Decision Tree in R

Data

> data(iris)

> attach(iris)

> str(iris)

> summary(iris)

Data Partition

> set.seed(555)

> ind<-sample(2,nrow(iris),replace = TRUE,prob = c(0.8,0.2))

> train<-iris[ind==1,]

> test<-iris[ind==2,]

Decision Tree With Party Package

> library(party)

> tree<-ctree(Species~.,train)

> print(tree)

Visualization of Decision Tree

> plot(tree)

> plot(tree,type="simple")

> head(train)

For Numerical (Quantitative variable)

> tree1<-ctree(Sepal.Length~.,train)

> plot(tree1)

> tree1<-ctree(Species~.,train,controls = ctree\_control(mincriterion=0.9999,minsplit = 6))

> plot(tree1)

Prediction

> predict(tree,train)

> predict(tree1,train,type="prob")

Misclassification Error for Training data

> p1<-predict(tree1,train)

> tab1<-table(Predicted=p1,Actual=train$Species)

> tab1

> 1-sum(diag(tab1))/sum(tab1)

Misclassification for test data

> p2<-predict(tree1,test)

> tab2<-table(Predicted=p2,Actual=test$Species)

> tab2

> 1-sum(diag(tab2))/sum(tab2)

Optimization Techniques

>prod.sol1<-lp("max",obj.fun,constr,constr.dir,int.vec = 1:2,rhs,compute.sens = TRUE)

obj1.fun<-c(4,4)

> constr1<-matrix(c(5,5,0,1,1,0),ncol = 2,byrow = TRUE)

> constr1.dir<-c("<=","<=","<=")

> rhs1<-c(25,4,3)

> library(lpSolve)

> sol<-lp("max",obj1.fun,constr1,constr1.dir,rhs1,compute.sens = TRUE)

> sol

Success: the objective function is 20

> sol$solution

[1] 3 2

>