# 决策树

不同于朴素贝叶斯同时联立变量，而是先选取最重要的特征，再选取次要特征，知道能分类为止。稳健性比朴素贝叶斯强很多，不会被诸多异常干扰。

选取的标准是信息熵，选取信息熵下降最快的特征，亦称该特征的信息增益，即CD3.0算法。

CD4.5算法先选取超平均信息增益，再按照信息增益率选取。CART使用基尼不纯度代替信息熵。

决策树还很方便地提高稳健性、剪枝、深度，处理缺失值、缺失值可以影响到特征值上去，也可以作为一种类别，或者多试几个类别选择。