AI Summer School 2024
Medical Imaging Informatics

University of Pittsburgh

Deep Learning Computer Vision

Instructor: Ahmad P. Tafti, PhD, FAMIA

Pitt SHRS School of Health and Rehabilitation Sciences

Computational Pathology & AI Center of Excellence

IEEE COMPUTER SOCIETY

Learning Objectives

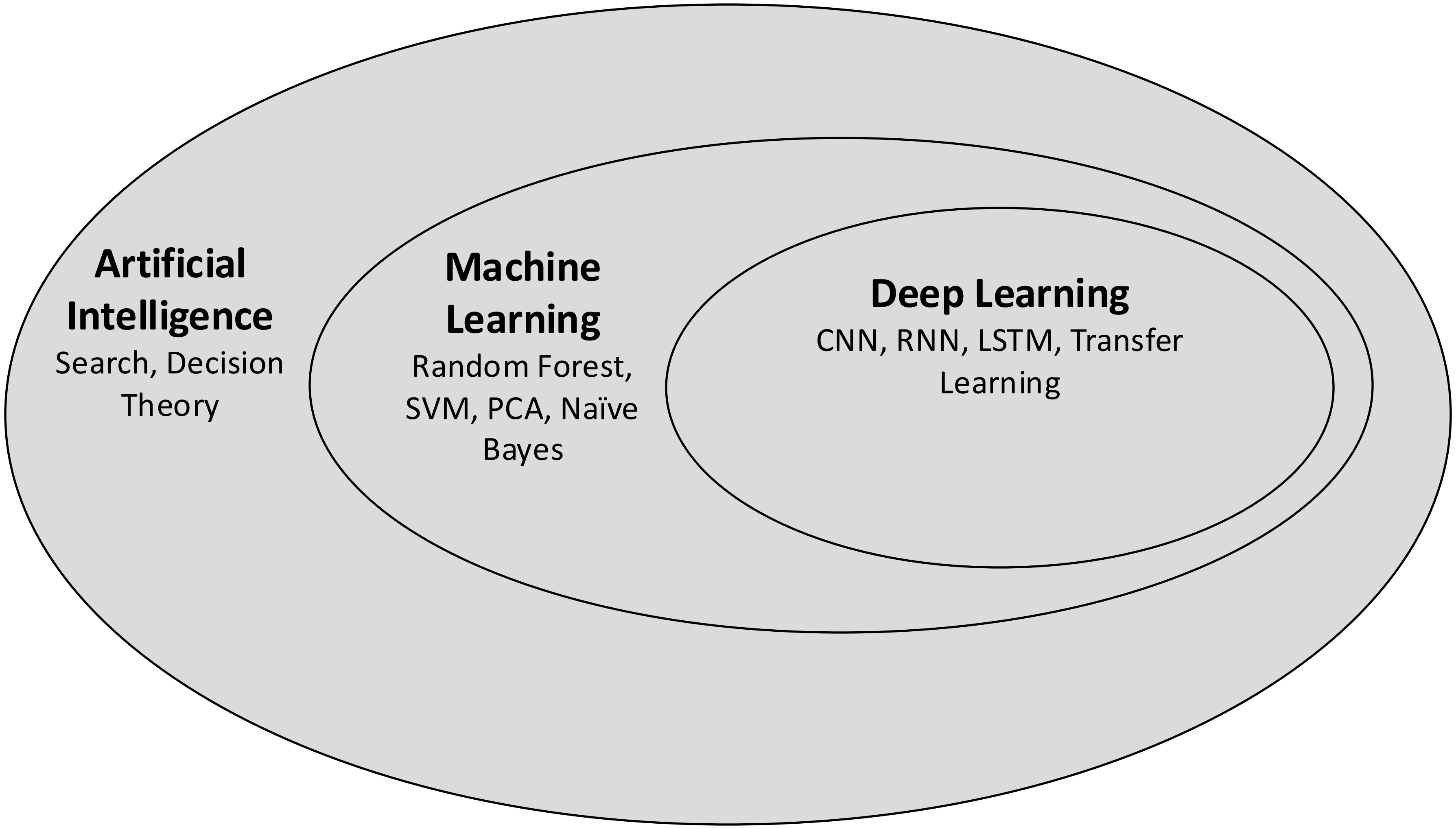
After completing this lecture, you should be able to:

- Understand what is deep learning and why we are using it?
- Demonstrate the Artificial Neural Networks and their applications
- Explain linear and non-linear classification
- Demonstrate shallow and deep neural networks
- Explain deep learning algorithms:
    - Auto-Encoders
    - Generative Adversarial Networks (GANs)
    - Convolutional Neural Networks (CNNs)

Outline

- Deep Learning; What and Why?
- Artificial neural networks (ANNs)
- Linear vs. Non-Linear Classification
- Perceptron Algorithm
- Auto-Encoders
- Generative Adversarial Networks (GANs)
- Convolutional Neural Networks (CNNs)

Deep Learning; What and Why?



Deep Learning; What and Why?

Machine Learning Tasks

Descriptive Machine Learning Algorithms (What happened?)

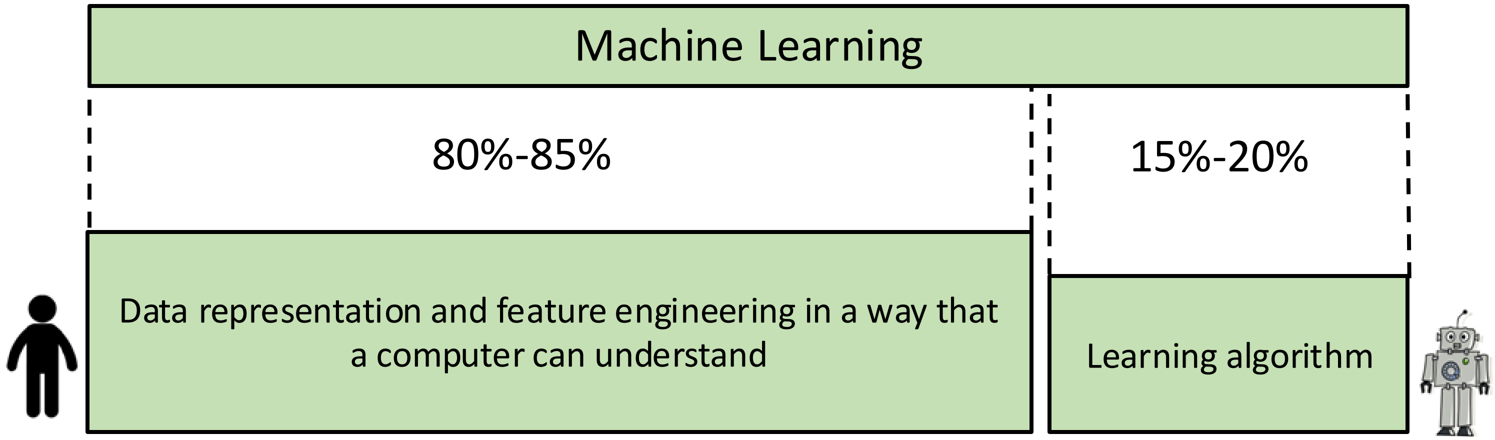
- **Clustering:** Grouping of samples (instances) given un-labeled data.
- **Summarization:** Finding a compact description for a data set.
- **Association Rules:** Discovering interesting relations between variables in a large DB.

Predictive Machine Learning Algorithms (What will be happened?)

- **Regression:** Attempting to predict a continuous attribute.
- **Classification:** Predicting the sample (instance) class from pre-labeled samples.

Deep Learning; What and Why?

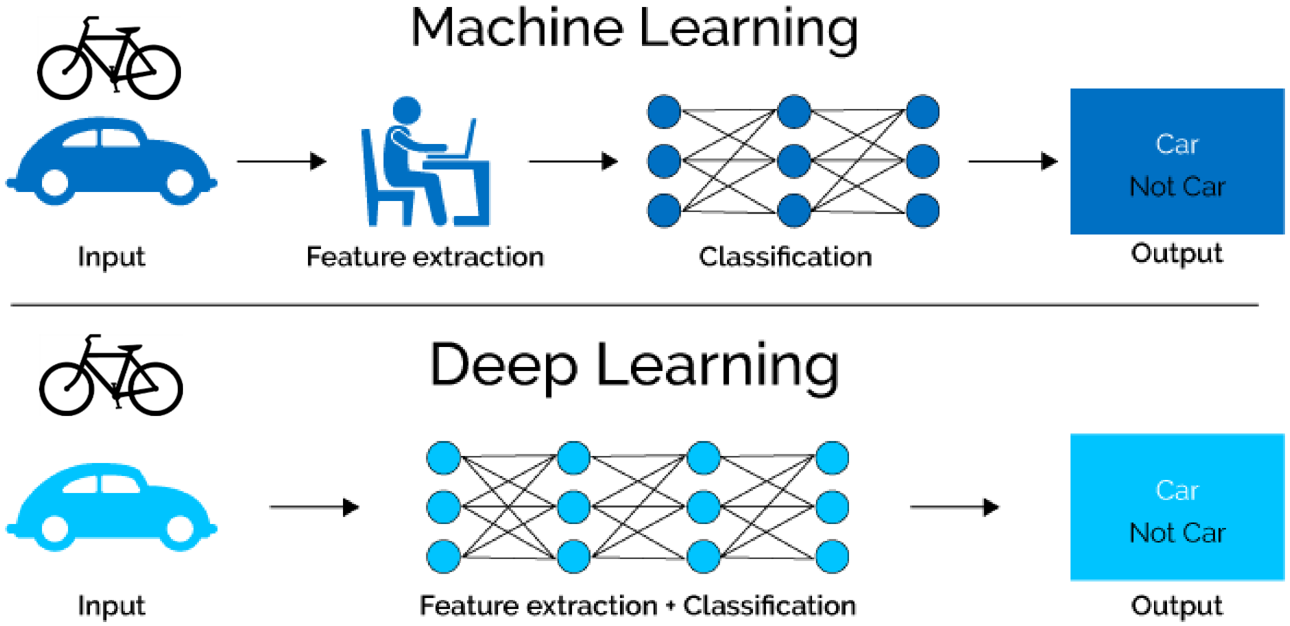
- Most **machine learning** methods work well because of **human-designed representation** and **input features**.
- **Machine learning** becomes **just optimizing weights** to best make a final prediction.



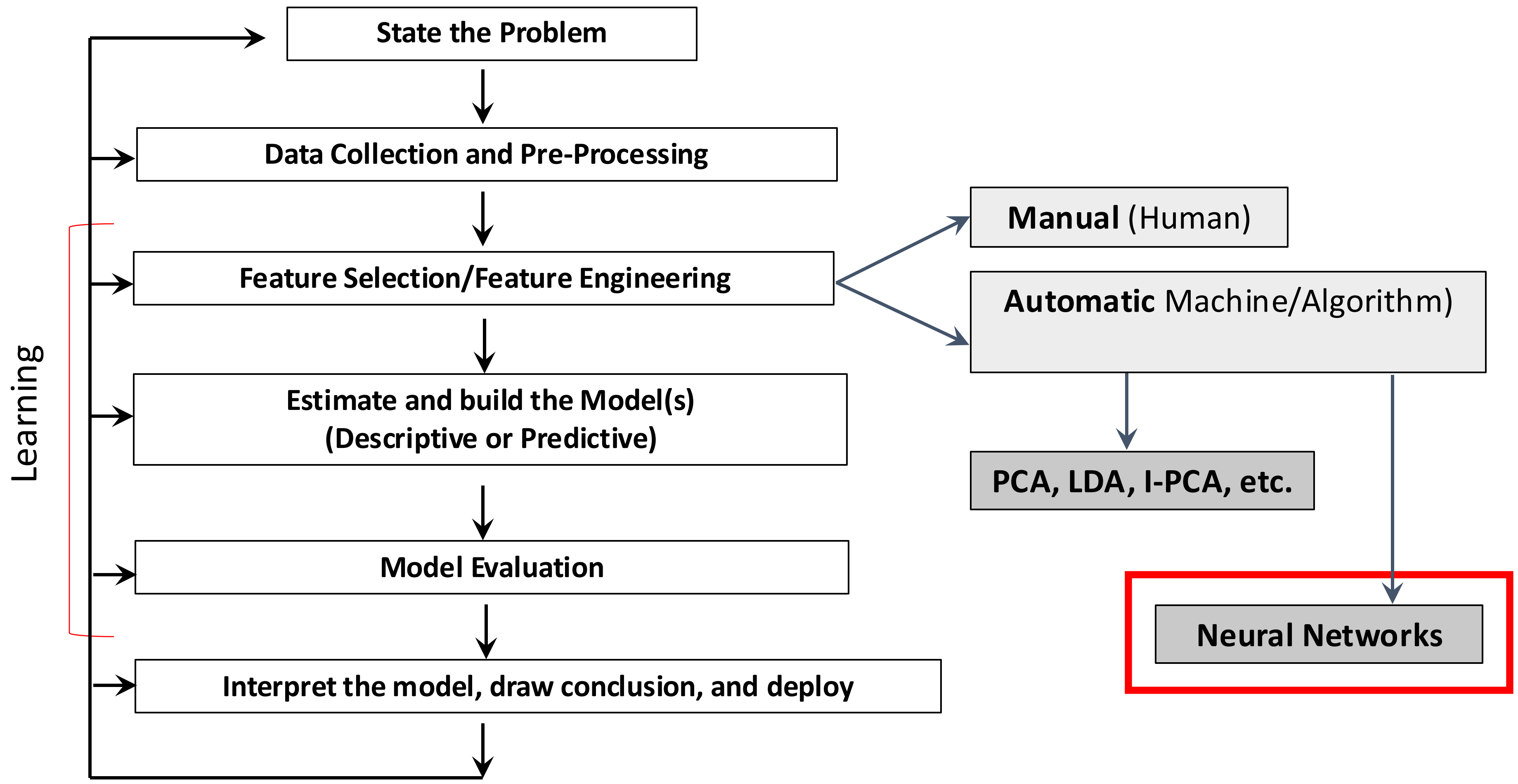
- <u>Problem</u>: **Manually designed features** are often **over-specified**, **incomplete**, and take a **long time** to **design** and **validate**.

Deep Learning; What and Why?

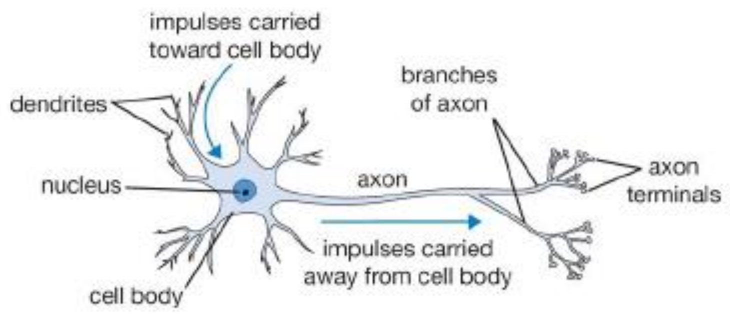
- **Deep Learning algorithms** attempt to **automatically learn good features** or **representation**.
- **Deep Learning** provides a very **flexible** and **universal** learnable framework for representing a variety of data types, such as visual data, linguistics, audio streams, and time series.

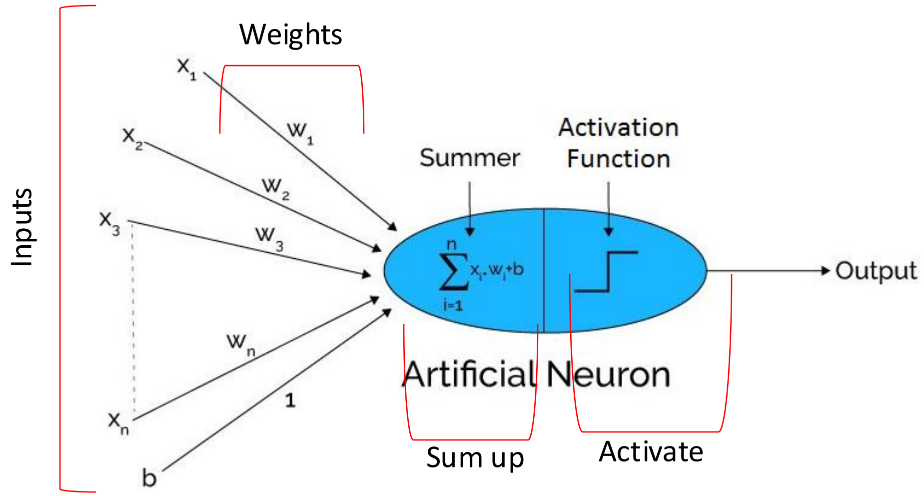


Deep Learning; What and Why?



Artificial Neural Networks (ANNs)





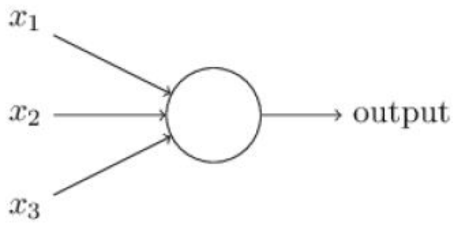
Neuron: Computational building block for the "Brain"
Human Brain: ~100 to 1000 trillion synapses

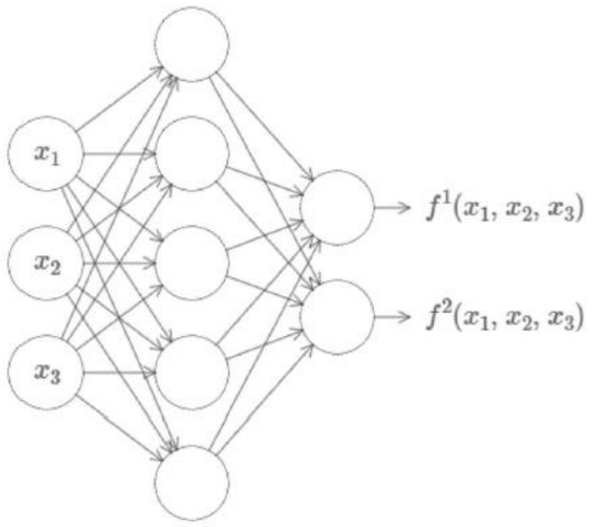
Artificial Neuron: Computational building block for the "Neural Networks"
Neural Network: ~1 to 10 billion synapses

Human brains have ~10,000 computational power than the best computer brain

1 0 0
1 1 1 0 0 0 0 0 0
1 1 0 0 0 0

Artificial Neural Networks (ANNs) are amazing!!!

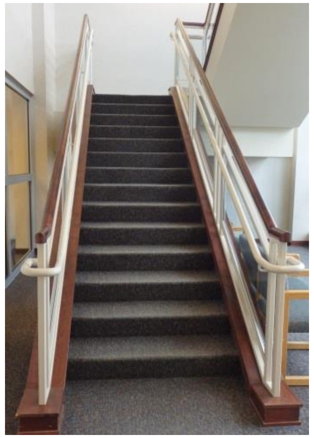




Universality: for any arbitrary function f(x), there exists a neural network that closely approximates it for any input x.

Universality is an incredible property to neural networks, and it holds for just 1 hidden layer.

Linear vs. Non-Linear Classification



[40 * 60 * 3]

Input image

Parameters

f (x, W) = W x +b

3 * 1

3 * 7200

7200 * 1

3 * 1

Stair: 5.4	
Door: -3.7	
Sign: 2.9	

Linear vs. Non-Linear Classification

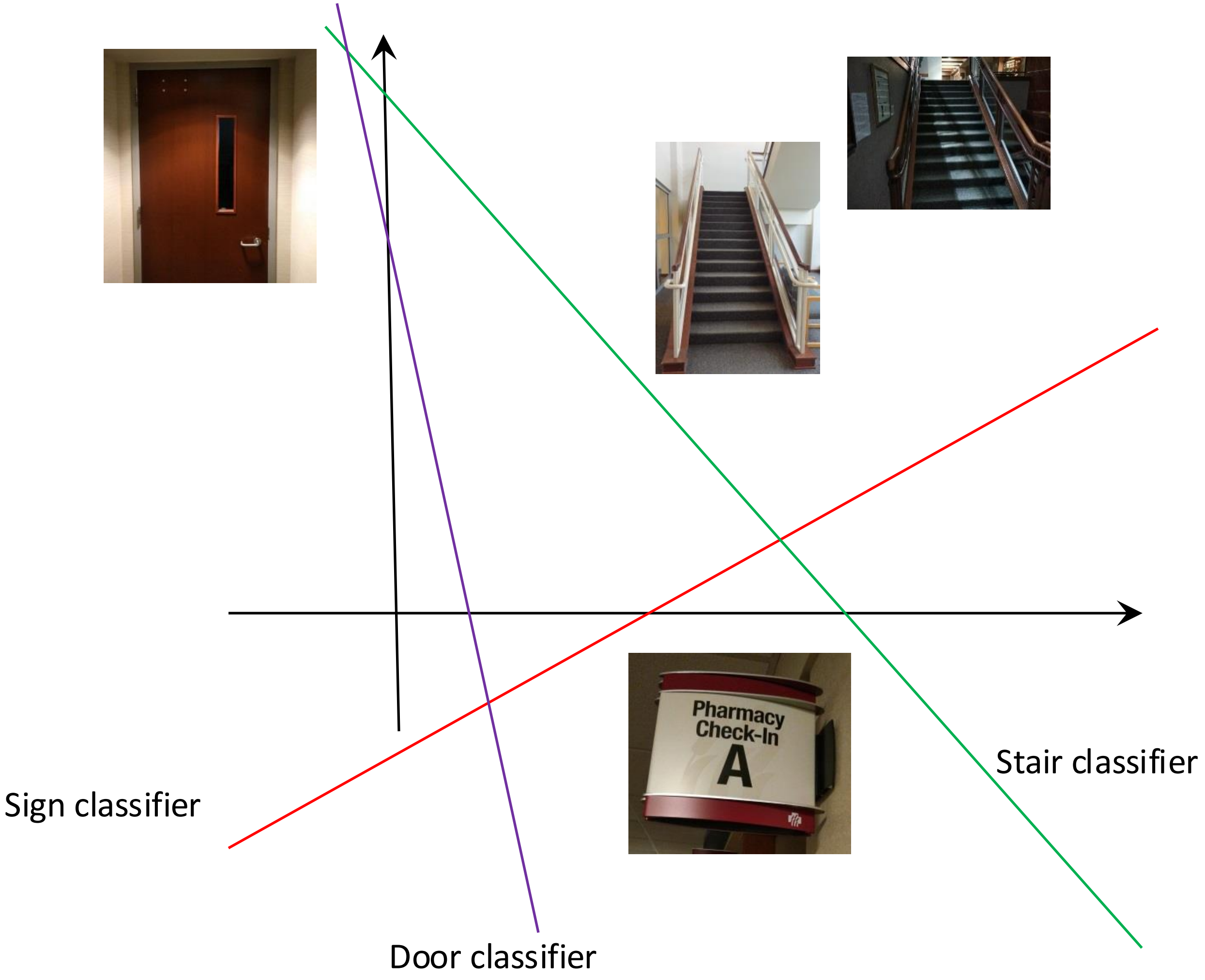
3 Classes AND 4 Features

Input Image

$$\underbrace{\begin{bmatrix} 0.2 & -0.5 & 1 & 2.0 \\ 1.5 & 1.3 & 2.1 & 0.0 \\ 0.0 & 0.25 & 0.2 & -0.3 \end{bmatrix}}_{W} * \underbrace{\begin{bmatrix} 56 \\ 231 \\ 24 \\ 2 \end{bmatrix}}_{x_i} + \underbrace{\begin{bmatrix} 1.1 \\ 3.2 \\ -1.2 \end{bmatrix}}_{b} \rightarrow \underbrace{\begin{bmatrix} -75.2 \\ \mathbf{437.9} \\ 60.75 \end{bmatrix}}_{f\,(x_i;\, W,\, b)} \begin{array}{l} \text{Stair} \\ \text{Door} \\ \text{Sign} \end{array}$$

Linear vs. Non-Linear Classification

$$f\,(x_i;\, W,\, b) = Wx_i + b$$



Perceptron Algorithm

Provide training set of (input, output) pairs and run:

1- Initialize the neural network with random weights

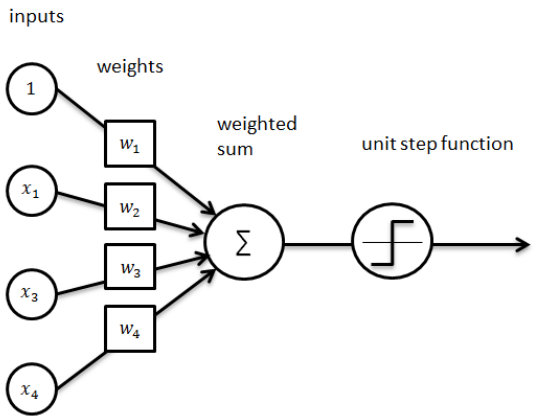
2- For the inputs of an example in the training set, compute the neural network output.

3- If the output does not match the output that is known to be correct for the example:

- If the output should have been 0, but was 1, decrease the weights that had an input of 1
- If the output should have been 1, but was 0, increase the weights that had an input of 1

4- Go to the next example in the training set and repeat steps 2-4 until the neural network makes no mistakes.

F1	F2	F3	F4	Output
7.0	12.25	2.35	11	**0**
10	0.54	17.12	23.05	**1**
…	…	…	…	**1**
…	…	…	…	**1**
…	…	…	…	**0**



Activation Functions

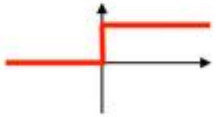
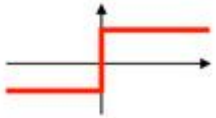
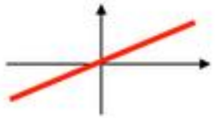
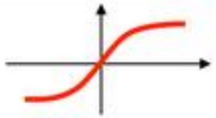
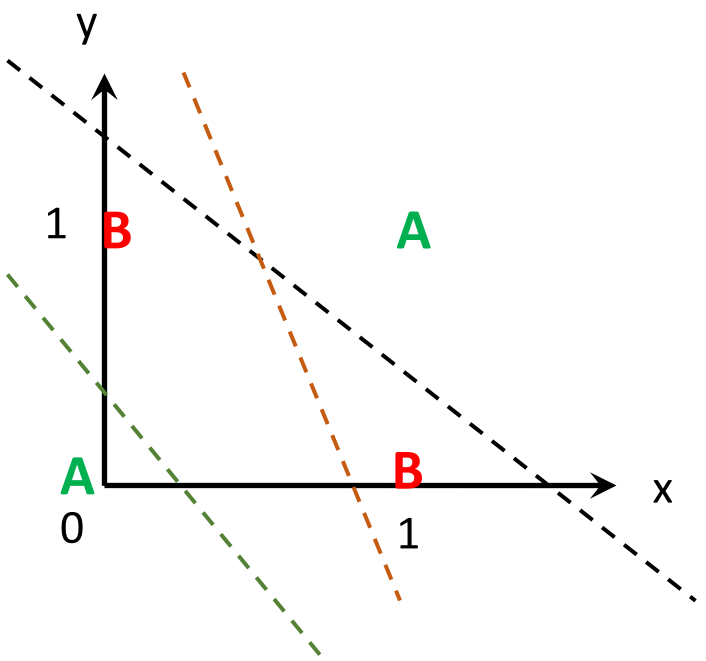
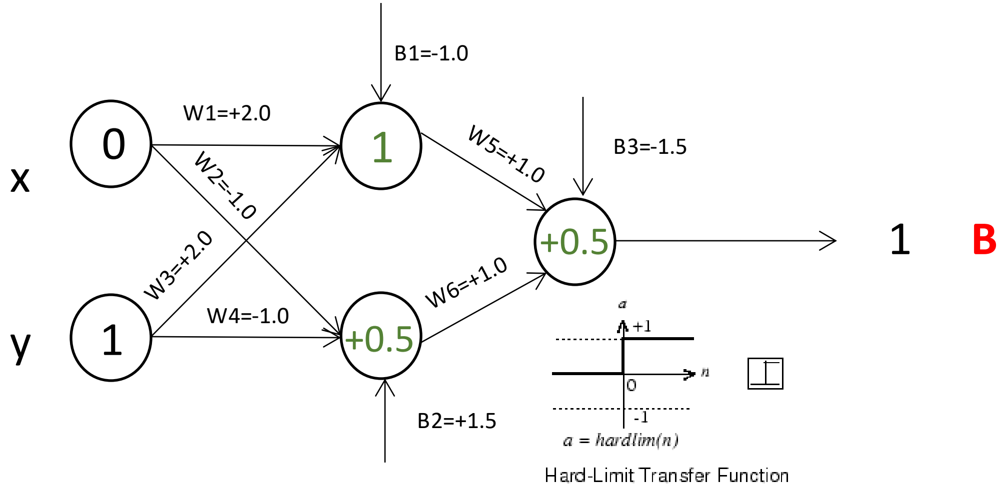
Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = max(0, z)$	Multi-layer Neural Networks	

Image from: http://rasbt.github.io

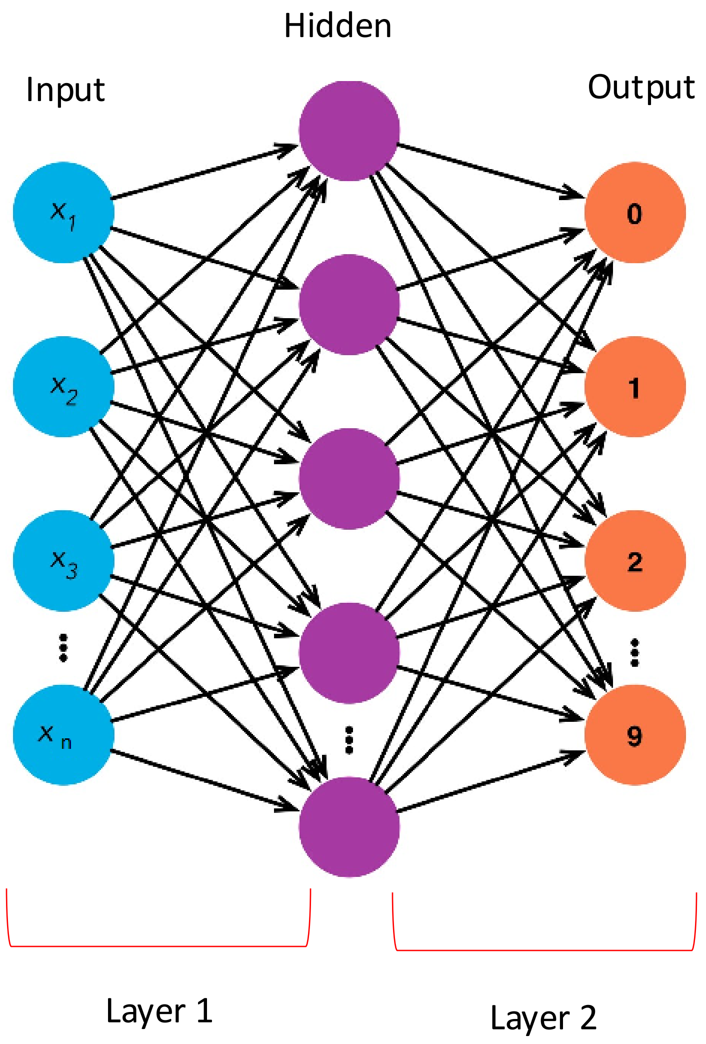
Non-Linear Classification using ANN

x	y	x XOR y	Class
0	0	0	A
0	1	1	B
1	0	1	B
1	1	0	A

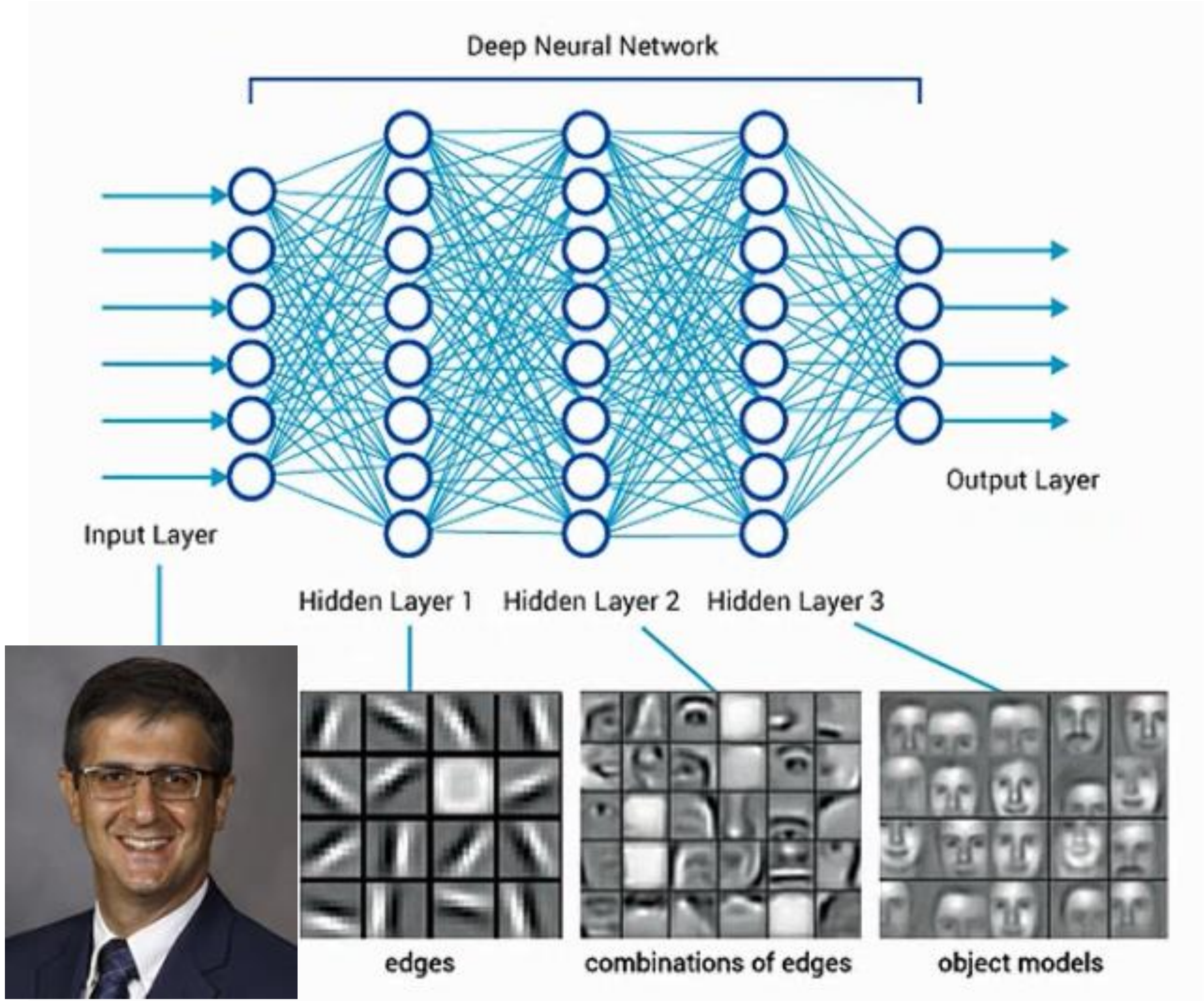




Shallow and Deep ANNs



Shallow and Deep ANNs



Deep Learning Models:

- AE (Auto Encoder)
- CNN (Convolutional Neural Network)
- RNN (Recurrent Neural Network)
- DBF (Deep Belief Network)

Deep ANNs

The first **CNN**, LeNet, to read and understand hand-wri
checks in

199
http://yann

The first **RNN**

Deep Learning stagnation and inactivity.
Reasons:
- Lack of large-scale training data
- Lack of high performance

Hinton, G.E., Osindero, S. and

CNN won several competitions in Image Classification, Object Recognition.

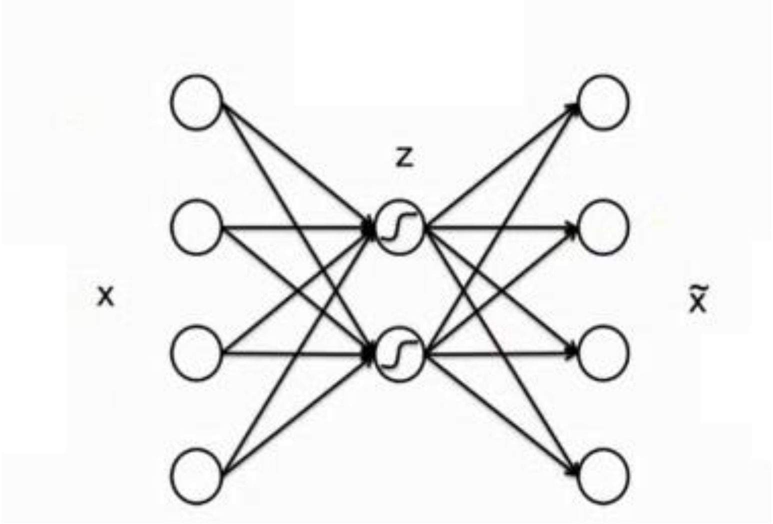
IMAGENET

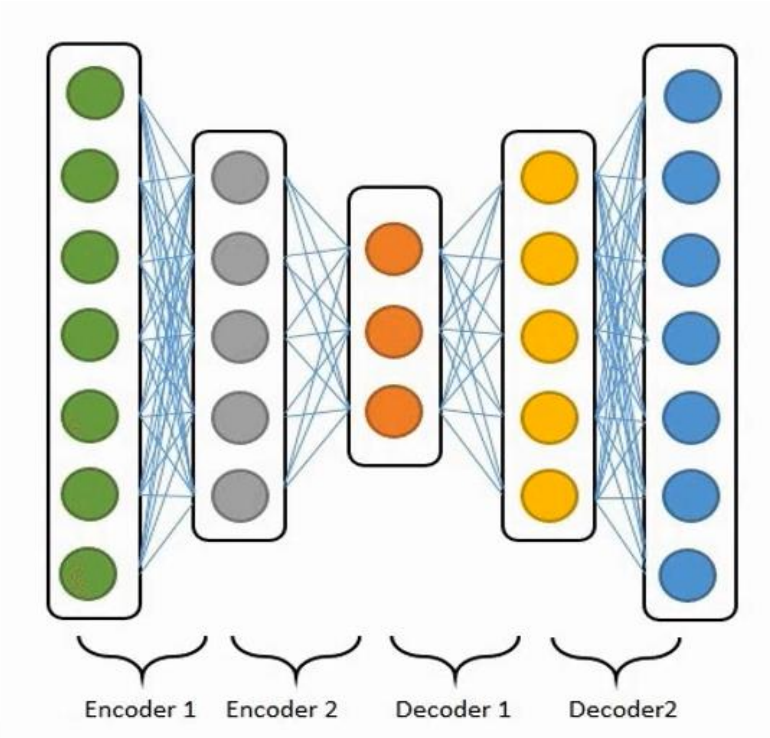
Year 2012

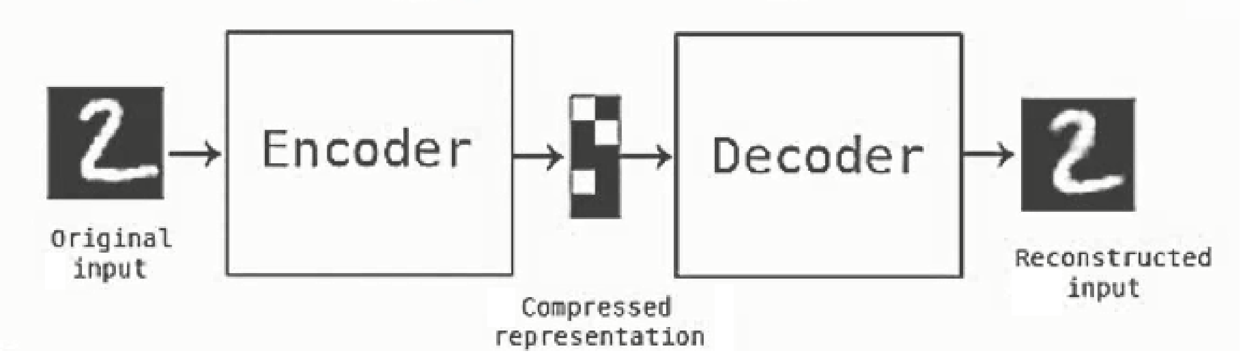
Industries

Universities

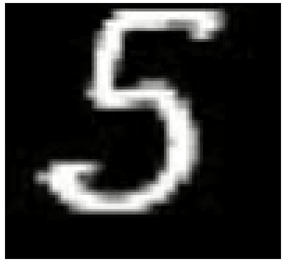
Auto-Encoders





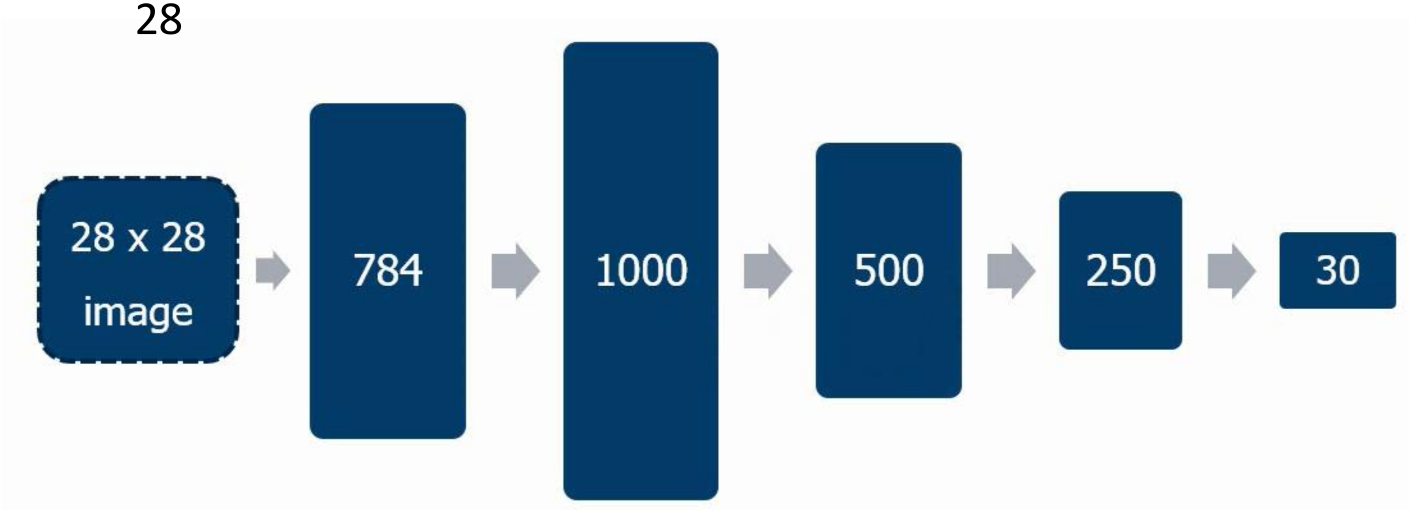


Auto-Encoders: Example

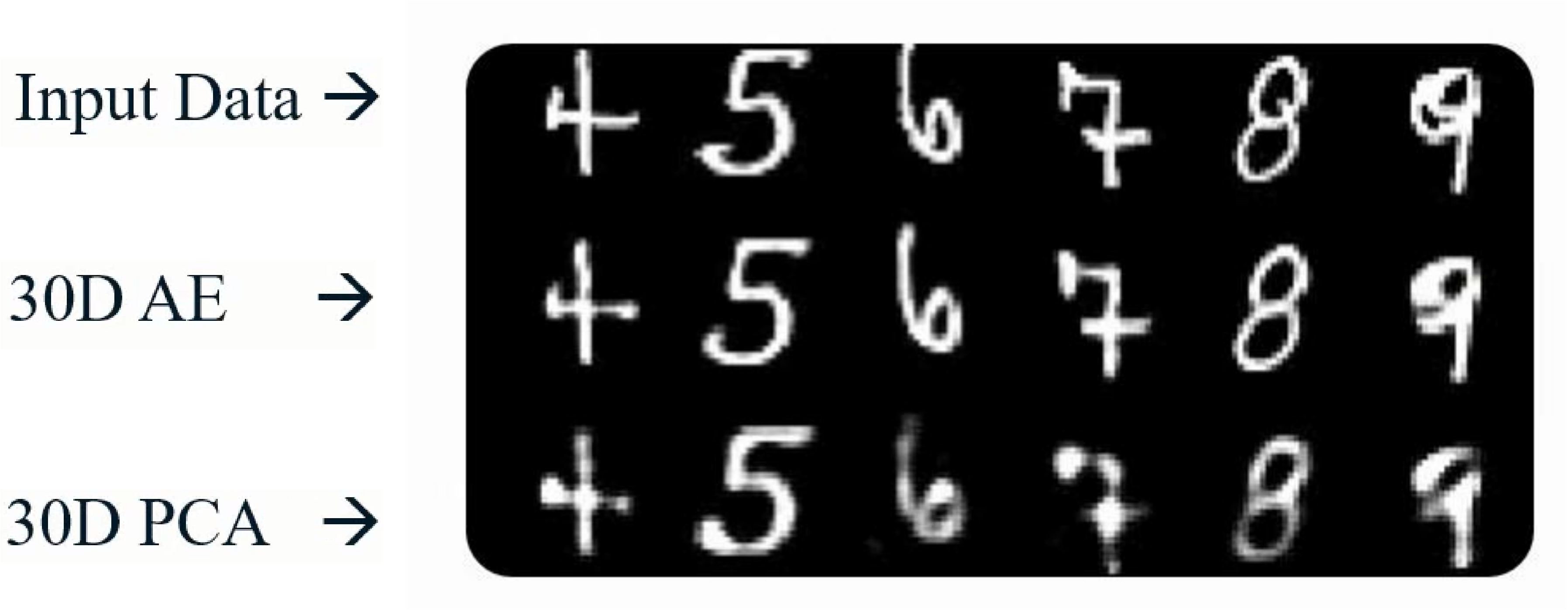


28

28



Auto-Encoders: Example



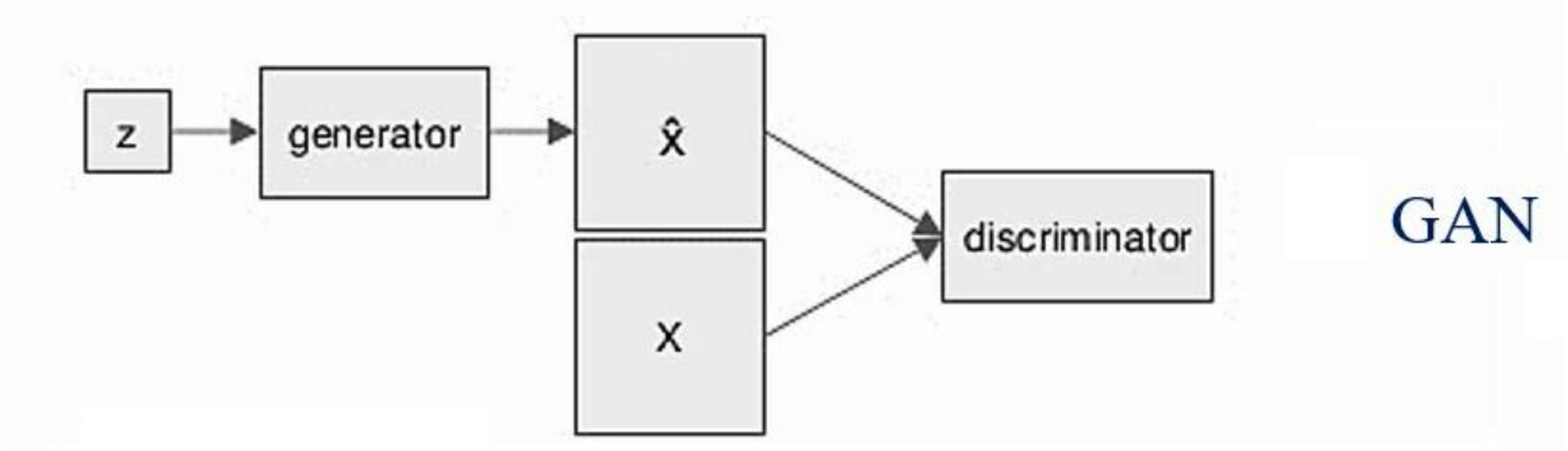
Auto-Encoders: Applications

- **Denoising**
- **Data Compression**
- **Clustering**
- **Manifold Learning**

Auto-Encoders:

- Stacked Auto-Encoder (SAE)
- Denoising Auto-Encoder (DAE)
- Sparse Auto-Encoder (SPAE)
- Contractive Auto-Encoder (CAE)
- Convolutional Auto-Encoder (CNAE)
- Variational Auto-Encoder (VAE)

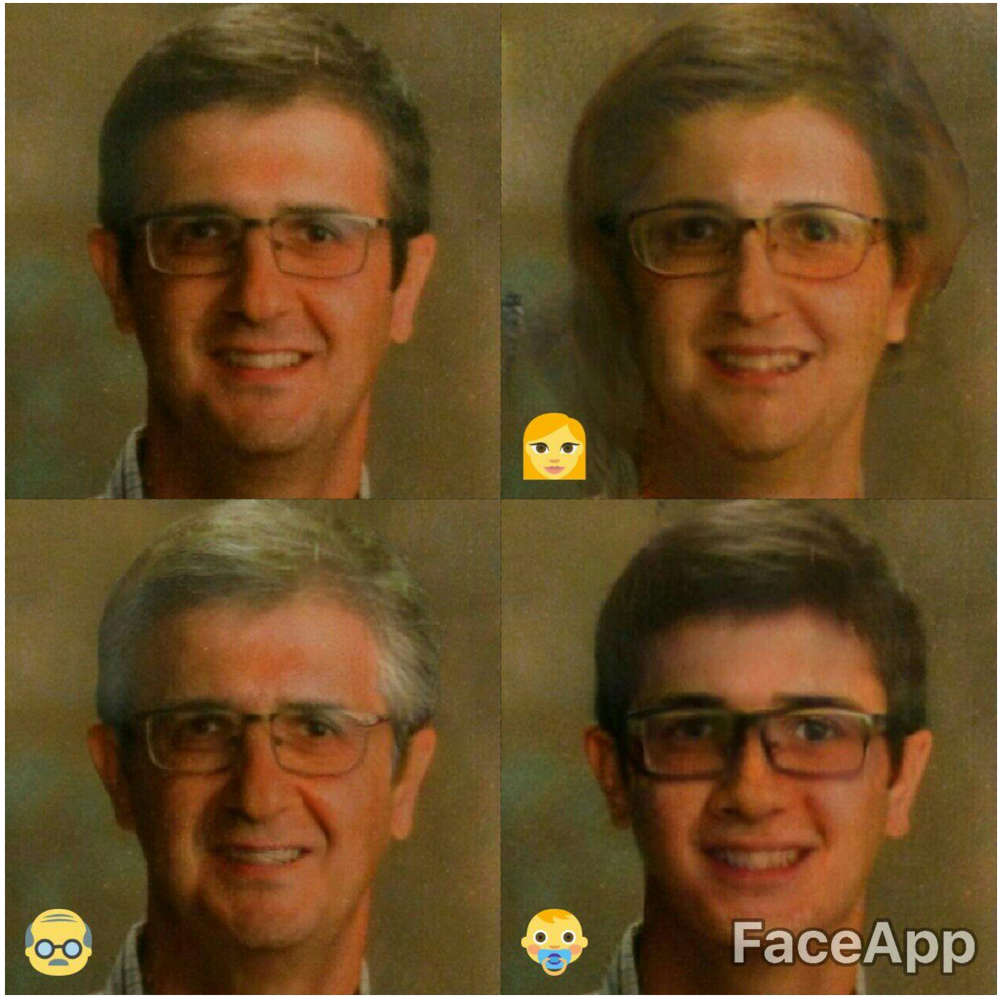
Generative Adversarial Networks (GANs)



Generative Adversarial Networks (GANs)

Real Data →

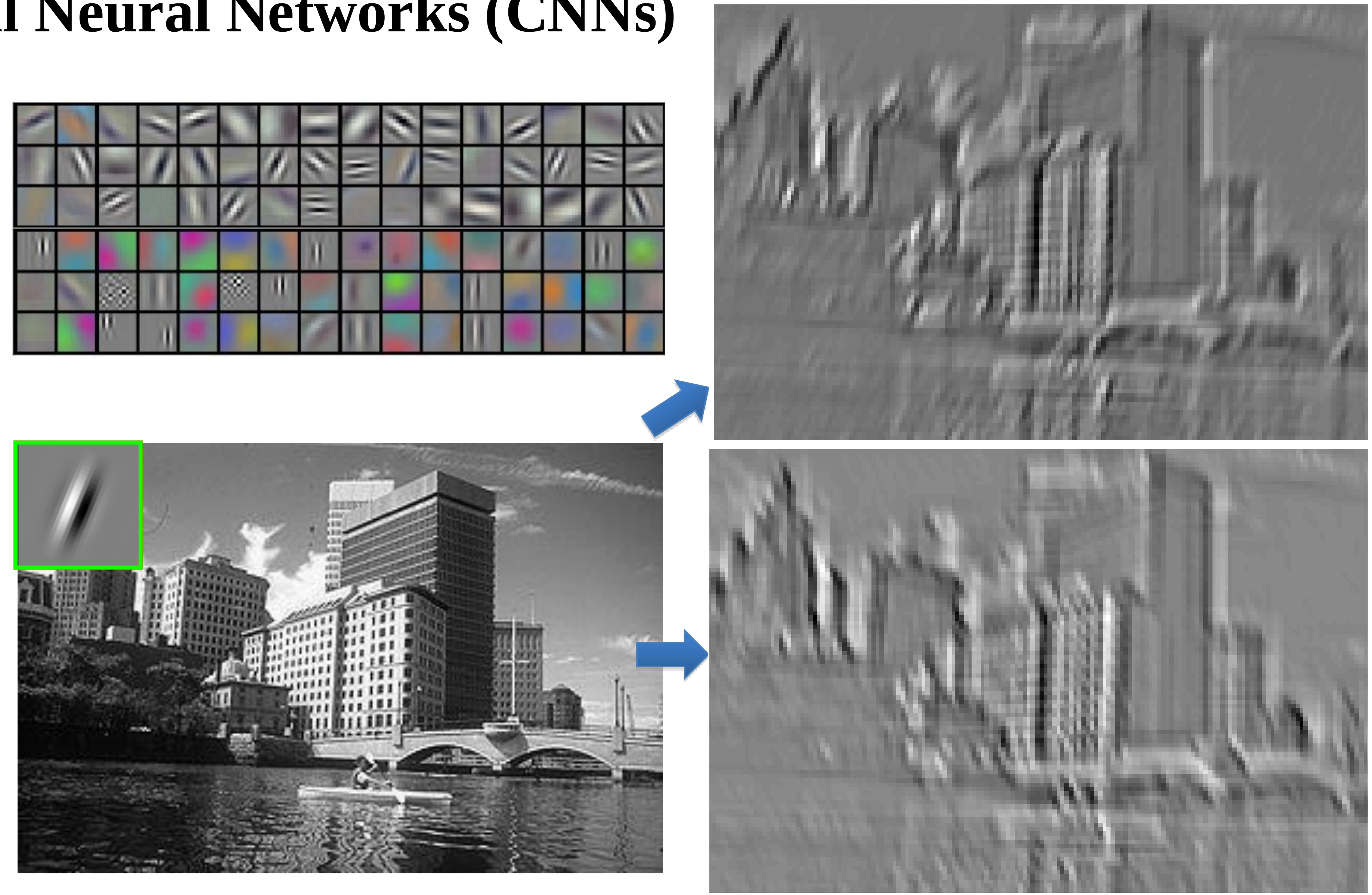
Synthetic data →



← Synthetic data

← Synthetic data

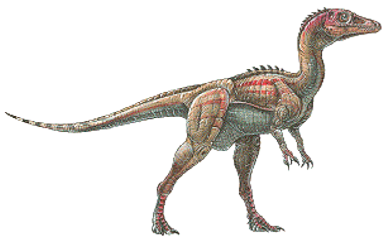
Convolutional Neural Networks (CNNs)



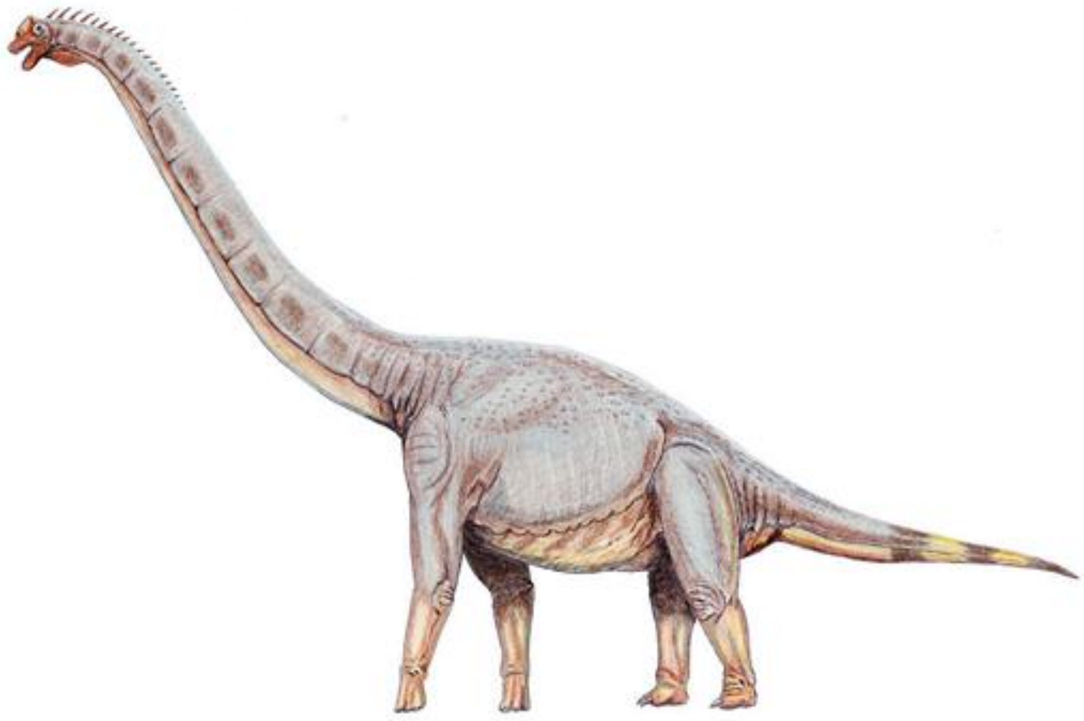
The convolution operation (slide adopted from [1])

1. Fergus, R., el al. (2012) *Deep learning methods for vision*. In: CVPR http://cs.nyu.edu/~fergus/tutorials/deep_learning_cvpr12/

Deep Learning VS. Machine Learning



Machine Learning



Deep Learning

Deep Learning VS. Machine Learning

Advantages	Disadvantages
Automatic feature extraction: It reduces the need for feature engineering, one of the most time-consuming parts of machine learning practice.	It requires a large amount of data. If we only have thousands of examples, deep learning is unlikely to outperform other approaches.
Multi-layer feature representation/learning	It is extremely computationally expensive to train. Complex models take weeks to train. We do need GPUs to speed up the process.
More accurate learning methods	Deep learning algorithms do not have much in the way of strong theoretical foundation
Can be adapted to new problems relatively easily	What is learned is not easy to comprehend. Other classifiers (e.g., decision trees, logistic regression, etc.) make it much easier to understand what's going on.

Thank you!

Questions!