Natural Language Processing with DeepLearning

week 16

The course

1. Recap: language models
2. What does a LM know?
3. Techniques to add knowledge to LMs
    1. Add pretrained entity embedding
    2. Use an external memory
    3. Modify the training data
4. Evaluating knowledge in LMs

<Language Model >

일반적인 language model은 text가 주어졌을 때, 다음 단어를 예측하고 그 확률을 계산한다. 최근에는 masked language model이 text의 일부분을 [mask] token으로 대체하고, 이를 bidirectional context를 이용해 masked token을 예측함

두 방식의 모델 둘다, 많은 양의 unlabeled text를 이용해 학습가능
지금까지의 LM들은 주어진 문장에서의 어떤 특정한 확률을 계산하거나, text를 생성하는 task에 사용됨. 오늘날의 LM들은 NLP task를 수행하기 위해, 언어를 이해하는 방식으로 encoding된 text의 미리 학습된 형태를 생성 (언어, 언어 구조를 이해하고 이를 활용해 답을 도출)

=> 일반적인 지식 기반으로 LM이 사용될 수 있나?

<What does a language model know?>

일반적인 LM으로는 특정한 지식을 묻는 질문에 대해 문법, 문맥상 말이 되는 답을 얻을 수는 얻지만, 지식에 기반한 정확한 정답은 얻기 힘들다. 그 이유는?

- Unseen facts: 학습 시에 보지 못한 정보라서
- Rare facts: LM은 학습동안 사실 정보를 기억할만큼 충분한 예시를 보지 않음
- Model sensitivity: LM은 답을 내놓을 때 prompt에 민감하다.
-

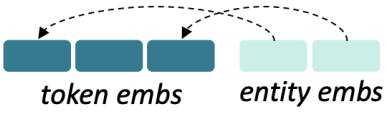
안정적으로 지식을 얻을 수 있는 능력이 부재함 → LM이 마주한 가장 큰 문제
그래서 LM이 궁극적으로 전통적인 지식 기반을 대체할 수 있을까?

<Advantages of language models over traditional KBs>

- LM은 많은 양의 unstructured, unlabeled data로 학습된다
- LM은 좀 더 flexible하다
- but, hard to interpret, hard to trust, hard to modify

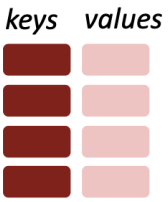
<Techniques to add knowledge to LMs>

Techniques to add knowledge to LMs



Add pretrained entity embeddings

- ERNIE
- KnowBERT



Use an external memory

- KGLM
- kNN-LM

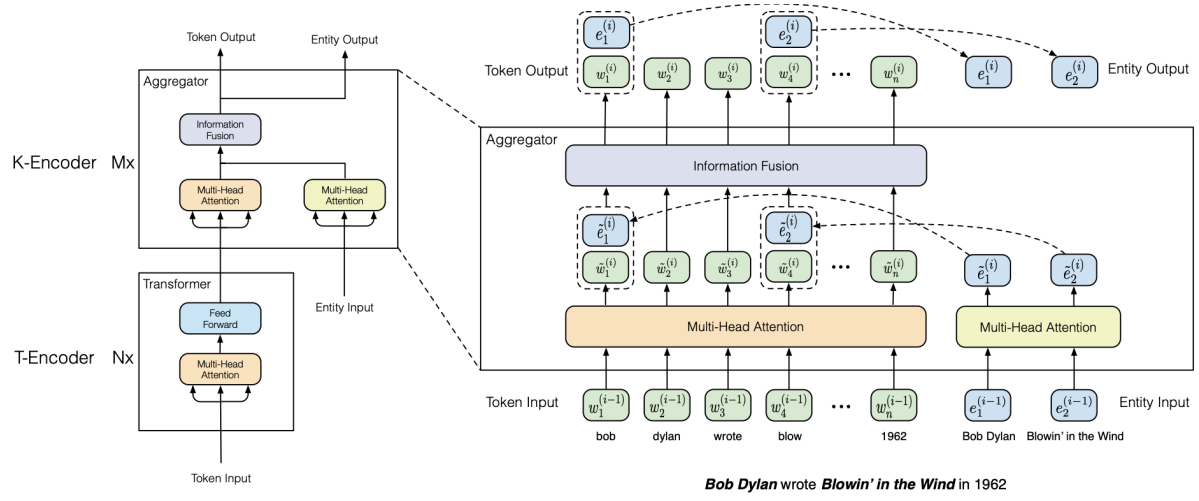
Modify the training data

- WKLM
- ERNIE (another!), salient span masking

<Method 1: Add pretrained entity embeddings>
세상의 사실 정보들은 entities로 이루어져 있음. pre-trained word embeddings는 다양한 entities들에 대한 하나의 개념만을 가지고 있지 않음. 그렇다면 이를 어떻게 나눠서 할당할 것인가?
=> entity linking
word embeddings와 비슷하게 entity embeddings이라는 것도 있음, 그럼 어떻게 다른 embedding space에 있는 entity embeddings를 만들 수 있을까?
=> usion layer를 추가

$$\boldsymbol{h}_j = F\left(\boldsymbol{W}_t\boldsymbol{w}_j + \boldsymbol{W}_e\boldsymbol{e}_k + \boldsymbol{b}\right)$$

<ERNIE>
text encoder + knowledge encoder



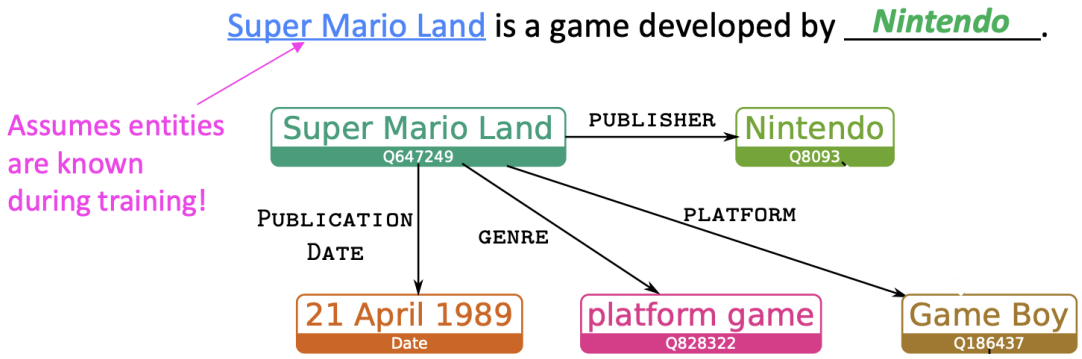
(a) Model Achitecture (b) Aggregator

masked language model, next sentence prediction, knowledge pretraining task로 pretrain knowledge-driven task에서 성능이 향상되었지만, input으로 entities가 annotate된 text input이 필요하다는 점, 또 다른 pretraining이 요구된다는 점이 단점
<KnowBERT>
BERT에 덧붙여서 integrated entity linker를 pretraining한다. downstream task에서 EL은 entities를 예측하기 때문에, entity annotation을 필요로 하지 않음 EL을 학습하는 것이 지식을 더 잘 encoding할 수도 있음
<Method 2: Use an external memory>
이전의 방법들은 factual knowledge를 encoding하기 위해 pretrained된 entity embedding에 의존

그런데 좀 더 직접적인 방법은 없을까?
=> external memory을 가지도록 하면 됨

- pretraining을 하지 않아도 되어서 사실 정보를 업데이트 하거나 새로 넣기 쉬워짐
- 좀 더 해석 가능

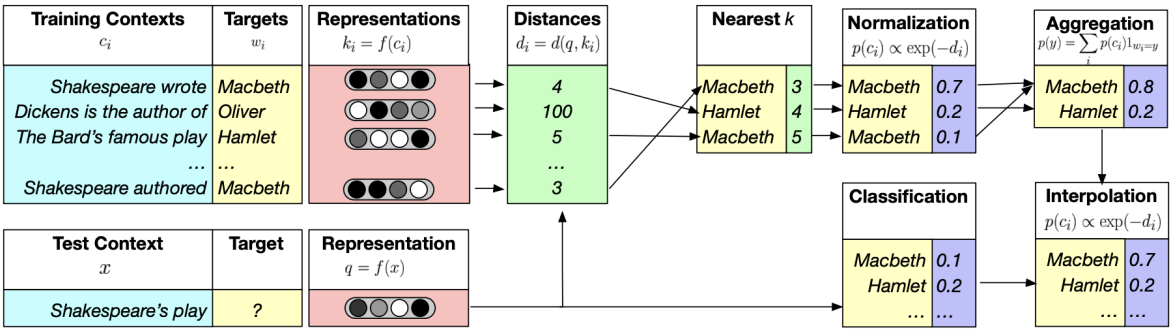
<KGLM>
language modeling처럼, entity information을 활용해서 다음 단어를 예측 이때, local knowledge graph를 만들어서 사용



LSTM을 사용해서 다음 단어의 type을 예측 (related entity, new entity, not an entity)
fact completion task에서 GPT-2, AWD-LSTM을 넘는 성능을 보임, 특히 GPT-2는 일반적인 token을 내뱉는 반면에, KGLM은 좀 더 정확하고 세밀한 token을 내뱉음, 또한 정보를 수정할 수 있음

<kNN-LM>
다음 단어를 예측하는 것보다, text sequences 사이의 유사도를 학습하는 것이 더 쉽다 라는 생각에서 출발. nearest neighbor datastore에 text sequences의 모든 표현을 저장함

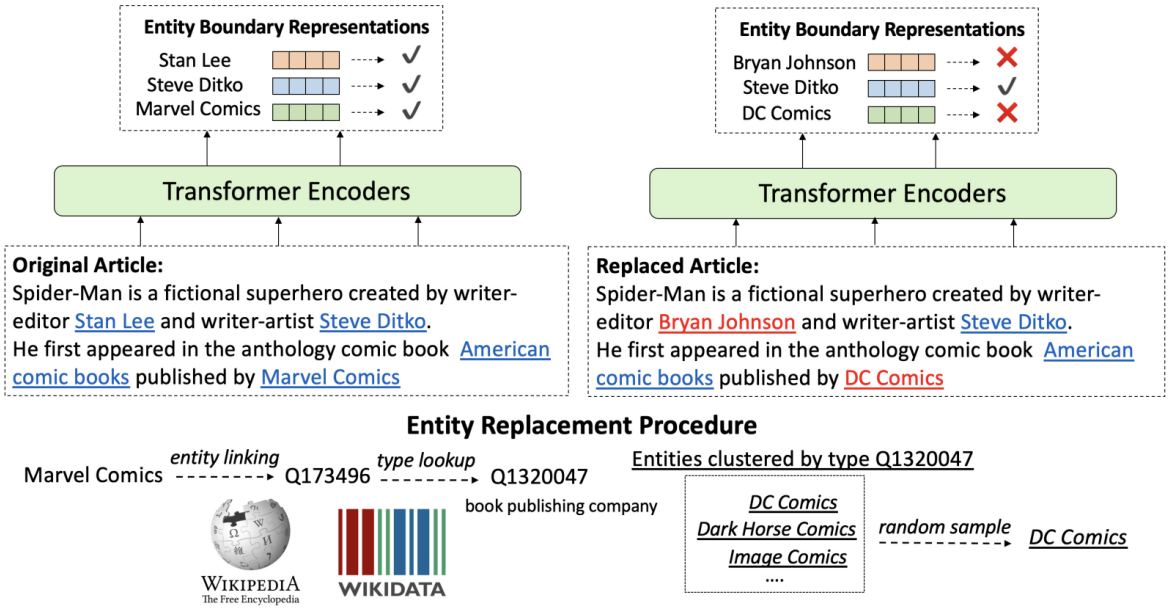
1. datastore에서 text의 k개의 가장 비슷한 sequences를 찾음
2. k sequences의 corresponing values를 검색
3. kNN 확률과 LM 확률을 최종 예측을 위해 합침



<Method 3: Modify the training data>
이전의 방법들은 pretrained embeddings나 external memory를 사용해서 explicitly하게 지식을 표현했다면, implicitly하게 표현할 수는 없을까?
=> data를 mask, corrupt

<WKLM>
true, false knowledge를 구분하도록 모델을 학습

- text내의 mention을 같은 타입을 가진 다른 entities로 교체하여 negative knowledge statements를 만듬
- 그래서 모델이 이것이 진짜인지 아닌지를 예측하도록 학습이 됨
- type을 고정시켰다는 것이 문법적으로 말이 되도록 만들어 줌



MLM은 token-level에서, entRep는 entity-level에서

<Salient span masking>
salient span masking은 T5에서 사용된 masking 기법으로 다른 masking 기법보다 더 좋은 성능을 보임

<LAMA>
얼마나 LM이 합리적인 지식을 가지고 있는가를 측정하는 분석. 수동으로 만든 cloze statements로 구성되어있음
KG triples와 question-answer pair에서 cloze statements를 생성

LM을 supervised relation extraction과 question answering systems과 비교해서, pretrained LM에 내재되어 있는 지식을 평가
(이는 얼마나 다양한 corpus로 학습했는지를 확인할 수 있다.)
LAMA의 단점

- 왜 모델이 잘 작동하는지 알 수가 없음.
- phrasing에 민감하다.
-

<LAMA-UHN>
relational knowledge없이 답할 수 있는 예제를 제거했다. 이로 인해 발견한 사실 → BERT는 예측할 때, entities의 표면에 의존하는 경향이 있음
BERT의 score가 LAMA-UHN을 사용했을 때, 8%까지 떨어질 수 있음
Developing better prompts to query knowledge in LMs
LM은 사실 정보를 알고 있을 수 있지만, LAMA 쿼리 자체로 성능이 떨어지는 것을 볼 수가 있음
= 즉 query에 따라 모델 성능이 크게 좌우된다, 민감하다라고 할 수 있다. 더 많은 LAMA prompts를 생성하고 prompt ensemble을 통해 context의 다양성을 늘리면 성능이 높아짐

ID	Modifications	Acc. Gain
P413	x plays in→at y position	+23.2
P495	x was created→made in y	+10.8
P495	x was→is created in y	+10.0
P361	x is a part of y	+2.7
P413	x plays ~~in~~ y position	+2.2