

幽默文本检测

19335301 庄鹏标

2021 年 11 月 16 日

目录

1	概述	1
2	实验原理	1
2.1	Feedback Network 和 Back Propagation	1
2.1.1	前向神经网络	1
2.1.2	后向传播过程	2
3	伪代码/实验过程	3
3.1	Feedback Network 和 Back Propagation	3
4	实验结果与分析	3
4.1	分类	3
4.2	回归预测	3
5	总结	3
6	参考资料	3
7	分工	3

1 概述

2 实验原理

2.1 Feedback Network 和 Back Propagation

2.1.1 前向神经网络

我们知道,将输入值的集合唯一映射为输出集合的关系,叫做函数: $X \xrightarrow{f} Y$, 其中 X 可以是一段文字, Y 可以是这段文字对应的幽默程度估计, 这种估计关系就是一种函数。但这种关系非常难用准确的数学函数式子表示, 所以我们希望:

- $f(x)$ 本身是一个黑箱子, 它可以自行调整自己的内部结构
- 它可以处理不限于线性的函数映射, 对于非线性的映射它也可以有一定的处理能力

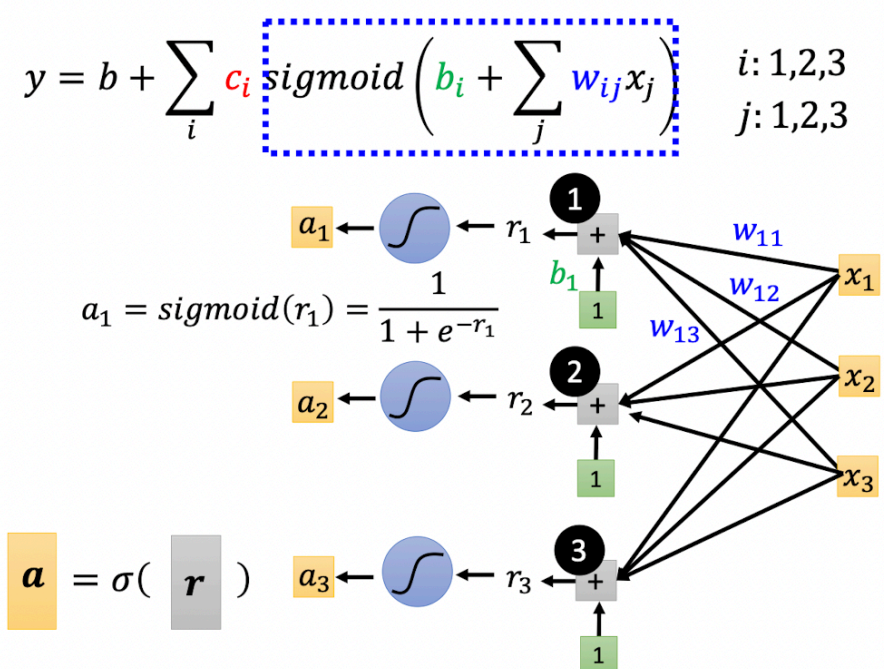
对于上面的问题, 我们可以分别用下面的方法解决:

- 自行调整: 我们可以借鉴逻辑回归的经验, 借助梯度下降这一方法来实现
- 线性预测: 我们可以简单地用 $Y = Wx + b$ 来很表达线性的函数结构, 其中的 W, b 就是需要用梯度下降让模型自行学习的参数
- 非线性: 引入非线性的函数, 与上一步中的 Y 串联输出, 自此, 我们就完成了从 $X \rightarrow Y$ 的简单非线性映射
- 双层叠加: 单次的线性预测 + 非线性输出并不足以表现非线性 (因为最终结果和中间的线性预测结果仍然是唯一相关的), 因此, 我们采用一种简单的方式: 将该结果, 重复上面两个过程, 这样就可以将第一次的非线性信息利用起来, 让我们的黑箱子打破非线性的结构, 正式拥有非线性的预测能力

形式化来讲: 我们的黑箱子中的线性预测单元称为: 神经元; 神经元输出通过的非线性函数称为激活函数 (这一过程称为激活); 我们把通过一次线性预测 + 一次激活称为一层中间层/隐藏层。

但不难想象的是，随着从 $X \rightarrow Y$ 的实际映射复杂度上升，有限的双层结构不能充分表达实际映射，此时我们有两种做法以增加复杂度：（1）将层内神经元的规模变大：但神经元的运算是线性的，这样并不足以弥补非线性部分。（2）增加层数：模仿双层叠加，将更多层串联在一起，这样便能不断增加黑箱子的非线性能力。

至此，我们得以初步窥视神经网络的基础结构：以神经元预测和激活过程为重复单元：给出图示如下：



2.1.2 后向传播过程

后向传播过程关注的是利用神经网络的结果进行梯度推导，我们用到的核心公式为链式法则：

$$\frac{\partial Z}{\partial X} = \frac{\partial Y}{\partial X} \frac{\partial Z}{\partial Y} \quad (1)$$

3 伪代码/实验过程

3.1 Feedback Network 和 Back Propagation

4 实验结果与分析

4.1 分类

4.2 回归预测

5 总结

6 参考资料

7 分工