

深度学习是机器学习研究中的一个新的领域，其动机在于建立、模拟人脑进行分析学习的神经网络，它模仿人脑的机制来解释数据，例如图像，声音和文本。

同机器学习方法一样，深度机器学习方法也有监督学习与无监督学习之分。不同的学习框架下建立的学习

模型很是不同。例如，卷积神经网络（Convolutional neural networks，简称 CNNs）就是一种深度的监督学习下的机器学习模型，而深度置信网（Deep Belief Nets，简称 DBNs）就是一种无监督学习下的机器学习模型。

目录

- 1 简介
- 2 基础概念
- ? 深度
- ? 解决问题
- 3 核心思想
- 4 例题
- 5 转折点
- 6 成功应用

1 简介

深度学习的概念源于人工神经网络的研究。含多隐层的多层感知器就是一种深度学习结构。深度学习通过组合低层特征形成更加抽象的高层表示属性类别或特征，以发现数据的分布式特征表示。^[2]

深度学习的概念由 Hinton 等人于 2006 年提出。基于深信度网 (DBN) 提出非监督贪心逐层训练算法，为解决深层结构相关的优化难题带来希望，随后提出

多层自动编码器深层结构。此外 Lecun 等人提出的卷积神经网络是第一个真正多层结构学习算法，它利用空间相对关系减少参数数目以提高训练性能。 [2]

2 基础概念

深度：从一个输入中产生一个输出所涉及的计算可以通过一个流向图 (flow graph) 来表示：流向图是一种能够表示计算的图，在这种图中每一个节点表示一个基本的计算并且一个计算

深度学习的值 (计算的结果被应用到这个节点的孩子节点的值)。考虑这样一个计算集合，它可以被允许在每一个节点和可能的图结构中，并定义了一个函数族。输入节点没有孩子，输出节点没有父亲。

这种流向图的一个特别属性是深度 (depth)：从一个输入到一个输出的最长路径的长度。

传统的前馈神经网络能够被看做拥有等于层数的深度 (比如对于输出层为隐层数加 1)。SVMs 有深度 2(一个对应于核输出或者特征空间，另一个对应于所产生输出的线性混合)。 [3]

解决问题：

需要使用深度学习解决的问题有以下特征：

深度不足会出现问题。

人脑具有一个深度结构。

认知过程逐层进行，逐步抽象。

深度不足会出现问题

在许多情形中深度 2 就足够表示任何一个带有给定目标精度的函数。 但是其代价是：图中所需要的节点数（比如计算和参数数量）可能变的非常大。理论结果证实那些事实上所需要的节点数随着输入的大小指数增长的函数族是存在的。

我们可以将深度架构看做一种因子分解。 大部分随机选择的函数不能被有效地表示，无论是用深的或者浅的架构。 但是许多能够有效地被深度架构表示的却不能被用浅的架构高效表示。 一个紧的和深度的表示的存在意味着在潜在的可被表示的函数中存在某种结构。 如果不存在任何结构，那将不可能很好地泛化。

大脑有一个深度架构

例如，视觉皮质得到了很好的研究， 并显示出一系列的区域， 在每一个这种区域中包含一个输入的代表和从一个到另一个的信号流（这里忽略了在一些层次并行路径上的关联，因此更复杂）。 这个特征层次的每一层表示在一个不同的抽象层上的输入，并在层次的更上层有着更多的抽象特征， 他们根据低层特征定义。

需要注意的是大脑中的表示是在中间紧密分布并且纯局部：他们是稀疏的：1%的神经元是同时活动的。 给定大量的神经元，仍然有一个非常高效地（指数级高效）表示。

认知过程逐层进行，逐步抽象

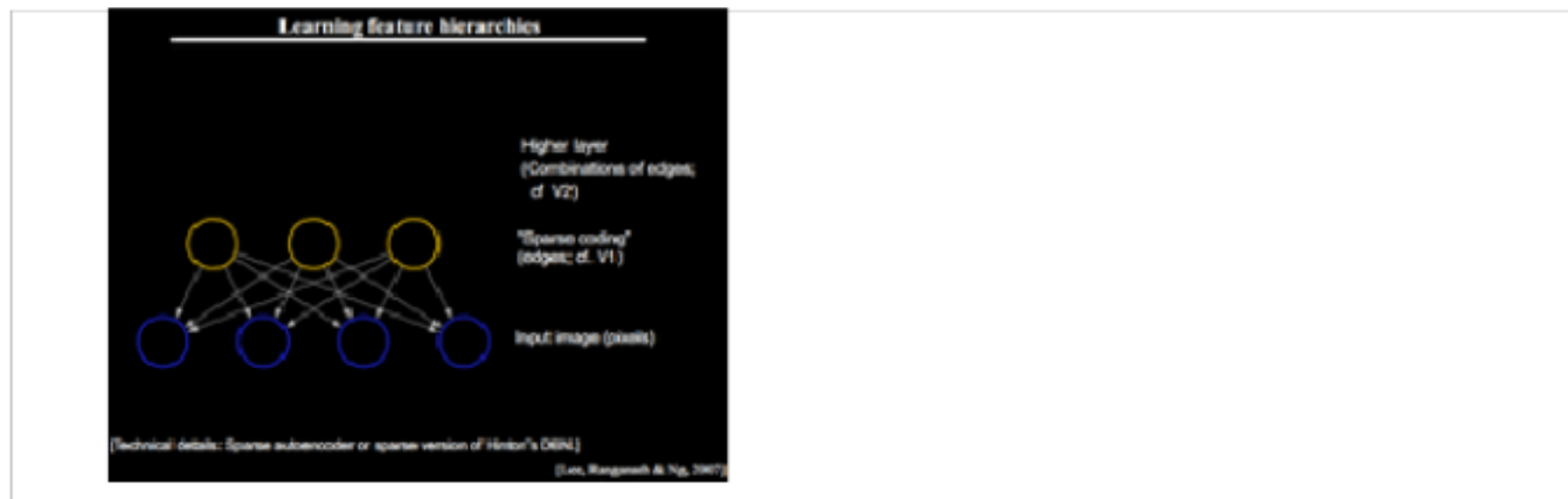
人类层次化地组织思想和概念；

人类首先学习简单的概念，然后用他们去表示更抽象的；

工程师将任务分解成多个抽象层次去处理；

学习/发现这些概念（知识工程由于没有反省而失败？）是很美好的。对语言可表达的概念的反省也建议我们一个稀疏的表示：仅所有可能单词 / 概念中的一个小的部分是可被应用到一个特别的输入（一个视觉场景）。

3 核心思想



深度学习的核心思想

把学习结构看作一个网络，则深度学习的核心思路如下：

无监督学习用于每一层网络的 pre-train ；

每次用无监督学习只训练一层，将其训练结果作为其高一层的输入；

用自顶而下的监督算法去调整所有层

4 例题

a). AutoEncoder

最简单的一种方法是利用人工神经网络的特点，人工神经网络（ ANN ）本身就是具有层次结构的系统，如果给定一个神经网络，我们假设其输出与输入是相同的，然后训练调整其参数，得到每一层中的权重，自然地，我们就得到了输入 I 的几种不同表示（每一层代表一种表示），这些表示就是特征，在研究中可以发现，如果在原有的特征中加入这些自动学习得到的特征可以大大提高精确度，甚至在分类问题中比目前最好的分类算法效果还要好！这种方法称为 AutoEncoder 。当然，我们还可以继续加上一些约束条件得到新的 Deep Learning 方法，如如果在 AutoEncoder 的基础上加上 L1 的 Regularity 限制(L1

主要是约束每一层中的节点中大部分都要为 0,只有少数不为 0,这就是 Sparse 名字的来源),我们就可以得到 Sparse AutoEncoder 方法。

b). Sparse Coding

如果我们把输出必须和输入相等的限制放松,同时利用线性代数中基的概念,即 $O = w_1*B_1 + W_2*B_2+\dots+ W_n*B_n$, B_i 是基, W_i 是系数,我们可以得到这样一个优化问题:

$$\text{Min } |I - O|$$

通过求解这个最优化式子,我们可以求得系数 W_i 和基 B_i ,这些系数和基础就是输入的另外一种近似表达,因此,可以用它们来特征表达输入 I ,这个过程也是自动学习得到的。如果我们在上述式子上加上 L1 的 Regularity 限制,得到:

$$\text{Min } |I - O| + u*(|W_1| + |W_2| + \dots + |W_n|)$$

这种方法被称为 Sparse Coding。

c) Restrict Boltzmann Machine (RBM)

假设有一个二部图,每一层的节点之间没有链接,一层是可视层,即输入数据层 (v),一层是隐藏层 (h),如果假设所有的节点都是二值变量节点(只能取 0 或者 1 值),同时假设全概率分布 $p(v, h)$ 满足 Boltzmann 分布,我们称这个模型是 Restrict Boltzmann Machine (RBM)。下面我们来看看为什么它是 Deep Learning 方法。首先,这个模型因为是二部图,所以在已知 v 的情况下,所有的隐藏节点之间是条件独立的,即 $p(h|v) = p(h_1|v)\dots p(h_n|v)$ 。同理,在已知隐藏层 h 的情况下,所有的可视节点都是条件独立的,同时又由于所有的 v 和 h 满足 Boltzmann 分布,因此,当输入 v 的时候,通过 $p(h|v)$ 可以得到隐藏层 h ,而得到隐藏层 h 之后,通过 $p(v|h)$ 又能得到可视层,通过调整参数,

我们就是要使得从隐藏层得到的可视层 v_1 与原来的可视层 v 如果一样，那么得到的隐藏层就是可视层另外一种表达，因此隐藏层可以作为可视层输入数据的特征，所以它就是一种 Deep Learning 方法。

如果，我们把隐藏层的层数增加，我们可以得到 Deep Boltzmann Machine (DBM)；如果我们在靠近可视层的部分使用贝叶斯信念网络（即有向图模型，当然这里依然限制层中节点之间没有链接），而在最远离可视层的部分使用 Restrict Boltzmann Machine，我们可以得到 Deep Belief Net（DBN）。

当然，还有其它的一些 Deep Learning 方法。总之，Deep Learning 能够自动地学习出数据的另外一种表示方法，这种表示可以作为特征加入原有问题的特征集合中，从而提高学习方法的效果，是业界的 research 热点。

5 转折点

2006 年前，尝试训练深度架构都失败了：训练一个深度有监督前馈神经网络趋向于产生坏的结果（同时在训练和测试误差中），然后将其变浅为 1(1 或者 2 个隐层)。

2006 年的 3 篇论文改变了这种状况，由 Hinton 的革命性的在深度信念网 (Deep Belief Networks, DBNs) 上的工作所引领：

Hinton, G. E., Osindero, S. and Teh, Y., A fast learning algorithm for deep belief nets. *Neural Computation* 18:1527-1554, 2006

Yoshua Bengio, Pascal Lamblin, Dan Popovici and Hugo Larochelle, Greedy LayerWise Training of Deep Networks, in J. Platt et al. (Eds), *Advances in Neural Information Processing Systems* 19 (NIPS 2006), pp. 153-160, MIT Press, 2007

Marc ' Aurelio Ranzato, Christopher Poultney, Sumit Chopra and Yann LeCun Efficient Learning of Sparse Representations with an Energy-Based Model, in J. Platt et al. (Eds), Advances in Neural Information Processing Systems (NIPS 2006), MIT Press, 2007

在这三篇论文中以下主要原理被发现：

表示的无监督学习被用于 (预)训练每一层；

在一个时间里的一个层次的无监督训练，接着之前训练的层次。在每一层学习到的表示作为下一层的输入；

用有监督训练来调整所有层 (加上一个或者更多的用于产生预测的附加层)；

DBNs 在每一层中利用用于表示的无监督学习 RBMs。Bengio et al paper 探讨和对比了 RBMs 和 auto-encoders(通过一个表示的瓶颈内在层预测输入的神经网络)。Ranzato et al paper 在一个 convolutional 架构的上下文中使用稀疏 auto-encoders(类似于稀疏编码)。Auto-encoders 和 convolutional 架构将在以后的课程中讲解。

从 2006 年以来，大量的关于深度学习的论文被发表。

6 成功应用

1、计算机视觉

ImageNet Classification with Deep Convolutional Neural Networks , Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton, NIPS 2012.

Learning Hierarchical Features for Scene Labeling , Clement Farabet, Camille Couprie, Laurent Najman and Yann LeCun, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013.

Learning Convolutional Feature Hierarchies for Visual Recognition

Koray Kavukcuoglu, Pierre Sermanet, Y-Lan Boureau, Karol Gregor, Michaël Mathieu and Yann LeCun, Advances in Neural Information Processing Systems (NIPS 2010), 23, 2010.

2、语音识别

微软研究人员通过与 Microsoft Research 合作，首先将 RBM 和 DBN 引入到语音识别声学模型训练中，并且在大规模语音识别系统中获得巨大成功，使得语音识别的错误率相对减低 30%。但是，DNN 还没有有效的并行快速算法，很多研究机构都是在利用大规模数据语料通过 GPU 平台提高 DNN 声学模型的训练效率。

在国际上，IBM、google 等公司都快速进行了 DNN 语音识别的研究，并且速度飞快。

国内方面，阿里巴巴，科大讯飞、百度、中科院自动化所等公司或研究单位，也在进行深度学习在语音识别上的研究。

3、自然语言处理等其他领域

很多机构在开展研究，但深度学习在自然语言处理方面还没有产生系统性的突破。

THANKS !!!

致力为企业和个人提供合同协议， 策划案计划书， 学习课件等等

打造全网一站式需求

欢迎您的下载， 资料仅供参考