



# 首届致远学术节 学生科研成果展示

## 基于信息泄露和对抗训练的长文本生成

### 作者及导师

论文作者:

郭家贤 (第一作者)  
卢思迪 (第二作者)  
蔡涵 (第三作者)

导师 (共同作者):

张伟楠 (上海交通大学电子信息与电气工程学院助理教授, 通讯作者)  
俞勇 (上海交通大学电子信息与电气工程学院教授)  
汪军 (伦敦大学学院教授)

### 研究背景

自动生成连贯一致、有明确语义的文本这一技术在很多领域都有着广泛应用。例如, 机器翻译、自动看图写话等问题, 都依赖于文本生成。

近期, 结合基于策略梯度法的强化学习和生成式对抗网络 (Generative Adversarial Nets, GANs) 两个工作, 文本生成这一问题得到了一些初步的有希望的进展, 这一工作被称为序列生成式对抗网络 (SeqGAN)。

SeqGAN使用一个基于循环神经网络的序列式生成器 (RNN-based Sequential Generator) 和一个文本二分类网络 (Binary Classifier Based Discriminator) 作为判别器。SeqGAN在对抗训练时, 其生成器使用强化学习基于其判别器给出的二分类打分来对其输出的结果进行调整, 同时判别器被训练去区分当前的生成器能够输出的文本和真实文本。

然而, 由于SeqGAN在更新其语言模型的策略时从判别器能够获取的信息相比较整个语言的隐分布而言极为匮乏和稀疏, 它并不能充分利用判别器的潜力来对生成器给出足够的改进。这使得生成器最终结束训练时趋向于选择保守的方案, 也就是尽可能输出那些虽然短但是没有丰富语义的句子。

### 研究目的

我们希望找到一个新的框架, 以解决在试图生成较长文本时使用对抗模型所遇到的问题。具体而言, 我们希望找到一个模型和相应的训练算法, 使得我们可以更充分地利用判别器对于目标语言的理解信息。

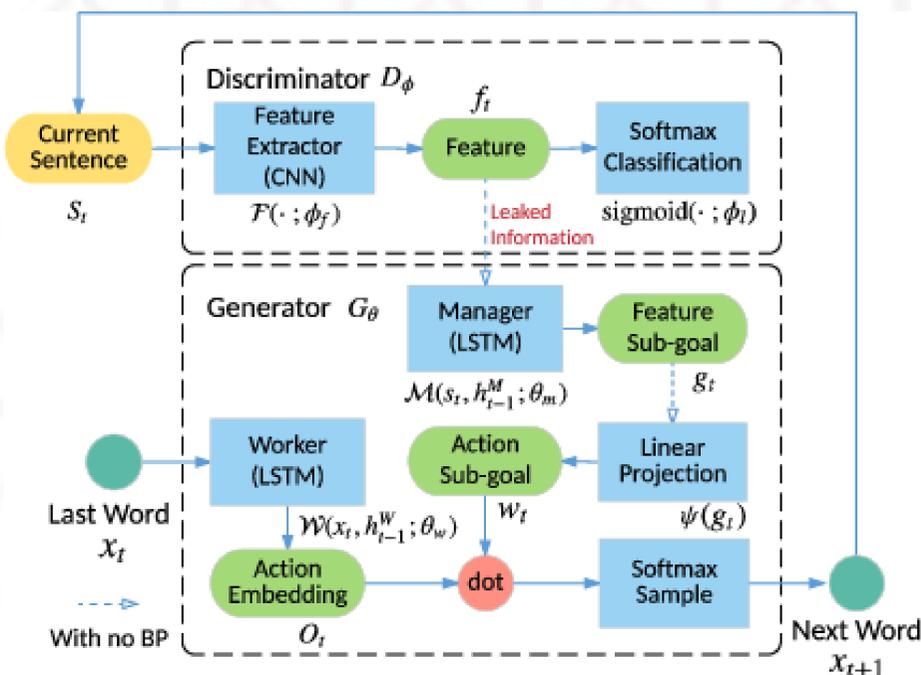
### 研究方法

我们把目光投向了若干地方。

首先, 如果我们打开判别器的黑箱子, 在判别器输出其二分类结果之前, 它已经拥有了一个线性可分的关于当前语言的语义和真实文本的语义分布的建模。如果我们使用一个神经网络去对我们当前的语境在判别器建立的模型中的编码进行降维、差分、拟合的话, 我们是否有希望拿到一个我们该如何在当前状态下使得我们的文本尽可能接近真实分布的指导信息的低维表示呢?

其次, 人类在生成文本的时候不是一蹴而就的。通常, 当人们构思一句话的时候, 人们的逻辑常常具有一定的层次性。这种层次性会导致每一个词在不同的语境下起到的作用常常会很不一样。如果我们能够计算出来每一个词在当前的语境下的语义的一个低维表示, 是否可以帮助我们的训练呢?

再次, 有的时候因为判别器远远强于生成器, 数值上而言, 不管生成器输出什么内容, 判别器给出的打分数量级都非常小, 这不利于生成器改善它的输出。我们可以想象, 如果有一位非常严格的老师, 对于学生考试的能力的提高本该是有好处的, 为什么在这种情况下老师很严格反而是件坏事呢? 当判别器远远强于生成器时, 如何利用这位“严格的老师”来帮助我们的生成器这位“后进”的学生?



### 结果

针对我们目光所及的几个立论点, 我们分别做出了一些改进, 这些改进被以一种非常自然而高效的方式耦合在一起。

针对我们的前两个观察, 我们重构了语言模型中原来的那个基于循环神经网络的序列式生成器, 把它拆成两部分。

第一部分, 仿照我们之前拜读过的一篇文章, 我们称之为“员工” (Worker)。“员工”的工作并不是去最优化语言模型的似然度, 仅仅是去针对当前语境 (使用循环神经网络编码成一个向量) 对每一个可能的词如果被选中会对当前语境造成什么语义影响做出一个估计。

第二部分, 仍然向之前那篇文章致敬, 我们称之为“经理” (Manager)。“经理”也不直接关心语言模型的似然度, 它只负责利用从判别器里面“泄露”出来的那个线性可分的关于当前已生成的文本的特征表示来计算一个低维方向向量, 这个方向向量近似编码了在当前语境下应当往哪个 (低维空间上的) 方向前进。

这两部分的输出会以内积的方式被组合起来, 最终一起去计算似然度。为了探索到更好的最优点, 我们使用了分层估计损失以避免“经理”和“员工”落到一些“妥协”点 (鞍点) 去。

针对第二个问题, 我们仿照批规范化处理 (Batch Normalization) 设计了批排序激活 (Batch Ranking Activation), 使用每个样本在各自小批次的质量估计的排名来代替它拿到的原始分数。这样, 对于每一个小批次来说, 拿到的奖励的均值方差在批排序激活后都是常数, 这大大加速和稳定了训练。

### 结论

我们提出了一种分层的、利用了泄露信息且基于对抗训练的语言模型和一些训练技巧, 并在一些评测指标和模拟真实场景中进行了测试, 结果显示我们的模型成为了所有参评模型中稳定最好的那一个。

### 关键词

自然语言处理 自然语言生成  
对抗模型 对抗信息泄露  
深度强化学习 分层强化学习

个人信息: 卢思迪, 2015级, 计算机  
邮箱: desire2020@sjtu.edu.cn