

Generative Adversarial Networks | GAN

Generative Models

- Produces an image that does not exist but is likely to exist.
- A statistical model of the joint probability distribution.
- An architecture to generate new data instances.

Generative Adversarial Network

- generator와 discriminator, 두개의 네트워크 합성
- objective function

$$\min_G \max_D V(D, G) = \underbrace{E_{x \sim p_{data}(x)} [\log D(x)]}_{\text{discriminator 합성}} + \underbrace{E_{z \sim p_z(z)} [\log (1 - D(G(z)))]}_{\text{generator 합성}}$$

기존의 샘플 (노이즈) latent vector (노이즈)

합성이 안되면 1/2이 됨 (자판기가 자판 뽑아줌)

- D, G 한쪽만 가면서 합성
- KL divergence : 두개의 분포가 얼마나 차이가 나는지 나타내는 공식
- Jensen Shannon divergence : 분포의 distance 구하는 데 사용

★ G의 loss가 높은 이유 : discriminator가 진짜 레이어를 진짜로 판별할 때 loss가 떨어짐. 합성 초기에는 discriminator가 다름에 따라 진짜를 판별할 수 있어서 discriminator의 loss는 낮고, generator의 loss는 높음.

★ Markov chain : 추후적인 시퀀스이고, 미래의 상태가 오직 현재 상태에만 의존하며 과거의 상태는 의존하지 않는 특성을 가진 확률 과정. 이러한 generative model들이 많이 사용된다고 함.

★ inference network : 레이어로 복귀를 통해 구조를 제공하는 모델. 이런 generative model들이 많이 사용된다고 함.

★ probabilistic computation : 결과를 직접 결정짓고 (항상 동일한 출력을 제공하는) 하지 않고, 대신 확률 분포를 사용하여 여러가능한 결과 중 하나를 제공한다.

★ generator가 레이어의 density에 대한 정보를 가지고 있다 = generator가 각 레이어 포인트가 생성될 확률이 얼마나 되는지를 알고 있다

★ 가우시안 분포 = 정규분포