

Conditional Generative Adversarial Network for Music

EPFL

Alex Rusnak

What is a GAN?

- Simply: it is a way to learn the characteristics of a dataset, and then generate fake samples that mimic these characteristics.
- Complicated: Stay tuned!

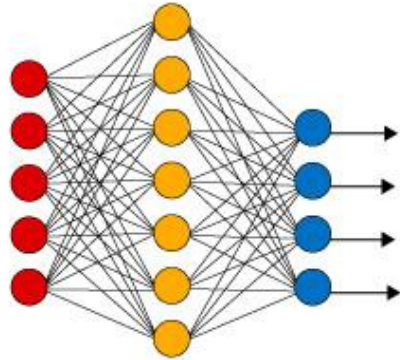
Why should I care?

- Synthetic data is useful in many contexts in Data Science / STEM (and music is cool!)
 - Creating more samples of challenging outliers in a dataset
 - Changing distributions of datasets for ethical or functional reasons
 - Useful anonymous data

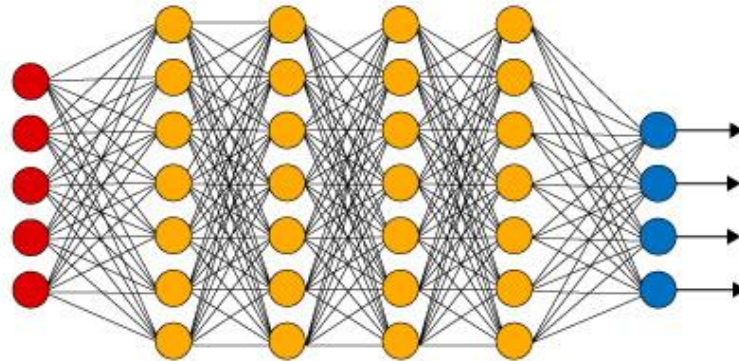
Building Blocks

- GANs are an architecture of neural network
 - Utilizing convolutional layers for up / downscaling

Simple Neural Network



Deep Learning Neural Network

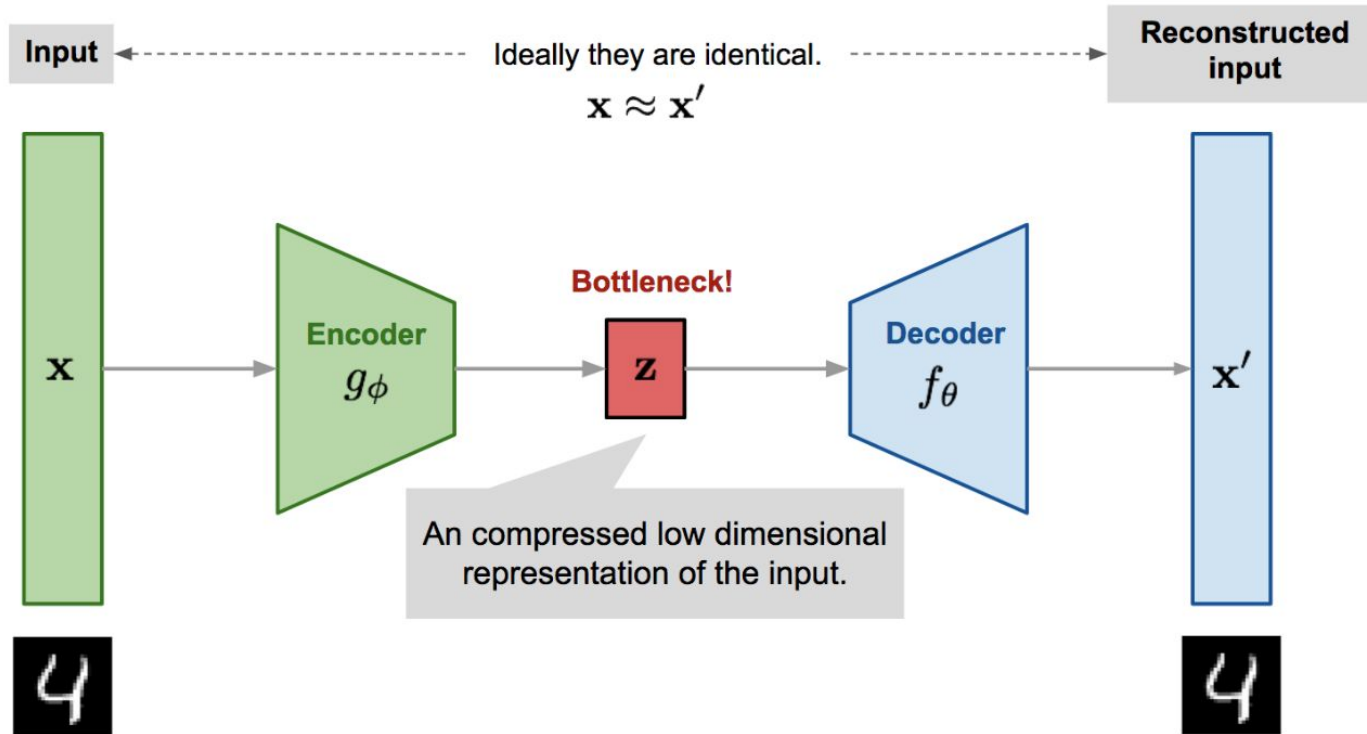


● Input Layer

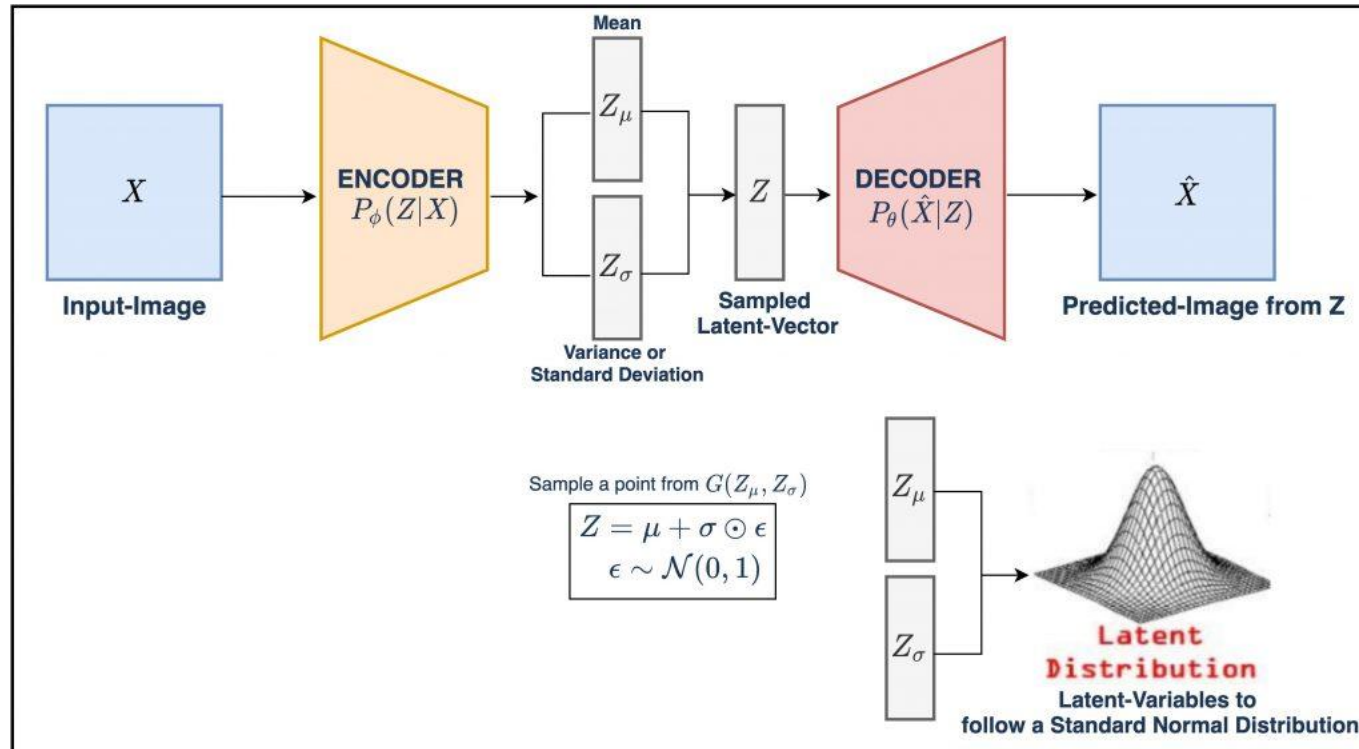
● Hidden Layer

● Output Layer

Autoencoder

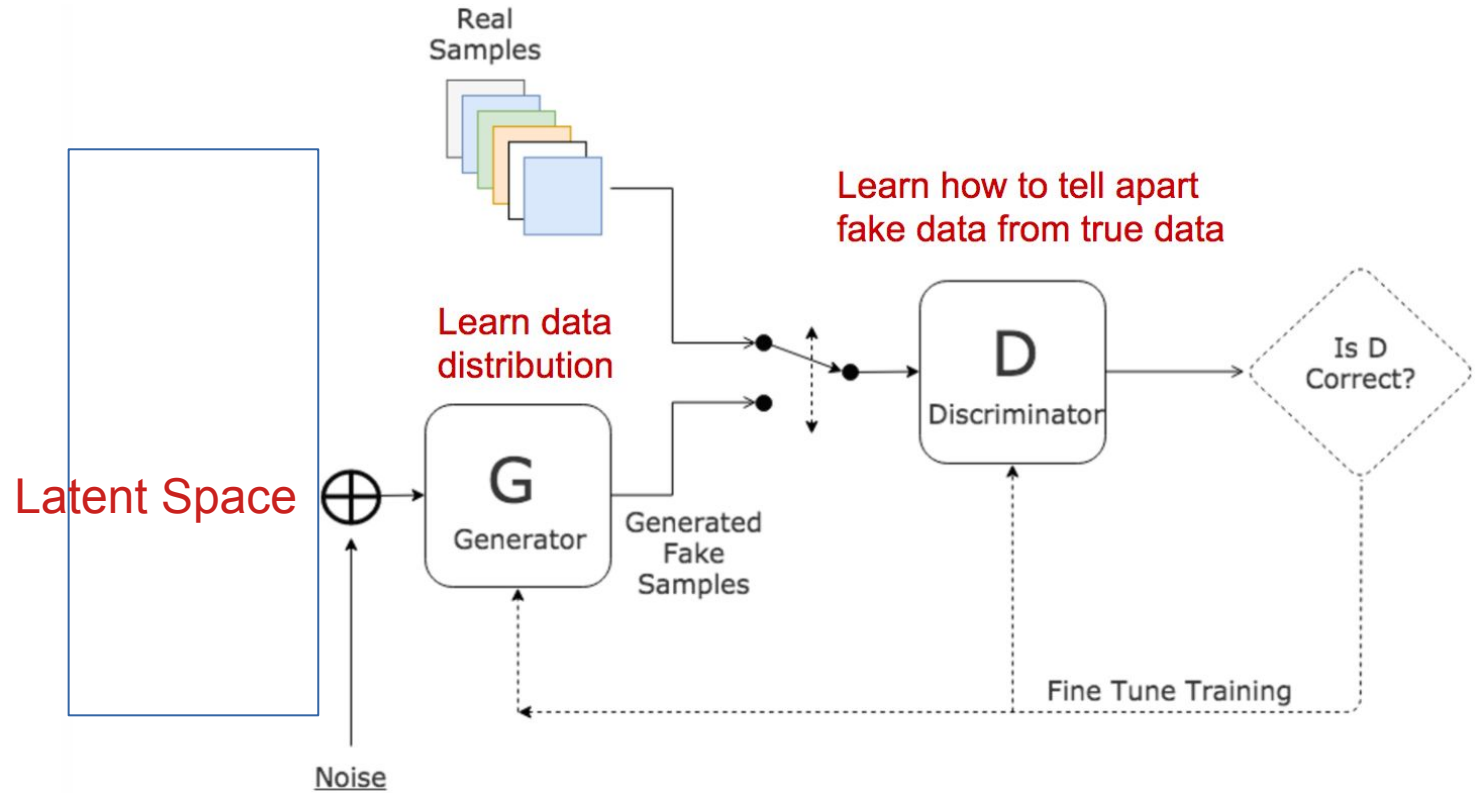


Variational Autoencoder



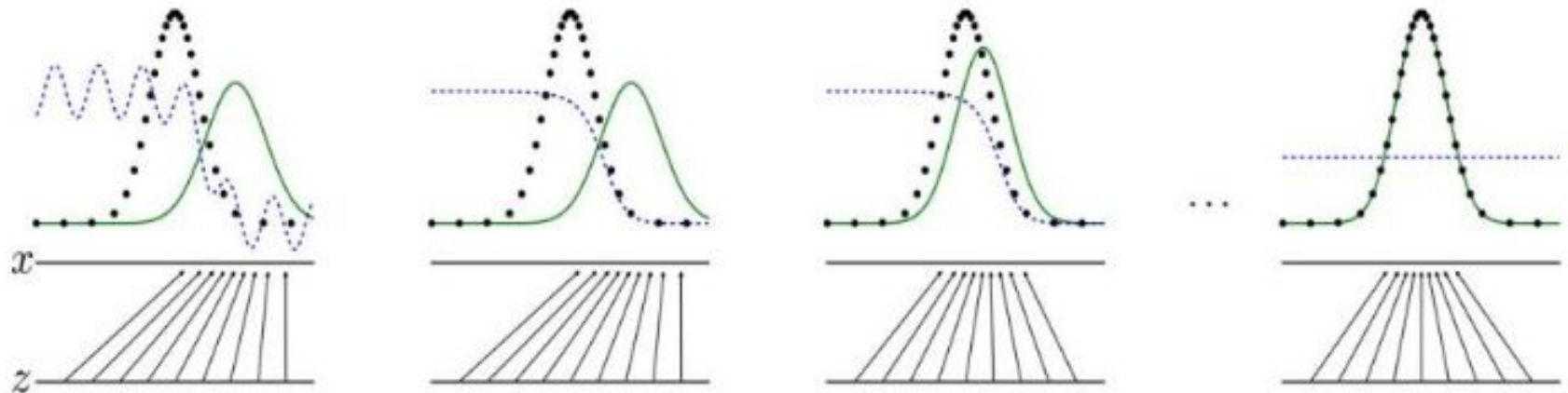
GAN

Simple generator is basically just the distribution and the decoder of the VAE



Training

- Black dotted line is the real data distribution.
- Green line is the generated distribution.
- Blue dotted line is the discriminator success relative to that part of the distribution.
- Z is the latent space, mapping to x (the generated sample)



Conditional?

- Generator receives prior series along with Z as input
- Discriminator receives either real prior and real series or real prior and fake series
- Allows sequential (autoregressive generation)

R = Real, F = Fake, Z = Latent Noise, n = total sample size

$R[0 \dots n/2]$ $Z[\dim Z] \rightarrow \text{Generator} \rightarrow F[n/2 \dots n]$

$R[0 \dots n/2]$ $R[n/2 \dots n] \rightarrow \text{Discriminator} \rightarrow \text{Classification}$

$R[0 \dots n/2]$ $F[n/2 \dots n] \rightarrow \text{Discriminator} \rightarrow \text{Classification}$

Sequence:

$R[0 \dots n/2] \leftarrow F[n/2 \dots n] \leftarrow F[n/2 \dots n] \leftarrow F[n/2 \dots n] \dots \text{etc}$

Let's dig into some code!



When your laptop freezes
so you put it in the oven

