ConvGAN: A Context-Aware Conversation GAN

Zejiang Shen & Ruochen Zhang | zejiang_shen@brown.edu, ruochen_zhang@brown.edu

BROWN Computer Science

Motivation

Context

What's unique about conversations? To me, it's about context: the sequential and turn-based mechanism in dialogues requires a high awareness of the context information for the next utterance. We define context as the dialogue does not divert from the ongoing topics given the presented sentences.

Inconsistent Pairs

Generative models are usually focused on mimic the distribution of the outputs to produce context consistent pairs. However, it neglects another important supervision signal from inconsistent pairs. Possibly, being able to learn to discriminate between context consistent and inconsistent pairs will speed up training and bring better generalization results.

Language GANs

Generative adversarial networks (GANs) have achieved promising results in natural language generation models recently. They are able to produce high-quality sentences by overcoming the gradient issues[1, 2, 3]. Additionally, with unsupervised training, they can avoid the exposure bias[4]. It's also easy to construct the context discriminator.

Methods

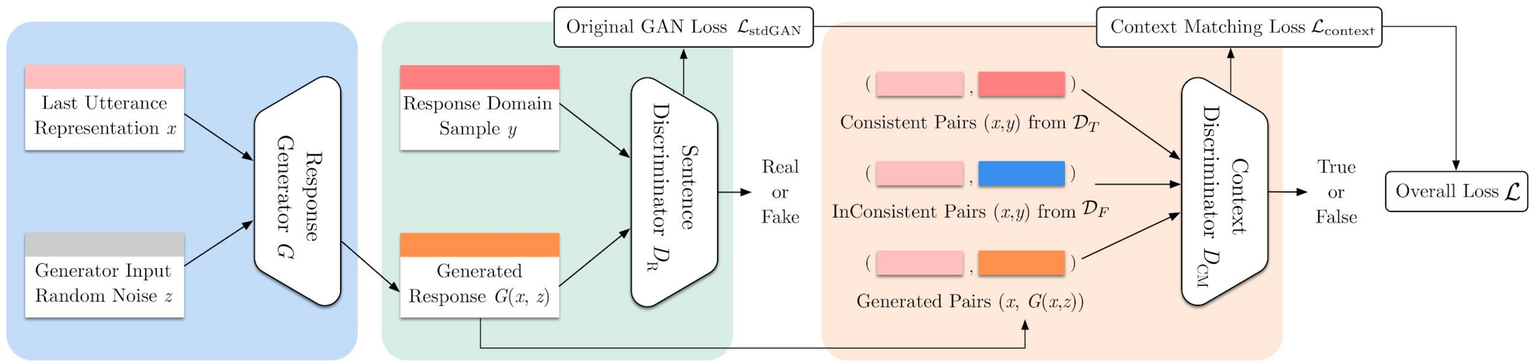
The ConvGAN Architecture

> Generator G: learns to generate response

> Language Discriminator D_R: learns to discriminate a real or a generated sentence

> Context Matching Discriminator D_{CM}: learns to identify whether input sentences pairs are context consistent

> Overall Loss L: considers both sentence generation loss and context loss



i) D_R and D_{CM} share the same encoder weights to encode the input sentence to the latent representation

ii) FMD[1] is used in optimizing the GANs

iii) D_T and D_F are sets contains context matching and mismatching sentence pairs

iv) The GAN is firstly trained with MLE for stable convergence

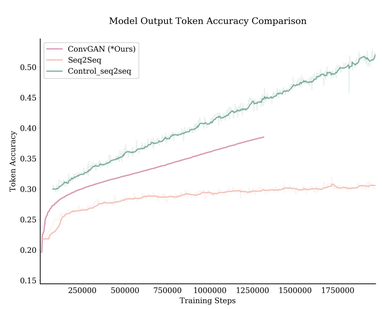
Experiments

Datasets

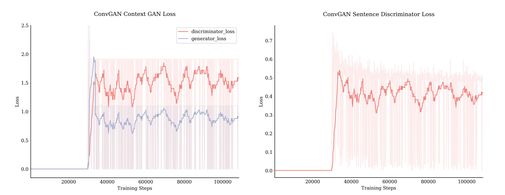
We tested our model on the ConvAI2 dataset[5], where there are more than 10,000 dialogues.

Model Specifications

We implemented the model using ParlAI and PyTorch. The generator was built upon a GRU with 2 layers and a hidden size of 128. Both of the discriminator took in the sentence representations from the GRU encoder's output.

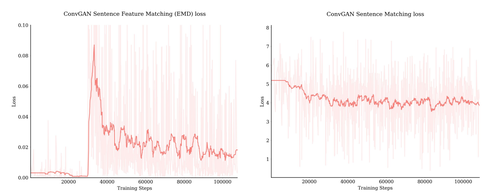


▲ *Our model outperforms the baseline seq2seq model in both converging speed and token accuracy in the limited training epochs.*



▲ *The fluctuating GAN losses indicate they are achieving equilibrium in their min max games.*

Last Utterance	Response in Dataset	Response Generated by Model
i think they are alright . what food do you like ?	i had lunch at taco bell . so exciting .	i enjoy fantasy pizza , being long hair . how bitten plans has wave 30th
yes . i went to a lot of different schools in 12 years	that would be fun . who do you want to see	m on joyfulness for your days with an english .



▲ *EMD Sentence Feature Matching loss is decaying faster: that's what we want!*

◀ *What's in our model's mind? In the first test case, it seems our model has an appetite for pizza ;) . But in the second text case, our model seems a bit confusing about the dialogue in the dataset.*

[1] Liqun Chen, Shuyang Dai, Chenyang Tao, Haichao Zhang, Zhe Gan, Dinghan Shen, Yizhe Zhang, Guoyin Wang, Ruiyi Zhang, and Lawrence Carin. Adversarial text generation via feature-mover's distance. In Advances in Neural Information Processing Systems, pages 4666–4677, 2018. **[2]** Yizhe Zhang, Zhe Gan, Kai Fan, Zhi Chen, Ricardo Henao, Dinghan Shen, and Lawrence Carin. Adversarial feature matching for text generation. In Proceedings of the 34th International Conference on Machine Learning-Volume 70, pages 4006–4015. JMLR. org, 2017. **[3]** Lantao Yu, Weinan Zhang, Jun Wang, and Yong Yu. Seqgan: Sequence generative adversarial nets with policy gradient. In Thirty-First AAAI Conference on Artificial Intelligence, 2017. **[4]** Marc'Aurelio Ranzato, Sumit Chopra, Michael Auli, and Wojciech Zaremba. Sequence level training with recurrent neural networks. arXiv preprint arXiv:1511.06732, 2015. **[5]** Dinan E, Logacheva V, Malykh V, et al. The second conversational intelligence challenge (convai2). arXiv preprint arXiv:1902.00098, 2019.