Paper: Generative Adversarial Nets, Ian J. Goodfellow et al.

Reviewer: Neeraj Vashistha

The author proposes a new methodology to estimate generative models(GM) by using the existing discriminative models(DM). GMs are used to capture the distribution of data. Due to intractable probabilistic computation it is difficult to approximate the maximum likelihood of the distribution. And unlike DMs where we use hierarchical network with piecewise linear unit and back propagation, the early GMs used Markov chains. The author gives concrete proof by sharing the code and experiments to build a new generative model estimation procedure that overcomes these difficulties.

The author claims to address the issue seen in SML/PCD methodologies proposed by Younes, L. (1999) Tieleman, T. (2008) where Markov chains were used to perform inference approximation and the chances of Markov chain burnout was high. The author coins a new term "adversarial nets" to solve the above problem. The adversarial nets are a combination of generative and discriminative models each having multilayer perceptrons to compete against each other in a way to improve one another. The GM in an adversarial network uses a forward propagation algorithm while DM employs backpropagation. Unlike undirected graphical models where Markov chain Monte Carlo (MCMC) methods are used for learning which are intrinsically difficult to estimate, the author avers to solve this issue of learning problem. In addition to learning, the author suggests that we can use any differentiable function to design the models, which is similar to the research published by Bengio, Y. et al 2014, on Generative Stochastic Network(GSN) which suffered from unbounded activation in feedback loop. The cause of this issue as claimed by the author is use of Markov chains for sampling is still unclear.

One of the main problems of using Generative Adversarial Networks (GANs) and the author fails to address is the convergence issue. GANs use adversarial learning methodology, but convergence of the GM and DM models and existence of equilibrium points have not been proved yet. Another problem the author fails to address is synchronisation of two networks, because of this it is difficult to obtain a stable training process and good training result. Since GANs are based on neural networks, they inherit a common defect of interpretability. Although GANs generate diverse samples but this is followed by another problem, mode collapse problem. In another research, this was highlighted by the author in 2016. Mode collapse refers to the scenario where a generative model generates multiple images with little or no difference to human perception. Martin Arjovsky et al 2017, highlighted another issue pertaining to GANs, he showed that if DM is very good, then GM network would fail due to vanishing gradients, because an optimised DM does not provide sufficient information for the generator to make progress.

The main trend revolves around the research on two challenges, firstly the mode collapse problem and building a better evaluation strategy. Martin Arjovsky et al 2017, suggested the use of Wasserstein loss, as it prevents the mode collapse by letting the discriminator train to optimality without worrying about vanishing gradients. Another approach, suggested by Luke Metz and Ben Poole in 2017 was use of unrolled GANs, the study suggested the use of current and future discriminator classification in the generator loss function. This would prevent over-optimisation of single discriminator by GM. Both the approaches try to force the generator to broaden its scope and prevent it from optimizing for a single fixed discriminator. The evaluation of GANs is a measure of dissimilarity between the generated distribution and the real distribution. There exists very few evaluation metrics for proper estimation of real distribution, which is why the comparison of two distribution is also difficult. The author uses a parzen window based log-likelihood approach proposed by Breuleux et al. 2011, for which exact likelihood are not tractable and is suggested by author as the best known approach of the time. In recent times, Salimans et al., 2016 developed Inception scores and further work was carried out by Upchurch et al. (2017), is the most popular quantitative score used, which assesses the quality and diversity of the generated images using an external model. And the latest work by David Lopez-Paz et al in 2018 for 1-nearest neighbor (1-NN)-based Classifier Two-Sample Tests (C2ST) which learn a suitable representation of the data and return test statistics in interpretable units, measures how generated distribution and the real distribution differ. Mode Score (Che et al., 2016) , Kernel MMD (Gret-ton et al., 2007), Wasserstein distance, Fréchet Inception Distance (FID) (Heusel et al., 2017) are other sample based evaluation metrics for quantitative scores. Other trends involved in working towards resolving the vanishing gradient problem in GM are WGAN by Arjovsky et al 2017, LSGAN by Mao et al, LS-GAN by Qi 2017, RGAN by Martineau 2018 and SN-GAN by Miyato et al 2018. These are some of the improvements in this area.

In recent times, the author's work has been discussed extensively and a lot of important and noteworthy research has been undertaken. GANs have been explored in various fields such as computer vision, game development, music, and natural language processing. Some of the noteworthy ideas on which recent research has been carried are text to image by Han Zhang et al 2017 and Scott Reed 2016, where multiple images are generated which fit a textual description.

Most of the ideas are contained in the computer vision field, from generating high resolution images from low resolution images (SRGAN by Christian Ledig 2017) to generating photo realistic images of an input semantic layout (GauGAN by Taesung Park et al 2019). This list is extensive, "Progressive growing of GANs" by Tero Karras et al 2018 which enhanced the realism in generated images used GANs in a different approach, he used divide and conquer strategy instead of detective and counterfeiters, StarGAN by Yunjey Choi et al 2018, showed us the capabilities of changing emotions in an input image, CycleGAN or cross domain transfer GANs by Zhu et al 2017 worked on images to transform from one domain (say real world) to another domain (Monet paintings or Van Gogh).

In the game development research field, GANs are used not only used to solve optimisation problems but also for content generation one such example is CESAGAN (COnditional EMbedding Self-Attention Generative Adversarial Network) by Torrado R R et al 2019 has used GANs for level generation where MCMC is usually used. Other fields where GANs are popular are music (speech and audio), bio-medical, and information retrieval.

Of the many challenges discussed above, the standardised of evaluation metrics is most important and critical to examine. There are a couple dozens of evaluation metrics proposed each as a quantitative and qualitative measure but each of them have their own limitations. This is crucial to the working of GANs because the evaluation metrics measure the dissimilarity between the generated distribution p_g and the real distribution p_r. Unfortunately, the accurate estimation of p_r is not possible. Thus, it is challenging to produce good estimations of the correspondence between p_r and p_g. If the evaluation metrics can be improved then other problems such as vanishing gradient in generator estimation will automatically be resolved as we do not want our discriminator to be too good. This also solves another issue where the generator over-optimises a particular class as the discriminator fails to learn but memorises the distribution. This problem is known as overfitting in GANs or Mode collapse and good evaluation metrics would always prevent such occurrence as it would penalise discriminator heavily. This kind of method would address the current issues in GANs and without increasing the complexity and side-effects for use of different evaluation metrics.