Opportunities and challenges in deep generative models

Evgeny I. Nikolaev
Institute of Information Technologies and Telecommunications
North-Caucasus Federal University, Stavropol, Russia
notdeveloper@gmail.com

Abstract

A Generative Model is a powerful way of learning any kind of data distribution using unsupervised learning and it has achieved tremendous success in just few years. Though there are several approaches to design information systems for generating synthetic data, wich are referred to as Deep Generative Model (DGM). Since then, DGM has become a trending topic both in academic literature and industrial applications. It is also receiving increasing attention in machine learning competitions. This paper aims to provide an overview of the current progress towards DGM, as well as discussing its various applications and open problems for future research. Moreover, we discuss some research we conducted during last years that may extend the existing state of the art approaches in synthetic data generation or improving existing deep models.

1 Introduction

Generative models have a long history and recent methods have combined the generality of probabilistic reasoning with the scalability of deep learning to develop learning algorithms that have been applied to a wide variety of problems giving state-of-the-art results in image generation, text-to-speech synthesis, image captioning and data augmentation, amongst many others. Advances in deep generative models are at the forefront of deep learning research because of the promise they offer for allowing data-efficient learning, and for model-based reinforcement learning. The latest advances in generative modeling include following types of models: Markov models, latent variable models and implicit models. These models can be scaled to high-dimensional data.

From the historical perspective it is important to mention research linked with Boltzmann machine [Hinton84, Hinton86], Restricted Boltsmann Machine [Smol86, Long10], Deep Belief Network [Hinton06, Hinton08], Deep Boltzmann Machine [Salakh09], Convolutional Boltzmann Machine [Desj08]. In recent years, there has been resurgence of interest in deep generative models. Emerging approaches such as Variational Autoencoders [Kingma13, Rezend14], Generative Adversarial Networks [Goodfellow16], auto-regressive networks (pixelRNNs [Oord16], RNN language models [Zarem14]), and many of their variants and extensions have led to impressive results in different applications. Researchers are making great progress in generating realistic high-resolution images [Woo17], manipulating and changing text, learning interpretable data representations, automatically augmenting data for training the models.

In: Marco Schaerf, Massimo Mecella, Drozdova Viktoria Igorevna, Kalmykov Igor Anatolievich (eds.): Proceedings of REMS 2018 – Russian Federation & Europe Multidisciplinary Symposium on Computer Science and ICT, Stavropol – Dombay, Russia, 15–20 October 2018, published at http://ceur-ws.org

2 Generative models

All types of generative models aim at learning the true data distribution of the training set so as to generate new data points with some variations. But it is not always possible to learn the exact distribution of our data either implicitly or explicitly and so we try to model a distribution which is as similar as possible to the true data distribution. For this, we can leverage the power of neural networks to learn a function which can approximate the model distribution to the true distribution.

Two of the most commonly used and efficient approaches are Variational Autoencoders (VAE) and Generative Adversarial Networks (GAN). VAE aims at maximizing the lower bound of the data log-likelihood and GAN aims at achieving an equilibrium between Generator and Discriminator. Many researches attempt to compile a unified view: new formulation of GANs and VAEs, and linked back to the classic variational inference algorithm and the wake-sleep algorothm.

2.1 Variational Autoencoder

The variational autoencoder [Kingma13, Rezend14] is a directed model that uses learned approximate inference and can be trained purely with gradient-based methods. An autoencoder can be used to encode an input image to a much smaller dimensional representation which can store latent information about the input data distribution. The variational autoencoder approach is theoretically pleasing and simple to implement. It also obtains excellent results and is among the state-of-the-art approaches to generative modeling.

One of the main features is that samples from VAEs trained on images tend to be somewhat blurry. The causes of this phenomenon are not yet known. The key idea of VAE are shown in Fig. 1.

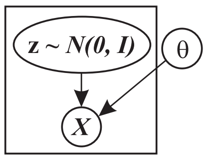

Figure 1: Mapping latent vector to data distribution using parameter

The primary objective is to model the data X with some parameters which maximizes the likelihood of training data X. In short, we are assuming that a low-dimensional latent vector has generated our data $x(x \in X)$ and we can map this latent vector to data x using a deterministic function $f(z;\theta)$ parametrized by theta θ which we need to evaluate. During generative process, our aim is to maximize the probability of each data in X.

2.2 Generative Adversarial Network

Generative adversarial networks, or GANs [Goodfellow16], are another generative modeling approach based on differentiable generator networks. Adversarial training has completely changed the way we train the artificial neural networks. GAN dont link with any explicit density estimation like VAE. GAN is based on game theory approach with an objective to find equilibrium between the two networks: generator and discriminator. The aim is to sample from a simple distribution and then learn to transform this noise to data distribution using approximators such as neural networks. This approach is shown in Fig 2

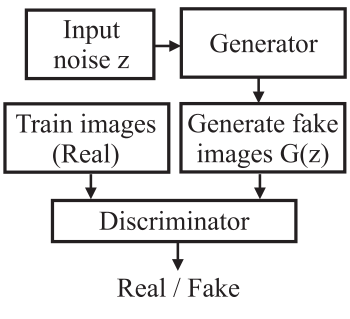

Figure 2: Training Generative Adversarial Network

We can formulate learning in GAN as a zero-sum game, in which a function $v(\theta^{(g)}, \theta^{(d)})$ determines the payoff of the discriminator. During learning, each player attempts to maximize its own payo, so that at convergence $g^* = \arg\min_g \max_d v(g, d)$. One of the earliest model on GAN employing Convolutional Neural Network (CNN) is Deep Convolutional Generative Adversarial Networks (DCGAN) [Radf17].

3 Conclusions and Future Work

Unsupervised learning is a next frontier in artificial intelligence. One of the main advantages of Generative Models is a possibility of the training in semi-supervised manner. Such models can be applied for solving complex problems: text-to-image translation, synthetic images generation, solving problems with multimodal data distribution, drug discovery, visual marks retrieval from images. DGM is a way to improve existing discriminative models. GANs help to solve the one of the main challenges in deep learning: huge amount of labelled data. These models help in building a better future for machine learning.

References

[Hinton84] G. E. Hinton, T. J. Sejnowski and D. H. Ackley Boltzmann Machines: Constraint satisfaction networks that learn. *Technical Report CMU-CS-84-119, Carnegie-Mellon University.*

[Hinton86] G. E. Hinton, T. J. Sejnowski Learning and relearning in Boltzmann machines. *In Rumelhart, D. E. and McClelland, J. L., editors, Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Vol. 1: Foundations, MIT Press, Cambridge, MA. pp 282-317

[Smol86] P. Smolensky Information processing in dynamical systems: Foundations of harmony theory. *Parallel distributed processing: Explorations in the microstructure of cognition, MIT Press, Cambridge, MA, (1986).*

[Long10] P. Long, R. Servedio. Restricted Boltzmann Machines are Hard to Approximately Evaluate or Simulate. *27th International Conference on Machine Learning (ICML).* 2010.

[Hinton06] G. E. Hinton, S. Osindero and Y. Teh A fast learning algorithm for deep belief nets. *Neural Computation.* Vol. 18, 2006, pp. 1527-1554.

[Hinton08] G. E. Hinton, R. Salakhutdinov Using Deep Belief Nets to Learn Covariance Kernels for Gaussian Processes. *Advances in Neural Information Processing Systems.* Vol. 20, 2008.

[Salakh09] G. E. Hinton, R. Salakhutdinov Deep Boltzmann Machines. *To appear in Artificial Intelligence and Statistics.* 2009.

[Desj08] G. Desjardins, Y. Bengio Empirical evaluation of convolutional RBMs for vision. *Technical Report 1327, Dpartement dInformatique et de Recherche Opra-tionnelle, Universit de Montral.* 2008.

[Kingma13] D. Kingma, M. Welling Auto-Encoding Variational Bayes. *Proc. of the 2nd Int. Conf. on Learning Representations (ICLR)*, arXiv: 1312.6114 [stat. ML].

[Rezend14] D. J. Rezende, S. Mohamed, D. Wierstra Stochastic Backpropagation and Approximate Inference in Deep Generative Models. *Proc. of the 31st Int. Conf. on Machine Learning (ICML)*, vol. 32, 2014.

[Goodfellow16] Goodfellow Nips 2016 tutorial: Generative adversarial networks. *NIPS 2016*, arXiv:1701.00160, 2016.

[Oord16] Aron van den Oord, N. Kalchbrenner, K. Kavukcuoglu Pixel Recurrent Neural Networks. *International Conference on Machine Learning (ICML).* 2016.

[Zarem14] W. Zaremba, I. Sutskever and O. Vinyals Recurrent Neural Network Regularization. *arXiv:1409.2329 [cs.NE].* 2014.

[Woo17] B. Wu, H. Duan, Zh. Liu, G. Sun SRPGAN: Perceptual Generative Adversarial Network for Single Image Super Resolution. *arXiv:1712.05927 [cs.CV].* 2017.

[Radf17] A. Radford, L. Metz, S. Chintala Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *Int. Conf. on Learning Representations (ICLR).* 2016.