# CS5242 Tutorial

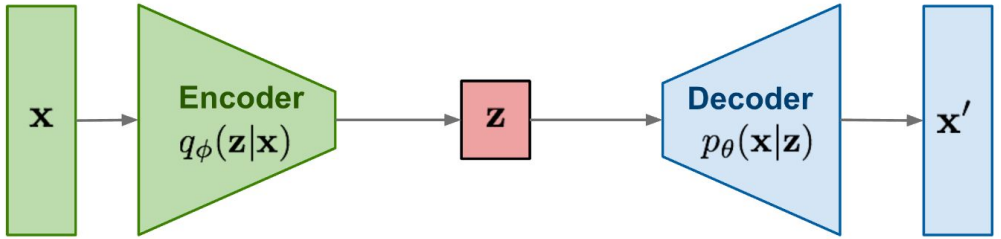
Week11

Homework 5

# VAE

- VAE(Variational Autoencoder) compress and de-compress video for Sora.


**VAE:** maximize variational lower bound

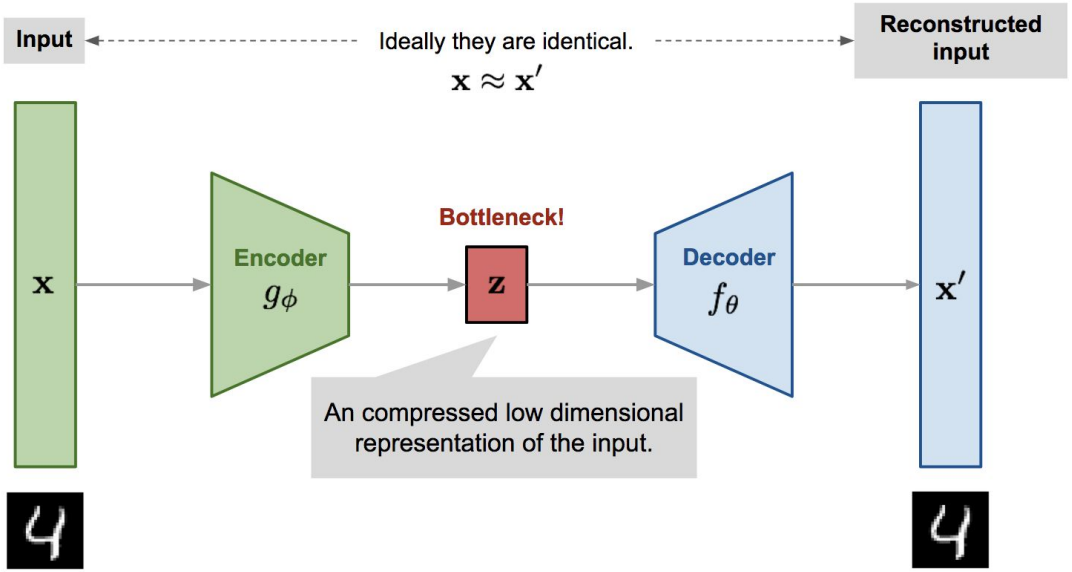


# Assignment 4

- Task: Implement AE(Autoencoder) and VAE (Variational Autoencoder) on images.
- DDL: 23:59 11 Apr (in 10 days)
- Tutorial: Introduce AE and VAE

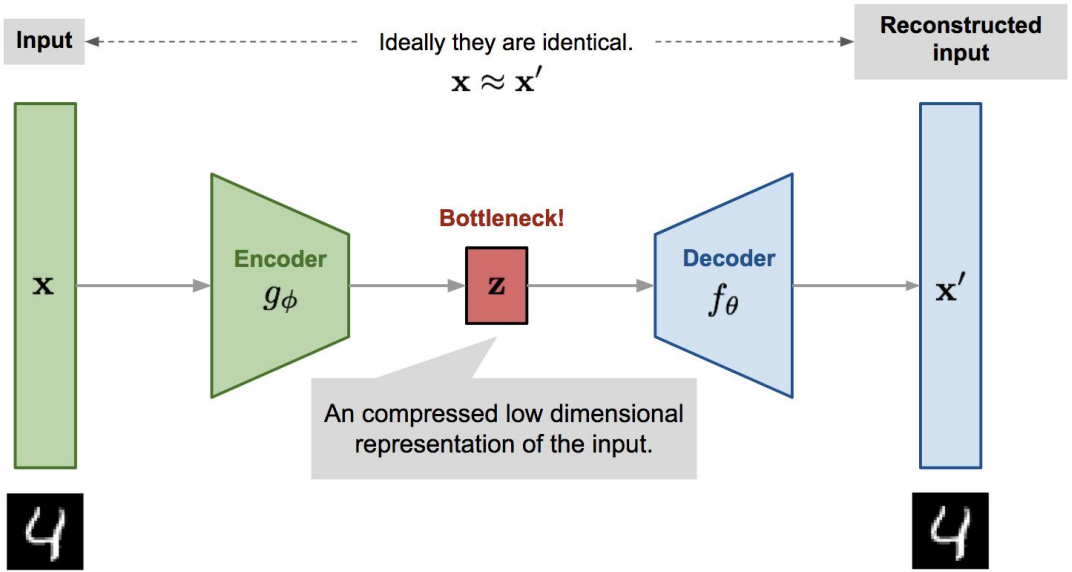
# AE

- Autoencoder: An unsupervised method to compress data.
- Encoder: Maps the original high-dimensional data to a lower-dimensional feature space.
- Decoder: Reconstructs the original data based on the compressed low-dimensional features.



# AE

- Purpose of decoder: make sure encoder learns important information losslessly
- How to train? -> target: x'- x = 0
- Loss function: MSE $L_{\mathrm{AE}}(\theta, \phi) = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{x}^{(i)} - f_\theta(g_\phi(\mathbf{x}^{(i)})))^2$

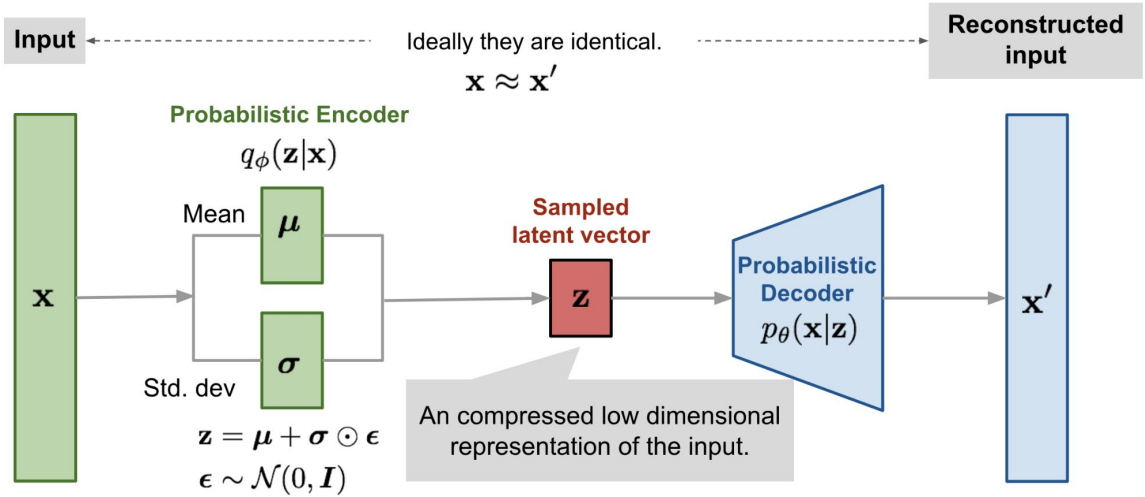


# VAE

- AE: compress data
- Problem:
  - Only see images in the training dataset. Poor generalization.
  - Cannot generate new content.
- VAE

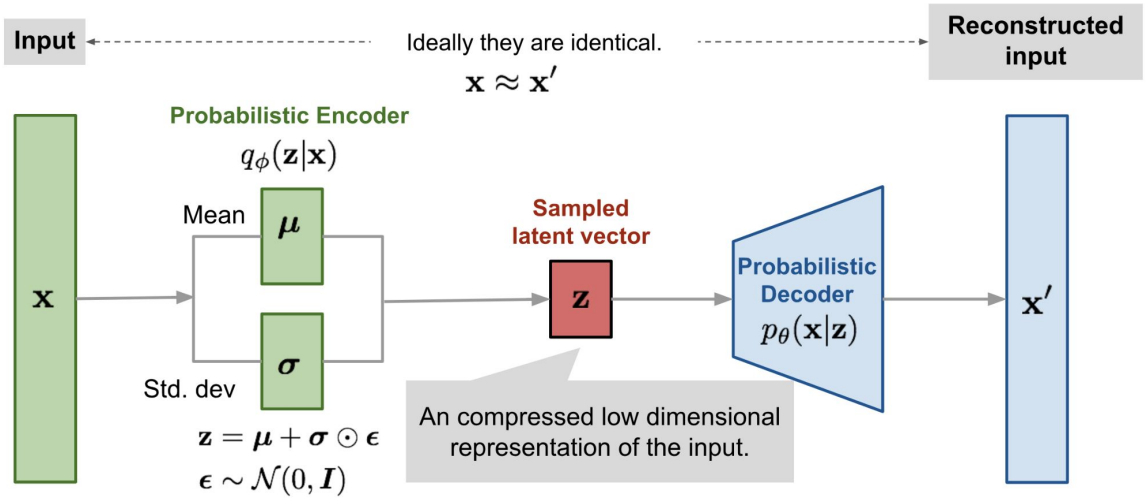
# VAE

- VAE: A probabilistic model based on variational inference.
- Intuitive guide in tutorial, more detailed mathmathic thoery materials are given at page 11.



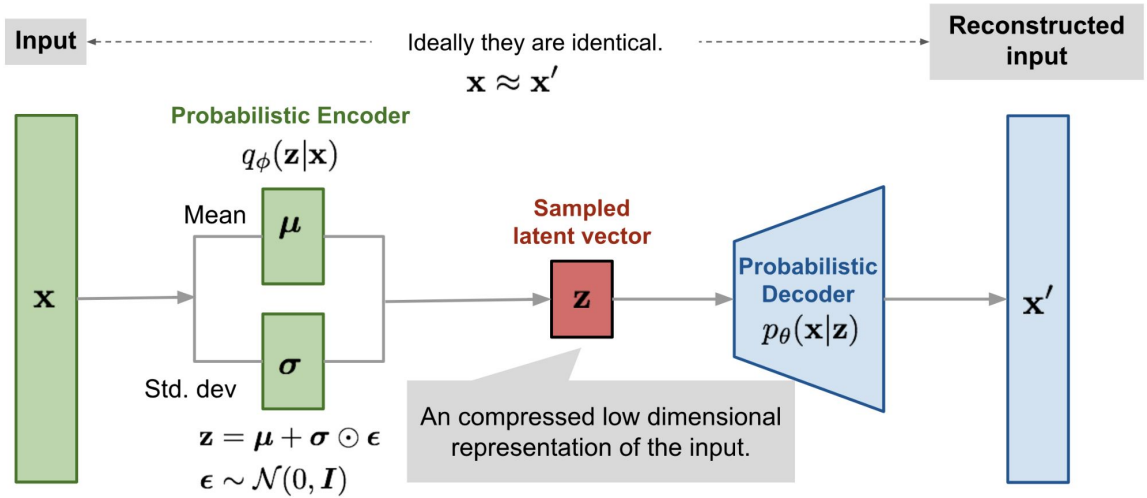
# VAE

- Encoder: Sample a Normal distribution from input.
- Decoder: Reconstructs the original data based on latent vector.



# VAE - Encoder

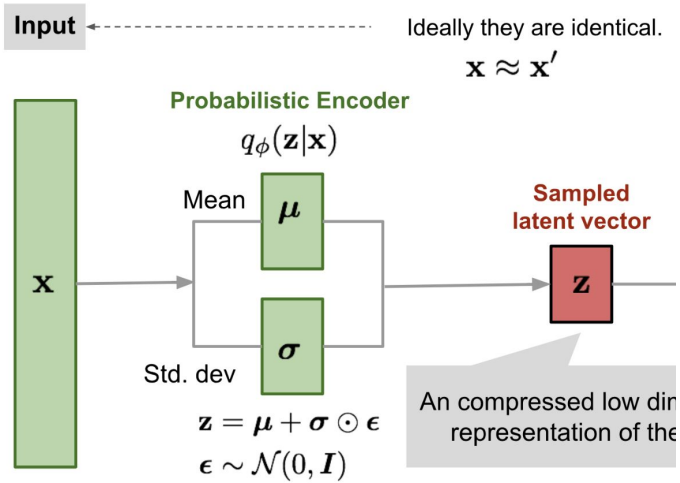
- Generate vector of mean and std based on inputs, apply std on a randomly normal distribution initialized vector, and sum them.



# VAE - Encoder

- z = mean + std * N(0,1)
- mean is the original latent vector in AE
- std * N(0,1) is the noise

- Benefits:
- force the mean close to N(0,1), so that we can generate images from N(0,1)
- learn to reconstruct the image from noise to be more general



# VAE

Reading materials:

https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73

https://lilianweng.github.io/posts/2018-08-12-vae/