COMP 532
Machine Learning and BioInspired Optimization

Lecture 9: Deep Learning

Dr. Shan Luo

Department of Computer Science

shan.luo@liverpool.ac.uk

Deep Learning - Overview

- Reading material
  - **Main reading**:
    CS231n: CNNs for Visual Recognition[1]
  - **Optional reading**:
    Deep Learning book by Goodfellow, Bengio & Courville[2]
- We will discuss:
  - Basics of Artificial Neural Networks
  - Backpropagation
  - Convolutional Networks (Deep)
  - Example: Deep Q-learning, Alpha Go

[1] http://cs231n.github.io/
[2] http://www.deeplearningbook.org/

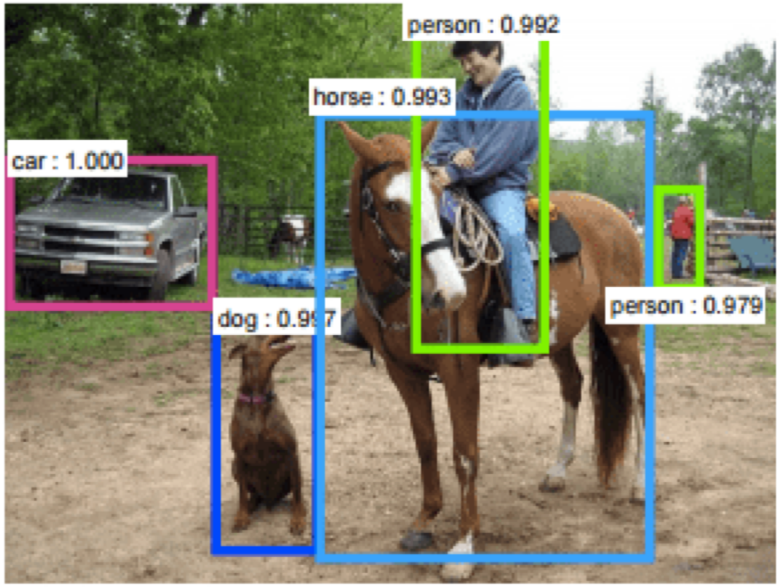
Lecture Overview

- Introduction to Deep Learning

- Background: Neural Networks
  - Biological inspirations
  - History of Artificial Neural Networks
  - From real to artificial neurons

Applications of Deep Learning

Computer Vision



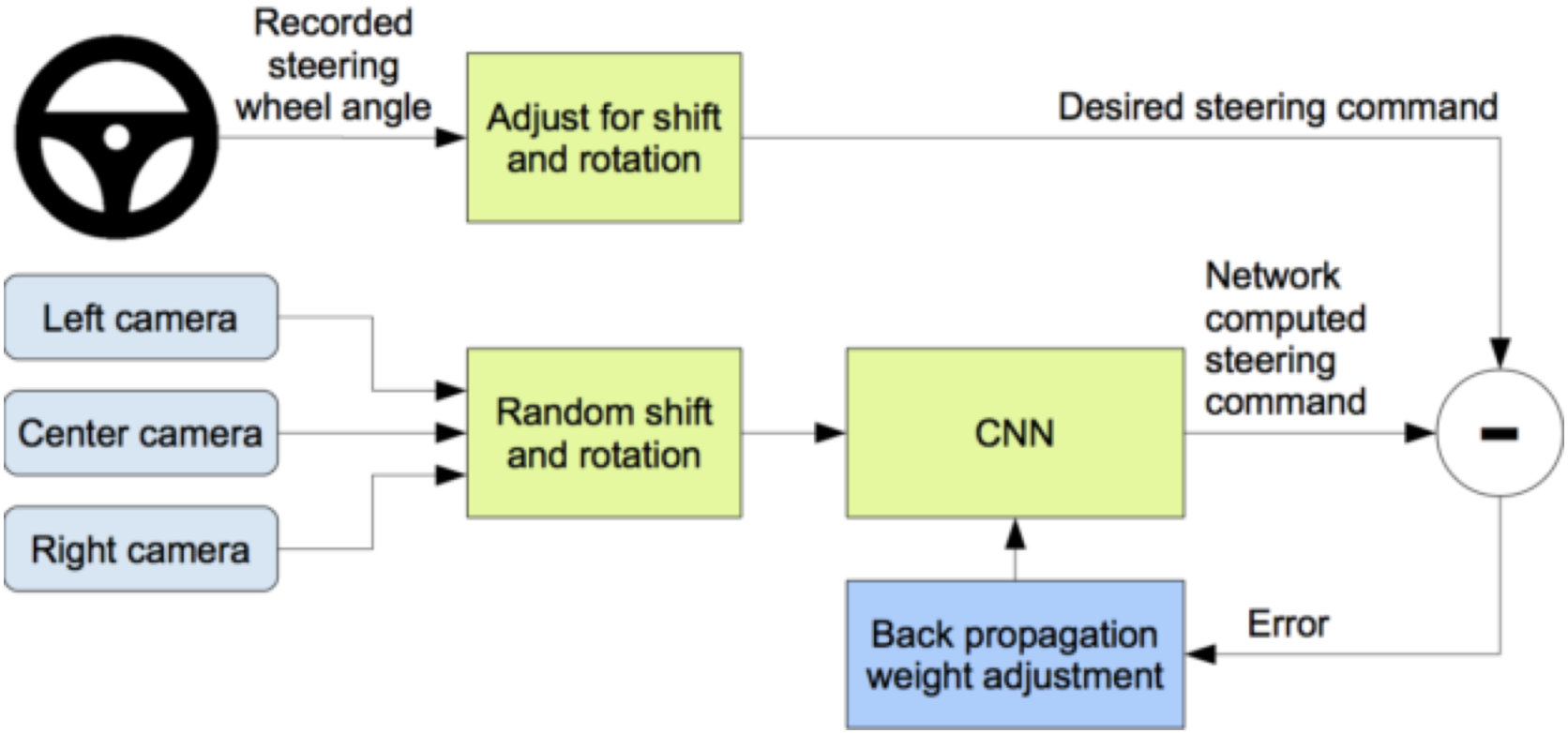
Object recognition/detection

Scene segmentation

3D face reconstruction

Applications of Deep Learning

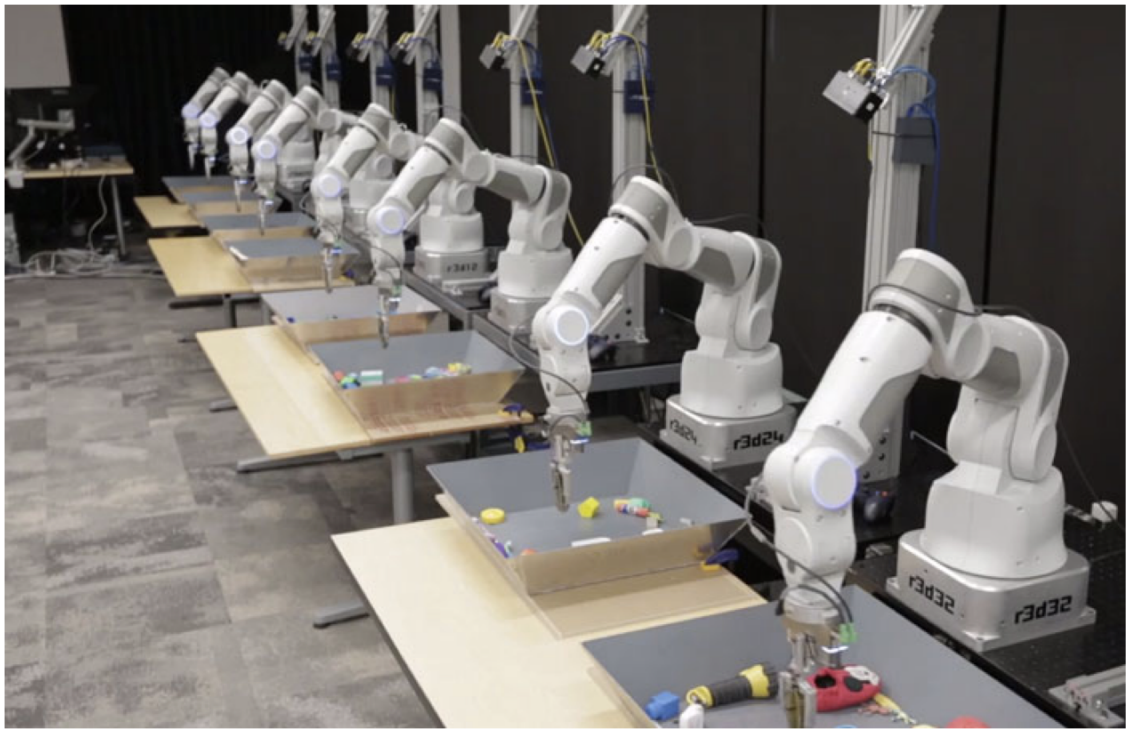
Self-driving cars



https://devblogs.nvidia.com/deep-learning-self-driving-cars/

Applications of Deep Learning

Robotics



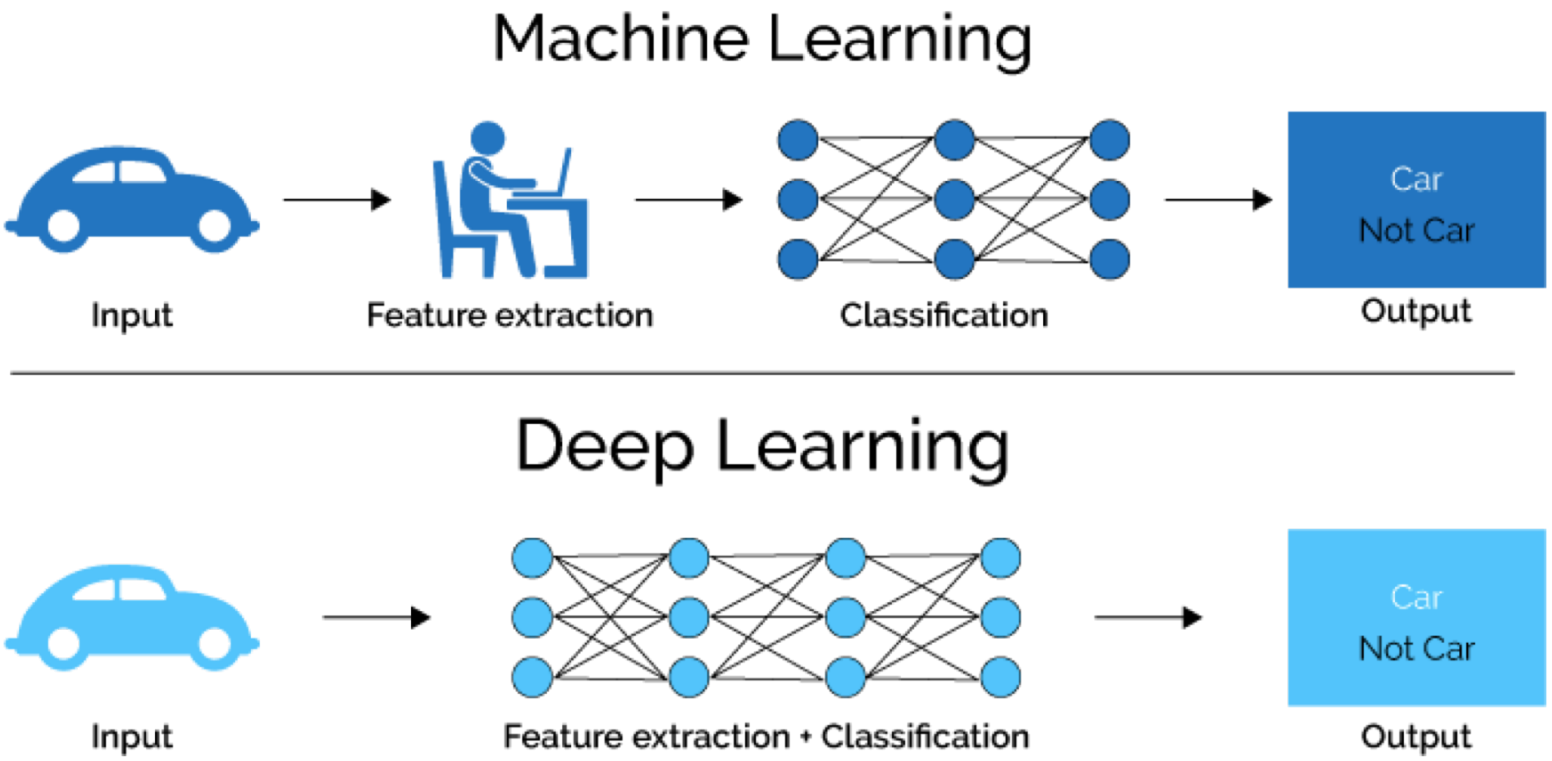
Levine, Sergey, et al. "Learning hand-eye coordination for robotic grasping with large-scale data collection." ISER, 2016.

Applications of Deep Learning

And many more…

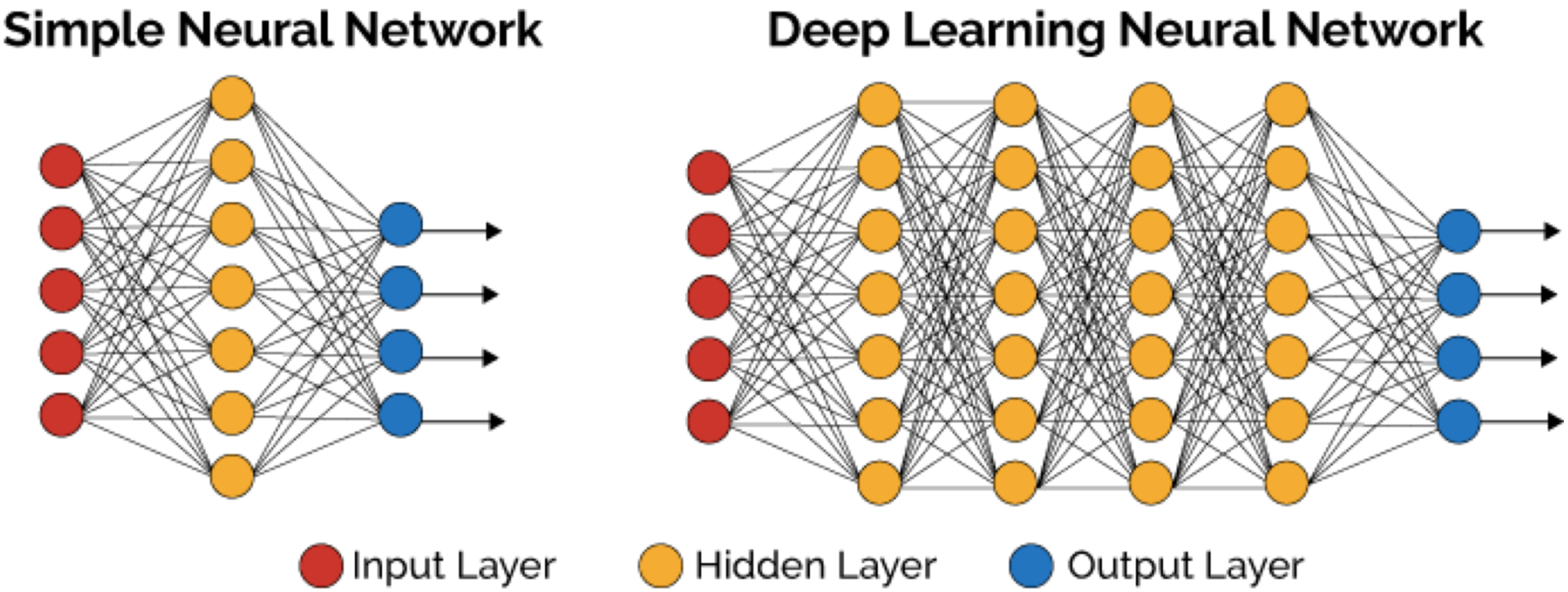
What is Deep Learning?

Deep Learning is a family of methods that uses deep architectures to learn high-level feature representations



What is Deep Learning?

Deep Learning is a family of methods that uses deep architectures to learn high-level feature representations

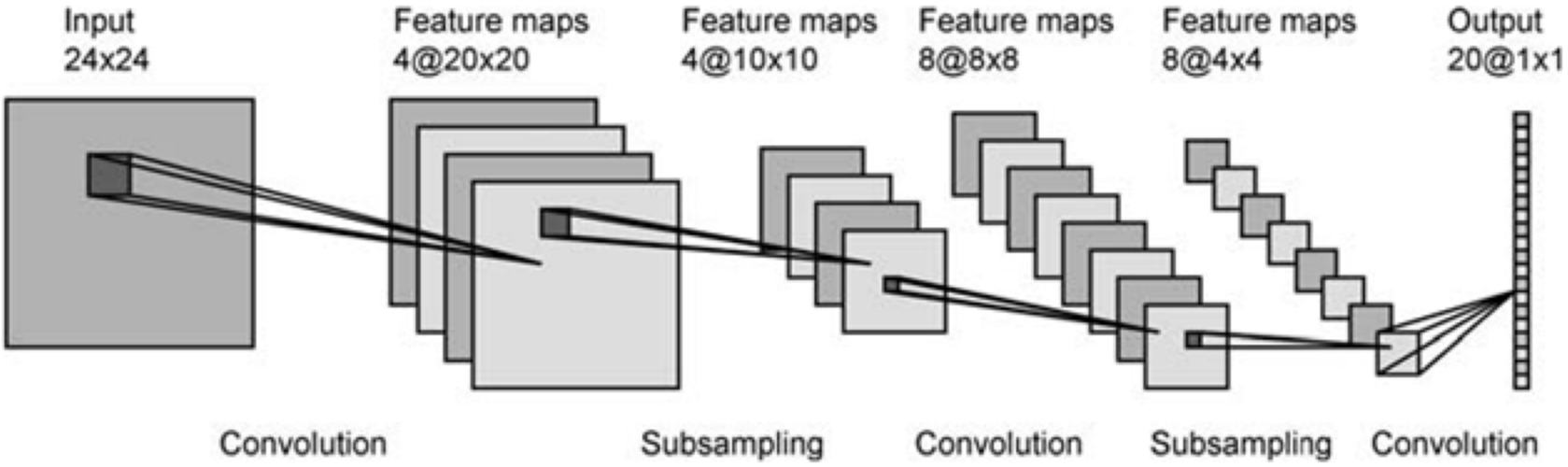


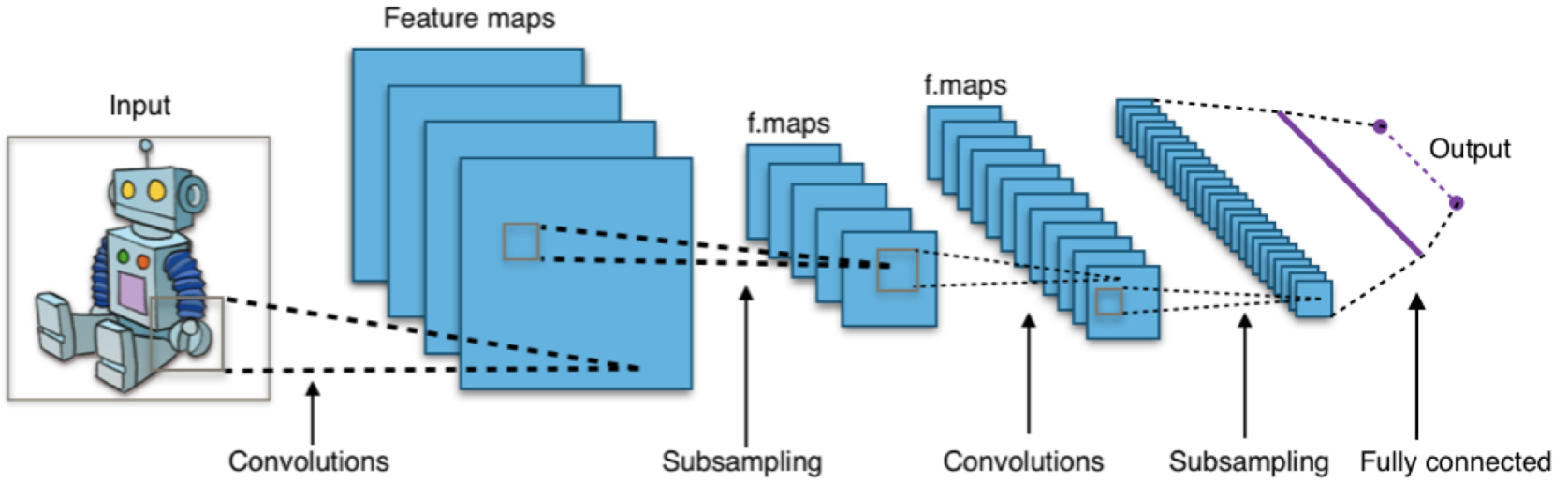
What is Deep Learning?

Some different types of deep architectures:

- **Convolutional Networks**
  - for e.g. image classification
- **Auto Encoders**
  - for dimensionality reduction
- **Deep Belief Networks**
  - for image recognition and generation
- **Recurrent Networks**
  - for learning patterns in sequential data

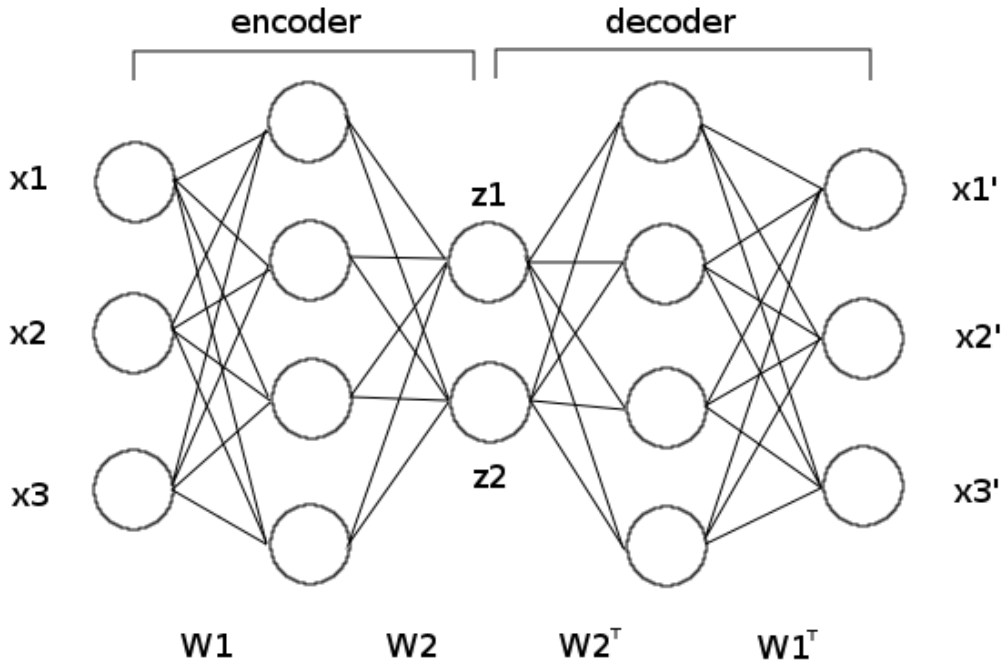
Convolutional Networks





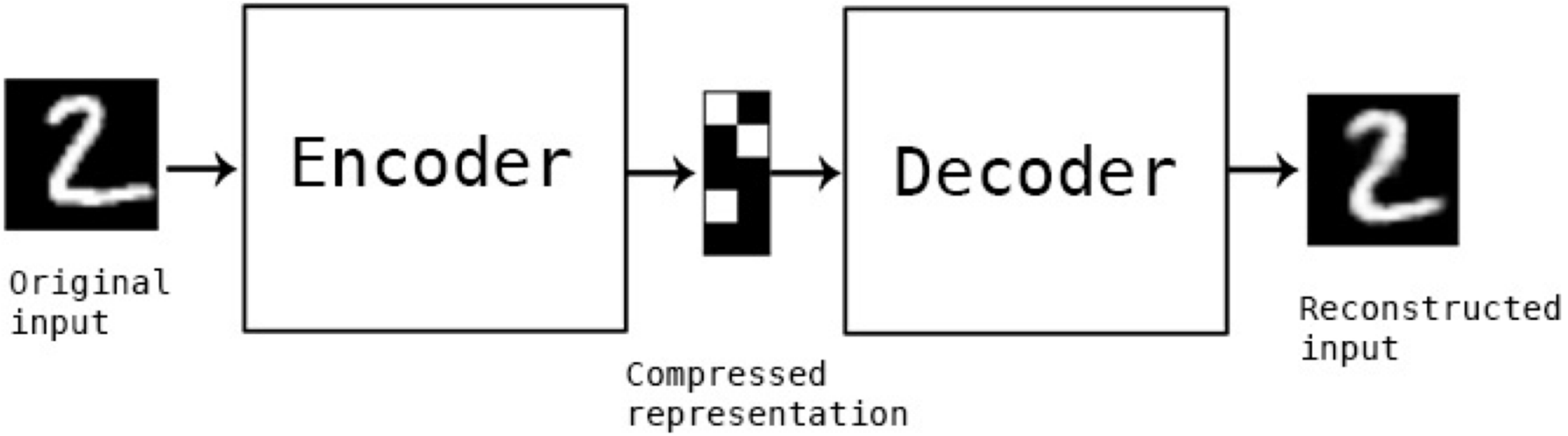
Auto Encoders

Auto Encoders learn to “reproduce” their own input after passing it to a lower dimensional layer

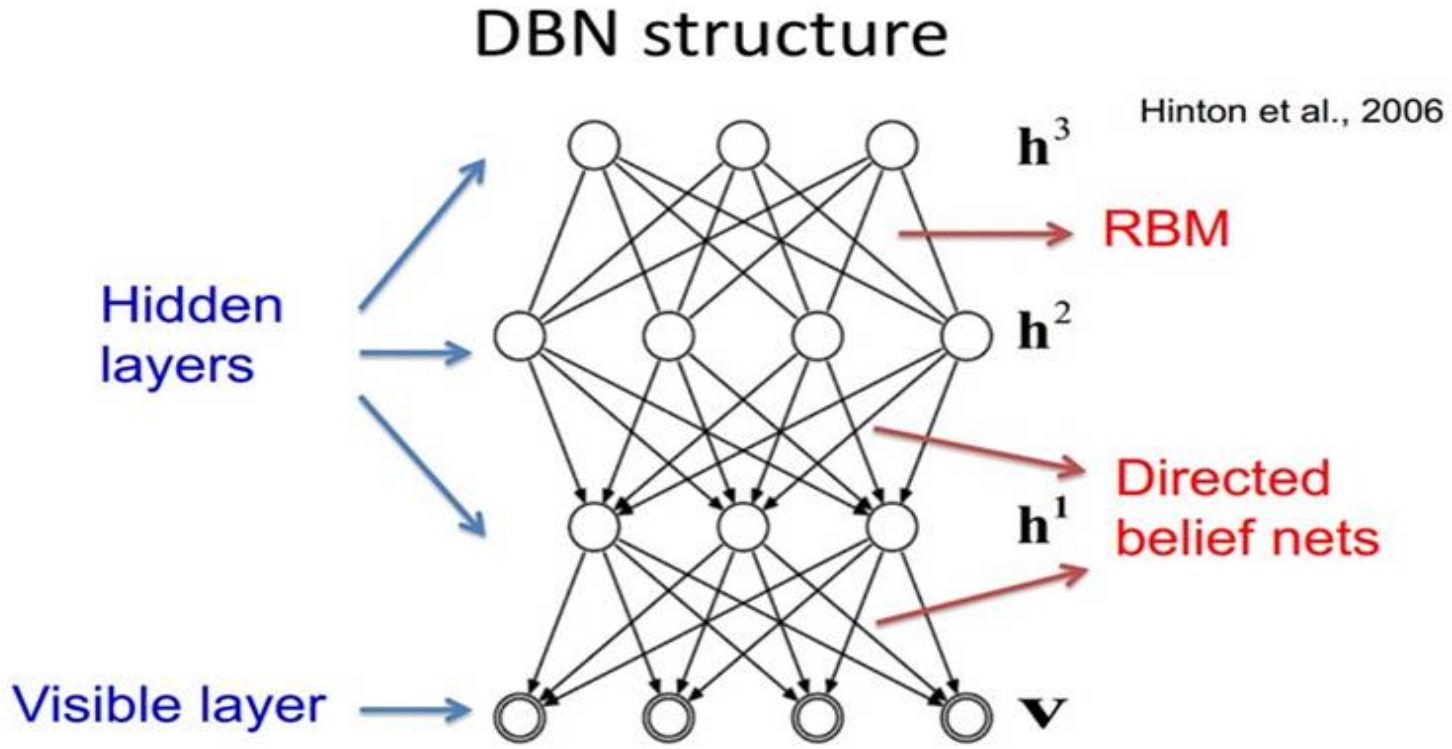


Auto Encoders

Auto Encoders learn to “reproduce” their own input after passing it to a lower dimensional layer



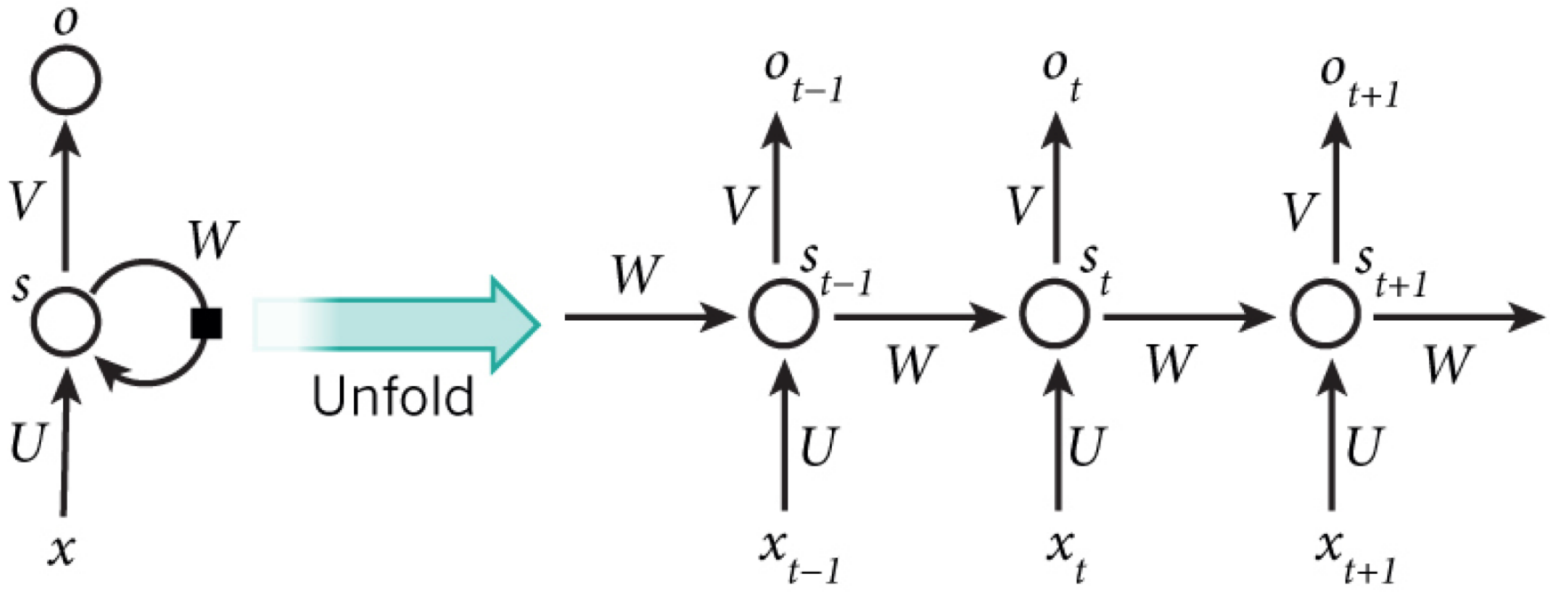
Deep Belief Networks



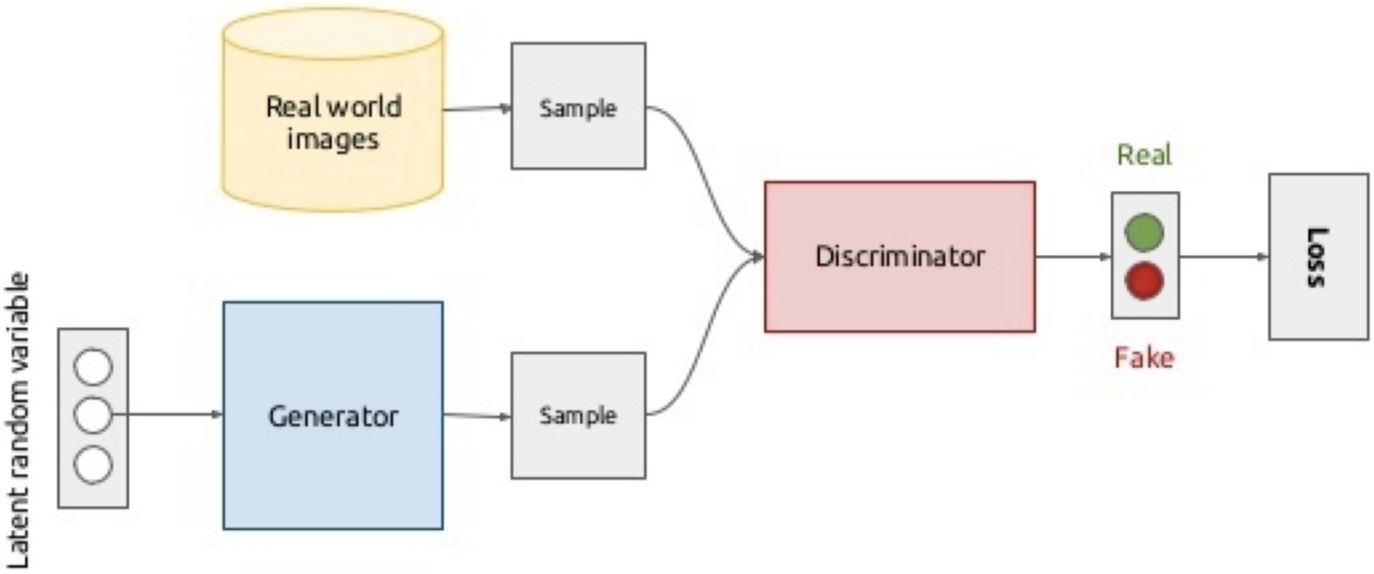
$$P(\mathbf{v},\mathbf{h}^1,\mathbf{h}^2,\ldots,\mathbf{h}^l) = P(\mathbf{v}\,|\,\mathbf{h}^1)P(\mathbf{h}^1\,|\,\mathbf{h}^2)\ldots P(\mathbf{h}^{l-2}\,|\,\mathbf{h}^{l-1})P(\mathbf{h}^{l-1},\mathbf{h}^l)$$

Recurrent Networks

- Output is fed back into hidden layer, serves as "memory"
- Useful to analyze sequences of data, e.g. speech recognition



Generative adversarial networks (GAN)



5

https://deeplearning4j.org/generative-adversarial-network
https://github.com/nightrome/really-awesome-gan#applied-vision

Breakthrough

- There were always good algorithms for learning weights in networks with 1 hidden layer, but…

  …these algorithms are not good at learning the weights for networks with more hidden layers

- **The big breakthrough:**
  algorithms for training many-layer networks

BACKGROUND

What are Neural Networks?

- Animals are able to **react adaptively** to changes in their external and internal environment, and they use their **nervous system** to perform these behaviours.
- An **appropriate model/simulation** of the nervous system should be able to produce similar responses and behaviours in artificial systems.
- The nervous system is built by **relatively simple units**, the neurons, so **copying their behaviour and functionality should be the solution**.

What are Artificial Neural Networks?

- Models of the brain and nervous system
- Highly parallel
  - Process information much more like the brain than a serial computer
- Able to learn complex functions

Very simple principles - Very complex behaviours!

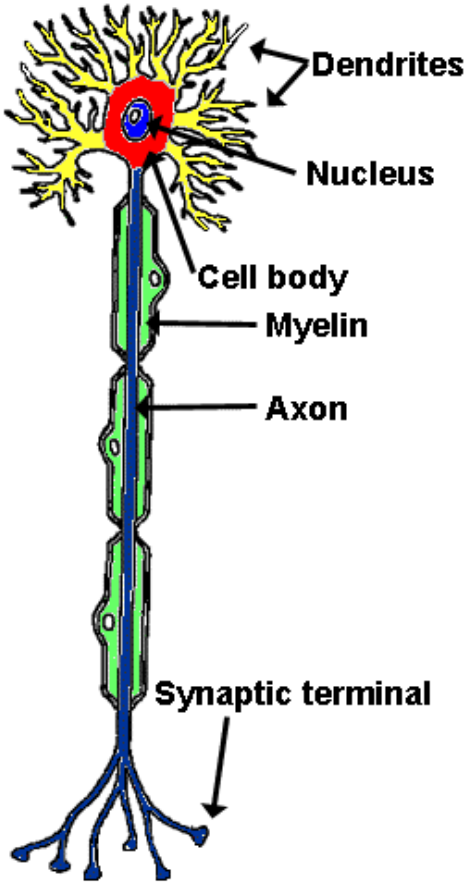
How do NNs and ANNs work?

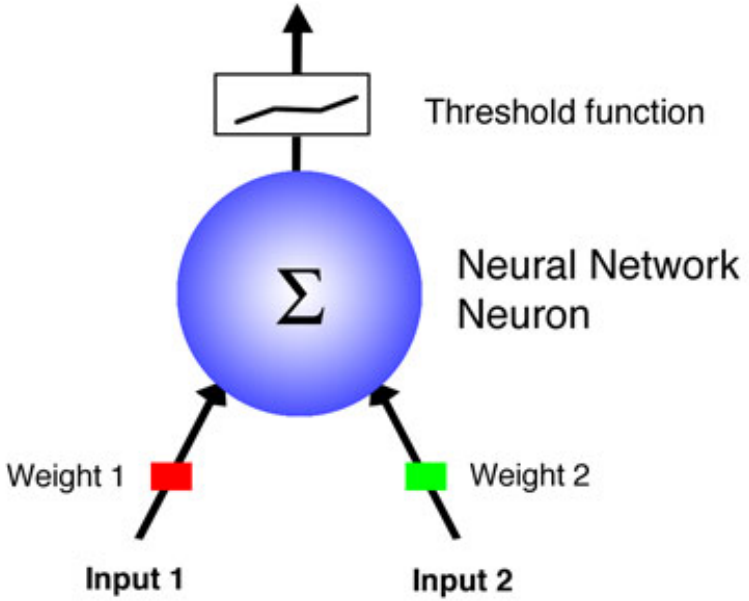
- The “building blocks” of neural networks are the **neurons**.
- They are also referred to as **units** or **nodes**.
- Basically, each neuron
  - receives **input** from many other neurons,
  - changes its internal state (**activation**) based on the current input,
  - sends **one output signal** to many other neurons, possibly including its input neurons (recurrent network)

How do NNs and ANNs work?

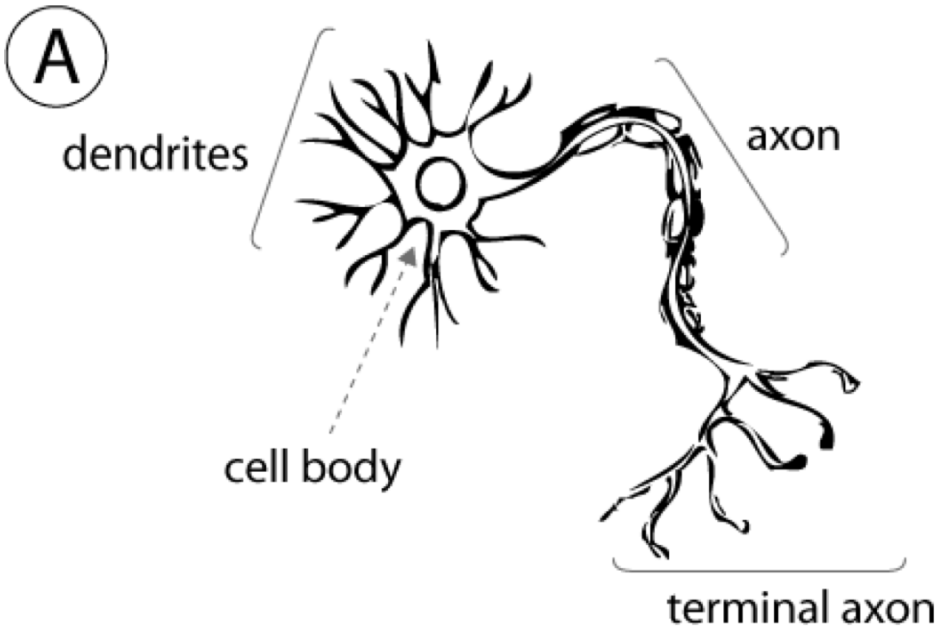
- Information is transmitted as a series of electric impulses, so-called **spikes**.
- The **frequency** and **phase** of these spikes encodes the information.
- In biological systems, one neuron can be connected to as many as **10,000** other neurons.
- Usually, a neuron receives its information from other neurons in a confined area, it's so-called **receptive field**.

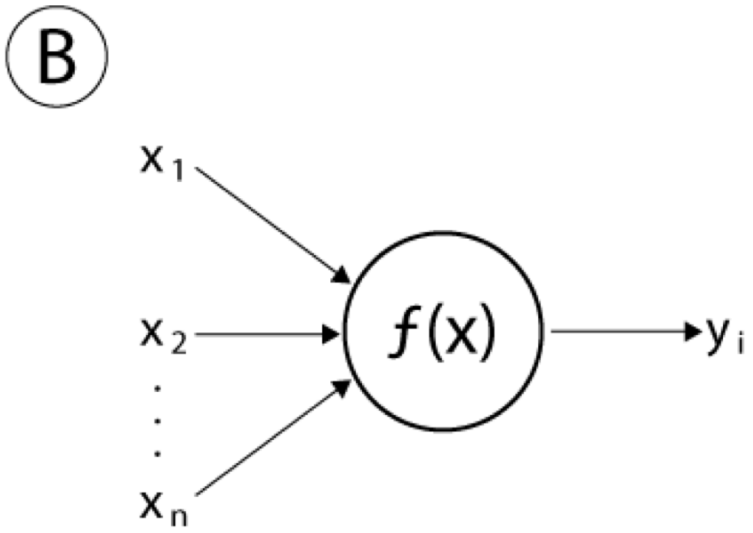
From Real to Artificial Neurons

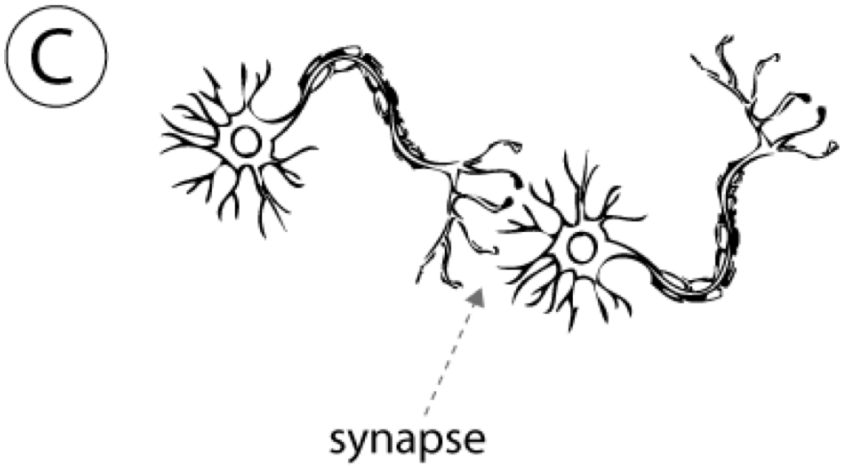


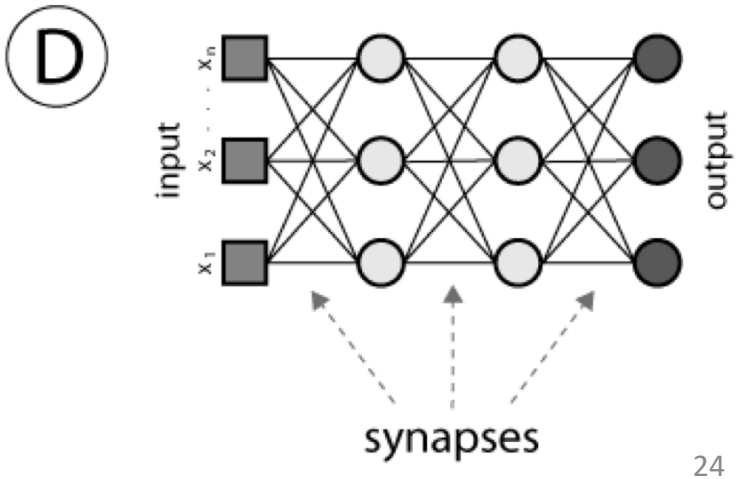


From Real NNs to Artificial NNs





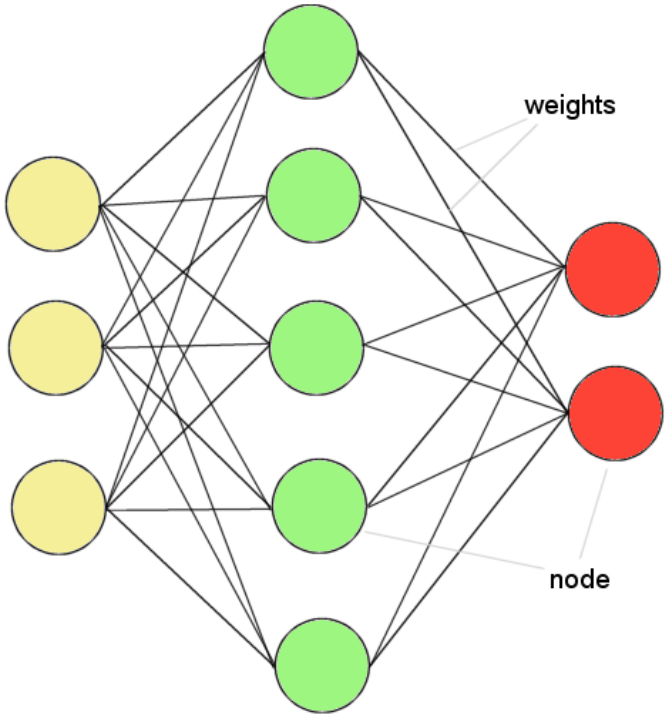




ANNs – The Basics

- ANNs incorporate the two fundamental components of biological neural nets:

1. Neurons (nodes)
2. Synapses (weights)



History of Artificial Neural Networks

1943 McCulloch & Pitts propose the first mathematical model for biological neurons

1949 Hebb proposes his learning rule: Repeated activation of one neuron by another strengthens their connection

1958 Rosenblatt invents the perceptron by basically adding a learning algorithm to the McCulloch & Pitts model

History of Artificial Neural Networks

1960 Widrow & Hoff introduce the ADALINE, a simple network trained through gradient descent

1961 Rosenblatt proposes a scheme for training multilayer networks, but his algorithm is weak because of non-differentiable node functions

1969 Minsky & Papert show that perceptrons are not computationally universal; interest in neural network research decreases

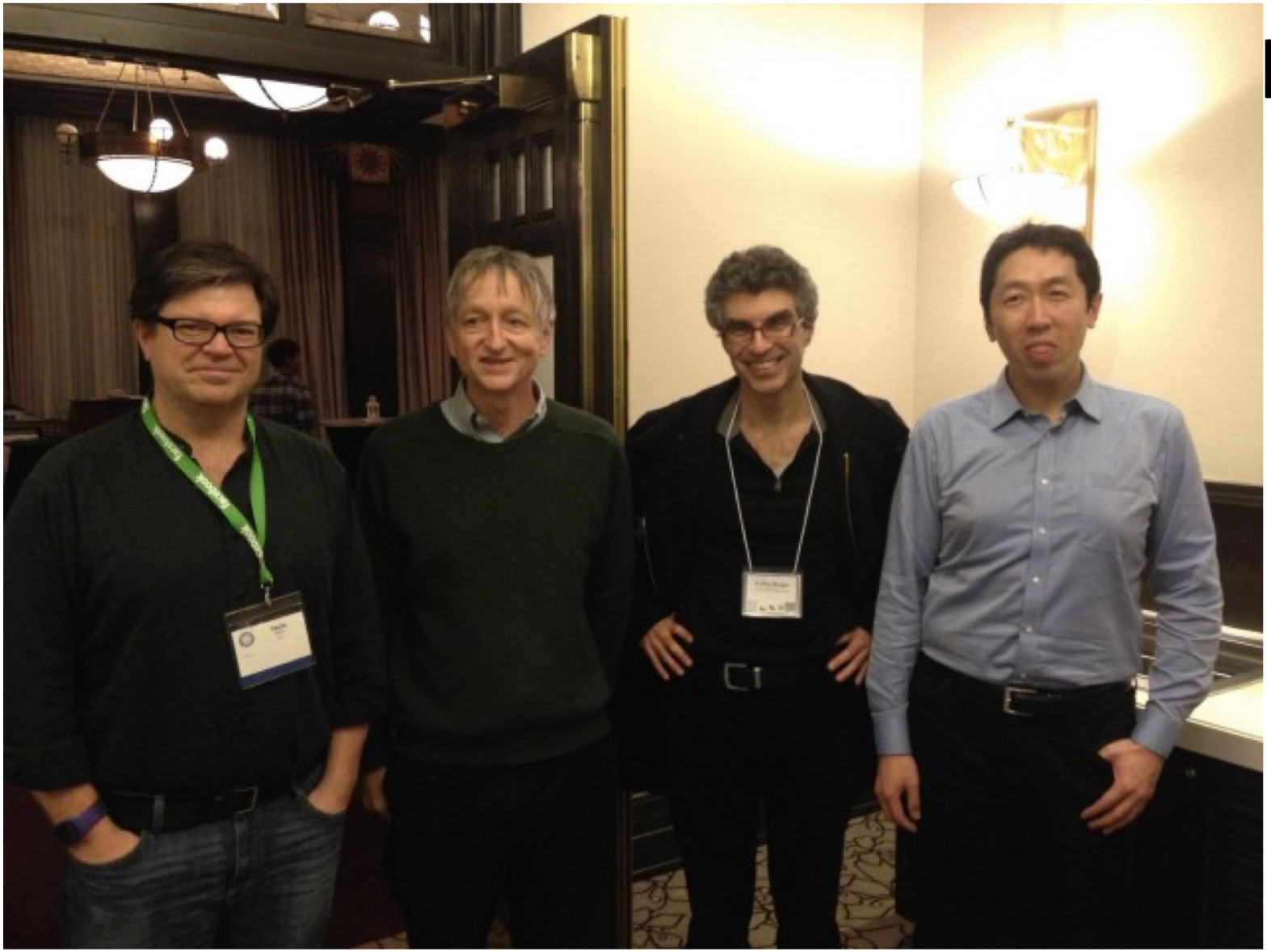
History of Artificial Neural Networks

1982	Hopfield develops his auto-association network
1982	Kohonen proposes the self-organizing map
1985	Ackley, Hinton & Sejnowski devise a stochastic network named Boltzmann machine
1986	Rumelhart, Hinton & Williams provide the backpropagation algorithm in its modern form, triggering new interest in the field

History of Artificial Neural Networks

Since then, research on artificial neural networks has remained active, leading to numerous new network types and variants, as well as hybrid algorithms and hardware for neural information processing.

ks



LeCun, Hinton, Bengio, Ng