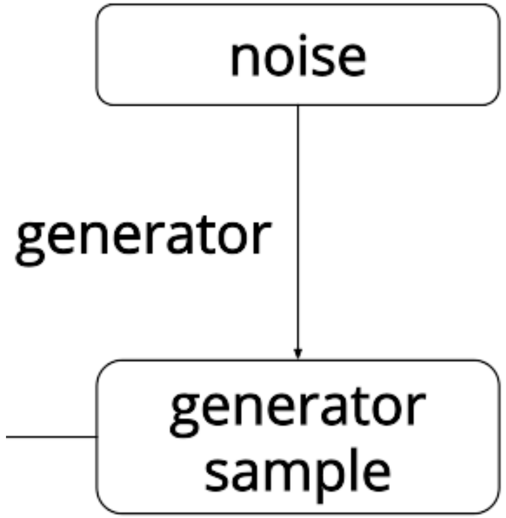
Generative Adversarial Networks

Forest Agostinelli
University of South Carolina

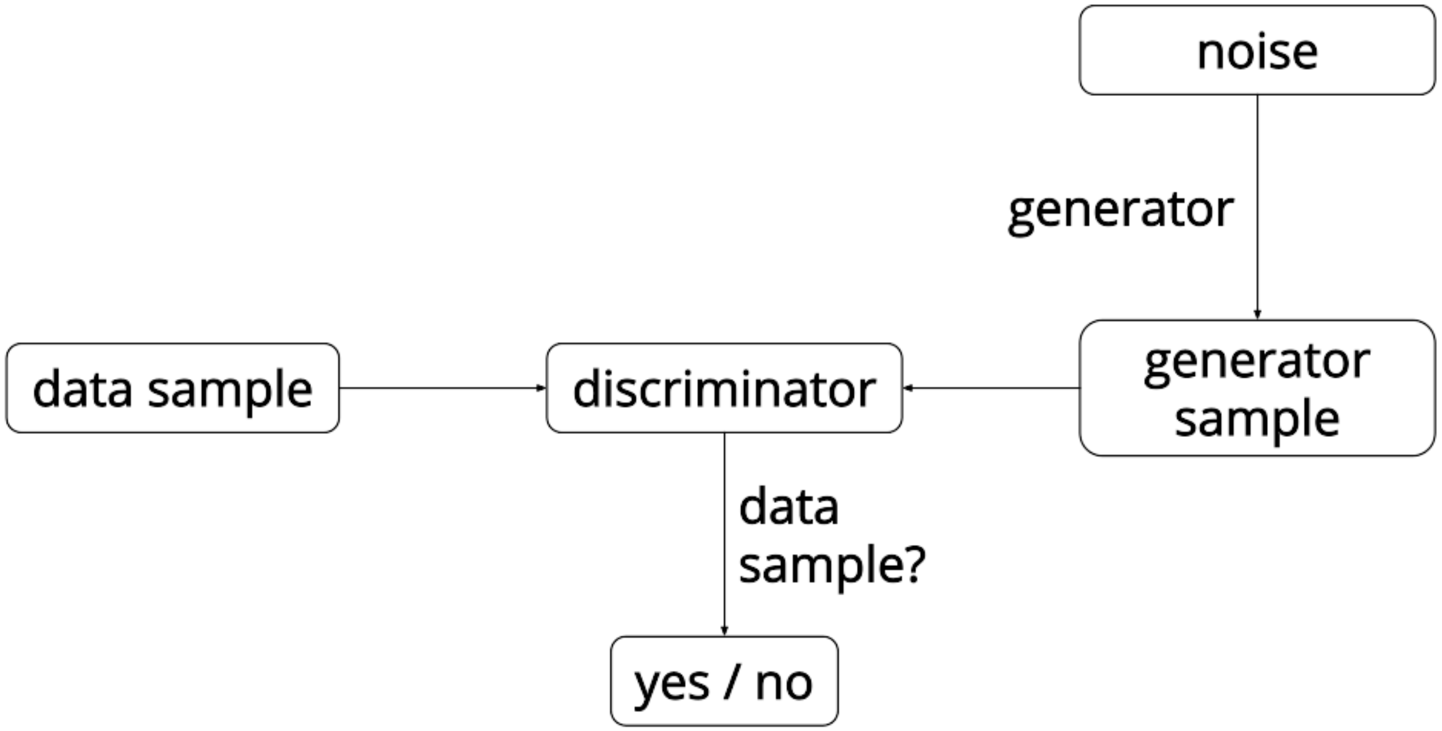
Generating Data from Random Noise

- Variational autoencoders train a network to encode data to a latent space that follows a given distribution and then decode this latent space to match the data
- Instead, we could train a network to map inputs from a given distribution to a distribution that maps that of the data
- We could then generate new data by sampling from our given distribution and obtaining the output of our network
- How can we then enforce that the distribution of that output match the data's distribution?
  - Remember, the distribution of the data may be difficult to formally characterize, especially in high dimensional spaces



Generative Adversarial Networks (GANs)

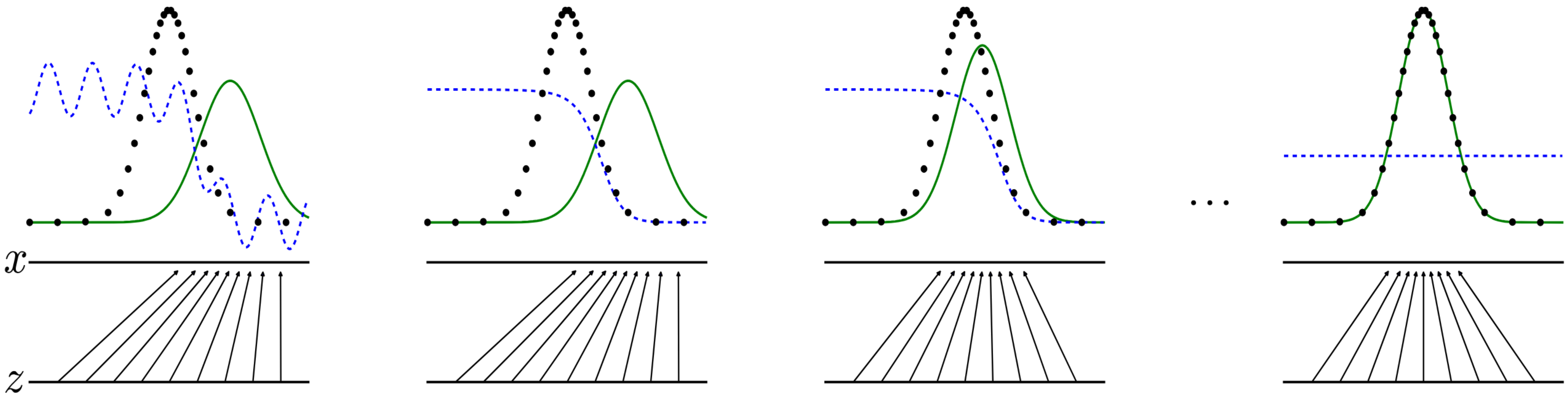
- We can train a discriminator to differentiate between generated data and data from the given dataset
- Will this result in the generator generating realistic data?



Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. In *NIPS* (pp. 2672-2680).
http://blog.aylien.com/introduction-generative-adversarial-networks-code-tensorflow/

Convergence of GANs

- The distribution of the real data is shown in black
- The distribution the discriminator learns is in blue
- The distribution of the generator is shown in green
- Ideally, the generator converges to the real distribution and the discriminator has no choice but to give an output of 0.5 all the time



(a) (b) (c) (d)

Improvement of GANs

- GANs can be difficult to train, in practice, due to their adversarial nature
    - However, many improvements have been made to stabilize training

2014 2015 2016 2017

Goodfellow, Ian, et al. "Generative adversarial networks." Communications of the ACM 63.11 (2020): 139-144.

Generative Adversarial Networks

- Unpaired domain transfer

Monet → photo

photo → Monet

zebra → horse

horse → zebra

J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle consistent adversarial networks. arXiv preprint arXiv:1703.10593, 2017.