Networked Models

School of Electronic and Computer Engineering
Peking University

Wang Wenmin

12. Models in Machine Learning

Contents:

- □ 12.1. Probabilistic Models
- □ 12.2. Geometric Models
- □ 12.3. Logical Models
- □ 12.4. Networked Models

What are Networked Models 什么是网络化模型

□ The networked models here refer to as the models of artificial neural network (ANN).

这里的网络化模型指的是**人工神经网络模型 (ANN)**。

□ An ANN is an artificial representation of the human brain that tries to simulate its learning processing.

一个ANN是**人脑的一种人工表征**，试图模拟人类的学习过程。

□ ANN can be constructed a system by interconnected “neurons” which send messages to each other.

ANN可以通过互联的“神经元”构建一个系统，神经元之间相互发送消息。

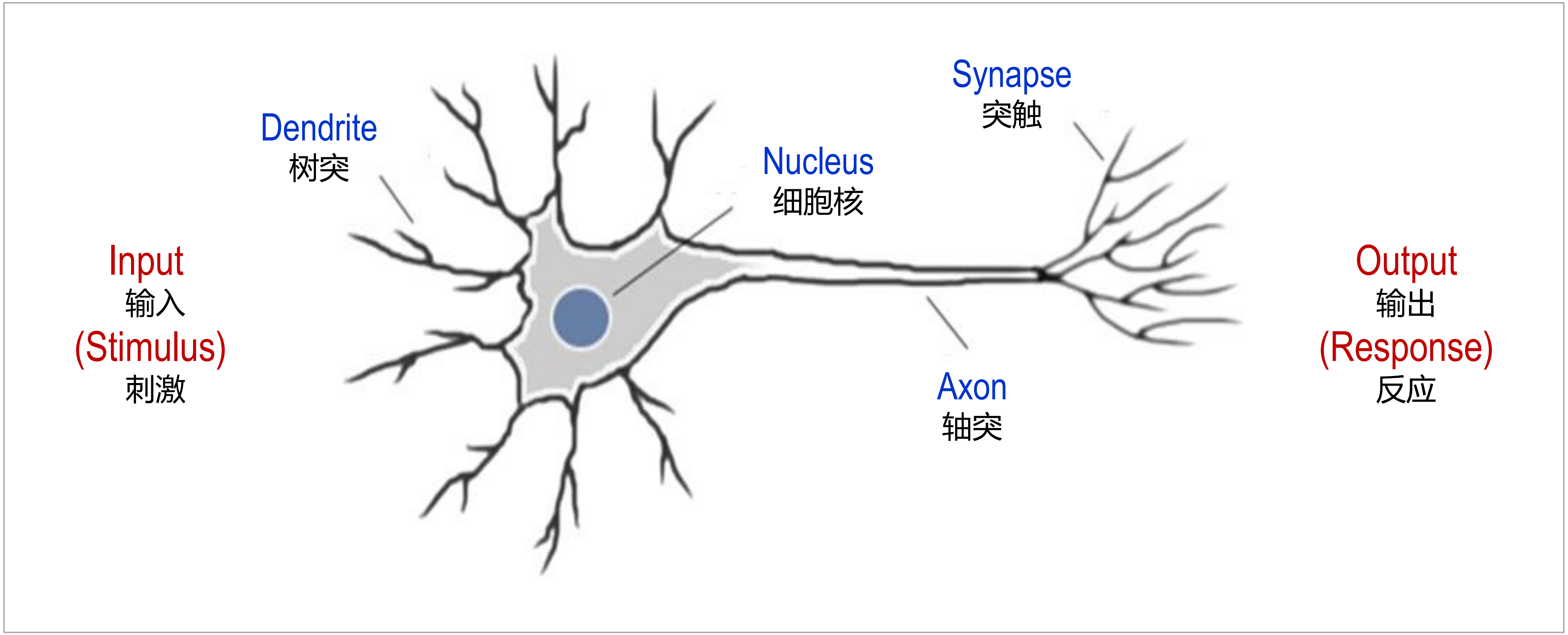
□ The connections between neurons have numeric weights that can be tuned based on experience, making ANN adaptive to inputs and capable of learning.

神经元之间的连接具有数值权重，可以通过经验调整，使ANN适应输入并且能够学习。

Contents:

- □ 12.4.1. Artificial Neural Networks
- □ 12.4.2. Deep Neural Networks

Biological Neuron 生物神经元



Input
输入
(Stimulus)
刺激

Dendrite
树突

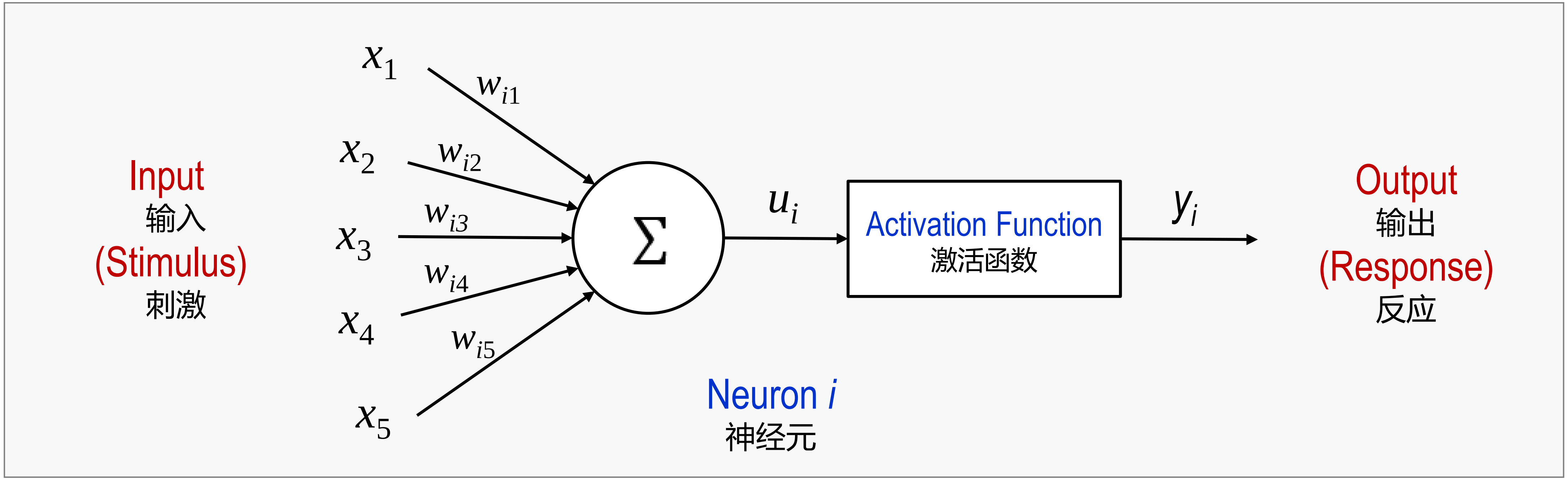
Nucleus
细胞核

Synapse
突触

Axon
轴突

Output
输出
(Response)
反应

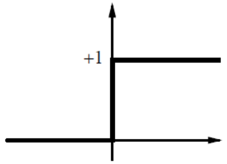
Artificial Neuron 人工神经元

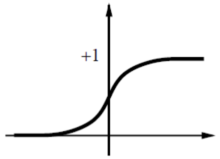


$$u_i = \sum_j w_{ij} \cdot x_j$$

$$y_i = f(u_i) = \begin{cases} +1, \text{ if } u_i \geq \omega_0 \\ -1, \text{ if } u_i < \omega_0 \end{cases} \quad \text{or} \quad \frac{1}{1 + e^{-u_i}}$$

Activation function:
激活函数

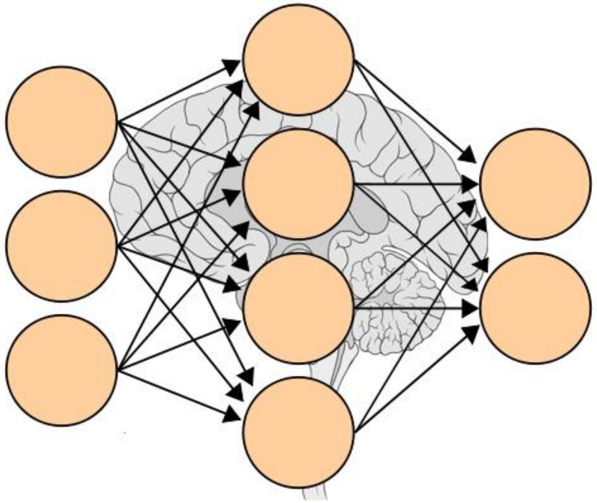




Artificial Neural Network (ANN) 人工神经网络

Input
输入
(Stimulus)
刺激

Input layers
输入层



Output layers
输出层

Output
输出
(Response)
反应

Hidden layers
隐藏层

ANN is a family of learning models inspired by biological neural networks ■ ANN是受生物神经网络启发的一系列学习模型

The interconnection between the different layers of neurons ■ 不同的神经元层次之间互联

The learning process for updating the weights of the interconnections ■ 学习过程是为了更新互联权重

The activation function that converts a neuron's weighted input to its output ■ 激活函数将神经元的加权输入转换为其输出

History of Artificial Neural Networks 人工神经网络的发展史

1943, McCulloch and Pitts
马卡洛和匹茨

- created a **computational model for neural networks** based on mathematics and algorithms called threshold logic.
  基于称之为阈值逻辑的数学和算法创建了**神经网络的计算模型**。

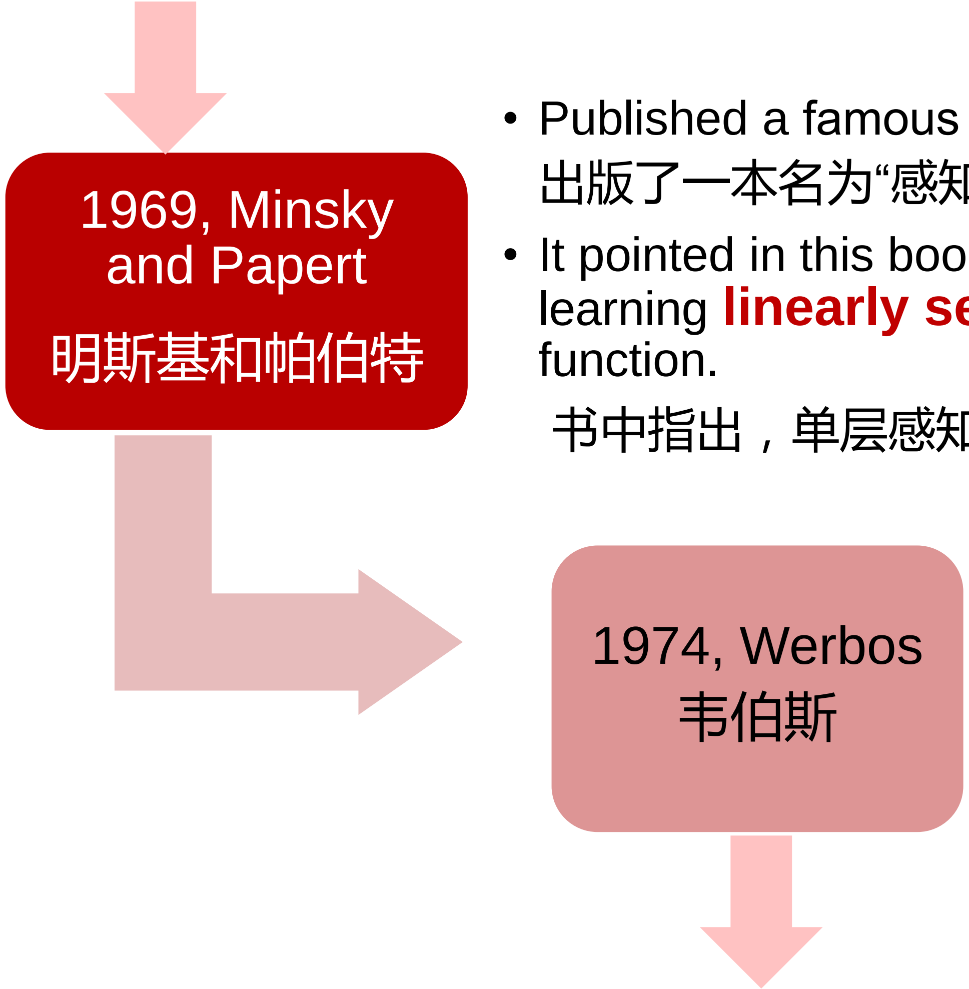
1954, Farley and Clark
法利和克拉克

- first used computational machines, then called calculators, to simulate a **Hebbian network**.
  首次利用计算的机器、后来称其为计算器，来仿真**赫布网络**。

1958, Rosenblatt
罗森布莱特

- created **perceptron**, an algorithm for pattern recognition, which is with only one output layer, so also called “single layer perceptron”.
  创建了**感知机**，一种模式识别算法，它仅有一个输出层，也被称为“单层感知机”。

History of Artificial Neural Networks 人工神经网络的发展史



1969, Minsky and Papert

明斯基和帕伯特

- Published a famous book entitled “Perceptrons”.
  出版了一本名为“感知机”的著名书籍。
- It pointed in this book that the single layer perceptrons are only capable of learning **linearly separable patterns**, but not possible to learn an XOR function.

  书中指出，单层感知机仅能学习**线性可分模式**，而不能用于学习异或功能。

1974, Werbos

韦伯斯

- Proposed the **back-propagation** algorithm, a method for training ANNs and used in conjunction with an optimization method such as gradient descent.

  提出了**反向传播算法**，一种用于训练ANNs的方法，并且与梯度下降等优化方法结合使用。
- Regenerates interest in the 1980s.

  1980年代才引起重视。

History of Artificial Neural Networks 人工神经网络的发展史

1989, Yann LeCun et al
雅恩·勒昆等人

- Published LeNet-5, a **pioneering 7-level convolutional neural network** (CNN) is applied to recognize hand-written numbers on checks.

  发表了LeNet-5，一种**开拓性的7层卷积神经网络** (CNN)，用于检查支票上的手写数字。

1992, Schmidhuber
施米德胡贝

- Proposed **recurrent neural network** (RNN), this creates an internal state which allows it to exhibit dynamic temporal behavior.

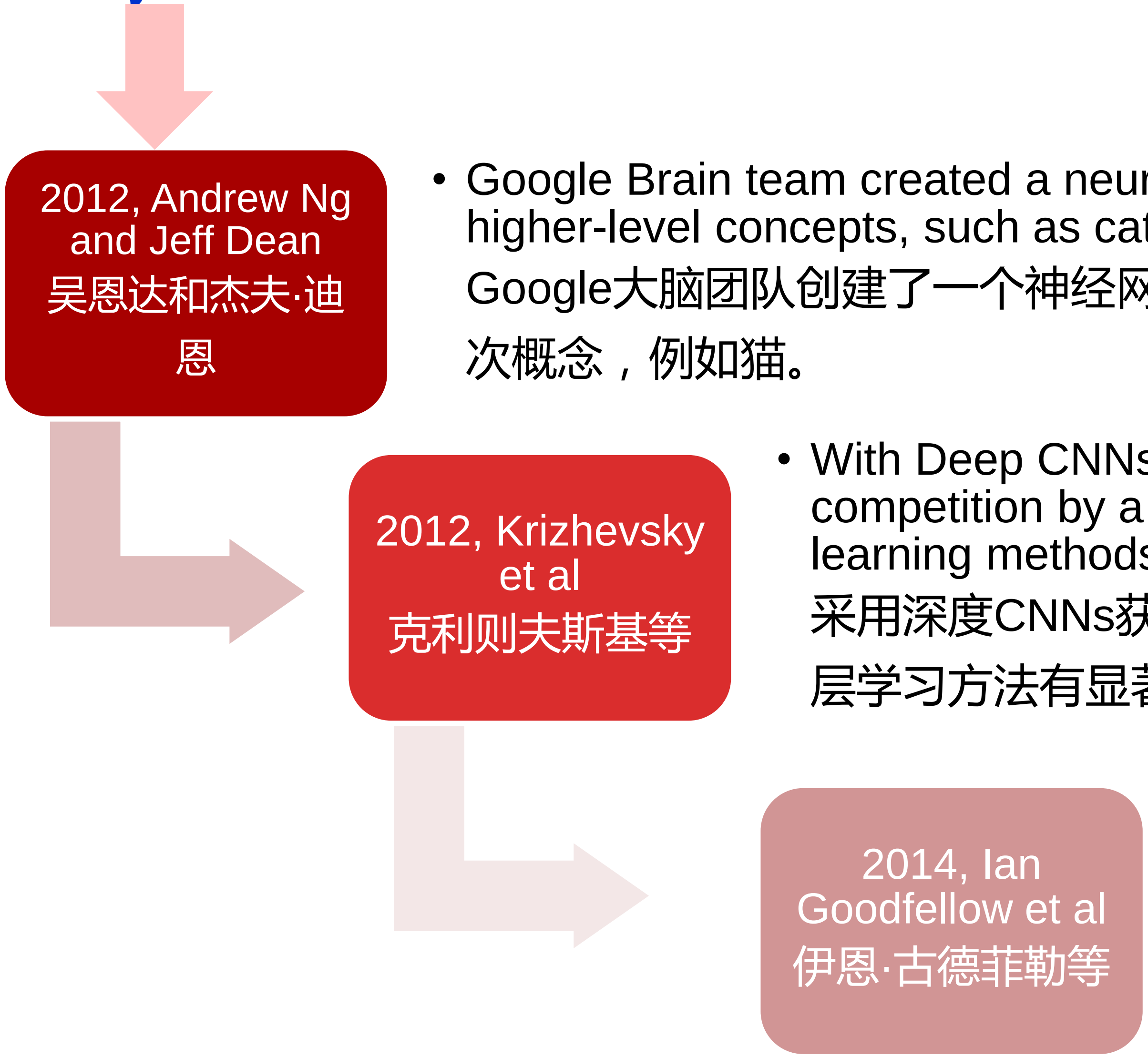
  提出了**循环神经网络**，它创建网络的内部状态，得以展现动态时间行为。

2006, Hinton and Salakhutdinov
辛顿和萨拉赫丁诺夫

- Renewed interest in **neural nets** was sparked by the advent of deep learning.

  深度学习的出现，再次引发了对**神经网络**的兴趣。

History of Artificial Neural Networks 人工神经网络的发展史



2012, Andrew Ng and Jeff Dean
吴恩达和杰夫·迪恩

- Google Brain team created a neural network that learned to recognize higher-level concepts, such as cats, from watching unlabeled images.
  Google大脑团队创建了一个神经网络，学会观看未标注图像来识别高层次概念，例如猫。

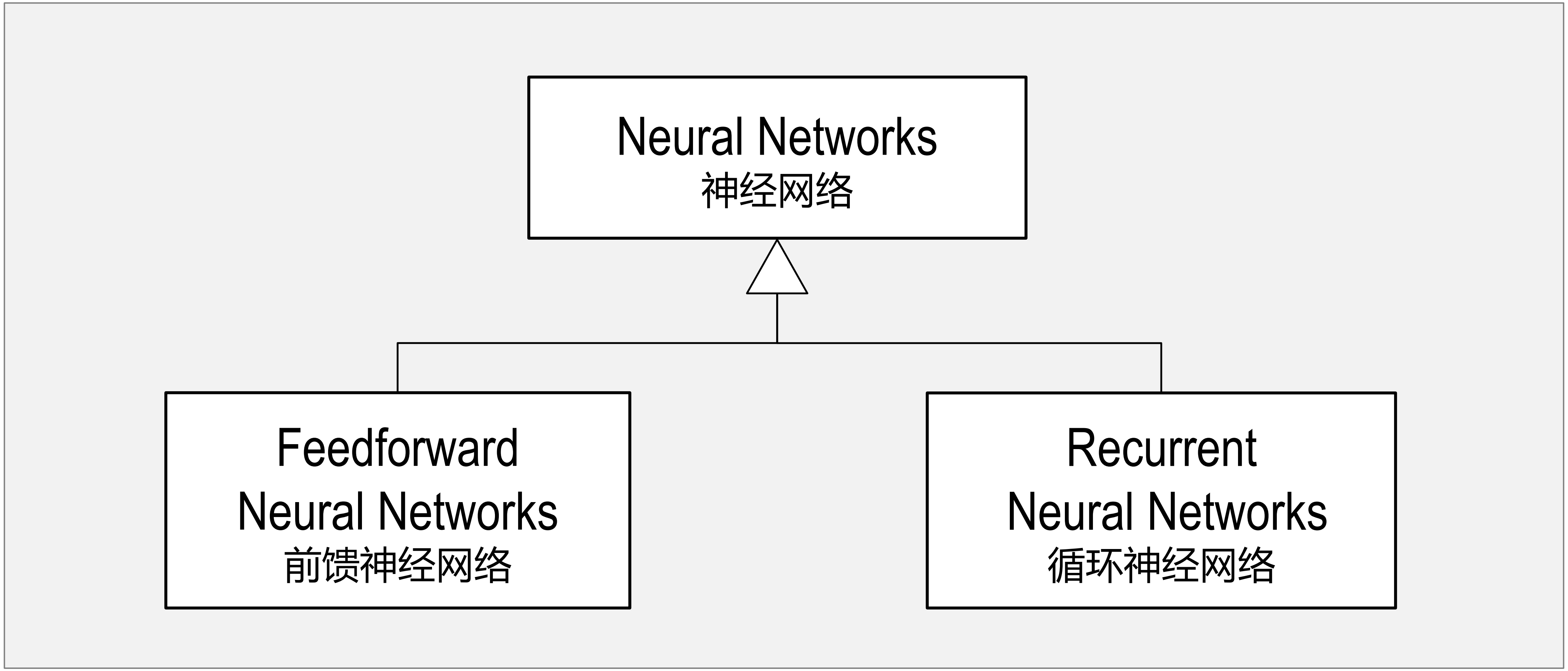
2012, Krizhevsky et al
克利则夫斯基等

- With Deep CNNs won the large-scale ImageNet competition by a significant margin over shallow machine learning methods.
  采用深度CNNs获得了大规模ImageNet比赛的胜利，比浅层学习方法有显著优势。

2014, Ian Goodfellow et al
伊恩·古德菲勒等

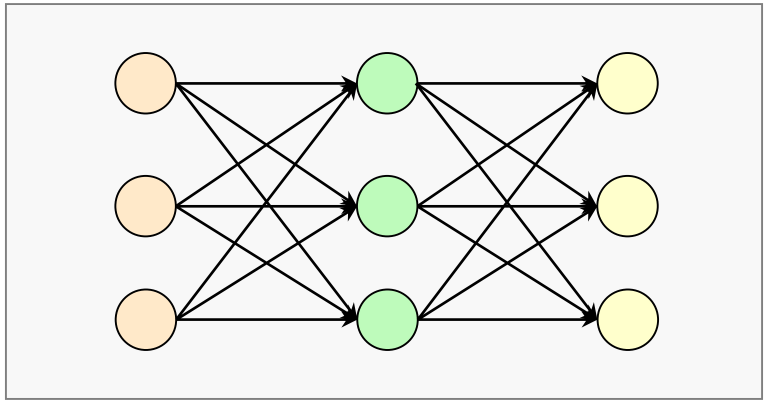
- Proposed **generative adversarial network** (GAN) which has two neural networks competing against each other in a zero-sum game framework.
  提出了**生成对抗网络** (GAN)，其中有两个神经网络，彼此以“零和”博弈方式相互竞争。

Structures of Neural Networks 神经网络的结构



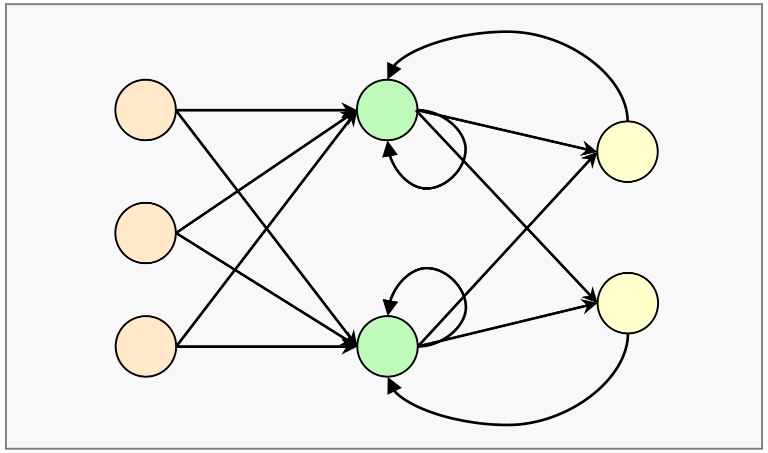
Structures of Neural Network Models 神经网络模型的结构

- □ Feedforward neural network 前馈神经网络
  - ■ information moves in only one direction, forward, from input nodes, through hidden nodes and to the output nodes.
    信息从输入结点仅仅以一个方向，即前进方向，穿过隐藏层并抵达输出节点。



Feedforward neural network 前馈神经网络

- □ Recurrent neural network 循环神经网络
  - ■ connections form a directed cycle.
    连接形成有向循环。
  - ■ creating an internal state of the network which allows it to exhibit dynamic temporal behavior.
    建立网络的内部状态，使之展现动态的时间特性。



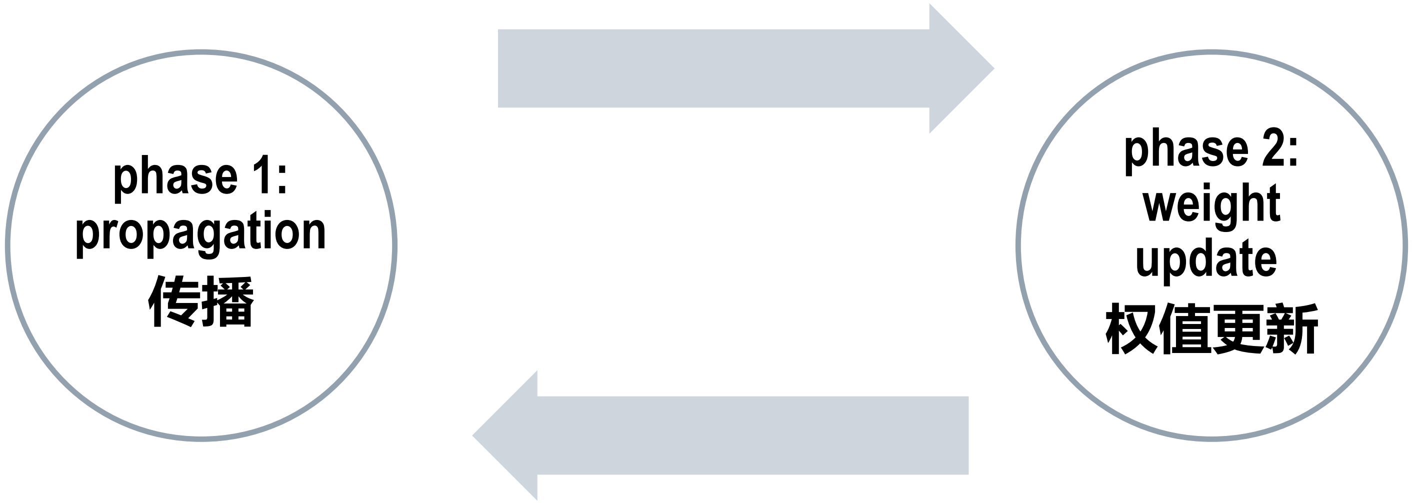
Recurrent neural network 循环神经网络

Back-propagation 反向传播

□ Back-propagation (BP) is an abbreviation for "**Backward propagation of errors**".
反向传播 (BP) 是 "**反向误差传播**" 的缩略语。

□ It is a common method of training Artificial Neural Networks, and used in conjunction with an optimization method such as gradient descent.
是训练人工神经网络的常用方法，与梯度下降优化方法结合使用。

□ The algorithm repeats a two phase cycle: 该算法重复两个阶段的循环：



Repeat phase 1 and phase 2 until the performance of the network is satisfactory.
重复阶段1和阶段2的操作，直到网络的性能得到满足。

Algorithm of Back-propagation 反向传播算法

□ Phase 1: **Propagation** 第1阶段：传播

- **Feedforward propagation** 前馈传播

  the input of training data through the neural network in order to generate *output* **activations**.

  输入的训练数据穿过神经网络，从而生成输出**激活值**。

- **Back-propagation** 反向传播

  the output activations through the neural network using the training data target in order to generate the **deltas** of all output and hidden neurons.

  输出激活再使用训练数据目标穿过神经网络，生成所有的输出层和隐藏层神经元的差值。

  - ➢ deltas = expected output - actual output values

    差值 = 期待输出 - 实际输出

Algorithm of Back-propagation 反向传播算法

□ Phase 2: **Weight update** 第2阶段：权值更新

For each weight: 对每个权值：

- Multiply its output delta and input activation, to get the gradient of the weight.
  将其输出差值与输入激活相乘，以便得到该权值梯度。
- Subtract a ***ratio*** (percentage) of the gradient from the weight. The ratio is called *learning rate*.
  从权值中减去梯度的比值（百分比）。该比值被称为学习率。
  - The greater the ratio, the faster the neuron trains;
    比值越大，神经元训练越快；
  - the lower the ratio, the more accurate the training is.
    比值越低，训练精度越高。

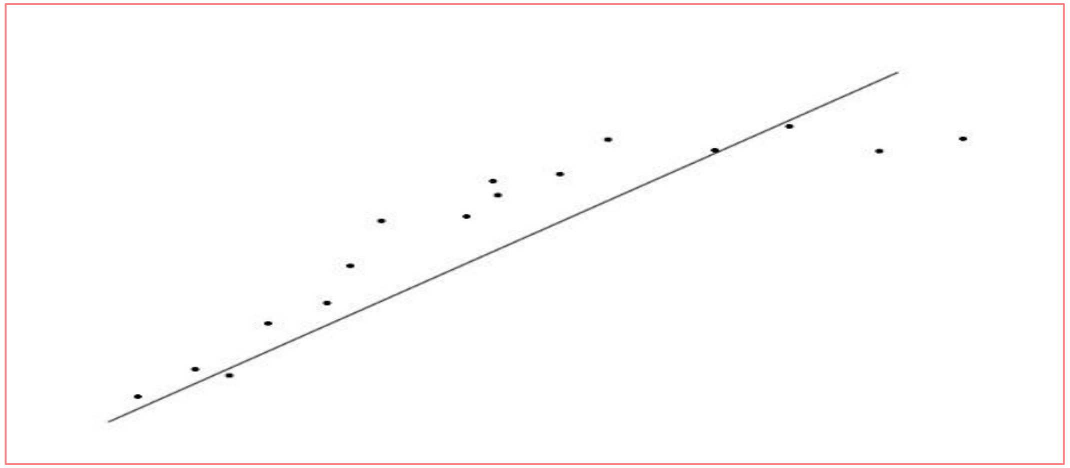
A Stochastic Gradient Descent Algorithm 随机梯度下降算法

```
function STOCHASTIC-GRADIENT-DESCENT() return the network
    initialize network weights (often small random values)
    do
        for each training example named ex
            prediction = neural-net-output(network, ex)  // forward pass
            actual = teacher-output(ex)
            compute error (prediction - actual) at the output units
            compute $\Delta w_h$ for all weights from hidden layer to output layer  // backward pass
            compute $\Delta w_i$ for all weights from input layer to hidden layer
            update network weights  // input layer not modified by error estimate
    until all examples classified correctly or another stopping criterion satisfied
    return the network
```

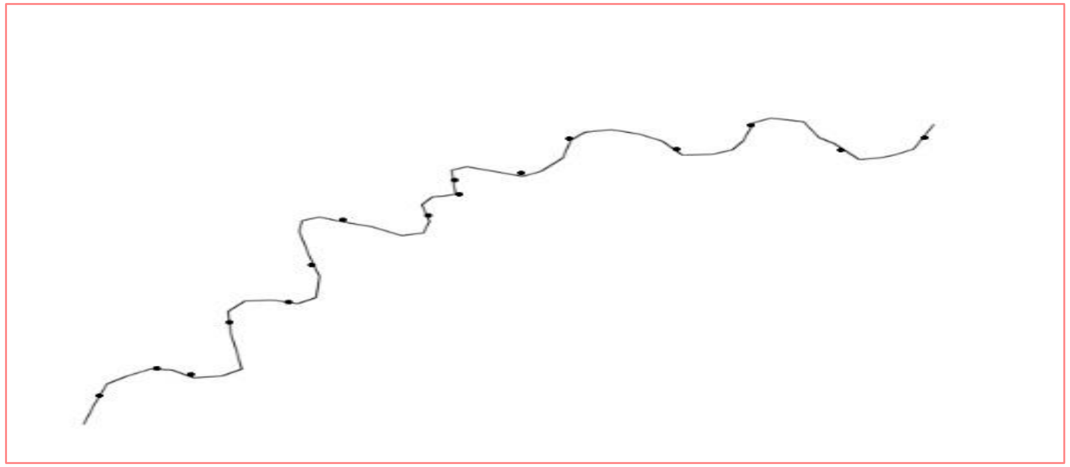
For training a three-layer network (only one hidden layer)

用于训练一个三层网络（仅有一个隐藏层）

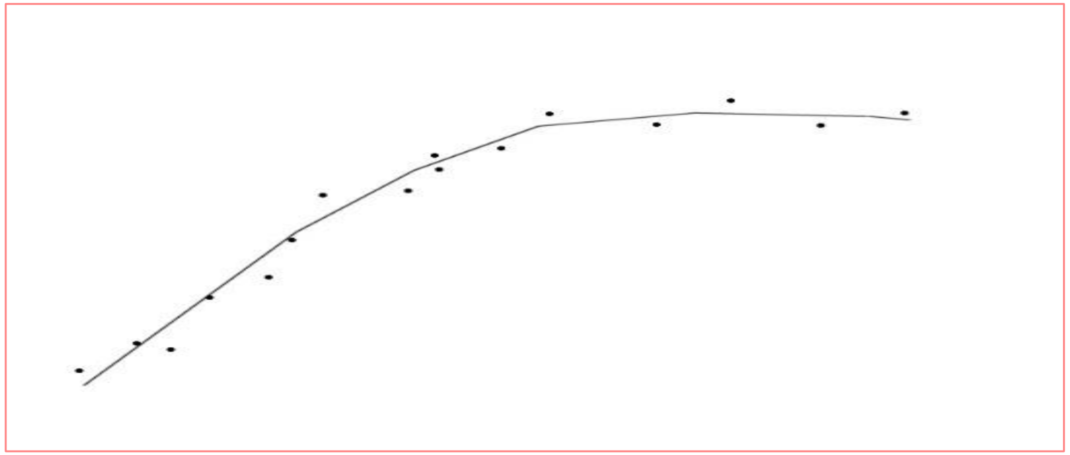
Comparison of Training Results 训练结果的比较



(a) Under-fit of the data
数据低拟合



(b) Over-fit of the data
数据过拟合

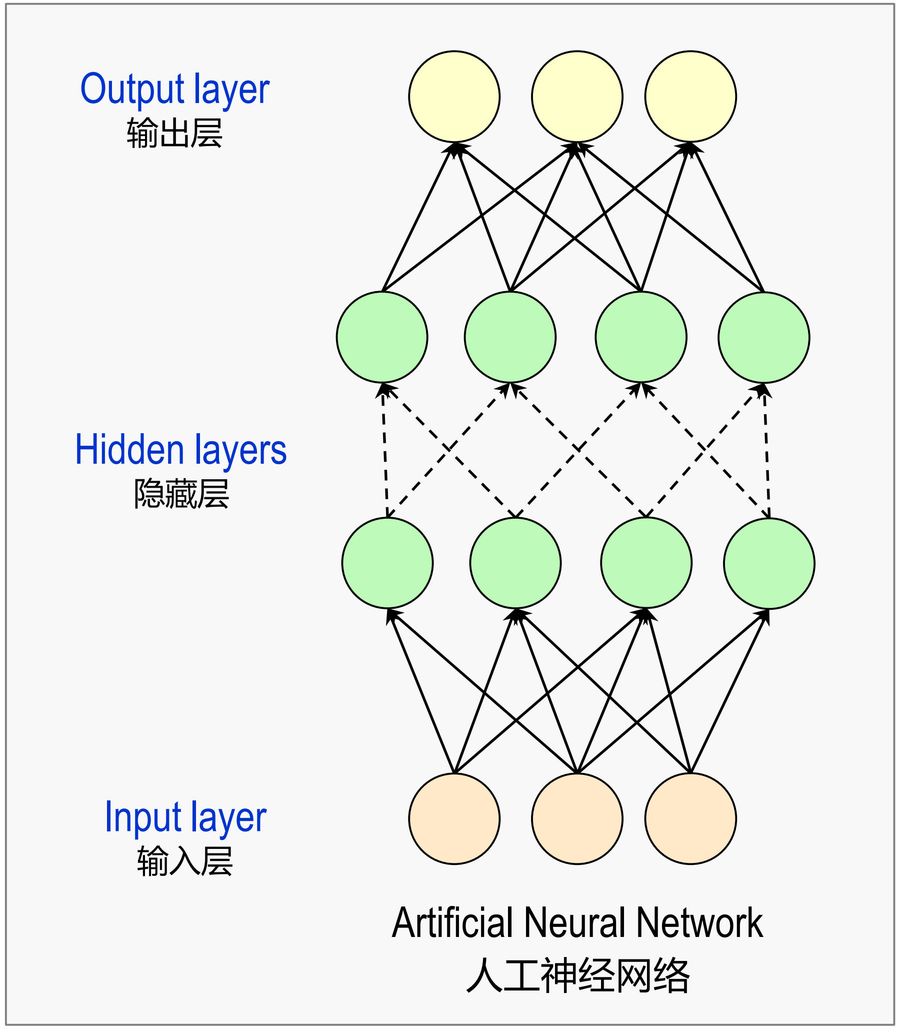


(c) Good fit of the data
数据良拟合

Shallow vs. Deep Neural Network 浅层与深层神经网络

- □ There is no universally agreed upon threshold of depth dividing shallow neural networks from deep neural networks.
就划分浅层神经网络与深层神经网络的深度而言，尚未有公认的观点。

- □ But most researchers agree that **deep neural networks have more than 2 of hidden layers, and hidden layers > 10 to be very deep neural networks**.
但大多数研究人员认为，**深度神经网络的隐藏层超过2、而隐藏层大于10的为超深度神经网络**。



Artificial Neural Network
人工神经网络

12.4. Networked Models

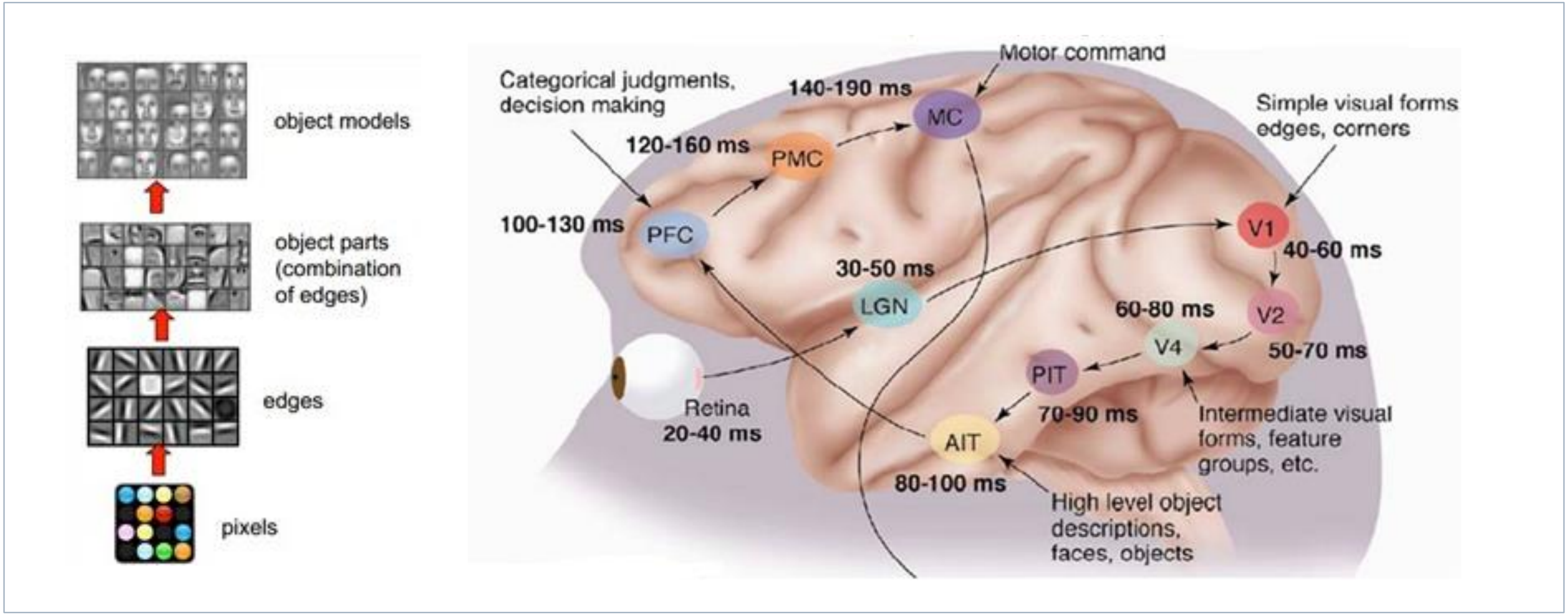
Contents:

- □ 12.4.1. Artificial Neural Networks
- □ 12.4.2. Deep Neural Networks

Why Deep Hierarchy 为什么深度层次

- □ Biological: Visual cortex is Deep Hierarchical

  生物学：视觉皮层是深层次的


object models
object parts (combination of edges)
edges
pixels
Motor command
Categorical judgments, decision making
140-190 ms
MC
120-160 ms
PMC
Simple visual forms edges, corners
100-130 ms
PFC
V1
40-60 ms
30-50 ms
LGN
60-80 ms
V2
V4
50-70 ms
PIT
Retina
20-40 ms
70-90 ms
Intermediate visual forms, feature groups, etc.
AIT
80-100 ms
High level object descriptions, faces, objects

Deep Neural Networks (DNNs) 深度神经网络

□ DNNs use many layers of nonlinear processing units for **feature extraction and transformation**.

DNNs使用许多层非线性处理单元，用于**特征提取和转换**。

□ Able to learn multiple levels of features or representations of the data. Higher level features are derived from lower level features.

能够学习数据的**多层特征或表征**。高层特征来自于低层特征。

□ Be part of the broader machine learning field: **learning representations of data**.

成为更广泛的机器学习领域的一部分：**学习数据表征**。

□ Learning multiple levels of representations that correspond to different levels of abstraction; the levels form a hierarchy of concepts.

学习多层级表征，对应于不同的抽象层级；这种层级形成了一种概念的层次结构。

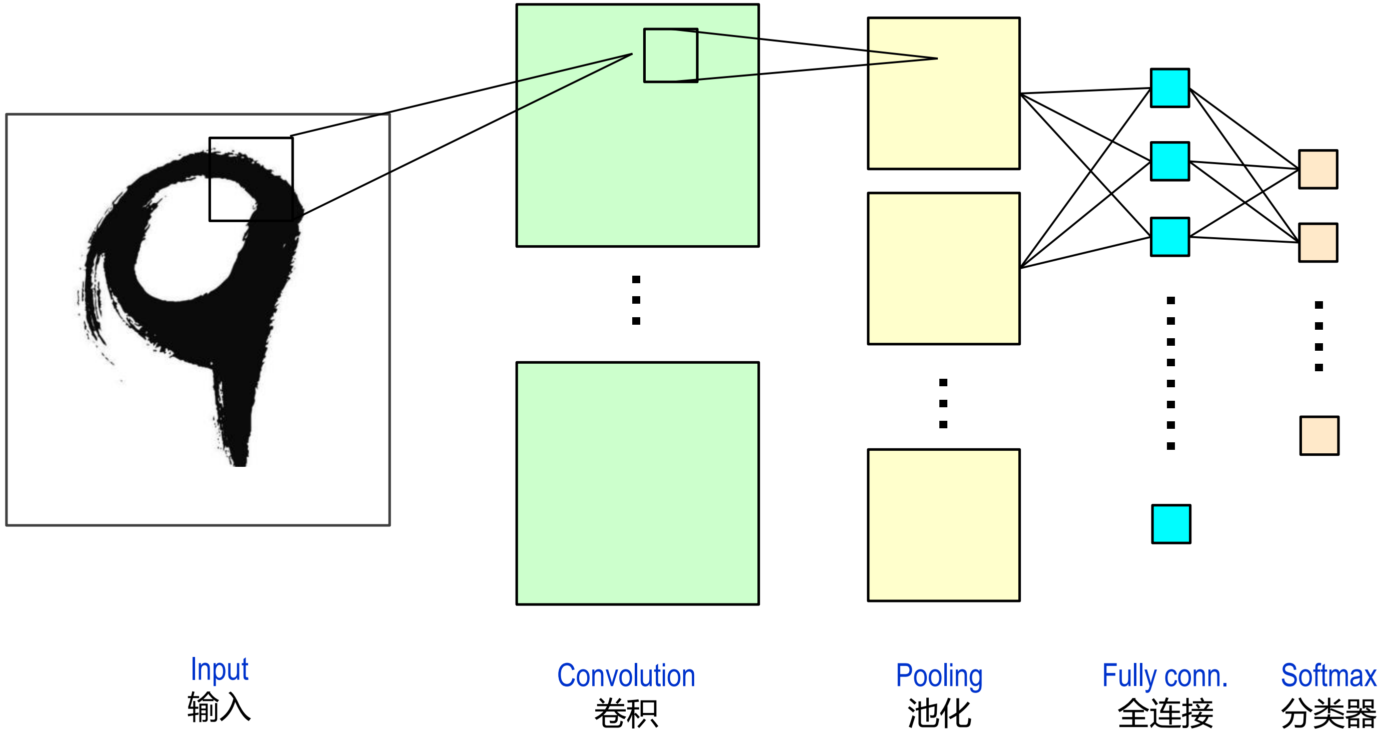
Typical Deep Neural Networks 代表性的深度神经网络

Deep belief networks (DBN)	□ 深度信念网络 (DBN)
Convolutional neural networks (CNN)	□ 卷积神经网络 (CNN)
Deep Boltzmann machines (DBM)	□ 深度波兹曼机 (DBM)
Recurrent neural networks (RNN)	□ 循环网络 (RNN)
Long short-term memory (LSTM)	□ 长短期记忆 (LSTM)
Auto-Encoders	□ 自动编码器
Generative Adversarial Network (GAN)	□ 生成对抗网络 (GAN)

Case Study: Convolutional neural network (CNN) 卷积神经网络

- □ CNN is a type of feed-forward artificial neural network that uses at least one of convolution in place of general matrix multiplication.
  CNN是一种前馈式人工神经网络，使用至少一个卷积层来代替一般的矩阵乘法。

- □ Four key ideas: 四个关键思想：
  - local connections (convolution)
    局部连接（卷积）
  - shared weights
    共享权值
  - pooling (sampling)
    池化（采样）
  - many layers.
    多层



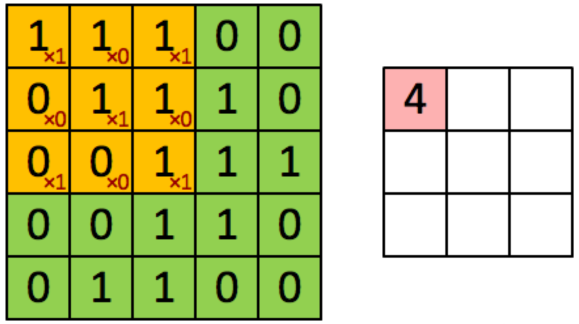
Case Study: Convolutional neural network (CNN) 卷积神经网络

□ Convolution layer 卷积层

Consist of a set of learnable filters, each filter is convolved across the width and height of the input volume, computing the *dot product* between the entries of the filter and the input, and producing a 2-dimensional activation map of that filter.

包含一组学习滤波器，每个滤波器对输入的宽和高进行卷积，计算滤波器和输入之间的点积，生成一个该滤波器的2维活动图。

Source: http://deeplearning.net/tutorial/lenet.html



$1_{\times1}$	$1_{\times0}$	$1_{\times1}$	0	0
$0_{\times0}$	$1_{\times1}$	$1_{\times0}$	1	0
$0_{\times1}$	$0_{\times0}$	$1_{\times1}$	1	1
0	0	1	1	0
0	1	1	0	0

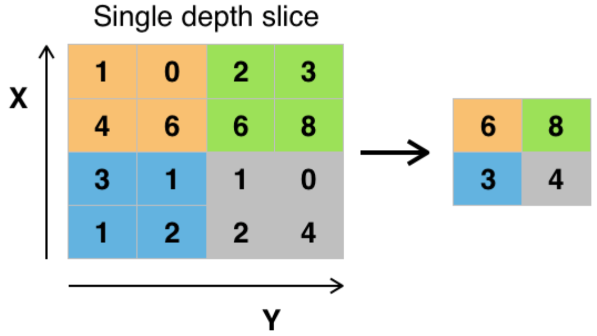
4		

Convolution with a 3x3 filter
用3x3滤波器进行卷积

□ Pooling layer 池化层

A form of non-linear down-sampling. It partitions the input image into a set of non-overlapping rectangles and, for each such sub-region, outputs the maximum.

是一种非线性下采样。它将输入图像分割成一组不重叠的矩形，对每个这样的子区域，再产出其最大值。



1	0	2	3
4	6	6	8
3	1	1	0
1	2	2	4

6	8
3	4

Max pooling with a 2x2 filter and stride = 2
用2x2滤波器进行最大池化，步长=2

Case Study: Generative Adversarial Network (GAN) 生成对抗网络

- □ GAN is pioneered by Ian Goodfellow et al at University of Montreal in 2014.

  GAN是由Goodfellow等人于2014年在蒙特利尔大学开创的。

- □ Adversarial Network, inspired by *Adversarial game.*

  对抗网络，灵感源于对抗博弈。

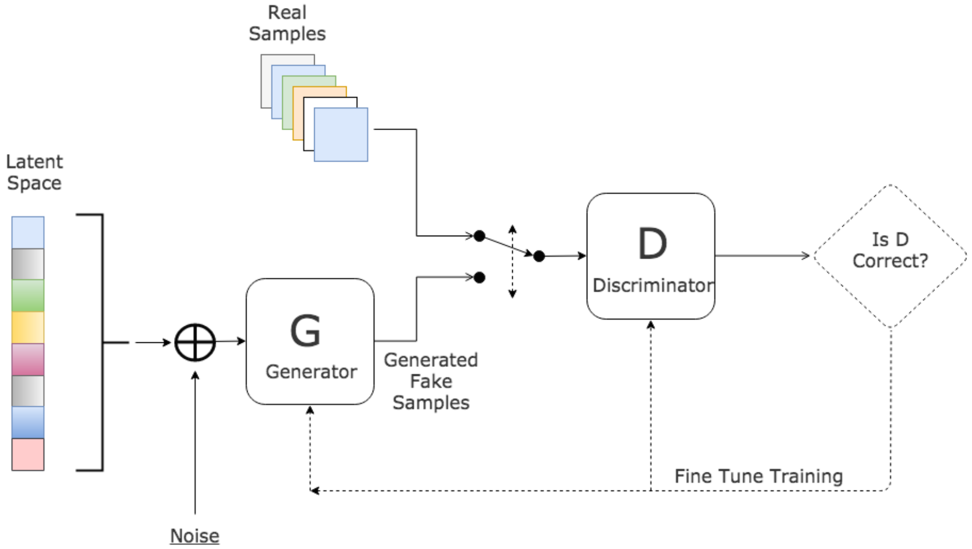
- □ Generator maps from a latent space to a particular data distribution of interest.

  生成器从潜在空间映射到所关注的特定数据分布。

- □ Discriminator discriminate between real samples and samples produced by generator.

  判别器在真实样本和生成样本之间进行判别。

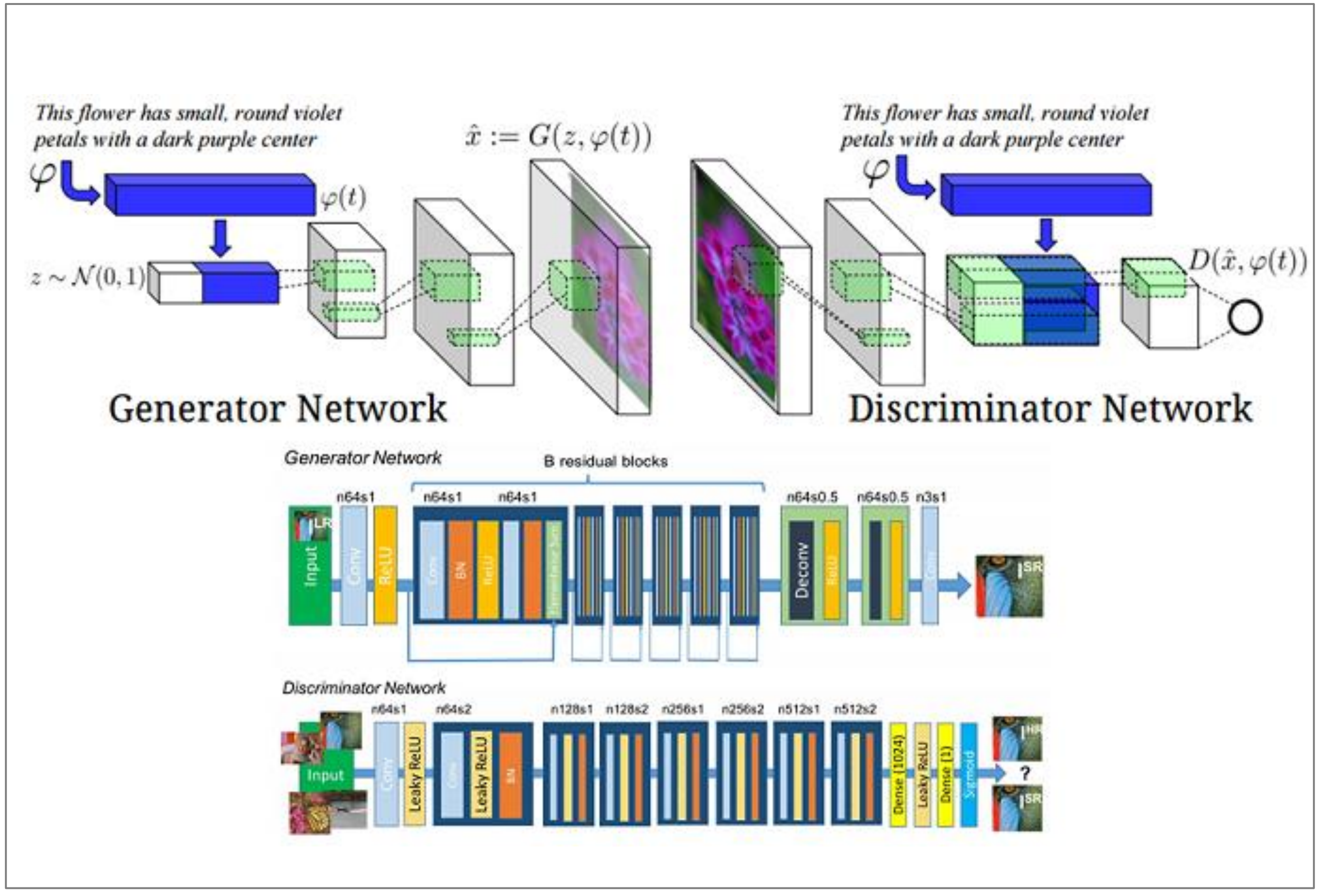
Source: Goodfellow @NIPS 2016 Workshop



Training a model in a worst-case scenario, with inputs chosen by an adversary.

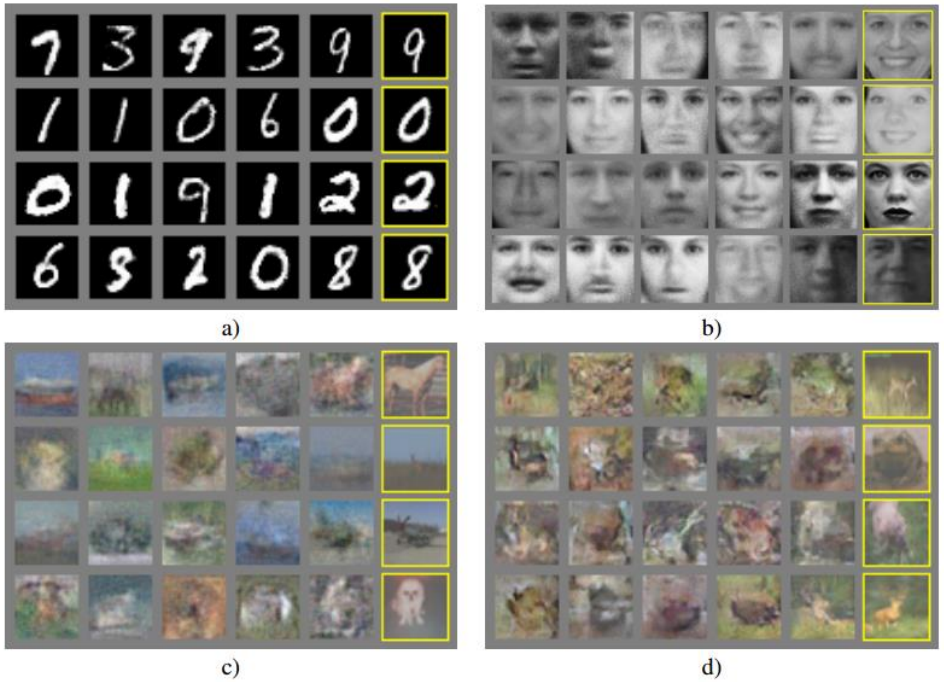
在最坏场景下训练模型，由adversary选择输入。

Case Study: Generative Adversarial Network (GAN) 生成对抗网络



Generative Adversarial Network (GAN)

生成对抗网络



The generator worked well with digits (a) and faces (b), but it created very fuzzy and vague images (c) and (d) when using the CIFAR-10 dataset.

该生成器对数字(a)和人脸(b)效果好，但使用CIFAR-10数据集时，却生成了非常模糊和含糊的图像(c)和(d)。

Typical Applications of Deep Neural Networks 深度神经网络的主要应用

Speech recognition	□	语音识别
Object recognition	□	物体识别
Image retrieval	□	图像检索
Image understanding	□	图像理解
Natural language processing	□	自然语言处理
Recommendation systems	□	推荐系统
Drug discovery	□	药物发现
Biomedical informatics	□	生物医学信息学

Postscript 后记

Deep Learning is a new area of Machine Learning research, which has been introduced with the objective of moving Machine Learning closer to one of its original goals: Artificial Intelligence.

深度学习是机器学习研究的一个新领域，其研究目的是使机器学习更接近其原始目标之一：**人工智能。**

Source: http://deeplearning.net/

Thank you for your attention!

AI