

# Self-Critic 기반 기계독해 알고리즘 일반화에 관한 연구

\*고성운, 박석, 정슬기, 주찬웅  
모두의 연구소 Deep Learning College  
e-mail : seongwoon.ko@modulab.ai, surk.park@modulab.ai,  
seulki.jung@modulab.ai, jcwleo@modulab.ai

## Generalization of Machine Reading Comprehension Using Self-Critic Based Learning

\*Seongwoon Ko, Surk Park, Seulki Jung, Chanwoong Joo  
Deep Learning College  
ModuLabs

### 요약

본 논문은 Stanford Question Answering Dataset (SQuAD)의 기계독해 성능 향상에 강화학습 알고리즘 Self-Critic의 일반화 가능성에 대하여 연구하였다. 주어진 텍스트를 읽고 이해하며 질문에 적합한 답을 추론하는 기계 독해는 기존 앙상블 등 다양한 머신러닝 알고리즘으로 적용 되어왔다. 본 연구는 기존 사용된 알고리즘 RNet, FusionNet, Conduct Net, Menemoric Reader에 Self-Critic을 적용하여 자연어 처리 및 기계독해 성능을 개선을 하였다. 기존 모델들은 질문의 정답을 지문의 위치로 나타내며 위치가 다를 경우 학습에 문제가 되어 성능 저하를 나타내는데 이 문제를 해결하기 위해 Self-Critic을 적용하였다. 실험 결과, Self-Critic이 적용된 네가지 알고리즘은 기존 SQuAD에 보여진 결과보다 평균적으로 우수한 성능을 보였으며 따라서 본 논문은 강화학습 알고리즘이 기계독해 성능 향상에 기여한다는 연구결과의 일반화 가능성을 주장한다.

### I. 서론

자연어 처리는 워드임베딩, 감성분석, 언어번역 등 다양한 분야에서 쓰이고 있으며 머신러닝 기술이 발전함에 따라 뛰어난 성능을 보이고 있다. 특히나 자연어 기반 질의응답(Question and Answering) 기술이 빠른 성장을 보이고 있는데 이를 위해 다양한 종류의 데이터가 제공되고 있다. 예를 들어 구글 답마인드는 질문하는 context의 적절한 대답을 빈칸에 메우는 스타일의 데이터 셋을 제공하며 스탠포드의 SQuAD는 질문의 정답으로 예상되는 단어를 찾아주는 형태로 구성되어 있다.

문제를 풀기 위해 기계는 주어진 데이터를 이해하고 추론을 해야하는데 이 과정을 기계독해( Machine

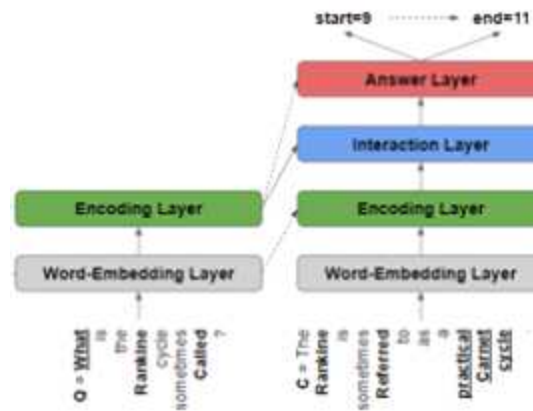
Reading Comprehension) 라고 한다. 현재 FusionNet, R-Net, Memoric Reader, DCN+ 등 과 같은 알고리즘들이 embedding, encoding, attention, answer Point 등 다양한 메카니즘을 사용하여 기계독해 성능을 개선하고있다. 본 논문은 스탠포드의 SQuAD 2.0 데이터 셋을 사용하며 위치 기반으로 학습을 시키는 기존 알고리즘들의 문제점을 개선하기 위해 도입된 Self-Critic을 다양한 알고리즘에 적용해 본다. 이를 통해 기존 모델과의 성능차이를 비교하며 Self-Critic의 적용이 기계독해에 일반적인 접근방법이 될 수 있음을 주장한다.

### II. 본론

#### 2.1 기계독해

최근 SQuAD 데이터셋의 기계독해 문제를 해결하기

위한 여러 아키텍처들이 연구됨에 따라 기계독해의 성능이 점점 인간의 성과와 가까워 지고 있다. 일반적인 QA 시스템은 word embedding, encoding, interaction 그리고 answer 레이어로 구성되어 있다. 질문과 지문간의 다양한 attention 매커니즘의 조합을 통해 질문 단어 임베딩과 지문 단어 임베딩 단위에서, 혹은 RNN으로 한번 처리가 된 단위에 attention을 적용하여 지문에서 중요한 부분을 찾고(참고: BiDaF Mnemonic Reader, R-net, FUSION), 이렇게 합쳐진 인코딩을 self-matching등을 적용하여 문장사이의 깊은 연관성을 파악하는 연구도 많이 진행되었다 (참고: Fast QA, Mnemonic Reader, R-net, FUSION).



참고 2. 기본 QA 기본도

## 2.2 강화학습을 적용한 기계독해

SQuAD 데이터셋에서는 metric이 string 기반으로 채점이 되는 반면, 실제로 모델을 training 할때 질문의 정답을 지문의 위치로 표현하기 때문에 cross entropy optimization과 실제 metric간에 조금 차이가 있다. 예를 들어 예측 위치가 실제 정답에서 벗어나 있는데 실제 단어는 정확히 일치하는 경우, EM과 F1 둘다 정답으로 인정되지만, cross entropy optimization은 (CE) 두 케이스 모두 틀린케이스로 인지하여 파라미터를 업데이트 한다. 이런 실제 metric의 선호도를 학습 시키기 위해, DQN+ 에서 처음 self-critical RL objective를 적용하여 metric과 차이를 극복하려는 시도를 했으며, Mnemonic Reader에서는 RL의 리워드가 항상 양수가 되도록 변경하여 더 안정적인 학습에 성공하였다.

$$l(\theta) = -E_{p_r} [ \log p_r(s, e; s, \hat{e}_T; \theta) ] - E_{p_r} [ F_1(ans(\hat{s}_T, \hat{e}_T), ans(s, e)) - F_1(ans(s_T, e_T), ans(s, e)) ]$$

참고 1. Self-Critical Reinforcement Learning Loss<sup>1)</sup>

최근 Deep Learning 기반 ‘Extractive QA’ 의 일반적인 아키텍처는 4가지 Layer로 구성되어 있고, 각 Layer의 성능을 높이기 위해, Layer의 Input으로 들어가는 데이터의 Representation Scope(Character-level, Word-level, Sentence-level)을 추가적으로 Feeding하거나 각 Layer에 사용되는 RNN 알고리즘(RNN-basic, LSTM, bi-LSTM, GRU 등)을 다양화하는 연구가 진행되고 있지만, 기본적인 아키텍처는 4Layer 로 구성된 아키텍처로 일반화 되어 있다.

1) C. Xiong, et al. DCN+ Mixed objective and deep residual coattention for question answering.

## 2.3 Deep Reinforcement Learning

Language Model에서의 Recurrent Neural Network은 RL의 Environment를 정의한다. State 는 각 ‘hidden state’, ‘Current input(현재 질문)’, ‘Source context(지문)’ 이 되고, Action은 ‘Next Input(다음 질문)’, ‘Source(컨텍스트 전체)’ 가 되며, Reward는 Answer layer가 예측한 Answer의 시작 위치 와 종료 위치에 대한 probability distribution이 된다. Supervised Learning을 수행하는 Language Model 자체가 Agent이며 REINFORCE, A2C, A3C 와 같은 Policy Gradient 알고리즘을 Deep Reinforcement Learning 기반 Neural Network에 학습하는 Approach를 적용한다.

## III. 실험

스텐포드에서 제공되어지는 SQuAD 2.0을 사용하였으며 Mnemonic Reader, Fusion Net, Rnet, Conductor Net 총 4가지 알고리즘에 Self-Critic 알고리즘을 적용하여 비교분석 하였다.

## IV. 결론 및 향후 연구 방향

본 논문에서는 기존 모델에 Self-Critic을 적용하여 강화학습이 기계독해 성능개선에 일반적으로 사용될수 있음을 제안하였고 Mnemonic Reader, R-net, Fusion-net, Conductor-net에 대한 비교 실험을 진행하였다. 실험 결과, 본 논문에서 제안한 Self-Critic이 적용된 모델들은 공통적으로 성능이 증가됨을 보였다.

이러한 연구 결과에 근거하여, SQuAD 2.0 버전에 새로이 등장한 unanswerable question 유형에 대해 강화학

습을 통한 기계독해 성능 향상을 향후 연구과제로 제언하고 있다.

## 감사의 글

이 논문은 2018년도 정부(서울시)와 모두의 연구소의 지원을 받아 수행된 결과물이다.

## 참고문헌

- [1] M. Hu, et al. Reinforced Mnemonic Reader for Machine Learning Comprehension. arXiv: 1705.02798, 2018
- [2] C. Xiong, et al. DCN+ Mixed objective and deep residual coattention for question answering. arXiv: 1711.00106, 2017
- [3] Microsoft Asia Natural Language Computing Group. R-Net: Machine Reading Comprehension with Self-Matching Networks. 2017
- [4] C. Tan, et al. S-Net: From Answer Extraction to Answer Generation for Machine Reading Comprehension. arXiv: 1706.04815, 2018
- [5] H. Huang, et al. FusionNet: Fusing via Fully Aware Attention with Application to Machine Comprehension. arXiv: 1711.07341, 2018
- [6] X. Liu, et al. Stochastic Answer Networks for Machine Reading Comprehension. arXiv: 1712.03556, 2018
- [7] P. Rajpurkar, et al. Squad: 100,000+ questions for machine comprehension of text. arXiv: 1606.05250, 2016
- [8] P. Rajpurkar, et al. Know what you don't know: unanswerable questions for SQuAD. arXiv: 1806.03822, 2018