

Autoencoder e aplicações

Bruno Guimarães de Oliveira

19/11/2021

Sumário

1. O que são?

- Definição
- Arquitetura

2. Tipos de autoencoders

- Sparse Autoencoders
- Denoising Autoencoders
- Variational Autoencoders

3. Aplicações

- Redução de dimensionalidade
- Remoção de ruído
- Recuperação de imagens
- Unsupervised pretraining
- Sistemas de recomendação

4. Referências

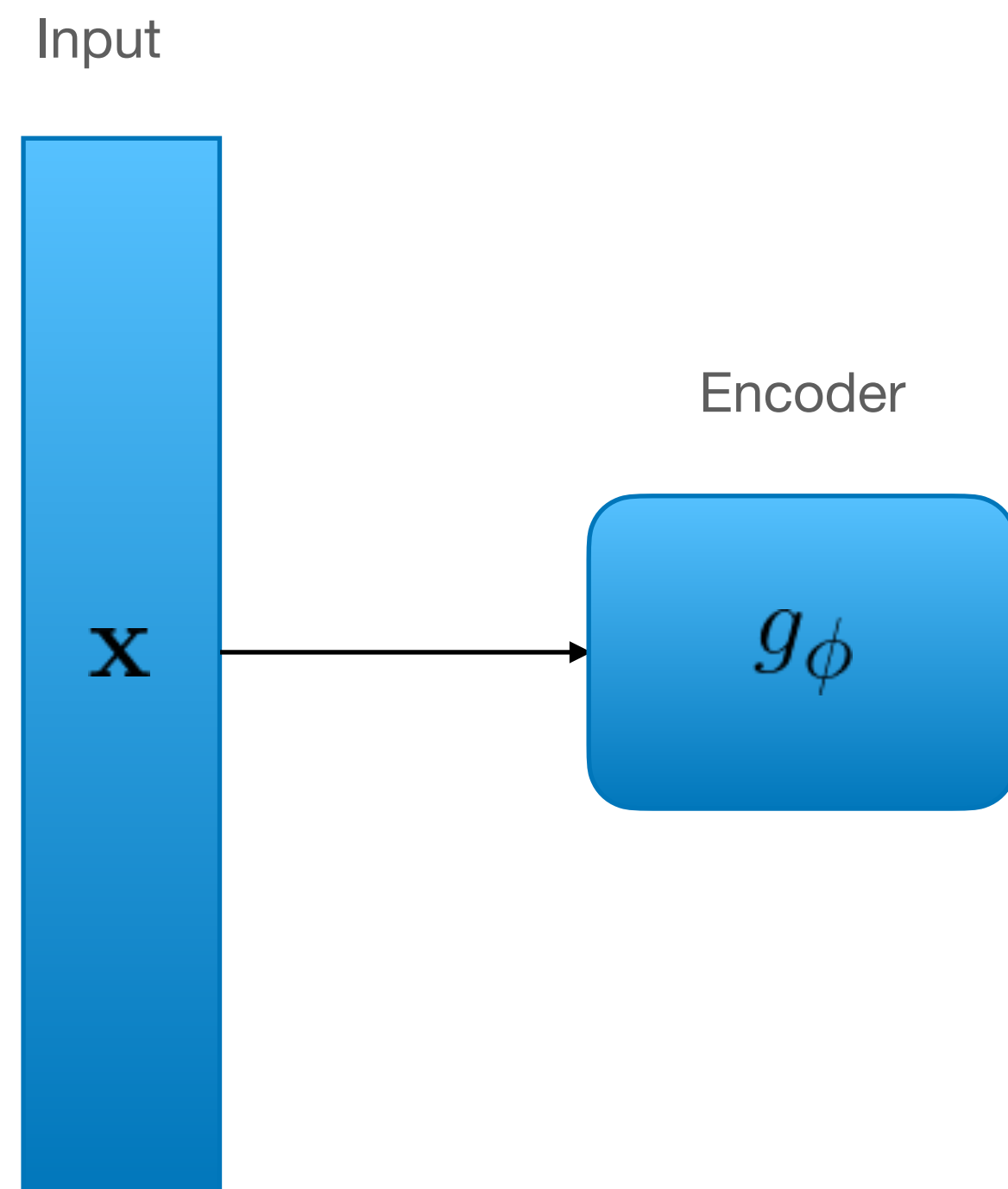
O que são?

Definição

- Um Autoencoder é uma rede neural que é treinada para ter como output uma cópia do seu input, ou seja, a camada de saída tem a mesma dimensionalidade da camada de entrada.
- A ideia é que essa rede neural consiga reconstruir cada dimensão do input.
- É composta de três partes:
 - Encoder: Rede neural feedforward e densa que recebe um input e comprime ele em uma representação, no chamado espaço latente, tentando manter ao máximo as características necessárias para essa representação poder ser usada para reconstruir o input original.
 - Vetor latente: Compressão do input.
 - Decoder: Rede neural feedforward que recebe os embeddings gerados pelo encoder e tem como missão reconstruir o input original.

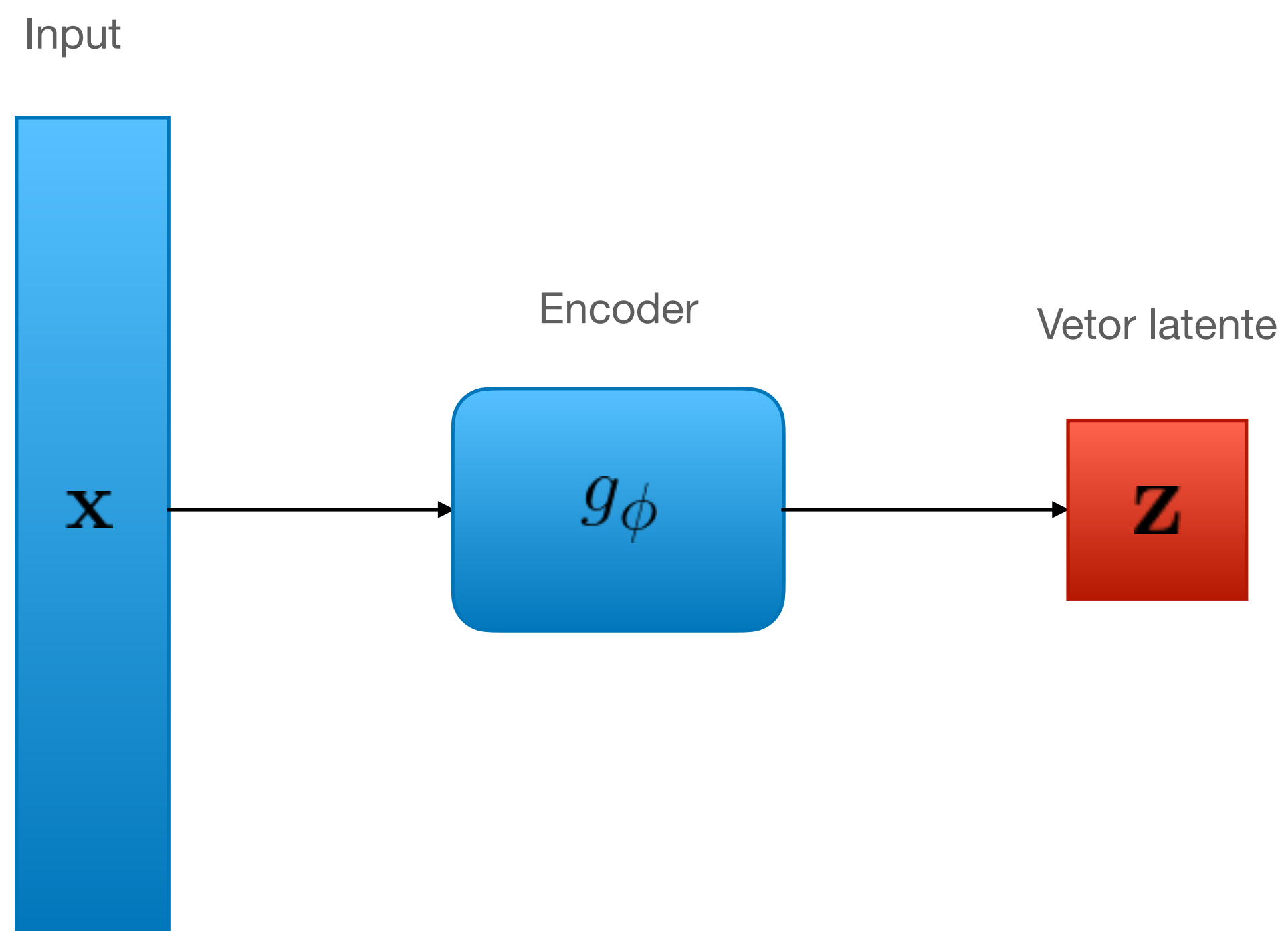
O que são?

Arquitetura



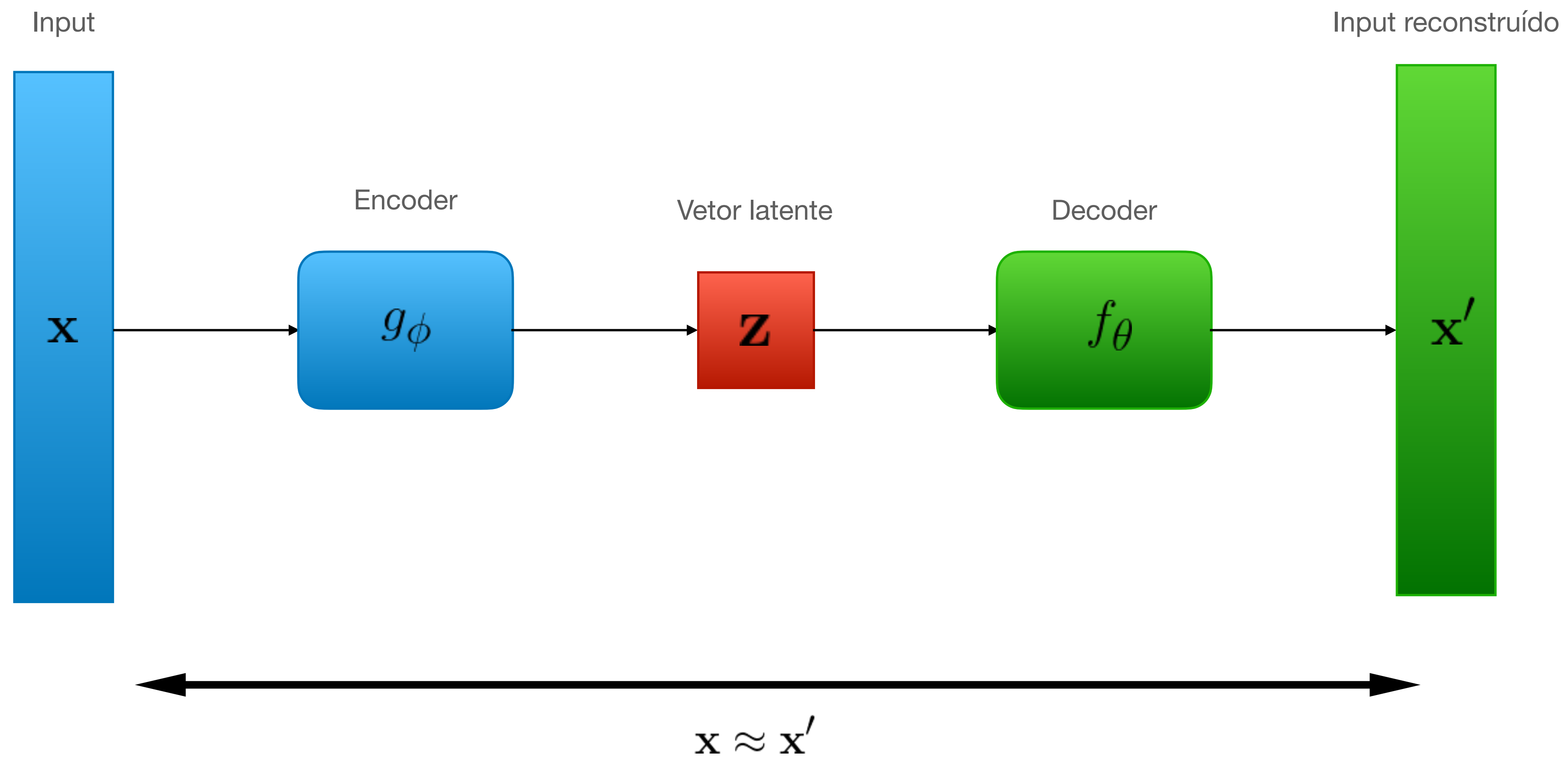
O que são?

Arquitetura



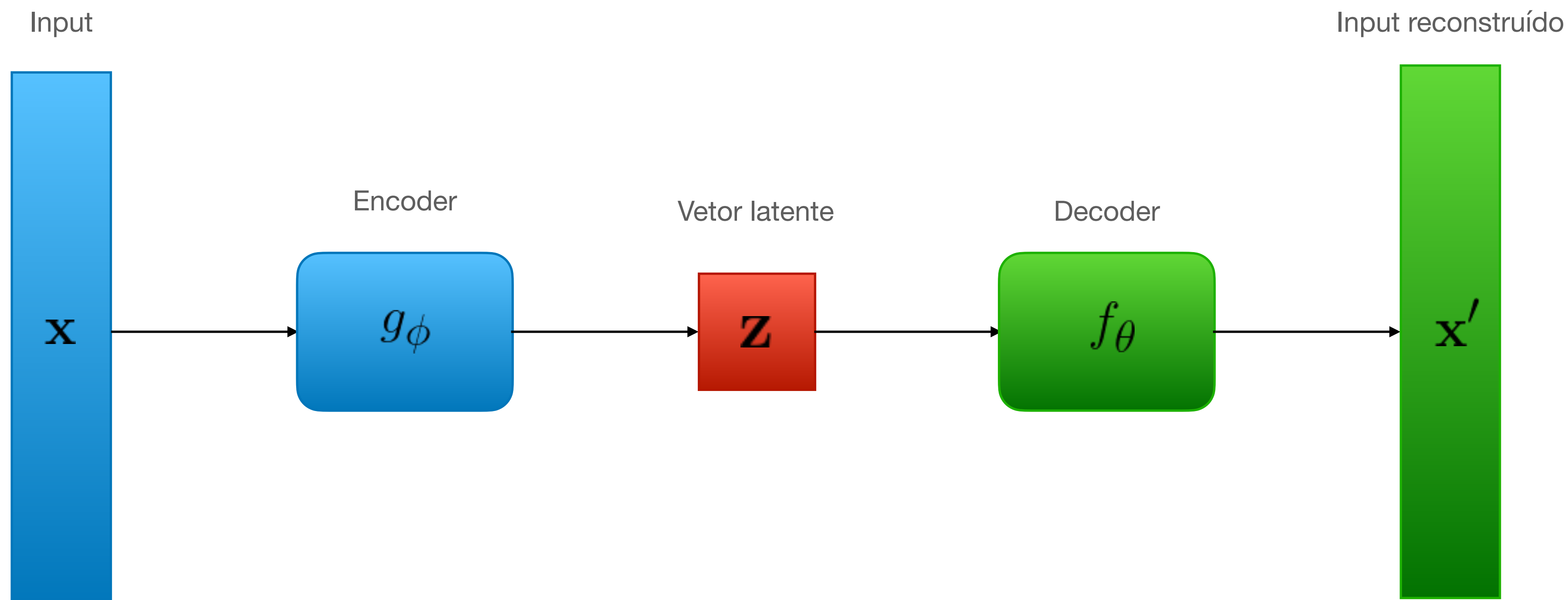
O que são?

Arquitetura



O que são?

Arquitetura



$$L(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|_2 = \frac{1}{n} \sum_{i=1}^n \left(\mathbf{x}^{(i)} - \mathbf{x}'^{(i)} \right)^2 \quad \text{Onde} \quad \mathbf{x}' = f_\theta(g_\phi(\mathbf{x}))$$

O que são?

Arquitetura

- Dada essa estrutura, podemos ter dois casos:
 - Undercomplete: $\dim(\mathbf{z}) \leq \dim(\mathbf{x})$
 - Overcomplete: $\dim(\mathbf{z}) \geq \dim(\mathbf{x})$
- Em ambos os casos, é importante controlar a complexidade do modelo.
- Podemos usar algum tipo de regularização.

Tipos de Autoencoder

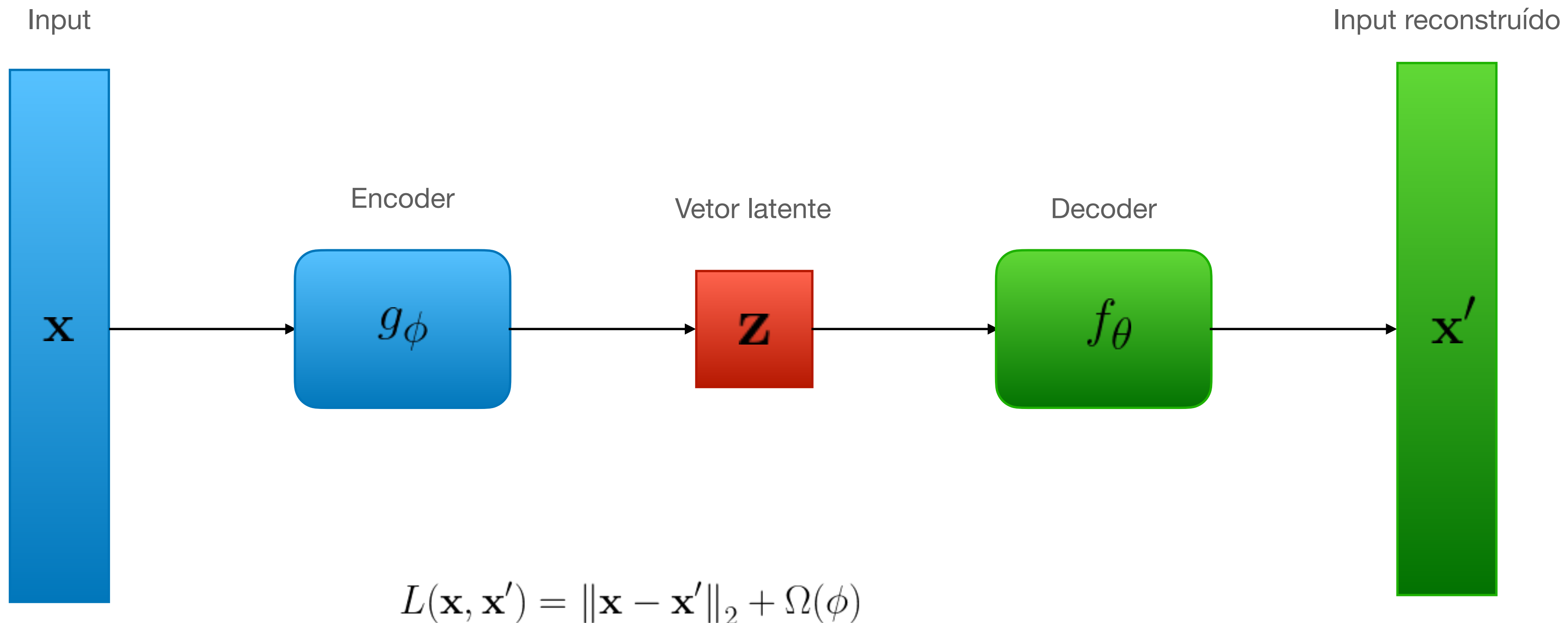
Sparse Autoencoders

- Temos a mesma estrutura de um autoencoder tradicional porém agora adicionamos regularização em nossa função de perda afim.

Tipos de Autoencoder

Sparse Autoencoders

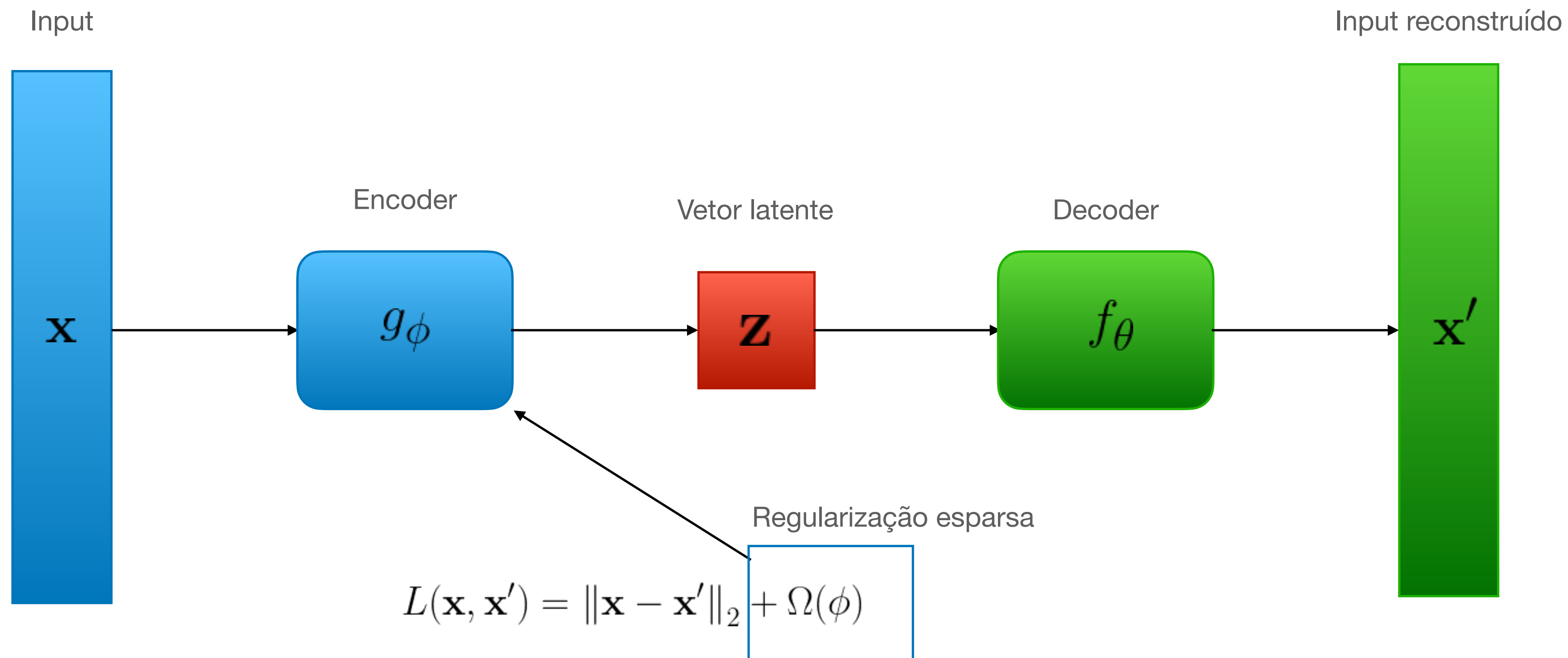
- Temos a mesma estrutura de um autoencoder tradicional porém agora adicionamos regularização em nossa função de perda afim.



Tipos de Autoencoder

Sparse Autoencoders

- Temos a mesma estrutura de um autoencoder tradicional porém agora adicionamos regularização em nossa função de perda afim.



Tipos de Autoencoder

Denoising autoencoder

- O input é corrompido por um ruído ou mascarando alguns valores do input de maneira estocástica.

Input

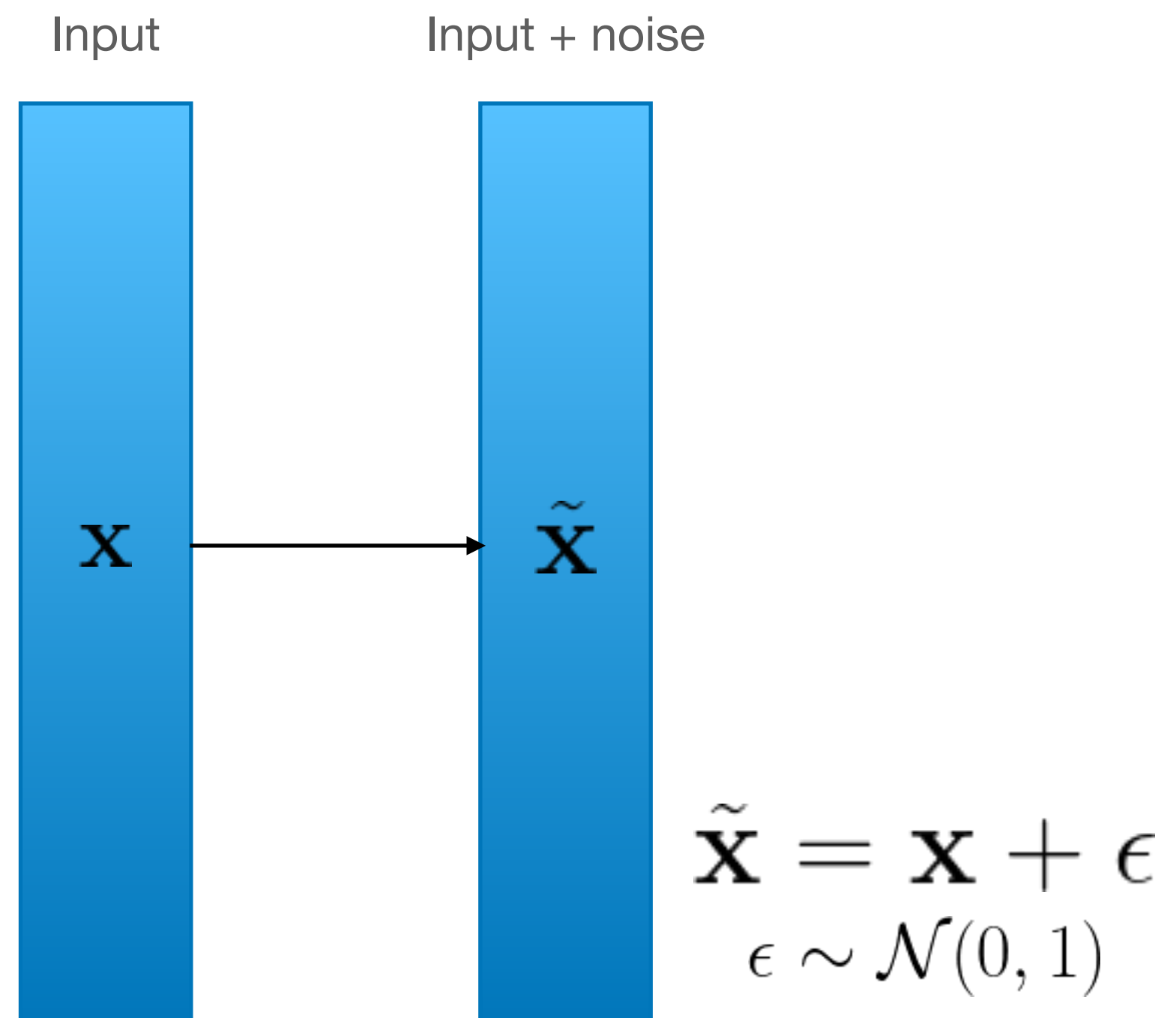


X

Tipos de Autoencoder

Denoising autoencoder

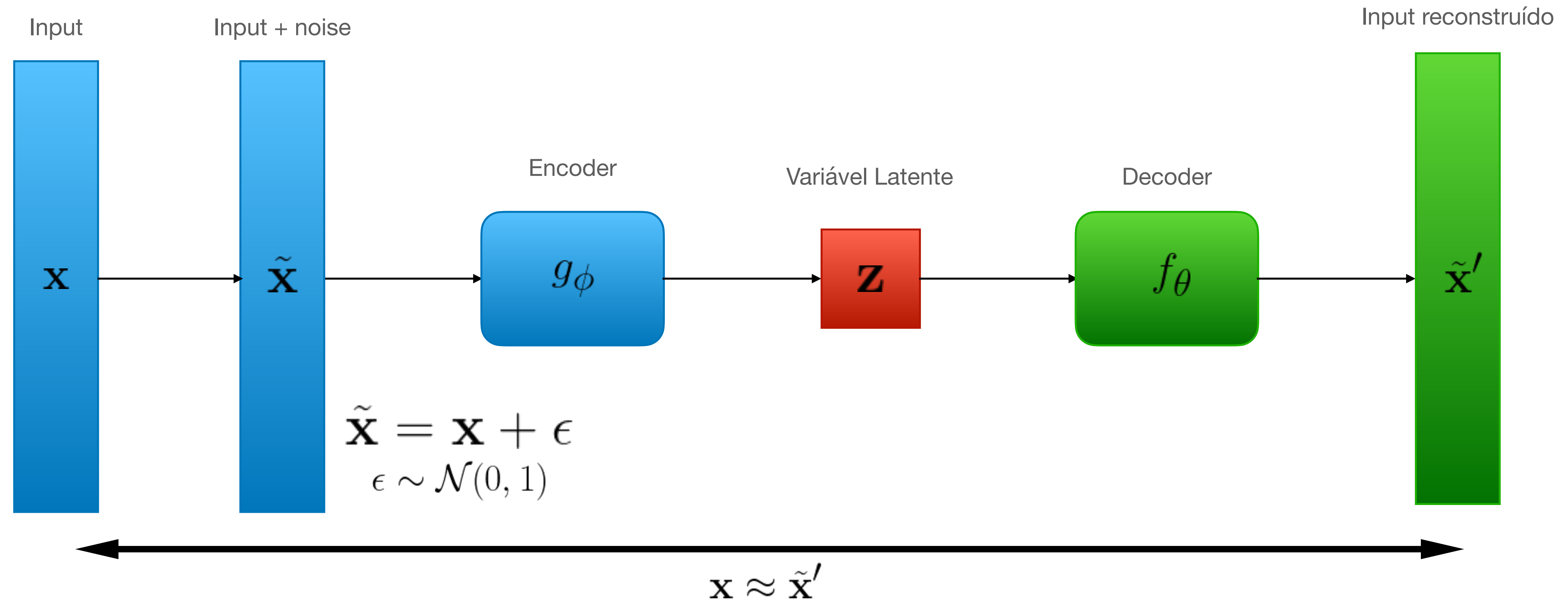
- O input é corrompido por um ruído ou mascarando alguns valores do input de maneira estocástica.



Tipos de Autoencoder

Denoising autoencoder

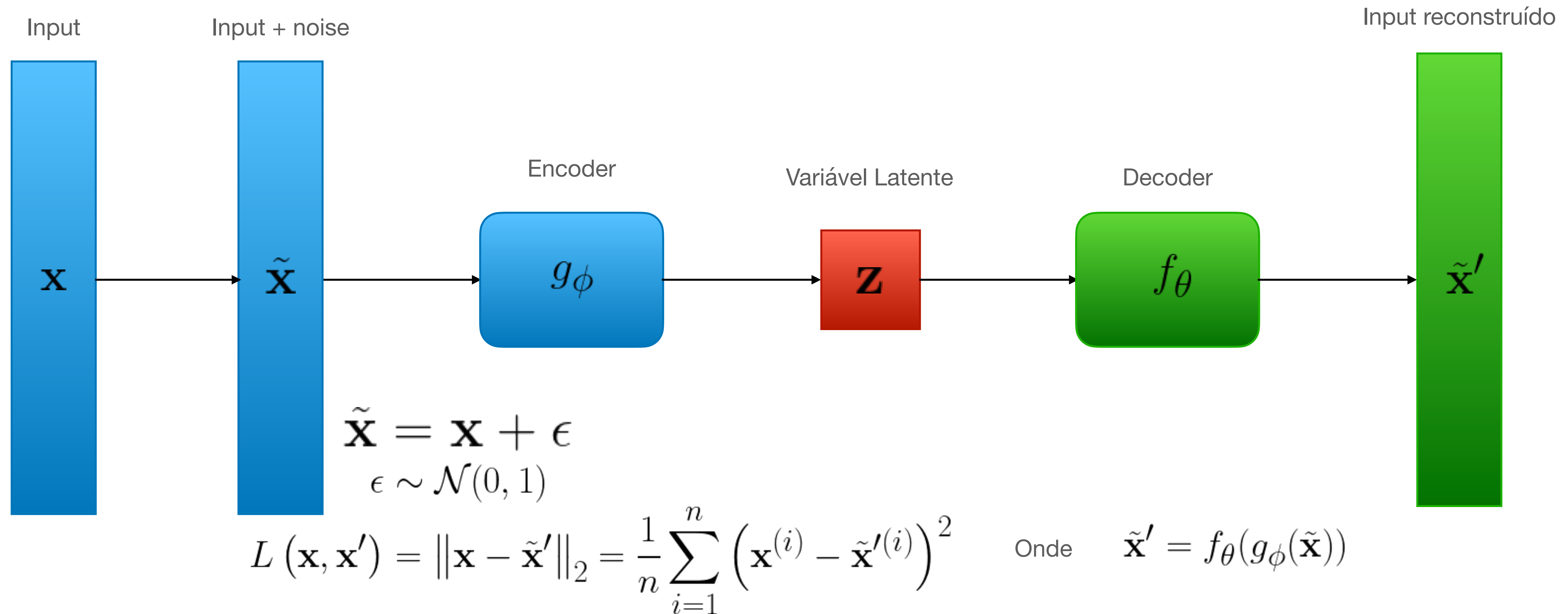
- O input é corrompido por um ruído ou mascarando alguns valores do input de maneira estocástica.



Tipos de Autoencoder

Denoising autoencoder

- Perda é computada levando em consideração o input original não o input ruidoso.



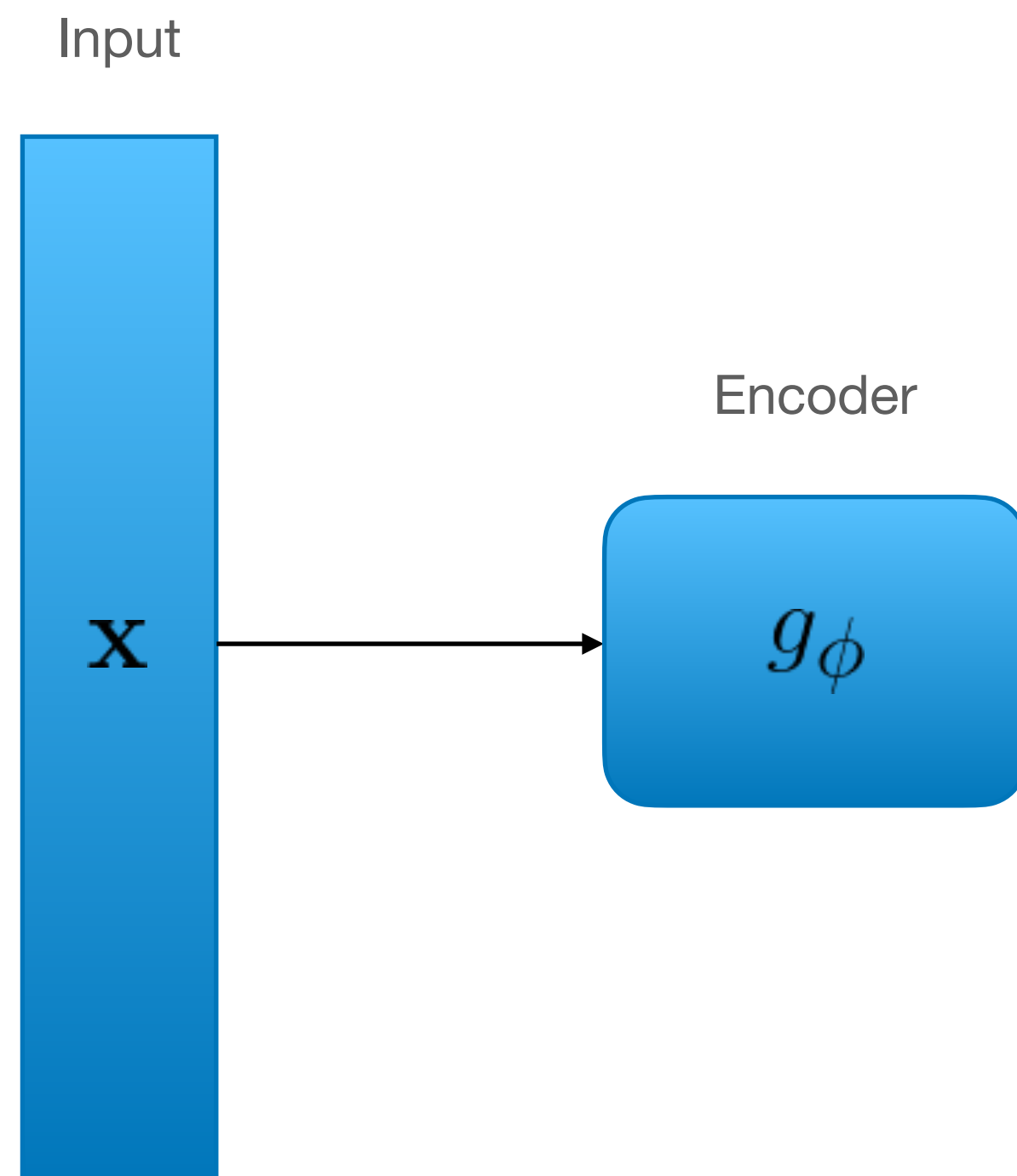
Tipos de Autoencoder

Variational Autoencoder

- Modelo generativo, ou seja, em contrapartida ao Autoencoder tradicional conseguimos criar dados novos.
- Em vez de mapear os dados de input para um vetor fixo, queremos agora mapear os dados em uma distribuição, ou seja, o output do encoder do VAE são os parâmetros de uma distribuição pré definida no espaço latente.
- O VAE impõe uma restrição ao espaço latente forçando essa distribuição a ser normal padrão, podemos pensar nessa restrição como sendo uma espécie de regularização.

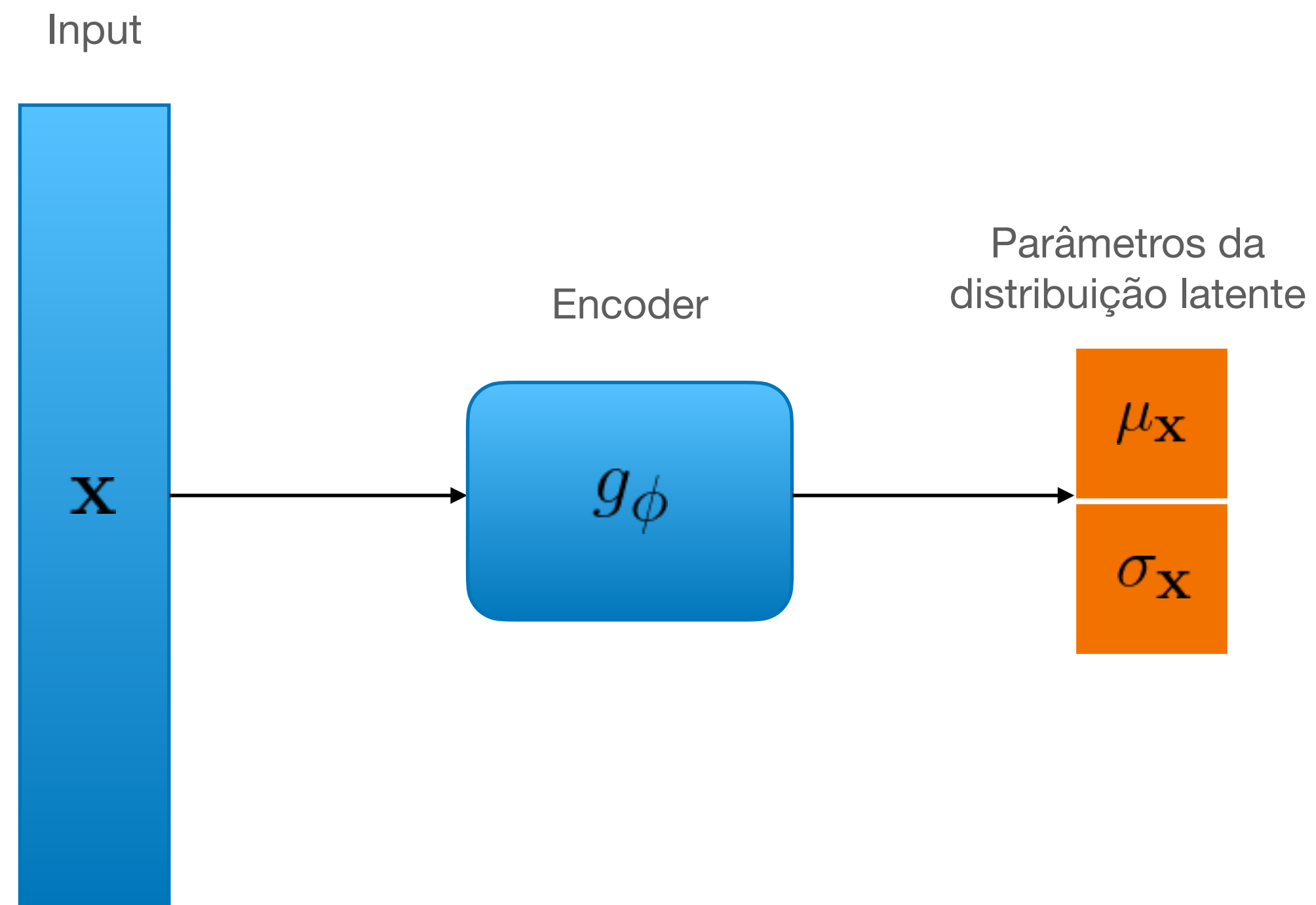
Tipos de Autoencoder

Variational Autoencoder



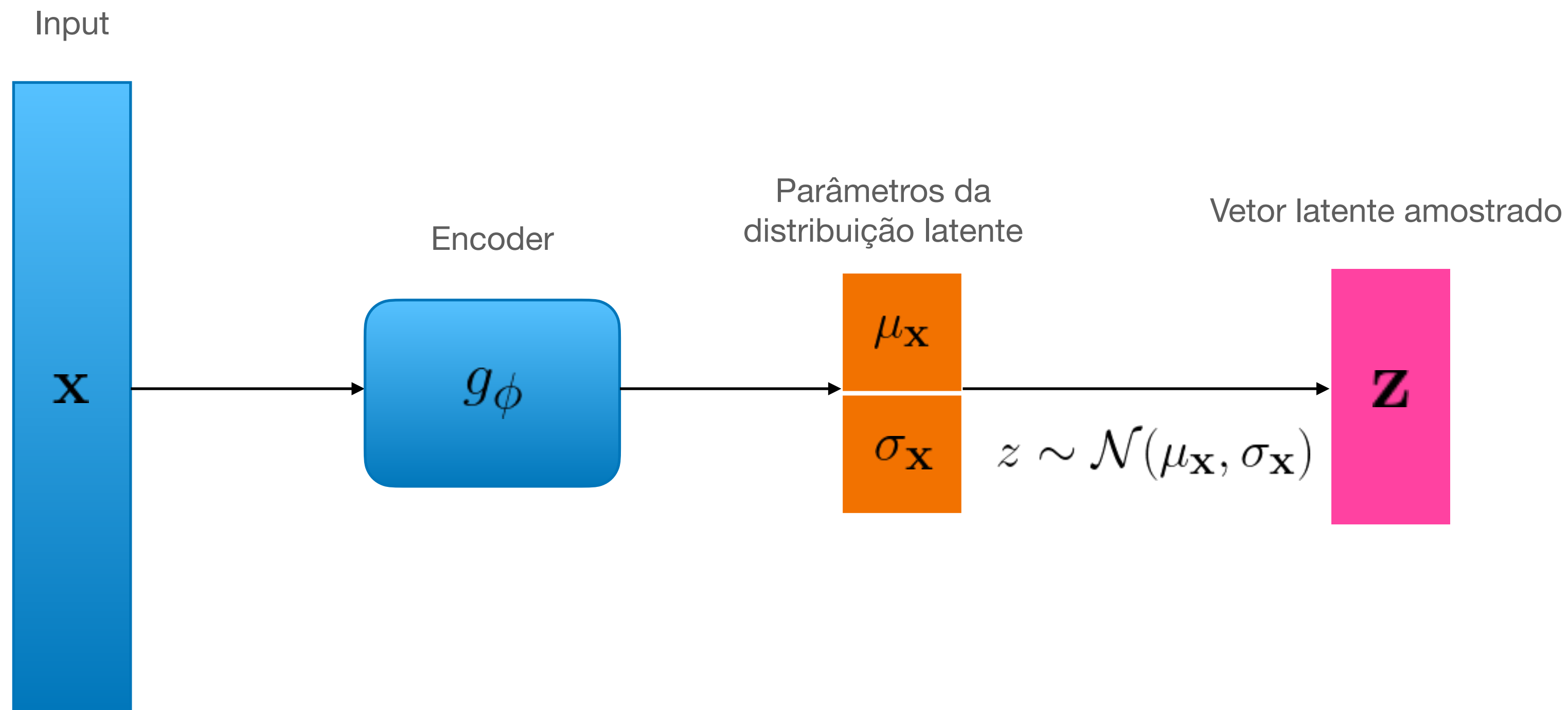
Tipos de Autoencoder

Variational Autoencoder



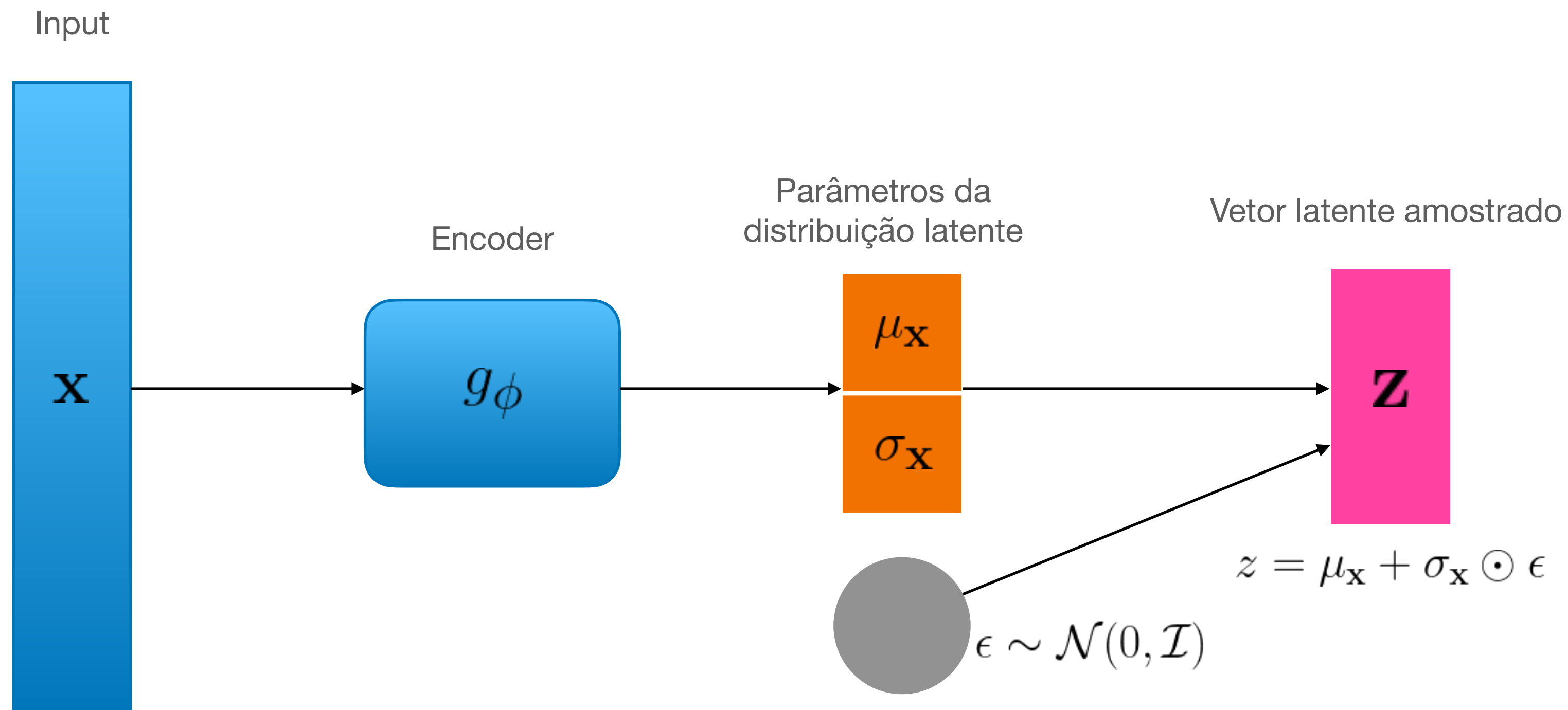
Tipos de Autoencoder

Variational Autoencoder



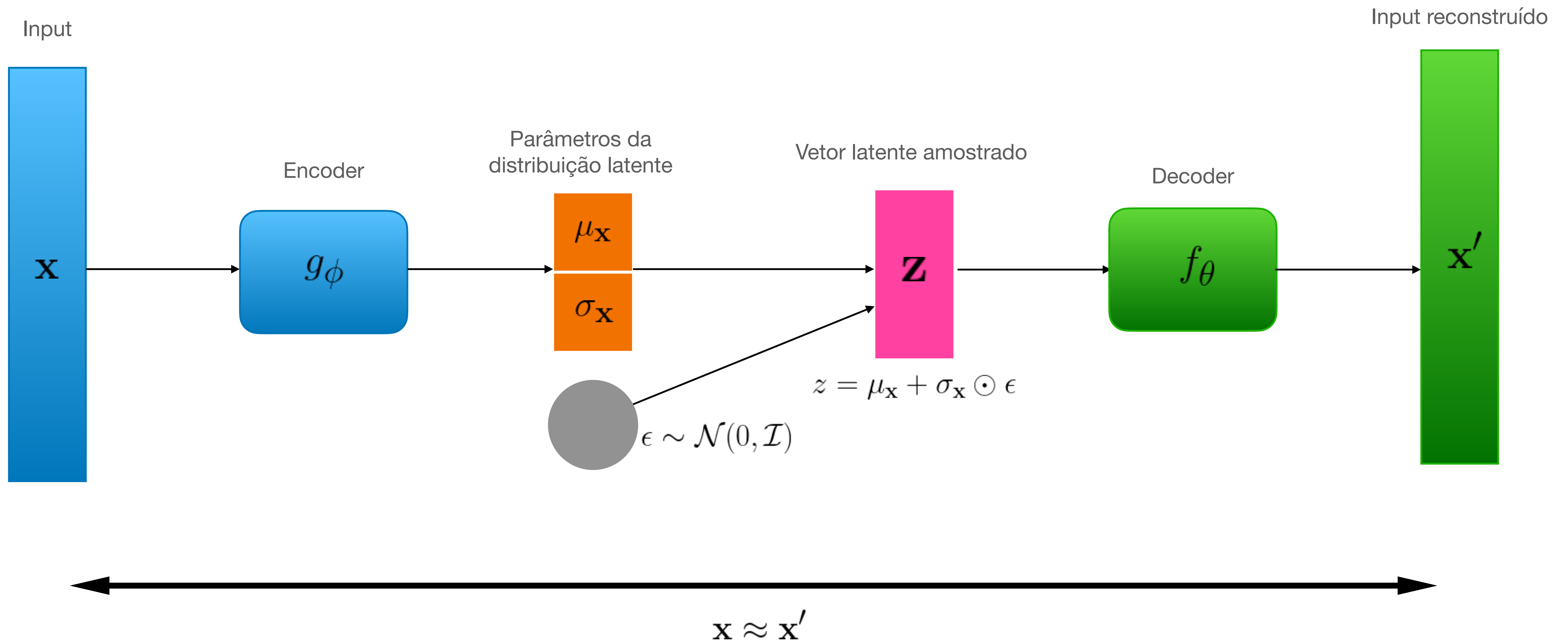
Tipos de Autoencoder

Variational Autoencoder



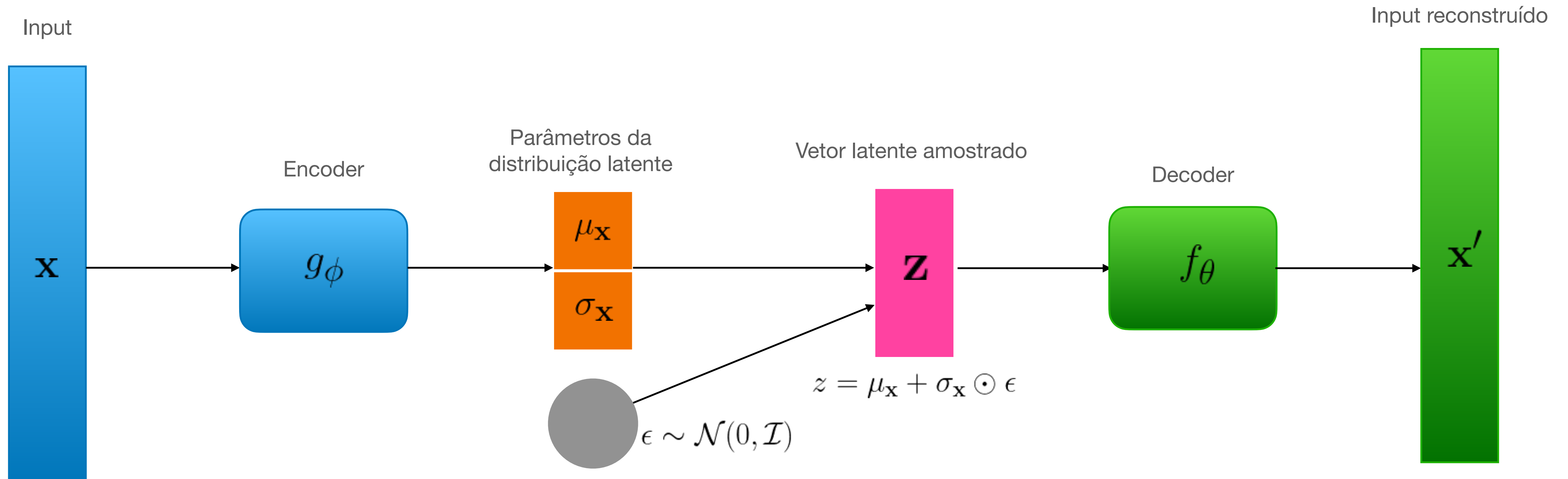
Tipos de Autoencoder

Variational Autoencoder



Tipos de Autoencoder

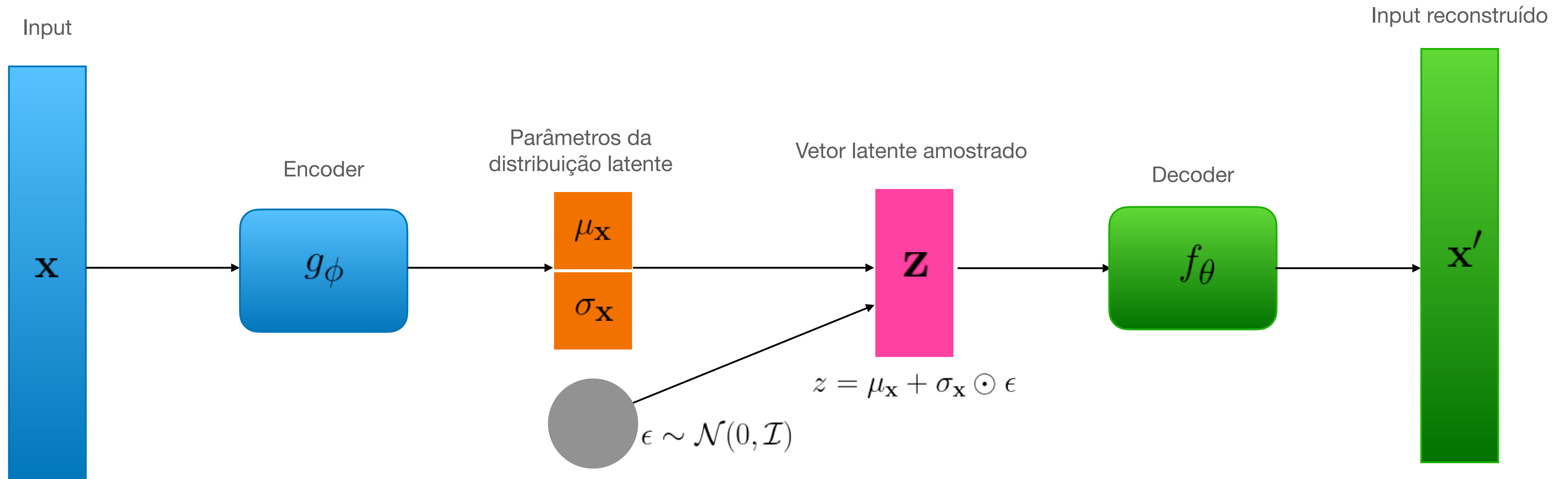
Variational Autoencoder



$$L(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|_2 + D_{KL}(\mathcal{N}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}) \parallel \mathcal{N}(0, \mathcal{I}))$$

Tipos de Autoencoder

Variational Autoencoder

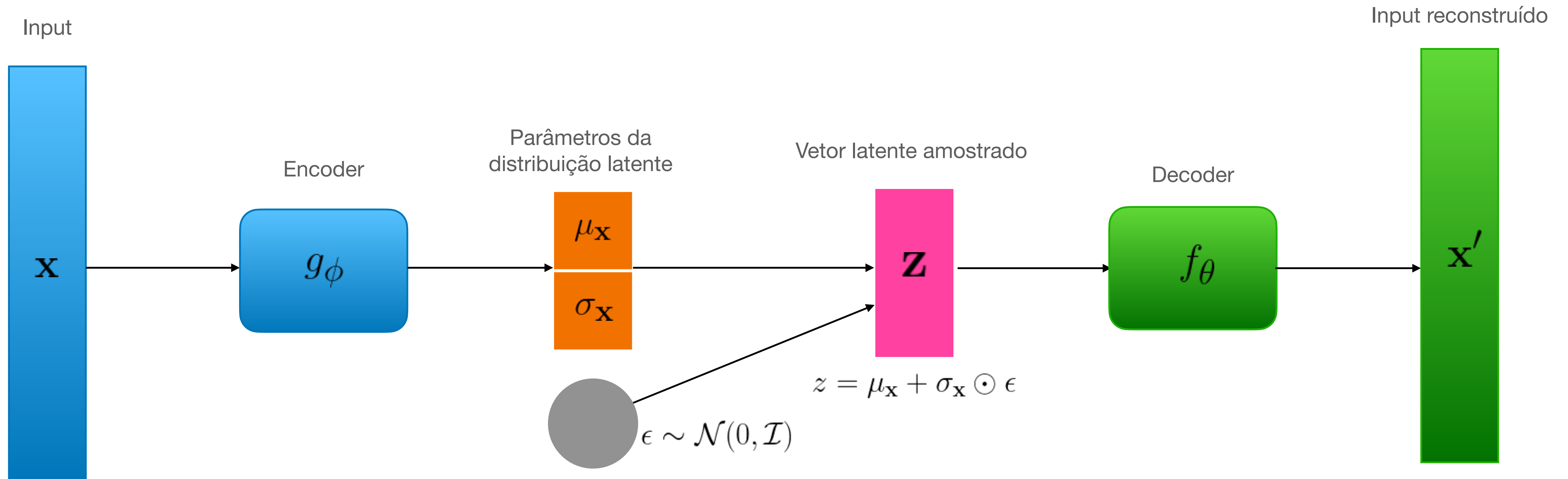


Função perda de reconstrução

$$L(\mathbf{x}, \mathbf{x}') = \boxed{\|\mathbf{x} - \mathbf{x}'\|_2} + D_{KL}(\mathcal{N}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}) \parallel \mathcal{N}(0, \mathcal{I}))$$

Tipos de Autoencoder

Variational Autoencoder



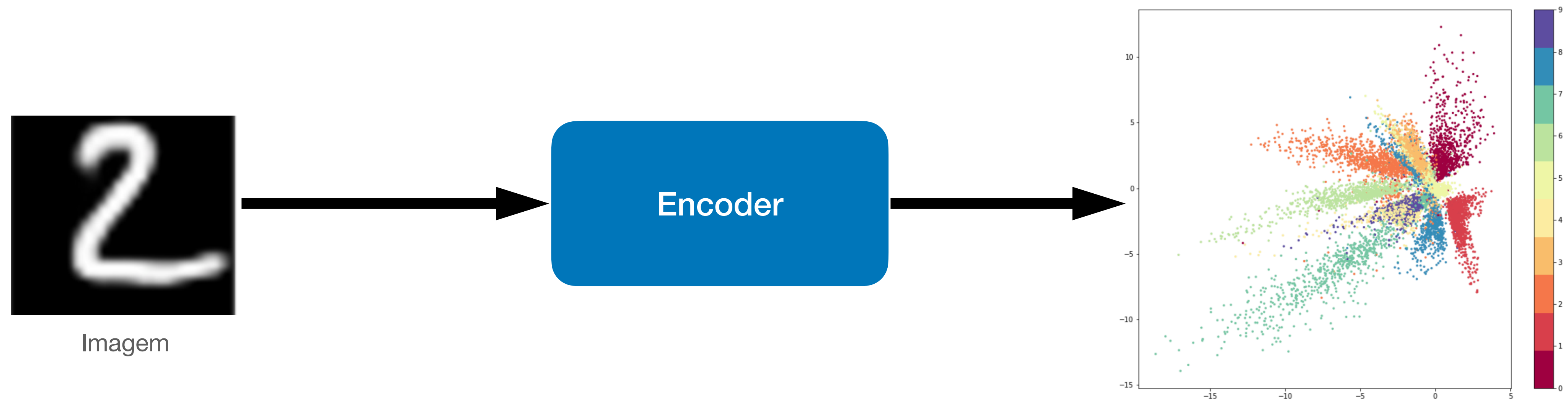
Função perda de similaridade

$$L(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|_2 + D_{KL}(\mathcal{N}(\mu_{\mathbf{x}}, \sigma_{\mathbf{x}}) \| \mathcal{N}(0, \mathcal{I}))$$

Aplicações

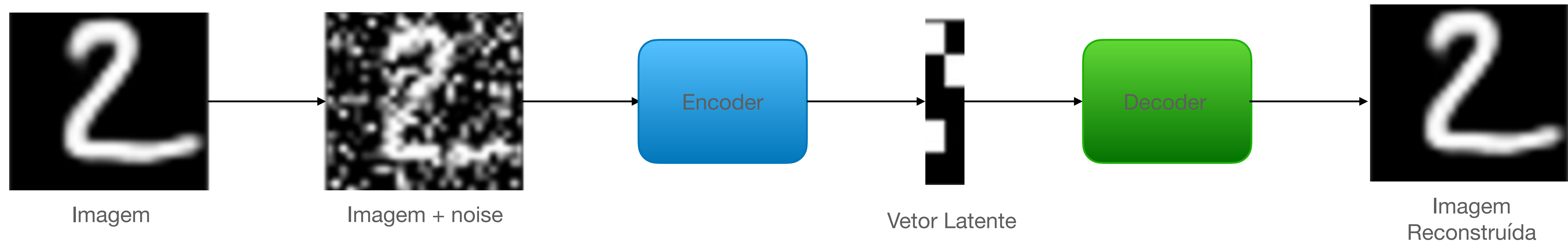
Redução de dimensionalidade

- Utilizamos apenas o encoder já treinado.
- Diferente do PCA, consegue capturar não linearidade nos dados, pois sempre conseguimos usar funções de ativações não lineares e adicionar mais camadas.



Aplicações

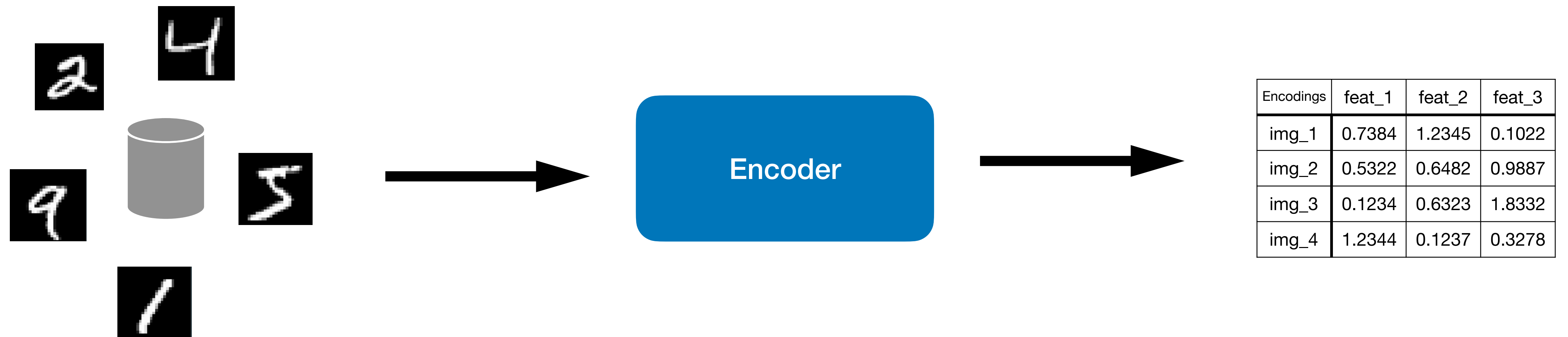
Remoção de ruído



Aplicações

Recuperação de Imagens

- Utilizamos apenas o encoder já treinado e extraímos todos os encodings de nosso database de imagens.



Aplicações

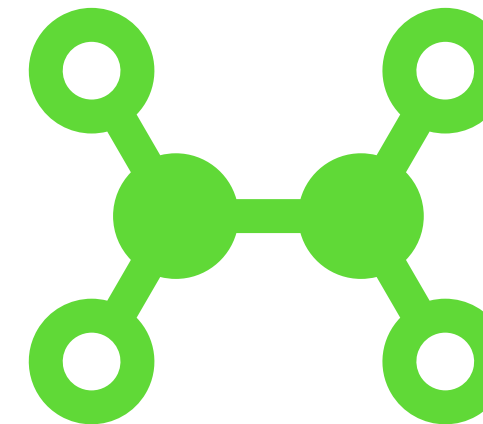
Recuperação de Imagens

- Treinamos um KNN com nosso novo dataset de encodings.

Encodings	feat_1	feat_2	feat_3
img_1	0.7384	1.2345	0.1022
img_2	0.5322	0.6482	0.9887
img_3	0.1234	0.6323	1.8332
img_4	1.2344	0.1237	0.3278



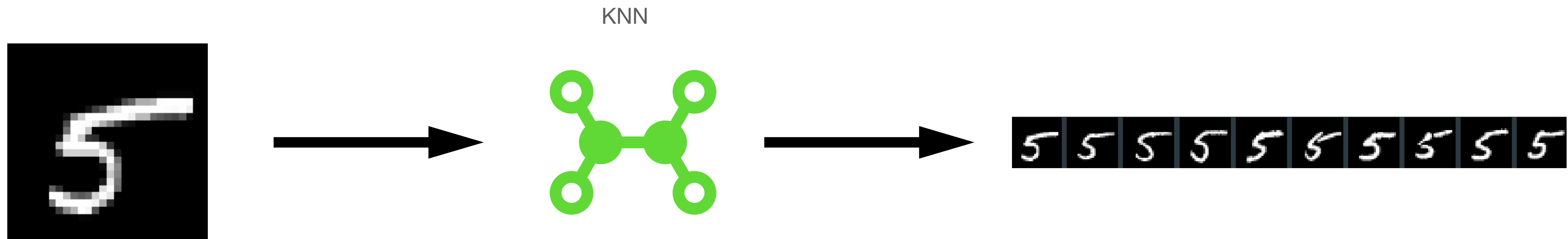
KNN



Aplicações

Recuperação de Imagens

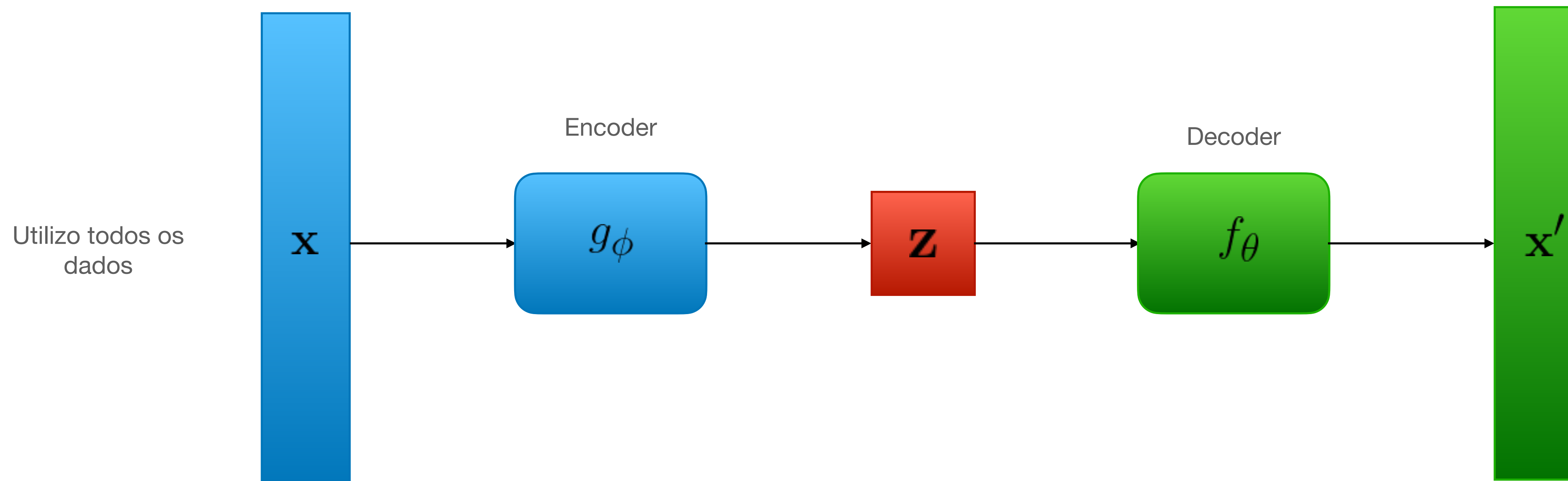
- Com uma imagem de referência conseguimos recuperar as imagens mais parecidas em nosso database.



Aplicações

Unsupervised pretraining

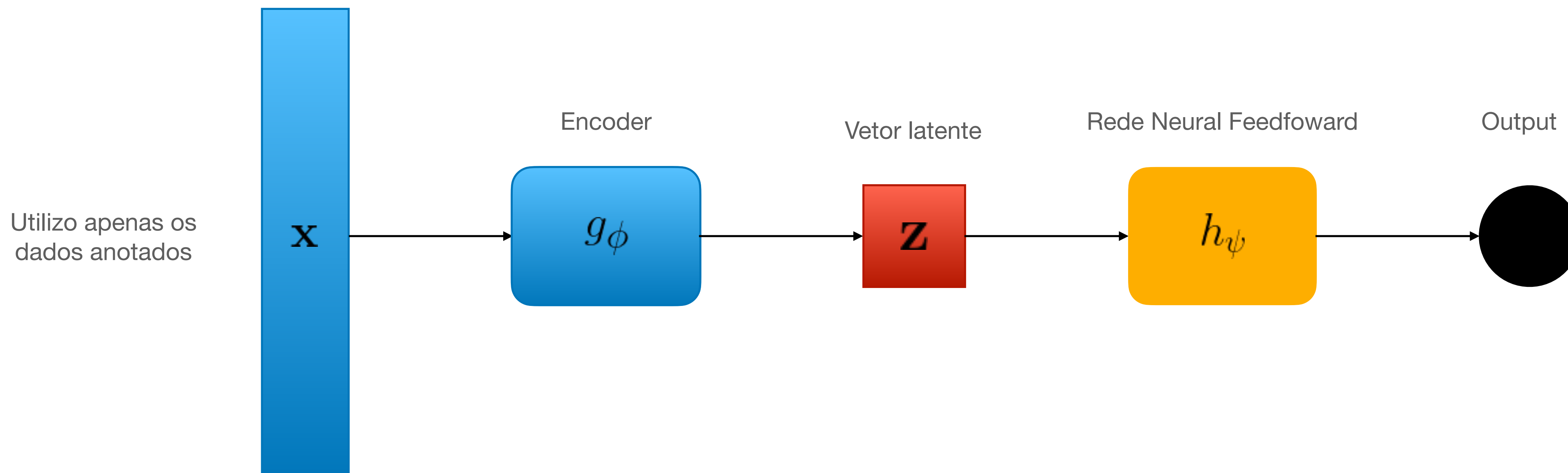
- Imagine um cenário em que você tenha muitos dados pra resolver uma task específica de Machine Learning porém apenas uma pequena porção desses dados estão anotados.



Aplicações

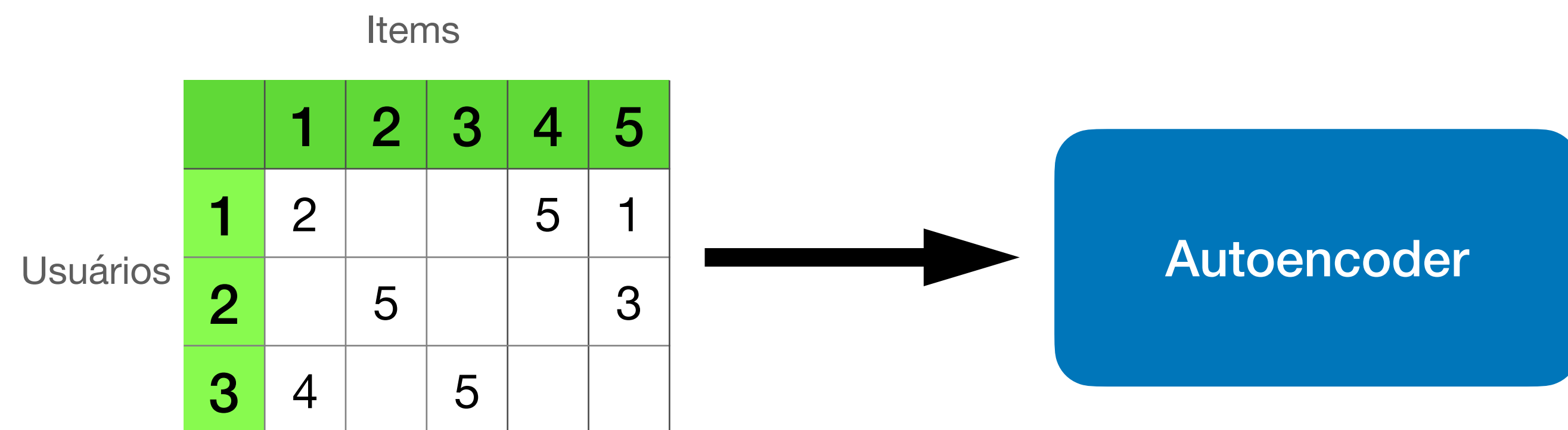
Unsupervised pretaining

- Imagine um cenário em que você tenha muitos dados pra resolver uma task específica de Machine Learning porém apenas uma pequena porção desses dados estão anotados.



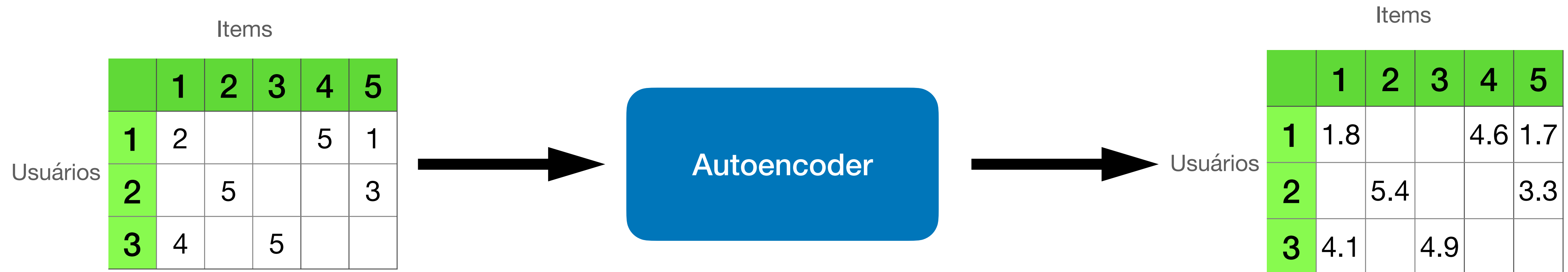
Aplicações

Sistemas de Recomendação



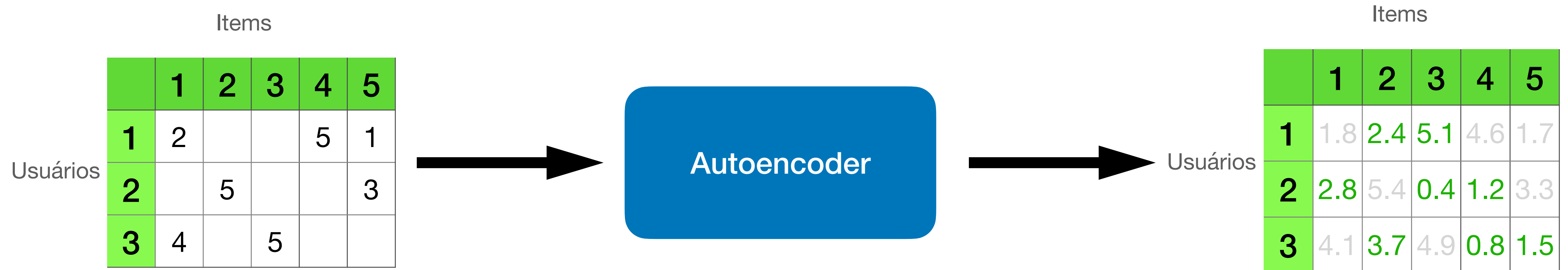
Aplicações

Sistemas de Recomendação



Aplicações

Sistemas de Recomendação



Referências

- [1] Ian Goodfellow, Deep Learning.
- [2] Aurélien Geron, Hands-On Machine Learning with Scikit-Learn, Keras, and Tensorflow.
- [3] Ferreira, D.; Silva, S.; Abelha, A.; Machado, J. Recommendation System Using Autoencoders. *Appl. Sci.* **2020**, *10*, 5510.
- [4] <https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73>
- [5] <https://becominghuman.ai/using-variational-autoencoder-vae-to-generate-new-images-14328877e88d>
- [6] <https://stats.stackexchange.com/questions/409995/why-is-random-sampling-a-non-differentiable-operation>
- [7] <https://www.pyimagesearch.com/2020/02/24/denoising-autoencoders-with-keras-tensorflow-and-deep-learning/>

Obrigado