NPTEL ONLINE CERTIFICATION COURSES

Course Name: Deep Learning

Faculty Name: Prof. P. K. Biswas
Department : E & ECE, IIT Kharagpur

Topic

Lecture 30: Autoencoder vs. PCA

CONCEPTS COVERED

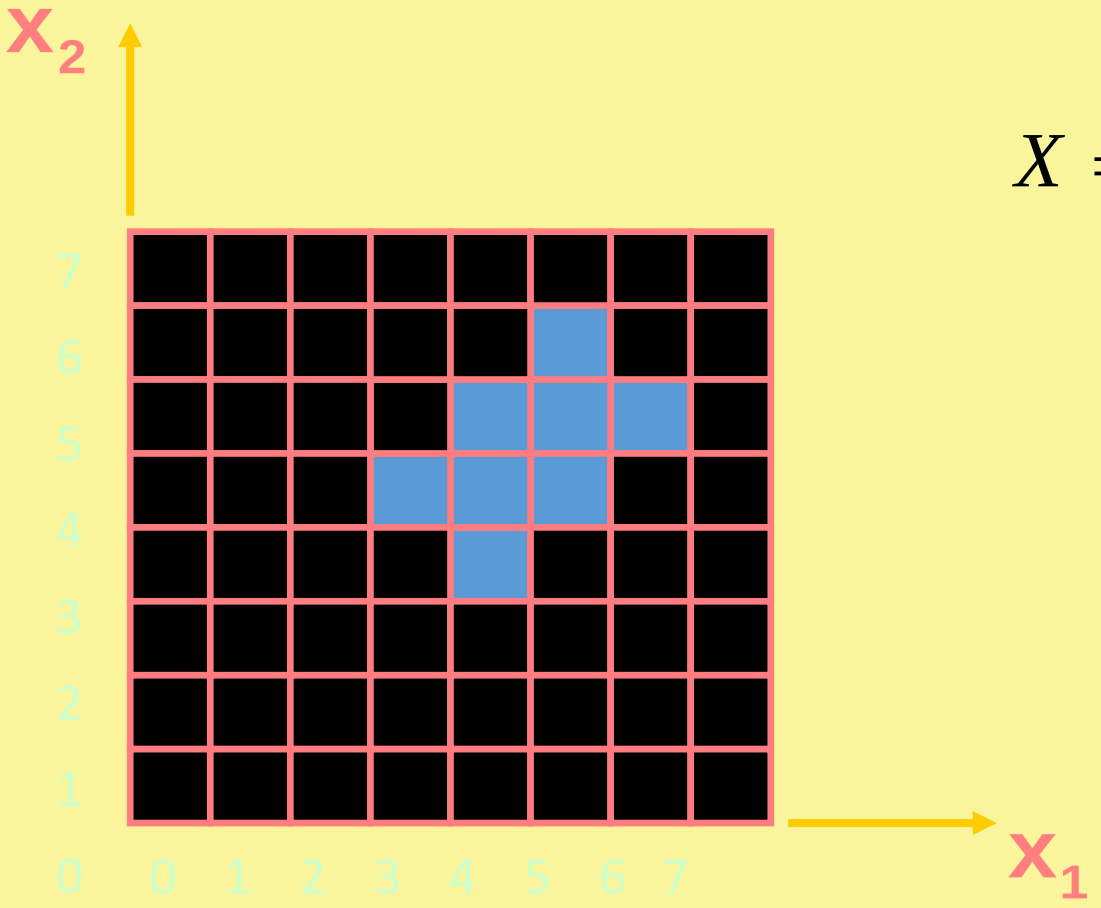
Concepts Covered:

- Autoencoder
  - Undercomplete Autoencoder
  - Autoencoder vs. PCA
  - Deep Autoencoder Training
  - Sparse Autoencoder
  - Denoising Autoencoder
  - Contractive Autoencoder
  - Convolution Autoencoder

Autoencoder
vs.
PCA

What is PCA?

PCA



$$X = \left\{ \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 6 \\ 5 \end{bmatrix} \right\}$$

$$\mu_X = \begin{bmatrix} 4.5 \\ 4.5 \end{bmatrix}$$

PCA

$$X = \left\{ \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 6 \\ 5 \end{bmatrix} \right\} \qquad \mu_X = \begin{bmatrix} 4.5 \\ 4.5 \end{bmatrix}$$

$$(X_1 - \mu_X)(X_1 - \mu_X)^t = \begin{bmatrix} -1.5 \\ -0.5 \end{bmatrix} \begin{bmatrix} -1.5 & -0.5 \end{bmatrix} = \begin{bmatrix} 2.25 & 0.75 \\ 0.75 & 0.25 \end{bmatrix}$$

$$(X_2 - \mu_X)(X_2 - \mu_X)^t = \begin{bmatrix} 0.25 & 0.75 \\ 0.75 & 2.25 \end{bmatrix}$$

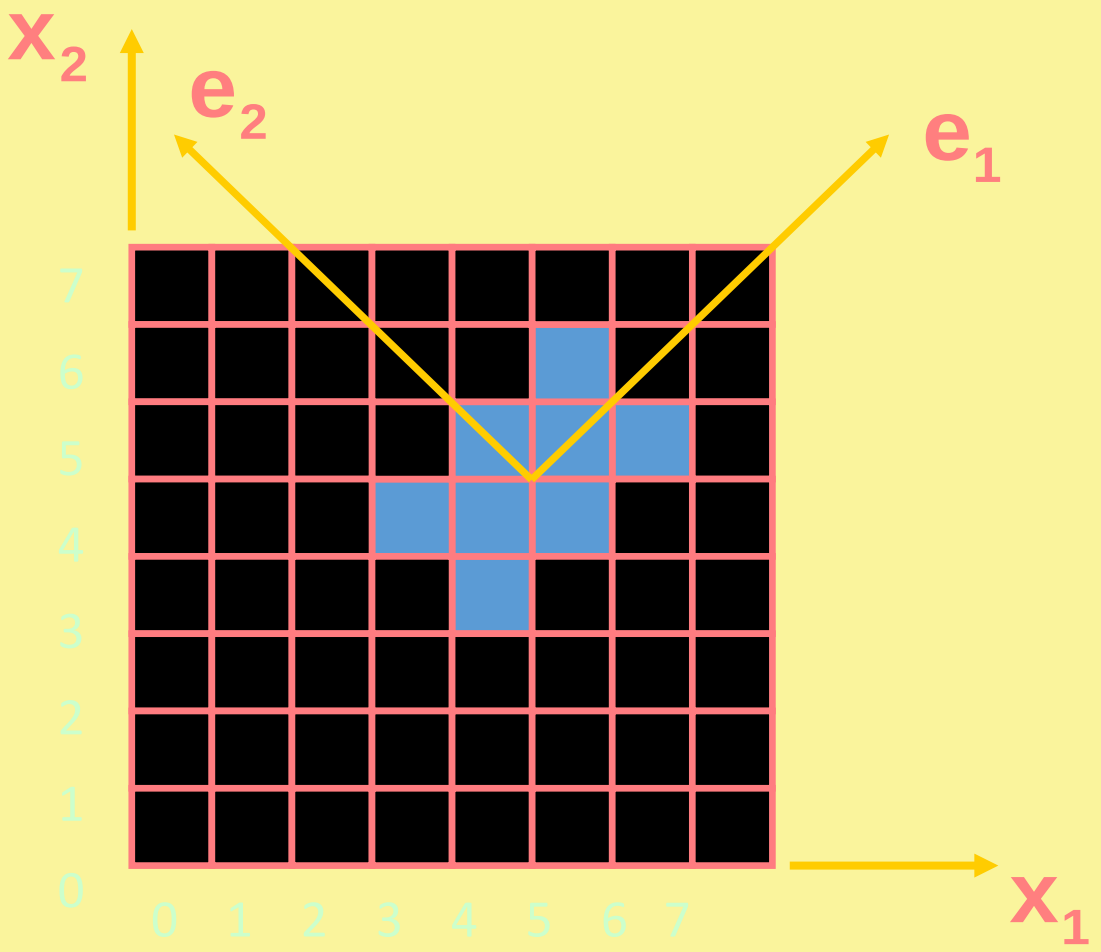
PCA

$$C_X = \begin{bmatrix} 0.75 & 0.375 \\ 0.375 & 0.75 \end{bmatrix}$$

$$\begin{vmatrix} 0.75-\lambda & 0.375 \\ 0.375 & 0.75-\lambda \end{vmatrix} = 0 \Rightarrow \lambda_1 = 1.125 \;\&\; \lambda_2 = 0.375$$

$$\lambda_1 \Rightarrow e_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \lambda_2 \Rightarrow e_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

PCA



PCA

ORIGINAL

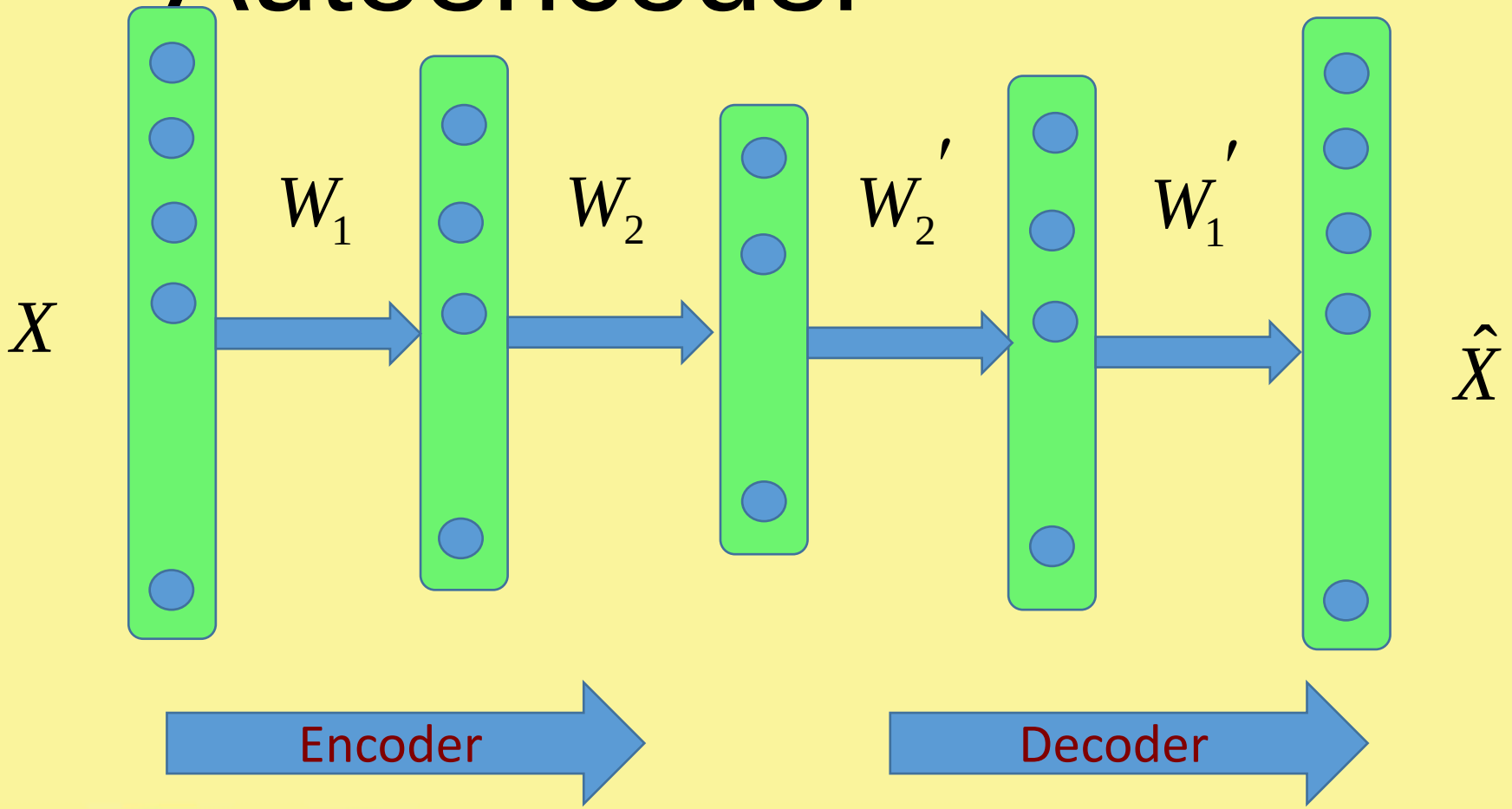
1

5

25

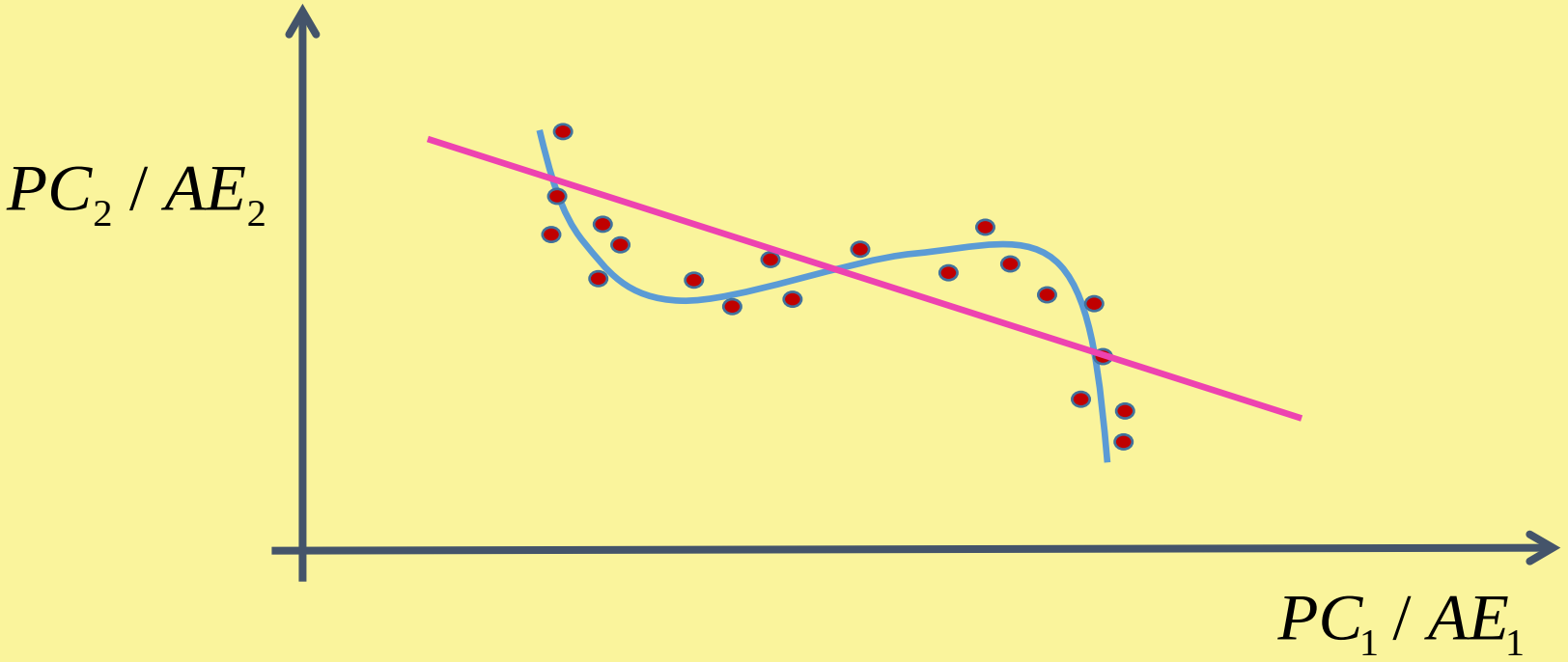
What does PCA do?

Deep Autoencoder



$$L(X, \hat{X}) = \frac{1}{2} \sum_{N} \left\| X - \hat{X} \right\|^2$$

Autoencoder vs. PCA



Experimental Setup for Dimensionality Reduction

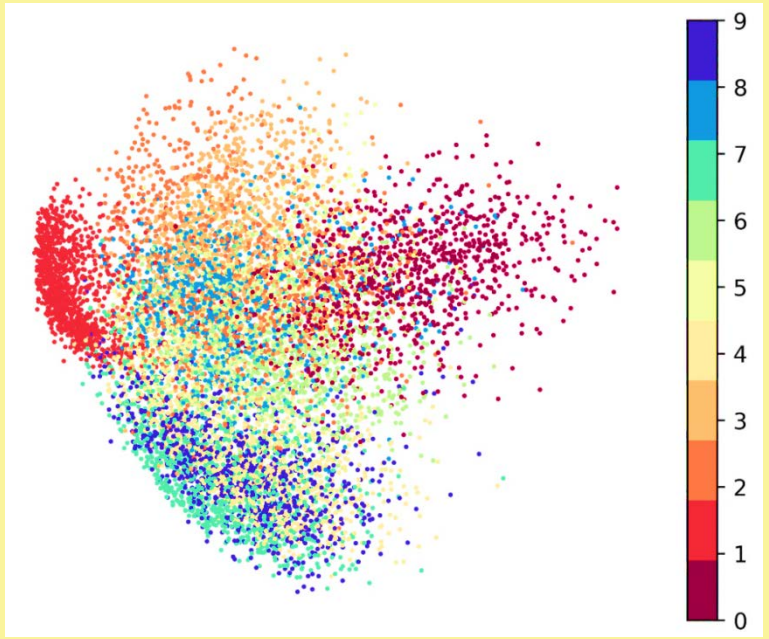
- Dataset used: MNIST (a large database of handwritten digits)
  - Total train Images: $60{,}000$
  - Total test Images: $10{,}000$
  - Image dimension: $28 \times 28\ (784)$
  - Dimensionality reduction: $784 \rightarrow 2$
  - Reconstruction: $784 \rightarrow 30$
- Optimizer used: Adam (Learning rate-10^{-4})
- Loss Function: Mean Squared Error
- Trained for 100 iterations

Source: G. E. Hilton and R. R. Salakhutdinov: “Reducing the Dimensionality of Data with Neural Networks”, Science, Vol 313, July 2006, pp. 504-507.

MNIST Data set: Example

Autoencoder converges to PCA

PCA



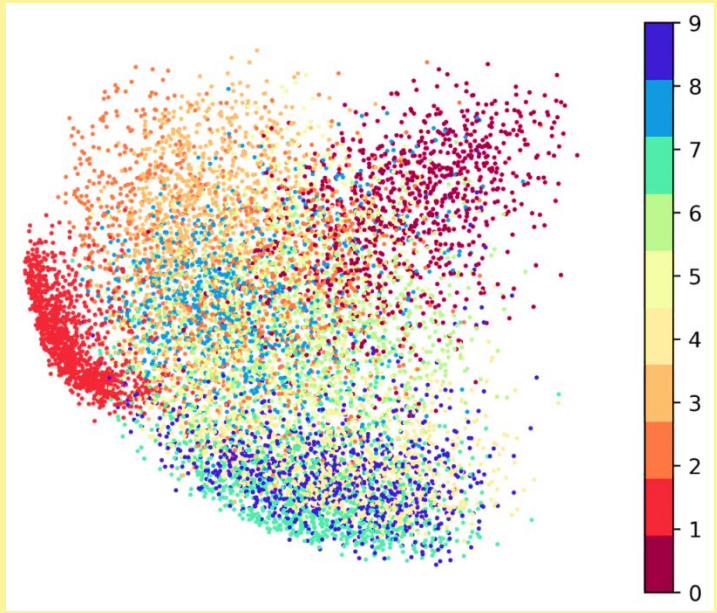
2-Layer Autoencoder

784
↓
2
↓
784

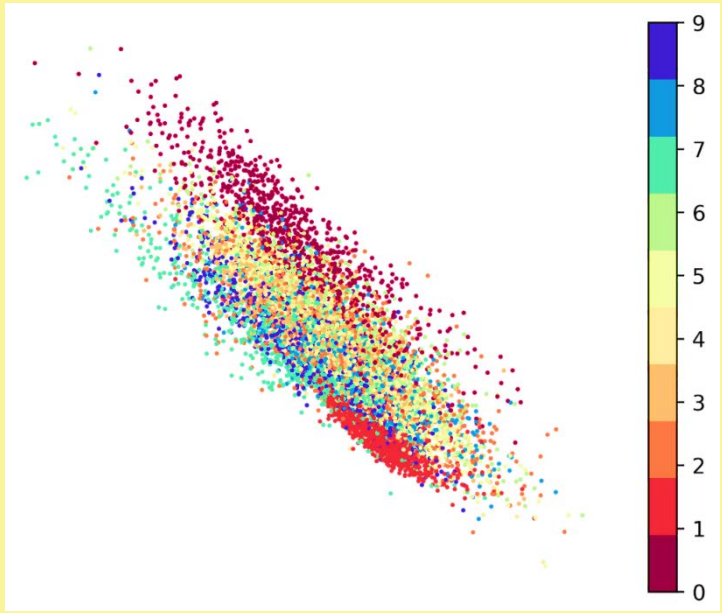
Source: G. E. Hilton and R. R. Salakhutdinov: “Reducing the Dimensionality of Data with Neural Networks”, Science, Vol 313, July 2006, pp. 504-507.

Deep vs. Shallow autoencoder

784
↓
2 (Sigmoid)
↓
784



2-Layered AE with
Sigmoid Non-Linearity



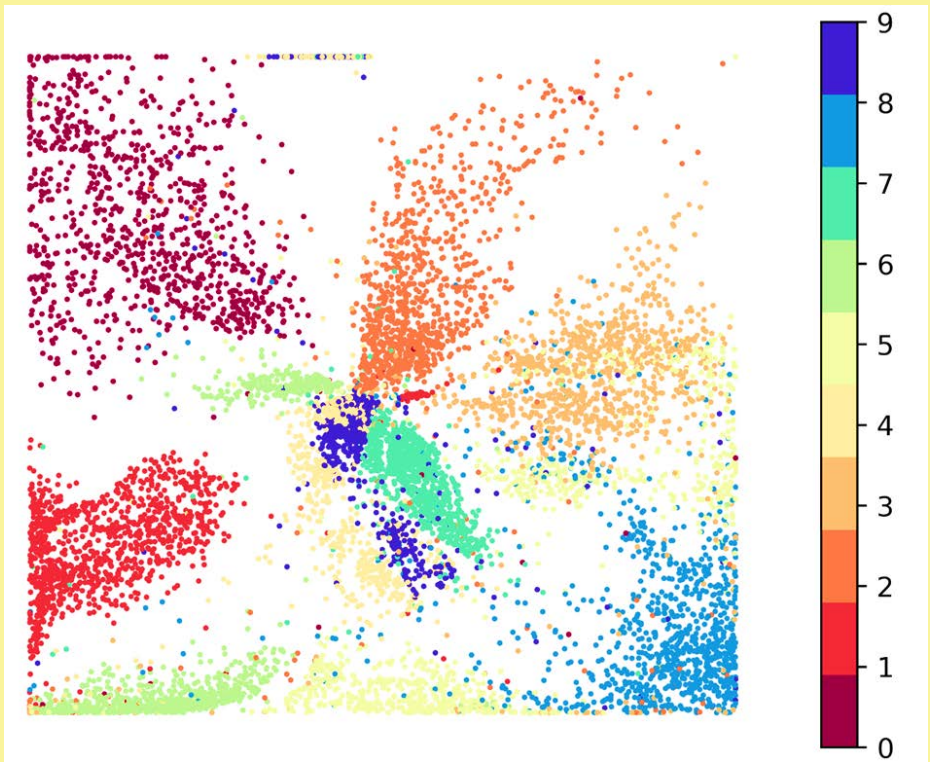
Deep AE without any
Non-Linear Activations

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

Source: G. E. Hilton and R. R. Salakhutdinov: “Reducing the Dimensionality of Data with Neural Networks”, Science, Vol 313, July 2006, pp. 504-507.

Deep Autoencoder with Non-Linear Activations

Principal component analysis (PCA)



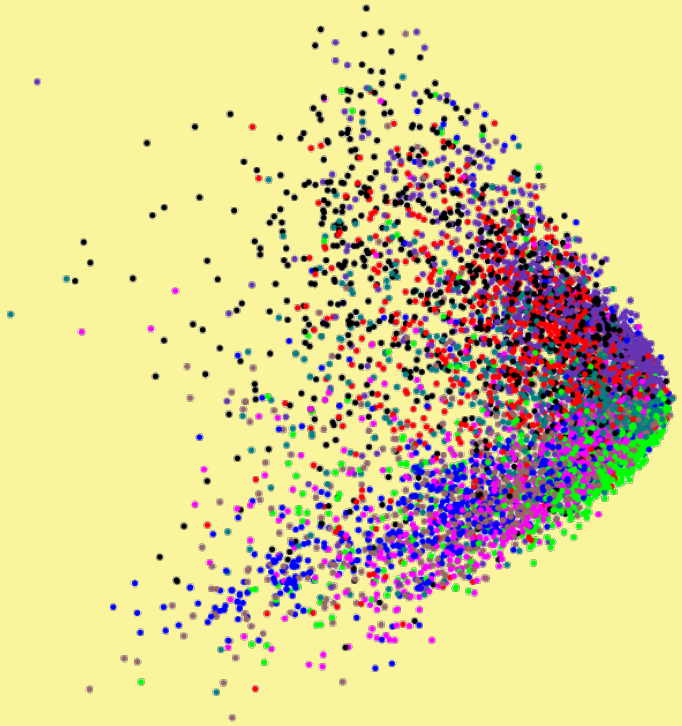
Deep Autoencoder with Non-Linear Activations

$$
\begin{aligned}
&784\ (ReLU)\\
&\downarrow\\
&1000\ (ReLU)\\
&\downarrow\\
&500\ (ReLU)\\
&\downarrow\\
&250\ (ReLU)\\
&\downarrow\\
&\mathbf{2}\ (\boldsymbol{Sigmoid})\\
&\downarrow\\
&250\ (ReLU)\\
&\downarrow\\
&500\ (ReLU)\\
&\downarrow\\
&1000\ (ReLU)\\
&\downarrow\\
&784
\end{aligned}
$$

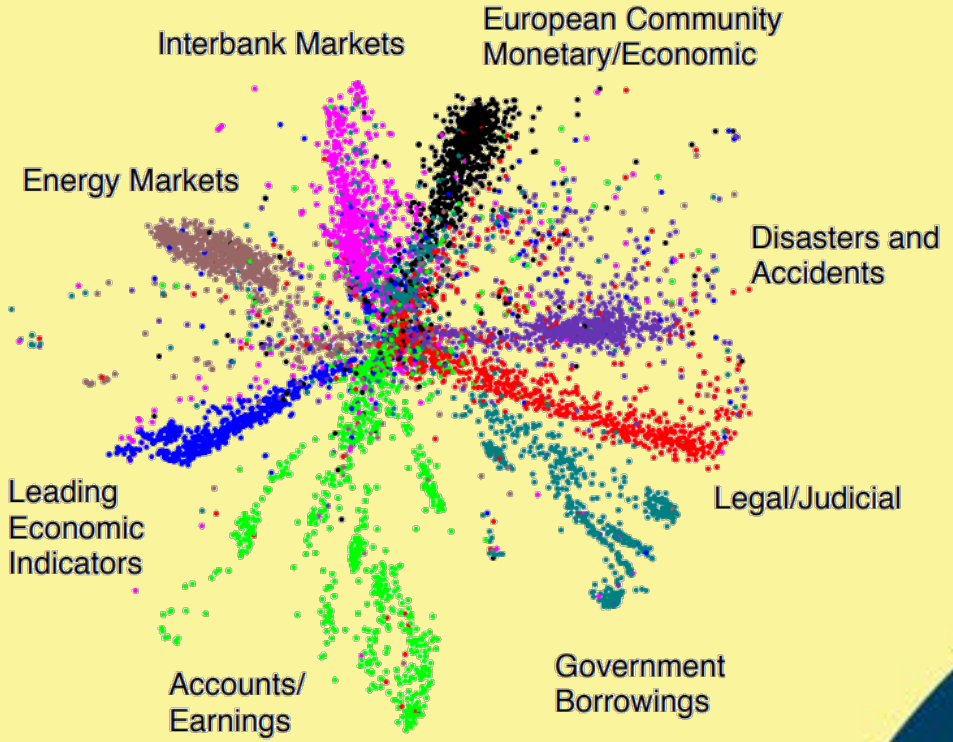
Source: G. E. Hilton and R. R. Salakhutdinov: “Reducing the Dimensionality of Data with Neural Networks”, Science, Vol 313, July 2006, pp. 504-507.

Autoencoder for Dimensionality Reduction

Articles from Reuter corpus were mapped to a 2 dimensional vector



PCA

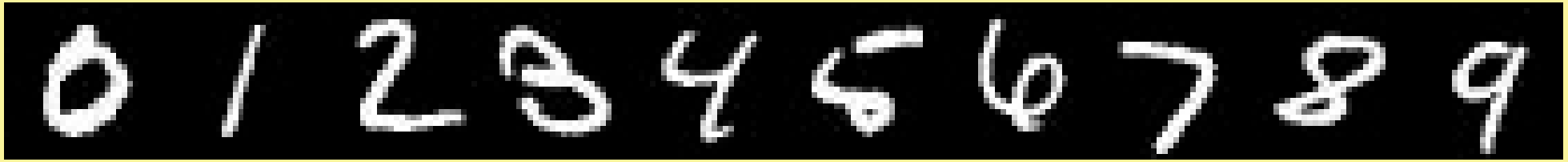


Autoencoder

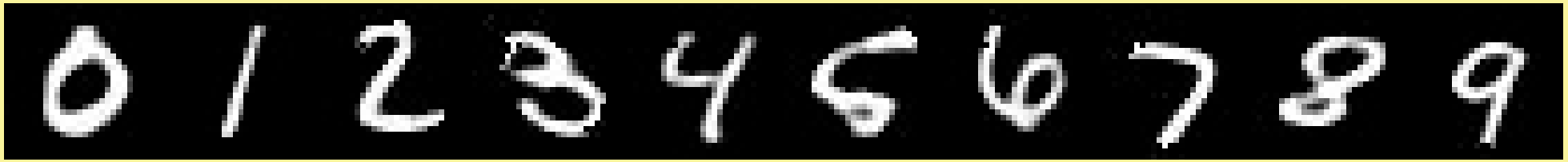
Source: G. E. Hilton and R. R. Salakhutdinov: “Reducing the Dimensionality of Data with Neural Networks”, Science, Vol 313, July 2006, pp. 504-507.

Reconstruction from Latent Space

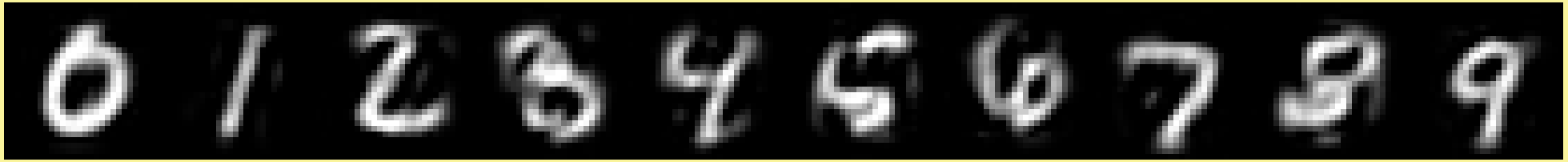
Original



Autoencoder



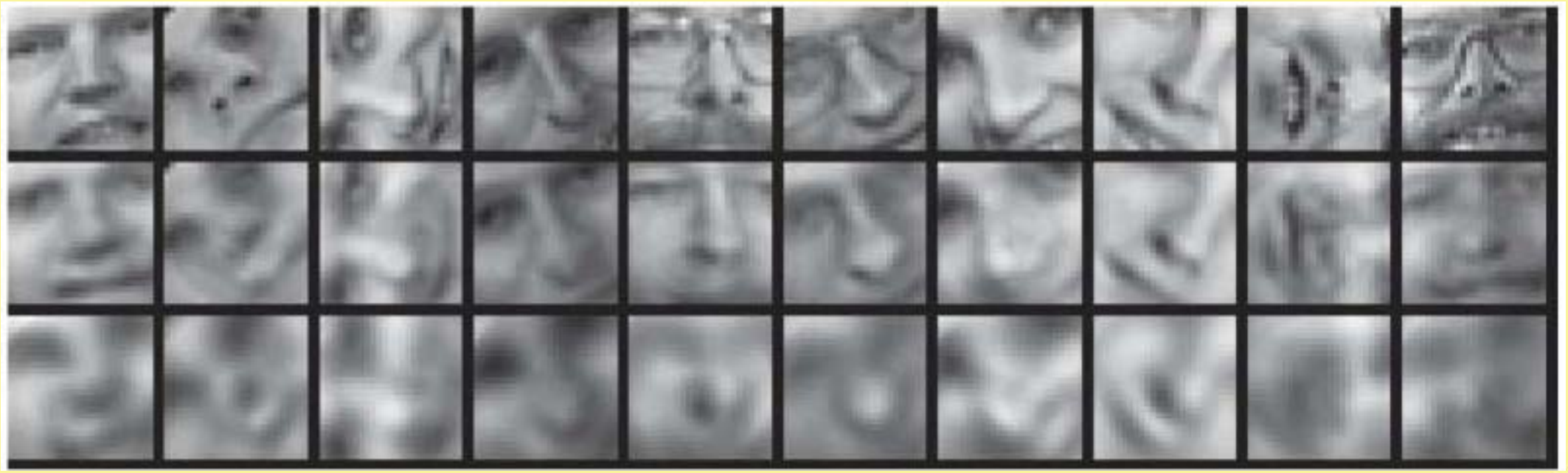
PCA



784 ($ReLU$)
↓
1000 ($ReLU$)
↓
500 ($ReLU$)
↓
250 ($ReLU$)
↓
30 ($\boldsymbol{ReLU}$)
↓
250 ($ReLU$)
↓
500 ($ReLU$)
↓
1000 ($ReLU$)
↓
784

Source: G. E. Hilton and R. R. Salakhutdinov: “Reducing the Dimensionality of Data with Neural Networks”, Science, Vol 313, July 2006, pp. 504-507.

Reconstruction from Latent Space



(a)

(b)

(c)

(a) Original (b) 30-D AE (c) 30-D PC

Source: G. E. Hilton and R. R. Salakhutdinov: “Reducing the Dimensionality of Data with Neural Networks”, Science, Vol 313, July 2006, pp. 504-507.

NPTEL ONLINE CERTIFICATION COURSES

Thank you