



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0065390  
(43) 공개일자 2021년06월04일

(51) 국제특허분류(Int. Cl.)

G06F 40/20 (2020.01)

(52) CPC특허분류

G06F 40/211 (2020.01)

G06F 40/284 (2020.01)

(21) 출원번호 10-2019-0154046

(22) 출원일자 2019년11월27일

심사청구일자 2019년11월27일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

한요섭

서울특별시 은평구 진관1로 77-8, 403동 204호 (진관동, 은평뉴타운폭포동아파트)

이주형

서울특별시 서초구 신반포로 45, 70동 306호(반포동, 반포아파트)

박준우

경기도 부천시 소향로 11, B동 2002호(상동, 코오롱이데아폴리스)

(74) 대리인

민영준

전체 청구항 수 : 총 10 항

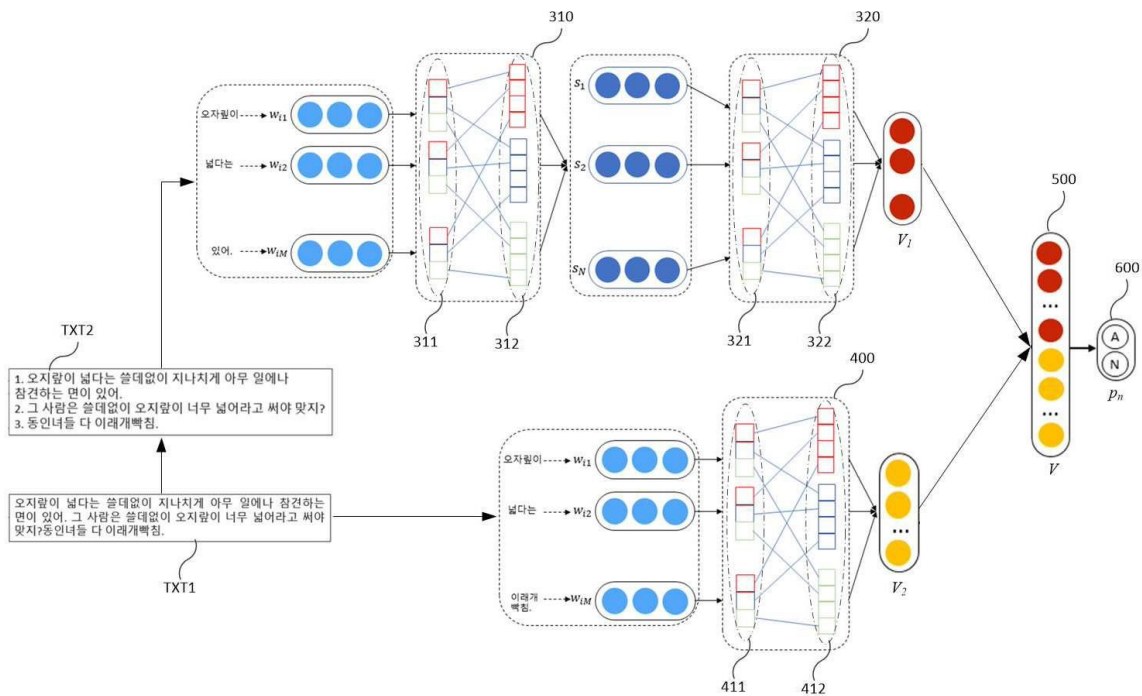
(54) 발명의 명칭 텍스트의 계층적 특성을 반영한 문맥적 비유리 탐지 장치 및 방법

### (57) 요약

본 발명은 입력되는 텍스트에 대해 기지정된 전처리 작업을 수행하고, 문장을 구분하는 전처리부, 전처리된 텍스트 내의 각 단어를 미리 학습된 패턴 추정 방식에 따라 벡터화하여 다수의 단어 벡터를 획득하는 벡터 변환부, 텍스트의 다수의 단어 벡터를 문장 단위로 구분하여 인가받고, 문장단위로 구분되어 인가된 다수의 단어 벡터를

(뒷면에 계속)

### 대표도



미리 학습된 패턴 추정 방식에 따라 인코딩하여 문장 문맥 벡터를 추출하고, 텍스트의 다수의 문장 각각에 대한 문장 문맥 벡터를 미리 학습된 패턴 추정 방식에 따라 인코딩하여 계층적 문맥 벡터를 획득하는 계층적 문맥 벡터 추출부, 텍스트의 다수의 단어 벡터를 일괄적으로 인가받아 미리 학습된 패턴 추정 방식에 따라 인코딩하여 포괄적 문맥 벡터를 획득하는 포괄적 문맥 벡터 추출부, 계층적 문맥 벡터와 포괄적 문맥 벡터를 결합하여 문맥 벡터를 획득하는 문맥 벡터 결합부 및 문맥 벡터를 인가받아 미리 학습된 패턴 분류 방식에 따라 분류하여 텍스트의 비윤리성을 판별하는 비윤리 판별부를 포함하는 비윤리 탐지 장치 및 방법을 제공할 수 있다.

(52) CPC특허분류

*G06F 40/30* (2020.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	2016-0-00562
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원(한국연구재단부설)
연구사업명	정보통신방송연구개발사업
연구과제명	[주관/한국과학기술원] 상대방의 감성을 추론, 판단하여 그에 맞추어 대화하고 대응할 수 있는 감성 지능 연구개발 (4/5)
기 여 율	1/1
과제수행기관명	한국과학기술원
연구기간	2019.05.01 ~ 2020.02.29

---

## 명세서

### 청구범위

#### 청구항 1

입력되는 텍스트에 대해 기지정된 전처리 작업을 수행하고, 문장을 구분하는 전처리부;

상기 전처리된 텍스트 내의 각 단어를 미리 학습된 패턴 추정 방식에 따라 벡터화하여 다수의 단어 벡터를 획득하는 벡터 변환부;

상기 텍스트의 다수의 단어 벡터를 문장 단위로 구분하여 인가받고, 문장단위로 구분되어 인가된 다수의 단어 벡터를 미리 학습된 패턴 추정 방식에 따라 인코딩하여 문장 문맥 벡터를 추출하고, 상기 텍스트의 다수의 문장 각각에 대한 문장 문맥 벡터를 미리 학습된 패턴 추정 방식에 따라 인코딩하여 계층적 문맥 벡터를 획득하는 계층적 문맥 벡터 추출부;

상기 텍스트의 다수의 단어 벡터를 일괄적으로 인가받아 미리 학습된 패턴 추정 방식에 따라 인코딩하여 포괄적 문맥 벡터를 획득하는 포괄적 문맥 벡터 추출부;

상기 계층적 문맥 벡터와 상기 포괄적 문맥 벡터를 결합하여 문맥 벡터를 획득하는 문맥 벡터 결합부; 및

상기 문맥 벡터를 인가받아 미리 학습된 패턴 분류 방식에 따라 분류하여 텍스트의 비윤리성을 판별하는 비윤리 판별부를 포함하는 비윤리 탐지 장치.

#### 청구항 2

제1 항에 있어서, 상기 계층적 문맥 벡터 추출부는

미리 학습된 패턴 추정 방식에 따라 문장 단위로 구분되어 전달되는 상기 단어 벡터의 특징을 추출하고, 문장 내에서 추출된 단어 벡터 각각의 특징이 서로 연관되는 문맥 특징을 추출하여 상기 문장 문맥 벡터를 획득하는 단어 인코더; 및

미리 학습된 패턴 추정 방식에 따라 상기 문장 문맥 벡터의 특징을 추출하고, 텍스트 내에서 다수의 문장 각각에 대한 문장 문맥 벡터의 특징이 서로 연관되는 문맥 특징을 추출하여 상기 계층적 문맥 벡터를 획득하는 문장 인코더를 포함하는 비윤리 탐지 장치.

#### 청구항 3

제2 항에 있어서, 상기 포괄적 문맥 벡터 추출부는

미리 학습된 패턴 추정 방식에 따라 상기 텍스트 전체의 다수의 단어 벡터를 일괄적으로 인가받아 특징을 추출하고, 추출된 단어 벡터 각각의 특징이 서로 연관되는 문맥 특징을 추출하여 상기 포괄적 문맥 벡터를 획득하는 텍스트 인코더로 구현되는 비윤리 탐지 장치.

#### 청구항 4

제3 항에 있어서, 상기 단어 인코더와 상기 문장 인코더 및 상기 텍스트 인코더는 C-LSTM 으로 구현되는 비윤리 탐지 장치.

#### 청구항 5

제4 항에 있어서, 상기 비윤리 탐지 장치는

학습시에 윤리 또는 비윤리가 레이블된 학습용 텍스트를 입력받고, 상기 비윤리 판별부의 판별결과와 상기 학습용 텍스트의 레이블을 비교하여 오차를 역전파함으로써, 상기 단어 인코더와 상기 문장 인코더 및 상기 텍스트 인코더가 학습되는 비윤리 탐지 장치.

#### 청구항 6

입력되는 텍스트에 대해 기지정된 전처리 작업을 수행하고, 문장을 구분하는 단계;

상기 전처리된 텍스트 내의 각 단어를 미리 학습된 패턴 추정 방식에 따라 벡터화하여 다수의 단어 벡터를 획득하는 단계;

상기 텍스트의 다수의 단어 벡터를 문장 단위로 구분하여 인가받고, 문장단위로 구분되어 인가된 다수의 단어 벡터를 미리 학습된 패턴 추정 방식에 따라 인코딩하여 문장 문맥 벡터를 추출하는 단계;

상기 텍스트의 다수의 문장 각각에 대한 문장 문맥 벡터를 미리 학습된 패턴 추정 방식에 따라 인코딩하여 계층적 문맥 벡터를 획득하는 단계;

상기 텍스트의 다수의 단어 벡터를 일괄적으로 인가받아 미리 학습된 패턴 추정 방식에 따라 인코딩하여 포괄적 문맥 벡터를 획득하는 단계;

상기 계층적 문맥 벡터와 상기 포괄적 문맥 벡터를 결합하여 문맥 벡터를 획득하는 단계; 및

상기 문맥 벡터를 인가받아 미리 학습된 패턴 분류 방식에 따라 분류하여 텍스트의 비윤리성을 판별하는 단계를 포함하는 비윤리 탐지 방법.

## 청구항 7

제6 항에 있어서, 상기 문장 문맥 벡터를 추출하는 단계는

미리 학습된 패턴 추정 방식에 따라 문장 단위로 구분되어 전달되는 상기 단어 벡터의 특징을 추출하는 단계; 및

미리 학습된 패턴 추정 방식에 따라 문장 내에서 추출된 단어 벡터 각각의 특징이 서로 연관되는 문맥 특징을 추출하여 상기 문장 문맥 벡터를 획득하는 단계를 포함하는 비윤리 탐지 방법.

## 청구항 8

제7 항에 있어서, 상기 계층적 문맥 벡터를 획득하는 단계는

미리 학습된 패턴 추정 방식에 따라 상기 문장 문맥 벡터의 특징을 추출하는 단계; 및

미리 학습된 패턴 추정 방식에 따라 텍스트 내에서 다수의 문장 각각에 대한 문장 문맥 벡터의 특징이 서로 연관되는 문맥 특징을 추출하여 상기 계층적 문맥 벡터를 획득하는 단계를 포함하는 비윤리 탐지 방법.

## 청구항 9

제8 항에 있어서, 상기 포괄적 문맥 벡터를 획득하는 단계는

미리 학습된 패턴 추정 방식에 따라 상기 텍스트 전체의 다수의 단어 벡터를 일괄적으로 인가받아 특징을 추출하는 단계; 및

미리 학습된 패턴 추정 방식에 따라 추출된 단어 벡터 각각의 특징이 서로 연관되는 문맥 특징을 추출하여 상기 포괄적 문맥 벡터를 획득하는 단계를 포함하는 비윤리 탐지 방법.

## 청구항 10

제9 항에 있어서, 상기 문장 문맥 벡터를 추출하는 단계와 상기 계층적 문맥 벡터를 획득하는 단계 및 상기 포괄적 문맥 벡터를 획득하는 단계는 각각 미리 학습된 C-LSTM에 의해 수행되고,

상기 비윤리 탐지 방법은

학습시에 윤리 또는 비윤리가 레이블된 학습용 텍스트를 입력받고, 상기 비윤리성을 판별하는 단계의 판별결과와 상기 학습용 텍스트의 레이블을 비교하여 오차를 상기 C-LSTM으로 역전파하는 학습 단계를 더 포함하는 비윤리 탐지 방법.

## 발명의 설명

## 기술 분야

본 발명은 비윤리 탐지 장치 및 방법에 관한 것으로, 텍스트의 계층적 특성을 반영한 문맥적 비윤리 탐지 장치

[0001]

및 방법에 관한 것이다.

## 배경 기술

- [0002] 온라인 상에서의 커뮤니케이션이 활성화됨에 따라 각종 온라인 커뮤니티, SNS, 게시판 등과 같은 다양한 커뮤니케이션 수단을 이용하여 사용자들이 상호 의견을 교환하고 있다.
- [0003] 온라인 환경에서는 많은 비속어가 사용되고 있으며, 또한 빠르게 전파되는 특징을 갖는다. 따라서 온라인 상에서 비속어를 탐지하여 필터링하기 위한 다양한 방법이 연구되어 왔다. 다만 온라인에서는 비속어뿐만 아니라 비윤리적 언어 또한 빈번하게 사용되고 있으며, 이에 최근에는 비윤리어를 탐지하기 위한 연구가 활발하게 수행되고 있으나, 단어 수준에서의 탐지되는 비속어에 비해 비윤리어의 경우, 문장의 문맥 흐름을 고려하여 탐지해야만 한다. 따라서 비속어가 포함되지 않은 비윤리 문장의 경우, 텍스트 전체를 단어 단위로 구분하여 벡터화하고, 벡터화된 단어를 기반으로 단순히 탐색을 수행하는 기존의 방식으로는 길이가 긴 텍스트에서 거리가 이격된 단어들 사이의 연계성에 대한 특징을 추출할 수 없어 탐지가 용이하지 않다는 한계가 있다.
- [0004] 따라서 텍스트의 길이로 인한 성능 저하가 적으면서 비속어가 포함되지 않은 비윤리적 언어를 탐지하여 필터링할 수 있도록 기술이 요구된다.

## 선행기술문헌

### 특허문헌

- [0005] (특허문헌 0001) 한국 공개 특허 제10-2016-0089210호 (2016.07.27 공개)

## 발명의 내용

### 해결하려는 과제

- [0006] 본 발명의 목적은 입력되는 텍스트의 길이에 무관하게 비윤리 문장을 용이하게 탐지할 수 있는 비윤리 탐지 장치 및 방법을 제공하는데 있다.
- [0007] 본 발명의 다른 목적은 다양한 길이로 입력되는 텍스트의 계층적 특성을 반영하여 문맥적 비윤리성을 탐지하여 정확하게 비윤리 문장을 탐지할 수 있는 비윤리 탐지 장치 및 방법을 제공하는데 있다.

### 과제의 해결 수단

- [0008] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 비윤리 탐지 장치는 입력되는 텍스트에 대해 기지정된 전처리 작업을 수행하고, 문장을 구분하는 전처리부; 상기 전처리된 텍스트 내의 각 단어를 미리 학습된 패턴 추정 방식에 따라 벡터화하여 다수의 단어 벡터를 획득하는 벡터 변환부; 상기 텍스트의 다수의 단어 벡터를 문장 단위로 구분하여 인가받고, 문장단위로 구분되어 인가된 다수의 단어 벡터를 미리 학습된 패턴 추정 방식에 따라 인코딩하여 문장 문맥 벡터를 추출하고, 상기 텍스트의 다수의 문장 각각에 대한 문장 문맥 벡터를 미리 학습된 패턴 추정 방식에 따라 인코딩하여 계층적 문맥 벡터를 획득하는 계층적 문맥 벡터 추출부; 상기 텍스트의 다수의 단어 벡터를 일괄적으로 인가받아 미리 학습된 패턴 추정 방식에 따라 인코딩하여 포괄적 문맥 벡터를 획득하는 포괄적 문맥 벡터 추출부; 상기 계층적 문맥 벡터와 상기 포괄적 문맥 벡터를 결합하여 문맥 벡터를 획득하는 문맥 벡터 결합부; 및 상기 문맥 벡터를 인가받아 미리 학습된 패턴 분류 방식에 따라 분류하여 텍스트의 비윤리성을 판별하는 비윤리 판별부를 포함한다.
- [0009] 상기 계층적 문맥 벡터 추출부는 미리 학습된 패턴 추정 방식에 따라 문장 단위로 구분되어 전달되는 상기 단어 벡터의 특징을 추출하고, 문장 내에서 추출된 단어 벡터 각각의 특징이 서로 연관되는 문맥 특징을 추출하여 상기 문장 문맥 벡터를 획득하는 단어 인코더; 및 미리 학습된 패턴 추정 방식에 따라 상기 문장 문맥 벡터의 특징을 추출하고, 텍스트 내에서 다수의 문장 각각에 대한 문장 문맥 벡터의 특징이 서로 연관되는 문맥 특징을 추출하여 상기 계층적 문맥 벡터를 획득하는 문장 인코더를 포함할 수 있다.
- [0010] 상기 포괄적 문맥 벡터 추출부는 미리 학습된 패턴 추정 방식에 따라 상기 텍스트 전체의 다수의 단어 벡터를 일괄적으로 인가받아 특징을 추출하고, 추출된 단어 벡터 각각의 특징이 서로 연관되는 문맥 특징을 추출하여 상기 포괄적 문맥 벡터를 획득하는 텍스트 인코더로 구현될 수 있다.

- [0011] 상기 단어 인코더와 상기 문장 인코더 및 상기 텍스트 인코더는 C-LSTM 으로 구현될 수 있다.
- [0012] 상기 비윤리 탐지 장치는 학습시에 윤리 또는 비윤리가 레이블된 학습용 텍스트를 입력받고, 상기 비윤리 판별부의 판별결과와 상기 학습용 텍스트의 레이블을 비교하여 오차를 역전파함으로써, 상기 단어 인코더와 상기 문장 인코더 및 상기 텍스트 인코더가 학습될 수 있다.
- [0013] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 비윤리 탐지 방법은 입력되는 텍스트에 대해 기지정된 전처리 작업을 수행하고, 문장을 구분하는 단계; 상기 전처리된 텍스트 내의 각 단어를 미리 학습된 패턴 추정 방식에 따라 벡터화하여 다수의 단어 벡터를 획득하는 단계; 상기 텍스트의 다수의 단어 벡터를 문장 단위로 구분하여 인가받고, 문장단위로 구분되어 인가된 다수의 단어 벡터를 미리 학습된 패턴 추정 방식에 따라 인코딩하여 문장 문맥 벡터를 추출하는 단계; 상기 텍스트의 다수의 문장 각각에 대한 문장 문맥 벡터를 미리 학습된 패턴 추정 방식에 따라 인코딩하여 계층적 문맥 벡터를 획득하는 단계; 상기 텍스트의 다수의 단어 벡터를 일괄적으로 인가받아 미리 학습된 패턴 추정 방식에 따라 인코딩하여 포괄적 문맥 벡터를 획득하는 단계; 상기 계층적 문맥 벡터와 상기 포괄적 문맥 벡터를 결합하여 문맥 벡터를 획득하는 단계; 및 상기 문맥 벡터를 인가받아 미리 학습된 패턴 분류 방식에 따라 분류하여 텍스트의 비윤리성을 판별하는 단계를 포함한다.

### 발명의 효과

- [0014] 따라서, 본 발명의 실시예에 따른 비윤리 탐지 장치 및 방법은 입력되는 텍스트에 대해 텍스트 내의 각 문장에 대한 문맥 벡터를 각각 추출한 후 텍스트 전체에 대해 문맥 벡터를 획득하여 계층적 문맥 벡터를 추출함 동시에 텍스트 전체에 대한 포괄적 문맥 벡터를 별도로 추출한 후, 추출된 계층적 문맥 벡터와 포괄적 문맥 벡터를 결합하여 비윤리어를 탐지함으로써, 텍스트의 길이에 무관하게 비윤리 문장을 용이하게 탐지할 수 있다.

### 도면의 간단한 설명

- [0015] 도 1은 본 발명의 일 실시예에 따른 비윤리 탐지 장치의 개략적 구조를 나타낸다.
- 도 2는 도 1의 비윤리 탐지 장치에서 계층적 문맥 벡터 추출부와 포괄적 문맥 벡터 추출부를 상세하게 나타낸 도면이다.
- 도 3은 본 발명의 일 실시예에 따른 비윤리 탐지 방법을 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

- [0016] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시예에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.
- [0017] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.
- [0018] 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. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0019] 도 1은 본 발명의 일 실시예에 따른 비윤리 탐지 장치의 개략적 구조를 나타내고, 도 2는 도 1의 비윤리 탐지 장치에서 계층적 문맥 벡터 추출부와 포괄적 문맥 벡터 추출부를 상세하게 나타낸 도면이다.
- [0020] 도 1 및 도 2를 참조하여, 본 실시예에 따른 비윤리 탐지 장치를 설명하면, 비윤리 탐지 장치는 전처리부(100), 벡터 변환부(200), 계층적 문맥 벡터 추출부(300), 포괄적 문맥 벡터 추출부(400), 문맥 벡터 결합부(500) 및 비윤리 판별부(600)를 포함한다.
- [0021] 전처리부(100)는 비윤리어를 탐지할 텍스트를 인가받고, 텍스트 내에서 문자, 공백, 구두점 등과 같이 문장을 구성하는 문장 구성 요소 이외에 나머지 구성 요소인 특수 문자, URL, SNS 지정 특성 문자(# 해시태그, @ 언급)등의 부가 구성 요소 모두 제거한다. 이는 문자와 문장 기호 및 공백과 같이 문장을 구성하는 문장 구성 요소 이외의 부가 구성 요소들은 비속어나 비윤리어로 이용될 가능성이 거의 없으므로 탐지 대상에서 배제하기 위



해서이다. 그리고 전처리부(100)는 부가 구성 요소가 제거된 텍스트 내에 포함된 각 문장을 구분한다.

- [0022] 벡터 변환부(200)는 전처리부(100)에서 부가 구성 요소가 제거되고 문장 단위로 구분된 텍스트를 인가받아 단어 단위로 구분하고, 구분된 단어 각각을 임베딩하여 벡터화함으로써 다수의 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 획득한다. 여기서  $i$ 는 텍스트의  $i$ 번째 문장을 나타내고,  $k$ 는  $i$ 번째 문장의  $k$ 번째 단어를 나타낸다.
- [0023] 벡터 변환부(200)는 미리 학습된 임베딩 모델을 이용하여 단어 각각을 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )로 변환할 수 있다. 단어를 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )로 변환하기 위한 다양한 임베딩 모델이 이미 공개되어 있으므로, 본 실시예의 벡터 변환부(200) 또한 기존에 공개된 임베딩 모델을 이용하여 단어를 단어 벡터로 변환할 수 있으며, 여기서는 일례로 fastText 로 알려진 임베딩 모델을 이용하여 단어를 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )로 변환하는 것으로 가정한다.
- [0024] 벡터 변환부(200)는 텍스트의 각 단어가 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )로 변환되면, 변환된 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 계층적 문맥 벡터 추출부(300)와 포괄적 문맥 벡터 추출부(400)로 각각 전달한다.
- [0025] 이때 벡터 변환부(200)는 포괄적 문맥 벡터 추출부(400)로는 텍스트 내의 모든 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 일괄적으로 전달하는 반면, 계층적 문맥 벡터 추출부(300)는 텍스트 내에서 구분된 각 문장 단위로 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 순차적으로 전달한다.
- [0026] 일례로 도 2에 도시된 바와 같이 벡터 변환부(200)는 전처리된 입력 텍스트(TXT1) 전체에 대한 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 함께 포괄적 문맥 벡터 추출부(400)로 전달한다. 이와 함께 벡터 변환부(200)는 전처리된 입력 텍스트(TXT1)의 각 문장을 문장 텍스트(TXT2)와 같이 구분하고, 문장 텍스트(TXT2)의 구분된 문장 각각에 대한 단어 벡터( $(w_{11}, w_{12}, \dots, w_{1,k}), (w_{21}, w_{22}, \dots, w_{2,k}), (w_{31}, w_{32}, \dots, w_{3,k})$ )를 순차적으로 계층적 문맥 벡터 추출부(300)로 전달한다.
- [0027] 계층적 문맥 벡터 추출부(300)는 미리 학습된 패턴 추정 방식에 따라 벡터 변환부(200)로부터 문장 단위로 구분되어 전달되는 단어 벡터( $(w_{11}, w_{12}, \dots, w_{1,k}), (w_{21}, w_{22}, \dots, w_{2,k}), (w_{31}, w_{32}, \dots, w_{3,k})$ )에 대해 인코딩하여, 각 문장에 대한 문장 문맥 벡터( $s_i$ )를 추출하고, 추출된 문장 문맥 벡터들( $s_i$ )에 대해 다시 미리 학습된 패턴 추정 방식에 따라 인코딩하여 계층적 문맥 벡터( $V_1$ )를 추출한다.
- [0028] 계층적 문맥 벡터 추출부(300)는 문장 단위로 구분되어 전달되는 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 인코딩하여 문장 문맥 벡터( $s_i$ )를 추출하는 단어 인코더(310)와 문장 문맥 벡터들( $s_i$ )을 인코딩하여 계층적 문맥 벡터( $V_1$ )를 추출하는 문장 인코더(320)를 포함할 수 있다.
- [0029] 그리고 단어 인코더(310)와 문장 인코더(320) 각각은 컨볼루션 신경망(Convolutional Neural Network: 이하 CNN)과 LSTM(Long Short Term Memory)가 결합된 C-LSTM 모델로 구현될 수 있다. 이는 CNN의 경우, 입력되는 벡터에 대한 특징을 추출하기에 용이하지만, 이전 추출된 특징이 반영되지 않는데 반해, LSTM은 순환 신경망(Recurrent Neural Network: RNN)이 장기간(Long Term) 특징을 반영할 수 있도록 개선된 구조를 갖는 신경망으로서 이전 추출된 특징이 누적 반영됨으로써 문맥 특징을 획득하기 용이하다는 장점이 있기 때문이다.
- [0030] 이에 도 2에 도시된 바와 같이, 단어 인코더(310)는 CNN으로 구현되는 단어 특징 추출부(311)와 LSTM으로 구현되는 문장 문맥 벡터 획득부(312)로 구성될 수 있으며, 문장 인코더(320) 또한 CNN으로 구현되는 문장 특징 추출부(321)와 LSTM으로 구현되는 계층적 문맥 벡터 획득부(322)로 구성될 수 있다.
- [0031] 단어 인코더(310)에서 CNN으로 구현되는 단어 특징 추출부(311)는 입력된 문장 단위로 구분되어 전달되는 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ ) 각각의 특징을 추출한다. 그리고 LSTM으로 구현되는 문장 문맥 벡터 획득부(312)는 단어 특징 추출부(311)된 문장 내에서 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ ) 각각의 특징이 서로 연관되는 문맥 특징을 추출하여 문장 문맥 벡터( $s_i$ )를 획득한다.
- [0032] 한편, 문장 인코더(320)에서 CNN으로 구현되는 문장 특징 추출부(321)는 단어 인코더(310)에서 전달되는 다수의 문장 문맥 벡터( $s_i$ ) 각각의 특징을 추출하고, LSTM으로 구현되는 계층적 문맥 벡터 획득부(322)는 텍스트 내의

다수의 문장 각각의 특징이 서로 연관되는 텍스트 전체의 문맥 특징을 추출하여 계층적 문맥 벡터( $V_1$ )를 획득할 수 있다.

[0033] 여기서 계층적 문맥 벡터 획득부(322)가 추출한 텍스트 전체의 문맥 특징을 계층적 문맥 벡터( $V_1$ )라 하는 것은 단어 인코더(310)에 의해 하위 계층에 해당하는 각각의 문장 내의 문장 문맥 벡터( $s_i$ )가 우선 획득되고, 획득된 문장 문맥 벡터( $s_i$ )를 이용하여 상위 계층에 해당하는 텍스트 전체의 문맥 특징이 추출되었기 때문이다.

[0034] 즉 계층적 문맥 벡터 추출부(300)는 텍스트를 하위 계층인 문장과 상위 계층인 텍스트로 구분하고, 구분된 계층에 따라 순차적으로 문맥 특징을 추출하여, 계층적 문맥 벡터( $V_1$ )를 획득한다.

[0035] 한편, 포괄적 문맥 벡터 추출부(400)는 전처리된 입력 텍스트(TXT1) 전체에 대한 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 일괄적으로 전달받고, 미리 학습된 패턴 추정 방식에 따라 일괄적으로 전달된 입력 텍스트(TXT1) 전체에 대한 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 인코딩하여 포괄적 문맥 벡터( $V_2$ )를 추출한다.

[0036] 포괄적 문맥 벡터 추출부(400) 또한 C-LSTM 모델로 구현되는 인코더로서 텍스트 전체에 대해 인코딩을 수행하므로, 텍스트 인코더로 볼 수 있다. 즉 포괄적 문맥 벡터 추출부(400)는 CNN으로 구현되는 텍스트 특징 추출부(411)와 LSTM으로 구현되는 포괄적 문맥 벡터 획득부(412)로 구성될 수 있다.

[0037] CNN으로 구현되는 텍스트 특징 추출부(411)는 미리 학습된 패턴 추정 방식에 따라 벡터 변환부(200)로부터 전달되는 텍스트(TXT1) 전체의 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ ) 각각의 특징을 추출한다. 즉 단어 인코더(310)의 단어 특징 추출부(311)와 유사한 동작을 수행한다. 다만, 텍스트 특징 추출부(411)는 문장 단위가 아닌 텍스트 전체에 대해 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ ) 각각의 특징을 추출하는 점에서 상이하다.

[0038] 그리고 LSTM으로 구현되는 포괄적 문맥 벡터 획득부(412)는 텍스트 내의 각 단어의 특징이 서로 연관되는 텍스트 전체의 문맥 특징을 추출하여 포괄적 문맥 벡터( $V_2$ )를 획득한다.

[0039] 포괄적 문맥 벡터 획득부(412)는 계층적 문맥 벡터 추출부(300)와 달리 문장의 구분없이 텍스트의 모든 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 이용하여 일괄적으로 텍스트 전체의 문맥 특징이 추출하므로, 텍스트(TXT1)에 대한 포괄적인 문맥 벡터를 추출한 것으로 볼 수 있으며, 이에 포괄적 문맥 벡터( $V_2$ )라고 한다.

[0040] 계층적 문맥 벡터( $V_1$ )와 포괄적 문맥 벡터( $V_2$ )가 획득되면, 문맥 벡터 결합부(500)는 획득된 계층적 문맥 벡터( $V_1$ )와 포괄적 문맥 벡터( $V_2$ )를 결합(concatenate)하여, 문맥 벡터( $V$ )를 획득한다.

[0041] 본 실시예에서 벡터 변환부(200)는 포괄적 문맥 벡터 추출부(400)로 전달하는 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )와 계층적 문맥 벡터 추출부(300)로 전달하는 단어 벡터( $(w_{11}, w_{12}, \dots, w_{1,k}), (w_{21}, w_{22}, \dots, w_{2,k}), (w_{31}, w_{32}, \dots, w_{3,k})$ )의 차원을 조절하여 전달할 수 있다. 일례로 벡터 변환부(200)는 입력 텍스트(TXT1)의 각각의 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 1 X 300 크기의 행렬에 대입하고, 문장 텍스트(TXT2)의 구분된 문장에서의 각 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 1 X 100 크기의 행렬에 대입하여 2차원 행렬 형태로 변환하여, 각각 포괄적 문맥 벡터 추출부(400)와 계층적 문맥 벡터 추출부(300)로 전달할 수 있다. 이는 포괄적 문맥 벡터 추출부(400)와 계층적 문맥 벡터 추출부(300)의 CNN이 컨볼루션 연산을 수행할 수 있도록 하고, 서로 차원을 동일하게 하여 계층적 문맥 벡터( $V_1$ )와 포괄적 문맥 벡터( $V_2$ )를 결합이 용이하도록 하기 위해서이다.

[0042] 그리고 비윤리 판별부(600)는 계층적 문맥 벡터( $V_1$ )와 포괄적 문맥 벡터( $V_2$ )가 결합되어 획득된 문맥 벡터( $V$ )를 인가받고, 미리 학습된 패턴 분류 방식에 따라 문맥 벡터( $V$ )를 분류하여, 텍스트가 비윤리 텍스트일 비윤리 확률(A)과 윤리 텍스트일 윤리 확률(N)을 출력할 수 있다. 경우에 따라서 비윤리 판별부(600)는 비윤리 확률(A)과 윤리 확률(N)에 기반하여 텍스트가 비윤리 텍스트인지 윤리 텍스트인지를 결정하여 하나의 결과만을 출력하도록 구성될 수도 있다. 또한 비윤리 확률(A)과 윤리 확률(N)을 추출하지 않고 곧바로 비윤리 텍스트인지 윤리 텍스트로 판별하여 출력하도록 구성될 수도 있다.

[0043] 본 실시예에 따른 비윤리 탐지 장치가 계층적 문맥 벡터 추출부(300)와 포괄적 문맥 벡터 추출부(400)를 구비하는 것은 텍스트의 길이에 무관하게 비윤리어를 정확하게 검출할 수 있도록 하기 위함이다.



- [0044] 만일 텍스트 전체의 단어 벡터를 기반으로 포괄적 문맥 벡터( $V_2$ )를 추출하는 포괄적 문맥 벡터 추출부(400)만을 구비하는 경우, 비록 C-LSTM을 이용할지라도 문장의 길이가 길어지게 되면, 텍스트 전체에 대한 정확한 문맥 벡터를 추출하기 어렵다. 즉 포괄적 문맥 벡터 추출부(400)의 경우 짧은 텍스트에는 적합하지만 길이가 긴 텍스트의 경우에는 탐지 성능이 크게 하락한다.
- [0045] 반면, 계층적 문맥 벡터 추출부(300)만을 포함하면, 텍스트에 포함된 문장 단위로 우선 문장 문맥 벡터를 추출한 후 전체 텍스트에 대한 문맥 벡터를 추출하므로, 길이가 긴 문장에서도 정확한 탐지 성능을 나타낼 수 있다. 다만 텍스트의 전체적인 문맥에서의 검출 성능은 포괄적 문맥 벡터 추출부(400)에 비해 떨어질 수 있다. 특히 짧은 텍스트의 경우에는 문장을 구분하는 것보다 텍스트 전체적으로 문맥을 분석하여 탐지하는 것이 더욱 우수한 성능을 나타낼 수 있다.
- [0046] 이에 본 실시예에 따른 비윤리 탐지 장치는 계층적 문맥 벡터 추출부(300)와 포괄적 문맥 벡터 추출부(400)를 함께 구비함으로써, 길이가 긴 텍스트나 짧은 텍스트 모두에서 우수한 성능으로 비윤리 텍스트를 탐지할 수 있다.
- [0047] 한편 본 실시예에서 비윤리 탐지 장치의 계층적 문맥 벡터 추출부(300)와 포괄적 문맥 벡터 추출부(400) 및 비윤리 판별부(600)는 미리 학습이 수행되어야 하며, 학습은 미리 윤리 또는 비윤리가 레이블된 다수의 학습용 텍스트를 입력하고, 비윤리 판별부(600)의 판별 결과와 레이블을 비교하여 오차를 역전파함으로써 수행될 수 있다.
- [0048] 도 3은 본 발명의 일 실시예에 따른 비윤리 탐지 방법을 나타낸다.
- [0049] 도 1 및 도 2를 참조하여, 도 3의 비윤리 탐지 방법을 설명하면, 우선 입력된 텍스트에 대해 기지정된 전처리 작업을 수행한다(S10). 여기서 전처리는 텍스트 내에서 문자, 공백, 구두점 등과 같이 문장을 구성하는 문장 구성 요소 이외에 나머지 부가 구성 요소를 모두 제거하고 전처리된 텍스트에서 각 문장을 구분하여 수행된다. 그리고 미리 학습된 임베딩 모델을 이용하여 전처리된 텍스트의 각 단어를 벡터화하여 다수의 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 획득한다(S20).
- [0050] 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )가 획득되면, 미리 학습된 패턴 추정 방식에 따라 획득된 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 문장 단위로 구분하여 인코딩하여, 각 문장에 대한 문장 문맥 벡터( $s_i$ )를 획득한다(S30).
- [0051] 이때 문장 문맥 벡터( $s_i$ )는 미리 학습된 패턴 추정 방식에 따라 문장 단위로 획득된 다수의 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )의 특징을 추출하고, 추출된 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ ) 각각의 특징이 서로 연관되는 문장 문맥 특징을 추출하여 획득될 수 있다.
- [0052] 그리고 텍스트 내의 다수의 문장에 대한 문장 문맥 벡터( $s_i$ )가 획득되면, 미리 학습된 패턴 추정 방식에 따라 획득된 다수의 문장 문맥 벡터( $s_i$ )를 함께 인코딩하여 계층적 문맥 벡터( $V_1$ )를 획득한다(S40).
- [0053] 여기서 계층적 문맥 벡터( $V_1$ )도 미리 학습된 패턴 추정 방식에 따라 다수의 문장 문맥 벡터( $s_i$ ) 각각의 특징을 추출하고, 추출된 다수의 문장 문맥 벡터( $s_i$ )의 특징이 서로 연관되는 텍스트 전체의 텍스트 문맥 특징을 추출하여 획득될 수 있다.
- [0054] 한편, 문장 문맥 벡터( $s_i$ )를 획득(S30)하고, 계층적 문맥 벡터( $V_1$ )를 획득(S40)함과 동시에 텍스트에서 획득된 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )를 미리 학습된 패턴 추정 방식에 따라 함께 인코딩하여 포괄적 문맥 벡터( $V_2$ )를 획득한다.
- [0055] 포괄적 문맥 벡터( $V_2$ )는 미리 학습된 패턴 추정 방식에 따라 텍스트 전체의 다수의 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ )의 특징을 추출하고, 추출된 텍스트 전체 단어 벡터( $w_{i1}, w_{i2}, \dots, w_{i,k}$ ) 각각의 특징이 서로 연관되는 텍스트 문맥 특징을 추출하여 획득될 수 있다.
- [0056] 텍스트에 대한 계층적 문맥 벡터( $V_1$ )와 포괄적 문맥 벡터( $V_2$ )가 획득되면, 획득된 계층적 문맥 벡터( $V_1$ )와 포괄적 문맥 벡터( $V_2$ )를 결합하여 문맥 벡터( $V$ )를 획득한다(S60).
- [0057] 그리고 획득된 문맥 벡터( $V$ )를 미리 학습된 패턴 분류 방식에 따라 분류하여 텍스트의 비윤리성을 판별한다

(S70). 이때 텍스트의 비윤리성은 확률 형태로 비윤리 확률(A)과 윤리 확률(N)이 각각 판별될 수 있으나, 비윤리 텍스트 또는 윤리 텍스트의 결과값으로 출력되도록 학습될 수도 있다.

[0058] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.

[0059] 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.

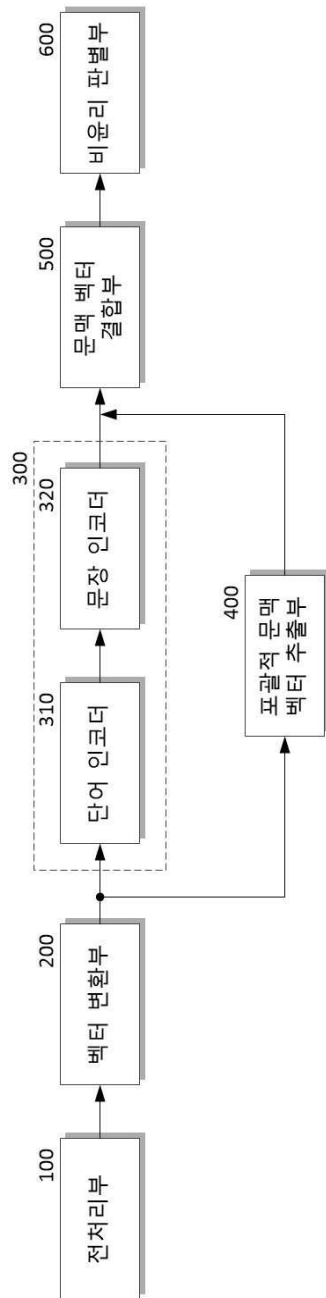
[0060] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

### 부호의 설명

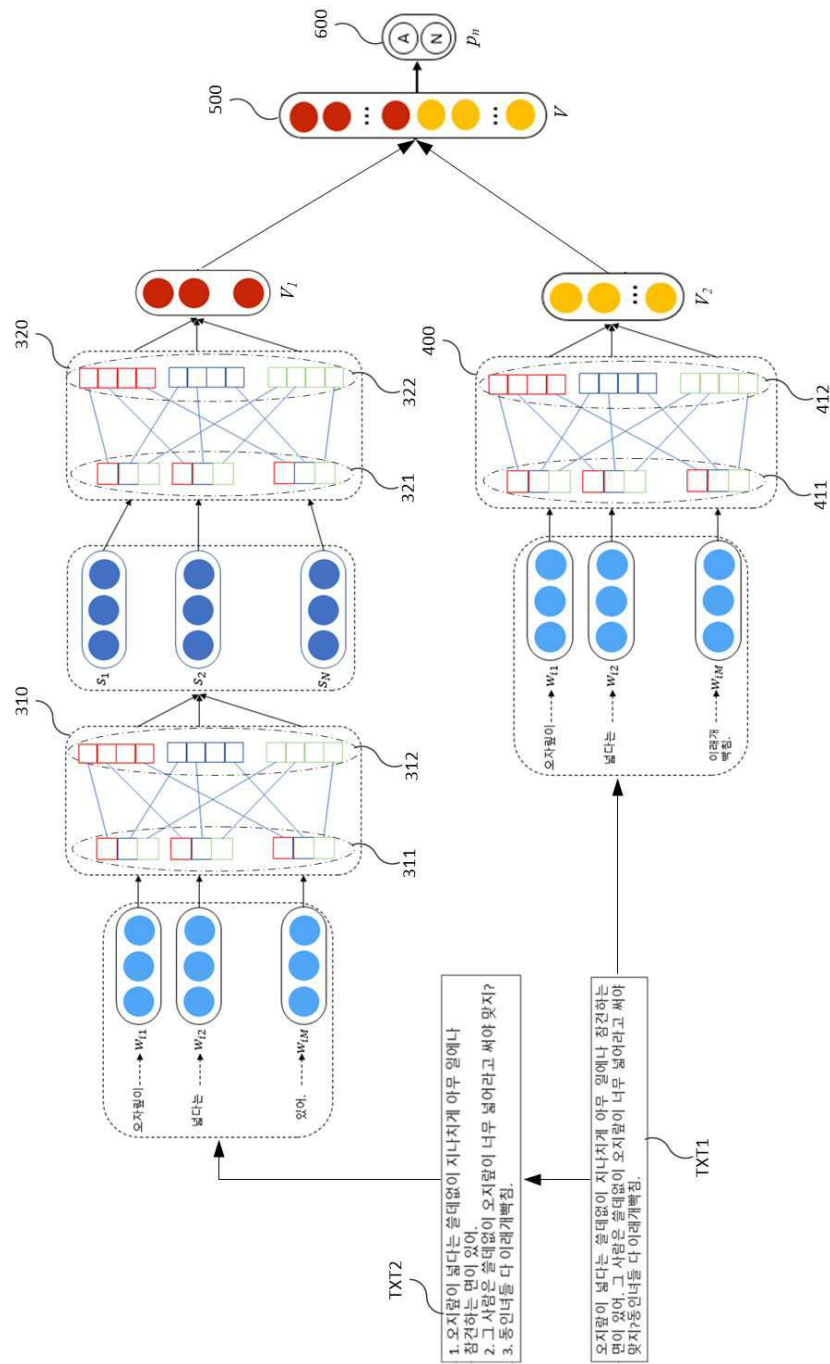
[0061] 100: 전처리부 200: 벡터 변환부  
300: 계층적 문맥 벡터 추출부 310: 단어 인코더  
320: 문장 인코더 400: 포괄적 문맥 벡터 추출부  
500: 문맥 벡터 결합부 600: 비윤리 판별부

도면

도면1



도면2



도면3

