

遍历所有样本空间检测决策树是否一致

结论：

编码={实数编码，One-hot编码}，选择={基尼系数，信息增益}

1. 不同编码，相同选择，构建的决策树，虽然样子不一样（由于编码不同），但是运行流程结果是一样的。
2. 相同编码，不同选择，实数编码两种选择结果一致，One-hot不一致.
3. 独热码的信息增益构建出的树与其他三类不同,其他三类保持一致.

原因:

在计算信息增益时，属性数量急剧增加会导致信息增益的值更容易受到噪声的影响，因为每个二元属性可能只包含非常少量的样本，因此其信息增益可能会被高估。这会使得决策树选择某些不太重要的属性来进行划分，从而导致决策树的复杂度增加。