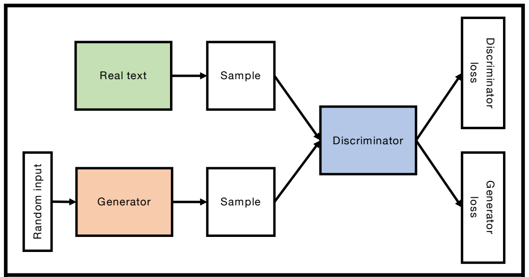
TweetGAN

Natural Language Generation using Tweets

Connor Hainje
Kathryn Leung

Generative Adversarial Networks (GANs)

- Consist of two competing neural networks: a *generator* and a *discriminator*.
- The generator is trained to generate data.
- The discriminator is trained to differentiate between real and generated data samples.
- GANs use *adversarial training*; the discriminator is rewarded for successfully identifying generated samples, and the generator is rewarded for fooling the discriminator.
- After sufficient training, the generator will (ideally) become advanced enough that the discriminator can no longer differentiate.



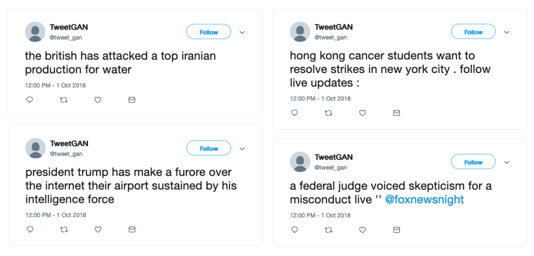
Adversarial training. After pre-training, the generator and discriminator are used to train *each other*, where one network's success penalizes the other.

Text Generation

- Text generation is difficult with GANs, since text is a form of *discrete* data, but most GAN models use a gradient model that relies on the *continuity* of the input and output data
- Solution: SeqGAN! The first successful GAN model to use sequences of discrete data.
  - Models the generator network as a stochastic policy in reinforcement learning

News Tweets

To test our GAN's performance on tweets of a consistent style, we began with a dataset of tweets from news outlets, as they tend to have a characteristic headline-esque style and specific structure. We found that the GAN was able to successfully pick up on the use of phrases like "Follow live updates:". Further, it seems to have generally captured the style of news headlines.



TweetGAN @tweet_gan
the british has attacked a top iranian production for water
12:00 PM - 1 Oct 2018

TweetGAN @tweet_gan
president trump has make a furore over the internet their airport sustained by his intelligence force
12:00 PM - 1 Oct 2018

TweetGAN @tweet_gan
hong kong cancer students want to resolve strikes in new york city . follow live updates :
12:00 PM - 1 Oct 2018

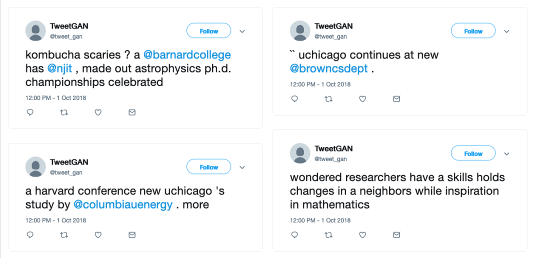
TweetGAN @tweet_gan
a federal judge voiced skepticism for a misconduct live '' @foxnewsnight
12:00 PM - 1 Oct 2018

News samples. Sample tweets generated by our GAN after training on our *news* dataset. We see that the GAN has learned the style of news headlines, especially with tweets like those on the right.

(Fake Twitter screenshots made with tweetgen.com.)

University Tweets

Another source that we tested our GAN's performance with were tweets from official university accounts. These should generally be written with decently formal English, but in a clear style that's less rigid than news headlines. We found, though, that the diversity in topics that university accounts tweeted about left the GAN rather confused, as it struggled to keep various topics separate. It does, however, tend to use various colleges' and departments' Twitter handles when introducing subjects, as done in many university tweets.



TweetGAN @tweet_gan
kombucha scaries ? a @barnardcollege has @njit , made out astrophysics ph.d. championships celebrated
12:00 PM - 1 Oct 2018

TweetGAN @tweet_gan
a harvard conference new uchicago 's study by @columbiauenergy . more
12:00 PM - 1 Oct 2018

TweetGAN @tweet_gan
`` uchicago continues at new @browncsdept .
12:00 PM - 1 Oct 2018

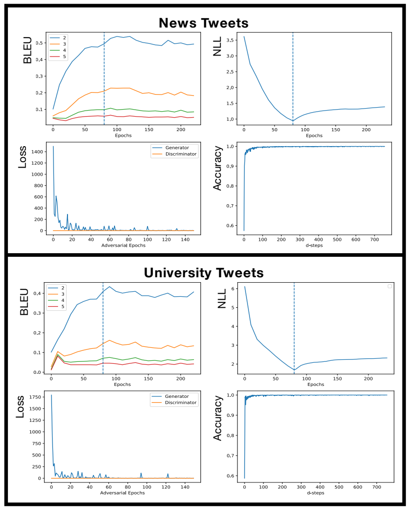
TweetGAN @tweet_gan
wondered researchers have a skills holds changes in a neighbors while inspiration in mathematics
12:00 PM - 1 Oct 2018

University samples. Sample tweets generated by our GAN after training on our *universities* dataset. We find that the GAN generally performs less well here, as the more varied styles and subjects left the GAN confused.

(Fake Twitter screenshots made with tweetgen.com.)

Evaluation

We recorded the BLEU scores and the negative log likelihood (NLL) of the generated samples every ten epochs of pre-training and every epoch of adversarial training. Every adversarial epoch, we also recorded the losses of both networks. At every discriminator training step, we recorded the discriminator's accuracy. For these two plots, the epoch at which adversarial training begins is marked with a dashed vertical line.



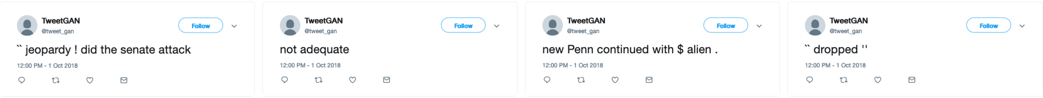
Conclusions

From the evaluation metrics, it appears that, with our parameters, the GAN does not benefit much from the adversarial training. The discriminator easily determines which samples are generated and which are real, and the generator never benefits much from the training. In future studies, we would explore a far wider range of training parameters, as well as datasets restricted to more specific topics.

Acknowledgments

We would like to thank Karthik Narasimhan for his advising on this project, as well as the rest of the COS 484 staff for their aid over the course of the semester.

Honorable Mentions



TweetGAN @tweet_gan
`` jeopardy ! did the senate attack
12:00 PM - 1 Oct 2018

TweetGAN @tweet_gan
not adequate
12:00 PM - 1 Oct 2018

TweetGAN @tweet_gan
new Penn continued with $ alien .
12:00 PM - 1 Oct 2018

TweetGAN @tweet_gan
`` dropped ''
12:00 PM - 1 Oct 2018

PRINCETON UNIVERSITY