一、阅读范围

《统计学习方法》第八章提升方法

二、负责讲解部分主要知识梳理

前面章节中分别介绍了AdaBoost算法以及训练误差分析。8.3节介绍的是对AdaBoost算法的另一种解释，即为：模型是加法模型、损失函数为指数函数、学习算法为前向分布算法时的二类分类学习方法。

8.3.1节主要介绍了前向分布算法。

考虑加法模型

其中，为基函数，为基函数的系数。显然，AdaBoost算法构建的基本分类器的线性组合式（8.6）是一个加法模型。

在给定训练数据及损失函数的条件下，学习加法模型成为经验风险最小化即损失函数最小化问题：

前向分布算法求解这一优化问题的想法是：因为学习的是加法模型，如果能够从前向后，每一步只学习一个基函数及其系数，逐步逼近优化目标函数式，那么就可以简化优化的复杂度。具体地，每步只需优化如下损失函数：

接下来书本给出了详细的算法 算法8.2（前向分布算法）

三、问题

1. 为什么提升树模型中没有权值？

答：这里的提升树算法其实已经不同于我们在最初章节里面学习到的AdaBoost算法，所以其实8.3节在介绍AdaBoost算法的解释的时候实际上是在对AdaBoost算法进行解构，为提升树算法做铺垫。其实AdaBoost算法的内核就是前向分布算法，本质上是一个迭代学习的过程，是在迭代优化中降低误分类可能。因此残差的引入其实已经能够解决这个问题，所以可以不需要权重的加入。

2. 在梯度提升构建回归树的算法中，(2)b步骤不是已经获得了cmj吗？为什么不采用这个而是保留了叶结点区域但重新用另一种方式获得cmj？

答：（2）b步应该仅得到了第m棵树的叶结点区域

（2）c步根据所得的计算得出相应的

3. 在梯度提升构建回归树的算法中，为什么书上M个子回归树的区域数都一样？

答：

答：个人感觉，应该是不一样的，所以J也不是个定值。

4. AdaBoost如何拓展至多分类问题？

答：初始的AdaBoost算法应用范围应该是二类分类。但后续的提升树算法可以适用于多分类情况，只是这里书上没有具体展开，只是解决了回归问题。

5. 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？

答：可能出现过拟合现象。为防止过拟合现象，在每一次迭代基函数时候引入变量与基函数相乘来控制迭代次数。通常用步长和迭代最大次数一起来决定算法的拟合效果。

6. 梯度提升算法第一步初始化为什么要估计损失函数极小化的常数值？

答：这里的残差变了，梯度提升算法中是用负梯度拟合残差，如果初始化值仍为0，估计第一步的梯度较之意义不明显。

四、下周预计读书进度

第九章 EM算法及其推广