寒假第二次读书会周报——郭昊南

**一、进度**

**本周进度：**《统计学习方法》第八章，负责8.4.2节分享

**下周进度：**《统计学习方法》第九章

**二、分享内容**

1、算法步骤（前向分步）：

1）初始化f0

2）计算下一棵提升树



需使经验风险最小化



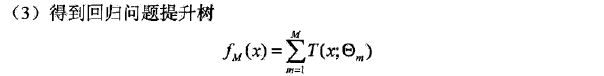
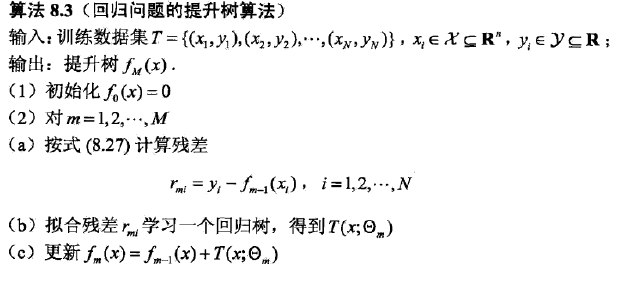
2、损失函数的选择

分类问题——指数损失函数

回归问题——平方误差损失函数

一般决策问题——一般损失函数

3、回归问题提升树



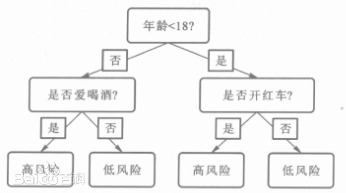
除第一棵树以外，后面的树拟合的是前面树的残差，即不断修正之前的树

个人理解：

回归树——根据x的值确定y的值

分类树——根据x的值确定该落入哪一类

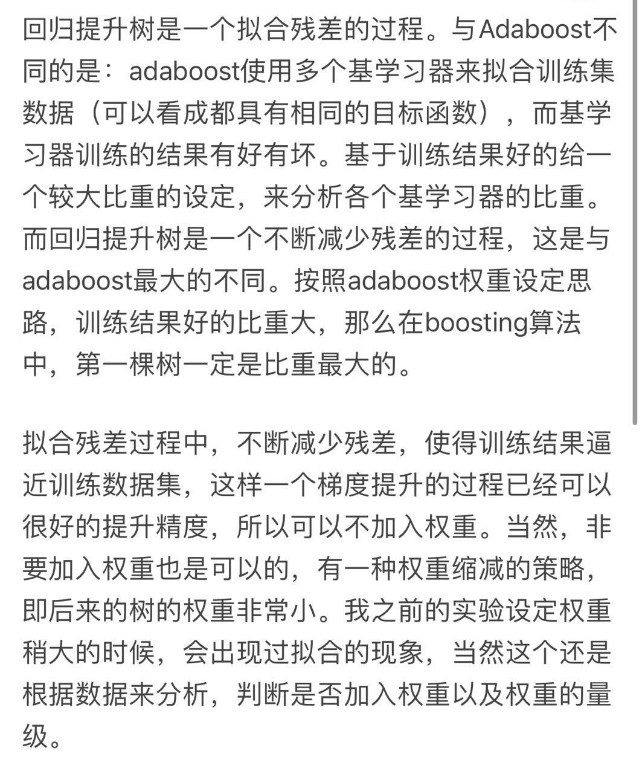
分类提升树书中并没写具体算法，个人理解其实对于一颗完整的分类树（如下图这种），每一层都可以单独看作一个提升树（树桩型），从顶层向下不断提升，可以看作前向分步提升的过程。



**三、讨论问题：**

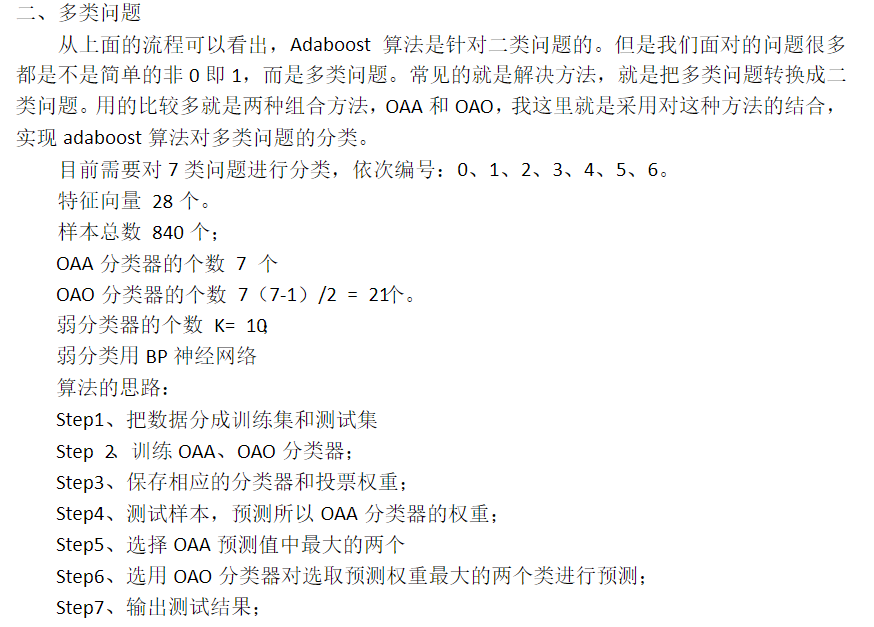
**问题1：为什么提升树没有权值？**

**解释：**



**问题2：AdaBoost如何拓展至多分类问题？**

**解释：**常规的AdaBoost只适合解决二分类问题，因为计算结果正负两种情况分别对应两类，但是也可以通过将多分类问题转化为二分类问题来实现多分类的AdaBoost。



<https://blog.csdn.net/on2way/article/details/48006539>

**问题3：. 提升方法中分类器过多是否会出现过拟合状况？有正则化方法吗？**

**解释：** 由于提升方法中是通过前向分步逐步提升准确率的，如果分步过多，最后会训练出完全匹配训练数据的模型，所以是可能会出现过拟合现象的。解决方法是适当调整分步步数，达到规定的精度后就可以停止。

