

8.5.1 神经网络的理论

2023 年 9 月 18 日

神经网络的理论

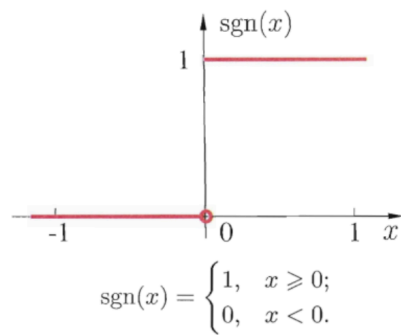
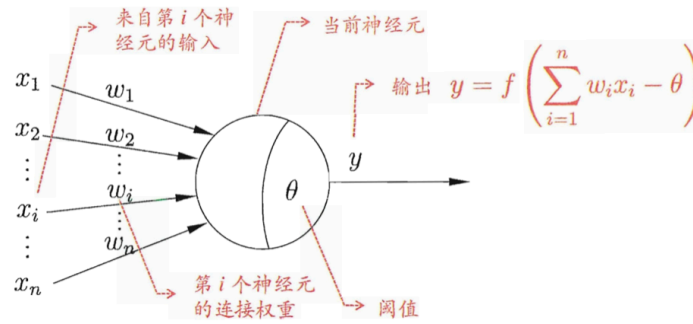
“神经网络”是具有适应性的简单单元组成的广泛并行互连的网络，它的组织能够模拟生物神经系统对真实世界物体所做出的交互反应。——Kohonen(1988)

神经网络中最基本的成分是神经元模型，每个神经元与其他神经元相连，当“它”兴奋时，就会向相连的神经元发送化学物质，从而改变这些神经元的电位；如果某些神经元的电位超过一个“阈值”，那么它就被激活，向其他神经元发送化学物质。

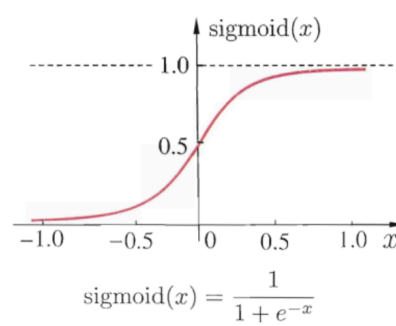


1. 单个神经元

神经网络又称为多层感知机 (multi-layer perceptron)，训练得到一个映射函数 $f(\cdot) : R^m \rightarrow R^o$ ，其中 m 是输入自变量的维度， o 是输出因变量的维度。该近似函数可以是非线性的，既可以用于回归也可以用于分类。与逻辑回归不同，在输入层和输出层之间可以有多个非线性的隐藏层，如下图所示为包含一个隐藏层的神经网络。



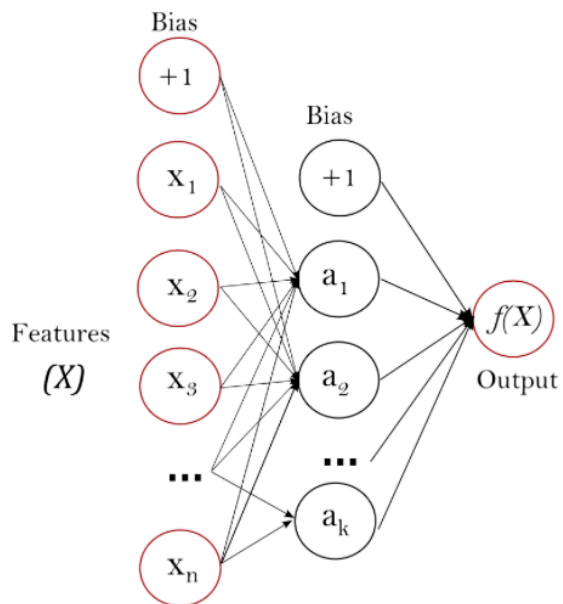
(a) 阶跃函数



(b) Sigmoid 函数

2. 多层神经元

最左侧的层称为输入层，由一系列表示输入特征的神经元 $\{x_i | x_1, x_2, \dots, x_m\}$ 组成。隐藏层中的每个神经元都计算上一层神经元输出的线性组合 $w_1 x_1 + w_2 x_2 + \dots + w_m x_m$ ，这里的 $\{w_i | w_1, w_2, \dots, w_m\}$ 为相邻层神经元的连接权重，并紧接着一个非线性激活函数 $g(\cdot) : R \rightarrow R$ ，如反正切函数。同样的，输出层接收最后一个隐藏层的输出。



神经网络模型的优点包括：

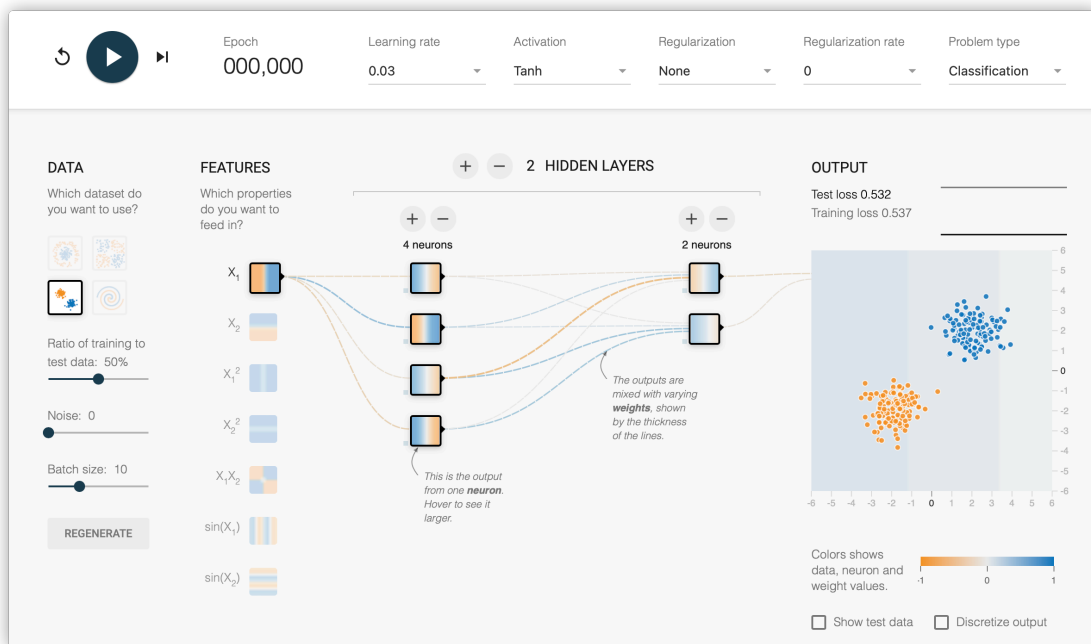
- 能够学习非线性模式；
- 能够实时在线学习，即基于部分新的实时增量数据，在原模型基础上增量学习，而不需要基于全量数据重新学习；

神经网络模型的缺点包括：

- 损失函数是非凸的，存在多个局部最优点，因此不同的初始化会产生不同的训练结果；
- 对超参数调优的要求较高，如隐藏层等；
- 对输入数据的尺度非常敏感，需要做标准化。

3. 神经网络可视化

神经网络可视化演示 [备用链接](#)



扩展阅读

1. Explaining neural networks in raw Python: <https://bronwojtek.github.io/neuralnets-in-raw-python/docs/index.html>