Reviewer #1

Questions

1. Please provide an "overall score" for this submission.
7: A good submission; an accept. I vote for accepting this submission, although I would not be upset if it were rejected.

2. Please provide a "confidence score" for your assessment of this submission.
3: You are fairly confident in your assessment. It is possible that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work. Math/other details were not carefully checked.

3. Please provide detailed comments that explain your "overall score" and "confidence score" for this submission. You should summarize the main ideas of the submission and relate these ideas to previous work at NIPS and in other archival conferences and journals. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: quality, clarity, originality, and significance.

The paper describes a language generation approach based on training a GAN with model based reinforcement learning (RL). The RL algorithm gets reward signals from feature matching with a separate guiding LSTM that learns to predict the future c words.

Empirically, on unconditional text generation, the authors showed that their approach (FMGAN) achieves comparable test-bleu scores with LeakGAN and outperforms other GAN approaches such as SeqGAN, RankGAN, GSGAN and TextGAN. Moreover, the authors showed that their approach overfits less than LeakGAN as it achieves a lower self-bleu score. They also conducted human evaluation on a scale of 1 to 5, and FMGAN achieves 3.95 compared to the second best LeakGAN which achieves 3.47.

On conditional generation, they showed that their approach outperforms
i) recent papers for image captioning [12], [31], [32], [33], [34] and SCST [24], and
ii) [27] and [11] on sentiment style transfer.

Strengths:
Quality: the paper proposes a novel GAN formulation for generating text sequences, with strong empirical results.

Originality: The feature matching with predicted futures seem novel. This is different from previous work on feature matching GAN [36] which matches features from ground truth and synthetic sentences.

Significance: Empirically, the authors showed that their approach outperforms a number of recent papers for both conditional and unconditional text generation.

Weaknesses:
Clarity: the paper is clearly written in general, but I have a few comments:
i) I believe two of the baselines, textgan and gsgan, are not defined in the paper.
ii) quick definitions of test-bleu and self-bleu (i.e. bleu between generated samples with training and test set respectively) would be useful.
iii) I can’t seem to find the value of c in the main paper.

4. How confident are you that this submission could be reproduced by others, assuming equal access to data and resources?
2: Somewhat confident

Reviewer #2

Questions

1. Please provide an "overall score" for this submission.
6: Marginally above the acceptance threshold. I tend to vote for accepting this submission, but rejecting it would not be that bad.

2. Please provide a "confidence score" for your assessment of this submission.
4: You are confident in your assessment, but not absolutely certain. It is unlikely, but not impossible, that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work.

3. Please provide detailed comments that explain your "overall score" and "confidence score" for this submission. You should summarize the main ideas of the submission and relate these ideas to previous work at NIPS and in other archival conferences and journals. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: quality, clarity, originality, and significance.

This paper proposes a model-based RL framework for enhancing the sequence generation performance by introducing a guider network to model the environment. The guider network provides a plan-ahead mechanism for selecting the next word. Feature-matching rewards are used to conduct the optimization of the generator via policy-gradient method. Experimental results on the tasks of adversarial text generation, image captioning, and style transfer demonstrate that the proposed framework can obtain better performance than the baseline approaches.

The main contribution of this work is introducing a guider network to overcome the sparse-reward problem in the RL training process for sequence generation models. This problem is well-motivated, and the experimental comparisons and analysis are sufficient and complete. From the results, we can see that the proposed framework can indeed improve the sequence generation performance.

One problem is that technical content is a little difficult to follow. The authors need to add more details to describe the main technical part of the proposed framework. Here are some of my comments/questions in a list form:

- Why use CNN as the encoder to model the sequence considering that the guider network and the decoder are all LSTM based sequence modeling component? How is the performance of employing LSTM as the encoder if you have tried it?
- $f_t = \text{Enc}(Y_{1...t})$, so $f_1$ is obtained based on $Y_1$? Do you have $f_0$ here?
- The output of $G(s_{t-1}, f_t)$ is the hidden state of $Y_{t+1}$?
- During training, the next word is generated based on the information from decoder and the feature from the guider
network. Because the guider network contains the information from the ground truth (i.e., \( f_t = \text{Enc}(Y_{1\ldots t}) \)), so how to conduct the generation during the testing stage considering that there is no ground truth in this stage? Is there any inconsistency between the training and testing?

- Do you use teacher forcing in the decoder during the training stage?

- In the guider network optimization, why you need to look ahead \( c \) steps? Is one step enough? How to conduct the \( c \)-steps look ahead operation?

- In Equation (3), you use the cosine similarity as the reward signal which should be maximized, so why you mention that this objective need to be minimized in line 131 and 132?

Overall, I agree with the motivation of this paper. The general idea of the proposed framework is reasonable and feasible. More technical details will make this work better.

4. How confident are you that this submission could be reproduced by others, assuming equal access to data and resources?
1: Not confident

Reviewer #3

Questions

1. Please provide an "overall score" for this submission.
6: Marginally above the acceptance threshold. I tend to vote for accepting this submission, but rejecting it would not be that bad.

2. Please provide a "confidence score" for your assessment of this submission.
4: You are confident in your assessment, but not absolutely certain. It is unlikely, but not impossible, that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work.

3. Please provide detailed comments that explain your "overall score" and "confidence score" for this submission. You should summarize the main ideas of the submission and relate these ideas to previous work at NIPS and in other archival conferences and journals. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: quality, clarity, originality, and significance.

This paper focuses on model-based RL method for sequence level training, and proposes a guider network to model the sequence-generation environment in the feature space of sequence tokens. Such a guider network can assist next-word prediction and provide intermediate rewards for generator optimization. Besides, a new type of self-attention mechanism is also proposed to assist the guider network in providing high-level planning-ahead information. Experimental results on both unconditional and conditional sequence generation (including adversarial text generation, image captioning and style transfer) show the effectiveness of the proposed method.

Pros:

The definition of the model and the overall architecture of this paper are both clear and straightforward, making this paper easy to understand.

The authors proposed a guider network for sequence-level training, which can provide a plan-ahead mechanism for next-word selection. The idea makes sense to me.
The proposed model is validated on several sequence generation tasks, including adversarial text generation, image caption and style transfer, and the experimental results show improved performance on these tasks.

Cons:

This paper is missing an important related paper ([Dutil+2017NIPS]). The author should add this paper as a baseline model in the conditional sequence generation part.

This paper uses a convolutional neural network as the encoder. I wonder why not adopt RNN-based encoder, which is a widely-used architecture for sequence encoding. I think the author should clarify this point and present experimental results.

Why self-BLEU scores of FMGAN are slightly lower than MLE on COCO image caption, but significantly higher than MLE on EMNLP2017 WMT?
Why the author only conducts human evaluation on WMT news? What about other datasets (especially for the style transfer task)?

Ref:
@proceedings{Dutil+2017NIPS,
author = {Isabelle Guyon and Ulrike von Luxburg and Samy Bengio and Hanna M. Wallach and Rob Fergus and S. V. N. Vishwanathan and Roman Garnett},
title = {Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, Long Beach, CA, {USA}},
year = {2017} }

4. How confident are you that this submission could be reproduced by others, assuming equal access to data and resources?
2: Somewhat confident