

《The lottery ticket hypothesis: Finding Sparse, Trainable Neural Networks》
ICLR2019 Best Paper

很难从头开始训练

网络剪枝

重要的不是子网络结构，而是网络权值的初始化

通过剪枝发现的结构很难从头开始训练

剪枝剪的越厉害，网络权值越稀疏，学习就会越慢，最后的acc_rate也会越低

一般是用一个大参数的模型去解决一个问题，因为我们很难知道在网络问题中的网络复杂度

commensurate training time

commensurate accuracy

fewer parameters

网络往往是过参数化的

- 1.目前取得的state of the art性能的模型，大多都是过参数化的，当然，很多方式可以进行抑制，但是对于过参数化的模型，其实剪枝也不乏是一个很有效且合理的选择。
- 2.其实很多地方可以发散和讨论，论文本身除了抛出一些观点，并且用实验结果进行了一些作证，其实并没有抛出更多的值得深思的方向和话题。
- 3.大网络中的小网络、SGD会自动寻找到一个合理初始化的小网络，密集网络比稀疏网络更容易优化（更多的小网络）

《Understanding the difficulty of training deep feedforward neural networks》

观点：很深的神经网络很难训练的地方在于层数越深，权值的方差在以1/3的速率衰减。

Bengio之前指出的初始化方法【 $-1/\sqrt{n}$, $1/\sqrt{n}$ 】，就会出现权值的衰减。
学生打脸老师。

所以用【 $-\sqrt{6/(n_i+n_{i+1})}$, $\sqrt{6/(n_i+n_{i+1})}$ 】

《Neural ...》

一篇关于NAS的文章，emmm名字我已经记不清了

NAS的工作现在有用RL的，有用RNN的。类似于拼乐高，搭积木。

800个GPU训练了25天，感觉文章的实际用途是在炫富。就这搜索效率，有个毛用啊。

想到之前公司嘉泰写的信号搜索器，不知道和这个会不会有异曲同工之妙。

不过据说那个信号搜索器还是挺有用的，就是有点难用。

VALSE Web Seminar有个关于借鉴类脑作自动导航的Pre。斯白露。

NAS: Neural Architecture Search

AutoML

有空可以看看O'Keefe教授在Nature上的工作。栅格细胞。

还是一个关键问题。

SNN的网络连续性不强，更为的离散，很难利用梯度进行训练，目前而言看来更多的还是类梯度。所以训练起来会有很大的难度。

之前想到用cnn训练的权值直接赋给snn的思路不知道如何了，有空可以看看唐老师的lab发的文章。如何变离散为连续，是个永远值得深思的问题啊！

感觉需要补一补数学知识了。

最近可能要着手把数学分析粗略的学习一下，顺带要看看随机过程。

接下来可以查查文献看看关于网络剪枝的内容。

希望实验室配的pc快点到位吧。TMD用MBP真的好别扭啊。有点怀恋之前公司的办公环境了。至少电脑配置32个g内存，那时真的牛逼。

就先写这么多吧。