#### 决策树面试问题汇总

1. 请简要描述决策树的基本思想。

决策树根据启发函数选择并划分特征，并在此步骤进行多步，直到达到某种临界值。

1. 给定如下表格：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Age | 长相 | Salary | Coding | Target |
| 1 | old | handsome | high | no | 0 |
| 2 | young | ordinary | medium | yes | 1 |
| 3 | young | ugly | high | no | 0 |
| 4 | young | ordinary | high | yes | 1 |
| 5 | young | ordinary | low | no | 0 |

1. 请计算特征Salary的信息增益：

答：

1. 请计算Salary的信息增益比：
2. 请计算Gini(D|Age=young)：

以上三种启发函数分别对应ID3, C4.5和CART树。

1. 请简要谈谈剪枝方法。
2. 预剪枝：提前规定最大叶子节点数、最大深度或最小中间节点所含样本数。
3. 后剪枝：就是构建完决策树之后，假如不分裂这部分和之前比怎么样。如果这部分不分裂效果更好或影响不大，就将这部分剪掉。