## 2020.4.25 第十三次读书报告

09018330 孙毅远

### 一、自己提出的问题的理解

#### 1. 如何理解强学习是弱学习的充要条件？

理解：理论上如果一个分类可以被多项式算法学习达到弱学习条件，即学习正确率仅比随机猜测好，就一定存在另一个多项式算法可以达到强学习的条件，证明过程应该毕竟复杂。

#### 2. AdaBoost算法的迭代次数如何判断

理解：达到精度要求或是时间要求即可停止，弱分类器数目较多的情况下可能时间会很长

### 二、别人提出的问题的理解

#### 1. AdaBoost算法中的 如果是负值如何理解？

理解：在二分类的情况下可以反转使得为正，多分类情况可能存在小于0的情况，满足弱学戏条件即可

#### 2.提升树与AdaBoost的关系

理解：都属于提升方法，但具体的损失函数有区别

### 三、读书计划

#### 本周 学完第八章

#### 下周 学完第九章

### 四、读书笔记

#### 8.1 提升方法 AdaBoost算法

* + 基本思路
    - 一个概念是强可学习的充分必要条件是这个概念是弱可学习的
    - 如何将容易发现的弱学习方法提升为强学习方法
      * 在每一轮如何改变数据的权值或概率分布
      * 提高被前一轮弱分类器错误分类样本的权值，二降低那些被正确分类的权值
      * 如何将弱分类器组合为一个强分类器
      * 采取加权多数表决的方法，具体是加大分类误差率小的弱分类器的权值，使其在表决内起到较大的作用
  + AdaBoost 算法

#### 8.2 AdaBoost算法训练误差分析

* + 定理8.1(AdaBoost的训练误差界)
  + 二类分类问题AdaBoost 的训练误差界

#### 8.3 AdaBoost算法的解释

* + 可以认为AdaBoost算法是模型为加法模型，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
  + 前向分布算法
  + 给定训练数据及损失函数的情况下，问题转化为损失函数极小化问题，前向分布算法()的想法是从前向后，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，就可以简化优化的复杂度
  + AdaBoost算法是前向分布加法算法的特例，这时模型是由基本分类器组成的加法模型，损失函数是指数函数

#### 8.4 提升树

* + 采用加法模型与前向分布算法，以决策树为基函数的提升方法称为提升树
  + 提升树算法


  + 梯度提升