



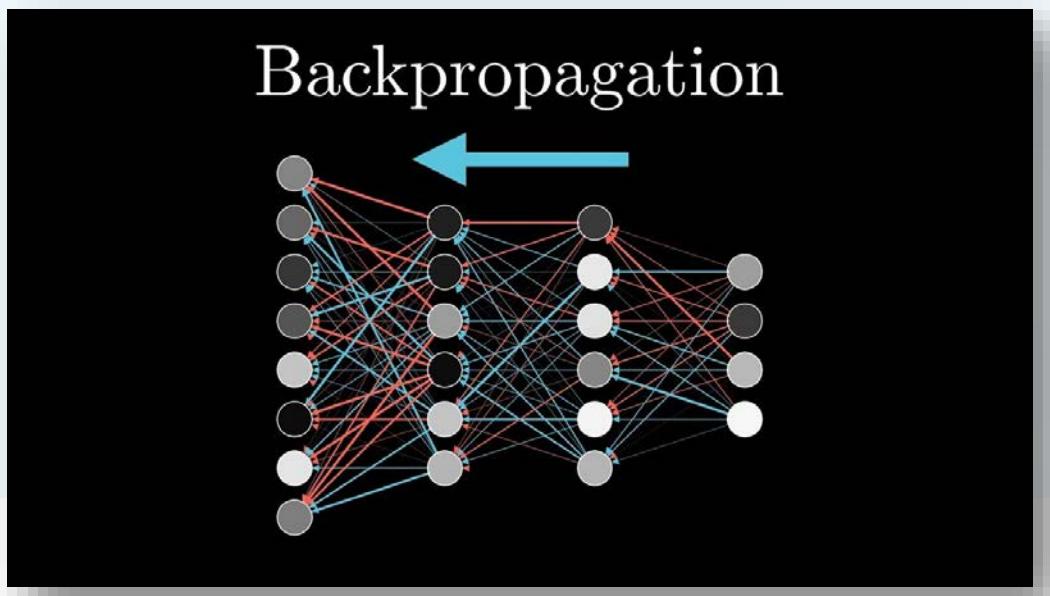
人工智能与信息社会

基于神经网络的智能系统I：反向传播

陈斌 北京大学 gischen@pku.edu.cn

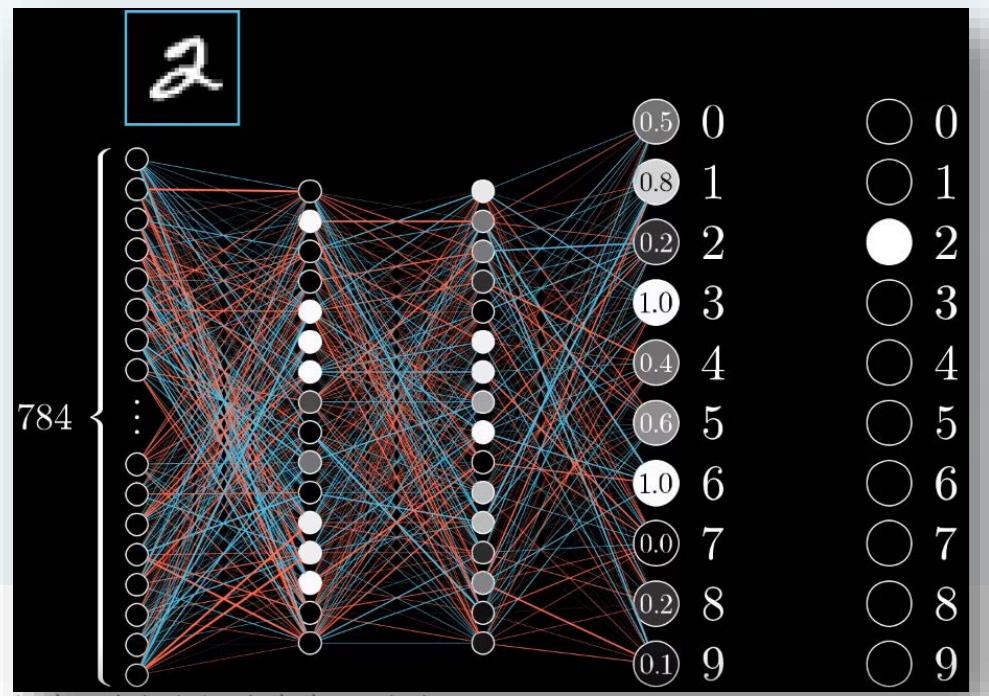
反向传播

- 通过多层神经网络得出的结果与实际标签的结果作比较，得到了损失函数。
- 根据梯度下降的方法，将当前的损失函数反馈给之前各层的神经网络，并调整各层网络参数的权值，这个过程称为反向传播。



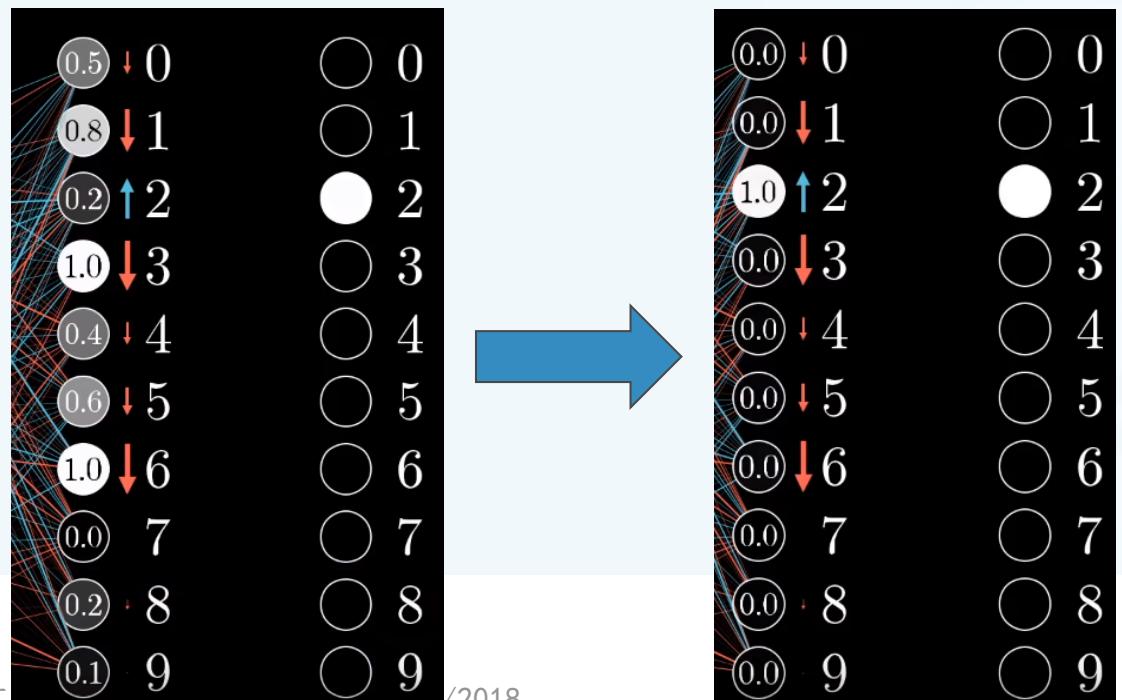
结果比较

- 对于数字2来说，未训练好的网络可能得到如下的一个结果。
- 我们不能直接调整输出的结果值，我们能改的只有网络中的各个参数值。



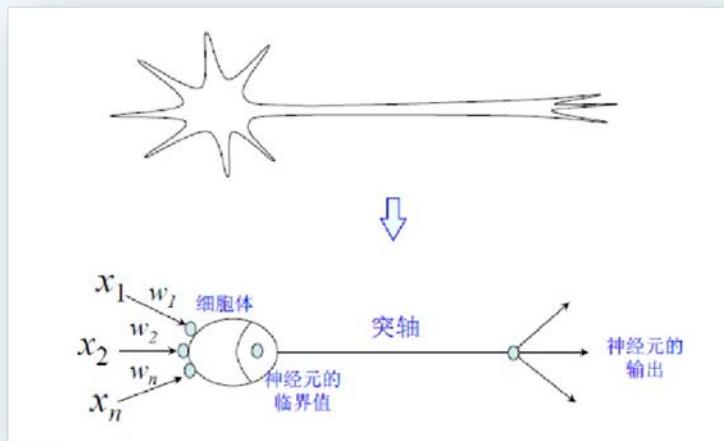
结果比较

- 对于这个结果，我们希望代表2的神经元的值提高，而把其他的神经元值降低。
- 提高和降低的程度和它现有的值成比例，提高2的值就比降低8的值来的重要。



神经元的参数

- 一个神经元的输入输出可以类比为一个多元一次的线性方程。
- 二元一次： $y = w_1x_1 + w_2x_2 + b$
- 可以通过调整w、x、b的值来调整y的值。

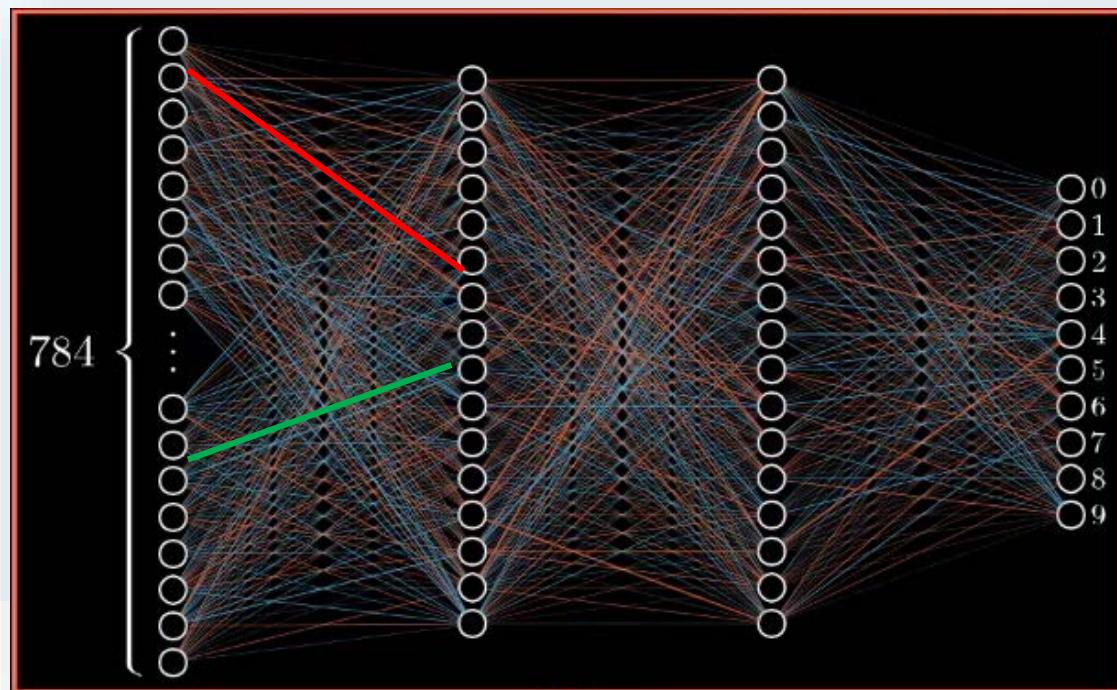


调整参数

- 根据梯度大小不同，调整不同的参数对损失函数影响不同。

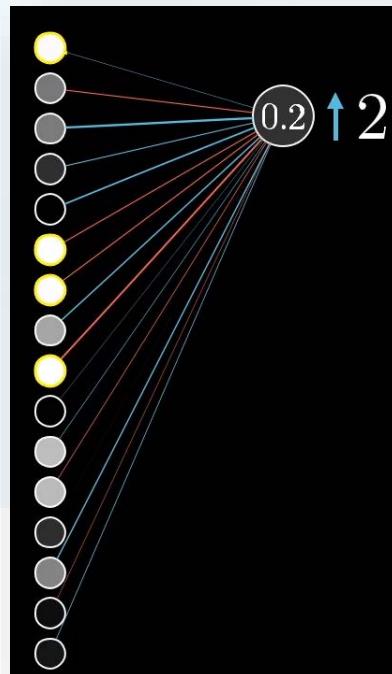
梯度值：3.6 调整它的影响大

梯度值：0.1 调整它的影响小



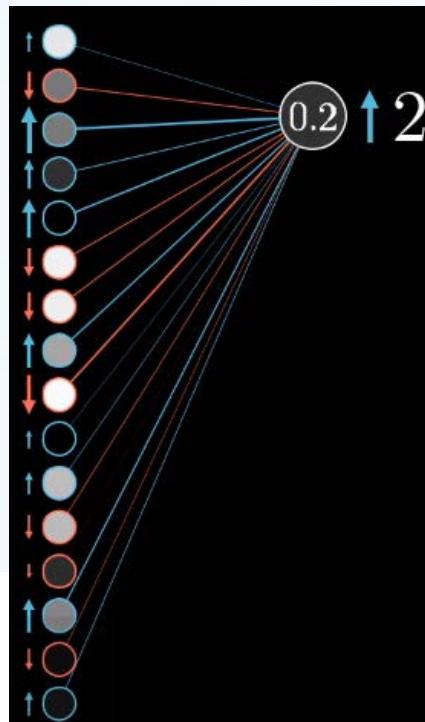
调整系数w

- › 调整输出 y 的一个方法是调整系数 w 。
- › 前一层的神经元激活值越大，表示越有可能“看到”某些与 2 相关的特征。
- › 调整激活值大的神经元的参数，影响比调整激活值小的参数要大。



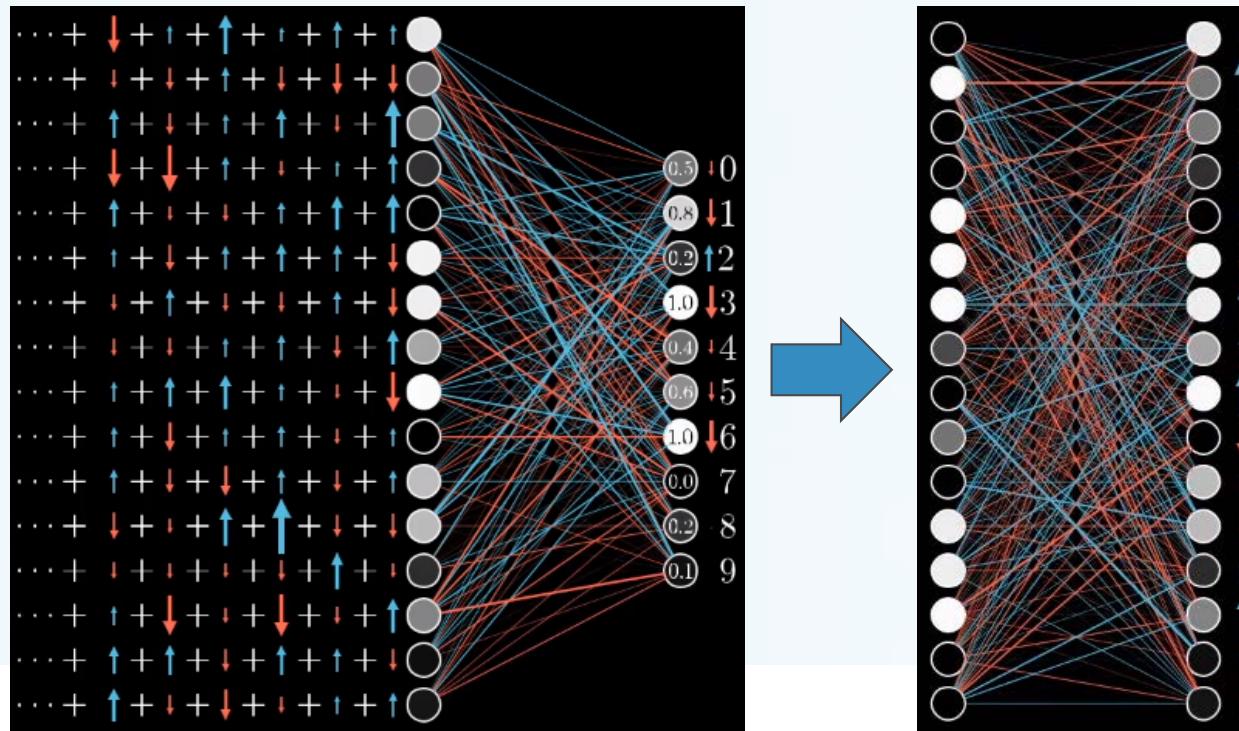
调整输入x

- › 调整输出y的另一个方法是调整输入x。
- › 增加原始参数w是正值的神经元的激活值，
减少原始参数w是负值的神经元的激活值，
增加和减少的程度与参数w大小相关。



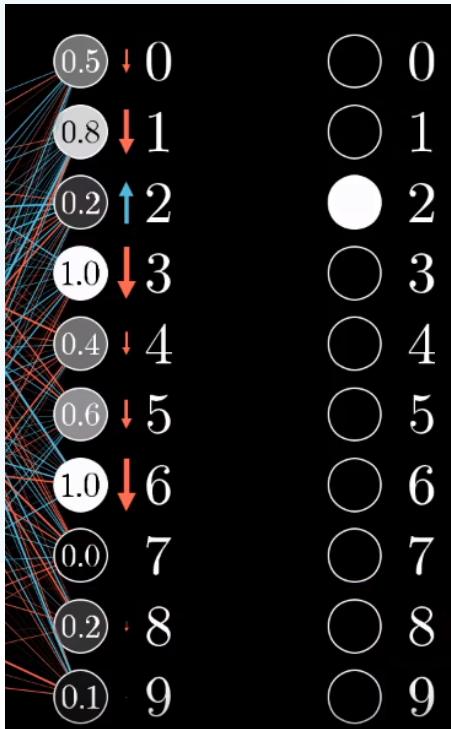
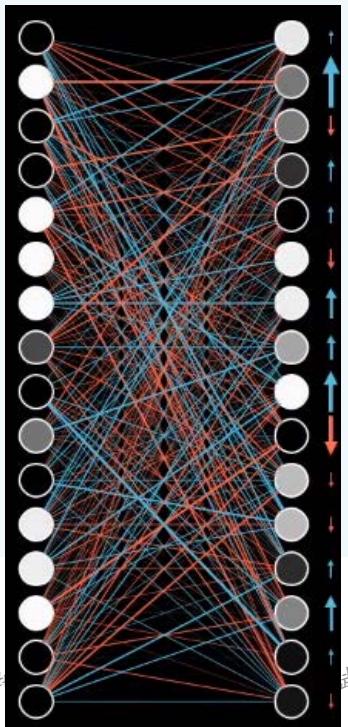
调整输入x

› 但我们无法直接改变每个神经元的激活值。
将每个输出需要改变的大小叠加之后，能够
得到总的改变量。



反向传播

- 通过刚才的过程，我们实现了从最后一层到倒数第二层的传播。
- 这样不断层层传播回去，最后我们就能知道该如何修改整个网络的参数。



训练结果

从随机参数开始，经过梯度下降法和反向传播训练，手写数字识别准确率从不到10%提高到了93%

