**决策树课后习题**

1.根据表1中所给的训练数据集，利用信息增益比（C4.5算法）生成决策树。表1是一个由15个样本组成的贷款申请训练数据。数据包括贷款申请人的4个特征（属性）：第1个特征是年龄，有3个可能值：青年，中年，老年；第2个特征是有工作，有2个可能值：是，否；第3个特征是有自己的房子，有2个可能值：是，否；第4个特征是信贷情况，有3个可能值：非常好，好，一般。表的最后一列是类别，是否同意贷款，取2个值：是，否。

表1 贷款申请样本数据表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 年龄 | 有工作 | 有房子 | 信用 | 类别 |
| 0 | 青年 | 否 | 否 | 一般 | 否 |
| 1 | 青年 | 否 | 否 | 好 | 否 |
| 2 | 青年 | 否 | 是 | 好 | 是 |
| 3 | 青年 | 是 | 是 | 一般 | 是 |
| 4 | 青年 | 否 | 否 | 一般 | 否 |
| 5 | 中年 | 否 | 否 | 一般 | 否 |
| 6 | 中年 | 否 | 否 | 好 | 否 |
| 7 | 中年 | 是 | 是 | 好 | 是 |
| 8 | 中年 | 是 | 否 | 非常好 | 是 |
| 9 | 中年 | 是 | 是 | 非常好 | 是 |
| 10 | 老年 | 否 | 是 | 非常好 | 是 |
| 11 | 老年 | 否 | 是 | 好 | 是 |
| 12 | 老年 | 是 | 否 | 好 | 是 |
| 13 | 老年 | 是 | 否 | 非常好 | 是 |
| 14 | 老年 | 否 | 否 | 一般 | 否 |