决策树实验报告

191250111 裴为东

决策树算法原理：

1.首先将所有的特征看成一个个节点，创建出根节点；

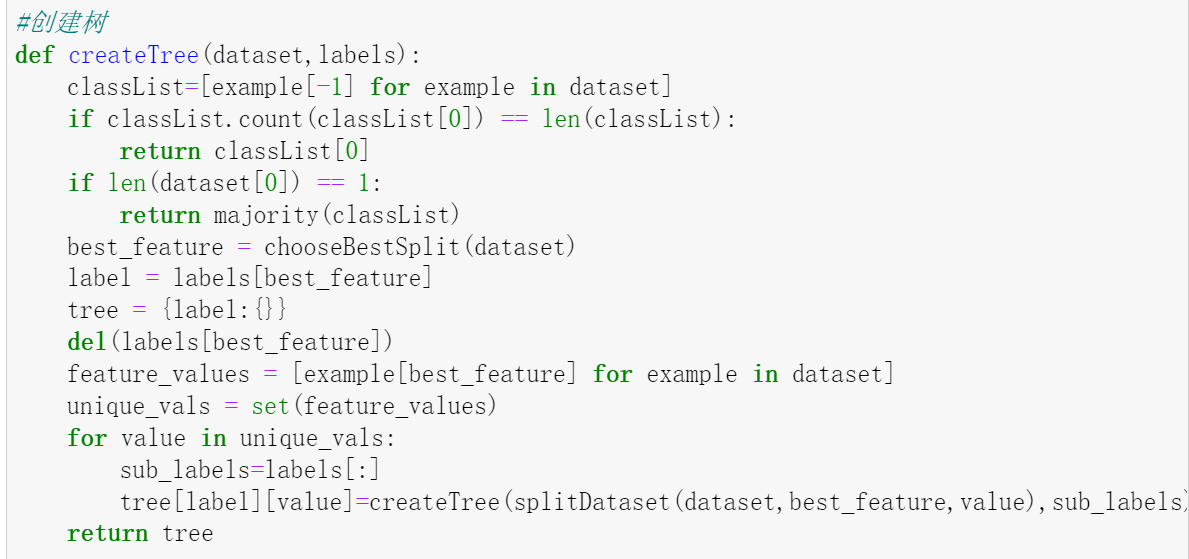
2.然后遍历所有的特征，在每一次到某个特征时遍历当前特征的所有分割方式，找到最好的分割点，讲数据划分为不同的子节点，计算划分后子节点的信息熵

3.在遍历的所以特征中，比较寻找最优的特征以及最优特征的最优划分方式，选择信息增益最高的特征作为节点进行分割操作，产生子树

4．重复2-3步，直到子节点中只有一种类型或为空，或者当前节点中样本数小于某个值，同时迭代次数达到指定值

核心代码：

创建决策树



计算香浓熵



选择最优的划分



实验结果：

