# 黑盒仿制算法算法流程及细节

## 基于ENAS的结构搜索

ENAS目的是在给定大小的图中，搜索出性能最佳的子图。

ENAS方法需要给定控制向量的形式以及节点数量，二者共同表示了EANS的搜索空间。

### DAG和子模型

构建一个拥有个节点的有向无环图(DAG)，每个节点表示网络中一个隐含层。该DAG是所有子模型(SubModel)的叠加(Superposition)，参数集为。

节点之间的连接方式和节点的类型通过个控制向量组成的控制矩阵表示。，其中，表示有数据从流向；表示的网络类型等超参数。目前还没有想好的具体形式，暂时将其忽略，令。网络层的输入，其中表示网络层的输出。

节点的输入只有来自输入层。所有叶子节点输出的求和作为输出层的输入。

根据矩阵可以从DAG生成子模型(SubModel) ，该子模型参数集为。不同子模型的参数集会重合，所以权重能在各个子模型之间进行共享。

### 控制器

控制器(Controller)负责生成控制矩阵。即控制子模型结构。我们希望Controller所生成的结构能够使得子模型在验证集上的损失最小。控制器的参数集为。

控制器(Controller)生成控制矩阵的过程如下：

|  |
| --- |
| 将向量的所有可能的取值采用one-hot编码，编码为动作向量。  For i =1 to N:  // 表示动作向量的概率分布  // 分类任务取概率最大的动作，这里是按概率抽样得到动作  //计算交叉熵损失  // 对动作向量进行embedding  // 根据编码表查找控制向量的值 |

### 训练过程

模型使用的训练数据,表示文本，表示类别。

对于每个MiniBatch执行如下两步训练过程

1. 更新DAG的权重集。（使用蒙特卡洛方法计算梯度）执行M次控制器，生成M个子模型。使用这些子模型对训练集进行预测并计算梯度。使用子模型的平均梯度用于更新DAG的权重集。
2. （已经使用步骤1更新过ω了）更新控制器的权重集。执行M次控制器，生成M个子模型。使用这些子模型对验证集进行预测并计算验证集交叉熵损失，计算回报，为常量。使用作为损失计算的梯度并进行更新。

解释：控制器生成的动作的概率分布，但是动作是抽样确定的，所以控制器可能生成概率小的动作。如果使用更新则会使LSTM生成的概率分布接近抽样结果。Reward这里看作一个权重，如果Reward高则说明概率分布应该接近抽样结果。

通过循环交替执行以上过程即可同时训练权重和结构。最后，我们使用完成训练的控制器，生成若干个权重互相独立的子模型，并执行一次完整的训练。对这几个模型进行评估选取性能最好的。

## 基于遗传算法的模型扩充

### 使用遗传算法确定子搜索空间

遗传算法目的是，在一个用户给定的完整搜索空间中，给ENAS算法提供子搜索空间，即控制向量的形式以及节点数量。

|  |
| --- |
| 根据完整搜索空间进行确定染色体的编码表示（具体方式还没确定）  随机生成若干个个体组成种群，每个个体就是一个子搜索空间  执行以下循环直到满足终止条件:  使用ENAS算法在子搜索空间进行搜索和评估  只保留验证集上性能最高的前若干个子搜索空间  种群中的子搜索空间进行交叉变异 |

### 使用权重继承层级扩充网络

ENAS算法得到的单个训练好的子模型的模型容量可能不足，这时候可以使用上述遗传算法生成新的搜索空间，再使用ENAS算法在先前训练好的子模型后拼接网络并进行结构搜索。示意图如下：

