



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0008480
(43) 공개일자 2021년01월22일

(51) 국제특허분류(Int. Cl.)
G06N 3/04 (2006.01) G06N 20/10 (2019.01)
G06N 3/08 (2006.01) G06N 5/00 (2019.01)

(52) CPC특허분류
G06N 3/0454 (2013.01)
G06N 20/10 (2019.01)

(21) 출원번호 10-2020-7032320
(22) 출원일자(국제) 2019년04월27일
심사청구일자 없음
(85) 번역문제출일자 2020년11월09일
(86) 국제출원번호 PCT/US2019/029532
(87) 국제공개번호 WO 2019/217113
국제공개일자 2019년11월14일
(30) 우선권주장
15/976,514 2018년05월10일 미국(US)

(71) 출원인
마이크로소프트 테크놀로지 라이선싱, 엘엘씨
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이

(72) 발명자
아미자데 사이드
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 마이크로소프트 테크놀로지
라이선싱, 엘엘씨
양 지
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 마이크로소프트 테크놀로지
라이선싱, 엘엘씨
(뒷면에 계속)

(74) 대리인
제일특허법인(유)

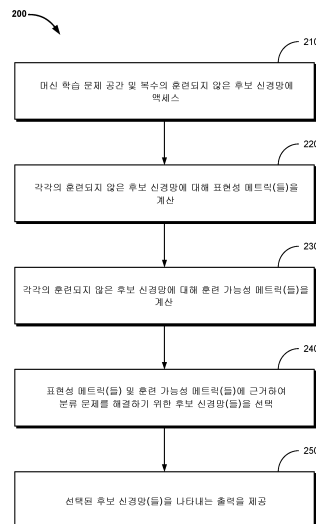
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 지도 머신 학습 문제를 위한 신경망 아키텍처 선택

(57) 요약

머신 학습 문제를 위한 신경망을 선택하는 시스템 및 방법이 개시된다. 방법은 입력 매트릭스에 액세스하는 것을 포함한다. 방법은 머신 학습 문제와 관련된 머신 학습 문제 공간 및 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것을 포함한다. 방법은 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 표현성(expressivity)을 캡처하는 적어도 하나의 표현성 측도를 계산하는 것을 포함한다. 방법은 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 훈련 가능성(trainability)을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산하는 것을 포함한다. 방법은 적어도 하나의 표현성 측도 및 적어도 하나의 훈련 가능성 측도에 근거하여 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것을 포함한다.

대표도 - 도2



(52) CPC특허분류

G06N 3/0445 (2013.01)

G06N 3/0472 (2013.01)

G06N 3/08 (2013.01)

G06N 5/003 (2019.01)

(72) 발명자

후시 니콜로

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 마이크로소프트 테크놀로지 라이선싱, 엘엘씨

카살레 프란체스코 파올로

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 마이크로소프트 테크놀로지 라이선싱, 엘엘씨

명세서

청구범위

청구항 1

프로세싱 하드웨어 및 상기 프로세싱 하드웨어로 하여금 동작을 수행하도록 하는 명령어를 저장하는 메모리를 포함하는 시스템으로서,

상기 동작은,

머신 학습 문제와 관련된 머신 학습 문제 공간 및 상기 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것과,

각각의 훈련되지 않은 후보 신경망에 대해 상기 머신 학습 문제에 대한 상기 후보 신경망의 표현성(expressivity)을 캡처하는 적어도 하나의 표현성 측도(measure)를 계산하는 것과,

각각의 훈련되지 않은 후보 신경망에 대해 상기 머신 학습 문제에 대한 상기 후보 신경망의 훈련 가능성(trainability)을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산하는 것과,

상기 적어도 하나의 표현성 측도 및 상기 적어도 하나의 훈련 가능성 측도에 근거하여 상기 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것과,

상기 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공하는 것을 포함하는,

시스템.

청구항 2

제 1 항에 있어서,

상기 적어도 하나의 표현성 측도는 상기 훈련되지 않은 후보 신경망에 의한 상기 머신 학습 문제 공간으로부터 샘플의 분리의 측도를 나타내는 것인, 시스템.

청구항 3

제 2 항에 있어서,

상기 분리의 측도는 크기인, 시스템.

청구항 4

제 2 항에 있어서,

상기 분리의 측도는 각도인, 시스템.

청구항 5

제 1 항에 있어서,

상기 적어도 하나의 훈련 가능성 측도는 훈련의 제 1 단계 동안 상기 후보 신경망에서 가중치의 확률적 경사 하강을 나타내는, 시스템.

청구항 6

제 1 항에 있어서,

상기 머신 학습 문제를 해결하기 위한 상기 적어도 하나의 후보 신경망을 선택하는 것은, 임계값을 초과하는 상기 적어도 하나의 표현성 측도 및 범위 내의 상기 적어도 하나의 훈련 가능성을 갖는 상기 적어도 하나의 후보 신경망을 선택하는 것을 포함하고, 상기 범위는 최소 범위 및 최대 범위에 의해 정의되는, 시스템.

청구항 7

제 1 항에 있어서,

상기 동작은 상기 머신 학습 문제를 해결하기 위해 상기 적어도 하나의 후보 신경망을 훈련하는 것을 더 포함하는, 시스템.

청구항 8

제 7 항에 있어서,

상기 동작은, 상기 머신 학습 문제를 해결하기 위해 상기 머신 학습 문제 공간에서 상기 훈련된 적어도 하나의 후보 신경망을 실행하는 것과, 상기 훈련된 적어도 하나의 후보 신경망에 의해 생성된 상기 머신 학습 문제에 대한 해결책을 제공하는 것을 더 포함하는, 시스템.

청구항 9

하나 이상의 머신으로 하여금 동작을 수행하도록 하는 명령어를 저장하는 머신 판독 가능 매체로서,

상기 동작은,

머신 학습 문제와 관련된 머신 학습 문제 공간 및 상기 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것과,

각각의 훈련되지 않은 후보 신경망에 대해 상기 머신 학습 문제에 대한 상기 후보 신경망의 표현성을 캡처하는 적어도 하나의 표현성 측도를 계산하는 것과,

각각의 훈련되지 않은 후보 신경망에 대해 상기 머신 학습 문제에 대한 상기 후보 신경망의 훈련 가능성을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산하는 것과,

상기 적어도 하나의 표현성 측도 및 상기 적어도 하나의 훈련 가능성 측도에 근거하여 상기 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것과,

상기 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공하는 것을 포함하는,

머신 판독 가능 매체.

청구항 10

제 9 항에 있어서,

상기 적어도 하나의 표현성 측도는 상기 훈련되지 않은 후보 신경망에 의한 상기 머신 학습 문제 공간으로부터 샘플의 분리의 측도를 나타내는, 머신 판독 가능 매체.

청구항 11

제 10 항에 있어서,
상기 분리의 측도는 크기인, 머신 판독 가능 매체.

청구항 12

제 10 항에 있어서,
상기 분리의 측도는 각도인, 머신 판독 가능 매체.

청구항 13

제 9 항에 있어서,
상기 적어도 하나의 훈련 가능성 측도는 훈련의 제 1 단계 동안 상기 후보 신경망에서 가중치의 확률적 경사 하강을 나타내는, 머신 판독 가능 매체.

청구항 14

제 9 항에 있어서,
상기 머신 학습 문제를 해결하기 위한 상기 적어도 하나의 후보 신경망을 선택하는 것은, 임계값을 초과하는 상기 적어도 하나의 표현성 측도 및 범위 내의 상기 적어도 하나의 훈련 가능성 측도를 갖는 상기 적어도 하나의 후보 신경망을 선택하는 것을 포함하고, 상기 범위는 최소 범위 및 최대 범위에 의해 정의되는, 머신 판독 가능 매체.

청구항 15

머신 학습 문제와 관련된 머신 학습 문제 공간 및 상기 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것과,

각각의 훈련되지 않은 후보 신경망에 대해 상기 머신 학습 문제에 대한 상기 후보 신경망의 표현성을 캡처하는 적어도 하나의 표현성 측도를 계산하는 것과,

각각의 훈련되지 않은 후보 신경망에 대해 상기 머신 학습 문제에 대한 상기 후보 신경망의 훈련 가능성을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산하는 것과,

상기 적어도 하나의 표현성 측도 및 상기 적어도 하나의 훈련 가능성 측도에 근거하여 상기 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것과,

상기 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공하는 것을 포함하는,
방법.

발명의 설명

기술 분야

배경 기술

[0001] 복수의 다른 타입의 신경망 아키텍처(예를 들어, 컨볼루션 신경망, 피드포워드 신경망 등)가 알려져 있다. 주어진 머신 학습 문제를 해결하기 위한 신경망 아키텍처 (및 주어진 아키텍처 타입 내의 서브-아키텍처)를 선택하는 것은 어려울 수 있다.

발명의 내용

해결하려는 과제

과제의 해결 수단

- [0002] 본 개시 내용은 일반적으로 머신 학습 문제를 해결하기 위한 신경망 아키텍처를 선택하도록 구성된 머신(이러한 특수 목적 머신의 컴퓨터화된 이형 및 이러한 이형에 대한, 그리고 이러한 특수 목적 머신이 신경망을 위한 기술을 제공하는 다른 특수 목적 머신에 비해 개선되게 되는 기술에 대한 개선을 포함함)에 관련된다. 특히, 본 개시는 주어진 머신 학습 문제 공간에서 머신 학습 문제를 해결하기 위한 신경망 아키텍처를 선택하는 시스템 및 방법을 다룬다.
- [0003] 본 명세서에 설명된 기술의 일부 측면에 따르면, 시스템은 프로세싱 하드웨어 및 메모리를 포함한다. 메모리는 프로세싱 하드웨어에 의해 실행될 때 프로세싱 하드웨어로 하여금 동작을 수행하도록 하는 명령어를 저장한다. 동작은 머신 학습 문제와 관련된 머신 학습 문제 공간 및 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것을 포함한다. 동작은 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 표현성(expressivity)을 캡처하는 적어도 하나의 표현성 측도(measure)를 계산하는 것을 포함한다. 동작은 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 훈련 가능성(trainability)과 관련된 적어도 하나의 훈련 가능성 측도를 계산하는 것을 포함한다. 동작은 적어도 하나의 표현성 측도 및 적어도 하나의 훈련 가능성 측도에 근거하여 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것을 포함한다. 동작은 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공하는 것을 포함한다.
- [0004] 본 명세서에 설명된 기술의 일부 측면에 따르면, 머신 관독 가능 매체는 하나 이상의 머신에 의해 실행될 때 하나 이상의 머신으로 하여금 동작을 수행하도록 하는 명령어를 저장한다. 동작은 머신 학습 문제와 관련된 머신 학습 문제 공간 및 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것을 포함한다. 동작은 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 표현성을 캡처하는 적어도 하나의 표현성 측도를 계산하는 것을 포함한다. 동작은 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 훈련 가능성을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산하는 것을 포함한다. 동작은 적어도 하나의 표현성 측도 및 적어도 하나의 훈련 가능성 측도에 근거하여 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것을 포함한다. 동작은 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공하는 것을 포함한다.
- [0005] 본 명세서에 설명된 기술의 일부 측면에 따르면, 방법은 입력 매트릭스에 액세스하는 것을 포함한다. 방법은 머신 학습 문제와 관련된 머신 학습 문제 공간 및 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것을 포함한다. 방법은 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 표현성을 캡처하는 적어도 하나의 표현성 측도를 계산하는 것을 포함한다. 방법은 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 훈련 가능성을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산하는 것을 포함한다. 방법은 적어도 하나의 표현성 측도 및 적어도 하나의 훈련 가능성 측도에 근거하여 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것을 포함한다. 방법은 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공하는 것을 포함한다.

도면의 간단한 설명

- [0006] 기술의 일부 실시예는 첨부된 도면에서 한정이 아닌 예시로서 도시된다.
- 도 1은 일부 실시예에 따라 머신 학습 문제를 해결하기 위한 신경망 아키텍처를 선택하는 것이 구현될 수 있는 예시적인 시스템을 도시한다.
- 도 2는 일부 실시예에 따라 머신 학습 문제를 해결하기 위한 신경망 아키텍처를 선택하기 위한 예시적인 방법에 대한 흐름도를 도시한다.
- 도 3은 일부 실시예에 따라 오류율을 감소시키기 위한 예시적인 방법의 흐름도를 도시한다.

도 4a 및 도 4b는 일부 실시예에 따라 블록들 중에서 셀을 선택하기 위한 예시적인 방법의 흐름도를 도시한다.

도 5는 일부 실시예에 따라 머신 관독 가능 매체로부터 명령어를 관독할 수 있고 여기에서 설명된 방법 중 임의의 것을 수행할 수 있는 머신의 컴포넌트를 도시하는 블록도이다.

발명을 실시하기 위한 구체적인 내용

- [0007] 개관
- [0008] 본 개시는, 다른 무엇보다도, 개별적으로 다양한 기능을 제공하는 방법, 시스템 및 컴퓨터 프로그램 제품을 설명한다. 다음의 설명에서, 설명의 목적으로, 본 개시의 다른 실시예의 다양한 측면의 철저한 이해를 제공하기 위해서 복수의 특정 세부 사항이 설명된다. 그러나, 본 개시가 특정 세부 사항의 모두가 없더라도 실시될 수 있음은 통상의 기술자에게 명백할 것이다.
- [0009] 전술된 것과 같이, 복수의 다른 타입의 신경망 아키텍처(예를 들어, 피드포워드 신경망, 컨볼루션 신경망, 순환 신경망 등)가 알려져 있다. 주어진 머신 학습 문제(예를 들어, 회귀 문제, 순위 문제 또는 예를 들어, 새의 이미지를 이미지의 새의 타입에 따라 분류하는 것과 같이 주어진 공간에서 데이터를 분류하는 것과 같은 분류 문제)를 해결하기 위한 높은 성과의 신경망 아키텍처를 선택하는 것은 어려울 수 있다.
- [0010] 본 명세서에 설명된 기술의 일부 측면은 신경망 아키텍처가 훈련되기 이전에 신경망 아키텍처의 세트로부터 주어진 머신 학습 문제를 해결하기 위한 신경망 아키텍처를 선택하는 기술적 문제를 해결하는 것에 관한 것이다. 일부 측면의 결과로서, 고성능 신경망 아키텍처는 주어진 머신 학습 문제를 해결하기 위해 훈련되고, 덜 바람직한 아키텍처는 훈련되지 않으므로 효율적이다. 이는 높은 성과를 내지 않는 신경망을 사용하지 않으므로 계산 시간을 줄이고 효율성을 높일 수 있다.
- [0011] 일부 경우에, 이 문제에 대한 해결책은 서버에서 구현된다. 서버는 데이터 저장소를 통해 머신 학습 문제와 관련된 머신 학습 문제 공간 및 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스한다. 서버는 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 표현성을 캡처하는 적어도 하나의 표현성 측도를 계산한다. 서버는 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 훈련 가능성을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산한다. 서버는 적어도 하나의 표현성 측도, 적어도 하나의 훈련 가능성 측도 및 후보 신경망의 아키텍처에 근거하여 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택한다. 서버는 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공한다.
- [0012] 일부 경우에, 선택된 적어도 하나의 후보 신경망은 머신 학습 문제를 해결하기 위해 부분적으로 또는 완전히 훈련된다. 여기에 사용되는 것과 같이, "부분적으로 또는 완전히 훈련된" 신경망은 수렴의 일부 지표가 충족될 때까지 몇 개의 에포크(epoch) 동안 훈련되고 있거나 훈련된 것을 포함할 수 있다. 훈련된 적어도 하나의 후보 신경망은 머신 학습 문제를 해결하기 위해 머신 학습 문제 공간에서 실행된다. 이후, 서버는 훈련된 적어도 하나의 후보 신경망에 의해 생성된 머신 학습 문제에 대한 해결책을 제공한다.
- [0013] 일부 예에 따르면, 적어도 하나의 표현성 측도는 훈련되지 않은 후보 신경망에 의한 분류 문제 공간으로부터 샘플의 분리의 측도(예를 들어, 크기 또는 각도)를 나타낸다. 일부 예에 따르면, 적어도 하나의 훈련 가능성 측도는 머신 학습 문제 공간으로부터 샘플이 제공된 마지막 계층 및 첫 번째 계층에서 경사(gradient)의 함수를 나타낸다. 일부 예에 따르면, 표현성 및 훈련 가능성 측도는 표현성 및 훈련 가능성과 같은 아키텍처의 다른 속성을 캡처하는 신경망의 수량, 측도 또는 통계를 포함할 수 있다.
- [0014] 도 1은 일부 실시예에 따라 머신 학습 문제를 해결하기 위한 신경망 아키텍처를 선택하는 것이 구현될 수 있는 예시적인 시스템(100)을 도시한다. 도시된 것과 같이, 시스템(100)은 네트워크(140)를 통해 서로 연결된 서버(110), 데이터 저장소(120) 및 클라이언트 장치(130)를 포함한다. 네트워크(140)는 인터넷, 인트라넷, 근거리 네트워크, 광역 네트워크, 유선 네트워크, 무선 네트워크, 셀룰러 네트워크, WiFi 네트워크 등의 하나 이상을 포함한다.
- [0015] 클라이언트 장치(130)는 랩톱 컴퓨터, 데스크톱 컴퓨터, 이동 전화기, 태블릿 컴퓨터, 프로세서와 메모리를 구비한 스마트 텔레비전, 스마트 워치 등일 수 있다. 클라이언트 장치(130)는 사용자에게 출력을 디스플레이하거나 또는 사용자로부터 입력을 수신하는데 사용될 수 있다.
- [0016] 데이터 저장소(120)는 데이터베이스 또는 임의의 다른 데이터 스토리지 구조로서 구현될 수 있다. 도시된 것과

같이, 데이터 저장소는 머신 학습 문제 공간(122)을 저장한다. 머신 학습 문제 공간(122)은 신경망에 의해 분류될 데이터를 포함한다. 예를 들어, 머신 학습 문제 공간(122)은 새의 타입에 따라 분류될 새의 사진 또는 "중요한 이메일", "중요하지 않은 이메일" 또는 "스팸"으로 분류될 이메일 메시지를 포함할 수 있다.

[0017] 서버(110)는 하나 이상의 서버를 포함할 수 있다. 서버(110)는 복수의 서버를 포함하는 서버 팜(server farm)으로 구현될 수 있다. 도시된 것과 같이, 서버(110)는 훈련되지 않은 후보 신경망(112.1-n)(여기서 n은 2보다 크거나 같은 양의 정수임), 선택 모듈(114) 및 훈련 모듈(116)을 저장한다. 훈련되지 않은 후보 신경망(112.1-n)은 다양한 분류 작업에 사용될 수 있는 신경망이다. 예를 들어, 훈련되지 않은 후보 신경망(112.1-n)은 컨볼루션 신경망 또는 피드포워드 신경망의 훈련되지 않은 버전을 포함할 수 있다.

[0018] 선택 모듈(114)은 머신 학습 문제 공간(122)과 관련된 머신 학습 문제를 해결하기 위한 훈련을 위해 훈련되지 않은 후보 신경망(112.1-n) 중 적어도 하나를 선택한다. 선택 모듈의 예시적인 동작의 더 많은 세부 사항은 도 2와 함께 제공된다. 훈련 모듈(116)은 머신 학습 문제를 해결하기 위해 (선택 모듈(114)에 의해) (훈련되지 않은 후보 신경망(112.1-n)으로부터) 선택된 신경망(들)을 훈련한다. 훈련 이후, 훈련된 신경망은 머신 학습 문제 공간(122)(또는 다른 문제 공간)에서 데이터를 분류함으로써 머신 학습 문제를 해결하는데 사용될 수 있다.

[0019] 도 2는 일부 실시예에 따라 머신 학습 문제를 해결하기 위한 신경망 아키텍처를 선택하기 위한 예시적인 방법(200)에 대한 흐름도를 도시한다. 후술되는 것과 같이, 방법(200)은 도 1의 서버(110)의 선택 모듈(114)을 사용하여 구현된다. 그러나, 방법(200)은 시스템(100)의 아키텍처로 한정되지 않고 다른 아키텍처 또는 다른 시스템으로 구현될 수 있다.

[0020] 동작(210)에서, 선택 모듈(114)은 (예를 들어, 네트워크(140)를 통해) 해결해야 할 머신 학습 문제와 관련된 머신 학습 문제 공간(122)에 액세스한다. 선택 모듈(114)은 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망(112.1-n)에 액세스한다.

[0021] 동작(220)에서, 선택 모듈(114)은 각각의 훈련되지 않은 후보 신경망(112.1-n)에 대해 머신 학습 문제에 대한 후보 신경망의 표현성과 관련된 표현성 메트릭(들)을 계산한다. 표현성 메트릭(들)은 훈련되지 않은 후보 신경망에 의한 머신 학습 문제 공간(122)으로부터 샘플의 분리의 측도를 나타낸다. 분리의 측도는 크기 또는 각도일 수 있다.

[0022] 동작(230)에서, 선택 모듈(114)은 훈련되지 않은 후보 신경망(112.1-n)에 대해 머신 학습 문제에 대한 후보 신경망의 훈련 가능성과 관련된 훈련 가능성 메트릭(들)을 계산한다.

[0023] 동작(240)에서, 선택 모듈(114)은 표현성 측도(들) 및 훈련 가능성 측도(들)에 근거하여, 머신 학습 문제를 해결하기 위한 후보 신경망(들)을 선택한다. 선택은 또한 후보 신경망(들)의 아키텍처(들)에 근거한다. 후보 신경망(들)은 복수의 훈련되지 않은 후보 신경망(112.1-n)으로부터 선택된다. 일부 경우에, 선택 모듈(114)은 임계치를 초과하는 표현성 측도(들) 및 범위 내의 훈련 가능성 측도(들)를 갖는 후보 신경망(들)을 선택한다. 범위는 최소 범위와 최대 범위에 의해 정의된다.

[0024] 동작(250)에서, 선택 모듈(114)은 선택된 후보 신경망(들)을 나타내는 출력을 제공한다. 일부 경우에, 훈련 모듈(116)은 머신 학습 문제를 해결하기 위해 선택된 후보 신경망(들)을 훈련한다. 훈련 모듈(116)은 머신 학습 문제를 해결하기 위해 머신 학습 문제 공간(122)에서 훈련된 후보 신경망(들)을 실행한다. 서버(110)는 훈련된 후보 신경망(들)에 의해 생성된 머신 학습 문제에 대한 해결책을 (예를 들어, 디스플레이를 위해 클라이언트 장치(130)에, 또는 저장을 위해 데이터 저장소(120)에) 제공한다.

[0025] 본 명세서에 설명된 기술의 하나의 목표는 주어진 작업에 대한 신경망 아키텍처를 자동으로 선택하고 구성하는 것이다. 보다 구체적으로, 데이터 세트가 주어지면, 일부 측면은 계층 타입(예를 들어, 컨볼루션, 맥스풀, 완전 연결 등), 하이퍼파라미터(예를 들어, 스트라이드(stride) 크기, 컨볼루션 크기), 다른 모든 계층에 대한 연결 및 계층의 총 개수를 자동으로 식별한다. 일부 측면은 또한 사용할 훈련 알고리즘(예를 들어, 확률적 경사 하강법, RMSProp, Adam 등)과 가중치를 초기화하는 방법(예를 들어, Glorot, Normal, Laplace, Uniform 등)을 식별한다.

[0026] 일부 방식은 (1) 아키텍처 상에 검색 공간을 정의하는 방법(예를 들어, 제한되지 않거나 또는 제한됨); (2) 공간을 탐색하는 방법(예를 들어, 강화 학습, Monte Carlo Tree Search); (3) 검색을 안내하는데 사용하는 예측 모델(예를 들어, SMBO(sequential model based optimization), RNN(recurrent neural network), 유전(genetic) 알고리즘); 및 (4) 검색을 보다 효율적으로 안내하기 위해 저렴한 대리 기능을 사용하는지 여부에 근거하여 그룹화된다. 유전 알고리즘과 강화 학습의 경우, 위의 (2)와 (3)은, 공동으로 예측 모델을 학습하고 공

간을 탐색하기 때문에 하나로 합쳐진다.

[0027] 일부 실시예에서, 검색 공간은 SMBO에 의해 탐색된다. SMBO의 예시적인 모델은 Bayesian 순환 신경망일 수 있다. 랜덤 신경망의 특성의 세트는 신경망의 실제 성능에 대해 저렴한 대리(또는 통계) 역할을 하도록 정의된다. 일부 경우에, 예측과 예측에 대한 불확실성을 모두 출력하는 다른 모델(예를 들어, Gaussian 처리)이 사용된다. 일부 경우에, 저렴한 대리(또는 통계)가 모델의 입력으로서 제공되지 않지만 출력으로 간주된다. 저렴한 대리는 전술된 훈련 가능성 및 표현성 측도를 포함할 수 있다.

[0028] 일부 확장된 실시예에서, 검색 공간은 제한되지 않는다. 강화 학습 또는 Monte Carlo Tree Search는 검색 공간을 탐색하는데 사용된다. 일부 경우에, 예측을 돕기 위해 아키텍처 자체의 표현을 추가로 사용할 수도 있다.

[0029] 일부 측면은 훈련 이전에 훈련 이후의 데이터 세트에 대한 초기 모델의 성능을 예측하는 2가지 주요 측도에 관한 것이다. 2가지 측도는 모두 데이터 세트로부터 랜덤 배치(batch)에서 수집된 모델의 통계이다. 일부 예에서, 배치는 n개의 포인트 $\{(x_i, y_i)\}_{i=1}^n$ 로 구성되고; 일부 측면은 레이블 y_i 을 무시한다. 모델은 일반적으로 2개의 컴포넌트, 즉, 입력 공간을 잠재적 공간에 임베딩하는 심층 신경망과 임베딩을 가능한 레이블의 세트에 대한 확률 분포로 변환하는 소프트맥스(softmax)가 후속하는 완전히 연결된(fully connected) 선형 계층을 포함한다. f 를 전자, 즉 임베딩으로 하고, L 계층을 갖는다고 가정한다.

[0030] 메트릭 표현성은 식 1에 따라 정의되고, 식 1은 식 2로 근사한다. 식 1에서, $\{x_1, \dots, x_n\}$ 은 입력의 배치이다.

[0031]
$$\frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \frac{\|f(x_i) - f(x_j)\|^2}{\|x_i - x_j\|^2}$$
 식 1

[0032]
$$E_{x, x' \sim P} \frac{\|f(x) - f(x')\|^2}{\|x - x'\|^2}$$
 식 2

[0033] 식 2에서, E는 예상값을 나타내고 P는 x s에 대한 데이터 분포를 나타낸다. 직관적으로, 이 측도는 입력 공간을 확장하고 포인트를 분리하는 f 의 경향을 나타낸다. 더 큰 메트릭 표현성은 훈련 이후 더 나은 성능과 상관이 있어야 한다. 이 측도의 변형은 공통 배치의 모든 쌍을 선택하는 대신 (x, x') 쌍을 샘플링하고, 2가 아닌 다른 거

듭 제곱을 선택(예를 들어, 일부 $p > 0$ 에 대해 $\frac{\|f(x_i) - f(x_j)\|^p}{\|x_i - x_j\|^p}$)하고, 신경망의 "확장성"(입력 공간을 확장하는 경향)을 테스트하는 다른 방법을 포함한다.

[0034] 메트릭 표현성 특징에 더하여, 경사 변형은 각각의 이전 계층에서 경사 벡터를 얻기 위해 마지막 계층에 대한 랜덤 경사 벡터를 샘플링하고 각각의 샘플 x_i 에 대한 순방향 계산 이후 역방향 전파를 수행하는 것을 포함한다.

이러한 고정된 입력 x_i 및 고정된 마지막 계층 경사 벡터를 가정하면, 계층 l 에서 파라미터 w 에 대한 경사 벡터

는 $\frac{\partial E}{\partial w^{(l)}}|_{x=x_i}$ 으로 표시된다. 이후, 파라미터 w 의 경사 변형은 식 3으로 정의된다.

[0035]
$$\frac{1}{n} \sum_{i=1}^n \left| \log \left\| \frac{\partial E}{\partial w^{(L)}} \Big|_{x=x_i} \right\| - \log \left\| \frac{\partial E}{\partial w^{(1)}} \Big|_{x=x_i} \right\| \right|$$
 식 3

[0036] 즉, 경사 변형은 전형적인 데이터 포인트에서 얼마나 많은 경사 폭발 또는 소멸이 일어나는지에 대한 측도이다. 이 값이 클수록 (SGD(stochastic gradient descent)을 통한) 훈련이 어려울 것으로 예상되기 때문에 훈련된 성능이 더 나빠질 것으로 예상된다. 이 측도의 변형은 모든 x_i 에 대해 새로운 마지막 계층 경사 벡터를 샘플링하고, 합계를 식 4로 대체하고, 경사 폭발/소멸을 측도하는 다른 방법을 포함한다. 고정된 아키텍처 및 랜덤 초기화의 경우 위의 측도는 잔여 네트워크 성능을 예측한다.

[0037]
$$\max \left(\left\| \frac{\partial E}{\partial w^{(L)}} \Big|_{x=x_i} \right\| / \left\| \frac{\partial E}{\partial w^{(1)}} \Big|_{x=x_1} \right\|, \left\| \frac{\partial E}{\partial w^{(1)}} \Big|_{x=x_1} \right\| / \left\| \frac{\partial E}{\partial w^{(L)}} \Big|_{x=x_1} \right\| \right)$$
 식 4

[0038] 메트릭 표현성의 경우와 동일한 설정이 주어지면, 각도 표현성은 식 5로 정의된다.

[0039]
$$\frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} \frac{f(x_i) \cdot f(x_j)}{\|f(x_i)\| \|f(x_j)\|}$$
 식 5

[0040] 각도 표현성은 그들의 각도를 분리한다는 의미에서 f 가 입력 벡터와 얼마나 "비상관"하는지에 대한 측도이다. 따라서, 큰 각도 표현성은 더 나은 성능과 상관할 것으로 예상된다. 일부 방법에서, 예측 수량은 f 의 깊이가

$\frac{f(x_i) \cdot f(x_j)}{\|f(x_i)\| \|f(x_j)\|}$ 무한대로 갈수록 그 점근적 한계와의 $\frac{f(x_i) \cdot f(x_j)}{\|f(x_i)\| \|f(x_j)\|}$ 의 편차이다. 이 점근적 편차의 프록시(proxy)는 Cauchy 오차 $C_1 - C_{(1+1)}$ 이고, 여기서 C_1 은 1번째 계층까지의 네트워크의 각도 표현성이다.

[0041] 일부 경우에, 통계가 자동으로 학습될 수 있다. 본 명세서에 설명된 데이터 기반 통계는 모델의 최종 성능과 크게 상관되지만, 주어진 모델의 일반화 능력을 예측하기 위해 추출할 수 있는 가장 예측성이 좋은 통계가 아닐 수도 있다. 가장 예측성이 좋은 통계는 원시 통계(예를 들어, 이전 섹션에서 통계를 계산하는데 사용된 임베딩 및 경사 측도)의 복잡한 비선형 함수일 수 있다.

[0042] 그럼에도 불구하고, 머신 학습의 임의의 다른 함수 근사 문제와 마찬가지로, 훈련 데이터가 충분히 존재하는 경우, 이러한 복잡한 함수를 학습하는 것이 가능할 수 있다. 이것은 서버(110)가 원시 데이터로부터 예측성이 좋은 통계를 학습하는 또 다른 버전의 프레임워크에 동기를 부여할 것이다. 특히, 일부 측면은 입력이 마지막 계층(즉, 위의 $f(x)$)에서의 데이터 표현, 마지막 및 첫 번째 계층의 경사 등인 신경망을 사용할 수 있다.

[0043] 일부 측면의 일반적인 버전에서, 알고리즘 1의 절차는 미리 결정된 원하는 오류율에 도달하거나 또는 미리 결정된 총 계산 비용에 도달할 때까지 반복된다.

[0044] **알고리즘 1**

[0045] 1. 복수의 DNN(deep neural network) 아키텍처를 제안한다. DNN 아키텍처는 미리 정의된 아키텍처의 공간으로부터 샘플링되거나 또는 일반적인 빌딩 블록을 사용하여 구축될 수 있다.

[0046] 2. Glorot Normal 초기화를 사용하여 아키텍처의 가중치를 초기화한다. 독립적인 Normal 분포와 Laplace 분포가 다른 실시예에서 사용될 수 있다.

[0047] 3. 각각의 아키텍처에 대한 메트릭 표현성 및 경사 변형을 계산한다.

[0048] 4. 각각의 아키텍처의 메트릭 표현성 및 경사 변형을 Bayesian RNN(recurrent neural network)에 대한 입력으로 포함시켜 아키텍처의 성능을 예측한다. 이 작업을 위해 Gaussian 처리가 또한 사용될 수 있다. 원칙적으로 예측에 대한 분포를 출력하는 임의의 확률적 머신 학습 알고리즘이 사용될 수 있다.

[0049] 5. 각각의 아키텍처에 대해 Bayesian RNN을 사용하여 계산된 사후 평균 및 분산을 예상되는 개선, 신뢰 상한(upper confidence bound) 또는 Thompson 샘플링과 같은 획득 함수에 대한 입력으로 사용한다.

[0050] 6. 아키텍처를 임계값을 초과하는 획득 함수에 대한 값으로 샘플링한다.

[0051] 7. 아키텍처를 임계값을 초과하는 획득 함수의 값으로 훈련하고, 제공된 세트에 대한 성능을 평가한다. 동작 4에서 사용된 모델을 관찰된 성능으로 업데이트한다. 동작 1로 복귀한다.

[0052] 알고리즘 1은 도 3에 요약되어 있다. 도 3은 일부 실시예에 따라 오류율을 감소시키기 위한 예시적인 방법의 흐름도(300)를 도시한다.

[0053] 동작(310)에서, 서버(예를 들어, 서버(110))는 복수의 DNN 아키텍처를 제안한다. DNN 아키텍처는 미리 정의된 아키텍처의 공간으로부터 샘플링되거나 또는 일반적인 빌딩 블록을 사용하여 구축될 수 있다.

[0054] 동작(320)에서, 서버는 Glorot Normal 초기화를 사용하여 아키텍처의 가중치를 초기화한다. 독립적인 Normal 분포와 Laplace 분포가 다른 실시예에서 사용될 수 있다.

[0055] 동작(330)에서, 서버는 각각의 아키텍처의 메트릭 표현성 및 경사 변형을 계산한다.

[0056] 동작(340)에서, 서버는 각각의 아키텍처의 메트릭 표현성 및 경사 변형을 Bayesian RNN에 대한 입력으로 포함시켜 아키텍처의 성능을 예측한다. 이 작업을 위해 Gaussian 처리가 또한 사용될 수 있다. 원칙적으로 예측에 대한 분포를 출력하는 임의의 확률적 머신 학습 알고리즘이 사용될 수 있다.

- [0057] 동작(350)에서, 서버는 각각의 아키텍처에 대해 Bayesian RNN을 사용하여 계산된 사후 평균 및 분산을 예상되는 개선, 신뢰 상한 또는 Thompson 샘플링과 같은 획득 함수에 대한 입력으로서 사용한다.
- [0058] 동작(360)에서, 서버는 아키텍처(들)를 임계값을 초과하는 획득 함수에 대한 값(들)으로 샘플링한다.
- [0059] 동작(370)에서, 서버는 아키텍처(들)를 임계값을 초과하는 획득 함수에 대한 값(들)으로 훈련하고 제공된 세트에 대한 성능을 평가한다. 동작(340)의 Bayesian RNN 모델은 관찰된 성능에 근거하여 업데이트된다. 동작(370) 이후, (동작(360)의) 획득 함수에 대한 값이 충분한 경우, 방법(300)은 종료된다. 그렇지 않으면, 방법(300)은 동작(310)으로 복귀한다.
- [0060] 특정 구현예가 여기에서 설명된다. 반복적인 기본 단위(셀)가 복수 회 반복되는 고정된 심층 신경망 아키텍처에서, 일부 측면은 가장 높은 정확도로 셀을 추론하는 것에 관한 것이다. 먼저, 셀은 블록의 특정 개수의 방향성 비순환 그래프로 정의된다. 각각의 블록은 2개의 입력(I_1, I_2)을 취하고, 이들 각각에 대해 동작(O_1, O_2)을 수행하고, 이들 두 동작의 출력의 합계를 반환한다. 블록에 대해 가능한 입력의 세트는 셀 내의 모든 이전 블록의 출력과 이전 2개 셀의 출력이다. 이들 입력에 대한 동작은 컨볼루션, 맥스풀링 등과 같은 표준 신경망 계층이다. 심층 신경망에서 이러한 기본 단위의 자동 및 정확한 검색은 일반적으로 더 큰 데이터 세트에 대한 애플리케이션에서 그들의 성능이 일반화되기 때문에 중요하다. B 블록을 갖는 최상의 셀을 찾기 위해, 알고리즘 2가 사용된다.
- [0061] 일부 경우에, 하나의 블록(C)에서 셀의 개수가 적을 수 있으므로 알고리즘 2의 검색은 철저할 수 있다. 예를 들어, 8개의 가능한 동작에 대해 검색 공간은 $8 * 8 * 2 * 2 = 256$ 이고, 예를 들어 256개의 GPU(graphics processing unit)를 병렬로 사용하여 철저하게 탐색될 수 있다.
- [0062] **알고리즘 2**
- [0063] 1. 하나의 블록에서 셀의 세트(C)를 고려한다.
- [0064] 2. C의 각각의 셀에 대해:
 - [0065] · Glorot Normal 초기화를 사용하여 아키텍처의 가중치를 초기화한다. 동일하게 분포된(identically distributed; iid) Normal 분포 및 Laplace 분포가 다른 실시예에서 사용될 수 있다.
 - [0066] · 메트릭 표현성과 경사 변형을 계산한다. 이것은 복수의 초기화를 통해 수행될 수 있고 평균 및 분산이 보고될 수 있다.
 - [0067] · 고정된 개수의 에포크 동안 신경망을 훈련한다.
 - [0068] · 테스트 세트에 대한 정확도를 계산한다.
- [0069] 3. C의 셀의 블록 개수가 B인 경우, 가장 높은 테스트 정확도를 갖는 셀에서 중지하고 반환한다.
- [0070] 4. 셀 아키텍처와 랜덤 통계를 입력으로 취하는 BRNN(Bayesian recurrent neural network)을 훈련하고 테스트 세트 정확도를 예측한다.
- [0071] 5. 새로운 블록을 추가하고 C의 모든 셀을 확장하여 얻은 셀을 포함하는, 세트 D를 고려한다.
- [0072] 6. D의 셀에 대한 메트릭 표현성과 경사 변형을 계산한다. (예를 들어, D의 셀의 개수가 많기 때문에) 이 계산의 비용이 너무 많은 경우, (i) 단지 셀 구조만을 취하는 추가 BRNN을 훈련하고 테스트 세트 정확도를 예측하며, (ii) 예측된 테스트 정확도에 근거하여 세트 D를 필터링하는데 이를 사용한다.
- [0073] 7. 동작 4에서 BRNN을 사용하여 다음 반복에서 고려할 셀의 세트 $\epsilon \subset D$ 를 결정한다. 이 결정은 Bayesian 최적화 프레임워크 내에서 이용과 탐색을 절충하여 이루어진다.
- [0074] 8. $\epsilon \rightarrow C$ 하고, 2단계로 이동한다.
- [0075] 알고리즘 2는 도 4a 및 도 4b에 요약되어 있다. 도 4a 및 도 4b는 일부 실시예에 따라 블록들 중에서 셀을 선택하기 위한 예시적인 방법(400)의 흐름도를 도시한다.
- [0076] 도 4a에 도시된 것과 같이, 동작(410)은 하위 동작(412-418)을 포함한다. 동작(410)에서, 서버(예를 들어, 서버(110))는 블록 C의 각각의 셀에 대해 하위 동작(412-418)을 수행한다.
- [0077] 하위 동작(412)에서, 서버는 Glorot Normal 초기화를 사용하여 아키텍처의 가중치를 초기화한다. 동일 분포된

(iid) Normal 분포 및 Laplace 분포가 다른 실시예에서 사용될 수 있다.

- [0078] 하위 동작(414)에서, 서버는 메트릭 표현성 및 경사 변형을 계산한다. 이것은 복수의 초기화를 통해 수행될 수 있고, 평균 및 분산이 보고될 수 있다.
- [0079] 하위 동작(416)에서, 서버는 고정된 개수의 에포크 동안 신경망을 훈련한다.
- [0080] 하위 동작(418)에서, 서버는 테스트 세트에 대한 정확도를 계산한다. C의 각각의 셀에 대해 동작(410)이 완료된 이후, 방법(400)은 동작(420)으로 진행한다.
- [0081] 동작(420)에서, 서버는 C의 블록의 개수가 미리 정해진 범위 내에 있는지 여부를 결정한다. 만약 그렇다면, 방법(400)은 동작(430)으로 진행한다. 그렇지 않다면, 방법(400)은 동작(440)으로 진행한다.
- [0082] 동작(430)에서, C의 블록의 개수가 미리 정해진 범위 내에 있는 경우, 서버는 테스트 정확도가 가장 높은 C의 셀을 반환한다. 동작(430) 이후, 방법(400)은 종료된다.
- [0083] 동작(440)에서, C의 블록의 개수가 미리 정해진 범위 내에 있지 않는 경우, 서버는 셀 아키텍처와 랜덤 통계를 입력으로 취하는 BRNN을 훈련하고 테스트 세트 정확도를 예측한다. 동작(440) 이후, 방법(400)은 도 4b에 도시된 동작(450)으로 진행한다.
- [0084] 도 4b에 도시된 것과 같이, 동작(450)에서, 서버는 새로운 블록을 추가하고 C의 모든 셀을 확장하여 획득한 셀을 포함하는 세트 D를 생성한다.
- [0085] 동작(460)에서, 서버는 D의 셀에 대한 메트릭 표현성 및 경사 변형을 계산한다. (예를 들어, D의 셀의 개수가 많기 때문에) 이 계산의 비용이 너무 많은 경우, 서버는 (i) 셀 구조만 취하는 추가 BRNN을 훈련하여 테스트 세트 정확도를 예측하고, (ii) 추가 BRNN을 사용하여 예측된 테스트 정확도에 근거하여 세트 D를 필터링한다.
- [0086] 동작(470)에서, 서버는 동작(440)에서 훈련된 BRNN을 사용하여 다음 반복에서 고려할 셀들의 세트 $\varepsilon \subset D$ 를 결정한다. 이 결정은 Bayesian 최적화 프레임워크 내에서 이용과 탐색을 절충하여 이루어진다.
- [0087] 동작(480)에서 서버는 ε 를 C로 저장한다. 동작(480) 이후, 방법(400)은 도 4a의 동작(420)으로 복귀한다.
- [0088] 번호가 붙은 예
- [0089] 특정 실시예는 여기에서 번호가 붙은 예 1, 2, 3 등으로 설명된다. 이들 번호가 붙은 예는 단지 예로서 제공되며 본 기술을 한정하지 않는다.
- [0090] 예 1은, 프로세싱 하드웨어 및 프로세싱 하드웨어로 하여금 동작을 수행하도록 하는 명령어를 저장하는 메모리를 포함하는 시스템으로서, 동작은, 머신 학습 문제와 관련된 머신 학습 문제 공간 및 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것과, 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 표현성(expressivity)을 캡처하는 적어도 하나의 표현성 측도를 계산하는 것과, 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 훈련 가능성(trainability)을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산하는 것과, 적어도 하나의 표현성 측도 및 적어도 하나의 훈련 가능성 측도에 근거하여 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것과, 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공하는 것을 포함한다.
- [0091] 예 2에서, 예 1의 대상은, 적어도 하나의 표현성 측도는 훈련되지 않은 후보 신경망에 의한 머신 학습 문제 공간으로부터 샘플의 분리의 측도를 나타내는 것임을, 포함한다.
- [0092] 예 3에서, 예 2의 대상은, 분리의 측도는 크기임을, 포함한다.
- [0093] 예 4에서, 예 2 및 예 3의 대상은, 분리의 측도는 각도임을, 포함한다.
- [0094] 예 5에서, 예 1 내지 예 4의 대상은, 적어도 하나의 훈련 가능성 측도는 훈련의 제 1 단계 동안 후보 신경망에서 가중치의 확률적 경사 하강을 나타내는 것임을, 포함한다.
- [0095] 예 6에서, 예 1 내지 예 5의 대상은, 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것은, 임계값을 초과하는 적어도 하나의 표현성 측도 및 범위 내의 적어도 하나의 훈련 가능성 측도를 갖는 적어도 하나의 후보 신경망을 선택하는 것을 포함하고, 범위는 최소 범위 및 최대 범위에 의해 정의되는 것임을, 포함한다.
- [0096] 예 7에서, 예 1 내지 예 6의 대상은, 동작은 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 훈

련하는 것을 더 포함하는 것임을, 포함한다.

- [0097] 예 8에서, 예 7의 대상은, 동작은 머신 학습 문제를 해결하기 위해 머신 학습 문제 공간에서 훈련된 적어도 하나의 후보 신경망을 실행하는 것 및 훈련된 적어도 하나의 후보 신경망에 의해 생성된 머신 학습 문제에 대한 해결책을 제공하는 것을 더 포함하는 것임을, 포함한다.
- [0098] 예 9는 하나 이상의 머신으로 하여금 동작을 수행하도록 하는 명령어를 저장하는 비일시적 머신 판독 가능 매체로서, 동작은, 머신 학습 문제와 관련된 머신 학습 문제 공간 및 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것과, 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 표현성을 캡처하는 적어도 하나의 표현성 측도를 계산하는 것과, 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 훈련 가능성을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산하는 것과, 적어도 하나의 표현성 측도 및 적어도 하나의 훈련 가능성 측도에 근거하여 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것과, 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공하는 것을 포함한다.
- [0099] 예 10에서, 예 9의 대상은, 적어도 하나의 표현성 측도는 훈련되지 않은 후보 신경망에 의한 머신 학습 문제 공간으로부터 샘플의 분리의 측도를 나타내는 것임을, 포함한다.
- [0100] 예 11에서, 예 10의 대상은, 분리의 측도는 크기임을, 포함한다.
- [0101] 예 12에서, 예 10 및 예 11의 대상은, 분리의 측도는 각도임을, 포함한다.
- [0102] 예 13에서, 예 9 내지 예 12의 대상은, 적어도 하나의 훈련 가능성 측도는 훈련의 제 1 단계 동안 후보 신경망에서 가중치의 확률적 경사 하강을 나타내는 것임을, 포함한다.
- [0103] 예 14에서, 예 9 내지 13의 대상은, 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것은, 임계값을 초과하는 적어도 하나의 표현성 측도 및 범위 내의 적어도 하나의 훈련 가능성 측도를 갖는 적어도 하나의 후보 신경망을 선택하는 것을 포함하고, 범위는 최소 범위 및 최대 범위에 의해 정의되는 것임을, 포함한다.
- [0104] 예 15는, 머신 학습 문제와 관련된 머신 학습 문제 공간 및 머신 학습 문제를 해결하기 위한 복수의 훈련되지 않은 후보 신경망에 액세스하는 것과, 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 표현성을 캡처하는 적어도 하나의 표현성 측도를 계산하는 것과, 각각의 훈련되지 않은 후보 신경망에 대해 머신 학습 문제에 대한 후보 신경망의 훈련 가능성을 캡처하는 적어도 하나의 훈련 가능성 측도를 계산하는 것과, 적어도 하나의 표현성 측도 및 적어도 하나의 훈련 가능성 측도에 근거하여 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것과, 선택된 적어도 하나의 후보 신경망을 나타내는 출력을 제공하는 것을 포함하는 방법이다.
- [0105] 예 16에서, 예 15의 대상은, 적어도 하나의 표현성 측도는 훈련되지 않은 후보 신경망에 의한 머신 학습 문제 공간으로부터 샘플의 분리의 측도를 나타내는 것임을, 포함한다.
- [0106] 예 17에서, 예 16의 대상은, 분리의 측도는 크기임을, 포함한다.
- [0107] 예 18에서, 예 16 및 예 17의 대상은, 분리의 측도는 각도임을, 포함한다.
- [0108] 예 19에서, 예 15 내지 예 18의 대상은, 적어도 하나의 훈련 가능성 측도는 훈련의 제 1 단계 동안 후보 신경망에서 가중치의 확률적 경사 하강을 나타내는 것임을, 포함한다.
- [0109] 예 20에서, 예 15 내지 예 19의 대상은, 머신 학습 문제를 해결하기 위한 적어도 하나의 후보 신경망을 선택하는 것은, 임계값을 초과하는 적어도 하나의 표현성 측도 및 범위 내의 적어도 하나의 훈련 가능성 측도를 갖는 적어도 하나의 후보 신경망을 선택하는 것을 포함하고, 범위는 최소 범위 및 최대 범위에 의해 정의되는 것임을, 포함한다.
- [0110] 예 21은 프로세싱 회로에 의해 실행될 때 프로세싱 회로로 하여금 예 1 내지 예 20 중 임의의 것을 구현하는 동작을 수행하도록 하는 명령어를 포함하는 적어도 하나의 머신 판독 가능 매체이다.
- [0111] 예 22는 예 1 내지 예 20 중 임의의 것을 구현하는 수단을 포함하는 장치이다.
- [0112] 예 23은 예 1 내지 예 20 중 임의의 것을 구현하는 시스템이다.

- [0113] 예 24는 예 1 내지 예 20 중 임의의 것을 구현하는 방법이다.
- [0114] 컴포넌트 및 로직
- [0115] 본 문서에서 특정 실시예는 로직(logic) 또는 복수의 컴포넌트 또는 메커니즘을 포함하는 것으로서 설명된다. 컴포넌트는 소프트웨어 컴포넌트(예를 들어, 머신 판독 가능 매체 상에 체현된 코드) 또는 하드웨어 컴포넌트를 구성할 수 있다. "하드웨어 컴포넌트"는 특정 동작을 수행할 수 있는 유형적인 유닛(tangible unit)이고 특정 물리적 방식으로 구성되거나 또는 배열될 수 있다. 다양한 예시적 실시예에서, 하나 이상의 컴퓨터 시스템(예를 들어, 독립형 컴퓨터 시스템, 클라이언트 컴퓨터 시스템, 또는 서버 컴퓨터 시스템) 또는 컴퓨터 시스템의 하나 이상의 하드웨어 컴포넌트(예를 들어, 프로세서 또는 프로세서의 그룹)은 본 명세서에 설명된 것과 같이 특정 동작을 수행하도록 동작하는 하드웨어 컴포넌트로서 소프트웨어(예를 들어, 애플리케이션 또는 애플리케이션 부분)에 의해 구성될 수 있다.
- [0116] 일부 실시예에서, 하드웨어 컴포넌트는 기계적으로, 전자적으로, 또는 이들의 임의의 적합한 조합으로 구현될 수 있다. 예를 들어, 하드웨어 컴포넌트는 특정 동작을 수행하도록 영속적으로 구성된(permanently configured) 전용 회로 또는 로직을 포함할 수 있다. 예를 들어, 하드웨어 컴포넌트는 FPGA(Field-Programmable Gate Array) 또는 ASIC(Application Specific Integrated Circuit)와 같은 특수 목적 프로세서(special-purpose processor)일 수 있다. 하드웨어 컴포넌트는 또한 특정 동작을 수행하도록 소프트웨어에 의해 임시적으로 구성된(temporarily configured) 프로그램 가능 로직 또는 회로를 포함할 수 있다. 예를 들어, 하드웨어 컴포넌트는 범용 프로세서(general-purpose processor) 또는 다른 프로그램 가능 프로세서에 의해 실행되는 소프트웨어를 포함할 수 있다. 일단 이러한 소프트웨어에 의해 구성되면, 하드웨어 컴포넌트는 구성된 기능을 수행하도록 고유하게 맞춤화된(tailored) 특정 머신(또는 머신의 특정 컴포넌트)이 되고 더 이상 범용 프로세서가 아니다. 하드웨어 컴포넌트를 기계적으로, 전용이고 영속적으로 구성된 회로로, 또는 (예를 들어, 소프트웨어에 의해 구성된) 임시적으로 구성된 회로로 구현한다는 결정은 비용 및 시간을 고려하여 추진될(driven) 수 있음이 인식될 것이다.
- [0117] 따라서, 문구 "하드웨어 컴포넌트"는 본 명세서에 설명된 특정 방식으로 동작하거나 또는 특정 동작을 수행하도록 물리적으로 구축되거나(physically constructed), 영속적으로 구성되거나(예를 들어, 고정 배선됨(hardwired)), 임시적으로 구성된(예를 들어, 프로그래밍된(programmed)) 레코드(record)인, 유형적인 레코드를 포함하는 것으로 이해되어야 한다. 여기에서 사용되는 것과 같이, "하드웨어 구현 컴포넌트(hardware-implemented component)"는 하드웨어 컴포넌트를 가리킨다. 하드웨어 컴포넌트가 임시적으로 구성된(예를 들어, 프로그래밍된) 실시예를 고려하면, 하드웨어 컴포넌트 각각이 시간에서의 임의의 하나의 인스턴스(instance)에서 구성되거나 또는 인스턴스화될(instantiated) 필요는 없다. 예를 들어, 하드웨어 컴포넌트가, 특수 목적 프로세서가 되도록 소프트웨어에 의해 구성되는 범용 프로세서를 포함하는 경우, 범용 프로세서는 다른 시간에 각각 다른 특수 목적 프로세서(예를 들어, 다른 하드웨어 컴포넌트를 포함함)로서 구성될 수 있다. 이에 따라 소프트웨어는 특정한 프로세서 또는 프로세서들을 구성하여, 예를 들어 시간의 하나의 인스턴스에서 특정한 하드웨어 컴포넌트를 구성하고 시간의 다른 인스턴스에서 다른 하드웨어 컴포넌트를 구성한다.
- [0118] 하드웨어 컴포넌트는 다른 하드웨어 컴포넌트에 정보를 제공하고 이로부터 정보를 수신할 수 있다. 따라서, 설명된 하드웨어 컴포넌트는 통신 가능하게 커플링된(communicatively coupled) 것으로 간주될 수 있다. 여러 하드웨어 컴포넌트가 동시적으로(contemporaneously) 존재하는 경우에, 통신은 하드웨어 컴포넌트 중 둘 이상 간의 (예를 들어, 적절한 회로 및 버스를 통한) 신호 전송을 통해서 달성될 수 있다. 복수의 하드웨어 컴포넌트가 다른 시간에 구성되거나 또는 인스턴스화되는 실시예에서, 이러한 하드웨어 컴포넌트 사이의 통신은, 예를 들어, 복수의 하드웨어 컴포넌트가 액세스하는 메모리 구조 내의 정보의 저장 및 인출(retrieval)을 통해서, 달성될 수 있다. 예를 들어, 하나의 하드웨어 컴포넌트가 동작을 수행하고 그 동작의 출력을 그것이 통신 가능하게 커플링된 메모리 장치에 저장할 수 있다. 그러면 다른 하드웨어 컴포넌트가 나중에 메모리 장치에 액세스하여 저장된 출력을 인출하고 처리할 수 있다. 하드웨어 컴포넌트는 또한 입력 또는 출력 장치와의 통신을 개시할 수 있고, 리소스(예를 들어, 정보의 집합(collection))에 대해 동작할 수 있다.
- [0119] 본 명세서에 설명된 예시적 방법의 다양한 동작은, 적어도 부분적으로, 관련된 동작을 수행하도록 (예를 들어, 소프트웨어에 의해) 임시적으로 구성되거나 또는 영속적으로 구성된 하나 이상의 프로세서에 의해 수행될 수 있다. 임시적으로 구성되든 또는 영속적으로 구성되든, 이러한 프로세서는 본 명세서에 설명된 하나 이상의 동작 또는 기능을 수행하도록 동작하는 프로세서 구현(processor-implemented) 컴포넌트를 구성할 수 있다. 여기에 사용되는 것과 같이, "프로세서 구현 컴포넌트"는 하나 이상의 프로세서를 사용하여 구현된 하드웨어 컴포넌트

를 가리킨다.

- [0120] 유사하게, 본 명세서에 설명된 방법은 적어도 부분적으로 프로세서 구현형일 수 있는데, 특정한 프로세서 또는 프로세서들은 하드웨어의 예이다. 예를 들어, 방법의 동작 중 적어도 일부는 하나 이상의 프로세서 또는 프로세서 구현 컴포넌트에 의해 수행될 수 있다. 더욱이, 하나 이상의 프로세서는 또한 "클라우드 컴퓨팅(cloud computing)" 환경에서 또는 "SaaS(Software as a Service)"로서 관련 동작의 수행을 지원하도록 동작할 수 있다. 예를 들어, 동작 중 적어도 일부는 (프로세서를 포함하는 머신의 예로서) 컴퓨터의 그룹에 의해 수행될 수 있는데, 이들 동작은 네트워크(예를 들어, 인터넷)를 통해, 그리고 하나 이상의 적절한 인터페이스(예를 들어, API)를 통해 액세스될 수 있다.
- [0121] 동작 중 특정의 것의 수행은, 단일 머신 내에 상주하는(residing) 것뿐만 아니라, 복수의 머신에 걸쳐 배치된 프로세서들 사이에 분산될 수 있다. 일부 예시적 실시예에서, 프로세서 또는 프로세서 구현 컴포넌트는 단일의 지리적 위치(예를 들어, 가정 환경, 사무실 환경 또는 서버 팜 내)에 위치될 수 있다. 다른 예시적 실시예에서, 프로세서 또는 프로세서 구현 컴포넌트는 복수의 지리적 위치에 걸쳐 분산될 수 있다.
- [0122] 예시적인 머신 및 소프트웨어 아키텍처
- [0123] 도 1 내지 도 4와 함께 설명된 컴포넌트, 방법, 애플리케이션 등은 일부 실시예에서 머신 및 관련 소프트웨어 아키텍처의 맥락에서 구현된다. 이하의 섹션은 개시된 실시예와 함께 사용하는데 적합한 대표적인 소프트웨어 아키텍처(들) 및 머신(예를 들어, 하드웨어) 아키텍처(들)를 설명한다.
- [0124] 특정한 목적에 맞춤화된 장치 및 머신을 창출하기 위해 하드웨어 아키텍처와 함께 소프트웨어 아키텍처가 사용된다. 예를 들어, 특정한 소프트웨어 아키텍처와 커플링된 특정한 하드웨어 아키텍처는 모바일 전화, 태블릿 장치 등과 같은 모바일 장치를 창출할 것이다. 약간 다른 하드웨어 및 소프트웨어 아키텍처가 "사물 인터넷(internet of things)"에서 사용되는 스마트 장치(smart device)를 산출할 수 있으나, 또 다른 조합은 클라우드 컴퓨팅 아키텍처 내에서 사용되는 서버 컴퓨터를 만들어낸다. 통상의 기술자는 개시된 대상을 여기에 포함된 개시와는 다른 맥락에서 어떻게 구현할 것인지를 쉽게 이해할 수 있으므로, 이러한 소프트웨어 및 하드웨어 아키텍처의 모든 조합이 여기에 제시되는 것은 아니다.
- [0125] 도 5는, 일부 예시적 실시예에 따라, 머신 관독 가능 매체(예를 들어, 머신 관독 가능 저장 매체)로부터 명령어를 관독하고 본 명세서에 설명된 방법 중 임의의 하나 이상을 수행할 수 있는 머신(500)의 컴포넌트들을 도시하는 블록도이다. 구체적으로, 도 5는, 머신(500)으로 하여금 본 명세서에 설명된 방법 중 임의의 하나 이상을 수행하도록 하는 명령어(516)(예를 들어, 소프트웨어, 프로그램, 애플리케이션, 애플릿(applet), 앱(app) 또는 다른 실행 가능 코드(executable code))가 내부에서 실행될 수 있는 머신(500)의 도식적 표현을 컴퓨터 시스템의 예시적 형태로 도시한다. 명령어(516)는 일반적인, 프로그래밍되지 않은(non-programmed) 머신을, 설명되고 도시된 기능을 설명된 방식으로 수행하도록 프로그래밍된 특정한 머신으로 변형한다. 이와 다른 실시예에서, 머신(500)은 독립형 장치(standalone device)로서 동작하거나, 다른 머신에 커플링될(예를 들어, 네트워크링될(networked)) 수 있다. 네트워크링된 배치에서, 머신(500)은 서버-클라이언트 네트워크 환경(server-client network environment) 내의 서버 머신 또는 클라이언트 머신의 입장에서, 또는 피어 대 피어(peer-to-peer) (또는 분산형(distributed)) 네트워크 환경 내의 피어 머신(peer machine)으로서 동작할 수 있다. 머신(500)은 서버 컴퓨터, 클라이언트 컴퓨터, PC, 태블릿 컴퓨터, 랩톱 컴퓨터, 넷북(netbook), PDA(personal digital assistant), 엔터테인먼트 미디어 시스템(entertainment media system), 셀룰러 전화, 스마트 폰, 모바일 장치, 웨어러블 장치(wearable device)(예를 들어, 스마트 워치(smart watch)), 스마트 홈 장치(smart home device)(예를 들어, 스마트 기기(smart appliance)), 다른 스마트 장치, 웹 기기(web appliance), 네트워크 라우터(network router), 네트워크 스위치(network switch) 또는 네트워크 브릿지(network bridge), 또는 머신(500)에 의해 취해질 행동을 지정하는 명령어(516)를 순차적으로 또는 다르게 실행할 수 있는 임의의 머신을 포함할 수 있지만, 이에 한정되는 것은 아니다. 또한, 오직 단일 머신(500)이 도시되나, 용어 "머신"은 또한 본 명세서에 설명된 방법 중 임의의 하나 이상을 수행하기 위해 개별적으로 또는 공동으로 명령어(516)를 실행하는 머신(500)의 집합을 포함하도록 취해질 것이다.
- [0126] 머신(500)은 프로세서(510), 메모리/스토리지(530) 및 I/O 컴포넌트(550)를 포함할 수 있는데, 이는, 예를 들어 버스(502)를 통해서 서로 통신하도록 구성될 수 있다. 예시적 실시예에서, 프로세서(510)(예를 들어, CPU(Central Processing Unit), RISC(Reduced Instruction Set Computing) 프로세서, CISC(Complex Instruction Set Computing) 프로세서, GPU(Graphics Processing Unit), DSP(Digital Signal Processor), ASIC, RFIC(Radio-Frequency Integrated Circuit), 다른 프로세서, 또는 이들의 임의의 적합한 조합)는, 예를

들어, 명령어(516)를 실행할 수 있는 프로세서(512) 및 프로세서(514)를 포함할 수 있다. 용어 "프로세서"는 동시적으로 명령어를 실행할 수 있는 2 이상의 독립적인 프로세서(때때로 "코어"로 지칭됨)를 포함할 수 있는 다중 코어(multi-core) 프로세서를 포함하도록 의도된다. 도 5가 복수의 프로세서(510)를 도시하나, 머신(500)은 단일 코어를 가진 단일 프로세서, 복수의 코어를 가진 단일 프로세서(예를 들어, 다중 코어 프로세서), 단일 코어를 가진 복수의 프로세서, 복수의 코어를 가진 복수의 프로세서, 또는 이들의 임의의 조합을 포함할 수 있다.

[0127] 메모리/스토리지(530)는 주 메모리(main memory) 또는 다른 메모리와 같은 메모리(532) 및 스토리지 유닛(536)을 포함할 수 있고, 양자 모두, 예를 들어 버스(502)를 통해 프로세서(510)에 액세스할 수 있다. 스토리지 유닛(536) 및 메모리(532)는 본 명세서에 설명된 방법 또는 기능 중 임의의 하나 이상을 체현하는 명령어(516)를 저장한다. 명령어(516)는 또한, 머신(500)에 의한 실행 동안에, 메모리(532) 내에, 스토리지 유닛(536) 내에, 프로세서(510) 중 적어도 하나 내에(예를 들어, 프로세서의 캐시 메모리(cache memory) 내에), 또는 이들의 임의의 적합한 조합에, 완전히 또는 부분적으로, 상주할 수 있다. 따라서, 메모리(532), 스토리지 유닛(536) 및 프로세서(510)의 메모리는 머신 관독 가능 매체의 예이다.

[0128] 여기에 사용되는 것과 같이, "머신 관독 가능 매체"는 임시적으로 또는 영속적으로 명령어(예를 들어, 명령어(516)) 및 데이터를 저장할 수 있는 장치를 의미하고, RAM(random-access memory), ROM(read-only memory), 버퍼 메모리, 플래시 메모리, 광학 매체, 자기적 매체, 캐시 메모리, 다른 타입의 스토리지(예를 들어, EEPROM(Erasable Programmable Read-Only Memory)) 및/또는 이들의 임의의 적합한 조합을 포함할 수 있지만, 이에 한정되는 것은 아니다. 용어 "머신 관독 가능 매체"는 명령어(516)를 저장할 수 있는 단일 매체 또는 복수의 매체(예를 들어, 중앙화된 또는 분산된 데이터베이스, 또는 관련된 캐시 및 서버)를 포함하는 것으로 이해되어야 한다. 용어 "머신 관독 가능 매체"는 또한 머신(예를 들어, 머신(500))에 의한 실행을 위한 명령어(예를 들어, 명령어(516))를, 명령어가 머신의 하나 이상의 프로세서(예를 들어, 프로세서(510))에 의해 실행되는 경우 머신으로 하여금 본 명세서에 설명된 방법 중 임의의 하나 이상을 수행하게 하도록, 저장할 수 있는 임의의 매체, 또는 복수 매체의 조합을 포함하도록 이해되어야 한다. 따라서, "머신 관독 가능 매체"는 단일 스토리지 기구 또는 장치는 물론, 복수의 스토리지 기구 또는 장치를 포함하는 "클라우드 기반"(cloud-based) 스토리지 시스템 또는 스토리지 네트워크도 가리킨다. 용어 "머신 관독 가능 매체"는 신호 그 자체는 배제한다.

[0129] I/O 컴포넌트(550)는 입력을 수신하고, 출력을 제공하고, 출력을 만들어내고, 정보를 전송하고, 정보를 교환하고, 측도를 캡처하는 등의 여러 다양한 컴포넌트를 포함할 수 있다. 특정한 머신에 포함된 특정 I/O 컴포넌트(550)는 머신의 타입에 좌우될 것이다. 예를 들어, 모바일 전화와 같은 휴대용(portable) 머신은 터치 입력 장치 또는 다른 이러한 입력 메커니즘을 포함할 가능성이 있고, 헤드리스 서버 머신(headless server machine)은 이러한 터치 입력 장치를 포함하지 않을 가능성이 있다. I/O 컴포넌트(550)는 도 5에 도시되지 않은 많은 다른 컴포넌트를 포함할 수 있음이 이해될 것이다. I/O 컴포넌트(550)는 단지 이하의 설명을 단순화하기 위해 기능에 따라 그룹화되며 그룹화는 결코 한정적이지 않다. 다양한 예시적 실시예에서, I/O 컴포넌트(550)는 출력 컴포넌트(552) 및 입력 컴포넌트(554)를 포함할 수 있다. 출력 컴포넌트(552)는 시각(visual) 컴포넌트(예를 들어, PDP(plasma display panel), LED(light emitting diode) 디스플레이, LCD(liquid crystal display), 프로젝터(projector), 또는 CRT(cathode ray tube)와 같은 디스플레이), 음향(acoustic) 컴포넌트(예를 들어, 스피커), 촉각(haptic) 컴포넌트(예를 들어, 진동 모터, 저항(resistance) 메커니즘), 다른 신호 생성기 등을 포함할 수 있다. 입력 컴포넌트(554)는 문자숫자(alphanumeric) 입력 컴포넌트(예를 들어, 키보드, 문자숫자 입력을 수신하도록 구성된 터치 스크린, 광전(photo-optical) 키보드, 또는 다른 문자숫자 입력 컴포넌트), 포인트 기반(point based) 입력 컴포넌트(예를 들어, 마우스, 터치패드, 트랙볼, 조이스틱, 모션 센서, 또는 다른 포인팅 기구), 감촉성(tactile) 입력 컴포넌트(예를 들어, 물리적 버튼, 터치 또는 터치 제스처의 위치 및/또는 힘을 제공하는 터치 스크린, 또는 다른 감촉성 입력 컴포넌트), 오디오(audio) 입력 컴포넌트(예를 들어, 마이크(microphone)) 등을 포함할 수 있다.

[0130] 다른 예시적인 실시예에서, I/O 컴포넌트(550)는 광범위한 다른 컴포넌트 중에서도 생체측정(biometric) 컴포넌트(556), 모션(motion) 컴포넌트(558), 환경적(environmental) 컴포넌트(560), 또는 포지션(position) 컴포넌트(562)를 포함할 수 있다. 예를 들어, 생체측정 컴포넌트(556)는 표현(예를 들어, 손 표현, 안면 표정, 목소리 표현, 신체 제스처, 또는 눈 추적(eye tracking))을 검출하는 컴포넌트, 생체신호(예를 들어, 혈압, 심박동수, 체온, 발한, 또는 뇌파)를 측정하는 컴포넌트, 운동 관련 메트릭(예를 들어, 이동된 거리, 움직임의 빠르기, 또는 운동하는 데에 쓰인 시간)을 측정하는 컴포넌트, 사람을 식별하는(예를 들어, 음성 식별, 망막 식별, 안면 식별, 지문 식별, 또는 뇌전도 기반 식별) 컴포넌트 등을 포함할 수 있다. 모션 컴포넌트(558)는 가

속도 센서 컴포넌트(예를 들어, 가속도계(accelerometer)), 중력 센서 컴포넌트, 회전 센서 컴포넌트(예를 들어, 자이로스코프(gyroscope)) 등을 포함할 수 있다. 환경적 컴포넌트(560)는, 예를 들어 조명 센서 컴포넌트(예를 들어, 광도계(photometer)), 온도 센서 컴포넌트(예를 들어, 주변 온도를 검출하는 하나 이상의 온도계(thermometer)), 습도 센서 컴포넌트, 압력 센서 컴포넌트(예를 들어, 기압계(barometer)), 음향 센서 컴포넌트(예를 들어, 배경 잡음을 검출하는 하나 이상의 마이크), 근접성 센서(proximity sensor) 컴포넌트(예를 들어, 인근의 객체를 검출하는 적외선 센서), 가스 센서(예를 들어, 안전을 위해 위험한 가스의 농도를 검출하거나 또는 대기 내의 오염물을 측정하는 가스 검출 센서), 또는 주위의 물리적 환경에 대응하는 표시(indication), 측정 또는 신호를 제공할 수 있는 다른 컴포넌트를 포함할 수 있다. 포지션 컴포넌트(562)는 위치 센서 컴포넌트(예를 들어, GPS(Global Position System) 수신기 컴포넌트), 고도 센서 컴포넌트(예를 들어, 고도계(altimeter), 또는 고도가 도출될 수 있는 기압을 검출하는 기압계), 배향(orientation) 센서 컴포넌트(예를 들어, 자력계(magnetometer)) 등을 포함할 수 있다.

[0131] 여러 다양한 기술을 사용하여 통신이 구현될 수 있다. I/O 컴포넌트(550)는 각각 커플링(coupling)(582) 및 커플링(572)을 통해 머신(500)을 네트워크(580) 또는 장치(570)에 커플링하도록 동작 가능한 통신 컴포넌트(564)를 포함할 수 있다. 예를 들어, 통신 컴포넌트(564)는 네트워크(580)와 인터페이스하는 네트워크 인터페이스 컴포넌트 또는 다른 적합한 장치를 포함할 수 있다. 다른 예에서, 통신 컴포넌트(564)는 유선 통신 컴포넌트, 무선 통신 컴포넌트, 셀룰러 통신 컴포넌트, NFC(Near Field Communication) 컴포넌트, 블루투스(Bluetooth®) 컴포넌트(예를 들어, 블루투스 저에너지(Bluetooth®Low Energy)), 와이파이(Wi-Fi®) 컴포넌트, 그리고 다른 양태를 통해 통신을 제공하는 다른 통신 컴포넌트를 포함할 수 있다. 장치(570)는 다른 머신이거나 또는 여러 다양한 주변 장치(예를 들어, USB를 통해 커플링된 주변 장치) 중 임의의 것일 수 있다.

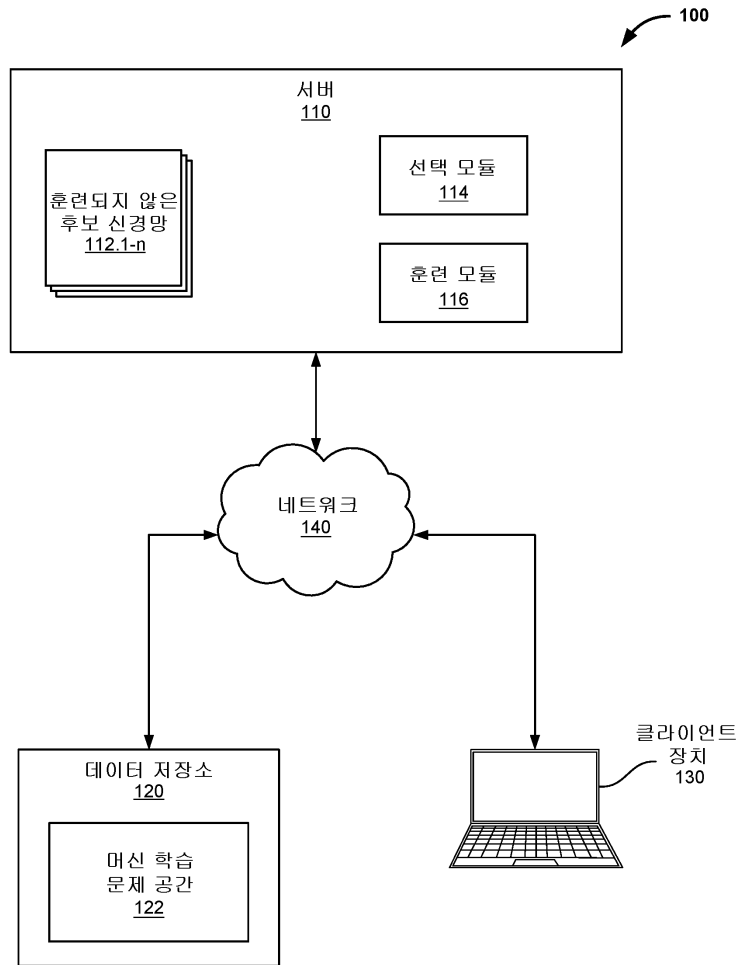
[0132] 더욱이, 통신 컴포넌트(564)는 식별자를 검출하거나 또는 식별자를 검출하도록 동작 가능한 컴포넌트를 포함할 수 있다. 예를 들어, 통신 컴포넌트(564)는 무선 RFID(Radio Frequency Identification) 태그 판독기(tag reader) 컴포넌트, NFC 스마트 태그 검출 컴포넌트, 광학 판독기(optical reader) 컴포넌트, 또는 음향 검출 컴포넌트(예를 들어, 태깅된(tagged) 오디오 신호를 식별하는 마이크)를 포함할 수 있다. 또한, 통신 컴포넌트(564)를 통해 IP(Internet Protocol) 지리위치(geolocation)를 통한 위치, 와이파이(Wi-Fi®) 신호 삼각 측량(triangulation)을 통한 위치, 특정한 위치를 나타낼 수 있는 NFC 비콘(beacon) 신호를 검출하는 것을 통한 위치 등과 같은 다양한 정보가 도출될 수 있다.

[0133] 다양한 예시적인 실시예에서, 네트워크(580)의 하나 이상의 부분은 애드혹 네트워크(ad hoc network), 인트라넷(intranet), 엑스트라넷(extranet), VPN(virtual private network), LAN(local area network), WLAN(Wireless LAN), WAN, WWAN(Wireless WAN), MAN(metropolitan area network), 인터넷, 인터넷의 일부, PSTN(Public Switched Telephone Network)의 일부분, POTS(plain old telephone service) 네트워크, 셀룰러 전화 네트워크, 무선 네트워크, 와이파이(Wi-Fi®) 네트워크, 다른 타입의 네트워크, 또는 둘 이상의 이러한 네트워크의 조합일 수 있다. 예를 들어, 네트워크(580) 또는 네트워크(580)의 일부는 무선 또는 셀룰러 네트워크를 포함할 수 있고, 커플링(582)은 CDMA(Code Division Multiple Access) 연결, GSM(Global System for Mobile communications), 또는 다른 타입의 셀룰러 또는 무선 커플링을 포함할 수 있다. 이 예에서, 커플링(582)은 1xRTT(Single Carrier Radio Transmission Technology), EVDO(Evolution-Data Optimized) 기술, GPRS(General Packet Radio Service) 기술, EDGE(Enhanced Data rates for GSM Evolution) 기술, 5G를 포함하는 3GPP(third Generation Partnership Project), 4세대 무선(4G) 네트워크, UMTS(Universal Mobile Telecommunications System), HSPA(High Speed Packet Access), WiMAX(Worldwide Interoperability for Microwave Access), LTE(Long Term Evolution) 표준, 다양한 표준 설정 기구에 의해 정의된 다른 것, 다른 장거리(long range) 프로토콜, 또는 다른 데이터 전송 기술과 같은, 다양한 타입의 데이터 전송 기술 중 임의의 것을 구현할 수 있다.

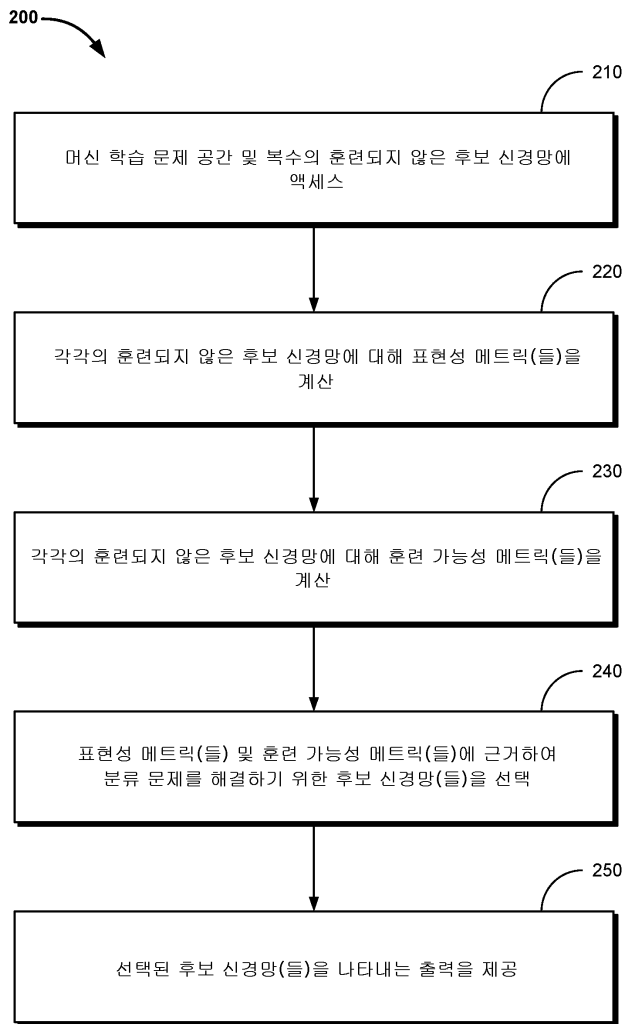
[0134] 명령어(516)는 네트워크 인터페이스 장치(예를 들어, 통신 컴포넌트(564)에 포함된 네트워크 인터페이스 컴포넌트)를 통해 전송 매체를 사용하여, 그리고 복수의 잘 알려진 전송 프로토콜(예를 들어, HTTP) 중 임의의 것을 활용하여 네트워크(580)를 통하여 전송되거나 또는 수신될 수 있다. 유사하게, 명령어(516)는 장치(570)와의 커플링(572)(예를 들어, 피어 대 피어 커플링)을 통해 전송 매체를 사용하여 전송되거나 또는 수신될 수 있다. 용어 "전송 매체"는 머신(500)에 의한 실행을 위한 명령어(516)를 저장, 인코딩(encoding) 또는 전달하는 것이 가능하고, 그러한 소프트웨어의 통신을 가능하게 하는 디지털 또는 아날로그 통신 신호 또는 다른 무형적 매체를 포함하는, 임의의 무형적(intangible) 매체를 포함하는 것으로 이해될 것이다.

도면

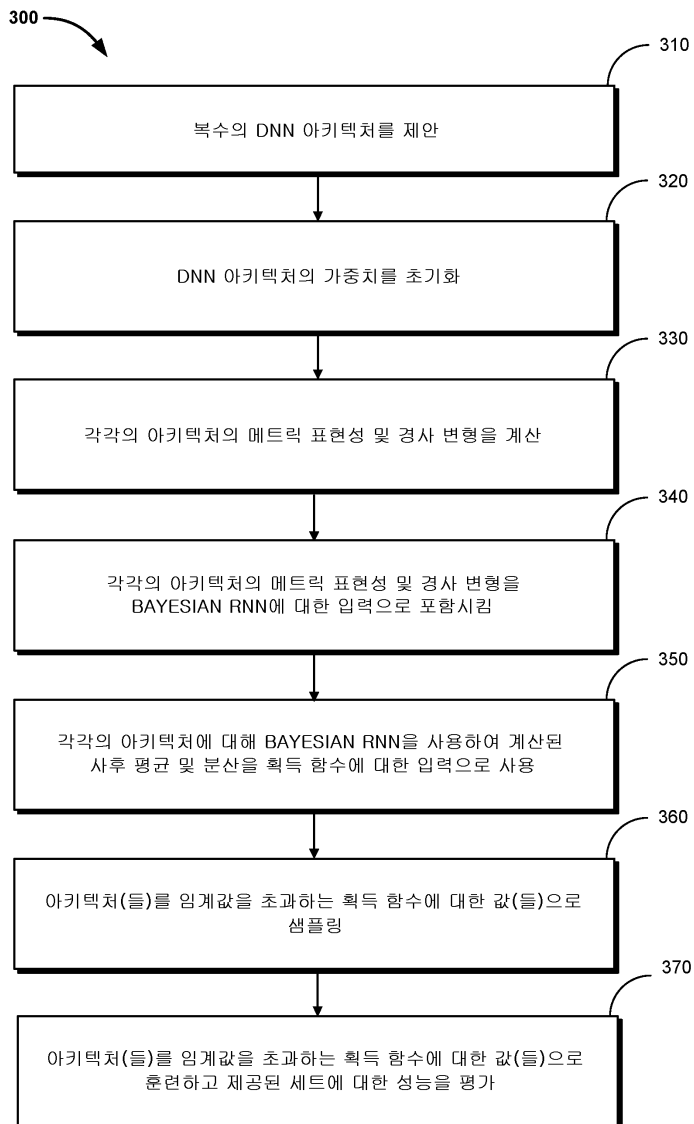
도면1



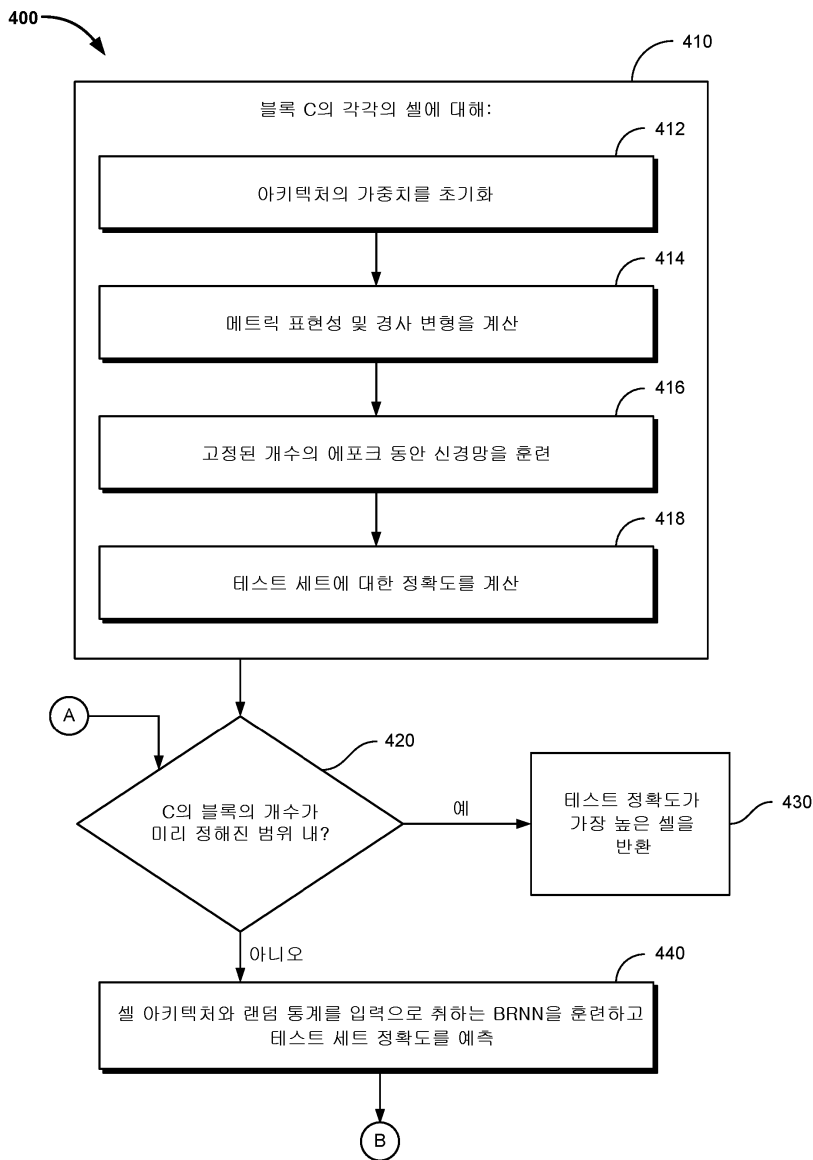
도면2



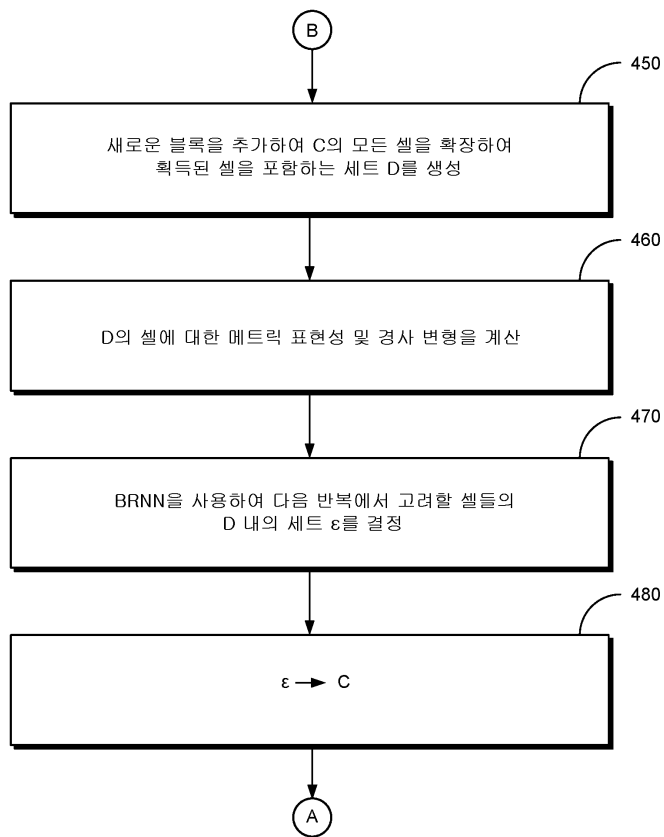
도면3



도면4a



도면4b



도면5

