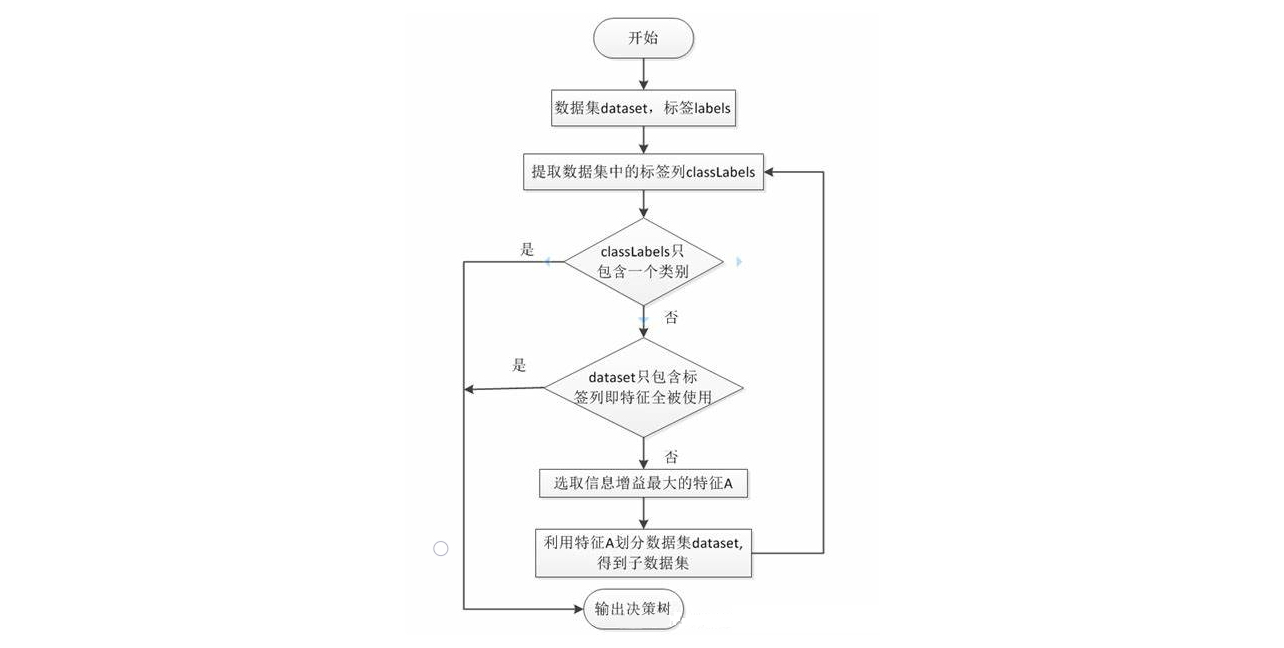
**核心算法ID3算法（基于信息增益的算法）**

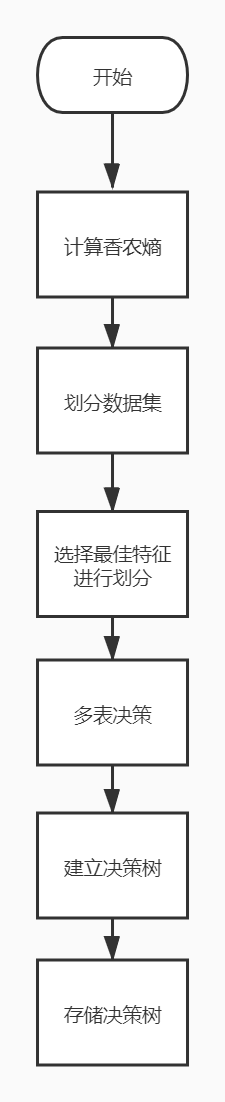
基本思想是：以信息熵为度量，用于决策树节点的属性选择，每次优先选取信息量最多的属性，亦即能使熵值变为最小的属性，以构造一颗熵值下降最快的决策树，到叶子节点处的熵值为0。此时，每个叶子节点对应的实例集中的实例属于同一类。

流程如下：

* 输入：训练数据集D和特征A；
* 输出：特征A对训练数据集D的信息增益g(D,A)
* 1、计算数据集D的经验熵H(D)
* 
* 2、计算特征A对数据集D的经验条件熵H(D|A)
* 
* 3、计算信息增益
* 

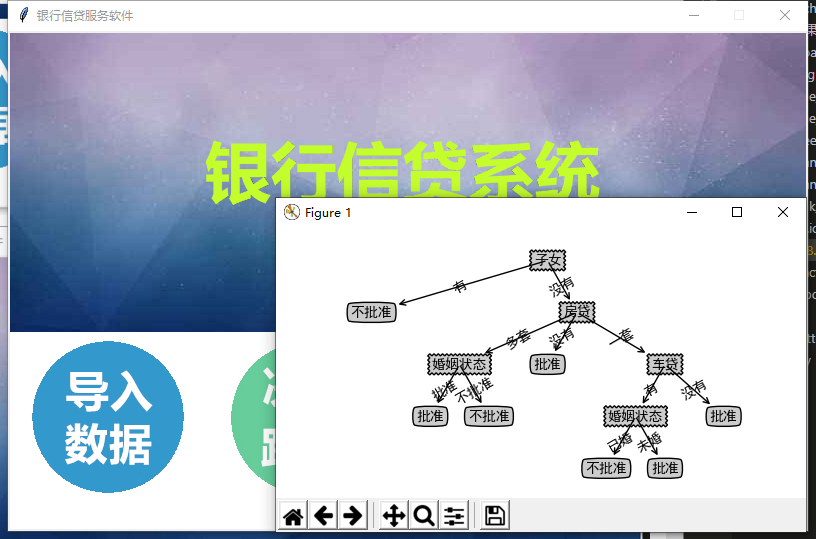
**ID3算法流程图**



实现过程

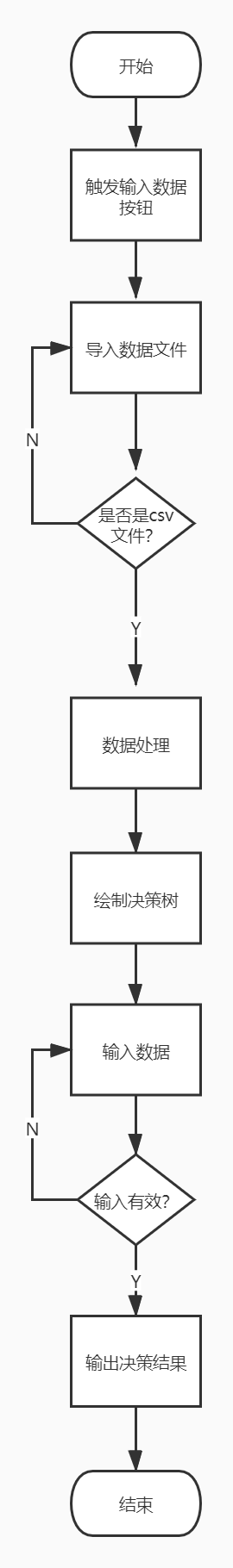










流程图绘制：