出现NAN的原因

NaN表示not a number. 总结可能出现的原因如下：

1. 梯度爆炸：观察log，注意每一轮迭代后的loss。loss随着每轮迭代越来越大，最终超过了浮点型表示的范围，就变成了NaN。
2. 不当的损失函数：有时候损失层中loss的计算可能导致NaN的出现，观察到训练产生的log时一开始并不能看到异常，loss也在逐步的降低，但是突然之间NaN就出现了。比如交叉熵loss = ，其中y’是模型输出的结果，y是真实值，所以，一般y取值1或0，y’取值为[0,1]，所以当y’足够小时,logy’就可能是NaN。
3. 错误的输入：输入中就包含有NaN，每当学习的过程中碰到这个错误的输入，就会变成NaN。观察log的时候也许不能察觉任何异常，loss也逐步降低，但是突然就变成了NaN了。