

# 机器学习导论

## 第二次大作业

班级, 学号, 作者姓名

### 1 训练一个决策树

考虑下面的训练集：共计 6 个训练样本，每个训练样本有三个维度的特征属性和标记信息。详细信息如表1所示。

请通过训练集中的数据训练一棵决策树，要求通过“信息增益” (information gain) 为准则来选择划分属性。请参考书中图 4.4，给出详细的计算过程并画出最终的决策树。

表 1: 训练集信息

| 序号 | 特征 A | 特征 B | 特征 C | 标记 |
|----|------|------|------|----|
| 1  | 0    | 1    | 1    | 0  |
| 2  | 1    | 1    | 1    | 0  |
| 3  | 0    | 0    | 0    | 0  |
| 4  | 1    | 1    | 0    | 1  |
| 5  | 0    | 1    | 0    | 1  |
| 6  | 1    | 0    | 1    | 1  |