

DialoGPT

전자공학과 220210031 유현우

1. DialoGPT는 text generation의 하위 분야인 neural response generation을 다룸
2. GPT-2와 같이 DIALOGPT는 autoregressive(AR) language model이며 모델 구성으로 multi-layer transformer를 사용
 - A. GPT-2와 다른 점은 Reddit discussion chain에서 추출된 대규모 대화 pair/session에서 학습된다는 것
 - B. 이 대규모 대화 pairs/session이 DIALOGPT가 대화 플로우에서 $P(\text{Target}, \text{Source})$ 에 대한 joint distribution를 얻을 수 있도록 함
3. Dataset
 - A. Dataset은 위와 같이 Reddit에서 얻어진 comment chain에서 추출
 - i. 이때 thread에 대한 response thread는 하위 thread의 root node를 형성하기 때문에 자연스럽게 tree-structured 기반의 response chain으로 이루어짐
 - ii. Root node에서 leaf node까지의 각각의 path(하위 thread)를 대화의 multi turn을 가진 학습 instance로 사용
 - iii. 그리고 이 데이터셋을 갖고 filtering을 하는데 그 경우는
 1. (1) URL이 있는 source나 target
 2. (2) 3개 이상의 단어 반복이 target에 포함된 경우
 3. (3) 응답이 가장 자주 사용하는 top 50 단어에 적어도 하나 이상 포함하지 않은 경우(예를 들면 the, of, a)(이는 영어가 아닐 수도 있기 때문)
 4. (4) 응답에 "[" 또는 "]"이 포함 된 경우 (이는 markup 언어 일 수도 있기 때문)
 5. (5) source와 target 시퀀스가 합쳐서 200 단어보다 긴 경우
 6. (6) target이 offensive language를 포함한 경우 (대용량 blcoklist에 매칭하는 방법으로 필터링)
 7. (7) 하위 레딧에 많은 수가 offensive한 내용을 포함할 가능성이 많다고 인식되는 경우

8. (8) 단조로운 문장 적극적으로 배제 (1,000번 이상 본 tri-gram의 90%가 포함된 응답)

4. Method

A. 최적화

- i.
$$p(T|S) = \prod_{n=m+1}^N p(x_n|x_1, \dots, x_{n-1})$$
- ii. 이때 multi turn dialogue session은 $p(T_K, \dots, T_2|T_1)$ 로 볼 수 있고, 이는 사실 $p(T_i|T_1, \dots, T_{i-1})$ 에서 i 가 $m+1$ 일 때 조건부 확률을 product한 것
- iii. 결과적으로 $p(T_K, \dots, T_2|T_1)$ 을 최적화하는 것은 모든 $p(T_i|T_1, \dots, T_{i-1})$ source-target 페어를 최적화 하는 과정으로 볼 수 있음

B. Mutual Information Maximization

- i.
$$\hat{T} = \arg \max_T \{(1 - \lambda) \log p(T|S) + \lambda \log p(S|T)\}$$
- ii. 위의 수식에서와 같이 target이 주어졌을 때 source에 대한 값과 source가 주어졌을 때 target에 대한 값을 고려해 줌으로써 mutual information을 강화할 수 있음