

深度学习

兰韵诗



DaSE
Data Science
& Engineering

授课教师

- 姓名：兰韵诗
- 电子邮箱：yslan@dase.ecnu.edu.cn
- 个人主页：<https://lanyunshi.github.io>
- 单位：华东师范大学 数据科学与工程学院
- 研究方向：自然语言处理，智能问答，计算教育学等

助教

- 陈沁文：研究生一年级
- 电子邮箱：
51275903075@stu.ecnu.edu.cn
- 研究方向：智能问答

- 王千予：研究生一年级
- 电子邮箱：
51275903082@stu.ecnu.edu.cn
- 研究方向：语法纠错

考核方式

- 理论课随堂表现 (10%)
- 实践作业 (50%)
- 期末考察 (40%)

理论课随堂表现

- 不缺席（每堂课随机点名）
- 积极回答课堂问题，参与互动

考察题目

- 题目类型

- 普通程序题：代码需要符合**指定要求**并且可以**顺利执行**。经助教要求修改后才能执行会被扣分
- 期末考察题：在**指定框架内编写程序代码**，解决一个实际应用问题。上传代码到平台，可以马上得知成果，需要与平台**排行榜的分数比较**

- 备注

- 普通程序题和竞赛作业均在“水杉在线”平台上完成
- 禁止抄作业，我们会对上交的代码进行抽查
- 答题出色的同学可能会被邀请在课堂上给大家讲解

考察题目

- 作业一(10%): 普通程序题, 一周内提交
- 作业二(10%): 普通程序题, 一周内提交
- 作业三(10%): 普通程序题, 一周内提交
- 作业四(10%): 普通程序题, 一周内提交
- 作业五(10%): 普通程序题, 一周内提交
- 期末考察题(40%): 五月初公布, 一个半月内提交



讨论

- 添加微信课程群
 - 将自己的名字改成真实姓名
- 有问题可以在微信上发言
 - 建议在工作时间发言
 - 如果有同学知道答案请帮忙回答
- 助教会记录好的问题、答案、发言



群聊: 本科深度学习 2025
春季

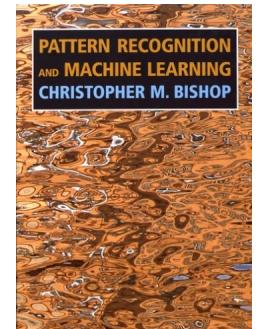
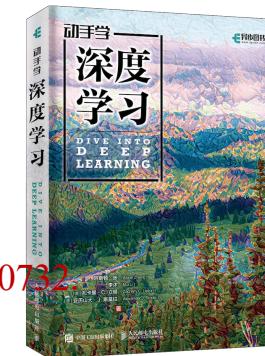
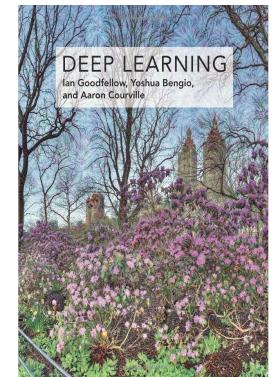
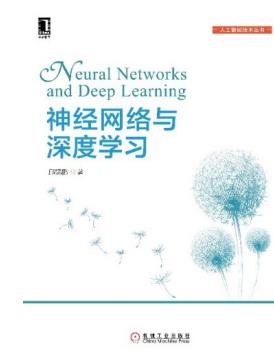


该二维码7天内(2月23日前)有效, 重新进入将更新

保存图片

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://mndl.github.io/>
- ▶ Ian Goodfellow等, 深度学习, ISBN 9780262035613
 - ▶ https://github.com/MingchaoZhu/DeepLearning/releases/download/v0.0.1/DL_cn.pdf
- ▶ 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
- ▶ 张奇、桂韬、郑锐、黄萱菁, 大语言模型理论与实践, 2024
 - ▶ <https://intro-llm.github.io/>
- ▶ 赵鑫、李军毅、周昆、唐天一、文继荣, 大语言模型, 2024
 - ▶ <https://llmbook-zh.github.io/>



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/>
- ▶ 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/slides/Tutorial_HYLee_Deep.pptx

推荐材料

- 斯坦福大学CS324: Large Language Models
 - <https://stanford-cs324.github.io/winter2022/>
 - Percy Liang更加理论的讲解大语言模型的工作原理
- CMU11-711: Advanced NLP Spring 2024
 - <https://www.phontron.com/class/anlp2024/lectures/>
 - Graham Neubig从自然语言处理应用的角度探讨大语言模型
- 普林斯顿COS 597R (Fall 2024): Deep Dive into Large Language Models
 - <https://princeton-cos597r.github.io/>
 - Danqi Chen和Sanjeev Arora结合前沿的发表论文对大语言模型的热门话题进行讨论

顶会论文

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

FAQ

- Q : 上这门课我需要有哪些基础知识 ?
- A : 基本的数学理论基础, 如 : 线性代数、概率统计等。基本的 python 编程能力。
- Q : 和统计机器学习有何不同 ?
- A : 部分理论和统计机器学习有一些衔接, 但重叠很少。
- Q : 和当代人工智能有何不同 ?
- A : 少量内容和模型有重复, 但和当代人工智能的侧重点不同。主要以全面掌握深度学习为目的, 随着大语言模型的发展, 后期加入少量大语言模型相关的课程内容。

《神经网络与深度学习》

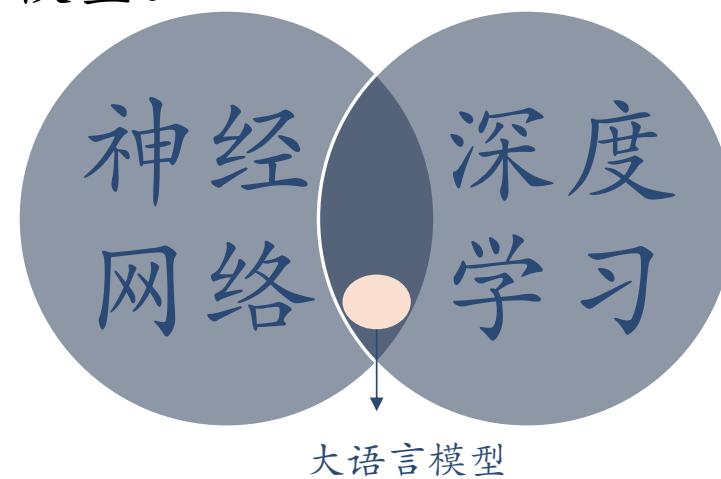


绪论

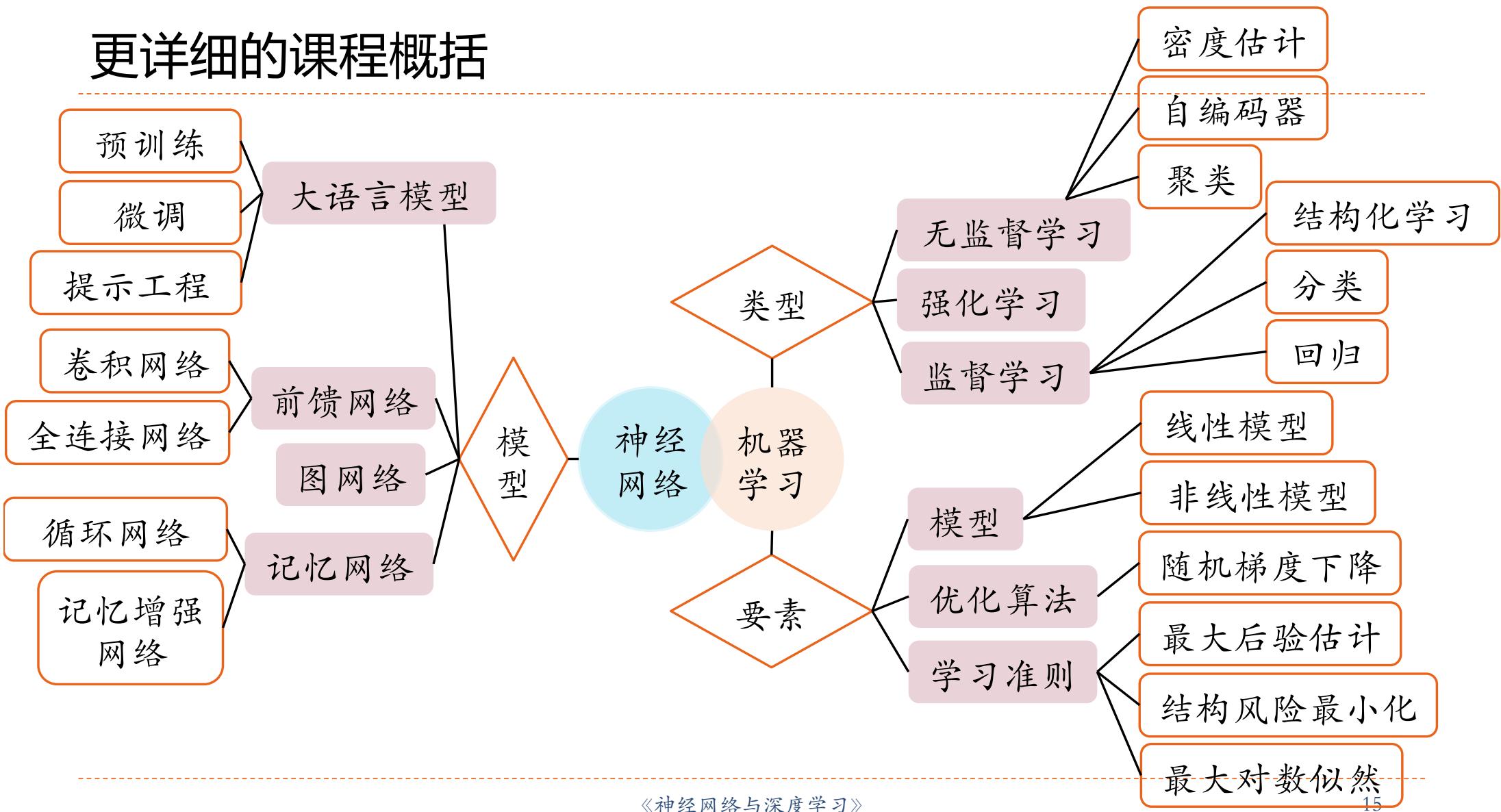
<https://nndl.github.io/>

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。
- ▶ 大语言模型：建立在深度学习之上的，以参数量大、训练语料量大为特点的深度学习模型。



更详细的课程概括



课程大纲

- ▶ 深度学习概述
- ▶ 线性模型
- ▶ 前馈神经网络(MLP, ...)
- ▶ 卷积神经网络(CNN, LeNet, ...)
- ▶ 循环神经网络(GRU, LSTM, ...)
- ▶ 序列生成模型(Seq2seq)
- ▶ 网络优化与正则化(Adagrad, Adam, ...)
- ▶ 独立学习方式 (ensemble, joint-learning, ...)
- ▶ 记忆与注意力机制(Attention Mechanism, Memory Network, ...)
- ▶ 无监督学习(PCA, Auto-Encoder, ...)
- ▶ 深度生成模型(VAE, GAN, diffusion model)
- ▶ ~~深度强化学习 (DQN, policy gradient, ...)~~
- ▶ 预训练语言模型概要
- ▶ 大语言模型架构和工作原理
- ▶ 大语言模型的提示工程(prompt learning, chain of thought...)
- ▶ 大语言模型的微调 (LoRA...)
- ▶ 大语言模型在语言教育中的应用实践

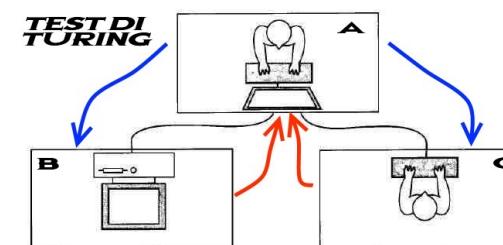
图灵测试

“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

---Alan Turing [1950]
«Computing Machinery and Intelligence»



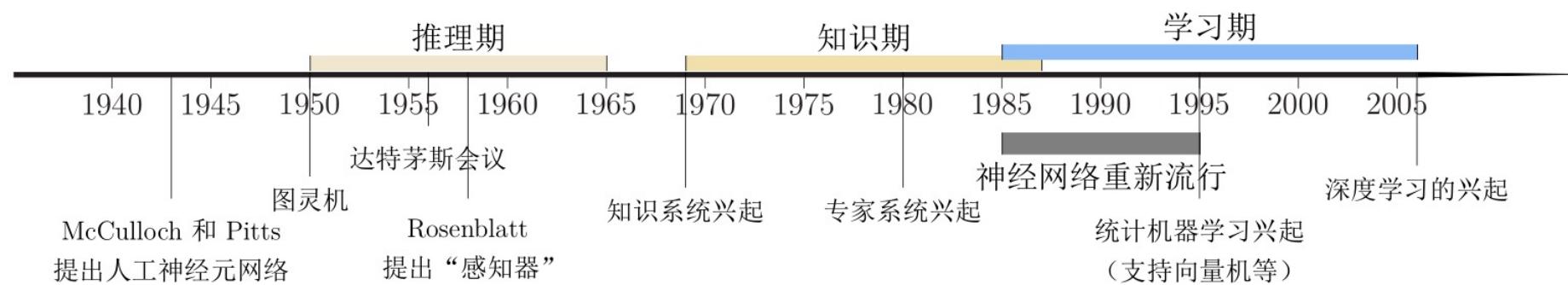
deepseek



人工智能的研究领域

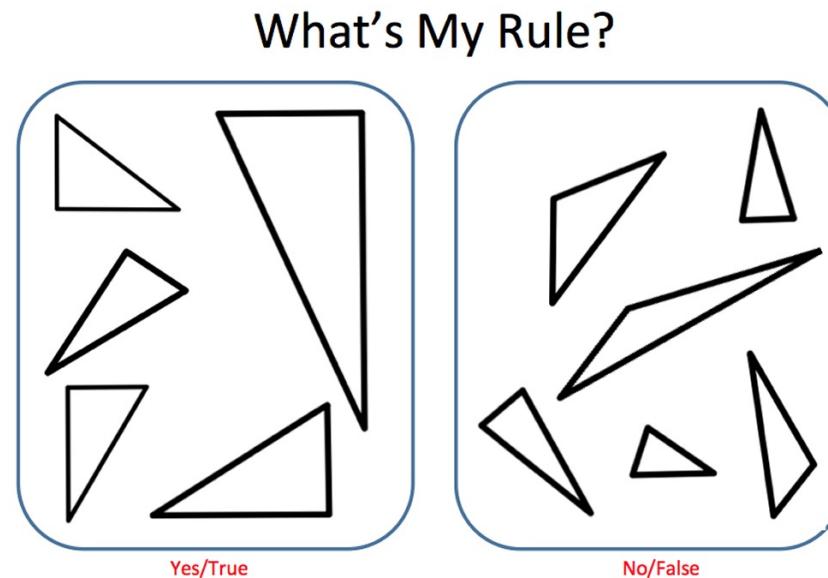
- ▶ 让机器具有人类的智能
 - ▶ 机器感知（计算机视觉、语音信息处理）
 - ▶ 学习（模式识别、机器学习、强化学习）
 - ▶ 语言（自然语言处理）
 - ▶ 记忆（知识表示）
 - ▶ 决策（规划、数据挖掘）

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?



机器学习

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 ≈ 构建一个映射函数

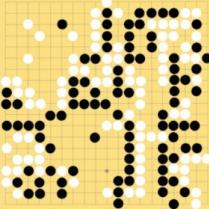
► 语音识别

$$f(\text{}) = \text{“你好”}$$

► 图像识别

$$f(\text{}) = \text{“9”}$$

► 围棋

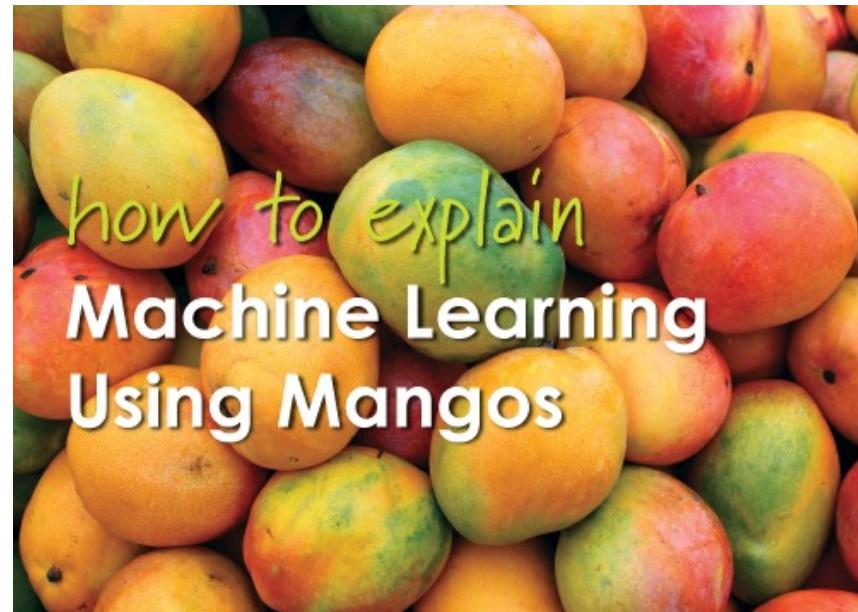
$$f(\text{}) = \text{“6-5” (落子位置)}$$

► 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？

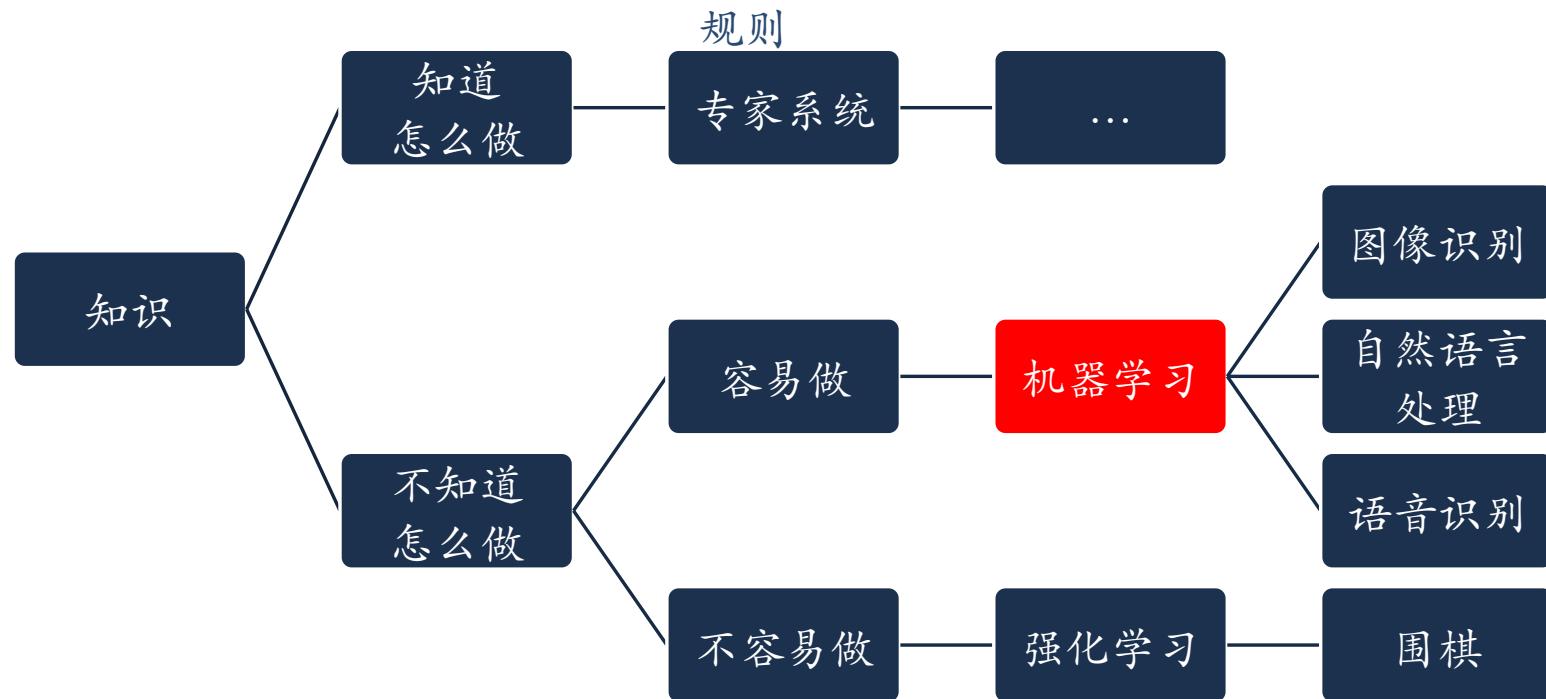


<https://www.quora.com/How-do-you-explain-Machine-Learning-and-Data-Mining-to-non-Computer-Science-people>

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。

如何开发一个人工智能系统？



为什么需要算法工程师

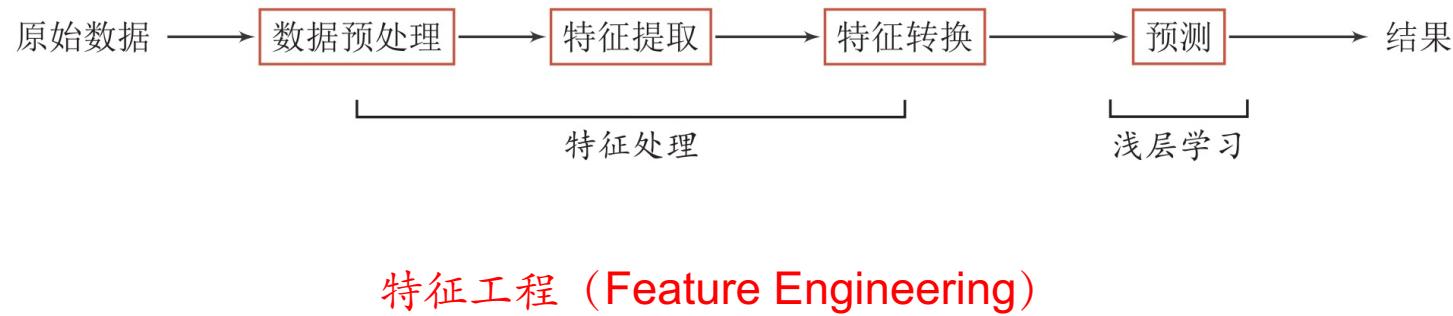
► 算法工程师

- 选择正确的模型，损失函数
 - 不同的模型，损失函数适用于解决不同的问题
- 有时很难找到最优模型
 - 有经验的算法工程师在建模和调试模型上都具有经验



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：



- 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

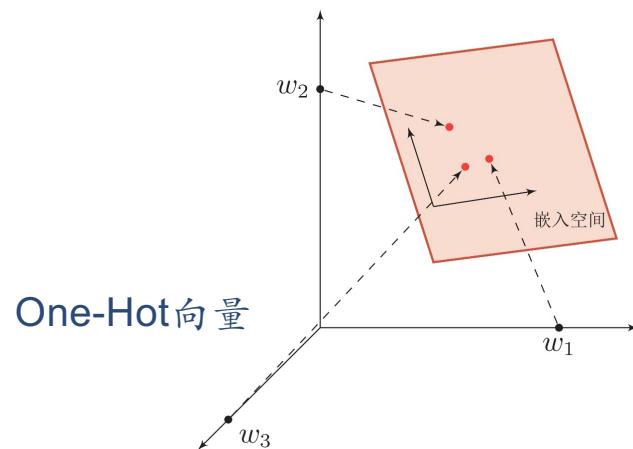
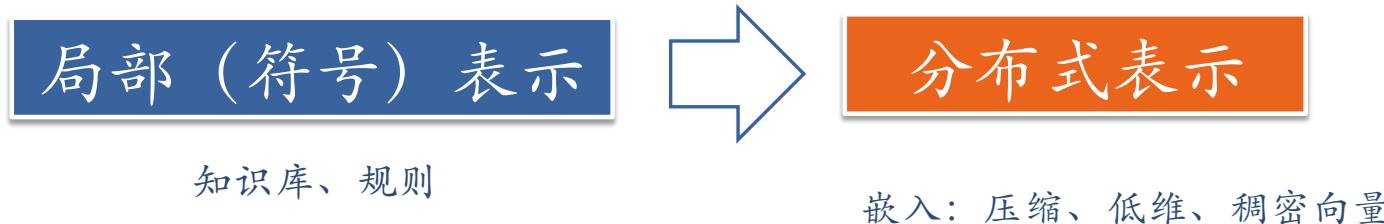
- ▶ 表示学习
 - ▶ 如何自动从数据中学习好的表示
- ▶ 难点
 - ▶ 没有明确的目标

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。

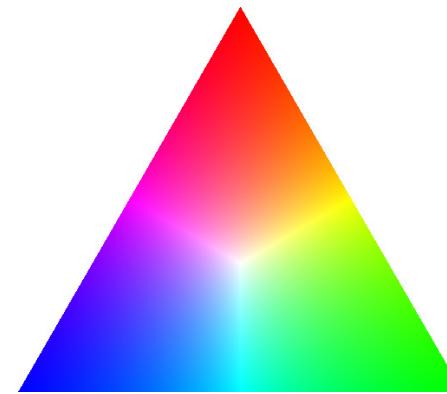
语义表示

► 如何在计算机中表示语义?



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

▶ 可解释性、用于线性模型效率较高

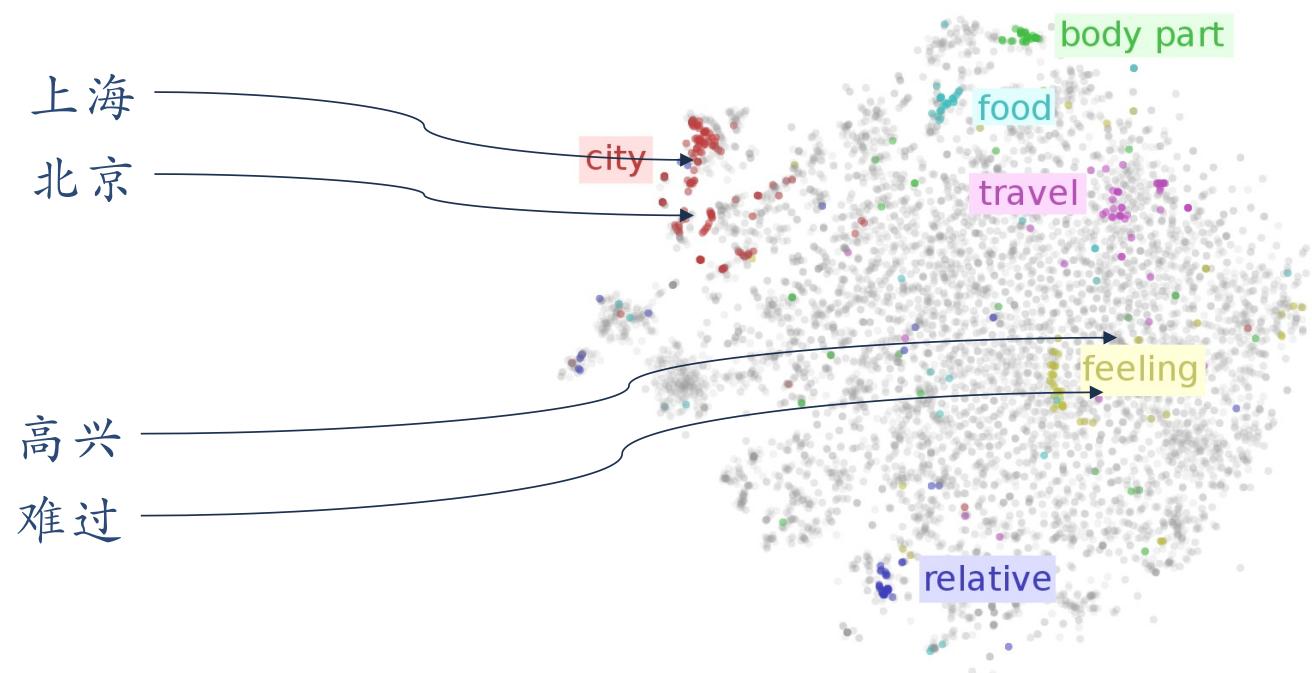
▶ 分布式(distributed)表示

▶ 压缩、低维、稠密向量

▶ 维度低，可以捕捉特征间的相似关系

	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]

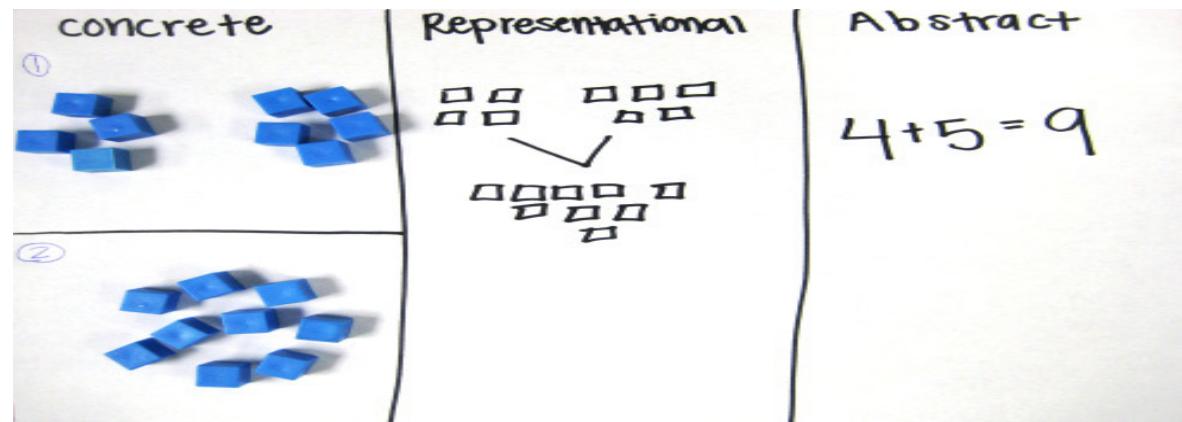
词嵌入 (Word Embeddings)



<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



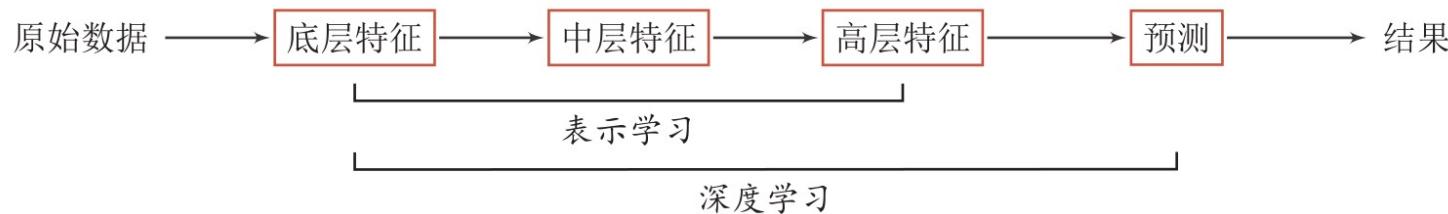
<https://mathteachingstrategies.wordpress.com/2008/11/24/concrete-and-abstract-representations-using-mathematical-tools/>

传统的特征提取

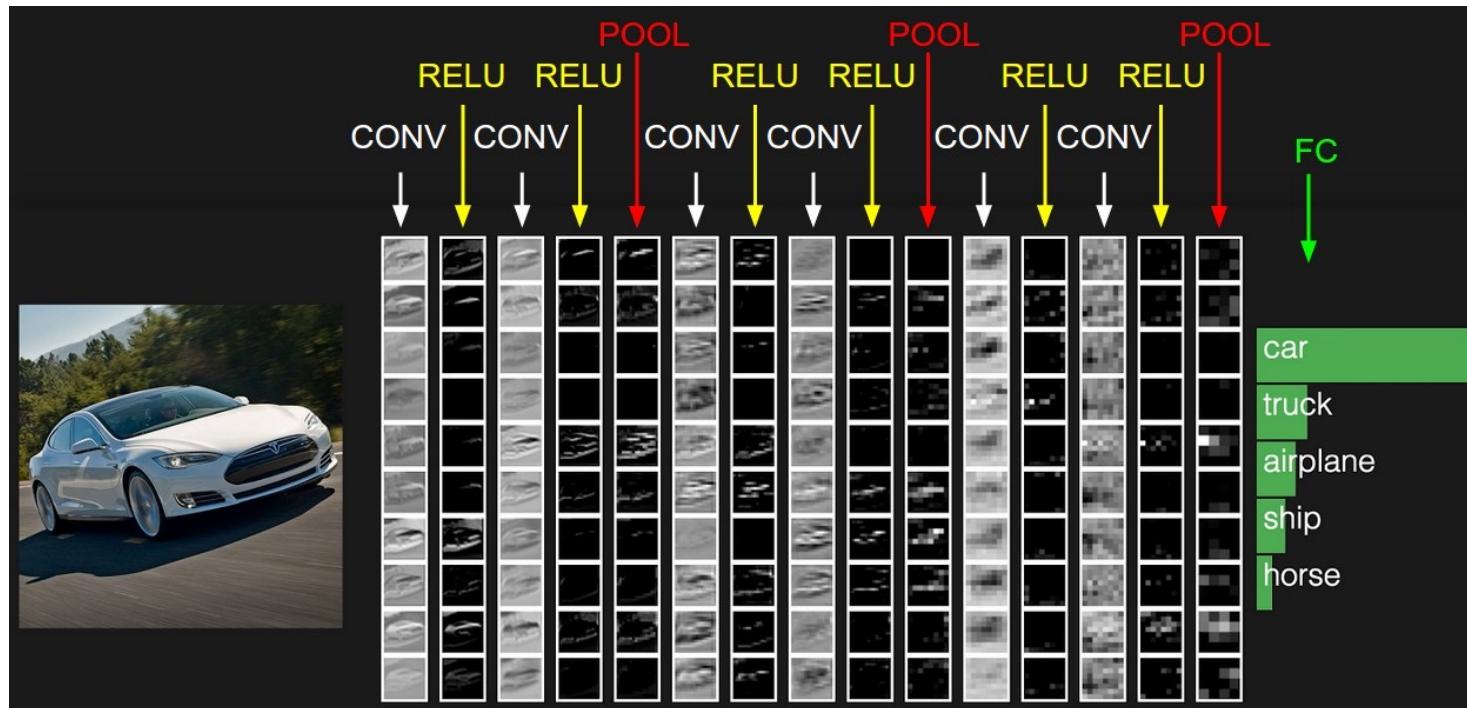
- ▶ 特征提取
 - ▶ 线性投影（子空间）
 - ▶ PCA（主成分分析）、LDA（线性判别分析）
 - ▶ 非线性嵌入
 - ▶ LLE（局部线性嵌入）、Isomap（等距映射）、谱方法
 - ▶ 自编码器
- ▶ 特征提取VS表示学习
 - ▶ 特征提取：基于任务或先验对去除无用特征
 - ▶ 表示学习：通过深度模型学习高层语义特征

深度学习

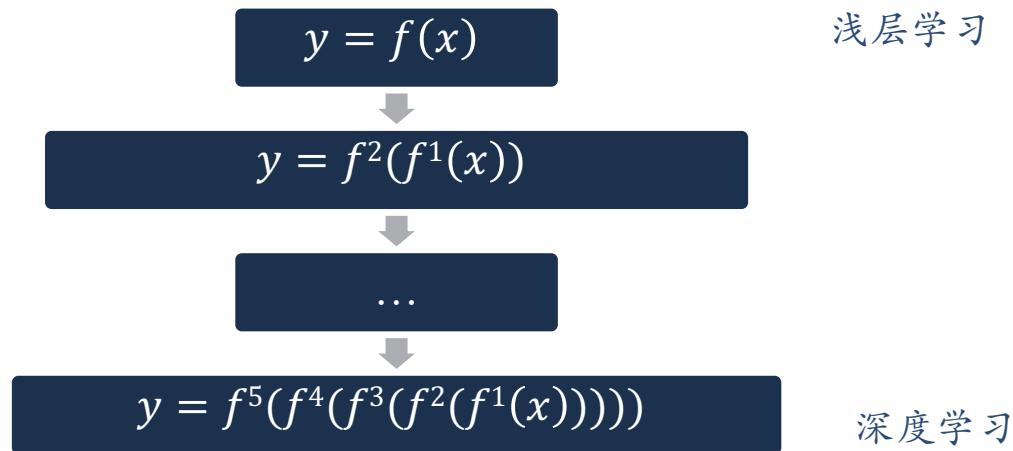
- 通过构建具有一定“深度”的模型，可以让模型来自动学习好的特征表示（从底层特征，到中层特征，再到高层特征），从而最终提升预测或识别的准确性。



表示学习与深度学习



深度学习的数学描述

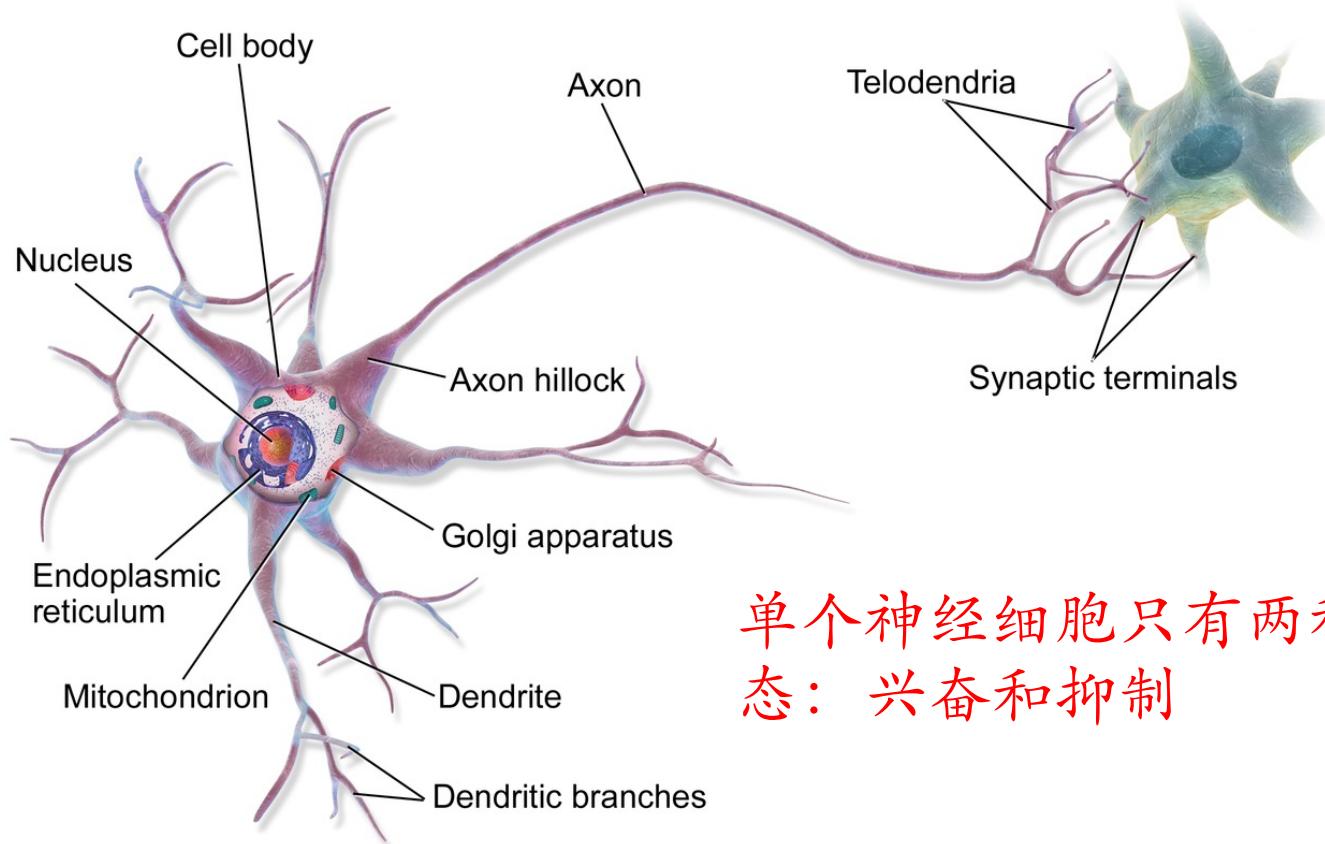




生物神经元

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？

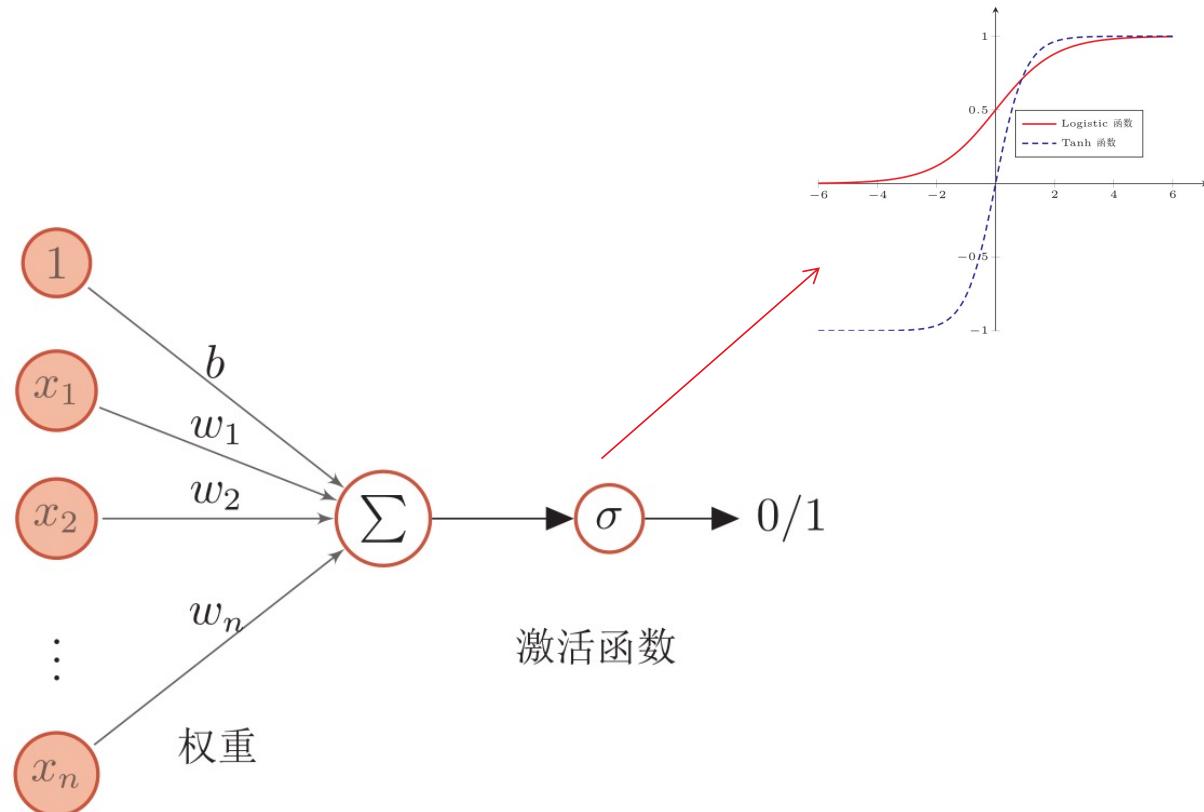
► 赫布法则 Hebb's Rule

► “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

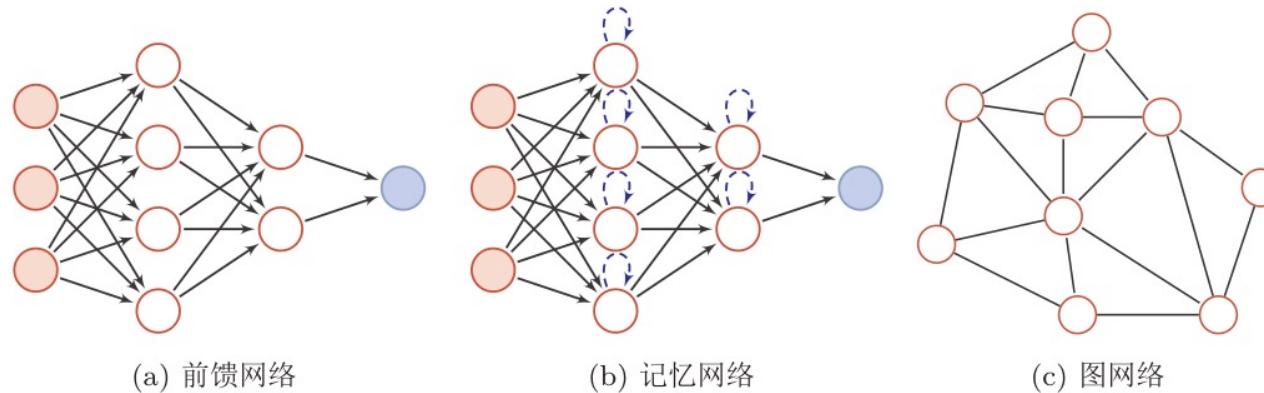


人工神经网络

- ▶ 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

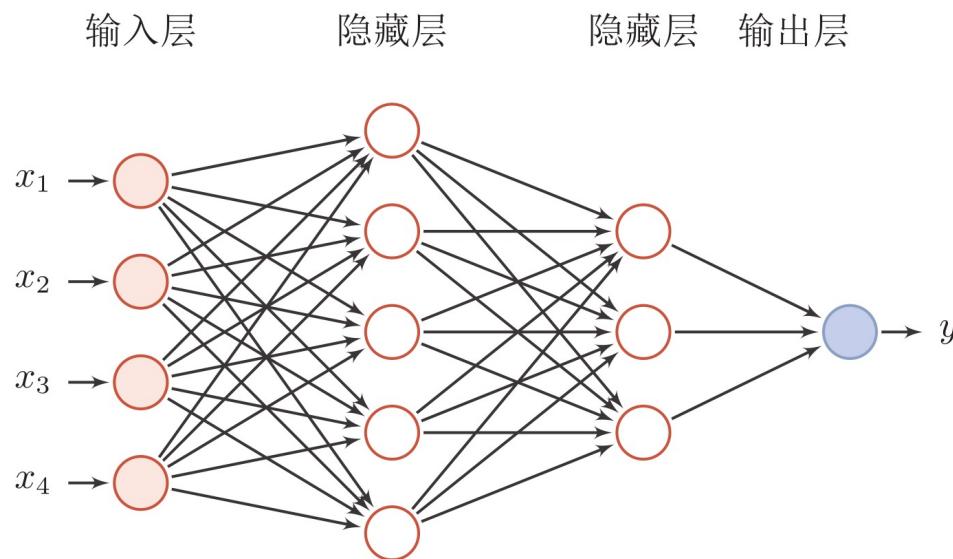
- ▶ 人工神经网络由神经元模型构成，这种由许多神经元组成的信息处理网络具有并行分布结构。
- ▶ 虽然这里将神经网络结构大体上分为三种类型，但是大多数网络都是复合型结构，即一个神经网络中包括多种网络结构。



神经网络

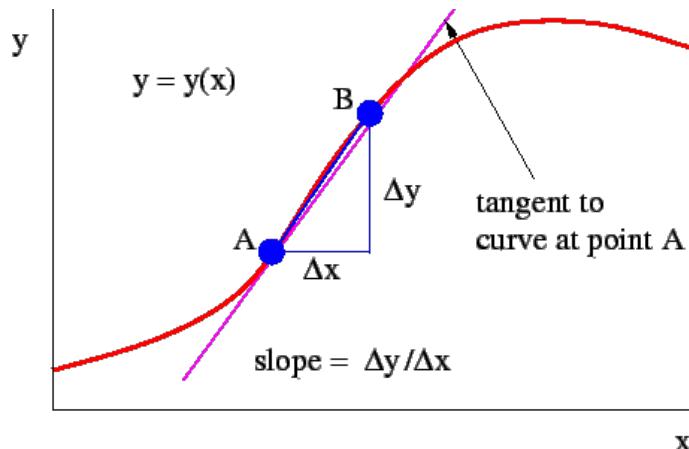
$$y = (f^3(f^2(f^1(x))))$$

$$f^l(x) = \sigma(W^l x)$$



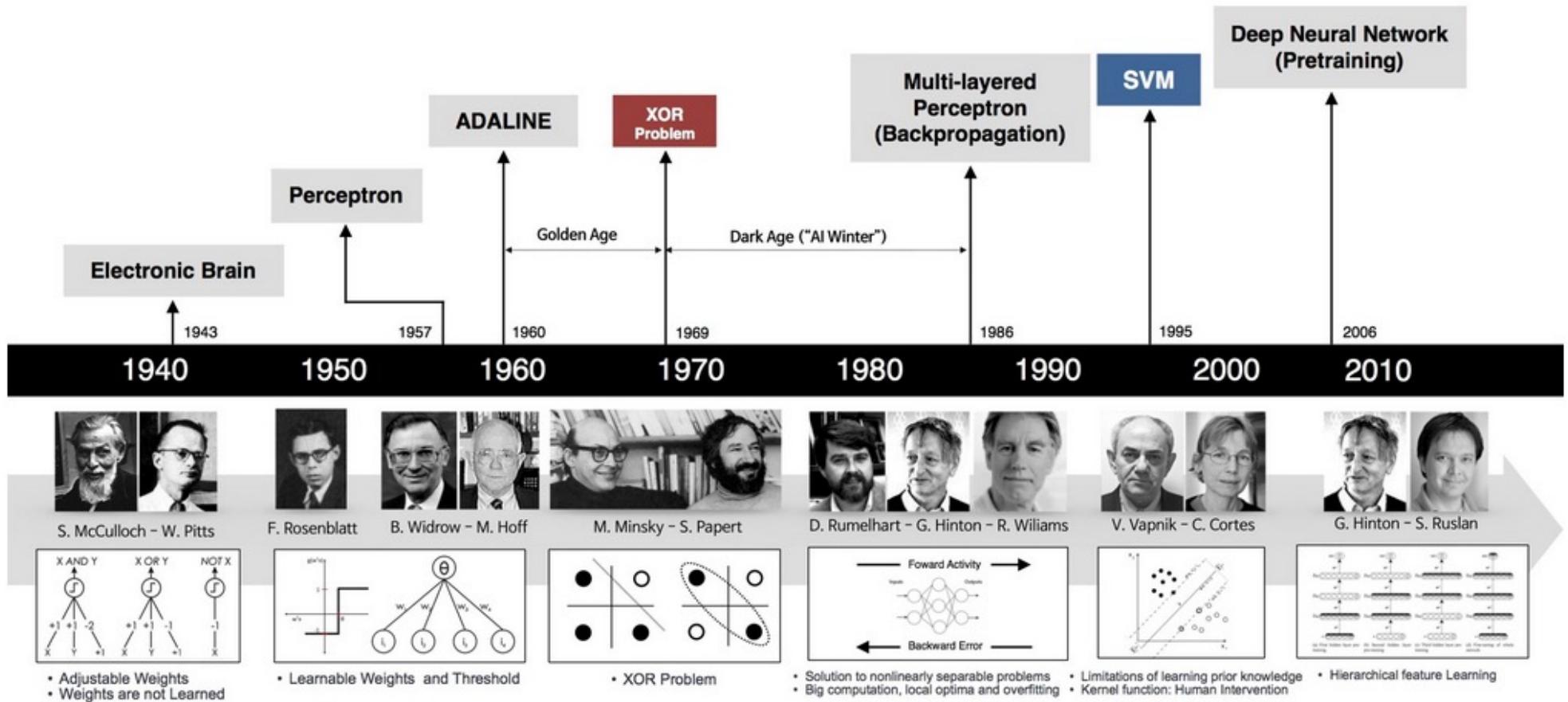
如果解决贡献度分配问题？

► 偏导数



► 贡献度

$$\frac{\partial y}{\partial x} = \frac{y(x + \Delta x) - y(x)}{\Delta x}$$



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 (XOR)。
 - 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 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

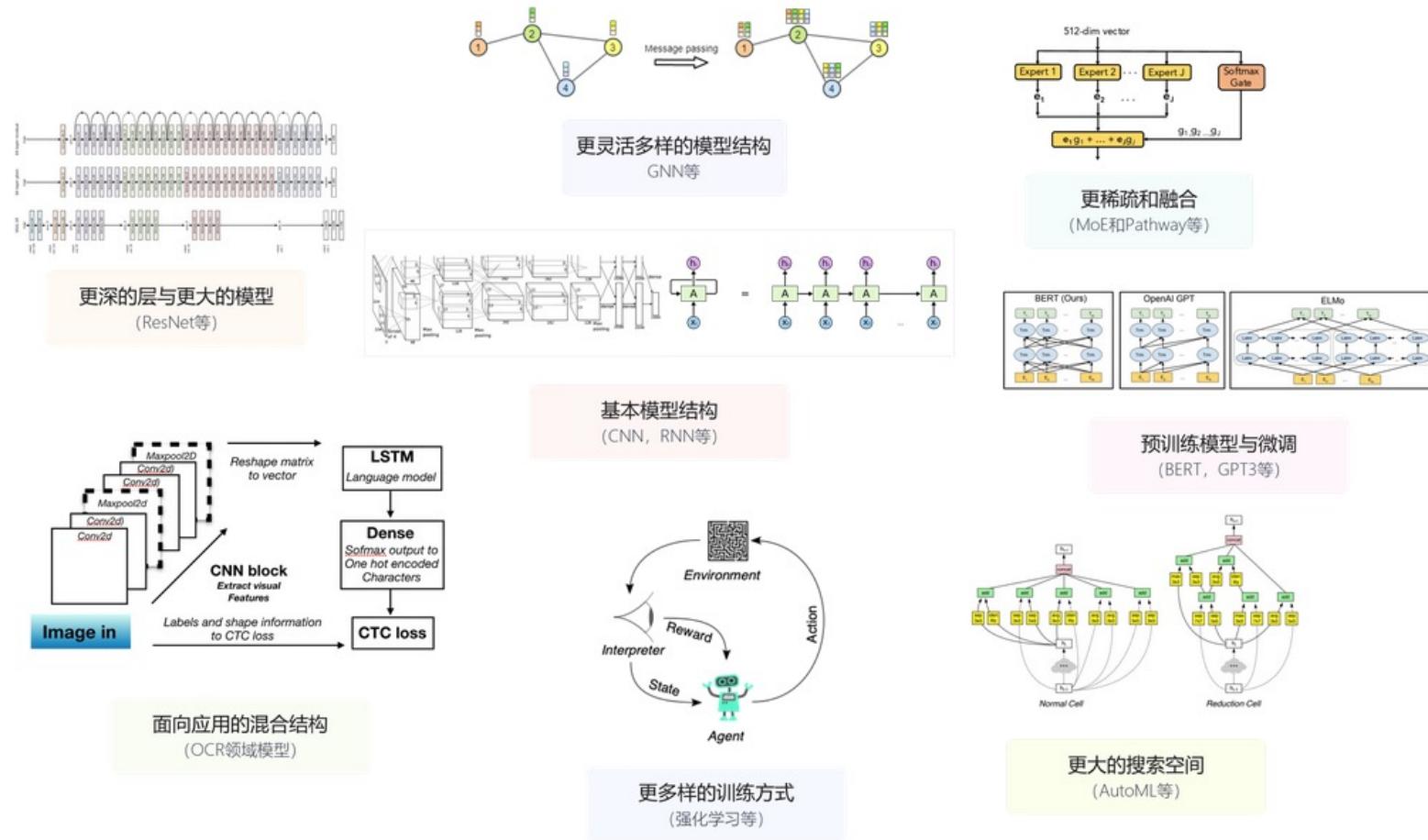
► 第四阶段：流行度降低

- 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。
- 2023年，ChatGPT：OpenAI提出的大型语言模型，百模大战。



 NOBELPRISET I FYSIK 2024
THE NOBEL PRIZE IN PHYSICS 2024

KUNGL.
VETENSKAPS-
AKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES


John J. Hopfield
Princeton University, NJ, USA


Geoffrey E. Hinton
University of Toronto, Canada

"för grundläggande upptäckter och uppfinningar som möjliggör maskininlärning med artificiella neuronnätverk"
"for foundational discoveries and inventions that enable machine learning with artificial neural networks"

KUNGL.
VETENSKAPS-
AKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

KUNGL.
VETENSKAPS-
AKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

ELLEN MOONS
HANS ELLEGREN
ANDERS HOLMÅK



深度学习三巨头共同获得2018年度图灵奖

日期：2019年03月29日 来源：news.sciencenet.cn

【大 中 小】 【打印】 【关闭】



【网易智能讯 3月27日消息】今日晚间，ACM（国际计算机学会）宣布，有“深度学习三巨头”之称的Yoshua Bengio、Yann LeCun、Geoffrey Hinton共同获得了2018年的图灵奖，这是图灵奖1966年建立以来少有的一年颁奖给三位获奖者。ACM同时宣布，将于2019年6月15日在旧金山举行年度颁奖晚宴，届时正式给获奖者颁奖，奖金100万美元。

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



谢 谢