

1

1.a

1.b

Question: What happens if you use different learning rates for the discriminator and generator, or if you train the generator multiple times for every iteration of the discriminator? Describe the best training scheme, and any problems you encountered during training.

Answer: _____

1.c

1.d

1.e

Question: Run your Conditional GAN on Fashion MNIST. How do these generated images compare to your initial GAN, both in quality and in the relative abundance of each class?

Answer: providing additional information significantly improves the quality of the images. For the regular GAN most pictures look like blurry mixtures of shirts and trousers. For the conditional GAN, it is clearer which object is depicted. Furthermore, it is more balanced: We have more shoes and bags than with the regular GAN.

1.f

Question: Describe the theoretical differences between a generative adversarial network (GAN) and a variational autoencoder (VAE). What differences did you notice in practice? What tasks would be more suited for a VAE, and what tasks require a GAN?

Answer: GANs consist of two neural networks: the generator network and the discriminator network. The generator network generates fake samples, while the discriminator network tries to distinguish between the generated samples and real samples from the dataset. When we use a GAN to learn a distribution, we gain implicit access to the distribution: we can generate samples from the distribution but do not obtain direct access to the distribution density.

what did I notice

In contrast, VAEs are based on the encoder-decoder architecture, where the encoder maps input data to a lower-dimensional latent space, while the decoder maps points in the latent space back to the input space. VAEs aim to model the probability distribution of the data in the latent space in parametric form. Hence we obtain direct access to the probability density.

In practice, GANs are generally better at generating realistic-looking samples. VAEs are well-suited for tasks that involve modeling the probability distribution of data.

1.g

Question: Why do we often choose the input to a GAN (z) to be samples from a Gaussian? Can you think of any potential problems with this?

Answer: The Gaussian distribution is a simple and easy-to-sample distribution (even for high dimensions). The Gaussian distribution is also continuous and has an unbounded domain, making it a good choice for the input to the generator network, which typically produces continuous data. Furthermore, the Gaussian distribution forms an exponential family.

A limitation: the Gaussian distribution has a bell-shaped curve with a single peak. This can lead to mode collapse where the generator produces a limited set of similar samples, rather than diverse samples that cover the full range of the target distribution.

1.h

Question: In class we talked about using GANs for the problem of mapping from one domain to another (e.g. faces with black hair to faces with blond hair). A simple model for this would learn two generators: one that takes the first domain as input and produces output in the second domain as judged by a discriminator, and vice versa for the other domain. What are some of the reasons the DiscoGAN/CycleGAN perform better at this task than the simpler model?

Answer: DiscoGAN artificially doubles the number of generators. In the example from above: it generates from blond hair to black hair back to blond hair, and generates from black hair to blond hair to black hair. This indirectly increases the number of training samples and hence generalizability. The two generators "black hair back to blond hair" share weights and the two generators "blond hair back to black hair" share weights. This reduces number of weights to train and facilitates optimisation.