1. 深度学习

最近几年，深度学习(deep learning)在很多领域上都取得了很好的效果，如图像识别、语音识别、自然语言处理。深度学习从而成为近年来学术会议的焦点， 工业界的炫技利器， 甚至在创业公司中都掀起了一波深度学习热潮， 拍衣搜商品、 智能机器人、 Siri语音助手、无人车都借助这一技术提升性能。 可能你还是巴望在深度学习门槛前的一个，在读完本章后，让我们看看“深度学习”有多简单。

* 1. 横空处世？

什么是深度学习？

深度学习其实在很早的时候叫“神经网络”， 但是由于缺乏训练数据和 计算能力瓶颈， 一直很难训练好一个神经网络。 要么网络大了（网络中的参数相对数据多了），模型过拟合，要么网络小了，模型不能处理复杂问题。最近随着大数据和计算能力（GPU）的快速发展，才使神经网络改头换面为“深度学习”并变成很多任务的winner，所以说，**深度学习的成功，只是由于有了更多数据和更多计算资源而已**。

神经网络的雏形为perceptron,

最后，再下面介绍具体模型之前，我不得不稍微介绍几个深度学习领域的鼻祖（字典序），肯定不能完全覆盖，但至少下面看到他们名字的时候能够有大概印象。

Yann Lecun： 先后就职于多伦多大学（博士后），AT&T图像处理研究部门（从小弟到老大），纽约大学教授，目前到了Facebook做AI Research（人工智能研究院）的老大。

* 1. Convolutional Neural Network (CNN)

在深度神经网络的众多类型中，卷积神经网络（Convolutional Neural Network， CNN）算是最广为研究的一种了，原因一是CNN结构简单，方便实现，二是CNN可以很好地处理图像相关问题。 在本节中我们首先介绍CNN的基本结构，然后深入看看近些年来都有哪些工作在CNN上进行改进。

CNN最早是Yann LeCun在98年提出用来做手写数字（0-9）识别的，在这个任务中，他们建立了一个叫LeNet-5的网络，用来从原始图像得到可以分辨图像类别的特征，此后随着CNN的发展又出现了很多CNN网络结构的变种，但其基本机构都是相似的。基本CNN网络结构包含三种类型的层：卷积层（convolutional layer）, 池化层（pooling layer）和全连接层（fully connected layer）， 以输入为图像为例。

* 卷积层

卷积层由很多特征图（feature map）组成，每个feature map内的神经元与上一层中几个相邻的神经元相连，这几个相邻的神经元就叫做感受野（receptive field）。**，**是一个卷积核（kernel）对输入图像进行卷积操作的输出。

**卷积神经网络也是一个层级结构的前向神经网络。除了输入层和输出层，中间的隐层的目标是从原始数据学得一系列非线性组合，使得可以很好地结合当前任务表征信号。以图像为例，输入为像素级数据，每个像素本身不含有太多信息，但是组合起来所得到的更高级别的特征可以表示物体类别等信息。尽管我们的神经信号不同与图像信号，但是文本分析[147]，语音识别等其他任务都受了图像的启发，因其方便可视化，所以我们采用图像进行解释。以输入层和第一层隐层中的一个神经元节点连接为例，如图**3.7**所示。图中黄色区域表示输入，为**32\*32\*3 **大小的数据，比如可表示一张深度为**3**（**3 **通道），宽、高均为**32 **的图像。图中的蓝色圆点表示与输入层相连的一个隐层节点，连接项由权重参数组成。和大脑中视觉皮层的接收域（**receptive fields**）一样，我们在卷积神经网络的隐层中也设置接收域，该接收域可以视作一个在输入中进行扫描的线性滤波器，为了**

* 池化层
* 全连接层
  1. Deeper than deeper?

上一节中我们已提到，从12年到15年，图像分类竞赛ImageNet冠军依次使用了越来越深的网络，那么我们要问，是不是只要不过拟合，网络越深越好使？拍拍脑袋想好像是这样，网络越深，参数拟合能力越强， 但在Highway Network一文中作者提出了一个相反的现象，［1］中称此现象为degradation, 如图所示，

* 1. Recurrent Neural Network (RNN)

c) Convolutional Neural Network

参考文献：

［1］.