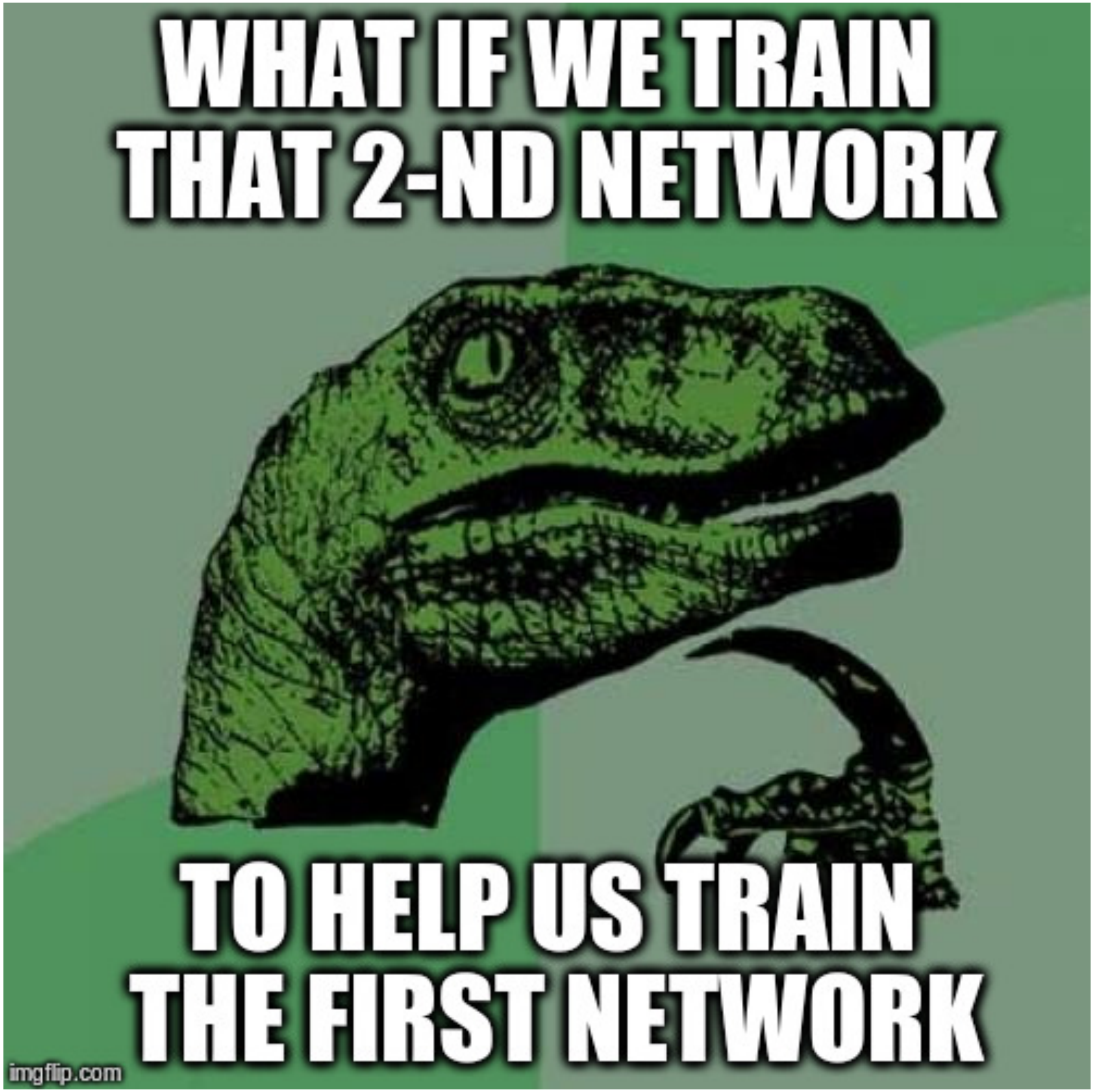
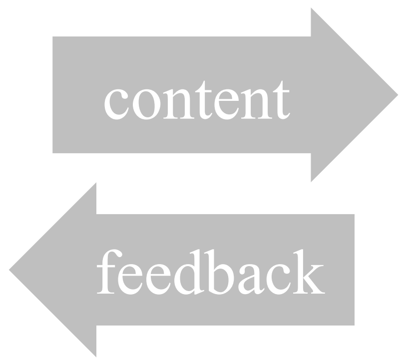
Lets train a model to tell us whether a generated image is good enough or not.



Generative Adversatial Networks

This is the high level overview.
GENERATOR VS
DISCRIMINATOR.

Generator



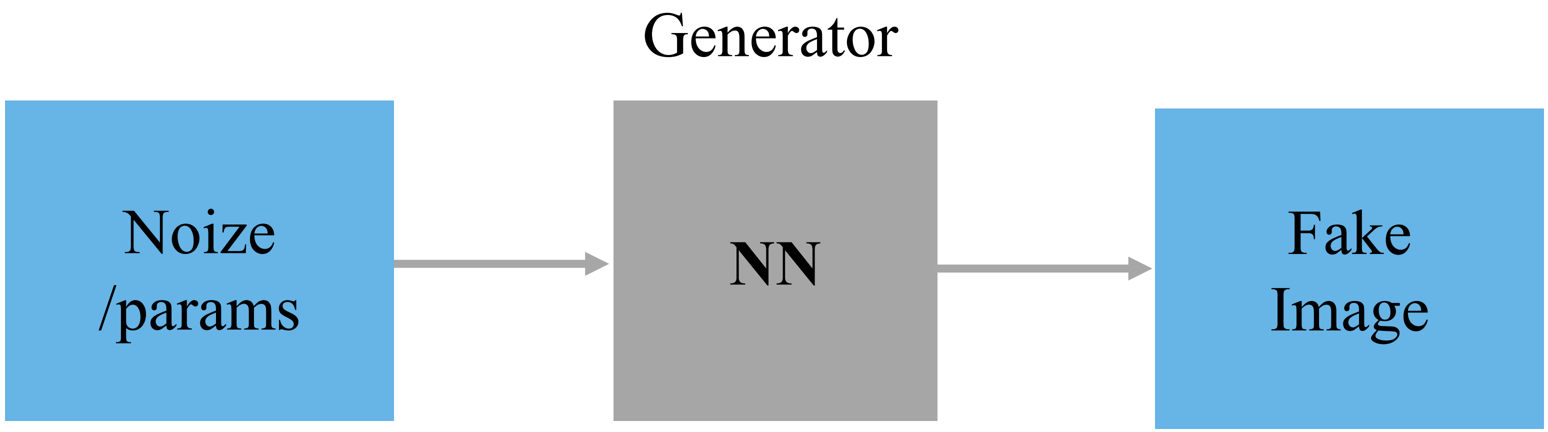
Discriminator

Generate image
(should be plausible)

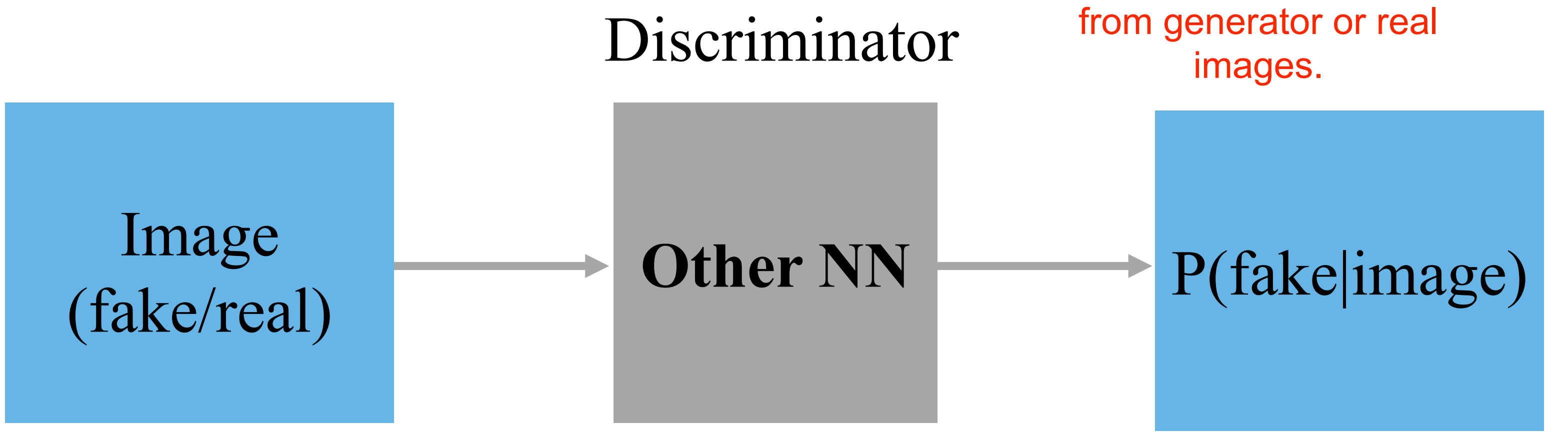
Tell if image is plausible
(image) → P(fake)

called discriminator
because it discriminates
between real images and
the generated images.

Generative Adversarial Networks



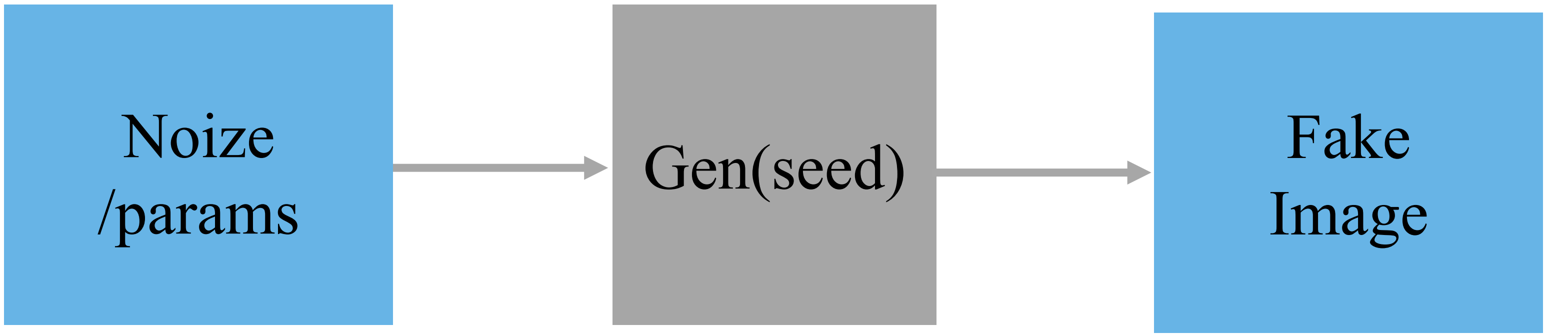
This one can train using fake images outputted from generator or real images.



The discriminator can train on positive examples from all real images, and negative examples using all generated fake images.

Generative Adversarial Networks

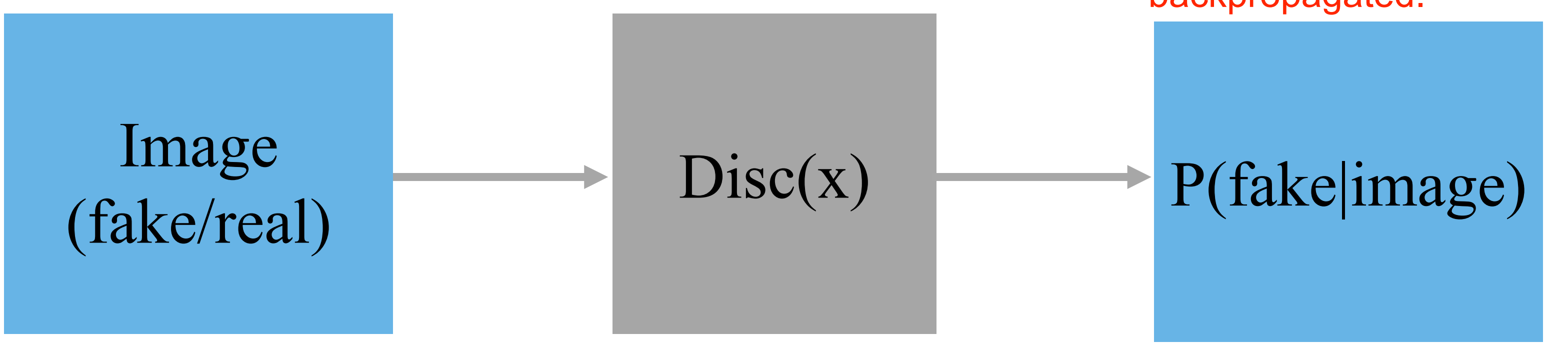
Generator



$$L_G = -log\left[1 - Disc\left(Gen(seed)\right)\right]$$

discriminator is built with differentiable operators which makes it able to be backpropagated.

Discriminator



$$L_D = -log[1 - Disc(realdata)] - logDisc\left(Gen(seed)\right)$$

Generative Adversarial Networks

Algorithm

- sample noise **z** and images **x**
- for k in 1…K
  - Train discriminator(**x**), discriminator(generator(**z**))

^these are the positive samples.

^these are the negative samples.

- For m in 1…M
  - Train generator(z)

Generative models are unstable. You have two models that 'hate each other'. And if one of them wins, you have to start the process all over again.

If discriminator wins (can train faster than generator), then the gradients vanish - the sigmoids that are used to compute P(real | image) are really close to 1 or 0. Thus, they have very small gradient.

If generator wins (constantly train faster than discriminator), then it can start learning the wrong things.
The generator can learn non-sensical stuff.

training steps:
(0). initialize generator and discriminator weights randomly.

(1). Train discriminator to clasify actual images against images generated by (untrained) generator

(2). Train generator to generate images that fool discriminator into believing they're real.

repeat (1) and then (2) again.

This cycle continues. This is why it's called ADVERSARIAL.
it is as if they are competing against each other.

Generative Adversarial Networks