



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0066707
(43) 공개일자 2023년05월16일

- | | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (51) 국제특허분류(Int. Cl.)
G06F 16/332 (2019.01) G06F 16/2457 (2019.01)
G06F 16/33 (2019.01) G06N 20/00 (2019.01)
G06N 3/04 (2023.01) G06N 5/04 (2023.01)
(52) CPC특허분류
G06F 16/3329 (2019.01)
G06F 16/24578 (2019.01)
(21) 출원번호 10-2021-0151973
(22) 출원일자 2021년11월08일
심사청구일자 2021년11월08일 | (71) 출원인
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
(72) 발명자
이근배
서울특별시 서초구 서운로 221, 103동 1203호
이지형
서울특별시 은평구 불광로2길 16, 110동 702호
(74) 대리인
특허법인이상 |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

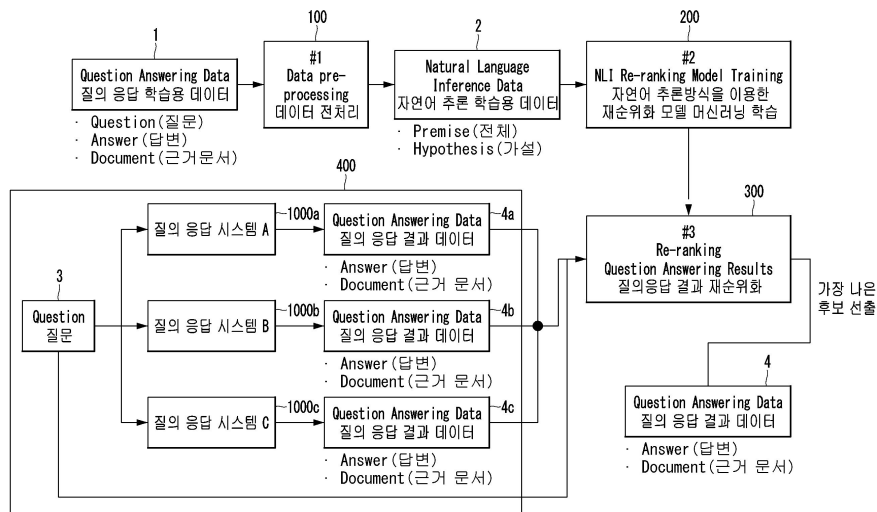
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치 및 방법

(57) 요약

본 발명의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치는, 프로세서(processor); 및 프로세서를 통해 실행되는 적어도 하나의 명령이 저장된 메모리(memory); 를 포함하고, 상기 적어도 하나의 명령은 상기 프로세서가: 질의응답 데이터를 자연어 추론 형식으로 변환하는 단계, 자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계, 및 훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계를 수행하도록 구성된다.

대표도



(52) CPC특허분류

G06F 16/3344 (2019.01)

G06N 20/00 (2021.08)

G06N 3/044 (2023.01)

G06N 5/04 (2023.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711134561
과제번호	2021-0-00354-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	차세대인공지능핵심원천기술개발(R&D)
연구과제명	비정형 텍스트를 학습하여 쟁점별 사실과 논리적 근거 추론이 가능한 인공지능 원천

기술

기 여 율	1/2
과제수행기관명	창원대학교산학협력단
연구기간	2021.04.01 ~ 2021.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711125943
과제번호	2019-0-01906-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성(R&D)
연구과제명	인공지능대학원지원(포항공과대학교)

기 여 율	1/2
과제수행기관명	포항공과대학교 산학협력단
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

프로세서(processor); 및

프로세서를 통해 실행되는 적어도 하나의 명령이 저장된 메모리(memory); 를 포함하고,

상기 적어도 하나의 명령은 상기 프로세서가:

질의응답 데이터를 자연어 추론 형식으로 변환하는 단계;

자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계; 및

훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계; 를 수행하도록 구성되는,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치.

청구항 2

청구항 1에 있어서, 상기 장치는,

질의응답 학습용 데이터의 질문(Question), 답변(Answer), 근거 문서(Document) 데이터를 입력받아서 데이터 전처리(Data-preprocessing)를 실행하는 데이터 전처리부; 를 포함하는,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치.

청구항 3

청구항 2에 있어서, 질의응답 학습용 데이터는,

특정 질의에 대한 정답 및 근거문서가 전부 표시되어 있는 신경망 모델 학습용 데이터이고,

멀티-홉 추론을 위한 질의 응답 학습용 데이터인 HotpotQA 데이터는 특정 질문에 대한 근거 문서와 답변을 포함하는,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치.

청구항 4

청구항 1에 있어서, 상기 장치는,

데이터 전처리부에서 데이터 전처리(Data-preprocessing)되어 자연어 추론 형식으로 변환된 자연어 추론 학습용 데이터를 입력받아서, 자연어 추론방식을 이용하여 재순위화 모델 머신러닝 학습을 수행하는 재순위화 모델 머신러닝 학습부; 를 포함하는,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치.

청구항 5

청구항 4에 있어서, 자연어 추론 학습용 데이터는,

자연어 추론 형식으로 변환된 질의응답 결과물로서 질의 - 근거문서 - 정답 - 라벨 쌍으로 이루어져 있으며, 라벨은 포함함(entailment), 및 포함하지 않음(non-entailment) 두 종류로 구성된,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치.

청구항 6

청구항 1에 있어서, 상기 장치는,

재순위화 모델 머신러닝 학습부에서 학습된 질의응답 결과 재순위 모델을 이용하여 기존의 질의 응답 결과 데이터를 재순위화하여 최종적인 질의 응답 결과 데이터를 산출하는 질의응답 결과 재순위화부; 를 포함하는, 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치.

청구항 7

청구항 6에 있어서, 기존의 질의 응답 결과 데이터는, 질문(Question)에 대한 제1, 제2, 제3 질의 응답 시스템에서 각각 생성한 제1, 제2, 제3 질의 응답 결과 데이터인, 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치.

청구항 8

청구항 1에 있어서, 상기 장치는, 자연어 추론 학습 시, 사전 학습된 언어 모델로 ELECTRA를 이용하고, 근거 문서 SD 와 질문 SQ, 답변 SA 을 CLS, SEP 토큰을 덧붙여 수학적 1에 따라 ELECTRA 인코더에 입력하는, 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치.

(수학적 1)

$$x = ([CLS], SD, [SEP], SQ, SA)$$

청구항 9

청구항 1에 있어서, 상기 장치는, 인코더의 출력 결과에서 CLS에 해당하는 첫 번째 벡터를 풀링 후 Classifier에 입력하여 문서와 질문, 정답의 관계를 함의 혹은 중립 두 가지의 라벨로 판별한 뒤 정확도를 향상하는 훈련으로 학습하고, 재순위화 시에는 질의응답 시스템이 생성한 질문, 답변, 근거문서를 학습한 자연어 추론 모델에 입력하여 함의 라벨의 점수에 따라 최고점의 답변 후보를 최종 산출하는, 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치.

청구항 10

특정한 질의에 대해서 복수의 질의응답 시스템이 생성한 근거 문서 및 답변 후보들 중에서 적어도 하나의 후보를 선정하는 자연어 추론(NLI) 시스템을 이용한 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법에 있어서, 질의응답 데이터를 자연어 추론 형식으로 변환하는 단계; 자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계; 및 훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계; 를 포함하는, 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법.

청구항 11

청구항 10에 있어서, 질의응답 데이터를 자연어 추론 형식으로 변환하는 단계는, 데이터 전처리부가 질의응답 학습용 데이터의 질문(Question), 답변(Answer), 근거 문서(Document) 데이터를 입력받아서 데이터 전처리(Data-precessing)를 실행하는 단계인, 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법.

청구항 12

청구항 11에 있어서, 질의응답 학습용 데이터는,

특정 질의에 대한 정답 및 근거문서가 전부 표시되어 있는 신경망 모델 학습용 데이터이고,

멀티-홉 추론을 위한 질의 응답 학습용 데이터인 HotpotQA 데이터는 특정 질문에 대한 근거 문서와 답변을 포함하는,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법.

청구항 13

청구항 10에 있어서, 자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계는,

재순위화 모델 머신러닝 학습부가 데이터 전처리부에서 데이터 전처리(Data-precessing)되어 자연어 추론 형식으로 변환된 자연어 추론 학습용 데이터를 입력받아서, 자연어 추론방식을 이용하여 재순위화 모델 머신러닝 학습을 수행하는 단계인,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법.

청구항 14

청구항 13에 있어서, 자연어 추론 학습용 데이터는,

자연어 추론 형식으로 변환된 질의응답 결과물로서 질의 - 근거문서 - 정답 - 라벨 쌍으로 이루어져 있으며,

라벨은 포함함(entailment), 및 포함하지 않음(non-entailment) 두 종류로 구성된,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법.

청구항 15

청구항 10에 있어서, 훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계는,

질의응답 결과 재순위화부가 재순위화 모델 머신러닝 학습부에서 학습된 질의응답 결과 재순위 모델을 이용하여 기존의 질의 응답 결과 데이터를 재순위화하여 최종적인 질의 응답 결과 데이터를 산출하는 단계인,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법.

청구항 16

청구항 15에 있어서, 기존의 질의 응답 결과 데이터는,

질문(Question)에 대한 제1, 제2, 제3 질의 응답 시스템에서 각각 생성한 제1, 제2, 제3 질의 응답 결과 데이터인,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법.

청구항 17

청구항 8에 있어서, 상기 방법은,

자연어 추론 학습 시, 사전 학습된 언어 모델로 ELECTRA를 이용하고,

근거 문서 SD 와 질문 SQ, 답변 SA 을 CLS, SEP 토큰을 덧붙여 수학적 1에 따라 ELECTRA 인코더에 입력하는,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법.

(수학적 1)

$x = ([CLS], SD, [SEP], SQ, SA)$

청구항 18

청구항 8에 있어서, 상기 방법은,

인코더의 출력 결과에서 CLS에 해당하는 첫 번째 벡터를 풀링 후 Classifier에 입력하여 문서와 질문, 정답의 관계를 합의 혹은 중립 두 가지의 라벨로 판별한 뒤 정확도를 향상하는 훈련으로 학습하고,

재순위화 시에는 질의응답 시스템이 생성한 질문, 답변, 근거문서를 학습한 자연어 추론 모델에 입력하여 합의 라벨의 점수에 따라 최고점의 답변 후보를 최종 산출하는,

자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법.

청구항 19

청구항 9 내지 청구항 16 중 어느 한 항의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법을 구현하기 위한 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

청구항 20

청구항 9 내지 청구항 16 중 어느 한 항의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법의 프로그램을 구현하기 위한 컴퓨터 판독 가능한 기록매체.

발명의 설명

기술 분야

[0001] 본 발명은 기계 학습을 활용한 질의 응답 결과 재순위화 기술에 대한 것으로, 자연어 추론 모델을 이용하여 질의 응답 결과를 검증할 수 있는 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치 및 방법에 관한 것이다.

배경 기술

[0002] 질의 응답 결과 재순위화 기술은 하나의 질의에 대해 생성된 여러 정답을 비교하여 가장 나은 후보를 뽑는 것이다.

[0003] 기존의 질의 응답 결과 재순위화 기술은 대체로 하나의 질의 응답 모델이 도출한 여러 정답 중에서 가장 나은 후보를 뽑는 경우가 많았다.

[0004] 그런데 만약 여러 질의 응답 시스템이 생성한 결과물들을 전부 이용하여 재순위화를 한다면, 각 질의응답 시스템이 가지고 있는 약점을 상호 보완하여 더욱 나은 성과를 거둘 수 있을 것이다.

[0005] 자연어 추론 모델은 특정 텍스트 가설이 주어진 전제에 맞는 정보를 포함하는지를 평가한다.

[0006] 이 때, 질의 응답 시스템 결과 검증에 자연어 추론 모델을 이용한다면 질의응답 시스템이 정답을 추출하는데 도움이 된다고 판단한 근거문서가 최종적으로 도출된 정답 및 질의에 대한 정보를 올바르게 포함하고 있는지 검증할 수 있다.

[0007] 이러한 검증 시스템을 통하여 각 질의 응답 결과물에 점수를 매김으로써 여러 질의 응답 시스템이 도출한 정답 및 근거문서를 재순위화 할 수 있다.

선행기술문헌

비특허문헌

[0008] (비특허문헌 0001) Y. Fang, S. Sun, Z. Gan, R. Pillai, S. Wang, and J. Liu, “Hierarchical graph network for multi-hop question answering,” arXiv preprint arXiv:1911.03631, 2019.)

(비특허문헌 0002) B. Kratzwald, A. Eigenmann, and S. Feuerriegel, “Rankqa: Neural question answering with answer reranking,” Annual Meeting of the Association for Computational Linguistics (ACL), 2019.

(비특허문헌 0003) Z. Yang, P. Qi, S. Zhang, Y. Bengio, W. W. Cohen, R. Salakhutdinov, and C. D. Manning, “HotpotQA: A dataset for diverse, explainable multi-hop question answering,” Conference on Empirical Methods in Natural Language Processing (EMNLP), 2018.

(비특허문헌 0004) D. Demszky, K. Guu, and P. Liang, “Transforming question answering datasets into natural language inference datasets,” arXiv preprint arXiv:1809.02922, 2018.

(비특허문헌 0005) K. Clark, M.-T. Luong, Q. V. Le, and C. D. Manning, “ELECTRA: Pre-training text encoders as discriminators rather than generators,” ICLR, 2020. [Online]. Available: <https://openreview.net/pdf?id=rlxMH1BtvB>

(비특허문헌 0006) A. Asai, K. Hashimoto, H. Hajishirzi, R. Socher, and C. Xiong, “Learning to retrieve reasoning paths over wikipedia graph for question answering,” International Conference on Learning Representations, 2020.

(비특허문헌 0007) W. Xiong, X. L. Li, S. Iyer, J. Du, P. Lewis, W. Y. Wang, Y. Mehdad, W.-t. Yih, S. Riedel, D. Kiela, and B. O˘guz, “Answering complex open-domain questions

(비특허문헌 0008) M. Tu, K. Huang, G. Wang, J. Huang, X. He, and B. Zhou, “Select, answer and explain: Interpretable multi-hop reading comprehension over multiple documents,” AAAI 2020 (accepted), 2020.)

발명의 내용

해결하려는 과제

[0009] 상기와 같은 문제점을 해결하기 위한 본 발명의 목적은, 본 특허에서는 자연어 추론 모델이 질의응답 결과물을 평가할 수 있다는 점에 착안하여, 자연어 추론 모델을 이용하여 질의에 대해 생성한 정답이 근거 문서 내에 포함되어 있는지 아닌지의 여부를 따짐으로써 여러 질의 응답 시스템이 도출한 정답 및 근거문서를 동시에 재순위화 하는 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치 및 방법을 제공하는 것이다.

과제의 해결 수단

[0010] 본 발명의 목적을 달성하기 위한 일 실시예에 따른 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치는, 프로세서(processor); 및 프로세서를 통해 실행되는 적어도 하나의 명령이 저장된 메모리(memory); 를 포함하고, 상기 적어도 하나의 명령은 상기 프로세서가: 질의응답 데이터를 자연어 추론 형식으로 변환하는 단계; 자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계; 및 훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계; 를 수행하도록 구성될 수 있다.

[0011] 상기 장치는, 질의응답 학습용 데이터의 질문(Question), 답변(Answer), 근거 문서(Document) 데이터를 입력받아서 데이터 전처리(Data-precessing)를 실행하는 데이터 전처리부; 포함할 수 있다.

[0012] 질의응답 학습용 데이터는, 특정 질의에 대한 정답 및 근거문서가 전부 표시되어 있는 신경망 모델 학습용 데이터이고, 멀티-홉 추론을 위한 질의 응답 학습용 데이터인 HotpotQA 데이터는 특정 질문에 대한 근거 문서와 답변을 포함할 수 있다.

[0013] 상기 장치는, 데이터 전처리부에서 데이터 전처리(Data-precessing)되어 자연어 추론 형식으로 변환된 자연어 추론 학습용 데이터를 입력받아서, 자연어 추론방식을 이용하여 재순위화 모델 머신러닝 학습을 수행하는 재순위화 모델 머신러닝 학습부; 를 포함할 수 있다.

[0014] 자연어 추론 학습용 데이터는, 자연어 추론 형식으로 변환된 질의응답 결과물로서 질의 - 근거문서 - 정답 - 라벨 쌍으로 이루어져 있으며, 라벨은 포함함(entailment), 및 포함하지 않음(non-entailment) 두 종류로 구성될 수 있다.

[0015] 상기 장치는, 재순위화 모델 머신러닝 학습부에서 학습된 질의응답 결과 재순위 모델을 이용하여 기존의 질의응답 결과 데이터를 재순위화하여 최종적인 질의 응답 결과 데이터를 산출하는 질의응답 결과 재순위화부; 를 포함할 수 있다.

[0016] 기존의 질의 응답 결과 데이터는, 질문(Question)에 대한 제1, 제2, 제3 질의 응답 시스템에서 각각 생성한 제1, 제2, 제3 질의 응답 결과 데이터일 수 있다.

- [0017] 상기 장치는, 자연어 추론 학습 시, 사전 학습된 언어 모델로 ELECTRA를 이용하고, 근거 문서 SD 와 질문 SQ, 답변 SA 을 CLS, SEP 토큰을 덧붙여 수학적 식 1에 따라 ELECTRA 인코더에 입력할 수 있다.
- [0018] 상기 장치는, 인코더의 출력 결과에서 CLS에 해당하는 첫 번째 벡터를 풀링 후 Classifier에 입력하여 문서와 질문, 정답의 관계를 함의 혹은 중립 두 가지의 라벨로 판별한 뒤 정확도를 향상하는 훈련으로 학습하고, 재순위화 시에는 질의응답 시스템이 생성한 질문, 답변, 근거문서를 학습한 자연어 추론 모델에 입력하여 함의 라벨의 점수에 따라 최고점의 답변 후보를 최종 산출할 수 있다.
- [0019] 본 발명의 다른 목적을 달성하기 위한 일 실시예의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법은, 특정한 질의에 대해서 복수의 질의응답 시스템이 생성한 근거 문서 및 답변 후보들 중에서 적어도 하나의 후보를 선정하는 자연어 추론(NLI) 시스템을 이용한 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법에 있어서, 질의응답 데이터를 자연어 추론 형식으로 변환하는 단계; 자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계; 및 훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계; 를 포함할 수 있다.
- [0020] 질의응답 데이터를 자연어 추론 형식으로 변환하는 단계는, 데이터 전처리부가 질의응답 학습용 데이터의 질문 (Question), 답변(Answer), 근거 문서(Document) 데이터를 입력받아서 데이터 전처리(Data-precessing)를 실행하는 단계일 수 있다.
- [0021] 질의응답 학습용 데이터는, 특정 질의에 대한 정답 및 근거문서가 전부 표시되어 있는 신경망 모델 학습용 데이터이고, 멀티-홉 추론을 위한 질의 응답 학습용 데이터인 HotpotQA 데이터는 특정 질문에 대한 근거 문서와 답변을 포함할 수 있다.
- [0022] 자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계는, 재순위화 모델 머신러닝 학습부가 데이터 전처리부에서 데이터 전처리(Data-precessing)되어 자연어 추론 형식으로 변환된 자연어 추론 학습용 데이터를 입력받아서, 자연어 추론방식을 이용하여 재순위화 모델 머신러닝 학습을 수행하는 단계일 수 있다.
- [0023] 자연어 추론 학습용 데이터는, 자연어 추론 형식으로 변환된 질의응답 결과물로서 질의 - 근거문서 - 정답 - 라벨 쌍으로 이루어져 있으며, 라벨은 포함함(entailment), 및 포함하지 않음(non-entailment) 두 종류로 구성될 수 있다.
- [0024] 훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계는, 질의응답 결과 재순위화부가 재순위화 모델 머신러닝 학습부에서 학습된 질의응답 결과 재순위화 모델을 이용하여 기존의 질의 응답 결과 데이터를 재순위화하여 최종적인 질의 응답 결과 데이터를 산출하는 단계일 수 있다.
- [0025] 기존의 질의 응답 결과 데이터는, 질문(Question)에 대한 제1, 제2, 제3 질의 응답 시스템에서 각각 생성한 제1, 제2, 제3 질의 응답 결과 데이터일 수 있다.
- [0026] 상기 방법은, 자연어 추론 학습 시, 사전 학습된 언어 모델로 ELECTRA를 이용하고, 근거 문서 SD 와 질문 SQ, 답변 SA 을 CLS, SEP 토큰을 덧붙여 수학적 식 1에 따라 ELECTRA 인코더에 입력할 수 있다.
- [0027] 상기 방법은, 인코더의 출력 결과에서 CLS에 해당하는 첫 번째 벡터를 풀링 후 Classifier에 입력하여 문서와 질문, 정답의 관계를 함의 혹은 중립 두 가지의 라벨로 판별한 뒤 정확도를 향상하는 훈련으로 학습하고, 재순위화 시에는 질의응답 시스템이 생성한 질문, 답변, 근거문서를 학습한 자연어 추론 모델에 입력하여 함의 라벨의 점수에 따라 최고점의 답변 후보를 최종 산출할 수 있다.
- [0028] 전술한 항 중 어느 한 항의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법을 구현하기 위한 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램일 수 있다.
- [0029] 전술한 항 중 어느 한 항의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법의 프로그램을 구현하기 위한 컴퓨터 판독 가능한 기록매체일 수 있다.

발명의 효과

- [0030] 본 발명을 통해, 자연어 추론 모델을 이용하기 때문에 자연어 텍스트의 심층적인 의미를 판단할 수 있으며, 멀티-홉 질의응답 시스템과 같은 높은 난이도의 질의응답 결과에 대해서도 유연하게 대처할 수 있다.

도면의 간단한 설명

- [0031] 도 1은 본 발명의 일 실시예에 따른 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치의 구성도이다.
- 도 2는 본 발명의 일 실시예에 따른 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치의 질의응답 결과 재순위화부의 구성도이다.
- 도 3은 본 발명의 일 실시예의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법의 순서도이다.
- 도 4는 본 발명의 HotpotQA 데이터 학습을 위한 자연어 추론 모델을 도시한 구조도이다.
- 도 5는 본 발명의 일 실시예의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치의 구성도이다.

발명을 실시하기 위한 구체적인 내용

- [0032] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하여 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.
- [0033] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는 데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. "및/또는"이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0034] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [0035] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 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.
- [0036] 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 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0037] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다. 본 발명을 설명함에 있어 전체적인 이해를 용이하게 하기 위하여 도면상의 동일한 구성요소에 대해서는 동일한 참조부호를 사용하고 동일한 구성요소에 대해서 중복된 설명은 생략한다.
- [0039] 도 1은 본 발명의 일 실시예에 따른 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)를 설명하기 위한 구성도이다.
- [0040] 도 1을 참조하면, 본 발명의 일 실시예에 따른 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)는, 데이터 전처리부(100), 재순위화 모델 머신러닝 학습부(200), 질의응답 결과 재순위화부(300)를 포함하여 구성된다.
- [0041] 데이터 전처리부(100)는, 질의응답 학습용 데이터(1)의 질문(Question), 답변(Answer), 근거 문서(Document) 데이터를 입력받아서 데이터 전처리(Data-precessing) 한다.

- [0042] 질의응답 학습용 데이터(Question Answering Data)(1)는 특정 질의에 대한 정답 및 근거문서가 전부 표시되어 있는 신경망 모델 학습용 데이터이고, 그 예시로 HotpotQA 데이터(<https://hotpotqa.github.io/>)가 있다. HotpotQA 데이터는 멀티-홉 추론을 위한 질의 응답 데이터로, 특정 질문에 대한 근거 문서와 답변을 포함한다.
- [0043] 재순위화 모델 머신러닝 학습부(200)는, 데이터 전처리부(100)에서 데이터 전처리(Data-precessing)되어 자연어 추론 형식으로 변환된 자연어 추론 학습용 데이터(2)를 입력받아서, 자연어 추론방식을 이용하여 재순위화 모델 머신러닝 학습(NLI Re-ranking Model Training)을 수행한다.
- [0044] 자연어 추론 학습용 데이터(natural Language Inference Data)(2)는, 자연어 추론 형식으로 변환된 질의응답 결과물로서 질의 - 근거문서 - 정답 - 라벨 쌍으로 이루어져 있으며, 라벨은 포함함(entailment), 및 포함하지 않음(non-entailment) 두 종류로 이루어져 있다.
- [0045] 질의응답 결과 재순위화부(300)는, 재순위화 모델 머신러닝 학습부(200)에서 학습된 질의응답 결과 재순위 모델을 이용하여 질문(Question)(3)에 대한 기존 시스템(400)의 질의 응답 시스템 A, B, C(1000a, 1000b, 1000c)에서 각각 생성한 질의 응답 결과 데이터(4a, 4b, 4c)를 재순위화(re-ranking Question Answering Results)하여 최종적인 질의 응답 결과 데이터(Question Answering Data)(4)를 산출한다.
- [0046] 기존 시스템(400)의 질의 응답 시스템(1000a, 1000b, 1000c)과 응답 결과 데이터(4a, 4b, 4c)는 적어도 하나 이상으로 두 개일 수 있으며, 세 개로 한정되는 것은 아니고, 세 개 이상 복수 개일 수 있다.
- [0048] 도 2는 본 발명의 일 실시예에 따른 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)의 질의응답 결과 재순위화부(300)의 구성도이다.
- [0049] 도 2를 참조하면, 본 발명의 일 실시예에 따른 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)의 질의응답 결과 재순위화부(300)는, 자연어 추론 표본 생성 모듈(310), 질문 정답 근거문서 인코더(320), 자연어 추론 결과 판별기(330), 질의 응답 결과물 점수 판별기(340)를 포함한다.
- [0050] 자연어 추론 표본 생성 모듈(310)은, 질의응답 데이터를 자연어 추론 학습 데이터형식으로 변환하는 모듈이다. 해당 모듈에서는 특정 질문-정답 쌍에 대해 관련이 있는 근거문장들을 전부 이어 붙여 포함함(entailment) 라벨을 붙이고, 관련이 없는 근거문장들을 임의로 골라 이어 붙여 포함하지 않음(non-entailment) 라벨을 붙여 자연어 추론 형식의 학습 표본을 생성한다.
- [0051] 질문 정답 근거문서 인코더(320)는, BERT, RoBERTa, ELECTRA등의 트랜스포머(transformer) 기반의 사전학습 인코더를 기반으로 한 신경망 모델이며, 질의, 근거문서, 정답을 인코딩하여 라벨 추론용 임베딩 벡터로 만드는 모듈이다. 인코더에는 [CLS] 토큰, 근거문서, [SEP] 토큰, 질의, 정답이 차례대로 이어 붙여져 입력된다.
- [0052] 자연어 추론 결과 판별기(330)는 질의, 정답, 근거문서 인코더에서 나온 임베딩 벡터의 CLS 레이어를 풀링(Pooling)하여 최종적으로 질의 및 정답이 근거문서와 관련이 있는지(포함함, entailment), 없는지(포함하지 않음, non-entailment)의 라벨을 판단한다. 이 때 자연어 추론 결과 판별기(330)는 기계 학습 방식 중 하나인 Feed forward network를 활용한다.
- [0053] 질의 응답 결과물 점수 판별기(340)는, 여러 질의응답 시스템의 결과물을 입력으로 받아, 포함함(entailment) 라벨의 점수를 각각 도출한 뒤 가장 높은 점수를 지닌 후보를 산출하여 질의응답 결과를 재순위화 한다. 질의 응답 결과물 점수 판별기(340)는 질문, 정답, 근거문서 인코더 및 자연어 추론 결과 판별기(330)의 학습된 가중치와 모델 구조를 그대로 이용한다.
- [0054] 본 발명의 일 실시예에 따른 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)의 질의응답 결과 재순위화부(300)는, 근거 문서가 포함된 질의응답 학습 데이터 저장부(10), 자연어 추론 형식으로 변환된 질의응답 데이터 저장부(20), 질의응답 시스템의 결과물 저장부(40)를 더 포함할 수 있다.
- [0055] 근거 문서가 포함된 질의응답 학습 데이터 저장부(10)는 질의 응답 학습용 데이터(1)를 저장한다.
- [0056] 자연어 추론 형식으로 변환된 질의응답 데이터 저장부(20)는 자연어 추론 학습용 데이터(2)를 저장한다.
- [0057] 질의응답 시스템의 결과물 저장부(40)는 질의 응답 결과 데이터(4)를 저장한다.

- [0059] 도 3은 본 발명의 일 실시예의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법의 순서도이다.
- [0060] 도 3을 참조하면, 본 발명의 일 실시예의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법은 다음과 같은 S100 단계 내지 S300 단계를 포함한다.
- [0061] 본 발명의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법은, 특정한 질의에 대해서 복수의 질의응답 시스템이 생성한 근거 문서 및 답변 후보들 중에서 적어도 하나의 후보를 선정하는 자연어 추론(NLI) 시스템을 이용한 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법에 있어서, 질의응답 데이터를 자연어 추론 형식으로 변환하는 단계(S100), 자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계(S200) 및 훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계(S300)를 포함한다.
- [0062] S100 단계에서, 질의응답 데이터를 자연어 추론 형식으로 변환하는 단계는, 데이터 전처리부(100)가 질의응답 학습용 데이터(1)의 질문(Question), 답변(Answer), 근거 문서(Document) 데이터를 입력받아서 데이터 전처리(Data-precessing)를 실행하는 단계이다.
- [0063] S200 단계에서, 자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계는, 재순위화 모델 머신러닝 학습부(200)가 데이터 전처리부(100)에서 데이터 전처리(Data-precessing)되어 자연어 추론 형식으로 변환된 자연어 추론 학습용 데이터(2)를 입력받아서, 자연어 추론방식을 이용하여 재순위화 모델 머신러닝 학습을 수행하는 단계이다.
- [0064] S300 단계에서, 훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계는, 질의응답 결과 재순위화부(300)가 재순위화 모델 머신러닝 학습부(200)에서 학습된 질의응답 결과 재순위 모델을 이용하여 기존의 질의 응답 결과 데이터(4a, 4b, 4c)를 재순위화하여 최종적인 질의 응답 결과 데이터(4)를 산출하는 단계이다.
- [0066] 본 발명의 일 실시예에 따른 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법의 작동 과정을 설명하면 다음과 같다.
- [0067] 자연어 추론(Natural Language Inference)은 전제 및 가설 두 자연어 문장이 주어졌을 때 문장 사이의 관계를 함의(Entailment), 모순(Contradiction), 중립(Neutral) 혹은 그 외의 관계로 추론하는 과제이다. 자연어 추론은 자연어 텍스트의 문법적 관계뿐만 아니라 문장 안에 내재한 의미 또한 파악해야 하기 때문에, 가장 기본적인 과제로서 여러 자연어 모델의 주요 평가 벤치마크로 이용되며 다른 자연어 하위 모델이 합리적인 결과를 도출하였는지 검증하는 수단으로도 이용될 수 있다.
- [0068] 예를 들어, 질의응답 모델은 질의에 대해 생성한 정답이 근거 문서 내에 올바르게 포함되어있는지 아닌지, 요약 모델은 요약된 결과물이 원본 문서의 핵심적인 내용을 포함하고 있는지 아닌지를 자연어 추론 모델을 이용하여 평가할 수 있다. 자연어 추론 모델은 문장의 심층적인 의미를 판단할 수 있어야 하므로, 자연어 추론 데이터도 단순히 단어 대조로는 전제와 가설 사이의 관계를 추측할 수 없는 경우가 많다. 또한, 일반적으로 각 전제와 가설이 여러 문단이 아닌 하나의 문단, 혹은 하나의 문장으로 이루어져 있다. 대규모 자연어 추론 데이터셋으로 SNLI, MultiNLI 등이 공개되어있으며, 질의응답 데이터셋을 자연어 추론 형식으로 변환한 사례도 존재한다. SciTail은 일반적인 자연어 추론 데이터셋이 특정 과제에 도움이 되기 힘들다는 점을 지적하고, 과학 문제의 질문 및 정답을 합성해 가설을 만들고 근거 문서를 전제로써 활용해 자연어 추론 데이터셋을 만들었다. 만약 여러 질의응답 시스템이 특정 질의에 대응하여 생성한 정답 중 가장 나은 후보를 고를 수 있다면, 결과물의 정확도를 더욱 높일 수 있다. 그리고 질의응답 시스템이 도출한 정답 및 근거 문서는 자연어 추론 모델을 통해서 얼마나 질의에 합당한 지 검증할 수 있다. 따라서, 본 발명에서는 멀티 홉 질의응답 데이터인 HotpotQA 데이터셋을 자연어 추론 데이터셋으로 변환한 뒤 이를 이용해 추론 모델을 훈련하여 여러 질의응답 시스템이 생성한 결과물을 재순위화 할 수 있는 모델을 제안한다.
- [0069] 질의 응답 모델의 정답 도출 과정은 크게 질의에 관련된 문서 검색(Passage Retrieval) 및 정답 판별(Answer Classification)로 나뉜다. 기존 질의 응답 연구는 가장 나은 결과를 도출하기 위하여 문서 검색 과정에서 여러 문서를 뽑아 재순위화 하거나 정답 판별 과정에서 나온 정답 후보에 우선순위를 매기는 등의 방법을 시도하였다. 그러나 위와 같은 방법은 같은 질의 응답 모델에서 도출된 후보만을 이용하거나, 정답 혹은 근거 문서만을 재순위화 하였다. 본 발명에서는 (1) 각기 다른 질의 응답 시스템이 생성한 (2) 답변 및 근거 문서 후보를

동시에 재순위화하고자 하여 종래의 방법과 차별화한다.

표 1

문서 1. 맥플라이	[0] 맥플라이(McFly)는 2003년 런던에서 결성된 영국의 밴드이다. [1] 이 밴드는 백투더 퓨처 캐릭터인 마티 맥플라이에서 이름을 따왔다. [2] 밴드는 톰 플래처, 대니 존스, 더기 파인더, 해리 저드로 구성되어 있다. [3] 그들은 2004년 출시부터 2007년 12월까지 아일랜드 레코드와 계약을 맺었고, 그 후 그들만의 레이블인 슈퍼 레코드를 만들었다.
문서 2. 파이어호스	[0] 파이어호스는 미국의 얼터너티브 록 밴드이다. [1] 그들은 1986년부터 1994년까지 활동했고, 2012년에 재결합했다.
질문	맥플라이와 파이어호스 중 어떤 밴드가 먼저 결성되었는가?
정답	파이어호스
근거 문서	[맥플라이, 0], [파이어 호스, 0], [파이어 호스, 1]

[0070]

[0072] 표 1은 HotpotQA 데이터의 번역된 예시이다. HotpotQA는 멀티 홉(Multi-Hop) 데이터, 즉 하나의 문서에서 정답을 찾을 수 있는 내용이 아니라, 여러 문서를 연관지어 그 안에 내재한 의미를 파악해야만 정답을 도출할 수 있는 질의응답 데이터셋이다. HotpotQA 데이터셋은 표 1처럼 특정 질의에 대해 여러 근거문서를 주며, 질의와 관련된 근거 문서 및 문장을 전부 찾아낸 뒤 정답을 추출해야 한다. 해당 과제는 정답뿐만 아니라 질의응답 시스템이 검색한 근거 문서의 정밀도(Precision) 및 재현율(Recall)을 전부 평가하기 때문에, 문서 검색 과정에서도 핵심 근거 문장만을 도출할 것을 필요로 한다. HotpotQA는 단일 문서 내에서 정답을 찾을 수 있는 일반적인 질의응답 데이터셋과 달리, 근거 문서를 여러 단계로 추론해야 하므로 더 높은 자연어 분석 및 추론 능력이 필요하다. 본 발명은 자연어 추론 모델이 문장에 내재한 의미를 분석하는데 뛰어나다는 점에 착안하여, 질의응답 시스템이 질의에 대해 생성한 답변 및 근거문서를 자연어 추론 모델로 검증하고자 하였다.

[0073] 자연어 추론 데이터는 전제(Premise) 및 가설(Hypothesis), 그리고 두 자연어 텍스트 사이의 관계로 이루어져 있다. 이때, 전제는 질의응답 데이터의 근거 문서에 있는 문장들을 요약하거나, 혹은 그대로 이용할 수 있다. 가설의 경우, 기존 질의응답 데이터를 자연어 추론 형식으로 변환한 사례는 수동, 혹은 자동으로 질문과 정답을 합성하여 하나의 가설 H를 생성하였다. 예를 들어, '철수와 영희는 무슨 관계인가?'라는 질문이 있고, 정답이 '남매'일 때는 '철수와 영희는 남매다'라는 새로운 문장을 생성하였다. 그러나 질문과 정답이 복잡해질 경우 가설 합성 과정에서 전제와의 관계를 판단하는 데에 필수적인 정보가 손실되는 경우가 있었기에, 본 발명에서는 새로운 문장을 만들어내기 보다는 질문과 정답을 이어붙여(Concatenating) 가설 H로 이용하였다.

[0074] 정답 근거문서 - 질문 - 답변 쌍을 함의(Entailment) 표본으로 이용하였고, 정답 질문 및 답변에 대한 오답 근거문서 샘플을 추가하여 중립(Neutral) 표본으로 이용하고자 하였다. 이때 중립 표본은 근거 문서에서 정답 근거문장이 아닌 다른 문장 두 개를 골라 생성하였다. 변환 과정에서 오류가 있던 몇 샘플을 제외한 90,447개의 HotpotQA 데이터에서 함의 샘플과 중립 샘플이 하나씩 나오기 때문에, 총 180,894개의 자연어 추론 데이터가 생성되었다.

[0076] 도 4는 본 발명의 HotpotQA 데이터 학습을 위한 자연어 추론 모델을 도시한 구조도이다.

[0077] 도 4를 참조하면, 자연어 추론 학습 시, 사전 학습된 언어 모델로써 ELECTRA를 이용하였다. 모델 구조는 도 4와 같으며, 근거 문서 SD 와 질문 SQ, 답변 SA 을 [CLS], [SEP] 토큰을 덧붙여 다음과 같이 ELECTRA 인코더에 입력한다.

[0079] (수학식 1)

[0080] $x = ([CLS], SD, [SEP], SQ, SA)$

[0082] 이후, 인코더의 출력 결과에서 [CLS]에 해당하는 첫 번째 벡터를 풀링 후 Classifier에 입력하여 문서와 질문, 정답의 관계를 함의 혹은 중립 두 가지의 라벨로 판별한 뒤 정확도를 높이도록 훈련한다. 재순위화 시에는 질의 응답 시스템이 생성한 질문, 답변, 근거문서를 위에서 학습한 자연어 추론 모델에 입력하며, 함의 라벨의 점수에 따라 가장 나은 답변 후보를 최종 산출한다.

[0084] <실험예>

[0085] 본 발명의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법의 일 실험예에서, 학습 시 사용된 사전학습 모델 ELECTRA는 base 모델을 이용하였으며, 1일동안 Fine-tuning하였다. 실험 환경에서 학습한 모델의 Hyperparameter는 Hidden Size는 768, Attention Head Size는 12, 레이어 수는 12, 문장 최대 길이는 256 토큰, 배치 크기는 32, dropout은 0.1, Learning Rate는 1×10^5 이다.

[0086] 또한, 질의응답 결과 재순위화 이후의 성적을 측정하기 위하여 HotpotQA Fullwiki, Distractor 설정에서 상위 성적을 보여 준 4개의 오픈 소스 질의응답 시스템을 이용하였다. 오픈 소스 코드를 통해 재현한 질의응답 결과물은 표 2와 같은 결과가 측정되었다. HGN(선행기술문헌 1)의 경우 두 사전학습 모델 Albert 및 Roberta를 기반으로 모델을 제공하고 있었기에 두 가지의 결과물을 얻을 수 있었다.

[0087] 재순위화 모델 평가 기준으로 HotpotQA의 평가 기준인 Exact Match와 F1 점수를 이용하였다. EM 및 F1 점수는 정답, 근거 문서, 정답과 근거문서의 joint score를 전부 측정하였다. 평가 데이터로는 7,405개의 HotpotQA Development 데이터를 이용하였고, 재순위화 이전 및 이후의 성적은 표 2와 같다.

표 2

HotpotQA Dev 점수 (재현)	정답		근거 문서		Joint	
	EM	F1	EM	F1	EM	F1
단일 모델 성적						
Recurrent Retriever [9]	60.50	73.31	49.17	76.05	35.83	61.44
MDR [10]	62.36	75.07	56.54	79.42	42.05	66.33
SAE [11]	67.67	80.74	63.26	87.38	46.78	72.74
HGN-Roberta [12]	68.95	82.2	63.11	88.59	46.5	74.36
HGN-Albert [12]	70.18	83.44	63.17	89.19	47.01	75.74
재순위화 이후 성적						
Recurrent Retriever + MDR	65.12	78.34	59.8	83.9	44.11	69.8
Recurrent Retriever + MDR + SAE	67.72	81.33	63.89	88.23	46.73	73.59
MDR + SAE	68.43	81.84	64.5	88.59	47.48	74.27
Recurrent Retriever + MDR + SAE + HGN-Robert + HGN-Albert	66.98	81.3	61.77	88.52	44.37	73.57
SAE + HGN-Robert + HGN-Albert	68.45	82.16	62.55	89.03	45.73	74.59

[0089]

[0091] 표 2는 재순위화 이전과 이후의 질의응답 결과물 Exact Match 및 F1 점수이다. 각 단일 모델에 대해 재순위화 이후에 성적이 향상된 경우는 굵은 글씨로 표시하였다. 여기서, [9]는 선행기술문헌 2, [10]은 선행기술문헌 3, [11]은 선행기술문헌 4, [12]는 선행기술문헌 5 이다.

[0092] Recurrent Retriever(선행기술문헌 6), MDR(선행기술문헌 7), SAE(선행기술문헌 8)의 경우 자연어 추론 모델을 이용한 재순위화 이후 성적이 더 올라간 모습을 보여주었다. 이는 학습된 추론 모델이 질의에 대해 더 나은 근거문서와 정답을 선택하였다는 뜻으로, 제안한 재순위화 방법이 질의응답 시스템의 성적을 성공적으로 개선하였다고 할 수 있다. 그러나 타 질의응답 시스템에 비해 높은 성적을 보여준 HGN의 결과물을 같이 재순위화 할 경우, HGN 단일 모델의 점수보다 재순위화 이후 점수가 오히려 떨어졌다. 해당 결과는 학습된 추론 모델이 일정 성능 이상의 결과물을 판단하는데 한계를 지니고 있음을 시사한다. 또, 다섯 가지 질의응답 시스템의 결과물을

한꺼번에 재순위화 하는 것 보다 소수의 상위 성적의 결과물만 재순위화 하는 것이 더 높은 점수를 보여주었다. 이는 낮은 성적의 결과물이 오히려 노이즈(noise)로 작용한다는 것을 드러내는 예시로서, 현재 구축한 추론 모델은 후보가 많을수록 가장 나은 결과물을 선택하는 능력이 떨어진다는 것을 보여준다. 제안한 방법으로 훈련한 추론 모델로 재순위화한 결과물은 기존 개별 질의응답 모델의 성적을 크게 웃도는 성적을 보여주었다. 그러나, 일정 수준 이상의 질문-대답 쌍에 대해서는 무엇이 올바른 후보인지 제대로 판단하지 못하고, 다수의 선택지가 주어졌을 때 최상의 샘플을 명확히 구분하지 못한다는 한계를 지닌다.

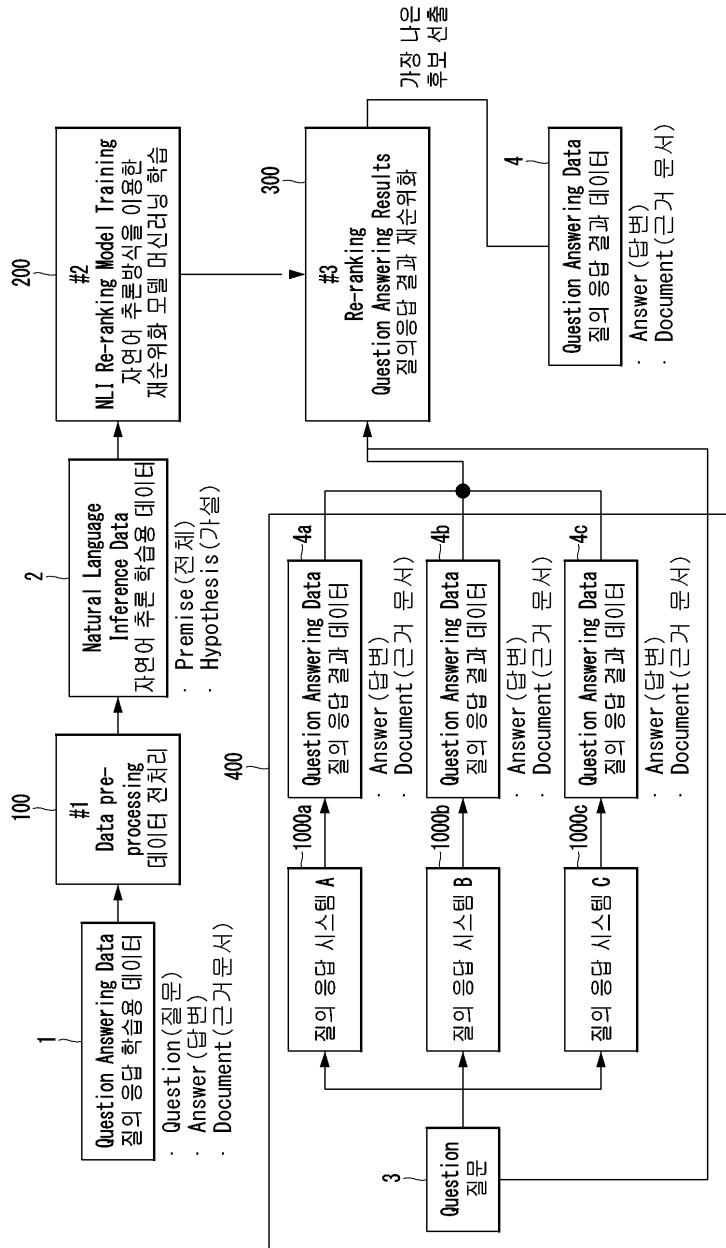
- [0093] 본 발명의 실험예에서는 HotpotQA 질의응답 데이터를 자연어 추론 데이터로 변환하여 추론 모델을 학습하고, 해당 모델을 이용해 여러 질의응답 시스템의 결과물을 재순위화하여 질의응답 결과의 성능 향상을 도모하였다. 일부 질의응답 시스템의 재순위화 결과는 제안한 방법으로 성능이 향상되어, 자연어 추론 모델을 통한 재순위화가 가능함을 증명하였다.
- [0095] 도 5는 본 발명의 일 실시예의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)의 구성도이다.
- [0096] 도 5를 참조하면, 본 발명의 일 실시예의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)는, 프로세서(1100), 메모리(1200), 송수신 장치(transceiver, 1300), 입력 인터페이스 장치(1400), 출력 인터페이스 장치(1500), 저장 장치(1600) 및 버스(bus)(1700)를 포함하여 구성될 수 있다.
- [0097] 본 발명의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)는, 프로세서(processor)(1100) 및 프로세서(1100)를 통해 실행되는 적어도 하나의 명령이 저장된 메모리(memory)(1200)를 포함하되, 적어도 하나의 명령은 상기 프로세서(1100)가, 특정한 질의에 대해서 복수의 질의응답 시스템이 생성한 근거 문서 및 답변 후보들 중에서 적어도 하나의 후보를 선정하는 자연어 추론(NLI) 시스템을 이용한 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 방법에 있어서, 질의응답 데이터를 자연어 추론 형식으로 변환하는 단계(S100), 자연어 추론 형식으로 변환된 데이터를 이용해 머신러닝 방식으로 재순위화 시스템을 구축하는 단계(S200) 및 훈련된 재순위화 시스템을 통해 질의응답 시스템이 생성한 후보를 재순위화 하는 단계(S300)를 수행하도록 구성된다.
- [0098] 프로세서(1100)는 중앙 처리 장치(central processing unit, CPU), 그래픽 처리 장치(graphics processing unit, GPU), 또는 본 발명의 실시예들에 따른 방법들이 수행되는 전용의 프로세서를 의미할 수 있다.
- [0099] 메모리(1200) 및 저장 장치(1600) 각각은 휘발성 저장 매체 및 비휘발성 저장 매체 중에서 적어도 하나로 구성될 수 있다. 예를 들어, 메모리(1200)는 읽기 전용 메모리(read only memory, ROM) 및 랜덤 액세스 메모리(random access memory, RAM) 중에서 적어도 하나로 구성될 수 있다.
- [0100] 또한, 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)는 무선 네트워크를 통해 통신을 수행하는 송수신 장치(transceiver)(1300)를 포함할 수 있다.
- [0101] 또한, 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)는 입력 인터페이스 장치(1400), 출력 인터페이스 장치(1500), 저장 장치(1600) 등을 더 포함할 수 있다.
- [0102] 또한, 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)에 포함된 각각의 구성 요소들은 버스(bus)(1700)에 의해 연결되어 서로 통신을 수행할 수 있다.
- [0103] 본 발명의 자연어 추론 신경망 모델을 이용한 질의응답 결과 재순위화 장치(1000)의 예를 들면, 통신 가능한 데스크탑 컴퓨터(desktop computer), 랩탑 컴퓨터(laptop computer), 노트북(notebook), 스마트폰(smart phone), 태블릿 PC(tablet PC), 모바일폰(mobile phone), 스마트 워치(smart watch), 스마트 글래스(smart glass), e-book 리더기, PMP(portable multimedia player), 휴대용 게임기, 네비게이션(navigation) 장치, 디지털 카메라(digital camera), DMB(digital multimedia broadcasting) 재생기, 디지털 음성 녹음기(digital audio recorder), 디지털 음성 재생기(digital audio player), 디지털 동영상 녹화기(digital video recorder), 디지털 동영상 재생기(digital video player), PDA(Personal Digital Assistant) 등일 수 있다.
- [0105] 본 발명의 실시예에 따른 방법의 동작은 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 프로그램 또는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의해 읽혀질 수 있는

데이터가 저장되는 모든 종류의 기록장치를 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어 분산 방식으로 컴퓨터로 읽을 수 있는 프로그램 또는 코드가 저장되고 실행될 수 있다.

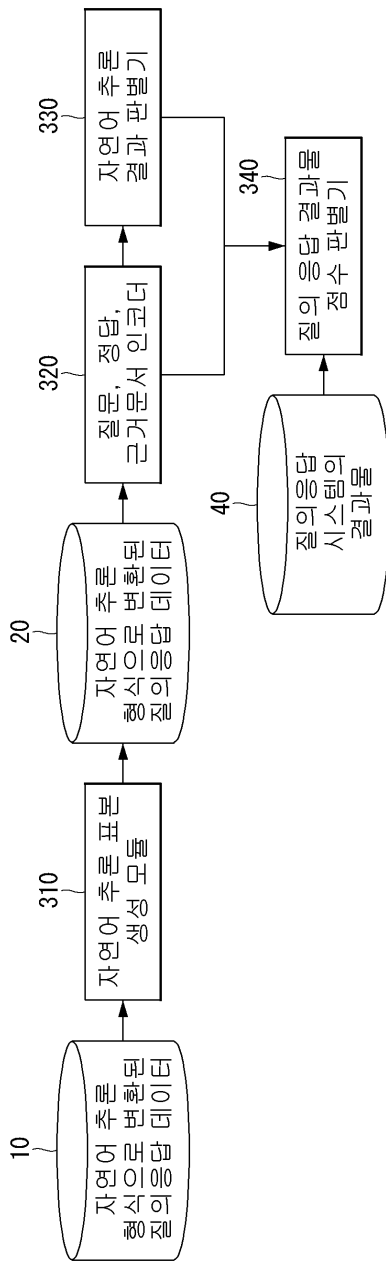
- [0106] 또한, 컴퓨터가 읽을 수 있는 기록매체는 롬(rom), 램(ram), 플래시 메모리(flash memory) 등과 같이 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치를 포함할 수 있다. 프로그램 명령은 컴파일러(compiler)에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터(interpreter) 등을 사용해서 컴퓨터에 의해 실행될 수 있는 고급 언어 코드를 포함할 수 있다.
- [0107] 본 발명의 일부 측면들은 장치의 문맥에서 설명되었으나, 그것은 상응하는 방법에 따른 설명 또한 나타낼 수 있고, 여기서 블록 또는 장치는 방법 단계 또는 방법 단계의 특징에 상응한다. 유사하게, 방법의 문맥에서 설명된 측면들은 또한 상응하는 블록 또는 아이템 또는 상응하는 장치의 특징으로 나타낼 수 있다. 방법 단계들의 몇몇 또는 전부는 예를 들어, 마이크로프로세서, 프로그램 가능한 컴퓨터 또는 전자 회로와 같은 하드웨어 장치에 의해(또는 이용하여) 수행될 수 있다. 몇몇의 실시예에서, 가장 중요한 방법 단계들의 하나 이상은 이와 같은 장치에 의해 수행될 수 있다.
- [0108] 실시예들에서, 프로그램 가능한 로직 장치(예를 들어, 필드 프로그램블 게이트 어레이)가 여기서 설명된 방법들의 기능의 일부 또는 전부를 수행하기 위해 사용될 수 있다. 실시예들에서, 필드 프로그램블 게이트 어레이는 여기서 설명된 방법들 중 하나를 수행하기 위한 마이크로프로세서와 함께 작동할 수 있다. 일반적으로, 방법들은 어떤 하드웨어 장치에 의해 수행되는 것이 바람직하다.
- [0109] 이상 본 발명의 바람직한 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

도면

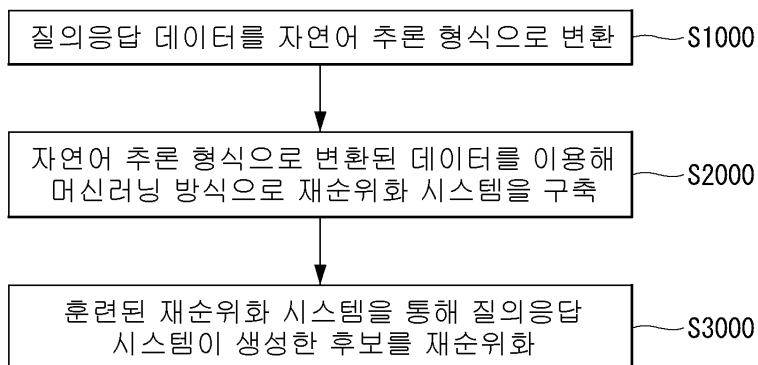
도면1



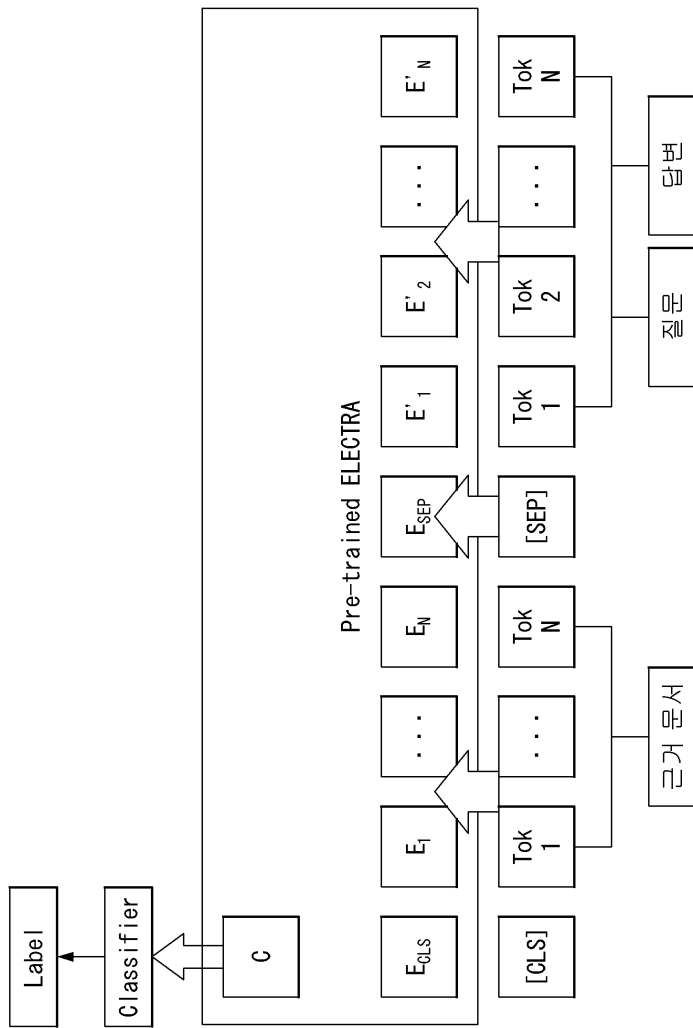
도면2



도면3



도면4



도면5

