

作业二

(10月22日上课前提交纸质版)

1. 采用信息增益准则，基于表 4.2 中编号为 1、2、3、6、7、9、10、14、15、16、17 的 11 个样本的色泽、根蒂、敲声、文理属性构建决策树。（本次作业可以用笔算，鼓励编程实现，但都需要列出主要步骤，其中
 $\log_2(3)=1.585, \log_2(5)=2.322, \log_2(6)=2.585, \log_2(7)=2.807, \log_2(9)=3.17, \log_2(10)=3.322, \log_2(11)=3.459$ ）
2. 用表 4.2 中编号为 4、5、8、11、12、13 的样本做测试集，对上题的训练数据采用预剪枝策略构建决策树，并汇报验证集精度。
3. 用表 4.2 中编号为 4、5、8、11、12、13 的样本做测试集，对题 1 所构建的决策树进行后剪枝，并汇报验证集精度。