

그래프 신경망 기반 질의응답 시스템에서 그래프 병합을 활용한 재추론 기법

이필원*, 김상훈**, 신용태***
*송실대학교 컴퓨터학과,
**송실대학교 정보보안학과,
***송실대학교 컴퓨터학부
*pwlee@soongsil.ac.kr

Re-Inference Method using Graph Merging in Graph Neural Network based Question Answering System

Pil-Won Lee*, Sang-Hoon Kim*, Yong-Tae Shin**
*Department of Computer Science, Soongsil Univ.
**Department of Information Security, Soongsil Univ.
***Dept of Computer Science and Engineering, Soongsil Univ.

요 약

최근 다수의 문서를 고려해야하는 다중홉(multi-hop) 추론과 같은 복잡한 문제를 해결하기 위해 계층적 그래프 신경망기반 질의응답 시스템이 제안되었다. 계층적 그래프 신경망 기반 질의응답 시스템은 사람의 정확도를 뛰어넘었으나 제한된 문서를 통해 추론을 진행하기 때문에 문서에 충분한 정보가 없을 경우 추론에 실패할 가능성이 존재한다. 따라서 본 논문에서는 위 문제를 해결하기 위해 정보를 재탐색하고 기존의 그래프 정보와 병합하여 기존의 정보와 새로운 정보를 고려하여 재추론할 수 있는 그래프 병합 기법을 제안한다. 제안하는 그래프 병합 기법은 사전에 정의된 규칙에 의해 수행되며 노드의 병합 및 연결을 통해 새로운 그래프를 도출한다. 새로운 그래프는 그래프 신경망을 통해 추론을 진행하여 기존 정보와 새로운 정보를 고려한 정답을 도출할 수 있다.

1. 서론

최근 다중홉(multi-hop) 추론과 같은 복잡한 문제에서 계층적 그래프 신경망(Hierarchical Graph Network) 기반 질의응답 시스템이 사람의 정확도를 뛰어 넘었다[1]. 계층적 그래프 신경망 기반 질의응답 시스템은 흩어져있는 단서를 집계하기 위해 텍스트를 계층적 그래프로 표현하고 그래프 신경망을 통해 추론을 수행하여 답변을 도출하는 질의응답 시스템이다. 그래프 신경망 기반 질의응답 시스템은 이미 사람의 정확도를 뛰어넘었으나 제한된 문서를 통해 추론을 진행하기 때문에 문서에 충분한 정보가 포함되어있지 않다면 추론에 실패할 가능성이 여전히 남아있다. 따라서 문서의 정보가 불충분 하더라도 재 추론을 통해 답변을 도출할 수 있는 실패에 강인한(Robust) 질의응답 시스템이 필요하다. 그러나 단순히 정보를 재탐색 하여 재 추론하는 시스템은 기존의 정보를 활용하지 못한다는 문제점이 있다. 본 논문에서는 위 문제를 해결하기 위해 추론

에 실패할 시 정보를 다시 탐색하고 기존의 그래프 정보와 병합하는 기법을 제안한다. 본 논문의 구성은 다음과 같다 2장 관련연구에서 기존 계층적 그래프 신경망과 그래프 병합 기법에 대해 서술한다. 3장 제안하는 그래프 병합 기법에서는 기존의 그래프 정보와 재탐색한 정보를 병합하는 기법에 대해 서술한다. 4장 결론에서는 제안하는 기법에 대한 종합적인 요약과 향후 연구 방향을 제시한다.

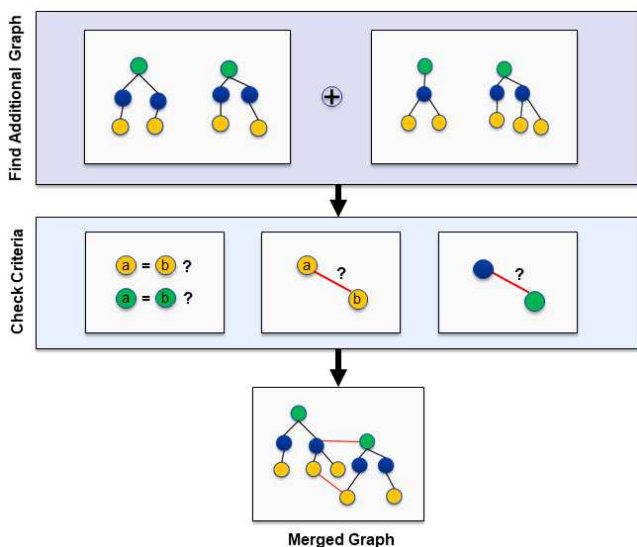
2. 관련연구

계층적 그래프 신경망 기반 질의응답 시스템 계층적 그래프 신경망은 제안된 연구 별로 다른 방식으로 계층적 그래프를 구성하지만 문장 수준 병로 계층을 구성한 HGN[1]모델이 HotpotQA[2] 데이터 세트를 기준으로 공개된 모델 중 가장 높은 정확도로 평가되었다. HGN에서 문장 수준 별 계층 그래프는 문단, 문장, 개체 수준으로 계층을 나누어 그래프를 구성한다. 위 과정을 통해 다수의 문서의 정보를 고려한 추론이 가능하다. 수준 별로 나누어진 계층

그래프는 GAT[3](Graph Attention Network)를 통해 주변 노드의 정보를 전파하고 각각의 노드의 상태를 수정한다. 위 과정을 통해 도출한 임베딩 벡터 값을 정답 예측 모듈을 통해 정답 예측을 수행한다. 다수의 문서를 고려해야하는 다중홉 질문에서 HGN은 82%의 높은 정확도로 평가되었다. 그러나 제한된 문서를 고려한다는 점에서 문서의 정보가 불충분하다면 추론에 실패할 가능성이 존재한다.

3. 제안하는 그래프 병합 기법

제안하는 그래프 병합 기법은 그래프 신경망 기반 질의응답 시스템에서 추론이 실패했을 경우를 전제한다. 이를 해결하기 위해 본 논문에서는 질문과 연관된 정보를 다시 탐색하고 도출한 정보를 그래프로 구성하고 기존의 그래프 정보와 병합하는 기법을 설계한다.



(그림 1) 그래프 병합 기법 도식화

(그림 1)은 그래프 병합 기법에 대한 프로세스를 도식화 한 것이다. 그래프 병합 기법은 기존의 그래프와 새로 입력된 그래프를 사전에 정의한 기준에 따라 연결하거나 병합하는 과정을 수행한다. 사전에 정의한 기준은 3가지가 존재하며 (1) 동일한 개체 노드 병합, (2) 노드의 관계를 인식하여 노드를 연결하는 개체 노드 연결, (3) 문단과 문장에 하이퍼링크가 존재한 경우 연결하는 문단 문장 연결이 있다.

3-1 개체 노드 병합

개체 노드 병합은 동일한 개체 노드를 병합하는 것을 의미한다. 개체 노드의 정보는 언어 모델을 통해

임베딩 벡터 값을 알 수 있기 때문에 의미하는 개체가 같은 개체인지 판별이 가능하다. 만약 두 개의 개체가 같은 개체라고 판별되면 각각의 개체에 연결된 모든 연결을 통합한다.

3-2 개체 노드 연결

개체 노드 연결은 개체의 연관성을 판단하여 연결을 수행한다. 각각의 개체는 임베딩 벡터로 표현될 수 있기 때문에 코사인 유사도(Cosine Similarity)로 연관성을 판단한다. 또한 연결 여부를 결정하기 위해 활성화함수를 활용한다. 이 때 활용하는 활성화함수는 코사인 유사도의 범위를 고려하여 시그모이드(sigmoid)를 활용한다. 이를 수식으로 표현하면 수식(1)과 같다.

$$y = \frac{1}{1 + \exp(-\text{CosineSimilarity}(\text{Node}_1, \text{Node}_2))} \quad (1)$$

수식(1)에서 Node1, Node2는 연관성을 판단하려는 노드의 임베딩 벡터를 의미한다.

3-3 문단 문장 연결

문단 문장 연결은 문장에 다른 문단으로 통하는 하이퍼링크가 존재할 때 노드를 연결하는 것을 의미한다. 문장과 문장사이를 연결하지 않는 것은 그래프에 너무 많은 연결이 존재하게 되면 그래프 구조에 노이즈(noise)가 발생할 가능성이 높아지기 때문이다. 따라서 오직 문장에 다른 문단으로 통하는 하이퍼링크가 존재할 경우만 연결한다.

그래프 병합 기법은 위 3가지 조건에 부합하는 노드를 병합 및 연결하여 새로운 그래프를 생성한다. 이를 통해 그래프 신경망을 통해 다시 추론을 수행하면 기존의 정보와 새로운 정보를 고려한 새로운 정답을 도출 수 있다.

4. 결론

기존의 계층적 그래프 기반 질의응답 시스템은 제한된 정보를 활용하여 정답을 도출하므로 정보가 충분하지 않을 경우 추론에 실패할 가능성이 존재한다. 이를 해결하기 위해 단순히 정보를 다시 탐색하는 것이 아니라 기존의 정보를 활용하기 위해 기존의 그래프와 새로 탐색한 정보 그래프를 병합하는 기법은 본 논문에서 제시했다. 병합한 새로운 그래

프는 그래프 신경망을 활용하여 기존 정보와 새로운 정보를 고려하여 추론을 수행할 수 있다. 향후 그래프 병합을 통한 재추론 기법이 질의응답 시스템에 적용되었을 때 추론 실패에 강인한 정도를 실험을 통해 측정할 예정이다.

ACKNOWLEDGMENT

"본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터지원사업의 연구결과로 수행되었음" (IITP-2020-2020-0-01602)

참고문헌

- [1] Yuwei Fang, Siqi Sun, Zhe Gan, Rohit Pillai, Shuohang Wang, and Jingjing Liu, "Hierarchical graph network for multi-hop question answering". In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020.
- [2] Zhilin Yang , PengQi, SaizhengZhang, YoshuaBengio, William Cohen, RuslanSalakhutdinov, and Christopher D.Manning, "HotpotQA: A dataset for diverse, explainable multi-hop question answering". In Empirical Methods in Natural Language Processing (EMNLP), 2018.
- [3] Velickovic, Petar, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio' and Yoshua Bengio. "Graph Attention Networks." 2018.