

激活函数的原理和类别

- 作用
 - 用来加入非线性因素的，解决线性模型所不能解决的问题
 - 如何直观理解？

sigmoid激活函数： $S(x) = \frac{1}{1+e^{-x}}$

则 $S(x^3) = \frac{1}{1+e^{-x^3}}$

泰勒展开：

$$e^x = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + o(x^3)$$

$$(w1 * x1 + w1 * x2)^2 = (w1 * w1 * x1 * x1 + w2 * w2 * x2 * x2 + w1 * x1 * x2 * w2)$$

激活函数种类

- Sigmoid
- Tanh
- ReLU
- LeakyReLU
- Maxout

Sigmoid

应用范围：二分类

公式： $f(z) = \frac{1}{1+exp^{-z}}$

缺点：梯度消失,计算量大

求导： $Sigmoid'(x) = \frac{1}{1+exp^{-x}} * \frac{exp^{-x}}{1+exp^{-x}} = Sigmoid(x) * (1 - Sigmoid(x))$

Tanh

公式： $f(z) = tanh(z) = \frac{e^z-e^{-z}}{e^z+e^{-z}}$
 $tanh(x) = 2sigmoid(2x) - 1$

求导： $tanh'(x) = (1 - tanh(x)^2)$

ReLU修正线性单元

公式： $\phi(x) = max(0, x)$
输入信号<0时，输出都是0，>0时，输出等于输入

- 优点：使用ReLU得到SGD的收敛速度会比simoid/tanh快很多
- 缺点：训练的时候很“脆弱”，很容易就“die”了

Softmax

Softmax-用于分类神经网络输出

公式： $\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$