

# 《神经网络与深度学习》

## 绪论

周晓飞  
自动化学院

# 研究小组介绍

---

▶ 主要主要从事基于视觉学习与认知的智能视觉信息处理研究，包括图像和视频的视觉显著目标检测与分割研究，显著目标检测在图像/视频分割、编辑、质量评价、医疗诊断、自动驾驶、缺陷检测、遥感等领域的应用研究，以及人工智能安全。具体而言，目前我的研究方向主要分为以下几个：

(1) 基于视觉感知的物体检测：

- a) 单幅图像中有效信息的智能处理：遥感/高分辨图像显著目标检测、伪装目标检测
- b) 多幅相关图像中有效信息的智能处理：弱监督/无监督物体检测
- c) 视频中有效信息的智能处理：视频显著目标检测、视频分割

(2) 基于视觉感知特性的图像/视频质量评价：用户生成内容的质量评价

(3) 基于机器视觉的缺陷检测：钢铁表面缺陷检测

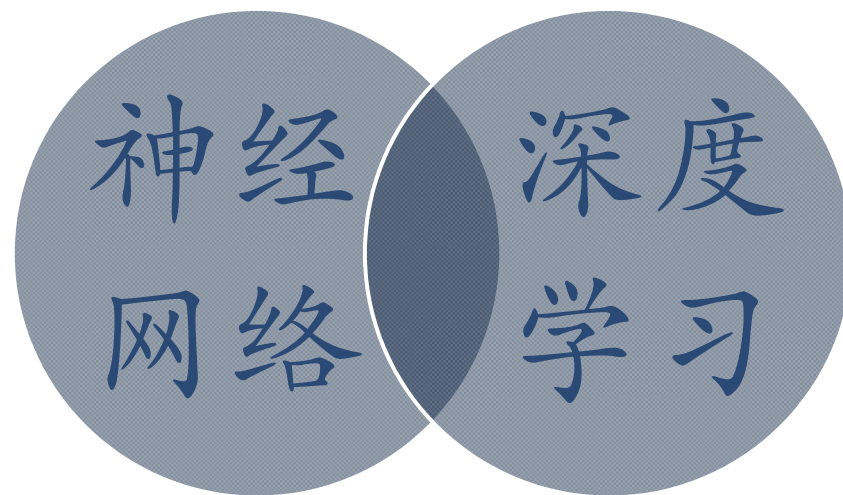
<https://zxforchid.github.io/index.html>

# 关于本课程

---

## ▶ 人工智能的一个子领域

- ▶ 神经网络：一种以（人工）神经元为基本单元的模型
- ▶ 深度学习：一类机器学习问题，主要解决贡献度分配问题（Credit Assignment Problem, CAP [Minsky,1961]）。



# 课程大纲 (32课时, 16次课)

---

## ▶ 概述

- ▶ 机器学习概述
- ▶ 线性模型

## ▶ 基础网络模型

- ▶ 前馈神经网络
- ▶ 卷积神经网络
- ▶ 循环神经网络
- ▶ 网络优化与正则化
- ▶ 记忆与注意力机制
- ▶ 无监督学习

## ▶ 进阶模型

- ▶ 概率图模型
- ▶ 玻尔兹曼机
- ▶ 深度信念网络
- ▶ 深度生成模型
- ▶ 深度强化学习
- ▶ 序列生成模型

# 预备知识

---

- ▶ 线性代数
- ▶ 微积分
- ▶ 数学优化
- ▶ 概率论
- ▶ 信息论

# 教材与推荐书籍

---

- ▶ 《神经网络与深度学习》，机械工业出版社，邱锡鹏
- ▶ 《Pattern recognition and machine learning》 CM Bishop.  
， Springer, 2006
- ▶ 《深度学习》， Ian Goodfellow等著，赵申剑、张志华  
等译， 2017
- ▶ 《神经网络与深度学习》， Michael Mielsen著， Zhu  
Xiaohui, Zhang Freeman 译， 2016

# 推荐课程

---

- ▶ 斯坦福大学CS224d: Deep Learning for Natural Language Processing
  - ▶ <http://cs224d.stanford.edu/>
  - ▶ Richard Socher 主要讲解自然语言处理领域的各种深度学习模型
- ▶ 斯坦福大学CS231n: Convolutional Neural Networks for Visual Recognition
  - ▶ <http://cs231n.stanford.edu/>
  - ▶ Fei-Fei Li Andrej Karpathy 主要讲解CNN、RNN在图像领域的应用
- ▶ 加州大学伯克利分校 CS 294: Deep Reinforcement Learning
  - ▶ <http://rail.eecs.berkeley.edu/deeprlcourse/>

# 推荐材料

---

▶ 林轩田 《机器学习基石》 《机器学习技法》

▶ <https://www.csie.ntu.edu.tw/~htlin/mooc/>

▶ 李宏毅 《1天搞懂深度学习》

▶ [http://speech.ee.ntu.edu.tw/~tlkagk/slide/Tutorial\\_HYLee\\_Deep.pptx](http://speech.ee.ntu.edu.tw/~tlkagk/slide/Tutorial_HYLee_Deep.pptx)

▶ 李宏毅 《Generative Adversarial Network (GAN)》

▶ [http://speech.ee.ntu.edu.tw/~tlkagk/slide/Tutorial\\_HYLee\\_GAN.pptx](http://speech.ee.ntu.edu.tw/~tlkagk/slide/Tutorial_HYLee_GAN.pptx)



# 顶会/顶刊论文

---

- ▶ NIPS、ICLR、ICML、AAAI、IJCAI
- ▶ ACL、EMNLP
- ▶ CVPR、ICCV
- ▶ IEEE T-IP/ T-MM/ T-CSVT .....

# 成绩评定

---

- ▶ 平时成绩

  - ▶ 40%

- ▶ 期末成绩

  - ▶ 60%

- ▶ 编程语言：Python

# 人工智能

---

- ▶ 人工智能（artificial intelligence, AI）就是让机器具有人类的智能。
  - ▶ “计算机控制” + “智能行为”
- ▶ 1950年，阿兰·图灵提出“图灵测试”
- ▶ 人工智能这个学科的诞生有着明确的标志性事件，就是1956年的达特茅斯（Dartmouth）会议。在这次会议上，“人工智能”被提出并作为本研究领域的名称。

人工智能就是要让机器的行为看起来就像是人所表现出的智能行为一样。

John McCarthy (1927-2011)

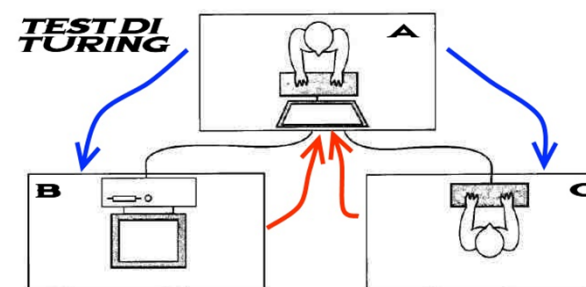
# 图灵测试

“一个人在不接触对方的情况下，通过一种特殊的方式，和对方进行一系列的问答。如果在相当长时间内，他无法根据这些问题判断对方是人还是计算机，那么就可以认为这个计算机是智能的”。

---Alan Turing [1950]  
《Computing Machinery and Intelligence》



Alan Turing



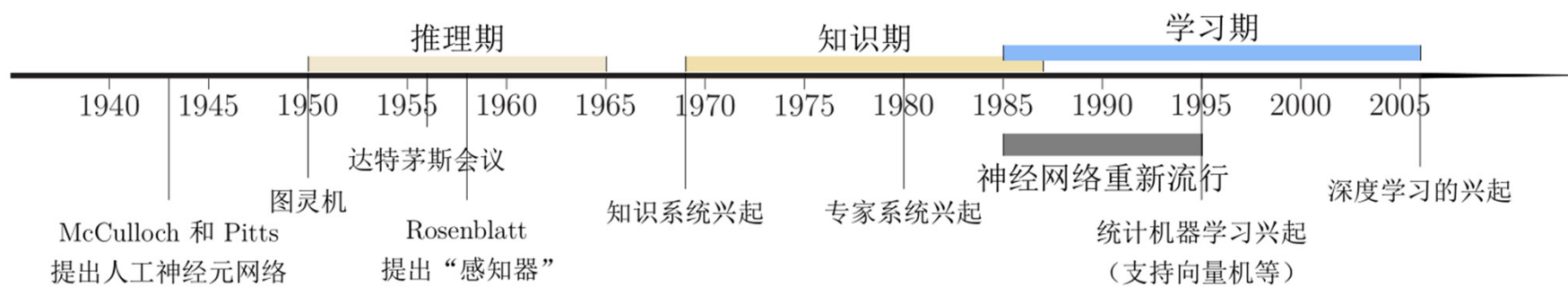
# 人工智能的研究领域

---

## ▶ 让机器具有人类的智能

- ▶ 机器感知（计算机视觉、语音信息处理）
- ▶ 学习（模式识别、机器学习、强化学习）
- ▶ 语言（自然语言处理）
- ▶ 记忆（知识表示）
- ▶ 决策（规划、数据挖掘）

# 发展历史



# 人工智能的流派

---

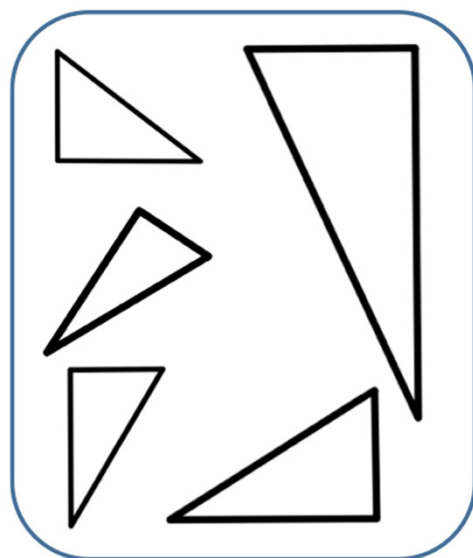
- ▶ 尚无通用理论来指导构建人工智能系统，主流方法可归结为以下两种：
  - ▶ 1) 符号主义 (symbolism) 逻辑主义、心理学派或计算机学派，盛行于人工智能的推理期和知识期；可解释
  - ▶ 2) 连接主义 (connectionism) 仿生学派或心理学派，主要结构是由大量简单的信息处理单元组成的互联网络，非线性、分布式、并行化、局部性计算以及自适应

# 如何开发一个人工智能系统?

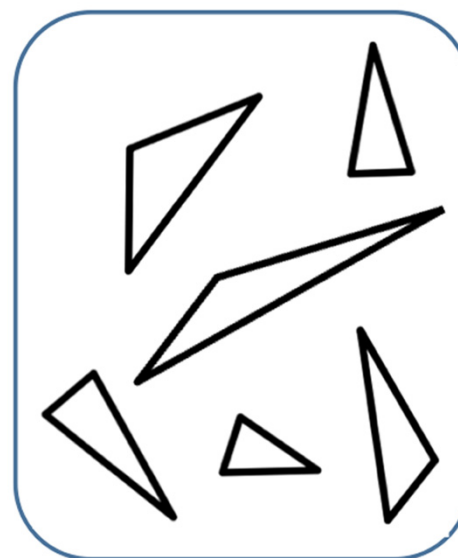
---

## ▶ 专家知识 (人工规则)

What's My Rule?



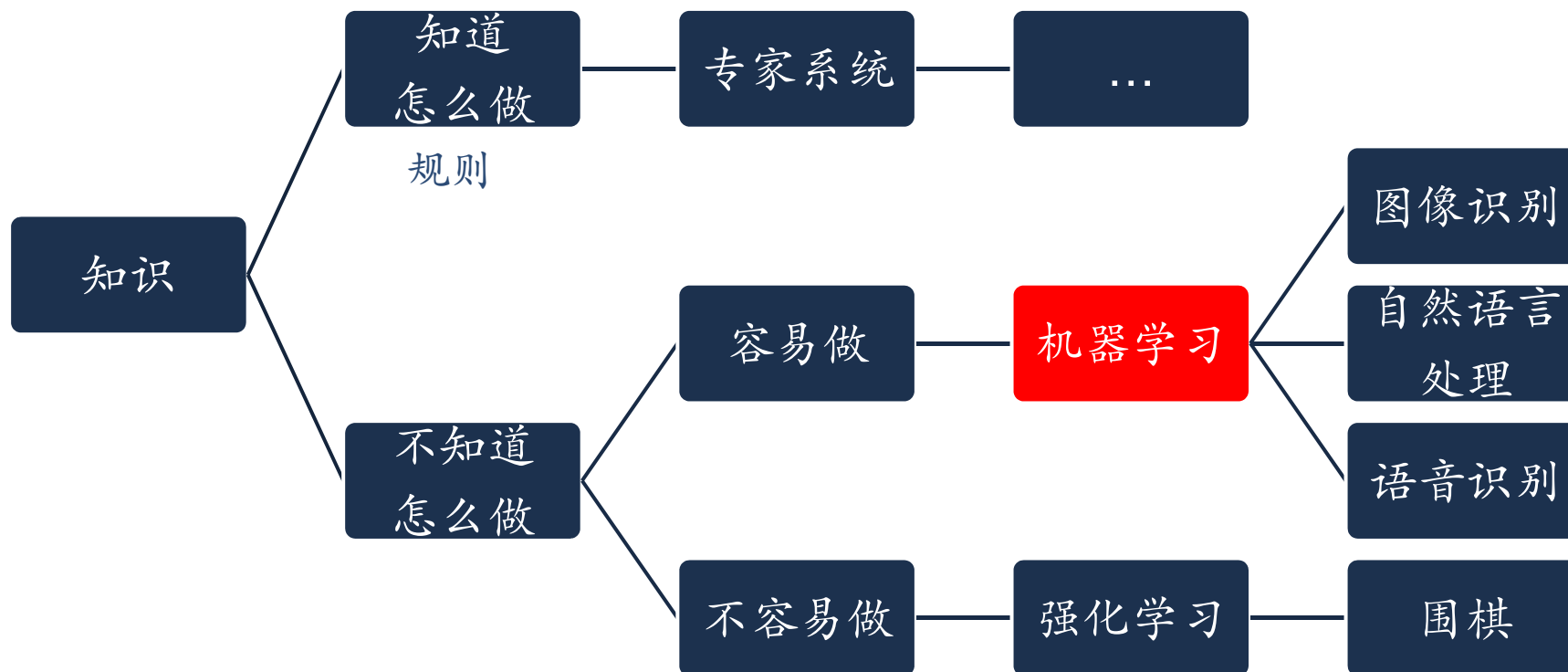
Yes/True



No/False



# 如何开发一个人工智能系统?



# What's the Rule?

---



2	6	8	9	3	4	7	5	6
3	4	7	9	5	5	6	7	2
5	8	7	0	9	4	3	5	4
5	2	3	4	9	5	6	7	8

## ▶ 机器学习

### ▶ Machine Learning ML

从有限的观测数据中学习出具有一般性的规律，并利用这些规律对未知数据进行预测的方法。

# 机器学习 $\approx$ 构建一个映射函数

---

▶ 语音识别

$$f(\text{语音波形}) = \text{“你好”}$$

▶ 图像识别

$$f(\text{数字9}) = \text{“9”}$$

▶ 围棋

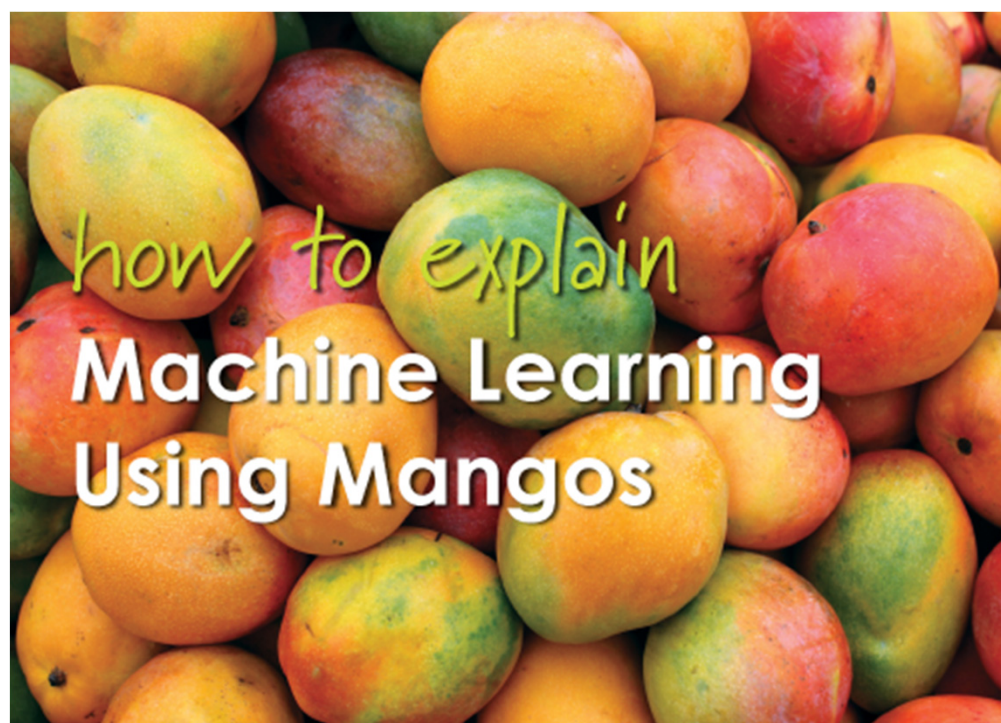
$$f(\text{围棋棋盘}) = \text{“6-5” (落子位置)}$$

▶ 机器翻译

$$f(\text{“你好!”}) = \text{“Hello!”}$$

# 芒果机器学习

如果判断芒果是否甜蜜？



<https://www.quora.com/How-do-you-explain-Machine-Learning-and-Data-Mining-to-non-Computer-Science-people>

# 芒果机器学习

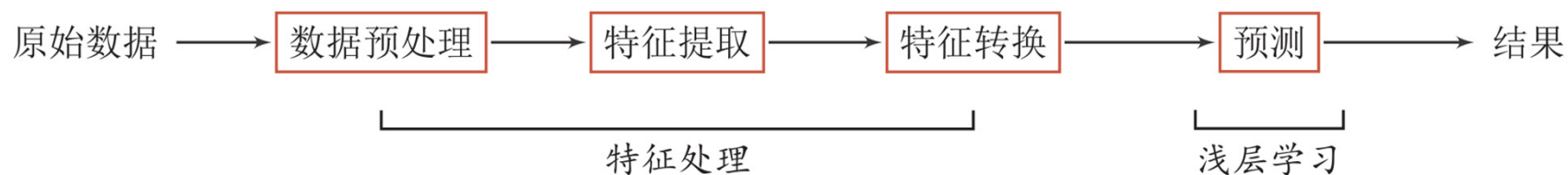
---

- ▶ 从市场上随机选取的芒果样本（**训练数据**），列出每个芒果的所有**特征**：
  - ▶ 如颜色，大小，形状，产地，品牌
- ▶ 以及芒果质量（**输出变量**）：
  - ▶ 甜蜜，多汁，成熟度。
  
- ▶ 设计一个**学习算法**来学习芒果的**特征与输出变量**之间的**相关性模型**。
  
- ▶ 下次从市场上买芒果时，可以根据芒果（**测试数据**）的特征，使用前面计算的**模型**来预测芒果的质量。

# 机器学习

---

- ▶ 当我们用机器学习来解决一些模式识别任务时，一般的流程包含以下几个步骤：



## 特征工程 (Feature Engineering)

- ▶ 浅层学习 (Shallow Learning)：不涉及特征学习，其特征主要靠人工经验或特征转换方法来抽取。

# 表示学习

Bengio, Yoshua, Aaron Courville, and Pascal Vincent.  
"Representation learning: A review and new perspectives."  
IEEE transactions on pattern analysis and machine intelligence  
35.8 (2013): 1798-1828.

---

- ▶ 数据表示是机器学习的核心问题。
  - ▶ 特征工程：需要借助人脑智能
- ▶ 表示学习 (representation learning)
  - ▶ 如何自动从数据中学习好的表示
- ▶ 难点
  - ▶ 没有明确的目标

# 语义鸿沟：人工智能的挑战之一

---

## ▶ 底层特征 VS 高层语义

- ▶ 人们对文本、图像的理解无法从字符串或者图像的底层特征直接获得



床前明月光，  
疑是地上霜。  
举头望明月，  
低头思故乡。



# 什么是好的数据表示？

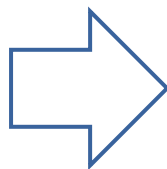
---

- ▶ “好的表示” 是一个非常主观的概念，没有一个明确的标准。
- ▶ 但一般而言，一个好的表示具有以下几个优点：
  - ▶ 应该具有很强的表示能力
  - ▶ 应该使后续的学习任务变得简单，包含更高层的语义信息
  - ▶ 应该具有一般性，是任务或领域独立的，可以迁移到其他任务上
- ▶ 机器学习任务中，经常采用两种方式来表示特征：  
局部表示(Local Representation) & 分布式表示 (Distributed Representation)

# 语义表示

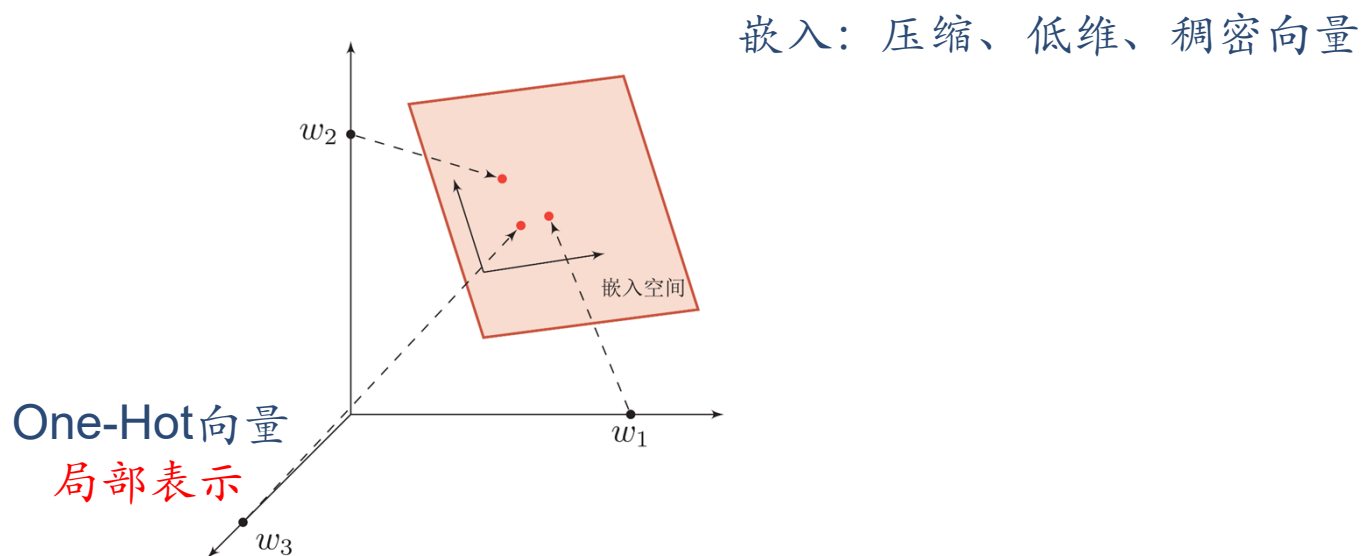
► 如何在计算机中表示语义?

符号表示



分布式表示

知识库  
规则



# 表示形式

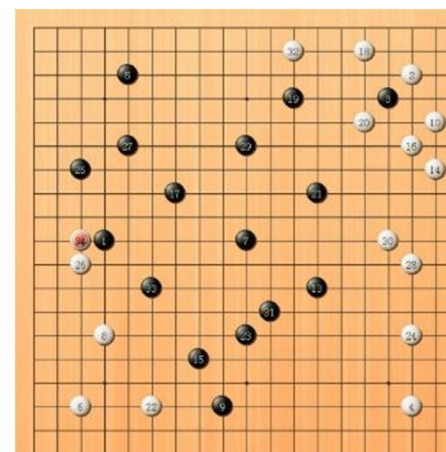
## ▶ 符号表示

- ▶ 离散表示、局部表示
- ▶ One-Hot向量

	离散表示	连续表示
A	[1 0 0 0]	[0.25 0.5]
B	[0 1 0 0]	[0.2 0.9]
C	[0 0 1 0]	[0.8 0.2]
D	[0 0 0 1]	[0.9 0.1]

## ▶ 分布式(distributed)表示

- ▶ 压缩、低维、稠密向量
- ▶ 用 $O(N)$ 个参数表示  $O(2^k)$  区间
  - ▶  $k$ 为非0参数,  $k < N$



分布式表示

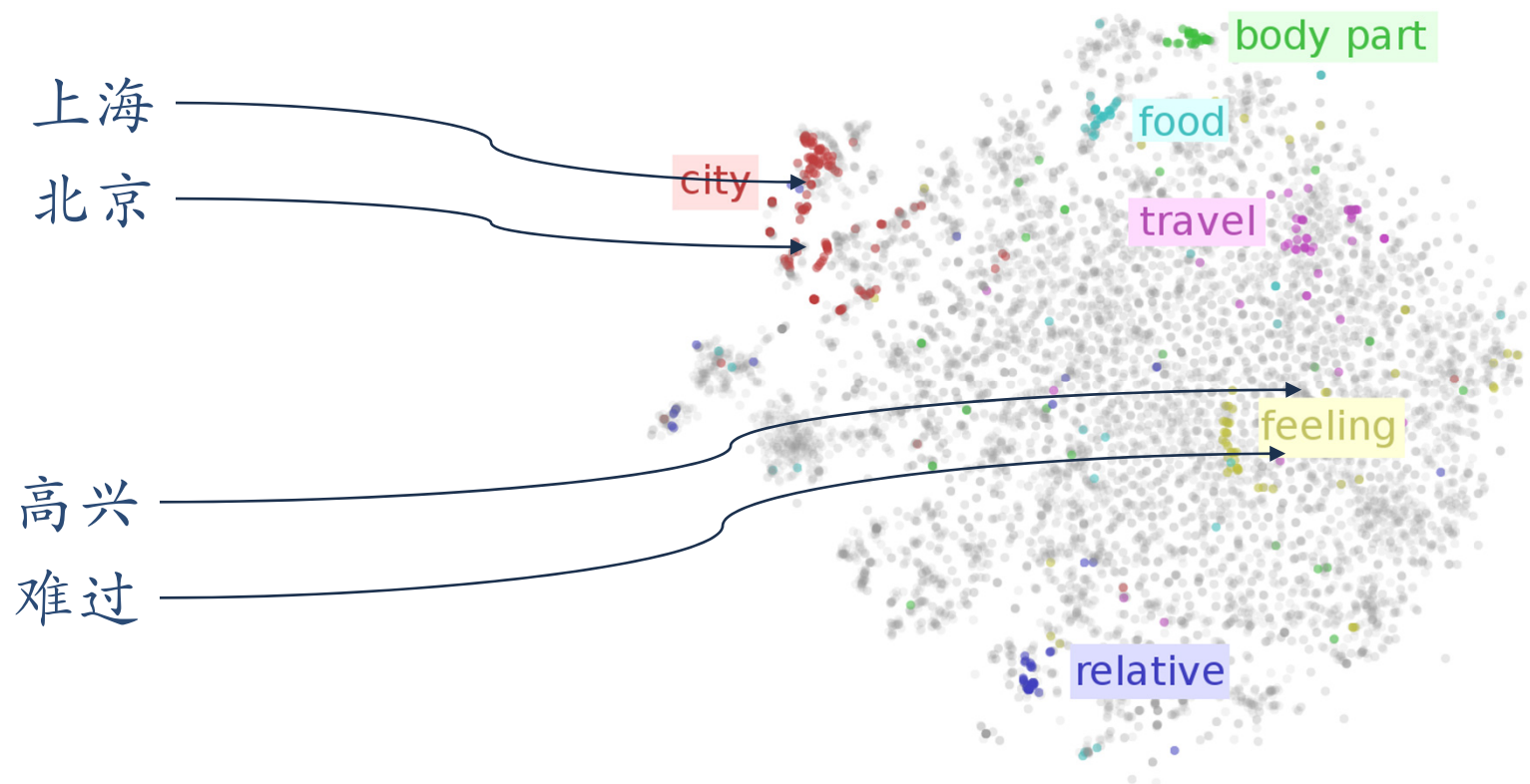
# 一个生活中的例子：颜色

颜色	局部表示	分布式表示
琥珀色	$[1, 0, 0, 0]^T$	$[1.00, 0.75, 0.00]^T$
天蓝色	$[0, 1, 0, 0]^T$	$[0.00, 0.5, 1.00]^T$
中国红	$[0, 0, 1, 0]^T$	$[0.67, 0.22, 0.12]^T$
咖啡色	$[0, 0, 0, 1]^T$	$[0.44, 0.31, 0.22]^T$



# 词嵌入 (Word Embeddings)

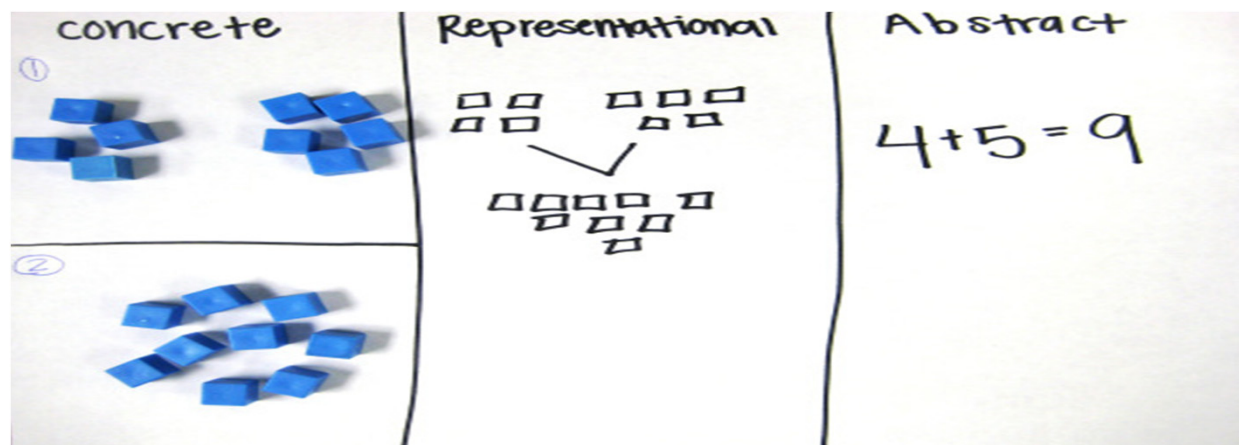
---



<https://indico.io/blog/visualizing-with-t-sne/>

# 表示学习与深度学习

- ▶ 一个好的表示学习策略必须具备一定的深度
  - ▶ 特征重用（深层结构）
    - ▶ 指数级的表示能力
  - ▶ 抽象表示与不变性
    - ▶ 抽象表示需要多步的构造



<https://mathteachingstrategies.wordpress.com/2008/11/24/concrete-and-abstract-representations-using-mathematical-tools/>

# 传统的特征提取

---

## ▶ 特征提取

### ▶ 线性投影（子空间）

▶ PCA、LDA

### ▶ 非线性嵌入

▶ LLE、Isomap、谱方法

### ▶ 自编码器

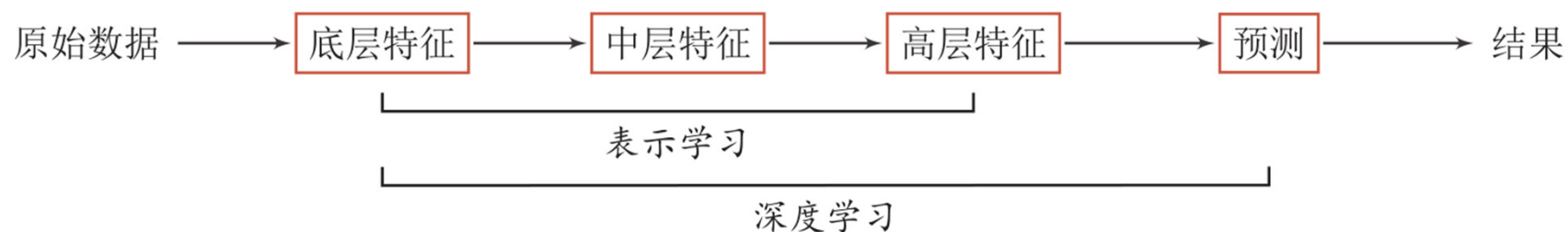
## ▶ 特征提取VS表示学习

▶ 特征提取：基于任务或先验对去除无用特征

▶ 表示学习：通过深度模型学习高层语义特征

# 深度学习

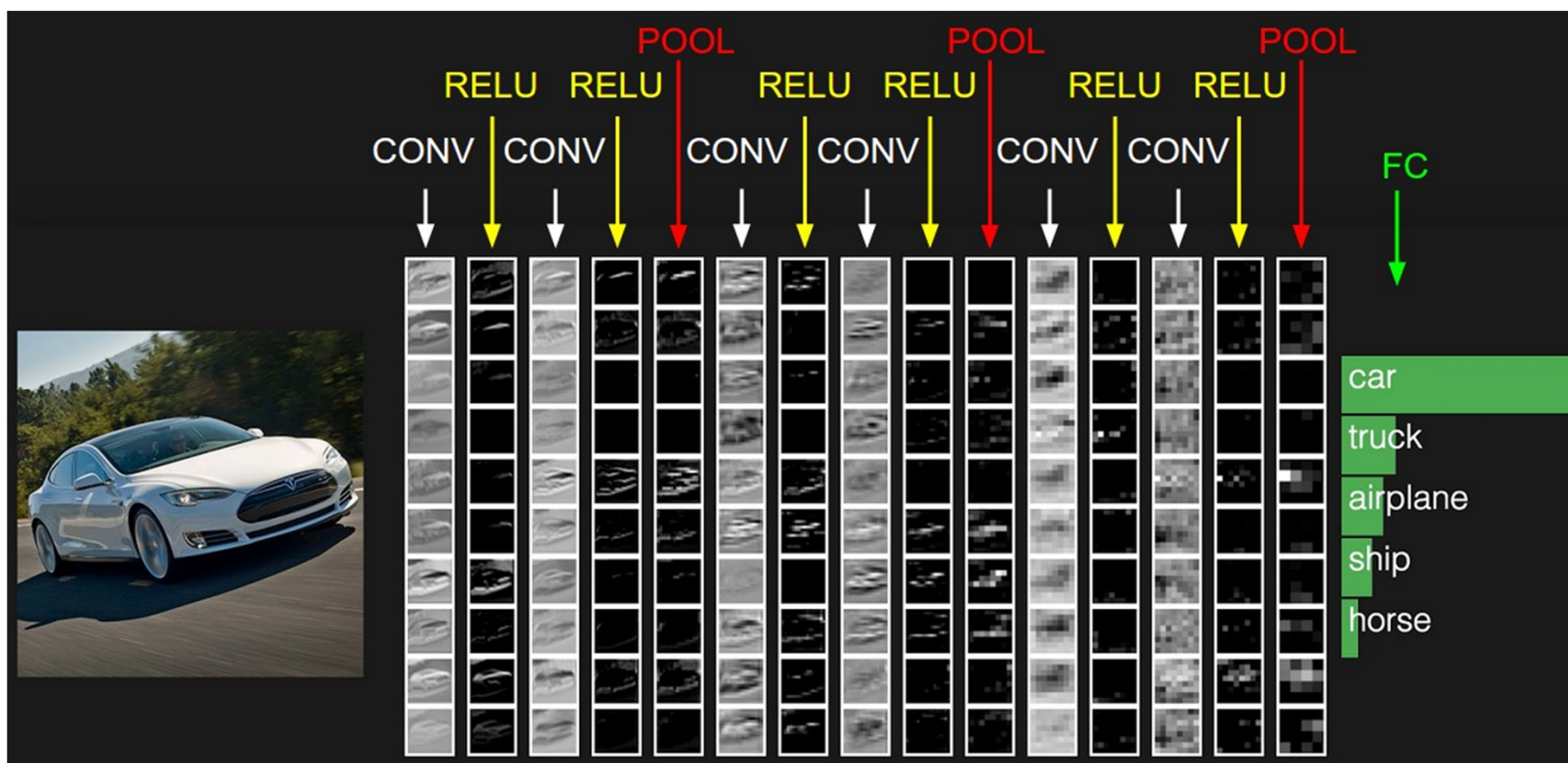
- ▶ 通过构建具有一定“深度”的模型，可以让模型来自动学习好的特征表示（从底层特征，到中层特征，再到高层特征），从而最终提升预测或识别的准确性。



- ▶ 端到端学习（端到端训练）：在学习过程中不进行分模块或者分阶段训练，直接优化任务的总体目标。中间过程无需人为干预，目前大部分采用神经网络的深度学习可以看作是一种端到端的学习

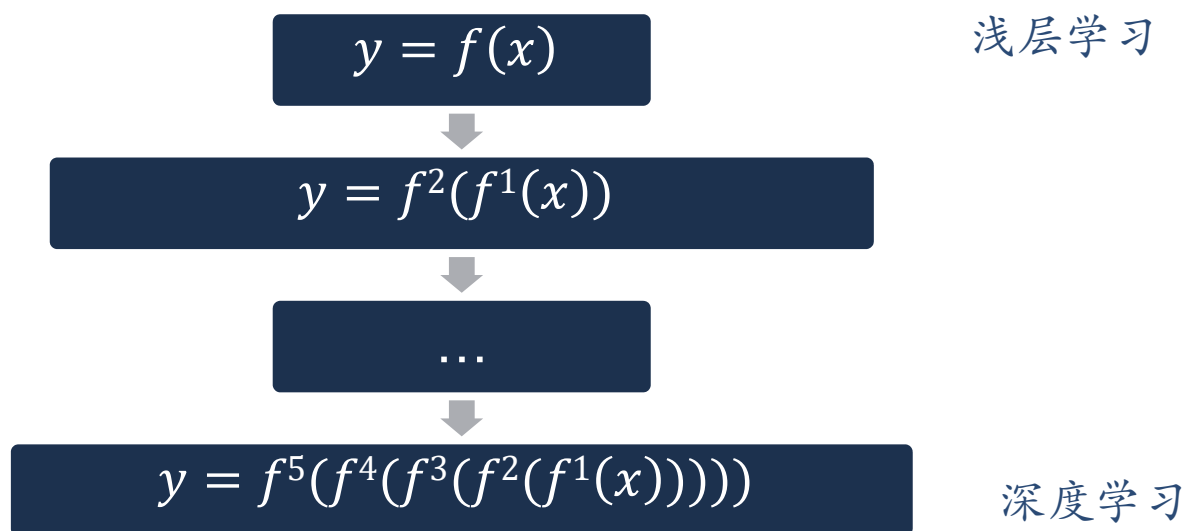


# 表示学习与深度学习



# 深度学习的数学描述

---



$f^1(x)$ 为非线性函数，不一定连续。

当 $f^1(x)$ 连续时为神经网络！ 比如 $f^1(x) = \sigma(W^1 f^{1-1}(\cdot))$

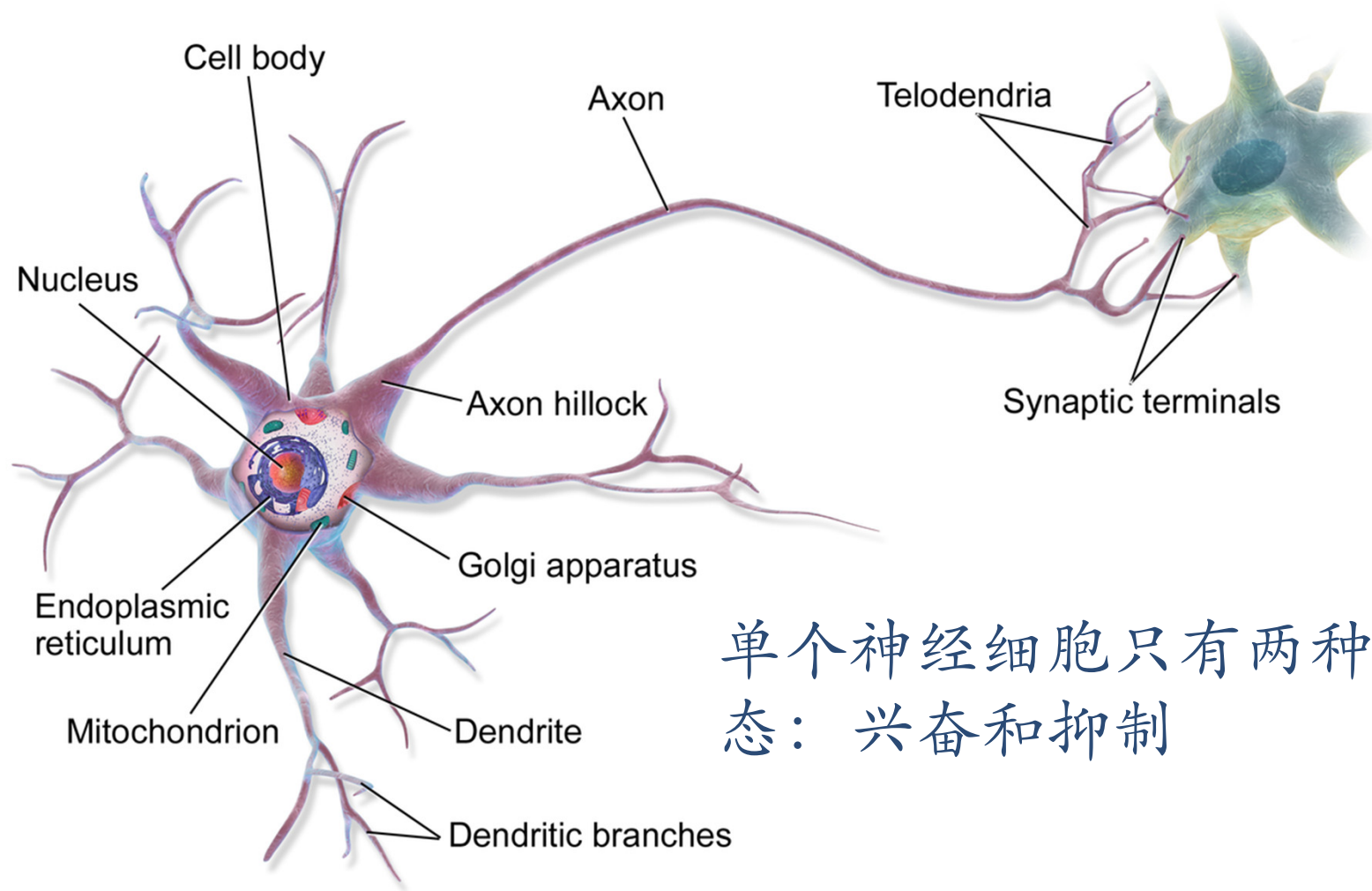
---

# 神经网络

# 生物神经元

人脑有860亿个神经元

[video: structure of brain](#)



单个神经细胞只有两种状态：兴奋和抑制

# 神经网络如何学习?

---

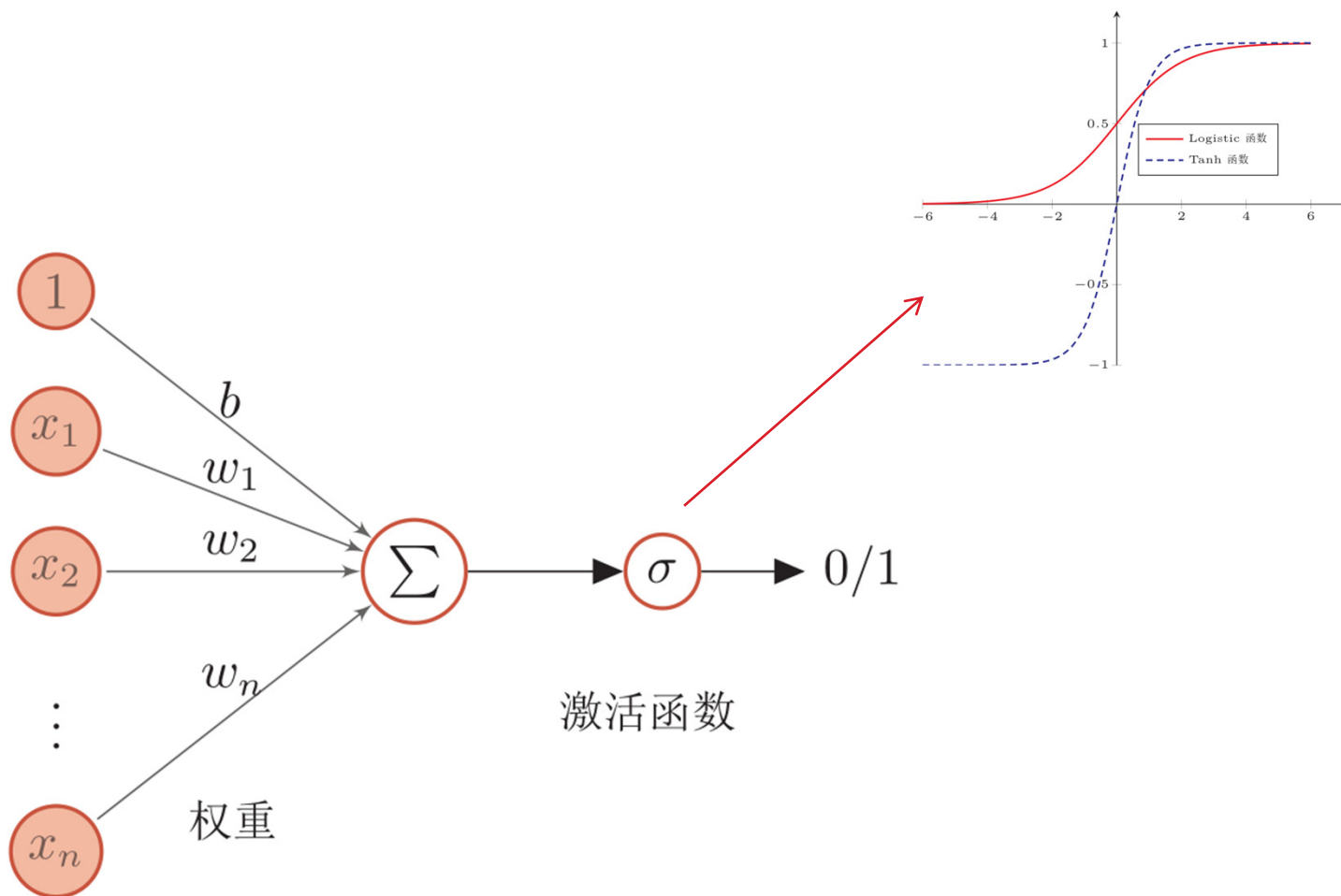
## ▶ 赫布法则 Hebb's Rule

- ▶ “当神经元 A 的一个轴突和神经元 B 很近，足以对它产生影响，并且持续地、重复地参与了对神经元 B 的兴奋，那么在这两个神经元或其中之一会发生某种生长过程或新陈代谢变化，以致于神经元 A 作为能使神经元 B 兴奋的细胞之一，它的效能加强了。”

----加拿大心理学家 Donald Hebb,  
《行为的组织》，1949

- ▶ 人脑有两种记忆：**长期记忆**和**短期记忆**。短期记忆持续时间不超过一分钟。如果一个经验重复足够的次数，此经验就可储存在长期记忆中。
- ▶ 短期记忆转化为长期记忆的过程就称为**凝固作用**。
- ▶ 人脑中的海马区为大脑结构凝固作用的核心区域。

# 人工神经元



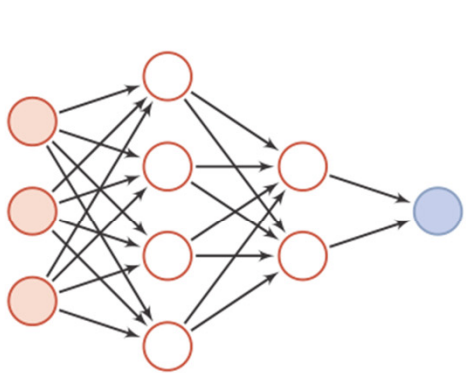
# 人工神经网络

---

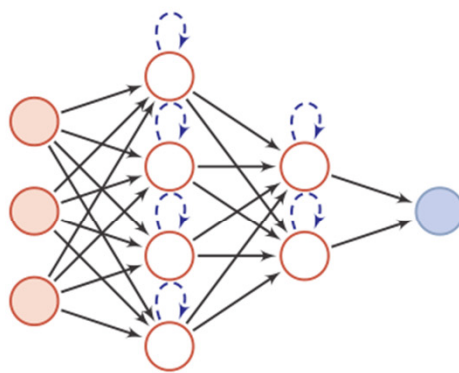
- ▶ 人工神经网络主要由大量的神经元以及它们之间的有向连接构成。因此考虑三方面：
  - ▶ 神经元的激活规则
    - ▶ 主要是指神经元输入到输出之间的映射关系，一般为非线性函数。
  - ▶ 网络的拓扑结构
    - ▶ 不同神经元之间的连接关系。
  - ▶ 学习算法
    - ▶ 通过训练数据来学习神经网络的参数。

# 人工神经网络

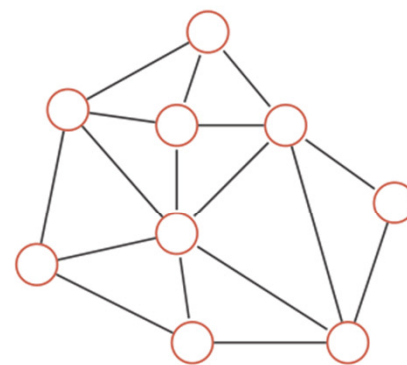
- ▶ 人工神经网络由神经元模型构成，这种由许多神经元组成的信息处理网络具有并行分布结构。
- ▶ 虽然这里将神经网络结构大体上分为三种类型，但是大多数网络都是**复合型结构**，即一个神经网络中包括多种网络结构。



(a) 前馈网络



(b) 记忆网络



(c) 图网络

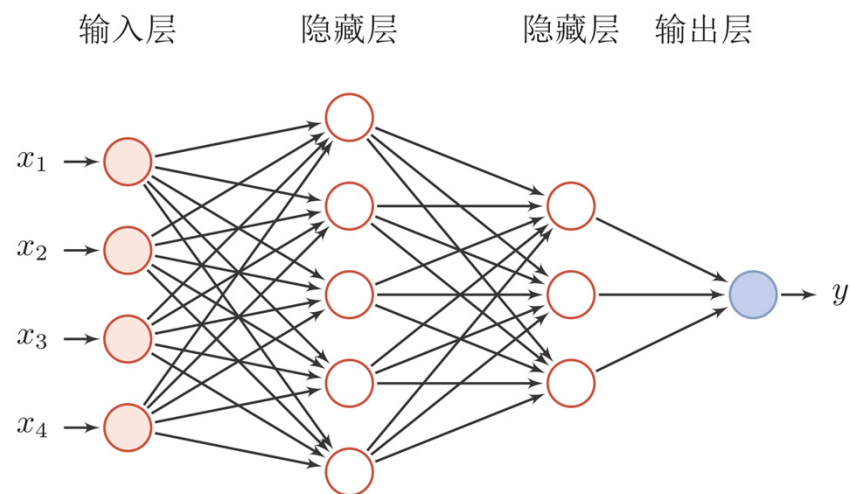


# 神经网络

---

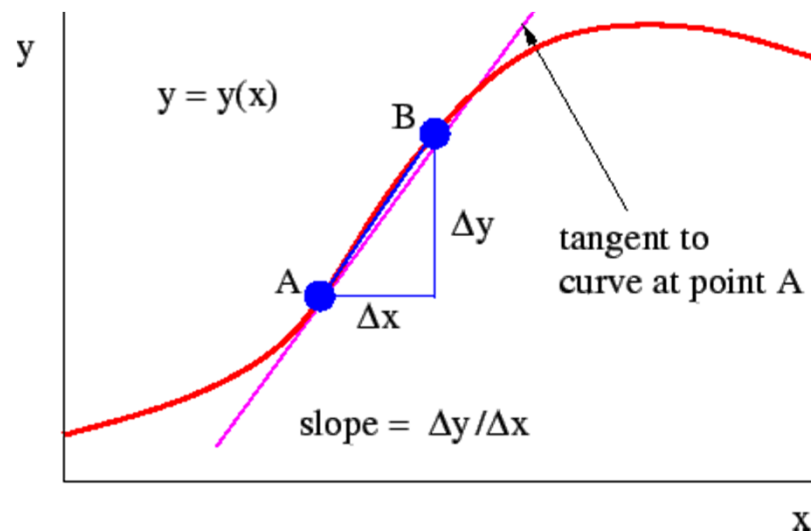
$$y = (f^3(f^2(f^1(x))))$$

$$f^1(x) = \sigma(W^1 f^{1-1}(\cdot))$$



# 如果解决贡献度分配问题?

## ▶ 偏导数



## ▶ 贡献度

$$\frac{\partial y}{\partial W^l} = \frac{y(W^l + \Delta W^l) - y(W^l)}{\Delta W^l}$$

# 神经网络发展史

---

- ▶ 神经网络的发展大致经过五个阶段。
- ▶ 第一阶段：模型提出（1943~1969年）
  - ▶ 在1943年，心理学家Warren McCulloch和数学家Walter Pitts和最早描述了一种理想化的人工神经网络，并构建了一种基于简单逻辑运算的计算机制。他们提出的神经网络模型称为MP模型。
  - ▶ 阿兰·图灵在1948年的论文中描述了一种“B型图灵机”。(赫布型学习)
  - ▶ 1951年，McCulloch和Pitts的学生Marvin Minsky建造了第一台神经网络机，称为SNARC。
  - ▶ Rosenblatt [1958]最早提出可以模拟人类感知能力的神经网络模型，并称之为感知器（Perceptron），并提出了一种接近于人类学习过程（迭代、试错）的学习算法。

# 神经网络发展史

---

## ▶ 第二阶段：冰河期（1969~1983年）

- ▶ 1969年，Marvin Minsky出版《感知机》一书，书中论断直接将神经网络打入冷宫，导致神经网络十多年的“冰河期”。他们发现了神经网络的两个关键问题。
  - ▶ 第一是基本感知机无法处理异或回路。
  - ▶ 第二个重要的问题是电脑没有足够的能力来处理大型神经网络所需要的很长的计算时间。
- ▶ 1974年，哈佛大学的Paul Webos发明反向传播算法，但当时未受到应有的重视。
- ▶ 1980年，Kunihiko Fukushima（福岛邦彦）提出了一种带卷积和子采样操作的多层神经网络：新知机（Neocognitron）

# 神经网络发展史

---

## ▶ 第三阶段：反向传播算法引起的复兴（1983~1995年）

- ▶ 1983年，加州理工学院的物理学家John Hopfield对神经网络引入能量函数的概念，并提出了用于联想记忆和优化计算的神经网络（称为Hopfield网络），在旅行商问题上获得当时最好结果，引起轰动。
- ▶ 1984年，Geoffrey Hinton提出一种随机化版本的Hopfield网络，即玻尔兹曼机。
- ▶ 1986年，David Rumelhart和James McClelland对于联结主义在计算机模拟神经活动中的应用提供了全面的论述，并重新发明了反向传播算法。
- ▶ Geoffrey Hinton[1986]等人将引入到多层感知器
- ▶ LeCun et al. [1989]将反向传播算法引入了卷积神经网络，并在手写体数字识别上取得了很大的成功。

# 神经网络发展史

---

## ▶ 第四阶段：流行度降低（1995~2006年）

- ▶ 在20世纪90年代中期，统计学习理论和以支持向量机为代表的机器学习模型开始兴起。
- ▶ 相比之下，神经网络的理论基础不清晰、优化困难、可解释性差等缺点更加凸显，神经网络的研究又一次陷入低潮。

# 神经网络发展史

---

## ▶ 第五阶段：深度学习的崛起（2006年~至今）

- ▶ 2006年，Hinton and Salakhutdinov [2006]发现多层前馈神经网络可以先通过逐层预训练，再用反向传播算法进行精调的方式进行有效学习。
- ▶ 深度的人工神经网络在语音识别和图像分类等任务上的巨大成功。
- ▶ 随着大规模并行计算以及GPU设备的普及，计算机的计算能力得以大幅提高。此外，可供机器学习的数据规模也越来越大。在计算能力和数据规模的支持下，计算机已经可以训练大规模的人工神经网络。

# 深度学习革命

---

## ▶ AI领域

- ▶ 语音识别：可以使得词错误率从1/4下降到1/8
- ▶ 计算机视觉：目标识别、图像分类等
- ▶ 自然语言处理：分布式表示、机器翻译、问题回答等
- ▶ 信息检索、社会化网络

## ▶ 三个Deep:

- ▶ Deep Blue
- ▶ Deep QA
- ▶ Deep Learning



# 学术机构

---

- ▶ Toronto 大学
  - ▶ Hinton 75 年 Edinburgh 大学博士
- ▶ NYU
  - ▶ Lecun (Now Facebook) 87 年 Hinton 博士后
- ▶ Montreal 大学
  - ▶ Bengio 91 年 M. Jordan 博士后
- ▶ Stanford 大学
  - ▶ Ng (Now Baidu) 03 年 UC Berkeley 大学 M. Jordan 博士
- ▶ IDSIA
  - ▶ Jürgen Schmidhuber



# 常用的深度学习框架

- 1. 简易和快速的原型设计
- 2. 自动梯度计算
- 3. 无缝CPU和GPU切换

 PaddlePaddle

  
TensorFlow

  
Chainer

 Microsoft  
CNTK

 Keras

 Caffe2

theano

PYTORCH

 APACHE  
mxnet

# 课后作业

---

- ▶ 复习数学知识（课本后面的数学基础）
- ▶ 编程练习
  - ▶ 熟悉基本的Numpy操作
    - ▶ Numpy是Python中对于矩阵处理很实用的工具包
  - ▶ <https://github.com/nndl/exercise>
    - ▶ [chap1\\_warmup](#)