ECE 6254 Project

Shenao Zhang, Tingyu Zhang, Haidong Zhi, Xueren Ge, Xinghao Chen

1 Project Summary

We choose generative models as the topic of our project. Specifically, we focus on two widely used frameworks: estimating generative models via an adversarial process and estimating intractable posterior via stochastic variational inference.

1.1 Why is it interesting?

We have seen datasets that contain real-world images in class, e.g., MNIST. But due to the expensive costs to take pictures that are in good condition, it will be very exciting to see generated samples that are similar with the images in the original datasets. In other words, estimate the distribution of the datasets, and sample from it to generate unreal but lifelike images. In fact, this is already a hot topic in machine learning and the two most popular frameworks are GAN and variational inference based methods. We think generative models will further benefit data-driven machine learning researches, and the probabilistic contents (e.g., MLE), the qualitative analysis (e.g., dimensionality reduction analysis) are closely related to our Statistical Machine Learning course.

1.2 What the proposed project will consist of?

We will first get familiar with GAN [4] and VAE [5], understanding the methods and formulas. Then we will conduct both quantitative and qualitative experiments. Experiments consist of the observations in the training process (e.g., training curves and generated samples), visualizations (e.g., t-SNE and dimensionality reduction), testing analysis (e.g., the quality of generated images),and ablation studies (e.g., the effects of hyper-parameter selection and data augmentation). Training datasets consist of MNIST and CIFAR10. Based on the drawbacks of the GAN and VAE papers and observations in experiments, we will read more papers to improve them. We will summarize all the relevant works that we encounter. We'll reformula the training process and formulas, analyzing what leads to improvement. If we have extra time, we'll reimplement the methods and compare the experimental results to give a better understanding.

2 Necessary background information and notation

The main purpose of generative models in images is to model the probability of image distributions in the (potentially large) dataset. One direct idea that comes to mind is maximum likelihood estimation (MLE), which we are familiar with. However, the probabilistic computations are usually intractable. For example, if we model the dataset images x with distribution $p(x)$, we can choose mixture models to separately fit two tractable (e.g., Gaussian) models $p(z)$ and $p(x|z)$, and calculate $p(z) = \int p(x|z)p(z)dz$, where z is the latent variable. From MLE, $\theta = \arg\max_\theta \frac{1}{N}\sum_i log p_\theta(x_i) =$

$\arg\max_{\theta} \frac{1}{N}\sum_i logp_{\theta}(\int p_{\theta}(x_i|z)p(z)dz)$. However, we don't have any information about the prior $p(z)$ and the generative model $p_{\theta}(x_i|z)$.

Estimating intractable posterior via stochastic variational inference (VAE framework) helps us bound the objective. We can sample from a estimated distribution $q_i(z)$, using importance sampling to get $logp(x_i) = log\mathbb{E}_{z\sim q_i(z)}[\frac{p(x_i|z)p(z)}{q_i(z)}] \geq \mathbb{E}_{z\sim q_i(z)}[log\frac{p(x_i|z)p(z)}{q_i(z)}]$, from Jensen's inequality and $p(z)$ is the prior, e.g., uniform distribution. Then we can use amortized inference to fit $q_i(z)$ with function estimation to get the whole process tractable.

Another simple way to do this is estimating generative models via an adversarial process, i.e., the framework of GAN [4], which consists of two main components: a generative model and a discriminative model. By trying to fool the discriminator by generating similar enough images as the original dataset, the generator improves its generative ability. Meanwhile, the discriminator updates its parameters to better classify whether the image is fake or not. Iteratively, the two components converge to equilibrium and achieve good generating ability.

3 Detailed Project Description

We will first read papers of GAN [4] and VAE [5] and get familiar with these two basic mechanisms, understanding the learning process and probabilistic formulas within them. Then we will conduct both quantitative and qualitative experiments. Quantitative experiments contain different metrics in the training process that are adopted widely in the community, e.g., the training loss curves and score metrics, and the scores (or quality) of the generated images. To better understand the whole generative processes, we will also conduct visualization analysis (including dimensionality reduction that are covered in class and t-SNE, manifold analysis etc. that are used in the machine learning community) as qualitative experiments. Ablation studies will also be conducted, e.g., the effects of hyper-parameter selection and data augmentation. We will mainly conduct experiments on the fundamental datasets to make more analysis, e.g., MNIST and CIFAR10. Based on the drawbacks of methods in GAN and VAE, as well as observations in our experiments, we will select and read the papers that improve them. This way, we will go deeper into this field and make more analysis, both theoretically and experimentally. We will analyze and summarize more relevant works that we encounter, giving comparisons, e.g., InfoGAN [3] Wasserstein GAN [1], IWAE [2]. Specifically, we will reformula the training process and architectures (formulas) with the same set of notations, comparing them from the machine learning perspective. The mathematical reason why improvement is obtained will also be analyzed. If we have extra time, we'll reimplement the methods of these related works, comparing the experimental results to give a better understanding.

3.1 List of tasks/collaboration plan

1. – Task: Understand papers: GAN [4] and VAE [5]
   – Leader(s): Tingyu Zhang
   – Deadline: 15-Oct-2020
   – Importance: It is critical since they contain the fundamental theory and learning processes that are needed when conducting experiments and summarizing more related papers.
   – Potential challenges: If the original papers are hard to fully understand, we will find slides or notes about these topics from other graduate courses.

2. – Task: Conduct experiments of GAN [4] and VAE [5]
   – Leader(s): Xueren Ge

– Deadline: 15-Nov-2020
– Importance: It is critical since we need to reimplement them and give analysis. This will help us understand the papers and find the drawbacks of the methods.
– Potential challenges: For GAN, the unsynchronized training of generative and discriminative models will lead to 'the Helvetica scenario' where the generative model will fail to model the data. Hence, synchronized training must be ensured. For VAE, the stochastic gradient descend method may not minimize the error function as well as other methods. Other minimization methods will be conducted and their results will be compared.

3. – Task: Read more related papers that improve GAN and VAE
   – Leader(s): Shenao Zhang
   – Deadline: 5-Nov-2020
   – Importance: It is critical to make comprehensive analysis.
   – Potential challenges: Modification for GAN and VAE requires a deep understanding of optimization theory. Instead of improving the optimization method, our group can focus on other areas, such as determining the optimal sample size for the best performance.

4. – Task: Conduct experiments of the above related papers.
   – Leader(s): Haidong Zhi
   – Deadline: 20-Nov-2020
   – Importance: It is optional, but will help us understand these approaches deeper.
   – Potential challenges: A balance between precision and cost has to be determined since Some improvements may increase the performances at the expense of computational cost.

References

[1] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein gan. *arXiv preprint arXiv:1701.07875*, 2017.

[2] Y. Burda, R. Grosse, and R. Salakhutdinov. Importance weighted autoencoders. *arXiv preprint arXiv:1509.00519*, 2015.

[3] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, and P. Abbeel. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In *Advances in neural information processing systems*, pages 2172–2180, 2016.

[4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[5] D. P. Kingma and M. Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.