



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년02월16일
(11) 등록번호 10-2215978
(24) 등록일자 2021년02월08일

(51) 국제특허분류(Int. Cl.)
G06N 20/20 (2019.01) G06N 5/04 (2006.01)
H04L 29/08 (2006.01)
(52) CPC특허분류
G06N 20/20 (2019.01)
G06N 5/04 (2013.01)
(21) 출원번호 10-2019-0114346
(22) 출원일자 2019년09월17일
심사청구일자 2019년09월17일
(56) 선행기술조사문헌
KR101828503 B1*
KR1020190032433 A*
US20150019214 A1
JP2017068710 A
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
주식회사 라인웍스
서울특별시 강남구 테헤란로14길 5, 8층(역삼동)
(72) 발명자
조용현
서울특별시 서초구 서초대로 13, 1동 1404호(방배동, 방배마에스트로아파트)
홍원준
경기도 고양시 덕양구 용현로 10, 502동 1203호(행신동, 무원마을5단지아파트)
(74) 대리인
특허법인스마트

전체 청구항 수 : 총 6 항

심사관 : 박승철

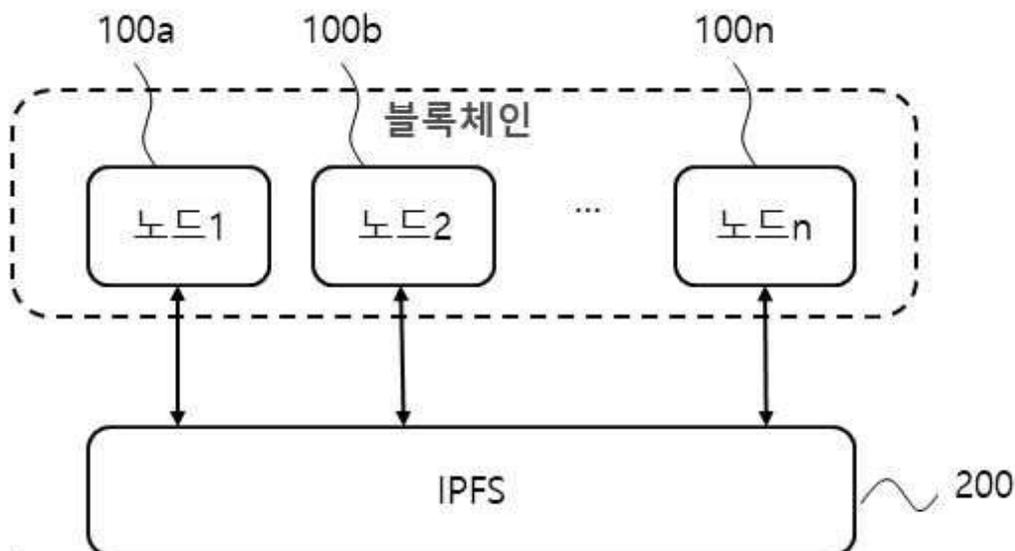
(54) 발명의 명칭 **블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템 및 그 방법**

(57) 요약

본 발명은 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템에 관한 것이다.

보다 구체적으로, 원장에 기록된 작업수행내용 및 작업수행상태를 모니터링하며, 미수행된 작업을 수행하여 특정 목적변수에 대한 확률을 출력하는 학습모델을 생성하는 복수의 작업수행노드 및 상기 원장에 작업수행내용을 기록하여 작업수행을 요청하고, 상기 복수의 작업수행노드에 의해 생성된 학습모델의 확률간의 가중치를 앙상블 학습하여 상기 특정 목적변수에 대한 결과를 추론하는 작업요청노드를 포함하는 것을 특징으로 하는 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템에 관한 것이다.

대표도 - 도1



(52) CPC특허분류
H04L 67/1097 (2013.01)

명세서

청구범위

청구항 1

원장에 기록된 작업수행내용 및 작업수행상태를 모니터링하며, 서로 다른 미수행된 작업을 수행하여 특정 목적 변수에 대한 확률을 출력하는 학습모델을 생성하는 복수의 작업수행노드; 및

상기 원장에 작업수행내용을 기록하여 작업수행을 요청하고, 상기 복수의 작업수행노드에 의해 생성된 학습모델의 확률 간의 가중치를 앙상블학습하여 상기 특정 목적변수에 대한 결과를 추론하는 작업요청노드;를 포함하고, 상기 작업요청노드는,

블록체인 노드의 자원을 공유하는 분산형파일시스템(IPFS:InterPlanetary File System)에 학습데이터 및 학습 알고리즘을 저장하고,

목적변수, 학습데이터 위치, 학습모델 생성 개수(m) 및 알고리즘 위치를 포함하는 작업수행내용을 블록체인의 상기 원장에 기록하여 작업을 요청하는 것을 특징으로 하는 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템.

청구항 2

삭제

청구항 3

제1항에 있어서,

상기 복수의 작업수행노드는,

원장을 실시간 모니터링하여 미수행작업이 있으면 작업수행상태를 상기 원장에 실시간 기록하며 학습을 수행하고, 생성된 학습모델을 상기 IPFS에 저장하고, 저장위치를 상기 원장에 기록하는 것을 특징으로 하는 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템.

청구항 4

제1항에 있어서,

상기 복수의 작업수행노드는, 서로 다른 미수행 작업에 대해 병렬적으로 작업을 수행하되, 하나의 작업수행노드는 특정 미수행 작업 수행 후 다른 미수행 작업을 수행하는 것을 특징으로 하는 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템.

청구항 5

제1항에 있어서,

상기 작업요청노드는, 앙상블모델을 통해 상기 학습모델의 확률간의 가중치를 적용하여 상기 결과를 추론하되, 학습 시작시점에는 기준 가중치를 적용하고, 완료 학습의 업데이트시에는 이전학습모델의 가중치를 적용하여 추론하는 것을 특징으로 하는 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템.

청구항 6

블록체인 노드 중 하나의 작업요청노드가, 원장에 작업수행내용을 기록하여 작업수행을 요청하는 단계;

상기 블록체인 노드 중 복수의 작업수행노드가, 원장에 기록된 작업수행내용 및 작업수행상태를 모니터링하는 단계;

상기 복수의 작업수행노드가, 상기 모니터링을 통해 서로 다른 미수행된 작업을 수행하여 특정 목적변수에 대한 확률을 출력하는 학습모델을 생성하는 단계;

상기 작업요청노드가, 상기 복수의 작업수행노드에 의해 생성된 학습모델의 확률 간의 가중치를 앙상블학습하여 상기 특정 목적변수에 대한 결과를 추론하는 단계를 포함하는 것을 특징으로 하는 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 방법.

청구항 7

하드웨어인 컴퓨터와 결합되어, 제6항의 방법을 컴퓨터에 의해 실행시키기 위하여 컴퓨터 관독 가능 기록매체에 저장된 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템에 관한 것이다.

[0002] 보다 구체적으로, 원장에 기록된 작업수행내용 및 작업수행상태를 모니터링하며, 미수행된 작업을 수행하여 특정 목적변수에 대한 확률을 출력하는 학습모델을 생성하는 복수의 작업수행노드 및 상기 원장에 작업수행내용을 기록하여 작업수행을 요청하고, 상기 복수의 작업수행노드에 의해 생성된 학습모델의 확률간의 가중치를 앙상블 학습하여 상기 특정 목적변수에 대한 결과를 추론하는 작업요청노드를 포함하는 것을 특징으로 하는 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템에 관한 것이다.

배경 기술

[0004] 기계 학습 또는 머신 러닝(machine learning)은 인공 지능의 한 분야로서 컴퓨터가 특정 문제에 대한 올바른 답을 도출할 수 있도록 학습 데이터를 통해 문제와 답 사이의 연관 관계를 스스로 학습하게 하는 기술을 의미한다.

[0006] 최근에는 동일한 학습 데이터에 여러 가지 기계 학습 알고리즘을 적용하여 생성한 다양한 예측 모델을 조합하여 목적변수 결과를 도출하는 앙상블 알고리즘에 대한 연구가 활발히 진행되고 있다.

[0008] 앙상블 알고리즘을 이용하여 목적변수 결과를 도출하는 선행문헌으로, 한국공개특허 제2018-0070103호가 있다. 상기 선행문헌은 앙상블모델에 기반하여 입력 데이터로부터 앙상블 모델에 대응하는 특징 벡터를 결정하고, 특징 벡터에 기초하여 입력 데이터를 인식하는 특징에 관한 것이다. 선행문헌은 앙상블 알고리즘을 이용함으로써 보다 향상된 데이터 인식을 도출할 수 있다.

[0010] 하지만, 선행문헌은 하나의 프로세서가 데이터 학습을 통해 다수의 예측모델을 생성하고, 생성된 예측모델들을 기반으로 앙상블 학습을 수행하며, 추론을 통해 데이터 인식 결과를 도출한다.

[0011] 즉, 모든 과정이 하나의 프로세서에서 순차 진행됨으로써 목적변수 결과 도출시까지 시간이 많이 소요되고, 자원 활용 측면에서도 비효율적이다. 또한, 동기성/직렬성 수행방법에 의해 락킹(locking)이 발생할 수밖에 없다.

[0013] 이에, 본 출원인은 블록체인을 기반으로 공유된 자원을 활용하여 비동기 분산 병렬적으로 앙상블 학습 및 추론을 수행할 수 있는 시스템 및 방법을 제안하고자 한다.

선행기술문헌

특허문헌

[0015] (특허문헌 0001) 1. 한국공개특허 제2018-0070103호(2018.06.26. 공개)

발명의 내용

해결하려는 과제

[0016] 본 발명의 목적은, 원장에 기록된 작업을 모니터링하는 블록체인 노드들을 통해 복수의 학습알고리즘을 비동기 분산 병렬적으로 학습하여 학습모델을 생성하고, 이를 앙상블 학습 및 추론함으로써 목적변수 결과를 출력하는 블록체인망 상 비동기 분산 병렬형 앙상블모델 학습 및 추론 시스템 및 그 방법을 제공하는 데 있다.

과제의 해결 수단

[0018] 상기한 바와 같은 목적을 달성하기 위한 본 발명의 일 실시 예에 따른 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템은 원장에 기록된 작업수행내용 및 작업수행상태를 모니터링하며, 서로 다른 미수행된 작업을 수행하여 특정 목적변수에 대한 확률을 출력하는 학습모델을 생성하는 복수의 작업수행노드, 상기 원장에 작업수행내용을 기록하여 작업수행을 요청하고, 상기 복수의 작업수행노드에 의해 생성된 학습모델의 확률 간의 가중치를 앙상블학습하여 상기 특정 목적변수에 대한 결과를 추론하는 작업요청노드를 포함할 수 있다.

[0020] 또한, 상기 작업요청노드는, 블록체인 멤버십 노드의 자원을 공유하는 분산형파일시스템(IPFS: InterPlanetary File System)에 학습데이터 및 학습 알고리즘을 저장하고, 목적변수, 학습데이터 위치, 학습모델 생성 개수(m) 및 알고리즘 위치를 포함하는 작업수행내용을 블록체인의 상기 원장에 기록하여 작업을 요청할 수 있다.

[0022] 또한, 상기 복수의 작업수행노드는, 원장을 실시간 모니터링하여 미수행작업이 있으면 작업수행상태를 상기 원장에 실시간 기록하며 학습을 수행하고, 생성된 학습모델을 상기 IPFS에 저장하고, 저장위치를 상기 원장에 기록할 수 있다.

[0024] 또한, 상기 복수의 작업수행노드는, 서로 다른 미수행 작업에 대해 병렬적으로 작업을 수행하되, 하나의 작업수행노드는 특정 미수행 작업 수행 후 다른 미수행 작업을 수행할 수 있다.

[0026] 또한, 상기 작업요청노드는, 앙상블모델을 통해 상기 학습모델의 확률간의 가중치를 적용하여 상기 결과를 추론하되, 학습 시작시점에는 기준 가중치를 적용하고, 완료 학습의 업데이트시에는 이전학습모델의 가중치를 적용하여 추론할 수 있다.

[0028] 한편, 본 발명의 일 실시 예에 따르면, 하는 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 방법은 블록체인 노드 중 하나의 작업수행노드가, 원장에 작업수행내용을 기록하여 작업수행을 요청하는 단계, 상기 블록체인 노드 중 복수의 작업수행노드가, 원장에 기록된 작업수행내용 및 작업수행상태를 모니터링하는 단계, 상기 복수의 작업수행노드가, 상기 모니터링을 통해 서로 다른 미수행된 작업을 수행하여 특정 목적변수에 대한 확률을 출력하는 학습모델을 생성하는 단계, 상기 작업수행노드가, 상기 복수의 작업수행노드에 의해 생성된 학습모델의 확률 간의 가중치를 앙상블학습하여 상기 특정 목적변수에 대한 결과를 추론하는 단계를 포함할 수 있다.

발명의 효과

[0030] 이상에서 설명한 바와 같이, 본 발명의 블록체인망 상 비동기 분산 병렬형 앙상블모델 학습 및 추론 시스템 및 방법은 블록체인 노드들을 이용하여 비동기 분산 병렬적으로 앙상블 학습을 수행 및 추론함으로써 정확한 목적변수 결과를 도출할 수 있다.

[0032] 또한, 블록체인의 공동원장을 통해 작업수행내용 및 작업상태를 모니터링하여 작업을 수행함으로써 복수의 노드가 비동기 분산 병렬적으로 학습을 수행할 수 있고, 다양한 학습알고리즘을 이용한 학습을 통해 복수의 학습모델을 생성하고, 생성된 학습모델의 확률간 가중치를 추출할 수 있다.

[0034] 또한, 앙상블 학습 중에도 추론을 수행할 수 있도록 가중치 적용을 설정함으로써 락킹(locking)을 방지할 수 있다.

[0036] 또한, 작업요청노드는 블록체인 노드들의 자원공유를 통해 학습 및 추론을 수행함으로써 자원을 효율적으로 이용할 수 있다.

도면의 간단한 설명

- [0038] 도 1은 본 발명의 일 실시 예에 따른 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템의 개략적인 구성을 나타내는 도면이다.
- 도 2는 도 1의 노드의 구성을 나타내는 구성도이다.
- 도 3은 작업요청모드의 노드 구성을 나타내는 구성도이다.
- 도 4는 작업수행모드의 노드구성을 나타내는 구성도이다.
- 도 5는 본 발명의 일 실시 예에 따른 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템을 설명하기 위한 도면이다.
- 도 6은 본 발명의 일 실시 예에 따른 시점별 작업수행을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0039] 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이거나 사전적인 의미로 한정해서 해석되어서는 안 되며, 발명자는 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다.
- [0041] 따라서 본 명세서에 기재된 실시예와 도면에 도시된 구성은 본 발명의 가장 바람직한 일 실시예에 불과할 뿐이고 본 발명의 기술적 사상을 모두 대변하는 것은 아니므로, 본 출원시점에 있어서 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있음을 이해하여야 한다.
- [0043] 이하, 도면을 참조하여 설명하기에 앞서, 본 발명의 요지를 드러내기 위해서 필요하지 않은 사항 즉 통상의 지식을 가진 당업자가 자명하게 부가할 수 있는 공지 구성에 대해서는 도시하지 않거나, 구체적으로 기술하지 않았음을 밝혀둔다.
- [0045] 도 1은 본 발명의 일 실시 예에 따른 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템(이하, 학습 및 추론 시스템)의 개략적인 구성을 나타내는 도면이다. 도 1을 참고하면, 본 발명의 학습 및 추론 시스템은 블록체인 기반의 복수의 노드(100a~100n)로 구성되며, 복수의 노드들은 인증된 멤버십 노드들이 될 수 있다.
- [0047] 이때, 복수의 노드들(100a~100n)은 자원을 공유하여 IPFS(200)를 형성할 수 있으며, 복수의 노드들(100a~100n) 중 작업을 요청하는 노드는 작업요청노드(일례로, 100a)가 되고, 그 외에 작업을 수행하는 노드들(100b~100n)은 작업수행노드가 될 수 있다. 여기서, 작업은 목적변수를 예측하기 위한 학습수행에 따른 학습모델을 생성하는 것이 될 수 있다.
- [0049] 여기서, IPFS(200)는 블록체인 멤버십 노드의 자원을 공유하는 분산형과일시스템으로, 작업내용에 따른 학습데이터, 학습시키려는 모델 알고리즘과 파라미터, 앙상블 모델 알고리즘 등이 작업요청노드(일례로, 100a)에 의해 저장될 수 있다.
- [0051] 또한, IPFS(200)는 작업수행에 따라 학습된 학습데이터, 학습에 따라 생성된 학습모델 등이 작업수행노드(일례로, 100b~100n)에 의해 저장될 수 있다. 이때, 저장되는 모든 정보들은 저장위치를 가지며, ID를 생성하여 식별될 수 있다.
- [0053] 도 2는 도 1의 노드의 구성을 나타내는 구성도이다. 도 2를 참고하면, 본 발명의 일 실시 예에 따른 노드(100a~100n)들은 원장(110), 모니터링부(120), 작업요청부(130), 학습모델 생성부(140), 앙상블학습부(150) 및 추론부(160)를 포함할 수 있다.
- [0055] 이때, 노드(100a~100n)의 작업수행 또는 작업요청에 따라 작업수행모드 및 작업요청모드로 구분되어 동작될 수 있으며, 작업요청모드로 동작하는 노드는 도 3과 같이 작업요청노드가 될 수 있고, 작업수행모드로 동작하는 노드는 도 4와 같이 작업수행노드가 될 수 있다. 설명의 이해를 위해, 도 3과 도 4를 통해 작업요청노드(100a)와 작업수행노드(100b~100n)의 동작을 구분하여 설명할 수 있다.
- [0057] 도 3은 작업요청모드의 노드(이하, 작업요청노드) 구성을 나타내는 구성도이다. 본 발명의 일 실시 예에 따른 작업요청노드는 원장(111), 모니터링부(121), 작업요청부(130), 앙상블학습부(150) 및 추론부(160)를 포함할 수

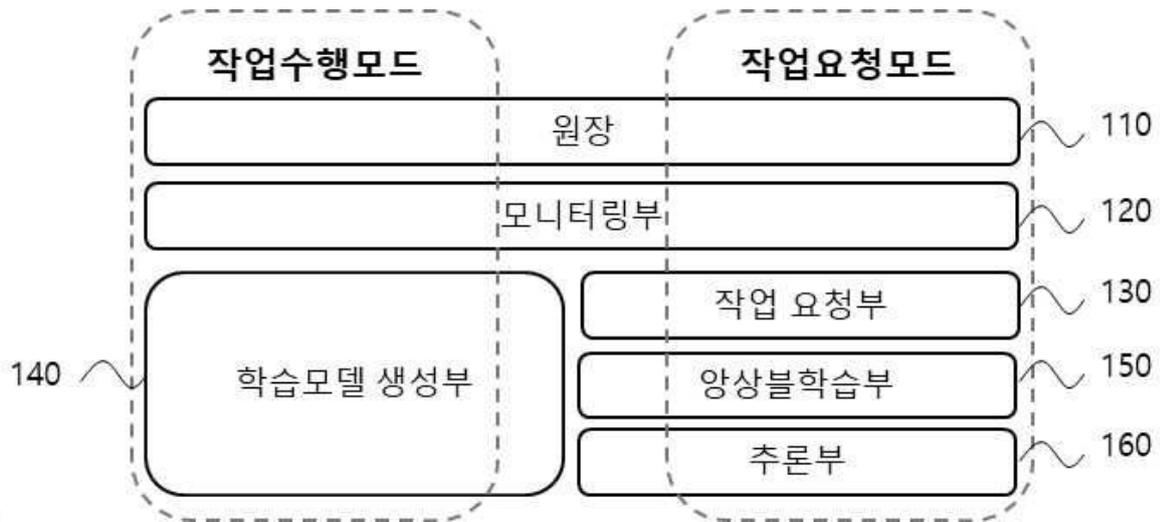
있다.

- [0059] 원장(111)은 블록체인 원장으로 작업요청노드(100a)의 작업요청부(130)에 따른 작업요청내용이 기록될 수 있다. 또한, 원장(111)에는 작업수행노드에 의한 작업상태, 작업에 따른 학습모델의 위치 등이 기록될 수 있다.
- [0060] 여기서, 작업수행내용은 목적변수, 학습데이터 위치, 예측을 위한 학습모델 생성 개수(n), 복수 개의 학습 알고리즘 위치 등이 될 수 있다.
- [0062] 즉, 작업은 목적변수 예측을 위해 IPFS(200)에 저장된 학습데이터를 작업요청노드(100a)가 요청한 학습 알고리즘별로 학습시켜 각각의 학습모델을 생성하는 것이 될 수 있다.
- [0064] 또한, 목적변수는 작업수행을 요청한 노드(100a)의 사용자 목적에 따라 결정되는 것으로, 일 예로 물체인식, 카드 이상사용 여부, 네트워크 로그 이상탐지 등이 될 수 있으며, 한정되지 않는다.
- [0065]
- [0066] 또한, 목적변수 예측을 위한 학습 알고리즘은 logistic regression, multi-layer perceptron, stochastic gradient descent, bagging, k-nearest neighbor, random forest, support vector machine 등이 될 수 있고, 이 중에서 작업요청노드(100a)에 의해 복수 개가 선택되어 작업을 요청할 수 있다.
- [0068] 모니터링부(121)는 원장(111)에 기록되는 작업요청내용 및 작업상태를 실시간으로 모니터링할 수 있다.
- [0069] 작업요청노드(100a)는 작업수행노드(100b~100n)들의 학습모델별 작업상태를 모니터링하여 완료된 학습모델을 확인할 수 있다. 또한, 완료된 학습모델을 이용하여 앙상블학습을 수행할 수 있다. 한편, 앙상블학습 및 이에 따른 추론은 작업수행노드(100b~100n)의 동작설명 후 설명하도록 한다.
- [0071] 도 4는 작업수행모드의 노드(이하, 작업수행노드) 구성을 나타내는 구성도이다. 본 발명의 일 실시 예에 따른 작업수행노드는 원장(112), 모니터링부(122) 및 학습모델 생성부(140)를 포함할 수 있다.
- [0072] 여기서, 원장(112)은 도 3의 원장(111)과 공통 분산원장이므로 원장(111)과 기록되는 내용은 동일하다.
- [0074] 모니터링부(122)는 원장(112)을 실시간모니터링하여 작업수행내용 중 미작업 상태인 학습이 있으면, 미작업 상태인 학습에 대해 원장(112)에 작업상태를 수행중으로 업데이트하고 해당 학습을 학습모델 생성부(140)에서 수행할 수 있게 한다.
- [0076] 일 예로, 다른 노드(100c)가 제1학습알고리즘을 이용하여 학습모델1을 작업 수행중이면, 노드(100b)는 미수행중인 제2학습알고리즘을 이용하여 학습모델2에 대한 작업을 수행할 수 있다. 이때, 노드(100b)는 학습모델2에 대한 작업을 수행중으로 원장(111)에 업데이트 하고 작업을 수행할 수 있다.
- [0078] 즉, 본 발명의 일 실시 예에 따른 작업수행노드(100b~100n)들은 작업수행내용 중 서로 다른 학습모델에 대한 작업을 동시에, 병렬적으로 수행할 수 있다. 또한, 작업수행노드(100b~100n)는 작업수행 중인 학습모델을 생성하여 작업을 완료하면, 다른 미수행 작업을 모니터링하여 작업을 수행할 수 있다.
- [0080] 학습모델 생성부(140)는 모니터링부(122)에서 미수행 작업에 대한 작업상태를 수행중으로 변경한 작업을 수행할 수 있다. 학습모델 생성부(140)는 작업요청내용에 포함된 학습데이터의 (IPFS)위치로 접근하여 해당 학습데이터에 대한 샘플셋을 생성하고, 대응되는 학습 알고리즘을 이용하여 샘플셋을 학습시킨 학습모델을 생성할 수 있다.
- [0081] 여기서, 생성된 학습모델은 작업요청노드(100a)가 설정한 특정 목적변수에 대한 확률을 출력할 수 있다.
- [0083] 도 5는 본 발명의 일 실시 예에 따른 블록체인망 상 비동기 분산 병렬형 앙상블 모델 학습 및 추론 시스템의 학습 및 추론을 설명하기 위한 도면이다. 도 5를 참고하면, 작업요청노드(100a)가 요청한 작업수행내용에 포함된 학습데이터(도5에서, 제1학습데이터)에 대하여, 제1작업수행노드의 학습모델 생성부(140)가 샘플셋1을 구성하고, 구성된 샘플셋1을 기설정된 제1학습 알고리즘을 이용하여 학습된 학습모델1을 생성할 수 있다.
- [0085] 또한, 제2작업수행노드가 학습모델 생성부(140)가 샘플셋2를 구성하고, 구성된 샘플셋2를 기설정된 제2학습 알고리즘을 이용하여 학습된 학습모델1을 생성할 수 있다.
- [0087] 또한, 또 다른 작업 미수행중인 작업수행노드 또는 학습모델1을 완료한 제1작업수행노드 또는 학습모델2를 완료한 제2작업수행노드 중 하나의 작업수행노드의 학습모델 생성부(140)가 샘플셋3을 구성하고, 구성된 샘플셋3을 기설정된 제3학습 알고리즘을 이용하여 학습된 학습모델3을 생성할 수 있다. 동일한 방법으로 학습모델m까

지 생성할 수 있으며, 이때 작업은 순차적인 처리가 아닌 병렬적으로 처리될 수 있다.

- [0089] 이때, 학습모델 생성부(140)는 학습이 완료된 학습모델을 IPFS(200)에 저장하고, 저장한 위치 및 학습모델에 대한 완료(작업상태)를 원장에 기록할 수 있다.
- [0091] 모니터링부(122)는 원장(121)이 업데이트될 때마다 해당 노드(자신)의 작업상태가 다른 작업을 수행할 수 있는 상태이면, 다른 미수행작업에 대한 작업상태를 수행중으로 업데이트하고 학습모델 생성부(140)에서 작업을 수행할 수 있게 한다.
- [0093] 작업요청노드의 모니터링부(121)는 작업요청내용 관련 원장(111)에 기록된 작업상태를 모니터링하며, 작업완료된 학습모델들의 (IPFS)위치를 파악하여 앙상블학습부(150)가 학습모델 생성부(140)에서 생성한 학습모델을 이용하여 앙상블학습을 수행할 수 있게 한다.
- [0095] 앙상블 학습부(150)는 복수의 작업수행노드(100b~100n)에 의해 생성된 학습모델들의 확률 간의 가중치를 학습시켜 앙상블모델을 생성할 수 있다.
- [0096] 이때, 앙상블 학습부(150)는 작업요청노드(100a)에 의해 설정된 학습데이터 또는 학습데이터와 동일한 목적변수에 대한 정보를 가지는 데이터(도 5에서, 제2학습데이터)를 작업요청노드들(100b~100n)에 의해 생성된 학습모델에 입력하고, 학습모델에서 출력된 확률들을 앙상블모델에 입력하여 확률간의 가중치를 학습할 수 있다.
- [0098] 추론부(160)는 앙상블 학습부(150)의 앙상블 학습 결과를 추론하여 앙상블모델의 출력으로 목적변수 결과를 얻을 수 있다.
- [0099] 일반적으로, 앙상블모델을 이용한 목적변수 도출시, 앙상블학습이 완료된 후 추론을 통해 목적변수를 얻을 수 있다. 이 경우, 락킹(locking ; 정체)이 발생된다. 이에, 본원발명은 학습 중에도 추론을 수행할 수 있도록 가중치를 설정할 수 있다. 즉, 학습과 추론은 병렬적으로 수행될 수 있다.
- [0101] 추론부(160)는 작업요청노드(100a)의 사용자(작업요청자)에 의해 추론이 요청되면, 추론이 요청된 시점을 기반으로 가중치를 부여하여 앙상블모델 추론값을 예측할 수 있다. 여기서, 추론값은 추론 요청 시점에서의 목적변수 결과가 될 수 있다.
- [0103] 추론부(160)는 앙상블 학습 초기시점에 추론이 요청되면, 확률간의 가중치가 학습되지 않은 상태이므로, 앙상블 학습에 이용되는 학습모델의 개수에 따라 기준값을 가중치로 적용할 수 있다.
- [0105] 도 6은 본 발명의 일 실시 예에 따른 시점별 작업수행을 설명하기 위한 도면이다. 일 예로, 도 6을 참고하면, 추론 요청 시점(t)이 t_0일 때, 원장의 작업상태가 학습모델1 및 학습모델2는 작업 수행중이고, 학습모델3은 미수행인 상태이면, 앙상블학습 초기시점이므로 학습모델에 가중치로 동일한 기준값(1/학습모델 개수)을 적용하여 추론값(E)을 출력할 수 있다.
- [0107] 또한, 추론부(160)는 학습모델의 작업상태가 완료되면, 학습모델의 확률간 가중치를 학습하고, 학습모델의 학습 완료시점에 추론이 요청되면(도 6의 t_1 시점) 학습된 가중치를 적용한 추론값(E)을 출력할 수 있다.
- [0109] 한편, 추론부(160)는 추론을 요청한 시점이 작업완료된 학습모델에 대해 업데이트가 발생한 시점이면(도 6의 t_3 시점), 해당 업데이트 대상 학습모델의 가중치로 이전 학습모델(업데이트 전 작업된 학습모델; 도 6에서 w_M_3_old)의 가중치를 적용한 추론값(E)을 출력할 수 있다. 여기서, 업데이트는 학습데이터의 추가 등이 될 수 있다.
- [0111] 추론부(160)는 시점별 학습모델의 작업상태에 따라 추론시 가중치를 적용함으로써 학습 수행 중에도 추론을 수행할 수 있고, 락킹 발생을 방지할 수 있다.
- [0113] 또한, 추론부(160)는 앙상블모델을 통해 출력된 추론값(E)과 임계치를 비교하여 클래스를 분류함으로써 목적변수 결과를 추론할 수 있다. 이에 따라, 하나의 학습모델을 이용한 목적변수 예측시 보다 정확한 목적변수 결과를 예측할 수 있다.
- [0115] 본 발명의 실시예들은, 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다.
- [0116] 상기 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다.

도면2



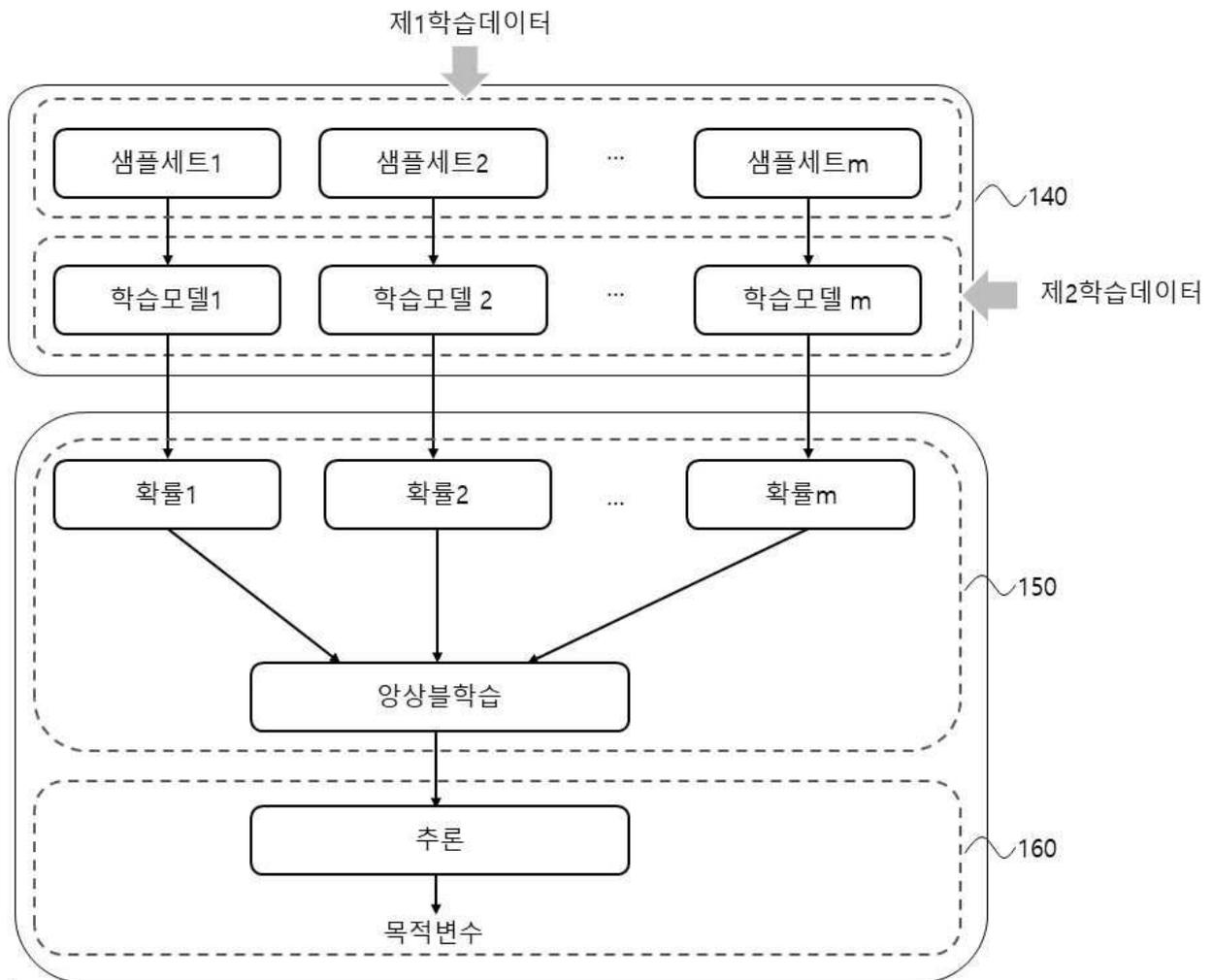
도면3



도면4



도면5



도면6

시점(t)	원장		시점(t)의 가중치	추론
t ₀	학습모델	작업상태	가중치	$E = 0.33 * M_1 + 0.33 * M_2 + 0.33 * M_3$
	학습모델1(M ₁)	수행중	기준값(0.33)	
	학습모델2(M ₂)	수행중	기준값(0.33)	
	학습모델3(M ₃)	미수행	기준값(0.33)	
t ₁	학습모델	작업상태	가중치	$E = w_{M_1} * M_1 + w_{M_2} * M_2 + w_{M_3} * M_3$
	학습모델1(M ₁)	완료	w _{M_1}	
	학습모델2(M ₂)	완료	w _{M_2}	
	학습모델3(M ₃)	완료	w _{M_3}	
t ₂	학습모델	작업상태	가중치	$E = w_{M_1} * M_1 + w_{M_2} * M_2 + w_{M_3_old} * M_3$
	학습모델1(M ₁)	완료	w _{M_1}	
	학습모델2(M ₂)	완료	w _{M_2}	
	학습모델3(M ₃)	업데이트	w _{M_3_old}	
	학습모델4(M ₄)	수행중	-	