Learn
Latent
Representations

Unsupervised and self-supervised learning

Auto-Encoders

Jonathon Hare

Vision, Learning and Control
University of Southampton

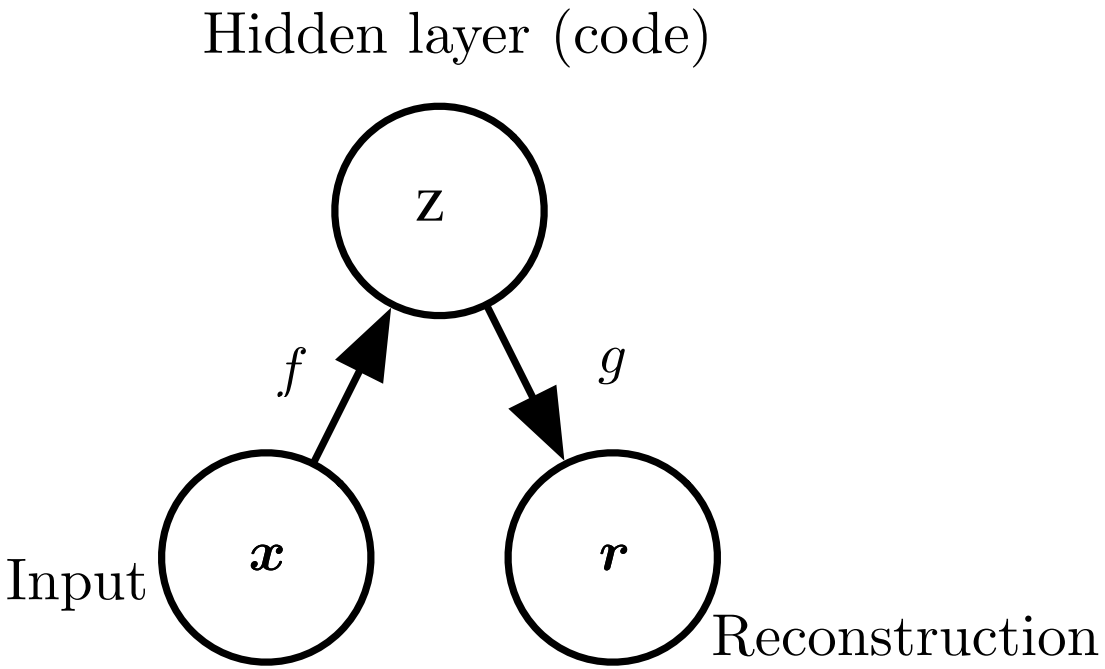
Compressed and Latent Representations

Compression

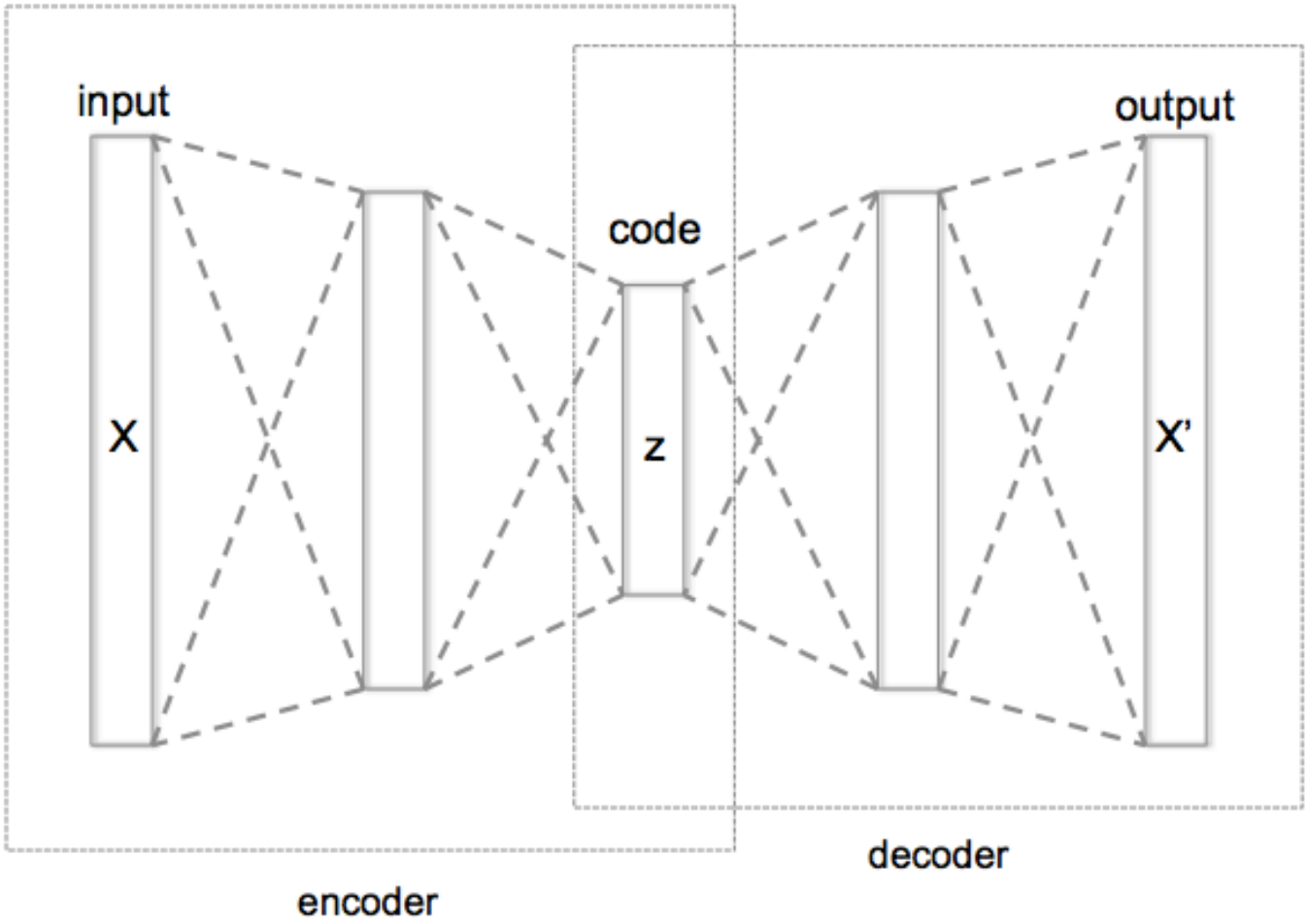
Generative Modelling from Latent Representations

Autoencoders

The aim of an autoencoder is to learn a representation (encoding) for a set of data, typically for dimensionality reduction, by training the network to ignore signal noise.



Deep Autoencoders



1

[1] Image taken from wikipedia

Denoising Autoencoders

- Denoising autoencoders take a partially corrupted input and train to recover the original undistorted input.

Denoising Autoencoders

- Denoising autoencoders take a partially corrupted input and train to recover the original undistorted input.
- To train an autoencoder to denoise data, it is necessary to perform a preliminary stochastic mapping to corrupt the data ($x \rightarrow \tilde{x}$).

Denoising Autoencoders

- Denoising autoencoders take a partially corrupted input and train to recover the original undistorted input.
- To train an autoencoder to denoise data, it is necessary to perform a preliminary stochastic mapping to corrupt the data ($x \rightarrow \tilde{x}$).
- A normal autoencoder is used with $\tilde{x}$ is used as input and x as output.

Denoising Autoencoders

- Denoising autoencoders take a partially corrupted input and train to recover the original undistorted input.
- To train an autoencoder to denoise data, it is necessary to perform a preliminary stochastic mapping to corrupt the data ($x \rightarrow \tilde{x}$).
- A normal autoencoder is used with $\tilde{x}$ is used as input and x as output.
- In a denoising autoencoder, the loss should be computed on $\mathcal{L}(x, \hat{x})$ as opposed to $\mathcal{L}(\tilde{x}, \hat{x})$.

Sparse Autoencoders

- In a sparse autoencoder, there are more hidden units than inputs, but only a small number of the hidden units are allowed to be active at the same time.

Convolutional Autoencoders

Autoencoder Applications

- Any basic AE (or its variant) is used to learn a compact representation of data.

Autoencoder Applications

- Any basic AE (or its variant) is used to learn a compact representation of data.
- You can learn automatic features from data.

Autoencoder Applications

- Any basic AE (or its variant) is used to learn a compact representation of data.
- You can learn automatic features from data.
- Denoising can help generalise over the test set since the data is distorted by adding noise.

Autoencoder Applications

- Any basic AE (or its variant) is used to learn a compact representation of data.
- You can learn automatic features from data.
- Denoising can help generalise over the test set since the data is distorted by adding noise.
- Pretraining networks by learning your network weights using a stacked AE.