

Abstract

본 논문은 강화학습이 Stanford Question Answering Dataset (SQuAD)의 기계독해 성능 향상에 주는 영향도에 대하여 연구하였다. 주어진 텍스트를 읽고 이해하고 질문에 적합한 답을 추론하는 기계 독해는 기존 앙상블 등 머신러닝 지도학습 알고리즘으로 적용 되어왔다. 본 연구는 기존 사용되었던 알고리즘에 강화학습을 더하여 자연어 처리 및 기계독해 성능을 개선하였으며 추론 정확성을 높였다. 새로이 적용된 XX는 기존 SQuAD에 시도된 바 없으므로 기계독해 알고리즘에 추가 적용하여 높은 정확도와 재현율을 얻음으로 써 강화학습 알고리즘이 자연어 처리 성능 향상에 기여함을 제시해 보고자 한다.

Introduction

4V(Volume, Variety, Velocity, Veracity)의 특징을 가지는 빅데이터는 최근 Deep Learning 기술의 발전을 가져온 가장 큰 이유 중 하나로 언급(Andrew Ng, Geoffrey Hinton, Yann LeCun, Yoshua Bengio)되고 있으며, 글로벌 시장조사 기관인 IDC에 의하면 2025년 전 세계의 데이터 생산량은 약 163ZB 정도의 크기가 될 것으로 예측되며, 이는 2016년에 생성된 16.1ZB보다 약 10배정도 늘어난 규모이다. IBM은 빅데이터의 80%가 비정형데이터(Text, Audio, Image)라고 조사하였으며 2020년에 의료분야의 지식은 73일 주기로 2배씩 폭증할 것으로 예상했다. 빅데이터의 20%에 해당하는 정형데이터로부터 추출하는 가치보다 ‘80%의 비정형데이터로부터 누가 먼저 가치를 추출해 낼 것인가?’가 매우 중요한 시대가 되었다. 본 논문에서는 비정형데이터 중 많은 부분을 차지하는 Text로 부터 Natural Language Processing을 하여 가치있는 정형데이터를 분석하는 여러가지 시도 중 특히, ‘자연어 기반 질의응답 기술’에 대해서 논의하고자 한다.

최근 자연어처리의 다양한 문제를 해결하기 위해 Deep Learning에 기반한 접근방법이 시도되고 있는 추세이며 자연어처리의 Preprocessing에 사용되는 형태소분석, PoS Tagging, NER과 같은 상대적으로 쉬운 분야뿐만 아니라 WordEmbedding(Word2Vec, GLoVe, fasttext, ELMo, CoVe 등), CharacterEmbedding, SentenceEmbedding과 같은 Presentation 영역에도 사용되어 LT(Language Translation), Sentiment Analysis 등과 같은 자연어처리의 어려운 영역에서도 놀라운 결과를 얻어내고 있다. 하지만 Summarization, QA(Question Answer)와 같은 자연어처리 영역은 Supervised Learning과 Unsupervised Learning, Reinforcement Learning 등의 Deep Learning 접근방법이 시도되고 있지만 여전히 해결하기 매우 어려운 영역으로 손꼽히고 있다.

본 논문에서는 자연어처리의 여러가지 문제를 해결하기 위한 Language Model 중 ‘Extractive QA’에 대해 다룰 것이다. ‘Extractive QA’는 ‘자연어 기반 질의응답 기술’ 중의 하나이고 스티uart 러셀 교수(버클리대)는 인공지능 기술을 검색산업에 도입하면 현재 1조 달러의 검색산업이 10조 달러 규모로 성장할 것이라고 예측하였다.(WEF 다보스 포럼, 2016) 또한 International Data Corporation에서는 2019년에 Robotics와 이와 관련된 서비스에 투자되는 돈이 135.4 billion에 이를 것이라고 예측하여 ‘자연어 기반 질의응답 기술’과 같은 Cognitive기술이 미래의 ‘인간과 인간’, ‘인간과 기계’ 사이의 Interface를 획기적으로 바꾸는 매우 중요한 기술적 가치를 가지고 있다고 보았다.

‘자연어 기반 질의응답 기술’에 대한 기술적 가치를 미리 보고 글로벌에서는 국가차원에서 또한 글로벌 기업 주도의 대형 장기 프로젝트가 진행되어 왔다. 미국의 방위고등연구계획국(DARPA)에서는 자연어 처리·질의응답 기술을 개발, 애플 시리 및 IBM 왓슨 QA의 모태가 되었다. 애플 시리는 2003년부터 5년 간 SRI인터내셔널이 주도하고 DARPA가 2억 달러를 투자한 인공지능 연구 프로젝트 CALO(Cognitive Assistant that Learns and Organizes)에서 출발한 기술이다. IBM에서는 DARPA의 Advanced QA를 연구하는 AQUAINT 프로젝트의 결과물을 기반으로 왓슨 QA 개발이 시작된바 있다. DARPA는 자연어 의미 이해를 위한 DEFT(Deep Exploration and Filtering of Text) 프로젝트 추진하였다. 결과가 다소 좋지 않은 사례들도 있다. 일본에서는 국립정보과학연구소(National Institute of Informatics)와 후지쯔가 공동으로 슈퍼컴퓨터를 이용하여 2021년 동경대 입시 합격을 목표로 Todai 로봇 프로젝트를 2011년부터 추진하다가 작년에 언어의 독해력 기술개발의 높은 한계에 직면하여 포기를 선언하였으며, 마이크로소프트는 2016년 3월 딥러닝 기반의 챗봇서비스 ‘테이’를 공개했으나 통제되지 않은 학습과 대화 품질 문제로 서비스를 중단하는 해프닝이 있기도 하였다. 이와 같은 인간의 도전에도 불구하고 여전히 기계가 사람언어의 의미를 이해하는 것은 아직 초보적인 단계라고 봐야 할 것이다.

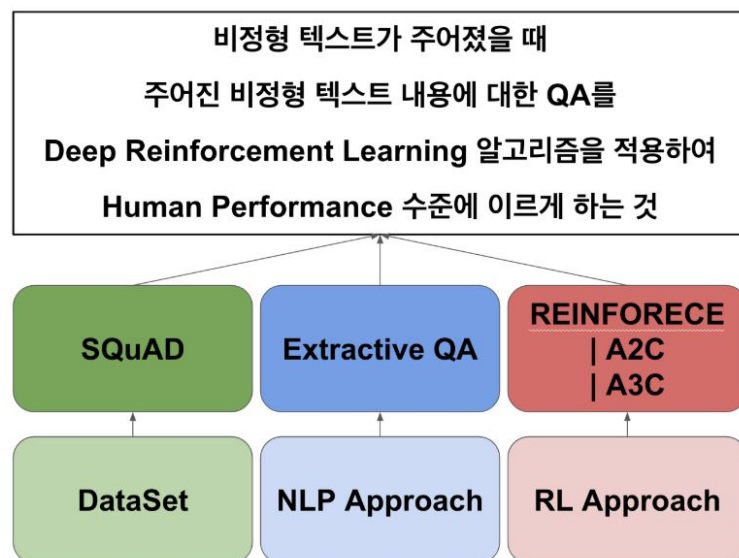
Deep Learning 기술의 발전의 뒤에 데이터가 있었던 것과 같이 ‘Extractive QA’와 같은 ‘자연어 기반 질의응답 기술’의 급격한 발전의 뒤에는 ‘질의응답 DataSet과 이와 관련된 연구가

있었다. 예를들어, 스탠포드 대학 뿐만 아니라 IBM, 구글, 마이크로소프트, 페이스북 등의 글로벌 기업이 질의응답 DataSet 개발에 선두자하고 이에 기반한 딥러닝 기반의 질의응답 기술을 개발하고 있다.

	데이터셋	질문 구축	질문 수	문서 수(문서유형)	URL
구글 답마인드	CNN	빈칸 메우기 질문 생성	387,420 (CNN 뉴스)	92,579 (CNN 뉴스)	http://cs.nyu.edu/~kcho/DMQA/
구글 답마인드	Daily Mail	빈칸 메우기 질문 생성	997,467 (Daily Mail 뉴스)	219,506 (Daily Mail 뉴스)	http://cs.nyu.edu/~kcho/DMQA/
스텐포드 대학	SQuAD	크라우드 소싱	107,785	536 (위키피디아)	https://rajpurkar.github.io/SQuAD-explorer/
마이크로소프트	MS MARCO	빙 검색로그	100,000	1M 정답단락 200,000+ 웹문서 링크	http://www.msmarco.org/
Maluuba	NewsQA	크라우드 소싱	119,633	12,744건 (CNN 뉴스)	https://datasets.maluuba.com/NewsQA

Deep Learning 접근방법이 ‘자연어 기반 질의응답 기술’에 적용되기 전에는 크게 2가지 접근 방법이 연구되었다. 첫번째는 ‘KB(Knowledge Base)기반 연역적 방법’으로 정답에 대한 신뢰성은 높지만 커버링할 수 있는 질문의 양은 적은 특징이 있었다. 이 접근방법은 ‘지식 표현 커버리지를 어떻게 높힐 것인가?’, ‘지식 구축 신뢰도를 어떻게 높힐 것인가?’가 연구방향이며 큰 도전이었다. 두번째는 IR(Information Retrieve)기반 귀납적 방법으로 정답에 대한 신뢰성이 낮지만 커버할 수 있는 질문의 양은 많은 특징이 있다. 이 접근방법은 KB에는 없는 지식도 Web상에 퍼져 있는 비정형데이터에는 있다는 가정과 검색결과를 제공한 후 사용자에게 추론을 맡긴다는 다소 적극적이지 않은 방식에서 출발했다. ‘자연어 기반 질의응답’ 문제는 질의타입과 응답타입에 따라 질의응답의 형태가 매우 다양하며 이를 해결하기 위해 귀납적방법과 연역적방법 혹은 두가지를 함께 사용하는 Hybrid 방법이 적용되었지만 여전히 해결되지 않은 문제로 남았었다.

본 논문에서는 SQuAD DataSet에 ‘Extractive QA’ 즉, ‘비정형텍스트가 주어졌을 때, 주어진 비정형 텍스트 내용에 대한 QA’를 Deep Learning 기술이 적용된 자연어처리기술인 Embedding, Encoding, Attention, Answer Point 메카니즘을 사용한 아키텍처에 Deep Reinforcement Learning 알고리즘(REINFORCE, A2C, A3C)을 적용하는 것이 “Extractive QA’ 구현하기 위한 일반적인 접근방법이 될 수 있음을 주장한다.



Related Work

Machine Reading Comprehension (MRC)

최근 SQuAD 데이터셋 MRC 문제를 해결하기 위한 여러 아키텍처들이 연구됨에 따라 MRC의 성능이 점점 Human Performance에 가까워지고 있다. Answer Pointer 모듈을 마지막 layer로 구성하는 것이 Match LSTM에서 제안되었었다 [Match-LSTM]. 이후 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[BiDaF Mnemonic Reader, R-net, FUSION], 이렇게 합쳐진 인코딩을 self-matching등을 적용하여 문장사이의 깊은 연관성을 파악하는 연구도 많이 진행되었었다 [Fast QA, Mnemonic Reader, R-net, FUSION].

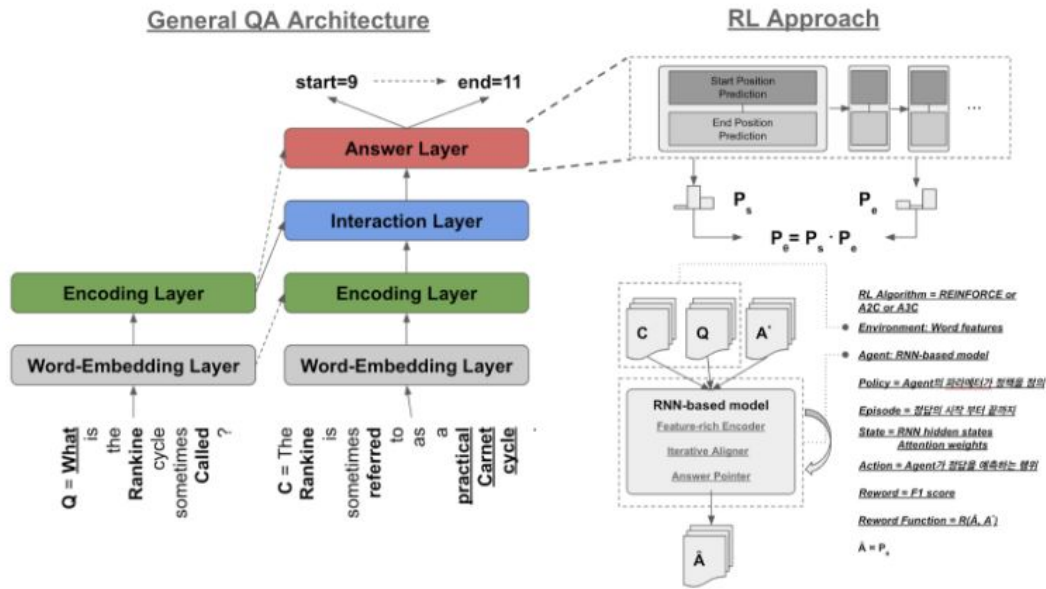
Reinforcement Learning in MRC

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

Intelligent QA

최근 Deep Learning 기반 ‘Extractive QA’의 일반적인 아키텍처는 4가지 Layer로 구성되어 있고, 각 Layer의 성능을 높이기 위해, Layer의 Input으로 들어가는 데이터의 Presentation 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 4Layer로 구성된 아키텍처로 일반화되어 있다.

Intelligent QA Architecture



Embedding Layer

자연어는 결국 Token들의 Sequence로 표현되는데 기존 방법론에서는 Token들의 분포가 연속적인 값을 가지지 않고 Sparse한 특징이 있었고 이로 인한 한계에 직면했다. 이를 뛰어 넘기 위해 자연어처리 과정의 기본적인 단계인 Preprocessing 과정을 거치고 나면 가장 중요한 단계인 Embedding 단계를 거쳐서 Token의 Presentation을 벡터로 바꾸게 된다. 최근에는 Word-level Embedding의 단점인 Unseen Word에 대한 한계를 뛰어 넘기 위해 Embedding시 Character-level Embedding도 병행하는 연구도 진행되고 있다.

Encoding Layer

Encoding Layer는 Sentence-level의 표현을 강화시키는 단계로 기존의 Word-Embedding으로 대변되는 Lexical Feature외에 PoS, NER, Query Category 등의 Syntactic Feature 들을 함께 Feeding 한다.

Interaction Layer

기존 RNN 모델이 long-term 메모리에 대한 한계를 어느 정도 해결하였지만 RNN의 마지막 시점의 벡터에 모든 정보를 담기 버거우므로 모든 입력시퀀스의 정보를 조합하고자하는 시도가 연구되었고 이것이 Attention 메카니즘의 첫번째 아이디어였다. Attention의 두번째 아이디어는 Question의 feature를 RNN에 함께 피딩하여 필요없는 단어에 대해서는 Attention을 주지 않고 필요한 단어에만 집중할 수 있게 한다. 결과적으로 Interaction Layer 는 첫째, Query 와 Context 사이의 Interaction 정보를 기억하며, 둘째, Context와 Context간의 Interaction 정보를 기억한다.

Answer Layer

결국 'Extractive QA'는 Answer의 시작 위치 와 종료 위치에 대한 probability distribution 을 계산하는 것이 목적이고 이 단계에서 RL이 적용된다. RL을 적용하지 않은 기존 모델들이 Multi-Sentence Reasoning을 해결 못하고 'Why' Query와 같은 Answer Boundary가 모호하고

긴 경우에 성능이 매우 낮았다. 이것은 오랫동안 정보를 변질없이 보전할 수 있는 LSTM Algorithm조차 Long-Term Memory를 처리하기에 완벽하지 않으며 문장별로 LSTM을 통과한 후 그 OUTPUT 정보를 메모리에 저장하여 이를 질문에 따라 불러 사용하는 Memory 메커니즘을 사용하기에 이르렀다. 총 4가지의 Layer를 거치면서 'Extractive QA'에 성능은 Human Performance 수준에 근접하였지만, 이를 뛰어 넘기 위해 RL을 Answer Layer에 적용하고자 한다.

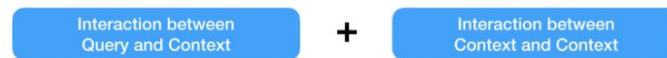
1. Embedding layer : Word-level 표현 강화



2. Encoding layer : Sentence-level 표현 강화



3. Interaction layer : Context 표현 강화



4. Answer layer : Answer 성능 강화



Deep Reinforcement Learning

Language Model에서의 Recurrent Neural Network은 RL의 Environment를 정의한다. State는 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 위치와 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.

