



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0106331  
(43) 공개일자 2022년07월29일

- |   |   |
|---|---|
| <p>(51) 국제특허분류(Int. Cl.)<br/>G06N 3/08 (2006.01) G06N 3/04 (2006.01)</p> <p>(52) CPC특허분류<br/>G06N 3/08 (2013.01)<br/>G06F 40/253 (2020.01)</p> <p>(21) 출원번호 10-2021-0009146<br/>(22) 출원일자 2021년01월22일<br/>심사청구일자 없음</p> | <p>(71) 출원인<br/>삼성전자주식회사<br/>경기도 수원시 영통구 삼성로 129 (매탄동)<br/>포항공과대학교 산학협력단<br/>경상북도 포항시 남구 청암로 77 (지곡동)</p> <p>(72) 발명자<br/>권순철<br/>경상북도 포항시 남구 청암로 77(지곡동)<br/>이근배<br/>경상북도 포항시 남구 청암로 77(지곡동)<br/>(뒷면에 계속)</p> <p>(74) 대리인<br/>정홍식, 김태현</p> |
|---|---|

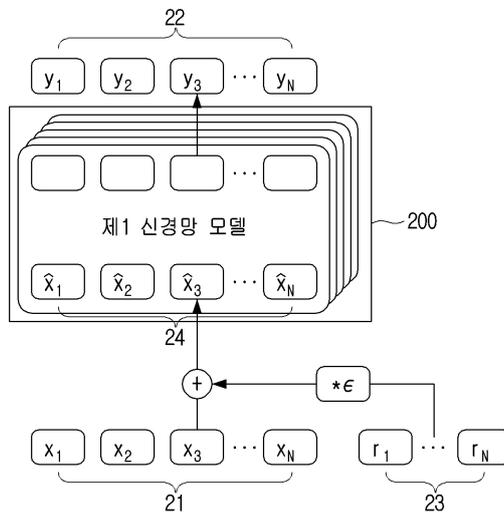
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 전자 장치 및 그 제어 방법

(57) 요약

전자 장치가 개시된다. 본 개시에 따른 전자 장치는 적어도 하나의 인스트럭션을 저장하는 메모리, 및 프로세서를 포함하고, 프로세서는, 입력 문장에 포함된 문법 오류를 교정하도록 학습된 제1 신경망 모델의 학습을 위한 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 포함하는 학습 데이터를 획득하고, 학습 데이터를 바탕으로 제1 입력 문장에 대한 정보에 대응되는 변동값을 획득하고, 제1 입력 문장에 대한 정보 및 변동값을 바탕으로 제2 입력 문장에 대한 정보를 획득하고, 제2 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 바탕으로 제1 신경망 모델을 학습시킨다.

대표도 - 도2



(52) CPC특허분류  
*G06N 3/0454* (2013.01)

(72) 발명자

**이해준**

경기도 수원시 영통구 삼성로 129(매탄동)

**유원호**

경기도 수원시 영통구 삼성로 129(매탄동)

## 명세서

### 청구범위

#### 청구항 1

전자 장치에 있어서,  
적어도 하나의 인스트럭션을 저장하는 메모리; 및  
프로세서;를 포함하고,  
상기 프로세서는,  
입력 문장에 포함된 문법 오류를 교정하도록 학습된 제1 신경망 모델의 학습을 위한 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 포함하는 학습 데이터를 획득하고,  
상기 학습 데이터를 바탕으로 상기 제1 입력 문장에 대한 정보에 대응되는 변동값을 획득하고,  
상기 제1 입력 문장에 대한 정보 및 상기 변동값을 바탕으로 제2 입력 문장에 대한 정보를 획득하고,  
상기 제2 입력 문장에 대한 정보 및 상기 제1 출력 문장에 대한 정보를 바탕으로 상기 제1 신경망 모델을 학습시키는  
전자 장치.

#### 청구항 2

제1 항에 있어서,  
상기 프로세서는,  
상기 제1 신경망 모델에 대한 손실 함수의 값이 기설정된 값보다 커지도록하는 값을 상기 변동값으로 획득하는  
전자 장치.

#### 청구항 3

제2 항에 있어서,  
상기 프로세서는,  
상기 제1 입력 문장에 대한 정보 및 상기 제1 출력 문장에 대한 정보를 상기 손실 함수에 입력하여 제1 손실 값을 획득하고,  
상기 제1 손실 값을 미분하여 상기 변동값을 획득하는  
전자 장치.

#### 청구항 4

제1 항에 있어서,  
상기 프로세서는,  
상기 제1 입력 문장에 대한 정보를 입력 문장에 대한 정보를 바탕으로 변동값을 획득하도록 학습된 제2 신경망 모델에 입력하여 상기 변동값을 획득하는  
전자 장치.

#### 청구항 5

제1 항에 있어서,

상기 프로세서는,

상기 변동값 및 상기 변동값의 크기를 조절하기 위한 하이퍼파라미터를 바탕으로 상기 제2 입력 문장에 대한 정보를 획득하는

전자 장치.

#### 청구항 6

제1 항에 있어서,

상기 프로세서는,

상기 제2 입력 문장에 대한 정보를 상기 제1 신경망 모델에 입력하여 제2 출력 문장에 대한 정보를 획득하고,

상기 제1 출력 문장에 대한 정보와 상기 제2 출력 문장에 대한 정보의 차이가 기설정된 값보다 작아지도록 상기 제1 신경망 모델을 학습시키는

전자 장치.

#### 청구항 7

전자 장치의 제어 방법에 있어서,

입력 문장에 포함된 문법 오류를 교정하도록 학습된 제1 신경망 모델의 학습을 위한 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 포함하는 학습 데이터를 획득하는 단계;

상기 학습 데이터를 바탕으로 상기 제1 입력 문장에 대한 정보에 대응되는 변동값을 획득하는 단계;

상기 제1 입력 문장에 대한 정보 및 상기 변동값을 바탕으로 제2 입력 문장에 대한 정보를 획득하는 단계; 및

상기 제2 입력 문장에 대한 정보 및 상기 제1 출력 문장에 대한 정보를 바탕으로 상기 제1 신경망 모델을 학습시키는 단계;를 포함하는

제어 방법.

#### 청구항 8

제7 항에 있어서,

상기 변동값을 획득하는 단계는,

상기 제1 신경망 모델에 대한 손실 함수의 값이 기설정된 값보다 커지도록하는 값을 상기 변동값으로 획득하는

제어 방법.

#### 청구항 9

제8 항에 있어서,

상기 변동값을 획득하는 단계는,

상기 제1 입력 문장에 대한 정보 및 상기 제1 출력 문장에 대한 정보를 상기 손실 함수에 입력하여 제1 손실 값을 획득하고,

상기 제1 손실 값을 미분하여 상기 변동값을 획득하는

제어 방법.

#### 청구항 10

제7 항에 있어서,

상기 변동값을 획득하는 단계는,

상기 제1 입력 문장에 대한 정보를 입력 문장에 대한 정보를 바탕으로 변동값을 획득하도록 학습된 제2 신경망

모델에 입력하여 상기 변동값을 획득하는 제어 방법.

**청구항 11**

제7 항에 있어서,

상기 제2 입력 문장에 대한 정보를 획득하는 단계는,

상기 변동값 및 상기 변동값의 크기를 조절하기 위한 하이퍼파라미터를 바탕으로 상기 제2 입력 문장에 대한 정보를 획득하는

제어 방법.

**청구항 12**

제7 항에 있어서,

상기 제1 신경망 모델을 학습시키는 단계는,

상기 제2 입력 문장에 대한 정보를 상기 제1 신경망 모델에 입력하여 제2 출력 문장에 대한 정보를 획득하고,

상기 제1 출력 문장에 대한 정보와 상기 제2 출력 문장에 대한 정보의 차이가 기설정된 값보다 작아지도록 상기 제1 신경망 모델을 학습시키는

제어 방법.

**청구항 13**

제7 항 내지 제12 항 중에 어느 한 항의 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 비일시적인(non-transitory) 기록매체.

**발명의 설명**

**기술 분야**

[0001] 본 개시는 전자 장치 및 그 제어 방법으로, 보다 상세하게는, 입력 문장에 포함된 문법 오류를 교정하는 신경망 모델을 학습시키는 전자 장치 및 그 제어 방법에 관한 것이다.

**배경 기술**

[0002] 인공지능(Artificial Intelligence, AI) 시스템은 인간 수준의 지능을 구현하는 컴퓨터 시스템이며, 기존 Rule 기반 스마트 시스템과 달리 기계가 스스로 학습하고 판단하며 똑똑해지는 시스템이다. 인공지능 시스템은 사용할수록 인식률이 향상되고 사용자 취향을 보다 정확하게 이해할 수 있게 되어, 기존 Rule 기반 스마트 시스템은 점차 딥러닝 기반 인공지능 시스템으로 대체되고 있다.

[0003] 최근에는, 인공지능을 이용한 문법 오류 교정(GEC: Grammatical error correction)에 대한 연구가 활발히 진행되고 있다. 문법 오류 교정이란 문법적 오류를 포함하는 문장을 입력으로 받아서, 문장에 포함된 문법 오류들을 교정한 문장을 출력하는 동작을 의미한다. 인공지능에 기초한 대표적인 문법 오류 교정 모델로는 ‘문법적으로 틀린 언어’를 바탕으로 ‘문법적으로 올바른 언어’를 출력하도록 학습되는 기계 번역 모델(machine translation model)이 있다.

[0004] 한편, 기계 번역 모델의 성능을 향상시키기 위해서는 대량의 학습 데이터가 필요하다. 특히, 학습 데이터는 문법적 오류를 포함하는 문장과 문법적 오류가 교정된 문장이 한 쌍으로 구성되는데, 이러한 학습 데이터를 수집하기 위해서는 사람이 직접 문법적 오류를 교정해야한다. 따라서, 학습 데이터 수집을 위해 많은 시간이 소요되며 대량 수집이 어렵다는 문제가 있다.

[0005] 이에 따라, 새로운 학습 데이터를 획득하는 방법에 대한 필요성이 대두된다.

**발명의 내용**

**해결하려는 과제**

[0006] 본 발명이 해결하고자 하는 일 기술적 과제는, 새로운 학습 데이터를 획득할 수 있는 전자 장치를 제공하는 것이다.

[0007] 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명의 기술분야에서의 통상의 기술자에게 명확하게 이해 될 수 있을 것이다.

**과제의 해결 수단**

[0008] 상술한 기술적 과제를 해결하기 위한 본 개시의 예시적인 일 실시 예에 따르면, 전자 장치에 있어서, 적어도 하나의 인스트럭션을 저장하는 메모리; 및 프로세서;를 포함하고, 상기 프로세서는, 입력 문장에 포함된 문법 오류를 교정하도록 학습된 제1 신경망 모델의 학습을 위한 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 포함하는 학습 데이터를 획득하고, 상기 학습 데이터를 바탕으로 상기 제1 입력 문장에 대한 정보에 대응되는 변동값을 획득하고, 상기 제1 입력 문장에 대한 정보 및 상기 변동값을 바탕으로 제2 입력 문장에 대한 정보를 획득하고, 상기 제2 입력 문장에 대한 정보 및 상기 제1 출력 문장에 대한 정보를 바탕으로 상기 제1 신경망 모델을 학습시키는 전자 장치가 제공될 수 있다.

[0009] 상술한 기술적 과제를 해결하기 위한 본 개시의 예시적인 다른 일 실시 예에 따르면, 전자 장치의 제어 방법에 있어서, 입력 문장에 포함된 문법 오류를 교정하도록 학습된 제1 신경망 모델의 학습을 위한 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 포함하는 학습 데이터를 획득하는 단계; 상기 학습 데이터를 바탕으로 상기 제1 입력 문장에 대한 정보에 대응되는 변동값을 획득하는 단계; 상기 제1 입력 문장에 대한 정보 및 상기 변동값을 바탕으로 제2 입력 문장에 대한 정보를 획득하는 단계; 및 상기 제2 입력 문장에 대한 정보 및 상기 제1 출력 문장에 대한 정보를 바탕으로 상기 제1 신경망 모델을 학습시키는 단계;를 포함하는 제어 방법이 제공될 수 있다.

[0010] 본 개시의 과제의 해결 수단이 상술한 해결 수단들로 제한되는 것은 아니며, 언급되지 아니한 해결 수단들은 본 명세서 및 첨부된 도면으로부터 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

**발명의 효과**

[0011] 이상과 같은 본 개시의 다양한 실시 예에 따르면, 전자 장치는 새로운 학습 데이터를 획득할 수 있다. 이에 따라, 문법 오류 교정 모델의 과적합(overfitting)이 방지될 수 있으며, 문법 오류 교정 모델의 성능이 향상될 수 있다

[0012] 그 외에 본 개시의 실시 예로 인하여 얻을 수 있거나 예측되는 효과에 대해서는 본 개시의 실시 예에 대한 상세한 설명에서 직접적 또는 암시적으로 개시하도록 한다. 예컨대, 본 개시의 실시 예에 따라 예측되는 다양한 효과에 대해서는 후술될 상세한 설명 내에서 개시될 것이다.

**도면의 간단한 설명**

- [0013] 도 1은 문법 오류 교정 모델의 학습 방법 및 동작을 설명하기 위한 도면이다.
- 도 2는 본 개시의 일 실시 예에 따른 전자 장치의 동작을 설명하기 위한 도면이다.
- 도 3은 본 개시의 다른 일 실시 예에 따른 변동 값을 획득하는 방법을 설명하기 위한 도면이다.
- 도 4는 본 개시의 일 실시 예에 따른 전자 장치의 구성을 도시한 블록도이다.
- 도 5는 본 개시의 일 실시 예에 따른 전자 장치의 제어 방법을 도시한 순서도이다.

**발명을 실시하기 위한 구체적인 내용**

[0014] 본 명세서에서 사용되는 용어에 대해 간략히 설명하고, 본 개시에 대해 구체적으로 설명하기로 한다.

[0015] 본 개시의 실시 예에서 사용되는 용어는 본 개시에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인

용어들을 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 개시의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 개시에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 개시의 전반에 걸친 내용을 토대로 정의되어야 한다.

- [0016] 본 개시의 실시 예들은 다양한 변환을 가할 수 있고 여러 가지 실시 예를 가질 수 있는바, 특정 실시 예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나 이는 특정한 실시 형태에 대해 범위를 한정하려는 것이 아니며, 개시된 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 실시 예들을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.
- [0017] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 구성요소들은 용어들에 의해 한정되어서는 안 된다. 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.
- [0018] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "구성되다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0019] 아래에서는 첨부한 도면을 참고하여 본 개시의 실시 예에 대하여 본 개시가 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 개시는 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시 예에 한정되지 않는다. 그리고 도면에서 본 개시를 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0021] 도 1은 문법 오류 교정 모델의 학습 방법 및 동작을 설명하기 위한 도면이다. 문법 오류 교정 모델(100)은 입력 문장(11)에 포함된 문법 오류를 교정하기 위한 모델이다. 문법 오류 교정 모델(100)은 서로 대응되는 입력 문장(11) 및 출력 문장(12)을 포함하는 학습 데이터를 바탕으로 학습될 수 있다. 구체적으로, 문법 오류 교정 모델(100)은 입력 문장(11)을 입력 받아, 입력 문장(11)에 포함된 문법적 오류가 교정된 출력 문장(12)을 획득하도록 학습될 수 있다. 학습된 문법 오류 교정 모델(100)은 입력 문장(11)에 포함된 복수의 단어 ( $x_1, x_2, x_3, \dots, x_N$ ) 중 적어도 하나를 다른 단어로 변경할 수 있다. 예로, 제1 단어( $x_1$ )가 제2 단어( $y_1$ )로 변경될 수 있다. 한편, 입력 문장(11) 및 출력 문장(12)에 포함된 각각의 단어(x, y)는 텍스트 또는 벡터일 수 있다.
- [0022] 한편, 문법 오류 교정 모델(100)의 성능을 향상시키려면 대량의 학습 데이터를 수집하는 것이 중요하다. 문법 오류 교정 모델(100)의 학습 데이터가 충분하지 못하면, 문법 오류 교정 모델(100)이 과적합되어 학습 데이터에 포함되지 않은 새로운 데이터가 입력되었을 때, 이를 잘 처리하지 못할 수도 있기 때문이다.
- [0023] 이항에서는, 문법 오류 교정 모델(100)의 학습을 위한 학습 데이터를 획득하는 방법에 대하여 설명하도록 한다. 한편, 이하에서, 문법 오류 교정 모델(100)은 제1 신경망 모델로 지칭될 수 있다.
- [0024] 도 2는 본 개시의 일 실시 예에 따른 전자 장치의 동작을 설명하기 위한 도면이다. 전자 장치(400)는 제1 입력 문장에 대한 정보(21) 및 제1 출력 문장에 대한 정보(22)를 포함하는 학습 데이터를 바탕으로 제1 신경망 모델(200)을 학습시킬 수 있다. 여기서, 제1 입력 문장에 대한 정보(21) 및 제1 출력 문장에 대한 정보(22)는 전자 장치(400)에 미리 저장되어 있을 수 있다. 또한, 제1 입력 문장에 대한 정보(21)는 제1 입력 문장에 포함된 단어에 대응되는 벡터를 포함할 수 있다. 그리고, 제1 출력 문장에 대한 정보(22)는 제1 출력 문장에 포함된 단어에 대응되는 벡터를 포함할 수 있다.
- [0025] 한편, 전자 장치(400)는 제1 신경망 모델(200)의 학습을 위한 새로운 입력 문장에 대한 정보를 획득할 수 있다. 예로, 전자 장치(400)는 제1 입력 문장에 대한 정보(21)에 변동 값(perturbation value)(23)을 부가하여 제2 입력 문장에 대한 정보(24)를 획득할 수 있다. 여기서, 변동 값(perturbation value)(23)은 임의의 값을 갖는 벡터일 수 있다. 한편, 전자 장치(400)는 제1 신경망 모델(200)에 대한 비용 함수의 값이 증가하도록 하는 변동 값(23)을 획득할 수 있다. 이에 따라, 학습 데이터 획득 모듈(210)은 제1 신경망 모델(200)이 잘 처리하지 못하는 제2 입력 문장에 대한 정보(24)를 획득할 수 있다.

[0026] 한편, 전자 장치(400)는 변동 값(23)에 하이퍼파라미터( $\epsilon$ )를 곱한 값을 제1 입력 문장에 대한 정보(21)에 더하여 제2 입력 문장에 대한 정보(24)를 획득할 수 있다. 하이퍼파라미터( $\epsilon$ )는 변동 값(23)의 크기를 조절하기 위한 0에서 1 사이의 값일 수 있다. 전자 장치(400)는 제1 입력 문장에 대한 정보(21)와 제2 입력 문장에 대한 정보(24)의 차이가 기 설정된 범위 이내가 되도록 하는 하이퍼파라미터( $\epsilon$ )를 결정할 수 있다.

[0027] 이와 같이, 새로운 입력 문장인 제2 입력 문장에 대한 정보(24)가 획득되면, 전자 장치(400)는 제2 입력 문장에 대한 정보(24)를 포함하는 학습 데이터를 바탕으로 제1 신경망 모델(200)을 학습시킬 수 있다. 구체적으로, 전자 장치(400)는 제1 신경망 모델(200)이 제2 입력 문장에 대한 정보(24)를 입력 받아 제1 출력 문장에 대한 정보(22)를 출력하도록 제1 신경망 모델(200)을 학습시킬 수 있다. 학습 데이터에 제2 입력 문장에 대한 정보(24)가 추가되어 전체 학습 데이터의 양이 증가하므로, 제1 신경망 모델(200)의 성능은 향상될 수 있다.

[0028] 한편, 전자 장치(400)는 다양한 방법으로 변동 값(23) 및 제2 입력 문장(24)에 대한 정보를 획득할 수 있다. 전자 장치(400)는 [수학식 1], [수학식 2] 및 [수학식 3]에 기초하여 변동 값(23)을 획득할 수 있다.

[0029] [수학식 1]

[0030] 
$$\hat{x} = x + \epsilon r$$

[0031] [수학식 2]

[0032] 
$$\lambda = - \sum_{i=1 \dots N} \log(P(y_i | x + \epsilon r))$$

[0033] [수학식 3]

[0034] 
$$r = \text{sign}(\nabla_x \lambda)$$

[0035] 여기서,  $\hat{x}$ 는 제2 입력 문장에 대한 정보,  $x$ 는 제1 입력 문장에 대한 정보,  $\lambda$ 는 제1 신경망 모델(200)에 대한 손실 함수,  $y_i$ 는 제1 출력 문장에 대한 정보,  $r$ 은 변동 값을 의미할 수 있다.

[0036] 전자 장치(400)는 제1 입력 문장에 대한 정보(21) 및 제1 출력 문장에 대한 정보(22)를 손실 함수( $\lambda$ )에 입력하여 손실 값을 획득할 수 있다. 그리고, 전자 장치(400)는 손실 값을 제1 입력 문장에 대해 미분하여 변동 값(23)을 획득할 수 있다.

[0037] 도 3은 본 개시의 다른 일 실시 예에 따른 변동 값을 획득하는 방법을 설명하기 위한 도면이다. 전자 장치(400)는 제1 입력 문장에 대한 정보(31)를 제2 신경망 모델(320)에 입력하여 변동 값(33)을 획득할 수 있다. 여기서, 제2 신경망 모델(320)은 제1 입력 문장에 대한 정보(31)를 바탕으로 손실 함수( $\lambda$ )의 값이 기설정된 값보다 커지도록 하는 변동 값(33)을 출력하도록 학습될 수 있다.

[0038] 전자 장치(400)는 변동 값(33)을 바탕으로 제2 입력 문장에 대한 정보(34)를 획득할 수 있다. 예로, 전자 장치(400)는 [수학식 1]에 기초하여 제2 입력 문장에 대한 정보(34)를 획득할 수 있다. 그리고, 전자 장치(400)는 제2 입력 문장에 대한 정보(34) 및 제1 출력 문장에 대한 정보(32)를 바탕으로 제1 신경망 모델(310)을 학습시킬 수 있다. 예를 들어, 전자 장치(400)는 제2 입력 문장에 대한 정보(34)를 제1 신경망 모델(310)에 입력하여 제2 출력 문장에 대한 정보(미도시)를 획득할 수 있다. 그리고, 전자 장치(400)는 제1 출력 문장에 대한 정보(32)와 제2 출력 문장에 대한 정보의 차이가 기설정된 값보다 작아지도록 제1 신경망 모델(310)을 학습시킬 수 있다.

[0039] 한편, 도 3에서는 제1 신경망 모델(310) 및 제2 신경망 모델(320)을 분리하여 각각 도시하였으나, 제1 신경망 모델(310) 및 제2 신경망 모델(320)은 통합된 하나의 신경망 모델로 구현될 수 있다.

[0040] 도 4는 본 개시의 일 실시 예에 따른 전자 장치의 구성을 도시한 블록도이다. 전자 장치(400)는 메모리(410) 및 프로세서(420)를 포함할 수 있다. 예로, 전자 장치(400)는 서버일 수 있다.

- [0041] 메모리(410)는 전자 장치(400)의 구성요소들의 전반적인 동작을 제어하기 위한 운영체제(OS: Operating System) 및 전자 장치(400)의 구성요소와 관련된 명령 또는 데이터를 저장할 수 있다. 이를 위해 메모리(410)는 비휘발성 메모리(ex: 하드 디스크, SSD(Solid state drive), 플래시 메모리), 휘발성 메모리 등으로 구현될 수 있다. 특히, 메모리(410)는 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 포함하는 학습 데이터를 저장할 수 있다. 또한, 메모리(410)는 문법 오류 교정을 위한 제1 신경망 모델 및 변동 값 생성을 위한 제2 신경망 모델을 저장할 수 있다.
- [0042] 프로세서(420)는 전자 장치(400)의 전반적인 동작을 제어할 수 있다.
- [0043] 예로, 프로세서(420)는 제1 신경망 모델의 학습을 위한 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 포함하는 학습 데이터를 획득할 수 있다. 여기서, 제1 신경망 모델은 입력 문장에 포함된 문법 오류를 교정하기 위한 모델일 수 있다.
- [0044] 프로세서(420)는 학습 데이터를 바탕으로 제1 입력 문장에 대한 정보에 대응되는 변동 값을 획득할 수 있다. 특히, 프로세서(420)는 제1 신경망 모델에 대한 손실 함수의 값이 기설정된 값보다 커지도록 하는 값을 변동값으로 획득할 수 있다. 일 예로, 프로세서(420)는 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 제1 신경망 모델에 대한 손실 함수에 입력하여 제1 손실 값을 획득할 수 있다. 그리고, 프로세서(420)는 제1 손실 값을 미분하여 변동 값을 획득할 수 있다. 다른 일 예로, 프로세서(420)는 제1 입력 문장에 대한 정보를 제2 신경망 모델에 입력하여 변동 값을 획득할 수 있다. 여기서, 제2 신경망 모델은 입력 문장에 대한 정보를 바탕으로 변동 값을 출력하도록 학습된 CNN (Convolutional Neural Network)일 수 있다. 다만, 이는 일 실시 예에 불과하며, 제2 신경망 모델은 RNN (Recurrent Neural Network)으로 구현될 수도 있다.
- [0045] 프로세서(420)는 제1 입력 문장에 대한 정보 및 변동 값을 바탕으로 제2 입력 문장에 대한 정보를 획득할 수 있다. 예를 들어, 프로세서(420)는 변동 값에 하이퍼파라미터를 곱하고, 하이퍼파라미터가 반영된 변동 값을 제1 입력 문장에 대한 정보에 부가하여 제2 입력 문장에 대한 정보를 획득할 수 있다.
- [0046] 프로세서(420)는 제2 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 바탕으로 제1 신경망 모델을 학습시킬 수 있다. 예를 들어, 프로세서(420)는 제2 입력 문장에 대한 정보를 제1 신경망 모델에 입력하여 제2 출력 문장에 대한 정보를 획득할 수 있다. 그리고, 프로세서(420)는 제1 출력 문장에 대한 정보와 제2 출력 문장에 대한 정보의 차이가 기설정된 값보다 작아지도록 제1 신경망 모델을 학습시킬 수 있다.
- [0047] 한편, 본 개시에 따른 인공지능과 관련된 기능은 프로세서(420)와 메모리(410)를 통해 동작된다. 프로세서(420)는 하나 또는 복수의 프로세서로 구성될 수 있다. 이때, 하나 또는 복수의 프로세서는 CPU, AP, DSP(Digital Signal Processor) 등과 같은 범용 프로세서, GPU, VPU(Vision Processing Unit)와 같은 그래픽 전용 프로세서 또는 NPU와 같은 인공지능 전용 프로세서일 수 있다. 하나 또는 복수의 프로세서는, 메모리(410)에 저장된 기 정의된 동작 규칙 또는 인공지능 모델에 따라, 입력 데이터를 처리하도록 제어한다. 또는, 하나 또는 복수의 프로세서가 인공지능 전용 프로세서인 경우, 인공지능 전용 프로세서는, 특정 인공지능 모델의 처리에 특화된 하드웨어 구조로 설계될 수 있다.
- [0048] 기 정의된 동작 규칙 또는 인공지능 모델은 학습을 통해 만들어진 것을 특징으로 한다. 여기서, 학습을 통해 만들어진다는 것은, 기본 인공지능 모델이 학습 알고리즘에 의하여 다수의 학습 데이터들을 이용하여 학습됨으로써, 원하는 특성(또는, 목적)을 수행하도록 설정된 기 정의된 동작 규칙 또는 인공지능 모델이 만들어짐을 의미한다. 이러한 학습은 본 개시에 따른 인공지능이 수행되는 기기 자체에서 이루어질 수도 있고, 별도의 서버 및/또는 시스템을 통해 이루어 질 수도 있다. 학습 알고리즘의 예로는, 지도형 학습(supervised learning), 비지도형 학습(unsupervised learning), 준지도형 학습(semi-supervised learning) 또는 강화 학습(reinforcement learning)이 있으나, 전술한 예에 한정되지 않는다.
- [0049] 인공지능 모델은 학습을 통해 만들어 질 수 있다. 여기서, 학습을 통해 만들어진다는 것은, 기본 인공지능 모델이 학습 알고리즘에 의하여 다수의 학습 데이터들을 이용하여 학습됨으로써, 원하는 특성(또는, 목적)을 수행하도록 설정된 기 정의된 동작 규칙 또는 인공지능 모델이 만들어짐을 의미한다. 인공지능 모델은, 복수의 신경망 레이어들로 구성될 수 있다. 복수의 신경망 레이어들 각각은 복수의 가중치들(weight values)을 갖고 있으며, 이전(previous) 레이어의 연산 결과와 복수의 가중치들 간의 연산을 통해 신경망 연산을 수행한다. 복수의 신경망 레이어들이 갖고 있는 복수의 가중치들은 인공지능 모델의 학습 결과에 의해 최적화될 수 있다. 예를 들어, 학습 과정 동안 인공지능 모델에서 획득한 로스(loss) 값 또는 코스트(cost) 값이 감소 또는 최소화되도록 복수의 가중치들이 갱신될 수 있다.

- [0050] 인공 신경망은 심층 신경망(DNN:Deep Neural Network)를 포함할 수 있으며, 예를 들어, CNN (Convolutional Neural Network), DNN (Deep Neural Network), RNN (Recurrent Neural Network), GAN (Generative Adversarial Network), RBM (Restricted Boltzmann Machine), DBN (Deep Belief Network), BRDNN(Bidirectional Recurrent Deep Neural Network), 심층 Q-네트워크 (Deep Q-Networks) 또는 LSTM (Long Short Term Memory), Transformer 등이 있으나, 전술한 예에 한정되지 않는다.
- [0051] 도 5는 본 개시의 일 실시 예에 따른 전자 장치의 제어 방법을 도시한 순서도이다.
- [0052] 전자 장치(400)는 제1 신경망 모델의 학습을 위한 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 포함하는 학습 데이터를 획득할 수 있다(S510). 전자 장치(400)는 메모리에 미리 저장된 학습 데이터를 획득할 수 있다. 여기서, 제1 신경망 모델은 입력 문장에 포함된 문법 오류를 교정하도록 학습되는 모델이다.
- [0053] 전자 장치(400)는 학습 데이터를 바탕으로 변동 값을 획득할 수 있다(S520). 전자 장치(400)는 제1 신경망 모델에 대한 손실 함수의 값이 기설정된 값보다 커지도록하는 변동값을 획득할 수 있다. 예를 들어, 전자 장치(400)는 제1 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 손실 함수에 입력하여 제1 손실 값을 획득하고, 제1 손실 값을 미분하여 변동값을 획득할 수 있다. 또는, 전자 장치(400)는 제1 입력 문장에 대한 정보를 입력 문장에 대한 정보를 바탕으로 변동값을 획득하도록 학습된 제2 신경망 모델에 입력하여 변동값을 획득할 수 있다.
- [0054] 전자 장치(400)는 제1 입력 문장에 대한 정보 및 변동 값을 바탕으로 제2 입력 문장에 대한 정보를 획득할 수 있다(S530). 전자 장치(400)는 변동값에 하이퍼파라미터를 곱하고 제1 입력 문장에 대한 정보에 더하여 제2 입력 문장에 대한 정보를 획득할 수 있다. 여기서, 제1 입력 문장에 대한 정보 및 제2 입력 문장에 대한 정보는 각각 벡터일 수 있다.
- [0055] 전자 장치(400)는 제2 입력 문장에 대한 정보 및 제1 출력 문장에 대한 정보를 바탕으로 제1 신경망 모델을 학습할 수 있다(S540). 전자 장치(400)는 제2 입력 문장에 대한 정보를 제1 신경망 모델에 입력하여 제2 출력 문장에 대한 정보를 획득할 수 있다. 그리고, 전자 장치(400)는 제1 출력 문장에 대한 정보와 제2 출력 문장에 대한 정보의 차이가 기설정된 값보다 작아지도록 제1 신경망 모델을 학습시킬 수 있다.
- [0056] 한편, 이상에서 설명된 다양한 실시 예들은 소프트웨어(software), 하드웨어(hardware) 또는 이들의 조합을 이용하여 컴퓨터(computer) 또는 이와 유사한 장치로 읽을 수 있는 기록 매체 내에서 구현될 수 있다. 일부 경우에 있어 본 명세서에서 설명되는 실시 예들이 프로세서 자체로 구현될 수 있다. 소프트웨어적인 구현에 의하면, 본 명세서에서 설명되는 절차 및 기능과 같은 실시 예들은 별도의 소프트웨어 모듈들로 구현될 수 있다. 소프트웨어 모듈들 각각은 본 명세서에서 설명되는 하나 이상의 기능 및 작동을 수행할 수 있다.
- [0057] 한편, 상술한 본 개시의 다양한 실시 예들에 따른 처리 동작을 수행하기 위한 컴퓨터 명령어(computer instructions)는 비일시적 컴퓨터 판독 가능 매체(non-transitory computer-readable medium)에 저장될 수 있다. 이러한 비일시적 컴퓨터 판독 가능 매체에 저장된 컴퓨터 명령어는 프로세서에 의해 실행되었을 때 상술한 다양한 실시 예에 따른 처리 동작을 특정 기기가 수행하도록 할 수 있다.
- [0058] 비일시적 컴퓨터 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 비일시적 컴퓨터 판독 가능 매체의 구체적인 예로는, CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM 등이 있을 수 있다.
- [0059] 한편, 기기로 읽을 수 있는 저장매체는, 비일시적(non-transitory) 저장매체의 형태로 제공될 수 있다. 여기서, '비일시적 저장매체'는 실재(tangible)하는 장치이고, 신호(signal)(예: 전자기파)를 포함하지 않는다는 것을 의미할 뿐이며, 이 용어는 데이터가 저장매체에 반영구적으로 저장되는 경우와 임시적으로 저장되는 경우를 구분하지 않는다. 예로, '비일시적 저장매체'는 데이터가 임시적으로 저장되는 버퍼를 포함할 수 있다.
- [0060] 일 실시예에 따르면, 본 문서에 개시된 다양한 실시예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다. 컴퓨터 프로그램 제품은 상품으로서 판매자 및 구매자 간에 거래될 수 있다. 컴퓨터 프로그램 제품은 기기로 읽을 수 있는 저장 매체(예: compact disc read only memory (CD-ROM))의 형태로 배포되거나, 또는 어플리케이션 스토어(예: 플레이 스토어™)를 통해 또는 두개의 사용자 장치들(예: 스마트폰들) 간에 직접, 온라인으로 배포(예: 다운로드 또는 업로드)될 수 있다. 온라인 배포의 경우에, 컴퓨터 프로그램 제품(예: 다운로드할 앱(downloadable app))의 적어도 일부는 제조사의 서버, 어플리케이션 스토어의

서버, 또는 중계 서버의 메모리와 같은 기기로 읽을 수 있는 저장 매체에 적어도 일시 저장되거나, 임시적으로 생성될 수 있다.

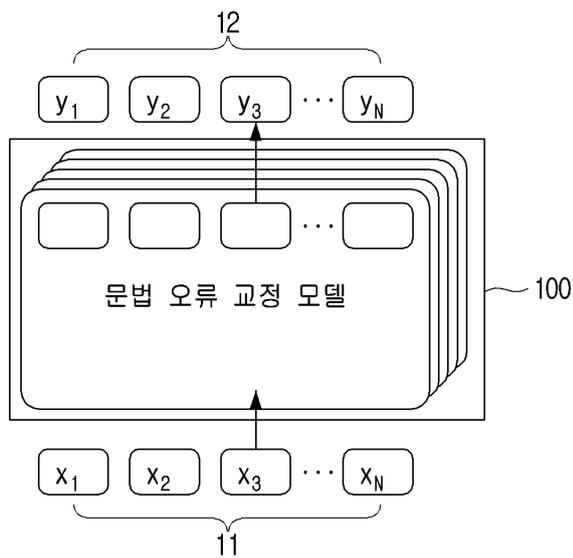
[0061] 이상에서는 본 개시의 바람직한 실시 예에 대하여 도시하고 설명하였지만, 본 개시는 상술한 특정의 실시 예에 한정되지 아니하며, 청구범위에서 청구하는 본 개시의 요지를 벗어남이 없이 당해 개시에 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 개시의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

**부호의 설명**

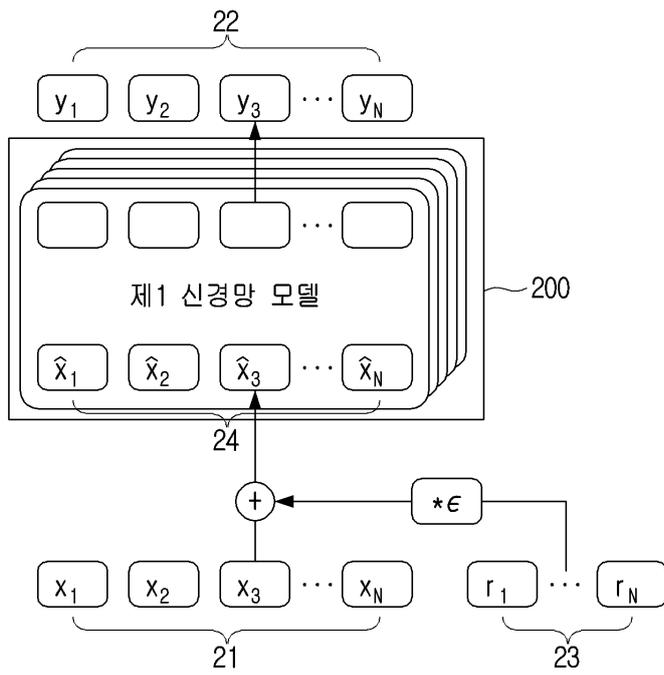
[0062] 400: 전자 장치 410: 메모리  
420: 프로세서

**도면**

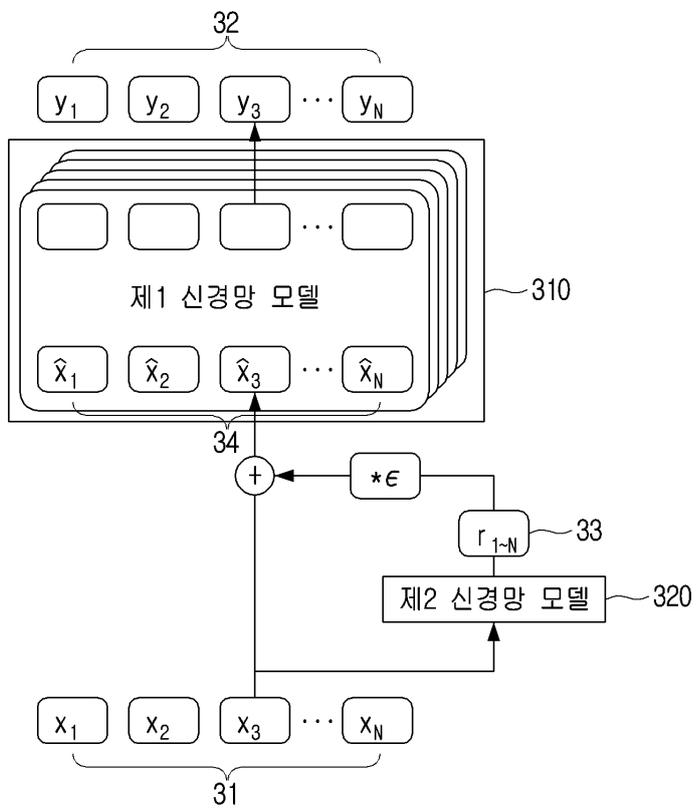
**도면1**



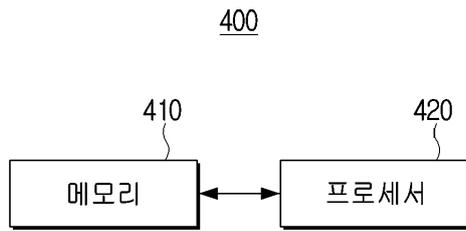
도면2



도면3



도면4



도면5

