Review Report on “Generative Adversarial Nets”

# Introduction:

Generative Adversarial Networks is a class in machine learning systems which generates new data with similar statistics as of training data. It involves 2 sets of data models, the generative model which synthesis the trained data and the discriminative network that evaluates the generative model. It is basically the game contest of two neural networks. The generative training data aims at increasing the error rate of the discriminative network by fooling it to synthesize the data that are the part of true data distribution. Typically the generative model is seeded with the random data that is sampled from the functional space. The discriminator evaluates the candidates synthesized by the generative model. Backpropagation is applied in both the networks such that the generator improves the syntheses and the discriminator becomes more skilled in committing the mistake(Wikipedia Jul 2019).

# Content:

This paper proposes a new framework of estimating the generative model by adversarial networks. In this technique, simultaneously two models are trained: a generative model G which is the clone of the training dataset and the discriminative model which estimates the probability of the sample from the training dataset rather than G. This framework exists with a unique solution which lies somewhere between the functions, G and D. Both the models are defined by multi-layer perceptron and hence both can be trained by backpropagation. Therefore the Markov chains(describing the probability sequence of events from the past occurred events(Henry Maltby 2019)) or approximate inference networks techniques are not needed. Relevant examples are quoted in the paper to establish the quality and quantity of the generated samples from this framework.

This framework is proposed to overcome the difficulties of the generative model. The deep generative model had less impact on the computation because it had the challenge to approximate many intractable probabilistic computations that aroused during the maximum estimation strategies and also due to the difficulty in generating the benefits of piecewise linear units in the generative models.

The research focuses on competing the generative models against an adversarial training dataset called discriminative network, which comprises the sample from either model distribution or data distribution. It mentions that the competition between both the models that help to improve their model until the imitations become imperceptible to the genuine articles. The paper simply states that the generative model having associated with the team of fakers producing the fake currency should align with the discriminative model, which is analogous to the police trying to detect the fake currency. To establish the generative model, a special case called adversarial nets is referred, in which random noise has been seeded through the multi-layer perceptron, and the samples are generated. Both the models can be trained using highly successful backpropagation, and dropout algorithms and the generated samples from the generator are trained with the forward propagation algorithm. This framework can produce a specific training model for different varieties of model and optimization algorithm.

# Innovation:

The background to this paper mentions that there are many other algorithms which incur computational difficulties associated with both direct and undirected graphical models, and it is not possible to derive the manageable unnormalized probability density. For example, restricted Boltzmann machines (RBMs), deep Boltzmann machines (DBMs) which are undirected graphical models, and are represented as the unnormalized functions which can be normalized by the summation of overall states of the random variables. The model and the gradient are unmanageable but can be estimated by Markov Chain Monte Carlo (MCMC) methods. But during estimation, the mixing of the variables can be problematic for the learning algorithm. Deep belief networks (DBNs) containing the fast layer-wise training of single directed and several undirected layers encounter the difficulty of computation in both direct and undirect models.

Alternative criteria have also been proposed, such as score matching and noise-contrastive estimation(NCE), which do not approximate the samples. In this case of NCE, there is no separate discriminative network incorporated into the model, a discriminative model is fitted into the generative model, hence the samples are generated from the generative model using a fixed noise distribution, because of which the generative learning dramatically slows down and the model has approximately learned the correct distribution with the observed variables from the dataset.

Furthermore, it states that some techniques do not involve finding the probability distribution explicitly; instead, the generative model is used to derive the samples from the desired data distribution. Thus the machines can be designed to be trained by backpropagation. Such kind of approach is used in generative stochastic network(GSN) framework, which uses the Markov chain in defining the parameters such that one learns the parameters of the machine that performs one step in the Markov chain followed by the next step. Thus to overcome this difficulty of feedback loops during the generation of samples the idea of adversarial networks framework is introduced which can leverage the piecewise linear units thus improving the performance of the backpropagation but encounters the problem when used in a feedback loop with unbounded activation.

# Technical Quality:

This new framework involves the interpretation of theoretical and practical experiments showing the data distribution with enough capacity in a non-parametric limit. In practical experiments, it is implemented using the iterative approach by optimizing the distribution network of k steps and optimizing generating data network to one step. This results in maintaining the optimal solution between the generative and distributive networks. These techniques are very well documented in the paper with a number of theorems and practical research. The paper clearly states the number of advantages of this technique that outweighs the disadvantages. The only disadvantage is that the generative functional space should be well synchronized with the distributive network and should not be trained too much without updating the discriminative space to avoid the redundancy of data. The advantage is that the Markov chain, the inference is not needed during training, only backpropagation is used, and also wide varieties of functions can be integrated into the model.

There are many statistical advantages as well that are technically drafted to support this theory which mentions that the adversarial networks have the advantage of generator networks not being directly updated with the data samples, instead uses the gradient from backpropagation used for the discriminator. The other main advantage cited is that the adversarial networks interpret very sharp images, whereas the Markov chains method needs the distribution to be blurry because the chain should be missed with the modes for the proper output.

# Application and X-factor:

The overall content of the paper is exciting, promising, and it led me to understand the difference between adversarial network and other networks for the data modelling. The framework can be used to implement a stochastic extension of the deterministic Boltzmann machines. The application domain is well suited to this technique, and this technique can also be used to other application domain like Deep belief networks (DBNs), Markov Chain Monte Carlo (MCMC), noise-contrastive estimation(NCE), deep and restricted Boltzmann machines. This framework can be used to generate conditional generative model, learned approximate inference algorithm, semi-supervised learning features from discriminator network or inference networks which can improve the performance of the classifiers when there is an availability of limited labelled data, efficiency improvements of the training which could be accelerated dramatically by producing better methods for coordinating discriminative and generative networks and also for the data distribution to samples from training data. The study to automate the synchronization of the discriminator and generator, which can normalize the redundancy in data should be considered in future work. The efficiency of the network should also be maintained simultaneously, which is a critical factor in data modelling. This framework is impressive, and it may spark a discussion in the tutorial as there are many theorems and algorithms used to support this technique.

# Presentation:

The paper is presented in a well-organized manner. Overall the writing is crisp, clear, and it has very knowledgeable fact and theories which helped me to understand the generative adversarial networks. The sections of the paper are so intense and deep that it helped me to understand the concepts very easily. The establishment of the algorithms, theorems, experiments, practical proposition was so informative that the thesis stands well acknowledged and well interpreted. The author combined many different direct and undirected graphical models to describe the adversarial networks descriptively. The depth of each model was interesting, and the connection between these models and adversarial networks were so keen that the flow of the work was easily understood.

# References:

Henry Maltby, W.P., Jeremy Jackson, Adrian Hernandez, Christopher Williams, Calvin Lin, Jimin Khim 2019, *Markov Chains*, viewed 23 Aug 2019, <<https://brilliant.org/wiki/markov-chains/>>.

Wikipedia Jul 2019, *Generative adversarial network*, viewed 23 Aug 2019, <<https://en.wikipedia.org/wiki/Generative_adversarial_network>>.