

从感知机到神经网络

神经网络的一个重要性质是它可以自动地从数据中学习到合适的权重参数

基本神经网络的架构

三层结构：输入层、隐藏层、输出层。每个层都可以由多个神经元构成。神经元之间进行连接。

激活函数：决定如何来激活输入信号的总和（处理**神经元**最后的输出权重，不一定是输出层）

激活函数**必须是非线性函数**，否则网络的层次没有意义

常用的激活函数（0 - 1 区间）：

- 阶跃函数 -- $u(x - TH)$
- Sigmoid -- $1/(1+\exp(-x))$ -- 优势：平滑
- ReLU -- $x > 0 ? x : 0$

一些关于Numpy的知识

shape与ndim

$A.shape = (3, 2)$ 表示 第一个维度有3个元素 第二个维度有2个元素

$np.ndim(A) = 2$ 表示 有两个维度

即 $len(shape) = ndim$

矩阵乘法

- 星号(*)乘：对应位置相乘 / 扩展相乘
- $np.dot$ 乘：一般意义的矩阵乘法

输出层的设计

神经网络可以用在分类问题和回归问题上，不过需要根据情况改变输出层的激活函数。一般而言，回归问题用恒等函数，分类问题用Softmax函数

Softmax -- $y_k = \exp(a_k) / \sum(\exp(a_i))$ -- 有几个类别，输出层就有几个神经元

Softmax的重要特性 -- 输出是0.0 到1.0之间的实数，函数的输出值的总和是1 -- **可以代表概率**

预处理

正则化 -- 把数据限定到某个范围内的处理（例如像素灰度的0-1正则化：除以255）

Batch（批处理） -- 利用矩阵运算的性质，每次输入一批而不是一个数据去过网络，加快训练速度

Dropout -- 在传播过程中进行随机丢弃