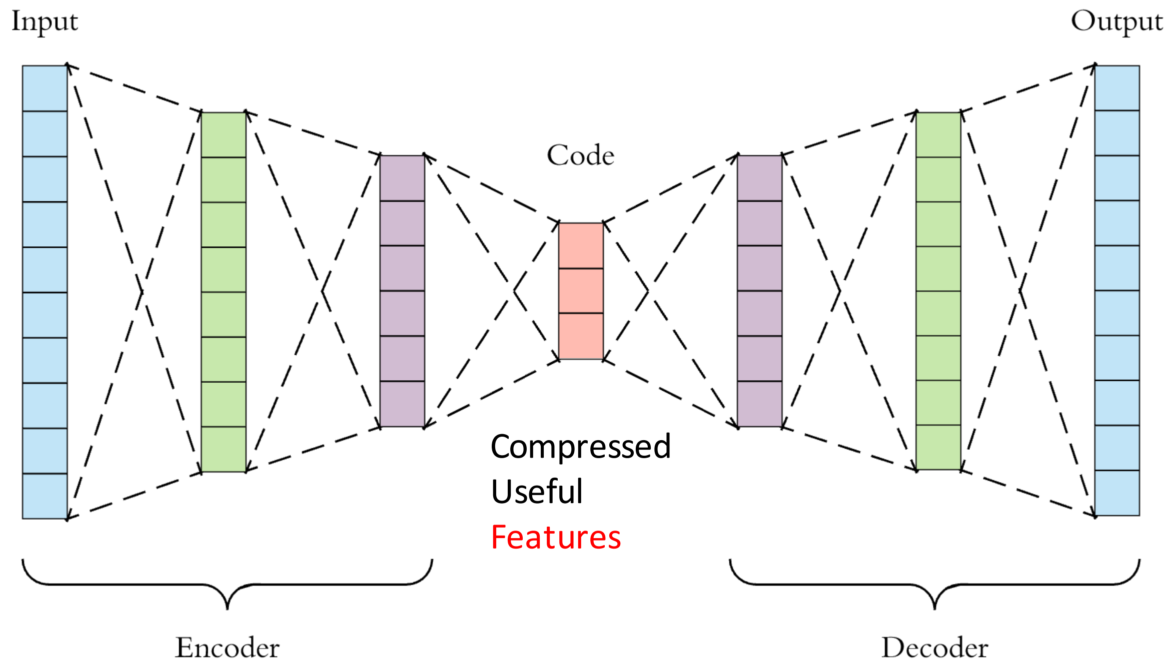
CSCE 636 Neural Networks (Deep Learning)

Lecture 16: Auto-Encoder

Anxiao (Andrew) Jiang

Based on the interesting lecture of Prof. Hung-yi Lee "Unsupervised Learning: Deep Auto-Encoder"
https://www.youtube.com/watch?v=Tk5B4seA-AU&list=PLJV_el3uVTsPy9oCRY30oBPNLCo89yu49&index=25

Auto-Encoder

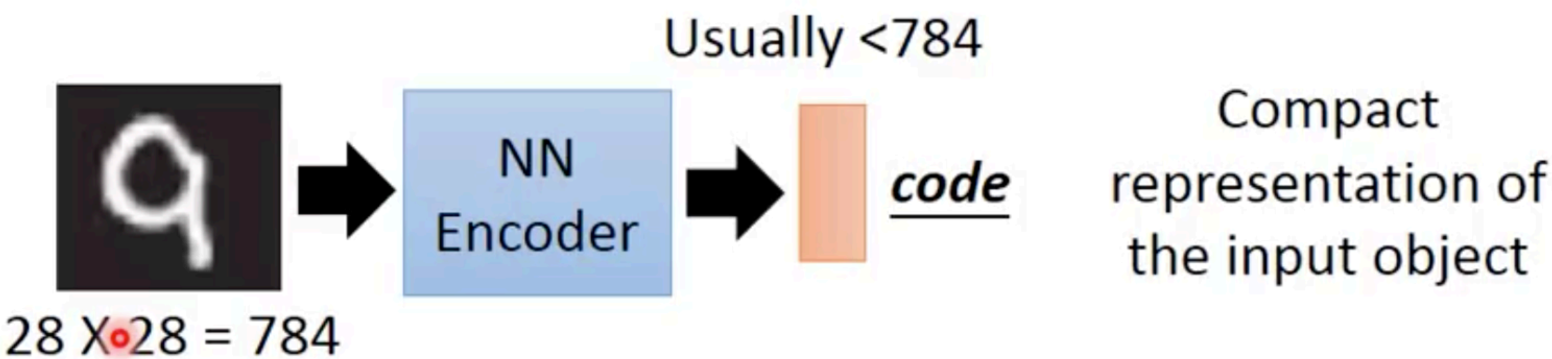


Methods: (1) Make output data be as close to input data as possible
(2) Limit the size of the encoder's output

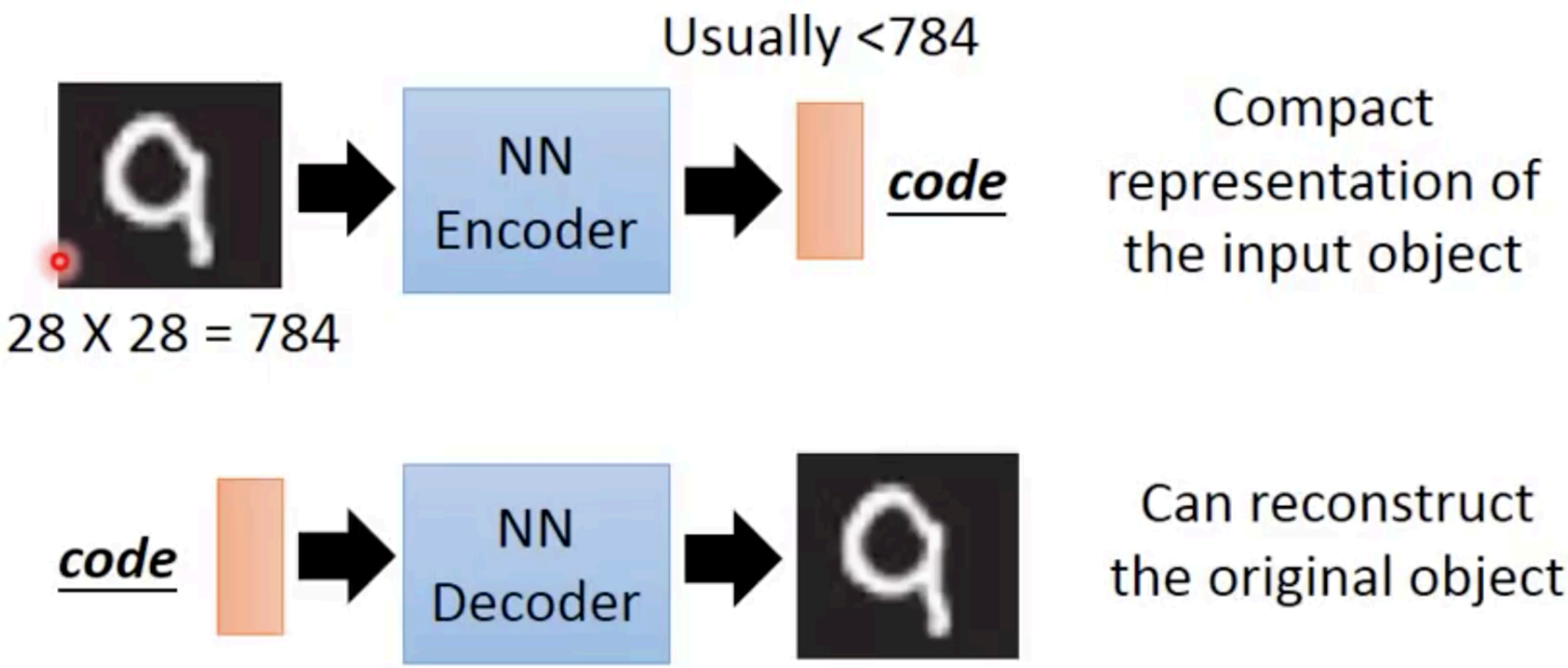
Use of encoder: it can extract useful features from input data; the useful features can be used by many applications.
Use of decoder: it can generate realistic data from random inputs.

Nice property of auto-encoder: it can be trained without labels for data.

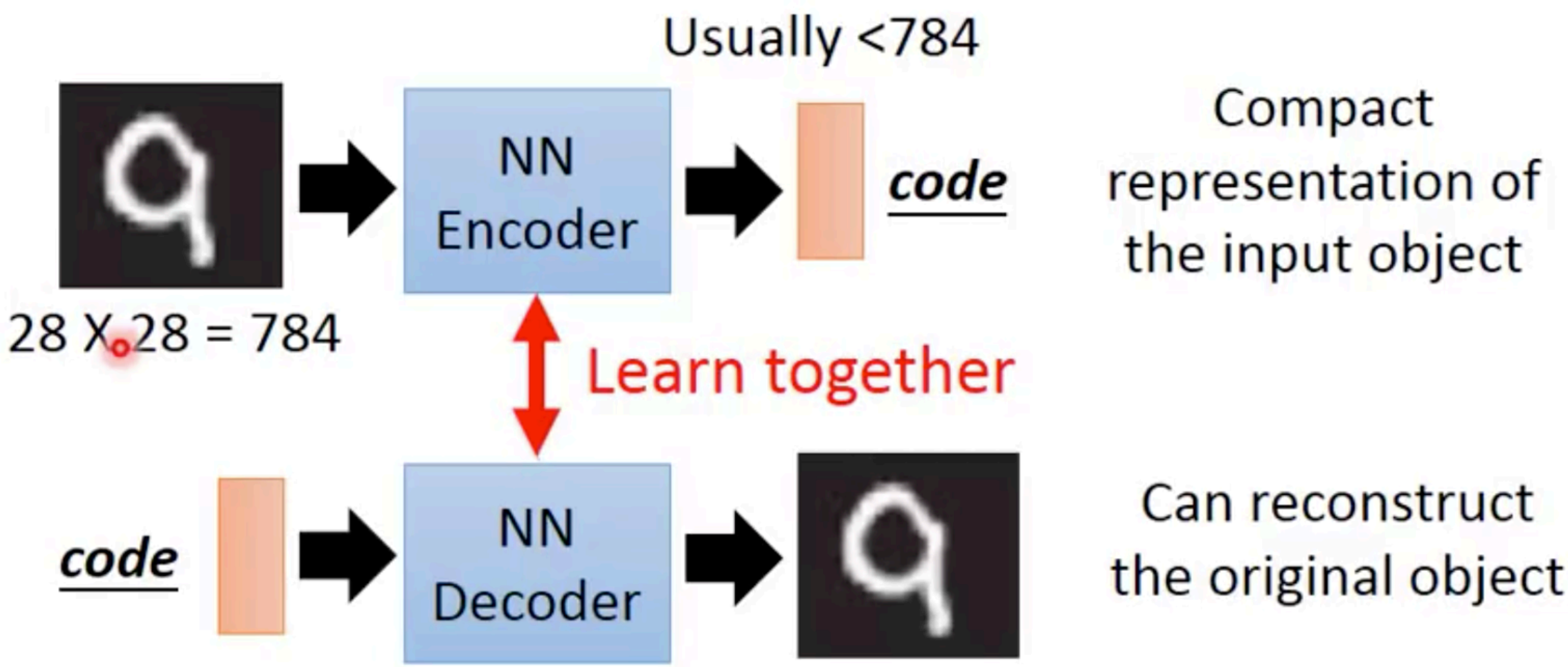
Auto-encoder



Auto-encoder

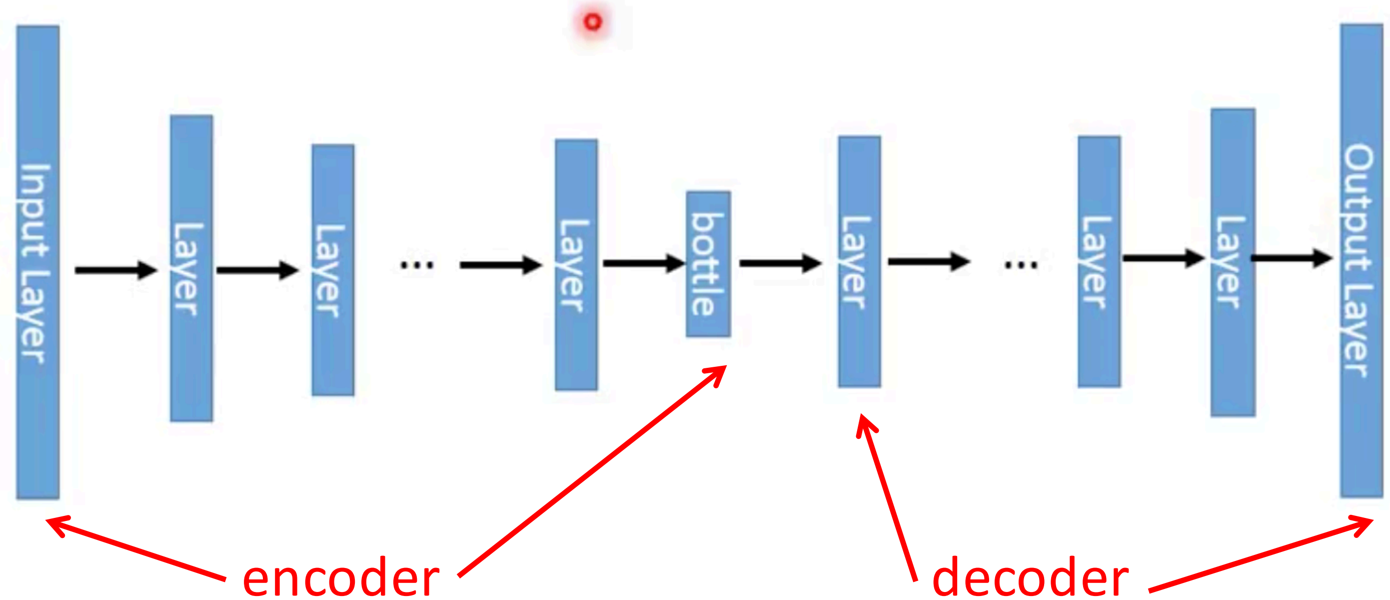


Auto-encoder



Deep Auto-encoder

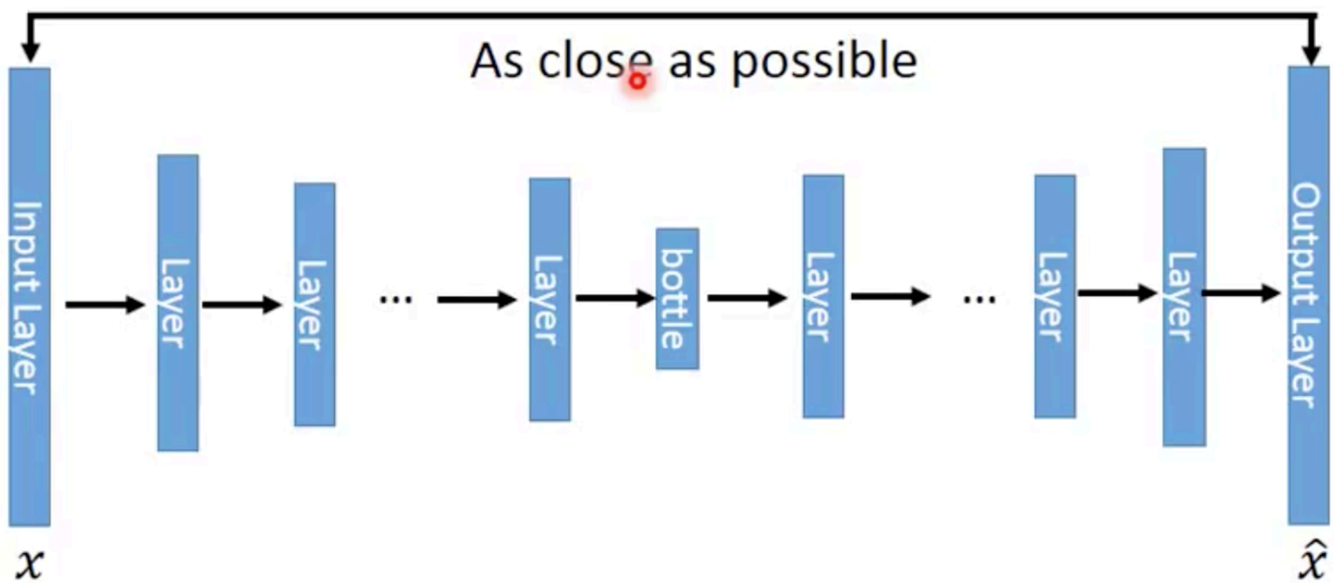
- Of course, the auto-encoder can be deep



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

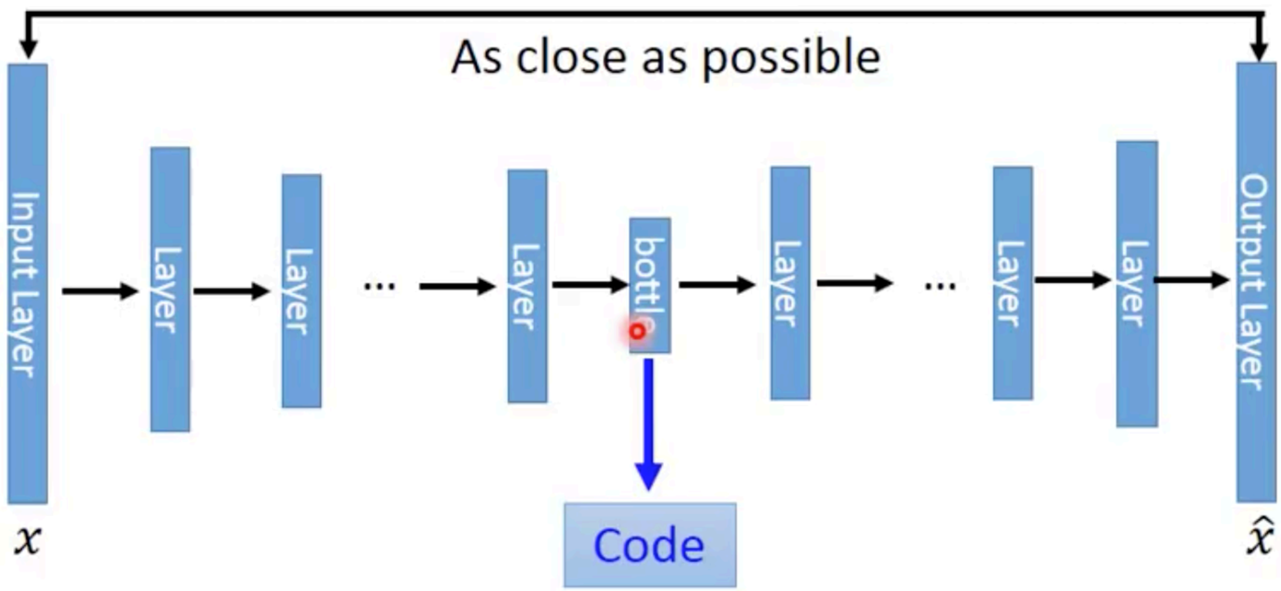
- Of course, the auto-encoder can be deep



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

- Of course, the auto-encoder can be deep

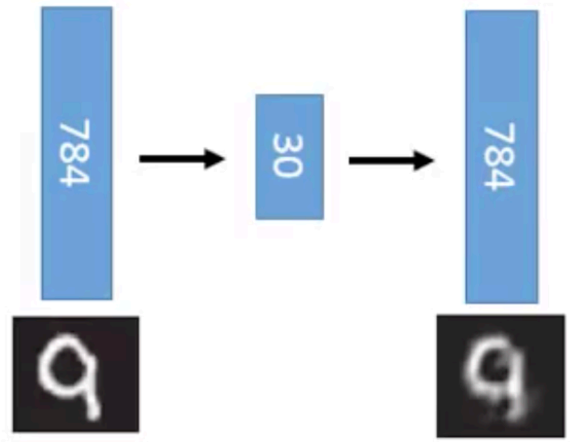


Reference: Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." *Science* 313.5786 (2006): 504-507

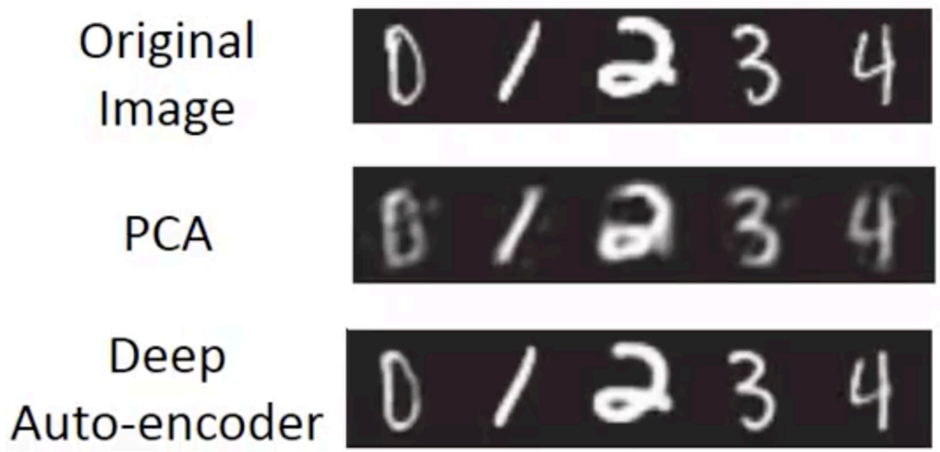
Deep Auto-encoder

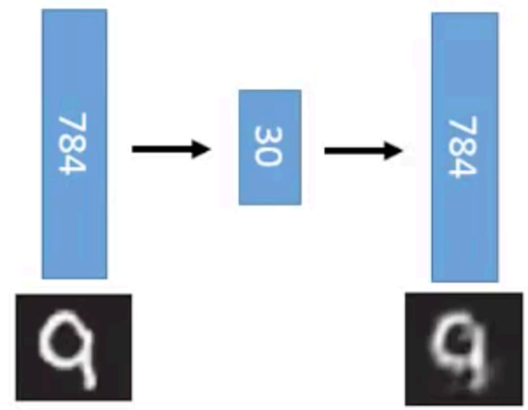
Original Image

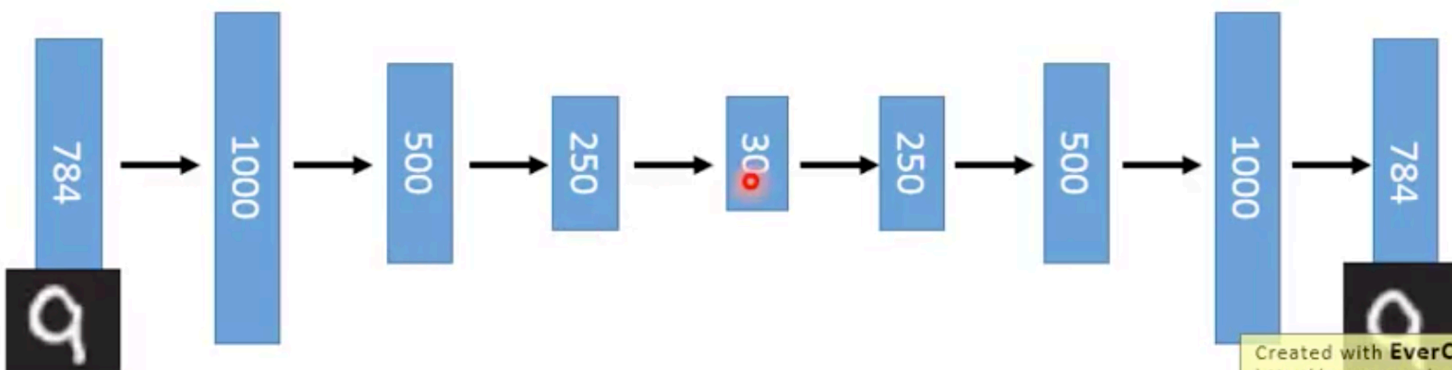
PCA

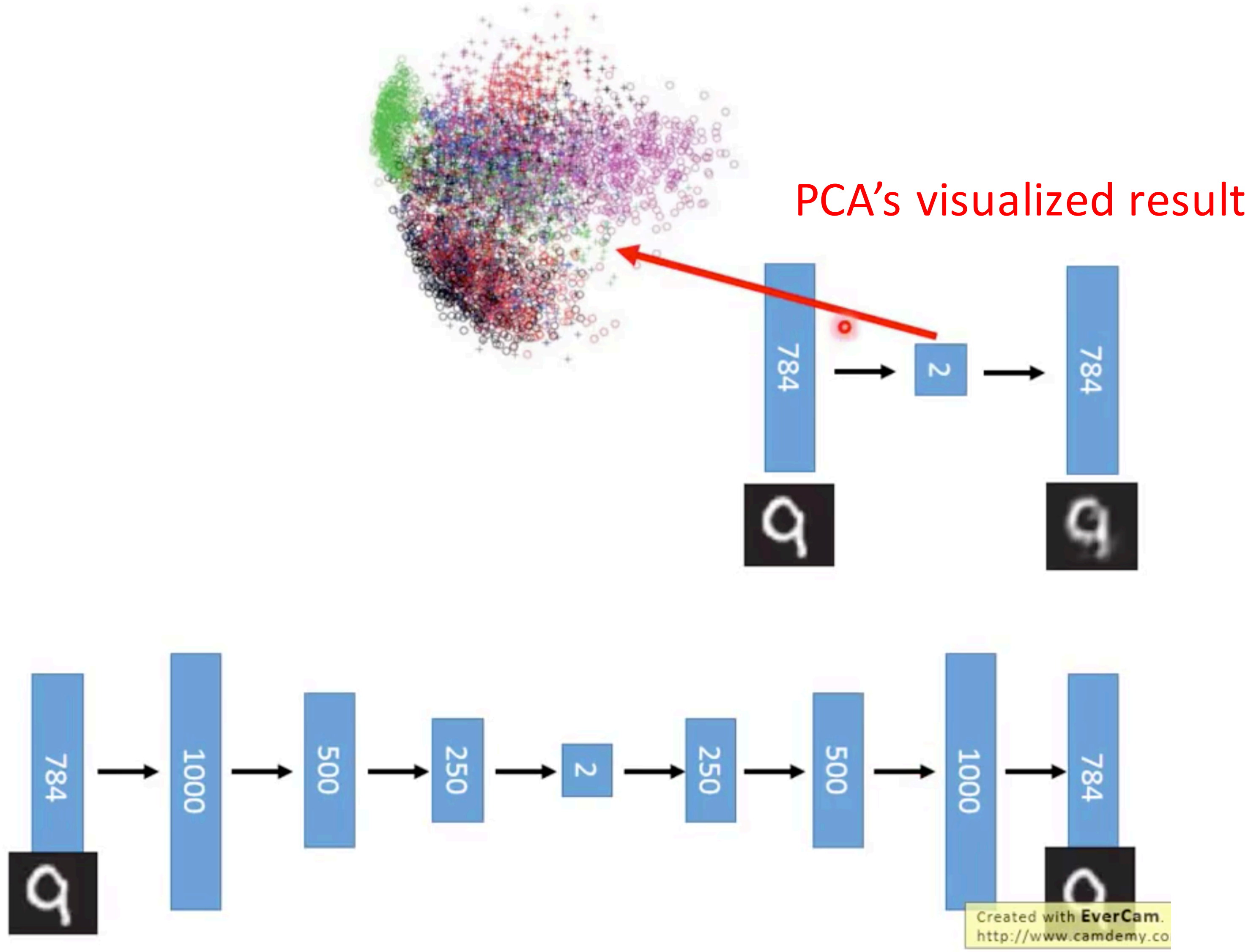


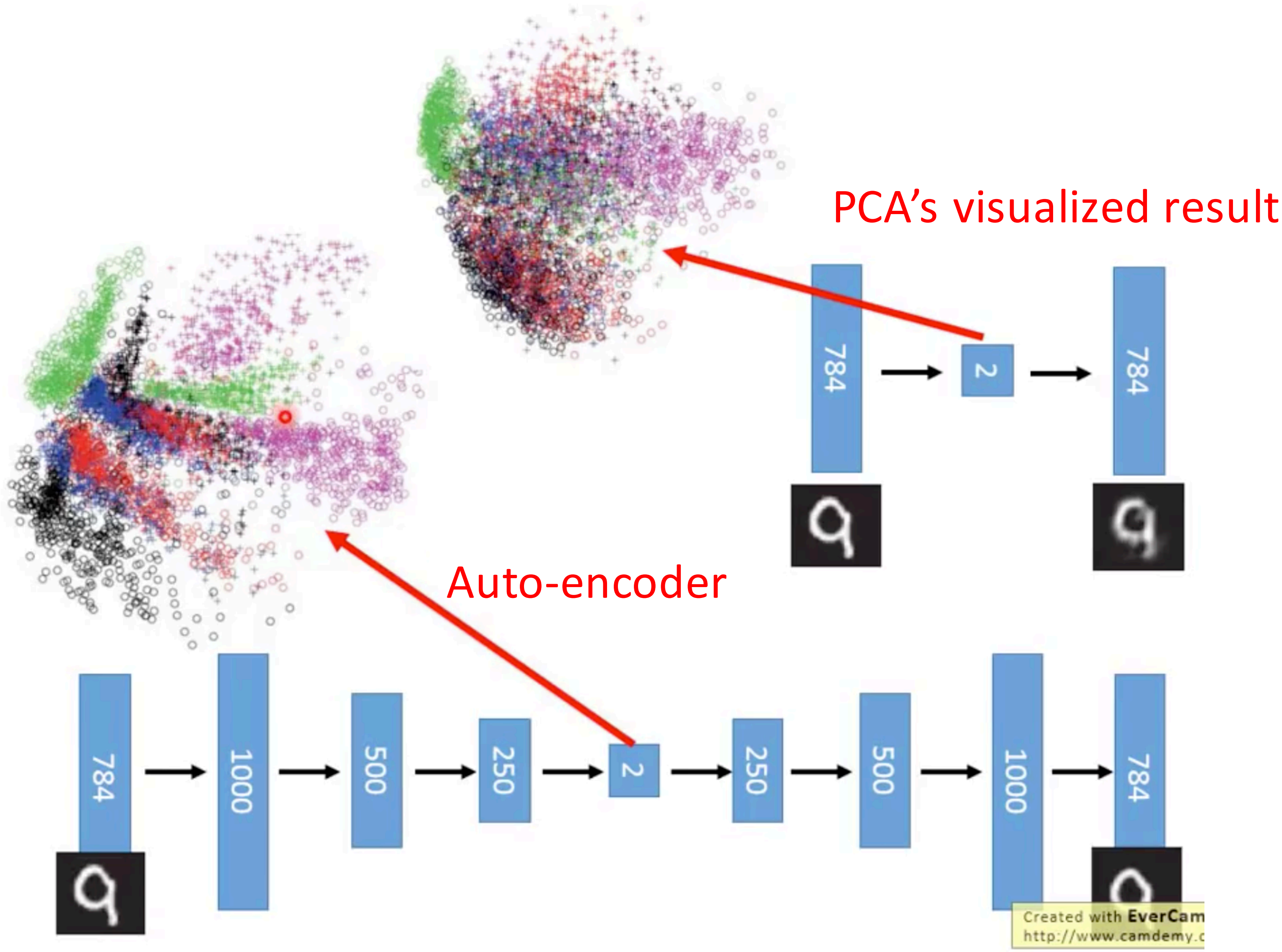
Deep Auto-encoder







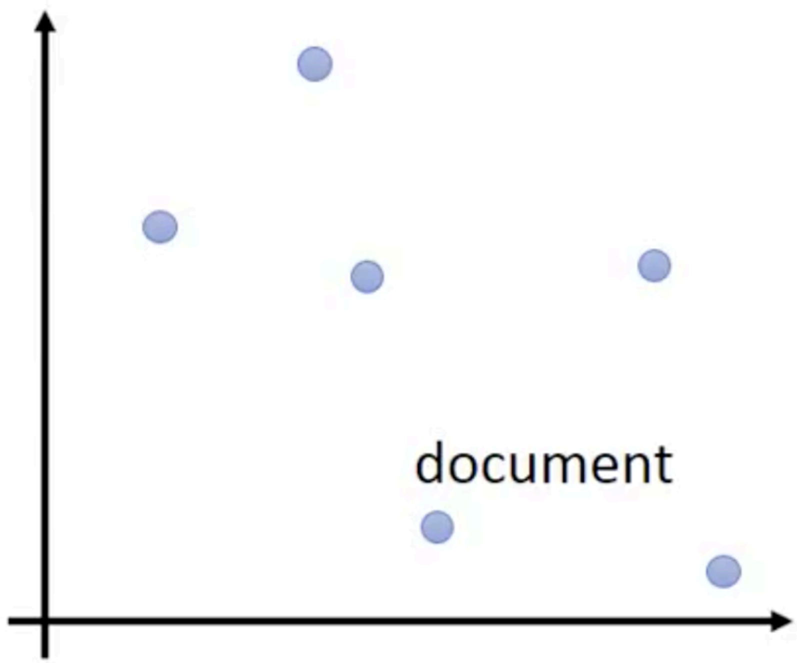

PCA’s visualized result
784
2
784
784
1000
500
250
2
250
500
1000
784
Created with EverCam.
http://www.camdemy.co


PCA's visualized result
784
2
784
Auto-encoder
784
1000
500
250
2
250
500
1000
784
Created with EverCam
http://www.camdemy.c

Auto-encoder – Text Retrieval

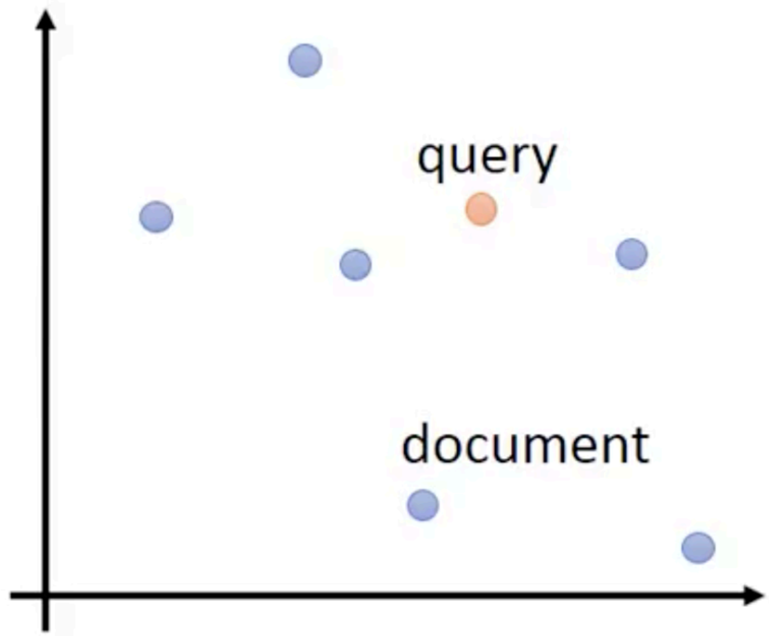
Auto-encoder – Text Retrieval

Vector Space Model



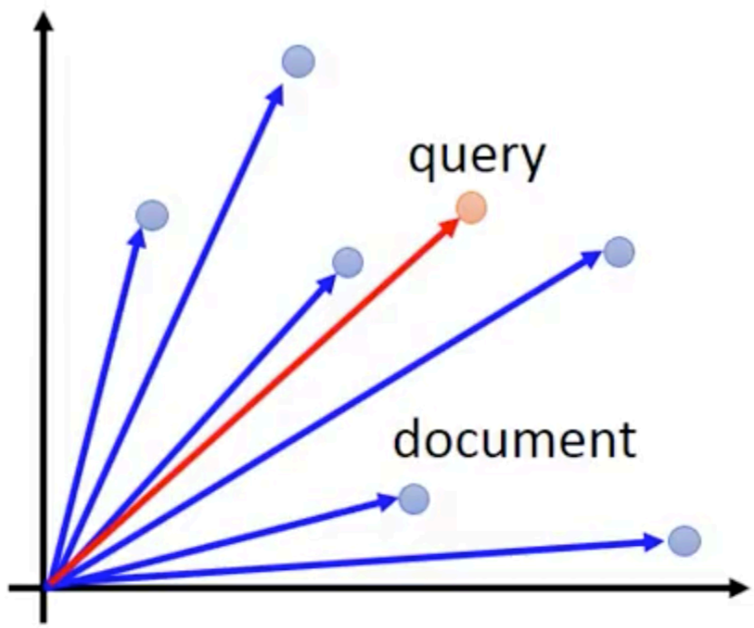
Auto-encoder – Text Retrieval

Vector Space Model



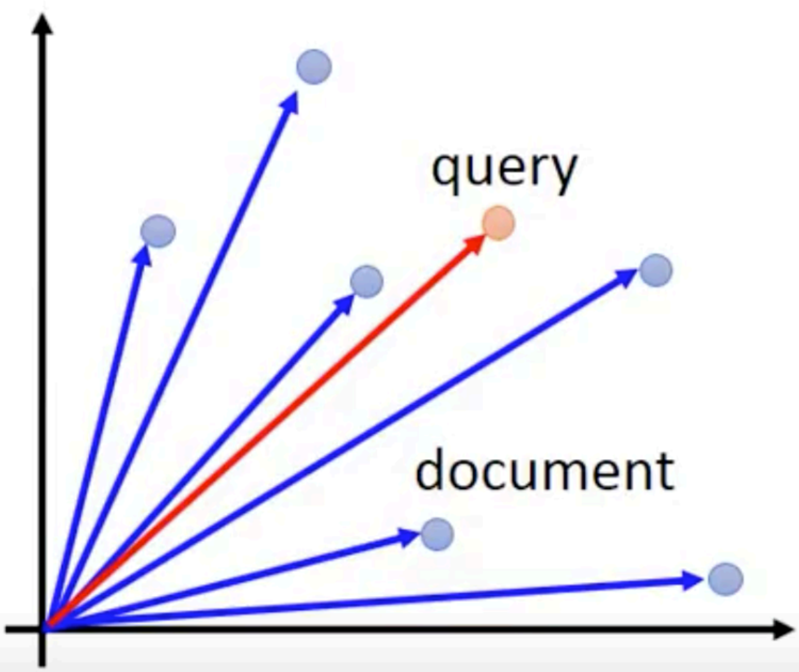
Auto-encoder – Text Retrieval

Vector Space Model



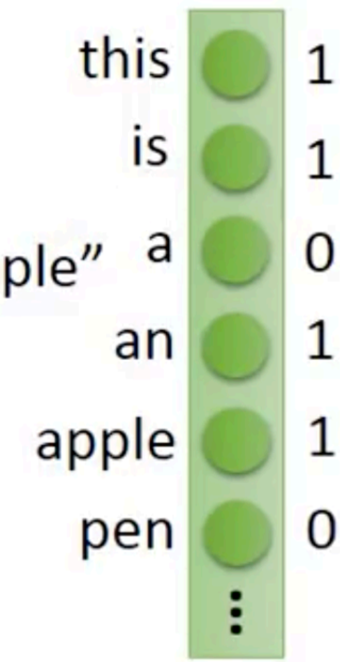
Auto-encoder – Text Retrieval

Vector Space Model



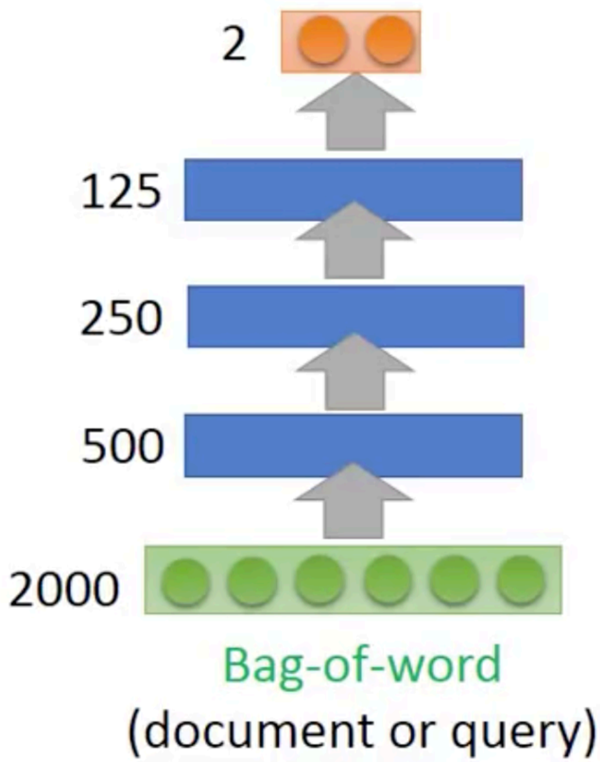
Bag-of-word

word string:
“This is an apple”



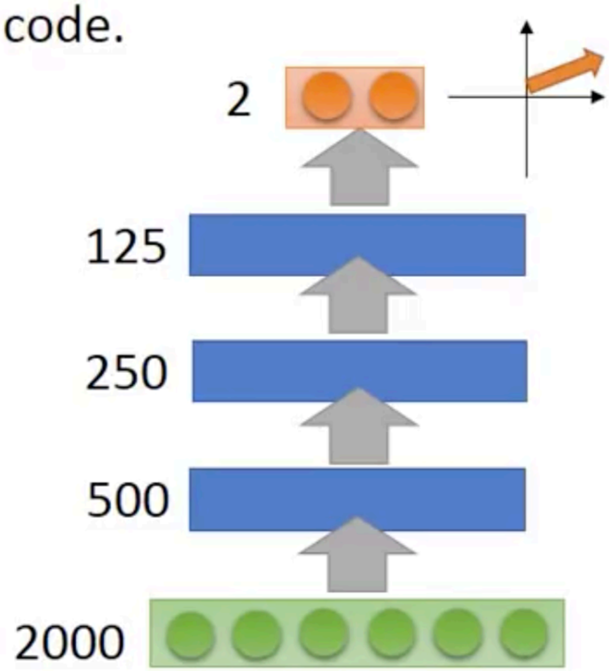
Auto-encoder – Text Retrieval

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

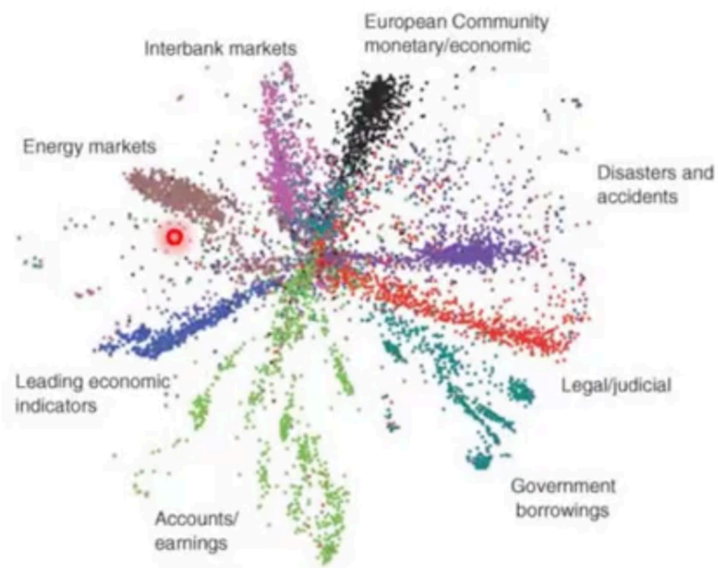


Auto-encoder – Text Retrieval

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

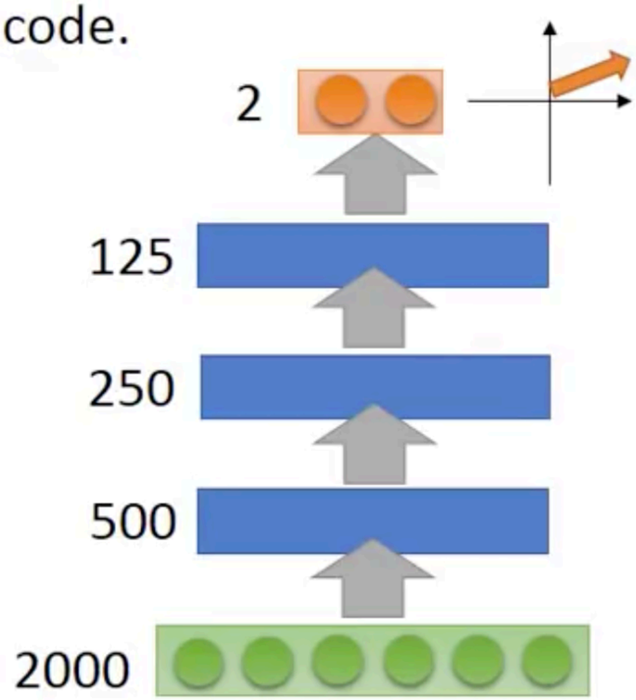


Bag-of-word
(document or query)

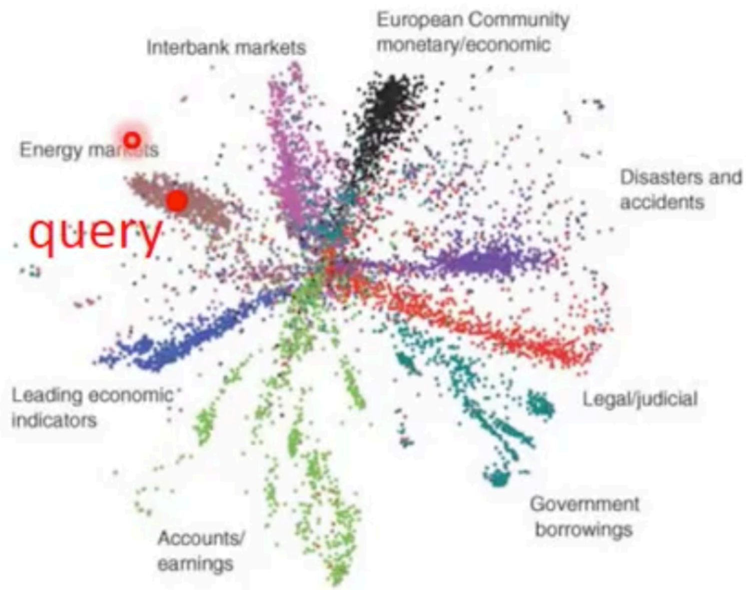


Auto-encoder – Text Retrieval

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



Bag-of-word
(document or query)



Auto-encoder – Similar Image Search

Retrieved using Euclidean distance in pixel intensity space

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

Auto-encoder – Similar Image Search

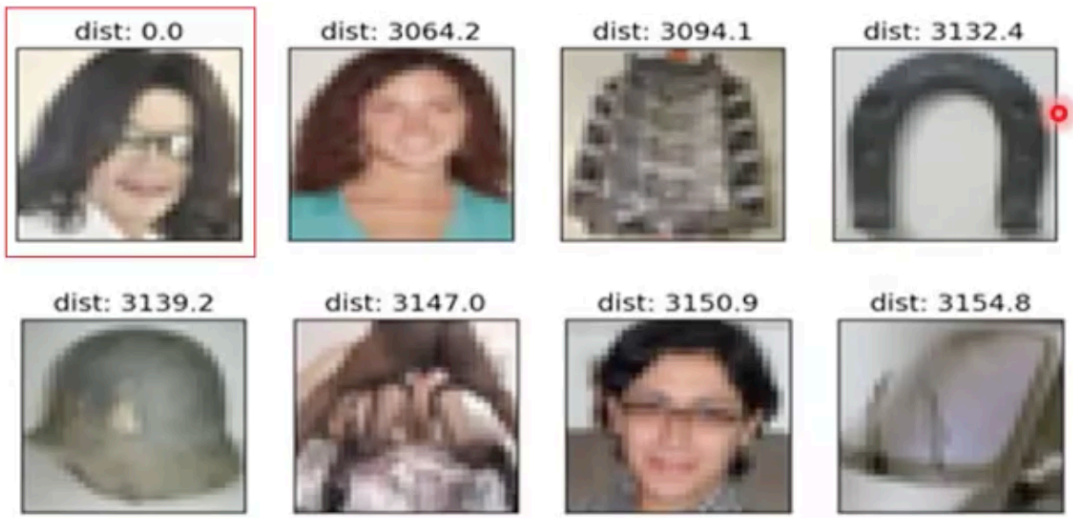
Retrieved using Euclidean distance in pixel intensity space

(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

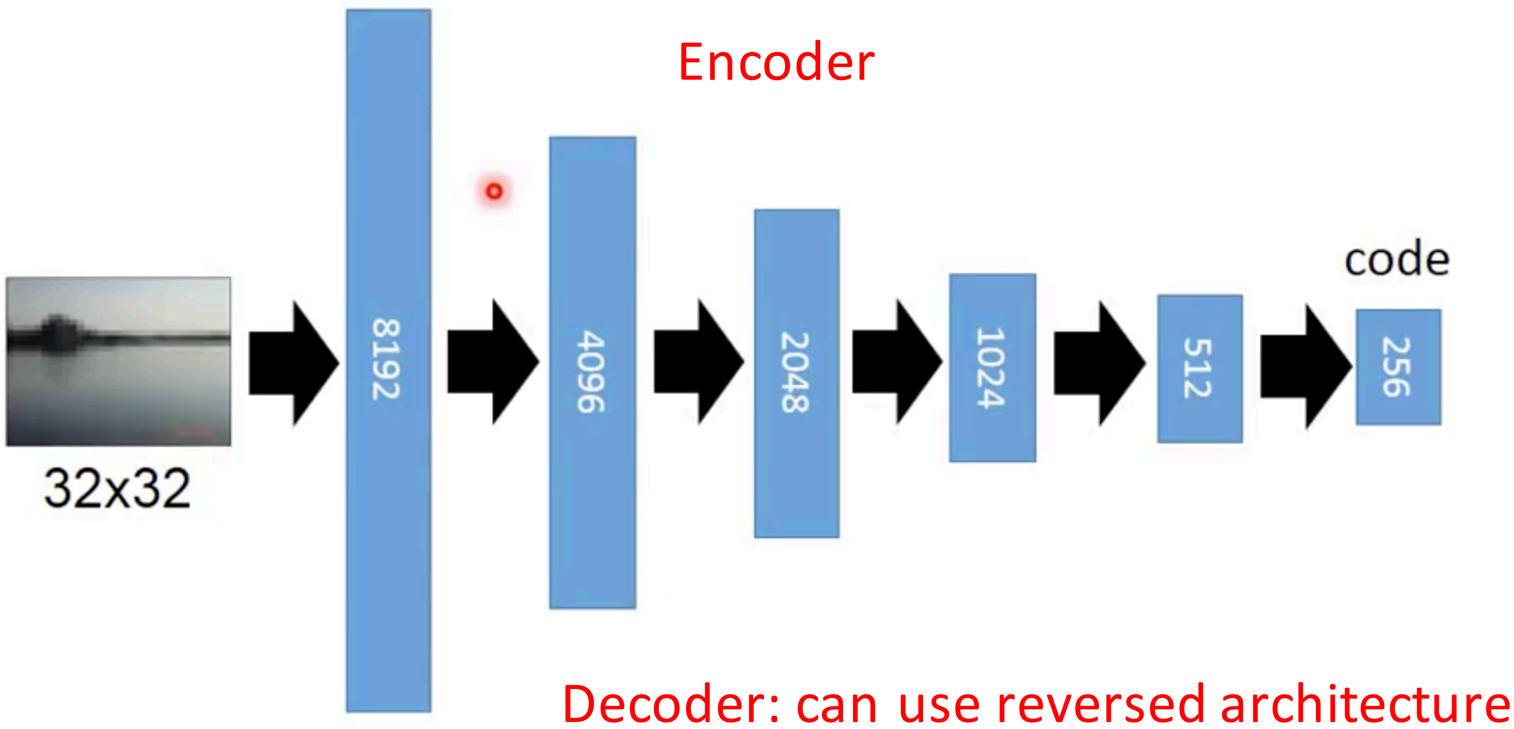
Retrieved using Euclidean distance in pixel intensity space



(Images from Hinton's slides on Coursera)

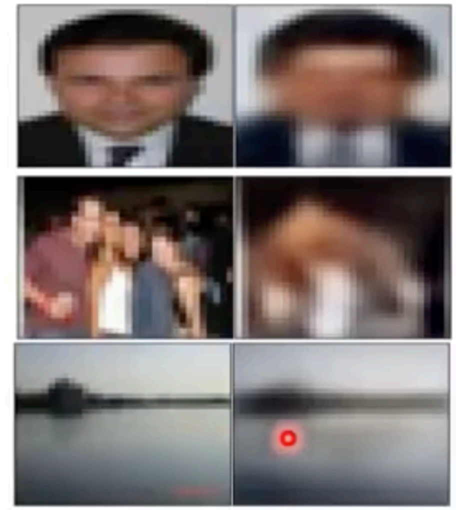
Reference: Krizhevsky, Alex, and Geoffrey E. Hinton. "Using very deep autoencoders for content-based image retrieval." *ESANN*. 2011.

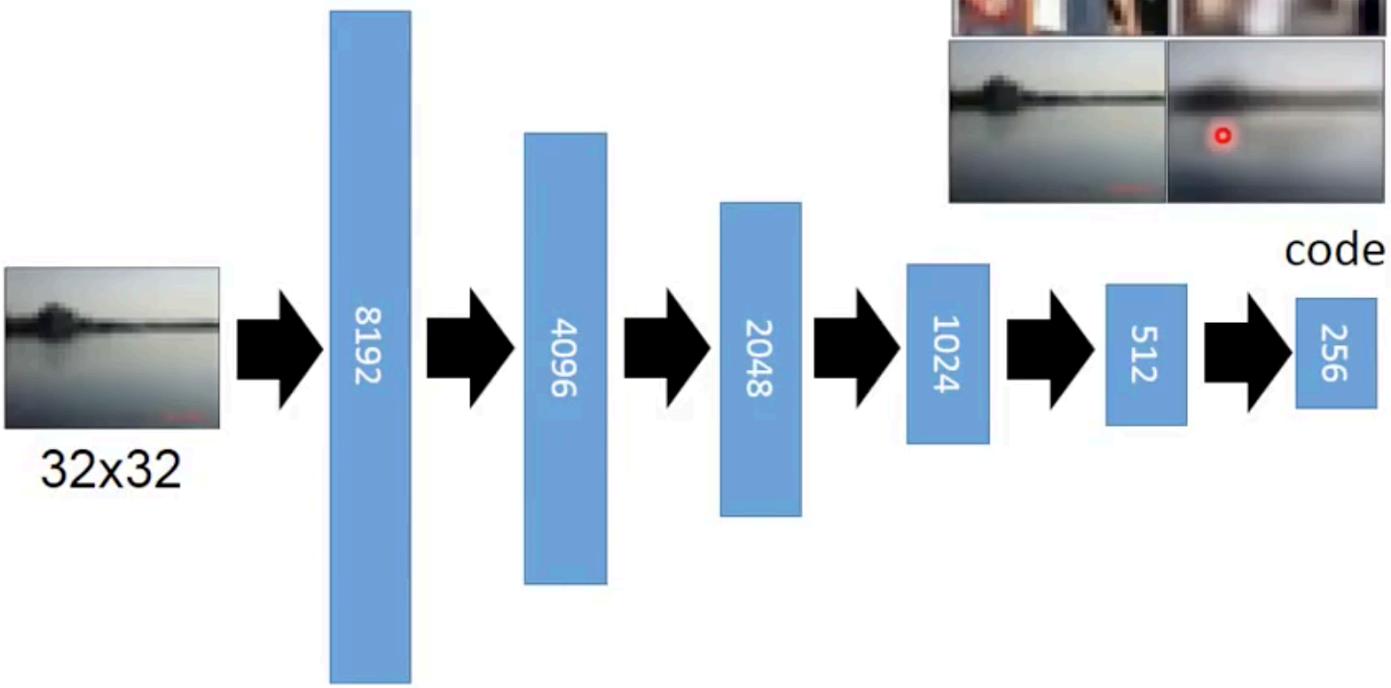
Auto-encoder –
Similar Image Search



(crawl millions of images from the Internet)

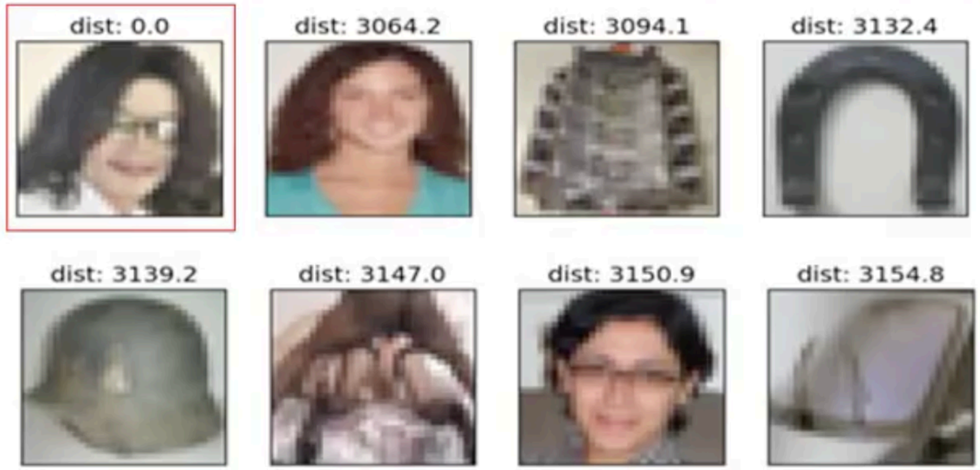
Auto-encoder – Similar Image Search



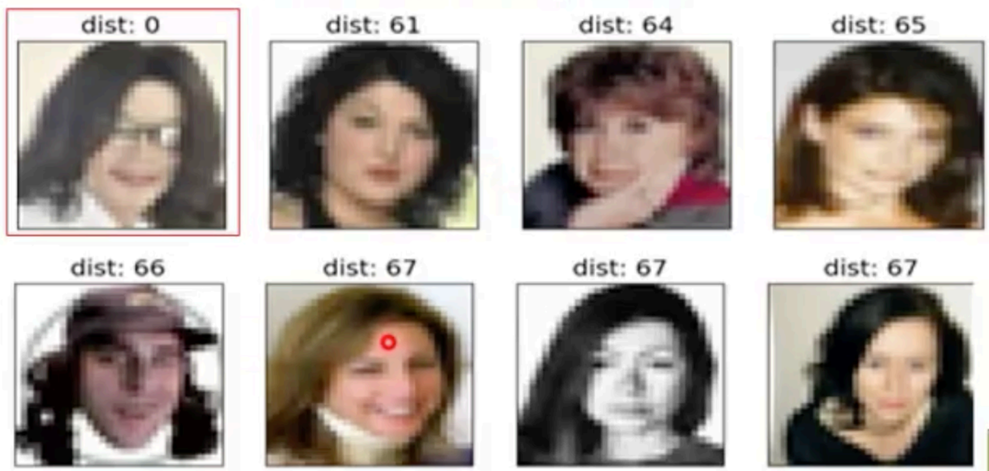


(crawl millions of images from the Internet)

Retrieved using Euclidean distance in pixel intensity space

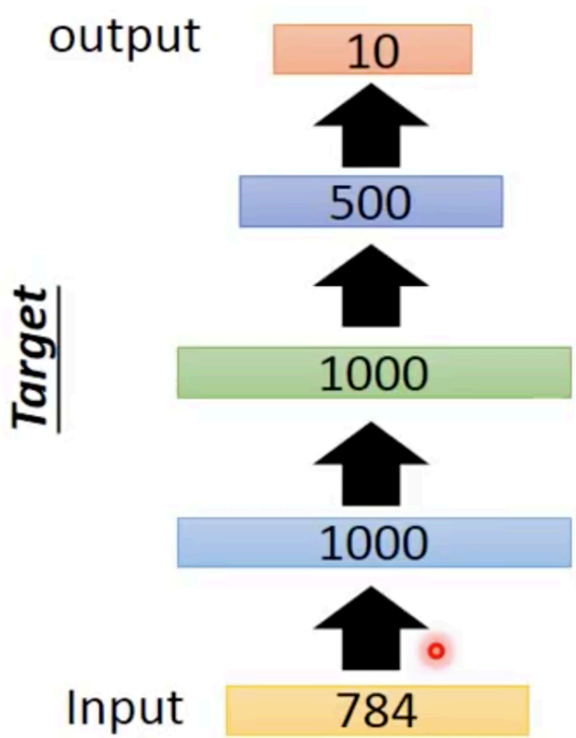


retrieved using 256 codes of auto-encoder



Auto-encoder – Pre-training DNN

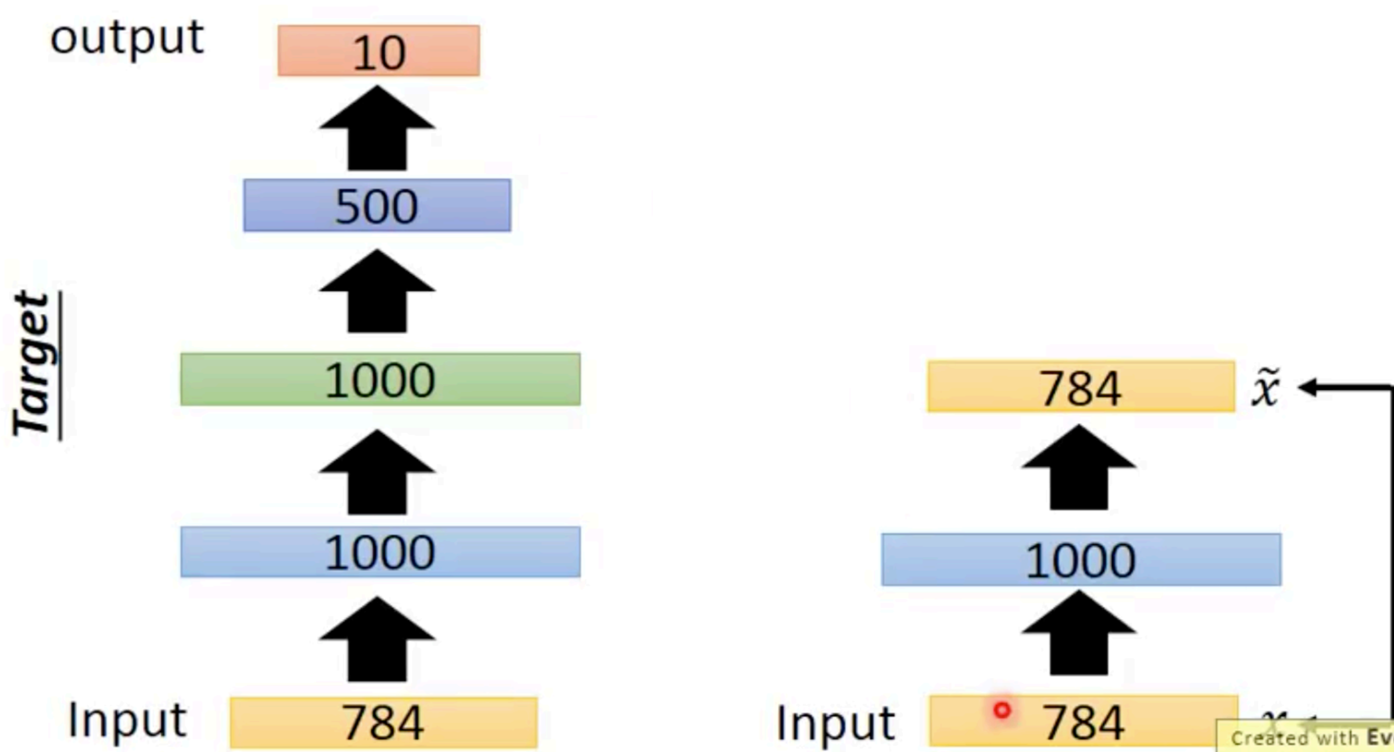
- Greedy Layer-wise Pre-training *again*



Pre-training:
find good initial values for weights

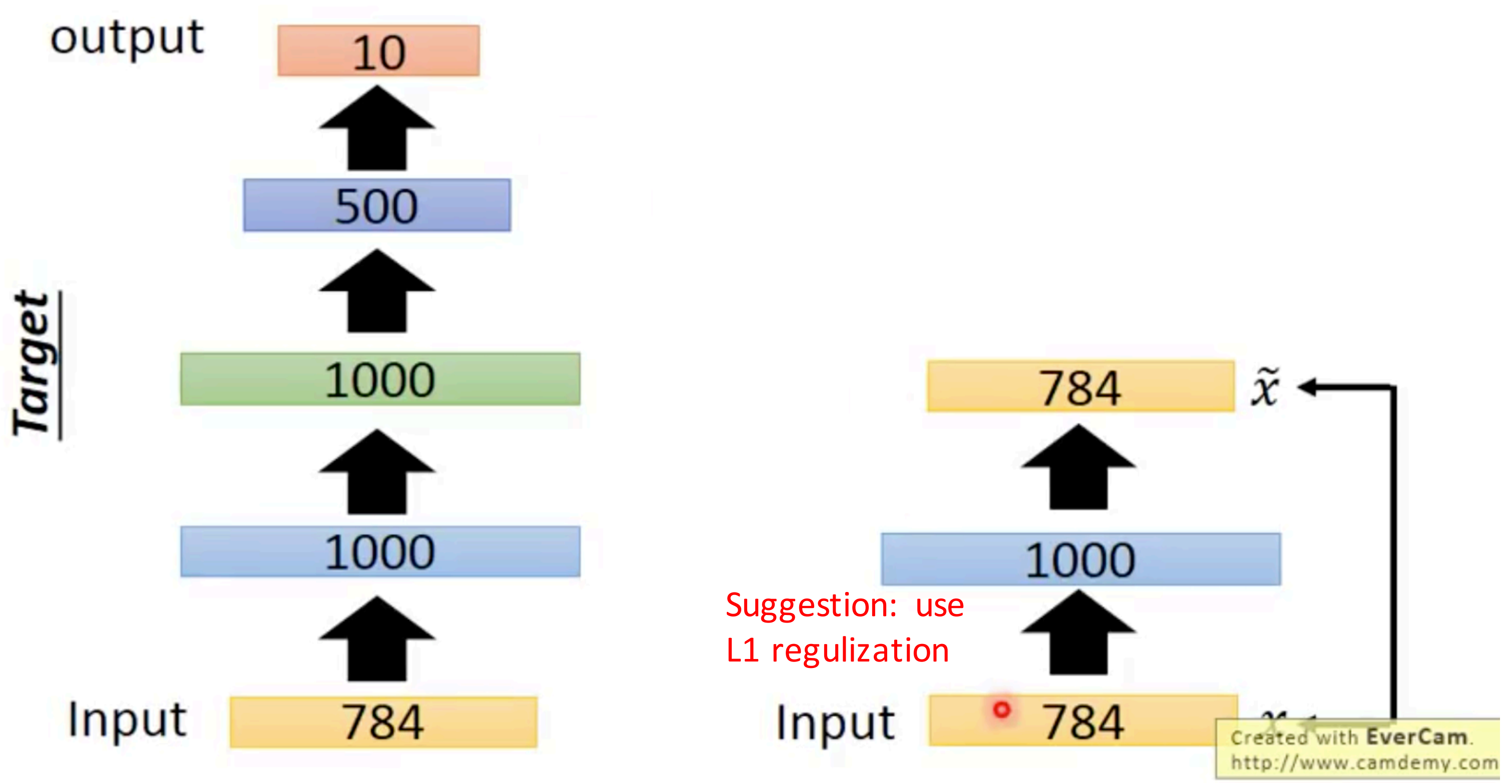
Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*



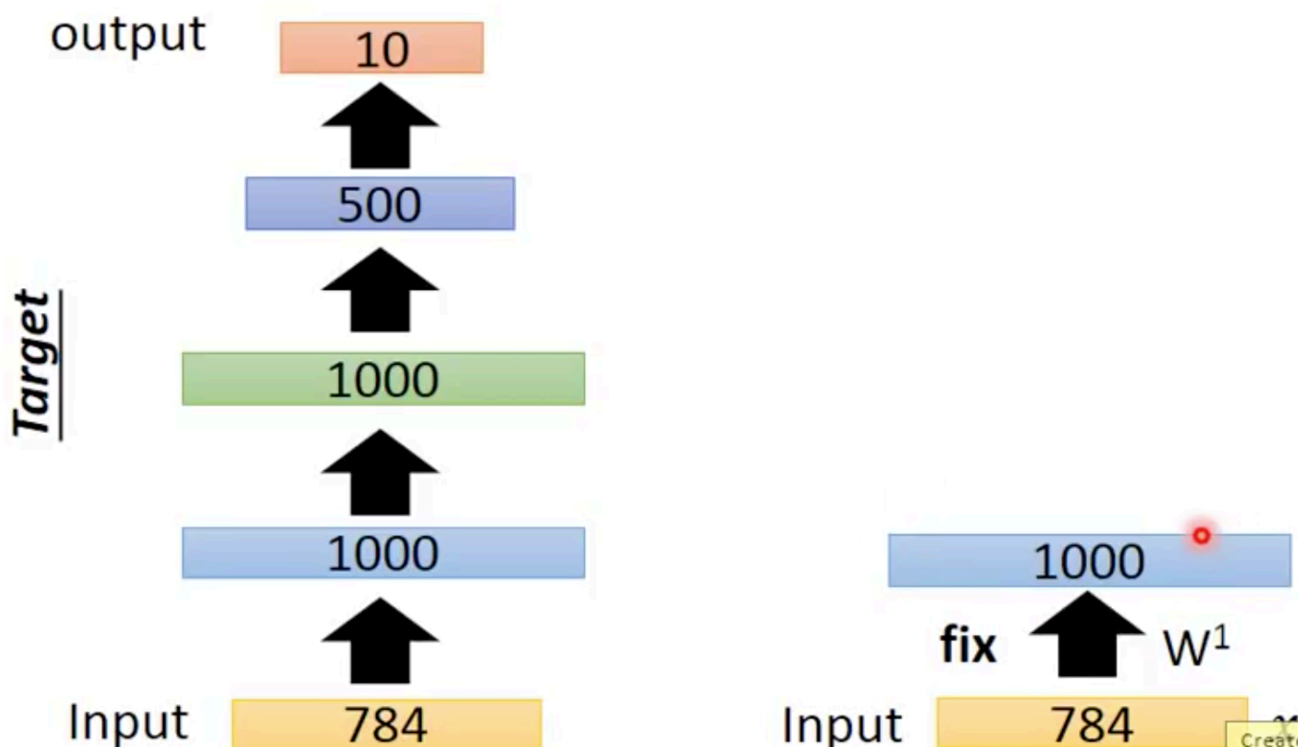
Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*



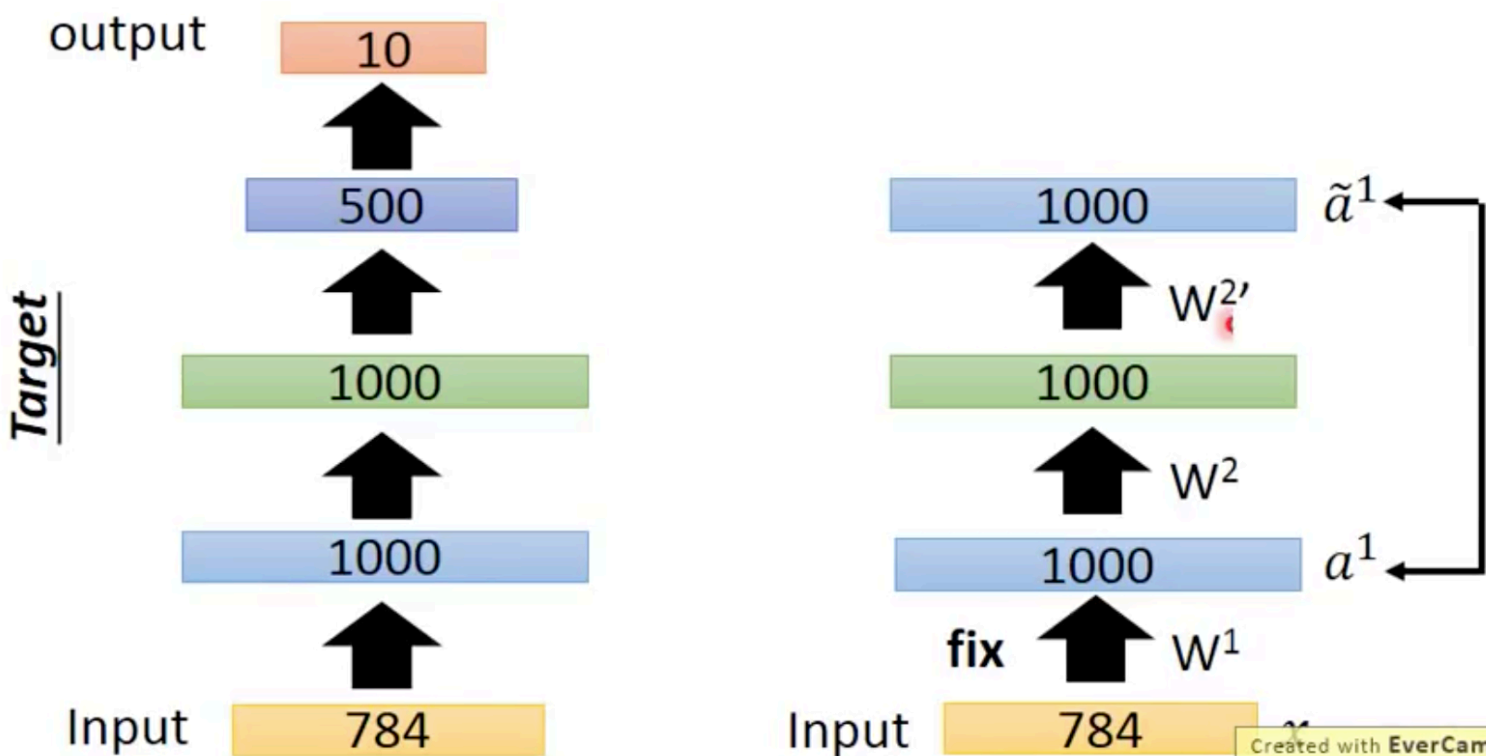
Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*



Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*

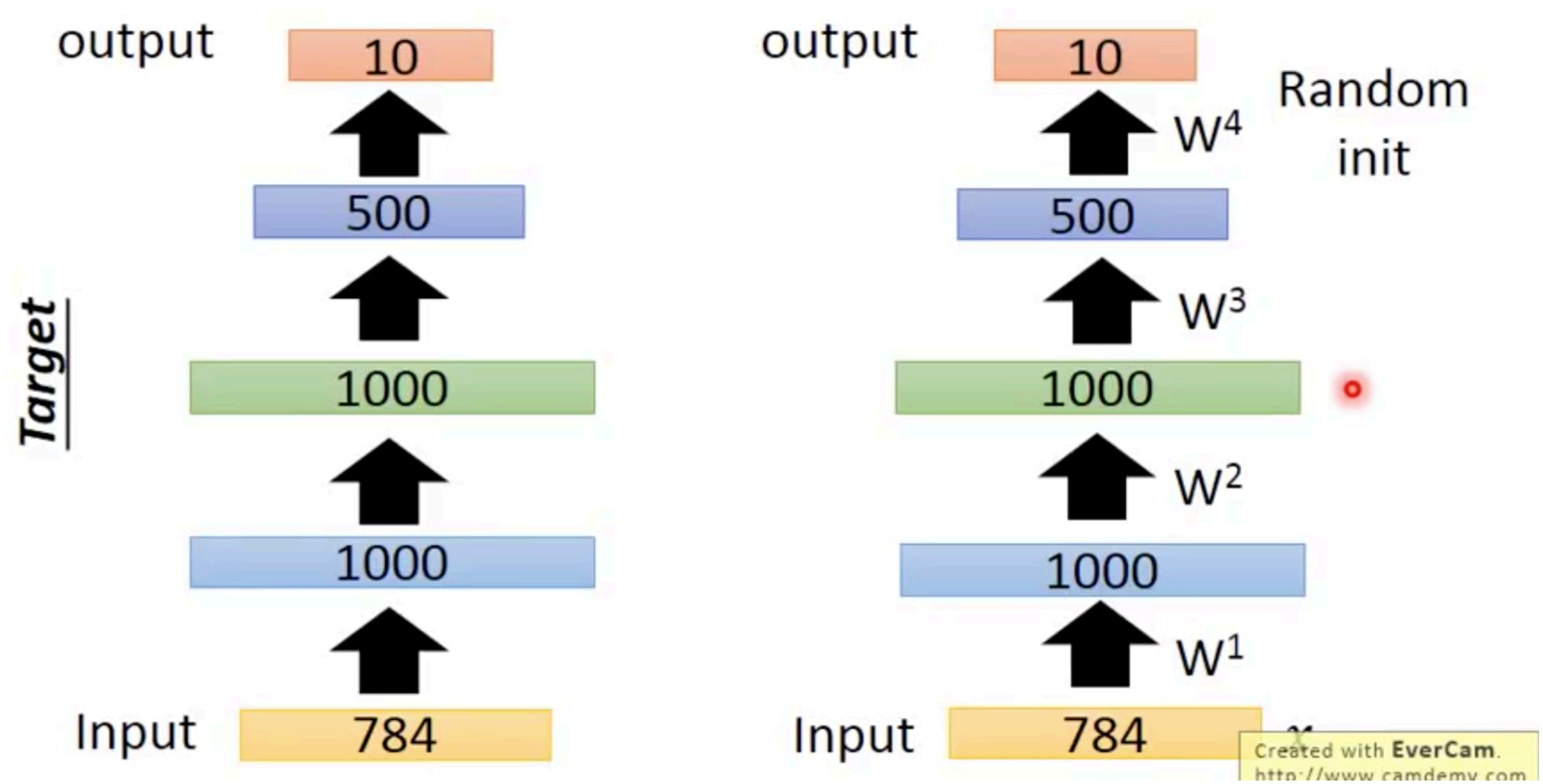


Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*

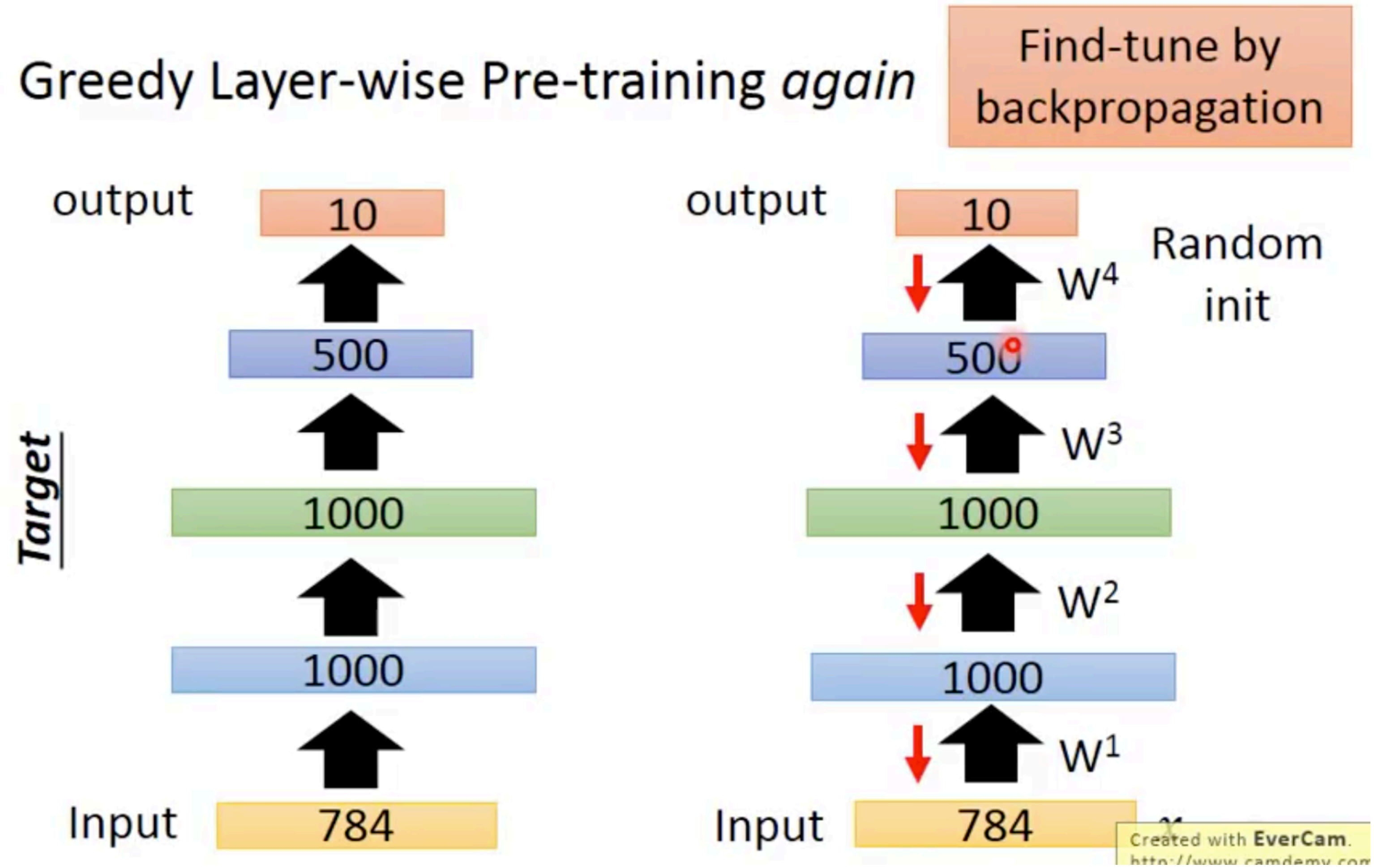
Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*



Auto-encoder – Pre-training DNN

- Greedy Layer-wise Pre-training *again*

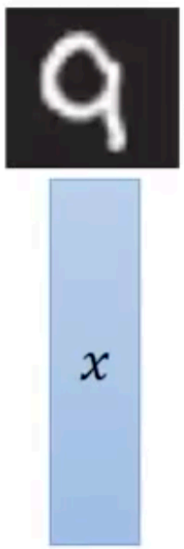


Pre-training is used much less often than before,
Because today we have back-propagation algorithms that can train very deep networks.

However, if we have a large set of un-labelled data and only a small set of labelled data, we can use the large set of un-labelled data to pre-train all the layers other than the last layer, and then use the small set of labelled data to train the last layer and fine tune weights in all layers.

Auto-encoder

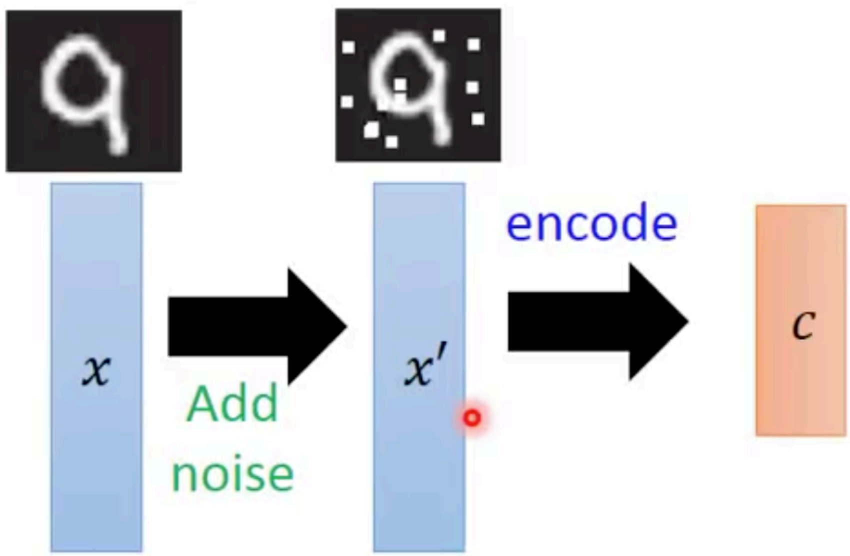
- De-noising auto-encoder



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

Auto-encoder

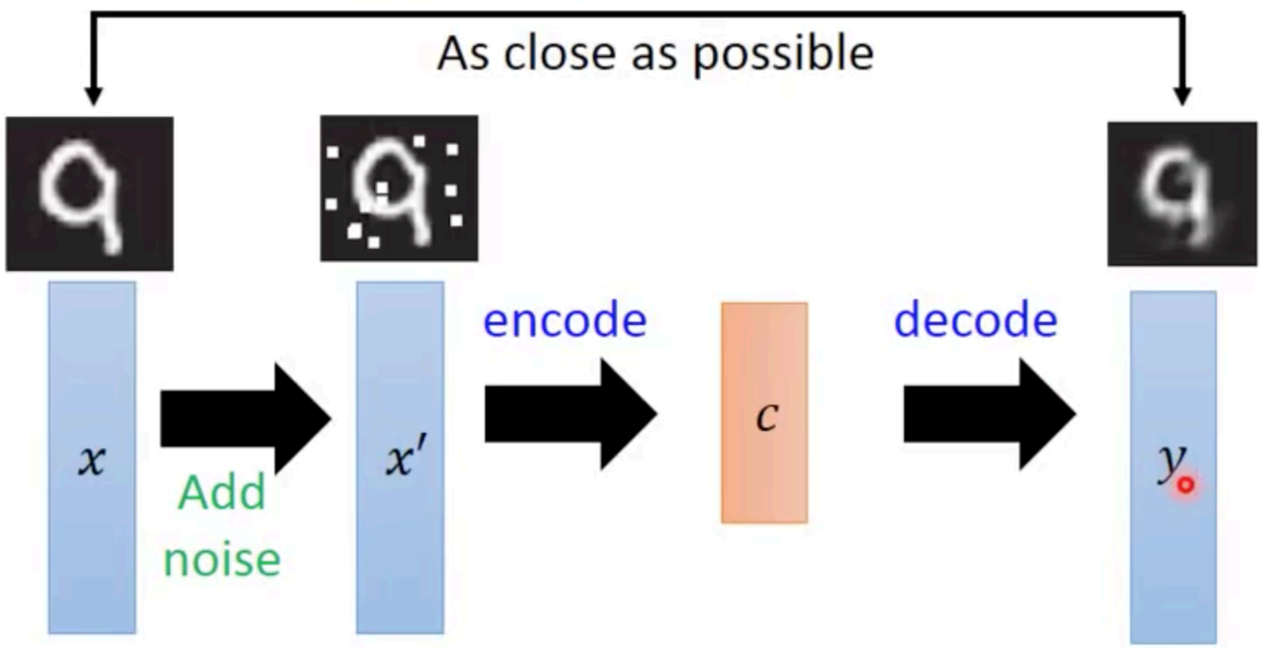
- De-noising auto-encoder



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

Auto-encoder

- De-noising auto-encoder



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

Learning More - Restricted Boltzmann Machine

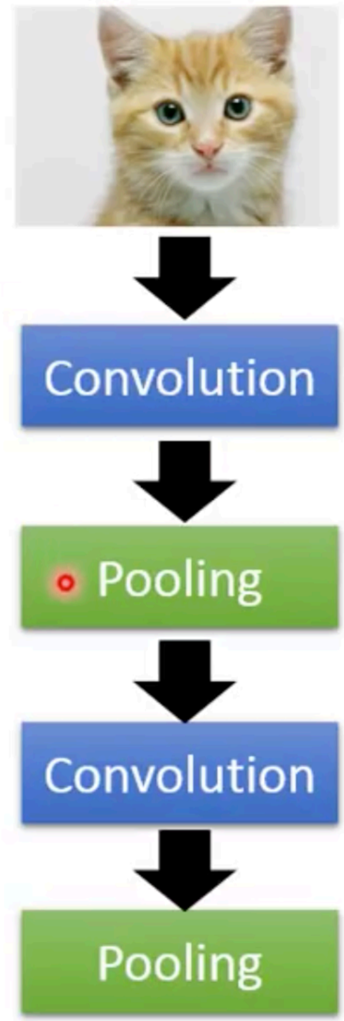
Methods for non-linear dimension-reduction

- Neural networks [5.1] : Restricted Boltzmann machine – definition
  - https://www.youtube.com/watch?v=p4Vh_zMw-HQ&index=36&list=PL6Xpj9I5qXYEcOhn7TqghAJ6NAPrNmUBH
- Neural networks [5.2] : Restricted Boltzmann machine – inference
  - https://www.youtube.com/watch?v=lekCh_i32iE&list=PL6Xpj9I5qXYEcOhn7TqghAJ6NAPrNmUBH&index=37
- Neural networks [5.3] : Restricted Boltzmann machine - free energy
  - https://www.youtube.com/watch?v=e0Ts_7Y6hZU&list=PL6Xpj9I5qXYEcOhn7TqghAJ6NAPrNmUBH&index=38

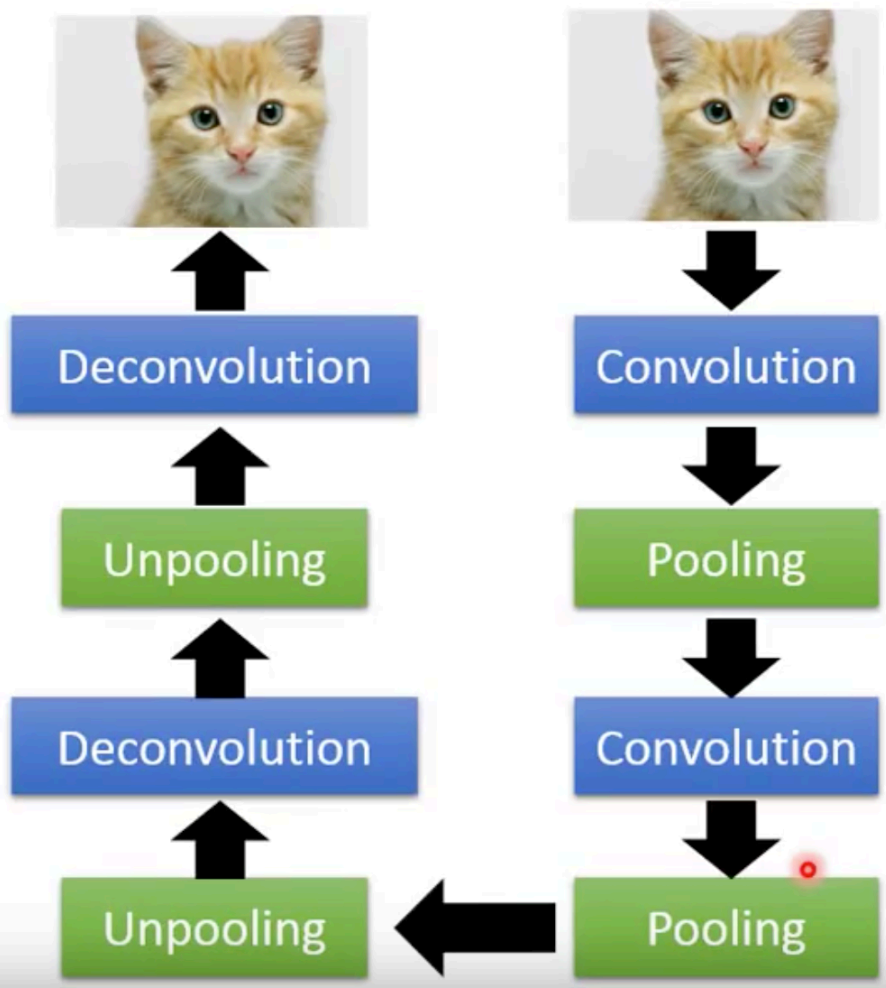
Learning More
- Deep Belief Network

- Neural networks [7.7] : Deep learning - deep belief network
  - https://www.youtube.com/watch?v=vkb6AWYXZ5I&list=PL6Xpj9I5qXYEcOhn7TqghAJ6NAPrNmUBH&index=57
- Neural networks [7.8] : Deep learning - variational bound
  - https://www.youtube.com/watch?v=pStDscJh2Wo&list=PL6Xpj9I5qXYEcOhn7TqghAJ6NAPrNmUBH&index=58
- Neural networks [7.9] : Deep learning - DBN pre-training
  - https://www.youtube.com/watch?v=35MUlYCColk&list=PL6Xpj9I5qXYEcOhn7TqghAJ6NAPrNmUBH&index=59

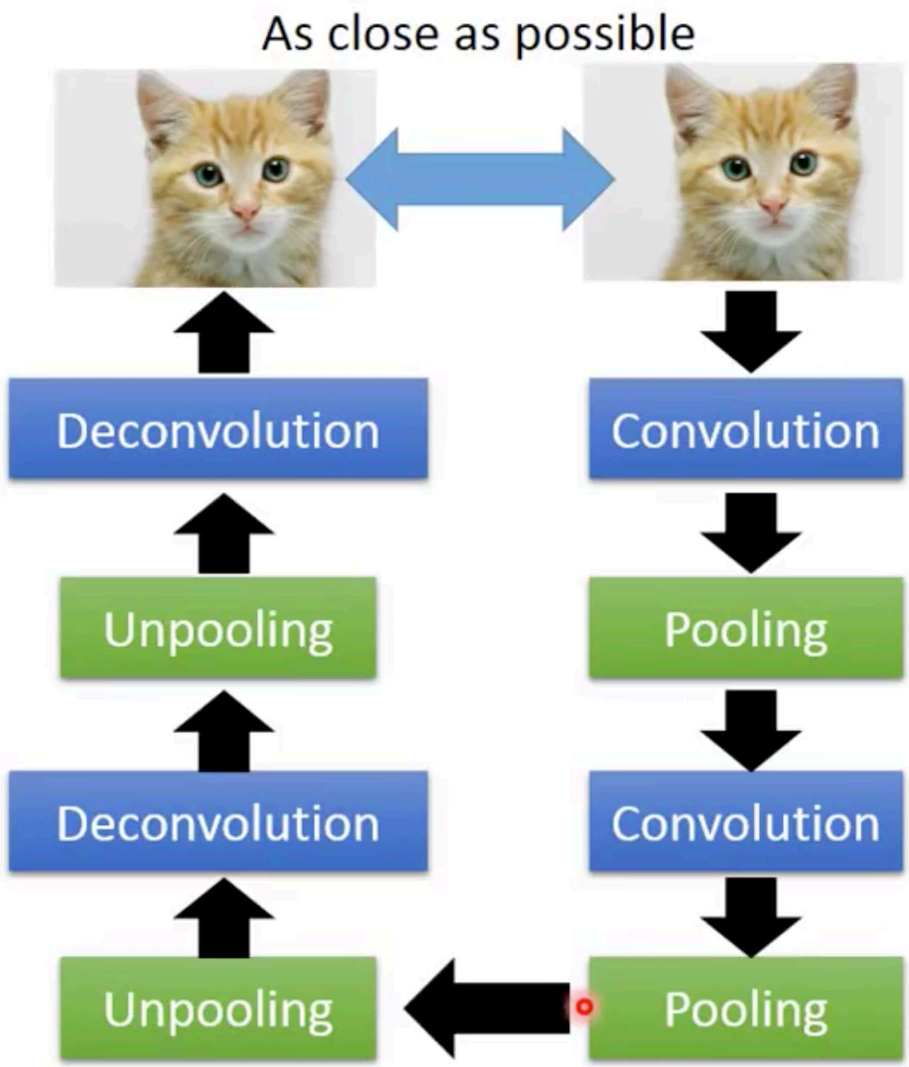
Auto-encoder for CNN



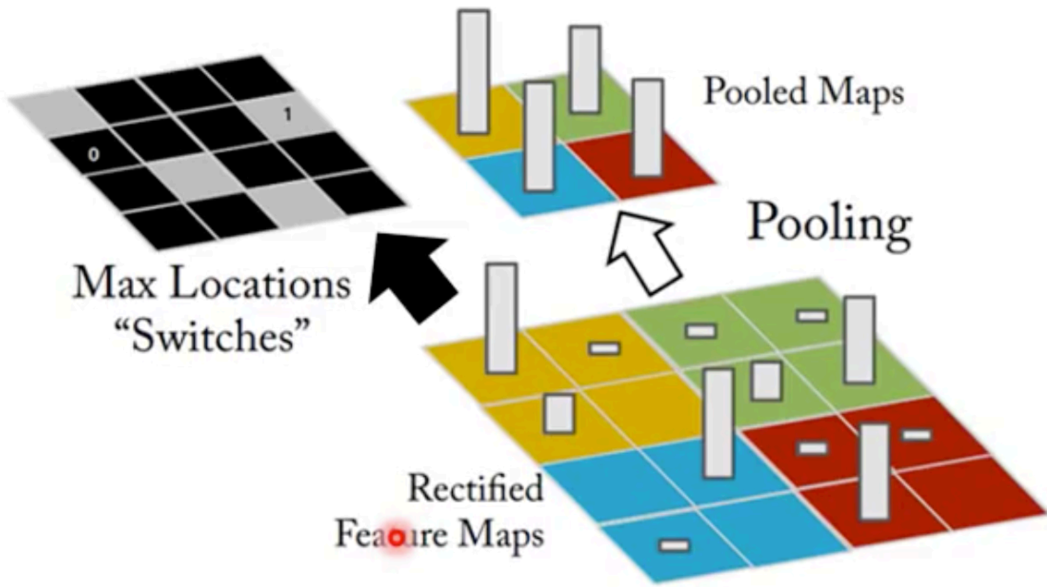
Auto-encoder for CNN



Auto-encoder for CNN

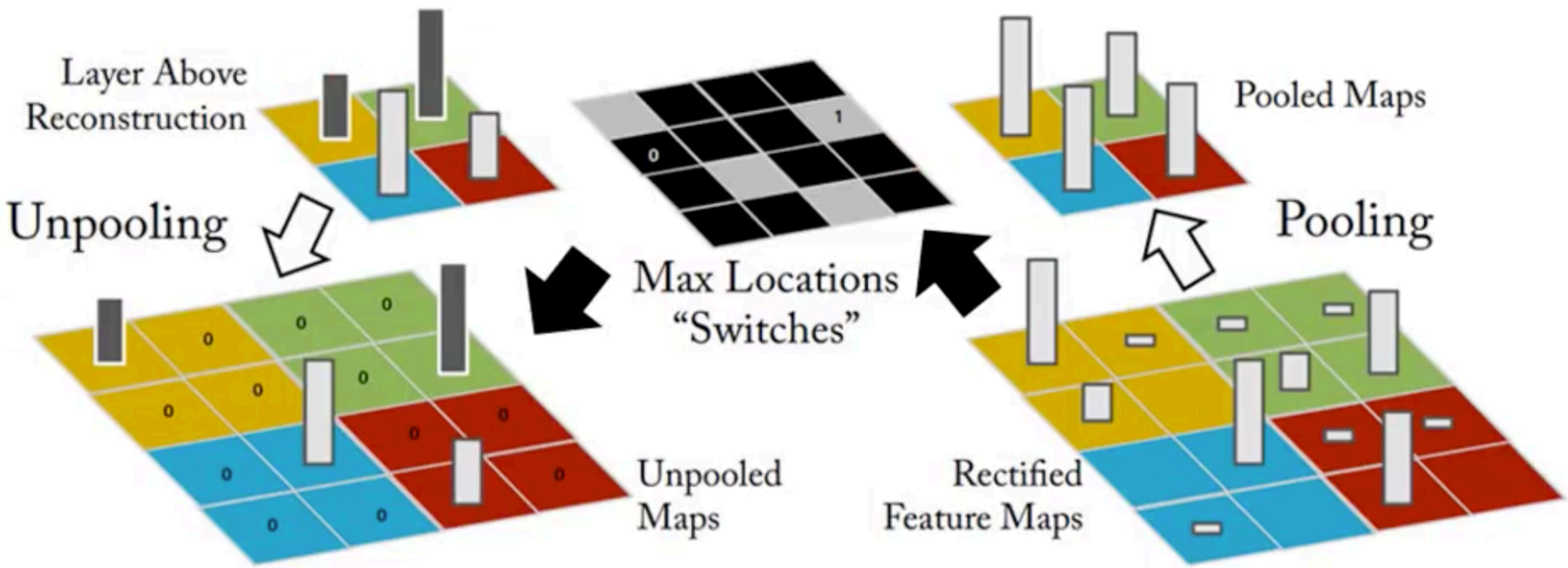


CNN -Unpooling

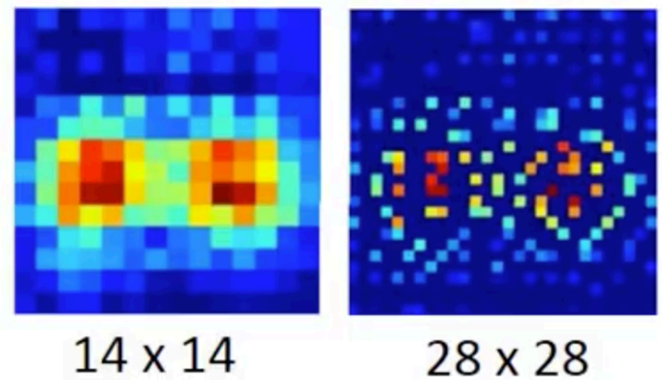


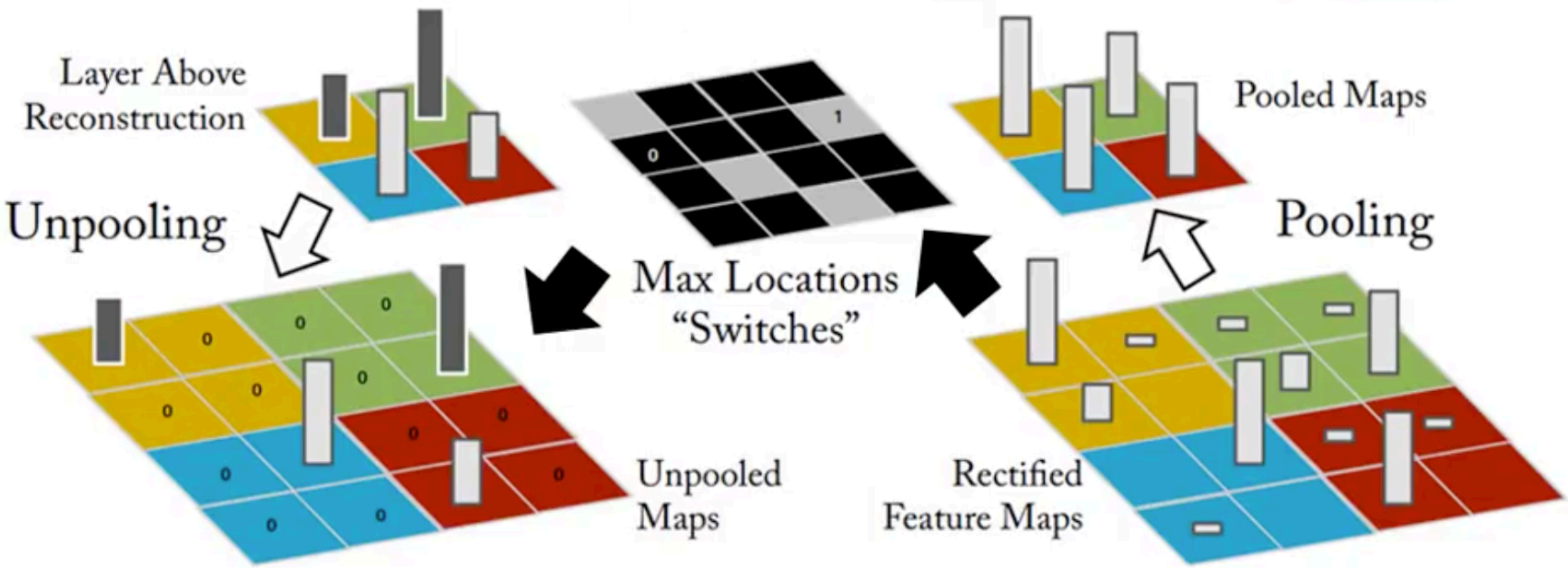
Source of image :
https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/image_segmentat

CNN -Unpooling

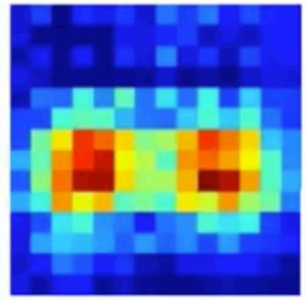


CNN -Unpooling

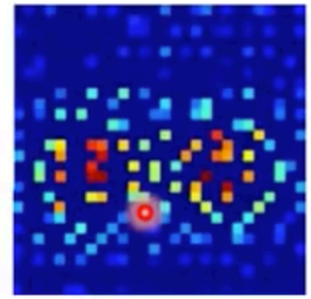




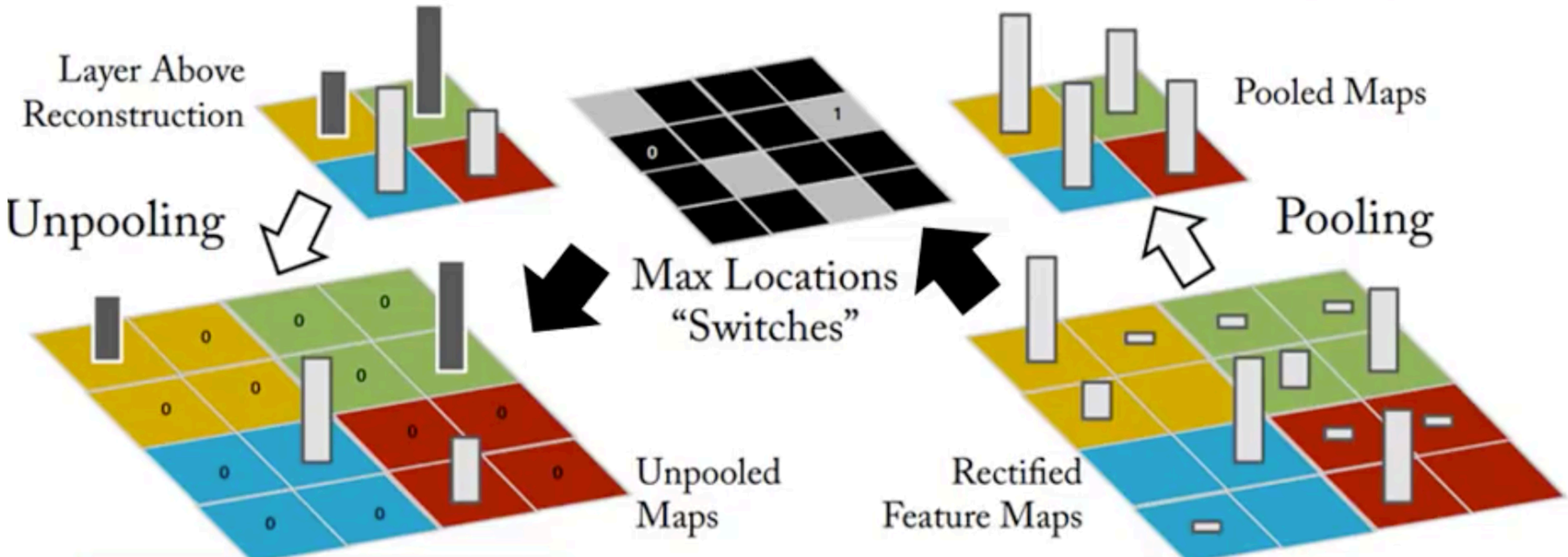
CNN -Unpooling



14 x 14



28 x 28



Keras

Alternative: simply repeat the values

No need to remember "Max locations"

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

CNN
- Deconvolution

Actually, deconvolution is convolution.

CNN
- Deconvolution

Actually, deconvolution is convolution.

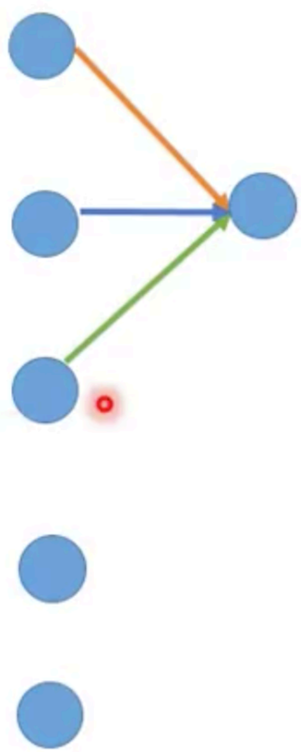
CNN
- Deconvolution

Example: 1-dimensional convolution

Actually, deconvolution is convolution.

CNN
- Deconvolution

convolution

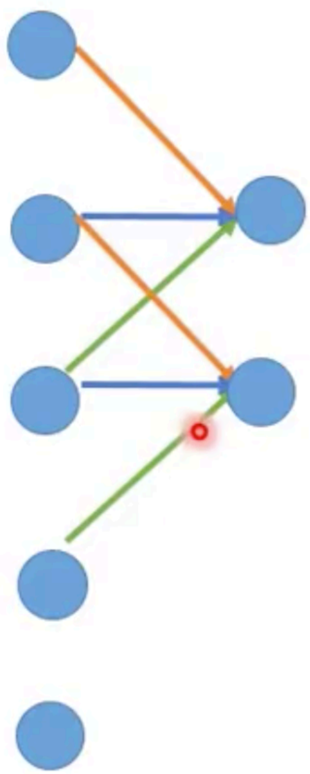


Filter size: 3

CNN
- Deconvolution

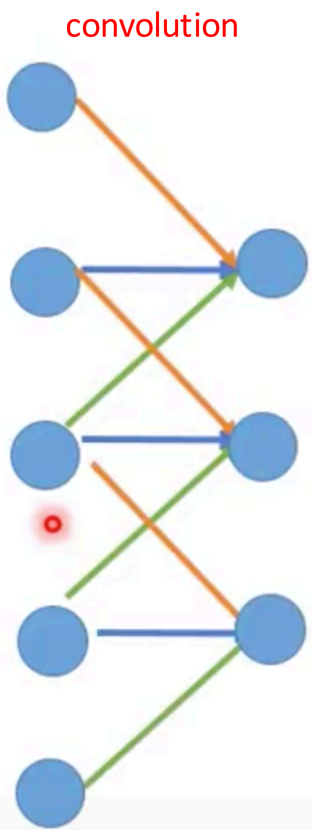
Actually, deconvolution is convolution.

convolution



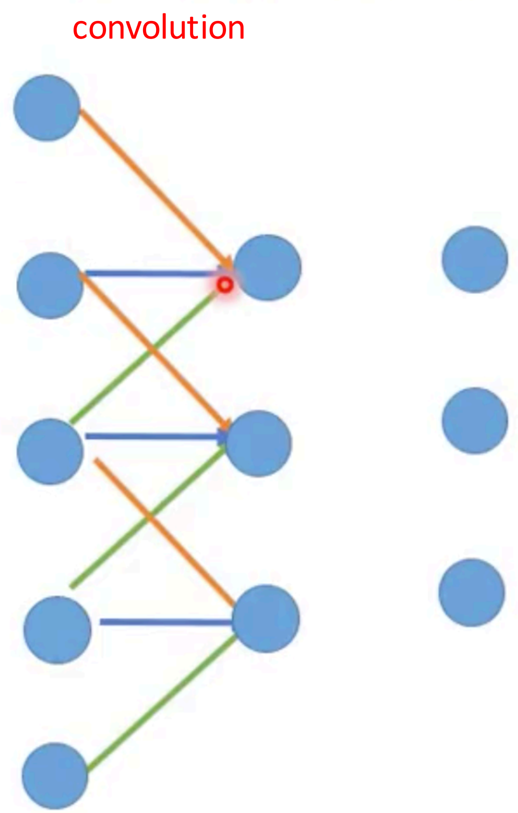
CNN
- Deconvolution

Actually, deconvolution is convolution.



CNN
- Deconvolution

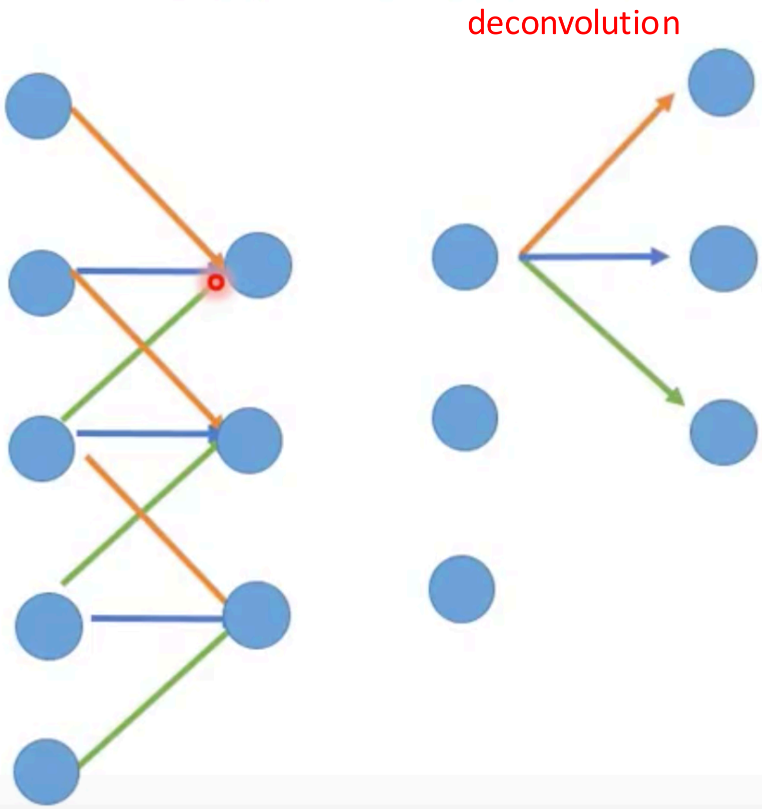
Actually, deconvolution is convolution.



Intuitively, in deconvolution,
every left node should correspond to
3 right node.

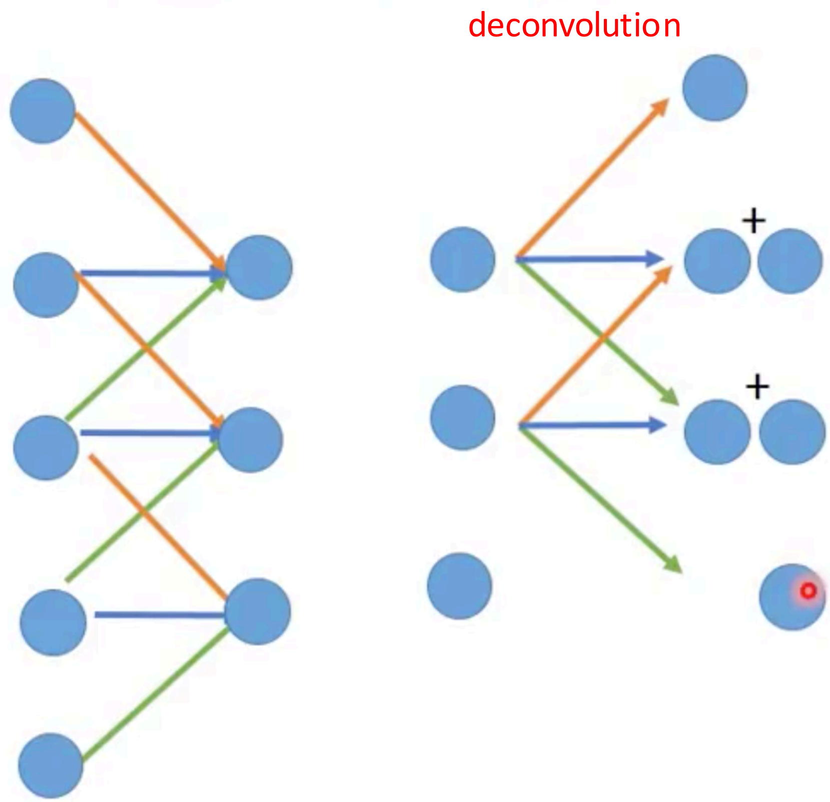
CNN
- Deconvolution

Actually, deconvolution is convolution.



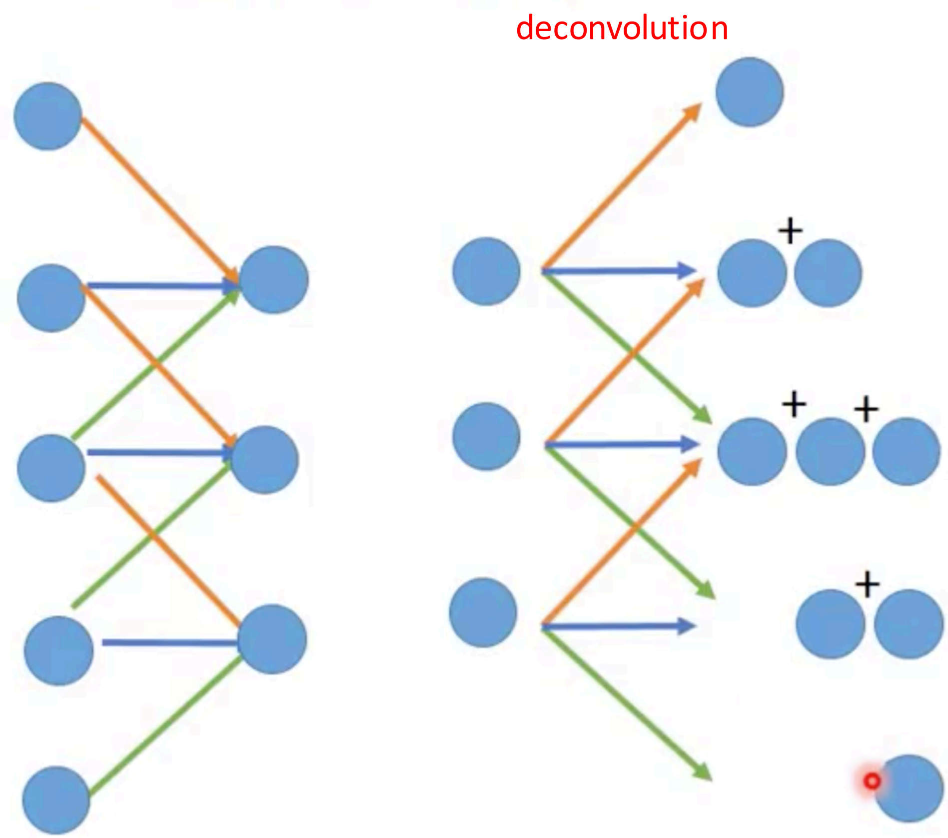
CNN
- Deconvolution

Actually, deconvolution is convolution.



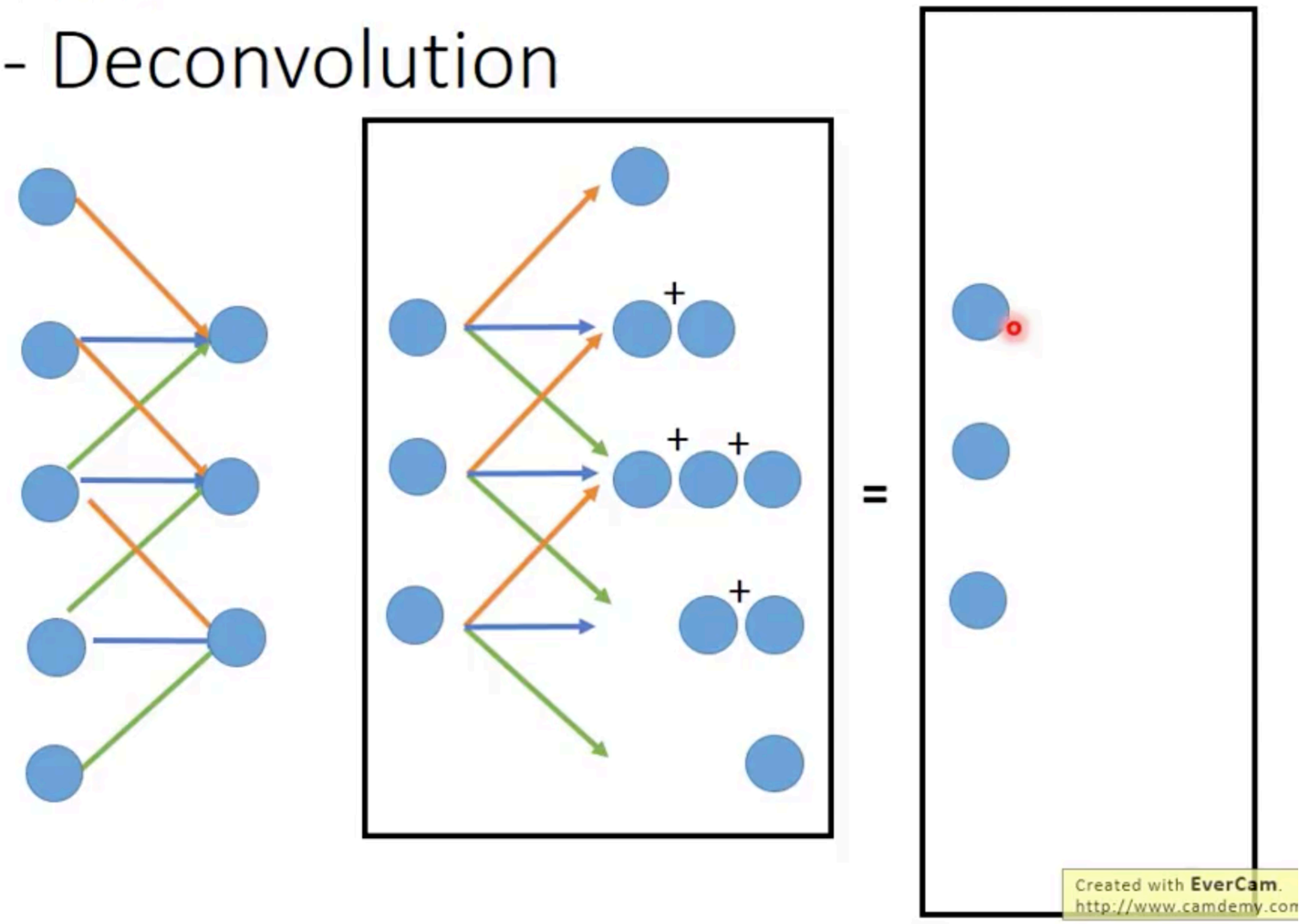
CNN
- Deconvolution

Actually, deconvolution is convolution.



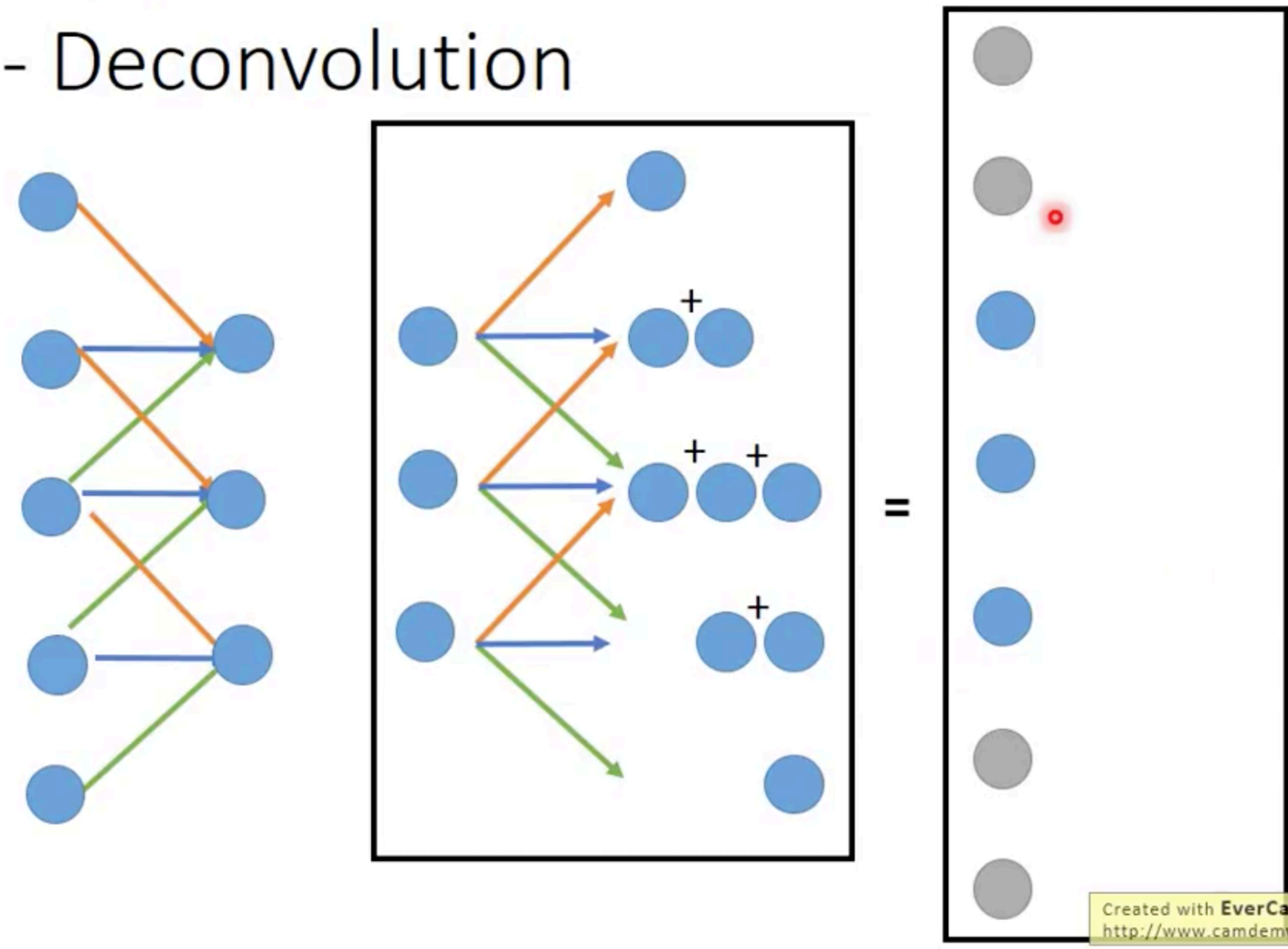
CNN
- Deconvolution

Actually, deconvolution is convolution.



CNN
- Deconvolution

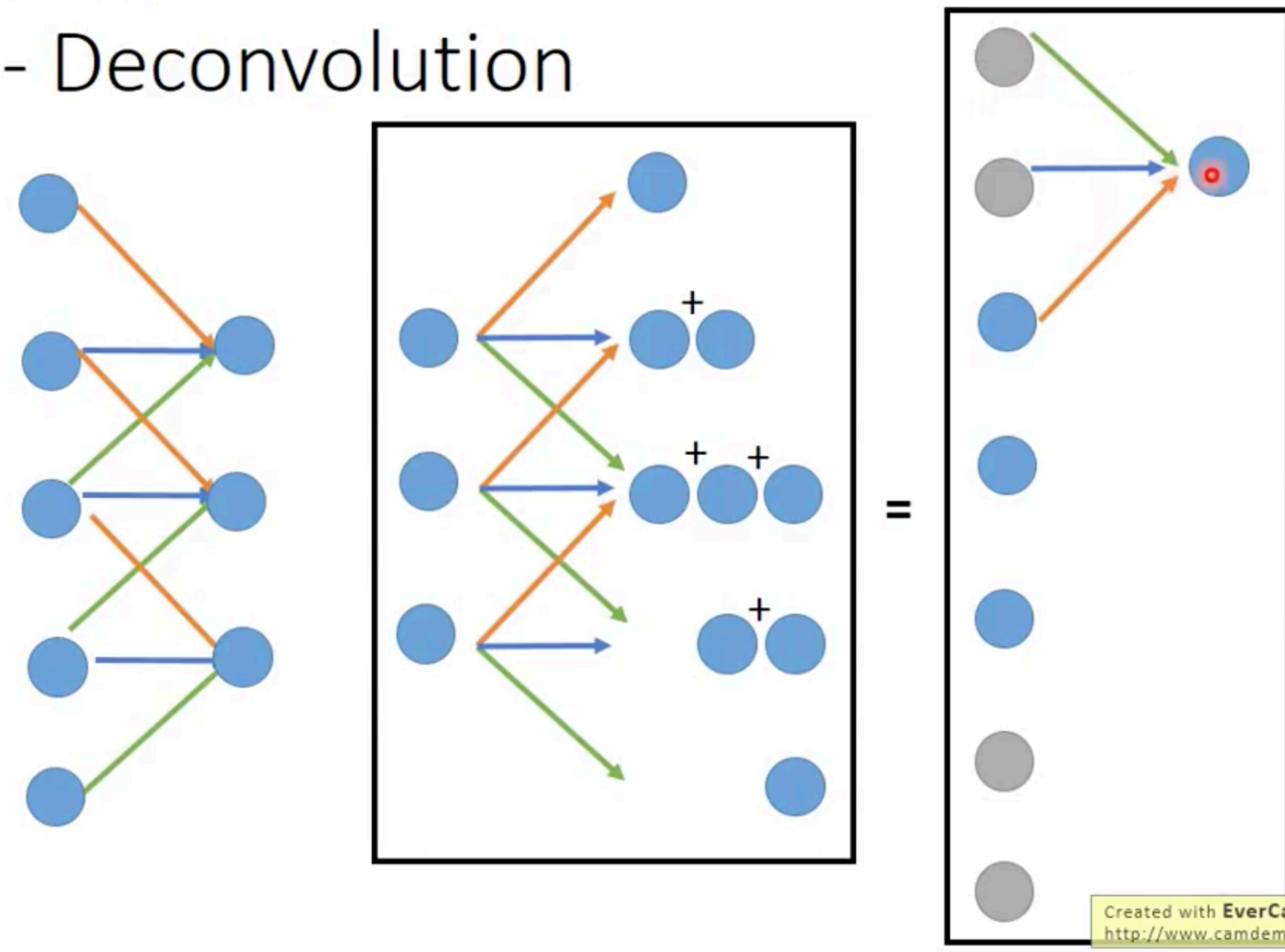
Actually, deconvolution is convolution.



Pad 0s on both ends

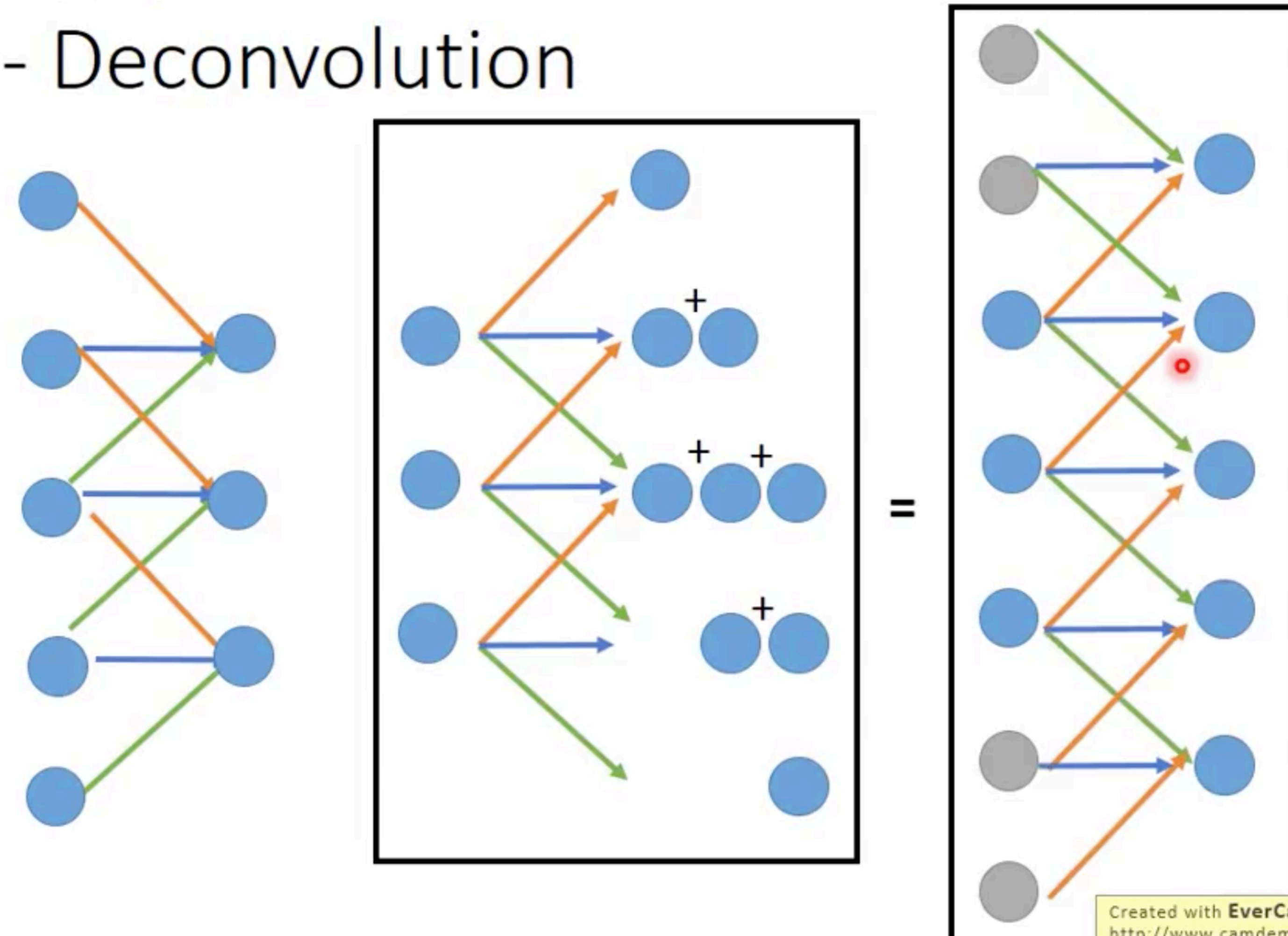
CNN
- Deconvolution

Actually, deconvolution is convolution.

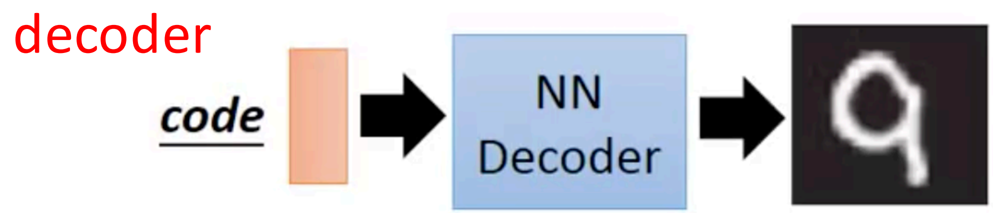


CNN
- Deconvolution

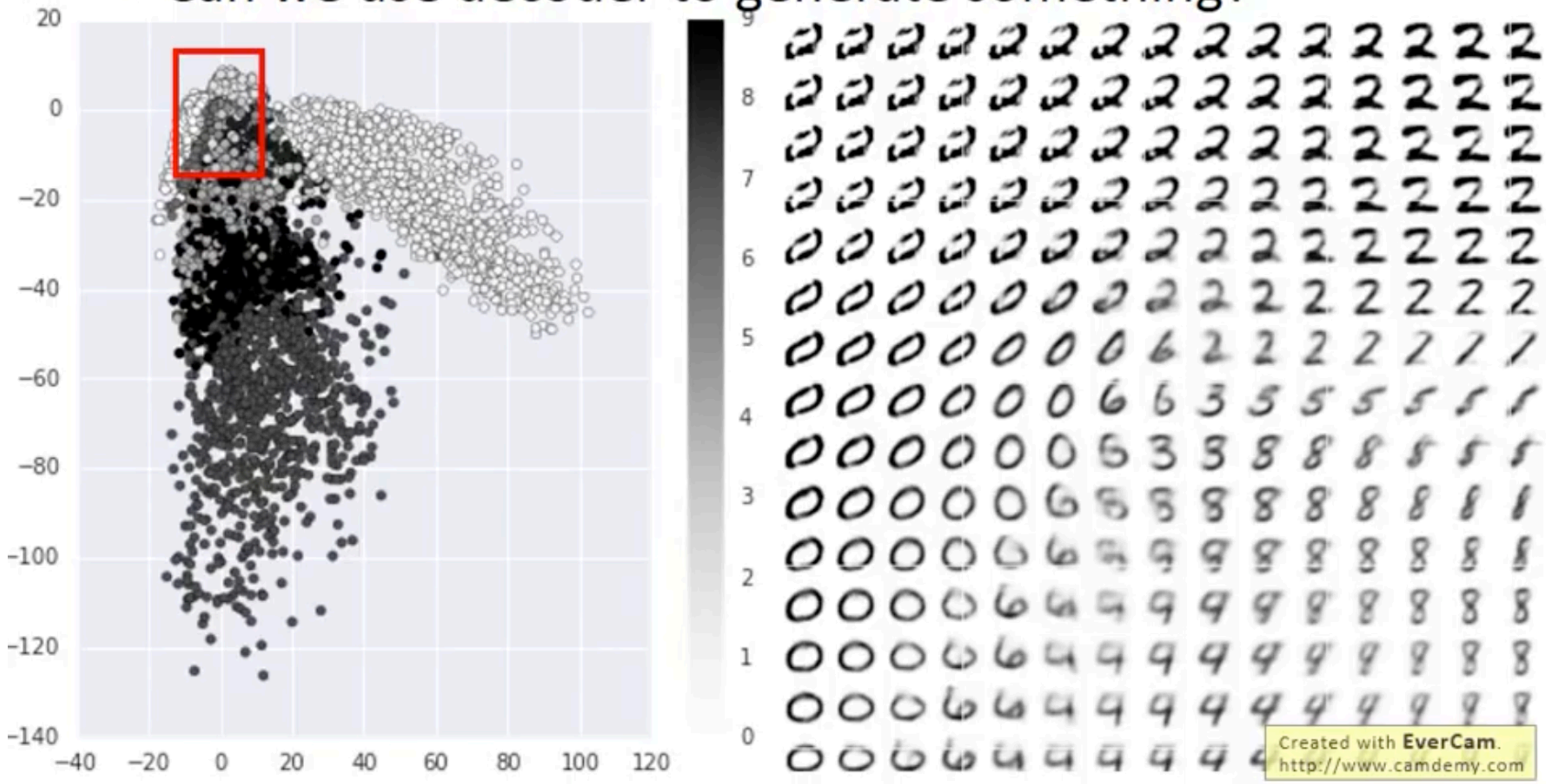
Actually, deconvolution is convolution.



Next .....

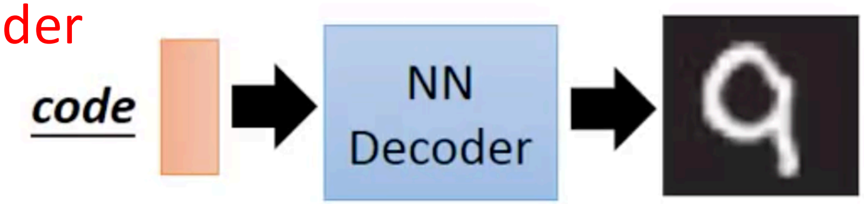


- Can we use decoder to generate something?



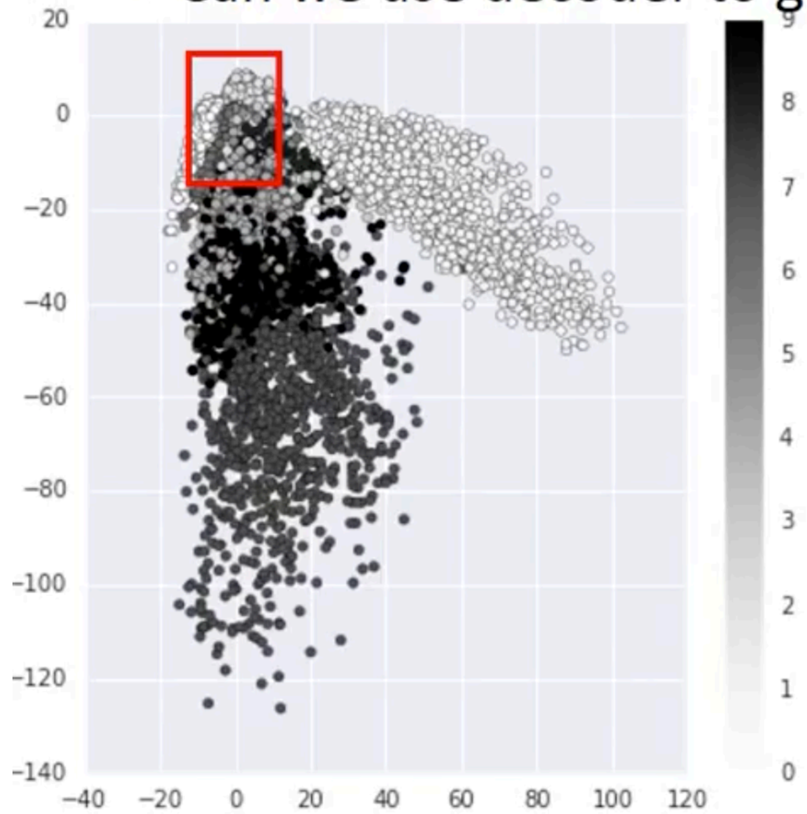
Next .....

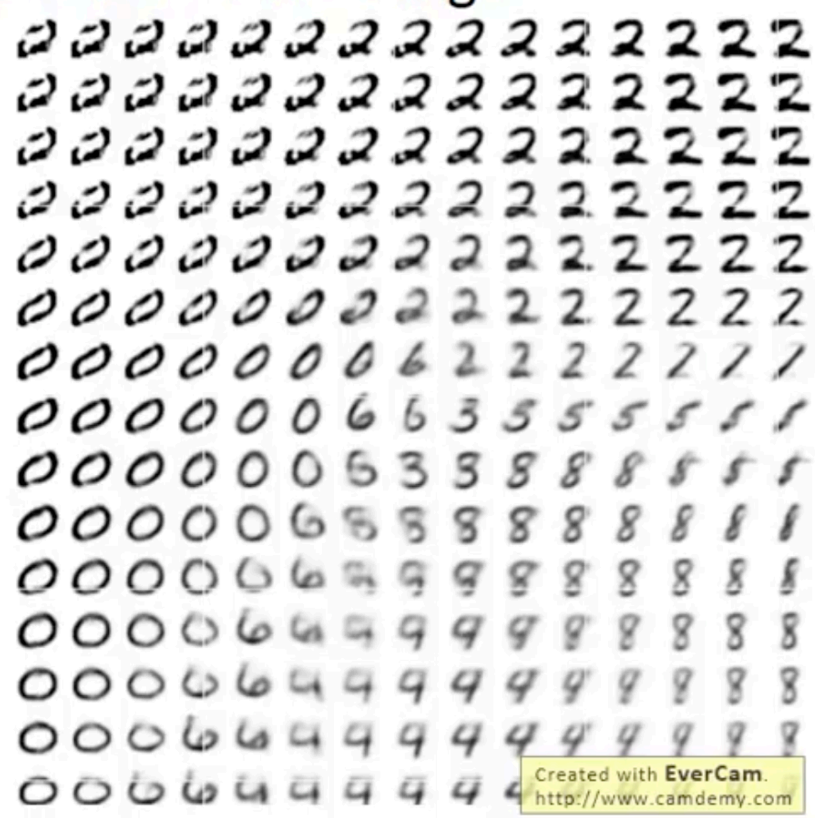
decoder



MNIST Dataset: Encoder compresses the 28 x 28 = 784-dimensional input image to 2-dimensional data.

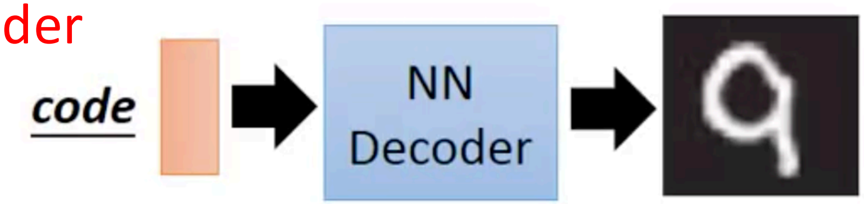
- Can we use decoder to generate something?





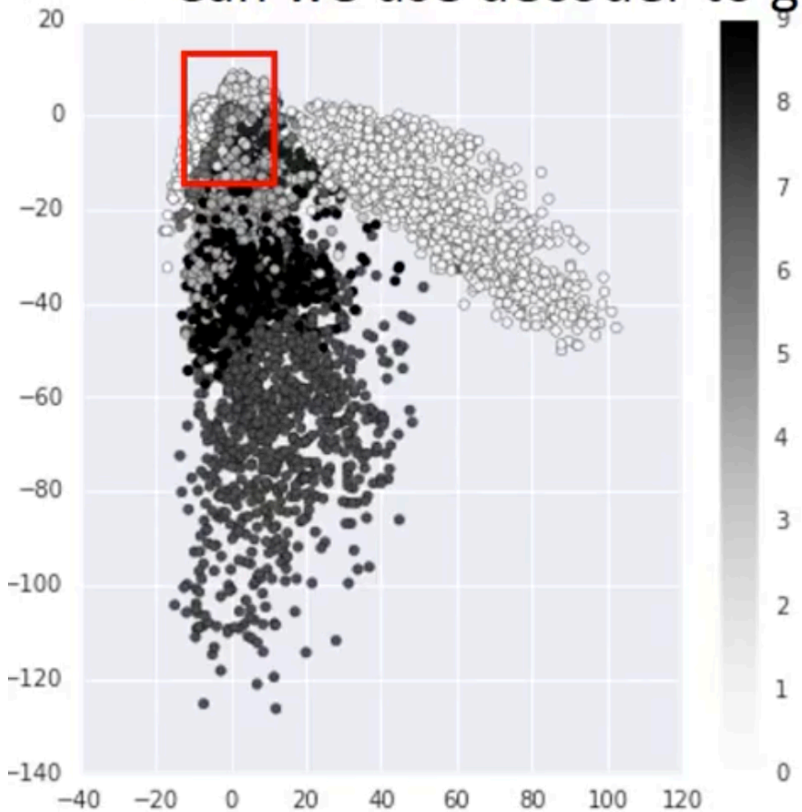
Next .....

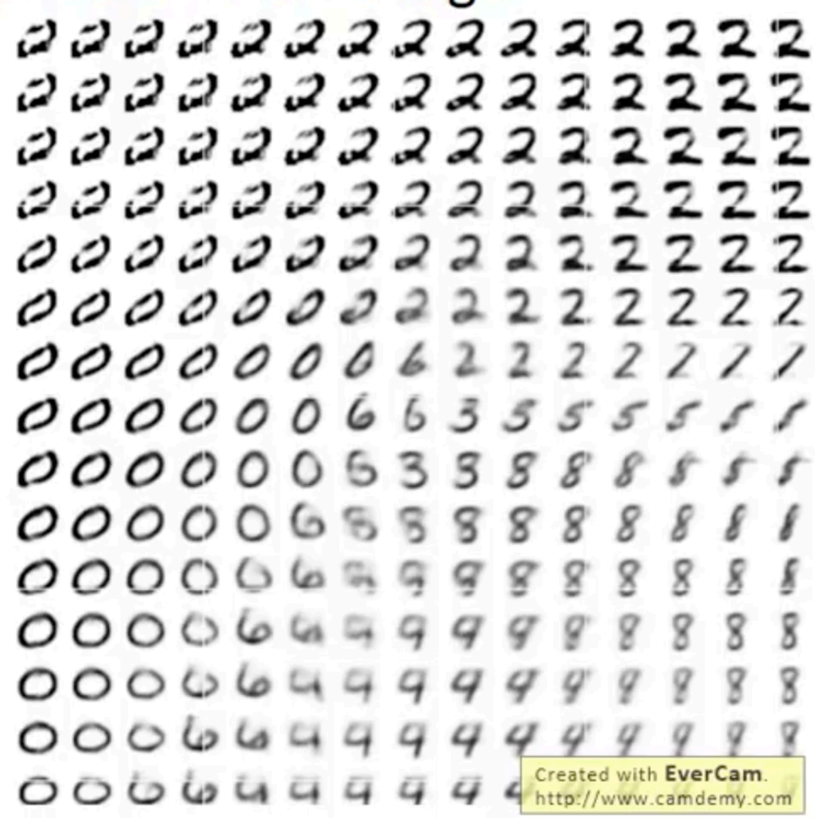
decoder



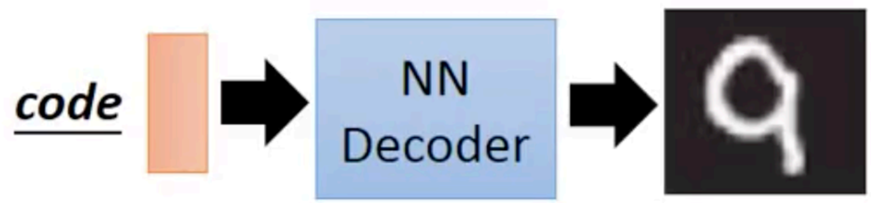
How to know which region to sample, if the code has more than 2 dimensions? We can use L2 regularization during training,

- Can we use decoder to generate something? so that the data are all around 0.





Next .....



- Can we use decoder to generate something?

Auto-encoder
trained with
L2 regularization