1、大规模的神经网络有两个缺点：

* 费时
* 容易过拟合

这两个缺点真是抱在深度学习大腿上的两个大包袱，一左一右，相得益彰，额不，臭气相投。过拟合是很多机器学习的通病，过拟合了，得到的模型基本就废了。而为了解决过拟合问题，一般会采用ensemble方法，即训练多个模型做组合，此时，费时就成为一个大问题，不仅训练起来费时，测试起来多个模型也很费时。总之，几乎形成了一个死锁。

2、dropout是指在深度学习网络的训练过程中，对于神经网络单元，按照一定的概率将其暂时从网络中丢弃。注意是暂时，对于随机梯度下降来说，由于是随机丢弃，故而每一个mini-batch都在训练不同的网络。

3、Dropout的出现很好的可以解决这个问题，每次做完dropout，相当于从原始的网络中找到一个更瘦的网络，如下图所示：

