

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2023-0080824
(43) 공개일자 2023년06월07일

- (51) 국제특허분류(Int. Cl.)
G06Q 50/20 (2012.01) G06F 16/332 (2019.01)
G06N 3/08 (2023.01)
- (52) CPC특허분류
G06Q 50/20 (2013.01)
G06F 16/3329 (2019.01)
- (21) 출원번호 10-2021-0168322
(22) 출원일자 2021년11월30일
심사청구일자 2021년11월30일
- (71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
- (72) 발명자
박성배
서울특별시 송파구 오금로 396, 1402호 (가락동, 에스케이허브파크)
- 김소연
경기도 용인시 기흥구 서천동로21번길 6-1, 301호 (서천동)
(뒷면에 계속)
- (74) 대리인
유미특허법인

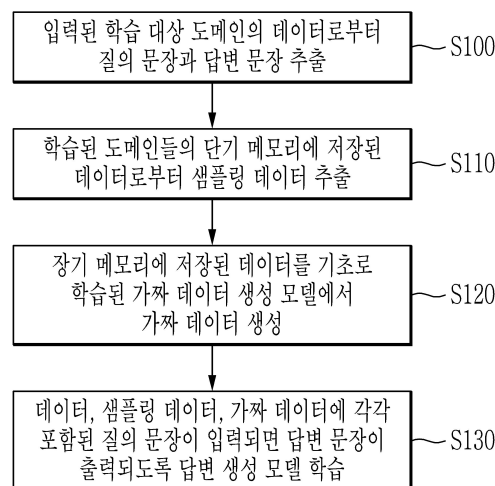
전체 청구항 수 : 총 12 항

(54) 발명의 명칭 연속 학습 시스템 및 이를 이용한 연속 학습 방법

(57) 요약

연속 학습 시스템이 답변 생성 모델을 연속 학습하는 방법으로서, 학습 대상 도메인의 복수의 데이터들로부터 각각 질의 문장과 답변 문장을 추출하고, 학습 대상 도메인에 앞서 답변 생성 모델을 학습한 후 단기 메모리에 저장된 학습된 도메인들 각각의 복수의 데이터들로부터 복수의 샘플링 데이터를 추출하고, 각 샘플링 데이터에서 샘플 질의 문장과 샘플 답변 문장을 추출한다. 그리고 단기 메모리에 저장된 학습된 도메인의 복수의 데이터들 중 일부 데이터이며 장기 메모리에 저장된 데이터들을 이용하여 생성한 가짜 데이터로부터 가짜 질의 문장과 가짜 답변 문장을 추출하고, 질의 문장, 샘플 질의 문장 및 가짜 질의 문장을 입력하면, 답변 문장, 샘플 답변 문장 및 가짜 답변 문장을 출력하도록 답변 생성 모델을 학습시킨다.

대표도 - 도3



(52) CPC특허분류

G06N 3/08 (2023.01)

(72) 발명자

이준범

경기도 용인시 기흥구 서천동로21번길 6-1, 301호
(서천동)

임연수

서울특별시 강서구 강서로56나길 110, 104동 504호
(등촌동, 부영아파트)

김봉민

경기도 용인시 기흥구 서그대로53번길 3-17, 304호
(서천동, 서천빌리지)

이 발명을 지원한 국가연구개발사업

과제고유번호

1711054934

과제번호

2013-0-00109

부처명

과학기술정보통신부

과제관리(전문)기관명

정보통신기획평가원

연구사업명

혁신성장동력프로젝트(R&D)(과기정통부)(13111015015511001158411)

연구과제명

(엑소브레인-2세부) WiseKB: 빅데이터 이해 기반 자가학습형 지식베이스 및 추론 기

술 개발

기 여 율

1/1

과제수행기관명

주식회사 솔트룩스

연구기간

2020.01.01~2023.02.28

명세서

청구범위

청구항 1

연속 학습 시스템이 답변 생성 모델을 연속 학습하는 방법으로서,

학습 대상 도메인의 복수의 데이터들로부터 각각 질의 문장과 답변 문장을 추출하는 단계,

상기 학습 대상 도메인에 앞서 상기 답변 생성 모델을 학습한 후 단기 메모리에 저장된 학습된 도메인들 각각의 복수의 데이터들로부터 복수의 샘플링 데이터를 추출하고, 각 샘플링 데이터에서 샘플 질의 문장과 샘플 답변 문장을 추출하는 단계,

상기 단기 메모리에 저장된 학습된 도메인의 복수의 데이터들 중 일부 데이터이며 장기 메모리에 저장된 데이터들을 이용하여 생성한 가짜 데이터로부터 가짜 질의 문장과 가짜 답변 문장을 추출하는 단계, 그리고

상기 질의 문장, 샘플 질의 문장 및 가짜 질의 문장을 입력하면, 상기 답변 문장, 상기 샘플 답변 문장 및 상기 가짜 답변 문장을 출력하도록 상기 답변 생성 모델을 학습시키는 단계

를 포함하는, 연속 학습 방법.

청구항 2

제1항에 있어서,

상기 단기 메모리에 저장된 각 데이터는, 각 데이터가 매핑되어 있는 도메인의 도메인 정보가 삽입된 도메인 레이블, 상기 답변 생성 모델을 학습시킨 횟수에 대한 반복 횟수 레이블, 그리고 상기 단기 메모리에 저장된 저장 시점인 시간 레이블을 포함하는, 연속 학습 방법.

청구항 3

제2항에 있어서,

상기 샘플 답변 문장을 추출하는 단계는,

상기 학습 대상 도메인의 각 데이터와 상기 단기 메모리에 저장된 각 데이터의 유사도를 계산하는 단계, 그리고

상기 단기 메모리에 저장된 데이터들 중 계산한 유사도가 미리 설정한 임계값 이상인 데이터의 상기 반복 횟수 레이블에 삽입된 횟수를 하나 증가시키는 단계

를 포함하는, 연속 학습 방법.

청구항 4

제3항에 있어서,

상기 샘플 답변 문장을 추출하는 단계는,

상기 단기 메모리에 저장된 각 데이터의 반복 횟수 레이블을 기초로, 각 데이터의 횟수가 미리 설정된 임계값 이상인 적어도 하나의 데이터를 확인하는 단계, 그리고

상기 임계값 이상의 횟수로 반복 횟수 레이블이 삽입된 데이터는 상기 단기 메모리에서 상기 장기 메모리로 이동시키는 단계

를 포함하는, 연속 학습 방법.

청구항 5

제2항에 있어서,

상기 샘플 답변 문장을 추출하는 단계는,

상기 단기 메모리에 저장된 각 데이터의 시간 레이블을 확인하는 단계, 그리고

상기 시간 레이블에 삽입된 시간이 미리 설정한 임계 시간을 초과한 데이터는 상기 단기 메모리에서 삭제하는 단계

를 포함하는, 연속 학습 방법.

청구항 6

제1항에 있어서,

상기 가짜 답변 문장을 추출하는 단계는,

상기 장기 메모리에 저장된 데이터들을 이용하여 가짜 데이터 생성 모델을 학습시키는 단계

를 더 포함하는, 연속 학습 방법.

청구항 7

제6항에 있어서,

상기 가짜 데이터 생성 모델을 학습시키는 단계 이후에,

상기 가짜 데이터 생성 모델을 학습시킨 데이터들을 상기 장기 메모리로부터 삭제하는 단계

를 포함하는, 연속 학습 방법.

청구항 8

질의 문장이 입력되면 상기 질의 문장에 대한 답변 문장을 생성하는 답변 생성 모델을 학습시키는 연속 학습 시스템으로서,

적어도 하나의 명령어를 포함하고, 상기 답변 생성 모델을 학습시킨 학습된 도메인들의 복수의 데이터들을 저장하는 단기 메모리,

상기 단기 메모리에 저장된 복수의 데이터들 중 상기 답변 생성 모델을 미리 설정한 횟수 이상 학습시킨 일부 데이터들을 저장되는 장기 메모리, 그리고

프로세서

를 포함하고,

상기 프로세서는,

학습 대상 도메인의 데이터들로부터 각각 각각 질의 문장과 답변 문장을 추출하고, 상기 단기 메모리에 저장된 복수의 데이터들로부터 샘플링 데이터를 추출하여 샘플 질의 문장과 샘플 답변 문장을 추출하며, 상기 장기 메모리에 저장된 일부 데이터들을 기초로 생성된 가짜 데이터로부터 가짜 질의 문장과 가짜 답변 문장을 추출하고, 상기 질의 문장, 샘플 질의 문장, 가짜 질의 문장이 입력되면 상기 답변 문장, 샘플 답변 문장, 가짜 답변 문장이 출력되도록 상기 답변 생성 모델을 학습시키는, 연속 학습 시스템.

청구항 9

제8항에 있어서,

상기 프로세서는,

상기 학습 대상 도메인의 각 데이터와 상기 단기 메모리에 저장된 각 데이터의 유사도를 계산하고, 상기 계산한 유사도가 미리 설정한 임계값 이상인 데이터의 반복 횟수 레이블에 삽입된 횟수를 하나 증가시키는, 연속 학습 시스템.

청구항 10

제9항에 있어서,

상기 프로세서는,

상기 단기 메모리에 저장된 각 데이터의 반복 횟수 레이블에 삽입된 횟수를 확인하고, 횟수가 미리 설정한 임계 값 이상인 적어도 하나의 데이터를 상기 장기 메모리로 이동시키는, 연속 학습 시스템.

청구항 11

제10항에 있어서,

상기 프로세서는,

상기 단기 메모리에 저장된 각 데이터의 시간 레이블을 확인하고, 상기 시간 레이블에 삽입된 시간이 임계 시간을 초과한 데이터는 상기 단기 메모리에서 삭제하는, 연속 학습 시스템.

청구항 12

제11항에 있어서,

상기 프로세서는,

상기 장기 메모리에 저장된 데이터들을 이용하여 가짜 데이터 생성 모델을 학습시키는, 연속 학습 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 다중 도메인 답변 생성 모델에서 발생하는 파괴적 망각 문제를 해결하기 위한 연속 학습 시스템 및 이를 이용한 연속 학습 방법에 관한 것이다.

배경 기술

[0002] 답변 생성기는 다양한 도메인의 데이터로 답변 생성 모듈을 학습시키고, 어느 하나의 도메인에 대한 질의가 입력되면, 질의에 대한 답변을 생성하여 제공한다. 이러한 종래의 답변 생성기를 이용한 답변 생성 방식은, 새로운 도메인이 추가될 때마다 기존에 답변 생성 모델을 학습시킨 도메인의 데이터와 새로운 도메인에 대한 데이터를 이용하여 답변 생성 모델을 학습시켜야 한다.

[0003] 이와 같은 비효율적 학습을 회피하면서 다중 도메인에 대한 질의가 입력되더라도 답변 생성기가 답변을 생성할 수 있도록, 복수의 도메인들의 데이터로 답변 생성 모델을 순차적으로 학습시키는 접근 방식들이 연구되고 있다. 순차적으로 학습시키는 방식 중에는 메모리-리플레이 기반의 연속 학습 기법이 있다.

[0004] 답변 생성 모델을 기 학습시킨 기존 도메인에 대한 답변 성능이 저하되는 파괴적 망각 현상을 완화하기 위하여, 메모리-리플레이 기반의 연속 학습 기법은 답변 생성 모델을 새로운 도메인으로 학습시킬 때 기존 도메인의 데이터의 일부를 이용하여 같이 재학습시킨다. 또한, 데이터 생성기가 만든 기존 도메인의 데이터와 유사한 데이터로 함께 답변 생성 모델을 학습시키기도 한다.

[0005] 이때, 답변 생성 모델을 새로운 도메인으로 학습시키면서 기존 도메인의 데이터의 일부를 이용하여 재학습시키기 위해서는, 기존 도메인의 데이터의 일부를 저장하기 위한 메모리 모듈이 구비되어야 한다. 따라서, 답변 생성 모델을 학습시킨 도메인의 수에 비례하여 메모리 모듈의 크기가 계속 증가하는 문제점이 발생한다.

[0006] 또한, 데이터 생성기가 만든 가짜 데이터로 답변 생성 모델을 학습시키는 방법을 이용할 경우, 생성된 가짜 데이터의 품질에 따라 답변 생성 모델의 성능이 좌우되는 문제점이 발생한다.

발명의 내용

해결하려는 과제

[0007] 따라서, 본 발명은 다중 도메인 답변 생성기의 학습을 위한 인간의 기억 시스템을 모방하는 연속 학습 시스템을 이용하여, 답변 생성기가 여러 도메인을 학습할 때 기존 도메인에 대한 성능이 저하되는 파괴적 망각 현상을 줄일 수 있는 연속 학습 시스템 및 이를 이용한 연속 학습 방법 제공한다.

과제의 해결 수단

[0008] 상기 본 발명의 기술적 과제를 달성하기 위한 본 발명의 하나의 특징인 연속 학습 시스템이 답변 생성 모델을

연속 학습하는 방법으로서,

- [0009] 학습 대상 도메인의 복수의 데이터들로부터 각각 질의 문장과 답변 문장을 추출하는 단계, 상기 학습 대상 도메인에 앞서 상기 답변 생성 모델을 학습한 후 단기 메모리에 저장된 학습된 도메인들 각각의 복수의 데이터들로부터 복수의 샘플링 데이터를 추출하고, 각 샘플링 데이터에서 샘플 질의 문장과 샘플 답변 문장을 추출하는 단계, 상기 단기 메모리에 저장된 학습된 도메인의 복수의 데이터들 중 일부 데이터이며 장기 메모리에 저장된 데이터들을 이용하여 생성한 가짜 데이터로부터 가짜 질의 문장과 가짜 답변 문장을 추출하는 단계, 그리고 상기 질의 문장, 샘플 질의 문장 및 가짜 질의 문장을 입력하면, 상기 답변 문장, 상기 샘플 답변 문장 및 상기 가짜 답변 문장을 출력하도록 상기 답변 생성 모델을 학습시키는 단계를 포함한다.
- [0010] 상기 단기 메모리에 저장된 각 데이터는, 각 데이터가 매핑되어 있는 도메인의 도메인 정보가 삽입된 도메인 레이블, 상기 답변 생성 모델을 학습시킨 횟수에 대한 반복 횟수 레이블, 그리고 상기 단기 메모리에 저장된 저장 시점인 시간 레이블을 포함할 수 있다.
- [0011] 상기 샘플 답변 문장을 추출하는 단계는, 상기 학습 대상 도메인의 각 데이터와 상기 단기 메모리에 저장된 각 데이터의 유사도를 계산하는 단계, 그리고 상기 단기 메모리에 저장된 데이터들 중 계산한 유사도가 미리 설정한 임계값 이상인 데이터의 상기 반복 횟수 레이블에 삽입된 횟수를 하나 증가시키는 단계를 포함할 수 있다.
- [0012] 상기 샘플 답변 문장을 추출하는 단계는, 상기 단기 메모리에 저장된 각 데이터의 반복 횟수 레이블을 기초로, 각 데이터의 횟수가 미리 설정된 임계값 이상인 적어도 하나의 데이터를 확인하는 단계, 그리고 상기 임계값 이상의 횟수로 반복 횟수 레이블이 삽입된 데이터는 상기 단기 메모리에서 상기 장기 메모리로 이동시키는 단계를 포함할 수 있다.
- [0013] 상기 샘플 답변 문장을 추출하는 단계는, 상기 단기 메모리에 저장된 각 데이터의 시간 레이블을 확인하는 단계, 그리고 상기 시간 레이블에 삽입된 시간이 미리 설정한 임계 시간을 초과한 데이터는 상기 단기 메모리에서 삭제하는 단계를 포함할 수 있다.
- [0014] 상기 가짜 답변 문장을 추출하는 단계는, 상기 장기 메모리에 저장된 데이터들을 이용하여 가짜 데이터 생성 모델을 학습시키는 단계를 더 포함할 수 있다.
- [0015] 상기 가짜 데이터 생성 모델을 학습시키는 단계 이후에, 상기 가짜 데이터 생성 모델을 학습시킨 데이터들을 상기 장기 메모리로부터 삭제하는 단계를 포함할 수 있다.
- [0016] 상기 본 발명의 기술적 과제를 달성하기 위한 본 발명의 또 다른 특징인 질의 문장이 입력되면 상기 질의 문장에 대한 답변 문장을 생성하는 답변 생성 모델을 학습시키는 연속 학습 시스템으로서,
- [0017] 적어도 하나의 명령어를 포함하고, 상기 답변 생성 모델을 학습시킨 학습된 도메인들의 복수의 데이터들을 저장하는 단기 메모리, 상기 단기 메모리에 저장된 복수의 데이터들 중 상기 답변 생성 모델을 미리 설정한 횟수 이상 학습시킨 일부 데이터들을 저장되는 장기 메모리, 그리고 프로세서를 포함하고, 상기 프로세서는, 학습 대상 도메인의 데이터들로부터 각각 각각 질의 문장과 답변 문장을 추출하고, 상기 단기 메모리에 저장된 복수의 데이터들로부터 샘플링 데이터를 추출하여 샘플 질의 문장과 샘플 답변 문장을 추출하며, 상기 장기 메모리에 저장된 일부 데이터들을 기초로 생성된 가짜 데이터로부터 가짜 질의 문장과 가짜 답변 문장을 추출하고, 상기 질의 문장, 샘플 질의 문장, 가짜 질의 문장이 입력되면 상기 답변 문장, 샘플 답변 문장, 가짜 답변 문장이 출력되도록 상기 답변 생성 모델을 학습시킨다.
- [0018] 상기 프로세서는, 상기 학습 대상 도메인의 각 데이터와 상기 단기 메모리에 저장된 각 데이터의 유사도를 계산하고, 상기 계산한 유사도가 미리 설정한 임계값 이상인 데이터의 반복 횟수 레이블에 삽입된 횟수를 하나 증가시킬 수 있다.
- [0019] 상기 프로세서는, 상기 단기 메모리에 저장된 각 데이터의 반복 횟수 레이블에 삽입된 횟수를 확인하고, 횟수가 미리 설정한 임계값 이상인 적어도 하나의 데이터를 상기 장기 메모리로 이동시킬 수 있다.
- [0020] 상기 프로세서는, 상기 단기 메모리에 저장된 각 데이터의 시간 레이블을 확인하고, 상기 시간 레이블에 삽입된 시간이 임계 시간을 초과한 데이터는 상기 단기 메모리에서 삭제할 수 있다.
- [0021] 상기 프로세서는, 상기 장기 메모리에 저장된 데이터들을 이용하여 가짜 데이터 생성 모델을 학습시킬 수 있다.

발명의 효과

[0022] 본 발명에 따르면, 인간의 기억 시스템을 모방하여 답변 생성 모델이 학습한 도메인의 수에 비례하여 메모리 모듈의 크기가 계속하여 증가하는 문제점을 해결할 수 있다.

[0023] 또한, 장기 메모리에 저장된 학습된 도메인의 데이터들을 기초로 가짜 데이터를 생성하므로, 가짜 데이터의 품질에 따라 모델의 성능이 좌우되는 문제점을 회피할 수 있다.

도면의 간단한 설명

[0024] 도 1은 본 발명의 실시예에 따른 연속 학습 시스템의 구조도이다.

도 2는 본 발명의 실시예에 따른 가짜 데이터 생성기의 예시도이다.

도 3은 본 발명의 실시예에 따른 연속 학습 방법에 대한 흐름도이다.

도 4는 본 발명의 실시예에 따른 메모리 관리 방법에 대한 흐름도이다.

도 5는 본 발명의 실시예에 따른 컴퓨팅 장치의 구조도이다.

발명을 실시하기 위한 구체적인 내용

[0025] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0026] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[0027] 이하, 도면을 참조로 하여 본 발명의 실시예에 따른 연속 학습 시스템 및 이를 이용한 연속 학습 방법에 대해 상세히 설명한다.

[0028] 도 1은 본 발명의 실시예에 따른 연속 학습 시스템의 구조도이다.

[0029] 도 1에 도시된 바와 같이, 연속 학습 시스템(100)은 답변 생성기(110), 단기 메모리(120), 장기 메모리(130), 가짜 데이터 생성기(140), 그리고 제어기(150)를 포함한다.

[0030] 답변 생성기(110)는 질의가 포함된 발화 음성이 입력되면, 발화 음성으로부터 질의를 추출하고, 추출한 질의에 대한 답변을 생성하여 제공한다. 이를 위해, 답변 생성기(110)는 질의가 입력되면 해당 질의에 대한 답변을 생성하는 답변 생성 모델(도면 미도시)을 포함한다.

[0031] 여기서, 답변 생성 모델이 질의에 대한 답변을 생성할 수 있도록, 답변 생성기(110)는 도메인 데이터를 이용하여 답변 생성 모델을 학습시킨다. 즉, 답변 생성기(110)는 학습시킬 도메인(이하, '학습 대상 도메인'이라 지칭함)의 질의에 해당하는 문장(이하, '질의 문장'이라 지칭함)과 질의에 대한 답변(이하, '답변 문장'이라 지칭함)으로 구성된 데이터가 입력되면, 질의 문장이 답변 생성 모델로 입력하면 답변 문장이 출력되도록 답변 생성 모델을 학습시킨다.

[0032] 이때, 답변 생성기(110)는 샘플링 데이터를 단기 메모리(120)로부터 수신하여, 학습 대상 도메인으로 답변 생성 모델을 학습시킬 때 함께 이용한다. 또한, 답변 생성기(110)는 이후 설명할 가짜 데이터 생성기(140)에서 생성한 가짜 데이터도 수신하여, 학습 대상 도메인으로 답변 생성 모델을 학습시킬 때 함께 이용한다.

[0033] 여기서, 샘플링 데이터는, 학습 대상 도메인으로 답변 생성 모델을 학습시키기 앞서, 답변 생성 모델을 학습시킨 도메인(이하, '학습된 도메인'이라 지칭함)의 데이터 중 일부 데이터로, 단기 메모리(120)에 저장된 학습된 도메인의 데이터들에서 일정 수 샘플링한 데이터를 의미한다. 그리고, 가짜 데이터는, 단기 메모리(120)에서 장기 메모리(130)로 이동된 일부 학습된 도메인의 데이터들을 기초로 생성된 가짜 질의 문장-응답 문장으로 구성된 데이터를 의미한다. 샘플링 데이터와 가짜 데이터는 이후 상세히 설명한다.

[0034] 답변 생성기(110)는 학습 대상 도메인의 데이터로 답변 생성 모델을 학습시킨 후, 학습 대상 도메인의 데이터를 단기 메모리(120)에 저장한다.

[0035] 단기 메모리(120)는 답변 생성기(110)가 답변 생성 모델을 학습시킨 학습 대상 도메인의 데이터를 저장한다. 이

때 데이터는 질의 문장과 답변 문장을 이어 붙인 하나의 문장으로 구성된다.

- [0036] 그리고, 데이터는 질의 문장과 답변 문장을 구분하기 위해, 질의 문장과 답변 문장 사이에 구분자를 포함한다. 구분자는 다양한 문자, 기호, 특수기호 등이 사용될 수 있으며, 본 발명의 실시예에서는 어느 하나의 형태로 한정하지 않는다.
- [0037] 또한, 데이터는 질의 문장이 어떤 도메인에서의 질의인지 나타내기 위해, 도메인 레이블을 포함한다. 또한, 데이터는 단기 메모리(120)에 저장된 이후 답변 생성기(110)가 새로 몇 개의 도메인을 학습했는지를 파악할 수 있도록 시간 레이블을 포함한다. 이를 위해, 답변 생성 모델을 학습시킨 후 답변 생성기(110)가 데이터를 단기 메모리(120)에 저장할 때, 저장 시점을 시간 레이블로 포함시키는 것을 예로 하여 설명한다.
- [0038] 데이터는 반복 횟수 레이블을 포함한다. 반복 횟수 레이블은 단기 메모리(120)에 저장된 후 답변 생성기(110)가 답변 생성 모델을 새로 학습시키는 새로운 도메인의 데이터 중에서 학습된 도메인의 데이터와의 유사도가 임계값 이상인 데이터가 존재하는 도메인이 몇 회 등장했는지를 나타내는 정보이다. 여기서 유사도는 제어기(150)에서 계산하며, 두 문장의 유사도를 계산하는 방법은 다양한 방법으로 실행될 수 있으므로 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다.
- [0039] 또한, 단기 메모리(120)는 제어기(150)의 제어에 따라 저장된 데이터를 삭제하거나, 일부 데이터는 장기 메모리(130)로 전달한다.
- [0040] 장기 메모리(130)는 단기 메모리(120)로부터 학습된 도메인의 데이터를 전달받아 저장한다. 이때, 인간의 장기 기억에 저장된 정보의 저장 시간이 아주 길다는 것에서 착안하되, 답변 생성 모델이 학습한 도메인의 수에 비례하여 메모리의 크기가 계속하여 증가하는 문제점을 해결하기 위하여, 본 발명의 실시예에서는 장기 메모리(130)에 학습된 도메인의 데이터를 계속 저장하지 않는다.
- [0041] 즉, 장기 메모리(130)의 데이터와 유사한 데이터를 생성할 수 있는 가짜 데이터 생성기(140)를 장기 메모리(130)에 저장된 데이터로 학습시킨 뒤, 장기 메모리의 데이터를 삭제한다. 이를 통해, 장기 메모리(130)의 데이터에 대한 정보는 가짜 데이터 생성기(140)가 유지하되, 메모리의 크기가 계속하여 증가하지 않도록 한다.
- [0042] 가짜 데이터 생성기(140)는 장기 메모리(130)에 저장된 데이터로 데이터 생성 모델을 학습시켜 가짜 질의 문장-가짜 응답 문장으로 구성된 가짜 데이터를 생성한다. 가짜 데이터 생성기(140)는 생성한 가짜 데이터를 답변 생성기(110)로 제공하여, 답변 생성기(110)가 학습 대상 도메인의 데이터로 답변 생성 모델을 학습시킬 때 함께 사용하도록 제공한다.
- [0043] 이때, 가짜 데이터 생성기(140)는 장기 메모리(130)에 저장되어 있는 학습된 도메인의 데이터를 기초로 가짜 데이터를 생성하도록 가짜 데이터 생성 모델을 학습시키기 때문에, 학습된 데이터와 유사한 가짜 데이터를 생성할 수 있다.
- [0044] 제어기(140)는 단기 메모리(120) 및 장기 메모리(130)에 저장된 데이터들의 반복횟수 레이블과 시간 레이블을 바탕으로 단기 메모리(120)에서 삭제할 데이터(이하, '삭제 대상 데이터'로 지칭함)와 장기 메모리(130)로 전달할 데이터(이하, '저장 대상 데이터'라 지칭함)를 결정한다.
- [0045] 즉, 인간의 단기 기억에서 수 분이 지나면 정보가 사라지는 것과 유사하게, 데이터의 시간 레이블이 미리 설정한 임계값을 초과할 경우, 단기 메모리(120)는 임계값이 초과된 데이터를 삭제 대상 데이터로 설정한다. 그리고 제어기(150)는 설정한 삭제 대상 데이터의 도메인 레이블과 질의 문장-답변 문장을 포함하여 단기 메모리(120)에서 해당 데이터가 삭제되도록 제어 신호를 생성한다.
- [0046] 또한, 인간의 단기 기억에 유사한 정보가 반복되는 경우 해당 정보의 저장 시간이 길어지는 것과 유사하게, 제어기(150)는 데이터의 반복횟수 레이블의 값에 따라 시간 레이블의 임계값을 증가시킨다. 또한, 인간의 단기 기억에 장기간 저장된 정보가 장기 기억으로 전달되는 것과 유사하게, 제어기(140)는 데이터의 반복 횟수 레이블이 미리 설정한 임계값을 초과할 경우, 해당 데이터를 장기 메모리(130)로 전달하도록 제어 신호를 생성한다.
- [0047] 상술한 연속 학습 시스템(100)의 구조에서 장기 메모리(130)에 저장된 데이터로 가짜 데이터를 생성하는 가짜 데이터 생성기(140)에 대해 도 2를 참조로 설명한다.
- [0048] 도 2는 본 발명의 실시예에 따른 가짜 데이터 생성기의 예시도이다.
- [0049] 가짜 데이터 생성기(140)는 가짜 데이터를 생성하는 사전 학습된 언어 모델(GPT-2)을 미세 조정(fine-tuning)하여 사용한다. 여기서, 언어 모델(GPT2)은 문장 생성 시작 토큰(GEN)이 입력되면, 다음에 등장할 단어의 확률을

계산하고, 확률이 가장 높은 단어를 선택하는 방법으로 문장을 생성하도록 사전 학습되어 있는 모델이다. 언어 모델(GPT2)을 사전 학습시키기 위한 데이터 입력은 GEN + sentence + EOS로, 언어 모델(GPT2)은 문장의 종료를 나타내는 EOS 토큰이 등장할 때까지 한 개의 문장을 생성한다.

- [0050] 가짜 데이터 생성기(140)는 이와 같이 사전 학습된 언어 모델(GPT2)이 한 개의 문장이 아닌 '질의 문장-답변 문장'과 같이 두 개의 문장을 한 번에 생성할 수 있도록 하기 위해서, 구분 토큰을 추가한 학습 데이터(GEN + 질의 문장 + SEP + 답변 문장 + EOS)로 언어 모델(GPT2)을 추가 학습시킨다. 이를 기초로, 언어 모델(GPT2)이 한 쌍의 질의 문장-답변 문장으로 구성된 데이터를 생성할 수 있다.
- [0051] 즉, 도 2의 (a)에 도시된 바와 같이, 문장 생성을 위한 시작 토큰(GEN)과 질문 문장, 질문 문장과 답변 문장을 구분하기 위한 구분 토큰(SEP), 답변 문장, 문장의 종료를 알리는 종료 토큰(EOS)을 포함한 데이터(GEN + source + SEP + target + EOS)를 사용하여 언어 모델(GPT-2)을 미세 조정하여 학습시킨다.
- [0052] 다시 말해, 언어 모델(GPT2)은 GEN 토큰부터 순서대로 단어들을 생성하는데, 이를 위해 매 순간 지금까지 생성된 단어를 조건으로 다음에 어떤 단어가 생성될 지 조건부 확률을 계산한다. 언어 모델(GPT2)은 종료 토큰(EOS)이 생성될 때 까지 문장을 생성하고, 언어 모델(GPT2)이 생성한 문장과 정답 데이터로 주어진 데이터(GEN + 질의 문장 + SEP + 답변 문장 + EOS)와의 차이를 비교하여, 조건부 확률 계산 모듈(도면 미도시)을 업데이트한다.
- [0053] 그리고, 도 2의 (b)에 도시된 바와 같이, 가짜 데이터 생성기(140)가 가짜 데이터를 생성할 때 문장 생성 시작 토큰(GEN)만을 입력으로 주어, 구분 토큰(SEP)으로 분리된 질문 문장과 답변 문장을 함께 생성할 수 있다. 즉, 가짜 데이터 생성 단계에서는 GEN 토큰만을 입력으로 주어, 언어 모델(GPT2)이 업데이트된 조건부 확률 계산 모듈을 바탕으로 EOS가 생성될 때 까지 문장을 생성한다.
- [0054] 다음은 연속 학습 시스템(100)을 이용하여 답변 생성 모델을 학습시키는 연속 학습 방법에 대해 도 3을 참조로 설명한다.
- [0055] 도 3은 본 발명의 실시예에 따른 연속 학습 방법에 대한 흐름도이다.
- [0056] 도 3에 도시된 바와 같이, 연속 학습 시스템(100)은 학습 대상 도메인의 데이터가 입력되면, 데이터로부터 질의 문장과 답변 문장을 추출한다(S100).
- [0057] 또한, 연속 학습 시스템(100)은 단기 메모리(120)에 저장되어 있는 기 학습된 적어도 하나 이상의 도메인들의 데이터로부터 샘플링 데이터를 추출한다(S110). 샘플링 데이터를 추출하는 예로서, 단기 메모리(120)에 저장되어 있는 데이터의 도메인 레이블을 활용하여, 단기 메모리(120)에 존재하는 도메인 별로 동일한 수의 데이터를 샘플링한다. 예를 들어 N개의 데이터를 샘플링하고자 하는데, 단기 메모리(120)에 M개의 도메인이 존재한다면, 도메인별로 N/M개의 데이터를 무작위로 샘플링한다.
- [0058] 본 발명의 실시예에서는 단기 메모리(120)에 저장되어 있는 데이터들 중 랜덤하게 데이터를 샘플링 데이터로 추출하는 것을 예로 하여 설명하나, 반드시 이와 같이 한정되는 것은 아니다. 또한, 단기 메모리(120)에 저장되어 있는 학습된 도메인들의 수를 어느 하나로 한정하지 않는다.
- [0059] 연속 학습 시스템(100)은 장기 메모리에 저장된 데이터들로 학습시킨 가짜 데이터 생성 모델에서 가짜 데이터를 생성한다(S120). 연속 학습 시스템(100)은 S100 단계에서 추출한 학습 대상 도메인의 질의 문장과 답변 문장, S110 단계에서 추출한 샘플링 데이터의 질의 문장과 답변 문장, 그리고 S120 단계에서 생성한 가짜 데이터의 질의 문장과 답변 문장을 이용하여 답변 생성 모델을 학습시킨다(S130). 즉, 각각의 질의 문장이 답변 생성 모델로 입력되면, 각 질의 문장에 대응하는 답변 문장이 출력되도록 답변 생성 모델을 학습시킨다.
- [0060] 다음은 연속 학습 시스템(100)에 포함되어 있는 단기 메모리(120)와 장기 메모리(130)를 관리하는 메모리 관리 방법에 대해 도 4를 참조로 설명한다.
- [0061] 도 4는 본 발명의 실시예에 따른 메모리 관리 방법에 대한 흐름도이다.
- [0062] 도 4에 도시된 바와 같이, 상술한 도 3의 절차에 따라 S100 단계에서 수신한 학습 대상 도메인의 학습이 완료되면, 연속 학습 시스템(100)은 학습 대상 도메인의 도메인 레이블에 학습 대상 도메인의 식별 정보를 기입하여 단기 메모리(120)에 저장한다(S200). 이때, 연속 학습 시스템(100)은 단기 메모리(120)에 데이터를 저장한 시간 정보를, 데이터에 포함되어 있는 시간 레이블에 설정한다(S210).
- [0063] 그리고, 연속 학습 시스템(100)은 단기 메모리(120)에 저장된 데이터들 중 시간 레이블에 삽입된 시간 정보를

기초로 미리 설정된 임계 시간 이상 지난 데이터가 있는지 확인한다. 만약 임계 시간 이상 지난 데이터가 있다면, 연속 학습 시스템(100)은 해당 데이터를 단기 메모리(120)에서 삭제한다(S220).

[0064] 한편, 연속 학습 시스템(100)은 S200 단계에서 새로 저장된 학습 대상 도메인의 데이터와 기 저장되어 있는 적어도 하나 이상의 학습된 도메인들의 데이터 사이에 유사도를 계산한다(S230). 즉, 도메인은 상이하더라도 유사한 질의 문장과 답변 문장으로 구성된 데이터가 있을 수 있으므로, 연속 학습 시스템(100)은 학습 대상 도메인의 데이터와 학습된 도메인들의 데이터 사이에 유사도를 계산한다.

[0065] 그리고, 유사도가 기 설정된 임계값 이상으로 계산된 학습된 도메인들의 데이터가 있다면, 해당 데이터에 포함되어 있는 반복 횟수 레이블에 기재되어 있는 횟수를 하나 증가시킨다(S240).

[0066] 한편, 연속 학습 시스템(100)은 단기 메모리(120)에 저장되어 있는 모든 데이터들의 반복 횟수 레이블을 갱신한 후, 각 데이터의 반복 횟수 레이블에 삽입된 횟수가 미리 설정한 임계값 이상인 데이터가 있는지 확인한다. 만약, 임계값 이상인 데이터가 있다면, 연속 학습 시스템(100)은 해당 데이터를 단기 메모리(120)에서 장기 메모리(130)로 이동시킨다(S250).

[0067] 한편, 연속 학습 시스템(100)은 S250 단계에서 단기 메모리(120)에서 장기 메모리(130)로 이동된 데이터와 장기 메모리(130)에 기 저장되어 있는 데이터를 이용하여, 가짜 데이터 생성 모델을 학습시킨다(S260). 연속 학습 시스템(100)이 장기 메모리(130)에 저장된 데이터들로 가짜 데이터 생성 모델을 학습시키는 방법은 다양한 방법으로 실행할 수 있으므로, 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다.

[0068] 그리고, 연속 학습 시스템(100)은 가짜 데이터 생성 모델을 학습시킨 데이터들을 장기 메모리(140)로부터 삭제시킨다(S270). 이때, 가짜 데이터 생성 모델을 학습시킨 모든 데이터를 삭제시킬 수도 있고, 미리 설정한 기준에 따라 장기 메모리(140)에 먼저 저장된 데이터 순서대로 또는 반복 횟수 레이블에 기록된 횟수에 따라 일정 데이터를 삭제시킬 수도 있다.

[0069] 이와 같은 메모리 관리 방법을 통해, 연속 학습 시스템(100)은 도메인 수에 비례하여 메모리의 크기를 계속 증가시키지 않더라도 단기 메모리(120)와 장기 메모리(130)를 이용하여 데이터를 관리할 수 있다.

[0070] 도 5는 본 발명의 실시예에 따른 컴퓨팅 장치의 구조도이다.

[0071] 도 5에 도시된 바와 같이, 물체 감지 시스템(100)은 적어도 하나의 프로세서에 의해 동작하는 컴퓨팅 장치(200)에서, 본 발명의 동작을 실행하도록 기술된 명령들(instructions)이 포함된 프로그램을 실행한다.

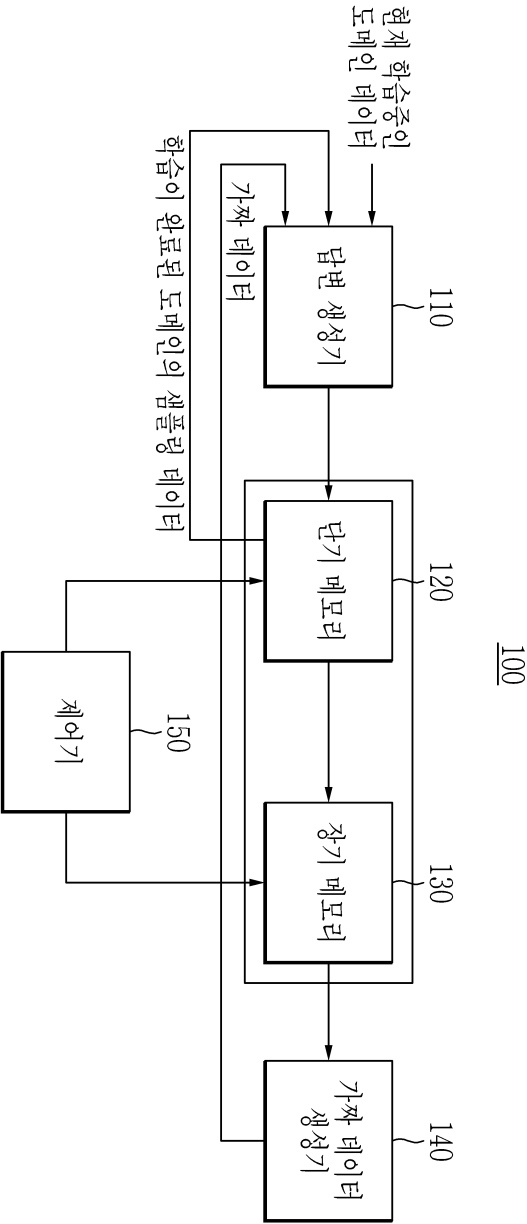
[0072] 컴퓨팅 장치(200)의 하드웨어는 적어도 하나의 프로세서(210), 메모리(220), 스토리지(230), 통신 인터페이스(240)를 포함할 수 있고, 버스를 통해 연결될 수 있다. 이외에도 입력 장치 및 출력 장치 등의 하드웨어가 포함될 수 있다. 컴퓨팅 장치(200)는 프로그램을 구동할 수 있는 운영 체제를 비롯한 각종 소프트웨어가 탑재될 수 있다.

[0073] 프로세서(210)는 컴퓨팅 장치(200)의 동작을 제어하는 장치로서, 프로그램에 포함된 명령들을 처리하는 다양한 형태의 프로세서(210)일 수 있고, 예를 들면, CPU(Central Processing Unit), MPU(Micro Processor Unit), MCU(Micro Controller Unit), GPU(Graphic Processing Unit) 등 일 수 있다. 메모리(220)는 본 발명의 동작을 실행하도록 기술된 명령들이 프로세서(210)에 의해 처리되도록 해당 프로그램을 로드한다. 메모리(220)는 예를 들면, ROM(read only memory), RAM(random access memory) 등 일 수 있다. 스토리지(230)는 본 발명의 동작을 실행하는데 요구되는 각종 데이터, 프로그램 등을 저장한다. 통신 인터페이스(240)는 유/무선 통신 모듈일 수 있다.

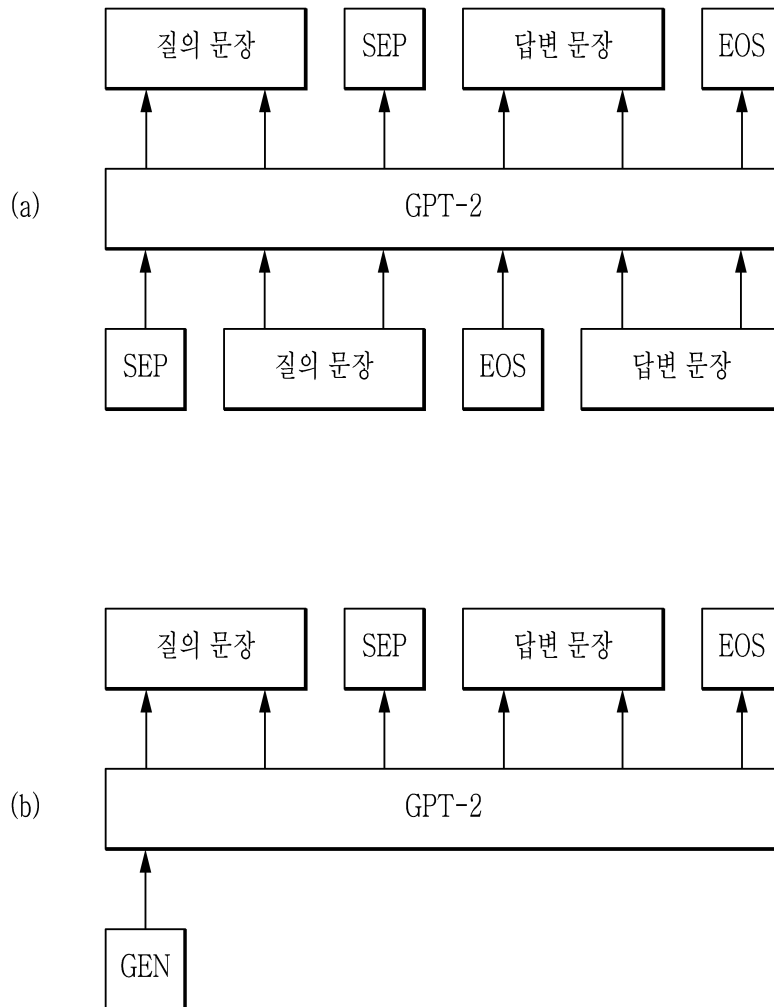
[0074] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

도면

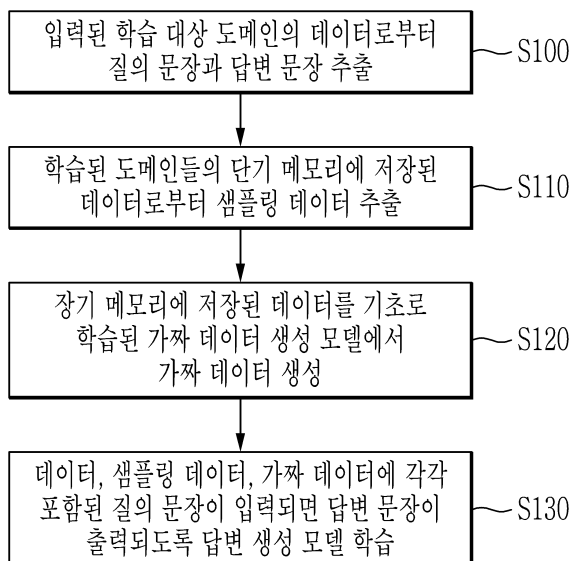
도면1



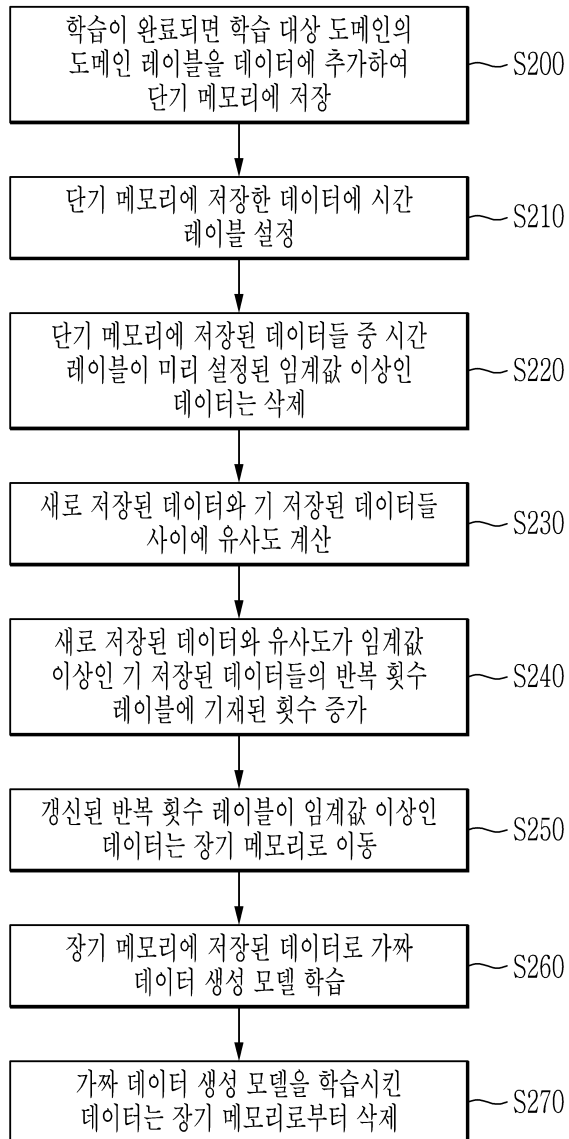
도면2



도면3



도면4



도면5

