# 2024 年诺贝尔物理学奖——人工神经网络实现机器学习



霍普菲尔德像

辛顿像

原文地址：<https://www.nobelprize.org/prizes/physics/2024/popular-information/>，使用 DeepSeek 进行翻译，略加修改。

2024 年诺贝尔物理学奖授予约翰·J·霍普菲尔德（John J. Hopfield）和杰弗里·E·辛顿（Geoffrey E. Hinton），以表彰他们“在实现使用人工神经网络进行机器学习方面的基础性发现和发明。

今年的获奖者利用物理学工具构建了为当今强大机器学习奠定基础的方法。约翰·霍普菲尔德（John Hopfield）创造了一种能存储和重建信息的结构；杰弗里·辛顿（Geoffrey Hinton）发明了一种能自主发现数据特性的方法，这对当前广泛应用的大型人工神经网络至关重要。

## 他们利用物理学在信息中寻找模式

许多人都体验过计算机如何实现语言翻译、图像解读甚至合理对话。但鲜为人知的是，这类技术长期以来对科学研究同样至关重要，包括海量数据的分类与分析。过去 15 至 20 年间，机器学习的发展突飞猛进，其核心是一种称为人工神经网络的结构。如今当我们谈论人工智能时，所指的往往就是这类技术。

尽管计算机无法思考，但如今的机器已能模拟记忆与学习等功能。今年的物理学奖得主为这一突破作出了关键贡献。他们基于物理学的基本概念和方法，开发出利用网络结构处理信息的技术。

机器学习与传统软件不同。传统软件如同菜谱：接收数据后按明确步骤处理并输出结果，就像人们收集食材并按菜谱步骤制作蛋糕。而机器学习则让计算机通过示例学习，使其能解决过于模糊或复杂而无法用逐步指令处理的问题，例如识别图片中的物体。

### 模仿大脑

人工神经网络通过整体网络结构处理信息。其灵感最初源于理解大脑运作方式的渴望。20 世纪 40 年代，研究者开始探索大脑神经元与突触网络的数学基础。另一关键线索来自心理学——神经科学家唐纳德·赫布（Donald Hebb）提出假说：神经元因协同工作而增强连接，这正是学习发生的机制。

随后，研究者尝试通过计算机模拟构建人工神经网络，以复现大脑网络的功能。在这些网络中，神经元被赋予不同值的节点模拟，突触则由节点间可强化或弱化的连接表示。唐纳德·赫布的假说至今仍是人工神经网络通过“训练”过程更新的基本规则之一。

突触

增强

减弱

神经元

大脑的神经网络由活细胞——神经元构成，这些神经元拥有复杂的内部机制。它们能够通过突触彼此传递信号。当我们学习新事物时，某些神经元之间的连接会增强，而其他连接则会减弱。

自然和人工神经元

1

1

0

节点

减弱

加强

人工神经网络由带有编码值的节点构成。这些节点相互连接，当网络被训练时，同时激活的节点之间的连接会增强，否则连接会减弱。

20 世纪 60 年代末，一些令人沮丧的理论结果使许多研究者怀疑神经网络终将无用。然而，20 世纪 80 年代，包括今年获奖者工作在内的多项重要成果重新激发了人们对人工神经网络的兴趣。

### 联想记忆

试想你要回忆一个生僻词，比如描述影院或阶梯教室中倾斜地板的词汇。你在记忆中搜索：类似 *ramp*（坡道）…也许是 *rad…ial*？不，不对。是 *rake*（倾斜度）！

这种通过相似词搜索目标的过程，让人联想到物理学家约翰·霍普菲尔德于 1982 年发现的**联想记忆**。霍普菲尔德网络能存储模式并重建它们：当输入不完整或轻微失真的模式时，该方法能找出最接近的存储模式。

霍普菲尔德曾利用自己的物理学背景探索分子生物学理论问题。在一次神经科学会议上，他接触到关于大脑结构的研究，深受启发并开始思考简单神经网络的动力学特性。当神经元协同作用时，会产生仅观察孤立组件时无法显现的新颖且强大的特性。

1980年，霍普菲尔德离开普林斯顿大学——当时他的研究兴趣已超出物理系同事的领域——横跨北美大陆，接受了加州理工学院（California Institute of Technology）化学与生物学教授职位。在那里，他获得可自由实验的计算机资源，用于发展神经网络思想。

然而，他并未抛弃物理学根基——正是在这一领域，他获得了理解“大量微小组件协同作用如何催生新奇现象”的灵感。他从磁性材料的研究中尤其受益：这类材料的特性源于原子自旋（一种使每个原子成为微型磁体的属性）。相邻原子的自旋会相互影响，从而形成自旋方向一致的磁畴。通过运用描述自旋相互作用下材料演化的物理原理，他成功构建了由节点和连接组成的模型网络。

### 网络将图像存储于“能量景观”

霍普菲尔德构建的网络中，所有节点通过不同强度的连接相互关联。每个节点可存储独立值——在霍普菲尔德的早期工作中，这仅为 0 或 1，类似黑白图像的像素。

霍普菲尔德用类似物理学中自旋系统能量的属性描述网络整体状态：能量通过节点值与连接强度的公式计算。当图像输入节点（赋予0 或 1 值）后，网络通过调整连接强度使存储图像处于低能量状态。当新模式输入网络时，系统按规则逐个检查节点：若改变节点值能否降低能量？例如，若将黑色像素变白可减少能量，则更改颜色。此过程持续直至无法进一步优化，此时网络通常已复现训练时的原始图像。

若仅存储单一模式，这或许并不惊人。你可能会想：为何不直接保存图像本身并与测试图像比对？但霍普菲尔德方法的独特之处在于，它能同时存储多个图像，且网络通常可精准区分它们。

霍普菲尔德将网络搜索存储状态的过程比喻为让球滚过峰谷地貌（摩擦力使其减速）。若球被置于某处，它将滚入最近山谷并停驻。同理，若输入接近存储模式的信号，网络将持续调整直至抵达能量景观的谷底，从而找到记忆中最接近的模式。

霍普菲尔德网络可用于重建含噪声或部分擦除的数据。

约翰·霍普菲尔德的联想记忆以一种类似于塑造地形的方式存储信息。当训练网络时，它会在虚拟能量景观中创建一个对应的山谷。

记忆存储于一片虚拟的能量景观之中

球会自然地滚向最近的谷底。网络也会朝能量较低的方向前进，找到最匹配的存储模式。

当网络接收一个扭曲或不完整的图像时，这个过程可以比作将一颗球滚下山坡。

输入的图案

保存的图案

1

2

能

量

霍普菲尔德及其他研究者持续完善了霍普菲尔德网络的功能细节，例如允许节点存储任意数值（而不仅是0或1）。若将节点视为图像中的像素，它们可呈现不同色彩（而不仅是黑白）。改进后的方法使存储更多图像成为可能，即使图像高度相似也能有效区分。只要信息由众多数据点构成，识别或重建任何内容均成为可能。

### 利用 19 世纪物理学的分类

记住图像是一回事，解读其内容则需更进一步。

即使幼儿也能指着不同动物自信区分狗、猫或松鼠。他们或许偶尔出错，但很快就能几乎全对。儿童无需物种或哺乳动物等概念图表或解释，仅通过观察示例即可建立分类。

当霍普菲尔德发表关于联想记忆的论文时，杰弗里·辛顿正在美国匹兹堡的卡内基梅隆大学工作。他此前曾在英格兰和苏格兰学习实验心理学与人工智能，并思考机器能否以类似人类的方式处理模式——即自主发现分类与解释信息的规则。辛顿与同事特伦斯·谢诺夫斯基（Terrence Sejnowski）以霍普菲尔德网络为基础，结合统计物理学思想构建了新模型。

统计物理学研究由大量相似元素组成的系统（例如气体分子）。虽然追踪每个单独分子极其困难，但可通过其集体行为确定气体宏观性质（如压力或温度）。气体分子能以不同速度分布于容器中，却仍呈现相同的宏观特性。

统计物理学可分析系统组件的共存状态及其发生概率。某些状态的出现概率更高，这取决于可用能量——其关系由 19 世纪物理学家路德维希·玻尔兹曼的方程描述。辛顿的网络应用了该方程，其方法于 1985 年以玻尔兹曼机之名发表。

### 识别同类新示例

玻尔兹曼机通常包含两类节点：接收信息的可见节点与构成隐藏层的节点。隐藏层的节点值和连接强度共同影响网络整体能量。

该机器通过逐节点更新值的规则运行，最终进入一种状态：节点模式可变，但网络整体属性稳定。此时，每个可能模式的出现概率由玻尔兹曼方程根据网络能量计算确定。停机时，机器生成新模式，使其成为早期生成模型的范例。

霍普菲尔德网络

玻尔兹曼机

受限玻尔兹曼机

可见

节点

约翰·霍普菲尔德的联想记忆模型构建了一个所有节点相互连接的网络。信息从所有节点输入，并从所有节点读取输出。

杰弗里·辛顿的玻尔兹曼机通常由两层构成，其中信息通过一层可见节点输入和读取。这些可见节点与隐藏节点相连，隐藏节点影响着整个网络的运行方式。

隐藏

节点

在受限玻尔兹曼机中，同一层内的节点之间没有连接。这些机器通常以链式结构依次连接使用。在训练第一个受限玻尔兹曼机后，其隐藏节点的内容会被用来训练下一个机器，以此类推。

不同类型的网络

玻尔兹曼机的学习不依赖指令，而是通过示例训练。训练时调整网络连接强度，使输入可见节点的示例模式在运行时出现概率最大化。若某模式在训练中多次重复，其概率将进一步提升。训练还会影响机器输出与示例相似新模式的概率。

训练后的玻尔兹曼机能识别未见信息中的熟悉特征。例如，当看到朋友的兄弟姐妹时，你一眼能认出他们的亲缘关系。类似地，若新示例属于训练材料中的类别，机器可准确识别并区分异类数据。

原始玻尔兹曼机效率较低且耗时长。辛顿通过多种改进使其焕发活力：例如删减部分单元间的连接（“稀疏化”），反而提升了运行效率。

20 世纪 90 年代，尽管许多研究者对人工神经网络失去兴趣，辛顿仍坚持深耕该领域。2006 年，他与同事西蒙·奥辛德罗（Simon Osindero）、黄伟（Yee Whye Teh）和鲁斯兰·萨拉赫丁诺夫（Ruslan Salakhutdinov）提出分层玻尔兹曼机预训练法——通过逐层堆叠玻尔兹曼机为网络连接提供更优初始值，从而优化图像特征识别训练。

玻尔兹曼机常作为大型网络的一部分，例如根据用户偏好推荐影视内容。

### 机器学习：当下与未来

约翰·霍普菲尔德与杰弗里·辛顿自 20 世纪 80 年代起的研究，为 2010 年左右开启的机器学习革命奠定了基础。

当前的发展得益于海量训练数据的获取与计算能力的指数级提升。如今的人工神经网络往往规模庞大且层级复杂，被称为深度神经网络，其训练方法则称为深度学习。

回望霍普菲尔德 1982 年关于联想记忆的论文，可清晰感知技术发展的脉络：他当时使用的网络仅含 30 个节点，全连接模式下共有 435 条连接。节点值、连接强度等参数总计不足 500 个。他曾尝试构建 100 节点的网络，但受限于当时的计算机性能而难以实现。与之对比，当今大型语言模型的参数规模已超 1 万亿（1,000,000,000,000）。

如今，研究者正不断拓展机器学习的应用疆域。哪些领域最具前景尚未可知，而围绕技术开发与使用的伦理讨论亦日益广泛。

物理学为机器学习提供了工具，而机器学习亦反哺物理学研究：从筛选希格斯粒子海量数据、降低黑洞碰撞引力波噪声，到搜寻系外行星，机器学习已深度应用于多项诺贝尔物理学奖相关领域。

近年来，该技术更延伸至分子与材料特性计算——例如预测决定蛋白质功能的分子结构，或设计具备高效太阳能转换特性的新型材料。

## 获奖者简介

约翰·J·霍普菲尔德：1933 年出生于美国伊利诺伊州芝加哥市，1958 年获得美国康奈尔大学博士学位，现任美国普林斯顿大学教授。

杰弗里·E·辛顿：1947 年出生于英国伦敦，1978 年获得英国爱丁堡大学博士学位，现任加拿大多伦多大学教授。

[官网链接](https://www.nobelprize.org/prizes/physics/2024/summary/)，[霍普菲尔德讲稿链接](https://www.nobelprize.org/uploads/2025/01/hopfield-lecture.pdf)，[辛顿讲稿链接](https://www.nobelprize.org/uploads/2024/12/hinton-lecture-1.pdf)。