先剪枝（之前提到过，运用得较少）

后剪枝：加入正则化项α|Ti|,α为剪枝强度。

预剪枝：设定阈值，缺点是会提前终止。

后剪枝

1 错误率降低剪枝：从下至上，遍历所有的可能的子树，直到在测试集上没有提升时，停止

2 悲观剪枝：上到下会提前终止。比较 被替换子树的错误数-标准差（由二项分布计算） 和 新叶子节点错误数

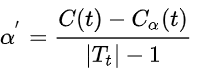
3.代价复杂度剪枝 （CCP， Cost-Complexity Pruning)

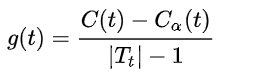
决策树的损失函数：

Cα=C（T）+α|T|

将α视为变量，当极小时，最初的决策树就是最优解，当其极大时，只能使用最简单的决策树，也就是根节点作为最优解。所以，当α固定时，可以找到一个最优的决策树结构。

令剪枝前后的损失相等，可以求得



计算：

找到最小的g(t)，剪去。