

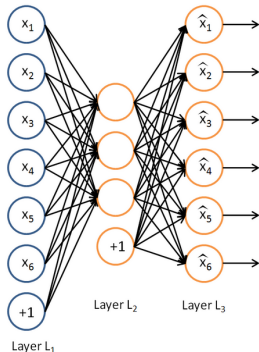
Neural Networks and Deep Learning: Unsupervised Learning

Nicolas Thome

Conservatoire National des Arts et Métiers (Cnam)
Département Informatique

Unsupervised Learning

- ▶ Standard criterion for unsupervised training: reconstruction error, e.g. Mean Squared Error (MSE), Maximum likelihood *etc*
- ▶ Ex: Auto-encoders: $\mathbf{z} = f(\mathbf{W}\mathbf{x})$, $\tilde{\mathbf{x}} = g(\mathbf{W}^t\mathbf{x})$
 - ▶ Auto-encoder objective function: $C = \sum_{i=1}^N \|\mathbf{x}_i - \tilde{\mathbf{x}}\|^2$

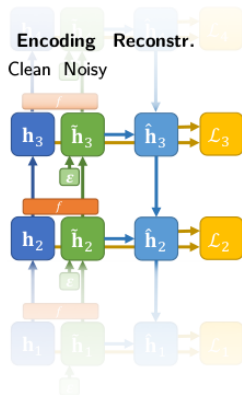


Unsupervised Learning

- ▶ Success of deep learning essentially for supervised tasks, e.g. classification
- ▶ Unsupervised deep learning no comparable breakthrough, **WHY?**
 - ⇒ Classification: clear objective (discrimination) vs
 - ⇒ Reconstruction: questionable
 - ▶ Fitting data well: what if ultimate goal is classification, generalization to a set of examples ?
 - ▶ Reconstruction is not required, or even not a good idea
 - ▶ Deeper representation \Leftrightarrow more abstract representations \Leftrightarrow generalization \Leftrightarrow loss of information
- ▶ Two current alternatives to unsupervised learning:
 1. Objective without reconstruction
 2. Casting unsupervised training as classification

Beyond Reconstruction: Ladder Networks [Rasmus et al., 2015]

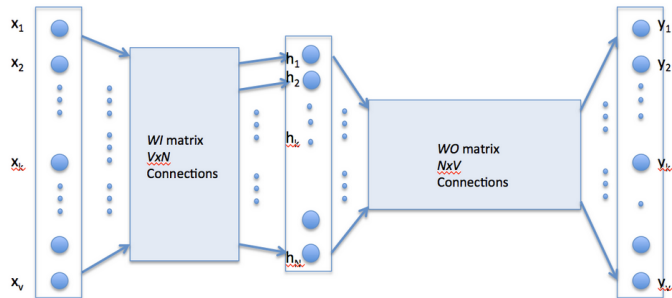
- ▶ "An autoencoder which can discard information"
- ▶ Layer above does not reconstruct layer below only with its activation
- ▶ Solution: Provide the details to learn only the abstract features
 - ▶ Decoder has a noisy version of the input to reconstruct



Auto-Supervision & Predictive Learning

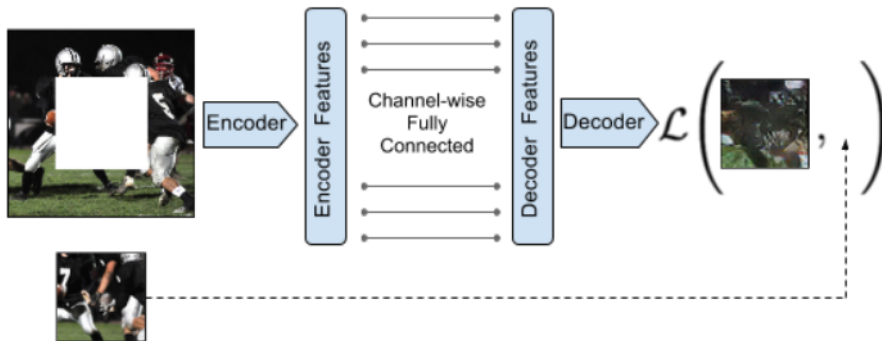
- ▶ Transformed unsupervised problem to a supervised one
- ▶ Automatically creating labels, exploiting "neighborhood", e.g.
 - ▶ Spatial
 - ▶ Temporal

Word2Vec [Mikolov et al., 2013]



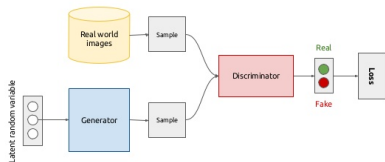
- ▶ Embedding of words, i.e. assign each one-hot word $\in \mathbb{R}^V$ a vector $\in \mathbb{R}^N$
- ▶ Word2Vec principle: predict a word given its context
 - ▶ Assumption: similar words appears in similar contexts
 - ▶ Input: Bag of Words of context
 - ▶ Project to a given space, apply soft max to classify the central word

Context-Encoders [Pathak et al., 2016]: Word2Vec for Images



Generative Adversarial Networks (GAN) [Goodfellow et al., 2014]

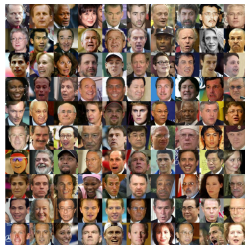
Generative adversarial networks (conceptual)



Noise $\sim N(0,1)$



Generative Model

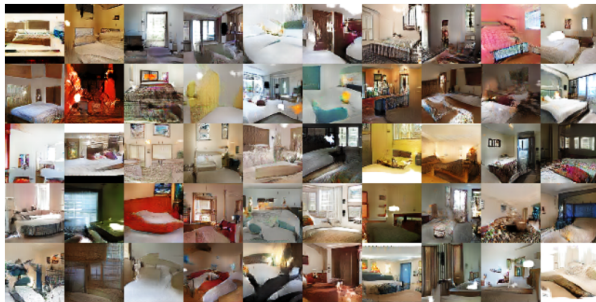
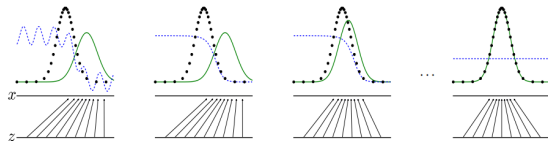


5

- Other trendy auto-supervised method: Generative Adversarial Networks (GAN) [Goodfellow et al., 2014]

Generative Adversarial Networks (GAN) [Goodfellow et al., 2014]

- ▶ Unsupervised problem \Rightarrow 2-player game theory problem
- ▶ Interesting results: optimal generator learns data distribution



Unsupervised Learning: Conclusion

References I



Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014).
Generative adversarial nets.

In Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 27*, pages 2672–2680. Curran Associates, Inc.



Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013).

Distributed representations of words and phrases and their compositionality.

In Burges, C. J. C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 26*, pages 3111–3119. Curran Associates, Inc.



Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T., and Efros, A. (2016).

Context encoders: Feature learning by inpainting.



Rasmus, A., Valpola, H., Honkala, M., Berglund, M., and Raiko, T. (2015).

Semi-supervised learning with ladder networks.

In *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'15, pages 3546–3554, Cambridge, MA, USA. MIT Press.