

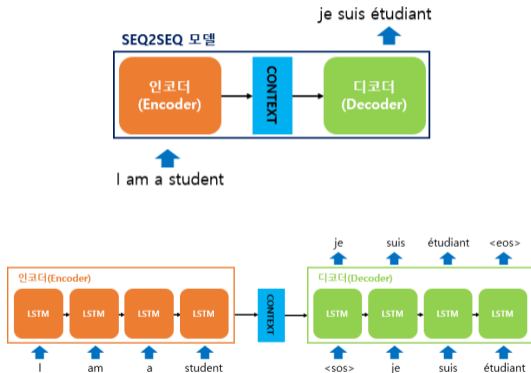
# AutoEncoder

by

김유겸

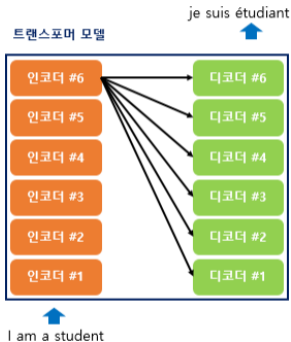
2024.10.08

- ✓ 입력 시퀀스를 받아 출력 시퀀스를 받는 구조의 모델
- ✓ 인코더는 입력을 순차적으로 받아 context vector를 반환
- ✓ 디코더는 context vector와 이전 단계에서 예측한 단어를 통해 순차적으로 단어를 예측

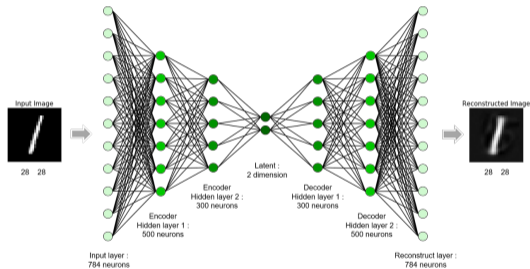


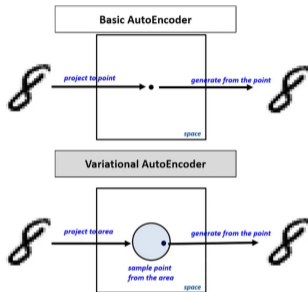
# Transformer

- ✓ Seq 2 Seq 모델에서는, 입력 시퀀스의 정보가 context vector로 압축되는 과정에서 손실이 발생
  - attention을 통하여 보정
- ✓ attention을 보정의 용도로 사용하는 것이 아닌, attention만으로 인코더와 디코더를 구성
- ✓ RNN기반 구조와 다르게 순차적으로 입력을 받지 않음



- ✓ Autoencoder는, 입력된 값을 압축한 뒤 다시 원본으로 복원하는 모델
- ✓ 압축 부분을 Encoder, 복구 부분을 Decoder
- ✓ 앞에서 본 context vector와 같은, latent vector가 존재
- ✓ Encoder는 데이터를 latent로 잘 나타내도록 학습, Decoder는 데이터를 잘 복원하도록 학습

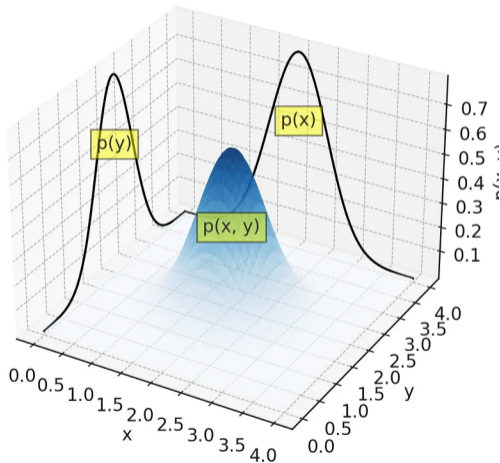




- ✓ Variational AutoEncoder(VAE)는 Autoencoder 구조에 확률 분포를 추가한 것
- ✓ Encoder를 통해 만들어진 latent vector에 가우시안 분포를 이용해 영역을 디자인

# Variational AutoEncoder

- ✓ 가우시안 분포들의 평균과 표준편차를 통해, 영역을 만들 수 있음
- ✓ VAE의 Encoder는 latent vector를 만드는 것이 아닌, 가우시안 분포들의 평균과 표준편차를 학습
- ✓ Decoder는 해당 영역에서 임의의 샘플링을 통하여 latent vector를 얻어내고, 데이터를 복원함



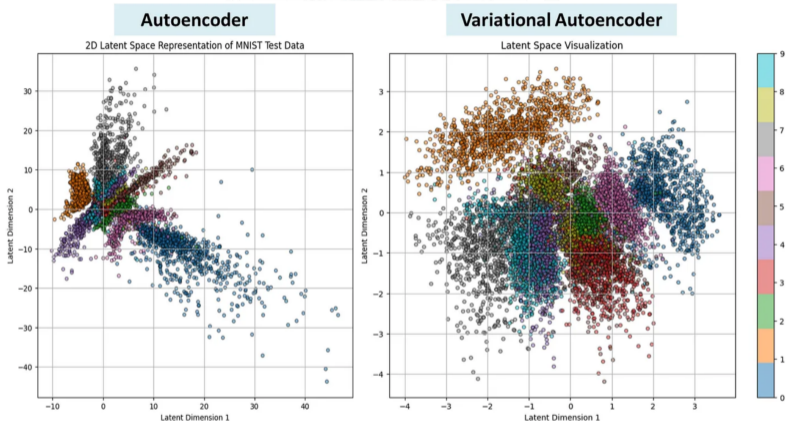
- ✓ Decoder가 임의 샘플링을 통해 결과물을 만들어내면, 역전파는 어떻게 디코더까지 전파되는가?
- ✓ Reparametrization Trick을 사용
- ✓ 먼저, latent vector  $z$ 는  $N(\mu, \sigma^2)$ 를 따르고, 이는 미분 불가능함
- ✓  $z = \mu + \sigma \cdot \epsilon$ , ( $\epsilon$ 은  $N(0, 1)$ 에서 샘플링한 상수)으로 바꾸어 쓴다
- ✓  $\epsilon$ 의 정규분포를 알고 있기 때문에, 몬테카를로 방법을 사용하여 미분 가능한 식으로 근사시킬 수 있음

- ✓ VAE의 loss function으로는 Mean Squared Error나 Binary Cross-Entropy를 사용
- ✓ 이미지 복원의 정확함 뿐만이 아니라, Encoder가 적절한 가우시안 분포를 만드느냐를 판정해야 함
- ✓  $\mu$ 와  $\sigma$ 에서의 확률분포가 정규분포와 얼마나 가까운지를 계산하는 데 KL Divergence를 사용

$$\text{KL Divergence} = D_{KL}[q(z|x)||p(z)]$$

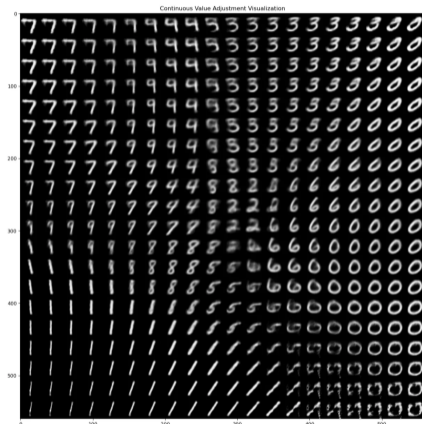
- ✓ 어떠한 분포  $q(z|x)$ 와 지정한 분포  $p(z)$ 와의 차이를 측정

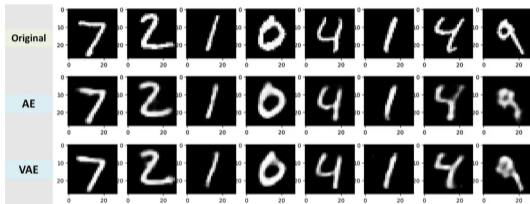




- ✓ AutoEncoder와 Variational AutoEncoder의 latent space 비교
- ✓ VAE가 좀 더 정규분포에 가까운 공간을 만드는 것을 볼 수 있음

- ✓ 2차원 latent space를 시각화 한 이미지
- ✓ 비슷한 숫자끼리 인접한 것을 알 수 있음





- ✓ AE와 비교해서, VAE의 복원이 입력 데이터에 더 근접한 것을 알 수 있음