

《神经网络与深度学习》



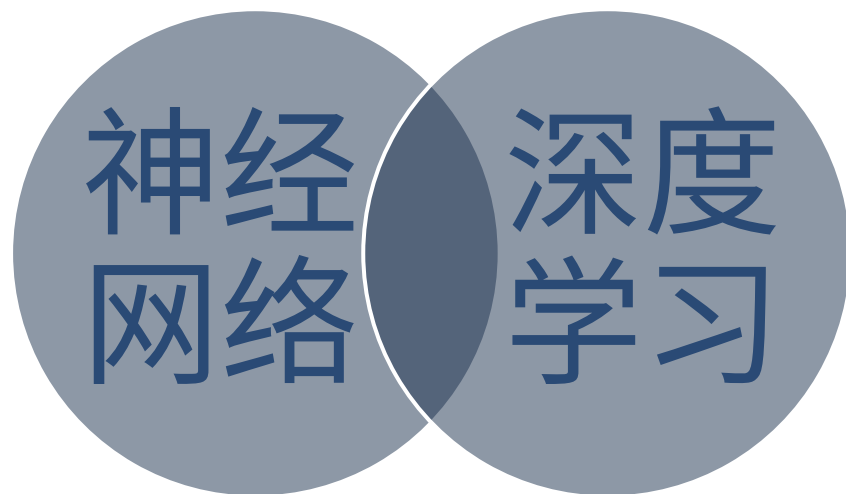
绪论

<https://nndl.github.io/>

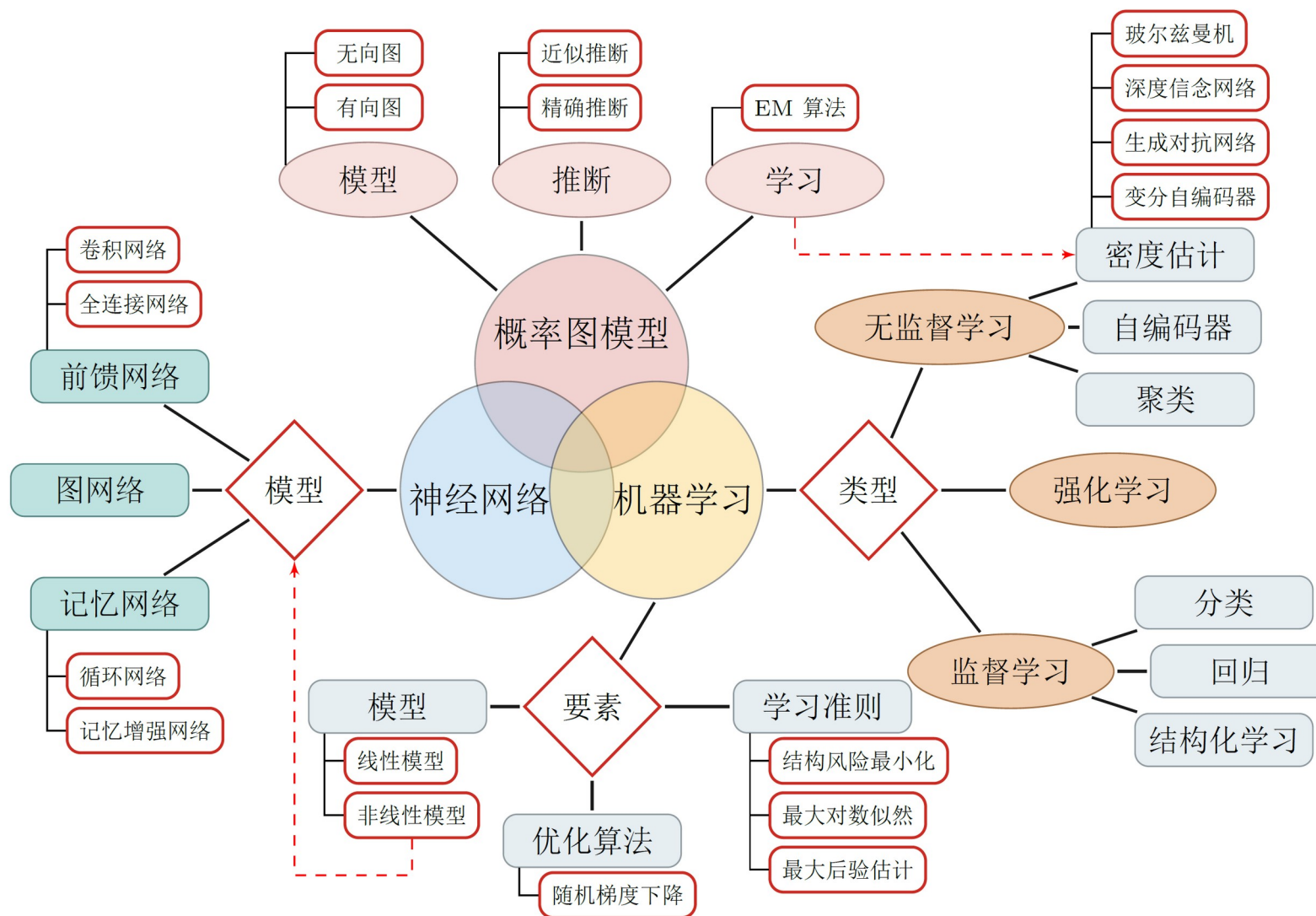
关于本课程

► 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。



更详细的课程概括



课程大纲

► 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

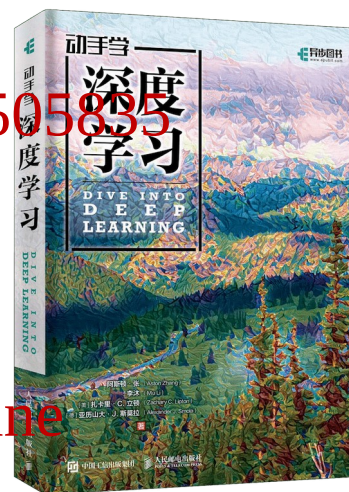
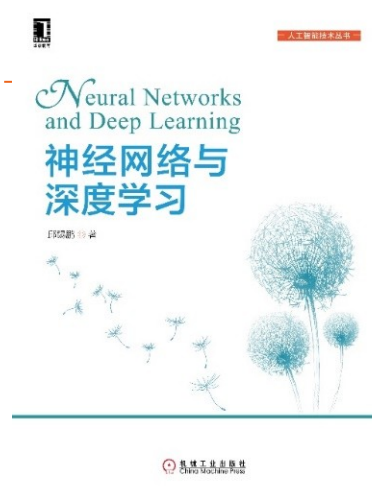
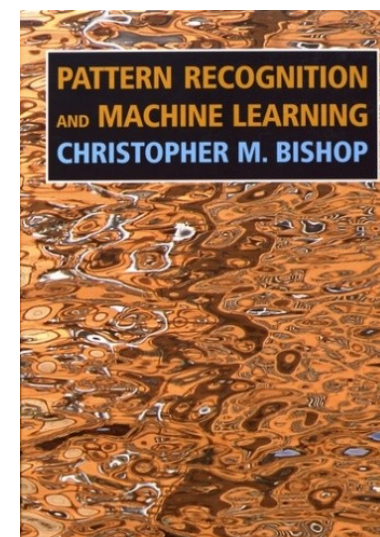
预备知识

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

<https://nndl.github.io/>
《数学基础》

推荐教材

- ▶ 邱锡鹏，神经网络与深度学习，机械工业出版社，2020，ISBN 9787111649687
 - ▶ <https://nndl.github.io/>
 - ▶ 提供配套练习
- ▶ 阿斯顿·张等，动手学深度学习，ISBN: 9787115505835
 - ▶ <https://d2l.ai/>
 - ▶ 有 PyTorch 版
- ▶ Bishop, C.M. (2006). Pattern recognition and Machine Learning. Springer. ISBN 9780387310732.
 - ▶ 网上有中文版 by 马春鹏



推荐教材

- ▶ Wright, S., & Nocedal, J. (1999). Numerical optimization. Springer Science, 35(67-68), 7.
- ▶ Boyd, S., & Vandenberghe, L. (2004). Convex optimization. Cambridge university press.

推荐课程

- ▶ **斯坦福大学 CS224n: Deep Learning for Natural Language Processing**
 - ▶ <https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1194/>
 - ▶ Chris Manning 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

- ▶ 李宏毅 “机器学习 2020 ”

- ▶ <https://www.bilibili.com/video/av94519857/>

顶会论文

- ▶ NeurIPS 、 ICLR 、 ICML 、 AAAI 、 IJCAI
- ▶ ACL 、 EMNLP
- ▶ CVPR 、 ICCV
- ▶ ...

成绩（研究生）

- ▶ 课程建设

 - ▶ 40%

- ▶ 期末作业

 - ▶ 60%

- ▶ 2 人一组

- ▶ 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”
- ▶ 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）

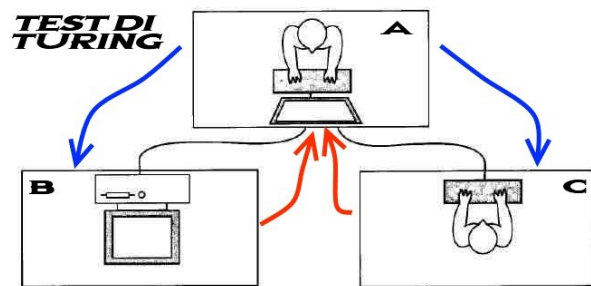
图灵测试

“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

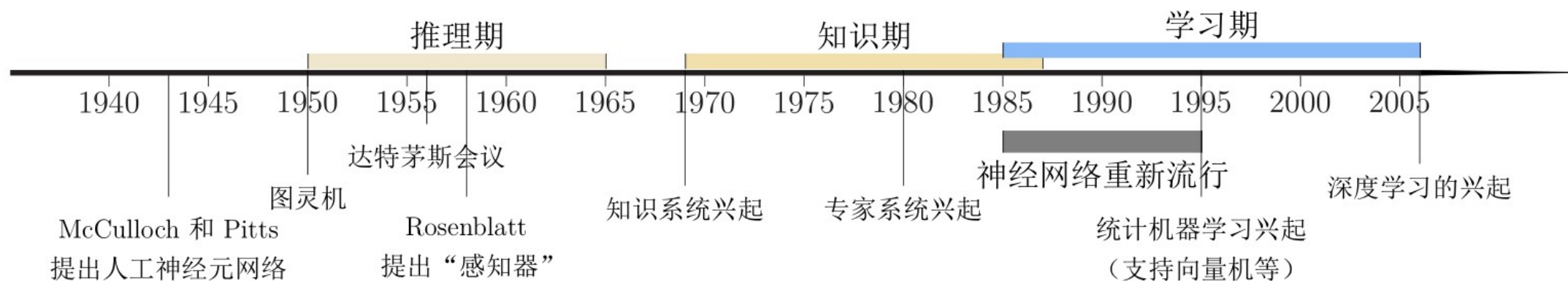


人工智能的研究领域

▶ 让机器具有人类的智能

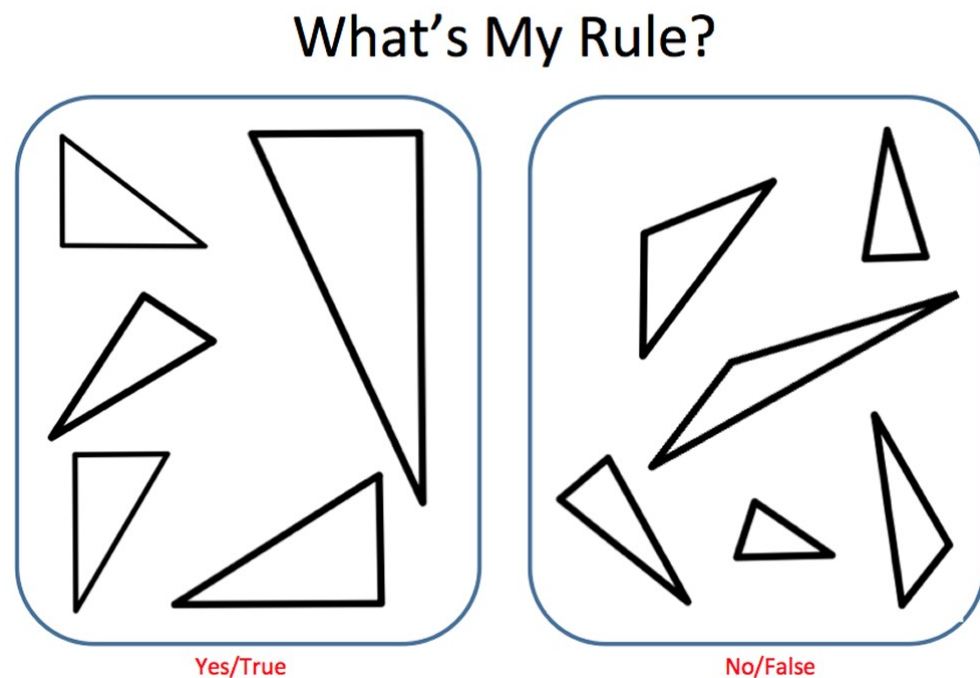
- ▶ 机器感知（计算机视觉、语音信息处理）
- ▶ 学习（模式识别、机器学习、强化学习）
- ▶ 语言（自然语言处理）
- ▶ 记忆（知识表示）
- ▶ 决策（规划、数据挖掘）

发展历史



如何开发一个人工智能系统？

► 专家知识（人工规则）



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

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

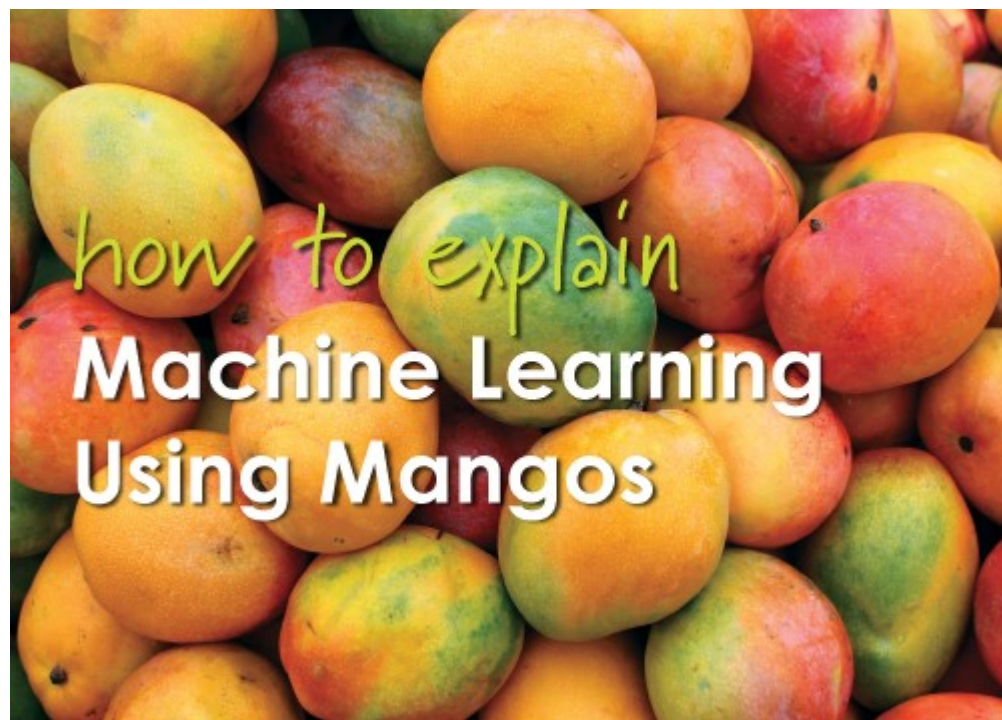
$$f(\text{数字9}) = \text{“9”}$$

► 围棋

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

► 机器翻译

$$f(\text{“你好!”}) = \text{“Hello!”}$$

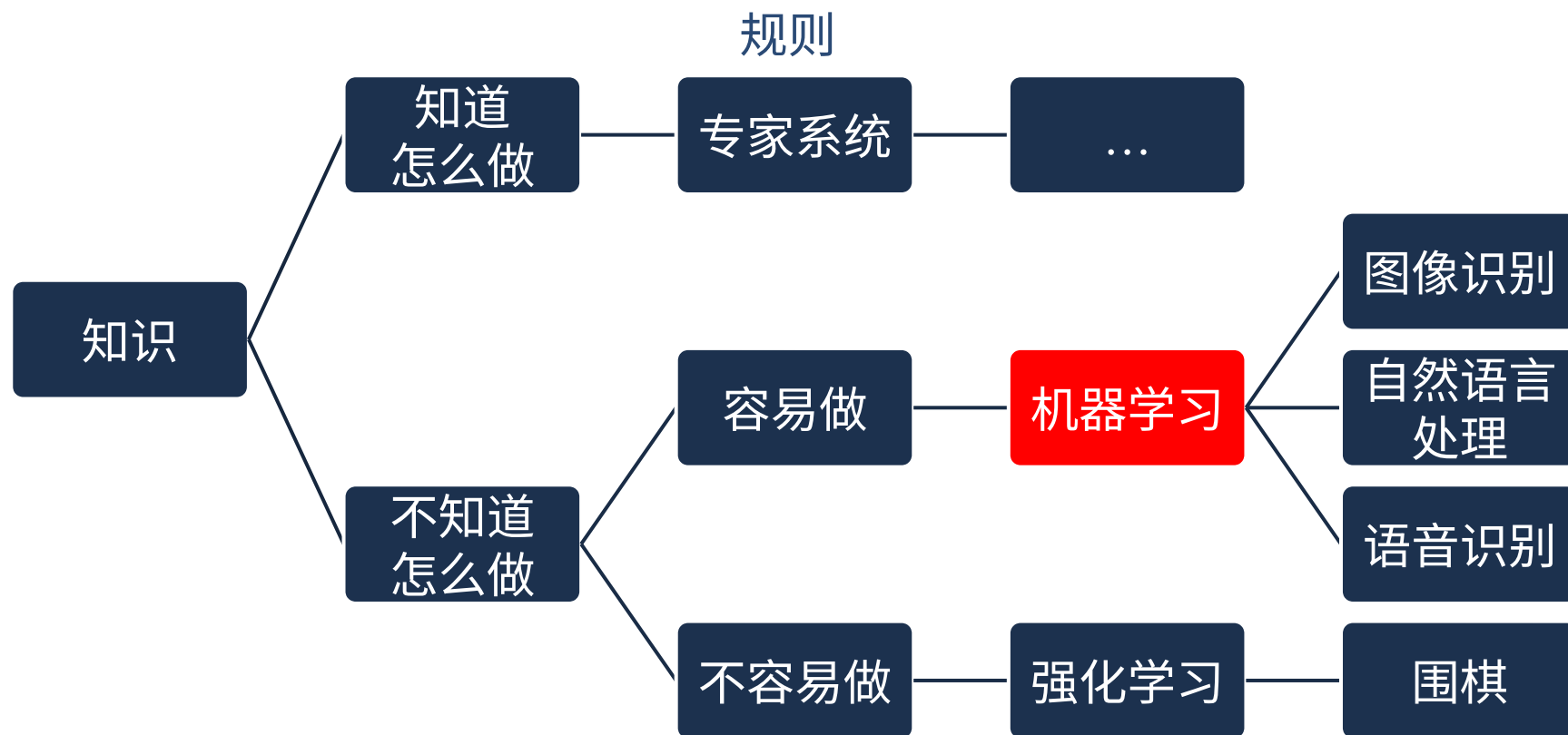


<https://www.quora.com/How-do-you-explain-Machine-Learning-and-Data-Mining-to-non-Computer-Science-people>

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。

如何开发一个人工智能系统？

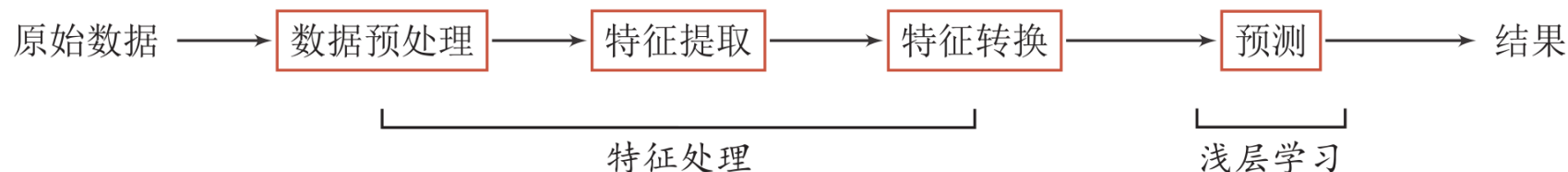




深度学习

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。

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。

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：需要借助人类智能

- ▶ **表示学习**

- ▶ 如何自动从数据中学习好的表示

- ▶ **难点**

- ▶ 没有明确的目标

什么是好的数据表示？

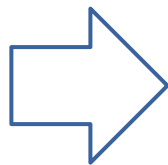
- ▶ “好的表示” 是一个非常主观的概念，没有一个明确的标准。
- ▶ 但一般而言，一个好的表示具有以下几个优点：
 - ▶ 应该具有很强的表示能力。
 - ▶ 应该使后续的学习任务变得简单。
 - ▶ 应该具有一般性，是任务或领域独立的。

语义表示

► 如何在计算机中表示语义？

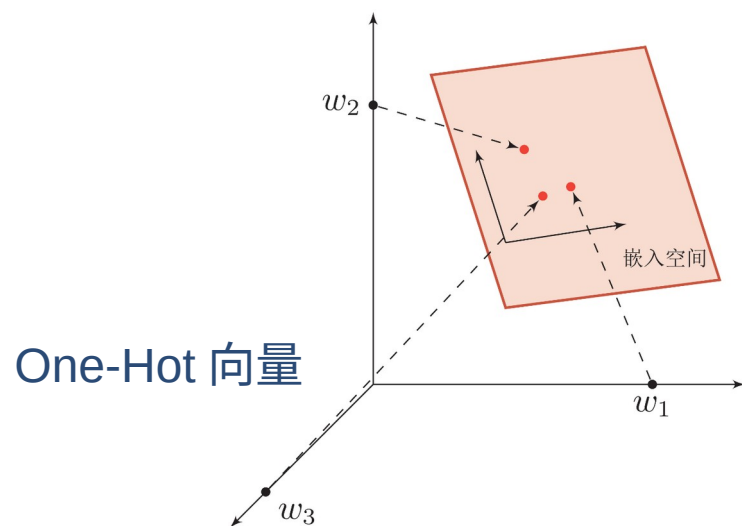
局部（符号）表示

知识库、规则



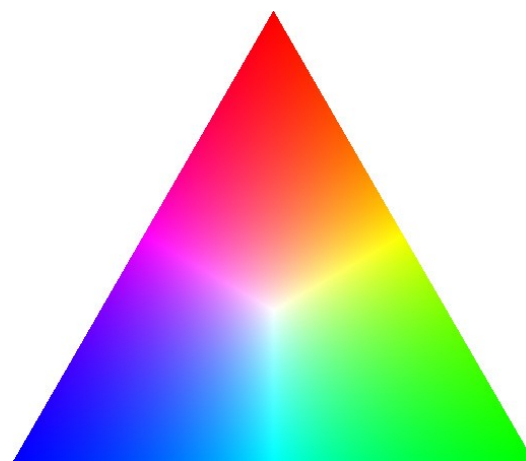
分布式表示

嵌入：压缩、低维、稠密向量



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	$[0, 0, 0, 1]^T$	$[0.44, 0.31, 0.22]^T$



表示形式

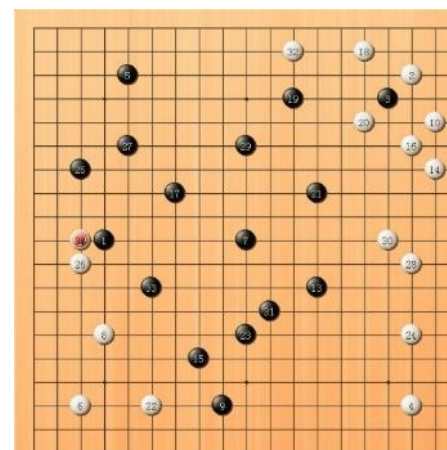
▶ 局部表示

- ▶ 离散表示、符号表示
- ▶ One-Hot 向量

	局部表示	分布式表示
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]

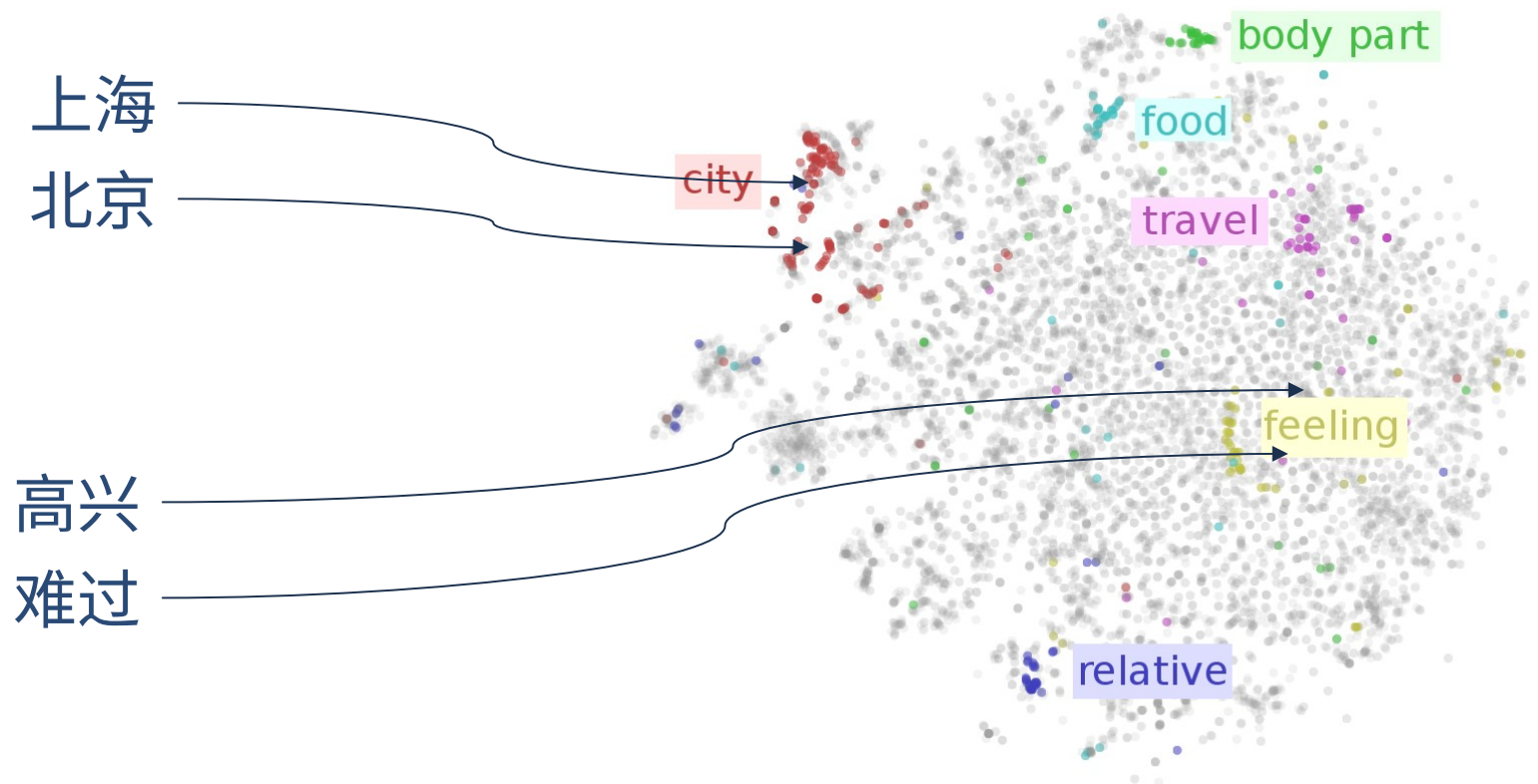
▶ 分布式 (distributed) 表示

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



分布式表示

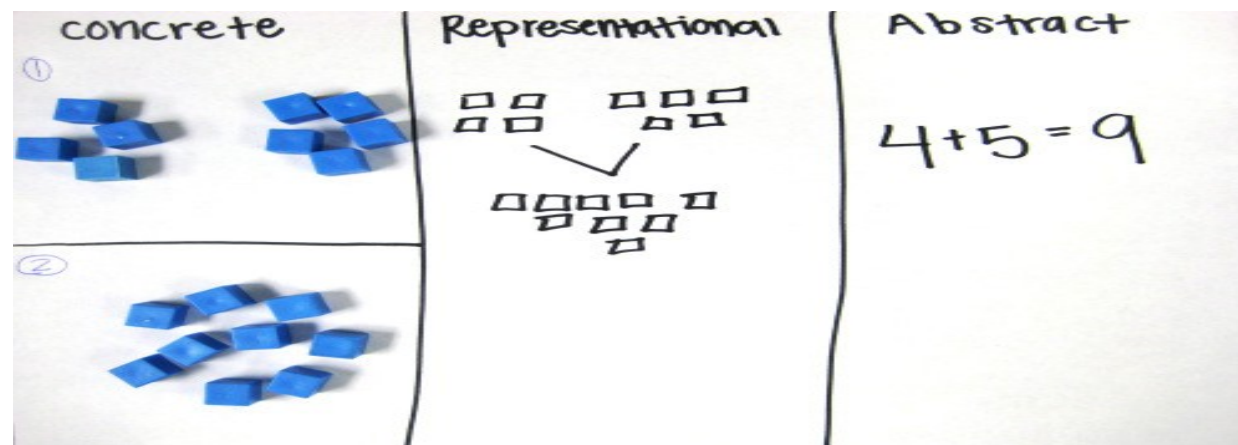
词嵌入 (Word Embeddings)



<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
 - ▶ 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

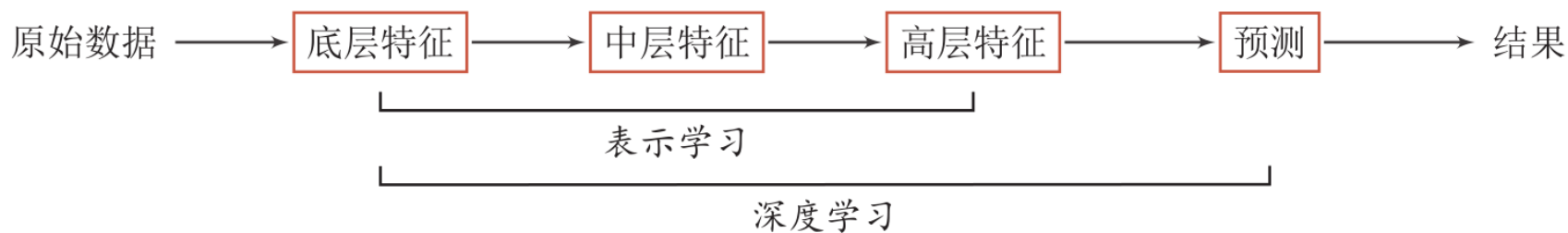
- ▶线性投影（子空间）
 - ▶PCA、LDA
- ▶非线性嵌入
 - ▶LLE、Isomap、谱方法
- ▶自编码器

▶特征提取 VS 表示学习

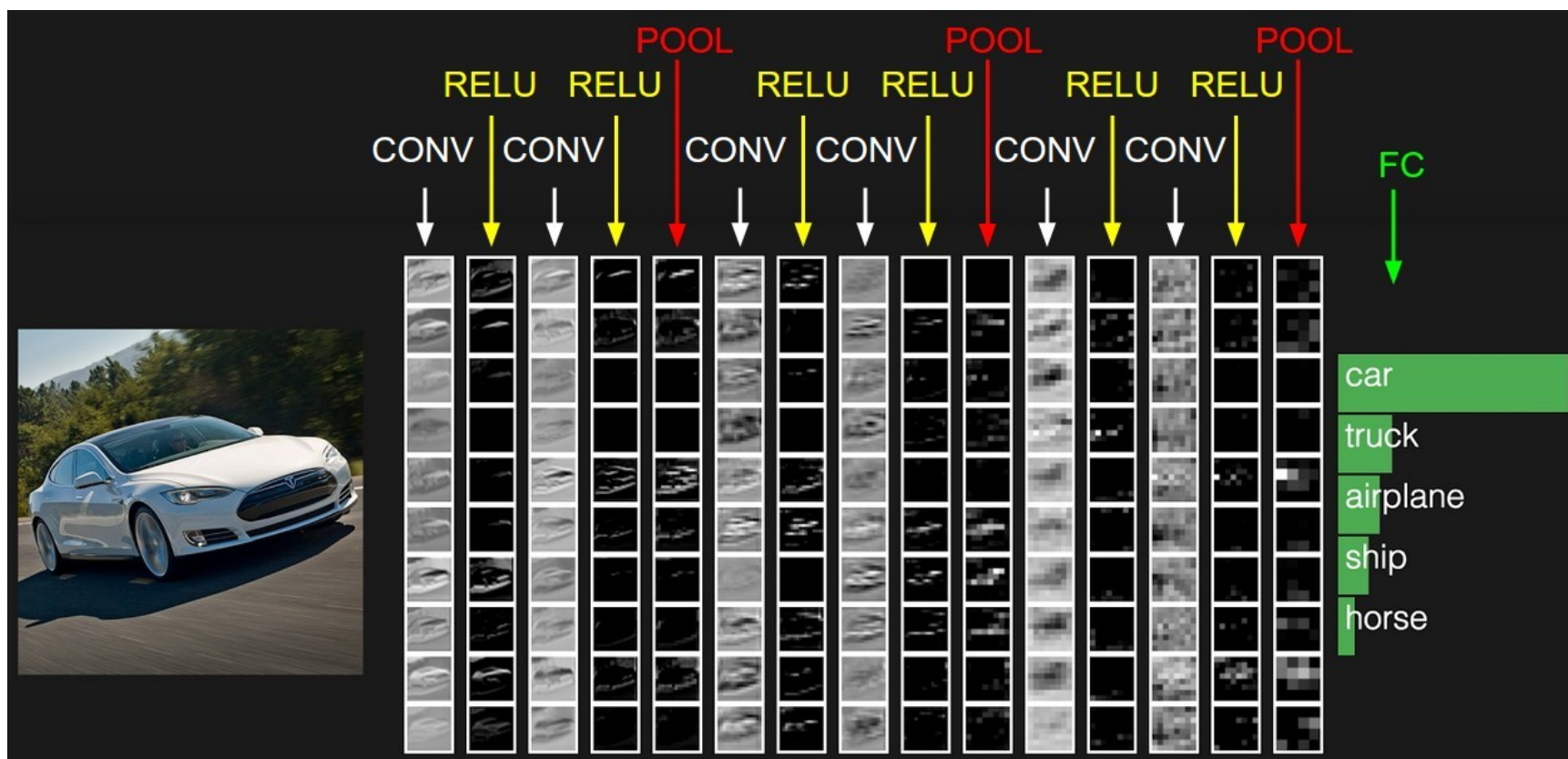
- ▶特征提取：基于任务或先验对去除无用特征
- ▶表示学习：通过深度模型学习高层语义特征

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。



表示学习与深度学习



深度学习的数学描述



$f^l(x)$ 为非线性函数，不一定连续。

当 $f^l(x)$ 连续时，比如 $f^l(x) = \sigma(W^l f^{l-1}(x))$ ，这个复合函数称为神经网络。

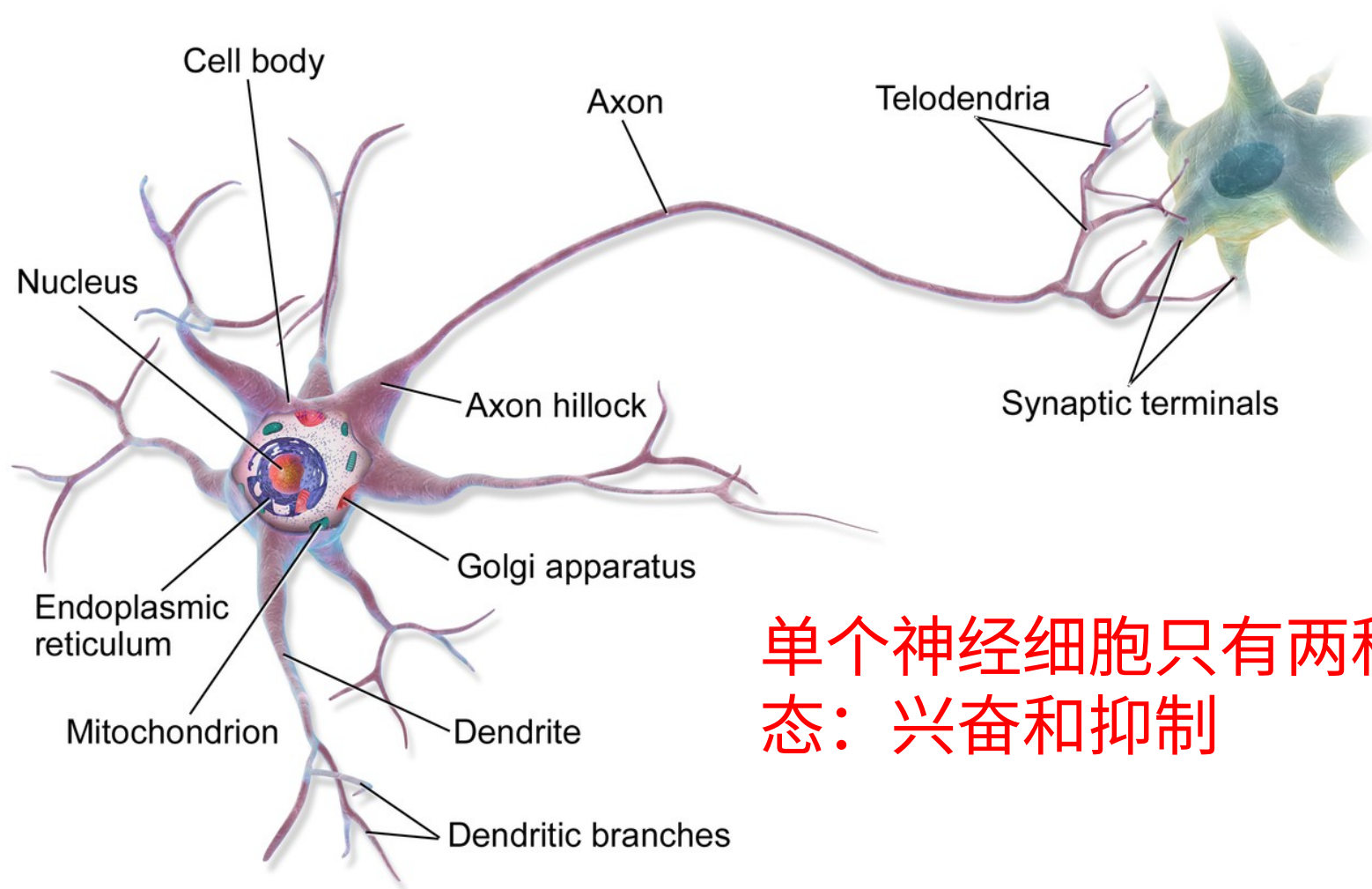


神经网络

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？

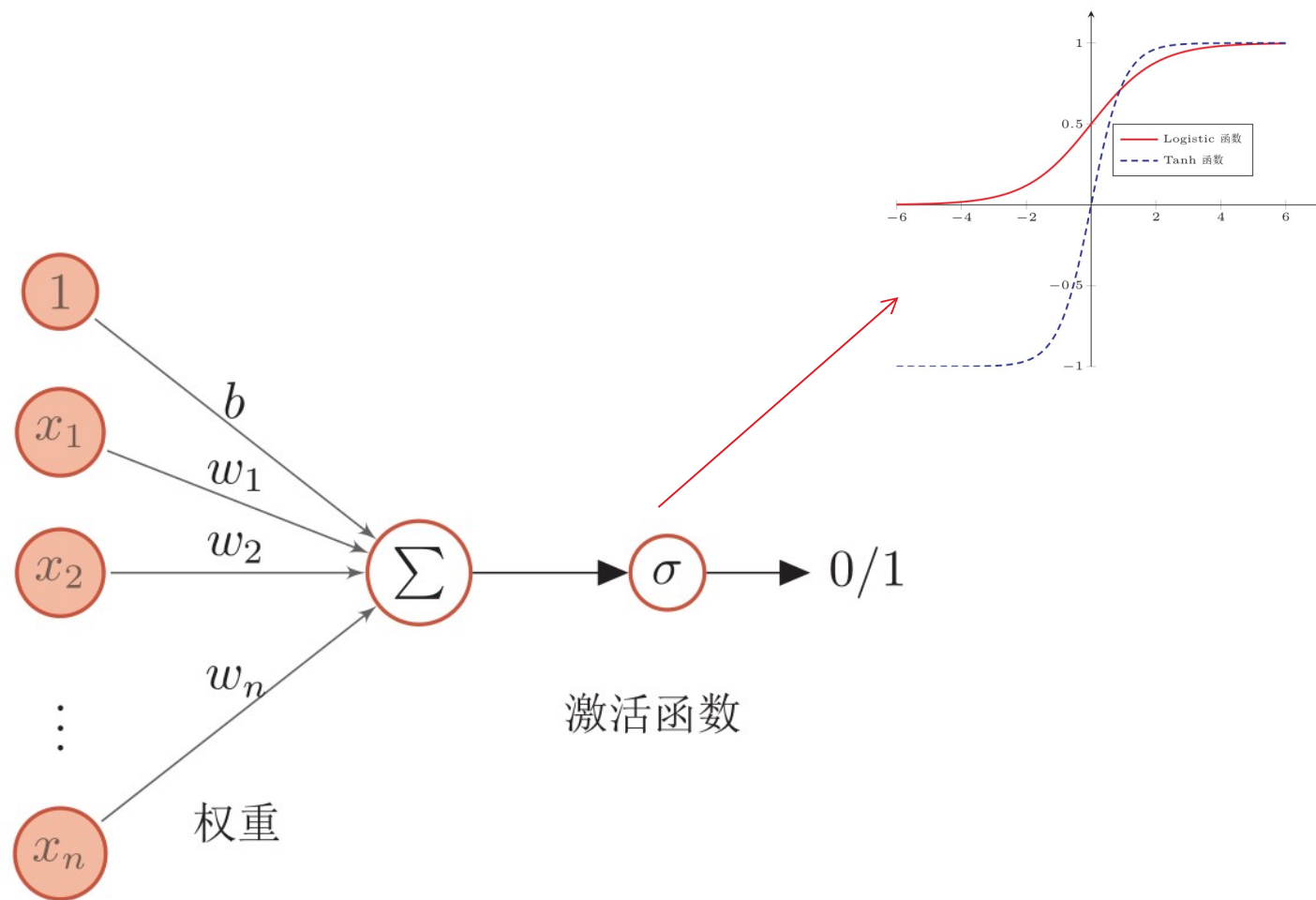
► 赫布法则 Hebb's Rule

- “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

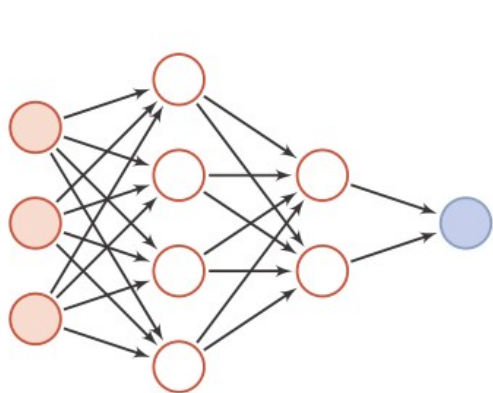


人工神经网络

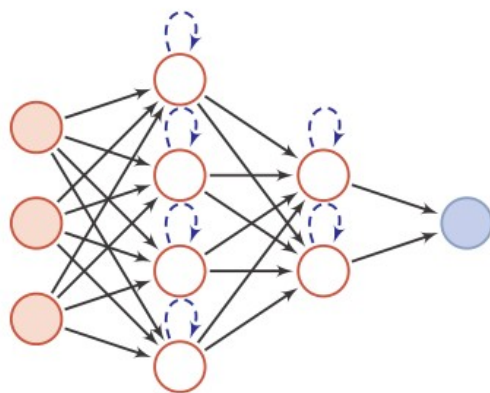
- ▶ 人工神经网络主要由大量的神经元以及它们之间的有向连接构成。因此考虑三方面：
- ▶ 神经元的激活规则
 - ▶ 主要是指神经元输入到输出之间的映射关系，一般为非线性函数。
- ▶ 网络的拓扑结构
 - ▶ 不同神经元之间的连接关系。
- ▶ 学习算法
 - ▶ 通过训练数据来学习神经网络的参数。

人工神经网络

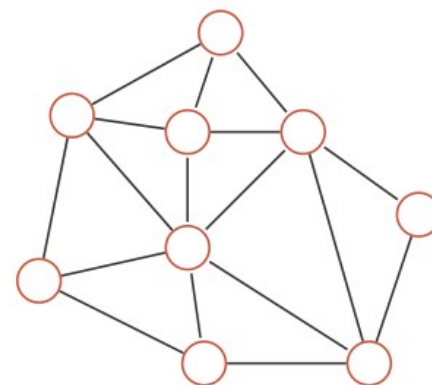
- ▶ 人工神经网络由神经元模型构成，这种由许多神经元组成的信息处理网络具有并行分布结构。
- ▶ 虽然这里将神经网络结构大体上分为三种类型，但是大多数网络都是复合型结构，即一个神经网络中包括多种网络结构。



(a) 前馈网络



(b) 记忆网络

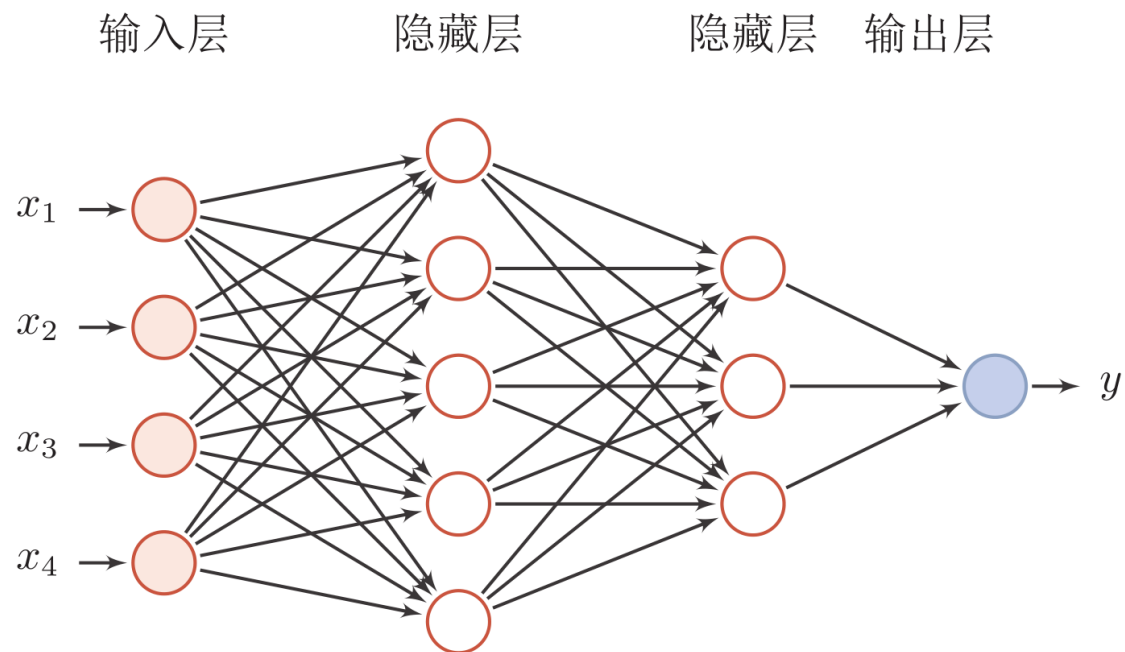


(c) 图网络

神经网络

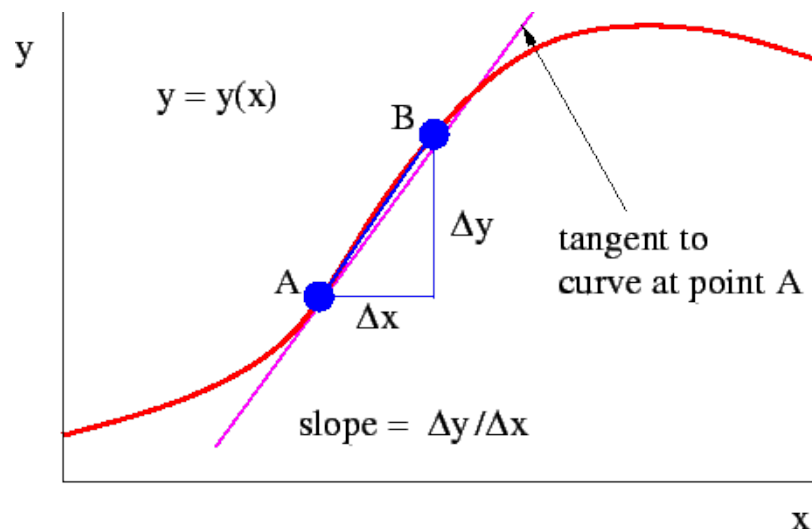
$$y = (f^3(f^2(f^1(x))))$$

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



如果解决贡献度分配问题？

► 偏导数



► 贡献度

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

神经网络天然不是深度学习，但深度学习天然神经网络。

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**
 - ▶ 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 年， Marvin Minsky 出版《感知器》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：
 - ▶ 1) 基本感知器无法处理异或回路。
 - ▶ 2) 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 Werbos 发明反向传播算法，但当时未受到应有的重视。
- ▶ 1980 年， Kunihiro 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 年，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。
- ▶ 1986 年，Geoffrey Hinton 等人将引入反向传播算法到多层感知器
- ▶ 1989 年，LeCun 等人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

- 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，Hinton 等人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。
 - 深度神经网络在语音识别和图像分类等任务上的巨大成功。
- 2013 年，AlexNet：第一个现代深度卷积网络模型，是深度学习技术在图像分类上取得真正突破的开端。
 - AlexNet 不用预训练和逐层训练，首次使用了很多现代深度网络的技术
- 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

早期的学术机构

► 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

► IDSIA

► Jürgen Schmidhuber

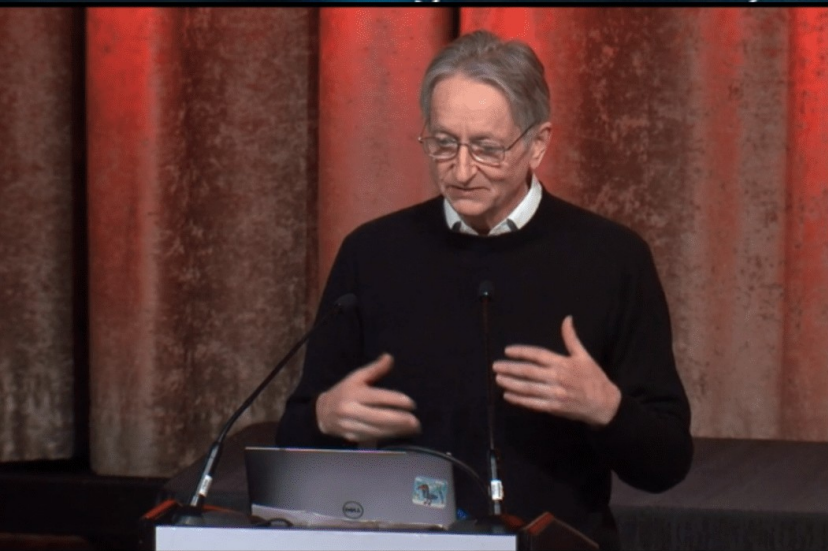


2018 图灵奖获得者

AAAI-20

Thirty-Fourth AAAI Conference on Artificial Intelligence

February 7-12 2020, Hilton New York Midtown, New York, New York USA

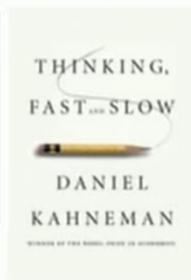


SYSTEM 1 VS. SYSTEM 2 COGNITION

2 systems (and categories of cognitive tasks):

System 1

- Intuitive, fast, **UNCONSCIOUS**, non-linguistic, habitual
- Current DL



System 2

- Slow, logical, sequential, **CONSCIOUS**, linguistic, algorithmic, planning, reasoning
- Future DL



Manipulates high-level / semantic concepts, which can be recombined combinatorially



humans reason not necessarily machines
as we train as we program them and and

常用的深度学习框架

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



课后作业

▶ 复习数学知识

- ▶ <https://nndl.github.io/> 附录 《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](#)