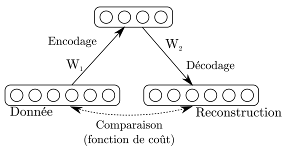
Autoencoders in a nutshell

Julien Velcin
M2 MALIA - Université Lumière Lyon 2

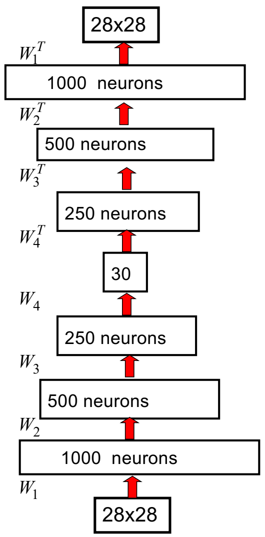
Autoencoder

- What's the objective of AE?
  - extract low-dimensional representation
- How ?
  - reconstruct (decode) data from the internal (latent) code
- Some features
  - PCA can be seen as a linear AE
  (see: https://arxiv.org/pdf/1804.10253.pdf)
  - possibility to stack encoding layers
  - improve the latent space (c.f., VAE)

Architecture of AE (1)

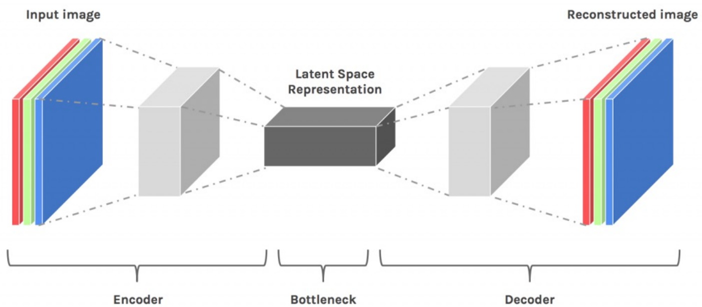


make is deep ->

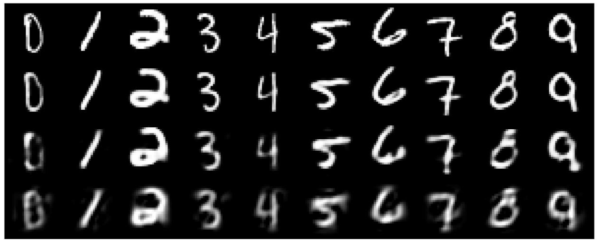


encoder : $p_{encode}(h|x)$
decoder: $p_{decode}(x|h)$

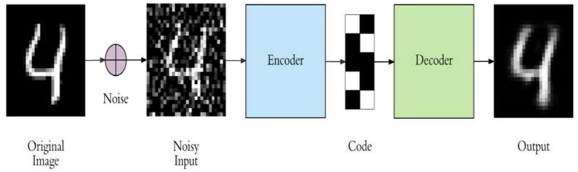
Architecture of AE (2)



Comparison
taken from CSC2535: 2013 (G. Hinton)



Denoising AE



source: https://medium.com/@harishr2301/denoising-autoencoders-996e866e5cd0

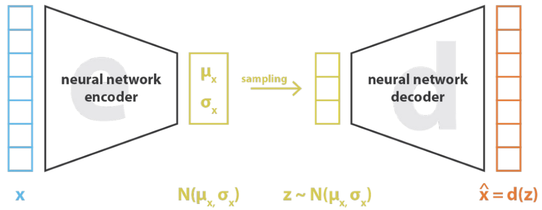
Sparse autoencoder

Adding a penalty on the code **h** :

$$L(x, g(f(x))) + \Omega(h)$$

For instance, $\Omega(h) = \lambda \sum_i |h_i|$

Variational AE



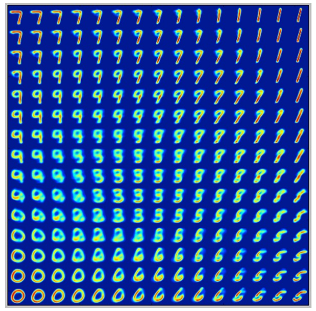
$$\text{loss} = \| x - \hat{x} \|^2 + \mathrm{KL}[N(\mu_x, \sigma_x), N(0, I)] = \| x - d(z) \|^2 + \mathrm{KL}[N(\mu_x, \sigma_x), N(0, I)]$$

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

VAE (con’t)

- Solution to get better latent spaces therefore better generation
- Two parts of the architecture:
  - encoder: p(z|x)
  - decoder: p(x|z)
- Optimization based on Variational Inference (VI)
- Implemented with the « reparametrization trick »
- The internal code (z) should follow a « simple » law (e.g., Gaussian)

VAE (con’t)



source: https://blog.keras.io/building-autoencoders-in-keras.html