

Generative Adversarial Network (GAN), Ian J. Goodfellow et. al. (2014)

- **Basics**
- **A simple Implementation—generating the MINIST dataset**
- **Cost function**
- **Optimization**
- **The training algorithm**

- Refer to wonderful talks by Ahlad Kumar, paper of Ian Goodfellow et.al. [gan](#)
- GAN is the most interesting idea in the last 10 years—Yann LeCun

Generative Adversarial Network

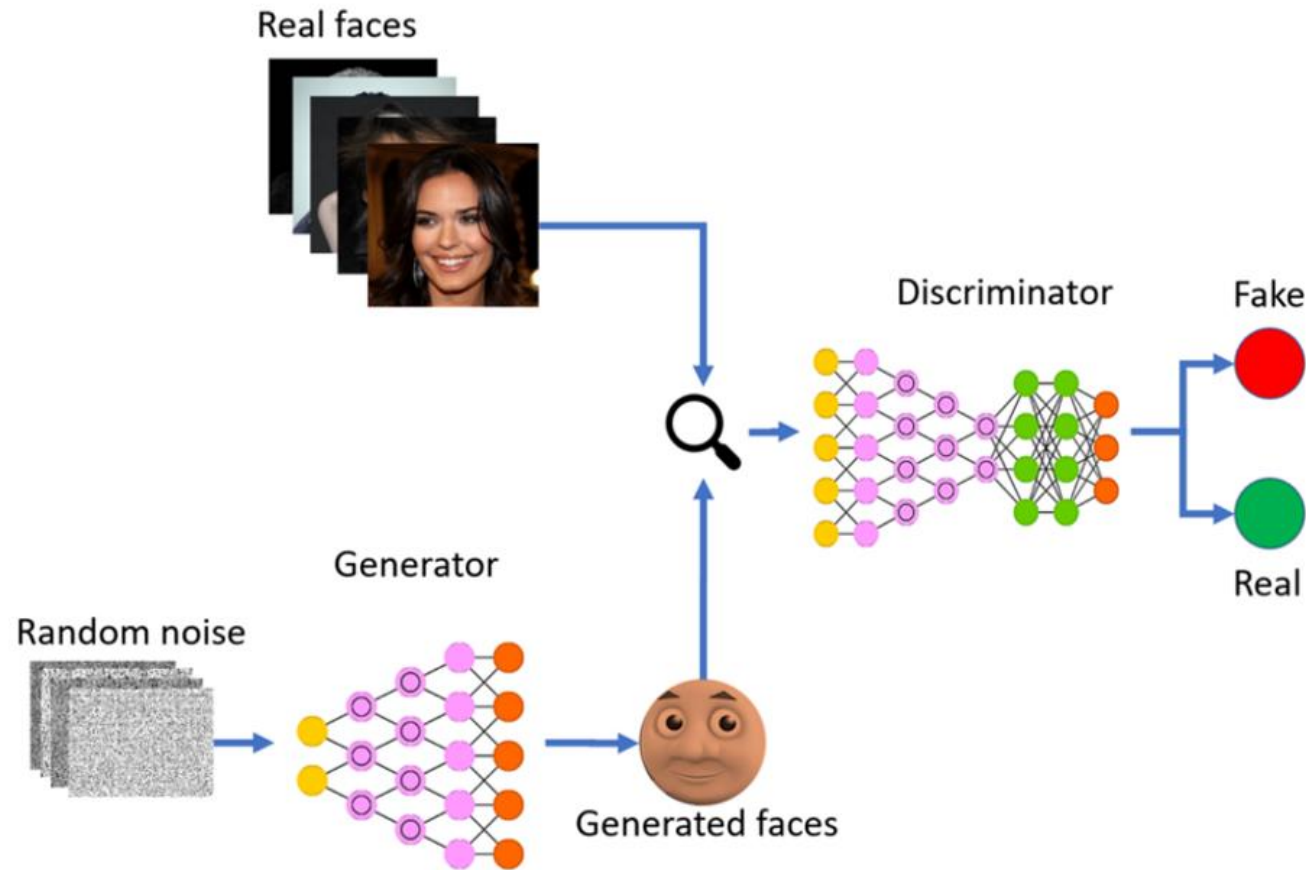
Definition: GANs are a deep neural net architectures comprising of two networks, competing one against the other

- **Adversarial**—conflict, opposition, two networks compete with each other in training, generator (G) network & discriminator (D) network
- GAN are neural networks that are trained in an adversarial manner to generate data mimicking the distribution of a given data set.
- To generate fake data that look like real! To transform the style of an art work! (For example, a book) “Little Red hat” to poetry style
- In a nutshell, **learning the probability distribution of a given data set.**

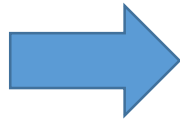
Two types of model in a GAN

- Discriminative model (D) – It is the one that distinguishes between real data samples from fake data samples.
- Generative model (G) – The generative model G is trained to generate fake dataset that has the same distribution (pdf) as that of a true dataset.

The initial state, G is not good at generating!



After training, images become so real!



Before: fuzzy digits, After: photorealistic faces

Insane task! convert zebras into horses

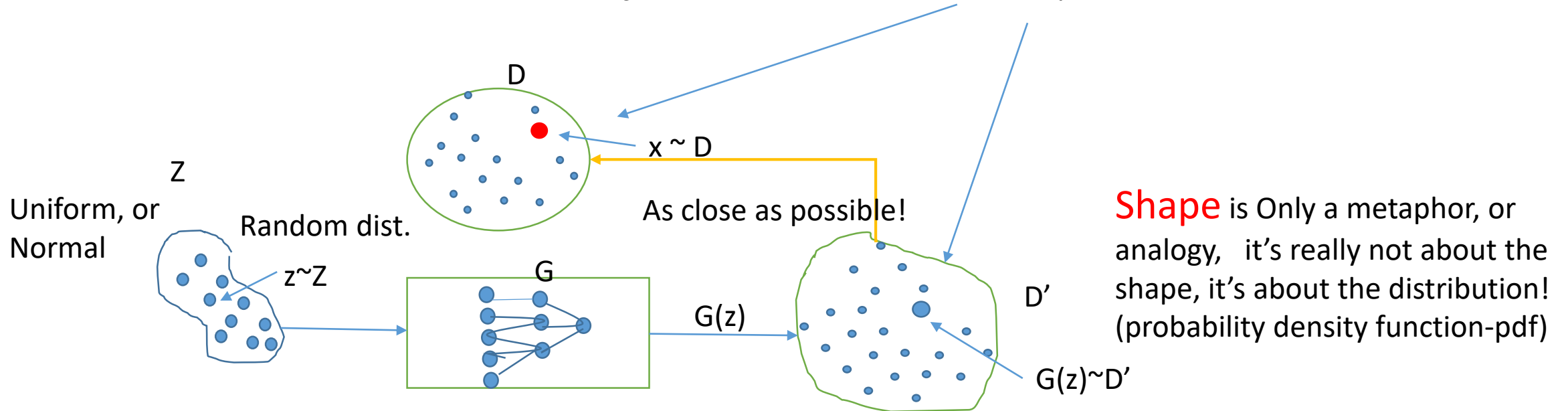


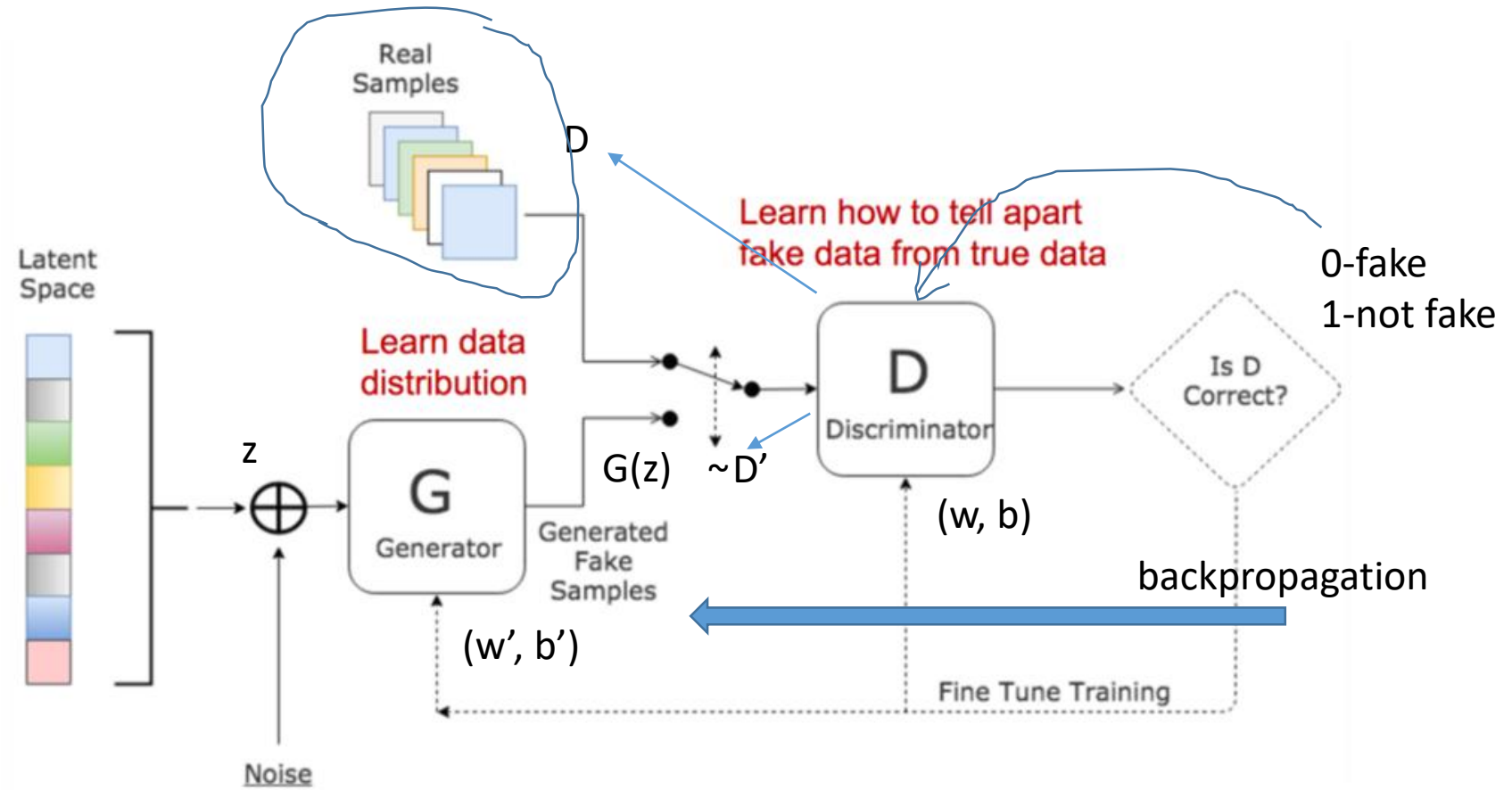
Generative model G

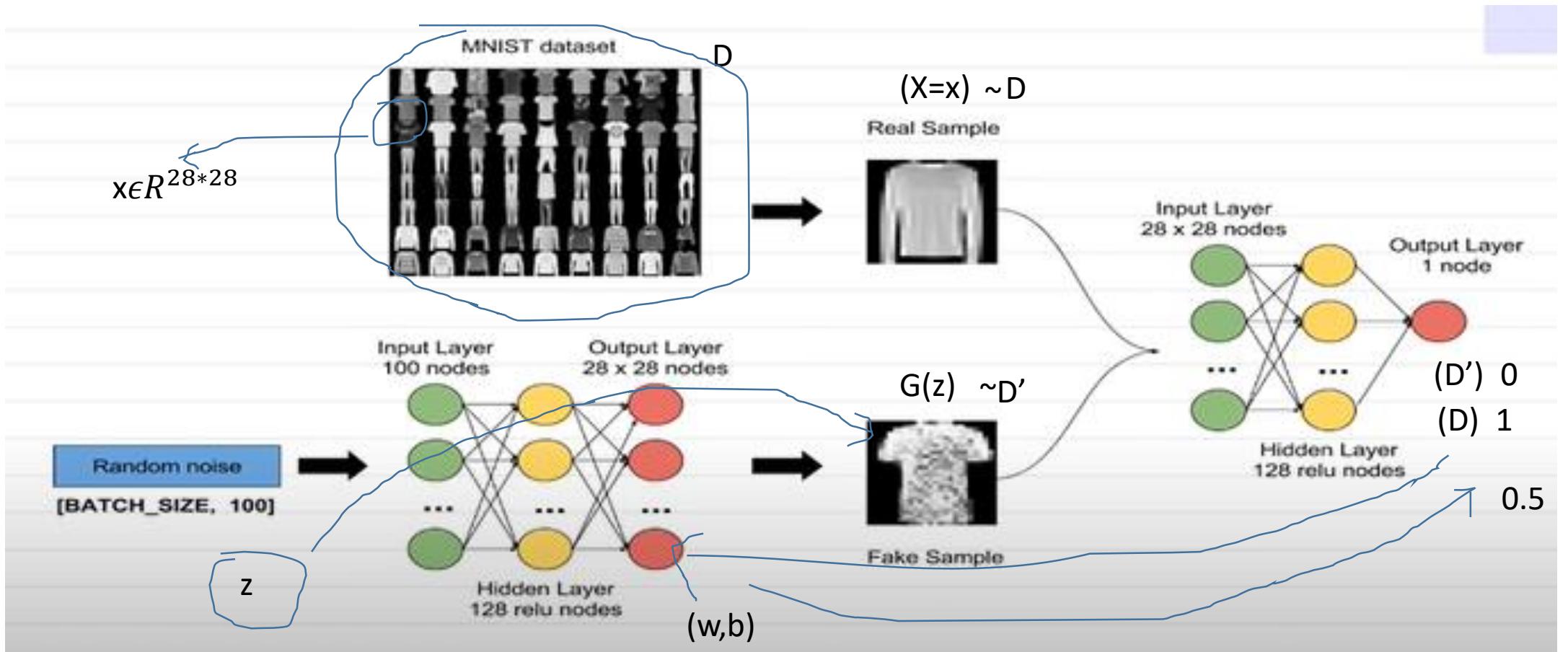
Z produces a distribution D' which is close to D according to some closeness metric. Mathematically,

$z \sim Z$ maps to $G(z) \sim D'$

Objective: D' as close as possible to D

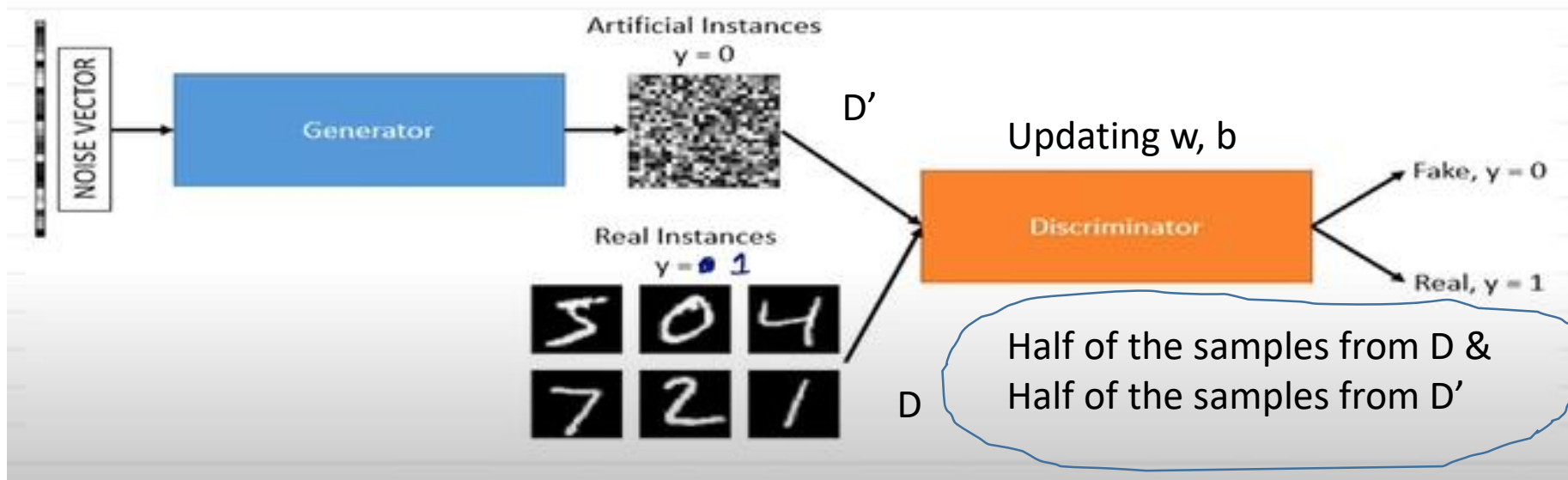




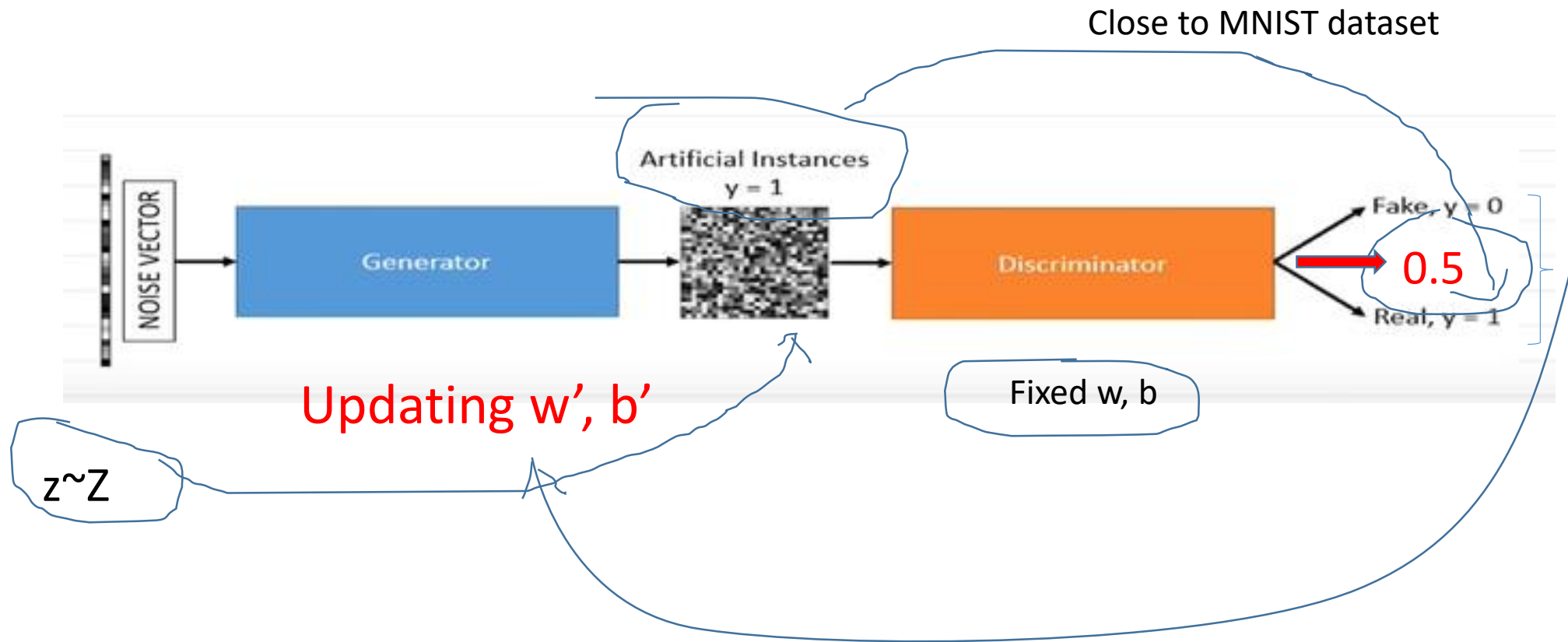


Learning mechanism

- Training the discriminator



Training the generator



The math

- [Part 1](#) (Loss function)
- [Part 2](#) (Optimization)
- [Part 3](#) (The GAN training algorithm)
- [Part 0](#) (Introduction)