MATHEMATICS BEHIND GAU

Harshvardhan

Flow of content

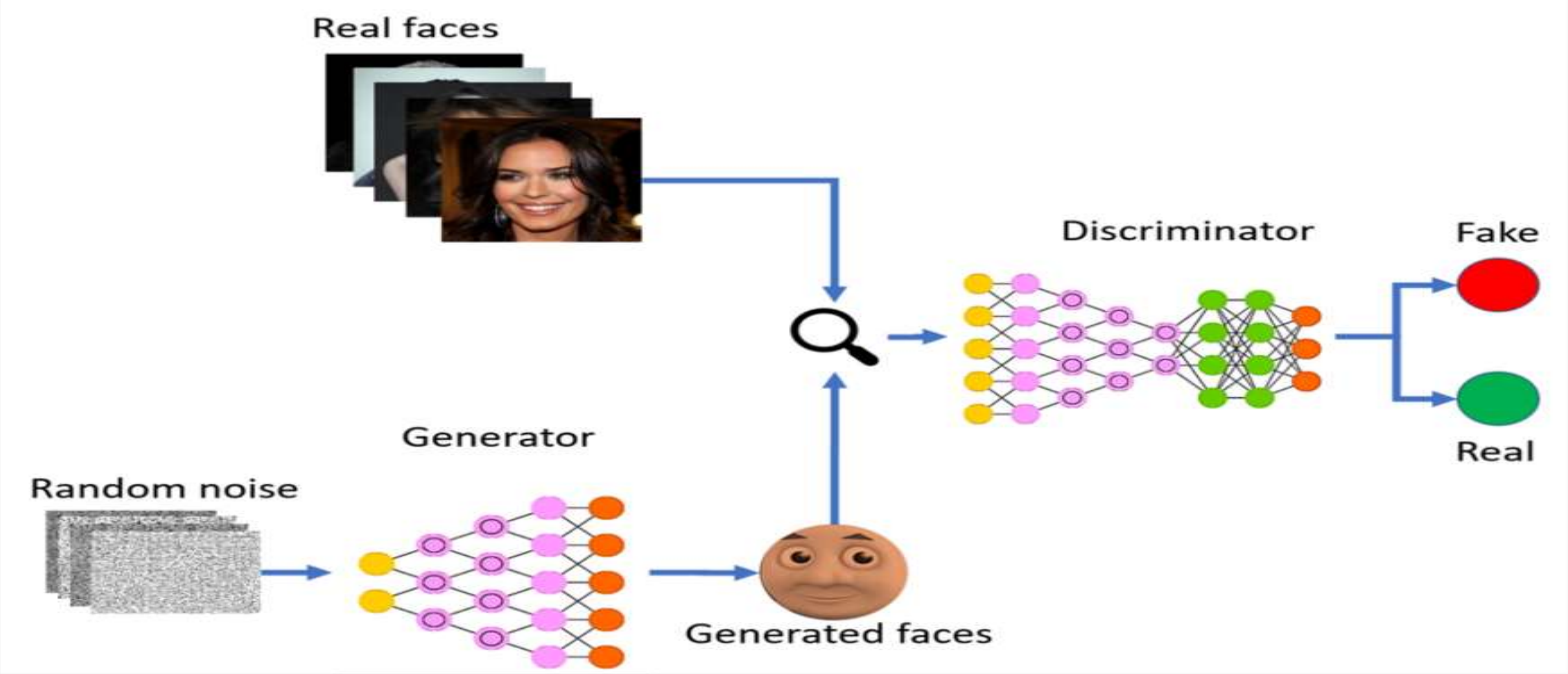
WHAT IS GAN.

MATHS BEHIND.

WHY IT IS USED.

WHAT IS GAN

It is like a game of tug of war between Generator and Discriminator.



GAN working process

- It take noise for generator N.N. and make a false input.
- Simentaniously we provide True and false data to discriminator, for which it calculate the lose for both fake input and real input and give there combine lose as D lose.
- G lose is also there calculated by gen. only by fake image.
- They tries to minimise there loses by suitable algorithm like Gradient decent.
- And G become stonger and stringer in generating images like real while discriminator become stronger to tell the difference.

Maths

DIS. Is a binary function which produces a high probability for the real image and a low for the fake image.

Dis Tries to maximise the probability of real image i.e. 1, $D(x)$. And simultaneously decrease the probability of fake data, $D(G(z))$.

In its counterpart Generator tries to increase the probability of fake data.

At last, they reach a Nash equilibrium

While doing what we said above we reach the following equation of losses for generator and Discriminator.

At Discriminator D

$Dloss_{real}$ = log (D(x))

$Dloss_{fake}$ = log (1-D (G(z)))

$$Dloss = Dloss_{real} + Dloss_{fake}$$

log (D(x)) + log (1-D (G(z)))

The total cost is

$$\frac{1}{m}\sum_{i=1}^{m} \log(\mathrm{D}(x^i)) + \log(1 - \mathrm{D}(\mathrm{G}(z^i)))$$

At Generator G

$Gloss$ = log (1-D (G(z))) or - log (D (G(z)))

The total cost is

$$\frac{1}{m}\sum_{i=1}^{m} \log(1 - \mathrm{D}(\mathrm{G}(z^i)))$$

or

$$\frac{1}{m}\sum_{i=1}^{m} -\log(\mathrm{D}(\mathrm{G}(z^i)))$$

D and G play the following two-player mini-max game with value function V (G, D)

- E is an EXPECTATION VALUE/ Probability distribution

$$\min_{G} \max_{D} V(D, G) = \mathbb{E}_{\boldsymbol{x} \sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

$$\max_D V(D) = \underbrace{\mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})]}_{\text{recognize real images better}} + \underbrace{\mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]}_{\text{recognize generated images better}}$$

$$\min_G V(G) = \underbrace{\mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]}$$

Optimize G that can fool the discriminator the most.

TRANING LOOP
fix the learning rate G
take m data samples with m fake samples
update Theta d by gradient decent

$$\frac{\sigma}{\sigma\theta d} * 1/m(\ln(D(x) + \ln(1 - D(G(z)))$$

fix the learning rate D
take m fake samples
update Theta d by gradient decent

$$\frac{\sigma}{\sigma\theta g} * 1/m(\ln(1 - D(G(z)))$$

Pg converge to Pd if our value function find the global minimum of min-max function.

USES

- Generate Examples for Image Datasets
- Generate Photographs of Human Faces
- Generate Realistic Photographs
- Generate Cartoon Characters
- Image-to-Image Translation
- Text-to-Image Translation
- Semantic-Image-to-Photo Translation
- Face Frontal View Generation
- Generate New Human Poses
- Photos to Emojis
- Photograph Editing
- Face Aging
- Photo Blending
- Super Resolution
- Photo Inpainting
- Clothing Translation
- Video Prediction
- 3D Object Generation