1. **DecisionTree<-function(traindata,trainrowname = NULL, traincolname = NULL, testdata = NULL, testrowname = NULL, testcolname = NULL, yname=NULL, xname=NULL, plotstr = NULL, treename = NULL)**

功能：

决策树。用树状模型来模拟分类决策过程，其中每个内部节点表示一个属性上的测试，每个分支代表一个测试输出，每个叶节点代表一种类别。适合二分类和多分类问题，需要加载rpart包和rpart.plot包。本脚决策树采用CART算法。

输入：

**traindata:**。训练数据集合，两列及以上。

其中yname指定的列是类标号，类型为字符串向量，数值型向量 会被自动转为字符串型。自变量个数不宜过少，一般至少2个。

**testdata** ：测试数据集合，当模型训练完后，输入测试集数据，用来预测其类标号。要求testdata的列名和类型要与xname追定的训练集保持一致。例如，traindata 中xname = [‘’age”, “sex”], 则testdata列也要是[‘’age”, “sex”]这两个向量，且数据类型保持一致。如果testdata为空，则不做预测。特别的，对于testdata 中字符串型的自变量的取值范围不能超过traintest对应的字符串型的自变量的取值范围，否则报错。 如traintest 中 job 自变量取值范围是[“医生”，’教师’，’工程师’]，则testdata 中 job 自变量取值范围不能超过[“医生”，’教师’，’工程师’]。

**trainrowname**：训练集行名。

**traincolname**：训练集列名。

**testrowname**：测试集行名。

**testcolname**：测试集列名。

**yname**：因变量名称，类型为字符串，不可缺省。

**xname**：自变量名称，类型为字符串向量，不可缺省。

**scale:** 是否标准化dataset数据（减去均值，除以标准差），缺省标准化。

**plotstr**：图片输出文件夹目录，类型为字符串，缺省不输出。

**treename**：决策树图输出名称，类型为字符串，缺省不输出。

输出：

**ConfusionRowName**：训练集上的分类混淆矩阵的行名，类型为字符串向量

**ConfusionColName**：训练集上的分类混淆矩阵的列名，类型为字符串向量

**Confusion**：训练集上的分类混淆矩阵。

如果testdata不为空，则输出预测结果：

**PredictionRowName**：预测结果的的行名（即testrowname），类型为字符串向量，。

**PredictionColName**：预测结果的列名，类型为字符串向量。

**pred**：预测结果。