Decsison Tree—Gini impurity

c1p

Gini impurity

Gini不纯度计算方法：

例：

X ，P(X=0) = 0.5 ,P(X=1)=0.5

Y ，P(Y=0) = 0.1 ,P(Y=1)=0.9

Gini(X)不纯度计算：

P(X=0)\*(1 - P(X=0)) +P(X=1)\*(1 - P(X=1)) == 0.5

Gini(Y)不纯度计算：

P(Y=0)\*(1 - P(Y=0)) + P(Y=1)\*(1 - P(Y=1)) = = 0.18

在Gini计算下X这种情况下，X中的Gini impurity 较大，分析看X中的数值出现的概率平均，所以这个X中不好确定是1还是0。

在Gini计算下Y的情况下，Y中的Gini impurity 较小，分析Y中的概率，Y = 1 的概率0.9很高，Y选择比X相对稳定