

## § 12 集成学习（Adaboost&决策树）

### 12.1 Adaboost

（1）集成学习 **Boosting** 算法能否实现“三个臭皮匠顶个诸葛亮”的类似功能？试分析原因。同时说明要达到此功能有什么先决条件；（思考）

（2）**AdaBoost** 算法的基本性质是能在学习过程中不断减少训练误差，关于 **AdaBoost** 训练误差上界的证明？（学习 ppt 课件“自习”部分内容）（思考）

（3）给定如表所示训练数据。假设弱分类器由  $x < v$  或  $x > v$  产生，其阈值  $v$  使该分类器在训练集上分类误差率最低。试用 **AdaBoost** 算法学习一个强分类器。（习题）

训练数据集

序号	1	2	3	4	5	6	7	8	9	10
x	0	1	2	3	4	5	6	7	8	9
y	1	1	1	-1	-1	-1	1	1	1	-1

说明及要求：

- （a）采用标准的 **Adaboost** 样本权重更新规则进行样本权重的更新（关键步骤不能省略）
- （b）先写算法流程
- （c）对上述算法过程和结论试通过编程实现进行对比。（有时间作为补充编程练习）

解：