Autoencoders

MLST PRESENTATION – 08/11/2017

FANNY ROCHE

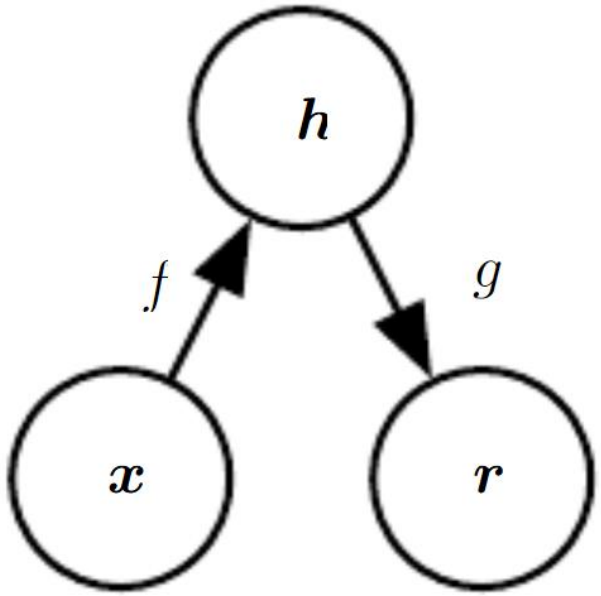
Autoencoders

Basic principle

- Objective : $r = g\big(f(x)\big) \approx x$
- BUT $r \neq x$!

  ➔ no Identity function !

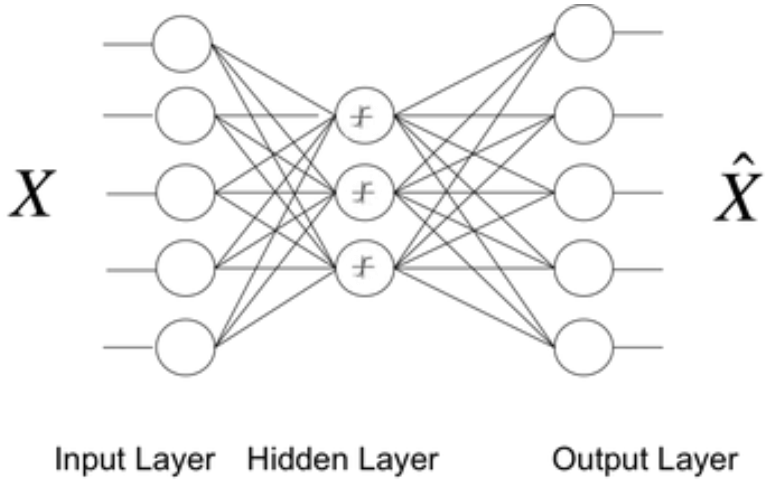
- f : encoder function
- g : decoder function
- h : latent features



Autoencoders

How does it work ?

- Encoder : extract useful properties of the data
- Decoder : reconstruct the input
- Undercomplete AE



Autoencoders

How does it work ?

- Special case of **feedforward** networks
- Use **gradient descent backpropagation**
- Comparison to **PCA** :
  - Linear + MSE
    → PCA decomposition
  - Nonlinear
    → more powerful nonlinear generalization

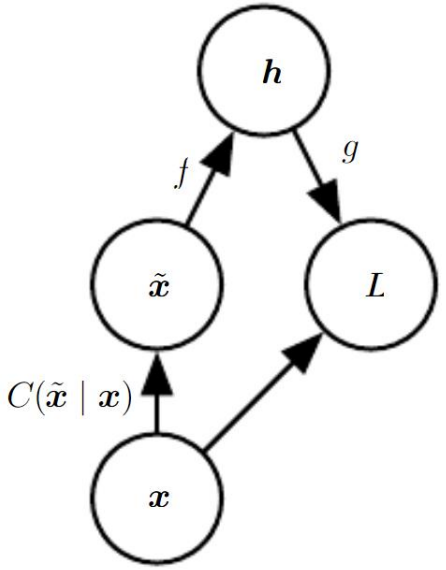
Autoencoders

History

- **Before** :
  - Dimensionality reduction
  - Feature Learning
- **Now :**
  - Generative Models

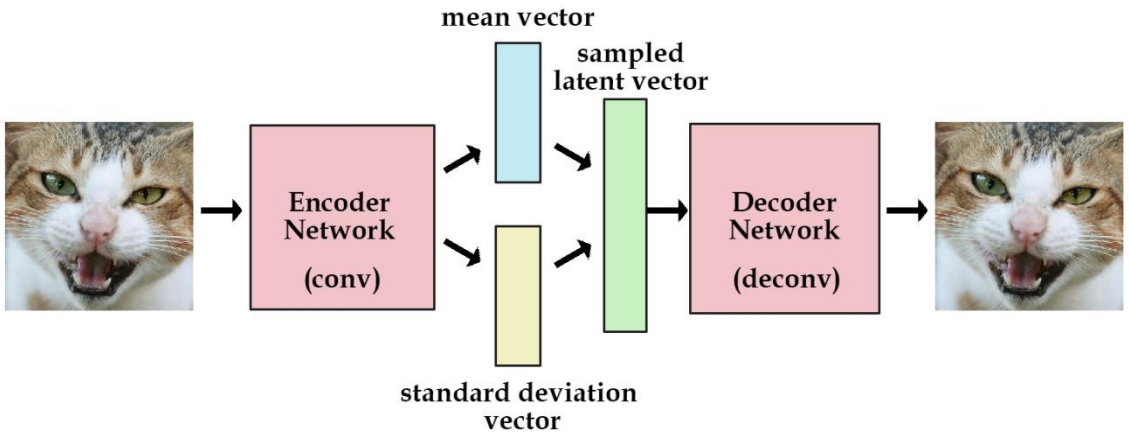
Variations of Autoencoder

- Denoising Autoencoder (DAE)
  - Add **noise** to the input
  - Loss : $\mathcal{L}\left(x, f\left(g(\tilde{x})\right)\right)$
- Sparse Autoencoder (SAE)
  - Sparsity constraint on h



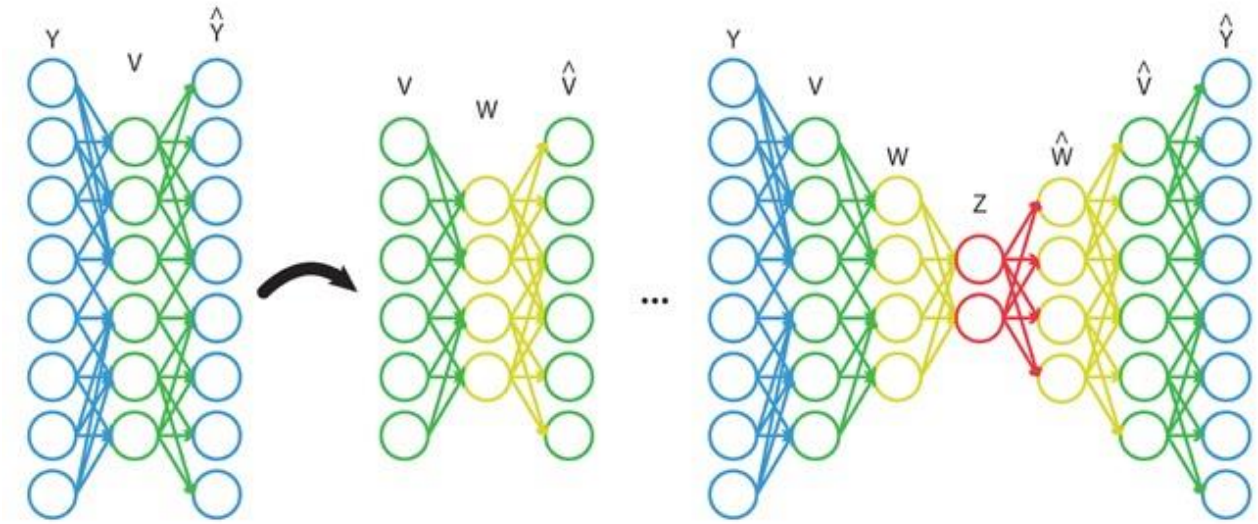
Variations of Autoencoder

- Contractive Autoencoders (CAE)
  - $\frac{\partial f}{\partial x}$ as small as possible
    → contractive penalty
- Variational Autoencoder (VAE)
  - $h \sim N(\mu, \sigma)$
    → variational penalty



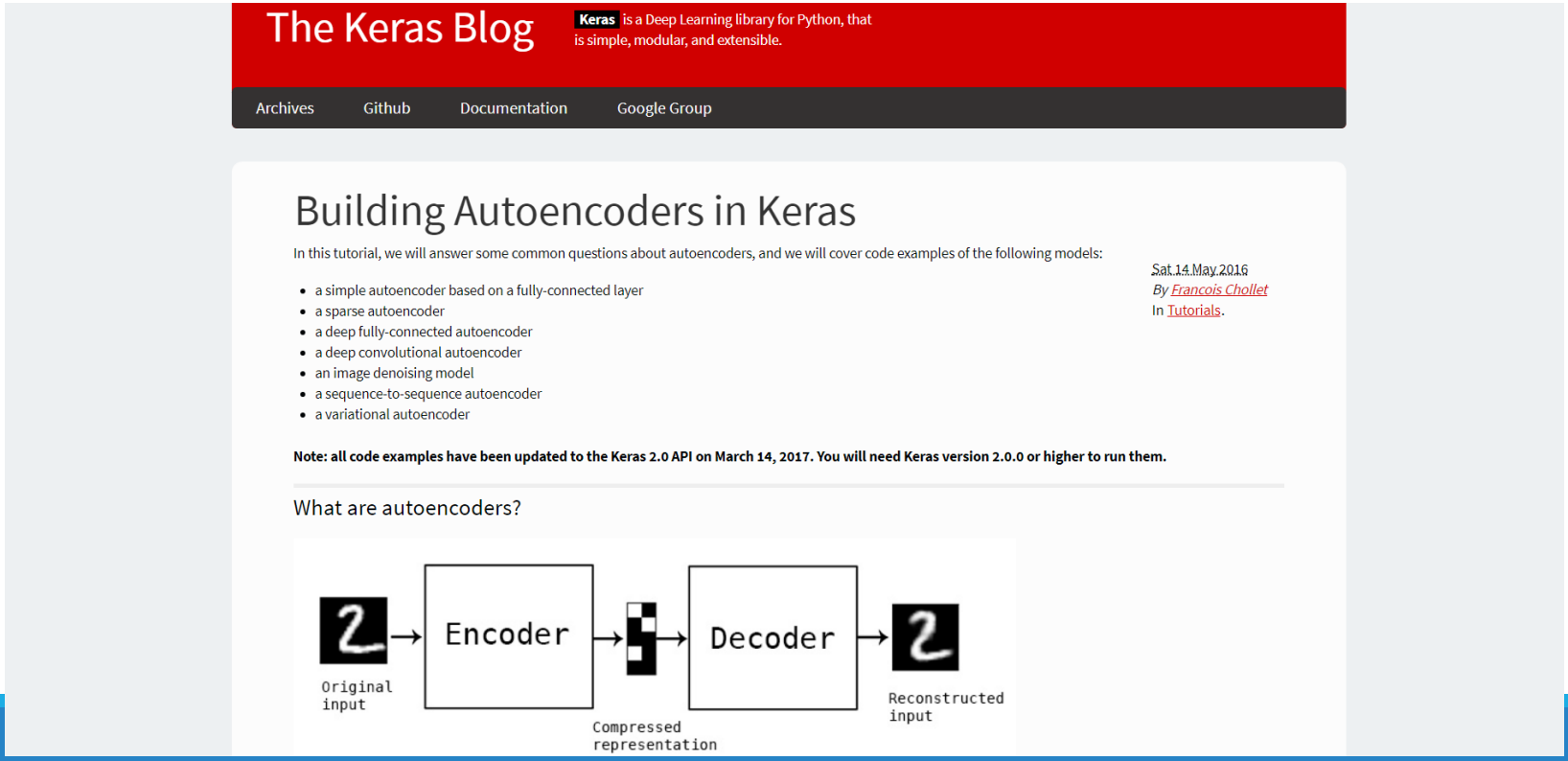
Deep Autoencoders
(or Stacked Autoencoders)

- Much better compression
- Greedy layer-wise pre-training



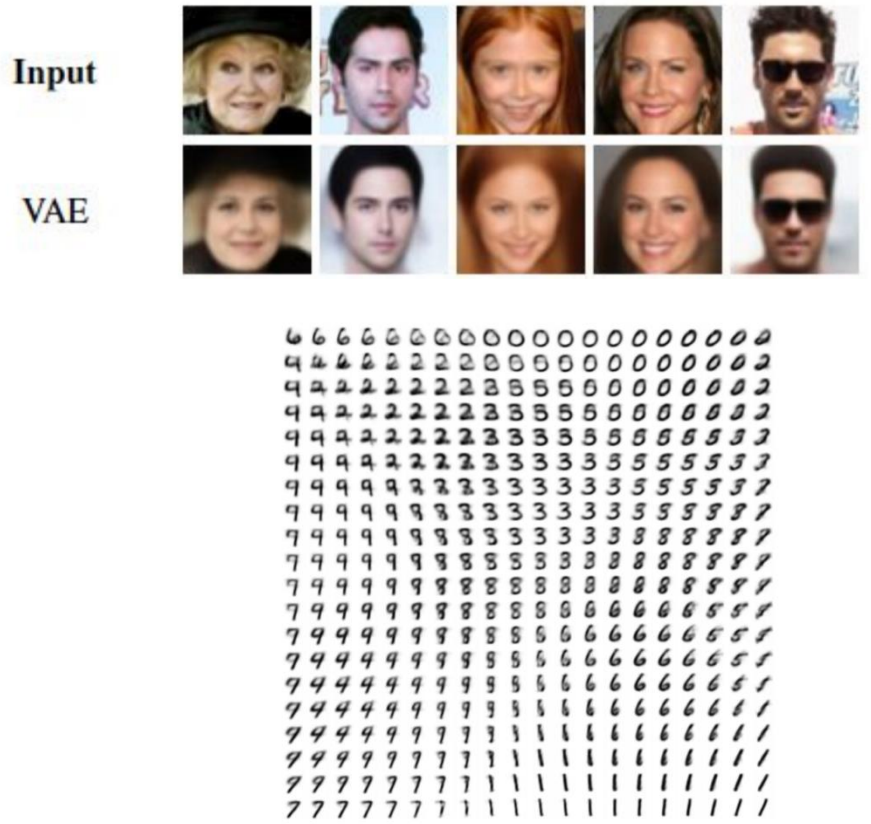
Autoencoder Tutorial
(with Keras by François Chollet)

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



Examples of applications
(non exhaustive)

- Dimensionality reduction
- Information Retrieval
- Image generation
- Voice conversion
- Instrument interpolation
- …



Bibliography

[1] Yoshua Bengio, Pascal Lamblin, Dan Popovici, and Hugo Larochelle. **Greedy layer-wise training of deep networks**. In *P. B. Schölkopf, J. C. Platt, and T. Hoffman, editors, Advances in Neural Information Processing Systems 19*, pages 153–160. MIT Press, 2007.

[2] Kevin Frans. **Variational autoencoders explained.** Blog post, 2016. http://kvfrans.com/variational-autoencoders-explained/.

[3] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[4] Geoffrey Hinton and Ruslan Salakhutdinov. **Reducing the dimensionality of data with neural networks**. *Science*, 313(5786) :504–507, 2006.

[5] Diederik P. Kingma and Max Welling. **Auto-encoding variational bayes**. *CoRR*, abs/1312.6114, 2013.

[6] Anders Boesen Lindbo Larsen, Søren Kaae Sønderby, Hugo Larochelle, and Ole Winther. **Autoencoding beyond pixels using a learned similarity metric**. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning – Volume 48*, ICML'16, pages 1558–1566. JMLR.org, 2016.