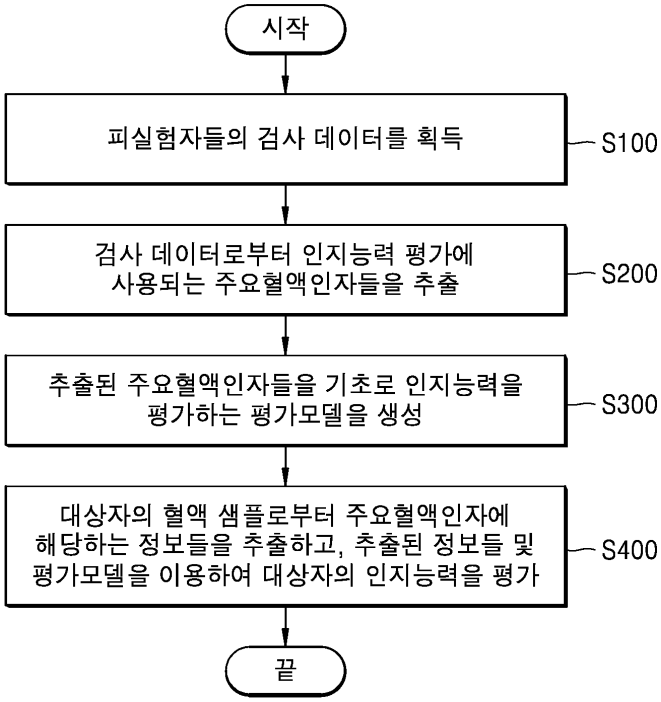
도면2


20
검사 데이터
23
BMI
. . .
22
인지검사 데이터
Attention
Language
Verbal Memory
Visual Memory
Visuospatial
Frontal
21
혈액검사 데이터
HDL-Cholesterol
CI
C-peptide
PCSK9
. . .

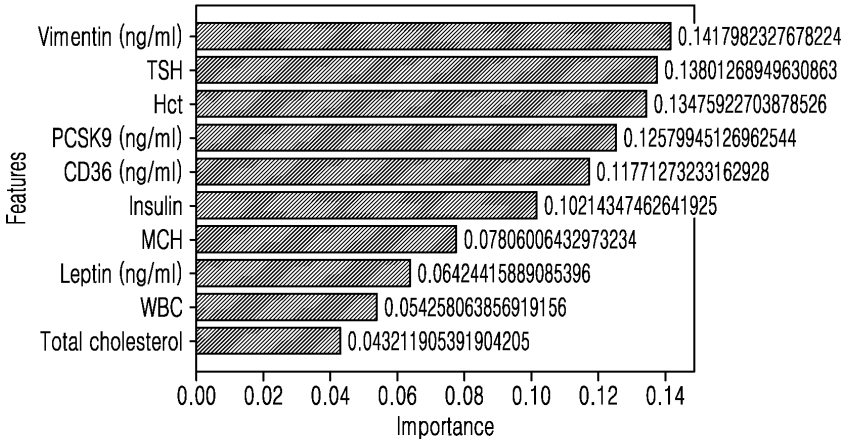
도면3


100
120
프로세서
110
통신부
130
메모리
140
입출력 인터페이스

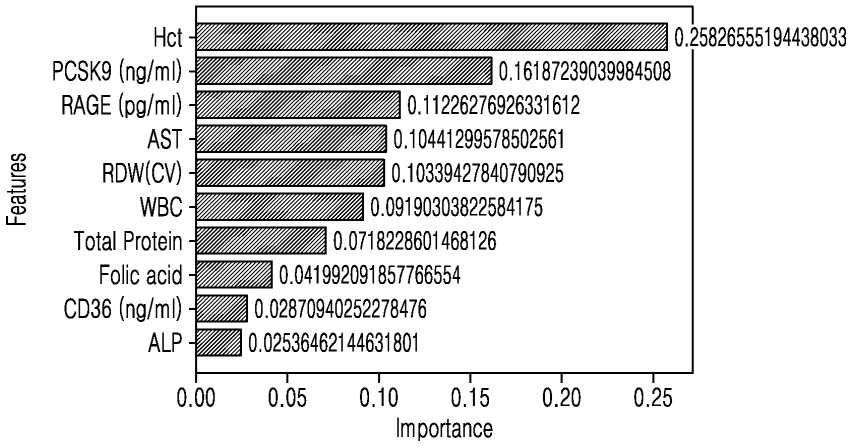
도면4



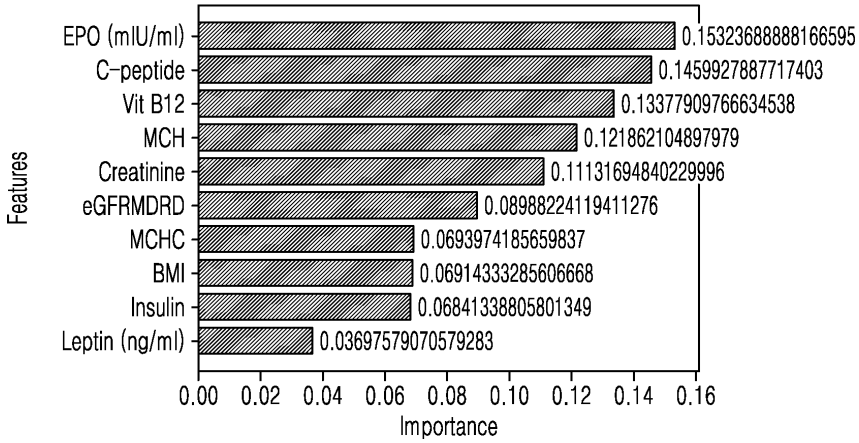
도면5a



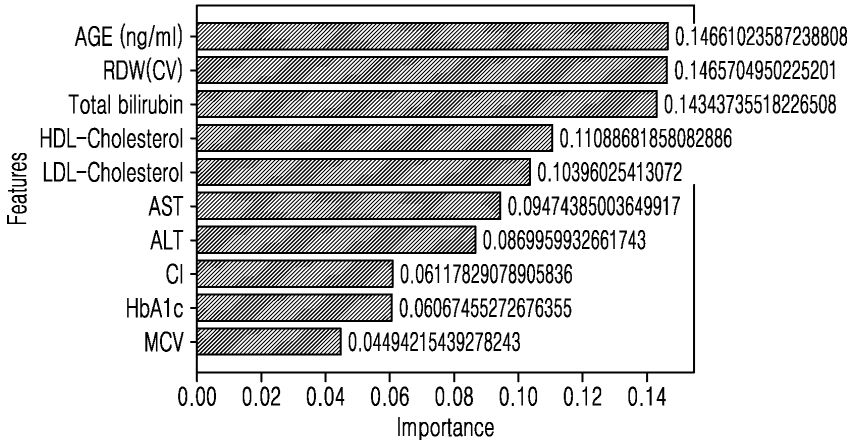
도면5b



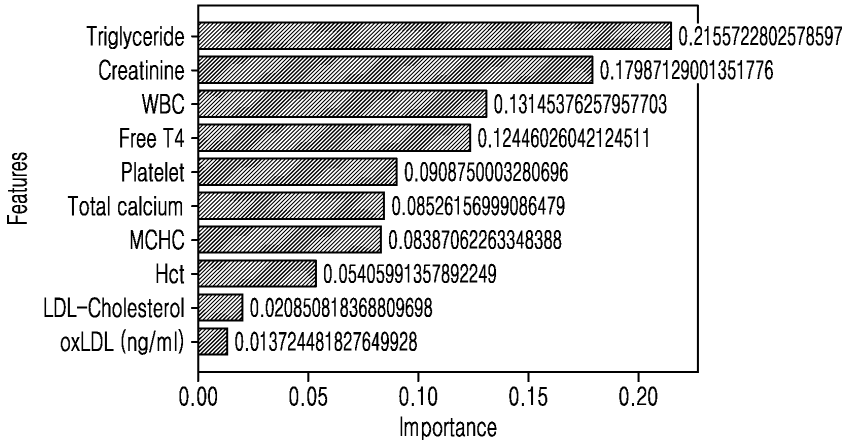
도면5c



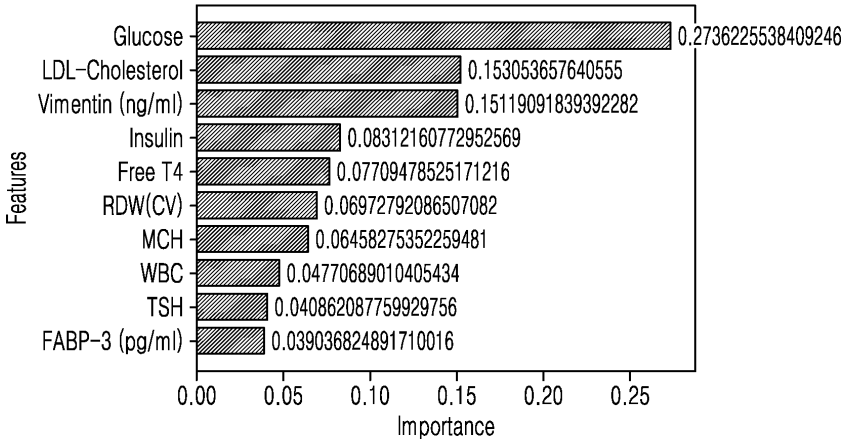
도면5d



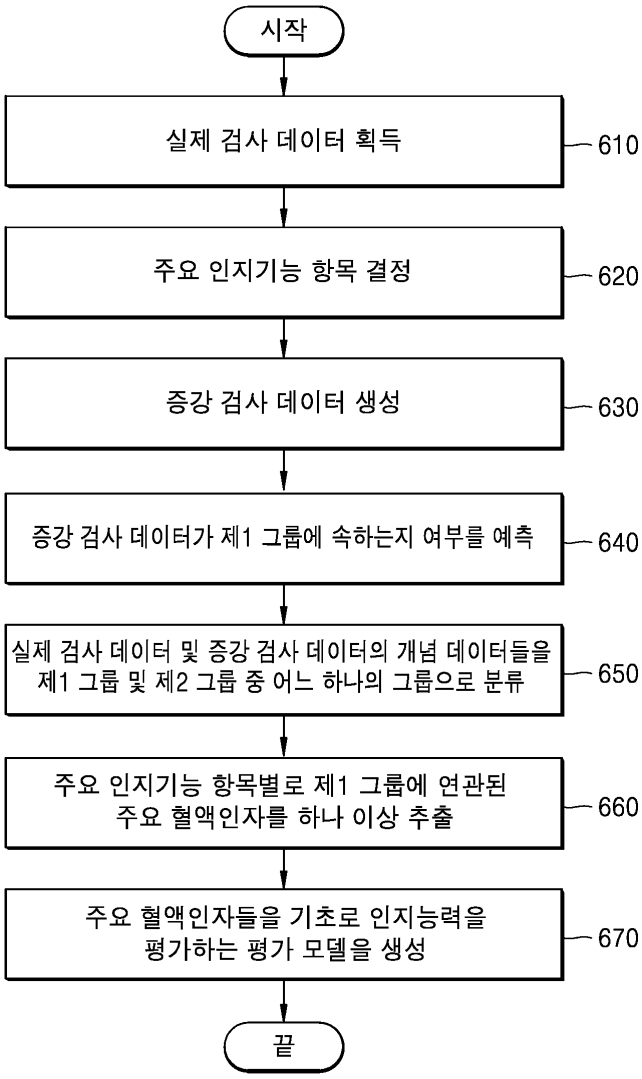
도면5e



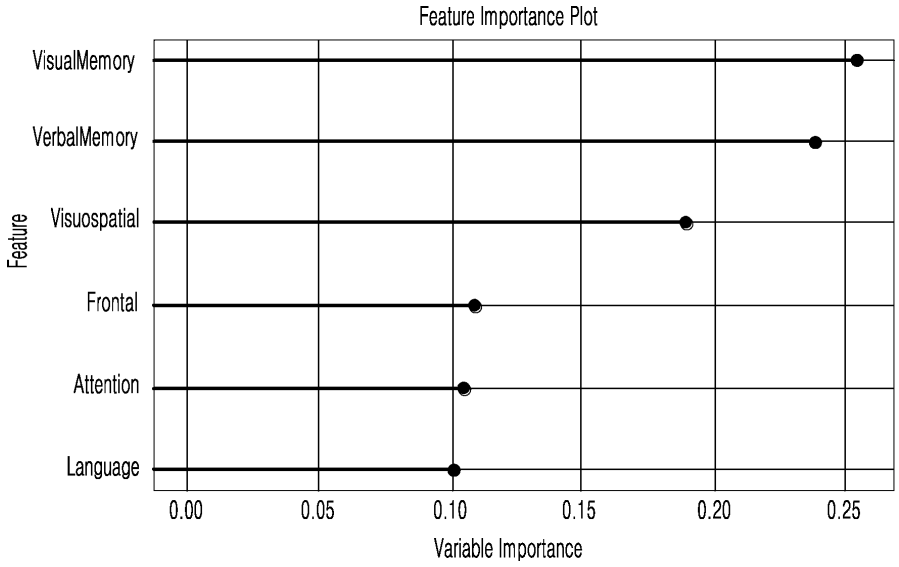
도면5f


Glucose
LDL-Cholesterol
Vimentin (ng/ml)
Insulin
Free T4
RDW(CV)
MCH
WBC
TSH
FABP-3 (pg/ml)
Features
0.2736225538409246
0.153053657640555
0.15119091839392282
0.08312160772952569
0.07709478525171216
0.06972792086507082
0.06458275352259481
0.04770689010405434
0.040862087759929756
0.039036824891710016
0.00
0.05
0.10
0.15
0.20
0.25
Importance

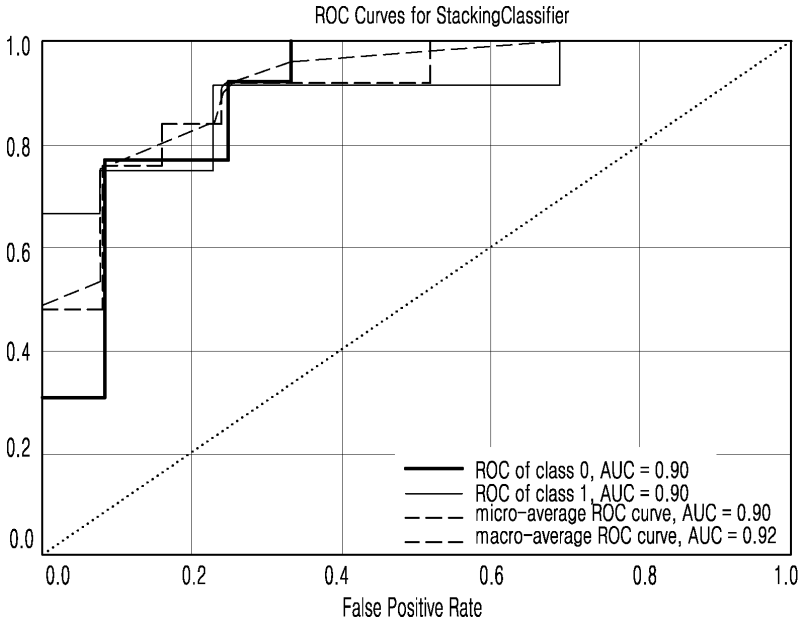
도면6


시작
실제 검사 데이터 획득
610
주요 인지기능 항목 결정
620
증강 검사 데이터 생성
630
증강 검사 데이터가 제1 그룹에 속하는지 여부를 예측
640
실제 검사 데이터 및 증강 검사 데이터의 개념 데이터들을 제1 그룹 및 제2 그룹 중 어느 하나의 그룹으로 분류
650
주요 인지기능 항목별로 제1 그룹에 연관된 주요 혈액인자를 하나 이상 추출
660
주요 혈액인자들을 기초로 인지능력을 평가하는 평가 모델을 생성
670
끝

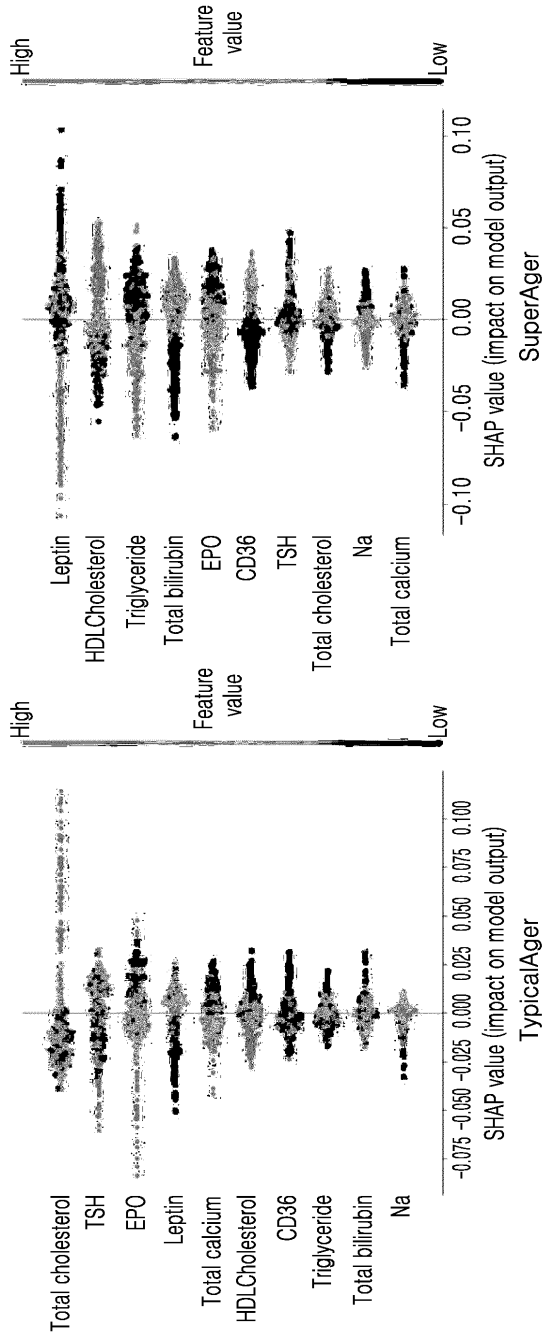
도면7


Feature Importance Plot
VisualMemory
VerbalMemory
Visuospatial
Frontal
Attention
Language
Feature
0.00
0.05
0.10
0.15
0.20
0.25
Variable Importance

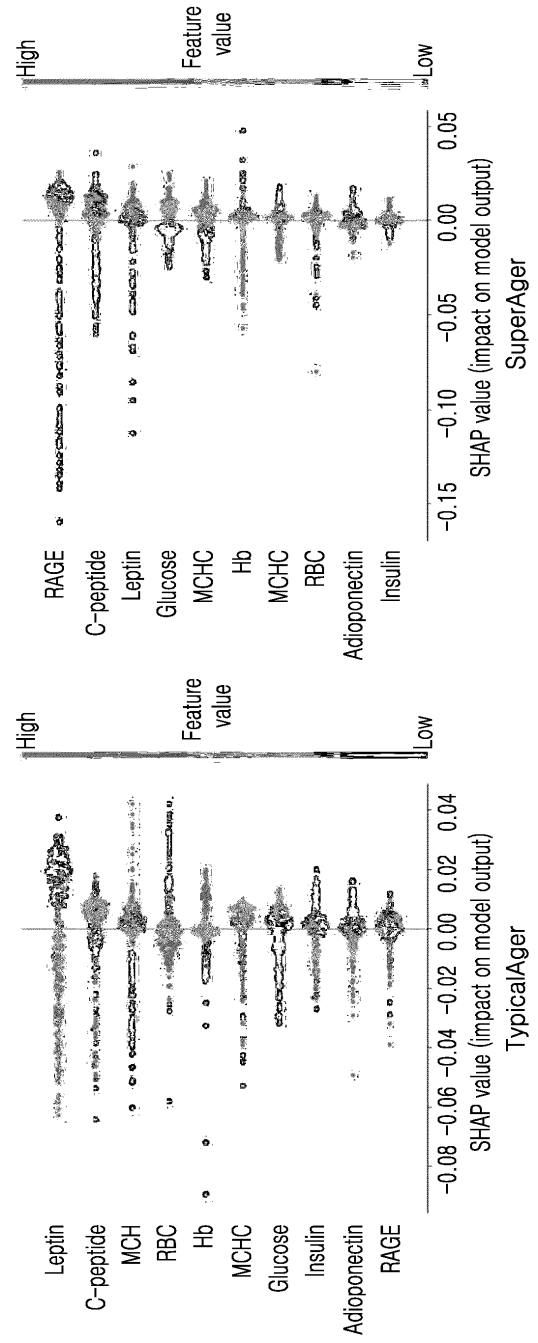
도면8


ROC Curves for StackingClassifier
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate
ROC of class 0, AUC = 0.90
ROC of class 1, AUC = 0.90
micro-average ROC curve, AUC = 0.90
macro-average ROC curve, AUC = 0.92

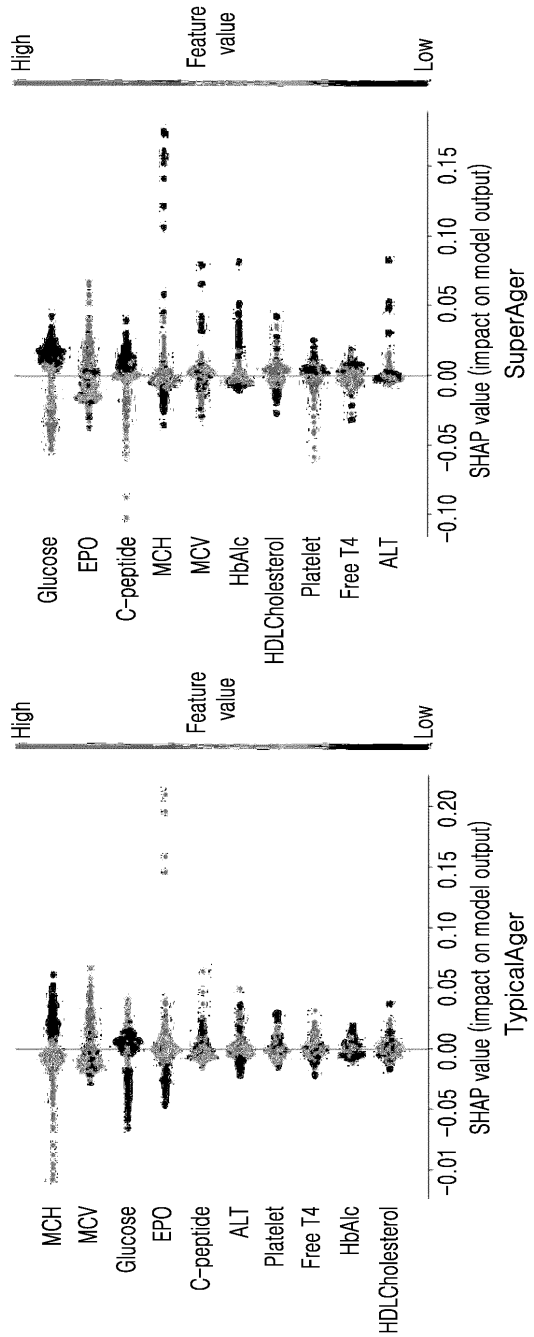
도면9a


High
Feature value
Low
Leptin
HDLCholesterol
Triglyceride
Total bilirubin
EPO
CD36
TSH
Total cholesterol
Na
Total calcium
−0.10
−0.05
0.00
0.05
0.10
SHAP value (impact on model output)
SuperAger
High
Feature value
Low
Total cholesterol
TSH
EPO
Leptin
Total calcium
HDLCholesterol
CD36
Triglyceride
Total bilirubin
Na
−0.075
−0.050
−0.025
0.000
0.025
0.050
0.075
0.100
SHAP value (impact on model output)
TypicalAger

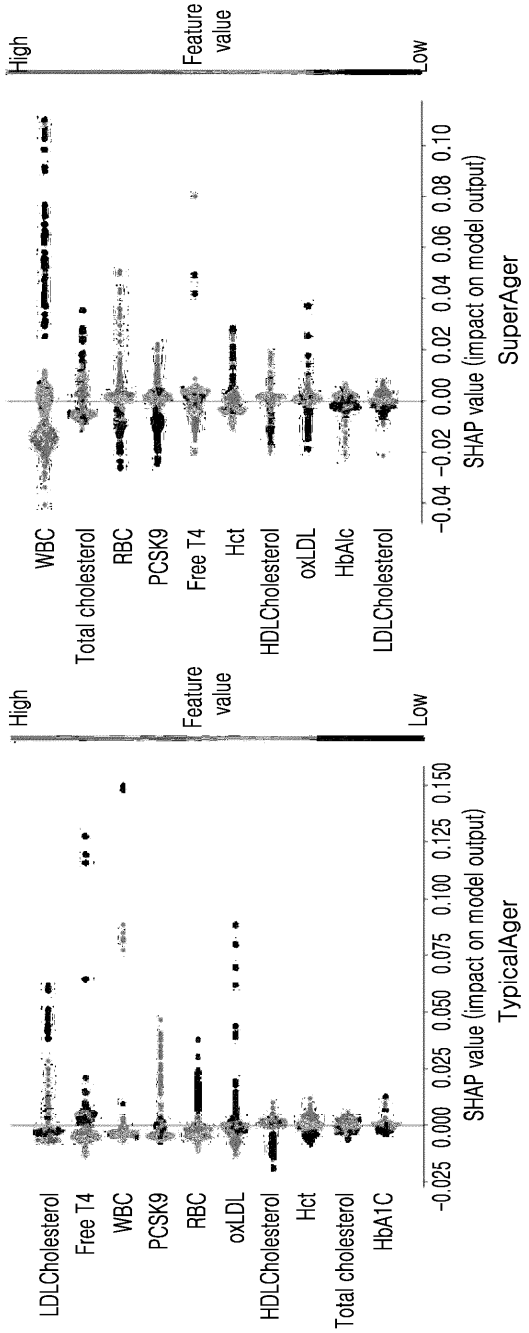
도면9b


High
Feature value
Low
RAGE
C-peptide
Leptin
Glucose
MCHC
Hb
MCHC
RBC
Adioponectin
Insulin
−0.15 −0.10 −0.05 0.00 0.05
SHAP value (impact on model output)
SuperAger
High
Feature value
Low
Leptin
C-peptide
MCH
RBC
Hb
MCHC
Glucose
Insulin
Adioponectin
RAGE
−0.08 −0.06 −0.04 −0.02 0.00 0.02 0.04
SHAP value (impact on model output)
TypicalAger

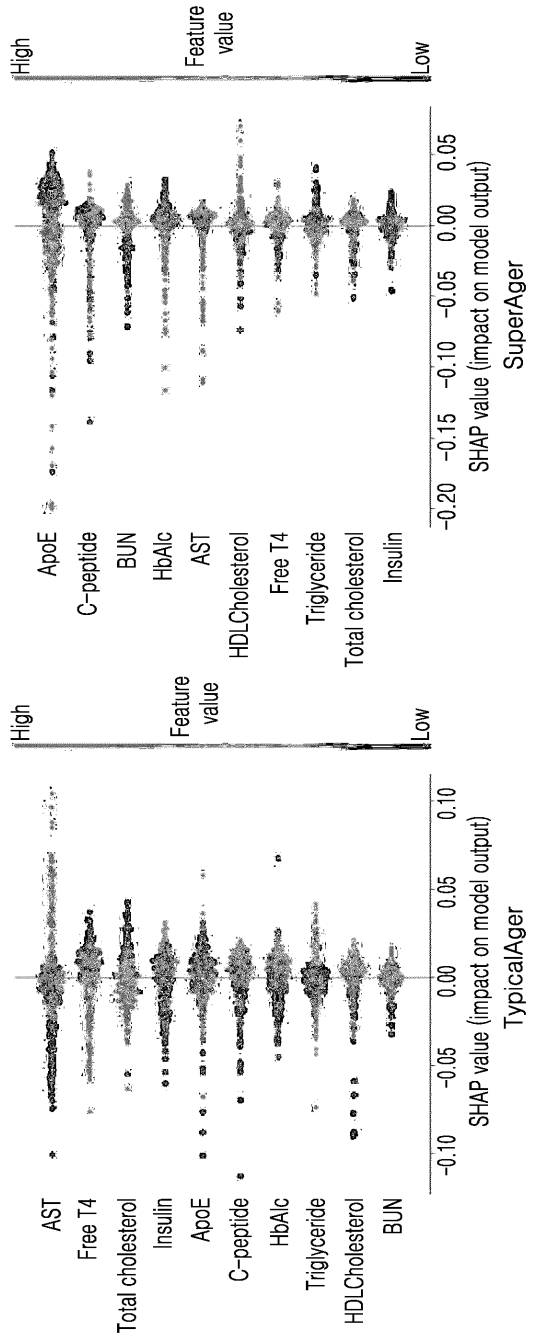
도면9c


High
Feature value
Low
Glucose
EPO
C-peptide
MCH
MCV
HbAlc
HDLCholesterol
Platelet
Free T4
ALT
−0.10 −0.05 0.00 0.05 0.10 0.15
SHAP value (impact on model output)
SuperAger
High
Feature value
Low
MCH
MCV
Glucose
EPO
C-peptide
ALT
Platelet
Free T4
HbAlc
HDLCholesterol
−0.01 −0.05 0.00 0.05 0.10 0.15 0.20
SHAP value (impact on model output)
TypicalAger

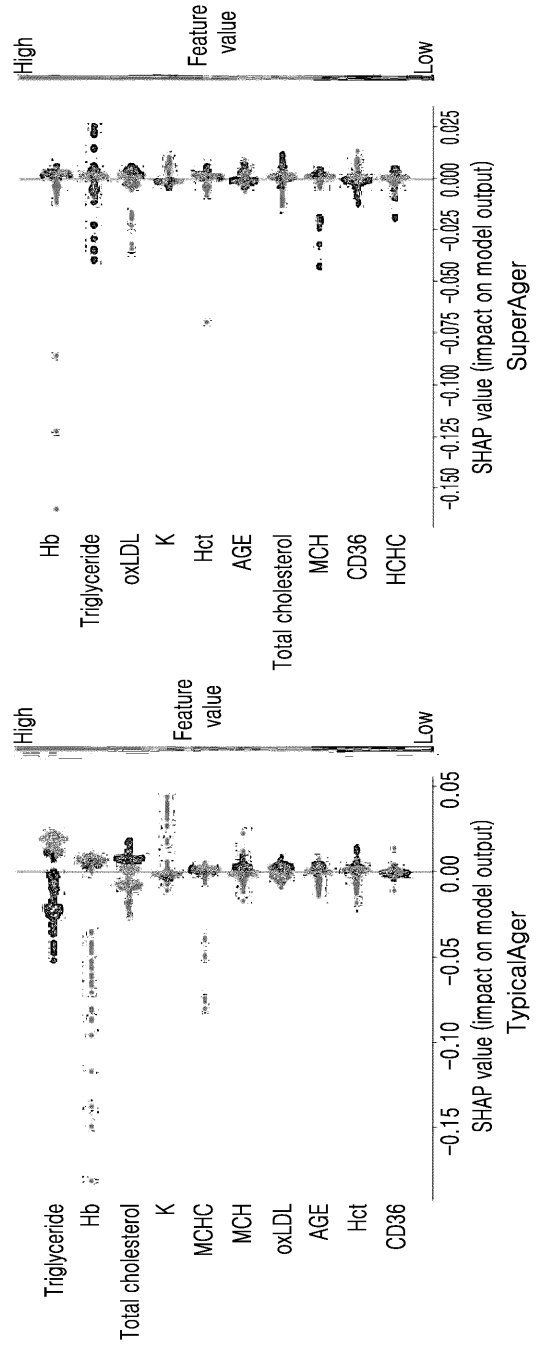
도면9d


High
Feature value
Low
WBC
Total cholesterol
RBC
PCSK9
Free T4
Hct
HDLCholesterol
oxLDL
HbAlc
LDLCholesterol
−0.04 −0.02 0.00 0.02 0.04 0.06 0.08 0.10
SHAP value (impact on model output)
SuperAger
High
Feature value
Low
LDLCholesterol
Free T4
WBC
PCSK9
RBC
oxLDL
HDLCholesterol
Hct
Total cholesterol
HbA1C
−0.025 0.000 0.025 0.050 0.075 0.100 0.125 0.150
SHAP value (impact on model output)
TypicalAger

도면9e


High
Feature value
Low
ApoE
C-peptide
BUN
HbAlc
AST
HDLCholesterol
Free T4
Triglyceride
Total cholesterol
Insulin
−0.20 −0.15 −0.10 −0.05 0.00 0.05
SHAP value (impact on model output)
SuperAger
High
Feature value
Low
AST
Free T4
Total cholesterol
Insulin
ApoE
C-peptide
HbAlc
Triglyceride
HDLCholesterol
BUN
−0.10 −0.05 0.00 0.05 0.10
SHAP value (impact on model output)
TypicalAger

도면9f


High
Feature value
Low
Hb
Triglyceride
oxLDL
K
Hct
AGE
Total cholesterol
MCH
CD36
HCHC
−0.150 −0.125 −0.100 −0.075 −0.050 −0.025 0.000 0.025
SHAP value (impact on model output)
SuperAger
High
Feature value
Low
Triglyceride
Hb
Total cholesterol
K
MCHC
MCH
oxLDL
AGE
Hct
CD36
−0.15 −0.10 −0.05 0.00 0.05
SHAP value (impact on model output)
TypicalAger