

Generative Adversarial Networks (GANs)

Advanced Institute for Artificial Intelligence – AI2

<https://advancedinstitute.ai>

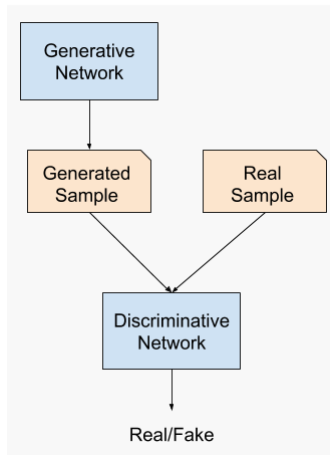
GANs são um tipo de arquitetura de Redes Neurais para o treinamento de modelos generativos

- **Modelagem Generativa** consiste em criar um modelo que gera novos exemplos seguindo uma "distribuição" específica
 - Por exemplo, gerando fotografias realísticas mas diferentes das fotos utilizadas durante o treinamento

Uma GAN é um modelo generativo treinado utilizando 2 Redes Neurais.

- **Gerador:** Aprende a gerar novos exemplos plausíveis
- **Discriminador:** Aprende a diferenciar exemplos gerados de exemplos reais.

O *Gerador* e o *Discriminador* competem em "um jogo".



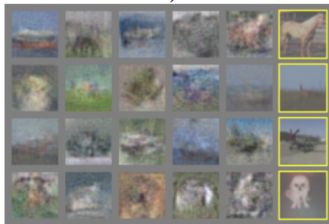
- Após o treinamento, o *Gerador* pode ser utilizado para criar novos exemplos realísticos.



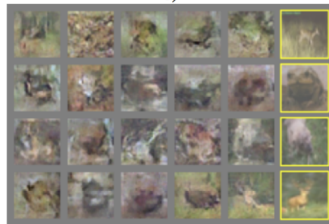
a)



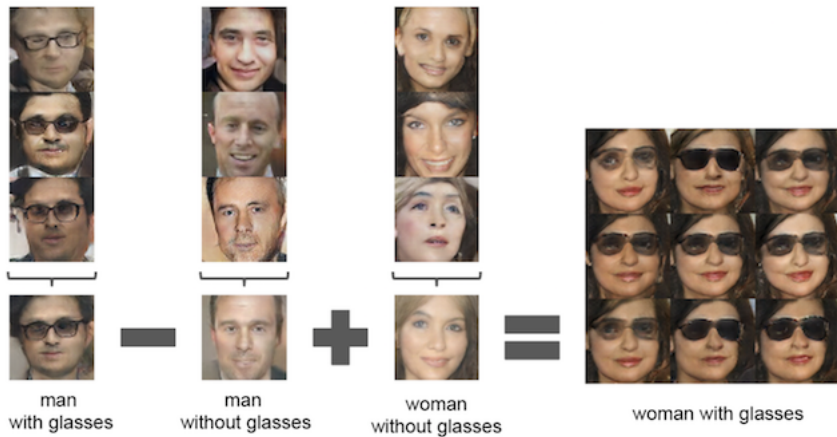
b)



c)



d)



Input



Output



Input



Output



Real

Input

Ours

NN



Como as GANs são treinadas?

- O **Discriminador** $D(x)$ deve aprender a distinguir um exemplo real de um exemplo gerado:

$$\mathcal{L}_D = -\frac{1}{2}\mathbb{E}_{\mathbf{x}}\log D(\mathbf{x}) - \frac{1}{2}\mathbb{E}_z\log(1 - D(G(z)))$$

, onde $G()$ é o modelo gerador e z é um vetor aleatório.

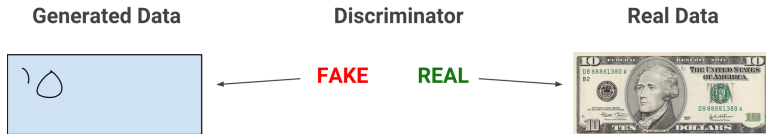
- Com essa função de custo, o *Discriminador* tenta aprender a distinguir exemplos sintéticos dos reais.
- Em contraste, o **Gerador** tenta gerar exemplos sintéticos indistinguíveis dos reais.
- A forma mais simples de modelar essa dinâmica é através de um *jogo de soma zero*.

$$\mathcal{L}_D = -\mathcal{L}_G \quad (1)$$

- O que significa que o *Gerador* vai tentar enganar o *Discriminador*.

O que acontece durante o treinamento?

Início



Após um tempo treinando



Após convergência



O Treinamento geralmente é realizado em ciclos:

- 1 Treina o $D(x)$ para que este consiga classificar samples
- 2 Treina $G(x)$ para que gere samples que enganem D .

Treinar GANs é muito desafiador, dado que é difícil para o gerador descobrir a "direção" em que seus updates devem ir para enganar o discriminador.

Dicas Avançadas: <https://github.com/soumith/ganhacks>