



糊糊

[NNDesign]第一篇:基本概念及标注符号

📖 [读书笔记] 👤 zoeey ⌚ 8个月前 (02-01) 👁 554次浏览 💬 0条评论

r

前言

NNDesign 全称Neural Network Design,是一本神经网络非常好的入门书。这本书是一些知乎大神推荐的,我没有找到中文版,在适应了几天阅读英文版之后发现也不怎么难理解,中间有几章还特意复习了一些线代知识,像我这种数学基础不怎么好的人也能看懂。作为一只被赶鸭子上架的战五渣,还是有必要老老实实的整理一些笔记出来,以供日后复习,希望能坚持下去,虽然对于未知的事物充满了畏惧,但是我还是会很不要脸的说,来啊,快活啊,反正有大把时光~ 😊

标注符号

在书中使用的符号主要有:

标量 — 小写斜体字母: a, b, c

向量 — 小写加粗非斜体字母: $\mathbf{a}, \mathbf{b}, \mathbf{c}$

矩阵 — 大写加粗非斜体字母: $\mathbf{A}, \mathbf{B}, \mathbf{C}$

单输入神经元

神经网络,就是使用计算机模拟人体神经元细胞的工作模式,实现分类、预测等任务。训练神经网络的方式主要有3种:监督学习、无监督学习和强化学习。其中最为常用的是监督学习(给定一组输入输出,当输入应用于网络时,将输出与给定输出作比较,以此来调整权值与偏差)。

典型的神经元实例——单输入神经元示意图如下:

其中, p 代表输入向量(在这里只是一个值), w 为权值向量(对单输入神经元来说也只是一个值), b 代表偏差,那么中间值 $n = wp + b$, Σ 表示“求和”。然后输出 $a = f(wp + b)$, f 的含义是激活函数(activation function),又叫传递函数(transfer function)。激活函数的作用是压缩输出空间,将乱七八糟的输出规范在一个范围或者特定值。

常见的激活函数有:

举个栗子,例如权值 $w=3, p=2, b=-1.5$,那么输出 $a = f(3 \times 2 - 1.5) = f(4.5)$,至于 a 最终的值是什么,就要看选取什么样的激活函数了。如果 f 为 hardlims ,那么 $a = +1$ 。

上图的Icon一栏就是常见的激活函数图像简化图,下图可以看得更清晰些:



多输入神经元

多输入神经元示意图如下：

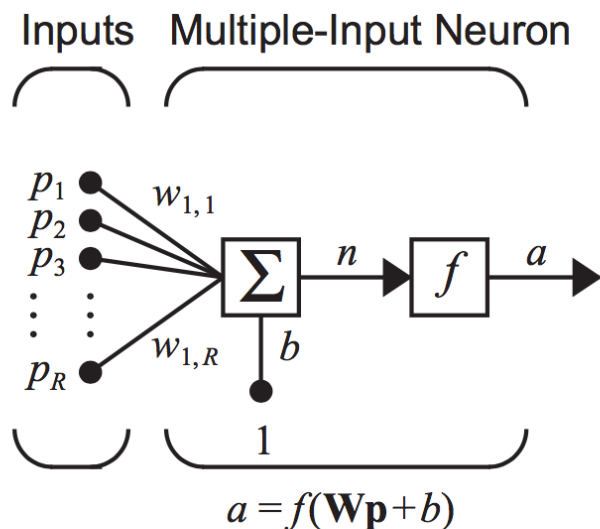


Figure 2.5 Multiple-Input Neuron

与单神经元相比， \mathbf{p} 和 \mathbf{w} 是真正的向量而不是单个值，其余的并没有什么变化，只是从数字运算变成了向量运算。中间值 $n = \mathbf{W}\mathbf{p} + b$

需要注意的是，在单个神经元的条件下，权值矩阵只有一行，因为输入向量 \mathbf{p} 只有一列。上图所示的 $w_{1,R}$ 中的1就指的是第1个神经元(当然在这里还没有涉及到)， R 就代表 p_R 所对应的权值，所以中间值

$$n = w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b.$$

如果输入过多，那么示意图画起来特别复杂，所以书中用一种简单的画法来表示神经元：

R 就是输入空间， $R \times 1$ 就是指 \mathbf{p} 向量有 R 行1列，权值矩阵 \mathbf{W} 的规模是 $1 \times R$ 。需要注意的是可以将偏差 b 看做一个特殊的权值，只不过输入永远为1，这一点后面会用到。在单个神经元的条件下，输出是一个标量，但是如果多个神经元的话，输出就是一个向量了。

神经元层

多个神经元并列就可以成为一个“层”，示意图：

这个时候 \mathbf{W} 就是多行多列向量了， $w_{S,R}$ 的含义是权值矩阵的第 S 行第 R 个，也就是第 S 个神经元的第 R 个权值。输出是一个向量，不再是一个值。

权值矩阵的直观图：

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \dots & w_{1,R} \\ w_{2,1} & w_{2,2} & \dots & w_{2,R} \\ \vdots & \vdots & & \vdots \end{bmatrix}$$





$$\begin{bmatrix} w_{S,1} & w_{S,2} & \cdots & w_{S,R} \end{bmatrix}$$

神经元层的简化图如图：

与多输入单神经元的示意图相比，其实就是W从向量进化成了矩阵，b从标量进化成了向量。

多层神经网络

乍一看略吓人，其实仔细瞅瞅还是换汤不换药的。这里你会发现多了一个右上角的上标，这个上标代表着第几层神经元，例如， $w_{S^1,R}^1$ 就代表第1个权值矩阵的第 S^1 行第R列，含义就是第一层的第S个神经元的第R个权值。这么说好像有点混乱，没关系，等后面遇到了问题再说。

这里有一个“隐层”的概念，网络中除了输出层就是隐层，再上图中，第1、2层是隐层。
书的后面还介绍了循环网络(Recurrent Network),我看完之后发现云里雾里的，所以不打算记录了，还是循序渐进吧。

习题

下面有一道习题(这本书真是良心，课后习题带答案的，虽然只有一部分)

挺简单的，套用多输入单神经元的公式就可以了。

先计算出中间值n:

然后套用激活函数：

$$\text{i. } a = \text{hardlims}(-1.8) = -1$$

$$\text{ii. } a = \text{satlin}(-1.8) = 0$$

总结

输入向量p的规模为 $n \times 1$ ，则权值矩阵W的规模为 $1 \times n$ 。

神经元有m个，则权值矩阵的规模为 $m \times n$,b的规模为 $n \times 1$ 。

有S个神经元层，则有S个权值矩阵。

各路神仙发现谬误请指出，战五渣感激不尽。o(*≧▽≦)ツ





♡ 喜欢 (0)

↪ 分享 (0)

神经网络

« 我有一壶酒，足以慰风尘 —— 新年札记

[NNDesign]第二篇：感知器与学习规则 »



[NNDesign]第二篇：感知器
与学习规则



[NNDesign]第三篇：Hebb
Rule

— [NNDesign]第二篇：感知器与学习规则

— [NNDesign]第三篇：Hebb Rule

声明

资源来自栩栩
或引用本站文章
议。如果有侵犯版权的资
系我，我会在24h内删
资源。

友情链接

龙猪 | i Huhu
静觅 | 静静寻觅生活的美好

联系方式

邮箱：646689983@qq.com

关于我

欢迎志趣相投的小伙伴一起来



Theme By 云落

