Diversity-Promoting GAN: A Cross-Entropy Based Generative Adversarial Network for Diversified Text Generation

Jingjing Xu, Xuancheng Ren, Junyang Lin, Xu Sun

MOE Key Lab of Computational Linguistics, School of EECS, Peking University

jingjingxu,renxc,linjunyang,xusun@pku.edu.cn

Abstract

- **Task**:
  - Text Generation.
- **Problem**:
  - Existing text generation methods tend to produce **repeated and "boring"** expressions.
- **Proposal**:
  - We propose a new model, called Diversity-Promoting Generative Adversarial Network (**DP-GAN**) for diversified text generation.
  - The proposed model assigns low reward for repeatedly generated text and high reward for "novel" and fluent text.
  - Moreover, we propose a novel **language model based discriminator** without the saturation problem.

Approach

Our model contains a **generator** and a **discriminator**. The sketch is shown in Figure 1.

- The **generator** G_θ is responsible for generating text, which is based on a sequence-to-sequence structure.
- The **discriminator** D_ϕ is a language model.
- We define **cross entropy** as the reward to train the generator.

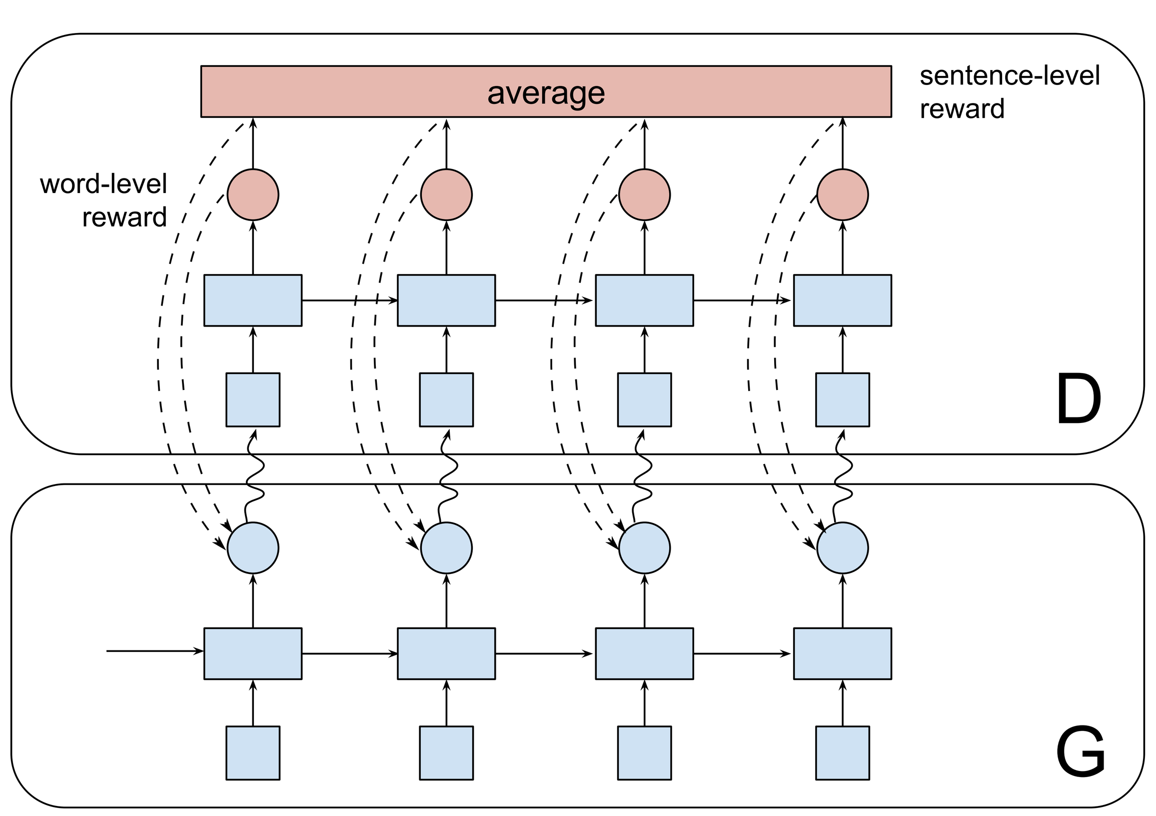

Figure 1: Illustration of DP-GAN. Lower: The generator. Upper: The discriminator.

Rewards

Our reward function consists of two parts:

- **Sentence-Level Reward**

For a sentence y_t of K words, the reward at the sentence level is the averaged reward of each word:

$$R(y_t) = -\frac{1}{K}\sum_{k=1}^{K} \log D_\phi(y_{t,k}|y_{t,<k}) \qquad (1)$$

- **Word-Level Reward**

Considering that the reward for different words in a sentence y_t should be different, we further propose to use the reward at the word level as follows:

$$R(y_{t,k}|y_{t,<k}) = -\log D_\phi(y_{t,k}|y_{t,<k}) \qquad (2)$$

Policy Gradient Training

Adversarial reinforcement training:

```
Initialize G_θ, D_φ with random weights θ, φ
Pre-train G_θ using MLE on a sequence dataset 𝒟 = (X, Y)
Generate samples using G_θ for training D_φ
Pre-train D_φ
N = number of training iterations
M = number of training generator
K = number of training discriminator
for each i = 1, 2, ..., N do
    for each j = 1, 2, ..., M do
        Generate a sequence Y_{1:T} ~ G_θ
        Update generator via policy gradient
        Sample a sequence Y_{1:T} ~ 𝒟
        Update generator parameters
    end for
    for each j = 1, 2, ..., K do
        Generate samples using G_θ
        Train discriminator D_φ
    end for
end for
```

Experiment Dataset

Task:

- Review generation: Yelp & Amazon.
- Dialogue generation: OpenSub.

Results

Automatic evaluation results:

Yelp	Token	Dist-1	Dist-2	Dist-3	Dist-S
MLE	151.2K	1.2K	3.9K	6.6K	3.9K
PG-BLEU	131.1K	1.1K	3.3K	5.5K	3.1K
SeqGAN	140.5K	1.1K	3.5K	6.1K	3.6K
DP-GAN	406.8K	**3.4K**	**22.3K**	**49.6K**	**17.3K**
Dialogue	Token	Dist-1	Dist-2	Dist-3	Dist-S
MLE	81.1K	1.4K	4.4K	6.3K	4.1K
PG-BLEU	97.9K	1.2K	3.9K	5.5K	3.3K
SeqGAN	83.4K	1.4K	4.5K	6.5K	4.5K
DP-GAN	97.3K	**2.1K**	**10.8K**	**19.1K**	**8.0K**

Human evaluation results:

Yelp	Relevance	Diversity	Fluency	All
MLE	1.49	1.73	1.78	1.89
PG-BLEU	1.47	2.59	**1.38**	2.22
SeqGAN	1.48	2.40	1.54	2.12
DP-GAN	**1.32**	**1.23**	1.66	**1.51**
Dialogue	Relevance	Diversity	Fluency	All
MLE	1.19	1.84	1.37	1.87
PG-BLEU	**1.13**	1.85	1.21	1.75
SeqGAN	**1.13**	1.71	**1.20**	1.64
DP-GAN	**1.13**	**1.50**	1.30	**1.55**

Case Study

- Examples generated by our method and the state-of-art models.

Input: *One of my favorite places to eat.*
MLE: *Service is great.*
PG-BLEU: *Service is always good.*
SeqGAN: *Love the chicken and waffles. Service is always great.*
DP-GAN: ***Love the fact that they have a large selection of food. Service is always great and the food is always fresh. I've been to this place a few times and have never been disappointed.***

Input: *I don't think it's too spicy, so i add a little hot curry spice to the sauce.*
MLE: *It's great for cooking. I have to say, I'm not sure if it is the same.*
PG-BLEU: *Love it!*
SeqGAN: *Love it! Love this stuff. I have used it for years and it is very good.*
DP-GAN:***I've tried many different brands of hot sauces but this one is the best. I've also used it in soups and stews. I also like that it is organic and has a great flavor.***

Analysis

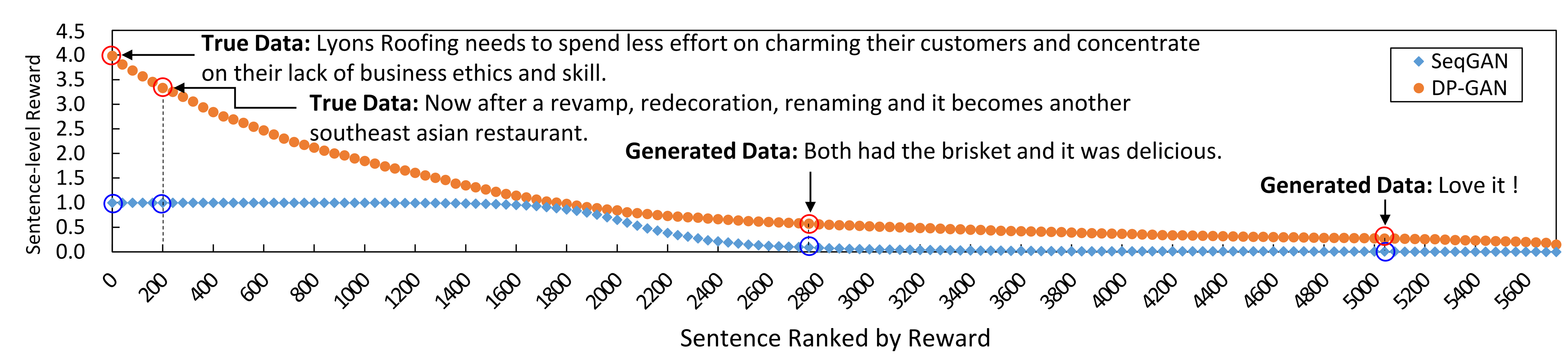

Figure 2: Distribution of rewards between SeqGAN and DP-GAN. The upper two sentences are sampled from the real-world data and the lower two sentences are sampled from the generated data.

Contributions

- We propose a new model, called DP-GAN, for diversified text generation, which assigns low reward for repeated text and high reward for novel and fluent text.
- We propose a novel language-model based discriminator that can better distinguish novel text from repeated text without the saturation problem.
- The experimental results on review generation and dialogue generation tasks show that our method can generate substantially more diverse and informative text than existing methods.