

人工智能概要

马兴军

复旦大学 计算机科学技术学院

2022年2月26日

课程介绍

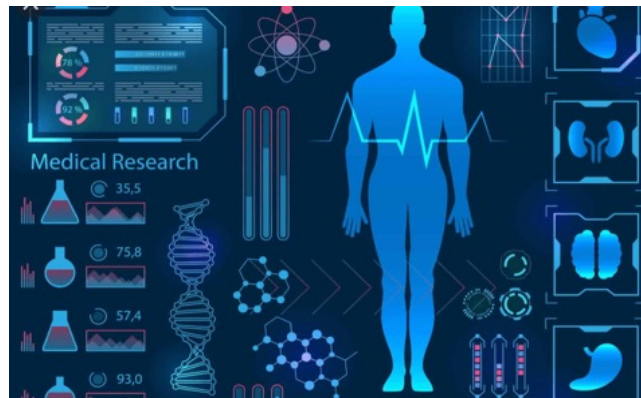
课程网站：<https://highschool.opentai.org/>



人工智能



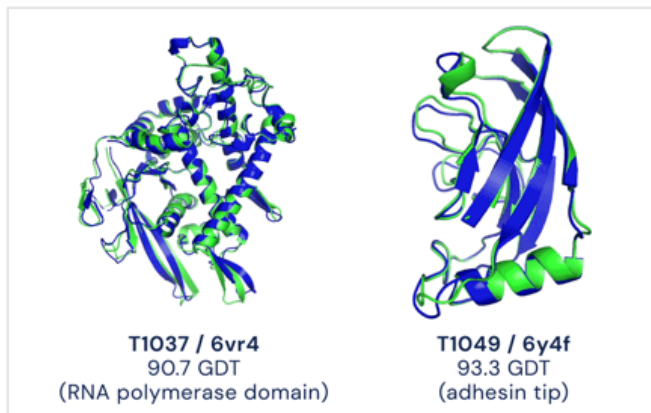
智慧教育



智慧医疗



自动驾驶



生物信息



智能制造



智慧金融

人工智能 | 定义

1956年8月的达特茅斯(Dartmouth)会议上，
“**人工智能(Artificial Intelligence, AI)**” 这个概念被提出并作为本研究领域的名称。

- “人工智能就是机器展现出现的智能。”
——维基百科
- “人工智能是制造智能机器，尤其是智能计算机程序的科学技术。”
——图灵
- “人工智能是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。”
——《人工智能标准化白皮书（2018版）》



达特茅斯会议主要参会者合影

人工智能 | 定义

➤ 人工智能就是**让机器拥有人类的智能**。

- “计算机控制”
- “智能行为”

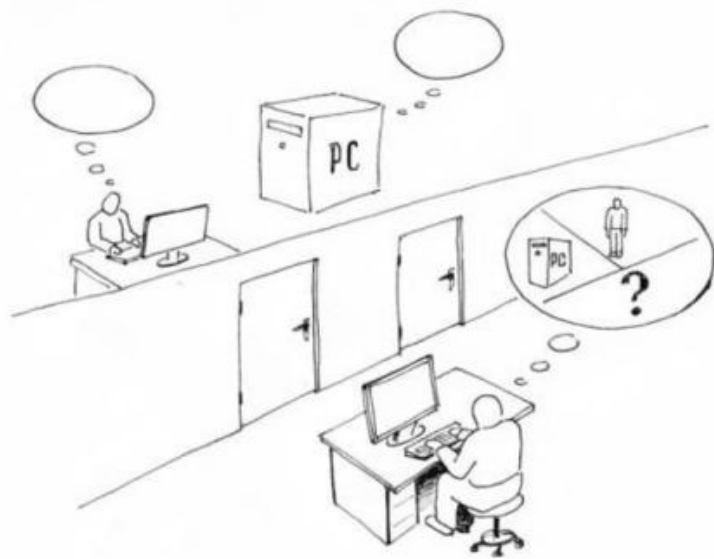
人工智能就是要让机器的行为看起来
就像是人所表现出的智能行为一样。
——约翰·麦卡锡（人工智能之父）



约翰·麦卡锡(John McCarthy)

图灵测试

- 提出目的：为了对“智能”做出定义
- 内容：测试者与被测试者（一个人和一台机器）隔开的情况下，通过一些装置（如键盘）向被测试者随意提问。进行多次测试后，如果机器让平均每个参与者做出超过30%的误判，那么这台机器就通过了测试，并被认为具有人类智能。



图灵测试示意图



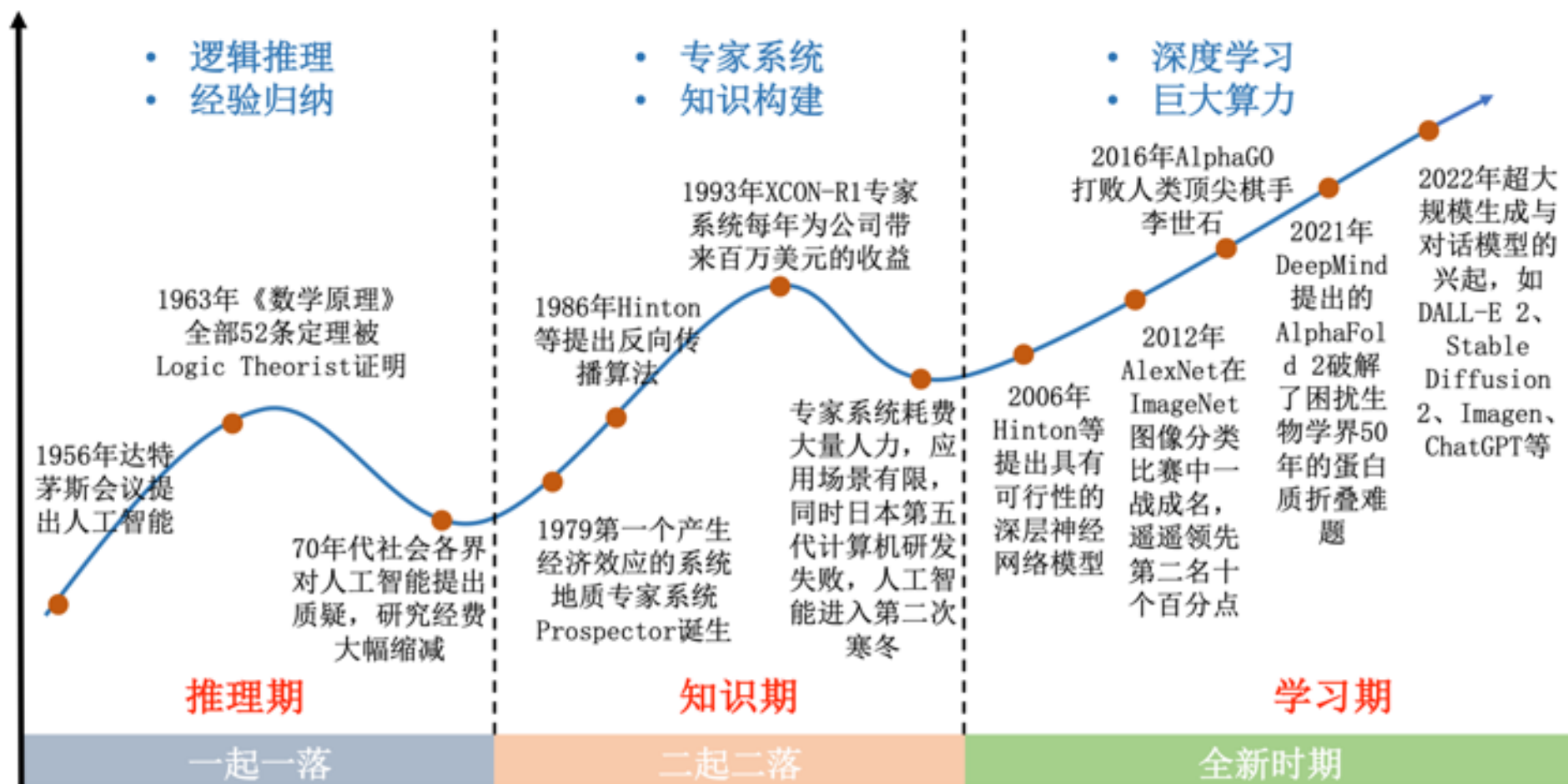
艾伦·麦席森·图灵

人工智能 | 研究领域

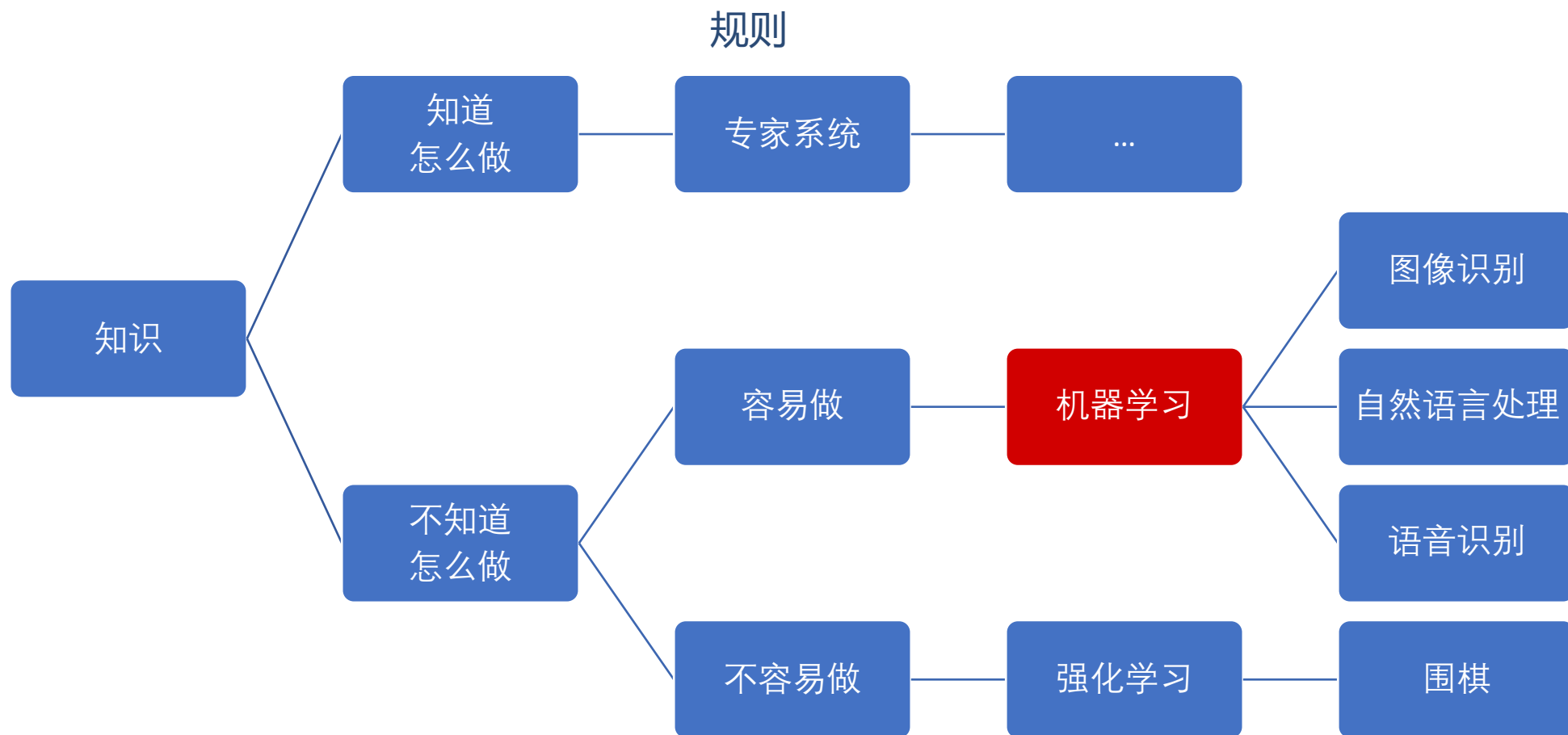
- 让机器具有人类的智能
 - 机器感知（计算机视觉、语音信息处理）
 - 学习（模式识别、机器学习、强化学习）
 - 语言（自然语言处理）
 - 记忆（知识表示）
 - 决策（规划、数据挖掘）



人工智能 | 发展历史



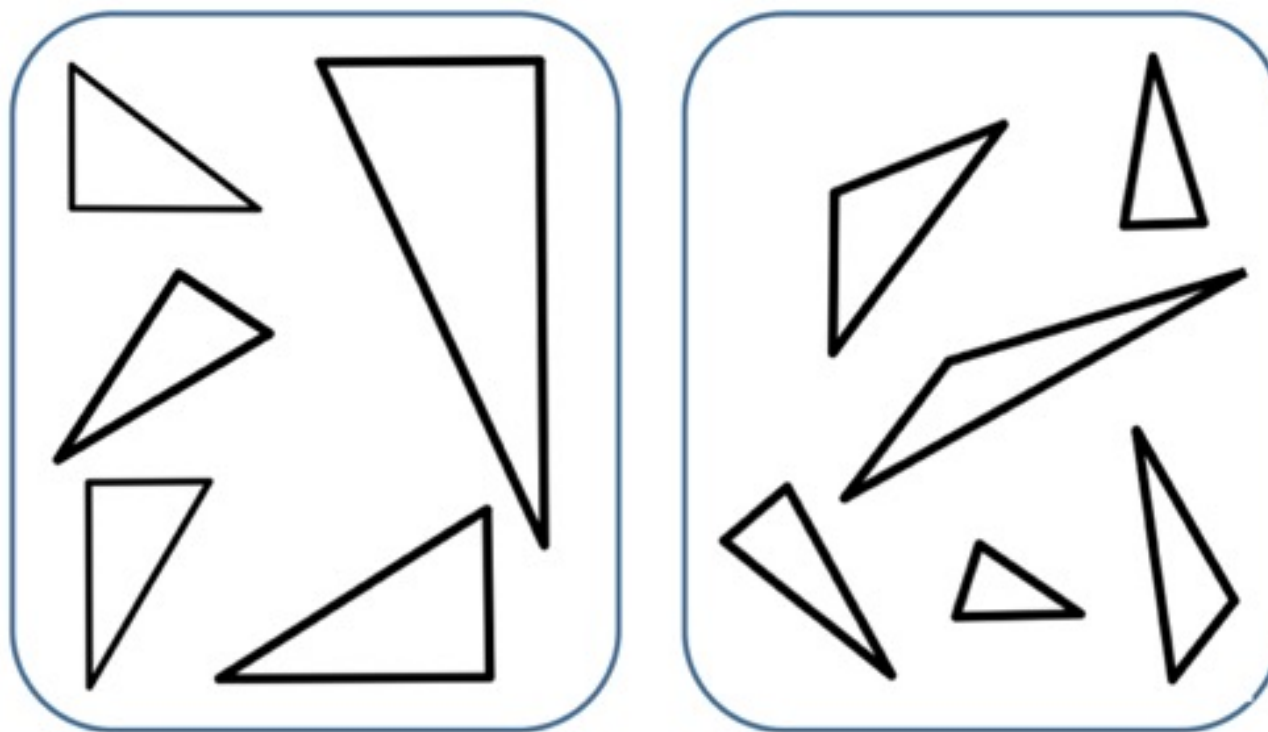
如何开发一个人工智能系统？



如何开发一个人工智能系统？

➤ 专家知识（人工规则）

采用什么规则？



首款聊天机器人Eliza

1966年，世界上首款聊天机器人伊莉莎（Eliza）横空出世。伊莉莎可以用英语和测试者进行交谈，甚至可以给人以夸赞和安慰。其背后是基于**关键词匹配**的。



姑娘：男人们全都一样。
伊莉莎：在哪些方面？
姑娘：他们总找借口躲避我们
伊莉莎：你能举个例子吗？
姑娘：好吧，我的男朋友要我到你这里来。
伊莉莎：哦，你的男朋友要你到我这里来。
姑娘：他说我很烦闷。
伊莉莎：你很烦闷，这使我很难过。

Eliza与一位姑娘的谈话摘要



深蓝击败国际象棋冠军

1997年，由 IBM 制造的超级计算机 “深蓝(Deep Blue)” 在标准的国际象棋比赛中再次挑战当时的国际象棋世界冠军卡斯帕罗夫，并以 3.5:2.5 战胜卡斯帕罗夫，成为第一个在标准比赛时间内击败国际象棋世界冠军的计算机系统。



“深蓝” 与国际象棋大师对决



IBM “深蓝” 超级计算机

如何开发一个人工智能系统？

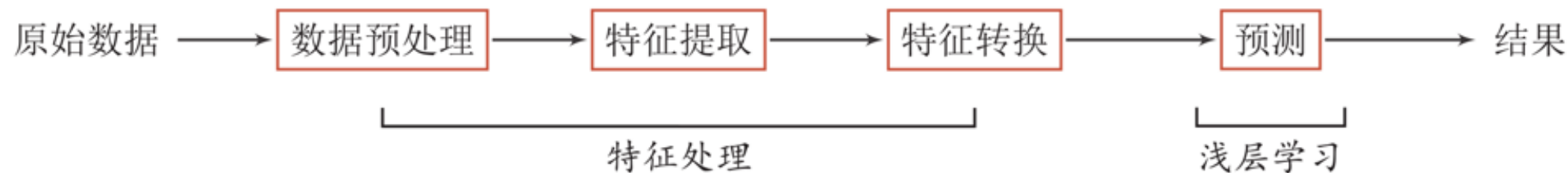


2	6	8	9	3	4	7	5	6
3	4	7	9	5	5	6	7	2
5	8	7	0	9	4	3	5	4
5	2	3	4	9	5	6	7	8

机器学习
Machine Learning , ML

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

➤ 表示学习

- 可以自动地从数据中学习出好的表示，并提高最终机器学习模型的性能

➤ 什么是好的表示？

没有明确的标准，但一般而言，一个好的表示具有以下优点：

- 很强的表示能力，同样大小的向量可以表示更多信息
- 可以使后续的学习任务变得简单，即需要包含更高层的语义信息
- 具有一般性，是任务/领域独立的。

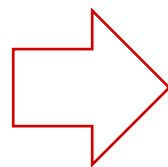


语义表示

- 如何在计算机中表示语义？

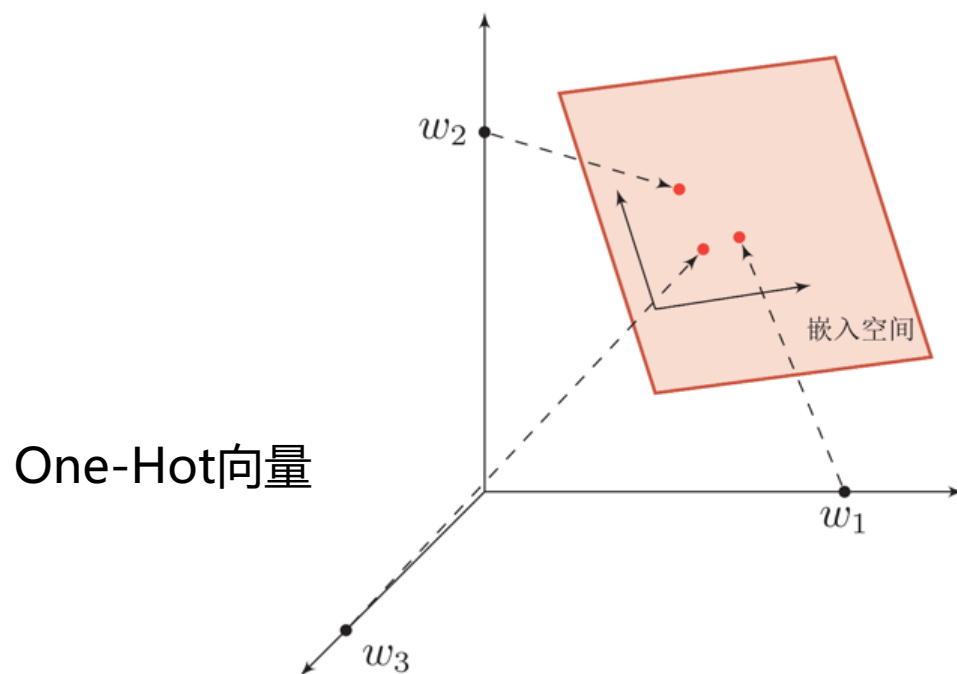
局部（符号）表示

知识库、规则



分布式表示

嵌入：压缩、低维、稠密向量



表示形式

➤ 局部表示

- 离散表示、符号表示
- One-hot 向量

➤ 分布式表示

- 压缩、低维、稠密向量
- 用 $O(N)$ 个参数表示 $O(2^k)$ 区间，其中 k 为非零参数， $k < N$

	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]



例子：颜色表示

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



➤ 局部表示

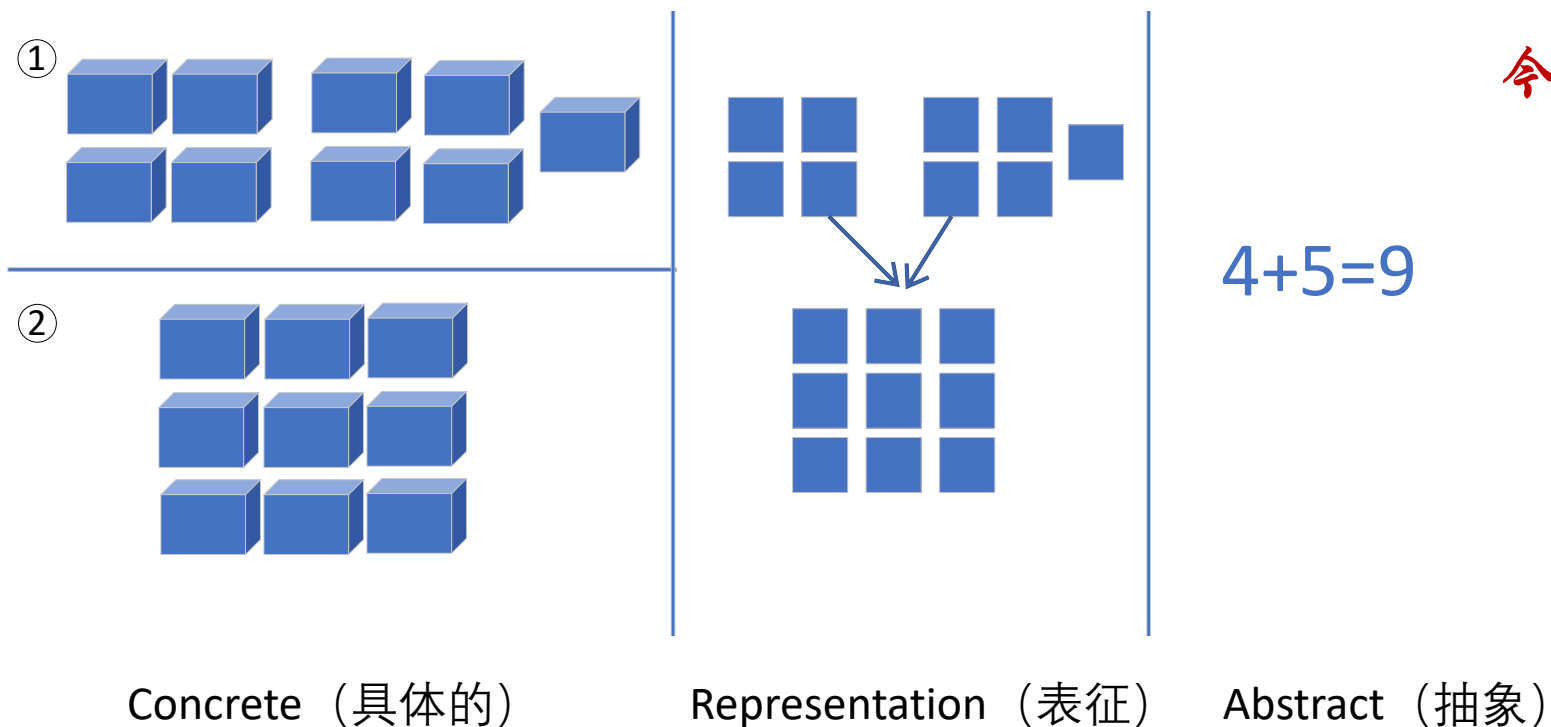
- One-hot向量维数很高。若有新的颜色，还需要添加一维才能表示
- 无法表示数据之间的关联。

➤ 分布式表示

- 低维
- 每个数据的向量表示包含了数据间的相互关联信息

表示学习与深度学习

- 一个好的表示学习的关键是构建具有一定深度的多层特征表示
- 特征重用，指数级地增加表示能力
 - 抽象表示与不变性

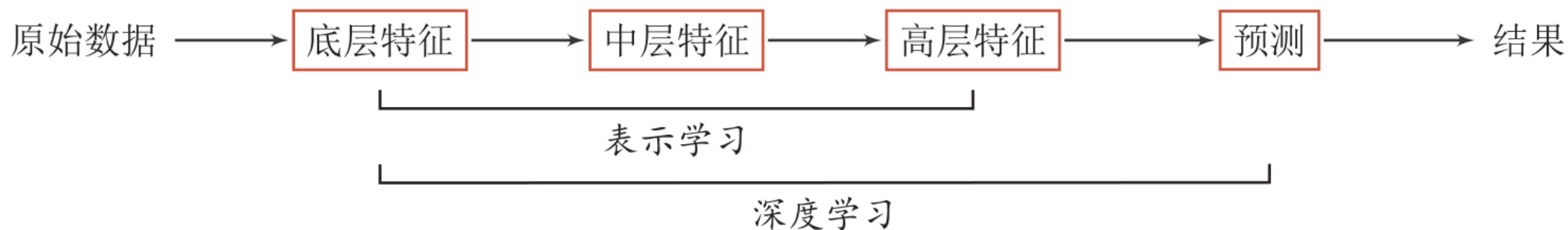


今天你搬了几块砖?



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”自动学习深层特征，无需人工特征

“端到端”是指从输入直接到结果



传统的特征提取

➤ 特征提取

- 尺度不变特征变换 (SIFT)
- 加速鲁棒特征 (SURF)
- 方向梯度直方图 (HoG)

➤ 特征提取 VS 表示学习

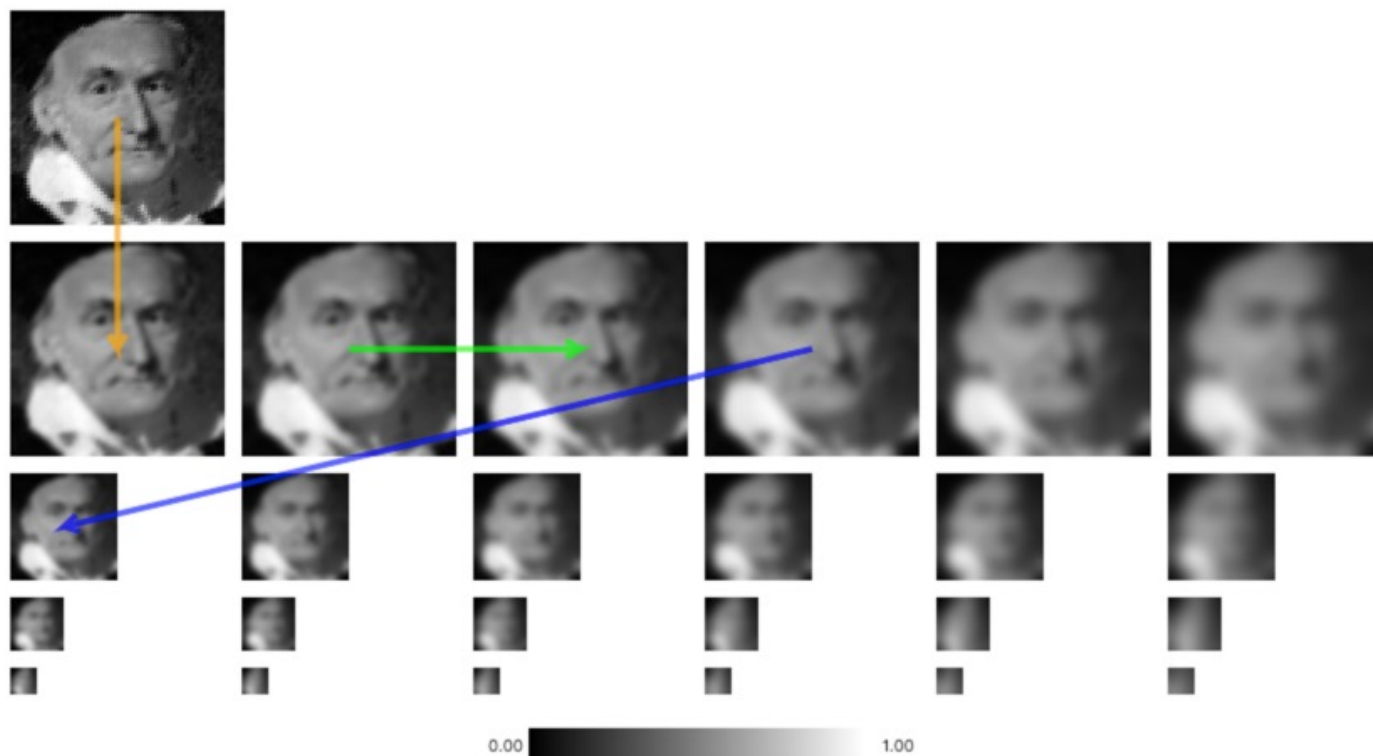
- 特征提取：基于任务或先验对去除无用特征
- 表示学习：通过深度模型学习高层语义特征



传统的特征提取



[This page should work with recent desktop versions of [Chrome](#) and [Firefox](#). I haven't tested with other browsers.]



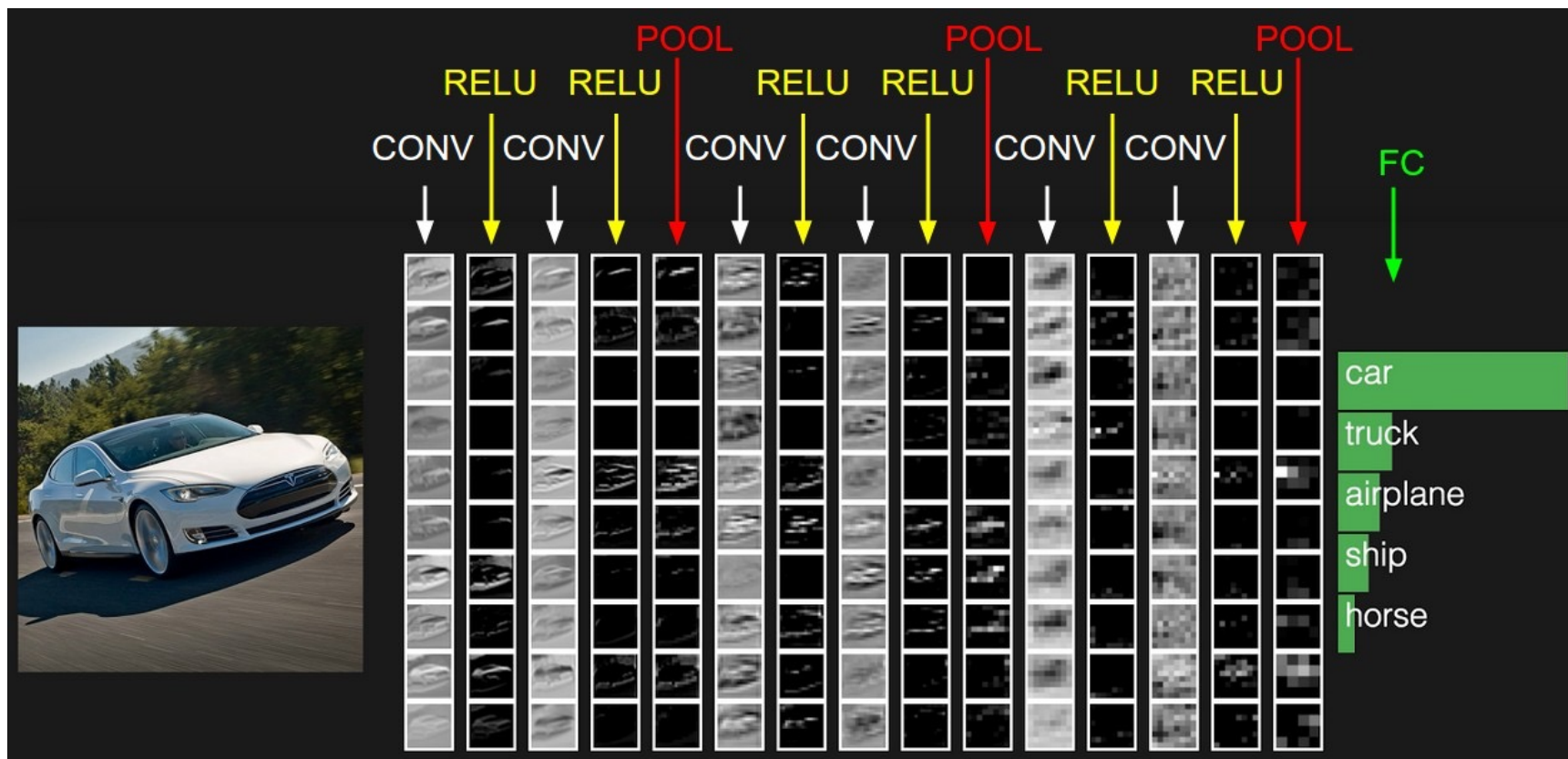
尺度不变特征变换（SIFT）提取：

- 将图像变成64x64大小后再增大一倍（第一行）
- 执行多次模糊化（第二行）
- 从第二行中取一个再模糊化
- 重复上述两步到无法进行

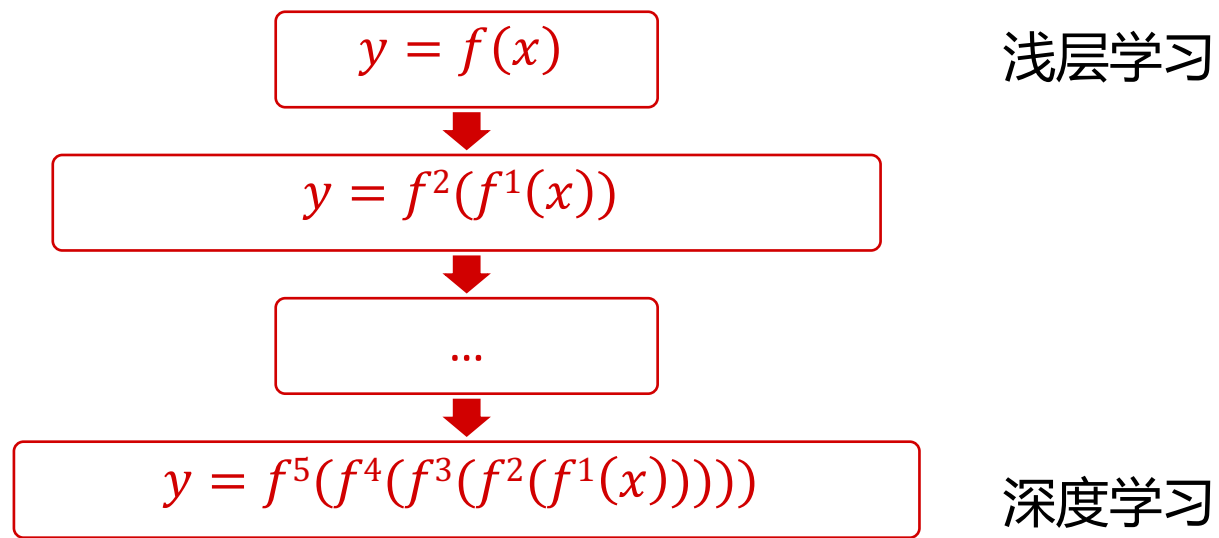
模拟
不同
观测
距离



表示学习与深度学习



深度学习的数学描述

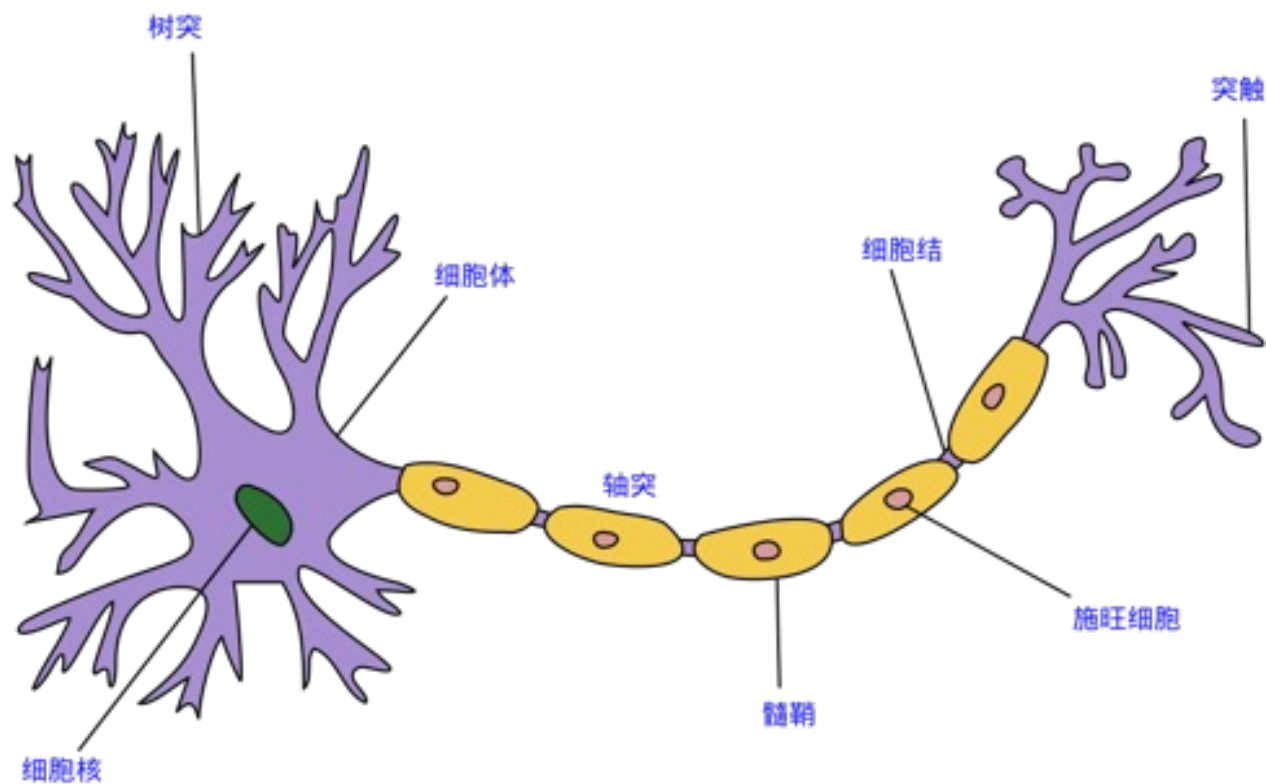


$f^1(x)$ 为非线性函数，不一定连续。

当 $f^l(x)$ 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**。

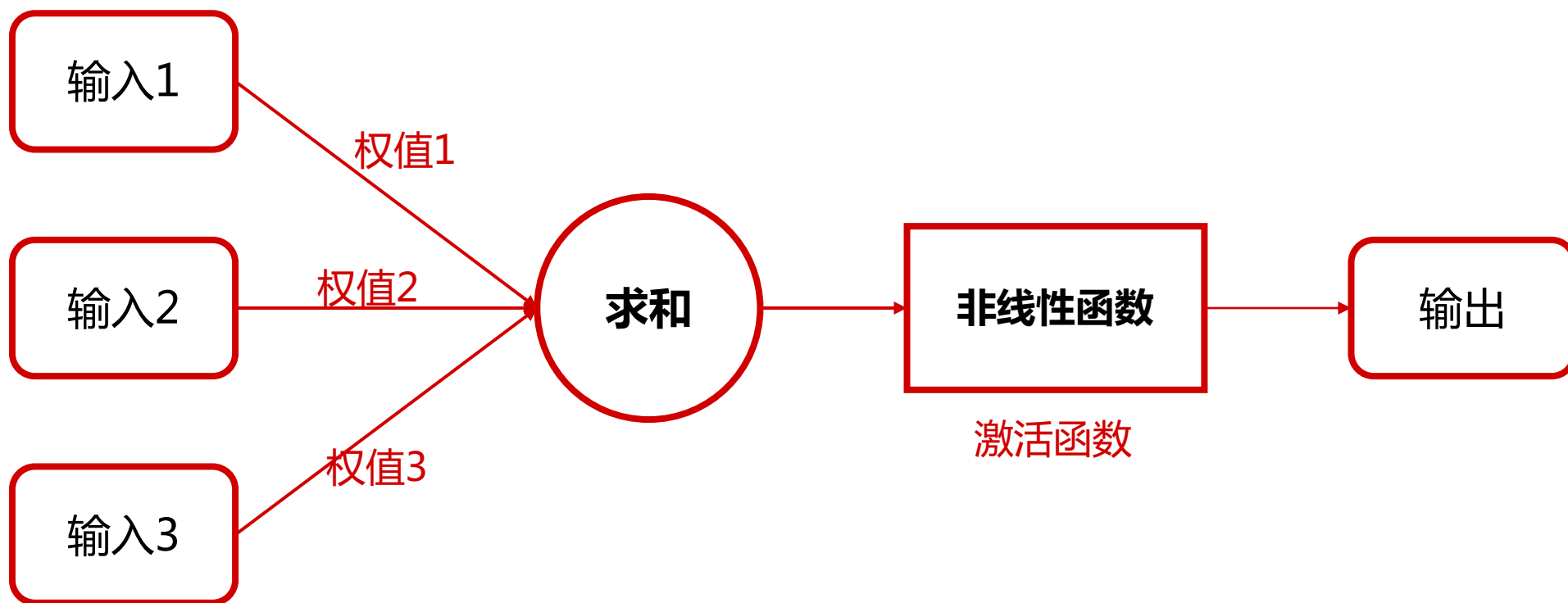


生物神经元



一个神经元通常具有多个**树突**，主要用来接受传入信息；而**轴突**只有一条，轴突尾端有许多轴突末梢可以给其他多个神经元传递信息。轴突末梢跟其他神经元的树突产生连接，从而传递信号。这个连接的位置在生物学上叫做“**突触**”。

人工神经元



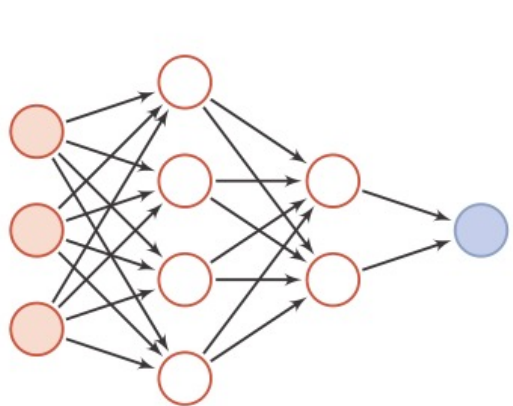
人工神经网络

- 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。

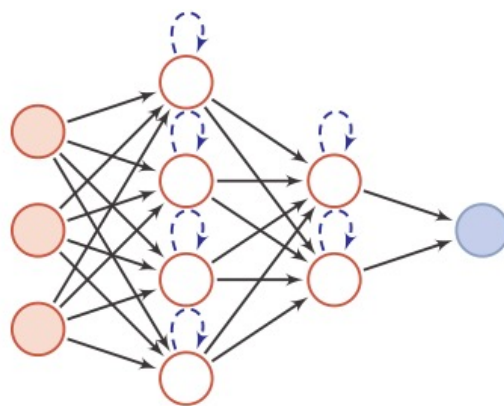


人工神经网络

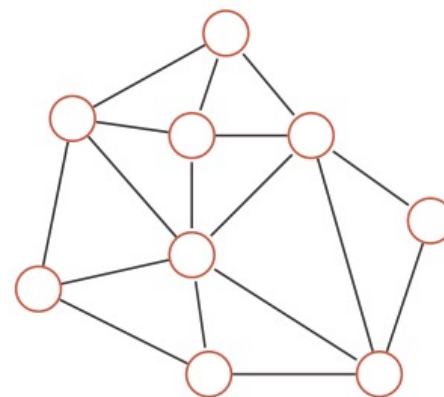
- 人工神经网络由神经元模型构成，这种由许多神经元组成的信息处理网络具有并行分布结构。
 - 虽然这里将神经网络结构大体上分为三种类型，但是大多数网络都是**复合型结构**，即一个神经网络中包括多种网络结构。



(a) 前馈网络



(b) 记忆网络



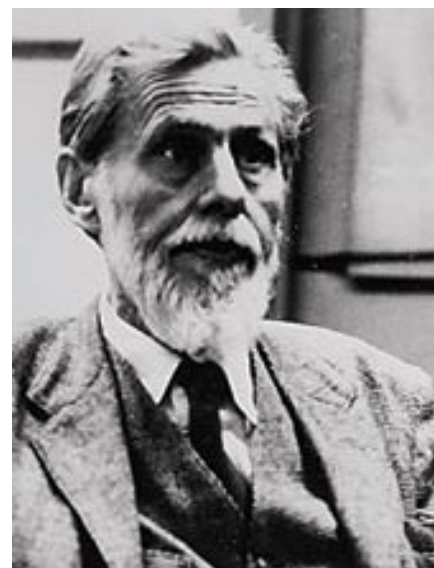
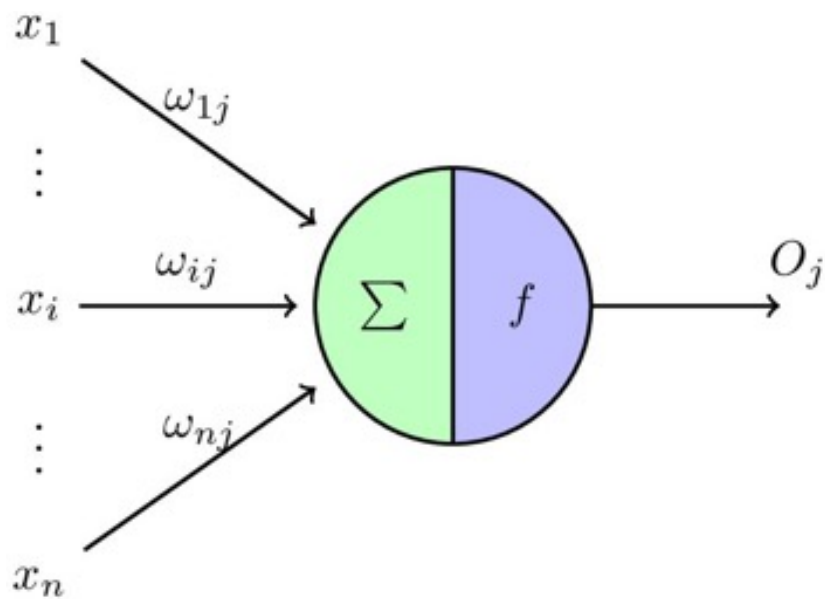
(c) 图网络

神经网络发展史 | 模型提出

➤ 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，沃伦·麦卡洛克(Warren McCulloch)和沃尔特·皮茨(Walter Pitts)最早描述了一种理想化的人工神经网络模型：**M-P模型**。



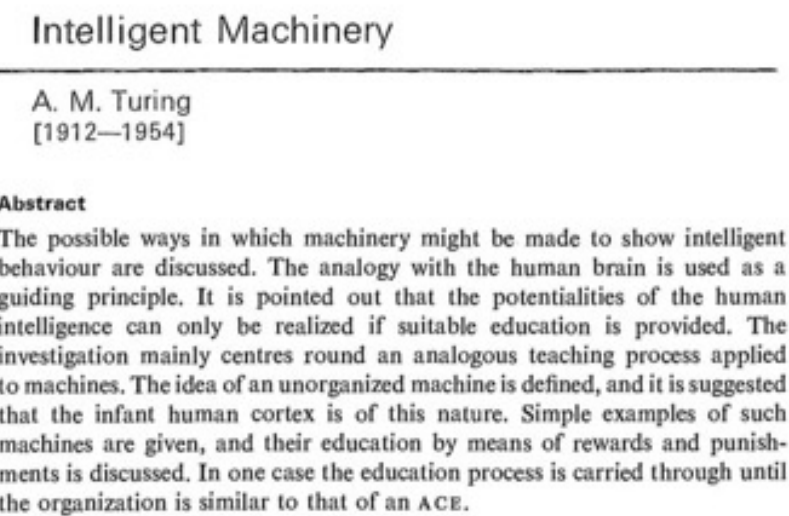
沃伦·麦卡洛克



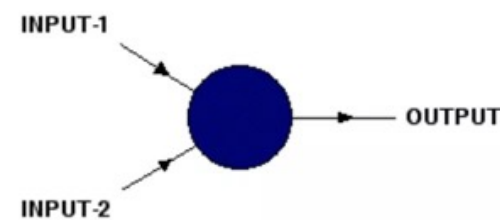
沃尔特·皮茨

神经网络发展史 | 模型提出

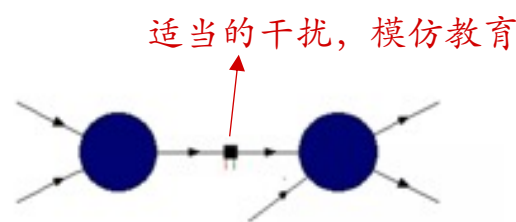
- 艾伦·图灵在1948年的论文中描述了一种 “B型无组织机器” 。



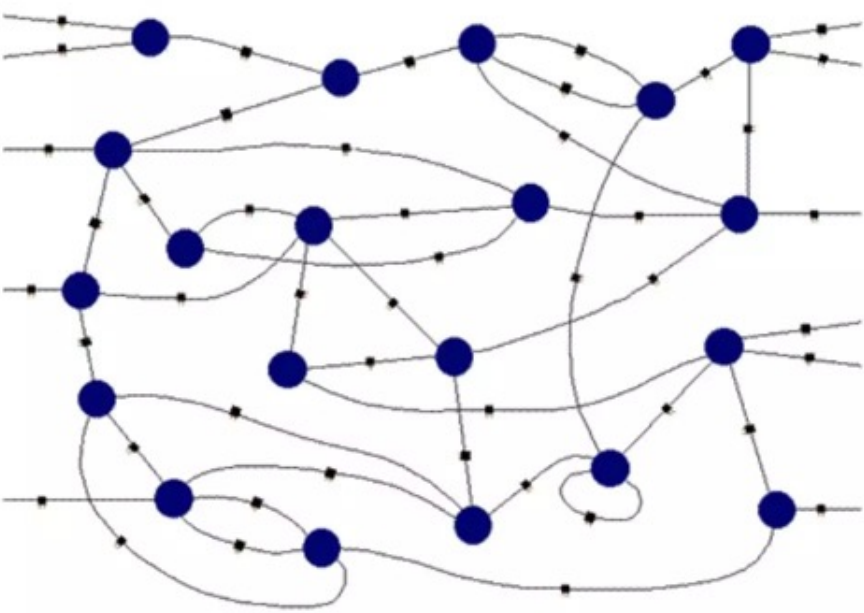
论文《智能机器》



单个神经元(与非门)



B 型网络中的两个神经元



大型初始随机 B 型网络的一部分



神经网络发展史 | 模型提出

- 1951年，麦卡洛克和皮茨的学生马文·明斯基(Marvin Minsky)建造了第一台神经网络机，称为SNARC。



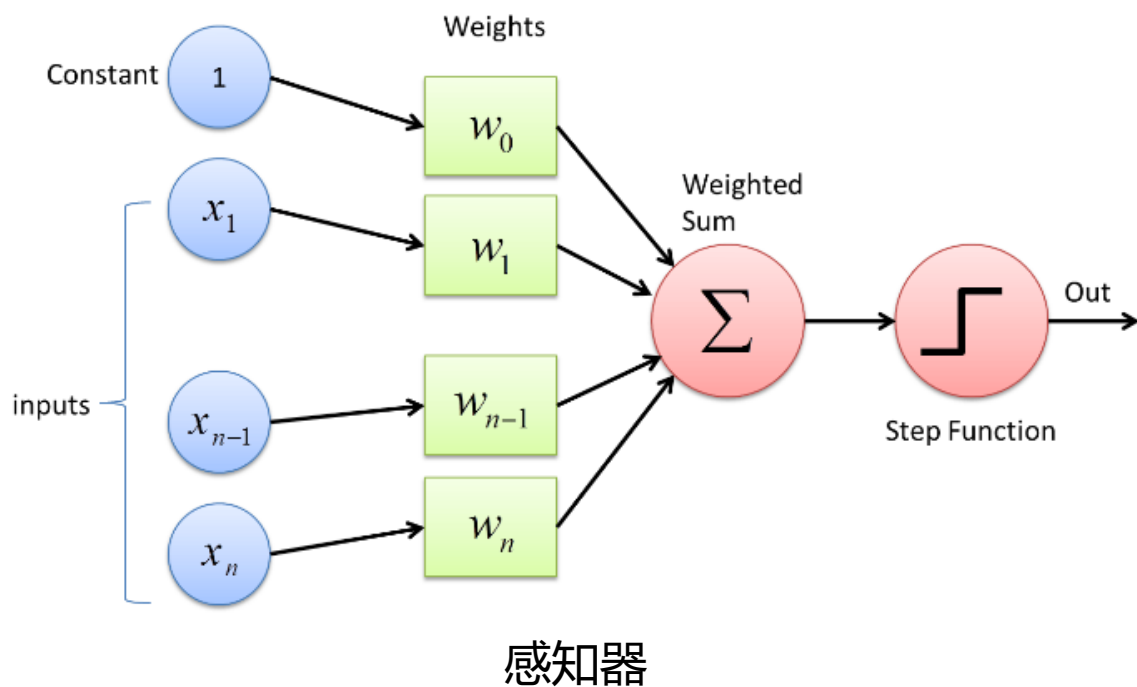
一个SNARC神经元的元件



马文·明斯基

神经网络发展史 | 模型提出

- 弗兰克·罗森布拉特(Frank Rosenblatt)最早提出可以模拟人类感知能力的神经网络模型：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。

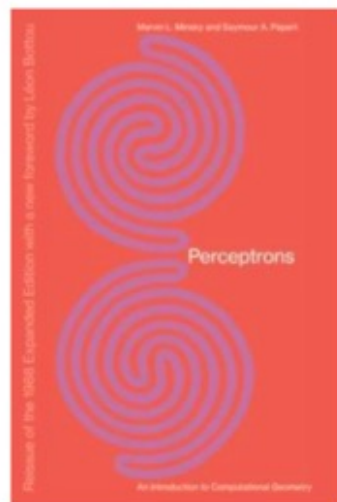


弗兰克·罗森布拉特

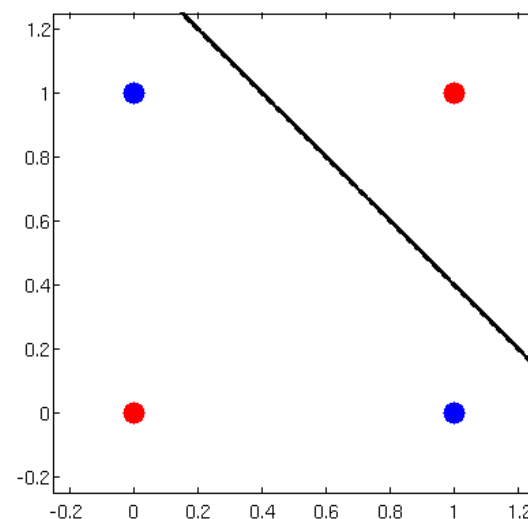
神经网络发展史 | 冰河期

➤ 第二阶段：冰河期

- 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。
 - 第一个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。



《感知机》



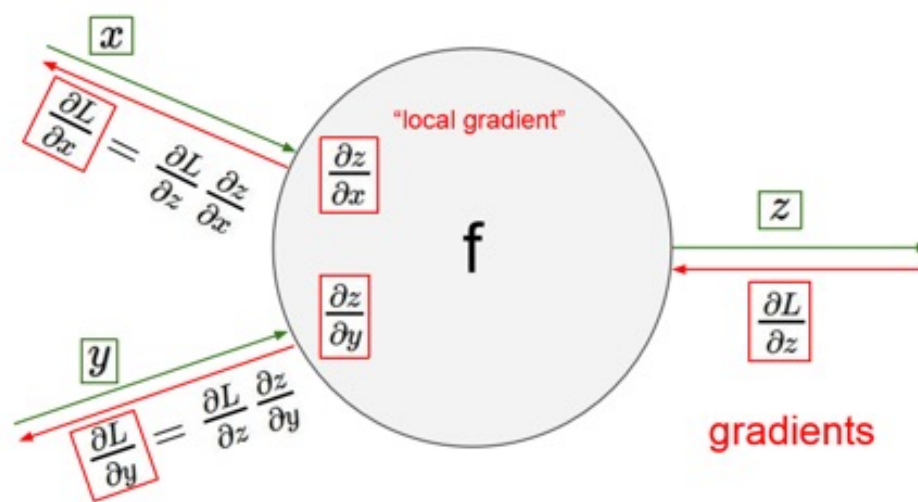
无法解决异或问题

神经网络发展史 | 冰河期

- 1974年，哈佛大学的保罗·沃伯斯(Paul Werbos)发明**反向传播算法**，但当时未受到应有的重视。
- 1980年，福岛邦彦(Kunihiko Fukushima)提出了一种带卷积和子采样操作的多层神经网络：新知网 (Neocognitron)。



保罗·沃伯斯

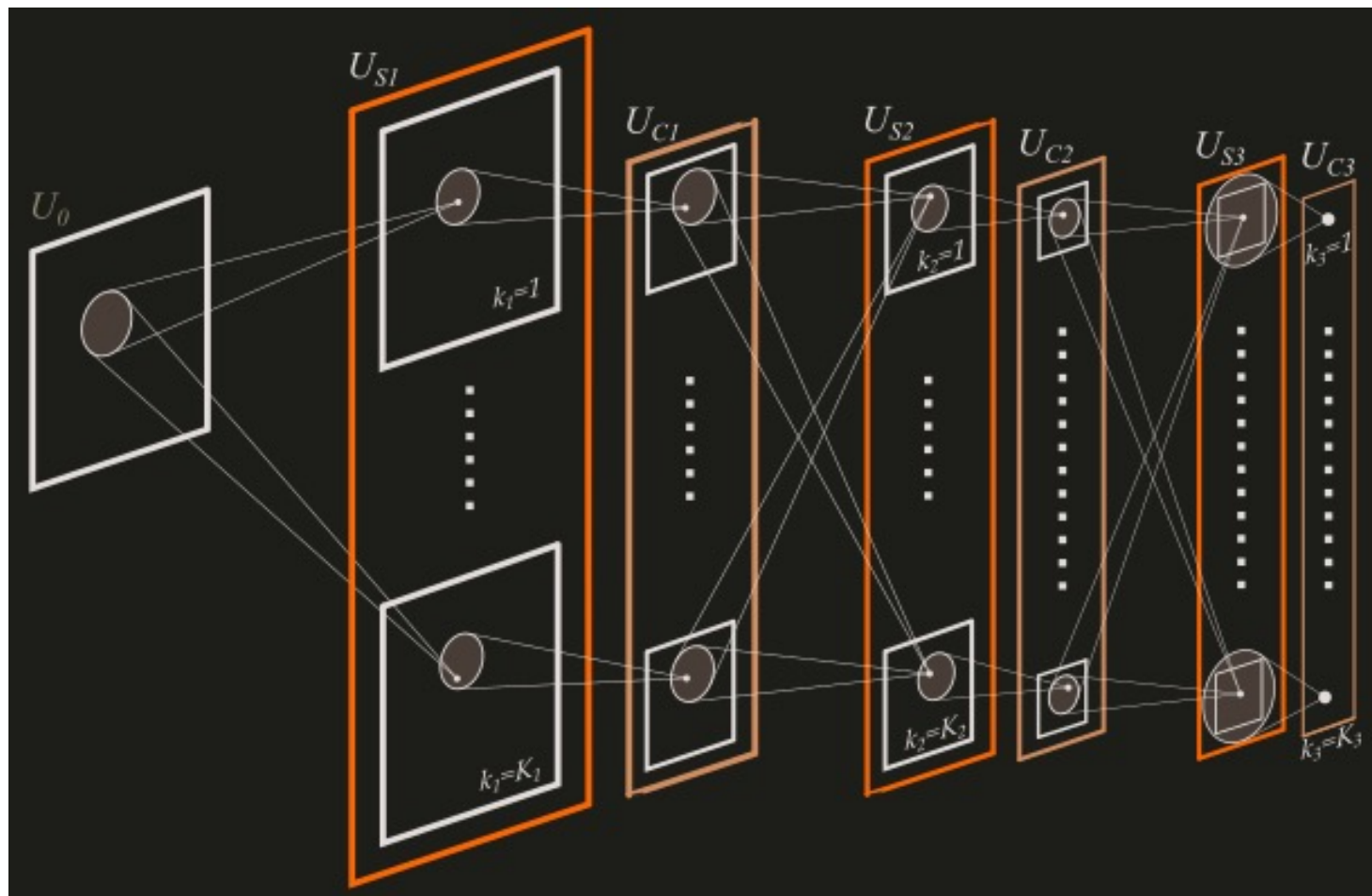


反向传播算法



福岛邦彦

神经网络发展史 | 冰河期

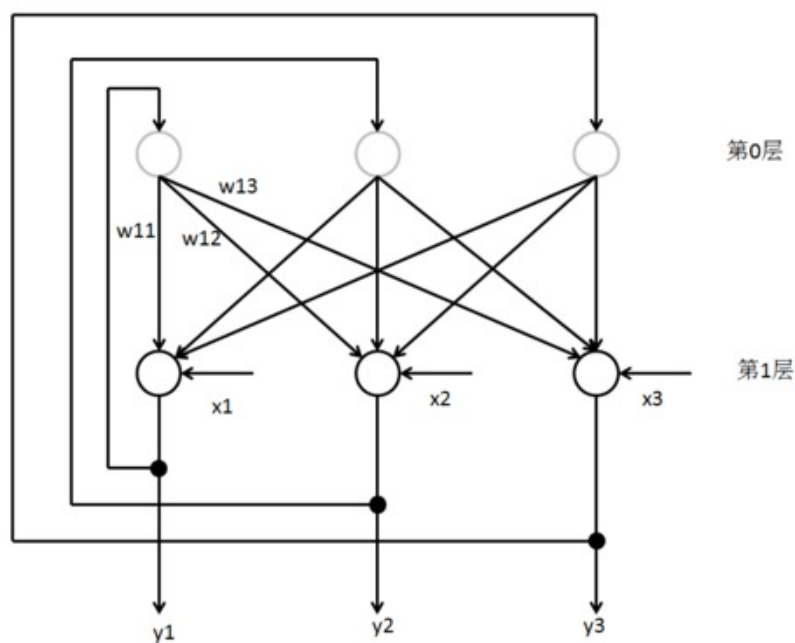


新知机结构 Neocognitron

神经网络发展史 | 反向传播引起的复兴

➤ 第三阶段：反向传播算法引起的复兴

- 1983年，加州理工学院的物理学家约翰·霍普菲尔德(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。



Hopfield网络



约翰·霍普菲尔德

神经网络发展史 | 反向传播引起的复兴

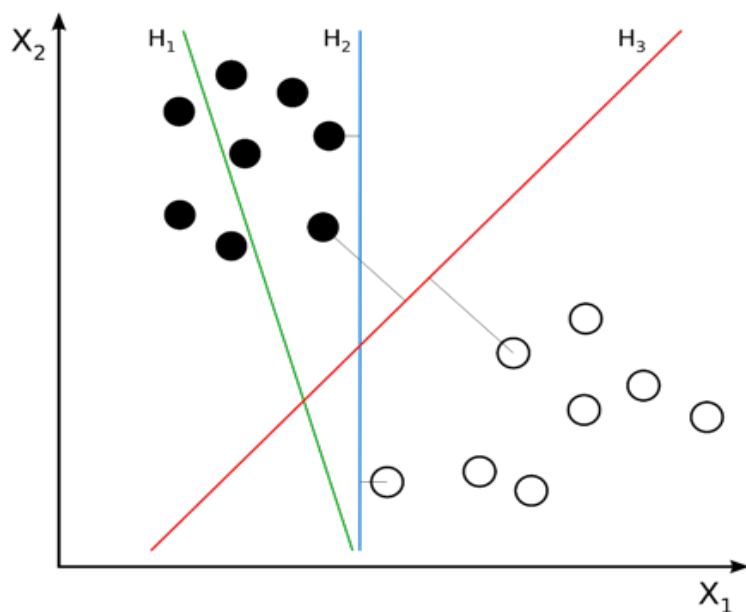
- 1984年，Geoffrey Hinton提出一种随机化版本的Hopfield网络，即玻尔兹曼机。
- 1986年，David Rumelhart和James McClelland对于联结主义在计算机模拟神经活动中的应用提供了全面的论述，并重新发明了反向传播算法。
- Geoffrey Hinton等人将反向传播算法引入到多层感知器
- LeCun等人将反向传播算法引入了卷积神经网络，并在手写体数字识别上取得了很大的成功。



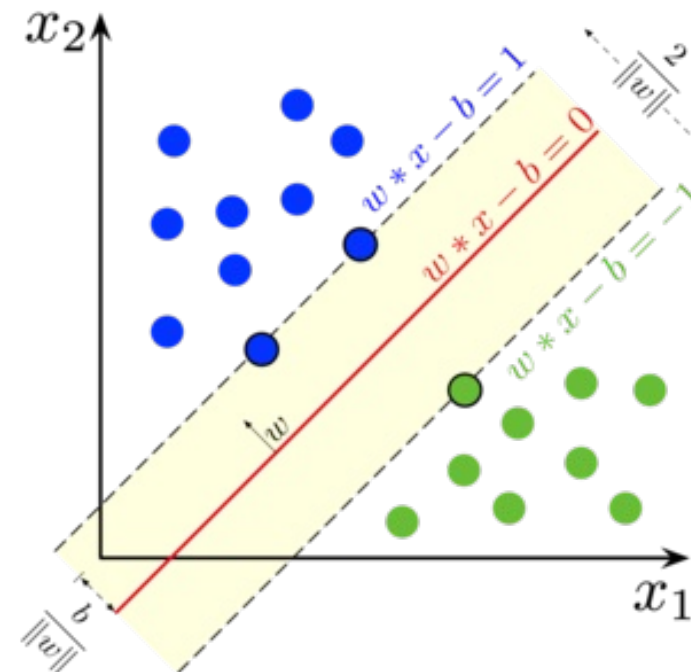
神经网络发展史

➤ 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 (support vector machines, SVM)**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。



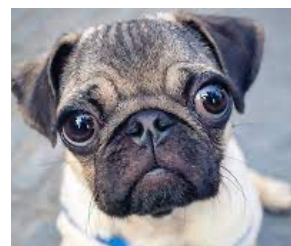
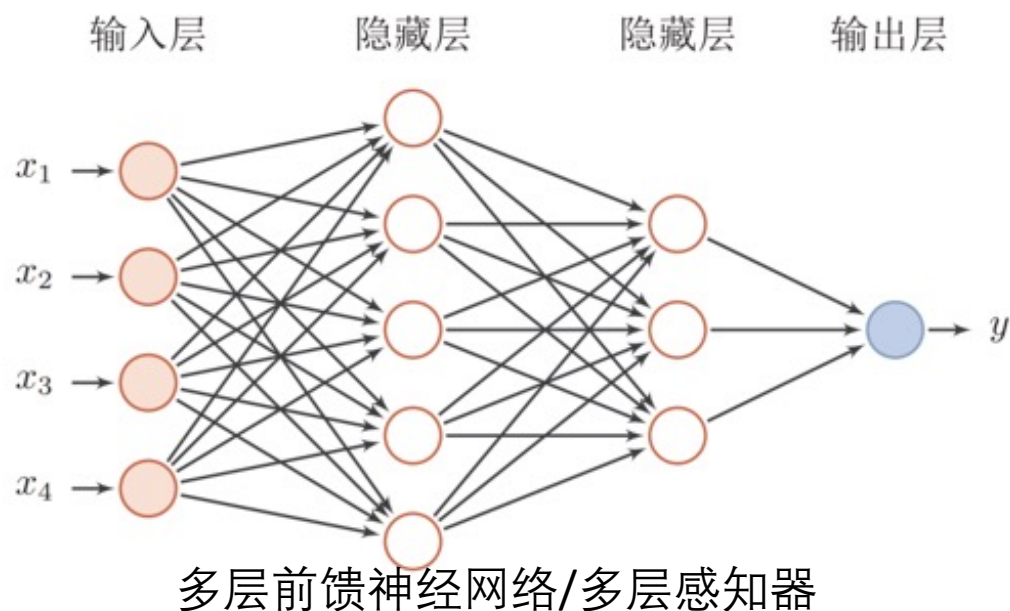
哪一个超平面更好？



神经网络发展史 | 深度学习的崛起

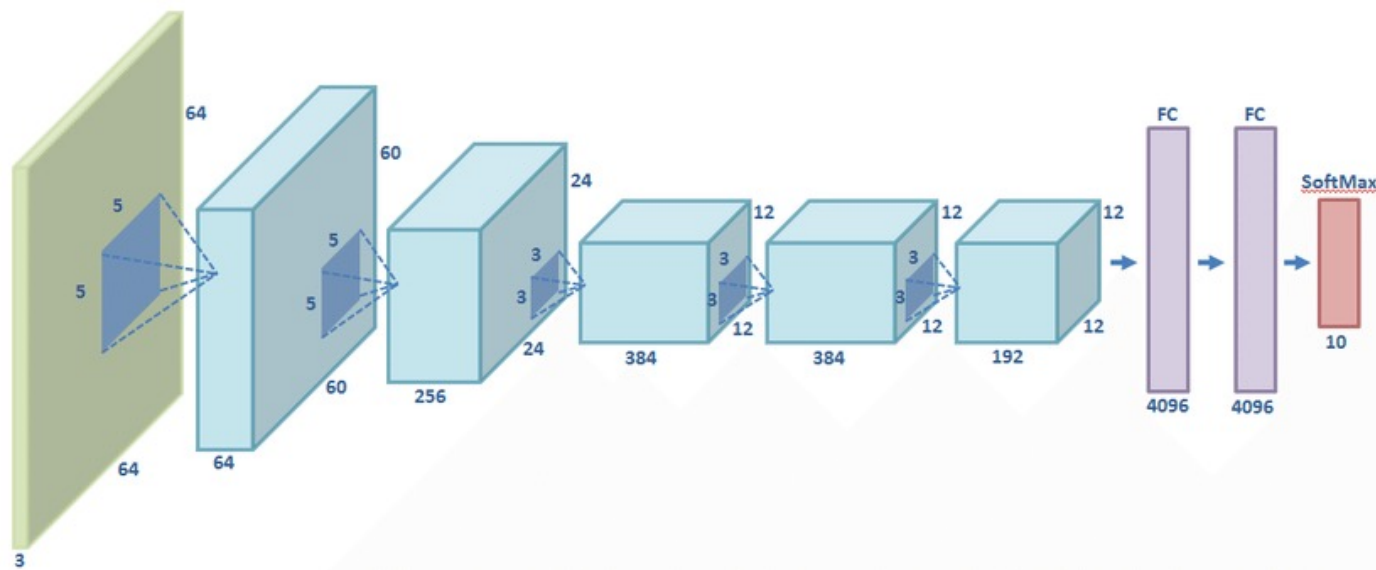
➤ 第五阶段：深度学习的崛起

- 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。



神经网络发展史 | 深度学习的崛起

- 2012年，AlexNet：第一个现代深度卷积神经网络模型，是深度学习技术在图像分类上取得真正突破的开端。
- 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。



AlexNet

ImageNet



李飞飞

李飞飞，女，1976年出生于中国北京，美国国家工程院院士、美国国家医学院院士、美国艺术与科学院院士，美国斯坦福大学首位红杉讲席教授，以人为本人工智能研究院（HAI）院长，AI4ALL联合创始人及主席，Twitter公司董事会独立董事。

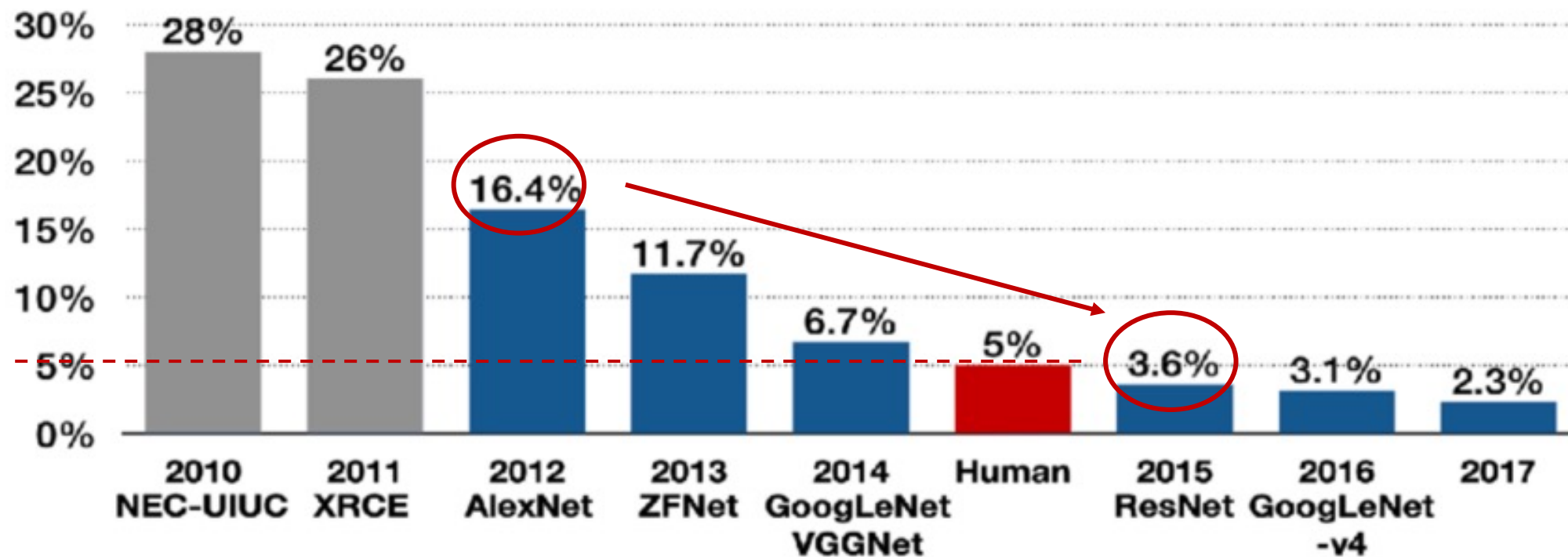
2006年，李飞飞发现整个学术圈和人工智能行业都在苦心研究同一个概念：通过更好的算法来制定决策，但却并不关心数据。她认为：**如果使用的数据无法反映真实世界的状况，即便是最好的算法也无济于事。**

她的解决方案是构建一个更好的数据集。



深度学习模型在图像识别任务上超越人类

Top-5 error



不同模型在ImageNet图像识别挑战赛任务上的错误率



三位图灵奖得主

2019年3月27日，ACM（国际计算机学会）宣布，三位“深度学习之父”杨立昆(Yann LeCun)、杰弗里·辛顿(Geoffrey Hinton)和约书亚·本吉奥(Yoshua Bengio)共同获得了2018年图灵奖。



Yann LeCun



Geoffrey Hinton



Yoshua Bengio

三位图灵奖得主 | 杨立昆

➤ **杨立昆 (Yann LeCun)** , Facebook首席人工智能科学家和纽约大学教授。

➤ **主要贡献**

- **卷积神经网络**: 在20世纪80年代, LeCun开发了卷积神经网络。他是第一个训练卷积神经网络系统处理手写数字图像的人。
- **改进反向传播算法**: LeCun提出了早期版本的反向传播算法, 并基于变分原理对其进行了清晰的推导, 他的工作加速了反向传播算法, 包括两种加速学习时间的简单方法。
- **拓宽神经网络**: LeCun还为神经网络开拓了广阔的空间, 将其作为处理广泛任务的一种计算模型, 并在早期的工作中引入了一些现在广为人知的基本概念。
- 他和Léon Bottou一起提出了一个被应用于每一个现代深度学习软件中的理念, 即学习系统可以被构建为复杂的模块网络, 在这个网络中, 反向传播通过自动分化来执行。他们还提出了能够控制结构化数据的深度学习体系结构。



三位图灵奖得主 | 约书亚·本吉奥

➤ **约书亚·本吉奥(Yoshua Bengio)**，蒙特利尔大学教授，也是魁北克人工智能研究所Mila的科学主任。

➤ 主要贡献

- **序列概率模型**:20世纪90年代，Bengio将神经网络与概率模型相结合。它被用在AT&T/NCR阅读手写支票的系统里，这被认为是20世纪90年代神经网络研究的巅峰。
- **高维词嵌入和注意力模型**:2000年，Bengio发表了具有里程碑意义的论文《神经概率语言模型（A Neural Probabilistic Language Model）》，引入**高维词嵌入**作为词的意义表示，Bengio的研究对自然语言处理任务产生了巨大而持久的影响，其中包括语言翻译、问和视觉回答。他的团队还引入了一种注意力机制，这种**注意力机制**在机器翻译方向取得了突破，成为深度学习中顺序处理的关键组成部分。
- **生成对抗网络**:自2010年以来，Bengio关于生成深度学习的论文，特别是与Ian Goodfellow共同开发的**生成对抗网络**(GANs)，在计算机视觉和计算机图形学领域引发了一场革命。



三位图灵奖得主 | 杰弗里·辛顿

➤ **杰弗里·辛顿(Geoffrey Hinton)**，谷歌副总裁兼工程研究员、Vector研究所首席科学顾问、多伦多大学名誉教授，亦是伦敦大学学院（UCL）盖茨比计算神经科学中心的创立者

➤ 主要贡献

- **反向传播算法**：在1986年的一篇论文“Learning Internal Representations by Error Propagation”中，Hinton与David Rumelhart和Ronald Williams证明了反向传播算法可以帮助神经网络发现数据的内部表示，使得神经网络可以用来解决从前无法解决的问题，反向传播算法是目前大多数神经网络的标准算法。
- **玻尔兹曼机 (Boltzmann Machines)**：1983年，Hinton与泰伦斯·谢诺夫斯基(Terrence Sejnowski)一起发明了玻尔兹曼机，这是第一个能够学习不属于输入或输出的神经元内部表征的神经网络。
- **卷积神经网络的改进**：2012年，Hinton和他的学生Alex Krizhevsky、Ilya Sutskever一起，利用分段线性神经元和dropout正则化改进了卷积神经网络。在著名的ImageNet比赛中，Hinton和他的学生们使物体识别的错误率几乎减半，从而重塑了计算机视觉领域。

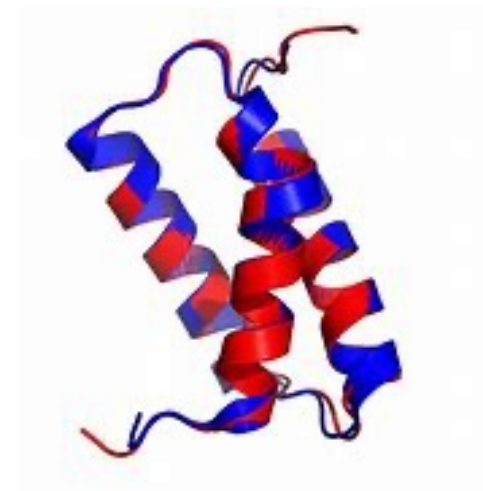
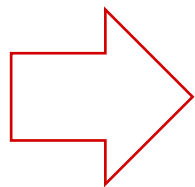




AlphaFold2解决了蛋白质折叠问题

2020 年 11 月底DeepMind 提出的 **AlphaFold** 算法，破解了困扰科学家 50 年之久的蛋白质分子折叠问题。

蛋白质的
氨基酸序列

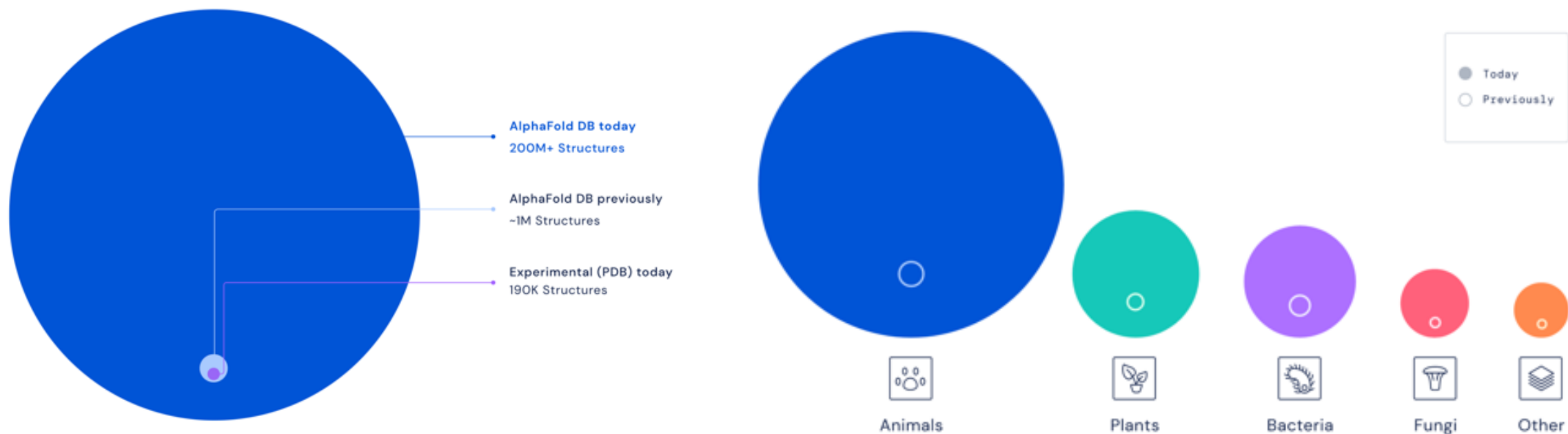


将改变医学，改变研究，改变生物工程，甚至改变一切。

——安德烈·鲁帕斯



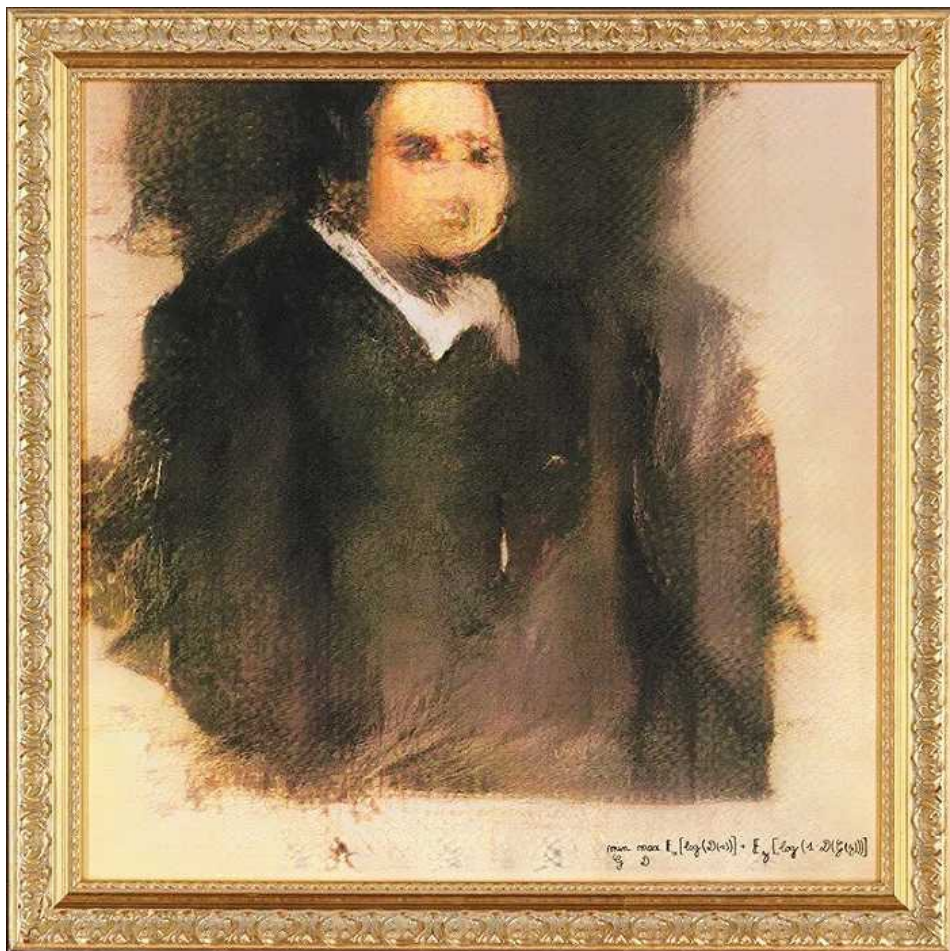
AlphaFold2解决了蛋白质折叠问题



如今的AlphaFold 2:

- ✓ 2.1亿蛋白质结构，其中35%高度准确
- ✓ 100万个物种

生成模型和AI创作 (AIGC)



《埃德蒙德·贝拉米的肖像》

2018年10月25日，由人工智能创作的艺术作品以432000美元（约合300万人民币）的高价成功拍卖，这幅肖像作品是第一个在大型拍卖会上成功交易的人工智能艺术品。

AIGC : AI-generated content (人工智能创作内容)

生成模型和AI创作（AIGC）

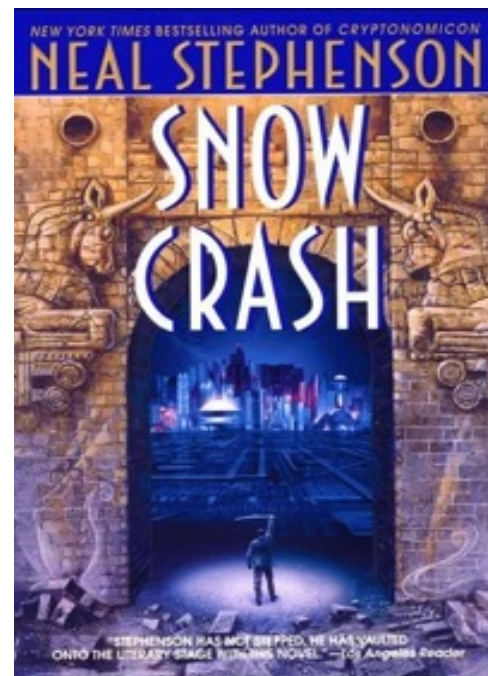


《太空歌剧院》

2022年8月，美国科罗拉多州举办艺术博览会，一幅名为《太空歌剧院》的画作最终获得数字艺术类别冠军。该幅画作是由AI绘图工具Midjourney生成，再经Photoshop润色而来。

元宇宙

- **元宇宙（Metaverse）**，是人类运用数字技术构建的，由现实世界映射或超越现实世界，可与现实世界交互的虚拟世界，具备新型社会体系的数字生活空间。
- “元宇宙”本身并不是新技术，而是集成了一大批现有技术，包括5G、云计算、人工智能、虚拟现实、区块链、数字货币、物联网、人机交互等。



“元宇宙”一词最早出自《雪崩》

区块链

➤ **定义：**区块链（blockchain）是一个共享的、不可篡改的账本，旨在促进业务网络中的交易记录和资产跟踪流程。**资产**可以有形的（如房屋、汽车、现金、土地），也可以是无形的（如知识产权、专利、版权、品牌）。几乎任何有价值的东西都可以在区块链网络上跟踪和交易，从而降低各方面的风险和成本。



区块链



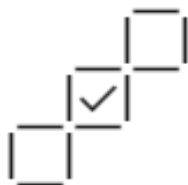
➤ 每个交易发生时，都会被记录为一个数据“块”

这些交易表明资产的流动，资产可以是有形的（如产品），也可以是无形的（如知识产权）。数据块可以记录您的选择信息：何人、何物、何时、何地、数量甚至条件，例如食品运输温度等。



➤ 每个块都连接到位于它前后的块

随着资产位置的改变或所有权的变更，这些数据块形成了数据链。数据块可以确认交易的确切时间和顺序，通过将数据块安全地链接在一起，防止任何数据块被篡改，或防止在两个现有数据块之间插入其他数据块。



➤ 交易被封闭在不可逆的链中：区块链

每添加一个数据块都会加强前一个块的验证，从而增强整个区块链。这使得区块链能够防止篡改，提供不可更改的关键优势。这消除了恶意人员进行篡改的可能性，还建立了您和其他网络成员可以信任的交易账本。

NFT(Non-fungible Token, 非同质通证)

- **NFT**就是一种基于区块链技术的标记技术，可以标记数字资产的所有权，是数字资产的身份证。



#5822
8KΞ (\$23.7M)
Feb 12, 2022



#7804
4.2KΞ (\$7.57M)
Mar 11, 2021



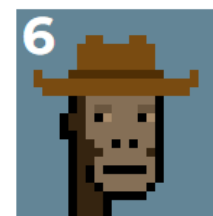
#3100
4.2KΞ (\$7.58M)
Mar 11, 2021



#2924
3.3KΞ (\$4.45M)
Sep 28, 2022



#4156
2.69KΞ (\$3.31M)
Jul 15, 2022



#5577
2.5KΞ (\$7.7M)
Feb 09, 2022



#4464
2.5KΞ (\$2.6M)
Jul 12, 2022



#4156
2.5KΞ (\$10.26M)
Dec 09, 2021



#5217
2.25KΞ (\$5.45M)
Jul 30, 2021



#8857
2KΞ (\$6.63M)
Sep 11, 2021



#2140
1.6KΞ (\$3.76M)
Jul 30, 2021



#7252
1.6KΞ (\$5.33M)
Aug 24, 2021



无聊猿

- **Bored Ape Yacht Club (无聊猿俱乐部 , BAYC)** 是由一万个猿猴NFT组成的数字收藏品合集。包括了帽子 , 眼睛 , 神态 , 服装 , 背景等170个稀有度不同的属性。
- 它们通过编程方式随机组合生成了1万个独一无二的猿猴 , 每个猿猴表情神态穿着各异 , 于 2021年4 月 30 日正式启动。购买一只无聊猿需要0.08ETH , 而当一万枚NFT售罄后 , 人们想再购买它就只能通过Opensea一类平台 , 从最早一批拥有者手中高价收购。



人工智能：顶级国际会议

- **国际表示学习会议**（International Conference on Learning Representations, ICLR）：主要聚焦于深度学习。
- **神经信息处理系统年会**（Annual Conference on Neural Information Processing Systems, NIPS）：交叉学科会议，但偏重于机器学习。主要包括神经信息处理，统计方法，学习理论以及应用等。
- **国际机器学习会议**（International Conference on Machine Learning, ICML）：机器学习顶级会议，深度学习作为近年来的热点，也占据了 ICML 的很大比例。
- **国际人工智能联合会议**（International Joint Conference on Artificial Intelligence, IJCAI）：人工智能领域最顶尖的综合性会议。
- **美国人工智能协会年会**（AAAI Conference on Artificial Intelligence, AAAI）：人工智能领域的顶级会议，每年二月份左右召开,地点一般在北美。



人工智能：顶级国际会议

➤在人工智能的很好应用方向的子领域都有非常好的专业学术会议。

➤计算机视觉领域

- 有计算机视觉与模式识别大会 (IEEE Conference on Computer Vision and Pattern Recognition , CVPR)
- 国际计算机视觉会议 (International Conference on Computer Vision , ICCV)

➤自然语言处理领域

- 有计算语言学年会 (Annual Meeting of the Association for Computational Linguistics , ACL)
- 自然语言处理实证方法大会 (Conference on Empirical Methods in Natural Language Processing , EMNLP)



谢谢！

