Paper: Generative Adversarial Nets

Summary

In this paper, the authors propose a new framework for estimating generative models via an adversarial process. The authors simultaneously train two models. They are generative model G and discriminative model D. The generative model generates the complex random data distribution and the discriminative model estimates the probability that whether the sample came from the training data or the generative model. Therefore, the generative model G is trained to maximize the probability of D making mistake.

The generative model is essentially pitted against an adversary i.e. a discriminative model that learns to determine whether a sample is from the model distribution or the data distribution. The network architecture can be envisioned as a minmax two-player game. Competition in this game drives both teams to improve their methods until the discriminator can no longer tell the difference. The generative and discriminative models are multilayer perceptron. The models are trained using backpropagation and dropout algorithms and sample from the generative model using only forward propagation. Unlike previous generative models, GAN does not require approximate inference or Markov chains.

The generative network G takes a random input z with density pz and returns an output xg that should follow the targeted probability distribution. On the other hand, a discriminative network D takes an input x and returns the probability D(x) of x to be data from the true distribution. This minmax game has a global optimum for pg=pdata, where pdata is the distribution of true data.

The authors trained adversarial nets on a range of datasets including MNIST, the Toronto Face Database, and CIFAR-10. The generator used a mixture of rectified linear activations and sigmoid activations, while the discriminator used maxout activations. They estimated the probability of the test set data under pg by fitting a Gaussian Parzen window to the samples generated with G and reporting the log-likelihood under this distribution.

The author discusses some advantages and disadvantages of using this framework. The advantage is that Markov chains and inference are not needed during learning and a wide variety of functions can be incorporated into the model. Statistically, the adversarial models gain advantage from not being updated directly with data examples, but only with gradients flowing through the discriminator. However, the disadvantage of such a framework is that there is no explicit representation of pg and that D must be synchronized well with G during training.

Strengths

- The analogies that the authors give to explain their model helps the reader understand the model intuitively.
- The paper talks not just about the advantage of such a network, but also disadvantages and ways to improve it in the future.

Weaknesses

- I think the authors made this paper a bit complicated than needed with theorems and weakly explained mathematical equations.

Confusions

- I had a hard time understanding how they evaluated the performance of their model.
- I did not understand some of their visualizations, for example, the arrow pointed in figure 1.

Discussions

- Does the order of min and max function in equation 1 matter?
- How has the concept of GAN been used in modern-day? Are they being used for data augmentation?
- They mentioned that they "alternate between k steps of optimizing D and one step of optimizing G", which they show in algorithm included in the paper. Why does it have to be this way? Can it be done in reverse i.e. switching the place of D and G?