王鹏-200301

总体理解：

Adaboost算法，我认为是算法的算法，其全称是Adaptive Boost，也就是算法具有自适应性，它适应什么呢，适应的就是弱分类器各自的训练误差率。而这次学的提升树，就是把基本分类器换成了决策树，而Freidman提出的梯度提升方法将残差近似成梯度，降低了一般损失函数优化的复杂性。

1. 我的问题
2. 提升树模型第1次和第n次的损失函数有什么区别？(n>1)

第一次的损失函数是样本点对应y值的平均误差，而第n次是样本残差的平均误差，残差就是，是m-1个基本分类器求和。

1. 梯度提升算法改进了什么地方，有什么值得学习的应用吗？

不直接计算残差，而是通过对每个基本分类器求偏导，作为残差近似值，拟合回归树。

有一个方法GBDT，是现在各个比赛很常用的分类模型，值得学习，同时包括随机森林算法。

1. 梯度提升法对比提升树，为什么更work一点？

根据算法8.4底部的叙述，对于平方损失函数，负梯度就是残差；而对于一般损失函数，就可以视为残差的估计值。那么我们就可以认为我们降低残差也就是降低方差。这其实也是提升方法神奇的地方，既能降低方差，避免欠拟合；又可以减低方差，避免过拟合。

1. 组员问题
2. 计划安排

本周8.4

下周9