

DLGNet: A Transformer-based Model for Dialogue Response Generation

이명재

이전 작업 역사

- 1 turn
 - 사용자 발화가 발생하고, 봇이 응답을 하는 1번의 과정
- 초창기 – Single turn Seq2seq
 - 대화 턴에 대한 장기적 종속성이 없음
 - Multi turn (대화) 에 대한 응답에 실패하는 경향
 - 단조롭고 일반적인 응답
- → HRED (hierarchical recurrent encoder decoder, 계층적 반복 인코더 디코더) 아키텍처와 같은 multi turn model 등장

문제점 지적

- 신경 대화 모델(Neural dialogue models – multi turn model)은 여전히 생성된 응답이 **관련성, 다양성, 일관성 부족**
- 이러한 문제가 발생하는 이유
 - 제한된 종속성 문제 (recurrent model)
 - MLE (maximum likelihood estimation) 의 한계
 - concave entropy profile of dialogue datasets
 - Profile = 윤곽, 옆모습, 단면도
 - out-of-vocabulary 에 의해 다수의 <UNK> 토큰 발생
- 이 네 가지를 봇이 응답을 짧게 하거나, 정보가 없는(무의미한) 일반적인 응답을 생성하는 주요 원인으로 지적

문제점 해결을 위한 우리의 기여 (1)

- **Recurrent model (이전 모델) → Transformer (decoding)**
 - recurrent architecture 의 한계 (gradient vanishing) → turns 과 word 의 제한
 - Transformer 로 변경하여 장기 시간적 종속성을 캡처
 - Vanishing gradient
 - RNN 과 Transformer 는 Auto-regressive 에서 차이가 있음
 - RNN 은 decoder 에서 이전 출력 한 개만 먹여치면서 생성함
 - Transformer 는 모든 이전 출력을 가지고 생성함
 - 이는 multi-turn 시나리오에서 대화 모델링에 매우 유용
 - 대화 맥락에서 더 일관된 응답을 생성

문제점 해결을 위한 우리의 기여 (2)

- **Concave entropy profile 을 막기위해 random informative paddings 을 sequence 앞 뒤에 넣음, MLE 를 보완하는 MMI 사용하여 redundancy 회피**
 - 인간 대화에는 정보가 없는(무의미한) 일반적인 응답이 포함 되어있어, word-level syntactic and utterance-level semantic redundancy(중복) 이 발생
 - 이런 redundancy 은 토큰 위치와 관련된 concave nonuniform sequence entropy profile 에서 온 것이 분명하며,
 - 이 모양은 시작과 끝에 있는 토큰이 중간에 있는 토큰의 엔트로피보다 낮은 엔트로피를 갖습니다
 - 이는 initial positive energy gradient 가 learning barriers 를 생성하여 output distribution 의 구경이 poor 해져서 일반 응답을 함
 - 랜덤 패딩 주입이 없는 모델은, 훈련 데이터의 낮은 엔트로피 영역에서 신속하게 overfit 되어 일반 또는 짧은 응답을 함
- random informative paddings 을 앞 뒤로 넣어 이 오목한 엔트로피 모양을 파괴시켜 일반적인 응답을 피하려 함
- MLE 를 보완하는 MMI 사용하여 컨텍스트에 대한 응답의 조건부 분포 대신 컨텍스트와 응답의 공동 분포를 모델링을 하여 redundancy(중복) 을 회피
 - MMI = maximum mutual information objective
- **짧거나 일반적인 대답 생성 \leftarrow redundancy \leftarrow concave entropy profile**

Concave entropy profile

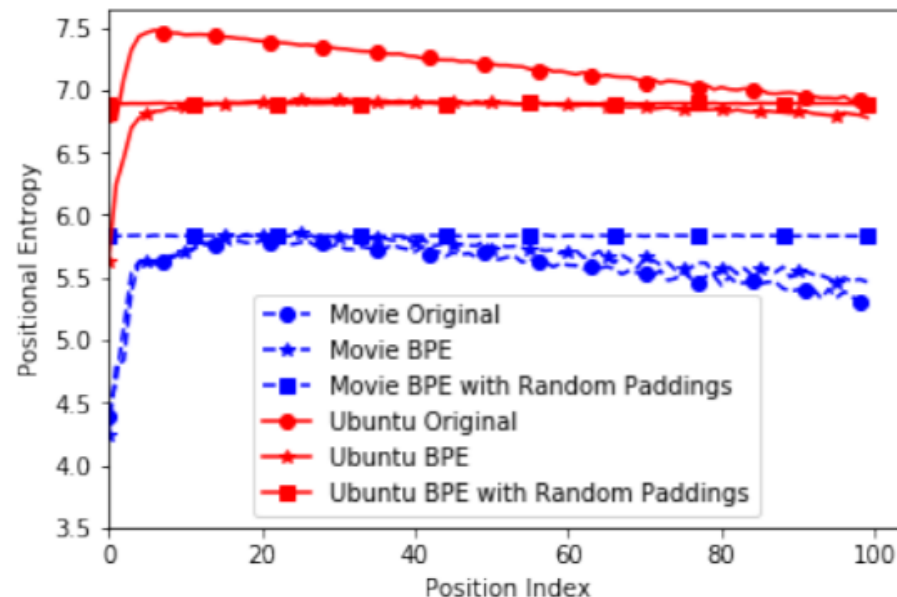


Figure 1: **Positional Entropy for Movie and Ubuntu datasets** - Applying a greedy training objective to the original and BPE datasets can achieve low overall entropy just by overfitting to low entropy regions, resulting in short and generic responses. Injecting random paddings into the data does not suffer from this problem and can be used to train transformer architectures due to their lack of recurrent propagations.

문제점 해결을 위한 우리의 기여 (3)

- 낮은 빈도 단어 및 전문용어가 UNK를 대량 발생 → 데이터 전처리를 제거하고 BPE 로 모든 토큰 100% 활용
 - 대화 모델에서 less frequent words 의 전처리도 word-level tokenization and lowercasing 를 이용하고 이는 <UNK> 로 맵핑 되어, 입/출력 텍스트의 공간을 제한하게 모델링 됨
 - Close domain 의 전문용어 → less frequent words → UNK → 입출력 표현의 한계
 - 이는 데이터 전처리의 문제로 지적
- UNK를 없애기 위해, 기존 데이터 전처리를 전부 제거하고 BPE 로 모든 토큰 100% 활용

문제점 해결을 위한 우리의 기여 (4)

- **Exposure Bias → 대책 없음**

- 학습과 추론의 방법이 달라서 생기는 문제
 - 학습 → teacher forcing
 - 추론 → auto-regressive
- 훈련과 추론 사이의 이러한 불일치는 노출 편향 문제로 알려져 있으며 추론 동안 디코딩 오류가 빠르게 발생하기 때문에 응답의 정보성 (informativeness of the responses) 을 크게 제한

Task 기술 (이전 작업)

$$P(y_i|\mathbf{x}_i) = \prod_{j=2}^{T_i} P(y_i^j|y_i^{1:j-1}, \mathbf{x}_i) \quad (1)$$

where $y_i^{1:j-1} = (y_i^1, \dots, y_i^{j-1})$.

The MLE objective based on the conditional distribution of (1) can be expressed as

$$L_{Cond} = -\log P_{\theta}(y_i|\mathbf{x}_i) = -\sum_{j=2}^{T_i} \log P_{\theta}(y_i^j|y_i^{1:j-1}, \mathbf{x}_i) \quad (2)$$

where θ are the model parameters.

- 컨텍스트에 대한 응답의 조건부 분포
- 이전 대화 turn에서 생산된 대화 y 를 다음 input 으로 넣어 응답 y 를 생성
- X 는 dialogue history

Task 기술 (본 논문)

$$P(y_i, \mathbf{x}_i) = P(y_i|\mathbf{x}_i)P(\mathbf{x}_i) \quad (3)$$

$$L_{Joint} = -\log P_{\theta}(y_i|\mathbf{x}_i) - \log P_{\theta}(\mathbf{x}_i) \quad (4)$$

$$P(\mathbf{x}_i^a, y_i, \mathbf{x}_i, \mathbf{x}_i^b) = P(\mathbf{x}_i^a)P(y_i|\mathbf{x}_i)P(\mathbf{x}_i)P(\mathbf{x}_i^b) \quad (5)$$

and

$$L_{DLGNet} = -\log P_{\theta}(\mathbf{x}_i^a) - \log P_{\theta}(y_i|\mathbf{x}_i) \\ - \log P_{\theta}(\mathbf{x}_i) - \log P_{\theta}(\mathbf{x}_i^b) \quad (6)$$

- Context 과 respond 를 공동으로 모델링 할 것을 제안
- mutual information objective 등 장!

- 4번 식에 input 앞 뒤에 random informative paddings을 넣어 5번 식이 됨

- 6번 식은 5번에 log 를 씌어 loss function 만듦

5번식 참고

- X_a, X_b = 랜덤 정보 패딩 토큰
- X_i 는 발화 문장

설정

- 파라미터 초기화
 - Random noise parameter
 - Pretrained language model parameter
- 학습 속도가 0.001 인 Adam
- 최대 시퀀스 길이는 1024
- GPU 메모리 제한으로 인해 배치 크기 2를 사용
- 5 회 반복하여 그라디언트를 누적
- 배치 크기 10
- perplexity 가 안정 상태에 도달 할 때까지 두 모델을 훈련

데이터 전처리

- BPE 로 100% 토큰화
- 대화의 응답에 맞는 target conversation 을 하나 샘플링
- 무작위 target conversation 에 맞는 2개의 padding chunk를 생성하여 샘플된 target conversation 앞 뒤에 붙여 최대 입력 시퀀스 길이를 채움
- 또한 발화들을 구분 하기 위해 another unique token 을 사용하고 있음

데이터

- The Movie dataset
 - 철자 오류가 거의 없음
 - 광범위한 주제
 - 3 turn 의 240,000 대화 데이터
 - Relevance-diversity (관련성-다양성) 의 trade off 을 연구하기 좋음
- Ubuntu dialog dataset
 - 평균 5 turn 의 1.85 백만 대화 데이터
 - domain-specific expert knowledge/recommendation 에 대한 대화 모델 학습에 적합

탐욕 알고리즘

- 이전 작업 모델들은 출력을 생성하기 위해 greedily sample 을 사용
- DLGNet 은 Top K 사용
- 하지만 $K=1$ 로 도출되어 greedily sample 와 같아짐

정성적 평가

- 응답이 높은 수준의 일관성 보임
- 단기 및 장기 의존성을 모두 포착
- 토론 주제와 관련된 답변을 제공하고, 답변을 선택할 수 있습니다.
- 일반적인 문구를 생성하지 않음

Table 1: Example of generated responses from DLGNet models

Item	Utterance
Movie	
Context_0	says he wanted food . <person> . he wanted the gold .
Context_1	how ' s he going to want the gold ? he couldn ' t even know we had it .
Groundtruth	he . he could have been watching us .
DLGNet-117M ¹	he knew we did .
DLGNet-345M	<person> ' s just say he doesn ' t have enough to buy a trailer and seeds . i ' ll write him up as a decorated veteran of the united states army .
DLGNet-117M_Fresh ²	<person> , you know the break-in , i might not be able to afford it .
DLGNet-345M_Fresh	are you going to preach now ?
Context_0	what are you thinking ?
Context_1	that i want you to take me home .
Groundtruth	now ? we just –
DLGNet-117M	now ? we ' re going now ? what about tomorrow ?
DLGNet-345M	now ? we ' re on the run !
DLGNet-117M_Fresh	now ?
DLGNet-345M_Fresh	no . keep me company . <person> offers its rewards . <person> ap- petit .
Ubuntu	
Context_0	The netboot one is suppose to download packages from the net.
Context_1	like the ones to be installed? or the installed to be run?
Groundtruth	Installed. The netbook also features the non-graphical installer.
DLGNet-117M	the installed to be run.
DLGNet-345M	the ones to be installed.
DLGNet-117M_Fresh	I think the netboot one is the one that is installed to the net.
DLGNet-345M_Fresh	the ones to be installed to.
Context_0	DJones: update manager won't detect 12.04.1 as a new version if you are already running 12.04, because 12.04.1 = 12.04 + lots of package updates
Groundtruth	DJones I am running 10.04.
DLGNet-117M	I'm not sure what you mean by "update manager won't detect"
DLGNet-345M	I see. I guess I'll just do a fresh install.
DLGNet-117M_Fresh	I'm running 12.04.1, and I'm not running 12.04.1.
DLGNet-345M_Fresh	I'm not sure what you mean by "update manager"

정량적 평가

Table 2: Automatic Evaluation of Model Performance

Model	Movie				Ubuntu			
	Relevance		Diversity		Relevance		Diversity	
	BLEU	ROUGE	DIST-1/2	NASL	BLEU	ROUGE	DIST-1/2	NASL
HRED	0.0474	0.0384	0.0026/0.0056	0.535	0.0177	0.0483	0.0203/0.0466	0.892
VHRED	0.0606	0.1181	0.0048/0.0163	0.831	0.0171	0.0855	0.0297/0.0890	0.873
hredGAN_u	0.0493	0.2416	0.0167/0.1306	0.884	0.0137	0.0716	0.0260/0.0847	1.379
hredGAN_w	0.0613	0.3244	0.0179/0.1720	1.540	0.0216	0.1168	0.0516/0.1821	1.098
DAIM	0.0155	0.0077	0.0005/0.0006	0.721	0.0015	0.0131	0.0013/0.0048	1.626
aBoots_u_cat	0.0880	0.4063	0.0624/0.3417	0.918	0.0210	0.1491	0.0523/0.1795	1.040
aBoots_w_cat	0.0940	0.3973	0.0613/0.3476	1.016	0.0233	0.2292	0.1288/0.5190	1.208
DLGNet-117M_Fresh	0.1796	0.4338	0.1198/0.4578	1.011	0.0215	0.1978	0.1827/0.4074	0.829
DLGNet-345M_Fresh	0.2682	0.4881	0.1286/0.4612	0.907	0.0315	0.2041	0.1927/0.4468	0.794
DLGNet-117M	0.1872	0.4346	0.1232/0.4506	0.982	0.0279	0.2191	0.2228/0.4953	0.746
DLGNet-345M	0.2742	0.4945	0.1282/0.4736	0.895	0.0309	0.2409	0.2436/0.5632	0.759

- Fresh = Model without pretraining

Ablation studies (1)

- 테이블 2 참조
- Open domain 이 Close domain 보다 점수가 높음
 - Close domain 의 대화 응답 학습이 더 어렵다는 것을 알 수 있음
- 사전훈련은 크게 도움이 되지 않음
 - 사전 훈련은 Open, close domain 모두 도움이 되기는 함
 - 사전 교육을 받지 않은 모델은 사전 교육을 받은 모델에 비해 평균적으로 더 긴 응답을 생성하는 경향이 있음
 - 이는 모델 사전 훈련이 관련성-다양성 트레이드 오프에 영향을 미침

Ablation studies (2)

- Relevance – diversity trade off
- 더 길고 다양한 응답을 생성 할 수 있는 가능성을 살펴보고 관련성 점수에 미치는 영향을 추정
- top k (Radford et al. 2019) and top p nucleus (Holtzman et al. 2019; Zellers et al. 2019) sampling strategies on the validation sets.
- 상위 k 샘플링의 값을 늘리면, 두 데이터 세트 모두에 대해 BLEU와 같은 관련성 메트릭을 희생하여 응답 길이가 증가
- Ubuntu의 ROGUE-2 점수가 상위 k 값이 증가함에 따라 증가

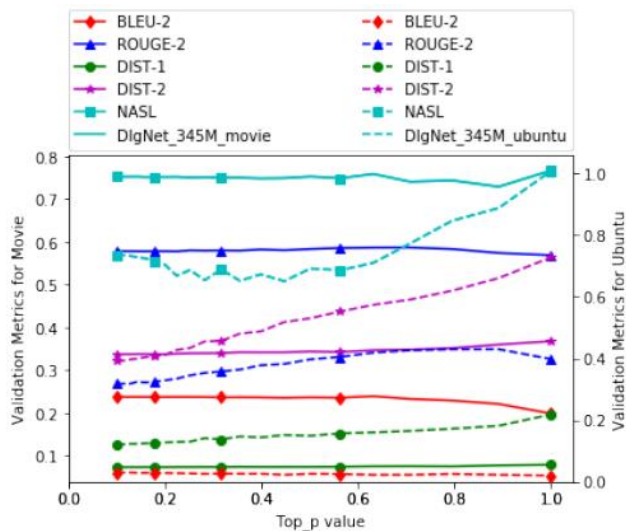


Figure 3: Relevance vs. diversity tradeoff with top_p sampling for DLGNet-345M models.

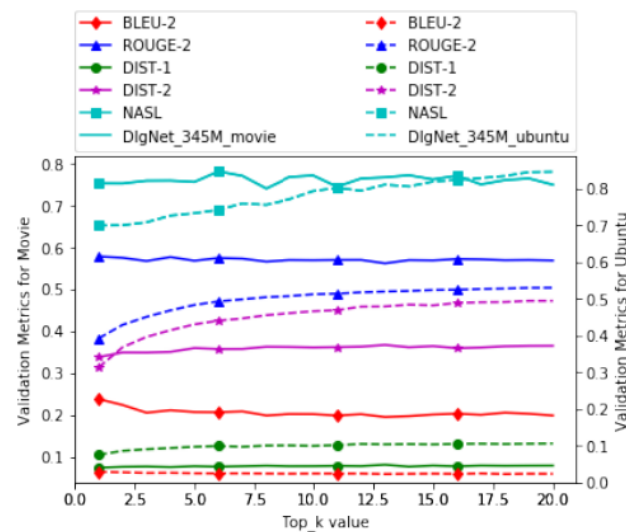


Figure 2: Relevance vs. diversity tradeoff with top_k sampling for DLGNet-345M models.

