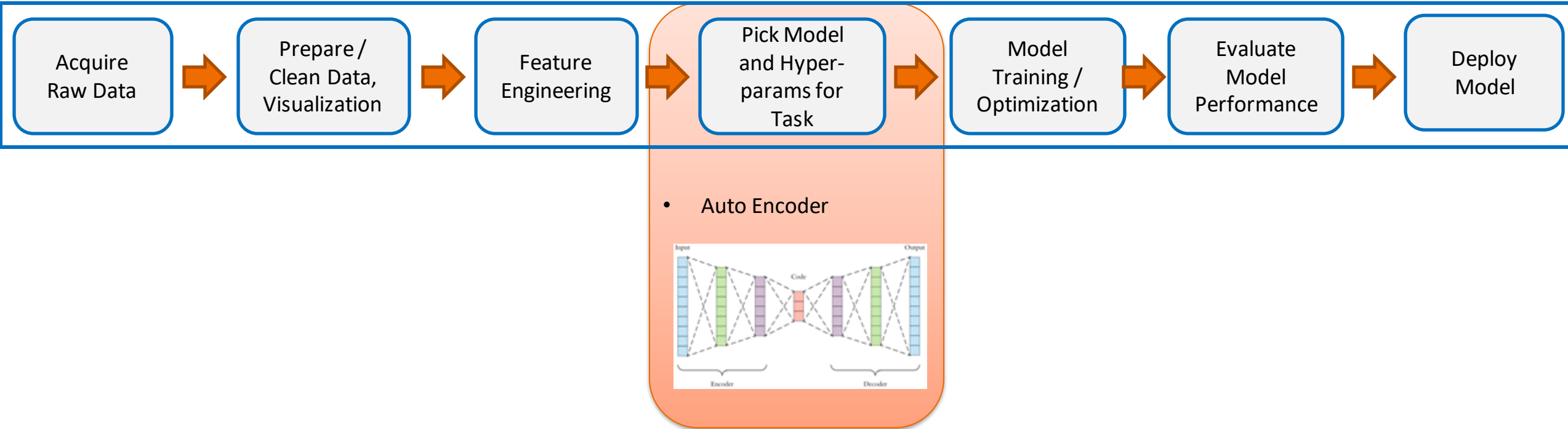


Focus for this lecture

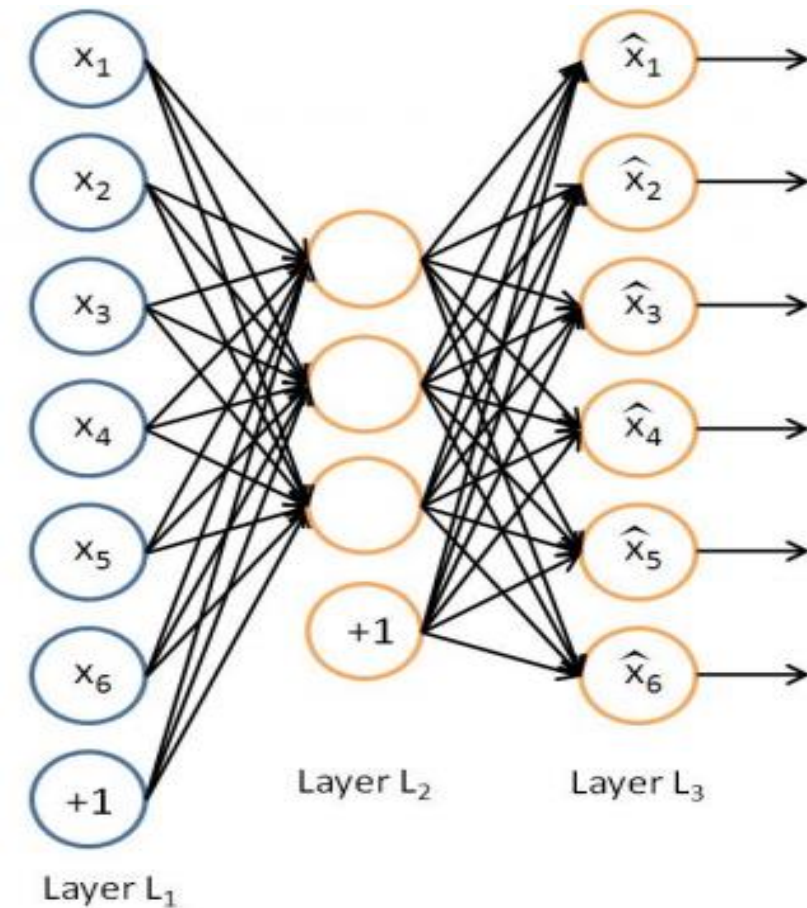


Auto Encoder

— What if we do not have labels? —

Auto-encoder

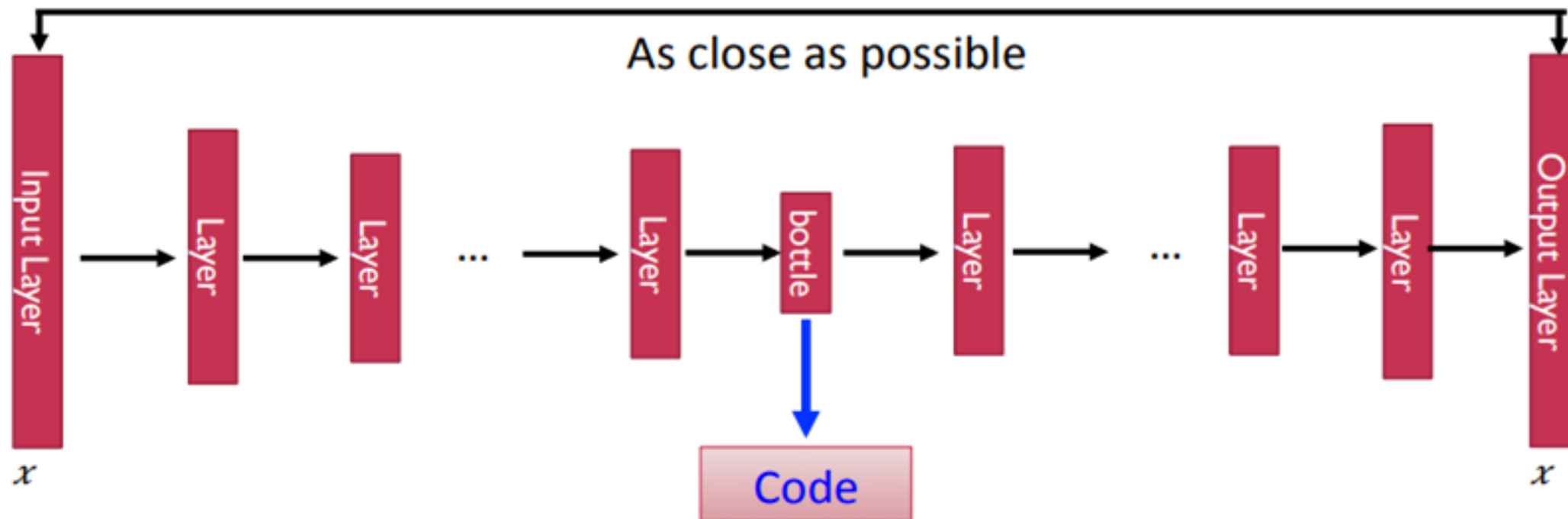
- Similar to MLP
- Input is the same as the output
- Network learns to reconstruct.
- “Bottleneck” layer learns a compact representation.



Deep Auto-encoder

- Of course, the auto-encoder can be deep

Symmetric is not necessary

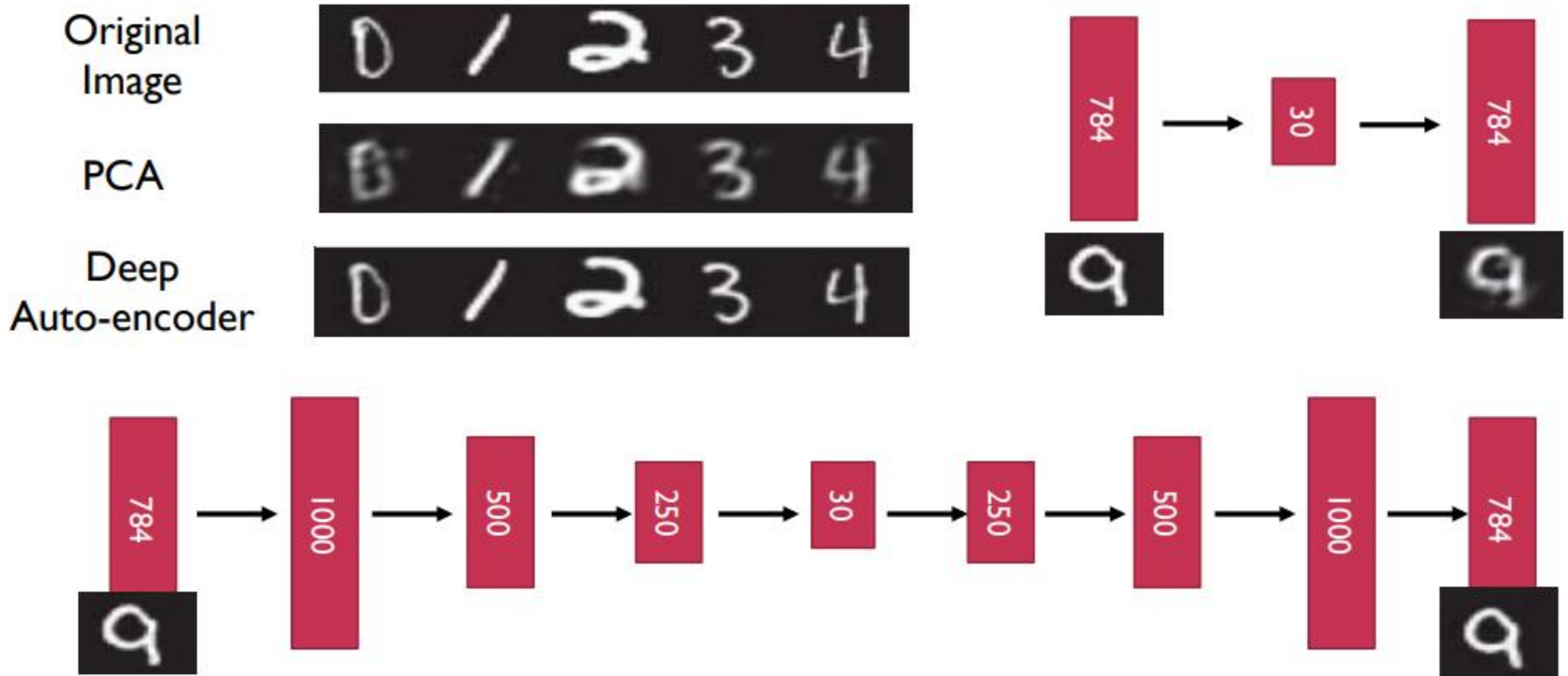


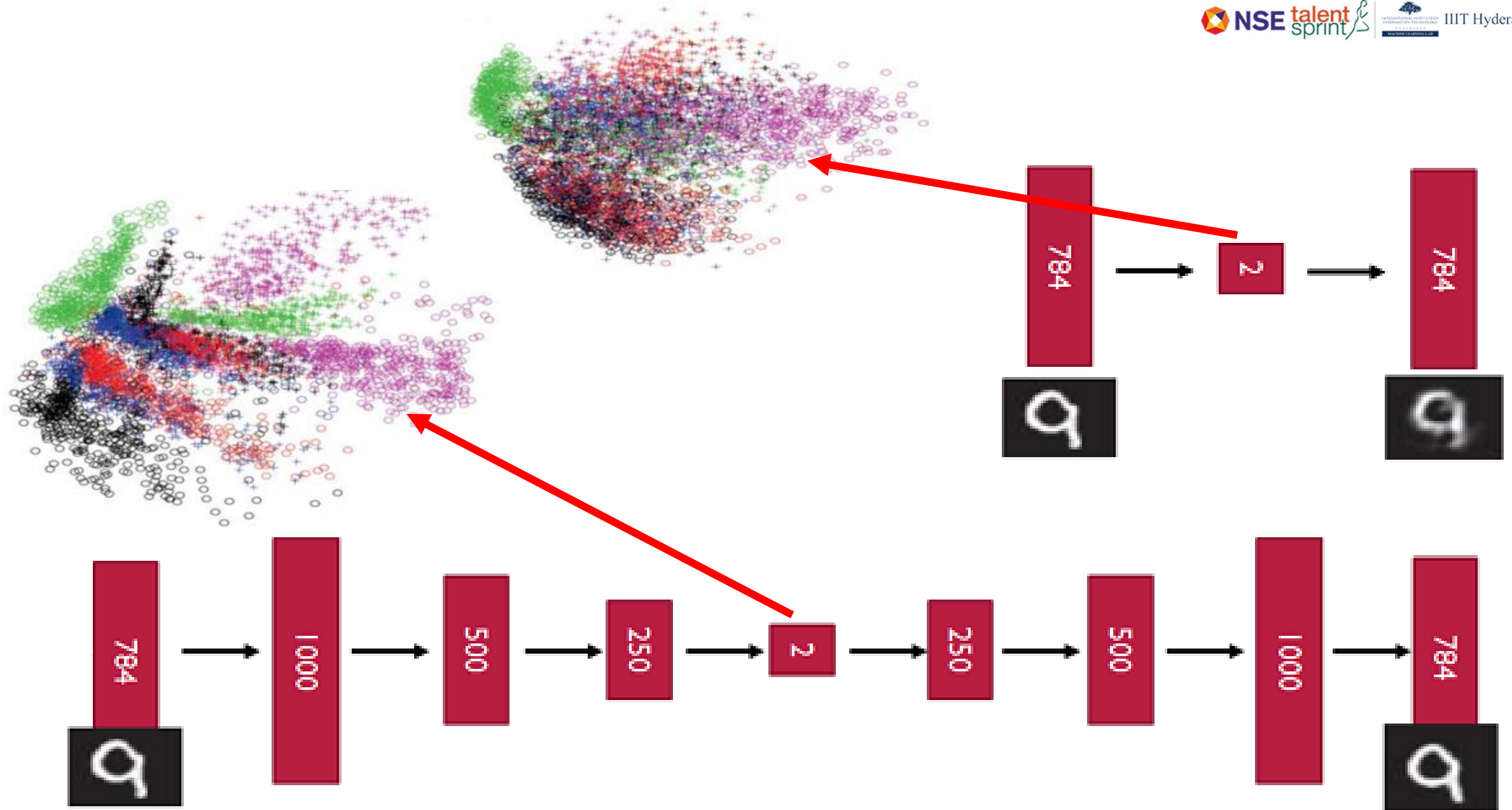
Reference: Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *Science* 313.5786 (2006): 504-507

View Points

- Nonlinear PCA (Dimensionality Reduction)
- Unsupervised Learning
- Data compression

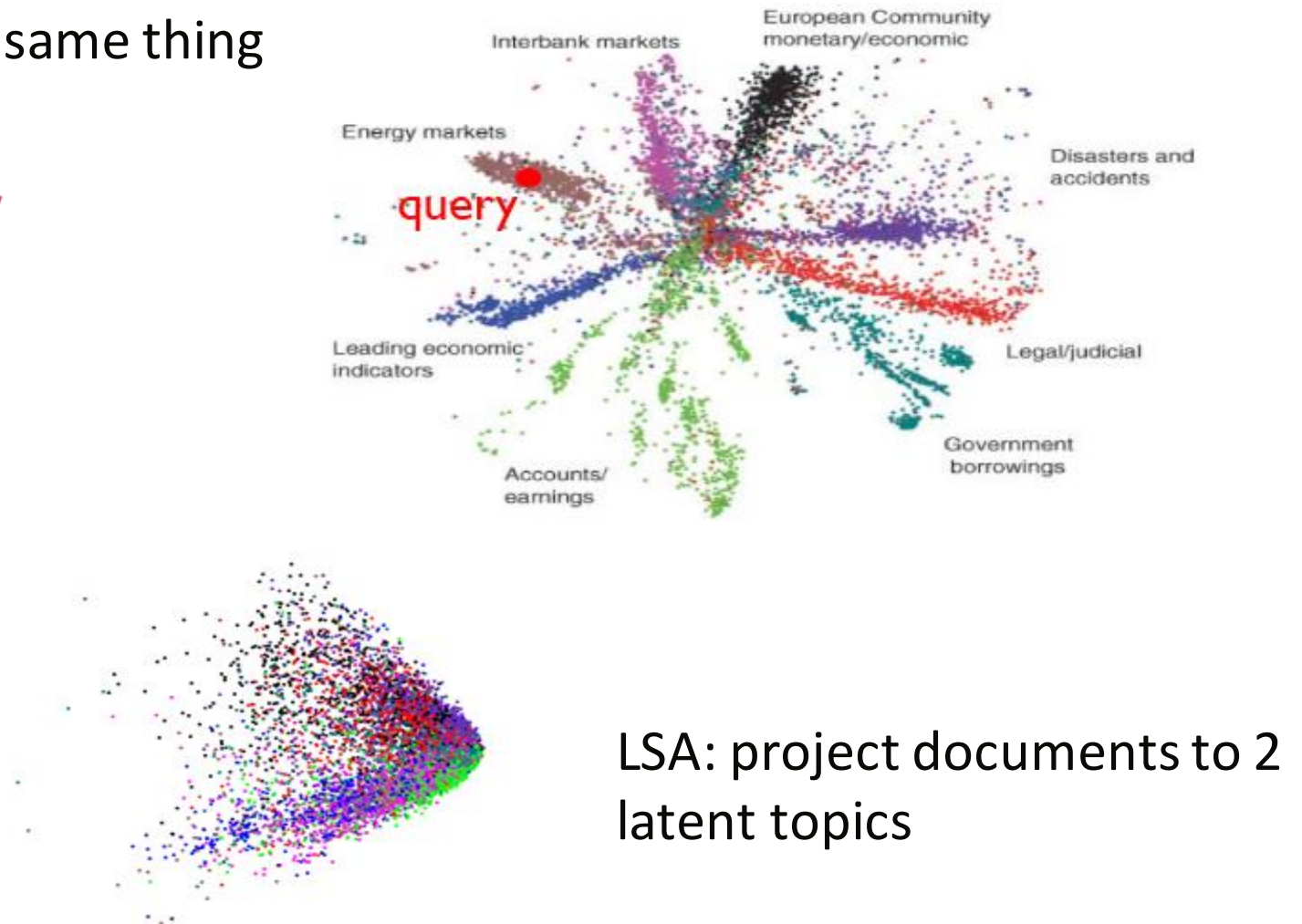
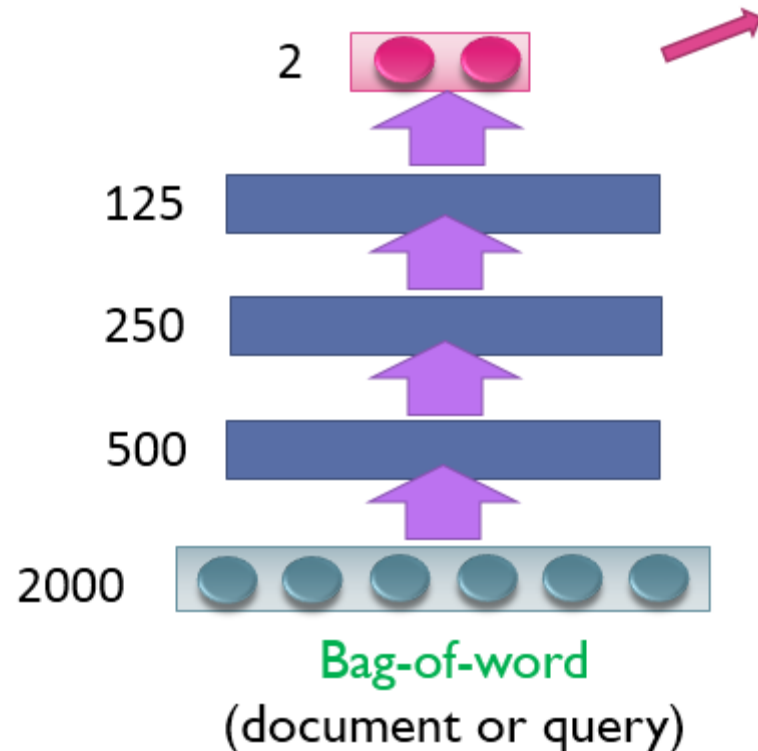
Deep Auto-encoder





Auto-encoder – Text Retrieval

The documents talking about the same thing will have close code.



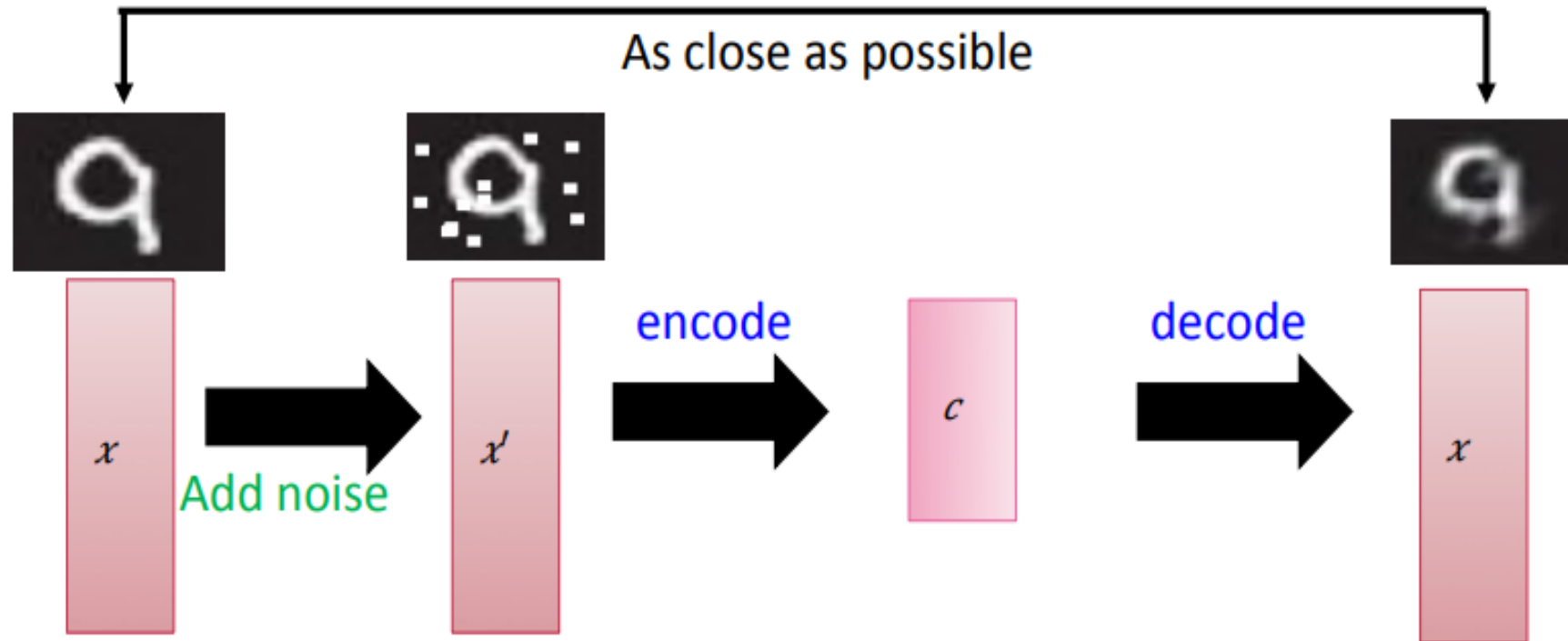
LSA: project documents to 2 latent topics

Variants

- De-noising Auto encoder
- Sparse Auto encoder

Auto-encoder

- De-noising auto-encoder



Vincent, Pascal, et al. "Extracting and composing robust features with denoising autoencoders." *ICML*, 2008.

Unsupervised Learning

Unsupervised Learning

- “We expect unsupervised learning to become far more important in the longer term. Human and animal learning is largely unsupervised: we discover the structure of the world by observing it, not by being told the name of every object.”
 - LeCun, Bengio, Hinton, Nature 2015
- As I've said in previous statements: most of human and animal learning is unsupervised learning. If intelligence was a cake, unsupervised learning would be the cake, supervised learning would be the icing on the cake, and reinforcement learning would be the cherry on the cake. We know how to make the icing and the cherry, but we don't know how to make the cake.
 - Yann LeCun, March 14, 2016 (Facebook)

Old and New

2016

■ "Pure" Reinforcement Learning (cherry)

- ▶ The machine predicts a scalar reward given once in a while.
- ▶ **A few bits for some samples**

■ Supervised Learning (icing)

- ▶ The machine predicts a category or a few numbers for each input
- ▶ Predicting human-supplied data
- ▶ **10→10,000 bits per sample**

■ Unsupervised/Predictive Learning (cake)

- ▶ The machine predicts any part of its input for any observed part.
- ▶ Predicts future frames in videos
- ▶ **Millions of bits per sample**

■ (Yes, I know, this picture is slightly offensive to RL folks. But I'll make it up)



2019

How Much Information is the Machine Given during Learning?

Y. LeCun

▶ "Pure" Reinforcement Learning (cherry)

- ▶ The machine predicts a scalar reward given once in a while.
- ▶ **A few bits for some samples**

▶ Supervised Learning (icing)

- ▶ The machine predicts a category or a few numbers for each input
- ▶ Predicting human-supplied data
- ▶ **10→10,000 bits per sample**

▶ Self-Supervised Learning (cake génoise)

- ▶ The machine predicts any part of its input for any observed part.
- ▶ Predicts future frames in videos
- ▶ **Millions of bits per sample**



© 2019 IEEE International Solid-State Circuits Conference

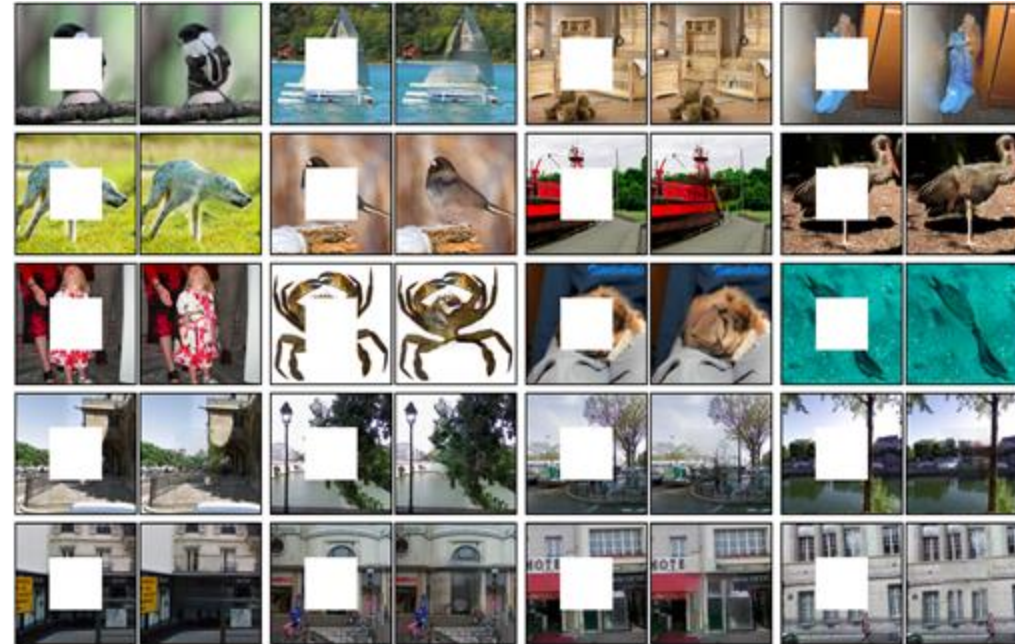
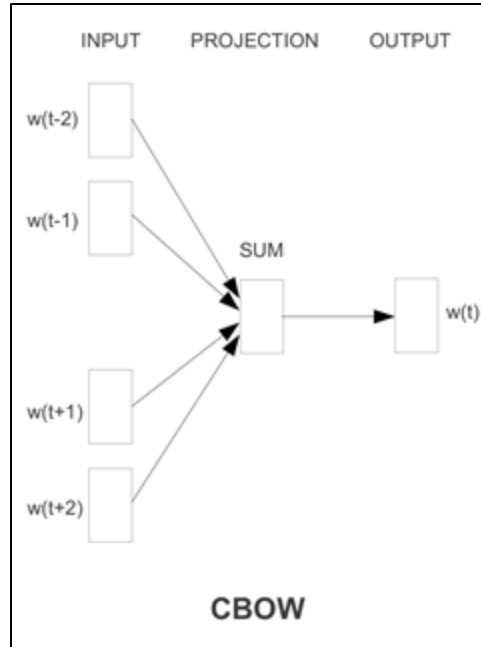
1.1: Deep Learning Hardware: Past, Present, & Future

59

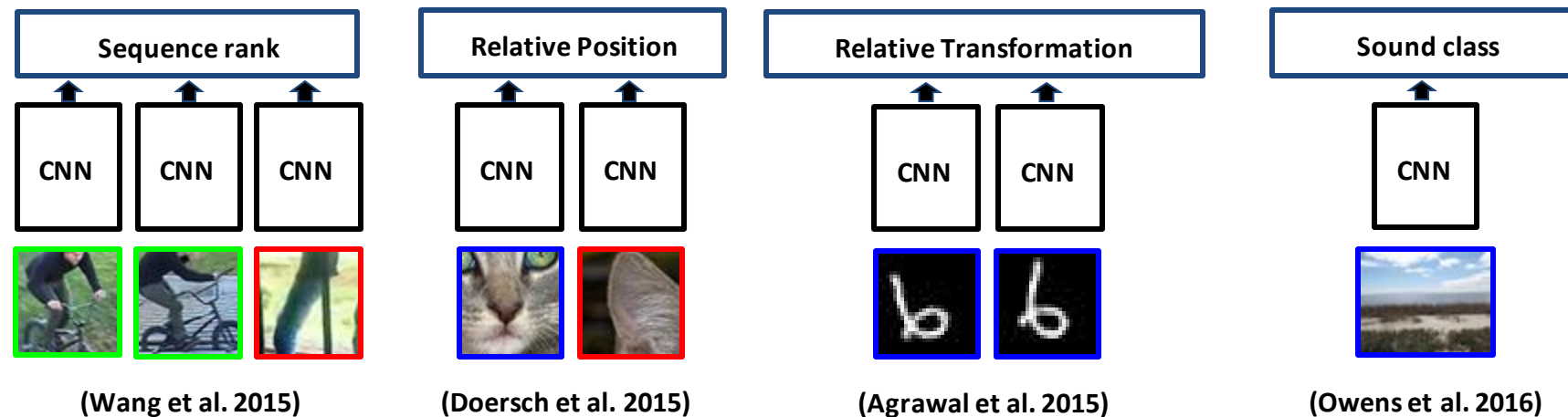
Source: Y. LeCun at NIPS 2016

Self Supervised Learning

Word2Vec
Mikolov 2013



Pathak et al, 2016



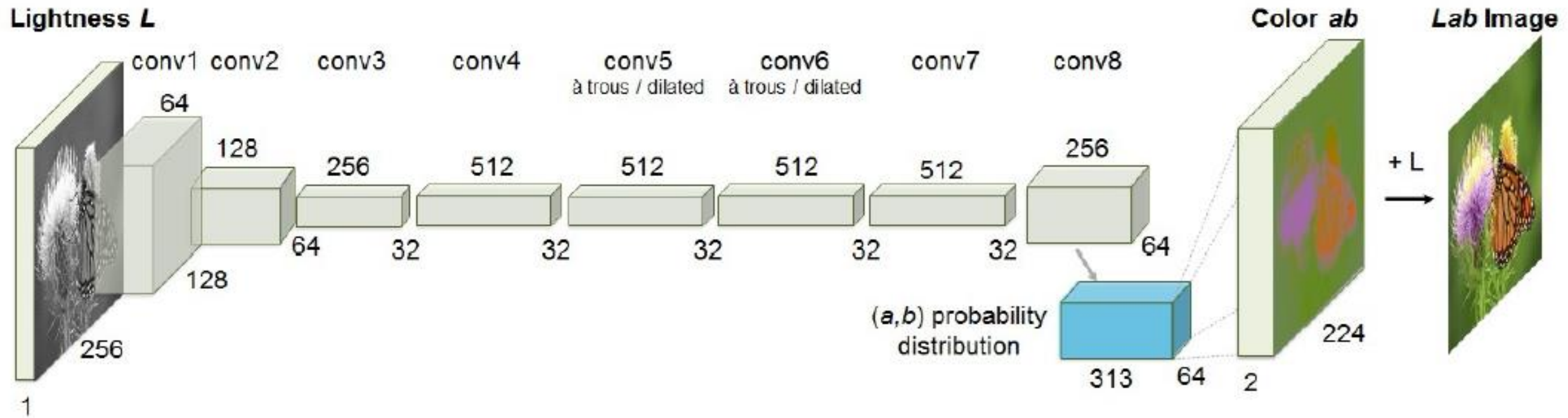
Case Study: Colorization

Encoder – Decoder Networks
(Not just an AutoEncoder)

Colorization



Network Architecture

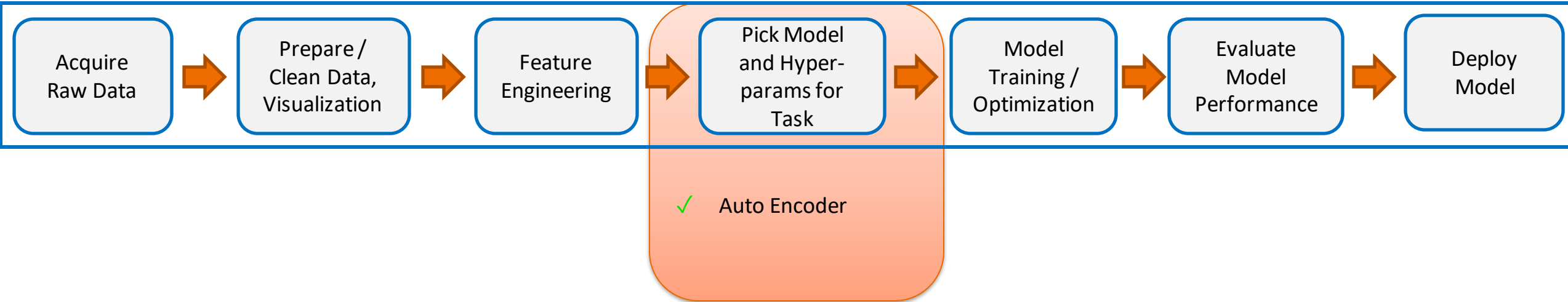


Source: From Zhang et al. [2016]

Colorization: Old Indian B/W Video



Focus for this lecture



Thanks!!

Questions?