



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0080932
(43) 공개일자 2024년06월07일

(51) 국제특허분류(Int. Cl.)

G06F 11/34 (2006.01) G06F 11/14 (2006.01)

G06F 11/22 (2017.01) G06F 11/36 (2006.01)

G06N 20/00 (2019.01)

(52) CPC특허분류

G06F 11/3428 (2013.01)

G06F 11/1476 (2013.01)

(21) 출원번호 10-2022-0164735

(22) 출원일자 2022년11월30일

심사청구일자 2022년11월30일

(71) 출원인

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

임희석

서울특별시 성북구 안암로 145 자연계캠퍼스 애기능생활관 311호 NLP&AI연구실

구선민

경기도 의정부시 오목로 171, 204동 1504호

(74) 대리인

김동용

전체 청구항 수 : 총 9 항

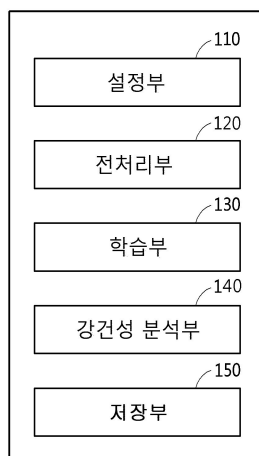
(54) 발명의 명칭 손실 조정을 이용한 벤치마크 데이터 셋의 강건성 분석 장치 및 방법

(57) 요약

손실 조정을 이용한 벤치마크 데이터 셋의 강건성 분석 장치가 개시된다. 상기 강건성 분석 장치는 사전학습 언어 모델(Pre-trained Language Model, PLM), 윈도우 사이즈(window size), 강건성 평가의 대상인 벤치마크 데이터 셋(benchmark dataset), 및 스코어 함수 중 적어도 하나에 대한 설정 동작을 수행하는 설정부, 상기 벤치마크 데이터 셋을 전처리하여 상기 배치(batch)를 구성하는 전처리부, 상기 전처리부에 의해 구성된 배치 구성에 따라 상기 PLM을 학습시키는 학습부, 및 상기 PLM에 대한 학습 과정에서 도출되는 손실(loss) 값에 기초하여 상기 벤치마크 데이터 셋의 강건성을 분석하는 강건성 분석부를 포함한다.

대 표 도 - 도1

100



(52) CPC특허분류

G06F 11/2263 (2013.01)

G06F 11/3688 (2013.01)

G06F 11/3692 (2013.01)

G06N 20/00 (2021.08)

이 발명을 지원한 국가연구개발사업

과제고유번호	1345349433
과제번호	2021R1A6A1A03045425
부처명	교육부
과제관리(전문)기관명	한국연구재단
연구사업명	이공학학술연구기반구축
연구과제명	Human-inspired AI 연구소
기 여 율	1/2
과제수행기관명	고려대학교
연구기간	2022.03.01 ~ 2023.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711159969
과제번호	2018-0-01405-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	Human-inspired 복합지능 원천기술 연구 및 인력양성
기 여 율	1/2
과제수행기관명	고려대학교산학협력단
연구기간	2022.01.01 ~ 2022.12.31

명세서

청구범위

청구항 1

사전학습 언어 모델(Pre-trained Language Model, PLM), 윈도우 사이즈(window size), 강건성 평가의 대상인 벤치마크 데이터 셋(benchmark dataset), 및 스코어 함수 중 적어도 하나에 대한 설정 동작을 수행하는 설정부;
상기 벤치마크 데이터 셋을 전처리하여 상기 배치(batch)를 구성하는 전처리부;
상기 전처리부에 의해 구성된 배치 구성에 따라 상기 PLM을 학습시키는 학습부; 및
상기 PLM에 대한 학습 과정에서 도출되는 손실(loss) 값에 기초하여 상기 벤치마크 데이터 셋의 강건성을 분석하는 강건성 분석부를 포함하는 강건성 분석 장치.

청구항 2

제1항에 있어서,
상기 스코어 함수는 MLM 스코어(Masked Language Model scoring) 함수 또는 ML 스코어(Language Model based sentences scoring) 함수인,
강건성 분석 장치.

청구항 3

제2항에 있어서,
상기 전처리부는 상기 벤치마크 데이터 셋에 포함된 각 데이터에 대한 스코어 값을 산출하고, 산출된 스코어 값에 기초하여 상기 배치를 구성하는,
강건성 분석 장치.

청구항 4

제3항에 있어서,
상기 전처리부는 상기 벤치마크 데이터 셋의 각 데이터를 엔트로피가 낮은 순서대로 정렬한 후 상기 배치를 구성하는,
강건성 분석 장치.

청구항 5

제4항에 있어서,
상기 전처리부는 상기 벤치마크 데이터 셋의 각 데이터를 산출된 스코어 값의 내림차순으로 정렬하는,
강건성 분석 장치.

청구항 6

제5항에 있어서,
상기 전처리부는 엔트로피가 낮은 순서대로 정렬된 상기 벤치마크 데이터 셋의 각 데이터를 상기 윈도우 사이즈 만큼 반복적으로 학습하도록 상기 배치를 구성하는,
강건성 분석 장치.

청구항 7

제5항에 있어서,

상기 전처리부는 엔트로피가 낮은 순서대로 정렬된 상기 벤치마크 데이터 셋의 각 데이터를 상기 윈도우 사이즈 만큼 반복적으로 학습하도록 상기 배치를 구성하되, 반복적으로 구성된 배치 사이에 이전 배치의 데이터와 이후 배치의 데이터가 혼합되도록 슬라이딩 방식을 적용하여 상기 배치를 구성하는,

강건성 분석 장치.

청구항 8

제5항에 있어서,

상기 전처리부는 엔트로피가 낮은 순서대로 정렬된 상기 벤치마크 데이터 셋의 각 데이터를 상기 윈도우 사이즈 만큼 반복적으로 학습하도록 상기 배치를 구성하되, 반복적으로 구성된 배치 사이에 이전 배치의 데이터와 이후 배치의 데이터가 혼합되도록 슬라이딩 방식을 적용하여 배치를 구성하고, 이전 배치를 포함하여 이어서 학습하도록 상기 배치를 구성하는,

강건성 분석 장치.

청구항 9

제5항에 있어서,

상기 강건성 분석부는 상기 손실의 수렴 시간을 기초로 상기 벤치마크 데이터 셋의 강건성을 분석하는,

강건성 분석 장치.

발명의 설명

기술 분야

[0001] 본 발명은 벤치마크 데이터 셋의 안정성(stability) 및 강건성(robustness)을 분석하는 기법에 관한 것이다.

배경 기술

[0002] 벤치마크 데이터 셋(benchmark dataset)은 머신러닝 모델을 평가하기 위한 데이터 셋으로 학습 셋(train set)과 테스트 셋(test set)으로 구성되어 있다. 벤치마크 데이터 셋은 해당 데이터 셋을 사용하여 모델을 학습하고 성능을 측정할 수 있다. 연구자들은 공통된 데이터 셋에 대해 평가를 수행하기 때문에 객관적으로 모델의 성능을 측정 및 비교할 수 있다. 즉, 벤치마크 데이터 셋은 연구의 발전 정도에 대한 기준으로서의 역할을 하기 때문에 데이터가 안정적이고 강건해야 한다. 따라서 데이터 셋 및 벤치마크 데이터 셋을 구성 및 구축할 때 편향되지 않은 강건한 데이터를 만들기 위해 다양한 방법론들을 사용하여 보다 안정적인 데이터 셋을 구축하자 하는 시도들이 있었다.

[0003] 기존의 데이터 셋 및 벤치마크 데이터 셋은 구축 단계에서 안정성 및 강건성을 고려하여 구축한다. 구축 과정에서의 강건성 분석은 주로 회귀 분석과 같은 통계적 기법을 사용한다. 예를 들어 데이터 셋 안에서 라벨이 한쪽에 편향되지 않고 고르게 분포되어 있는 지 등을 분석하는 방식이다. 기존 연구 방법들은 대부분 구축 과정에서 통계적 기법에 집중되어 있고, 구축된 이후의 안정성 및 강건성에 대한 연구는 비교적 부족한 편이다. 즉, 연구들 간의 비교를 위해 벤치마크 데이터 셋을 사용하고 있으나 데이터 자체의 안정성 및 강건성에 대한 분석은 활발하지 않은 편이다.

[0004] 따라서, 본 발명에서는 이러한 문제점을 해결하기 위하여 벤치마크 데이터셋의 특징을 바탕으로 안정성 및 강건성을 모델링 관점에서 해결하고자 한다. 벤치마크 데이터셋은 안정적이어야 하기 때문에 쉽게 과적합(overfitting) 혹은 과소적합(underfitting)이 발생하면 안된다. 우리는 이 중 과적합 부분에 집중한다. 데이터 양이 적고 모델이 데이터를 학습하는 횟수가 증가할수록 과적합이 일어나기 쉽다는 조건을 이용하여 학습 방법에 변화를 주어 데이터 셋 안에서 엔트로피와 과적합의 시작 지점을 파악하여 이를 통해 벤치마크 데이터 셋의 안정성 및 강건성을 분석할 수 있는 모델을 만들고자 한다.

선행기술문헌

특허문헌

- [0005] (특허문헌 0001) 대한민국 공개특허 제2022-0092256호 (2022.07.01. 공개)
(특허문헌 0002) 대한민국 공개특허 제2018-0120061호 (2018.11.05. 공개)
(특허문헌 0003) 대한민국 공개특허 제2022-0009662호 (2022.01.25. 공개)

발명의 내용

해결하려는 과제

- [0007] 본 발명이 이루고자 하는 기술적인 과제는 손실(loss) 조정을 이용하여 벤치마크 데이터 셋의 강건성 및/또는 안정성을 분석하는 모델을 제공하는 것이다.

과제의 해결 수단

- [0008] 본 발명의 일 실시예에 따른 강건성 분석 장치는 사전학습 언어 모델(Pre-trained Language Model, PLM), 윈도우 사이즈(window size), 강건성 평가의 대상인 벤치마크 데이터 셋(benchmark dataset), 및 스코어 함수 중 적어도 하나에 대한 설정 동작을 수행하는 설정부, 상기 벤치마크 데이터 셋을 전처리하여 상기 배치(batch)를 구성하는 전처리부, 상기 전처리부에 의해 구성된 배치 구성에 따라 상기 PLM을 학습시키는 학습부, 및 상기 PLM에 대한 학습 과정에서 도출되는 손실(loss) 값에 기초하여 상기 벤치마크 데이터 셋의 강건성을 분석하는 강건성 분석부를 포함한다.

발명의 효과

- [0009] 본 발명은 기존의 데이터 셋의 안정성 및 강건성에 대한 분석이 데이터 셋 구축과정에서 주로 이루어져 있고 구축 후에 이미 존재하는 벤치마크 데이터 셋에 대한 연구가 부족하며 특히 모델링 관점의 분석이 부족하다는 문제점을 해결할 수 있을 것이라고 기대한다.
- [0010] 간단한 통계적 기법이 아닌 학습 과정에서 데이터 셋의 안정성 및 강건성을 분석할 수 있기 때문에 실제 연구 환경 및 실험 환경에서 분석하고자 하는 데이터 셋이 안정적인지 대한 인사이트를 제공할 수 있을 것으로 기대한다.

도면의 간단한 설명

- [0011] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.
- 도 1은 본 발명의 일 실시예에 따른 강건성 분석 장치의 기능 블록도이다.
- 도 2는 도 1에 도시된 강건성 분석 장치의 동작을 설명하기 위한 개념도이다.
- 도 3은 도 1에 도시된 강건성 분석 장치에 의해 수행되는 강건성 분석 방법을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0012] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.
- [0013] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.
- [0014] 제1 또는 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1 구성 요소는 제2 구성 요소로 명명

될 수 있고 유사하게 제2 구성 요소는 제1 구성 요소로도 명명될 수 있다.

- [0015] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등도 마찬가지로 해석되어야 한다.
- [0016] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0017] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0018] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시예들을 상세히 설명한다. 그러나, 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0019] 도 1은 본 발명의 일 실시예에 따른 강건성 분석 장치의 기능 블록도이고, 도 2는 도 1에 도시된 강건성 분석 장치의 동작을 설명하기 위한 개념도이다.
- [0020] 도 1과 도 2를 참조하면, 벤치마크 데이터 셋의 강건성 분석 장치(100)는 설정부(110), 전처리부(120), 학습부(130), 및 강건성 판단부(140)를 포함한다. 실시예에 따라, 강건성 분석 장치(100)는 저장부(150)를 더 포함할 수 있다. 강건성 분석 장치(100)는 적어도 프로세서(processor) 및/또는 메모리(memory)를 포함하는 컴퓨팅 장치로 구현될 수 있다. 컴퓨팅 장치는 PC(Personal Computer), 서버(server), 태블릿 PC, 노트북 등을 포함할 수 있다. 또한, 컴퓨팅 장치는 하나의 물리적 장치로 구현되거나 복수의 물리적 장치로 구현될 수 있다.
- [0021] 설정부(110)는 강건성 분석에 필요한 설정 동작을 수행할 수 있다. 구체적으로, 설정부(110)는 학습 모델(사전 학습 언어 모델(Pre-trained Language Model, PLM)을 의미할 수 있음), 학습 방법(또는 배치 구성 방법), 윈도우 사이즈(window size), 강건성 평가의 대상인 벤치마크 데이터 셋, 및 스코어 함수(scoring function) 중 적어도 하나를 결정(또는 설정)할 수 있다. 일 예로, 설정부(110)는 소정의 입출력 인터페이스를 통해 수신되는 사용자의 입력에 기초하여, 설정 동작을 수행할 수 있다.
- [0022] 본 발명에서는, 학습 모델로 Google의 언어 모델인 BERT를 사용한다. 그러나, 본 발명의 권리범위가 이에 제한되는 것은 아니며, 사용되는 딥러닝 모델(학습 모델)은 실시예에 따라 상이할 수 있다. 또한, 설정부(110)는 복수의 학습 모델들 중에서 어느 하나의 학습 모델을 선택(또는 설정)할 수도 있다.
- [0023] 학습 방법은 배치(batch) 구성 방법을 의미할 수 있다. 예시적으로, 본 발명에서는 3가지의 학습 방법(배치 구성 방법)을 제안한다. 따라서, 설정부(110)는 복수의 학습 방법들(배치 구성 방법들) 중에서 어느 하나의 학습 방법을 선택(또는 설정)할 수 있다.
- [0024] 스코어 함수는 제1 스코어 함수와 제2 스코어 함수를 포함할 수 있다. 제1 스코어 함수는 MLM 스코어(Masked Language Model Scoring) 함수로써, 벤치마크 데이터 셋을 이루는 각 데이터에 대한 MLM 스코어 값을 산출할 수 있다. 제1 스코어 함수에 대하여는 Julian Salazar 등의 논문(Julian Salazar, Davis Liang, Toan Q. Nguyen, and Katrin Kirchhoff, "Masked Language Model Scoring," Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 2699-2712, July 5-10, 2020)등이 참조될 수 있다. 제2 스코어 함수는 ML 스코어(Language Model based sentences scoring) 함수로써, 벤치마크 데이터 셋을 이루는 각 데이터에 대한 ML 스코어 값을 산출할 수 있다. 제2 스코어 함수에 대하여는 Alec Radford 등의 논문(Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever, "Language Models are

Unsupervised Multitask Learners") 등이 참조될 수 있다. 설정부(110)는 복수의 스코어 함수들 중에서 사용할 스코어 함수를 선택(또는 설정)할 수 있다.

[0025] 전처리부(120)는 벤치마크 데이터 셋을 이루는 각 데이터에 대한 전처리 동작을 수행할 수 있다. 여기서, 각 데이터는 하나 또는 두 개의 문장과 이에 대응하는 라벨(label)을 포함하는데, 전처리부(120)는 각 데이터에 대한 전처리 동작을 통하여, MLM 스코어 값, LM 스코어 값, 및 토큰 길이 중 적어도 하나를 산출할 수 있다.

[0026] 또한, 전처리부(120)는 벤치마크 데이터 셋의 각 데이터를 엔트로피가 낮은 순서대로 정렬한 후 배치를 구성하는 인덱스를 조절함으로써 학습 모델의 배치를 구성할 수 있다. 엔트로피가 낮은 순서라 함은 MLM 스코어 값의 내림차순, LM 스코어 값의 내림차순, 및 토큰 길이의 내림차순 중 어느 하나를 의미할 수 있다. 벤치마크 데이터 셋의 배치 구성에 관한 구체적인 설명은 후술하기로 한다.

[0027] 학습부(130)는 전처리부(120)에 의한 배치 구성에 따라 학습 모델에 대한 학습을 진행한다. 이때, 학습부(130)는 학습이 진행됨에 따른 손실(loss) 값을 산출할 수 있다. 학습은 사전학습 언어 모델을 사용하여 미세조정(fine-tuning)하는 방식으로 이루어진다.

[0028] 강건성 판단부(140)는 학습부(130)에 의해 산출된 손실 값에 기초하여 벤치마크 데이터 셋에 대한 강건성 또는 강건성을 판단하기 위한 지표 값을 산출할 수 있다. 일 실시예에 따르면, 강건성 판단부(140)는 학습 모델의 과적합(overfitting) 여부를 판단하고, 과적합 시점을 기초로 강건성을 판단할 수 있다. 예컨대, 과적합 시점 자체를 강건성에 대한 지표 값으로 산출하거나, 과적합 시점에 소정의 가중치를 부여함으로써 강건성에 대한 지표 값을 산출하거나, 과적합 시점을 미리 정해진 기준과 비교하여 강건성에 대한 정도를 판단할 수 있다. 여기서 과적합 시점은 손실 값이 특정 값으로 수렴하는 시점을 의미할 수 있다. 수렴 시점은 손실의 변화율이 미리 정해진 기준치 미만인 시점 등을 의미할 수 있으나, 본 발명이 이에 제한되는 것은 아니며 판단의 방법은 다양할 수 있다.

[0029] 저장부(150)에는 강건성 분석 장치(100)의 동작을 위한 OS(Operating System, 운영체제), 프로그램, 어플리케이션, 앱, 소스 코드 등이 저장되어 있을 수 있다. 또한, 저장부(150)에는 설정부(110)에 의한 설정 결과, 전처리부(120)에 의한 전처리 결과, 학습 모델, 벤치마크 데이터 셋, 학습부(130)에 의한 학습 결과, 강건성 판단부(140)에 의한 강건성 판단의 결과 등이 저장될 수 있다.

[0030] 이하에서는, 학습의 방법에 해당하는 배치 구성 방법을 설명한다.

[0031] 본 발명에서는 제안하는 3가지의 방법 중 제1 방법은 정렬된 데이터의 배치를 구성할 때 앞에서부터 설정한 윈도우 사이즈(예컨대, window_size=2)에 해당하는 횟수만큼 반복하여 구성하는 방법이다. 구체적인 예로, x (x 는 임의의 자연수)번째 배치의 구성은 " $11 + 22 + 33 + \dots + xx$ "와 같은 형식으로 학습을 수행한다.

[0032] 제2 방법은 윈도우 사이즈만큼 배치를 이동하며 학습하는데 지나간 배치는 여러번 학습하지 않는 방법이다. 구체적인 예로, " $11 + 12 + 22 + 23 + \dots + (x-1)x + xx$ "와 같은 배치 구성으로 학습을 수행하는 방법이다. 다시 말하면, 엔트로피가 낮은 순서대로 정렬된 벤치마크 데이터 셋의 각 데이터를 윈도우 사이즈만큼 반복적으로 학습하도록 배치를 구성하되, 반복적으로 구성된 배치 사이에 이전 배치의 데이터와 이후 배치의 데이터가 혼합되도록 슬라이딩 방식을 적용하여 배치를 구성할 수 있다.

[0033] 제3 방법은 윈도우 사이즈만큼 배치를 학습하되, 이전 배치를 이어서 학습하는 방식이다. 구체적인 예로, " $(11) + (11+12) + (11+12+22) + \dots + (11+12+12+\dots+xx)$ "와 같은 배치 구성으로 학습을 수행하는 방법이다. 다시 말하면, 엔트로피가 낮은 순서대로 정렬된 벤치마크 데이터 셋의 각 데이터를 윈도우 사이즈만큼 반복적으로 학습하도록 배치를 구성하되, 반복적으로 구성된 배치 사이에 이전 배치의 데이터와 이후 배치의 데이터가 혼합되도록 슬라이딩 방식을 적용하여 배치를 구성하고, 이전 배치를 포함하여 이어서 학습하도록 배치를 구성할 수 있다.

[0034] 이와 같이, 배치를 구성하는 인덱스를 조절하여 학습 모델이 데이터를 보는 횟수를 조절함으로써, 학습 모델이 기존의 일반적인 학습 방법 보다 빠르게 과적합에 이르게 할 수 있다. 즉, 데이터의 안정성 및 강건성을 분석하기 위한 기준 중 하나인 과적합이 시작되는 지점을 빠르게 볼 수 있게 된다. 즉, 본 발명에 의하면, 학습 모델은 벤치마크 데이터 셋이 입력으로 주어졌을 때, 기준과 달리 한 에포크(epoch) 당 데이터를 한 번 보는 것이 아닌 데이터 학습 횟수에 변형을 주어 학습함으로써 과적합이 빠르게 발생할 수 있도록 한다. 이를 하는 이유는 벤치마크 데이터 셋은 과적합이 잘 일어나지 않는 안정성 있는 데이터여야 객관적인 성능 평가에 도움이 되기 때문이다.

[0035] 도 3은 도 1에 도시된 강건성 분석 장치에 의해 수행되는 강건성 분석 방법을 설명하기 위한 흐름도이다. 도 3

의 강건성 분석 방법의 단계들 중 적어도 일부는 도 1의 강건성 분석 장치에 포함되는 프로세서의 동작으로 이해될 수도 있다. 이하에서, 강건성 분석 방법을 설명함에 있어, 앞선 기재와 중복되는 내용에 관하여는 그 구체적인 설명을 생략하기로 한다.

- [0036] 도 1 내지 도 3을 참조하면, 강건성 분석 장치(100)의 설정부(110)에 의한 설정 동작이 수행된다(S110). 설정 동작에서는, 학습 모델, 학습 방법(또는 배치 구성 방법), 윈도우 사이즈, 벤치마크 데이터 셋, 및 스코어 함수 중 적어도 하나가 설정될 수 있다.
- [0037] 설정된(또는 선택된) 벤치마크 데이터 셋의 각 데이터에 대한 전처리 동작이 수행된다(S120). 전처리 동작의 결과로, 학습을 위한 배치 구성이 완료된다. 배치 구성은 상술한 3 가지의 방법 중 어느 하나의 방법을 통해 구성될 수 있다.
- [0038] 설정된 배치 구성에 따라 학습 모델에 대한 학습(또는 미세조정)이 수행된다(S130). 학습 모델에 대한 학습이 수행되는 도중에 시간에 따른 손실 값이 도출될 수 있다.
- [0039] 도출된 손실 값에 기초하여 벤치마크 데이터 셋에 대한 강건성이 판단될 수 있다(S140). 실시예에 따르면, 손실 값에 기초하여 벤치마크 데이터 셋에 대한 강건성을 판단하거나 강건성을 판단하기 위한 지표 값을 산출할 수 있다.
- [0040] 이상에서 설명된 장치는 하드웨어 구성 요소, 소프트웨어 구성 요소, 및/또는 하드웨어 구성 요소 및 소프트웨어 구성 요소의 집합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성 요소는, 예를 들어, 프로세서, 컨트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크로컴퓨터, FPA(Field Programmable array), PLU(Programmable Logic Unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(Operation System, OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술 분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(Processing Element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(Parallel Processor)와 같은, 다른 처리 구성(Processing Configuration)도 가능하다.
- [0041] 소프트웨어는 컴퓨터 프로그램(Computer Program), 코드(Code), 명령(Instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(Collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성 요소(Component), 물리적 장치, 가상 장치(Virtual Equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(Signal Wave)에 영구적으로, 또는 일시적으로 구체화(Embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0042] 실시예에 따른 방법은 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. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(Magnetic Media), CD-ROM, DVD와 같은 광기록 매체(Optical Media), 플롭티컬 디스크(Floptical Disk)와 같은 자기-광 매체(Magneto-optical Media), 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0043] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 예를 들어,

설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성 요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성 요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

[0044]

100 : 강건성 판단 장치

110 : 설정부

120 : 전처리부

130 : 학습부

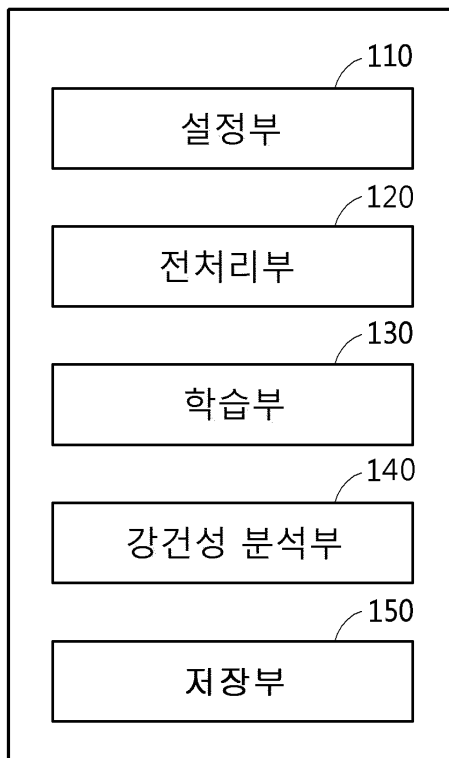
140 : 강건성 분석부

150 : 저장부

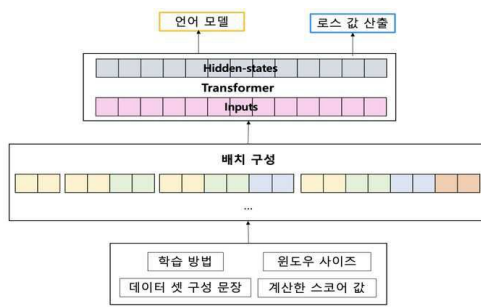
도면

도면1

100



도면2



도면3

