

第一章 引言

1.0.1 参考书目及课程资源

参考书目如下：

- 《动手学深度学习》
- 《深度学习》(花书)

视频资源：

- 《动手学深度学习》
- 深度学习——李宏毅

代码资源：

- 《动手学深度学习》

1.0.2 深度学习与机器学习的区别与联系

机器学习是什么？ 并不是所有的解决方案都可以通过直接编写程序来解决，比如，编写一个识别猫的图片的程序。机器学习尝试利用**数据**，并结合**模型假设**，再基于一定的**优化方案**获得相对最优的模型。这一过程中，通常需要借用人类的认知知识，比如，标注好的猫的图片，因此，“学习”的过程就有了，所以称为机器学习 (machine learning)。机器学习是现阶段实现人工智能的有效方式。

从程序设计的角度而言，这一区别于直接编写程序，而是通过数据学习的方式来构建解决方案，也因此被称为“**用数据编程**”。

从数学的角度而言，机器学习与统计学有着密切联系，即，基于数据的分析技术，比如，最小二乘线性回归。

深度学习中的模型假设。 通常人们所说的深度学习是基于深度神经网络模型的学习框架。这种形式灵活多变的模型构建方法，使得模型假设也变得特别一般化 (general)，随着网络模型宽度与深度的增加，模型的表达能力 (capacity) 可以很大很大。从网络模型这样的模型假设来看，机器学习的一些经典模型都可以看做是浅层神经网络，比如，线性回归，逻辑斯蒂回归，支持向量机。。。

表示学习 (representation learning)。 深度学习成功引起人们的重视，是源自于2012年计算机视觉领域中的 ImageNet 竞赛，AlexNet 模型取得了远超传统机器学习模型的性能优势。视觉领域中，特征表示学习起着至关重要的作用，回顾深度模型之前的视觉领域，大量的科研关注点是设计更加有效的特征抽取方法，经典的工作包括适用于行人识别的 HOG 特征，适用于纹理细节描述的 LBP、SIFT、Gabor 等特征，以及为解决多尺度目标识别的特征金字塔构建方案。

深度网络模型中从原始图像像素输入层层到分类决策层的层层堆叠过程中，可以看成是在逐级构建不同抽象语义级别的特征表示，比如，底层在抽取图像中的边、角信息，中层在抽取部件、边框等信息，高层在抽取对象级上的语义信息。在终端任务驱动下，自动学习出的特征表示如下图 1.1 所示。正是因为这种分层逐级抽象的特征表示方式，也被称为“分布式表示学习(distributed representation learning)”。

1.0.3 深度学习能够兴起的时代背景

优化技术。 从浅层神经网络到深度神经网络的想法并不是多么石破天惊，早有人尝试过，但是难以学习到有效的模型，主要原因在于基于梯度下降算法的优化过程中存在严重的梯度消失/梯度爆炸问题，以及权重矩阵退化问题，即，多数权重趋于 0。聚焦于权重初始化，数据预处理及归一化，以及 dropout 以及动量梯度优化技术，训练深层神经网络成为可能。

大数据。 从 ORL 人脸数据集，到 CIFAR，到 ImageNet，大规模精细标注的数据为模型及算法的研究提供了学习基础。

高算力。 NVIDIA 推出了用于 GPU 并行计算的 CUDA 工具包为深度神经网络模型的并行训练提供了便利。

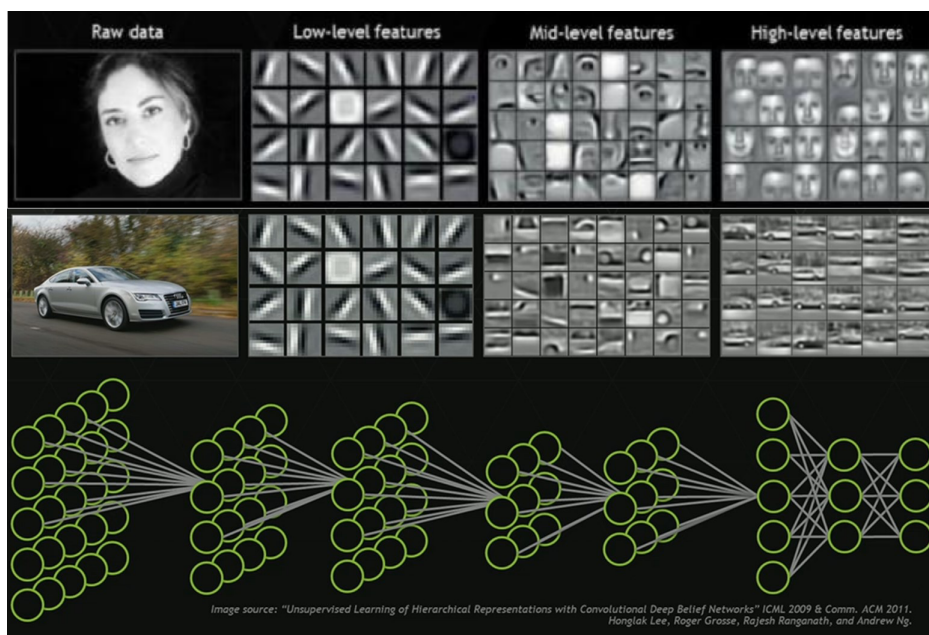


图 1.1: 表示学习学到了什么?

