相比其它方法，GAN 在生成超仿真图像的领域取得了巨大成功，但是它在文本序列领域并没有广泛的应用。这归结于自然文本的离散性，使得标准 GAN 中的传播梯度（propagate gradient）从判别器回到生成器变得难以实现。我们通过使用强化学习（RL）解决了这一问题，利用最大似然和随机梯度下降方法，在训练判别器的同时进行生成器的训练。

GAN 也通常受到诸如训练不稳定性和模式下降（mode dropping）等问题，这两个问题在文本情况下都更加恶化。模式下降会在某些形式的训练集中由生成器很罕见地引起，例如，使得所有生成的火山图像变成同一个火山的多个变体。这在文本生成中变成一个重要的问题，因为数据含有许多复杂模式，从 bigram 到短语到长习语。训练稳定性也是一个问题，因为不同于图像，文本是自回归生成的，从而只有在完整的句子生成之后才能从判别器观察到损失函数的具体值。当生成越来越长的句子时，这个问题会复杂化。

Policy Gradients: <https://www.jianshu.com/p/2ccbab48414b>

Actor-Critic: <https://zhuanlan.zhihu.com/p/25831658>

<https://yq.aliyun.com/articles/176854>

尽管大部分时候，有监督学习比无监督的能获得更好的训练效果。但真实世界中，有监督学习需要的数据标注（label）是相对少的。所以研究者们从未放弃去探索更好的无监督学习策略，希望能从海量的无标注数据中学到对于这个真实世界的表示（representation）甚至知识，从而去更好地理解我们的真实世界。

评价无监督学习好坏的方式有很多，其中生成任务就是最直接的一个。只有当我们能生成/创造我们的真实世界，才能说明我们是完完全全理解了它。然而，生成任务所依赖的生成式模型（generative models）往往会遇到两大困难。

首先是我们需要大量的先验知识去对真实世界进行建模，其中包括选择什么样的先验、什么样的分布等等。而建模的好坏直接影响着我们的生成模型的表现。

另一个困难是，真实世界的数据往往很复杂，我们要用来拟合模型的计算量往往非常庞大，甚至难以承受。

而在过去一两年中，有一个让人兴奋的新模型，则很好地避开了这两大困难。这个模型叫做 Generative Adversarial Networks（GAN），由 [1] 提出。在原始的 GAN paper [1] 中，作者是用博弈论来阐释了 GAN 框架背后的思想。每一个 GAN 框架，都包含着一对模型 —— 一个生成模型（G）和一个判别模型（D）。因为 D 的存在，才使得 GAN 中的 G 不再需要对于真实数据的先验知识和复杂建模，也能学习去逼近真实数据，最终让其生成的数据达到以假乱真的地步 —— D 也无法分别 —— 从而 G 和 D 达到了某种纳什均衡。

**产生的问题**

然而，GAN 虽然不再需要预先建模，但这个优点同时也带来了一些麻烦。那就是尽管它用一个 noise z 作为先验，但生成模型如何利用这个 z，是无法控制的。也就是说，GAN 的学习模式太过于自由了，使得 GAN 的训练过程和训练结果很多时候都不太可控。为了稳定 GAN ，后来的研究者们分别从 heuristic 、 模型改进和理论分析的角度上提出了许多训练技巧和改进方法。

比如在原始 GAN 论文 [1] 中，每次学习参数的更新过程，被设为 D 更新 k 回， G 才更新 1 回，就是出于减少 G 的 “自由度” 的考虑。

1、训练经验角度的改良

另一篇重量级的关于 GAN 训练技巧的研究的工作便是 Deep Convolutional Generative Adversarial Networks（DCGAN）[6] 。[6] 中总结了许多对于 GAN 这的网络结构设计和针对 CNN 这种网络的训练经验。比如，他们用 strided convolutional networks 替代传统 CNN 中的 pooling 层，从而将 GAN 中的生成模型 （G）变成了 fully differentiable 的，结果使得 GAN 的训练更加稳定和可控。

2、加入监督机制

为了提高训练的稳定性，另一个很自然的角度就是改变学习方法。把纯无监督的 GAN 变成半监督或者有监督的。这便可以为 GAN 的训练加上一点点束缚，或者说加上一点点目标。[2] 中提出的 Conditional Generative Adversarial Nets （CGAN）便是十分直接的模型改变，在生成模型（G）和判别模型（D）的建模中均引入 conditional variable y，这个 y 就是数据的一种 label。也因此，CGAN 可以看做把无监督的 GAN 变成有监督的模型的一种改进。这个简单直接的改进被证明非常有效，并广泛用于后续的相关工作中。

3、循序渐进

第三种改进 GAN 过于自由的思路，和第一种会比较相似。既然太难控制 GAN 的学习，不如我们就拆解一下，不要让 GAN 一次学完全部的数据，而是让 GAN 一步步完成这个学习过程。具体到图片生成来说就是，不要让 GAN 中的生成模型（G）每次都直接生成一整张图片，而是让它生成图片的一部分。这个思想可以认为是 DeepMind 也很有名的工作 DRAW 的一种变形。DRAW 的论文 [3] 开篇就说，我们人类在绘制一张图片时，很少是一笔完成的。既然我们人类都不是这样，为什么我们要寄希望于机器可以做到呢？

论文 [4] 中提出的 LAPGAN 就是基于这个思想，将 GAN 的学习过程变成了 sequential “序列式” 的。 具体上，LAPGAN 采用了 Laplacian Pyramid 实现了 “序列化” ，也因此起名做 LAPGAN 。值得一提的是，这个 LAPGAN 中也有 “残差” 学习的思想（与后来大火的 ResNet 也算是有一点关联）。在学习序列中，LAPGAN 不断地进行 downsample 和 upsample 操作，然后在每一个 Pyramid level 中，只将残差传递给判别模型（D）进行判断。这样的 sequential + 残差结合的方式，能有效减少 GAN 需要学习的内容和难度，从而达到了 “辅助” GAN 学习的目的。

4、GRAN

另一个基于 sequential 思想去改进 GAN 的工作来自于 [5] 中的 GRAN。与 LAPGAN [4] 每一个 sequential step（Pyramid level）都是独立训练的不同的是，GRAN 把 GAN 和 LSTM 结合，让 sequence 中的每一步学习和生成能充分利用上一步的结果。具体上来看，GRAN 的每一步都有一个像 LSTM 中的 cell，C\_t，它决定了每一步生成的内容和结果；GRAN 中的 h\_{c,t} 也如 LSTM 一样，代表着 hidden states 。既然是结合 LSTM 和 GAN，那么说完了 LSTM 方面的引入，便是 GAN 方面的了。GRAN 将 GAN 中生成模型（G）的先验也进行了建模，变成了 hidden of prior h\_z；然后将 h\_z 和 h\_{c,t} 拼接（concatenate）之后传递给每一步的 C\_t。

5、Info GAN

最后一种改进 GAN 的训练稳定性的方式则更加贴近本质，也是最新的研究成果。这便是号称 openAI 近期五大突破之一的 infoGAN [7] 。InfoGAN [7] 的出发点是，既然 GAN 的自由度是由于仅有一个 noise z，而无法控制 GAN 如何利用这个 z。那么我们就尽量去想办法在 “如何利用 z” 上做文章。于是，[7] 中将 z 做了拆解，认为 GAN 中生成模型（G）应该包含的 “先验” 分成两种：

**Adversarial Generation of Natural Language [cs.CL] 31 May 2017**

Typically, recurrent and convolutional language models are trained to maximize the likelihood of observing a word or character given the previous observations in the sequence.

These models are commonly trained using a technique called teacher forcing where the inputs to the network are fixed and the model is trained to predict only the next item in the sequence given all previous observations.

This corresponds to maximum-likelihood training of these models. However this one-step ahead prediction during training makes the model prone to exposure bias (Ranzato et al., 2015; Bengio et al., 2015). Exposure bias occurs when a model is only trained conditioned on groundtruth contexts and is not exposed to its own errors (Wiseman and Rush, 2016). An important consequence to exposure bias is that generated sequences can degenerate as small errors accumulate.

Many important problems in NLP such as machine translation and abstractive summarization are trained via a maximum-likelihood training objective (Bahdanau et al., 2014; Rush et al., 2015), but require the generation of extended sequences and are evaluated based on sequence-level metrics such as BLEU (Papineni et al., 2002) and ROUGE (Lin, 2004).

介绍：

1什么是teacher forcing

2 什么是exposure bias

3 ML training会导致什么后果

Generative Adversarial Nets [stat.ML] 10 Jun 2014

**Generating Sentences from a Continuous Space**  [cs.LG] 12 May 2016

Connecting Generative Adversarial Networks and Actor-CriticMethods [cs.LG] 18 Jan 2017

**Adversarial Generation of Natural Language** [cs.CL] 31 May 2017

将GAN用在了离散的output space上，并且用CGAN在sentiment和questions条件下生成句子

批判者：https://www.zhihu.com/question/60902505

LeCun和Yoav Goldberg的大战：

https://blog.csdn.net/dQCFKyQDXYm3F8rB0/article/details/78994637

Goldberg接连在Twitter和medium上言辞激烈地撰文称自己极其不喜欢这篇文章，并且以这篇文章为例，

批判了Arxiv对稿件不加审核的制度，以及深度学习领域生成文本实验不靠谱的普遍现象。

着重讲

* 怎么解决文本discrete的问题
* 用了不同的GAN，效果如何
* 为什么有人批判这篇论文？

**SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient** [cs.LG] 25 Aug 2017

https://blog.csdn.net/sinat\_26917383/article/details/54233599

为什么Ian在reddit上说GAN做不了是因为word embedding加减无意义就做不了呢？既然这样，我在latent vector上做加减不就行了吗？这个方法看上去可以，实际上很难work。

使用generative model解决language generation最大的问题在于latent space存在非常多的desert hole。在training的时候，text的latent vector有聚拢的倾向。不过解决的方法也是有很多的。最简单的方法是用VAE而不是用GAN。GAN本身的训练方式是非常依赖连续空间的。在训练的时候，我们的目标就是连续空间上的pixel值。在这一点上，VAE就没有这个假设。因此VAE是自然的选择。另外一个方法是结合policy gradient，把它做成一个RL的问题。

本文通过把word选择由softmax output选择变成policy选择，作者巧妙的避开了GAN和word embedding不兼容的问题。

Adversarial Learning for Neural Dialogue Generation [cs.CL] 24 Sep 2017

Improved Training of Wasserstein GANs [cs.LG] 25 Dec 2017

MASKGAN: BETTER TEXT GENERATION VIA FILLING IN THE \_\_\_\_\_\_ [stat.ML] 1 Mar 2018

讲清楚mode collapse

https://www.jianshu.com/p/29948a5c5560

为了让G学得更好一点，能不能让G多更新？

- Goodfellow认为是G更新太多次会导致diversity不足，意思就是G独自更新的话容易mode collapse

——————————————————————————————————————

为什么GAN不适合NLP？

作者：随风

链接：https://www.zhihu.com/question/54463527/answer/140864385

解释一下用到latent space的思路

我谈下自己的看法。个人认为GAN不适合NLP还是由于discrete output space的缘故。首先考虑continuous output space (e.g image)， 我们可以直接把generator的output当做input送到discriminator里，从而产生smooth的gradient。对于文本，尤其是当generator的output是句子，需要重新encode这个output，然后送到discriminator，这个结构不是differentiable，只能用REINFORCE（high variance）来train。此外，如果把generator的output distribution按timestep 去decompose；然后在每个time step放一个discriminator，即这个min-max game简化到词的层面时，问题已经和NCE区别不大（当然还是有些细节不同的）。所以，大多GAN在NLP的应用，例如Sam Bowman的文章，都是把discriminator用到latent space(which is continuous)。用到output时需要REINFORCE，个人认为不是个好方向。

——————————————————————————————————————