



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0071821

(43) 공개일자 2020년06월22일

(51) 국제특허분류(Int. Cl.)

G06F 40/20 (2020.01)

(52) CPC특허분류

G06F 40/226 (2020.01)

G06N 3/08 (2013.01)

(21) 출원번호 10-2018-0152658

(22) 출원일자 2018년11월30일

심사청구일자 2018년11월30일

(71) 출원인

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

정창성

서울특별시 노원구 노원로 214, 6동 905호(하계동, 삼익선경아파트)

서영경

서울특별시 동작구 동작대로29가길 15(사당동)

(74) 대리인

윤귀상

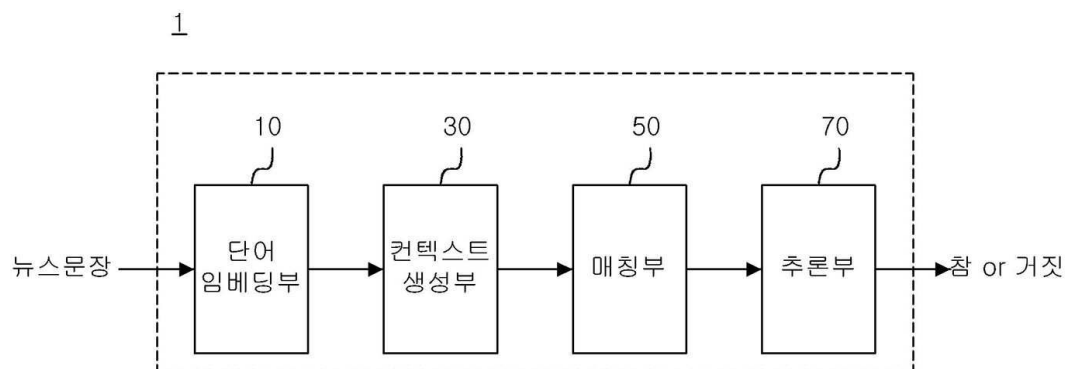
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법, 이를 수행하기 위한 기록매체 및 장치

(57) 요약

신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법, 이를 수행하기 위한 기록매체 및 장치가 개시된다. 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법은 뉴스 문장에 의해 참 또는 거짓이 판별되는 명제(proposition) 문장을 단어(word) 벡터로 임베딩하는 단계, 상기 단어 벡터를 소정의 자연어 처리 신경망에 입력하여 컨텍스트(context) 벡터를 생성하는 단계, 상기 컨텍스트 벡터를 소정의 자연어 처리 신경망에 입력하여 상기 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장을 생성하는 단계 및 상기 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 단계를 포함한다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 1711073495

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터

연구사업명 인공지능산업원천기술개발

연구과제명 가짜뉴스 판독을 위한 딥러닝 기반 자연어 콘텐츠 가치 판별 기술 개발

기 여 율 1/1

주관기관 주식회사 누아

연구기간 2018.04.01 ~ 2018.12.31

명세서

청구범위

청구항 1

뉴스 문장에 의해 참 또는 거짓이 판별되는 명제(proposition) 문장을 단어(word) 벡터로 임베딩하는 단계;

상기 단어 벡터를 소정의 자연어 처리 신경망에 입력하여 컨텍스트(context) 벡터를 생성하는 단계;

상기 컨텍스트 벡터를 소정의 자연어 처리 신경망에 입력하여 상기 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장을 생성하는 단계; 및

상기 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 단계를 포함하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법.

청구항 2

제1항에 있어서,

상기 뉴스 문장에 의해 참 또는 거짓이 판별되는 명제(proposition) 문장을 단어(word) 벡터로 임베딩하는 단계는,

원-핫 인코딩(One-hot encoding) 방식을 이용하여 상기 명제 문장을 구성하는 각 단어들을 벡터로 변환하는 단계를 포함하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법.

청구항 3

제1항에 있어서,

상기 단어 벡터를 소정의 자연어 처리 신경망에 입력하여 컨텍스트(context) 벡터를 생성하는 단계는,

상기 단어 벡터를 LSTM(Long Short Term Memory) 신경망에 입력하여 상기 컨텍스트 벡터를 생성하는 단계를 포함하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법.

청구항 4

제1항에 있어서,

상기 컨텍스트 벡터를 소정의 자연어 처리 신경망에 입력하여 상기 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장을 생성하는 단계는,

시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 어텐션 메커니즘에 따라 상기 컨텍스트 벡터의 가중 합(Weighted sum)을 계산하여 어텐션(attention) 벡터를 생성하는 단계;

상기 어텐션 벡터와 상기 컨텍스트 벡터의 생성 시 발생하는 히든 스테이트(hidden state) 벡터를 비교하여 매칭 벡터를 예측하는 단계; 및

상기 매칭 벡터를 LSTM(Long Short Term Memory) 신경망에 입력하여 상기 후보 문장을 생성하는 단계를 포함하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법.

청구항 5

제1항에 있어서,

상기 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 단계는,

상기 후보 문장을 구성하는 각 단어에 해당하는 벡터를 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 소프트맥스(softmax) 함수에 입력하여 주어진 개수의 단어를 선정 및 출력하는 단계;

빔 서치 디코더를 이용하여 상기 소프트맥스(softmax) 함수의 출력 단어를 조합한 최종 후보 문장을 생성하는

단계; 및

상기 최종 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 단계를 포함하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법.

청구항 6

제5항에 있어서,

상기 최종 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 단계는,

상기 최종 후보 문장 및 상기 뉴스 문장을 임베딩하는 단계;

임베딩한 상기 최종 후보 문장 및 상기 뉴스 문장의 코사인 유사도를 계산하는 단계; 및

상기 코사인 유사도를 이용하여 상기 뉴스 문장에 의한 상기 명제 문장의 참 또는 거짓을 판별하는 단계를 포함하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법.

청구항 7

제1항 내지 제6항 중 어느 하나의 항에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법을 수행하기 위한, 컴퓨터 프로그램이 기록된 컴퓨터로 판독 가능한 기록 매체.

청구항 8

뉴스 문장에 의해 참 또는 거짓이 판별되는 명제(proposition) 문장을 단어(word) 벡터로 임베딩하는 단어 임베딩부;

상기 단어 벡터를 소정의 자연어 처리 신경망에 입력하여 컨텍스트(context) 벡터를 생성하는 컨텍스트 생성부;

상기 컨텍스트 벡터를 소정의 자연어 처리 신경망에 입력하여 상기 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장을 생성하는 매칭부; 및

상기 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 추론부를 포함하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치.

청구항 9

제8항에 있어서,

상기 단어 임베딩부는,

원-핫 인코딩(One-hot encoding) 방식을 이용하여 상기 명제 문장을 구성하는 각 단어들을 벡터로 변환하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치.

청구항 10

제8항에 있어서,

상기 컨텍스트 생성부는,

상기 단어 벡터를 LSTM(Long Short Term Memory) 신경망에 입력하여 상기 컨텍스트 벡터를 생성하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치.

청구항 11

제8항에 있어서,

상기 매칭부는,

시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 어텐션 매커니즘에 따라 상기 컨텍스트 벡터의 가중 합(Weighted sum)을 계산하여 어텐션(attention) 벡터를 생성하고, 상기 어텐션 벡터와 상기 컨텍스트 벡터의 생성 시 발생하는 히든 스테이트(hidden state) 벡터를 비교하여 매칭 벡터를 예측하며, 상기 매칭 벡터를

LSTM(Long Short Term Memory) 신경망에 입력하여 상기 후보 문장을 생성하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치.

청구항 12

제8항에 있어서,

상기 추론부는,

상기 후보 문장을 구성하는 각 단어에 해당하는 벡터를 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 소프트맥스(softmax) 함수에 입력하여 주어진 개수의 단어를 선정 및 출력하고, 빔 서치 디코더를 이용하여 상기 소프트맥스(softmax) 함수의 출력 단어를 조합한 최종 후보 문장을 생성하며, 상기 최종 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치.

청구항 13

제12항에 있어서,

상기 추론부는,

상기 최종 후보 문장 및 상기 뉴스 문장을 임베딩하고, 임베딩한 상기 최종 후보 문장 및 상기 뉴스 문장의 코사인 유사도를 계산하며, 상기 코사인 유사도를 이용하여 상기 뉴스 문장에 의한 상기 명제 문장의 참 또는 거짓을 판별하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치.

발명의 설명

기술 분야

[0001] 본 발명은 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법, 이를 수행하기 위한 기록매체 및 장치에 관한 것으로, 보다 상세하게는 딥러닝에 기반하여 문장을 생성하고, 이를 이용하여 뉴스의 진위 여부를 판별하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법, 이를 수행하기 위한 기록매체 및 장치에 관한 것이다.

배경 기술

[0002] 고속 이동통신망 및 스마트폰이 널리 보급된 모바일 시대가 도래하면서, 각종 사회관계망서비스(SNS)의 사용이 급속도로 증가하고 있다. 특히, 근래에 들어 블로그(blog), 카카오톡(KakaoTalk), 라인(Line), 페이스북(Facebook), 트위터(Twitter), 인스타그램(Instagram) 및 텀블러(Tumblr)와 같은 SNS의 사용이 급증하면서, 각종 SNS를 통한 정보 또는 뉴스의 전달 역시 폭발적으로 증가하고 있다.

[0003] 그러나, 선거와 같은 정치적 행사가 있을 때마다 각종 SNS를 통해 잘못된 정보 또는 뉴스가 퍼지는 경우 또한 적지 않게 발생하는 것이 현실이다. 또한, 더욱 심각한 것은 특정한 목적을 가지고 의도적으로 각종 SNS를 통해 잘못된 정보 또는 뉴스를 전파하는 경우도 빈번히 발생하고 있다는 점인데, 이와 같은 이유로 SNS를 통한 정보 또는 뉴스의 전달이 향후 중요한 사회적 문제가 될 가능성이 높다고 볼 수 있다.

[0004] 한편 위와 같이 특정한 목적을 가지고 의도적으로 조작된 뉴스인 가짜뉴스는 대부분 사람에 의해 탐색되고, 그 진위 여부가 판별된다. 이는 많은 시간과 노력을 필요로 하므로, 사람의 판단을 요구하지 않으며, 가짜뉴스를 실시간으로 선별할 수 있는 새로운 가짜뉴스 탐색 모델이 필요한 실정이다.

발명의 내용

해결하려는 과제

[0005] 본 발명은 딥러닝에 기반하여 명제(proposition) 문장의 문법을 변형시킨 후보 문장을 생성하고, 후보 문장과 뉴스 문장을 비교하여 뉴스의 진위 여부를 판별하는 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법, 이를 수행하기 위한 기록매체 및 장치를 제공한다.

과제의 해결 수단

- [0006] 본 발명에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법은 뉴스 문장에 의해 참 또는 거짓이 판별되는 명제(proposition) 문장을 단어(word) 벡터로 임베딩하는 단계, 상기 단어 벡터를 소정의 자연어 처리 신경망에 입력하여 컨텍스트(context) 벡터를 생성하는 단계, 상기 컨텍스트 벡터를 소정의 자연어 처리 신경망에 입력하여 상기 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장을 생성하는 단계 및 상기 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 단계를 포함한다.
- [0007] 한편, 상기 뉴스 문장에 의해 참 또는 거짓이 판별되는 명제(proposition) 문장을 단어(word) 벡터로 임베딩하는 단계는, 원-핫 인코딩(One-hot encoding) 방식을 이용하여 상기 명제 문장을 구성하는 각 단어들을 벡터로 변환하는 단계를 포함할 수 있다.
- [0008] 또한, 상기 단어 벡터를 소정의 자연어 처리 신경망에 입력하여 컨텍스트(context) 벡터를 생성하는 단계는, 상기 단어 벡터를 LSTM(Long Short Term Memory) 신경망에 입력하여 상기 컨텍스트 벡터를 생성하는 단계를 포함할 수 있다.
- [0009] 또한, 상기 컨텍스트 벡터를 소정의 자연어 처리 신경망에 입력하여 상기 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장을 생성하는 단계는, 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 어텐션 매커니즘에 따라 상기 컨텍스트 벡터의 가중 합(Weighted sum)을 계산하여 어텐션(attention) 벡터를 생성하는 단계, 상기 어텐션 벡터와 상기 컨텍스트 벡터의 생성 시 발생하는 히든 스테이트(hidden state) 벡터를 비교하여 매칭 벡터를 예측하는 단계 및 상기 매칭 벡터를 LSTM(Long Short Term Memory) 신경망에 입력하여 상기 후보 문장을 생성하는 단계를 포함할 수 있다.
- [0010] 또한, 상기 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 단계는, 상기 후보 문장을 구성하는 각 단어에 해당하는 벡터를 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 소프트맥스(softmax) 함수에 입력하여 주어진 개수의 단어를 선정 및 출력하는 단계, 빔 서치 디코더를 이용하여 상기 소프트맥스(softmax) 함수의 출력 단어를 조합한 최종 후보 문장을 생성하는 단계 및 상기 최종 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 단계를 포함할 수 있다.
- [0011] 또한, 상기 최종 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 단계는, 상기 최종 후보 문장 및 상기 뉴스 문장을 임베딩하는 단계, 임베딩한 상기 최종 후보 문장 및 상기 뉴스 문장의 코사인 유사도를 계산하는 단계 및 상기 코사인 유사도를 이용하여 상기 뉴스 문장에 의한 상기 명제 문장의 참 또는 거짓을 판별하는 단계를 포함할 수 있다.
- [0012] 또한, 상기 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법을 수행하기 위한, 컴퓨터 프로그램이 기록된 컴퓨터로 판독 가능한 기록 매체일 수 있다.
- [0013] 한편, 본 발명에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치는 뉴스 문장에 의해 참 또는 거짓이 판별되는 명제(proposition) 문장을 단어(word) 벡터로 임베딩하는 단어 임베딩부, 상기 단어 벡터를 소정의 자연어 처리 신경망에 입력하여 컨텍스트(context) 벡터를 생성하는 컨텍스트 생성부, 상기 컨텍스트 벡터를 소정의 자연어 처리 신경망에 입력하여 상기 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장을 생성하는 매칭부 및 상기 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜뉴스에 해당하는지를 탐색하는 추론부를 포함한다.
- [0014] 한편, 상기 단어 임베딩부는, 원-핫 인코딩(One-hot encoding) 방식을 이용하여 상기 명제 문장을 구성하는 각 단어들을 벡터로 변환할 수 있다.
- [0015] 또한, 상기 컨텍스트 생성부는, 상기 단어 벡터를 LSTM(Long Short Term Memory) 신경망에 입력하여 상기 컨텍스트 벡터를 생성할 수 있다.
- [0016] 또한, 상기 매칭부는, 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 어텐션 매커니즘에 따라 상기 컨텍스트 벡터의 가중 합(Weighted sum)을 계산하여 어텐션(attention) 벡터를 생성하고, 상기 어텐션 벡터와 상기 컨텍스트 벡터의 생성 시 발생하는 히든 스테이트(hidden state) 벡터를 비교하여 매칭 벡터를 예측하며, 상기 매칭 벡터를 LSTM(Long Short Term Memory) 신경망에 입력하여 상기 후보 문장을 생성할 수 있다.

[0017] 또한, 상기 추론부는, 상기 후보 문장을 구성하는 각 단어에 해당하는 벡터를 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 소프트맥스(softmax) 함수에 입력하여 주어진 개수의 단어를 선정 및 출력하고, 빔 서치 디코더를 이용하여 상기 소프트맥스(softmax) 함수의 출력 단어를 조합한 최종 후보 문장을 생성하며, 상기 최종 후보 문장과 상기 뉴스 문장을 비교하여 상기 명제 문장의 참 또는 거짓을 판별하고 상기 뉴스 문장이 가짜 뉴스에 해당하는지를 탐색할 수 있다.

[0018] 또한, 상기 추론부는, 상기 최종 후보 문장 및 상기 뉴스 문장을 임베딩하고, 임베딩한 상기 최종 후보 문장 및 상기 뉴스 문장의 코사인 유사도를 계산하며, 상기 코사인 유사도를 이용하여 상기 뉴스 문장에 의한 상기 명제 문장의 참 또는 거짓을 판별할 수 있다.

발명의 효과

[0019] 본 발명에 따르면 사람의 판단을 필요로 하지 않으며, 실시간으로 가짜뉴스의 선별 처리가 가능하다. 아울러, 가짜뉴스를 선별함으로써 가짜뉴스가 전파되지 않도록 하여 가짜뉴스로부터 야기될 수 있는 사회적 혼란을 예방할 수 있다.

도면의 간단한 설명

[0020] 도 1은 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치의 블록도이다.

도 2는 도 1에 도시된 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치에 의해 구성되는 LSTM 기반의 시퀀스-투-시퀀스(sequence to sequence) 학습 모델을 보여주는 도면이다.

도 3은 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법의 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0021] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.

[0022] 본 명세서에서 사용된 용어는 실시예들을 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 및/또는 "포함하는(comprising)"은 언급된 구성요소, 단계 및 동작은 하나 이상의 다른 구성요소, 단계 및 동작의 존재 또는 추가를 배제하지 않는다.

[0024] 도 1은 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치의 블록도이다.

[0025] 도 1을 참조하면, 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)는 단어 임베딩부(10), 컨텍스트 생성부(30), 매칭부(50) 및 추론부(70)를 포함할 수 있다.

[0026] 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)는 뉴스 문장의 참 또는 거짓을 판별할 수 있는데, 딥러닝에 기반하여 명제(proposition) 문장의 문법을 변형시킨 후보 문장을 생성하고 이를 이용하여 뉴스의 진위 여부를 판별할 수 있다. 따라서 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)는 사람의 개입 없이 빠르고 효율적인 가짜뉴스 탐색 모델을 제시할 수 있다.

[0027] 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)는 가짜뉴스 탐색을 위한 소프트웨어(어플리케이션)가 설치되어 실행될 수 있으며, 단어 임베딩부(10), 컨텍스트 생성부(30), 매칭부(50) 및 추론부(70)는 가짜뉴스 탐색을 위한 소프트웨어에 의해 제어될 수 있다.

[0028] 단어 임베딩부(10), 컨텍스트 생성부(30), 매칭부(50) 및 추론부(70)의 구성은 통합 모듈로 형성되거나, 하나 이상의 모듈로 이루어질 수 있다. 그러나, 이와 반대로 각 구성은 별도의 모듈로 이루어질 수도 있다.

[0029] 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)는 이동성을 갖거나 고정될 수 있다. 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)는 컴퓨터

(computer), 서버(server) 또는 엔진(engine) 형태일 수 있으며, 디바이스(device), 기구(apparatus), 단말(terminal), UE(user equipment), MS(mobile station), MT(mobile terminal), UT(user terminal), SS(subscriber station), 무선기기(wireless device), PDA(personal digital assistant), 무선 모뎀(wireless modem), 휴대기기(handheld device) 등 다른 용어로 불릴 수 있다.

- [0030] 이하, 도 1에 도시된 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)의 각 구성에 대해 구체적으로 설명한다.
- [0031] 단어 임베딩부(10)는 뉴스 문장으로부터 참 또는 거짓이 판별되는 명제 문장을 단어(word) 벡터로 임베딩할 수 있다.
- [0032] 여기서, 명제 문장은 본 실시예에서의 탐색 대상인 뉴스의 진위 여부를 판별하기 위한 문장으로, 예를 들어, 뉴스 문장에 의해 명제 문장이 거짓인 것으로 판별되는 경우 해당 뉴스는 가짜뉴스로 판별될 수 있다.
- [0033] 단어 임베딩부(10)는 임베딩을 통해 명제 문장으로부터 단어 벡터를 생성할 수 있다. 단어 임베딩부(10)는 임베딩 방식 중 원-핫 인코딩(One-hot encoding) 방식을 채택하여 명제 문장을 구성하는 각 단어들을 벡터로 변환할 수 있다. 이는 본 실시예에서 신경망 문법 변형을 이용하여 명제 문장과 의미는 동일하되 문법은 다른 문장인 후보 문장들을 생성하는데 신경망 알고리즘에 문자 자체를 입력하는 것이 불가능하기 때문이다.
- [0034] 한편, 본 실시예에서 신경망 문법 변형 시 LSTM(Long Short Term Memory) 신경망 기반의 시퀀스-투-시퀀스(sequence to sequence) 학습 모델을 이용할 수 있다. 후술하는 컨텍스트 생성부(30), 매칭부(50) 및 추론부(70)는 LSTM 신경망 기반의 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 각 계층을 구성하여 가짜뉴스 탐색을 수행할 수 있다.
- [0035] LSTM 신경망은 문자와 단어들 사이의 흐름을 처리하는 데에 널리 사용되고 있는 RNN(Recurrent Neural Network) 신경망의 히든 스테이트(hidden state)에 셀 스테이트(cell state)를 추가한 구조이다. RNN 신경망은 히든 스테이트의 히든 노드가 방향을 갖는 엣지로 연결되어 순환구조를 이루는 인공신경망의 한 종류이다. LSTM 신경망에 추가되는 셀 스테이트는 스테이트가 오래 경과하더라도 그라디언트가 비교적 잘 전파될 수 있게 하는 수식을 사용하여 다음 셀에 영향을 줄 수 있다.
- [0036] 시퀀스-투-시퀀스(sequence to sequence) 학습 모델은 LSTM 신경망을 활용하여 구현되는 딥러닝 알고리즘으로, 텍스트와 같이 입력의 차원과 출력의 차원이 고정되지 않은 데이터를 처리하는 데 용이하다. 시퀀스-투-시퀀스(sequence to sequence) 학습 모델은 입력으로 소스 시퀀스(source sequence) 및 타겟 시퀀스(target sequence)를 받을 수 있으며, 이러한 입력은 크게 인코더 및 디코더의 두 파트로 나누어 처리할 수 있다. 인코더는 소스 시퀀스를 고정 크기의 벡터로 변환하고, 디코더는 인코더의 벡터를 타겟 시퀀스로 변환할 수 있다. 이때, 타겟 시퀀스의 예측 과정에서 디코더는 직전에 예측한 결과를 다음 단계의 입력으로 넣어 예측할 수 있다. 시퀀스-투-시퀀스(sequence to sequence) 학습 모델에서의 시퀀스 예측 정확도를 높이기 위해, 중요한 단어에 임팩트를 주는 어텐션(attention) 매커니즘을 이용할 수도 있다.
- [0037] 또한, 시퀀스-투-시퀀스(sequence to sequence) 학습 모델은 소프트맥스(softmax) 계층을 포함하여 최종적으로 단어를 디코딩하여 문장을 생성하는데, 이때, 디코더를 사용하여 높은 확률의 단어로 이루어지는 문장들을 출력할 수 있다. 이러한 디코더의 종류는 여러 가지 제안된 바 있으나, 본 실시예에 적용되는 시퀀스-투-시퀀스(sequence to sequence) 학습 모델은 빔 서치 디코더(Beam Search Decoder)를 포함하여 구현될 수 있다. 빔 서치 디코더는 빔 사이즈에 따라 가장 높은 확률의 단어들을 주어진 개수만큼 생성하여 최적 조합의 문장을 만들 수 있다.
- [0038] 컨텍스트 생성부(30)는 단어 벡터를 소정의 자연어 처리 신경망에 입력하여 컨텍스트(context) 벡터를 생성할 수 있다. 여기서 자연어 처리 신경망은 상술한 것처럼 LSTM 신경망일 수 있다.
- [0039] 매칭부(50)는 컨텍스트 벡터를 이용하여 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장들을 생성할 수 있다.
- [0040] 매칭부(50)는 후보 문장 생성의 정확도를 높일 수 있도록 상술한 바와 같은 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 어텐션 매커니즘을 적용할 수 있다.
- [0041] 즉, 매칭부(50)는 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 어텐션 매커니즘을 이용하여 컨텍스트 벡터로부터 어텐션 벡터를 생성할 수 있다. 매칭부(50)는 어텐션 매커니즘에 따라 컨텍스트 벡터의 가중 합

(weighted sum)을 계산하여 중요한 단어에 임팩트를 준 어텐션 벡터를 생성할 수 있다.

- [0042] 매칭부(50)는 어텐션 벡터 및 컨텍스트 생성부(30)에서 컨텍스트 벡터의 생성 시 발생하는 히든 스테이트 벡터를 비교하는 매칭 동작을 수행하여 매칭 벡터를 예측할 수 있다. 상술한 바와 같이 컨텍스트 생성부(30)는 LSTM 신경망을 이용하여 컨텍스트 벡터를 생성하는데, LSTM 신경망의 히든 스테이트에서 히든 스테이트 벡터가 생성될 수 있다. 매칭부(50)는 어텐션 벡터 및 히든 스테이트 벡터를 비교하는 매칭 동작을 수행하여 매칭 벡터를 예측할 수 있다.
- [0043] 매칭부(50)는 매칭 벡터를 소정의 자연어 처리 신경망에 입력하여 후보 문장을 생성할 수 있다. 여기서 자연어 처리 신경망은 상술한 것처럼 LSTM 신경망일 수 있다.
- [0044] 추론부(70)는 후보 문장과 뉴스 문장을 비교하여 명제 문장의 참 또는 거짓을 판별하고 뉴스 문장이 가짜뉴스에 해당하는지를 탐색할 수 있다.
- [0045] 추론부(70)는 최적 조합의 후보 문장과 뉴스 문장을 비교하여 가짜뉴스 선별의 정확도를 높일 수 있도록 상술한 바와 같은 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 소프트맥스(softmax) 계층을 이용할 수 있다.
- [0046] 즉, 추론부(70)는 후보 문장을 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 소프트맥스(softmax) 함수에 입력하여 주어진 개수의 단어를 선정 및 출력할 수 있다. 여기서 후보 문장은 매칭부(50)의 출력인 연계된(Concatenated) 합계 매칭(aggregated matching) 벡터에 해당할 수 있으며, 추론부(70)는 후보 문장을 구성하는 각 단어에 해당하는 합계 매칭 벡터를 소프트맥스(softmax) 함수에 입력하여 주어진 개수의 단어를 선정 및 출력할 수 있다. 이때 출력되는 단어는 뉴스 문장에 나타날 확률이 높은 단어일 수 있다.
- [0047] 추론부(70)는 소프트맥스(softmax) 함수의 출력을 빔 서치 디코더에 입력하여 최종 후보 문장을 생성할 수 있다. 상술한 것처럼 빔 서치 디코더는 소프트맥스(softmax) 함수의 출력인 단어들을 완성도 있는 조합으로 구성하여 문장을 생성할 수 있다.
- [0048] 추론부(70)는 이와 같이 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 소프트맥스(softmax) 계층을 구성하여 보다 완성도 있는 후보 문장을 생성할 수 있다.
- [0049] 추론부(70)는 뉴스 문장에 의한 명제 문장의 참 또는 거짓을 판별하기 위해, 뉴스 문장과 명제 문장의 문법이 변형된 문장인 후보 문장을 비교할 수 있다.
- [0050] 이를 위해 추론부(70)는 뉴스 문장 및 빔 서치 디코더의 출력인 후보 문장을 임베딩할 수 있다. 예를 들어, 추론부(70)는 Doc2Vec 모델을 이용하여 뉴스 문장 및 후보 문장을 임베딩하여 벡터로 변환할 수 있다. 이때, 빔 서치 디코더의 출력인 후보 문장은 복수 개일 수 있으며, 이와 같은 경우 추론부(70)는 복수 개의 후보 문장들을 모아 문단(sentence) 그룹을 구성할 수 있다.
- [0051] 추론부(70)는 임베딩한 뉴스 문장 및 후보 문장의 코사인 유사도를 계산할 수 있다. 추론부(70)는 코사인 유사도를 이용하여 뉴스 문장에 의한 명제 문장의 참 또는 거짓을 판별할 수 있다. 예를 들어, 코사인 유사도는 0~1 사이의 값을 가질 수 있으며, 추론부(70)는 코사인 유사도가 0.5 이상인 경우, 명제 문장을 참으로 판별할 수 있다.
- [0052] 추론부(70)는 문단 그룹을 구성한 경우, 문단 그룹의 각 후보 문장들과 뉴스 문장의 코사인 유사도를 모두 계산할 수 있다. 추론부(70)는 문단 그룹의 각 후보 문장들과 뉴스 문장의 코사인 유사도 중 가장 높은 코사인 유사도를 이용하여 뉴스 문장에 의한 명제 문장의 참 또는 거짓을 판별할 수 있다.
- [0053] 추론부(70)는 뉴스 문장에 의한 명제 문장이 참으로 판별된 경우, 해당 뉴스 또한 진실인 것으로 선별할 수 있으며, 뉴스 문장에 의한 명제 문장이 거짓으로 판별된 경우, 해당 뉴스 또한 가짜뉴스인 것으로 선별할 수 있다.
- [0054] 이와 같이, 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)는 LSTM 기반의 시퀀스-투-시퀀스(sequence to sequence) 학습 모델을 구성하여, 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장들을 생성하고, 후보 문장들과 뉴스 문장을 비교하여 가짜뉴스를 선별할 수 있다.
- [0055] 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)는 사람의 판단을 필요로 하지 않으며, 실시간으로 가짜뉴스 선별 처리가 가능하다.

- [0057] 도 2는 도 1에 도시된 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치에 의해 구성되는 LSTM 기반의 시퀀스-투-시퀀스(sequence to sequence) 학습 모델을 보여주는 도면이다.
- [0058] 도 2를 참조하면, LSTM 기반의 시퀀스-투-시퀀스(sequence to sequence) 학습 모델은 단어 임베딩 레이어(Word Embedding Layer), 컨텍스트 생성 레이어(Context Generation Layer), 매칭 레이어(Matching Layer) 및 추론 레이어(Inference Layer)의 4 개의 계층으로 이루어질 수 있다.
- [0059] 단어 임베딩부(10)는 단어 임베딩 레이어(Word Embedding Layer)를 구성하며, 뉴스 문장으로부터 참 또는 거짓 이 판별되는 명제 문장을 단어 벡터로 임베딩할 수 있다.
- [0060] 컨텍스트 생성부(30)는 컨텍스트 생성 레이어(Context Generation Layer)를 구성하며, 단어 벡터를 LSTM 신경망에 입력하여 컨텍스트 벡터를 생성할 수 있다.
- [0061] 매칭부(50)는 매칭 레이어(Matching Layer)를 구성하며, 컨텍스트 벡터에 대한 매칭 동작을 수행하여 매칭 벡터를 생성할 수 있다. 이때, 매칭부(50)는 컨텍스트 벡터를 어텐션 함수에 입력하여 어텐션 벡터를 생성하고, 어텐션 벡터와 컨텍스트 생성 레이어(Context Generation Layer)에 포함되는 LSTM 신경망의 히든 스테이트 벡터를 비교하는 매칭 동작을 수행하여 매칭 벡터를 생성할 수 있다. 그리고, 매칭부(50)는 매칭 벡터를 LSTM 신경망에 입력하여 명제 문장과 의미는 동일하되 문법이 다른 문장인 후보 문장에 해당하는 연계된 합계 매칭 벡터를 생성할 수 있다.
- [0062] 추론부(70)는 추론 레이어(Inference Layer)를 구성하며, 연계된 합계 매칭 벡터로부터 합계 매칭 벡터를 획득하고 이를 소프트맥스(softmax) 함수에 입력하여 주어진 개수의 단어를 선정 및 출력할 수 있다. 그리고, 추론부(70)는 소프트맥스(softmax) 함수의 출력을 빔 서치 디코더에 입력하여 복수의 후보 문장들을 생성할 수 있다. 추론부(70)는 복수의 후보 문장들과 뉴스 문장들을 비교하여 뉴스 문장에 의한 명제 문장의 참 또는 거짓을 판별하고, 그 결과에 따라 가짜뉴스를 선별할 수 있다.
- [0064] 이하에서는 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법에 대해 설명한다.
- [0065] 도 3은 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법의 흐름도이다.
- [0066] 본 발명의 일 실시예에 따른 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 방법은 도 1에 도시된 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)와 실질적으로 동일한 구성에서 실행될 수 있다. 따라서 도 1에 도시된 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치(1)와 동일한 구성요소는 동일한 도면부호를 부여하고, 반복되는 설명은 생략한다.
- [0067] 도 3을 참조하면, 단어 임베딩부(10)는 명제 문장을 단어 벡터로 임베딩할 수 있다(S100).
- [0068] 명제 문장은 뉴스 문장에 의해 참 또는 거짓이 판별될 문장에 해당한다. 단어 임베딩부(10)는 임베딩 방식 중 원-핫 인코딩(One-hot encoding) 방식을 채택하여 명제 문장을 구성하는 각 단어들을 벡터로 변환할 수 있다.
- [0069] 컨텍스트 생성부(30)는 단어 벡터를 신경망에 입력하여 컨텍스트 벡터를 생성할 수 있다(S200).
- [0070] 여기서 신경망은 자연어 처리 신경망으로, LSTM 신경망일 수 있다.
- [0071] 매칭부(50)는 컨텍스트 벡터를 신경망에 입력하여 후보 문장을 생성할 수 있다(S300).
- [0072] 후보 문장은 명제 문장과 의미는 동일하되 문법이 다른 문장에 해당한다.
- [0073] 매칭부(50)는 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 어텐션 매커니즘을 이용하여 컨텍스트 벡터로부터 어텐션 벡터를 생성할 수 있다. 매칭부(50)는 어텐션 매커니즘에 따라 컨텍스트 벡터의 가중 합(weighted sum)을 계산하여 중요한 단어에 임팩트를 준 어텐션 벡터를 생성할 수 있다.
- [0074] 매칭부(50)는 어텐션 벡터 및 컨텍스트 벡터의 생성 시 발생하는 히든 스테이트 벡터를 비교하는 매칭 동작을 수행하여 매칭 벡터를 예측할 수 있다.
- [0075] 매칭부(50)는 매칭 벡터를 LSTM 신경망에 입력하여 후보 문장을 생성할 수 있다.
- [0076] 추론부(70)는 후보 문장과 뉴스 문장을 비교하여 가짜뉴스를 탐색할 수 있다(S400).
- [0077] 추론부(70)는 후보 문장과 뉴스 문장을 비교하여 명제 문장의 참 또는 거짓을 판별하고 뉴스 문장이 가짜뉴스에

해당하는지를 탐색할 수 있다.

- [0078] 추론부(70)는 후보 문장을 시퀀스-투-시퀀스(sequence to sequence) 학습 모델의 소프트맥스(softmax) 함수에 입력하여 주어진 개수의 단어를 선정 및 출력할 수 있다. 여기서 후보 문장은 매칭부(50)의 출력인 연계된(Concatenated) 합계 매칭(aggregated matching) 벡터에 해당할 수 있으며, 추론부(70)는 후보 문장을 구성하는 각 단어에 해당하는 합계 매칭 벡터를 소프트맥스(softmax) 함수에 입력하여 주어진 개수의 단어를 선정 및 출력할 수 있다. 이때 출력되는 단어는 뉴스 문장에 나타날 확률이 높은 단어일 수 있다.
- [0079] 추론부(70)는 소프트맥스(softmax) 함수의 출력을 빔 서치 디코더에 입력하여 최종 후보 문장을 생성할 수 있다. 상술한 것처럼 빔 서치 디코더는 소프트맥스(softmax) 함수의 출력인 단어들을 완성도 있는 조합으로 구성하여 문장을 생성할 수 있다.
- [0080] 추론부(70)는 뉴스 문장 및 빔 서치 디코더의 출력인 후보 문장을 임베딩할 수 있다. 예를 들어, 추론부(70)는 Doc2Vec 모델을 이용하여 뉴스 문장 및 후보 문장을 임베딩할 수 있다. 이때, 빔 서치 디코더의 출력인 후보 문장은 복수 개일 수 있으며, 이와 같은 경우 추론부(70)는 복수 개의 후보 문장들을 모아 문단(sentence) 그룹을 구성할 수 있다.
- [0081] 추론부(70)는 임베딩한 뉴스 문장 및 후보 문장의 코사인 유사도를 계산할 수 있다. 추론부(70)는 코사인 유사도를 이용하여 뉴스 문장에 의한 명제 문장의 참 또는 거짓을 판별할 수 있다. 예를 들어, 코사인 유사도는 0~1 사이의 값을 가질 수 있으며, 추론부(70)는 코사인 유사도가 0.5 이상인 경우, 명제 문장을 참으로 판별할 수 있다.
- [0082] 추론부(70)는 문단 그룹을 구성한 경우, 문단 그룹의 각 후보 문장들과 뉴스 문장의 코사인 유사도를 모두 계산할 수 있다. 추론부(70)는 문단 그룹의 각 후보 문장들과 뉴스 문장의 코사인 유사도 중 가장 높은 코사인 유사도를 이용하여 뉴스 문장에 의한 명제 문장의 참 또는 거짓을 판별할 수 있다.
- [0084] 이와 같은 본 발명의 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색방법은 다양한 컴퓨터 구성요소를 통하여 수행될 수 있는 프로그램 명령어의 형태로 구현되어 컴퓨터 판독 가능한 기록 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능한 기록 매체는 프로그램 명령어, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다.
- [0085] 상기 컴퓨터 판독 가능한 기록 매체에 기록되는 프로그램 명령어는 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 분야의 당업자에게 공지되어 사용 가능한 것일 수도 있다.
- [0086] 컴퓨터 판독 가능한 기록 매체의 예에는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM, DVD 와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 ROM, RAM, 플래시 메모리 등과 같은 프로그램 명령어를 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다.
- [0087] 프로그램 명령어의 예에는, 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드도 포함된다. 상기 하드웨어 장치는 본 발명에 따른 처리를 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0088] 이상 첨부된 도면을 참조하여 본 발명의 실시예를 설명하였지만, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 본 발명이 그 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다.

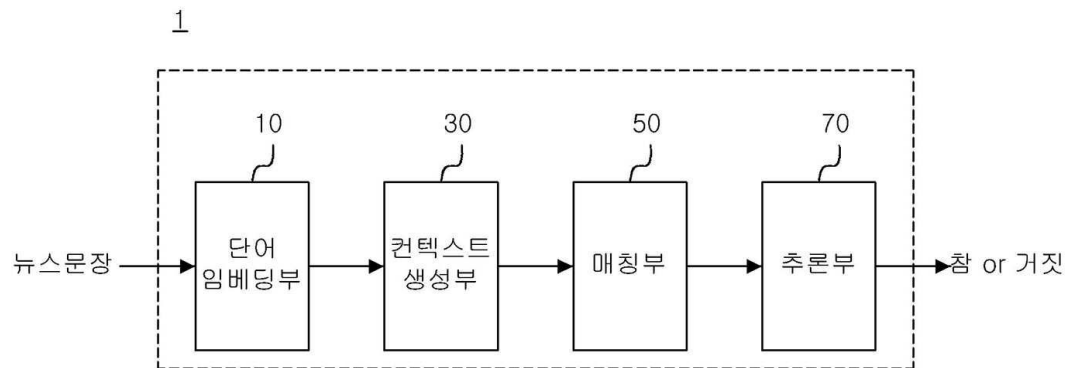
부호의 설명

- [0089] 1: 신경망 상의 문법 변형을 이용한 가짜뉴스 탐색 장치
10: 단어 임베딩부
30: 컨텍스트 생성부
50: 매칭부

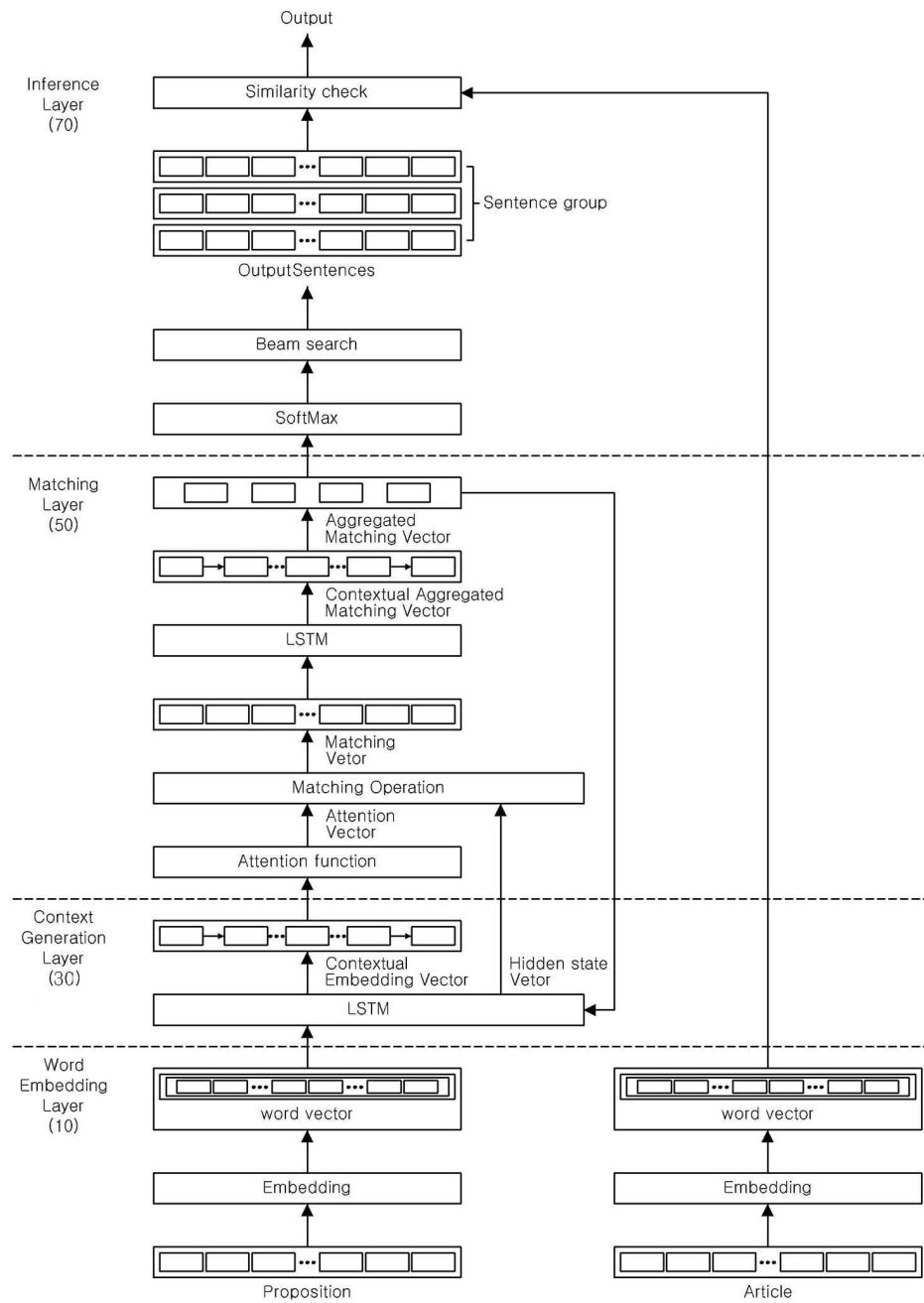
70: 추론부

도면

도면1



도면2



도면3

