**第1章**

**P001**

各位同学，各位读者大家好，当你打开本书的时候呢，就说明你想开始本领域的学习了，你能选择我们的这本教材作为学习资料，这是让我们非常高兴的一件事。所以，作为本书的作者，我要谢谢大家！

先说一下，本书是一共19章，前十章为理论篇，后九章为实践篇，这也是本书的一个特色，就是说我们不只是介绍神经网络的理论，我们还要给大家介绍网络的设计和应用。网络实验主要是基于matlab的神经网络工具箱和深度学习工具箱。对硬件的要求也不高，一台PC机就可以了。由于matlab提供了丰富的调用函数，让我们使用起来简洁方便，哪怕是matlab零基础的同学也没有关系，很快就能上手的。

本书作为中国海洋大学核心通识课的配套教材，选课的学生中刚入学的大一的零基础的文理科同学，因此，为了照顾不同基础的同学，我们在每一章都添加了一些跟本章内容有关的数学理论。大家都知道了，本书的名字是《神经网络与深度学习基础》，顾名思义，我们讲的是一些神经网络与深度学习的基础模型，像现在比较流行的一些复杂神经网络，比如，NAS，就是神经网络结构搜索，等动态模型；还有由Attention机制组成的Transformer模型,在本书里面都没有被涉及到，为什么呢？因为现在的这些大规模的神经网络，它作为一些基础网络的一种组合也好，一种优化也好，一种拓展也好，只有当你学会了基础模型，你才会容易理解、容易掌握这些比较流行的复杂模型。我们计划于2022年出版一本海大研究生课程的配套教材《神经网络与深度学习》，里面会有对目前比较流行的大规模复杂神经网络的介绍。另外，为方便大家学习，我们还计划将这两本教材建设成立体教材，除了教材本身提供的文字和二维码音频讲解之外，拟将两门课程完整的视频上传到网络平台，初步选择“智慧树”网络平台，请感兴趣的同学和读者给予关注和支持。

下面说一说第一章的内容。第一章是由初识人工神经网络、人工神经网络的发展历史、人工神经网络的类型和人工神经网络的应用领域这4个部分组成的。那么，第一部分呢，其实就是要告诉大家什么是人工神经网络，我们由两个经典的例子给大家引出神经网络的概念，即，人工神经网络是对生物神经网络的结构和工作机理的模拟，可以由电子线路或仿真软件来实现。阿法狗围棋软件就是一款神经网络的仿真软件；曼切斯特大学开发的目前最强的人工大脑spiking就是由电子线路搭建而成的，是由100万个CPU（中央处理器）和1200块相互连接的电路板组成的超级计算机。

再说神经网络的发展历史，书上是先讲的人工智能的发展历史，然后再讲神经网络的发展历史，因为神经网络和人工智能是密不可分的，其实，它就是人工智能的一个新技术。至于神经网络的类型，是讲基础模型的类型，像对抗模型、动态模型等会在第二本教材里进行介绍。神经网络的应用领域呢，是介绍几个比较经典的传统的应用领域，现在的应用领域越来越广泛了，肯定不止这些了，但是，我们就是想让同学们知道，神经网络是很有用的，而且现在是发展的如火如荼。通过本章的学习，你会对神经网络的概念、历史、应用有一个简单的了解，但是神经网络的数学模型和学习算法等内容，你得往下学才会明白。

P004

此处要着重解释的就是，神经网络与传统的人工智能是有区别的。传统的人工智能呢，就是我们常说的把大脑视为黑箱，不研究大脑的运行机理，只模拟大脑的功能，即，以演绎逻辑的符号系统为理论框架，以冯诺依曼计算机为执行环境的人工智能，而且这些传统的人工智能已经在很多领域都取得了成功的应用。至于，为什么把这个传统的人工智能叫symbolism呢？那是因为演绎逻辑本身需要一个符号系统的，比如说，数学上的推理，必须有一套严谨的数学符号系统，所以呢，这个基于符号系统的演绎逻辑推理的人工智能，就被叫做symbolism,就是在英文单词符号symbol后面加一个ism，有翻译成符号主义的，也有翻译成符号机制的。与传统人工智能不同，人工神经网络研究大脑的运行机制，而组成大脑的生物神经元是连接成网、并行处理和传递信息的，这种分布并行的信息处理方式与符号系统的表达方式完全不同，它强调连接机制，因此，这种新生的人工智能就被叫作connectionism,就是在英文单词连接connection后面加一个ism，可以叫被作连接机制或连接主义。在这里，我们还要强调一下，人工智能AI（artificial intelligence）这个单词，是在1956年达特茅斯会议上提出来的，但是有关人工智能的科学研究早就开始了，人工神经网络的研究也绝对不是1956年之后才有的，比如，在1943年就有了MP模型，所以，这一点，同学们不要误解了好吗？

**第2章**

P014

本章开始介绍神经元和神经网络的数学模型，首先呢，在基础部分，介绍了点积的定义与计算。有的同学可能会说，这么简单的东西在线型代数里面都学过了，为什么还要讲啊？大家别忘了，我们这是核心通识课的配套教材，选课的同学里面有大一刚入学的，他们是没有学过线性代数的，向量啊，点积啊，矩阵运算啊，这些对他们来讲，都是陌生的，所以我们在介绍数学模型之前必须补充一点基础知识。但是，我们用到的线性代数的内容并不多，后面的章节也只会用到矩阵的乘法和转置运算，而且，书上都会介绍，所以，没有学过线型代数的同学完全不用担心自己的学习基础。

本章第二节是介绍生物神经元模型。为什么要介绍生物神经元呢？因为人工神经元是受生物神经元的启发而构建的数学模型，所以，要分析生物神经元所具备的特征，以及由生物神经元连接成的生物神经网络的特征，人工神经元和人工神经网络必须具有生物神经元和生物神经网络的最基本的特征，才能具有人工智能。

在介绍完生物神经元模型之后，自然，就要介绍人工神经元模型了，人工神经元的数学模型是由连接函数和激活函数（作用函数或激励函数）组成的。在这个数学模型里面，激活函数起到了一个非常重要的作用，可以说不同的激活函数决定着神经元的不同性能。本章的这一部分，介绍了各种激活函数，这个是非常重要的，同学们一定要认真学习。

人工神经元的数学模型介绍完了，自然就要介绍由人工神经元组成的人工神经网络的数学模型，包括单层前馈网络、多层前馈网络、侧反馈网络、循环网络的结构和数学模型。可能有的同学会想，第2章就能够把这么多的网络都介绍完吗？不是这样的，每一个网络有很多内容要学习，本章只是把网络的数学模型和结构展现给你，至于网络的算法、算法的推导、网络的训练，网络的应用、网络的性能评估等等，都是要在后面章节介绍的。

介绍了几种类型的网络之后，本章又介绍了一些常用的网络约定，这部分内容难度不大，但是同学们是一定要记住的，因为后面的学习是经常要用到的。本章最后介绍了神经网络的学习方式，就是监督学习和无监督学习两种学习方式，可能有的同学会说，还有半监督学习呢？那是随着深度学习兴起之后才有的，也是在这两种基本的学习方式之上的一些改变吧。

同学们，通过本章的学习，你会对神经元和神经网络有更进一步的认识，尤其是你学习了他们的数学模型之后，对神经网络会有一个理性上的认识。

P018

本章在第2节总结了生物神经元和生物神经网络的6个基本特征，其中前3个基本特征是生物神经网络应该具有的，后3个基本特征是生物神经元所具有的。由于人工神经元是在生物神经元的基础上提炼出来的数学模型。所以，生物神经元所具有的3个基本特征，人工神经元也应该具有。图2.3所示的人工神经元的结构，是心理学家麦克卡罗斯和数学家批次在1943年提出的，即著名的M-P模型。神经元模型的外部输入X1，X2 ，XN，可以自于其他神经元的输出，也可以是来自网络外部的信息，每一个输入会有一个连接强度的，就是说每个输入会乘以一个系数，这个系数叫做连接权重。

人工神经元的数学模型是由两部分组成的，连接函数的输出作为激活函数的输入。连接函数就是神经元的外部输入和各自的连接权重的一个加权和，如果用向量形式表示，连接函数就是输入向量和权重训练的点积，其输是一个标量，用net表示。将net输入到激活函数（最早为阶跃函数），得到的输出，用o表示，就是这个神经元的输出，输出可以为正，也可以为负。

人工神经元的输出可以为正，也可以为负，可以理解成是对“生物神经元的输出可以起刺激作用，也可以起抑制作用”的一种模拟。连接函数就是对人工神经元的所有输入的累积。最早的激活函数是阶跃函数，其实就是一个阈值函数，就是模仿生物神经元的阈值。讲到这，不用我说大家也都知道了，人工神经元具备了生物神经元的3个基本特征。

可能有的同学会想说，人工神经元的数学模型就是通过观察生物神经元得到的呀？是呀，我们就是通过观察生物神经元，提炼出了人工神经元的数学模型。你别小看这个简单的数学模型，它是组成神经网络的基本单元，如果取不同的激活函数，这种单元会有很多种类，然后再由他们组成各种各样的网络，就构成了我们今天这个庞大的人工神经网络家族。

P027

此处给同学们思考问题提一点线索，比如说第1题，人工神经元满足了生物神经网络的几个基本特征，前面都是分析人工神经元满足了生物神经元的哪些基本特征，这里，把生物神经元看成生物神经网络的一部分，就好了。

再一个,就说这个sigmoid函数作为神经元的激活函数，会使后面神经元的输入发生偏移，这句话是怎么理解呢？你从sigmoid函数的输出范围上看，一般来讲神经元的输入都有正有负，对吧？而且输入范围是很大的。但是sigmoid函数作为激活函数的神经元，它的输出范围是什么样呢？它的输出作为下一层与它相连的神经元的输入，那是不是就把下一层神经元的输入限制在一定的范围内了呢？

再说这个ReLU函数,你从它的图形可以看出它是左侧单向饱和的,而Sigmoid函数，是两端饱和的。当神经元的网络输入net的绝对值比较大的时候，无论为负数还是正数，Sigmoid函数进入饱和区，当函数的输入发生变化时，函数的输出变化得很小，即，下降的梯度非常小，因此会产生梯度弥撒现象。但是ReLU函数，是单方向左侧饱和的，它的右侧并不是饱和的。大家可以从这个角度去理解问题，好吗？可能有的同学会说老师现在还没讲网络呢，只讲到神经元的模型，这个问题是不是应该往后放一放啊？嗯，我当时也这么考虑过，但是想到这一块儿都是在讲激活函数的，就提了一个这样的问题，后面还会讲到梯度弥撒和梯度爆炸问题，现在，担心同学们因为没学后面的内容对此问题理解得不是很清楚，所以我就给大家提个醒了。

**第3章 感知器**

P036

本章开始介绍感知器，感知器是最早出现的人工神经网络，那么,在对感知器的数学模型和训练的介绍中，会用到矩阵的乘法以及矩阵的转置等运算，所以，本章在基础知识部分就给大家先介绍了矩阵的一些基本运算。

其实最早出现的离散单输出感知器，就是一个神经元，严格上讲还算不上一个网络，但是呢，它是神经网络的一个了不起的开始，已经具备了神经网络所拥有的一些特征，它采用监督学习，有自己的训练算法，能够基于数据集训练权重，也就是说已经具有了一定自主学习的能力。当然，由于离散单输出感知器的结构过于简单，它的性能有一定的局限性，比如，解决不了非线性分类问题，最典型的例子就是解决不了“异或”问题，当时明斯基还主观地断言，即使多层感知器也解决不了“异或”问题，那么到底能不能解决呢？大家往下学就知道了。

在离散单输出感知器之后，又出现了离散多输出感知器、连续多层感知器，感知器的性能逐步得到了提高。

本书在每章开始，都会提出“本章问题”，是让大家带着问题去学习本章的内容，学完后，能够找到问题的答案。我相信大家通过本章的学习，会找到四个本章问题的答案的。

P051

通过前面的学习，大家都知道了，离散单输出感知器是解决不了异或问题的，那么,是不是像明斯基所说的那样，多层感知器也解决不了异或问题呢？请大家看图3.8，图3.8展示的是一个两层的感知器，神经元的权重和阈值都已在图上标明,3个神经元的连接函数都是输入的加权和，激活函数都单位阶跃函数。现在我们就要看一看，这样的一个两层感知器能不能解决异或问题，也就是说，能不能实现表3.1所示的异或问题的真值表里面的内容。

当X1为0，X2为0，即，神经网络的输入为[0,0]的时候，神经元AN1连接函数的输出值net=0×1+0×（-1）-0.75=-0.75，因为net小于零，所以，经过激活函数后，神经元AN1的输出为0。同样的道理，神经元AN2的输出也为0。此时，AN1和AN2的输出作为神经元AN0的输入，AN0的连接函数的输出值net=0×1+0×1-0.75=-0.75，那么经过单位阶跃函数之后，神经元AN0的输出仍为0。

当X1为0，X2为1，即，神经网络的输入为[0,1]的时候，神经元AN1连接函数的输出值net=0×1+1×（-1）-0.75=-1.75，因为net小于零，所以，经过激活函数后，神经元AN1的输出为0。神经元AN2的连接函数的输出值net=0×（-1）+1×1-0.75=0.25，因为net大于零，所以，经过激活函数后，神经元AN2的输出为1。此时，神经元AN0的输入是AN1和A2的输出，它的连接函数的输出值net=0×1+1×1-0.75=0.25，那么经过单位阶跃函数之后，神经元AN0的输出为1。

同样的道理我们可以得到，当X1为1，X2为0的时候，网络的输出值，也可以得到，X1为1，X2也为1时，网络的输出值。这两种情况留给大家们自己算一下吧。当你做出来之后，对比一下表3.1的内容。这个双层感知器是不是实现了异或运算的真值表里面的内容呢？如果实现了，那是不是就说明双层感知器能够解决异或问题呢？明斯基的断言是不是过于主观了呢？

**第4章**

P059

BP神经网络,即误差反传神经网络是一种经典的前馈神经网络，也可以说，BP神经网络就是一种多层感知器，只不过，与最早出现的感知器相比，BP网络的神经元的激活函数不再使用单位阶跃函数，而是经常使用sigmoid函数或者双曲正切函数，如果将BP网络用于分类，其分类面不再是平面或者超平面，这样，就比较容易实现非线性分类。

BP网络最精华的部分是BP算法，而BP算法是基于梯度下降算法的一种误差反传算法，所以，本章在基础知识部分先介绍了梯度下降算法。但是基于梯度下降的算法也有它的弊病，就是容易出现局部极小点问题，它也会导致了BP网络的训练陷入局部最小点，从而，影响网络的精度，这也成为了BP网络被人诟病的一个缺点。后来，为了克服或者缓解局部极小点问题，又出现了一些改进的BP算法，那么，大家对此有没有什么好的想法呢？

P067

BP算法的主要思想就是用输出层的误差调整输出层的权重矩阵，并用此误差估计输出层的前一层的误差，并调整前一层的权重矩阵，再用前一层的误差估计更前一层的误差，并调整更前一层的权重矩阵。这种与输入信号相反方向的逐层计算误差并调整权重的过程，就叫做误差反传。

可能有人会说，为什么非要从后面一层一层地往前推？这多麻烦呀。可是你想一下，采用监督学习的神经网络，它的误差是输出层神经元的输出值和理想值的误差，隐含层的误差是不能直接求出来的，那么没有误差，基于误差的梯度下降算法怎么实现呢？所以，只能由后一层的误差去估计前一层的误差。直到现在，深度卷积神经网络和深度循环神经网络也是采用这种误差反传的方式去调整权重的，只是调整公式和BP网不一样，误差反传是一直在用的。

具体的算法推导从公式4.15开始，从公式4.15可以看到，权重的调整量是与误差的下降梯度成正比的，也就是说，权重的调整要使误差越来越小。但是误差对权重的偏导是不能直接算出来的，所以我们就采用了复合函数求导的方法.现在，假设权重是输出层神经元j与它前一层的神经元i的连接权重，正如公式4.16，将误差对权重的偏导变成误差对神经元j的网络输入netj的偏导，再乘以netj对权重的偏导。由公式4.18可以看到，netj对权重的偏导是很容易求到的，下面就看如何计算误差对netj的偏导。将网络误差对神经元的网络输入net的偏导定义为灵敏度函数δ。当神经元为输出层神经元时，它的灵敏度函数值是很容易由公式4.21和公式4.22求得。但是，如果神经元是隐含层的神经元，那么，它的灵敏度函数值就不能直接求出来了，必须由它后一层的神经元的灵敏度函数值估算出来。所以，误差反传也可以说是灵敏度反传。

**第5章**

P081

1988年，Ｍoody和Ｄarken根据生物神经元具有局部响应这一特点，将径向基函数引入到神经网络设计中，提出了一种以径向基函数为激活函数的神经网络结构，即径向基神经网络.

RBF神经网络的结构类似于BP网络，是由输入层、隐含层和输出层组成的前馈网络。输入层的神经元只负责将输入信息传送到隐含层；隐含层神经元的模型与BP网络的神经元不同，它使用欧式距离作为连接函数，高斯函数作为激活函数；输出层神经元为线性神经元。

径向基神经网络的学习总体上被认为是监督学习，即，网络在训练时需要理想输出。但是，与BP网络的训练不同，径向基神经网络不需要训练，它的输出层的权重是由伪逆矩阵求解法计算得到的，不像BP网络的权重是通过误差反传算法不断迭代调整的，所以径向基神经网络的运行速度是非常快的。这也是它的一个优势。

P089

这里要强调的是，径向基神经网络总的来讲是采用监督学习的，但是，它是分两个阶段的，第一个阶段是确定径向基函数的中心，第二个阶段是确定输出层神经元的连接权重。第一个阶段是采用无监督学习方式，用聚类的方法，当然也可以用其他的方法，来确定隐含层神经元的径向基函数（激活函数）的中心**ci**。**ci**是一个向量，它的维数是跟输入向量的维数是一样的，也就是说，与输入层神经元的个数是一样的，你可以把**ci**理解成隐含层神经元与输入层神经元的连接权重组成的向量。第二个阶段是采用监督学习的，由伪逆矩阵求解法输出层神经元的连接权重，这时，是需要理想输出值的。

**第6章**

P099

本章介绍的反馈神经网络与前几章介绍的前馈神经网络不同，前馈神经网络的输出仅由当前的网络输入和权重决定，而与网络先前的输出无关。反馈神经网络的突出特点是具有联想记忆功能，即，网络的输出与网络先前的输出有关，所以，反馈神经网络的结构与前馈神经网络不同，最早的反馈神经网络Hopfield网络是一个单层网，每个神经元的输出都反馈到其它神经元的输入端，这样网络就能够对输入信号逐渐进行“修复”和“加强”。Hopfield神经网络的权重是由样本数据计算得来的，并不是像BP网那样通过一次次迭代估算出来的。如果说Hopfield网络是采用一种监督学习的话，那么，它的标签儿就应该是被网络事先记忆的那些数据样本；对于以后的输入，网络会输出与之相近似的记忆样本，即实现了联想记忆。

Hopfield网络的一个典型的应用就是求解旅行商问题，旅行商问题是组合优化中的一个ＮＰ难题，所以本章在基础知识部分对ＮＰ问题进行了介绍。

本章最后介绍了一种典型的局部回归网络，即,Ｅlman神经网络。Ｅlman神经网络可以看作是一个具有局部记忆单元和局部反馈连接的前馈神经网络。这个网络非常重要，是组成深度循环神经网络的基础网络。

P105

Hopfield网络的稳定性证明是本章的一个难点,尽管Ｃohen和Ｇrossberg于1983年，给出了Hopfield网络的稳定性定理, 但是，本节对Hopfield网络的稳定性分析采用了另一种方法。

公式6.7是Hopfield网络的能量函数，也叫网络的稳定性度量函数，是选用了著名的Ｌyapunov函数。

下面就是证明，当一个神经元K的状态发生变化的时候，它引起的能量函数变化总是小于0的，也就是说新能量E撇总是小于原来的能量E，那么就可以说：随着神经元状态的变化，网络的能量函数总是下降的，网络最后趋于稳定。

首先，将公式6.7里的3项进行拆项，把和神经元K有关的项单独拿出来，由于在新的能量函数中，只有和神经元K有关的项发生了变化，其他项都没有发生变化，那么，新旧能量的差就是和神经元K有关的那些项的差。这样，就很容易证明，能量函数变化总是小于0的，由于，神经元K为任意神经元，因此，可以说，网络最后趋于稳定。

**第7章**

P114

前面几章介绍的神经网络均采用监督学习方式，但是，人在认知过程中，很多情况下，常常是靠“无师自通”，即，通过对客观事物的反复观察、分析比较，自行挖掘出其内在的规律，从而对具有共同特征的事物进行正确归类。本章介绍的自组织映射网络（SOM神经网络）正是基于这种学习方式，即，不需要理想输出作为参考的无监督学习方式。

本章最后介绍了lvq学习量化神经网络。学习向量化算法是由Ｋohonen提出的一种有监督的学习算法，lvq神经网络允许对输入样本的分类进行指定，即，网络训练是有理想输出的。LVQ网络的学习分为两部分，第一部分，类似于SOM神经网络，采用无监督学习方式对数据样本进行聚类；第二部分，就是竞争层到输出层的连接权重的确认，可以理解成是一个有监督的学习过程。由于lvq神经网络结构的特殊性，使得网络能够以较小的计算量来处理大量的输入样本，所以，非常适合大数据的在线学习。

P120

SOM网络采用的学习算法称为Kohonen算法。大家要注意的是，Kohonen算法与胜者为王算法是不同的。在胜者为王算法中，获胜神经元对周围神经元的抑制是封杀似的，只有获胜神经元才有机会调整其权重向量，从而更接近输入向量，其它神经元没有调整其权重向量的机会；但是，Kohonen算法采用了一个优胜域思想，在这个优胜域里的神经元都可以调整自己的权重向量，还都有机会成为获胜神经元，所以Kohonen算法相比胜者为王算法要柔和一些，胜者为王算法有些太暴力了。

**第8章**

P130

前几章所介绍的神经网络的训练方法都是确定的方法，但是，生物神经网络是按照概率运行的，本章介绍非确定的方法，又称为统计方法。

大家是否还记得我们在学习第四章BP神经网络的时候，我问过大家一个问题，就是，BP神经网络的训练是采用基于梯度下的降误差反传算法，网络训练容易陷入局部极小点。当时，我问大家怎么才能逃出局部极小点呢？大家有什么好方法吗？其实，针对于局部极小点问题，现在出现了很多的改进方法，其中应用最广泛的方法就是基于概率的一些方法，也就是本章介绍的不确定方法。

本章介绍的模拟退火算法，引进了波兹曼概率，添加了不确定因素，以网络性能的暂时变坏为代价，使网络训练有可能跳出局部极小点，最后达到全局最小点。

本章后来又介绍了柯西分布，在模拟退火算法中，用柯西分布取代波兹曼分布之后，温度可以下降的更快。

对于那些由于计算过于复杂，难以得到解析解或者根本没有解析解的问题，蒙特卡罗方法是一个比较有效的求数值解的方法，最常见的就是蒙特卡罗积分了。在神经网络的学习算法中遇到的积分问题也可采用蒙特卡罗方法求解。

本章最后介绍的波兹曼机是深度波兹曼机的基础网络，是非常重要的。

P138

模拟退火算法其实是一个很繁琐的算法，是挺费时间和空间的。但是，现在的机器是越来越强了，费点儿内存和运行时间，好像也不算什么了。算法8.3展示的网络训练的模拟退火算法，只是一个训练周期内的算法，这个大家要注意一下，实际上，我们在真正运行网络的时候，设置几十个、上百个、上千个周期都是有的。大家可以看到，算法8.3引进了波斯曼概率，当网络误差增大的时候，权重的调整量，还是要按一定的概率来接受的，这样做的目的是希望网络训练能够跳出局部极小点，当然啦，网络的性能是暂时变坏了，但是，一旦跳出局部极小点，就会向全局最小点逼近。不过，运气不好的话，跳出一个局部极小点，可能陷入另一个局部极小点，这种可能也是有的。再要强调的就是，波兹曼分布并不是唯一的选择，也选高斯分布或者柯西分布，总之，效果好才是王道。

**第9章**

P153

深度卷积神经网络和传统神经网络有很大的不同，它是由卷积层、池化层和全连接层组成的前馈神经网络，采用监督学习方式，具有一定程度上的平移、缩放和旋转不变性。

这里要强调的就是，本章介绍的卷积神经网络的卷积层的操作，并没有“翻转”的内容，有人就会说，既然叫卷积为什么没有翻转呢？实质上，这里的卷积就是一个互相关操作，你也可以把卷积层的操作理解成是翻转后的操作。毕竟，人家一开始就叫卷积网络，咱们改成互相关网络也没必要。

再有，为什么说卷积神经网络具有平移、伸缩、旋转的不变性？我理解，主要是池化层的作用，池化层的操作是求平均或求最大值，当输入平移一点、伸缩一点、倾斜一点，对它的结果影响并不大，对吧？当然，卷积层也会起一些作用的。

P162

卷积神经网络的信息传播其实与BP网络很像，它也是分两个阶段，一个正向传播和一个反向传播。正向传播，就是将输入信息，通过卷积层、池化层、全连接层等传送到输出层。反向传播其实就是一个通过误差反传算法调整权重的过程。在这里，要强调的是，卷积层神经元的权重就是卷积核里的参数，卷积层神经元的权重和偏置都是要调整的，但是，池化层的参数无需调整。

可能有的人会说，不对呀，池化层里也有参数要调整的。是这样的，如果池化操作之后，还要经历一个激活函数，那会有一个比例系数和一个神经元的偏置，在这种情况下，池化层是有少量参数需要调整的；但是，大部分情况下，只有一个池化操作，那就不需要调整参数了，所以，大家在这个地方不要感到困惑。

卷积神经网络的参数调整，与BP神经网络一样，都是采用误差反传算法，只是，卷积神经网络的池化层是一个下采样层，所以，往回误差反传的时候，就要有一个上采样的过程。这些在本章都有详细的介绍，这是与BP网络的参数调整不太一样的地方。

**第10章**

P169

本书在第6章介绍的反馈型神经网络是有记忆功能的神经网络，与前馈型神经网络不同的是神经元当前时刻的状态，不仅仅和当前的网络输入有关，也和上一时刻的神经元状态有关。那么本章讲解的循环神经网络和第6章介绍的反馈型神经网络有什么相似和不同之处呢？

其实，第6章介绍的hopfeild神经网络和elman神经网络，它们本身就是经典的循环网络。反馈型神经网络强调的是同层神经元之间的连接，即，具有侧反馈；而循环神经网络强调的是输出层的输出要反馈到网络的输入层，但是如果网络只有一层的话，那么反馈型神经网络不就是循环神经网络嘛，对吧？所以，这一点，大家不要困惑。本章将要介绍的深度学习的神经网络就是以第六章介绍的elman神经网络为基础的多层循环神经网络，它具有记忆功能，不要求输入向量和输出向量的维数固定，也就是说的，输入序列的长度跟输出序列的长度都是可以任意改变的，所以，它是一种能力更强的网络，尤其是在处理时序数据时。

P173

循环神经网络的参数学习是通过随时间反传算法，就是BPTT算法，来实现的，其实质还是BP算法，其核心还是梯度下降算法，那么，既然叫做随时间的反向传播算法，与传统的BP算法有什么不同呢？

这与循环神经网络特殊的网络结构有关，因为循环神经网络的神经元的状态与它前面时刻的状态有关。为了方便讲解，可以把网络按时间展开成一个前馈神经网络，即，把每一时刻的隐含层看作是前馈神经网络的一层，这样，网络就具有了若干个隐含层。这样，循环神经网络就可以按照前馈神经网络中的错误反传算法来计算参数梯度了。因此，参数的真实梯度是所有展开层的参数梯度之和。 明白了这一点，详细的内容看书上的介绍就可以了。

还有，这是本书的最后一个音频讲解，本人的普通话不是很标准，也想过请发音标准的同学帮忙录制音频，但还是觉得自己写的内容由自己来讲解会，感觉更好一些，读者也想听到作者的声音，所以，还是自己来录制了。录制得不完美还请大家谅解，希望每位读者通过本书的学习都有很多收获，再见！