CSC2547 Learning Discrete Latent Structure: Generative Adversarial Networks via Gradient Estimators

Aparna Balagopalan*
University of Toronto
27 King's College Circle
Toronto, ON M5S
aparna@cs.toronto.edu

Satya Krishna Gorti*
University of Toronto
27 King's College Circle
Toronto, ON M5S
satyag@cs.toronto.edu

Mathieu Ravaut*
University of Toronto
27 King's College Circle
Toronto, ON M5S
mravox@cs.toronto.edu

Raeid Saqur*
University of Toronto
27 King's College Circle
Toronto, ON M5S
raeidsaqur@cs.toronto.edu

Abstract

In this work, we consider the problem of sequence generation using Generative Adversarial Networks (GANs). Sequential data generation using Generative Adversarial Networks presents two structural issues: backpropagating gradients through discrete generator outputs, and waiting for an entire output sequence to be generated before computing the performance score. In this project, we aim to circumvent the above mentioned issues using policy gradients. We build on SeqGAN and compare the performance of the following gradient estimators on token generation: REINFORCE, REBAR, and RELAX.

1 Introduction

The task of meaningful text generation is of crucial interest. To generate long sequences that make sense from end to end, a machine learning model must learn to keep and manage a lot of context and abstract features. In the past years, RNN-based neural networks have been trained successfully to predict the next character or word from a sequence of previous characters or words. However, directly producing full sequences, in natural language for instance, remains an active research area.

In the last few years, Generative Adversarial Networks (GANs) have shown great performance on several generation tasks, especially image generation. However, less work has been done regarding the task of generating sequential data, mainly because of two structural aspects of GANs. First, in GANs, the generator updates by backpropagating gradients from the outputs of the discriminator, which makes little sense when the generated data is made of discrete tokens. Secondly, GANs can only give a score once an entire sequence has been generated, which is a problem with large sequences of tokens.

To alleviate these two issues, SeqGAN takes a reinforcement learning approach while generating tokens, and considers the generative model as a stochastic parametrized policy. This way, the gradient can pass back from the discriminator loss to the generator's weights. When a token is generated, the rest of the sequence is drawn with Monte Carlo sampling, so that the model bypasses the need to wait for a full sequence generation to get a score. It can get a score at each token generation via averaging these Monte Carlo samples. The reward is given by the discriminator's output when fed with the generated sequence. In this paper, we follow the same approach and use the discriminator as our reward function.

* indicates equal contribution

2 Related Work

3 Gradient Estimators

3.1 REINFORCE

3.2 REBAR

3.3 RELAX

4 SeqGAN with policy gradient

5 Backpropagation through time

6 Evaluation

7 Experiments

7.1 Toy data

7.2 Pre-training with MLE

7.3 Comparison of gradient estimators

7.4 Quality of generated sequences

7.5 Training schedule

8 Conclusion