

机器学习神经网络神经科学深度学习 (Deep Learning)

关注者52被浏览22,855

神经网络的历史?

了解一下神经网络的由来，主要分为两个方面：1. Pattern Recognition and Machine Learning的第五章开篇写了 ...显示全部

关注问题

写回答

+ 邀请回答

好问题 2

添加评论

分享

...

你可以邀请下面用户，快速获得回答

搜索你想邀请的人

委托刘看山邀请
开启后为你智能邀请潜在的回答者

开启邀请

镜湖三百里
回答过相似问题

邀请回答

Adam
最近回答过该领域问题

邀请回答

jacob benseal
回答过相似问题

邀请回答

Simon Jia
回答过相似问题

邀请回答

更多推荐结果

3 个回答默认排序

Frankenstein
我心光明，夫复何言。

63 人赞同了该回答

人工神经网络的发展大致经历了三次高潮：20世纪40年代-60年代的控制论、20世纪80年代到90年代中期的联结主义以及2006年以来的深度学习。

1943年，神经科学家和控制论专家Warren McCulloch和逻辑学家Walter Pitts基于数学和阈值逻辑算法创造了一种神经网络计算模型。这一线性模型通过测试 $f(x, w)$ 是正是负来识别两种不同类别的输入。神经网络的研究由此分为了对大脑中生物过程的研究及对把神经网络应用于人工智能的研究（人工神经网络）。

20世纪40年代后期，心理学家Donald Hebb根据神经可塑性的机制提出了一种学习假说：

我们可以假定，反射活动的持续与重复会导致神经元稳定性的持久性提升.....当神经元A的轴突与神经元B很近并参与了对B的重复持续的兴奋时，这两个神经元或其中一个便会发生某些生长过程或代谢变化，致使A作为能使B兴奋的细胞之一，它的效能增强了。

即 $w_{ij} = x_i x_j$, w_{ij} 是神经元 i 与神经元 j 之间的权重， x_i 是神经元 i 的输入。

Hebb的假说后来被称为**赫布型学习**，被认为是一种典型的无监督学习规则。它后来的变种是长期增强作用（由于同步刺激两个神经元而使得两个神经元信号传输持久增强的现象）的早期模型。

1948年，研究人员将这种计算模型的思想应用到了B型图灵机上。

1954年，Farley和物理学家Wesley A. Clark在MIT首次用计算机模拟了一个**赫布网络**

赞同 4

添加评论

分享

收藏

喜欢

...

1969年，Marvin Minsky和Seymour Papert发现了神经网络的两个重大缺陷：其一，基本感知机无法处理异或回路。其二，当时计算机的计算能力不足以用来处理大型神经网络。神经网络的研究就此停滞不前。

1974年，Paul Werbos在博士论文中提出了用**误差反向传导**来训练人工神经网络，有效解决了异或回路问题，使得训练多层神经网络成为可能。

1979年，Kunihiko Fukushima受Hubel & Wiesel的工作启发，提出了**Neocognitron**。在这个工作中他已经有了诸如卷积、池化的想法。

1980年代中期，以**联结主义**的名义，分布式并行处理流行起来。

在20世纪剩下的时间里，支持向量机和其它更简单的算法（如线性分类器）的流行程度逐步超过了神经网络。

1988年，Yann LeCun提出了一种用反向传导进行更新的**卷积神经网络**，称为LeNet。

继续浏览内容

知乎

发现更大的世界

打开

Chrome

继续

2006年，Geoffrey Hinton用**贪婪逐层预训练**（greedy layer-wise pretraining）有效训练了一个深度信念网络。这一技巧随后被研究人员推广到了许多不同的神经网络上，大大提高了模型在测试集上的泛化效果。以Geoffrey Hinton为代表的加拿大高等研究院附属机构的研究人员开始将人工神经网络/联结主义重新包装为了**深度学习**并进行推广。

2009-2012年，瑞士人工智能实验室IDSIA的Jürgen Schmidhuber带领研究小组发展了**递归神经网络**和**深前馈神经网络**。

2012年，Geoffrey Hinton组的研究人员在ImageNet 2012上夺冠，他们图像分类的效果远远超过了第二名，深度学习的热潮由此开始并一直持续到现在。

关于神经网络和生物学、神经科学的关系，我摘几段Ian Goodfellow、Yoshua Bengio和Aaron Courville的Deep Learning一书中的内容：

虽然人工神经网络有时被用来理解大脑机制，但它们总的来说并不是被设计来模拟生物机制。神经科学观点下的深度学习来源于两个想法。一个观点是，大脑用实际例子证明了智能行为是可能的。一个在概念上直截了当的用于创建智能的路径是对大脑背后的计算原理进行逆向工程并且复制其机制。另一个观点是，理解大脑和人类智能背后的原理是非常有趣的，因此和这些基础科学问题相关联的机器学习模型在解决工程问题之外也是有用的。

如今，神经科学被认为是深度学习研究人员的重要灵感来源之一，但它不再是这一领域的决定性指导方针。

神经科学在深度学习领域失去影响力的主要原因是我们没有关于大脑的足够信息。

神经科学给了我们一个理由来希望单一的深度学习算法可以解决许多不同的任务。神经科学家已经发现如果对雪貂的大脑重新连线，把视觉信号发送到他们的听觉处理区域，他们可以用听觉处理区域来学习“看”（Von Melchner et al. 2000）。这暗示很多哺乳动物的大脑可以用单一的算法来解决大脑负责的大部分不同任务。在这一假设之前，机器学习应用方面的研究更加分散，不同的研究群体分别进行自然语言处理、计算机视觉、运动规划和语音识别领域的研究。而今，这些研究群体虽然仍然分散，但是深度学习研究小组已经习惯于同时研究很多甚至全部的应用。

现代深度学习从很多领域汲取灵感，特别是应用数学的基础如线性代数、概率、信息论和数值优化。当一些深度学习研究人员把神经科学作为一个重要灵感来源时，他们通常是在寻找一些可以借鉴的生物学原理。

算神经科学两个领域是非常普遍的。

关于第二个问题，我对统计学一窍不通，不知道有没有高人指点一下。我个人感觉神经网络发展之初和统计学的关联比较有限。各机器学习算法的兴衰似乎主要还是在算法实施的难易程度、模型拟合训练集的能力以及模型的泛化能力。

关于过拟合的问题，一般认为非参数和非线性的机器学习模型更容易过拟合。而神经网络恰恰是非线性的。

关于神经网络在统计学术性质谷歌上应该能找到一些相关的研究，比如这篇[STATISTICAL PROPERTIES OF ARTIFICIAL NEURAL NETWORKS](#)。

参考资料：

- 1. [Artificial neural network](#)
- 2. [Hebbian theory - Wikipedia](#)

继续浏览内容

知乎

发现更大的世界

打开

Chrome

继续

感谢@刘间或指出，维数灾祸（the curse of dimensionality）在机器学习里是指随着样本空间和参数空间维数的增加，潜在参数组合的数量指数型增长，相同精确度的预测会需要多得多的训练样本。当训练样本不变时，预测的精度与空间维数的增加成反比。

Yoshua Bengio在What is the curse of dimensionality? 指出由此带来的真正困难是要学习的目标方程的复杂度以及其复杂度是否合理，能否通过学习少量的样本就拥有足够出色的泛化能力。一个由此而来的研究问题是在何种程度上算法有潜力以这样非本地的方式泛化（在与训练集很不同的样本上泛化）。他认为深度学习有这样的潜力。参考他和Yann Lecun写的[Scaling Learning Algorithms towards AI](#)。

编辑于 2017-12-26

赞同 63 4 条评论 分享 收藏 喜欢 收起

史博

机器学习话题下的优秀答主

4 人赞同了该回答

- 1. 早期对神经网络贡献最大的学科可能未必是生物科学、甚至神经科学，反而是逻辑控制，电子电路，认知科学等领域。可能模仿对象是神经元，但是并非自然而然的衍生，而是在另外一个领域的再造。至于命名，有点类似宇宙飞船命名成X船，甚至嫦娥一样。更多细节可以参考：[深度神经的诞生](#)
- 2. 过拟合、维灾受到重视更多源于早期统计机器学习领域的工作，但是致力于过拟合、维灾的机器学习中核心的计算学习理论要比神经网络理论诞生晚点，但是理论上却蛮美的。就初学的话，个人觉得先不要和神经网络放在一起讨论，因为神经网络、尤其深度神经网络相关的理论解释，还缺乏很完美的公认的结论，或许这样搞深度学习又被称为炼丹。

发布于 2018-12-11

赞同 4 添加评论 分享 收藏 喜欢

微尘-黄含驰

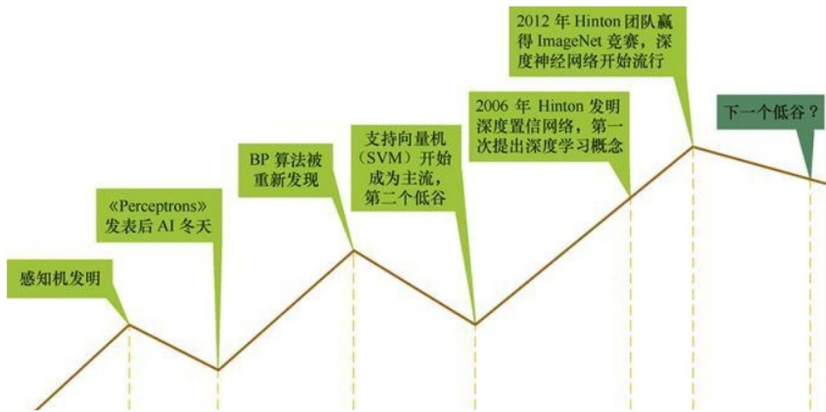
强化学习+组合优化+MAB瞎盐ing;微信jinuxjp5x

4 人赞同了该回答

我觉得这本书的介绍很简洁形象，搬一下砖~（[百面机器学习：算法工程师带你去面试 - 知乎书店](#)）

赞同 4 添加评论 分享 收藏 喜欢

中，感知机也被认为是单层的人工神经网络，尽管结构简单，却能够学习并解决相当复杂的问题。下图是神经网络的发展历史。



继续浏览内容

知乎

发现更大的世界

打开

Chrome

继续

神经网络这个可导函数的问题，就必须发展多层感知机，即中间至少包含一个隐藏的多层神经网络，但是当时根本找不到运用在多层神经网络上的有效算法。至此，学术权威开始质疑神经网络，悲观主义开始蔓延。

从现在看，突破性的误差反向传播算法，即著名的反向传播算法，开启训练多层神经网络的“钥匙”，其实那个时候已经存在了。冰冻10年中，尽管 Paul Werbos在1974年的博士毕业论文中深刻分析了将反向传播算法运用于神经网络方面的可能性，成为美国第一位提出可以将神经网络的研究人员，但是他没有发表将反向传播算法用于神经网络这方面的研究，因为圈子大体已经失去解决这些问题的信念，这时候，LeCun Yann(他给自己取了个中文名杨立昆)大侠上场了，20世纪80年代他在Hinton实验室做博士后期间，提出了神经网络的反向传播算法原型，186年，Rumelhart、Hinton和Williams合著Learning representations by backpropagating errors，反向传播算法开始流行开来。

Lecun Yann和其他人发展的神经网络正开始被热捧的时候，他一生较劲的对象Vapnik(尔实验室的同事)出现了。20世纪90年代中期，由Vapnik等人发明的支持向量机诞生，它同样解决了线性不可分问题，但是对比神经网络有全方位优势，比如，高效，可以快速训练，无须调参，没有梯度消失问题，泛化性能好，过拟合风险小，支持向量机迅速打败多层神经网络成为主流。后来一度发展到，只要你的论文中包含神经网络相关的字眼，就很容易被拒稿，学术界那时对神经网络的态度可想而知，神经网络再次堕入黑暗。10年沉寂中，只有几个学者仍然在坚持研究，比如一再提及的Hinton教授。

2006年，Hinton在《Science》和相关期刊上发表了论文，首次提出了“深度置信网络”的概念。与传统的训练方式不同，深度信念网络有一个“预训练”的过程，它的作用是让神经网络权值找到一个接近最优解的值，之后再使用“微调”技术，使用反向传播算法或者其他算法作为调优的手段，来对整个网络进行优化训练，这两个技术的运用大幅度提升了模型的性能，且少了训练多层神经网络的时间。他给多屋神经网络相关的学习方法赋予了一个新名词——深度学习。

后面的故事我们都知道了，2012年Hinton的团队用LeCun赖以成名的卷积神经网络，和自己在深度置信网络的调优技术，碾压了其他机器学习办法。至此，深度学习开始垄断人工智能的新闻报道，Hinton、LeCun和他们的学生像摇滚明星一般受到追捧，学者们的态度也来了个180度大转变，现在没有和深度学习沾上边的文章很难发表了。除了名，还有利，Google等大公司不但把学术界人物挖了个通，更是重金收购深度学习大佬们所创建的公司。坐了几十年冷板凳的人忽然一夜之间身价暴涨、财务自由。不过，现在主导Facebook AI实验室的LeCun Yann则不断呼吁学术界对深度学习保持冷静。

编辑于 2019-11-16

继续浏览内容

 知乎

发现更大的世界

打开

 Chrome

继续