

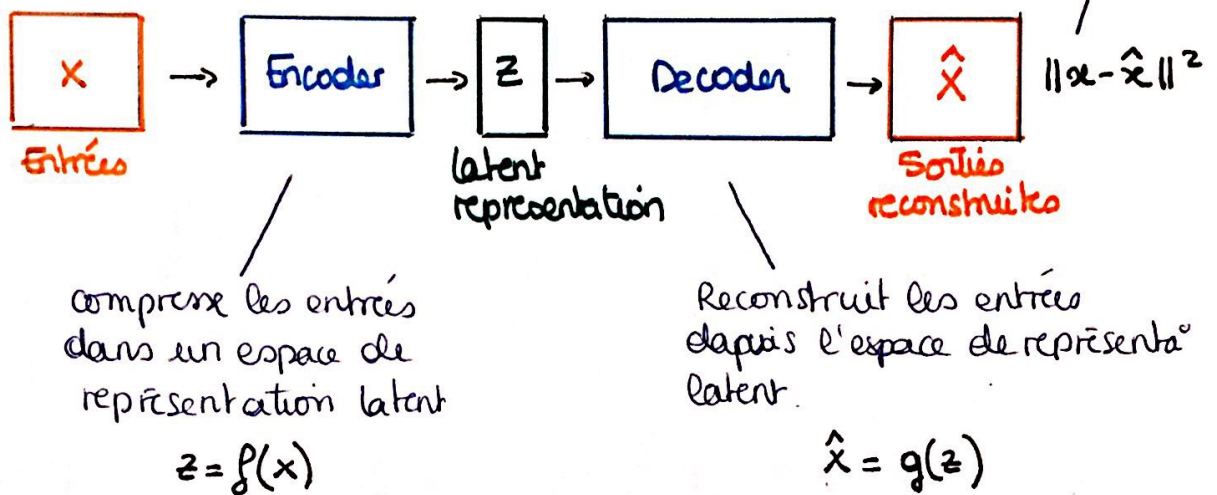
# Autoencoder (AE)

=> d'autoencoder est un réseau de neurones qui a pour but d'apprendre une représentation des données en compressant les entrées dans un espace latent réduit, puis en reconstruisant les entrées à partir de cette représentation réduite.

On ne peut pas :

- estimer la probabilité d'une donnée
- générer des données selon la distribution.

on veut minimiser l'erreur de reconstruction



On aimerait que  $z$  contienne des propriétés intéressantes de  $x$  :

- Contraindre  $z$  à avoir des dimensions plus petites que  $x$  (undercomplete)

=> on apprend les caractéristiques les + marquantes.

- Si la dimension de  $z$  est la même que  $x$  ou que la dim de la représentation latente est supérieure à celle de l'entrée  
=> même un codeur linéaire et un décodeur linéaire peuvent apprendre à copier l'entrée sur la sortie sans apprendre quoi que ce soit d'utile sur la distribution des données.

Architecture = spécifique à chaque problème

- Encoder = ConvNet (marche très bien).
- Decoder = Déconvolution (soit par upsampling, soit par transposition convolution).

Utilisation =

- Denoising autoencoder
- Fine tuning.
- Détection d'outliers.