(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0129966
(43) 공개일자 2024년08월28일

(51) 국제특허분류(Int. Cl.)
G06F 21/62 (2013.01) *G06F 40/279* (2020.01)
(52) CPC특허분류
G06F 21/6245 (2013.01)
G06F 40/279 (2020.01)
(21) 출원번호 10-2023-0023236
(22) 출원일자 2023년02월21일
심사청구일자 2023년02월21일

(71) 출원인
울산과학기술원
울산광역시 울주군 언양읍 유니스트길 50
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
(72) 발명자
이정혜
울산광역시 울주군 언양읍 유니스트길 50 (반연리, 울산과학기술원)
전치혁
경상북도 포항시 남구 청암로 77 (지곡동, 포항공과대학교)
(뒷면에 계속)
(74) 대리인
특허법인 수

전체 청구항 수 : 총 22 항

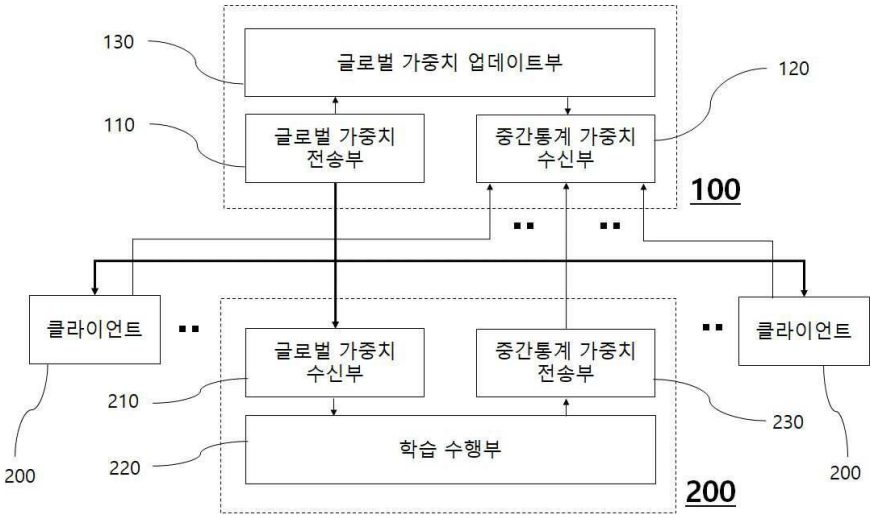
(54) 발명의 명칭 **Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법**

(57) 요 약

본 발명은 개인정보에 민감한 특징을 숨기는 데 중점을 둔 Word2Vec 기반의 효율적인 개인정보 보호 연합표현학습을 수행하고, 임베딩 업데이트 노출로 인해 개인정보에 민감한 특징이 직접적으로 유출되는 것을 방지하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법에 관한 것이다.

본 발명에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치는 Word2Vec 학습 모델의 학습을 수행하는 클라이언트, 및 클라이언트로 Word2Vec 학습 모델의 글로벌 가중치를 전송하고 클라이언트에서 학습된 중간통계 가중치를 수신받아 연합학습을 수행하는 서버로 이루어진다.

대 표 도 - 도4



(72) 발명자

김수현

울산광역시 울주군 언양읍 유니스트길 50 (반연리, 울산과학기술원)

이택호

경상북도 포항시 남구 청암로 77 (지곡동, 포항공과대학교)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711165204
과제번호	2020R1C1C1011063
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	딥러닝 기반의 개인정보보호가 보장되는 인공지능 연합학습 플랫폼 개발
기 여 율	25.00/100
과제수행기관명	울산과학기술원
연구기간	2022.03.01 ~ 2023.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1425170850
과제번호	P0016191
부처명	중소벤처기업부
과제관리(전문)기관명	한국산업기술진흥원
연구사업명	규제자유특구실증기반조성(R&D)
연구과제명	지능형 바이오 오믹스 분석기술 개발
기 여 율	25.00/100
과제수행기관명	울산과학기술원
연구기간	2022.01.01 ~ 2023.03.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1345354535
과제번호	2022R1I1A1A01073861
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	이공학학술연구기반구축
연구과제명	교차-기기 연합 학습 환경에서 개인정보보호가 보장되는 공유 표현 학습을 이용한 개별화 모형 개발
기 여 율	50.00/100
과제수행기관명	포항공과대학교
연구기간	2022.06.01 ~ 2023.02.28

명 세 서

청구범위

청구항 1

Word2Vec 학습 모델의 학습을 수행하는 클라이언트;

상기 클라이언트로 상기 Word2Vec 학습 모델의 글로벌 가중치를 전송하고 상기 클라이언트에서 학습된 중간통계 가중치를 수신받아 연합학습을 수행하는 서버;를 포함하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 2

제 1항에 있어서,

상기 서버는,

상기 클라이언트로 상기 Word2Vec 학습 모델의 상기 글로벌 가중치를 전송하는 글로벌 가중치 전송부;

적어도 두 개 이상의 상기 클라이언트로부터 상기 중간통계 가중치를 수신받는 중간통계 가중치 수신부; 및

상기 중간통계 가중치를 취합하여 상기 글로벌 가중치를 업데이트하는 글로벌 가중치 업데이트부;를 포함하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 3

제 2항에 있어서,

상기 취합은, 상기 가중 평균을 이용해 하나의 가중평균 중간통계 가중치를 생성하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 4

제 1항에 있어서,

상기 클라이언트는,

상기 서버로부터 Word2Vec 학습 모델의 상기 글로벌 가중치를 수신받는 글로벌 가중치 수신부;

상기 글로벌 가중치 정보를 토대로 Word2Vec의 학습을 수행하는 학습 수행부; 및

상기 Word2Vec의 상기 중간통계 가중치를 상기 서버로 전송하는 중간통계 가중치 전송부;를 포함하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 5

제 4항에 있어서,

상기 학습 수행부는, 상기 클라이언트 내의 로컬 데이터 통계를 토대로 학습을 수행하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 6

제 5항에 있어서,

상기 로컬 데이터 통계는 개인정보를 포함하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 7

제 4항에 있어서,

상기 학습 수행부는,

중심단어가 저장되어 있는 입력층;

상기 중심단어와 가까운 단어인 문맥단어가 저장되어 있는 출력층; 및

상기 입력층과 상기 출력층 간 단일층으로 이루어진 은닉층;을 포함하여 학습을 수행하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 8

제 7항에 있어서,

상기 문맥단어의 개수는 윈도우 사이즈로 나타내게 되는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 9

제 8항에 있어서,

상기 학습 수행부는, 상기 윈도우 사이즈를 벗어나는 가짜단어를 사용하여 학습을 수행하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 10

제 9항에 있어서,

상기 가짜단어는, 상기 문맥단어와 유사하게 업데이트되도록 선정하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 11

제 9항에 있어서,

상기 가짜단어는, 업데이트가 기준값 이하의 상기 중심단어에 대해 추가하여 학습을 수행하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치.

청구항 12

클라이언트에서 Word2Vec 학습 모델의 학습을 수행하는 Word2Vec 학습 모델 학습 수행단계;

상기 클라이언트로 상기 Word2Vec 학습 모델의 글로벌 가중치를 전송하고 상기 클라이언트에서 학습된 중간통계 가중치를 수신받아 서버에서 연합학습을 수행하는 연합학습 수행단계;를 포함하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 13

제 12항에 있어서,

상기 연합학습 수행단계는,

글로벌 가중치 전송부에서 상기 클라이언트로 상기 Word2Vec 학습 모델의 상기 글로벌 가중치를 전송하는 글로벌 가중치 전송단계;

중간통계 가중치 수신부에서 적어도 두 개 이상의 상기 클라이언트로부터 상기 중간통계 가중치를 수신받는 중간통계 가중치 수신단계; 및

글로벌 가중치 업데이트부에서 상기 중간통계 가중치를 취합하여 상기 글로벌 가중치를 업데이트하는 글로벌 가중치 업데이트단계;를 포함하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 14

제 13항에 있어서,

상기 취합은, 상기 가중 평균을 이용해 하나의 가중평균 중간통계 가중치를 생성하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 15

제 12항에 있어서,

상기 Word2Vec 학습 모델 학습 수행단계는,

글로벌 가중치 수신부에서 상기 서버로부터 상기 Word2Vec 학습 모델의 상기 글로벌 가중치를 수신받는 글로벌 가중치 수신단계;

학습 수행부에서 상기 글로벌 가중치 정보를 토대로 Word2Vec의 학습을 수행하는 학습 수행단계; 및

중간통계 가중치 전송부에서 상기 Word2Vec의 상기 중간통계 가중치를 상기 서버로 전송하는 중간통계 가중치 전송단계;를 포함하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 16

제 15항에 있어서,

상기 학습 수행단계에서는, 상기 클라이언트 내의 로컬 데이터 통계를 토대로 학습을 수행하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 17

제 16항에 있어서,

상기 로컬 데이터 통계는 개인정보를 포함하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 18

제 15항에 있어서,

상기 학습 수행단계는,

중심단어가 저장되어 있는 입력층;

상기 중심단어와 가까운 단어인 문맥단어가 저장되어 있는 출력층; 및

상기 입력층과 상기 출력층 간 단일층으로 이루어진 은닉층;을 포함하여 학습을 수행하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 19

제 18항에 있어서,

상기 문맥단어의 개수는 윈도우 사이즈로 나타내게 되는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 20

제 19항에 있어서,

상기 학습 수행단계는, 상기 윈도우 사이즈를 벗어나는 가짜단어를 사용하여 학습을 수행하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 21

제 20항에 있어서,

상기 가짜단어는, 상기 문맥단어와 유사하게 업데이트되도록 선정하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

청구항 22

제 20항에 있어서,

상기 가짜단어는, 업데이트가 기준값 이하의 상기 중심단어에 대해 추가하여 학습을 수행하는 것을 특징으로 하는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법에 관한 것으로, 상세하게는, 연합학습(federated learning)에서 유출 우려가 있는 개인정보의 보호를 강화하는 것이다.

[0002] 또한, 본 발명은 Word2Vec의 특징을 이용하여 개인정보와 같은 민감한 정보와 일반 정보를 구분하여 민감한 정보가 직접적으로 유출되는 것을 방지할 수 있는 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법에 관한 것이다.

배 경 기 술

[0003] 인터넷과 스마트 통신 기술의 발달로 인해 사용자는 손쉽게 상품에 대한 정보를 얻을 수 있게 되었다. 반면, 이러한 기술의 발달은 사용자에게 지나치게 많은 양의 정보를 제공함으로 인해 상품 선택에 장시간이 소요되는 문제를 발생시켰다. 이에 따라, 사용자에게 불필요한 정보를 필터링 하여 구매율이 높은 상품만을 제공하는 개인 맞춤형 기술 및 추천 시스템이 요구되고 있다.

[0004] 과거의 추천 시스템은 단순한 규칙 기반 알고리즘을 활용하여 사용자의 기호 정보를 고려한 상품을 추천하는 수준에 불과했다. 이후, 기술의 발달로 추천 시스템은 기계 학습(machine learning) 및 데이터 마이닝(data mining) 기법을 활용하여 사용자의 행동 패턴과 성향을 분석하게 되었으며, 이를 통해 추천 시스템은 사용자가 선호할 것으로 예상되는 정보 및 컨텐츠를 추천하게 되었다.

[0005] 현재, 이러한 추천 시스템을 구현하기 위해 다양한 기법들이 개발되고 있는데, 이때 고도로 분산된 환경에서 충분한 데이터를 수집/저장하기 위해 많은 통신/저장비용이 발생하고 있으므로 중앙 집중 환경에서 기계 학습을 활용한 추천 시스템 구현은 어려워지고 있다. 특히, 최근 들어 데이터 공유는 개인정보유출을 방지하기 위한 규제에 의해 제한되는 경우가 많아 개인정보의 유출을 보호하는 다양한 기법들이 연구되어 왔다.

[0006] 그 일례로, 대한민국 공개특허공보 제10-2015-7024126호에서는 행렬 분해 및 리지 회귀에 기초하여 평가 기여 사용자들 및 비-기여 사용자들에게 프라이버시-보호 및 블라인드 방식으로 추천들을 제공하는 방법을 제안하였다.

[0007] 그러나, 이 경우에도 개인정보를 보호하는 방법 중 행렬분해(Matrix Factorization; MF)와 같은 전통적인 알고리즘에 의존함으로 인해 추천 정확성이 감소하는 단점이 있다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 대한민국 공개특허공보 제10-2015-7024126호(2014.05.01)

발명의 내용

해결하려는 과제

[0009] 본 발명의 목적은, 개인정보에 민감한 특징을 숨기는 데 중점을 둔 Word2Vec 기반의 효율적인 개인정보 보호 연합표현학습을 수행하는, Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법을 제공하는 것이다.

[0010] 본 발명은 임베딩(embedding) 업데이트 노출로 인해 개인정보에 민감한 특징이 직접적으로 유출되는 것을 방지하는, Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법을 제공하는데 또 다른 목적이 있다.

과제의 해결 수단

[0011] 본 발명의 일 실시예에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치는 Word2Vec 학습 모델의 학습을 수행하는 클라이언트, 및 클라이언트로 Word2Vec 학습 모델의 글로벌 가중치를 전송하고 클라이언트에서 학습된 중간통계 가중치를 수신받아 연합학습을 수행하는 서버를 포함할 수 있다.

[0012] 여기서, 서버는 클라이언트로 Word2Vec 학습 모델의 글로벌 가중치를 전송하는 글로벌 가중치 전송부, 적어도 두 개 이상의 클라이언트로부터 중간통계 가중치를 수신받는 중간통계 가중치 수신부, 및 중간통계 가중치를 취합하여 글로벌 가중치를 업데이트하는 글로벌 가중치 업데이트부를 포함할 수 있다.

[0013] 또한, 취합은 가중 평균을 이용해 하나의 가중평균 중간통계 가중치를 생성할 수 있다.

[0014] 여기서, 클라이언트는 서버로부터 Word2Vec 학습 모델의 글로벌 가중치를 수신받는 글로벌 가중치 수신부, 글로벌 가중치 정보를 토대로 Word2Vec의 학습을 수행하는 학습 수행부, 및 Word2Vec의 중간통계 가중치를 서버로 전송하는 중간통계 가중치 전송부를 포함할 수 있다.

[0015] 또한, 학습 수행부는 클라이언트 내의 로컬 데이터 통계를 토대로 학습을 수행할 수 있다.

[0016] 여기서, 로컬 데이터 통계는 개인정보를 포함할 수 있다.

[0017] 또한, 학습 수행부는 중심단어가 저장되어 있는 입력층, 중심단어와 가까운 단어인 문맥단어가 저장되어 있는 출력층, 및 입력층과 출력층 간 단일층으로 이루어진 은닉층을 포함하여 학습을 수행할 수 있다.

[0018] 여기서, 문맥단어의 개수는 윈도우 사이즈로 나타내게 될 수 있다.

[0019] 또한, 학습 수행부는 윈도우 사이즈를 벗어나는 가짜단어를 사용하여 학습을 수행할 수 있다.

[0020] 여기서, 가짜단어는 문맥단어와 유사하게 업데이트되도록 선정할 수 있다.

[0021] 또한, 가짜단어는 업데이트가 기준값 이하의 중심단어에 대해 추가하여 학습을 수행할 수 있다.

[0022] 본 발명의 다른 실시예에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법은 클라이언트에서 Word2Vec 학습 모델의 학습을 수행하는 Word2Vec 학습 모델 학습 수행단계 및 클라이언트로 Word2Vec 학습 모델의 글로벌 가중치를 전송하고 클라이언트에서 학습된 중간통계 가중치를 수신받아 서버에서 연합학습을 수행하는 연합학습 수행단계를 포함할 수 있다.

[0023] 여기서, 연합학습 수행단계는 글로벌 가중치 전송부에서 클라이언트로 Word2Vec 학습 모델의 글로벌 가중치를 전송하는 글로벌 가중치 전송단계, 중간통계 가중치 수신부에서 적어도 두 개 이상의 클라이언트로부터 중간통계 가중치를 수신받는 중간통계 가중치 수신단계, 및 글로벌 가중치 업데이트부에서 중간통계 가중치를 취합하여 글로벌 가중치를 업데이트하는 글로벌 가중치 업데이트단계를 포함할 수 있다.

[0024] 또한, 취합은 가중 평균을 이용해 하나의 가중평균 중간통계 가중치를 생성할 수 있다.

[0025] 여기서, Word2Vec 학습 모델 학습 수행단계는 글로벌 가중치 수신부에서 서버로부터 Word2Vec 학습 모델의 글로벌 가중치를 수신받는 글로벌 가중치 수신단계, 학습 수행부에서 글로벌 가중치 정보를 토대로 Word2Vec의 학습을 수행하는 학습 수행단계, 및 중간통계 가중치 전송부에서 Word2Vec의 중간통계 가중치를 서버로 전송하는 중간통계 가중치 전송단계를 포함할 수 있다.

[0026] 또한, 학습 수행단계에서는 클라이언트 내의 로컬 데이터 통계를 토대로 학습을 수행할 수 있다.

[0027] 여기서, 로컬 데이터 통계는 개인정보를 포함할 수 있다.

[0028] 또한, 학습 수행단계는 중심단어가 저장되어 있는 입력층, 중심단어와 가까운 단어인 문맥단어가 저장되어 있는

출력층, 및 입력층과 출력층 간 단일층으로 이루어진 은닉층을 포함하여 학습을 수행할 수 있다.

[0029] 여기서, 문맥단어의 개수는 윈도우 사이즈로 나타내게 될 수 있다.

[0030] 또한, 학습 수행단계는 윈도우 사이즈를 벗어나는 가짜단어를 사용하여 학습을 수행할 수 있다.

[0031] 여기서, 가짜단어는 문맥단어와 유사하게 업데이트되도록 선정할 수 있다.

[0032] 또한, 가짜단어는 업데이트가 기준값 이하의 중심단어에 대해 추가하여 학습을 수행할 수 있다.

발명의 효과

[0033] 본 발명에 의한 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법은 개인정보에 민감한 특징을 숨기는 데 중점을 둔 Word2Vec 기반의 효율적인 개인정보 보호 연합 표현학습을 수행하는 장점이 있다.

[0034] 또한, 본 발명에 의한 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법은 임베딩(Embedding) 업데이트 노출로 인해 개인정보에 민감한 특징이 직접적으로 유출되는 것을 방지할 수 있는 장점이 있다.

도면의 간단한 설명

[0035] 도 1은 본 발명의 일 실시예에 따른 서버를 나타낸 구성도이다.

도 2는 본 발명의 다른 실시예에 따른 클라이언트를 나타낸 구성도이다.

도 3은 도 2의 학습 수행부를 상세히 나타낸 구성도이다.

도 4는 본 발명의 또다른 실시예에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치를 나타낸 구성도이다.

도 5는 본 발명의 일 실시예에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법을 나타낸 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0036] 이하, 본 발명의 실시를 위한 구체적인 실시예를 첨부된 도면들을 참조하여 설명한다.

[0037] 본 발명을 설명함에 있어서 제 1, 제 2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 구성요소들은 용어들에 의해 한정되지 않을 수 있다. 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제 1 구성요소는 제 2 구성요소로 명명될 수 있고, 유사하게 제 2 구성요소도 제 1 구성요소로 명명될 수 있다.

[0038] 어떤 구성요소가 다른 구성요소에 연결되어 있다거나 접속되어 있다고 언급되는 경우는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해될 수 있다.

[0039] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다.

[0040] 본 명세서에서, 포함하다 또는 구비하다 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것으로서, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해될 수 있다.

[0041] 또한, 도면에서의 요소들의 형상 및 크기 등은 보다 명확한 설명을 위해 과장될 수 있다.

[0042] 이하, 첨부된 도면을 참조하여 본 발명에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법에 대해 상세히 설명한다.

[0044] 도 1은 본 발명의 일 실시예에 따른 서버(100)를 나타낸 구성도이다.

[0045] 최근 연구들은 중간 통계만 주고 받아 모형을 학습하는 연합학습(Federated learning)을 채택하여 단일 글로벌 모델을 분산방식으로 훈련하는 프로토콜을 제공하고 있는데, 이는 스마트폰, 웨어러블기기 등 IoT(Internet of Things)의 발달로 인해 점점 더 많은 데이터가 고도로 분산된 환경에서 생성 및 저장되어 기기간 연합학습(Cross-device federated learning)이 가능하기 때문이다.

[0046] 한편, 정확한 협업 필터링 기반 추천시스템(Recommendation system) 구축을 위해 충분한 양의 사용자 데이터 수집이 불가피하며, 기존 추천시스템은 모든 사용자를 식별할 수 있고, 사용자의 기록이 중앙서버에 수집되어 있어 공유된 중간통계치(Intermediate statistics)를 통해서도 여전히 개인정보 유출의 여지가 남아있다.

[0047] 또한, 최근 연구에 따르면 공격자는 사용자와 서버간 전달되는 기울기와 같은 중간통계를 분석하여 학습 데이터를 재구성하거나 속성을 추론할 수 있으므로, 개인정보에 민감한 특징들은 중간 통계로부터 누출되지 않도록 숨겨야 할 필요가 있다.

[0048] 예를 들어, 드물게 평가된 영화일수록 해당 영화 평가 사실을 이용해 사용자를 특정하기 쉬운 것과 마찬가지인데, 기존의 연구를 보면, 개인정보를 보호하는 여러 추천방법들이 사용하는 차등개인정보(Differential privacy) 또는 암호화(Cryptography)기법들은 모형 학습의 효율성을 감소시키는 단점이 있다.

[0049] 또한, 개인정보를 보호하는 많은 방법들은 행렬분해(Matrix Factorization; MF)와 같은 전통적인 알고리즘에 의존하나, 최근 일반적인 추천시스템에서는 오히려 전통적 알고리즘과 달리 사용자 데이터가 발생한 순서정보를 활용하기 위해 잘 학습된 표현을 이용하여 높은 추천성능을 달성하고 있다.

[0050] 본 발명에 따른 서버(100)는 이러한 개인정보 유출을 방지하기 위한 구조를 제공하고 있다.

[0051] 도 1에서 알 수 있는 바와 같이, 본 발명의 서버(100)는 클라이언트(200)로 Word2Vec 학습 모델의 글로벌 가중치를 전송하는 글로벌 가중치 전송부(110), 적어도 두 개 이상의 클라이언트(200)로부터 중간통계 가중치를 수신받는 중간통계 가중치 수신부(120), 및 중간통계 가중치를 취합하여 글로벌 가중치를 업데이트하는 글로벌 가중치 업데이트부(130)로 이루어진다.

[0052] 즉, 본 발명은 Word2Vec 학습 모델의 구조를 이용하여 서버(100)와 클라이언트(200) 간 개인정보와 같은 민감한 정보가 노출되지 않도록 하고 있으며, 특히 취합은 다수의 클라이언트(200)에서 전송한 중간통계 가중치의 가중평균을 이용해 하나의 가중평균 중간통계 가중치를 생성함으로써, 기기간 연합학습을 수행한다.

[0054] 도 2는 본 발명의 다른 실시예에 따른 클라이언트(200)를 나타낸 구성도이며, 도 3은 도 2를 상세히 설명하기 위한 세부 구성도이다.

[0055] 먼저, 도 2를 참조하면, 본 발명의 다른 실시예에 따른 클라이언트(200)는 서버(100)로부터 Word2Vec 학습 모델의 글로벌 가중치를 수신받는 글로벌 가중치 수신부(210), 글로벌 가중치 정보를 토대로 Word2Vec의 학습을 수행하는 학습 수행부(220), 및 Word2Vec의 중간통계 가중치를 서버(100)로 전송하는 중간통계 가중치 전송부(230)로 이루어진다.

[0056] 즉, 본 발명은 Word2Vec 학습 모델의 구조를 이용하여 서버(100)와 클라이언트(200) 간 개인정보와 같은 민감한 정보가 노출되지 않도록 한다.

[0057] 이때, 학습 수행부(220)는 서버(100)에서 수신받은 글로벌 가중치를 이용해 클라이언트(200) 내의 로컬 데이터 통계를 토대로 학습을 수행한다. 여기서, 로컬 데이터 통계는 개인정보를 포함할 수 있어 이러한 민감한 정보가 노출되지 않도록 구성할 수 있다.

[0058] 본 발명에서 이러한 민감한 정보가 노출되지 않도록 하는 방법에 대해 이하 도 3에서 상세히 설명한다.

[0060] 도 3은 도 2의 학습 수행부(220)를 상세히 나타낸 구성도이다.

[0061] 도 3에서 알 수 있는 바와 같이, 본 발명의 학습 수행부(220)는 중심단어(W310)가 저장되어 있는 입력층(221), 중심단어(W310)와 가까운 단어인 문맥단어(W320)가 저장되어 있는 출력층(223), 및 입력층(221)과 출력층(223) 간 단일층으로 이루어진 은닉층(222)을 포함하여 학습을 수행한다.

[0062] 여기서, 문맥단어(W320)의 개수는 윈도우 사이즈로 나타내게 되는데, 예를 들어 윈도우 사이즈가 1일 경우 중심

단어(W310)와 가장 가까운 전,후 단어 1개 씩의 단어를 포함할 수 있다.

[0063] 한편, 윈도우 사이즈를 벗어나는 단어를 가짜단어(negative word)(W330)라 하며, 본 발명에서는 이를 혼합하여 학습을 수행한다. 본 발명에서는 이러한 가짜단어(W330)를 혼합함으로써 개인정보와 같은 민감한 정보가 가짜단어(W330)와 구분하기 어렵도록 반영할 수 있다.

[0064] 예를 들어, 'The fat cat sat on the bench.'라는 문장에서 'sat'을 중심단어(W310)라 할 때 윈도우 사이즈가 1일 경우 'cat'과 'on'은 문맥단어(W320)며, 이 외 단어는 가짜단어(W330)이다.

[0065] 즉, 자주 등장하지 않는 민감한 정보는 업데이트 비율이 낮으므로 업데이트가 낮은 중심단어(W310)에 대해서는 자주 등장할 수 있는 가짜단어(W330)를 선택하여 혼합하고, 이때 가짜단어(W330)는 문맥단어(W320)와 유사하게 업데이트되도록 선정할 수 있다.

[0066] 따라서, 민감한 정보인 중심단어(W310)와 다른 문맥단어(W320)의 업데이트 비율이 유사하게 됨으로써, 민감한 정보가 유출되지 않도록 할 수 있다. 즉, 본 발명에서는 업데이트가 기준값 이하인 중심단어(W310)에 대해 가짜단어(W330)를 추가하여 학습을 수행함으로써, 민감한 정보가 유출되지 않도록 할 수 있다.

[0068] 도 4는 본 발명의 또다른 실시예에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치를 나타낸 구성도이다.

[0069] 도 4에서 볼 수 있는 바와 같이, 본 발명의 또다른 실시예에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치는 Word2Vec 학습 모델의 학습을 수행하는 클라이언트(200) 및 클라이언트(200)로 Word2Vec 학습 모델의 글로벌 가중치를 전송하고 클라이언트(200)에서 학습된 중간통계 가중치를 수신받아 연합학습을 수행하는 서버(100)로 이루어진다.

[0070] 여기서, 서버(100)는 클라이언트(200)로 Word2Vec 학습 모델의 글로벌 가중치를 전송하는 글로벌 가중치 전송부(110), 적어도 두 개 이상의 클라이언트(200)로부터 중간통계 가중치를 수신받는 중간통계 가중치 수신부(120), 및 중간통계 가중치를 취합하여 글로벌 가중치를 업데이트하는 글로벌 가중치 업데이트부(130)로 이루어진다.

[0071] 즉, 본 발명의 서버(100)는 Word2Vec 학습 모델의 구조를 이용하여 서버(100)와 클라이언트(200) 간 개인정보와 같은 민감한 정보가 노출되지 않도록 하고, 특히, 취합은 다수의 클라이언트(200)에서 전송한 중간통계 가중치의 가중 평균을 이용해 하나의 가중평균 중간통계 가중치를 생성함으로써, 기기간 연합학습을 수행할 수 있다.

[0072] 또한, 클라이언트(200)는 서버(100)로부터 Word2Vec 학습 모델의 글로벌 가중치를 수신받는 글로벌 가중치 수신부(210), 글로벌 가중치 정보를 토대로 Word2Vec의 학습을 수행하는 학습 수행부(220), 및 Word2Vec의 중간통계 가중치를 서버(100)로 전송하는 중간통계 가중치 전송부(230)로 이루어진다.

[0073] 즉, 본 발명은 Word2Vec 학습 모델의 구조를 이용하여 서버(100)와 클라이언트(200) 간 개인정보와 같은 민감한 정보가 노출되지 않도록 한다.

[0074] 여기서, 학습 수행부(220)는 서버(100)에서 수신받은 글로벌 가중치를 이용해 클라이언트(200) 내의 로컬 데이터 통계를 토대로 학습을 수행한다. 이때, 로컬 데이터 통계는 개인정보를 포함할 수 있어, 이러한 민감한 정보가 노출되지 않도록 구성할 수 있다.

[0075] 즉, 학습 수행부(220)는 중심단어(W310)가 저장되어 있는 입력층(221), 중심단어(W310)와 가까운 단어인 문맥단어(W320)가 저장되어 있는 출력층(223), 및 입력층(221)과 출력층(223) 간 단일층으로 이루어진 은닉층(222)을 포함하여 학습을 수행한다.

[0076] 여기서, 문맥단어(W320)의 개수는 윈도우 사이즈로 나타내게 되는데, 예를 들어 윈도우 사이즈가 1일 경우 중심단어(W310)와 가장 가까운 전,후 단어 1개 씩의 단어를 포함할 수 있다.

[0077] 한편, 윈도우 사이즈를 벗어나는 단어를 가짜단어(negative word)(W330)라 하며, 이를 혼합하여 학습을 수행할 수 있다. 본 발명에서는 이러한 가짜단어(W330)를 혼합함으로써, 개인정보와 같은 민감한 정보는 가짜단어(W330)와 구분하기 어렵도록 반영할 수 있다.

[0078] 이러한 가짜단어(W330)를 혼합하여 학습을 수행함으로써 개인정보와 같은 민감한 정보가 가짜단어(W330)와 구분하기 어렵도록 반영하는 방법은, 상기 도 3 에서와 동일한 방법을 사용할 수 있으며 이에 대한 상세한 설명은 생략한다.

[0079] 이상과 같은 본 발명에 의한 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치는, 임베딩(Embedding) 업데이트 노출로 인해 개인정보에 민감한 특징이 직접적으로 유출되는 것을 방지하기 위해 가짜단어와 같은 가짜샘플(Negative samples)의 업데이트를 혼합하고 추론 공격을 방지하기 위해 드물게 나타나는 기능(개인정보에 민감한 특징)의 업데이트를 조정함으로써, 개인정보의 유출을 방지할 수 있다.

[0081] 도 5는 본 발명의 일 실시예에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법을 나타낸 순서도이다.

[0082] 도 5에서 알 수 있는 바와 같이, Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법은, 클라이언트(200)에서 Word2Vec 학습 모델의 학습을 수행하는 Word2Vec 학습 모델 학습 수행단계(S200) 및 클라이언트(200)로 Word2Vec 학습 모델의 글로벌 가중치를 전송하고 클라이언트(200)에서 학습된 중간통계 가중치를 수신받아 서버(100)에서 연합학습을 수행하는 연합학습 수행단계(S100)로 이루어진다.

[0083] 여기서, 연합학습 수행단계(S100)는 글로벌 가중치 전송부(110)에서 클라이언트(200)로 Word2Vec 학습 모델의 글로벌 가중치를 전송하는 글로벌 가중치 전송단계(S110), 중간통계 가중치 수신부(120)에서 적어도 두 개 이상의 클라이언트(200)로부터 중간통계 가중치를 수신받는 중간통계 가중치 수신단계(S120), 및 글로벌 가중치 업데이트부(130)에서 중간통계 가중치를 취합하여 글로벌 가중치를 업데이트하는 글로벌 가중치 업데이트단계(S130)를 포함한다.

[0084] 즉, 본 발명은 Word2Vec 학습 모델의 구조를 이용하여 서버(100)와 클라이언트(200) 간 개인정보와 같은 민감한 정보가 노출되지 않도록 하고, 특히 취합은 다수의 클라이언트(200)에서 전송한 중간통계 가중치의 가중 평균을 이용해 하나의 가중평균 중간통계 가중치를 생성함으로써, 기기간 연합학습을 수행할 수 있다.

[0085] 여기서, Word2Vec 학습 모델 학습 수행단계(S200)는 글로벌 가중치 수신부(210)에서 서버(100)로부터 Word2Vec 학습 모델의 글로벌 가중치를 수신받는 글로벌 가중치 수신단계(S210), 학습 수행부(220)에서 글로벌 가중치 정보를 토대로 Word2Vec의 학습을 수행하는 학습 수행단계(S220), 및 중간통계 가중치 전송부(230)에서 Word2Vec의 중간통계 가중치를 서버(100)로 전송하는 중간통계 가중치 전송단계(S230)를 포함한다. 즉, Word2Vec 학습 모델의 구조를 이용하여 서버(100)와 클라이언트(200) 간 개인정보와 같은 민감한 정보가 노출되지 않도록 한다.

[0086] 또한, 학습 수행단계(S220)에서는 클라이언트(200) 내의 로컬 데이터 통계를 토대로 학습을 수행한다. 이때, 로컬 데이터 통계는 개인정보를 포함할 수 있어, 이러한 민감한 정보가 노출되지 않도록 구성할 수 있다.

[0087] 또한, 학습 수행단계(S220)는 중심단어(W310)가 저장되어 있는 입력층(221), 중심단어(W310)와 가까운 단어인 문맥단어(W320)가 저장되어 있는 출력층(223), 및 입력층(221)과 출력층(223) 간 단일층으로 이루어진 은닉층(222)을 포함하여 학습을 수행한다.

[0088] 여기서, 문맥단어(W320)의 개수는 윈도우 사이즈로 나타내게 되는데, 예를 들어 윈도우 사이즈가 1일 경우 중심단어(W310)와 가장 가까운 전,후 단어 1개 씩의 단어를 포함할 수 있다.

[0089] 한편, 윈도우 사이즈를 벗어나는 단어를 가짜단어(negative word)(W330)라 하며, 이를 혼합하여 학습을 수행할 수 있다. 이때, 가짜단어(W330)를 혼합함으로써, 개인정보와 같은 민감한 정보는 가짜단어(W330)와 구분하기 어렵도록 반영할 수 있다.

[0090] 이러한 가짜단어(W330)를 혼합하여 학습을 수행함으로써 개인정보와 같은 민감한 정보가 가짜단어(W330)와 구분하기 어렵도록 반영하는 방법은, 상기 도 3 에서와 동일한 방법을 사용할 수 있으며 이에 대한 상세한 설명은 생략한다.

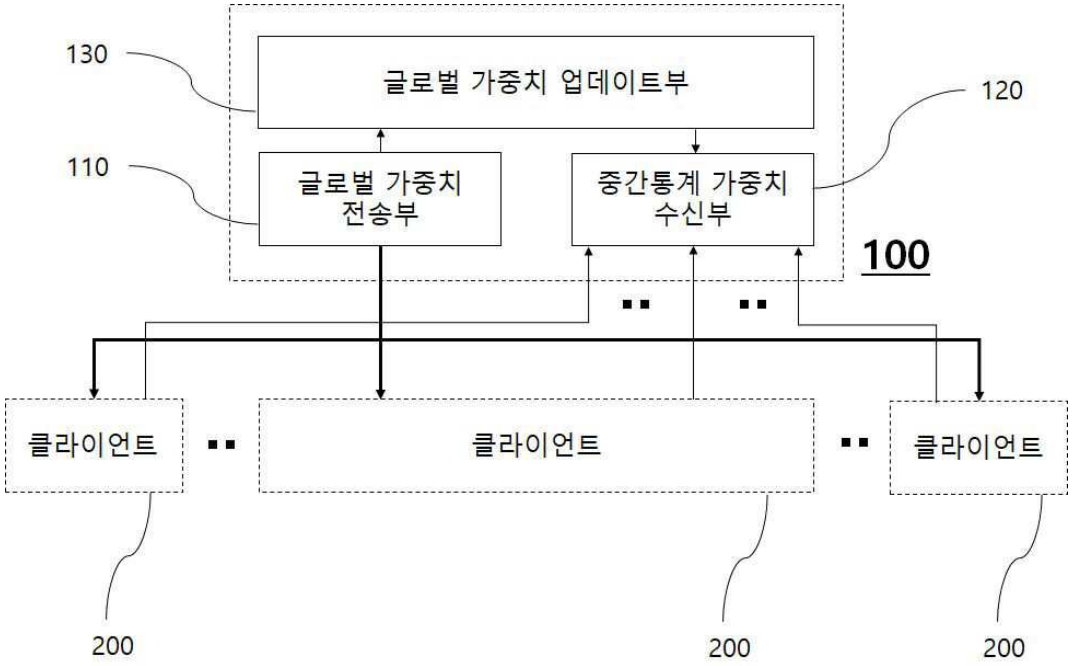
[0091] 이와 같은 본 발명에 의한 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 방법은 임베딩(Embedding) 업데이트 노출로 인해 개인정보에 민감한 특징이 직접적으로 유출되는 것을 방지하기 위해 가짜단어와 같은 가짜샘플(Negative samples)의 업데이트를 혼합하고 추론 공격을 방지하기 위해 드물게 나타나는 기능(개인정보에 민감한 특징)의 업데이트를 조정함으로써, 개인정보의 유출을 방지할 수 있다.

[0093] 이상과 같이 본 발명에 따른 Word2Vec 기반의 교차-기기 연합학습에서 효율적으로 개인정보 보호를 보장하는 공유표현학습 장치 및 방법은 개인정보에 민감한 특징을 숨기는 데 중점을 둔 Word2Vec 기반의 효율적인 개인정보 보호 연합표현학습을 수행할 수 있는 효과가 있으며, 또한 임베딩(Embedding) 업데이트 노출로 인해 개인정보에 민감한 특징이 직접적으로 유출되는 것을 방지할 수 있는 효과가 있다.

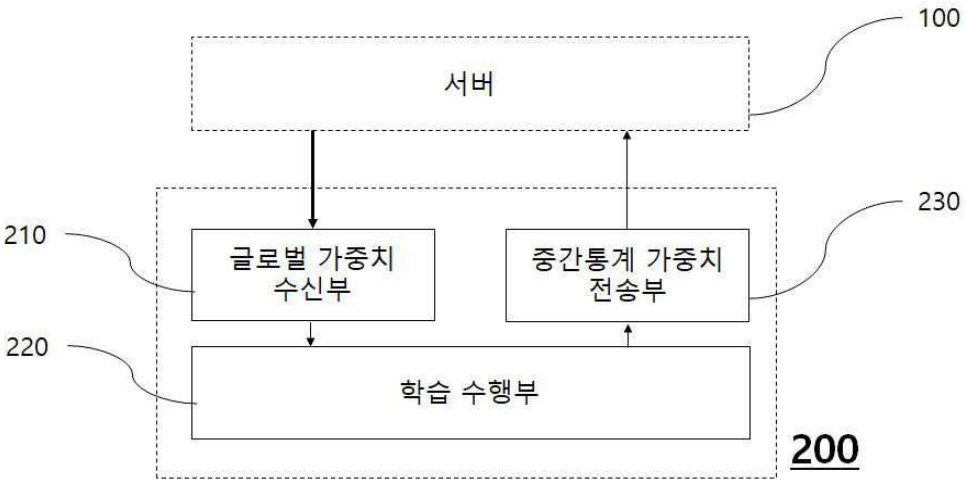
[0095] 상술한 것은 하나 이상의 실시예의 실례를 포함한다. 물론, 상술한 실시예들을 설명할 목적으로 컴포넌트들 또는 방법들의 가능한 모든 조합을 기술할 수 있는 것이 아니라, 당업자들은 다양한 실시예의 많은 추가 조합 및 치환할 수 있음을 인식할 수 있다. 따라서 설명한 실시예들은 첨부된 청구범위의 진의 및 범위 내에 있는 모든 대안, 변형 및 개조를 포함하는 것이다.

도면

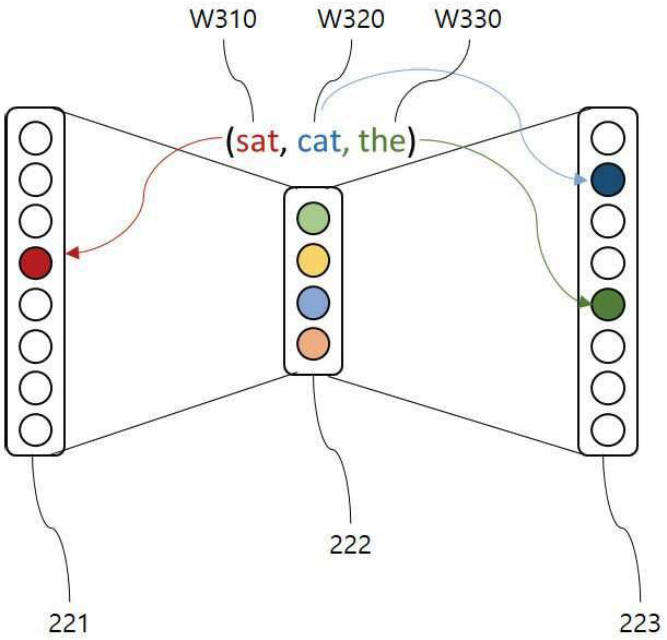
도면1



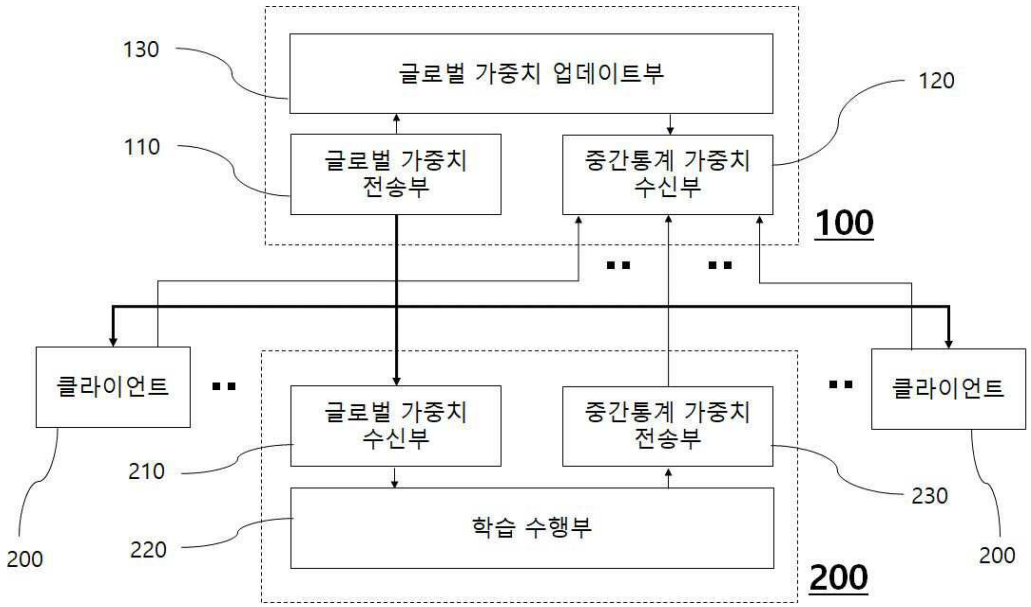
도면2



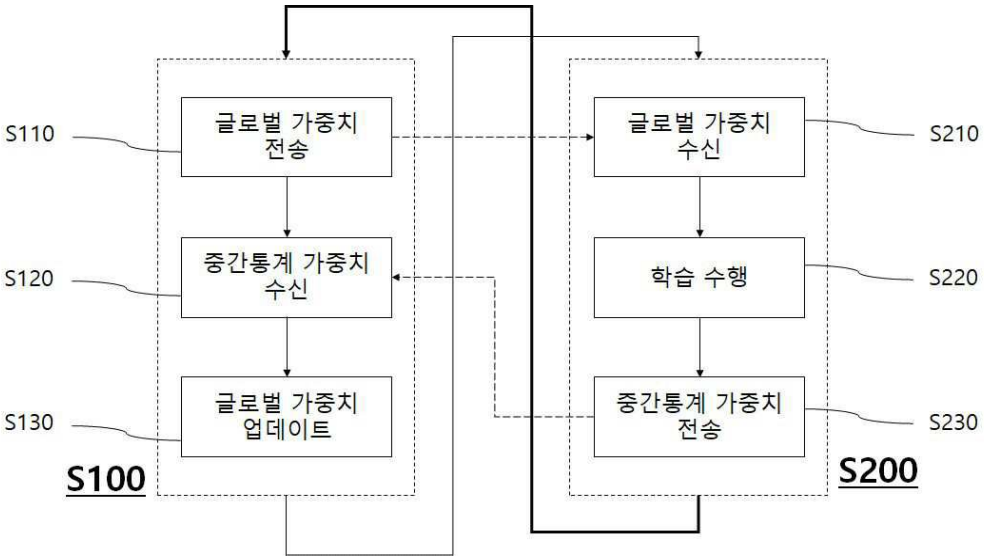
도면3


W310
W320
W330
(sat, cat, the)
222
221
223

도면4


130
글로벌 가중치 업데이트부
120
110
글로벌 가중치
전송부
중간통계 가중치
수신부
100
클라이언트
글로벌 가중치
수신부
중간통계 가중치
전송부
클라이언트
210
230
학습 수행부
200
220
200
200

도면5


S110
글로벌 가중치 전송
S120
중간통계 가중치 수신
S130
글로벌 가중치 업데이트
S100
글로벌 가중치 수신
S210
학습 수행
S220
중간통계 가중치 전송
S230
S200