

secondlabtry

2025-09-29

```
X=vector("numeric",length=10)
for(i in 1:10){
  X[i]=rbinom(1,1,0.8)
}
print(X)
```

```
## [1] 1 1 1 1 1 1 1 1 0 0
```

```
sm=mean(X)
if(sm>0.5){
  print(1)
}else{
  print(0)
}
```

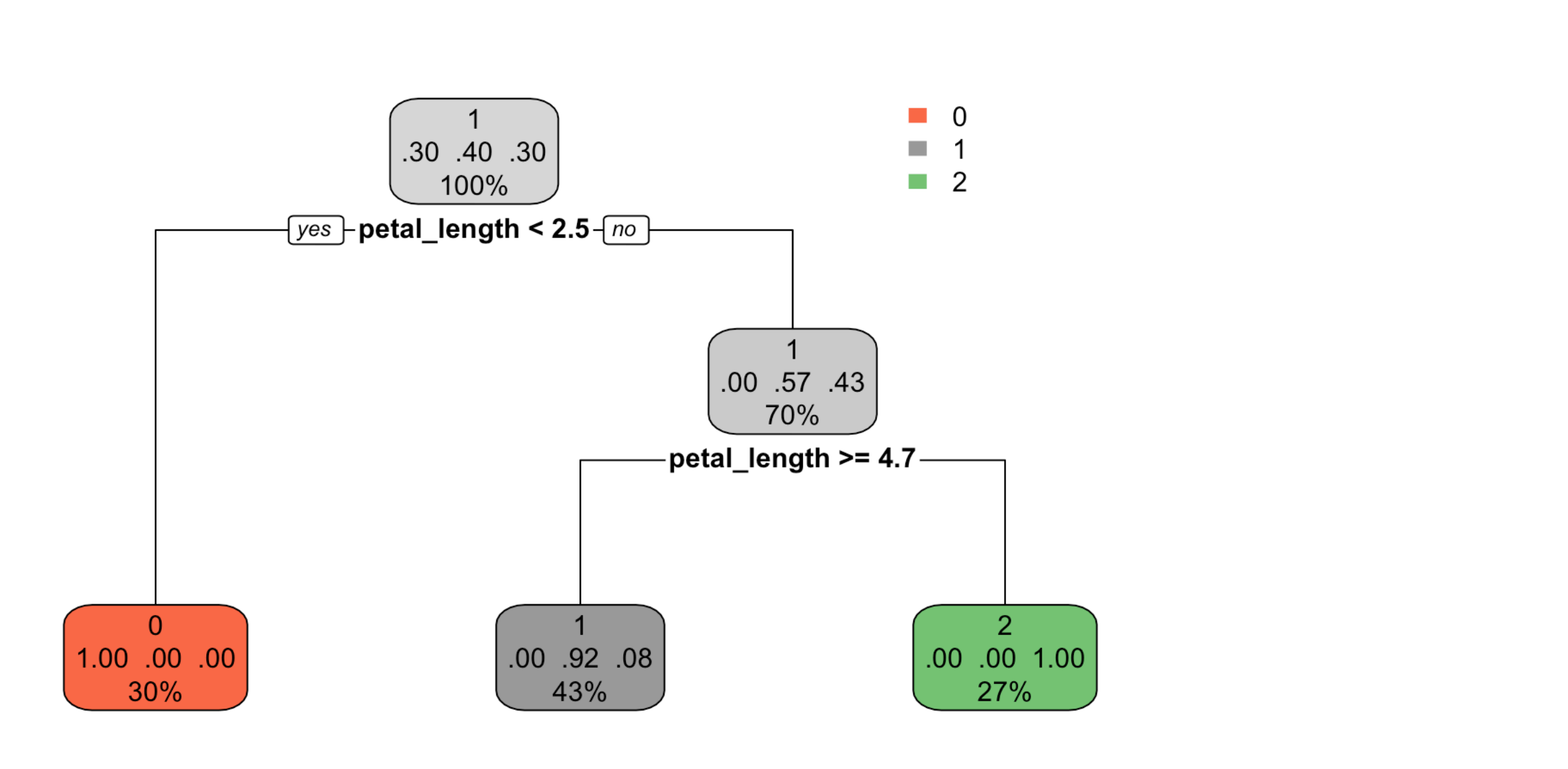
```
## [1] 1
```

```
ccount=0
for(j in 1:10000){
  X=vector("numeric",length=10)
  for(i in 1:10){
    X[i]=rbinom(1,1,0.8)
  }
  sm=mean(X)
  if(sm>0.5){
    ccount=ccount+1
  }
}
acc=ccount/10000
print(acc)
```

```
## [1] 0.9678
```

```
library(rpart)
df=read.csv("iris.csv")
nr=nrow(df)
mclass=vector("numeric",length = nr)
for(i in 1:nr){
  if(df[i,5]=="setosa"){
    mclass[i]=0
  }else if(df[i,5]=='virginica'){
    mclass[i]=1
  }else{
    mclass[i]=2
  }
}
df=cbind(df,mclass)
```

```
s=sample(x=1:nr,size=0.2*nr,replace = FALSE)
train=df[s,-5]
test=df[-s,-5]
mfit=rpart(mclass~,data=train,method="class")
library(rpart.plot)
rpart.plot(mfit)
```



```
ypredtrain=predict(mfit,newdata=train,type="class")
ypredtest=predict(mfit,newdata=test,type="class")
trainacc=mean(train[,5]==ypredtrain)
testacc=mean(test[,5]==ypredtest)
print(trainacc)
```

```
## [1] 0.9666667
```

```
print(testacc)
```

```
## [1] 0.9083333
```

```
ntrees=10
vtrainacc=vector("numeric",length=ntrees)
vtestacc=vector("numeric",length=ntrees)
forest=vector("list",length=ntrees)
for(i in 1:ntrees){
  s=sample(x=1:nr,size=0.2*nr,replace = FALSE)
  train=df[s,-5]
  test=df[-s,-5]
  mfit=rpart(mclass~,data=train,method="class")
  forest[i]=list(mfit=mfit)
  ypredtrain=predict(mfit,newdata = train,type="class")
  ypredtest=predict(mfit,newdata = test,type="class")
  vtrainacc[i]=mean(train[,5]==ypredtrain)
  vtestacc[i]=mean(test[,5]==ypredtest)
}
print(vtrainacc)
```

```
## [1] 0.7000000 1.0000000 0.9333333 0.9666667 0.7666667 1.0000000 0.9666667
## [8] 0.8000000 0.9666667 0.9333333
```

```
print(vtestacc)
```

```
## [1] 0.6583333 0.9333333 0.8500000 0.9416667 0.6000000 0.9416667 0.9083333
## [8] 0.5833333 0.9500000 0.9583333
```

```
sntrain=data.frame(C(1:30))
sntest=data.frame(c(1:120))
for(j in 1:ntrees){
  ypredtrain=predict(forest[[j]],newdata=train,type="class")
  ypredtest=predict(forest[[j]],newdata=test,type="class")
  sntrain=cbind(sntrain,ypredtrain)
  sntest=cbind(sntest,ypredtest)
}
```

```
library(DescTools)
get_value=function(x){
  sm=Mode(x)
  if(length(sm)==1){
    return(sm)
  }else{
    return(sm[1])
  }
}
fypredtrain=apply(sntrain[,2:11],1,get_value)
fypredtest=apply(sntest[,2:11],1,get_value)
ftrainacc=mean(train[,5]==fypredtrain)
ftestacc=mean(test[,5]==fypredtest)
print(ftrainacc)
```

```
## [1] 0.9333333
```

```
print(ftestacc)
```

```
## [1] 0.9666667
```