Summary 17: Generative Adversarial Nets

Susan Cherry

This paper proposes a new framework for estimating generative models through an adversarial process. The authors train two models: a generative model G that captures the data distribution and a discriminative model D that estimates the probability that a sample came from the training data and not G. The intuition behind the framework is that the generative model is a team of counterfeiters who are trying to produce fake currency. The discriminative model can be thought of as the police, who are trying to detect the counterfeit money. Competition between the counterfeiters and the police drive both sides to continuously improve their methods. Specifically, this article explores the case where the generative model generates samples by passing random noise through a perceptron. The discriminative model is also a perceptron.

The paper begins by describing adversarial nets, in which both models are multilayer perceptrons. D and G play a two player minimiax game as follows:

$$min_G max_D V(D, G) = E[logD(x)] + E[log(1 - D(G(z)))]$$

Figure 1 explains how to implement this game iteratively and numerically. They alternate between k steps of optimizing D and one step of optimizing G, resulting in D being maintained near its optimal solution as long as G changes slowly enough.

The authors present theoretical results in the next section and describe Algorithm 1, which is minibatch stochastic gradient descent training of generative adversarial nets. They first discuss global optimality of $p_g = p_{data}$ and outline the optimal discriminatory D for a fixed G. They show that the global minimum is achieved if and only if $p_g = p_{data}$. They the show the conditions under which p_g converges to p_{data}.

Then the paper moves on to experiments. They train adversarial nets on a range of datasets. They claim that samples drawn from the generator net after training are at least competitive with the generative models in the literature. Next the authors discuss the advantages and disadvantages of their new framework. The disadvantages are there there is no explicit representation of $p_g(x)$ and that D must be synchronized well with G during training. The advantages are that Markov chains are never needed, only backpropagation is used to obtain gradients, no inference is needed during learning, and many functions can be incorporated into the model. Adversarial models might also gain statistical advantage since the generator network is only updated with gradients flowing through the discriminator, not with data examples. Adversarial networks can also represent very sharp distributions that methods based on Markov chains cannot.

The paper concludes by suggesting so extensions: a conditional generative model, learned approximate inference, semi-supervised learning, and efficiency improvements. Overall, this paper shows that adversarial modeling framework can be a good option and suggests useful research directions to the reader.