Генеративные модели. Часть 1

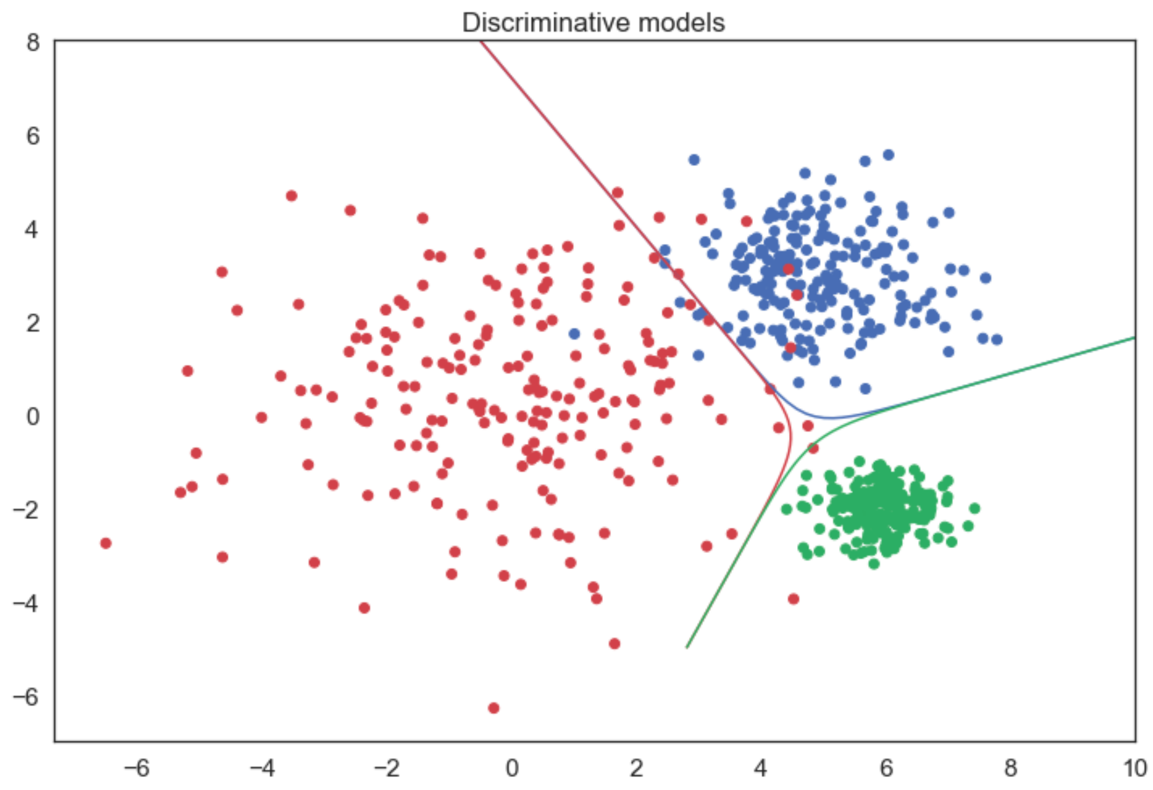
Никита Балаганский

Генерация изображений

Discriminative / Generative models

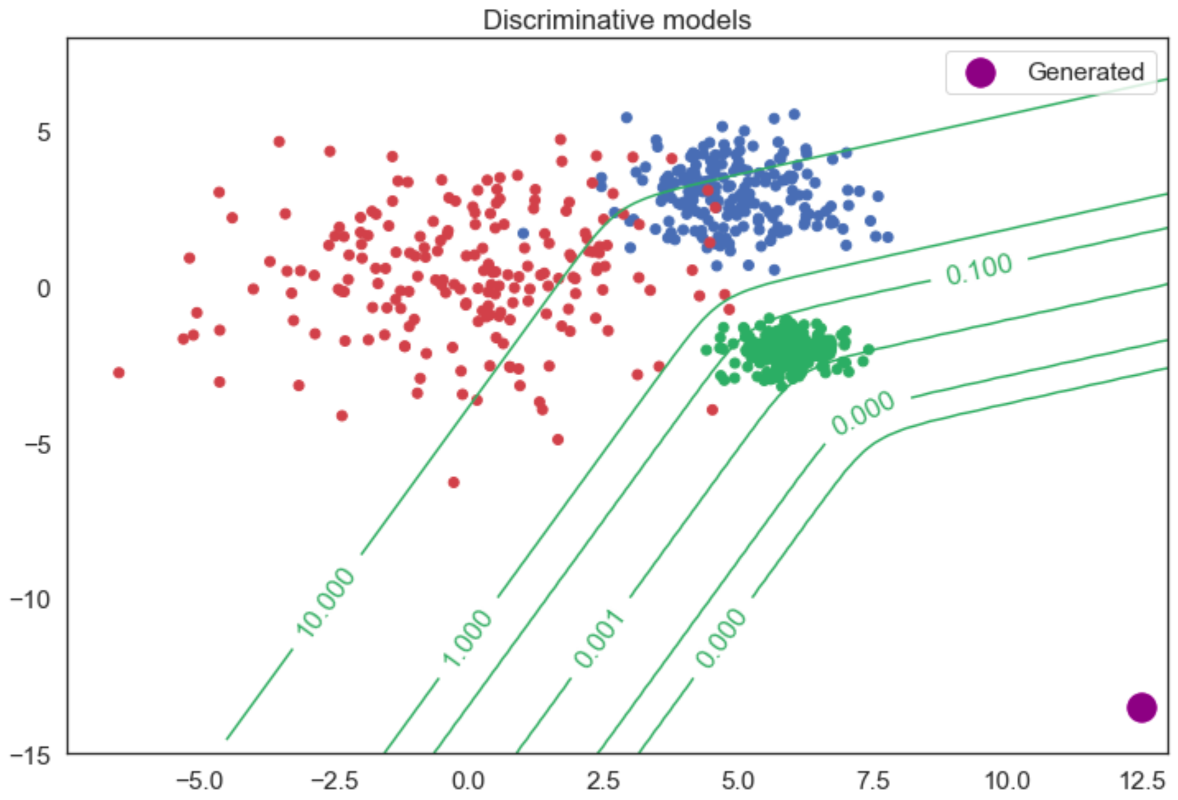
Discriminative models

Большинство изученных ранее моделей — дискриминативные. С помощью них можно выучить распределение P(y|x).



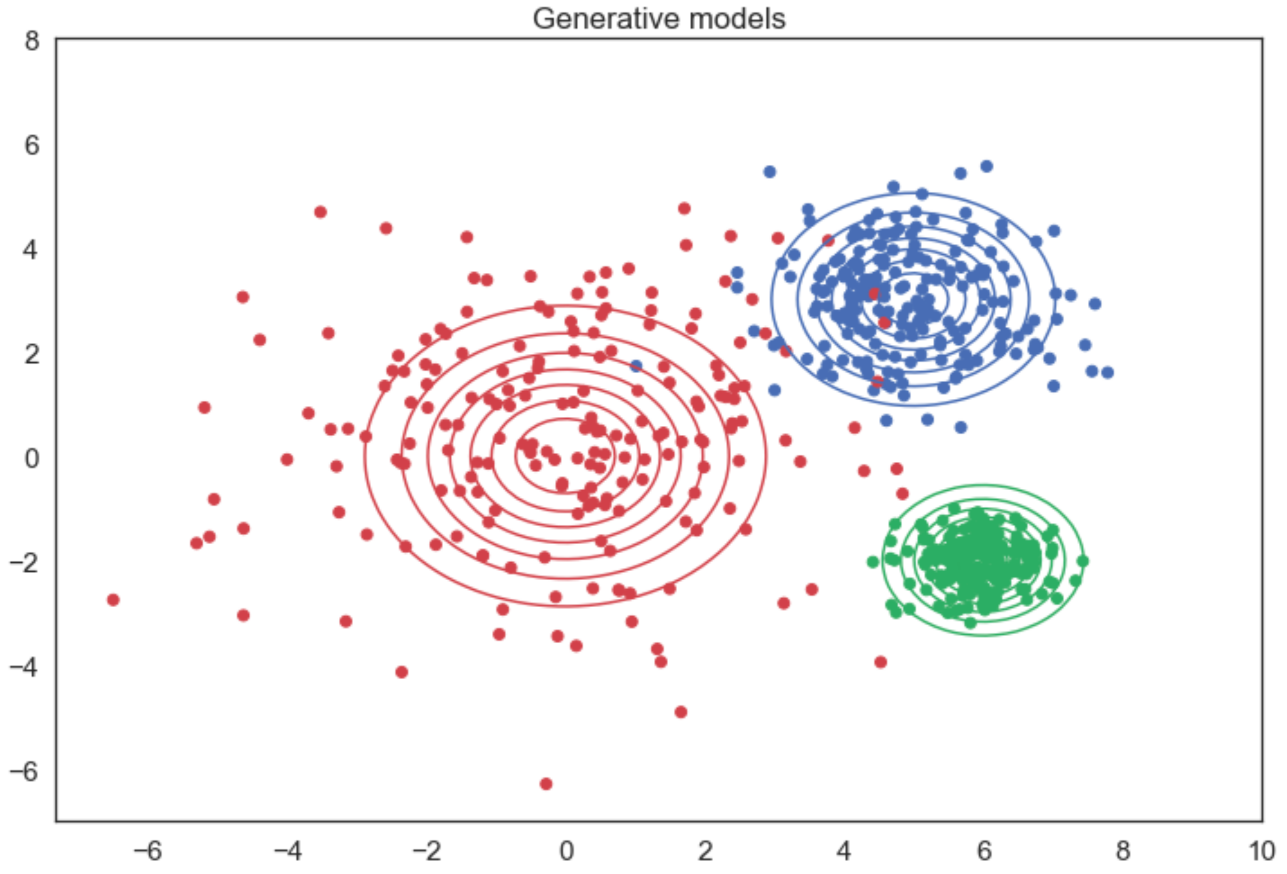
Discriminative models

Попробуем сгенерировать объект зеленого класса. Возьмем $\log(P(y|x))$. И для градиентного спуска добавим минус перед логарифмом.



Generative models

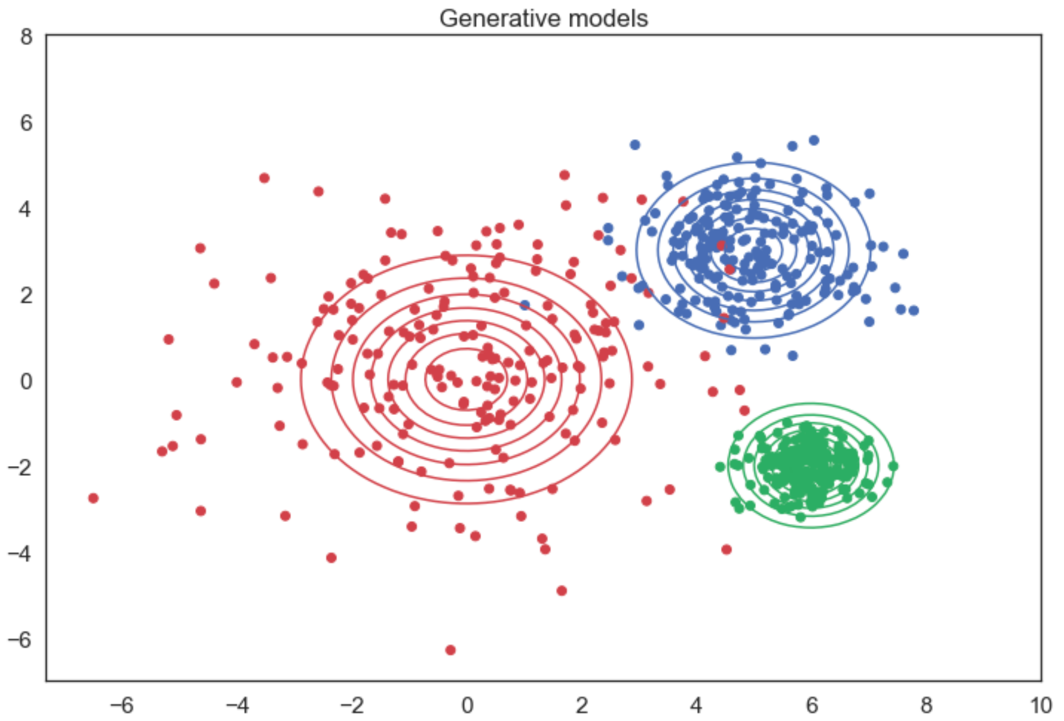
Для генерации нужен другой класс моделей.



Generative models

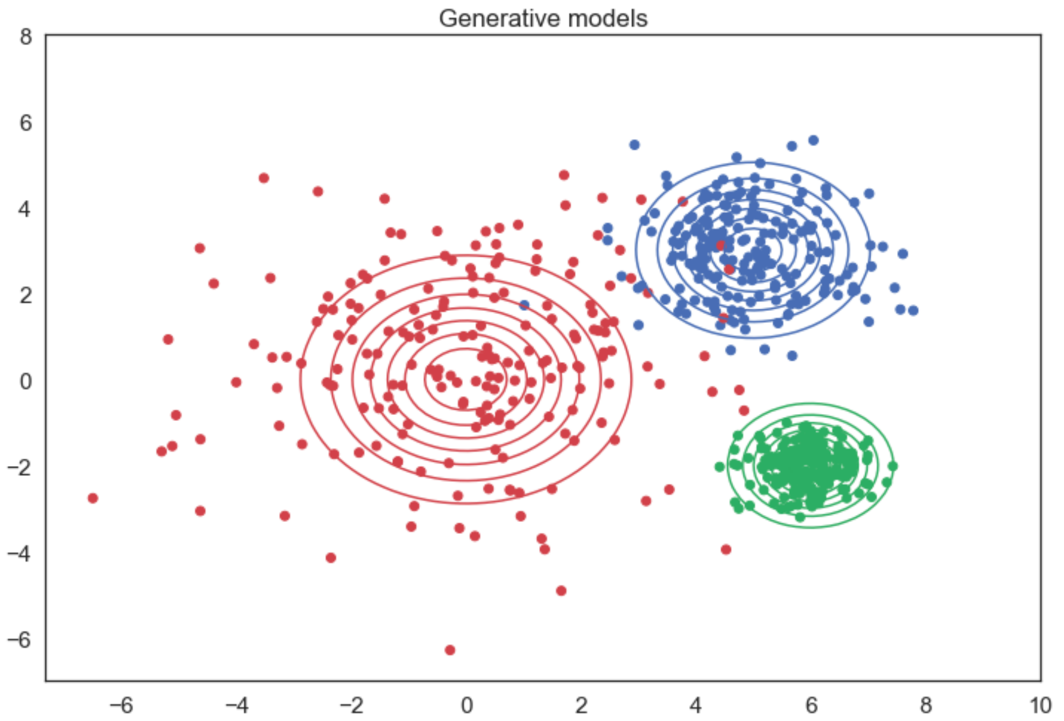
Для генерации нужен другой класс моделей.

Проблема: данные не всегда распределены унимодально в исходном пространстве.



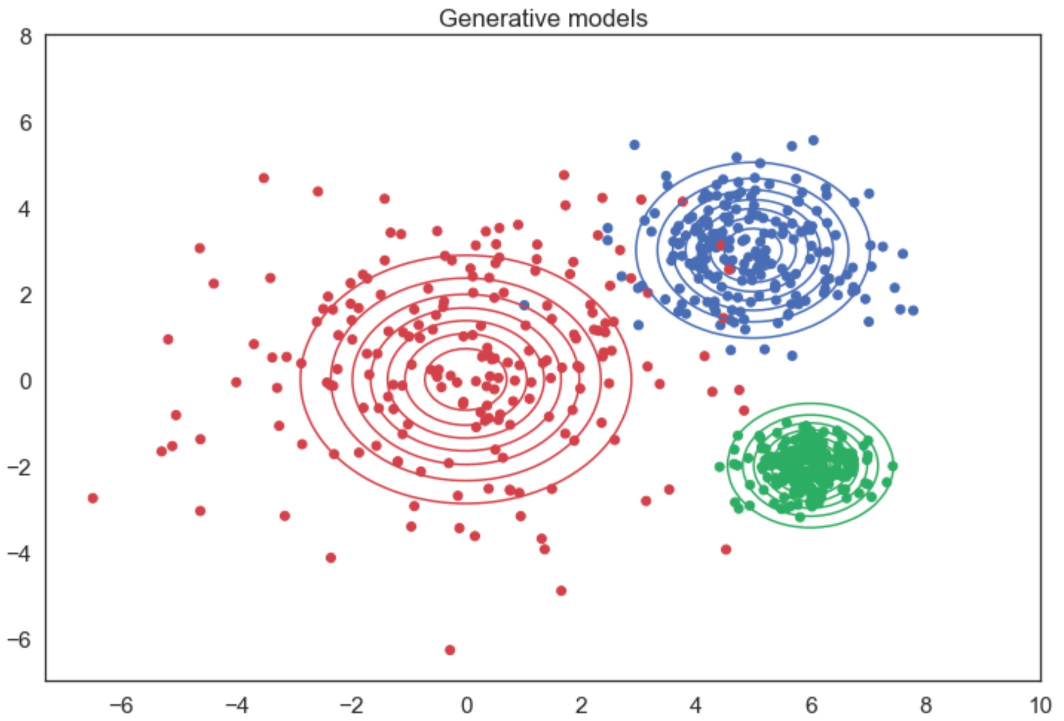
Generative models

Предположим что существует функция f такая что величина $f(X)$ будет распределена нормально, причем как правило можно взять $f : R^n \rightarrow R^m,\ m < n$ (в случае одномерных данных обычно наоборот). Такое пространство размерности m называется латентным.



Generative models

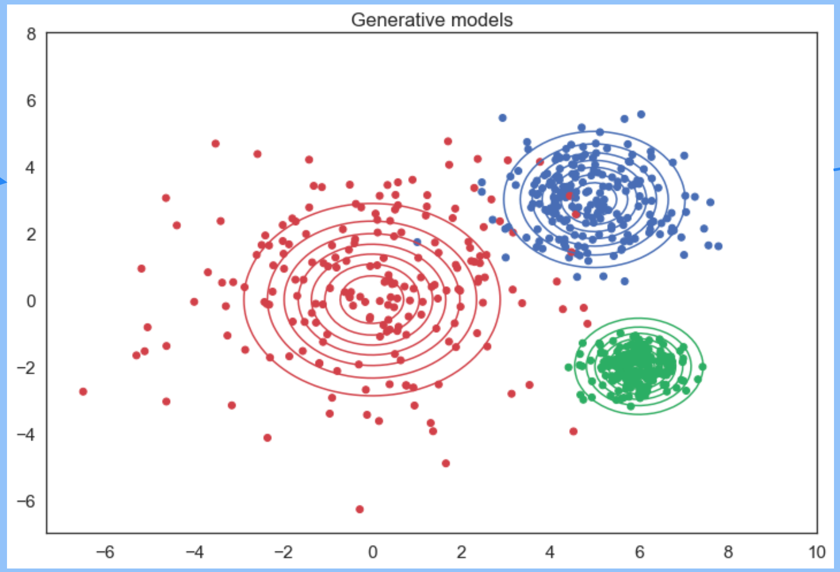
Предположим что существует функция f такая что величина $f(X)$ будет распределена нормально, причем как правило можно взять $f: R^n \to R^m,\ m < n$ (в случае одномерных данных обычно наоборот). Такое пространство размерности m называется латентным.
Для генерации нам необходимо найти и другую функцию g, такую, чтобы $g(f(x))$ снова была распределена как и исходные данные.



Generative models

X

f



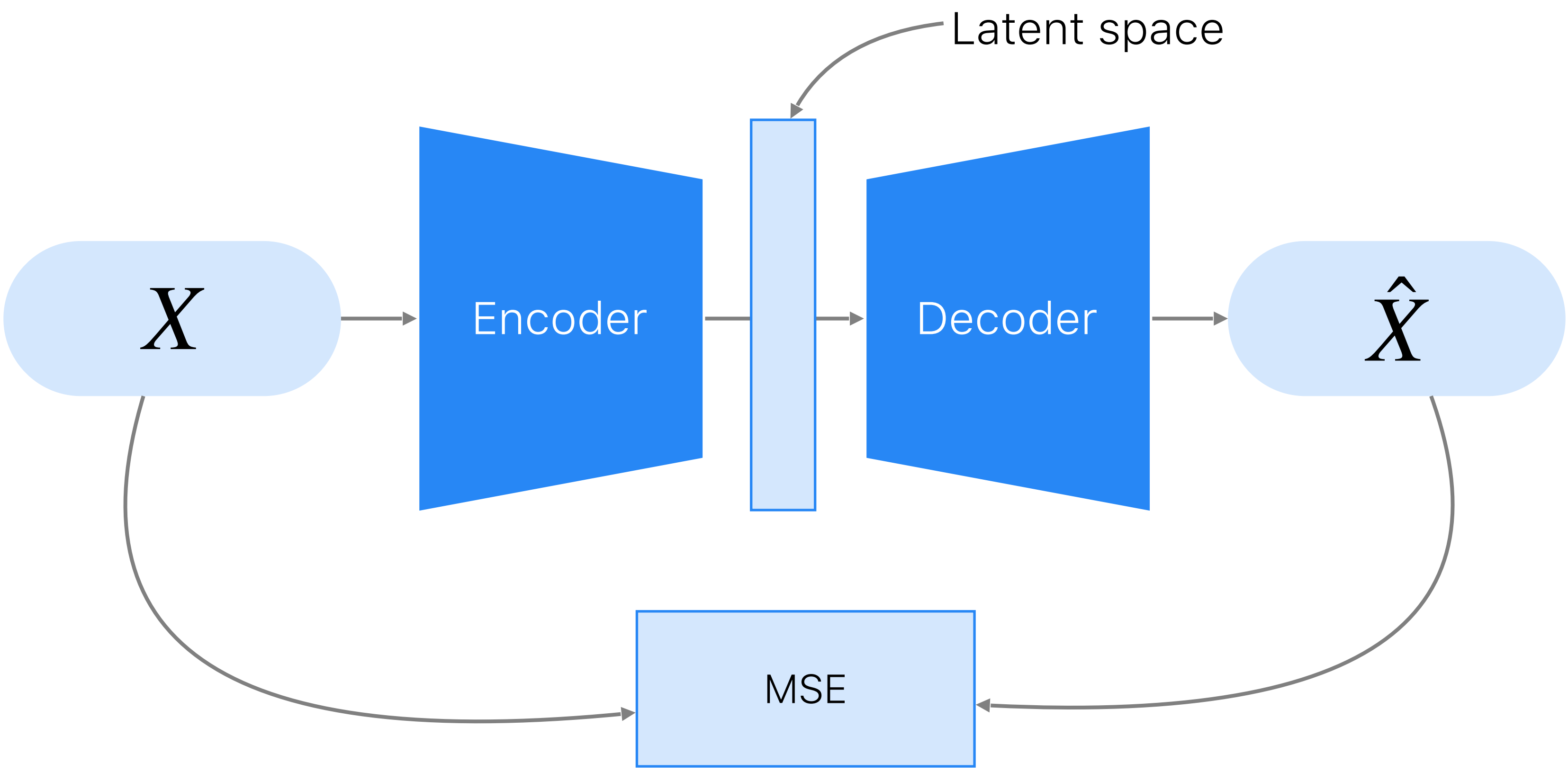
Латентное пространство
(Latent space)

g

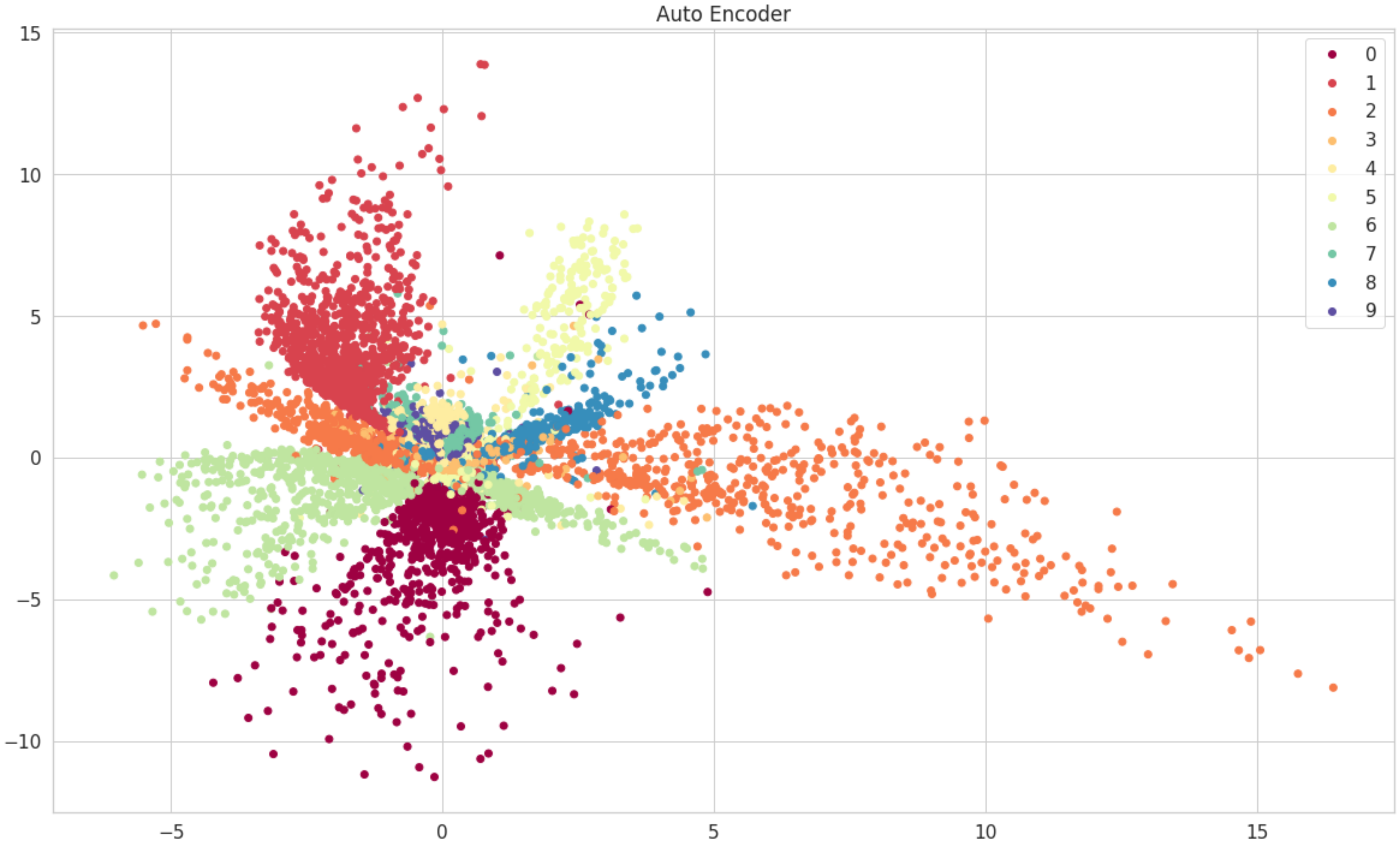
$f(g(X))$

Auto Encoder

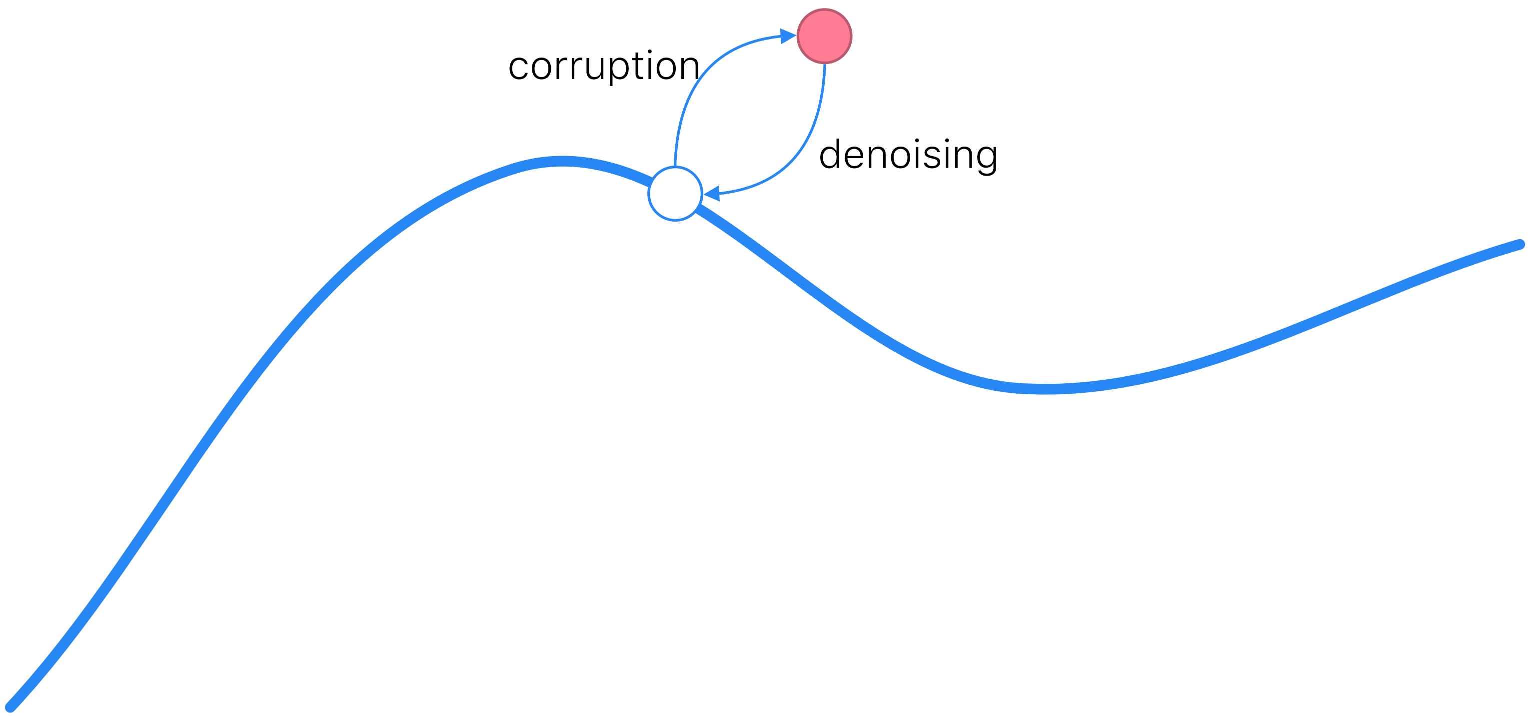
Auto Encoder



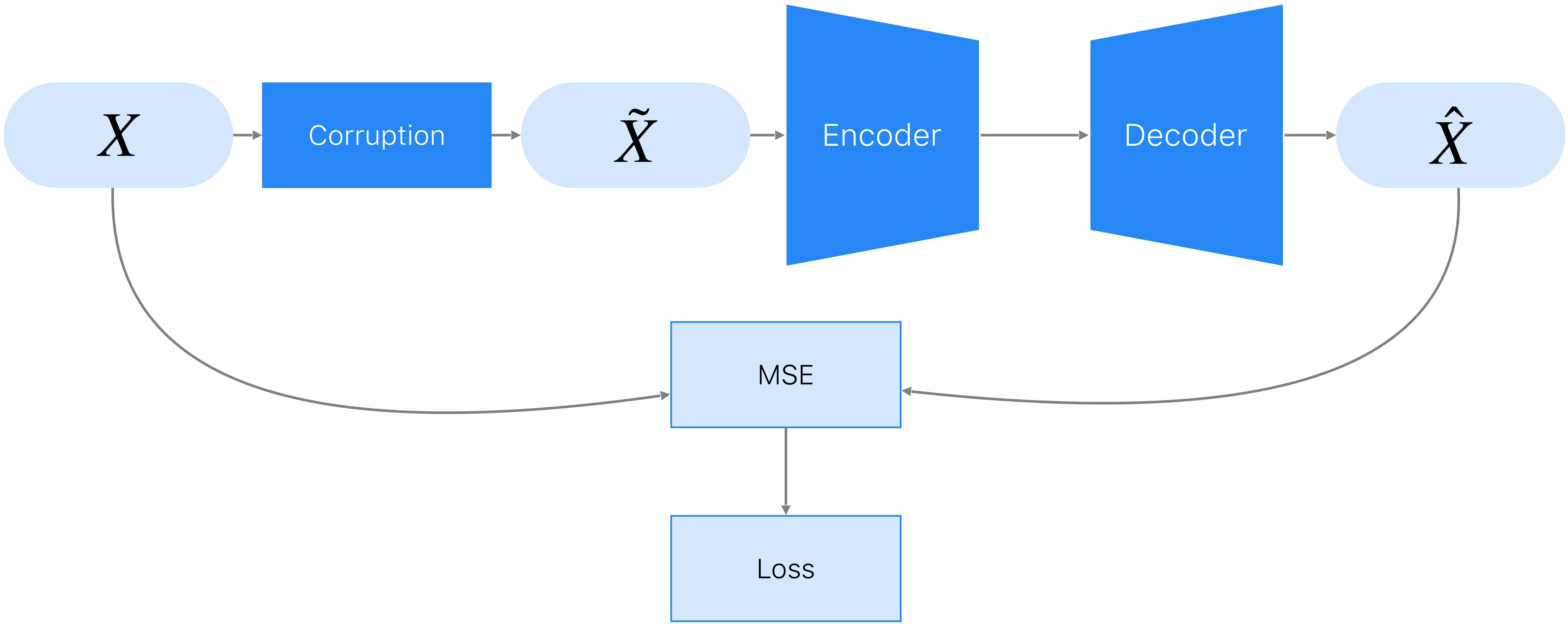
Dimensionality reduction



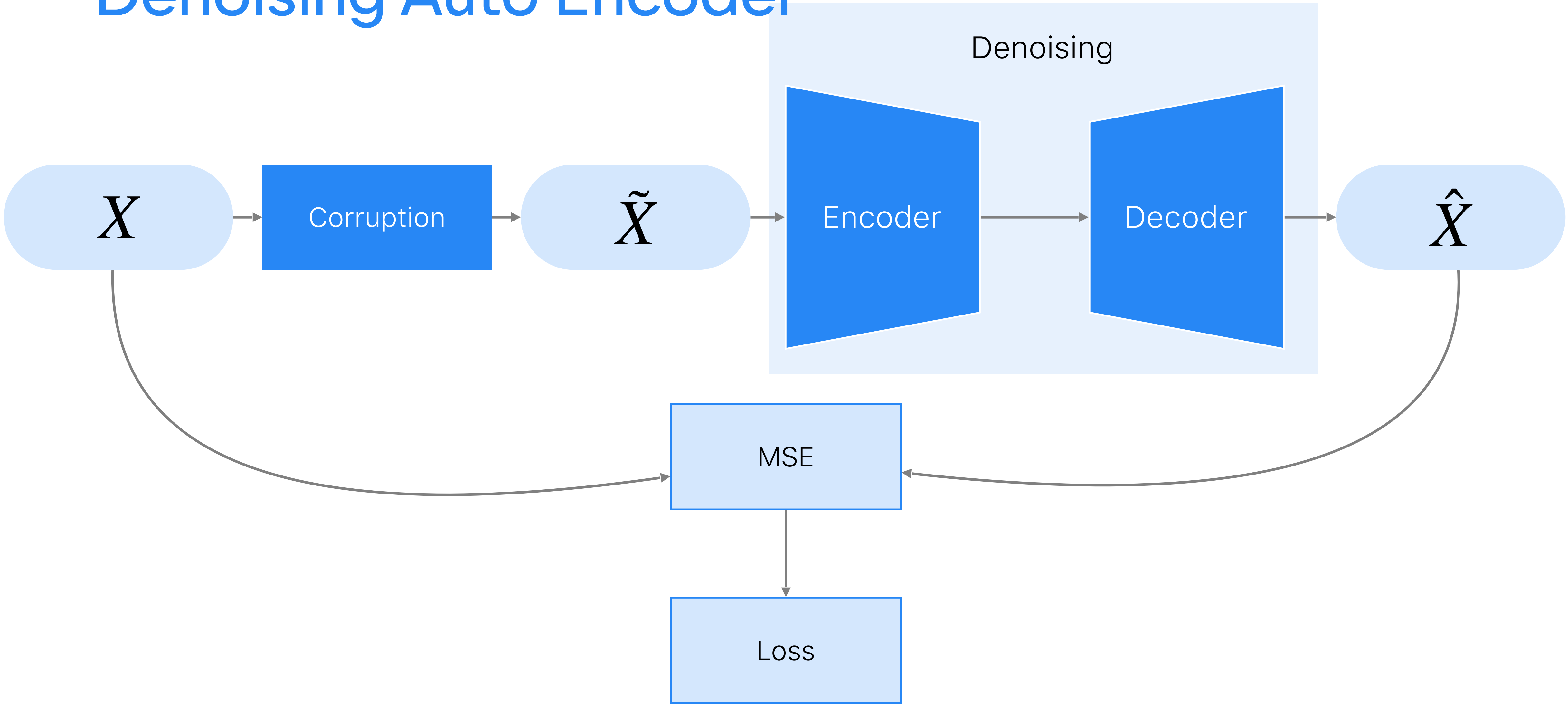
Denoising Auto Encoder



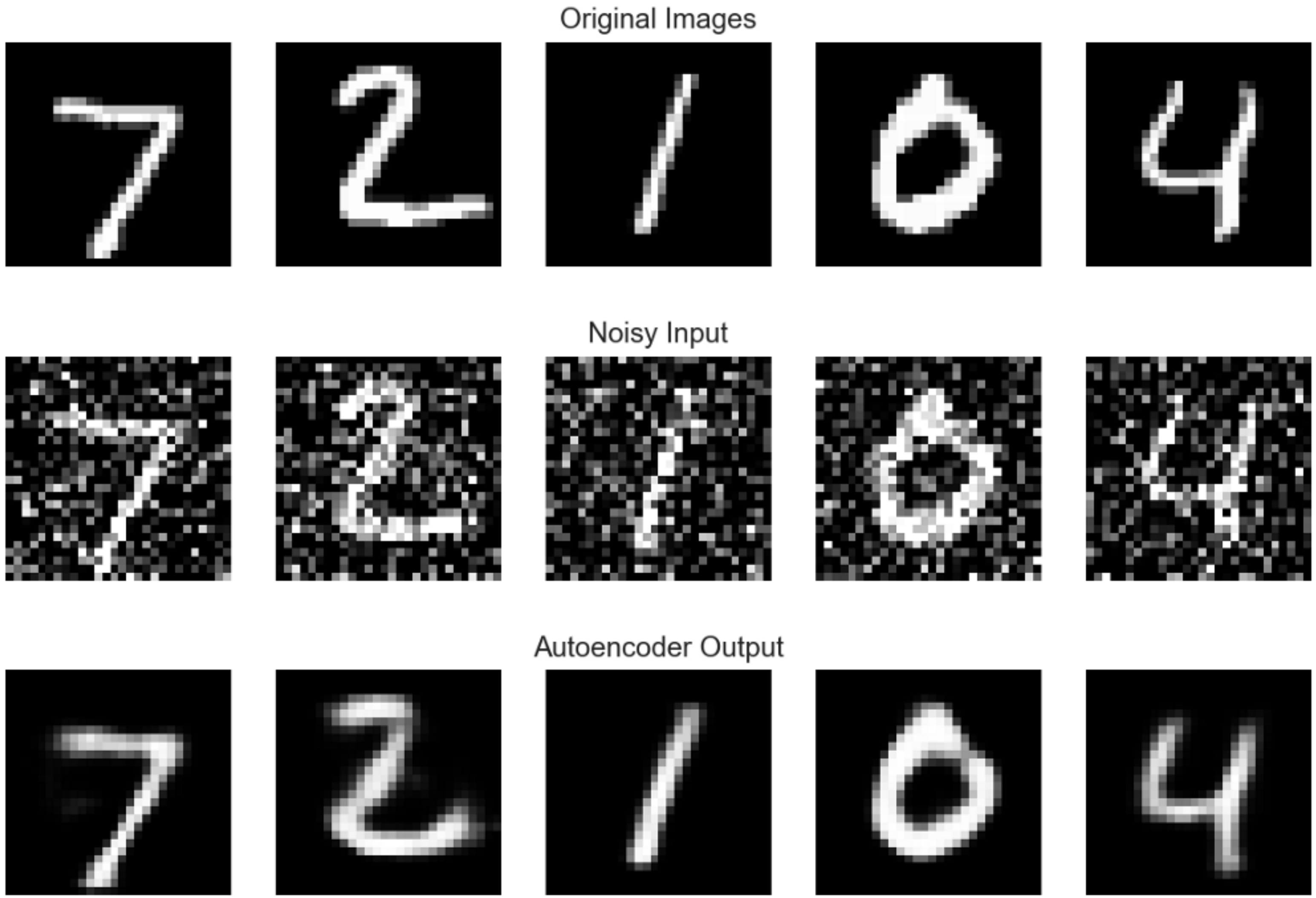
Denoising Auto Encoder



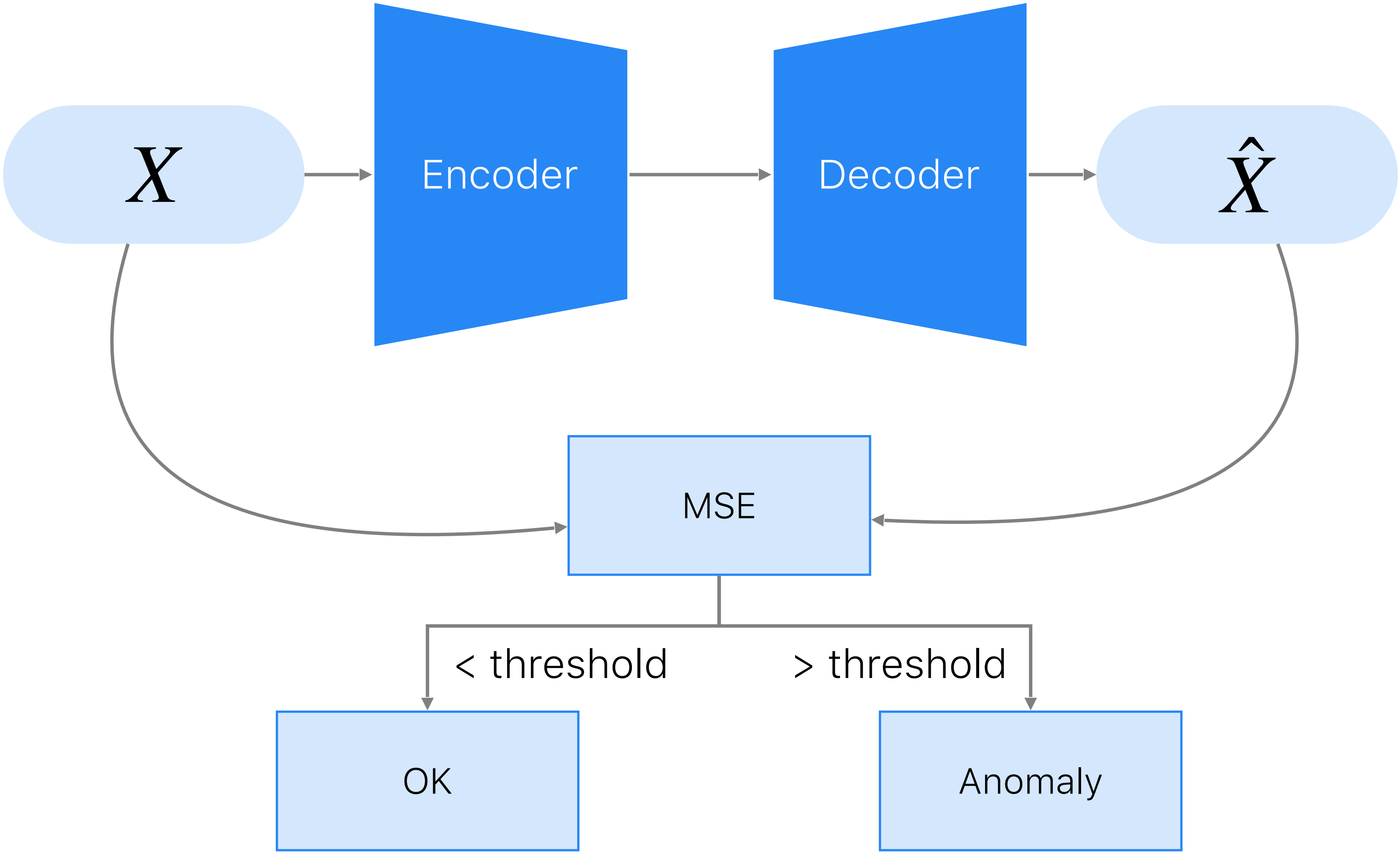
Denoising Auto Encoder



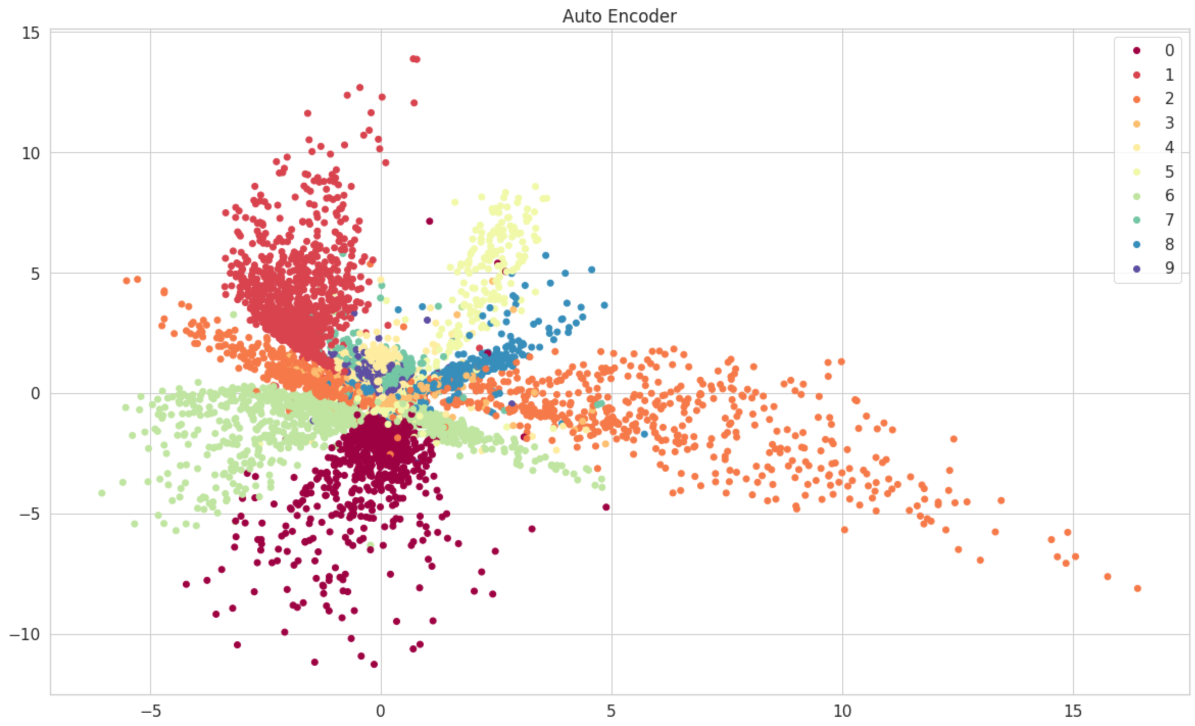
Denoising Auto Encoder

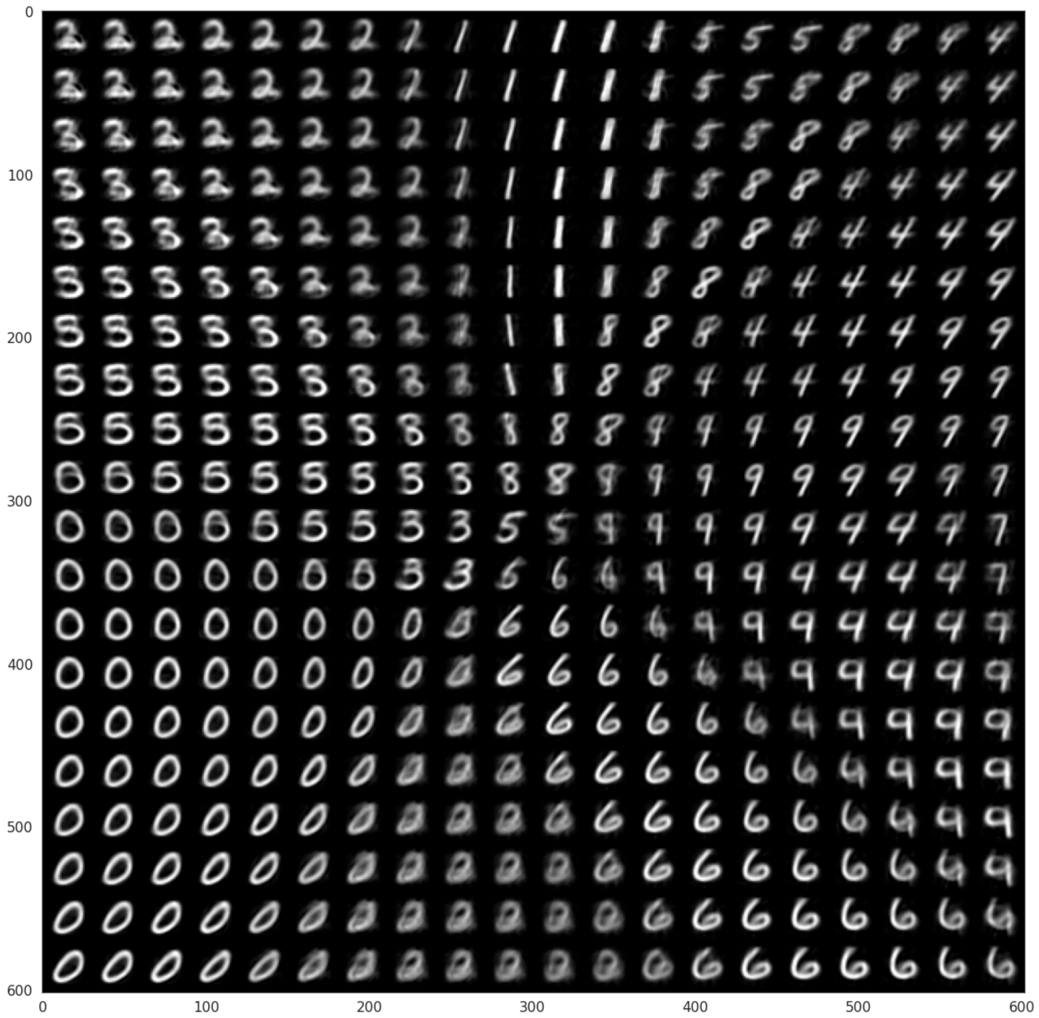


Anomaly detection with AE

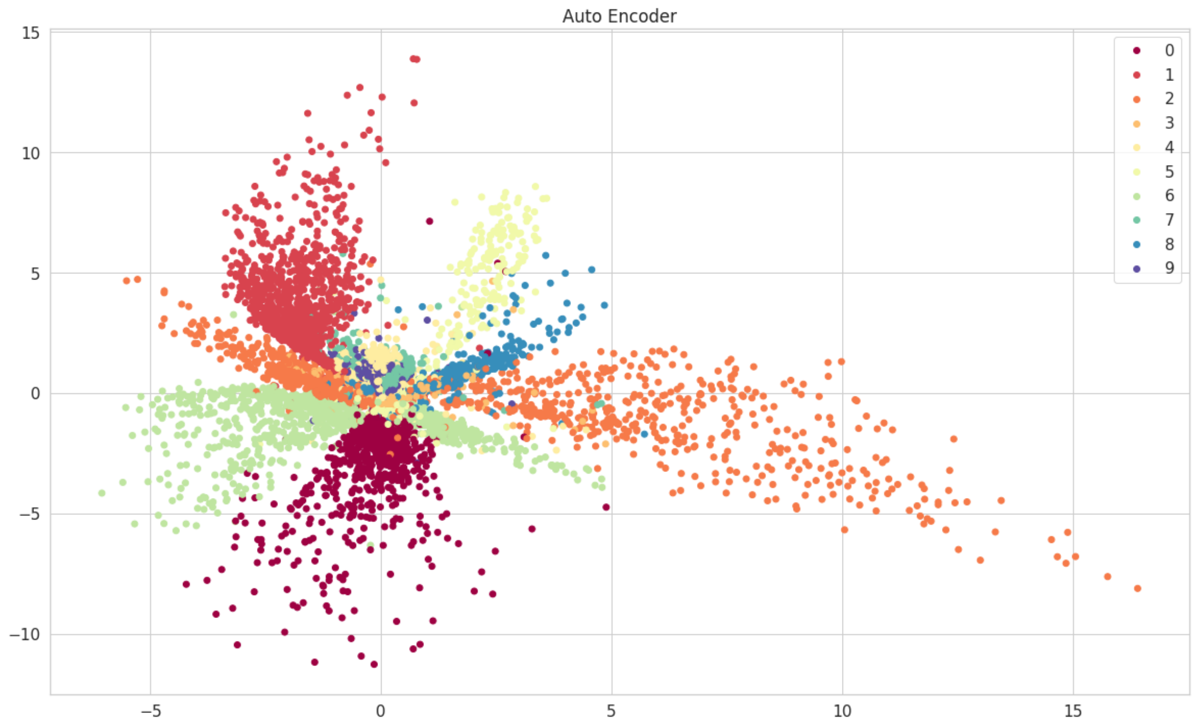


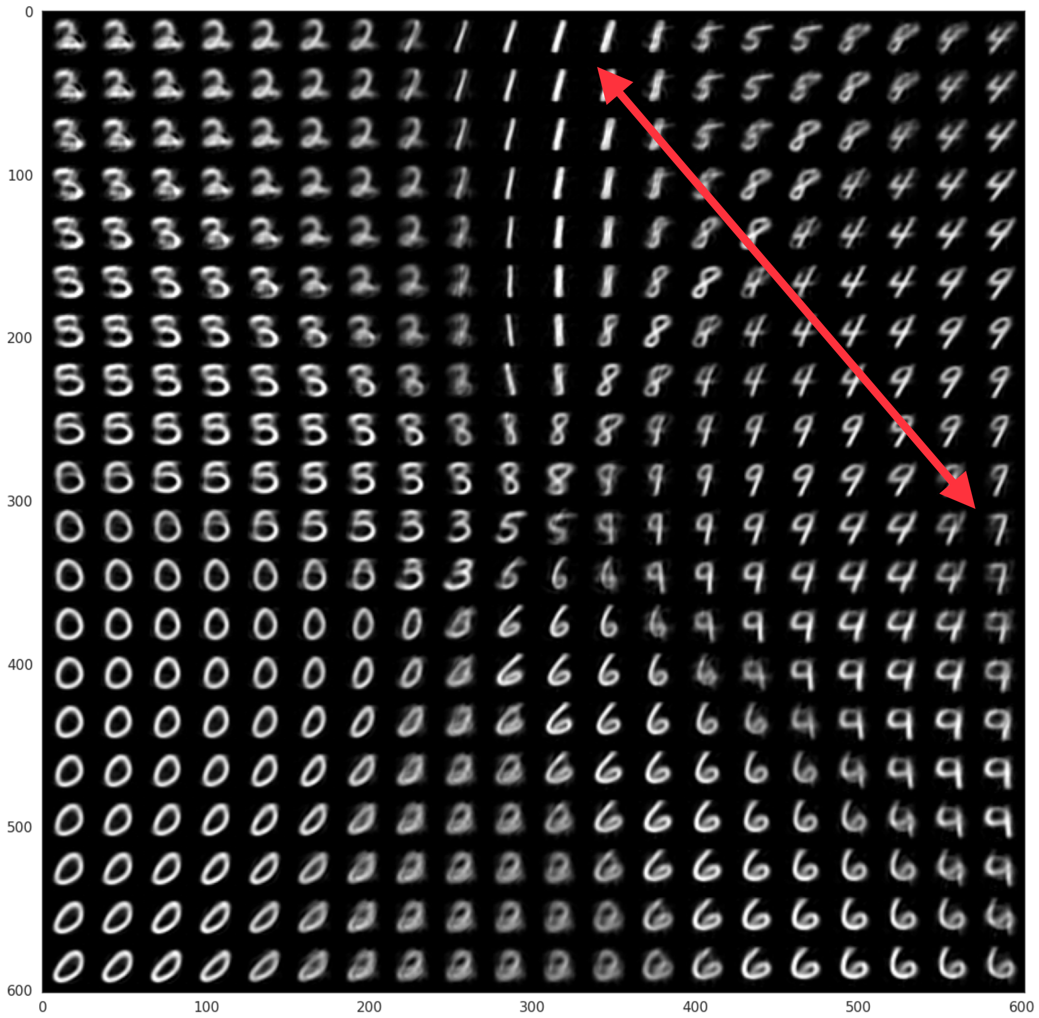
Как генерировать?



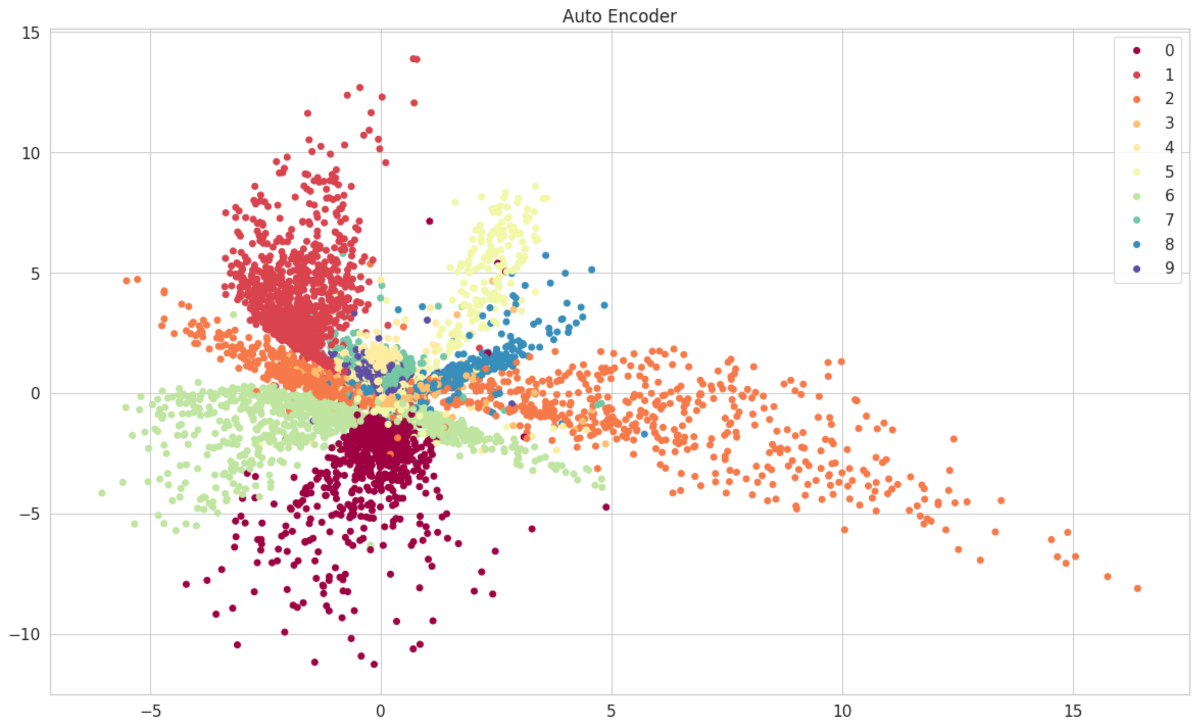


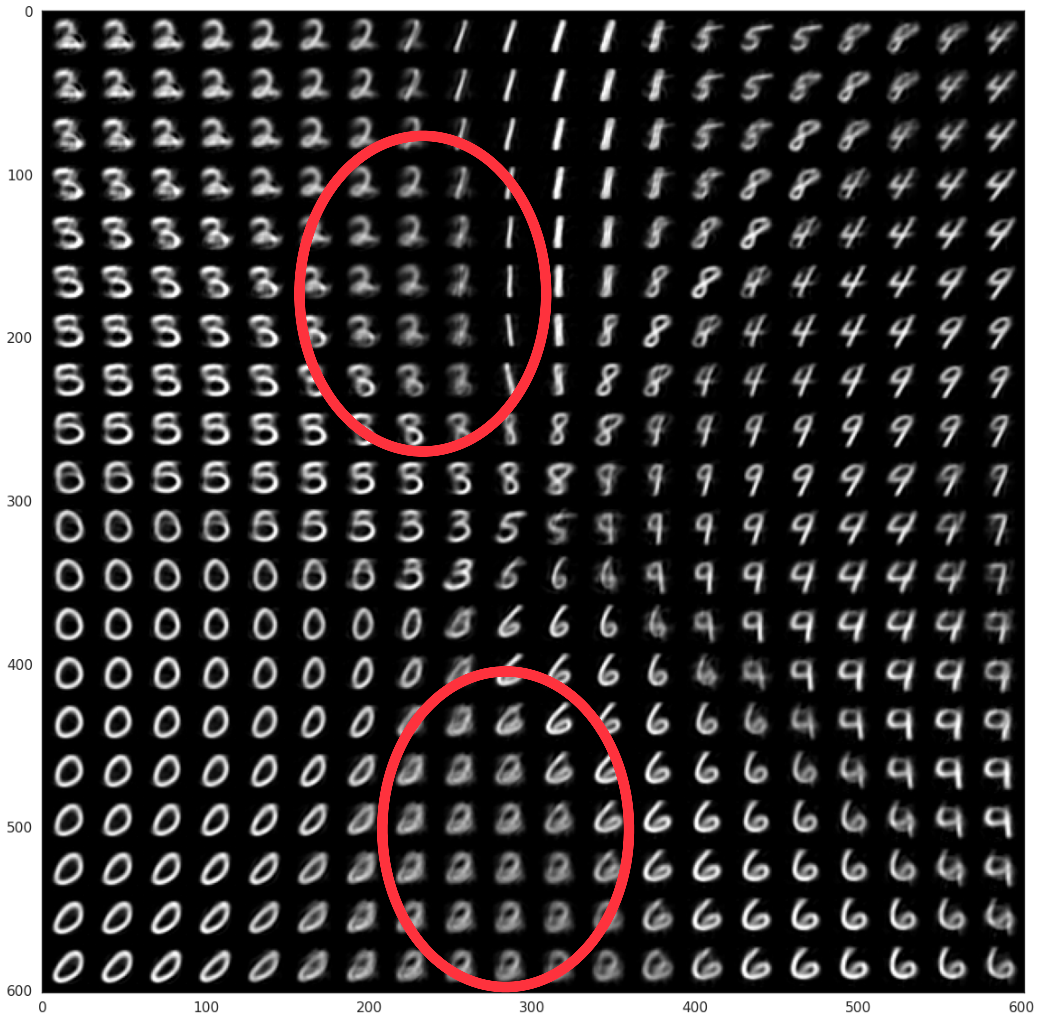
Как генерировать?





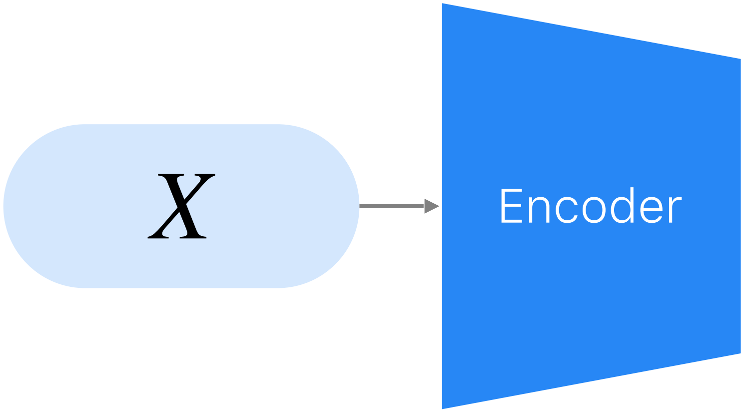
Как генерировать?



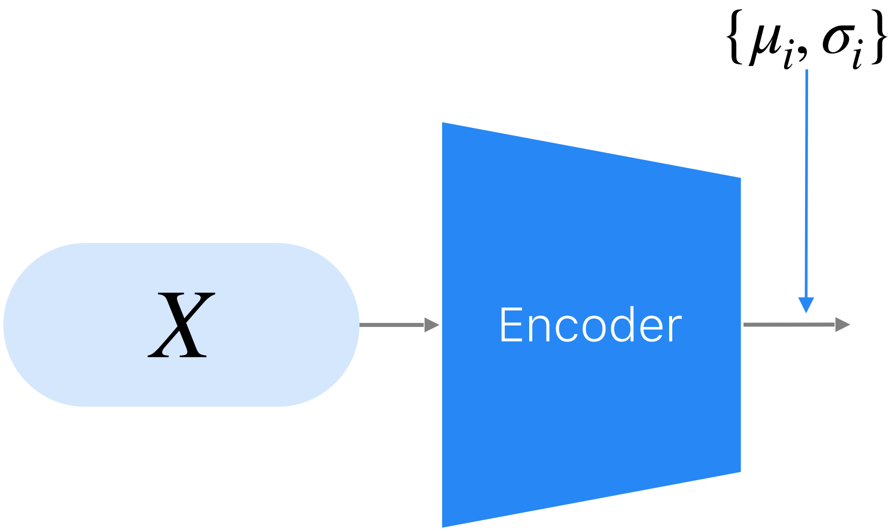


Variational Auto Encoder

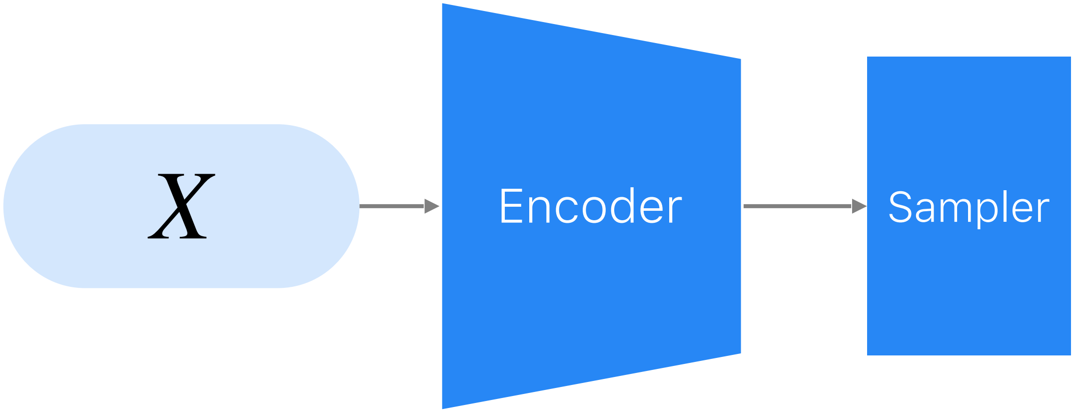
Variational Auto Encoder



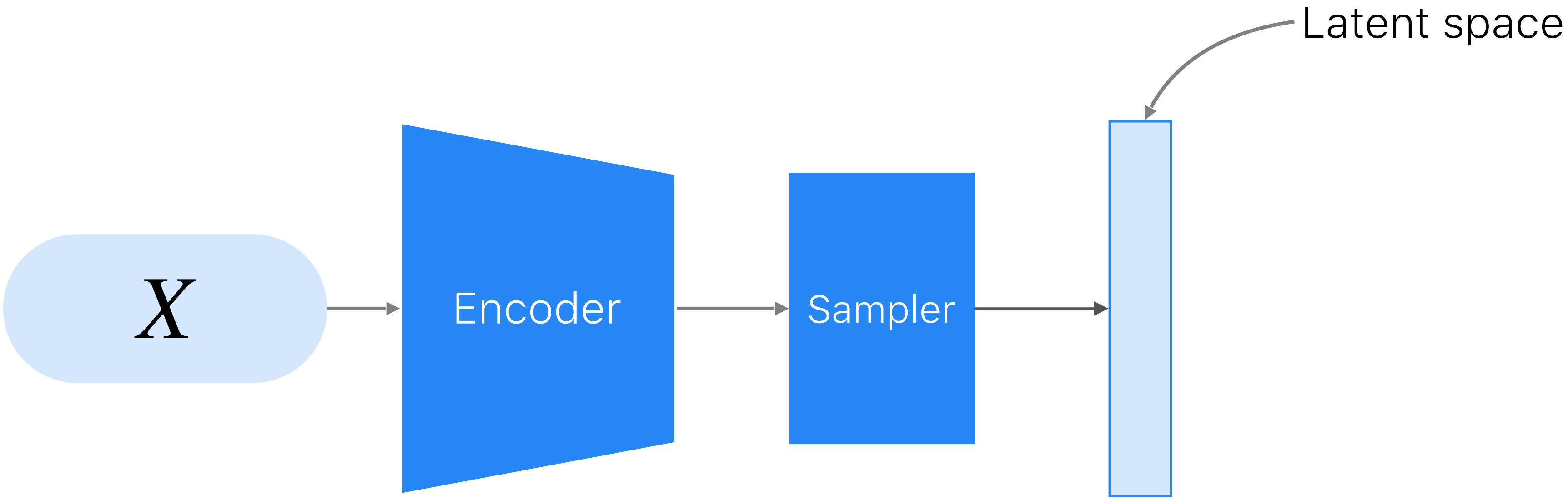
Variational Auto Encoder



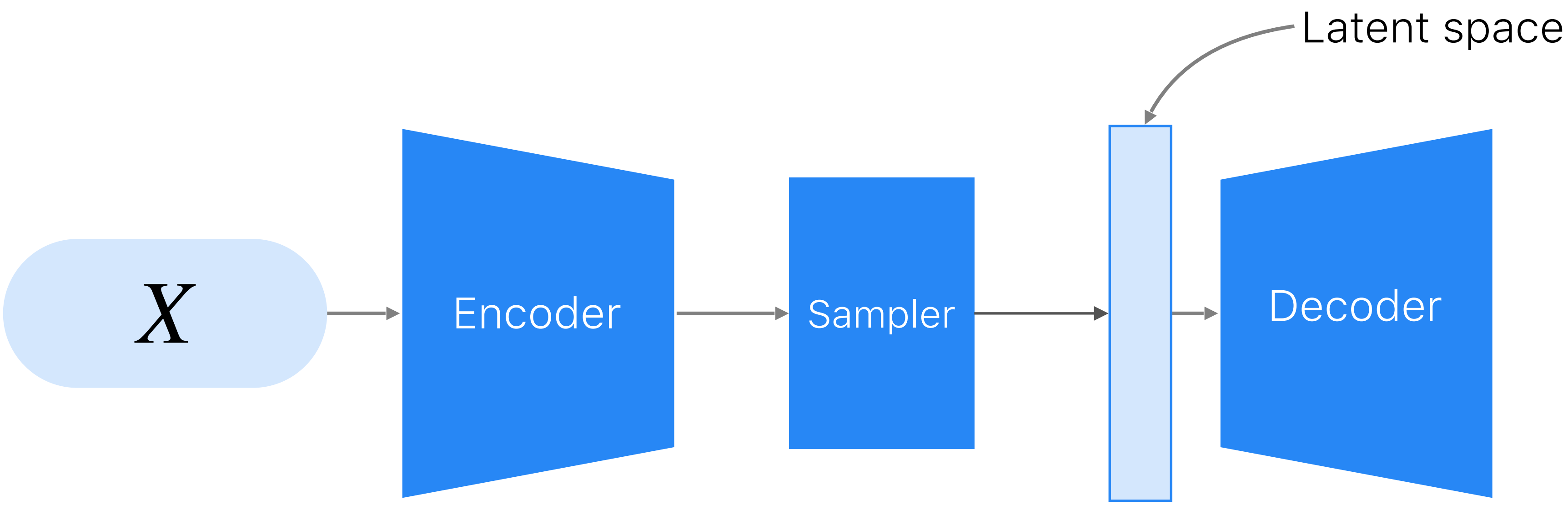
Variational Auto Encoder



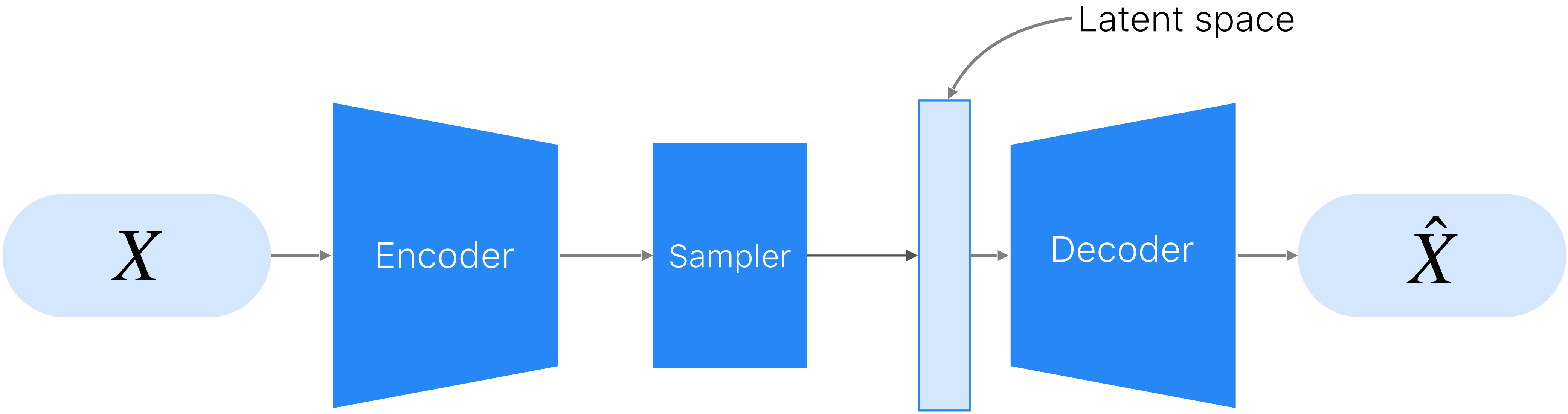
Variational Auto Encoder



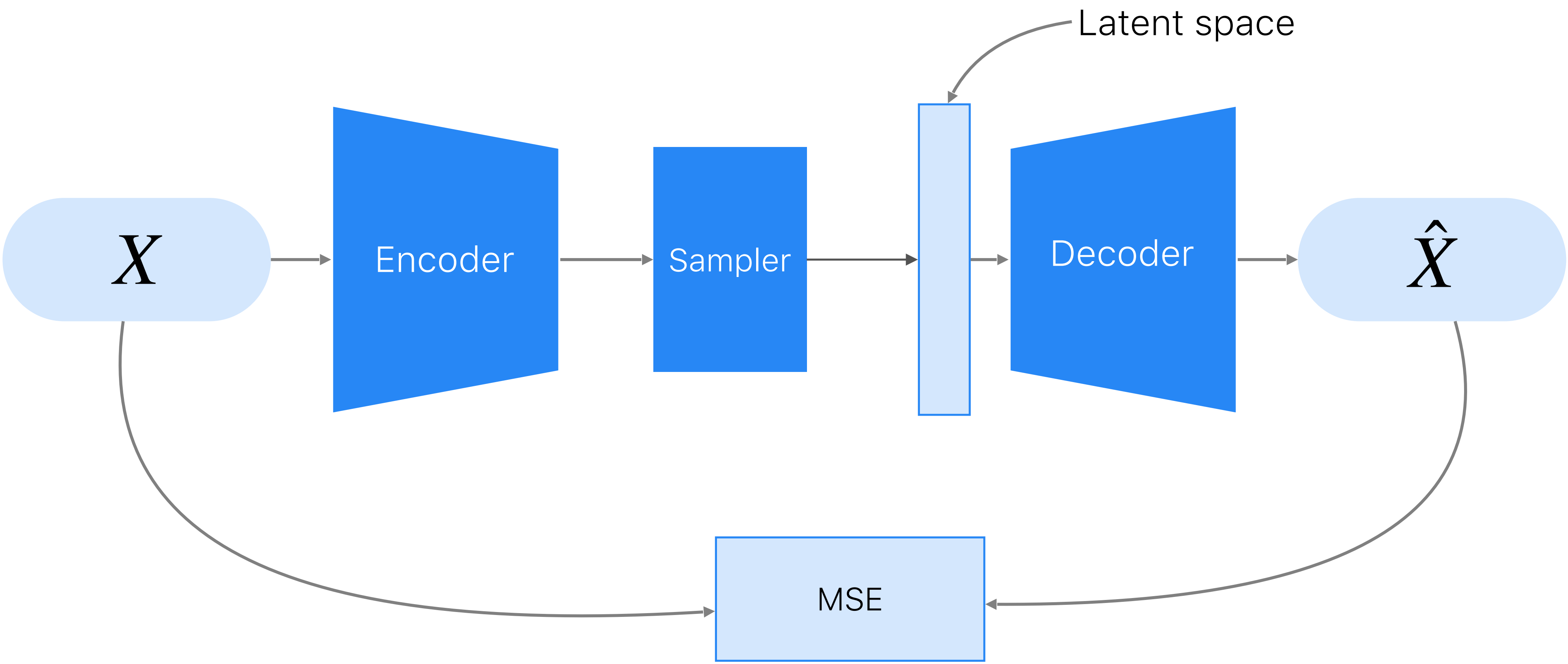
Variational Auto Encoder



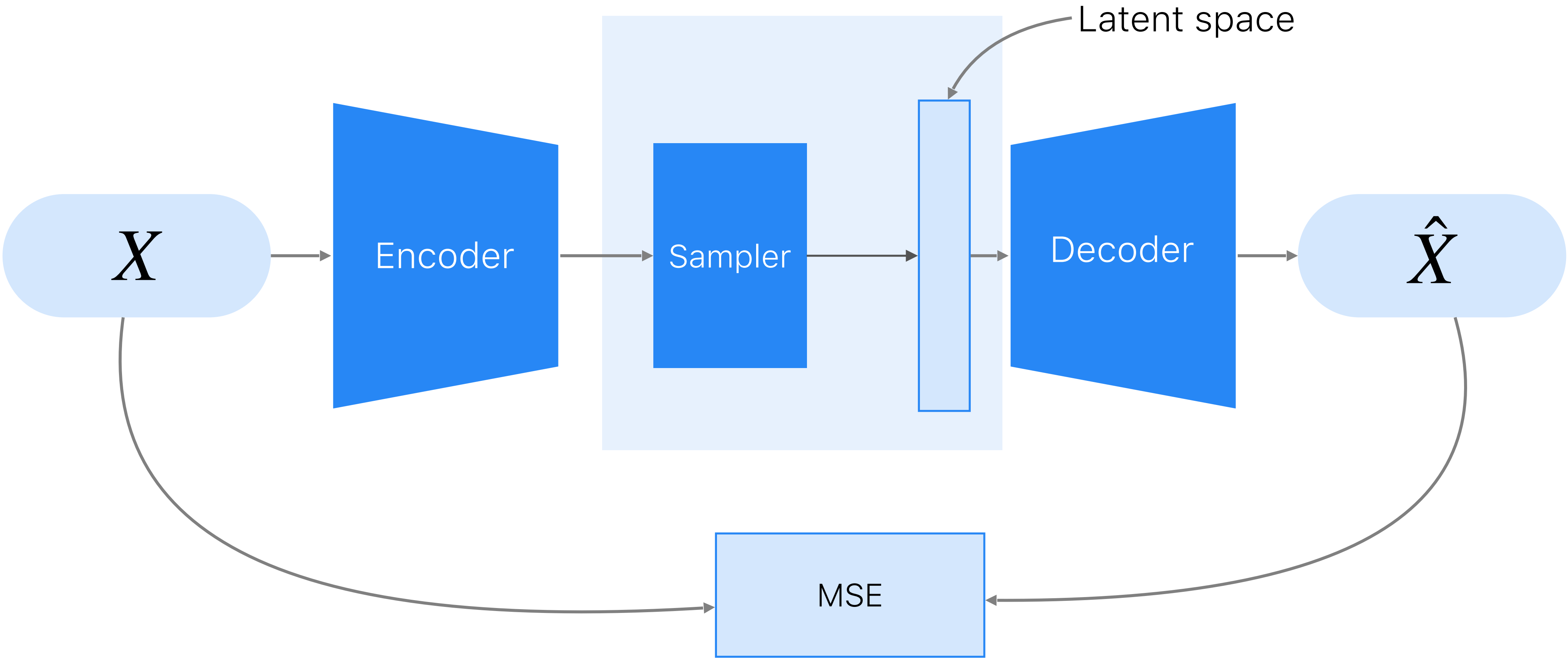
Variational Auto Encoder



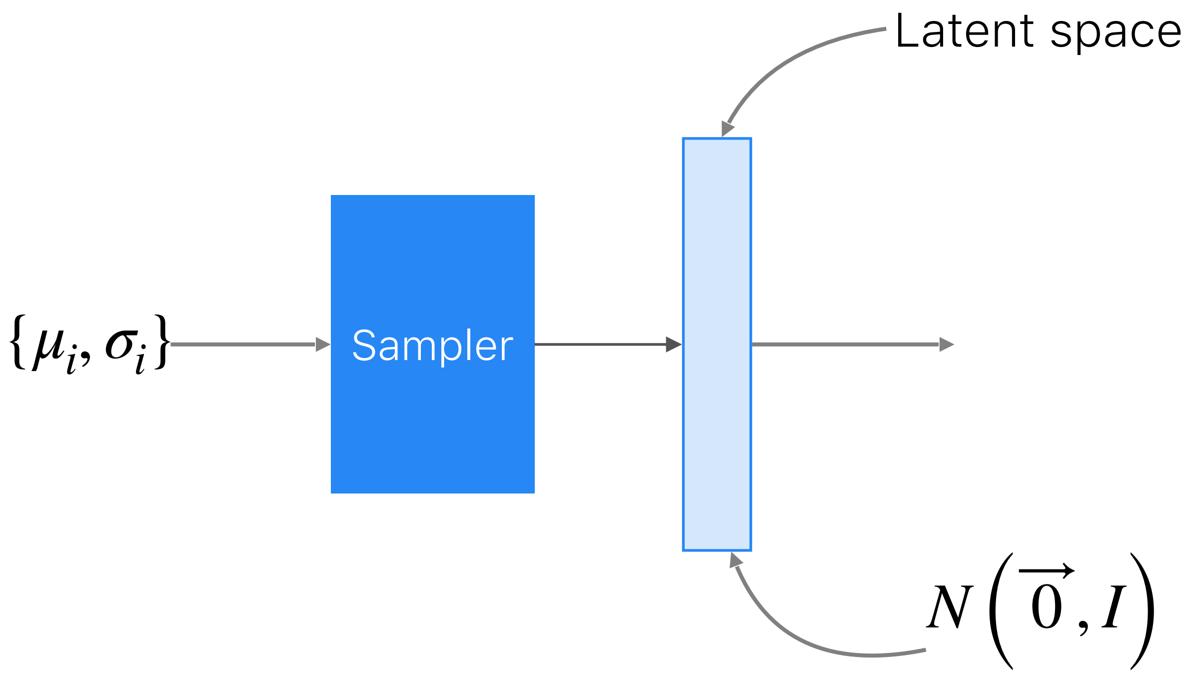
Variational Auto Encoder



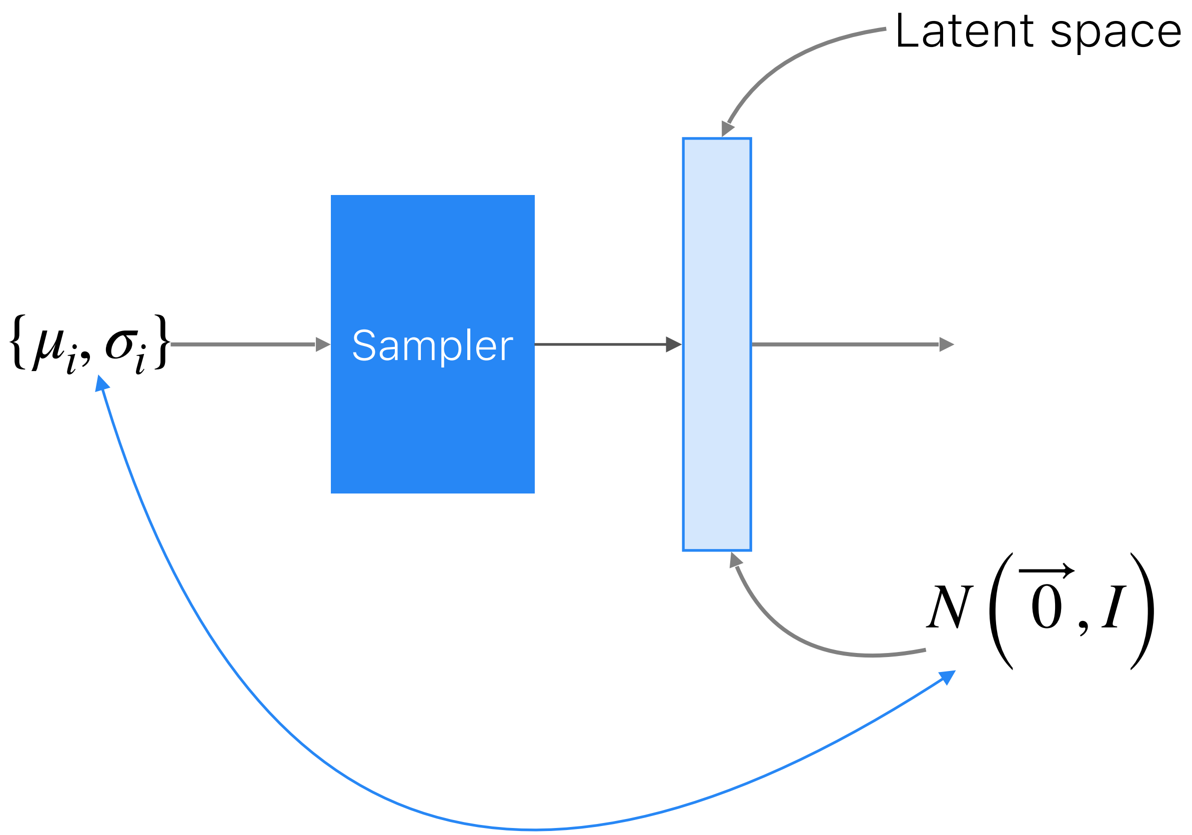
Variational Auto Encoder



Variational Auto Encoder



Variational Auto Encoder

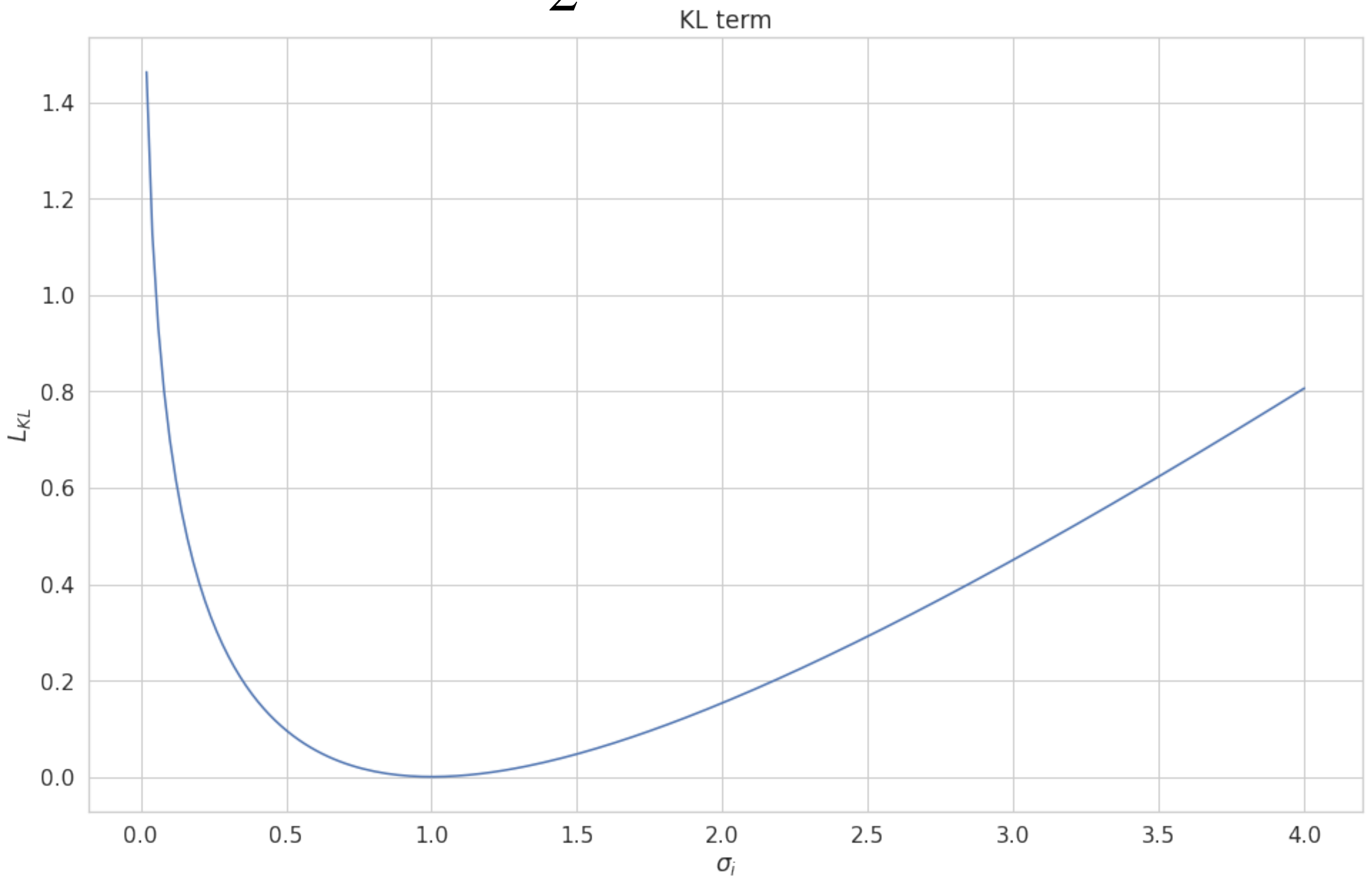


Variational Auto Encoder

$$L_{KL} = \frac{1}{2}\left(\sigma_i - \log \sigma_i - 1 + \mu_i^2\right)$$

Variational Auto Encoder

$$L_{KL} = \frac{1}{2}\left(\sigma_i - \log \sigma_i - 1\right)$$

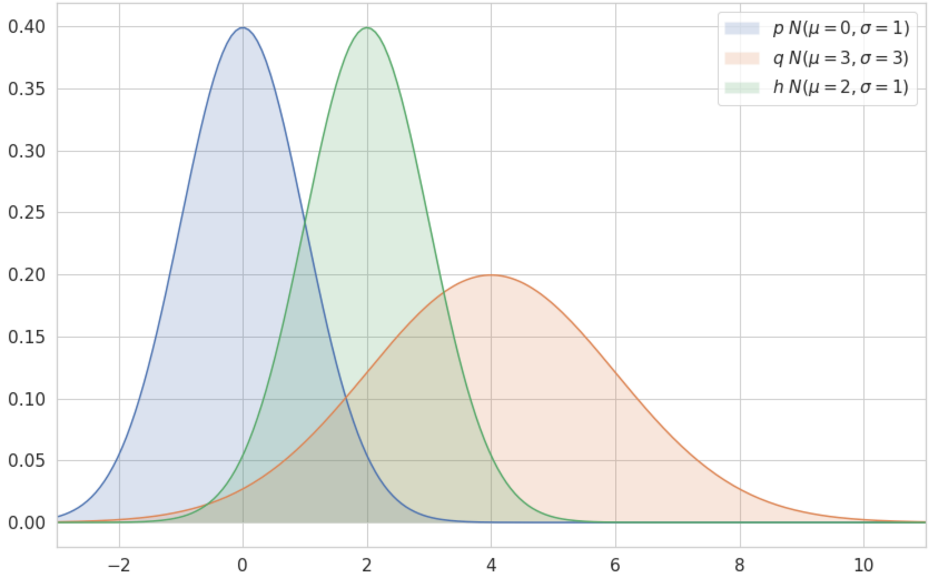


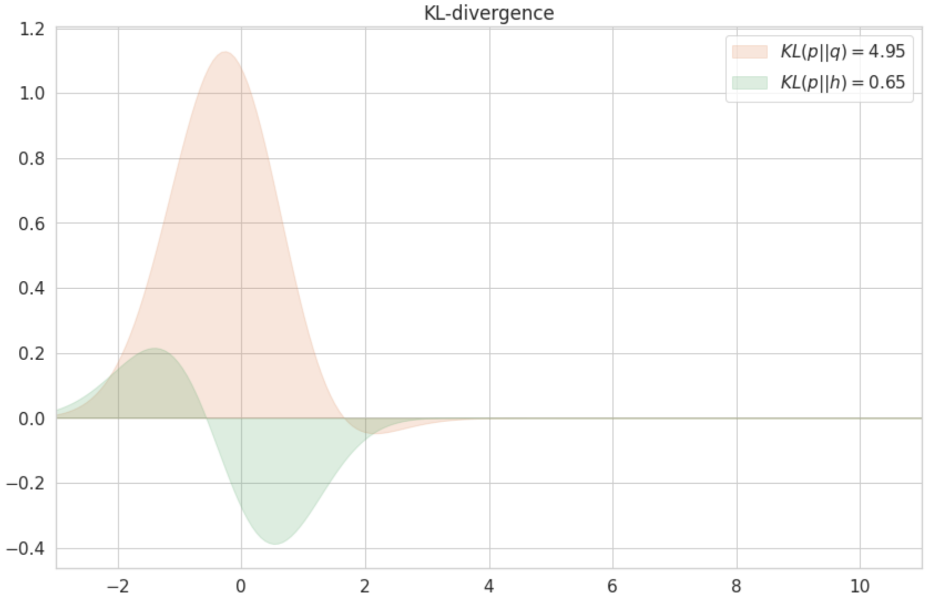
Дивергенция

- $D(p||q)$ — мера схожести двух распределений
- $D(p||p) = 0$
- $D(p||q) \geq 0$
- $D(p||q) + D(q||h) \nleq D(p||h)$
- $D(p||q) \neq D(q||p)$

KL-дивергенция

$$KL(p\,||\,q) = -\sum_{x\in\mathcal{X}} p(x)\log\left(\frac{q(x)}{p(x)}\right)$$





KL-дивергенция

$$KL(p||q) = -\sum_{x \in \mathcal{X}} p(x) \log \left(\frac{q(x)}{p(x)} \right)$$

$N(0,1)$ $\quad p = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2} \right)$

$q = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(x^2 - \mu_i)}{2\sigma_i^2} \right)$ $\quad N(\mu_i, \sigma_i)$

KL-дивергенция

$$KL(p||q) = -\sum_{x \in \mathcal{X}} p(x) \log\left(\frac{q(x)}{p(x)}\right)$$

$$p = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right)$$

$$q = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(x^2 - \mu_i)}{2\sigma_i^2}\right)$$

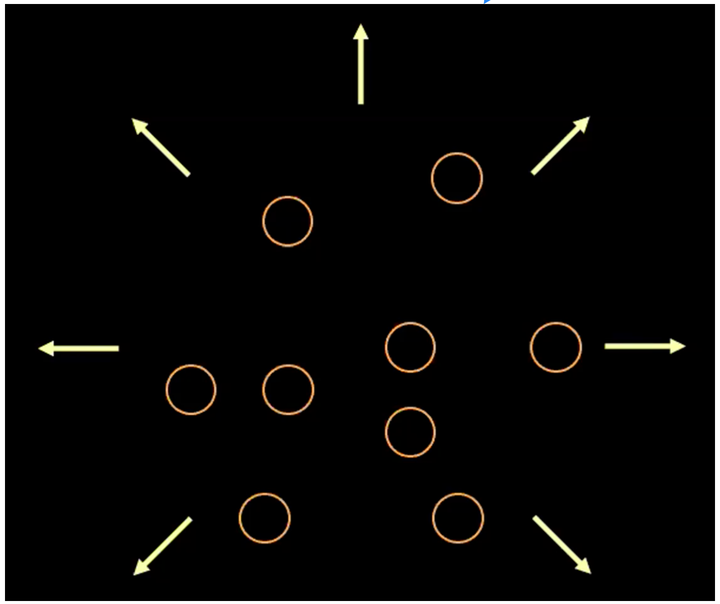
$$KL(p||q) = \frac{1}{2}\left(\sigma_i - \log\sigma_i - 1 + \mu_i^2\right)$$

VAE Loss

$$L = L_{rec} + \beta \cdot L_{KL}$$

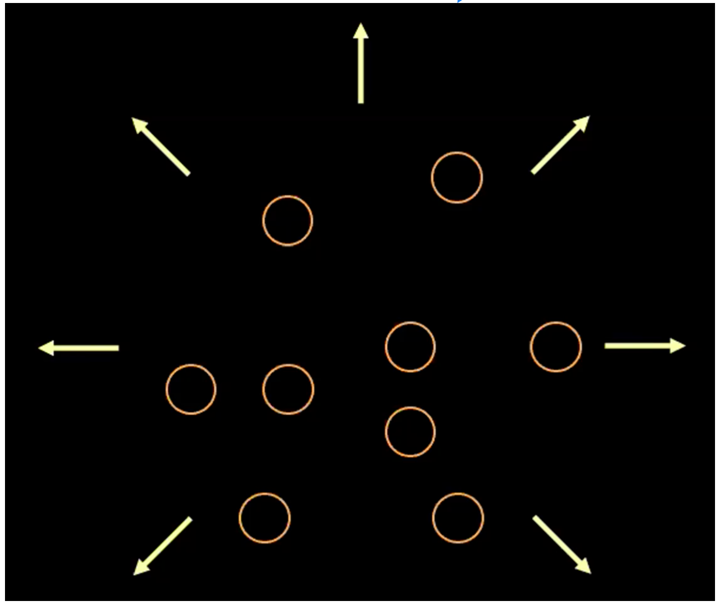
VAE Loss

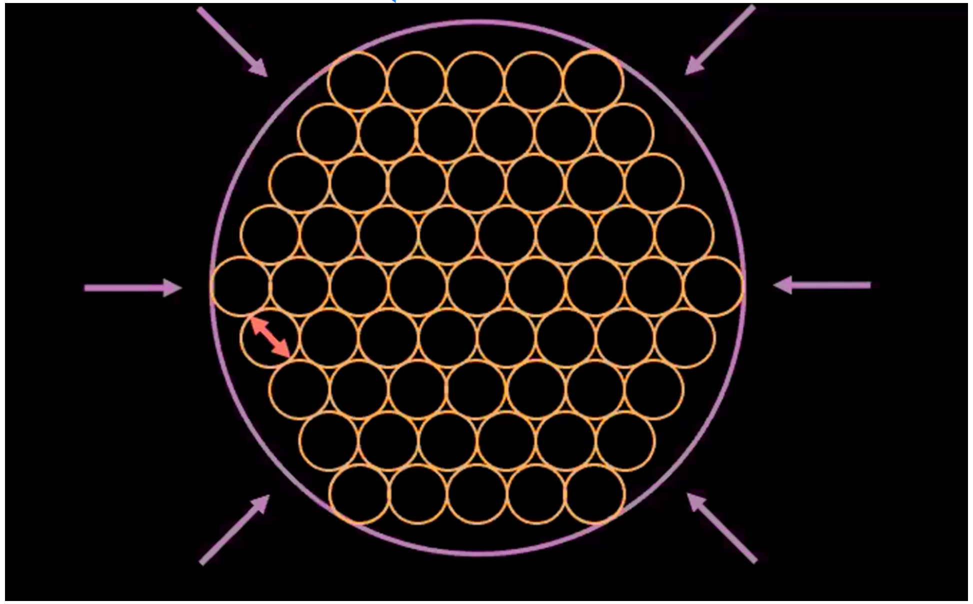
$$L = \boxed{L_{rec}} + \beta \cdot L_{KL}$$



VAE Loss

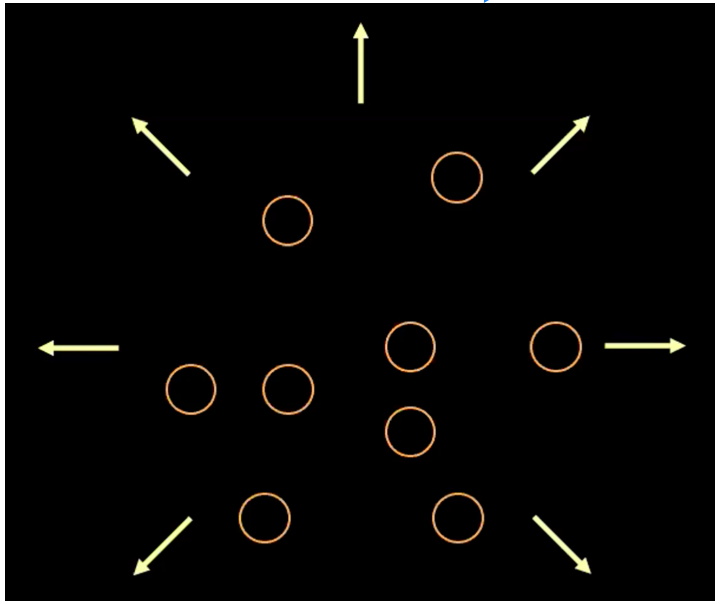
$$L = L_{rec} + \boxed{\beta \cdot L_{KL}}$$

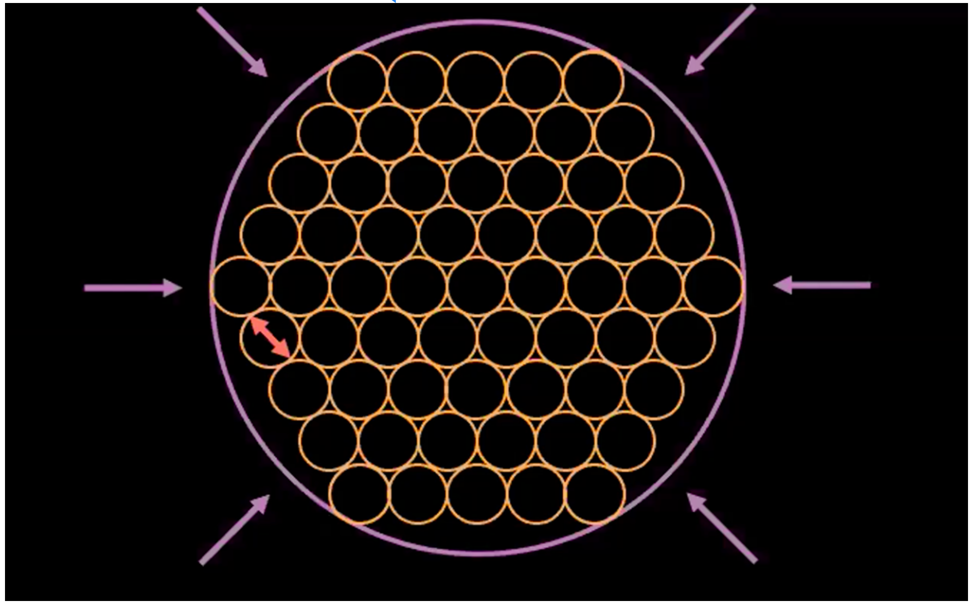




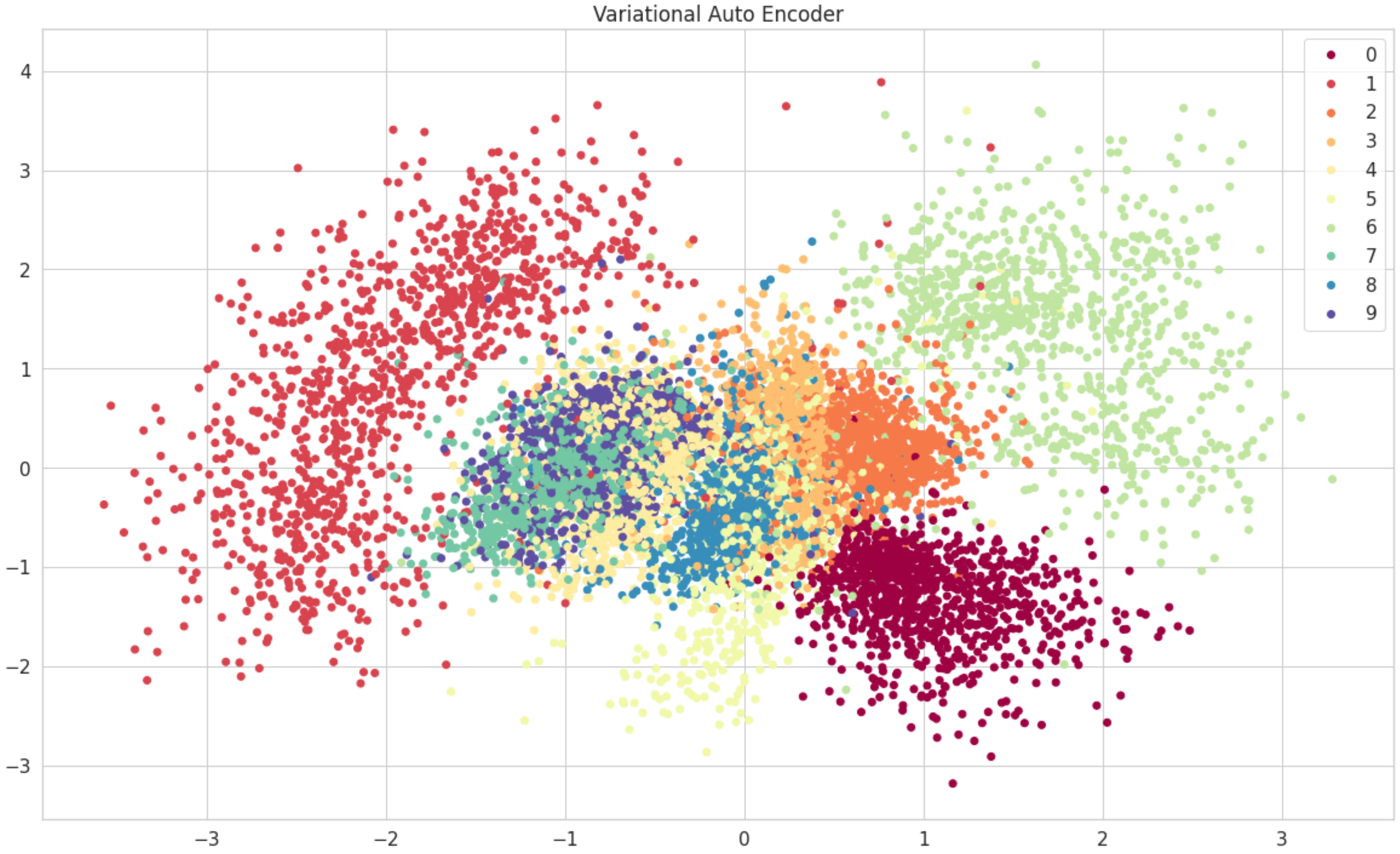
VAE Loss

$$L = L_{rec} + \beta \cdot L_{KL}$$

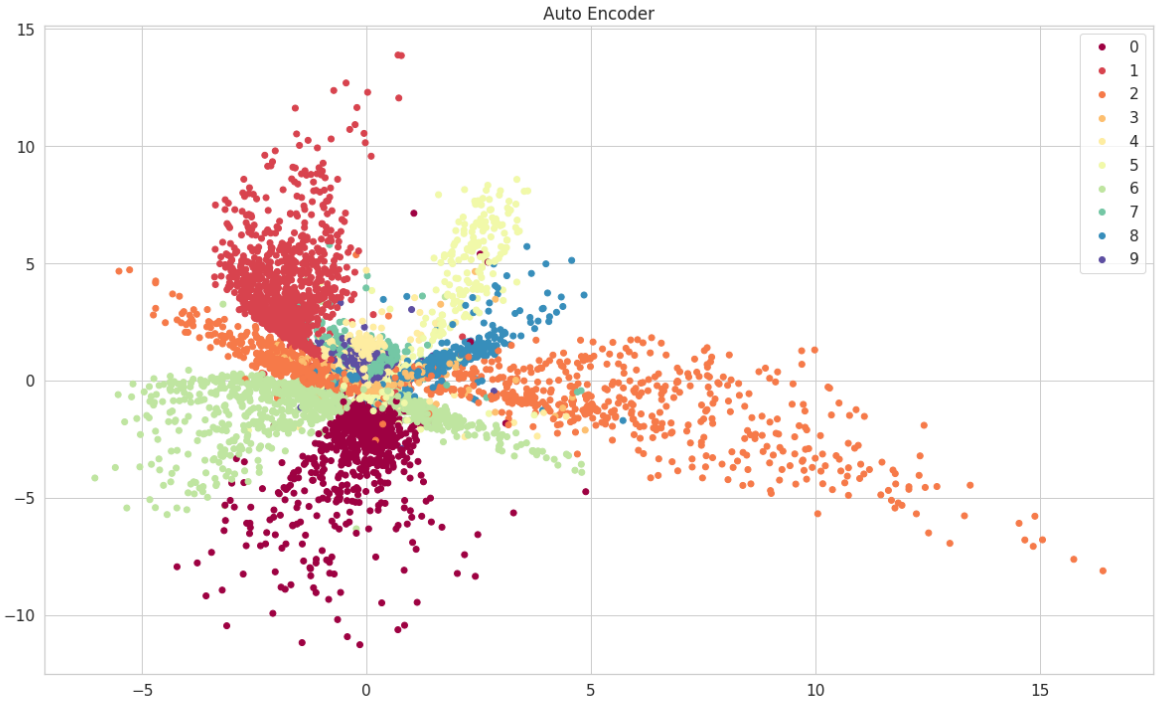


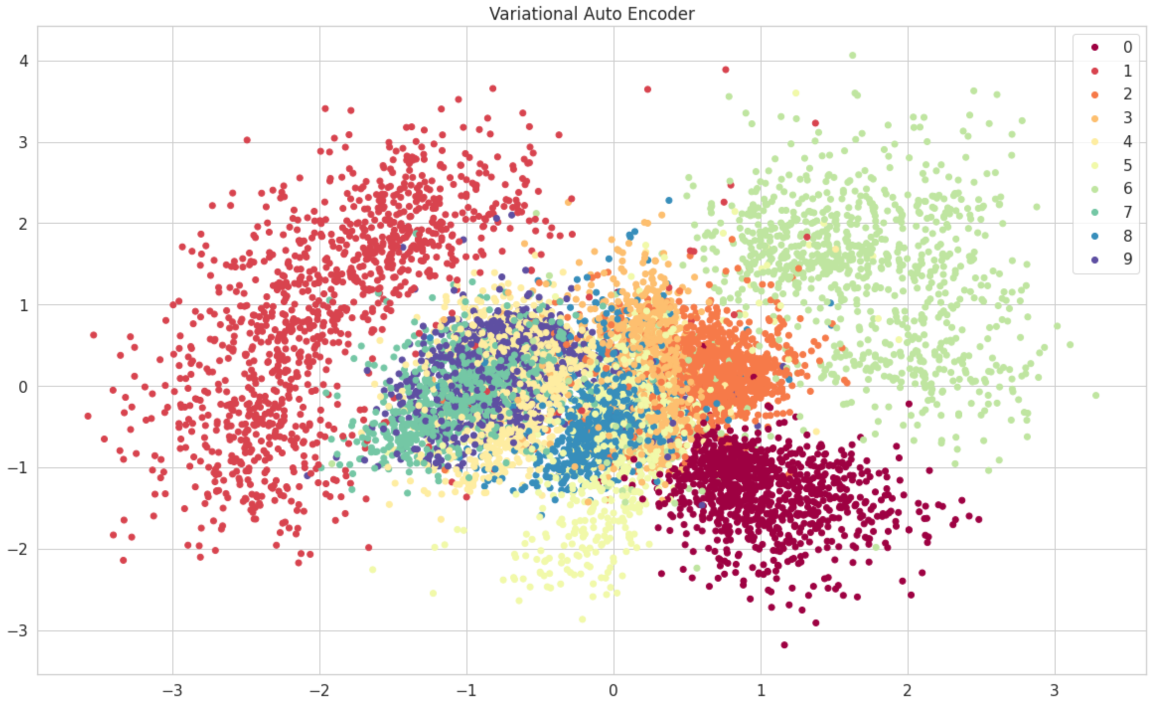


Latent Space

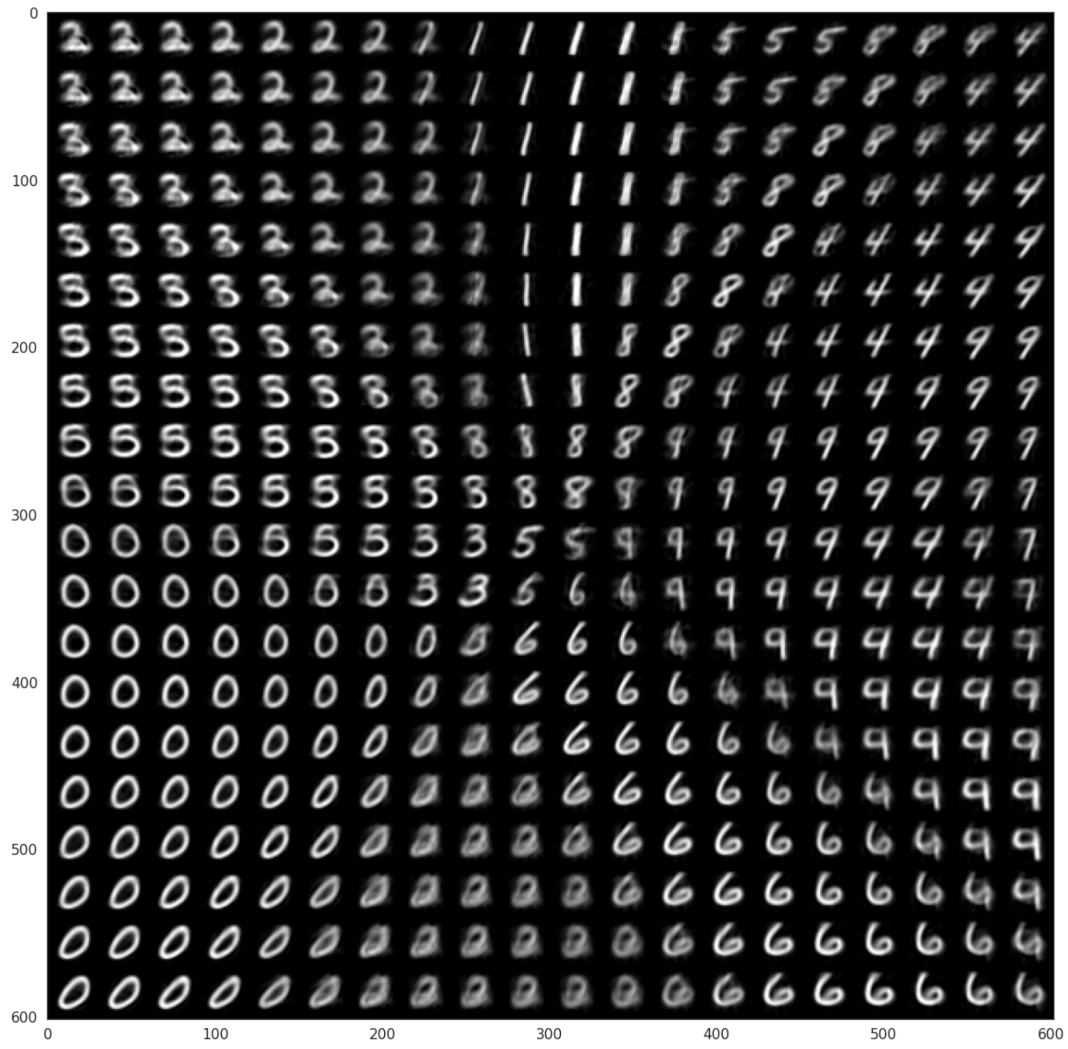


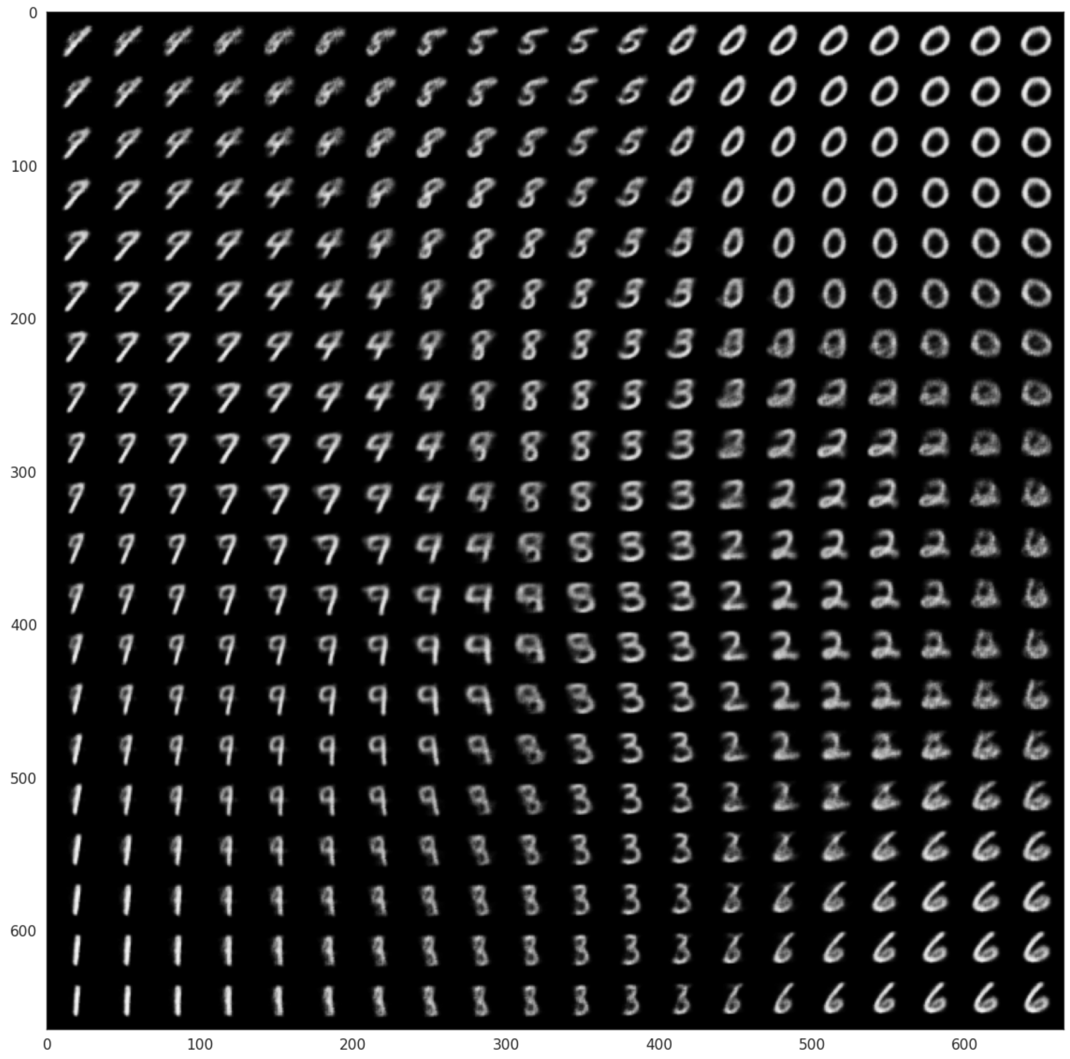
Latent Space. AE vs VAE.





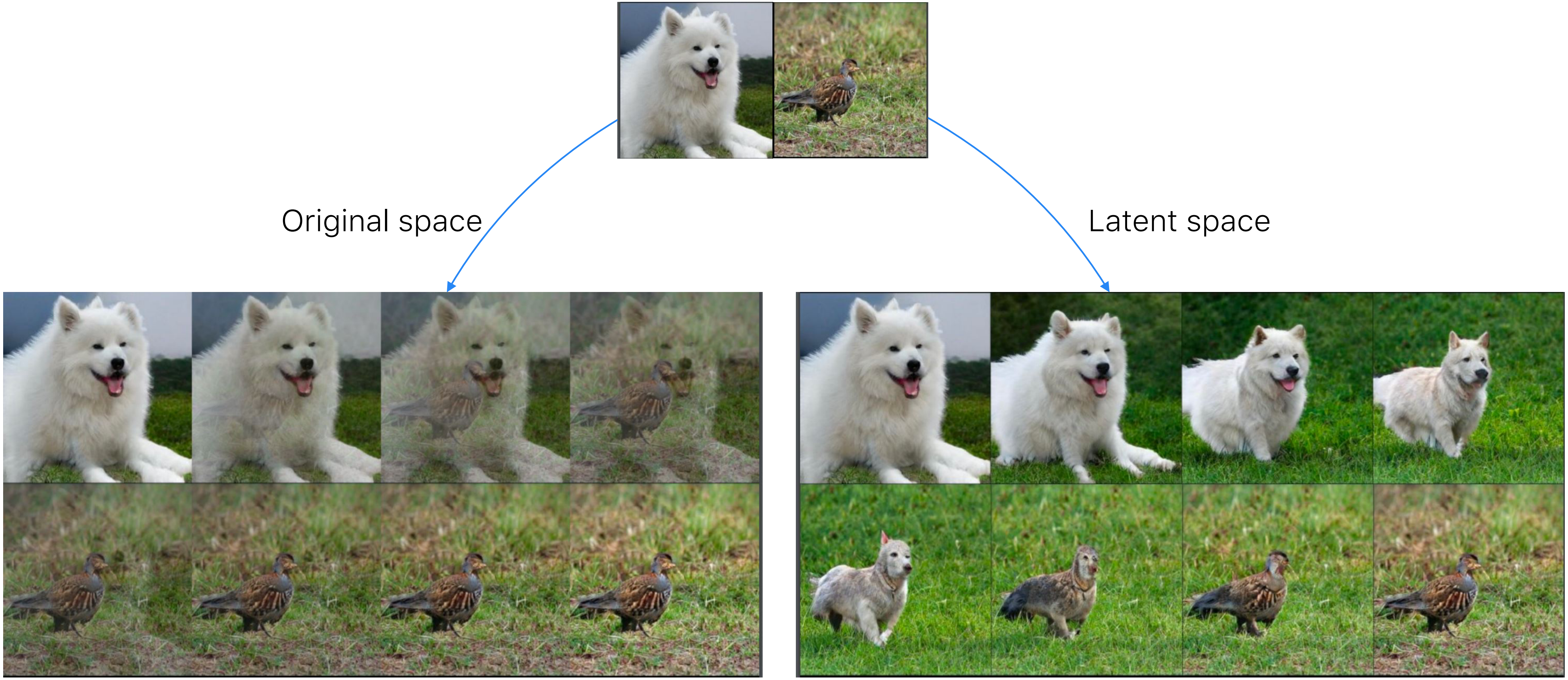
Latent Space. AE vs VAE.



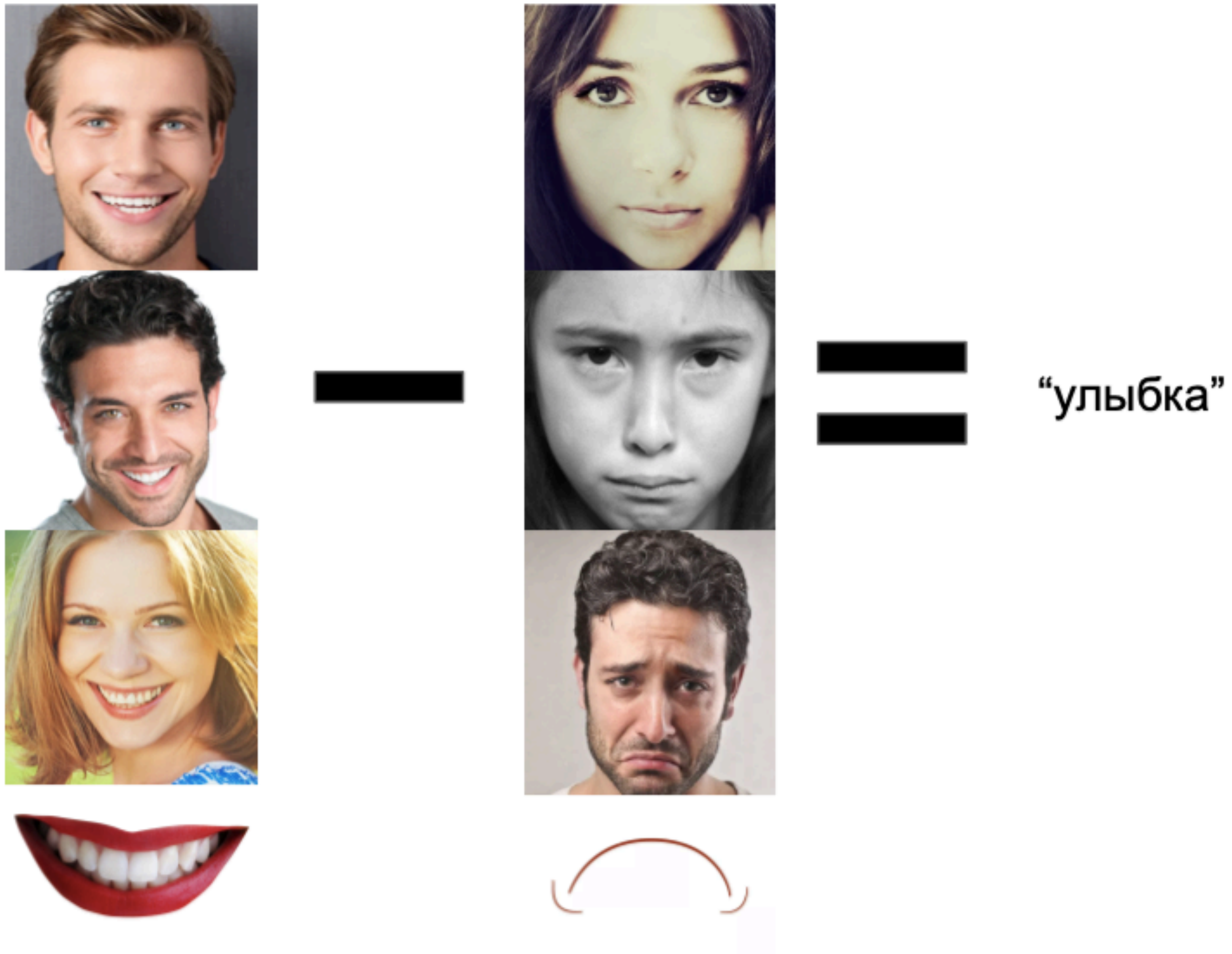


Latent Space


Latent Space



Latent space



Latent space