Machine Learning

Deep Learning with Auto-Encoding

Instructor: Prof. Yi Fang

yfang@nyu.edu

Acknowledge: The slides are partially referred to the online materials by Prof. Hung-yi Lee from NTU, Auto-encoder session, and also other online materials

Note that some figures and contents used in the presentation are from public available resources, the original authors own the copyright of the figures and content

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)

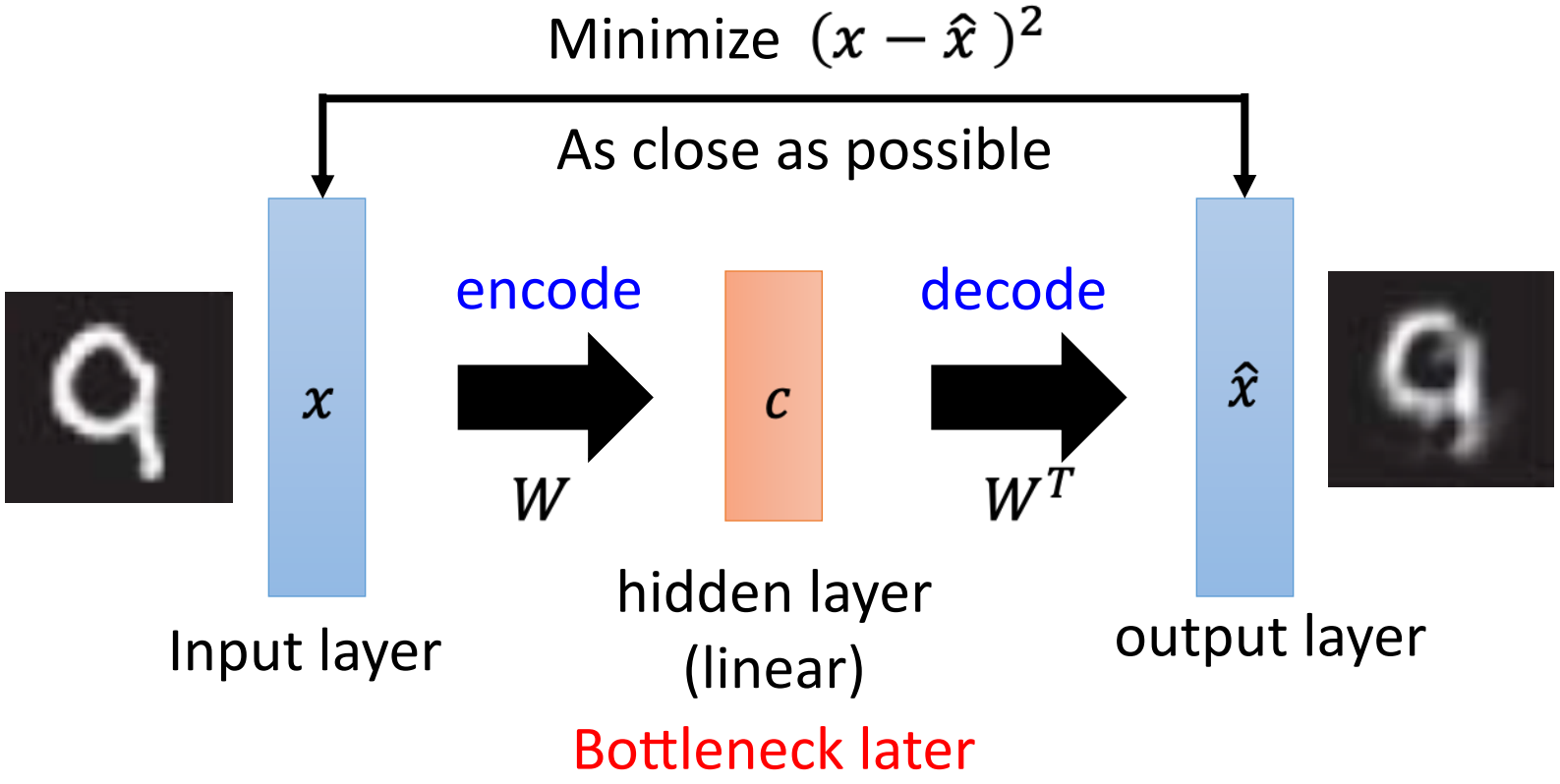
Unsupervised Learning

What we have learnt:

- Manifold Learning: PCA, LLE, MDS, etc.

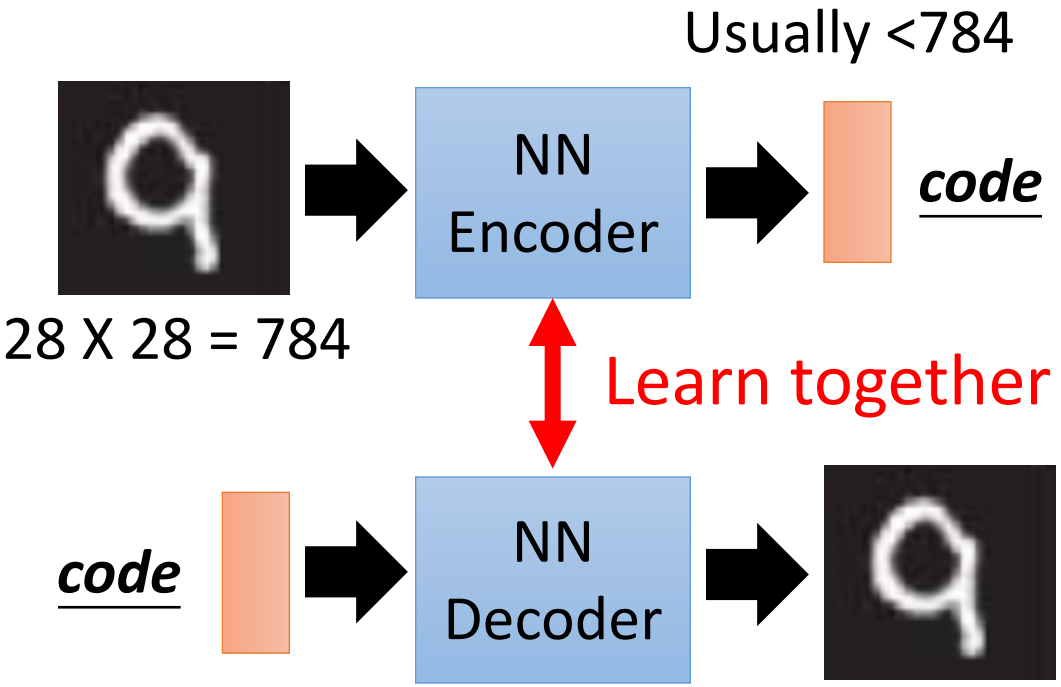
Acknowledge: The slides are partially referred to the online materials by Taegyun Joen, https://www.slideshare.net/TaegyunJeon1/pr12-you-only-look-once-yolo-unified-realtime-object-detection and online YOLO paper and other materials (from ECS289g by Prof. Lee)

Recap: PCA



Output of the hidden layer is the code

Auto-encoder



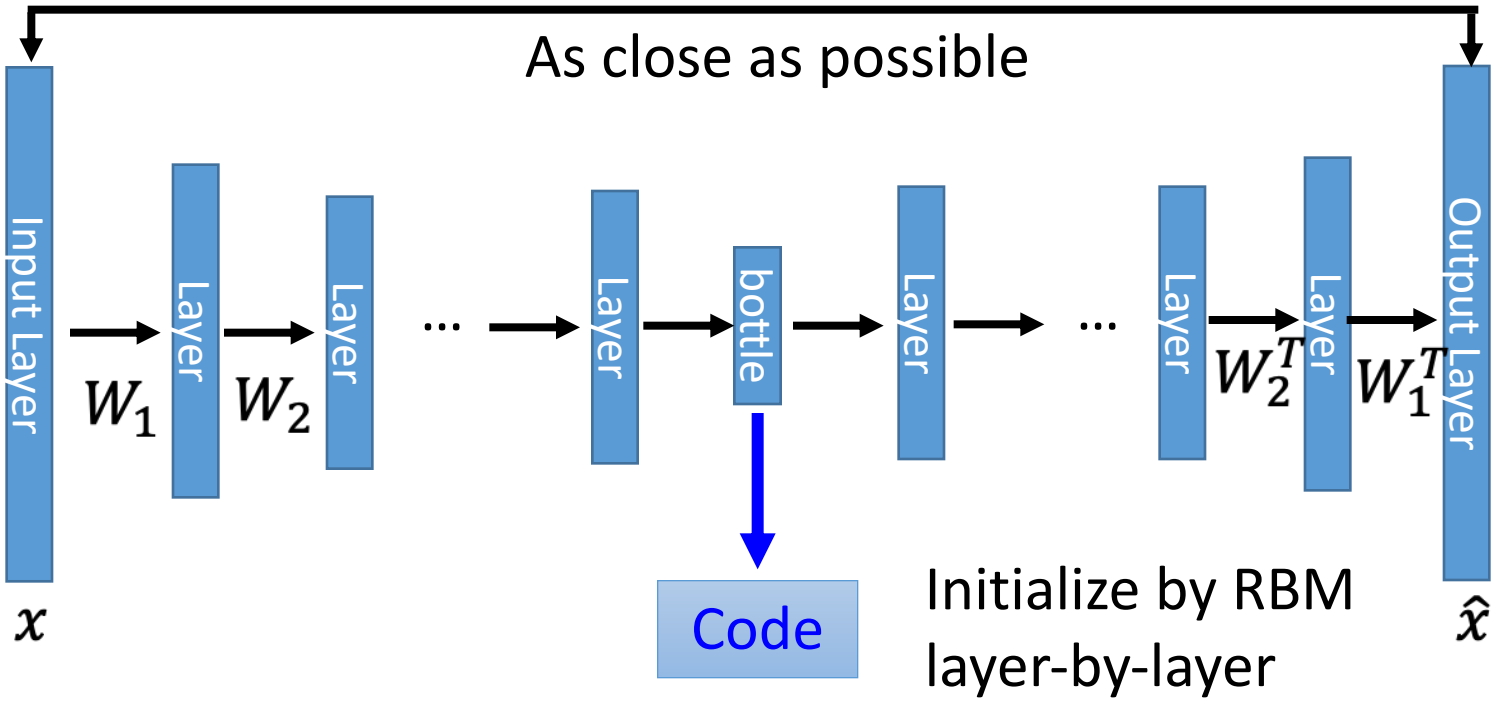
Compact representation of the input object

Can reconstruct the original object

Deep Auto-encoder

Symmetric is not necessary.

- Of course, the auto-encoder can be deep



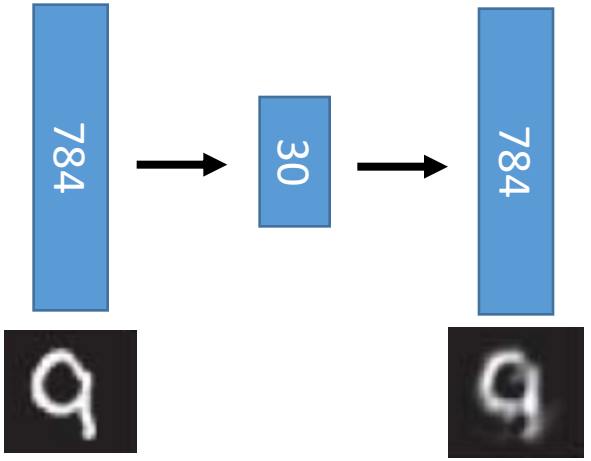
Reference: Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *Science* 313.5786 (2006): 504-507

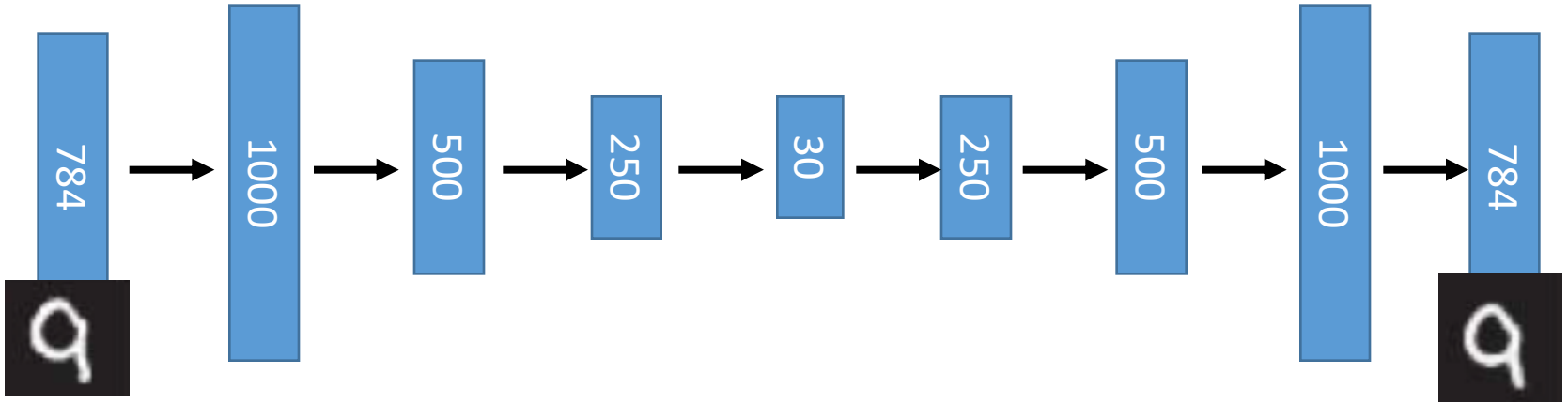
Deep Auto-encoder

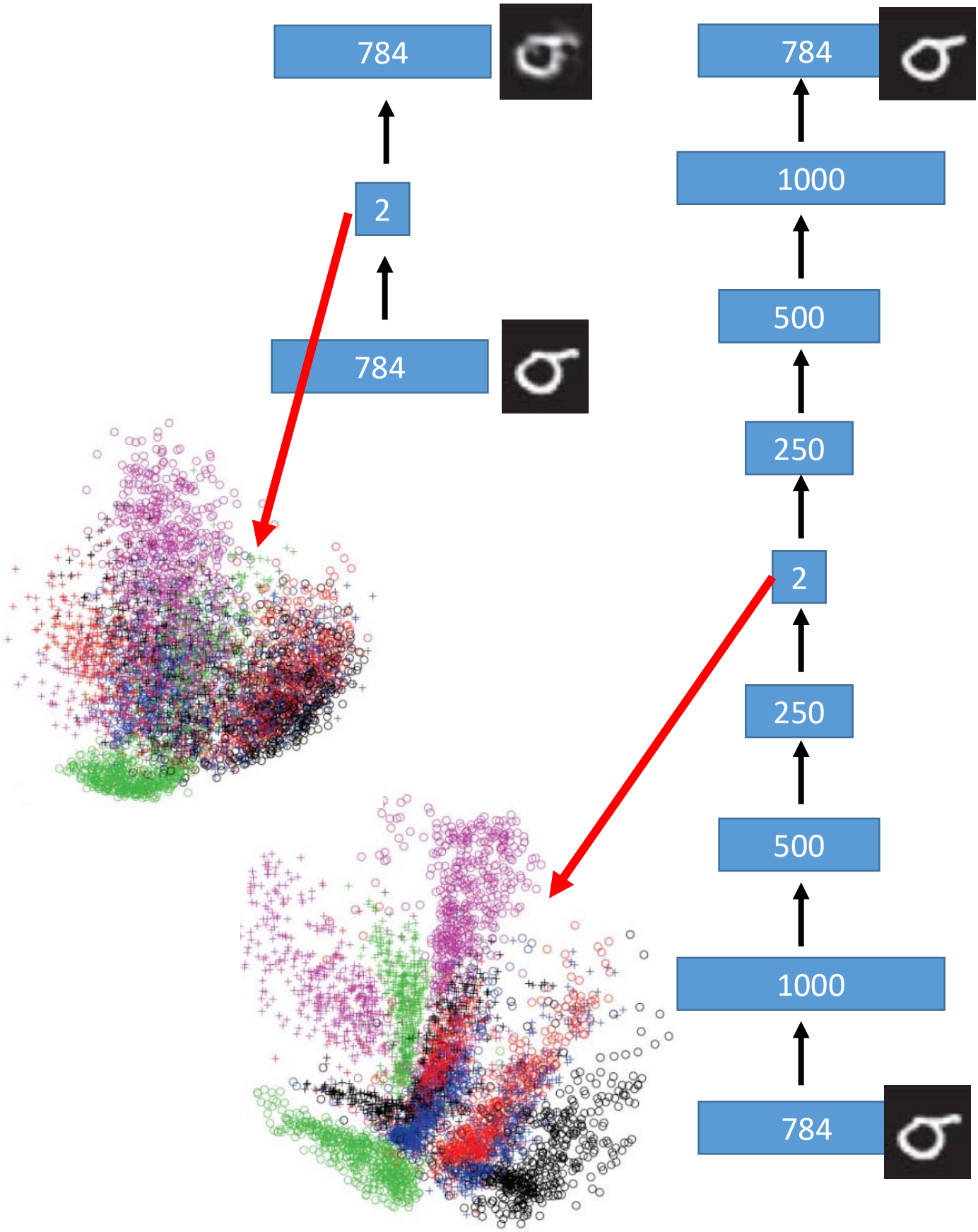
Original Image

PCA

Deep Auto-encoder





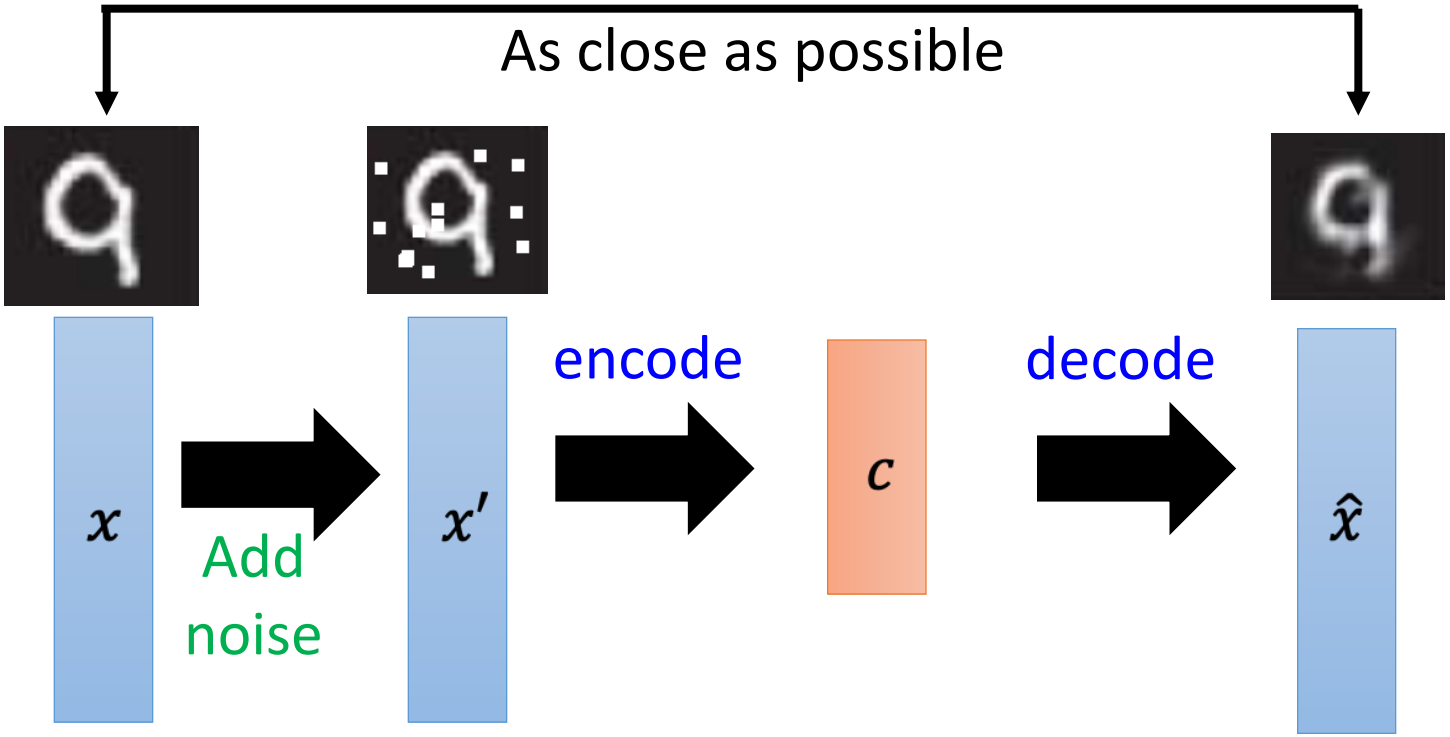

784
2
784
784
1000
500
250
2
250
500
1000
784

More: Contractive auto-encoder

Auto-encoder

Ref: Rifai, Salah, et al. "Contractive auto-encoders: Explicit invariance during feature extraction." *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*. 2011.

- De-noising auto-encoder

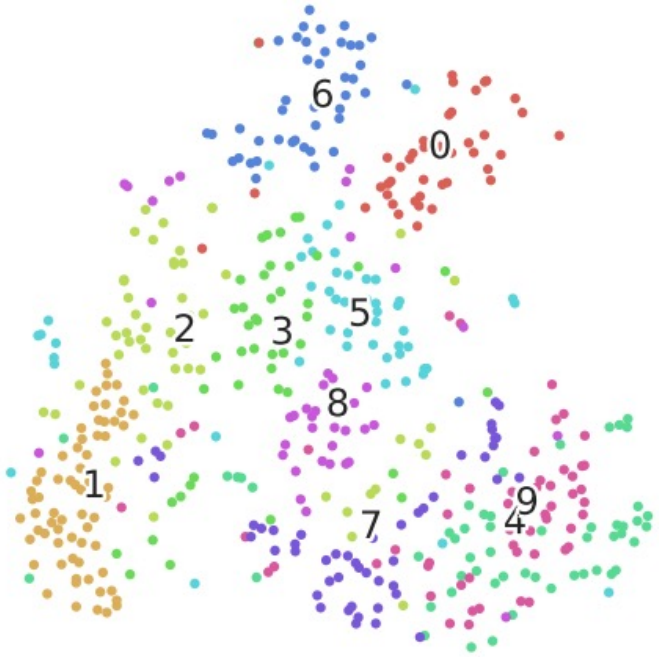


Vincent, Pascal, et al. "Extracting and composing robust features with denoising autoencoders." *ICML,* 2008.

Deep Auto-encoder - Example

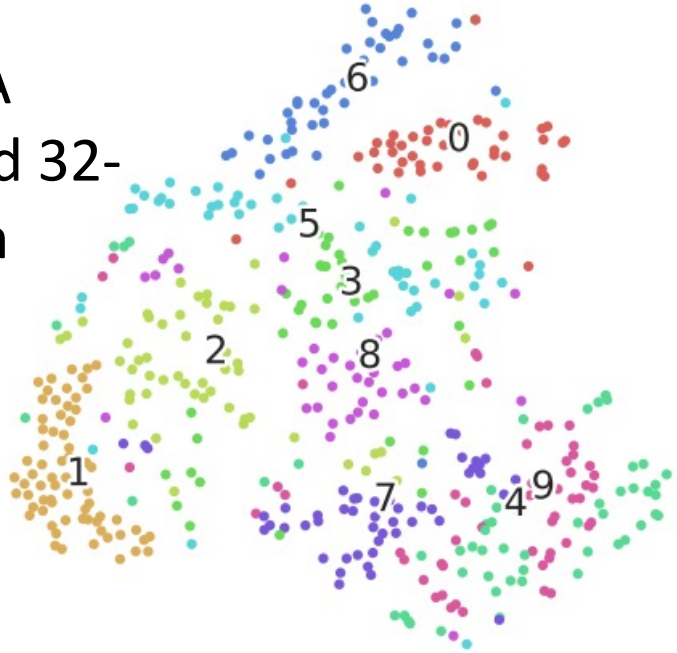
NN Encoder

c



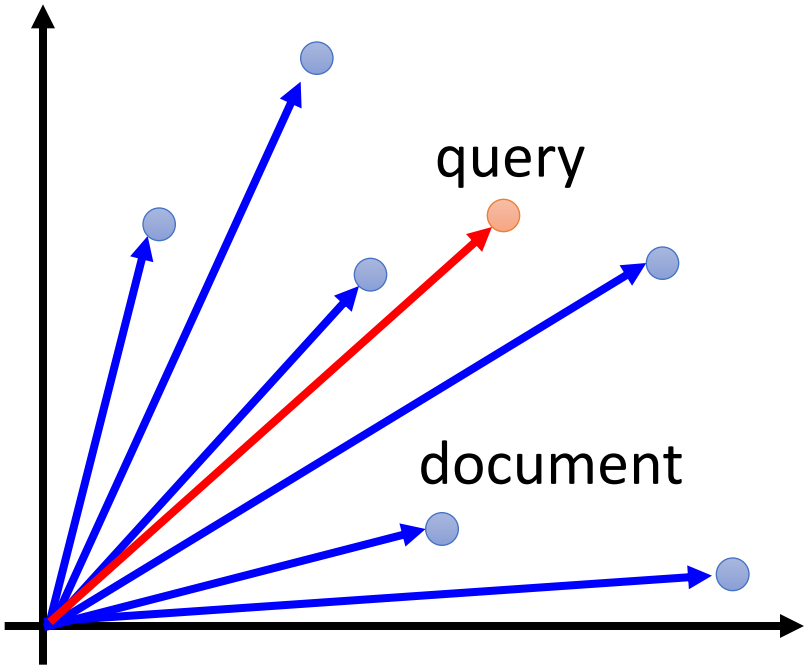
Pixel -> tSNE

PCA Reduced 32-dim



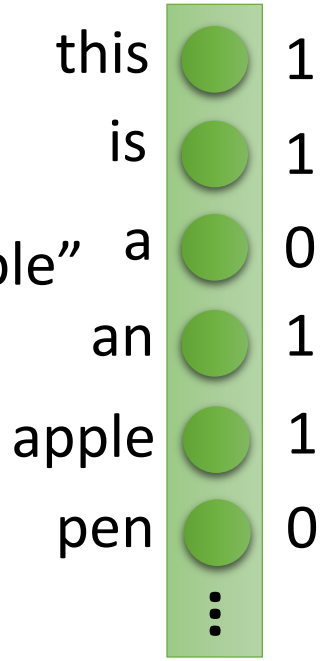
Auto-encoder – Text Retrieval

Vector Space Model



Bag-of-word

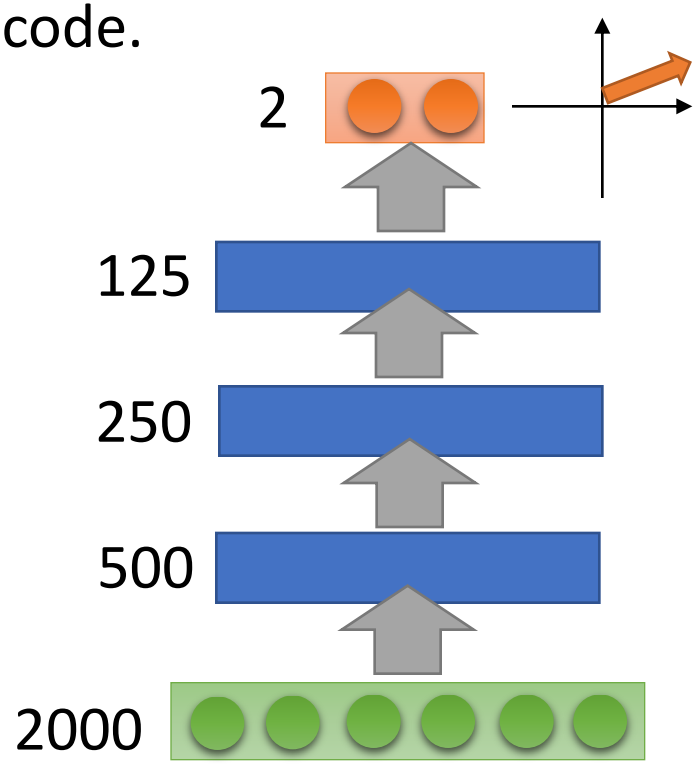
word string:
"This is an apple"



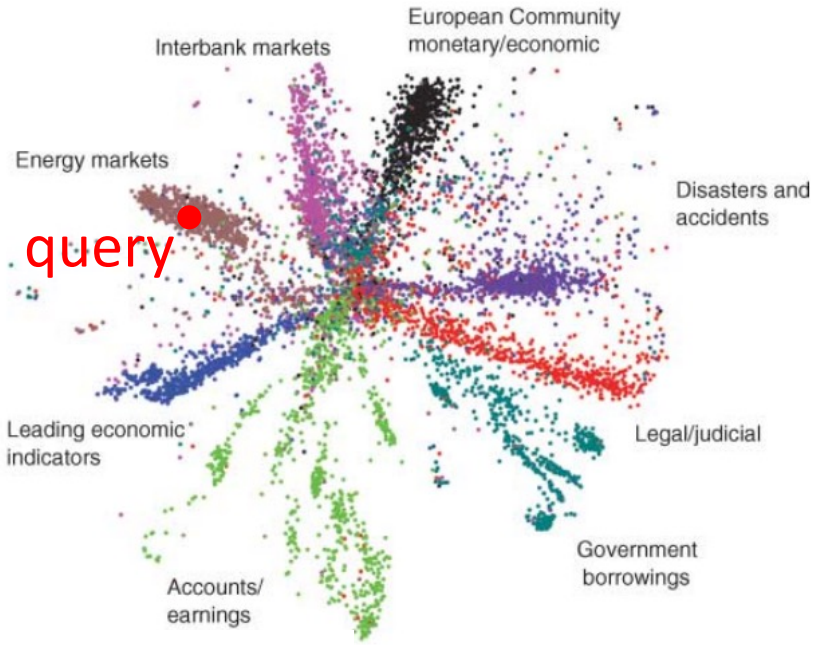
Semantics are not considered.

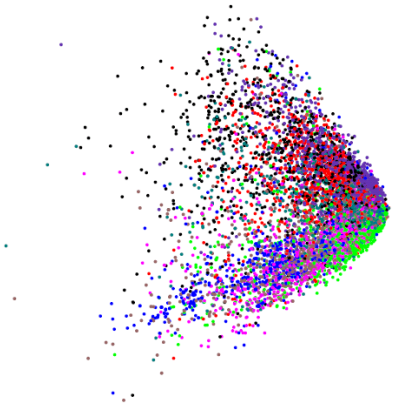
Auto-encoder – Text Retrieval

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



Bag-of-word
(document or query)





LSA: project documents to 2 latent topics

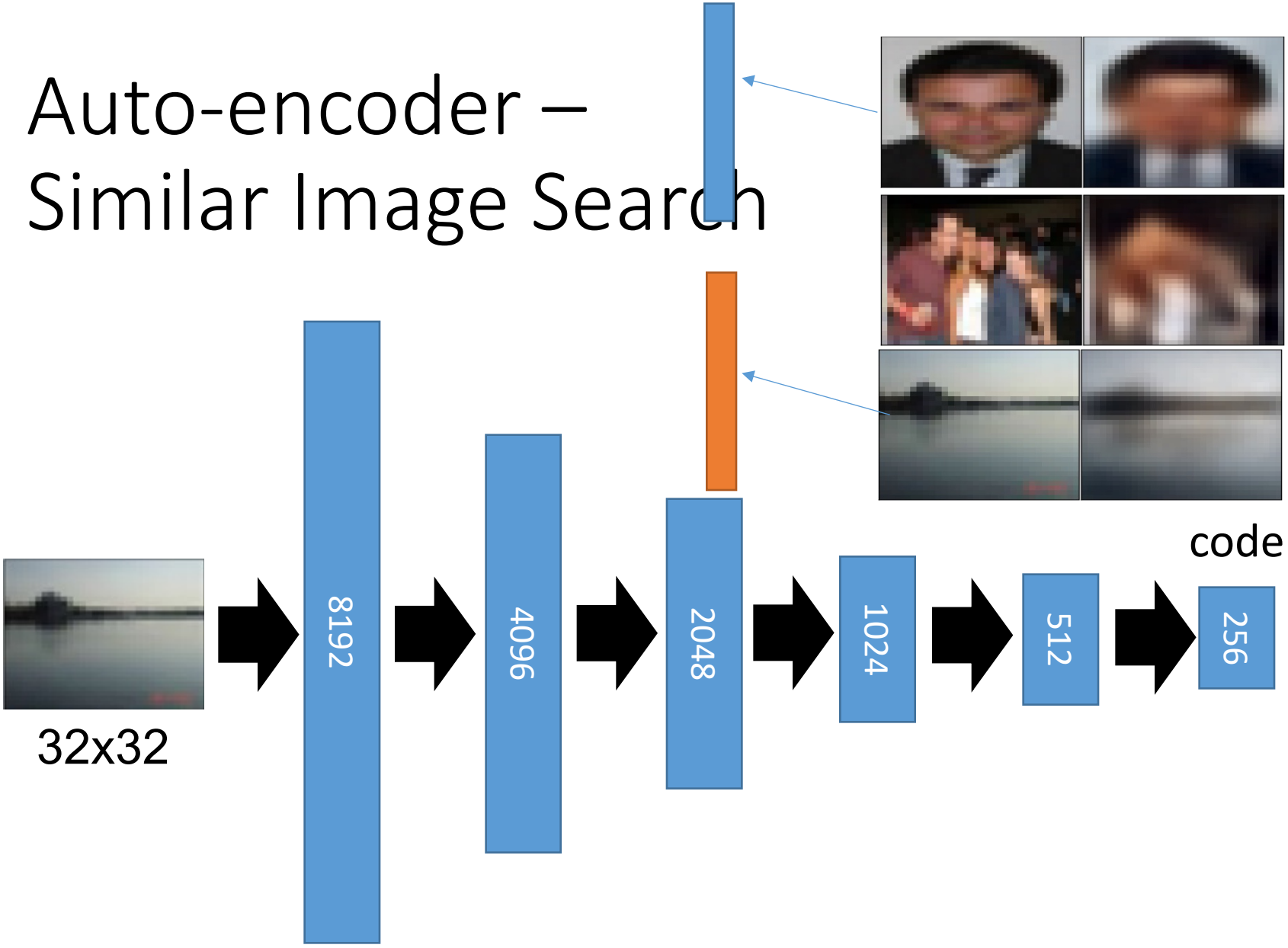
Auto-encoder – Similar Image Search

Retrieved using Euclidean distance in pixel intensity space

(Images from Hinton's slides on Coursera)

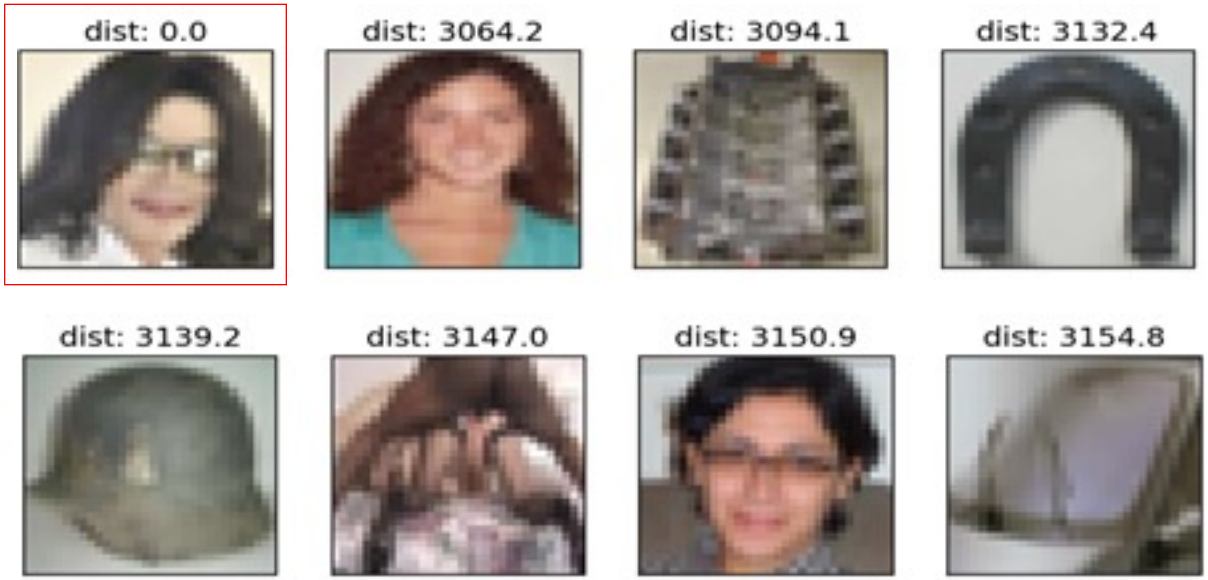
Reference: Krizhevsky, Alex, and Geoffrey E. Hinton. "Using very deep autoencoders for content-based image retrieval." *ESANN*. 2011.

Auto-encoder – Similar Image Search



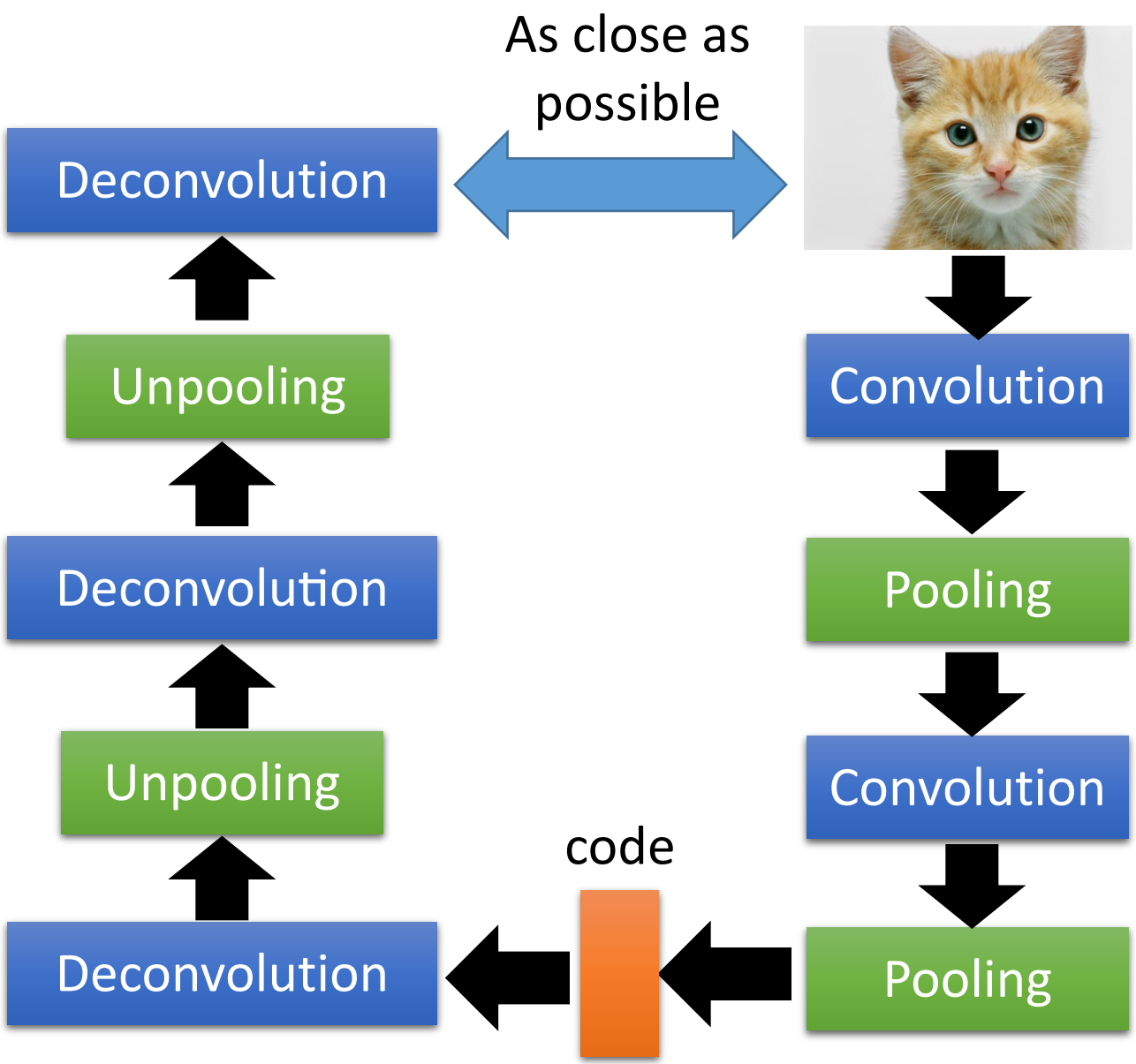
(crawl millions of images from the Internet)

Retrieved using Euclidean distance in pixel intensity space

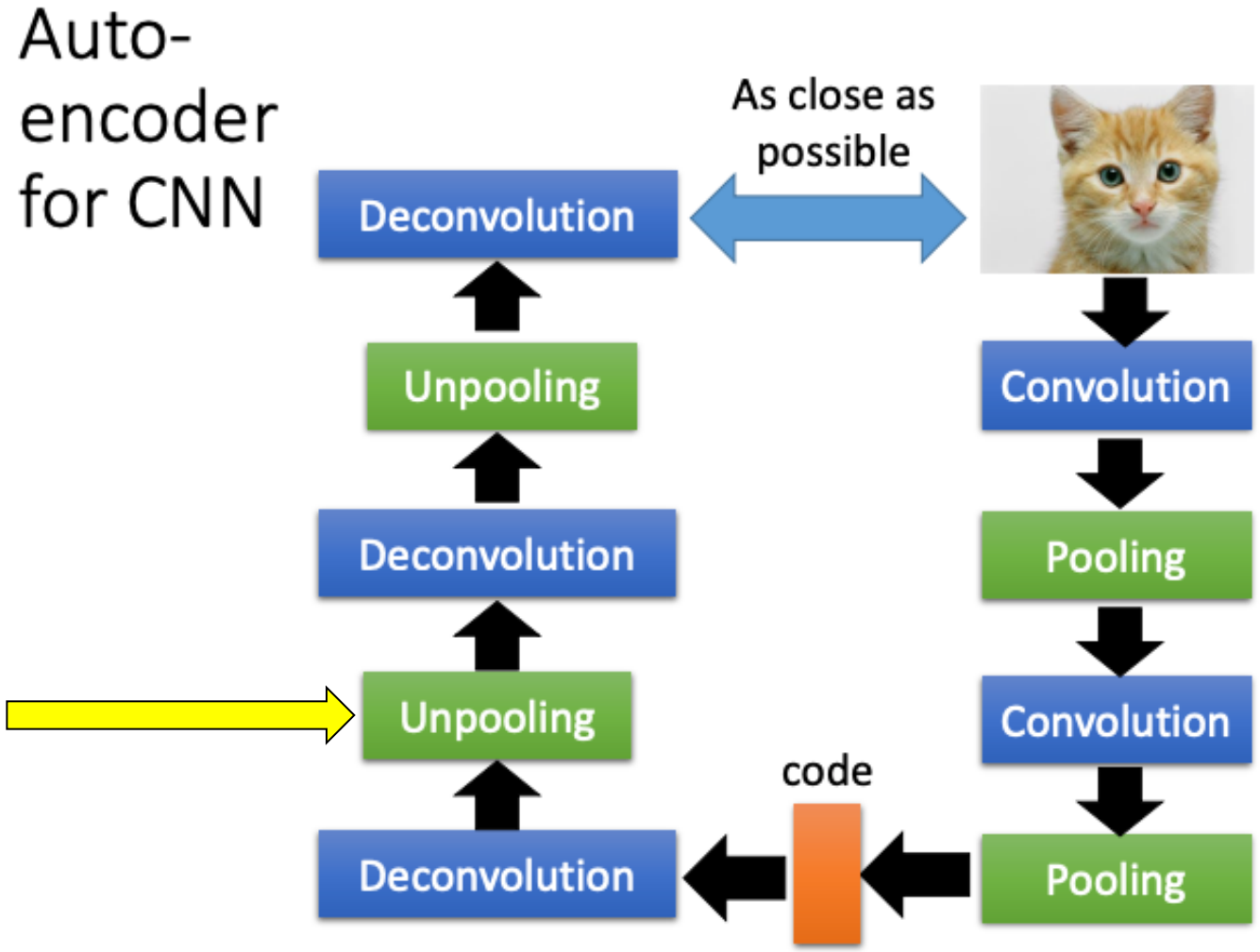


retrieved using 256 codes

Auto-encoder for CNN



CNN For Auto-Encoder



Unpooling

Nearest Neighbor

1	2
3	4

→

1	1	2	2
1	1	2	2
3	3	4	4
3	3	4	4

Input: 2 x 2 Output: 4 x 4

"Bed of Nails"

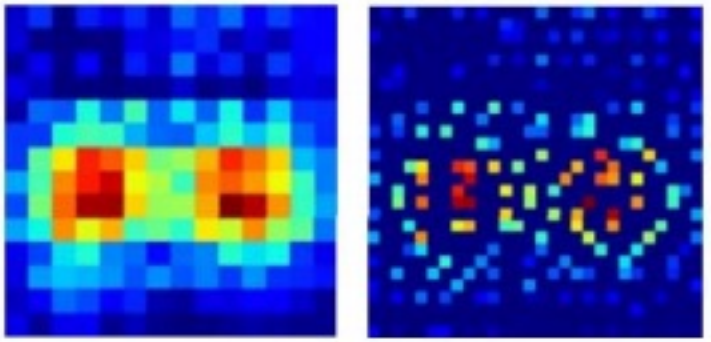
1	2
3	4

→

1	0	2	0
0	0	0	0
3	0	4	0
0	0	0	0

Input: 2 x 2 Output: 4 x 4

Unpooling Result



14 x 14 28 x 28

Source: https://jinzequn.github.io/2018/01/28/deconv-and-unpool/

Max Unpooling

Max Pooling
Remember which element was max!

1	2	6	3
3	5	2	1
1	2	2	1
7	3	4	8

Input: 4 x 4

5	6
7	8

Output: 2 x 2

Rest of the network

Max Unpooling
Use positions from pooling layer

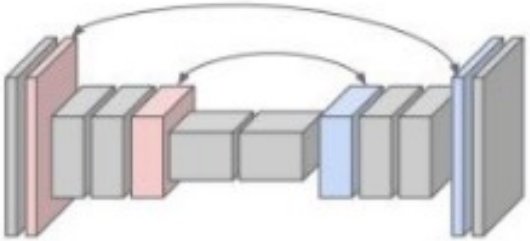
1	2
3	4

Input: 2 x 2

0	0	2	0
0	1	0	0
0	0	0	0
3	0	0	4

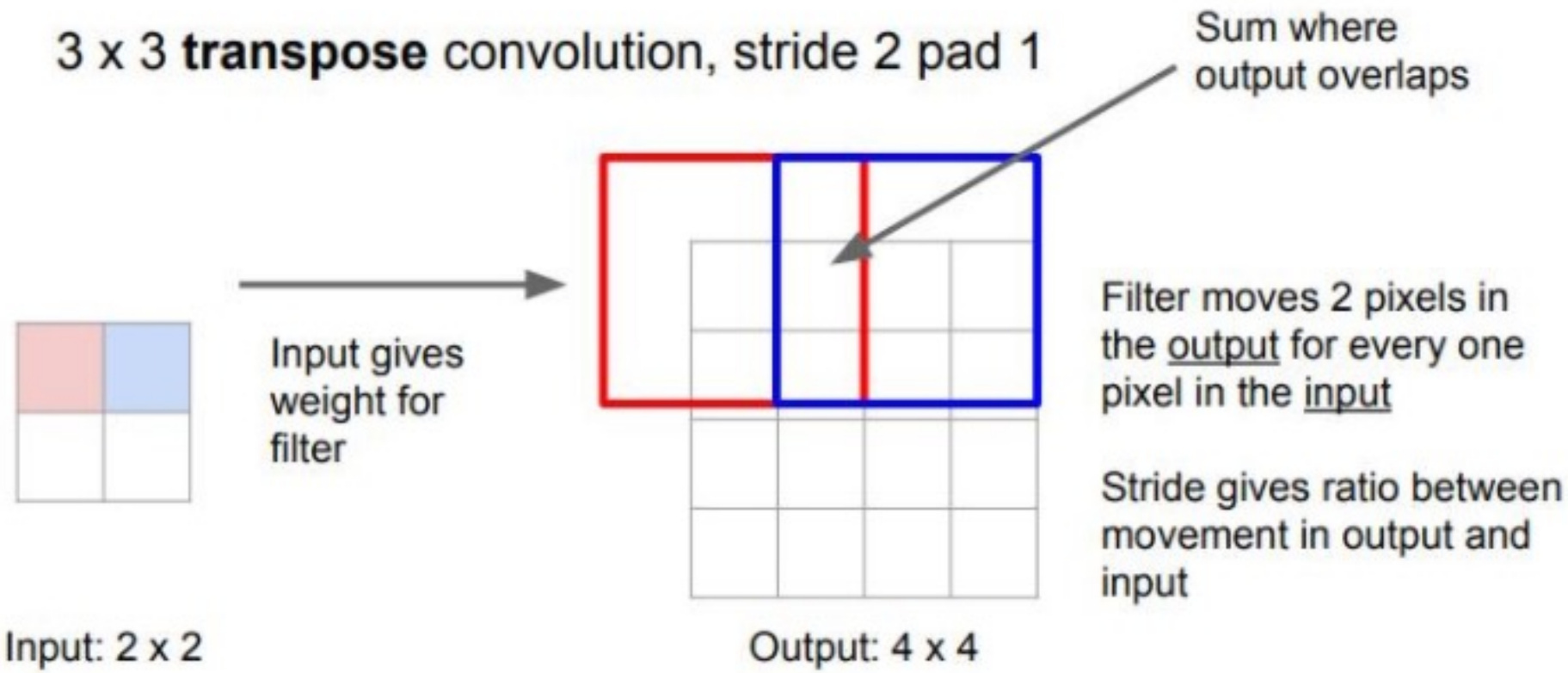
Output: 4 x 4

Corresponding pairs of downsampling and upsampling layers

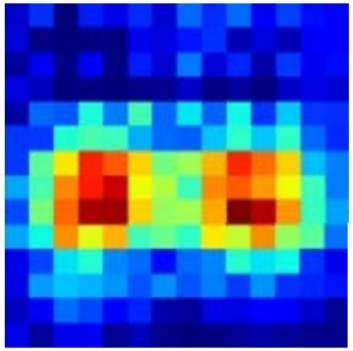


Source: https://jinzequn.github.io/2018/01/28/deconv-and-unpool/

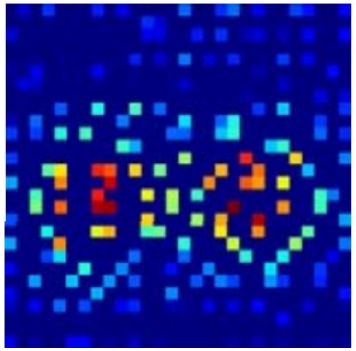
Transpose Convolution



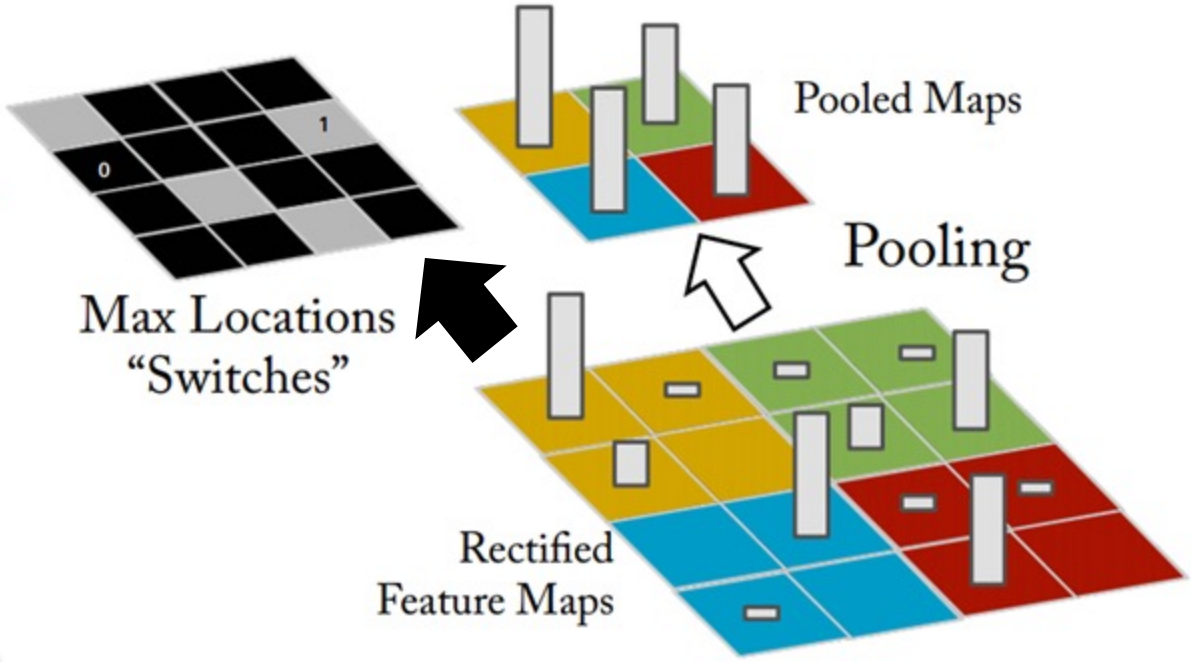
CNN -Unpooling



14 x 14



28 x 28

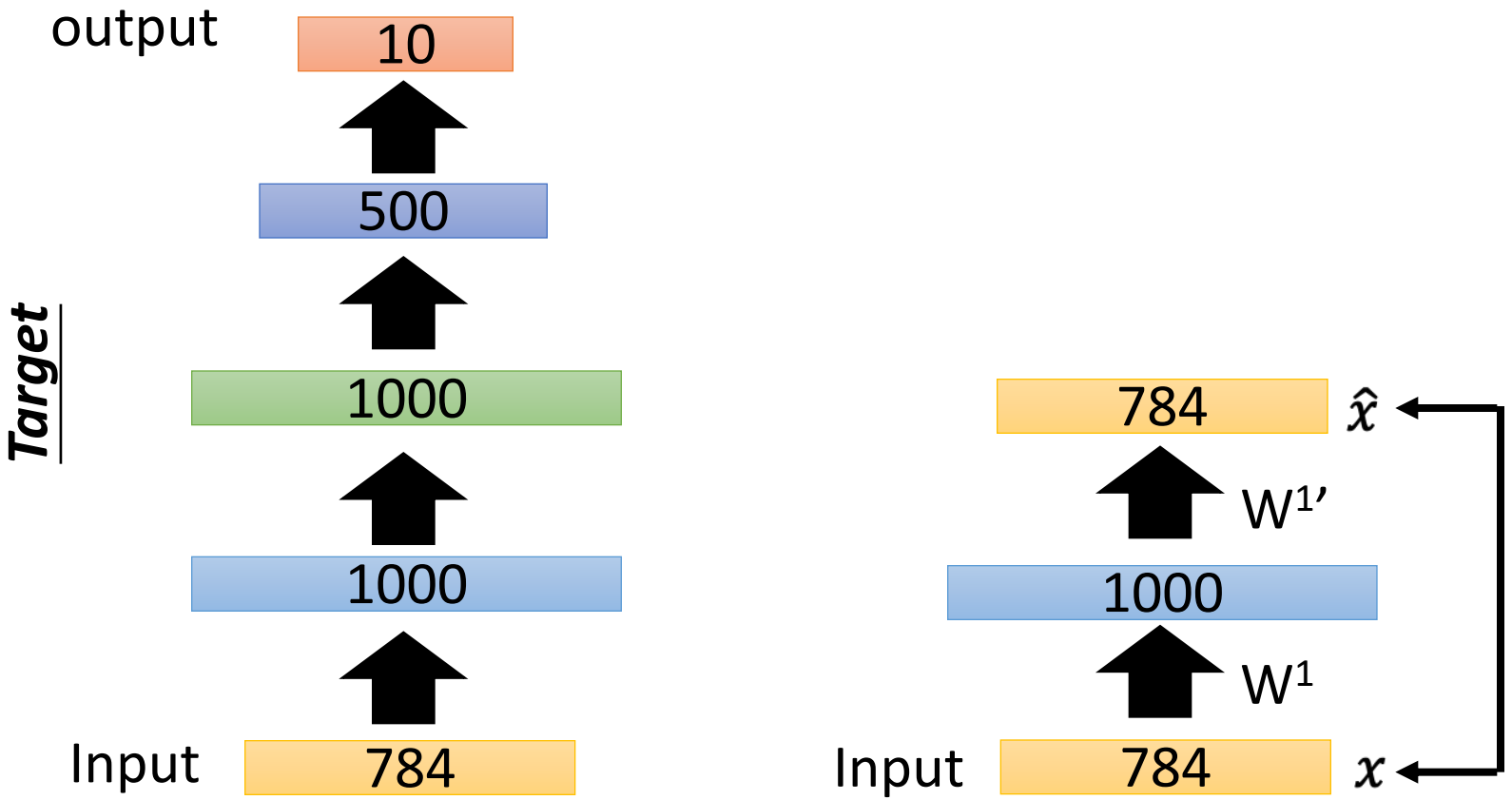


Alternative: simply repeat the values

Source of image : https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/image_segmentation.html

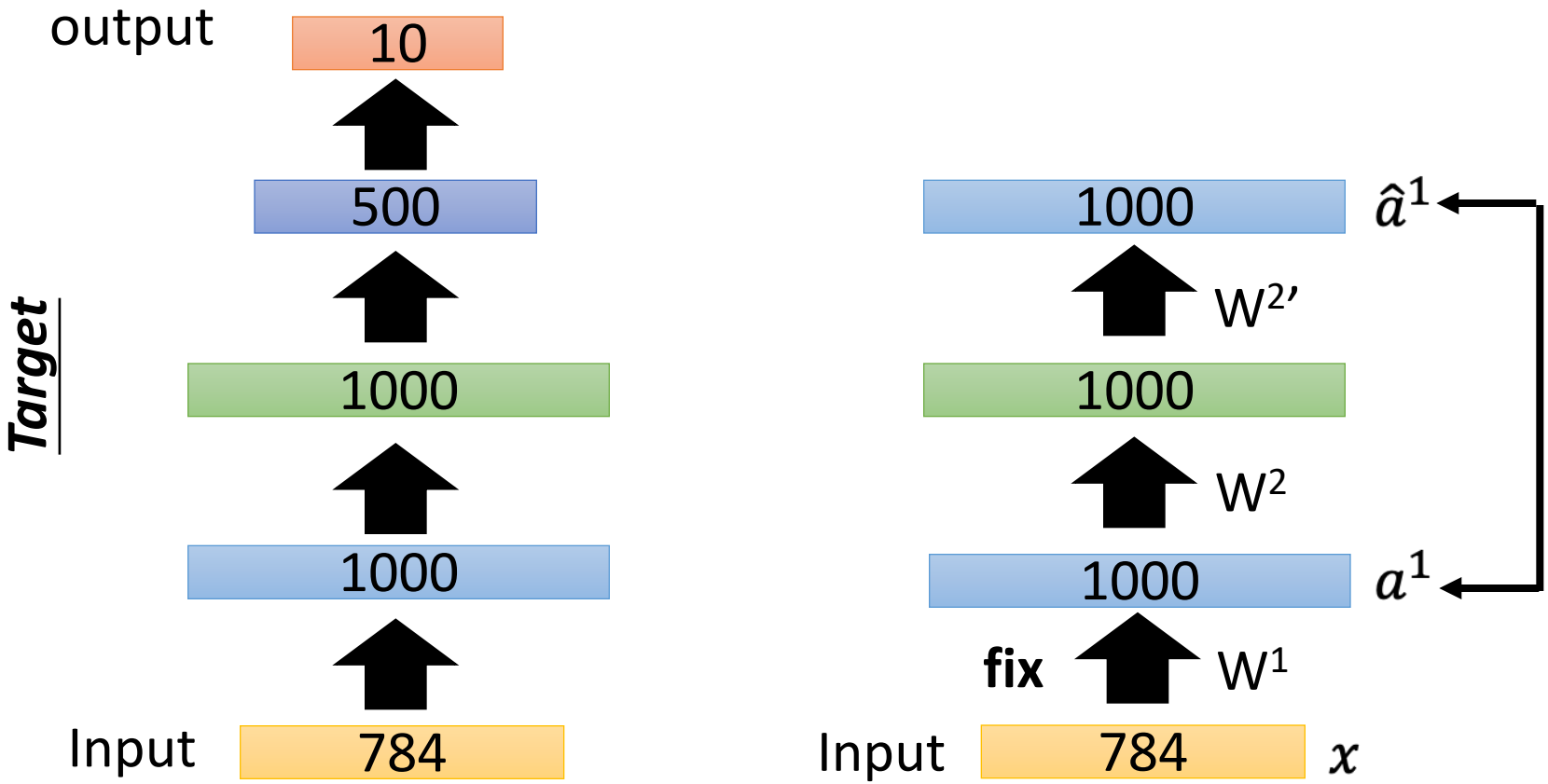
Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*



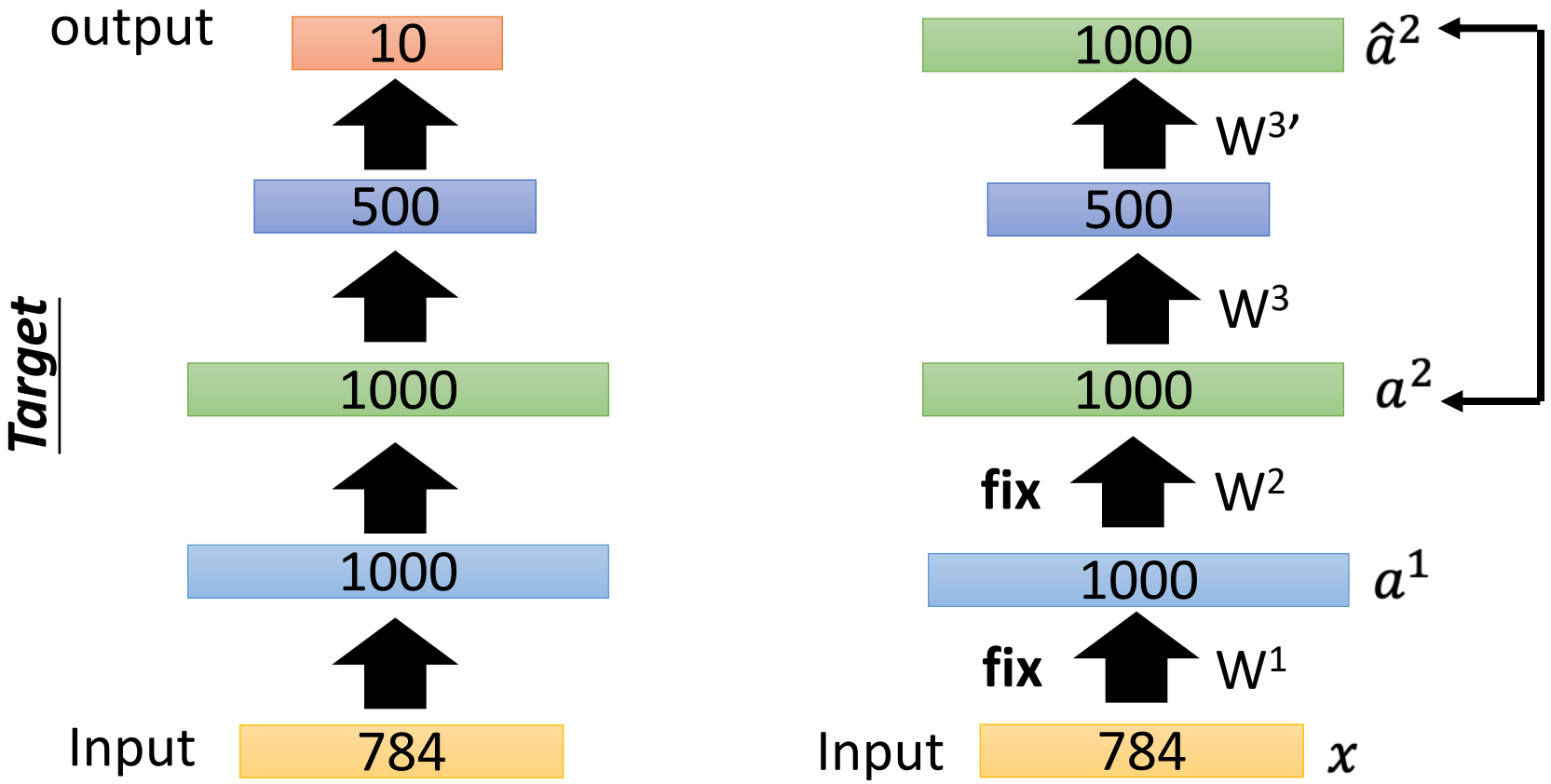
Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*



Auto-encoder – Pre-training DNN

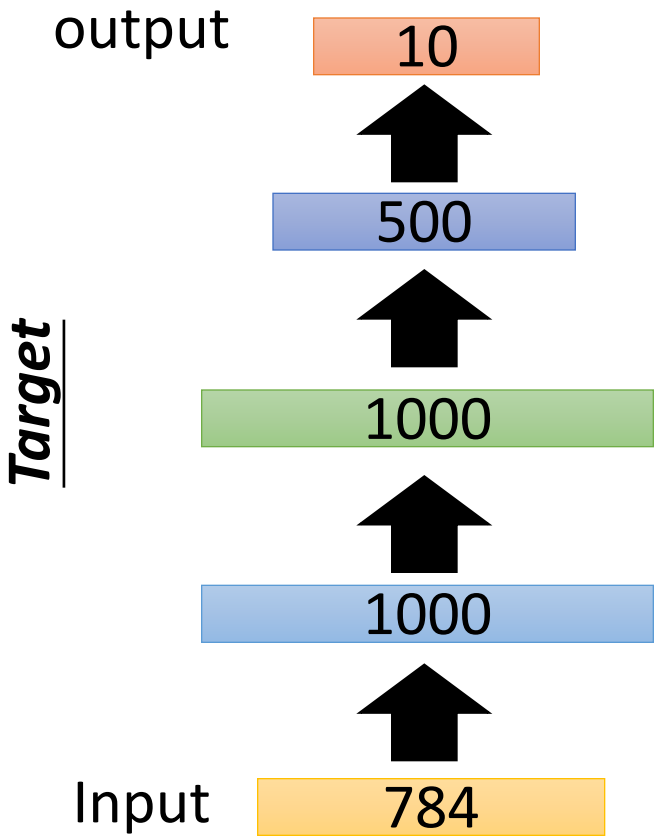
- Greedy Layer-wise Pre-training *again*

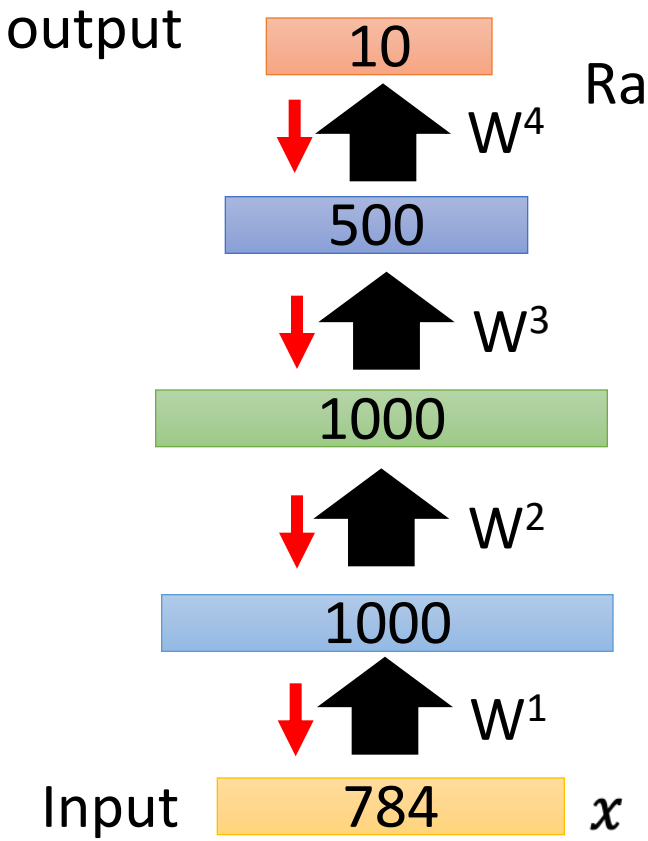


Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*

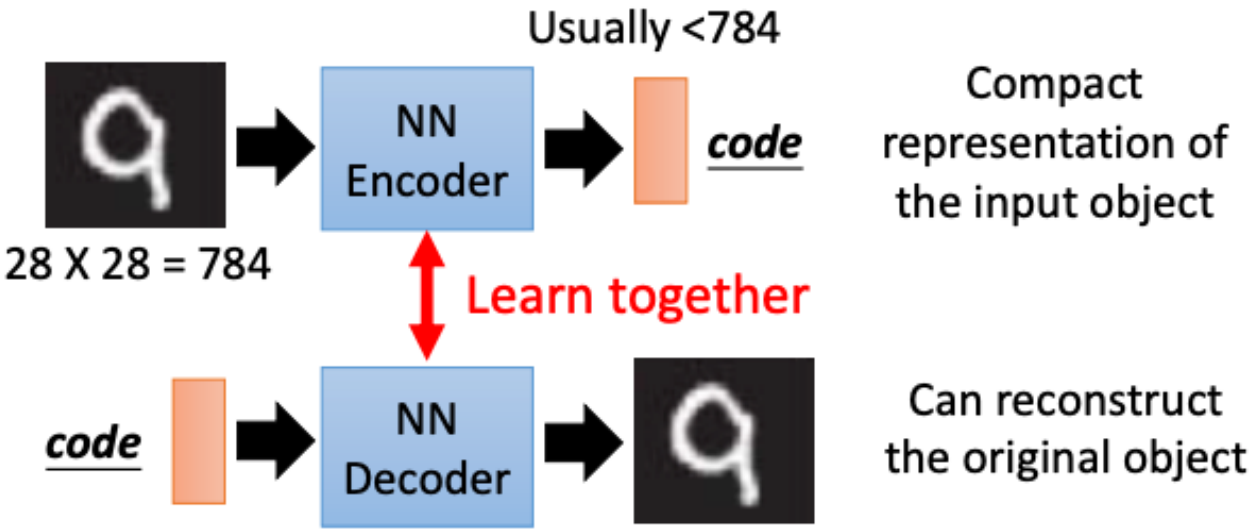
Find-tune by backpropagation





Can we use Decoder to Generate Something?

Training



Generator

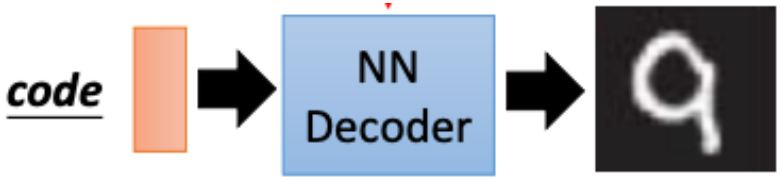
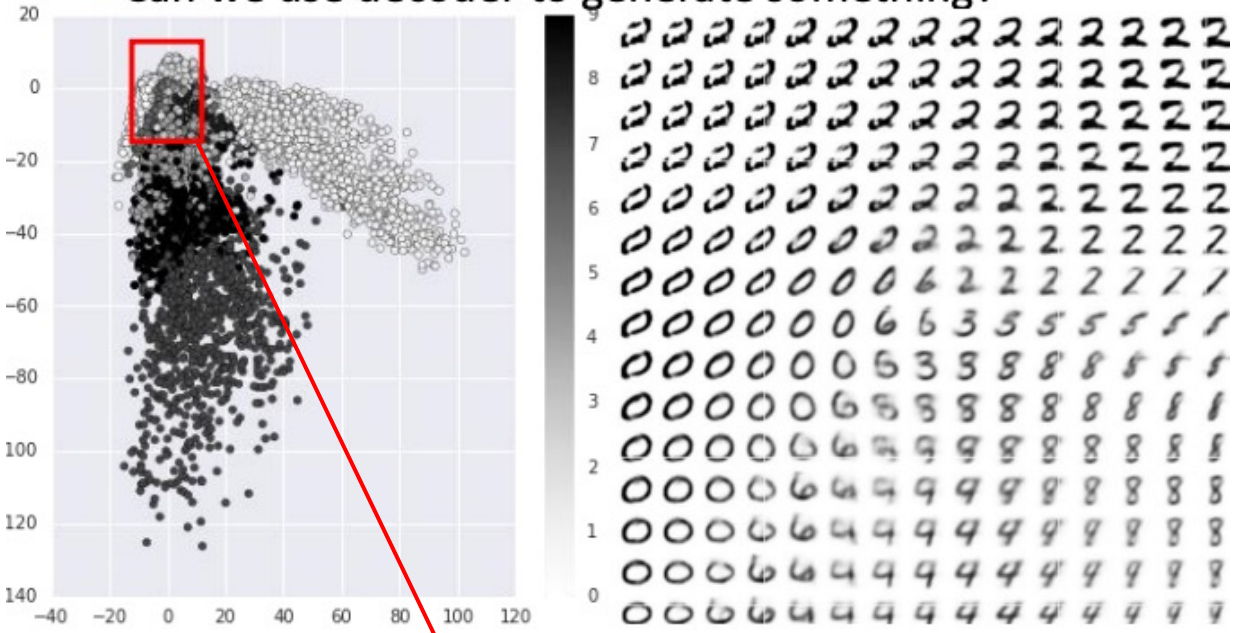


Image Generator

- Can we use decoder to generate something?



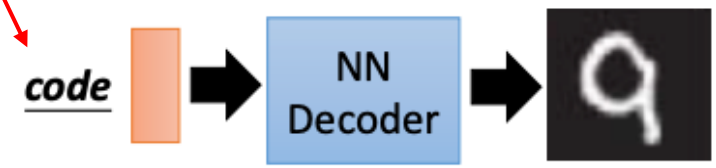


Image source: