(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0099938
(43) 공개일자 2023년07월05일

(51) 국제특허분류(Int. Cl.)
G16H 10/20 (2018.01) *A61B 5/00* (2021.01)
G06N 3/08 (2023.01) *G16H 10/60* (2018.01)
G16H 50/20 (2018.01)
(52) CPC특허분류
G16H 10/20 (2021.08)
A61B 5/4806 (2013.01)
(21) 출원번호 10-2021-0189390
(22) 출원일자 2021년12월28일
심사청구일자 2021년12월28일

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
이재동
서울특별시 강남구 밤고개로23길 22(율현동)
이현훈
서울특별시 동대문구
(74) 대리인
위병갑

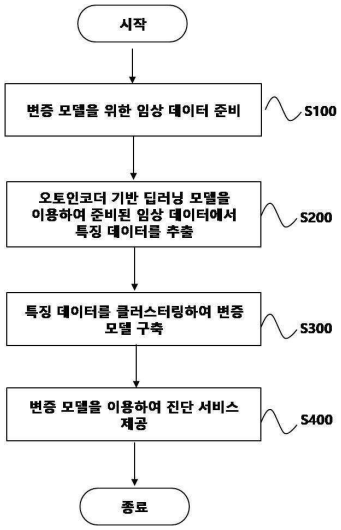
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **변증 진단 방법 및 그 시스템**

(57) 요 약

변증 진단 방법 및 그 시스템이 제공된다. 본 개시의 몇몇 실시예들에 따른 변증 진단 방법은, 오토인코더 기반 딥러닝 모델을 이용하여 임상 데이터에서 특징 데이터를 추출하고 추출된 특징 데이터를 클러스터링하여 변증 모델을 구축하는 단계를 포함할 수 있다. 이렇게 구축된 변증 모델은 환자의 변증 유형을 정확하고 객관적으로 판별할 수 있으며, 그 결과로 사용자에게 고품질의 한의학적 진단 서비스가 제공될 수 있다.

대 표 도 - 도2



(52) CPC특허분류
A61B 5/4854 (2013.01)
A61B 5/7264 (2013.01)
G06N 3/08 (2023.01)
G16H 10/60 (2021.08)
G16H 50/20 (2018.01)

명 세 서

청구범위

청구항 1

하나 이상의 컴퓨팅 장치에 의해 수행되는 변증 진단 방법으로서,

변증 모델을 위한 임상 데이터를 준비하는 단계;

오토인코더 기반 딥러닝 모델을 이용하여 상기 준비된 임상 데이터에서 특징 데이터를 추출하고, 상기 추출된 특징 데이터를 클러스터링하여 상기 변증 모델을 구축하는 단계; 및

상기 구축된 변증 모델을 이용하여 진단 서비스를 제공하는 단계를 포함하는,

변증 진단 방법.

청구항 2

제1항에 있어서,

상기 변증 모델은 수면 장애에 관한 변증 모델이고,

상기 임상 데이터는 환자의 인구통계학적 정보, 병력 정보, 수면 정보, 식이영양 정보 및 위장관 정보를 포함하는,

변증 진단 방법.

청구항 3

제1항에 있어서,

상기 임상 데이터를 준비하는 단계는,

복수개의 샘플을 포함하는 원시 임상 데이터를 획득하는 단계;

정상 데이터를 이용하여 오토인코더 기반의 다른 딥러닝 모델을 트레이닝하는 단계;

상기 트레이닝된 다른 딥러닝 모델을 이용하여 상기 복수개의 샘플에 대한 복원 손실(reconstruction loss) 값을 산출하는 단계;

상기 복수개의 샘플 중에서 상기 산출된 복원 손실 값이 기준치 이하인 샘플을 제거하는 단계;

상기 복수개의 샘플 중에서 상기 정상 데이터에 대한 클러스터 중심과의 거리가 기준치 이하인 샘플을 더 제거하는 단계;

상기 트레이닝된 다른 딥러닝 모델을 이용하여 상기 복수개의 샘플에 대한 잠재 벡터(latent vector)를 추출하는 단계; 및

상기 복수개의 샘플 중에서 상기 추출된 잠재 벡터에 대한 클러스터 중심과의 거리가 기준치 이하인 샘플을 더 제거하는 단계를 포함하는,

변증 진단 방법.

청구항 4

제1항에 있어서,

상기 임상 데이터를 준비하는 단계는,

복수개의 샘플을 포함하는 원시 임상 데이터를 획득하는 단계;

정상 데이터와 환자 데이터를 이용하여 오토인코더 기반의 다른 딥러닝 모델을 트레이닝하는 단계;

상기 트레이닝된 다른 딥러닝 모델을 이용하여 상기 정상 데이터와 상기 환자 데이터에 대한 잠재 벡터(latent vector)를 추출하는 단계;

정상 클래스와 환자 클래스를 분류하는 분류기를 이용하여 상기 추출된 잠재 벡터에 대한 분류 손실(classification loss) 값을 산출하는 단계;

상기 산출된 분류 손실 값을 이용하여 상기 분류기를 트레이닝하는 단계;

상기 트레이닝된 다른 딥러닝 모델의 디코더(freezing)를 프리징한 상태에서, 상기 산출된 분류 손실 값을 이용하여 상기 다른 딥러닝 모델의 인코더를 더 트레이닝하고, 상기 복수개의 샘플에 대한 잠재 벡터를 추출하는 단계;

상기 추출된 잠재 벡터를 상기 트레이닝된 분류기에 입력하여 환자 클래스에 대한 컨피던스 스코어를 획득하는 단계; 및

상기 복수개의 샘플 중에서 상기 획득된 컨피던스 스코어가 기준치 이하인 샘플을 제거하는 단계를 포함하는,

변증 진단 방법.

청구항 5

제1항에 있어서,

상기 특징 데이터는 상기 딥러닝 모델에 의해 상기 준비된 임상 데이터에서 추출된 잠재 벡터(latent vector)이거나 상기 잠재 벡터로부터 기인한 데이터이고,

상기 특징 데이터를 추출하는 단계는,

상기 딥러닝 모델을 트레이닝하는 단계; 및

상기 트레이닝된 딥러닝 모델을 이용하여 상기 잠재 벡터를 추출하는 단계를 포함하는,

변증 진단 방법.

청구항 6

제5항에 있어서,

상기 특징 데이터를 추출하는 단계는,

상기 추출된 잠재 벡터에 대해 주성분 분석을 수행하는 단계를 더 포함하는,

변증 진단 방법.

청구항 7

제5항에 있어서,

상기 딥러닝 모델은 인코더 및 디코더를 포함하고,

상기 딥러닝 모델을 트레이닝하는 단계는,

복원 손실(reconstruction loss) 값을 이용하여 상기 딥러닝 모델을 트레이닝하는 단계; 및

상기 디코더를 프리징(freezing)한 상태에서, 클러스터링 손실(clustering loss) 값을 이용하여 상기 인코더를 더 트레이닝하는 단계를 포함하며,

상기 클러스터링 손실 값은 상기 잠재 벡터에 대한 클러스터링 평가 점수에 기초하여 산출되는 것인,

변증 진단 방법.

청구항 8

제1항에 있어서,

상기 변증 모델을 구축하는 단계는,

상기 딥러닝 모델을 이용하여 제1 특징 데이터를 추출하는 단계;

상기 제1 특징 데이터를 클러스터링하여 제1 후보 변증 모델을 구축하는 단계;

상기 딥러닝 모델의 구성을 변경하여 상기 제1 특징 데이터와 차원수가 상이한 제2 특징 데이터를 추출하는 단계;

상기 제2 특징 데이터를 클러스터링하여 제2 후보 변증 모델을 구축하는 단계; 및

상기 제1 후보 변증 모델과 상기 제2 후보 변증 모델의 평가 결과에 기초하여 상기 구축된 후보 변증 모델 중에서 상기 변증 모델을 선정하는 단계를 포함하는,

변증 진단 방법.

청구항 9

제8항에 있어서,

상기 임상 데이터는 트레이닝용 데이터와 검증용 데이터를 포함하고,

상기 특징 데이터는 상기 트레이닝용 데이터에서 추출된 것이며,

상기 평가 결과는 상기 구축된 후보 변증 모델의 클러스터 별로 상기 검증용 데이터에 대한 평균과 분산을 산출하고 비교함으로써 얻어진 평가 결과를 포함하는,

변증 진단 방법.

청구항 10

제1항에 있어서,

상기 변증 모델을 구축하는 단계는,

상기 추출된 특징 데이터를 클러스터링하여 제1 후보 변증 모델을 구축하는 단계;

상기 추출된 특징 데이터를 클러스터링하여 상기 제1 후보 변증 모델과 클러스터 개수가 상이한 제2 후보 변증 모델을 구축하는 단계; 및

상기 제1 후보 변증 모델과 상기 제2 후보 변증 모델의 평가 결과에 기초하여 상기 구축된 후보 변증 모델 중에서 상기 변증 모델을 선정하는 단계를 포함하는,

변증 진단 방법.

청구항 11

제1항에 있어서,

상기 진단 서비스를 제공하는 단계는,

상기 구축된 변증 모델을 이용하여 환자의 변증 유형을 판별하는 단계; 및

상기 판별된 변증 유형에 따른 경혈(acupuncture point) 정보를 제공하는 단계를 포함하는,

변증 진단 방법.

청구항 12

하나 이상의 프로세서; 및

하나 이상의 인스트럭션을 저장하는 메모리를 포함하고,

상기 하나 이상의 프로세서는,

상기 저장된 하나 이상의 인스트럭션을 실행함으로써,

변증 모델을 위한 임상 데이터를 준비하는 동작,

오토인코더 기반 딥러닝 모델을 이용하여 상기 준비된 임상 데이터에서 특징 데이터를 추출하고, 상기 추출된 특징 데이터를 클러스터링하여 상기 변증 모델을 구축하는 동작 및

상기 구축된 변증 모델을 이용하여 진단 서비스를 제공하는 동작을 포함하는,

변증 진단 시스템.

청구항 13

컴퓨팅 장치와 결합되어,

변증 모델을 위한 임상 데이터를 준비하는 단계;

오토인코더 기반 딥러닝 모델을 이용하여 상기 준비된 임상 데이터에서 특징 데이터를 추출하고, 상기 추출된 특징 데이터를 클러스터링하여 상기 변증 모델을 구축하는 단계; 및

상기 구축된 변증 모델을 이용하여 진단 서비스를 제공하는 단계를 실행시키기 위하여 컴퓨터 판독가능한 기록매체에 저장된,

컴퓨터 프로그램.

발명의 설명

기 술 분 야

[0001] 본 개시는 변증 진단 방법 및 그 시스템에 관한 것으로, 보다 자세하게는, 변증 유형을 판별하는 모델을 구축하고 구축된 변증 모델을 이용하여 한의학적 진단 서비스를 제공하는 방법 및 그 방법을 수행하는 시스템에 관한 것이다.

배 경 기 술

[0002] 한의학은 예로부터 한국인의 몸을 진단하고 치유하던 의학으로서 서양 의학과는 다르게 우리의 몸에 맞는 데이터가 수백 년에 걸쳐 쌓여 왔으며, 한국인 고유의 신체적 특성에 맞게 병을 진단하고 처방할 수 있는 과학적인 의학이다. 한의학은 서양 의학과는 다르게 널리 보편화가 되지 않아 일반인에게는 거리감이 있었으나, 최근 들어 대체의학에 대한 관심이 높아지면서 세계적으로 관심이 커지고 있다.

[0003] 한편, 침구 치료는 한의학에서 이용되는 주된 치료 방법 중 하나인데, 한의사는 보통 망문문절(望聞問切) 등을 통해 환자의 증상 정보를 수집하고, 수집된 증상 정보를 토대로 변증(pattern identification)을 수행하며, 변증 결과를 기초로 침구 치료가 행해질 경혈(acupuncture point)을 선택하게 된다.

[0004] 그런데, 대부분의 질병은 복합적이고 복잡한 증상을 일으키기 때문에, 숙련되지 않은 한의사가 변증을 정확하게 수행하는 것은 매우 어려운 일이다. 뿐만 아니라, 증상에 대한 환자의 주관적 표현으로 인해 숙련된 한의사라고 할지라도 변증을 정확하게 수행하는 것이 쉽지 않으며, 변증 결과의 부정확성은 침구 치료의 효과를 떨어뜨리는 주된 요인이 될 수 있다.

[0005] 또한, 변증 과정 중에 한의사의 주관적 판단이 개입될 수 있기 때문에, 같은 환자라도 경우에 따라 변증 결과가 다를 수 있고, 이는 궁극적으로 침구 치료에 대한 환자의 신뢰도를 떨어뜨릴 수 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국공개특허 제10-2020-0022113호(20.03.03. 공개)

발명의 내용

해결하려는 과제

[0007] 본 개시의 몇몇 실시예들을 통해 해결하고자 하는 기술적 과제는, 환자의 변증 유형을 정확하고 객관적으로 판

별하기 위한 변증 모델을 구축할 수 있는 방법 및 그 방법을 수행하는 시스템을 제공하는 것이다.

[0008] 본 개시의 몇몇 실시예들을 통해 해결하고자 하는 다른 기술적 과제는, 주어진 임상 데이터에서 변증 유형을 판별하기 위한 특징 데이터를 정확하게 추출할 수 있는 방법 및 그 방법을 수행하는 시스템을 제공하는 것이다.

[0009] 본 개시의 몇몇 실시예들을 통해 해결하고자 하는 또 다른 기술적 과제는, 주어진 임상 데이터에서 이상치를 정확하게 제거할 수 있는 방법 및 그 방법을 수행하는 시스템을 제공하는 것이다.

[0010] 본 개시의 몇몇 실시예들을 통해 해결하고자 하는 또 다른 기술적 과제는, 구축된 변증 모델을 이용하여 진단 서비스를 제공하는 방법 및 그 방법을 수행하는 시스템을 제공하는 것이다.

[0011] 본 개시의 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 개시의 기술분야에서의 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0012] 상기 기술적 과제를 해결하기 위한, 본 개시의 몇몇 실시예들에 따른 변증 진단 방법은, 하나 이상의 컴퓨팅 장치에 의해 수행되는 변증 진단 방법으로서, 변증 모델을 위한 임상 데이터를 준비하는 단계, 오토인코더 기반 딥러닝 모델을 이용하여 상기 준비된 임상 데이터에서 특징 데이터를 추출하고, 상기 추출된 특징 데이터를 클러스터링하여 상기 변증 모델을 구축하는 단계 및 상기 구축된 변증 모델을 이용하여 진단 서비스를 제공하는 단계를 포함할 수 있다.

[0013] 몇몇 실시예들에서, 상기 변증 모델은 수면 장애에 관한 변증 모델이고, 상기 임상 데이터는 환자의 인구통계학적 정보, 병력 정보, 수면 정보, 식이영양 정보 및 위장관 정보를 포함할 수 있다.

[0014] 몇몇 실시예들에서, 상기 특징 데이터는 상기 딥러닝 모델에 의해 상기 준비된 임상 데이터에서 추출된 잠재 벡터(latent vector)이거나 상기 잠재 벡터로부터 기인한 데이터이고, 상기 특징 데이터를 추출하는 단계는, 상기 딥러닝 모델을 트레이닝하는 단계 및 상기 트레이닝된 딥러닝 모델을 이용하여 상기 잠재 벡터를 추출하는 단계를 포함할 수 있다.

[0015] 몇몇 실시예들에서, 상기 진단 서비스를 제공하는 단계는, 상기 구축된 변증 모델을 이용하여 환자의 변증 유형을 판별하는 단계 및 상기 판별된 변증 유형에 따른 경혈(acupuncture point) 정보를 제공하는 단계를 포함할 수 있다.

[0016] 상술한 기술적 과제를 해결하기 위한, 본 개시의 몇몇 실시예들에 따른 변증 진단 시스템은, 하나 이상의 프로세서 및 하나 이상의 인스트럭션을 저장하는 메모리를 포함하고, 상기 하나 이상의 프로세서는, 상기 저장된 하나 이상의 인스트럭션을 실행함으로써, 변증 모델을 위한 임상 데이터를 준비하는 동작, 오토인코더 기반 딥러닝 모델을 이용하여 상기 준비된 임상 데이터에서 특징 데이터를 추출하고, 상기 추출된 특징 데이터를 클러스터링하여 상기 변증 모델을 구축하는 동작 및 상기 구축된 변증 모델을 이용하여 진단 서비스를 제공하는 동작을 포함할 수 있다.

[0017] 상술한 기술적 과제를 해결하기 위한, 본 개시의 몇몇 실시예들에 따른 컴퓨터 프로그램은, 컴퓨팅 장치와 결합되어, 변증 모델을 위한 임상 데이터를 준비하는 단계, 오토인코더 기반 딥러닝 모델을 이용하여 상기 준비된 임상 데이터에서 특징 데이터를 추출하고, 상기 추출된 특징 데이터를 클러스터링하여 상기 변증 모델을 구축하는 단계 및 상기 구축된 변증 모델을 이용하여 진단 서비스를 제공하는 단계를 실행시키기 위하여 컴퓨터 판독 가능한 기록매체에 저장될 수 있다.

발명의 효과

[0018] 상술한 본 개시의 몇몇 실시예들에 따르면, 오토인코더 기반의 딥러닝 모델을 통해 임상 데이터에서 특징 데이터가 추출되고, 추출된 특징 데이터를 클러스터링함으로써 변증 모델이 구축될 수 있다. 오토인코더 기반의 딥러닝 모델은 임상 데이터를 종합적으로 고려하여 노이즈를 제거하고 특징적인 부분만을 압축하여 특징 데이터로 제공할 수 있고, 클러스터링 기법은 제공된 특징 데이터를 유형 별로 잘 구분할 수 있다. 이에 따라, 변증 유형을 객관적이고 정확하게 판별할 수 있는 고성능의 변증 모델이 구축될 수 있으며, 구축된 변증 모델을 이용하여 사용자에게 고품질의 진단 서비스가 제공될 수 있다.

[0019] 또한, 오토인코더 기반의 딥러닝 모델을 이용하여 원시 임상 데이터에서 불필요한 샘플(e.g. 정상 샘플, 정상에 가까운 샘플, 이상치 샘플 등)이 정확하게 제거될 수 있다. 그리고, 정제된 임상 데이터를 이용하여 변증 모델

을 구축함으로써, 변증 모델의 정확도가 더욱 향상될 수 있다.

[0020] 또한, 특징 데이터의 차원수, 클러스터의 개수 등을 변경해가며 다수의 후보 변증 모델을 구축하고, 후보 변증 모델을 다양한 방식으로 평가함으로써 고성능의 변증 모델이 용이하게 구축될 수 있다.

[0021] 본 개시의 기술적 사상에 따른 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0022] 도 1은 본 개시의 몇몇 실시예들에 따른 변증 진단 시스템과 그의 서비스 제공 환경을 설명하기 위한 예시적인 도면이다.

도 2는 본 개시의 몇몇 실시예들에 따른 변증 진단 방법을 개략적으로 나타내는 예시적인 흐름도이다.

도 3은 도 2에 도시된 임상 데이터 준비 단계 S100을 부연 설명하기 위한 예시적인 흐름도이다.

도 4 및 도 5는 도 2에 도시된 특징 데이터 추출 단계 S200을 부연 설명하기 위한 예시적인 도면이다.

도 6은 도 2에 도시된 변증 모델 구축 단계 S300을 부연 설명하기 위한 예시적인 도면이다.

도 7 및 도 8은 본 개시의 몇몇 실시예들에 따른 데이터 정제 방법을 설명하기 위한 예시적인 도면이다.

도 9는 본 개시의 다른 몇몇 실시예들에 따른 데이터 정제 방법을 설명하기 위한 예시적인 도면이다.

도 10은 본 개시의 몇몇 실시예들에 따른 변증 모델 구축 방법을 설명하기 위한 예시적인 흐름도이다.

도 11은 본 개시의 다른 몇몇 실시예들에 따른 변증 모델 구축 방법을 설명하기 위한 예시적인 흐름도이다.

도 12는 본 개시의 몇몇 실시예들에 따른 변증 진단 시스템을 구현할 수 있는 예시적인 컴퓨팅 장치를 도시한다.

도 13은 본 개시의 몇몇 구현예들에 따른 변증 모델의 클러스터 개수 결정 과정을 설명하기 위한 예시적인 도면이다.

발명을 실시하기 위한 구체적인 내용

[0023] 이하, 첨부된 도면을 참조하여 본 개시의 바람직한 실시예들을 상세히 설명한다. 본 개시의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 개시의 기술적 사상은 이하의 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 이하의 실시예들은 본 개시의 기술적 사상을 완전하도록 하고, 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자에게 본 개시의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 개시의 기술적 사상은 청구항의 범주에 의해 정의될 뿐이다.

[0024] 각 도면의 구성요소들에 참조부호를 부가함에 있어서, 동일한 구성요소들에 대해서는 비록 다른 도면상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 개시를 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 개시의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.

[0025] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어(기술 및 과학적 용어를 포함)는 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 사용될 수 있다. 또 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다. 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 개시를 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다.

[0026] 또한, 본 개시의 구성 요소를 설명하는 데 있어서, 제1, 제2, A, B, (a), (b) 등의 용어를 사용할 수 있다. 이러한 용어는 그 구성 요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 용어에 의해 해당 구성 요소의 본질이나 차례 또는 순서 등이 한정되지 않는다. 어떤 구성 요소가 다른 구성요소에 "연결", "결합" 또는 "접속"된다고 기재된 경우, 그 구성 요소는 그 다른 구성요소에 직접적으로 연결되거나 또는 접속될 수 있지만, 각 구성 요소 사이에 또 다른 구성 요소가 "연결", "결합" 또는 "접속"될 수도 있다고 이해되어야 할 것이다.

[0027] 본 개시에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성 요소, 단계, 동작

및/또는 소자는 하나 이상의 다른 구성 요소, 단계, 동작 및/또는 소자의 존재 또는 추가를 배제하지 않는다.

[0028] 이하, 본 개시의 다양한 실시예들에 대하여 첨부된 도면에 따라 상세하게 설명한다.

[0029] 도 1은 본 개시의 몇몇 실시예들에 따른 변증 진단 시스템(10)과 그의 서비스 제공 환경을 설명하기 위한 예시적인 도면이다.

[0030] 도 1에 도시된 바와 같이, 변증 진단 시스템(10)은 변증(pattern identification)에 기반한 한의학적 진단을 수행하거나 한의학적 진단 서비스를 제공하는 시스템일 수 있다. 여기서, 변증(辨證)이란 질병으로 인해 복합적으로 나타나는 여러 병상(病狀)으로부터 한의학적 진단의 기초가 되는 증(證)을 분별하는 행위, 과정 또는 그 행위의 결과를 의미할 수 있다. 이하에서는, 설명의 편의상, 변증 진단 시스템(10)을 "진단 시스템(10)"으로 약칭하도록 한다.

[0031] 구체적으로, 진단 시스템(10)은 딥러닝 기법을 통해 임상 데이터(clinic data)에서 변증을 위한 특징 데이터를 추출하고, 추출된 특징 데이터에 기반하여 변증 모델을 구축할 수 있다. 또한, 진단 시스템(10)은 구축된 변증 모델을 이용하여 사용자(e.g. 환자)에게 진단 서비스를 제공할 수 있다. 가령, 진단 시스템(10)은 딥러닝 기법을 통해 수면 장애와 연관된 임상 데이터에서 특징 데이터를 추출하고, 추출된 특징 데이터에 기반하여 수면 장애의 변증 유형을 판별하는 모델을 구축할 수 있다. 또한, 도시된 바와 같이, 진단 시스템(10)은 사용자 단말(1)의 진단 요청에 응답하여 구축된 변증 모델에 기반한 진단 서비스를 제공할 수 있다. 이를테면, 진단 시스템(10)은 변증 모델을 통해 사용자의 변증 유형을 판별하고, 판별된 변증 유형에 따른 경혈(acupuncture point) 정보를 사용자 단말(1)에게 제공할 수 있다. 진단 시스템(10)이 변증 모델을 구축하는 방법에 관하여서는 추후 도 2 이하의 도면을 참조하여 상세하게 설명하도록 한다.

[0032] 진단 시스템(10)은 하나 이상의 컴퓨팅 장치로 구현될 수 있다. 다시 말해, 진단 시스템(10)은 단일 컴퓨팅 장치로 구현될 수도 있고, 복수의 컴퓨팅 장치로 구현될 수도 있다. 구체적인 예를 들면, 진단 시스템(10)은 복수의 컴퓨팅 장치로 구현되되, 진단 시스템(10)의 제1 기능은 제1 컴퓨팅 장치에서 구현되고 제2 기능은 제2 컴퓨팅 장치에서 구현될 수 있다. 또는, 진단 시스템(10)의 특정 기능이 복수의 컴퓨팅 장치에서 구현될 수도 있다.

[0033] 컴퓨팅 장치는 예를 들어 노트북, 데스크톱(desktop), 랩탑(laptop) 등이 될 수 있으나, 이에 국한되는 것은 아니며 컴퓨팅 기능이 구비된 모든 종류의 장치를 포함할 수 있다. 컴퓨팅 장치의 일 예시는 도 12를 참조하도록 한다.

[0034] 사용자 단말(1)은 사용자(e.g. 환자)가 소지한 단말로서 어떠한 컴퓨팅 장치로 구현되더라도 무방하다.

[0035] 도시된 바와 같이, 사용자 단말(1)과 진단 시스템(10)은 네트워크를 통해 통신할 수 있다. 여기서, 네트워크는 근거리 통신망(Local Area Network; LAN), 광역 통신망(Wide Area Network; WAN), 이동 통신망(mobile radio communication network), Wibro(Wireless Broadband Internet) 등과 같은 모든 종류의 유/무선 네트워크로 구현될 수 있다.

[0036] 지금까지 도 1을 참조하여 본 개시의 몇몇 실시예들에 따른 진단 시스템(10)에 대하여 설명하였다. 이하에서는, 도 1에 예시된 진단 시스템(10)에서 수행될 수 있는 변증 진단 방법에 대하여 설명하도록 한다.

[0037] 이하에서 후술될 변증 진단 방법은 하나 이상의 컴퓨팅 장치에 의해 수행될 수 있다. 다시 말해, 후술될 변증 진단 방법은 컴퓨팅 장치의 프로세서에 의해 실행될 수 있는 하나 이상의 인스트럭션(instruction)으로 구현될 수 있다. 이하에서는, 이해의 편의를 제공하기 위해 후술될 변증 진단 방법의 모든 단계가 도 1에 예시된 진단 시스템(10)에서 수행되는 것을 가정하여 설명을 이어가도록 한다. 따라서, 특정 단계(동작)의 주어가 생략된 경우, 도 1의 진단 시스템(10)에 의해 수행되는 것으로 이해될 수 있다. 물론, 경우에 따라 후술될 변증 진단 방법의 일부 단계는 다른 컴퓨팅 장치에서 수행될 수도 있다.

[0038] 도 2는 본 개시의 몇몇 실시예들에 따른 변증 진단 방법을 개략적으로 나타내는 예시적인 흐름도이다. 단, 이는 본 개시의 목적을 달성하기 위한 바람직한 실시예일뿐이며, 필요에 따라 일부 단계가 추가되거나 삭제될 수 있음은 물론이다.

[0039] 도 2에 도시된 바와 같이, 본 실시예는 변증 모델을 위한 임상 데이터를 준비하는 단계 S100에서 시작될 수 있다. 본 단계의 세부 과정은 도 3에 도시되어 있다.

[0040] 도 3에 도시된 바와 같이, 임상 데이터 준비 단계 S100은 원시(raw) 임상 데이터를 획득하는 단계 S110과 데이터 정제 단계 S120를 포함할 수 있다.

[0041] 단계 S110에서, 원시 임상 데이터를 획득하는 방식은 어떠한 방식이 되더라도 무방하다.

[0042] 원시 임상 데이터는 다양한 유형의 정보를 포함할 수 있다. 다만, 그 구체적인 정보의 유형은 대상 질병 또는 실시예에 따라 달라질 수 있다.

[0043] 예를 들어, 대상 질병이 수면 장애(sleep disturbance)라고 가정하자. 다시 말해, 구축하고자 하는 변증 모델이 수면 장애의 변증 유형을 판별하기 위한 모델이라고 가정하자. 이러한 경우, 원시 임상 데이터는 환자의 인구통계학적 정보, 근무 정보, 병력 정보, 수면 정보, 식이영양 정보 및 위장관 정보 등을 포함할 수 있다. 인구통계학적 특성, 근무 특성, 병력, 식습관, 영양 상태, 위장관 상태 등은 모두 환자의 수면 상태에 큰 영향을 미치는 요인이기 때문에, 이러한 정보로 구성된 원시 임상 데이터를 이용하면 정확도 높은 변증 모델이 구축될 수 있다.

[0044] 인구통계학적 정보는 예를 들어 인종, 생년월일, 키, 몸무게, 학력, 신체 등급(e.g. 신체 나이), 흡연력, 음주력, 운동력 등의 정보를 포함할 수 있을 것이나, 이에 한정되는 것은 아니다. 환자의 소속이 군인인 경우, 인구통계학적 정보는 입대일, 군번, 계급, 소속부대 등의 정보를 더 포함할 수도 있다.

[0045] 다음으로, 근무 정보는 예를 들어 근무 유형, 야간 근무 형태(e.g. 근무 취침이 있는 근무와 없는 근무), 야간 근무 횟수, 야간 근무 시간, 야간 근무가 수면 및/또는 피로도에 미치는 영향, 주간 근무 횟수, 주간 근무 시간 등의 정보를 포함할 수 있을 것이나, 이에 한정되는 것은 아니다. 환자의 소속이 군인인 경우, 근무 정보는 병과, 보직 등의 정보를 더 포함할 수도 있다.

[0046] 다음으로, 병력 정보는 예를 들어 수면 장애, 위장관 장애, 고혈압, 당뇨, 고지혈증, 심장질환 등에 관한 병력, 스트레스 상태, 약물력 등의 정보를 포함할 수 있을 것이나, 이에 한정되는 것은 아니다.

[0047] 다음으로, 수면 정보는 예를 들어 PSQI(Pittsburgh Sleep Quality Index), 베를린(Berlin) 설문 결과, 수면 환경 등의 정보를 포함할 수 있을 것이나, 이에 한정되는 것은 아니다.

[0048] 다음으로, 식이영양 정보는 예를 들어 식습관, 영양지수(NQ; Nutrition Quotient) 등의 정보를 포함할 수 있을 것이나, 이에 한정되는 것은 아니다.

[0049] 다음으로, 위장관 정보는 예를 들어 GSRS(Gastrointestinal Symptom Rating Scale), BSS(Bristol Stool Scale) 등의 정보를 포함할 수 있을 것이나, 이에 한정되는 것은 아니다.

[0050] 단계 S120에서, 원시 임상 데이터에 대한 정제(cleaning)가 수행될 수 있다. 가령, 진단 시스템(10)은 원시 임상 데이터를 구성하는 복수의 샘플 중에서 중복 샘플, 결측치가 존재하는 샘플, 이상치가 존재하는 샘플 등을 제거할 수 있다. 또는, 진단 시스템(10)은 원시 임상 데이터에서 정상(즉, 정상인의 샘플) 또는 정상에 가까운 샘플을 제거할 수 있다. 구체적인 예로써, 수면 장애에 관한 변증 모델을 구축하는 경우, 진단 시스템(10)은 PSQI가 약 5 이하인 샘플을 제거할 수 있다.

[0051] 몇몇 실시예들에서는, 오토인코더 기반의 딥러닝 모델을 이용하여 원시 임상 데이터가 정제될 수 있다. 예를 들어, 진단 시스템(10)은 정상 데이터로 트레이닝된 오토인코더 기반의 딥러닝 모델을 이용하여 불필요한 샘플(e.g. 정상 또는 정상에 가까운 샘플, 이상치 샘플 등)을 제거할 수 있다. 그렇게 함으로써, 변증 모델 구축에 이용되는 임상 데이터의 품질이 향상되고, 궁극적으로 변증 모델의 정확도가 크게 향상될 수 있다. 본 실시예와 관련하여서는 추후 도 7 내지 도 9를 참조하여 상세하게 설명하도록 한다.

[0052] 몇몇 실시예들에서는, 오토인코더 기반의 딥러닝 모델을 이용하여 샘플의 결측치가 보정될 수 있다. 구체적으로, 진단 시스템(10)은 환자 데이터(즉, 복수개의 환자 샘플로 구성된 데이터)를 이용하여 오토인코더 기반의 딥러닝 모델을 트레이닝하고, 트레이닝된 딥러닝 모델을 이용하여 샘플의 결측치를 추정할 수 있다. 가령, 진단 시스템(10)은 샘플의 결측치에 서로 다른 값을 할당해가며 트레이닝된 딥러닝 모델을 통해 해당 샘플에 대한 복원 손실(reconstruction loss) 값을 반복적으로 산출할 수 있다. 그리고, 진단 시스템(10)은 복원 손실 값이 기준치 이하인 경우의 할당 값으로 해당 샘플의 결측치를 보정할 수 있다. 복원 손실 값이 작다는 것은 딥러닝 모델이 해당 샘플을 잘 복원한다는 것을 의미하고, 이는 할당된 값이 환자 샘플에 적합할 확률이 높다는 것을 의미하기 때문이다. 복원 손실 값이 기준치 이하로 떨어지지 않는 경우라면, 진단 시스템(10)은 원시 임상 데이터에서 해당 샘플을 제거할 수 있다. 다른 예로서, 결측치 샘플의 복원 손실 값(e.g. 결측치에 0을 할당하여 산출된 복원 손실 값)과 결측치에 특정 값을 할당한 경우의 복원 손실 값과의 차이가 기준치 이상인 경우, 진단 시스템(10)은 할당된 값으로 해당 샘플의 결측치를 보정할 수도 있다. 본 실시예에 따르면, 환자 데이터의 특성을 학습한 오토인코더 기반의 딥러닝 모델을 이용하여 샘플의 결측치가 정확하게 보정될 수 있다.

[0053] 참고로, 도 2 또는 도 3에 명확하게 도시되어 있지 않으나, 딥러닝 모델을 이용하기 전에, 원시 임상 데이터 또는 데이터 정제를 통해 준비된 임상 데이터는 수치화된 형태로 변환될 수 있다. 가령, 임상 데이터는 당해 기술 분야에서 널리 알려진 벡터화 기법을 통해 수치형 벡터로 변환될 수 있다.

[0054] 다시 도 2를 참조하여 설명한다.

[0055] 단계 S200에서, 오토인코더 기반 딥러닝 모델을 이용하여 준비된 임상 데이터에서 특징 데이터가 추출될 수 있다. 여기서, 오토인코더 기반 딥러닝 모델은 기본 오토인코더(Auto-Encoder)와 그의 다양한 변형물(variation)을 포함하는 것일 수 있다. 예를 들어, 오토인코더 기반 딥러닝 모델은 기본 오토인코더, 변이형 오토인코더(Variational Auto-Encoder; VAE) 등을 포함할 수 있다.

[0056] 이해의 편의를 제공하기 위해, 본 단계 S200에 대한 구체적인 설명에 앞서, 도 4를 참조하여 기본 오토인코더 구조의 딥러닝 모델에 대하여 간략하게 설명하도록 한다. 도 4는 잠재 벡터(23; 또는 병목 레이어)의 차원수(d)가 3인 오토인코더 구조를 예시하고 있다.

[0057] 도 4에 도시된 바와 같이, 기본 오토인코더 구조의 딥러닝 모델은 인코더(22) 및 디코더(24)로 구성될 수 있으며, 입력 데이터(21)를 압축(인코딩) 및 복원(디코딩)하는 동작을 수행할 수 있다. 구체적으로, 인코더(22)는 입력 데이터(21)를 인코딩(압축)하여 잠재 벡터(23)를 출력할 수 있고, 디코더(24)는 잠재 벡터(23)를 디코딩하여 복원된 데이터(25)를 출력할 수 있다. 다른 방식으로 표현하면, 인코더(22)는 입력 데이터(21)를 잠재 공간(latent space) 상의 벡터(23)로 매핑시킬 수 있고, 디코더(24)는 역으로 잠재 벡터(23)로부터 매핑된 데이터(25)를 복원할 수 있다. 인코더(22) 및 디코더(24)는 입력 데이터(21)와 복원된 데이터(25) 간의 복원 손실(reconstruction loss) 값을 이용하여 트레이닝됨으로써, 이러한 매핑 관계를 학습할 수 있다(이른바 "매니폴드 러닝"). 당해 기술 분야의 종사자라면, 오토인코더의 구성 및 동작 원리에 대해 이미 숙지하고 있을 것인 바, 더 이상의 자세한 설명은 생략하도록 한다.

[0058] 이하, 단계 S200의 세부 과정에 대하여 상세하게 설명하되, 도 4에 예시된 딥러닝 모델을 통해 특징 데이터를 추출하는 것을 가정하여 설명하도록 한다.

[0059] 먼저, 준비된 임상 데이터를 이용하여 오토인코더 기반 딥러닝 모델이 트레이닝될 수 있다. 가령, 진단 시스템(10)은 임상 데이터를 구성하는 복수의 샘플(e.g. 21)에 대한 복원 손실 값을 이용하여 딥러닝 모델을 트레이닝할 수 있다. 트레이닝 과정 중에, 인코더(22)와 디코더(24)의 가중치 파라미터는 복원 손실 값을 감소시키는 방향으로 업데이트될 수 있다. 모델의 트레이닝은 정해진 기준(e.g. 시간, 에폭, 손실 값 등에 기초하여 정해진 기준)을 만족할 때까지 반복적으로 수행될 수 있으며, 트레이닝이 완료되면 임상 데이터를 구성하는 각 샘플(e.g. 21)은 잠재 공간 상에 매핑될 수 있고 유사한 샘플은 유사한 위치에 매핑되게 된다.

[0060] 다음으로, 트레이닝된 딥러닝 모델을 이용하여 임상 데이터를 구성하는 각 샘플의 잠재 벡터가 추출될 수 있다. 가령, 도 5에 도시된 바와 같이, 진단 시스템(10)은 각 샘플(26)을 트레이닝된 딥러닝 모델의 인코더(22)에 입력하여 잠재 벡터(27)를 추출할 수 있다.

[0061] 이렇게 추출된 잠재 벡터(27) 또는 잠재 벡터(27)로부터 기인한 데이터가 특징 데이터로 이용될 수 있다. 잠재 벡터(27)는 입력된 샘플(26)의 특징적인 부분만을 압축한 데이터이기 때문이다. 가령, 특정 샘플(26)에서 추출된 잠재 벡터(27)가 해당 샘플(26)의 특징 벡터가 될 수 있고, 임상 데이터에서 추출된 특징 데이터는 하나 이상의 특징 벡터를 칭하는 것일 수 있다.

[0062] 몇몇 실시예들에서는, 복수의 잠재 벡터(e.g. 27)에 대해 주성분 분석(PCA; Principal Component Analysis)이 수행될 수 있고, 주성분 분석을 통해 차원이 축소된 벡터가 특징 데이터가 될 수 있다. 또는, 임상 데이터에 대해 주성분 분석이 먼저 수행되고, 그 이후에 오토 인코더 기반의 딥러닝 모델을 통해 잠재 벡터가 추출될 수도 있다.

[0063] 한편, 몇몇 실시예들에서는, 오토인코더 기반의 딥러닝 모델이 복원 손실 값뿐만 아니라 클러스터링 손실(clustering loss) 값을 더 이용하여 트레이닝될 수 있다. 가령, 진단 시스템(10)은 디코더(e.g. 24)를 프리징(freezing)한 상태에서 잠재 벡터(e.g. 23)에 대한 클러스터링 손실 값을 이용하여 인코더(e.g. 22)를 더 트레이닝할 수 있다. 이러한 경우, 딥러닝 모델의 인코더(e.g. 22)가 클러스터링에 적합한 잠재 벡터(e.g. 27)를 추출하도록 학습될 수 있으며, 이에 따라 보다 향상된(e.g. 클러스터 내 응집도와 클러스터 간 분리도가 향상된) 성능(e.g. 정확도)을 갖는 변증 모델이 구축될 수 있다. 본 실시예에서, 클러스터링 손실 값은 잠재 벡터(e.g. 27)에 대한 클러스터링 평가 점수에 기초하여 산출될 수 있으며, 구체적인 산출 방식은 어떠한 방식이 되더라도

무방하다.

[0064] 다시 도 2를 참조하여 설명한다.

[0065] 단계 S300에서, 추출된 특징 데이터를 클러스터링하여 변증 모델이 구축될 수 있다. 가령, 도 6에 도시된 바와 같이, 진단 시스템(10)은 추출된 특징 데이터를 클러스터링함으로써 하나 이상의 클러스터(31 내지 33)를 포함하는 변증 모델을 구축할 수 있다. 여기서, 각 클러스터(31 내지 33)는 특정 변증 유형에 대응되는 것일 수 있다. 진단 시스템(10)은 K-평균 클러스터링 기법과 같이 당해 기술 분야에서 널리 알려진 클러스터링 기법을 이용하여 추출된 특징 데이터를 클러스터링할 수 있을 것이나, 본 개시의 범위가 이에 한정되는 것은 아니다.

[0066] 한편, 몇몇 실시예들에서는, 진단 시스템(10)이 복수개의 후보 변증 모델을 구축하고, 구축된 후보 변증 모델을 평가하며, 평가 결과에 기초하여 후보 변증 모델 중에서 서비스용 변증 모델을 선정할 수 있다. 예를 들어, 진단 시스템(10)은 오토인코더 기반 딥러닝 모델의 구성 변경을 통해 서로 다른 차원수를 갖는 특징 데이터를 추출하고, 추출된 특징 데이터를 클러스터링하여 복수의 후보 변증 모델을 구축할 수 있다. 또는, 진단 시스템(10)은 클러스터의 개수를 변경해가며 복수의 후보 변증 모델을 구축할 수 있다. 그렇게 함으로써, 정확도 높은 변증 모델이 용이하게 구축될 수 있다. 본 실시예에 관하여서는 추후 도 10 및 도 11을 참조하여 보다 상세하게 설명하도록 한다.

[0067] 단계 S400에서, 구축된 변증 모델을 이용하여 한의학적 진단 서비스가 제공될 수 있다. 예를 들어, 진단 시스템(10)은 구축된 변증 모델을 이용하여 환자의 변증 유형을 판별할 수 있고, 판별된 변증 유형에 따른 경혈 정보를 해당 환자에게 제공할 수 있다. 보다 구체적으로, 진단 시스템(10)은 변증 모델을 구성하는 복수의 클러스터 중에서 환자 데이터에서 추출된 특징 데이터에 대응되는 클러스터를 결정하고, 결정된 클러스터의 변증 유형을 환자의 변증 유형으로 결정할 수 있다.

[0068] 지금까지 도 2 내지 도 6을 참조하여 본 개시의 몇몇 실시예들에 따른 변증 진단 방법에 대하여 설명하였다. 상술한 방법에 따르면, 오토인코더 기반의 딥러닝 모델을 통해 임상 데이터에서 특징 데이터가 추출되고, 추출된 특징 데이터를 클러스터링함으로써 변증 모델이 구축될 수 있다. 오토인코더 기반의 딥러닝 모델은 임상 데이터를 종합적으로 고려하여 노이즈를 제거하고 특징적인 부분만을 압축하여 특징 데이터로 제공할 수 있고, 클러스터링 기법은 제공된 특징 데이터를 유형 별로 잘 구분할 수 있다. 이에 따라, 변증 유형을 객관적이고 정확하게 판별할 수 있는 고성능의 변증 모델이 구축될 수 있으며, 구축된 변증 모델을 이용하여 사용자에게 고품질의 진단 서비스가 제공될 수 있다.

[0069] 이하에서는, 도 7 내지 도 9를 참조하여 데이터 정제 방법에 관한 다양한 실시예들에 대하여 설명하도록 한다.

[0070] 먼저, 도 7 및 도 8을 참조하여 본 개시의 몇몇 실시예들에 따른 데이터 정제 방법에 대하여 설명하도록 한다.

[0071] 도 7 및 도 8에 도시된 바와 같이, 본 실시예는 오토인코더 기반의 딥러닝 모델(40)을 이용하여 임상 데이터를 정제하는 방법에 관한 것이다. 보다 자세하게는, 정상 데이터(44)로 트레이닝된 딥러닝 모델(40)을 이용하여 원시 임상 데이터에서 대상 질병과 관련 없는 정상 샘플(또는 이상치 샘플 등)을 제거하는 방법에 관한 것이다. 이하에서는, 이해의 편의를 제공하기 위해, 딥러닝 모델(40)이 기본 오토인코더 구조로 이루어진 것을 가정하여 설명을 이어가도록 한다. 다만, 본 개시의 범위가 이에 한정되는 것은 아니다.

[0072] 도 7에 도시된 바와 같이, 딥러닝 모델(40)은 인코더(41) 및 디코더(42)를 포함하도록 구성될 수 있고, 정상 데이터(44)를 이용하여 트레이닝될 수 있다. 가령, 진단 시스템(10)은 정상 데이터(44)를 구성하는 각각의 샘플(45)을 딥러닝 모델(40)에 입력하고, 입력된 샘플(45)과 복원된 데이터(46) 간의 복원 손실 값을 이용하여 딥러닝 모델(40)을 트레이닝할 수 있다. 이때, 딥러닝 모델(40)의 내부에서는 인코더(41)를 통해 입력된 샘플(45)을 잠재 벡터(43)로 인코딩하고, 디코더(42)를 통해 잠재 벡터(43)를 디코딩하는 과정이 수행될 수 있고, 복원 손실 값이 역전파(back-propagation)됨에 따라 인코더(41) 및 디코더(42)의 가중치 파라미터가 업데이트될 수 있다. 이러한 트레이닝 과정이 반복됨에 따라, 딥러닝 모델(40)은 입력된 정상 샘플(e.g. 45)을 압축(인코딩)하고 다시 복원(디코딩)하는 능력을 구비하게 된다.

[0073] 다음으로, 도 8에 도시된 바와 같이, 트레이닝된 딥러닝 모델(40)을 이용하여 원시 임상 데이터(47)를 구성하는 각각의 샘플(48)에 대해 복원 손실 값이 산출될 수 있다. 그리고, 산출된 복원 손실 값에 기초하여 데이터 정제가 이루어질 수 있다. 가령, 진단 시스템(10)은 복원 손실 값이 기준치 이하인 샘플을 원시 임상 데이터(47)에서 제거할 수 있다. 복원 손실 값이 기준치 이하라는 것은 딥러닝 모델(40)을 통해 잘 복원되는 샘플이라는 것을 의미하고, 이는 곧 해당 샘플이 정상 샘플(또는 이상치 샘플 등)일 확률이 높다는 것을 의미하기 때문이다.

[0074] 상기 기준치는 미리 설정된 고정 값 또는 상황에 따라 변동되는 변동 값일 수 있다. 예를 들어, 기준치는 정상 샘플의 복원 손실 값과 환자 샘플의 복원 손실 값의 평균(e.g. 가중평균 등)에 기초하여 미리 설정된 값일 수 있다. 다른 예로서, 기준치는 원시 임상 데이터의 샘플 개수, 변증 모델의 목표 성능(e.g. 정확도), 원시 임상 데이터의 신뢰도(e.g. 설문지 응답자의 신뢰도)에 기초하여 변동되는 값일 수도 있다.

[0075] 몇몇 실시예들에서는, 정상 데이터(44) 또는 정상 데이터(44)와 기 확보된 환자의 임상 데이터에 대해 클러스터링이 수행될 수 있다. 그리고, 원시 임상 데이터(47)를 구성하는 복수의 샘플(e.g. 48) 중에서 정상 데이터에 대한 클러스터 중심과의 거리가 기준치 이하인 샘플이 더 제거될 수 있다. 이러한 샘플 또한 정상 샘플(또는 이상치 샘플 등)일 확률이 높기 때문이다.

[0076] 또한, 몇몇 실시예들에서는, 트레이닝된 딥러닝 모델(40)을 이용하여 정상 데이터(44)에 대한 잠재 벡터(e.g. 43)가 추출되고, 잠재 벡터(e.g. 43)에 대해 클러스터링이 수행될 수 있다. 그리고, 원시 임상 데이터(47)를 구성하는 복수개의 샘플(e.g. 48) 중에 클러스터 중심(즉, 정상 잠재 벡터의 클러스터)과의 거리가 기준치 이하인 샘플이 더 제거될 수 있다. 이러한 샘플 또한 정상 샘플(또는 이상치 샘플 등)일 확률이 높기 때문이다.

[0077] 이하에서는, 도 9를 참조하여 본 개시의 다른 몇몇 실시예들에 따른 데이터 정제 방법에 대하여 설명하도록 한다.

[0078] 도 9에 도시된 바와 같이, 본 실시예 또한 오토 인코더 기반의 딥러닝 모델(50)을 이용하여 원시 임상 데이터를 정제하는 방법에 관한 것이다. 다만, 딥러닝 모델(50)의 구성과 트레이닝 방법이 앞선 실시예와 상이하다.

[0079] 구체적으로, 본 실시예에 따른 딥러닝 모델(50)은 인코더(51)와 디코더(52) 외에도 분류기(54)를 더 포함할 수 있다. 이때, 분류기(54)는 입력된 잠재 벡터(53)를 정상 클래스(normal)와 환자 클래스(abnormal)로 분류하는 역할을 담당할 수 있다.

[0080] 딥러닝 모델(50)은 정상 데이터와 환자 데이터를 포함하는 트레이닝 데이터(55)를 이용하여 트레이닝될 수 있다. 가령, 진단 시스템(10)은 트레이닝 데이터(55)의 복원 손실 값을 이용하여 인코더(51)와 디코더(52)를 트레이닝할 수 있다. 상술한 바와 같이, 복원 손실 값은 입력된 샘플(56)과 복원된 데이터(57) 간의 차이에 기초하여 산출될 수 있다. 또한, 진단 시스템(10)은 트레이닝 데이터(55)를 구성하는 각각의 샘플(56)에서 잠재 벡터(53)를 추출하고, 분류기(54)를 이용하여 추출된 잠재 벡터(53)에 대한 분류 손실(classification loss) 값을 산출할 수 있다. 분류 손실 값은 분류기(54)를 통해 예측된 클래스와 정답 클래스의 차이에 기초하여 산출될 수 있다. 또한, 진단 시스템(10)은 디코더(52)를 프리징한 상태에서 산출된 분류 손실 값을 이용하여 분류기(54)와 인코더(51)를 업데이트할 수 있다. 그렇게 함으로써, 인코더(51)가 클래스 분류에 적합한 잠재 벡터(53)를 추출하도록 학습될 수 있다.

[0081] 다음으로, 진단 시스템(10)은 트레이닝된 딥러닝 모델(50)을 이용하여 원시 임상 데이터(e.g. 47)를 구성하는 각각의 샘플(e.g. 48)로부터 잠재 벡터(e.g. 53)를 추출할 수 있다. 그리고, 진단 시스템(10)은 추출된 잠재 벡터(e.g. 53)를 분류기(54)에 입력하여 환자 클래스에 대한 컨피던스 스코어(confidence score)를 획득할 수 있고, 원시 임상 데이터(e.g. 47)에서 컨피던스 스코어가 기준치 이하인 샘플을 제거할 수 있다. 환자 클래스에 대한 컨피던스 스코어가 낮다는 것은 해당 샘플이 정상 샘플(또는 이상치 샘플 등)일 확률이 높다는 것을 의미하기 때문이다.

[0082] 지금까지 도 7 내지 도 9를 참조하여 본 개시의 다양한 실시예들에 따른 데이터 정제 방법에 대하여 설명하였다. 상술한 방법에 따르면, 오토인코더 기반의 딥러닝 모델을 이용하여 원시 임상 데이터에서 불필요한 샘플(e.g. 정상 샘플, 정상에 가까운 샘플, 이상치 샘플 등)이 정확하게 제거될 수 있다. 또한, 정제된 임상 데이터를 이용하여 변증 모델을 구축함으로써, 변증 모델의 정확도가 더욱 향상될 수 있다.

[0083] 이하에서는, 도 10 및 도 11을 참조하여 고성능의 변증 모델을 구축하는 방법에 관한 다양한 실시예들에 대하여 설명하도록 한다.

[0084] 도 10은 본 개시의 몇몇 실시예들에 따른 변증 모델 구축 방법을 나타내는 예시적인 흐름도이다. 단, 이는 본 개시의 목적을 달성하기 위한 바람직한 실시예일뿐이며, 필요에 따라 일부 단계가 추가되거나 삭제될 수 있음은 물론이다.

[0085] 도 10에 도시된 바와 같이, 본 실시예는 준비된 임상 데이터를 이용하여 복수개의 후보 변증 모델을 구축하는 단계 S510에서 시작될 수 있다. 다만, 후보 변증 모델을 구축하는 방식은 실시예에 따라 달라질 수 있다.

[0086] 몇몇 실시예들에서는, 오토인코더 기반 딥러닝 모델의 구성을 변경(e.g. 잠재벡터의 차원수가 변경되도록 구성

을 변경)함으로써 임상 데이터에서 서로 다른 차원수의 특징 데이터를 추출하고, 추출된 특징 데이터를 클러스터링하여 복수개의 후보 변증 모델이 구축될 수 있다. 가령, 진단 시스템(10)은 트레이닝된 딥러닝 모델을 이용하여 임상 데이터에서 제1 특징 데이터를 추출하고, 제1 특징 데이터를 클러스터링하여 제1 후보 변증 모델을 구축할 수 있다. 또한, 진단 시스템(10)은 딥러닝 모델의 구성을 변경하여 다시 트레이닝하고, 트레이닝된 딥러닝 모델을 이용하여 제1 특징 데이터와 차원수가 상이한 제2 특징 데이터를 추출할 수 있다. 그리고, 진단 시스템(10)은 제2 특징 데이터를 클러스터링하여 제2 후보 변증 모델을 구축할 수 있다.

[0087] 몇몇 실시예들에서는, 클러스터의 개수를 변경해가며 복수의 후보 변증 모델이 구축될 수 있다. 가령, 진단 시스템(10)은 오토인코더 기반 딥러닝 모델을 이용하여 추출된 특징 데이터를 클러스터링하여 제1 후보 변증 모델을 구축하고, 특징 데이터를 다시 클러스터링하여 제1 후보 변증 모델과 클러스터 개수가 상이한 제2 후보 변증 모델을 구축할 수 있다.

[0088] 단계 S520에서, 복수개의 후보 변증 모델에 대한 평가가 수행될 수 있다. 모델에 대한 평가는 다양한 방식으로 수행될 수 있다. 가령, 진단 시스템(10)은 칼린스키-하라바츠 지수(Calinski-Harabasz index), 실루엣 계수(Silhouette coefficient), WCSS(Within Cluster Sum of Squares) 등의 평가 지표를 이용하여 후보 변증 모델을 평가할 수 있다. 또는, 진단 시스템(10)은 각 클러스터에 속한 데이터 요소 값의 평균과 분산을 산출하고, 클러스터 별로 산출된 평균과 분산을 비교함으로써 후보 변증 모델을 평가할 수 있다. 이를테면, 진단 시스템(10)은 클러스터 내의 분산 값이 작을수록, 클러스터 간의 평균 값 차이가 클수록 높은 평가 점수를 부여할 수 있다. 보다 구체적인 예로서, 수면 장애에 관한 변증 모델을 구축하는 경우, 진단 시스템(10)은 클러스터 별로 PSQI, 베를린 설문 결과, GSRS, 영양지수 등에 관한 평균과 분산을 산출하고, 산출된 평균과 분산을 비교함으로써 후보 변증 모델을 평가할 수 있다.

[0089] 단계 S530에서, 평가 결과에 기초하여 복수개의 후보 변증 모델 중에서 서비스용 변증 모델(즉, 실제 진단 서비스에 이용될 변증 모델)이 선정될 수 있다. 가령, 진단 시스템(10)은 평가 결과(e.g. 클러스터링 평가 점수)가 가장 우수한 후보 변증 모델 또는 소정의 평가 기준에 부합하는 하나 이상의 후보 변증 모델을 서비스용 변증 모델로 선정할 수 있다.

[0090] 이하에서는, 도 11을 참조하여 본 개시의 다른 몇몇 실시예들에 따른 변증 모델 구축 방법에 대하여 설명하도록 한다. 다만, 본 개시의 명료함을 위해 앞선 실시예와 중복되는 내용에 대한 설명은 생략하도록 한다.

[0091] 도 11에 도시된 바와 같이, 본 실시예는 준비된 임상 데이터를 트레이닝용과 검증용으로 구분하는 단계 S610에서 시작될 수 있다. 가령, 진단 시스템(10)은 미리 설정된 비율(e.g. 트레이닝용:검증용 = 8:2)에 따라 준비된 임상 데이터를 구분할 수 있다.

[0092] 단계 S620에서, 트레이닝용 데이터를 이용하여 복수개의 후보 변증 모델이 구축될 수 있다. 가령, 진단 시스템(10)은 오토인코더 기반의 딥러닝 모델을 이용하여 트레이닝 데이터에서 특징 데이터를 추출하고, 추출된 특징 데이터를 클러스터링하여 후보 변증 모델을 구축할 수 있다. 본 단계에 대한 자세한 설명은 상술한 단계 S510을 더 참조하도록 한다.

[0093] 단계 S630에서, 검증용 데이터를 이용하여 복수개의 후보 변증 모델에 대한 평가가 수행될 수 있다. 가령, 진단 시스템(10)은 후보 변증 모델의 클러스터 별로 검증용 데이터에 대한 평균과 분산을 산출하고 비교함으로써 평가를 수행할 수 있다. 검증용 데이터는 후보 변증 모델 구축에 이용되지 않은 데이터이기 때문에, 이러한 데이터를 이용하면 후보 변증 모델이 더욱 정확하게 평가될 수 있다. 평가 방식에 대한 자세한 설명은 상술한 단계 S520을 더 참조하도록 한다.

[0094] 단계 S640에서, 평가 결과에 기초하여 복수개의 후보 변증 모델 중에서 서비스용 변증 모델이 선정될 수 있다. 본 단계에 대한 자세한 설명은 상술한 단계 S530을 참조하도록 한다.

[0095] 지금까지 도 10 및 도 11을 참조하여 본 개시의 다양한 실시예들에 따른 변증 모델 구축 방법에 대하여 설명하였다. 상술한 방법에 따르면, 특징 데이터의 차원수, 클러스터의 개수 등을 변경해가며 다수의 후보 변증 모델을 구축하고, 후보 변증 모델을 다양한 방식으로 평가함으로써 고성능의 변증 모델이 용이하게 구축될 수 있다.

[0096] 이하에서는, 본 개시의 몇몇 실시예들에 따른 진단 시스템(10)을 구현할 수 있는 예시적인 컴퓨팅 장치(100)에 대하여 설명하도록 한다.

[0097] 도 12는 컴퓨팅 장치(100)를 나타내는 예시적인 하드웨어 구성도이다.

[0098] 도 12에 도시된 바와 같이, 컴퓨팅 장치(100)는 하나 이상의 프로세서(110), 버스(130), 통신 인터페이스(140),

프로세서(110)에 의하여 수행되는 컴퓨터 프로그램을 로드(load)하는 메모리(120)와, 컴퓨터 프로그램(160)를 저장하는 스토리지(150)를 포함할 수 있다. 다만, 도 12에는 본 개시의 실시예와 관련 있는 구성요소들만이 도시되어 있다. 따라서, 본 개시가 속한 기술분야의 통상의 기술자라면 도 12에 도시된 구성요소들 외에 다른 범용적인 구성 요소들이 더 포함될 수 있음을 알 수 있다. 즉, 컴퓨팅 장치(100)에는, 도 12에 도시된 구성 요소 이외에도 다양한 구성 요소가 더 포함될 수 있다. 또한, 경우에 따라, 도 12에 도시된 구성요소들 중 일부가 생략된 형태로 컴퓨팅 장치(100)가 구성될 수도 있다. 이하, 컴퓨팅 장치(100)의 각 구성요소에 대하여 설명한다.

[0099] 프로세서(110)는 컴퓨팅 장치(100)의 각 구성의 전반적인 동작을 제어할 수 있다. 프로세서(110)는 CPU(Central Processing Unit), MPU(Micro Processor Unit), MCU(Micro Controller Unit), GPU(Graphic Processing Unit) 또는 본 개시의 기술 분야에 잘 알려진 임의의 형태의 프로세서 중 적어도 하나를 포함하여 구성될 수 있다. 또한, 프로세서(110)는 본 개시의 실시예들에 따른 동작/방법을 실행하기 위한 적어도 하나의 애플리케이션 또는 프로그램에 대한 연산을 수행할 수 있다. 컴퓨팅 장치(100)는 하나 이상의 프로세서를 구비할 수 있다.

[0100] 다음으로, 메모리(120)는 각종 데이터, 명령 및/또는 정보를 저장할 수 있다. 메모리(120)는 본 개시의 실시예들에 따른 동작/방법을 실행하기 위하여 스토리지(150)로부터 컴퓨터 프로그램(160)을 로드할 수 있다. 메모리(120)는 RAM과 같은 휘발성 메모리로 구현될 수 있을 것이나, 본 개시의 기술적 범위가 이에 한정되는 것은 아니다.

[0101] 다음으로, 버스(130)는 컴퓨팅 장치(100)의 구성 요소 간 통신 기능을 제공할 수 있다. 버스(130)는 주소 버스(Address Bus), 데이터 버스(Data Bus) 및 제어 버스(Control Bus) 등 다양한 형태의 버스로 구현될 수 있다.

[0102] 다음으로, 통신 인터페이스(140)는 컴퓨팅 장치(100)의 유무선 인터넷 통신을 지원할 수 있다. 또한, 통신 인터페이스(140)는 인터넷 통신 외의 다양한 통신 방식을 지원할 수도 있다. 이를 위해, 통신 인터페이스(140)는 본 개시의 기술 분야에 잘 알려진 통신 모듈을 포함하여 구성될 수 있다.

[0103] 다음으로, 스토리지(150)는 상기 하나 이상의 컴퓨터 프로그램(160)을 비임시적으로 저장할 수 있다. 스토리지(150)는 ROM(Read Only Memory), EPROM(Erasable Programmable ROM), EEPROM(Electrically Erasable Programmable ROM), 플래시 메모리 등과 같은 비휘발성 메모리, 하드 디스크, 착탈형 디스크, 또는 본 개시가 속하는 기술 분야에서 잘 알려진 임의의 형태의 컴퓨터로 읽을 수 있는 기록 매체를 포함하여 구성될 수 있다.

[0104] 다음으로, 컴퓨터 프로그램(160)은 메모리(120)에 로드될 때 프로세서(110)로 하여금 본 개시의 다양한 실시예들에 따른 동작/방법을 수행하도록 하는 하나 이상의 인스트럭션을 포함할 수 있다. 즉, 프로세서(110)는 상기 하나 이상의 인스트럭션을 실행함으로써, 본 개시의 다양한 실시예에 따른 동작/방법을 수행할 수 있다.

[0105] 예를 들어, 컴퓨터 프로그램(160)은 변증 모델을 위한 임상 데이터를 준비하는 동작, 오토인코더 기반 딥러닝 모델을 이용하여 준비된 임상 데이터에서 특징 데이터를 추출하고, 추출된 특징 데이터를 클러스터링하여 변증 모델을 구축하는 동작 및 구축된 변증 모델을 이용하여 진단 서비스를 제공하는 동작을 수행하는 하나 이상의 인스트럭션을 포함할 수 있다. 이와 같은 경우, 컴퓨팅 장치(100)를 통해 본 개시의 몇몇 실시예들에 따른 진단 시스템(10)이 구현될 수 있다.

[0106] 지금까지 도 12를 참조하여 본 개시의 몇몇 실시예들에 따른 진단 시스템(10)을 구현할 수 있는 예시적인 컴퓨팅 장치(100)에 대하여 설명하였다. 이하에서는, 본 개시의 기술적 사상을 검증하기 위해 수행된 실험예에 대하여 간략하게 소개하도록 한다.

[0107] 먼저, 본 개시의 발명자들은 설문을 통해 육군 병사들로부터 수면 장애에 관한 임상 데이터를 수집하였다. 수집된 임상 데이터는 약 2579개의 샘플로 구성되었고, 하기의 표 1에 기재된 정보로 구성되었다.

표 1

[0108]

구분	세부 정보(데이터 요소)
인구통계학적 정보	인종, 생년월일, 키, 몸무게, 학력, 신체 등급, 흡연력, 음주력, 운동력, 입대일, 군번, 계급, 소속부대
근무 정보	야간 근무 형태, 야간 근무 횟수, 야간 근무 시간, 야간 근무가 수면 및/또는 피로도에 미치는 영향, 병과, 보직
병력 정보	수면 장애, 위장관 장애, 고혈압, 당뇨, 고지혈증, 심장질환에 관한 병력, 스트레스 상태, 약물력
수면 정보	PSQI, 베를린 설문, 수면 환경
식이영양 정보	식습관, 영양지수
위장관 정보	GSRS, BSS

[0109] 다음으로, 본 발명자들은 임상 데이터의 각 샘플을 160차원의 수치화된 벡터로 변환하고, 도 2에 예시된 방법에 따라 오토인코더 기반의 딥러닝 모델을 이용하여 변증 모델을 구축하였다. 구체적으로는, 하기의 표 2에 기재된 바와 같이, 도 4에 예시된 구조의 딥러닝 모델에서 추출되는 잠재 벡터의 차원수(d)와 클러스터 개수를 변경해 가며 구현예1 내지 구현예10에 따른 변증 모델을 구축하였다. 또한, 성능 비교를 위해 변환된 임상 데이터를 그대로 클러스터링하여 비교예1에 따른 변증 모델을 구축하고, PCA를 통해 추출된 특징 벡터를 클러스터링하여 비교예2에 따른 변증 모델을 구축하였다. 변증 모델 구축은 K-평균 클러스터링 기법을 통해 수행되었다.

[0110] 다음으로, 본 발명자들은 실루엣 계수를 평가 지표로 이용하여 구현예들과 비교예들에 따른 변증 모델을 평가하였고, 평가 결과는 하기의 표 2에 기재되어 있다.

표 2

[0111]

구분	클러스터 개수						
	2	3	4	5	6	7	8
비교예1 (raw data)	0.090	0.028	0.010	0.005	0.006	0.002	0.005
비교예2(PCA)	0.277	0.202	0.187	0.186	0.189	0.187	0.185
구현예1(AE:d=1)	0.628	0.574	0.568	0.586	0.579	0.564	0.568
구현예2(AE:d=2)	0.704	0.614	0.586	0.560	0.524	0.523	0.504
구현예3(AE:d=3)	0.620	0.527	0.497	0.473	0.455	0.423	0.391
구현예4(AE:d=4)	0.550	0.463	0.390	0.371	0.326	0.322	0.294
구현예5(AE:d=5)	0.538	0.477	0.405	0.344	0.316	0.302	0.295
구현예6(AE:d=6)	0.524	0.412	0.348	0.312	0.259	0.234	0.236
구현예7(AE:d=7)	0.532	0.436	0.366	0.296	0.258	0.255	0.239
구현예8(AE:d=8)	0.487	0.371	0.307	0.277	0.215	0.211	0.197
구현예9(AE:d=9)	0.503	0.370	0.339	0.267	0.243	0.194	0.226
구현예10(AE:d=10)	0.475	0.347	0.291	0.249	0.200	0.191	0.177

[0112] 상기 표 2에 기재된 바와 같이, 구현예들에 따른 실루엣 계수의 값이 비교예들을 크게 상회하는 것을 확인할 수 있다. 이는 오토인코더 기반의 딥러닝 모델을 이용한 특징 추출 방식이 비교예들에 비해 훨씬 우수하고 클러스터링 기법에 적합한 방식이라는 것을 의미하는 것으로 판단된다.

[0113] 또한, 본 발명자들은 평가 결과가 우수한 구현예2에 따른 딥러닝 모델을 이용하여 엘보우(elbow) 기법에 따라 최적 클러스터 개수를 구하는 실험을 더 진행하였다. 실험 결과는 도 13에 도시되어 있다.

[0114] 도 13에 도시된 바와 같이, 클러스터의 개수가 3에서 4로 증가할 때, WCSS의 감소 비율이 급격하게 줄어드는 것을 확인할 수 있다. 이러한 실험 결과에 따라, 본 발명자들은 최적 클러스터의 개수를 "3개"로 결정하고, 3개의 클러스터에 속한 임상 데이터를 분석하여 각 클러스터가 수면 장애에 관한 변증 유형과 실제로 매칭되는지를 확인하였다.

[0115] 확인 결과, 하기의 표 3에 기재된 바와 같이, 각각의 클러스터가 서로 다른 변증 유형에 매칭되는 것을 확인하였는데, 이는 상술한 변증 진단 방법에 따라 구축된 변증 모델이 환자의 변증 유형을 정확하게 판별할 수 있다는 것을 의미하고, 아래와 같이 각 변증 유형에 따른 경혈 정보를 이용하여 고품질의 진단 서비스를 제공할 수 있다는 것을 의미한다.

표 3

[0116]

구분	클러스터#1	클러스터#2	클러스터#3
변증 유형	위중불화(胃中不和)	담연울결 (痰涎鬱結)	심비양허 (心脾兩虛)
경혈	ST36, CV12, ST25	ST40, CV12	CV12, ST36, ST40

[0117] 지금까지 본 개시의 기술적 사상을 검증하기 위해 수행된 실험예에 대하여 간략하게 소개하였다.

[0118] 지금까지 도 1 내지 도 13을 참조하여 설명된 본 개시의 기술적 사상은 컴퓨터가 읽을 수 있는 매체 상에 컴퓨

터가 읽을 수 있는 코드로 구현될 수 있다. 상기 컴퓨터로 읽을 수 있는 기록 매체는, 예를 들어 이동형 기록 매체(CD, DVD, 블루레이 디스크, USB 저장 장치, 이동식 하드 디스크)이거나, 고정식 기록 매체(ROM, RAM, 컴퓨터 구비 형 하드 디스크)일 수 있다. 상기 컴퓨터로 읽을 수 있는 기록 매체에 기록된 상기 컴퓨터 프로그램은 인터넷 등의 네트워크를 통하여 다른 컴퓨팅 장치에 전송되어 상기 다른 컴퓨팅 장치에 설치될 수 있고, 이로써 상기 다른 컴퓨팅 장치에서 사용될 수 있다.

[0119] 이상에서, 본 개시의 실시예를 구성하는 모든 구성 요소들이 하나로 결합되거나 결합되어 동작하는 것으로 설명되었다고 해서, 본 개시의 기술적 사상이 반드시 이러한 실시예에 한정되는 것은 아니다. 즉, 본 개시의 목적 범위 안에서라면, 그 모든 구성요소들이 하나 이상으로 선택적으로 결합하여 동작할 수도 있다.

[0120] 도면에서 동작들이 특정한 순서로 도시되어 있지만, 반드시 동작들이 도시된 특정한 순서로 또는 순차적 순서로 실행되어야만 하거나 또는 모든 도시 된 동작들이 실행되어야만 원하는 결과를 얻을 수 있는 것으로 이해되어서는 안 된다. 특정 상황에서는, 멀티태스킹 및 병렬 처리가 유리할 수도 있다. 더욱이, 위에 설명한 실시예들에서 다양한 구성들의 분리는 그러한 분리가 반드시 필요한 것으로 이해되어서는 안 되고, 설명된 프로그램 컴포넌트들 및 시스템들은 일반적으로 단일 소프트웨어 제품으로 함께 통합되거나 다수의 소프트웨어 제품으로 패키지 될 수 있음을 이해하여야 한다.

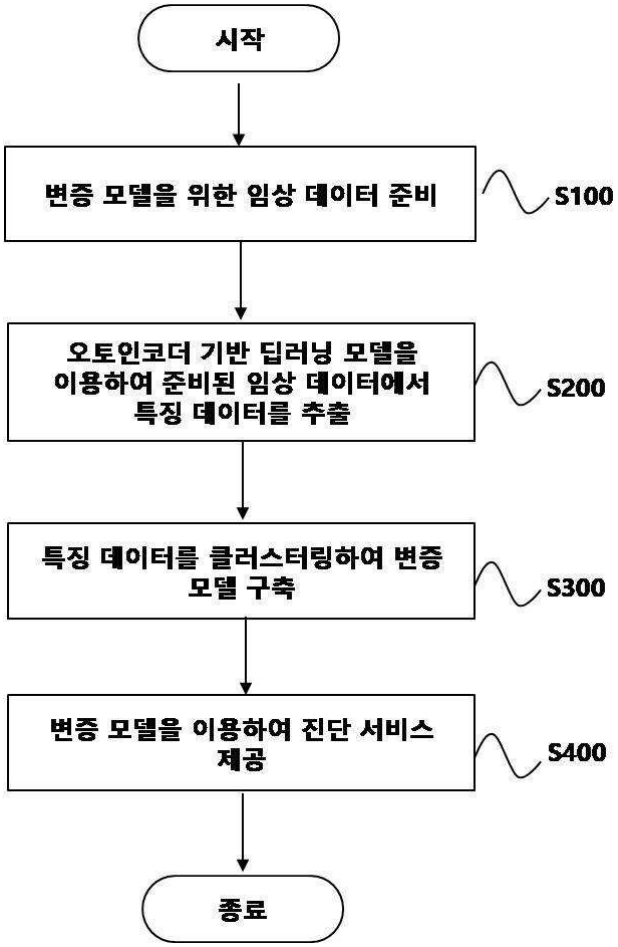
[0121] 이상 첨부된 도면을 참조하여 본 개시의 실시예들을 설명하였지만, 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자는 그 기술적 사상이나 필수적인 특징을 변경하지 않고서 본 개시가 다른 구체적인 형태로도 실시될 수 있다는 것을 이해할 수 있다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적인 것이 아닌 것으로 이해해야만 한다. 본 개시의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 개시에 의해 정의되는 기술적 사상의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

도면

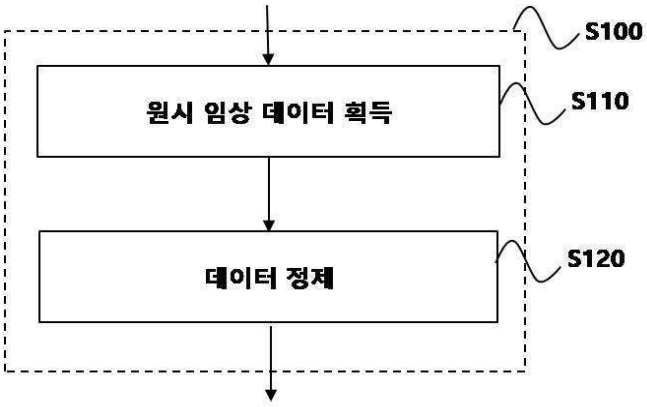
도면1



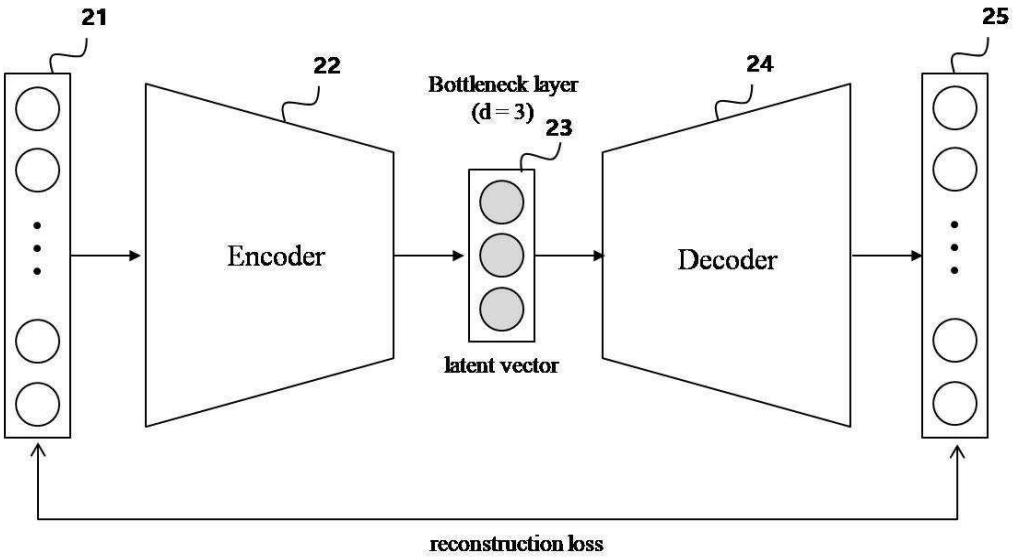
도면2


시작
변증 모델을 위한 임상 데이터 준비
S100
오토인코더 기반 딥러닝 모델을 이용하여 준비된 임상 데이터에서 특징 데이터를 추출
S200
특징 데이터를 클러스터링하여 변증 모델 구축
S300
변증 모델을 이용하여 진단 서비스 제공
S400
종료

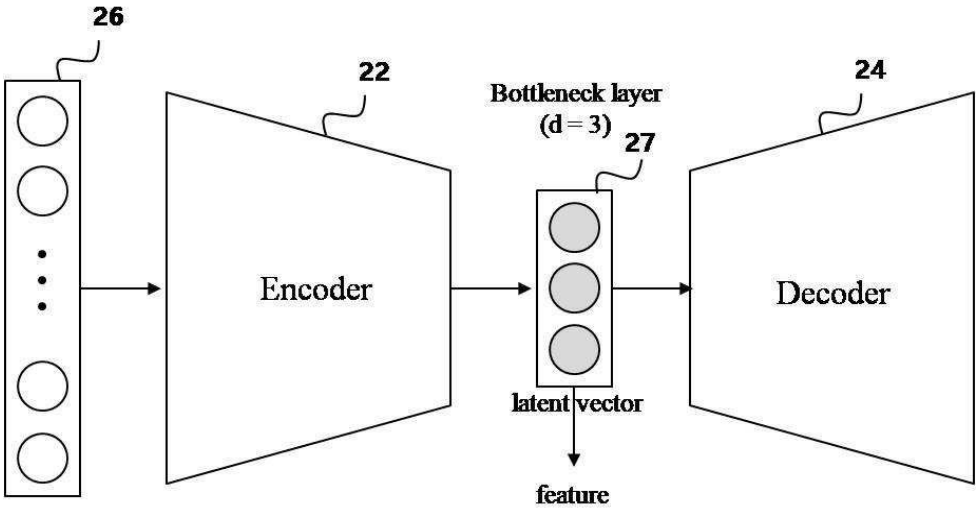
도면3


S100
원시 임상 데이터 획득
S110
데이터 정제
S120

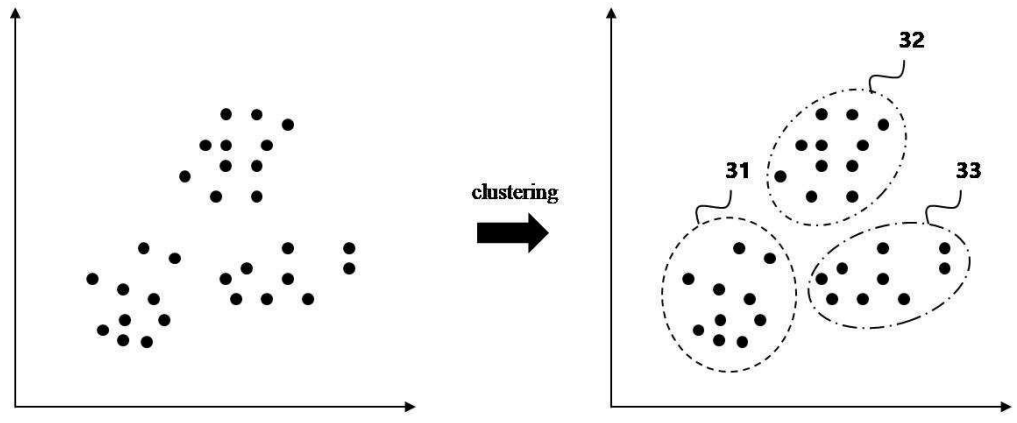
도면4


21
22
Bottleneck layer
(d = 3)
23
24
25
Encoder
Decoder
latent vector
reconstruction loss

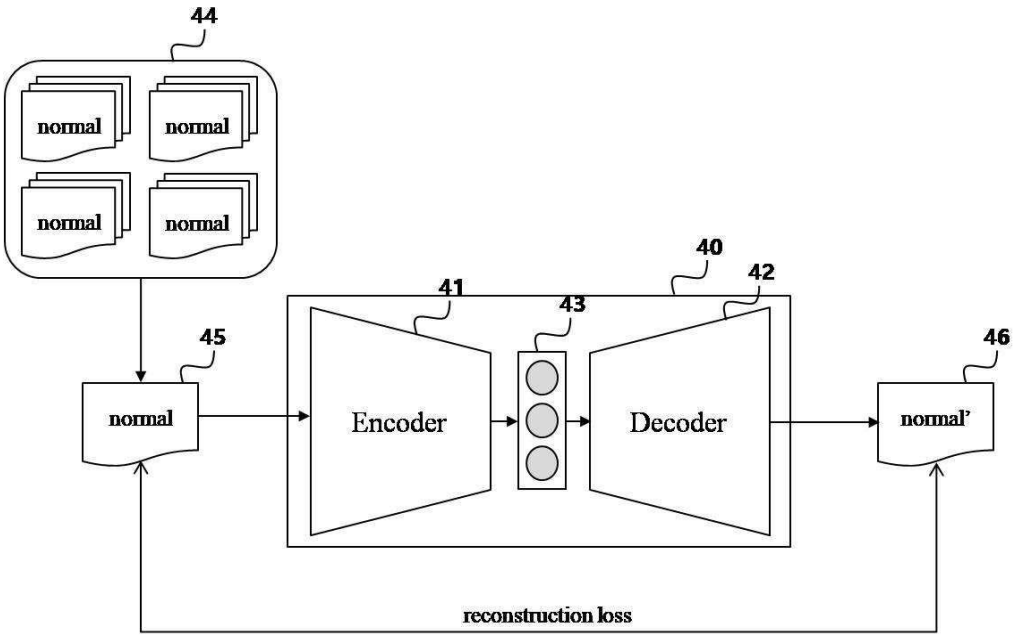
도면5


26
22
Bottleneck layer
(d = 3)
27
24
Encoder
Decoder
latent vector
feature

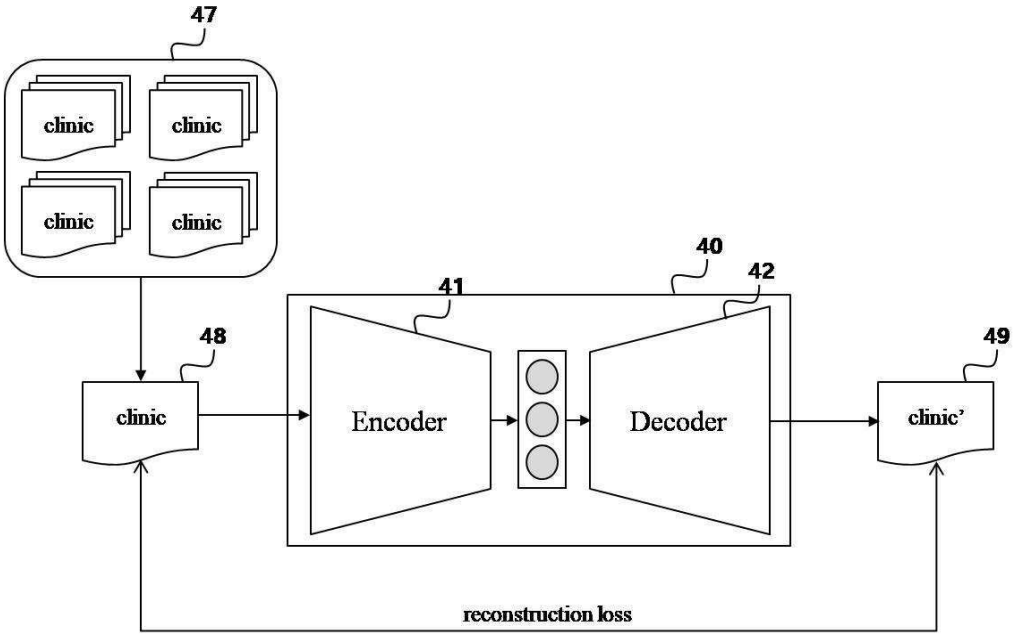
도면6


clustering
32
31
33

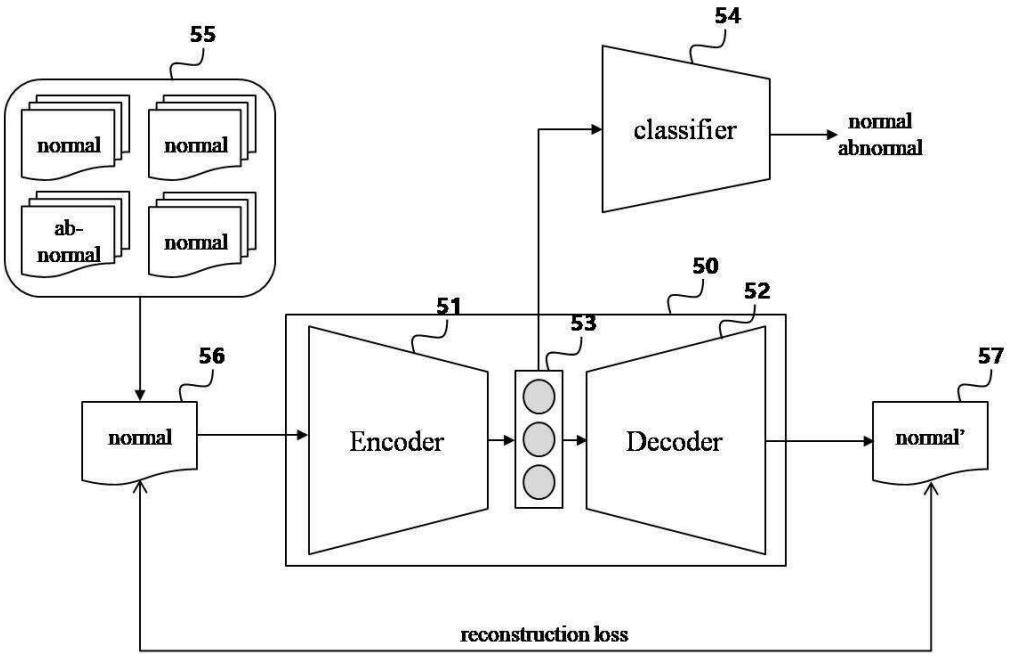
도면7


44
normal
normal
normal
normal
40
42
41
43
45
46
normal
Encoder
Decoder
normal'
reconstruction loss

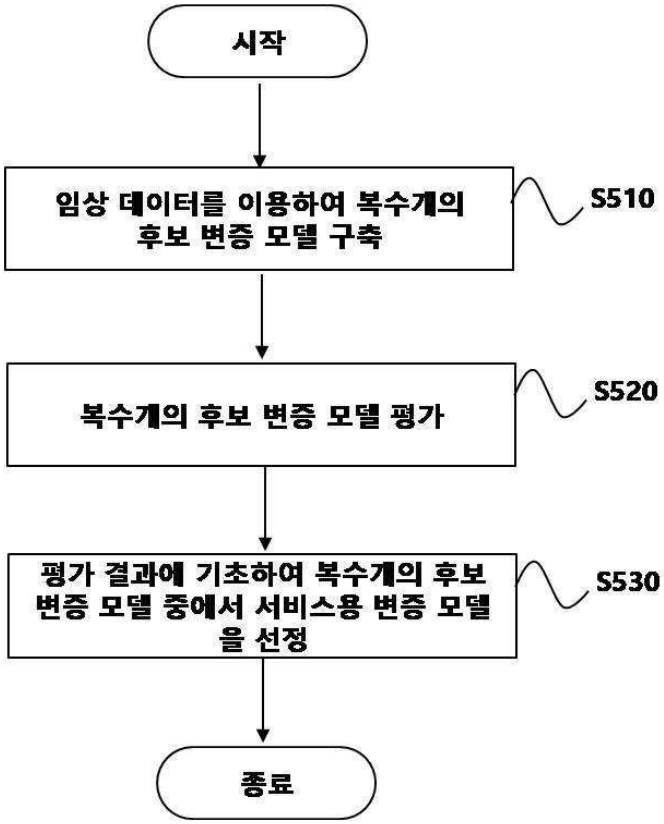
도면8


47
clinic
clinic
clinic
clinic
40
42
41
48
49
clinic
Encoder
Decoder
clinic'
reconstruction loss

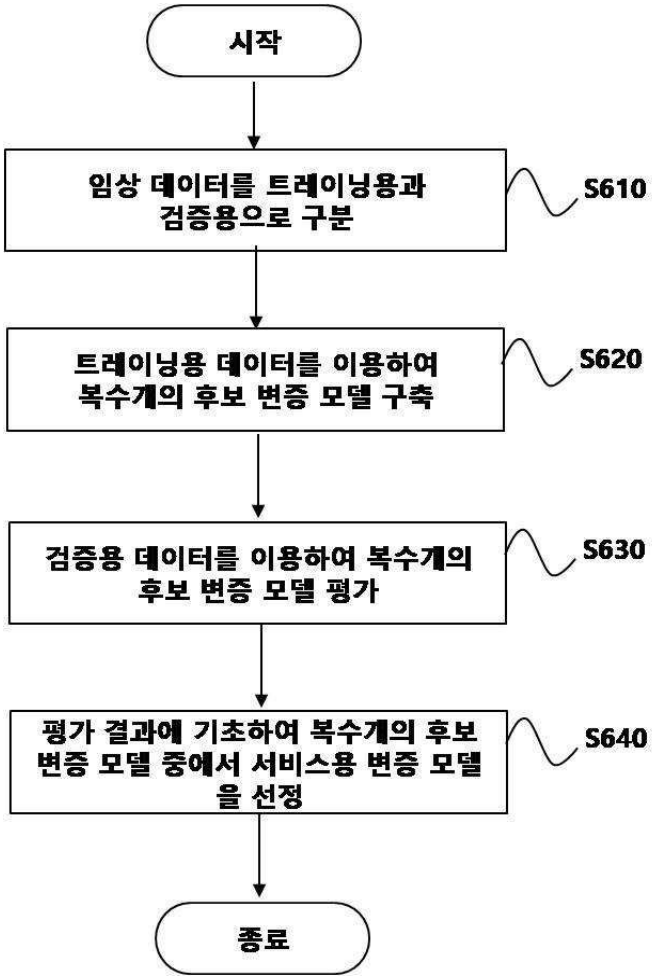
도면9


55
54
normal
normal
ab-
normal
normal
classifier
normal
abnormal
50
52
51
53
56
57
normal
Encoder
Decoder
normal'
reconstruction loss

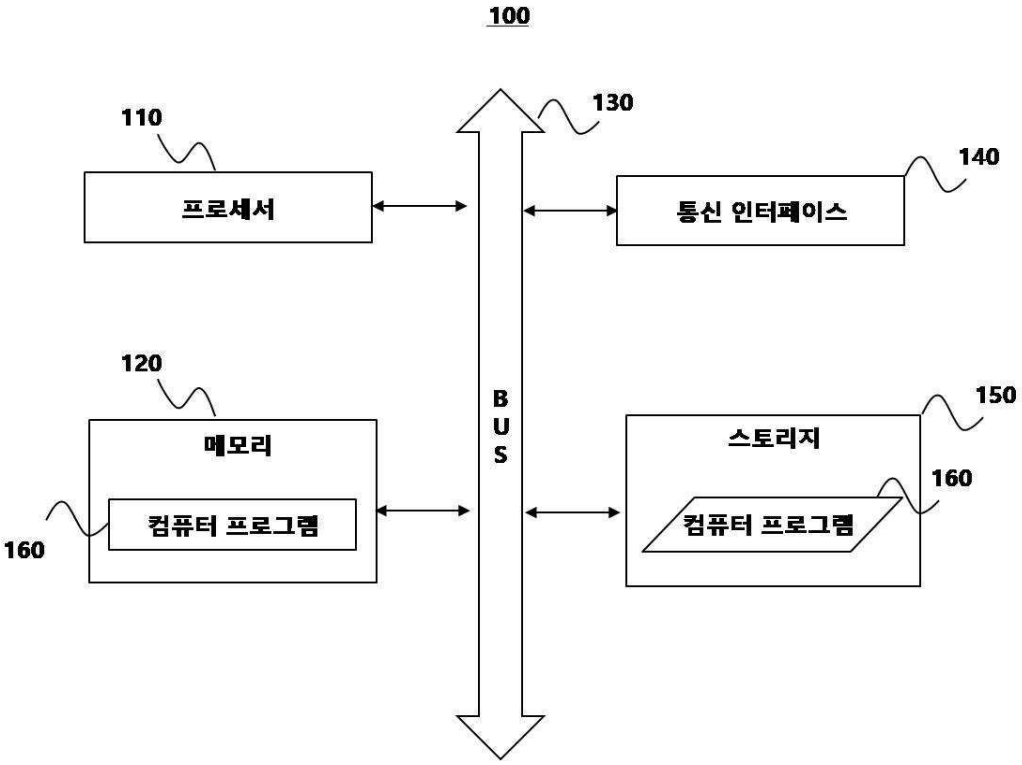
도면10


시작
임상 데이터를 이용하여 복수개의 후보 변증 모델 구축
S510
복수개의 후보 변증 모델 평가
S520
평가 결과에 기초하여 복수개의 후보 변증 모델 중에서 서비스용 변증 모델을 선정
S530
종료

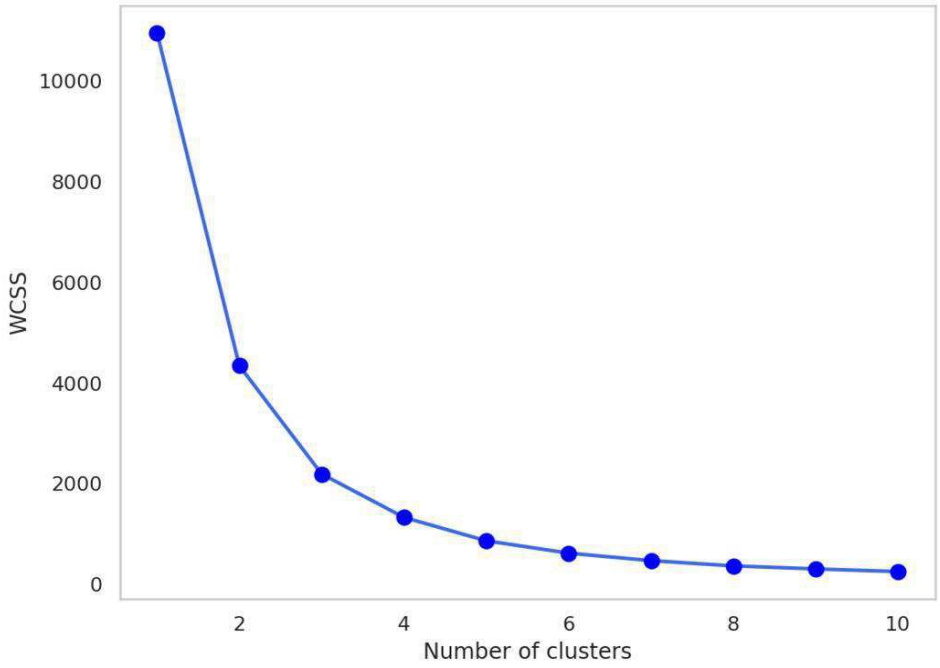
도면11


시작
임상 데이터를 트레이닝용과 검증용으로 구분
S610
트레이닝용 데이터를 이용하여 복수개의 후보 변증 모델 구축
S620
검증용 데이터를 이용하여 복수개의 후보 변증 모델 평가
S630
평가 결과에 기초하여 복수개의 후보 변증 모델 중에서 서비스용 변증 모델을 선정
S640
종료

도면12


100
110
프로세서
130
140
통신 인터페이스
120
메모리
160
컴퓨터 프로그램
BUS
150
스토리지
160
컴퓨터 프로그램

도면13


10000
8000
6000
4000
2000
0
WCSS
2
4
6
8
10
Number of clusters