

这里我们用的 PPT 是原作者的 Lecture Sildes (<http://www.deeplearningbook.org/>)。

Introduction

Lecture slides for Chapter 1 of *Deep Learning*

www.deeplearningbook.org

Ian Goodfellow

2016-09-26

第一章的 PPT 作者是 Ian Goodfellow。

1997 年, IBM 的“深蓝”在国际象棋项目中超越了人类。这似乎看起来人工智能就要崛起了。

Ironically, abstract and formal tasks that are among the most difficult mental undertakings for a human being are among the easiest for a computer.

有些嘲讽的是, 那些对于人类很难解决的规则化和抽象的任务对于计算机来说恰恰是比较容易的。

作者这句话写的确实有意思, 我们的确知道最近几年才刚刚“勉强”使计算机在识别图像和语音上能稍微和人类的平均水平匹敌。我们人类从生下来就在不断地获取信息, 通过信息来不

断地学习，以此来更加充分的认识与了解这个大千世界。对于计算机来说，第一个难点就来了，我们如何将信息进行一些简单的处理来让计算机更好地理解他们？

Representations Matter

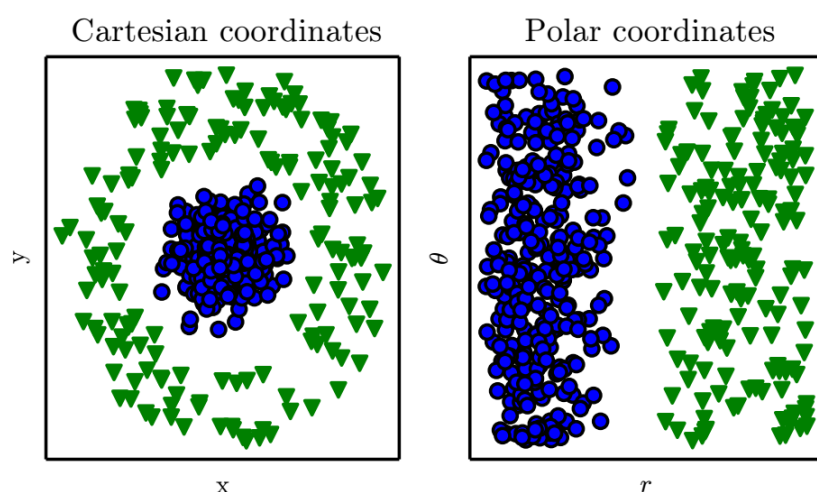


Figure 1.1

(Goodfellow 2016)

上图（图 1.1）就是一个很好的例子，蓝色和绿色分别是两个类别的数据，他们分布在笛卡尔坐标系中（左），我们要做的就是来画一条直线，把蓝色和绿色的点分开。显然这是不可能的。我们可以观察出来貌似这些数据的分布好像是个圆，如果我们以圆心为坐标原点，建立一个极坐标系，问题会不会简单些？很明显是这样的（PPT 里面都画出来了我在这么问有点马后炮的感觉）。我们会发现当问题被转化到极坐标系下，我们就可以直接用一条直线将蓝色和绿色数据分离。这个问题瞬间变得很容易。

这时就有人会说了，那我们在给计算机输入数据之前，都这么弄一下不就好了么，简单的很（hin）啊。

However, for many tasks, it is difficult to know what features should be extracted.

遗憾的是，我们生活中的各种数据，并没有这么理想，即使是我们人类也很难去定义这些特征。

比如，我们需要在一张照片中检测车辆。我们知道汽车有轮子，所以我们可能会把轮子作为汽车检测的一个特征。不幸的是，在像素的级别上描述轮子是非常困难的。轮子的几何形状非常简单，但是轮子上的阴影、金属部分的折射、保险杠的影响或前景物体的遮挡等让描述它变得非常复杂。

对于这种问题，传统方法并没有什么好的解决办法，于是深度学习（或许这里我们叫他神经网络更为合适）的优势就显现出来了。

Deep learning solves this central problem in representation learning by introducing representations that are expressed in terms of other, simpler representations. Deep learning allows the computer to build complex concepts out of simpler concepts.

“深度学习”通过表征的层级组合解决了表征学习中的核心问题。深度学习允许计算机通过学习一系列简单的概念来组成复杂的概念。

Depth: Repeated Composition

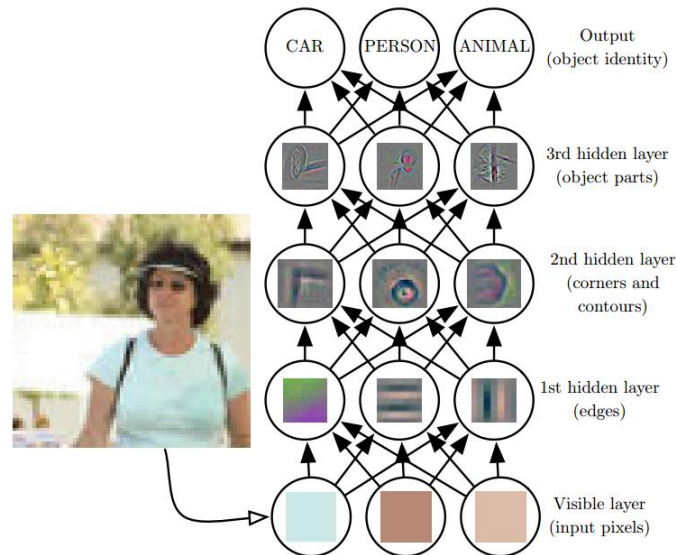


Figure 1.2

(Goodfellow 2016)

上图 (图 1.2) 是深度学习的一个示意图。给定一个输入的图片, 我们想知道他里面的内容是什么 (小朋友们, 这张照片里装了什么吖?)。我们确实一眼就能看出这里面是个人 (这个问题我们假设输出只能是车, 人, 动物, 也就是说这是个选择题)。但是如果 we 强求计算机“一眼”看出来是什么的话, 计算机只能告诉你, 什么鬼??? 这里深度学习就出马了 (图中用得是卷积神经网络), 通过逐层卷积, 计算机已经能够大概的提取一些图片的特征了。将我们人类看到的“视觉层” (也就是原图的 RGB 通道) 作为输入, 第一个隐层可以十分容易的通过相邻像素之间的明暗对比抽取出边缘的信息, 第二层则可以通过轮廓信息学习出角点和轮廓的信息, 第三层则可以通过角点和轮廓的信息寻找到指定物体的一些组成部件 (头和身子?) 信息。最终, 通过对物体的组成部件描述, 我们可以识别出图像中的物体。这里计算机就可以回答我们, 它看到的是一个人 (报告老师我看到了一个大妈)。

Computational Graphs

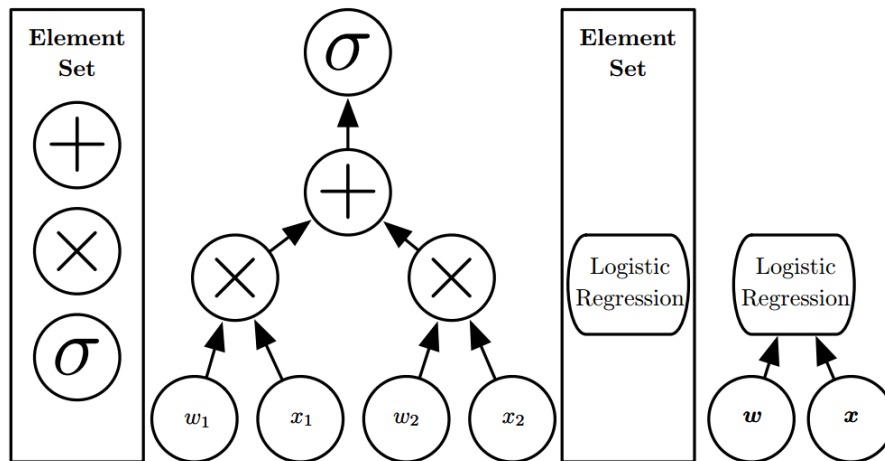


Figure 1.3

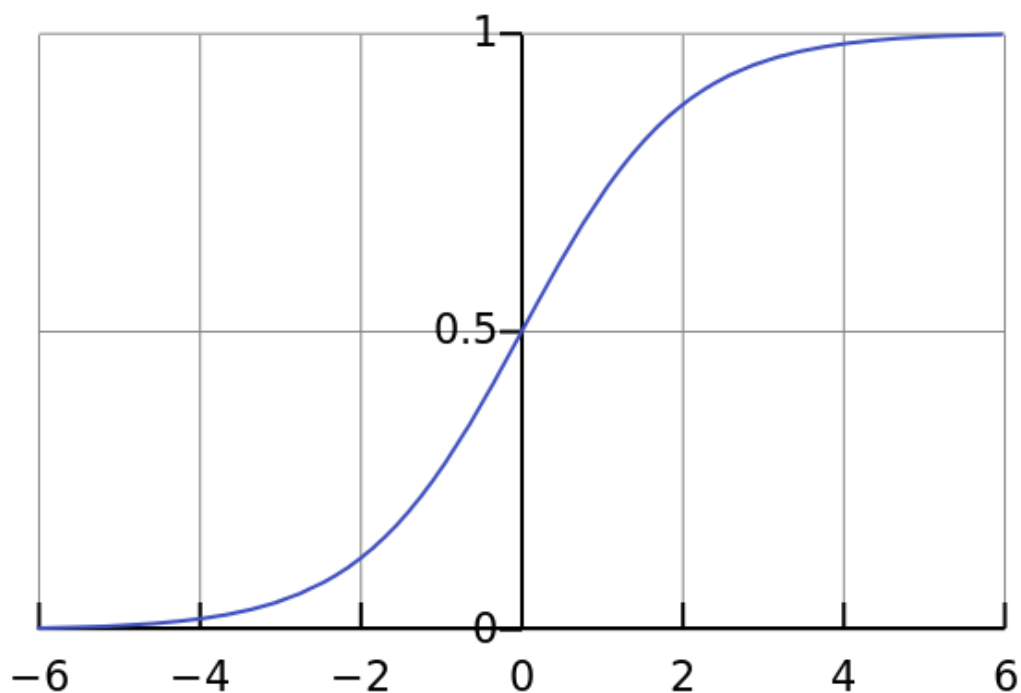
(Goodfellow 2016)

上图 (图 1.3) 给出了一个深度学习的基本组成单位 (神经网络的神经元节点) 的结构 , 这

里就不多废话了直接上公式 , 假定节点输入为 , 输出为 :

$$y = \sigma(\omega^T \mathbf{x}) = \frac{1}{1 + e^{-(\omega_1 x_1 + \omega_2 x_2)}}$$

这里的 $\sigma(\cdot)$ 被称作激活函数 (Active Function) , 这里我们选用 Sigmoid 函数 :



我们可以看到，深度学习的基本组成单元（神经元节点）只有加法，乘法和激活函数。由许多个神经元节点相互连接形成深度神经网络，也就是本书中的深度学习模型。

Machine Learning and AI

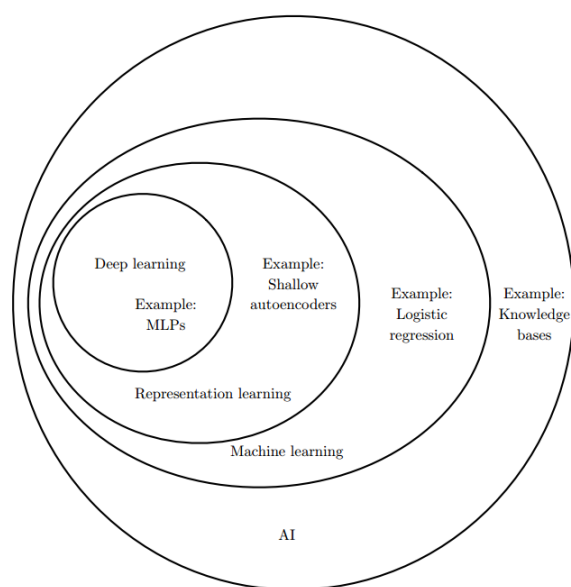


Figure 1.4

(Goodfellow 2016)

从深度学习发展至今，有许多人把深度学习叫做人工智能（Artificial Intelligence）。那 Deep

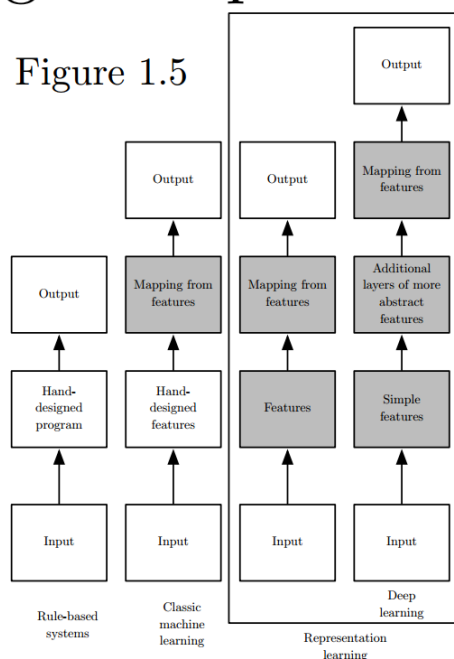
Learning 和 AI 到底有什么关系呢？

To summarize, deep learning, the subject of this book, is an approach to AI. Specifically, it is a type of machine learning, a technique that allows computer systems to improve with experience and data.

概括的说，本书的主题，深度学习，只是一种人工智能的计算方法。具体来说，它是一种机器学习（Machine Learning）方法，一种允许计算机系统根据经验和数据改进的技术。

Learning Multiple Components

Figure 1.5



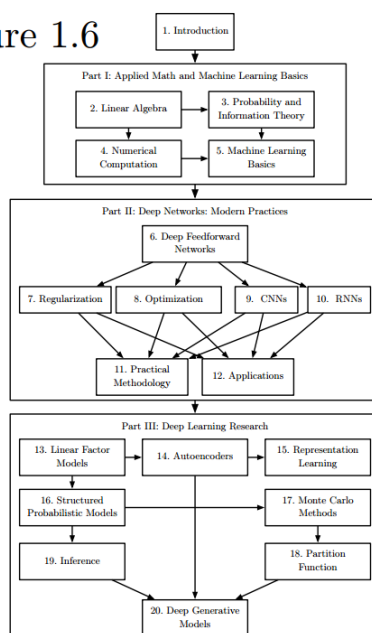
(Goodfellow 2016)

上图（图 1.5）给出了一些机器学习方法的流程图（其中也包含了我们的深度学习）。最早的基于规则的系统（专家系统），我们只能人为的给出一些规则来辅助计算机判断（比如玩超级玛丽，看见前面来了一个板栗小兵就要跳起来不然就会挂掉）。然后后续的进入了机器学习的时代，我们可以简单地提取一些特征来辅助我们的专家系统（例如设备故障诊断中将时

域信号经过傅里叶变换转换为频域信号进行判断，以及图 1.1 的坐标系变换)，将数据转换为我们关注的特征，然后用专家系统进行判断。进入了表征学习阶段后，很庆幸我们不需要在人为的规定一些专家系统了（因为好多规则我们自己都说不明白），我们告诉计算机问题和对应的答案（监督学习，也叫做有教师学习），然后让计算机自己去学习，至于怎么学的我们不管。然后根据学习的好坏来优化我们的参数（学会了就喂个枣，学错了就一顿毒打，像不像养小狗），这种根据计算结果的误差来调整网络参数的方法也被叫做反向传播（Back Propagation）。到了深度学习，我们的网络除了一些简单的特征以外，计算机可以学习到一些更加抽象的特征（如 Neural Style 里面学习的画风），使得我们的系统越来越“智能”。

Organization of the Book

Figure 1.6



(Goodfellow 2016)

上图（图 1.6）是本书的章节安排，我们由浅及深地将深度学习拆成了三个部分，第一阶段面向小白读者，讲解本书中所用到的最本地数学理论，第二部分面向只是想应用深度学习的读者（比如你只想应用卷积神经网络，那你可以直接看 6-9-12 章），而第三部分，正如他的名字，提供给深度学习的研究者们，介绍了被认为对深度学习的进一步研究十分重要的更

有趣的想法。

Historical Waves

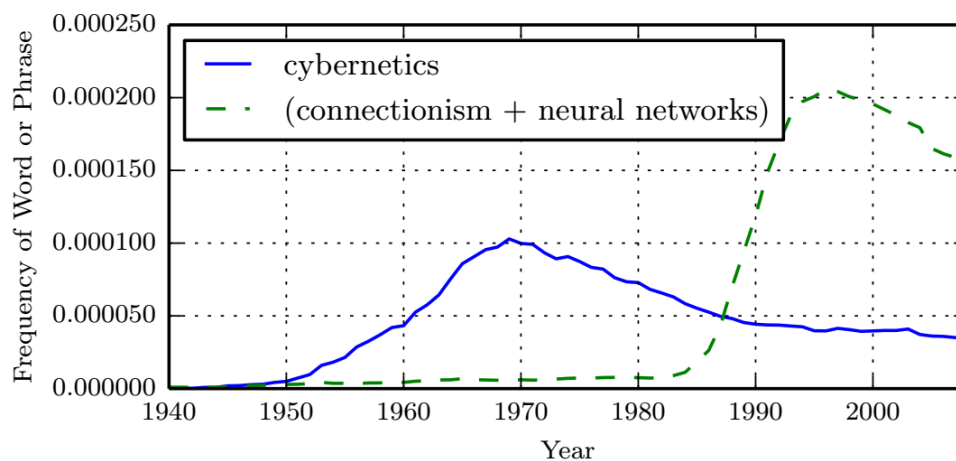


Figure 1.7

(Goodfellow 2016)

图 1.7 展示了人工神经网络研究的发展史上三次浪潮中的两次。我们统计了三个词的词频，控制论（cybernetics）、联结学（connectionism）以及神经网络（neural networks），以此来代表深度学习流派的发展。

The first wave started with cybernetics in the 1940s–1960s, with the development of theories of biological learning (McCulloch and Pitts, 1943; Hebb, 1949) and implementations of the first models such as the perceptron (Rosenblatt, 1958) allowing the training of a single neuron. The second wave started with the connectionist approach of the 1980–1995 period, with back-propagation (Rumelhart et al., 1986a) to train a neural network with one or two hidden layers. The current and third wave, deep learning, started around 2006 (Hinton et al., 2006; Bengio et al., 2007; Ranzato et al., 2007a), and is just

now appearing in book form as of 2016. The other two waves similarly appeared in book form much later than the corresponding scientific activity occurred.

第一次浪潮是在 1940-1960 间出现的控制论，伴随着生物学习理论的发展和第一个允许训练的单个神经元的感知元模型的建立。第二次浪潮是 1980-1995 年出现的联结学，那时出现了反向传播算法，我们可以训练拥有一到两个隐层的神经网络了。现在是第三次浪潮，从 2006 年开始的深度学习，直到 2016 年才被写进书中。不过另外两次浪潮也一样，也是滞后于相关科技的发展很长时间才被收录。

Historical Trends: Growing Datasets

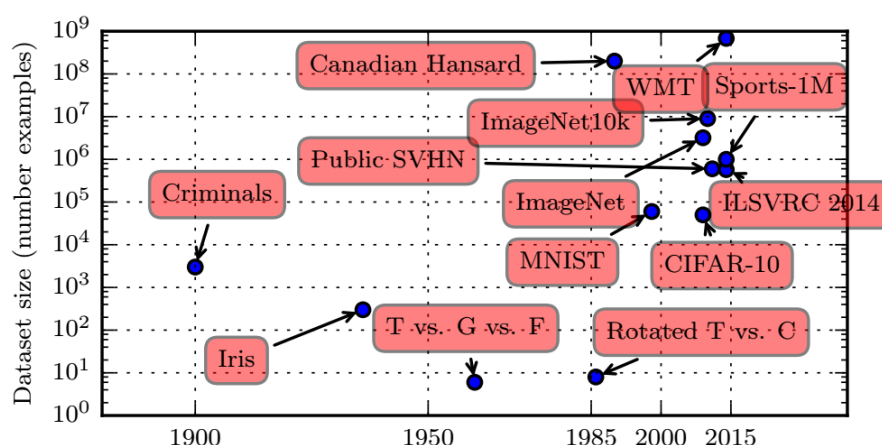
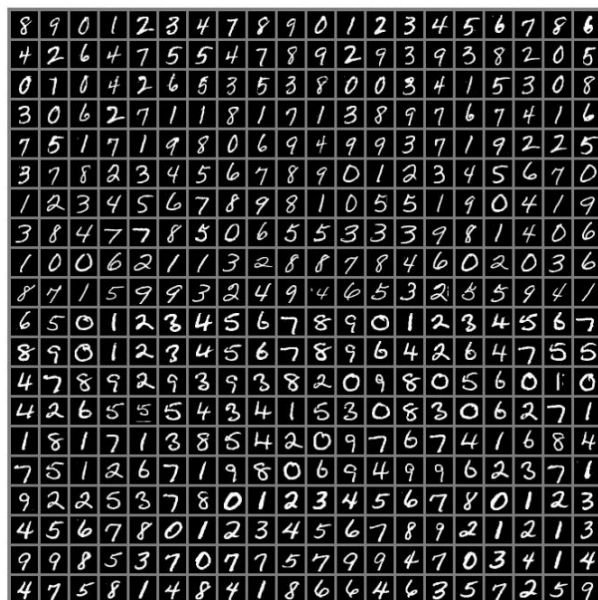


Figure 1.8

(Goodfellow 2016)

一个数据驱动的学科的发展必然伴随着数据的激增。从图 1.8 可以看出，进入 21 世纪后，尤其是 2006 年后（深度学习的第三次浪潮），诞生了许多从十万到千万量级的数据集，这些数据集也反过来促进了深度学习的发展。

The MNIST Dataset



8	9	0	1	2	3	4	7	8	9	0	1	2	3	4	5	6	7	8	6
4	2	6	4	7	5	5	4	7	8	9	2	9	3	9	3	8	2	0	5
0	7	0	4	2	6	5	3	5	3	8	0	0	3	4	1	5	3	0	8
3	0	6	2	7	1	1	8	1	7	1	3	8	9	7	6	7	4	1	6
7	5	1	7	1	9	8	0	6	9	4	9	9	3	7	1	9	2	2	5
3	7	8	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	0
1	2	3	4	5	6	7	8	9	8	1	0	5	5	1	9	0	4	1	9
3	8	4	7	7	8	5	0	6	5	5	3	3	3	9	8	1	4	0	6
1	0	0	6	2	1	1	3	2	8	8	7	8	4	6	0	2	0	3	6
8	7	1	5	9	9	3	2	4	9	4	6	5	3	2	3	5	9	4	1
6	5	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7
8	9	0	1	2	3	4	5	6	7	8	9	6	4	2	6	4	7	5	5
4	7	8	9	2	9	3	9	3	8	2	0	9	8	0	5	6	0	1	0
4	2	6	5	5	5	4	3	4	1	5	3	0	8	3	0	6	2	7	1
1	8	1	7	1	3	8	5	4	2	0	9	7	6	7	4	1	6	8	4
7	5	1	2	6	7	1	9	8	0	6	9	4	9	9	6	2	3	7	1
9	2	2	5	3	7	8	0	1	2	3	4	5	6	7	8	0	1	2	3
4	5	6	7	8	0	1	2	3	4	5	6	7	8	9	2	1	2	1	3
9	9	8	5	3	7	0	7	7	5	7	9	9	4	7	0	3	4	1	4
4	7	5	8	1	4	8	4	1	8	6	6	4	6	3	5	7	2	5	9

Figure 1.9

(Goodfellow 2016)

图 1.9 是 Yann LeCun 建立的 MNIST 数据集，MNIST 数据集中有 0-9 的手写数字的扫描数据以及相应的标签。这个简单的分类是深度学习研究中最简单也最为广泛使用的测试。尽管这个任务对现代的技术来说很简单，但仍然很流行。Geoffrey Hinton 称其为“机器学习的果蝇”，意味着机器学习研究者可以在可控的实验条件下研究他们的算法，正如生物学家经常使用果蝇来帮助研究一样。

Connections per Neuron

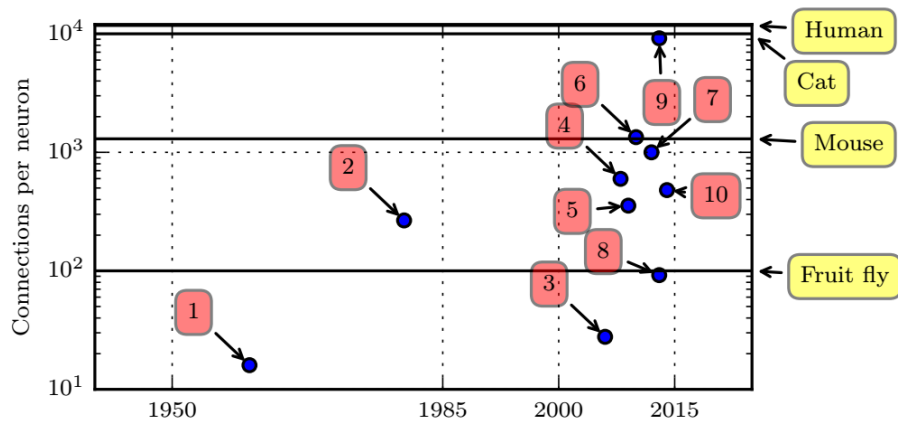


Figure 1.10

(Goodfellow 2016)

最初人工神经网络的单个神经元的连接数的设定受限于硬件的能力，现在单个神经元连接数设定主要是出于网络设计的考量。我们可以看到现有的技术和硬件设备足以支持我们搭建一个连接数媲美人类的神经元。

图 1.10 中的数字分别为：

1. 自适应线性元
2. 认知机
3. GPU 加速卷积网络
4. 深度玻尔兹曼机
5. 无监督卷积网络
6. GPU 加速多层感知机
7. 分布式自编码器
8. 多 GPU 卷积网络

9. COTS HPC 无监督卷积网络

10. GoogLeNet

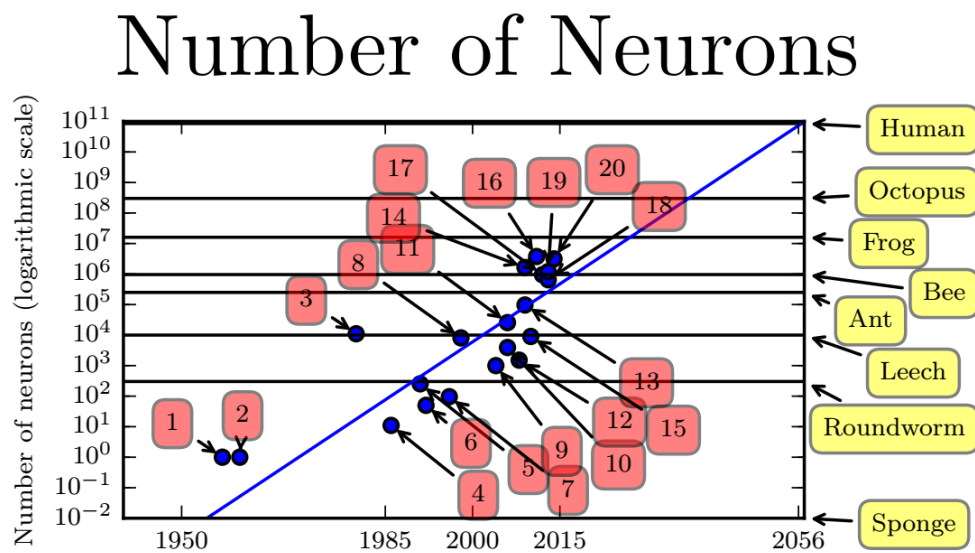


Figure 1.11

(Goodfellow 2016)

自从引入了隐藏元 ,人工神经网络的规模每 2.4 年会提升一倍 ,但是由于收到硬件的局限 ,
我们的神经元总数距离人类仍有很大差距。不过已经能够解决许多问题了。

图 1.11 中的数字分别为 :

1. 感知机
2. 自适应线性元
3. 认知机
4. 早期的反向传播网络
5. 用于语音识别的递归神经网络
6. 用于语音识别的多层感知机

7. 平均场 sigmoid 信念网络
8. LeNet-5
9. Echo state 网络
10. 深度信念网络
11. GPU 加速卷积网络
12. 深度玻尔兹曼机
13. GPU 加速深度信念网络
14. 无监督卷积网络
15. GPU 加速多层感知机
16. OMP-1 网络
17. 分布式自编码器
18. 多 GPU 卷积网络
19. COTS HPC 无监督卷积网络
20. GoogLeNet

Solving Object Recognition

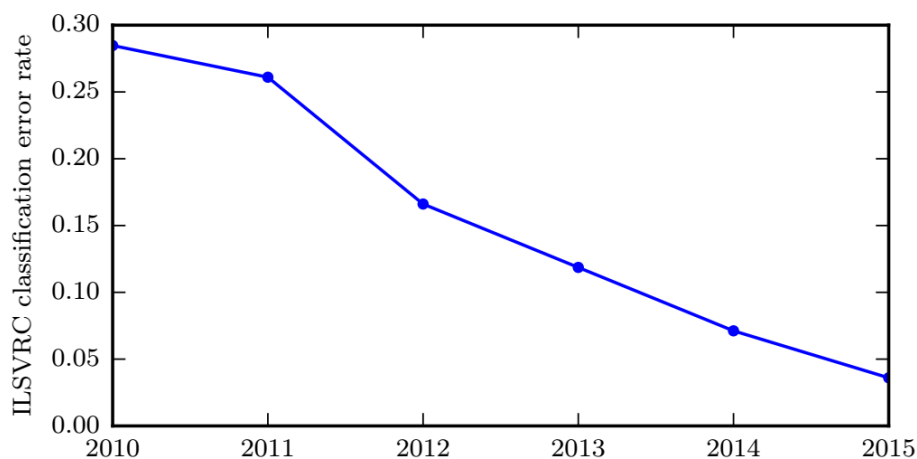


Figure 1.12

(Goodfellow 2016)

图 1.12 是深度学习在图像分类问题的错误率,自从深度学习参与 ImageNet 竞赛(AlexNet)以来,已经连续获得多年冠军,且错误率逐年降低,已经超越了人类的水平。

深度学习在其他的科学领域也做出了贡献。现代的用于物体识别的卷积神经网络提供了神经学家可以研究的视觉处理模型。深度学习也为科技领域中处理海量的数据、做出有用的预测提供了工具。深度学习通过预测分子之间的交互来帮助制药企业设计新的药品、也曾被用来寻找亚原子和自动解析用于构建人脑三维图像的显微镜图像。我们期望深度学习未来能在更多的领域生根发芽。

现在,我们有 Nvidia 的 Pascal GPU, 我们有 Tensorflow、Caffe、Troch、MXNET 等优秀的框架,我们很幸运生在了一个开源的时代。深度学习的未来充满着挑战和机遇,而它也将去探索新的边界。感谢大家支持 AG-Group, 感谢大家支持极视角的极市平台。

关注极市平台获取更多干货及在线分享信息



关注 AG Group 获取好玩，很酷的科技

