神经网络
作业三

181220076, 周韧哲, zhourz@smail.nju.edu.cn

2020 年 11 月 24 日

Problem 1

如下函数适合作为神经网络的激活函数吗? 请说明理由。

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Solution. .

不适合，因为该函数在 $x = 0$ 处不可微，在其他定义域的梯度永远为 0，不适合反向传播来调整误差。

Problem 2

有如下图1神经网络, 隐藏层使用 relu 激活函数, 输出层使用 sigmoid 激活函数且没有偏置。对于输入 $X_i = \{i_1, i_2\}$，其对应输出记为 Y_i ，该数据的真实标签记为 Y_i，这里 $i \in \{1, 2, ..., n\}$，损失函数定义为

$$E = \sum_{i=1,\cdots,n} (Y_i' - Y_i)^2$$

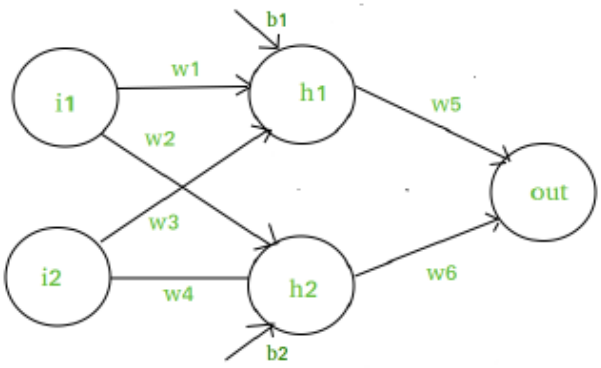

图 1: 题 2

请推导损失函数对于 w_1, b_2, w_5 的偏导。当 $w_1, w_2, w_3, w_4, w_5, w_6$ 分别取值 $0.4, 0.5, 0.2, 0.4, 3.5, 0.6$ 且 $b1 = 0.3, b2 = 0.8$，且只有一条输入数据 $X_1 = \{0.3, 2.8\}, Y_1 = 5.6$ 的时候，计算该损失函数对于 w_3 的具体值，结果保留两位小数。

Solution. .

令 $f_1(x) = Relu(x), f_2(x) = Sigmoid(x)$，且有 $f_2'(x) = f_2(x)(1 - f_2(x))$，再令

$$g(x) = f_1'(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

令 $E_i = (Y_i' - Y_i)^2$，将网络前向过程写成：

$$\begin{aligned} h_1 &= f_1(v_1) = f_1(w_1 i_1 + w_3 i_2 + b_1) \\ h_2 &= f_1(v_2) = f_1(w_2 i_1 + w_4 i_2 + b_2) \\ Y_i' &= f_2(v_3) = f_2(w_5 h_1 + w_6 h_2) \end{aligned}$$

则

$$\begin{aligned} \frac{\partial E_i}{\partial w_5} &= \frac{\partial E_i}{\partial Y_i'} \frac{\partial Y_i'}{\partial v_3} \frac{\partial v_3}{\partial w_5} = 2(Y_i' - Y_i) Y_i'(1 - Y_i') h_1 \\ \frac{\partial E_i}{\partial b_2} &= \frac{\partial E_i}{\partial Y_i'} \frac{\partial Y_i'}{\partial v_3} \frac{\partial v_3}{\partial h_2} \frac{\partial h_2}{\partial v_2} \frac{\partial v_2}{\partial b_2} = 2(Y_i' - Y_i) Y_i'(1 - Y_i') w_6 g(v_2) \\ \frac{\partial E_i}{\partial w_1} &= \frac{\partial E_i}{\partial Y_i'} \frac{\partial Y_i'}{\partial v_3} \frac{\partial v_3}{\partial h_1} \frac{\partial h_1}{\partial v_1} \frac{\partial v_1}{\partial w_1} = 2(Y_i' - Y_i) Y_i'(1 - Y_i') w_5 g(v_1) i_1 \end{aligned}$$

从而，

$$\frac{\partial E}{\partial w_5} = \sum_{i=1}^{n} \frac{\partial E_i}{\partial w_5}, \quad \frac{\partial E}{\partial b_2} = \sum_{i=1}^{n} \frac{\partial E_i}{\partial b_2}, \quad \frac{\partial E}{\partial w_1} = \sum_{i=1}^{n} \frac{\partial E_i}{\partial w_1}$$

当只有一条数据时，易知

$$\frac{\partial E}{\partial w_3} = \frac{\partial E}{\partial Y_i'} \frac{\partial Y_i'}{\partial v_3} \frac{\partial v_3}{\partial h_1} \frac{\partial h_1}{\partial v_1} \frac{\partial v_1}{\partial w_3} = 2(Y_i' - Y_i) Y_i'(1 - Y_i') w_5 g(v_1) i_2$$

计算得 $v_1 = h_1 = 0.98, v_2 = h_2 = 2.07, v_3 = 4.672, Y_1' = 0.991$，从而得到

$$\frac{\partial E}{\partial w_3} = 2 \times (0.991 - 5.6) \times 0.991 \times (1 - 0.991) \times 3.5 \times 1 \times 2.8 = -0.81$$

Problem3

能否用一个神经元拟合二次曲线? 如果能, 请给出实例。如果不能, 请说明至少需要多少个神经元才能拟合二次曲线。

Solution. .

单个神经元无法拟合二次曲线，至少需要 2 个神经元才能拟合二次曲线，一个隐藏层神经元，一个输出层神经元。只要神经元个数足够多，理论上单隐层的神经网络具有拟合任何函数的能力。