
Analytics

The Delta Mission

Increase access and usefulness of data to **empower changemakers** working to better their communities

We accomplish this through our two core programs

Data Service Grant

Bridging the nonprofit skill gap by enabling nonprofits to **accelerate their impact** through leveraging **skilled data professionals** on project opportunities.

Machine Learning for Good

Empower anyone, anywhere to leverage data for good in their communities by **building technical capacity** around the world.

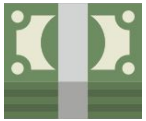
Since 2013...

Completed **~50 projects** with nonprofits and social impact organizations

Open-sourced and taught our machine learning curriculum to students globally

Recruited **200+ professionals** to work on projects

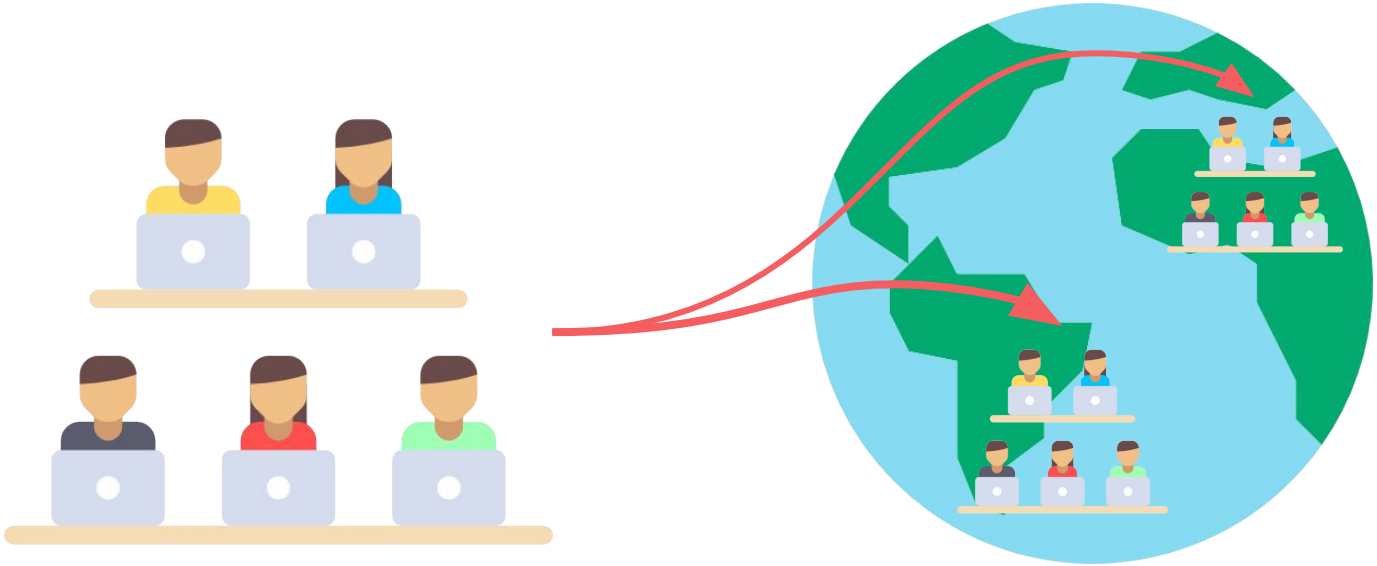
Donated over **20,000 hours**



$0 charged for services

Teaching Fellows

Delta has been working since 2013 to help non-profits around the world with their data. In 2017, we started the Teaching Fellows program.



Our key initiatives.

2017

Built our curriculum from scratch
Taught our pilot program in Nairobi, Kenya

2018

2019

2020

Our key initiatives.

2017

Built our curriculum from scratch
Taught our pilot program in Nairobi, Kenya

2018

Beta tested, revised, and open-sourced our curriculum
Taught subsequent classes in Morocco + the SF Bay Area

2019

2020

Our key initiatives.

2017

Built our curriculum from scratch
Taught our pilot program in Nairobi, Kenya

2018

Beta tested, revised, and open-sourced our curriculum
Taught subsequent classes in Morocco + the SF Bay Area

2019

Explored supplementing our curriculum with active learning techniques

2020

Now, we are going global.

Why Global Teaching Fellows?

Alicia Tsai Follow
Feb 26 · 5 min read

We just opened up our Global Teaching Fellows program. This is a big change from requiring that all our teaching fellows are local in the Bay Area. We wanted to explain why we made this change and what makes us excited about the 2020 teaching fellow program.

Delta Analytics is a 501(c)(3) nonprofit organization that was founded 6 years ago with a focus on helping nonprofits and social impact organizations leverage their data for good.

Three years ago, Delta Analytics started a Teaching Fellows Initiative. The goal of the program is to build technical capacity around the world. In 2017, we developed our pilot course, which focused on explaining the fundamentals of machine learning.

Our goal: support educators all around the world.

Meet our 2020 global teaching fellows!

Machine Learning Virtual Talks

@ Delta Analytics Teaching Fellows

Krish

Leke

Dina

Chima

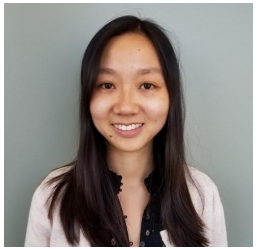
Aseda

Celebrate our teaching fellows as they wrap up the teaching fellow program with a guest lecture on the ML topic of their choice.

Talks will be streamed live on the Delta Analytics YouTube channel.

A few thank you:

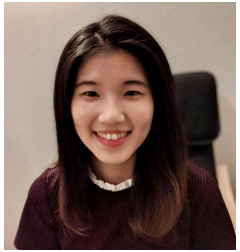
Sara Hooker
Co-Director



Amanda Su
Co-Director

Brian Spiering
Code Lead

Melissa Fabros
Teaching Lead



Alicia Tsai
Teaching Fellow

Raul Maldonado
Teaching Fellow

Allie Wang
Teaching Fellow

Analytics

Today's lecture ...

Introducing…

Leke Onilude

Data Science Nigeria

Generative Adversarial Networks

By Gbemileke Onilude

Outline

Motivation

GANs are a very recent breakthrough.

A generative adversarial network (GAN) is a class of machine learning frameworks designed by Ian Goodfellow and his colleagues in 2014.

Paper: https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

Image: https://en.wikipedia.org/wiki/Ian_Goodfellow

Fake or Real ?

1 2 3

Fake or Real ?

1

Fake

2

Fake

3

Fake

Image: https://rosebud.ai/

GANs aim to make real and fake indistinguishable.

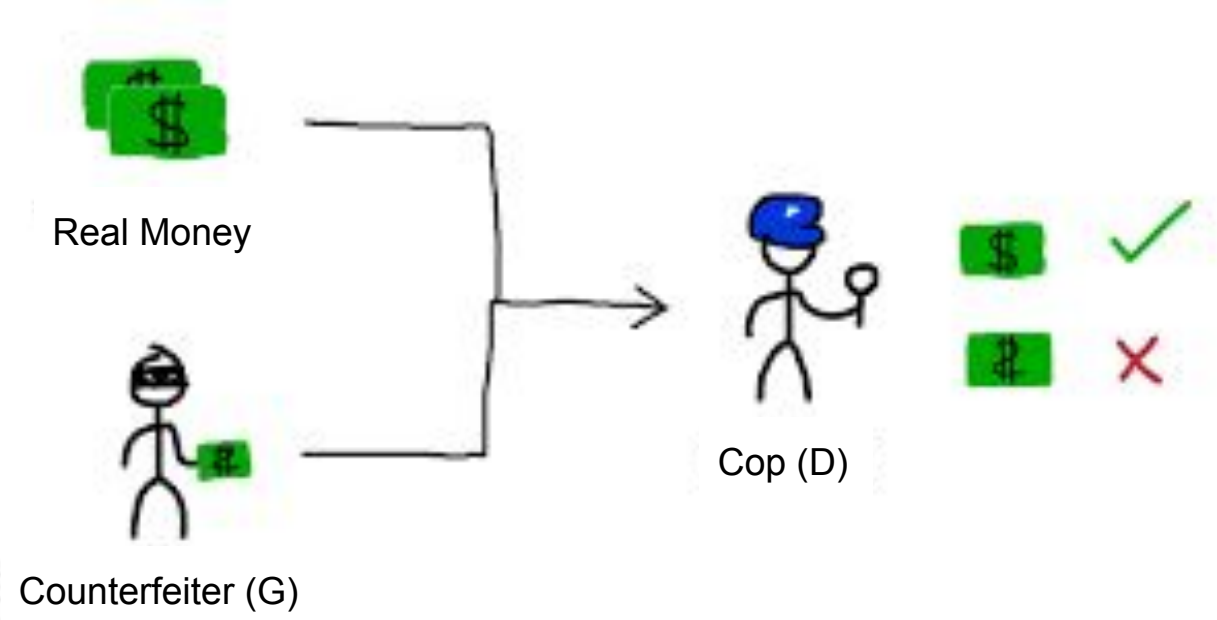


Image: https://www.cfml.se/blog/generative_adversarial_networks/

What is GAN

Generative Adversarial Network (GAN) is a setup of two neural networks, the two neural networks contest with each other in a game (in the sense of game theory, often in the form of a zero-sum game).

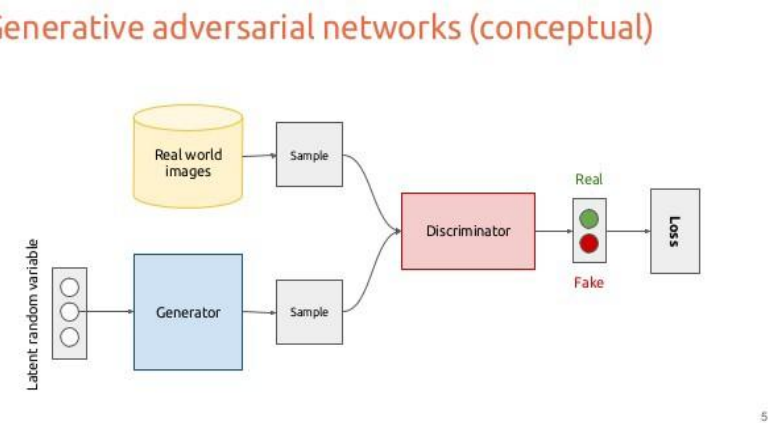
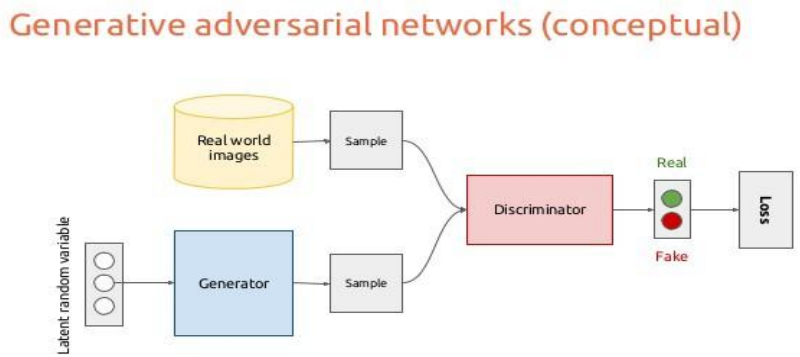


Image: https://medium.com/@saibharath897/generative-adversarial-networks-gans-560c5c988128

What is GAN

One of the neural networks create fake data while the other try to predict if the data is fake or real.

The goal where is by making the two model compete you can come up with a two good model, one which is good at creating fake data and other which is also good at distinguishing between fake and real data



Why Should I care about GAN

- GAN + Art
- GANs have plenty of real-world use cases like image generation, artwork generation, music generation, and video generation.
- GANs can enhance the quality of your images, stylize or colorize your images, generate faces and can perform many more interesting tasks.
- GANs can generate an Image based on a given text

Real World Application

Image generations

Real World Application

Bringing famous portrait back to life

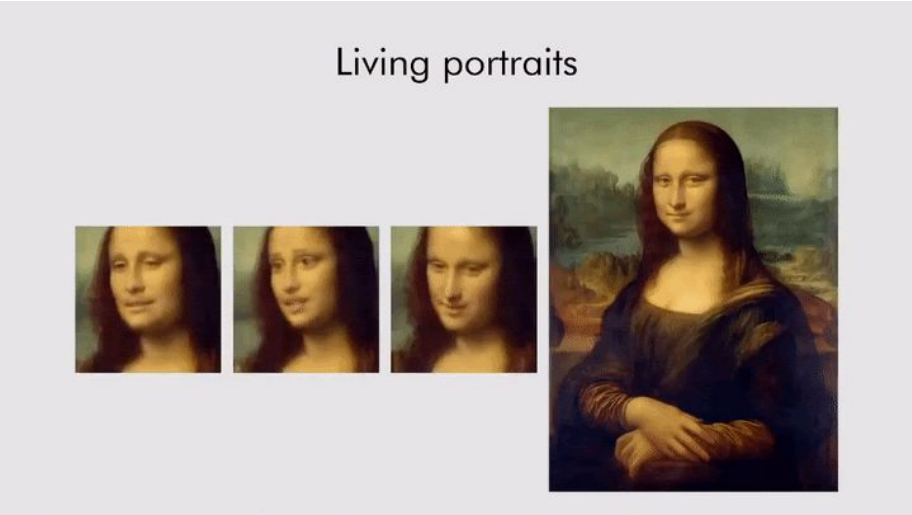


Image: https://www.bbc.com/news/technology-48395521

Creating Deep Fake Videos

Image: https://makeagif.com/gif/obama-peele-deep-fake-NUyaQL

Real World Application

Generating animation models

Image: https://towardsdatascience.com/animating-ganime-with-stylegan-part-1-4cf764578e

Increase Image resolution

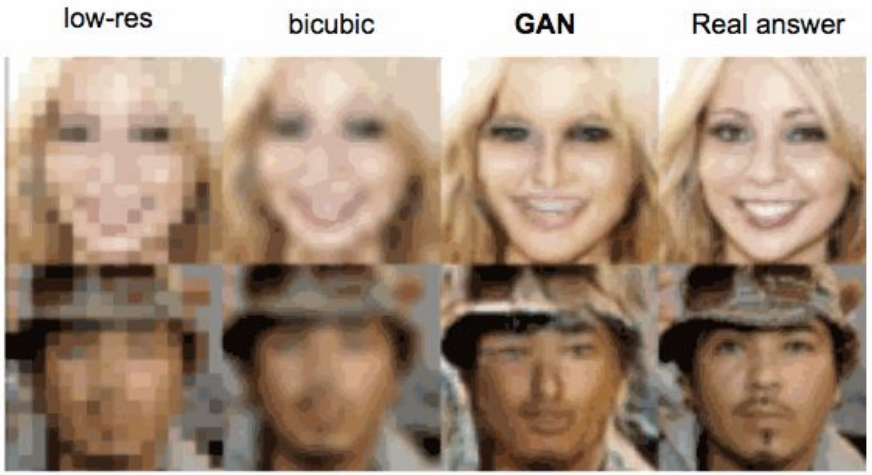
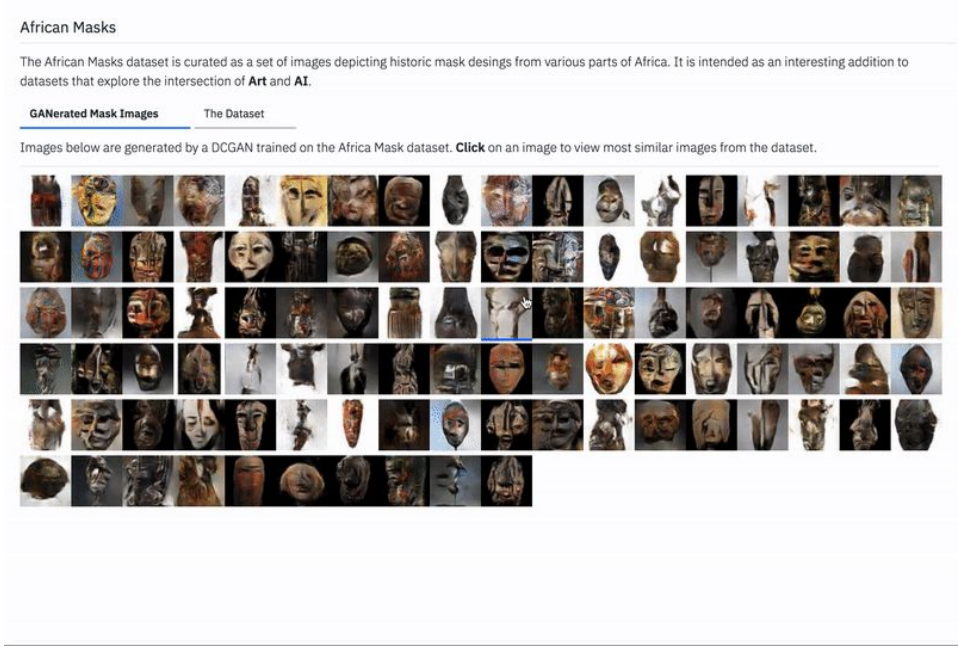


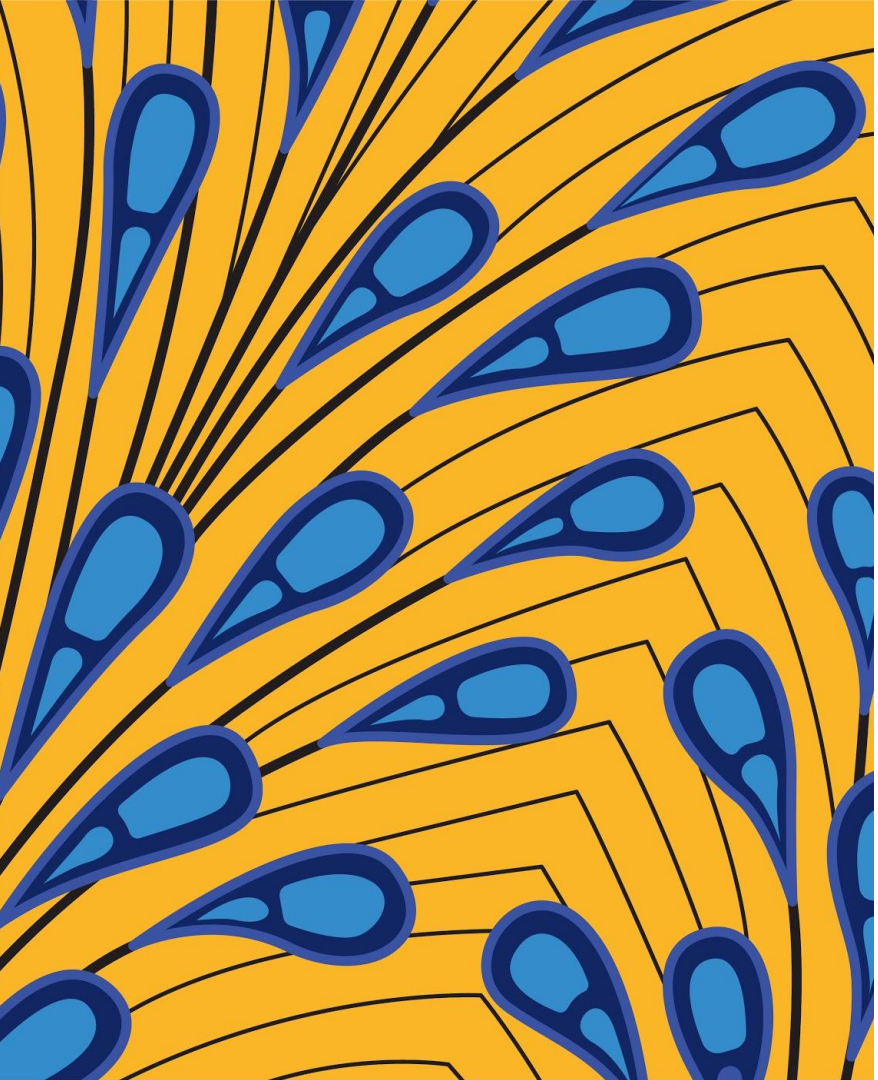
Image: https://www.bbc.com/news/technology-48395521

GANs used to generate African Masks

African Masks generation



Web demo:
https://victordibia.github.io/cocoafrica/#masks



Huge potential for GANs in an African setting.

1. Solving data scarcity
2. Generation of different cloth design for Africa dresses
3. Generation of new dance steps
4. Generation of new melody based on africa instructment

Fake or Real ?

1

2

3

Image: https://rosebud.ai/

How did we do?

1

Fake

2

Fake

3

Real

Check for Understanding

True or False

1. GAN consists of two neural networks

2. GAN machine learning framework was designed by Ian Goodfellow and his colleagues in 1970.

How did we do?

1. GAN consists of two neural networks

True

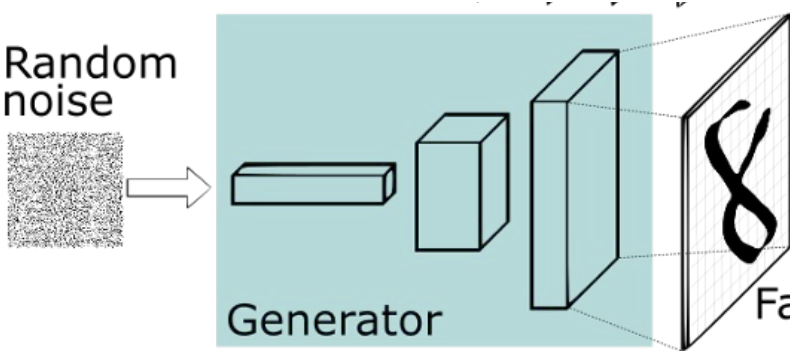
2. GAN machine learning framework was designed by Ian Goodfellow and his colleagues in 2015.

False

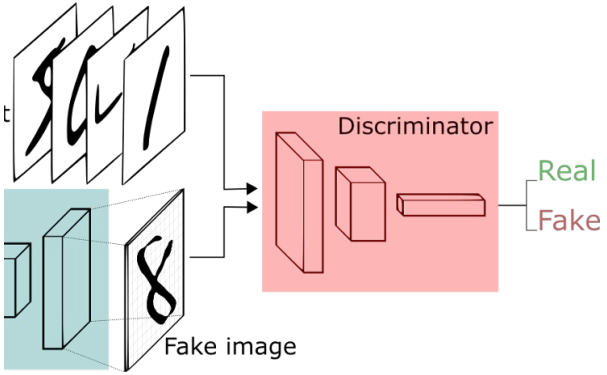
Basic Knowledge

Basic Architecture

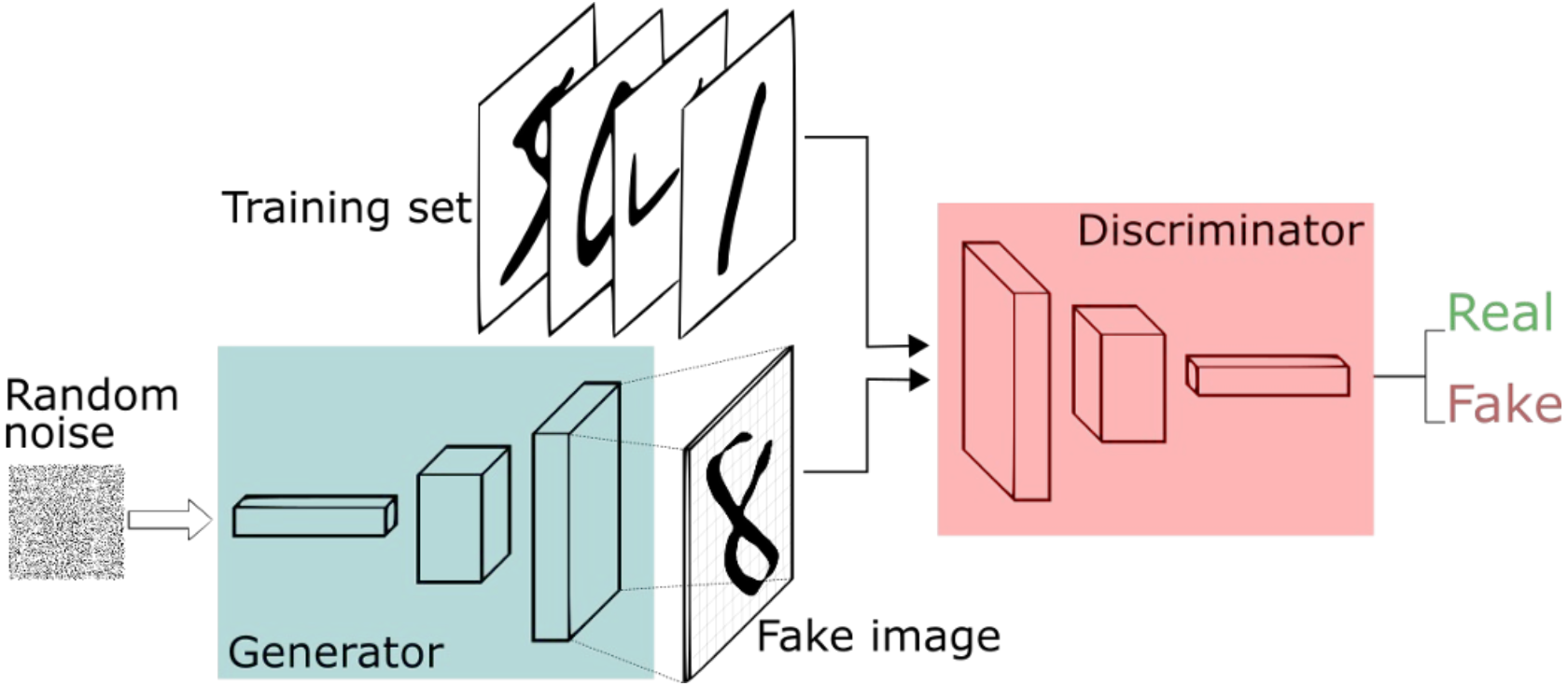
Generator: learns to generate plausible data. The generated instances become negative training examples for the discriminator.



Discriminator: learns to distinguish the generator's fake data from real data. The discriminator penalizes the generator for producing implausible results.



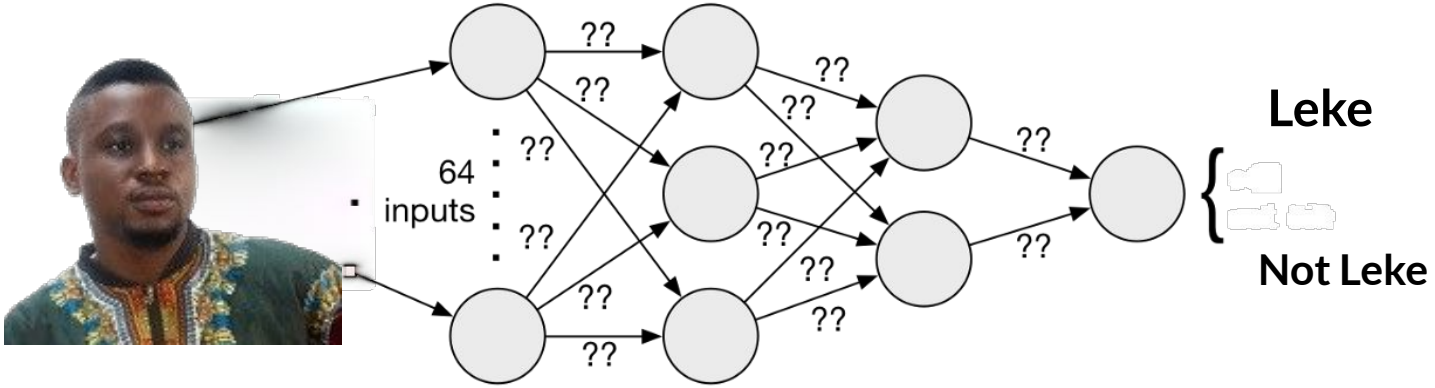
The discriminator and generator are jointly trained.



Let build a GAN model.

1. That can successfully generate images that look like Leke. Here, the discriminator that can distinguish between a generated image of Leke and a real image of Leke.

Discriminator is just a classifier.



Discriminator is tasked with distinguishing between the fake data generated by a Generator and real images.

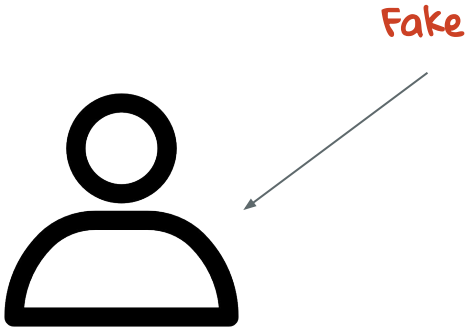
Example task:

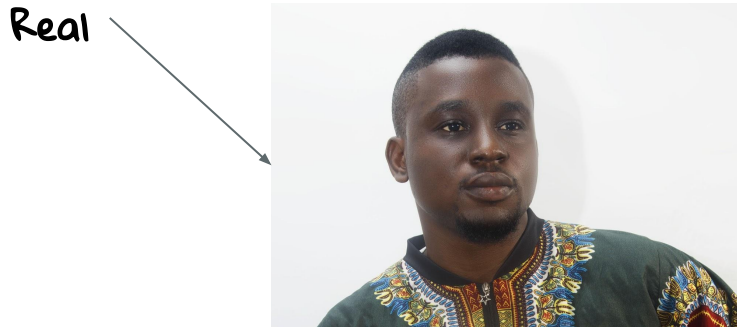
Is this an image a picture of Leke or not?

Discriminator is trained on both real and fake data.

Real data: Dataset from real event, people e.t.c. The Discriminator uses these instances as positive examples during training.

Fake data: Dataset created by the generator. The Discriminator uses these instances as negative examples during training.

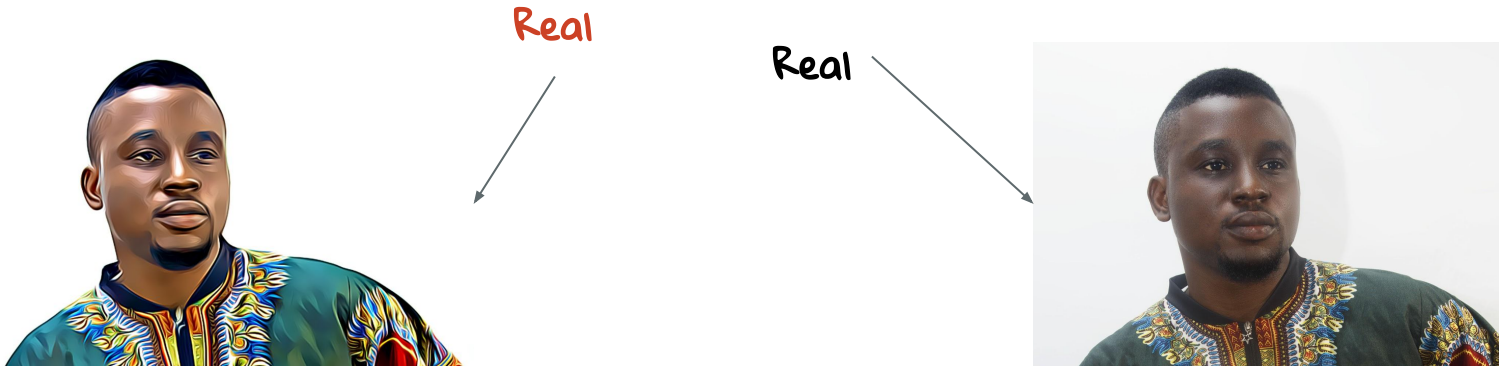




Generator attempts to fool the Discriminator

Generator learns to create fake dataset based upon the feedback from the Discriminator.

Generator aims to create a fake data point that would fool the Discriminator.

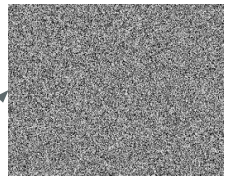


Input to GANs is random gaussian noise.

- Neural networks need some form of input
- Random gaussian noise
- Experiments suggest that the distribution of the noise doesn't matter much

Noise sampled is usually of smaller dimension than the dimensionality of the output space

Noise 32 X 32

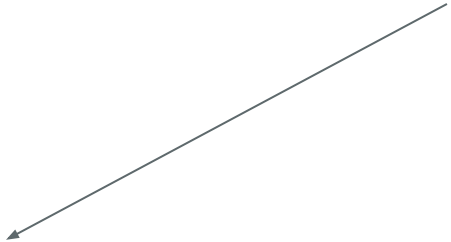


Output 120 X 120

How to train a GAN?

The GAN architecture must be able to jointly train the two model effectively. It need to be able to achieve convergence.

Loss Function



$$\min_G \max_D \mathbb{E}_{x \sim q_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z \sim p(z)}[\log(1 - D(G(z)))],$$

Simplify Loss Function

$$\log D(x) + \log\big(1 - D(G(z))\big)$$

- D(x) is the discriminator estimate of the probability that real data instance x is real.
- G(z) is the generator's output when given noise z.
- D(G(z)) is the discriminator estimate of the probability that a fake instance is real.

Simplify Loss Function

$$\min_{G} \max_{D} \log D(x) + \log\left(1 - D(G(z))\right)$$

- The values of D(x) can be between [0, 1].

  0 = Fake

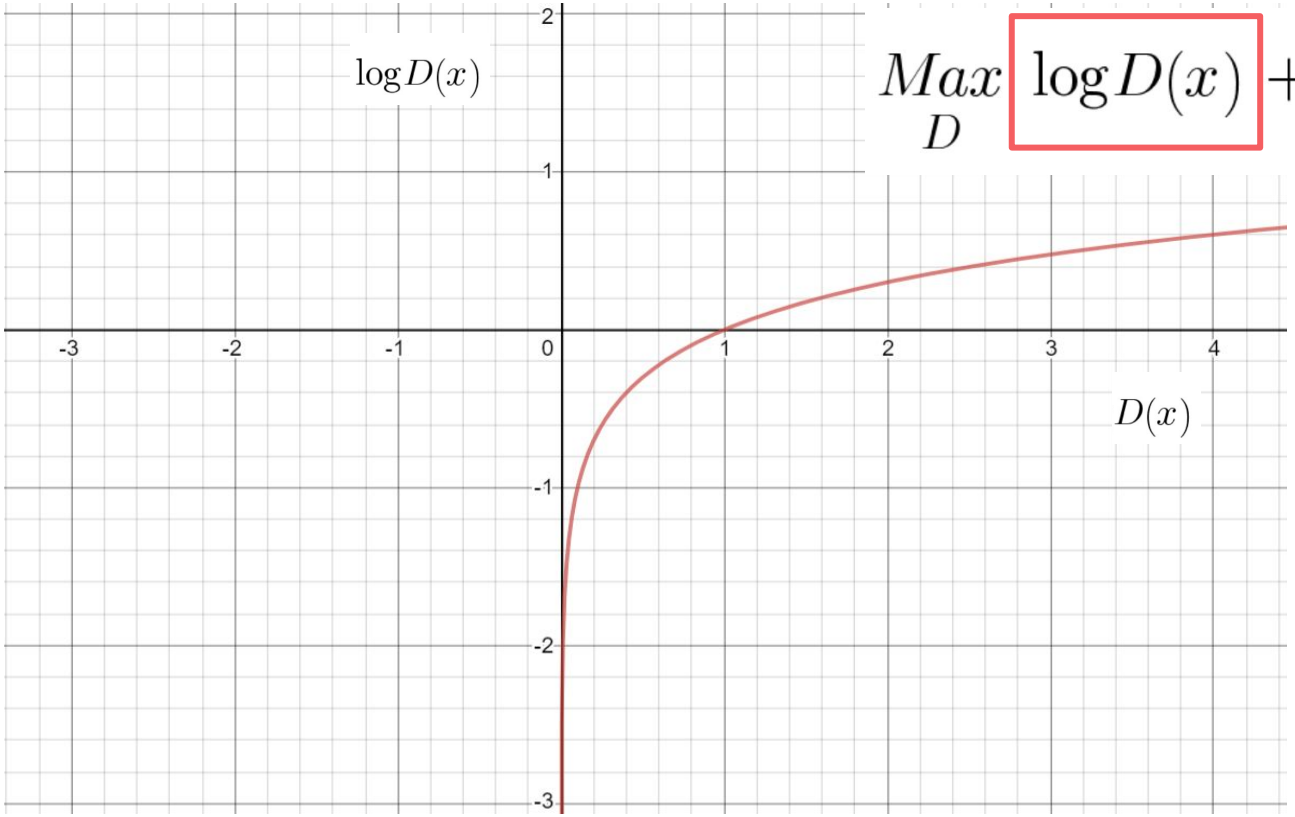
  1 = Real

- The values of D(G(z)) can be between [0, 1].

  0 = The Generator was unable to fool the discriminator

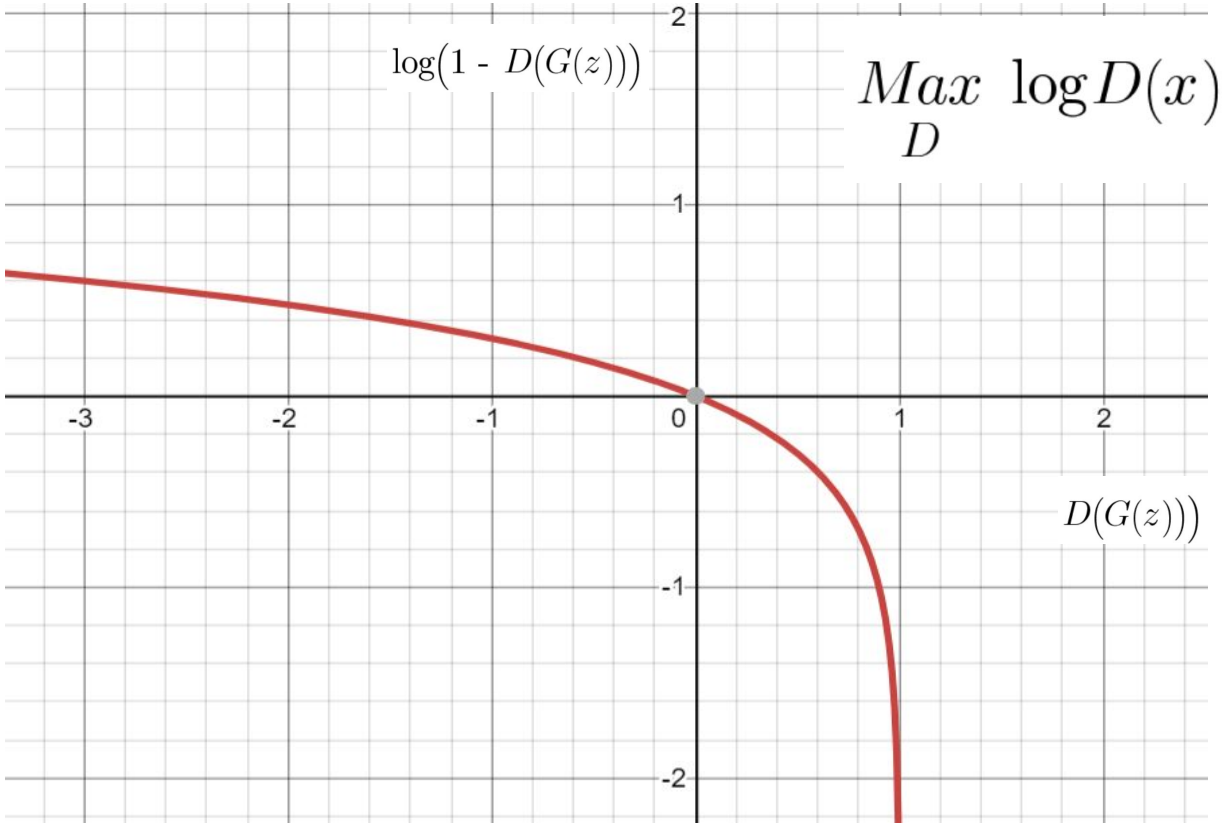
  1 = The Generator successfully fool the discriminator

Max Loss Function (Discriminator)



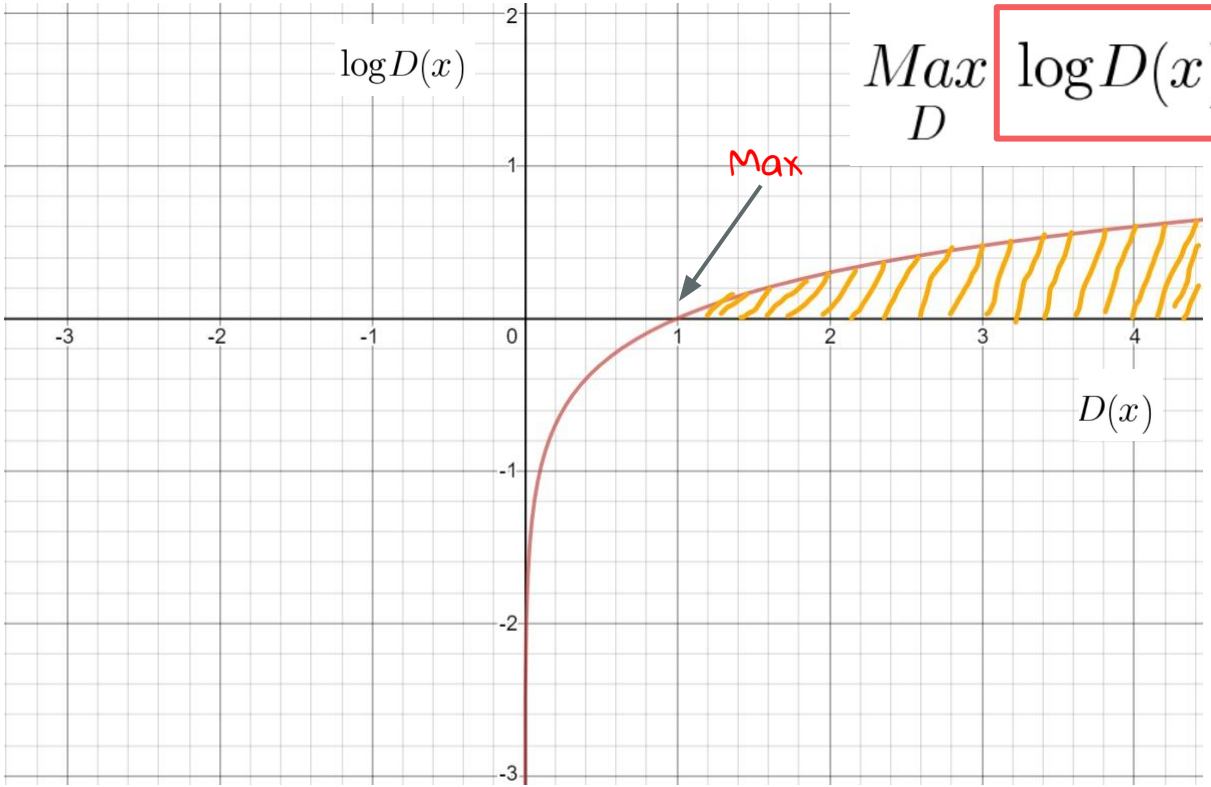
$$\underset{D}{Max} \boxed{\log D(x)} + \log\left(1 - D(G(z))\right)$$

Max Loss Function (Discriminator)



$$\max_{D} \log D(x) + \log\big(1 - D(G(z))\big)$$

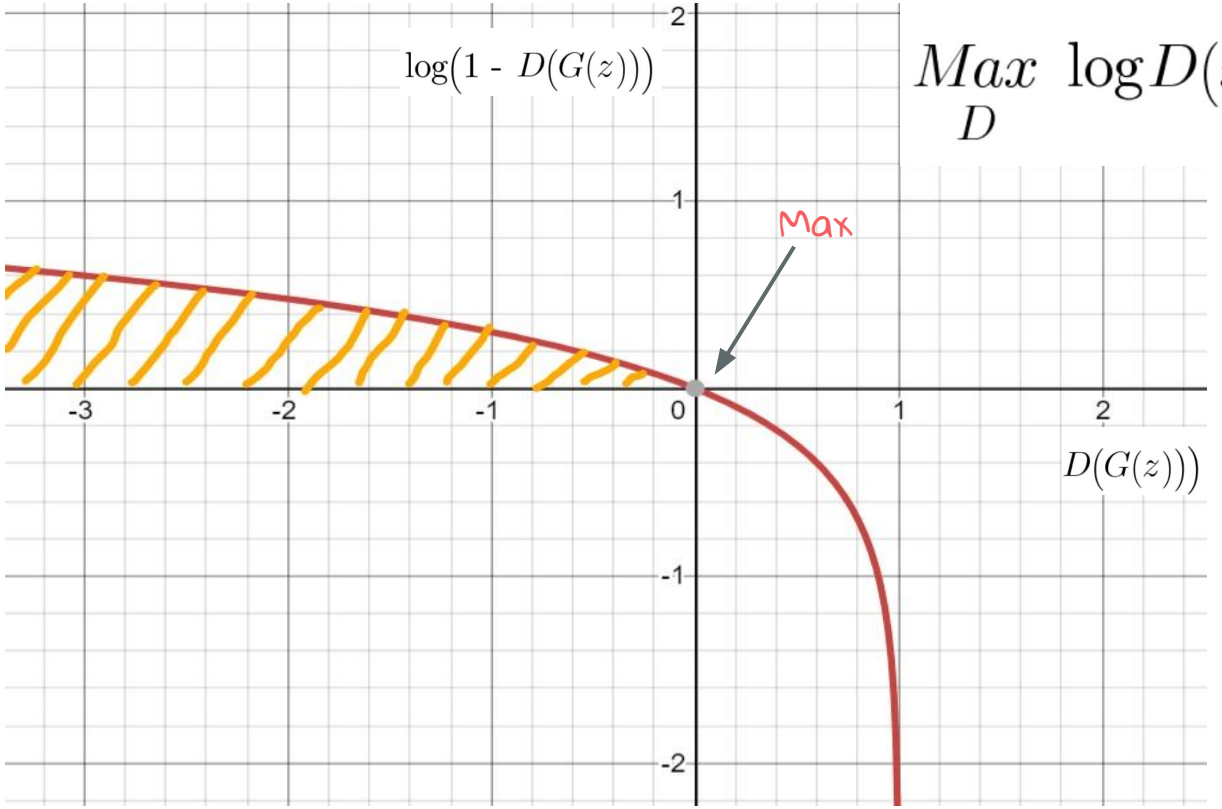
Max Loss Function (Discriminator)



$$\underset{D}{Max} \boxed{\log D(x)} + \log\left(1 - D(G(z))\right)$$

- Max D(x) = 1

Max Loss Function (Discriminator)



$$\underset{D}{Max}\ \log D(x) + \log\left(1 - D(G(z))\right)$$

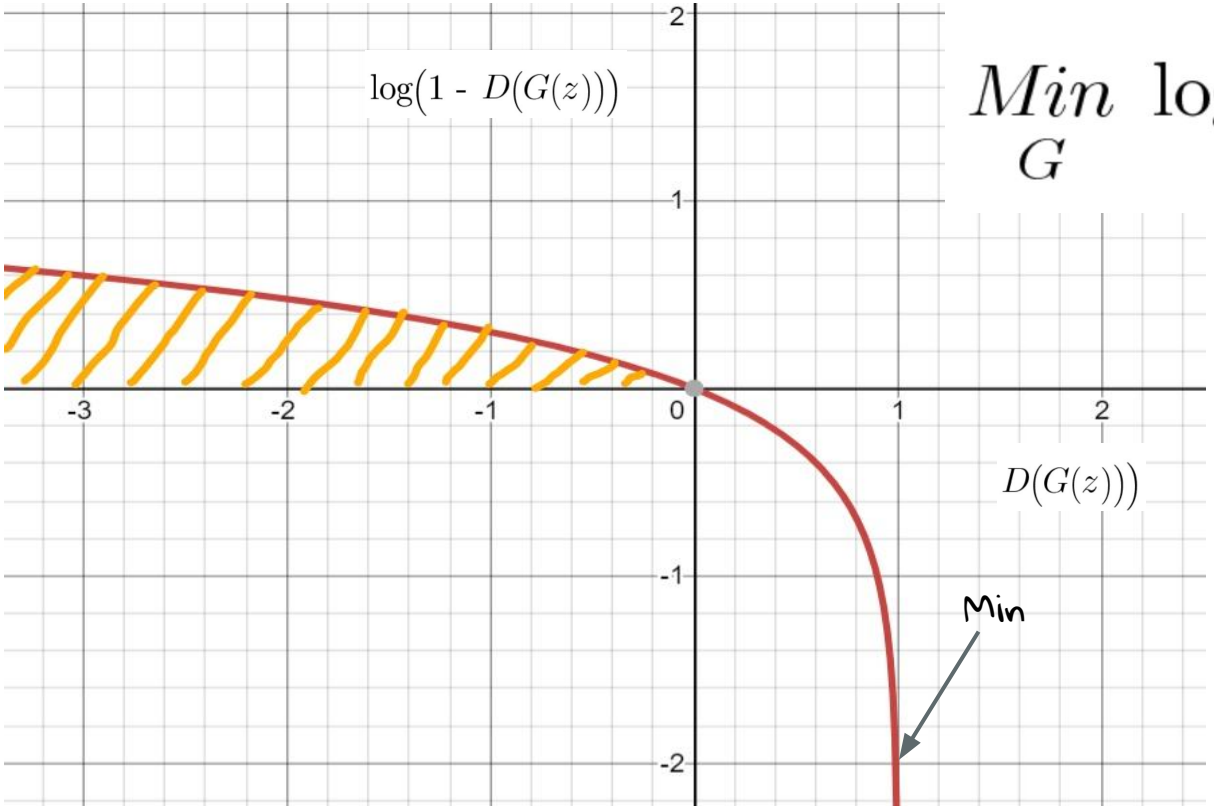
- Max D(G(z))= 0

Min Loss Function (Generator)

$$\min_G \log D(x) + \log\big(1 - D(G(z))\big)$$

$$\min_G \log\big(1 - D(G(z))\big)$$

Min Loss Function (Generator)



$$\min_G \log(1 - D(G(z)))$$

- Min D(G(z))= 1

Putting it all together

$$\min_G \max_D \mathbb{E}_{x \sim q_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim p(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))],$$

Check for Understanding

True or False

1. Discriminator creates fake data.

2. The input of a generator is random noise.

3. Negative examples during training are real examples.

How did we do?

1. Discriminator creates fake data.

   False

2. The input of a generator is random noise.

   True

3. Negative examples during training are real examples.

   False

State Of The Art Research Work

Original GAN Model

https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

KeyPoint

- 32 X 32 Images where used
- Generator and Discriminator in (a) and (b) were MLPs (not convolutional)

Conditional GAN

https://arxiv.org/pdf/1411.1784.pdf

KeyPoint

- A conditional element is inputted with the data to specify the kind of data to be generated

  - a category ID ("cat", "dog", e.t.c)

  - an input image from another domain

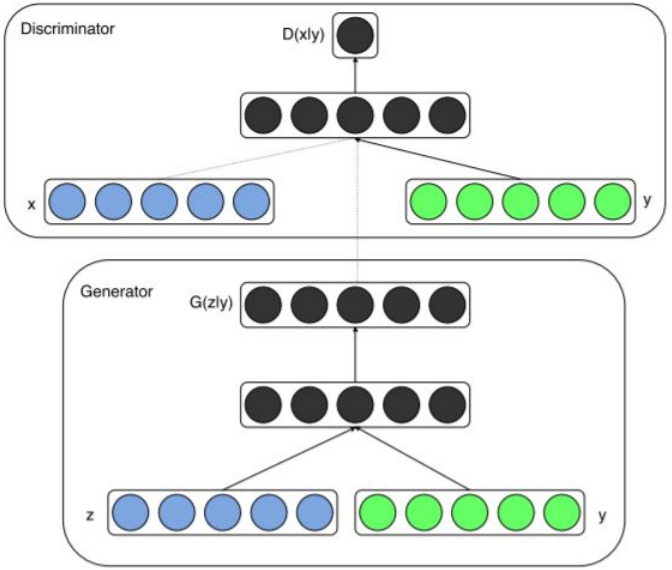


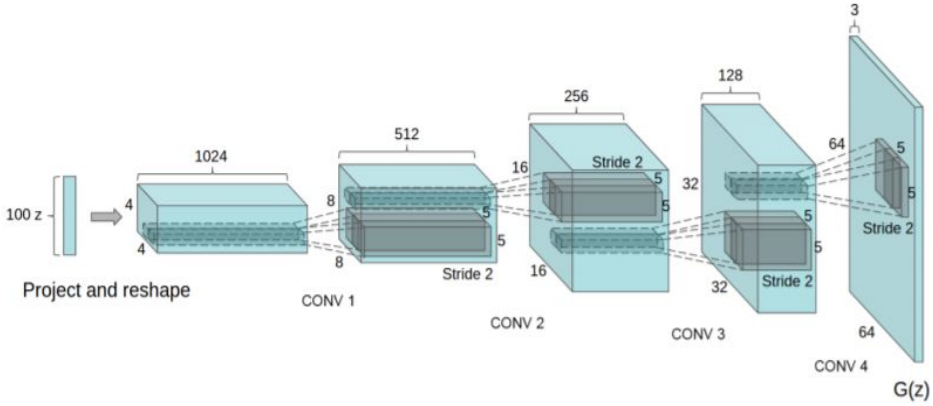
Figure 1: Conditional adversarial net

DCGAN

https://arxiv.org/pdf/1511.06434.pdf

KeyPoint

- Convolutional Neural Network for G and D
- Made use of batch normalization to stabilise the learning process



BigGAN

https://arxiv.org/pdf/1809.11096.pdf

KeyPoint

- Studied the reliability of GAN when trained on large dataset
- Changing the scale of noise sample during training

1 INTRODUCTION

Figure 1: Class-conditional samples generated by our model.

Style GAN

https://arxiv.org/pdf/1812.04948.pdf

KeyPoint

- Style GAN can generate photo realistic face images
- Redesign the generator architecture in a way that add spatial pixel noise at each layer

Thank you. Questions?

Thank you.

A big thank you to Sara, you are the best.

Sara

Thank you to everyone who provided feedback along the way.

Sasha Sheng

Facebook

Rosanne Liu

Machine learning Collective

Roberto Sanchis

Spotify

Go Further

1. How to build a GAN with Python
    a. https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/tutorials/generative/dcgan.ipynb
    b. https://machinelearningmastery.com/how-to-develop-a-generative-adversarial-network-for-an-mnist-handwritten-digits-from-scratch-in-keras/

2. Online lecture by researchers
    a. https://www.youtube.com/watch?v=wFsI2WqUfdA&list=PLqYmG7hTraZCDxZ44o4p3N5Anz3lLRVZF&index=9
    b. https://www.youtube.com/watch?v=9JpdAg6uMXs

More questions?

Email: lekeonilude@gmail.com

Twitter: @lekeonilude