

Adversarial Text Generation via Feature-Mover's Distance

Author: Liqun Chen

Duke University

Presenter: Faizan Ahmad

<https://qdata.github.io/deep2Read>

Outline

- 1 Introduction
- 2 Related Work
- 3 FM-GAN
- 4 Evaluation and Results
- 5 Takeaways

Introduction

- Unconditional text generation - important problem
- Two limitations of log-likelihood based methods
 - Exposure bias
 - Redundant Text
- GANs to the rescue

- Two types of GANs for text generation
 - Reinforcement Learning based - discriminator is trained via policy gradients
 - High Variance
 - Soft approximations e.g Gumbel Softmax
 - Low variance but generic samples and low diversity
- Feature Matching approaches to the rescue
 - TextGAN - Match feature distribution using Maximum Mean Discrepancy
 - Short sentences, kernel choice

FM-GAN

Feature Matching GAN

- Basic Idea: Adversarial Distribution Matching
 - Bring real and synthetic distribution of sentences closer to each other
- How?
 - Earth Mover Distance from Optimal Transport
 - Differentiable
 - Easy to compute

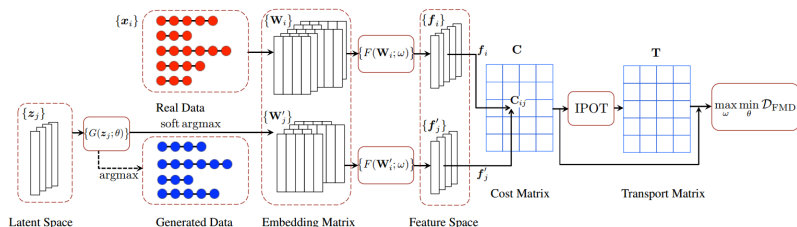


Figure: Proposed System

Adversarial Distribution Matching

- Match synthetic and actual text distribution using GANs loss function
 - p_d - real samples
 - p_G - generated samples
 - D - discriminator

$$\min_G \max_D \mathbb{V}(p_d, p_G; D)$$

- Use Earth Mover Distance (EMD) for training

$$\mathcal{D}_{\text{EMD}}(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \mathbb{E}_{(x, y) \sim \gamma} c(x, y)$$

Complete Algorithm

Algorithm 2 Adversarial text generation via FMD.

- 1: **Input:** batch size n , dataset \mathbf{X} , learning rate η , maximum number of iterations N .
- 2: **for** $\text{itr} = 1, \dots, N$ **do**
- 3: **for** $j = 1, \dots, J$ **do**
- 4: Sample a mini-batch of $\{\mathbf{x}_i\}_1^n \sim \mathbf{X}$ and $\{\mathbf{z}_i\}_1^n \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$;
- 5: Extract sentence features $\mathbf{F} = \{F(\mathbf{W}_e \mathbf{x}_i; \phi)\}_1^n$ and $\mathbf{F}' = \{F(G(\mathbf{z}_i; \theta); \phi)\}_1^n$;
- 6: Update the feature extractor $F(\cdot; \phi)$ by maximizing:

$$\mathcal{L}_{\text{FM-GAN}}(\{\mathbf{x}_i\}_1^n, \{\mathbf{z}_i\}_1^n; \phi) = \mathcal{D}_{\text{FMD}}(\mathbf{F}, \mathbf{F}'; \phi)$$

- 7: **end for**
- 8: Repeat Step 4 and 5;
- 9: Update the generator $G(\cdot; \theta)$ by minimizing:

$$\mathcal{L}_{\text{FM-GAN}}(\{\mathbf{x}_i\}_1^n, \{\mathbf{z}_i\}_1^n; \theta) = \mathcal{D}_{\text{FMD}}(\mathbf{F}, \mathbf{F}'; \theta)$$

- 10: **end for**
-

Dataset	Train	Test	Vocabulary	average length
CUB captions	100,000	10,000	4,391	15
MS COCO captions	120,000	10,000	27,842	11
EMNLP2017 WMT News	278,686	10,000	5,728	28

Table 1: Summary statistics for the datasets used in the generic text generation experiments.

Figure: Datasets used for experiments

- Evaluation Metrics
 - Test-BLEU-n
 - Self-BLEU

Results

Comparison Benchmarks

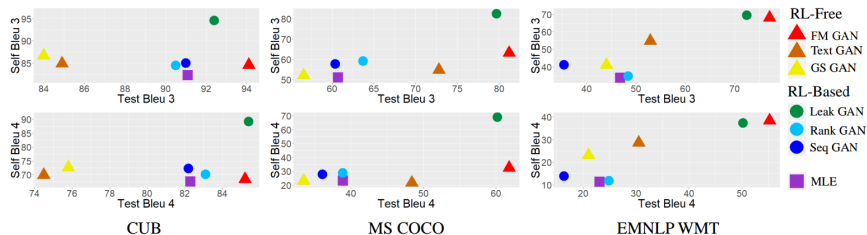


Figure 2: Test-BLEU score (higher value implies better quality) vs self-BLEU score (lower value implies better diversity). Upper panel is BLEU-3 and lower panel is BLEU-4.

Qualitative Results

- FM-GAN:**
- I think about the quality of life and we ' ve got to see you just , and you ' re not happy with you and I ' m sure .
 - The company said the " biggest decision is to be used to provide its own content , which is the biggest test . "
 - I am not in connection with her parents at the age of the girl and she was walking home with her father .
 - It is a huge opportunity to have a lot of time in the world , but we have to see the family of the community in it .
 - It was a great thing for us to be in the case , but I am not sure it was .
 - In the past , they are not aware of the case , which they are not available to them , according to the researchers read .
 - The US , South Korea has been able to provide a new social network from the community and also working together .
 - Women ' s health services have also seen as a report from the UK ' s largest population in 2015 , according to Reuters .
 - Officials who have already had to support from the state to meet with them in the past few years , he has already said .
 - It ' s been clear from this year , so we ' ve been working with many people .
 - I have to be able to make sure you are to make a difference in the community , where they are working with their own care of what they can .
 - The video was a video of a study found that the child had been diagnosed with the same , but the video is found .
 - Clinton has been far more than 200 million in July since 2012 , according to the US Secretary of State .
 - Russia is expected to be a significant increase in the EU referendum on the EU , which has been announced by the EU referendum .
 - " I think that people are trying to find out of the incident , " she told The Sun on the phone and not to be heard .

Takeaways

- Advancing text generation by GANs
- But do we really need GANs for text generation?
 - *OpenAI's new language system writes really great fake news*
 - *Some pretty impressive machine-learning generated poetry courtesy of GPT-2*
- Better applications of these models include style transfer, adversarial samples etc
- Feature matching is an interesting idea in GANs, works well across different domains