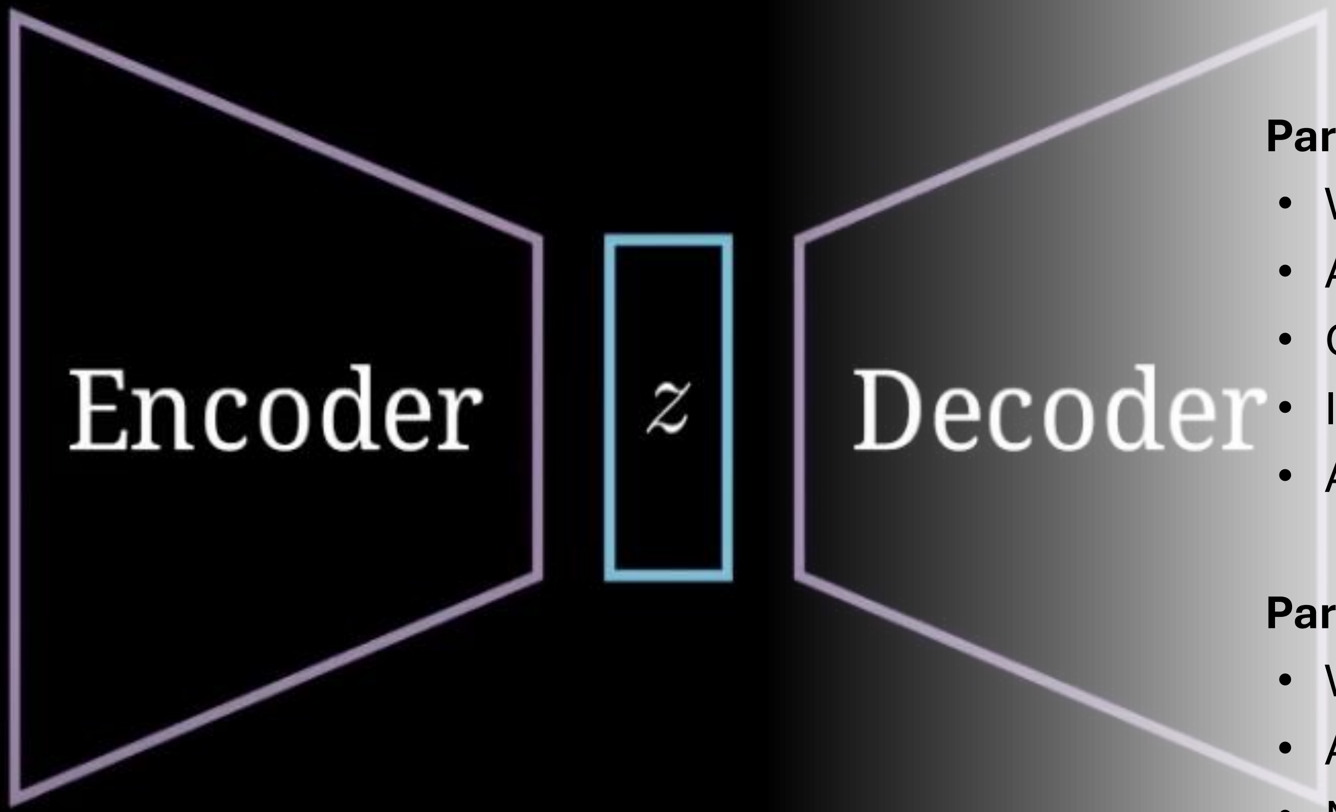
Autoencoders (Basic) & Variational autoencoders (VAE)

Sanjaya Edirisinghe

Sanjayae@bu.edu



AGENDA

Part 1 – Basic Autoencoders

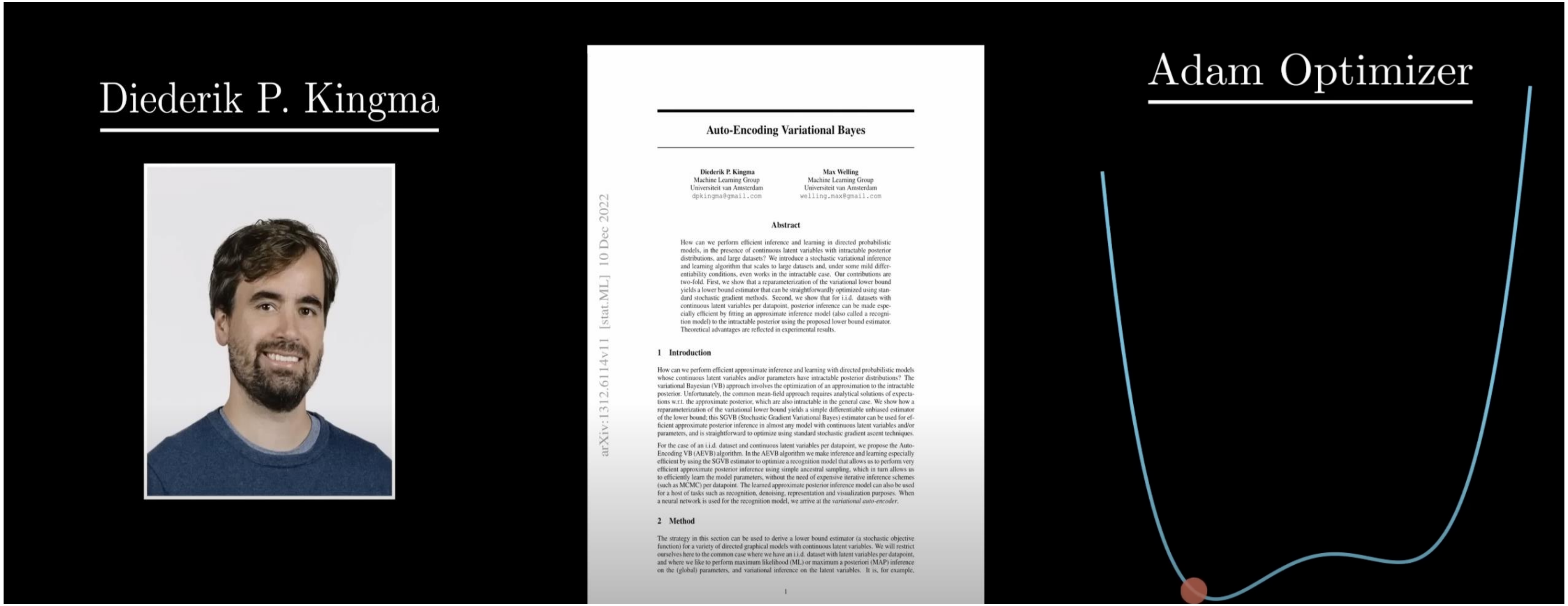
- Why Autoencoders and the use of it
- Architecture of Autoencoders
- Components of the Autoencoders
- Information loss of Autoencoders
- Applications or Autoencoders

Part 2 - Variational Autoencoders (VAEs)

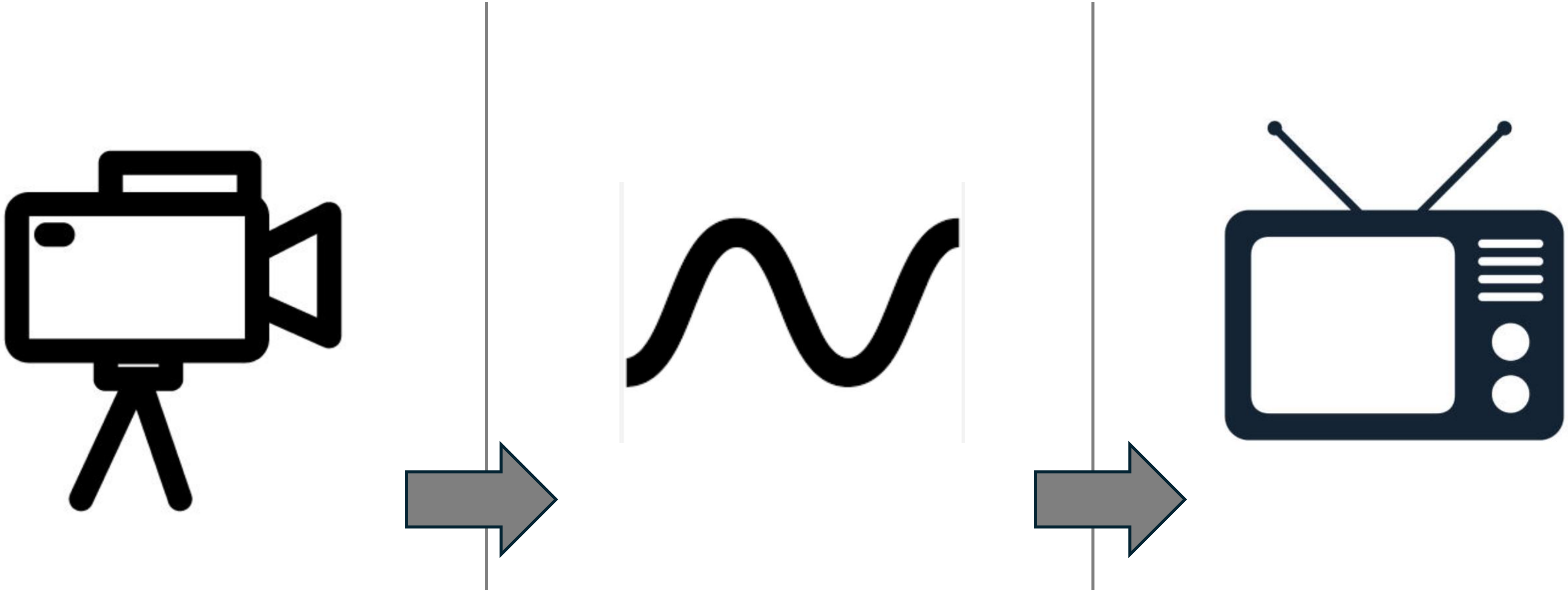
- Why VAEs and their purpose
- Architecture of VAEs
- Mathematical representation of VAEs
- How VAE's works
- Limitations of VAE's
  Different types of Autoencoders

Sanjayae@bu.edu

The inventor of Autoencoder – 2013



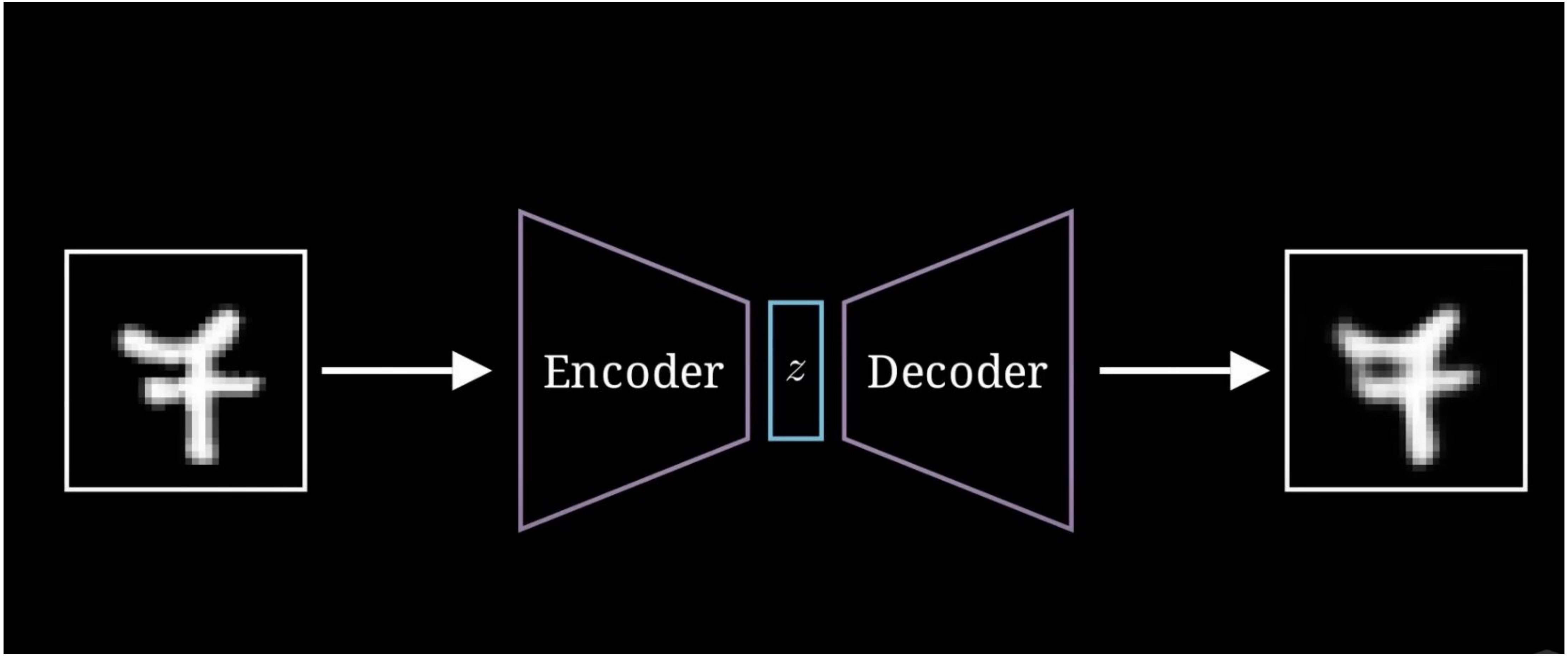
Do you know how TV is working ..



High Dimensions — Low Dimensions — High Dimensions

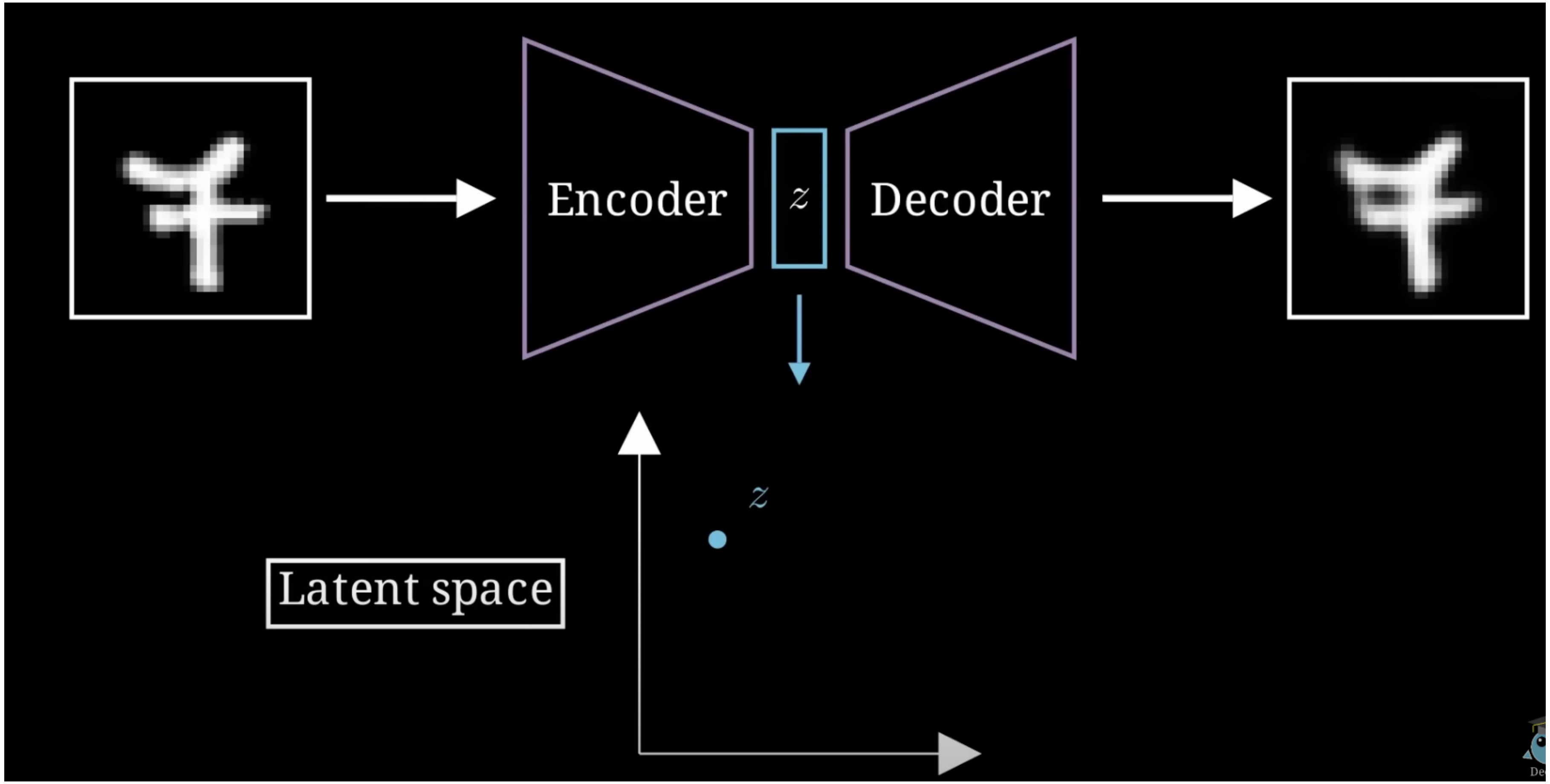
Sanjayae@bu.edu

Why Autoencoders and what does it do ?



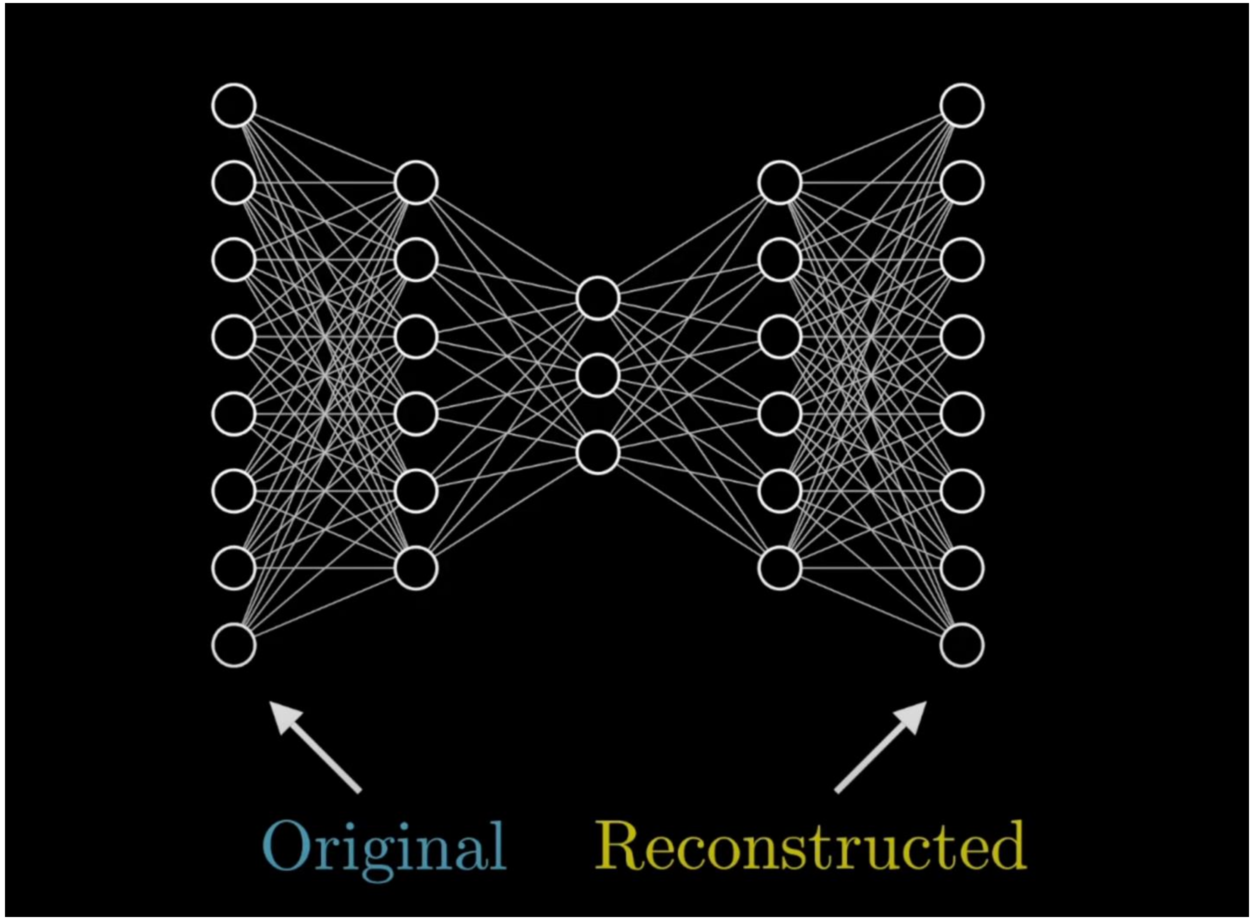
Sanjayae@bu.edu

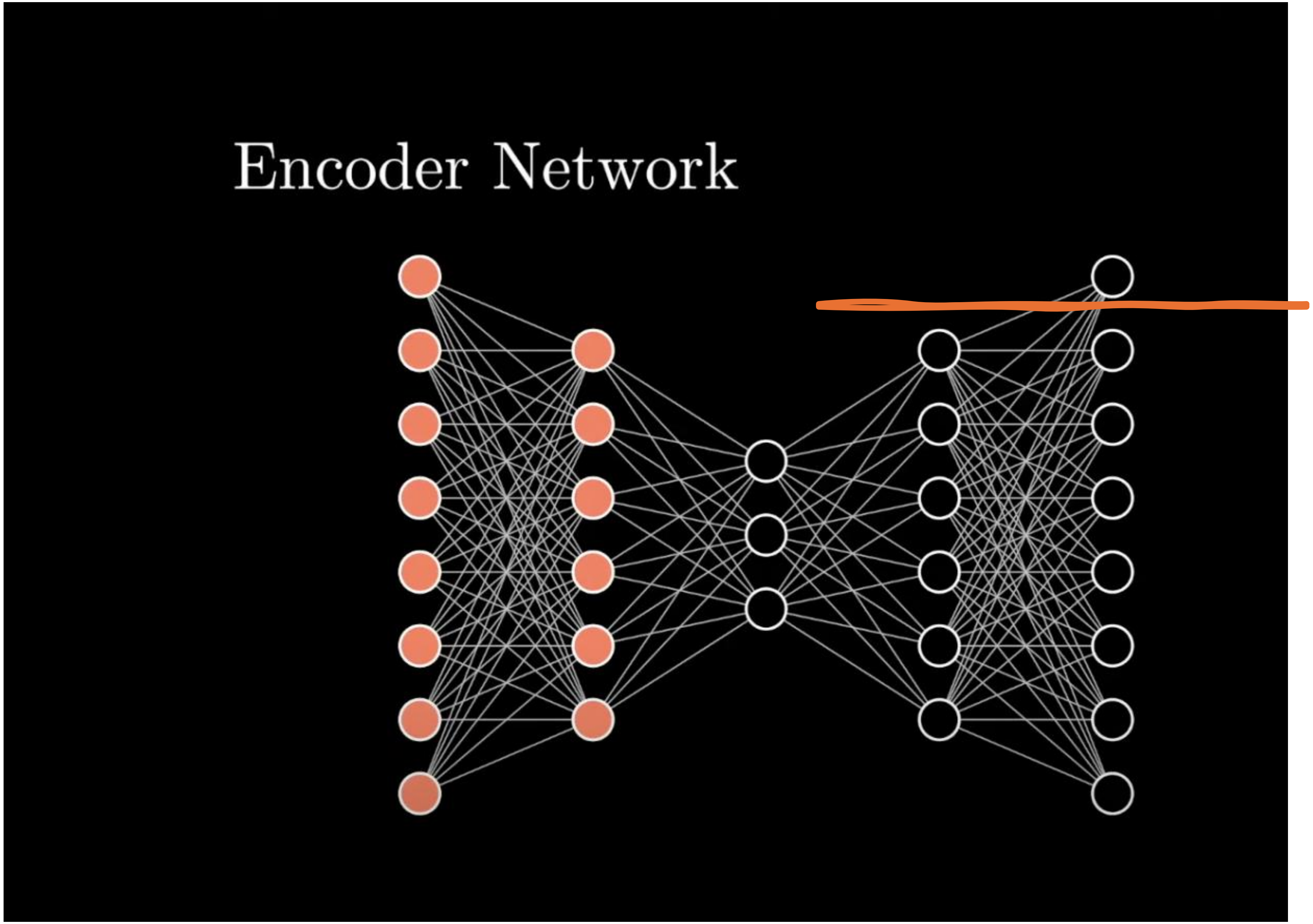
Components of Autoencoders ...



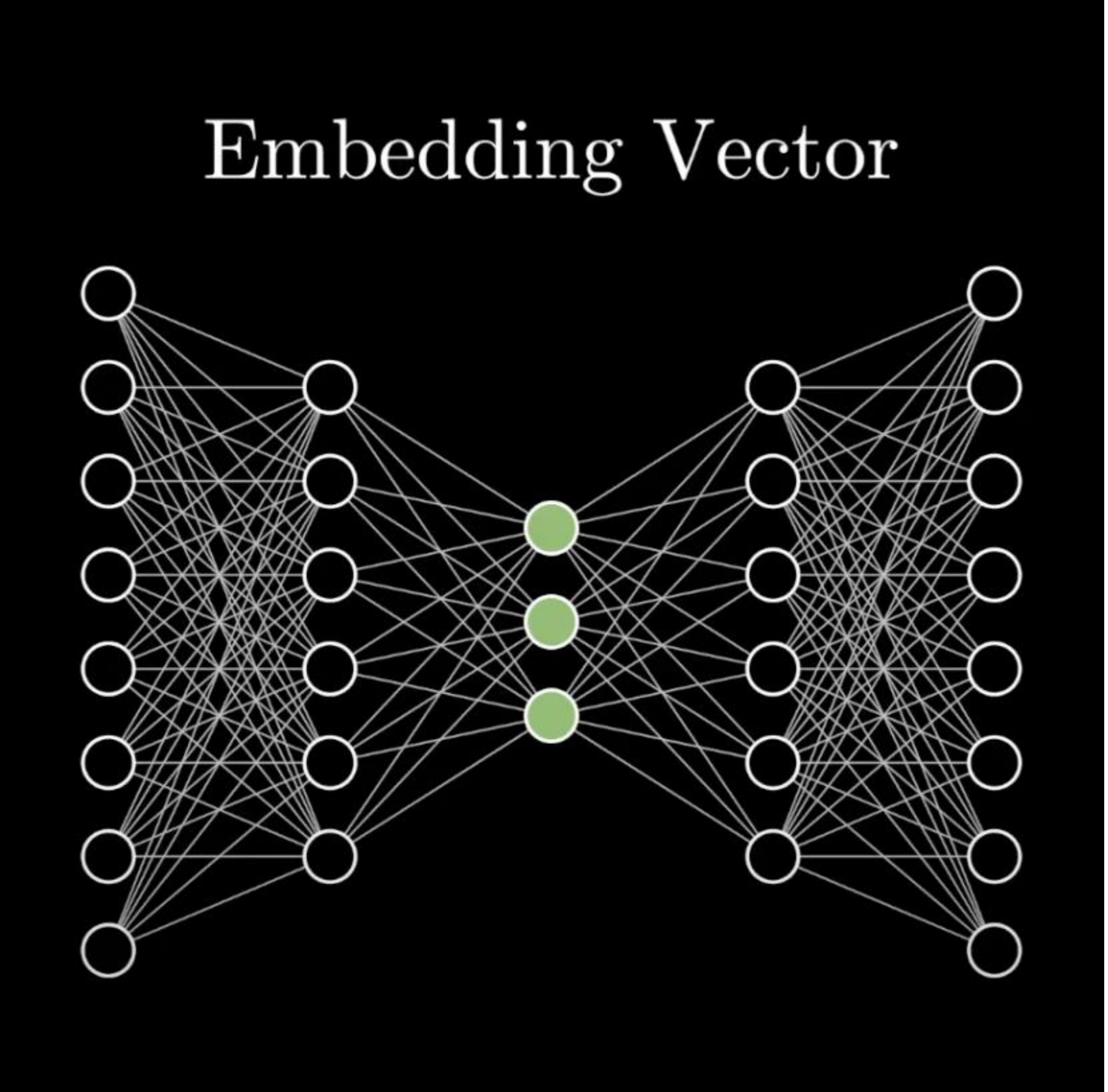
Sanjayae@bu.edu

Autoencoders are NN based deeplearning models ..

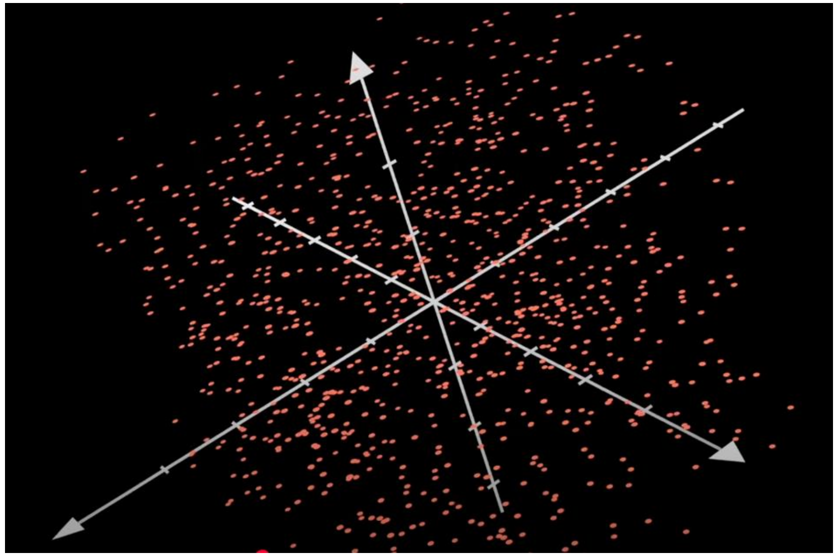




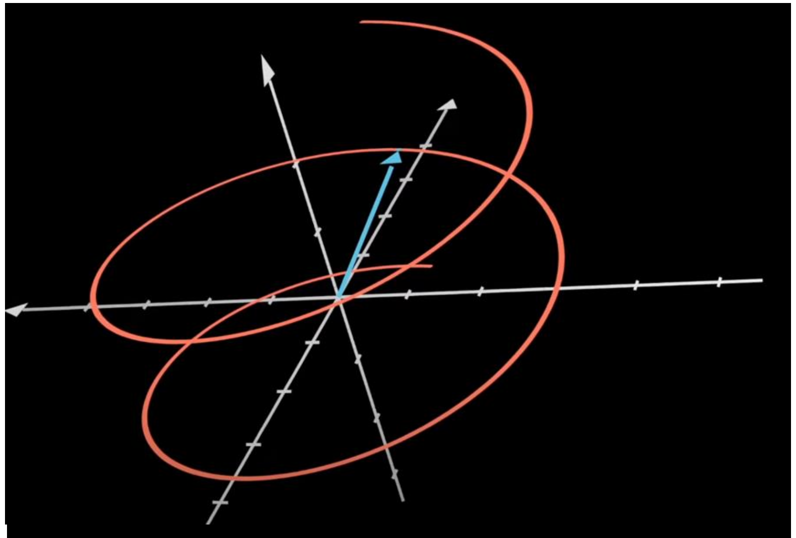
Sanjayae@bu.edu



- Bottleneck
- Low Dimensions



Random Distribution

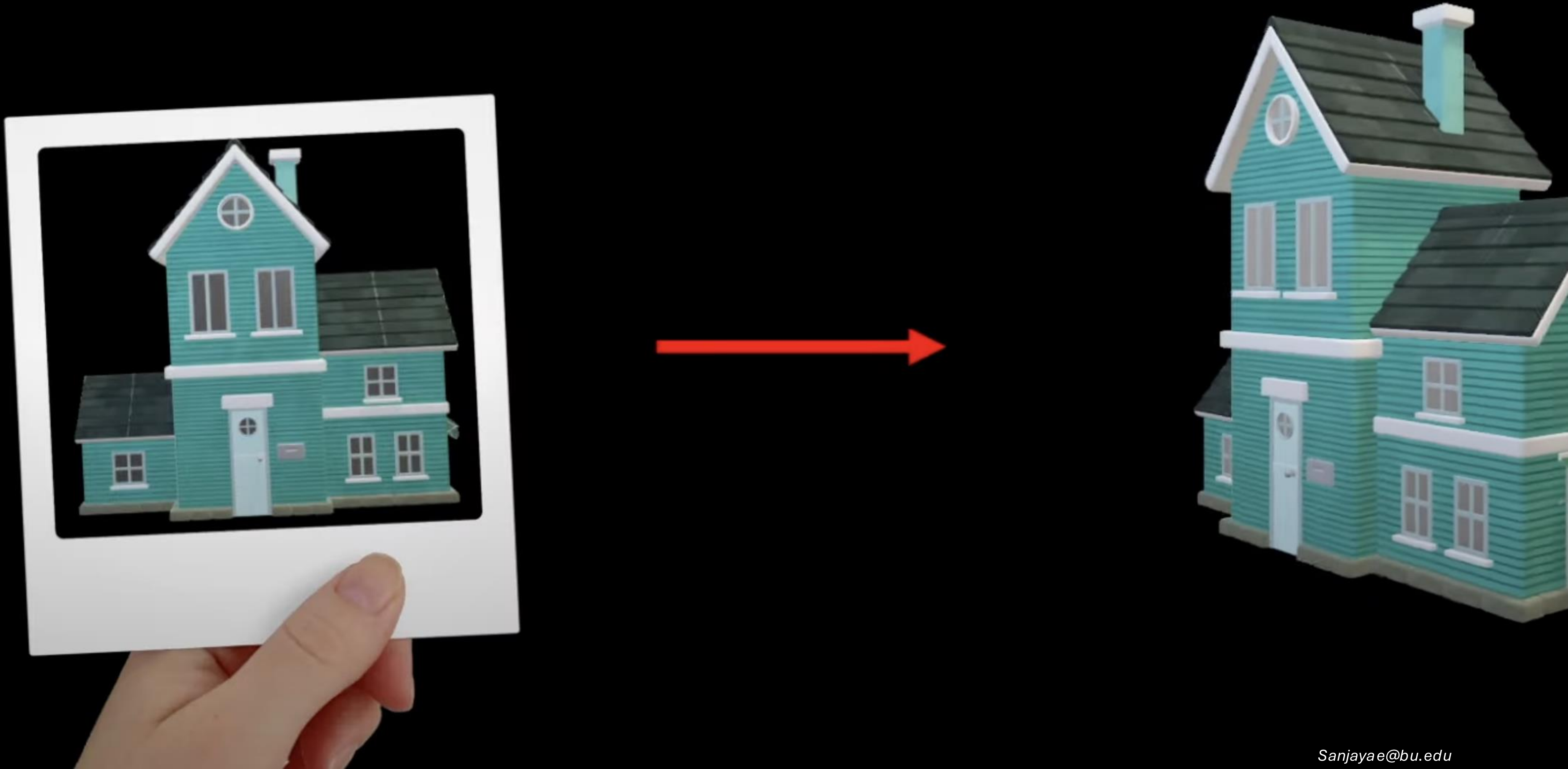


$$x = 3\cos\theta \qquad \theta = 2\pi \Leftrightarrow x = 3$$
$$y = 3\ \sin\theta \qquad y = 0$$
$$z = {}^{\theta}/_{\pi} - 1.5 \qquad z = 0.5$$

Nature has a structure, therefore most of the natural data followed a pattern behind them

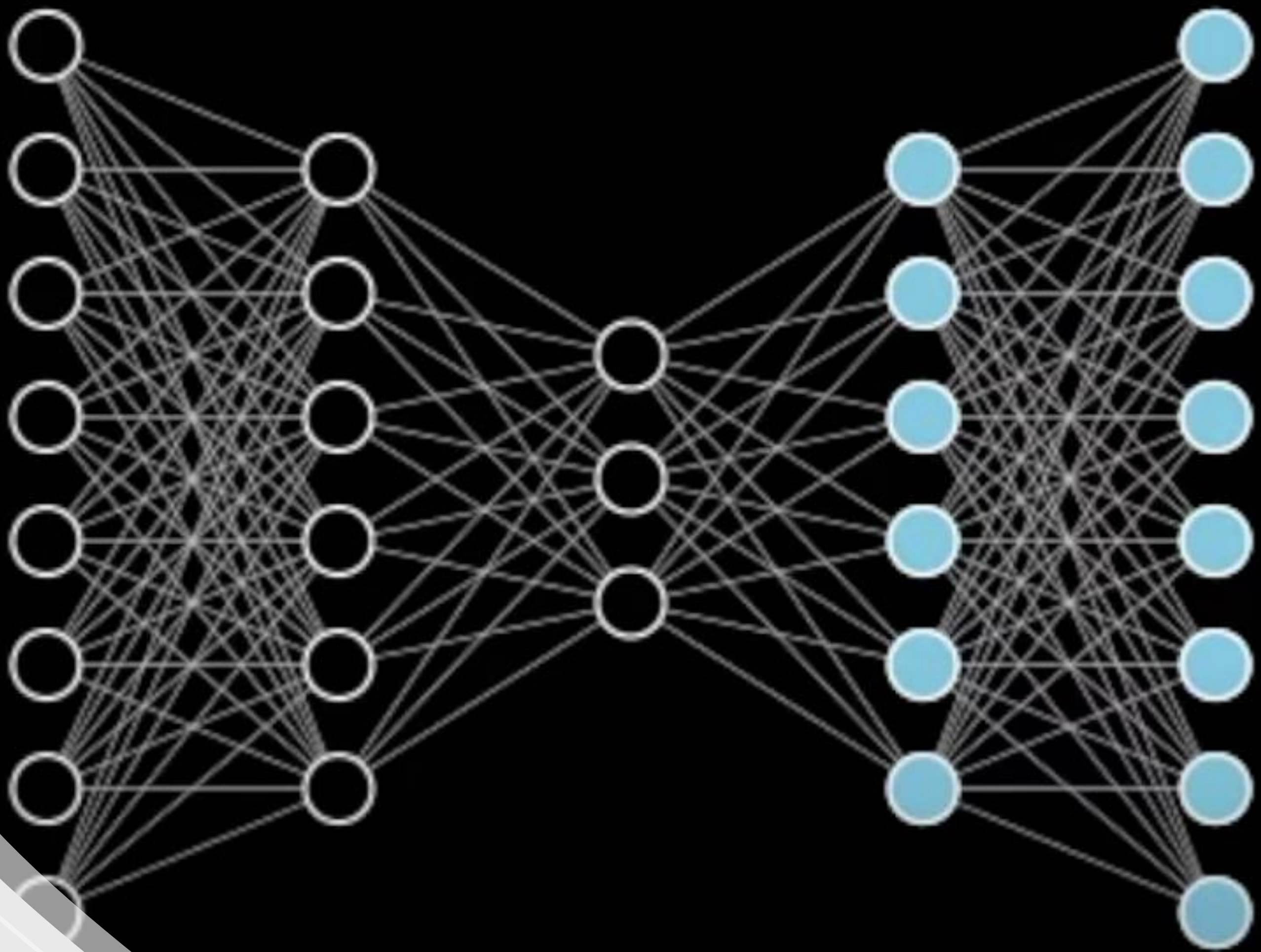
(City, Country)

- (Tokyo, Japan)
- (Paris, France)
- (Hong Kong, Spain) ?

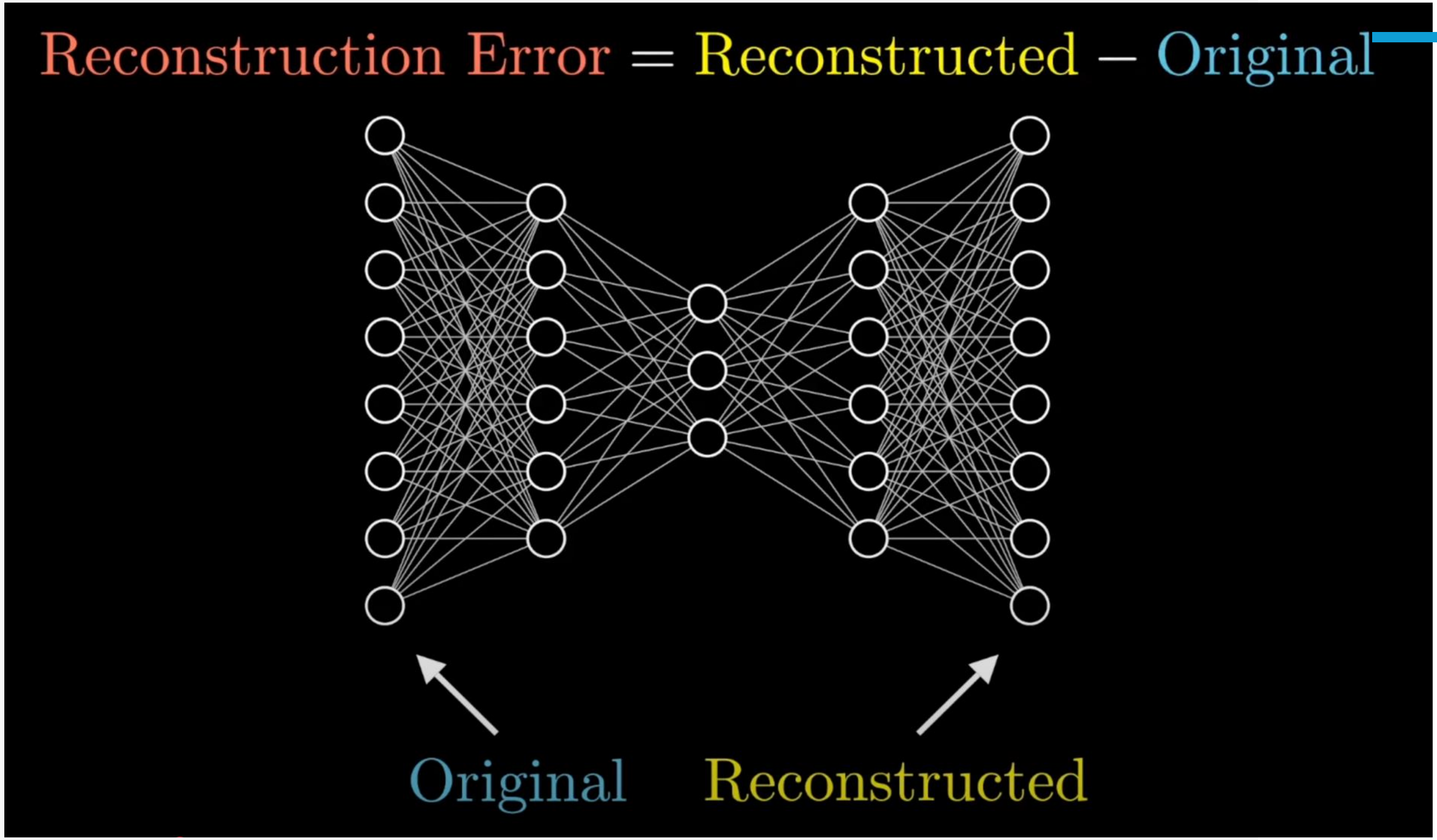


Sanjayae@bu.edu

Decoder Network

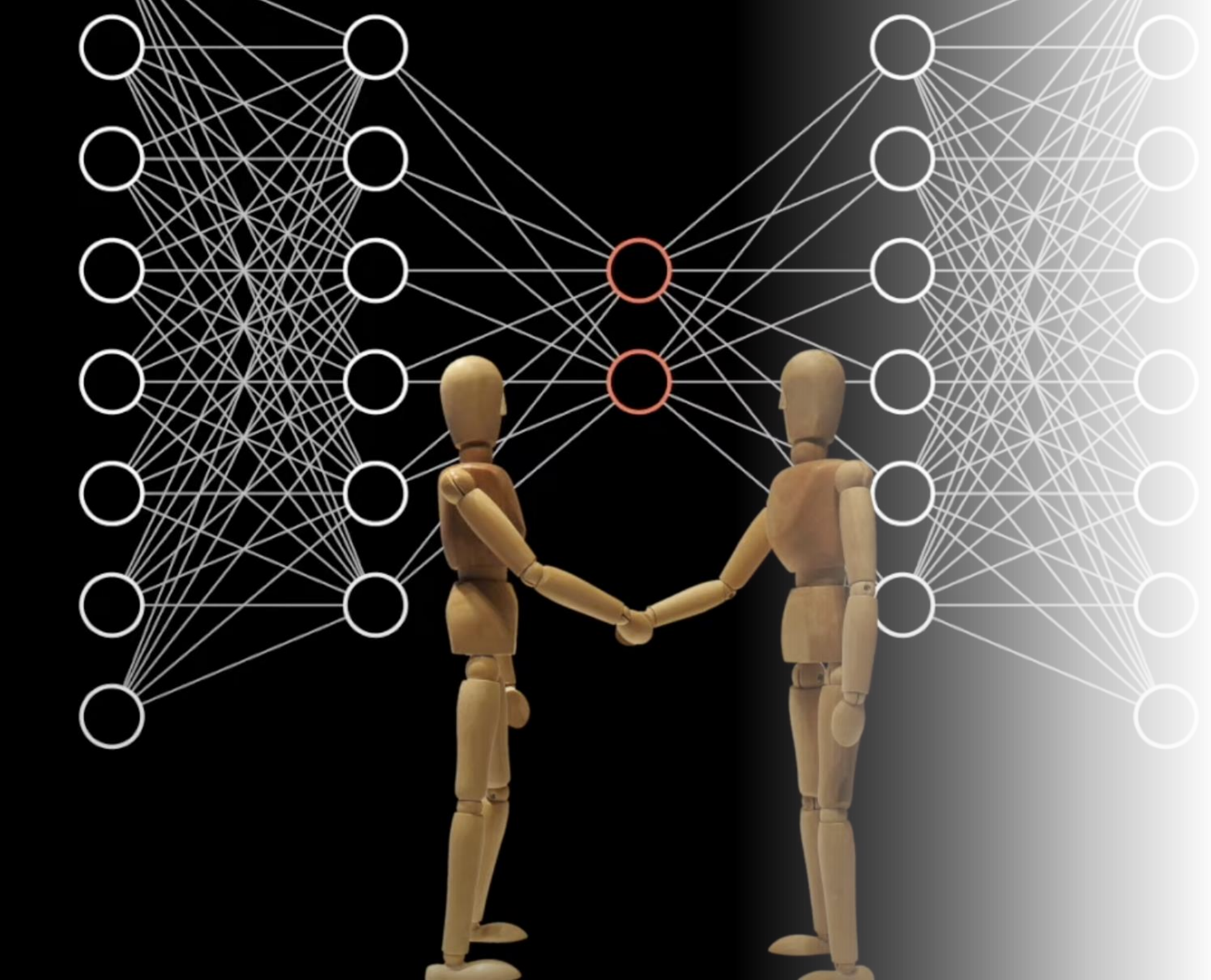


Sanjayae@bu.edu



Autoencoders have information losses

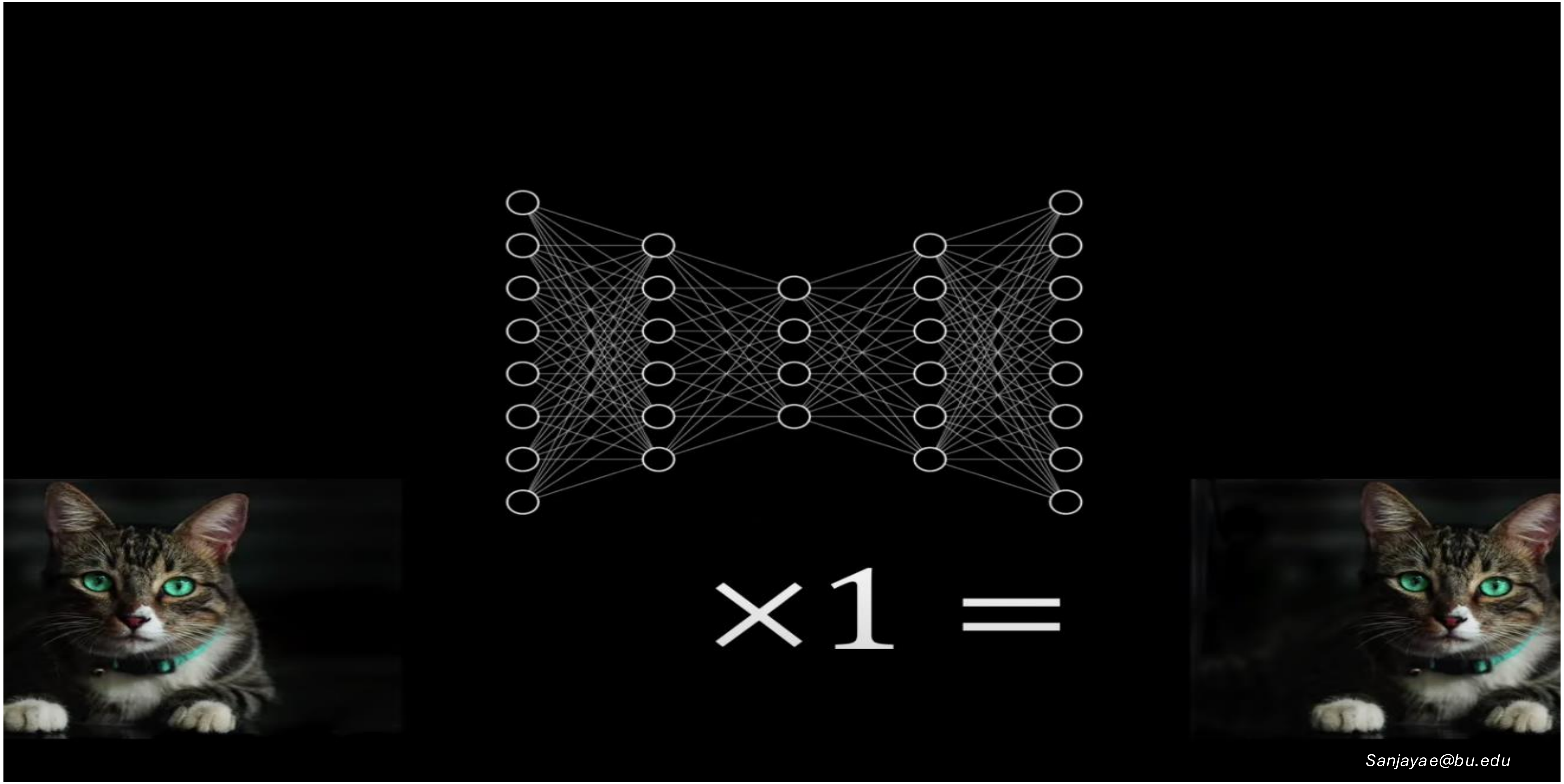
Sanjayae@bu.edu



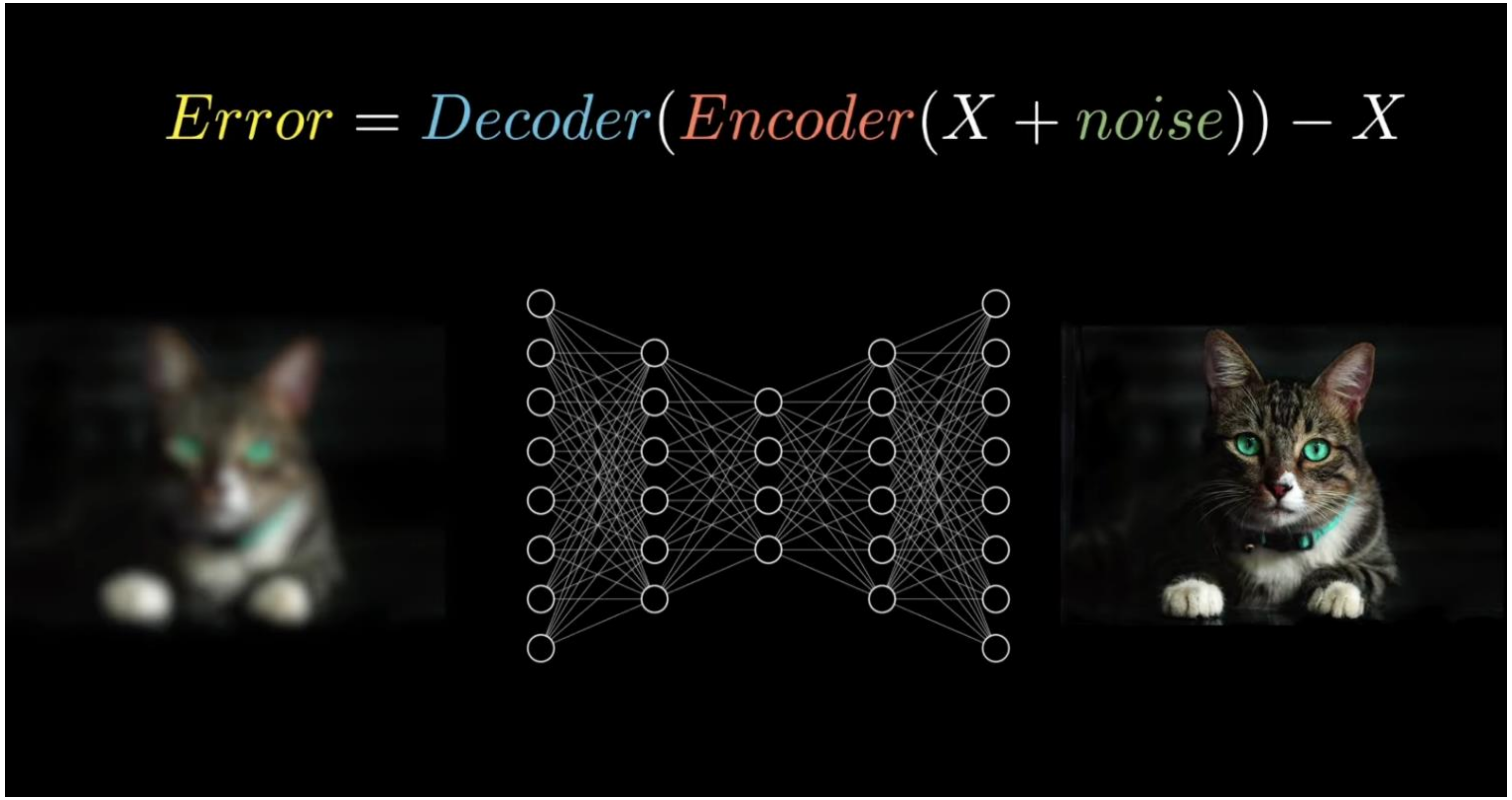
Encoder and Decoder works together to minimize the information loss

Sanjayae@bu.edu

Standard (simple) autoencoders are X 1pass through system and cannot generate new data



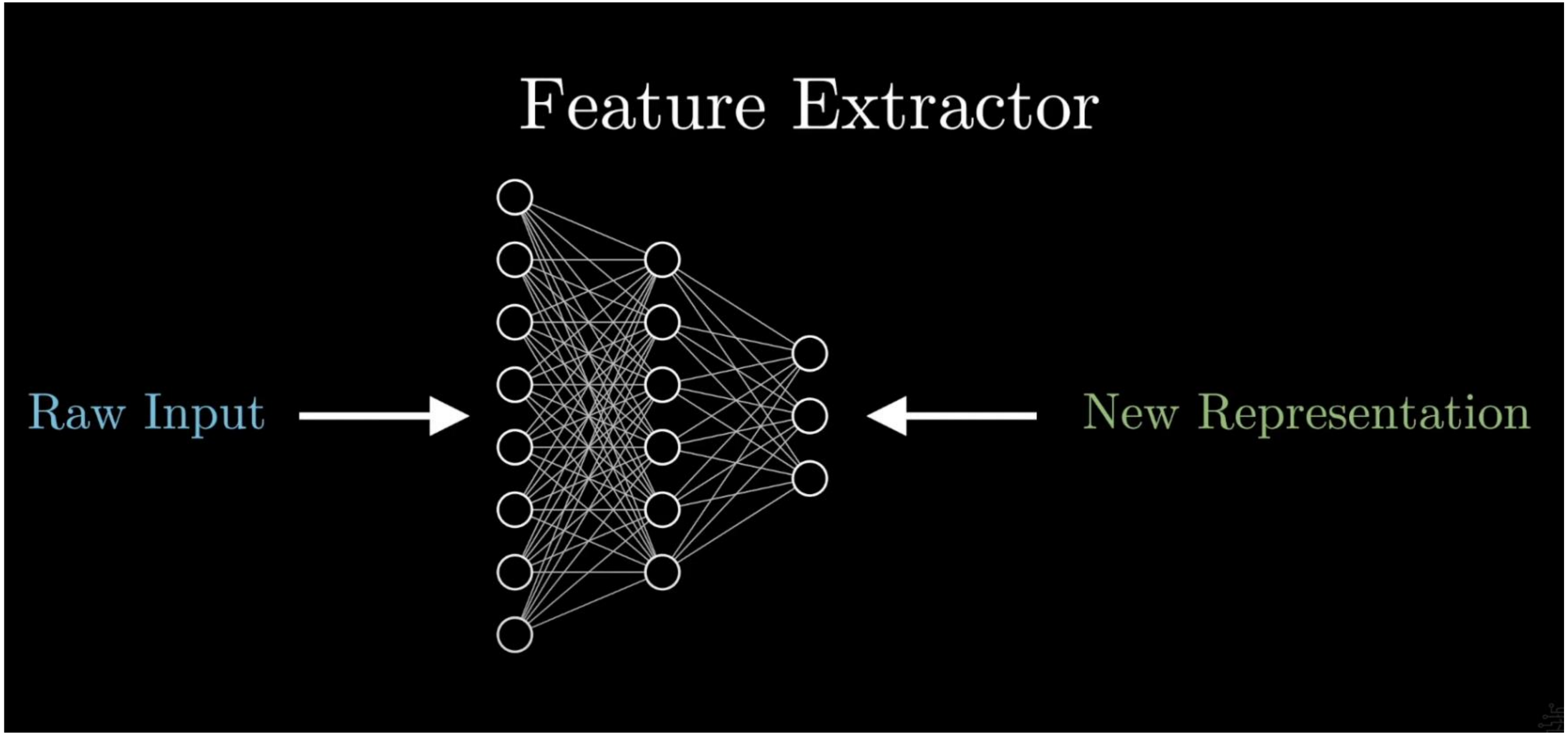
Applications of Autoencoders

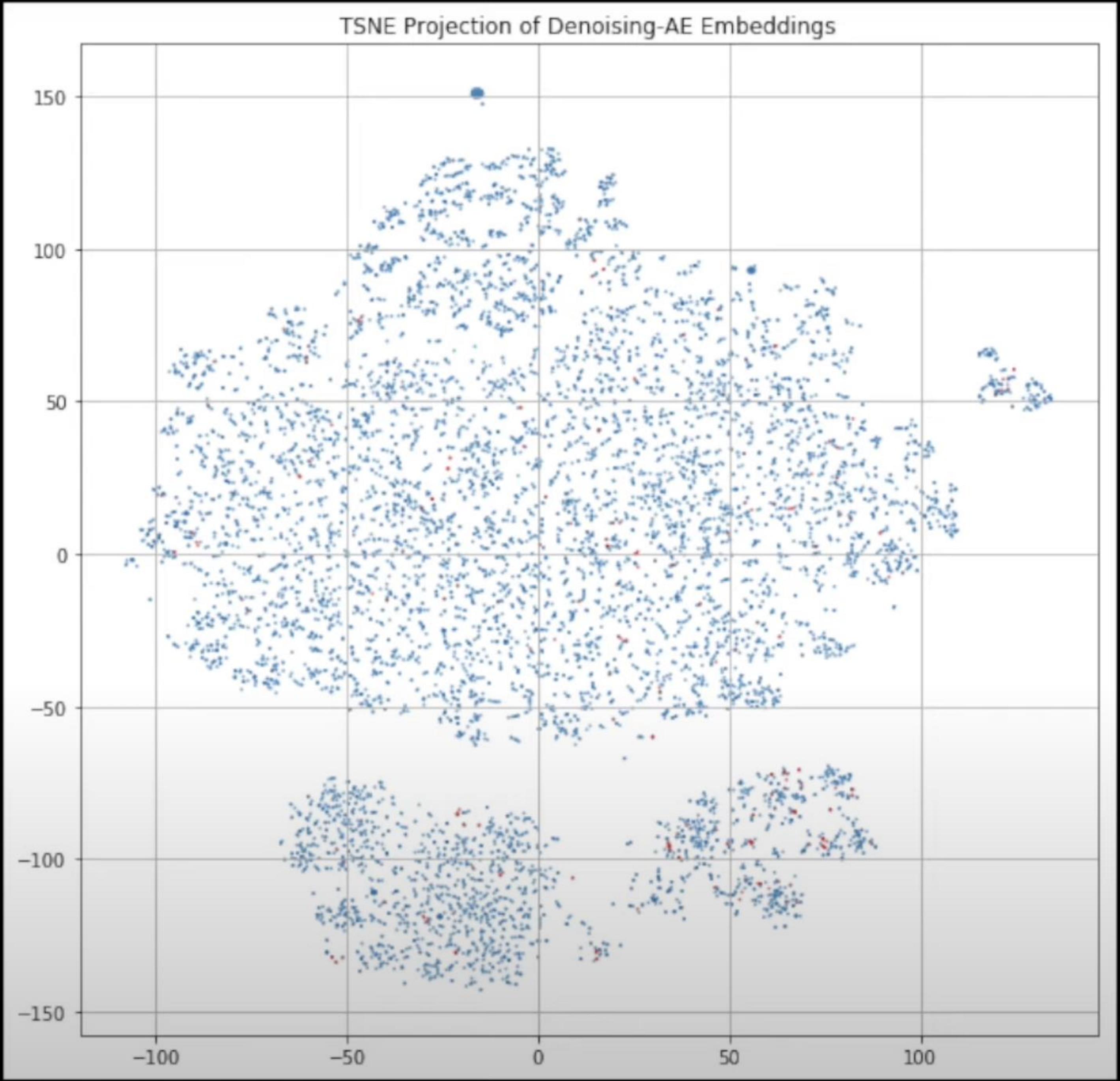


Denoising Autoencoders

- Trained to reconstruct the original input from a **corrupted** version.
- Learns robust features by handling noisy data.
- **Use case**: Noise reduction, data denoising.

Applications of Autoencoders

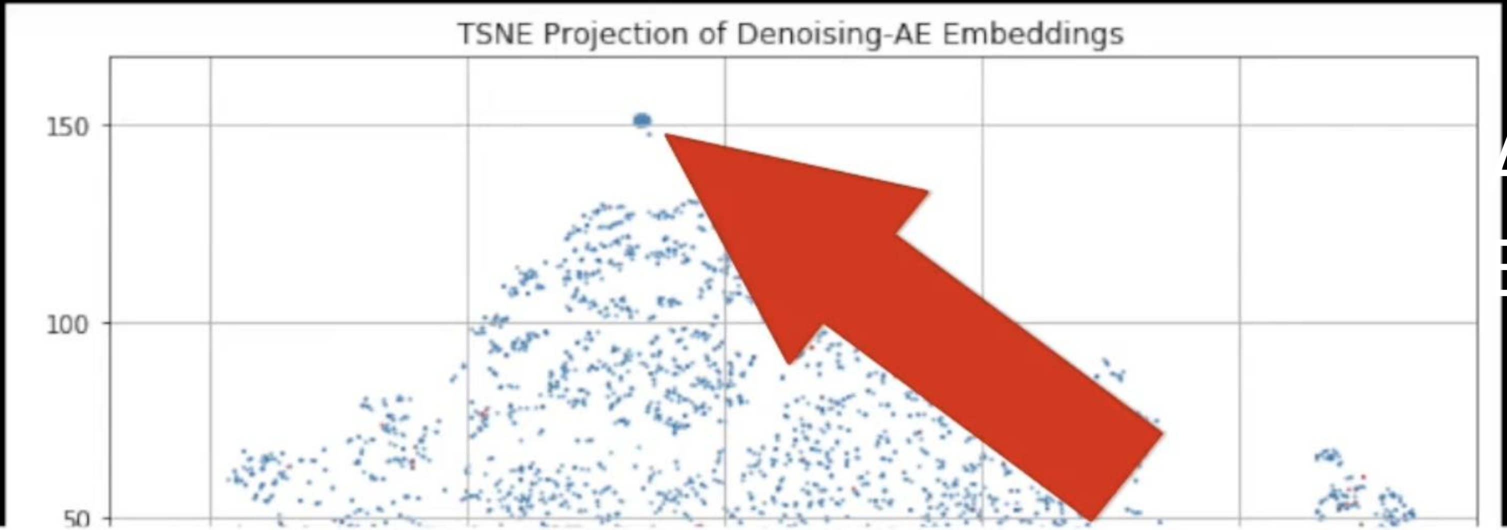




Applications of Autoencoders

Fraud Detection Systems

Sanjayae@bu.edu



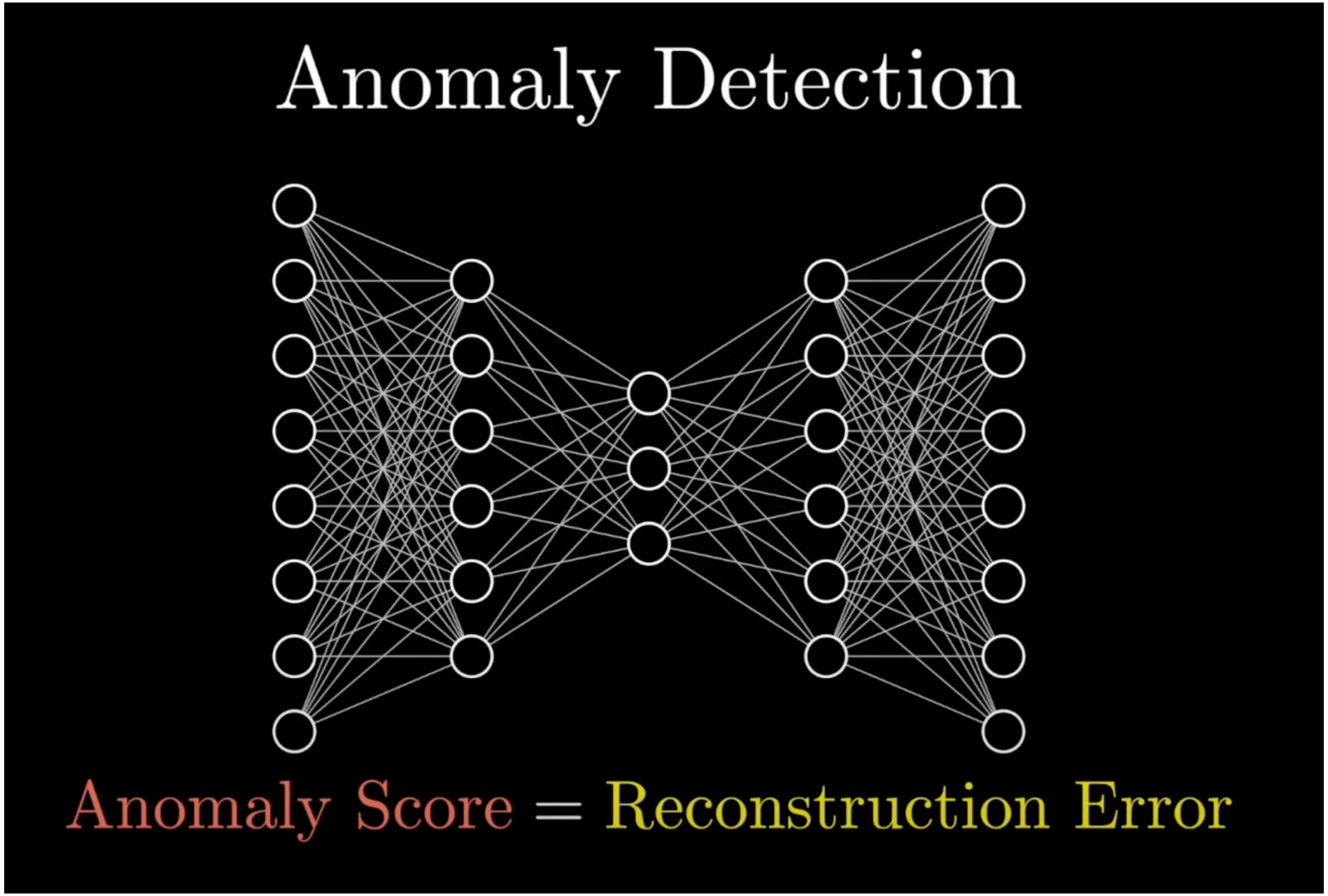
Applications of Auto Encoders

Fraud Detection

	ProductCD	card1	card2	card3	card4	card5	card6	addr1	addr2	P_emaildomain	R_emaildomain	M1	M2	M3
940	W	9323	111.0	150.0	visa	226.0	debit	191.0	87.0	charter.net	<UNK>	<UNK>	<UNK>	<UNK>
977	W	9323	111.0	150.0	visa	226.0	debit	191.0	87.0	charter.net	<UNK>	<UNK>	<UNK>	<UNK>
999	W	9323	111.0	150.0	visa	226.0	debit	191.0	87.0	charter.net	<UNK>	<UNK>	<UNK>	<UNK>
059	W	9323	111.0	150.0	visa	226.0	debit	191.0	87.0	charter.net	<UNK>	<UNK>	<UNK>	<UNK>
114	W	9323	111.0	150.0	visa	226.0	debit	191.0	87.0	charter.net	<UNK>	<UNK>	<UNK>	<UNK>
385	W	9323	111.0	150.0	visa	226.0	debit	191.0	87.0	charter.net	<UNK>	<UNK>	<UNK>	<UNK>
405	W	9323	111.0	150.0	visa	226.0	debit	191.0	87.0	charter.net	<UNK>	<UNK>	<UNK>	<UNK>

Sanjayae@bu.edu

Applications of Autoencoders

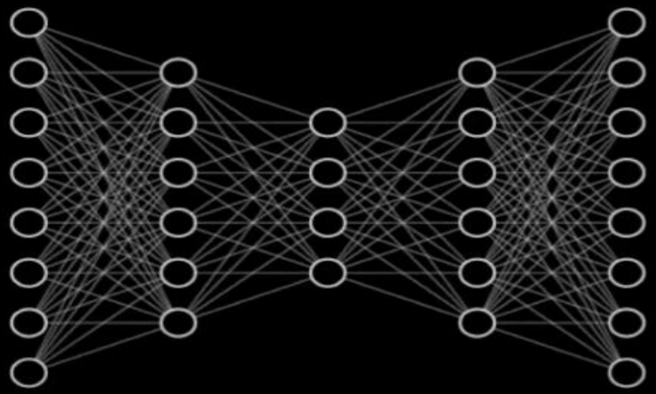


Applications of Autoencoders

Missing value recreation

RU	-15.0	yandex.ru
US	8.3	gmail.com
IN	11.2	rediff.com
RU	<MISSING>	yandex.ru

<MISSING>	-15.0	yandex.ru
US	<MISSING>	gmail.com
IN	11.2	<MISSING>



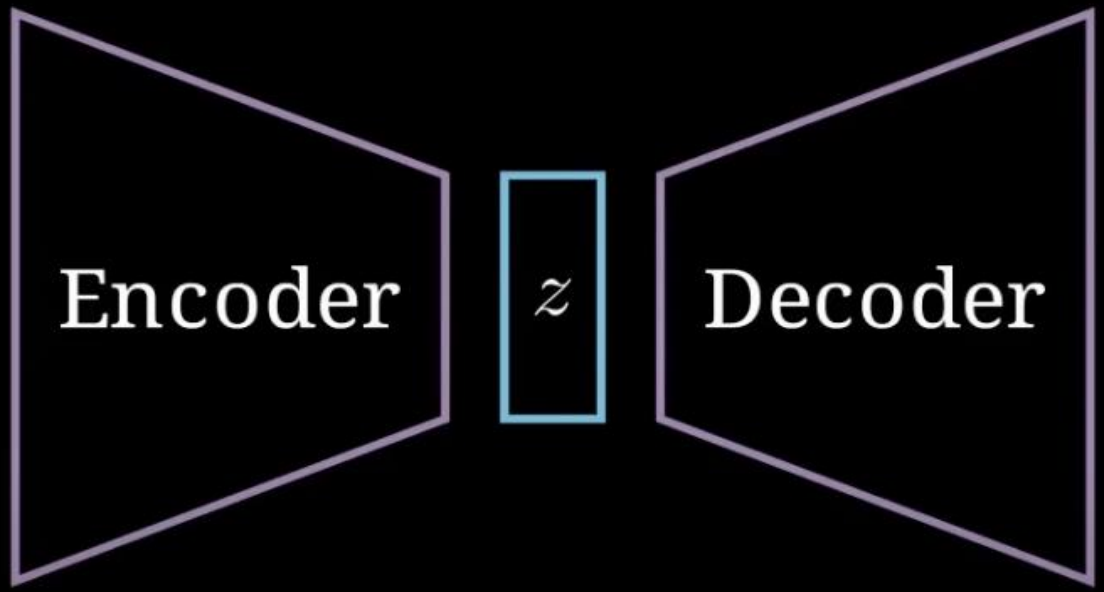
RU	-15.0	yandex.ru
US	8.3	gmail.com
IN	11.2	rediff.com

Sanjayae@bu.edu

Variational Autoencoder (VAE)

Sanjayae@bu.edu

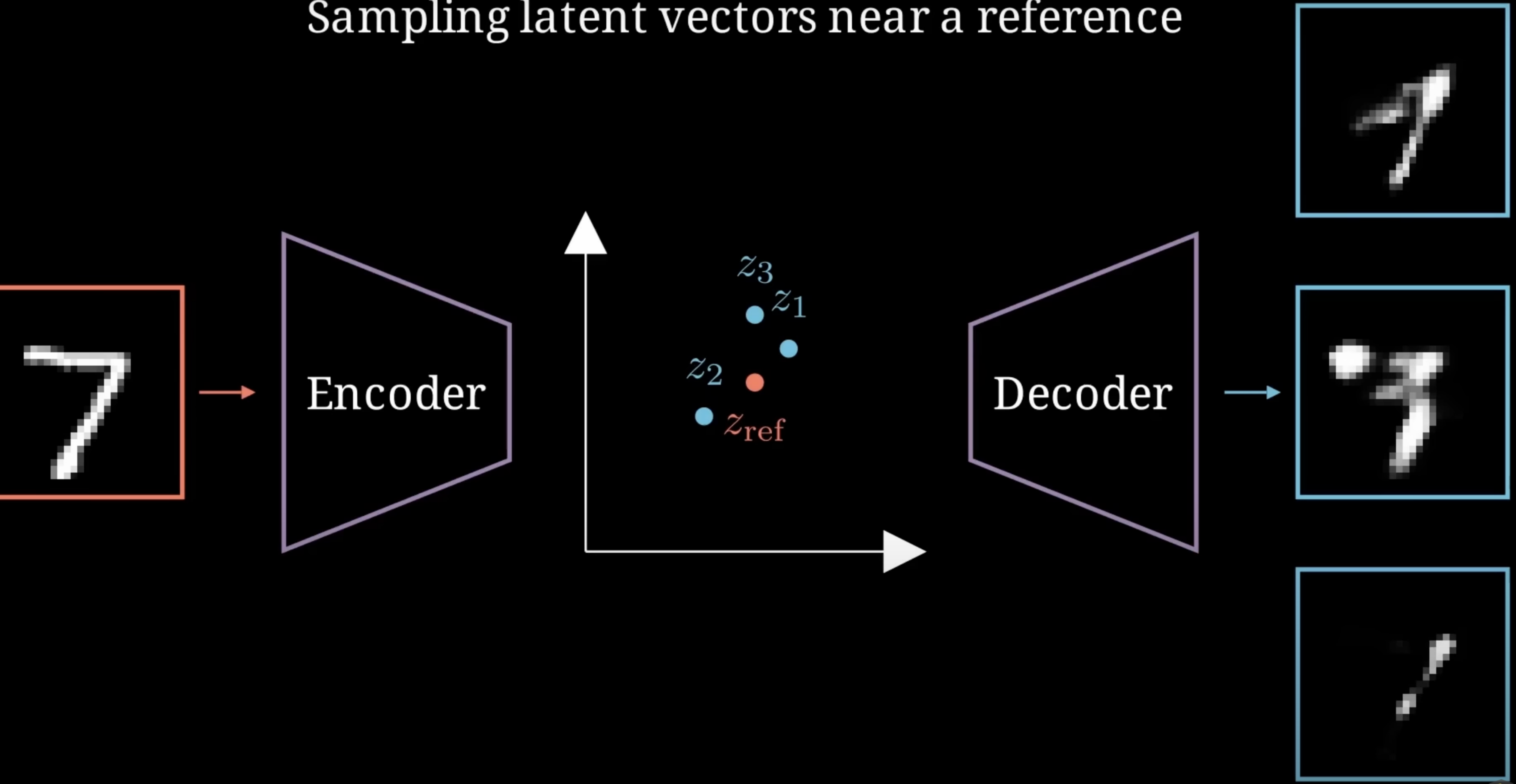
Why Variational Autoencoder (VAE) ?



Standard Autoencoders can't generate new data :(

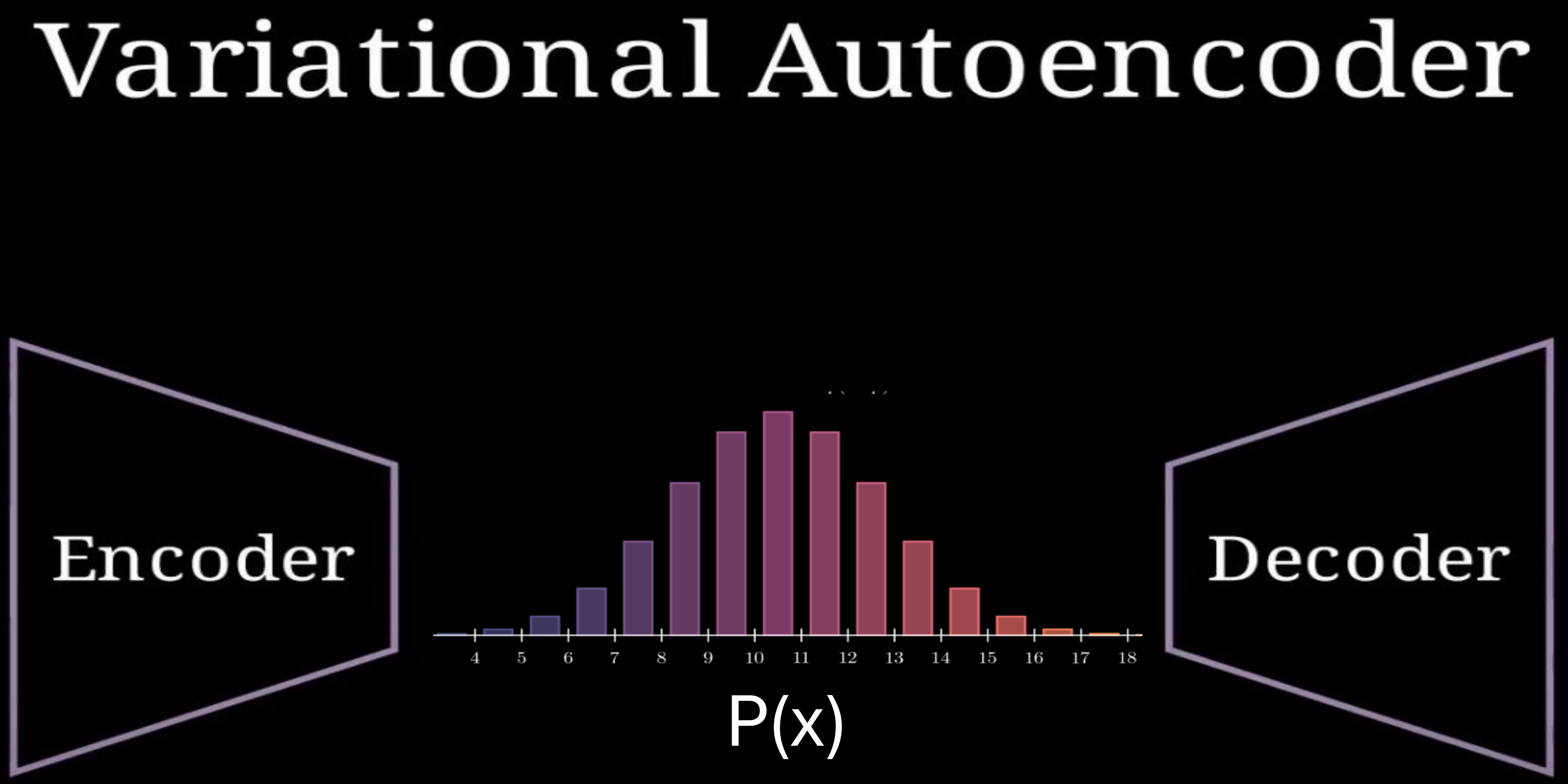
Sanjayae@bu.edu

Sampling latent vectors near a reference

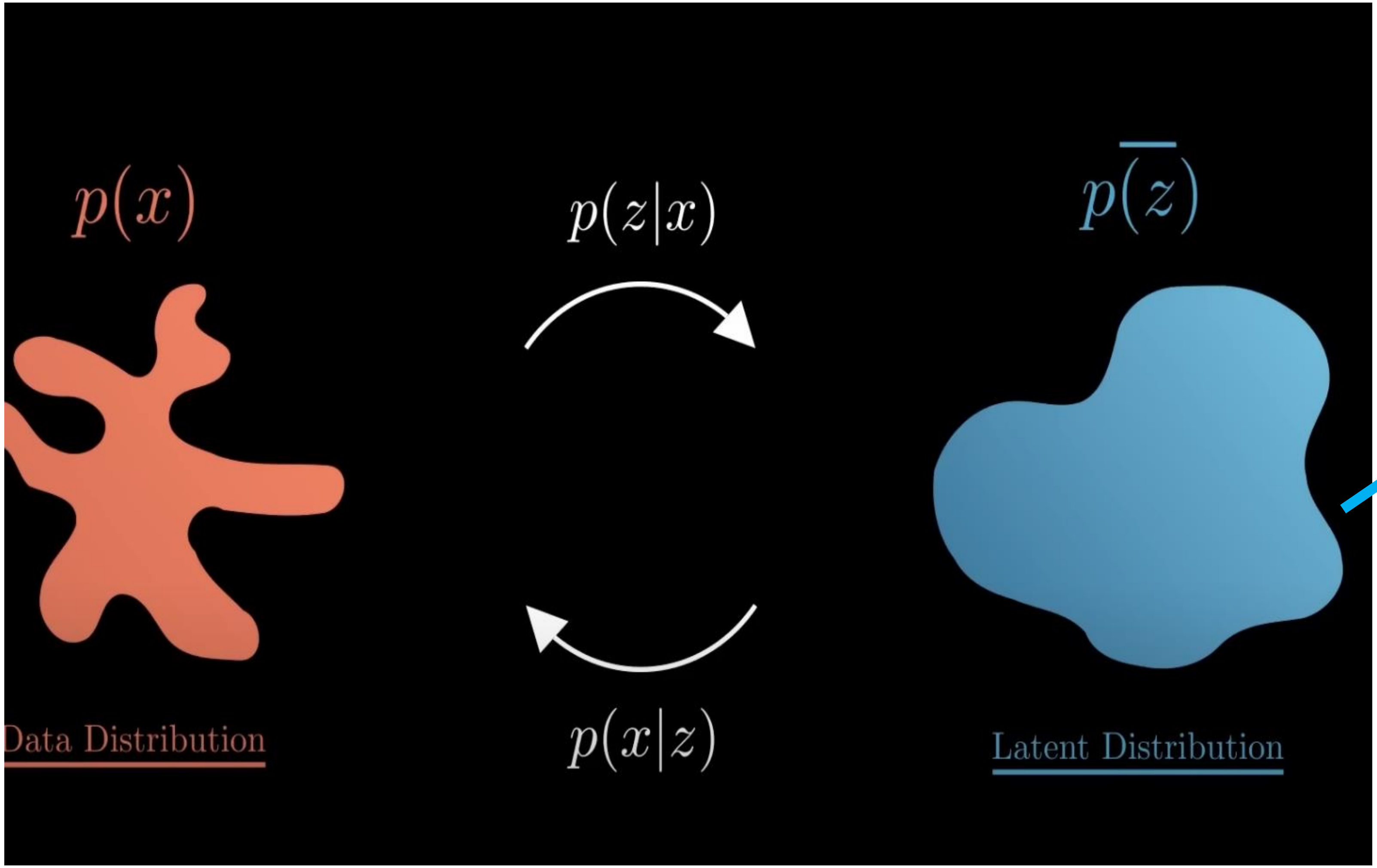


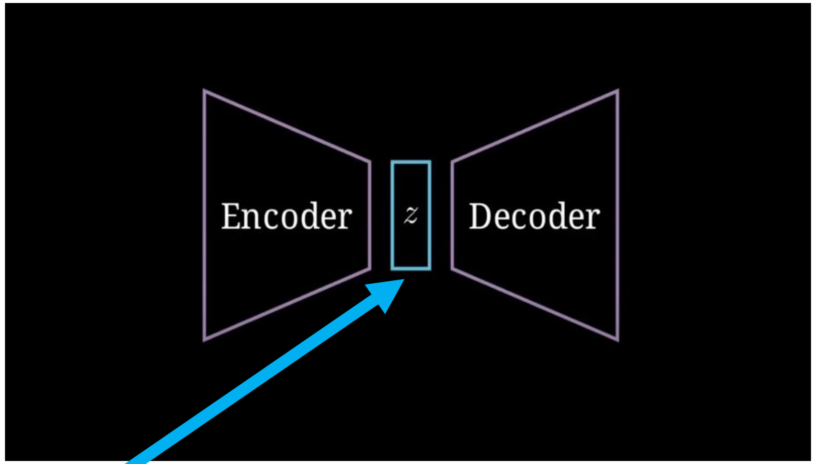
Sanjayae@bu.edu

Variational Autoencoder



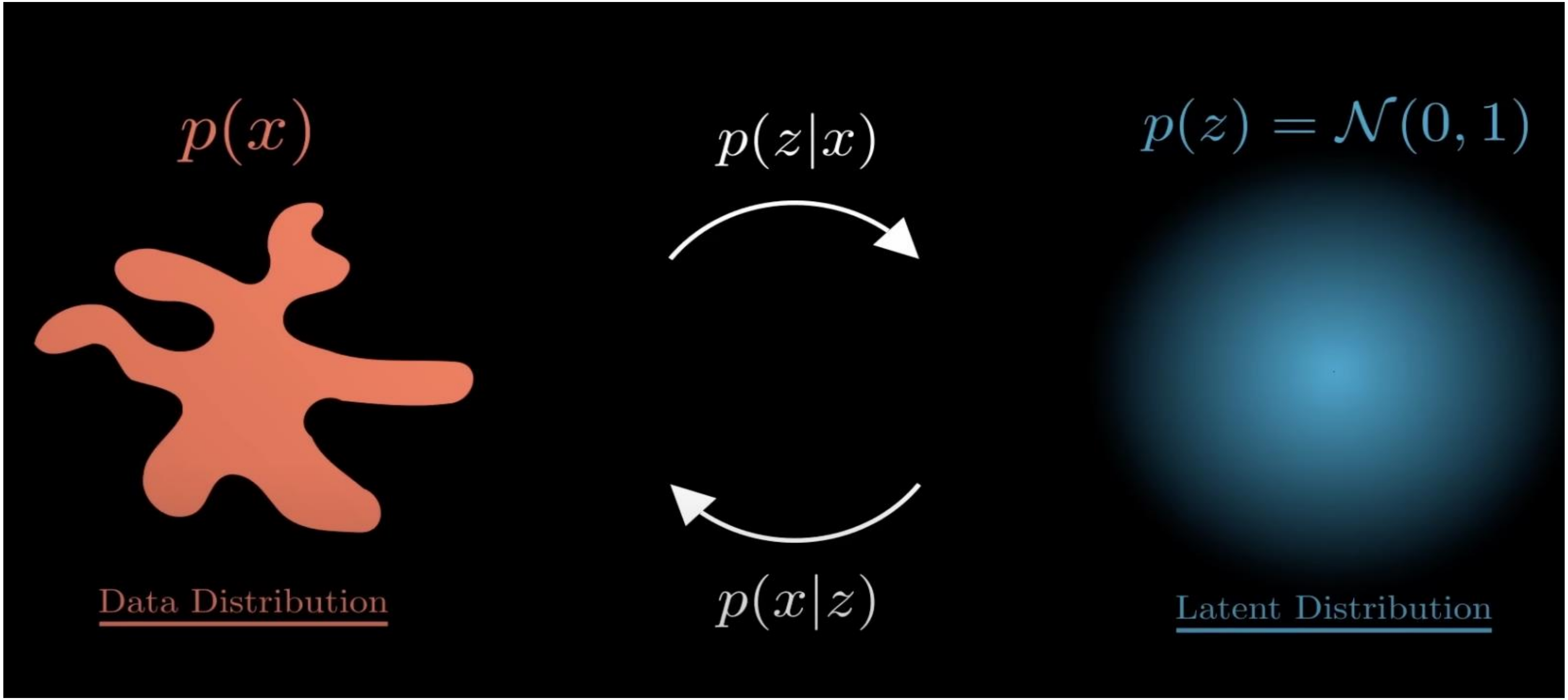
we know the x, so we can calculate the p(z/x) and if we know the P(x/z) . We can recreate the image .. but ...



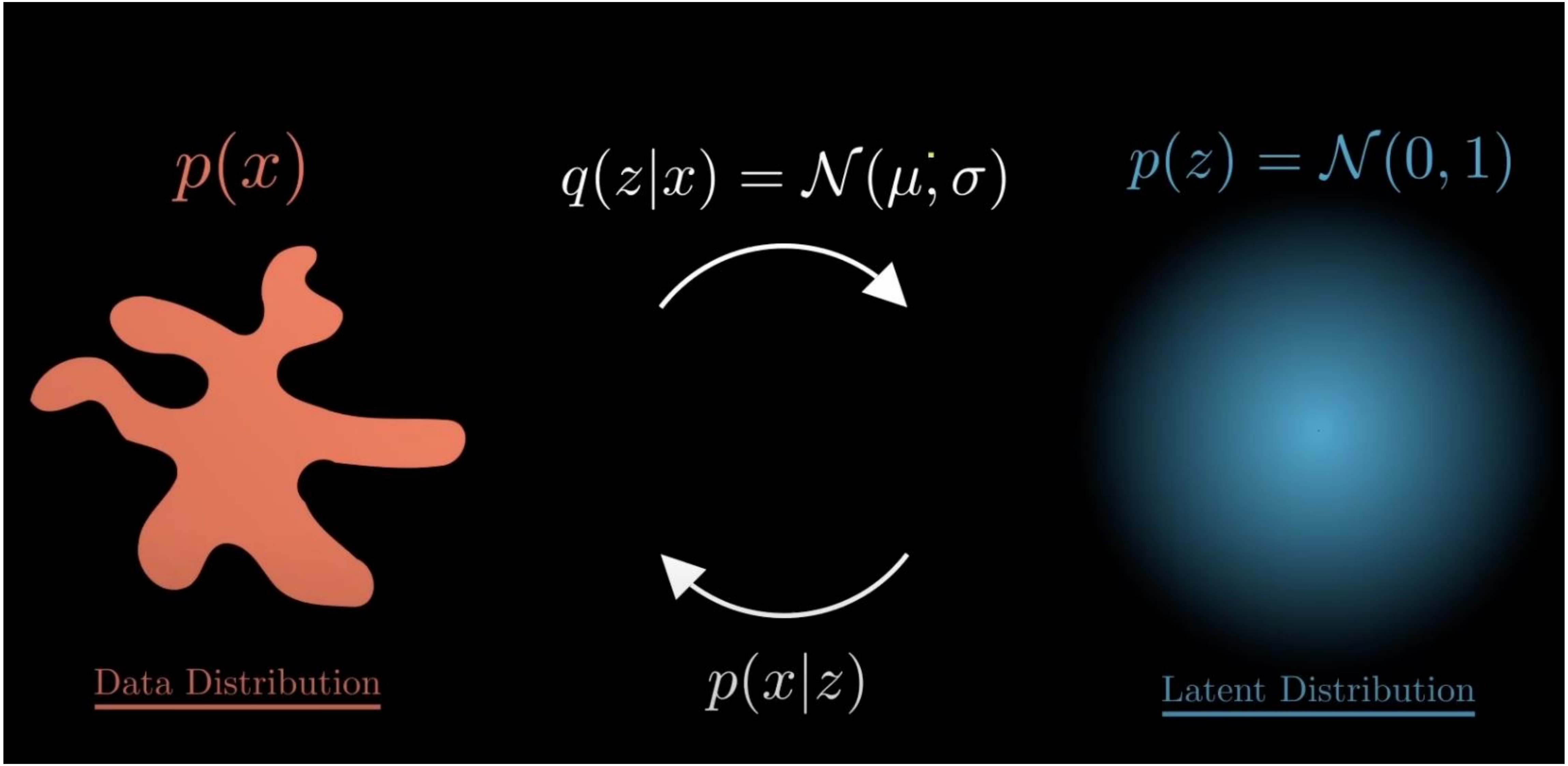


How do we know the probability distribution of P(z) ???/

We can assume that p(z) will have a Normal (Gaussian) distribution

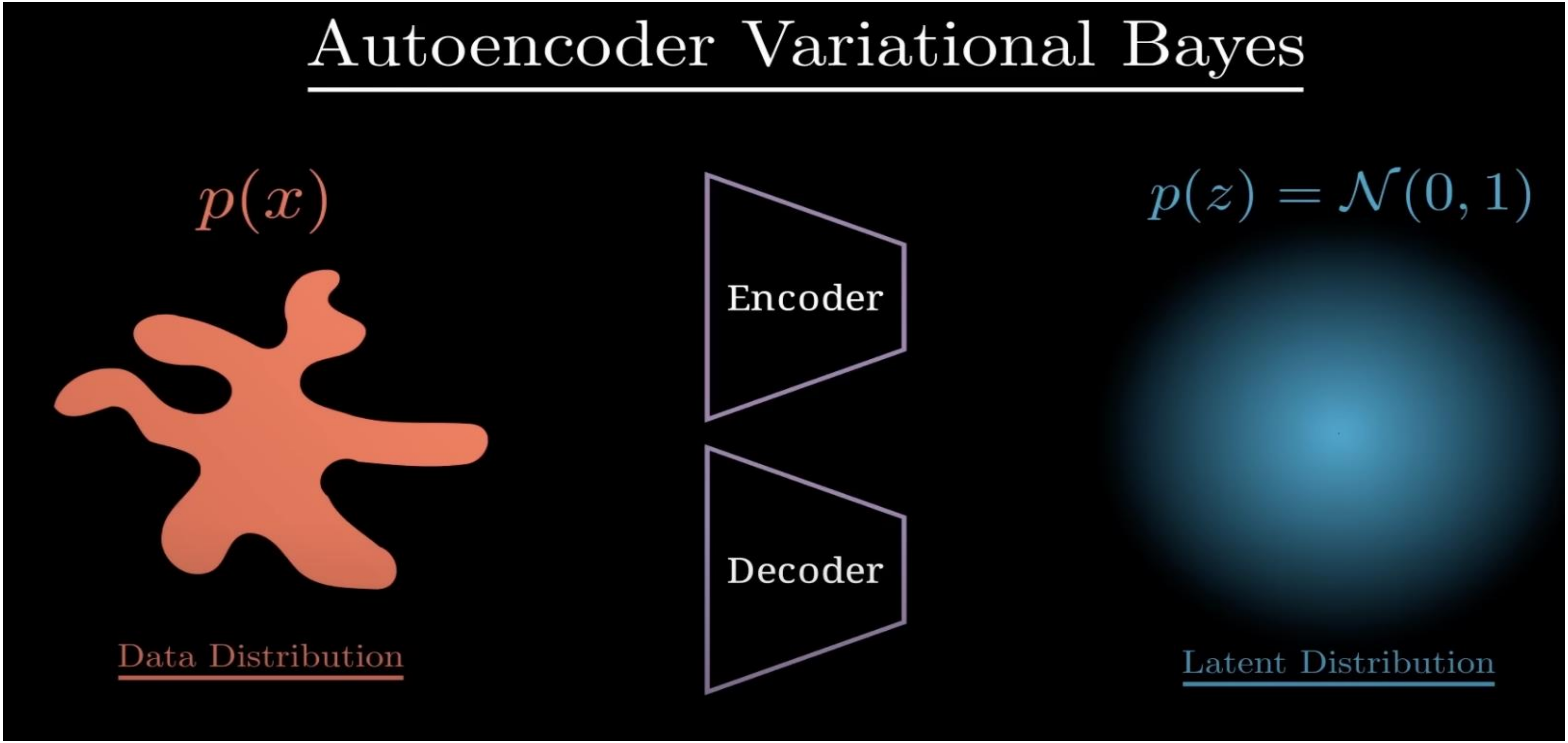


We can use the medium and standard deviation of X distribution to create P(z) using training parameter tuning....



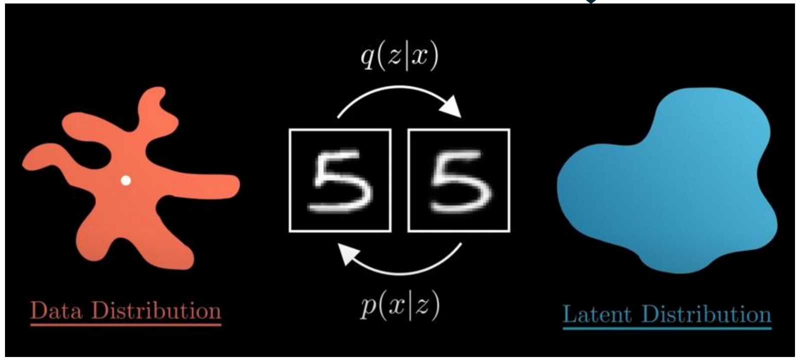
Sanjayae@bu.edu

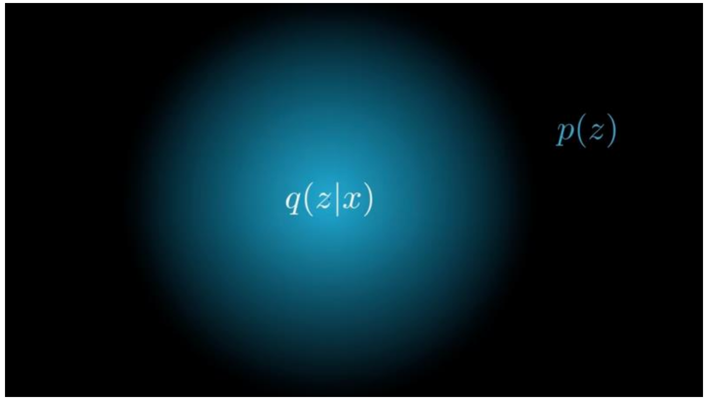
We can use encoder and decoder to train the autoencoder end to end and fine tune the weights (parameters)



We can represent this process and calculate the loss funtion with the following equetion ..

$$\mathcal{L}(x) = \underbrace{\mathbb{E}_{q(z|x)}\left[\log p(x|z)\right]}_{\text{Data consistency}} - \overbrace{\mathrm{KL}(q(z|x) \mid p(z))}^{\text{Regularization}}$$



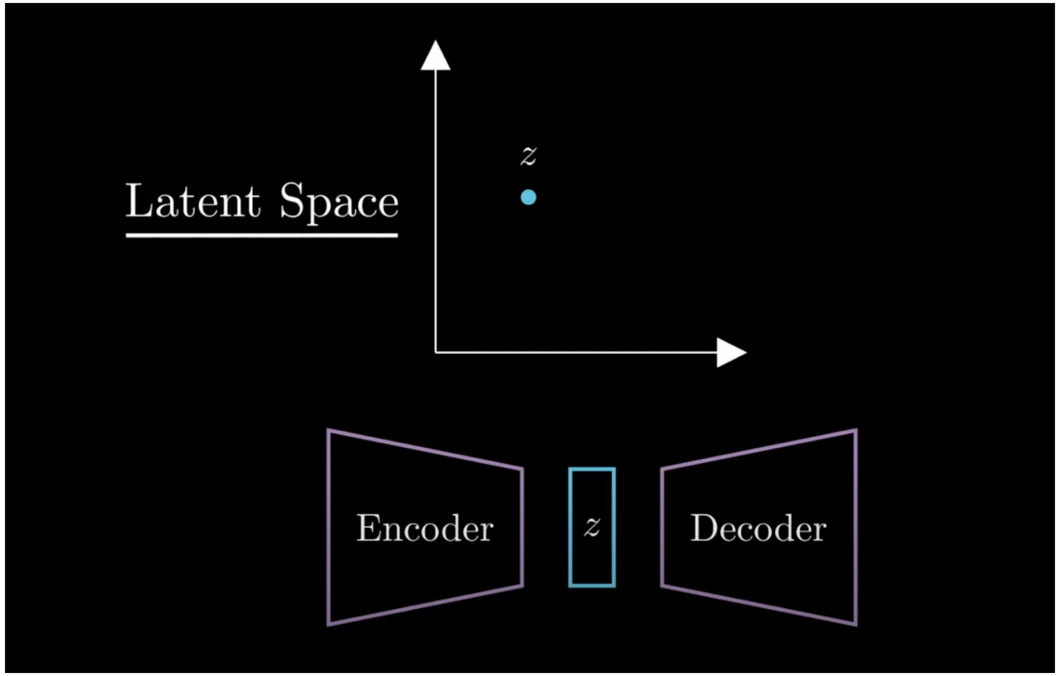


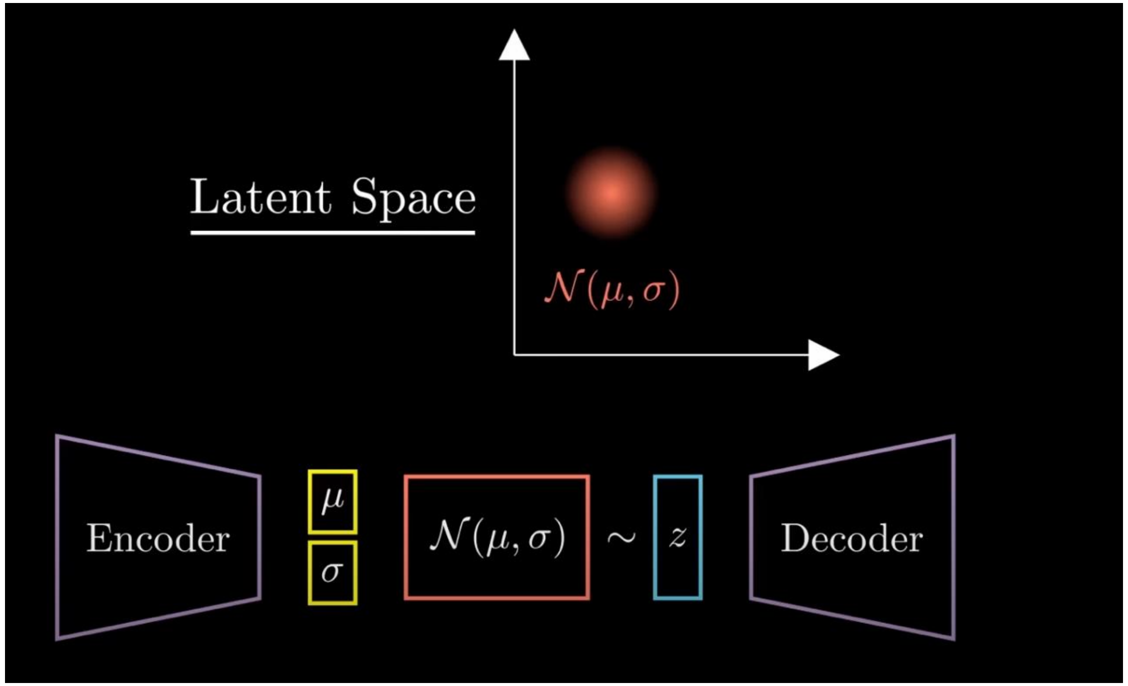
L2 Difference between two dimensional spaces

Measure how close the two probability distribution

Let's look at how Variational Autoencoders practically works ?

How this works in practically ?

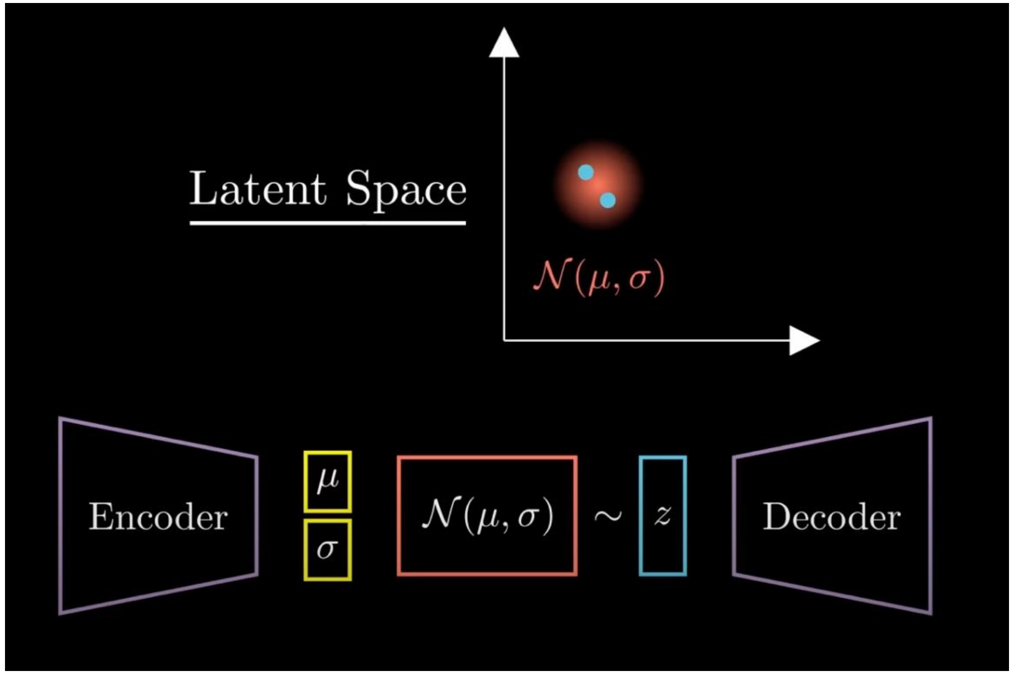


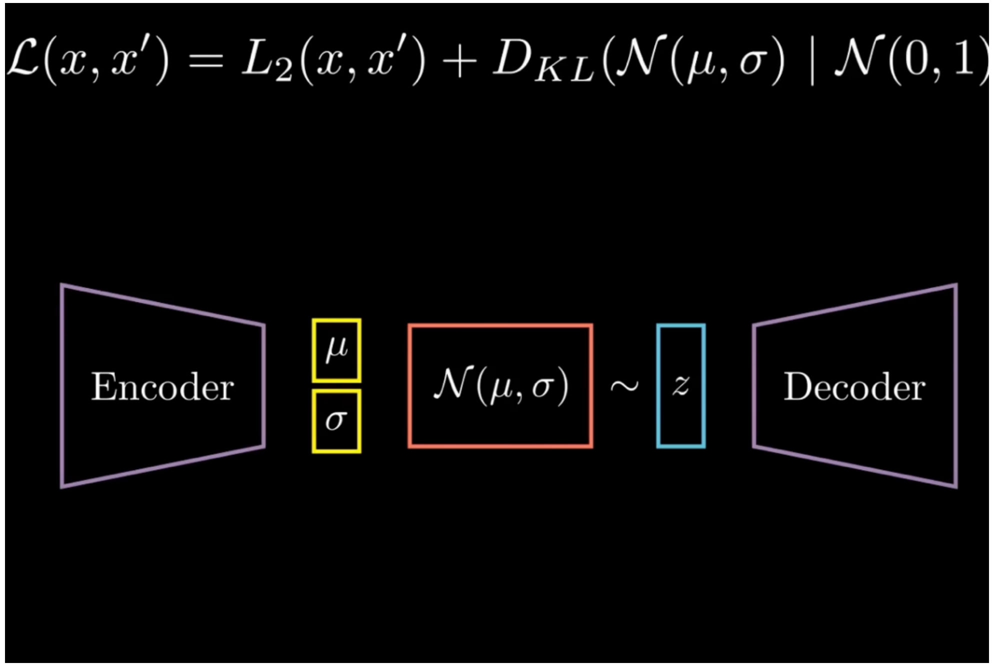


In a **standard autoencoder** an input (e.g. image) is a 1 point in the latent space

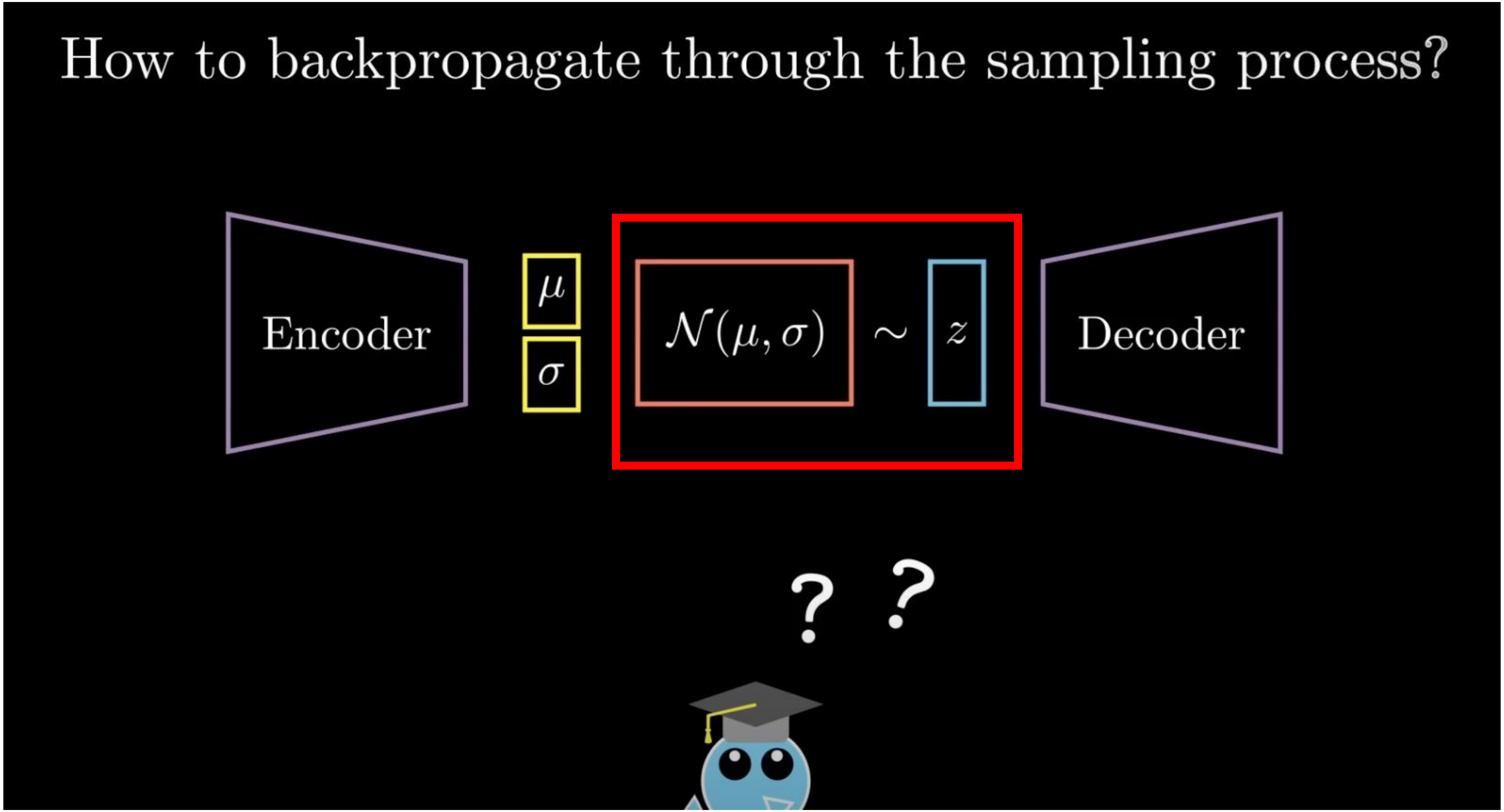
In a **Variational Autoencoder** an input (e.g. image) represented as a Gaussian distribution with parameters defined.

How this works in practically ?





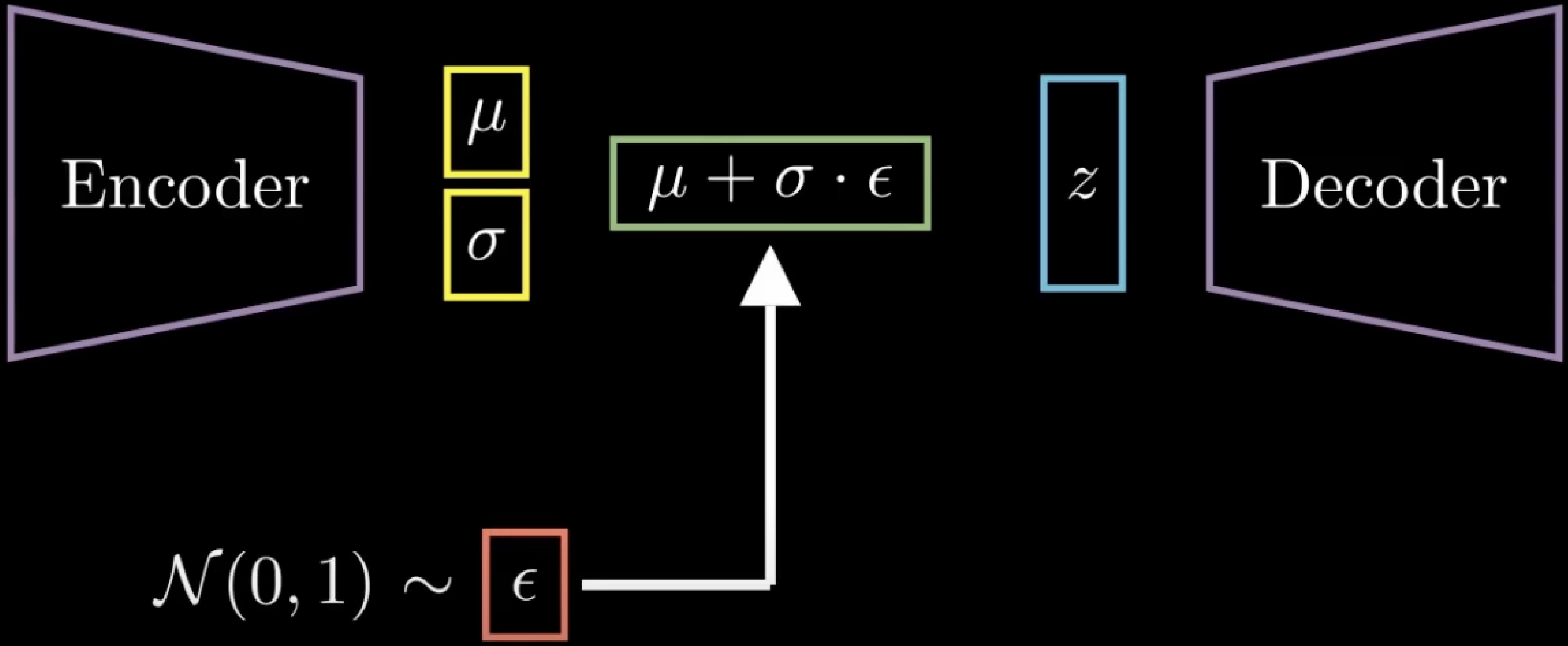
Select two samples (x , x') from the latent space and calcite the loss and backpropagate to the input space



We can't backpropagate through a distribution

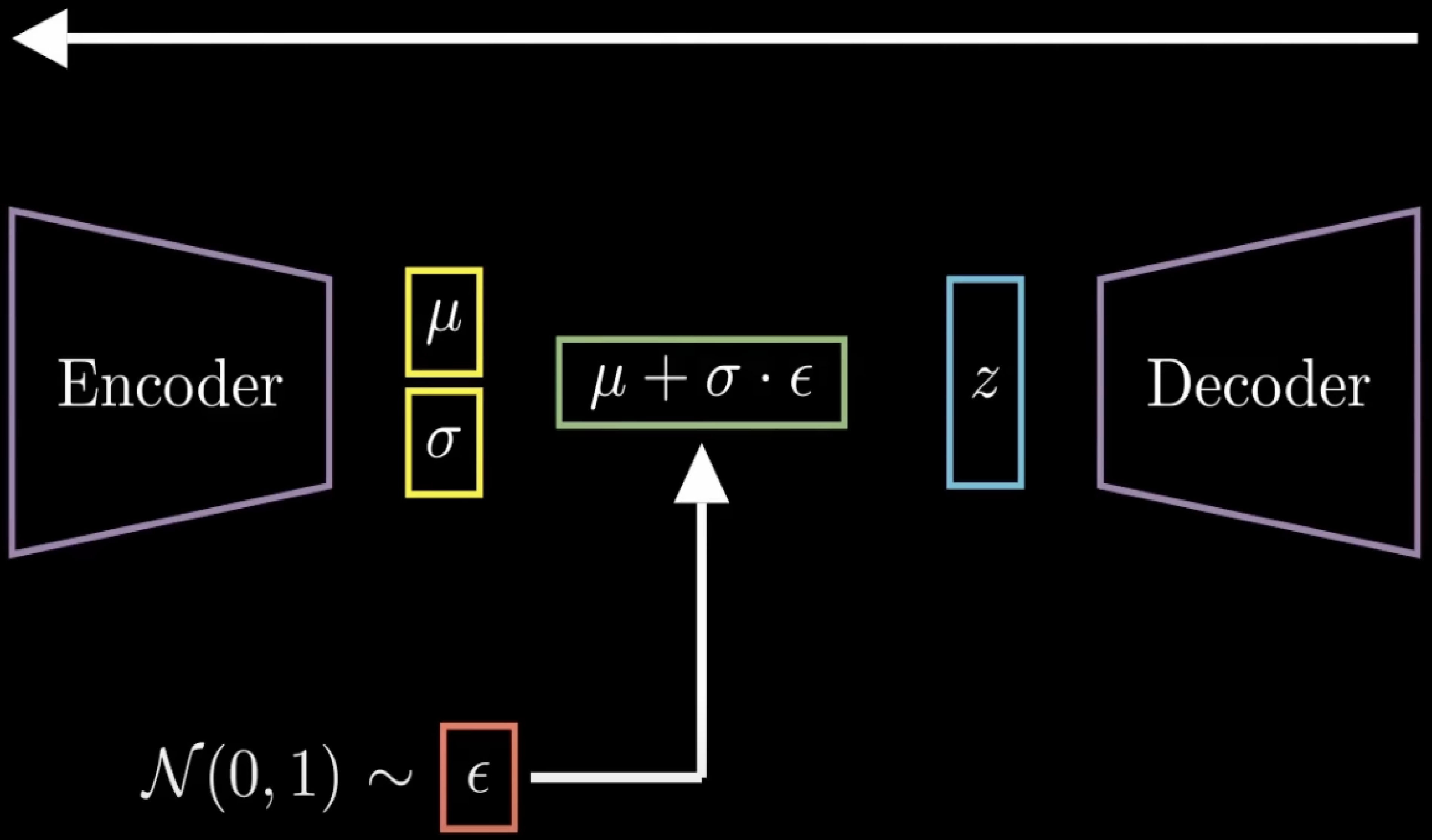
Sanjayae@bu.edu

Reparameterization Trick

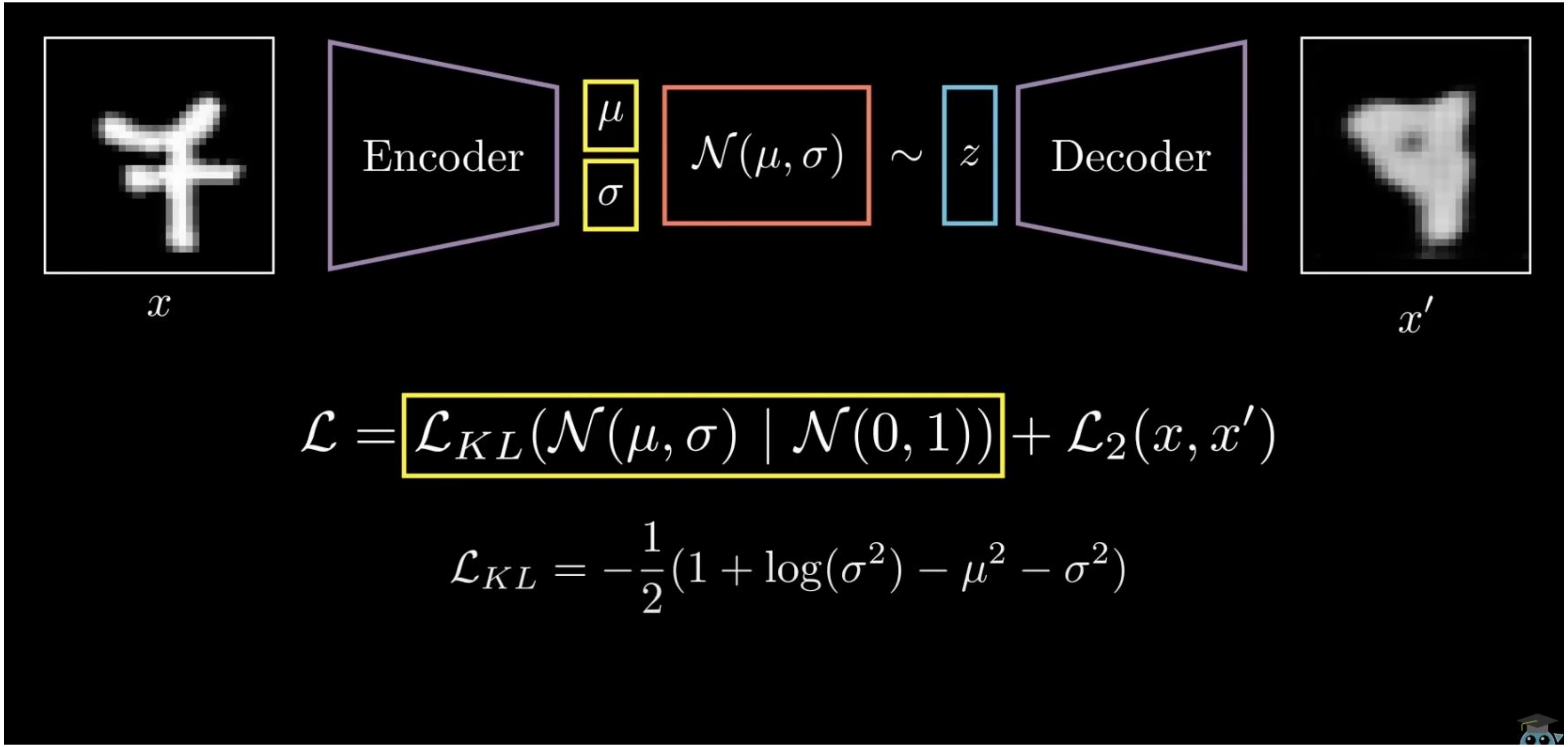


Sanjayae@bu.edu

Backpropagation

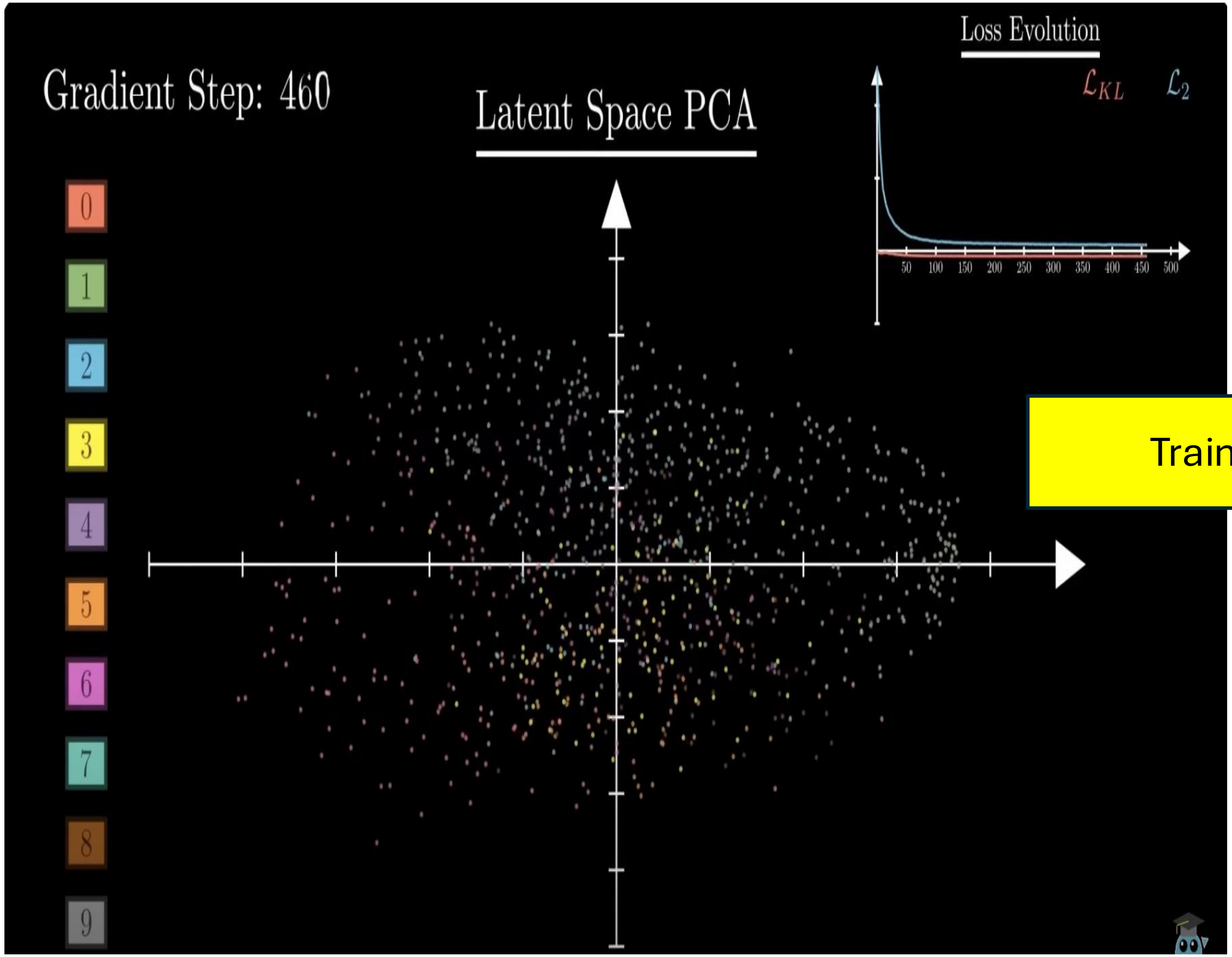


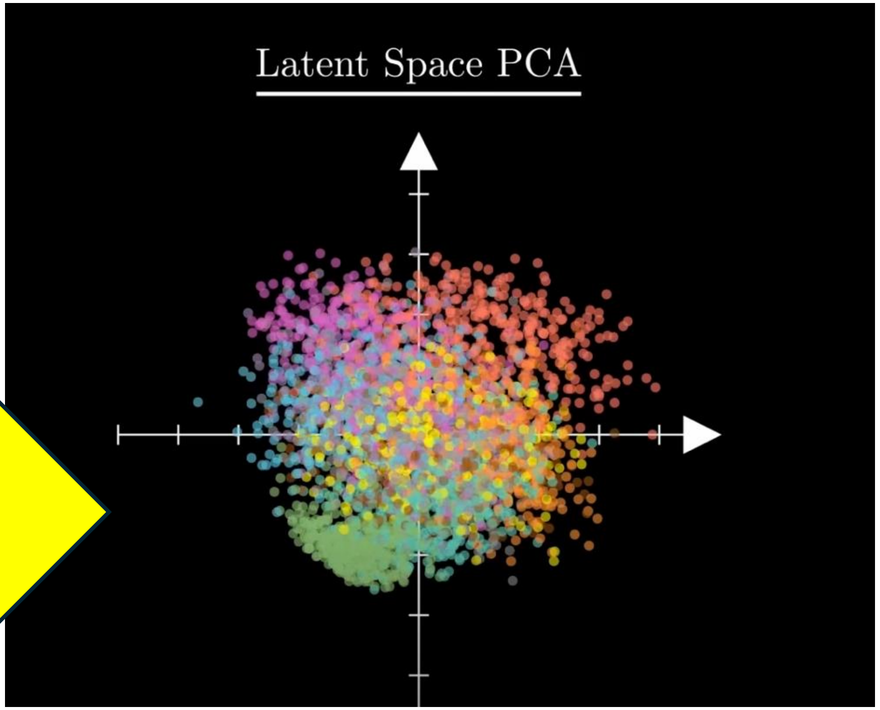
Sanjayae@bu.edu



Now we can train the VAE similar to a standard AE , and finetune the parameters for each image (input) of the data set.

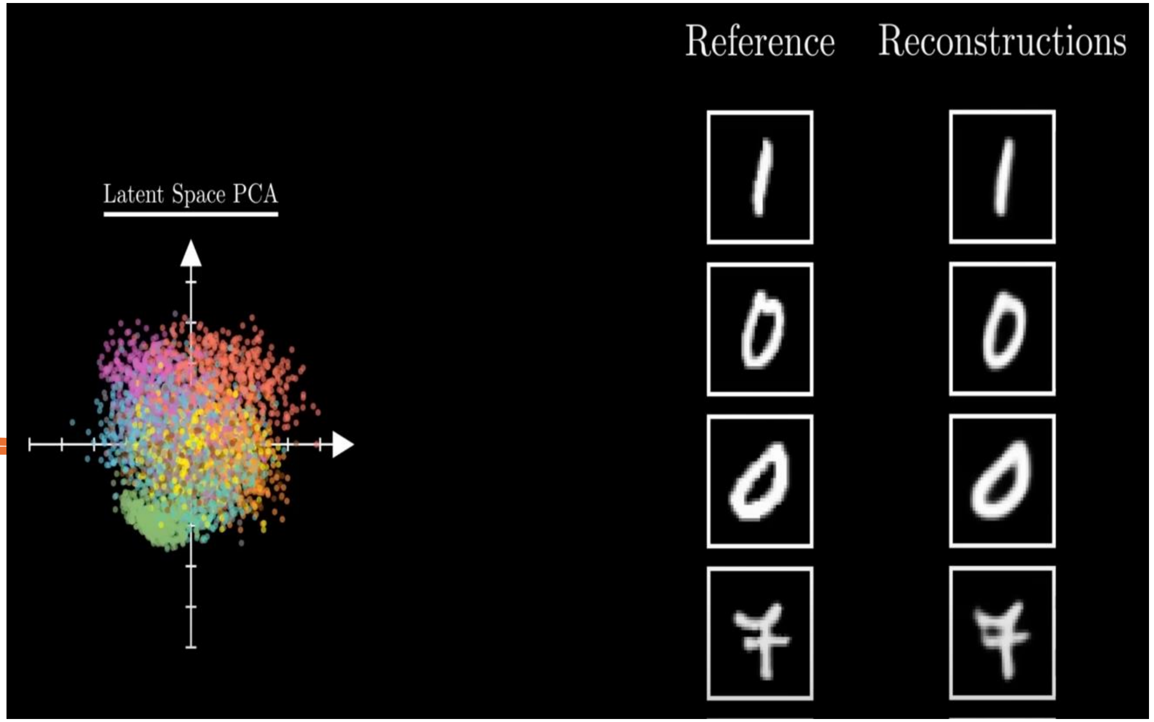
Sanjayae@bu.edu





Initially the VAE latent space is more spread (high variance) and distributed a lot.

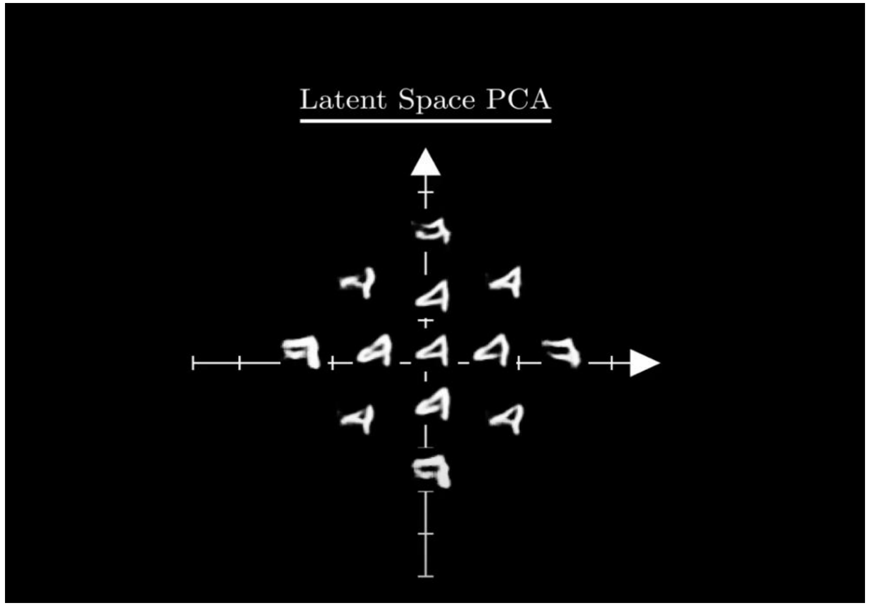
Once the training is completed, VAE latent space is compact, high dense vector embeddings with proximity to each other. This means, most of the points in latent space is present a real data in the input space. Additionally this will enable generating new data

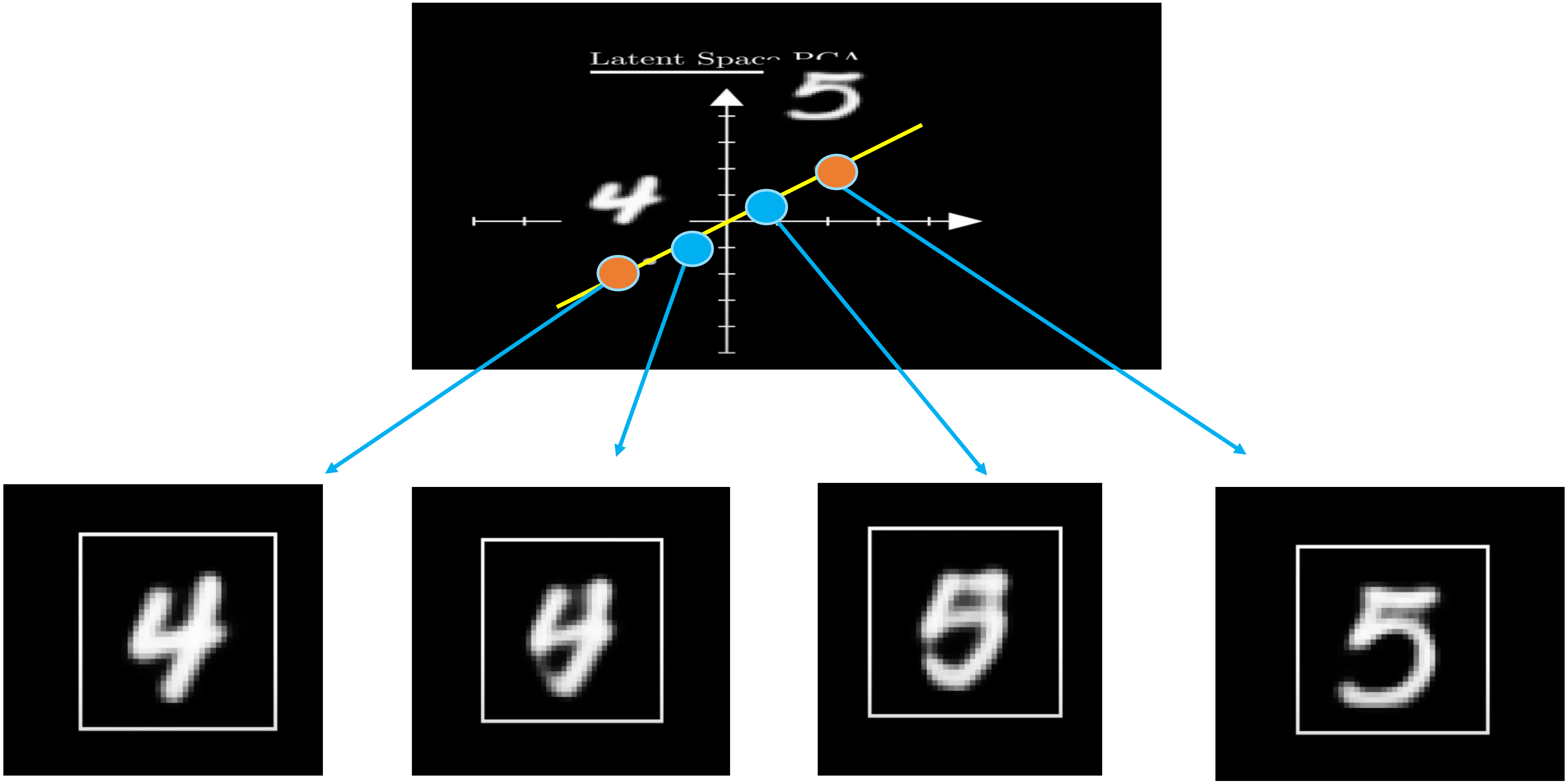


- Most of the points in latent space is present a real data in the input space. Additionally this will enable generating new data

- Very rarely we can have non-meaning less reconstructions in well trained VAE's , this will be the cases where we have hallucination and completely new creativity

When the distance increase from the original image latent space point , the image get different than its original creation.





Sanjayae@bu.edu

Limitations with VAEs ...

- They create blurry images

- How do we decide which image to generate ?

- Advanced VAE's like conditional VAE's address this with class specified image generations.

- *Sanjayae@bu.edu*

VAE Samples

Diffusion samples

GAN Samples

Sanjayae@bu.edu

Various types of Autoencoders

Autoencoder Type	Key Feature	Common Use Cases
Basic Autoencoder	Simple encoder-decoder	Dimensionality reduction
VAE	Probabilistic latent space	Generative modeling
Denoising Autoencoder	Learns to remove noise	Image & audio denoising
Sparse Autoencoder	Sparse activations	Feature learning, anomaly detection
Contractive Autoencoder	Regularized representation	Robust feature learning
Convolutional Autoencoder	Uses CNN layers	Image compression & feature extraction
Stacked Autoencoder	Deep multi-layer model	Deep feature extraction
AAE	Autoencoder + GAN	High-quality data generation
LSTM Autoencoder	LSTM-based sequence learning	Time series anomaly detection
Transformer Autoencoder	Uses transformer layers	NLP feature extraction