1. （必填）自己提出的问题的理解（罗列全部）：
2. 提出的问题1：我没有理解60页的“剪枝之后的规则集合可能不再是互不相交且完全覆盖的”，因为我感觉示例里面P59 3.7(b)的剪枝后的决策树就是互不相交且完全覆盖的，而且我还觉得在这个图里对其他任意的树节点剪枝，得到的结果还是互不相交且完全覆盖的。

讨论后的理解：剪枝可以是自顶向上或者自底向下的，如果生成了规则，可以剪一开始节点的规则。那么剪开始节点的限制条件可以导致该分支在二维空间上代表的区域扩张，和其他区域重合。在三叉树里面剪去一个树的分支，可以使该树不完全覆盖。



1. （必填）别人提出的问题的理解（选择几个问题罗列，并给出理解）：
2. 问题2：结合西瓜书第二章阅读，西瓜书p34页倒数第三段为什么，前一个标记为（x，y），若当前为真正例，则为（x，y+1/m+），若为假正例，则为（x+1/m-，y）？

自己的理解：这个是可以数学证明的，因为如果是真正例，那么真阳率增加1/m+，假阳率不变。如果是假正例，那么真阳率不变，假阳率增加1/m-

1. 问题3：在ROC曲线中，为什么主对角线（代表随机猜测，即每个类分成正例为一个固定比例）上测试数据的假阳率FPR总是和真阳率TPR相等？

自己的理解：这又是一个数学证明题。首先，在分类为正例为固定比例的情况下TP/FP=FN/TN，那么TP/（TP+FN）=FP/ (TN+FP),也就是真阳率=假阳率，画出来是一条对角线。

4、问题4：P59如何理解后剪枝操作中“如果一个树节点的错误估计要比其子树的错误估计小，则其子树将被剪枝”？错误估计的思路是什么？

自己的理解：在树的分支已经过拟合的情况下，用这棵树和加上子树用测试数据进行测试，会导致原树的正确率反而更高，这就是‘树节点的错误估计要比其子树的错误估计小’，所以这个时候可以确认过拟合已经产生，要剪枝。

1. （必填）读书计划

1、本周完成的内容章节：3.1-3.5

2、下周计划：3.6-3.10